

(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2020-534795
(P2020-534795A)

(43) 公表日 令和2年12月3日(2020.12.3)

(51) Int.Cl.	F I	テーマコード (参考)
C 1 2 N 15/55 (2006.01)	C 1 2 N 15/55	4 B 0 5 0
C 0 7 K 19/00 (2006.01)	C 0 7 K 19/00 Z N A	4 B 0 6 5
C 1 2 N 9/78 (2006.01)	C 1 2 N 9/78	4 C 0 7 6
C 1 2 N 15/85 (2006.01)	C 1 2 N 15/85 Z	4 C 0 8 4
C 1 2 N 1/15 (2006.01)	C 1 2 N 1/15	4 C 0 8 7
審査請求 未請求 予備審査請求 未請求 (全 224 頁) 最終頁に続く		

(21) 出願番号 特願2020-504311 (P2020-504311)
 (86) (22) 出願日 平成30年7月27日 (2018. 7. 27)
 (85) 翻訳文提出日 令和2年3月30日 (2020. 3. 30)
 (86) 国際出願番号 PCT/US2018/044242
 (87) 国際公開番号 WO2019/023680
 (87) 国際公開日 平成31年1月31日 (2019. 1. 31)
 (31) 優先権主張番号 62/538, 380
 (32) 優先日 平成29年7月28日 (2017. 7. 28)
 (33) 優先権主張国・地域又は機関 米国 (US)

(71) 出願人 507044516
 プレジデント アンド フェローズ オブ
 ハーバード カレッジ
 アメリカ合衆国 マサチューセッツ 02
 138, ケンブリッジ, クインシー
 ストリート 17
 (74) 代理人 100102842
 弁理士 葛和 清司
 (72) 発明者 ザーロニー, ベン
 アメリカ合衆国 マサチューセッツ州 0
 2138, ケンブリッジ, グリスウォール
 ド ストリート 3

最終頁に続く

(54) 【発明の名称】 ファージによって支援される連続的進化 (PACE) を用いて塩基編集因子を進化させるための方法および組成物

(57) 【要約】

本明細書は進化した塩基編集因子を提供し、これらは、当分野のもの欠陥を克服し(増大した効率、および/または編集部位における特異的な配列コンテキストの減少した要求を包含する)、これらは、ファージによって支援される連続的進化(PACE)システムの結果として得られる。特に、本明細書は進化したシチジン塩基編集因子(例えば、APOBEC1、CDA、またはAIDシチジンデアミナーゼドメインに基づく)を提供し、これらは、当分野のもの欠陥を克服し(増大した効率、および/または編集部位における特異的な配列コンテキストの減少した要求を包含する)、これらは、ファージによって支援される連続的進化(PACE)システムの結果として得られる。

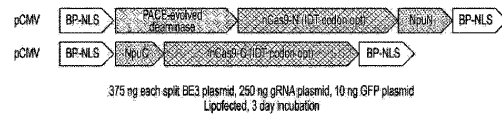


Figure 79

【特許請求の範囲】

【請求項 1】

配列番号1のアミノ酸残基2-162と少なくとも80%、85%、90%、95%、98%、99%、または99.5%同一であるアミノ酸配列を含むシチジンデアミナーゼであって、

シチジンデアミナーゼが、配列番号1と比べてH102X₁、D104X₂、およびV115X₃からなる群から選択される1つ以上の変異または別のシチジンデアミナーゼの対応する変異（単数もしくは複数）を含み、X₁がH以外のいずれかのアミノ酸であり、X₂がD以外のいずれかのアミノ酸であり、X₃がV以外のいずれかのアミノ酸である、

前記シチジンデアミナーゼ。

【請求項 2】

10

配列番号1のアミノ酸残基2-162を含むシチジンデアミナーゼであって、

シチジンデアミナーゼが、配列番号1と比べてH102X₁、D104X₂、およびV115X₃からなる群から選択される1つ以上の変異を含み、X₁がH以外のいずれかのアミノ酸であり、X₂がD以外のいずれかのアミノ酸であり、X₃がV以外のいずれかのアミノ酸である、前記シチジンデアミナーゼ。

【請求項 3】

シチジンデアミナーゼが、配列番号1と比べてH102X₁、D104X₂、およびV115X₃からなる群から選択される2もしくは3つ全ての変異または別のシチジンデアミナーゼの対応する変異を含み、X₁がH以外のいずれかのアミノ酸であり、X₂がD以外のいずれかのアミノ酸であり、X₃がV以外のいずれかのアミノ酸である、請求項1または2に記載のシチジンデアミナーゼ。

20

【請求項 4】

X₁がPである、請求項1～3のいずれか一項に記載のシチジンデアミナーゼ。

【請求項 5】

X₂がNである、請求項1～4のいずれか一項に記載のシチジンデアミナーゼ。

【請求項 6】

X₃がMである、請求項1～5のいずれか一項に記載のシチジンデアミナーゼ。

【請求項 7】

シチジンデアミナーゼが、配列番号1と比べてH102PおよびD104N変異を含む、請求項1～6のいずれか一項に記載のシチジンデアミナーゼ。

30

【請求項 8】

シチジンデアミナーゼが、配列番号5のアミノ酸残基2-162を含む、請求項1～7のいずれか一項に記載のシチジンデアミナーゼ。

【請求項 9】

シチジンデアミナーゼが、N末端メチオニン（M）アミノ酸残基をさらに含む、請求項1～8のいずれか一項に記載のシチジンデアミナーゼ。

【請求項 10】

シチジンデアミナーゼが、配列番号5のアミノ酸配列を含む、請求項1～9のいずれか一項に記載のシチジンデアミナーゼ。

【請求項 11】

40

配列番号2のアミノ酸残基3-229と少なくとも80%、85%、90%、95%、98%、99%、または99.5%同一であるアミノ酸配列を含むシチジンデアミナーゼであって、

シチジンデアミナーゼが、配列番号2と比べてE4X₁、V10X₂、E31X₃、Y40X₄、E95X₅、H109X₆、H122X₇、D124X₈、R126X₉、R154X₁₀、N158X₁₁、A165X₁₂、P201X₁₃、F205X₁₄、およびI208X₁₅からなる群から選択される1つ以上の変異または別のシチジンデアミナーゼの対応する変異（単数もしくは複数）を含み、X₁、X₃、およびX₅がE以外のいずれかのアミノ酸であり、X₂がV以外のいずれかのアミノ酸であり、X₄がY以外のいずれかのアミノ酸であり、X₆およびX₇がH以外のいずれかのアミノ酸であり、X₈がD以外のいずれかのアミノ酸であり、X₉およびX₁₀がR以外のいずれかのアミノ酸であり、X₁₁がN以外のいずれかのアミノ酸であり、X₁₂がA以外のいずれかのアミノ酸であり、X₁₃がP以外のいずれかのアミノ酸で

50

あり、 X_{14} がF以外のいずれかのアミノ酸であり、 X_{15} がI以外のいずれかのアミノ酸である、前記シチジンデアミナーゼ。

【請求項12】

配列番号2のアミノ酸残基3-229を含むシチジンデアミナーゼであって、

シチジンデアミナーゼが、配列番号2と比べてE4 X_1 、V10 X_2 、E31 X_3 、Y40 X_4 、E95 X_5 、H109 X_6 、H122 X_7 、D124 X_8 、R126 X_9 、R154 X_{10} 、N158 X_{11} 、A165 X_{12} 、P201 X_{13} 、F205 X_{14} 、およびI208 X_{15} からなる群から選択される1つ以上の変異を含み、 X_1 、 X_3 、および X_5 がE以外のいずれかのアミノ酸であり、 X_2 がV以外のいずれかのアミノ酸であり、 X_4 がY以外のいずれかのアミノ酸であり、 X_6 および X_7 がH以外のいずれかのアミノ酸であり、 X_8 がD以外のいずれかのアミノ酸であり、 X_9 および X_{10} がR以外のいずれかのアミノ酸であり、 X_{11} がN以外のいずれかのアミノ酸であり、 X_{12} がA以外のいずれかのアミノ酸であり、 X_{13} がP以外のいずれかのアミノ酸であり、 X_{14} がF以外のいずれかのアミノ酸であり、 X_{15} がI以外のいずれかのアミノ酸である、前記シチジンデアミナーゼ。

10

【請求項13】

シチジンデアミナーゼが、配列番号2と比べてE4 X_1 、V10 X_2 、E31 X_3 、Y40 X_4 、E95 X_5 、H109 X_6 、H122 X_7 、D124 X_8 、R126 X_9 、R154 X_{10} 、N158 X_{11} 、A165 X_{12} 、P201 X_{13} 、F205 X_{14} 、およびI208 X_{15} からなる群から選択される2、3、4、5、6、7、8、9、10、11、12、13、14、または15個全ての変異を含み、 X_1 、 X_3 、および X_5 がE以外のいずれかのアミノ酸であり、 X_2 がV以外のいずれかのアミノ酸であり、 X_4 がY以外のいずれかのアミノ酸であり、 X_6 および X_7 がH以外のいずれかのアミノ酸であり、 X_8 がD以外のいずれかのアミノ酸であり、 X_9 および X_{10} がR以外のいずれかのアミノ酸であり、 X_{11} がN以外のいずれかのアミノ酸であり、 X_{12} がA以外のいずれかのアミノ酸であり、 X_{13} がP以外のいずれかのアミノ酸であり、 X_{14} がF以外のいずれかのアミノ酸であり、 X_{15} がI以外のいずれかのアミノ酸である、請求項11または12に記載のシチジンデアミナーゼ。

20

【請求項14】

X_1 がKである、請求項11～13のいずれか一項に記載のシチジンデアミナーゼ。

【請求項15】

X_2 および X_5 がAである、請求項11～14のいずれか一項に記載のシチジンデアミナーゼ。

【請求項16】

X_3 がVである、請求項11～15のいずれか一項に記載のシチジンデアミナーゼ。

30

【請求項17】

X_4 がCである、請求項11～16のいずれか一項に記載のシチジンデアミナーゼ。

【請求項18】

X_6 および X_8 がNである、請求項11～17のいずれか一項に記載のシチジンデアミナーゼ。

【請求項19】

X_7 および X_{15} がLである、請求項11～18のいずれか一項に記載のシチジンデアミナーゼ。

【請求項20】

X_9 および X_{10} がHである、請求項11～19のいずれか一項に記載のシチジンデアミナーゼ。

40

【請求項21】

X_{11} 、 X_{12} 、 X_{13} 、および X_{14} がSである、請求項11～20のいずれか一項に記載のシチジンデアミナーゼ。

【請求項22】

1つ以上の変異が、配列番号2と比べてE4 X_1 、H109 X_6 、H122 X_7 、D124 X_8 、R154 X_{10} 、A165 X_{12} 、P201 X_{13} 、およびF205 X_{14} からなる群から選択され、 X_1 がE以外のいずれかのアミノ酸であり、 X_6 および X_7 がH以外のいずれかのアミノ酸であり、 X_8 がD以外のいずれかのアミノ酸であり、 X_{10} がR以外のいずれかのアミノ酸であり、 X_{12} がA以外のいずれかのアミノ酸で

50

あり、 X_{13} がP以外のいずれかのアミノ酸であり、 X_{14} がF以外のいずれかのアミノ酸である、請求項11～21のいずれか一項に記載のシチジンデアミナーゼ。

【請求項23】

1つ以上の変異が、配列番号2と比べてE4K、H109N、H122L、D124N、R154H、A165S、P201S、およびF205Sからなる群から選択される、請求項11～22のいずれか一項に記載のシチジンデアミナーゼ。

【請求項24】

シチジンデアミナーゼが、配列番号2と比べてE4K、H109N、H122L、D124N、R154H、A165S、P201S、およびF205Sからなる群から選択される2、3、4、5、6、7、または8つ全ての変異を含む、請求項11～23のいずれか一項に記載のシチジンデアミナーゼ。

10

【請求項25】

シチジンデアミナーゼが、配列番号2と比べて変異E4K、H109N、H122L、D124N、R154H、A165S、P201S、およびF205Sを含む、請求項11～24のいずれか一項に記載のシチジンデアミナーゼ。

【請求項26】

シチジンデアミナーゼが、配列番号6のアミノ酸残基3-229を含む、請求項11～25のいずれか一項に記載のシチジンデアミナーゼ。

【請求項27】

シチジンデアミナーゼが、N末端メチオニン(M)アミノ酸残基をさらに含む、請求項11～26のいずれか一項に記載のシチジンデアミナーゼ。

20

【請求項28】

シチジンデアミナーゼが、2つのN末端アミノ酸残基をさらに含み、これらがMおよびSである、請求項11～26のいずれか一項に記載のシチジンデアミナーゼ。

【請求項29】

シチジンデアミナーゼが、配列番号6のアミノ酸配列を含む、請求項11～28のいずれか一項に記載のシチジンデアミナーゼ。

【請求項30】

配列番号3のアミノ酸残基2-208と少なくとも80%、85%、90%、95%、98%、99%、または99.5%同一であるアミノ酸配列を含むシチジンデアミナーゼであって、

シチジンデアミナーゼが、配列番号3と比べて $H10X_1$ 、 $F23X_2$ 、 $V75X_3$ 、 $K120X_4$ 、 $A123X_5$ 、 $C158X_6$ 、 $I193X_7$ 、 $I195X_8$ 、および $V197X_9$ からなる群から選択される1つ以上の変異または別のシチジンデアミナーゼの対応する変異(単数もしくは複数)を含み、 X_1 がH以外のいずれかのアミノ酸であり、 X_2 がF以外のいずれかのアミノ酸であり、 X_3 および X_9 がV以外のいずれかのアミノ酸であり、 X_4 がK以外のいずれかのアミノ酸であり、 X_5 がA以外のいずれかのアミノ酸であり、 X_6 がC以外のいずれかのアミノ酸であり、 X_7 および X_8 がI以外のいずれかのアミノ酸である、前記シチジンデアミナーゼ。

30

【請求項31】

配列番号3のアミノ酸残基2-208を含むシチジンデアミナーゼであって、

シチジンデアミナーゼが、配列番号3と比べて $H10X_1$ 、 $F23X_2$ 、 $V75X_3$ 、 $K120X_4$ 、 $A123X_5$ 、 $C158X_6$ 、 $I193X_7$ 、 $I195X_8$ 、および $V197X_9$ からなる群から選択される1つ以上の変異を含み、 X_1 がH以外のいずれかのアミノ酸であり、 X_2 がF以外のいずれかのアミノ酸であり、 X_3 および X_9 がV以外のいずれかのアミノ酸であり、 X_4 がK以外のいずれかのアミノ酸であり、 X_5 がA以外のいずれかのアミノ酸であり、 X_6 がC以外のいずれかのアミノ酸であり、 X_7 および X_8 がI以外のいずれかのアミノ酸である、前記シチジンデアミナーゼ。

40

【請求項32】

シチジンデアミナーゼが、配列番号3と比べて $H10X_1$ 、 $F23X_2$ 、 $V75X_3$ 、 $K120X_4$ 、 $A123X_5$ 、 $C158X_6$ 、 $I193X_7$ 、 $I195X_8$ 、および $V197X_9$ からなる群から選択される2、3、4、5、6、7、8、または全ての9つの変異を含み、 X_1 がH以外のいずれかのアミノ酸であり、 X_2 がF以外のいずれかのアミノ酸であり、 X_3 および X_9 がV以外のいずれかのアミノ酸であり、 X_4 がK以外のいずれかのアミノ酸であり、 X_5 がA以外のいずれかのアミノ酸であり、 X_6 がC以外のい

50

れかのアミノ酸であり、 X_7 および X_8 がI以外のいずれかのアミノ酸である、請求項30または31に記載のシチジンデアミナーゼ。

【請求項33】

X_1 がYである、請求項30～32のいずれか一項に記載のシチジンデアミナーゼ。

【請求項34】

X_2 がSである、請求項30～33のいずれか一項に記載のシチジンデアミナーゼ。

【請求項35】

X_3 がIである、請求項30～34のいずれか一項に記載のシチジンデアミナーゼ。

【請求項36】

X_4 および X_6 がRである、請求項30～35のいずれか一項に記載のシチジンデアミナーゼ。 10

【請求項37】

X_5 がVである、請求項30～36のいずれか一項に記載のシチジンデアミナーゼ。

【請求項38】

X_7 がTである、請求項30～37のいずれか一項に記載のシチジンデアミナーゼ。

【請求項39】

X_8 がFまたはTである、請求項30～38のいずれか一項に記載のシチジンデアミナーゼ。

【請求項40】

X_9 がAである、請求項30～39のいずれか一項に記載のシチジンデアミナーゼ。 20

【請求項41】

1つ以上の変異が、配列番号3と比べてF23 X_2 、A123 X_5 、およびI195 X_8 からなる群から選択され、 X_2 がF以外のいずれかのアミノ酸であり、 X_5 がA以外のいずれかのアミノ酸であり、 X_8 がI以外のいずれかのアミノ酸である、請求項30～40のいずれか一項に記載のシチジンデアミナーゼ。

【請求項42】

1つ以上の変異が、配列番号3と比べてF23S、A123V、およびI195Fからなる群から選択される、請求項30～41のいずれか一項に記載のシチジンデアミナーゼ。

【請求項43】

シチジンデアミナーゼが、配列番号3と比べてF23S、A123V、およびI195Fからなる群から選択される2または3つ全ての変異を含む、請求項30～42のいずれか一項に記載のシチジンデアミナーゼ。 30

【請求項44】

シチジンデアミナーゼが、配列番号3と比べて変異F23S、A123V、およびI195Fを含む、請求項30～43のいずれか一項に記載のシチジンデアミナーゼ。

【請求項45】

シチジンデアミナーゼが、配列番号7のアミノ酸残基2-208を含む、請求項30～44のいずれか一項に記載のシチジンデアミナーゼ。

【請求項46】

シチジンデアミナーゼが、N末端メチオニン(M)アミノ酸残基をさらに含む、請求項30～45のいずれか一項に記載のシチジンデアミナーゼ。 40

【請求項47】

シチジンデアミナーゼが、配列番号7のアミノ酸配列を含む、請求項30～46のいずれか一項に記載のシチジンデアミナーゼ。

【請求項48】

配列番号4のアミノ酸残基3-229と少なくとも80%、85%、90%、95%、98%、99%、または99.5%同一であるアミノ酸配列を含むシチジンデアミナーゼであって、

シチジンデアミナーゼが、配列番号4と比べてE4 X_1 、H122 X_2 、D124 X_3 、R154 X_4 、A165 X_5 、P201 X_6 、およびF205 X_7 からなる群から選択される1つ以上の変異または別のシチジンデアミナーゼの対応する変異(単数もしくは複数)を含み、 X_1 がE以外のいずれかのアミノ 50

酸であり、 X_2 がH以外のいずれかのアミノ酸であり、 X_3 がD以外のいずれかのアミノ酸であり、 X_4 がR以外のいずれかのアミノ酸であり、 X_5 がA以外のいずれかのアミノ酸であり、 X_6 がP以外のいずれかのアミノ酸であり、 X_7 がF以外のいずれかのアミノ酸である、前記シチジンデアミナーゼ。

【請求項 49】

配列番号4のアミノ酸残基3-229を含むシチジンデアミナーゼであって、

シチジンデアミナーゼが、配列番号4と比べてE4 X_1 、H122 X_2 、D124 X_3 、R154 X_4 、A165 X_5 、P201 X_6 、およびF205 X_7 からなる群から選択される1つ以上の変異を含み、 X_1 がE以外のいずれかのアミノ酸であり、 X_2 がH以外のいずれかのアミノ酸であり、 X_3 がD以外のいずれかのアミノ酸であり、 X_4 がR以外のいずれかのアミノ酸であり、 X_5 がA以外のいずれかのアミノ酸であり、 X_6 がP以外のいずれかのアミノ酸であり、 X_7 がF以外のいずれかのアミノ酸である、前記シチジンデアミナーゼ。

10

【請求項 50】

シチジンデアミナーゼが、配列番号4と比べてE4 X_1 、H122 X_2 、D124 X_3 、R154 X_4 、A165 X_5 、P201 X_6 、およびF205 X_7 からなる群から選択される2、3、4、5、6、または7つ全ての変異を含み、 X_1 がE以外のいずれかのアミノ酸であり、 X_2 がH以外のいずれかのアミノ酸であり、 X_3 がD以外のいずれかのアミノ酸であり、 X_4 がR以外のいずれかのアミノ酸であり、 X_5 がA以外のいずれかのアミノ酸であり、 X_6 がP以外のいずれかのアミノ酸であり、 X_7 がF以外のいずれかのアミノ酸である、請求項48または49に記載のシチジンデアミナーゼ。

【請求項 51】

X_1 がKである、請求項48～50のいずれか一項に記載のシチジンデアミナーゼ。

20

【請求項 52】

X_2 がLである、請求項48～51のいずれか一項に記載のシチジンデアミナーゼ。

【請求項 53】

X_3 がNである、請求項48～52のいずれか一項に記載のシチジンデアミナーゼ。

【請求項 54】

X_4 がHである、請求項48～53のいずれか一項に記載のシチジンデアミナーゼ。

【請求項 55】

X_5 がSである、請求項48～54のいずれか一項に記載のシチジンデアミナーゼ。

【請求項 56】

X_6 がSである、請求項48～55のいずれか一項に記載のシチジンデアミナーゼ。

30

【請求項 57】

X_7 がSである、請求項48～56のいずれか一項に記載のシチジンデアミナーゼ。

【請求項 58】

1つ以上の変異が、配列番号4と比べてE4K、H122L、D124N、R154H、A165S、P201S、およびF205Sからなる群から選択される、請求項48-57のいずれか一項に記載のシチジンデアミナーゼ。

【請求項 59】

シチジンデアミナーゼが、配列番号4と比べてE4K、H122L、D124N、R154H、A165S、P201S、およびF205Sからなる群から選択される2、3、4、5、6、または7つ全ての変異を含む、請求項48～58のいずれか一項に記載のシチジンデアミナーゼ。

40

【請求項 60】

シチジンデアミナーゼが、配列番号4と比べて変異E4K、H122L、D124N、R154H、A165S、P201S、およびF205Sを含む、請求項48～59のいずれか一項に記載のシチジンデアミナーゼ。

【請求項 61】

シチジンデアミナーゼが、配列番号8のアミノ酸残基3-229を含む、請求項48～60のいずれか一項に記載のシチジンデアミナーゼ。

【請求項 62】

シチジンデアミナーゼが、N末端メチオニン(M)アミノ酸残基をさらに含む、請求項4

50

8 ~ 6 1 のいずれか一項に記載のシチジンデアミナーゼ。

【請求項 6 3】

シチジンデアミナーゼが、2つのN末端アミノ酸残基をさらに含み、これらがMおよびSである、請求項 4 8 ~ 6 2 のいずれか一項に記載のシチジンデアミナーゼ。

【請求項 6 4】

シチジンデアミナーゼが、配列番号8のアミノ酸配列を含む、請求項 4 8 ~ 6 3 のいずれか一項に記載のシチジンデアミナーゼ。

【請求項 6 5】

- (i) 核酸によってプログラム可能なDNA結合蛋白質 (napDNAbp) と；
 - (ii) 請求項 1 ~ 6 4 のいずれか一項に記載のシチジンデアミナーゼと；
 - (iii) ウラシルグリコシラーゼ阻害因子ドメイン (UGI) と、
- を含む融合蛋白質。

10

【請求項 6 6】

融合蛋白質が、2、3、4、または5つのUGIドメインを含む、請求項 6 5 に記載の融合蛋白質。

【請求項 6 7】

核酸によってプログラム可能なDNA結合蛋白質 (napDNAbp) が、Cas9ドメインである、請求項 6 5 または 6 6 に記載の融合蛋白質。

【請求項 6 8】

Cas9ドメインが、ヌクレアーゼ活性型Cas9、ヌクレアーゼ不活性型Cas9 (dCas9)、またはCas9ニッカーゼ (nCas9) である、請求項 6 5 ~ 6 7 のいずれか一項に記載の融合蛋白質。

20

【請求項 6 9】

Cas9ドメインが、Cas9ニッカーゼ (nCas9) である、請求項 6 5 ~ 6 8 のいずれか一項に記載の融合蛋白質。

【請求項 7 0】

nCas9が、アミノ酸配列

【化 1 - 1】

DKKYSIGLAIGTNSVGVAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEA
 TRCLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHRLEESFLVEEDKKHERHPIFGNI
 VDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDV

30

【化 1 - 2】

DKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLI
 ALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAIL
 LSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAG
 YIDGGASQEEFYKFIKPILEKMDGTEELLVKNLREDLLRKQRTFDNGSIPHQIHLGELHAI
 LRRQEDFYFPLKDNREKIEKILTRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVD
 KGASAQSFIERMTNFDKNLPNEKVLPHKSLLYEYFTVYNELTKVKYVTEGMRKPAFLSG
 EQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIK
 DKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGR
 LSRKLINGIRDKQSGKTILDFLKSDFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLH
 EHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRER
 MKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVD
 HIVPQSFLKDDSIDNKVLTRSDKNRGSNDNPSEEVVKKMKNYWRQLLNAKLITQRKF
 DNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVIT
 LKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVY
 DVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKG
 RDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK KYGGFDS
 PTVAYSVLVVAKVEKGKSKLKS VKELLGITIMERSSEKPNIDFLEAKGYKEVKKDLIIK
 LPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLGKSPEDNEQK
 QLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKHRDKPIREQAENIIHLFTLTN
 LGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQ SITGLYETRIDLSQLGGD (配列番号
 9).

10

20

30

と少なくとも80%、85%、90%、95%、98%、99%、99.5%、または99.9%同一であるアミノ酸配列を含む、請求項65～69のいずれか一項に記載の融合蛋白質。

【請求項71】

nCas9が、配列番号9に提示されているアミノ酸配列を含む、請求項65～70のいずれか一項に記載の融合蛋白質。

【請求項72】

napDNAbpが、CasX、CasY、Cpf1、C2c1、C2c2、C2c3、またはアルゴノート蛋白質である、請求項65～66のいずれか一項に記載の融合蛋白質。

【請求項73】

UGIドメインが、UDG活性を阻害する能力があるドメインを含む、請求項65～72のいずれか一項に記載の融合蛋白質。

40

【請求項74】

UGIドメインが、アミノ酸配列

【化2】

MTNLSDIIEKETGKQLVIQESILMLPEEVVEEVIGNKPESDILVHTAYDESTDENVMLLTSDA
 PEYKPWALVIQDSNGENKIKML (配列番号 10).

50

と少なくとも80%、85%、90%、95%、98%、99%、99.5%、または99.9%同一であるアミノ酸配列を含む、請求項65～73のいずれか一項に記載の融合蛋白質。

【請求項75】

UGIドメインが、配列番号10に提示されているアミノ酸配列を含む、請求項65～74のいずれか一項に記載の融合蛋白質。

【請求項76】

融合蛋白質が、構造：

NH_2 -[シチジンデアミナーゼ]-[napDNAbp]-[UGI]-COOH；

NH_2 -[シチジンデアミナーゼ]-[UGI]-[napDNAbp]-COOH；

NH_2 -[napDNAbp]-[UGI]-[シチジンデアミナーゼ]-COOH；

NH_2 -[napDNAbp]-[シチジンデアミナーゼ]-[UGI]-COOH；

NH_2 -[UGI]-[シチジンデアミナーゼ]-[napDNAbp]-COOH；および

NH_2 -[UGI]-[napDNAbp]-[シチジンデアミナーゼ]-COOH；

を含み、

シチジンデアミナーゼが請求項1～64のいずれか一項に記載のシチジンデアミナーゼであり、UGIがUGIドメインであり、「-」の各個が任意のリンカーを含む、

請求項65～75のいずれか一項に記載の融合蛋白質。

【請求項77】

融合蛋白質が、構造：

NH_2 -[シチジンデアミナーゼ]-[napDNAbp]-[UGI]-[UGI]-COOH；

NH_2 -[シチジンデアミナーゼ]-[UGI]-[napDNAbp]-[UGI]-COOH；

NH_2 -[UGI]-[シチジンデアミナーゼ]-[napDNAbp]-[UGI]-COOH；

NH_2 -[シチジンデアミナーゼ]-[UGI]-[napDNAbp]-[UGI]-COOH；

NH_2 -[シチジンデアミナーゼ]-[UGI]-[UGI]-[napDNAbp]-COOH；

NH_2 -[UGI]-[シチジンデアミナーゼ]-[UGI]-[napDNAbp]-COOH；

NH_2 -[UGI]-[napDNAbp]-[UGI]-[シチジンデアミナーゼ]-COOH；

NH_2 -[napDNAbp]-[UGI]-[UGI]-[シチジンデアミナーゼ]-COOH；

NH_2 -[napDNAbp]-[UGI]-[シチジンデアミナーゼ]-[UGI]-COOH；

NH_2 -[napDNAbp]-[シチジンデアミナーゼ]-[UGI]-[UGI]-COOH；

NH_2 -[napDNAbp]-[UGI]-[シチジンデアミナーゼ]-[UGI]-COOH；

NH_2 -[UGI]-[napDNAbp]-[シチジンデアミナーゼ]-[UGI]-COOH；

NH_2 -[UGI]-[シチジンデアミナーゼ]-[napDNAbp]-[UGI]-COOH；

NH_2 -[UGI]-[シチジンデアミナーゼ]-[UGI]-[napDNAbp]-COOH；

NH_2 -[UGI]-[UGI]-[シチジンデアミナーゼ]-[napDNAbp]-COOH；

NH_2 -[UGI]-[napDNAbp]-[シチジンデアミナーゼ]-[UGI]-COOH；

NH_2 -[UGI]-[napDNAbp]-[UGI]-[シチジンデアミナーゼ]-COOH；および

NH_2 -[UGI]-[UGI]-[napDNAbp]-[シチジンデアミナーゼ]-COOH；

を含み、

シチジンデアミナーゼが請求項1～64のいずれか一項に記載のシチジンデアミナーゼであり、UGIがUGIドメインであり、「-」の各個が任意のリンカーを含む、

請求項65～76のいずれか一項に記載の融合蛋白質。

【請求項78】

融合蛋白質が、構造：

NH_2 -[シチジンデアミナーゼ]-[napDNAbp]-[UGI]-[UGI]-COOH；

を含み、

シチジンデアミナーゼが請求項1～64のいずれか一項に記載のシチジンデアミナーゼであり、UGIがUGIドメインであり、「-」の各個が任意のリンカーを含む、

請求項65～77のいずれか一項に記載の融合蛋白質。

【請求項79】

シチジンデアミナーゼおよびnapDNAbpが、アミノ酸配列SGGSSGGSSGSETPGTSESATPESSGGS

10

20

30

40

50

SGGS (配列番号11) を含むリンカーを介して融合されている、請求項 7 8 に記載の融合蛋白質。

【請求項 8 0】

napDNAbp および UGI ドメインが、アミノ酸配列 SGGSGGSGGS (配列番号12) を含むリンカーを介して融合されている、請求項 7 8 または 7 9 に記載の融合蛋白質。

【請求項 8 1】

UGI ドメインが、アミノ酸配列 SGGSGGSGGS (配列番号12) を含むリンカーを介して融合されている、請求項 7 7 ~ 8 0 のいずれか一項に記載の融合蛋白質。

【請求項 8 2】

融合蛋白質が、核局在配列 (NLS) をさらに含む、請求項 7 7 ~ 8 1 のいずれか一項に記載の融合蛋白質。

10

【請求項 8 3】

NLS が、アミノ酸配列 KRTADGSEFEPKKKRKV (配列番号13) を含む、請求項 8 2 に記載の融合蛋白質。

【請求項 8 4】

融合蛋白質が、2つの核局在配列 (NLS) を含む、請求項 8 2 または 8 3 に記載の融合蛋白質。

【請求項 8 5】

融合蛋白質が、構造：

NH_2 -[NLS]-[シチジンデアミナーゼ]-[napDNAbp]-[UGI]-[UGI]-[NLS]-COOH；

20

を含み、

シチジンデアミナーゼが請求項 1 ~ 6 4 のいずれか一項に記載のシチジンデアミナーゼであり、UGI が UGI ドメインであり、「-」の各個が任意のリンカーを含む、

請求項 8 2 ~ 8 4 のいずれか一項に記載の融合蛋白質。

【請求項 8 6】

UGI および NLS が、アミノ酸配列 SGGGS (配列番号14) を含むリンカーを介して融合されている、請求項 8 5 に記載の融合蛋白質。

【請求項 8 7】

融合蛋白質が、配列番号 1 5 ~ 2 0 のいずれか 1 つに提示されているアミノ酸配列のいずれか 1 つと少なくとも 80%、85%、90%、95%、98%、99%、または 99.5% 同一であるアミノ酸配列を含む、請求項 6 5 ~ 8 6 のいずれか一項に記載の融合蛋白質。

30

【請求項 8 8】

融合蛋白質が、配列番号 1 5 ~ 2 0 のいずれか 1 つに提示されているアミノ酸配列のいずれか 1 つを含む、請求項 6 5 ~ 8 6 のいずれか一項に記載の融合蛋白質。

【請求項 8 9】

請求項 1 ~ 6 4 のいずれか一項に記載のシチジンデアミナーゼまたは請求項 6 5 ~ 8 8 のいずれか一項に記載の融合蛋白質をコードする、核酸配列。

【請求項 9 0】

核酸配列が、哺乳類細胞における発現のためにコドン最適化されている、請求項 8 9 に記載の核酸。

40

【請求項 9 1】

哺乳類細胞が、HEK293T 細胞である、請求項 9 0 に記載の核酸。

【請求項 9 2】

請求項 8 9 ~ 9 1 のいずれか一項に記載の核酸を含むベクター。

【請求項 9 3】

ベクターが、核酸の発現を駆動する異種プロモーターを含む、請求項 9 2 に記載のベクター。

【請求項 9 4】

請求項 6 5 ~ 8 8 のいずれか一項に記載の融合蛋白と融合蛋白質の napDNAbp に結合した RNA とを含む複合体。

50

【請求項 9 5】

RNAが、ガイドRNA (gRNA) である、請求項 9 4 に記載の複合体。

【請求項 9 6】

RNAが、シングルガイドRNA (sgRNA) である、請求項 9 4 または 9 5 に記載の複合体。

【請求項 9 7】

sgRNAが、

gttttagagctagaaatagcaagttaaaataaggctagtcggttatcaacttgaaaaagtgccaccgagtcggtgctttttttの配列番号132のバックボーン配列を含み、任意に、tの各個がウラシル(u)である、請求項 7 0 に記載の複合体。

【請求項 9 8】

RNAが、10~100ヌクレオチドの長さであり、標的配列に対して相補的である少なくとも10個の一続きのヌクレオチドの配列を含む、請求項 9 4 ~ 9 7 のいずれか一項に記載の複合体。

【請求項 9 9】

RNAが、標的配列に対して相補的である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、または40個の一続きのヌクレオチドの配列を含む、請求項 9 4 ~ 9 8 のいずれか一項に記載の複合体。

【請求項 1 0 0】

標的配列が、DNA配列である、請求項 9 4 ~ 9 9 のいずれか一項に記載の複合体。

【請求項 1 0 1】

RNAが、ApoEまたはTMC1をコードする核酸配列に対して相補的である少なくとも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、または40個の一続きのヌクレオチドを含む核酸配列を含む、請求項 9 4 ~ 1 0 0 のいずれか一項に記載の複合体。

【請求項 1 0 2】

標的が、生物のゲノム中にある、請求項 9 8 ~ 1 0 1 のいずれか一項に記載の複合体。

【請求項 1 0 3】

生物が原核生物である、請求項 1 0 2 に記載の複合体。

【請求項 1 0 4】

生物が真核生物である、請求項 1 0 2 に記載の複合体。

【請求項 1 0 5】

生物が脊椎動物である、請求項 1 0 4 に記載の複合体。

【請求項 1 0 6】

脊椎動物が哺乳動物である、請求項 1 0 5 に記載の複合体。

【請求項 1 0 7】

哺乳動物がヒトである、請求項 1 0 6 に記載の複合体。

【請求項 1 0 8】

請求項 1 ~ 6 4 のいずれか一項に記載のシチジンデアミナーゼまたは請求項 6 5 ~ 8 8 のいずれか一項に記載の融合蛋白質を含む、細胞。

【請求項 1 0 9】

請求項 8 9 ~ 9 1 のいずれか一項に記載の核酸を含む細胞。

【請求項 1 1 0】

請求項 9 2 または 9 3 に記載のベクターを含む細胞。

【請求項 1 1 1】

請求項 9 4 ~ 1 0 7 のいずれか一項に記載の複合体を含む細胞。

【請求項 1 1 2】

請求項 6 5 ~ 8 8 のいずれか一項に記載の融合蛋白質をコードする核酸配列と；融合蛋白質の発現を駆動する異種プロモーターとを含む核酸コンストラクトを含む、キット。

【請求項 1 1 3】

	10
	20
	30
	40
	50

さらに、ガイドRNAバックボーンをコードする発現コンストラクトを含み、コンストラクトが、ガイドRNAバックボーンへの標的配列と同一または相補的な核酸配列のクローニングを許すように位置したクローニング部位を含む、請求項112に記載のキット。

【請求項114】

請求項65～88のいずれか一項に記載の融合蛋白質を含む医薬組成物。

【請求項115】

請求項94～107のいずれか一項に記載の複合体を含む医薬組成物。

【請求項116】

請求項89～91のいずれか一項に記載の核酸を含む医薬組成物。

【請求項117】

請求項92または93に記載のベクターを含む医薬組成物。

【請求項118】

さらに、薬学的に許容し得る賦形剤を含む、請求項114～117のいずれか一項に記載の医薬組成物。

【請求項119】

さらに脂質を含む、請求項114～118のいずれか一項に記載の医薬組成物。

【請求項120】

脂質がカチオン性脂質である、請求項119に記載の医薬組成物。

【請求項121】

さらにポリマーを含む、請求項114～120のいずれか一項に記載の医薬組成物。

【請求項122】

核酸分子を請求項94～107のいずれか一項に記載の複合体と接触させることを含む、方法。

【請求項123】

核酸がDNAである、請求項122に記載の方法。

【請求項124】

核酸が二本鎖DNAである、請求項123に記載の方法。

【請求項125】

核酸が、疾患または異常に関連する標的配列を含む、請求項122～124のいずれか一項に記載の方法。

【請求項126】

標的配列が、疾患または異常に関連する点変異を含む、請求項125に記載の方法。

【請求項127】

標的配列が、疾患または異常に関連するT CまたはA G点変異を含み、変異体C塩基の脱アミノ化が、疾患または異常に関連しない配列をもたらす、請求項125～126のいずれか一項に記載の方法。

【請求項128】

標的配列が、蛋白質をコードし、点変異がコドンにあり、野生型コドンと比較して変異体コドンによってコードされるアミノ酸の変化をもたらす、請求項125～127のいずれか一項に記載の方法。

【請求項129】

標的配列が、スプライシング部位にあり、点変異が、野生型転写物と比較してmRNA転写物のスプライシングの変化をもたらす、請求項125～127のいずれか一項に記載の方法。

【請求項130】

標的配列が、遺伝子のプロモーターにあり、点変異が遺伝子の増大した発現をもたらす、請求項125～127に記載の方法。

【請求項131】

標的配列が、遺伝子のプロモーターにあり、点変異が遺伝子の減少した発現をもたらす、請求項125～127に記載の方法。

10

20

30

40

50

【請求項 1 3 2】

変異体Cの脱アミノ化が、変異体コドンによってコードされるアミノ酸の変化をもたらす、請求項 1 2 7 ~ 1 3 1 のいずれか一項に記載の方法。

【請求項 1 3 3】

変異体Cの脱アミノ化が、野生型アミノ酸をコードするコドンをもたらす、請求項 1 2 7 ~ 1 3 1 のいずれか一項に記載の方法。

【請求項 1 3 4】

変異体Cの脱アミノ化が、mRNA転写物の変化をもたらす、請求項 1 2 9 に記載の方法。

【請求項 1 3 5】

変異体Cの脱アミノ化が、野生型mRNA転写物をもたらす、請求項 1 2 7 ~ 1 3 4 のいずれか一項に記載の方法。 10

【請求項 1 3 6】

変異体Cの脱アミノ化が、遺伝子の増大した発現をもたらす、請求項 1 2 7 ~ 1 3 5 のいずれか一項に記載の方法。

【請求項 1 3 7】

変異体Cの脱アミノ化が、遺伝子の減少した発現をもたらす、請求項 1 2 7 ~ 1 3 5 のいずれか一項に記載の方法。

【請求項 1 3 8】

接触させることが、in vitroで実行される、請求項 1 2 2 ~ 1 3 7 のいずれか一項に記載の方法。 20

【請求項 1 3 9】

接触させることが、対象においてin vivoで実行される、請求項 1 2 2 ~ 1 3 7 のいずれか一項に記載の方法。

【請求項 1 4 0】

対象が、疾患または異常と診断されている、請求項 1 3 9 に記載の方法。

【請求項 1 4 2】

疾患または異常が、ApoE遺伝子またはTmc1遺伝子の点変異に関連し、任意に、点変異がApoE C112RまたはApoE C158R変異であり、任意に、点変異がTmc1 Y182C変異である、請求項 8 9 ~ 1 0 6 のいずれか一項に記載の方法。

【請求項 1 4 3】 30

任意に、デグロンタグが、AANDENYNYALAA（配列番号134）と少なくとも80%、85%、90%、95%、98%、99%、または100%同一であるアミノ酸配列を含む、デグロンタグに融合されているRNAポリメラーゼを含む融合蛋白質。

【請求項 1 4 4】

RNAポリメラーゼが、T7 RNAポリメラーゼであり、任意に、デグロンタグが、

【化 3】

MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFRKMFERQL
KAGEVADNAAAKPLITLLPKMIARINDWFEEVKAKRGRPTAFQFLQEIKPEAVAYITIK
TTLACLTSADNTTVQAVASAIGRAIEDEARFGRIRDLKAKHFKNVVEQLNKRVGHVYK
KAFMQVVEADMLSKGLLGGEAWSSWHKEDSIHVGVRCIEMLIESTGMVSLHRQNAGV
VGQDSETIELAPEYAEAIATRAGALAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLAL
VRTHSKKALMRYEDVYMPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAI
EREELPMKPEDIDMNPEALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKAI
WFPYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLVKIHGANCAGV
DKVPPPERIKFIEENHENIMACAKSPLNTWWAEQDSPFCFLAFCFEYAGVQHHGLSYNC
SLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYGIVAKKVNEILQADAINGT
DNEVVTVTDENTGEISEKVKLGTKALAGQWLAYGVTRSVTKRSVMTLAYGSKEFGFRQ
QVLEDTIQPAIDSGKGLMFTQPNQAAGYMAKLIWESVSVTVVAAVEAMNWLKSAKLL
AAEVKDKKTGEILRKRCVHWVTPDGFVWQEYKKPIQTRLNLMFLGQFRLQPTINTN
KDSEIDAHKQESGIAPNFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANL
FKAVRETMDVTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRDILESDFAF
(配列番号 133).

10

20

と少なくとも80%、85%、90%、95%、98%、99%、または100%同一であるアミノ酸配列を含む、請求項143に記載の融合蛋白質。

【請求項145】

デグロンタグが、RNAポリメラーゼのC末端に融合されており、任意に、融合体が、

【化4】

MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFRKMFERQL
 KAGEVADNAAAKPLITLLPKMIARINDWFEEVKAKRGKRPTAFQFLQEIKPEAVAYITIK
 TTLACLTSADNTTVQAVASAIGRAIEDEARFGRIRDLKAKHFKNVEEQLNKRVGHVYK
 KAFMQVVEADMLSKGLLGGEAWSSWHKEDSIHVGVRCEMLIESTGMVSLHRQNAGV
 VGQDSETIELAPEYAEAIATRAGALAGISPMFQPCVVPKPWTGITGGGYWANGRRPLAL
 VRTHSKKALMRYEDVYMPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAI
 EREELPMKPEDIDMNPEALTAWKRAAAVYRKDKARKSRRISLEFMLEQANKFANHKAI
 WFPYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLVKIHGANCAGV
 DKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAGVQHHGLSYNC
 SLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYGIVAKKVNILQADAINGT
 DNEVVTVTDENTGEISEKVKLGTKALAGQWLAYGVTRSVTKRSVMTLAYGSKEFGFRQ
 QVLEDTIQPAIDSGKGLMFTQPNQAAGYMAKLIWESVSVTVVA AVEAMNWLKSAAKLL
 AA EVKDKKTGEILRKRC AVHWVTPDGFVWQEYK KPIQTRLNLMFLGQFRLQPTINTN
 KDSEIDAHKQESGIAPNFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTIPADAANL
 FKAVRETMVDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRDILESDFAF
 WTRAANDENYNYALAA (配列番号 135)

10

20

と少なくとも80%、85%、90%、95%、98%、99%、または100%同一であるアミノ酸配列を含む、請求項143または144に記載の融合蛋白質。

【請求項146】

請求項143～145のいずれか一項に記載の融合蛋白質をコードする核酸。

【請求項147】

請求項146に記載の核酸を含むベクター。

30

【請求項148】

任意に、融合体が、

【化5】

CLSYETEILTVEYGLLPYGKIVEKRIECTVYSVDNNGNIYTQPVAQWHDRGEQEVFEYCLE
 DGSLIRATKDHKFMTVDGQMLPIDEIFERELDLMRVDNLPN* (配列番号 136).

と少なくとも80%、85%、90%、95%、98%、99%、または100%同一であるアミノ酸配列を含む、Nインテインに融合されているシチジンデアミナーゼを含む、融合蛋白質。

40

【請求項149】

シチジンデアミナーゼが、APOBECデアミナーゼである、請求項148に記載の融合蛋白質。

【請求項150】

シチジンデアミナーゼが、請求項1～64のいずれか一項に記載のシチジンデアミナーゼである、請求項148または149に記載の融合蛋白質。

【請求項151】

NインテインがシチジンデアミナーゼのC末端に融合されており、任意に、融合体が、

【化6】

MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELKRETCCLLYEINWGGRHSIWRHTSQNTNK
 HVEVNFIEKFTTTERYFCPNTRCSITWFLSWSPCGECSRAITEFLSRYPHVTLFIYIARLYHH
 ADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFVNYSNEAHWPRYPHLWVRLYVLE
 LYCIILGLPPCLNILRRKQPQLTFFTIALQSCHYQRLPPHILWATGLKSGSETPGTSESATPE
 CLSYETEILTVEYGLLPKIVEKRIECTVYSVDNNGNIYTQPVAQWHDRGEQEVFEYCLE
 DGSLIRATKDHKFM TVDGMPLPIDEIFERELDLMRVDNLPN (配列番号 137).

10

と少なくとも80%、85%、90%、95%、98%、99%、または100%同一であるアミノ酸配列を含む、請求項148～150のいずれか一項に記載の融合蛋白質。

【請求項152】

請求項148～151のいずれか一項に記載の融合蛋白質をコードする核酸。

【請求項153】

任意に、ベクターが、ファージミドであり、さらに任意に、ファージミドが、M13ファージミドである、請求項152に記載の核酸を含むベクター。

【請求項154】

任意に、融合体が、MIKIATRKYLGKQNVYDIGVERDHNFKNGFIASNC (配列番号138) と少なくとも80%、85%、90%、95%、98%、99%、または100%同一であるアミノ酸配列を含む、Cas9ドメイン、ウラシルDNAグリコシラーゼ阻害因子(UGI)ドメイン、およびCインテインを含む融合蛋白質。

20

【請求項155】

Cas9ドメインが、ヌクレアーゼ不活性型Cas9ドメイン(dCas9)またはCas9ニックアーゼ(nCas9)である、請求項154に記載の融合蛋白質。

【請求項156】

Cインテインが、Cas9ドメインのN末端に融合されており、UGIドメインがCas9ドメインのC末端に融合されており、任意に、融合体が、

【化7-1】

30

MIKIATRKYLGKQNVYDIGVERDHNFKNGFIASNCFNKYSIGLAIGTNSVGWAVITDE
 YKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQE
 IFSNEMAKVDDSSFHRLSEESFLVEEDKKHERHPHIFGNIVDEVAYHEKYPTIYHLRKKLVDS
 TDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASG
 VDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIASLGLTPNFKSNFDLAEDAQLQ
 LSKD TYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRY
 DEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDG
 TEELLVKNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPFLKDNREKIEKILFR
 IPYYVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGGASAQSFIERMTNFDKNLPNEK
 VLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLK
 EDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIKDKDFLDNEENEDILEDIVLTLTLFE
 DREMIEERLKTYAHLFDDKVMKQLKRRRYTGWRKLSRKLINGIRDKQSGKTILDFLKSD

40

【化 7 - 2】

GFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIAKKGILQTVKVVDE
 LVKVMGRHHPENIVIAMARENQTTQKGQKNSRERMKRIIEGKELGSQILKEHPVENTQ
 LQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVAIVPQSFLKDDSIDNKVLRSDKNR
 GKSDNVPSEEVVKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLV
 ETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVVREINNYH
 HAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSN
 IMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNVKKTVEVQ
 TGGFSKESILPKRNSDKLIARKKDWDPKKGFFSPTVAYSVLVVAKVEKGGKSKKLSV
 KELLGITIMERSSEFEKNPIDFLEAKGYKEVKKDLIKLPKYSLELENGRKRMLASAGELQ
 KGNELALPSKYVNFLYLASHYEKLGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVIL
 ADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEV
 LDATLIHQSTGLYETRIDLSQLGGDSGGSMTNLSDIIEKETGKQLVIQESILMLPEEVEEVI
 GNKPESDILVHTAYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKML (配列番号
 139).

10

20

と少なくとも80%、85%、90%、95%、98%、99%、または100%同一であるアミノ酸配列を含む、請求項154または155に記載の融合蛋白質。

【請求項157】

請求項154～156のいずれか一項に記載の融合蛋白質をコードする核酸。

【請求項158】

任意に、ベクターが、ファージミドであり、任意に、ファージミドが、M13ファージミドである、請求項157に記載の核酸を含むベクター。

【請求項159】

ウラシルDNA結合蛋白質と転写活性化因子とを含む融合蛋白質。

30

【請求項160】

転写活性化因子が、RpoZである、請求項159に記載の融合蛋白質。

【請求項161】

ウラシルDNA結合蛋白質が、UdgXである、請求項159または160に記載の融合蛋白質。

【請求項162】

請求項159～161のいずれか一項に記載の融合蛋白質をコードする核酸。

【請求項163】

請求項162に記載の核酸を含むベクター。

40

【請求項164】

ファージに基づく連続的な指向的進化のためのベクターシステムであって、以下：

a. シトシン、またはインティンをコードするその一部を脱アミノ化する能力がある塩基編集因子蛋白質をコードする核酸を含むベクターと；

b. 請求項147に記載のベクターと；

c. luxABおよびガイドRNA (gRNA) をコードするベクターと、

を含む、前記ベクターシステム。

【請求項165】

a. 請求項147に記載のベクターと；

b. 請求項153に記載のベクターと；

50

c. 請求項 1 5 4 に記載のベクターと；
 d. gIII、luxAB、およびガイドRNA (gRNA) をコードする核酸を含むベクターと、
 を含むベクターシステム。

【請求項 1 6 6】

a. シトシンを脱アミノ化する能力がある塩基編集因子蛋白質をコードする核酸を含むベクターと；

b. 請求項 1 6 3 に記載のベクターと；

c. luxABおよびgRNAをコードする核酸を含むベクターと、
 を含むベクターシステム。

【請求項 1 6 7】

luxABおよびgRNAをコードする核酸が、ドミナントネガティブgIII蛋白質バリエーション、
 例えばpIII-negをコードする、請求項 1 6 6 に記載のベクターシステム。

【請求項 1 6 8】

請求項 1 6 4 ~ 1 6 6 のいずれか一項に記載のベクターシステムを含む細胞。

【請求項 1 6 9】

請求項 1 4 7、1 5 3、1 5 8、または 1 6 3 のいずれか一項に記載のベクターを含む
 細胞またはファージ。

【請求項 1 7 0】

核酸の連続的進化の方法であって、以下：

(i) 請求項 1 5 3 に記載のベクターを含むセレクションファージミド (SP) をラグーン中の細菌宿主細胞のフローに導入することと (宿主細胞は、セレクションファージミドを感染性のファージ粒子中へとパッケージングするために要求されるファージ遺伝子を含み、セレクションファージミドをファージ粒子中へとパッケージングするために要求される少なくとも1つの遺伝子が欠損しており、セレクションファージミドを感染性のファージ粒子中へとパッケージングするために要求される少なくとも1つの遺伝子は、宿主細胞において進化させられるべき遺伝子の発現にตอบสนองして発現され、ラグーン中の宿主細胞の流量は、ラグーン中の宿主細胞ではなくファージミドの複製を許す)；

(ii) ホスト細胞のフロー中のファージミドを複製および変異させることと；

(iii) 進化させられるべき変異した遺伝子を含むファージミドを細胞のフローから単離することと、

を含む、前記方法。

【請求項 1 7 1】

宿主細胞が、宿主細胞には欠損しているセレクションファージミドをファージ粒子へとパッケージングするために要求される遺伝子を含む第1のアクセサリプラスミド (AP) を含み、ここで遺伝子が、SPによってコードされる融合蛋白質の発現にตอบสนองしてアクセサリプラスミドから発現される、請求項 1 7 0 に記載の方法。

【請求項 1 7 2】

第1のAPが、gIII、luxAB、およびガイドRNA (gRNA) をコードする核酸を含む、請求項 1 7 1 に記載の方法。

【請求項 1 7 3】

宿主細胞が、第2のAPを含み、ここで第2のAPが請求項 1 4 7 に記載のベクターを含む、請求項 1 7 2 に記載の方法。

【請求項 1 7 4】

宿主細胞が、第3のAPを含み、ここで第3のAPが請求項 1 5 8 に記載のベクターを含む、請求項 1 7 3 に記載の方法。

【請求項 1 7 5】

請求項 1 7 0 ~ 1 7 4 のいずれか一項に記載の方法によって調製される塩基編集因子。

【請求項 1 7 6】

塩基編集因子がシチジンデアミナーゼ、Cas9ドメイン、およびUGIドメインを含む、請求項 1 7 5 に記載の塩基編集因子。

10

20

30

40

50

【発明の詳細な説明】

【技術分野】

【0001】

関連出願および参照による組み込み

本願は、2017年7月28日出願のU.S.仮特許出願No.62/538,380の優先権を主張し、この内容全体は参照によって本明細書に組み込まれる。

【背景技術】

【0002】

発明の背景

核酸配列の標的化された編集、例えばゲノムDNAの標的化された切断または特異的改変の標的化された導入は、遺伝子機能の研究のための高度に有望なアプローチであり、ヒト遺伝子疾患、例えば点変異によって引き起こされるものための新たな治療を提供するポテンシャルをもまた有する。点変異は、疾患に関連する公知のヒト遺伝子バリエーションの大多数にあたる(1)。よって、点変異を導入および修正するためのロバストな方法を開発することは、遺伝学的構成要素を有する疾患を理解および処置するための重要な課題である。

10

【0003】

操作された塩基編集因子が最近開発されてきた(2,3)。塩基編集因子は、触媒欠損したCas部分と核酸塩基修飾酵素(例えば、天然のまたは進化した核酸塩基デアミナーゼ、例えば、APOBEC1(「アポリポ蛋白質B mRNA編集酵素触媒ポリペプチド1」)、CDA(「シチジンデアミナーゼ」)、またはAID(「活性化誘導シチジンデアミナーゼ」)を包含するシチジンデアミナーゼ)ドメインとの融合体である。いくつかのケースでは、塩基編集因子は、もたらされる1ヌクレオチド変化の効率および安定性を増大させるように細胞のDNA修復プロセスを変調させる蛋白質、例えばUGIドメインをもまた包含し得る(2,3)。

20

【0004】

現在のところ、塩基編集因子の2つのクラスが一般的に記載されてきた:シチジン塩基編集因子は標的C-G塩基対をT-A塩基対に変換し、アデニン塩基編集因子はA-T塩基対をG-C塩基対に変換する。まとめると、塩基編集因子のこれらの2つのクラスは全ての4つのトランジション変異(C T、G A、A G、およびT C)の標的化されたインストールを可能化し、まとめると、これらはClinVarデータベース中の公知のヒト病原性小ヌクレオチド多型(SNP)の約61%を占める。特に、C T塩基編集因子は、シチジンデアミナーゼを用いて、Cas9によって開かれた一本鎖DNAループ中のシチジンをウリジンに変換する。反対鎖はCas9によってニックングされて、編集された鎖を鋳型として用いるDNA修復メカニズムを刺激する。一方で、融合したウラシルグリコシラーゼ阻害因子(示されていない)は編集された塩基の切り出しを遅くする。最終的に、DNA修復がC・G T・C塩基対変換に至る。

30

【0005】

塩基編集因子は多くの標的を高い効率で編集し得、多くの場合には、1回の処置後に濃縮なしに細胞の30-70%の編集を達成する。しかしながら、残念なことに、塩基編集の有用性はいくつかの制約によって限定され、用いられる特定のCas部分(例えば、S. pyogenesからの天然に存在するCas9、またはその改変されたバージョン、またはそのホモログ)によって課されるPAM要件、所望の編集部位の近傍の非標的ヌクレオチドのオフターゲット塩基編集、望まれない編集されたゲノム副産物(例えば、インデル)の産生、および低い総体的な編集効率を包含する。例えば、現行のシチジン塩基編集因子活性は標的ヌクレオチドの周囲の塩基に依存する。APOBEC1シチジンデアミナーゼに基づくC T編集因子は、

40

【数1】

TC

モチーフを編集することへの選好性を有し、ほとんどの

【数 2】

GC

塩基を敬遠する。この選好性は、古典的な編集ウィンドウ外の位置における

【数 3】

TC

の編集と、それらが最適に位置しているときでさえも

10

【数 4】

GC

標的の不良な編集とに至り得る。CDAまたはAIDなどの他のデアミナーゼを用いるC T編集因子はある

【数 5】

GC

部位については有効な代替を提供し得るが、APOBEC1 BEと比較して、より低い総体的な編集効率を有する。

20

【0006】

「次世代」塩基編集因子の開発はこれらの限定のいくつかに対処し始めており、異なるかまたは拡大されたPAM適合性を有する塩基編集因子(19-21)、縮減されたオフターゲット活性を有する高フィデリティ塩基編集因子(20,22-25)、より狭い編集ウィンドウ(正常では約5ヌクレオチド幅)を有する塩基編集因子(19)、緩められた配列コンテキスト選好性を有する塩基編集因子、および縮減された副産物を有するシチジン塩基編集因子(BE4)(6)を包含する。それでも、これらの最近の進歩にもかかわらず、塩基編集因子による塩基編集の効率は他の因子のなかでも細胞型および標的座位によって幅広く変わる。それゆえに、当分野においては、引き続いて、配列コンテキストアグノスティックな塩基編集活性による改善された編集効率を有する塩基編集因子の開発の有意な必要がある。本願は、高い効率を有しかつ配列コンテキストアグノスティックである進化した塩基編集因子を開発および産生するためのファージによって支援される連続的進化システムを記載し、それゆえに当分野の問題に対処する。

30

【発明の概要】

【0007】

発明の概要

本明細書は進化した塩基編集因子を提供し、これらは当分野のものの欠陥を克服し(増大した効率、および/または編集部位における特異的な配列コンテキストの減少した要求を包含する)、これらはファージによって支援される連続的進化(PACE)システムの結果として得られる。特に、本明細書は進化したシチジン塩基編集因子(例えば、APOBEC1、CDA、またはAIDシチジンデアミナーゼドメインに基づく)を提供し、これらは当分野のものの欠陥を克服し(増大した効率、および/または編集部位における特異的な配列コンテキストの減少した要求を包含する)、これらはファージによって支援される連続的進化(PACE)システムの結果として得られる。加えて、本明細書は、本明細書に記載される進化した塩基編集因子をコードおよび/または発現する核酸分子、ならびに本明細書に記載される進化した塩基編集因子を発現するための発現ベクターまたはコンストラクト、前記核酸分子および発現ベクターを含む宿主細胞、ならびに本明細書に記載される核酸に基づく態様を送達および/または投与するための組成物を提供する。加えて、本開示は、単離された進化した塩基編集因子、および前記単離された進化した塩基編集因子を含む組成物

40

50

を、本明細書に記載される通り提供する。なお、さらに、本開示は、進化した塩基編集因子を作る方法、および技術水準を形成する塩基編集因子と比較して改善された効率で、好ましくは配列コンテキストアグノスティックな（すなわち、所望の編集部位が特異的な配列コンテキストを要求しない）様式で、核酸分子、例えばゲノムを編集することを包含する用途に、進化した塩基編集因子または進化した塩基編集因子をコードする核酸分子を用いる方法を提供する。いくつかの態様では、本明細書において提供される作る方法は、改善されたファージによって支援される連続的進化（PACE）システムであり、これは、塩基編集因子の1つ以上の構成要素（例えば、Cas9ドメインまたはシチジンデアミナーゼドメイン）を迅速で連続的な様式で進化させるために利用され得る。本明細書は、本明細書に記載される塩基編集システムによって（例えば、本明細書に記載される単離された進化した塩基編集因子またはそれをコードするベクターもしくはコンストラクトの形態で）、標的核酸分子、例えばゲノムの1個の核酸塩基を効率的に編集し、塩基（based）編集を、好ましくは配列コンテキストアグノスティックな様式でとり行うための方法をもまた提供する。なお、さらに、本明細書は、標的核酸分子、例えばゲノムを塩基編集システムと（例えば、単離された進化した塩基編集因子蛋白質またはそれをコードするベクターの形態で）接触させることと、遺伝子疾患を処置する、および/または遺伝学的形質（例えば、目の色）を変化させるために塩基（based）編集をとり行うこととによる遺伝子疾患を処置するためおよび/または遺伝学的形質もしくは状態を変調もしくは変化させるための治療方法を提供する。

10

20

【0008】

驚くべきことに、塩基編集因子の1つ以上のドメインまたは構成要素の機能を迅速に改善するために用いられ得る有効な連続的な進化的変異導入プロセス（すなわちPACE）を開発することによって、本発明者らは、改善された（improve）塩基編集因子（例えばシチジン（cytidine）塩基編集因子）を発見した。

【0009】

それゆえに、1つの側面では、本明細書は、配列番号1のアミノ酸残基2-162と少なくとも80%、85%、90%、95%、98%、99%、または99.5%同一であるアミノ酸配列を含むシチジンデアミナーゼを提供し、シチジンデアミナーゼは、配列番号1と比べてH102X₁、D104X₂、およびV115X₃からなる群から選択される1つ以上の変異または別のシチジンデアミナーゼの対応する変異（単数もしくは複数）を含み、X₁はH以外のいずれかのアミノ酸であり（例えばH102P）、X₂はD以外のいずれかのアミノ酸であり（例えばD104N）、X₃はV以外のいずれかのアミノ酸である（例えばM）。1つの態様では、シチジンデアミナーゼは配列番号5の残基2-162を含む。シチジンデアミナーゼはN末端メチオニン（M）アミノ酸残基をもまた包含し得る。

30

40

【0010】

別の側面では、本明細書は、配列番号2のアミノ酸残基3-229と少なくとも80%、85%、90%、95%、98%、99%、または99.5%同一であるアミノ酸配列を含むシチジンデアミナーゼを提供し、シチジンデアミナーゼは、配列番号2と比べてE4X₁、V10X₂、E31X₃、Y40X₄、E95X₅、H109X₆、H122X₇、D124X₈、R126X₉、R154X₁₀、N158X₁₁、A165X₁₂、P201X₁₃、F205X₁₄、およびI208X₁₅からなる群から選択される1つ以上の変異または別のシチジンデアミナーゼの対応する変異（単数もしくは複数）を含み、X₁、X₃、およびX₅はE以外のいずれかのアミノ酸であり、X₂はV以外のいずれかのアミノ酸であり、X₄はY以外のいずれかのアミノ酸であり、X₆およびX₇はH以外のいずれかのアミノ酸であり、X₈はD以外のいずれかのアミノ酸であり、X₉およびX₁₀はR以外のいずれかのアミノ酸であり、X₁₁はN以外のいずれかのアミノ酸であり、X₁₂はA以外のいずれかのアミノ酸であり、X₁₃はP以外のいずれかのアミノ酸であり、X₁₄はF以外のいずれかのアミノ酸であり、X₁₅はI以外のいずれかのアミノ酸である。種々の態様では：X₁はKであり得；X₂およびX₅はAであり得；X₃はVであり得、X₄はCであり得；X₆およびX₈はNであり得；X₇およびX₁₅はLであり得；X₉およびX₁₀はHであり得；X₁₁、X₁₂、X₁₃；およびX₁₄はSであり得る。種々の態様では、シチジンデアミナーゼは、配列番号2と比べてE4K、H109N、H122L、D124N、R154H、A165S、P201S、およびF205S

50

を包含し得る2、3、4、5、6、7、または全ての8つの変異を含み得る。1つの態様では、シチジンデアミナーゼは配列番号6のアミノ酸配列を含む。

【0011】

なお別の態様では、本明細書は、配列番号3のアミノ酸残基2-208と少なくとも80%、85%、90%、95%、98%、99%、または99.5%同一であるアミノ酸配列を含むシチジンデアミナーゼを提供し、シチジンデアミナーゼは、配列番号3と比べてH10X₁、F23X₂、V75X₃、K120X₄、A123X₅、C158X₆、I193X₇、I195X₈、およびV197X₉からなる群から選択される1つ以上の変異または別のシチジンデアミナーゼの対応する変異（単数もしくは複数）を含み、X₁はH以外のいずれかのアミノ酸であり、X₂はF以外のいずれかのアミノ酸であり、X₃およびX₉はV以外のいずれかのアミノ酸であり、X₄はK以外のいずれかのアミノ酸であり、X₅はA以外のいずれかのアミノ酸であり、X₆はC以外のいずれかのアミノ酸であり、X₇およびX₈はI以外のいずれかのアミノ酸である。種々の態様では：X₁はYであり得；X₂はSであり得；X₃はIであり得；X₄およびX₆はRであり得；X₅はVであり得；X₇はTであり得；X₈はFまたはTであり得；X₉はAであり得る。一態様では、シチジンデアミナーゼは配列番号3と比べて変異F23S、A123V、およびI195Fを含む。他の態様では、シチジン（cytidine）デアミナーゼはN末端メチオニン（M）アミノ酸残基を含む。好ましい態様では、シチジンデアミナーゼは配列番号7のアミノ酸配列を含む。

10

【0012】

なお別の側面では、本明細書は、配列番号4のアミノ酸残基3-229と少なくとも80%、85%、90%、95%、98%、99%、または99.5%同一であるアミノ酸配列を含むシチジンデアミナーゼを提供し、シチジンデアミナーゼは、配列番号4と比べてE4X₁、H122X₂、D124X₃、R154X₄、A165X₅、P201X₆、およびF205X₇からなる群から選択される1つ以上の変異または別のシチジンデアミナーゼの対応する変異（単数もしくは複数）を含み、X₁はE以外のいずれかのアミノ酸であり、X₂はH以外のいずれかのアミノ酸であり、X₃はD以外のいずれかのアミノ酸であり、X₄はR以外のいずれかのアミノ酸であり、X₅はA以外のいずれかのアミノ酸であり、X₆はP以外のいずれかのアミノ酸であり、X₇はF以外のいずれかのアミノ酸である。種々の態様では、シチジンデアミナーゼは、配列番号4と比べてE4X₁、H122X₂、D124X₃、R154X₄、A165X₅、P201X₆、およびF205X₇からなる群から選択される2、3、4、5、6、または全ての7つの変異を含み、X₁はE以外のいずれかのアミノ酸であり、X₂はH以外のいずれかのアミノ酸であり、X₃はD以外のいずれかのアミノ酸であり、X₄はR以外のいずれかのアミノ酸であり、X₅はA以外のいずれかのアミノ酸であり、X₆はP以外のいずれかのアミノ酸であり、X₇はF以外のいずれかのアミノ酸である。ある態様では：X₁はKであり得；X₂はLであり得；X₃はNであり得；X₄はHであり得；X₅はSであり得；X₆はSであり得；および/またはX₇はSであり得る。他の態様では、1つ以上の変異は、配列番号4と比べてE4K、H122L、D124N、R154H、A165S、P201S、およびF205Sからなる群から選択され得る。別の態様では、デアミナーゼは配列番号89、または配列番号8のアミノ酸残基3-229であり得、N末端メチオニン（M）アミノ酸残基または2つのN末端アミノ酸残基を含み得、これらはMおよびSである。

20

30

【0013】

なお別の側面では、本明細書は進化した塩基編集因子融合蛋白質を提供し：(i) 核酸によってプログラム可能なDNA結合蛋白質（napDNAbp）と；(ii) 本明細書において上の側面のいずれかに記載されているシチジンデアミナーゼと；(iii) ウラシルグリコシラーゼ阻害因子ドメイン（UGI）とを含む。融合蛋白質は2、3、4、または5つのUGIドメインを有し得る。核酸によってプログラム可能なDNA結合蛋白質（napDNAbp）はCas9ドメインであり得る。napDNAbpはCasX、CasY、Cpf1、C2c1、C2c2、C2c3、またはアルゴノート蛋白質でもまたあり得る。Cas9ドメインはヌクレアーゼ活性型Cas9、ヌクレアーゼ不活性型Cas9（dCas9）、またはCas9ニッカーゼ（nCas9）であり得る。

40

【0014】

進化した塩基編集因子融合蛋白質の種々の態様では、nCas9は、アミノ酸配列：

【化 1 - 1】

DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEA
 TRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLLEESFLVEEDKKHERHPIFGN
 IVDEVAYHEKYPTIYHLRKKLV DSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDV
 DKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLI
 ALSLGLTPNFKSNFDLAEDAQLQSKDQYDDDLNLLAQIGDQYADLFLAAKNLS DAIL
 LSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAG
 YIDGGASQEEFYKFIKPILEKMDGTEELLV KLNREDLLRKQRTFDNGSIPHQIHLGELHAI
 LRRQEDFYFPLKDNREKIEKILTRIPYVVGPLARGNSRFAWMTRKSEETITPWNFEVV
 DKGASAQSFIERMTNFDKNLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLS
 GEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLKII
 KDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWG
 RLSRKLINGIRDKQSGKTILDFLKSDFANRNFMLIHDSDLTKEDIQKAQVSGQGD SL
 HEHIANLAGSPAIKKILQTVKVVDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRE
 RMKRIIEGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDV

10

20

【化 1 - 2】

DHIVPQSFLKDDSIDNKVLTRSDKNRGSNDNPSEEVVKKMKNYWRQLLNAKLITQRK
 FDNLTKAERGGSELKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVI
 TLKSKLVSDFRKDFQFYKREINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYK
 VYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD
 KGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK KYGG
 FDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSEKPNIDFLEAKGYKEVKK
 DLIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLGSPED
 NEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKHRDKPIREQAENIHL
 FTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGD (

30

配列番号 9).

と少なくとも80%、85%、90%、95%、98%、99%、99.5%、または99.9%同一であるアミノ酸配列を有し得る。

40

【0015】

他の態様では、進化した塩基編集因子融合蛋白質のUGIドメインは、UDG活性を阻害する能力があるドメインを含み得る。UGIドメインは、アミノ酸配列：

MTNLSDIIEKETGKQLV IQESILMLPEEVVEVIGNKPESDILVHTAYDESTDENVMLLTSDAPEYKPWALVIQDSNGENK
 IKML (配列番号 10) と少なくとも80%、85%、90%、95%、98%、99%、99.5%、または99.9%同一であるアミノ酸配列を含む。

【0016】

種々の態様では、本明細書に記載される進化した塩基編集因子融合蛋白質は次の構造のいずれかを含み得る：NH₂-[シチジンデアミナーゼ]-[napDNAbp]-[UGI]-COOH；NH₂-[シチジンデアミナーゼ]-[UGI]-[napDNAbp]-COOH；NH₂-[napDNAbp]-[UGI]-[シチジンデアミナーゼ]-COOH；NH₂-[napDNAbp]-[シチジンデアミナーゼ]-[UGI]-COOH；NH₂-[UGI]-[シチジン

50

デアミナーゼ]-[napDNAbp]-COOH；およびNH₂-[UGI]-[napDNAbp]-[シチジンデアミナーゼ]-COOH；シチジンデアミナーゼは、上の側面において記載されている進化したシチジンデアミナーゼであり得、UGIはUGIドメインであり、「-」の各個は任意のリンカーを含む。

【0017】

他の態様では、本明細書に記載される進化した塩基編集因子融合蛋白質は次の構造のいずれかを含み得：NH₂-[シチジンデアミナーゼ]-[napDNAbp]-[UGI]-[UGI]-COOH；NH₂-[シチジンデアミナーゼ]-[UGI]-[napDNAbp]-[UGI]-COOH；NH₂-[シチジンデアミナーゼ]-[napDNAbp]-[UGI]-COOH；NH₂-[シチジンデアミナーゼ]-[UGI]-[napDNAbp]-[UGI]-COOH；NH₂-[シチジンデアミナーゼ]-[UGI]-[UGI]-[napDNAbp]-COOH；NH₂-[シチジンデアミナーゼ]-[UGI]-[napDNAbp]-COOH；NH₂-[シチジンデアミナーゼ]-[UGI]-[UGI]-[シチジンデアミナーゼ]-COOH；NH₂-[napDNAbp]-[UGI]-[UGI]-[シチジンデアミナーゼ]-COOH；NH₂-[napDNAbp]-[UGI]-[シチジンデアミナーゼ]-[UGI]-COOH；NH₂-[napDNAbp]-[シチジンデアミナーゼ]-[UGI]-COOH；NH₂-[シチジンデアミナーゼ]-[UGI]-COOH；NH₂-[シチジンデアミナーゼ]-[napDNAbp]-[UGI]-COOH；NH₂-[シチジンデアミナーゼ]-[UGI]-[napDNAbp]-COOH；NH₂-[シチジンデアミナーゼ]-[UGI]-[UGI]-[シチジンデアミナーゼ]-[napDNAbp]-COOH；NH₂-[シチジンデアミナーゼ]-[UGI]-COOH；NH₂-[シチジンデアミナーゼ]-[napDNAbp]-[UGI]-[シチジンデアミナーゼ]-COOH；およびNH₂-[シチジンデアミナーゼ]-[UGI]-[UGI]-[napDNAbp]-[シチジンデアミナーゼ]-COOH；シチジンデアミナーゼは、上の側面において記載されている進化したシチジンデアミナーゼであり、UGIはUGIドメインであり、「-」の各個は任意のリンカーを含む。

10

20

【0018】

リンカーは次のアミノ酸配列のいずれかを含み得る：SGGSSGGSSGSETPGTSESATPESSGGSSGS（配列番号11）；SGGSSGGSSGS（配列番号12）；SGGS（配列番号14）；またはSGGSSGGSSGS（配列番号12）。

【0019】

種々の態様では、進化した塩基編集因子融合蛋白質は、1つ以上の（例えば、2、3、4、5、6、またはより多くの）核局在配列（NLS）、例えばKRTADGSEFEPKKRKRK（配列番号13）をさらに包含し得る。

【0020】

種々の態様では、進化した塩基編集因子融合蛋白質は、配列番号15-20のいずれか1つに提示されているアミノ酸配列のいずれか1つと少なくとも80%、85%、90%、95%、98%、99%、または99.5%同一であるアミノ酸配列を含み得る。

30

【0021】

種々の他の態様では、本明細書は、進化した塩基編集因子融合蛋白質またはそれらのドメインのいずれかをコードする核酸分子を提供する。核酸配列は、哺乳類細胞、例えばHEK293Tによる発現のためにコドン最適化され得る。

【0022】

なお他の態様では、本明細書は、進化した塩基編集因子（またはそれらの1つ以上の（one more more）個々の構成要素）をコードする核酸配列の発現を駆動するための適切なプロモーターを有するベクターを提供する。

40

【0023】

種々の態様では、本発明の連続的進化/PACE法は、進化した塩基編集因子融合蛋白質の発現を2つ以上の発現ベクターへと分割することを企図し、各発現単位は進化した塩基編集因子融合蛋白質のある部分をコードする。発現される部分は分割型インテンドメインを包含し、これらは、完全に形成された進化した塩基編集因子の自己形成を蛋白質スプライシングのプロセスによって駆動する。

【0024】

他の側面では、本明細書は、複合体、本明細書に記載される進化した塩基編集因子融合蛋白質と融合蛋白質のnapDNAbpに結合したRNA、例えばガイドRNA（gRNA）とを提供する。ガイドRNAは1つのガイドRNAまたは複数のガイドRNAであり得る。RNAは長さが10-100ヌク

50

レオチドであり得、標的配列に対して相補的である少なくとも10個の一続きの (contiguous)ヌクレオチド、または標的配列に対して相補的である少なくとも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、もしくは40ヌクレオチドを含み得る。

【0025】

いくつかの態様では、その標的配列 (編集されるべき配列) はDNA配列であり、生物のゲノムを包含する。生物は原核生物、あるいは真核生物、または脊椎動物、哺乳動物、もしくはヒトであり得る。

【0026】

他の態様では、本明細書は細胞を提供し、これらは、本明細書において開示される進化したシチジンデアミナーゼ、進化した塩基編集因子融合体、本明細書において開示される複合体、それらをコードする核酸分子、または核酸分子を含むベクターを含む。

10

【0027】

なお他の態様では、本明細書はキットを提供し：本明細書において開示される融合蛋白質をコードする核酸配列と；融合蛋白質の発現を駆動する異種プロモーターとを含む核酸の核酸コンストラクトを含む。キットはガイドRNAバックボーンをコードする発現コンストラクトをもまた含み得、コンストラクトは、ガイドRNAバックボーンへの標的配列と同一または相補的な核酸配列のクローニングを許すように位置したクローニング部位を含む。

【0028】

20

他の態様では、本明細書は医薬組成物を提供し、本明細書に記載される進化した塩基編集因子融合蛋白質および薬学的に許容し得る賦形剤、ならびに任意にカチオン性脂質などの脂質を含む。医薬組成物はポリマーをもまた (allos) 含み得る。

【0029】

本明細書は、標的ヌクレオチド配列 (例えばゲノム) を編集するために、進化した塩基編集因子またはそれをコードする核酸分子を包含する本明細書に記載される組成物を用いる方法をもまた提供する。標的ヌクレオチド配列は疾患または異常に関連する標的配列 (例えば点変異) を含み得る。標的配列は疾患または異常に関連するT C点変異を含み得、変異体C塩基の脱アミノ化は疾患または異常に関連しない配列をもたらず。標的配列は蛋白質をコードし得、ここで、点変異はコドンにあり、野生型コドンと比較して変異体コドンによってコードされるアミノ酸の変化をもたらず。標的配列はスプライシング部位にもまたあり得、点変異は野生型転写物と比較してmRNA転写物のスプライシングの変化をもたらず。加えて、標的は遺伝子のプロモーターにあり得、点変異は遺伝子の増大または減少した発現をもたらず。

30

【0030】

種々の態様では、進化した塩基編集因子は標的部位の脱アミノ化をもたらず。いくつかのケースでは、変異体Cの脱アミノ化は変異体コドンによってコードされるアミノ酸の変化をもたらし、これは、いくつかのケースでは、野生型アミノ酸の発現をもたらし得る。脱アミノ化はmRNA転写物の変化をもまたもたらし得、mRNA転写物を野生型状態へと回復さえもさせる。

40

【0031】

進化した塩基編集因子を標的ヌクレオチド配列と接触させることが関わる本明細書に記載される方法は、*in vitro*でまたは*in vivo*で対象に起こり得る。対象は、ApoE遺伝子の点変異に関連する疾患または異常などの疾患または異常と診断された誰かであり得る。

【0032】

本発明の方法は、デグロンタグに融合させられているRNAポリメラーゼ (例えば、T7 RNAポリメラーゼ) にもまた関する。

【0033】

本発明の方法は、NインテインまたはCインテインに融合させられた塩基編集因子のシチ

50

ジンデアミナーゼと、相補的な (complementary) NインテインまたはCインテインドメインに融合される塩基編集因子の残りを含む独立して発現される融合蛋白質とを含む融合蛋白質にもまた関し、その結果、NおよびCインテインドメインに作用する細胞の蛋白質スプライシング機構の結果として、同じ細胞によるシチジン (cytidine) デアミナーゼの発現および塩基編集因子の残りの発現は完全な塩基編集因子をもたらす。シチジン (cytidine) デアミナーゼはAPOBECデアミナーゼであり得る。

【0034】

本明細書は、分割型インテインドメインに融合される完全な塩基編集因子の第1の部分をコードする核酸分子を含む第1の発現ベクターと、コグネイトな分割型インテインドメインに融合される完全な塩基編集因子の第2の部分をコードする核酸分子を含む第2の発現ベクターとをもまた提供し、細胞によって発現されるときに、細胞の蛋白質スプライシング機構は、分割型インテインドメイン同士をつなぎ合わせることでそれからその爾後の取り除きとによって、完全な塩基編集因子を形成し、第1および第2の部分をつなぎ合わせる残留するペプチド結合を後に残す。

10

【0035】

本明細書は、ファージに基づく連続的な指向的進化のためのベクターシステムをもまた提供し：

【0036】

a. シトシンを脱アミノ化する能力がある塩基編集因子蛋白質をコードする核酸を含むベクターと；

20

【0037】

b. 完全な塩基編集因子の残りの部分をコードする第2のベクターと；

【0038】

c. luxABおよびガイドRNA (gRNA) をコードするベクターと、を含む。

【0039】

本明細書は、核酸の連続的進化の方法をもまた提供し：(i) 進化させられるべき目当ての遺伝子 (例えば、デアミナーゼ) を含むセレクションファージミド (SP) をラグーン中の細菌宿主細胞のフローに導入することと

【0040】

(宿主細胞は、セレクションファージミドを感染性のファージ粒子へとパッケージングするために要求されるファージ遺伝子を含み、セレクションファージミドをファージ粒子へとパッケージングするために要求される少なくとも1つの遺伝子が欠損しており、

30

【0041】

セレクションファージミドを感染性のファージ粒子へとパッケージングするために要求される少なくとも1つの遺伝子は、宿主細胞において進化させられるべき遺伝子の発現に応答して発現され、

【0042】

ラグーン中の宿主細胞の流量は、ラグーン中の宿主細胞ではなくファージミドの複製を許す)；

40

【0043】

(ii) ホスト細胞のフロー中のファージミドを複製および変異させることと；

【0044】

(iii) 進化させられるべき変異した遺伝子を含むファージミドを細胞のフローから単離することと、を含む。

【0045】

宿主細胞は、宿主細胞には欠損しているセレクションファージミドをファージ粒子へとパッケージングするために要求される遺伝子を含む第1のアクセサリプラスミド (AP) を含み得、遺伝子は、SPによってコードされる融合蛋白質の発現に応答してアクセサ

50

リープラスミドから発現される。

【0046】

ある態様では、標的は生物のゲノム中である。

【0047】

ある態様では、生物は原核生物である。ある態様では、生物は真核生物である。ある態様では、生物は脊椎動物である。ある態様では、脊椎動物は哺乳動物である。ある態様では、哺乳動物はヒトである。

【0048】

1つの側面では、本明細書は、本開示の進化した塩基編集因子融合蛋白質のいずれか1つを含む細胞を開示する。

【0049】

1つの側面では、本明細書は、本開示の核酸のいずれか1つを含む細胞を開示する。

【0050】

1つの側面では、本明細書は、本開示のベクターのいずれか1つを含む細胞を開示する。

。

【0051】

1つの側面では、本明細書は、本開示の複合体のいずれか1つを含む細胞を開示する。

【0052】

1つの側面では、本明細書は、核酸分子を本開示の複合体のいずれかと接触させることを含む方法を開示する。ある態様では、核酸はDNAである。ある態様では、核酸は二本鎖DNAである。ある態様では、核酸は疾患または異常に関連する標的配列を含む。ある態様では、標的配列は疾患または異常に関連する点変異を含む。

【0053】

ある態様では、標的配列は疾患または異常に関連するT C点変異を含み、変異体C塩基の脱アミノ化は疾患または異常に関連しない配列をもたらす。ある態様では、標的配列は疾患または異常に関連するG A点変異を含み、変異体C塩基の脱アミノ化は疾患または異常に関連しない配列をもたらす。

【0054】

ある態様では、所望の編集が起こるべき標的配列は配列アグノスティックである。つまり、本明細書に記載される進化した塩基編集因子は、標的編集部位における特異的な配列コンテキストを要求することなしに、効率的で正確な編集を行い得る (may carrying out)。

【0055】

ある態様では、標的配列は蛋白質をコードし、点変異はコドンにあり、野生型コドンと比較して変異体コドンによってコードされるアミノ酸の変化をもたらす。ある態様では、標的配列はスプライシング部位にあり、点変異は野生型転写物と比較してmRNA転写物のスプライシングの変化をもたらす。ある態様では、標的配列は遺伝子のプロモーターにあり、点変異は遺伝子の増大した発現をもたらす。ある態様では、標的配列は遺伝子のプロモーターにあり、点変異は遺伝子の減少した発現をもたらす。

【0056】

ある態様では、変異体Cまたは変異体Aの脱アミノ化は変異体コドンによってコードされるアミノ酸の変化をもたらす。ある態様では、変異体Cまたは変異体Aの脱アミノ化は野生型アミノ酸をコードするコドンをもたらす。ある態様では、変異体Cまたは変異体Aの脱アミノ化はmRNA転写物の変化をもたらす。ある態様では、変異体Cまたは変異体Aの脱アミノ化は野生型のmRNA転写物をもたらす。ある態様では、変異体Cまたは変異体Aの脱アミノ化は遺伝子の増大した発現をもたらす。ある態様では、変異体Cまたは変異体Aの脱アミノ化は遺伝子の減少した発現をもたらす。

【0057】

ある態様では、接触させることは *in vitro* で実行される。ある態様では、接触させることは *in vivo* で対象に実行される。

10

20

30

40

50

【 0 0 5 8 】

ある態様では、対象は疾患または異常と診断されている。ある態様では、疾患または異常は、先天性糖化異常症1f型、家族性肢端紅痛症、発作性激痛症、慢性無痛症、鎌状赤血球貧血、およびサラセミアからなる群から選択される。ある態様では、疾患または異常はMDPU1遺伝子の点変異に関連する。ある態様では、疾患または異常はSCN9a遺伝子の点変異に関連する。ある態様では、疾患または異常はHBG1および/またはHBG2遺伝子の発現を増大させることによって処置され得る。

【 0 0 5 9 】

1つの側面では、本明細書はキットを開示し、(a)本開示の融合蛋白質のいずれか1つをコードする核酸配列と；(b)(a)の配列の発現を駆動する異種プロモーターとを含む核酸コンストラクトを含む。ある態様では、キットは、ガイドRNAバックボーンをコードする発現コンストラクトをさらに含み、コンストラクトは、ガイドRNAバックボーンへの標的配列と同一または相補的な核酸配列のクローニングを許すように位置したクローニング部位を含む。

10

【 0 0 6 0 】

1つの側面では、本明細書は本開示の融合蛋白質のいずれか1つを含む医薬組成物を開示する。

【 0 0 6 1 】

1つの側面では、本明細書は本開示の複合体のいずれか1つを含む医薬組成物を開示する。

20

【 0 0 6 2 】

1つの側面では、本明細書は本開示の核酸のいずれか1つを含む医薬組成物を開示する。

【 0 0 6 3 】

1つの側面では、本明細書は本開示のベクターのいずれか1つを含む医薬組成物を開示する。ある態様では、医薬組成物はさらに薬学的に許容し得る賦形剤を含む。ある態様では、医薬組成物はさらに脂質を含む。ある態様では、脂質はカチオン性脂質である。ある態様では、医薬組成物はさらにポリマーを含む。

【 0 0 6 4 】

ある態様では、融合蛋白質は構造： NH_2 -[本開示のシチジンデアミナーゼのいずれか]-[Cas9ドメイン]-COOHを含み、「-」の各個は任意のリンカーを含む。ある態様では、融合蛋白質は構造： NH_2 -[本開示のシチジンデアミナーゼのいずれか]-[Cas9ドメイン]-[UGIドメイン]-COOHを含み、「-」の各個は任意のリンカーを含む。ある態様では、融合蛋白質は構造： NH_2 -[本開示のシチジンデアミナーゼのいずれか]-[Cas9ドメイン]-[第1のUGIドメイン]-[第2のUGIドメイン]-COOHを含み、「-」の各個は任意のリンカーを含む。ある態様では、融合蛋白質は構造： NH_2 -[本開示のシチジンデアミナーゼのいずれか]-[Cas9ドメイン]-[核局在配列]-COOHを含み、「-」の各個は任意のリンカーを含む。ある態様では、融合蛋白質は構造： NH_2 -[第1の核局在配列]-[本開示のシチジンデアミナーゼのいずれか]-[Cas9ドメイン]-[第2の核局在配列]-COOHを含み、「-」の各個は任意のリンカーを含む。ある態様では、融合蛋白質は構造： NH_2 -[第1の核局在配列]-[本開示のシチジンデアミナーゼのいずれか]-[Cas9ドメイン]-[第1のUGIドメイン]-[第2のUGIドメイン]-[第2の核局在配列]-COOHを含み、「-」の各個は任意のリンカーを含む。

30

40

【 0 0 6 5 】

1つの側面では、本明細書は本開示の融合蛋白質のいずれかをコードする核酸を開示する。ある態様では、核酸は本開示の核酸のいずれかを含む。

【 0 0 6 6 】

本開示はこれについては限定されないもので、上述の概念および下で論ぜられる追加の概念はいずれかの好適な組み合わせで構成され得るといふことは了解されるはずである。さらに、本開示の他の利点および新規の特徴は、付随する図と併せて考慮されるときに、種々の限定しない態様の次の詳細な記載から明らかになるであろう。

50

- 【図面の簡単な説明】
- 【0067】
- 【図1】図1はCas9ゲノム編集およびC T塩基編集の概略図を示す。
- 【0068】
- 【図2】図2はPACEの概略図を示す。
- 【0069】
- 【図3】図3は塩基編集なし対塩基編集の概略図を示す。
- 【0070】
- 【図4】図4はT7RNAP発現レベルを示す。
- 【0071】 10
- 【図5】図5は分割型インテインを用いる改変されたファージバックボーンによるT7RNAP発現を示す。
- 【0072】
- 【図6】図6はAPOBECファージ増殖が24時間以内に優勢になるということを示す。
- 【0073】
- 【図7-1】図7は低ストリンジェンシー回路上におけるPACEにおける増殖を示す。
- 【図7-2】図7は低ストリンジェンシー回路上におけるPACEにおける増殖を示す。
- 【0074】
- 【図8】図8はgIII組換え体野生型対T7RNAP組換え体の概略図を示す。
- 【0075】 20
- 【図9】図9は塩基編集の概略図を示す。
- 【0076】
- 【図10-1】図10はEdgX-RpoZ融合体によるスクリーニングを示す。
- 【図10-2】図10はEdgX-RpoZ融合体によるスクリーニングを示す。
- 【0077】
- 【図11】図11はClinVARデータベース病原性バリエーションを示す。
- 【0078】
- 【図12】図12はCas9ゲノム編集およびC T塩基編集の概略図を示す。
- 【0079】
- 【図13】図13は塩基編集結果を示す。
- 【0080】 30
- 【図14】図14はPam特異性の概略図を示す。
- 【0081】
- 【図15】図15は発現および活性の概略図を示す。
- 【0082】
- 【図16】図16は提案される塩基進化方法の概略図を示す。
- 【0083】
- 【図17】図17は進化プロセスの概略図を示す。
- 【0084】
- 【図18】図18は*in vivo*の連続的な指向的進化の概略図を示す。
- 【0085】 40
- 【図19】図19は*in vivo*の連続的な指向的進化の概略図を示す。
- 【0086】
- 【図20】図20はファージによって支援される連続的進化の概略図を示す。
- 【0087】
- 【図21】図21はファージによって支援される連続的進化の概略図を示す。
- 【0088】
- 【図22】図22は塩基編集への使用のためのPACEセレクションの概略図を示す。
- 【0089】
- 【図23】図23は塩基編集への使用のためのPACEセレクションの概略図を示す。
- 50

- 【 0 0 9 0 】
- 【 図 2 4 】 図24は塩基編集への使用のためのPACEセレクションの概略図を示す。
- 【 0 0 9 1 】
- 【 図 2 5 】 図25はT7 RNAP活性のC末端デグロンタグカットの概略図を示す。
- 【 0 0 9 2 】
- 【 図 2 6 】 図26はT7 RNAPのC末端の蛋白質モデルを示す。
- 【 0 0 9 3 】
- 【 図 2 7 】 図27は塩基編集なしおよび塩基編集両方におけるC末端デグロンタグの概略図を示す。
- 【 0 0 9 4 】 10
- 【 図 2 8 】 図28はT7RNAPの活性化および塩基編集結果の概略図を示す。
- 【 0 0 9 5 】
- 【 図 2 9 】 図29は恒常的なT7RNAP発現および塩基編集結果の概略図を示す。
- 【 0 0 9 6 】
- 【 図 3 0 】 図30は塩基編集因子ファージの概略図を示す。
- 【 0 0 9 7 】
- 【 図 3 1 】 図31はGolden Gateクローニングを用いる塩基編集因子ファージアセンブリの概略図を示す。
- 【 0 0 9 8 】
- 【 図 3 2 】 図32は第1世代SPを示す。 20
- 【 0 0 9 9 】
- 【 図 3 3 】 図33はセレクションおよび最適化の概略図を示す。
- 【 0 1 0 0 】
- 【 図 3 4 】 図34はインテインによって分割された塩基編集因子およびgRNA依存的な増殖の概略図を示す。
- 【 0 1 0 1 】
- 【 図 3 5 】 図35はPACEにおけるインテインによって分割された塩基編集因子の概略図を示す。
- 【 0 1 0 2 】
- 【 図 3 6 】 図36は活性な編集因子の結果を示す。 30
- 【 0 1 0 3 】
- 【 図 3 7 - 1 】 図37はインテインによって分割された塩基編集因子に観察された変異を示す。
- 【 図 3 7 - 2 】 図37はインテインによって分割された塩基編集因子に観察された変異を示す。
- 【 0 1 0 4 】
- 【 図 3 8 】 図38は組換えによって発生する裏切り体の概略図を示す。
- 【 0 1 0 5 】
- 【 図 3 9 】 図39は組換えによって発生する裏切り体の概略図を示す。
- 【 0 1 0 6 】 40
- 【 図 4 0 - 1 】 図40はPACEデータを示す。
- 【 図 4 0 - 2 】 図40はPACEデータを示す。
- 【 0 1 0 7 】
- 【 図 4 1 】 図41はルシフェラーゼ読み出しによってアッセイされたファージ塩基編集活性を示す。
- 【 0 1 0 8 】
- 【 図 4 2 】 図42はルシフェラーゼ時系列アッセイを示す。
- 【 0 1 0 9 】
- 【 図 4 3 】 図43はPACE 5最上位15クローンのルシフェラーゼ時系列アッセイを示す。
- 【 0 1 1 0 】 50

- 【図44】図44はPACE 5最上位8クローンのルシフェラーゼ時系列アッセイを示す。
【0111】
- 【図45】図45は概要および将来の方向性を示す。
【0112】
- 【図46】図46はPACEアウトプットを「クリーニング」することの概略図を示す。
【0113】
- 【図47】図47は「1.5ハイブリッド」セレクションの概略図を示す。
【0114】
- 【図48】図48は「1.5ハイブリッド」セレクションの概略図を示す。
【0115】 10
- 【図49】図49は転写活性化因子リクルートの概略図を示す。蛋白質配列は配列番号46に対応する。
【0116】
- 【図50】図50は転写活性化因子リクルートの概略図を示す。
【0117】
- 【図51】図51は転写活性化因子リクルートの概略図を示す。
【0118】
- 【図52】図52は編集に対して応答性の塩基を示す。凡例中の蛋白質配列は上から下に配列番号99、100、および99に対応する。
【0119】 20
- 【図53】図53は編集に対して応答性の塩基を示す。蛋白質配列は配列番号100に対応する。
【0120】
- 【図54】図54は編集に対して応答性の塩基を示す。蛋白質配列は配列番号99に対応する。
【0121】
- 【図55】図55は編集に対して応答性の塩基を示す。蛋白質配列は配列番号99に対応する。
【0122】
- 【図56】図56は編集に対して応答性の塩基を示す。蛋白質配列は配列番号14に対応する。
【0123】 30
- 【図57】図57は高い活性化因子発現による最大の応答を示す。蛋白質配列は配列番号14に対応する。
【0124】
- 【図58】図58はPACEによって進化したファージ回路最適化を示す。
【0125】
- 【図59】図59は鋳型鎖 (stand) 編集の概略図を示す。
【0126】
- 【図60】図60は鋳型鎖 (stand) 編集の概略図を示す。
【0127】 40
- 【図61】図61は鋳型鎖 (stand) 編集の概略図を示す。
【0128】
- 【図62】図62は鋳型鎖 (stand) 編集の概略図を示す。
【0129】
- 【図63】図63は鋳型鎖 (stand) 編集の概略図を示す。
【0130】
- 【図64】図64は鋳型鎖脱アミノ化によって観察されるサンプルコード変異を示す。配列は上から下に配列番号95-98に対応する。
【0131】 50

- 【図65】図65はより強いT7RNAP発現の結果を示す。
【0132】
- 【図66】図66はE158K変異の概略図を示す。
【0133】
- 【図67】図67はE158K変異の概略図を示す。
【0134】
- 【図68】図68はC末端T7リゾチーム-デグロン融合体を示す。
【0135】
- 【図69】図69はより小さい編集因子パリアントの概略図および結果を示す。
【0136】 10
- 【図70】図70はプラスミドによってコードされる編集因子の概略図および結果を示す。
【0137】
- 【図71】図71はファージによってコードされる編集因子の概略図を示す。
【0138】
- 【図72】図72はファージゲノム対プラスミドからの発現の比較を示す。
【0139】
- 【図73】図73はT7 RNAP-デグロン発現を縮減することの結果を示しており、より高いターンオンに至る。
【0140】
- 【図74】図74はファージバックボーン変異が有意な役割を演ずるということを示す。
【0141】 20
- 【図75】図75は進化したファージがより良く増殖するということを示す。
【0142】
- 【図76】図76は非鋳型鎖編集が回路を活性化せずであり得るということを示す。
【0143】
- 【図77】図77は5'AGがCDAではなくAPOBEC編集を縮減するということを示す。
【0144】
- 【図78】図78は5'塩基アグノスティックなAPOBECを進化させるPCE 12を示す。
【0145】
- 【図79】図79は分割型BEコンストラクトを用いる初期のPACEによって進化したデアミナーゼのHEK細胞編集を示しており、野生型BEからの改善を示す。
【0146】 30
- 【図80】図80は分割型BEコンストラクトを用いる初期のPACEによって進化したデアミナーゼのHEK細胞編集を示しており、野生型BEからの改善を示す。
【0147】
- 【図81】図81は分割型BEコンストラクトを用いる初期のPACEによって進化したデアミナーゼのHEK細胞編集を示しており、野生型BEからの改善を示す。
【0148】
- 【図82】図82は分割型BEコンストラクトを用いる初期のPACEによって進化したデアミナーゼのHEK細胞編集を示しており、野生型BEからの改善を示す。
【0149】 40
- 【図83】図83は分割型BEコンストラクトを用いる初期のPACEによって進化したデアミナーゼのHEK細胞編集を示しており、野生型BEからの改善を示す。
【0150】
- 【図84】図84は細菌回路活性化が改善されたGCC編集を示すということを示す。
【0151】
- 【図85】図85はPACEによって進化したデアミナーゼの実験的な詳細を示す。これらはHEK細胞編集の改善を示す。
【0152】
- 【図86】図86はNC標的を示す。
【0153】 50

【 0 1 5 3 】

【 図 8 7 】 図87はNC標的を示す。

【 0 1 5 4 】

【 図 8 8 】 図88はEMX1データを示しており、低いトランスフェクションドーズ（30ng）でのHEK細胞における進化したvs.野生型のデアミナーゼの改善された活性を示す。

【 0 1 5 5 】

【 図 8 9 】 図89はRNF2データを示しており、低いトランスフェクションドーズ（30ng）でのHEK細胞における進化したvs.野生型のデアミナーゼの改善された活性を示す。

【 0 1 5 6 】

【 図 9 0 】 図90はHEK4データを示しており、低いトランスフェクションドーズ（30ng）でのHEK細胞における進化したvs.野生型のデアミナーゼの改善された活性を示す。

10

【 0 1 5 7 】

【 図 9 1 】 図91はHEK2データを示しており、低いトランスフェクションドーズ（30ng）でのHEK細胞における進化したvs.野生型のデアミナーゼの改善された活性を示す。

【 0 1 5 8 】

【 図 9 2 】 図92はHEK3データを示しており、低いトランスフェクションドーズ（30ng）でのHEK細胞における進化したvs.野生型のデアミナーゼの改善された活性を示す。

【 0 1 5 9 】

【 図 9 3 】 図93はEMX1 GC₃について750ngトランスフェクションHEK細胞編集を示す。HEK細胞編集：750ng塩基編集因子、250ngガイドRNA；リポフェクション；3日インキュベーション。指示されているところを例外としてN=3；エラーバー=SD。矢印はさらなる研究のために選択された野生型および「evo」遺伝子型を示す。

20

【 0 1 6 0 】

【 図 9 4 】 図94はEMX1 CC₂について750ngトランスフェクションHEK細胞編集を示す。HEK細胞編集：750ng塩基編集因子、250ngガイドRNA；リポフェクション；3日インキュベーション。指示されているところを例外としてN=3；エラーバー=SD。矢印はさらなる研究のために選択された野生型および「evo」遺伝子型を示す。

【 0 1 6 1 】

【 図 9 5 】 図95はEMX1 TC₅について750ngトランスフェクションHEK細胞編集を示す。HEK細胞編集：750ng塩基編集因子、250ngガイドRNA；リポフェクション；3日インキュベーション。指示されているところを例外としてN=3；エラーバー=SD。矢印はさらなる研究のために選択された野生型および「evo」遺伝子型を示す。

30

【 0 1 6 2 】

【 図 9 6 】 図96はEMX1 CC₆について750ngトランスフェクションHEK細胞編集を示す。HEK細胞編集：750ng塩基編集因子、250ngガイドRNA；リポフェクション；3日インキュベーション。指示されているところを例外としてN=3；エラーバー=SD。矢印はさらなる研究のために選択された野生型および「evo」遺伝子型を示す。

【 0 1 6 3 】

【 図 9 7 】 図97はEMX1 GC₁₀について750ngトランスフェクションHEK細胞編集を示す。HEK細胞編集：750ng塩基編集因子、250ngガイドRNA；リポフェクション；3日インキュベーション。指示されているところを例外としてN=3；エラーバー=SD。矢印はさらなる研究のために選択された野生型および「evo」遺伝子型を示す。

40

【 0 1 6 4 】

【 図 9 8 】 図98はHEK2 AC₃について750ngトランスフェクションHEK細胞編集を示す。HEK細胞編集：750ng塩基編集因子、250ngガイドRNA；リポフェクション；3日インキュベーション。指示されているところを例外としてN=3；エラーバー=SD。矢印はさらなる研究のために選択された野生型および「evo」遺伝子型を示す。

【 0 1 6 5 】

【 図 9 9 】 図99はHEK2 AC₄について750ngトランスフェクションHEK細胞編集を示す。HEK細胞編集：750ng塩基編集因子、250ngガイドRNA；リポフェクション；3日インキュベーション

50

ョン。指示されているところを例外としてN=3；エラーバー=SD。矢印はさらなる研究のために選択された野生型および「evo」遺伝子型を示す。

【0166】

【図100】図100はHEK2 AC₆について750ngトランスフェクションHEK細胞編集を示す。HEK細胞編集：750ng塩基編集因子、250ngガイドRNA；リポフェクション；3日インキュベーション。指示されているところを例外としてN=3；エラーバー=SD。矢印はさらなる研究のために選択された野生型および「evo」遺伝子型を示す。

【0167】

【図101】図101はHEK2 GC₁₁について750ngトランスフェクションHEK細胞編集を示す。HEK細胞編集：750ng塩基編集因子、250ngガイドRNA；リポフェクション；3日インキュベーション。指示されているところを例外としてN=3；エラーバー=SD。矢印はさらなる研究のために選択された野生型および「evo」遺伝子型を示す。

10

【0168】

【図102】図102はHEK3 GC₃について750ngトランスフェクションHEK細胞編集を示す。HEK細胞編集：750ng塩基編集因子、250ngガイドRNA；リポフェクション；3日インキュベーション。指示されているところを例外としてN=3；エラーバー=SD。矢印はさらなる研究のために選択された野生型および「evo」遺伝子型を示す。

【0169】

【図103】図103はHEK3 CC₄について750ngトランスフェクションHEK細胞編集を示す。HEK細胞編集：750ng塩基編集因子、250ngガイドRNA；リポフェクション；3日インキュベーション。指示されているところを例外としてN=3；エラーバー=SD。矢印はさらなる研究のために選択された野生型および「evo」遺伝子型を示す。

20

【0170】

【図104】図104はHEK3 CC₅について750ngトランスフェクションHEK細胞編集を示す。HEK細胞編集：750ng塩基編集因子、250ngガイドRNA；リポフェクション；3日インキュベーション。指示されているところを例外としてN=3；エラーバー=SD。矢印はさらなる研究のために選択された野生型および「evo」遺伝子型を示す。

【0171】

【図105】図105はHEK3 AC₉について750ngトランスフェクションHEK細胞編集を示す。HEK細胞編集：750ng塩基編集因子、250ngガイドRNA；リポフェクション；3日インキュベーション。指示されているところを例外としてN=3；エラーバー=SD。矢印はさらなる研究のために選択された野生型および「evo」遺伝子型を示す。

30

【0172】

【図106】図106はHEK3 GC₁₄について750ngトランスフェクションHEK細胞編集を示す。HEK細胞編集：750ng塩基編集因子、250ngガイドRNA；リポフェクション；3日インキュベーション。指示されているところを例外としてN=3；エラーバー=SD。矢印はさらなる研究のために選択された野生型および「evo」遺伝子型を示す。

【0173】

【図107】図107はHEK3 AC₁₆について750ngトランスフェクションHEK細胞編集を示す。HEK細胞編集：750ng塩基編集因子、250ngガイドRNA；リポフェクション；3日インキュベーション。指示されているところを例外としてN=3；エラーバー=SD。矢印はさらなる研究のために選択された野生型および「evo」遺伝子型を示す。

40

【0174】

【図108】図108はHEK4 GC₃について750ngトランスフェクションHEK細胞編集を示す。HEK細胞編集：750ng塩基編集因子、250ngガイドRNA；リポフェクション；3日インキュベーション。指示されているところを例外としてN=3；エラーバー=SD。矢印はさらなる研究のために選択された野生型および「evo」遺伝子型を示す。

【0175】

【図109】図109はHEK4 AC₅について750ngトランスフェクションHEK細胞編集を示す。HEK細胞編集：750ng塩基編集因子、250ngガイドRNA；リポフェクション；3日インキュベ

50

ション。指示されているところを例外としてN=3；エラーバー=SD。矢印はさらなる研究のために選択された野生型および「evo」遺伝子型を示す。

【0176】

【図110】図110はHEK4 GC₈について750ngトランスフェクションHEK細胞編集を示す。HEK細胞編集：750ng塩基編集因子、250ngガイドRNA；リポフェクション；3日インキュベーション。指示されているところを例外としてN=3；エラーバー=SD。矢印はさらなる研究のために選択された野生型および「evo」遺伝子型を示す。

【0177】

【図111】図111はHEK4 GC₁₁について750ngトランスフェクションHEK細胞編集を示す。HEK細胞編集：750ng塩基編集因子、250ngガイドRNA；リポフェクション；3日インキュベーション。指示されているところを例外としてN=3；エラーバー=SD。矢印はさらなる研究のために選択された野生型および「evo」遺伝子型を示す。

10

【0178】

【図112】図112はRNF2 TC₃について750ngトランスフェクションHEK細胞編集を示す。HEK細胞編集：750ng塩基編集因子、250ngガイドRNA；リポフェクション；3日インキュベーション。指示されているところを例外としてN=3；エラーバー=SD。矢印はさらなる研究のために選択された野生型および「evo」遺伝子型を示す。

【0179】

【図113】図113はRNF2 TC₆について750ngトランスフェクションHEK細胞編集を示す。HEK細胞編集：750ng塩基編集因子、250ngガイドRNA；リポフェクション；3日インキュベーション。指示されているところを例外としてN=3；エラーバー=SD。矢印はさらなる研究のために選択された野生型および「evo」遺伝子型を示す。

20

【0180】

【図114】図114はRNF2 TC₁₂について750ngトランスフェクションHEK細胞編集を示す。HEK細胞編集：750ng塩基編集因子、250ngガイドRNA；リポフェクション；3日インキュベーション。指示されているところを例外としてN=3；エラーバー=SD。矢印はさらなる研究のために選択された野生型および「evo」遺伝子型を示す。

【0181】

【図115】図115はEMX1 GC₃について750ngトランスフェクションHEK細胞編集を示す。HEK細胞編集：750ng塩基編集因子、250ngガイドRNA；リポフェクション；3日インキュベーション。指示されているところを例外としてN=3；エラーバー=SD。矢印はさらなる研究のために選択された野生型および「evo」遺伝子型を示す。

30

【0182】

【図116】図116はEMX1 CC₂について750ngトランスフェクションHEK細胞編集を示す。HEK細胞編集：750ng塩基編集因子、250ngガイドRNA；リポフェクション；3日インキュベーション。指示されているところを例外としてN=3；エラーバー=SD。矢印はさらなる研究のために選択された野生型および「evo」遺伝子型を示す。

【0183】

【図117】図117はEMX1 TC₅について750ngトランスフェクションHEK細胞編集を示す。HEK細胞編集：750ng塩基編集因子、250ngガイドRNA；リポフェクション；3日インキュベーション。指示されているところを例外としてN=3；エラーバー=SD。矢印はさらなる研究のために選択された野生型および「evo」遺伝子型を示す。

40

【0184】

【図118】図118はEMX1 CC₆について750ngトランスフェクションHEK細胞編集を示す。HEK細胞編集：750ng塩基編集因子、250ngガイドRNA；リポフェクション；3日インキュベーション。指示されているところを例外としてN=3；エラーバー=SD。矢印はさらなる研究のために選択された野生型および「evo」遺伝子型を示す。

【0185】

【図119】図119はEMX1 GC₁₀について750ngトランスフェクションHEK細胞編集を示す。HEK細胞編集：750ng塩基編集因子、250ngガイドRNA；リポフェクション；3日インキュベ

50

ーション。指示されているところを例外としてN=3；エラーバー=SD。矢印はさらなる研究のために選択された野生型および「evo」遺伝子型を示す。

【0186】

【図120】図120はHEK2 AC₃について750ngトランスフェクションHEK細胞編集を示す。HEK細胞編集：750ng塩基編集因子、250ngガイドRNA；リポフェクション；3日インキュベーション。指示されているところを例外としてN=3；エラーバー=SD。矢印はさらなる研究のために選択された野生型および「evo」遺伝子型を示す。

【0187】

【図121】図121はHEK2 AC₄について750ngトランスフェクションHEK細胞編集を示す。HEK細胞編集：750ng塩基編集因子、250ngガイドRNA；リポフェクション；3日インキュベーション。指示されているところを例外としてN=3；エラーバー=SD。矢印はさらなる研究のために選択された野生型および「evo」遺伝子型を示す。

10

【0188】

【図122】図122はHEK2 AC₆について750ngトランスフェクションHEK細胞編集を示す。HEK細胞編集：750ng塩基編集因子、250ngガイドRNA；リポフェクション；3日インキュベーション。指示されているところを例外としてN=3；エラーバー=SD。矢印はさらなる研究のために選択された野生型および「evo」遺伝子型を示す。

【0189】

【図123】図123はHEK2 GC₁₁について750ngトランスフェクションHEK細胞編集を示す。HEK細胞編集：750ng塩基編集因子、250ngガイドRNA；リポフェクション；3日インキュベーション。指示されているところを例外としてN=3；エラーバー=SD。矢印はさらなる研究のために選択された野生型および「evo」遺伝子型を示す。

20

【0190】

【図124】図124はHEK3 GC₃について750ngトランスフェクションHEK細胞編集を示す。HEK細胞編集：750ng塩基編集因子、250ngガイドRNA；リポフェクション；3日インキュベーション。指示されているところを例外としてN=3；エラーバー=SD。矢印はさらなる研究のために選択された野生型および「evo」遺伝子型を示す。

【0191】

【図125】図125はHEK3 CC₄について750ngトランスフェクションHEK細胞編集を示す。HEK細胞編集：750ng塩基編集因子、250ngガイドRNA；リポフェクション；3日インキュベーション。指示されているところを例外としてN=3；エラーバー=SD。矢印はさらなる研究のために選択された野生型および「evo」遺伝子型を示す。

30

【0192】

【図126】図126はHEK3 CC₅について750ngトランスフェクションHEK細胞編集を示す。HEK細胞編集：750ng塩基編集因子、250ngガイドRNA；リポフェクション；3日インキュベーション。指示されているところを例外としてN=3；エラーバー=SD。矢印はさらなる研究のために選択された野生型および「evo」遺伝子型を示す。

【0193】

【図127】図127はHEK3 AC₉について750ngトランスフェクションHEK細胞編集を示す。HEK細胞編集：750ng塩基編集因子、250ngガイドRNA；リポフェクション；3日インキュベーション。指示されているところを例外としてN=3；エラーバー=SD。矢印はさらなる研究のために選択された野生型および「evo」遺伝子型を示す。

40

【0194】

【図128】図128はHEK3 GC₁₄について750ngトランスフェクションHEK細胞編集を示す。HEK細胞編集：750ng塩基編集因子、250ngガイドRNA；リポフェクション；3日インキュベーション。指示されているところを例外としてN=3；エラーバー=SD。矢印はさらなる研究のために選択された野生型および「evo」遺伝子型を示す。

【0195】

【図129】図129はHEK3 AC₁₆について750ngトランスフェクションHEK細胞編集を示す。HEK細胞編集：750ng塩基編集因子、250ngガイドRNA；リポフェクション；3日インキュベ

50

ーション。指示されているところを例外としてN=3；エラーバー=SD。矢印はさらなる研究のために選択された野生型および「evo」遺伝子型を示す。

【0196】

【図130】図130はHEK4 GC₃について750ngトランスフェクションHEK細胞編集を示す。HEK細胞編集：750ng塩基編集因子、250ngガイドRNA；リポフェクション；3日インキュベーション。指示されているところを例外としてN=3；エラーバー=SD。矢印はさらなる研究のために選択された野生型および「evo」遺伝子型を示す。

【0197】

【図131】図131はHEK4 AC₅について750ngトランスフェクションHEK細胞編集を示す。HEK細胞編集：750ng塩基編集因子、250ngガイドRNA；リポフェクション；3日インキュベーション。指示されているところを例外としてN=3；エラーバー=SD。矢印はさらなる研究のために選択された野生型および「evo」遺伝子型を示す。

10

【0198】

【図132】図132はHEK4 GC₈について750ngトランスフェクションHEK細胞編集を示す。HEK細胞編集：750ng塩基編集因子、250ngガイドRNA；リポフェクション；3日インキュベーション。指示されているところを例外としてN=3；エラーバー=SD。矢印はさらなる研究のために選択された野生型および「evo」遺伝子型を示す。

【0199】

【図133】図133はHEK4 GC₁₁について750ngトランスフェクションHEK細胞編集を示す。HEK細胞編集：750ng塩基編集因子、250ngガイドRNA；リポフェクション；3日インキュベーション。指示されているところを例外としてN=3；エラーバー=SD。矢印はさらなる研究のために選択された野生型および「evo」遺伝子型を示す。

20

【0200】

【図134】図134はRNF2 TC₃について750ngトランスフェクションHEK細胞編集を示す。HEK細胞編集：750ng塩基編集因子、250ngガイドRNA；リポフェクション；3日インキュベーション。指示されているところを例外としてN=3；エラーバー=SD。矢印はさらなる研究のために選択された野生型および「evo」遺伝子型を示す。

【0201】

【図135】図135はRNF2 TC₆について750ngトランスフェクションHEK細胞編集を示す。HEK細胞編集：750ng塩基編集因子、250ngガイドRNA；リポフェクション；3日インキュベーション。指示されているところを例外としてN=3；エラーバー=SD。矢印はさらなる研究のために選択された野生型および「evo」遺伝子型を示す。

30

【0202】

【図136】図136はRNF2 TC₁₂について750ngトランスフェクションHEK細胞編集を示す。HEK細胞編集：750ng塩基編集因子、250ngガイドRNA；リポフェクション；3日インキュベーション。指示されているところを例外としてN=3；エラーバー=SD。矢印はさらなる研究のために選択された野生型および「evo」遺伝子型を示す。

【0203】

【図137】図137は、編集ウィンドウをシフトするかまたは幅広くするためのPACEセレクションの拡張を示す。N=6の独立した宿主細胞培養物が、デアミナーゼ-インテイン融合体を含むバックボーンアイソジェニックファージによって感染された。感染後2-3時間の間の直線期における回路活性化速度（分あたりOD600あたりのルミネッセンス）が各回路についてwt-APOBECに対して正規化されている。エラーバーはSEMを示す。

40

【0204】

【図138】図138は、胚細胞における選択された進化デアミナーゼおよびコンストラクト（construct）の組み合わせの概略図を示す。上から下に、配列は配列番号126-128に対応する。

【0205】

【図139】図139は、evo-デアミナーゼが、以前には編集可能でない部位における塩基編集を許すということを示しており；BE-CDAによる編集は進化後に（8%から33%に）増大

50

した。

【0206】

【図140】図140はApoE蛋白質ドメイン構造、ApoE4 R112 sgRNA、およびApoE4 E3 (R112 C112) 編集の概略図を示す。中心のセクションでは、配列は上から下に配列番号129-131に対応する。

【0207】

【図141】図141は核酸塩基改変酵素による1塩基変化の概略図を示す。

【0208】

【図142】図142：PACEの間には、目当ての活性は、その活性をそれらのゲノム中にコードするバクテリオファージの増殖にカップリングされる(1)。このカップリングは、E. coliホスト細胞によってコードされる遺伝子回路によって奏され(2)、それらは固定体積ベッセル内に連続的に圧送され、そこでそれらはファージに感染する(3)。回路を活性化するファージがコードする遺伝子はファージ遺伝子IIIの発現に至り(4)、これはそれらが繁殖することを許し、一方で、不活性な遺伝子を有するファージはしない。感染したホスト細胞およびファージは連続的にベッセルから流出し(5)、不活性な遺伝子を希釈除去する。活性なファージゲノムは強い人工変異導入によって複製し(6)、それから感染性粒子として培養培地中に放出され(7)、その結果、それらは新たなホスト細胞に感染し得る。セクション下にある活性をコードするファージのみが十分に速く増殖して希釈速度を克服し得る。

【0209】

【図143-1】図143は、塩基編集セレクション回路の活性化が全ての回路構成要素と全長の塩基編集因子蛋白質およびオンターゲットガイドRNAの存在とに依存するというを示す。

【図143-2】図143は、塩基編集セレクション回路の活性化が全ての回路構成要素と全長の塩基編集因子蛋白質およびオンターゲットガイドRNAの存在とに依存するというを示す。

【0210】

【図144】図144はセレクション回路上におけるTCCおよびGCC標的による進化したデアミナーゼファージの相対的活性を示す。

【0211】

定義

本明細書および請求項において用いられる単数形「a」、「an」、および「the」は、コンテキストが明瞭に別様に指示しない限り、単数および複数の参照を包含する。それゆえに、例えば、「薬剤」の参照は1個の薬剤および複数のかかる薬剤を包含する。

【0212】

アクセサリーブラスミド

本明細書において用いられる用語「アクセサリーブラスミド」は、条件付きプロモーターのコントロール下における感染性のウイルス粒子の生成のために要求される遺伝子を含むプラスミドを言う。遺伝子の連続的進化のコンテキストにおいて、アクセサリーブラスミドの条件付きプロモーターからの転写は、典型的には、進化させられるべき遺伝子の機能によって直接的または間接的に活性化される。従って、条件付きプロモーターを活性化することができるかまたは進化させられるべき遺伝子の他のバージョンよりも強く条件付きプロモーターを活性化することができる、進化させられるべき遺伝子のバージョンを持つウイルスベクターの所与の集団においては、アクセサリーブラスミドは、それらのウイルスベクターに競争上の利点を授ける機能を供する。いくつかの態様では、進化させられるべき遺伝子の「活性化」バージョンを持つウイルスベクターのみが、ホスト細胞において感染性のウイルス粒子を生成するために要求される遺伝子の発現を誘導することができ、それゆえに、ホスト細胞のフロー中においてウイルスゲノムのパッケージングおよび増殖を許すであろう。他方で、進化させられるべき遺伝子の活性化しないバージョンを持つベクターは、感染性のウイルスベクターを生成するために要求される遺伝子の発現を誘導

せず、それゆえに、フレッシュな宿主細胞に感染し得るウイルス粒子へとパッケージングされないであろう。

【0213】

祖先配列再構築 (ASR)

祖先配列再構築 (ASR) は、ASRアルゴリズムを用いて、進化的 / 系統的コンテキストにおいて現代の配列を分析して、ある樹の特定のノードの祖先配列を推定するプロセスである。ASRアルゴリズムは当分野において公知である。

【0214】

塩基編集

塩基編集は、標的化されたゲノム座位における特異的な核酸塩基から別のものへの変換が関わるゲノム編集テクノロジーである。ある側面では、これは二本鎖DNA切断 (DSB) を要求することなしに達成され得る。多くの遺伝子疾患は点変異から生起するので、このテクノロジーはヒトの健康および疾患の研究に重要な含意を有する。

10

【0215】

現在のところ、CRISPRに基づくシステムを包含する他のゲノム編集技術は、目当ての座位におけるDSBの導入によって始まる。爾後に、細胞のDNA修復酵素が切断を繕い、普通は、DSBの部位における塩基のランダムな挿入または欠失 (インデル) をもたらす。しかしながら、遺伝子全体の確率的な破壊よりもむしろ標的座位における点変異の導入または修正が所望であるときには、これらのゲノム編集技術は不適である。なぜなら、修正率が低く (例えば、典型的には0.1%から5%)、主要なゲノム編集産物はインデルであるからである。ランダムなインデルを同時に導入することなしに遺伝子修正の効率を増大させるために、本発明者らは、以前に、DSB形成なしに1つのDNA塩基を別のものに直接変換するようにCRISPR / Cas9システムを改変した。

20

【0216】

塩基編集因子

本明細書において用いられる用語「塩基編集因子 (BE)」または「核酸塩基編集因子 (NBE)」は、本明細書に記載される改善されたCas融合蛋白質を言う。いくつかの態様では、融合蛋白質はデアミナーゼに融合させられたヌクレアーゼ不活性型Cas9 (dCas9) を含み、これはガイドRNAによってプログラムされた様式でRループの形成を介してDNAになお結合するが、DNAバックボーンを切断しない。例えば、その全体が参照によって本明細書に組み込まれるPCT/US2016/058344 (WO2017/070632として公開) に記載されている通り、融合蛋白質のdCas9はD10AおよびH840A変異を含み得る (これは、Cas9が核酸二重鎖の1つの鎖のみを切断する能力があるようにする)。いくつかの態様では、融合蛋白質は、デアミナーゼ、例えばDNA塩基シトシンをウラシルに変換するシチジンデアミナーゼ (rAPOBEC1) に融合させられたCas9ニッカーゼを含む。1つのかかる塩基編集因子は文献では「BE1」と言われる。いくつかの態様では、融合蛋白質は、デアミナーゼに融合させられかつさらにUGIドメイン (ウラシルDNAグリコシラーゼ阻害因子。これは爾後のU:Gミスマッチが再びC:G塩基対へと修復されることを防止する) に融合させられたヌクレアーゼ不活性型Cas9を含む。1つのかかる塩基編集因子は文献では「BE2」と言われる。他の態様では、塩基編集効率を改善するために、BE2のCas9 HNHドメインの位置840の触媒His残基が回復し (be restore) 得る (文献に記載されている通り「BE3」をもたらす)。これは非編集鎖のみをニッキングし、新たに合成されたDNAを模倣し、所望のU:A産物に至る。他の態様では、dCas9はPCT/US2017/045381 (WO2018/027078として公開) に開示または記載されているいずれかのdCas9であり、これはその全体が参照によって本明細書に組み込まれる。用語「核酸塩基編集因子 (NBE)」および「塩基編集因子 (BE)」は交換可能に用いられ得る。用語「塩基編集因子」は、本出願の時に当分野において公知または記載されているいずれかの塩基編集因子を包摂するが、本明細書に記載される進化した塩基編集因子をもまた包摂する。編集効率を改善するための本明細書に記載される方法および戦略によって改変され得る技術水準において公知の塩基編集因子は、例えばBE1、BE2、BE3、またはBE4を包含する。

30

40

50

【0217】

Cas9またはCas9部分もしくはCas9ドメイン

用語「Cas9」または「Cas9ヌクレアーゼ」または「Cas9部分」または「Cas9ドメイン」は、CRISPR関連蛋白質9またはその機能的断片を言い、いずれかの生物からのいずれかの天然に存在するCas9、いずれかの天然に存在するCas9同等物またはその機能的断片、いずれかの生物からのいずれかのCas9ホモログ、オーソログ、またはパラログ、および天然に存在するかまたは操作されたCas9のいずれかの変異体またはバリエーションを包摂する。より広くは、Cas9は、ある型の「RNAによってプログラム可能なヌクレアーゼ」もしくは「RNAによってガイドされるヌクレアーゼ」、またはより広くはある型の「核酸によってプログラム可能なDNA結合蛋白質 (napDNAbp)」である。用語Cas9は特に限定的であることを意味されず、「Cas9または同等物」と言われ得る。例示的なCas9蛋白質はさらに本明細書に記載および/または当分野において記載されており、参照によって本明細書に組み込まれる。本開示は、本発明の進化した塩基編集因子に採用される特定のCas9については限定がない。

10

dCas9

本明細書において用いられる用語「dCas9」はヌクレアーゼ不活性型Cas9もしくはヌクレアーゼデッドCas9またはその機能的断片を言い、いずれかの生物からのいずれかの天然に存在するdCas9、いずれかの天然に存在するdCas9同等物またはその機能的断片、いずれかの生物からのいずれかのdCas9ホモログ、オーソログ、またはパラログ、および天然に存在するかまたは操作されたdCas9のいずれかの変異体またはバリエーションを包摂する。用語dCas9は特に限定的であることを意味されず、「dCas9または同等物」と言われ得る。例示的なdCas9蛋白質およびdCas9蛋白質を作るための方法がさらに本明細書に記載および/または当分野において記載されており、参照によって本明細書に組み込まれる。

20

【0218】

キメラ蛋白質

用語「キメラ蛋白質」は、第1の蛋白質部分および第2の蛋白質部分が異なる種に由来する融合蛋白質を言う。

【0219】

連続的進化

本明細書において用いられる用語「連続的進化」は、核酸の集団が(a)複製、(b)変異、および(c)セレクションという複数のラウンドに付されて所望の進化した産物、例えば所望の活性を有する蛋白質をコードする核酸を産生する進化手続きを言い、複数のラウンドは研究者の相互作用なしに実行され得、(a)-(c)のプロセスは同時に行われ得る。典型的には、進化手続きは、*in vitro*で、例えば培養の細胞を宿主細胞として用いて行われる。一般的に、本明細書において提供される連続的進化プロセスは、目当ての遺伝子が核酸ベクターによって提供され、これが宿主細胞中における複製と別の宿主細胞への伝播とを包含するライフサイクルを経験するシステムに依拠し、ライフサイクルのある決定的な構成要素は失活しており、構成要素の再活性化は目当ての遺伝子の所望の変異に依存する。

30

【0220】

いくつかの態様では、目当ての遺伝子は目当ての遺伝子の活性に依存的な様式で細胞から細胞に伝播する。いくつかの態様では、トランスファーベクターは、細胞に感染するウイルス、例えばバクテリオファージまたはレトロウイルスベクターである。いくつかの態様では、ウイルスベクターはファージベクターであり、細菌宿主細胞に感染する。いくつかの態様では、トランスファーベクターはレトロウイルスベクター、例えばレンチウイルスベクターまたは水胞性口炎ウイルスベクターであり、ヒトまたはマウス細胞に感染する。いくつかの態様では、トランスファーベクターはドナー細菌細胞からレシピエント細菌細胞に伝播する接合プラスミドである。

40

【0221】

いくつかの態様では、目当ての遺伝子を含む核酸ベクターは、ファージ、ウイルスベク

50

ター、または裸のDNA（例えば、可動化プラスミド）である。いくつかの態様では、細胞から細胞への目当ての遺伝子の伝播は感染、トランスフェクション、トランスダクション、接合、または裸のDNAの取り込みによってであり、細胞から細胞への伝播の効率（例えば、伝播率）は目当ての遺伝子によってコードされる産物の活性に依存する。例えば、いくつかの態様では、核酸ベクターは目当ての遺伝子を抱くファージであり、ファージ伝播の効率（感染による）は、ファージ粒子の生成のために要求される蛋白質（例えば、M13ファージのpIII）が目当ての遺伝子の所望の活性の存在下でのみ宿主細胞によって発現されるということで、目当ての遺伝子の活性に依存する。別の例では、核酸ベクターは、レトロウイルスベクター、例えば目当ての遺伝子を抱くレンチウイルスまたは水胞性口炎ウイルスベクターであり、細胞から細胞へのウイルス伝播の効率は、ウイルス粒子の生成のために要求される蛋白質（例えばエンベロープ蛋白質、例えばVSV-g）が目当ての遺伝子の所望の活性の存在下でのみ宿主細胞によって発現されるということで、目当ての遺伝子の活性に依存する。別の例では、核酸ベクターは、例えば接合によって細菌宿主細胞間で伝播する目当ての遺伝子を含む可動化プラスミドDNAの形態の、DNAベクターであり、細胞から細胞への接合によって媒介される伝播の効率は、接合によって媒介される伝播のために要求される蛋白質（例えば、traAまたはtraQ）が目当ての遺伝子の所望の活性の存在下でのみ宿主細胞によって発現されるということで、目当ての遺伝子の活性に依存する。宿主細胞はそれらの遺伝子の1つまたは両方を欠くFプラスミドを含有する。

10

【0222】

例えば、いくつかの態様は、進化させられるべき目当ての遺伝子を含むウイルスベクターの集団が、宿主細胞のフロー、例えばラグーンのフロー中で複製する連続的進化システムを提供し、ウイルスベクターは、感染性のウイルス粒子の生成に必須である蛋白質をコードする遺伝子を欠損しており、遺伝子は、目当ての遺伝子またはその変異したバージョンによってコードされる遺伝子産物によって活性化され得る条件付きプロモーターのコントロール下で宿主細胞内に含まれる。いくつかの態様では、条件付きプロモーターの活性は、目当ての遺伝子によってコードされる遺伝子産物の所望の機能に依存する。目当ての遺伝子が所望の機能を付与する変異を獲得していないウイルスベクターは、条件付きプロモーターを活性化しないか、または最小限の活性化のみを達成するであろう。一方で、所望の変異を付与する目当ての遺伝子のいずれかの変異は、条件付きプロモーターの活性化をもたらすであろう。条件付きプロモーターはウイルスライフサイクルの必須蛋白質をコントロールするので、有利な変異を獲得したそれらのベクターにとって、このプロモーターの活性化はウイルス伝染および複製の上での利点に直接的に対応する。

20

30

【0223】

シチジンデアミナーゼ

本明細書において用いられる、CDA遺伝子によってコードされる「シチジンデアミナーゼ」は、シチジン（すなわち、リボース環に取り付けられているときには塩基シトシン）からウリジン（C U）およびデオキシシチジンからデオキシウリジン（C U）へのアミノ基の取り除きを触媒する酵素である。シチジンデアミナーゼの限定しない例はAPOBEC1（「アポリポ蛋白質B mRNA編集酵素触媒ポリペプチド1」）である。別の例はAID（「活性化誘導シチジンデアミナーゼ」）である。標準的なWatson-Crick水素結合対形成下では、シトシン塩基はグアニン塩基と水素結合する。シチジンがウリジンに変換される（またはデオキシシチジンがデオキシウリジンに変換される）ときには、ウリジン（またはウリジンのウラシル塩基）は塩基アデニンとの水素結合対形成を経験する。それゆえに、シチジンデアミナーゼによる「C」からウリジン（「U」）への変換は細胞の修復および/または複製プロセスの間に「G」の代わりに「A」の挿入を引き起こすであろう。アデニン「A」はチミン「T」と対形成するので、シチジンデアミナーゼはDNA複製と連携して二本鎖DNA分子中にC-G対形成からT-A対形成への変換を引き起こす。

40

【0224】

CRISPR

CRISPRは、原核生物に侵入したウイルスによる先立つ感染のスニペットにあたる、細菌

50

および古細菌におけるDNA配列のあるファミリーである（すなわち、CRISPRクラスター）。DNAのスニペットは原核生物細胞によって用いられて、類似のウイルスによる爾後の攻撃からのDNAを検出および破壊し、CRISPR関連蛋白質（Cas9およびそのホモログを包含する）およびCRISPR関連RNAのアレイと共に、原核生物免疫防御システムを有効に組成する。天然では、CRISPRクラスターは転写され、CRISPR RNA（crRNA）へとプロセシングされる。ある型のCRISPRシステム（例えば、II型CRISPRシステム）では、pre-crRNAの正しいプロセシングはtrans-encoded small RNA（tracrRNA）、内在性リボヌクレアーゼ3（rnc）、およびCas9蛋白質を要求する。tracrRNAはpre-crRNAのリボヌクレアーゼ3によって支援されるプロセシングのためのガイドとしての用をなす。爾後に、Cas9/crRNA/tracrRNAはRNAに対して相補的な線形または環状dsDNA標的をエンド的に切断する。具体的には、crRNAに対して相補的ではない標的鎖が第1にエンド的に切られ、それからエキソ的に3'-5'トリミングされる。天然では、DNA結合および切断は典型的には蛋白質および両方のRNAを要求する。しかしながら、crRNAおよびtracrRNA両方の側面を1個のRNA種 - ガイドRNAに組み込むように、シングルガイドRNA（「sgRNA」または単純に「gRNA」）が操作され得る。例えば、Jinek M., Chylinski K., Fonfara I., Hauer M., Doudna J. A., Charpentier E. *Science* 337: 816-821 (2012)を参照する。この内容全体は参照によってここに組み込まれる。Cas9は、CRISPRリピート配列中の短いモチーフ（PAMまたはプロトスペーサー隣接モチーフ）を認識して、自己対非自己を区別することを助ける。CRISPRの生物学、ならびにCas9ヌクレアーゼ配列および構造は当業者に周知である（例えば、"Complete genome sequence of an M1 strain of *Streptococcus pyogenes*." Ferretti et al., J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K., Primeaux C, Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P., Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J., Yuan X., Clifton S.W., Roe B.A., McLaughlin R.E., *Proc. Natl. Acad. Sci. U.S.A.* 98: 4658-4663 (2001); "CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III." Deltcheva E., Chylinski K., Sharma CM., Gonzales K., Chao Y., Pirzada Z.A., Eckert M.R., Vogel J., Charpentier E., *Nature* 471: 602-607 (2011); および"A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity." Jinek M., Chylinski K., Fonfara I., Hauer M., Doudna J. A., Charpentier E. *Science* 337: 816-821 (2012)を参照する。これらのそれぞれの内容全体は参照によって本明細書に組み込まれる。Cas9オーソログは種々の種において記載されており、*S. pyogenes*および*S. thermophilus*を包含するが、これらに限定されない。追加の好適なCas9ヌクレアーゼおよび配列は本開示に基づいて当業者には明らかであろう。かかるCas9ヌクレアーゼおよび配列は、Chylinski, Rhun, and Charpentier, "The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems" (2013) *RNA Biology* 10: 5, 726-737に開示されている生物および座位からのCas9配列を包含する；この内容全体は参照によって本明細書に組み込まれる。

10

20

30

40

50

【0225】

デアミナーゼまたはデアミナーゼドメイン

本明細書において用いられる用語「デアミナーゼ」または「デアミナーゼドメイン」または「デアミナーゼ部分」は、脱アミノ化反応を触媒する蛋白質または酵素を言う。いくつかの態様では、デアミナーゼはアデノシンデアミナーゼであり、これはアデニンまたはアデノシンの加水分解的脱アミノ化を触媒する（例えば、DNA中のアデノシンを脱アミノ化する操作されたアデノシンデアミナーゼ）。いくつかの態様では、デアミナーゼまたはデアミナーゼドメインはシチジンデアミナーゼであり、シチジンまたはデオキシシチジンからそれぞれウリジンまたはデオキシウリジンへの加水分解的脱アミノ化を触媒する。いくつかの態様では、デアミナーゼまたはデアミナーゼドメインはシチジンデアミナーゼドメインであり、シトシンからウラシルへの加水分解的脱アミノ化を触媒する。いくつかの態様では、デアミナーゼまたはデアミナーゼドメインはヒト、チンパンジー、ゴリラ、サル、ウシ、イヌ、ラット、またはマウスなどの生物からの天然に存在するデアミナーゼである。いくつかの態様では、デアミナーゼまたはデアミナーゼドメインは、ある生物から

の天然に存在するデアミナーゼの天然に存在しないバリエーションである。例えば、いくつかの態様では、デアミナーゼまたはデアミナーゼドメインは、ある生物からの天然に存在するデアミナーゼと少なくとも50%、少なくとも55%、少なくとも60%、少なくとも65%、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、または少なくとも99.5%同一である。用語デアミナーゼは、野生型カウンターパートのデアミナーゼと比べて1つ以上の変化を含むアミノ酸配列を有するバリエーションデアミナーゼをもたらず遺伝子改変（例えば1つ以上の変異）を含み得る、いずれかの遺伝子操作されたデアミナーゼをもまた包摂する。デアミナーゼの例が本明細書において与えられ、用語は限定的であることを意味されない。

10

【0226】

有効量

本明細書において用いられる用語「有効量」は、所望の生物学的応答を誘起するために十分である生物活性の薬剤の量を言う。例えば、いくつかの態様では、塩基編集因子の有効量は、標的部位ヌクレオチド配列、例えばゲノムを編集するために十分である塩基編集因子の量を言い得る。いくつかの態様では、本明細書において提供される塩基編集因子の、例えばヌクレアーゼ不活性型Cas9ドメインと核酸編集ドメイン（例えばデアミナーゼドメイン）とを含む融合蛋白質の有効量は、融合蛋白質によって特異的に結合および編集される標的部位の編集を誘導するために十分である融合蛋白質の量を言い得る。当業者によって了解されるであろう通り、薬剤、例えば融合蛋白質、ヌクレアーゼ、デアミナーゼ、

20

【0227】

進化した塩基編集因子

用語「進化した塩基編集因子」または「進化した塩基編集因子バリエーション」は、参照のまたは出発点の塩基編集因子（またはその構成要素もしくはドメイン）を連続的進化方法（例えばPACE）によって変異させる結果として形成される塩基編集因子を言い、進化した塩基編集因子は、参照のまたは出発点の塩基編集因子のアミノ酸配列と比べて、そのアミノ酸配列に導入された1つ以上のアミノ酸バリエーションを有する。例えば、コード配列中のいずれかの特定の位置のコドンの変化、1つ以上のアミノ酸の欠失（例えば、短縮された蛋白質）、1つ以上のアミノ酸の挿入、または上述のいずれかの組み合わせをもたらず塩基編集因子をコードするヌクレオチド配列の変化の結果として、アミノ酸配列バリエーションは参照塩基編集因子のアミノ酸配列中に1つ以上の変異した残基を包含し得る。いくつかの態様では、進化した塩基編集因子は、参照塩基編集因子と少なくとも50%、少なくとも55%、少なくとも60%、少なくとも65%、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、または少なくとも99.5%同一である。進化した塩基編集因子は、塩基編集因子の1つ以上の構成要素またはドメインのバリエーションを包含し得る（例えば、Cas9ドメイン、デアミナーゼドメイン、もしくはUGIドメインに導入されたバリエーション、またはこれらのドメインの組み合わせに導入されたバリエーション）。

30

40

【0228】

フロー

宿主細胞のコンテキストにおいて本明細書において用いられる用語「フロー」は宿主細胞の流れを言い、フレッシュな宿主細胞が宿主細胞集団、例えばラグーンの宿主細胞集団に導入されていき、限定された時間に渡って集団内に残り、それから宿主細胞集団から取り除かれる。単純な形態では、宿主細胞フローは、例えばコントロールされた速度における、チューブまたはチャンネル中のフローであり得る。他の態様では、宿主細胞のフローは、ある体積の細胞培養培地を保持するラグーン中を導かれ、インフロー

50

およびアウトフローを含む。フレッシュな宿主細胞の導入は連続的または間欠的であり得、取り除きは、例えばオーバーフローによって受動的、または例えば能動的サイフォン送液もしくは圧送によって能動的であり得る。さらに、取り除きはランダムであり得る。例えば、宿主細胞の攪拌懸濁培養物が提供される場合には、取り除かれる液体培養培地は、フレッシュに導入された宿主細胞と、何らかの時間に渡ってラグーン内において宿主細胞集団の構成員であった細胞とを含有するであろう。理論上は、細胞はラグーンからの取り除きを際限なく脱出し得るが、平均的な宿主細胞は、限定された時間に渡ってのみラグーン内に残るであろう。これは主にラグーン中の培養培地（および懸濁された細胞）の流量によって決定される。

【0229】

ウイルスベクターは宿主細胞のフロー中において複製し、その中では、フレッシュな未感染の宿主細胞が提供され、一方で感染した細胞は取り除かれるので、複数の継続したウイルスライフサイクルが研究者の相互作用なしに起こり得、これは1回の進化実験で複数の有利な変異の蓄積を許す。

【0230】

目当ての遺伝子

本明細書において用いられる用語「目当ての遺伝子」は、目当ての遺伝子産物、例えば、本明細書において提供される連続的進化プロセスによって進化させられるべき遺伝子産物（例えば、塩基編集因子またはその構成要素/ドメイン）をコードするヌクレオチド配列を含む核酸コンストラクトを言う。用語は、本明細書において提供される方法に従う連続的進化プロセスの結果である目当ての遺伝子のいずれかのバリエーションを包含する。例えば、いくつかの態様では、目当ての遺伝子は、コード配列の発現がウイルスゲノム中の1つ以上のプロモーターのコントロール下にあるように、ウイルスベクター、例えばファージゲノムにクローニングされた、進化させられるべき蛋白質をコードするヌクレオチド配列を含む核酸コンストラクトである。他の態様では、目当ての遺伝子は、進化させられるべき蛋白質をコードするヌクレオチド配列とコード配列に作動可能に連結されたプロモーターとを含む核酸コンストラクトである。ウイルスベクター、例えばファージゲノムにクローニングされるときに、かかる目当ての遺伝子のコード配列の発現は異種プロモーターのコントロール下にあり、いくつかの態様では、ウイルスゲノムに含まれる1つ以上のプロモーターによってもまた影響され得る。

【0231】

目当ての遺伝子の機能

用語「目当ての遺伝子の活性」と交換可能に用いられる用語「目当ての遺伝子の機能」は、目当ての遺伝子によってコードされる遺伝子産物、例えば核酸または蛋白質の機能または活性を言う。例えば、目当ての遺伝子の機能は、酵素活性（例えば、反応産物の生成をもたらす酵素活性、リン酸化活性、ホスファターゼ活性など）、転写を活性化する能力（例えば、特異的なプロモーター配列を標的化した転写活性化活性）、結合形成活性（例えば、共有結合的な結合の形成をもたらす酵素活性）、または結合活性（例えば、蛋白質、DNA、もしくはRNA結合活性）であり得る。

【0232】

融合蛋白質

本明細書において用いられる用語「融合蛋白質」は、少なくとも2つの異なる蛋白質からの蛋白質ドメインを含むハイブリッドポリペプチドを言う。1つの蛋白質は融合蛋白質のアミノ末端（N末端）部分にまたはカルボキシ末端（C末端）蛋白質に見つかり得、それゆえにそれぞれ「アミノ末端融合蛋白質」または「カルボキシ末端融合蛋白質」を形成する。蛋白質は、核酸編集蛋白質の異なるドメイン、例えば核酸結合ドメイン（例えば、標的部位への蛋白質の結合を導くCas9のgRNA結合ドメイン）および核酸切断ドメインまたは触媒ドメインを含み得る。本明細書において提供される蛋白質のいずれかは当分野において公知のいずれかの方法によって産生され得る。例えば、本明細書において提供される蛋白質は組換え体蛋白質発現および精製によって産生され得、これはとりわけペプチドリン

10

20

30

40

50

カーを含む融合蛋白質に適する。組換え体蛋白質発現および精製のための方法は周知であり、Green and Sambrook, *Molecular Cloning: A Laboratory Manual* (4th ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (2012))に記載されているものを包含し、これの内容全体は参照によって本明細書に組み込まれる。

【0233】

ヘルパーファージ

本明細書において用語「ヘルパーファージミド」および「ヘルパープラスミド」と交換可能に用いられる用語「ヘルパーファージ」は、ファージライフサイクルのために要求されるファージ遺伝子または複数のかかる遺伝子を含むが、ファージ粒子へのゲノムパッケージングのために要求される構造的要素を欠く核酸コンストラクトを言う。例えば、ヘルパーファージはファージ複製起点を欠く野生型ファージゲノムを提供し得る。いくつかの態様では、ファージ粒子の生成のために要求される遺伝子を含むが、感染性の粒子の生成のために要求される遺伝子、例えば全長pIII遺伝子を欠くヘルパーファージが提供される。いくつかの態様では、ヘルパーファージは、ファージ粒子の生成のために要求される全てではなくいくつかの遺伝子のみを提供する。ヘルパーファージは、ファージ粒子の生成のために要求される遺伝子を欠く改変されたファージが宿主細胞内でファージライフサイクルを完遂することを許すために有用である。典型的には、ヘルパーファージは、ファージゲノムに欠けているファージ粒子の生成のために要求される遺伝子を含み、それゆえにファージゲノムを相補するであろう。連続的進化コンテキストでは、ヘルパーファージは典型的にはセレクションファージを相補するが、両方は、感染性のファージ粒子の産生のために要求されるあるファージ遺伝子を欠く。

10

20

【0234】

宿主細胞

本明細書において用いられる用語「宿主細胞」は、本明細書において提供される連続的進化プロセスにとって有用なファージベクターを宿主、複製、および伝播し得る細胞を言う。ベクターがウイルスベクターである態様では、好適な宿主細胞は、ウイルスベクターによって感染され得る、それを複製し得る、およびフレッシュな宿主細胞に感染し得るウイルス粒子へとそれをパッケージングし得る細胞である。それがウイルスベクターの遺伝子の発現、ウイルスゲノムの複製、および/またはウイルス粒子の生成を支持し得る場合には、細胞はウイルスベクターを宿主し得る。細胞が所与のウイルスベクターにとって好適な宿主細胞であるかどうかを決定するための1つの基準は、細胞がウイルスベクターが由来する野生型ウイルスゲノムのウイルスライフサイクルを支持し得るかどうかを決定することである。例えば、本明細書に記載されるいくつかの態様において提供される通り、ウイルスベクターが改変されたM13ファージゲノムである場合には、好適な宿主細胞は野生型M13ファージライフサイクルを支持し得るいずれかの細胞であろう。連続的進化プロセスに有用なウイルスベクターにとって好適な宿主細胞は当業者に周知であり、本開示はこれについて限定されない。いくつかの態様では、ウイルスベクターはファージであり、宿主細胞は細菌細胞である。いくつかの態様では、宿主細胞はE. coli細胞である。好適なE. coli宿主株は当業者には明らかであろう。New England Biolabs (NEB) Turbo, Top10F', DH12S, ER2738, ER2267、およびXL1-Blue MRF'を包含するが、これらに限定されない。これらの株名は当分野において認識されており、これらの株の遺伝子型は良くキャラクタリゼーションされている。上の株は例示的であるのみであるということと、本発明はこれについて限定されないということとは理解されるはずである。本明細書において宿主細胞のコンテキストにおいて用語「非感染」または「未感染」と交換可能に用いられる用語「フレッシュ」は、本明細書において提供される連続的進化プロセスに用いられる目当ての遺伝子を含むウイルスベクターによって感染されていない宿主細胞を言う。しかしながら、フレッシュな宿主細胞は、進化させられるべきベクターに無関係なウイルスベクターによってまたは同じかもしくは類似の型だが目当ての遺伝子を持たないベクターによって感染して得る。

30

40

【0235】

50

いくつかの態様では、ホスト細胞は原核生物細胞、例えば細菌細胞である。いくつかの態様では、ホスト細胞はE. coli細胞である。いくつかの態様では、ホスト細胞は真核生物細胞、例えば酵母細胞、昆虫細胞、または哺乳類細胞である。当然のことながら、ホスト細胞の型は採用されるウイルスベクターに依存し、好適なホスト細胞/ウイルスベクターの組み合わせは当業者には難なく明らかであろう。

【0236】

いくつかのPACE態様では、例えばM13セクションファージを採用する態様では、ホスト細胞は、普通はF因子、性決定因子、またはFプラスミドともまた言われる稔性因子を発現するE. coli細胞である。F因子は、細菌が接合に必要な性線毛を産生することを許す細菌DNA配列であり、ある種のファージによる、例えばM13ファージによるE. coli細胞の感染に必須である。例えば、いくつかの態様では、M13-PACEのためのホスト細胞は遺伝子型 F'proA⁺B⁺ (lacIZY) zzf::Tn10(TetR)/ endA1 recA1 galE15 galK16 nupG rpsL lacI ZYA araD 139 (ara, leu)7697 mcrA (mrr-hsdRMS-mcrBC) proBA::pir116 である。

10

【0237】

感染性のウイルス粒子

本明細書において用いられる用語「感染性のウイルス粒子」は、それが含むウイルスゲノムを好適なホスト細胞内に輸送することができるウイルス粒子を言う。全てのウイルス粒子がウイルスゲノムを好適なホスト細胞に伝播させることができるわけではない。これを奏することができない粒子は非感染性のウイルス粒子と言われる。いくつかの態様では、ウイルス粒子は複数の異なるコート蛋白質を含み、コート蛋白質の1つまたはいくつかはウイルス粒子の構造を損なうことなしに省かれ得る。いくつかの態様では、少なくとも1つのコート蛋白質が感染性の喪失なしには省かれ得ないウイルス粒子が提供される。ウイルス粒子が感染性を付与する蛋白質を欠く場合には、ウイルス粒子は感染性ではない。例えば、ファージ蛋白質(例えばpVIII)のコートによってパッケージングされたファージゲノムを含むpIII(蛋白質III)を欠くM13ファージ粒子は、非感染性のM13ファージ粒子である。なぜなら、pIIIはM13ファージ粒子の感染特性に必須であるからである。

20

【0238】

塩基修復の阻害因子

用語「塩基修復の阻害因子」または「IBR」は、核酸修復酵素、例えば塩基除去修復酵素の活性を阻害する能力がある蛋白質を言う。いくつかの態様では、IBRはイノシン塩基除去修復の阻害因子である。塩基修復の例示的な阻害因子は、APE1、Endo III、Endo IV、Endo V、Endo VIII、Fpg、hOGG1、hNEIL1、T7 Endo1、T4PDG、UDG、hSMUG1、およびhAAGの阻害因子を包含する。いくつかの態様では、IBRはEndo VまたはhAAGの阻害因子である。いくつかの態様では、IBRは触媒的に不活性なEndoVまたは触媒的に不活性なhAAGである。

30

【0239】

インテイン

本明細書において用いられる用語「インテイン」は、生命の全てのドメインからの生物に見出される自己プロセシングポリペプチドドメインを言う。インテイン(介在蛋白質)は蛋白質スプライシングとして公知のユニークな自己プロセシング事象を行い、これにおいては、それは、それ自体をより大きい前駆体ポリペプチドから2つのペプチド結合の切断によって切り出し、プロセスにおいて、新たなペプチド結合の形成によってフランキングエクステイン(外側の蛋白質)配列同士をライゲーションする。インテイン遺伝子は他の蛋白質コード遺伝子とインフレームで埋め込まれて見出されるので、この再配列は翻訳後に(またはおそらく翻訳同時的に)起こる。さらにその上、インテインによって媒介される蛋白質スプライシングは自発的である;それは外側からの因子またはエネルギー源を要求せず、インテインドメインの折りたたみのみを要求する。「分割型インテイン」によるトランス蛋白質スプライシングの天然プロセスと対比して、このプロセスはシス蛋白質スプライシングとしてもまた公知である。インテインは自己スプライシングRNAイントロンの蛋白質同等物であり(Perler et al., Nucleic Acids Res. 22:1125-1127 (1994))

40

50

を参照する)、これらは前駆体蛋白質からのそれら自体の切り出しを触媒し、エクステインとして公知のフランキング蛋白質配列同士の付随した融合を有する(Perler et al., Curr. Opin. Chem. Biol. 1: 292-299 (1997); Perler, F. B. Cell 92(1): 1-4 (1998); Xu et al., EMBO J. 15 (19): 5146-5153 (1996)に概説)。

【0240】

ラグーン、セルスタット、およびタービドスタット

本明細書において用いられる用語「ラグーン」は、その中を宿主細胞のフローが導かれるベッセルを言う。本明細書において提供される連続的進化プロセスについて用いられるときには、ラグーンは、典型的には、宿主細胞の集団と宿主細胞集団内で複製するウイルスベクターの集団とを保持し、ラグーンは、宿主細胞がラグーンから取り除かれるアウトフローと、フレッシュな宿主細胞がラグーンに導入され、それゆえに宿主細胞集団を補充するインフローとを含む。いくつかの態様では、ラグーン中の細胞のフローは、ラグーン内に本質的に一定数の宿主細胞をもたらすように制御される。いくつかの態様では、ラグーン中の細胞のフローはラグーン内に本質的に一定数のフレッシュな宿主細胞をもたらすように制御される。

10

【0241】

本明細書において用いられる用語「セルスタット」は宿主細胞を含む培養ベッセルを言い、その中の細胞数は経時的に実質的に一定である。

【0242】

本明細書において用いられる用語「タービドスタット」は懸濁培養の宿主細胞を含む培養ベッセルを言い、その中の培養培地の濁度は経時的に実質的に本質的に一定である。いくつかの態様では、懸濁培養物の、例えば細菌細胞の濁度は培養培地中の細胞密度の尺度である。いくつかの態様では、タービドスタットは、フレッシュな培地のインフローおよびアウトフローと、タービドスタット内の懸濁培養物の濁度に基づいてタービドスタット内へのおよび/またはから外へのフローを制御するコントローラとを含む。

20

【0243】

リンカー

本明細書において用いられる用語「リンカー」は、2つの分子または部分、例えばヌクレアーゼの結合ドメインおよび切断ドメインを連結する化学基または分子を言う。いくつかの態様では、リンカーは、RNAによってプログラム可能なヌクレアーゼのgRNA結合ドメインとリコンビナーゼの触媒ドメインとをつなぎ合わせる。いくつかの態様では、リンカーはdCas9と塩基編集因子部分(例えば、シチジンまたはアデノシンデアミナーゼ)とをつなぎ合わせる。典型的には、リンカーは2つの基、分子、または他の部分の間に位置するかまたはそれらによってフランキングされ、共有結合的な結合を介してそれぞれのものに接続され、それゆえに2つを接続する。いくつかの態様では、リンカーはアミノ酸または複数のアミノ酸(例えば、ペプチドまたは蛋白質)である。いくつかの態様では、リンカーは有機分子、基、ポリマー、またはケミカル部分である。いくつかの態様では、リンカーは長さが5-100アミノ酸、例えば、長さが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、30-35、35-40、40-45、45-50、50-60、60-70、70-80、80-90、90-100、100-150、または150-200アミノ酸である。より長いまたはより短いリンカーもまた企図される。

30

40

【0244】

変異原

本明細書において用いられる用語「変異原」は、所与の生物学的システム、例えば宿主細胞において、そのシステムの変異の天然に存在するレベルよりも上のレベルまで変異を誘導するかまたは変異速度を増大させる薬剤を言う。連続的進化プロセスでは、変異原は校正能を欠くDNAポリメラーゼであり得る。

【0245】

変異導入プラスミド

本明細書において用いられる用語「変異導入プラスミド」は、変異原として作用する遺

50

伝子産物をコードする遺伝子を含むプラスミドを言う。いくつかの態様では、遺伝子は校正能を欠くDNAポリメラーゼをコードする。いくつかの態様では、遺伝子は、細菌SOSストレス応答に関わる遺伝子、例えばUmuC、UmuD'、またはRecA遺伝子である。いくつかの態様では、遺伝子はGATCメチラーゼ遺伝子、例えばデオキシアデノシンメチラーゼ（damメチラーゼ）遺伝子である。いくつかの態様では、遺伝子はヘミメチル化されたGATC配列の結合に関わり、例えばseqA遺伝子である。いくつかの態様では、遺伝子は変異性核酸塩基搬出の抑制に関わり、例えばemrRである。変異導入プラスミド（変異導入コンストラクトともまた言われる）は、例えば2016年10月20日にWO2016/168631として公開された2016年4月16日出願の国際特許出願PCT/US2016/027795によって記載されており、これらの内容全体は参照によって本明細書に組み込まれる。

10

【0246】

変異

本明細書において用いられる用語「変異」は、配列、例えば核酸もしくはアミノ酸配列内の残基の別の残基による置換、または配列内の1つ以上の残基の欠失もしくは挿入を言う。典型的には、変異は、本明細書においては、元々の残基、次に配列内の残基の位置、次に新たに置換された残基のアイデンティティーを同定することによって記載される。本明細書において提供されるアミノ酸置換（変異）をなすための種々の方法は当分野において周知であり、例えばGreen and Sambrook, *Molecular Cloning: A Laboratory Manual* (4th ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (2012)) によって提供される。変異は、1塩基多型、微小重複領域、インデル、および逆位などの種々のカテゴリーを包含し得、決して限定的であることを意味されない。変異は「機能喪失」変異を包含し得、これは蛋白質活性を縮減または廃絶する変異の正常な結果である。ほとんどの機能喪失変異は劣性である。なぜなら、ヘテロ接合体においては、第2の染色体コピーが、その存在が変異の効果を補償する完全に機能的な蛋白質をコードする遺伝子の未変異のバージョンを持つからである。機能喪失変異が優性であるいくつかの例外があり、1つの例はハプロ不全であり、そこでは、生物は、ヘテロ接合体が被る蛋白質活性のおよそ50%の縮減を忍容することができない。これはヒトの数個の遺伝子疾患の説明であり、フィブリリンと呼ばれる結合組織蛋白質の遺伝子の変異からもたらされるマルファン症候群を包含する。変異は「機能獲得」変異をもまた包摂し、これは、さもなければ正常な状態では存在しない異常な活性を蛋白質または細胞に付与するものである。多くの機能獲得変異はコード領域よりもむしろ制御配列にあり、よって、いくつもの帰結を有し得る。例えば、変異は1つ以上の遺伝子が間違った組織で発現されることに至り得、これらの組織はそれらが正常では欠く機能を獲得する。代替的には、変異は細胞周期のコントロールに関わる1つ以上の遺伝子の過剰発現に至り得、それゆえに、コントロールされていない細胞分裂、ゆえに癌に至る。それらの本質ゆえに、機能獲得変異は通常は優性である。

20

30

【0247】

天然に存在しないまたは操作された

用語「天然に存在しない」または「操作された」は交換可能に用いられ、人の手の関わりを指示する。用語は、核酸分子またはポリペプチド（例えば、Cas9またはデアミナーゼ）を参照するときには、核酸分子またはポリペプチドが、それらが天然に結びつけられているおよび/または共に天然に見出される少なくとも1つの他の構成要素を少なくとも実質的に不含であるということの意味する（例えば、天然に見出されないアミノ酸配列）。

40

【0248】

核酸 / 核酸分子

本明細書において用いられる用語「核酸」はヌクレオチドのポリマーを言う。ポリマーは、天然のヌクレオシド（例えば、アデノシン、チミジン、グアノシン、シチジン、ウリジン、デオキシアデノシン、デオキシチミジン、デオキシグアノシン、およびデオキシシチジン）；ヌクレオシドアナログ（例えば、2-アミノアデノシン、2-チオチミジン、イノシン、ピロロ-ピリミジン、3-メチルアデノシン、5-メチルシチジン、C5-プロモウリジン、C5-フルオロウリジン、C5-ヨードウリジン、C5-プロピニル-ウリジン、C5-プロピニル-

50

シチジン、C5-メチルシチジン、7-デアザアデノシン、7-デアザグアノシン、8-オキソアデノシン、8-オキソグアノシン、0(6)-メチルグアニン、4-アセチルシチジン、5-(カルボキシヒドロキシメチル)ウリジン、ジヒドロウリジン、メチルシュードウリジン、1-メチルアデノシン、1-メチルグアノシン、N6-メチルアデノシン、および2-チオシチジン)、化学修飾塩基、生物学的修飾塩基(例えばメチル化塩基)、インターカレーション塩基、修飾糖(例えば、2'-フルオロリボース、リボース、2'-デオキシリボース、2'-O-メチルシチジン、アラビノース、およびヘキソース);あるいは修飾リン酸基(例えば、ホスホチオエートおよび5'-N-ホスホロアミダイト連結部)を包含し得る。

【0249】

核酸によってプログラム可能なR/DNA結合蛋白質(napR/DNAbp)

用語「核酸によってプログラム可能なD/RNA結合蛋白質(napR/DNAbp)」は、蛋白質と結びついた1つ以上の核酸分子(またはその部分もしくは領域)に対して相補的である特異的な標的ヌクレオチド配列(例えば、ゲノムの遺伝子座位)に局在するように、蛋白質を導くかまたは別様にプログラムし、それによって蛋白質が特異的な標的部位のヌクレオチド配列に結合することを引き起こす1つ以上の核酸分子(すなわち、これは広く「napR/DNAbpをプログラムする核酸分子」と言われ得、例えばCasシステムのケースのガイドRNAを包含する)と結びつき(例えば、複合体を形成し)得る、いずれかの蛋白質を言う。この用語napR/DNAbpは、天然に存在するかまたは天然に存在しないか(例えば、操作された、もしくは組換え体の)にかかわらず、CRISPR Cas9蛋白質、およびCas9同等物、ホモログ、オーソログ、またはパラログを包摂し、いずれかの型のCRISPRシステム(例えば、II、V、VI型)からのCas9同等物を包含し得、Cpf1(V型CRISPR-Casシステム)、C2c1(V型CRISPR-Casシステム)、C2c2(VI型CRISPR-Casシステム)、およびC2c3(V型CRISPR-Casシステム)を包含する。さらなるCas同等物はMakarova et al., "C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector," *Science* 2016; 353 (6299)に記載されており、この内容は参照によって本明細書に組み込まれる。しかしながら、本発明とのつながりで用いられ得る核酸によってプログラム可能なDNA結合蛋白質(napDNAbp)はCRISPR-Casシステムに限定されない。本発明は、いずれかのかかるプログラム可能な蛋白質、例えばNatronobacterium gregoryiからのArgonaute蛋白質(NgAgo)を包摂し、これもまたDNAによってガイドされるゲノム編集に用いられ得る。NgAgoガイドDNAシステムはPAM配列またはガイドRNA分子を要求せず、これは、ゲノム編集が、単純にジェネリックなNgAgo蛋白質の発現および合成オリゴヌクレオチドの導入によって、いずれかのゲノム配列に対して実行され得るということを意味する。Gao F, Shen XZ, Jiang F, Wu Y, Han C. DNA-guided genome editing using the Natronobacterium gregoryi Argonaute. *Nat Biotechnol* 2016; 34 (7): 768-73を参照する。これは参照によって本明細書に組み込まれる。

【0250】

napR/DNAbpをプログラムする核酸分子またはガイド配列

用語「napR/DNAbpをプログラムする核酸分子」または同等に「ガイド配列」は、napR/DNAbp蛋白質と結びつき、蛋白質と結びついた1つ以上の核酸分子(またはその部分もしくは領域)に対して相補的である特異的な標的ヌクレオチド配列(例えば、ゲノムの遺伝子座位)に局在するように導くかまたは別様にプログラムし、それによってnapR/DNAbp蛋白質が特異的な標的部位のヌクレオチド配列に結合することを引き起こす、1つ以上の核酸分子を言う。限定しない例はCRISPR-Casゲノム編集システムのCas蛋白質のガイドRNAである。

【0251】

核局在シグナル(NLS)

核局在シグナルまたは配列(NLS)は、核輸送による細胞核への搬入のために蛋白質をタグ付けする、指名する、または別様にマークするアミノ酸配列である。典型的には、このシグナルは蛋白質表面に暴露した正荷電のリジンまたはアルギニンの1つ以上の短い配列からなる。異なる核局在蛋白質は同じNLSを共有し得る。NLSは、蛋白質を核から外へと

10

20

30

40

50

標的化する核外搬出シグナル (NES) の反対の機能を有する。それゆえに、1個の核局在シグナルは、それが結びつけられている実体を細胞の核へと導き得る。かかる配列は、いずれかのサイズおよび組成、例えば25、25、15、12、10、8、7、6、5、または4アミノ酸よりも多くであり得るが、好ましくは、核局在シグナル (NLS) として機能することが公知の少なくとも4から8アミノ酸の配列を含むであろう。

【0252】

核酸塩基改変部分または核酸エフェクタードメイン

本明細書において用いられる用語「核酸塩基改変部分」または同等に「核酸エフェクタードメイン」は、DNAまたはRNA分子を改変する能力があるいずれかの蛋白質、酵素、またはポリペプチド（またはその機能的断片）を包摂する。核酸塩基改変部分は天然に存在し得るか、または組換え体であり得る。例えば、核酸塩基改変部分は、例えば1つ以上のDNA修復酵素、および塩基除去修復 (BER)、ヌクレオチド除去修復 (NER)、相同性依存的な組換え修復 (HR)、非相同末端結合修復 (NHEJ)、マイクロホモロジー媒介末端結合修復 (MMEJ)、ミスマッチ修復 (MMR)、直接的な復帰修復、または他の公知のDNA修復経路に関わる酵素または蛋白質を包含し得る。核酸塩基改変部分は1つ以上の型の酵素活性を有し得、エンドヌクレアーゼ活性、ポリメラーゼ活性、リガーゼ活性、複製活性、校正活性を包含するが、これらに限定されない。核酸塩基改変部分はDNAもしくはRNA修飾酵素および/または変異性の酵素、例えばDNAメチラーゼおよび脱アミノ化酵素（すなわち、シチジンデアミナーゼおよびアデノシンデアミナーゼを包含するデアミナーゼ。全て上で定義されている）をもまた包含し得、これらは核酸塩基を脱アミノ化し、いくつかのケースでは、正常な細胞のDNA修復および複製プロセスによって変異性の修正に至る。本明細書において用いられる「核酸エフェクタードメイン」（例えば、DNAエフェクタードメインまたはRNAエフェクタードメイン）もまた、核酸（例えば、DNAまたはRNA）に1つ以上の改変（例えば、シチジン残基の脱アミノ化）をなす能力がある蛋白質または酵素を言い得る。例示的な核酸編集ドメインは、デアミナーゼ、ヌクレアーゼ、ニッカーゼ、リコンビナーゼ、メチルトランスフェラーゼ、メチラーゼ、アセチラーゼ、アセチルトランスフェラーゼ、転写活性化因子、または転写抑制因子ドメインを包含するが、これらに限定されない。いくつかの態様では、核酸編集ドメインはデアミナーゼである（例えばシチジンデアミナーゼ、例えばAPOBECまたはAIDデアミナーゼ）。

【0253】

オリゴヌクレオチド/ポリヌクレオチド

本明細書において用いられる用語「オリゴヌクレオチド」および「ポリヌクレオチド」は、ヌクレオチドのポリマー（例えば、少なくとも3ヌクレオチドのストリング）を言うために交換可能に用いられ得る。いくつかの態様では、「核酸」はRNAならびに一本および/または二本鎖DNAを包摂する。核酸は、例えばゲノム、転写物、mRNA、tRNA、rRNA、siRNA、snRNA、プラスミド、コスミド、染色体、クロマチド、または他の天然に存在する核酸分子のコンテキストにおいて、天然に存在し得る。他方で、核酸分子は、天然に存在しない分子、例えば組換えDNAもしくはRNA、人工染色体、操作されたゲノム、またはその断片、あるいは合成DNA、RNA、DNA/RNAハイブリッドであり得、あるいは天然に存在しないヌクレオチドまたはヌクレオシドを包含してい得る。さらにその上、用語「核酸」、「DNA」、「RNA」、および/または類似の用語は、核酸アナログ、例えばホスホジエステルバックボーン以外を有するアナログを包含する。核酸は、天然のソースから精製、組換え発現システムを用いて産生および任意に精製、化学合成などされ得る。適切などころでは、例えば化学合成された分子のケースでは、核酸は、ヌクレオシドアナログ、例えば化学修飾された塩基または糖を有するアナログ、ならびにバックボーン修飾を含み得る。別様に指示されない限り、核酸配列は5'から3'の方向に提出される。いくつかの態様では、核酸は、天然のヌクレオシド（例えば、アデノシン、チミジン、グアノシン、シチジン、ウリジン、デオキシアデノシン、デオキシチミジン、デオキシグアノシン、およびデオキシシチジン）；ヌクレオシドアナログ（例えば、2-アミノアデノシン、2-チオチミジン、イノシン、ピロロ-ピリミジン、3-メチルアデノシン、5-メチルシチジン、2-アミノアデノ

シン、C5-プロモウリジン、C5-フルオロウリジン、C5-ヨードウリジン、C5-プロピニルウリジン、C5-プロピニル-シチジン、C5-メチルシチジン、2-アミノアデノシン、7-デアザアデノシン、7-デアザグアノシン、8-オキソアデノシン、8-オキソグアノシン、0(6)-メチルグアニン、および2-チオシチジン)；化学修飾塩基；生物学的修飾塩基(例えばメチル化塩基)；インターカレーション塩基；修飾糖(例えば、2'-フルオロリボース、リボース、2'-デオキシリボース、アラビノース、およびヘキソース)；および/または修飾リン酸基(例えば、ホスホロチオエートおよび5'-N-ホスホロアミダイト連結部)であるか、またはそれを含む。

【0254】

PACE (ファージによって支援される連続的進化)

本明細書において用いられる用語「ファージによって支援される連続的進化(PACE)」はウイルスベクターとしてファージを採用する連続的進化を言う。PACEテクノロジーの一般的概念は、例えば、2010年3月11日にWO2010/028347として公開された2009年9月8日出願の国際PCT出願PCT/US2009/056194；2012年6月28日にWO2012/088381として公開された2011年12月22日出願の国際PCT出願PCT/US2011/066747；2015年5月5日発行のU.S.出願U.S.特許No.9,023,594、2015年9月11日にWO2015/134121として公開された2015年1月20日出願の国際PCT出願PCT/US2015/012022、および2016年10月20日にWO2016/168631として公開された2016年4月15日出願の国際PCT出願PCT/US2016/027795に記載されている。これらのそれぞれの内容全体は参照によって本明細書に組み込まれる。

【0255】

プロモーター

用語「プロモーター」は当分野において認識されており、細胞の転写機構によって認識され、下流の遺伝子の転写を開始することができる配列を有する核酸分子を言う。プロモーターは恒常的に活性であり得、プロモーターが所与の細胞コンテキストにおいて常に活性であるということの意味し、または条件付きで活性であり得、プロモーターが具体的な条件の存在下でのみ活性であるということの意味する。例えば、条件付きプロモーターは、プロモーター中の制御エレメントと結びついた蛋白質を基本転写機構に接続する特異的な蛋白質の存在下でのみ、または阻害分子の不在下でのみ活性である。条件付きで活性なプロモーターのサブクラスは、小分子「誘導因子」の存在を活性のために要求する誘導性プロモーターである。誘導性プロモーターの例は、アラビノース誘導性プロモーター、Tet-onプロモーター、およびタモキシフェン誘導性プロモーターを包含するが、これらに限定されない。種々の恒常的、条件付き、および誘導性プロモーターが当業者に周知であり、当業者は本発明を行うことに有用な種々のかかるプロモーターを確かめることができるであろう。これはこれについて限定されない。

【0256】

ファージ

本明細書において用語「バクテリオファージ」と交換可能に用いられる用語「ファージ」は細菌細胞に感染するウイルスを言う。典型的には、ファージは遺伝子材料を包み込む外縁蛋白質カプシドからなる。遺伝子材料は線形または環状形態のssRNA、dsRNA、ssDNA、またはdsDNAであり得る。ファージおよびファージベクターは当業者に周知であり、本明細書において提供される方法を行うために有用であるファージの限定しない例は(Lysogen)、T2、T4、T7、T12、R17、M13、MS2、G4、P1、P2、P4、Phi X174、N4、6、および29である。ある態様では、本発明に利用されるファージはM13である。追加の好適なファージおよび宿主細胞は当業者には明らかであろう。かつ本発明はこの側面において限定されない。追加の好適なファージおよび宿主細胞の例示的な記載は、Elizabeth Kutter and Alexander Sulakvelidze: Bacteriophages: Biology and Applications. CRC Press; 1st edition (December 2004), ISBN: 0849313368; Martha R. J. Clokie and Andrew M. Kropinski: Bacteriophages: Methods and Protocols, Volume 1: Isolation, Characterization, and Interactions (Methods in Molecular Biology) Humana Press; 1st edition (December, 2008), ISBN: 1588296822; Martha R. J. Clokie and Andrew M.

10

20

30

40

50

Kropinski: Bacteriophages: Methods and Protocols, Volume 2: Molecular and Applied Aspects (Methods in Molecular Biology) Humana Press; 1st edition (December 2008), ISBN: 1603275649を参照する；これらの全ては、好適なファージおよび宿主細胞ならびにかかるファージの単離、培養、および操作のための方法およびプロトコールの開示について、それらの全体が参照によって本明細書に組み込まれる)。

【0257】

いくつかの態様では、ファージは繊維状ファージである。いくつかの態様では、ファージはM13ファージである。M13ファージは当業者に周知であり、M13ファージの生物学は鋭意に研究されている。野生型M13ゲノムの概略図は図16に提供されている。野生型M13ファージ粒子はおよそ6.4kbの環状の一本鎖ゲノムを含む。野生型(wilt-type)ゲノムは10個の遺伝子gI-gXを包含し、これらは、翻って、それぞれ10個のM13蛋白質pI-pXをコードする。gVIIIはpVIIIをコードし、多くの場合にはファージ粒子の主要構造蛋白質ともまた言われる。一方で、gIIIはpIIIをコードし、マイナーコート蛋白質ともまた言われ、これはM13ファージ粒子の感染性のために要求される。

10

【0258】

M13ライフサイクルは、pIII蛋白質を介する好適な細菌宿主細胞の性線毛へのファージの取り付けと、宿主細胞へのファージゲノムの挿入とを包含する。それから、環状の一本鎖ファージゲノムが複製型(RF)ともまた言われる環状の二本鎖DNAに変換され、これからファージ遺伝子転写が開始される。野生型M13ゲノムは9つのプロモーターおよび2つの転写ターミネーター、ならびに複製起点を含む。この一連のプロモーターは転写の勾配を提供し、その結果、2つの転写ターミネーターに最も近い遺伝子(gVIIIおよびIV)は最も高いレベルで転写される。野生型M13ファージでは、全ての10個の遺伝子の転写は同じ方向に進む。ファージによってコードされる(encode)蛋白質の1つpIIが宿主細胞内において線形の一本鎖ファージゲノムの生成を開始し、これらは爾後に環状化され、pVによって結合および安定化される。それから、環状化した一本鎖M13ゲノムはpVIIIによって結合される。一方で、pVはゲノムから引き剥がされ、これはパッケージングプロセスを開始する。パッケージングプロセスの終わりに、pIIIの複数コピーが野生型M13粒子に取り付けられ、それゆえに、別の宿主細胞に感染する用意ができた感染性のファージを生成し、ライフサイクルを締めくくる。

20

【0259】

M13ファージゲノムは、例えば、野生型遺伝子の1つ以上を欠失させることおよび/または異種核酸コンストラクトをゲノムに挿入することによって操作され得る。M13は厳格なゲノムサイズ制限を有さず、最高で42kbの挿入が報告されている。これは、関わる遺伝子の長さに限定を課すことなしに、M13ファージベクターが目当ての遺伝子を進化させるための連続的進化実験に用いられることを許す。

30

【0260】

M13ファージは良くキャラクター化されており、M13のゲノム配列は報告されている。代表的なM13ゲノム配列はパブリックデータベースから取得され得、例示的な配列はNational Center for Biotechnology Information (NCBI) データベース (www.ncbi.nlm.nih.gov) のエントリーV00604によって提供される：

40

【0261】

ファージM13ゲノム：

【0262】

>gi|56713234|emb|V00604.2|ファージM13ゲノム

【0263】

【化 2 - 1】

AACGCTACTACTATTAGTAGAATTGATGCCACCTTTTCAGCTCGCGCCC
CAAATGAAAATATAGCTAAACAGGTTATTGACCATTTGCGAAATGTATCTAATGGTC
AAACTAAATCTACTCGTTCGCAGAATTGGGAATCAACTGTTACATGGAATGAACTT
CCAGACACCGTACTTTAGTTGCATATTTAAAACATGTTGAGCTACAGCACCAGATTC
AGCAATTAAGCTCTAAGCCATCCGCAAAAATGACCTCTTATCAAAGGAGCAATTA
AAGGTACTCTCTAATCCTGACCTGTTGGAGTTTGCTTCCGGTCTGGTTCGCTTTGAAG
CTCGAATTAACGCGATATTTGAAGTCTTTCGGGCTTCCTCTTAATCTTTTTGATGC
AATCCGCTTTGCTTCTGACTATAATAGTCAGGGTAAAGACCTGATTTTTGATTTATGG
TCATTCTCGTTTTCTGAACTGTTTAAAGCATTTGAGGGGGATTCAATGAATATTTATG
ACGATTCCGCAGTATTGGACGCTATCCAGTCTAAACATTTTACTATTACCCCTCTGG
CAAACTTCTTTTGCAAAAGCCTCTCGCTATTTTGGTTTTTATCGTCGTCTGGTAAAC
GAGGGTTATGATAGTGTTGCTCTTACTATGCCTCGTAATTCCTTTTGGCGTTATGTAT
CTGCATTAGTTGAATGTGGTATTCCTAAATCTCAACTGATGAATCTTCTACCTGTAA
TAATGTTGTTCCGTTAGTTCGTTTTATTAACGTAGATTTTTCTTCCCAACGTCCTGACT

10

20

【化 2 - 2】

GGTATAATGAGCCAGTTCCTAAAATCGCATAAGGTAATTCACAATGATTAAAGTTGA
AATTAACCATCTCAAGCCCAATTTACTACTCGTTCTGGTGTTCCTCGTCAGGGCAA
GCCTTATTCAGTGAATGAGCAGCTTTGTTACGTTGATTTGGGTAATGAATATCCGGTT
CTTGTCAGATTACTCTTGATGAAGGTCAGCCAGCCTATGCGCCTGGTCTGTACACC
GTTTCATCTGTCCTCTTTCAAAGTTGGTCAGTTCGGTTCCTTATGATTGACCGTCTGC
GCCTCGTTCGGGCTAAGTAACATGGAGCAGGTCGCGGATTCGACACAATTTATCAG
GCGATGATACAAATCTCCGTTGTACTTTGTTTCGCGCTTGGTATAATCGCTGGGGGTC
AAAGATGAGTGTTTTAGTGTATTCTTTCGCTCTTTCGTTTTAGGTTGGTGCCTTCGT
AGTGGCATTACGTATTTTACCCGTTTAAATGGAACTTCCTCATGAAAAAGTCTTTAGT
CCTCAAAGCCTCTGTAGCCGTTGCTACCCTCGTTCGATGCTGTCTTTCGCTGCTGAG
GGTGACGATCCCGCAAAGCGGCCTTTAACTCCCTGCAAGCCTCAGCGACCGAATA
TATCGGTTATGCGTGGGCGATGGTTGTTGTCATTGTTCGGCGCAACTATCGGTATCAA
GCTGTTTAAAGAAATTCACCTCGAAAGCAAGCTGATAAACCGATACAATTAAGGCT
CCTTTTGGAGCCTTTTTTTTTGGAGATTTTCAACATGAAAAAATTATTATTCGCAATT
CCTTTAGTTGTTCTTTCTATTCTCACTCCGCTGAAACTGTTGAAAGTTGTTTAGCAA
AACCCCATACAGAAAATTCATTTACTAACGTCTGGAAAGACGACAAAACCTTTAGATC
GTTACGCTAACTATGAGGGTGTCTGTGGAATGCTACAGGCGTTGTAGTTTGTACTG
GTGACGAAACTCAGTGTACGGTACATGGGTTCCCTATTGGGCTTGCTATCCCTGAAA
ATGAGGGTGGTGGCTCTGAGGGTGGCGGTTCTGAGGGTGGCGGTTCTGAGGGTGGC
GGTACTAAACCTCCTGAGTACGGTGATACACCTATTCCGGGCTATACTTATATCAAC
CCTCTCGACGGCACTTATCCGCCTGGTACTGAGCAAAACCCCGCTAATCCTAATCCT
TCTCTTGAGGAGTCTCAGCCTCTTAATACTTTCATGTTTCAGAATAATAGGTTCCGAA
ATAGGCAGGGGGCATTAACTGTTTATACGGGCACTGTTACTCAAGGCACTGACCCCG
TTAAAACCTTATTACCAGTACACTCCTGTATCATCAAAAGCCATGTATGACGCTTACT
GGAACGGTAAATTCAGAGACTGCGCTTTCATTCTGGCTTTAATGAGGATCCATTCTG
TTTGTGAATATCAAGGCCAATCGTCTGACCTGCCTCAACCTCCTGTCAATGCTGGCG
GCGGCTCTGGTGGTGGTTCTGGTGGCGGCTCTGAGGGTGGTGGCTCTGAGGGTGGCG
GTTCTGAGGGTGGCGGCTCTGAGGGAGGCGGTTCCGGTGGTGGCTCTGGTTCCGGTG
ATTTTGATTATGAAAAGATGGCAAACGCTAATAAGGGGGCTATGACCGAAAATGCC
GATGAAAACGCGCTACAGTCTGACGCTAAAGGCAAACCTTGATTCTGTCGCTACTGAT

10

20

30

40

【化 2 - 3】

TACGGTGCCTGCTATCGATGGTTTCATTGGTGACGTTTCCGGCCTTGCTAATGGTAATG
GTGCTACTGGTGATTTTGCTGGCTCTAATCCCAAATGGCTCAAGTCGGTGACGGTG
ATAATCACCTTTAATGAATAATTTCCGTCAATATTTACCTCCCTCCCTCAATCGGT
TGAATGTGCCCCCTTTTGTCTTTAGCGCTGGTAAACCATATGAATTTTCTATTGATTGT
GACAAAATAAACTTATTCCGTGGTGTCTTTGCGTTTCTTTTATATGTTGCCACCTTTA
TGTATGTATTTTCTACGTTTGCTAACATACTGCGTAATAAGGAGTCTTAATCATGCCA
GTTCTTTTGGGTATTCCGTTATTATTGCGTTTCCTCGGTTTCCTTCTGGTAACTTTGTT
CGGCTATCTGCTTACTTTTCTTAAAAAGGGCTTCGGTAAGATAGCTATTGCTATTCA
TTGTTTCTTGCTCTTATTATTGGGCTTAACTCAATTCTTGTGGGTTATCTCTCTGATAT
TAGCGCTCAATTACCCTCTGACTTTGTTTCAGGGTGTTCAGTTAATTCTCCCGTCTAAT
GCGCTTCCCTGTTTTTATGTTATTCTCTCTGTAAAGGCTGCTATTTTCATTTTTGACGT
TAAACAAAAAATCGTTTCTTATTTGGATTGGGATAAATAATATGGCTGTTTATTTTGT
AACTGGCAAATTAGGCTCTGGAAAGACGCTCGTTAGCGTTGGTAAGATTCAGGATA
AAATTGTAGCTGGGTGCAAAATAGCAACTAATCTTGATTTAAGGCTTCAAAACCTCC
CGCAAGTCGGGAGGTTTCGCTAAAACGCCTCGCGTTCTTAGAATACCGGATAAGCCTT
CTATATCTGATTTGCTTGCTATTGGGCGCGGTAATGATTCCTACGATGAAAATAAAA
ACGGCTTGCTTGTCTCGATGAGTGCGGTACTTGGTTTAATACCCGTTCTTGGAAATGA
TAAGGAAAGACAGCCGATTATTGATTGGTTTCTACATGCTCGTAAATTAGGATGGGA
TATTATTTTTCTTGTTCAGGACTTATCTATTGTTGATAAACAGGCGCGTTCTGCATTA
GCTGAACATGTTGTTTATTGTCGTCGTCTGGACAGAATTACTTTACCTTTTGTGCGGTA
CTTTATATTCTTTATTACTGGCTCGAAAATGCCTCTGCCTAAATTACATGTTGGCGT
TGTTAAATATGGCGATTCTCAATTAAGCCCTACTGTTGAGCGTTGGCTTTATACTGGT
AAGAATTTGTATAACGCATATGATACTAAACAGGCTTTTTCTAGTAATTATGATTCC
GGTGTTTATTCTTATTTAACGCCTTATTTATCACACGGTCGGTATTTCAAACCATTAA
ATTTAGGTCAGAAGATGAAATTAATACTAAAATATATTTGAAAAAGTTTTCTCGCGTTC
TTGTCTTGCGATTGGATTTGCATCAGCATTACATATAGTTATATAACCCAACCTAA
GCCGGAGGTTAAAAAGGTAGTCTCTCAGACCTATGATTTTGATAAATTCACTATTGA
CTCTTCTCAGCGTCTTAATCTAAGCTATCGCTATGTTTTCAAGGATTCTAAGGGAAAA
TTAATTAATAGCGACGATTTACAGAAGCAAGGTTATTCACTCACATATATTGATTTA
TGTACTGTTTCCATTAATAAAAAGGTAATTCAAATGAAATTGTTAAATGTAATTAATTTT

10

20

30

40

【化 2 - 4】

GTTTTCTTGATGTTTGTTCATCATCTTCTTTTGCTCAGGTAATTGAAATGAATAATTC
GCCTCTGCGCGATTTTGTAACTTGGTATTCAAAGCAATCAGGCGAATCCGTTATTGTT
TCTCCCGATGTAAAAGGTAAGTACTGTTACTGTATATTCATCTGACGTTAAACCTGAAAAT
CTACGCAATTTCTTTATTTCTGTTTTACGTGCTAATAATTTTGATATGGTTGGTTCAAT
TCCTTCCATAATTCAGAAGTATAATCCAAACAATCAGGATTATATTGATGAATTGCC
ATCATCTGATAATCAGGAATATGATGATAATTCGCTCCTTCTGGTGGTTTCTTTGTT
CCGCAAAATGATAATGTTACTCAAACCTTTTAAAATTAATAACGTTCCGGGCAAAGGAT
TTAATACGAGTTGTGAATGTTTGTAAAGTCTAATACTTCTAAATCCTCAAATGTAT
TATCTATTGACGGCTCTAATCTATTAGTTGTTAGTGCACCTAAAGATATTTTAGATAA
CCTTCCTCAATTCCTTTCTACTGTTGATTTGCCAACTGACCAGATATTGATTGAGGGT
TTGATATTTGAGGTTCCAGCAAGGTGATGCTTTAGATTTTTCATTTGCTGCTGGCTCTC
AGCGTGGCACTGTTGCAGGCGGTGTTAATACTGACCGCCTCACCTCTGTTTTATCTTC
TGCTGGTGGTTCGTTCCGGTATTTTTAATGGCGATGTTTTAGGGCTATCAGTTCGCGCA
TTAAAGACTAATAGCCATTCAAAAATATTGTCTGTGCCACGTATTCTTACGCTTTCAG
GTCAGAAGGGTCTATCTCTGTTGGCCAGAATGTCCCTTTTATTACTGGTCGTGTGAC
TGGTGAATCTGCCAATGTAAATAATCCATTTCCAGACGATTGAGCGTCAAAATGTAGG
TATTTCCATGAGCGTTTTTCCCTGTTGCAATGGCTGGCGGTAATATTGTTCTGGATATT
ACCAGCAAGGCCGATAGTTTGAGTTCTTCTACTCAGGCAAGTGATGTTATTAATAAT
CAAAGAAGTATTGCTACAACGGTTAATTTGCGTGATGGACAGACTCTTTTACTCGGT
GGCCTCACTGATTATAAAAACACTTCTCAAGATTCTGGCGTACCGTTCCTGTCTAAA
ATCCCTTTAATCGGCCTCCTGTTTAGCTCCCGCTCTGATTCCAACGAGGAAAGCACG
TTATACGTGCTCGTCAAAGCAACCATAGTACGCGCCCTGTAGCGGCGCATTAAAGCGC
GGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGC
CCGCTCCTTTCGCTTCTTCCCTTCTTCTCGCCACGTTCCGCCGGCTTTCCCGTCAA
GCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGAC
CCCAAAAACCTTGATTTGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACG
GTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAA
CTGGAACAACACTCAACCCTATCTCGGGCTATTCTTTTGATTTATAAGGGATTTTGCC
GATTTCCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTT
TAACAAAATATTAACGTTTACAATTTAAATATTTGCTTATACAATCTTCCCTGTTTTTG

10

20

30

40

【化 2 - 5】

GGGCTTTTCTGATTATCAACCGGGGTACATATGATTGACATGCTAGTTTTACGATTAC
 CGTTCATCGATTCTCTTGTGTTGCTCCAGACTCTCAGGCAATGACCTGATAGCCTTTGT
 AGACCTCTCAAAAATAGCTACCCTCTCCGGCATGAATTTATCAGCTAGAACGGTTGA
 ATATCATATTGATGGTGATTTGACTGTCTCCGGCCTTTCTCACCCCTTTTGAATCTTTA
 CCTACACATTACTCAGGCATTGCATTTAAAATATATGAGGGTTCTAAAAATTTTTATC
 CTTGCGTTGAAATAAAGGCTTCTCCCGCAAAGTATTACAGGGTCATAATGTTTTTG
 GTACAACCGATTTAGCTTTATGCTCTGAGGCTTTATTGCTTAATTTTGCTAATTCTTT
 GCCTTGCCTGTATGATTTATTGGATGTT (配列番号 21)

10

【 0 2 6 4 】

遺伝子II : 連結 (6006..6407,1..831)

【 0 2 6 5 】

【化 3】

翻訳 =MIDMLVLRLPFIDSLVCSRLSGNDLIAFVDLSKIATLSGMNLSA
 RTVEYHIDGDLTVSGLSHPFESLPTHYSGLAFKIYEGSKNFYPCVEIKASPAKVLQGHNVF
 GTTDLALCSEALLNFANSLPCLYDLLDVNATTISRIDATFSARAPNENIAKQVIDHLRNV
 SNGQTKSTRSQNWESTVTWNETSRRHTLVAYLKHVELQHQQQLSSKPSAKMTSYQKE
 QLKVLSNPDLLEFASGLVRFEARIKTRYLKSFGLPLNLFDAIRFASDYN SQGKDLIFDLW
 SFSFSELFKAFEGDSMNIYDDSAVLDAIQSKHFTITPSGKTSFAKASRYFGFYRRLVNEGY
 DSVALTMPRNSFWRYVSALVECGIPKSQLMNLSTCANNVPLVRFINVD FSSQRPDWYNE
 PVLKIA (配列番号 22)

20

【 0 2 6 6 】

遺伝子X : 496..831

【 0 2 6 7 】

【化 4】

翻訳 =MNIYDDSAVLDAIQSKHFTITPSGKTSFAKASRYFGFYRRLVN
 EGYDSVALTMPRNSFWRYVSALVECGIPKSQLMNLSTCANNVPLVRFINVD FSSQRPDW
 YNEPVLKIA (配列番号 23)

30

【 0 2 6 8 】

遺伝子V : 843..1106

【 0 2 6 9 】

【化 5】

翻訳 =MIKVEIKPSQAQFTTRSGVSRQGKPYSLNEQLCYVDLGNEYPV
 LVKITLDEGQPAYAPGLYTVHLSSFKVGQFGLMIDRLRLVPAK (配列番号 24)

40

【 0 2 7 0 】

遺伝子VII : 1108..1209

【 0 2 7 1 】

翻訳=MEQVADFDTIYQAMIQISVVLCFALGIIAGGQR (配列番号 25)

【 0 2 7 2 】

遺伝子IX : 1206..1304

50

【 0 2 7 3 】

翻訳=MSVLVYSFASFVLGWCLRSGITYFTRLMETSS (配列番号26)

【 0 2 7 4 】

遺伝子VIII : 1301..1522

【 0 2 7 5 】

【 化 6 】

翻訳 =MKKSLVLKASVAVATLVPMLSFAAEGDDPAKAAFNSLQASA

TEYIGYAWAMVVVIVGATIGIKLFFKKFTSKAS (配列番号 27)

10

【 0 2 7 6 】

遺伝子III : 1579..2853

【 0 2 7 7 】

【 化 7 】

翻訳 =MKKLLFAIPLVVPFYSHSAETVESCLAKPHTENSFTNVWKDD

KTLDRYANYEGCLWNATGVVVCTGDETQCYGTWVPIGLAIPENEGGGSEGGGSEGGGS

EGGGTKPPEYGDTPIPGYTYINPLDGTYPGTEQNPANPNPSLEESQPLNTFMFQNNRFR

NRQGALTVYTGTVTQGTDPVKTYQYTPVSSKAMYDAYWNGKFRDCAFHSGFNEDPF

VCEYQGQSSDLPQPPVNAGGGSGGGSGGGSEGGGSEGGGSEGGGSGGGSGSGD

FDYEKMANANKGAMTENADENALQSDAKGKLDVATDYGAIDGFIGDVSGLANGNG

ATGDFAGSNSQMAQVGDGDN SPLMNNFRQYLSPLQSV ECRPFVFSAGKPYEFSIDCDK

INLFRGVFAFLLYVATFMYVFSTFANILRNKES (配列番号 28)

20

【 0 2 7 8 】

遺伝子VI : 2856.3194

【 0 2 7 9 】

【 化 8 】

翻訳 =MPVLLGIPLLLRFLGFLLVTLFGYLLTFLKKGFGKIAIAISLFLA

LIIGLNSILVGYLSDISAQLPSDFVQGVQLILPSNALPCFYVILSVKAAIFIFDVKQKIVSYL

DWDK (配列番号 29)

30

【 0 2 8 0 】

遺伝子I : 3196..4242

【 0 2 8 1 】

【 化 9 】

翻訳 =MAVYFVTGKLGSGKTLVSVGKIQDKIVAGCKIATNLDLRLQN

LPQVGRFAKTPRVLRIPDKPSISDLLAIGRGNDSYDENKNGLLVLDECGTWFNTRSWND

KERQPIIDWFLHARKLGWDIIFLVQDLSIVDKQARSALAEHVVYCRRLDRITLFPVGTLY

SLITGSKMPLPKLHVGVVVKYGDSQLSPTVERWLYTGKNLYNAYDTKQAFSSNYDSGVY

SYLTPYLSHGRYFKPLNLGQKMKLTKIYLKKFSRVLCLAIGFASAFTYSYITQPKPEVKK

VVSQTYDFDKFTIDSSQRLNLSYRYVFKDSKGLINSDDLQKQGYSLTYIDLCTVSIKKG

NSNEIVKCN (配列番号 30)

40

【 0 2 8 2 】

遺伝子IV : 4220..5500

50

【 0 2 8 3 】

【 化 1 0 】

翻訳 =MKLLNVINFLVFLMFVSSSSFAQVIEMNNSPLRDFVTWYSKQSG
 ESVIVSPDVKGTVTVYSSDVKPENLRNFFISVLRANNFDMVGSIPSIIQKYNPNNQDYIDE
 LPSSDNQEYDDNSAPSGGFFVPQNDNVTQTFKINNVRAKDLIRVVELFVKSNTSKSSNVL
 SIDGSNLLVVSAPKDILDNLPQFLSTVDLPTDQILIEGLIFEVQQGDALDFSFAAGSQRGT
 VAGGVNTDRLTSLSSAGGSFGIFNGDVLGLSVRALKTNSHSHKILSVPRILTLGQKGSIS
 VGQNVPFITGRVTGESANVNNPFQTIERQNVGISMSVFPVAMAGGNIVLDITSKADSLSS
 STQASDVITNQRSIATTVNLRDGTLLGLTDYKNTSQDSGVPFLSKIPLIGLLFSSRSDS
 NEESTLYVLVKATIVRAL (配列番号 31)

10

【 0 2 8 4 】

蛋白質 / ペプチド / ポリペプチド

用語「蛋白質」、「ペプチド」、および「ポリペプチド」は本明細書において交換可能に用いられ、ペプチド(アミド)結合によって一緒になって連結されたアミノ酸残基同士のポリマーを言う。用語は、いずれかのサイズ、構造、または機能の蛋白質、ペプチド、またはポリペプチドを言う。典型的には、蛋白質、ペプチド、またはポリペプチドは少なくとも3アミノ酸の長さであろう。蛋白質、ペプチド、またはポリペプチドは、個々の蛋白質または蛋白質の集まりを言い得る。蛋白質、ペプチド、またはポリペプチド中のアミノ酸の1つ以上は修飾され得、例えば、化学的実体、例えば炭水化物基、ヒドロキシル基、リン酸基、ファルネシル基、イソファルネシル基、脂肪酸基、コンジュゲーション、官能化、または他の改変のためのリンカーなどの追加による。蛋白質、ペプチド、またはポリペプチドは単分子でもまたあり得、または多分子複合体であり得る。蛋白質、ペプチド、またはポリペプチドは、天然に存在する蛋白質またはペプチドの単に断片であり得る。蛋白質、ペプチド、またはポリペプチドは、天然に存在するか、組換え体か、もしくは合成か、またはそれらのいずれかの組み合わせであり得る。本明細書において用いられる用語「融合蛋白質」は、少なくとも2つの異なる蛋白質からの蛋白質ドメインを含むハイブリッドポリペプチドを言う。1つの蛋白質は融合蛋白質のアミノ末端(N末端)部分にまたはカルボキシ末端(C末端)部分(protein)に見つかり得、それゆえにそれぞれ「アミノ末端融合蛋白質」または「カルボキシ末端融合蛋白質」を形成する。蛋白質は、異なるドメイン、例えば核酸結合ドメイン(例えば、標的部位への蛋白質の結合を導くCas9のgRNA結合ドメイン)およびリコンビナーゼの核酸切断ドメインまたは触媒ドメインを含み得る。いくつかの態様では、蛋白質は、蛋白質部分、例えば核酸結合ドメインを構成するアミノ酸配列と、有機化合物、例えば核酸切断薬剤として作用し得る化合物とを含む。いくつかの態様では、蛋白質は、核酸、例えばRNAとの複合体であるかまたは結び付けられている。本明細書において提供される蛋白質のいずれかは当分野において公知のいずれかの方法によって産生され得る。例えば、本明細書において提供される蛋白質は組換え蛋白質発現および精製によって産生され得、これはペプチドリナーを含む融合蛋白質にとりわけ適している。組換え体蛋白質発現および精製のための方法は周知であり、Green and Sambrook, *Molecular Cloning: A Laboratory Manual* (4th ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (2012))によって記載されているものを包含し、これの内容全体は参照によって本明細書に組み込まれる。

20

30

40

【 0 2 8 5 】

蛋白質スプライシング

50

本明細書において用いられる用語「蛋白質スプライシング」は、前駆体蛋白質の内側領域（インテイン）が切り出され、蛋白質のフランキング領域（エクステイン）同士がライゲーションされて成熟蛋白質を形成するプロセスを言う。この天然プロセスは原核生物および真核生物両方からの数々の蛋白質で観察されている（Perler, F. B., Xu, M. Q., Paulus, H. Current Opinion in Chemical Biology 1997, 1, 292-299; Perler, F. B. Nucleic Acids Research 1999, 27, 346-347）。インテイン単位は蛋白質スプライシングを触媒するために必要とされる必要な構成要素を含有し、多くの場合には、インテイン可動性に関するエンドヌクレアーゼドメインを含有する（Perler, F. B., Davis, E. O., Dean, G. E., Gimble, F. S., Jack, W. E., Neff, N., Noren, C. J., Thomer, J., Belfort, M. Nucleic Acids Research 1994, 22, 1127-1127）。もたらされる蛋白質同士は連結され、別個の蛋白質としては発現されない。蛋白質スプライシングはトランスにもまたとり行われ得、別個のポリペプチドとして発現された分割型インテインが自発的に組み合わさって1個のインテインを形成し、それから、これは蛋白質スプライシングプロセスを経験して別個の蛋白質をつなぎ合わせる。

10

【0286】

組換え体

蛋白質または核酸のコンテキストにおいて本明細書において用いられる用語「組換え体」は、天然には存在せず、ヒトの操作している産物である蛋白質または核酸を言う。例えば、いくつかの態様では、組換え体蛋白質または核酸分子は、いずれかの天然に存在する配列と比較して少なくとも1、少なくとも2、少なくとも3、少なくとも4、少なくとも5、

20

【0287】

参照塩基編集因子

本明細書において用いられる用語「参照塩基編集因子」は、進化した塩基編集因子を達成するための連続的進化プロセス、例えばPACEの出発点として用いられる塩基編集因子のバージョンを言う。参照塩基編集因子は天然に存在するポリペプチド配列を包含し得る。参照塩基編集因子は、天然に存在しないポリペプチド配列、例えば、アミノ酸配列の1つ以上の変化（例えば、野生型または古典的なポリペプチドと比べて1つ以上の変異した残基、1つ以上のアミノ酸の挿入、または1つ以上のアミノ酸の欠失）を有する塩基編集因子をもまた包含し得る。換言すると、参照塩基編集因子は、天然に存在する塩基編集因子構成要素（例えば、デアミナーゼおよびCas9）を含み得るか（例えば、野生型ヒト、マウス、ラット、ウマ、もしくはウサギポリペプチド配列またはそれらの天然に存在するバリエーション）、あるいは、それらは、天然に存在する配列と比べてすでに改変されており、かつ本明細書に記載される連続的進化プロセス、例えばPACEを用いてさらに進化および/または変化および/または改善されることが所望である塩基編集因子をもまた包含し得る。塩基編集因子の個々の構成要素を参照するときには、同様の定義が遵守されるであろう。例えば、「参照Cas9ドメイン」または「参照デアミナーゼ」または「参照UGI」または塩基編集因子の他のかかる個々の構成要素は、その構成要素またはドメインの進化したバージョンまたはバリエーションを達成するかまたは得るための連続的進化プロセス、例えばPACE

30

40

【0288】

RNAによってプログラム可能なヌクレアーゼ / RNAによってガイドされるヌクレアーゼ

用語「RNAによってプログラム可能なヌクレアーゼ」および「RNAによってガイドされるヌクレアーゼ」は本明細書において交換可能に用いられ、切断の標的ではない1つ以上のRNAと複合体を形成する（例えば、結合または会合する）ヌクレアーゼを言う（例えば、Cas9またはそのホモログもしくはバリエーション）。いくつかの態様では、RNAによってプログラム可能なヌクレアーゼは、RNAとの複合体であるときには、ヌクレアーゼ:RNA複合体と言われ得る。典型的には、結合されるRNA（単数または複数）はガイドRNA（gRNA）と言われる。gRNAは2つ以上のRNAの複合体としてまたは1個のRNA分子として存在し得る。1個のR

50

NA分子として存在するgRNAはシングルガイドRNA (sgRNA) と言われ得るが、「gRNA」は、1個の分子としてまたは2つ以上の分子の複合体としてどちらかで存在するガイドRNAを言うために交換可能に (interchangeably) 用いられる。典型的には、1個のRNA種として存在するgRNAは2つのドメインを含む：(1) 標的核酸に対して相同性を共有する (例えば、かつ標的へのCas9 (または同等物) 複合体の結合を導く) ドメイン；および(2) Cas9蛋白質に結合するドメイン。いくつかの態様では、ドメイン(2) はtracrRNAとして公知の配列に対応し、ステムループ構造を含む。例えば、いくつかの態様では、ドメイン(2) はJinek et al., *Science* 337: 816-821 (2012)の図1Eに図示されている通りtracrRNAに対して相同的であり、これの内容全体は参照によって本明細書に組み込まれる。gRNAの他の例 (例えば、ドメイン2を包含するもの) は、「Switchable Cas9 Nucleases And Uses Thereof」と題する2013年9月6日出願のU.S. 仮特許出願U.S.S.N.61/874,682および「Delivery System For Functional Nucleases」と題する2013年9月6日出願のU.S. 仮特許出願U.S.S.N.61/874,746に見出され得、それぞれの内容全体はそれらの全体が参照によってここに組み込まれる。いくつかの態様では、gRNAはドメイン(1) および(2) の2つ以上を含み、「拡張型gRNA」と言われ得る。例えば、拡張型gRNAは例えば2つ以上のCas9蛋白質に結合し、本明細書に記載される通り、2つ以上の別個の領域において標的核酸に結合するであろう。gRNAは標的部位を相補するヌクレオチド配列を含み、これは前記標的部位へのヌクレアーゼ/RNA複合体の結合を媒介し、ヌクレアーゼ:RNA複合体の配列特異性を提供する。いくつかの態様では、RNAによってプログラム可能なヌクレアーゼは、(CRISPR関連システム) Cas9エンドヌクレアーゼ、例えば*Streptococcus pyogenes*からのCas9 (Csn1) である (例えば、"Complete genome sequence of an M1 strain of *Streptococcus pyogenes*." Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K., Priemaux C, Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P., Qian Y., Jia H. G., Najar F.Z., Ren Q., Zhu H., Song L., White J., Yuan X., Clifton S.W., Roe B. A., McLaughlin R.E., *Proc. Natl. Acad. Sci. U.S.A.* 98: 4658-4663 (2001) ; "CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III." Deltchev a E., Chylinski K., Sharma CM., Gonzales K., Chao Y., Pirzada Z.A., Eckert M.R., Vogel J., Charpentier E., *Nature* 471: 602-607 (2011) ; および "A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity." Jinek M., Chylinski K., Fonfara I., Hauer M., Doudna J.A., Charpentier E. *Science* 337: 816-821 (2012) を参照する。これらのそれぞれの内容全体は参照によって本明細書に組み込まれる)。

10

20

30

【 0 2 8 9 】

RNAによってプログラム可能なヌクレアーゼ (例えばCas9) はRNA : DNAハイブリダイゼーションを用いてDNA切断部位を標的化するので、これらの蛋白質は、原理的には、ガイドRNAによって規定されるいずれかの配列へと標的化されることが可能である。RNAによってプログラム可能なヌクレアーゼ、例えばCas9を部位特異的切断のために (例えば、ゲノムを改変するために) 用いる方法は当分野において公知である (例えば、Cong, L. et al. *Multiplex genome engineering using CRISPR/Cas systems.* *Science* 339, 819-823 (2013) ; Mali, P. et al. *RNA-guided human genome engineering via Cas9.* *Science* 339, 823-826 (2013) ; Hwang, W.Y. et al. *Efficient genome editing in zebrafish using a CRISPR-Cas system.* *Nature biotechnology* 31, 227-229 (2013) ; Jinek, M. et al. *RNA-programmed genome editing in human cells.* *eLife* 2, e00471 (2013) ; Dicarlo, J.E. et al. *Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems.* *Nucleic Acids research* (2013) ; Jiang, W. et al. *RNA-guided editing of bacterial genomes using CRISPR-Cas systems.* *Nature biotechnology* 31, 233-239 (2013) を参照する ; これらのそれぞれの内容全体は参照によって本明細書に組み込まれる) 。

40

【 0 2 9 0 】

セレクションファージ

本明細書におい用語「セレクションプラスミド」と交換可能に用いられる用語用語「セ

50

レクシオンファージ」は、進化させられるべきtRNAシンターゼをコードする核酸配列を含み、かつ感染性のファージ粒子の生成のために要求される蛋白質をコードする全長遺伝子を欠く改変されたファージを言う。例えば、本明細書において提供されるいくつかのM13レクシオンファージは、進化させられるべき遺伝子をコードする核酸配列を例えばM13プロモーターのコントロール下に含み、かつ感染性のファージ粒子の生成のために要求されるある蛋白質をコードするファージ遺伝子、例えば、gI、gII、gIII、gIV、gV、gVI、gVII、gVIII、gIX、もしくはgX、またはそれらのいずれかの組み合わせの全てまたは一部を欠く。例えば、本明細書において提供されるいくつかのM13レクシオンファージは、進化させられるべきtRNAシンターゼ蛋白質をコードする核酸配列を例えばM13プロモーターのコントロール下に含み、かつ感染性ファージ粒子の生成のために要求されるある蛋白質をコードする遺伝子、例えばpIII蛋白質をコードするgIII遺伝子の全てまたは一部を欠く。

10

【0291】

配列コンテキストアグノスティック

本明細書において用いられる用語「配列コンテキストアグノスティック」は、所望の標的編集部位の近位の（上流のおよび/または下流の）配列が、進化した塩基編集因子が所望の標的編集部位を編集する効率へのインパクトまたは効果を有さないかまたはほとんど有さない、本明細書に記載される進化した塩基編集因子の所望の特性または性質を言う。

【0292】

分割型インテイン

同定されたインテイン遺伝子の小さい画分（5%未満）は分割型インテインをコードする⁹。より普通の一続きのインテインとは違って、これらは2つの別個のポリペプチド、NインテインおよびCインテインとして転写および翻訳され、それぞれは1つのエクステインに融合させられている。翻訳によって、インテイン断片同士は古典的なインテイン構造へと自発的にかつ非共有結合的にアセンブリして、トランスの蛋白質スプライシングを行う。

20

【0293】

対象

本明細書において用いられる用語「対象」は個々の生物、例えば個々の哺乳動物を言う。いくつかの態様では、対象はヒトである。いくつかの態様では、対象は非ヒト哺乳動物である。いくつかの態様では、対象は非ヒト霊長類である。いくつかの態様では、対象は齧歯類である。いくつかの態様では、対象はヒツジ、ヤギ、畜牛、ネコ、またはイヌである。いくつかの態様では、対象は脊椎動物、両生類、爬虫類、魚類、昆虫、ハエ、または線虫である。いくつかの態様では、対象は研究動物である。いくつかの態様では、対象は遺伝子操作されており、例えば遺伝子操作された非ヒト対象である。対象はどちらかの性別および発達のいずれかのステージであり得る。

30

【0294】

標的部位

用語「標的部位」は、デアミナーゼまたはデアミナーゼを含む融合蛋白質（例えば、本明細書において提供されるdCas9-デアミナーゼ融合蛋白質）によって脱アミノ化される核酸分子内の配列を言う。

40

【0295】

ベクター

本明細書において用いられる用語「ベクター」は、目当ての遺伝子をコードするように改変され得、かつ宿主細胞に侵入し、宿主細胞内で変異および複製し、それから複製された形態のベクターを別の宿主細胞に伝播させることができる核酸を言う。例示的な好適なベクターは、ウイルスベクター、例えばレトロウイルスベクターまたはバクテリオファージおよび繊維状ファージ、ならびに接合プラスミドを包含する。追加の好適なベクターは本開示に基づいて当業者には明らかであろう。

50

【0296】

ウイルスライフサイクル

本明細書において用いられる用語「ウイルスライフサイクル」はウイルス繁殖サイクルを言い、宿主細胞へのウイルスゲノムの挿入、宿主細胞内のウイルスゲノムの複製、および宿主細胞によるウイルス粒子へのウイルスゲノムの複製産物のパッケージングを含む。

【0297】

ウイルス粒子

本明細書において用いられる用語「ウイルス粒子」は、ウイルス蛋白質（単数または複数）のコートと、かついくつかのケースでは脂質のエンベロープと結びつけられているウイルスゲノム、例えばDNAまたはRNAゲノムを言う。例えば、ファージ粒子は、野生型ファージゲノムによってコードされる蛋白質中にパッケージングされたファージゲノムを含む。

10

【0298】

ウイルスベクター

本明細書において用いられる用語「ウイルスベクター」は、好適な宿主細胞に導入されたときに、ウイルスゲノムを別の宿主細胞に伝播させることができるウイルス粒子へと複製およびパッケージングされ得る、ウイルスゲノムを含む核酸を言う。用語ウイルスベクターは、短縮されたまたは部分的なウイルスゲノムを含むベクターまで拡張する。例えば、いくつかの態様では、感染性のウイルス粒子の生成に必要な蛋白質をコードする遺伝子を欠くウイルスベクターが提供される。しかしながら、好適な宿主細胞、例えば、欠けている遺伝子を条件付きプロモーターのコントロール下に含む宿主細胞内においては、かかる短縮されたウイルスベクターは複製し得、短縮されたウイルスゲノムを別の宿主細胞に伝播させることができるウイルス粒子を生成し得る。いくつかの態様では、ウイルスベクターは、ファージ、例えば繊維状ファージ（例えばM13ファージ）である。いくつかの態様では、進化させられるべき目当ての遺伝子を含むウイルスベクター、例えばファージベクターが提供される。

20

【0299】

ウラシルグリコシラーゼ阻害因子またはUGI

本明細書において用いられる用語「ウラシルグリコシラーゼ阻害因子」または「UGI」は、ウラシルDNAグリコシラーゼ塩基除去修復酵素を阻害する能力がある蛋白質を言う。いくつかの態様では、UGIドメインは野生型UGIまたは配列番号10に提示されているUGIを含む。いくつかの態様では、本明細書において提供されるUGI蛋白質は、UGIの断片、およびUGIまたはUGI断片に対して相同的な蛋白質を包含する。例えば、いくつかの態様では、UGIドメインは配列番号10に提示されているアミノ酸配列の断片を含む。いくつかの態様では、UGI断片は、配列番号10に提示されているアミノ酸配列の少なくとも60%、少なくとも65%、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、または少なくとも99.5%を含むアミノ酸配列を含む。いくつかの態様では、UGIは、配列番号10に提示されているアミノ酸配列に対して相同的なアミノ酸配列、または配列番号10に提示されているアミノ酸配列の断片に対して相同的なアミノ酸配列を含む。いくつかの態様では、UGIもしくはUGIの断片またはUGIもしくはUGI断片のホモログを含む蛋白質は「UGIパリアント」と言われる。UGIパリアントはUGIまたはその断片に対する相同性を共有する。例えば、UGIパリアントは、野生型UGIまたは配列番号10に提示されているUGIと少なくとも70%同一、少なくとも75%同一、少なくとも80%同一、少なくとも85%同一、少なくとも90%同一、少なくとも95%同一、少なくとも96%同一、少なくとも97%同一、少なくとも98%同一、少なくとも99%同一、少なくとも99.5%同一、または少なくとも99.9%同一である。いくつかの態様では、UGIパリアントは、断片が野生型UGIまたは配列番号10に提示されているUGIの対応する断片と少なくとも70%同一、少なくとも80%同一、少なくとも90%同一、少なくとも95%同一、少なくとも96%同一、少なくとも97%同一、少なくとも98%同一、少なくとも99%同一、少なくとも99.5%同一、または少なくとも99.9%であるような、UGIの断片

30

40

50

を含む。いくつかの態様では、UGIは次のアミノ酸配列を含む：MTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKML（配列番号10）（P14739|UNGI_BPPB2ウラシルDNAグリコシラーゼ阻害因子）。

【0300】

処置

用語「処置」、「処置する」、および「処置すること」は、本明細書に記載される通り、疾患もしくは異常またはその1つ以上の症状を逆転させるか、軽減するか、その始まりを遅延させるか、またはその進行を阻害することを狙った臨床的介入を言う。本明細書において用いられる用語「処置」、「処置する」、および「処置すること」は、本明細書に記載される通り、疾患もしくは異常またはその1つ以上の症状を逆行させるか、軽減するか、その始まりを遅延させるか、またはその進行を阻害することを狙った臨床的介入を言う。いくつかの態様では、処置は、1つ以上の症状が発生した後におよび/または疾患が診断された後に投与され得る。他の態様では、処置は、例えば症状の始まりを防止するかもしくは遅延させるかまたは疾患の始まりもしくは進行を阻害するために、症状の不在下において投与され得る。例えば、処置は、（例えば、症状の既往に照らしておよび/または遺伝学的もしくは他の感受性因子に照らして）症状の始まりに先立って感受性の個人に投与され得る。処置は、例えばそれらの再発を防止するかまたは遅延させるために、症状が解消した後にもまた引き続き得る。

【0301】

バリエーション

本明細書において用いられる用語「バリエーション」は、天然に存在するものから逸脱するパターンを有する品質の呈示を意味すると取られるべきである。例えば、バリエーションCas9は、野生型Cas9アミノ酸配列と比較してアミノ酸残基に1つ以上の変化を含むCas9である。

【0302】

野生型

本明細書において用いられる用語「野生型」は、当業者によって理解される当分野の用語であり、変異体またはバリエーション形態から区別される、それが天然に存在する通りの生物、株、遺伝子、または性質の典型的な形態を意味する。

【発明を実施するための形態】

【0303】

本発明者らは、驚くべきことに、塩基編集因子の1つ以上のドメインまたは構成要素の機能を迅速に改善するために用いられ得る有効な連続的な進化的な変異導入プロセス（すなわちPACE）を開発することによって、改善された（improve）塩基編集因子（例えば、シチジン（cytidine）塩基編集因子）を発見した。

【0304】

本明細書は進化した塩基編集因子を提供し、これらは当分野のものの欠陥を克服し（増大した効率、および/または編集部位における特異的な配列コンテキストの減少した要求を包含する）、これらはファージによって支援される連続的な進化（PACE）システムの結果として得られる。特に、本明細書は進化したシチジン塩基編集因子（例えば、APOBEC1、CDA、またはAIDシチジンデアミナーゼドメインに基づく）を提供し、これらは当分野のものの欠陥を克服し（増大した効率、および/または編集部位における特異的な配列コンテキストの減少した要求を包含する）、これらはファージによって支援される連続的な進化（PACE）システムの結果として得られる。加えて、本明細書は、本明細書に記載される進化した塩基編集因子をコードおよび/または発現する核酸分子、ならびに本明細書に記載される進化した塩基編集因子を発現するための発現ベクターまたはコンストラクト、前記核酸分子および発現ベクターを含む宿主細胞、ならびに本明細書に記載される核酸に基づく態様を送達および/または投与するための組成物を提供する。加えて、本開示は、単離された進化した塩基編集因子、および前記単離された進化した塩基編集因子を含む組成物を、本明細書に記載される通り提供する。なお、さらに、本開示は、進化した塩基編集因

10

20

30

40

50

子を作る方法、および技術水準を形成する塩基編集因子と比較して改善された効率で、好ましくは配列コンテキストアグノスティックな（すなわち、所望の編集部位が特異的な配列コンテキストを要求しない）様式で、核酸分子、例えばゲノムを編集することを包含する用途に、進化した塩基編集因子または進化した塩基編集因子をコードする核酸分子を用いる方法を提供する。いくつかの態様では、本明細書において提供される作る方法は、改善されたファージによって支援される連続的進化（PACE）システムであり、これは、塩基編集因子の1つ以上の構成要素（例えば、Cas9ドメインまたはシチジンデアミナーゼドメイン）を迅速で連続的な様式で進化させるために利用され得る。本明細書は、本明細書に記載される塩基編集システムによって（例えば、本明細書に記載される単離された進化した塩基編集因子またはそれをコードするベクターもしくはコンストラクトの形態で）、標的核酸分子、例えばゲノムの1個の核酸塩基を効率的に編集し、塩基（based）編集を好ましくは配列コンテキストアグノスティックな様式でとり行うための方法をもまた提供する。なお、さらに、本明細書は、遺伝子疾患を処置するためおよび/または遺伝学的形質もしくは状態を変調もしくは変化させるための治療方法を提供し、標的核酸分子、例えばゲノムを塩基編集システムと（例えば、単離された進化した塩基編集因子蛋白質またはそれをコードするベクターの形態で）接触させることと、遺伝子疾患を処置するおよび/または遺伝学的形質（例えば、目の色）を変化させるために塩基（based）編集をとり行うこととによる。

10

【0305】

1. 進化した塩基編集因子

20

種々の側面では、本開示は、参照のまたは出発点の塩基編集因子（またはその構成要素もしくはドメイン）を連続的進化方法（例えばPACE）によって変異させる結果として、進化した塩基編集因子を提供する。種々の態様では、本開示は、参照のまたは出発点の塩基編集因子のアミノ酸配列と比べてそのアミノ酸配列に導入された1つ以上のアミノ酸バリエーションを有する進化した塩基編集因子を提供する。例えば、コード配列中のいずれかの特定の位置のコドンの変化、1つ以上のアミノ酸の欠失（例えば、短縮された蛋白質）、1つ以上のアミノ酸の挿入、または上述のいずれかの組み合わせをもたらす塩基編集因子をコードするヌクレオチド配列の変化の結果として、アミノ酸配列バリエーションは参照塩基編集因子のアミノ酸配列中に1つ以上の変異した残基を包含し得る。いくつかの態様では、進化した塩基編集因子は参照塩基編集因子と少なくとも50%、少なくとも55%、少なくとも60%、少なくとも65%、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも8%、少なくとも99%、または少なくとも99.5%同一である。進化した塩基編集因子は、塩基編集因子の1つ以上の構成要素またはドメインにバリエーションを包含し得る（例えば、Cas9ドメイン、デアミナーゼドメイン、もしくはUGIドメインに導入されたバリエーション、またはこれらのドメインの組み合わせに導入されたバリエーション）。

30

【0306】

ある側面では、塩基編集因子を進化させるための本明細書に記載される方法は、当分野において公知の塩基編集因子によって始まる。技術水準は本出願現在で数々の塩基編集因子を記載している。塩基編集因子を改善するために本明細書に記載される方法およびアプローチは、いずれかの以前に公知の塩基編集因子に、または将来（further）開発され得るが本方法および改変アプローチによって付与される有益な性質を欠く塩基編集因子に適用され得る。本発明の進化した塩基編集因子を達成するために本明細書に記載される方法によって改変され得る例示的な塩基編集因子は、例えば次の参照および/または特許公開に記載されているものを包含し得、これらのそれぞれはそれらの全体が参照によって組み込まれる：（a）PCT/US2014/070038（2015年6月18日にWO2015/089406として公開）およびUS内のまたは世界的なその同等物；（b）PCT/US2016/058344（2017年4月27日にWO2017/070632として公開）およびUS内のまたは世界的なその同等物；（c）PCT/US2016/058345（2017年4月27日にWO2017/070633として公開）およびUS内のまたは世界的なその同等物；（d）PCT/US2017/045381（2018年2月8日にWO2018/027078として公開）およびUS内のまたは世

40

50

界的なその同等物；(e) PCT/US2017/056671 (2018年4月19日にWO2018/071868として公開) およびUS内のまたは世界的なその同等物；PCT/2017/048390 (WO2017/048390、2017年3月23日) およびUS内のまたは世界的なその同等物；(f) PCT/US2017/068114 (未公開) およびUS内のまたは世界的なその同等物；(g) PCT/US2017/068105 (未公開) およびUS内のまたは世界的なその同等物；(h) PCT/US2017/046144 (WO2018/031683、2018年2月15日) およびUS内のまたは世界的なその同等物；(i) PCT/US2018/024208 (未公開) およびUS内のまたは世界的なその同等物；(j) PCT/2018/021878 (WO2018/021878、2018年2月1日) およびUS内のまたは世界的なその同等物；(k) Komor, A.C., Kim, Y.B., Packer, M.S., Zuris, J.A. & Liu, D.R. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533, 420- (2016)；(l) Gaudelli, N.M. et al. Programmable base editing of A.T to G.C in genomic DNA without DNA cleavage. *Nature* 551, 464- (2017)；(m) 当分野において公知の塩基編集因子を報告または記載している「参照」と題する本明細書に列記されている参照のいずれか。

10

【0307】

種々の側面では、本明細書に記載される進化したまたは改変された塩基編集因子は次の一般化された構造を有し：A-B-C、「A」はCas部分またはnapDNAbpであり、「B」は核酸エフェクタードメインであり(例えばデアミナーゼ、例えばシチジンまたはアデノシンデアミナーゼ)、「C」は任意の追加の塩基編集因子機能ドメインにあたる(例えば、UGIドメインまたはNLSドメイン)。加えて、「-」は、部分A、B、およびCを共有結合的につなぎ合わせるリンカーにあたる。リンカーはいずれかの好適な型(例えば、アミノ酸配列もしくは他のバイオポリマー、または部分が互いにバイオコンジュゲーションされるケースでは合成の化学的連結部)または長さであり得る。加えて、本発明の機能的な改善された塩基編集因子は、修正されるべき特異的な部位を標的化するための塩基編集因子のR/DNAによってプログラム可能な機能性を果たすための、1つ以上の「R」またはガイド配列(例えば、Cas9またはCas9同等物のケースではガイドRNA)をもまた包含し得る。

20

【0308】

特定の部分の要素の構成が機能的な塩基編集因子を産生する限り、部分の連結部の順序は、特に限定的であることを意味されない。つまり、本発明の進化した塩基編集因子は、次の構造によって表される編集因子をもまた包含し得る：B-A-C；B-C-A；C-B-A；C-A-B；およびA-C-B。種々の態様では、進化した塩基編集因子は、連続的進化プロセス(例えばPACE)によって進化させられた、進化した塩基編集因子の少なくとも1つのドメイン(例えば、Cas9ドメインまたはデアミナーゼドメイン)を含み得る。それゆえに、1つの態様では、本明細書は、参照Cas9ドメインに対して進化したCas9ドメインを含むが、塩基編集因子の他のドメインは進化させられていない、進化した塩基編集因子を提供する。別の態様では、本明細書は、進化したデアミナーゼドメイン(例えば、APOBEC1、AID、またはCDADドメイン)を含むが、塩基編集因子の他のドメインは進化させられていない、進化した塩基編集因子を提供する。また別の態様では、本明細書は、進化したUGIドメインを含むが、塩基編集因子の他のドメインは進化させられていない、進化した塩基編集因子を提供する。なお他の態様では、進化した塩基編集因子は、本明細書に記載される連続的進化プロセスによって進化させられているドメイン同士の組み合わせを含み得る。

30

40

【0309】

1つの態様では、進化した塩基編集因子は：(i) 核酸によってプログラム可能なDNA結合蛋白質(napDNAbp)と；(ii) シチジンデアミナーゼと；(iii) ウラシルグリコシラーゼ阻害因子ドメイン(UGI)とを含む融合蛋白質を含み得、(i)、(ii)、または(iii)の少なくとも1つは本明細書に記載される連続的進化プロセス(例えばPACE)を用いて進化させられている。種々の態様では、融合蛋白質は2、3、4、もしくは5、またはより多くのUGIドメインを含み得る。ある態様では、核酸によってプログラム可能なDNA結合蛋白質(napDNAbp)はCas9ドメインである。種々の態様では、Cas9ドメインはヌクレアーゼ活性型Cas9、ヌクレアーゼ不活性型Cas9(dCas9)、またはCas9ニックアーゼ(nCas9)であり得る。種々の他の態様では、napDNAbpはCasX、CasY、Cpf1、C2c1、C2c2、C2c3、または

50

Argonaute蛋白質である。

【0310】

(i)、(ii)、または(iii)の少なくとも1つが本明細書に記載される連続的進化プロセス(例えばPACE)によって進化させられた(has be evolved)、(i)核酸によってプログラム可能なDNA結合蛋白質(napDNAbp)と；(ii)シチジンデアミナーゼと；(iii)ウラシルグリコシラーゼ阻害因子ドメイン(UGI)とを含む進化した塩基編集因子は、種々の配設で構造的に構成され得、これらは：

NH₂-[シチジンデアミナーゼ]-[napDNAbp]-[UGI]-COOH；
 NH₂-[シチジンデアミナーゼ]-[UGI]-[napDNAbp]-COOH；
 NH₂-[napDNAbp]-[UGI]-[シチジンデアミナーゼ]-COOH；
 NH₂-[napDNAbp]-[シチジンデアミナーゼ]-[UGI]-COOH；
 NH₂-[UGI]-[シチジンデアミナーゼ]-[napDNAbp]-COOH；または
 NH₂-[UGI]-[napDNAbp]-[シチジンデアミナーゼ]-COOH、

を包含するが、これらに限定されず、「-」の各個は任意のリンカーを含む。

【0311】

他の態様では、(i)、(ii)、または(iii)の少なくとも1つが本明細書に記載される連続的進化プロセス(例えばPACE)によって進化させられた(has be evolved)、(i)核酸によってプログラム可能なDNA結合蛋白質(napDNAbp)と；(ii)シチジンデアミナーゼと；(iii)ウラシルグリコシラーゼ阻害因子ドメイン(UGI)とを含む進化した塩基編集因子は、種々の配設で構造的に構成され得、これらは：

NH₂-[シチジンデアミナーゼ]-[napDNAbp]-[UGI]-[UGI]-COOH；
 NH₂-[シチジンデアミナーゼ]-[UGI]-[napDNAbp]-[UGI]-COOH；
 NH₂-[UGI]-[シチジンデアミナーゼ]-[napDNAbp]-[UGI]-COOH；
 NH₂-[シチジンデアミナーゼ]-[UGI]-[napDNAbp]-[UGI]-COOH；
 NH₂-[シチジンデアミナーゼ]-[UGI]-[UGI]-[napDNAbp]-COOH；
 NH₂-[UGI]-[シチジンデアミナーゼ]-[UGI]-[napDNAbp]-COOH；
 NH₂-[UGI]-[napDNAbp]-[UGI]-[シチジンデアミナーゼ]-COOH；
 NH₂-[napDNAbp]-[UGI]-[UGI]-[シチジンデアミナーゼ]-COOH；
 NH₂-[napDNAbp]-[UGI]-[シチジンデアミナーゼ]-[UGI]-COOH；
 NH₂-[napDNAbp]-[シチジンデアミナーゼ]-[UGI]-[UGI]-COOH；
 NH₂-[napDNAbp]-[UGI]-[シチジンデアミナーゼ]-[UGI]-COOH；
 NH₂-[UGI]-[napDNAbp]-[シチジンデアミナーゼ]-[UGI]-COOH；
 NH₂-[UGI]-[シチジンデアミナーゼ]-[napDNAbp]-[UGI]-COOH；
 NH₂-[UGI]-[シチジンデアミナーゼ]-[UGI]-[napDNAbp]-COOH；
 NH₂-[UGI]-[UGI]-[シチジンデアミナーゼ]-[napDNAbp]-COOH；
 NH₂-[UGI]-[napDNAbp]-[シチジンデアミナーゼ]-[UGI]-COOH；
 NH₂-[UGI]-[napDNAbp]-[UGI]-[シチジンデアミナーゼ]-COOH；および
 NH₂-[UGI]-[UGI]-[napDNAbp]-[シチジンデアミナーゼ]-COOH、

を包含するが、これらに限定されず；「-」の各個は任意のリンカーを含む。

【0312】

いくつかの態様(embodiments)では、塩基編集因子は参照塩基編集因子を用いて進化させられ、これらは進化した塩基編集因子の対応する例と併せて下で要約されている。

【0313】

下の全ての配列では、テキストフォーマットは塩基編集因子構成要素のアイデンティティを次のように指示する：SV40 BPNLS-デアミナーゼ-リンカー-nCas9-リンカー-UGI-リンカー-UGI-SV40 BPNLS。

【0314】

次の塩基編集因子(配列番号15)は野生型rAPOBEC1を包含し、これは参照塩基編集因子として用いられ得る。この塩基編集因子が進化させられて、(配列番号16)として下に示されているevoAPOBEC塩基編集因子を生成した。

野生型APOBEC BE4Maxの全長アミノ酸配列

【化11-1】

MKRTADGSEFESP**KKR****KV**SSETGPVAVDPTLRRRIEPHEFEVFFDPRELK**ET**CLLYE
INWGGRHSIWRHTSQNTNKHVEVNFIEKFTT**ERY**FCPNTRCSITWFLSWSPCGECSRAIT**E**FLS
RYPHVTLFIYIARLYHHADPRNRQGLRDLISSGV**TI**QIMTE**Q**ESGYCWRNFVN**Y**SPSNEAHWPR
YPHLWVRLYVLELYCIILGLPPCLNILRR**K**QPQLTFF**T**I**A**LQ**S**CHYQRLPPHILWATGLK**S**GG**S**
 SGGSSG**SE**TPGT**SE**SAT**P**ESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYK**V**PSK**K**FKVLGN
 TDRHSIKK**N**LIGALLFDSGETAEATRLK**R**TARRRYTRRKNR**I**CYLQ**E**IFSNEMAKVDD**S**FFHRL
 EESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLR**K**KLV**D**STDKADLRLIYLALAHMIK**F**R
 GHFLIEGDLNPDNSDV**D**KLFIQLVQTY**N**QLFEENPINASGVDAK**A**ILSARLSKSRLENL**I**AQL
 PGEK**K**NGLFGNLIALSLGLTPNFKS**N**FDLAEDAKLQLSKD**T**YDDDLNLLAQIGDQYADLFL**A**A
 K**N**LS**D**AILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVR**Q**QLPEKYKEIFFDQ**S**KN
 GYAGYIDGGASQEEFYK**F**IKPILEKMDGTEELLV**K**LNREDLLR**K**QRTFDNGSIP**H**QIHLGELHA
 ILRRQEDFY**P**FLKDNREKIEKIL**T**FRIPY**V**GPLARGNSRF**A**WMTRKSEETITPWN**F**EEV**D**KG
 ASAQ**S**FIERMTNFDKNLPNEK**V**LPKHSLLYEYFTVYNELTK**V**KYVTEGMR**K**PAFLS**G**E**Q**KA**I**V
 DLL**F**KTN**R**KVT**V**KQLKEDY**F**KKIECFDSVEISGVEDRFNASLGTYHDLL**K**IKDKDFLDNEENE
 DILEDIVLTLTLFEDREMI**E**ERLKYAHLFDDK**V**M**Q**LKRRRYTG**W**GRLSRKLINGIRDK**Q**SG**K**
 TILD**F**LKSDGFANR**N**F**M**QLIHDDSLTFKEDI**Q**KAQ**V**SG**Q**DSLHEHIANLAGSPA**I**KK**G**IL**Q**TV
 K**V**DEL**V**K**M**GR**H**KPENIV**I**EMAREN**Q**TT**Q**KG**Q**NSRERM**K**RI**E**E**G**IKELGS**Q**ILKEHP**V**ENT**Q**
 L**Q**NEKLYLYL**Q**NGRDMYVD**Q**ELDINRLSDYD**V**DH**I**VP**Q**SFLKDDSIDNK**V**L**T**RS**D**KNR**G**KS**D**N
 VPSEEV**V**KK**M**KNYWR**Q**LLNAKLIT**Q**RKFDNL**T**KAERGGLSELDKAG**F**IK**R**QL**V**ET**R**QITKH**V**AQ

10

20

【化11-2】

ILDSRMNTKYDENDK**L**IRE**V**K**V**ITL**K**SKLVSDFR**K**DF**Q**FYK**V**REINNYHHAHDAYLNAV**V**GTAL
 IK**K**Y**P**KL**E**SE**F**VYGDYK**V**YDVR**K**MI**A**K**S**E**Q**E**I**G**K**ATA**K**Y**F**F**Y**SN**I**M**N**F**F**K**T**E**I**T**L**AN**G**E**I**R**K**RP
 LI**E**T**N**GET**G**E**I**V**D**K**G**R**D**FAT**V**R**K**VL**S**MP**Q**V**N**IV**K**K**T**E**V**Q**T**GG**F**SK**E**S**I**L**P**K**R**NS**D**K**L**I**A**R**K**D
 W**D**PK**K**Y**G**G**F**DS**P**TV**A**YS**V**L**V**V**A**K**V**E**K**G**K**SK**L**K**S**V**K**ELL**G**IT**I**M**E**R**S**S**F**E**K**NP**I**D**F**LE**A**K**G**Y**K**E
 V**K**K**D**L**I**IK**L**PK**Y**SL**F**E**L**ENGR**K**R**M**L**A**S**A**G**E**L**Q**K**G**N**E**L**A**L**P**S**K**Y**V**N**F**LYL**A**SH**Y**E**K**L**K**G**S**P**E**D**N**E
 Q**K**Q**L**F**V**E**Q**H**K**H**Y**L**D**E**I**E**Q**I**S**E**F**SK**R**V**I**L**A**D**A**N**L**D**K**V**L**S**A**Y**N**K**H**R**D**K**P**I**R**E**Q**A**E**N**I**I**H**L**F**T**L**T**N**
 L**G**A**P**A**A**F**K**Y**F**D**T**T**I**D**R**K**R**Y**T**S**T**K**E**V**L**D**A**T**L**I**H**Q**S**I**T**G**L**Y**E**T**R**I**D**L**S**Q**L**G**G**D**S**G**G**S**G**S**G**S**T**N**L**
SD**I**E**K**E**T**G**K**Q**L**V**I**Q**E**S**I**L**M**L**P**E**E**V**E**V**I**G**N**K**P**E**S**D**I**L**V**H**T**A**Y**D**E**S**T**D**E**N**V**M**L**L**T**S**D**A**P**E**Y**K**P**W
AL**V**I**Q**D**S**N**G**E**N**K**I**K**M**L**S**G**G**S**G**S**G**S**T**N**L****S**D**I**E**K**E**T**G**K**Q**L**V**I**Q**E**S**I**L**M**L**P**E**E**V**E**V**I**G**N**K**P**E**S**
DI**L**V**H**T**A**Y**D**E**S**T**D**E**N**V**M**L**L**T**S**D**A**P**E**Y**K**P**W**A**L**V**I**Q**D**S**N**G**E**N**K**I**K**M**L**S**G**G**S**K**R**T**A**D**G**S**E**F**E**P**K**K**R**
KV* (配列番号 15)

30

40

【0315】

次の塩基編集因子はevoAPOBECを包含する。これは上で提供されている塩基編集因子（配列番号15）に基づいて進化させられた。

evoAPOBEC BE4Maxの全長アミノ酸配列

【化 1 2 - 1】

MKRTADGSEFESPKKKRKVSSKTGPVAVDPTLRRRIEPHEFEVFFDPRELRKETCLLYE
INWGGRHSIWRHTSQNTNKHVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSRAITEFLS
RYPNVTLFIYIARLYHLANPRNRQGLRDLISSGVTIQIMTEQESGYCWHNFVNYSNESHWR
YPHLWVRLYVLELYCIILGLPPCLNILRRKQSOLTSFTIALQSCHYQRLPPHILWATGLKSGGS
 SGGSSGSETPGTSESATPESGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPKSKFKVLGN
 TDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRL
 EESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFR
 GHFLIEGDLNPDNSDVKLFIQLVQTYNQLEENPINASGVDAKAILSARLSKSRLENLIAQL
 PGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDXYDDDLNLLAQIGDQYADFLAA
 KNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKN
 GYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKNLREDLLRKQRTFDNGSIPHQIHLGELHA
 ILRRQEDFYFPLKDNREKIEKILTFRIPYVVGPLARGNSRFAMTRKSEETITPWNFEVVDKG
 ASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIV
 DLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIKDKDFLDNEENE
 DILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGK
 TILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILOTV

10

20

【化 1 2 - 2】

KVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQ
 LQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRKGSDN
 VPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQ
 ILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTA
 LKPKYKLESEFVYGDYKVDVRKMIKSEQEI GKATAKYFFYSNIMNFFKTEITLANGEIRKRP
 LIETNGETGEIVWDKGRDFATVRKVL SMPQVNI VKKTEVQTGGFSKESILPKRNSDKLIARKKD
 WDPKKYGGFDSPTVAYSVLVAKVEKKGSKKLSVKELLGITIMERSSEKPNIDFLEAKGYKE
 VKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLGSPEDNE
 QKQLFVEQHKHYLDEIEQISEFSKRVILADANLDKVL SAYNKHRDKPIREQAENI IHLFTLTN
 LGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSI TGLYETRIDLSQLGGDSGGSGGSGG**TNL**
SDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLTSDAPEYKPW
ALVIQDSNGENKIKMLSGGSGGSGGTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPES****
DILVHTAYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSKRTADGSEFEPKKR
KV* (配列番号 16)

30

40

【0 3 1 6】

次の塩基編集因子（配列番号17）は野生型pmCDA1を包含する。これは参照塩基編集因子として用いられ得る。この塩基編集因子が進化させられて、（配列番号18）として下に示されているevoCDA塩基編集因子を生成した。

wt-CDA-BE4Maxの全長アミノ酸配列

50

【化 1 3 - 1】

MKRTADGSEFESPKKKRKVSTDAEYVRIHEKLDIYTFKKQFFNNKKS^VSHRCYVLFELK
RRGERRACFWGYAVNKPQSGTERGIHAEIFSIRKVEEYLRDNPQGFTINWYSSWSPCADCAEKI
LEWYNQELRGNGHTLKIWACKLYYEKNARNQIGLWNLRDNGVGLNVMVSEHYQCCRKIFIQSSH
NQLNENRWLEKTLKRAEKRRSELSIMIQVKILHTTKSPAVSGGSSGGSSGSETPGTSESATPES
 SGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSSKFKVLGNTDRHSIKKNLIGALLFDSGE
 TAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHRLEESFLVEEDKKHERHPIFGN
 IVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLF
 IQLVQTYNQLFEEPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLT
 PNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEIT
 KAPLSASMIKRYDEHHQDLTLLKALVRQQLPKEYKEIFFDQSKNGYAGYIDGGASQEEFYKFIK
 PILEKMDGTEELLVKNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIE
 KILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEVVVDKGASAQSFIERMTNFDKNLPNE

10

【化 1 3 - 2】

KVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYF
 KKIECFDSVEISGVEDRFNASLGTYHDLKI IKDKDFLDNEENEDILEDIVLTLTLFEDREMIE
 ERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMLI
 HDDSLTFKEDIQKAQVSGQDLSLHEHIANLAGSPAIKKILQTVKVDELVKVMGRHKPENIVI
 EMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVD
 QELDINRLSDYDVDHIVPQSFLKDDSIDNKVLRSDKNRGSNDNVPSEEVVKMKNYWRQLLNA
 KLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREV
 KVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGITALIKKYPKLESEFVYGDYKVD
 VRKMIKSEQEI GKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFA
 TVRKVLSMPQVNIKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKGFDSPVAVSVLV
 VAKVEKGKSKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDLIIKLPKYSLFELENG
 RKRMLASAGELQKGNELALPSKYVNFYLYLASHYEKLGKSPEDNEQKQLFVEQHKHYLDEIIEQI
 SEFSKRVLADANLDKVL SAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYT
 STKEVL DATLIHQSI TGLYETRIDLSQLGGDGGSGGSGG**TNLSDIIEKETGKQLVIQESILM**
LPEEVEEVIGNKPESDILVHTAYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGS
GGSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLT
SDAPEYKPWALVIQDSNGENKIKMLSGGSKRTADGSEFEPKKKRKV* (配列番号 17)

20

30

40

【0 3 1 7】

次の塩基編集因子はevoCDAを包含する。これは上で提供されている塩基編集因子（配列番号17）に基づいて進化させられた。
 evoCDA-BE4Maxの全長アミノ酸配列

【化14-1】

MKRTADGSEFESPKKKRKVSTDAEYVRIHEKLDIYTFKKQFSNNKKS^VSHRCYVLFELK
RRGERRACFWGYAVNKPQSGTERGIHAEIFSRKVEEYLRDNPQGFTINWYSSWSPCADCAEKI
LEWYNQELRGNGHTLKIWVCKLYYEKNARNQIGLWNLRDNGVGLNVMVSEHYQCCRKIFIQSSH
NQLNENRWLEKTLKRAEKRRSELSIMFQVKILHTTKSPAVSGGSSGGSSGSETPGTSESATPES
 SGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPKFKVVLGNTDRHSIKKNLIGALLFDSGE
 TAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGN
 IVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLF
 IQLVQTYNQLEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIASLGLT
 PNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEIT
 KAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIK

10

【化14-2】

PILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFLKDNREKIE
 KILTRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEVVVDKGASAQSFIERMTNFDKNLPNE
 KVLPKHSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYF
 KKIECFDSVEISGVEDRFNASLGTYHDLKI IKDKDFLDNEENEDILEDIVLTLTLFEDREMIE
 ERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDFANRNFMQLI
 HDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVI
 EMARENQTTQKGQKNSRERMKRIE EGIKELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVD
 QELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNA
 KLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREV
 KVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTA LIKKYPKLESEFVYGDYKVYD
 VRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFA
 TVRKVLSMPQVNIKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLV
 VAKVEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDLIIKLPKYSLFELENG
 RKRMLASAGELQKGNELALPSKYVNFYLYLASHYEKLGSPEDNEQKQLFVEQHKKHYLDEIEEQI
 SEFSKRVLADANLDKVL SAYNKHRDKPIREQAENI IHLFTLTNLGAPAAFKYFDTTIDRKRYT
 STKEVLDATLIHQSI TGLYETRIDLSQLGGDSGGSSGGSS**TNLSDIIEKETGKQLV IQESILM**
LPEEVEEVIGNKPESDILVHTAYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGS
GGSSGNLSDIIEKETGKQLV IQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLT
SDAPEYKPWALVIQDSNGENKIKMLSGGSKRTADGSEFEPKKKRKV* (配列番号 18)

20

30

40

【0318】

次の塩基編集因子（配列番号19）は野生型FERNYを包含する。これは参照塩基編集因子として用いられ得る。この塩基編集因子が進化させられて、（配列番号20）として下で示されているevoFERNY塩基編集因子を生成した。

FERNY-BE4Maxの全長アミノ酸配列

【化15-1】

MKRTADGSEFESPKKKRKVSFERNYDPRELRKETYLLYEIKWGKSGKLWRHWCQNNRTQ
HAEVYFLENIFNARRFNPSTHCSITWYLSWSPCAECSQKIVDFLKEHPNVNLEIYVARLYYHED
ERNRQGLRDLVNSGVTIRIMDLDPDYNWCWTFVSDQGGDEYWPGHFAPWIKQYSLKLSGGSSG
 GSSGSETPGTSESATPESGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVP SKKFKVLGNTD
 RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLOEIFSNEMAKVDDSSFFHRL
 SFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLRLIYLALAHMIKFRGH
 FLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLLENLIAQLPG

10

【化15-2】

EKKNLFGNLIASLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKN
 LSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGY
 AGYIDGGASQEEFYKFIKPILEKMDGTEELLVKNLREDLLRKQRTFDNGSIPHQIHLGELHAIL
 RRQEDFYFPLKDNREKIEKILTFRIPYVVGPLARGNSRFAMTRKSEETITPWNFEVVDKGAS
 AQSFIERMTNFDKNLPNEKVLPHKSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDL
 LFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLKIIKDKDFLDNEENEDI
 LEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTI
 LDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKILQTVKV
 VDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIE EGIKELGSQILKEHPVENTQLQ
 NEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVP
 SEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELKAGFIKRQLVETRQITKHVAQIL
 DSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVVREINNYHHAHDAYLNAVVG TALIK
 KYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLI
 ETNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWD
 PKKYGGFDSPTVAYSVLVVAKEGKSKKLKSVKELLGITIMERSSEKPNIDFLEAKGYKEVK
 KDLI IKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLGSPEDNEQK
 QLFVEQHKKHYLDEIEEQISEFSKRVI LADANLDKVL SAYNKHRDKPIREQAENI IHLFTLTNLG
 APAAFKYFDTTIDRKRYTSTKEVLDATLIHQSI TGLYETRIDLSQLGGDSGGSGGSGG**TNLS**
IIKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLTSDAPEYKPWAL
VIQDSNGENKIKMLSGGSGGSGGTNLS**DIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDI**
LVHTAYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSKRTADGSEFEPKKKRKV

20

30

40

* (配列番号 19)

【0319】

次の塩基編集因子はevoFERNYを包含する。これは上で提供されている塩基編集因子（配列番号19）に基づいて進化させられた。

evoFERNY-BE4Maxの全長アミノ酸配列

【化 1 6 - 1】

MKRTADGSEFESPKKRKVSFERNYDPRELKETYLLYEIKWKGSGKLRHWCQNNRTQ
HAEVYFLENIFNARRFNPSTHCSITWYLSWSPCAECSQKIVDFLKEHPNVNLEIYVARLYYPEN
ERNRQGLRDLVNSGVTIRIMDLDPDYNWCWTFVSDQGGDEYWPGHFAPWIKQYSLKLSGGSSG
GSSGSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSSKFKVLGNTD
RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLLEE

【化 1 6 - 2】

10

SFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGH
FLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPG
EKKNGLFGNLIASLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADFLAAKN
LSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGY
AGYIDGGASQEEFYKFIKPILEKMDGTEELLVKNLREDLLRKQRTFDNGSIPHQIHLGELHAIL
RRQEDFYFPLKDNREKIEKILTFRIPYVVGPLARGNSRFAMTRKSEETITPWNFEEVVDKGAS
AQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDL
LFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLKIIKDKDFLDNEENEDI
LEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTI
LDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKV
VDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQ
NEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVP
SEEVVKMKKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQIL
DSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKPREINNYHHAHDAYLNAVVGITALIK
KYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLI
ETNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTVEVQTGGFSKESILPKRNSDKLIARKDWD
PKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSEKPNIDFLEAKGYKEVK
KDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLGKSPEDNEQK
QLFVEQHKHYLDEIEEQISEFSKRVI LADANLDKVL SAYNKHRDKPIREQAENI IHLFTLTNLG
APAAFKYFDTTIDRKRYTSTKEVL DATLIHQSI TGLYETRIDLSQLGGDSGGSGGSGG**TNLS**
IIKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLTSDAPEYKPWAL
VIQDSNGENKIKMLSGGSGGSGGTNLS**DIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDI**
LVHTAYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSKRTADGSEFEPKKRKV

20

30

40

* (配列番号 20)

【0320】

いくつかの態様では、本明細書において提供される塩基編集因子蛋白質のいずれかは、例えばイノシン塩基除去修復の阻害因子（例えば、ウラシルグリコシラーゼ阻害因子（UGI）ドメインまたは触媒的に不活性なイノシン特異的ヌクレアーゼ（dISN））などの1つ以上の追加の核酸エフェクター部分をさらに包含し得る。いずれかの特定の理論によって拘束されることを欲することなしに、UGIドメインまたはdISNは脱アミノ化されたアデノシン残基（例えばイノシン）の塩基除去修復を阻害または防止し得、これは塩基編集因子

50

の活性または効率を改善し得る。追加の塩基編集因子機能がさらに本明細書に記載されている。

【0321】

(A) Cas9ドメイン

本明細書によって提供される進化した塩基編集因子は、いずれかの好適なCas9部分または同等な蛋白質、例えばCRISPR関連蛋白質9またはその機能的断片を包含し、いずれかの生物からのいずれかの天然に存在するCas9、いずれかの天然に存在するCas9同等物またはその機能的断片、いずれかの生物からのいずれかのCas9ホモログ、オソログ、またはパラログ、および天然に存在するかまたは操作されているCas9のいずれかの変異体またはバリエーションを包摂する。これらのCas9部分 (moieties) または同等な蛋白質は本明細書に記載される連続的進化方法 (例えばPACE) を用いて進化させられ得る。進化した塩基編集因子は、Cas9部分のみがPACEを用いて進化させられているもの、またはCas9部分が1つ以上の他の塩基編集因子ドメイン (例えばデアミナーゼ) と併せて進化させられているものを包含する。本明細書に記載される進化した塩基編集因子は、Cas9部分またはドメインがPACEを用いて進化させられていないが、1つ以上の他の塩基編集因子ドメイン (例えばデアミナーゼドメイン) がPACEを用いて進化させられている、融合蛋白質をもまた包含し得る。

10

【0322】

より広くは、Cas9、ある型の「RNAによってプログラム可能なヌクレアーゼ」または「RNAによってガイドされるヌクレアーゼ」または「核酸によってプログラム可能なDNA結合蛋白質」である。用語napR/DNAbpまたはCas9は特に限定的であることを意味されない。本開示は、本発明の進化した塩基編集因子に採用される特定のnapR/DNAbp、Cas9、またはCas9同等物については限定がない。

20

【0323】

本開示のコンテキストから理解されるであろう通り、いずれかのCas9ドメインは、一般的に、本明細書に記載される連続的進化方法 (例えばPACE) を用いる処理のための可能な参照ポリペプチド (すなわち出発点) と見なされるべきである。さもなければ、本明細書に記載される連続的進化方法を用いて進化させられたCas9ドメインはかかるものとして指示される。

【0324】

いくつかの態様では、napR/DNAbpはCas部分である。

30

【0325】

種々の態様では、Cas部分は*S. pyogenes* Cas9であり、これはゲノム工学のツールとして概ね幅広く用いられている。このCas9蛋白質は2つの別個のヌクレアーゼドメインを含有する大きいマルチドメイン蛋白質である。点変異がヌクレアーゼ活性を廃絶するようにCas9に導入されて、デッドCas9 (dCas9) をもたらし得る。これはsgRNAによってプログラムされた様式でDNAに結合するその能力をなお持ち続ける。原理的には、別の蛋白質またはドメインに融合させられたときには、単純に適切なsgRNAとの共発現によって、dCas9はその蛋白質を事実上いずれかのDNA配列へと標的化し得る。

40

【0326】

他の態様では、Cas部分は：*Corynebacterium ulcerans* (NCBI Ref : NC_015683.1, NC_017317.1) ; *Corynebacterium diphtheria* (NCBI Ref : NC_016782.1, NC_016786.1) ; *Spiroplasma syrrhodicola* (NCBI Ref : NC_021284.1) ; *Prevotella intermedia* (NCBI Ref : NC_017861.1) ; *Spiroplasma taiwanense* (NCBI Ref : NC_021846.1) ; *Streptococcus iniae* (NCBI Ref : NC_021314.1) ; *Belliella baltica* (NCBI Ref : NC_018010.1) ; *Psychroflexus torquis* (NCBI Ref : NC_018721.1) ; *Streptococcus thermophilus* (NCBI Ref : YP_820832.1) ; *Listeria innocua* (NCBI Ref : NP_472073.1) ; *Campylobacter jejuni* (NCBI Ref : YP_002344900.1) ; または *Neisseria meningitidis* (NCBI Ref : YP_002342100.1) からのCas9である。

【0327】

50

なお他の態様では、Cas部分はいずれかのCRISPR関連蛋白質を包含し得、Cas1、Cas1B、Cas2、Cas3、Cas4、Cas5、Cas6、Cas7、Cas8、Cas9 (Csn1およびCsx12としてもまた公知)、Cas10、Csy1、Csy2、Csy3、Cse1、Cse2、Csc1、Csc2、Csa5、Csn2、Csm2、Csm3、Csm4、Csm5、Csm6、Cmr1、Cmr3、Cmr4、Cmr5、Cmr6、Csb1、Csb2、Csb3、Csx17、Csx14、Csx10、Csx16、CsaX、Csx3、Csx1、Csx15、Csf1、Csf2、Csf3、Csf4、それらのホモログ、またはそれらの改変されたバージョンを包含するが、これらに限定されない。これらの酵素は公知である；例えば、*S. pyogenes* Cas9蛋白質のアミノ酸配列はSwissProtデータベースにおいてアクセッション番号Q99ZW2で見出され得る。いくつかの態様では、Cas9などの未改変のCRISPR酵素はDNA切断活性を有する。いくつかの態様では、CRISPR酵素はCas9であり、*S. pyogenes*または*S. pneumoniae*からのCas9であり得る。いくつかの態様では、CRISPR酵素は、例えば標的配列内においておよび/または標的配列の相補物内において、標的配列の場所の1つのまたは両方の鎖の切断を導く。いくつかの態様では、CRISPR酵素は、標的配列の第1のまたは最後のヌクレオチドから約1、2、3、4、5、6、7、8、9、10、15、20、25、50、100、200、500、またはより多くの塩基対以内において、1つのまたは両方の鎖の切断を導く。いくつかの態様では、ベクターは、変異したCRISPR酵素が、標的配列を含有する標的ポリヌクレオチドの1つのまたは両方の鎖を切断する能力を欠くように、対応する野生型酵素に対して変異しているCRISPR酵素をコードする。例えば、*S. pyogenes*からのCas9のRuvC1触媒ドメインのアスパラギン酸アラニン置換(D10A)は、Cas9を両方の鎖を切断するヌクレアーゼからニッカーゼ(一本の鎖を切断する)に変換する。Cas9をニッカーゼにする変異の他の例は、限定なしに、H840A、N854A、およびN863Aを包含する。

10

20

【0328】

Cas部分はcasn1ヌクレアーゼまたはCRISPR (clustered regularly interspaced short palindromic repeat) 関連ヌクレアーゼともまた言われ得る。上で概観されている通り、CRISPRは可動遺伝子エレメント(ウイルス、転移エレメント、および接合プラスミド)からの保護を提供する適応免疫システムである。CRISPRクラスターは、スペーサー、祖先由来の可動性エレメントに対して相補的な配列を含有し、侵入核酸を標的化する。CRISPRクラスターは転写され、CRISPR RNA (crRNA) へとプロセシングされる。II型CRISPRシステムでは、pre-crRNAの正しいプロセシングはtrans-encoded small RNA (tracrRNA)、内在性リボヌクレアーゼ3 (rnc)、およびCas9蛋白質を要求する。tracrRNAはpre-crRNAのリボヌクレアーゼ3によって支援されるプロセシングのためのガイドとしての用をなす。爾後に、Cas9/crRNA/tracrRNAはスペーサーに対して相補的な線形または環状dsDNA標的をエンド的に切断する。crRNAに対して相補的でない標的鎖が第1にエンド的に切られ、それから3'-5'エキソ的にトリミングされる。天然では、DNA結合および切断は典型的には蛋白質および両方のRNAを要求する。しかしながら、crRNAおよびtracrRNA両方の側面を1個のRNA種に組み込むように、シングルガイドRNA(「sgRNA」または単純に「gRNA」)が操作され得る。例えば、Jinek M., Chylinski K., Fonfara I., Hauer M., Doudna J.A., Charpentier E. *Science* 337: 816-821 (2012)を参照する。これらの内容全体は参照によってここに組み込まれる。

30

【0329】

Cas9および同等物は、CRISPRリピート配列中の短いモチーフ(PAMまたはプロトスペーサー隣接モチーフ)を認識して、自己対非自己を区別することを助ける。本明細書に注記されている通り、Cas9ヌクレアーゼ配列および構造は当業者に周知である(例えば、"Complete genome sequence of an M1 strain of *Streptococcus pyogenes*." Ferretti et al., J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K., Primeaux C, Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P., Qian Y., Jia H.G., Najjar F. Z., Ren Q., Zhu H., Song L., White J., Yuan X., Clifton S.W., Roe B.A., McLaughlin R.E., *Proc. Natl. Acad. Sci. U.S.A.* 98: 4658-4663 (2001); "CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III." Deltcheva E., Chylinski K., Sharma CM., Gonzales K., Chao Y., Pirzada Z.A., Eckert M.R., Vogel J., Ch

40

50

arpentier E., Nature 471: 602-607 (2011); および "A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity." Jinek M., Chylinski K., Fonfara I., Hauer M., Doudna J. A., Charpentier E. Science 337: 816-821 (2012)を参照する。これらのそれぞれの内容全体は参照によって本明細書に組み込まれる。

【0330】

Cas部分はいずれかの好適なホモログおよび/またはオーソログを包含し得る。Cas9ホモログおよび/またはオーソログは種々の種において記載されており、*S. pyogenes*および*S. thermophilus*を包含するが、これらに限定されない。追加の好適なCas9ヌクレアーゼおよび配列は本開示に基づいて当業者には明らかであろう。かかるCas9ヌクレアーゼおよび配列は、Chylinski, Rhun, and Charpentier, "The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems" (2013) RNA Biology 10: 5, 726-737に開示されている生物および座位からのCas9配列を包含する; これの内容全体は参照によって本明細書に組み込まれる。いくつかの態様では、Cas9ヌクレアーゼは不活性化な(例えば、不活性化された)DNA切断ドメインを有し、つまり、Cas9はニッカーゼである。

10

【0331】

種々の態様では、進化した塩基編集因子はヌクレアーゼ不活性化型Cas蛋白質を含み得、交換可能に、(ヌクレアーゼ「デッド」Cas9として)「dCas」または「dCas9」蛋白質と言われ得る。不活性化DNA切断ドメインを有するCas9蛋白質(またはその断片)を生成するための方法は公知である(例えば、Jinek et al., Science. 337: 816-821 (2012); Qi et al., "Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression" (2013) Cell. 28; 152(5): 1173-83を参照する。これらのそれぞれの内容全体は参照によって本明細書に組み込まれる)。例えば、Cas9のDNA切断ドメインは2つのサブドメイン、HNHヌクレアーゼサブドメインおよびRuvC1サブドメインを包含することが公知である。HNHサブドメインはgRNAに対して相補的な鎖を切断するが、RuvC1サブドメインは非相補鎖を切断する。これらのサブドメインの変異はCas9のヌクレアーゼ活性をサイレンシングし得る。例えば、変異D10AおよびH840Aは*S. pyogenes* Cas9のヌクレアーゼ活性を完全に不活性化する(Jinek et al., Science. 337: 816-821 (2012); Qi et al., Cell. 28; 152(5): 1173-83 (2013))。いくつかの態様では、Cas9の断片を含む蛋白質が提供される。例えば、いくつかの態様では、蛋白質は2つのCas9ドメインの1つを含む:(1) Cas9のgRNA結合ドメイン; または(2) Cas9のDNA切断ドメイン。

20

30

【0332】

いくつかの態様では、Cas9またはその断片を含む蛋白質は「Cas9バリエーション」と言われる。Cas9バリエーションはCas9またはその断片に対する相同性を共有する。例えば、Cas9バリエーションは、野生型Cas9と少なくとも約70%同一、少なくとも約80%同一、少なくとも約90%同一、少なくとも約95%同一、少なくとも約96%同一、少なくとも約97%同一、少なくとも約98%同一、少なくとも約99%同一、少なくとも約99.5%同一、または少なくとも約99.9%同一である。いくつかの態様では、Cas9バリエーションは、野生型Cas9と比較して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、またはより多くのアミノ酸変化を有し得る。いくつかの態様では、Cas9バリエーションは、断片が、野生型Cas9の対応する断片と少なくとも約70%同一、少なくとも約80%同一、少なくとも約90%同一、少なくとも約95%同一、少なくとも約96%同一、少なくとも約97%同一、少なくとも約98%同一、少なくとも約99%同一、少なくとも約99.5%同一、または少なくとも約99.9%同一であるような、Cas9の断片(例えば、gRNA結合ドメインまたはDNA切断ドメイン)を含む。いくつかの態様では、断片は、対応する野生型Cas9のアミノ酸長さの少なくとも30%、少なくとも35%、少なくとも40%、少なくとも45%、少なくとも50%、少なくとも55%、少なくとも60%、少なくとも65%、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも95%同一、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、または少なくとも99.5%である。

40

【0333】

50

いくつかの態様では、Cas9断片は長さが少なくとも100アミノ酸である。いくつかの態様では、断片は長さが少なくとも100、150、200、250、300、350、400、450、500、550、600、650、700、750、800、850、900、950、1000、1050、1100、1150、1200、1250、または少なくとも1300アミノ酸である。いくつかの態様では、野生型Cas9はStreptococcus pyogenesからのCas9に対応する（NCBI参照配列：NC_017053.1）。他の態様では、野生型Cas9はStreptococcus pyogenesからのCas9に対応する（NCBI参照配列：NC_002737.2）。なお他の態様では、dCas9は、部分的にまたは完全に、Cas9ヌクレアーゼ活性を不活性化する1つ以上の変異を有するCas9アミノ酸配列に対応するか、またはそれを含む。

【0334】

いくつかの態様では、Cas9ドメインは、D10A変異、一方で、Streptococcus pyogenesからのCas9などの野生型配列と比べて位置840の残基を含む（NCBI参照配列：NC_017053.1）。

【0335】

いずれかの特定の理論によって拘束されることを欲することなしに、触媒残基H840の存在は、標的化されるCと反対のGを含有する非編集（例えば、非脱アミノ化）鎖を切断するCas9の活性を回復させる。H840の（例えばA840からの）回復はCを含有する標的鎖の切断をもたらさない。かかるCas9パリアントは、gRNAによって定義される標的配列に基づいて特異的な場所に一本鎖DNA切断（ニック）を生成することができ、非編集鎖の修復に至り、最終的には非編集鎖のG A変化をもたらす。簡潔には、C-G塩基対のCがデアミナーゼ、例えばAPOBECデアミナーゼによってUへと脱アミノ化され得る。Gを有する非編集鎖をニックキングすることはミスマッチ修復メカニズムによるGの取り除きを助長する。UGIはUDGを阻害し、これはUの取り除きを防止する。

【0336】

他の態様では、D10AおよびH840A以外の変異を有するdCas9パリアントが提供され、これらは例えばヌクレアーゼ不活性化型Cas9（dCas9）をもたらす。かかる変異は、例として、Streptococcus pyogenesからのCas9（NCBI参照配列：NC_017053.1）などの野生型配列の参照によって、D10およびH820における他のアミノ酸置換またはCas9のヌクレアーゼドメイン内の他の置換（例えば、HNHヌクレアーゼサブドメインおよび/またはRuvC1サブドメイン内の置換）を包含する。いくつかの態様では、dCas9のパリアントまたはホモログ（例えば、Streptococcus pyogenesからのCas9（NCBI参照配列：NC_017053.1）のパリアント）が提供され、これらはNCBI参照配列：NC_017053.1と少なくとも約70%同一、少なくとも約80%同一、少なくとも約90%同一、少なくとも約95%同一、少なくとも約98%同一、少なくとも約99%同一、少なくとも約99.5%同一、または少なくとも約99.9%同一である。いくつかの態様では、dCas9のパリアント（例えば、NCBI参照配列：NC_017053.1のパリアント）が提供され、NC_017053.1よりも約5アミノ酸、約10アミノ酸、約15アミノ酸、約20アミノ酸、約25アミノ酸、約30アミノ酸、約40アミノ酸、約50アミノ酸、約75アミノ酸、約100アミノ酸、またはより多くだけ短いまたは長いアミノ酸配列を有する。

【0337】

いくつかの態様では、本明細書において提供される進化した塩基編集因子は、Cas9蛋白質の全長アミノ酸配列、例えば本明細書において提供されるCas9配列の1つを含む。しかしながら、他の態様では、本明細書において提供される融合蛋白質は全長Cas9配列を含まず、その断片のみを含む。例えば、いくつかの態様では、本明細書において提供されるCas9融合蛋白質は、断片がcrRNAおよびtracrRNAまたはsgRNAに結合するCas9断片を含むが、例えばそれがヌクレアーゼドメインの短縮バージョンのみを含むかまたはヌクレアーゼドメインを少しも含まないということで、機能的なヌクレアーゼドメインを含まない。好適なCas9ドメインおよびCas9断片の例示的なアミノ酸配列は本明細書において提供されており、Cas9ドメインおよび断片の追加の好適な配列は当業者には明らかであろう。

【0338】

それらのパリアントおよびホモログを包含する追加のCas9蛋白質（例えば、ヌクレアーゼデッドCas9（dCas9）、Cas9ニッカーゼ（nCAs9）、またはヌクレアーゼ活性型Cas9）は

10

20

30

40

50

本開示の範囲内であるということは了解されるはずである。例示的なCas9蛋白質は、限定なしに、下で提供されるものを包含する。いくつかの態様では、Cas9蛋白質はヌクレアーゼデッドCas9 (dCas9) である。いくつかの態様では、dCas9はアミノ酸配列 (配列番号32) を含む。いくつかの態様では、Cas9蛋白質はCas9ニッカーゼ (nCas9) である。

【 0 3 3 9 】

ある態様では、本発明の進化した塩基編集因子は、次の参照配列を有する触媒的に不活性なCas9 (dCas9) :

【 0 3 4 0 】

【 化 1 7 - 1 】

DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGAL 10
 LFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHRLEESFLVEEDK
 KHERHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLRLIYLALAHMIKFRGHFLIE
 GDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPG
 EKKNGLFGNLIASLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLF
 LAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIF
 FDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIP
 HQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEET 20
 ITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPHKSHLLYEYFTVYNELTKVKYVT
 EGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNAS
 LGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQ
 LKRRRYTGWRLSRKLINGIRDKQSGKTILDFLKSDFANRNFMQLIHDDSLTFKEDIQK
 AQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIAMARENQT
 TQKGQNSRERMKRIIEGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQE
 LDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKMKNYWRQ 30
 LLNAKLITQRKFDNLTKAERGGLSELKAGFIKRQLVETRQITKHVAQILDSRMNTKYD
 ENDKLIREVKVITLKSCLVSDFRKDFQFYKREINNYHHAHDAYLNAVVG TALIKKYPK
 LESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIE
 TNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARK
 KDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLSVKELLGITIMERSSEFEKNPIDFL
 EAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASH
 YEKLGKSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVL SAYNKHRDKP 40
 IREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQ SITGLYETRIDL
 S
 QLGGD (配列番号 32)

、または本明細書に記載される連続的進化プロセス (例えばPACE) を用いて進化させられたその進化したバリエーションを包含し得る。

【 0 3 4 1 】

他の態様では、進化した塩基編集因子は :

【 0 3 4 2 】

10

20

30

40

50

【化 1 8】

DKKYSIGLAIGTNSVGWAVITDEYKVPSSKFKVLGNTDRHSIKKNLIGAL
 LFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDDSFFHRLEESFLVEEDK
 KHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIE
 GDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPG
 EKKNGLFGNLIASLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLF
 LAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIF
 FDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKNLREDLLRKQRTFDNGSIP
 HQIHLGELHAILRRQEDFYKFLKDNREKIEKILTRIPYYVGPLARGNSRFAWMTRKSEET
 ITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPHSLLYEYFTVYNELTKVKYVT
 EGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNAS
 LGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQ
 LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMLIHDDSLTFKEDIQK
 AQVSGQGDSLHEHIANLAGSPAIKKGIQTVKVVDELVKVMGRHHPENIVIAMARENQT
 TQKGQKNSRERMKRIIEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQE
 LDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQ
 LLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYD
 ENDKLIREVKVITLKSCLVSDFRKDFQFYKVINNYHHAHDAYLNAVVGTAALIKKYPK
 LESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIE
 TNGETGEIVWDKGRDFATVRKVLSPQVNVIVKKTEVQTGGFSKESILPKRNSDKLIARK
 KDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLSVKELLGITIMERSSEKPNIDFL
 EAKGYKEVKKDLIILPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASH
 YEKLKGSPEDEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVL SAYNKHRDKP
 IREQAENIIHLFTLTNLGAPAAFYFDTTIDRKRYTSTKEVLDATLIHQSI TGLYETRIDL S
 QLGGD (配列番号 9)

10

20

30

のアミノ酸配列と少なくとも80%、85%、90%、95%、98%、または99%同一であるアミノ酸配列を含むCas9ニッカーゼ (nCas9) を含み得、その進化したバージョンであり得る。

【0343】

なお他の態様では、進化した塩基編集因子は：

【0344】

40

【化 1 9】

DKKYSIGLDIGTNSVGVAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGAL
 LFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFHRLSEESFLVEEDK
 KHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIE
 GDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPG
 EKKNGLFGNLIASLGLTPNFKSNDLAEDAQLSKDQYDDDLNLLAQIGDQYADLF
 LAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIF
 FDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIP 10
 HQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYVVGPLARGNSRFAWMTRKSEET
 ITPWNFEVVDKGASQAQSFIERMTNFDKNLPNEKVLPHSLLYEYFTVYNELTKVKYVT
 EGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNAS
 LGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQ
 LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDFANRNFQMQLIHDDSLTFKEDIQK
 AQVSGQGDSLHEHIANLAGSPAIKKILQTVKVVDELVKVMGRHKPENIVIAMARENQT
 TQKGQKNSRERMKRIEIEGKELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQE 20
 LDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNPSEEVVKKMKNYWRQ
 LLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYD
 ENDKLIREVKVITLKSCLVSDFRKDFQFYKREINNYHHAHDAYLNAVVGTAALIKKYPK
 LESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIE
 TNGETGEIVWDKGRDFATVRKVLSPQVNVKKTQEVQTGGFSKESILPKRNSDKLIARK
 KDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSEFEKNPIDFL
 EAKGYKEVKKDLIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFYLYLASH
 YEKLGKSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKHRDKP 30
 IREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDL
 QLGGD (配列番号 33).

のアミノ酸配列と少なくとも80%、85%、90%、95%、98%、または99%同一であるアミノ酸配列を含む触媒的に活性なCas9を含み得る。

【 0 3 4 5】

いくつかの態様では、Cas部分は、単細胞原核生物微生物のあるドメインおよび界を構成する古細菌（例えば、ナノ古細菌）からのCas9またはCas9ホモログを言う。いくつかの態様では、Cas9はCasXまたはCasYを言い、これらは例えばBurststein et al., "New CRISPR-Cas systems from uncultivated microbes." *Cell Res.* 2017 Feb 21. doi: 10.1038/cr 40
 .2017.21に記載されており、この内容全体は参照によってここに組み込まれる。ゲノム解像度のメタゲノミクスを用いて、生命の古細菌ドメインにおける第1に報告されたCas9を包含するいくつかのCRISPR-Casシステムが同定された。この分岐したCas9蛋白質は、ほとんど研究されていないナノ古細菌において、活性なCRISPR-Casシステムの一部として見出された。細菌では、2つの以前には未知のシステム、CRISPR-CasXおよびCRISPR-CasYが発見された。これらはこれまでに発見された最もコンパクトなシステムの一つである。いくつかの態様では、Cas9はCasXまたはCasXのパリアントを言う。いくつかの態様では、Cas9はCasYまたはCasYのパリアントを言う。他のRNAによってガイドされるDNA結合蛋白質が、核酸によってプログラム可能なDNA結合蛋白質（napDNAbp）として用いられ得、本開示の範囲内であるということは了解されるはずである。 50

【0346】

いくつかの態様では、Cas9部分は、本明細書において提供される融合蛋白質のいずれかの核酸によってプログラム可能なDNA結合蛋白質 (napDNAbp) であり、CasXまたはCasY蛋白質であり得る。いくつかの態様では、napDNAbpはCasX蛋白質である。いくつかの態様では、napDNAbpはCasY蛋白質である。いくつかの態様では、napDNAbpは、天然に存在するCasXまたはCasY蛋白質と少なくとも85%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、または少なくとも99.5%同一であるアミノ酸配列を含む。いくつかの態様では、napDNAbpは天然に存在するCasXまたはCasY蛋白質である。いくつかの態様では、napDNAbpは、野生型Cas部分または本明細書において提供されるいずれかのCas部分と少なくとも85%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、または少なくとも99.5%同一であるアミノ酸配列を含む。

10

【0347】

種々の態様では、核酸によってプログラム可能なDNA結合蛋白質は、限定なしに、Cas9 (例えば、dCas9およびnCas9)、CasX、CasY、Cpf1、C2c1、C2c2、C2c3、およびArgonaut eを包含する。Cas9とは異なるPAM特異性を有する核酸によってプログラム可能なDNA結合蛋白質の1つの例はClustered Regularly Interspaced Short Palindromic Repeats from *Prevotella* and *Francisella* 1 (Cpf1) である。Cas9に類似に、Cpf1もまたクラス2 CRISPRエフェクターである。Cpf1はロバストなDNA干渉を媒介し、Cas9とは別々の特質を有するということを示されている。Cpf1は、tracrRNAを欠く1個のRNAによってガイドされるエンドヌクレアーゼであり、それはトリッチなプロトスペーサー隣接モチーフ (TTN、TTTN、またはYTN) を利用する。さらにその上、Cpf1は互い違いのDNA二本鎖切断によってDNAを切断する。16個のCpf1ファミリー蛋白質のうち、*Acidaminococcus* および *Lachnospiraceae* からの2つの酵素はヒト細胞において効率的なゲノム編集活性を有することが示される。Cpf1蛋白質は当分野において公知であり、以前に記載されている。例えばYamano et al., "Crystal structure of Cpf1 in complex with guide RNA and target DNA." *Cell* (165) 2016, p. 949-962; これの内容全体は参照によってここに組み込まれる。

20

【0348】

ガイドヌクレオチド配列によってプログラム可能なDNA結合蛋白質ドメインとして用いられ得るヌクレアーゼ不活性型Cpf1 (dCpf1) バリエーションもまた本組成物および方法に有用である。Cpf1蛋白質は、Cas9のRuvCドメインに類似であるRuvC様エンドヌクレアーゼドメインを有するが、HNHエンドヌクレアーゼドメインを有さず、Cpf1のN末端はCas9のアルファヘリックス状の認識ローブを有さない。Zetsche et al., *Cell*, 163, 759-771, 2015 (これは参照によって本明細書に組み込まれる) においては、Cpf1のRuvC様ドメインは両方のDNA鎖を切断することを担っており、RuvC様ドメインの不活性化はCpf1ヌクレアーゼ活性を不活性化するということが示されている。例えば、*Francisella novicida* Cpf1のD917A、E1006A、またはD1255Aに対応する変異はCpf1ヌクレアーゼ活性を不活性化する。いくつかの態様では、本開示のdCpf1は、配列番号34のD917A、E1006A、D1255A、D917A/E1006A、D917A/D1255A、E1006A/D1255A、またはD917A/E1006A/D1255Aに対応する変異を含む。Cpf1のRuvCドメインを不活性化するいずれかの変異、例えば、置換変異、欠失、または挿入が本開示に従って用いられ得るということは理解されるべきである。

30

40

【0349】

いくつかの態様では、本明細書において提供される融合蛋白質のいずれかの核酸によってプログラム可能なDNA結合蛋白質 (napDNAbp) はCpf1蛋白質であり得る。いくつかの態様では、Cpf1蛋白質はCpf1ニッカーゼ (nCpf1) である。いくつかの態様では、Cpf1蛋白質はヌクレアーゼ不活性型Cpf1 (dCpf1) である。いくつかの態様では、Cpf1、nCpf1、またはdCpf1は、配列番号34-41のいずれか1つと少なくとも85%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、または少なくとも99.5%同一であるア

50

ミノ酸配列を含む。いくつかの態様では、dCpf1は、配列番号34-41のいずれか1つと少なくとも85%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、または少なくとも99.5%同一であるアミノ酸配列を含み、配列番号34のD917A、E1006A、D1255A、D917A/E1006A、D917A/D1255A、E1006A/D1255A、またはD917A/E1006A/D1255Aに対応する変異を含む。他の細菌種からのCpf1もまた本開示に従って用いられ得るということは了解されるはずである。

【0350】

野生型Francisella novicida Cpf1 (配列番号34) (D917、E1006、およびD1255は太字かつ下線付きである)

【0351】

【化20】

MSIYQEFV NKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAK
 QIIDKYHQFFIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYI
 KDSEKFKNLFNQNLIDAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIIKSFKG
 WTTYFKGFHENRKNVYSSNDIPTSIIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIK
 KDLAEEELTFDIDYKTSEVNQRVFSLDEVFEIANFN NYLNQSGITKFNTIIGGK FVNGENTK
 RKGINEYINLYSQQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTM QSFY
 EQIAAFKTVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSSQQVFDDYSVIGTA
 VLEYITQQIAPKNLDNPSKKEQELIAKKTEKAKYLSLETIKLAL EEFNKHRDIDKQCRFEE
 ILANFAAIPMIFDEIAQNKDNL AQISIKYQNQGKDLLQASAEDDVKA IKDLLDQTNLL
 HKLKIFHISQSEDKANILDKDEHFYLVFEECYFELANIVPLYNKIRNYITQKPYSDEKFKL
 NFENSTLANGW DKNKEPDNTAILFIKDDKY YLGVMNKKNNKIFDDKA IKENKGEYKK
 IVYKLLPGANKMLPKVFFSAKSIK FYNPSEDILRIRNHSTHTKNGSPQKGYEKFEFNIEDC
 RKFIDFYKQSISKHPEWKDFGFRFSDTQR YNSIDEFYREVENQGYKLT FENISESYIDSVV
 NQGKLYL FQIYNKDFSA YSKGRPNLHTLYWKALFDERNLQD VVYKLNGEAELFYRKQS
 IPKKITHPAKEAIANKNDNP KESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFN
 DEINLLLKEKANDVHILSIDR GERHLAYYTLVDGKGNIIKQDTFNIIGNDRMKTNYHDKL
 AAIEKDRDSARKDWK KINNIKEMKEGYLSQVVHEIAKLVIEYNAIVVFEDLNFGFKRGR
 FKVEKQVYQKLEKMLIEKLN YLVFKDNEFDKTGGVLRAYQLTAPFETFKKMGKQTGII
 YYVPAGFTSKICPVTGFVNQLYPKYESVSKSQEFFSKFDKICYNLDKGYFEFSFDYKNFG
 DKA AKGKWTIASFGSRLINFRNSDKNHNWD TREVYPTKELEKLLKDYSIEYGHGECIKA
 AICGESDKKFFAKLTSVLNTILQMRNSKTGT ELDYLISPVADVNGNFFDSRQAPKNMPQ
 DADANGAYHIGLKGLM LLGRIKNNQEGKKNLVIKNEEYFEFVQNRNN (配列番号
 34)

10

20

30

40

【0352】

Francisella novicida Cpf1 D917A (配列番号35) (A917、E1006、およびD1255は太字かつ下線付きである)

【0353】

【化 2 1】

MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAK
 QIIDKYHQFFIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYI
 KDSEKFKNLFNQNLDIAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIIKSFKG
 WTTYFKGFHENRKNVYSSNDIPTSHYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIK
 KDLAEELTFDIDYKTSEVNQRVFLDEVFIEANFNNYLNQSGITKFNTIIGGK FVNGENTK
 RKGINEYINLYSQQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTTMQSFY
 EQIAAFKTVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLDLSQQVFDDYSVIGTA 10
 VLEYITQQIAPKNLDNPSKKEQELIAKKTAKAYLSLETIKLALAEFNKHRDIDKQCRFEE
 ILANFAAIPMIFDEIAQNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIKDLLDQTNLL
 HKLKIFHISQSEDKANILDKDEHFYLVFEECYFELANIVPLYNKIRNYITQKPYSDEKFKL
 NFENSTLANGWDKNKEPDNTAILFIKDDKY YLGV MNKKNNKIFDDKA IKENKGEGYKK
 IVYKLLPGANKMLPKVFFSAKSIKFYNPSEDILRIRNHSTHTKNGSPQKGYEKFEFNIEDC
 RKFIDFYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYREVENQGYKLT FENISESYIDSVV
 NQGKLYLFQIYNKDFSAYSKGRPNLHTLYWKALFDERNLQDVVYKLNGEAELFYRKQS 20
 IPKKITHPAKEAIAANKNDNPKKESVF EYDLIKDKRFTEDKFFFHCPITINFKSSGANKFN
 DEINLLLKEKANDVHILSIARGERHLAYYTLVDGKGNIIKQDTFNIIGNDRMKTNYHDKL
 AAIEKDRDSARKDWKKNINIKEMKEGYLSQVVHEIAKL VIEYNAIVVFEDLNFGFKRGR
 FKVEKQVYQKLEKMLIEKLNLYLVFKDNEFDKTGGVLRAYQLTAPFETFKKMGKQTGII
 YYVPAGFTSKICPVTGFVNQLYPKYESVSKSQEFFSKFDKICYNLDKGYFEFSFDYKNFG
 DKAAGKWTIASFGSRLINFRNSDKNHNWDTRE VYPTKELEKLLKDYSIEYGHGECIKA
 AICGESDKKFFAKLTSVLNTILQMRNSKTGT ELDYLISPVADVNGNFFDSRQAPKNMPQ 30
 DADANGAYHIGLKGLMLLGRIKNNQEGKLNLVIKNEEYFEFVQNRNN (配列番号
 35)

【 0 3 5 4】

Francisella novicida Cpf1 E1006A (配列番号36) (D917、A1006、およびD1255は太字か
 つ下線付きである)

【 0 3 5 5】

【化 2 2】

MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAK
 QIIDKYHQFFIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYI
 KDSEKFKNLFNQNLIIDAKKGGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIIKSFKG
 WTTYFKGFHENRKNVYSSNDIPTSIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIK
 KDLAEELTFDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKFNTIIGGKRVNGENTK
 RKGINEYINLYSQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTMQSFY
 EQIAAFKTVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDYSVIGTA
 VLEYITQQIAPKNLDNPSKKEQELIAKKTEKAKYLSLETIKLALAEFNKHRDIDKQCRFEE
 ILANFAAIPMIFDEIAQNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIKDLLDQTNNLL
 HKLKIFHISQSEDKANILDKDEHFYLVFEECYFELANIVPLYNKIRNYITQKPYSDEKFKL
 NFENSTLANGWDKNKEPDNTAILFIKDDKY YLGVMNKKNNKIFDDKAIKENKGEYKK
 IVYKLLPGANKMLPKVFFSAKSIKFNPSIEDILRIRNHSTHTKNGSPQKGYEKFENIEDC
 RKFIDFYKQSISKHPEWKDFGFRSDTQRYNSIDEFYREVENQGYKLTFFENISESYIDSVV
 NQGKLYLFQIYNKDFSAYSKGRPNLHTLYWKALFDERNLQDVVYKLNGEAELFYRKQS
 IPKKITHPAKEAIAANKNDNPKKESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFN
 DEINLLLKEKANDVHILSIDRGERHLAYYTLVDGKGNIIKQDTFNIIGNDRMKTNYHDKL
 AAIEKDRDSARKDWKKINNIKEMKEGYLSQVVHEIAKLVIEYNAIVVFA^ADLNFGFKRGR
 FKVEKQVYQKLEKMLIEKLNLYLVFKDNEFDKTGGVLRAYQLTAPFETFKKMGKQTGII
 YYVPAGFTSKICPVTGFVNQLYPKYESVSKSQEFFSKFDKICYNLDKGYFEFSFDYKNFG
 DKAAGKWTIASFGSRLINFRNSDKNHNWDTREVVYPTKELEKLLKDYSIEYGHGECIKA
 AICGESDKKFFAKLTSVLNTILQMRNSKTGTELDYLISPVADVNGNFFDSRQAPKNMPQ
 DAD^AANGAYHIGLKGLMLLGRIKNNQEGKLNLVIKNEEYFEFVQNRNN (配列番号
 36)

10

20

30

【 0 3 5 6】

Francisella novicida Cpf1 D1255A (配列番号37) (D917、E1006、およびA1255は太字か
 つ下線付きである)

【 0 3 5 7】

【化 2 3 - 1】

MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAK
 QIIDKYHQFFIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYI
 KDSEKFKNLFNQNLIIDAKKGGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIIKSFKG
 WTTYFKGFHENRKNVYSSNDIPTSIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIK
 KDLAEELTFDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKFNTIIGGKRVNGENTK

40

【化 2 3 - 2】

RKGINEYINLYSQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTMQSFY
 EQIAAFKTVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDYSVIGTA
 VLEYITQQIAPKNLDNPSKKEQELIAKKTEKAKYLSLETIKLALAEFNKHRDIDKQCRFEE
 ILANFAAIPMIFDEIAQNKDNLAQISIKYQNQGKDLLQASAEDDVKAIKDLLDQTNLL
 HKLKIFHISQSEDKANILDKDEHFYLVFEECYFELANIVPLYNKIRNYITQKPYSDEKFKL
 NFENSTLANGWDKNKEPDNTAILFIKDDKY YLGVMNKKNNKIFDDKAIKENKGEYKK
 IVYKLLPGANKMLPKVFFSAKSIKFYNPSEDILRIRNHSTHTKNGSPQKGYEKFENIEDC
 RKFIDFYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYREVENQGYKLT FENISESYIDSVV
 NQGKLYLFQIYNKDFSAYS KGRPNLHTLYWKALFDERNLQDVVYKLNGEAELFYRKQS
 IPKKITHPAKEAIANKNDNPKKESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFN
 DEINLLLKEKANDVHILSIDRGERHLA YYTLVDGKGNIIKQDTFNIIGNDRMKTNYHDKL
 AAIEKDRDSARKDWKKINNIKEMKEGYLSQVVHEIAKL VIEYNAIVVFEDLNFGFKRGR
 FKVEKQVYQKLEKMLIEKLNLYLVFKDNEFDKTGGVLRAYQLTAPFETFKKMGKQTGII
 YYVPAGFTSKICPVTGFVNQLYPKYESVSKSQEFFSKFDKICYNLDKGYFEFSFDYKNFG
 DKAAGKWTIASFGSRLINFRNSDKNHNWDTREVYPTKELEKLLKDYSIEYGHGECIKA
 AICGESDKKFFAKLTSVLNTILQMRNSKTGTELDYLISPVADVNGNFFDSRQAPKNMPQ
 DAAANGAYHIGLKGLMLLGRIKNNQEGKKLNLVIKNEEYFEFVQNRNN (配列番号
 37)

10

20

【 0 3 5 8】

Francisella novicida Cpf1 D917A/E1006A (配列番号38) (A917、A1006、およびD1255は
 太字かつ下線付きである)

【 0 3 5 9】

30

【化 2 4 - 1】

MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAK
 QIIDKYHQFFIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYI
 KDSEKFKNLFNQNLIDAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIISFKG
 WTTYFKGFHENRKNVYSSNDIPTSIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIK
 KDLAEELTFDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKFNTIIGGKFFVNGENTK
 RKGINEYINLYSQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTMQSFY
 EQIAAFKTVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDYSVIGTA
 VLEYITQQIAPKNLDNPSKKEQELIAKKTEKAKYLSLETIKLALAEFNKHRDIDKQCRFEE
 ILANFAAIPMIFDEIAQNKDNLAQISIKYQNQGKDLLQASAEDDVKAIKDLLDQTNLL
 HKLKIFHISQSEDKANILDKDEHFYLVFEECYFELANIVPLYNKIRNYITQKPYSDEKFKL

40

【化 2 4 - 2】

NFENSTLANGWDKNKEPDNTAILFIKDDKY YLGVMNKKNNKIFDDKAIKENKGEGYKK
 IVYKLLPGANKMLPKVFFSAKSIKFYNPSEDILRIRNHSTHTKNGSPQKGYEKFEFNIEDC
 RKFIDFYKQSISKHPEWKDFGFRFSDTQRYSIDEFYREVENQGYKLTFFENISESYIDSVV
 NQGKLYLFQIYNKDFSAYSKGRPNLHTLYWKALFDERNLQDVVYKLNGEAELFYRKQS
 IPPKITHPAKEAIANKNDNPKKESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFN
 DEINLLLKEKANDVHILSIARGERHLAYYTLVDGKGNIKQDTFNIIGNDRMKTNYHDKL
 AAIEKDRDSARKDWKKINNIKEMKEGYLSQVVHEIAKLVIEYNAIVVFADLNFGFKRGR
 FKVEKQVYQKLEKMLIEKLNLYLVFKDNEFDKTGGVLRAYQLTAPFETFKKMGKQTGII
 YYVPAGFTSKICPVTGFVNQLYPKYESVSKSQEFFSKFDKICYNLDKGYFEFSFDYKNFG
 DKAAGKWTIASFGSRLINFRNSDKNHNWDTREVPYPTKELEKLLKDYSIEYGHGECIKA
 AICGESDKKFFAKLTSVLNTILQMRNSKTGTELDYLISPVADVNGNFFDSRQAPKNMPQ
 DADANGAYHIGLKGLMLLGRINNKQEGKKNLVIKNEEYFEFVQNRNN (配列番号
 38)

10

【 0 3 6 0】

Francisella novicida Cpf1 D917A/D1255A (配列番号39) (A917、E1006、およびA 1255
 は太字かつ下線付きである)

20

【 0 3 6 1】

【化 2 5 - 1】

MSIQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAK
 QIIDKYHQFFIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYI
 KDSEKFKNLFNQNLIDAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIISFKG
 WTTYFKGFHENRKNVYSSNDIPTSIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIK
 KDLAEEELTFDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKFNTHGGKVFVNGENTK
 RKGINEYINLYSQQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTMMSFY
 EQIAAFKTVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSSQVFDYDYSVIGTA
 VLEYITQQIAPKNLDNPSKKEQELIAKKTEKAKYLSLETIKLALFEFNKHRDIDKQCRFEE
 ILANFAAIPMIFDEIAQNKDNLAQISIKYQNQGKDLLQASAEDDVKAIKDLLDQTNLL
 HKLKIFHISQSEDKANILDKDEHFYLVFEECYFELANIVPLYNKIRNYITQKPYSDEKFKL
 NFENSTLANGWDKNKEPDNTAILFIKDDKY YLGVMNKKNNKIFDDKAIKENKGEGYKK
 IVYKLLPGANKMLPKVFFSAKSIKFYNPSEDILRIRNHSTHTKNGSPQKGYEKFEFNIEDC
 RKFIDFYKQSISKHPEWKDFGFRFSDTQRYSIDEFYREVENQGYKLTFFENISESYIDSVV
 NQGKLYLFQIYNKDFSAYSKGRPNLHTLYWKALFDERNLQDVVYKLNGEAELFYRKQS
 IPPKITHPAKEAIANKNDNPKKESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFN

30

40

【化 2 5 - 2】

DEINLLLKEKANDVHILSIARGERHLAYYTLDVGKGNIIKQDTFNIIGNDRMKTNYHDKL
 AAIEKDRDSARKDWKINNIKEMKEGYLSQVVHEIAKLVIEYNAIVVFEDLNFGFKRGR
 FKVEKQVYQKLEKMLIEKLNLYLVFKDNEFDKTGGVLRAYQLTAPFETFKKMGKQTGII
 YYVPAGFTSKICPVTGFVNQLYPKYESVSKSQEFFSKFDKICYNLDKGYFEFSFDYKNFG
 DKAAGKWTIASFGSRLINFRNSDKNHNWDTREVVYPTKELEKLLKDYSIEYGHGECIKA
 AICGESDKKFFAKLTSVLNTILQMRNSKTGTELDYLISPVADVNGNFFDSRQAPKNMPQ
 DAAANGAYHIGLKGLMLLGRIKNNQEGKKNLVIKNEEYFEFVQNRNN (配列番号
 39)

10

【 0 3 6 2】

Francisella novicida Cpf1 E1006A/D1255A (配列番号40) (D917、A1006、およびA 1255
 は太字かつ下線付きである)

【 0 3 6 3】

【化 2 6】

MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAK
 QIIDKYHQFFIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYI
 KDSEKFKNLFNQNLDIAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIIKSFKG
 WTTYFKGFHENRKNVYSSNDIPTSIIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIK
 KDLAEELTFDIDYKTSEVNQRVFSLDEVFEIANFNLYLNQSGITKFNTIIGGKFVNGENTK
 RKGINEYINLYSQQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTTMQSFY
 EQIAAFKTVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDYSVIGTA
 VLEYITQQIAPKNLDNPSKKEQELIAKKTEKAKYLSLETIKLALAEFNKHRDIDKQCRFEE
 ILANFAAIPMIFDEIAQNKDNLAQISIKYQNQGKDLLQASAEDDVKAIKDLLDQTNLL
 HKLKIFHISQSEDKANILDKDEHFYLVFEECYFELANIVPLYNKIRNYITQKPYSDEKFKL
 NFENSTLANGWDKNKEPDNTAILFIKDDKYYLGMNKKNNKIFDDKAIKENKGEYKK
 IVYKLLPGANKMLPKVFFSAKSIKFYNPSEDILRIRNHSTHTKNGSPQKGYEKFENIEDC
 RKFIDFYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYREVENQGYKLTFFENISESYIDSVV
 NQGKLYLFIYNKDFSAYSKGRPNLHTLYWKALFDERNLQDVVYKLNGEAELFYRKQS
 IPKKITHPAKEAIANKNDNPCKESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFN
 DEINLLLKEKANDVHILSIDRGERHLAYYTLDVGKGNIIKQDTFNIIGNDRMKTNYHDKL
 AAIEKDRDSARKDWKINNIKEMKEGYLSQVVHEIAKLVIEYNAIVVFADLNFGFKRGR
 FKVEKQVYQKLEKMLIEKLNLYLVFKDNEFDKTGGVLRAYQLTAPFETFKKMGKQTGII
 YYVPAGFTSKICPVTGFVNQLYPKYESVSKSQEFFSKFDKICYNLDKGYFEFSFDYKNFG
 DKAAGKWTIASFGSRLINFRNSDKNHNWDTREVVYPTKELEKLLKDYSIEYGHGECIKA
 AICGESDKKFFAKLTSVLNTILQMRNSKTGTELDYLISPVADVNGNFFDSRQAPKNMPQ
 DAAANGAYHIGLKGLMLLGRIKNNQEGKKNLVIKNEEYFEFVQNRNN (配列番号
 40)

20

30

40

50

【 0 3 6 4 】

Francisella novicida Cpf1 D917A/E1006A/D1255A (配列番号41) (A917、A1006、およびA1255は太字かつ下線付きである)

【 0 3 6 5 】

【 化 2 7 】

MSIYQEFV NKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAK
 QIIDKYHQFFIEEILSSVCISEDLLQNYSDVYFKLKKSSDDNLQKDFKSAKDTIKKQISEYI
 KDSEKFKNLFNQNLIDAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIIKSFKG
 WTTYFKGFHENRKNVYSSNDIPTSIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIK 10
 KDLAEEITFDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKFNTIIGGKFVNGENTK
 RKGINEYINLYSQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTMQSFY
 EQIAAFKTVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSSQVFDYVIGTA
 VLEYITQQIAPKNLDNPSKKEQELIAKKTEKAKYLSLETIKLALIEFNKHRDIDKQCRFEE
 ILANFAAIPMIFDEIAQNKDNLAQISIKYQNQGKDLLQASAEDDVKAIKDLLDQTNLL
 HKLKIFHISQSEDKANILDKDEHFYLVFEECYFELANIVPLYNKIRNYITQKPYSDEKFKL
 NFENSTLANGWDKNKEPDNTAILFIKDDKY YLGVMNKKNNKIFDDKAIKENKGEYKK 20
 IVYKLLPGANKMLPKVFFSAKSIKFNPSDILRIRNHSTHTKNGSPQKGYEKFEFNIEDC
 RKFIDFYKQSISKHPEWKDFGFRFSDTQRYSIDEFYREVENQGYKLT FENISESYIDSVV
 NQGKLYLFQIYNKDFSAYS KGRPNLHTLYWKALFDERNLQDVVYKLNGEAELFYRKQS
 IPKKITHPAKEAIANKNDNPKKESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFN
 DEINLLLKEKANDVHILSIARGERHLA YYTLVDGKGNIKQDTFNIIGNDRMKTNYHDKL
 AAIEKDRDSARKDWKINNIKEMKEGYLSQVVHEIAKLVIEYNAIVVFADLNFGFKRGR
 FKVEKQVYQKLEKMLIEKLNLYLVFKDNEFDKTGGVLRAYQLTAPFETFKKMGKQTGII
 YYVPAGFTSKICPVTGFVNQLYPKYESVSKSQEFFSKFDKICYNLDKGYFEFSFDYKNFG 30
 DKAAGKWTIASFGSRLINFRNSDKNHNWDTREVVYPTKELEKLLKDYSIEYGHGECIKA
 AICGESDKKFFAKLTSVLNTILQMRNSKTGTELDYLISPVADVNGNFFDSRQAPKNMPQ
 DAAANGAYHIGLKGLMLLGRIKNNQEGKLLNLVIKNEEYFEFVQNRNN (配列番号
 41)

【 0 3 6 6 】

いくつかの態様では、核酸によってプログラム可能なDNA結合蛋白質 (napDNAbp) は、古典的な (NGG) PAM配列を要求しない核酸によってプログラム可能なDNA結合蛋白質である。いくつかの態様では、napDNAbpはargonaute蛋白質である。かかる核酸によってプログラム可能なDNA結合蛋白質の1つの例は*Natronobacterium gregoryi*からのArgonaute蛋白質 (NgAgo) である。NgAgoはssDNAによってガイドされるエンドヌクレアーゼである。NgAgoは、それをその標的部位へとガイドするための約24ヌクレオチドの5'リン酸化されたssDNA (gDNA) に結合し、gDNA部位においてDNA二本鎖切断を作るであろう。Cas9とは対照的に、NgAgo-gDNAシステムはプロトスペーサー隣接モチーフ (PAM) を要求しない。ヌクレアーゼ不活性型NgAgo (dNgAgo) を用いることは、標的化され得る塩基を多大に拡大し得る。NgAgoのキャラクタリゼーションおよび使用はGao et al., *Nat Biotechnol.*, 2016 Jul;34 (7): 768-73. PubMed PMID: 27136078 ; Swarts et al, *Nature*. 507 (7491) (2014): 258-61 ; およびSwarts et al., *Nucleic Acids Res.* 43 (10) (2015):5120-9に記載されており、これらのそれぞれは参照によって本明細書に組み込まれる。*Natronobacteri* 40 50

um gregoryi Argonauteの配列は配列番号42によって提供されている。

【 0 3 6 7 】

野生型Natronobacterium gregoryi Argonaute (配列番号42)

【 0 3 6 8 】

【 化 2 8 】

```

MTVIDLDSTTTADELTSGHTYDISVTLTGVDNTDEQHPRMSLAFEQDNG
ERRYTILWKNTTPKDVFTYDYATGSTYIFTNIDYEVKDGyenLTATYQTTVENATAQEV
GTTDEDETFAGGEPLDHHLLDDALNETPDDAETESDSGHVMTSFASRDQLPEWTLHTYTL
TATDGAKTDTHEYARRTLAYTVRQELYTDHDAAPVATDGLMLLTPEPLGETPLDLDCGV
RVEADETRTLDYTTAKDRLLARELVEEGLKRSLWDDYLVRGIDEVLSKEPVLTCEFDL
HERYDLSVEVGHSGRAYLHINFRHRFVPKLTLADIDDDNIYPGLRVKTTYRPRRGHIVW
GLRDECATDSLNTLGNQSVVAYHRNNQTPINTDLLDAIEAADRRVVETRRQGHGDDAV
SFPQELLAVEPNTHQIKQFASDGFHQQARSKTRLSASRCSEKAQAFERLDPVRLNGSTV
EFSSEFFTGNNEQQLRLLYENGESVLTFRDGARGAHPDETFSKGIVNPPESFEVAVVLPE
QQADTCKAQWDTMADLLNQAGAPPTRSETVQYDAFSSPESISLNVAGAIIDPSEVDAAF
VVLPPDQEGFADLASPTETYDELKKALANMGIYSQMA YFDRFRDAKIFYTRNVALGLL
AAAGGVAFTTEHAMPGDADMFIGIDVSRYPEDGASGQINIAATATAVYKDGITLGHSS
TRPQLGEKLQSTDVRDIMKNAILGYQQVTGESPTHIVHRDGMNEDLDPATEFLNEQG
VEYDIVEIRKQPQTRLLAVSDVQYDTPVKSIAAINQNEPRATVATFGAPEYLA TRDGGGL
PRPIQIERVAGETDIETLTRQVYLLSQSHIQVHNSTARLPITTAYADQASTHATKGYLVQT
GAFESNVGFL ( 配列番号 42)

```

10

20

【 0 3 6 9 】

いくつかの態様では、napDNAbpはArgonaute蛋白質の原核生物ホモログである。Argonaute蛋白質の原核生物ホモログは公知であり、例えば、Makarova K., et al., "Prokaryotic homologs of Argonaute proteins are predicted to function as key components of a novel system of defense against mobile genetic elements", Biol Direct. 2009 Aug 25; 4: 29. doi: 10.1186/1745-6150-4-29に記載されており、この内容全体は参照によってここに組み込まれる。いくつかの態様では、napDNAbpはMarinitoga piezophila Argonaute (MpAgo) 蛋白質である。CRISPR関連Marinitoga piezophila Argonaute (MpAgo) 蛋白質は5'リン酸化されたガイドを用いて一本鎖標的配列を切断する。5'ガイドは全ての公知のArgonauteによって用いられる。MpAgo-RNA複合体の結晶構造は、5'リン酸相互作用をブロックする残基を含むガイド鎖結合部位を示す。このデータは5'-ヒドロキシル化ガイドへの非古典的な特異性を有するArgonauteサブクラスの進化を示唆している。例えば、Kaya et al., "A bacterial Argonaute with non-canonical guide RNA specificity", Proc Natl Acad Sci USA. 2016 Apr 12; 113 (15): 4057-62を参照する。この内容全体は参照によってここに組み込まれる)。他のargonaute蛋白質が用いられ得、本開示の範囲内であるということは了解されるはずである。

30

40

【 0 3 7 0 】

いくつかの態様では、核酸によってプログラム可能なDNA結合蛋白質 (napDNAbp) は微生物CRISPR-Casシステムの1個のエフェクターである。微生物CRISPR-Casシステムの1個のエフェクターは、限定なしに、Cas9、Cpf1、C2c1、C2c2、およびC2c3を包含する。典型的には、微生物CRISPR-Casシステムはクラス1およびクラス2システムに分けられる。クラス1システムはマルチサブユニットエフェクター複合体を有する。一方で、クラス2システム

50

は1個の蛋白質エフェクターを有する。例えば、Cas9およびCpf1はクラス2エフェクターである。Cas9およびCpf1に加えて、3つの別個のクラス2 CRISPR-Casシステム (C2c1、C2c2、およびC2c3) がShmakov et al., "Discovery and Functional Characterization of Diverse Class 2 CRISPR Cas Systems", *Mol. Cell*, 2015 Nov 5; 60 (3): 385-397によって記載されており、この内容全体は参照によってここに組み込まれる。システムの2つのエフェクターC2c1およびC2c3は、Cpf1に近縁のRuvC様エンドヌクレアーゼドメインを含有する。第3のシステムC2c2は、2つの予測されるHEPN RNaseドメインを有するエフェクターを含有する。C2c1によるCRISPR RNAの産生とは違って、成熟CRISPR RNAの産生はtracrRNA非依存的である。C2c1はDNA切断のためにはCRISPR RNAおよびtracrRNA両方に依存する。細菌のC2c2は、そのRNAによって活性化される一本鎖RNA分解活性とは別個のCRISPR RNA成熟のためのユニークなRNase活性を所有することが示されている。これらのRNase機能は、互いとは、かつCpf1のCRISPR RNAプロセッシング挙動とは異なる。例えば、East-Seletsky, et al., "Two distinct RNase activities of CRISPR-C2c2 enable guide-RNA processing and RNA detection", *Nature*, 2016 Oct 13; 538 (7624): 270-273を参照する。この内容全体は参照によってここに組み込まれる。Leptotrichia shahiiのC2c2の *in vitro* 生化学分析は、C2c2が1個のCRISPR RNAによってガイドされ、相補的プロトスペーサーを持つssRNA標的を切断するようにプログラムされ得るということを示す。2つの保存されたHEPNドメインの触媒残基が切断を媒介する。触媒残基の変異は触媒的に不活性なRNA結合蛋白質を生成する。例えば、Abudayyeh et al., "C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector", *Science*, 2016 Aug 5; 353 (6299)を参照する。この内容全体は参照によってここに組み込まれる。

10

20

【0371】

Alicyclobacillus acidoterrastris C2c1 (AacC2c1) の結晶構造がキメラ1分子ガイドRNA (sgRNA) との複合体によって報告されている。例えばLiu et al., "C2c1-sgRNA Complex Structure Reveals RNA-Guided DNA Cleavage Mechanism", *Mol. Cell*, 2017 Jan 19; 65(2): 310-322を参照する。この内容全体は参照によってここに組み込まれる。結晶構造は三者複合体として標的DNAに結合した*Alicyclobacillus acidoterrestris* C2c1によってもまた報告されている。例えば、Yang et al., "PAM-dependent Target DNA Recognition and Cleavage by C2C1 CRISPR-Cas endonuclease", *Cell*, 2016 Dec 15; 167(7): 1814-1828を参照する。この内容全体は参照によってここに組み込まれる。標的および非標的DNA鎖両方によるAacC2c1の触媒的にコンピテントな立体配座が、独立して1個のRuvC触媒ポケット内に位置してキャプチャーされており、C2c1によって媒介される切断は標的DNAの互い違いの7ヌクレオチド切断をもたらす。C2c1三者複合体ならびに以前に同定されたCas9およびCpf1カウンターパートの間の構造比較は、CRISPR-Cas9システムによって用いられるメカニズムの多様性を実証している。

30

【0372】

いくつかの態様では、本明細書において提供される融合蛋白質のいずれかの核酸によってプログラム可能なDNA結合蛋白質 (napDNAbp) はC2c1、C2c2、またはC2c3蛋白質であり得る。いくつかの態様では、napDNAbpはC2c1蛋白質である。いくつかの態様では、napDNAbpはC2c2蛋白質である。いくつかの態様では、napDNAbpはC2c3蛋白質である。いくつかの態様では、napDNAbpは、天然に存在するC2c1、C2c2、またはC2c3蛋白質と少なくとも85%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、または少なくとも99.5%同一であるアミノ酸配列を含む。いくつかの態様では、napDNAbpは天然に存在するC2c1、C2c2、またはC2c3蛋白質である。いくつかの態様では、napDNAbpは、配列番号43または44のいずれか1つと少なくとも85%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、または少なくとも99.5%同一であるアミノ酸配列を含む。いくつかの態様では、napDNAbpはいずれか1つの配列番号43または44のアミノ酸配列を含む。他の細菌種からのC2c1、C2c2、またはC2c3もまた本開示に従って用いられ

40

50

得るということは了解されるはずである。

【 0 3 7 3 】

C2c1 (uniprot.org/uniprot/T0D7 A2#)

sp|T0D7A2|C2C1_ALIAG CRISPR関連エンドヌクレアーゼC2c1

OS=Alicyclobacillus acidoterrestris (株ATCC 49025 / DSM 3922 / CIP 106132 / NCIM B 13137 / GD3B) GN=c2c1 PE=1 SV=1

【 0 3 7 4 】

【 化 2 9 】

MAVKSIVKLRLLDDMPEIRAGLWKLHKEVNAGVRYYTEWLSLLRQENL 10
 YRRSPNGDGEQECDKTAEECKAELLERLRARQVENGHRGPAGSDDELLQLARQLYELL
 VPQAIGAKGDAQQIARKFLSPLADKDAVGGLGIAKAGNKPRWVRMREAGEPGWEEEK
 EKAETRKSADRTADVLRALADFGKPLMRVYTDSEMSSVEWKPLRKGQAVRTWDRD
 MFQQAIERMMSWESWNQRVQGEYAKLVEQKNRFEQKNFVQGEHLVHLVNQLQQDM
 KEASPGLESKEQTAHYVTGRALRGSDKVFEEKWGKLPDAPFDLYDAEIKNVQRRNTRR
 FGSHDLFAKLAPEYQALWREDASFLTRYAVYNSILRKLNHAKMFATFTLPDATAHPIW
 TRFDKLGGNLHQYTFLFNEFGERRHAIRFHKLLKVENGVAREVDDVTVPISMSEQLDNL
 LPRDPNEPIALYFRDYGAEQHFTGEFGGAKIQCRRDQLAHMHRRRGARDVYLNVSVRV 20
 QSQSEARGERRPPYAAVFRLVGDNHRAFVHFDKLSDYLAEHPDDGKLGSEGLLSGLRV
 MSVDLGLRTSASISVFRVARKDELKPNKGRVPFFFPKGNLVAVHERSQLLKLPGET
 ESKDLRAIREERQRTLRLQRTQLAYLRLLVRCGSEDVGRRERSWAKLIEQPVDAANHMT
 PDWREAFENELQKLKSLHGICSDKEWMDAVYESVRRVWRHMGKQVRDWRKDVRSGE
 RPKIRGYAKDVVGGNSIEQIEYLERQYKFLKSWFFGKVSGQVIRAEKGSRAITLREHID
 HAKEDRLKKLADRIIMEALGYVYALDERGKGGKWWAKYPPCQLILLEELSEYQFNDRPP
 SENNQLMQWSHRGVFQELINQAQVHDLVGTMYAAFSSRFDARTGAPGIRCRRVPARC 30
 TQEHNPEPFPWWLNKFVVEHTLDACPLRADDLIPTGEGEIFVSPFSAEEGDFHQIHADLN
 AAQNLQQLWSDFDISQIRLRCDWGEVDGELVLIPRLTGKRTADSYSNKVFYTTNGVTY
 YERERGGKKRRKVFAQEKLSEEEAELLVEADEAREKSVVLMRDPGSIINRGNWTRQKEF
 WSMVNQRIEGLVVKQIRSRVPLQDSACENTGDI (配列番号 43)

【 0 3 7 5 】

C2c2 (uniprot.org/uniprot/P0DOC6)

> sp|P0DOC6|C2C2_LEPSD CRISPR関連エンドリボヌクレアーゼC2c2

OS=Leptotrichia shahii (株DSM 19757 / CCUG 47503 / CIP 107916 / JCM 16776 / LB37) GN=c2c2 PE=1 SV=1

【 0 3 7 6 】

10

20

30

40

【化 3 0】

MGNLFGHKRWYEVDRDKKDFKIKRKKVVKRNYDGNKYILNINENNKEK
 IDNNKFIRKYINYKKNDNILKEFTRKFHAGNILFKLKGKEGIIRIENDDFLETEEVVLYIE
 AYGKSEKALKALGITKKKIIDEAIRQGITKDDKKIEIKRQENEEEEIEIDIRDEYTNKTLNDCSI
 ILRIIENDELETKKSIYEIFKNINMSLYKIIKIIENETEKVFENRYEYEEHLREKLLKDDKID
 VILTNFMEIREKIKSNLEILGFVKFYLVNVDGDKKSKNKKMLVEKILNINVDLTVEDIAD
 FVIKELEFWNITKRIEKVKKVNNEFLEKRRNRITYIKSYVLLDKHEKFKIERENKKDKIVK
 FFVENIKNNSIKEKIEKILAEFKIDELIKLEKELKKGNCDETEIFGIFKKHYKVNFDKSKKFS
 KKSDEEKELYKIIYRYLKGRIEKILVNEQKVRLLKMEKIEIEKILNESILSEKILKRVKQYT
 LEHIMYLGKLRHNDIDMTTVNTDDFSRLHAKEELDLELITFFASTNMELNKIFSRENINN
 DENIDFFGGDREKNYVLDKKILNSKIKIIRDLDFIDNKNNITNNFIRKFTKIGTNERNRILH
 AISKERDLQGTQDDYNKVINIIQNLKISDEEVSKALNLDVVFVKDKKNITKINDIKISEENN
 NDIKYLPSFSKVLPEILNLYRNNPKNEPFDTIETEKIVLNALIYVNKELYKLLILEDDLEEN
 ESKNIFLQELKKTGLNIDEIDENIENYKNAQISASKGNNKAIKKYQKKVIECYIGYLRK
 NYEELFDFSDFKMNIQEIKKQIKDINDNKTYERITVKTSDKTIVINDDFEYIISIFALLNSNA
 VINKIRNRFFATSVWLNTSEYQNIIDILDEIMQLNTRLNECITENWNLNLEEFIQKMKEIEK
 DFDDFKIQTKKEIFNNYEDIKNNILTEFKDDINGCDVLEKKLEKIVIFDDETKFEIDKKS
 ILQDEQRKLSNINKKDLKKKVDQYIKDKDQEIKSKILCRIIFNSDFLKKYKKEIDNLIEDM
 ESEENKQFQEIYYPKERKNELYIYKKNLFLNIGNPNFDKIYGLISNDIKMADAKFLFNIDG
 KNIRKNKISEIDAILKNLNDKLNKNGYSKEYKEYIKKLENDFFAKNIQNKNYKSFEKDY
 NRVSEYKKIRDLVEFNLYLNKIESYLIDINWKLAIQMARFERDMHYIVNGLRELGIKLSG
 YNTGISRAYPKRNGSDGFYTTTAYYKFFDEESYKKFEKICYGFGIDLSENSEINKPENESI
 RNYISHFYIVRNPFADYSIAEQIDRVSNLLSYSTRYNNSTYASVFEVFKKDVNLDYDELK
 KFKLIGNNDILERLMKPKKVSVLELESYNSDYIKNLIHELLTKIENTNDTL (配列番号
 44)

10

20

30

【 0 3 7 7】

本開示のいくつかの側面は、異なるPAM特異性を有するCas9ドメインを提供する。典型的には、Cas9蛋白質、例えば*S. pyogenes*からのCas9 (spCas9) は、特定の核酸領域に結合するために古典的なNGG PAM配列を要求する。これはゲノム中の所望の塩基を編集する能力を限定し得る。いくつかの態様では、本明細書において提供される塩基編集融合蛋白質は精確な場所に置かれることを必要とし得、例えば、標的塩基は、PAMのおよそ15塩基上流にある4塩基領域 (例えば「脱アミノ化ウィンドウ」) 内に置かれる。Komor, A.C., et al., "Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage" *Nature* 533, 420-424 (2016)を参照する。この内容全体は参照によってここに組み込まれる。従って、いくつかの態様では、本明細書において提供される融合蛋白質のいずれかは、古典的な (例えばNGG) PAM配列を含有しないヌクレオチド配列に結合する能力があるCas9ドメインを含有し得る。非古典的なPAM配列に結合するCas9ドメインは当分野において記載されており、当業者には明らかであろう。例えば、非古典的なPAM配列に結合するCas9ドメインはKleinstiver, B. P., et al., "Engineered CRISPR-Cas9 nucleases with altered PAM specificities" *Nature* 523, 481-485 (2015); およびKleinstiver, B. P., et al., "Broadening the targeting range of Staphylococcus a

40

50

ureus CRISPR-Cas9 by modifying PAM recognition" Nature Biotechnology 33, 1293-1298 (2015)に記載されており；それぞれの内容全体は参照によってここに組み込まれる。

【0378】

いくつかの態様では、Cas9ドメインはStaphylococcus aureusからのCas9ドメイン (SaCas9) である。いくつかの態様では、SaCas9ドメインはヌクレアーゼ活性型SaCas9、ヌクレアーゼ不活性型SaCas9 (SaCas9d)、またはSaCas9ニッカーゼ (SaCas9n) である。いくつかの態様では、SaCas9はアミノ酸配列配列番号45を含む。いくつかの態様では、SaCas9は配列番号45のN579X変異を含み、XはNを例外とするいずれかのアミノ酸である。いくつかの態様では、SaCas9は配列番号45のN579A変異を含む。いくつかの態様では、SaCas9ドメイン、SaCas9dドメイン、またはSaCas9nドメインは非古典的なPAMを有する核酸配列に結合し得る。いくつかの態様では、SaCas9ドメイン、SaCas9dドメイン、またはSaCas9nドメインはNNGRRT PAM配列を有する核酸配列に結合し得る。

いくつかの態様では、本明細書において提供される融合蛋白質のいずれかのCas9ドメインは、配列番号45と少なくとも60%、少なくとも65%、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、または少なくとも99.5%同一であるアミノ酸配列を含む。いくつかの態様では、本明細書において提供される融合蛋白質のいずれかのCas9ドメインは配列番号45のアミノ酸配列を含む。いくつかの態様では、本明細書において提供される融合蛋白質のいずれかのCas9ドメインは配列番号45のアミノ酸配列からなる。

【0379】

例示的なSaCas9アミノ酸配列は：

【化31】

```

KRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRG
ARRLKRRRRHRIQRVKKLLFDYNLLTDHSELSGINPYEARVKGLSQKLSEEEFSAALLHL
AKRRGVHNVNEVEEDTGNELSTKEQISRNSKALEEKYVAELQLERLKKDGEVRGSINRF
KTSDYVKEAKQLLKVQKAYHQLDQSFIDTYIDLLETRRYYEGPGEGSPFGWKDIKEW
YEMLMGHCTYFPEELRSVKYAYNADLYNALNDLNNLVITRDENEKLEYEKFQIENVF
KQKKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEFTNLKVYHDIKDITARKEIENAEILLD
QIAKILTIYQSSEDIQEELTNLNSLTQEIEQISNLKGYTGTHNLSLKAINLILDELWHTN
DNQIAIFNRLKLVPKKVDLSQKQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIKKYGLPNDII
IELAREKNSKDAQKMINEMQKRNRQTNERIEEIRTTGKENAKYLIEKIKLHDMQEGKCL
YSLEAIPLEDLLNNPFNYEVDHIIPRSVSFDNSFNKVLVKQEENSKKGNRTPFQYLSSSD
SKISYETFKKHILNLAGKGRISKTKKEYLLEERDINRFSVQKDFINRNLVDTRYATRGL
MNLRSYFRVNNLDVVKVKSINGGFTSFLRRKWKFKKERNKGYKHAEDALIIANADFIF
KEWKKLDKAKKVMENQMFEEKQAESMPEIETEQEYKEIFITPHQIKHIKDFKDYKYSHR
VDKKPNRELINDTLYSTRKDDKGNTLIVNNLNGLYDKDNDKLLKLINKSPEKLLMYHH
DPQTYQKLKLIMEQYGDEKNPLYKYEEETGNYLTKYSKKDNGPVIKKIKYGGNKLNAH
LDITDDYPNSRNKVVKLSLKPFRFDVYLDNGVYKFVTVKNLDVIKKENYEVNSKCYE
EAKLKKISNQAEFIASFYNNDLIKINGELRYVIGVNNDLLNRIEVMIDITYREYLENMN
DKRPPRIIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKG (配列番号 45)

```

である。

【0380】

50

(B) デアミナーゼドメイン

種々の態様では、本明細書において提供される進化した塩基編集因子は1つ以上の核酸エフェクタードメイン（例えばシチジンデアミナーゼ）を含み、これは任意に本明細書に記載される連続的進化プロセス（例えばPACE）を用いて進化させられ得る。

種々の態様では、核酸エフェクタードメインは、DNAまたはRNA分子を改変する能力があるいずれかの蛋白質、酵素、またはポリペプチド（またはその機能的断片）であり得る。核酸塩基改変部分は天然に存在し得るかまたは組換え体であり得る。例えば、核酸塩基改変部分は、例えば1つ以上のDNA修復酵素、および塩基除去修復（BER）、ヌクレオチド除去修復（NER）、相同性依存的な組換え修復（HR）、非相同末端結合修復（NHEJ）、マイクロホモロジー媒介末端結合修復（MMEJ）、ミスマッチ修復（MMR）、直接的な復帰修復、または他の公知のDNA修復経路に関わる酵素または蛋白質を包含し得る。核酸塩基改変部分は1つ以上の型の酵素活性を有し得、エンドヌクレアーゼ活性、ポリメラーゼ活性、リガーゼ活性、複製活性、校正活性を包含するが、これらに限定されない。核酸塩基改変部分はDNAもしくはRNA修飾酵素および/または変異性の酵素、例えばDNAメチラーゼおよび脱アミノ化酵素（すなわち、シチジンデアミナーゼおよびアデノシンデアミナーゼを包含するデアミナーゼ。全て上で定義されている）をもまた包含し得、これらは核酸塩基を脱アミノ化し、いくつかのケースでは、正常な細胞のDNA修復および複製プロセスによって変異性の修正に至る。本明細書において用いられる「核酸エフェクタードメイン」（例えば、DNAエフェクタードメインまたはRNAエフェクタードメイン）は、核酸（例えば、DNAまたはRNA）に1つ以上の改変（例えば、シチジン残基の脱アミノ化）をなす能力がある蛋白質または酵素をもまた言い得る。例示的な核酸編集ドメインは、デアミナーゼ、ヌクレアーゼ、ニッカーゼ、リコンビナーゼ、メチルトランスフェラーゼ、メチラーゼ、アセチラーゼ、アセチルトランスフェラーゼ、転写活性化因子、または転写抑制因子ドメインを包含するが、これらに限定されない。いくつかの態様では、核酸編集ドメインはデアミナーゼ（例えばシチジンデアミナーゼ、例えばAPOBECまたはAIDデアミナーゼ）またはその進化したバージョンである。

【0381】

いくつかの態様では、核酸編集ドメインはデアミナーゼを含む。いくつかの態様では、デアミナーゼはシチジンデアミナーゼである。他の態様では、デアミナーゼはアデノシンデアミナーゼである。いくつかの態様では、デアミナーゼはアポリポ蛋白質B mRNA編集複合体（APOBEC）ファミリーデアミナーゼである。いくつかの態様では、デアミナーゼはAPOBEC1デアミナーゼ、APOBEC2デアミナーゼ、APOBEC3Aデアミナーゼ、APOBEC3Bデアミナーゼ、APOBEC3Cデアミナーゼ、APOBEC3Dデアミナーゼ、APOBEC3Fデアミナーゼ、APOBEC3Gデアミナーゼ、APOBEC3Hデアミナーゼ、またはAPOBEC4デアミナーゼである。いくつかの態様では、デアミナーゼは活性化誘導デアミナーゼ（AID）である。いくつかの態様では、デアミナーゼはLamprey CDA1（pmCDA1）デアミナーゼである。いくつかの態様では、デアミナーゼはヒト、チンパンジー、ゴリラ、サル、乳牛、イヌ、ラット、またはマウスからである。いくつかの態様では、デアミナーゼはヒトからである。いくつかの態様では、デアミナーゼはラットからである。いくつかの態様では、デアミナーゼはヒトAPOBEC1デアミナーゼである。いくつかの態様では、デアミナーゼはpmCDA1である。いくつかの態様では、デアミナーゼはヒトAPOBEC3Gである。いくつかの態様では、デアミナーゼはヒトAPOBEC3Gパリアントである。いくつかの態様では、デアミナーゼは本明細書において提示されるAPOBECアミノ酸配列のいずれか1つと少なくとも80%、少なくとも85%、少なくとも90%、少なくとも92%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、または少なくとも99.5%同一である。

本開示の側面に従ってCas9ドメインに融合させられ得るいくつかの例示的な好適な核酸編集ドメイン、例えばデアミナーゼおよびデアミナーゼドメインは下で提供されている。いくつかの態様では、それぞれの配列の活性ドメイン、例えば局在シグナル（核局在配列、核外搬出シグナルなし、細胞質局在シグナル）なしのドメインが用いられ得るといふことは理解されるはずである。

10

20

30

40

50

【 0 3 8 2 】

ヒトAID:

【 0 3 8 3 】

【 化 3 2 】

MDSLLMNRRKFLYQFKNVRWAKGRRETYLCYVVKRRDSATSFSLDFGY
 LRNKGCHVELLFLRYISDWDLDPGRCYRVTWFTSWSPCYDCARHVADFLRGNPNLSL
 RIFTARLYFCEDRKAPEGLRRLHRAGVQIAIMTFKDYFYCWNTFVENHERTFKAWEGL
 HENSVRLSRQLRRILLPLYEVDDLRDAFRTLGL (配列番号 47)

10

【 0 3 8 4 】

(下線 : 核局在配列 ; 二重下線 : 核外搬出シグナル)

【 0 3 8 5 】

マウスAID:

【 0 3 8 6 】

【 化 3 3 】

MDSLLMKQKKFLYHFKNVRWAKGRHETYLCYVVKRRDSATSCSLDFGH
 LRNKGCHVELLFLRYISDWDLDPGRCYRVTWFTSWSPCYDCARHVAEFLRWPNLSL
 RIFTARLYFCEDRKAPEGLRRLHRAGVQIGIMTFKDYFYCWNTFVENRERTFKAWEGL
 HENSVRLTRQLRRILLPLYEVDDLRDAFRMLGF (配列番号 48)

20

【 0 3 8 7 】

(下線 : 核局在配列 ; 二重下線 : 核外搬出シグナル)

【 0 3 8 8 】

イヌAID:

【 0 3 8 9 】

【 化 3 4 】

MDSLLMKQRKFLYHFKNVRWAKGRHETYLCYVVKRRDSATSFSLDFGH
 LRNKGCHVELLFLRYISDWDLDPGRCYRVTWFTSWSPCYDCARHVADFLRGYPNLSL
 RIFAARLYFCEDRKAPEGLRRLHRAGVQIAIMTFKDYFYCWNTFVENREKTFKAWEGL
 HENSVRLSRQLRRILLPLYEVDDLRDAFRTLGL (配列番号 49)

30

【 0 3 9 0 】

(下線 : 核局在配列 ; 二重下線 : 核外搬出シグナル)

【 0 3 9 1 】

ウシAID:

【 0 3 9 2 】

【 化 3 5 】

MDSLLKKQRQFLYQFKNVRWAKGRHETYLCYVVKRRDSPTSFSLDFGHL
 RNKAGCHVELLFLRYISDWDLDPGRCYRVTWFTSWSPCYDCARHVADFLRGYPNLSLR
 IFTARLYFCDKERKAPEGLRRLHRAGVQIAIMTFKDYFYCWNTFVENHERTFKAWEGL
 HENSVRLSRQLRRILLPLYEVDDLRDAFRTLGL (配列番号 50)

40

【 0 3 9 3 】

(下線 : 核局在配列 ; 二重下線 : 核外搬出シグナル)

【 0 3 9 4 】

ラット : AID:

50

【化36】

MAVGSKPKAALVGPHWERERIWCFLCSTGLGTQQTGQTSRWLR
 PAATQDPVSPPRSLLMKQRKFLYHFKNVRWAKGRHETYLCYVVKRRDSATSFSLDFGY
 LRNKSCHVELLFLRYISDWDLDPGRCYRVTWFTSWSPCYDCARHVADFLRGNPNLSL
 RIFTARLTGWGALPAGLMSPARPSDYFYCWNTFVENHERTEFKAWEGLEHENSRLSRRL
RRILLPLYEVDDLRLDAFRTLGL (配列番号 51)

【0395】

(下線：核局在配列；二重下線：核外搬出シグナル)

【0396】

マウスAPOBEC-3：

【0397】

【化37】

MGPFCLGC~~SHRKC~~CYSPIRNLISQETFKFHFKNLGYAKGRKDTFLCYEVTR
 KDCDSPVSLHHGVFKNKDNIHAEICFLYWFHDKVLKVLSPREEFKITWYMSWSPCFECAEQ
 IVRFLATHHNLSLDIFSSRLYNVQDPETQQNLCRLVQEGAQVAAMDLYEFKCCWKKFV
 DNGGRRFRPWKRLLTNFRYQDSKLQEILRPCYIPVPSSSSSTLSNICLTKGLPETRFCVEG
 RRM~~DPLSEEEFY~~SQFYNQ~~RVKHLCYYHRM~~KPYLCYQLEQFNGQAPLKGCLLSEK~~GKQ~~
 HAEILFLDKIRSMELSQVITCYLTWSPCPNCAWQLAAFKRDRPDLILHIYTSRLYFHWKRP
 FQKGLCSLWQSGILVDVMDLPQFTDCWTNFVNPKRPFWPWKGLEIISRRTQRRLRRIKE
 SWGLQDLVNDFGNLQLGPPMS (配列番号 52)

【0398】

(斜体：核酸編集ドメイン)

【0399】

ラットAPOBEC-3：

【0400】

【化38】

MGPFCLGC~~SHRKC~~CYSPIRNLISQETFKFHFKNLRYAIDRKDTFLCYEVTRK
 DCDSPVSLHHGVFKNKDNIHAEICFLYWFHDKVLKVLSPREEFKITWYMSWSPCFECAEQV
 LRFLATHHNLSLDIFSSRLYNIRDPENQQNLCRLVQEGAQVAAMDLYEFKCCWKKFVD
 NGGRRFRPWKLLTNFRYQDSKLQEILRPCYIPVPSSSSSTLSNICLTKGLPETRFCVERR
 RVHLLSEEEFY~~SQFYNQ~~RVKHLCYYHG~~VKPYLCYQLEQ~~FNGQAPLKGCLLSEK~~GKQ~~HA
 EILFLDKIRSMELSQVIITCYLTWSPCPNCAWQLAAFKRDRPDLILHIYTSRLYFHWKRPFQ
 KGLCSLWQSGILVDVMDLPQFTDCWTNFVNPKRPFWPWKGLEIISRRTQRRLHRIKESW
 GLQDLVNDFGNLQLGPPMS (配列番号 53)

【0401】

(斜体：核酸編集ドメイン)

【0402】

Rhesus macaque APOBEC-3G：

【0403】

10

20

30

40

【化39】

MVEPMDPRTFVSNFNRRPILSGLNTVWLCCEVKT KDPSGPPLDAKIFQGK
VYSKAKYHPEMRFLRWFKWRQLHHDQEYKVTWYVSWSPCTRCANSVATFLAKDPKVT
 LTIFVARLYYFWKPDYQQALRILCQKRGPHATMKIMNYNEFQDCWNKFVDGRGKPF
 KPRNNLPKHHTLLQATLGELLRHLMDPGTFTSNFNKPVVSGQHETYLCYKVERLHND
 TWVPLNQHRGFLRNQAPNIHGFPKGRHAELCFLDLIPFWKLDGQQYRVTCFTSWSPCFSC
 AQEMAKFISNNEHVSLCIFAARIYDDQGRYQEGLRALHRDGAKIAMMNYSEFEYCWDT
 FVDRQGRPFQPWDGLDEHSQALSGRLRAI (配列番号 54)

10

【0404】

(斜体 : 核酸編集ドメイン ; 下線 : 細胞質局在シグナル)

【0405】

チンパンジー-APOBEC-3G :

【0406】

【化40】

MKPHFRNPVERMYQDTFSDNFYNRPILSHRNTVWLCYEVKTKGPSRPPL
DAKIFRGQVYSKLYHPEMRFFHWFSKWRKLHRDQEYEVTVYISWSPCTKCTRDRVATFLA
 EDPKVTLTIFVARLYYFWDPDYQEALRSLCQKRDGPRATMKIMNYDEFQHCWSKVFYS
 QRELFEWNNLPKYIILLHIMLGEILRHSMDPPTFTSNFNELWVRGRHETYLCYEVERL
 HNDTWVLLNQRRGFLCNQAPHKHGFLEGRHAELCFLDVIPFWKLDLHQDYRVTCFTSW
 PCFSCAQEMAKFISNNKHVSLCIFAARIYDDQGRCQEGLRTLAKAGAKISIMTYSEFKHC
 WDTFVDHQGCPFPWDGLEEHSQALSGRLRAILQNQGN (配列番号 55)

20

【0407】

(斜体 : 核酸編集ドメイン ; 下線 : 細胞質局在シグナル)

【0408】

グリーンモンキー-APOBEC-3G :

【0409】

【化41】

MNPQIRNMVEQMEPDIFVYFNNRPILSGRNTVWLCYEVKTKDPSGPPLD
ANIFQGLYPEAKDHPKFLHWFRKWRQLHRDQEYEVTVYVSWSPCTRCANSVATFLA
 EDPKVTLTIFVARLYYFWKPDYQQALRILCQERGGPHATMKIMNYNEFQHCWNEFVDG
 QGKPFKPRKNLPKHHTLLHATLGELLRHVMDPGTFTSNFNKPVVSGQRETYLCYKVE
 RSHNDTWVLLNQHRGFLRNQAPDRHGFPKGRHAELCFLDLIPFWKLDGQQYRVTCFTSW
 SPCFSCAQKMAKFISNNKHVSLCIFAARIYDDQGRCQEGLRTLHRDGAKIAVMNYSEFE
 YCWDTFVDRQGRPFQPWDGLDEHSQALSGRLRAI (配列番号 56)

40

【0410】

(斜体 : 核酸編集ドメイン ; 下線 : 細胞質局在シグナル)

【0411】

ヒト-APOBEC-3G :

【0412】

【化42】

MKPHFRNTVERMYRDTFSYNFYNRPILSRRNTVWLCYEVKTKGPSRPPL
DAKIFRGQVYSELKYHPEMRFFHWF~~SKWRKLHRDQEY~~EVTWYISWSPCTKCTRDMATFLA
 EDPKVTLTIFVARLYYFWDPDYQEALRSLCQKRDGPRATMKIMNYDEFQHCWSKFVYS
 QRELFEPWNNLPKYIILLHIMLGEILRHSMDPPTFTFNFNNEPWVRGRHETYLCYEVEVER
 MHNDTWVLLNQRRGFLCNQAPHKHGFLEGRHAELCFLDVIPFWKLDLDQDYRVTCFTSW
 SPCFSCAQEMAKFISKXHVSLCIFTARIYDDQGRCEGLRTLAEAGAKISIMTYSEFKH
 CWDTFVDHQGCPFQPWDGLDEHSQDLSGRLRAILQNQN (配列番号 57)

10

【0413】

(斜体: 核酸編集ドメイン; 下線: 細胞質局在シグナル)

【0414】

ヒトAPOBEC-3F:

【0415】

【化43】

MKPHFRNTVERMYRDTFSYNFYNRPILSRRNTVWLCYEVKTKGPSRPRL
 DAKIFRGQVYSQPEHHAEMCFLSWFCGNQLPAYKCFQITWFVSWTPCPDCVAKLAEFLAE
 HPNVTLTISAARLYYYWERDYRRALCRLSQAGARVKIMDDEEFAYCWENFVYSEGQPF
 MPWYKFDDNYAFLHRTLKEILRNPMEAMYPHIFYHFKNLRKAYGRNESWLCFTMEV
 VKHHSPVSWKRGVFRNQVDPETHCHAERCFLSWFCDDILSPNTNYEVTWYTSWSPCPECA
 GEVAEFLARHSNVNLTIFTARLYYFWDTDYQEGLRSLSQEGASVEIMGYKDFKYCWEN
 FVYNDDEPFKPKWGLKYNFLFLDSKLQEILE (配列番号 58)

20

【0416】

(斜体: 核酸編集ドメイン)

【0417】

ヒトAPOBEC-3B:

【0418】

【化44】

MNPQIRNPMERMYRDTFYDNFENEPILYGRSYTWLCYEVKIKRGRSNLL
 WDTGVFRGQVYFKPQYHAEMCFLSWFCGNQLPAYKCFQITWFVSWTPCPDCVAKLAEFL
 SEHPNVTLTISAARLYYYWERDYRRALCRLSQAGARVTIMDYEEFAYCWENFVYNEGQ
 QFMPWYKFDENYAFLHRTLKEILRYLMDPDTFTFNFNNDPLVLRRRQTYLCYEVEERLD
 NGTWVLMQHMGLCNEAKNLLCGFYGRHAELRFLDLVPSLQLDPAQIYRVTFISWSPC
 FSWGCAGEVRAFLQENTHVRLRIFAARIYDYDPLYKEALQMLRDAGAQVSIMTYDEFE
 YCWDTFVYRQGCPFQPWDGLEEHSQALSRLRAILQNQGN (配列番号 59)

40

【0419】

(斜体: 核酸編集ドメイン)

【0420】

ラットAPOBEC-3B:

【0421】

【化 4 5】

MQPQGLGPNAGMGPVCLGCSHRRPYSPIRNPLKKLYQQTFFYHFKNVRY
 AWGRKNNFLCYEVNGMDCALPVPLRQGVFRKQGHHAELCFIYWFHDKVLRVLS PME
 EFKVTWYMSWSPCSKCAEQVARFLAAHRNLSLAIFSSRLYYYLRNPNYQQKLCRLIQEG
 VHVAAMDLPFKKCNKFDNDGQPFPRWMLRINFSFYDCKLQEIFSRMNLREDVF
 YLQFNNSHRVKPVQNRYYRRKSYLCYQLERANGQEPLKGYLLYKKGEQHVEILFLEKM
 RSMELSQVRITCYLTWSPCPNCARQLAAFKKDHPDLILRIYTSRLYFYWRKKFQKGLCT
 LWRSGIHVDVMDLPQFADCWTNFNPNQRPFRPWNELEKNSWRIQRRLRRIKESWGL

10

(配列番号 60)

【 0 4 2 2】

ウシ APOBEC-3B :

【 0 4 2 3】

【化 4 6】

DGWEVAFRSGTVLKAGVLGVSMTTEGWAGSGHPGQGACVWTPGTRNTM
 NLLREVLFKQQFGNQPRVPAPYYRRKTYLCYQLKQRNDLTLDRGCFRNKKQRHAEIRFI
 DKINSLDLNPSQSYKIICYITWSPCPNCANELVNFITRNNHLKLEIFASRLYFHWIKSFKM
 GLQDLQNA GISVAVMTHTEFEDCWEQFVDNQS RPFQPWDKLEQYSASIRRLQRILTAP

20

I (配列番号 61)

【 0 4 2 4】

チンパンジー APOBEC-3B :

【 0 4 2 5】

【化 4 7】

MNPQIRNPMEWMYQRTFYYNFENEPILYGRSYTWLCYEVKIRRGHSNLL
 WDTGVFRGQMYSQPEHHAEMCFLSWFCGNQLSAYKCFQITWFVSWTPCPDCVAKLAK
 FLAEHPNVTLTISAARLYYYWERDYRRALCRLSQAGARVKIMDDEEFAYCWENFVYNE
 GQPFMPWYKFDDNYAFLHRTLKEIIRHLMDPDTFTFNFNNDPLVLRRHQTYLCYEVERL
 DNGTWVLMQHMGLFCNEAKNLLCGFYGRHAELRFLDLVPSLQLDPAQIYRVTWFIS
 WSPCFSWGCAGQVRAFLQENTHVRLRIFAARIYDYDPLYKEALQMLRDAGAQVSIMTY
 DEFAYCWDTFVYRQGCPFPWDGLEEHSQALSGRLRAILQVRASSLCMVPHRPPPPQSS
 PGPCPLCSEPPLGSLPTGRPAPSLPFLLTASFSPPPASLPPLPSLSLSPGHLPVPSFHSLT

30

SCSIQPPCSSRIRETEGWASVSKEGRDLG (配列番号 62)

40

【 0 4 2 6】

ヒト APOBEC-3C :

【 0 4 2 7】

【化48】

MNPQIRNPMKAMYPGTFYFQFKNLWEANDRNETWLCFTVEGIKRRSVVS
 WKTGVFRNQVDSETH*CHAERCFLSWFCDDILSPNTKYQVTWYTSWSPDC*CAGEVAEFLA
 RHSNVNLTIFTARLYYFQYPCYQEGLRSLSQEGVAVEIMDYEDFKYCWENFVYNDNEPF
 KPWKGLKTNFRLLRRLRESLQ (配列番号 63)

【0428】

(斜体 : 核酸編集ドメイン)

【0429】

ゴリラ APOBEC3C :

【化49】

MNPQIRNPMKAMYPGTFYFQFKNLWEANDRNETWLCFTVEGIKRRSVVSWKTGVFRN
 QVDSETH*CHAERCFLSWFCDDILSPNTNYQVTWYTSWSPCPEC*CAGEVAEFLARHSNVNLT
 FTARLYYFQDQDYQEGLRSLSQEGVAVKIMDYKDFKYCWENFVYNDDEPFKPWKGLK
 YNFRFLKRRRLQEILE (配列番号 64)

10

20

【0430】

(斜体 : 核酸編集ドメイン)

【0431】

ヒト APOBEC-3A :

【0432】

【化50】

MEASPASGPRHLMDPHIFTSNFNNGIGRHKTYLCYEVERLDNGTSVKMD
 QHRGFLHNQAKNLLCGFYGRHAELRFLDLVPSLQLDPAQIYRVTWFISWSPCFSWGCAGE
 VRAFLQENTHVRLRIFAARIYDYDPLYKEALQMLRDAGAQVSIMTYDEFKHCWDTFVD
 HQGCPFQPWDGLDEHSQALSGRLRAILQNQGN (配列番号 65)

30

【0433】

(斜体 : 核酸編集ドメイン)

【0434】

Rhesus macaque APOBEC-3A :

【化51】

MDGSPASRPRHLMDPNTFTFNFNNDLSVRGRHQTYLCYEVERLDNGTWVPMDERRGF
 LCNKAKNVPCGDYGCHVELRFLCEVPSWQLDPAQTYRVTWFISWSPCFRRGCAGQVRVFL
 QENKHVRLRIFAARIYDYDPLYQEALRTLRLDAGAQVSIMTYEEFKHCWDTFVDRQGRP
 FQPWDGLDEHSQALSGRLRAILQNQGN (配列番号 66)

40

【0435】

(斜体 : 核酸編集ドメイン)

【0436】

ウシ APOBEC-3A :

【化52】

MDEYTFTENFNQGWPSKTYLCYEMERLDGDATIPLDEYKGFVRNKGLDQPEKPCHAE
 LYFLGKIHSWNLDRNQHYRLTCFISWSPCYDCAQKLTTFLENHHISLHILASRIYTHNRFG
 CHQSGLCCELQAAGARITIMTFEDFKHCWETFVDHKGKPFQPWEGLNVKSQALCTELQAI
 LKTQQN (配列番号 67)

【0437】

(斜体 : 核酸編集ドメイン)

【0438】

10

ヒト APOBEC-3H :

【0439】

【化53】

MALLTAETFRLQFNNKRRLRRPYPRKALLCYQLTPQNGSTPTRGYFENK
 KKCHAEICFINEIKSMGLDETQCYQVTCYLTWSPCSSCAWELVDFIKAHDHLNLGIFASRLY
 YHWCKPQKGLRLLCGSQVPVEVMGFPKFADCWENFVDHEKPLSFNPYKMLEELDKN
 SRAIKRRLERIKIPGVRAQGRYMDILCDAEV (配列番号 68)

【0440】

20

(斜体 : 核酸編集ドメイン)

【0441】

Rhesus macaque APOBEC-3H :

【0442】

【化54】

MALLTAKTFSLQFNNKRRVKNPYPRKALLCYQLTPQNGSTPTRGHLKN
 KKKDHAEIRFINKIKSMGLDETQCYQVTCYLTWSPCPCAGELVDFIKAHRHLNLRIFAS
 RLYYHWRPNYQEGLLLLCGSQVPVEVMGLPEFTDCWENFVDHKEPPSFNPSEKLEELD
 KNSQAIKRRLERIKSRSVDVLENGLRSLQLGPVTPSSSIRNSR (配列番号 69)

30

【0443】

ヒト APOBEC-3D :

【0444】

【化55】

MNPQIRNPMERMYRDTFYDNFENEPILYGRSYTWLCYEVKIKRGRSNLL
 WDTGVFRGPVLPKRQSNHRQEVYFRFENHAEMCFLSWFCGNRLPANRRFQITWVSWNP
 CLPCVVKVTKFLAEHPNVTLTISAARLYYYRDRDWRWVLLRLHKAGARVKIMDYEDF
 AYCWENFVCNEGQPFMPWYKFDDNYASLHRTLKEILRNPMEAMYPHIFYFHFKNLLKA
 CGRNESWLCFTMEVTKHSAVFRKRGVFRNQVDPETHCHAERCFLSWFCDDILSPNTNY
 EVTWYTSWSPCECAGEVAEFLARHSNVNLTIFTARLCYFWDTDYQEGLCSLSQEGASV
 KIMGYKDFVSCWKNFVYSDDEPFKPKWGLQTNFRLLKRRRLREILQ (配列番号 70)

40

【0445】

(斜体 : 核酸編集ドメイン)

【0446】

ヒト APOBEC-1 :

【0447】

50

【化56】

MTSEKGPSTGDPTLRRRIEPWEFDVFDPRELRKEACLLYEIKWGMSRKI
 WRSSGKNTTNHVEVNFIEKFTSERDFHPSMSCSITWFLSWSPCWECQAIREFLSRHPGV
 TLVIYVARLFWHMDQQNRQGLRDLVNSGVTIQIMRASEYYHCWRNFVNYPGDEAHW
 PQYPPLWMMLYALELHCILSLPPCLKISRRWQNHLTFFRLHLQNCYQTIPPHILLATGL
 IHPSVAWR (配列番号 71)

【0448】

マウスAPOBEC-1 :

【0449】

【化57】

MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELKTCCLYEINWGGRHSV
 WRHTSQNTSNHVEVNFLEKFTTERYFRPNTRCSITWFLSWSPCGECSRAITEFLSRHPYV
 TLFYIARLYHHTDQRNRQGLRDLISSGVTIQIMTEQEYCYCWRNFVNYPSPNEAYWPR
 YPHLWVKLYVLELYCIIILGLPPCLKILRRKQPQLTFFTITLQTCHYQRIPPHLLWATGLK
 (配列番号 72)

【0450】

ラットAPOBEC-1 :

【0451】

【化58】

MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELKTCCLYEINWGGRHSIW
 RHTSQNTNKHVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSRAITEFLSRYPHVTL
 FIYIARLYHHADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFVNYSNEAHWPRYPH
 LWVRLYVLELYCIIILGLPPCLNILRRKQPQLTFFTIALQSCHYQRLPPHILWATGLK (

配列番号 73)

【0452】

ヒトAPOBEC-2 :

【0453】

【化59】

MAQKEEAAVATEAASQNGEDLENLDDPEKCLKELIELPPFEIVTGERLPANF
 FKFQFRNVEYSSGRNKTFLCYVVEAQGGQVQASRGYLEDEHAAAHAEAEFFNTILP
 AFDPALRYNVTWYVSSSPCAACADRIIKTLSKTKNLRLILVGRLFMWEEPEIQAAALKKL
 KEAGCKLRIMKPQDFEYVWQNFVEQEEGESKAFQPWEDIQENFLYYEEKLADILK

(配列番号 74)

【0454】

マウスAPOBEC-2 :

【0455】

10

20

30

40

【化60】

MAQKEEAAEAAAPASQNGDDLENLEDPEKCLKELIDLPPFEIVTGVRLPVN
 FFKFQFRNVEYSSGRNKTFLCYVVEVQSKGGQAQATQGYLEDEHAGAHAEAEAFFNTILP
 AFDPALKYNVTWYVSSSPCAACADRILKTLKTKNLRLLILVSRLFMWEEPEVQAALKK
 LKEAGCKLRIMKPQDFEYIWQNFVEQEEGESKAFEPWEDIQENFLYYEEKLADILK

(配列番号 75)

【0456】

ラットAPOBEC-2 :

10

【0457】

【化61】

MAQKEEAAEAAAPASQNGDDLENLEDPEKCLKELIDLPPFEIVTGVRLPVN
 FFKFQFRNVEYSSGRNKTFLCYVVEAQSKGGQVQATQGYLEDEHAGAHAEAEAFFNTILP
 AFDPALKYNVTWYVSSSPCAACADRILKTLKTKNLRLLILVSRLFMWEEPEVQAALKK
 LKEAGCKLRIMKPQDFEYLWQNFVEQEEGESKAFEPWEDIQENFLYYEEKLADILK

(配列番号 76)

20

【0458】

ウシAPOBEC-2 :

【0459】

【化62】

MAQKEEAAAAAEPASQNGEEVENLEDPEKCLKELIELPPFEIVTGERLPAHY
 FKFQFRNVEYSSGRNKTFLCYVVEAQSKGGQVQASRGYLEDEHATNHAEAEAFFNSIMPT
 FDPALRYMVTWYVSSSPCAACADRIVKTLNKTKNLRLLILVGRLFMWEEPEIQAALRKL
 KEAGCRLRIMKPQDFEYIWQNFVEQEEGESKAFEPWEDIQENFLYYEEKLADILK (

30

配列番号77)

【0460】

Petromyzon marinus CDA1 (pmCDA1)

【0461】

【化63】

MTDAEYVRIHEKLDIYTFKKQFFNNKKS VSHRCYVLFELKRRGERRACF
 WGYAVNKPQSGTERGIHAEIFSIRKVEEYLRDNPQGFTINWYSSWSPCADCAEKILEWY
 NQELRGNGHTLKIWACKLYYEKNARNQIGLWNLRDNGVGLNVMVSEHYQCCRKIFIQS
 SHNQLNENRWLEKTLKRAEKRRSELSIMI QVKILHTTKSPAV (配列番号 78)

40

【0462】

ヒトAPOBEC3G D316R_D317R

【0463】

【化64】

MKPHFRNTVERMYRDTFSYNFYNRPILSRRNTVWLCYEVKTKGPSRPL
 DAKIFRGQVYSELKYHPEMRFFHWFSKWRKLHRDQEYEVWYISWSPCTKCTRDMAT
 FLAEDPKVTLTIFVARLYYFWDPDYQEALRSLCQKRDGPRATMKIMNYDEFQHCWSKF
 VYSQRELFEPWNNLPKYIILLHIMLGEILRHSMDPPTFTFNFNNEPWVRGRHETLYLCYE
 VERMHNDTWVLLNQRRGFLCNQAPHKHGFLEGRHAELCFLDVIPFWKLDLDQDYRV
 TCFTSWSPCFSCAQEMAKFISKKNHVS LCIFTARIYRRQGRQCQGLR TLAEAGAKISIMTY
 SEFKHCWDTFVDHQQGCPFPWDGLDEHSQDLSGRLRAILQNQEN (配列番号 79)

10

【0464】

ヒトAPOBEC3G A鎖

【0465】

【化65】

MDPPTFTFNFNNEPWVRGRHETLYLCYEVERMHNDTWVLLNQRRGFLCN
 QAPHKHGFLEGRHAELCFLDVIPFWKLDLDQDYRVTCFTSWSPCFSCAQEMAKFISKKN
 HVS LCIFTARIYDDQGRQCQGLR TLAEAGAKISIMTYSEFKHCWDTFVDHQQGCPFPWD
 GLDEHSQDLSGRLRAILQ (配列番号 80)

20

【0466】

ヒトAPOBEC3G A鎖 D120R_D121R

【0467】

【化66】

MDPPTFTFNFNNEPWVRGRHETLYLCYEVERMHNDTWVLLNQRRGFLCN
 QAPHKHGFLEGRHAELCFLDVIPFWKLDLDQDYRVTCFTSWSPCFSCAQEMAKFISKKN
 HVS LCIFTARIYRRQGRQCQGLR TLAEAGAKISIMTYSEFKHCWDTFVDHQQGCPFPWD
 GLDEHSQDLSGRLRAILQ (配列番号 81)

30

【0468】

先述のDNAエフェクタードメインのいずれかは本明細書に記載される連続的進化プロセス(例えばPACE)に付され得る。

【0469】

本開示のいくつかの側面はシチジンデアミナーゼを提供し、これらのいずれかは本明細書に記載される連続的進化プロセス(例えばPACE)に付され得る。

【0470】

いくつかの態様では、第2の蛋白質は核酸編集ドメインを含む。いくつかの態様では、核酸編集ドメインはC U塩基変化を触媒し得る。いくつかの態様では、核酸編集ドメインはデアミナーゼドメインである。いくつかの態様では、デアミナーゼはシチジンデアミナーゼまたはシチジンデアミナーゼである。いくつかの態様では、デアミナーゼはアポリボ蛋白質B mRNA編集複合体(APOBEC)ファミリーデアミナーゼである。いくつかの態様では、デアミナーゼはAPOBEC1デアミナーゼである。いくつかの態様では、デアミナーゼはAPOBEC2デアミナーゼである。いくつかの態様では、デアミナーゼはAPOBEC3デアミナーゼである。いくつかの態様では、デアミナーゼはAPOBEC3Aデアミナーゼである。いくつかの態様では、デアミナーゼはAPOBEC3Bデアミナーゼである。いくつかの態様では、デアミナーゼはAPOBEC3Cデアミナーゼである。いくつかの態様では、デアミナーゼはAPOBEC3Dデアミナーゼである。いくつかの態様では、デアミナーゼはAPOBEC3Eデアミナーゼである。いくつかの態様では、デアミナーゼはAPOBEC3Fデアミナーゼである。いくつかの態様では、デ

40

50

アミナーゼはAPOBEC3Gデアミナーゼである。いくつかの態様では、デアミナーゼはAPOBEC3Hデアミナーゼである。いくつかの態様では、デアミナーゼはAPOBEC4デアミナーゼである。いくつかの態様では、デアミナーゼは活性化誘導デアミナーゼ(AID)である。いくつかの態様では、デアミナーゼは脊椎動物デアミナーゼである。いくつかの態様では、デアミナーゼは無脊椎動物デアミナーゼである。いくつかの態様では、デアミナーゼはヒト、チンパンジー、ゴリラ、サル、乳牛、イヌ、ラット、またはマウスデアミナーゼである。いくつかの態様では、デアミナーゼはヒトデアミナーゼである。いくつかの態様では、デアミナーゼはラットデアミナーゼ、例えばrAPOBEC1である。

【0471】

いくつかの態様では、核酸編集ドメインは、上で開示されているデアミナーゼ配列のいずれか1つのデアミナーゼドメインと少なくとも80%、少なくとも85%、少なくとも90%、少なくとも92%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、または少なくとも99.5%同一である。

10

【0472】

本開示のいくつかの側面は、例えばデアミナーゼドメインに点変異を作ることによって、本明細書において提供される融合蛋白質のいずれかのデアミナーゼドメイン触媒活性を調節することが、融合蛋白質(例えば塩基編集因子)の処理性能を左右するという認識に基づく。例えば、塩基編集融合蛋白質内のデアミナーゼドメインの触媒活性を縮減するが抹消はしない変異は、デアミナーゼドメインが標的残基に隣接する残基の脱アミノ化を触媒するであろうということをより蓋然的でなくし得、それによって脱アミノ化ウィンドウを狭める。脱アミノ化ウィンドウを狭める能力は、特異的な標的残基に隣接する残基の不要の脱アミノ化を防止し得、これはオフターゲット効果を減少させ得るかまたは防止し得る。

20

【0473】

いくつかの態様では、本明細書において提供される融合蛋白質のいずれかは、縮減された触媒デアミナーゼ活性を有するデアミナーゼドメイン(例えばシチジンデアミナーゼドメイン)を含む。いくつかの態様では、本明細書において提供される融合蛋白質のいずれかは、適切なコントロールと比較して縮減されたデアミナーゼ触媒活性を有するデアミナーゼドメイン(例えばシチジンデアミナーゼドメイン)を含む。例えば、適切なコントロールは、1つ以上の変異をデアミナーゼに導入することに先立つデアミナーゼのデアミナーゼ活性であり得る。他の態様では、適切なコントロールは野生型デアミナーゼであり得る。いくつかの態様では、適切なコントロールは野生型アポリポ蛋白質B mRNA編集複合体(APOBEC)ファミリーデアミナーゼである。いくつかの態様では、適切なコントロールはAPOBEC1デアミナーゼ、APOBEC2デアミナーゼ、APOBEC3Aデアミナーゼ、APOBEC3Bデアミナーゼ、APOBEC3Cデアミナーゼ、APOBEC3Dデアミナーゼ、APOBEC3Fデアミナーゼ、APOBEC3Gデアミナーゼ、またはAPOBEC3Hデアミナーゼである。いくつかの態様では、適切なコントロールは活性化誘導デアミナーゼ(AID)である。いくつかの態様では、適切なコントロールはPetromyzon marinmからのシチジンデアミナーゼ1(pmCDA1)である。いくつかの態様では、デアミナーゼ(deaminase)ドメインは、適切なコントロールと比較して少なくとも1%、少なくとも5%、少なくとも15%、少なくとも20%、少なくとも25%、少なくとも30%、少なくとも40%、少なくとも50%、少なくとも60%、少なくとも(at least)70%、少なくとも80%、少なくとも90%、または少なくとも95%少ない触媒デアミナーゼ活性を有するデアミナーゼドメインであり得る。

30

40

【0474】

シチジンデアミナーゼ酵素のアポリポ蛋白質B mRNA編集複合体(APOBEC)ファミリーは、コントロールされたかつ有益な様式で変異導入を開始する用をなす11個の蛋白質を包摂する。1つのファミリー構成員、活性化誘導シチジンデアミナーゼ(AID)は、転写依存的な鎖偏向的な様式でssDNA中のシトシンをウラシルに変換することによって抗体の成熟を担う。アポリポ蛋白質B編集複合体3(APOBEC3)酵素は、逆転写されたウイルスssDNA中のシトシンの脱アミノ化によってある種のHIV-1株からのヒト細胞の保護を提供する。こ

50

これらの蛋白質は全て触媒活性のためにはZn²⁺配位モチーフ (His-X-Glu-X₂₃₋₂₆-Pro-Cys-X₂₋₄-Cys ; 配列番号82) および結合した水分子を要求する。Glu残基は、脱アミノ化反応における求核攻撃のために水分子を水酸化亜鉛へと活性化するように作用する。各ファミリー構成員は、hAIDのWRC (WはAまたはTであり、RはAまたはGである) からhAPOBEC3FのTTCまでの範囲であるそれ自体の特定の「ホットスポット」において選好的に脱アミノ化する。APOBEC3Gの触媒ドメインの最近の結晶構造は、6つのヘリックスによってフランキングされる5鎖シートコアを含む二次構造を解明した。これはファミリー全体で保存されていると信じられている。活性中心ループはssDNA結合と「ホットスポット」のアイデンティティを決定することとを両方担うことが示されている。これらの酵素の過剰発現はゲノム不安定性および癌に連結されており、それゆえに配列特異的な標的化の重要性を強調している。

10

【0475】

本開示のいくつかの側面は、APOBEC酵素などのシチジンデアミナーゼ酵素の活性がゲノムDNA中の特異的な部位へと導かれ得るという認識に関する。いずれかの特定の理論によって拘束されることを欲することなしに、Cas9を認識薬剤として用いる利点は、(1) Cas9の配列特異性が単純にsgRNA配列を変化させることによって容易に変調され得るということと；(2) Cas9がdsDNAを変性させることによってその標的配列に結合し、一本鎖、よってデアミナーゼの現実的な基質であるDNAのストレッチをもたらすということとを包含する。他の触媒ドメイン、または他のデアミナーゼからの触媒ドメインもまたCas9との融合蛋白質を生成するために用いられ得るということと、本開示はこれについて限定されない

20

【0476】

本開示のいくつかの側面は、Cas9:デアミナーゼ融合蛋白質がヌクレオチドを効率的に脱アミノ化し得るという認識に基づく。Cas9:デアミナーゼ融合蛋白質によって標的化され得るヌクレオチドについて本明細書において提供される結果に鑑みて、当業者は、融合蛋白質を脱アミノ化されるべきヌクレオチドを含む標的配列へと標的化するための好適なガイドRNAを設計することができるであろう。

【0477】

ある態様では、参照シチジンデアミナーゼドメインは、次の通り、配列番号1に従うアミノ酸配列、または配列番号1と少なくとも60%、65%、70%、75%、80%、85%、90%、95%、98%、99%、もしくは99.5%同一であるアミノ酸配列を有する「FERNY」ポリペプチドを含む：

30

【0478】

【化67】

MFERNYDPRELRKETYLLYEIKWKGSGKLWRHWCQNNRTQHAEVYFLE

NIFNARRFNPSTHCSITWYLSWSPCAECSQKIVDFLKEHPNVNLEIYVARLYYHEDERNR

QGLRDLVNSGVTIRIMDLDPDYNWCWKTFSVSDQGGDEDYWPGHFAPWIKQYSLKL (

配列番号1)

【0479】

ある他の態様では、進化したシチジンデアミナーゼドメイン (すなわち、本明細書に記載される連続的進化プロセスの結果として) は、次の通り、配列番号5に従うアミノ酸配列、または配列番号5と少なくとも60%、65%、70%、75%、80%、85%、90%、95%、98%、99%、もしくは99.5%同一であるアミノ酸配列を有する「evoFERNY」ポリペプチドを含み、H102PおよびD104N置換を含む：

40

【0480】

【化68】

MFERNYDPRELRLKETYLLYEIKWGKSGKLWRHWCQNNRTQHAEVYFLE
NIFNARRFNPSTHCSITWYLSWSPCAECSQKIVDFLKEHPNVNLEIYVARLYYPENERNR
QGLRDLVNSGVITIRIMDLDPDYNWCWKTFSVSDQGGDEDYWPGHFAPWIKQYSLKL (

配列番号 5)

【0481】

他の態様では、参照シチジンデアミナーゼドメインは、次の通り、配列番号2に従うアミノ酸配列、または配列番号2と少なくとも60%、65%、70%、75%、80%、85%、90%、95、98%、99%、もしくは99.5%同一であるアミノ酸配列を有する「ラットAPOBEC-1」ポリペプチドを含む：

10

【0482】

【化69】

MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELRLKETCLLYEINWGGRHSIW
RHTSQNTNKHVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSRAITEFLSRYPHVTL
FIYIARLYHHADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFVNYSNEAHWPRYPH
LWVRLYVLELYCIILGLPPCLNILRRKQPQLTFFTIALQSCHYQRLPPHILWATGLK (

20

配列番号 2)

【0483】

ある他の態様では、進化したシチジンデアミナーゼドメイン（すなわち、本明細書に記載される連続的進化プロセスの結果として）は、次の通り、配列番号6に従うアミノ酸配列、または配列番号6と少なくとも60%、65%、70%、75%、80%、85%、90%、95、98%、99%、もしくは99.5%同一であるアミノ酸配列を有し、かつ置換E4K；H109N；H122L；D124N；R154H；A165S；P201S；F205Sを含む「evoAPOBEC」ポリペプチドを含む：

【0484】

【化70】

MSSKTGPVAVDPTLRRRIEPHEFEVFFDPRELRLKETCLLYEINWGGRHSIW
RHTSQNTNKHVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSRAITEFLSRYPNVTL
FIYIARLYHLANPRNRQGLRDLISSGVTIQIMTEQESGYCWHNFVNYSNESHWPRYPH
LWVRLYVLELYCIILGLPPCLNILRRKQSQLTSFTIALQSCHYQRLPPHILWATGLK (

30

配列番号 6)

【0485】

なお他の態様では、参照シチジンデアミナーゼドメインは、次の通り、配列番号3に従うアミノ酸配列、または配列番号3と少なくとも60%、65%、70%、75%、80%、85%、90%、95、98%、99%、もしくは99.5%同一であるアミノ酸配列を有する「Petromyzon marinus CDA1 (pmCDA1)」ポリペプチドを含む：

40

【0486】

【化71】

MTDAEYVRIHEKLDIYTFKKQFFNKKSVSHRCYVLFELKRRGERRACF
WGYAVNKPQSGTERGIHAEIFSIRKVEEYLRDNPQGFTINWYSSWSPCADCAEKILEWY
NQELRGNGHTLKIWACKLYYEKNARNQIGLWNLRDNGVGLNVMVSEHYQCCRKIFIQS
SHNQLNENRWLEKTLKRAEKRRSELSIMIQVKILHTTKSPAV (配列番号 3)

50

【0487】

他の態様では、進化したシチジンデアミナーゼドメイン（すなわち、本明細書に記載される連続的進化プロセスの結果として）は、次の通り、配列番号7に従うアミノ酸配列、または配列番号7と少なくとも60%、65%、70%、75%、80%、85%、90%、95%、98%、99%、もしくは99.5%同一であるアミノ酸配列を有し、かつ置換F23S；A123V；I195Fを含む「evoCDA」ポリペプチドを含む：

【0488】

【化72】

MTDAEYVRIHEKLDIYTFKKQFSNNKKS^UVSHRCYVLFELKRRGERRACF
 WGYAVNKPQSGTERGIHAEIFSIRKVEEYLRDNPQGFTINWYSSWSPCADCAEKILEWY
 NQELRGNGHTLKIWV^YCKLYYEKNARNQIGLWNLRDNGVGLNVMVSEHYQCCRKIFIQS
 SHNQLNENRWLEKTLKRAEKRRSELSIM^FEQVKILHTTKSPAV (配列番号 7)

10

【0489】

まだ他の態様では、参照シチジンデアミナーゼドメインは、次の通り、配列番号4に従うアミノ酸配列、または配列番号4と少なくとも60%、65%、70%、75%、80%、85%、90%、95%、98%、99%、もしくは99.5%同一であるアミノ酸配列を有する「Anc689 APOBEC」ポリペプチドを含む：

20

【0490】

【化73】

MSSETGPVAVDPTLRRRIEPHEFEVFFDPREL^RRKETCLLYEIKWGTSHKIW
 RHSSKNTTKHVEVNFIEKFTSERHFPCSTSCSITWFLSWSPCGECSKAITEFLSQHPNVTL
 VIYVARLYHHMDQQNRQGLRDLVNSGVTIQIMTAPEYDYCWRNFVNYPPGKEAHWPR
 YPPLWMKLYALELHAGILGLPPCLNILRRKQPQLTFFTIALQSCHYQRLPPHILWATGLK
 (配列番号 4)

30

【0491】

他の態様では、進化したシチジンデアミナーゼドメイン（すなわち、本明細書に記載される連続的進化プロセスの結果として）は、次の通り、配列番号8に従うアミノ酸配列、または配列番号8と少なくとも60%、65%、70%、75%、80%、85%、90%、95%、98%、99%、もしくは99.5%同一であるアミノ酸配列を有し、かつ置換E4K；H122L；D124N；R154H；A165S；P201S；F205Sを含む「evoAnc689 APOBEC」ポリペプチドを含む：

【0492】

【化74】

MSS^KTGPVAVDPTLRRRIEPHEFEVFFDPREL^RRKETCLLYEIKWGTSHKIW
 RHSSKNTTKHVEVNFIEKFTSERHFPCSTSCSITWFLSWSPCGECSKAITEFLSQHPNVTL
 VIYVARLYH^LM^NQQNRQGLRDLVNSGVTIQIMTAPEYDYCW^HNFVNYPPGKESH^SHWPR
 YPPLWMKLYALELHAGILGLPPCLNILRRKQ^SQLT^SFTIALQSCHYQRLPPHILWATGLK
 (配列番号 8)

40

【0493】

いくつかの側面では、本明細書は進化したシチジンデアミナーゼを提供し、これらは改善された特性を有する塩基編集因子を構築するために用いられる。例えば、進化したシチジンデアミナーゼ、例えば本明細書において提供されるものは、周囲の配列にかかわらず

50

塩基をより効率的に編集するために、塩基編集効率を改善および/または塩基編集因子の能力を改善する能力がある。例えば、いくつかの側面では、本開示は進化したAPOBECデアミナーゼ（例えば進化したAPOBEC1）を提供し、5'-G-3'のコンテキストにおいて、それが標的塩基（例えばC）の5'であるときに、改善された塩基編集効率を有する。いくつかの態様では、本開示は、本明細書において提供される進化したシチジンデアミナーゼのいずれかを含む塩基編集因子を提供する。本明細書において提供される進化したシチジン（cytidine）デアミナーゼのいずれかが、塩基編集因子蛋白質、例えば本明細書において提供される塩基編集因子のいずれかにデアミナーゼとして用いられ得るということは了解されるはずである。本開示は、本明細書において提供される変異のいずれか、例えば実施例セクションに記載される変異のいずれかを有するシチジンデアミナーゼを企図するという

10

【0494】

(C) UGIドメイン

他の態様では、本明細書に記載される塩基編集因子は1つ以上のウラシルグリコシラーゼ阻害因子を含み得、これらは任意に本明細書に記載される連続的進化プロセス（例えばPACE）を用いて進化させられ得る。

【0495】

本明細書において用いられる用語「ウラシルグリコシラーゼ阻害因子」または「UGI」は、ウラシルDNAグリコシラーゼ塩基除去修復酵素を阻害する能力がある蛋白質を言う。いくつかの態様では、UGIドメインは野生型UGIまたは配列番号10に提示されているUGIを含む。いくつかの態様では、本明細書において提供されるUGI蛋白質は、UGIの断片、およびUGIまたはUGI断片に対して相同的な蛋白質を包含する。例えば、いくつかの態様では、UGIドメインは配列番号10に提示されているアミノ酸配列の断片を含む。いくつかの態様では、UGI断片は、配列番号10に提示されているアミノ酸配列の少なくとも60%、少なくとも65%、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、または少なくとも99.5%を含むアミノ酸配列を含む。いくつかの態様では、UGIは、配列番号10に提示されているアミノ酸配列に対して相同的なアミノ酸配列、または配列番号10に提示されているアミノ酸配列の断片に対して相同的なアミノ酸配列を含む。いくつかの態様では、UGIもしくはUGIの断片またはUGIもしくはUGI断片のホモログを含む蛋白質は「UGI

20

30

40

【0496】

> sp|P14739|UNGI_BPPB2ウラシルDNAグリコシラーゼ阻害因子

【化75】

MTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLTSD

APEYKPWALVIQDSNGENKIKML (配列番号 10).

【0497】

本明細書に記載される塩基編集因子は1つよりも多くのUGIドメインを含み得、これらは本明細書に記載される1つ以上のリンカーによって分離され得る。

50

【0498】

(D) 分割型インテインドメイン

本明細書に記載される種々の態様では、連続的進化方法（例えばPACE）は塩基編集因子の第1の部分を進化させるために用いられ得る。第1の部分は、1個の構成要素またはドメイン、例えばCas9ドメイン、デアミナーゼドメイン、またはUGIドメインを包含し得る。それから、別個に進化した構成要素またはドメインは、分割型インテインポリペプチドドメインを有する進化した部分および残りの未進化の部分両方を別個に発現すること（express）によって、細胞内で塩基編集因子の残りの部分に融合させられ得る。より広くは、第1の部分は、本明細書に記載される連続的進化方法を用いて進化させられることが所望である塩基編集因子のいずれかの第1のアミノ酸部分を包含し得る。この態様では、第2の部分は、本明細書の方法を用いて進化させられない塩基編集因子の残りのアミノ酸部分を言うであろう。塩基編集因子の進化した第1の部分および第2の部分はそれぞれ分割型インテインポリペプチドドメインと共に細胞内で発現され得る。細胞の天然の蛋白質スプライシングメカニズムは、進化した第1の部分および進化していない第2の部分を再アセンブリして、1個の融合蛋白質の進化した塩基編集因子を形成するであろう。進化した第1の部分は1個の融合蛋白質のNまたはC末端部分どちらかを含み得る。同様の様式で、第2の直交的なトランススプライシングインテイン対の使用は、進化した第1の部分が1個の融合蛋白質の内部部分を含むことを許し得る。

10

【0499】

それゆえに、本明細書に記載される塩基編集因子の進化したおよび進化していない構成要素のいずれかは、細胞内において、進化したおよび進化していない構成要素を含む完全な塩基編集因子の形成を助長するために、分割型インテインタグと共に発現され得る。

20

【0500】

蛋白質スプライシングプロセスのメカニズムは多大に詳細に研究されており（Chong, et al., J. Biol. Chem. 1996, 271, 22159-22168 ; Xu, M-Q & Perler, F. B. EMBO Journal, 1996, 15, 5146-5153）、保存されたアミノ酸がインテインおよびエクステインスプライシング点に見出されている（Xu, et al., EMBO Journal, 1994, 13 5517-522）。本明細書に記載されるコンストラクトは、第1の遺伝子（例えば、塩基編集因子の進化した部分）の5'末端に融合させられたインテイン配列を含有する。好適なインテイン配列は、蛋白質スプライシングエレメントを含有することが公知の蛋白質のいずれかから選択され得る。全ての公知のインテインを含有するデータベースがワールドワイドウェブ上に見出され得る（Perler, F. B. Nucleic Acids Research, 1999, 27, 346-347）。インテイン配列は3'末端において第2の遺伝子の5'末端に融合させられる。この遺伝子のある種のオルガネラへと標的化するためには、ペプチドシグナルが遺伝子のコード配列に融合させられ得る。同じ細胞による複数の蛋白質の発現のためには、第2の遺伝子の後に、インテイン-遺伝子配列が所望なだけ何度も繰り返され得る。マルチインテイン含有コンストラクトでは、異なるソースからのインテインエレメントを用いることが有用であり得る。発現されるべき最後の遺伝子の配列の後には、転写終結配列が挿入されなければならない。1つの態様では、それがインテインからのエクステインの切り出しを触媒、およびエクステインのライゲーションを防止両方をし得るように、改変されたインテインスプライシング単位が設計される。Pyrococcus species GB-D DNAポリメラーゼのC末端エクステインジャンクションの変異導入は、エクステインおよびインテインの切断を誘導するがエクステインの爾後のライゲーションを防止する変調したスプライシングエレメントを産生することが見出された（Xu, M-Q & Perler, F. B. EMBO Journal, 1996, 15, 5146-5153）。セリン538からアラニンまたはグリシンどちらかへの変異は、切断を誘導するがライゲーションを防止した。インテインへのC末端エクステインジャンクションにおけるアミノ酸の保存が原因で、他のインテインスプライシング単位における同等な残基の変異もまたエクステインライゲーションを防止するはずである。エンドヌクレアーゼドメインを含有しない好ましいインテインはMycobacterium xenopi GyrA蛋白質である（Telenti, et al. J. Bacteriol. 1997, 179, 6378-6382）。他のものは天然に見出されているか、またはエンド

30

40

50

ヌクレアーゼ含有インテインからエンドヌクレアーゼドメインを取り除くことによって人工的に作出されている (Chong, et al. J. Biol. Chem. 1997, 272, 15587-15590)。好ましい態様では、インテインは、それが、スプライシング機能を実行するために必要とされる最小限の数のアミノ酸からなるように選択され、例えば *Mycobacterium xenopi* GyrA 蛋白質からのインテインである (Telenti, A., et al., J. Bacteriol. 1997, 179, 6378-6382)。代替的な態様では、エンドヌクレアーゼドメインを取り除くように改変された *Saccharomyces cerevisiae* VMAインテインまたは *Mycobacterium xenopi* GyrA蛋白質からのインテインなどの、エンドヌクレアーゼ活性なしのインテインが選択される (Chong, 1997)。インテインスプライシング単位のさらなる改変は、切断反応の反応速度が変調されることを許し得、蛋白質用量が単純にスプライシング単位の遺伝子配列を改変することによってコントロールされることを許す。

10

【0501】

インテインは、2つの別個に転写および翻訳される遺伝子によってコードされる2つの断片としてもまた存在し得る。これらのいわゆる分割型インテインは自己会合し、トランスに蛋白質スプライシング活性を触媒する。分割型インテインは多様なシアノバクテリアおよび古細菌において同定されているが (Caspi et al, Mol Microbiol. 50 : 1569-1577 (2003) ; Choi J. et al, J Mol Biol. 556 : 1093-1106 (2006.) ; Dassa B. et al, Biochemistry. 46 : 322-330 (2007.) ; Liu X. and Yang J., J Biol Chem. 275 : 26315-26318 (2003) ; Wu H. et al.

20

【0502】

Proc Natl Acad Sci USA. 95 : 9226-9231 (1998.) ; および Zettler J. et al, FEBS Letters. 553 : 909-914 (2009))、ここまでは真核生物では見出されていない。最近、環境メタゲノミクスデータのパイオインフォマティクス分析が、新規のゲノム構成を有する26個の異なる座位を解明した。各座位では、保存された酵素コード領域に分割型インテインが割り込み、インテインサブドメインをコードするセクション間にフリースタンディングなエンドヌクレアーゼ遺伝子が挿入されている。それらのうち、5つの座位が完全にアセンブリされた : DNAヘリカーゼ (gp41-1, gp41-8) ; イノシン-5'-リン酸デヒドロゲナーゼ (IMPDH-1) ; およびリボヌクレオチドレダクターゼ触媒サブユニット (NrdA-2およびNrdJ-1)。このフラクチャー化した遺伝子編成は主にファージに存在するように見える (Dassa et al, Nucleic Acids Research. 57: 2560-2573 (2009))。

30

【0503】

分割型インテインNpu DnaEは、蛋白質トランススプライシング反応について報告された最も高い速度を有するとキャラクター化されている。加えて、Npu DnaE蛋白質スプライシング反応は、異なるエクステイン配列、6から37の温度、および最高で6M尿素の存在についてロバストかつ高収量だと考慮されている (Zettler J. et al, FEBS Letters. 553: 909-914 (2009) ; Iwai I. et al, FEBS Letters 550: 1853-1858 (2006))。予想された通り、これらのインテインのNドメインのCys1 Ala変異が導入されたときには、最初のN-S-アシルシフト、よって、蛋白質スプライシングはブロックされた。残念なことに、C末端切断反応もまたほぼ完全に阻害された。C末端スプライシングジャンクションにおけるアスパラギン環化の、被切断性のN末端ペプチド結合におけるアシルシフトへの依存は、天然に分割型のDnaEインテインアレルに共通のユニークな特性であるように見える (Zettler J. et al. FEBS Letters. 555: 909-914 (2009))。

40

【0504】

蛋白質スプライシングのメカニズムは典型的には4つのステップを有する [29-30] : 1) 上流のペプチド結合を切断し、Nエクステインとインテインの第1のアミノ酸の側鎖 (CysまたはSer) との間にエステル結合を形成する、インテインN末端におけるN-SまたはN-Oアシルシフト ; 2) NエクステインをインテインC末端に移動し、NエクステインをCエクステインの第1のアミノ酸の側鎖 (Cys, Ser, またはThr) に連結する新たなエステル結合を形成する、エステル交換反応 ; 3) インテインおよびCエクステインの間のペプチド結合を切断するAsn環化 ; ならびに4) NエクステインおよびCエクステインの間のエステル結合を

50

ペプチド結合によって置き換える、S-NまたはO-Nアシルシフト。

【0505】

分割型インテインによって触媒される蛋白質トランススプライシングは、蛋白質ライゲーションのための全面的に酵素的な方法を提供する[31]。分割型インテインは、本質的には、それぞれNインテインおよびCインテインと名付けられた2つのピースへと分割された一続きのインテイン(例えば、ミニインテイン)である。分割型インテインのNインテインおよびCインテインは非共有結合的に会合して活性なインテインを形成し、一続きのインテインがするのと本質的に同じやり方でスプライシング反応を触媒し得る。分割型インテインは天然に見出されており、実験室においてもまた操作されている[31-35]。本明細書において用いられる用語「分割型インテイン」は、N末端およびC末端配列が別個の分子になり、これらがトランススプライシング反応について機能的であるインテインへと非共有結合的に再会合または再構成し得るように、1つ以上のペプチド結合切断がN末端およびC末端アミノ酸配列の間に存在するいずれかのインテインを言う。いずれかの触媒的に活性なインテインまたはその断片が、本発明の方法への使用のための分割型インテインを派生させるために用いられ得る。例えば、1つの側面では、分割型インテインは真核生物インテインに由来し得る。別の側面では、分割型インテインは細菌インテインに由来し得る。別の側面では、分割型インテインは古細菌インテインに由来し得る。好ましくは、そう由来する分割型インテインは、トランススプライシング反応を触媒するために必須のアミノ酸配列のみを所有するであろう。

10

【0506】

本明細書において用いられる「N末端分割型インテイン(In)」は、トランススプライシング反応について機能的であるN末端アミノ酸配列を含むいずれかのインテイン配列を言う。Inは、それゆえに、トランススプライシングが起こるときにスプライシングアウトされる配列をもまた含む。Inは、天然に存在するインテイン配列のN末端部分の改変である配列を含み得る。例えば、Inは、かかる追加のおよび/または変異した残基の包含がInをトランススプライシングにおいて非機能的にしない限り、追加のアミノ酸残基および/または変異した残基を含み得る。好ましくは、追加のおよび/または変異した残基の包含はInのトランススプライシング活性を改善または強化する。

20

【0507】

本明細書において用いられる「C末端分割型インテイン(Ic)」は、トランススプライシング反応について機能的であるC末端アミノ酸配列を含むいずれかのインテイン配列を言う。1つの側面では、Icは、その少なくとも4アミノ酸がそれが由来したインテインの最後の鎖からである4から7つの一続きのアミノ酸残基を含む。Icは、それゆえに、トランススプライシングが起こるときにスプライシングアウトされる配列をもまた含む。Icは、天然に存在するインテイン配列のC末端部分の改変である配列を含み得る。例えば、Icは、かかる追加のおよび/または変異した残基の包含がInをトランススプライシングにおいて非機能的にしない限り、追加のアミノ酸残基および/または変異した残基を含み得る。好ましくは、追加のおよび/または変異した残基の包含はIcのトランススプライシング活性を改善または強化する。

30

【0508】

本発明のいくつかの態様では、IcまたはInに連結されるペプチドは追加のケミカル部分を含み得、中でも、蛍光基、ビオチン、ポリエチレングリコール(PEG)、アミノ酸アナログ、非天然アミノ酸、リン酸基、グリコシル基、放射性同位体標識、および医薬分子を包含する。他の態様では、Icに連結されるペプチドは1つ以上の化学的に反応性の基を含み得、中でも、ケトン、アルデヒド、Cys残基、およびLys残基を包含する。分割型インテインのNインテインおよびCインテインは、非共有結合的に会合して活性なインテインを形成し、スプライシング反応を触媒し得る。このときには「インテインスプライシングポリペプチド(ISP)」が存在する。本明細書において用いられる「インテインスプライシングポリペプチド(ISP)」は、Ic、In、または両方が分割型インテインから取り除かれるときに残る分割型インテインのアミノ酸配列の部分の言う。ある態様では、InはISPを含

40

50

む。別の態様では、IcはISPを含む。まだ別の態様では、ISPはInにもIcにも共有結合的に連結されていない別個のペプチドである。

【0509】

分割型インテインは、ミニインテインの構造に見出される-12個の保存されたベータ鎖間の構造化されていないループまたは介在アミノ酸配列中に1つ以上の分割部位を操作することによって、一続きのインテインから作出され得る[25-28]。分割の作出が、インテイン、特に構造化されたベータ鎖の構造を、蛋白質スプライシング活性が失われる十分な程度まで破壊しないのであれば、ベータ鎖間の領域内における分割部位の位置には何らかのフレキシビリティが存在し得る。

【0510】

蛋白質トランススプライシングでは、1つの前駆体蛋白質はNエクステイン部分、次にNインテインからなり、別の前駆体蛋白質がCインテイン、次にCエクステイン部分からなり、トランススプライシング反応(一緒になってNおよびCインテインによって触媒される)は2つのインテイン配列を切り出し、2つのエクステイン配列をペプチド結合によって連結する。酵素反応である蛋白質トランススプライシングは蛋白質の非常に低い(例えばマイクロモル)濃度で働き得、生理条件において行われ得る。

【0511】

(E) 追加の塩基編集因子機能性

種々の態様では、本明細書において開示される塩基編集因子1つ以上の、好ましくは少なくとも2つの核局在シグナルをさらに含む。好ましい態様では、塩基編集因子は少なくとも2つのNLSを含む。少なくとも2つのNLSを有する態様では、NLSは同じNLSであり得るか、またはそれらは異なるNLSであり得る。加えて、NLSは塩基編集因子の残りの部分との融合蛋白質の一部として発現され得る。NLS融合の場所は塩基編集因子のN末端、C末端、または配列内であり得る(例えば、コードされるnapR/DNAbp構成要素(例えばCas9)およびDNAエフェクター部分(例えばデアミナーゼ)の間に挿入される)。

【0512】

NLSは当分野におけるいずれかの公知のNLS配列であり得る。NLSは、核局在のためのいずれかの将来発見されるNLSでもまたあり得る。NLSは、いずれかの天然に存在するNLSまたはいずれかの天然に存在しないNLS(例えば、1つ以上の所望の変異を有するNLS)でもまたあり得る。

【0513】

核局在シグナルまたは配列(NLS)は、核輸送による細胞核への搬入のために蛋白質をタグ付けする、指名する、または別様にマークするアミノ酸配列である。典型的には、このシグナルは蛋白質表面に暴露した正荷電のリジンまたはアルギニンの1つ以上の短い配列からなる。異なる核局在蛋白質は同じNLSを共有し得る。NLSは、蛋白質を核から外へと標的化する核外搬出シグナル(NES)の反対の機能を有する。核局在シグナルは細胞の外表面をもまた標的化し得る。それゆえに、1個の核局在シグナルは、それが結びつけられている実体を細胞の外表面および細胞の核へと導き得る。かかる配列は、いずれかのサイズおよび組成、例えば25よりも多く、25、15、12、10、8、7、6、5、または4アミノ酸であり得るが、好ましくは、核局在シグナル(NLS)として機能することが公知の少なくとも4から8アミノ酸の配列を含むであろう。

【0514】

用語「核局在配列」または「NLS」は、例えば核輸送による細胞核への蛋白質の搬入を促進するアミノ酸配列を言う。核局在配列は当分野において公知であり、当業者には明らかであろう。例えば、NLS配列は、2001年5月31日にWO/2001/038547として公開された2000年11月23日出願のPlank et al., 国際PCT出願PCT/EP2000/011690に記載されており、これの内容は、例示的な核局在配列のそれらの開示について参照によって本明細書に組み込まれる。いくつかの態様では、NLSはアミノ酸配列PKKKRKV(配列番号83)、MDSLLMNRKFLYQFKNVRWAKGRRETYLC(配列番号84)、KRTADGSEFESPKKKRKV(配列番号101)、またはKRTADGSEFESPKKKRKV(配列番号13)を含む。

10

20

30

40

50

【0515】

本発明の1つの側面では、塩基編集因子（例えば公知の塩基編集因子、例えばBE1、BE2、BE3、またはBE4）は1つ以上の核局在シグナル（NLS）、好ましくは少なくとも2つのNLSによって修飾され得る。好ましい態様では、塩基編集因子は2つ以上のNLSによって修飾される。本発明は、本発明の時に当分野において公知のいずれかの核局在シグナル、または本出願時の後の技術水準によって同定されるかもしくはさもなければ利用可能にされるいずれかの核局在シグナルの使用を企図する。代表的な核局在シグナルは、配列が発現される細胞の核へと蛋白質を導くペプチド配列である。核局在シグナルは圧倒的に塩基性であり、蛋白質のアミノ酸配列中のほぼどこにでも位置し得、一般的に4アミノ酸（Autieri & Agrawal, (1998) J. Biol. Chem. 273: 14731-37. 参照によって本明細書に組み込まれる）から8アミノ酸の短い配列を含み、典型的にはリジンおよびアルギニン残基がリッチである（Magin et al., (2000) Virology 274: 11-16. 参照によって本明細書に組み込まれる）。核局在シグナルは多くの場合にはプロリン残基を含む。種々の核局在シグナルが同定されており、細胞質から細胞の核への生物学的分子の輸送を成し遂げるために用いられている。例えば、Tinland et al., (1992) Proc. Natl. Acad. Sci. U.S.A. 89: 744 2-46; Moede et al., (1999) FEBS Lett. 461: 229-34を参照する。これは参照によって組み込まれる。移行には現行では核膜孔蛋白質が関わりと考えられている。

10

【0516】

ほとんどのNLSは3つの一般的な群に分類され得る：(i) SV40ラージT抗原NLSによって例示される単節型NLS（PKKKRKV配列番号83）；(ii) 変動する数のスペーサーアミノ酸によって分離された2つの塩基性ドメインからなり、XenopusヌクレオプラスミンNLSによって例示される、双節型モチーフ（KRXXXXXXXXXXXXKKL配列番号102）；および(iii) 非古典的な配列、例えばhnRNP A1蛋白質のM9、インフルエンザウイルス核蛋白質NLS、および酵母Gal4蛋白質NLS（Dingwall and Laskey 1991）。

20

【0517】

核局在シグナルは蛋白質のアミノ酸配列中の種々の点に現れる。NLSは蛋白質のN末端、C末端、および中央領域に同定されている。それゆえに、本明細書は、1つ以上のNLSによって塩基編集因子のC末端、N末端、および内部領域（regain）を修飾され得る塩基編集因子を提供する。構成要素NLS残基として機能しないより長い配列の残基が、核局在シグナルそれ自体に例えば強直的にまたは立体的に干渉しないように選択されるべきである。よって、実のところ、NLSを含む配列の組成の厳しい限定はないが、かかる配列は機能的には長さおよび組成が限定され得る。

30

【0518】

本開示は、1つ以上のNLSを包含するように塩基編集因子を改変するべきいずれかの好適な手段を企図する。1つの側面では、塩基編集因子は、そのN末端またはそのC末端（または両方）において1つ以上のNLSに翻訳的に融合させられている塩基編集因子蛋白質を発現するように、すなわち塩基編集因子-NLS融合コンストラクトを形成するように操作され得る。他の態様では、塩基編集因子をコードするヌクレオチド配列は、コードされる塩基編集因子の内部領域に1つ以上のNLSをコードするリーディングフレームを組み込むように遺伝子改変され得る。加えて、NLSは、塩基編集因子とN末端に、C末端に、または内部に取り付けられたNLSアミノ酸配列との間に、例えば蛋白質の中央領域にコードされる種々のアミノ酸リンカーまたはスペーサー領域を包含し得る。それゆえに、本開示は、塩基編集因子と1つ以上のNLSとを含む融合蛋白質を発現するためのヌクレオチドコンストラクト、ベクター、および宿主細胞をもまた提供する。

40

【0519】

本明細書に記載される進化した塩基編集因子は、1つ以上のリンカー、例えばポリマー、アミノ酸、核酸、多糖、ケミカル、または核酸リンカー要素を介して塩基編集因子に連結されている核局在シグナルをもまた含み得る。本開示の企図される範囲内のリンカーは、いずれかの限定を有することを意図されず、いずれかの好適な型の分子（例えば、ポリマー、アミノ酸、多糖、核酸、脂質、またはいずれかの合成ケミカルリンカー部分）であ

50

り得、塩基編集因子と1つ以上のNLSとの間の結合（例えば、共有結合連結部、水素結合）を形成することを成就するいずれかの好適な戦略によって塩基編集因子につながり得る。

【0520】

本明細書に記載される進化した塩基編集因子は1つ以上の追加の機能性をもまた包含し得る。ある態様では、追加の機能性は塩基修復のエフェクターを包含し得る。

【0521】

ある態様では、本明細書に記載される塩基編集因子は塩基修復の阻害因子を含み得る。用語「塩基修復の阻害因子」または「IBR」は、核酸修復酵素、例えば塩基除去修復酵素の活性を阻害する能力がある蛋白質を言う。いくつかの態様では、IBRはイノシン塩基除去修復の阻害因子である。塩基修復の例示的な阻害因子は、APE1、Endo III、Endo IV、Endo V、Endo VIII、Fpg、hOGG1、hNEIL1、T7 EndoI、T4PDG、UDG、hSMUG1、およびhAAGの阻害因子を包含する。いくつかの態様では、IBRはEndo VまたはhAAGの阻害因子である。いくつかの態様では、IBRは触媒的に不活性なEndoVまたは触媒的に不活性なhAAGである。

10

【0522】

いくつかの態様では、本明細書に記載される塩基編集因子は1つ以上の異種蛋白質ドメイン（例えば、塩基編集因子構成要素に加えて約1、2、3、4、5、6、7、8、9、10以上の、またはより多くのドメイン）を含み得る。塩基編集因子は、いずれかの追加の蛋白質配列、および任意にいずれかの2つのドメイン間のリンカー配列を含み得る。塩基編集因子またはその構成要素（例えば、napR/DNAbp部分、核酸エフェクター部分、またはNLS部分）に融合させられ得る蛋白質ドメインの例は、限定なしに、エピトープタグ、レポーター遺伝子配列、および次の活性の1つ以上を有する蛋白質ドメインを包含する：メチラーゼ活性、デメチラーゼ活性、転写活性化活性、転写抑制活性、転写解除因子活性、ヒストン修飾活性、RNA切断活性、および核酸結合活性。エピトープタグの限定しない例はヒスチジン（His）タグ、V5タグ、FLAGタグ、インフルエンザヘマグルチニン（HA）タグ、Mycタグ、VSV-Gタグ、およびチオレドキシシン（Trx）タグを包含する。レポーター遺伝子の例は、グルタチオン-S-トランスフェラーゼ（GST）、西洋ワサビペルオキシダーゼ（HRP）、クロラムフェニコールアセチルトランスフェラーゼ（CAT）、ベータ-ガラクトシダーゼ、ベータ-グルクロニダーゼ、ルシフェラーゼ、緑色蛍光蛋白質（GFP）、HcRed、DsRed、シアン蛍光蛋白質（CFP）、黄色蛍光蛋白質（YFP）、および青色蛍光蛋白質（BFP）を包含する自家蛍光蛋白質を包含するが、これらに限定されない。塩基編集因子は、DNA分子に結合するかまたは他の細胞の分子に結合する蛋白質または蛋白質の断片をコードする遺伝子配列に融合させられ得、マルトース結合蛋白質（MBP）、Sタグ、LexA DNA結合ドメイン（DBD）融合体、GAL4 DNA結合ドメイン融合体、および単純ヘルペスウイルス（HSV）BP16蛋白質融合体を包含するが、これらに限定されない。塩基編集因子を含む融合蛋白質の一部を形成し得る追加のドメインは、参照によって本明細書に組み込まれるUS20110059502に記載されている。いくつかの態様では、タグ付けされた塩基編集因子が用いられて、標的配列の場所を同定する。

20

30

【0523】

本発明のある側面では、グルタチオン-S-トランスフェラーゼ（GST）、西洋ワサビペルオキシダーゼ（HRP）、クロラムフェニコールアセチルトランスフェラーゼ（CAT）、ベータ-ガラクトシダーゼ、ベータ-グルクロニダーゼ、ルシフェラーゼ、緑色蛍光蛋白質（GFP）、HcRed、DsRed、シアン蛍光蛋白質（CFP）、黄色蛍光蛋白質（YFP）、および青色蛍光蛋白質（BFP）を包含する自家蛍光蛋白質を包含するが、これらに限定されないレポーター遺伝子が、遺伝子産物の発現の変調または改変を測定すべきマーカーとしての用をなす遺伝子産物をコードするように、細胞に導入され得る。本発明のさらなる態様では、遺伝子産物をコードするDNA分子はベクターによって細胞に導入され得る。本発明の好ましい態様では、遺伝子産物はルシフェラーゼである。本発明のさらなる態様では、遺伝子産物の発現が減少させられる。

40

【0524】

50

存在し得る他の例示的な特徴は、局在配列、例えば細胞質局在配列、搬出配列、例えば核外搬出配列、または他の局在配列、および融合蛋白質の可溶化、精製、もしくは検出にとって有用である配列タグである。本明細書において提供される好適な蛋白質タグは、ピオチンカルボキシラーゼキャリア蛋白質 (BCCP) タグ、mycタグ、カルモジュリンタグ、FLAGタグ、ヘマグルチニン (HA) タグ、ヒスチジンタグもしくはHisタグともまた言われるポリヒスチジンタグ、マルトース結合蛋白質 (MBP) タグ、nusタグ、グルタチオン-S-トランスフェラーゼ (GST) タグ、緑色蛍光蛋白質 (GFP) タグ、チオレドキシントグ、Sタグ、Softag (例えば、Softag1、Softag3)、strepタグ、ピオチンリガーゼタグ、FIAsHタグ、V5タグ、およびSBPタグを包含するが、これらに限定されない。追加の好適な配列は当業者には明らかであろう。いくつかの態様では、融合蛋白質は1つ以上のHisタグを含む。

10

【0525】

(F) ガイド配列 (例えばガイドRNA)

【0526】

種々の態様では、進化した塩基編集因子は、(例えば、いずれかの型の共有結合的なまたは非共有結合的な結合を介して) 1つ以上のガイド配列、すなわち塩基編集因子に結びつけられるかまたは結合されるようになり、ガイド配列またはその部分に対する相補性を有する特異的な標的配列へのその局在を導く配列と、複合体化、結合、または別様に結びつけられ得る。ガイド配列の特定の設計側面は、他の因子、例えばPAM配列の場所、標的配列のパーセントG/C含量、微小相同性領域の程度、二次構造などの中でも、目当てのゲノム標的部

20

【0527】

一般的に、ガイド配列は、標的配列とハイブリダイゼーションするように標的ポリヌクレオチド配列との十分な相補性を有するいずれかのポリヌクレオチド配列であり、標的配列へのnapR/DNAbp (例えば、Cas9、Cas9ホモログ、またはCas9バリエーション) の配列特異的結合を導く。いくつかの態様では、ガイド配列およびその対応する標的配列の間の相補性の程度は、好適なアライメントアルゴリズムを用いて最適にアライメントされたときには、約50%、60%、75%、80%、85%、90%、95%、97.5%、99%以上、またはより多くである。最適なアライメントは、配列をアライメントするためのいずれかの好適なアルゴリズムの使用によって決定され得、これの限定しない例はSmith-Watermanアルゴリズム、Needleman-Wunschアルゴリズム、Burrows-Wheeler変換に基づくアルゴリズム (例えば、Burrows Wheeler Aligner)、ClustalW、Clustal X、BLAT、Novoalign (Novocraft Technologies, ELAND (Illumina, サンディエゴ, Calif.)), SOAP (soap.genomics.org.cnで利用可能)、およびMaq (maq.sourceforge.netで利用可能) を包含する。いくつかの態様では、ガイド配列は長さが約5、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、35、40、45、50、75以上の、またはより多くのヌクレオチドである。

30

【0528】

いくつかの態様では、ガイド配列は長さが約75、50、45、40、35、30、25、20、15、12、またはより少数のヌクレオチド未満である。標的配列への塩基編集因子の配列特異的結合を導くガイド配列の能力はいずれかの好適なアッセイによって評価され得る。例えば、試験されるべきガイド配列を包含する塩基編集因子の構成要素が、例えば本明細書において開示される塩基編集因子の構成要素をコードするベクターによるトランスフェクションによって、対応する標的配列を有する宿主細胞に提供され得、次に、本明細書に記載される通り、例えばSurveyorアッセイによる標的配列内の選好的切断の評価である。類似に、標的ポリヌクレオチド配列の切断は、標的配列、試験されるべきガイド配列を包含する塩基編集因子の構成要素、および試験ガイド配列とは異なるコントロールガイド配列を提供することと、標的配列における結合または切断率を試験およびコントロールガイド配列反応間において比較することによって、試験管内で判定され得る。他のアッセイが可能

40

50

であり、当業者には思い浮かぶであろう。

【0529】

ガイド配列はいずれかの標的配列を標的化するように選択され得る。いくつかの態様では、標的配列は細胞のゲノム中の配列である。例示的な標的配列は、標的ゲノム中においてユニークであるものを包含する。例えば、*S. pyogenes* Cas9では、ゲノム中のユニークな標的配列は形態MMMMMMNNNNNNNNNNNNXGGのCas9標的部

位を包含し得、NNNNNNNNNNNNXGG (NはA、G、T、またはCであり；Xはどれでもあり得る)はゲノム中に1個存在する。ゲノム中のユニークな標的配列は形態MMMMMMNNNNNNNNNNNNXGGの*S. pyogenes* Cas9標的部

位を包含し得、NNNNNNNNNNNNXGG (NはA、G、T、またはCであり；Xはどれでもあり得る)はゲノム中に1個存在する。*S. thermophilus* CRISPR1 Cas9では、ゲノム中のユニークな標的

配列は形態MMMMMMNNNNNNNNNNNNXXAGAAWのCas9標的部

位を包含し得、NNNNNNNNNNNNXXAGAAW (NはA、G、T、またはCであり；Xはどれでもあり得；WはAまたはTである)はゲノム中に1個存在する。ゲノム中のユニークな標的配列は形態MMMMMMNNNNNNNNNNNNXXAGAAWの

S. thermophilus CRISPR1 Cas9標的部

位を包含し得、NNNNNNNNNNNNXXAGAAW (NはA、G、T、またはCであり；Xはどれでもあり得；WはAまたはTである)はゲノム中に1個存在する。*S. pyogenes* Cas9では、ゲノム中のユニークな標的配列は形態MMMMMMNNNNNNNNNNNNXGGXGの

Cas9標的部

位を包含し得、NNNNNNNNNNNNXGGXG (NはA、G、T、またはCであり；Xはどれでもあり得る)はゲノム中に1個存在する。ゲノム中のユニークな標的配列は形態MMMMMMNNNNNNNNNNNNXGGXGの

S. pyogenes Cas9標的部

位を包含し得、NNNNNNNNNNNNXGGXG (NはA、G、T、またはCであり；Xはどれでもあり得る)はゲノム中に1個存在する。これらの配列のそれぞれにおいて、「M」はA、G、T、またはCであり得、配列をユニークだと同定する上で考慮されることを必要とせず

10

20

【0530】

いくつかの態様では、ガイド配列は、ガイド配列中の二次構造の程度を縮減するように選択される。二次構造はいずれかの好適なポリヌクレオチド折りたたみアルゴリズムによって決定され得る。いくつかのプログラムは最小Gibbs自由エネルギーを計算することに基づく。1つのかかるアルゴリズムの例は、Zuker and Stiegler (Nucleic Acids Res. 9 (1981), 133-148)によって記載されているmFoldである。別の例の折りたたみアルゴリズムはウィーン大学のInstitute for Theoretical Chemistryで開発されたオンラインウェブサーバRNAfoldであり、セントロイド構造予測アルゴリズムを用いる(例えば、A. R. Gruber et al., 2008, Cell 106 (1): 23-24; およびPA Carr and GM Church, 2009, Nature Biotechnology 27 (12): 1151-62を参照する)。さらなるアルゴリズムはU.S.出願Ser.No.61/836,080; Broadの参照BI-2013/004Aに見出され得る); 参照によって本明細書に組み込まれる。

30

【0531】

一般的に、tracrメイト配列はtracr配列との十分な相補性を有するいずれかの配列を包含して、(1)対応するtracr配列を含有する細胞内における、tracrメイト配列によってフランキングされるガイド配列の切り出し; および(2)標的配列における複合体の形成の1つ以上を促進し、複合体はtracr配列にハイブリダイゼーションしたtracrメイト配列を含む。一般的に、相補性の程度は、2つの配列の短い方の長さに沿ったtracrメイト配列およびtracr配列の最適なアライメントの参照によってである。最適なアライメントはいずれかの好適なアライメントアルゴリズムによって決定され得、二次構造、例えばtracr配列またはtracrメイト配列どちらかの中の自己相補性をさらに説明し得る。いくつかの態様では、2つの短い方の長さに沿ったtracr配列およびtracrメイト配列の間の相補性の程度は、最適にアライメントされたときには、約25%、30%、40%、50%、60%、70%、80%、90%、95%、97.5%、99%以上、またはより高くである。いくつかの態様では、tracr配列は長さが約5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、25、30、40、50以上の、またはより多くのヌクレオチドである。いくつかの態様では、tracr配列およびtracrメイト配列は1個の転写物中に含有され、その結果、2つの間のハイブリダイゼーションは、ヘアピンなどの二次構造を有する転写物を産生する。ヘアピン構造への使用のため

40

50

に好ましいループ形成配列は長さが4ヌクレオチドであり、最も好ましくは配列GAAAを有する。しかしながら、より長いまたはより短いループ配列が代替的な配列として (as may) 用いられ得る。配列は好ましくはヌクレオチドトリプレット (例えばAAA) と追加のヌクレオチド (例えばCまたはG) とを包含する。ループ形成配列の例はCAAAおよびAAAGを包含する。本発明のある態様では、転写物または転写されるポリヌクレオチド配列は少なくとも2つ以上のヘアピンを有する。好ましい態様では、転写物は2、3、4、または5つのヘアピンを有する。本発明のさらなる態様では、転写物は多くても5つのヘアピンを有する。いくつかの態様では、1個の転写物はさらに転写終結配列を包含する; 好ましくは、これはポリT配列、例えば6つのTヌクレオチドである。ガイド配列とtracrメイト配列とtracr配列とを含む1個のポリヌクレオチドのさらなる限定しない例は次の通りであり (5'から3'に列記されている)、「N」はガイド配列の塩基にあたり、小文字の第1のブロックはtracrメイト配列にあたり、小文字の第2のブロックはtracr配列にあたり、最後のポリT配列は転写ターミネーターにあたる: (1) NNNNNNNNggttttgtactctcaagatttaGAAAataaatcttgagaagctacaaagataaggcttcatgccgaaatcaacaccctgtcattttatggcagggtgttttcggtattttaTTTTT (配列番号103); (2) NNNNNNNNNNNNNNNNNNggttttgtactctcaGAAAtgcagaagctacaaagataaggcttcatgccgaaatcaacaccctgtcattttatggcagggtgttttcggtattttaTTTTTT (配列番号104); (3) NNNNNNNNNNNNNNNNNNggttttgtactctcaGAAAtgcagaagctacaaagataaggcttcatgccgaaatcaacaccctgtcattttatggcagggtgtTTTTT (配列番号105); (4) NNNNNNNNNNNNNNNNNNggttttagagctaGAAAtagcaagttaaaataaggctagtcggttatcaacttgaaaaagtgccaccgagtcggtgcTTTTTT (配列番号106); (5) NNNNNNNNNNNNNNNNNNggttttagagctaGAAATAGcaagttaaaataaggctagtccgttatcaacttgaaaaagtgTTTTTTT (配列番号107); および (6) NNNNNNNNNNNNNNNNNNggttttagagctagAAATAGcaagttaaaataaggctagtcggttatcaTTTTTTTTT (配列番号108)。いくつかの態様では、配列 (1) から (3) は *S. thermophilus* CRISPR1からのCas9との組み合わせで用いられる。いくつかの態様では、配列 (4) から (6) は *S. pyogenes*からのCas9との組み合わせで用いられる。いくつかの態様では、tracr配列はtracrメイト配列を含む転写物とは別個の転写物である。

10

20

30

40

【0532】

本明細書において開示されるCas9ドメインとデアミナーゼとを含む融合蛋白質のいずれかを、標的部位、例えば編集されるべき点変異を含む部位へと標的化するためには、典型的には融合蛋白質をガイドRNA、例えばsgRNAと一緒に共発現することが必要であるという

【0533】

いくつかの態様では、ガイドRNAは構造5'-[ガイド配列]-guuuuagagcuagaaaauagcaaguuaaauuuuagguaguccguuaucaacuugaaaaaguggcaccgagucggugcuuuuu-3' (配列番号109) を含み、ガイド配列は標的配列に対して相補的である配列を含む。ガイド配列は典型的には20ヌクレオチドの長さである。Cas9: 核酸編集酵素/ドメイン融合蛋白質を特異的なゲノム標的部位へと標的化するために好適なガイドRNAの配列は、本開示に基づいて当業者には明らかであろう。かかる好適なガイドRNA配列は、典型的には、編集されるべき標的ヌクレオチドの上流または下流の50ヌクレオチド以内の核酸 (nucleic) 配列に対して相補的であるガイド配列を含む。提供される融合蛋白質のいずれかを特異的な標的配列へと標的化するために好適ないくつかの例示的なガイドRNA配列が本明細書において提供される。追加のガイド配列は当分野において周知であり、本明細書に記載される塩基編集因子に用いられ得る。

【0534】

(G) リンカー

ある態様では、リンカーが本発明のペプチドまたはペプチドドメインもしくは部分のいずれかを連結するために用いられ得る (例えば、部分Aが部分Bに共有結合的に連結され、これが部分Cに共有結合的に連結される)。

50

【0535】

上で定義されている通り、本明細書において用いられる用語「リンカー」は、2つの分子または部分、例えばヌクレアーゼの結合ドメインおよび切断ドメインを連結する化学基または分子を言う。いくつかの態様では、リンカーはRNAによってプログラム可能なヌクレアーゼのgRNA結合ドメインとリコンビナーゼの触媒ドメインとをつなぎ合わせる。いくつかの態様では、リンカーはdCas9と塩基編集因子部分（例えば、シチジンまたはアデノシンデアミナーゼ）とをつなぎ合わせる。典型的には、リンカーは2つの基、分子、または他の部分の間に位置するかまたはそれらによってフランキングされ、共有結合的な結合を介してそれぞれのものに接続され、それゆえに2つを接続する。いくつかの態様では、リンカーはアミノ酸または複数のアミノ酸（例えば、ペプチドもしくは蛋白質）である。いくつかの態様では、リンカーは有機分子、基、ポリマー、またはケミカル部分である。いくつかの態様では、リンカーは長さが5-100アミノ酸、例えば、長さが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、30-35、35-40、40-45、45-50、50-60、60-70、70-80、80-90、90-100、100-150、または150-200アミノ酸である。より長いまたはより短いリンカーもまた企図される。

10

【0536】

リンカーは共有結合的な結合ほどに単純であり得るか、またはそれは長さが多くの原子のポリマーリンカーであり得る。ある態様では、リンカーはポリペプチドであるか、またはアミノ酸に基づく。他の態様では、リンカーはペプチド様ではない。ある態様では、リンカーは共有結合的な結合（例えば、炭素-炭素結合、ジスルフィド結合、炭素-ヘテロ原子結合など）である。ある態様では、リンカーはアミド連結部の炭素-窒素結合である。ある態様では、リンカーは、環式または非環式の、置換または無置換の、分枝または非分枝の、脂肪族またはヘテロ脂肪族のリンカーである。ある態様では、リンカーはポリマー性である（例えば、ポリエチレン、ポリエチレングリコール、ポリアミド、ポリエステルなど）。ある態様では、リンカーはアミノアルカン酸の単量体、二量体、またはポリマーを含む。ある態様では、リンカーはアミノアルカン酸を含む（例えば、グリシン、エタン酸、アラニン、ベータ-アラニン、3-アミノプロパン酸、4-アミノブタン酸、5-ペントン酸など）。ある態様では、リンカーはアミノヘキサ酸（Ahx）の単量体、二量体、またはポリマーを含む。ある態様では、リンカーは炭素環式部分（例えばシクロペンタン、シクロヘキサン）に基づく。他の態様では、リンカーはポリエチレングリコール部分（PEG）を含む。他の態様では、リンカーはアミノ酸を含む。ある態様では、リンカーはペプチドを含む。ある態様では、リンカーはアリアルまたはヘテロアリアル部分を含む。ある態様では、リンカーはフェニル環に基づく。リンカーは、ペプチドからリンカーへの求核剤（例えば、チオール、アミノ）の取り付けを助長するように官能化された部分を包含し得る。いずれかの求電子剤がリンカーの一部として用いられ得る。例示的な求電子剤は、活性化エステル、活性化アミド、マイケル受容体、ハロゲン化アルキル、ハロゲン化アリアル、ハロゲン化アシル、およびイソチオシアネートを包含するが、これらに限定されない。

20

30

【0537】

いくつかの他の態様では、リンカーは、アミノ酸配列（GGGGS）_n（配列番号110）、（G）_n（配列番号111）、（EAAAK）_n（配列番号112）、（GGS）_n（配列番号113）、（SGGS）_n（配列番号114）、SGSETPGTSESATPES配列番号115）、（XP）_n（配列番号116）、またはそれらのいずれかの組み合わせを含み、nは独立して1および30の間の整数であり、Xはいずれかのアミノ酸である。いくつかの態様では、リンカーはアミノ酸配列（GGS）_n（配列番号117）を含み、nは1、3、または7である。いくつかの態様では、リンカーはアミノ酸配列SGSETPGTSESATPES（配列番号99）を含む。いくつかの態様では、リンカーはアミノ酸配列SGGSSGGSSGSETPGTSESATPES（配列番号11）を含む。いくつかの態様では、リンカーはアミノ酸配列SGGSSGGSSGGS（配列番号12）を含む。いくつかの態様では、リンカーはアミノ酸配列SGGS（配列番号14）を含む。

40

【0538】

50

いくつかの態様では、融合蛋白質は構造[核酸編集ドメイン]-[任意のリンカー配列]-[dCas9またはCas9ニッカーゼ]-[任意のリンカー配列]-[UGI]を含む。いくつかの態様では、融合蛋白質は、構造[核酸編集ドメイン]-[任意のリンカー配列]-[UGI]-[任意のリンカー配列]-[dCas9もしくはCas9ニッカーゼ]；[UGI]-[任意のリンカー配列]-[核酸編集ドメイン]-[任意のリンカー配列]-[dCas9もしくはCas9ニッカーゼ]；[UGI]-[任意のリンカー配列]-[任意のリンカー配列]-[dCas9もしくはCas9ニッカーゼ]-[任意のリンカー配列]-[核酸編集ドメイン]；[dCas9もしくはCas9ニッカーゼ]-[任意のリンカー配列]-[UGI]-[任意のリンカー配列]-[核酸編集ドメイン]；または[dCas9もしくはCas9ニッカーゼ]-[任意のリンカー配列]-[核酸編集ドメイン]-[任意のリンカー配列]-[UGI]を含む。

10

【0539】

塩基編集因子の連続的進化

【0540】

塩基編集因子の設計の最近の進歩にもかかわらず、塩基編集の効率は幅広く変わる。塩基編集効率を増大させるために、本発明者らは細胞における塩基編集効率を限定する因子を同定しようとした。驚くべきことに、ヒト細胞における発現および核局在は鍵となるボトルネックを編集効率に課すということが本発明者らによって見出された。本発明者らは、コドン使用頻度を最適化することと、改善された核局在配列(NLS)を用いることと、デアミナーゼの祖先再構築を実行することとが、多大に増大した編集効率を有する塩基編集因子をもたらし、多くの場合には未変更のカウンターパートの編集因子と比較して標的ヌクレオチド変換収量を倍化以上させるということ(that by)を発見した。もたらされた塩基編集因子は、実施例において実証されている通り、以前に記載された塩基編集因子よりもかなり効率的に、ヒト疾患に関係する点変異を種々の哺乳類細胞型にインストールすることが示された。これらの方法は、進化した塩基編集因子を提供するために用いられ得、これらは、当分野において公知の塩基編集因子と比較して劇的に改善されている様式で、核酸分子を効率的に編集するために用いられ得る。進化した塩基編集因子は、例えば疾患を引き起こす点変異を修正することによって、核酸分子、例えばゲノムを効率的に編集するために用いられ得る。

20

【0541】

それゆえに、本発明は、種々の側面では、操作の種々のモードによって開示される進化した塩基編集因子を作る方法に関し、これらは、細胞におけるより多大な発現レベルを達成するための塩基編集因子の構成要素の(例えば、デアミナーゼ)コドン最適化および祖先再構築の実行、ならびに細胞核内への発現された塩基編集因子の局在を増大させるための核局在配列(NLS)、好ましくは少なくとも2つのNLSの使用を包含するが、これらに限定されない。

30

【0542】

発現を増大させる

本明細書において企図される塩基編集因子は、コドン最適化および祖先再構築分析によって増大した発現をもたらす変更を包含し得る。

【0543】

いくつかの態様では、塩基編集因子(またはその構成要素)は真核生物細胞などの特定の細胞による発現のためにコドン最適化される。真核生物細胞は、ヒト、マウス、ラット、ウサギ、イヌ、または非ヒト霊長類を包含するが、これらに限定されない哺乳動物などの特定の生物のものであり得るか、またはそれに由来し得る。一般的に、コドン最適化は、固有の配列の少なくとも1つのコドン(例えば、約1、2、3、4、5、10、15、20、25、50以上の、またはより多くのコドン)を、その宿主細胞の遺伝子により頻繁にまたは最も頻繁に用いられているコドンによって置き換え、一方では固有のアミノ酸配列を維持することによって、目当ての宿主細胞における強化された発現のために核酸配列を変更するプロセスを言う。種々の種は特定のアミノ酸のある種のコドンについての特定の偏向を呈示する。コドン偏向(生物間のコドン使用頻度の違い)は多くの場合にはメッセンジャーRNA(mRNA)の翻訳の効率と相関し、これは、翻って、他の事の中でも、翻訳されよう

40

50

とするコドンの特性と特定のトランスファーRNA (tRNA) 分子の利用可能性とに依存すると信じられている。細胞内の選択されるtRNAの圧倒性は、一般的に、ペプチド合成に最も頻繁に用いられるコドンの反映である。従って、遺伝子は、コドン最適化に基づいて所与の生物における最適な遺伝子発現のために逃えられ得る。コドン使用頻度表は例えば「Codon Usage Database」において難なく利用可能であり、これらの表はいくつものやり方で適応させられ得る。Nakamura, Y., et al. "Codon usage tabulated from the international DNA sequence databases: status for the year 2000" Nucl. Acids Res. 28 : 292 (2000)を参照する。特定の配列を特定の宿主細胞による発現のためにコドン最適化するためのコンピュータアルゴリズムもまた利用可能であり、例えばGene Forge (Aptagen ; Jacobus, Pa.) もまた利用可能である。いくつかの態様では、CRISPR酵素をコードする配列中の1つ以上のコドン (例えば、1、2、3、4、5、10、15、20、25、50、またはより多くの、または全てのコドン) は、特定のアミノ酸について最も頻繁に用いられるコドンに対応する。

10

【 0 5 4 4 】

他の態様では、本発明の塩基編集因子は、祖先配列再構築分析の結果として改善された発現 (未変更のまたは技術水準カウンターパートの編集因子と比較して) を有する。祖先配列再構築 (ASR) は、進化的 / 系統的コンテキストにおいて現代の配列を分析して、ある樹の特定のノードの祖先配列を推定するプロセスである。それから、これらの古代の配列は、最も多くの場合には合成され、実験室微生物または細胞株によって組換え的に発現され、それからキャラクタリゼーションされて、絶滅した生体分子の古代の特性を解明する^{2, 3, 4, 5, 6}。このプロセスは分子適応および機能分岐のメカニズムについての膨大な知見を産生した⁷。かかる知見にもかかわらず、ASRの大いなる粗は、一般的に、実装されるアルゴリズムの精度をベンチマーキングする能力がないことである。ASRをベンチマーキングすることは多くの理由で困難である。特に、遺伝子材料は、ほとんどのASR研究を満足させるくらい十分に長い時間スケールでは化石中に保存されず (数百万から数十億年も昔)、サンプルを収集するために時間を遡って旅することはまだ物理的に可能ではない。Cai et al., "Reconstruction of ancestral protein sequences and its applications," BMC Evolutionary Biology 2004, 4: 33およびZakas et al., "Enhancing the pharmaceutical properties of protein drugs by ancestral sequence reconstruction," Nature Biotechnology, 35, pp. 35-37 (2017)の参照がなされ得、これらのそれぞれは参照によって本明細書に組み込まれる。

20

30

【 0 5 4 5 】

祖先状態再構築を実行し得る利用可能な多くのソフトウェアパッケージがある。一般的に、これらのソフトウェアパッケージは、関するフィールドの科学者の努力によって開発および維持されており、フリーソフトウェアライセンスでリリースされている。次のリストは全ての利用可能なパッケージの網羅的な一覧であることは意味されず、異なる強さおよび特徴を有する祖先再構築の方法を実装する広範な種々のパッケージの代表的なサンプルを提供する : PAML (Phylogenetic Analysis by Maximum Likelihood. //abacus.gene.ucl.ac.uk/software/paml.html利用可能)、BEAST (Bayesian evolutionary analysis by sampling trees. //www.beast2.org/wiki/index.php/Main_Pageで利用可能)、およびDiversitree (R. Methods in Ecology and EvolutionのFitzJohn RG, 2012. Diversitree: comparative phylogenetic analyses of diversification)、およびHyPhy (Hypothesis testing using phylogenies. //hyphy.org/w/index.php/Main_Pageで利用可能)。

40

【 0 5 4 6 】

[38] 実施例は、ASRを用いて、本明細書において開示される塩基編集因子の総体的な発現を増大させ、縮減された長さを有する機能的な塩基編集因子を生成するための1つの態様を実証する。

【 0 5 4 7 】

上の記載は、増大した発現およびそれによって増大した (increase) 編集効率を有する塩基編集因子を作ることについて、限定しないことを意味されている。

50

【0548】

核局在を増大させる

1つの側面では、本明細書は、例えばN末端またはC末端融合蛋白質として1つ以上の核局在シグナル(NLS)をその中に組み込むことによって、塩基編集因子を改善するための戦略を提供する。好ましくは、少なくとも2つのNLSが塩基編集因子に組み込まれる。実施例では、本発明者らは、最適とは言えない核局在が不良な編集効率の(or)根拠であり得るかどうかを調査した。本発明者らは、SV40 NLSまたは双節型NLS(bpNLS)どちらかへのNおよび/またはC末端融合体として、塩基編集因子「BE4」の6つの組み合わせを試験する。実施例に示されている通り、1または2つのbpNLSを用いる全てのバリエーションは編集効率の改善を示した。NおよびC末端両方におけるbpNLSの存在(本明細書では以後「ビスbpNLS」と言われる)は最も良い成績であり、5つの例示的な試験されたゲノム座位において、BE4によって媒介されるC・G → T・A編集効率の1.3倍の平均の改善をもたらした(BE4に用いられるC末端SV40 NLSの37+5.6%と比較して48+8.0%の平均編集)。これらの結果は、一緒になって、1つ以上の核局在シグナル、例えばビスbpNLSによって塩基編集因子を修飾することが、BE3およびBE4などの公知の塩基編集因子について以前に記載された編集効率を有意に改善し得るということを示唆している(6,7)。

10

【0549】

しかしながら、実施例は、限定的ではなく、1つ以上の核局在シグナル、好ましくは少なくとも2つのNLSによる塩基編集因子の修飾によって塩基編集因子効率を改善するためのより幅広い戦略を実証することのみを意図されている。本発明は、どのNLSが採用されるか、およびNLSが塩基編集因子に取り付けられるかまたは別様にカップリングされる様式について、限定的であることを意図されていない。NLS配列は当分野において公知であり、例が本明細書において開示される。

20

【0550】

ベクター

本発明の塩基編集因子を作ることおよび用いることのいくつかの側面は、1つ以上のベクターを含むベクターシステム、またはベクターそのものに関する。ベクターは本開示の進化した塩基編集因子をクローニングおよび/または発現するように設計され得る。ベクターは、本開示の進化した塩基編集因子を、1つ以上の細胞、例えば本明細書において開示される塩基編集因子システムおよび方法による処置のための標的有疾患真核生物細胞にトランスフェクションするようにもまた設計され得る。

30

【0551】

ベクターは、原核生物または真核生物細胞における塩基編集因子転写物(例えば、核酸転写物、蛋白質、または酵素)の発現のために設計され得る。例えば、塩基編集因子転写物は、細菌細胞、例えばEscherichia coli、昆虫細胞(バキュロウイルス発現ベクターを用いる)、酵母細胞、または哺乳類細胞によって発現され得る。好適なホスト細胞はGoedel, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990)においてさらに論ぜられている。代替的には、本明細書に記載される1つ以上の進化した塩基編集因子をコードする発現ベクターは、*in vitro*で、例えばT7プロモーター制御配列およびT7ポリメラーゼを用いて転写および翻訳され得る。

40

【0552】

ベクターは原核生物細胞に導入され、増殖させられ得る。いくつかの態様では、原核生物は、真核生物細胞に導入されるべきベクターのコピーを増幅するために、または真核生物細胞に導入されるべきベクターの産生の中間体ベクターとして用いられる(例えば、ウイルスベクターパッケージングシステムの一部としてプラスミドを増幅する)。いくつかの態様では、原核生物は、ベクターのコピーを増幅し、1つ以上の核酸を発現するために用いられて、例えば、ホスト細胞またはホスト生物への送達のための1つ以上の蛋白質のソースを提供する。原核生物における蛋白質の発現は、最も多くの場合には、融合または非融合蛋白質どちらかの発現を導く恒常的または誘導性プロモーターを含有するベクターによって、Escherichia coliにおいて行われる。

50

【0553】

融合体発現ベクターもまた、本開示の進化した塩基編集因子を発現するために用いられ得る。一般的に、かかるベクターは、その中にコードされる蛋白質に、例えば組換え体蛋白質のアミノ末端にいくつものアミノ酸を追加する。かかる融合ベクターは1つ以上の目的に供し得る。例えば：(i)組換え体蛋白質の発現を増大させるため；(ii)組換え体蛋白質の可溶性を増大させるため；および(iii)アフィニティー精製のリガンドとして作用することによって組換え体蛋白質の精製を支援するため。多くの場合には、融合体発現ベクターでは、蛋白質加水分解的な切断部位が融合体部分と組換え体蛋白質とのジャンクションに導入されて、融合蛋白質の精製の爾後の融合体部分からの組換え体蛋白質の分離を可能化する。かかる酵素およびそれらのコグネイトな認識配列は第Xa因子、トロンピン、およびエンテロキナーゼを包含する。融合体発現ベクターの例はpGEX (Pharmacia Biotech Inc; Smith and Johnson, 1988. Gene 67: 31-40)、pMAL (New England Biolabs, Beverly, Mass.)、およびpRIT5 (Pharmacia, Piscataway, N.J.)を包含し、これらはそれぞれグルタチオンS-トランスフェラーゼ(GST)、マルトースE結合蛋白質、またはプロテインAを標的の組換え体蛋白質に融合させる。

10

【0554】

好適な誘導性非融合体E. coli発現ベクターの例はpTrc (Amrann et al., (1988) Gene 69: 301-315) およびpET11d (Studier et al., GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990) 60-89)を包含する。

【0555】

いくつかの態様では、ベクターは、本明細書に記載される進化した塩基編集因子を発現するための酵母発現ベクターである。酵母Saccharomyces cerevisiaeによる発現のためのベクターの例はpYepSec1 (Baldari, et al., 1987. EMBO J. 6: 229-234)、pMFa (Kuijjan and Herskowitz, 1982. Cell 30: 933-943)、pJRY88 (Schultz et al., 1987. Gene 54: 113-123)、pYES2 (Invitrogen Corporation, San Diego, Calif.)、およびpicZ (Invitrogen Corp, San Diego, Calif.)を包含する。

20

【0556】

いくつかの態様では、ベクターはバキュロウイルス発現ベクターを用いて昆虫細胞による蛋白質発現を駆動する。培養昆虫細胞(例えば、SF9細胞)による蛋白質の発現のための利用可能なバキュロウイルスベクターはpAcシリーズ (Smith, et al., 1983. Mol. Cell Biol. 3: 2156-2165) およびpVLシリーズ (Lucklow and Summers, 1989. Virology 170: 31-39)を包含する。

30

【0557】

いくつかの態様では、ベクターは哺乳類発現ベクターを用いて哺乳類細胞による1つ以上の配列の発現を駆動する能力がある。哺乳類発現ベクターの例はpCDM8 (Seed, 1987. Nature 329: 840) およびpMT2PC (Kaufman, et al., 1987. EMBO J. 6: 187-195)を包含する。哺乳類細胞に用いられるときに、発現ベクターのコントロール機能は典型的には1つ以上の制御エレメントによって提供される。例えば、普通に用いられるプロモーターは、ポリオーマ、アデノウイルス2、サイトメガロウイルス、シミアンウイルス40、ならびに本明細書において開示されるおよび当分野において公知の他のものに由来する。原核生物および真核生物細胞両方についての他の好適な発現システムは、例えばSambrook, et al., MOLECULAR CLONING: A LABORATORY MANUAL. 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989の16および17章を参照する。

40

【0558】

いくつかの態様では、組換え体哺乳類発現ベクターは特定の細胞型において選好的に核酸の発現を導く能力がある(例えば、組織特異的な制御エレメントが核酸を発現するために用いられる)。組織特異的な制御エレメントは当分野において公知である。好適な組織特異的なプロモーターの限定しない例は、アルブミンプロモーター(肝臓特異的; Pinkert, et al., 1987. Genes Dev. 1: 268-277)、リンパ特異的なプロモーター (Calame and Eat

50

on, 1988. Adv. Immunol. 43: 235-275)、特にT細胞受容体(Winoto and Baltimore, 1989. EMBO J. 8: 729-733)およびイムノグロブリン(Baneiji, et al., 1983. Cell 33: 729-740; Queen and Baltimore, 1983. Cell 33: 741-748)のプロモーター、ニューロン特異的プロモーター(例えば、ニューロフィラメントプロモーター; Byrne and Ruddle, 1989. Proc. Natl. Acad. Sci. USA 86: 5473-5477)、膵臓特異的プロモーター(Edlund, et al., 1985. Science 230: 912-916)、および乳腺特異的プロモーター(例えば、乳清プロモーター、U.S.Pat.No.4,873,316および欧州出願公開No.264,166)を包含する。発生的に制御されるプロモーターもまた包摂され、例えばネズミhoxプロモーター(Kessel and Gruss, 1990. Science 249: 374-379)および-フェトプロテインプロモーター(Campes and Tilghman, 1989. Genes Dev. 3: 537-546)である。

10

【0559】

塩基編集因子効率を増大させる

本開示のいくつかの側面は、本明細書において提供される塩基編集因子のいずれかが、有意な割合のインデルを生成することなしに、特異的なヌクレオチド塩基を改変する能力があるという認識に基づく。本明細書において用いられる「インデル」は核酸中のヌクレオチド塩基の挿入または欠失を言う。かかる挿入または欠失は遺伝子のコード領域内のフレームシフト変異に至り得る。いくつかの態様では、核酸中に多数の挿入または欠失(すなわちインデル)を生成することなしに、核酸中の特異的なヌクレオチドを効率的に改変する(例えば変異させるかまたは脱アミノ化する)塩基編集因子を生成することが望ましい。ある態様では、本明細書において提供される塩基編集因子のいずれかは、インデルに対してより多大な割合の意図される改変(例えば点変異または脱アミノ化)を生成する能力がある。いくつかの態様では、本明細書において提供される塩基編集因子は、1:1よりも多大である意図される点変異対インデルの比を生成する能力がある。いくつかの態様では、本明細書において提供される塩基編集因子は、少なくとも1.5:1、少なくとも2:1、少なくとも2.5:1、少なくとも3:1、少なくとも3.5:1、少なくとも4:1、少なくとも4.5:1、少なくとも5:1、少なくとも5.5:1、少なくとも6:1、少なくとも6.5:1、少なくとも7:1、少なくとも7.5:1、少なくとも8:1、少なくとも10:1、少なくとも12:1、少なくとも15:1、少なくとも20:1、少なくとも25:1、少なくとも30:1、少なくとも40:1、少なくとも50:1、少なくとも100:1、少なくとも200:1、少なくとも300:1、少なくとも400:1、少なくとも500:1、少なくとも600:1、少なくとも700:1、少なくとも800:1、少なくとも900:1、もしくは少なくとも1000:1、またはより多くである意図される点変異対インデルの比を生成する能力がある。意図される変異およびインデルの数は、いずれかの好適な方法、例えば下の実施例に用いられている方法を用いて決定され得る。いくつかの態様では、インデル頻度を計算するためには、シーケンシングリードが、インデルが起こり得るウィンドウの両側をフランキングする2つの10bp配列に対する厳密なマッチについてスキャンされる。厳密なマッチが見つからない場合には、リードは分析から除外される。このインデルウィンドウの長さが厳密に参照配列にマッチする場合には、リードはインデルを含有していないと分類される。インデルウィンドウが参照配列よりも2塩基以上長いまたは短い場合には、シーケンシングリードはそれぞれ挿入または欠失として分類される。

20

30

40

【0560】

いくつかの態様では、本明細書において提供される塩基編集因子は核酸のある領域におけるインデルの形成を限定する能力がある。いくつかの態様では、領域は、塩基編集因子によって標的化されるヌクレオチド、または塩基編集因子によって標的化されるヌクレオチドから2、3、4、5、6、7、8、9、もしくは10ヌクレオチド以内の領域である。いくつかの態様では、本明細書において提供される塩基編集因子のいずれかは、核酸のある領域におけるインデルの形成を1%未満、1.5%未満、2%未満、2.5%未満、3%未満、3.5%未満、4%未満、4.5%未満、5%未満、6%未満、7%未満、8%未満、9%未満、10%未満、12%未満、15%未満、または20%未満に限定する能力がある。ある核酸領域において形成されるインデルの数は、核酸(例えば、細胞のゲノム内の核酸)が塩基編集因子に暴露される時間の量に依存

50

し得る。いくつかの態様では、インデルの数または割合は、少なくとも1時間、少なくとも2時間、少なくとも6時間、少なくとも12時間、少なくとも24時間、少なくとも36時間、少なくとも48時間、少なくとも3日、少なくとも4日、少なくとも5日、少なくとも7日、少なくとも10日、または少なくとも14日の、核酸（例えば、細胞のゲノム内の核酸）を塩基編集因子に暴露することの後に決定される。

【0561】

本開示のいくつかの側面は、本明細書において提供される塩基編集因子のいずれかが、有意な数の意図されない変異、例えば意図されない点変異を生成することなしに、意図される変異、例えば点変異を核酸（例えば、対象のゲノム内の核酸）中に効率的に生成する能力があるという認識に基づく。いくつかの態様では、意図される変異は、意図される変異を生成するように特異的に設計されたgRNAに結合した特異的な塩基編集因子によって生成される変異である。いくつかの態様では、意図される変異は疾患または異常に関連する変異である。いくつかの態様では、意図される変異は疾患または異常に関連するアデニン（A） グアニン（G）点変異である。いくつかの態様では、意図される変異は疾患または異常に関連するチミン（T） シトシン（C）点変異である。いくつかの態様では、意図される変異は遺伝子のコード領域内のアデニン（A） グアニン（G）点変異である。いくつかの態様では、意図される変異は遺伝子のコード領域内のチミン（T） シトシン（C）点変異である。いくつかの態様では、意図される変異は、遺伝子のコード領域内に終止コドン、例えば未成熟終止コドンを生成する点変異である。いくつかの態様では、意図される変異は終止コドンを抹消する変異である。いくつかの態様では、意図される変異は遺伝子のスプライシングを変調させる変異である。いくつかの態様では、意図される変異は遺伝子の制御配列（例えば、遺伝子プロモーターまたは遺伝子リプレッサー）を変調させる変異である。いくつかの態様では、本明細書において提供される塩基編集因子のいずれかは、1:1よりも多大である意図される変異対意図されない変異（例えば、意図される点変異：意図されない点変異）の比を生成する能力がある。いくつかの態様では、本明細書において提供される塩基編集因子のいずれかは、少なくとも1.5:1、少なくとも2:1、少なくとも2.5:1、少なくとも3:1、少なくとも3.5:1、少なくとも4:1、少なくとも4.5:1、少なくとも5:1、少なくとも5.5:1、少なくとも6:1、少なくとも6.5:1、少なくとも7:1、少なくとも7.5:1、少なくとも8:1、少なくとも10:1、少なくとも12:1、少なくとも15:1、少なくとも20:1、少なくとも25:1、少なくとも30:1、少なくとも40:1、少なくとも50:1、少なくとも100:1、少なくとも150:1、少なくとも200:1、少なくとも250:1、少なくとも500:1、もしくは少なくとも1000:1、またはより多くである、意図される変異対意図されない変異（例えば、意図される変異：意図されない点変異）の比を生成する能力がある。本明細書の「塩基編集因子効率」セクションに記載される塩基編集因子の性質は、本明細書において提供される融合蛋白質、または融合蛋白質を用いる方法のいずれかに適用され得るということは了解されるはずである。

II. 進化した塩基編集因子を作り用いる方法

【0562】

本開示のいくつかの側面は、本明細書において開示される進化した塩基編集因子、または1つ以上のnapR/DNAbpをプログラムする核酸分子（例えば、Cas9ガイドRNA）と本明細書において提供される核酸塩基編集因子とを含む塩基編集因子複合体を作る方法を提供する。加えて、本開示のいくつかの側面は、標的ヌクレオチド配列（例えばゲノム）を編集するための進化した塩基編集因子を用いる方法を提供する。

【0563】

連続的進化方法

本開示の種々の側面は、連続的進化方法およびシステム（例えば、適切なベクター、細胞、ファージ、フローベッセルなど）を提供することに関する。

【0564】

本明細書において提供される連続的進化方法は、ウイルスベクター中の進化させられるべき目当ての遺伝子（例えば、塩基編集因子遺伝子）が、ウイルスライフサイクルの複数

10

20

30

40

50

世代をかけて、宿主細胞のフロー中において、所望の機能または活性を獲得することを許す。

【0565】

本発明のいくつかの側面は、目当ての遺伝子の連続的進化の方法を提供し、(a) 宿主細胞の集団を目当ての遺伝子を含むウイルスベクターの集団と接触させることを含み、(1) 宿主細胞はウイルスベクターによる感染を受け入れ；(2) 宿主細胞は、ウイルス粒子の生成のために要求されるウイルス遺伝子を発現し；(3) 感染性のウイルス粒子の産生のために要求される少なくとも1つのウイルス遺伝子の発現は、目当ての遺伝子の機能に依存し；(4) ウイルスベクターは宿主細胞による蛋白質の発現を許し、宿主細胞によってウイルス粒子へと複製およびパッケージングされ得る。いくつかの態様では、方法は(b) 宿主細胞を変異原と接触させることを含み。いくつかの態様では、方法はさらに(c) ウイルス複製およびウイルス粒子の産生を許す条件下で宿主細胞の集団をインキュベーションすることを含み、宿主細胞が宿主細胞集団から取り除かれ、フレッシュな未感染宿主細胞が宿主細胞の集団に導入され、それゆえに宿主細胞の集団を補充し、宿主細胞のフローを作出する。細胞は、全ての態様において、目当ての遺伝子に変異を獲得することを許す条件下でインキュベーションされる。いくつかの態様では、方法はさらに(d) 進化した遺伝子産物(例えば蛋白質)をコードするウイルスベクターの変異したバージョンを宿主細胞の集団から単離することを含む。

10

【0566】

いくつかの態様では、ファージによって支援される連続的進化の方法が提供され、(a) 細菌宿主細胞の集団を、進化させられるべき目当ての遺伝子を含み、かつ感染性のファージの生成のために要求される遺伝子を欠損しているファージの集団と接触させることを含み、(1) ファージは宿主細胞による目当ての遺伝子の発現を許し；(2) 宿主細胞はファージ感染、複製、およびパッケージングにとって好適な宿主細胞であり；(3) 宿主細胞は、感染性のファージの生成のために要求される遺伝子をコードする発現コンストラクトを含み、遺伝子の発現は目当ての遺伝子の遺伝子産物の機能に依存する。いくつかの態様では、方法はさらに(b) 宿主細胞の集団を、目当ての遺伝子の変異、感染性のファージの産生、およびファージによる宿主細胞の感染を許す条件下でインキュベーションすることを含み、感染した細胞は宿主細胞の集団から取り除かれ、宿主細胞の集団は、ファージによって感染されていないフレッシュな宿主細胞を補充される。いくつかの態様では、方法はさらに(c) 進化した蛋白質をコードする変異したファージ複製産物を宿主細胞の集団から単離することを含む。

20

30

【0567】

いくつかの態様では、ウイルスベクターまたはファージは、繊維状ファージ、例えばM13ファージ、例えば本明細書において他所により詳細に記載されるM13セレクションファージである。いくつかのかかる態様では、感染性のウイルス粒子の産生のために要求される遺伝子はM13遺伝子III(gIII)である。

【0568】

いくつかの態様では、ウイルスベクターは哺乳類細胞に感染する。いくつかの態様では、ウイルスベクターはレトロウイルスベクターである。いくつかの態様では、ウイルスベクターは水胞性口炎ウイルス(VSV)ベクターである。dsRNAウイルスとして、VSVは高い変異速度を有し、長さが最高で4.5kbの目当ての遺伝子を包含するカーゴを持ち得る。感染性のVSV粒子の生成は、ホスファチジルセリン取り付けおよび細胞侵入を媒介するウイルス糖蛋白質、エンベロープ蛋白質VSV-Gを要求する。VSVは哺乳類および昆虫細胞を包含する広いスペクトルの宿主細胞に感染し得る。よって、VSVはヒト、マウス、または昆虫宿主細胞における連続的進化にとって高度に好適なベクターである。類似に、VSV-Gエンベロープ蛋白質によってシュードタイピングされ得る他のレトロウイルスベクターは、本明細書に記載される連続的進化プロセスにとって等しく好適である。

40

【0569】

多くのレトロウイルスベクター、例えばネズミ白血病ウイルスベクターまたはレンチウ

50

イルスベクターが、ウイルスの固有のエンベロープ蛋白質の代用としてVSV-Gエンベロープ蛋白質によって効率的にパッケージングされ得るということは当業者には公知である。いくつかの態様では、固有のエンベロープ (env) 蛋白質 (例えば、VSVベクター中のVSV-GまたはMLVベクター中のenv) がウイルスゲノムから欠失させられ、目当ての遺伝子は、所望の宿主細胞内で活性であるプロモーターのコントロール下でウイルスゲノムに挿入されるということで、かかるVSV-Gパッケージング可能なベクターは連続的進化システムへの使用のために適応させられる。翻って、宿主細胞は、VSV-G蛋白質、ベクターシュードタイピングにとって好適な別のenv蛋白質、またはウイルスベクターの固有のenv蛋白質を、その活性が目当ての遺伝子によってコードされる産物の活性に依存するプロモーターのコントロール下で発現し、その結果、目当ての遺伝子の増大した活性に至る変異を有するウイルスベクターは、ベースラインまたは機能喪失変異を有するベクターよりも高い効率でパッケージングされるであろう。

【0570】

いくつかの態様では、哺乳類宿主細胞が、ウイルスベクター、例えば目当ての遺伝子を含みかつVSV-Gをコードする遺伝子を欠くVSVベクターの連続的に進化する集団による感染に付され、宿主細胞はVSV-G蛋白質をコードする遺伝子を条件付きプロモーターのコントロール下に含む。かかるレトロウイルスに基づく (bases) システムは2ベクターシステムであり得 (ウイルスベクター、およびエンベロープ蛋白質をコードする遺伝子を含む発現コンストラクト)、または代替的にはヘルパーウイルス、例えばVSVヘルパーウイルスが採用され得る。ヘルパーウイルスは典型的には短縮されたウイルスゲノムを含み、ゲノムをウイルス粒子へとパッケージングするために要求される構造的要素を欠損しているが、宿主細胞内でのウイルスゲノムプロセッシングおよびウイルス粒子の生成のために要求される蛋白質をコードするウイルス遺伝子を包含する。かかる態様では、ウイルスベクターに基づくシステムは3ベクターシステムであり得る (ウイルスベクター、条件付きプロモーターによって駆動されるエンベロープ蛋白質を含む発現コンストラクト、およびウイルスゲノム増殖のために要求されるウイルス機能を含むがエンベロープ蛋白質を含まないヘルパーウイルス)。いくつかの態様では、宿主細胞内のヘルパーウイルスまたは発現コンストラクトからのVSVゲノムの5つの遺伝子の発現が、目当ての遺伝子を持つ感染性のウイルス粒子の産生を許し、不均衡な遺伝子発現がウイルス複製を縮減された速度で許すということを指示し、実に、VSV-Gの縮減された発現が効率的なウイルス産生における律速 (limiting) ステップとしての用をなすであろうということを示唆している。

【0571】

ヘルパーウイルスを用いる1つの利点は、ウイルスベクターが、ヘルパーウイルスによって提供される蛋白質または他の機能をコードする遺伝子を欠損し得、従ってより長い目当ての遺伝子を持ち得るということである。いくつかの態様では、ヘルパーウイルスはエンベロープ蛋白質を発現しない。なぜなら、ウイルスエンベロープ蛋白質の発現はいくつかのウイルスベクターによる宿主細胞の感染能力を受容体干渉によって縮減することが公知であるからである。連続的進化プロセスにとって好適なウイルスベクター、例えばレトロウイルスベクター、それらのそれぞれのエンベロープ蛋白質、およびかかるベクターのためのヘルパーウイルスは当業者に周知である。本明細書に記載される連続的進化手続にとって好適ないくつかの例示的なウイルスゲノム、ヘルパーウイルス、宿主細胞、およびエンベロープ蛋白質の概略は、Coffin et al., *Retroviruses*, CSHL Press 1997, IS BN0-87969-571-4を参照する。その全体は本明細書に組み込まれる。

【0572】

いくつかの態様では、宿主細胞のインキュベーションは、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100、少なくとも200、少なくとも300、少なくとも400、少なくとも500、少なくとも600、少なくとも700、少なくとも800、少なくとも900、少なくとも1000、少なくとも1250、少なくとも1500、少なくとも1750、少なくとも2000、少なくとも2500、少なくとも3000、少なくとも4000、少なくとも5000、少なくとも7500、少なくとも10000、またはより多くの継続したウイルスライフサ

イクルにとって十分な時間に渡ってである。ある態様では、ウイルスベクターはM13ファージであり、1回のウイルスライフサイクルの長さは約10-20分である。

【0573】

いくつかの態様では、細胞は懸濁培養によって接触および/またはインキュベーションされる。例えば、いくつかの態様では、細菌細胞が懸濁培養によって液体培養培地中でインキュベーションされる。細菌懸濁培養のための好適な培養培地は当業者には明らかであろう。本発明はこれについて限定されない。例えば、Molecular Cloning: A Laboratory Manual, 2nd Ed., ed. by Sambrook, Fritsch, and Maniatis (Cold Spring Harbor Laboratory Press: 1989); Elizabeth Kutter and Alexander Sulakvelidze: Bacteriophages: Biology and Applications. CRC Press; 1st edition (December 2004), ISBN: 0849313368; Martha R. J. Clokie and Andrew M. Kropinski: Bacteriophages: Methods and Protocols, Volume 1: Isolation, Characterization, and Interactions (Methods in Molecular Biology) Humana Press; 1st edition (December, 2008), ISBN: 1588296822; Martha R. J. Clokie and Andrew M. Kropinski: Bacteriophages: Methods and Protocols, Volume 2: Molecular and Applied Aspects (Methods in Molecular Biology) Humana Press; 1st edition (December 2008), ISBN: 1603275649を参照する; これらの全ては、細菌宿主細胞培養のための好適な培養培地の開示について、それらの全体が参照によって本明細書に組み込まれる)。懸濁培養は、典型的には、培養培地が連続的にまたは間欠的にどちらかで振盪されることを要求する。これは、いくつかの態様では、宿主細胞集団を含むベッセルを振盪または攪拌することによって達成される。いくつかの態様では、宿主細胞のアウトフローおよびフレッシュな宿主細胞のインフローは、懸濁した宿主細胞を維持するために十分である。これは、特に、ラグーン内へのおよび/またはから外への細胞の流量が高い場合にである。

10

20

【0574】

いくつかの態様では、ウイルスベクターのライフサイクルが宿主細胞の細胞分裂間の平均時間よりも有意に短い、ウイルスベクター/宿主細胞組み合わせが選ばれる。平均の細胞分裂時間およびウイルスベクターライフサイクル時間は多くの細胞型およびベクターについて当分野において周知であり、当業者がかかる宿主細胞/ベクター組み合わせを確かめることを許す。ある態様では、ある宿主細胞が取り除かれる前に宿主細胞集団に残る平均時間が宿主細胞の細胞分裂間の平均時間よりも短い、採用されるウイルスベクターの平均ライフサイクルよりは長いことをもたらす速度で、宿主細胞は、ウイルスベクターと接触させられた宿主細胞の集団から取り除かれていく。この結果は、宿主細胞が、平均して、宿主細胞集団内でのそれらの時間中に増殖するための十分な時間を有さないということであり、一方で、ウイルスベクターは、ある宿主細胞が細胞集団に残る時間中に、宿主細胞に感染、宿主細胞内で複製、および新たなウイルス粒子を生成するために十分な時間を有する。これは、宿主細胞集団中の唯一の複製する核酸はウイルスベクターであるということと、宿主細胞ゲノム、アクセサリープラスミド、またはいずれかの他の核酸コンストラクトは課される選択圧からの脱出を許す変異を獲得し得ないということとを保証する。

30

【0575】

例えば、いくつかの態様では、ある宿主細胞が宿主細胞集団中に残る平均時間は約10、約11、約12、約13、約14、約15、約16、約17、約18、約19、約20、約21、約22、約23、約24、約25、約30、約35、約40、約45、約50、約55、約60、約70、約80、約90、約100、約120、約150、または約180分である。

40

【0576】

いくつかの態様では、ある宿主細胞が宿主細胞集団中に残る平均時間は、どのくらい速く宿主細胞が分裂するか、およびどのくらい長くを感染(または接合)が要求するかに依存する。一般的に、流量は、細胞分裂に要求される平均時間よりも速いが、ウイルス(または接合)増殖を許すためには十分遅くあるべきである。前者は、例えば培地の型によって変わり、細胞分裂阻害抗生物質(E. coliのFtsZ阻害因子など)を追加すること

50

によって遅延させられ得る。連続的進化の律速 (limiting) ステップは細胞から細胞への遺伝子伝播のために要求される蛋白質の産生であるので、ベクターがウォッシュアウトする流量は目当ての遺伝子 (単数または複数) の現行の活性に依存するであろう。いくつかの態様では、本明細書に記載される感染性の粒子の生成のために要求される蛋白質のタイトレーション可能な産生がこの問題を緩和し得る。いくつかの態様では、ファージ感染のインジケータが、流量のコンピュータによってコントロールされた最適化を、現行の活性レベルについてリアルタイムで許す。

【0577】

いくつかの態様では、宿主細胞集団はフレッシュな未感染宿主細胞を連続的に補充される。いくつかの態様では、これは、宿主細胞の集団内へのフレッシュな宿主細胞の定常的な流れによって奏される。しかしながら、他の態様では、ラグーン内へのフレッシュな宿主細胞のインフローは半連続的または間欠的 (例えば、バッチ投入される) である。いくつかの態様では、細胞集団内へのフレッシュな宿主細胞インフローの速度は、宿主細胞集団からの細胞の取り除きの速度が補償されるようである。いくつかの態様では、この細胞フロー補償の結果は、細胞集団の細胞数が連続的進化手続きの時間中に実質的に一定であるということである。いくつかの態様では、細胞集団中のフレッシュな未感染細胞の部分は連続的進化手続きの時間中に実質的に一定である。例えば、いくつかの態様では、宿主細胞集団の細胞の約10%、約15%、約20%、約25%、約30%、約40%、約50%、約60%、約70%、約75%、約80%、または約90%はウイルスに感染していない。一般的には、宿主細胞の流量がより速いほど、感染している宿主細胞集団の細胞の部分はより小さいであろう。しかしながら、より速い流量は、より多くの伝播サイクル、例えばウイルスライフサイクルを、それゆえに進化したベクターのより多くの世代を所与の時間で許すであろう。一方で、より遅い流量は、宿主細胞集団中のより大きい部分の感染した宿主細胞、よって、より大きいライブラリーサイズを、より遅い進化と引き換えにもたらす。いくつかの態様では、有効な流量の範囲は、遅い端では細胞分裂時間を、高い端ではベクターウォッシュアウトを必ず限度とする。いくつかの態様では、例えば細胞培養培地の体積あたりの感染性のウイルス粒子として測定される、ウイルス負荷は、連続的進化手続きの時間中に実質的に一定である。

【0578】

いくつかの態様では、フレッシュな宿主細胞は、ウイルスベクターのセレクションのために要求されるアクセサリプラスミド、例えば、進化させられようとするファージに欠けている感染性のファージ粒子の生成のために要求される遺伝子を含むアクセサリプラスミドを含む。いくつかの態様では、宿主細胞は、未感染の宿主細胞を然るべきベクター、例えばアクセサリプラスミドおよび任意に変異導入プラスミドと接触させることと、連続的進化実験の宿主細胞集団の補充のために十分な量の宿主細胞を増殖させることとによって生成される。宿主細胞へのプラスミドおよび他の遺伝子コンストラクトの導入のための方法は当業者に周知であり、本発明はこれについて限定されない。細菌宿主細胞では、かかる方法はコンピテントセルのエレクトロポレーションおよび熱ショックを包含するが、これらに限定されない。いくつかの態様では、アクセサリプラスミドは、セレクションマーカー、例えば抗生物質耐性マーカーを含み、フレッシュな宿主細胞はそれぞれの抗生物質の存在下で増殖させられて、宿主細胞内のプラスミドの存在を保証する。複数のプラスミドが存在するところでは、典型的には異なるマーカーが用いられる。細胞培養におけるかかるセレクションマーカーおよびそれらの使用は当業者に公知であり、本発明はこれについて限定されない。

【0579】

いくつかの態様では、連続的進化実験中の宿主細胞集団は、並行した連続培養によって増殖するフレッシュな宿主細胞を補充される。いくつかの態様では、ウイルスベクターと接触させられた宿主細胞集団中の宿主細胞の細胞密度とフレッシュな宿主細胞集団の密度とは実質的に同じである。

【0580】

10

20

30

40

50

典型的には、ウイルスベクターと接触させられた細胞集団から取り除かれようとする細胞は、ウイルスベクターに感染している細胞および未感染細胞を含む。いくつかの態様では、細胞は、例えば集団からの細胞の連続的アウトフローを成し遂げることによって細胞集団から連続的に取り除かれていく。他の態様では、細胞は半連続的にまたは間欠的に集団から取り除かれる。いくつかの態様では、フレッシュな細胞の補充は、細胞集団からの細胞の取り除きのモードにマッチするであろう。例えば、細胞が連続的に取り除かれる場合には、フレッシュな細胞が連続的に導入されるであろう。しかしながら、いくつかの態様では、補充および取り除きのモードはミスマッチであり得、例えば、細胞集団はフレッシュな細胞を連続的に補充され得、細胞が半連続的にまたはバッチで取り除かれ得る。

【0581】

いくつかの態様では、フレッシュな宿主細胞補充の速度および/または宿主細胞取り除きの速度は、細胞集団中の宿主細胞を定量することに基づいて調整される。例えば、いくつかの態様では、宿主細胞集団を含む培養培地の濁度がモニタリングされ、濁度が閾値レベルよりも下に入る場合には、宿主細胞インフロー対宿主細胞アウトフローの比は、増大した細胞培養物濁度によって表明される集団中の宿主細胞数の増大を成し遂げるように調整される。他の態様では、濁度が閾値レベルよりも上に上がる場合には、宿主細胞インフロー対宿主細胞アウトフローの比は、減少した細胞培養物濁度によって表明される集団中の宿主細胞数の減少を成し遂げるように調整される。宿主細胞集団中の宿主細胞の密度を具体的な密度範囲内に維持することは、十分な宿主細胞が進化するウイルスベクター集団の宿主として利用可能であるということを保証し、ウイルスパッケージングと引き換えの栄養素の枯渇と、過密な培養物からの細胞起源の毒素の蓄積とを回避する。

【0582】

いくつかの態様では、宿主細胞集団の細胞密度および/またはインフローのフレッシュな宿主細胞密度は約 10^2 細胞/mlから約 10^{12} 細胞/mlである。いくつかの態様では、宿主細胞密度は約 10^2 細胞/ml、約 10^3 細胞/ml、約 10^4 細胞/ml、約 10^5 細胞/ml、約 $5 \cdot 10^5$ 細胞/ml、約 10^6 細胞/ml、約 $5 \cdot 10^6$ 細胞/ml、約 10^7 細胞/ml、約 $5 \cdot 10^7$ 細胞/ml、約 10^8 細胞/ml、約 $5 \cdot 10^8$ 細胞/ml、約 10^9 細胞/ml、約 $5 \cdot 10^9$ 細胞/ml、約 10^{10} 細胞/ml、または約 $5 \cdot 10^{10}$ 細胞/mlである。いくつかの態様では、宿主細胞密度は約 10^{10} 細胞/mlよりも多い。

【0583】

いくつかの態様では、宿主細胞集団は変異原と接触させられる。いくつかの態様では、ウイルスベクター（例えばファージ）と接触させられた細胞集団は、宿主細胞集団中にある間に、目当ての遺伝子の増大した変異速度を許すが変異原へのそれらの暴露中の宿主細胞にとって有意に毒性ではない濃度で、変異原に連続暴露される。他の態様では、宿主細胞集団は間欠的に変異原と接触させられ、増大した変異導入の、従って増大したウイルスベクター多様化のフェーズを作出する。例えば、いくつかの態様では、宿主細胞は、時間の約10%、約20%、約50%、または約75%に渡って、目当ての遺伝子の増大した変異導入速度を生成するために十分な濃度の変異原に暴露される。

【0584】

いくつかの態様では、宿主細胞は、変異導入発現コンストラクト、例えば細菌宿主細胞のケースでは変異導入プラスミドを含む。いくつかの態様では、変異導入プラスミドは、変異導入を促進する遺伝子産物、例えば校正が損なわれたDNAポリメラーゼをコードする遺伝子発現カセットを含む。他の態様では、変異導入プラスミドはSOSストレス応答に関わる遺伝子（例えば、UmuC、UmuD'、および/またはRecA）を包含する（including）。いくつかの態様では、変異導入を促進する遺伝子は誘導性プロモーターのコントロール下にある。好適な誘導性プロモーターは当業者には周知であり、例えば、アラビノース誘導性プロモーター、テトラサイクリンまたはドキシサイクリン誘導性プロモーター、およびタモキシフェン誘導性プロモーターを包含する。いくつかの態様では、宿主細胞集団は、変異導入の増大した速度を成し遂げるために十分な量で誘導性プロモーターの誘導因子と接触させられる。例えば、いくつかの態様では、細菌宿主細胞集団が提供され、こ

10

20

30

40

50

れの中の宿主細胞は、dnaQ926、UmuC、UmuD'、およびRecA発現カセットがアラビノース誘導性プロモーターによってコントロールされる変異導入プラスミドを含む。いくつかのかかる態様では、宿主細胞の集団は、誘導因子、例えばアラビノースと、増大した変異速度を誘導するために十分な量で接触させられる。

【0585】

誘導性変異導入プラスミドの使用は、誘導因子の不在下ではフレッシュな未感染宿主細胞の集団を生成することを許し、それゆえに、それらがウイルスベクターと接触させられた細胞の集団に導入される前には、フレッシュな宿主細胞の増大した変異速度を回避する。しかしながら、ひとたびこの集団に導入されると、これらの細胞は増大した変異速度をよように誘導され得、これは連続的進化のいくつかの態様に特に有用である。例えば、いくつかの態様では、宿主細胞は本明細書に記載される変異導入プラスミドを含み、dnaQ926、UmuC、UmuD'、およびRecA730の発現をpBADプロモーターから駆動するアラビノース誘導性プロモーターを含む（例えば、Khlebnikov A, Skaug T, Keasling JD. Modulation of gene expression from the arabinose-inducible araBAD promoter. *J Ind Microbiol Biotechnol.* 2002 Jul ; 29 (1): 34-7を参照する；pBADプロモーターの開示について参照によって本明細書に組み込まれる）。いくつかの態様では、宿主細胞が、セクションフェージの集団が複製するラグーンに到達するまで、フレッシュな宿主細胞はアラビノースに暴露されない。これは上で想定されている遺伝子の発現を活性化し、それゆえにアラビノースに暴露された細胞の変異速度を増大させる。従って、いくつかの態様では、それらがラグーンの宿主細胞集団の一部になり、そこでそれらが誘導因子（例えばアラビノース）に、それゆえに増大した変異導入に暴露されるまで、宿主細胞の変異速度は正常である。いくつかの態様では、変異導入によってウイルスベクターの集団を多様化させるフェーズを包含する連続的進化の方法が提供され、これにおいては、細胞は、進化した蛋白質をコードするウイルスベクターの変異した複製産物の厳格なセクションの不在下で、ウイルスベクターの変異導入にとって好適な条件下でインキュベーションされる。これは、進化させられるべき所望の機能がis単にすでに存在する機能の増大、例えば転写因子の転写活性化速度の増大ではなく、進化手続きの出だしには目当ての遺伝子に存在しない機能の獲得である態様に特に有用である。ウイルスベクター、例えばフェージの集団中の目当ての遺伝子の変異したバージョンのプールを多様化させるステップは、所望の機能を授ける変異を見出す見込みの増大を許す。

10

20

30

【0586】

いくつかの態様では、ウイルスベクター集団を多様化させることは、ウイルスベクターの集団の複製、変異導入、および増殖について目当ての遺伝子の機能獲得変異をセクションしない宿主細胞のフローを提供することによって達成される。いくつかの態様では、宿主細胞は、感染性のウイルス粒子の生成のために要求される全ての遺伝子を発現する宿主細胞、例えば、完全なヘルパーフェージを発現し、それゆえに目当ての遺伝子に選択圧を課さない細菌細胞である。他の態様では、宿主細胞は、目当ての遺伝子の有意な機能獲得変異の不在下でさえも、ウイルスベクター増殖を支持するために十分なベースライン活性を有する条件付きプロモーターを含むアクセサリープラスミドを含む。これは、「リーキーな」条件付きプロモーターを用いることによって、高コピー数アクセサリープラスミドを用い、それゆえにベースラインリーキーさを増幅することによって、および/または目当ての遺伝子の最初のバージョンが低いレベルの活性を成し遂げ、所望の機能獲得変異は有意により高い活性を成し遂げる条件付きプロモーターを用いることによって、達成され得る。

40

【0587】

例えば、実施例セクションにより詳細に記載される通り、いくつかの態様では、感染性のフェージ粒子の生成のために要求される遺伝子を有する高コピーのアクセサリープラスミドを含む宿主細胞の集団が、目当ての遺伝子を含むセクションフェージと接触させられ、アクセサリープラスミドは、活性が目当ての遺伝子によってコードされる遺伝子産物の活性に依存する条件付きプロモーターからの生成のために要求される遺伝子の発現を

50

駆動する条件付きプロモーターを含む。いくつかのかかる態様では、目当ての最初の遺伝子とそのプロモーターに対する何らかの活性を呈示するやり方で条件付きプロモーターを設計することによって、低ストリンジェンシーセレクションフェーズが達成され得る。例えば、転写活性化因子、例えばT7RNAPまたは転写因子が非固有の標的DNA配列（例えば、T7RNAPが活性を有さないT3RNAPプロモーター配列）を認識するように進化させられるべき場合には、低ストリンジェンシーアクセサリープラスミドは、標的配列が所望の性質を含むが、転写活性化因子が標的配列を認識することを、ただしその固有の標的配列よりも少ない効率で許す固有の認識配列の特徴をもまた持ち続ける、条件付きプロモーターを含むように設計され得る。ハイブリッド標的配列（例えば、T7/T3ハイブリッドプロモーター。終局的に所望の標的配列のいくつかの特徴と固有の標的配列のいくつかとを有する）を含むかかる低ストリンジェンシーアクセサリープラスミドへの最初の暴露は、アクセサリープラスミドの寛容な性格に基づいて、直ちにはセレクションされないファージベクターの集団が複数の変異を獲得することによって多様化することを許す。それから、ファージベクターのかかる多様化した集団は、厳格なセレクションアクセサリープラスミド、例えば固有の標的配列の特徴を持ち続けない終局的に所望の標的配列をその条件付きプロモーター内に含むプラスミドに暴露され得、それゆえに、所望の標的配列の認識を許す変異を獲得していないファージベクターに対して強い負の選択圧を生成する。

10

【0588】

いくつかの態様では、進化するウイルスベクターの集団と接触させられた最初の宿主細胞集団が、最初の集団の宿主細胞とは異なるフレッシュな宿主細胞を補充される。例えば、いくつかの態様では、最初の宿主細胞集団は、低ストリンジェンシーアクセサリープラスミドを含むかもしくはかかるプラスミドを少しも含まない宿主細胞から作られるか、またはウイルス感染および増殖について寛容である。いくつかの態様では、低ストリンジェンシーのまたはセレクションなしの宿主細胞集団においてウイルスベクターの集団を多様化させた後に、フレッシュな宿主細胞は、目当ての遺伝子の所望の機能についてより厳格な選択圧を課す宿主細胞集団に導入される。例えば、いくつかの態様では、二次のフレッシュな宿主細胞はウイルス複製および増殖についてもはや寛容ではない。いくつかの態様では、厳格に選択的な宿主細胞はアクセサリープラスミドを含み、これの中の条件付きプロモーターが最小限のベースライン活性のみを呈示するかもしくはは全然呈示せず、および/またはこれは低いかまたは非常に低いコピー数でのみ宿主細胞内に存在する。

20

30

【0589】

様々な選択ストリンジェンシーの宿主細胞が関わるかかる方法は、目当ての遺伝子の最初のバージョンには完全に不在である機能の進化のために、例えば、目当ての最初の遺伝子として用いられる固有の転写因子は少しも認識しない外来性の標的配列を認識する転写因子の進化のために、本明細書において提供される連続的進化方法の威力を操ることを許す。または、別の例では、所望の標的配列には結合しないか、またはそれへと導かれるいずれかの活性を呈示しないDNA結合蛋白質、リコンビナーゼ、ヌクレアーゼ、ジンクフィンガー蛋白質、またはRNAポリメラーゼによる、所望の標的配列の認識である。

【0590】

いくつかの態様では、望まれない活性を罰することによって、本明細書に記載される連続的進化方法の間に、ネガティブセレクションが適用される。いくつかの態様では、これは、望まれない活性がpIII産生に干渉することを引き起こすことによって達成される。例えば、gIII RBSおよび/または開始コドンに対して相補的なアンチセンスRNAの発現は、ネガティブセレクションを適用する1つのやり方であり、プロテアーゼ（例えば、TEV）を発現することおよびpIII中にプロテアーゼ認識部位を操作することは、別のものである。

40

【0591】

いくつかの態様では、進化した産物の望まれない活性を罰することによって、本明細書に記載される連続的進化方法の間に、ネガティブセレクションが適用される。これは、例

50

えば、所望の進化産物が、高い特異性を有する酵素、例えば変調しているが広がってはいない特異性を有する転写因子またはプロテアーゼである場合には有用である。いくつかの態様では、望まれない活性のネガティブセレクションは、望まれない活性がpIII産生に干渉することを引き起こし、それゆえに望まれない活性を有する遺伝子産物をコードするファージゲノムの増殖を阻害することによって達成される。いくつかの態様では、pIIIのドミナントネガティブバージョンの発現またはgIII RBSおよび/またはgIII開始コドンに対して相補的なアンチセンスRNAの発現が、望まれない活性の存在に連結される。いくつかの態様では、認識または切断が望まれないヌクレアーゼまたはプロテアーゼ切断部位が、それぞれpIII転写物配列またはpIIIAミノ酸配列中に挿入される。いくつかの態様では、pIIIのドミナントネガティブバリエーションの発現を抑制し、かつ認識または切断が望まれないプロテアーゼ切断部位を含む転写または翻訳抑制因子が用いられる。

10

【0592】

いくつかの態様では、望まれない進化した産物活性をファージ増殖の阻害に連結することによって、非標的基質に対する活性のカウンターセレクションが達成される。例えば、転写因子が望まれないオフターゲット配列ではなく特異的な標的配列を認識するように進化させられるいくつかの態様では、オフターゲット配列を含むプロモーターのコントロール下にpIIIのドミナントネガティブバージョン (pIII-neg) をコードする核酸配列を含むネガティブセレクションカセットが採用される。進化産物がオフターゲット配列を認識する場合には、もたらされるファージ粒子はpIII-negを組み込み、これはファージ感染力およびファージ増殖の阻害をもたらし、それゆえに、かかる活性を呈示しない進化した産物と比較して、望まれないオフターゲット活性を呈示する進化産物をコードするいずれかのファージゲノムにとっての選択的な不利を構成する。いくつかの態様では、連続的進化実験の間に、二重セレクション戦略が適用され、これにおいては、ポジティブセレクションおよびネガティブセレクションコンストラクト両方が宿主細胞内に存在する。いくつかのかかる態様では、ポジティブおよびネガティブセレクションコンストラクトは同じプラスミド上に在り、二重セレクションアクセサリプラスミドともまた言われる。

20

【0593】

例えば、いくつかの態様では、二重セレクションアクセサリプラスミドが採用され、標的核酸配列を含むプロモーターのコントロール下にpIIIコード配列を含むポジティブセレクションカセットと、オフターゲット核酸配列を含むプロモーターのコントロール下にpIII-negをコードするカセットを含むネガティブセレクションカセットとを含む。同時の二重セレクション戦略を用いる1つの利点は、セレクションストリンジェンシーが、ポジティブセレクションコンストラクトと比較してネガティブセレクションコンストラクトの活性または発現レベルに基づいてファインチューニングされ得るということである。二重セレクション戦略の別の利点は、セレクションが、所望のまたは望まれない活性の存在または不在ではなく、所望の活性対望まれない活性の比、それゆえに、それぞれのファージ粒子に組み込まれているpIIIおよびpIII-negのもたらされる比に依存するというところである。

30

【0594】

本発明のいくつかの側面は、pIIIのドミナントネガティブバリエーション (pIII-neg) を提供または利用する。これらの側面は、pIIIの2つのN末端ドメインと短縮された終結不能なC末端ドメインとを含むpIIIバリエーションが不活性であるのみならず、pIIIのドミナントネガティブバリエーションであるという驚くべき発見に基づく。pIIIの2つのN末端ドメインと短縮された終結不能なC末端ドメインとを含むpIIIバリエーションはBennett, N. J.; Rakonjac, J., Unlocking of the filamentous bacteriophage virion during infection is mediated by the C domain of pIII. *Journal of Molecular Biology* 2006, 356 (2), 266-73 に記載されている；この内容全体は参照によって本明細書に組み込まれる。しかしながら、かかるpIIIバリエーションのドミナントネガティブ特性は以前には記載されていない。本発明のいくつかの側面は、本明細書において提供されるpIII-negバリエーションがファージ粒子に効率的に組み込まれるが、それは感染の間に侵入のための粒子のアンロックを触媒せ

40

50

ず、野生型pIIIが同じファージ粒子に存在する場合でさえも、それぞれのファージを非感染性にするという驚くべき発見に基づく。従って、例えば、認識が望まれない認識モチーフを含むプロモーターのコントロール下にpIII-negバリエーションをコードする核酸配列を含む発現コンストラクトを提供することによって、かかるpIII-negバリエーションは、PACEのコンテキストにおいてネガティブセレクション戦略を工夫するために有用である。他の態様では、pIII-negはポジティブセレクション戦略に用いられる。例えば、pIII-negコード配列が、どちらかのものの認識が所望であるヌクレアーゼ標的的部位または抑制因子認識部位を含むプロモーターによってコントロールされる発現コンストラクトを提供することによる。

【0595】

さらに、ポジティブおよびネガティブセレクション戦略が非DNA指向的な活性をファージ増殖効率に連結するように設計され得る。例えば、所望の標的プロテアーゼ切断部位に対するプロテアーゼ活性は、標的的部位を認識するプロテアーゼによって不活性化され得る遺伝子発現の抑制因子を工夫することによって、pIII発現に連結され得る。いくつかの態様では、pIII発現は、かかる抑制因子の結合部位を含むプロモーターによって駆動される。好適な転写抑制因子は当業者に公知であり、1つの例示的な抑制因子はラムダ抑制因子蛋白質であり、これはラムダプロモーターpRを効率的に抑制し、所望のプロテアーゼ切断部位を包含するように改変され得る（例えば、Sices, H. J.; Kristie, T. M., A genetic screen for the isolation and characterization of site-specific proteases. *Proc Natl Acad Sci U S A* 1998, 95 (6), 2828-33; およびSices, H. J.; Leusink, M. D.; Pacheco, A.; Kristie, T. M., Rapid genetic selection of inhibitor-resistant protease mutants: clinically relevant and novel mutants of the HIV protease. *AIDS Res Hum Retroviruses* 2001, 17 (13), 1249-55を参照する。これらのそれぞれの内容全体は参照によって本明細書に組み込まれる）。ラムダ抑制因子(cI)はN末端DNA結合ドメインおよびC末端二量体化ドメインを含有する。これらの2つのドメインはフレキシブルなリンカーによって接続される。効率的な転写抑制はcIの二量体化を要求し、それゆえに、二量体化および結合ドメインを接続するリンカーの切断は、cIの抑制因子活性を廃絶することをもたらす。

【0596】

いくつかの態様は、pIIIの発現を駆動するpRプロモーター(cI結合部位を含有する)を含むpIII発現コンストラクトを提供する。二量体化および結合ドメインを接続するリンカー配列中に所望のプロテアーゼ切断部位を含む改変されたcIと一緒に発現されたときには、cI分子は所望のプロテアーゼ活性の不在下ではpIII転写を抑制し、この抑制はかかる活性の存在下では廃絶され、それゆえにプロテアーゼ切断活性とpIII発現の増大との間の連結を提供し、これはPACEプロテアーゼポジティブセレクションにとって有用である。いくつかの態様は、PACE進化産物中の望まれないプロテアーゼ活性に対するネガティブセレクション戦略を提供する。いくつかの態様では、ネガティブセレクションは、cIによって抑制されるプロモーターのコントロール下にpIII-negコード核酸を含む発現カセットによって付与される。望まれないプロテアーゼ切断部位を含むcI抑制因子蛋白質と共発現されたときには、pIII-negの発現は、望まれない標的的部位に対するプロテアーゼ活性を呈示するプロテアーゼを発現するファージを抱く細胞において起こり、それゆえにかかる望まれない進化した産物をコードするファージに対してネガティブセレクションするであろう。プロテアーゼ標的の特異性についての二重セレクションは、cIによって抑制可能なpIIIおよびpIII-negをコードする発現コンストラクトを、異なるDNA標的配列を認識する直交的なcIバリエーションと共発現し、それゆえに互いに干渉することなしに同時発現を許すことによって達成され得る。二量体化特異性およびDNA結合特異性両方において直交的なcIバリエーションは当業者に公知である（例えば、Wharton, R. P.; Ptashne, M., Changing the binding specificity of a repressor by redesigning an alpha-helix. *Nature* 1985, 316 (6029), 601-5; およびWharton, R. P.; Ptashne, M., A new-specificity mutant of 434 repressor that defines an aminoacid-base pair contact. *Nature* 1987, 326 (6116), 88

10

20

30

40

50

8-91を参照する。これらのそれぞれの内容全体は参照によって本明細書に組み込まれる）。

【0597】

所望の活性を有する遺伝子産物の他のセレクションスキームは、当業者には周知であるか、または本開示から明らかであろう。本明細書において提供される連続的進化プロセスおよび方法に用いられ得るセレクション戦略は、2ハイブリッドスクリーニングに有用なセレクション戦略を包含するが、これらに限定されない。例えば、本明細書において他所でより詳細に記載されているT7 RNAPセレクション戦略は、プロモーター認識セレクション戦略の例である。2ハイブリッドアクセサリプラスミドセットアップはさらに蛋白質-蛋白質相互作用の進化を許し、感染性のウイルス粒子の生成のために要求される蛋白質、例えばpIIIの産生のために部位特異的リコンビナーゼ活性を要求するアクセサリプラスミドは、リコンビナーゼがいずれかの所望の標的部位を認識するように進化させられることを許す。2ハイブリッドセットアップまたは近縁の1ハイブリッドセットアップはさらにDNA結合蛋白質を進化させるために用いられ得、3ハイブリッドセットアップはRNA-蛋白質相互作用を進化させ得る。

10

【0598】

所望の小分子の存在に対して応答性であるプロモーターまたはリボスイッチ（例えば、遺伝子III発現/翻訳をコントロールする）によって、小分子を産生する生合成経路もまた進化させられ得る。例えば、ブタノールの存在下でのみ転写されるプロモーターがアクセサリプラスミド上において遺伝子IIIの上流に置かれて、ブタノール合成の酵素をコードする生合成経路を最適化し得る。ブタノール合成を高める活性を獲得した目当ての遺伝子を持つファージベクターは、かかる機能獲得を獲得していない進化するファージ集団中の他のファージよりも、選択的な利点を有するであろう。代替的には、例えば参照によって本明細書に組み込まれるBaker and Cornish, PNAS, 2002に記載されている化学的相補システムが、結合形成反応の能力がある個々の蛋白質または酵素を進化させるために用いられ得る（）。他の態様では、それ自体を特定の標的配列へとスプライシングするように設計されたトランススプライシングイントロンが、標的配列によって先行されたアクセサリプラスミドからの遺伝子IIIの後ろの半分のみを発現させること、および別の半分（トランススプライシングイントロンに融合させられている）をセレクションファージ上に置くことによって、進化させられ得る。首尾良いスプライシングは全長pIIIをコードするmRNAを再構成するであろう。プロテアーゼ特異性および活性は、所望のプロテアーゼ認識部位を含有するリンカーによって分離されたアクセサリプラスミドからの大きい蛋白質に融合させられたpIIIを発現することによって、進化させられ得る。セレクションファージによってコードされる活性なプロテアーゼによるリンカーの切断は感染性のpIIIをもたらすが、未切断のpIIIはブロックする蛋白質が原因で結合することができないであろう。さらに、例えばMalmborg and Borrebaeck 1997によって記載されている通り、標的抗原が細菌のF線毛に融合させられ得、野生型pIIIを結合からブロックする。抗原に特異的な抗体を呈するファージは結合および感染し得、ファージディスプレイにおいては>1000倍の濃縮を生んだ。いくつかの態様では、アクセサリプラスミドが、tolA受容体に接触および実際の感染を実行するための野生型pIIIを産生するように設計されるが（なぜなら、抗体-pIII融合体は良く結合するが、低い効率で感染するので）、セレクションファージはpIII-抗体融合蛋白質をコードするというこ

このシステムは連続的進化のために適応させられ得る。両方の型のpIIIを含有する子孫ファージは、抗体-抗原相互作用によってF線毛に固く吸着し、野生型pIIIはtolAに接触し、高効率の感染を媒介する。最初の抗体-抗原相互作用が弱いときに増殖を許すために、宿主細胞の混合物がラグーンに流入し得る：小さい画分は野生型pIIIを発現し、活性にかかわらずいずれかのセレクションファージを増殖させる能力がある感染した細胞のリザーバーとしての用をなすが、細胞の大多数は首尾良い相互作用を要求し、それらの結合親和性を改善するいずれかの変異体への「報酬」としての用をなす。この最後のシステムは、いくつかの態様では、標的病原体に対して有効である新たな抗体が、病原体それ自体が進化し得るよりも速く進化し得る。なぜな

20

30

40

50

ら、本明細書に記載されるPACEおよび他のシステムの進化速度は、ヒト特異的な病原体のもの、例えばヒトウイルスのものよりも高いからである。

【0599】

本明細書に記載されるセレクション戦略を行うために好適な条件付きプロモーターを設計するための方法および戦略は当業者に周知である。いくつかの例示的な設計戦略が図3Bに要約されている。細胞-細胞遺伝子伝播のために要求される遺伝子、例えばgIIIの発現を駆動する条件付きプロモーターを設計するための例示的な好適なセレクション戦略および方法の概略は、その全体が参照によって組み込まれるVidal and Legrain, Yeast n-hybrid review, Nucleic Acids Research 27, 919 (1999)を参照する。

【0600】

10

連続的な進化のための装置

【0601】

本発明は核酸の連続的な進化のための装置をもまた提供する。かかる装置のコア要素は、ウイルスベクターの集団が複製および増殖し得る宿主細胞のフローの生成を許すラグーンである。いくつかの態様では、ラグーンは細胞培養ベッセルを含み、ウイルスベクター、例えば目当ての遺伝子を含むファージベクターの活発に複製する集団と宿主細胞、例えば細菌宿主細胞の集団とを含む。いくつかの態様では、ラグーンは、ラグーンへのフレッシュな宿主細胞の導入のためのインフローと、ラグーンからの宿主細胞の取り除きのためのアウトフローとを含む。いくつかの態様では、インフローはフレッシュな宿主細胞の培養物を含むタービドスタットに接続される。いくつかの態様では、アウトフローは廃液ベッセルまたはシンクに接続される。いくつかの態様では、ラグーンは、ラグーンへの変異原の導入のためのインフローをさらに含む。いくつかの態様では、そのインフローは変異原の溶液を保持するベッセルに接続される。いくつかの態様では、本明細書において他所でより詳細に記載される通り、ラグーンは、ラグーンへの遺伝子発現の誘導因子の、例えば、変異導入を促進する遺伝子の発現を駆動する宿主細胞内で誘導性プロモーターを活性化する誘導因子の（例えば、変異導入プラスミドの一部としての）導入のためのインフローを含む。いくつかの態様では、そのインフローは、誘導因子の溶液、例えばアラビノースの溶液を含むベッセルに接続される。

20

【0602】

いくつかの態様では、ラグーンはウイルスベクターの集団を含む。いくつかの態様では、ラグーンはウイルスベクターの集団を含む。いくつかの態様では、ウイルスベクターは、ファージ、例えば、本明細書に記載される通り、感染性のウイルス粒子の生成のために要求される遺伝子を欠損しているM13ファージである。いくつかのかかる態様では、宿主細胞は、ファージ感染、複製、およびファージの増殖を受け入れる原核生物細胞、例えば、本明細書に記載される通り、感染性のウイルス粒子の生成のために要求される遺伝子を条件付きプロモーターのコントロール下に含むアクセサリプラスミドを含む宿主細胞である。

30

【0603】

いくつかの態様では、ラグーンは、宿主細胞のインフローおよびアウトフロー速度、変異原のインフロー、ならびに/または誘導因子のインフローの制御のためのコントローラを含む。いくつかの態様では、ファージの存在の視覚的インジケータ、例えば蛍光マーカーが追跡され、流量を管理するために用いられ、トータルの感染集団を一定に保つ。いくつかの態様では、視覚的マーカーは、ファージゲノムによってコードされる蛍光蛋白質またはファージゲノムによってコードされる酵素であり、これは、ひとたび宿主細胞において発現されると、宿主細胞の視覚的に検出可能な変化をもたらす。いくつかの態様では、システムがベクターウォッシュアウトのリスクなしに可能な限り速く流れることを保つように流量を調整するために、感染した細胞の視覚的追跡が用いられる。

40

【0604】

いくつかの態様では、感染性の粒子の生成のために要求される遺伝子の発現はタイトレーション可能である。いくつかの態様では、これは、ラグーンに追加されるアンヒドロテ

50

トラサイクリンの量に比例する、pIIIを産生するアクセサリープラスミドによって奏される。他のいくつかの態様では、かかるタイトレーション可能な発現コンストラクトは本明細書に記載される別のアクセサリープラスミドと組み合わせられ得、活性についての同時セレクションとpIIIのタイトレーション可能なコントロールとを許す。これは、さもなければラグーン中に生き残るには弱すぎる活性の進化を許し、中立的な浮動が局所適応度ピークトラップを脱出することを許す。いくつかの態様では、望まれない活性を罰することによって、本明細書に記載される連続的進化方法の間に、ネガティブセレクションが適用される。いくつかの態様では、これは、望まれない活性がpIII産生に干渉することを引き起こすことによって達成される。例えば、gIII RBSおよび/または開始コドンに対して相補的なアンチセンスRNAの発現は、ネガティブセレクションを適用する1つのやり方であり、プロテアーゼ（例えば、TEV）を発現することおよびpIII中にプロテアーゼ認識部位を操作することは、別のものである。

10

【0605】

いくつかの態様では、装置はタービドスタットを含む。いくつかの態様では、タービドスタットは細胞培養ベッセルを含み、これらの中には、例えば液体懸濁培養物としてフレッシュな宿主細胞の集団が在る。いくつかの態様では、タービドスタットはラグーンのインフローに接続されているアウトフローを含み、タービドスタットからラグーンへのフレッシュな細胞の導入を許す。いくつかの態様では、タービドスタットはタービドスタットへのフレッシュな培養培地の導入のためのインフローを含む。いくつかの態様では、インフローは滅菌培養培地を含むベッセルに接続される。いくつかの態様では、タービドスタットは、タービドスタットからの宿主細胞の取り除きのためのアウトフローをさらに含む。いくつかの態様では、そのアウトフローは廃液ベッセルまたはドレーンに接続される。

20

【0606】

いくつかの態様では、タービドスタットは、タービドスタットのフレッシュな宿主細胞の培養物の濁度を測定するための濁度計を含む。いくつかの態様では、タービドスタットは、無菌液体培地のインフローおよび廃液ベッセルへのアウトフローをタービドスタット内の培養液体の濁度に基づいて制御する（regulated）コントローラを含む。

【0607】

いくつかの態様では、ラグーンおよび/またはタービドスタットは一定のまたは間欠的な振盪のためのシェーカーまたはアジテーター、例えばシェーカー、ミキサー、スターラー、またはバブラーを含み、宿主細胞の集団が連続的にまたは間欠的に振盪および酸素供給されることを許す。

30

【0608】

いくつかの態様では、コントローラは、ラグーンからのアウトフローの速度と実質的に同じ（体積/体積）であるように、ラグーンへのフレッシュな宿主細胞のインフローの速度を制御する。いくつかの態様では、ラグーンへのフレッシュな宿主細胞のインフローの速度および/またはラグーンからの宿主細胞のアウトフローの速度は、連続的進化実験の時間中に実質的に一定であるように制御される。いくつかの態様では、インフローの速度および/またはアウトフローの速度は時間あたり約0.1ラグーン体積から時間あたり約25ラグーン体積である。いくつかの態様では、インフローの速度および/またはアウトフローの速度は時間あたりおよそ0.1ラグーン体積（lv/h）、およそ0.2lv/h、およそ0.25lv/h、およそ0.3lv/h、およそ0.4lv/h、およそ0.5lv/h、およそ0.6lv/h、およそ0.7lv/h、およそ0.75lv/h、およそ0.8lv/h、およそ0.9lv/h、およそ1lv/h、およそ2lv/h、およそ2.5lv/h、およそ3lv/h、およそ4lv/h、およそ5lv/h、およそ7.5lv/h、およそ10lv/h、または10lv/hよりも多くである。

40

【0609】

いくつかの態様では、インフローおよびアウトフロー速度は、ラグーンの宿主細胞の集団の定量的評価に基づいて、例えば細胞数、細胞密度、体積あたりのウェットバイオマス重量、濁度、または細胞増殖速度を測定することによって、コントロールされる。いく

50

つかの態様では、ラグーンインフローおよび/またはアウトフロー速度は、約 10^2 細胞/mlから約 10^{12} 細胞/mlという宿主細胞密度をラグーン中に維持するようにコントロールされる。いくつかの態様では、インフローおよび/またはアウトフロー速度は、約 10^2 細胞/ml、約 10^3 細胞/ml、約 10^4 細胞/ml、約 10^5 細胞/ml、約 5×10^5 細胞/ml、約 10^6 細胞/ml、約 5×10^6 細胞/ml、約 10^7 細胞/ml、約 5×10^7 細胞/ml、約 10^8 細胞/ml、約 5×10^8 細胞/ml、約 10^9 細胞/ml、約 5×10^9 細胞/ml、約 10^{10} 細胞/ml、約 5×10^{10} 細胞/ml、または 5×10^{10} 細胞/mlよりも多くという宿主細胞密度をラグーン中に維持するようにコントロールされる。いくつかの態様では、タービドスタットのフレッシュな宿主細胞の密度およびラグーンの宿主細胞の密度は実質的に同一である。

【0610】

10

いくつかの態様では、ラグーンインフローおよびアウトフロー速度は、実質的に一定数の宿主細胞をラグーン中に維持するようにコントロールされる。いくつかの態様では、インフローおよびアウトフロー速度は、実質的に一定の頻度のフレッシュな宿主細胞をラグーン中に維持するようにコントロールされる。いくつかの態様では、宿主細胞の集団は、ファージによって感染されていないフレッシュな宿主細胞を連続的に補充される。いくつかの態様では、補充は半連続的、または細胞集団へのフレッシュな細胞のバッチ投入によってである。

【0611】

いくつかの態様では、ラグーン体積はおよそ1mlからおよそ100mlである。例えば、ラグーン体積は、およそ1ml、およそ10ml、およそ50ml、およそ100ml、およそ200ml、およそ250ml、およそ500ml、およそ750ml、およそ1l、およそ2ml、およそ2.5l、およそ3l、およそ4l、およそ5l、およそ10l、およそ1ml-10ml、およそ10ml-50ml、およそ50ml-100、およそ100ml-250ml、およそ250ml-500ml、およそ500ml-1l、およそ1l-2l、およそ2l-5l、およそ5l-10l、およそ10-50l、およそ50-100l、または100lよりも多くである。

20

【0612】

いくつかの態様では、ラグーンおよび/またはタービドスタットは温度をコントロールするヒーターおよびサーモスタットをさらに含む。いくつかの態様では、ラグーンおよび/またはタービドスタットの温度は約4 から約55、好ましくは約25 から約39、例えば約37 であるようにコントロールされる。

【0613】

30

いくつかの態様では、インフロー速度および/またはアウトフロー速度は、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100、少なくとも200、少なくとも300、少なくとも400、少なくとも500、少なくとも600、少なくとも700、少なくとも800、少なくとも900、少なくとも1000、少なくとも1250、少なくとも1500、少なくとも1750、少なくとも2000、少なくとも2500、少なくとも3000、少なくとも4000、少なくとも5000、少なくとも7500、少なくとも10000、またはより多くの継続したウイルスベクターまたはファージライフサイクルにとって十分な時間に渡って、宿主細胞の集団のインキュベーションおよび補充を許すようにコントロールされる。いくつかの態様では、1回のファージライフサイクルにとって十分な時間は約10分である。

【0614】

40

よって、いくつかの態様では、進化手続き全体の時間は約12時間、約18時間、約24時間、約36時間、約48時間、約50時間、約3日、約4日、約5日、約6日、約7日、約10日、約2週、約3週、約4週、または約5週である。

【0615】

例えば、いくつかの態様では、100mlまたは1l体積のラグーンを含むPACE装置が提供され、ラグーンは、0.5l、1l、または3l体積のタービドスタットに、および変異導入プラスミドの誘導因子、例えばアラビノースを含むベッセルに接続され、ラグーンおよびタービドスタットは、E. coli細胞の懸濁培養物を約 5×10^8 細胞/mlの濃度で含む。いくつかの態様では、ラグーン中の細胞のフローは時間あたり約3ラグーン体積に制御される。いくつかの態様では、細胞は、連続的圧送によって、例えばラグーンの流体の所望の体積(例え

50

ば約100ml)に対応するラグーンベッセルの高さにセットされた廃液針を用いることによって、ラグーンから取り除かれる。いくつかの態様では、宿主細胞は、F'プラスミドを含むE. coli細胞、例えば遺伝子型F'proA⁺B⁺ (lacIZY) zzzf::Tn10(TetR)/ endA1 recA1 galE15 galK16 nupG rpsL lacIZYA araD 139 (ara,leu)7697 mcrA (mrr-hsdRMS-mcrBC) proBA::pir116 の細胞である。いくつかの態様では、セレクトションファージは、pIIIコード領域またはその一部が、目当ての遺伝子、例えば野生型ファージプロモーターによって駆動されるコード領域によって置き換えられたM13ゲノムを含む。いくつかの態様では、宿主細胞は、感染性のファージ粒子の生成のために要求される蛋白質をコードする遺伝子、例えばM13 pIIIが、本明細書において他所により詳細に記載される条件付きプロモーターから発現されるアクセサリプラスミドを含む。いくつかの態様では、宿主細胞は、変異導入プラスミド、例えば、アラビノース誘導性プロモーターなどの誘導性プロモーターから変異導入を促進する蛋白質を発現する変異導入プラスミドをさらに含む。いくつかの態様では、装置は、約3-7日に渡って、時間あたり約2-4ラグーン体積の細胞のフローの生成のために、タービドスタットにフレッシュな培地を提供するようにセットアップされる。

10

【0616】

ベクターおよび試薬

【0617】

本発明は本発明の連続的進化プロセスのためのウイルスベクターを提供する。いくつかの態様では、ファージによって支援される連続的進化のためのファージベクターが提供される。いくつかの態様では、感染性のファージ粒子の生成のために要求される少なくとも1つの遺伝子を欠損しているファージゲノムと進化させられるべき目当ての遺伝子とを含むセレクトションファージが提供される。

20

【0618】

例えば、いくつかの態様では、セレクトションファージは、感染性のM13ファージ粒子の生成のために要求される遺伝子、例えば全長gIIIを欠損しているM13ファージゲノムを含む。いくつかの態様では、セレクトションファージは、感染性のファージ粒子の生成のために要求される遺伝子を例外として、ファージライフサイクルのために要求される全ての他のファージ機能を提供するファージゲノムを含む。いくつかの態様では、gI、gII、gIV、gV、gVI、gVII、gVIII、gIX、およびgX遺伝子を含むが全長gIIIを含まないM13セレクトションファージが提供される。いくつかの態様では、セレクトションファージはgIIIの3'断片を含むが、全長gIIIを含まない。gIIIの3'末端はプロモーターを含み(図16を参照する)、このプロモーター活性を持ち続けることは、いくつかの態様では、gIII 3'プロモーターの直ちに下流にあるgVIの増大した発現、または宿主細胞におけるファージ遺伝子の発現レベルのより均衡のとれた(野生型ファージ様の)比にとって有益であり、これは、翻って、より効率的なファージ産生に至り得る。いくつかの態様では、gIII遺伝子の3'断片は3'gIIIプロモーター配列を含む。いくつかの態様では、gIIIの3'断片はgIIIの最後の180bp、最後の150bp、最後の125bp、最後の100bp、最後の50bp、または最後の25bpを含む。いくつかの態様では、gIIIの3'断片はgIIIの最後の180bpを含む。

30

【0619】

例えばgVIII 3'ターミネーターの下流にかつgIII 3'プロモーターの上流に挿入された、目当ての遺伝子をファージゲノム中に含むM13セレクトションファージが提供される。いくつかの態様では、目当ての遺伝子をファージゲノム中にクローニングするためのマルチクローニング部位、例えばgVIII 3'ターミネーターの下流にかつgIII 3'プロモーターの上流に挿入されたマルチクローニング部位(MCS)を含む、M13セレクトションファージが提供される。

40

【0620】

本発明のいくつかの側面は連続的進化手続きのためのベクターシステムを提供し、ウイルスベクター、例えばセレクトションファージ、およびマッチするアクセサリプラスミドを含む。いくつかの態様では、ファージに基づく連続的な指向的進化のためのベクターシ

50

ステムが提供され、これは、(a) ファージゲノムが、感染性のファージを生成するために要求される遺伝子を欠損している、進化させられるべき目当ての遺伝子を含むセレクションファージ；および (b) 条件付きプロモーターが、目当ての遺伝子によってコードされる遺伝子産物の機能によって活性化される、感染性のファージ粒子を生成するために要求される遺伝子を条件付きプロモーターのコントロール下に含むアクセサリプラスミドを含む。

【0621】

いくつかの態様では、セレクションファージは本明細書に記載されるM13ファージである。例えば、いくつかの態様では、セレクションファージは、ファージ粒子の生成のために要求される全ての遺伝子、例えばgI、gII、gIV、gV、gVI、gVII、gVIII、gIX、およびgX遺伝子を包含するが、全長gIII遺伝子を包含しないM13ゲノムを含む。いくつかの態様では、セレクションファージゲノムはF1またはM13複製起点を含む。いくつかの態様では、セレクションファージゲノムはgIII遺伝子の3'断片を含む。いくつかの態様では、セレクションファージはgIII 3'プロモーターの上流かつgVIII 3'ターミネーターの下流にマルチクロニング部位を含む。

10

【0622】

いくつかの態様では、セレクションファージは全長gVIを含まない。GVIはgIIIと類似に感染のために要求され、それゆえに、本明細書でgIIIについて記載されるセレクションのために類似の様式で用いられ得る。しかしながら、pIIIの連続的発現はいくつかの宿主細胞をM13による感染に対して耐性にするということが見出された。従って、pIIIは感染後にのみ産生されることということが望ましい。これは、本明細書に記載される通り、誘導性プロモーター、例えばアラビノース誘導性プロモーターのコントロール下にpIIIをコードする遺伝子を提供することと、タービドスタットではなく、感染が生ずるラグーンにおいて、または感染が生ずる前に別様に、誘導因子を提供することによって達成され得る。いくつかの態様では、感染性のファージの生成のために要求される複数の遺伝子、例えばgIIIおよびgVIがセレクションファージゲノムから取り除かれ、宿主細胞によって、例えば本明細書に記載されるアクセサリプラスミドによって提供される。

20

【0623】

ベクターシステムはさらにヘルパーファージを含み得、セレクションファージはファージ粒子の生成のために要求される全ての遺伝子を含むわけではなく、ヘルパーファージはセレクションファージのゲノムを相補し、その結果、ヘルパーファージゲノムおよびセレクションファージゲノムは一緒になってファージ粒子の生成のために要求される全ての遺伝子の少なくとも1つの機能的なコピーを含むが、感染性のファージ粒子の生成のために要求される少なくとも1つの遺伝子を欠損している。

30

【0624】

いくつかの態様では、ベクターシステムのアクセサリプラスミドは、感染性のファージの生成のために要求される遺伝子を条件付きプロモーターのコントロール下に含む発現カセットを含む。いくつかの態様では、ベクターシステムのアクセサリプラスミドは、その活性が目当ての遺伝子の産物の機能に依存する条件付きプロモーターのコントロール下にpIIIをコードする遺伝子を含む。

40

【0625】

いくつかの態様では、ベクターシステムは、変異導入プラスミド、例えば本明細書に記載されるアラビノース誘導性の変異導入プラスミドをさらに含む。

【0626】

いくつかの態様では、ベクターシステムは、セレクションファージのファージゲノムにまたはアクセサリプラスミドに含まれないいずれかのファージ遺伝子の発現コンストラクトを提供するヘルパープラスミドをさらに含む。

【0627】

種々の態様では、本明細書において連続的進化プロセスに用いられるベクターは、次の構成要素をいずれかの組み合わせで包含し得る：

50

【 0 6 2 8 】

gRNAバックボーン

【 0 6 2 9 】

gttttagagctagaaatagcaagttaaaataaggctagtcggttatcaacttgaaaaagtggcaccgagtcggtgctttt
ttt (配列番号132)

【 0 6 3 0 】

T7 RNAポリメラーゼ

【 0 6 3 1 】

【 化 7 6 】

MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARF 10
RKMFERQLKAGEVADNAAAKPLITLLPKMIARINDWFEEVKAKRGKRPTAFQFLQEIK
PEAVAYITIKTTLACLTSADNTTVQAVASAIGRAIEDEARFGRIRDLKAKHFKNVVEEQL
NKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKEDSIHVGVRCIEMLIESTGMV
SLHRQNAGVVGQDSETIELAPEYAEAIATRAGALAGISPMFQPCVPPKPWTGITGGGY
WANGRRPLALVRTHSKKALMRYEDVYMPEVYKAINIAQNTAWKINKKVLAVANVITK
WKHCPVEDIPAIEREELPMKPEDIDMNPEALTAWKRAAAAVYRKDKARKSRRISLEFML
EQANKFANHKAIWFPYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGY 20
WLKIHGANCAGVDKVPFPERIKFIEENHENIMACAKSPLNTWWAEQDSPFCFLAFCFE
YAGVQHHGLSYNCSLPLAFDGSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYGIVAK
KVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYGVTRSVTKRSV
MTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGYMAKLIWESVSVTVVAA
VEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVWVTPDGFVWQEYKPIQTRLNL
MFLGQFRLQPTINTNKDSEIDAHKQESGIAPNFVHSQDGSHLRKTVVWAHEKYGIESFA
LIHDSFGTIPADAANLFKAVRETMVDTYESCDVLADFYDQFADQLHESQLDKMPALPA
KGNLNLRDILESDFAF* (配列番号 133) 30

【 0 6 3 2 】

デグロンタグ

【 0 6 3 3 】

AANDENYNYALAA (配列番号134)

【 0 6 3 4 】

融合配列は、

【 0 6 3 5 】

【化 7 7】

MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARF
 RKMFERQLKAGEVADNAAAKPLITLLPKMIARINDWFEEVKAKRGKRPTAFQFLQEIK
 PEAVAYITIKTTLACLTSADNTTVQAVASAIGRAIEDEARFGRIRDLKAKHFKNVEEQL
 NKRVGHVYKKAQVVEADMLSKGLLGGEAWSSWHKEDSIHVGVRCEMLIESTGMV
 SLHRQAGVVGQDSETIELAPEYAEAIATRAGALAGISPMFQPCVVPKPWTGITGGGY
 WANGRRPLALVRTHSKKALMRYEDVYMPEVYKAINIAQNTAWKINKKVLAVANVITK
 WKHCPVEDIPAIEREELPMKPEDIDMNPEALTAWKRAAAAVYRKDKARKSRRISLEFML
 EQANKFANHKAIFWPYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGY
 WLKIHGANCAGVDKVPFPERIKFIEENHENIMACAKSPLNTWWAEQDSPFCFLAFCFE
 YAGVQHHGLSYNCSLPLAFDGSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYGIVAK
 KVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYGVTRSVTKRSV
 MTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGYMAKLIWESVSVTVVAA
 VEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHWVTPDGFVWQEYKKPIQTRLNL
 MFLGQFRLQPTINTNKDSEIDAHKQESGIAPNFVHSQDGSHLRKTVVWAHEKYGIESFA
 LIHDSFGTIPADAANLFKAVRETMVDTYESCDVLADFYDQFADQLHESQLDKMPALPA
 KGNLNLRDILESDFAFWTRAANDENYNYALAA* (配列番号 135)

10

20

である。

【 0 6 3 6 】

DnaE インテイン (XTEN リンカー を介するデアミナーゼへの融合体)

【 0 6 3 7 】

【化 7 8】

CLSYETEILTVEYGLLPIGKIVEKRIECTVYSVDNNGNIYTQPVAQWHDRG
 EQEVFEYCLEDGSLIRATKDHKFM TVDGQMLPIDEIFERELDLMRVDNLPN* (配列番号
 136)

30

【 0 6 3 8 】

APOBECへの融合体

【 0 6 3 9 】

【化79】

MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELKTCCLYEINWGGRHSIW
RHTSQNTNKHVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSRAITEFLSRYPHTL
FIYIARLYHHADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFVNYSNEAHWPRYPH
LWVRLYVLELYCIILGLPPCLNILRRKQPQLTFFTIALQSCHYQRLPPHILWATGLKSGSE
TPGTSESATPECLSYETEILTVEYGLLPKIVEKRIECTVYSVDNNGNIYTQPVAQWHDR
GEQEVFEYCLEDGSLIRATKDHKFMVTDGQMLPIDEIFERELDLMRVDNLPN* (配列番号
137)

10

【0640】

cas9に融合させられたCインテイン

【0641】

MIKIATRKYLGKQNVYDIGVERDHNFKNGFIASNC (配列番号138)

【0642】

cas9への融合体

【0643】

20

【化 8 0】

MIKIATRKYLGKQNVYDIGVERDHNFALKNGFIASNCFNKYSIGLAIGTNS
 VGWAVITDEYKVPSSKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTR
 RKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTI
 YHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQ
 LFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIASLGLTPNFKSN
 FDLAEDAQLSKDQYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKA
 PLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYK
 FIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKD
 NREKIEKILTRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGGASAQSFIERM
 TNFDKNLPNEKVLPHKSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFK
 TNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED
 ILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLLINGIRDQK
 SGKTIIDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIK
 KGILQTVKVVDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIIEGKELGS
 QILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVAIVPQSFLKDDSI
 DNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLS
 ELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKD
 FQFYKVINNYHHAHDAYLNAVVGTAIHKKYPKLESEFVYGDYKVVYDVRKMIKSEQ
 EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLS
 MPQVNIVKKTEVQTTGGFSKESILPKRNSDKLIARKKDWDPKKGFFSPTVAYSVLVV
 AKVEKGKSKKLKSVKELLGITIMERSSEKPNIDFLEAKGYKEVKKDLIILPKYSLFELE
 NGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLGSPEDNEQQLFVEQHKH
 YLDEIIEQISEFSKRVLADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKY
 FDTTIDRKRYTSTKEVLDATLIHQSIITGLYETRIDLSQLGGDSGGSMTNLSDIIEKETGKQ
 LVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLTSDAPEYKPWALVIQDS
 NGENKIKML* (配列番号 139)

10

20

30

DNAまたはRNAを編集する

【0644】

本開示のいくつかの側面は、本明細書に記載される塩基編集因子を用いて核酸を編集する
 ための方法を提供する。いくつかの態様では、方法は、核酸の核酸塩基（例えば、二本
 鎖DNA配列の塩基対）を編集する方法である。いくつかの態様では、方法は：a）核酸（例
 えば、二本鎖DNA配列）の標的領域を、塩基編集因子（例えば、アデノシンデアミナーゼ
 に融合させられたCas9ドメイン）とガイド核酸（例えばgRNA）とを含む複合体と接触させ
 、標的領域は標的化された核酸塩基対を含むステップ、b）前記標的領域の鎖分離を誘導
 するステップ、c）標的領域の一本鎖中の前記標的核酸塩基対の第1の核酸塩基を第2の核
 酸塩基に変換するステップ、およびd）前記標的領域の1つの鎖だけを切り、第1の核酸塩
 基の塩基に対して相補的な第3の核酸塩基が、第2の核酸塩基に対して相補的な第4の核酸
 塩基によって置き換えられるステップを含む。いくつかの態様では、方法は20%未満のイン
 デル形成を核酸にもたらず。いくつかの態様では、ステップbが省かれるということは

40

50

了解されるはずである。いくつかの態様では、第1の核酸塩基はアデニンである。いくつかの態様では、第2の核酸塩基は脱アミノ化されたアデニン、またはイノシンである。いくつかの態様では、第3の核酸塩基はチミンである。いくつかの態様では、第4の核酸塩基はシトシンである。いくつかの態様では、方法は19%、18%、16%、14%、12%、10%、8%、6%、4%、2%、1%、0.5%、0.2%未満、または0.1%未満のインデル形成をもたらす。いくつかの態様では、方法は、第2の核酸塩基を第4の核酸塩基に対して相補的である第5の核酸塩基によって置き換えることをさらに含み、それによって、意図される編集された塩基対を生成する（例えば、A:T G:C）。いくつかの態様では、第5の核酸塩基はグアニンである。いくつかの態様では、意図される塩基対の少なくとも5%が編集される。いくつかの態様では、意図される塩基対の少なくとも10%、15%、20%、25%、30%、35%、40%、45%、または50%が編集される。

10

【0645】

いくつかの態様では、標的ヌクレオチドの意図される産物対意図されない産物の比は少なくとも2:1、5:1、10:1、20:1、30:1、40:1、50:1、60:1、70:1、80:1、90:1、100:1、もしくは200:1、またはより多くである。いくつかの態様では、意図される点変異対インデル形成の比は1:1、10:1、50:1、100:1、500:1、もしくは1000:1よりも多大、またはより多くである。いくつかの態様では、切られる一本鎖（ニックング鎖）はガイド核酸にハイブリダイゼーションさせられる。いくつかの態様では、切られる一本鎖は第1の核酸塩基を含む鎖とは反対である。いくつかの態様では、塩基編集因子はCas9ドメインを含む。いくつかの態様では、第1の塩基はアデニンであり、第2の塩基はG、C、A、またはTではない。いくつかの態様では、第2の塩基はイノシンである。いくつかの態様では、第1の塩基はアデニンである。いくつかの態様では、第2の塩基はG、C、A、またはTではない。いくつかの態様では、第2の塩基はイノシンである。いくつかの態様では、塩基編集因子は編集された鎖の塩基除去修復を阻害する。いくつかの態様では、塩基編集因子は非編集鎖を保護するかまたは結合する。いくつかの態様では、塩基編集因子はUGI活性を含む。いくつかの態様では、塩基編集因子は触媒的に不活性なイノシン特異的ヌクレアーゼを含む。いくつかの態様では、塩基編集因子はニッカーゼ活性を含む。いくつかの態様では、意図される編集された塩基対はPAM部位の上流にある。いくつかの態様では、意図される編集された塩基対はPAM部位の1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、または20ヌクレオチド上流にある。いくつかの態様では、意図される編集された塩基対はPAM部位の下流にある。いくつかの態様では、意図される編集された塩基対はPAM部位の1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、または20ヌクレオチド下流（downstream stream）にある。いくつかの態様では、方法は古典的な（例えばNGG）PAM部位を要求しない。いくつかの態様では、核酸塩基編集因子はリンカーを含む。いくつかの態様では、リンカーは長さが1-25アミノ酸である。いくつかの態様では、リンカーは長さが5-20アミノ酸である。いくつかの態様では、リンカーは長さが10、11、12、13、14、15、16、17、18、19、または20アミノ酸である。いくつかの態様では、標的領域は標的ウィンドウを含み、標的ウィンドウは標的核酸塩基対を含む。いくつかの態様では、標的ウィンドウは1-10ヌクレオチドを含む。いくつかの態様では、標的ウィンドウは長さが1-9、1-8、1-7、1-6、1-5、1-4、1-3、1-2、または1ヌクレオチドである。いくつかの態様では、標的ウィンドウは長さが1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、または20ヌクレオチドである。いくつかの態様では、意図される編集された塩基対は標的ウィンドウ内にある。いくつかの態様では、標的ウィンドウは意図される編集された塩基対を含む。いくつかの態様では、方法は本明細書において提供される塩基編集因子のいずれかを用いて実行される。いくつかの態様では、標的ウィンドウは脱アミノ化ウィンドウである。

20

30

40

【0646】

いくつかの態様では、本開示はヌクレオチドを編集するための方法を提供する。いくつかの態様では、本開示は二本鎖DNA配列の核酸塩基対を編集するための方法を提供する。いくつかの態様では、方法は、a) 二本鎖DNA配列の標的領域を、塩基編集因子とガイド核

50

酸（例えばgRNA）とを含む複合体と接触させ、標的領域は標的核酸塩基対を含むこと、b）前記標的領域の鎖分離を誘導すること、c）標的領域の一本鎖中の前記標的核酸塩基対の第1の核酸塩基を第2の核酸塩基に変換すること、d）前記標的領域の1つの鎖だけを切り、第1の核酸塩基の塩基に対して相補的な第3の核酸塩基が、第2の核酸塩基に対して相補的な第4の核酸塩基によって置き換えられ、第2の核酸塩基が、第4の核酸塩基に対して相補的である第5の核酸塩基によって置き換えられることを含み、それによって意図される編集された塩基対を生成し、意図される編集された塩基対を生成する効率は少なくとも5%である。いくつかの態様では、ステップbは省かれるということは了解されるはずである。いくつかの態様では、意図される塩基対の少なくとも5%が編集される。いくつかの態様では、意図される塩基対の少なくとも10%、15%、20%、25%、30%、35%、40%、45%、または50%が編集される。いくつかの態様では、方法は、19%、18%、16%、14%、12%、10%、8%、6%、4%、2%、1%、0.5%、0.2%未満、または0.1%未満のインデル形成を引き起こす。いくつかの態様では、標的ヌクレオチドにおける意図される産物対意図されない産物の比は少なくとも2:1、5:1、10:1、20:1、30:1、40:1、50:1、60:1、70:1、80:1、90:1、100:1、もしくは200:1、またはより多くである。いくつかの態様では、意図される点変異対インデル形成の比は1:1、10:1、50:1、100:1、500:1、もしくは1000:1よりも多大、またはより多くである。いくつかの態様では、切られる一本鎖はガイド核酸にハイブリダイゼーションさせられる。いくつかの態様では、切られる一本鎖は、第1の核酸塩基を含む鎖とは反対である。いくつかの態様では、第1の塩基はアデニンである。いくつかの態様では、第2の核酸塩基はG、C、A、またはTではない。いくつかの態様では、第2の塩基はイノシンである。いくつかの態様では、塩基編集因子は編集された鎖の塩基除去修復を阻害する。いくつかの態様では、塩基編集因子は非編集鎖を保護する（例えば、塩基除去修復を形成する）かまたは結合する。いくつかの態様では、核酸塩基編集因子はUGI活性を含む。いくつかの態様では、塩基編集因子は触媒的に不活性なイノシン特異的ヌクレアーゼを含む。いくつかの態様では、核酸塩基編集因子はニッカーゼ活性を含む。いくつかの態様では、意図される編集された塩基対はPAM部位の上流にある。いくつかの態様では、意図される編集された塩基対はPAM部位の1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、または20ヌクレオチド上流にある。いくつかの態様では、意図される編集された塩基対はPAM部位の下流にある。いくつかの態様では、意図される編集された塩基対はPAM部位1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、または20ヌクレオチド下流（downstream stream）にある。いくつかの態様では、方法は古典的な（例えばNGG）PAM部位を要求しない。いくつかの態様では、核酸塩基編集因子はリンカーを含む。いくつかの態様では、リンカーは長さが1-25アミノ酸である。いくつかの態様では、リンカーは長さが5-20アミノ酸である。いくつかの態様では、リンカーは長さが10、11、12、13、14、15、16、17、18、19、または20アミノ酸である。いくつかの態様では、標的領域は標的ウィンドウを含み、標的ウィンドウは標的核酸塩基対を含む。いくつかの態様では、標的ウィンドウは1-10ヌクレオチドを含む。いくつかの態様では、標的ウィンドウは長さが1-9、1-8、1-7、1-6、1-5、1-4、1-3、1-2、または1ヌクレオチドである。いくつかの態様では、標的ウィンドウは長さが1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、または20ヌクレオチドである。いくつかの態様では、意図される編集された塩基対は標的ウィンドウ内に存在する。いくつかの態様では、標的ウィンドウは意図される編集された塩基対を含む。いくつかの態様では、核酸塩基編集因子は本明細書において提供される塩基編集因子のいずれか1つである。

10

20

30

40

50

【0647】

別の態様では、本開示は編集方法を提供し、DNAまたはRNA分子を本明細書において提供される塩基編集因子のいずれかおよび少なくとも1つのガイド核酸（例えばガイドRNA）と接触させることを含み、ガイド核酸（例えばガイドRNA）は約15-100ヌクレオチドの長さであり、標的配列に対して相補的である少なくとも10個の一続きのヌクレオチドの配列を含む。いくつかの態様では、標的配列の3'末端は古典的なPAM配列（NGG）に直ちに隣接

する。いくつかの態様では、標的配列の3'末端は古典的なPAM配列(NGG)に直ちに隣接はしない。いくつかの態様では、標的配列の3'末端はAGC、GAG、TTT、GTG、またはCAA配列に直ちに隣接する。

【0648】

いくつかの態様では、標的DNA配列は疾患または異常に関連する配列を含む。いくつかの態様では、標的DNA配列は疾患または異常に関連する点変異を含む。いくつかの態様では、融合蛋白質(例えば、アデノシンデアミナーゼおよびCas9ドメインを含む)または複合体の活性は点変異の修正をもたらす。いくつかの態様では、標的DNA配列は疾患または異常に関連するG A点変異を含み、変異体A塩基の脱アミノ化は、疾患または異常に関連しない配列をもたらす。いくつかの態様では、標的DNA配列は蛋白質をコードし、点変異はコドンにあり、野生型コドンと比較して変異体コドンによってコードされるアミノ酸の変化をもたらす。いくつかの態様では、変異体Aの脱アミノ化は変異体コドンによってコードされるアミノ酸の変化をもたらす。いくつかの態様では、変異体Aの脱アミノ化は野生型アミノ酸をコードするコドンをもたらす。いくつかの態様では、接触させることは対象において *in vivo* である。いくつかの態様では、対象は疾患または異常を有するか、またはそれと診断されている。いくつかの態様では、疾患または異常はフェニルケトン尿症、フォン・ヴィレブランド病(vWD)、変異体PTENもしくはBRCA1に関連する新生物疾患、またはリ・フラウメニ症候群である。本明細書に記載される塩基編集因子を用いて処置され得る例示的な疾患および異常のリストが表1に示されている。表1は、標的遺伝子、修正されるべき変異、関する疾患、および随伴するプロトスペーサーおよびPAMのヌクレオチド配列を包含する。

【0649】

表1 - 本明細書に記載される塩基編集因子によって処置され得る例示的な疾患のリスト。プロトスペーサー内の編集されるべきAは下線によって指示されており、PAMは太字で指示されている。

【表1】

標的遺伝子	変異	ATCC細胞株	疾患	プロトスペーサーおよびPAM
PTEN	Cys136Tyr	HTB-128	癌素因	TATATGCATATTTATTACATCGG(配列番号85)
PTEN	Arg233Ter	HTB-13	癌素因	CCGTCATGTGGGTCCTGAATTGG(配列番号86)
TP53	Gln258Lys	HTB-65	癌素因	ACACTGAAAGACTCCAGGTCAGG(配列番号87)
BRCA1	Gly1738Arg	NA	癌素因	GTCAGAAAGAGATGTGGTCAATGG(配列番号88)
BRCA1	4097-1G>A	NA	癌素因	TTTAAAGTGAAGCAGCATCTGGG(配列番号89) ATTTAAAGTGAAGCAGCATCTGG(配列番号90)
PAH	Thr380Met	NA	フェニルケトン尿症	ACTCCATGACAGTGTAAATTTGG(配列番号91)
VWF	Ser1285Phe	NA	フォン・ヴィレブランド病(血友病)	GCCTGGAGAAGCCATCCAGCAGG(配列番号92)
VWF	Arg2535Ter	NA	フォン・ヴィレブランド病(血友病)	CTCAGACACACTCATTGATGAGG(配列番号93)
TP53	Arg175His	HCC1395	リ・フラウメニ症候群	GAGGCCTGCCCCACCATGAGCG(配列番号94)

106501

【0650】

いくつかの態様は、本明細書において提供される進化した塩基編集因子を用いるための方法を提供する。いくつかの態様では、塩基編集因子は、標的核酸塩基、例えばC残基を脱アミノ化することによって核酸に点変異を導入するために用いられる。いくつかの態様では、標的核酸塩基の脱アミノ化は、遺伝子の欠陥の修正、例えば遺伝子産物の機能喪失

10

20

30

40

50

に至る点変異の修正をもたらす。いくつかの態様では、遺伝子の欠陥は、疾患または異常、例えば、例えば1型糖尿病などのリソソーム蓄積症または代謝疾患に関連する。いくつかの態様では、本明細書において提供される方法は、疾患または異常に関連する遺伝子産物をコードする遺伝子またはアレルに、失活させる点変異を導入するために用いられる。例えば、いくつかの態様では、（例えば増殖性疾患の処置において）失活させる点変異を癌遺伝子に導入するためにDNA編集融合蛋白質を採用する方法が本明細書において提供される。失活させる変異は、いくつかの態様では、コード配列中に未成熟終止コドンを生じ得、これは、短縮された遺伝子産物、例えば全長蛋白質の機能を欠く短縮された蛋白質の発現をもたらす。

【0651】

いくつかの態様では、本明細書において提供される方法の目的は、機能不全遺伝子の機能をゲノム編集によって回復させることである。本明細書において提供される核酸塩基編集蛋白質は、遺伝子編集に基づくヒト治療学について、*in vitro*で、例えばヒト細胞培養物の疾患関連変異を修正することによって、検証され得る。本明細書において提供される核酸塩基編集蛋白質、例えば核酸によってプログラム可能なDNA結合蛋白質（例えば、Cas9）とアデノシンデアミナーゼドメインとを含む融合蛋白質は、いずれかの1個のG AまたはC T点変異を修正するために用いられ得るということは当業者によって理解されるであろう。第1のケースでは、変異体AからIへの脱アミノ化は変異を修正し、後者のケースでは、変異体Tと塩基対形成しているAの脱アミノ化、次に複製のラウンドが変異を修正する。修正され得る例示的な点変異は表1に列記されている。

【0652】

疾患関連遺伝子およびアレルの点変異の首尾良い修正は、治療学および基礎研究に用途を有する遺伝子修正のための新たな戦略を開く。核酸によってプログラム可能なDNA結合蛋白質とアデノシンデアミナーゼドメインとの開示される融合体のような、部位特異的1塩基改変システムは、「逆」遺伝子治療にもまた用途を有し、ある種の遺伝子機能が意図的に抑圧または廃絶される。これらのケースでは、蛋白質の不活性化する変異または蛋白質の機能を阻害する変異に至る残基を部位特異的に変異させることは、蛋白質機能を廃絶または阻害するために用いられ得る。

【0653】

処置の方法

本開示は、本明細書において提供されるDNA編集融合蛋白質によって修正され得る点変異に関連するかまたはそれによって引き起こされる疾患と診断された対象の処置のための方法を提供する。例えば、いくつかの態様では、かかる疾患、例えば上に記載されている点変異に関連する癌を有する対象に、点変異を修正するかまたは失活させる変異を疾患関連遺伝子に導入するアデノシンデアミナーゼ融合蛋白質の有効な量を投与することを含む方法が提供される。いくつかの態様では、疾患は増殖性疾患である。いくつかの態様では、疾患は遺伝子疾患である。いくつかの態様では、疾患は新生物疾患である。いくつかの態様では、疾患は代謝疾患である。いくつかの態様では、疾患はリソソーム蓄積症である。点変異を修正することまたは失活させる変異を疾患関連遺伝子に導入することによって処置され得る他の疾患は当業者には公知であろう。本開示はこれについて限定されない。

【0654】

本開示は、追加の疾患または異常、例えば、デアミナーゼによって媒介される遺伝子編集によって修正され得る点変異に関連するかまたはそれによって引き起こされる疾患または異常の処置のための方法を提供する。いくつかのかかる疾患は本明細書に記載され、本明細書において提供される戦略および融合蛋白質によって処置され得る追加の好適な疾患は本開示に基づいて当業者には明らかであろう。例示的な好適な疾患および異常が下に列記されている。それぞれの配列中の具体的な位置または残基の付番は、用いられる特定の蛋白質および付番スキームに依存するということは理解されるであろう。付番は例えば成熟蛋白質の前駆体および成熟蛋白質それ自体において異なり得、種から種の配列の違いが付番を左右し得る。当業者は、当分野において周知の方法によって、例えば配列アライメ

10

20

30

40

50

ントおよび相同残基の決定によって、いずれかの相同的蛋白質およびそれぞれのコード核酸中のそれぞれの残基を同定することができるであろう。例示的な好適な疾患および異常は、限定なしに：2-メチル-3-ヒドロキシ酪酸尿症；3ベータ-ヒドロキシステロイドデヒドロゲナーゼ欠損症；3-メチルグルタコン酸尿症；3-オキソ-5アルファ-ステロイドデルタ4-デヒドロゲナーゼ欠損症；46、XY性腺異形成1、3、および5型；5-オキソピロリナーゼ欠損症；6-ピルボイル-テトラヒドロプテリンシンターゼ欠損症；アールスコッグ症候群；アーゼ症候群；軟骨無発生症2型；色盲2および7；後天性QT延長症候群；肢先端脳梁症候群、シンゼル型；先端大腿骨頭異形成；先端異骨症2、ホルモン抵抗性ありまたはなし；手掌角化症；先端短肢異形成症；ACTH非依存性大結節性副腎皮質過形成2；活性化PI3Kデルタ症候群；急性間欠性ポルフィリン症；アシルCoAデヒドロゲナーゼファミリー、メンバー9欠損症；アダムズ・オリバー症候群5および6；アデニンホスホリボシルトランスフェラーゼ欠損症；アデニル酸キナーゼ欠損症；アデニロコハク酸リアーゼ欠損による溶血性貧血；思春期ネフロン癆；腎・肝・膵異形成；メッケル症候群7型；副腎白質ジストロフィー；成人接合部型表皮水疱症；表皮水疱症、接合部型、限局性バリエーション；成人型神経セロイドリポフスチン症；成人型神経セロイドリポフスチン症；眼球運動失行を伴う成人発症型失調症；ADULT症候群；無フィブリノゲン血症および先天性無フィブリノゲン血症；常染色体劣性無ガンマグロブリン血症2；加齢黄斑変性3、6、11、および12；アイカルディ・グティエール症候群1、4、および5；凍瘡状狼瘡1；アラジール症候群1および2；アレキサンダー病；アルカプトン尿症；アラン・ハーン・ダドリー症候群；先天性汎発性無毛症；アルパーズ脳症；アルファ-1-アンチトリプシン欠損症；常染色体優性、常染色体劣性、およびX連鎖劣性アルポート症候群；アルツハイマー病、家族性、3、瘵性対麻痺および失行を伴う；アルツハイマー病、1、3、および4型；低石灰化型および未成熟型、IIA1エナメル質形成不全；アミノアシラーゼ1欠損症；アーミッシュ小児てんかん症候群；アミロイド形成性トランスチレチンアミロイドーシス；アミロイド心筋症、トランスチレチン関連；心筋症；筋萎縮性側索硬化症1、6、15（前頭側頭型認知症ありまたはなし）、22（前頭側頭型認知症ありまたはなし）、および10型；TDP43封入体を伴う前頭側頭型認知症、TARDBP関連；アンダーマン症候群；アンダーセン・タウィル症候群；先天性QT延長症候群；貧血、非球状赤血球性溶血性、G6PD欠損が原因；エンジェルマン症候群；小脳症を伴う重症新生児発症脳症；自閉症感受性、X連鎖3；血管症、遺伝性、腎症、動脈瘤、および筋肉のけいれんを伴う；アンジオテンシンI変換酵素、良性の血清増加；無虹彩症、小脳性運動失調、および精神遅滞；アノキア；アンチトロンピンIII欠損症；性器異常およびステロイド合成異常を伴うアントレー・ピクスラー症候群；大動脈瘤、家族性胸部4、6、および9；胸部大動脈瘤および大動脈解離；全身性平滑筋機能障害症候群；もやもや病5；再生不良性貧血；見かけ上の鉍質コルチコイド過剰；アルギナーゼ欠損症；アルギニノサクシネートリアーゼ欠損症；アロマターゼ欠損症；催不整脈性右室心筋症5、8、および10型；原発性家族性肥大型心筋症；先天性多発性関節拘縮症、遠位、X連鎖；関節拘縮症腎臓機能障害胆汁うっ滞症候群；関節拘縮症、腎臓機能障害、および胆汁うっ滞2；アスパラギンシンターゼ欠損症；神経細胞移動異常；ビタミンE欠損を伴う運動失調；運動失調、感覚性、常染色体優性；毛細血管拡張性運動失調症候群；遺伝性癌素因症候群；無トランスフェリン血症；心房性細動、家族性、11、12、13、および16；心房中隔欠損2、4、および7（房室伝導障害ありまたはなし）；心房停止2；房室中隔欠損4；遺伝性眼球萎縮；ATR-X症候群；耳介顆状症候群2；自己免疫疾患、多系統、乳児性；自己免疫性リンパ増殖性症候群、1a型；常染色体優性低汗性外胚葉異形成症；常染色体優性進行性外眼筋麻痺、ミトコンドリアDNA欠失1および3を伴う；常染色体優性捻転ジストニア4；常染色体劣性中核性ミオパチー；常染色体劣性先天性魚鱗癬1、2、3、4A、および4B；常染色体劣性皮膚弛緩症1Aおよび1B型；常染色体劣性低発汗性外胚葉性異形成症候群；外胚葉異形成症11b；低汗/髪/歯型、常染色体劣性；常染色体劣性低リン血症性骨疾患；アクセンフェルト・リーガー症候群3型；ペインブリッジ・ローパー症候群；バナヤン・ライリー・ルバルカバ症候群；PTEN過誤腫瘍症候群；パイライト・ウィンター症候群1および2；パラカット症候群；バルデ・ビードル症候群1、11、16、および19；裸リンパ球症

10

20

30

40

50

候群2型、相補群E；パーター症候群出生前2型；パーター症候群3、低カルシウム尿症を伴う3、および4型；大脳基底核石灰化、特発性、4；連珠毛；良性家族性血尿症；良性家族性新生児けいれん1および2；けいれん、良性家族性新生児、1、および/またはミオキミア；けいれん、早期乳児てんかん性脳症7；良性家族性新生児乳児けいれん；良性遺伝性舞蹈病；心筋症を伴う良性肩甲骨腹膜筋ジストロフィー；ベルナル・スーリエ症候群、A1およびA2型（常染色体優性）；ベストロフィン症、常染色体劣性；ベータサラセミア；ベスレムミオパチーおよびベスレムミオパチー2；二角形結晶性角膜網膜ジストロフィー；胆汁酸合成異常、先天性、2；ピオチニダーゼ欠損症；パークバレル精神遅滞異形症候群；眼瞼裂狭小・眼瞼下垂・逆内眼角贅皮症候群；ブルーム症候群；ポージェソン・フォルスマン・レーマン症候群；プーシェノイハウザー症候群；短指症A1およびA2型；高血圧症を伴う短指症；出血を伴う脳小血管疾患；分岐鎖ケト酸デヒドロゲナーゼ欠損症；鰓症候群2および3；乳癌、早発性；乳-卵巣癌、家族性1、2、および4；脆性角膜症候群2；プロディミオパチー；気管支拡張症、高塩化汗ありまたはなし、3；Brown-Vialetto-Van Laere症候群およびBrown-Vialetto-Van Laere症候群2；ブルガダ症候群；ブルガダ症候群1；心室細動；発作性家族性心室細動；ブルガダ症候群およびブルガダ症候群4；QT延長症候群；突然心臓死；ブルズアイ黄斑ジストロフィー；スターガルト病4；桿体錐体ジストロフィー12；水疱性魚鱗癬赤紅皮症；Burn-Mckeown症候群；カンジダ症、家族性、2、5、6、および8；炭水化物欠乏糖蛋白質症候群IおよびII型；カルボニックアンヒドラーゼVA欠損が原因の高アンモニア血症；結腸癌；心不整脈；QT延長症候群、LQT1サブタイプ；心筋症、致死乳児性、チトクロムcオキシダーゼ欠損が原因；心臓・顔・皮膚症候群；心筋症；ダノン病；肥大型心筋症；左室非圧縮型心筋症；Carnevale症候群；カーニーコンプレックス、1型；カルニチン-アシルカルニチントランスポーター欠損症；カルニチンパルミトイルトランスフェラーゼI、II、II（遅発性）、およびII（乳児性）欠損症；白内障1、4、常染色体優性、常染色体優性、複数型、小角膜を伴う、coppock様、若年性、小角膜および糖尿病を伴う、および核、びまん性、非進行性；カテコラミン誘発多形性心室頻拍；尾部退行症候群；Cd8欠損症、家族性；セントラルコア病；染色体1、9および16のセントロメア不安定性および免疫不全；小脳性運動失調、乳児性、進行性の外眼筋麻痺および小脳性運動失調を伴う、精神遅滞、および平衡失調症候群2；脳アミロイド血管症、APP関連；脳常染色体優性および劣性動脈症、皮質下梗塞およびおよび白質脳症を伴う；脳海綿状奇形成2；脳・眼・顔・骨格症候群2；脳・眼・顔・骨格症候群症候群；脳網膜微小血管症、石灰化および嚢胞を伴う；神経セロイドリポフスチン症2、6、7、および10；Chediak-Higashi症候群、Chediak-Higashi症候群、成人型；シャルコー・マリー・トゥース病1B、2B2、2C、2F、21、2U（軸索性）、1C（脱髄性）、優性中間C、劣性中間A、2A2、4C、4D、4H、1F、1VF、およびX型；肩甲骨腹膜脊髄性筋萎縮症；遠位型脊髄性筋萎縮症、先天性非進行性；脊髄性筋萎縮症、遠位、常染色体劣性、5；CHARGE症候群；小児低ホスファターゼ症；成人低ホスファターゼ症；胆嚢炎；進行性家族性肝内胆汁うっ滞3；胆汁うっ滞、肝内、妊娠3；コレスタノール蓄積症；コレステロールモノオキシゲナーゼ（側鎖切断性）欠損症；軟骨異形成症、Blomstrand型；点状軟骨異形成症1、X連鎖劣性および2X連鎖優性；CHOPS症候群；慢性肉芽腫症、常染色体劣性、チトクロムb陰性、1および2型；Chudley-McCullough症候群；原発性線毛運動不全症7、11、15、20および22；シトルリン血症I型；シトルリン血症IおよびII型；鎖骨頭蓋形成不全；C様症候群；コケイン症候群、A型；補酵素Q10欠損症、原発性1、4、および7；コフィン・シリス/知的障害；コフィン・ローリー症候群；コーエン症候群；寒冷誘発性発汗症候群1；コール・カーペンター症候群2；混合性の細胞性および体液性免疫障害、肉芽腫を伴う；混合性のd-2-およびl-2-ヒドロキシグルタル酸尿症；混合性のマロン酸およびメチルマロン酸尿症；複合型酸化リン酸化異常1、3、4、12、15、および25；混合性の部分型および完全型17-アルファ-ヒドロキシラーゼ/17、20-リアーゼ欠損症；分類不能型免疫不全9；補体成分4部分欠損症、機能不全c1阻害因子が原因；補体B因子欠損症；桿体1色覚；桿体錐体ジストロフィー2および6；桿体錐体ジストロフィー、エナメル質形成不全症；先天性副腎皮質過形成および先天性副腎低形成、X連鎖；先天性血球減少性血小板減少症；先天性無虹彩症；

10

20

30

40

50

先天性中枢性低換気症候群；ヒルシスブルグ病3；先天性収縮性クモ膜症；四肢および顔面の先天性拘縮、筋緊張低下、および発達遅滞；先天性グリコシル化異常1B、1D、1G、1H、1J、1K、1N、1P、2C、2J、2K、1Im型；先天性赤血球造血性貧血、IおよびII型；顔面の先天性外胚葉異形成症；先天性赤血球造血性ポルフィリン症；先天性全身性リポジストロフィー2型；先天性心疾患、多発型、2；先天性心疾患；大動脈弓離断症；先天性脂肪腫過形成、血管奇形、および表皮母斑；非小細胞肺癌；卵巣の新生物；心臓伝導障害、非特異性；先天性微絨毛性萎縮症；先天性筋ジストロフィー；先天性筋ジストロフィー、部分的なLAMA2欠損が原因；脳および眼の異常を伴う先天性筋ジストロフィー-ジストログリカノパチー、A2、A7、A8、A11、およびA14型；精神遅滞を伴う先天性筋ジストロフィー-ジストログリカノパチー、B2、B3、B5、およびB15型；精神遅滞なしの先天性筋ジストロフィー-ジストログリカノパチー、B5型；先天性筋肥大・脳症候群；先天性筋無力症候群、アセタゾラミド応答性；線維タイプ不均等症を伴う先天性ミオパチー；先天性コロボーマ；先天停止性夜盲、1A、1B、1C、1E、1F、および2A型；コプロポルフィリア；扁平角膜2；角膜ジストロフィー、フックス内皮、4；角膜内皮ジストロフィー2型；角膜の脆弱性球状角膜（Corneal fragility keratoglobus）、青強膜および関節過度可動性；コルネリア・デ・ランゲ症候群1および5；冠状動脈性心疾患、常染色体優性2；冠状動脈性心疾患；高アルファリポタンパク血症2；他の脳奇形を伴う皮質形成異常、複雑型5および6；大脳皮質形成異常、後頭葉；コルチコステロイド結合グロブリン欠損症；コルチコステロンメチルオキシダーゼ2型欠損症；コステロ症候群；コーデン症候群1；扁平股；頭蓋骨幹異形成、常染色体優性；頭蓋骨癒合症1および4；頭蓋骨癒合症および歯の異常；クレアチン欠損症、X連鎖；Crouzon症候群；潜在眼球症候群；陰唇症、片側または両側；クッシング指節癒合症；皮膚悪性黒色腫1；骨ジストロフィーならびに重度の肺、胃腸、および泌尿器異常を伴う皮膚弛緩症；チアノーゼ、一過性新生児および非典型腎症性；嚢胞性線維症；シスチン尿症；チトクロムcオキシダーゼ1欠損症；チトクロムcオキシダーゼ欠損症；D-2-ヒドロキシグルタル酸尿症2；ダリエー病、分節型；難聴-迷路無発生-小耳-小歯（LAMM）；難聴、常染色体優性3a、4、12、13、15、常染色体優性非症候性感音性17、20、および65；難聴、常染色体劣性1A、2、3、6、8、9、12、15、16、18b、22、28、31、44、49、63、77、86、および89；難聴、蝸牛、近視および知的機能障害を伴う、前庭の関わりなし、常染色体優性、X連鎖2；2-メチルブチリルCoAデヒドロゲナーゼ欠損症；3-ヒドロキシアシルCoAデヒドロゲナーゼ欠損症；アルファ-マンノシダーゼ欠損症；芳香族L-アミノ酸デカルボキシラーゼ欠損症；ビスホスホグリセリン酸ムターゼ欠損症；ブチリルCoAデヒドロゲナーゼ欠損症；フェロキシダーゼ欠損症；ガラクトキナーゼ欠損症；グアニジノ酢酸メチルトランスフェラーゼ欠損症；ヒアルロノグルコサミニダーゼ欠損症；リボース-5-リン酸イソメラーゼ欠損症；ステロイド11-ベータ-モノオキシゲナーゼ欠損症；UDPグルコース-ヘキソース-1-リン酸ウリジリルトランスフェラーゼ欠損症；キサンチンオキシダーゼ欠損症；Dejerine-Sottas病；シャルコー・マリー・トゥース病、IDおよびIVF型；Dejerine-Sottas症候群、常染色体優性；樹状細胞、単球、Bリンパ球、およびナチュラルキラーリンパ球欠損症；Desbuquois異形成2；Desbuquois症候群；DFNA2非症候性難聴；視神経萎縮および難聴を伴う糖尿病および尿崩症；糖尿病、2型、およびインスリン依存性、20；ダイヤモンド・ブラックファン貧血1、5、8、および10；下痢3（ナトリウム分泌、先天性、症候群性）および5（タフティング腸疾患を伴う、先天性）；ジカルボン酸アミノ酸尿症；びまん性掌蹠角化症、ポスニア型；指腎脳症候群；ジヒドロプテリジンレダクターゼ欠損症；拡張型心筋症1A、1AA、1C、1G、1BB、1DD、1FF、1HH、1I、1KK、1N、1S、1Y、および3B；左室不全3；シトクロムp450オキシドレダクターゼ欠損が原因のステロイド合成異常；遠位関節喉頭症2B型；遠位遺伝性運動ニューロパチー2B型；遠位ミオパチーMarkesbery-Griggs型；遠位型脊髄性筋萎縮症、X連鎖3；二列睫毛-リンパ浮腫症候群；皮膚の不在を伴う優性ジストロフィー性表皮水疱症；優性遺伝性視神経萎縮症；ドンナイパロー症候群；ドーパミンベータヒドロキシラーゼ欠損症；ドーパミン受容体d2の脳内密度減少；Dowling-degos病4；Doyme蜂巣状網膜ジストロフィー；家族性ドルーゼン；Duane症候群2型；Dubin-Johnson症候群；デュシェンヌ筋ジスト

ロフィー；ベッカー筋ジストロフィー；異常フィブリノーゲン血症；先天性角化異常症、常染色体優性および常染色体劣性、3；先天性角化異常症、常染色体劣性、1、3、4、および5；先天性角化異常症、X連鎖；ジスキネジア、家族性、顔面けいれんを伴う；異常ブラスミノゲン血症；ジストニア2（捻転、常染色体劣性）、3（捻転、X連鎖）、5（ドーパ反応型）、10、12、16、25、26（ミオクロヌス性）；けいれん、良性家族性乳児性、2；早期乳児てんかん性脳症2、4、7、9、10、11、13、および14；非典型レット症候群；早期T細胞前駆体急性リンパ芽球性白血病；外胚葉異形成皮膚脆弱性症候群；外胚葉異形成合指症候群1；水晶体偏位、単独、常染色体劣性および優性；無指症、外胚葉異形成、および口唇口蓋裂症候群3；エーラース・ダンロス症候群7型（常染色体劣性）、古典型、2型（早老性）、ヒドロキシリジン欠乏症、4型、4型変異、およびテネイシンX欠乏症が原因である；イクスフェルト型先天性筋ジストロフィー；内分泌-脳骨異形成（Endocrine-cerebroosteodisplasy）；S錐体増強症候群；拡大前庭水道症候群；エンテロキナーゼ欠損症；表皮異形成疣贅型（verruciformis）；単発性表皮水疱症および四肢帯型筋ジストロフィー、斑状色素沈着を伴う単一性、幽門閉鎖症を伴う単一性、単一性、常染色体劣性、および幽門閉鎖を伴う；表皮剥離性掌蹠角皮症；家族熱性けいれん8；てんかん、小児期不在2、12（特発全身性、への感受性）5（夜間前頭葉）、夜間前頭葉1型、部分的、可変焦点を伴う、進行ミオクロヌス性3、およびX連鎖、可変の学習障害および行動異常を伴う；てんかん性脳症、小児期発症、乳児期早期、1、19、23、25、30および32；骨端異形成、多発、近視および伝導性難聴；エピソード失調症2型；偶発性疼痛症候群、家族性、3；エプスタイン症候群；フェヒトナー症候群；赤血球産生性プロトポルフィリン症；エストロゲン耐性；滲出性硝子体網膜症6；ファブリー病およびファブリー病、心臓性変異；H、VII、X、v、およびviii因子、2の合併欠乏症、xiii、サブユニット、欠乏症；家族性腺腫性ポリープ症1および3；じんましんおよび難聴を伴う家族性アミロイド腎症；家族性寒冷じんましん；小脳虫部の家族性形成不全；家族性良性天疱瘡；乳房の家族性がん；乳がん、への感受性；骨肉腫；膵臓がん3；家族性心筋症；家族性風邪自己炎症症候群2；家族性大腸がん；家族性滲出性硝子体網膜症、X連鎖；家族性片麻痺性片頭痛1型および2型；家族性高コレステロール血症；家族性肥大型心筋症1、2、3、4、7、10、23および24；家族性低カリウム血症？低マグネシウム血症；家族性低形成、糸球体腎炎；家族性乳児筋無力症；家族性若年性痛風；家族性地中海熱および家族性地中海熱、常染色体優性；家族性孔脳症；家族性晩発性皮膚ポルフィリン症；家族性肺毛細血管腫症；家族性腎糖尿病；家族性腎性低尿酸血症；家族性制限型心筋症1；家族性1型および3型の超リポ蛋白質血症；ファンコーニ貧血、補完グループE、I、N、およびO；ファンコーニピッケル症候群；ファビズム、への感受性；熱性けいれん、家族性、11；ファインゴールド症候群1；胎児ヘモグロビン量的形質遺伝子座1；FG症候群およびFG症候群4；眼球外筋の線維症、先天性、1、2、3a（眼球外病変を伴うまたはなしの）、3b；魚眼疾患；フレック角膜ジストロフィー；フローティングハーバー症候群；精神発達障害を伴うまたはなしの発話異常を伴う焦点発作；巣状分節性糸球体硬化症5；前脳の欠陥；フランクテルハール症候群；Borroni Di Rocco Crovato症候群；フラジエ症候群；ウィルムス腫瘍1；フリーマン・シールドン症候群；前頭骨幹端異形成1および3；前頭側頭型認知症；前頭側頭型認知症および/または筋萎縮性側索硬化症3および4；前頭側頭型認知症染色体3連鎖および前頭側頭型認知症ユビキチン陽性；フルクトース・ビホスファターゼ欠損症；フーマン症候群；ガンマアミノ酪酸トランスアミナーゼ欠損症；ガンストープ・ウォルフアルト症候群；ゴーシェ病1型および亜急性神経障害；進行性脊柱側弯症を伴う視線麻痺、遺伝子水平性；全身優性ジストロフィー性表皮水疱症；熱性けいれんプラス3を伴う遺伝子活性化てんかん、1型、2型；レノックスガスト型てんかん性脳症；巨大軸索型神経障害；グランツマン血栓症；緑内障1、開放角、e、F、およびG；緑内障3、原発性先天性、d；先天性緑内障、先天性緑内障、先天性緑内障；緑内障、一次開放角、若年発症；神経膠腫の感受性1；グルコーストランスポーター1型欠損症症候群；グルコース6リン酸輸送障害；GLUT1欠損症症候群2；てんかん、特発全身性、への感受性、12；グルタミン酸ホルミミノトランスフェラーゼ欠損症；グルタル酸血症IIAおよびIIB；グルタル酸尿症、1型；グルタチオン合成酵素欠損

症；グリコーゲン貯蔵疾患0（筋肉）、II（成人型）、IXa2、IXc、1A型；II型、IV型、IV型（肝臓の、および筋障害性合併型）、V型、およびVI型；ゴールドマン・ファール症候群；ゴードン症候群；ゴーリン症候群；全前脳症；全前脳症7；肉芽腫性疾患、慢性、X連鎖、変異；卵巣の顆粒膜細胞腫；グレー血小板症候群；Griscelli症候群3型；グレーノール角膜ジストロフィー1型；成長および精神遅滞、下顎顔面骨形成不全、小脳症、口蓋裂；下垂体異常を伴う成長ホルモン欠乏症；免疫不全を伴う成長ホルモン非感受性；GTPシクロヒドロラーゼI欠損症；ハイドゥーチェニー症候群；手足子宮症候群；聴覚障害；血管腫、毛細血管乳児；血液腫瘍；ヘモクロマトーシス1、2B、および3型；糖尿病7の微小血管合併症；トランスフェリン血清レベル量的形質遺伝子座2；ヘモグロビンH疾患、無欠失の；溶血性貧血、非球形性、グルコースリン酸イソメラーゼ欠損症が原因である；血球貪食性リンパ組織球増加症、家族性、2；血球貪食性リンパ組織球増加症、家族性、3；ヘパリン補因子II欠乏症；遺伝性腸性肢端皮膚炎；遺伝性乳がんおよび卵巣がん症候群；毛細血管拡張性運動失調症；遺伝性びまん性胃がん；スフェロイドを伴う遺伝性びまん性白質脳症；遺伝因子II、IX、VIII欠乏症疾患；遺伝性出血性毛細血管拡張症2型；無汗症を伴う疼痛への遺伝的非感受性；遺伝性リンパ浮腫I型；視神経萎縮を伴う遺伝性運動神経障害および感覚神経障害；早期呼吸不全を伴う遺伝性ミオパチー；遺伝性神経性筋萎縮症；遺伝性非ポリープ性結腸直腸腫瘍；リンチ症候群IおよびII；遺伝性膵炎；膵炎、慢性、への感受性；遺伝性感覚および自律神経障害IIB型およびIIA型；遺伝性の鉄芽球性貧血；ヘルマンスキー・パドラック症候群1、3、4、および6；内臓逆位、内臓、2、4、および6、常染色体；内臓逆位、内臓、X連鎖；異所形成；組織球性髄質網状症；組織球増殖症リンパ節腫大プラス症候群；ホロカルボキシラーゼ合成酵素欠損症、全前脳症2、3、7および9、およびホルト・オーラム症候群；MTHFR欠乏症が原因であるホモシステイン血症、CBS欠乏症、およびホモシステニン尿症、ピリドキシン応答性；コバラミン代謝における欠陥が原因であるホモシステニン尿症-巨赤芽球性貧血、cbIE補完型；ハウエル・エヴァンス（Howel-Evans）症候群；ハーラー症候群；ハチンソン・ギルフォード症候群；水頭症；高アンモニア血症、III型；高コレステロール血症および高コレステロール血症、常染色体劣性；過剰驚愕症2および過剰驚愕症遺伝性；高フェリチン血症白内障症候群；高グリシン尿症；周期熱を伴う高免疫グロブリンD；メバロン酸尿症；高免疫グロブリンE症候群；高インスリン血症性低血糖症家族性3、4、および5；高インスリン血症-高アンモニア血症症候群；高リジン血症；ジストニアを伴う高マンガン血症、赤血球増加症および肝硬変；高オルニチン血症-高アンモニア血症-ホモシトルリン尿症症候群；副甲状腺機能亢進症1および2；副甲状腺機能亢進症、新生児重症；高フェニルアラニン血症、bh4欠乏症、a、pts部分欠乏症が原因、BH4欠乏症、D、および非pku；精神発達障害症候群2、3、および4を伴う過リン酸血症症；多毛症骨軟骨ジストロフィー；低ベータリポ蛋白血症、家族性、a p o b 32に関連する；低カルシウム血症、常染色体優性1；低カルシウム尿性高カルシウム血症、家族性、1および3型；低軟骨形成；鉄過負荷を伴う低色素性小球性貧血；肝臓におけるグリコーゲンシンターゼの欠損症を伴う低血糖；無性腺症を伴うまたはなしの性腺機能低下性腺機能低下症11；免疫不全を伴う発汗低下外胚葉異形成；低汗性X連鎖外胚葉異形成；低カリウム性周期性麻痺1および2；低マグネシウム血症1、腸；低マグネシウム血症、けいれん、および精神発達障害；低髄鞘形成性白質ジストロフィー7；低形成性左心症候群；房室中隔欠損と一般的な房室接合部；尿道下裂1および2、X連鎖；甲状腺機能低下症、先天性、非興奮性、1；低トリコピー症8および12；多毛症-リンパ浮腫-毛細血管拡張症症候群；I式血液型；シーメンスの魚鱗癬角膜剥離症；剥脱性魚鱗癬；未熟児魚鱗癬症候群；特発性大脳基底核石灰化5；特発性線維化性肺炎、慢性形；先天性角化異常症、常染色体優性2および5；乳児の特発性高カルシウム血症；カルシウム流入障害2によるT細胞不活性化を伴う免疫機能障害；免疫不全15、16、19、30、31C、38、40、8、高IgM1および2型を伴う、cd3ゼータの欠陥が原因であり、およびマグネシウム欠陥を伴うX連鎖、エプスタイン・バーウイルス感染、および腫瘍症；免疫不全セントロメア不安定性-顔面異常症候群2；封入体ミオパチー2および3；埜中ミオパチー；乳児のけいれんおよび発作性コレオアテトーゼ、家族性；乳児皮質骨増殖症；乳児GM1ガ

10

20

30

40

50

ングリオシド蓄積症；乳児低ホスファターゼ症；乳児ネフロン癆；乳児眼振、X連鎖；乳児パーキンソニズムジストニア；多尾の精子および過度DNAに関連する不妊症；インシュリン抵抗性；インシュリン抵抗性糖尿病および黒色表皮腫；インシュリン依存性糖尿病分泌性下痢；間質性腎炎、巨大核性；子宮内発育遅延、骨幹端異形成、先天性副腎低形成、および性器異常；ヨードチロシルカップリング欠陥；IRAK4欠乏症；虹彩隅角異形成優性型および1型；脳における鉄蓄積；坐骨膝蓋骨異形成；腓島細胞過形成；17、20リアーゼ単独欠損症；ルトロピン単独欠損症；イソバレリルCoA脱水素酵素欠損症；ヤンコビッチリペラ症候群；ジャーベルおよびランゲニールセン症候群2；ジューベール症候群1、6、7、9/15（二性）、14、16、および17、および/または口顔指症候群xiv；ヘルリッツの接合部表皮水疱症；若年性GM1ガングリオシド症；若年性ポリープ症候群；若年性多発/遺伝性出血性毛細血管拡張症候群；若年性網膜分離症；歌舞伎メーキャップ症候群；カルマン症候群1、2、および6；思春期遅発；神崎病；カラク症候群；カルタ遺伝子症候群；ケニーカフェイ症候群2型；ケッペンルピンスキー症候群；円錐角膜1；濾胞性角化症；掌蹠線条体角化症；キンドラー症候群；L-2-ヒドロキシグルタル酸尿；ラーセン症候群、優性型；ラティス角膜ジストロフィーIII型；レーバー黒内障；ツェルヴェーガー症候群；ペルオキシソーム形成異常；ツェルヴェーガー症候群スペクトラム；レーバー先天性黒内障11、12、13、16、4、7、および9；レーバー視神経萎縮症；アミノグリコシド誘発性難聴；難聴、非症候感音性、ミトコンドリアの；左室不全5；左室緻密化障害；リー疾患；ミトコンドリア短鎖エノイルCoAヒドラターゼ欠損症；ミトコンドリア複合体欠乏が原因であるリー症候群；ライナー疾患；レリ・ワイル軟骨軟骨症；致死的先天性拘縮症候群6；白血球粘着不全症IおよびIII型、白質ジストロフィー、低髄鞘形成、11および6；運動失調を伴う白質脳症、脳幹および脊髄病変および乳酸上昇を伴う、白質消失を伴う、および進行性、卵巣不全を伴う；白斑症合計；レビー小体型認知症；リヒテンシュタインクノール症候群；リーフラウメニ症候群1；Lig4症候群；肢帯型筋ジストロフィー、タイプ1B、2A、2B、2D、C1、C5、C9、C14；脳および眼の異常を伴う先天性筋ジストロフィー、ジストログリカノパチー、A14およびB14型；リパーゼ欠損合併症；脂質蛋白症；脂肪異栄養症、家族性部分的、2および3型；滑脳症1、2（X連鎖）、3、6（小脳症あり）、X連鎖；皮質下層状異所形成、X連鎖；急性乳児肝不全；ロイス・ディーツ症候群1、2、3；QT延長症候群1、2、2/9、2/5、（二性）、3、5および5、後天性、への感受性；肺がん；リンパ浮腫、遺伝性、id；リンパ浮腫、原発性、骨髄性異型；リンパ増殖性症候群1、1（X連鎖）、および2；リソソーム酸性リパーゼ欠損症；大頭症、巨人症、顔面異形症候群；黄斑変性症、卵黄様、成人発症；悪性高熱感受性1型；悪性リンパ腫、非ホジキン；悪性黒色腫；前立腺の悪性腫瘍；下顎骨肢端形成異常；AまたはB型脂肪異栄養症を伴う、下顎骨肢端異形成；下顎顔面骨形成不全、トリーチャーコリン型、常染色体劣性；マンノース結合蛋白質欠乏症；メープルシロップ尿症1A型および3型；Marden Walker様症候群；マルファン症候群；マリネスコ・シェーグレン症候群；マルトソルフ症候群；若年者の成人発症糖尿病、1型、2型、11型、3型、および9型；メイ-ヘグリン異常；MYH9に関連する異常；セバスチャン症候群；マキューン・オルブライト症候群；成長ホルモン分泌腺腫；性索間質腫瘍；クッシング症候群；マックシク・カウフマン症候群；マクラウド神経アカントサイトーシス症候群；メッケル・グルーバー症候群；中鎖アシル補酵素A脱水素酵素欠損症；髄芽腫；皮質下嚢胞1および2aを伴う巨大脳白質脳症；先天性末梢血管拡張大理石様皮膚巨脳症；PIK3CA関連過成長スペクトラム；巨大脳症-小脳症-多指症-頭症症候群2；糖尿病および感音性難聴を伴う巨赤芽球性貧血、チアミン反応性；Meier-Gorlin症候群1および4；メリックニードルズ症候群；髄膜腫；精神発達障害、X連鎖、3、21、30、および72；脳橋や小脳低形成を伴う精神発達障害および小脳症；精神発達障害X連鎖症候群5；精神発達障害、上顎前突、および斜視；精神発達障害、常染色体優性12、13、15、24、3、30、4、5、6、および9；精神発達障害、常染色体劣性15、44、46、および5；精神発達障害、常同性運動、てんかん、および/または脳奇形成；精神発達障害、症候群性、クラウス・イエンセン型、X連鎖；精神発達障害、X連鎖、非特異的、症候群性、ヘデラ型、および症候群性、wu型；メロシン欠乏性先天性筋ジストロフィー；若年性異染性白質

ジストロフィー、乳児後期、および成人型；異染性白質ジストロフィー；変態異形成；メトヘモグロビン血症1型および2型；メチオニンアデノシルトランスフェラーゼ欠損症、常染色体優性；ホモシスチン尿症を伴うメチルマロン酸血症、メチルマロン酸尿症c b l B型、；メチルマロニルCoAムターゼ欠損症によるメチルマロン酸尿症メチルマロン酸、mut(0)型；小頭骨異形成性原発性小人症2型；網脈絡膜症を伴うまたはなしの小脳症、リンパ浮腫、または精神発達障害；ヘルニアおよびネフローゼ症候群；小脳症；脳梁の發育不全；痙性対麻痺50、常染色体劣性；全体の發育遅延；CNSミエリン形成不全；脳萎縮症；小脳症、正常知能および免疫不全；小脳症-毛細血管奇形症候群；小球性貧血；小眼球症候群5、7、および9；小眼球症、単独3、5、6、8、およびコロボーマ6；小球状水晶体；片頭痛、家族性脳底；ミラー症候群；外眼筋麻痺を伴うミニコアミオパチー；ミオパチー、コアと先天的；ミッチェルライリー症候群；ミトコンドリア3-ヒドロキシ-3-メチルグルタルルCoAシクターゼ欠損症；ミトコンドリア複合体I、II、III、III(核型2、4、または8)欠乏症；ミトコンドリアDNA枯渇症候群11、12(心筋症型)、2、4B(MNGIE型)、8B(MNGIE型)；ミトコンドリアDNA枯渇症候群3および7、肝脳型、および13(脳筋症型)；ミトコンドリアリン酸運搬体およびピルビン酸運搬体欠乏症；ミトコンドリア三機能的蛋白質欠乏症；長鎖3-ヒドロキシアシルCoA脱水素酵素欠損症；三好筋ジストロフィー1；前脛骨発症を伴うミオパチー、遠位；モールトラネピャエル症候群；モリブデン補因子欠乏症、相補群A；モワット-ウィルソン症候群；ムコリピドーシスIIIガンマ；粘膜多糖症VI型、VI型(重症)、およびVII型；粘膜多糖症、MPS-I-H/S、MPS-II、MPS-III-A、MPS-III-B、MPS-III-C、MPS-IV-A、MPS-IV-B；網膜色素変性症73；ガングリオシド症GM1型1(心臓の関与を伴う)；多中心性骨溶解性腎症；多中心性骨溶解症、結節症および関節症；複数の先天異常心房中隔欠損2；多発性先天異常-筋緊張低下-けいれん症候群3；複数の皮膚および粘膜の静脈形成；多発性内分泌腫瘍、1および4型；多発性骨端異形成5または優性；複数の消化管閉鎖症；多発性翼状片症候群、Escobar型；多種スルファターゼ欠損症；多発性癒合症候群3；筋AMPデアミナーゼ欠損症；筋・眼・脳病；筋ジストロフィー、先天性、巨人型；筋無力症、家族性乳児、1；筋無力症候群、先天性、11、アセチルコリン受容体欠乏症に関連する；筋無力症候群、先天性、17、2A(スローチャンネル)、4B(ファーストチャンネル)、および尿細管凝集塊なし；ミエロペルオキシダーゼ欠損症；MYH関連ポリープ症；子宮内膜がん；心筋梗塞1；ミオクローヌスジストニア；ミヨクローヌ性アトニーてんかん；赤色ぼろ線維・ミオクローヌステんかん症候群；筋原線維性ミオパチー1およびZASP関連；ミオグロビン尿症、急性再発性、常染色体劣性；筋神経性胃腸脳症候群；進行性外眼筋麻痺を伴う小脳性運動失調；ミトコンドリアDNA枯渇症候群4B、MNGIE型；ミオパチー、中心核、1、先天性、過剰な筋紡錘を伴う、遠位、1、乳酸アシドーシス、および側芽細胞性貧血1、先天性白内障を伴う進行性のミトコンドリア、難聴、および発達遅延、ならびに尿細管凝集、2；近視6；筋硬化症、常染色体劣性；先天性筋緊張症(Myotonia congenital)；先天性筋緊張症、常染色体優性および劣性形；爪膝蓋骨症候群；ナンスホラン症候群；ナノフトルモス2；ナバホ神経障害；ネマリン性ミオパチー3および9；新生児低緊張症；知的障害；けいれん；発話と言語の発達遅延；精神発達障害、常染色体優性31；シトリン欠乏により引き起こされる新生児肝内胆汁うっ滞；腎性尿崩症、腎性尿崩症、X連鎖；腎結石症/骨粗鬆症、低リン血症、2；ネフロン癆13、15、および4；不妊症；小脳橋-眼球-腎臓症候群(ネフロン癆、眼球運ドン失調症および小脳異常)；眼球異常を伴うまたはなしのネフローゼ症候群、3型、5型、7型、および9型；ネストールギジェルモ早老症候群；Neu-Laxova症候群1；鉄蓄積4および6を伴う神経変性；神経フェリチン症；神経線維腫症、1型および2型；神経線維肉腫；神経下垂体性尿崩症；神経障害、遺伝性感覚、IC型；中性1アミノ酸輸送欠陥；ミオパチーを伴う中性脂質貯蔵病；好中球免疫不全症候群；Nicolaidis-Baraitser症候群；ニーマンピック病C1、C2型、A型、およびC1型、成人形；非ケトン性高グリシン血症；ヌーナン症候群1および4、LEOPARD症候群1；若年性骨髄単球性白血病を伴うまたはなしのヌーナン症候群様異常；正常カリウム血性周期性麻痺、カリウム感受性；ノルム病；てんかん、難聴、および精神発達障害症候群；精神発達障害、X連鎖102、および症候群性13；肥満；眼白化症、タイプ

10

20

30

40

50

I ; 眼皮白化症1B型、3型、および4型 ; 眼歯指異形成 ; Odontohypophosphatasia ; Odontotrichomelic症候群 ; 小口病 ; 部分性無歯症-大腸がん症候群 ; オーピツG / BBB症候群 ; 視神経萎縮症9 ; 口腔・顔面・指趾症候群 ; オルニチンアミノトランスフェラーゼ欠損症 ; 口腔顔面裂11および7、口唇口蓋裂-外胚葉異形成症候群 ; Orstavik Lindemann Solberg症候群 ; 軽度の軟骨異形成を伴う変形性関節症 ; 離断性骨軟骨炎 ; 正常強膜を伴う骨形成不全症12型、5型、7型、8型、I型、III型、優性形、劣性周産期致死性 ; 頭蓋硬化症を伴うオステオパチア線条体 ; 骨粗鬆症常染色体優性1型および2型、劣性4、劣性1、劣性6 ; 偽神経膠腫を伴う骨粗鬆症 ; 耳口蓋指症候群、タイプIおよびII ; 卵巣形成異常1 ; 卵巣白質ジストロフィー ; 先天性爪肥厚症4および2型 ; 骨のパジェット病、家族性 ; パリスターホール症候群 ; 掌蹠角化症、表皮不溶解性、焦点またはびまん性 ; 焦点またはびまん性非形成および先天性心疾患 ; パピヨン・ルフェーヴル (Papillon-Lefevre) 症候群 ; 傍神経節腫3 ; ヴォン・オイレンブルク (von Eulenburg) の先天性パラミオトニー ; 副甲状腺癌 ; パーキンソン病14、15、19 (若年性発症)、2、20 (早期発症)、6、(常染色体劣性、早期発症、および9 ; 網膜色素上皮のパターン化ジストロフィー ; PC-K6a ; ペリツェウス・メルツパッハ病 ; ペンドレッド症候群 ; 末梢性脱髄性神経障害、中枢性髄鞘形成異常 ; ヒルシュスプルング病 ; 恒久的な新生児糖尿病 ; 神経学的特徴を伴う糖尿病、永続的な新生児 ; 新生児インスリン依存性糖尿病 ; 若年者の成人発症型糖尿病、2型 ; ペルオキシソーム形成異常14B、2A、4A、5B、6A、7A、および7B ; ペロー症候群4 ; ペリー症候群 ; 乳児期の持続性高インスリン血症性低血糖 ; 家族性高インスリン血症 ; 表現型 ; フェニルケトン尿症 ; 褐色細胞腫 ; 遺伝性傍神経節腫-褐色細胞腫症候群 ; 傍神経節腫1 ; 腸のカルチノイド腫瘍 ; コーデン症候群3 ; ホスホグリセリン酸脱水素酵素欠損症 ; ホスホグリセリン酸キナーゼ1欠損症 ; 感光性トリコチオジストロフィー ; フィタン酸貯蔵疾患 ; ピック病 ; ピアソン症候群 ; 色素性網膜ジストロフィー ; 色素性結節性副腎皮質疾患、原発性、1 ; 石灰化上皮腫 ; ピットホプキンス症候群 ; 下垂体依存性副腎皮質機能亢進症 ; 下垂体ホルモン欠乏症、1、2、3、および4合併症 ; プラスミノーゲン活性因子インヒビター1型欠乏症 ; プラスミノーゲン欠乏症、I型 ; 血小板型出血異常15および8 ; 多形皮膚萎縮症、腱拘縮を伴う遺伝性線維症、ミオパチー、および肺線維症 ; 多発性嚢胞腎2、成人型、および小児型 ; 硬化性白質脳症を伴う多嚢胞性脂肪膜性骨異形成症 ; 免疫不全を伴うまたはなしの、ポリグルコサンボディミオパチー1 ; 多小脳回症、非対称、両側前頭頭頂骨 ; 多発神経障害、難聴、運動失調、網膜色素変性症、および白内障 ; 小脳小脳低形成症4型 ; 膝窩部贅皮症候群 ; 孔脳症2 ; 汗孔角化症8 ; 光線性型 ; ボルホビリノーゲンシンターゼ欠損症 ; 晩発性皮膚ポルフィリン症 ; 色素性網膜炎を伴う後柱失調症 ; 後極白内障2型 ; プラダー・ヴィリ様症候群早発卵巣不全4、5、7、および9 ; 原発性常染色体劣性小脳症10、2、3、および5 ; 一次繊毛ジスキネジア24 ; 原発性拡張型心筋症 ; 膝窩翼状片症候群 ; 孔脳症2 ; ポロケラトーシス8、播種性表在性光線型 ; 左室不全6 ; 4、左心室圧迫10 ; 発作性心房細動 ; 原発性高シュウ酸尿症、I型、II型、およびIII型 ; 原発性肥大性骨関節症、常染色体劣性2 ; 原発性低マグネシウム血症、原発性若年性緑内障1 ; 原発性肺高血圧症 ; プリムローズ症候群 ; 進行性家族性心臓ブロック1B型 ; 進行性家族性肝内胆汁うっ滞2および3 ; 進行性肝内胆汁うっ滞 ; 運動失調を伴う進行性ミオクローヌステんかん ; 進行性偽関節リウマチ異形成 ; 進行性硬化性ポリジストロフィー ; プロリダーゼ欠損症 ; プロリン脱水素酵素欠損症 ; 統合失調症4 ; プロパージン欠乏症、X連鎖 ; プロピオン酸血症 ; プロ蛋白質変換酵素1 / 3欠損症 ; 前立腺癌、遺伝性、2 ; プロタン欠陥 ; 蛋白尿症 ; フィンランド先天性ネフローゼ症候群 ; プロテウス症候群 ; 乳腺がん ; 偽軟骨形成性脊椎骨端異形成症候群 (Pseudoachondroplastic spondyloepiphyseal dystrophy syndrome) ; 偽性低アルドステロン症1型常染色体優性および劣性および2型 ; 偽性副甲状腺機能低下症1A型、偽性偽性副甲状腺機能低下症 ; 偽性新生児副腎白質ジストロフィー ; 偽原発性高アルドステロン症 ; 弾性線維性仮性黄色腫 ; 乳児の全身動脈石灰化2 ; 多発性凝固因子欠乏症を伴う弾性線維性仮性黄色腫様の異常 ; 乾癬感受性2 ; PTEN過誤腫瘍症候群 ; 遺伝性出血性毛細血管拡張症に関連した肺高血圧症 ; 肺線維症および / または骨髄不全、テロメア関連、1および3 ; 遺伝性出血性毛細血管拡張症に関する肺高血圧症 ; 肺線

10

20

30

40

50

維症および/または骨髄不全、テロメア関連、1および3；遺伝性出血性毛細血管拡張症を伴う肺高血圧症、原発性、1；プリンスクレオシドホスホリラーゼ欠損症；ピルビン酸カルボキシラーゼ欠損症；ピルビン酸脱水素酵素E1-アルファ欠損症；赤血球のピルビン酸キナーゼ欠損症；レイン症候群ラソパシー劣性異栄養性表皮水疱症；爪異常、先天性非症候群性、8；ライフエンシュタイン症候群；腎異形成；腎カルニチン輸送障害；腎コロボーマ症候群；腎異形成；腎異形成、網膜色素変性症、小脳性運動失調症および骨格異形成；腎尿細管性アシドーシス、遠位、常染色体劣性、遅発性感音難聴、または溶血性貧血を伴う；尿細管性アシドーシス、眼の異常および精神発達障害を伴う；網膜錐体ジストロフィー-3B；網膜色素変性症；網膜色素変性症10、11、12、14、15、17、および19；網膜色素変性症2、20、25、35、36、38、39、4、40、43、45、48、66、7、70、72；網膜芽細胞腫；レット異常；ラブドイド腫瘍プレデポジション症候群2；裂孔原性網膜剥離、常染色体優性；根性点状軟骨異形成症2型および3型；ロバーツSCフォコメリア症候群；Robinow Soerauf症候群；ロビノー症候群、常染色体劣性、常染色体劣性、短合多指症を伴う；ロスメントムソン症候群；ラバジリーノ症候群；RRM2B関連ミトコンドリア病；ルピンスタインタイビ症候群；サラ病；サンドホフ病、成人および小児型；サルコイドーシス、早期発症；ブラウ症候群；シンドラー病、1型；裂脳症；統合失調症15；蝸牛様骨盤異形成；神経鞘腫症2；シュワルツジャンベル症候群1型；角膜硬化症、常染色体劣性；硬化症；第2のアリール甲状腺機能低下症；瀬川病症候群、常染色体劣性；シニアローケン症候群4および5、感覚運動失調性神経障害、構音障害、および眼麻痺；セピアプテリン還元酵素欠損症；セサミ（sesame）症候群；ADA欠損症が原因である重度合併免疫不全、小脳症を伴う、発育遅延、およびイオン化する放射線への感受性、非典型的、常染色体劣性、T細胞陰性、B細胞陽性、NK陽性のNK細胞陰性；部分的アデノシンデアミナーゼ欠損症；重度先天性好中球減少症；重度先天性好中球減少症3、乳児における重症ミオクロニートんかん；熱性けいれんプラスを伴う全身性てんかん、1および2型；重度X連鎖筋細管ミオパチー；QT短縮症候群3；非特異的な骨格異常を伴う低身長；低身長、聴神経管閉鎖症、下顎発育不全、骨格異常；低身長、爪異形成、顔面異形成、および貧毛症；原始性低身長症；多指症を伴うまたはなしの仮肋骨（Short-rib）胸郭異形成11または3；シアリドーシスI型およびII型；銀痙性対麻痺症候群；遅い神経伝導速度、常染色体優性；スミス・レムリ・オピッツ症候群；スナイダーロビンソン症候群；成長ホルモン分泌腺腫；プロラクチノーマ；家族性、下垂体腺腫プレデポジション、ソトス症候群1または2；痙性失調症5、常染色体劣性、シャルルヴォアサグネ型、1、10、または11、常染色体劣性；筋萎縮性側索硬化症5型；痙性対麻痺15、2、3、35、39、4、常染色体優性55、常染色体劣性、および5A；胆汁酸合成欠陥、先天性、3；精子形成不全11、3、および8；球状赤血球症4および5型；スフェロイド体ミオパチー；脊髄性筋萎縮症、下肢主（predominant）2、常染色体優性；脊髄性筋萎縮症、タイプII；脊髄小脳性運動失調症14、21、35、40、および6；脊髄小脳失調症常染色体劣性1および16；脾臓発育不全；脊椎骨底足骨癒合症候群；脊椎指異形成（Spondylocheirodysplasia）、エーラース・ダンロス症候群様、免疫失調症を伴う、アグリカン型、先天性関節脱臼、短手型、Sedaghatian型、錐体杆体ジストロフィー、およびコズロスキー型；擦れ小人症；シュタルガルト疾患1型；錐体杆体ジストロフィー-3；スティックラー症候群1型；クナリスト異形成；スティックラー症候群、1型（非症候眼性）および4；刺傷関連血管症、乳児期発症；ストルモーケン（Stormorken）症候群；スタージ・ウェーバー症候群、毛細血管奇形、先天性、1；サクシニルCoAアセトアセテートトランスフェラーゼ欠損症；スクラーゼ-イソマルターゼ欠損症；乳幼児突然死症候群；亜硫酸オキシダーゼ欠損症、単離された；大動脈弁上部（Supraaortic）狭窄症；界面活性剤代謝機能障害、肺、2および3；交感神経、近位、1b；合指症Cenani Lenz型；合指症3型；症候群性X連鎖精神発達障害16；タリペスタンジール病；TARP症候群；テイ・サックス病、B1変異、GM2ガングリオシドーシス（成人）、GM2ガングリオシドーシス（成人発症）；テタミー症候群；テノリオ症候群；末端骨異形成；テストステロン17-ベータ-脱水素酵素欠損症；テトラアメリア、常染色体劣性；ファロー四徴症；左心低形成症候2；総動脈幹；心臓および大血管の奇形；心室中隔欠損1；ティールベーンケ角膜ジストロフィー；胸

10

20

30

40

50

部大動脈瘤および大動脈解離；マルファン体質（Marfanoid habitus）；スリーM（Three M）症候群2；血小板減少症、血小板機能不全、溶血、および不均衡なグロビン合成；血小板減少症、X連鎖；血栓症、遺伝性、蛋白質C欠乏症が原因である、常染色体優性および劣性；甲状腺非形成；甲状腺がん、濾胞性；甲状腺ホルモン代謝、異常；甲状腺ホルモン耐性、全身性、常染色体優性；甲状腺中毒性周期性四肢麻痺および甲状腺中毒性周期性四肢麻痺2；甲状腺刺激ホルモン放出ホルモン抵抗性；ティモシー症候群；TNF受容体関連周期熱症候群（TRAPS）；歯の非形成、選択的、3および4；多形性心室頻拍；Townes-Brocks鰓弓耳腎様症候群；新生児の一過性水疱性皮膚炎；トリーチャー・コリンズ症候群；精神発達障害、小人症および網膜の色素変性症を伴う長睫毛症；毛髪鼻指節骨異形成I型；毛髪鼻指節骨症候群3型；トリメチルアミン尿症；結節性硬化症症候群；リンパ管筋腫症；結節性硬化症1および2；チロシナーゼ陰性眼皮膚白皮症；チロシナーゼ陽性眼皮膚白皮症；チロシン血症I型；UDPグルコース-4-エピメラーゼ欠損症；ウルリッヒ先天性筋ジストロフィー；尺骨および腓骨の不在、重度四肢欠損を伴う；血栓性血小板減少性紫斑病（Upshaw-Schulman syndrome）；ウロカニン酸ヒドラターゼ欠損症；アッシャー症候群、1型、1B、1D、1G、2A、2C、および2D；網膜色素変性症39；紫外線感受性症候群；ファン・デル・ウンデ症候群；Van Maldergem症候群2；ヘネカムリンパ管拡張症-リンパ浮腫症候群2；異型ポルフィリン症；嚢胞腎を伴う心室肥大；Verheij症候群；極長鎖アシルCoAデヒドロゲナーゼ欠損症；膀胱尿管逆流症8；内臓逆位5、常染色体；内臓ミオパチー；ビタミンD依存性くる病、1および2型；卵黄様ジストロフィー；フォン・ヴィレブランド病2M型および3型；ワールデンブルグ症候群1、4C、および2E型（神経系の関わりを伴う）；Klein-Wardenberg症候群；Walker-Warburg先天性筋ジストロフィー；Warburgミクロ症候群2および4；いぼ、低ガンマグロブリン血症、感染、および骨髄性溶血症；ウィーバー症候群；Weill-Marchesani症候群1および3；Weill-Marchesani様症候群；Weissenbacher-Zweymüller症候群；ウェルドニッヒ・ホフマン病；シャルコー・マリー・トゥース病；Werner症候群；WFS1関連異常；ヴィーデマン・スタイナー症候群；Wilson病；Wolfram様症候群、常染色体優性；Worth病；Van Buchem病2型；色素性乾皮症、相補群b、群D、群E、および群G；X連鎖無ガンマグロブリン血症；X連鎖遺伝性運動感覚ニューロパチー；ステリルスルファターゼ欠損を伴うX連鎖魚鱗癬；X連鎖脳室周囲異所性灰白質；耳口蓋指趾症候群1型；X連鎖重症複合免疫不全；Zimmermann-Laband症候群およびZimmermann-Laband症候群2；ならびに層間粉状白内障3を包含する。

10

20

30

40

50

【0655】

本開示は病原性G AまたはC T変異を含む遺伝子のリストを提供する。かかる病原性G AまたはC T変異は、本明細書において提供される方法および組成物を用いて、例えばAをGにおよび/またはTをCに変異させることによって修正され得、それによって遺伝子機能を回復させる。表2は本明細書に記載される塩基編集因子を用いて修正され得る例示的な変異を包含する。表2は、遺伝子記号、関連する表現型、修正されるべき変異、および変異を修正するために用いられ得る例示的なgRNA配列を包含する。表2に提供されているgRNA配列は、Cas9または本明細書において提供される塩基編集因子のいずれかを標的部位へと導き得るRNAをコードする配列である。例えば、表2に提供されているgRNA配列は、疾患に関する変異を修正するために、Cas9または本明細書において提供される塩基編集因子のいずれかを標的部位へと標的化するgRNAをコードするように、pFYFなどのgRNA発現ベクターにクローニングされ得る。しかしながら、G AまたはC T変異に関連する追加の疾患を処置するために、追加の変異が修正され得るということは了解されるはずである。さらにその上、追加のgRNAが、本開示、および当業者によって了解されるであろう当分野の知識に基づいて設計され得る。

【0656】

医薬組成物

本開示の他の側面は、本明細書に記載されるアデノシンデアミナーゼ、融合蛋白質、または融合蛋白質-gRNA複合体のいずれかを含む医薬組成物に関する。本明細書において用いられる用語「医薬組成物」は医薬使用のために調合される組成物を言う。いくつかの態

様では、医薬組成物は薬学的に許容し得る担体をさらに含む。いくつかの態様では、医薬組成物は追加の薬剤を含む（例えば、特異的送達のため、半減期を増大させる、または他の治療化合物）。

【0657】

ここで用いられる用語「薬学的に許容し得る担体」は、体の1つの部位（例えば送達部位）から別の部位（例えば、体の臓器、組織、または部分）に化合物を運ぶことまたは輸送することに関わる、薬学的に許容し得る材料、組成物、または基剤、例えば液体または固体のフィラー、希釈剤、賦形剤、製造助剤（例えば滑剤、タルクマグネシウム、ステアリン酸カルシウムもしくは亜鉛、またはステアリン酸（steric acid））、または溶媒封入材料を意味する。薬学的に許容し得る担体は、調合物の他の成分と適合性でありかつ対象の組織にとって有害でないという意味で「許容される」（例えば、生理学的に適合性、無菌、生理的なpHなど）。薬学的に許容し得る担体としての用をなし得る材料のいくつかの例は：（1）糖、例えばラクトース、グルコース、およびスクロース；（2）澱粉、例えばコーンスターチおよび馬鈴薯澱粉；（3）セルロースおよびその誘導体、例えばカルボキシメチルセルロースナトリウム、メチルセルロース、エチルセルロース、微結晶セルロース、および酢酸セルロース；（4）粉末化トラガカント；（5）麦芽；（6）ゼラチン；（7）潤滑剤、例えばステアリン酸マグネシウム、ラウリル硫酸ナトリウム、およびタルク；（8）賦形剤、例えばココアバターおよび座薬ワックス；（9）油、例えばピーナツオイル、綿実油、サフラワー油、ごま油、オリーブ油、コーン油、および大豆油；（10）グリコール、例えばプロピレングリコール；（11）ポリオール、例えばグリセリン、ソルビトール、マンニトール、およびポリエチレングリコール（PEG）；（12）エステル、例えばオレイン酸エチルおよびラウリン酸エチル；（13）寒天；（14）緩衝剤、例えば水酸化マグネシウムおよび水酸化アルミニウム；（15）アルギン酸；（16）パイロジェン不含水；（17）等張食塩水；（18）リンゲル液；（19）エチルアルコール；（20）pH緩衝溶液；（21）ポリエステル、ポリカーボネート、および/またはポリ酸無水物；（22）増量剤、例えばポリペプチドおよびアミノ酸；（23）血清構成要素、例えば血清アルブミン、HDL、およびLDL；（22）C2-C12アルコール、例えばエタノール；ならびに（23）医薬調合に採用される他の非毒性の適合性の物質を包含する。湿潤剤、着色料、離型剤、コーティング剤、甘味料、フレーバー剤、香料、保存料、および抗酸化剤もまた調合物中に存在し得る。「賦形剤」、「担体」、「薬学的に許容し得る担体」、または同類などの用語は本明細書において交換可能に用いられる。

【0658】

いくつかの態様では、医薬組成物は、対象への送達のために、例えば遺伝子編集のために調合される。本明細書に記載される医薬組成物を投与する好適な経路は、限定なしに：外用、皮下、経皮、皮内、病巣内、関節内、腹腔内、膀胱内、経粘膜、歯肉、歯内、蝸牛内、経鼓室、臓器内、硬膜外、髄腔内、筋肉内、静注、血管内、骨内、眼周辺、腫瘍内、脳内、および脳室内投与を包含する。

【0659】

いくつかの態様では、本明細書に記載される医薬組成物は有疾患部位（例えば腫瘍部位）に局所投与される。いくつかの態様では、本明細書に記載される医薬組成物は、注射によって、カテーテルの手段によって、座薬の手段によって、またはインプラントの手段によって対象に投与され、インプラントは多孔質、非多孔質、またはゼラチン質材料であり、膜、例えばシラスティック（sialastic）膜、またはファイバーを包含する。

【0660】

他の態様では、本明細書に記載される医薬組成物はコントロールされた放出システムによって送達される。1つの態様では、ポンプが用いられ得る（例えば、Langer, 1990, Science 249: 1527-1533; Sefton, 1989, CRC Crit. Ref. Biomed. Eng. 14: 201; Buchwald et al., 1980, Surgery 88: 507; Saudek et al., 1989, N. Engl. J. Med. 321: 574を参照する）。別の態様では、ポリマー材料が用いられ得る（例えば、Medical Applications of Controlled Release (Langer and Wise eds., CRC Press, Boca Raton, Fla., 1

10

20

30

40

50

974) ; Controlled Drug Bioavailability, Drug Product Design and Performance (Smol en and Ball eds., Wiley, New York, 1984) ; Ranger and Peppas, 1983, Macromol. Sci . Rev. Macromol. Chem. 23: 61を参照する。Levy et al., 1985, Science 228: 190 ; Du ring et al., 1989, Ann. Neurol. 25: 351 ; Howard et al., 1989, J. Neurosurg. 71: 105をもまた参照する)。他のコントロールされた放出システムは例えば上記Langerにおいて論ぜられている。

【 0 6 6 1 】

いくつかの態様では、医薬組成物は、慣例的な手続きに従って、対象、例えばヒトへの静注または皮下投与に適応した組成物として調合される。いくつかの態様では、注射による投与のための医薬組成物は無菌の等張水系緩衝液中の溶液である。必要などころでは、
10
医薬は、可溶化剤、および注射の部位の痛みを和らげるためのリグノカインなどの局所麻酔をもまた包含し得る。一般的に、成分は、例えば、活性な薬剤の数量を指示するアンプルまたはサシットなどの気密容器中の乾燥した凍結乾燥粉末または水不含濃縮物として、別個に、またはユニット剤形中で一緒に混合されてどちらかで供給される。医薬が輸液によって投与されるべきところでは、それは無菌の医薬グレードの水または食塩水を含有する輸液ボトルによって配液され得る。医薬組成物が注射によって投与される所ころでは、成分が投与に先立って混合され得るように、無菌の注射のための水または食塩水のアンプルが提供され得る。

【 0 6 6 2 】

全身投与のための医薬組成物は、液体、例えば無菌食塩水、乳酸加リンゲル液、またはハンクス液であり得る。加えて、医薬組成物は固体形態であり得、使用に直ちに先立って再溶解または懸濁され得る。凍結乾燥形態もまた企図される。
20

【 0 6 6 3 】

医薬組成物は、脂質粒子またはベシクル、例えばリボソームまたはマイクロ結晶内に含有され得、これもまた非経口投与にとって好適である。組成物がその中に含有される限り、粒子はいずれかの好適な、例えば単層ラメラまたは複層ラメラ構造であり得る。化合物は、膜融合性脂質ジオレオイルホスファチジルエタノールアミン (DOPE)、低いレベル (5-10mol%) のカチオン性脂質を含有し、ポリエチレングリコール (PEG) コーティングによって安定化された、「安定化されたプラスミド-脂質粒子」 (SPLP) 中に捕捉され得る (Zhang Y. P. et al, Gene Ther. 1999, 6: 1438-47)。正荷電脂質、例えばN-[1-(2,3-
30
ジオレオイルオキシ)プロピル]-N,N,N-トリメチル-アンモニウム (amonium) メチルサルフェートまたは「DOTAP」が、かかる粒子およびベシクルにとって特に好ましい。かかる脂質粒子の調製は周知である。例えばU.S.特許No.4,880,635 ; 4,906,477 ; 4,911,928 ; 4,917,951 ; 4,920,016 ; および4,921,757を参照する ; これらのそれぞれは参照によって本明細書に組み込まれる。

【 0 6 6 4 】

本明細書に記載される医薬組成物は例えばユニットドーズとして投与またはパッケージングされ得る。本開示の医薬組成物の参照に用いられるときに、用語「ユニットドーズ」は、対象のための単位的な用量として好適な物理的に別々になったユニットを言い、各ユニットは、要求される希釈剤 ; すなわち担体または基剤と結びつけられた、所望の治療効果
40
を産生するように計算された所定の数量の活性材料を含有する。

【 0 6 6 5 】

さらに、医薬組成物は、(a) 本発明の化合物を凍結乾燥形態で含有する容器と (b) 注射のための薬学的に許容し得る希釈剤 (例えば滅菌水) を含有する第2の容器とを含む医薬キットとして提供され得る。薬学的に許容し得る希釈剤は本発明の凍結乾燥化合物の再構成または希釈のために用いられ得る。任意に、かかる容器 (単数または複数) には、医薬または生物学的産物の製造、使用、または販売を規制する政府当局によって定められた形態の注意書きが随伴し得、この注意書きはヒト投与のための製造、使用、または販売の当局による認可を反映する。

【 0 6 6 6 】

10

20

30

40

50

別の側面では、上に記載されている疾患の処置のために有用な材料を含有する製造品が包含される。いくつかの態様では、製造品は容器およびラベルを含む。好適な容器は、例えば、ボトル、バイアル、シリンジ、および試験管を包含する。容器はガラスまたはプラスチックなどの種々の材料から形成され得る。いくつかの態様では、容器は、本明細書に記載される疾患を処置するために有効である組成物を保持し、無菌のアクセスポートを有し得る。例えば、容器は、皮下注射針によって刺通可能なストッパーを有する静注溶液バッグまたはバイアルであり得る。組成物中の活性な薬剤は本発明の化合物である。いくつかの態様では、容器上のまたは随伴するラベルは、組成物が選ばれた疾患を処置するために用いられるということを示す。製造品は、さらに、薬学的に許容し得る緩衝剤、例えばリン酸緩衝食塩水、リンゲル液、またはデキストロス溶液を含む第2の容器を含み得る。それは、商業のおよび使用者の立場から望ましい他の材料をさらに包含し得、他の緩衝剤、希釈剤、フィルター、針、シリンジ、および使用の説明書を有するパッケージインサートを包含する。

【0667】

送達方法

いくつかの側面では、本発明は、本明細書に記載される1つ以上のベクターなどの (such as or) 1つ以上のポリヌクレオチド、その1つ以上の転写物、および/またはそれから転写される1つ以上の (one or) 蛋白質を宿主細胞に送達することを含む方法を提供する。いくつかの側面では、本発明は、かかる方法によって産生される細胞、およびかかる細胞を含むかまたはそれから産生される生物 (例えば動物、植物、または真菌) をさらに提供する。いくつかの態様では、ガイド配列と組み合わせた (かつ任意に複合体化した) 本明細書に記載される塩基編集因子が細胞に送達される。従来のウイルスおよび非ウイルスに基づく遺伝子伝播方法が、核酸を哺乳類細胞または標的組織に導入するために用いられ得る。かかる方法は、塩基編集因子の構成要素をコードする核酸を培養物のまたは宿主生物の細胞に投与するために用いられ得る。非ウイルスベクター送達システムは、DNAプラスミド、RNA (例えば、本明細書に記載されるベクターの転写物)、裸の核酸、およびリポソームなどの送達基剤と複合体化した核酸を包含する。ウイルスベクター送達システムはDNAおよびRNAウイルスを包含し、これらは細胞への送達後にエピソーム性のまたはインテグレーションされたゲノムどちらかを有する。遺伝子治療手続きの概説は、Anderson, Science 256: 808-813 (1992); Nabel & Feigner, TIBTECH 11: 211-217 (1993); Mitani & Caskey, TIBTECH 11: 162-166 (1993); Dillon, TIBTECH 11: 167-175 (1993); Miller, Nature 357: 455-460 (1992); Van Brunt, Biotechnology 6 (10): 1149-1154 (1988); Vigne, Restorative Neurology and Neuroscience 8: 35-36 (1995); Kremer & Perricaudet, British Medical Bulletin 51 (1): 31-44 (1995); Haddada et al., in Current Topics in Microbiology and Immunology Doerfler and Bihm (eds) (1995); および Yu et al., Gene Therapy 1: 13-26 (1994) を参照する。

【0668】

核酸の非ウイルス送達の方法は、リポフェクション、ヌクレオフェクション、マイクロインジェクション、バイオリスティック、ピロソーム、リポソーム、イムノリポソーム、ポリカチオンまたは脂質：核酸コンジュゲート、裸のDNA、人工ビリオン、および薬剤によって強化されるDNAの取り込みを包含する。リポフェクションは例えばU.S. Pat. No. 5,049,386、4,946,787; および4,897,355) に記載されており、リポフェクション試薬は市販されている (例えば、Transfectam™ および Lipofectin™)。ポリヌクレオチドの効率的な受容体認識リポフェクションにとって好適であるカチオン性および中性脂質は、Feigner, WO91/17424; WO91/16024のものを包含する。送達は細胞 (例えば、in vitro またはエキスピボ投与) または標的組織 (例えば in vivo 投与) にであり得る。

【0669】

免疫脂質複合体などの標的化されたリポソームを包含する脂質：核酸複合体の調製は当業者に周知である (例えば、Crystal, Science 270: 404-410 (1995); Blaese et al., Cancer Gene Ther. 2: 291-297 (1995); Behr et al., Bioconjugate Chem. 5: 382-389 (

10

20

30

40

50

1994) ; Remy et al., Bioconjugate Chem. 5: 647-654 (1994) ; Gao et al., Gene Therapy 2: 710-722 (1995) ; Ahmad et al., Cancer Res. 52: 4817-4820 (1992) ; U.S.Pat.No. 4,186,183、4,217,344、4,235,871、4,261,975、4,485,054、4,501,728、4,774,085、4,837,028、および4,946,787を参照する)。

【0670】

核酸の送達のためのRNAまたはDNAウイルスに基づくシステムの使用は、ウイルスを体内の特異的な細胞へと標的化しウイルスペイロードを核へと運搬するために、高度に進化したプロセスを利する。ウイルスベクターは患者に直接投与され得 (*in vivo*)、またはそれらは *in vitro* の細胞を処置するために用いられ得、改変された細胞は任意に患者に投与され得る (エクスピボ)。従来のウイルスに基づくシステムは遺伝子伝播のためのレトロウイルス、レンチウイルス、アデノウイルス、アデノ随伴および単純ヘルペスウイルスベクターを包含し得る。レトロウイルス、レンチウイルス、およびアデノ随伴ウイルス遺伝子伝播方法では、ホストゲノムへのインテグレーションが可能であり、多くの場合には挿入されたトランスジーンの長期発現をもたらす。加えて、高いトランスダクション効率が多く異なる細胞型および標的組織で観察されている。

10

【0671】

ウイルスの指向性は、外来性のエンベロープ蛋白質を組み込むことによって変調させられ得、標的細胞のあり得る標的集団を拡大する。レンチウイルスベクターはレトロウイルスベクターであり、これらは非分裂細胞にトランスダクションまたは感染することができ、典型的には高いウイルス力価を産生する。よって、レトロウイルス遺伝子伝播システムのセレクションは標的組織に依存するであろう。レトロウイルスベクターはシスに作用する長い末端リピートを含み、最高で6-10kbの外来性の配列のパッケージング容量を有する。最小のシスに作用するLTRはベクターの複製およびパッケージングにとって十分であり、これらは、それから、治療遺伝子を標的細胞にインテグレーションして永久的なトランスジーン発現を提供するために用いられる。幅広く用いられるレトロウイルスベクターは、ネズミ白血病ウイルス (MuLV)、テナガザル白血病ウイルス (GaLV)、サル免疫不全ウイルス (SIV)、ヒト免疫不全ウイルス (HIV)、およびそれらの組み合わせに基づくものを包含する (例えば、Buchscher et al., J. Virol. 66: 2731-2739 (1992) ; Johann et al., J. Virol. 66: 1635-1640 (1992) ; Somnerfelt et al., Virol. 176: 58-59 (1990) ; Wilson et al., J. Virol. 63: 2374-2378 (1989) ; Miller et al., J. Virol. 65: 2220-2224 (1991) ; PCT/US 94/05700を参照する)。一過的発現が好ましい用途では、アデノウイルスに基づくシステムが用いられ得る。アデノウイルスに基づくベクターは多くの細胞型で非常に高いトランスダクション効率の能力があり、細胞分裂を要求しない。かかるベクターによって、高い力価および発現レベルが得られた。このベクターは比較的単純なシステムによって大量に産生され得る。アデノ随伴ウイルス (「AAV」) ベクターは、標的核酸によって細胞をトランスダクションするために、例えば核酸およびペプチドの *in vitro* 産生に、ならびに *in vivo* およびエクスピボの遺伝子治療手続きのためにもまた用いられ得る (例えば、West et al., Virology 160: 38-47 (1987) ; U.S.Pat.No.4,797,368 ; WO93/24641 ; Kotin, Human Gene Therapy 5: 793-801 (1994) ; Muzyczka, J. Clin. Invest. 94: 1351 (1994) を参照する。組換え体AAVベクターの構築はいくつもの公開物に記載されており、U.S.Pat.No.5,173,414 ; Tratschin et al., Mol. Cell. Biol. 5: 3251-3260 (1985) ; Tratschin, et al., Mol. Cell. Biol. 4: 2072-2081 (1984) ; Hermonat & Muzyczka, PNAS 81: 6466-6470 (1984) ; および Samulski et al., J. Virol. 63: 03822-3828 (1989) を包含する。

20

30

40

【0672】

典型的には、宿主細胞に感染する能力があるウイルス粒子を形成するために、パッケージング細胞が用いられる。かかる細胞は、アデノウイルスをパッケージングする293細胞およびレトロウイルスをパッケージングする2細胞またはPA317細胞を包含する。遺伝子治療に用いられるウイルスベクターは、通常は、核酸ベクターをウイルス粒子中にパッケージングする細胞株を産生することによって生成される。ベクターは、典型的には、パ

50

パッケージングおよびホストへの爾後のインテグレーションのために要求される最小限のウイルス配列を含有し、他のウイルス配列は、発現されるべきポリヌクレオチド（単数または複数）のための発現カセットによって置き換えられる。足りないウイルス機能は典型的にはパッケージング細胞株によってトランスで供給される。例えば、遺伝子治療に用いられるAAVベクターは、典型的には、パッケージングおよびホストゲノムへのインテグレーションのために要求されるAAVゲノムからのITR配列のみを所有する。ウイルスDNAは細胞株内でパッケージングされ、これは、他のAAV遺伝子、すなわちrepおよびcapをコードするがITR配列を欠くヘルパープラスミドを含有する。細胞株はヘルパーとしてアデノウイルスによってもまた感染され得る。ヘルパーウイルスはAAVベクターの複製とヘルパープラスミドからのAAV遺伝子の発現とを促進する。ITR配列の欠如が原因で、ヘルパープラスミドは有意な量ではパッケージングされない。アデノウイルスのコンタミネーションは、例えば、アデノウイルスがAAVよりも敏感である熱処置によって縮減され得る。細胞への核酸の送達のための追加の方法は当業者に公知である。例えば、参照によって本明細書に組み込まれるUS20030087817を参照する。

【0673】

キット、ベクター、細胞

本開示のいくつかの側面はキットを提供し、デオキシリボ核酸（DNA）分子中のアデノシンを脱アミノ化する能力があるアデノシンデアミナーゼをコードするヌクレオチド配列を含む核酸コンストラクトを含む。いくつかの態様では、ヌクレオチド配列は本明細書において提供されるアデノシンデアミナーゼのいずれかをコードする。いくつかの態様では、ヌクレオチド配列はアデノシンデアミナーゼの発現を駆動する異種プロモーターを含む。

【0674】

本開示のいくつかの側面はキットを提供し、(a)本明細書において提供される、アデノシンデアミナーゼに融合させられたnapDNAbp（例えばCas9ドメイン）、またはnapDNAbp（例えばCas9ドメイン）およびアデノシンデアミナーゼを含む融合蛋白質をコードするヌクレオチド配列と；(b)(a)の配列の発現を駆動する異種プロモーターとを含む核酸コンストラクトを含む。いくつかの態様では、キットは、ガイド核酸バックボーン（例えばガイドRNAバックボーン）をコードする発現コンストラクトをさらに含み、コンストラクトは、ガイド核酸（例えばガイドRNAバックボーン）への、標的配列と同一または相補的な核酸配列のクローニングを許すように位置したクローニング部位を含む。

【0675】

本開示のいくつかの側面は、本明細書において提供されるアデノシンデアミナーゼ、融合蛋白質、または複合体のいずれかを含有細胞を提供する。いくつかの態様では、細胞は、本明細書において提供されるアデノシンデアミナーゼまたは融合蛋白質のいずれかをコードするヌクレオチドを含む。いくつかの態様では、細胞は本明細書において提供されるヌクレオチドまたはベクターのいずれかを提供する。

【0676】

いくつかの態様では、ホスト細胞が、本明細書に記載される1つ以上のベクターによって一過的にまたは非一過的にトランスフェクションされる。いくつかの態様では、細胞は、それが天然に対象内に存在するままでトランスフェクションされる。いくつかの態様では、トランスフェクションされる細胞は対象から取られる。いくつかの態様では、細胞は、対象から取られた細胞、例えば細胞株に由来する。組織培養のための幅広い種々の細胞株が当分野において公知である。細胞株の例は、C8161、CCRF-CEM、MOLT、mIMCD-3、NHDF、HeLa-S3、Huh1、Huh4、Huh7、HUVEC、HASC、HEKn、HEKa、MiaPaCell、Panc1、PC-3、TF1、CTLL-2、C1R、Rat6、CV1、RPTE、A10、T24、J82、A375、ARH-77、Calu1、SW480、SW620、SKOV3、SK-UT、CaCo2、P388D1、SEM-K2、WEHI-231、HB56、TIB55、Jurkat、J45.01、LRMB、Bcl-1、BC-3、IC21、DLD2、Raw264.7、NRK、NRK-52E、MRC5、MEF、Hep G2、HeLa B、HeLa T4、COS、COS-1、COS-6、COS-M6A、BS-C-1サル腎上皮、BALB/3T3マウス胚線維芽細胞、3T3 Swiss、3T3-L1、132-d5ヒト胎児線維芽細胞；10.1マウス線維芽細胞、293-T、

10

20

30

40

50

3T3、721、9L、A2780、A2780ADR、A2780cis、A172、A20、A253、A431、A-549、ALC、B16、B35、BCP-1細胞、BEAS-2B、bEnd.3、BHK-21、BR293、BxPC3、C3H-10T1/2、C6/36、Cal-27、CHO、CHO-7、CHO-1R、CHO-K1、CHO-K2、CHO-T、CHO Dhfr -/-、COR-L23、COR-L23/CP R、COR-L23/5010、COR-L23/R23、COS-7、COV-434、CML T1、CMT、CT26、D17、DH82、DU145、DuCaP、EL4、EM2、EM3、EMT6/AR1、EMT6/AR10.0、FM3、H1299、H69、HB54、HB55、HCA2、HEK-293、HeLa、Hepa1c1c7、HL-60、HMEC、HT-29、Jurkat、JY細胞、K562細胞、Ku812、KCL22、KG1、KY01、LNCap、Ma-Me1 1-48、MC-38、MCF-7、MCF-10A、MDA-MB-231、MDA-MB-468、MDA-MB-435、MDCK 11、MDCK 11、MOR/0.2R、MONO-MAC 6、MTD-1A、MyEnd、NCI-H69/CPR、NCI-H69/LX10、NCI-H69/LX20、NCI-H69/LX4、NIH-3T3、NALM-1、NW-145、OPCN/OPCT細胞株、Peer、PNT-1A/PNT2、RenCa、RIN-5F、RMA/RMAS、Saos-2細胞、Sf-9、SkBr3、T2、T-47D、T84、THP1細胞株、U373、U87、U937、VCaP、Vero細胞、WM39、WT-49、X63、YAC-1、YAR、およびそれらのトランスジェニック変種を包含するが、これらに限定されない。細胞株は当業者に公知の種々のソースから利用可能である（例えばAmerican Type Culture Collection (ATCC) (マナッサス, Va.)を参照する)。いくつかの態様では、本明細書に記載される1つ以上のベクターによってトランスフェクションされた細胞は、1つ以上のベクター由来配列を含む新たな細胞株を樹立するために用いられる。いくつかの態様では、本明細書に記載されるCRISPRシステムの構成要素によって一過的にトランスフェクションされ（例えば、1つ以上のベクターの一過的トランスフェクション、またはRNAのトランスフェクションによる）、CRISPR複合体の活性によって改変された細胞は、改変を含有するがいずれかの他の外因性の配列を欠く細胞を含む新たな細胞株を樹立するために用いられる。いくつかの態様では、本明細書に記載される1つ以上のベクターによって一過的にもしくは非一過的にトランスフェクションされた細胞、またはかかる細胞に由来する細胞株は、1つ以上の試験化合物を評価することに用いられる。

10

20

【0677】

例

例1: PACEを用いる塩基編集因子またはその構成要素の連続的進化

CRISPR-Cas9ゲノム編集は合成生物学および遺伝医学の範囲を多大に拡大しており、その汎用性は大きい長所である。なぜなら、それは配列が欠失させられるかまたは置き換えられることを許すからである。しかしながら、ヒト医療のための目当ての多くの遺伝子標的は1塩基点変異である。これらの小さい変化を作るためには、なお、CRISPR-Cas9アプローチは、ゲノムを不安定化する二本鎖切断を導入しなければならず、それらを修復するための追加される鋳型DNAによる相同的組換えに依拠する。

30

【0678】

塩基編集は、二本鎖切断なしに所定位置のDNA塩基を改変する代替的なアプローチである。Cas9は、編集部位を標的化し、テザリングされたデアミナーゼに一本鎖DNA基質を提供するために用いられ、これはシチジンをウラシルに変換する。それから、下に示されている通り、正常なDNA修復がG-C対をA-Tによって置き換えることに至る。

【0679】

既存の塩基編集方法は、より汎用的なツールであるためには、とりわけ治療学として現実的であるためには改善を必要とする。なぜなら、安全性および有効性要件は高く、*in vivo*送達は難題であるからである。迅速な生体分子進化方法、ファージによって支援される連続的進化(PACE)が、高い活性および配列一般性ならびに精確な塩基標的化を有するより有効な塩基編集因子を開発するために用いられようとしている。

40

【0680】

PACEでは、セレクション下にある遺伝子はM13バクテリオファージゲノム上にコードされている。その活性は、活性なパリアントのみが感染性の子孫ファージを産生するように遺伝子IIIの発現をコントロールすることによってM13増殖に連結されている。ファージを連続的に増殖および変異させるが、変異は宿主またはそのセレクション回路ではなくファージゲノムのみ蓄積する。なぜなら、フレッシュな宿主細胞が連続的増殖ベッセルに流入(および流出)させられ、セレクションバックグラウンドを有効にリセットするか

50

らである。セレクションを通過するファージのみが繁殖し、連続的希釈プロセスを生き残り得る。

【0681】

新たなPACEセレクションの鍵は、遺伝子III発現を目当ての活性に連結することである。塩基編集がT7 RNAポリメラーゼを活性化し、これがgIIIを転写する低ストリンジェンシーセレクションを設計した。1個の編集事象が、編集されたDNAの転写によって直ちに高いアウトプット増幅に至り得る。

セレクション最適化および検証

【0682】

本明細書において概観される概念実証回路はT7 RNAポリメラーゼのガイドRNA依存的活性化を示したが、ターンオンは最適ではなかった。T7RNAPの発現レベルをプロモーター/RBSスキマニングによって最適化した。プラスミドによってコードされる編集因子発現は、最適化されたコンストラクトを用いて高い-倍の活性化を与えたが(3hで>200x)、ファージによってコードされる編集因子はターンオンを与えず、編集因子ファージは回路上において空のコントロールファージよりも強くは増殖しなかった。

10

【0683】

PACEセレクションでは珍しく、選択的な増殖は、より高活性の遺伝子バリエーションを用いるかまたはセレクションストリンジェンシーを低くすることよりもむしろファージを最適化することによって達成された。PACEによって数百の世代に渡って増殖させたものによって、ファージバックボーンを置き換えた。編集因子のDNAサイズの負担もまた分割型インテインを用いることによって縮減された。ここで、別個に発現されるNインテインおよびCインテインはつなぎ合わさって融合蛋白質を形成する。このスキームでは、リンカーおよびトランススプライシングインテインのN末端部分と併せて、N末端デアミナーゼのみがファージ上にコードされている。ファージ感染により、ホストによってコードされる恒常的に発現されるC末端dCas9およびUGIとのスプライシングによって、全長の塩基編集因子が再構成される。これは変異をデアミナーゼおよびリンカーに制限し、これはアクセス可能な標的を限定する。進化の間のファージバックボーンおよびデアミナーゼ活性の改善は、今や、全長の塩基編集因子がファージ上にコードされることを許し得る(下の表1)。

20

【0684】

活性な(APOBEC-Nインテイン)および不活性な(RFP)ファージの混合物を回路上において連続的なフローの下で増殖させることによって、セレクションを検証した。1000x過剰のRFPファージでさえも、APOBECファージはそれらの最初の力価にかかわらず24h以内に優勢になった。

30

【0685】

低ストリンジェンシー回路上におけるPACEによる増殖はロバストである(図7に示されている。同じインプットファージおよび回路について3つの流量スケジュール)。個々のクローンによる回路活性化についてのルシフェラーゼアッセイ(ラグーンあたり24、時点あたり8)は出発遺伝子型からの改善を示すが、これは今までのところは大体はバックボーン最適化が原因である。

【0686】

ファージ上への遺伝子IIIまたは活性化されたT7 RNAポリメラーゼの組換えは、回路上における塩基編集非依存的な増殖を許す。相同性を縮減するために再コードすること、ファージ上のT7活性を取り除くためのネガティブセレクションラウンド、および改善された塩基編集因子クローンは、これらの問題を縮減または抹消した。

40

代替的なセレクション：ウラシル結合「1ハイブリッド」

【0687】

低ストリンジェンシー-T7 RNAP活性化セレクションは、異なる標的および/またはネガティブセレクションに適応させるためには限定されたポテンシャルを有する。なぜなら、編集部位は機能的に敏感なコード配列内にあるからである。PACEによるTALENおよびCas9進化のために用いられる1ハイブリッドDNA結合ドメインセレクションに基づく代替的なス

50

キームを設計した。塩基編集（示されていない）は弱いプロモーターの上流にウラシルをインストールし、これは、転写活性化因子に融合させられたウラシルDNA結合蛋白質UdgXによって結合される。このアーキテクチャは、任意の標的配列仕様と、不要の塩基編集活性についての対称的な二重セクションとを許すであろう。

【0688】

転写活性化はRpoZ、RNAP、およびプロモーターの空間的編成に敏感である。種々の上流位置を標的化した編集によって、一連のUdgX-RpoZ融合体をスクリーニングした。いくつかの部位および活性化因子融合体は非常に控えめだが再現性あるガイドRNA依存的な転写活性化を示す。しかしながら、今までのところは、PACEからの最も活性なファージクローンは最適化された回路上において1.6×活性化を産生し得るのみであり、これは増殖を

10

支持するには十分ではない。

核酸塩基編集因子（塩基編集因子）のファージによって支援される連続的進化

【0689】

本明細書において概観される本発明は、高精度ゲノム改変のための方法としての塩基編集、ファージによって支援される連続的進化、塩基編集のための低ストリンジェンシーセクションの開発、セクションチューニングおよび検証、ならびに指向的進化の第1のステップを記載する。

【0690】

点変異は重要なゲノム編集標的である。1ヌクレオチド多型はヒト遺伝子バリエーションの約90%を占める。1塩基変化は多くのヒト疾患に対処し得る。インデルの代わりに終止コドンを導入することは、より均一なノックアウトおよび縮減された細胞死を与える。図11はClinVARデータベース病原性バリエーションを示す。

20

【0691】

図12に例示されている通り、塩基編集は二本鎖切断なしに点変異を導入する。

【0692】

塩基編集は改善され得る。「ウィンドウ」内の複数のCが編集され得る。配列コンテキストは編集効率を左右する（図13）。PAM特異性およびウィンドウは、精確にまたは少しでも編集され得る部位をコントロールする（図14）。送達が限定的であるときには、発現および活性（特異的な部位に対する）は決定的である（図15）。

【0693】

図16に示されている通り、塩基編集には不偏的な進化方法が有益であり得る。さらに、図17に例示されている通り、指向的進化の別々になったラウンドは時間を取る。本明細書に記載される連続的な指向的進化は全てのステップを *in vivo* で組み合わせる（図18および19を参照する）。

30

【0694】

ファージによって支援される連続的進化は進化する遺伝子に変異を限定する（図20および21）。

塩基編集因子のPACE

【0695】

塩基編集因子は大きく、不良に発現される。第1に、蛋白質発現レベルおよび成熟時間は問題であることが分かる。第2に、ファージゲノムサイズは適応度を損ない、裏切りを奨励する。編集は遅く起こる。哺乳類細胞では、編集は3-5日で最大レベルに起こる。細菌では、編集は40-60世代で最大レベルに起こる。それゆえに、DNA修復または複製が要求される場合には、>1世代が必要とされる。最後に、編集因子の結合は編集の読み出しに干渉し得る（CRISPRi）。Cas9滞在時間は長く、蛋白質コード配列を編集することはCRISPRiに至り得る。

40

【0696】

図22、23、および24に示されている通り、PACEセクションは塩基編集のために適応させられ得る。これは低いストリンジェンシーによる莫大な増幅という利益を提供し、それはチューニング可能である。

50

【0697】

原理的には、いずれかのPACE回路は、1) 正常ではセレクション下にあるファージ構成要素を宿主細胞へと動かすことと、2) 塩基編集によって修正され得るやり方で、その構成要素または回路のいずれかの一部を不活性化することによって、塩基編集のために適応させられ得る。図24では、T7 RNAポリメラーゼが塩基編集によって活性化される。他の同様の例は、1) 塩基編集によって、プロテアーゼを触媒的に活性化することもしくは非切断可能な基質を切断可能なものに変換すること、または2) 塩基編集によって、リコンビナーゼを触媒的に活性化することなどを包含する。

【0698】

図25に示されている通り、C末端デグロンタグはT7 RNAP活性をカットする。いくつかの態様では、デグロンタグは蛋白質を分解へと標的化するアミノ酸配列である。いくつかの態様では、デグロンタグはユビキチン依存的である。いくつかの態様では、デグロンタグはユビキチン非依存的である。さらに、T7 RNAポリメラーゼのC末端は触媒作用に参与する。T7 RNAポリメラーゼC末端のモデルが図26に示されている。図27はT7 RNAP活性へのC末端デグロンタグの効果を塩基編集および塩基編集なし両方について例示している。T7 RNAPはC末端デグロンタグの融合によって不活性化され、これはそれを蛋白質加水分解へと標的化し、そのC末端カルボン酸（これは触媒作用に参与する）をブロックする。鋳型鎖上のTrpコドンの編集はmRNA中のそれを終止に変換し（塩基編集の位置および数に依存してTAA、TGA、またはTAG）、野生型T7 RNAPを回復させる。

【0699】

塩基編集はT7 RNAPを活性化する。T7RNAPの活性化の概略図および塩基編集結果が図28に与えられている。編集因子のプラスミドに基づく発現は、1) ルシフェラーゼレポーターによる回路のガイドRNA依存的活性化、カラム2および3、ならびに2) 野生型T7 RNAPへの塩基編集因子標的化の回路アウトプットへの最小限のインパクトを示す。しかしながら、図29に示されている通り、恒常的なT7RNAP発現はより良く働く。T7 RNAPのための恒常的プロモーターの使用は回路ターンオンを > 10倍改善する。

【0700】

図30に示されている通り、塩基編集因子ファージは非常に大きい。Golden GateクローニングはPCRなしにクローニング的なファージアセンブリを許す。これは図31に例示されている。

【0701】

第1世代SPはC末端デグロン回路上で濃縮しない。第1世代SPの結果は図32に示されている。正（いずれかの感染可能な宿主細胞において増殖するgIIIファージ、またはセレクションを短絡させるT7 RNAPファージ）および負（塩基編集活性を有さない空/終止コドンファージ）のコントロールは予想される通り挙動するが、ファージ上の未進化の全長塩基編集因子は塩基編集回路上で増殖しない。しかしながら、塩基編集因子ファージは空ファージよりも効率的に増殖し、編集依存的な増殖が起こっていることを示唆している。図33に示されている通り、この方法の1つの利益は、異なるセレクションが最適化を達成するために用いられ得るということである。ステップは次の通りである：1) PACE最適化されたファージバックボーンに編集因子をクローニングし、2) T7RNAPの発現を縮減し、3) gIIIプラスミドのコピー数を増大させ、4) 編集因子をインテインによって分割し、dCas9.ugiを宿主プラスミド上に置く。

【0702】

図34に示されている通り、インテインによって分割された塩基編集因子はgRNA依存的な増殖を示す。2つの鍵となる変化が塩基編集因子のガイドRNA依存的な濃縮（一晚で10倍）を許す。1) 他のPACEプロジェクトからの約30個の蓄積した変異を有するファージバックボーンの使用は増殖を改善する。2) トランススプライシング分割型インテインを用いて塩基編集因子を分割することは、ファージカーゴがデアミナーゼおよびNインテインに限定されることを許し、Cインテイン、Cas9、およびUGIは宿主細胞によって発現される。これはファージゲノムサイズを実質的に縮減し、複製およびパッケージングをスピードア

ップする。それは進化をファージによってコードされるデアミナーゼに限定もまたし、これはPACEがデアミナーゼによって駆動される塩基編集因子の変化をより効率的に明らかにすることを許す。図35に示されている通り、インテインによって分割された塩基編集因子はPACEにおいてさらに増殖する。最適化された分割型塩基編集因子は連続フローのPACEにおいて存続し得る（力価は経時的に増大する）。

【0703】

図36に示されている通り、活性な編集因子はラグーン中において優勢になり得る。活性な分割型BEファージは、はるかに過剰の不活性な（RFP）ファージを播種されたラグーン中において、活性なファージの低い力価においてさえも（ $1e7$ pfu/mLのRFPファージ、 $1e4$ pfu/mLの活性なファージ）、優勢になる。

10

【0704】

図37に示されている通り、インテインによって分割された塩基編集因子は変異を蓄積する。

【0705】

図38および39に示されている通り、組換えは、野生型増殖を有する裏切り体を生成する。

【0706】

さらなるPACEは縮減されたgIII組換えを示す。PACEデータが図40に示されている。

【0707】

図41に示されている通り、ファージ塩基編集活性はルシフェラーゼ読み出しによってアッセイされ得る。ルシフェラーゼアッセイ時系列はBEファージ活性が見積もられることを許す。宿主細胞の感染後の2および3hの間に、ルミネッセンス/OD/時間の直線期を観察し、この傾きを回路活性化速度の（or）尺度として取る。図42に示されている通り、ルシフェラーゼ時系列アッセイは改善されたファージ適応度を示す。さらなるPACE実験は、親の遺伝子型（黒い太線）よりも速い回路活性化速度を有するファージ子孫（異なるラグーンおよび時点から個々にアッセイ。着色した線）を産生する。PACE 5最上位15クローンが図43に示されている。ファージ上へのT7RNAPの組換えは裏切り体を作成する。PACE 5最上位8クローンが図44に示されており、裏切り体は強調されている。

20

【0708】

概要として、T7RNAP活性化は活性な塩基編集因子をセレクションし、塩基編集因子はPACEによって増殖し得、最適化されたインプットはセレクションが働くようにし得る。将来の開発では、裏切りは改善よりも有効でないようにされ、全長塩基編集因子PACEが可能化されるであろう。加えて、代替的なセレクションが立ち上げられるであろう。これは図45に例示されている。

30

【0709】

図46に示されている通り、PACEアウトプットは「クリーニング」され得る。

【0710】

ウラシルDNA結合蛋白質は「1.5ハイブリッド」セレクションを可能化し得る。これは図47および48に例示されている。

【0711】

図49、50、および51に例示されている通り、転写活性化因子リクルートは空間的編成に敏感である。

40

【0712】

図52、53、54、55、および56に示されている通り、さらなるスクリーニングは、いくつかの位置が編集に対して応答性であるということを示している。

【0713】

図57に示されている通り、応答は高い活性化因子発現によって最大である。

【0714】

図58に例示されている通り、PACEによって進化したファージは最適化された回路の活性化を示す。

50

- 【0715】
 図59、60、61、62、および63に示されている通り、鋳型鎖編集はより速い発現レベルの応答に至る。
- 【0716】
 図64に示されている通り、鋳型鎖脱アミノ化は限定されたセットのコード変異に至る。
- 【0717】
 図65に示されている通り、より強いT7RNAP発現はターンオンを縮減し、毒性である。
- 【0718】
 図66に例示されている通り、E158K突然変異はバックグラウンドを縮減する。
- 【0719】
 図67に示されている通り、デグロンは活性を縮減することではなく毒性を縮減することに決定的である。 10
- 【0720】
 図68に示されている通り、C末端T7リゾチーム-デグロン融合体は縮減されたバックグラウンドを有さない。
- 【0721】
 図69に例示されている通り、より小さいかまたはより良く発現される編集因子バリエーションは役立ち得る。
- 【0722】
 図70に示されている通り、プラスミドによってコードされる編集因子はファージに感染した細胞内で機能的である。 20
- 【0723】
 図71に示されている通り、ファージによってコードされる編集因子は活性である。
- 【0724】
 ファージゲノムからの発現はプラスミドよりも低い。
- 【0725】
 T7 RNAP発現についてプロモーターおよびRBSの強さをスキャンニングすることは、回路活性化の最適化を許す（活性なvs.不活性な塩基編集因子をプラスミド上で発現するときのルシフェラーゼアウトプットの変化によって評価される）。元々の回路SD8 proAは、最適化された回路（R3 proB）よりもかなり低い-倍の活性化および高いバックグラウンドを有する。これは図73に示されている。 30
- 【0726】
 専らデアミナーゼ-インテイン融合体インサートよりもむしろファージバックボーンの変異は、ファージ活性の実質的な変化に至り得る（図74）。
- 【0727】
 PACEによって進化したファージはルシフェラーゼアッセイにおいて回路をより速く活性化するのみならず、それらはより厳格な回路上において未進化のファージ（右）よりも効率的に増殖し得る（R2およびR4 RBSはより低いT7 RNAP発現レベルに至り、よって、回路活性化のためにはR3よりも多くの編集事象を要求する）。これらのより厳格な回路は、R3回路上で強く濃縮するファージに対する連続的なPACEを許す（示されていない）。図75。 40
- 【0728】
 T7 RNAP活性化回路内に設計されるRNA重合のための鋳型鎖の編集は、PACE適合性の時間スケールでの活性化のために要求される（図76）。mRNAレベルの表現型に至るためにはプラスミド複製およびおそらくDNA修復を要求する非鋳型編集は、進化したファージによってさえも3h以内に活性化を示さないが、同様の鋳型鎖編集は強い活性化に至る。
- 【0729】
 全長BE2 PACEは今や実際的に見える。首尾良いPACEには>10倍の増殖が要求される。全長野生型塩基編集因子は進化したファージバックボーンによってさえも増殖することに失敗するが、分割型BE PACEからのデアミナーゼ変異を追加することは最高で1000×増殖に至る。全長BEのPACEはdCas9を包含する全てのBE構成要素の進化を許すであろう。 50

【0730】

表2. 進化したファージバックボーン（バックボーン#）およびデアミナーゼ変異（インサート）は全長塩基編集因子がBE回路上で増殖することを許し、BE全体が進化させられ得るということを示唆している。

【表2】

バックボーン	インサート	ストリンジェンシー1 5'TCCに対する増殖 (予備的)
137 (PACE7)	空	↓4000x
137	dCas9.ugi	↓50x
13	BE2	↓2000x
29	BE2	↓100x
137	BE2	↑2x
137	BE2 A165S F205S	↑80x
137	BE2 H109N A165S P201S F205S	↑1000x

10

20

【0731】

いくつかのデアミナーゼはPACEでは十分に活性である。BE3およびBE4はラットAPOBEC1に基づく。

【0732】

BE3およびBE4はラットAPOBEC1に基づく。それらは、高い総体的な活性、GC標的を編集する重度に損なわれた活性、およびTC標的の高い編集を有する。代替的なデアミナーゼが塩基編集因子として実証されている。AIDおよびCDA両方はGC標的に対して良く働くが、一般的にAPOBEC1よりも低い活性を有する。APOBEC3Gはこれらの全てよりも良く働かない (Komor, A. C. et al. Improved base excision repair inhibition and bacteriophage Mu Gam protein yields C:G-to-T:A base editors with higher efficiency and product purity. *Sci Adv* 3, eaao4774 (2017)). TARGET-AID 塩基編集実装はCDAを用いる (Nishida, K. et al. Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. *Science* 353, aaf8729-aaf8729 (2016)). 「FERNY」はAPOBECファミリー系統樹に基づくNおよびC末端を短縮された祖先配列再構築である。rAPOBEC1: 229aa; FERNY: 161aa。rAPOBEC1に対する配列類似性は55%である。祖先再構築技術は投稿論文に記載されているが (Koblan et al., *Nature Biotechnol* 投稿)、FERNY配列はそこに記載されていない。

30

【0733】

ルシフェラーゼアッセイは、AGCC標的がATCC標的と比較してrAPOBEC1による活性化速度を劇的にカットするが、pmCDA1は両方を活性化し得るということを示している (図77)。PACE進化は、GCC回路をTCC回路と同じだけ強く活性化し得るrAPOBEC1遺伝子型およびファージと、APOBEC1に匹敵する活性化速度を有するCDAおよびFERNYファージとを産生し得る (図78)。分割型BEコンストラクトを用いる初期のPACEによって進化したデアミナーゼのHEK細胞編集は、野生型BEからの改善を示す (図79-83、表3)。

40

【0734】

表3. 予備的なHEKアッセイにおける進化した塩基編集因子の遺伝子型。

【表 3】

プラスミド	デアミナーゼ	起源	-に対して進化	遺伝子型
pBT169	rAPOBEC1	野生型	5'TCC	
pBT170	rAPOBEC1	PAC7最上位クローン	5'TCC	H109N A165S P201S F205S
pBT171	rAPOBEC1	PAC7最上位クローン	5'TCC	A165S F205S
pBT173	pmCDA1	野生型	5'TCC	
pBT174	pmCDA1	PAC5最上位クローン	5'TCC	A123V, XTEN P6S
pBT175	pmCDA1	PAC10最上位クローン	5'TCC	F23S A123V C158R, XTEN P6S
pBT176	pmCDA1	PAC10最上位クローン	5'TCC	A123V G197A, XTEN P6S
pBT177	pmCDA1	PAC10最上位クローン	5'TCC	F23S A123V I195F, XTEN P6S
pBT178	pmCDA1	PAC10最上位クローン	5'TCC	A123V C158R I195T, XTEN P6S
pBT179	pmCDA1	PAC10最上位クローン	5'TCC	V75I K120R A123V C158R I193T, XTEN P6S

10

20

【0735】

HEK細胞における全長塩基編集因子としてのPACE由来APOBEC、CDA、およびFERNY遺伝子型の大きいセットの試験は、それらが野生型デアミナーゼと比較して改善された編集性質を有するというを示している（表4）。

【0736】

表4. PACE 12は改善されたデアミナーゼ遺伝子型を産生した。レギュラーテキストはPACE 5-PACE10であり、斜体テキストはPACE12である。

30

【表 4 - 1】

プラスミド	デアミナーゼ	-に対して進化	デアミナーゼ遺伝子型
pBT209	rAPOBEC1		wt
pBT210	pmCDA1		wt
pBT211	FERNY		"wt"
pBT212	hsAID		wt
pBT213	pmCDA1	TCC	<i>A123V</i>
pBT214	rAPOBEC1	TCC	H109N A165S P201S F205S

40

【表 4 - 2】

pBT215	rAPOBEC1	TCC	E4K H109N A165S P201S F205S
pBT216	rAPOBEC1	TCC	A165S F205S
pBT217	pmCDA1	TCC	F23S A123V C158R
pBT218	FERNY	TCC	V115M
pBT219	pmCDA1	TCC	A123V V197A
pBT220	pmCDA1	TCC	A123V C158R I195T
pBT221	pmCDA1	TCC	V75I K120R A123V C158R I193T
pBT222	pmCDA1	TCC	F23S A123V I195F
pBT223	rAPOBEC1	GCC	E4K H109N H122L D124N R154H A165S P201S F205S
pBT224	FERNY	TCC	H102P D104N
pBT226	rAPOBEC1	TCC	E4K V10A E95A H109N H122L D124N A165S P201S F205S
pBT227	pmCDA1	TCC	A123V D143N N149D I162V T202 -1 frameshift
pBT229	rAPOBEC1	GCC	E4K H109N H122L D124N A165S P201S F205S
pBT232	rAPOBEC1	TCC	E4K H109N D124N A165S P201S F205S
pBT233	rAPOBEC1	TCC	E4K Y40C H109N H122L D124N A165S P201S F205S
pBT234	rAPOBEC1	GCC	E4K H109N H122L D124N N158S A165S P201S F205S
pBT235	rAPOBEC1	GCC	H109N H122L R126H A165S P201S F205S
pBT236	rAPOBEC1	GCC	E4K H109N H122L R126H A165S P201S F205S
pBT237	rAPOBEC1	TCC	E4K E95A H109N H122L D124N A165S P201S F205S

10

【0737】

図84は、それらがアイソジェニックファージバックボーンにクローニングされている場合に、デアミナーゼは細菌においてBE回路活性化速度によってアッセイされ得るということを示している。デアミナーゼ遺伝子型のみが異なり、そのため、それらが活性化速度を決定する。図84では、進化したデアミナーゼを標準化されたファージバックボーンにサブクローニングすることによって、データを収集した。そのため、それらが主としてデアミナーゼの性質を反映する。

20

【0738】

PACEによって進化したデアミナーゼはHEK細胞編集の改善を示す(図85-87)。標的部位の各Cについて、編集%は野生型のおよび進化したデアミナーゼについて同じオーダーで示されている。

30

【0739】

4つの遺伝子型が傑出している(表5、図88-92)。2つのCDAバリエーションは本質的には同等な活性を有し、野生型CDAよりも高い編集とやや広がったウィンドウとを有する。進化したAPOBEC遺伝子型はGC標的に対して活性を有するが、野生型APOBECはGCに対して非常に低い活性を有する。進化したFERNY遺伝子型もまた高いGC活性を有し、より短い蛋白質であるにもかかわらずAPOBECに匹敵して活性である。

30

【0740】

表5. 4つの遺伝子型が傑出している。

【表 5】

プラスミド	デアミナーゼ	-に対して進化	PACE	デアミナーゼ遺伝子型	Abbr.	実際のトランスフェクション
pBT221	pmCDA1	TCC	10	V75I K120R A123V C158R I193T	IRVRT	38.5 ng
pBT222	pmCDA1	TCC	12	F23S A123V I195F	SVF	33.2 ng
pBT223	rAPOBEC1	GCC	12	E4K H109N H122L D124N R154H A165S P201S F205S	KNLNHSSS	30.0 ng
pBT224	FERNY	TCC	12	H102P D104N	PN	n.d.
pBT209	rAPOBEC1			wt		27.6 ng
pBT210	CDA			wt		28.6 ng

40

【0741】

PACEセクションは他の進化標的に適応させられ得る。これらの拡張はまだ実験的に検証されていない。CCA編集部位の5'の塩基は自由に変えられ得、特異的な5'配列コンテキストにおける編集についてのセクションを許す。ウィンドウ内のほとんどの位置におけ

50

る編集がセレクションされ得る（例えば+1における編集）。編集によってT7 RNAPの終止コドンにQまたはRに復帰させることによって、ABE（A G塩基編集因子）活性がセレクションされ得る。デアミナーゼ活性、ウィンドウ、またはコンテキスト特異性を改善/変調させるために、同様のポジティブセレクションがABEに対して行われ得る。全長塩基編集因子PACEは、PAM特異性、編集ウィンドウ、標的部局滞在などを左右するCas9の変異を許すであろう。直交的なプロモーター特異性（T3）および再コードされたC末端を有するT7 RNAPの第2のコピーを供給することによって、ネガティブセレクションが実装され得る。望まれない編集はT3バリエーションを活性化し（ポジティブセレクションと同じ様式でC末端デグロンを取り除くことによる）、pIII-negの産生を駆動し、これはファージ増殖を縮減する。ネガティブセレクションは、特異的な5'塩基または特異的なウィンドウ位置を有する標的に対して行われ得る。ポジティブおよびネガティブセレクションは同じ宿主細胞内で同時に起こり得る。このタンデムな二重セレクションの1つの例のユースケースは、位置5における編集に対して、位置1における編集をセレクションすることであり、編集ウィンドウがPAMからシフトすることを促すであろう。所与のデアミナーゼを用いるBEについてのウィンドウシフトは確実に達成されたことがない。

10

【0742】

次のデータ（表6、図93-114）は、コドン最適化されたBE4Maxのコンテキストにおいて、PACEに直接由来するデアミナーゼ遺伝子型をコドン最適化なしに用いている。それらは、高いトランスフェクション効率（750ng編集因子プラスミド）においては、いくつかの進化した遺伝子型が野生型デアミナーゼよりも優れたまたは等しい性能を有するという

20

【0743】

表6. PACEに直接由来するデアミナーゼ遺伝子型。

anc689のコンテキストに移した (ancBE4Max. Koblan, L. W. et al. Improving cytidine and adenine base editors by expression optimization and ancestral reconstruction. Nature Publishing Group 1-9 (2018). doi: 10.1038/nbt.4172)。

【 0 7 4 5 】

図115-136は、進化した変異を例外として組成がBE4Max (技術水準rAPOBEC1に基づく塩基編集因子) およびancBE4Max (いくつかの部位についてBE4Maxを上回る再構築された祖先anc689) と同等の、Genscriptコドン最適化されたデアミナーゼからの編集を示している。ドーズは高く (750ng)、よって、ウィンドウの中心における編集は以前のデータのように飽和している。

【 0 7 4 6 】

APOBEC / FERNYの特異性ループの2つの変異 (変異体表の点線) はGC活性にとって決定的である。これらの変異はそれらが進化したデアミナーゼコンテキストに特異的である; どちらのセットもanc689バックグラウンドにはポートしない。このループ領域は-1および-2塩基選好性に影響することが種々のシチジンデアミナーゼについて公知である (例えば、Kohli, R. M. et al. A portable hot spot recognition loop transfers sequence preferences from APOBEC family members to activation-induced cytidine deaminase. Journal of Biological Chemistry 284, 22898-22904 (2009))。APOBECまたはanc689をFERNYのサイズまで短縮することはそれらを完全に不活性化する。evoCDA変異の全ての3つは機能的に見えるが、F23S変異が最も重要に見える。evoCDAは約1-13という最大半量の活性ウィンドウを有する (vs.CDA 約1-9)。evoAPOBECは、どちらかの側に約0.5塩基拡大された、APOBECに類似のウィンドウを有する。evoAPOBECおよびevoFERNYは試験された全ての部位においてanc689 (現行の技術水準BE) よりも優れている。evoAPOBECは数個の標的においてのみevoFERNYよりも高い編集を有する。evoFERNYはAPOBEC/anc689の約227aaと比較して約161aaの長さであり、それをDNAサイズが制約される送達方法のためのより良い選択肢にする。

【 0 7 4 7 】

Genscriptコドン最適化は活性を改善するが、わずかにのみであり (PACE遺伝子型およびGenscriptコドン最適化されたトランスフェクションデータを比較せよ)、そのため、蛋白質配列が性能の一次的な決定因子である。

【 0 7 4 8 】

次は、デアミナーゼの編集ウィンドウをシフトするかまたは幅広くするためにBE PACEを用いることについての概念実証データである (図137)。CDAはAPOBEC1よりも幅広いウィンドウを有する (かつ各evoバージョンは野生型よりも幅広いウィンドウを有する)。ガイドRNAがGTCC編集標的を位置4および5または位置1および2に置くPACE回路を構築した。回路同士はさもなければ同一である。より幅広いウィンドウを有するデアミナーゼは、標的が1/2であるときには、4/5よりもかなり強く野生型APOBEC1を上回る。よって、1/2回路上での進化は、より幅広いウィンドウまたはPAMからシフトしたウィンドウを有するデアミナーゼを濃縮すると予想される。

簡潔な方法:

【 0 7 4 9 】

プラスミドは標準的な分子生物学技術を用いてUSERおよびGolden Gateクローニングによって構築した。PCRによって増幅された全てのDNAは使用に先立って配列確認した。

ルシフェラーゼレポーターアッセイ:

【 0 7 5 0 】

典型的なアッセイでは、適切なプラスミドを含有するS1030またはS2060細胞 [Carlson, J. C., Badran, A. H., Guggiana-Nilo, D. A. & Liu, D. R. Negative selection and stringency modulation in phage-assisted continuous evolution. Nature Chemical Biology 10, 216-222 (2014)] を必要に応じてカルベニシリン (50 µg/mL)、カナマイシン (30 µg/mL)、クロラムフェニコール (40 µg/mL)、および/またはスペクチノマイシン (100 µg/mL) を含有する1mLのDRM培地に接種し、プリーザブルトップフィルムシールを

10

20

30

40

50

装着したディープ96ウェルプレート容量2mL (Eppendorf) 上で37°Cで一晩増殖させた。独立した生物学的レプリケートは、1つ以上のプラスミドのフレッシュな形質転換からの別個のコロニー、または-80グリセロールストックの独立した一晩アウトグロースどちらかであった。一晩培養物を、抗生物質を有するフレッシュなDRMに50倍再希釈し、それから約1.5h (ファージによって発現される塩基編集因子のため) または2h (プラスミドによって発現される塩基編集因子のため) に渡ってディープウェルプレート上で増殖させた。ファージアッセイでは、135 μ Lの宿主細胞培養物を15 μ Lの高力価ファージストック ($> 1e10$ pfu/mL) とクリアボトムブラック96ウェルアッセイプレート (Costar) 上で混合した。プラスミドによって発現される塩基編集因子アッセイでは、培養物をアラビノース (10mMまたは示されている通り) によって誘導し、さらなる3hに渡って増殖させ、それからアッセイプレートに移した (ウェルあたり150-200 μ L)。OD600およびルミネッセンスはInfinite M1000 Proマイクロプレートリーダー (Tecan) を用いてモニタリングし、温度は37°Cにセットした。速度論的アッセイでは、読み取りはモニタリング期間中に3.5分毎になされ、プレートは読み取り間に30sに渡って振った。

増殖アッセイ (スライド11、13、19) :

【0751】

ルシフェラーゼアッセイについて記載されている通り調製したDRM中の対数期宿主細胞を、力価測定されたファージストックと約 $1e6$ pfu/mLの終濃度および1mLの体積に混合し、プリーザブルトップフィルムシールを有するディープウェルプレートによって一晩増殖させた。培養物を遠心して (3,600 \times g、10分) 細胞を取り除き、上清を以前に記載されている通り力価測定した [Badran, A. H. et al. Continuous evolution of *Bacillus thuringiensis* toxins overcomes insect resistance. *Nature* 533, 1-19 (2016)]。宿主細胞によって増殖したファージの力価を同じインプット濃度で細胞なしで一晩振ったファージの力価によって除算することによって、-倍の濃縮を計算した。

【0752】

PACE実験は本質的に記載されている通りに実行した [Badran, A. H. et al. Continuous evolution of *Bacillus thuringiensis* toxins overcomes insect resistance. *Nature* 533, 1-19 (2016)]。

【0753】

哺乳類細胞実験 : 塩基編集因子プラスミドをGolden Gateクローニングによって調製し、ZymoPure Midiキット (Zymo Research) を用いてミディプレップした。HEK細胞トランスフェクション、HTSライブラリー調製、およびシーケンシングは本質的に記載されている通り実行した (Komor, A. C. et al. Improved base excision repair inhibition and bacteriophage Mu Gam protein yields C:G-to-T:A base editors with higher efficiency and product purity. *Sci Adv* 3, eaao4774 (2017))。

Baringoマウスにおける編集の概要 : 聴覚難聴モデル

【0754】

APOBEC1配列特異性のイシューを克服し、Baringo部位を効率的に編集するために、活性化誘導 (シチジン) デアミナーゼ (AID) またはシチジンデアミナーゼ (CDA) を有するBE4バリエーションはBaringo変異を標的化することにおいて良く働くということが示された。図138では、進化デアミナーゼを組み合わせ、コンストラクトをBaringo胚細胞にヌクレオフェクションした。Evoデアミナーゼは以前には編集可能でない部位における塩基編集を許す。BE-CDAによる編集は進化後に (8%から33%に) 増大した (図139)。

【0755】

概要 : 塩基編集因子の連続的な指向的進化

ゲノム編集は生命科学に革命を起こしており、遺伝子疾患を治癒させるポテンシャルをオフにする。DNAの二本鎖切断またはドナー鋳型を要求することなしに、プログラム可能な様式で1つの標的塩基から別のものへの直接的な不可逆的変換を可能化するゲノム編集のための新たな戦略、塩基編集が最近開発された¹⁻⁴。疾患に関係する変異の大多数は1塩基変化であるので、塩基編集因子は遺伝学的構成要素を有する多くの疾患の研究および

10

20

30

40

50

処置に適用可能である。

【0756】

C T塩基編集因子は、シチジンデアミナーゼを用いて、Cas9によって開かれた一本鎖DNAループ中のシチジンをウリジンに変換する。反対鎖はCas9によってニックングされて、編集された鎖を鋳型として用いるDNA修復メカニズムを刺激する。一方で、融合したウラシルグリコシラーゼ阻害因子（示されていない）は編集された塩基の切り出しを遅くする。最終的に、DNA修復がC・G T・C塩基対変換に至る。

【0757】

塩基編集因子は多くの標的を高い効率で編集し得、多くの場合には、1回の処置後に濃縮なしに細胞の30-70%の編集を達成する。しかしながら、現行のシチジン塩基編集因子活性は標的ヌクレオチドの周囲の塩基に依存する。APOBEC1シチジンデアミナーゼに基づくC T編集因子はTCモチーフを編集することへの選好性を有し、ほとんどのGC塩基を敬遠する。この選好性は、古典的な編集ウィンドウの外部の位置におけるTCの編集と³、それらが最適に位置しているときでさえもGC標的の不良な編集とに至り得る。CDAまたはAIDなどの他のデアミナーゼを用いるC T編集因子はある種のGC部位については有効な代替を提供し得るが、APOBEC1 BEと比較して、より低い総体的な編集効率を有する^{3,5}。ゴールは、指向的進化を用いて、高度に活性な、配列コンテキストアグノスティックな塩基編集因子を生成することである。

【0758】

C T塩基編集因子は、3つの蛋白質構成要素をタンデムで用いて非固有の機能を実行する >1800アミノ酸を有する操作された酵素融合体であり、それについての非常に限定された構造および機械的情報がある。これらの特徴の全ては、不偏的な指向的進化を、塩基編集因子機能を改善するための魅力的なプラットフォームにする。ファージによって支援される連続的進化PACEが本明細書において用いられ、これは各実験において数百ラウンドの変異およびセレクションを行い得る。狙いは、塩基編集についてセレクションするためにこの強力なシステムを適応させることであった。

【0759】

塩基編集をファージ複製にカップリングすることは、1塩基変換によって強く活性化され得る回路を要求する。塩基編集がT7 RNAポリメラーゼの発現に至り、これが、それから (than) 遺伝子III (またはルシフェラーゼレポーター) をT7プロモーターから転写するP ACEセレクションを設計および検証した。転写鋳型鎖の塩基編集はmRNA中のTrpコドンを終止コドンに変換し、翻訳される酵素から蛋白質加水分解的な分解タグを取り除く。このアーキテクチャは幅広いダイナミックレンジを有し、T7 RNAポリメラーゼ転写を変調することによってチューニング可能であり、それは、E. coliおよび哺乳類細胞の間で異なる下流のDNA修復ステップから編集効率をデカップリングする。

【0760】

回路を用いてPACEを始めるために、塩基編集因子のデアミナーゼ部分に制限されるように進化を選んだ。これは、Cas9を包含することと比較して、ファージゲノムのサイズを縮減し、増殖をスピードアップし、より密に機能的な変異標的を作出する。これは、トランスプライシング分割型インテインに融合させられたファージ上のデアミナーゼをコードすることによって奏した。それから、編集因子の残り（ヌクレアーゼデッドCas9およびウラシルグリコシラーゼ阻害因子）もまた分割型インテイン融合体として宿主細胞によって発現され、ファージ感染後に、全長編集因子が蛋白質スプライシングによって再構成される。

【0761】

本セレクションは、塩基編集活性を改善するためにまたは標的CCAの5'の塩基のアイデンティティー (identify) を変えることによって5'配列コンテキスト適合性についてセレクションするために用いられ得る。このPACE塩基編集セレクションを適用して、BE3およびBE4などのAPOBEC1に基づく編集因子によっては不良に編集されるGCC標的のC T編集を改善した^{1,3}。APOBEC1ファージを、活性を最適化するために第1にTCC標的に対して、それ

10

20

30

40

50

からGCC標的に対してPACEに付した。PACEから出現した1つのAPOBEC1バリエーションは細菌ルシフェラーゼレポーターアッセイにおいてGCC標的に対して野生型APOBEC1からの180倍の活性増大を呈した。同じ回路を用いて、CDAのより高活性のバリエーションを進化させた。これは高い固有のGC活性を有するが、より低い総体的な効率を有する³。CDAの見かけ上の活性は細菌ルシフェラーゼアッセイでは3から4倍改善された。多様なゲノム標的を編集するための哺乳類細胞におけるこれらの進化したデアミナーゼの試験が進行中であり、論文が準備中である。

【0762】

塩基編集のためのPACEセレクションの利用可能性は、塩基編集因子機能を改善するための多くの可能性を開く。次のゴールは編集ウィンドウをシフトすることおよび狭めることである。これは近傍の「パイスタンダー」Cの編集なしに標的Cの正確な改変を許し、1個のアレルへのクリーンな変換に至るであろう。これは、所望のウィンドウ内のポジティブセレクションおよびその外部のネガティブセレクションによるタンデムな二重セレクションを実装することを要求するであろう（2018年のハーバード授業、Christine Zhengによる）。類似のアプローチが、所与の配列コンテキスト内のCのみを改変するコンテキスト特異的編集因子をセレクションし、再びパイスタンダーC改変を縮減するために用いられ得る。最後に、BE PACEセレクションは、A G塩基編集因子に¹¹、および操作されたCas9またはそのホモログを用いる新たな塩基編集因子バリエーションに適用されて、それらの性能を改善し得る。

参照

【0763】

1 Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A. & Liu, D. R. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533, 420-424, doi: 10.1038/nature17946 (2016).

【0764】

2 Kim, Y. B. et al. Increasing the genome-targeting scope and precision of base editing with engineered Cas9-cytidine deaminase fusions. *Nat. Biotech.* 35, 371-376, doi: 10.1038/nbt.3803 (2017).

【0765】

3 Komor, A. C. et al. Improved base excision repair inhibition and bacteriophage Mu Gam protein yields C:G-to-T:A base editors with higher efficiency and product purity. *Sci Adv* 3, eaao4774, doi: 10.1126/sciadv.aao4774 (2017).

【0766】

4 Rees, H. A. et al. Improving the DNA specificity and applicability of base editing through protein engineering and protein delivery. *Nat Commun* 8, 15790, doi: 10.1038/ncomms15790 (2017).

【0767】

5 Nishida, K. et al. Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. *Science* 353, doi: 10.1126/science.aaf8729 (2016).

【0768】

6 Esvelt, K. M., Carlson, J. C. & Liu, D. R. A system for the continuous directed evolution of biomolecules. *Nature* 412, 499-503, doi: 10.1038/nature09929 (2011).

【0769】

7 Carlson, J. C., Badran, A. H., Guggiana-Nilo, D. A. & Liu, D. R. Negative selection and stringency modulation in phage-assisted continuous evolution. *Nat Chem Biol* 10, 216-222, doi: 10.1038/nchembio.1453 (2014).

【0770】

8 Leconte, A. M. et al. A population-based experimental model for protein evolution

10

20

30

40

50

ion: effects of mutation rate and selection stringency on evolutionary outcomes. *Biochemistry* 52, 1490-1499, doi: 10.1021/bi3016185 (2013).

【0771】

9 Hubbard, B. P. et al. Continuous directed evolution of DNA-binding proteins to improve TALEN specificity. *Nat Methods* 12, 939-942, doi: 10.1038/nmeth.3515 (2015).

【0772】

10 Bryson, D. I. et al. Continuous directed evolution of aminoacyl-tRNA synthetases. *Nat Chem Biol*, doi: 10.1038/nchembio.2474 (2017).

【0773】

11 Gaudelli, N. M. et al. Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage. *Nature* 551, 464-471, doi: 10.1038/nature24644 (2017).

同等物および範囲

【0774】

請求項においては、反対に指示されないかまたはコンテキストから別様に明白でない限り、「a」、「an」、および「the」などの冠詞は1つまたは1つよりも多くを意味し得る。反対に指示されないかまたはコンテキストから別様に明白でない限り、「または」を群の1つ以上の構成員間に包含する請求項または記載は、群の構成員の1つ、1つよりも多く、または全てが所与の産物またはプロセスに存在するか、使用されるか、または別様に関係する場合には満足していると考えられる。本発明は、群の厳密に1つの構成員が所与の産物またはプロセスに存在するか、使用されるか、または別様に関係する態様を包含する。本発明は、群の構成員の1つよりも多くまたは全てが所与の産物またはプロセスに存在するか、使用されるか、または別様に関係する態様を包含する。

【0775】

さらにその上、本発明は、列記されている請求項の1つ以上からの1つ以上の限定、要素、節、または記述用語が別の請求項に導入される全ての変形、組み合わせ、および順列を包摂する。例えば、別の請求項に従属するいずれかの請求項は、同じ元の請求項に従属するいずれかの他の請求項に見出される1つ以上の限定を包含するように改変され得る。要素がリストとして、例えばマーカッシュ群フォーマットで提出されているところでは、要素の各下位群もまた開示されており、いずれかの要素（単数または複数）が群から取り除かれ得る。一般的に、本発明または本発明の側面が特定の要素および/または特徴を含むと言われるところでは、本発明のある態様または本発明の側面がかかる要素および/または特徴からなるかまたは本質的になるということは理解されるべきである。単純の目的のために、それらの態様は本明細書においていちいち具体的には提示しなかった。用語「含む」および「含有する」は開放的であることを意図されており、追加の要素またはステップの包含を許すということが注記される。範囲が与えられているところでは、エンドポイントが包含される。さらにその上、別様に指示されないかまたはコンテキストおよび当業者の理解から別様に明白でない限り、範囲として表現されている値は、コンテキストが明瞭に別様に述べていない限り、本発明の異なる態様において、書かれている範囲内のいずれかの具体的な値または部分範囲を、範囲の下限の単位の十分の一までとり得る。

【0776】

本願は種々の発行済み特許、公開特許出願、雑誌記事、および他の公開物を参照し、これらの全ては参照によって本明細書に組み込まれる。組み込まれた参照のいずれかおよび本明細書の間には矛盾がある場合には、本明細書が優先するものとする。加えて、従来技術のうちに入る本発明のいずれかの特定の態様は、請求項のいずれか1つ以上から明示的に除外され得る。かかる態様は当業者に公知であると見なされるので、それらは除外が本明細書において明示的に提示されていない場合でさえも除外され得る。本発明のいずれかの特定の態様は、従来技術の存在に関するか否かにかかわらず、いずれかの請求項からいずれかの理由で除外され得る。

【0777】

10

20

30

40

50

当業者は、本明細書に記載される具体的な態様の多くの同等物を認識するか、またはせいぜい慣例的な実験作業を用いて確かめることができるであろう。本明細書に記載される本態様の範囲は上の記載に限定されることを意図されず、むしろ、添付の請求項に提示されている通りである。当業者は、次の請求項において定義される本発明の趣旨または範囲から逸脱することなしに、本明細書の種々の変化および改変がなされ得るということを了解するであろう。

【 図 1 】

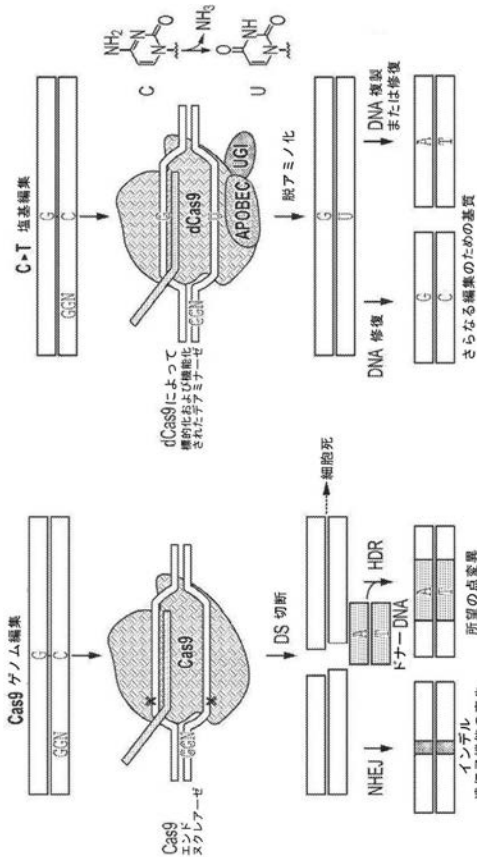


図 1

【 図 2 】

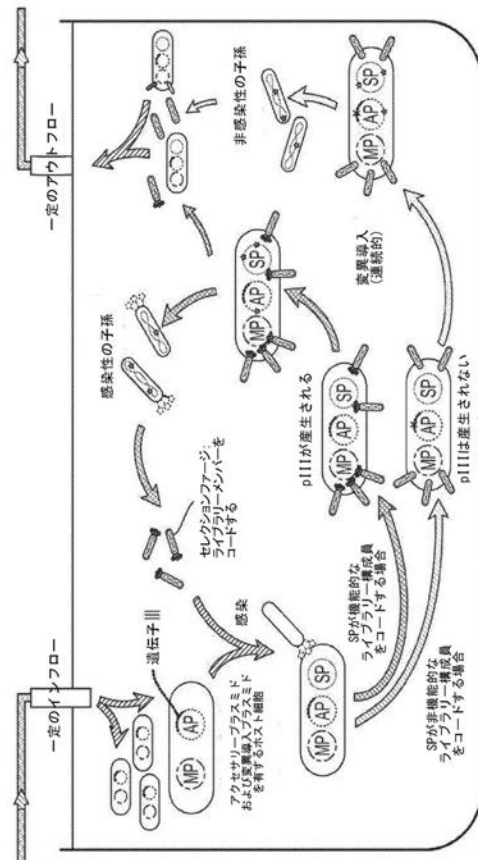


図 2

【 図 3 】

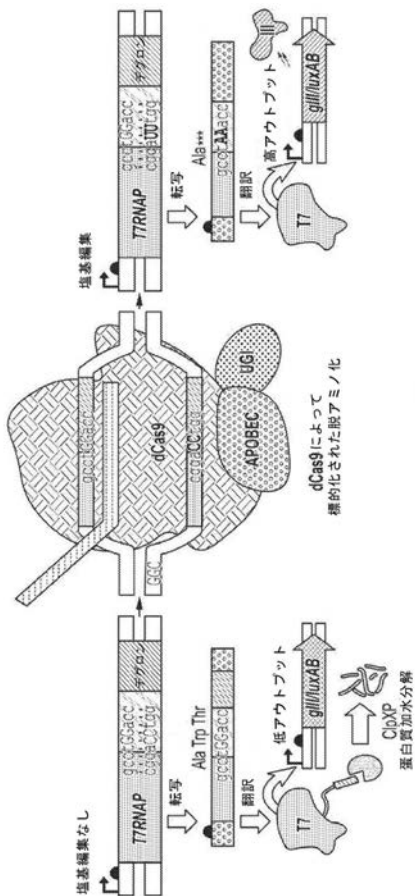


図 3

【 図 5 】

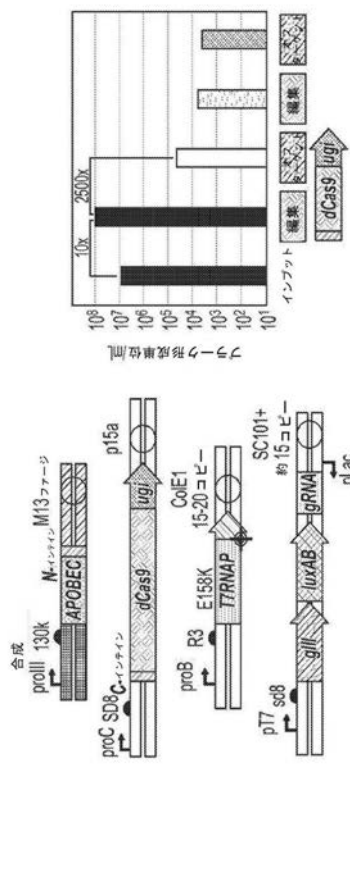


図 5

【 図 4 】

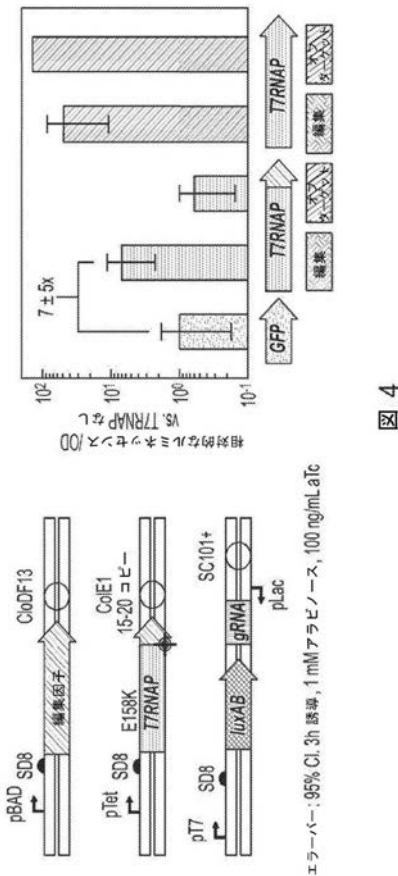


図 4

【 図 6 】

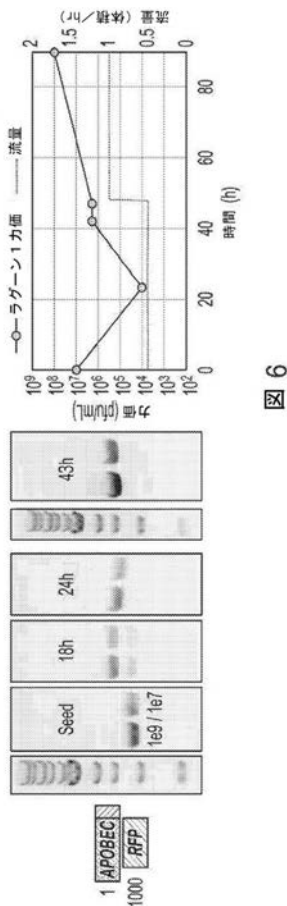


図 6

【 図 1 0 - 1 】

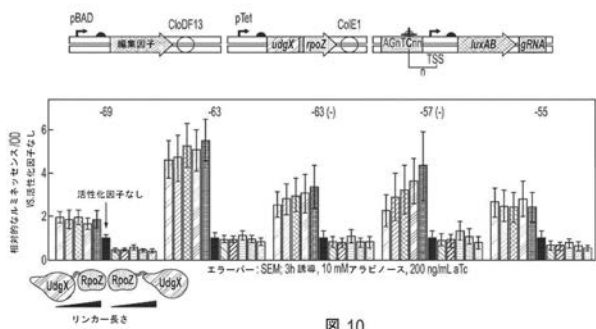


図 10

【 図 1 0 - 2 】

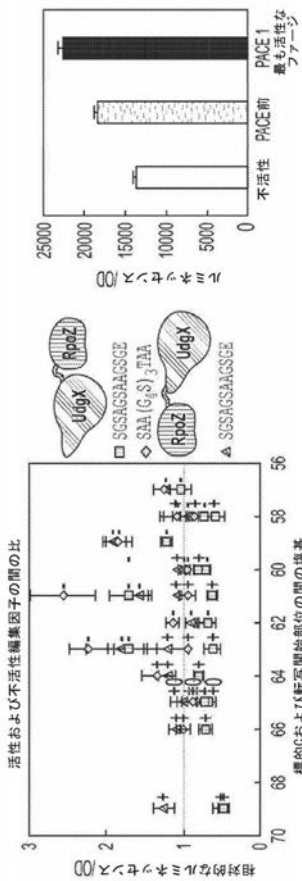


図 10 続き

【 図 1 1 】

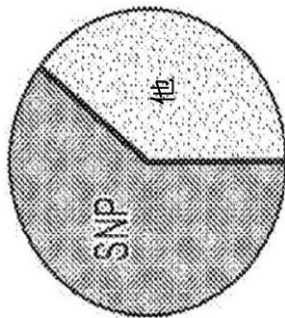


図 11

【 図 1 2 】

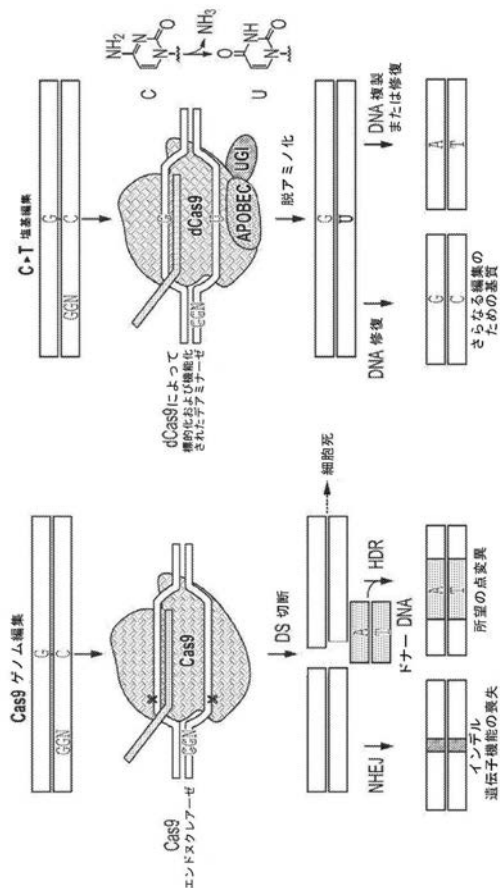


図 12

【 図 1 3 】

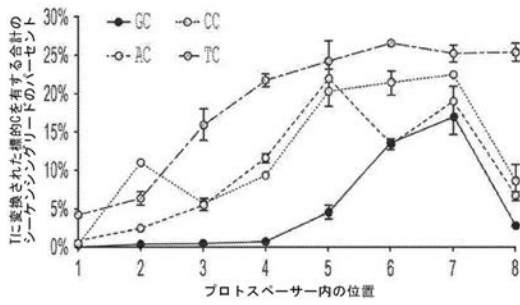


図 13

【 図 1 4 】

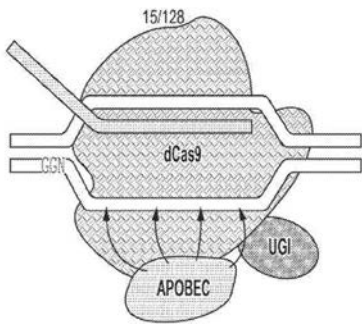


図 14

【 図 1 6 】

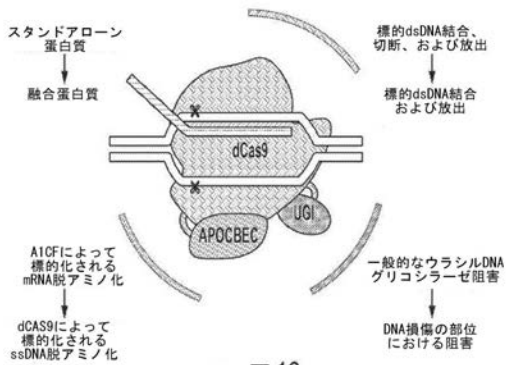


図 16

【 図 1 7 】

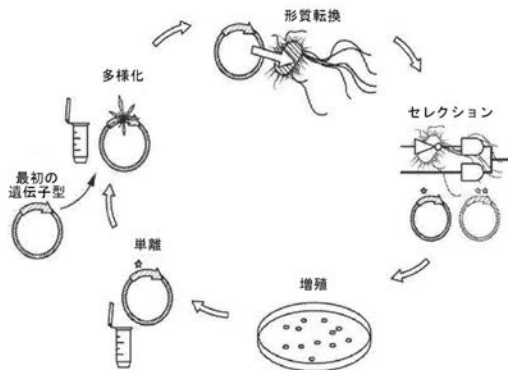


図 17

【 図 1 5 】

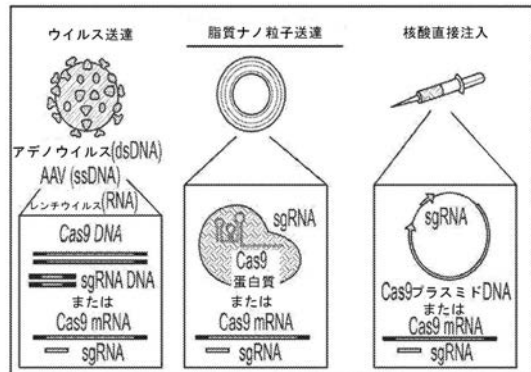
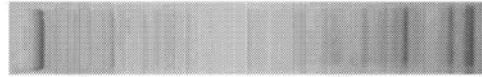


図 15

【 図 1 8 】

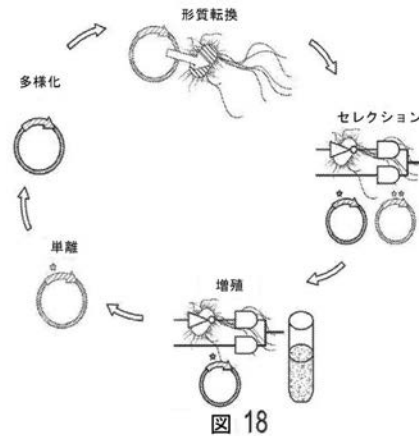


図 18

【 図 19 】

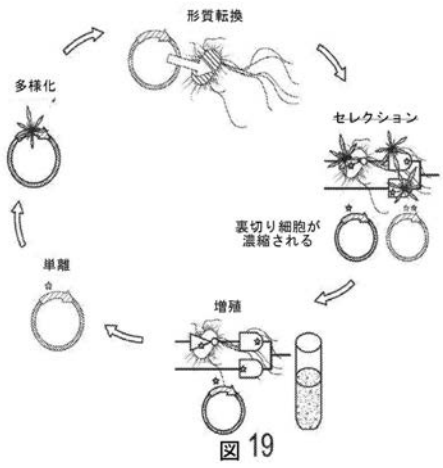


図 19

【 図 20 】

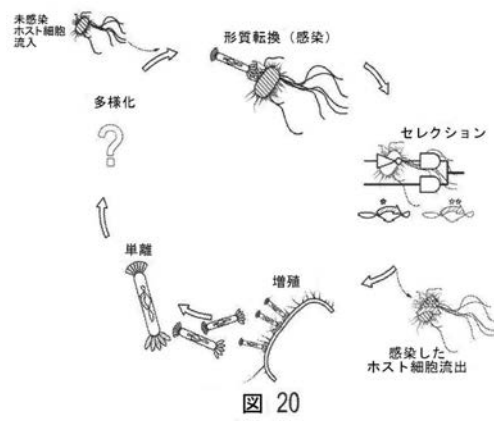


図 20

【 図 21 】

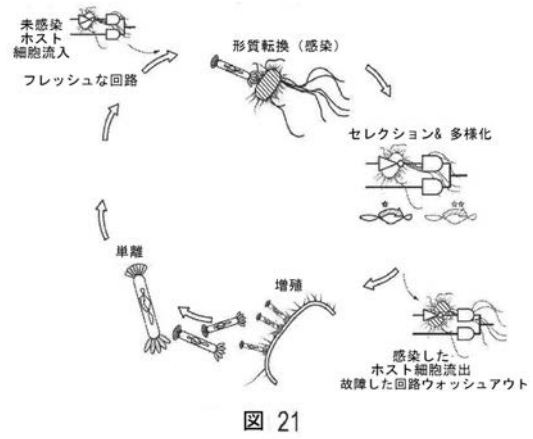


図 21

【 図 22 】

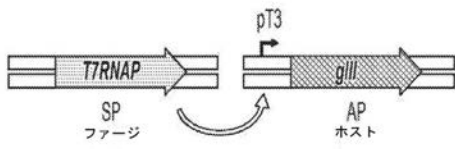


図 22

【 図 23 】

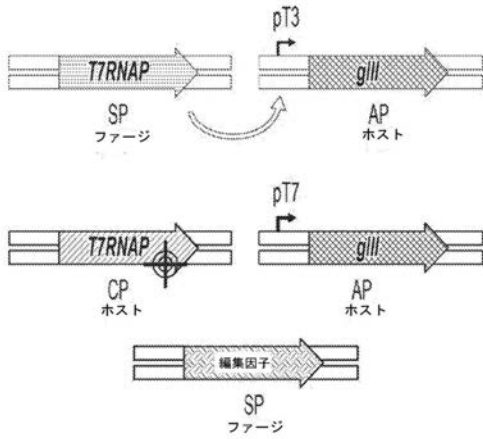


図 23

【 図 24 】

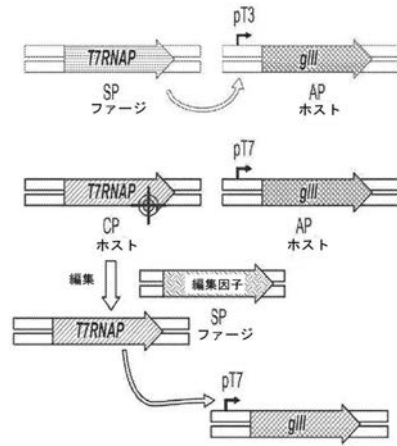


図 24

【 図 2 5 】

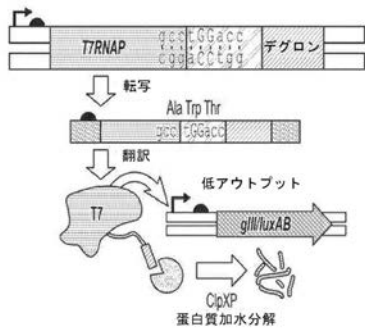


図 25

【 図 2 6 】

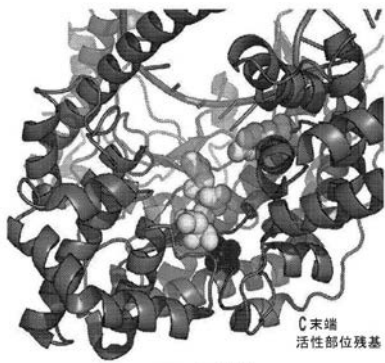


図 26

【 図 2 8 】

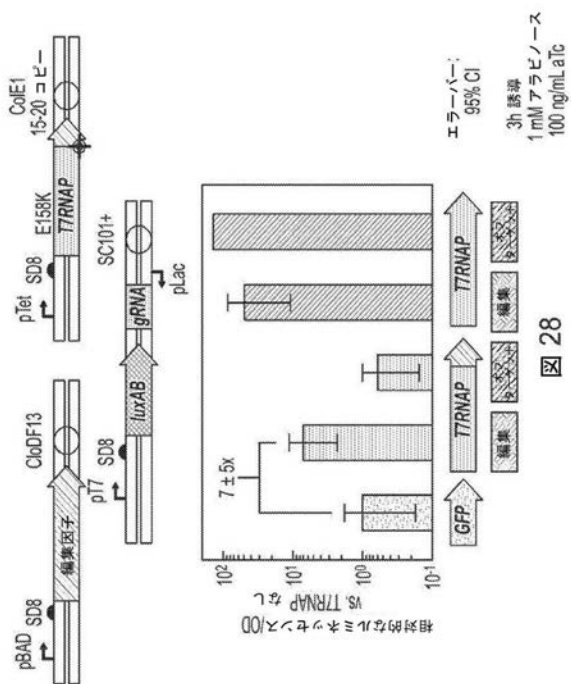


図 28

【 図 2 7 】

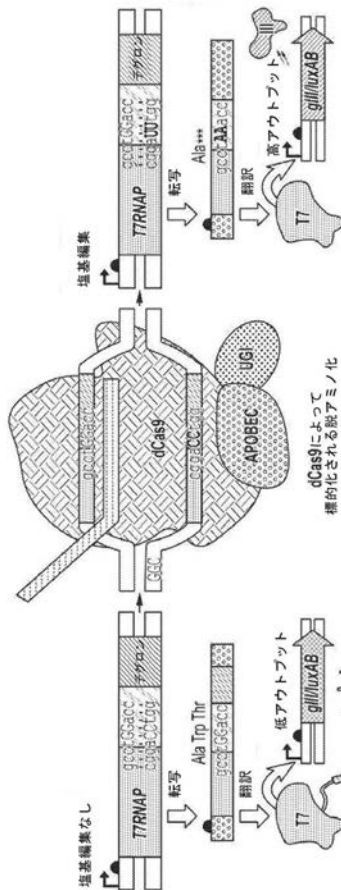


図 27

【 図 2 9 】

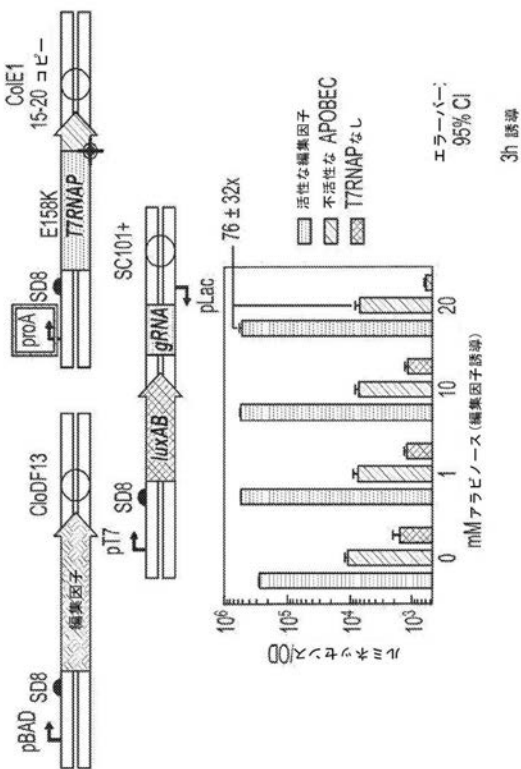
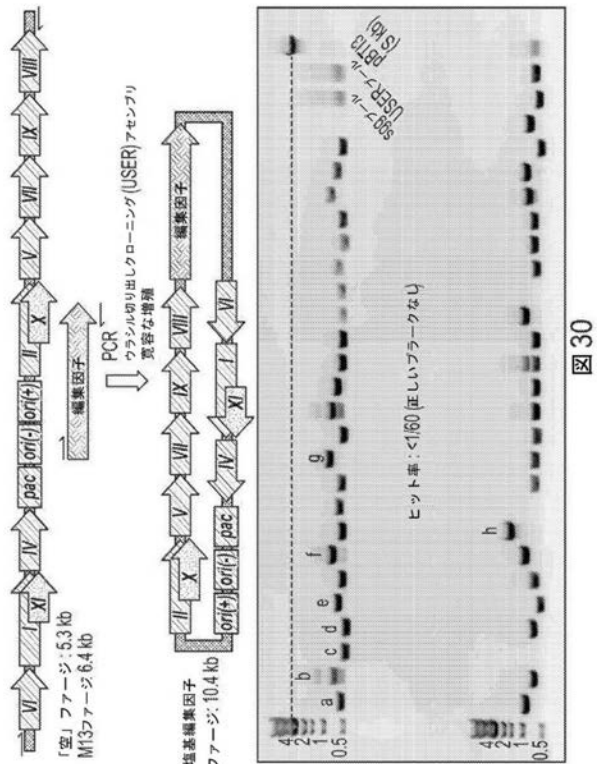


図 29

【 図 3 0 】



【 図 3 1 】

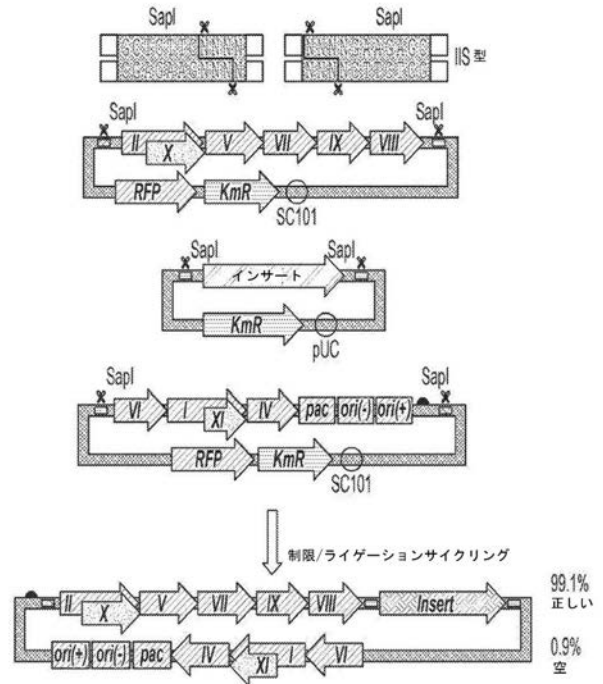
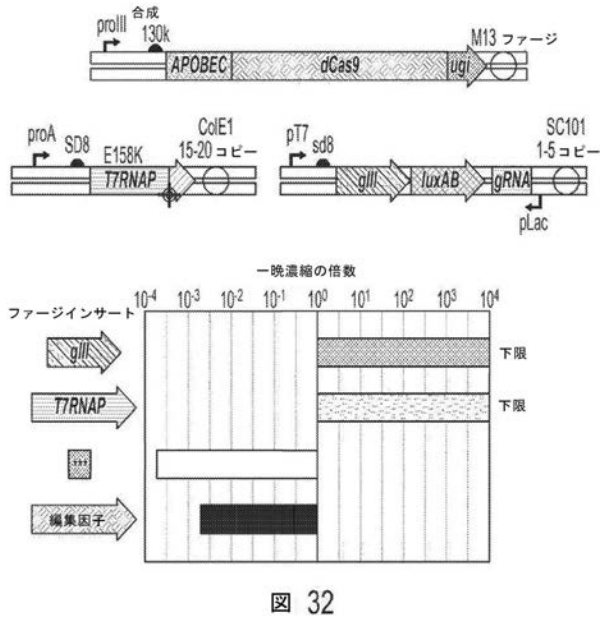
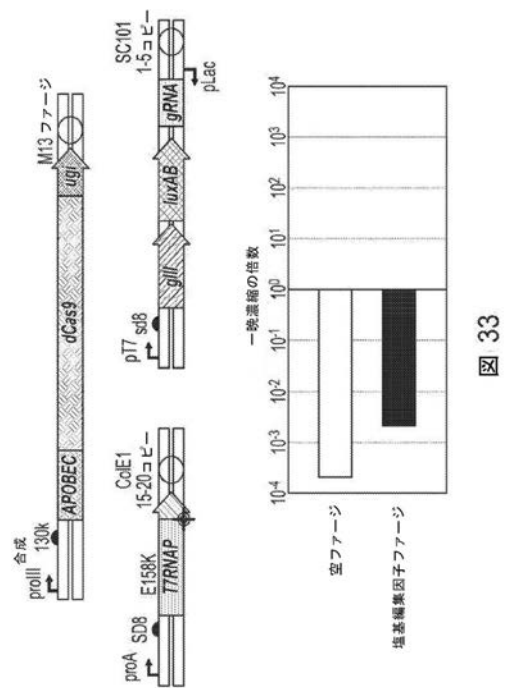


図 31

【 図 3 2 】



【 図 3 3 】



【 図 3 4 】

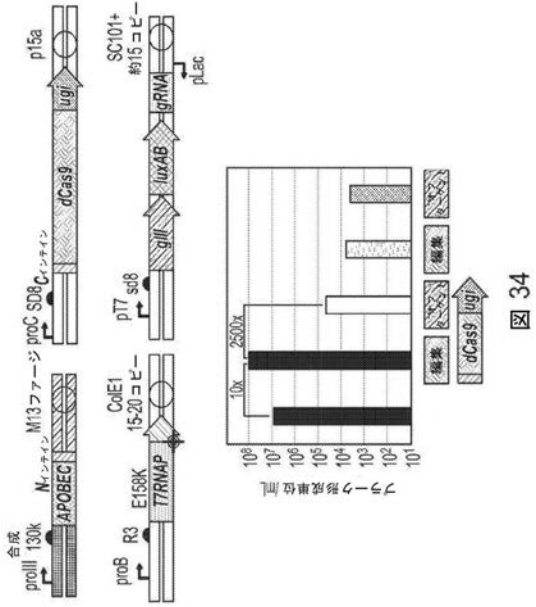


図 34

【 図 3 5 】

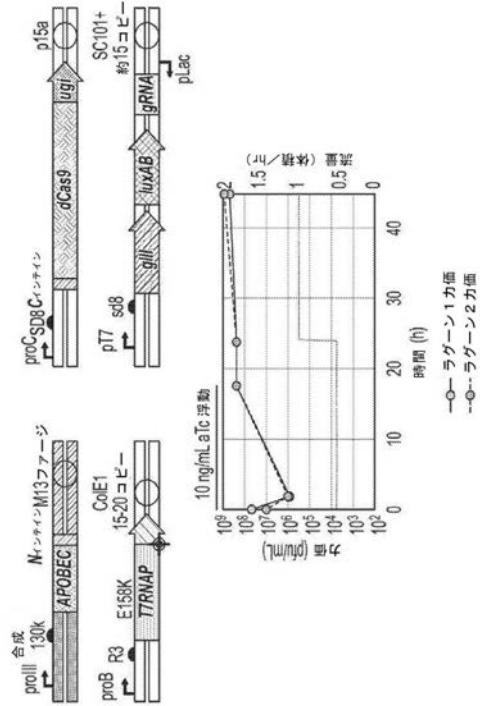


図 35

【 図 3 6 】

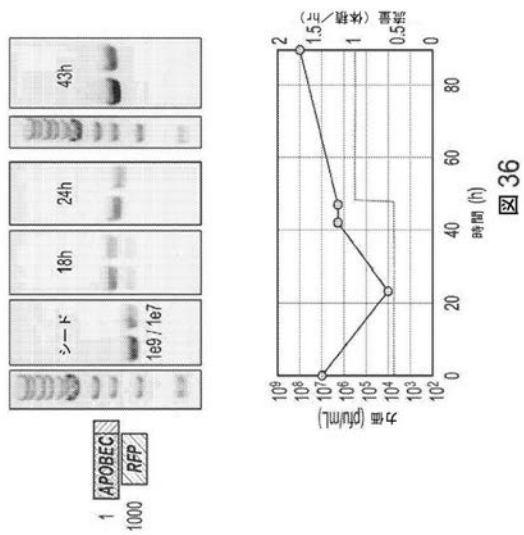


図 36

【 図 3 7 - 1 】

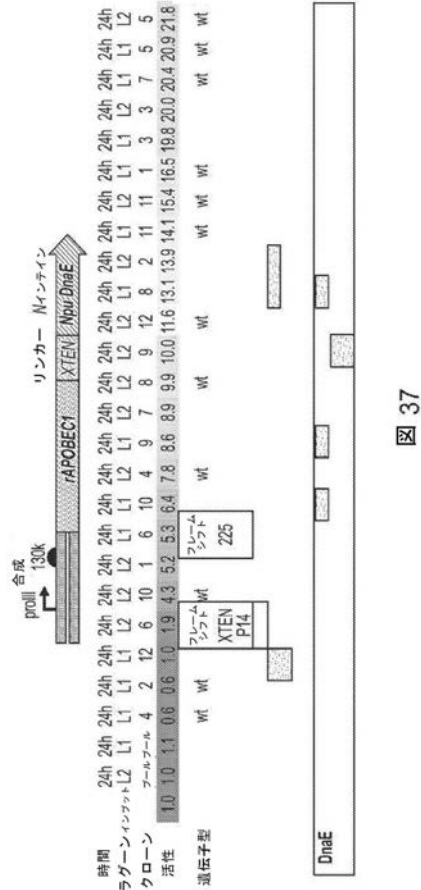


図 37

【 図 4 0 - 2 】

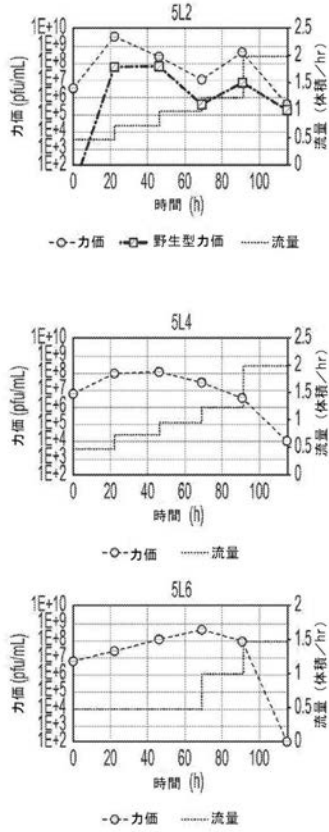


図 40 続き

【 図 4 2 】

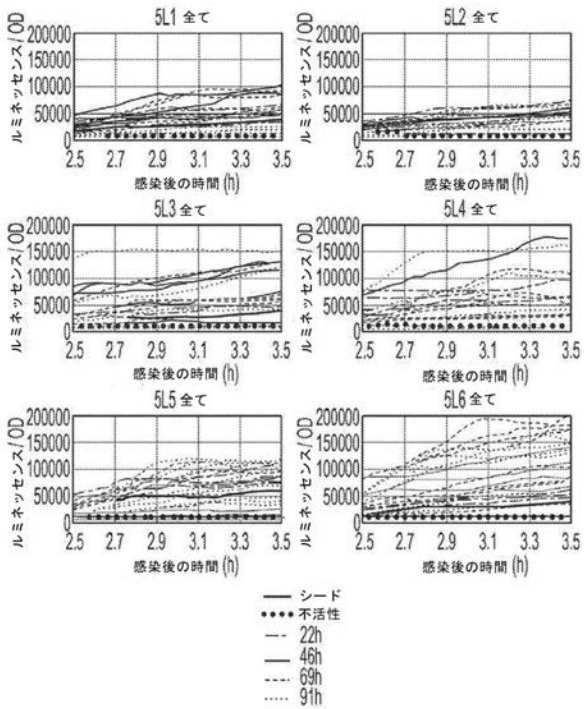


図 42

【 図 4 1 】

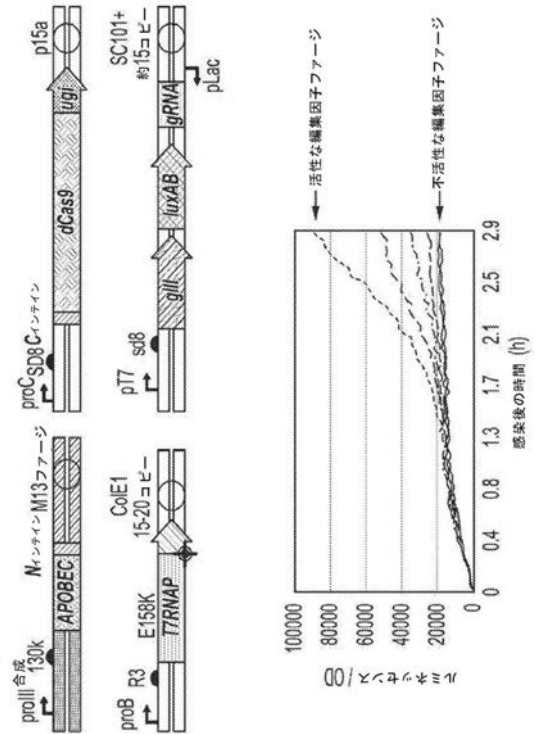


図 41

【 図 4 3 】

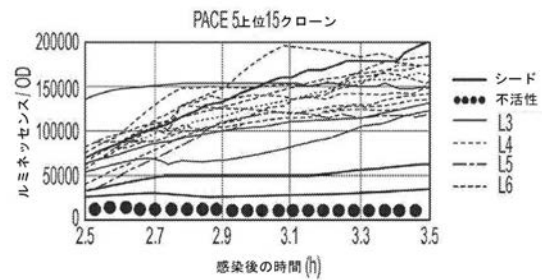


図 43

【 図 4 4 】

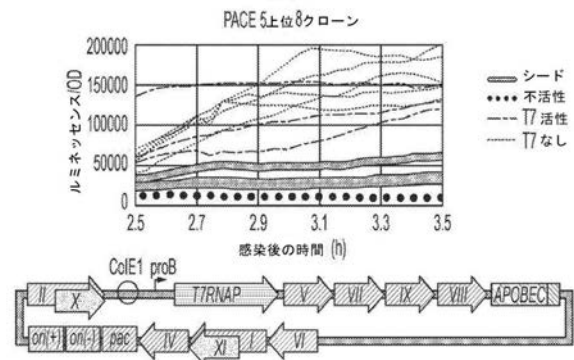


図 44

【 図 4 5 】

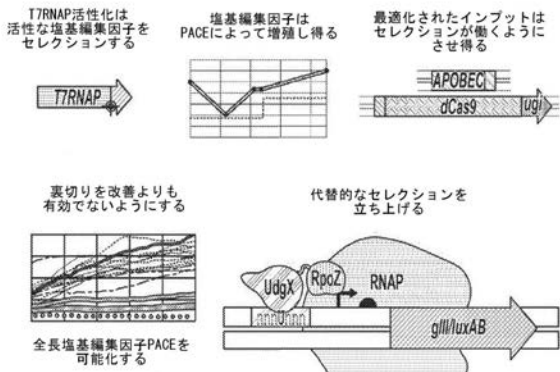


図 45

【 図 4 6 】

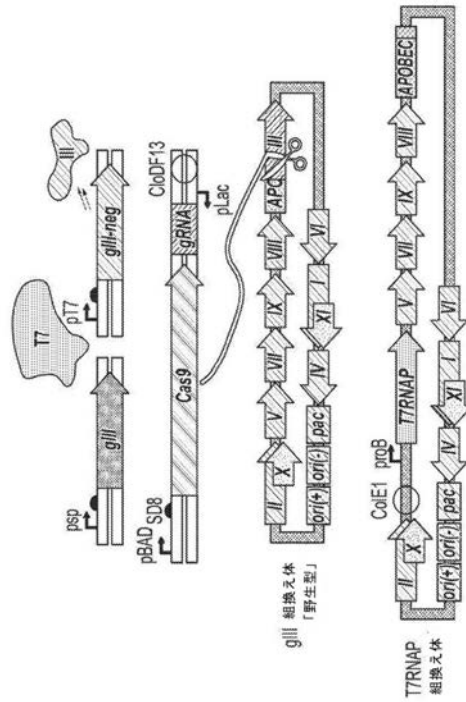


図 46

【 図 4 7 】

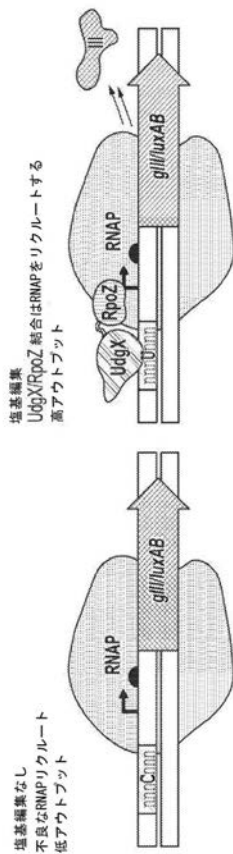


図 47

【 図 4 8 】

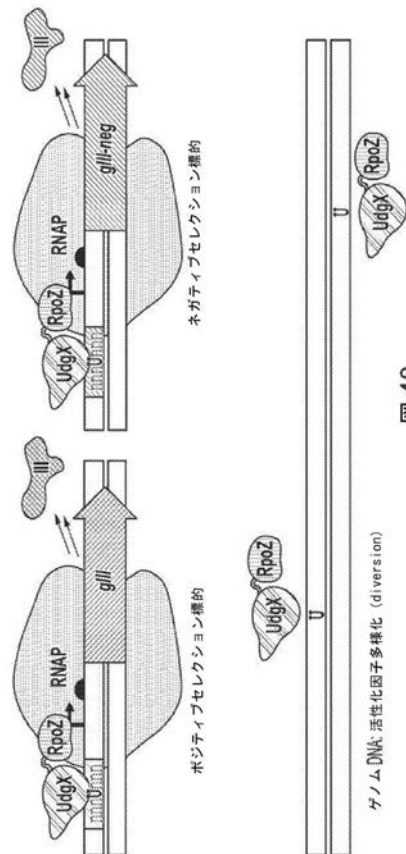


図 48

【 図 4 9 】

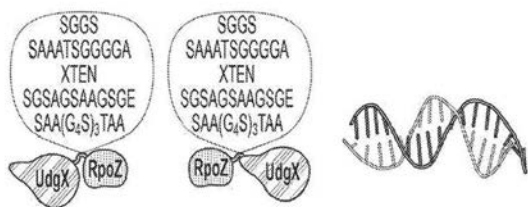


図 49

【 図 5 0 】

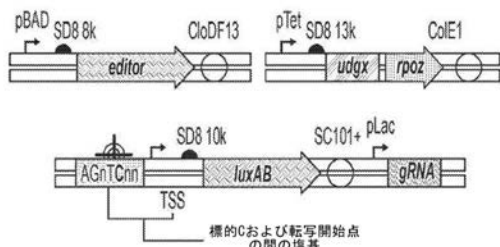


図 50

【 図 5 1 】

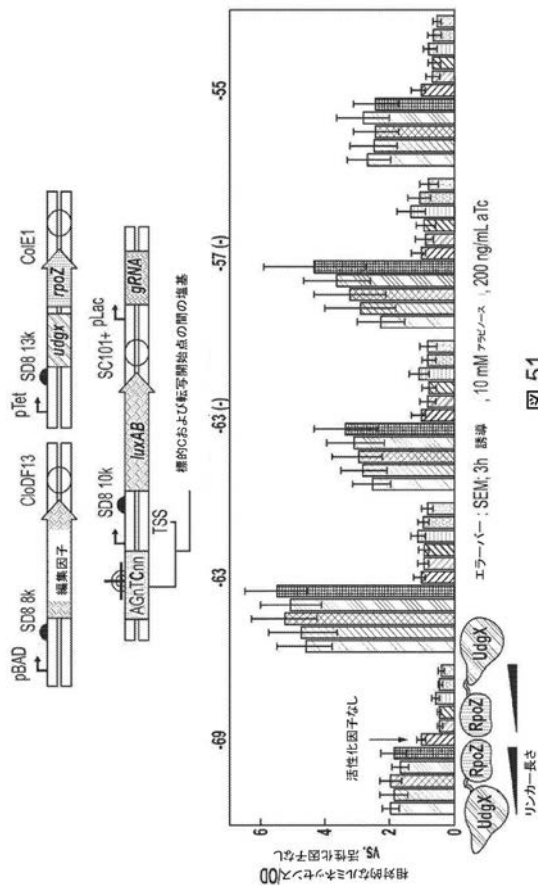


図 51

【 図 5 2 】

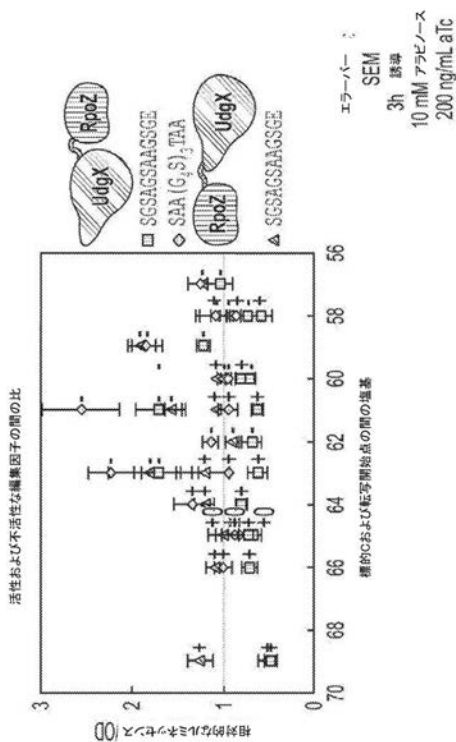


図 52

【 図 5 3 】

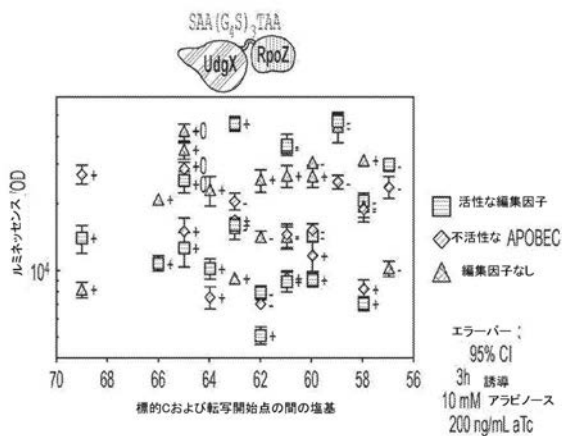


図 53

【 図 5 4 】

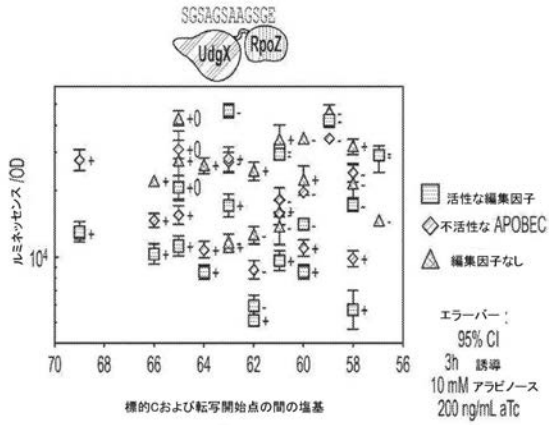


図 54

【 図 5 5 】

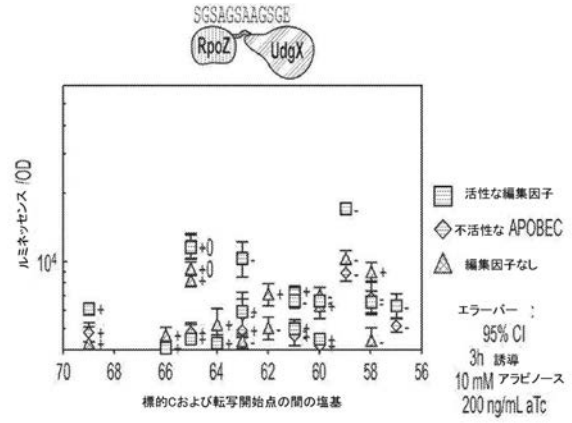


図 55

【 図 5 6 】

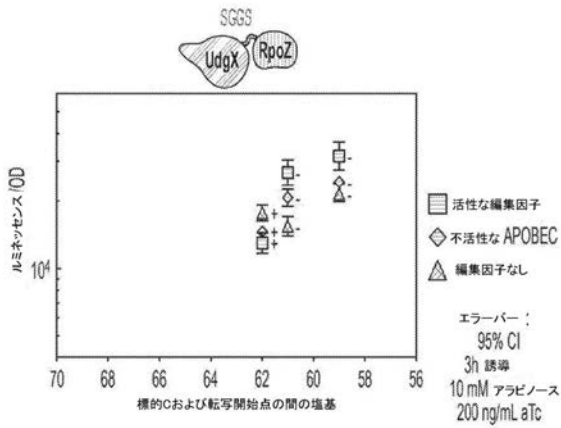


図 56

【 図 5 7 】

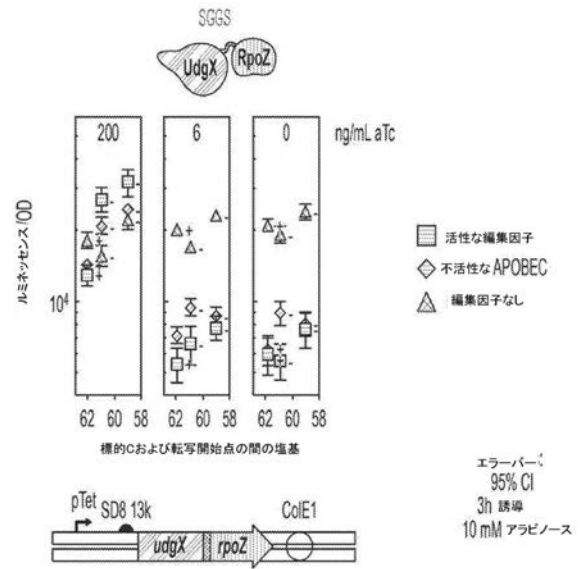


図 57

【 図 5 8 】

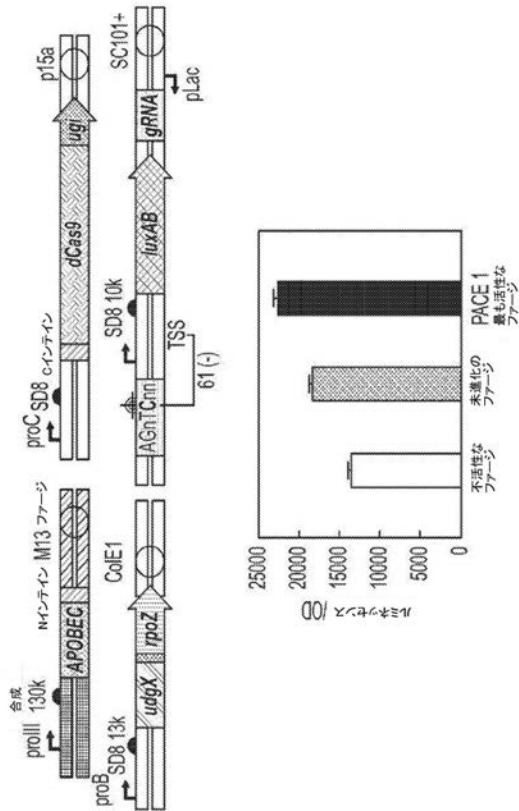


図 58

【 図 5 9 】

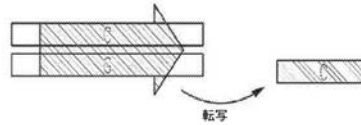


図 59

【 図 6 0 】

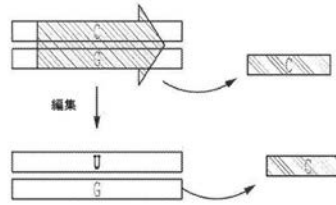


図 60

【 図 6 1 】

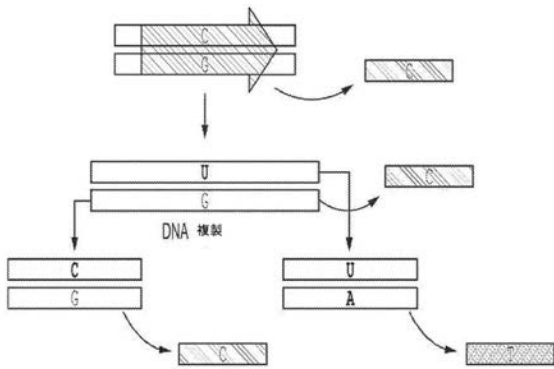


図 61

【 図 6 3 】

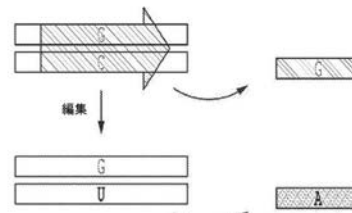


図 63

【 図 6 2 】

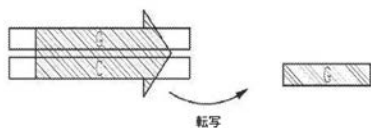


図 62

【 図 6 4 】

非機能的なGIによってコードされるAA	Arg	Glu	Cys	Ala	Ser	Asp	Arg	Val	Trp
機能的なAIによってコードされるAA	His	Lys	Tyr	Thr	Asn	Asn	Gln	Ile/Met	終止
必須のコードするGを有するAA (あり得るオフターゲット変化)	Ala	Asp	Glu	Cys	Gly	Arg	Val	Met	
	Thr	Asn	Lys	Tyr	Asn	Lys	Ile	Ile	
					Lys	His			
					Asp				
					Glu				

図 64

【 図 6 5 】

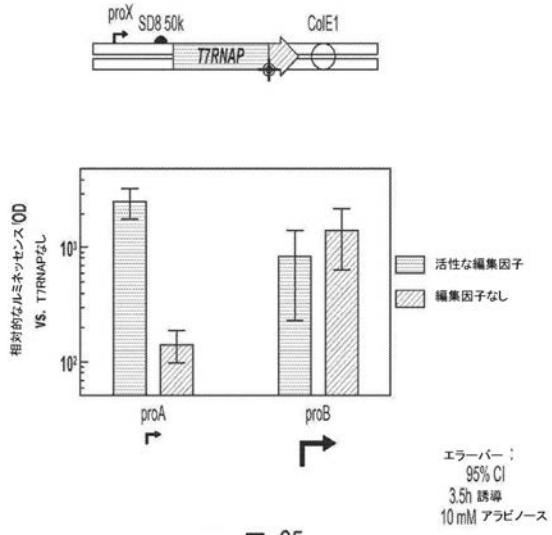


図 65

【 図 6 6 】

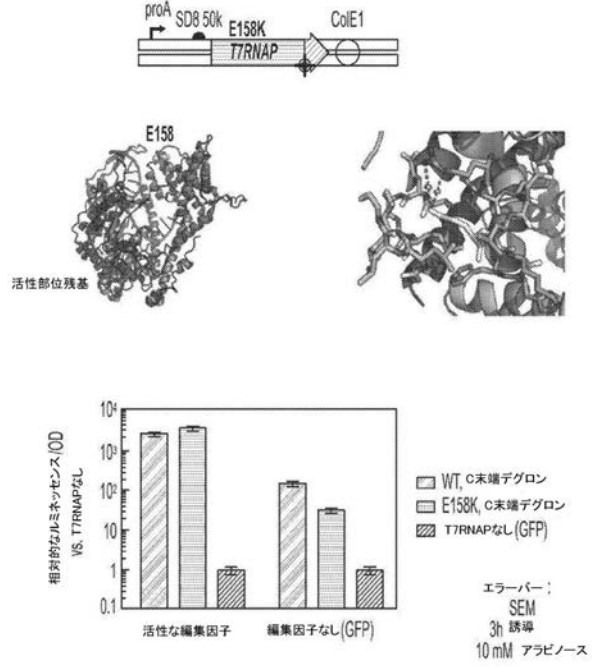


図 66

【 図 6 7 】

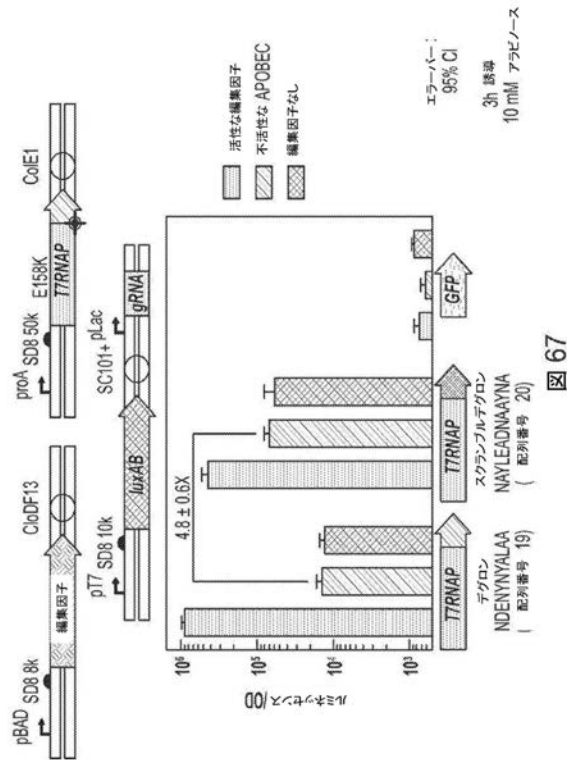


図 67

【 図 6 8 】

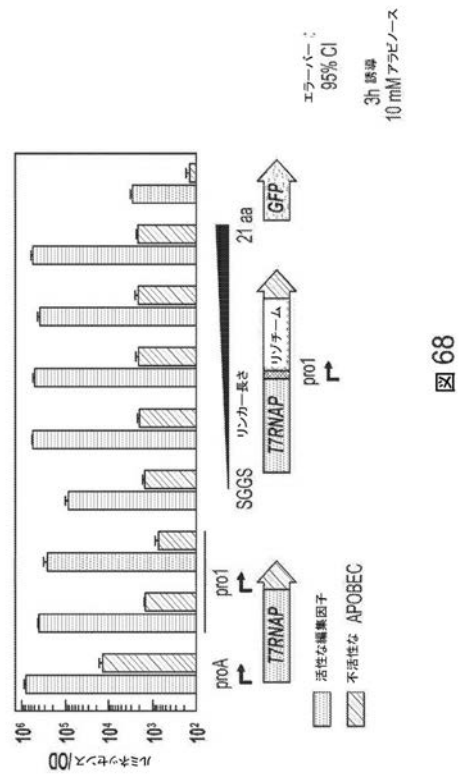


図 68

【 図 69 】

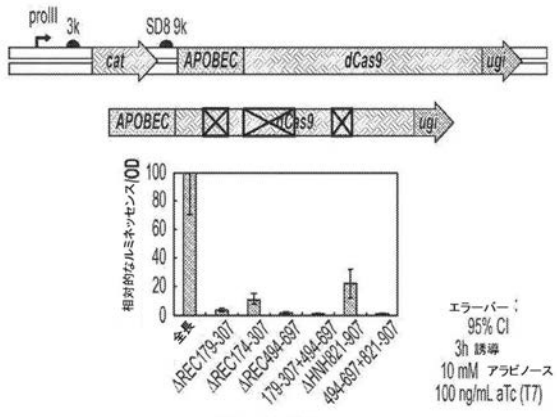


図 69

【 図 70 】

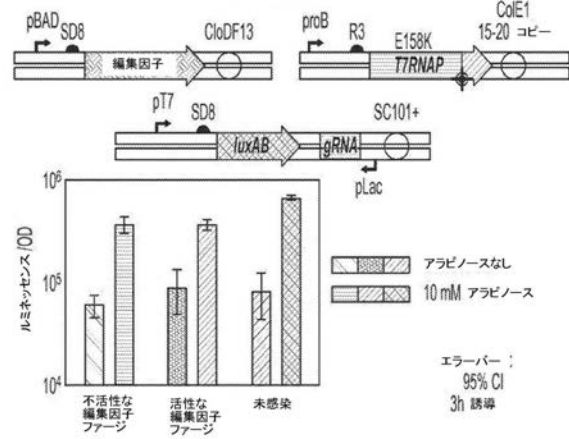


図 70

【 図 71 】

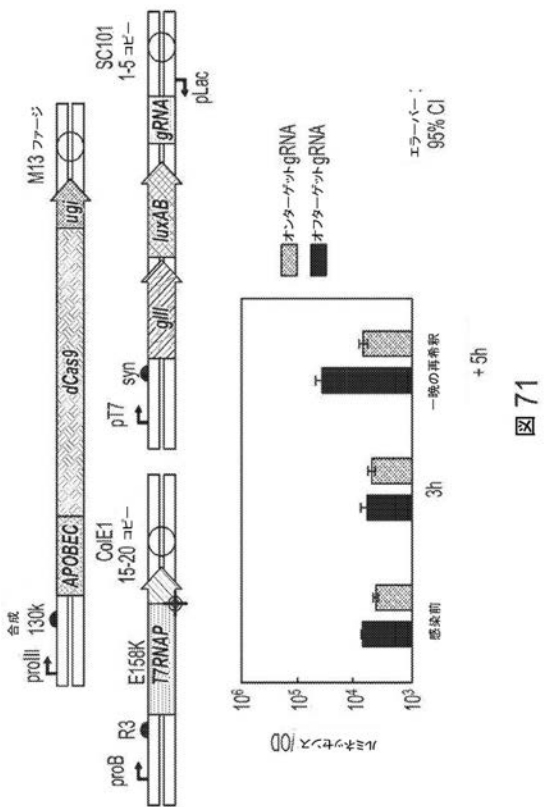


図 71

【 図 72 】

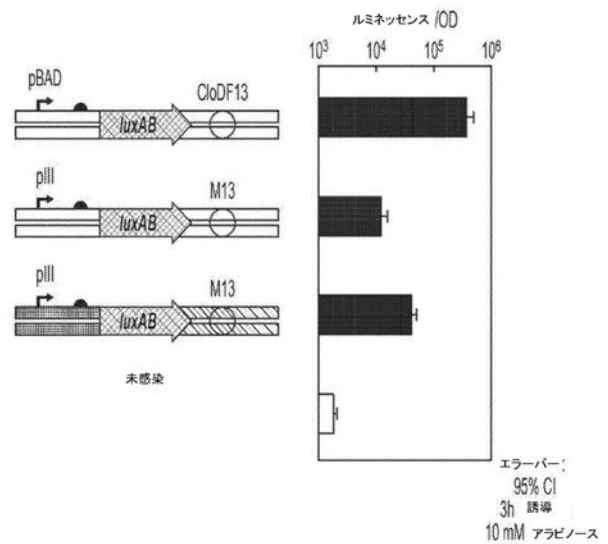


図 72

【 図 7 3 】

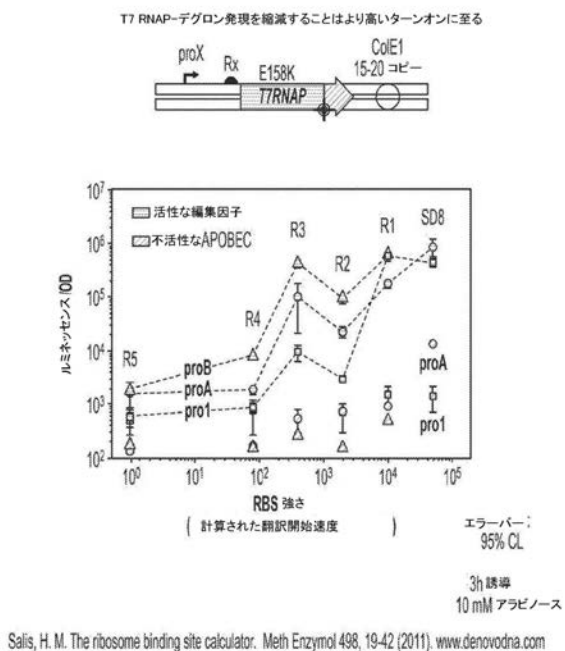


図 73

【 図 7 4 】

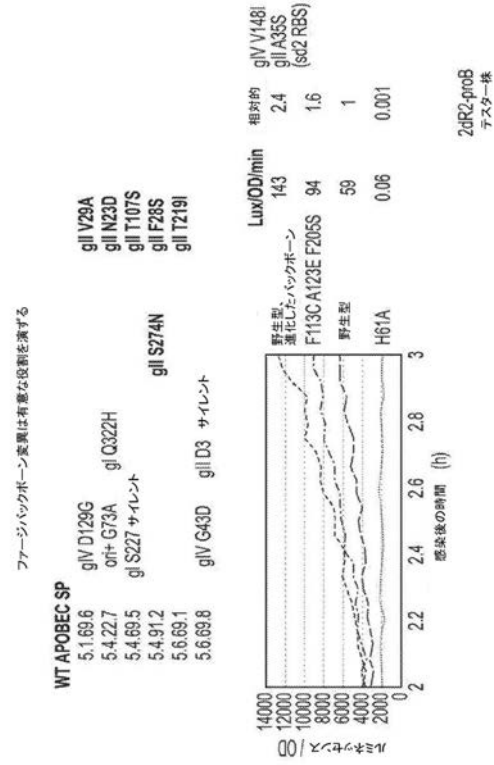


図 74

【 図 7 5 】

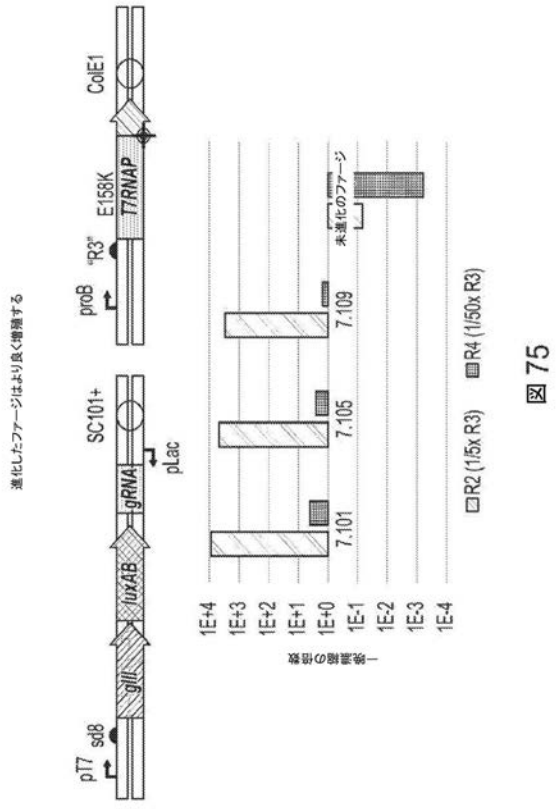


図 75

【 図 7 6 】

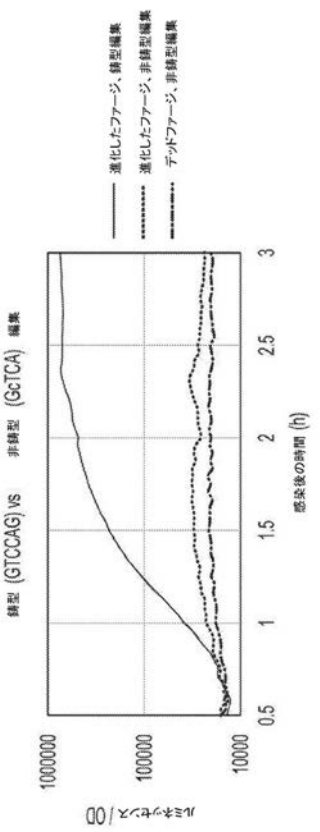


図 76

【 図 77 】

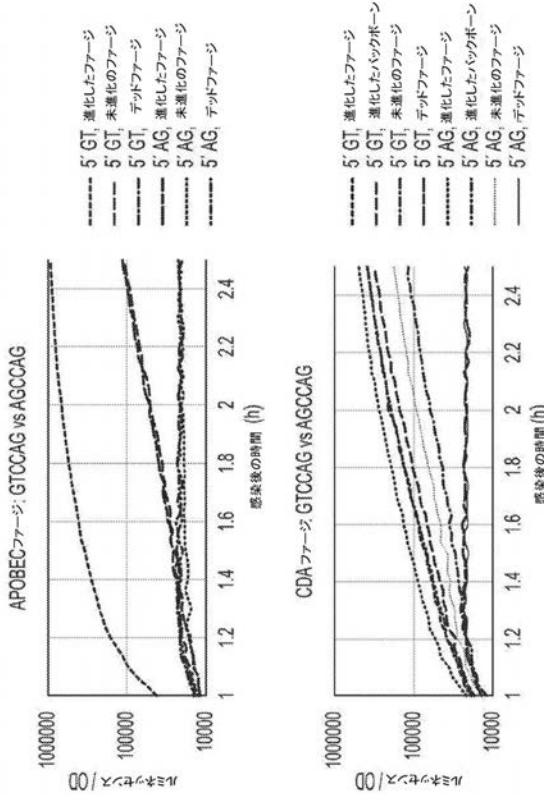


図 77

【 図 78 】

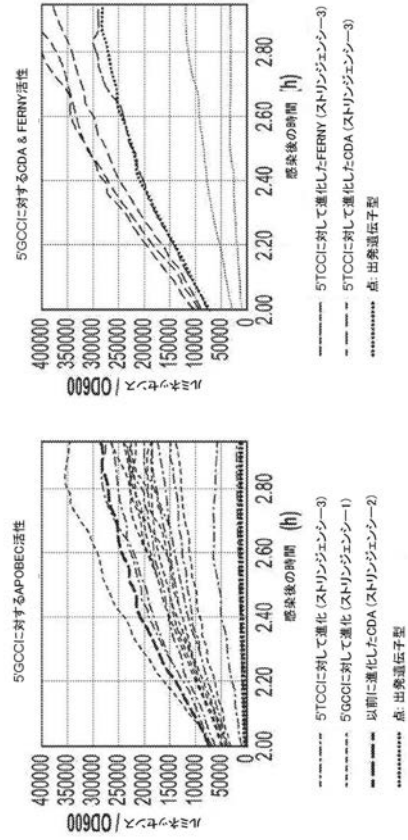


図 78

【 図 79 】

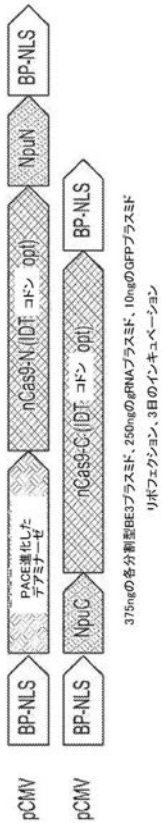


図 79

【 図 80 】

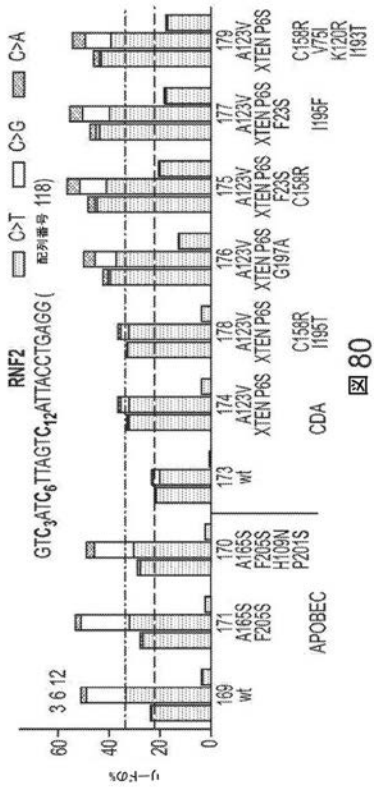


図 80

【 8 1 】

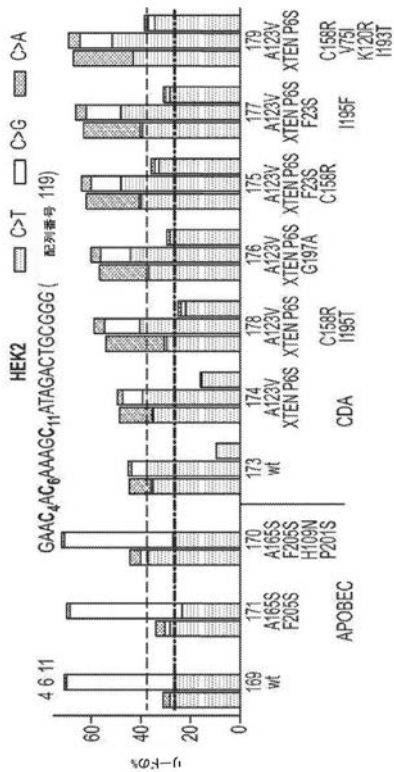


図 81

【 8 2 】

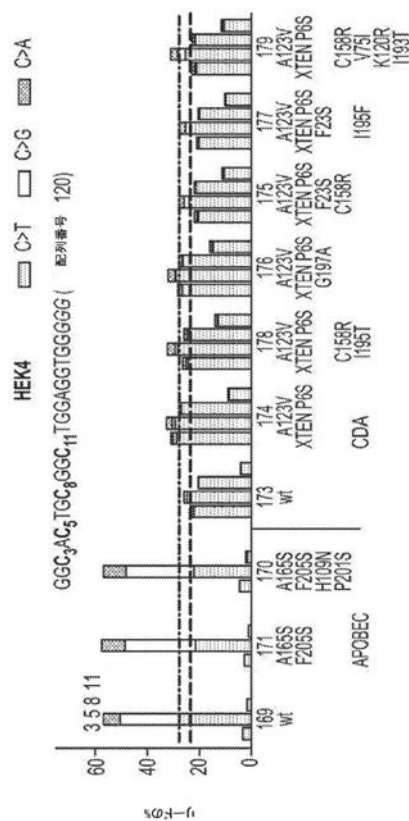


図 82

【 8 3 】

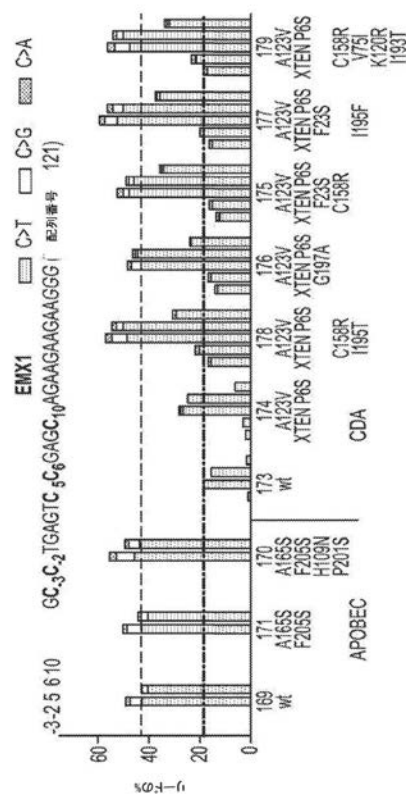


図 83

【 8 4 】

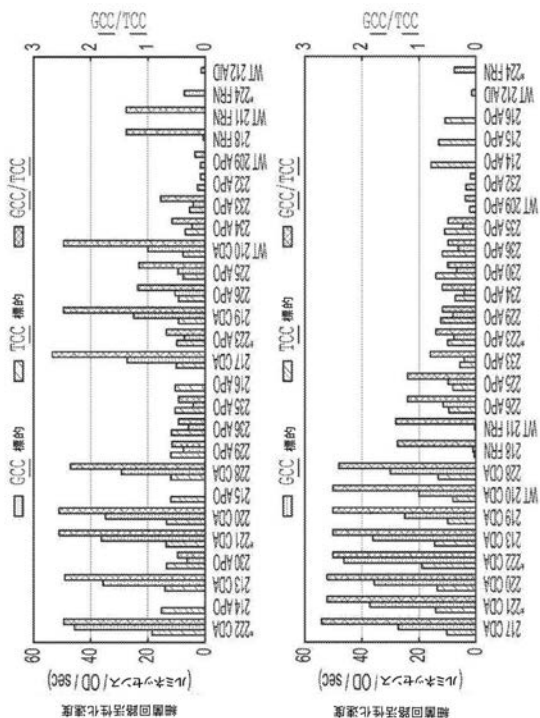
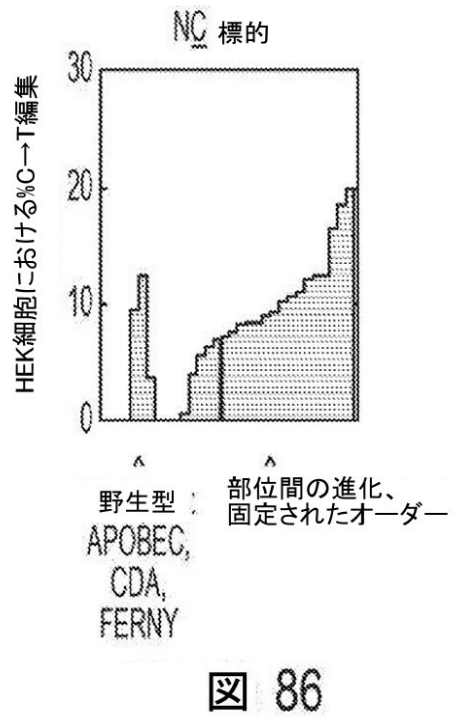


図 84

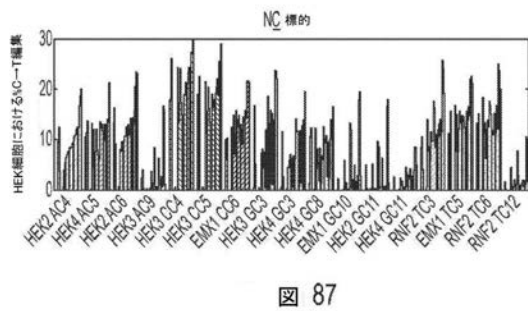
【 図 8 5 】



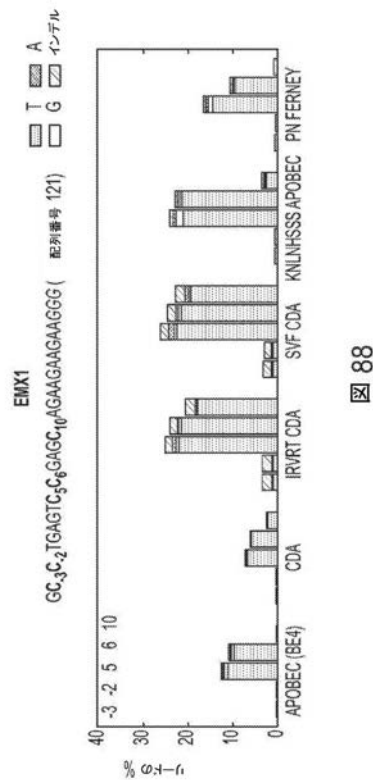
【 図 8 6 】



【 図 8 7 】



【 図 8 8 】



【 89 】

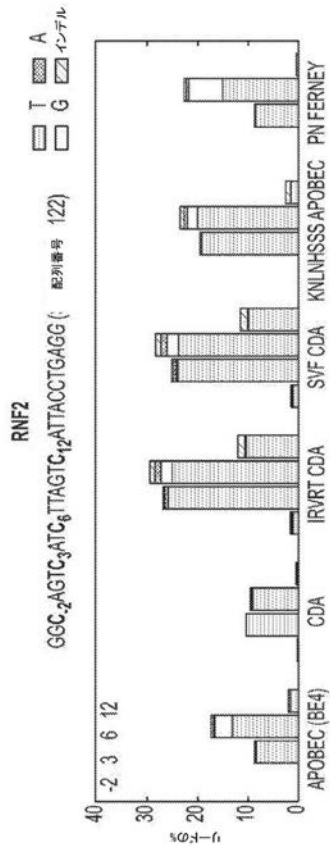


図 89

【 90 】

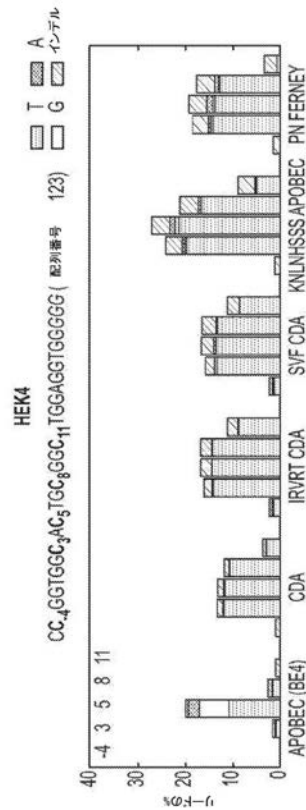


図 90

【 91 】

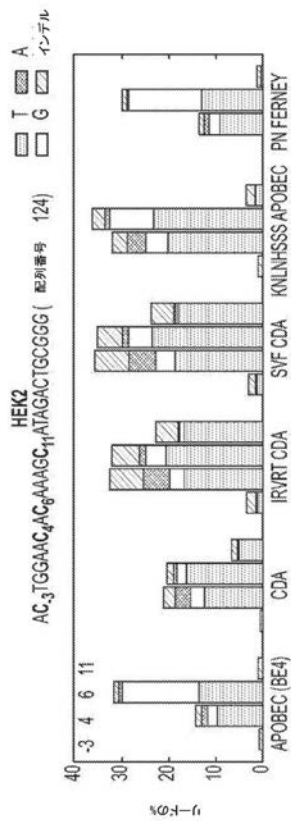


図 91

【 92 】

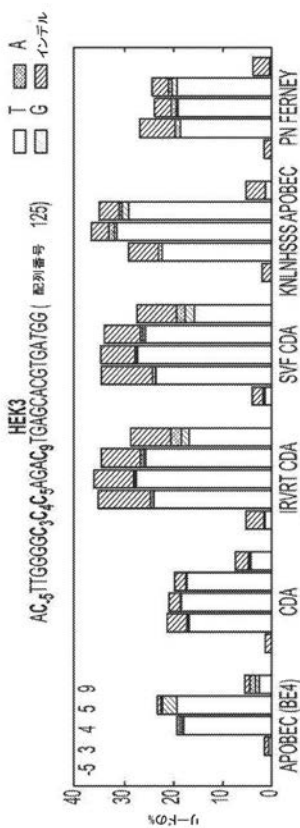
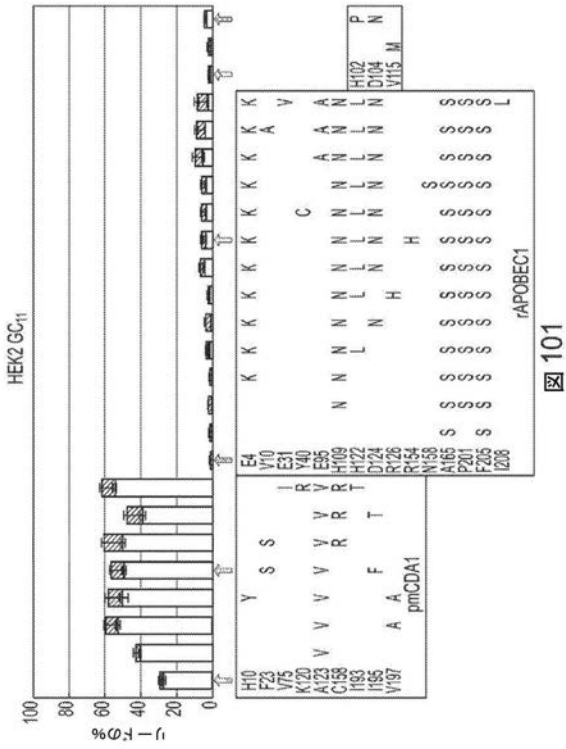


図 92

【 101 】



【 図 1 1 3 】

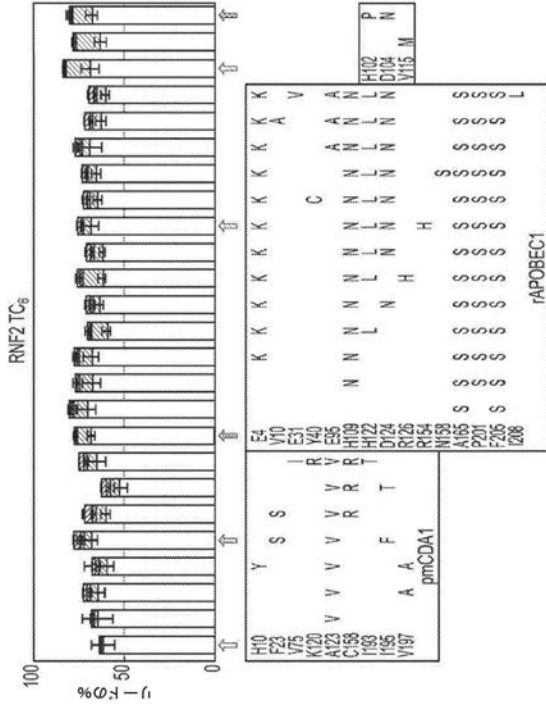


図 113

【 図 1 1 4 】

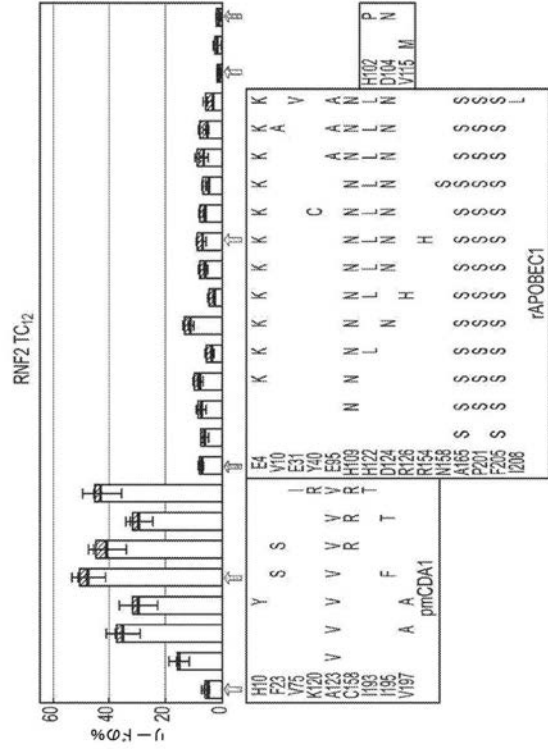
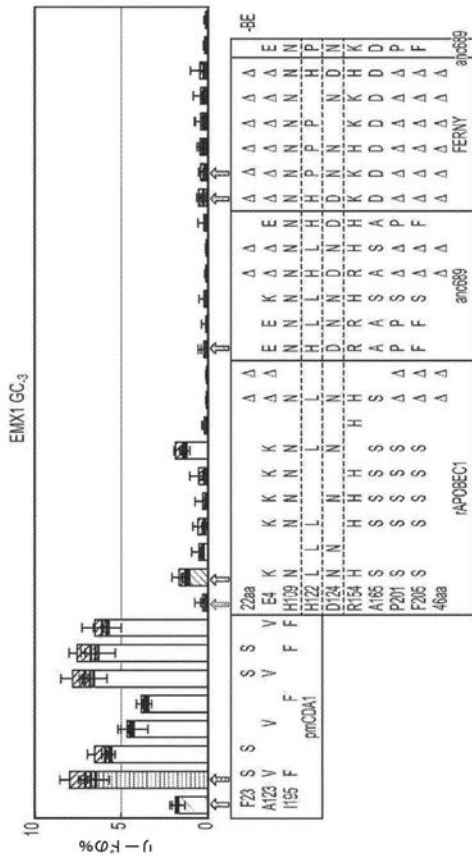
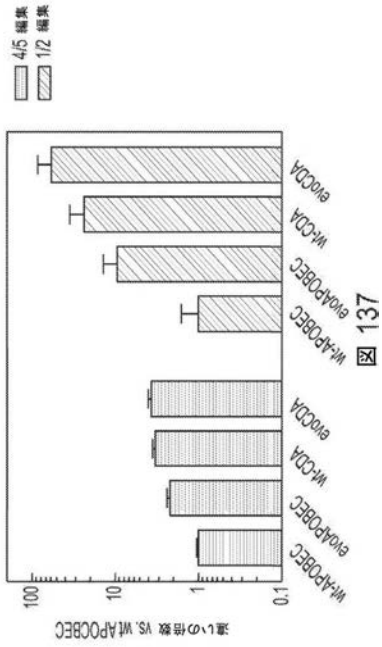


図 114

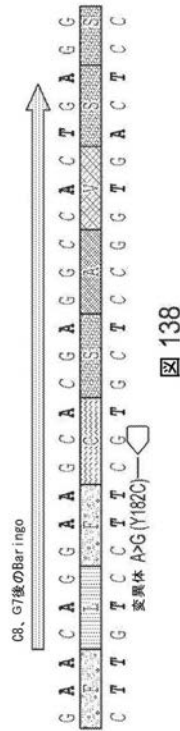
【 図 1 1 5 】



【 図 1 3 7 】



【 図 1 3 8 】



【 図 1 3 9 】

サイレント mut.	サイレント mut.	
	位置 1	位置 8
1	0.1%	1.9%
2	37.8%	33.4%
3	0.1%	1.1%
4	1.8%	1.9%
5	16.8%	8.0%
6	0.1%	0.2%
7	0.4%	5.5%
8	0.0%	0.1%
9	0.0%	0.0%

図 139

【 図 1 4 0 】

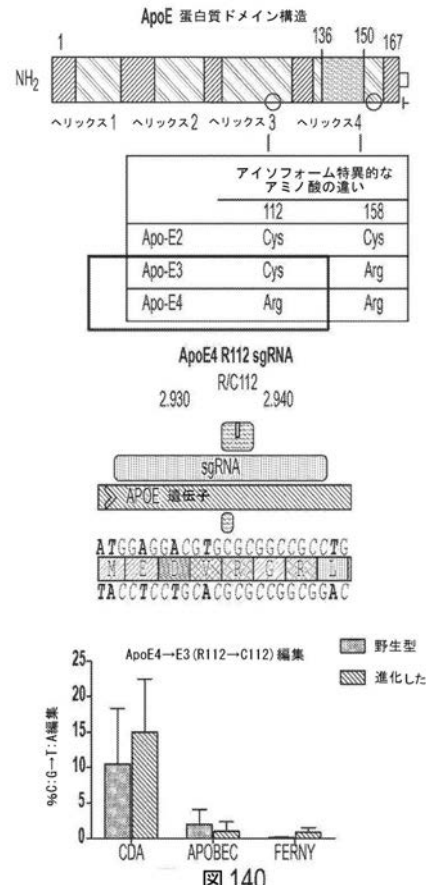
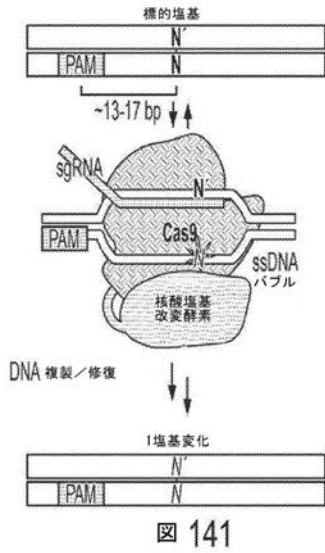
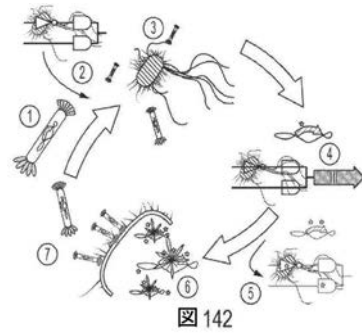


図 140

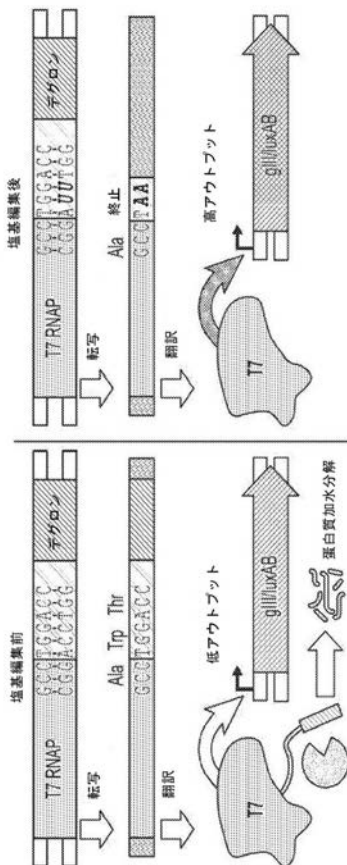
【 図 1 4 1 】



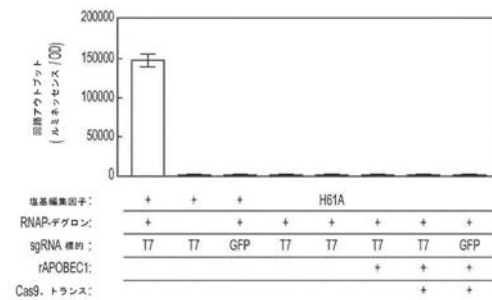
【 図 1 4 2 】



【 図 1 4 3 - 1 】



【 図 1 4 3 - 2 】



【 図 1 4 4 】

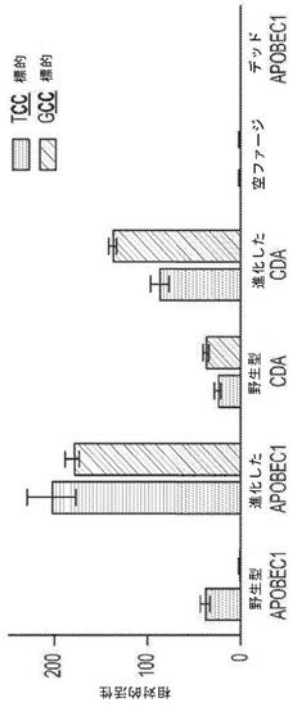


図 144

【 配 列 表 】

2020534795000001.app

【 国際調査報告 】

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2018/044242

Box No. 1 Nucleotide and/or amino acid sequence(s) (Continuation of Item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:
- a. forming part of the international application as filed:
- in the form of an Annex C/ST.25 text file.
- on paper or in the form of an image file.
- b. furnished together with the international application under PCT Rule 13ter.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.
- c. furnished subsequent to the international filing date for the purposes of international search only:
- in the form of an Annex C/ST.25 text file (Rule 13ter.1(a)).
- on paper or in the form of an image file (Rule 13ter.1(b) and Administrative Instructions, Section 713).
2. In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.
3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2018/044242**Box No. II Observations where certain claims were found unsearchable (Continuation of Item 2 of first sheet)**

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of Item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.

3. As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-10(completely); 89-93, 108, 110, 116-121, 142(partially)

Remark on Protest

- The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2018/044242

A. CLASSIFICATION OF SUBJECT MATTER		
INV. C07K14/32	C12N9/22	C12N9/24 C12N9/78
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)		
C07K C12N		
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, WPI Data, BIOSIS, Sequence Search, EMBASE, CHEM ABS Data		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	W0 2017/070632 A2 (HARVARD COLLEGE [US]) 27 April 2017 (2017-04-27) paragraph [0005] - paragraph [0284]; claims 1-302; sequences 1-5808	1-10, 89-93, 108,110, 116-121, 142
X	US 2015/165054 A1 (LIU DAVID R [US] ET AL) 18 June 2015 (2015-06-18) column 4 - column 13; examples 1-6 ----- -/--	1-10, 89-93, 108,110, 116-121, 142
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> 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 *Z* document member of the same patent family
Date of the actual completion of the international search		Date of mailing of the international search report
18 September 2018		21/11/2018
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 Seranski, Peter

2

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2018/044242

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	EP 3 115 457 A1 (NAT UNIV CORP KOBE UNIV [JP]) 11 January 2017 (2017-01-11) the whole document -----	1-10, 89-93, 108,110, 116-121, 142
A	ALEXIS C. KOMOR ET AL: "Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage", NATURE, vol. 533, no. 7603, 20 April 2016 (2016-04-20), pages 420-424, XP055343871, GB ISSN: 0028-0836, DOI: 10.1038/nature17946 the whole document -----	1-10, 89-93, 108,110, 116-121, 142

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2018/044242

Patent document cited in search report	Publication date	Patent family member(s)	Publication date		
WO 2017070632 A2	27-04-2017	AU 2016342380 A1	10-05-2018		
		CA 3002827 A1	27-04-2017		
		CN 108513575 A	07-09-2018		
		CN 108699116 A	23-10-2018		
		EP 3365356 A2	29-08-2018		
		EP 3365357 A2	29-08-2018		
		KR 20180069898 A	25-06-2018		
		SG 11201803173V A	30-05-2018		
		US 2017121693 A1	04-05-2017		
		US 2018312825 A1	01-11-2018		
		WO 2017070632 A2	27-04-2017		
		WO 2017070633 A2	27-04-2017		
		US 2015165054 A1	18-06-2015	AU 2014362208 A1	30-06-2016
				CA 2933625 A1	18-06-2015
CN 105934516 A	07-09-2016				
EP 3080265 A1	19-10-2016				
JP 2017500035 A	05-01-2017				
US 2015165054 A1	18-06-2015				
US 2015166980 A1	18-06-2015				
US 2015166981 A1	18-06-2015				
US 2015166982 A1	18-06-2015				
US 2015166983 A1	18-06-2015				
US 2015166984 A1	18-06-2015				
US 2015166985 A1	18-06-2015				
US 2016304846 A1	20-10-2016				
WO 2015089406 A1	18-06-2015				
EP 3115457 A1	11-01-2017	CA 2947941 A1	11-09-2015		
		CN 106459957 A	22-02-2017		
		EP 3115457 A1	11-01-2017		
		JP 6206893 B2	04-10-2017		
		JP 2018019705 A	08-02-2018		
		JP W02015133554 A1	06-04-2017		
		SG 11201609211V A	29-12-2016		
		US 2017073670 A1	16-03-2017		
		WO 2015133554 A1	11-09-2015		

International Application No. PCT/ US2018/ 044242

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. claims: 1-10(completely); 89-93, 108, 110, 116-121, 142(partially)

A cytidine deaminase comprising an amino acid sequence that is at least 80%, 85%, 90%, 95%, 98%, 99%, or 99.5% identical to amino acid residues 2-162 of SEQ ID NO: 1, wherein the cytidine deaminase comprises one or more mutations selected from the group consisting of H102Xi , D104X2 , and V115X3 relative to SEQ ID NO: 1, or a corresponding mutation(s) in another cytidine deaminase, wherein X1 is any amino acid other than H, X2 is any amino acid other than D, and X3 is any amino acid other than V; Nucleic acid encoding said protein; vector; cell; kit;

2. claims: 11-29(completely); 89-93, 108, 110, 116-121, 142(partially)

A cytidine deaminase comprising an amino acid sequence that is at least 80%, 85%, 90%, 95%, 98%, 99%, or 99.5% identical to amino acid residues 3-229 of SEQ ID NO: 2, wherein the cytidine deaminase comprises one or more mutations selected from the group consisting of E4Xi , V10X2 , E31X3 , Y40X4 , E95X5 , H109X6 , H122X7 , D124X8 , R126X9 , R154X10 , N158X11 , A165X12 , P201X13 , F205X14 , and I208X15 relative to SEQ ID NO: 2, or a corresponding mutation(s) in another cytidine deaminase, wherein X1 , X3 , and X5 are any amino acid other than E, X2 is any amino acid other than V, X4 is any amino acid other than Y, X6 and X7 are any amino acid other than H, X8 is any amino acid other than D, X9 and X10 are any amino acid other than R, X11 is any amino acid other than N, X12 is any amino acid other than A, X13 is any amino acid other than P, X14 is any amino acid other than F, and X15 is any amino acid other than I; Nucleic acid encoding said protein; vector; cell; kit;

3. claims: 30-47(completely); 89-93, 108, 110, 116-121, 142(partially)

A cytidine deaminase comprising an amino acid sequence that is at least 80%, 85%, 90%, 95%, 98%, 99%, or 99.5% identical to amino acid residues 2-208 of SEQ ID NO: 3, wherein the cytidine deaminase comprises one or more mutations selected from the group consisting of H10X1 , F23X2 , V75X3 , K120X4 , A123X5 , C158X6 , I193X7 , I195X8 , and V197X9 relative to SEQ ID NO: 3, or a corresponding mutation(s) in another cytidine deaminase, wherein X1 , is any amino acid other than H, X2 is any amino acid other than F, X3 and X9 are any amino acid other than V, X4 is any amino acid other than K, X5 is any amino acid other than A, X6 is any amino acid other than C, and X7 and X8 are any amino acid other than I.

International Application No. PCT/ US2018/ 044242

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

4. claims: 48-64(completely); 89-93, 108, 110, 116-121, 142(partially)

A cytidine deaminase comprising an amino acid sequence that is at least 80%, 85%, 90%, 95%, 98%, 99%, or 99.5% identical to amino acid residues 3-229 of SEQ ID NO: 4, wherein the cytidine deaminase comprises one or more mutations selected from the group consisting of E4X1, H122X2, D124X3, R154X4, A165X5, P201X6, and F205X7 relative to SEQ ID NO: 4, or a corresponding mutation(s) in another cytidine deaminase, wherein X, is any amino acid other than E, X2 is any amino acid other than H, X3 is any amino acid other than D, X4 is any amino acid other than R, X5 is any amino acid other than A, X6 is any amino acid other than P, and X7 is any amino acid other than F.

5. claims: 65-88, 94-107, 109, 111-115, 122-140(completely); 116-121, 142(partially)

A fusion protein comprising:(i) a nucleic acid programmable DNA binding protein (napDNAbp);(ii) the cytidine deaminase of any one of claims 1-64; and(iii) a uracil glycosylase inhibitor domain (UGI).

6. claims: 143-147, 164(completely); 165(partially)

A fusion protein comprising an RNA polymerase fused to a degron tag, optionally wherein the degron tag comprises an amino acid sequence that is at least 80%, 85%, 90%, 95%, 98%, 99%, or 100% identical to AANDENYNYALAA (SEQ ID NO: 134).

7. claims: 148-153(completely); 165-176(partially)

A fusion protein comprising a cytidine deaminase fused to an N-intein, optionally wherein the fusion comprises an amino acid sequence that is at least 80%, 85%, 90%, 95%, 98%, 99%, or 100% identical to Seq-ID-N°136

8. claims: 154-158(completely); 165-176(partially)

. A fusion protein comprising a Cas9 domain, a uracil DNA glycosylase inhibitor (UGI) domain, and a C-intein, optionally wherein the fusion comprises an amino acid sequence that is at least 80%, 85%, 90%, 95%, 98%, 99%, or 100% identical to Seq-ID-N°138

9. claims: 159-163

International Application No. PCT/ US2018/ 044242

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

A fusion protein comprising a uracil-DNA binding protein and
a transcriptional activator.

フロントページの続き

(51) Int.Cl.	F I	テーマコード(参考)
C 1 2 N 1/19 (2006.01)	C 1 2 N 1/19	4 H 0 4 5
C 1 2 N 1/21 (2006.01)	C 1 2 N 1/21	
C 1 2 N 5/10 (2006.01)	C 1 2 N 5/10	
A 6 1 K 38/16 (2006.01)	A 6 1 K 38/16	
A 6 1 K 47/64 (2017.01)	A 6 1 K 47/64	
A 6 1 K 48/00 (2006.01)	A 6 1 K 48/00	
A 6 1 P 43/00 (2006.01)	A 6 1 P 43/00	1 1 1
A 6 1 K 35/76 (2015.01)	A 6 1 K 35/76	
C 1 2 N 15/09 (2006.01)	C 1 2 N 15/09	1 1 0

(81) 指定国・地域 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, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, 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

(72) 発明者 ウィルソン, クリストファー, ジェラード
 アメリカ合衆国 マサチューセッツ州 0 2 4 5 1、ウォルサム、メイン ストリート 6 9 6、
 アpartment 3 1 1

(72) 発明者 リウ, デイビッド, アール.
 アメリカ合衆国 マサチューセッツ州 0 2 4 2 0、レキシントン、ホイットマン サークル 3

Fターム(参考) 4B050 CC04 CC05 CC10 DD11 LL01
 4B065 AA93X AA99Y AB01 AC14 BA02 CA24 CA27 CA44
 4C076 AA95 CC41 EE41
 4C084 AA02 AA03 AA13 BA02 BA22 BA44 MA02 NA13 NA14 ZC41
 4C087 AA01 AA02 BC83 CA12 MA02 NA13 NA14 ZC51
 4H045 AA10 AA30 BA10 BA41 BA50 CA40 DA89 EA20 FA74