

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7642531号
(P7642531)

(45)発行日 令和7年3月10日(2025.3.10)

(24)登録日 令和7年2月28日(2025.2.28)

(51)国際特許分類		F I	
C 1 2 N	15/09 (2006.01)	C 1 2 N	15/09 1 1 0
C 1 2 N	15/12 (2006.01)	C 1 2 N	15/12 Z N A
C 1 2 N	5/10 (2006.01)	C 1 2 N	5/10
C 1 2 N	5/078(2010.01)	C 1 2 N	5/078
C 1 2 N	5/0789(2010.01)	C 1 2 N	5/0789
請求項の数 13 (全230頁) 最終頁に続く			
(21)出願番号	特願2021-513763(P2021-513763)	(73)特許権者	520439645
(86)(22)出願日	令和1年5月11日(2019.5.11)		ビーム セラピューティクス インク.
(65)公表番号	特表2021-523737(P2021-523737 A)		アメリカ合衆国 0 2 1 4 2 マサチュー
(43)公表日	令和3年9月9日(2021.9.9)		セッツ州 ケンブリッジ メイン ストリ
(86)国際出願番号	PCT/US2019/031897	(74)代理人	100083806
(87)国際公開番号	WO2019/217942		弁理士 三好 秀和
(87)国際公開日	令和1年11月14日(2019.11.14)	(74)代理人	100111235
審査請求日	令和4年5月11日(2022.5.11)		弁理士 原 裕子
(31)優先権主張番号	62/670,521	(74)代理人	100195257
(32)優先日	平成30年5月11日(2018.5.11)		弁理士 大淵 一志
(33)優先権主張国・地域又は機関	米国(US)	(74)代理人	100201189
(31)優先権主張番号	62/670,539		弁理士 安原 二良
(32)優先日	平成30年5月11日(2018.5.11)	(72)発明者	ブライソン、 デイビッド
最終頁に続く			アメリカ合衆国 0 2 1 3 9 マサチュー
		最終頁に続く	

(54)【発明の名称】 プログラム可能塩基エディターシステムを用いて病原性アミノ酸を置換する方法

(57)【特許請求の範囲】
【請求項1】

鎌状赤血球症に関連する一塩基多型 (SNP) を含むβ-グロビン (HBB) ポリヌクレオチドを編集するインビトロまたはエクスピボの方法であって、前記SNPは前記HBBポリヌクレオチドにコードされたアミノ酸位置7にバリンを有するHBBポリペプチドの発現をもたらすものであり、前記方法は、前記HBBポリヌクレオチドを、1以上のガイドポリヌクレオチドと複合体を形成する塩基エディターと接触させることを含み、前記塩基エディターは、核酸配列5'-NGC-3'を含むプロトスペーサー隣接モチーフ (PAM) 特異性を有するSp Cas9の核酸プログラム可能なDNA結合タンパク質 (napDNAbp) ドメインおよびアデノシンデアミナーゼドメインを含み、前記1以上のガイドポリヌクレオチドが前記塩基エディターをターゲティングして前記SNPのA・TからG・Cへの改変をもたらし、前記SNPのA・TからG・Cへの改変が、前記HBBポリヌクレオチドにコードされたHBBポリペプチドのアミノ酸位置7におけるバリンをアラニンに変化させ、それにより前記バリンの代わりにアラニンを有するHBBポリペプチドの発現をもたらし、ここで、前記アミノ酸位置7にバリンを有するHBBポリペプチドは以下のアミノ酸配列：

1 mvhlt

[v]eks avtalwgkvn vdevggealg rllvvypwtq rffesfgdls tpdavmgnpk

61 vkahgkklvg afsdglahld nlkgtfatls elhcdklhvd penfrllgnv lvcvlahhfg

121 keftppvqaa yqkvvagvan alahkyh

を含み、[v]は前記アミノ酸位置7におけるバリンを示す、方法。

【請求項 2】

前記napDNAbpドメインが、ヌクレアーゼ不活性バリエーションもしくはニッカーゼバリエーションである、請求項 1 に記載の方法。

【請求項 3】

前記アデノシンデアミナーゼドメインが、以下のTadA*7.10アミノ酸配列MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNRRVIGEGWNRAIGLHDPTAHAEIMALRQGG LVMQNYRLIDATLYVTTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPG MNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTDまたはN末端メチオニンを欠く前記TadA*7.10アミノ酸配列を含む、請求項 1 または 2 に記載の方法。

【請求項 4】

核酸配列5'-NGC-3'を含むプロトスペーサー隣接モチーフ (PAM) 特異性を有するSpCas9の核酸プログラム可能なDNA結合タンパク質 (napDNAbp) ドメインおよびアデノシンデアミナーゼドメインを含む、塩基エディター、または前記塩基エディターをコードするポリヌクレオチド；および

前記塩基エディターをターゲティングしてβ-グロビン (HBB) ポリヌクレオチド中のSNPのA・TからG・Cへの改変をもたらす1つ以上のガイドポリヌクレオチドであって、前記SNPは鎌状赤血球症に関連するものであり、前記SNPは前記HBBポリヌクレオチドにコードされたアミノ酸位置7にバリンを有するHBBポリペプチドの発現をもたらすものである、ガイドポリヌクレオチド

を含む、細胞またはその前駆細胞であって、

前記SNPのA・TからG・Cへの改変が、前記HBBポリヌクレオチドにコードされたHBBポリペプチドのアミノ酸位置7におけるバリンをアラニンに変化させ、それによって前記バリンの代わりにアラニンを有するHBBポリペプチドの発現をもたらす、ここで、前記アミノ酸位置7にバリンを有するHBBポリペプチドは以下のアミノ酸配列：

1 mvhlt[p]v]eks avtalwgvkn vdevggealg rllvvypwtq rffesfgdls tpdavmgnpk
61 vkahgkklvg afsdglahld nlkgtfatls elhcdklhvd penfrllgnv lvcvlahhfg
121 keftppvqaa yqkvvagvan alahkyh

を含み、[v]は前記アミノ酸位置7におけるバリンを示す、

細胞またはその前駆細胞。

【請求項 5】

前記細胞が、造血幹細胞、骨髓系共通前駆細胞、前赤芽球、赤芽球、網状赤血球、または赤血球であり、前記細胞は鎌状赤血球症を有する対象由来のものである、請求項 4 に記載の細胞。

【請求項 6】

前記アデノシンデアミナーゼドメインは、以下のTadA*7.10アミノ酸配列MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNRRVIGEGWNRAIGLHDPTAHAEIMALRQGG LVMQNYRLIDATLYVTTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPG MNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTDまたはN末端メチオニンを欠く前記TadA*7.10アミノ酸配列を含む、請求項 4 または 5 に記載の細胞。

【請求項 7】

対象における鎌状赤血球症を治療する方法における使用のための請求項 4～6 のいずれか一項に記載の細胞を含む組成物であって、前記方法は前記対象に前記細胞を投与することを含む、組成物。

【請求項 8】

対象における鎌状赤血球症を治療する方法における使用のための、塩基エディターまたはそれをコードするポリヌクレオチドを含む組成物であって、前記方法は、前記対象に以下のもの：

核酸配列5'-NGC-3'を含むプロトスペーサー隣接モチーフ (PAM) 特異性を有するSpCas9の核酸プログラム可能なDNA結合タンパク質 (napDNAbp) ドメインおよびアデノシンデアミナーゼドメインを含む、前記塩基エディター、または前記塩基エディターをコー

10

20

30

40

50

ドするポリヌクレオチド；および

前記塩基エディターをターゲティングしてβ-グロビン（HBB）ポリヌクレオチド中のSNPのA・TからG・Cへの改変をもたらす、1つ以上のガイドポリヌクレオチドであって、前記SNPは鎌状赤血球症に関連するものであり、前記SNPは前記HBBポリヌクレオチドにコードされたアミノ酸位置7にバリンを有するHBBポリペプチドの発現をもたらす、ガイドポリヌクレオチド

を投与することを含み、

前記SNPのA・TからG・Cへの改変が、前記HBBポリヌクレオチドにコードされたHBBポリペプチドのアミノ酸位置7におけるバリンをアラニンに変化させ、それにより前記バリンの代わりにアラニンを有するHBBポリペプチドの発現をもたらす、ここで、前記アミノ酸位置7にバリンを有するHBBポリペプチドは以下のアミノ酸配列：

1 mvhlt[p]v]eks avtalwgkvn vdevggealg rllvvypwtq rffesfgdls tpdavmgnpk
61 vkahgkkvlg afsdglahld nlkgtfatls elhcdklhvd penfrllgnv lvcvlahhfg
121 keftppvqaa yqkvvagvan alahkyh

を含み、[v]は前記アミノ酸位置7におけるバリンを示す、組成物。

【請求項9】

前記対象は哺乳動物またはヒトであり；

前記方法は、前記塩基エディター、または前記塩基エディターをコードするポリヌクレオチド、および前記1つ以上のガイドポリヌクレオチドを、前記対象の細胞に送達することを含み；

前記細胞が、造血幹細胞、骨髓系共通前駆細胞、前赤芽球、または赤芽球である、

請求項8に記載の組成物。

【請求項10】

前記アデノシンデアミナーゼドメインが、以下のTadA*7.10アミノ酸配列MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNRRVIGEGWNRAIGLHDPTAHAEIMALRQGG LVMQNYRLIDATLYVTTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPG MNHRVEITEGILADECAALLCYFFRMQRQVFNAQKKAQSSTDまたはN末端メチオニンを欠く前記TadA*7.10アミノ酸配列を含む、請求項8または9に記載の組成物。

【請求項11】

赤血球またはその前駆細胞を産生するインビトロまたはエクスピボの方法であって：

鎌状赤血球症に関連するSNPを含む赤血球前駆細胞に、

核酸配列5'-NGC-3'を含むプロトスペーサー隣接モチーフ（PAM）特異性を有するSpCas9の核酸プログラム可能なDNA結合タンパク質（napDNAbp）ドメインおよびアデノシンデアミナーゼドメインを含む、塩基エディター、または前記塩基エディターをコードするポリヌクレオチドと；

前記塩基エディターをターゲティングしてβ-グロビン（HBB）ポリヌクレオチド中のSNPのA・TからG・Cへの改変をもたらす1以上のガイドポリヌクレオチドであって、前記SNPは鎌状赤血球症に関連するものであり、前記SNPは前記HBBポリヌクレオチドにコードされたアミノ酸位置7にバリンを有するHBBポリペプチドの発現をもたらすものである、ガイドポリヌクレオチドと

を導入し、

前記SNPのA・TからG・Cへの改変が、前記HBBポリヌクレオチドにコードされたHBBポリペプチドのアミノ酸位置7におけるバリンをアラニンに変化させ、それにより前記バリンの代わりにアラニンを有するHBBポリペプチドの発現をもたらすことを含み、ここで、前記アミノ酸位置7にバリンを有するHBBポリペプチドは以下のアミノ酸配列：

1 mvhlt[p]v]eks avtalwgkvn vdevggealg rllvvypwtq rffesfgdls tpdavmgnpk
61 vkahgkkvlg afsdglahld nlkgtfatls elhcdklhvd penfrllgnv lvcvlahhfg
121 keftppvqaa yqkvvagvan alahkyh

を含み、[v]は前記アミノ酸位置7におけるバリンを示す、方法。

【請求項12】

10

20

30

40

50

前記赤血球前駆細胞を、造血幹細胞、骨髓系共通前駆細胞、前赤芽球、赤芽球、網状赤血球、または赤血球のうちの一つ以上に分化させることをさらに含み；前記赤血球前駆細胞は、鎌状赤血球症を有する対象から得られるものである、請求項 1 1 に記載の方法。

【請求項 1 3】

前記アデノシンデアミナーゼドメインは、以下の TadA*7.10 アミノ酸配列 MSEVEFSHE YWMRHALTLAKRARDEREVPVGAVLV LNNRVIGEGWNRAIGLHDPTAHAEIMALRQGG LVMQNYRLIDATLYVTTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPG MNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTD または N 末端メチオニンを欠く前記 TadA*7.10 アミノ酸配列を含む、請求項 1 1 または 1 2 に記載の方法。

【発明の詳細な説明】

【技術分野】

【0 0 0 1】

関連出願

本出願は、2018年5月11日出願の米国仮出願第62/670,521号、2018年5月11日出願の米国仮出願第62/670,539号、および2018年12月17日出願の米国仮出願第62/780,890号の利益を主張するものであり、それぞれの内容全体が参照により本明細書に組み込まれる。

参照による組み込み

本明細書に記載されているすべての刊行物、特許、および特許出願は、それぞれの個々の刊行物、特許、または特許出願が具体的かつ個別に参照により組み込まれるように示されているのと同じ程度に、参照により本明細書に組み込まれる。別段の表示がない限り、本明細書に記載されている刊行物、特許および特許出願は、その全体が参照により本明細書に組み込まれる。

【0 0 0 2】

ほとんどの既知の遺伝性疾患では、疾患の根底にある原因を研究または対処するために、遺伝子の確率的破壊ではなく、標的遺伝子座における点突然変異の補正が必要となる。

【背景技術】

【0 0 0 3】

Clustered regularly interspaced short palindromic repeat (CRISPR) 系を用いた現在のゲノム編集技術は、遺伝子補正の第一段階として標的遺伝子座で二本鎖DNA切断を導入する。二本鎖DNA切断にตอบสนองして、細胞のDNA修復プロセスは、ほとんどの場合、DNA切断部位での非相同末端結合によるランダムな挿入または欠失（インデル）をもたらす。ほとんどの遺伝性疾患は点突然変異から生じるが、点突然変異補正に対する現在のアプローチは非効率的であり、典型的には、dsDNA切断に対する細胞応答から生じる標的遺伝子座でのランダムな挿入および欠失（インデル）を多く誘導する。したがって、より効率的で、確率的な挿入もしくは欠失（インデル）または転座などの望ましくない産物がはるかに少ない、改良された形態のゲノム編集が必要とされている。

【発明の概要】

【0 0 0 4】

本明細書では、対象における遺伝的障害を治療する方法が提供され、この方法は、塩基エディター、または塩基エディターをコードするポリヌクレオチドを対象に投与すること（ここで塩基エディターは、ポリヌクレオチドのプログラミングが可能なヌクレオチド結合ドメインおよびデアミナーゼドメインを含む）と；ガイドポリヌクレオチドを対象に投与すること（ここで、ガイドポリヌクレオチドは、塩基エディターを対象の標的ヌクレオチド配列に対してターゲティング（標的指向化）する）と；標的ヌクレオチド配列への塩基エディターのターゲティングに際して、核酸塩基を脱アミノ化することにより標的ヌクレオチド配列の核酸塩基を編集し、それにより核酸塩基を別の核酸塩基に変化させることにより遺伝的障害を治療することを含み、ここで、前記遺伝的障害はタンパク質中の病原性アミノ酸に起因し、別の核酸塩基が、前記病原性アミノ酸を、前記タンパク質の野生型アミノ酸とは異なる良性（benign）アミノ酸で置換する。

【0005】

本明細書では、対象における遺伝的障害を治療するための細胞、組織、または器官を産生する方法が提供され、該方法は、該細胞、組織、または器官を、塩基エディター、または該塩基エディターをコードするポリヌクレオチドと接触させること（ここで塩基エディターは、ポリヌクレオチドのプログラミングが可能なヌクレオチド結合ドメインおよびデアミナーゼドメインを含む）と；該細胞、組織、または器官を、ガイドポリヌクレオチドに接触させること（ここでガイドポリヌクレオチドは、該細胞、組織、または器官の標的ヌクレオチド配列に対して塩基エディターをターゲティングする）と；標的ヌクレオチド配列への塩基エディターのターゲティングに際して、核酸塩基を脱アミノ化することにより標的ヌクレオチド配列の核酸塩基を編集し、それにより、核酸塩基を別の核酸塩基に変化させることにより遺伝的障害を治療するための細胞、組織、または器官を産生することを含み、ここで、前記遺伝的障害はタンパク質中の病原性アミノ酸に起因し、別の核酸塩基が、前記病原性アミノ酸を、前記タンパク質の野生型アミノ酸とは異なる良性アミノ酸で置換する。いくつかの態様において、本方法は、細胞、組織、または器官を対象に投与することをさらに含む。ある態様において、細胞、組織、または器官は、対象にとって自家系（autologous）である。ある態様において、細胞、組織、または器官は、対象にとって同種異系（allogeneic）である。ある態様において、細胞、組織、または器官は、対象にとって異種系（xenogeneic）である。

10

【0006】

ある態様において、核酸塩基は、遺伝的障害の原因である遺伝子に位置する。いくつかの実施形態において、編集は、遺伝子内に位置する複数の核酸塩基を編集することを含み、複数の核酸塩基は遺伝的障害の原因ではない。いくつかの態様において、編集することは、少なくとも1つの他の遺伝子に位置する1つ以上のさらなる核酸塩基を編集することをさらに含む。ある態様において、遺伝子および少なくとも1つの他の遺伝子は、タンパク質の1つ以上のサブユニットをコードする。

20

【0007】

ある態様において、編集される核酸塩基は、表3Aまたは3Bに列挙された遺伝子にあり、編集は、表3Aまたは3Bに示された遺伝子によってコードされるタンパク質におけるアミノ酸変化をもたらす。ある態様において、遺伝的障害は、ACADM欠損症、鎌状赤血球症（SCD）、ヘモグロビン病、βサラセミア、ペンドレッド症候群、常染色体優性パーキンソン病、またはα-1アンチトリプシン欠損症（A1AD）である。

30

【0008】

一態様では、本発明は、プログラム可能な核酸塩基エディターを用いて病原性アミノ酸を置換するための組成物および方法の特徴とする。特に、β-グロビタンパク質の鎌状赤血球症バリエーション（Sickle HbS；E6V）の第六アミノ酸のコドンにおいてチミジン（T）をシチジン（C）核酸塩基に塩基編集し、それによってバリンをアラニンに置換する（E6A）ための組成物および方法を提供する。鎌状赤血球ヘモグロビンSの6位でバリンがアラニンに置換されると、鎌状赤血球症の表現型を欠くβグロビタンパク質バリエーションが生じる（例えば、正常β-グロビン蛋白質（HbA；E6）の性質を有し、病原性変異体HbS等の場合のように重合する可能性がない）。したがって、本発明の組成物および方法は、鎌状赤血球症の治療に有用である。一実施形態では、編集される核酸塩基は、ベータ（β）-グロビンをコードするHBB遺伝子内にあり、塩基の編集は、HBB遺伝子によってコードされるβ-グロビン（HBB）タンパク質のアミノ酸6において、バリン（Val）からアラニン（Ala）へのアミノ酸変化をもたらす（β6 Val→Ala）。ある態様において、遺伝的障害は、鎌状赤血球症またはヘモグロビン病である。ある態様において、塩基編集は、ヘモグロビンのβサブユニットにおいてE6V→E6Aアミノ酸変化をもたらす。

40

【0009】

別の態様において、本発明は、鎌状赤血球症に関連する一塩基多型（SNP）を含むHBBポリヌクレオチドを編集する方法を提供し、この方法は、HBBポリヌクレオチドを、一以上のガイドポリヌクレオチドと複合体をなす塩基エディターと接触させることを含み、塩

50

基エディターは、ポリヌクレオチドプログラミング可能なDNA結合ドメインおよびアデノシンデアミナーゼドメインを含み、一以上のガイドポリヌクレオチドは、鎌状赤血球症に関連するSNPのA・TからG・Cへの改変をもたらすために塩基エディターをターゲティングする。

【0010】

別の態様では、本発明は、塩基エディター、塩基エディターをコードするポリヌクレオチド（塩基エディターはプログラム可能なDNA結合ドメインおよびアデノシンデアミナーゼドメインを含む）と、鎌状赤血球症に関連するSNPのA・TからG・Cへの改変をもたらすために塩基エディターをターゲティングする1つ以上のガイドポリヌクレオチドとを、細胞またはその前駆体に導入することによって作製される細胞を提供する。

10

【0011】

別の態様において、本発明は、本明細書に記載されるいずれかの態様による細胞を、必要とする対象に投与することを含む、対象における鎌状赤血球症を治療する方法を提供する。

【0012】

別の態様において、本発明は、本明細書に記載された任意の態様による細胞から増殖または拡張された単離細胞または細胞集団を提供する。

【0013】

別の態様において、本発明は、対象における鎌状赤血球症を治療する方法を提供し、その方法は、塩基エディター、または塩基エディターをコードするポリヌクレオチドと（塩基エディターはプログラム可能なDNA結合ドメインおよびアデノシンデアミナーゼドメインを含む）；鎌状赤血球症に関連するSNPのA・TからG・Cへの改変をもたらすために塩基エディターをターゲティングする1つ以上のガイドポリヌクレオチドとを、必要とする対象に投与することを含む。

20

【0014】

もう1つの態様において、本発明は、赤血球（red blood cell、erythrocyte）またはその前駆細胞を産生する方法を提供し、該方法は、(a) 塩基エディター、または塩基エディターをコードするポリヌクレオチドと（塩基エディターはポリヌクレオチドプログラミング可能なヌクレオチド結合ドメインおよびアデノシンデアミナーゼドメインを含む）、鎌状赤血球症に関連するSNPのA・TからG・Cへの改変をもたらすために塩基エディターをターゲティングする1つ以上のガイドポリヌクレオチドとを、鎌状赤血球症に関連するSNPを含む赤血球前駆細胞に導入することと；(b) 赤血球前駆細胞を赤血球へと分化させることとを含む。

30

【0015】

別の局面において、本発明は、(i) *Streptococcus thermophilus* Cas 9 (St1Cas9) を含むポリヌクレオチドプログラミング可能なDNA結合ドメイン、および (ii) アデノシンデアミナーゼドメインを含む、塩基エディターを提供する。

【0016】

別の局面において、本発明は、CUUCUCCACAGGAGUCAGAU; ACUUCUCCACAGGAGUCAGAU; および GACUUCUCCACAGGAGUCAGAUから選択される核酸配列を含むガイドRNA (gRNA) を提供する。

40

【0017】

別の態様において、本発明は、(i) 改変された *Staphylococcus aureus* Cas 9 (SaCas9) を含むポリヌクレオチドプログラミング可能なDNA結合ドメイン、および (ii) アデノシンデアミナーゼドメインを含む、塩基エディターを提供する。

【0018】

別の局面において、本発明は、UCCACAGGAGUCAGAU GCAC および UCCACAGGAGUCAGAU GCACから選択される核酸配列を含むガイドRNA (gRNA) を提供する。

【0019】

別の局面において、本発明は、UUCUCCACAGGAGUCAGA; CUUCUCCACAGGAGUCA

50

GA; ACUUCUCCACAGGAGUCAGA; GACUUCUCCACAGGAGUCAGA; および AGACUUCUCCACAGGAGUCAGAから選択される核酸配列を含むガイドRNA (gRNA) を提供する。

【0020】

一実施形態によると、塩基編集の結果、SERPINA 1遺伝子がコードする α -1アンチトリプシンタンパク質のE342K> E342Gアミノ酸変化がもたらされる。一実施形態によると、遺伝的障害は、中鎖アシルCoAデヒドロゲナーゼ (ACADM) 欠損症である。一実施形態によると、塩基編集の結果、中鎖アシルCoAデヒドロゲナーゼ (ACADM) 遺伝子がコードするタンパク質においてK329E> K329Gアミノ酸変化が生じる。一実施形態によると、遺伝的障害はヘモグロビン病である。一実施形態では、塩基編集は、HBB遺伝子によってコードされるヘモグロビンの β サブユニットにおいてE26K> E26Gアミノ酸変化をもたらす。いくつかの態様において、遺伝的障害は、ペンドレッド症候群である。いくつかの態様において、塩基編集は、SLC26A4 (PDS遺伝子によってコードされる、Solute Carrier Family 26 Member 4 (PDS) タンパク質) におけるT416P> T416Fアミノ酸変化をもたらす。ある態様において、遺伝的障害は常染色体優性パーキンソン病である。ある態様において、編集は、SNCA遺伝子によってコードされるアルファシヌクレイン (SNCA) タンパク質におけるA30P> A30Lアミノ酸変化をもたらす。

10

【0021】

本明細書に記載する任意の態様の種々の実施形態において、鎌状赤血球症に関連するSNPにおけるA・TからG・Cへの改変は、HBBポリペプチドにおいてバリンをアラニンに変化させる。様々な実施形態において、鎌状赤血球症に関連するSNPは、アミノ酸位置6にバリンを有するHBBポリペプチドの発現をもたらす。様々な実施形態において、鎌状赤血球症に関連するSNPは、グルタミン酸をバリンで置換する。

20

【0022】

本明細書に記載するあらゆる態様の様々な実施形態において、接触は、細胞、真核細胞、哺乳動物細胞、またはヒト細胞内におけるものである。様々な実施形態において、対象は哺乳動物またはヒトである。様々な実施形態において、細胞はin vivoまたはex vivoである。様々な実施形態において、細胞またはその前駆体は、胚性幹細胞、人工多能性幹細胞造血幹細胞、骨髓系共通前駆細胞、前赤芽球、赤芽球、網状赤血球、または赤血球である。様々な実施形態において、造血幹細胞はCD34⁺細胞である。様々な態様において、細胞は、鎌状赤血球症を有する対象からのものである。様々な実施形態において、細胞は対象にとって自家性 (autologous) である。様々な実施形態において、細胞は、対象にとって同種異系 (allogeneic) または異種 (xenogeneic) である。本明細書に記載される任意の態様の様々な実施形態において、方法は、塩基エディタ、または塩基エディタをコードするポリヌクレオチド、および1つ以上のガイドポリヌクレオチドを、対象の細胞に送達することを含む。

30

【0023】

本明細書に記載されるあらゆる態様の様々な実施形態において、ポリヌクレオチドプログラミング可能DNA結合ドメインは、改変されたStaphylococcus aureus Cas 9 (SaCas9)、Streptococcus thermophilus 1 Cas 9 (St1Cas9)、改変されたStreptococcus pyogenes Cas 9 (SpCas9)、またはそれらのバリエーションである。種々の実施形態において、ポリヌクレオチドプログラミング可能DNA結合ドメインは、改変されたプロトスペーサー隣接モチーフ (PAM) 特異性を有する改変されたSaCas9を含む。様々な実施形態において、改変されたPAMは、核酸配列5'-NNNRRT-3'を含む。様々な実施形態において、改変されたSaCas9は、アミノ酸置換E782K、N968K、およびR1015Hまたはそれらに対応するアミノ酸置換を含む。

40

【0024】

様々な実施形態において、ポリヌクレオチドプログラミング可能なDNA結合ドメインは、改変されたプロトスペーサー隣接モチーフ (PAM) 特異性を有するSpCas9のバリエーションを含む。様々な実施形態において、改変されたPAMは、核酸配列5'-NGC-3'を含む。

【0025】

50

様々な実施形態において、改変されたSpCas9は、アミノ酸置換D1135M、S1136Q、G1218K、E1219F、A1322R、D1332A、R1335E、およびT1337Rまたはそれらに対応するアミノ酸置換を含む。様々な実施形態において、ポリヌクレオチドプログラミング可能なDNA結合ドメインは、ヌクレアーゼ不活性化またはニックアーゼバリエーションである。様々な実施形態において、ニックアーゼバリエーションは、アミノ酸置換D10Aまたはそれに対応するアミノ酸置換を含む。

【0026】

本明細書に記載する任意の態様の様々な実施形態では、塩基エディターは、ジンクフィンガードメインをさらに含む。様々な実施形態において、ジンクフィンガードメインは、認識ヘリックス配列RNEHLEV、QSTTLKR、およびRTEHLAR、または認識ヘリックス配列RGEHLRQ、QSGTLKR、およびRNDKLVPを含む。様々な、ジンクフィンガードメインは、zf1raまたはzf1rbの1つ以上である。

【0027】

本願明細書に記載される任意の態様の様々な実施形態において、アデノシンデアミナーゼドメインは、デオキシリボ核酸 (DNA) 中のアデニンを脱アミノ化することができる。様々な実施形態において、アデノシンデアミナーゼは、自然界には存在しない改変アデノシンデアミナーゼである。様々な実施形態において、アデノシンデアミナーゼは、TadAデアミナーゼである。様々な実施形態において、TadAデアミナーゼはTadA*7.10である。

【0028】

本願明細書に記載される任意の態様の様々な実施形態において、1つ以上のガイドRNAは、CRISPR RNA (crRNA) およびトランスコード低分子RNA (tracrRNA) を含み、ここで、crRNAは、鎌状赤血球症に関連するSNPを含むHBB核酸配列に相補的な核酸配列を含む。様々な実施形態において、塩基エディターは、鎌状赤血球症に関連するSNPを含むHBB核酸配列に相補的な核酸配列を含む単一ガイドRNA (sgRNA) と複合体をなしている。

【0029】

本明細書に記載される任意の態様の様々な実施形態において、St1Cas9は、以下のアミノ酸配列を含む：

SDLVLGLAIGIGSVGVGILNKVTGEIIHKNSRIFPAAQAENNLVRRRTNRQGRRLARRKKHR
RVRLNRLFEEGLITDFTKISINLNPYQLRVKGLTDELSNEELFIALKNMVKHRGISYLD
DASDDGNSSVGDYAQIVKENSQKLETKTPGQIQRLERYQTYGQLRGDFTVEKDGKKHRLINVF
PTSAYRSEALRILQTQQEFNPQITDEFINRYLEILTGRKRYHGPNGNEKSRDYGRTSG
ETLDNIFGILIGKCTFYFDEFRAAKASYTAQEFNLLNDLNNLTVPTETKKLSKEQKNQIIN
YVKNEKAMGPAKLFKYIAKLLSCDVADIKGYRIDKSGKAEIHTFEAYRKMKTLETLDIEQ
MDRETLDKLAYVLTNTEREGIQEALHEFADGSFSQKQVDELVQFRKANSSIFGKGWHN
FSVKLMMELIPELYETSEEQMTILTRLGKQKTSSSNKTKYIDEKLLTEEIYNPVAKSVR
QAIKIVNAAIKEYGDFDNIVIEMARETNEDDEKKAIQKIQKANKDEKDAAMLKAANQYNG
KAELPHSVFHHGKQLATKIRLWHQQGERCLYTGKTISIHDLINNSNQFEVDHILPLSITFD
DSLANKVLVYATANQEKQRTPYQALDSMDDAWSFRELKAFVRESKTLNKKKEYLLTE
EDISKFDVRKKFIERNLVDTLVYASRVVLNALQEHFRAHKIDTKVSVVRGQFTSQLRRHWG
IEKTRDTHHHAVDALIAASSQLNLWKKQKNTLVSYSEDQLLDIETGELISDDEYKESVF
KAPYQHFDVDTLKSKEFEDSILFSYQVDSKFNRKISDATIYATRQAKVGKDKADETYVLGKI
KDIYTQDGYDAFMKIYKKDKSKFLMYRHDPQTFEKVIEPILENYPNKQINDKGKEVPCNP
FLKYKEEHGYIRKYSKKGNGPEIKSLKYYDSKLGNHIDITPKDSNNKVVLQSVSPWRADV
YFNKTTGKYEILGLKYADLQFDKGTGTYKISQEKYNDIKKKEGVDSSEFKFTLYKNDLLL
VKDTETKEQQLFRFLSRTMPKQKHVELKPYDKQKFEGGEALIKVLGNVANSQGCKKGL
GKSNIYKVRTDVLGNQHIIKNEGDKPKLDF。

【0030】

種々の実施形態において、塩基エディターは、ポリヌクレオチドプログラミング可能なDNA結合ドメインとアデノシンデアミナーゼのドメインとの間にリンカーを含む。種々の実施形態において、リンカーは、アミノ酸配列：SGGSSGGSSGSETPGTSESATPESを含む

10

20

30

40

50

。種々の実施形態において、塩基エディターは、1以上の核局在化シグナルを含む。種々の態様において、塩基エディターは、以下のアミノ酸配列を含む：

MPKKKRKVSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGR
HDPTAHAEIMALRQGGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAK
TGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAKKKAQSSTDSG
GSSGGSSGSETPGTSESATPESSGGSSGSSEVEFSHEYWMRHALTLAKRARDEREVPVGA
VLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGGLVMQNYRLIDATLYVTLEPCVMCA
GAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFR
MPRQVFNAQKKAQSSTDSGGSSGGSSGSETPGTSESATPESDLVLGLAIGISVGVGILNK
VTGEIIHKNSRIFPAAQAENNLVRRNTNRQGRRLARRKKHRRVRLNRLFEESSLITDFTKISI
NLNPYQLRVKGLTDELSNEELFIALKNMVKHGRISYLDASDDGNSSVGDYAQIVKENS
KQLETKTPGQIQQLERYQTYGQLRGDFTVEKDGGKHRLINVFPTSAYRSEALRILQTQQEFN
PQITDEFINRYLEILTGKRKYHGPNGEKSRTDYGRYRTSGETLDNIFGILIGKCTFYPDEF
RAAKASYTAQEFNLLNDLNNLTVPETETKLSKEQKNQIINYVKNEKAMGPAKLFKYIAKL
LSCDVADIKGYRIDKSGKAEIHTFEAYRKMKTLTLDIEQMDRETLDKLAYVLTNTEREG
IQEALHEFADGSFSQKQVDELVQFRKANSSIFGKGWHNFSVKLMMELIPELYETSEEQM
TILTRLGKQKTTSSSNKTKYIDEKLLTEEIYNPVAKSVRQAIKIVNAAIKEYGDFDNIVIE
MARETNEDDEKKAIQIKANKDEKDAAMLKAANQYNGKAELPHSVFHGHKQLATKIRL
WHQQGERCLYTGTISIHDLINNSNQFEVDHILPLSITFDDSLANKVLVYATANQEKQR
TPYQALDSMDDAWSFRELKAFVRESKTLNKKKEYLLTEEDISKFDVRKKFIERNLVDTL
YASRVVLNALQEHFRAHKIDTKVSVVRGQFTSQLRRHWGIEKTRDTYHHHAVDALIIAAS
SQLNLWKKQKNTLVSYSEDQLLDIETGELISDDEYKESVFKAPYQHFDVTLKSKEFEDSIL
FSYQVDSKFNRKISDATIYATRQAKVGKDKADETYVLGKIKDIYTQDGYDAFMKIYKKDK
SKFLMYRHDPQTFEKVIENYPNKQINDKGKEVPCNPFLKYKEEHGYIRKYSKKGNG
PEIKSLKYYDSKLGHNHIDITPKDSNNKVVLQSVSPWRADVFNKTTGKYEILGLKYADLQ
FDKGTGTYKISQEKYNDIKKKEGVDSSEFKFTLYKNDLLLVKDTETKEQQLFRFLSRTMP
KQKHVELKPYDKQKFEGGEALIKVLGNVANSGQCKKGLGKSNISYKVRTDVLGNQHII
KNEGDKPKLDFPKKKRKVEGADKRTADGSEFESPKKKRKV。

【0031】

本願明細書に記載される任意の態様の様々な実施形態において、ガイドRNAは、以下の核酸配列をさらに含む：

GUUUUUGUACUCUCAAGAUUUUAAGUACUGUACAACGAAACUACACAGUUACUUA
AUCUUGCAGAAGCUACAAAGAUAAAGGCUUCAUGCCGAAAUCAACACCCUGUCAUUU
AUGGCAGGGUG。

【0032】

様々な実施形態において、ガイドRNAは、

CUUCUCCACAGGAGUCAGAUGUUUUUGUACUCUCAAGAUUUUAAGUACUGUACAACG
AAACUACACAGUUACUUAUUUUCUUGCAGAAGCUACAAAGAUAAAGGCUUCAUGCCG
AAUCAACACCCUGUCAUUUUUAUGGCAGGGUG;

ACUUCUCCACAGGAGUCAGAUGUUUUUGUACUCUCAAGAUUUUAAGUACUGUACAAC
GAAACUACACAGUUACUUAUUUUCUUGCAGAAGCUACAAAGAUAAAGGCUUCAUGCCG
AAAUCAACACCCUGUCAUUUUUAUGGCAGGGUG;または

GACUUCUCCACAGGAGUCAGAUGUUUUUGUACUCUCAAGAUUUUAAGUACUGUACAA
CGAAACUACACAGUUACUUAUUUUCUUGCAGAAGCUACAAAGAUAAAGGCUUCAUGCC
GAAAUCAACACCCUGUCAUUUUUAUGGCAGGGUG

から選択される核酸配列を含む。

【0033】

本願明細書に記載される任意の態様の様々な態様において、タンパク質核酸複合体は、本願明細書に記載される任意の態様による塩基エディターおよび本願明細書に記載される任意の態様によるガイドRNAを含む。

【0034】

本明細書に記載される任意の態様の様々な実施形態において、改変されたSaCas9は、アミノ酸置換E782K、N968K、およびR1015Hまたはそれらに対応するアミノ酸置換を含む。様々な実施形態において、SaCas9は、以下のアミノ酸配列を含む：

KRNYILGLAIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRRRRH
RIQRVKKLLFDYNLLTDHSELGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGVHNVNE
VEEDTGNELSTKEQISRNSKALEEKYVAELQLERLKKDGEVRGSINRFKTS DYVKEAKQLL
KVQKAYHQLDQSFIDTYIDLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEE
LRSVKYAYNADLYNALNDLNNLVITRDENEKLEYEKFQIENVFQKQKKPTLKQIAKEIL
VNEEDIKGYRVTSTGKPEFTNLKVYHDIKDITARKEIENAELLDQIAKILTIYQSSEDIQEE
LTNLNSELQTQEEIEQISNLKGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPPKV
DLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIKKYGLPNDIIELAREKNSKDAQKMI
NEMQKRNRQTNERIEEII RTTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNNPFN
YEVDHIIPRSVSFDNSFNNKVLVKQEENSKKGNRTPFQYLSSSDSKISYETFKKHILNLA
KGGRISKTKEYLLEERDINRFSVQKDFINRNLVDTRYATRGLMNLRSYFRVNNLDVKV
KSINGGFTSFLRRKWKFKKERNKGYKHAEDALIINANADFIFKEWKLDKAKKVMENQM
FEEKQAESMPEIETE QEYKEIFITPHQIKHIKDFKDYKYSHRVDKKPNRKLINDTLYSTRK
DDKGNTLIVNNLNGLYDKDNDKLKLINKSPEKLLMYHHDPTQTYQKLKLIMEQYGDEN
PLYKYYEETGNYLT KYSKKDN GPVIKKIKYYGNKLNAHL DITDDYPNSRNKVVKLSLKP
YRFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSKCYEEAKKLKKNQAEFIASFYKNDLIK
INGELYRVIGVNNDLLNRIEVMIDITYREYLENMNDKRPPHIIKTIASKTQSIKKYSTDIL
GNLYEVKSKKHPQIIKKG。

10

20

【0035】

本明細書に記載される任意の態様の様々な実施形態において、塩基エディターは、以下のアミノ酸配列を含む：

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAH
AEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAGS
LMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRREQEKAQKKAQSSTDSSGSSGSS
GSETPGTSESATPESSGGSSGSSGSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVN
RVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHS
RIGRVVFGVRNAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLCYFFRMPRQVF
NAQKKAQSSTDSSGSSGSSGSETPGTSESATPESSGGSSGSSGSKRNYILGLAIGITSVGYGII
DYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRRRRHRIQRVKKLLFDYNLLTDH
SELGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNS
KALEEKYVAELQLERLKKDGEVRGSINRFKTS DYVKEAKQLLKVQKAYHQLDQSFIDTYI
DLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALND
LNNLVITRDENEKLEYEKFQIENVFQKQKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEF
TNLKVYHDIKDITARKEIENAELLDQIAKILTIYQSSEDIQEELTNLNSELQTQEEIEQISNL
KGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPPKVDLSQQKEIPTTLVDDFILS
PVVKRSFIQSIKVINAIKKYGLPNDIIELAREKNSKDAQKMINEMQKRNRQTNERIEEII
RTTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNNPFNYEVDHIIPRSVSFDNSFNN
KVLVKQEENSKKGNRTPFQYLSSSDSKISYETFKKHILNLA KGGRISKTKEYLLEERDIN
RFSVQKDFINRNLVDTRYATRGLMNLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKE
RNKGYKHAEDALIINANADFIFKEWKLDKAKKVMENQMFEKQAESMPEIETE QEYKEI
FITPHQIKHIKDFKDYKYSHRVDKKPNRKLINDTLYSTRKDDKGNTLIVNNLNGLYDKDN
DKLKKLINKSPEKLLMYHHDPTQTYQKLKLIMEQYGDENPLYKYYEETGNYLT KYSKK
DNGPVIKKIKYYGNKLNAHL DITDDYPNSRNKVVKLSLKP YRFDVYLDNGVYKFVTVKNLD
VIKKENYYEVNSKCYEEAKKLKKNQAEFIASFYKNDLIKINGELYRVIGVNNDLLNRIE
VMIDITYREYLENMNDKRPPHIIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKGGA
DKRTADGSEFESPKKKRKV。

30

40

50

【0036】

種々の実施形態において、塩基エディターは、以下のアミノ酸配列を含む：

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAH
AEIMALRQGGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAGS
LMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAKKAQSSTDSSGSSGGSS
GSETPGTSESATPESSGGSSGSSGSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVNN
RVIGEGWNRAIGLHDPTAHAEIMALRQGGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHS
RIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVF
NAQKKAQSSTDSSGSSGGSSGSETPGTSESATPESSGGSSGGSKRNYILGLAIGITSVGYGII
DYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRRRRHRIQRVKKLLFDYNLLTDHS
ELSGINPYEARVKGLSQKLSEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNS
KALEEKYVAELQLERLKKDGEVRGSINRFKTSYVKEAKQLLKVKQAYHQLDQSFIDTYI
DLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALND
LNNLVITRDENEKLEYEYEFQIENVFQKQKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEF
TNLKVVYHDIKDITARKEIENAELLDQIAKILTIIYQSSEDIQEELTNLNSELTQEEIEQISNL
KGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILS
PVVKRSFIQSIKVINAIKKYGLPNDIIELAREKNSKDAQKMINEMQKRNRQTNERIEEIIIR
TTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNNPFNYEVDHIIIPRSVSFDNSFNN
KVLVKQEENSCKGNRTPFQYLSSSDSKISYETFKKHILNLAGKGRISKTKEYLLEERDIN
RFSVQKDFINRNLVDTRYATRGLMNLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKE
RNKGKHHAEALIINANADFIFKEWKKLDKAKKVMENQMFEKKAESMPEIETEYKEYKEI
FITPHQIKHIKDFKDYKYSHRVDKKPNRKLINDTLYSTRKDDKGNTLIVNNLNGLYDKDN
DKLKKLINKSPEKLLMYHHDPTQYQKLKLIMEQYGDEKNPLYKYEETGNLYLTKYSKKD
NGPVIKKIKYYGNKLNAHLDDITDYPNSRNKVVKLSLKPYRFDVYLDNGVYKFVTVKNLD
VIKKENYYEVNSKCYEEAKKLKISNQAEFIASFYKNDLIKINGELYRVIGVNNDDLNRIEV
NMIDITYREYLENMNDKRPPHIIKTASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKGEGA
DKRTADGSEFESPKKKRKVSSGNSNANSRGPSSFSSGLVPLSLRGSHSRPGERPFQCRICMR
NFSRNEHLEVHTRTHTGEKPFQCRICMRNFSQSTTLKRHLRTHTGEKPFQCRICMRNFSR
TEHLARHLKTHLRGSSAQ; または

【0037】

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAH
AEIMALRQGGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAGS
LMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAKKAQSSTDSSGSSGGSS
GSETPGTSESATPESSGGSSGSSGSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVNN
RVIGEGWNRAIGLHDPTAHAEIMALRQGGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHS
RIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVF
NAQKKAQSSTDSSGSSGGSSGSETPGTSESATPESSGGSSGGSKRNYILGLAIGITSVGYGII
DYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRRRRHRIQRVKKLLFDYNLLTDHS
ELSGINPYEARVKGLSQKLSEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNS
KALEEKYVAELQLERLKKDGEVRGSINRFKTSYVKEAKQLLKVKQAYHQLDQSFIDTYI
DLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALND
LNNLVITRDENEKLEYEYEFQIENVFQKQKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEF
TNLKVVYHDIKDITARKEIENAELLDQIAKILTIIYQSSEDIQEELTNLNSELTQEEIEQISNL
KGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILS
PVVKRSFIQSIKVINAIKKYGLPNDIIELAREKNSKDAQKMINEMQKRNRQTNERIEEIIIR
TTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNNPFNYEVDHIIIPRSVSFDNSFNN
KVLVKQEENSCKGNRTPFQYLSSSDSKISYETFKKHILNLAGKGRISKTKEYLLEERDIN
RFSVQKDFINRNLVDTRYATRGLMNLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKE
RNKGKHHAEALIINANADFIFKEWKKLDKAKKVMENQMFEKKAESMPEIETEYKEYKEI
FITPHQIKHIKDFKDYKYSHRVDKKPNRKLINDTLYSTRKDDKGNTLIVNNLNGLYDKDN

10

20

30

40

50

DKLKKLINKSPEKLLMYHHDPTQYQKLKLIMEQYGDENPLYKYYEETGNYLTKYSKKD
NGPVIKKIKYYGNKLNALHDITDDYPNSRNKVVKLSLKPYPYRFDVYLDNGVYKFVTVKNLD
VIKKENYYEVNSKCYEEAKKLKKISNQAEFIASFYKNDLIKINGELYRVIGVNNDLLNRIEV
NMIDITYREYLENMNDKRPPHIIKTISKYSTDILGNLYEVKSKKHPQIIKKGEGA
DKRTADGSEFESPKKKRKVSSGNSNANSRGPFSFSSGLVPLSLRGSHSRPGERPFQCRICMR
NFSRGEHLRQHTRTHTGEKPFQCRICMRNFSQSGTLKRHLRTHHTGEKPFQCRICMRNFSR
NDKLVPHLKTHLRGSSAQ.

【0038】

本願明細書に記載される任意の態様の様々な実施形態において、ガイドRNAは、核酸配列GUUUUAGUACUCUGUAAUGAAAAUACAGAAUCUACUAAAACAAGGCAAAAUGCCGUGUUUAUCUCGUCAACUUGUUGGCGAGAUUUUUUをさらに含む。

10

【0039】

様々な実施形態において、ガイドRNAは、核酸配列UCCACAGGAGUCAGAUGCACGUUUAUAGUACUCUGUAAUGAAAAUACAGAAUCUACUAAAACAAGGCAAAAUGCCGUGUUAUCUCGUCAACUUGUUGGCGAGAUUUUUU、または核酸配列CUCCACAGGAGUCAGAUGCACGUUUUAGUACUCUGUAAUGAAAAUACAGAAUCUACUAAAACAAGGCAAAU GCCGUGUUUAUCUCGUCAACUUGUUGGCGAGAUUUUUUを含む。

【0040】

いくつかの実施形態において、本明細書に提供される方法は、さらなる核酸塩基の第2の編集をさらに含む。一実施形態では、この追加の核酸塩基は遺伝的障害の原因ではない。別の実施形態では、追加の核酸塩基は遺伝的障害の原因である。

20

【0041】

もう1つの態様において、対象における遺伝的障害を治療する方法が提供され、この方法は、必要とする対象に塩基エディターを投与することと（ここで塩基エディターは、ガイドポリヌクレオチドと共に、ポリヌクレオチドプログラミング可能なヌクレオチド結合ドメインおよびデアミナーゼドメインを含む）；対象のポリヌクレオチドの標的ヌクレオチド配列にガイドポリヌクレオチドを結合させることと；標的ヌクレオチド配列へのガイドポリヌクレオチドの結合に際して核酸塩基を脱アミノ化することによって標的ヌクレオチド配列の核酸塩基を編集し、それによって核酸塩基を別の核酸塩基に変化させることにより遺伝的障害を治療することとを含み、ここで、該核酸塩基は遺伝子の調節エレメントまたは調節領域にある。

30

【0042】

もう1つの態様において、それを必要とする対象において遺伝的障害を治療するための細胞、組織、または器官を作製する方法が提供され、この方法は、該細胞、組織、または器官を塩基エディターと接触させることと（ここで塩基エディターは、ポリヌクレオチドプログラミング可能なヌクレオチド結合ドメインおよびデアミナーゼドメインをガイドポリヌクレオチドと共に含む）；細胞、組織または器官のポリヌクレオチドの標的ヌクレオチド配列にガイドポリヌクレオチドを結合させることと；ガイドポリヌクレオチドが標的ヌクレオチド配列に結合する際に核酸塩基を脱アミノ化することによって標的ヌクレオチド配列の核酸塩基を編集し、それによって核酸塩基を別の核酸塩基に変化させることによって、遺伝的障害を治療するための細胞、組織または器官を産生することとを含み、ここで、該核酸塩基は遺伝子の調節エレメント中にある。いくつかの態様において、本方法は、細胞、組織、または器官を対象に投与することをさらに含む。ある態様において、細胞、組織、または器官は、対象に対して自家性である。ある態様において、細胞、組織、または器官は、対象に対して同種異系である。ある態様において、細胞、組織、または器官は、対象に対して異種性である。

40

【0043】

上述した方法のいくつかの態様において、遺伝子は遺伝的障害の原因である。ある態様において、遺伝子は、遺伝的障害の原因ではない。ある態様において、編集は、遺伝子の

50

転写量の変化をもたらす。ある態様において、変化は、遺伝子の転写量の増加である。ある態様において、変化は、遺伝子の転写量の減少である。いくつかの実施形態において、編集は、少なくとも1つのタンパク質の調節エレメントへの結合パターンを改変する。ある態様において、調節エレメントは、遺伝子のプロモーター、エンハンサー、リプレッサー、サイレンサー、インスレーター、開始コドン、終止コドン、Kozakコンセンサス配列、スプライスアクセプター、スプライスドナー、スプライス部位、3'非翻訳領域(UTR)、5'非翻訳領域(UTR)、または遺伝子間領域である。一部の実施形態では、編集の結果、スプライス部位が除去される。一部の実施形態では、編集の結果、スプライス部位が追加される。いくつかの態様において、編集は、イントロン含有物をもたらす。一部の実施形態では、編集の結果、エクソンのスキップが生じる。いくつかの実施形態において、編集は、開始コドン、停止コドン、またはKozakコンセンサス配列の除去をもたらす。いくつかの実施形態において、編集は、開始コドン、停止コドン、またはKozakコンセンサス配列の付加をもたらす。いくつかの態様において、編集は、遺伝子の調節エレメントに位置する複数の核酸塩基を編集することを含む。

【0044】

上記に記載された方法のいくつかの態様において、編集は、複数の核酸塩基を編集することを含み、複数の核酸塩基の少なくとも1つの核酸塩基は、少なくとも1つのさらなる遺伝子の少なくとも1つのさらなる調節エレメントに位置する。ある態様において、遺伝子および少なくとも1つの追加の遺伝子は、少なくとも1つのタンパク質の1つ以上のサブユニットをコードする。

【0045】

上述の方法のいくつかの実施形態では、編集は、本明細書の表4に示される変化のいずれか1つから選択される。ある態様において、遺伝的障害は、鎌状赤血球貧血とも呼ばれる鎌状赤血球症(SCD)である。ある態様において、遺伝性障害は、胎児ヘモグロビンの遺伝的持続性(HPFH)である。一部の実施形態では、核酸塩基はHBG1/2のc. -114 ~ -102に位置する。ある態様において、核酸塩基は、HBG1/2のプロモーターに位置する。

【0046】

上記で記述された方法のいくつかの態様において、該方法は、少なくとも1つのさらなる核酸塩基の第二の編集を含み、ここで、該少なくとも1つのさらなる核酸塩基は、該遺伝子の調節エレメント内に存在しない。ある態様において、この追加的な核酸塩基は、タンパク質コード領域に位置する。

【0047】

上述した態様の方法の特定の態様において、デアミナーゼドメインは、アデノシンデアミナーゼドメインである。ある態様において、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る。ある態様において、アデノシンデアミナーゼドメインは、デオキシリボ核酸(DNA)中のアデニンを脱アミノ化することができる。ある態様において、ガイドポリヌクレオチドは、リボ核酸(RNA)またはデオキシリボ核酸(DNA)を含む。いくつかの実施形態において、ガイドポリヌクレオチドは、CRISPR RNA (crRNA) 配列、トランス活性化CRISPR RNA (tracrRNA) 配列、またはそれらの組み合わせを含む。

【0048】

いくつかの実施形態において、本明細書で提供される方法は、第2のガイドポリヌクレオチドをさらに含む。ある態様において、第2のガイドポリヌクレオチドは、リボ核酸(RNA)またはデオキシリボ核酸(DNA)を含む。いくつかの実施形態において、第2のガイドポリヌクレオチドは、CRISPR RNA (crRNA) 配列、トランス活性化CRISPR RNA (tracrRNA) 配列、またはそれらの組合せを含む。いくつかの実施形態において、第2のガイドポリヌクレオチドは、第2の標的ヌクレオチド配列に対して塩基エディターをターゲティングする。

【0049】

いくつかの実施形態において、ポリヌクレオチドプログラミング可能なDNA結合ドメインは、Cas9ドメイン、Cpf1ドメイン、CasXドメイン、CasYドメイン、Cas12b/C2c1

10

20

30

40

50

ドメイン、またはCas12c/C2c3ドメインを含む。ある態様において、ポリヌクレオチドプログラミング可能なDNA結合ドメインは、ヌクレアーゼ不活 (nuclease dead) である。いくつかの実施形態において、ポリヌクレオチドプログラミング可能なDNA結合ドメインは、ニックアーゼである。いくつかの実施形態において、ポリヌクレオチドプログラミング可能なDNA結合ドメインは、Cas9ドメインを含む。ある態様において、Cas9ドメインは、ヌクレアーゼ不活Cas9 (dCas9)、Cas9ニックアーゼ (nCAs9)、またはヌクレアーゼ活性Cas9を含む。いくつかの実施形態において、Cas9ドメインはCas9ニックアーゼを含む。いくつかの実施形態において、ポリヌクレオチドプログラミング可能なDNA結合ドメインは、操作された、または改変されたポリヌクレオチドプログラミング可能なDNA結合ドメインである。

10

【0050】

いくつかの実施形態では、本明細書で提供される方法のいずれかは、第2の塩基エディターをさらに含む。いくつかの実施形態では、第2の塩基エディターは、他の塩基エディターとは異なるデアミナーゼドメインを含む。

【0051】

いくつかの実施形態において、塩基編集は、20%未満のインデル形成をもたらす。いくつかの実施形態において、塩基編集は、15%未満のインデル形成をもたらす。いくつかの実施形態において、塩基編集は、10%未満のインデル形成をもたらす。いくつかの実施形態において、塩基編集は、5%未満のインデル形成をもたらす。いくつかの実施形態において、塩基編集は、4%未満のインデル形成をもたらす。いくつかの実施形態において、塩基編集は、3%未満のインデル形成をもたらす。いくつかの実施形態において、塩基編集は、2%未満のインデル形成をもたらす。いくつかの実施形態において、塩基編集は、1%未満のインデル形成をもたらす。いくつかの実施形態において、塩基編集は、0.5%未満のインデル形成をもたらす。いくつかの実施形態において、塩基編集は、0.1%未満のインデル形成をもたらす。一部の実施形態では、塩基編集は転座をもたらさない。

20

【0052】

本開示の特徴は、添付の特許請求の範囲に特に記載されている。本開示の特徴および利点のよりよい理解は、以下の詳細な説明を参照することによって得られる。この詳細な説明は、例示的な実施形態を示し、ここで、本開示の原理が説明され、利用され、添付の図面は以下の通りである。

30

【図面の簡単な説明】

【0053】

【図1】図1は、健常者とアンチトリプシン欠損症 (A1AD) 患者とを比較した模式図である。健常者では、 α -1アンチトリプシン (A1AT) タンパク質はプロテアーゼから肺を保護し、肝臓は α -1アンチトリプシンを血液中に放出する。 α 1アンチトリプシン欠損症 (A1AD) 患者において、正常な α 1アンチトリプシンの欠損は肺組織損傷をもたらす。肝臓の肝細胞に異常な α -1アンチトリプシンが蓄積すると肝硬変になる。

【0054】

【図2】図2は、異なる遺伝子型についての血清 α -1アンチトリプシン (A1AT) レベルの典型的な範囲を示す(標準 (MM); α -1アンチトリプシン欠損症のヘテロ接合体キャリア (MZ, SZ); ホモ接合体欠損症 (SS, ZZ))。血清 α -1アンチトリプシン (AAT) 濃度は、左「y」軸において μ Mで表されており、これが文献において一般的である。右の「y」軸は、臨床検査室および様々な測定技術(比濁法や放射免疫拡散法)によって一般的に報告されるmg/dL単位への血清AAT濃度のおおよその換算を示す。

40

【0055】

【図3】図3は、A1ATをコードするSERPINA 1遺伝子内のE342Kの補正のための標的部位的配列を示す。ハイライトされているのは、非標準的なspCas9 NGC PAM、ならびに編集によってE342Kの所望の補正がもたらされる標的A核酸塩基である。また、編集によってE342GまたはD341Gなどの良性アリルがもたらされ得るさらなるオフターゲットAも示されている。

50

【0056】

【図4】図4は、A1ATタンパク質の異なるバリエーションをコードするプラスミドで一時的にトランスフェクトされたHEK293Tの培養上清中の分泌タンパク質のレベルを示す棒グラフである。A1AT濃度は、既報の方法(Borel et al., 2017, “Alpha-1 Antitrypsin Deficiency: Methods and Protocols,” 10.1007/978-1-4939-7163-3)を用いてELISAにより測定した。A1ATの最も一般的な臨床的バリエーション(例えば、病原性突然変異)はE264V (PiSアリル)とE342K (PiZアリル)である。PiSおよびPiZタンパク質は、野生型タンパク質よりも生産量が少ない。D341Gタンパク質もE342Gタンパク質も野生型と同様のレベルで産生される。従って、本明細書に記載のアデニン塩基エディターおよび塩基編集法を用いて、肝細胞からのA1AT分泌を回復させ、同時に肝毒性を改善し、肺へのA1ATの循環を増加させることができるこれらの良性アリルが作製された。図中、A1AT: α -1アンチトリプシン; A1AD: α 1アンチトリプシン欠損症; 「Z突然変異」はE342K (PiZアリル)変異; 「S変異」はE264V (PiSアリル)である。

10

【0057】

【図5】図5は、TadAから出発してDNAデオキシアデノシンデアミナーゼを進化させる戦略を示す概略図である。E.coliのライブラリーは、dCas9に融合された変異ecTadA (TadA*) 遺伝子のプラスミドライブラリーと、抗生物質耐性遺伝子を修復するために標的化されたA・TからG・Cへの突然変異を必要とする選択プラスミドとを保有している。生き残ったTadA*バリエーションからの突然変異は、ヒト細胞における塩基編集のために、ABE構築物に移入された。

20

【0058】

【図6】図6は、正常HbA、病原性バリエーションである鎌状HbSおよびHbC、ならびに、表現型はHbAに類似しHbSのように重合しないHbG Makassarバリエーションを含む、成熟ヘモグロビン (Hb) の最初の8個のアミノ酸を示す表を提示する。図6には、Hbタイプの各々におけるアミノ酸位置6でコードされるアミノ酸、ならびにこれらのHbタンパク質の最初の8個のアミノ酸をコードするDNAおよびmRNA配列が示されている。

【0059】

【図7】図7Aおよび7Bは、異なるPAM配列を認識する種々のA→G塩基エディター (ABE)を用いて、HbSのアミノ酸位置6においてバリンをコードするコドンに相補的な配列 (CAC) 中の核酸塩基アデノシン (A) をグアノシン (G) に編集する実験結果を示す。図7Aは、所望の編集および潜在的なオフターゲット編集の位置を含む、テストされたHBB gRNAおよび対応するABEの特徴を説明する表である。図7Bは、鎌状赤血球標的的部位における塩基編集のためにABEを使用した結果を示すグラフである。

30

【0060】

【図8】図8A～8Gは、NNNRRTに対する寛容性を有するStaphylococcus aureus Cas9バリエーション (saKKH)を、単独で、または鎌状赤血球標的的部位における配列特異性を有するDNA結合ドメインに融合されたかたちで用いて、HbSのアミノ酸位置6におけるバリンをコードするコドン(CAC)中のアデノシン (A) をグアノシン (G) に編集する実験の結果を示す。図8Aは、saKKH ABE7.10、saKKH ABE7.10 zf1ra、およびsaKKH ABE7.10 zf1rbを含むポリペプチド内のドメインの構成を示す、ABE構築物の概略図を示す。図8Bは、鎌状赤血球標的的部位における核酸配列、ならびに下の線によって示されるようにガイドRNAの標的相補配列(g1およびg4)を示す。図8Cは、saKKH ABE7.10、saKKH ABE7.10 zf1ra、およびsaKKH ABE7.10 zf1rbを、鎌状赤血球症の標的的部位に相補的な20ヌクレオチド (nt) 長の核酸配列を有するガイドRNA g1と組み合わせて用いた結果を示すグラフである。図8Cのグラフの右側は、鎌状赤血球標的的部位における核酸配列およびg1ガイドRNAの標的相補配列である。図8Dは、saKKH ABE7.10、saKKH ABE7.10 zf1ra、およびsaKKH ABE7.10 zf1rbを、鎌状赤血球症標的的部位に相補的な長さ21 ntの核酸配列を有するガイドRNA g1と組み合わせて用いた結果を示すグラフである。図8Eは、saKKH ABE7.10、saKKH ABE7.10 zf1ra、およびsaKKH ABE7.10 zf1rbを、鎌状赤血球症標的的部位に相補的な長さ20 ntの核酸配列を有するガイドRNA g4と組み合わせて用いた結果

40

50

を示すグラフである。図8Eのグラフの右側は、鎌状赤血球症標的的部位における核酸配列およびg4ガイドRNAの標的相補配列である。図8Fは、saKKH ABE7.10、saKKH ABE7.10 zf1ra、およびsaKKH ABE7.10 zf1rbを、鎌状赤血球症標的的部位に相補的な長さ21 ntの核酸配列を有するガイドRNA g4と組み合わせて用いた結果を示すグラフである。図8Gは、制御HEK2サイトにおける塩基編集を示す。

【0061】

【図9】図9A～9Eは、鎌状赤血球症標的的部位での塩基編集のための*Streptococcus thermophilus* Cas9 (St1Cas9) DNA結合ドメインを有するアデノシン塩基エディター (ABE) の開発および評価を示す。図9Aは、ABE St1Cas9を、St1Cas9標準PAM配列、NNAGAA (TTCTAG;逆相補鎖)とともに使用する塩基編集を示す。下側の挿入図は、塩基編集部位における、ABE St1Cas9、St1 Cas9ヌクレアーゼ、および未処理を比較したインデル割合 (インデル%) を示す。図9Bは、St1Cas9標準PAM配列NNAGAAを有するABE St1Cas9を用いた塩基編集を示す。下の挿入図は、塩基編集部位における、ABE St1Cas9、St1Cas9ヌクレアーゼ、および未処理を比較したインデル割合を示す。図9Cは、ABE St1Cas9をSt1Cas9非標準PAM配列NNACCA (TG GTNN;逆相補鎖)と共に使用する塩基編集を示す。下側の挿入図は、塩基編集部位においてABE St1Cas9、St1Cas9ヌクレアーゼ、および未処理のものを比較したインデルの割合を示す。図9Dは、ABE St1Cas9をSt1Cas9非標準PAM配列NNACCA (TG GTNN;逆相補鎖)と共に使用する塩基編集を示す。下側の挿入図は、塩基編集部位においてABE St1Cas9、St1Cas9ヌクレアーゼ、および未処理のものを比較したインデルの割合を示す。図9Eは、鎌状赤血球症標的的部位においてSt1Cas9非標準PAM配列NNACCAを有するABE St1Cas9を用いた塩基編集を示す。矢印は、Hbの鎌状赤血球症標的的部位においてABE-St1Cas9塩基エディターによってA・TからG・Cへの突然変異 (Val→Ala)が誘導されたことを示す。

【0062】

【図10】図10は、NGC PAMを受け入れるように進化および操作されたSpCas9 DNA結合ドメインを有するABE (ngcABE)を用いた、鎌状赤血球症標的的部位での塩基編集のパーセントを示す。棒グラフの左端のバーは「Pro6Pro」を表し、中央のバーは「Val7Ala」、右端のバーは「Ser10Pro」を表す。

【0063】

【図11】図11は、HBG1/2遺伝子のプロモーター領域を表す概略図である。個々の紫色の三角形は、HPFH患者に自然に認められるSNPおよび欠失を示し、緑色の矢印、例えば「BCL11A」、「CCAAT」、「90 BCL11A」および「ZBTB7A」は、潜在的な転写結合部位を示す。HBG1/2配列の上下に集まった太い尖った線 (ピンク) は、これらの関心領域、例えば、遺伝子の標的配列を標的とすることができるガイドRNAを示す。

【0064】

【図12】図12は、示されたgRNAおよびCas9塩基エディターでトランスフェクトされた293T細胞におけるHBG1/2遺伝子における標的配列の標的塩基編集率を示す。塩基編集効力のパーセンテージはMiseqにより決定した。図に示されているのは、それぞれのタイプのgRNAを用いて293T細胞で起こった編集の割合であり、その遺伝子および標的配列は表4に示されている。「Cs」は、gRNAと合わせたCBEにより編集が行われるであろう位置をgRNAとの関係で示す。「As」は、それぞれのgRNAと共にABEが配列を編集するであろう位置をgRNAとの関係で示す。

【0065】

【図13】図13は、各タイプのgRNAによって実施された初代骨髄CD 34+細胞における編集のパーセンテージを示し、遺伝子および標的の配列は表4に示される。CD 34+細胞を、示されたgRNAおよび塩基エディターでトランスフェクトした。「Cs」は、gRNAと合わせたBE4などのCBEによる編集が行われるであろう位置をgRNAとの関係で示す。「As」は、ABEがそれぞれのgRNAとともに標的配列を編集する位置をgRNAとの関係において示す。HBG1およびHBG 2遺伝子座の両方における塩基編集のパーセンテージは、Miseqによって評価された。

10

20

30

40

50

【発明を実施するための形態】

【0066】

本明細書に記載するように、本発明は、プログラム可能な核酸塩基エディターを用いて病原性アミノ酸を置換するための組成物および方法の特徴とする。特定の局面において、記載された組成物および方法は、HBB遺伝子によってコードされるβ-グロビンタンパク質の6番目のアミノ酸におけるGlu→Val突然変異によって引き起こされる鎌状赤血球症の治療に有用である。遺伝子編集の分野におけるこれまでの多くの開発にもかかわらず、Glu→Valを復帰させるための疾患HBB遺伝子の精確な補正は依然として達成し難く、CRISPR/CasヌクレアーゼまたはCRISPR/Cas塩基編集アプローチのいずれを用いても達成されていない。

10

【0067】

CRISPR/Casヌクレアーゼアプローチを用いて患部ヌクレオチドを置換するHBB遺伝子のゲノム編集は、ゲノムDNAの切断を必要とするが、ゲノムDNAの切断は、塩基挿入/欠失（インデル）を生じるリスクを増加させ、これは、早期終止コドンの生成、コドン読み取り枠の変更などを含む、意図しない望ましくない結果を引き起こす可能性がある。さらに、β-グロビン遺伝子座での二本鎖切断の生成は、組換え事象を介して遺伝子座を根本的に改変してしまう可能性がある。β-グロビン遺伝子座には、互いに配列同一性をもつグロビン遺伝子（-5'-ε-；Gγ-；Aγ-；δ-；およびβ-グロビン-3'）のクラスターがある。β-グロビン遺伝子座のこの構造のために、遺伝子座内の二本鎖切断の組換え修復は、グロビン遺伝子間、例えばδ-グロビン遺伝子とβ-グロビン遺伝子間の介在配列の遺伝子喪失をもたらす可能性がある。遺伝子座の意図しない改変はサラセミアを引き起こすリスクももたらす。

20

【0068】

CRISPR/Cas塩基編集アプローチは、核酸塩基レベルで正確な変化を生成する能力があるという点で有望である。しかしながら、Val→Glu（GTG→GAG）の精確な補正は、現在存在することが知られていないT・AからA・Tへの変換エディターを必要とする。さらに、CRISPR/Casの塩基編集の特異性は、部分的に、CRISPR/CasがDNAに結合するときにR-ループ形成によって生じる、編集可能なヌクレオチドの限られたウィンドウによるものである。したがって、CRISPR/Casのターゲティングは、塩基編集を可能にするためには、鎌状赤血球部位またはその近くで起こらなければならない、ウィンドウ内での最適な編集のためのさらなる配列要件が存在する可能性がある。

30

【0069】

CRISPR/Casターゲティングのための一つの必要条件は、標的化される部位に隣接するプロトスペーサー隣接モチーフ（PAM）配列の存在である。例えば、多くの塩基編集は、PAM配列NGGを必要とするSpCas9に基づいている。T・AからA・Tへのトランスバージョンが可能であると仮定しても、このようなSpCas9塩基編集にとって望ましい位置にターゲット「A」を配置するNGG PAMは存在しない。利用可能なPAMのコレクションを拡大する多くの新しいCRISPR/Casタンパク質が発見または生成されたが、PAMの必要性は、CRISPR/Cas塩基エディターをゲノムの任意の位置の特定のヌクレオチドに導く能力における制限因子であり続ける。

40

【0070】

本発明は、鎌状赤血球貧血の治療のためのゲノム編集アプローチを提供するための前述の課題に対処する、本明細書に記載されるいくつかの発見に少なくとも部分的に基づく。一態様において、本発明は、鎌状赤血球症を引き起こすHbタンパク質のアミノ酸位置6のバリンをアラニンで置換し、それによって、鎌状赤血球表現型を生じないHbバリエーション（Hb Makassar）を生成する能力に部分的に基づく。正確な補正（GTG→GAG）は、T・AからA・Tへの変換塩基エディターなしでは不可能であるが、本明細書に記載される結果は、A・TからG・Cへの塩基エディター（ABE）を使用して、Val→Ala（GTG→GCG）置換（すなわち、Hb Makassarバリエーション）を生成し得るという知見を示す。これは、本明細書に提供されるように、新規の塩基エディターおよび新規の塩基編集戦略の開発によって部

50

分的に達成された。例えば、鎌状赤血球症の標的部位における最適な塩基編集のために隣接配列（例えばPAM配列;ジンクフィンガー結合配列）を利用する新規なABE塩基エディター（すなわち、アデノシンデアミナーゼドメインを有するもの）が開発された。

【0071】

β-グロビンタンパク質の鎌状赤血球症バリエーション(鎌状HbS; E6V)の第六アミノ酸のコドン中のチミジン(T)をシチジン(C)に塩基編集し、それによってこのアミノ酸位置のバリンアミノ酸残基をアラニンアミノ酸残基に置換(V6A)するための組成物および方法が本明細書に提供され、記載される。HbSの6位でバリンがアラニンに置換されると、鎌状赤血球の表現型をもたない(例えば、病原性バリエーションHbSの場合のように重合する可能性はない)βグロビンタンパク質バリエーションが生じる。従って、本発明の組成物および方法は、鎌状赤血球症の治療に有用である。

10

【0072】

本明細書に提供され、記載されるのは、本明細書の表3A、3B、または4に提供される遺伝子によって引き起こされるか、またはそれら遺伝子に関連する、疾患または障害を治療するための、本明細書に記載される塩基エディターおよび塩基エディターシステムを含む組成物および方法である。

【0073】

以下の説明および実施例は、本開示の実施形態を詳細に例示する。本開示は、本明細書に記載される特定の実施形態に限定されず、従って変化し得ることが理解されるべきである。本開示には、その範囲内に包含される多数の変形および修正があることを当業者は認識するであろう。

20

【0074】

全ての用語は、当業者によって理解されるものとして理解されることが意図されている。別段の定義がない限り、本明細書で使用される全ての技術的および科学的用語は、本開示が関係する分野の当業者によって一般的に理解されるものと同じ意味を有する。

【0075】

本明細書中で使用されるセクション見出しは、まとめる目的だけのためであり、記載される主題を限定するものと解釈されるべきではない。

【0076】

本開示の種々の特徴は、単一の実施形態に関連して説明され得るが、これらの特徴は、分離されて、または任意の適切な組み合わせで提供することもできる。逆に、本開示は、明確にするために別々の実施形態の文脈で説明され得るが、本開示はまた、単一の実施形態で実施することもできる。

30

【0077】

[定義]

別段の定義がない限り、本明細書で使用されるすべての技術的および科学的用語は、本発明が属する分野の当業者によって一般的に理解される意味を有する。以下の参考文献は、当業者に、本発明において使用される多くの用語の一般的な定義を提供する: Singleton et al., Dictionary of Microbiology and Molecular Biology (2nd ed. 1994); The Cambridge Dictionary of Science and Technology (Walker ed., 1988); The Glossary of Genetics, 5th Ed., R. Rieger et al. (eds.), Springer Verlag (1991); and Hale & Marham, The Harper Collins Dictionary of Biology (1991)。

40

【0078】

本出願において、単数形の使用は、特に断りのない限り、複数形を含む。本明細書で 사용되는ように、単数形「a」、「an」、および「the」は、文脈が明確にそうでないことを指示しない限り、複数の指示対象を含むことに留意されたい。本出願において、「又は」の使用は、別段の記載がない限り、「及び/又は」を意味する。さらに、用語「含む(including)」、ならびに「含む(include)」、「含む(includes)」、および「含まれる(included)」などの他の形態の使用は、非限定的である。

【0079】

50

本明細書及びクレームにおいて使用される場合、用語「含む(comprising)」(および「含む(comprise)」および「含む(comprises)」などのそのあらゆる形態)、「有する(having)」(「有する(have)」および「有する(has)」などのそのあらゆる形態)、「含む(including)」(「含む(include)」および「含む(includes)」などのそのあらゆる形態)又は「含む(containing)」(「含む(contains)」および「含む(contain)」などのそのあらゆる形態)は、包括的又は開放的であり、追加の、記載されていない要素又は方法工程を排除しない。本明細書で議論される任意の実施形態は、本開示の任意の方法または組成物に関して実施することができ、逆もまた同様であると考えられる。さらに、本開示の組成物を用いて、本開示の方法を達成することができる。

【0080】

用語「約(about)」または「およそ(approximately)」は、当業者によって決定されるように、特定の値について許容可能な誤差範囲内であることを意味し、これは、値がどのように測定または決定されるか、すなわち、測定システムの制限に部分的に依存する。例えば、「約」は、当該技術分野における実務によれば、1以内または1を超える標準偏差内であることを意味し得る。あるいは、「約」は、所定の値の最大20%、最大10%、最大5%、または最大1%の範囲を意味し得る。別法として、特に生物学的システムまたはプロセスに関して、この用語は、同じ桁以内、好ましくは5倍以内、より好ましくは2倍以内の値を意味することができる。特定の値が出願及び特許請求の範囲に記載されている場合、別段の記載がない限り、その特定の値について許容可能な誤差範囲内にあることを意味する「約」という用語が推定されるべきである。

【0081】

明細書における「いくつかの実施形態」、「ある実施形態」、「一実施形態」、または「他の実施形態」という言及は、その実施形態に関連して説明される特定の特徵、構造、または特性が、本開示の少なくともいくつかの実施形態に含まれるが、必ずしもすべての実施形態に含まれるとは限らないことを意味する。

【0082】

「アデノシンデアミナーゼ」とは、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ことができるポリペプチドまたはその断片を意味する。ある態様において、デアミナーゼまたはデアミナーゼドメインは、アデノシンからイノシン、またはデオキシアデノシンからデオキシイノシンへの加水分解的脱アミノ化を触媒するアデノシンデアミナーゼである。ある態様において、アデノシンデアミナーゼは、デオキシリボ核酸(DNA)中の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。本明細書中で提供されるアデノシンデアミナーゼ(例えば、遺伝子操作されたアデノシンデアミナーゼ、進化させたアデノシンデアミナーゼ)は、細菌などの任意の生物由来であり得る。

【0083】

「投与すること」は、本明細書に記載される1以上の製品または組成物を患者または対象に提供することとして本明細書で言及される。例として、限定するものではないが、製品または組成物の投与、例えば注射は、静脈内(i.v.)注射、皮下(s.c.)注射、皮内(i.d.)注射、腹腔内(i.p.)注射又は筋肉内(i.m.)注射によって行われ得る。1つ以上のそのような経路を用いることができる。非経口投与は、例えば、ボーラス注射によって、または経時的に徐々に灌流することによって行うことができる。あるいは、または同時に、投与は経口経路によることができる。限定するものではないが、鼻腔内、直腸、頭蓋内、腔内、頬側、胸部、皮内、経皮などの他の投与モードも想定される。

【0084】

「薬剤(agent)」とは、任意の小分子化合物、抗体、核酸分子、またはポリペプチド、またはそれらの断片を意味する。

【0085】

「改善する(ameliorate)」とは、疾患の発生または進行を減少させる、抑制する、減衰させる、減少させる、停止させる、または安定させることを意味する。

【0086】

「改変」とは、本明細書に記載されるような標準技術の公知の方法によって検出されるような、遺伝子またはポリペプチドの発現レベルまたは活性の変化（増加または減少）を意味する。本明細書で使用される場合、改変は、発現レベルの10%の変化、好ましくは25%の変化、より好ましくは40%の変化、最も好ましくは50%以上の発現レベルの変化を含む。

【0087】

「アナログ」とは、同一ではないが、類似の機能的または構造的特徴を有する分子を意味する。例えば、ポリペプチドアナログは、対応する天然ポリペプチドの生物学的活性を保持しながら、天然ポリペプチドと比較してそのアナログの機能を増強させる特定の生化学的修飾を有する。そのような生化学的修飾は、例えばリガンド結合を変化させることなく、アナログのプロテアーゼ抵抗性、膜透過性、または半減期を増加させることができる。アナログは、非天然アミノ酸を含み得る。

【0088】

「塩基エディター (BE)」または「核酸塩基エディター (NBE)」とは、ポリヌクレオチドに結合し、核酸塩基改変活性を有する薬剤を意味する。様々な実施形態において、塩基エディターは、核酸塩基改変ポリペプチド(例えばデアミナーゼ)と、ポリヌクレオチドプログラミング可能なヌクレオチド結合ドメインとを、ガイドポリヌクレオチド(例えばガイドRNA)と共に含む。様々な態様において、この薬剤は、塩基編集活性を有するタンパク質ドメイン、すなわち、核酸分子(例えばDNA)内の塩基(例えばA、T、C、G、U)を改変することができるドメインを含む生体分子複合体である。ある態様において、ポリヌクレオチドプログラミング可能なDNA結合ドメインは、デアミナーゼドメインに融合または連結される。一実施形態において、薬剤は、塩基編集活性を有するドメインを含む融合タンパク質である。別の態様において、塩基編集活性を有するタンパク質ドメインは、ガイドRNAに連結される(例えば、ガイドRNA上のRNA結合モチーフおよびデアミナーゼに融合されたRNA結合ドメインを介して)。ある態様において、塩基編集活性を有するドメインは、核酸分子内の塩基を脱アミノ化することができる。ある態様において、塩基エディターは、DNA分子内の塩基を脱アミノ化することができる。ある態様において、塩基エディターは、DNA内のシトシン (C) またはアデノシン (A) を脱アミノ化することができる。ある態様において、塩基エディターは、シチジン塩基エディター (CBE) である。一部の実施形態では、塩基編集はアデノシン塩基エディター (ABE) である。ある態様において、アデノシンデアミナーゼは、*TadA*から進化される。ある態様において、ポリヌクレオチドプログラミング可能なDNA結合ドメインは、CRISPR関連(例えばCasまたはCpf1)酵素である。ある態様において、塩基エディターは、デアミナーゼドメインに融合された、触媒的には死んだCas9 (dCas9) である。ある態様において、塩基エディターは、デアミナーゼドメインに融合されたCas9ニッカーゼ (nCas9) である。ある態様において、塩基エディターは、塩基除去修復 (BER) の阻害因子に融合される。ある態様において、塩基除去修復の阻害因子は、ウラシルDNAグリコシラーゼ阻害因子 (UGI) である。ある態様において、塩基除去修復の阻害因子は、イノシン塩基除去修復阻害因子である。塩基エディターの詳細は、国際PCT出願番号PCT/2017/045381(国際公開第2018/027078号)およびPCT/US 2016/058344(国際公開第2017/070632号)に記載されており、それぞれその全体が参照により本明細書に組み込まれる。Komor, A.C., et al., “Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage” *Nature* 533, 420-424 (2016); Gaudelli, N.M., et al., “Programmable base editing of A·T to G·C in genomic DNA without DNA cleavage” *Nature* 551, 464-471 (2017); 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” *Science Advances* 3:eaao4774 (2017), および Rees, H.A., et al., “Base editing: precision chemistry on the genome and transcriptome of living cells.” *Nat Rev Genet.* 2018 Dec;19(12):770-788. doi: 10.1038/s41576-018-0059-1も参照のこと(その内容全体を参照により本明細書に組み込む)。

10

20

30

40

50

【0089】

「シチジンデアミナーゼ」とは、アミノ基をカルボニル基に変換する脱アミノ化反応を触媒することができるポリペプチドまたはその断片を意味する。1つの実施形態において、シチジンデアミナーゼは、シトシンをウラシルに、または5-メチルシトシンをチミンに変換する。Petromyzon marinus由来のPmCDA1 (Petromyzon marinus cytosine deaminase 1、「PmCDA1」)、哺乳動物（例えば、ヒト、ブタ、ウシ、ウマ、サル等）由来のAID（活性化誘導シチジンデアミナーゼ; AICDA）、およびAPOBECは、例示的シチジンデアミナーゼである。

【0090】

例として、シチジン塩基エディタBE4は、以下の核酸配列を有する。BE4核酸配列に対して少なくとも95%以上の同一性を有するポリヌクレオチド配列も包含される。

ATGagctcagagactggcccagtggtgtggacccacattgagacggcggatcgagcccatgagtttgag
gtattcttcgatccgagagagctccgcaaggagacctgcctgctttacgaaattaattgggggggcccggcact
ccatttggcgacatacatcacagaacactaacaagcacgtcgaagtcaacttcacgagaagttcacgacaga
aagatatttctgtccgaacacaaggtgcagcattacctggttctcagctggagcccatcgggcgaatgtagta
gggccatcactgaattcctgtcaaggtatccccacgtcactctgtttatttacatcgcaaggctgtaccaccacg
ctgacccccgcaatcgacaaggcctgcgggatttgatctcttcaggtgtgactatccaaattatgactgagcag
gagtcaggatactgctggagaaactttgtgaattatagcccagtaataagcccactggcctaggtatcccc
atctgtgggtacgactgtacgttcttgaactgtactgcatcatactgggcctgcctccttgtctcaacattctgag
aaggaagcagccacagctgacattctttaccatcgctcttcagctctgtcattaccagcgactgccccacacat
tctctggggccaccgggttgaaatctggtggttctctggtggttctagcggcagcgagactcccgggacctcag
agtccgccacacccgaaagtctggtggttctctggtggttctgataaaaagtattctattggttttagccatcg
gcactaattccgttggatgggctgtcataaccgatgaatacaaaagtaccttcaaagaaatthaagggtgttgggg
aacacagaccgtcattcgattaaaaagaatcttatcggtgccctcctattcgatagtggcgaacggcagagg
cgactgcctgaaacgaaccgctcggagaaggtatacacgtcgcaagaaccgaatatgttacttacaagaaat
tttagcaatgagatggccaaagttgacgattctttcttcaccgtttggaagagtccttccttgtcgaagagga
caagaaacatgaacggcaccatcttttgaaacatagtagatgaggtggcatatcatgaaaagtacccaacg
atttatcacctcagaaaaaagctagttgactcaactgataaagcggacctgaggttaattctacttggctcttgcc
catatgataaagttccgtgggcactttctcattgaggggtgatctaaatccggacaactcggatgtcgacaaact
gttcatccagttagtaaaaacctataatcagttgtttgaagagaacctataaatgcaagtggtggtgagtcga
aggctattcttagcgcccgctctctaaatcccgacggctagaaaacctgatcgacacattaccggagagaa
gaaaaatgggttgttcggttaaccttatagcgctctcactaggcctgacaccaaattttaagtcgaacttcgactt
agctgaagatgccaaattgcagcttagtaaggacacgtacgatgacgatctcgacaattctactggcacaat
ggagatcagtatgcggacttatttttggtgctgcaaaaaccttagcgatgcaatcctcctatctgacatactgaga
gttaatactgagattaccaaggcgccgttatccgcttcaatgatcaaaaggtagatgaacatcaccaagactt
gacacttctcaaggccctagtcgtcagcaactgcctgagaaatataaggaaatattctttgatcagtcgaaaa
acgggtacgcaggttatattgacggcggagcagtgcaagagggaattctacaagtttatcaaaccatattaga
gaagatggatgggacggaagagttgcttgtaaaactcaatcgcaagatctactgcgaagcagcggactttc
gacaacggtagcattccacatcaaattccacttaggcgaattgcatgctatacttagaaggcaggaggatttta
tccgttctctaaagacaatcgtgaaaagattgagaaaatcctaactttcgcataccttactatgtgggacccct
ggcccaggggaactctcggttcgcatggatgacaagaaagtccgaagaaacgattactccatggaattttgag
gaagttgtcgataaagggtgcgtcagctcaatcggttcacgagaggatgaccaactttgacaagaatttaccgaa
cgaaaaagtattgcctaagcacagtttactttacaggtatttcacagtgtagaactcacgaaagttaagta
tgtcactgagggcatgcgtaaacccgcctttctaagcggagaacagaagaaagcaatagtagatctgttattc
aagaccaaccgcaaagtgcagtttaagcaattgaaagaggactactttaagaaaattgaatgcttcgattctgt
cgagatctccggggtagaagatcgatttaatgcgtcacttgggtacgtatcatgacctcctaaagataattaag
ataaggacttctggataacgaagagaatgaagatatcttagaagatatagttgactcttaccctctttgaag
atcgggaaatgattgaggaaagactaaaaacatacgctcacctgttcgacgataagggttatgaaacagttaaa
gaggcgtcgctatacgggctggggacgattgtcgcggaacttatcaacgggataagagacaagcaaatggtg
taaaactattctcgattttctaagagcgacggcttcgccaatagggaactttatgcagctgatccatgatgactc

10

20

30

40

50

ttaaacccttcaaagaggatatacaaaaaggcacaggtttccggacaaggggactcattgcacgaacatattgcg
aatcttgctggttcgccagccatcaaaaagggcatactccagacagtcgaaagtagtgatgagctagttaagg
tcatgggacgtcacaaaccggaaaacattgtaatcgagatggcacgcgaaaatcaaacgactcagaaggggc
aaaaaaacagtcgagagcggatgaagagaatagaagaggggtattaaagaactgggagccagatcttaag
gagcatcctgttgaaaatacccaattgcagaacgagaaactttacctctattacctaataaaatggaagggaca
tgtatgttgatcaggaactggacataaaccgtttatctgattacgacgtcgatcacattgtacccaatcctttt
gaaggacgattcaatcgacaataaagtgccttacacgctcggataagaaccgagggaaaagtgaacatgttcc
aagcagaggaagtgcgtaaagaaaatgaagaactattggcggcagctcctaaatgcgaaactgataacgcaaag
aaagttcgataacttaactaaagtgcgagaggggtggcttgcgtgaacttgacaaggccggatttattaaacgtc
agctcgttggaaccgccaaatcacaaagcatgttgacagatactagattcccgaatgaatacgaaatacga
cgagaacgataagctgattcgggaagtcaaagtaatcactttaaagtcaaaattgggtgctggacttcagaaag
gattttcaattctataaagttaggagataaataactaccacatgcgcacgacgcttatcttaatgccgtcgta
gggaccgcactcattaagaaatacccgaagctagaaagttagtttgcgtatgggtattacaaagtttatgacgt
ccgtaagatgatcgcaaaaagcgaacaggagataggcaaggctacagccaaatacttcttttattctaactta
tgaatttctttaagacggaaatcactctggcaaacggagagatacgcaaacgacctttaattgaaaccaatggg
gagacaggtgaaatcgtatgggataagggccgggacttcgcgacgggtgagaaaagttttgcctatgccccaa
gtcaacatagtaaagaaaactgaggtgcagaccggaggggtttcaaaggaatcgattcttccaaaaaggaata
gtgataagctcatcgctcgtaaaaaggactgggacccgaaaaagtagggtggcttcgatagccctacagttgc
ctattctgtcctagtagtggaagaaagttagaagggaaaatccaagaaactgaagtcagtcgaagaattattg
gggataacgattatggagcgcctcgtcttttgaagaaaccccatcgacttccttgaggcgaaaggttacaaag
aagtaaaaaaggatctcataatataactaccaaagtatagctgtttagtttagaaaatggccgaaaacggatg
ttggctagcgcggagagcttcaaaaagggaacgaactcgcactaccgtctaaatacgtgaatttctgtattt
agcgtcccattacgagaagttgaaaggttcacctgaagataacgaacagaagcaactttttgttgagcagcac
aaacattatctcgacgaaatcatagagcaaatctcgaattcagtaagagagtcctctagctgatgccaatct
ggacaaagtattaagcgcatatacaacagcacaggagataaaccatacgtgagcaggcggaatattatcca
tttgtttactcttaccacactcggcgtccagccgacttcaagtattttgacacaacgatagatcgcaaacgata
cacttctaccaaggaggtgctagacgcgacactgattcaccaatccatcacgggattatatgaaactcggata
gatttgtcacagcttgggggtgactctggtggttctggaggatctggtggttctactaatctgtcagatattattg
aaaaggagaccggtaagcaactggttatccaggaatccatcctcatgctcccagaggaggtggaagaagtca
ttgggaacaagccggaaagcgatatactcgtgcacaccgcctacgacgagagcaccgacgagaatgtcatgc
ttctgactagcgcacgcccctgaatacaagccttgggctctggtcatacaggatagcaacgggtgagaacaagat
taagatgctctctggtggttctggaggatctggtggttctactaatctgtcagatattattgaaaaggagaccgg
taagcaactggttatccaggaatccatcctcatgctcccagaggaggtggaagaagtcattgggaacaagccg
gaaagcgatatactcgtgcacaccgcctacgacgagagcaccgacgagaatgtcatgcttctgactagcgac
gcccctgaatacaagccttgggctctggtcatacaggatagcaacgggtgagaacaagattaagatgctctctg
gtggttctaaaaggacggcgagcggatcagagttcgagagtcgaaaaaaaacgaaaggtcgaataa

【 0 0 9 1 】

コドンをも最適化したBE4核酸配列は、以下に提供される:

atgtcatccgaaaccgggccagtgccgtagaccaacactcaggaggcggatagaaccccatgagtttgaa
gtgttcttcgacccagagagctgcgcaaagagacttgccctctgtatgaaataaattgggggggtcgccattc
aatgggagggcacactagccagaataactaacaacacgtggaggttaaattttatcgagaagtttaccaccgaa
agatacttttgcccaatacacgggtgttcaattacctggtttctgtcatggagtccatgtggagaatgtagtaga
gcgataactgagttcctgtctcgatatacctcacgtcacgttggtttatatacatcgctcggctttatcaccatgcgg
accgcggaaacaggcaaggtcttcgggacctcatatcctctgggggtgacctccagataatgacggagcaag
agagcggatactgctggcgaaactttgttaactacagccaaagcaatgaggcacactggcctagatatccgca
tctctgggttcgactgtatgtccttgaaactgtactgcataattctgggacttcgccatgcttgaaacattctgcgg
cggaacaaccacagctgacctttttcacgattgctctccaaagttgtcactaccagcgattgccaccccat
cttgtgggctactggactcaagtctggaggaagttcaggcggaagcagcgggtctgaaacgcccggaaacctc
agagagcgcaacgcccgaagctctggaggggtcaagtgggtgtagtgataagaaatactccatcggcctcgc
catcgggtacgaattctgtcgggtgggcccgttatcaccgatgagtacaaggtcccttctaagaaattcaaggtttt

10

20

30

40

50

gggcaatacagaccgccattctataaaaaaaaaaacctgatcggcgcccttttgttgacagtggtgagactgctg
aagcgactcgcctgaagcgaactgccaggaggcgggtatacaggcgaaaaaacgaatttgttacctccagg
agattttctcaaatagaatggccaaggtagatgatagttttttaccgcttggaagaaagttttctcgttgagg
aggacaaaaagcacgagaggcacccaatctttggcaacatagtcgatgaggtcgcataccatgagaaatatac
ctacgatctatcatctccgcaagaagctggctgatagcacggataaagctgacctccggctgatctaccttgct
cttgctcacatgattaaattcagggggccatttctgatagaaggagacctcaatcccgacaattctgatgtcgac
aaactgtttattcagctcgttcagacctataatcaactctttgaggagaaccccatcaatgcttcaggggtggac
gcaaaggccattttgtccgcgcgttgagtaatacagcagcctcgagaatttgatagctcaactgccgggtg
agaagaaaaacgggttgtttgggaatctcatagcgttgagtttgggacttacgccaaactttaagtctaactttg
atttgccgaagatgccaaattgcagctgtccaaagatacctatgatgacgacttggataaccttcttgccgag
attgggtgaccaatacgcggatctgtttcttgccgaaaaaatctgtccgacgccatactcttgccgatatactg
cgcgtcaatactgagataactaaggctccctcagcgcgtccatgattaaaagatacgatgagcaccaccaag
atctcactctgttgaaagccctgggttcgcccagcagcttcagagaagtataaggagatattttgcaccaatcta
aaaacggctatcggggttacattgacgggtggcgcctctcaagaagaattctacaagtttataaagccgatactt
gagaaaatggacgggtacagaggaattgttggttaagctcaatcgcgaggacttgttgagaaagcagcgcaca
tttgacaatggtagattccacaccagattcatctgggaggttgcatgccattcttagaagacaagaagatttt
tatccgtttctgaaagataacagagaaaagattgaaaagatacttacctttcgcataccgtattatgtaggtccc
ctggctagaggggaacagtcgcttcgcttggtgactcgaaaaatcagaagaaacaataacccctggaattttg
aagaagtggtagataaagggtgcgagtgcccaatcttttattgagcggatgacaaattttgacaagaatctgcct
aacgaaaagggtgcttccaagcattccctttgtatgaatactttacagtatataatgaactgactaaagtgaag
tacgttaccgaggggatgcgaaagccagcttttctcagtgggcgagcagaaaaagcaatagttgacctgctgt
tcaagacgaataggaaggttaccgtcaaacagctcaaagaagattactttaaaaagatcgaatgttttgattca
gttgagataagcggagtagaggatagatttaacgcaagtcttggaacttatcatgacctttgagatcatcaa
ggataaagattttttggacaacgaggagaatgaagatatcctggaagatatagttacttaccttgacgctttttga
agatcgagagatgatcgaggagcgaacttaagacgtacgcacatctctttgacgataagggttatgaaacaattg
aaacgccggcggtatactggctggggcaggctttctcgaagctgattaatgggtatccgcgataagcagctctg
gaaagacaatccttgactttctgaaaagtgtatggattttgcaaatagaaactttatgcagcttatacatgatgact
ctttgacgttcaaggaagacatccagaaggcacaggtatccggccaaggggatagcctccatgaacacatag
ccaacctggccggctcaccagctattaaaaagggaatattgcaaaccgttaagggttgttgacgaactcgttaag
gttatggggccgacacaaaccagagaaatctgtgattgagatggctaggagaaatcagaccactcaaaaagggt
cagaaaaattctcgcgaaaggatgaagcgaattgaagagggaatcaaagaacttggtctcaaattttgaaag
agcaccggtagaaaaactcagctgcagaatgaaaagctgtatctgtattatctgcagaatggctcgagatat
gtacgttgatcaggagctggatatcaataggctcagtgactacgatgtcgaccacatcgttcctcaatctttcct
gaaagatgactctatcgacaacaaagtgttgacgcgatcagataagaaccgggggaaaatccgacaatgtacc
ctcagaagaagttgtcaagaagatgaaaaactattggagacaattgctgaacgccaagctcataacacaacgc
aagttcgataacttgacgaaagccgaaagagggtgggtgtcagaattggacaaagctggctttattaagcgcc
aattgggtggagaccggcagattacgaaacacgtacacaaattttgattcacgaatgaataccaaatacga
cgaaaacgacaaattgatacgcgaggtgaaagtattacgcttaagagtaagttgggtttcgatttcaggaag
gattttcagttttacaaagtaagagaaataaacaactaccaccacgcccatgatgcttacctcaacgcggtagt
tggcacagctcttatcaaaaaatatccaaagctggaaagcgagttcgtttacgggtgactataaagtatacgacg
ttcggaagatgatagccaaatcagagcaggaaattgggaaggcaaccgcaaaatacttcttctattcaaacat
catgaacttctttaagacggagattacgctcgcgaacggcgaaatacgaagaggccctcatagagactaac
ggcgaaaccggggagatcgtatgggacaaaggacggggactttgcgaccgttagaaaagtactttcaatgcc
caagtgaatattgttaaaaagacagaagtacaaacaggggggttcagtaaggaaatccattttgccaaagcgga
acagtataaattgatagcaaggaaaaaagattgggaccctaagaagtaacggtggttcgactctcctaccgtt
gcatattcagtcctttagtttgcgaaagtggaaaaggggaaaagtaagaagcttaagagtgttaagagcttc
tgggcataaccataatggaacggtctagcttcgagaaaaatccatttgactttctcgaggctaaagggttacaag
gaggtaaaaaaggacctgataattaaactccaaagtacagtccttcgagttggagaatgggaggaagaga
atgttggcatctgcaggggagctccaaaagggggaacgagctggctctgccttcaaaatacgtgaactttctgta
cctggccagccactacgagaaactcaagggttctcctgaggataacgagcagaaacagctgtttgtagagca

10

20

30

40

50

gcacaagcattacctggacgagataattgagcaaattagtgagttctcaaaaagagtaatccttgacagcg
aatctggataaagtctttccgctataataagcaccgggacaagcctatacgagaacaagccgagaacatca
ttcacctctttacccttactaatctgggcgcgcggccgcttcaaatacttcgacaccacgatagacaggaaa
aggtatacgagtaccaaagaagtacttgacgccactctcatccaccagtctataacagggttgtaacgaacga
ggatagatttgtcccagctcggcggcgactcaggagggtcaggcggctccggtggatcaacgaatctttccga
cataatcgagaagaacacggcaaacagttggtgatccaagaatcaatcctgatgctgcctgaagaagtaga
agagggtgattggcaacaaacctgagcttgacattcttgtccacaccgcgtatgacgagagcacggacgagaa
cgttatgcttctcactagcgacgccccctgagtataaaccatgggcgctggtcatccaagattccaatggggaaa
acaagattaagatgcttagtggtgggtctggaggggagcgggtgggtccacgaacctcagcgacattattgaaaa
agagactggtaaacaacttgtaatacaagagtctattctgatgttgcctgaagagggtggaggagggtgattggg
aacaacccggagctctgatatacttgttcataaccgcttatgacgaatctactgatgagaatgtgatgcttttaacg
tcagacgctcccgagtacaaacctgggctctggtgattcaggacagcaatgggtgagaataagattaaaatgt
tgagtgggggctcaaagcgacggctgacggtagcgaatttgagagccccaaaaaaaaaacgaaagggtcgaa
taa

【0092】

別の、コドン最適化BE4核酸配列（ThermoFisher Scientific社のGeneArt）を以下に
提供する：

atgagcagcgagacaggccctgtggctgtggatcctacactgcggagaagaatcgagccccacgagttcgag
gtgttcttcgacccagagagctgcggaaagagacatgcctgctgtacgagatcaactggggcggcagacac
tctatctggcggcacacaagccagaacaccaacaagcacgtggaagtgaactttatcgagaagtttacgaccg
agcggctacttctgccccaacaccagatgcagcatcacctggtttctgagctggtcccccttgcggcgagtgcagc
agagccatcacccgagtttctgtccagatatccccacgtgacctgttcatctatatcgccggctgtaccaccac
gccgatcctagaaatagacagggactgcgcgacctgatcagcagcggagtgacctccagatcatgaccgag
caagagagcggctactgctggcggaaacttctgtaactacagccccagcaacgaagccactggcctagatat
cctcacctgtgggtccgactgtacgtgctggaactgtactgcatcatcctgggcctgctccatgctgaacat
cctgagaagaagcagcctcagctgaccttcttcacaatcgccctgcagagctgccactaccagagactgcct
ccacacatcctgtggggccaccggacttaagagcggaggatctagcggcggctctagcggatctgagacacct
ggcacaagcgagctctgccacacctgagagtagcggcggatcttctggcggctccgacaagaagtactctatc
ggactggccatcggcaccaactctgttggatgggcccgtgatcacccagcagtagaagggtgccagcaagaaa
ttcaagggtgtcgggaacaccgacccggcacagcatcaagaagaatctgatcggcgccctgctgttcgactctg
gcgaaacagccgaagccaccagactgaagagaaccgccaggcggagatacacccggcggagaaccggat
ctgctacctgcaagagatcttcagcaacgagatggccaagggtggacgacagcttcttcacagactggaaga
gtccttctctggtggaagaggacaagaagcacgagcggcaccccatcttcggcaacatcgtggatgaggtggc
ctaccacgagaagtacccaccatctaccacctgagaaagaactggtggacagcaccgacaaggccgacct
gagactgatctacctggctctggccacatgatcaagttccggggccactttctgatcagggcgatctgaacc
ccgacaacagcgacgtggacaagctgttcatccagctggtgcagacctacaaccagctgttcgaggaaaacc
ccatcaacgcctctggcgtggacgccaaggctatcctgtctgccagactgagcaagagcagaaggctggaaa
acctgatcggccagctgcctggcgagaagaagaatggcctgttcggcaacctgattgccctgagcctgggact
gaccttaacttcaagagcaacttcgacctggccgaggatgccaactgcagctgagcaaggacacctacga
cgacgacctggacaatctgctggccagatcgggcgatcagtagccgacttgtttctggccgccaagaacctg
tccgacgccatcctgtgagcgatatcctgagagtgaacaccgagatcacaaaggccccctctgagcgcctcta
tgatcaagagatacgacgagcaccaccaggatctgacctgtgaaggccctcgtagacagcagctgccag
agaagtacaaagagattttctcgatcagtccaagaacggctacgccggctacattgatggcggagccagcca
agaggaattctacaagttcatcaagcccatcctggaaaagatggacggcaccgaggaactgctggtcaagct
gaacagagaggacctgctgcggaagcagcggaccttcgacaatggctctatccctaccagatccacctggg
agagctgcacgccattctgcggagacaagaggacttttaccattcctgaaggacaaccgggaaaagatcga
gaagatcctgacctcaggatcccctactacgtgggaccactggccagaggcaatagcagattcgcttgatg
accagaaagagcggaggaaaccatcacacctggaaacttcgagggaagtgggtggacaagggcgccagcgctca
gtccttcatcgagcggatgaccaacttcgataagaacctgcctaacgagaagggtgctgccaagcactccctg
ctgtatgagtacttcaccgtgtacaacgagctgaccaagtgaatacgtgaccgagggaatgagaaagccc

10

20

30

40

50

gcctttctgagcggcgagcagaaaaaggccattgtggatctgctgttcaagaccaaccggaaagtaccgtga
 agcagctgaaagaggactacttcaagaaaatcgagtgttcgacagcgtggaaatcagcggcgtggaagatc
 ggttcaatgccagcctgggcacataccacgacctgctgaaaattatcaaggacaaggacttcttgacaacga
 agagaacgaggacatttctcgaggacatcgtgctgacctgacactgtttgaggacagagagatgatcgagga
 acggctgaaaacatacggccacctgttcgacgacaaaagtgatgaagcaactgaagcggaggcggtacacag
 gctggggcagactgttctcggaagctgatcaacggcatccgggataagcagtcgggcaagacaatcctggattt
 cctgaagtccgacggcttcgccaacagaaacttcatgcagctgatccacgacgacagcctgacctttaagag
 gacatccagaaaagcccagggtgtccggccaaggcgatttctgcacgagcacattgccaacctggccggatctc
 ccgccattaagaagggcacacctgcagacagtgaagggtgggtggacgagcttgtgaaagtgatgggcagacaca
 agcccgagaacatcgtgatcgaaatggccagagagaaccagaccacacagaaggggccagaagaacagccg
 cgagagaatgaagcggatcgaagagggcacaaagagctgggcagccagatcctgaaagaacaccccgtgg
 aaaacacccagctgcagaacgagaagctgtacctgtactacctgcagaatggacgggatatgtacgtggacc
 aagagctggacatcaaccggctgagcgactacgatgtggaccatatcgtgccccagagctttctgaaggacg
 actccatcgataacaaggctcctgaccagaagcgacaagaaccggggcaagagcgataacgtgccctccgaa
 gaggtggtcaagaagatgaagaactactggcgacagctgctgaacccaagctgattaccagcgggaagttc
 gataacctgaccaaggccgagagaggcgccctgagcgaacttgataaggccggcttcattaagcggcagctg
 gtggaaacccggcagatcaccaaacacgtggcacagattctggactcccgatgaacactaagtacgacgag
 aatgacaagctgatccgggaagtgaagtcatcacctgaagtctaagctggtgtccgatttccggaaggattt
 ccagttctacaaagtgcgggaaatcaacaactaccatcacgccacgacgcctacctgaatgccgttgttggaa
 acagccctgatcaagaagtatcccaagctggaaagcgagttcgtgtacggcgactacaaggtgtacgacgtg
 cggaagatgatcgccaagagcgaacaagagatcggcaaggctaccgccaagtacttttctacagcaacatc
 atgaactttttcaagacagagatcacctggccaacggcgagatccggaaaagacccctgatcgagacaaac
 ggcgaaacccggggagatcgtgtgggataaggggcagagattttgccacagtgcggaaagtgtgagcatgcc
 caagtgaatatcgtgaagaaaaccgaggtgcagacaggcggcttcagcaaagagtctatcctgcctaagcgg
 aacagcgataagctgatcgccagaaagaaggactgggaccctaagaagtacggcggcttcgatagccctacc
 gtggcctattctgtgctggtggtggccaaagtggaaaagggaagtccaaaaagtcaagagcgtgaaagag
 ctgctggggatcacatcatggaaagaagcagctttgagaagaacccgatcgactttctggaagccaagggt
 acaaagaagtcaagaaggacctcatcatcaagctcccaagtacagcctgttcgagctggaaaatggccgga
 agcggatgctggcctcagcaggcgaactgcagaaaggcaatgaactggccctgcctagcaaatacgtcaact
 tcctgtacctggccagccactatgagaagctgaagggcagccccgaggacaatgagcaaaagcagctgtttg
 tggaacagcacaagcactacctggacgagatcatcgagcagatcagcgagtttccaagagagtgtcctgg
 ccgacgctaacctggataagggtgctgtctgcctataacaagcaccgggacaagcctatcagagagcaggccg
 agaatatcatccacctgtttacctgaccaacctgggagccccctgccgccttcaagtacttcgacaccaccatc
 gaccggaagagggtacaccagcaccaaaagaggtgctggacgccacactgatccaccagtctatcacccggcctg
 tacgaaacccggatcgacctgtctcagctcggcggcgattctgggtggttctggcggaagtggcggtaccacca
 atctgagcgacatcatgaaaaagagacaggcaagcagctcgtgatccaagaatccatcctgatgtgcctga
 agaggttgaggaagtgatcggcaacaagcctgagtcgcgacatcctgggtgcacaccgcctacgatgagagcac
 cgatgagaacgtcatgctgctgacaagcgacgcccctgagtacaagccttgggctctcgtgattcaggacagc
 aatgggggagaacaagatcaagatgctgagcggaggttagcggaggcagtgggcggaagcacaacacctgtctga
 tatcattgaaaaaagaaaccgggaagcaactggtcattcaagagtccattctcatgtctccggaagaagtgcgag
 gaagtattgaaacaaacccgagagcgatattctggtccacacagcctatgacgagtctacagacgaaaac
 gtgatgtcctgacctctgacgctcccagagtataagccctgggcacttgttatccaggactctaacggggaaaa
 caaatcaaatgttgtccggcggcagcaagcggacagccgatggatctgagttcgagagccccaagaagaa
 acggaagggtggagtaa,

【0093】

「塩基編集活性」とは、ポリヌクレオチド内の塩基を化学的に改変するように作用することを意味する。一実施形態では、第1の塩基が第2の塩基に変換される。一実施形態では、塩基編集活性は、例えば標的C・GをT・Aに変換する、シチジンデアミナーゼ活性である。別の実施形態では、塩基編集活性は、例えばA・TをG・Cに変換する、アデノシンデアミナーゼ活性である。

10

20

30

40

50

【0094】

用語「塩基エディターシステム」は、標的ヌクレオチド配列の核酸塩基を編集するための系を指す。様々な実施形態において、塩基エディター (BE) システムは、(1)ポリヌクレオチドプログラミング可能なヌクレオチド結合ドメインと、核酸塩基を脱アミノ化するためのデアミナーゼドメイン、および(2)ポリヌクレオチドプログラム可能なヌクレオチド結合ドメインに伴うガイドポリヌクレオチド(例えばガイドRNA)を含む。いくつかの実施形態において、塩基エディターシステムは、(1)ポリヌクレオチドプログラミング可能なDNA結合ドメインと、核酸塩基を脱アミノ化するためのデアミナーゼドメインとを含む塩基エディター (BE)、および(2)ポリヌクレオチドプログラミング可能なDNA結合ドメインに伴うガイドRNAを含む。ある態様において、ポリヌクレオチドプログラミング可能なヌクレオチド結合ドメインは、ポリヌクレオチドプログラミング可能なDNA結合ドメインである。いくつかの実施形態では、塩基編集はシチジン塩基エディター (CBE) である。ある態様において、塩基エディターは、アデニンまたはアデノシン塩基エディター (ABE) である。

10

【0095】

「β-グロビン (HBB) タンパク質」とは、NCBI受付番号NP_000509のアミノ酸配列と少なくとも約95%のアミノ酸配列同一性を有するポリペプチドまたはその断片を意味する。特定の実施形態において、β-グロビタンパク質は、以下の参照配列に対して1以上の改変を含む。特定の一実施形態において、鎌状赤血球症に関連するβ-グロビタンパク質は、E6V (E7Vとも呼ばれる)突然変異を含む。例示的なβ-グロビンアミノ酸配列(例えば、参照配列)を以下に提供する。

20

```
1 mvhltpeeks avtalwgkvn vdevggealg rllvvypwtq rffesfgdls tpdavmgnpk
61 vkahgkklvg afsdglahld nlkgtfatls elhcdklhvd penfrllgnv lvcvlahhfg
121 keftppvqaa yqkvvagvan alahkyh
```

【0096】

「HBBポリヌクレオチド」とは、β-グロビタンパク質をコードする核酸分子またはその断片を意味する。NCBI受付番号NM_000518で入手可能な例示的HBBポリヌクレオチドの配列は、以下に提供される：

```
1 acatttgctt ctgacacaac tgtgttcact agcaacctca aacagacacc atggtgcatc
61 tgactcctga ggagaagtct gccgttactg ccctgtgggg caaggtgaac gtggatgaag
121 ttggtggtga ggccctgggc aggctgctgg tggctaccc ttggaccag aggttctttg
181 agtcctttgg ggatctgtcc actcctgatg ctgttatggg caaccctaag gtgaaggctc
241 atggcaagaa agtgctcggg gccttagtg atggcctggc tcacctggac aacctcaagg
301 gcacctttgc cacactgagt gagctgcact gtgacaagct gcacgtggat cctgagaact
361 tcaggctcct gggcaacgtg ctggtctgtg tgctggccca tcactttggc aaagaattca
421 cccaccagt gcaggctgcc tatcagaaag tgggtggctgg tgtggctaata gccctggccc
481 acaagtatca ctaagctgc tttcttgctg tccaatttct attaaagggt ctttgttcc
541 ctaagtccaa ctactaaact gggggatatt atgaagggcc ttgagcatct ggattctgcc
601 taataaaaaa catttatttt cattgcaa
```

30

【0097】

「HBG1タンパク質」、すなわちホモ・サピエンスヘモグロビン・サブユニットγ1 (HBG1) タンパク質」とは、NCBI参照配列番号NM_000559.2のアミノ酸配列と少なくとも約95%のアミノ酸配列同一性を有するポリペプチドまたはその断片を意味する。いくつかの実施形態において、HBG1タンパク質は、以下のアミノ酸配列に対して1つ以上の改変を含み得る。特定の実施形態では、本明細書に記載されるように、鎌状赤血球症を治療または改善するために、HBG1タンパク質に関連する調節領域、例えばプロモーターに対して編集が行われる。例示的なHBG1アミノ酸配列は、以下に提供される：

40

```
MGHFTEEDKATITSLWGKVNVEDAGGETLGRLLVVPWTQRFFDSFGNLSSASAIMGNP
KVKAHGKKVLTSLGDATKHLDDLKGTFAQLSELHCDKLHVDPENFKLLGNVLVTVLAIH
FGKEFTPEVQASWQKMVTAVASALSSRYH。
```

50

【0098】

「HBG1ポリヌクレオチド」とは、HBG1タンパク質またはその断片をコードする核酸分子を意味する。例示的なHBG1ポリヌクレオチドの核酸配列は、以下に提供される：

```

1 acactcgctt ctggaacgtc tgaggttatc aataagctcc tagtccagac gcc
tggttc
61 atttcacaga ggaggacaag gctactatca caagcctgtg gggcaagggtg aat
gtggaag
121 atgctggagg agaaaccctg ggaaggctcc tggttgtcta cccatggacc caga
ggttct
181 ttgacagctt tggcaacctg tcctctgcct ctgccatcat gggcaacccc aaagtc
aagg
241 cacatggcaa gaagggtgtg acttccttgg gagatgccac aaagcacctg gatg
atctca
301 agggcacctt tgcccagctg agtgaactgc actgtgacaa gctgcatgtg gatcc
tgaga
361 acttcaagct cctgggaaat gtgctggtga ccgttttggc aatccatttc ggcaaa
gaat
421 tcaccctga ggtgcaggct tcctggcaga agatggtgac tgcagtggcc agtgc
cctgt
481 cctccagata ccactgagct cactgcccac gattcagagc tttcaaggat aggct
ttatt
541 ctgcaagcaa tacaataat aaatctattc tgctgagaga tcac

```

10

20

【0099】

「HBG2タンパク質」、すなわち、ホモ・サピエンスヘモグロビン・サブユニットγ2 (HBG2) タンパク質とは、NCBI参照配列番号NM_000184.3のアミノ酸配列に対して少なくとも約95%のアミノ酸配列同一性を有するポリペプチドまたはその断片を意味する。いくつかの実施形態において、HBG2タンパク質は、以下のアミノ酸配列に対して1以上の改変を含み得る。特定の実施形態では、本明細書に記載されるように、鎌状赤血球症を治療または改善するために、HBG2タンパク質に関連する調節領域、例えばプロモーターに対して編集が行われる。例示的なHBG2アミノ酸配列は、以下に提供される：

30

```

MGHFTEEDKATITSLWGKVNVEDAGGETLGRLLVVPWTQRFDFSFGNLSSASAIMGNP
KVKAHGKKVLTSLGDAIKHLDDLKGTFAQLSELHCDKLHVDPENFKLLGNVLVTVLAIHF
GKEFTPEVQASWQKMVTGVASALSSRYH

```

【0100】

「HBG2ポリヌクレオチド」とは、HBG2タンパク質またはその断片をコードする核酸分子を意味する。例示的なHBG2ポリヌクレオチドの核酸配列は、以下に提供される：

```

1 acactcgctt ctggaacgtc tgaggttatc aataagctcc tagtccagac gccatgggtc
61 atttcacaga ggaggacaag gctactatca caagcctgtg gggcaagggtg aatgtggaag
121 atgctggagg agaaaccctg ggaaggctcc tggttgtcta cccatggacc cagaggttct
181 ttgacagctt tggcaacctg tcctctgcct ctgccatcat gggcaacccc aaagtcaagg
241 cacatggcaa gaagggtgtg acttccttgg gagatgccat aaagcacctg gatgatctca
301 agggcacctt tgcccagctg agtgaactgc actgtgacaa gctgcatgtg gatcctgaga
361 acttcaagct cctgggaaat gtgctggtga ccgttttggc aatccatttc ggcaagaat
421 tcaccctga ggtgcaggct tcctggcaga agatggtgac tggagtggcc agtgccctgt
481 cctccagata ccactgagct cactgcccac gatgcagagc tttcaaggat aggctttatt
541 ctgcaagcaa tacaataata aatctattct gctaagagat cacaca

```

40

【0101】

「ALAS1タンパク質」、すなわちホモ・サピエンス5'-アミノレブリン酸シンターゼ1 (ALAS1) タンパク質とは、NCBI参照配列番号NM_000688.6のアミノ酸配列と少なくとも約95%のアミノ酸配列同一性を有するポリペプチドまたはその断片を意味する。ある態

50

様において、ALAS1タンパク質は、以下のアミノ酸配列に対して1以上の改変を含み得る。特定の実施形態では、本明細書に記載されるように、鎌状赤血球症を治療または改善するために、ALAS1タンパク質と関連する調節領域、例えばプロモーターに対して編集が行われる。例示的なALAS1アミノ酸配列は、以下に提供される:

```
MESVVRRCFPLSRVPQAFLLQKAGKSLFLFYAQNCPKMMMEVGAKPAPRALSTAA
VHYQQIKETPPASEKDKTAKAKVQQTDPGSGQSPDGTQLPSGHPLPATSQGTG
SKCPFLAAQMNQRGSSVFCASLELQEDVQEMNAVRKEVAETSAGPSVSVSK
TDGGDPSGLLKNFQDIMQKQRPERSHLLQDNLPKSVSTFQYDRFFEKKIDEKK
NDHTYRVFKTVNRRAHIFPMADDYSDSLITKKQVSVWCSNDYLGMSRHRPRVCG
AVMDTLKQHGAGAGGTRNISGTSKFHVDLERELADLHGKDAALLFSSCFVAND
STLFTLAKMMPGCEIYSDSGNHASMIQGIRNSRVPKYIFRHNDVSHLRELLQRS
PSVPKIVAFETVHSMMDGAVCPLEELCDVAHEFGAITFVDEVHAVGLYGARGGGI
GDRDGVMPKMDIISGTLGKAFCVGGYIASTSSLIDTVRSYAAGFIFTTSLPPML
LAGALESVRILKSAEGRVLRQRHQRNVKLMRQMLMDAGLPVVHCPSHIIPVRV
ADAAKNTEVCDLMSRHNIYVQAINYPTVPRGEELLRIAPTPHHTPQMMNYFLE
NLLVTWKQVGLLELKPSSAECNFCRRPLHFVMSEREKSYFSGLSKLVSAQA
【0102】
```

「ALAS1ポリヌクレオチド」とは、ALAS1タンパク質またはその断片をコードする核酸分子を意味する。例示的なALAS1ポリヌクレオチドの核酸配列は、以下に提供される:

```
aggctgctcc cggacaaggg caacgagcgt ttcgtttgga cttctcgact tgagtccccg cctccttcgc
cgccgcctct gcagtcctca gcgcagttat gccagttct tcccgctgtg gggacacgac cacggagga
a tcttgcttc agggactcgg gaccctgctg gacccttcc tcgggttag gggatgtggg gaccagg
aga aagtcaggat ccctaagagt cttccctgcc tggatggatg agtggcttct tctccacctt gattctt
tcc acaggagcca gcatacttcc tgaacatgga gagtgttgtt
cgccgctgcc cattcttacc ccgagtcgcc caggccttcc tgcagaaagc aggcacatct ctgttgttct
atgccccaaa ctgccccaa atgatggaag ttggggccaa gccagcccct cgggcattgt ccactgca
gc agtacactac caacagatca aagaaacccc tccggccagt gagaaagaca aaactgctaa ggcca
aggct caacagactc ctgatggatc ccagcagagt ccagatggca cacagcttcc gtctggacac ccc
ttgcctg ccacaagcca gggcactgca agcaaatgcc ctttcttggc agcacagatg aatcagagag g
cagcagtggt cttctgcaaa gccagtcctg agcttcagga ggatgtgcag gaaatgaatg ccgtgaggaa
agagggtgct gaaacctcag caggccccag tgtggttagt gtgaaaaccg atggagggga tccagtg
ga ctgctgaaga acttccagga catcatgcaa aagcaaagac cagaaagagt gtctcatctt cttcaa
gata acttgccaaa atctgttcc acttttcagt atgatcggtt ctttgagaaa aaaattgatg agaaaaa
gaa tgaccacacc tatcgagttt taaaactgt gaaccggcga gcacacatct tcccatggc agatga
ctat tcagactccc tcatcaccaa aaagcaagtg tcagtctggt gcagtaatga ctacctagga atgag
tcgcc acccacgggt gtgtggggca gttatggaca ctttgaaaca acatggtgct ggggcagggt gta
ctagaaa ttttctgga actagtaaat tcatgttgga ctagagcgg gagctggcag acctccatgg ga
aagatgcc gcactctgt tttctcgtg cttgtggcc aatgactcaa cctcttcac cctggctaag atg
atgccag
gctgtgagat ttactctgat tctgggaacc atgcctccat gatccaaggg attcgaaaca gccgagtgcc
aaagtacatc ttccgccaca atgatgtcag ccacctcaga gaactgctgc aaagatctga cccctcagtc
ccaagattg tggcatttga aactgtccat tcaatggatg gggcggtgtg cccactggaa gagctgtgtg
atgtggccca tgagtttggg gcaatcacct tcgtggatga ggtccacgca gtggggcttt atggggctcg
aggcgagggt attggggatc gggatggagt catgcaaaa atggacatca tttctggaac acttgga
aa gcctttggtt gtgttgagg gtacatcgcc agcacgagtt ctctgattga caccgtacgg tcctatgc
tg ctggcttcat cttcaccacc tctctgccac ccatgctgct ggctggagcc ctggagtctg tgcggatc
ct gaagagcgct gagggacggg tgcttcgccg ccagcaccag cgcaacgtca aactcatgag acaga
tgcta atggatgccg gcctccctgt tgtccactgc cccagccaca tcatccctgt gcgggttgca gatg
ctgcta aaaacacaga agtctgtgat gaactaatga gcagacataa catctacgtg caagcaatca att
accttac ggtgccccgg ggagaagagc tcctacggat tgccccacc cctcaccaca cccccagat
```

10

20

30

40

50

gatgaactac ttccttgaga atctgctagt cacatggaag caagtggggc tggaactgaa gcctcattcc
tcagctgagt gcaacttctg caggaggcca ctgcattttg aagtgatgag tgaaagagag aagtcctatt
tctcaggctt gagcaagttg gtatctgctc aggcctgagc atgacctcaa ttatttcact
taaccccgagg ccattatcat atccagatgg tcttcagagt tgtctttata tgtgaattaa gttatattaa at
tttaatct atagtaaaaa catagtcctg gaaataaatt cttgcttaaa tgggtg

【0103】

「BCL11A」タンパク質、すなわち、ホモ・サピエンスB細胞CLL/リンパ腫11A (BCL11A) タンパク質 (ジンクフィンガータンパク質) は、GenBank受託番号ADL_14508.1のアミノ酸配列と少なくとも約95%のアミノ酸配列同一性を有するポリペプチドまたはその断片を意味する。いくつかの実施形態において、BCL11Aタンパク質は、以下のアミノ酸配列に対して1つ以上の改変を含み得る。特定の実施形態において、塩基編集は、BCL11Aタンパク質の、またはBCL11Aタンパク質と関連する、調節領域、例えばプロモーターにおいて起こり、例えば胎児ヘモグロビン産生を増加させることによって、ベータサラセミアおよび鎌状赤血球症 (SCD) などの疾患を治療または改善する。BCL11Aをコードする遺伝子はいくつかの造血系統で高度に発現し、胎児から成体への赤血球産生への移行の際の γ -グロビン発現から β -グロビン発現へのスイッチにおいて役割を果たしている。BCL11Aは胎児ヘモグロビン産生の抑制に役割を果たし得る。リンパ腫の病変形成にも関与している可能性がある；B細胞悪性病変に関連する転座がBCL11Aの発現を脱調節することが見出されている。ヒトBCL11Aのアミノ酸配列の例を以下に示す。

【0104】

MSRRKQGKPKQHLSKREFSPEPLEAILTDDEPDHGPLGAPEGDHDLLTCGQCQMNFPPLGD
ILIFIEHKRKQCNGSLCLEKAVDKPPSPSPSIEMKKASNPVEVGIQVTPEDDDCLSTSSRGIC
PKQEHIAADKLLHWRGLSSPRSAHGALIPTPGMSAEYAPQGICKDEPSSYTCTTCKQPFTS
AWFLLQHAQNTHGLRIYLESEHGSPLTPRVGIPSGLGAECPSQPPLHGIHIADNNPFNLL
RIPGSVSREASGLAEGFRPPTPPLFSPPPRHLDPHRIERLGAEEALATHHPSAFDRVLR
LNPMAMEPPAMDFSRRLRELANTSSPPLSPGRPSPMQRLLQPFQPGSKPPFLATPPLP
PLQSAPPPSQPPVKSKSCEFCGKTFKFQSNLVVHRRSHTGEKPYKCNLCDHACTQASKLK
RHMKTHMHKSSPMTVKSDDGLSTASSPEPGTSDLVGSASSALKSVVAKFKSENDPNLIPE
NGDEEEEEEDDEEEEEEEEEEEEEELTESERVVDYGFGLSLEAARHHENSSRGAVVGVGDESRA
LPDVMQGMVLSSMQHFSEAFHQVLGEKHKRGLAEAEHRDTCDEDSVAGESDRIDG
TVNGRGCSGESASGGLSKLLLGSPLSPFSKRIKLEKEFDLPPAAMPNTENVYSQWLA
GYAASRQLKDPFLSFGDSRQSPFASSEHSSENGSLRFSTPPGELDGGISGRSGTGSGGSTP
HISGPGPGRPSSKEGRRSDTCEYCGKVFKNCSNLTVHRRSHTGERPYKCELCNYACAQSS
KLTRHMKTHGQVGKDVKCEICKMPFSVYSTLEKHMKKWHSDRVLNNDIKTE

【0105】

「BCL11Aポリヌクレオチド」とは、BCL11Aタンパク質またはその断片をコードする核酸分子を意味する。例示的なヒトBCL11A (アイソフォーム1) ポリヌクレオチド、参照配列番号GU 324937.1の核酸配列は、以下に提供される：

atgtctcgccgcaagcaaggcaaaccaggcacttaagcaaacgggaattctcgcccgagcctcttgaagcca
ttcttacagatgatgaaccagaccacggcccgttgggagctccagaaggggatcatgacctcctcacctgtgg
gcagtgccagatgaacttccattgggggacatt cttatttttatcgagcacaacggaaacaatgcaatggc
agcctctgcttagaaaaagctgtggataagccaccttcccttcaccaatcgagatgaaaaagcatccaatcc
cgtggag gttggcatccaggtcacgccagaggatgacgattgtttatcaacgtcatctagaggaatt tgccc
caaacaggaacacatagcagataaacttctgcactggaggggcctctcctccctcgttctgcacatggagct
ctaattccccacgcctgggatgagtgagaatatgccccgcag ggtatttgtaaagatgagcccagcagctac
acatgtacaacttgcaaacagccattcacc agtgcattggtttcttgaacacgcacagaacactcatggatt
aagaatctacttagaaagcgaacacggaagtccctgaccccgcggttggtatcccttcaggactaggtgca
gaa tgtccttcccagccaccttccatgggattcatattgcagacaataacccctttaacctg ctaagaatac
caggatcagtatcgagagaggcttcggcctggcagaaggcgcttccaccactccccctgttagtcc
accaccgagacatcattggacccccaccgcatagagcgcctggggcggaagagatggccctggccaccc

atcacccgagtgcccttgacaggggtgctgcggttgaatccaatggctatggagcctcccgccatggatttctct
aggagacttagagagctggcaggggaacacgtctagcccaccgctgtccccaggccggccagccctatgcaa
aggttactgcaaccattccagccaggtagcaagccgccccttctggcgacgccccccctcctcctctgcaatc
cgccccctcctccctccagcccccggtcaagtccaagtcatgaggttctgcggaagacgttcaaatttcaga
gcaacctggtggtgcaccggcgagccacacgggcgagaagccctacaagtgaacctgtgacaccacgcg
tgcaaccagggcagcaagctgaagcgccacatgaagacgcacatgcacaaatcgtcccccatgacggtaag
tccgacgacggttctccaccgccagctccccggaacccggcaccagcgacttggtgggcagcgccagcagc
gcgctcaagtccgtggtggccaagttcaagagcgagaacgaccccaacctgatcccgagaaacggggacga
ggaggaagaggaggagcagcaggaagaggaagaagaggaggaagaggaggaggagctgacggaga
gcgagagggtggactacggcttcgggctgagcctggaggcgggcgccaccacgagaacagctcgcgggg
cgcggtcgtgggctgggacgagagccgcgcccctgccgacgtcatgcaggcatggtgctcagctccat
gcagcacttcagcagggccttccaccaggtcctgggcgagaagcataagcgcgccacctggccgaggccg
agggccacagggacacttgcgacgaagactcggtggccggcgagtcggaccgcatagacgatggcactgtt
aatggccgcggtgctccccgggagtcggcctcggggggctgtccaaaagctgctgctgggcagcccc
agctcgtgagcccccttcttaagcgcatcaagctcgagaaggagttcgacctgccccggcgcgatgcca
acacggagaacgtgtactcgagtggtcgcgggctacgcggcctccaggcagctcaaagatcccttccttag
cttcggagactccagacaatcgcccttgcctcctcgtcggagcactcctcgagaacgggagcttgcgcttct
ccacaccgcccggggagctggacggaggatctcggggcgagcggcacgggaagtggaggggagcagcc
ccatattagtgtccggggccgggagggccagctcaaaagagggcagacgcagcagcacttgtagtactg
tgggaaagtcttcaagaactgtagcaatctcactgtccacaggagaagccacacgggcgaaaggccttataaa
tgcgagctgtgcaactatgcctgtgccagagtagcaagctcaccaggcacatgaaaacgatggccaggtg
gggaaggacgtttacaaatgtgaaatttgtaagatgccttttagcgtgtacagtaccctggagaaacacatgaa
aaaatggcacagtgatcgagtggtgaataatgatataaaaaactgaatag。

【0106】

一部の実施形態では、核酸塩基エディターシステムは、複数の塩基編集コンポーネントを含むことができる。例えば、核酸塩基エディターシステムは、複数のデアミナーゼを含むことができる。ある態様において、ヌクレアーゼ塩基編集システムは、1以上のシチジンデアミナーゼおよび/または1以上のアデノシンデアミナーゼを含み得る。いくつかの実施形態において、単一のガイドポリヌクレオチドを利用して、異なるデアミナーゼを標的核酸配列にターゲティングすることができる。ある態様において、一対のガイドポリヌクレオチドが、異なるデアミナーゼを標的核酸配列にターゲティングするために利用され得る。

【0107】

塩基エディター系の核酸塩基成分およびポリヌクレオチドプログラミング可能なヌクレオチド結合成分は、互いに共有結合的または非共有結合的により結合され得る。例えば、いくつかの実施形態において、デアミナーゼドメインが、ポリヌクレオチドプログラミング可能なヌクレオチド結合ドメインによって標的ヌクレオチド配列にターゲティングされ得る。ある態様において、ポリヌクレオチドプログラミング可能なヌクレオチド結合ドメインは、デアミナーゼドメインに融合または連結され得る。いくつかの実施形態において、ポリヌクレオチドプログラミング可能なヌクレオチド結合ドメインは、デアミナーゼドメインと非共有結合的に相互作用または結合することによって、デアミナーゼドメインを標的ヌクレオチド配列にターゲティングすることができる。例えば、いくつかの実施形態において、核酸塩基編集成分、例えばデアミナーゼ成分は、ポリヌクレオチドプログラミング可能なヌクレオチド結合ドメインの一部をなすさらなる異種部分またはドメインと相互作用し、会合し、または複合体を形成することができる、さらなる異種部分またはドメインを含むことができる。いくつかの実施形態において、追加の異種部分は、ポリペプチドと結合し、相互作用し、会合し、または複合体を形成することができる。いくつかの実施形態において、追加の異種部分は、ガイドポリヌクレオチドに結合することができる。いくつかの実施形態において、追加

10

20

30

40

50

の異種部分は、ポリペプチドリinkerに結合することができる。いくつかの実施形態において、追加の異種部分は、ポリヌクレオチドリinkerに結合することができる。追加の異種部分は、タンパク質ドメインであってもよい。いくつかの実施形態において、追加の異種部分は、K相同 (KH) ドメイン、MS2コートタンパク質ドメイン、PP7コートタンパク質ドメイン、SfMu Comコートタンパク質ドメイン、ステリル α モチーフ、テロメラーゼKu結合モチーフおよびKuタンパク質、テロメラーゼSm7結合モチーフおよびSm7タンパク質、またはRNA認識モチーフであり得る。

【0108】

塩基編集システムは、ガイドポリヌクレオチド成分をさらに含むことができる。塩基編集システムの構成要素は、共有結合、非共有結合的相互作用、またはそれらの結合および相互作用の任意の組合せを介して互いに結合され得ることが理解されるべきである。ある状態において、デアミナーゼドメインは、ガイドポリヌクレオチドによって標的ヌクレオチド配列にターゲティングされ得る。例えば、いくつかの実施形態において、塩基エディター系の核酸塩基編集成分、例えばデアミナーゼ成分は、ガイドポリヌクレオチドの一部またはセグメント(例えばポリヌクレオチドモチーフ)と相互作用し、結合し、または複合体を形成し得るさらなる異種部分またはドメイン(例えばRNAまたはDNA結合タンパク質のようなポリヌクレオチド結合ドメイン)を含み得る。いくつかの実施形態において、追加の異種部分またはドメイン(例えばRNAまたはDNA結合タンパク質などのポリヌクレオチド結合ドメイン)は、デアミナーゼドメインに融合または連結され得る。いくつかの実施形態において、追加の異種部分は、ポリペプチドと結合し、相互作用し、会合し、またはポリペプチドと複合体を形成することができる。いくつかの実施形態において、追加の異種部分は、ポリヌクレオチドと結合し、相互作用し、会合し、またはポリヌクレオチドと複合体を形成することができる。いくつかの実施形態において、追加の異種部分は、ガイドポリヌクレオチドに結合することができる。いくつかの実施形態において、追加の異種部分は、ポリペプチドリinkerに結合することができる。いくつかの実施形態において、追加の異種部分は、ポリヌクレオチドリinkerに結合することができる。追加の異種部分は、タンパク質ドメインであってもよい。いくつかの実施形態において、追加の異種部分は、K相同 (KH) ドメイン、MS2コートタンパク質ドメイン、PP7コートタンパク質ドメイン、SfMu Comコートタンパク質ドメイン、無菌アルファモチーフ、テロメラーゼKu結合モチーフおよびKuタンパク質、テロメラーゼSm7結合モチーフおよびSm7タンパク質、またはRNA認識モチーフであり得る。

【0109】

ある状態において、塩基編集システムは、塩基除去修復 (BER) コンポーネントの阻害因子をさらに含むことができる。塩基編集システムの構成要素は、共有結合、非共有結合的相互作用、またはそれらの結合および相互作用の任意の組み合わせを介して互いに結合され得ることが理解されるべきである。BER成分の阻害因子は、塩基除去修復阻害因子を含み得る。ある状態において、塩基除去修復の阻害因子は、ウラシルDNAグリコシラーゼ阻害因子 (UGI) であり得る。ある状態において、塩基除去修復の阻害因子は、イノシン塩基除去修復阻害因子であり得る。ある状態において、塩基除去修復の阻害因子は、ポリヌクレオチドプログラミング可能なヌクレオチド結合ドメインにより標的ヌクレオチド配列にターゲティングされ得る。ある状態において、ポリヌクレオチドプログラミング可能なヌクレオチド結合ドメインは、塩基除去修復の阻害因子に融合または連結され得る。ある状態において、ポリヌクレオチドプログラミング可能なヌクレオチド結合ドメインは、デアミナーゼドメインおよび塩基除去修復の阻害因子に融合または連結され得る。いくつかの実施形態において、ポリヌクレオチドプログラミング可能なヌクレオチド結合ドメインは、塩基除去修復の阻害因子と非共有結合的に相互作用するか、または塩基除去修復の阻害因子と会合することによって、塩基除去修復の阻害因子を標的ヌクレオチド配列へとターゲティングすることができる。例えば、いくつかの実施形態において、塩基除去修復成分の阻害因子は、ポリヌクレオチドプログラミング可能なヌクレオチド結合ドメインの一部であるさらなる異種部分またはドメインと相互作用、会合、または複合体形成し得るさらなる

10

20

30

40

50

異種部分またはドメインを含み得る。ある態様において、塩基除去修復の阻害因子は、ガイドポリヌクレオチドにより標的ヌクレオチド配列にターゲティングされ得る。例えば、いくつかの実施形態において、塩基除去修復の阻害因子は、ガイドポリヌクレオチドの一部またはセグメント(例えば、ポリヌクレオチドモチーフ)と相互作用、会合、または複合体形成し得るさらなる異種部分またはドメイン(例えば、RNAまたはDNA結合タンパク質のようなポリヌクレオチド結合ドメイン)を含み得る。いくつかの実施形態において、ガイドポリヌクレオチドのさらなる異種部分またはドメイン(例えば、RNAまたはDNA結合タンパク質のようなポリヌクレオチド結合ドメイン)は、塩基除去修復の阻害因子に融合または連結され得る。いくつかの実施形態において、追加の異種部分は、ポリヌクレオチドと結合、相互作用、会合、または複合体形成することができる。いくつかの実施形態において、追加の異種部分は、ガイドポリヌクレオチドに結合することができる。いくつかの実施形態において、追加の異種部分は、ポリペプチドリンカーに結合することができる。いくつかの実施形態において、追加の異種部分は、タンパク質ドメインであってもよい。いくつかの実施形態において、追加の異種部分は、K相同 (KH) ドメイン、MS2コートタンパク質ドメイン、PP7コートタンパク質ドメイン、SfMu Comコートタンパク質ドメイン、無菌アルファモチーフ、テロメラーゼKu結合モチーフおよびKuタンパク質、テロメラーゼSm7結合モチーフおよびSm7タンパク質、またはRNA認識モチーフであり得る。

【0110】

用語「Cas9」または「Cas9ドメイン」とは、Cas9タンパク質またはその断片(例えばCas9の活性、不活性、または部分的に活性なDNA切断ドメイン、および/またはCas9のgRNA結合ドメインを含むタンパク質)を含むRNA誘導型ヌクレアーゼを表す。Cas9ヌクレアーゼは、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呼ばれる。例示的なCas9は、アミノ酸配列が以下に提供されるStreptococcus pyogenes Cas9である：

【0111】

10

20

30

40

50

MDKKYSIGLDIGTNSVGWAVITDDYKVPSSKFKVLGNTDRHSIKKNLIGALLFGSGET
 AEATRLKRTARRRYTRRKNRICYLQEI FSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNI
 VDEVAYHEKYPTIYHLRKKLADSTDKADLRLLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFI
 QLVQIYNQLFEENPINASRVDAKAILSARLSKSRLENLIAQLPGEKRNGLFGNLIALSLGLTP
 NFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNSEITK
 APLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKP
 ILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPFLKDNREKIEK
 ILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEK
 VLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFK
 KIECFDSVEISGVEDRFNASLGAYHDLKIIKDKDFLDNEENEDILEDIVLTLTLFEDRGMIEE
 RLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIH
 DDSLTFKEDIQKAQVSGQGHSLEQIANLAGSPAIKKGILQTVKIVDELVKVMGHKPENIVIE
 MARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQE
 LDINRLSDYDVDHIVPQSFIDKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKNYWRQLLNAKL
 ITQRKFDNLTKAERGGLSELDKAGFIKROLVETROITKHVAQILDSRMNTKYDENDKLIREVKV
 ITLKSCLVSDFRKDFQFYKVRINNYHHAHDAYLNAVVGITALIKKYPKLESEFVYGDYKVYDVR
 KMIKSEQEI GKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATV
 RKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVA
 KVEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDLIIKLPKYSLEFEENGRK
 RMLASAGELQKGNELALPSKYVNFYLYLASHYEKLKGSPEDEQKQLFVEQHKHYLDEIIIEQISE
 FSKRVILADANLDKVL SAYNKHDKPIREQAENIIHLFTLTNLGAPAAFYFDTTIDRKRYTST
 KEVLDTLIHQSIITGLYETRIDLSQLGGD

10

20

30

(一本下線：HNHドメイン；二重下線：RuvCドメイン)。

【0112】

「保存的アミノ酸置換」または「保存的変異」という用語は、あるアミノ酸が共通の特性を有する別のアミノ酸に置き換わることを指す。個々のアミノ酸間の共通の性質を定義するための機能的な方法は、相同的な生物の対応するタンパク質間のアミノ酸変化の正規化された頻度を分析することである(Schulz, G. E. and Schirmer, R. H., Principles of Protein Structure, Springer-Verlag, New York (1979))。このような分析によれば、グループ内のアミノ酸が互いに優先的に交換され、それゆえに全体的なタンパク質構造に対するそれらの影響において互いに最も類似しているところアミノ酸のグループを定義することができる(上記Schulz, G. E. and Schirmer, R. H.)。保存変異の非限定的な例としては、例えば正電荷を維持することができるアミノ酸アルギニンからリジンおよびその逆；負電荷を維持することができるアスパラギン酸からグルタミン酸およびその逆；遊離の-OHが維持されるトレオニンからセリン；および遊離NH₂を維持できるアスパラギンからグルタミンのようなアミノ酸置換が挙げられる。

40

【0113】

本明細書中で交換可能に使用される用語「コード配列」または「タンパク質コード配列」とは、タンパク質をコードするポリヌクレオチドのセグメントをいう。この領域または配列は、5'末端に近い方が開始コドンで境界され、3'末端に近い方が停止コドンで境界さ

50

れる。コード配列はオープンリーディングフレームとも呼ばれる。

【0114】

本明細書中で使用する場合、用語「デアミナーゼ」または「デアミナーゼドメイン」とは、脱アミノ化反応を触媒するタンパク質または酵素を指す。ある態様において、デアミナーゼまたはデアミナーゼドメインは、それぞれウリジンまたはデオキシウリジンへのシチジンまたはデオキシシチジンの加水分解的脱アミノ化を触媒するシチジンデアミナーゼである。ある態様において、デアミナーゼまたはデアミナーゼドメインはシトシンデアミナーゼであり、ウラシルへのシトシンの加水分解的脱アミノ化を触媒する。ある態様において、デアミナーゼは、ヒポキサンチンへのアデニンの加水分解的脱アミノ化を触媒するアデノシンデアミナーゼである。ある態様において、デアミナーゼは、アデノシンまたはアデニン (A) からイノシン (I) への加水分解的脱アミノ化を触媒するアデノシンデアミナーゼである。ある態様において、デアミナーゼまたはデアミナーゼドメインは、アデノシンデアミナーゼであり、それぞれイノシンまたはデオキシイノシンへのアデノシンまたはデオキシアデノシンの加水分解的脱アミノ化を触媒する。ある態様において、アデノシンデアミナーゼは、デオキシリボ核酸 (DNA) 中のアデノシンの加水分解的脱アミノ化を触媒する。本明細書中で提供されるアデノシンデアミナーゼ(例えば、遺伝子操作されたアデノシンデアミナーゼ、進化されたアデノシンデアミナーゼ)は、細菌などの任意の生物由来であり得る。ある態様において、アデノシンデアミナーゼは、*E. coli*、*S. aureus*、*S. typhi*、*S. putrefaciens*、*H. influenzae*、または*C. crescentus*などの細菌に由来する。ある態様において、アデノシンデアミナーゼは、*TadA*デアミナーゼである。ある態様において、デアミナーゼまたはデアミナーゼドメインは、例えばヒト、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由来の天然デアミナーゼのバリエーションである。ある態様において、デアミナーゼまたはデアミナーゼドメインは、非天然のものである。例えば、いくつかの実施形態において、デアミナーゼまたはデアミナーゼドメインは、天然のデアミナーゼに対して少なくとも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%、少なくとも99.1%、少なくとも99.2%、少なくとも99.3%、少なくとも99.4%、少なくとも99.5%、少なくとも99.6%、少なくとも99.7%、少なくとも99.8%、または少なくとも99.9%の同一性を有する。例えば、デアミナーゼドメインは、国際PCT出願番号PCT/2017/045381(国際公開第2018/027078号)およびPCT/US 2016/058344(国際公開第2017/070632号)に記載されており、これらの各々は、その全体が参照により本明細書に組み込まれる。Komor, A.C., et al., "Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage" *Nature* 533, 420-424 (2016); Gaudelli, N.M., et al., "Programmable base editing of A·T to G·C in genomic DNA without DNA cleavage" *Nature* 551, 464-471 (2017); 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" *Science Advances* 3:eaao4774 (2017), and Rees, H.A., et al., "Base editing: precision chemistry on the genome and transcriptome of living cells." *Nat Rev Genet.* 2018 Dec;19(12):770-788. doi: 10.1038/s41576-018-0059-1も参照(その内容全体を参照により本明細書に組み込む)。

【0115】

「検出可能な標識」とは、目的の分子に連結されたときに、分光学的、光化学的、生化学的、免疫化学的、または化学的手段を介してそれを検出可能にする組成物を意味する。例えば、有用な標識には、放射性同位体、磁気ビーズ、金属ビーズ、コロイド粒子、蛍光色素、高電子密度試薬、酵素(例えばELISAで一般的に使用されるもの)、ビオチン、ジゴキシゲニン、またはハプテンが含まれる。

【0116】

「疾患」とは、細胞、組織または器官の正常な機能を損傷または妨害するあらゆる状態または障害を意味する。疾患の例には、網膜色素変性症、Usher症候群、鎌状赤血球症、βサラセミア、胎児ヘモグロビン遺伝性持続症 (HPFH)、α1アンチトリプシン欠損症 (A1AD)、肝ポルフィリン症、中鎖アシルCoA脱水素酵素 (ACADM) 欠損症、リソソーム酸性リパーゼ (LAL) 欠損症、フェニルケトン尿症、ヘモクロマトーシス、Von Gierke病、Pompe病、Gaucher病、Hurler症候群、嚢胞性線維症、慢性疼痛などがある。一実施形態では、疾患はA1ADである。一実施形態では、疾患は、鎌状赤血球症 (SCD) であり、これは「鎌状赤血球貧血」とも呼ばれる。

【0117】

「有効量」とは、未治療の患者、または疾患のない個体すなわち健康な個体と比較して、必要とする対象または患者における疾患の症状を改善するために必要な、薬剤または活性化合物（例えば本明細書に記載の塩基エディター）の量を意味する。疾患の治療的処置のための記載された方法を実施するために使用される活性化合物の有効量は、投与態様、対象の年齢、体重、および全般的健康状態に依存して変化する。最終的には、主治医または獣医が適切な量および用量を決定する。このような量は「有効」量と呼ばれる。一実施形態では、有効量は、細胞（例えば、インビトロまたはインビボの細胞）内の目的の遺伝子に改変を導入するのに十分な本発明の塩基エディターの量である。一実施形態では、有効量は、治療効果を達成する（例えば、網膜色素変性、アッシャー症候群、鎌状赤血球症 (SCD)、β-サラセミア、胎児ヘモグロビン遺伝性持続症 (HPFH)、α-1アンチトリプシン欠損症 (A1AD)、肝ポルフィリン症、中鎖アシルCoAデヒドロゲナーゼ (ACADM) 欠損症、リソソーム酸性リパーゼ (LAL) 欠損症、フェニルケトン尿症、ヘモクロマトーシス、フォン・ギールケ病、ポンペ病、ゴーシェ病、ハーラー症候群、嚢胞性線維症、または慢性疼痛を低減または制御する）ために必要な塩基エディターの量である。このような治療効果は、対象、組織または器官の全ての細胞において病原性遺伝子を変化させるのに十分である必要はなく、対象、組織または器官に存在する細胞の約1%、5%、10%、25%、50%、75%またはそれ以上において病原性遺伝子を変化させるだけでよい。一実施形態では、有効量は、疾患（例えば、網膜色素変性、Usher症候群、鎌状赤血球症 (SCD)、βサラセミア、胎児ヘモグロビン遺伝性持続症 (HPFH)、α1アンチトリプシン欠損症 (A1AD)、肝ポルフィリン症、中鎖アシルCoA脱水素酵素 (ACADM) 欠損症、リソソーム酸性リパーゼ (LAL) 欠損症、フェニルケトン尿症、ヘモクロマトーシス、フォン・ゲルケ病、ポンペ病、ゴーシェ病、ハーラー症候群、嚢胞性線維症、または慢性疼痛)の一つ以上の症状を改善するのに十分である。

【0118】

「断片」とは、ポリペプチドまたは核酸分子の一部を意味する。この部分は、好ましくは、参照核酸分子またはポリペプチドの全長の少なくとも10%、20%、30%、40%、50%、60%、70%、80%、または90%を含む。断片は、10、20、30、40、50、60、70、80、90、または100、200、300、400、500、600、700、800、900、または1000のヌクレオチドまたはアミノ酸を含み得る。

【0119】

「ハイブリダイゼーション」は、相補的な核酸塩基間の水素結合を意味し、ワトソン-クリック、フーグスティーンまたは逆フーグスティーン水素結合であり得る。たとえば、アデニンとチミンは相補的な核酸塩基で、水素結合を形成して対を形成する。

【0120】

用語「塩基修復の阻害因子」または「IBR」とは、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ことができるタンパク質を指す。ある態様において、IBRは、イノシン塩基除去修復の阻害因子である。塩基修復の阻害因子の例としては、APE1、Endo II I、Endo IV、Endo V、Endo VIII、Fpg、hOGG1、hNEIL1、T7 Endo I、T4 PDG、UDG、hSMUGLおよびhAAGの阻害因子が挙げられる。ある実施形態では、IBRは、Endo VまたはhAAGの阻害因子である。ある実施形態では、IBRは、触媒的に不活性なEndo Vまたは触媒的に不活性なhAAGである。ある実施形態では、塩基修復阻害因子は、Endo Vま

たはhAAGの阻害因子である。ある実施形態では、塩基修復阻害因子は、触媒的に不活性なEndoVまたは触媒的に不活性なhAAGである。ある実施形態では、塩基修復阻害因子は、ウラシルグリコシラーゼ阻害因子 (UGI) である。UGIとは、ウラシル-DNAグリコシラーゼの塩基除去修復酵素を阻害することができるタンパク質を指す。いくつかの実施形態において、UGIドメインは、野生型UGIまたはその断片を含む。いくつかの実施形態において、本明細書で提供されるUGIタンパク質は、UGIの断片、および、UGIまたはUGI断片に相同的なタンパク質を含む。ある態様において、塩基修復阻害因子は、イノシン塩基除去修復の阻害因子である。いくつかの態様において、塩基修復阻害因子は、「触媒的に不活性なイノシン特異的ヌクレアーゼ」または「死んだイノシン特異的ヌクレアーゼ」である。いかなる特定の理論にも拘束されることを望まないが、触媒的に不活性なイノシングリコシラーゼ（例えば、アルキルアデニングリコシラーゼ (AAG)）は、イノシンに結合することができるが、脱塩基部位を作ることイノシンを除去することもできず、それによって、新たに形成されるイノシン部分をDNA損傷/修復機構から立体的にブロックする。いくつかの実施形態において、触媒的に不活性なイノシン特異的ヌクレアーゼは、核酸中のイノシンに結合することができるが、核酸を切断しない。代表的な、触媒的に不活性なイノシン特異的ヌクレアーゼの非限定的例としては、（例えばヒト由来の）触媒的に不活性なアルキルアデノシングリコシラーゼ(AAGヌクレアーゼ)、および（例えばE.coli由来の）触媒的に不活性なエンドヌクレアーゼV (EndoVヌクレアーゼ)が挙げられる。いくつかの実施形態において、触媒的に不活性なAAGヌクレアーゼは、E125Q突然変異または別のAAGヌクレアーゼにおける対応する突然変異を含む。

【0121】

用語「単離された」、「精製された」、または「生物学的に純粋な」とは、その天然状態で見出される場合に通常それに付随する成分が、様々な程度まで除去されている物質を指す。「単離」は、元の入手源又は周囲環境からの分離の程度を示す。「精製」は、単離よりも高い分離度を示す。「精製された」又は「生物学的に純粋な」タンパク質は、不純物がタンパク質の生物学的特性に実質的に影響を与えたり他の有害な結果を引き起こさないように、他の物質が十分に除去されている。すなわち、本発明の核酸またはペプチドは、組換えDNA技術によって産生された場合には細胞物質、ウイルス物質または培地を実質的に含まない場合、あるいは化学的に合成された場合には化学的前駆体その他の化学物質を実質的に含まない場合には、精製されている。純度および均一性は、典型的には、分析化学技術、例えば、ポリアクリルアミドゲル電気泳動または高速液体クロマトグラフィーを用いて決定される。用語「精製された」は、核酸またはタンパク質が電気泳動ゲルにおいて本質的に1つのバンドを生じることを意味し得る。修飾、例えば、リン酸化またはグリコシル化を受けることができるタンパク質については、異なる修飾は、別々に精製することができる異なる単離されたタンパク質を生じ得る。

【0122】

「単離ポリヌクレオチド」とは、本発明の核酸分子が由来する生物の天然ゲノムにおいて当該遺伝子に隣接する遺伝子を含まない核酸(例えばDNA)を意味する。したがって、この用語は、例えば、ベクターに組み込まれた；自律的に複製するプラスミドやウイルスに組み込まれた；原核生物や真核生物のゲノムDNAに組み込まれた；または他の配列とは独立した別の分子(例えば、PCRまたは制限エンドヌクレアーゼ消化によって生成されたcDNAまたはゲノムもしくはcDNA断片)として存在する組換えDNAを含む。さらに、この用語は、DNA分子から転写されるRNA分子、ならびに、さらなるポリペプチド配列をコードするハイブリッド遺伝子の一部である組換えDNAを含む。

【0123】

「単離ポリペプチド」とは、天然状態で付随する成分から分離された本発明のポリペプチドを意味する。典型的には、ポリペプチドは、それが天然状態で会合しているタンパク質および天然有機分子から重量で少なくとも60%フリーである場合に、単離されている。好ましくは、調製物は、重量で少なくとも75%、より好ましくは少なくとも90%、最も好ましくは少なくとも99%が本発明のポリペプチドである。本発明の単離されたポリペプチ

ドは、例えば、天然源からの抽出、そのようなポリペプチドをコードする組換え核酸の発現；または化学的にタンパク質を合成することにより得ることができる。純度は、任意の適切な方法、例えば、カラムクロマトグラフィー、ポリアクリルアミドゲル電気泳動、またはHPLC分析によって測定することができる。

【0124】

本明細書において使用される用語「リンカー」は、二つの分子もしくは部分（例えばタンパク質複合体もしくはリボ核複合体の二つの成分、または融合タンパク質の二つのドメイン、例えば、ポリヌクレオチドプログラミング可能DNA結合ドメイン(例dCas9)とデアミナーゼドメイン(例えばアデノシンデアミナーゼまたはシチジンデアミナーゼ))を連結する共有結合リンカー(例えば共有結合)、非共有結合リンカー、化学基、または分子を指すことができる。リンカーは、塩基エディターシステムの異なるコンポーネントまたはコンポーネントの異なる部分を繋げることができる。例えば、いくつかの実施形態において、リンカーは、ポリヌクレオチドプログラミング可能ヌクレオチド結合ドメインのガイドポリヌクレオチド結合ドメインおよびデアミナーゼの触媒ドメインを繋げることができる。ある態様において、リンカーは、CRISPRポリペプチドおよびデアミナーゼを結合することができる。ある態様において、リンカーは、Cas9およびデアミナーゼを繋げることができる。ある態様において、リンカーは、dCas9およびデアミナーゼを繋げることができる。ある態様において、リンカーは、nCAs9およびデアミナーゼを繋げることができる。ある態様において、リンカーは、ガイドポリヌクレオチドおよびデアミナーゼを繋げることができる。いくつかの実施形態において、リンカーは、塩基エディター系の脱アミノ化成分およびポリヌクレオチドプログラミング可能なヌクレオチド結合成分を繋げることができる。いくつかの実施形態において、リンカーは、塩基エディターシステムの脱アミノ化成分のRNA結合部分と、ポリヌクレオチドプログラミング可能なヌクレオチド結合成分とを繋げることができる。いくつかの実施形態において、リンカーは、塩基エディタ系の脱アミノ化成分のRNA結合部分と、ポリヌクレオチドプログラミング可能なヌクレオチド結合成分のRNA結合部分とを結合することができる。リンカーは、2つの基、分子、または他の部分の間に配置され、またはそれらによって挟まれ、共有結合または非共有結合相互作用を介してそれぞれに連結され、かくして2つを連結することができる。ある態様において、リンカーは、有機分子、基、ポリマー、または化学的部分であり得る。ある態様において、リンカーはポリヌクレオチドであり得る。ある態様において、リンカーはDNAリンカーであり得る。ある態様において、リンカーはRNAリンカーであり得る。ある態様において、リンカーは、リガンドに結合することができるアプタマーを含むことができる。ある態様において、リガンドは、炭水化物、ペプチド、タンパク質、または核酸であり得る。いくつかの実施形態において、リンカーは、リボスイッチに由来するアプタマーを含むことができる。アプタマーが由来するリボスイッチは、テオフィリンリボスイッチ、チアミンピロリン酸 (TPP) リボスイッチ、アデノシンコバラミン (AdoCbl) リボスイッチ、S-アデノシルメチオニン (SAM) リボスイッチ、SAHリボスイッチ、フラビンモノヌクレオチド (FMN) リボスイッチ、テトラヒドロ葉酸リボスイッチ、リジンリボスイッチ、グリシンリボスイッチ、プリンリボスイッチ、GlmSリボスイッチ、またはプレクエオン1 (PreQ1) リボスイッチから選択され得る。ある態様において、リンカーは、ポリペプチドまたはポリペプチドリガンドなどのタンパク質ドメインに結合したアプタマーを含み得る。ある態様において、ポリペプチドリガンドは、K相同 (KH) ドメイン、MS2コートタンパク質ドメイン、PP7コートタンパク質ドメイン、SfMu Comコートタンパク質ドメイン、無菌の α モチーフ、テロメラーゼKu結合モチーフおよびKuタンパク質、テロメラーゼSm7結合モチーフおよびSm7タンパク質、またはRNA認識モチーフであり得る。ある態様において、ポリペプチドリガンドは、塩基エディター系成分の一部であり得る。例えば、核酸塩基編集成分は、デアミナーゼドメインおよびRNA認識モチーフを含み得る。

【0125】

ある態様において、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タンパク質)であり得る。ある態様において、リンカーは、長さが約5~100アミノ酸、例

10

20

30

40

50

例えば、長さが約5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、20-30、30~40、40~50、50~60、60~70、70~80、80~90または90~100アミノ酸であり得る。ある態様において、リンカーは、長さが約100~150、150~200、200~250、250~300、300~350、350~400、400~450、または450~500アミノ酸であり得る。より長いまたはより短いリンカーも考えられる。

【0126】

ある態様において、リンカーは、Cas9ヌクレアーゼドメインを含むRNAプログラム可能ヌクレアーゼのgRNA結合ドメインと、核酸編集タンパク質(例えば、シチジンまたはアデノシンデアミナーゼ)の触媒ドメインとを繋げる。ある態様において、リンカーは、dCas9および核酸編集タンパク質を繋げる。例えば、リンカーは、2つの基、分子、または他の部分の間に配置されるか、または2つの基、分子、または他の部分によって隣接され、共有結合を介してそれぞれに連結され、したがって2つを連結する。ある態様において、リンカーは、アミノ酸または複数のアミノ酸(例えば、ペプチドまたはタンパク質)である。ある態様において、リンカーは、有機分子、基、ポリマー、または化学的部分である。ある態様において、リンカーは、長さが5~200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、25、35、45、50、55、60、60、65、70、70、75、80、85、90、90、95、100、101、102、103、104、105、110、120、130、140、150、160、175、180、190、または200アミノ酸である。より長いまたはより短いリンカーも考えられる。ある態様において、リンカーは、XTENリンカーとも呼ばれ得るアミノ酸配列SGSETPGTSESATPESを含む。ある態様において、リンカーは、アミノ酸配列SGGSを含む。ある態様において、リンカーは、(SGGS)_n、(GGGS)_n、(GGGGGS)_n、(G)_n、(EAAAK)_n、(GGS)_n、SGSETPGTSESATPES、または(XP)_nモチーフ、またはこれらの任意の組合せを含み、ここで、nは独立して1~30の整数であり、Xは任意のアミノ酸である。いくつかの実施態様において、nは、1、2、3、4、5、6、7、8、9、10、11、12、13、14又は15である。ある態様において、リンカーは、複数のプロリン残基を含み、長さが5~21、5~14、5~9、5~7アミノ酸であり、例えば、PAPAP、PAPAPA、PAPAPAP、PAPAPAPA、P(AP)₄、P(AP)₇、P(AP)₁₀である。このようなプロリンに富むリンカーはまた、「リジッド (rigid)」リンカーとも呼ばれる。

【0127】

ある態様において、塩基エディターのドメインは、SGGSSGSETPGTSESATPESGGGS、SGGSSGGSSGSETPGTSESATPESGGSSGGGS、またはGGSGGSPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGGSSGGGSのアミノ酸配列を含むリンカーを介して融合される。ある態様において、塩基エディターのドメインは、XTENリンカーとも呼ばれ得るアミノ酸配列SGSETPGTSESATPESを含むリンカーを介して融合される。ある態様において、リンカーは、長さが24アミノ酸である。ある態様において、リンカーは、アミノ酸配列SGGSSGGSSGSETPGTSESATPESを含む。ある態様において、リンカーは、40アミノ酸長である。ある態様において、リンカーは、アミノ酸配列SGGSSGGSSGSETPGTSESATPESGGSSGGSSGSSGGGSを含む。ある態様において、リンカーは、64アミノ酸長である。ある態様において、リンカーは、アミノ酸配列SGGSSGGSSGSETPGTSESATPESGGSSGGSSGGSSGGSSGSETPGTSESATPESGGSSGGGSを含む。ある態様において、リンカーは、長さが92アミノ酸である。いくつかの実施形態では、リンカーは、アミノ酸配列PGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGGSSGGGSを含む。

【0128】

本明細書中で使用される場合、用語「変異」とは、配列内、例えば核酸もしくはアミノ酸配列内の残基が、別の残基により置換されること、または配列内の1以上の残基の欠失もしくは挿入をいう。変異は、本明細書中において典型的には、元の残基を同定し、次いで配列内の残基の位置を同定し、そして新たに置換された残基を同定することによって、記載される。本明細書中に提供されるアミノ酸置換(変異)を作製するための種々の方法は

、当該技術分野においてよく知られており、例えば、Green and Sambrook, *Molecular Cloning: A Laboratory Manual* (4th ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (2012)) によって提供されている。いくつかの態様において、本開示の塩基エディターは、有意な数の非意図的な変異、例えば非意図的な点突然変異を生成することなく、核酸(例えば対象のゲノム内の核酸)において「意図された変異」例えば点突然変異を効率的に生成することができる。いくつかの実施形態において、意図された変異は、その意図された変異を生じさせるように特に設計された、ガイドポリヌクレオチド(例えばgRNA)に結合した特定の塩基エディター(例えばシチジン塩基エディターまたはアデノシン塩基エディター)によって生じる変異である。

【0129】

一般に、配列(例えば、本明細書に記載のアミノ酸配列)においてなされるか又は同定される変異は、参照(または野生型)配列、すなわち、その変異を含まない配列に関連して番号付けされる。当業者は、参照配列に対するアミノ酸および核酸配列における突然変異の位置を決定する方法を容易に理解するであろう。

【0130】

「核局在化配列」、「核局在化シグナル」または「NLS」という用語は、タンパク質の細胞核への移入を促進するアミノ酸配列を意味する。核局在化配列は当該技術分野において公知であり、例えば、2000年11月23日に出願され2001年5月31日にWO/2001/038547として公開された国際PCT出願PCT/EP 2000/011690のPlank et al.に記載されており、その内容は、例示的な核局在化配列の開示について参照により本明細書に組み入れられる。他の実施形態では、NLSは、例えばKoblan et al., *Nature Biotech.* 2018 doi:10.1038/nbt.4172によって記載された、最適化されたNLSである。ある態様において、NLSは、アミノ酸配列KRTADGSEFESPKKKRKV, KRPAATKKAGQAKKKK, KKTELQTNAENKTKKL, KRGINDRNFWRGENGRKTR, RKSGKIAAIVVKRPRK, PKKKRKV,またはMDSLLMNRKFLYQFKNVRWAKGRRETYLCを含む。

【0131】

用語「核酸塩基」、「窒素塩基」、または「塩基」は、本明細書中で互換的に使用され、ヌクレオシドを形成する窒素含有生物学的化合物を指し、ヌクレオシドはヌクレオチドの成分である。塩基対を形成し、互いに積み重なる核酸塩基の能力は、直接、リボ核酸(RNA) およびデオキシリボ核酸(DNA) のような長鎖らせん構造をもたらす。アデニン(A)、シトシン(C)、グアニン(G)、チミン(T)、ウラシル(U)の5種類の核酸塩基は、一次塩基または標準塩基と呼ばれる。アデニンとグアニンはプリンに由来し、シトシン、ウラシル、チミンはピリミジンに由来する。DNAとRNAは修飾された他の(非一次)塩基を含むこともある。非限定的な例示的修飾核酸塩基としては、ヒポキサンチン、キサンチン、7-メチルグアニン、5,6-ジヒドロウラシル、5-メチルシトシン(m5C)、および5-ヒドロメチルシトシンが挙げられる。ヒポキサンチンとキサンチンは変異原の存在によって生成され得、どちらも脱アミノ化(アミン基のカルボニル基への置換)によって生成される。ヒポキサンチンはアデニンから修飾され得る。キサンチンはグアニンから修飾され得る。ウラシルはシトシンの脱アミノ化によって生じ得る。「ヌクレオシド」は核酸塩基と五炭糖(リボースまたはデオキシリボース)からなる。ヌクレオシドの例としては、アデノシン、グアノシン、ウリジン、シチジン、5-メチルウリジン(m5U)、デオキシアデノシン、デオキシグアノシン、チミジン、デオキシウリジン、およびデオキシシチジンが挙げられる。修飾核酸塩基を有するヌクレオシドとしては、イノシン(I)、キサントシン(X)、7-メチルグアノシン(m7G)、ジヒドロウリジン(D)、5-メチルシチジン(m5C)、プソイドウリジン(Ψ)等が挙げられる。ヌクレオチドは、核酸塩基、五炭糖(リボースまたはデオキシリボース)、および少なくとも一つのリン酸基からなる。

【0132】

本明細書中で使用される場合、用語「核酸」および「核酸分子」とは、核酸塩基および酸性部分を含む化合物、例えばヌクレオシド、ヌクレオチド、またはヌクレオチドのポリマーをいう。典型的には、ポリマー核酸、例えば3つ以上のヌクレオチドを含む核酸分子

10

20

30

40

50

は、隣接するヌクレオチドがホスホジエステル結合を介して互いに連結されている直鎖分子である。ある態様において、「核酸」は、個々の核酸残基(例えばヌクレオチドおよび/またはヌクレオシド)を指す。ある態様において、「核酸」は、三つ以上の個々のヌクレオチド残基を含むオリゴヌクレオチド鎖をいう。本明細書中で使用される、用語「オリゴヌクレオチド」、「ポリヌクレオチド」、および「ポリ核酸」は、ヌクレオチドのポリマー(例えば少なくとも3つのヌクレオチドの鎖)を指すために交換可能に使用され得る。ある態様において、「核酸」は、RNAならびに一本鎖および/または二本鎖DNAを包含する。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、染色分体、または他の天然に存在する核酸分子との関連において、天然に存在し得る。他方、核酸分子は、例えば非天然に存在する分子、組換えDNAもしくはRNA、人工染色体、操作されたゲノム、もしくはその断片、または合成DNA、RNA、DNA/RNAハイブリッド、または非天然に存在するヌクレオチドもしくはヌクレオシドを含む、非天然に存在する分子であり得る。さらに、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骨格以外を有するアナログを含む。核酸は、天然源から精製される、組換え発現系を用いて産生され必要に応じて精製される、化学的に合成される、等が可能である。化学的に合成された分子の場合、核酸は、適切な場合には、例えば、化学的に修飾された塩基または糖、および骨格修飾を有するアナログなどのヌクレオシドアナログを含み得る。核酸配列は、特に示されない限り、5'〜3'方向に示される。ある態様において、核酸は、天然ヌクレオシド(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)；ヌクレオシド類似体(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-オキソグアニン、O(6)-メチルグアニン、および2-チオシチジン)；化学修飾塩基；生物学的に修飾された塩基(例えばメチル化塩基)；挿入塩基；修飾糖(例えば2'-フルオリボース、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結合)であるか、またはそれらを含む。

【0133】

「核酸プログラム可能DNA結合タンパク質」または「napDNAbp」という用語は、「ポリヌクレオチドプログラム可能ヌクレオチド結合ドメイン」と交換可能に使用され、napDNAbpを特定の核酸配列にガイドする、ガイド核酸などの核酸(例えばDNAまたはRNA)に付随するタンパク質を意味する。例えば、Cas9タンパク質は、ガイドRNAに相補的な特定のDNA配列にCas9タンパク質をガイドするガイドRNAと結合することができる。ある態様において、napDNAbpは、Cas9ドメイン、例えばヌクレアーゼ活性Cas9、Cas9ニッカーゼ(nCas9)、またはヌクレアーゼ不活性Cas9(dCas9)である。いくつかの実施形態において、Cas9ドメインは、本明細書に記載のアミノ酸配列のいずれか1つを含む。いくつかの実施形態において、Cas9ドメインは、本明細書に記載されたアミノ酸配列のいずれかに対して少なくとも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%の同一性を有するアミノ酸配列を含む。いくつかの実施形態において、Cas9ドメインは、本明細書に記載されるアミノ酸配列のいずれかと比較して、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またはそれ以上の変異を有するアミノ酸配列を含む。いくつかの実施形態において、Cas9ドメインは、本明細書に記載されたアミノ酸配列のいずれかと比較して、少なくとも10、少なくとも15、少なくとも20、少なくとも30、少なくとも40、少

なくとも50、少なくとも60、少なくとも70、少なくとも80、少なくとも90、少なくとも100、少なくとも150、少なくとも200、少なくとも250、少なくとも300、少なくとも350、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1100、または少なくとも1200の同一の連続したアミノ酸残基を有するアミノ酸配列を含む。

【0134】

核酸プログラム可能なDNA結合タンパク質の例としては、Cas9 (例:dCas9およびnCas9)、Cas12a/Cpf1、Cas12b/C2cl、Cas12c/C2c3、Cas12d/CasY、Cas12e/CasX、Cas12g、Cas12h、およびCas12iが挙げられるが、これらに限定されない。他の核酸プログラム可能なDNA結合タンパク質もまた、本開示の範囲内であるが、それらは本開示に具体的に記載されていないこともあり得る。例えば、Makarova et al. “Classification and Nomenclature of CRISPR-Cas Systems: Where from Here?” CRISPR J. 2018 Oct;1:325-336. doi: 10.1089/crispr.2018.0033; Yan et al., “Functionally diverse type V CRISPR-Cas systems” Science. 2019 Jan 4;363(6422):88-91. doi: 10.1126/science.aav7271に記載されているとおりであり、それぞれの内容全体を参照により本明細書に組み込む。

【0135】

「核酸塩基編集ドメイン」または「核酸塩基編集タンパク質」という用語は、本明細書において使用される場合、シトシン(もしくはシチジン)からウラシル(もしくはウリジン)またはチミン(もしくはチミジン)への、およびアデニン(もしくはアデノシン)からヒポキサンチン(もしくはイノシン)への脱アミノ化、ならびに非鋳型ヌクレオチド付加および挿入のような、RNAまたはDNAにおける核酸塩基修飾を触媒することができるタンパク質または酵素を指す。ある態様において、核酸塩基編集ドメインは、デアミナーゼドメイン(例えば、シチジンデアミナーゼ、シトシンデアミナーゼ、アデニンデアミナーゼ、またはアデノシンデアミナーゼ)である。ある態様において、核酸塩基編集ドメインは、天然に存在する核酸塩基編集ドメインであり得る。いくつかの実施形態において、核酸塩基編集ドメインは、天然に存在する核酸塩基編集ドメインから操作または進化された核酸塩基編集ドメインであり得る。核酸塩基編集ドメインは、細菌、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任意の生物由来であり得る。例えば、核酸塩基編集タンパク質は、国際PCT出願番号PCT/2017/045381(国際公開第2018/027078号)およびPCT/US 2016/058344(国際公開第2017/070632号)に記載されており、これらの各々は、その全体が参照により本明細書に組み込まれている。Komor, A.C., et al., “Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage” Nature 533, 420-424 (2016); Gaudelli, N.M., et al., “Programmable base editing of A · T to G · C in genomic DNA without DNA cleavage” Nature 551, 464-471 (2017); および 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” Science Advances 3:eaa04774 (2017)も参照(その内容全体が参照により本明細書に組み込まれる)。

【0136】

本明細書で使用される場合、「薬剤を取得する」におけるような「取得する」は、その薬剤を合成すること、購入すること、単離すること、または他の方法で獲得することを含む。

【0137】

本明細書で使用される「患者」または「対象」は、疾患または障害と診断されている、それらを生じるリスクがある、またはそれらを有するかもしれない生ずる疑いがある哺乳動物対象または個体を指す。ある態様において、用語「患者」は、疾患または障害を生ずる可能性が平均より高い哺乳動物対象を指す。例示的な患者は、ヒト、ヒト以外の霊長類、ネコ、イヌ、ブタ、ウシ、ネコ、ウマ、ラクダ、ラマ、ヤギ、ヒツジ、齧歯類(例えば、マウス、ウサギ、ラット、モルモット)および本明細書に開示される治療から利益を得ること

ができる他の哺乳動物であり得る。例示的なヒト患者は、男性および/または女性であり得る。

【0138】

「それを必要とする患者」または「それを必要とする対象」は、本明細書では、疾患または障害、例えば限定されないが鎌状赤血球症 (SCD) もしくは α -1アンチトリプシン欠損症 (A1AD)、または本明細書の表3A、3B、または4に列挙されている遺伝子に関連する疾患または障害と診断された、これらの疾患または障害を有するリスクがある、またはこれらの疾患または障害を有することが疑われる患者または対象を指す。

【0139】

「病原性変異」、「病原性バリエーション」、「疾患原因(または疾患関連)変異」、「疾患原因(または疾患関連)バリエーション」、「有害変異」または「素因となる変異」という用語は、特定の疾患または障害に対する固体の感受性または素因を増大させる遺伝子変化または突然変異を指す。ある態様において、病原性変異は、遺伝子によってコードされるタンパク質中の少なくとも1つの野生型アミノ酸が少なくとも1つの病原性アミノ酸によって置換されたものを含む。

【0140】

「非保存的変異」という用語は、異なるグループの間のアミノ酸置換を意味し、例えば、トリプトファンからリジン、またはセリンからフェニルアラニンなどである。この場合、非保存的アミノ酸置換は、機能性バリエーションの生物学的活性を妨害または阻害しないことが好ましい。非保存的アミノ酸置換は、機能的バリエーションの生物学的活性が野生型タンパク質と比較して増加するように、機能的バリエーションの生物学的活性を増強することができる。

【0141】

用語「タンパク質」、「ペプチド」、「ポリペプチド」、およびそれらの文法的等価物は、本明細書において交換可能に使用され、ペプチド (アミド) 結合によって互いに連結されたアミノ酸残基のポリマーを指す。この用語は、任意のサイズ、構造または機能のタンパク質、ペプチドまたはポリペプチドを指す。典型的には、タンパク質、ペプチド、またはポリペプチドは、少なくとも3アミノ酸長である。タンパク質、ペプチド、またはポリペプチドは、個々のタンパク質または一群のタンパク質を指すことができる。タンパク質、ペプチド、またはポリペプチド中の1以上のアミノ酸は、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、結合のためのリンカー、官能化、または他の修飾などの化学的実体の添加によって、修飾され得る。タンパク質、ペプチド、またはポリペプチドはまた、単一分子であってもよく、または多分子複合体であってもよい。タンパク質、ペプチド、またはポリペプチドは、天然に存在するタンパク質またはペプチドの単なる断片であり得る。タンパク質、ペプチド、またはポリペプチドは、天然に存在するもの、組換え体、もしくは合成のもの、またはそれらの任意の組み合わせであり得る。本明細書中で使用される場合、用語「融合タンパク質」とは、少なくとも二つの異なるタンパク質由来のタンパク質ドメインを含むハイブリッドポリペプチドをいう。1つのタンパク質は、融合タンパク質のアミノ末端 (N末端) 部分またはカルボキシ末端 (C末端) タンパク質に位置することができ、かくして、それぞれ、アミノ末端融合タンパク質またはカルボキシ末端融合タンパク質を形成する。タンパク質は、異なるドメイン、例えば、核酸結合ドメイン (例えば、標的部位へのタンパク質の結合を誘導する Cas9 の gRNA 結合ドメイン) および核酸切断ドメイン、あるいは核酸編集タンパク質の触媒ドメインを含み得る。ある態様において、タンパク質は、タンパク質性部分、例えば、核酸結合ドメインを構成するアミノ酸配列、および有機化合物、例えば、核酸切断剤として作用し得る化合物を含む。ある態様において、タンパク質は、核酸 (例えば、RNA または DNA) と複合体化されているか、または核酸と会合している。本明細書で提供される任意のタンパク質は、当技術分野で公知の任意の方法によって生産することができる。例えば、本明細書で提供されるタンパク質は、組換えタンパク質発現および精製を介して生産することができ、これは、ペプチドリンカーを含む融合タンパク質に特に適している。組換

10

20

30

40

50

えタンパク質の発現および精製のための方法はよく知られており、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含み、その全内容は参照により本明細書に組み入れられる。

【0142】

本明細書に開示されるポリペプチドおよびタンパク質(機能的部分およびその機能的バリエーションを含む)は、一つ以上の天然に存在するアミノ酸の代わりに合成アミノ酸を含むことができる。そのような合成アミノ酸は当該分野で公知であり、例えば、アミノシクロヘキサンカルボン酸、ノルロイシン、 α -アミノ-n-デカン酸、ホモセリン、S-アセチルアミノメチル-システイン、trans-3-およびtrans-4-ヒドロキシプロリン、4-アミノフェニルアラニン、4-ニトロフェニルアラニン、4-クロロフェニルアラニン、4-カルボキシフェニルアラニン、 β -フェニルセリン β -ヒドロキシフェニルアラニン、フェニルグリシン、 α -ナフチルアラニン、シクロヘキシルアラニン、シクロヘキシルグリシン、インドリン-2-カルボン酸、1,2,3,4-テトラヒドロイソキノリン-3-カルボン酸、アミノマロン酸、アミノマロン酸モノアミド、N'-ベンジル-N'-メチルリジン、N',N'-ジベンジル-リジン、6-ヒドロキシリジン、オルニチン、 α -アミノシクロペンタンカルボン酸、アミノシクロヘキサンカルボン酸、アミノシクロヘキサンカルボン酸、 α -アミノシクロヘプタンカルボン酸、 α -(2-アミノ-2-ノルボルナン)-カルボン酸、 α, γ -ジアミノ酪酸、 α, β -ジアミノプロピオン酸、ホモフェニルアラニン、および α -tert-ブチルグリシンが挙げられる。ポリペプチドおよびタンパク質は、ポリペプチド構築物の1つ以上のアミノ酸の翻訳後修飾と結合することができる。翻訳後修飾の非限定的な例としては、リン酸化、アセチル化およびホルミル化を含むアシル化、グリコシル化(N-リンクおよびO-リンクを含む)、アミド化、ヒドロキシル化、メチル化およびエチル化を含むアルキル化、ユビキチン化、ピロリドンカルボン酸の添加、ジスルフィド架橋の形成、硫酸化、ミリストイル化、パルミトイル化、イソプレニル化、ファルネシル化、ゲラニル化、グリビエーション、リポイル化およびヨード化が挙げられる。

【0143】

本明細書中で使用される場合、用語「遺伝子」とは、典型的には、タンパク質コード領域およびタンパク質非コード領域を含むポリヌクレオチドをいう。タンパク質非コード領域は、1以上の調節エレメントを含み得る。調節エレメントの非限定的な例は、プロモーター、エンハンサー、リプレッサー、サイレンサー、インスレーター、開始コドン、終止コドン、Kozakコンセンサス配列、スライス受容体、スプライスドナー、3'および/または5'非翻訳領域(UTR)、スライス部位、または遺伝子間領域を含む。ある態様において、調節エレメントは、遺伝的疾患または遺伝的障害の原因である遺伝子に位置する。遺伝的疾患または障害の原因である遺伝子に位置する調節エレメントの非限定的な例としては、開始コドン、終止コドン、Kozakコンセンサス配列、遺伝子間領域、3'UTR、または5'UTRなどが挙げられる。いくつかの態様において、調節エレメントは、遺伝的疾患または障害の原因である遺伝子に位置しない。遺伝的疾患の原因である遺伝子に位置しない調節エレメントの非限定的な例としては、エンハンサー、リプレッサー、またはインスレーターなどが挙げられる。

【0144】

用語「ポリヌクレオチドプログラム可能ヌクレオチド結合ドメイン」とは、ポリヌクレオチドプログラム可能なDNA結合ドメインを特定の核酸配列にガイドするガイドポリヌクレオチド(例えばガイドRNA)のような核酸(DNAやRNAなど)に付随するタンパク質をいう。ある態様において、ポリヌクレオチドプログラム可能なヌクレオチド結合ドメインは、ポリヌクレオチドプログラム可能なDNA結合ドメインである。ある態様において、ポリヌクレオチドプログラム可能なヌクレオチド結合ドメインは、ポリヌクレオチドプログラム可能なRNA結合ドメインである。ある態様において、ポリヌクレオチドプログラム可能なヌクレオチド結合ドメインは、Cas9タンパク質である。Cas9タンパク質は、ガイドRNAに相補的な特定のDNA配列にCas9タンパク質をガイドするガイドRNAと結合することが

できる。ある態様において、ポリヌクレオチドプログラム可能なヌクレオチド結合ドメインは、Cas9ドメイン、例えばヌクレアーゼ活性Cas9、Cas9ニッカーゼ (nCas9)、またはヌクレアーゼ不活性Cas9 (dCas9) である。核酸プログラム可能DNA結合タンパク質の非限定的な例は、Cas9 (例:dCas9およびnCas9)、Cas12a/Cpf1、Cas12b/C2cl、Cas12c/C2c3、Cas12d/CasY、Cas12e/CasX、Cas12g、Cas12h、およびCas12iを含む。Cas酵素の非限定的な例としては、Cas1、Cas1B、Cas2、Cas3、Cas4、Cas5、Cas5d、Cas5t、Cas5h、Cas5a、Cas6、Cas7、Cas8、Cas8a、Cas8b、Cas8c、Cas9(Csn1またはCsn12とも呼ばれる)、Cas10、Cas10d、Cas12a/Cpf1、Cas12b/C2cl、Cas12c/C2c3、Cas12d/CasY、Cas12e/CasX、Cas12g、Cas12h、Cas12i、Csy1、Csy2、Csy3、Csy4、Cse1、Csc1、Csc1、Csm2、Csm3、Csm3、Csm4、Csm6、Cmr1、Cmr3、Cmr4、Cmr5、Cmr6、Csb1、Csb2、Csb3、Csn17、Csn14、Csn10、Csn16、CsaX、Csn3、Csn1、Csn1S、Csn11、Csf1、Csf2、CsO、Csf4、Csd1、Csd2、Cst1、Cst2、Csh1、Csh2、Csa1、Csa2、Csa3、Csa4、Csa5、タイプII Casエフェクタータンパク質、タイプV Casエフェクタータンパク質、タイプVI Casエフェクタータンパク質、CARF、DinG、それらのホモログ、またはそれらの修飾もしくは改変バージョンが挙げられる。他の核酸プログラム可能なDNA結合タンパク質もまた、本開示の範囲内であるが、本開示に具体的に記載されていないこともあり得る。

【0145】

タンパク質または核酸に関連して本明細書中で使用される用語「組み換え体」とは、自然界には存在しないが、人間の工学の産物であるタンパク質または核酸を指す。例えば、いくつかの実施形態において、組換えタンパク質または核酸分子は、任意の天然に存在する配列と比較して、少なくとも1つ、少なくとも2つ、少なくとも3つ、少なくとも4つ、少なくとも5つ、少なくとも6つ、または少なくとも7つの変異を含むアミノ酸またはヌクレオチド配列を含む。

【0146】

「減少する」とは、少なくとも10%、25%、50%、75%、または100%の負の変化を意味する。

【0147】

「参照」は、標準又は対照条件を意味する。非限定的な例として、本明細書に記載されているような塩基編集、例えば良性または調節性の塩基編集後の遺伝子(および/またはそれにコードされるタンパク質産物)の活性または機能のアッセイは、良性または調節性の塩基編集を生じなかった遺伝子(および/またはそのコードされた産物)の活性または機能、または参照としての野生型遺伝子(および/またはそのコードされた産物)の活性または機能と比較される。1つの実施形態において、参照は野生型または健常な細胞である。

【0148】

「参照配列」は、配列比較の基礎として使用される定義済み配列である。参照配列は、特定の配列のサブセット又は全体であり得る；例えば、完全長のcDNAもしくは遺伝子配列のセグメント、または完全なcDNAまたは遺伝子配列。ポリペプチドについては、参照ポリペプチド配列の長さは、一般に、少なくとも約16アミノ酸、好ましくは少なくとも約20アミノ酸、より好ましくは少なくとも約25アミノ酸、さらにより好ましくは約35アミノ酸、約50アミノ酸、または約100アミノ酸である。核酸については、参照核酸配列の長さは、一般に、少なくとも約50ヌクレオチド、好ましくは少なくとも約60ヌクレオチド、より好ましくは少なくとも約75ヌクレオチド、さらにより好ましくは約100ヌクレオチドまたは約300ヌクレオチドまたはそれらの周辺もしくはそれらの間の任意の整数である。

【0149】

用語「RNAプログラム可能なヌクレアーゼ」および「RNA誘導ヌクレアーゼ」は、切断の標的ではない一つ以上のRNAをとともに使用される(例えば、それに結合または付随する)。ある態様において、RNAプログラム可能なヌクレアーゼは、RNAと複合体である場合、ヌクレアーゼ:RNA複合体と称され得る。典型的には、結合したRNAはガイドRNA (gRNA)

と呼ばれる。ガイドRNA (gRNA) は2つ以上のRNAの複合体として存在することもある。単一のRNA分子として存在することもある。単一のRNA分子として存在するgRNAは、単一ガイドRNA (sgRNA) と呼ばれることがあるが、「gRNA」は、単一の分子として、または二以上の分子の複合体として存在するガイドRNAのいずれかを指すために互換的に使用される。典型的には、単一のRNA種として存在するgRNAは、二つのドメイン: (1) 標的核酸と相同性を共有する(例えば、標的へのCas9複合体の結合を誘導する)ドメイン; および(2) Cas9タンパク質に結合するドメインを含む。ある態様において、ドメイン(2) は、tracrRNAとして知られる配列に対応し、ステム-ループ構造を含む。例えば、いくつかの実施形態において、ドメイン(2) は、Jinek et al., Science 337:816-821(2012)に提供されるtracrRNAと同一または相同であり、その全内容を参照により本明細書に援用する。gRNA (例えばドメイン2を含むもの)の他の例は、2013年9月6日に出願された「Switchable Cas9 Nucleases and Uses Thereof」という名称の米国仮特許出願第61/874,682号および2013年9月6日に出願された「Delivery System for Functional Nucleases」という名称の米国仮特許出願第61/874,746号に見出され得る。いくつかの実施形態において、gRNAは、二つ以上のドメイン(1) および(2)を含み、「伸長された (extended) gRNA」と称され得る。例えば、伸長されたgRNAは、本明細書に記載されるように、二つ以上のCas9タンパク質と結合し、二つ以上の異なる領域で標的核酸と結合する。gRNAは、ヌクレアーゼ/RNA複合体の標的部位への結合を媒介し、ヌクレアーゼ:RNA複合体の配列特異性を提供する、標的部位を相補するヌクレオチド配列を含む。ある態様において、RNAプログラム可能ヌクレアーゼは、(CRISPR関連システム) Cas9エンドヌクレアーゼ、例えば、*Streptococcus pyogenes*由来のCas9 (CsnI) 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., Primeaux 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." Deltcheva E., Chylinski K., Sharma C.M., Gonzales K., Chao Y., Pirzada Z.A., Eckert M.R., Vogel J., Charpentier E., Nature 471:602-607(2011)参照)。

【0150】

用語「一塩基多型 (SNP)」は、ゲノム中の特定の位置で起こる単一ヌクレオチドの変異であり、ここで、各変異は集団内で認識できるある程度(例> 1%)まで存在する。たとえば、ヒトゲノムの特定の塩基位置では、ほとんどの個体でCヌクレオチドが出現しうが、少数の個体ではその位置がAで占められている。これは、この特定の位置にSNPがあり、CまたはAという2つのヌクレオチドのバリエーションがこの位置のアリルであることを意味する。SNPは疾患に対する感受性の差の根底にある。病気の重症度や治療に対する体の反応も、遺伝的バリエーションの表れである。SNPは、遺伝子のコード領域、遺伝子の非コード領域、または遺伝子間領域(遺伝子と遺伝子の間の領域)に存在しうる。ある態様において、コード配列内のSNPは、遺伝子コードの縮重のために、産生されるタンパク質のアミノ酸配列を必ずしも変化させない。コード領域のSNPには、同義SNPと非同義SNPの2種類がある。同義SNPはタンパク質配列に影響しないが、非同義SNPはタンパク質のアミノ酸配列を変化させる。非同義SNPにはミスセンスとナンセンスの2種類がある。タンパク質をコードする領域にないSNPは、遺伝子のスプライシング、転写因子の結合、メッセンジャーRNAの分解、または非コードRNAの配列に影響を与えることがある。この種のSNPによって影響を受ける遺伝子発現はeSNP (発現SNP)と呼ばれ、遺伝子の上流または下流にあり得る。一塩基バリエーション (SNV) は、頻度に制限のない一塩基のバリエーションであり、体細胞で生じ得る。体細胞一塩基バリエーション(例えば癌により引き起こされるもの)は一塩基改変とも呼ばれ得る。

【0151】

「特異的に結合する」とは、本発明のポリペプチドおよび/または核酸分子を認識および

結合するが、試料（例えば生物学的試料）中の他の分子を実質的に認識および結合しない核酸分子、ポリペプチド、もしくはそれらの複合体（例えば、核酸プログラム可能なDNA結合ドメインおよびガイド核酸）、化合物、または分子を意味する。

【0152】

本発明の方法において有用な核酸分子は、本発明のポリペプチドまたはその断片をコードする任意の核酸分子を含む。そのような核酸分子は、内因性核酸配列と100%同一である必要はないが、典型的には実質的同一性を示す。内因性配列に対して「実質的同一性」を有するポリヌクレオチドは、典型的には、二本鎖核酸分子の少なくとも一つの鎖とハイブリダイズすることができる。本発明の方法において有用な核酸分子は、本発明のポリペプチドまたはその断片をコードする任意の核酸分子を含む。そのような核酸分子は、内因性核酸配列と100%同一である必要はないが、典型的には実質的同一性を示す。内因性配列に対して「実質的同一性」を有するポリヌクレオチドは、典型的には、二本鎖核酸分子の少なくとも一つの鎖とハイブリダイズすることができる。「ハイブリダイズする」とは、種々のストリンジェンシー条件下で相補的ポリヌクレオチド配列(例えば、本明細書に記載の遺伝子)の間、またはその一部の間に二本鎖分子を形成する対をなすことを意味する。(例えばWahl, G. M. and S. L. Berger (1987) *Methods Enzymol.* 152:399; Kimmel, A. R. (1987) *Methods Enzymol.* 152:507参照)。

【0153】

例えば、ストリンジェントな塩濃度は、通常、約750 mM未満 NaClおよび75 mMクエン酸三ナトリウム、好ましくは約500 mM未満 NaClおよび50 mMクエン酸三ナトリウム、より好ましくは約250 mM未満 NaClおよび25 mMクエン酸三ナトリウムである。低ストリンジェンシーハイブリダイゼーションは、有機溶媒、例えばホルムアミドの非存在下で得ることができ、一方、高ストリンジェンシーハイブリダイゼーションは、少なくとも約35%ホルムアミド、より好ましくは少なくとも約50%ホルムアミドの存在下で得ることができる。ストリンジェントな温度条件は、通常、少なくとも約30℃、より好ましくは少なくとも約37℃、最も好ましくは少なくとも約42℃の温度を含むであろう。ハイブリダイゼーション時間、界面活性剤(例えば、ドデシル硫酸ナトリウム(SDS))の濃度、および担体DNAの含入または排除などの様々な追加のパラメーターは、当業者によく知られている。必要に応じてこれら様々な条件を組み合わせることによって、様々なレベルのストリンジェンシーが達成される。ある実施形態では、ハイブリダイゼーションは、30℃で、750 mM NaCl、75 mMクエン酸三ナトリウムおよび1% SDS中で起こる。別の実施形態では、ハイブリダイゼーションは、37℃で、500 mM NaCl、50 mMクエン酸三ナトリウム、1% SDS、35%ホルムアミドおよび100 µg/ml変性サケ精子DNA(ssDNA)中で起こる。別の実施形態において、ハイブリダイゼーションは、42℃において、250 mM NaCl、25 mMクエン酸三ナトリウム、1% SDS、50%ホルムアミドおよび200 µg/ml ssDNA中で起こる。これらの条件の有用なバリエーションは、当業者には容易に明らかになるであろう。

【0154】

ほとんどの用途では、ハイブリダイゼーションに続く洗浄工程もまた、ストリンジェンシーが異なる。洗浄ストリンジェンシー条件は、塩濃度および温度によって定義することができる。上記のように、洗浄ストリンジェンシーは、塩濃度を低下させるか、または温度を上昇させることによって増加させることができる。例えば、洗浄工程のためのストリンジェントな塩濃度は、約30 mM未満 NaClおよび3 mMクエン酸三ナトリウムであり、約15 mM未満 NaClおよび1.5 mMクエン酸三ナトリウムであり得る。洗浄工程のためのストリンジェントな温度条件は、通常、少なくとも約25℃、より好ましくは少なくとも約42℃、さらにより好ましくは少なくとも約68℃の温度を含む。好ましい実施形態において、洗浄工程は、25℃で、30 mM NaCl、3 mMクエン酸三ナトリウム、および0.1% SDS中で行われる。より好ましい実施形態において、洗浄工程は、42℃で、15 mM NaCl、1.5 mMクエン酸三ナトリウム、および0.1% SDS中で行われる。より好ましい実施形態において、洗浄工程は、68℃で、15 mM NaCl、1.5 mMクエン酸三ナトリウム、および0.

1% SDS中で行われる。これらの条件のさらなるバリエーションは、当業者には容易に明らかになるであろう。ハイブリダイゼーション技術は、当業者によく知られており、例えば、Benton and Davis (Science 196:180, 1977); Grunstein and Hogness (Proc. Natl. Acad. Sci., USA 72:3961, 1975); Ausubel et al. (Current Protocols in Molecular Biology, Wiley Interscience, New York, 2001); Berger and Kimmel (Guide to Molecular Cloning Techniques, 1987, Academic Press, New York); および Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, New Yorkに記述されている。

【0155】

「実質的に同一である」とは、参照アミノ酸配列(例えば、本明細書に記載のアミノ酸配列のいずれか1つ)または核酸配列(例えば、本明細書に記載の核酸配列のいずれか1つ)に対して少なくとも50%の同一性を示すポリペプチドまたは核酸分子を意味する。好ましくは、そのような配列は、比較のために使用される配列とアミノ酸レベルまたは核酸において少なくとも60%、より好ましくは80%または85%、より好ましくは90%、95%または99%までもの同一性を有する。

【0156】

配列同一性は、典型的には、配列解析ソフトウェア(例えば、Genetics Computer Group, University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, Wis. 53705のSequence Analysis Software Package、BLAST、BESTFIT、COBALT、EMBOSS Needle、GAP、またはPILEUP/PRETTYBOXプログラム)を用いて測定される。そのようなソフトウェアは、種々の置換、欠失、および/または他の改変に相同性の程度を割り当てることによって、同一または類似の配列をマッチさせる。保存的置換は、典型的には、以下の群内の置換を含む:グリシン、アラニン; バリン、イソロイシン、ロイシン; アスパラギン酸、グルタミン酸、アスパラギン、グルタミン; セリン、トレオニン; リジン、アルギニン; フェニルアラニン、チロシン。同一性の程度を決定するための例示的なアプローチにおいて、BLASTプログラムを使用することができ、 e^{-3} と e^{-100} の間の確率スコアが密接に関連した配列を示す。COBALTは、たとえば次のパラメータとともに使用される:

- a) アラインメントパラメータ: Gap penalties-11,-1 and End-Gap penalties-5,-1,
- b) CDDパラメータ: Use RPS BLAST on; Blast E-value 0.003; Find Conserved columns and Recompute on
- c) クエリー・クラスタリング・パラメータ: Use query clusters on; Word Size 4; Max cluster distance 0.8; Alphabet Regular.

EMBOSS Needleは、たとえば次のパラメータで使用される。

- a) Matrix: BLOSUM62;
- b) GAP OPEN: 10;
- c) GAP EXTEND: 0.5;
- d) OUTPUT FORMAT: pair;
- e) END GAP PENALTY: false;
- f) END GAP OPEN: 10; and
- g) END GAP EXTEND: 0.5.

【0157】

「対象」とは、限定されるものではないが、ウシ、ウマ、イヌ、ヒツジまたはネコなどのヒトまたは非ヒト哺乳動物を含む哺乳動物を意味する。

【0158】

用語「標的部位」とは、核酸分子内の配列であって、核酸塩基エディターによって改変される配列をいう。一実施形態において、標的部位は、デアミナーゼまたはそれを含む融合タンパク質(例えば、シチジンまたはアデニンデアミナーゼ)によって脱アミノ化される。

【0159】

RNAプログラム可能なヌクレアーゼ(例:Cas9)は、DNA切断部位を標的とするためにRN

10

20

30

40

50

A:DNAハイブリダイゼーションを使用するので、これらのタンパク質は、原理的に、ガイドRNAによって指定されるあらゆる配列を標的とすることができる。部位特異的切断のためにCas9のようなRNAプログラム可能なヌクレアーゼを使用する方法(例えばゲノムを改変するために)は、当該技術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。これらの各々の全内容は参照により本明細書に組み入れられる)。

【0160】

本明細書中で使用される、用語「治療する(treat)」、「治療する(treating)」、「治療(treatment)」などは、疾患もしくは障害および/またはそれに関連する症状を軽減もしくは改善すること、または所望の薬理学的および/または生理学的効果を得ることを指す。障害または状態を治療することは、関連する障害、状態または症状を完全に除去することを必要としないことが理解されるであろう(完全な除去が除外されるわけでもない)。いくつかの態様において、効果は、治療的であり、すなわち、限定されるものではないが、効果は、疾患または障害および/または疾患または障害に起因する有害症状を部分的または完全に低減、減少、除去、軽減、緩和、強度低下、または治療する。ある態様において、効果は予防的であり、すなわち効果は、疾患、障害、または状態の発生または再発を保護または予防する。この目的のために、本開示の方法は、本明細書に記載されるような治療的に有効な量の組成物を投与することを含む。

【0161】

「ウラシルグリコシラーゼ阻害因子」とは、ウラシル除去修復系を阻害する因子を意味する。1つの実施形態において、該因子は、宿主ウラシル-DNAグリコシラーゼに結合してDNAからのウラシル残基の除去を妨げるタンパク質またはその断片である。

【0162】

本明細書で提供される範囲は、範囲内のすべての値(最初と最後の値を含む)、およびその間の値についての省略形であると理解される。例えば、1~50の範囲は、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, または50からなる群からの任意の数、数の組み合わせ、またはサブ範囲を含むと理解される。

【0163】

本明細書における変数の任意の定義における化学基のリストの記載は、リストされた基の任意の単一の基または組合せとしてのその変数の定義を含む。本明細書における変数または態様に関する実施形態の説明は、任意の単一の実施形態としての、または任意の他の実施形態もしくはその一部と組み合わせられた、その実施形態を含む。

【0164】

本明細書に提供される任意の組成物または方法は、本明細書に提供される他の任意の組成物および方法の1つ以上と組み合わせることができる。

【0165】

DNA編集は、遺伝子レベルで病原性変異を補正することにより疾患状態を修飾する実行可能な手段として出現してきた。最近まで、全てのDNA編集プラットフォームは、特定のゲノム部位でDNA二本鎖切断(DSB)を誘導することにより機能し、内因性DNA修復経路に依存して半確率的な態様で産物結果が決定され、遺伝的産物の複雑な集団を生じていた。

相同性誘導修復 (HDR) 経路を介すれば、ユーザーにより定義される正確な修復結果を達成できるが、多くの課題が、治療に関連する細胞型でのHDRを用いた高効率修復を妨げている。実際的にはこの経路は、競合する、誤りを起こしやすい非相同末端結合経路と比較して効率が悪い。さらに、HDRは細胞周期のG1期とS期に厳密に限定されており、有糸分裂後の細胞におけるDSBの正確な修復を妨げる。その結果、これらの集団において高い効率で、ユーザー定義のプログラム可能な態様でゲノム配列を改変することは困難または不可能であることと判明してきた。

【0166】

〔核酸塩基エディター〕

ポリヌクレオチドの標的ヌクレオチド配列を編集、修飾または改変するための塩基エディターまたは核酸塩基エディターが本明細書に開示される。本明細書に記載されるのは、ポリヌクレオチドプログラム可能なヌクレオチド結合ドメインおよび核酸塩基編集ドメインを含む核酸塩基エディターまたは塩基エディターである。ポリヌクレオチドプログラム可能なヌクレオチド結合ドメインは、結合されたガイドポリヌクレオチド(例えばgRNA)と一緒にある場合に、(結合されたガイド核酸の塩基と標的ポリヌクレオチド配列の塩基との間の相補的塩基対形成を介して) 標的ポリヌクレオチド配列に特異的に結合することができ、それによって、編集されることが所望される標的核酸配列に塩基エディターを局在化させることができる。或る実施態様では、標的ポリヌクレオチド配列は一本鎖DNAまたは二本鎖DNAを含む。或る実施態様では、標的ポリヌクレオチド配列はRNAを含む。或る実施態様では、標的ポリヌクレオチド配列はDNA-RNAハイブリッドを含む。

【0167】

〔ポリヌクレオチドプログラム可能なヌクレオチド結合ドメイン〕

用語「ポリヌクレオチドプログラム可能なヌクレオチド結合ドメイン」または「核酸プログラム可能なDNA結合タンパク質」とは、ガイドポリヌクレオチド(例えばガイドRNA)のような核酸(例えばDNAまたはRNA)と会合するタンパク質であって、その核酸がポリヌクレオチドプログラム可能なヌクレオチド結合ドメインを特定の核酸配列にガイドするところのものをいう。ある態様において、ポリヌクレオチドプログラム可能なヌクレオチド結合ドメインは、ポリヌクレオチドプログラム可能なDNA結合ドメインである。ある態様において、ポリヌクレオチドプログラム可能なヌクレオチド結合ドメインは、ポリヌクレオチドプログラム可能なRNA結合ドメインである。ある態様において、ポリヌクレオチドプログラム可能なヌクレオチド結合ドメインは、Cas9タンパク質である。ある態様において、ポリヌクレオチドプログラム可能なヌクレオチド結合ドメインは、Cpf1タンパク質である。

【0168】

CRISPRは、可動遺伝要素(ウイルス、転移因子、接合プラスミド)に対する防御を提供する適応免疫系である。CRISPRクラスターは、スペーサー、先行する可動要素に相補的な配列、および標的侵入核酸を含む。CRISPRクラスターは転写され、CRISPR RNA (crRNA) に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トランスコード小RNA (tracrRNA)、内因性リボヌクレアーゼ3 (rnc)、およびCas9タンパク質を必要とする。tracrRNAはリボヌクレアーゼ3によるpre-crRNAのプロセシングのガイドとなる。続いて、Cas9/crRNA/tracrRNAが、スペーサーに相補的な線状または環状のdsDNA標的をエンドヌクレアーゼで切断する。crRNAに相補的でない標的鎖は、最初にエンドヌクレアーゼ的に切断され、次にエキソヌクレアーゼ的に3' -5' にトリムされる。自然界では、DNA結合と切断にはタンパク質と両方のRNAが必要である。しかしながら、crRNAおよびtracrRNAの両方の側面を単一のRNA種に組み込むように、単一ガイドRNA (「sgRNA」、あるいは単に「gRNA」) を作製することができる。例えばJinek M., Chylinski K., Fonfara I., Hauer M., Doudna J. A., Charpentier E. Science 337:816-821(2012)を参照されたい(その内容全体が参照により本明細書に組み入れられる)。Cas9は、CRISPR反復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」と「非自己」を区別することを助ける。

【0169】

〔核酸塩基エディターのCas9ドメイン〕

Cas9ヌクレアーゼの配列および構造は、当業者によく知られている(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., 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)参照。その内容全体が参照により本明細書に組み入れられる。)。Cas9オソログは、限定されるものではないが、*S.pyogenes*および*S.thermophilus*を含む種々の種において記述されてきた。さらなる適切な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明らかとなり得、そのような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遺伝子座由来のCas9配列を含む。その全内容は参照により本明細書に組み込まれる。

【0170】

ある態様において、Cas9ヌクレアーゼは、不活性な(例えば不活化された) DNA切断ドメインを有し、すなわち、Cas9は、「nCas9」タンパク質(「nickase」Cas9の意)と呼ばれるニックアーゼである。ヌクレアーゼ不活性化Cas9タンパク質は、互換的に「dCas9」タンパク質(nuclease-dead Cas9の意)とも称され得る。不活性なDNA切断ドメインを有するCas9タンパク質(又はその断片)を生成する方法は公知である(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): 1173-83参照(各内容は参照により本明細書に組み込まれる))。例えば、Cas9のDNA切断ドメインは、HNHヌクレアーゼサブドメインとRuvC1サブドメインという2つのサブドメインを含むことが知られている。HNHサブドメインはgRNAに相補的な鎖を切断し、RuvC1サブドメインは非相補的な鎖を切断する。これらのサブドメイン内の変異はCas9のヌクレアーゼ活性を抑制し得る。例えば、突然変異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))。ある態様において、Cas9の断片を含むタンパク質が提供される。例えば、いくつかの実施形態において、タンパク質は、以下の2つのCas9ドメインのうちの1つを含む：(1) Cas9のgRNA結合ドメイン；(2) Cas9のDNA切断ドメイン。ある態様において、Cas9またはその断片を含むタンパク質は、「Cas9バリエーション」と称される。Cas9バリエーションは、Cas9またはその断片と相同性を共有する。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。いくつかの実施形態において、Cas9変異体は、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またはそれ以上のアミノ酸変化を有し得る。いくつかの実施形態において、Cas9バリエーションは、Cas9の断片(例えばgRNA結合ドメインまたはDNA切断ドメイン)を含み、その断片は、野生型Cas9の対応する断片と少なくとも約70%同一であり、少なくとも約80%同一であり、少なくとも約90%同一であり、少なくとも約95%同一であり、少なくとも約96%同一であり

、少なくとも約97%同一であり、少なくとも約98%同一であり、少なくとも約99%同一であり、少なくとも約99.5%同一であり、または少なくとも約99.9%同一である。ある態様において、断片は、対応する野生型Cas9のアミノ酸長の少なくとも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。

【0171】

ある態様において、断片は、長さが少なくとも100アミノ酸である。ある態様において、断片は、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アミノ酸の長さである。

【0172】

ある態様において、野生型Cas9は、*Streptococcus pyogenes*からのCas9に対応する(NCBI参照配列:NC_17053.1、以下のヌクレオチド及びアミノ酸配列):

```
ATGGATAAGAAATACTCAATAGGCTTAGATATCGGCACAAATAGCGTCGGATGGGCGGT
GATCACTGATGATTATAAGGTTCCGTCTAAAAAGTTCAAGGTTCTGGGAAATACAGACC
GCCACAGTATCAAAAAAATCTTATAGGGGCTCTTTTATTTGGCAGTGGAGAGACAGCG
GAAGCGACTCGTCTCAAACGGACAGCTCGTAGAAGGTATACACGTCGGAAGAATCGTAT
TTGTTATCTACAGGAGATTTTTTCAAATGAGATGGCGAAAGTAGATGATAGTTTCTTTC
ATCGACTTGAAGAGTCTTTTTTGGTGGAAGAAGACAAGAAGCATGAACGTCATCCTATT
TTTGGAAATATAGTAGATGAAGTTGCTTATCATGAGAAATATCCAACATCTATCATCT
GCGAAAAAATTGGCAGATTCTACTGATAAAGCGGATTTGCGCTTAATCTATTTGGCCT
TAGCGCATATGATTAAGTTTCGTGGTCATTTTTTGGATTGAGGGAGATTTAAATCCTGAT
AATAGTGATGTGGACAACTATTTATCCAGTTGGTACAAATCTACAATCAATTATTTGA
AGAAAACCCTATTAACGCAAGTAGAGTAGATGCTAAAGCGATTCTTTCTGCACGATTGA
GTAAATCAAGACGATTAGAAAATCTCATTGCTCAGCTCCCCGGTGAGAAGAGAAATGGC
TTGTTTGGGAATCTCATTGCTTTGTCATTGGGATTGACCCCTAATTTTAAATCAAATTT
GATTTGGCAGAAGATGCTAAATTACAGCTTTTCAAAGATACTTACGATGATGATTTAGA
TAATTTATTTGGCGCAAATTGGAGATCAATATGCTGATTTGTTTTTGGCAGCTAAGAATT
TATCAGATGCTATTTTACTTTTCAAGATATCCTAAGAGTAAATAGTGAAATAACTAAGGCT
CCCCTATCAGCTTCAATGATTAAGCGCTACGATGAACATCATCAAGACTTGACTCTTTT
AAAAGCTTTAGTTTCGACAACAACCTCCAGAAAAGTATAAAGAAATCTTTTTTGATCAAT
CAAAAAACGGATATGCAGGTTATATTGATGGGGGAGCTAGCCAAGAAGAATTTTATAAA
TTTATCAAACCAATTTTAGAAAAAATGGATGGTACTGAGGAATTATTGGTGAAACTAAA
TCGTGAAGATTTGCTGCGCAAGCAACGGACCTTTGACAACGGCTCTATTCCCCATCAAA
TTCACCTGGGTGAGCTGCATGCTATTTTGAGAAGACAAGAAGACTTTTATCCATTTTAA
AAAGACAATCGTGAGAAGATTGAAAAAATCTTGACTTTTCGAATTCCTTATTATGTTGG
TCCATTGGCGCGTGGAATAGTCGTTTTGCATGGATGACTCGGAAGTCTGAAGAAACAA
TTACCCCATGGAATTTTGAAGAAGTTGTCGATAAAGGTGCTTCAGCTCAATCATTTATT
GAACGCATGACAACTTTGATAAAAATCTTCCAATGAAAAAGTACTACCAAAACATAG
TTTGCTTTATGAGTATTTTACGGTTTATAACGAATTGACAAAGGTCAAATATGTTACTG
AGGGAATGCGAAAACCAGCATTTCTTTTCAAGGTGAACAGAAGAAAGCCATTGTTGATTTA
CTTTTCAAAACAAATCGAAAAGTAACCGTTAAGCAATTAAGAAGATTATTTCAAAAA
AATAGAATGTTTTGATAGTGTTGAAATTTCAAGGAGTTGAAGATAGATTTAATGCTTCAT
TAGGCGCCTACCATGATTTGCTAAAAATTATTAAGATAAAGATTTTTTGGATAATGAA
GAAATGAAGATATCTTAGAGGATATTGTTTTAACATTGACCTTATTTGAAGATAGGGG
GATGATTGAGGAAAGACTTAAACATATGCTCACCTCTTTGATGATAAGGTGATGAAAC
AGCTTAAACGTCGCCGTTATACTGGTTGGGGACGTTTGTCTCGAAAATTGATTAATGGT
ATTAGGGATAAGCAATCTGGCAAAACAATATTAGATTTTTTGAATCAGATGGTTTTGC
```

10

20

30

40

50

CAATCGCAATTTTATGCAGCTGATCCATGATGATAGTTTGACATTTAAAGAAGATATTC
AAAAAGCACAGGTGTCTGGACAAGGCCATAGTTTACATGAACAGATTGCTAACTTAGCT
GGCAGTCCTGCTATTAAAAAAGGTATTTTACAGACTGTAAAAATTGTTGATGAACTGGT
CAAAGTAATGGGGCATAAGCCAGAAAAATATCGTTATTGAAATGGCACGTGAAAATCAGA
CAACTCAAAAGGGCCAGAAAAATTTCGCGAGAGCGTATGAAACGAATCGAAGAAGGTATC
AAAGAATTAGGAAGTCAGATTCTTAAAGAGCATCCTGTTGAAAATACTCAATTGCAAAA
TGAAAAGCTCTATCTCTATTATCTACAAAATGGAAGAGACATGTATGTGGACCAAGAAT
TAGATATTAATCGTTTAAAGTGATTATGATGTCGATCACATTGTTCCACAAAGTTTCATT
AAAGACGATTCAATAGACAATAAGGTACTAACGCGTTCTGATAAAAATCGTGGTAAATC
GGATAACGTTCCAAGTGAAGAAGTAGTCAAAAAGATGAAAACTATTGGAGACAACCTTC
TAAACGCCAAGTTAATCACTCAACGTAAGTTTGATAATTTAACGAAAGCTGAACGTGGA
GTTTTGAGTGAACCTTGATAAAGCTGGTTTTATCAAACGCCAATTGGTTGAAACTCGCCA
AATCACTAAGCATGTGGCACAAATTTTGGATAGTCGCATGAATACTAAATACGATGAAA
ATGATAAACTTATTCGAGAGGTTAAAGTGATTACCTTAAAATCTAAATTAGTTTCTGAC
TTCCGAAAAGATTTCCAATTCTATAAAGTACGTGAGATTAACAATTACCATCATGCCCA
TGATGCGTATCTAAATGCCGTCGTTGGAACTGCTTTGATTAAGAAATATCCAAAACCTTG
AATCGGAGTTTGTCTATGGTGATTATAAAGTTTATGATGTTTCGTAAAATGATTGCTAAG
TCTGAGCAAGAAATAGGCAAAGCAACCGCAAAATATTTCTTTTACTCTAATATCATGAA
CTTCTTCAAAACAGAAATTACACTTGCAAATGGAGAGATTTCGCAAACGCCCTCTAATCG
AAACTAATGGGGAACTGGAGAAATTGTCTGGGATAAAGGGCGAGATTTTGCCACAGTG
CGCAAAGTATTGTCCATGCCCAAGTCAATATTGTCAAGAAAACAGAAGTACAGACAGG
CGGATTCTCCAAGGAGTCAATTTTACCAAAAAGAAATTCGGACAAGCTTATTGCTCGTA
AAAAAGACTGGGATCCAAAAAATATGGTGGTTTTGATAGTCCAACGGTAGCTTATTCA
GTCCTAGTGTTGCTAAGGTGGAAAAAGGGAAATCGAAGAAGTTAAAATCCGTTAAAGA
GTTACTAGGGATCACAATTATGGAAAGAAGTTCCTTTGAAAAAATCCGATTGACTTTT
TAGAAGCTAAAGGATATAAGGAAGTTAAAAAAGACTTAATCATTAAACTACCTAAATAT
AGTCTTTTTGAGTTAGAAAACGGTCGTAAACGGATGCTGGCTAGTGCCGGAGAATTACA
AAAAGGAAATGAGCTGGCTCTGCCAAGCAAATATGTGAATTTTTTATATTTAGCTAGTC
ATTATGAAAAGTTGAAGGGTAGTCCAGAAGATAACGAACAAAAACAATTGTTTGTGGAG
CAGCATAAGCATTATTTAGATGAGATTATTGAGCAAATCAGTGAATTTTCTAAGCGTGT
TATTTTAGCAGATGCCAATTTAGATAAAGTTCTTAGTG CATATAACAAACATAGAGACA
AACCAATACGTGAACAAGCAGAAAATATTATTCATTTATTTACGTTGACGAATCTTGGA
GCTCCCGCTGCTTTTAAATATTTTGATACAACAATTGATCGTAAACGATATACGTCTAC
AAAAGAAGTTTTAGATGCCACTCTTATCCATCAATCCATCACTGGTCTTTATGAAACAC
GCATTGATTTGAGTCAGCTAGGAGGTGACTGA.

10

20

30

40

50

MDKKYSIGLDIGTNSVGWAVITDDYKVPSKKFKVLGNTDRHSIKKNLIGALLFGSGETAETRL
 KRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVAY
 HEKYPTIYHLRKKLADSTDKADLRILIYALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQIY
 NQLFEENPINASRVDAKAILSARLSKSRLENLIAQLPGEKRNGLFGNLIALSLGLTPNFKSNF
 DLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNSEITKAPLSAS
 MIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMD
 GTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRI
 PYYVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGASQSFIERMTNFDKNLPNEKVLPHKS
 LLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFD
 SVEISGVEDRFNASLGAYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDRGMIEERLKTYA
 HLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDFANRNFQMQLIHDDSLTF
 KEDIQKAQVSGQGHS~~LHEQIANLAGSPA~~IKKGILQTVKIVDELVKVMGHKPENIVIE~~MARE~~NQT
 TQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRL
 SDYDVDHIVPQSFIKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKMKNYWRQLLNAKLITQRKF
 DNLTKAERGGLSEL~~DKAGFIKROLVETROITKHVAQILD~~SRMNTKYDENDKLI~~REVKVITL~~KSK
 LVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG~~TALIKKYPKLESEFVYG~~DYKVYDVRKMI~~AKS~~
 EQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL~~SM~~
 PQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEK~~KG~~
 SKKLKSVKELLGITIMERSSEFEKNPIDFLEAKGYKEVKKDIIKLPKYSLFELENG~~RKRMLASA~~
 GELQKGNELALPSKYVNFYLYLASHYEKLKGSPE~~DNEQKQLFVEQHKHYLDEIEQISEF~~SKRVI
 LADANLDKVL~~SAYNKH~~RD~~KPIREQAENI~~IHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDA
 TLIHQ~~SITGLYETRIDLSQL~~GGD

10

20

30

(単一下線：HNHドメイン；二重下線：RuvCドメイン)。

【0173】

いくつかの実施形態において、野生型Cas9は、以下のヌクレオチドおよび/またはアミノ酸配列に対応するか、またはこれらを含む：

ATGGATAAAAAGTATTCTATTGGTTTAGACATCGGCACTAATTCCGTTGGATGGGCTGT
 CATAACCGATGAATACAAAGTACCTTCAAAGAAATTTAAGGTGTTGGGGAACACAGACC
 GTCATTCGATTAAAAAGAATCTTATCGGTGCCCTCCTATTCGATAGTGGCGAAACGGCA
 GAGGCGACTCGCCTGAAACGAACCGCTCGGAGAAGGTATACACGTCGCAAGAACCGAAT
 ATGTTACTTACAAGAAATTTTTAGCAATGAGATGGCCAAAGTTGACGATTCTTTCTTTC
 ACCGTTTGGGAAGAGTCCTTCCTTGTCGAAGAGGACAAGAAACATGAACGGCACCCCATC
 TTTGGAAACATAGTAGATGAGGTGGCATATCATGAAAAGTACCCAACGATTTATCACCT
 CAGAAAAAAGCTAGTTGACTCAACTGATAAAGCGGACCTGAGGTTAATCTACTTGGCTC
 TTGCCCATATGATAAAGTTCCGTGGGCACTTTCTCATTGAGGGTGATCTAAATCCGGAC
 AACTCGGATGTCGACAACTGTTTCATCCAGTTAGTACAAACCTATAATCAGTTGTTTGA
 AGAGAACCCTATAAATGCAAGTGGCGTGGATGCGAAGGCTATTCTTAGCGCCCGCCTCT
 CTAAATCCCGACGGCTAGAAAACCTGATCGCACAAATTACCGGAGAGAAGAAAAATGGG
 TTGTTCCGGTAACCTTATAGCGCTCTCACTAGGCCTGACACCAAATTTTAAAGTCGAACTT
 CGACTTAGCTGAAGATGCCAAATTGCAGCTTAGTAAGGACACGTACGATGACGATCTCG

40

50

ACAATCTACTGGCACAAATTGGAGATCAGTATGCGGACTTATTTTTGGCTGCCAAAAAC
CTTAGCGATGCAATCCTCCTATCTGACATACTGAGAGTTAATACTGAGATTACCAAGGC
GCCGTTATCCGCTTCAATGATCAAAAGGTACGATGAACATCACCAAGACTTGACACTTC
TCAAGGCCCTAGTCCGTCAGCAACTGCCTGAGAAATATAAGGAAATATTCTTTGATCAG
TCGAAAAACGGGTACGCGAGGTTATATTGACGGCGGAGCGAGTCAAGAGGAATTCTACAA
GTTTATCAAACCCATATTAGAGAAGATGGATGGGACGGAAGAGTTGCTTGTAACACTCA
ATCGCGAAGATCTACTGCGAAAGCAGCGGACTTTTCGACAACGGTAGCATTCCACATCAA
ATCCACTTAGGCGAATTGCATGCTATACTTAGAAGGCAGGAGGATTTTTATCCGTTCT
CAAAGACAATCGTGAAAAGATTGAGAAAATCCTAACCTTTTCGCATACCTTACTATGTGG
GACCCCTGGCCCGAGGGAACCTCTCGGTTTCGCATGGATGACAAGAAAGTCCGAAGAAACG
ATTACTCCATGGAATTTTGAGGAAGTTGTCGATAAAGGTGCGTCAGCTCAATCGTTCAT
CGAGAGGATGACCAACTTTGACAAGAATTTACCGAACGAAAAAGTATTGCCTAAGCACA
GTTTACTTTACGAGTATTTACAGTGTACAATGAACTCACGAAAGTTAAGTATGTCACT
GAGGGCATGCGTAAACCCGCCTTTCTAAGCGGAGAACAGAAGAAAGCAATAGTAGATCT
GTTATTCAAGACCAACCGCAAAGTGACAGTTAAGCAATTGAAAGAGGACTACTTTAAGA
AAATTGAATGCTTCGATTCTGTGAGATCTCCGGGGTAGAAGATCGATTTAATGCGTCA
CTTGGTACGTATCATGACCTCCTAAAGATAATTAAAGATAAGGACTTCCTGGATAACGA
AGAGAATGAAGATATCTTAGAAGATATAGTGTTGACTCTTACCCTCTTTGAAGATCGGG
AAATGATTGAGGAAAGACTAAAAACATACGCTCACCTGTTTCGACGATAAGGTTATGAAA
CAGTTAAAGAGGCGTCGCTATACGGGCTGGGGACGATTGTCGCGGAAACTTATCAACGG
GATAAGAGACAAGCAAAGTGGTAAAACCTATTCTCGATTTTCTAAAGAGCGACGGCTTCG
CCAATAGGAACTTTATGCAGCTGATCCATGATGACTCTTTAACCTTCAAAGAGGATATA
CAAAAGGCACAGGTTTCCGGACAAGGGGACTCATTGCACGAACATATTGCGAATCTTGC
TGGTTCCGACCCATCAAAAAGGGCATACTCCAGACAGTCAAAGTAGTGGATGAGCTAG
TTAAGGTCATGGGACGTCACAAACCGGAAAACATTGTAATCGAGATGGCACGCGAAAAT
CAAACGACTCAGAAGGGGCAAAAAAACAGTCGAGAGCGGATGAAGAGAATAGAAGAGG
GTATTAAAGAACTGGGCAGCCAGATCTTAAAGGAGCATCCTGTGGAAAATACCCAATTG
CAGAACGAGAACTTTACCTCTATTACCTACAAAATGGAAGGGACATGTATGTTGATCA
GGAACCTGGACATAAACCGTTTATCTGATTACGACGTCGATCACATTGTACCCCAATCCT
TTTTGAAGGACGATTCAATCGACAATAAAGTGCTTACACGCTCGGATAAGAACCGAGGG
AAAAGTGACAATGTTCCAAGCGAGGAAGTCGTAAAGAAAATGAAGAACTATTGGCGGCA
GCTCCTAAATGCGAACTGATAACGCAAAGAAAGTTTCGATAACTTAACTAAAGCTGAGA
GGGGTGGCTTGTCTGAACTTGACAAGGCCGGATTTATTAAACGTCAGCTCGTGGAACCC
CGCCAAATCACAAAGCATGTTGCACAGATACTAGATTCCCGAATGAATACGAAATACGA
CGAGAACGATAAGCTGATTCGGGAAGTCAAAGTAATCACTTTAAAGTCAAATTTGGTGT
CGGACTTCAGAAAGGATTTTCAATTCTATAAAGTTAGGGAGATAAATAACTACCACCAT
GCGCACGACGCTTATCTTAATGCCGTCGTAGGGACCGCACTCATTAAGAAATACCCGAA
GCTAGAAAGTGAGTTTGTGTATGGTGATTACAAAGTTTATGACGTCGGTAAGATGATCG
CGAAAAGCGAACAGGAGATAGGCAAGGCTACAGCCAAATACTTCTTTTATTCTAACATT
ATGAATTTCTTTAAGACGGAAATCACTCTGGCAAACGGAGAGATACGCAAACGACCTTT
AATTGAAACCAATGGGGAGACAGGTGAAATCGTATGGGATAAGGGCCGGGACTTCGCGA
CGGTGAGAAAAGTTTTGTCCATGCCCAAGTCAACATAGTAAAGAAAACCTGAGGTGCAG
ACCGGAGGGTTTTCAAAGGAATCGATTCTTCAAAAAGGAATAGTGATAAGCTCATCGC
TCGTAAAAAGGACTGGGACCCGAAAAAGTACGGTGGCTTCGATAGCCCTACAGTTGCCT
ATTCTGTCTAGTAGTGGAAGTTGAGAAGGGGAAAATCCAAGAACTGAAGTCAGTC
AAAGAATTATTGGGGATAACGATTATGGAGCGCTCGTCTTTTGAAAAGAACCCCATCGA
CTTCCTTGAGGCGAAAGGTTACAAGGAAGTAAAAAAGGATCTCATAATTAACTACCAA
AGTATAGTCTGTTTGAGTTAGAAAATGGCCGAAAACGGATGTTGGCTAGCGCCGGAGAG
CTTCAAAAGGGGAACGAACTCGCACTACCGTCTAAATACGTGAATTTCTGTATTTAGC
GTCCCATTACGAGAAGTTGAAAGGTTACCTGAAGATAACGAACAGAAGCAACTTTTTG

10

20

30

40

50

TTGAGCAGCACAAACATTATCTCGACGAAATCATAGAGCAAATTTTCGGAATTCAGTAAG
AGAGTCATCCTAGCTGATGCCAATCTGGACAAAGTATTAAGCGCATAACAAGCACAG
GGATAAACCCATACGTGAGCAGGCGGAAAATATTATCCATTTGTTTACTCTTACCAACC
TCGGCGCTCCAGCCGCATTCAAGTATTTTGACACAACGATAGATCGCAAACGATACACT
TCTACCAAGGAGGTGCTAGACGCGACACTGATTCACCAATCCATCACGGGATTATATGA
AACTCGGATAGATTTGTACAGCTTGGGGGTGACGGATCCCCAAGAAGAAGAGGAAAG
TCTCGAGCGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAG
GATGACGATGACAAGGCTGCAGGA

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRL
KRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVAY
HEKYPTIYHLRKKLVDSTDKADLRLLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTY
NQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALLSLGLTPNFKSNF
DLAEDAKLQLSKDITYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSAS
MIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMD
GTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPFLKDNREKIEKILTFRI
PYYVGPLARGNSRFAWMTRKSEETITPWNFEVVVDKGASAQSFIERMTNFDKNLPNEKVLPHKS
LLEYEFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFD
SVEISGVEDRFNASLGTYHDLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYA
HLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTF
KEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILOTVKVVDDELVKVMGRHKPENIVIEMARENQ
TTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLONEKLYLYYLQNGRDMYVDQELDINR
LSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKKNYWRQLLNAKLITQRK
FDNLTKAERGGLSELDKAGFIKROLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS
KLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGITALIKKYPKLESEFVYGDYKVYDVRKMIAK
SEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL
MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVEKG
KSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKDLIIKLPKYSLFELENGRKRMLAS
AGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDEQKQLFVEQHKHYLDEIIIEQISEFSKRV
ILADANLDKVL SAYNKHDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLD
ATLIHQSIITGLYETRIDLSQLGGD

10

20

30

40

(単一下線：HNHドメイン；二重下線：RuvCドメイン)。

【0174】

ある態様において、野生型Cas9は、Streptococcus pyogenes由来のCas9に対応する
(NCBI参照配列:NC_2737.2 (以下のヌクレオチド配列)。およびUniprot参照配列:Q99
ZW2 (以下のアミノ酸配列)。

ATGGATAAGAAATACTCAATAGGCTTAGATATCGGCACAAATAGCGTCGGATGGGCGGT
GATCACTGATGAATATAAGGTTCCGTCTAAAAAGTTCAAGGTTCTGGGAAATACAGACC
GCCACAGTATCAAAAAAATCTTATAGGGGCTCTTTTATTTGACAGTGGAGAGACAGCG
GAAGCGACTCGTCTCAAACGGACAGCTCGTAGAAGGTATACACGTCGGAAGAATCGTAT
TTGTTATCTACAGGAGATTTTTTCAAATGAGATGGCGAAAGTAGATGATAGTTTCTTTC

50

ATCGACTTGAAGAGTCTTTTTTGGTGGGAAGAAGACAAGAAGCATGAACGTCATCCTATT
TTTGGAAATATAGTAGATGAAGTTGCTTATCATGAGAAATATCCAACCTATCTATCATCT
GCGAAAAAATTGGTAGATTCTACTGATAAAGCGGATTTGCGCTTAATCTATTTGGCCT
TAGCGCATATGATTAAGTTTCGTGGTCATTTTTTTGATTGAGGGAGATTTAAATCCTGAT
AATAGTGATGTGGACAACTATTTATCCAGTTGGTACAAACCTACAATCAATTATTTGA
AGAAAACCCTATTAACGCAAGTGGAGTAGATGCTAAAGCGATTCTTTCTGCACGATTGA
GTAAATCAAGACGATTAGAAAATCTCATTGCTCAGCTCCCCGGTGAGAAGAAAAATGGC
TTATTTGGGAATCTCATTGCTTTGTCATTGGGTTTGACCCCTAATTTTAAATCAAATTTT
GATTTGGCAGAAGATGCTAAATTACAGCTTTCAAAAGATACTTACGATGATGATTTAGA
TAATTTATTGGCGCAAATTGGAGATCAATATGCTGATTTGTTTTTGGCAGCTAAGAATT
TATCAGATGCTATTTTACTTTTCAAGATATCCTAAGAGTAAATACTGAAATAACTAAGGCT
CCCCTATCAGCTTCAATGATTAAACGCTACGATGAACATCATCAAGACTTGACTCTTTT
AAAAGCTTTAGTTTCGACAACAACCTCCAGAAAAGTATAAAGAAATCTTTTTTGATCAAT
CAAAAAACGGATATGCAGGTTATATTGATGGGGGAGCTAGCCAAGAAGAATTTTATAAA
TTTATCAAACCAATTTTAGAAAAAATGGATGGTACTGAGGAATTATTGGTGAAACTAAA
TCGTGAAGATTTGCTGCGCAAGCAACGGACCTTTGACAACGGCTCTATTCCCCATCAAA
TTCACCTGGGTGAGCTGCATGCTATTTTGAGAAGACAAGAAGACTTTTATCCATTTTAA
AAAGACAATCGTGAGAAGATTGAAAAAATCTTGACTTTTCGAATTCCTTATTATGTTGG
TCCATTGGCGCGTGGAATAGTCGTTTTGTCATGGATGACTCGGAAGTCTGAAGAAACAA
TTACCCCATGGAATTTTGAAGAAGTTGTCGATAAAGGTGCTTCAGCTCAATCATTTATT
GAACGCATGACAACTTTGATAAAAATCTTCAAATGAAAAAGTACTACCAAAACATAG
TTTGCTTTATGAGTATTTTACGGTTTATAACGAATTGACAAAGGTCAAATATGTTACTG
AAGGAATGCGAAAACCAGCATTTCTTTTCAAGGTGAACAGAAGAAAGCCATTGTTGATTAA
CTCTTCAAAACAAATCGAAAAGTAACCGTTAAGCAATTAAGAAGAGATTATTTCAAAAA
AATAGAATGTTTTGATAGTGTTGAAATTTTCAAGGAGTTGAAGATAGATTTAATGCTTCAT
TAGGTACCTACCATGATTTGCTAAAAATTATTAAGATAAAGATTTTTTGGATAATGAA
GAAAATGAAGATATCTTAGAGGATATTGTTTTAACATTGACCTTATTTGAAGATAGGGA
GATGATTGAGGAAAGACTTAAACATATGCTCACCTCTTTGATGATAAGGTGATGAAAC
AGCTTAAACGTCGCCGTTATACTGGTTGGGGACGTTTGTCTCGAAAATTGATTAATGGT
ATTAGGGATAAGCAATCTGGCAAAACAATATTAGATTTTTTGAATCAGATGGTTTTGCT
CAATCGCAATTTTATGCAGCTGATCCATGATGATAGTTTGACATTTAAAGAAGACATTC
AAAAAGCACAAGTGTCTGGACAAGGCGATAGTTTACATGAACATATTGCAAATTTAGCT
GGTAGCCCTGCTATTAAAAAAGGTATTTTACAGACTGTAAAAGTTGTTGATGAATTGGT
CAAAGTAATGGGGCGGCATAAGCCAGAAAATATCGTTATTGAAATGGCACGTGAAAATC
AGACAACCTCAAAGGGCCAGAAAATTCGCGAGAGCGTATGAAACGAATCGAAGAAGGT
ATCAAAGAATTAGGAAGTCAGATTCTTAAAGAGCATCCTGTTGAAAATACTCAATTGCA
AAATGAAAAGCTCTATCTCTATTATCTCCAAAATGGAAGAGACATGTATGTGGACCAAG
AATTAGATATTAATCGTTTAAAGTGATTATGATGTCGATCACATTGTTCCACAAAGTTTC
CTTAAAGACGATTCAATAGACAATAAGGTCTTAAACGCGTTCTGATAAAAATCGTGGTAA
ATCGGATAACGTTCCAAGTGAAGAAGTAGTCAAAAAGATGAAAACTATTGGAGACAAC
TTCTAAACGCCAAGTTAATCACTCAACGTAAGTTTGATAATTTAACGAAAGCTGAACGT
GGAGGTTTGAGTGAACCTTGATAAAGCTGGTTTTATCAAACGCCAATTGGTTGAAACTCG
CCAAATCACTAAGCATGTGGCACAAATTTTGGATAGTCGCATGAATACTAAATACGATG
AAAATGATAAACTTATTCGAGAGGTTAAAGTGATTACCTTAAATCTAAATTAGTTTCT
GACTTCCGAAAAGATTTCCAATTCTATAAAGTACGTGAGATTAACAATTACCATCATGC
CCATGATGCGTATCTAAATGCCGTCGTTGGAACCTGCTTTGATTAAGAAATATCCAAAAC
TTGAATCGGAGTTTGTCTATGGTGATTATAAAGTTTATGATGTTTCGTAAAATGATTGCT
AAGTCTGAGCAAGAAATAGGCCAAAGCAACCGCAAAATATTTCTTTTACTCTAATATCAT
GAACCTCTTCAAACAGAAATTACACTTGCAAATGGAGAGATTGCGAAACGCCCTCTAA
TCGAAACTAATGGGGAACTGGAGAAATTGTCTGGGATAAAGGGCGAGATTTTGCCACA

10

20

30

40

50

GTGCGCAAAGTATTGTCCATGCCCCAAGTCAATATTGTCAAGAAAACAGAAGTACAGAC
 AGGCGGATTCTCCAAGGAGTCAATTTTACCAAAAAGAAATTCGGACAAGCTTATTGCTC
 GTAAAAAAGACTGGGATCCAAAAAATATGGTGGTTTTGATAGTCCAACGGTAGCTTAT
 TCAGTCCTAGTGGTTGCTAAGGTGGAAAAAGGGAAATCGAAGAAGTTAAAATCCGTTAA
 AGAGTTACTAGGGATCACAATTATGGAAAGAAGTTCCTTTGAAAAAAATCCGATTGACT
 TTTTAGAAGCTAAAGGATATAAGGAAGTTAAAAAAGACTTAATCATTAACTACCTAAA
 TATAGTCTTTTTGAGTTAGAAAACGGTCGTAAACGGATGCTGGCTAGTGCCGGAGAATT
 AAAAAAGGAAATGAGCTGGCTCTGCCAAGCAAATATGTGAATTTTTTATATTTAGCTA
 GTCATTATGAAAAGTTGAAGGGTAGTCCAGAAGATAACGAACAAAAACAATTGTTTGTG
 GAGCAGCATAAGCATTATTTAGATGAGATTATTGAGCAAATCAGTGAATTTTCTAAGCG
 TGTTATTTTAGCAGATGCCAATTTAGATAAAGTTCTTAGTGCATATAACAAACATAGAG
 ACAAACCAATACGTGAACAAGCAGAAAATATTATTCAATTTATTTACGTTGACGAATCTT
 GGAGCTCCCGCTGCTTTTAAATATTTTGATACAACAATTGATCGTAAACGATATACGTC
 TACAAAAGAAGTTTTAGATGCCACTCTTATCCATCAATCCATCACTGGTCTTTATGAAA
 CACGCATTGATTTGAGTCAGCTAGGAGGTGACTGA

10

MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRL
 KRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVAY
 HEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTY
 NQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKS NF
 DLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSAS
 MIKRYDEHHQDLTLLKALVRQQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMD
 GTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRI
 PYYVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNEKVLPHKS
 LLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFD
 SVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTLFEDREMIEERLKTYA
 HLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDFANRNFQMQLIHDDSLTF
 KEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVDELVKVMGRHKPENIVIAMARENQ
 TTQKGQKNSRERMKRIEIEGIELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINR
 LSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKNYWRQLLNAKLITQRK
FDNLTKAERGGGLSELDKAGFIKRQLVETROI TKHVAQILDSRMNTKYDENDKLIREVKVITLKS
KLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIAK
SEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL
MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVEKG
 KSKKLKSVKELLGITIMERSSEKPNIDFLEAKGYKEVKDLIIKLPKYSLELENGRKRMLAS
 AGELQKGNELALPSKYVNFLYLASHYEKLKGS PEDNEQKQLFVEQHKHYLDEII EQISEFSKR
 VILADANLDKVL SAYNKH RDKPIREQAENI IHLFTLTNLGAPAAF KYFDTTIDRKRYTSTKEVLD
 ATLIHQSI TGLYETRIDLSQLGGD

20

30

40

(単一下線：HNHドメイン；二重下線：RuvCドメイン)。

【0175】

50

いくつかの実施形態において、Cas9は、*Corynebacterium ulcerans* (NCBI Refs: NC_015683.1, NC_017317.1); *Corynebacterium diphtheria* (NCBI Refs: NC_016782.1, NC_016786.1); *Spiroplasma syrphidicola* (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 altica* (NCBI Ref: NC_018010.1); *Psychroflexus torquus* (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) もしくは *Neisseria meningitidis* (NCBI Ref: YP_002342100.1) からのCas9を表し、または他の生物からのCas9を表す。

10

【0176】

いくつかの実施形態において、Cas9ドメインはD10A突然変異を含み、一方、上記で提供したアミノ酸配列における位置840における残基、または本明細書で提供されるアミノ酸配列のいずれかにおける対応する位置における残基は、ヒスチジンのままである。

【0177】

ある態様において、dCas9は、Cas9ヌクレアーゼ活性を不活性化する1以上の突然変異を有するCas9アミノ酸配列に対応するか、またはその一部もしくは全体を含む。例えば、いくつかの実施形態において、dCas9ドメインは、D10AおよびH840A突然変異または別のCas9における対応する突然変異を含む。いくつかの実施形態において、dCas9は、以下のdCas9 (D10AおよびH840A)のアミノ酸配列を含む：

20

30

40

50

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRL
KRTARRRYTRRKNRICYLQEIFSNEMAKVDDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAY
HEKYPTIYHLRKKLVDSTDKADLRILIYLAHAMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTY
NQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNF
DLAEDAKLQLSKDITYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSAS
MIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMD
GTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRI
PYYVGPLARGNSRFAMWTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPHKS
LLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFD
SVEISGVEDRFNASLGTYHDLKKIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYA
HLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTIIDFLKSDGFANRNFQMQLIHDDSLTF
KEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDLVKVMGRHKPENIVIEMARENO
TTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINR
LSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKMKNYWRQLLNAKLITQRK
FDNLTKAERGGLSELDKAGFIKRQLVETROITKHVAQILDSRMNTKYDENDKLIREVKVITLKS
KLVSDFRKDFQFYKVRREINNYHHAHDAYLNAVVGITALIKKYPKLESEFVYGDYKYVDVRKMIAK
SEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLS
MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEKG
KSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLEFENGRKRMLAS
AGELQKGNELALPSKYVNFLYLASHYEKLKGSPEQKQLFVEQHKHYLDEIIEQISEFSKRV
ILADANLDKVL SAYNKH RDKPIREQAENIIHLEFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLD
ATLIHQSIITGLYETRIDLSQLGGD

10

20

30

(単一下線：HNHドメイン；二重下線：RuvCドメイン)。

【0178】

いくつかの実施形態において、Cas9ドメインはD10A突然変異を含み、一方、上記で提供したアミノ酸配列における位置840における残基、または本明細書で提供されるアミノ酸配列のいずれかにおける対応する位置における残基は、ヒスチジンのままである。

【0179】

他の実施形態において、例えばヌクレアーゼ不活性化Cas9 (dCas9) をもたらす、D10AおよびH840A以外の突然変異を有するdCas9バリエーションが提供される。このような突然変異は、例えば、D10およびH840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サブドメインにおける置換)を含む。ある態様において、dCas9のバリエーションまたはホモログであって、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%の同一性を有するものが提供される。ある態様において、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アミノ酸またはそれ以上だけ短いまたは長いアミノ酸配列を有するdCas9のバリエーションが提供される。

40

【0180】

50

いくつかの実施形態において、本明細書に提供されるCas9融合タンパク質は、Cas9タンパク質の全長アミノ酸配列、例えば、本明細書に提供されるCas9配列の1つを含む。しかしながら、他の実施形態において、本明細書に提供される融合タンパク質は、全長Cas9配列を含まず、その1つ以上の断片のみを含む。好適なCas9ドメインおよびCas9断片の例示的アミノ酸配列が本明細書に提供され、Cas9ドメインおよび断片のさらなる好適な配列は、当業者には明らかであろう。

【0181】

Cas9タンパク質は、そのガイドRNAに相補的な特定のDNA配列にCas9タンパク質をガイドする、ガイドRNAと結合することができる。ある態様において、ポリヌクレオチドプログラム可能なヌクレオチド結合ドメインは、Cas9ドメイン、例えばヌクレアーゼ活性Cas9、Cas9ニッカーゼ (nCas9)、またはヌクレアーゼ不活性Cas9 (dCas9) である。核酸プログラム可能なDNA結合タンパク質の例としては、Cas9 (例:dCas9およびnCas9)、Cas12a/Cpf1、Cas12b/C2cl、Cas12c/C2c3、Cas12d/CasY、Cas12e/CasX、Cas12g、Cas12h、およびCas12iが挙げられるが、これらに限定されない。

【0182】

ヌクレアーゼ不活性化Cas9タンパク質は、互換的に、「dCas9」タンパク質 (ヌクレアーゼ「dead」Cas9の意)、または触媒的に不活性なCas9と称され得る。不活性なDNA切断ドメインを有するCas9タンパク質 (又はその断片) を生成する方法が知られている (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):1173-83参照 (各内容は参照により本明細書に組み込まれる))。例えば、Cas9のDNA切断ドメインは、HNHヌクレアーゼサブドメインとRuvC1サブドメインの2つのサブドメインを含むことが知られている。HNHサブドメインはgRNAに相補的な鎖を切断し、RuvC1サブドメインは非相補的な鎖を切断する。これらのサブドメイン内の変異はCas9のヌクレアーゼ活性を抑制する。例えば、突然変異D10Aと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))。

【0183】

ある態様において、Cas9ドメインは、Cas9ニッカーゼである。Cas9ニッカーゼは、二本鎖核酸分子 (例えば二本鎖DNA分子) の一方の鎖のみを切断することができるCas9タンパク質であり得る。いくつかの実施形態において、Cas9ニッカーゼは、二本鎖核酸分子の標的鎖を切断し、これは、Cas9ニッカーゼが、Cas9に結合しているgRNA (例えばsgRNA) と塩基対を形成している (相補的である) 鎖を切断することを意味する。ある態様において、Cas9ニッカーゼは、D10A突然変異を含み、位置840にヒスチジンを有する。いくつかの実施形態において、Cas9ニッカーゼは、二本鎖核酸分子の非標的、非塩基編集鎖を切断し、これは、Cas9ニッカーゼが、Cas9に結合しているgRNA (例えばsgRNA) と塩基対を形成していない鎖を切断することを意味する。ある態様において、Cas9ニッカーゼは、H840A突然変異を含み位置10にアスパラギン酸残基を有するか、または対応する突然変異。いくつかの実施形態において、Cas9ニッカーゼは、本明細書に提供されるCas9ニッカーゼのいずれかと少なくとも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%の同一性であるアミノ酸配列を含む。さらなる適切なCas9ニッカーゼは、本開示および当該分野における知識に基づいて当業者に明らかであり、本開示の範囲内である。

【0184】

ある態様において、Cas9ドメインは、ヌクレアーゼ不活性Cas9ドメイン (dCas9) である。例えば、dCas9ドメインは、二本鎖核酸分子のいずれの鎖も切断することなく、二本鎖核酸分子に結合し得る (例えば、gRNA分子を介して)。いくつかの実施形態において、ヌクレアーゼ不活性dCas9ドメインは、本明細書中に記載されたアミノ酸配列のD10X突然変異およびH840X突然変異、または本明細書中に提供されたアミノ酸配列のいずれかに

ける対応する突然変異を含み、Xは任意のアミノ酸変化である。いくつかの実施形態において、ヌクレアーゼ不活性dCas9ドメインは、本明細書に記載のアミノ酸配列のD10A突然変異およびH840A突然変異、または本明細書に記載のアミノ酸配列のいずれかにおける対応する突然変異を含む。一例として、ヌクレアーゼ不活性Cas9ドメインは、クローニングベクターpPlatTET-gRNA 2 (アクセス番号BAV54124)に記載の以下のアミノ酸配列を含む：

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEA
TRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNI
VDEVAYHEKYPTIYHLRKKLVDSTDKADRLRIYLALAHMIKFRGHFLIEGDLNPDNSDVD
KLFQILVQTYNQLFEENPINASGVDAKAILSARLSKSRRLLENLIAQLPGEKKNGLFGNLIAL
SLGLTPNFKSNFDLAEDAKLQLSKDQYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDI
LRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDG
GASQEEFYKFIKPILEKMDGTEELLVKLNRDRLRKQRTFDNGSIPHQIHLGELHAILRRQ
EDFYFPFLKDNREKIEKILTFRIPIYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGAS
AQSFIERMTNFDKNLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKA
IVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLD
NEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLIN
GIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA
GSPAIIKKGILQTVKVVDELVKVMGRHKPENIVIMARENQTTQKGQKNSRERMKRIEEGI
KELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVAIVPQSFLK
DDSIDNKVLTNRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGG
LSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRK
DFQFYKVRINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQ
EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLS
MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAK
VEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDLIIKLPKYSLELENGR
KRMLASAGELQKGNELALPSKYVNFYLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEII
EQISEFSKRVLADANLDKVL SAYNKHRRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTID
RKRYTSTKEVLDATLIHQSIITGLYETRIDLSQLGGD

(Qi et al., “Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression.” Cell. 2013; 152(5):1173-83参照。その全内容が参照により本明細書に組み入れられる。)

【0185】

追加的なCas9タンパク質(例えば、ヌクレアーゼ死Cas9 (dCas9)、Cas9ニッカーゼ (nCas9)、またはヌクレアーゼ活性Cas9)は、そのバリエーションおよびホモログを含めて、本開示の範囲内にあることを理解されたい。例示的なCas9タンパク質は、限定されるものではないが、以下に提供されるものを含む。ある態様において、Cas9タンパク質は、ヌクレアーゼ死Cas9 (dCas9) である。ある態様において、Cas9タンパク質は、Cas9ニッカーゼ (nCas9) である。いくつかの実施形態において、Cas9タンパク質はヌクレアーゼ活性Cas9である。

【0186】

触媒不活性Cas9 (dCas9) の例：

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEA
TRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNI
VDEVAYHEKYPTIYHLRKKLVDSTDKADRLRIYLALAHMIKFRGHFLIEGDLNPDNSDVD
KLFQILVQTYNQLFEENPINASGVDAKAILSARLSKSRRLLENLIAQLPGEKKNGLFGNLIAL
SLGLTPNFKSNFDLAEDAKLQLSKDQYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDI
LRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDG
GASQEEFYKFIKPILEKMDGTEELLVKLNRDRLRKQRTFDNGSIPHQIHLGELHAILRRQ
EDFYFPFLKDNREKIEKILTFRIPIYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGAS

AQSFIERMTNFDKNLPNEKVLPHKSHLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKA
IVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLD
NEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLIN
GIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA
GSPAIIKKGILQTVKVVDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGI
KELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDVAIVPQSFLK
DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGG
LSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRK
DFQFYKVRINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQ
EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLS
MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAK
VEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDLIIKLPKYSLELENGR
KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEII
EQISEFSKRVLADANLDKVL SAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTID
RKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGD

【0187】

Cas9ニッカ―ゼ (nCas9)の例を以下に記載する：

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEA
TRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNI
VDEVAYHEKYPTIYHLRKKLVDSTDKADRLIYLALAHMIKFRGHFLIEGDLNPDNSDVD
KLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIAL
SLGLTPNFKSNFDLAEDAKLQLSKD TYDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDI
LRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDG
GASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ
EDFYFPLKDNREKIEKILTFRIPIYVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGAS
AQSFIERMTNFDKNLPNEKVLPHKSHLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKA
IVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLD
NEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLIN
GIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA
GSPAIIKKGILQTVKVVDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGI
KELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK
DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGG
LSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRK
DFQFYKVRINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQ
EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLS
MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAK
VEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDLIIKLPKYSLELENGR
KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEII
EQISEFSKRVLADANLDKVL SAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTID
RKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGD。

【0188】

触媒的に活性なCas9の例を以下に記載する：

MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEA
TRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNI
VDEVAYHEKYPTIYHLRKKLVDSTDKADRLIYLALAHMIKFRGHFLIEGDLNPDNSDVD
KLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIAL
SLGLTPNFKSNFDLAEDAKLQLSKD TYDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDI
LRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDG
GASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ
EDFYFPLKDNREKIEKILTFRIPIYVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGAS

AQSFIERMTNFDKNLPNEKVLPHKSHLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKA
IVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLD
NEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLIN
GIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA
GSPAIIKKGILQTVKVVDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGI
KELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK
DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGG
LSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRK
DFQFYKVVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVDVRKMIKSEQ
EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLS
MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAK
VEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDLIIKLPKYSLELENGR
KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEII
EQISEFSKRVLADANLDKVL SAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTID
RKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGD。

【0189】

ある態様において、Cas9は、単細胞原核微生物のドメインおよび界を構成する古細菌(例えばナノアーキア)由来のCas9を指す。ある態様において、核酸プログラム可能DNA結合タンパク質は、例えば、Burstein et al., "New CRISPR-Cas systems from uncultivated microbes." Cell Res. 2017 Feb 21. doi: 10.1038/cr.2017.21に記載されているCasXまたはCasYを指し、その全体の内容は参照により本明細書に組み込まれる。ゲノム分解メタゲノミクスを用いて、生命の古細菌ドメインにおいて最初に報告されたCas9を含め、多くのCRISPR - Cas系が同定された。この分岐Cas9タンパク質は、ほとんど研究されていないナノアーキアにおいて、活性CRISPR - Cas系の一部として発見された。細菌では、それまで知られていなかった二つの系、CRISPR-CasXとCRISPR-CasYが発見され、それらは、これまでに発見された中でも最もコンパクトな系に入る。いくつかの実施形態において、本明細書に記載される塩基エディターシステムにおいて、Cas9は、CasXまたはCasXのバリエーションによって置き換えられる。いくつかの実施形態において、本明細書に記載される塩基エディターシステムにおいて、Cas9は、CasYまたはCasYのバリエーションによって置き換えられる。核酸プログラム可能DNA結合タンパク質 (napDNAbp) として他のRNA誘導DNA結合タンパク質も使用され得、本開示の範囲内であることが理解されるべきである。

【0190】

いくつかの実施形態において、本明細書で提供される融合タンパク質のいずれかの核酸プログラム可能DNA結合タンパク質 (napDNAbp) は、CasXまたはCasYタンパク質であり得る。いくつかの実施形態において、napDNAbpはCasXタンパク質である。いくつかの実施形態において、napDNAbpはCasYタンパク質である。いくつかの実施形態において、napDNAbpは、天然に存在するCasXまたはCasYタンパク質に対して少なくとも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%の同一性を有するアミノ酸配列を含む。いくつかの実施形態において、napDNAbpは、天然に存在するCasXまたはCasYタンパク質である。いくつかの実施形態において、napDNAbpは、本明細書に記載されるいずれかのCasXまたはCasYタンパク質に対して少なくとも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%の同一性を有するアミノ酸配列を含む。他の細菌種由来のCasXおよびCasYもまた、本開示に従って使用され得ることを理解されたい。

【0191】

例として、次のCas配列が提供される。

【0192】

CasX(uniprot.org/uniprot/F0NN87; uniprot.org/uniprot/F0NH53) tr | F0NN87 | F0NN87_SULIH CRISPR-associated Casx protein OS = Sulfolobus islandicus (strain HVE10/4) GN = SiH_0402 PE=4 SV=1:

MEVPLYNIFGDNYIIQVATEAENSTIYNNKVEIDDEELRNVNLAYKIAKNNEDAAAERRG
KAKKKKGEEGETTTSNIILPLSGNDKNPWTETLKCYNFPTTVALSEVFNFSQVKECEEVS
APSFVKPEFYEFGRSPGMVERTRRVKLEVEPHYLIIAAGWVLTRLGKAKVSEG DYVGVN
VFTPTRGILYSLIQNVNGIVPGIKPETAFLGLWIARKVVSSVTNPVSVVRIYTISDAVGQN
PTTINGGFSIDLTKLLEKRYLLSERLEAIARNALSISSNMRRERYIVLANIYIYEYLTG SKRLE
DLLYFANRDLIMNLSDDGKVRDLKLISAYVNGELIRGEG.

10

【0193】

> tr | F0NH53 | F0NH53_SULIR CRISPR associated protein, Casx OS = Sulfolobus islandicus (strain REY15A) GN=SiRe_0771 PE=4 SV=1:

MEVPLYNIFGDNYIIQVATEAENSTIYNNKVEIDDEELRNVNLAYKIAKNNEDAAAERRG
KAKKKKGEEGETTTSNIILPLSGNDKNPWTETLKCYNFPTTVALSEVFNFSQVKECEEVS
APSFVKPEFYKFGSPGMVERTRRVKLEVEPHYLIIMAAAGWVLTRLGKAKVSEG DYVGV
NVFTPTRGILYSLIQNVNGIVPGIKPETAFLGLWIARKVVSSVTNPVSVVSIYTISDAVGQ
NPTTINGGFSIDLTKLLEKRDLLSERLEAIARNALSISSNMRRERYIVLANIYIYEYLTGSKRLE
DLLYFANRDLIMNLSDDGKVRDLKLISAYVNGELIRGEG.

20

【0194】

Deltaproteobacteria CasX

MEKRINKIRKKLSADNATKPVSRSGPMKTLLVRVMTDDLKKRLEKRRKKPEVMPQVISN
NAANNLRMLDDYTKMKEAILQVYWQEFKDDHVGMLMCKFAQPASKKIDQNKLPKPEMDE
KGNLT TAGFACSQCQGPLFVYKLEQVSEK GKAYTNYFGRCNVAEHEKLILLAQLKPKV KDS
DEAVTYS LGKFGQRALDFYSIHVTKESTHPVKPLAQIAGNRYASGPVGKALSDACMG TIA
SFLSKYQDIIIEHQKVVKGNQKRLESLRELAKENLEYP SVTLPPQPHTKEGVDfAYNEVI
ARVRM WVNLNLWQKLKLSRDDAKPLLRLKGFPSPVVERRENEVDWWNTINEVKKLID
AKRDMGRVFWSGVTAEK RNTILEGYNL PNENDHKKREGSLENPKKPAKRQFGDLLYL
EKKYAGDWGKVFDEAWERIDKKIAGLTSHIEREEARNAEDAQSKAVLTDWLRAKASFVL
ERLKEMDEKEFYACEIQLQKWYGD LRGNPF AVEAENRVVDISGFSIGSDGHSIQYRNLLA
WKYLENGKREFYLLMNYGKKGRIRFTDGT DIKKS GKWQGLLYGGGKAKVIDLTFDPDDE
QLIILPLAFGTRQGREFIWN D LLSLETGLIKLANGRVIEKTIYNKKIGRDEPALFVALTFER
REVVDPSNIKPVN LIGVARGENIPAVIALTDPEGCPLPEFKDSSGGPTDILRIGEGYKEKQR
AIQAAKEVEQRRAGGYSRK FASKSRNLADDMVRNSARDLFYHAVTHDAVLVFANLSRGF
GRQGKRTFMTERQYTKMEDWLTAKLAYEGLTSKTYLSKTLAQYTSKTC SNCGFTITYAD
MDVMLVRLKKTSDGWATT LNNKELKA EYQITYYNRYKRQTVEKELSAELDRLSEESGNN
DISKWTKGRRDEALFLLKKRFSHRPVQE QFVCLDCGHEVHAAEQ AALNIARSWLFLNSN
STEFKSYKSGKQPFVGAWQAFYKRRLKEVWKPN A.

30

【0195】

CasY (ncbi.nlm.nih.gov/protein/APG80656.1) > APG80656.1 CRISPR-associated protein CasY [uncultured Parcubacteria group bacterium]:

MSKRHPRISGVKGYRLHAQRLEYTGKSGAMRTIKYPLYSSPSGGRTVPREIVSAINDDYV
GLYGLSNFDDLYNAEK RNEEKVYSVLD F WYDCVQYGA VFSYTAPGLLKNVAEVRGGSYE
LTKTLKGSHLYDELQIDKVIKFLNKKEISRANGSLDKLKKDIIDCFKA EYRERHKDQC NKL
ADDIKNAKKDAGASLGERQKKLFRDFFGISEQSENDKPSFTNPLNLTCCLLPFDTVNNR
NRGEVLFNKLKEYAQKLDKNEGSLEMWEYIGIGNSGTAFSNFLGEGFLGRLRENKITE LK
KAMMDITDAWRGQE QEEEELEKRLRILAALTIKLREP KFDNHWWGGYRS DINGKLSSWLQN
YINQTVKIKEDLKGHKKDLKKAKEMINRFGESDTKEEAVVSSLLESIEKIVPDDSADDEKP
DIPAIAIYRRFLSDGRLTLNR FVQREDVQEALIKERLEAEK KKKPKKRKKKSDAEDEKETID
FKELFPHLAKPLKLPNFYGD SKRELYKKYKNAAIYTDALWKAVEKIYKSAFSSSLKNSFF

40

50

DTDFDKDFFIKRLQKIFSVYRRFNTDKWKPIVKNSFAPYCDIVSLAENEVLYKPKQSRSRK
 SAAIDKNRVRLPSTENIAKAGIALARELSVAGFDWKDLLKKEEHEEYIDLIELHKTALALL
 LAVTETQLDISALDFVENGTVKDFMKT RDGNLVLEGRFLEMFSQSIVFSELRGLAGLMSR
 KEFITRSAIQTMNGKQAEELLYIPHEFQSAKITTPKEMSRAFLDLAPAEFATSLEPESEKS
 LLKLKQMRYYPHYFGYELTRTGQGIDGGVAENALRLEKSPVKKREIKCKQYKTLGRGQNK
 IVLYVRSSYYQTQFLEWFLHRPKNVQTDVAVSGSFLIDEKKVKTRWNYDALTVALEPVS
 GSERVFVSQPFTIFPEKSAAEEGQRYLGIDIGEYGIAYTALEITGDSAKILDQNFISDPQLKT
 LREEVKGLKLDQRRGTFAMPSTKIARIRESLVHSLRNRIHHLALKHKAKIVYELEVSRFEE
 GKQKIKKVYATLKKADVSEIDADKNLQTTVWGKLAVASEISASYTSQFCGACKKLWRAE
 MQVDETITTQELIGTVRVIKGGTLIDAIKDFMRPPIFDENDTFFPKYRDFCDKHHISKMR
 GNSCLFICPFCRANADADIQASQTIALLRVKEEKKVEDYFERFRKLKNIKVLGQMKKI
 【0196】

10

Cas12b/C2c1 (uniprot.org/uniprot/T0D7A2#2) sp | T0D7A2 | C2C1_ALIAG CRIS
 PR-associated endo-nuclease C2c1 OS = Alicyclobacillus acido-terrestris (stra
 in ATCC 49025 / DSM 3922/ CIP 106132 / NCIMB 13137/GD3B) GN=c2c1 PE
 =1SV=1:

MAVKSIVKLRLDDMPEIRAGLWKLHKEVNAGVRYYTEWLSLLRQENLYRRSPNGDGEQ
 ECDKTAECKAELLERLRARQVENGHARGPAGSDDELLQLARQLYELLVPQAIGAKGDAQ
 QIARKFLSPLADKDAVGGGLIAKAGNKPRWVRMREAGEPGWEEKEKAETRKSADRTAD
 VLRALADFGKPLMRVYTDSEMSSVEWKPLRKGQAVRTWDRDMFQQAERMMSWESW
 NQVRGQEYAKLVEQKNRFEQKNFVGQEHVHLVNQLQQDMKEASPGLESKEQTAHYVT
 GRALRGSDKVFEEKWGKLAPDAPFDLYDAEIKNVQRRNTRRFGSHDLFAKLAEPEYQALW
 REDASFLTRYAVYNSILRKLNHAKMFATFTLPDATAHPIWTRFDKLGGLNHQYTFLFNEF
 GERRHAIRFHKLKLVENGVAEVDVTVPIISMSEQLDNLLPRDPNEPIALYFRDYGAEQH
 FTGEFGGAKIQCRDQLAHMHRRRGARDVYLVNSVVRVQSQSEARGERRPPYAAVFRLVG
 DNHRAVFVHFDKLSDYLAEHPDDGKLGSEGLLSGLRVMSVDLGLRTSASISVFRVARKDEL
 KPNSKGRVPFFFPIKGNDNLVAVHERSOLLKLPGETESKDLRAIREERQRTLRLQLRTQLAY
 LRLLVRCGSEDVGRRERSWAKLIEQPVDAAANHMTDPDWREAFENELQKLKSLHGICSDKE
 WMDAVYESVRRVWRHMGKQVRDWRKDVRSGERPKIRGYAKDVVGGNSIEQIEYLERQY
 KFLKSWSSFGKVSQVIRAEKGSRAITLREHIDHAKEDRLKKLADRIIMEALGYVYALDE
 RGKGKWVAKYPPCQLILLEELSEYQFNNDRPPSENNQLMQWSHRGVFQELINQAQVHD
 LLVGTMYAAFSSRFDARTGAPGIRCRRVPARCTQEHNPFPWWLNKFVVEHTLDACPL
 RADDLIPTGEGEIFVSPFSAEEGDFHQIHADLNAAQNLQQRLWSDFDISQIRLRCDWGEV
 DGELVLIPRLTGKRTADSYSNKVFYTNVTGVTYYERERGGKRRKVFQAQEKLSSEEAELLVEA
 DEAREKSVVLMRDPGSIINRGNWTRQKEFWSMV NQRIEGLVKQIRSRVPLQDSACENT
 GDI.

20

30

【0197】

いくつかの実施形態において、融合タンパク質中に存在するCas9ドメインの1つは、PA
 M配列を必要としないガイドヌクレオチド配列プログラム可能DNA結合タンパク質ドメイ
 ンで置換され得る。

40

【0198】

ある態様において、核酸プログラム可能DNA結合タンパク質 (napDNAbp) は、微生物
 CRISPR-Cas系の単一のエフェクターである。微生物CRISPR-Cas系の単一エフェクター
 は、Cas9、Cpf1、Cas12b/C2c1、およびCas12c/C2c3を含むが、これらに限定されな
 い。典型的には、微生物CRISPR-Cas系はクラス1およびクラス2系に分けられる。クラス
 1の系は多サブユニットエフェクター複合体をもち、クラス2の系は単一のタンパク質エフ
 ェクターをもつ。たとえば、Cas9とCpf1はクラス2エフェクターである。Cas9およびCpf1
 に加えて、三つの異なるクラス2 CRISPR-Casシステム(Cas12b/C2c1およびCas12c/C
 2c3)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

50

記載されている(その全体の内容は参照により本明細書に組み込まれる)。2つの系、Cas12b/C2c1とCas12c/C2c3のエフェクターはCpf1に関連するRuvC様エンドヌクレアーゼ領域を含む。第3の系は、2つの予測されるHEPN RNaseドメインをもつエフェクターを含む。Cas12b/C2c1によるCRISPR RNAの産生とは異なり、成熟CRISPR RNAの産生がtracrRNA非依存性である。Cas12b/C2c1はDNA切断にCRISPRRNAとtracrRNAの両方に依存する。

【0199】

Alicyclobacillus acidoterrestris Cas12b/C2c1 (AacC2c1) の結晶構造が、キメラ単一分子ガイドRNA (sgRNA) との複合体として報告されている。例えば、Liu et al., “C2c1-sgRNA Complex Structure Reveals RNA-Guided DNA Cleavage Mechanism”, Mol. Cell, 2017 Jan. 19; 65(2):310-322参照(その内容全体を参照により本明細書に組み込む)。また、三元複合体として標的DNAに結合した*Alicyclobacillus acidoterrestris* C2c1においても結晶構造が報告されている。例えば、Yang et al., “PAM-dependent Target DNA Recognition and Cleavage by C2C1 CRISPR-Cas endonuclease”, Cell, 2016 Dec. 15; 167(7):1814-1828参照(その内容全体が参照により本明細書に組み込まれる)。標的DNA鎖および非標的DNA鎖の両方と共に、AacC2c1の触媒的にコンピテントな立体配座は、1つのRuvC触媒ポケット内に独立して捕捉され、Cas12b/C2c1媒介切断は標的DNAのスタガード7ヌクレオチドの切断を生じる。Cas12b/C2c1三元複合体と以前に同定されたCas9およびCpf1対応物の間の構造比較は、CRISPR-Cas9システムにより使用される機構の多様性を示す。

【0200】

いくつかの実施形態において、本明細書で提供される融合タンパク質のいずれかの核酸プログラム可能DNA結合タンパク質(napDNAbp)は、Cas12b/C2c1またはCas12c/C2c3タンパク質であり得る。いくつかの実施形態において、napDNAbpはCas12b/C2c1タンパク質である。いくつかの実施形態において、napDNAbpはCas12c/C2c3タンパク質である。いくつかの実施形態において、napDNAbpは、天然に存在するCas12b/C2c1またはCas12c/C2c3タンパク質に対して少なくとも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%の同一性を有するアミノ酸配列を含む。いくつかの実施形態において、napDNAbpは、天然に存在するCas12b/C2c1またはCas12c/C2c3タンパク質である。いくつかの実施形態において、napDNAbpは、本明細書に提供されるnapDNAbp配列のいずれかに対して少なくとも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%の同一性を有するアミノ酸配列を含む。他の細菌種由来のCas12b/C2c1またはCas12c/C2c3も本開示に従って使用することができることを理解されたい。

【0201】

Cas12b/C2c1 (uniprot.org/uniprot/T0D7A2#2) sp | T0D7A2 | /C2C1_ALIAG CRISPR-associated endo-nuclease C2c1 OS = *Alicyclobacillus acido-terrestris* (strain ATCC 49025 / DSM 3922 / CIP 106132 / NCIMB 13137/GD3B) GN=c2c1 PE=1 SV=1アミノ酸配列を以下に提供する：

MAVKSIVKVLRLDDMPDIRAGLWKLHKEVNAGVRYYTEWLSLLRQENLYRRSPNGDGEQ
ECDKTAEECKAELLRLRARQVENGHRGPAGSDDELLQLARQLYELLVPQAIGAKGDAQ
QIARKFLSPLADKDAVGGLGIAKAGNKPRWVRMREAGEPGWEEKEKAETRKSADRTAD
VLRALADFGKPLMRVYTDSEMSSVEWKPLRKGQAVRTWDRDMFQQAIERMMSWESW
NQRVGQEYAKLVEQKNRFEQKNFVGQEHVHLVNQLQQDMKEASPGLESKEQTAHYVT
GRALRGSDKVFKEKWGLAPDAFDLYDAEIKNVQRRNTRRFGSHDLFAKLAEPEYQALW
REDASFLTRYAVYNSILRKLNHAKMFATFTLPDATAHPIWTRFDKLGGLNHQYTFLFNEF
GERRHAIRFHKLLKVENGVAEVDVTVPIISMSEQLDNLLPRDPNEPIALYFRDYGAEQH

FTGEFGGAKIQCRDQLAHMHRRRGARDVYLVNSVRVQSQSEARGERRPPYAAVFRLVG
DNHRAVFVHFDKLSDYLAHPDDGKLGSEGLLSGLRVMSVDLGLRTSASISVFRVARKDEL
KPNSKGRVPFFFPIKGNDNLVAVHERSOLLKLPGETESKDLRAIREERQRTLRLRTQLAY
LRLLVRCGSEDVGRRERSWAKLIEQPVDAANHMTDPDWREAFENELQKLKSLHGICSDKE
WMDAVYESVRRVWRHMGKQVRDWRKDVRSGERPKIRGYAKDVVGGNSIEQIEYLERQY
KFLKSWSSFFGKVSGQVIRAEKGSRAITLREHIDHAKEDRLKKLADRIIMEALGYVYALDE
RGKGKWVAKYPPCQLILLEELSEYQFNNDRPPSENNQLMQWSHRGVFQELINQAQVHD
LLVGTMYAAFSSRFDARTGAPGIRCRRVPARCTQEHNPFPWWLNKFVVEHTLDACPL
RADDLIPTGEGEIFVSPFSAEEGDFHQIHADLNAAQNLQQRLWSDFDISQIRLRCDWGEV
DGELVLIPRLTGKRTADSYSNKVFYTNLTGVTYYERERGGKRRKVFQAQEKLSSEEAELLVEA
DEAREKSVVLMRDPGSIINRGNWTRQKEFWSMVNQRIEGLVKQIRSRVPLQDSACENT
GDI

10

【0202】

BhCas12b (*Bacillus hisashii*), NCBI Reference Sequence: WP_095142515の
アミノ酸配列を以下に提供する：

MAPKKKRKVGIGHGVPAAATRSFILKIEPNEEVKKGLWKTHEVLNHGIAYYMNLKLRQE
AIYEHHEQDPKNPKKVSKEAIQAELWDFVLKMQKCNSTHEVDKDEVFNILRELYEELVP
SSVEKKGEANQLSNKFLYPLVDPNSQSGKGTASSGRKPRWYNLKIAGDPSWEEKKKKWE
EDKKKDPLAKILGKLAEYGLIPLFIPYTDSEPIVKEIKWMEKSRNQSVRRLDKDMFIQAL
ERFLSWESWNLKVKEEYKVEKEYKTEERIKEDIQALKALEQYKERQEQLLRDTLNTN
EYRLSKRGLRGWREIIQKWLKMDENEPSEKYLEVFQDYQRKHPREAGDYSVYEFLSKKE
NHFIWRNHPEYPYLYATFCEIDKKKKDAKQATFTLADPINHPLWVRFEERSGSNLNKY
RILTEQLHTEKLKKKLTVQLDRLIYPTESGGWEEKGKVDIVLLPSRQFYNQIFLDIEEKGK
HAFTYKDESIKFLKGTGGARVQFDRDHLRRYPHKVESGNVGRIFYNMTVNIEPTESPV
SKSLKIHRDDFPKVVNFKPKELTEWIKDSKGKKLKSIGESLEIGLRVMSIDLQGRQAAAASI
FEVVDQKPDIEGKLPFPIKGTLEYAVHRASFNIKLPGETLVKSREVLKAREDNLKL MNQ
KLNFLRNVLHFQQFEDITEREKRVTKWISRQENS DVPLVYQDELIQIRELMYKPYKDWV
AFLKQLHKKRLEVEIGKEVKHWRKSLSDGRKGLYGISLKNIDEIDRTRKFLLRWSLRPTEPG
EVRRLEPGQRFAIDQLNHLNALKEDRLKKMANTIIMHALGYCYDVRKKKKWQAKNPACQI
ILFEDLSNYPYEERSRFENSKLMKWSRREIPRQVALQGEIYGLQVGEVGAQFSSRFHAK
TGSPGIRCSVVTKEKLQDNRRFFKNLQREGRLTLDKIAVLKEGDLYPDKGGEKFISLSKDRK
CVTTHADINAAQNLQKRFWTRTHGFYKVYCKAYQVDGQTVYIPESKDQKQKIIIEFGEG
YFILKDGVEYEWVNAGKLKIKKGSSKQSSSELVDSIDLKDSFDLASELKGEKLMLYRDP
SGNVFSPDKWMAAGVFFGKLERILISKLTNQYSISTIEDSSKQSMKRPAATKKAGQAKKKK.

20

30

【0203】

いくつかの実施形態において、Cas12bはBvCas12Bであり、これはBhBvCas12Bの
バリエーションであり、BhBvCas12Bに対して以下の変化を含む：S893R、K846R、およびE8
37G。

【0204】

BvCas12b (*Bacillus* sp. V3-13), NCBI Reference Sequence: WP_101661451.1
のアミノ酸配列を以下に提供する：

40

MAIRSIKLMKTNSGTDISIYLRKALWRTHQLINEGIAYYMNLTLRYQEAIGDKTKEAYQ
AELINIIRNQQRNNGSSEEHGSDQEILALLRQLYELIIPSSIGESGDANQLGNKFLYPLVDP
NSQSGKGTSNAGRKPRWKRLKEEGNPDWELEKKKDEERKAKDPTVKIFDNLNKYGLLPL
FPLFTNIQKDIEWLPLGKRQSVRKWDKDMFIQAIERLLSWESWNRRVADEYKQLKEKTE
SYYKEHLTGGEEWIEKIRKFEKERNMELEKNAFAPNDGYFITSRQIRGWDRVYEKWSKLP
ESASPEELWKVVAEQNKMSSEGFGDPKVFSSFLANREN RDIWRGHSERIYHIAAYNGLQK
KLSRTKEQATFTLPDAIEHPLWIRYESPGGTNLNLFKLEEKQKKNYYVTLSKIIWPSEK
WIEKENIEIPLAPSIQFNRQIKLKQHVKGKQEISFSDYSSRISLDGVLGGSRIQFNRKYIKN
HKELLGEGDIGPVFFNLVVDVAPLQETRNGRLQSPIGKALKVSSDFSKVIDYKPKELMD

50

WMNTGSASNSFGVASLLEGMRVMSIDMGQRTSASVSIFEVVKELPKDQEQKLFYSINDTE
LFAIHKRSFLLNLPGEVVTNNKQQRQERRKKRQFVRSQIRMLANVLRLETCKTPDERKK
AIHKLMEIVQSYDSWTASQKEVWEKELNLLTNMAAFNDEIWKESLVELHHRIEPYVGQI
VSKWRKGLSEGRKNLAGISMWNIDELEDTRLLISWSKRSRTPGEANRIETDEPFGSSLL
QHIQNVKDDRLKQMANLIIMTALGFKYDKEEKDRYKRWKETYPACQIILFENLNRYLFNL
DRSRRENSRLMKWAHRSIPRTVSMQGEMFGLQVGDRSEYSSRFHAKTGAPGIRCHALT
EEDLKAGSNTLKRLEDGFINESELAYLKKGDIIPSQGGELFVTLKRYKKDSNNELTVIH
ADINAAQNLQKRFWQQNSEVYRVPCQLARMGEDKLYIPKSQTETIKKYFGKGSFVKNNT
EQEVYKWEKSEKMKIKTDTTFDLQDLDFEDISKTIELAQQKKYLTMFRDPSGYFFNN
ETWRPQKEYWSIVNNIIKSLKKKILSNKVEL

10

【0205】

ポリヌクレオチドプログラム可能ヌクレオチド結合ドメインはまた、RNAに結合する核酸プログラム可能タンパク質を含むことができることを理解されたい。例えば、ポリヌクレオチドプログラム可能ヌクレオチド結合ドメインは、ポリヌクレオチドプログラム可能ヌクレオチド結合ドメインをRNAにガイドする核酸と結合され得る。他の核酸プログラム可能DNA結合タンパク質もまた、本開示の範囲内にあるが、それらは本開示には特に列記されていない。

【0206】

本明細書中で用いることができるCAsタンパク質は、クラス1およびクラス2を含む。Casタンパク質の非限定的な例としては、Cas1、Cas1B、Cas2、Cas3、Cas4、Cas5d、Cas5t、Cas5h、Cas5a、Cas6、Cas7、Cas8、Cas9(Csn1またはCsx12とも呼ばれる)、Cas10、Csy1、Csy2、Csy3、Csy4、Cse1、Cse2、Cse3、Cse4、Cse5、Csn1、Csn2、Csm2、Csm3、Csm4、Csm5、Csm6、Cmr1、Cmr3、Cmr4、Cmr5、Csb1、Csb2、Csb3、Csx17、Csx14、Csx10、Ccss16、Cx16、Cx、Csx3、Csx1、Csx1S、Csf1、Csf2、Cs0、Csf4、Csd1、Csd2、Cst1、Cst2、Csh1、Csh2、Csa1、Csa2、Csa3、Csa4、Csa5、Cas12a/Cpfl、Cas12b/C2cl、Cas12c/C2c3、Cas12d/CasY、Cas12e/CasX、Cas12g、Cas12h、およびCas12i、CARF、DinG、それらのホモログ、またはそれらの改変体が挙げられる。未改変のCRISPR酵素は、Cas9のように、2つの機能性エンドヌクレアーゼ領域、RuvCおよびHNHを有するDNA切断活性を有することができる。CRISPR酵素は、標的配列内および/または標的配列の相補鎖内などの標的配列において、一方または両方の鎖の切断を誘導することができる。例えば、CRISPR酵素は、標的配列の最初または最後のヌクレオチドから約1、2、3、4、5、6、7、8、9、10、15、20、25、50、100、200、500塩基対またはそれ以上にある一方または両方の鎖の切断を誘導することができる。

20

30

【0207】

標的配列を含む標的ポリヌクレオチドの一方または両方の鎖を切断する能力を欠失するように、対応する野生型酵素に対して変異されたCRISPR酵素をコードするベクターを用いることができる。Cas9は、野生型の例示的なCas9ポリペプチド(例えば*S. pyogenes*からのCas9)と少なくとも、または少なくともおよそ、50%、60%、70%、80%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、または100%の配列同一性および/または配列相同性を有するポリペプチドを指すことができる。Cas9は、野生型の例示的なCas9ポリペプチド(例えば*S. pyogenes*からのもの)に対して、最大で、または最大でおよそ、約50%、60%、70%、80%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、または100%の配列同一性および/または配列相同性を有するポリペプチドを指すことができる。Cas9は、野生型、または欠失、挿入、置換、バリエーション、突然変異、融合、キメラ、またはそれらの任意の組合せなどのアミノ酸変化を含み得るCas9タンパク質の改変型を指すことができる。

40

【0208】

いくつかの実施形態において、本明細書に記載される方法は、組み換え操作された(engineered) Casタンパク質を利用することができる。ガイドRNA (gRNA) は、Cas結合に

50

必要な足場配列と、修飾されるゲノム標的を規定するユーザー定義の約20塩基スパーサーとからなる短い合成RNAである。したがって、Casタンパク質特異性のゲノム標的の変化は、ゲノムの他の部分と比較したゲノム標的に対するgRNA標的化配列の特異性によって部分的に決定されることが理解されるであろう。

【0209】

Cas9ヌクレアーゼはRuvCとHNHという二つの機能性エンドヌクレアーゼ領域をもっている。Cas9は標的DNAに結合すると二番目のコンホメーション変化を起こし、これがヌクレアーゼドメインを位置付け、標的DNAの反対側の鎖を切断させる。Cas9を介したDNA切断の最終結果は、標的DNA内の二本鎖切断 (DSB) である (PAM配列の約3~4ヌクレオチド上流)。生じたDSBは、次の二つの一般的修復経路のうちの一つにより修復される: (1) 効率的だが誤りがちな非相同末端結合 (NHEJ) 経路; または (2) 効率は低いが忠実度の高い相同性誘導修復 (HDR) 経路。

【0210】

非相同末端結合 (NHEJ) および/または相同性指向修復 (HDR) の「効率」は、任意の簡便な方法によって計算することができる。例えば、ある場合には、効率は、成功したHDRのパーセンテージで表すことができる。例えば、検査用ヌクレアーゼ分析を用いて切断産物を生成し、基質に対する産物の比を用いてパーセンテージを計算することができる。たとえば、HDRが成功した結果として新たに組み込まれた制限配列を含むDNAを直接切断する測定用ヌクレアーゼ酵素を使用することができる。切断される基質が多いほどHDRの割合が高かった (HDRの効率がより高かった) ことを示す。例示的な例として、HDRの割合 (パーセンテージ) は、以下の式を用いて計算することができる $[(\text{切断産物}) / (\text{基質} + \text{切断産物})]$ (例えば、 $(b+c) / (a+b+c)$ ここで、「a」はDNA基質のバンド強度であり、「b」および「c」は切断生成物である。)

【0211】

場合によっては、効率はNHEJの成功率で表すことができる。例えば、T7エンドヌクレアーゼIアッセイを用いて切断産物を生成し、基質に対する産物の比を用いてNHEJのパーセンテージを計算することができる。T7エンドヌクレアーゼIは、野生型および突然変異体DNA鎖 (NHEJは最初の切断部位に小さなランダムな挿入や欠失 (インデル) を生じる) のハイブリダイゼーションから生じるミスマッチのヘテロ二本鎖DNAを切断する。切断が多いほどNHEJの割合が高いこと (NHEJの効率が低いこと) を示す。例示的な例として、NHEJの割合 (パーセンテージ) は、式 $(1 - (1 - (b+c)/(a+b+c))^{1/2}) \times 100$ を用いて計算することができ、ここで、「a」はDNA基質のバンド強度であり、「b」および「c」は切断生成物である (Ran et al., Cell. 2013 Sep. 12; 154(6):1380-9; and Ran et al., Nat Protoc. 2013 Nov.; 8(11): 2281-2308)。

【0212】

NHEJ修復経路は最も活性な修復機構であり、DSB部位での小ヌクレオチド挿入または欠失 (インデル) を頻繁に引き起こす。NHEJ媒介DSB修復のランダム性は、Cas9およびgRNAまたはガイドポリヌクレオチドを発現する細胞集団が多様な突然変異のアレイをもたらすので、重要な実地的意味を有する。ほとんどの場合、NHEJは標的DNAに小さなインデルを生じさせ、その結果、標的遺伝子のオープンリーディングフレーム (ORF) 内に未熟な終止コドンをもたらすアミノ酸欠失、挿入、またはフレームシフト突然変異が生じる。理想的な最終結果は、標的遺伝子内の機能喪失型変異である。

【0213】

NHEJ媒介DSB修復はしばしば遺伝子のオープンリーディングフレームを破壊するが、相同性指向修復 (HDR) は単一ヌクレオチド変化からフルオロフォアまたはタグの付加のような大きな挿入までの範囲の特異的ヌクレオチド変化を生成するために使用できる。

【0214】

HDRを遺伝子編集に利用するために、所望の配列を含むDNA修復テンプレートを、gRNAおよびCas9またはCas9ニッカーゼと共に目的の細胞型に送達することができる。修復テンプレートは、所望の編集ならびにその標的のすぐ上流および下流 (左右相同アームと呼ば

れる)にさらなる相同配列を含むことができる。各相同アームの長さは導入される変化の大きさに依存し得、より大きな挿入はより長い相同アームを必要とする。修復鑄型は、一本鎖オリゴヌクレオチド、二本鎖オリゴヌクレオチド、または二本鎖DNAプラスミドであり得る。HDRの効率は、Cas9、gRNAおよび外因性修復テンプレートを発現する細胞においても、一般的に低い(10%未満の修正アリル)。HDRは細胞周期のS期とG2期の間に起こるので、細胞を同調させることによってHDRの効率を高めることができる。NHEJに関与する遺伝子を化学的または遺伝的に阻害することもHDR頻度を増加させ得る。

【0215】

いくつかの実施態様において、Cas9は、修飾Cas9である。所与のgRNA標的配列は、ゲノム全体にわたり、部分的な相同性が存在するさらなる部位を有し得る。これらの部位はオフターゲットサイト(「オフターゲット」と呼ばれ、gRNAを設計する際に考慮する必要があるが、gRNAの設計を最適化することに加えて、Cas9を修飾することによってCRISPRの特異性を高めることもできる。Cas9は2つのヌクレアーゼドメイン、RuvCおよびHNHの組合せ活性を介して二本鎖切断(DSB)を生成する。SpCas9のD10A変異体であるCas9ニッカーゼは1つのヌクレアーゼドメインを保持し、DSBではなくDNAニックを生成する。特異的遺伝子編集のためのHDR媒介遺伝子編集にニッカーゼを組み合わせることもできる。

【0216】

ある場合には、Cas9はバリエーションCas9タンパク質である。バリエーションCas9ポリペプチドは、野生型Cas9タンパク質のアミノ酸配列と比較して、一アミノ酸単位で異なる(例えば、欠失、挿入、置換、融合を有する)アミノ酸配列を有する。いくつかの例において、バリエーションCas9ポリペプチドは、Cas9ポリペプチドのヌクレアーゼ活性を低下させるアミノ酸変化(例えば欠失、挿入、または置換)を有する。例えば、いくつかの例において、バリエーションCas9ポリペプチドは、対応する野生型Cas9タンパク質のヌクレアーゼ活性の50%未満、40%未満、30%未満、20%未満、10%未満、5%未満、または1%未満を有する。ある場合には、バリエーションCas9タンパク質は実質的なヌクレアーゼ活性をもたない。対象Cas9タンパク質が、実質的なヌクレアーゼ活性を有さないバリエーションCas9タンパク質である場合、それは「dCas9」と称され得る。

【0217】

ある場合には、バリエーションCas9タンパク質はヌクレアーゼ活性を低下させる。例えば、バリエーションCas9タンパク質は、野生型Cas9タンパク質(例えば野生型Cas9タンパク質)のエンドヌクレアーゼ活性の約20%未満、約15%未満、約10%未満、約5%未満、約1%未満、または約0.1%未満を示す。

【0218】

ある場合には、バリエーションCas9タンパク質は、ガイド標的配列の相補鎖を切断することができるが、二本鎖ガイド標的配列の非相補鎖を切断する能力が低下している。例えば、バリエーションCas9タンパク質は、RuvCドメインの機能を低下させる突然変異(アミノ酸置換)を有することができる。非限定的な例として、いくつかの実施形態において、バリエーションCas9タンパク質は、D10A(アミノ酸位置10におけるアスパラギン酸からアラニン)を有し、したがって、二本鎖ガイド標的配列の相補鎖を切断することができるが、二本鎖ガイド標的配列の非相補鎖を切断する能力が低下している(したがって、このバリエーションCas9タンパク質が二本鎖標的核酸を切断するとき、二本鎖切断(DSB)の代わりに一本鎖切断(SSB)を生じる)(例えばJinek et al., Science, 2012 Aug. 17; 337(6096):816-21参照)。

【0219】

ある場合には、バリエーションCas9タンパク質は、二本鎖ガイド標的配列の非相補鎖を切断することができるが、ガイド標的配列の相補鎖を切断する能力が低下している。例えば、バリエーションCas9タンパク質は、HNHドメインの機能を低下させる変異(アミノ酸置換)を有することができる(RuvC/HNH/RuvCドメインモチーフ)。非限定的な例として、いくつかの実施形態において、バリエーションCas9タンパク質は、H840A(アミノ酸位置840におけ

るヒスチジンからアラニン)突然変異を有し、したがって、ガイド標的配列の非相補的ストランドを切断することができるが、ガイド標的配列の相補的ストランドを切断する能力が低下している(したがって、このバリエーションCas9タンパク質が二本鎖ガイド標的配列を切断すると、DSBの代わりにSSBが生じる)。このようなCas9タンパク質は、ガイド標的配列(例えば一本鎖ガイド標的配列)を切断する能力が低下しているが、ガイド標的配列(例えば一本鎖ガイド標的配列)に結合する能力を保持している。

【0220】

ある場合には、バリエーションCas9タンパク質は、二本鎖標的DNAの相補鎖および非相補鎖の両方を切断する能力が低下している。非限定的な例として、ある場合には、バリエーションCas9タンパク質は、D10AおよびH840A突然変異の両方を有し、その結果、ポリペプチドは、二本鎖標的DNAの相補鎖および非相補鎖の両方を切断する能力が低下している。このようなCas9タンパク質は、標的DNA(例えば一本鎖標的DNA)を切断する能力が低下しているが、標的DNA(例えば一本鎖標的DNA)に結合する能力は保持している。

10

【0221】

別の非限定的な例として、いくつかの場合において、バリエーションCas9タンパク質は、W476AおよびW1126A突然変異を有し、その結果、ポリペプチドは、標的DNA(例えば一本鎖標的DNA)を切断する能力が低下しているが、標的DNA(例えば一本鎖標的DNA)に結合する能力は保持している。

【0222】

別の非限定的な例として、いくつかの場合において、バリエーションCas9タンパク質は、P475A、W476A、N477A、D1125A、W1126A、およびD1127A突然変異を有し、その結果、ポリペプチドは、標的DNAを切断する能力が低下している。そのようなCas9タンパク質は、標的DNAを切断する能力を低下しているが(例えば一本鎖標的DNA)、標的DNA(例えば一本鎖標的DNA)に結合する能力を保持している。

20

【0223】

別の非限定的な例として、いくつかの場合において、バリエーションCas9タンパク質は、H840A、W476A、およびW1126A突然変異を有し、その結果、ポリペプチドは、標的DNA(例えば一本鎖標的DNA)を切断する能力が低下しているが、標的DNA(例えば一本鎖標的DNA)に結合する能力は保持している。別の非限定的な例として、いくつかの場合において、バリエーションCas9タンパク質は、H840A、D10A、W476A、およびW1126A突然変異を有し、その結果、ポリペプチドは、標的DNAを切断する能力が低下している。このようなCas9タンパク質は、標的DNA(例えば一本鎖標的DNA)を切断する能力が低下しているが、標的DNA(例えば一本鎖標的DNA)に結合する能力は保持している。いくつかの実施形態において、バリエーションCas9は、Cas9 HNHドメインの位置840における触媒的His残基が回復されている(A840H)。

30

【0224】

別の非限定的な例として、いくつかの場合において、バリエーションCas9タンパク質は、H840A、P475A、W476A、N477A、D1125A、W1126A、およびD1127A突然変異を有し、その結果、ポリペプチドは、標的DNA(例えば一本鎖標的DNA)を切断する能力を低下させるが、標的DNA(例えば一本鎖標的DNA)に結合する能力を保持する。別の非限定的な例として、いくつかの場合において、バリエーションCas9タンパク質は、D10A、H840A、P475A、W476A、N477A、D1125A、W1126A、およびD1127A突然変異を有し、その結果、ポリペプチドは、標的DNAを切断する能力が低下している。そのようなCas9タンパク質は、標的DNA(例えば一本鎖標的DNA)を切断する能力を低下しているが、標的DNA(例えば一本鎖標的DNA)に結合する能力を保持している。バリエーションCas9タンパク質がW476AおよびW1126A変異を有する場合、またはバリエーションCas9タンパク質がP475A、W476A、N477A、D1125A、W1126A、およびD1127A変異を有する場合、バリエーションCas9タンパク質はPAM配列に効率的に結合しない。したがって、このような場合には、このようなバリエーションCas9タンパク質を結合の方法に用いると、この方法はPAM配列を必要としない。換言すれば、ある場合には、このようなバリエーションCas9タンパク質を

40

50

結合の方法に用いる場合、この方法はガイドRNAを含み得るが、この方法は、PAM配列の非存在下で行うことができる(したがって、結合の特異性はガイドRNAの標的セグメントによってもたらされる)。上記の効果を達成するために、他の残基を変異させ得る(すなわち一方または他方のヌクレアーゼ部分を不活性化する)。非限定的な例として、残基D10、G12、G17、E762、H840、N854、N863、H982、H983、A984、D986、および/またはA987を変更(すなわち置換)することができる。また、アラニン置換以外の変異も好適である。

【0225】

ある態様において、低下した触媒活性を有するバリエーションCas9タンパク質(例えばCas9タンパク質がD10、G12、G17、E762、H840、N854、N863、H982、H983、A984、D986、および/またはA987突然変異、例えばD10A、G12A、G17A、E762A、H840A、N854A、N863A、H982A、H983A、A984A、および/またはD986Aを有する場合)は、それがガイドRNAと相互作用する能力を保持する限り、部位特異的様式で標的DNAに結合することができる(ガイドRNAによって標的DNA配列に誘導されるからである)。

【0226】

S. pyogenes Cas9の代替としては、哺乳動物細胞において切断活性を示すCpf1ファミリー由来のRNA誘導エンドヌクレアーゼが挙げられ得る。PrevotellaおよびFrancisella由来のCRISPR (CRISPR/Cpf1) は、CRISPR/Cas9システムに類似したDNA編集技術である。Cpf1はクラスII CRISPR/Cas系のRNA誘導エンドヌクレアーゼである。この獲得免疫機構はPrevotellaやFrancisella細菌に見られる。したがって、Cpf1は、Cas9とは異なるPAM特異性を有する核酸プログラム可能DNA結合タンパク質の一例を表す。Cas9と同様に、Cpf1もクラス2のCRISPRエフェクターである。Cpf1はCas9とは異なる特徴で頑強なDNA干渉を媒介することが示されている。Cpf1は、tracrRNAを欠く単一RNA誘導エンドヌクレアーゼであり、Tに富むプロトスペーサ隣接モチーフ(TTN、TTTN、またはYTN)を利用する。さらに、Cpf1はDNAを互い違いの二本鎖切断で切断する。16のCpf1ファミリータンパク質のうち、AcidaminococcusとLachnospiraceaeからの二つの酵素が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ている。Cpf1タンパク質は、例えば、Yamano et al., “Crystal structure of Cpf1 in complex with guide RNA and target DNA.” Cell (165) 2016, p. 949-962に記載されており、その全体の内容は参照により本明細書に組み込まれる。

【0227】

Cpf1遺伝子はCRISPR遺伝子座に関連しており、ウイルスDNAを見出して切断するためにガイドRNAを用いるエンドヌクレアーゼをコードしている。Cpf1はCas9より小さく単純なエンドヌクレアーゼであるため、Cpf1はCRISPR/Cas9系の制限のいくつかを克服することができる。Cas9ヌクレアーゼとは異なり、Cpf1を介したDNA切断の結果は、短い3'突出を伴う二本鎖切断である。Cpf1の互い違いの切断パターンは、伝統的な制限酵素クローニングに類似した、方向性のある遺伝子導入の可能性を開くことができ、これは遺伝子編集の効率を高め得る。上述したCas9のバリエーションおよびオーソログと同様に、Cpf1は、CRISPRが標的とすることができる部位の数を、SpCas9が好むNGG PAM部位を欠くATに富む領域またはATに富むゲノムに拡大することもできる。Cpf1遺伝子座は α/β 混合ドメイン、RuvC-Iとそれに続くらせん領域、RuvC-IIおよびジンクフィンガー様ドメインを含む。Cpf1タンパク質は、Cas9のRuvCドメインに類似したRuvC様エンドヌクレアーゼドメインを有する。さらに、Cpf1はHNHエンドヌクレアーゼ領域をもたず、Cpf1のN末端はCas9の α ヘリックス認識ローブをもたない。Cpf1 CRISPR-Casドメイン構成は、Cpf1が機能的にユニークであり、クラス2、タイプV CRISPRシステムとして分類されることを示した。Cpf1遺伝子座は、II型系よりもI型およびIII型に類似したCas1、Cas2およびCas4タンパク質をコードしていた。機能的Cpf1はトランス活性化CRISPR RNA (tracrRNA) を必要としない;したがって、CRISPR (crRNA) だけを要する。Cpf1はCas9より小さいだけでなく、より小さいsgRNA分子(Cas9の約半分の数のヌクレオチド)を有するので、これはゲノム編集に有益である。Cas9が標的とするGリッチPAMとは対照的に、

Cpf1-crRNA複合体は、モチーフ5'-YTN-3'に隣接するプロトSpacerの同定によって標的DNAまたはRNAを切断する。PAMの同定後、Cpf1は、4または5ヌクレオチドの突出を有するスティッキーエンド様のDNA二本鎖切断を導入する。

【0228】

また、ガイドヌクレオチド配列プログラム可能DNA結合タンパク質ドメインとして使用され得るヌクレアーゼ不活性Cpf1 (dCpf1) バリエーションも、本発明の組成物および方法において有用である。Cpf1タンパク質は、Cas9のRuvCドメインに類似した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 α -ヘリックス認識ローブを有さない。Zetsche et al., Cell, 163, 759-771, 2015 (参照により本明細書に組み込まれる)は、Cpf1のRuvC様ドメインが両方のDNA鎖の切断を担い、RuvC様ドメインの不活性化はCpf1ヌクレアーゼ活性を不活化することを示した。例えば、Francisella novicida Cpf1におけるD917A、E1006A、またはD1255Aに対応する突然変異は、Cpf1ヌクレアーゼ活性を不活化する。いくつかの実施形態において、本開示のdCpf1は、D917A、E1006A、D1255A、D917A/E1006A、D917A/D1255A、E1006A/D1255A、またはD917A/E1006A/D1255Aに対応する突然変異を含む。Cpf1のRuvCドメインを不活性化する任意の突然変異、例えば、置換変異、欠失、または挿入が、本開示に従って使用され得ることを理解されたい。

【0229】

いくつかの実施形態において、本明細書で提供される融合タンパク質のいずれかの核酸プログラム可能DNA結合タンパク質 (napDNAbp) は、Cpf1タンパク質であり得る。いくつかの実施形態において、Cpf1タンパク質は、Cpf1ニッカーゼ (nCpf1) である。ある態様において、Cpf1タンパク質は、ヌクレアーゼ不活性Cpf1 (dCpf1) である。いくつかの実施形態において、Cpf1、nCpf1、またはdCpf1は、本明細書に開示されたCpf1配列に対して少なくとも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%の同一性を有するアミノ酸配列を含む。いくつかの実施形態において、dCpf1は、本明細書に開示されたCpf1配列に対して少なくとも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%の同一性を有するアミノ酸配列を含み、D917A、E1006A、D1255A、D917A/E1006A、D917A/D1255A、E1006A/D1255A、またはD917A/E1006A/D1255Aに対応する突然変異を含む。

【0230】

他の細菌種からのCpf1もまた、本開示に従って使用され得ることを理解されたい。従って、他の細菌種由来の以下の例示的Cpf1配列も、本開示に従って使用することができる：

【0231】

野生型Francisella novicida Cpf1 (D917、E1006、およびD1255を太字下線で示す。)

MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYHQFFI
EEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSEKFKNLFN
QNLIDAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIISFKGWTTYFKGFHEN
RKNVYSSNDIPTSIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIKKDLAEELTFDID
YKTSEVNQRVFSLDEVFEIANFNLYLNQSGITKFNTIIGGKFVNGENTKRKGINEYINLYS
QQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTTMQSFYEQIAAFKTVEEKS
IKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSSQVFDYVIGTAVLEYITQQIAPKNLD
NPSKKEQELIAKKTEKAKYLSLETIKLALAEFNKHRDIDKQCRFEEILANFAAIPMIFDEIA
QNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIKDLLDQTNNLLHKLKIFHISQSEDKAN
ILDKDEHFYLVFEECYFELANIVPLYNKIRNYITQKPYSDEKFKLNFENSTLANGWDKNKE
PDNTAILFIKDDKYYLGVNMNKKNNKIFDDKAIKENKGEYKKIVYKLLPGANKMLPKVFF
SAKSIKFYNPSEDILRIRNHSTHTKNGSPQKGYEKFEFNIEDCRKFIDFYKQSISKHPEWKD

FGFRFSDTQRYNSIDEFYREVENQGYKLTFENISESYIDSVVNQGKLYLFQIYNKDFSAYS
KGRPNLHTLYWKALFDERNLQDVVYKLNGEAEFYRKQSIPKKITHPAKEAIANKNDNP
KKESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFNDEINLLLKEKANDVHILSIDRG
ERHLAYYTLVDGKGNIQKQDTFNIIGNDRMKTNYHDKLAAIEKDRDSARKDWKKINNIKE
MKEGYLSQVVHEIAKLVEIYNAIVVFEDLNF~~GF~~KRGRFKVEKQVYQKLEKMLIEKLNLYLV
FKDNEFDKTGGVLRAYQLTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGTVNQLYPKY
ESVSKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKAAKGKWTIASFGSRLINFRNSDKN
HNWDTREVVYPTKELEKLLKDYSIEYGHGECIKAAICGESDKKFFAKLTSVLNTILQMRNSK
TGTELDYLISPVADVNGNFFDSRQAPKNMPQDADANGAYHIGLKGLMLLGRIKNNQEGK
KLNLVIKNEEYFEFVQNRNN

10

【0232】

Francisella novicida Cpf1 D917A (A917、E1006、およびD1255を太字下線で示す。)

MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYHQFFI
EEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSEKFNLFN
QNLIDAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIISFKGWTTYFKGFHEN
RKNVYSSNDIPTSIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIKKDLAEELTFDID
YKTSEVNQRVFSLDEVFEIANFNLYLNQSGITKFNTIIGGKFVNGENTKRKGINEYINLYS
QQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTMMQSFYEQIAAFKTVEEKS
IKETLSLLFDDLKAQKLDLSKIYFKNDKSLDLSQQVFDDYSVIGTAVLEYITQQIAPKNLD
NPSKKEQELIAKKTEKAKYLSLETIKLAL~~EE~~FNKHRDIDKQCRFEEILANFAAIPMIFDEIA
QNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIKDLLDQTNNLLHKLKIFHISQSEDKAN
ILDKDEHFYLVFEECYFELANIVPLYNKIRNYITQKPYSDEKFKLN~~FEN~~STLANGWDKNKE
PDNTAILFIKDDKYYLGVMNKKNNKIFDDKA~~IKEN~~KGEGYKKIVYKLLPGANKMLPKVFF
SAKSIKFYNPSEDILRIRNHSTHTKNGSPQKGYEKFEFNIEDCRKFIDFYKQSISKHPEWKD
FGFRFSDTQRYNSIDEFYREVENQGYKLTFENISESYIDSVVNQGKLYLFQIYNKDFSAYS
KGRPNLHTLYWKALFDERNLQDVVYKLNGEAEFYRKQSIPKKITHPAKEAIANKNDNP
KKESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFNDEINLLLKEKANDVHILSIARG
ERHLAYYTLVDGKGNIQKQDTFNIIGNDRMKTNYHDKLAAIEKDRDSARKDWKKINNIKE
MKEGYLSQVVHEIAKLVEIYNAIVVFEDLNF~~GF~~KRGRFKVEKQVYQKLEKMLIEKLNLYLV
FKDNEFDKTGGVLRAYQLTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGTVNQLYPKY
ESVSKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKAAKGKWTIASFGSRLINFRNSDKN
HNWDTREVVYPTKELEKLLKDYSIEYGHGECIKAAICGESDKKFFAKLTSVLNTILQMRNSK
TGTELDYLISPVADVNGNFFDSRQAPKNMPQDADANGAYHIGLKGLMLLGRIKNNQEGK
KLNLVIKNEEYFEFVQNRNN

20

30

【0233】

Francisella novicida Cpf1 E1006A (D917、A1006、およびD1255を太字下線で示す。)

MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYHQFFI
EEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSEKFNLFN
QNLIDAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIISFKGWTTYFKGFHEN
RKNVYSSNDIPTSIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIKKDLAEELTFDID
YKTSEVNQRVFSLDEVFEIANFNLYLNQSGITKFNTIIGGKFVNGENTKRKGINEYINLYS
QQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTMMQSFYEQIAAFKTVEEKS
IKETLSLLFDDLKAQKLDLSKIYFKNDKSLDLSQQVFDDYSVIGTAVLEYITQQIAPKNLD
NPSKKEQELIAKKTEKAKYLSLETIKLAL~~EE~~FNKHRDIDKQCRFEEILANFAAIPMIFDEIA
QNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIKDLLDQTNNLLHKLKIFHISQSEDKAN
ILDKDEHFYLVFEECYFELANIVPLYNKIRNYITQKPYSDEKFKLN~~FEN~~STLANGWDKNKE
PDNTAILFIKDDKYYLGVMNKKNNKIFDDKA~~IKEN~~KGEGYKKIVYKLLPGANKMLPKVFF
SAKSIKFYNPSEDILRIRNHSTHTKNGSPQKGYEKFEFNIEDCRKFIDFYKQSISKHPEWKD

40

50

FGFRFSDTQRYNSIDEFYREVENQGYKLTFENISESYIDSVVNQGKLYLFQIYNKDFSAYS
KGRPNLHTLYWKALFDERNLQDVVYKLNGEAEFYRKQSIPKKITHPAKEAIANKNDNP
KKESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFNDEINLLLKEKANDVHILSIDRG
ERHLAYYTLVDGKGNIQKQDTFNIIGNDRMKTNYHDKLAAIEKDRDSARKDWKKINNIKE
MKEGYLSQVVHEIAKLVEIYNAIVVFADLNF~~FG~~FKRGRFKVEKQVYQKLEKMLIEKLNLYLV
FKDNEFDKTGGVLRAYQLTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTG~~FN~~QLYPKY
ESVSKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKAAKGKWTIASFGSRLINFRNSDKN
HNWDTREVVYPTKELEKLLKDYSIEYGHGECIKAAICGESDKKFFAKLTSVLNTILQMRNSK
TGTELDYLISPVADVNGNFFDSRQAPKNMPQDADANGAYHIGLKGLMLLGRIKNNQEGK
KLNLVIKNEEYFEFVQNRNN

10

【0234】

Francisella novicida Cpf1 D1255A (D917、E1006、およびA1255を太字下線で示す。)

MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYHQFFI
EEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSEKFNLFN
QNLIDAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIISFKGWTTYFKGFHEN
RKNVYSSNDIPTSIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIKKDLAEELTFDID
YKTSEVNQRVFSLDEVFEIANFN~~NY~~LNQSGITKFNTIIGGKFVNGENTKRKGINEYINLYS
QQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTMMQSFYEQIAAFKTVEEKS
IKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDL~~S~~QQVFDDYSVIGTAVLEYITQQIAPKNLD
NPSKKEQELIAKKTEKAKYLSLETIKLAL~~EE~~FNKHRDIDKQCRFEEILANFAAIPMIFDEIA
QNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIKDLLDQTNNLLHKLKIFHISQSEDKAN
ILDKDEHFYLVFE~~EC~~YFELANIVPLYNKIRNYITQKPYSDEKFKLN~~FEN~~STLANGWDKNKE
PDNTAILFIKDDKY~~Y~~LGVMNKKNNKIFDDKA~~IKEN~~KGEGYKKIVYKLLPGANKMLPKVFF
SAKSIKFYNPSEDILRIRNHSTHTKNGSPQKG~~YE~~KFEFNIEDCRKFIDFYKQSISKHPEWKD
FGFRFSDTQRYNSIDEFYREVENQGYKLTFENISESYIDSVVNQGKLYLFQIYNKDFSAYS
KGRPNLHTLYWKALFDERNLQDVVYKLNGEAEFYRKQSIPKKITHPAKEAIANKNDNP
KKESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFNDEINLLLKEKANDVHILSIDRG
ERHLAYYTLVDGKGNIQKQDTFNIIGNDRMKTNYHDKLAAIEKDRDSARKDWKKINNIKE
MKEGYLSQVVHEIAKLVEIYNAIVVFEDLNF~~FG~~FKRGRFKVEKQVYQKLEKMLIEKLNLYLV
FKDNEFDKTGGVLRAYQLTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTG~~FN~~QLYPKY
ESVSKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKAAKGKWTIASFGSRLINFRNSDKN
HNWDTREVVYPTKELEKLLKDYSIEYGHGECIKAAICGESDKKFFAKLTSVLNTILQMRNSK
TGTELDYLISPVADVNGNFFDSRQAPKNMPQDAAANGAYHIGLKGLMLLGRIKNNQEGK
KLNLVIKNEEYFEFVQNRNN

20

30

【0235】

Francisella novicida Cpf1 D917A/E1006A (A917、A1006、D1255を太字下線で示す。)

MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYHQFFI
EEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSEKFNLFN
QNLIDAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIISFKGWTTYFKGFHEN
RKNVYSSNDIPTSIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIKKDLAEELTFDID
YKTSEVNQRVFSLDEVFEIANFN~~NY~~LNQSGITKFNTIIGGKFVNGENTKRKGINEYINLYS
QQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTMMQSFYEQIAAFKTVEEKS
IKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDL~~S~~QQVFDDYSVIGTAVLEYITQQIAPKNLD
NPSKKEQELIAKKTEKAKYLSLETIKLAL~~EE~~FNKHRDIDKQCRFEEILANFAAIPMIFDEIA
QNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIKDLLDQTNNLLHKLKIFHISQSEDKAN
ILDKDEHFYLVFE~~EC~~YFELANIVPLYNKIRNYITQKPYSDEKFKLN~~FEN~~STLANGWDKNKE
PDNTAILFIKDDKY~~Y~~LGVMNKKNNKIFDDKA~~IKEN~~KGEGYKKIVYKLLPGANKMLPKVFF
SAKSIKFYNPSEDILRIRNHSTHTKNGSPQKG~~YE~~KFEFNIEDCRKFIDFYKQSISKHPEWKD

40

50

FGFRFSDTQRYNSIDEFYREVENQGYKLTFENISESYIDSVVNQGKLYLFQIYNKDFSAYS
KGRPNLHTLYWKALFDERNLQDVVYKLNGEAEFYRKQSIPKKITHPAKEAIANKNDNP
KKESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFNDEINLLLKEKANDVHILSIARG
ERHLAYYTLVDGKGNIQKQDTFNIIGNDRMKTNYHDKLAAIEKDRDSARKDWKKINNIKE
MKEGYLSQVVHEIAKLVEIYNAIVVFADLNF~~FG~~FKRGRFKVEKQVYQKLEKMLIEKLNLYLV
FKDNEFDKTGGVLRAYQLTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGTVNQLYPKY
ESVSKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKAAKGKWTIASFGSRLINFRNSDKN
HNWDTREVVYPTKELEKLLKDYSIEYGHGECIKAAICGESDKKFFAKLTSVLNTILQMRNSK
TGTELDYLISPVADVNGNFFDSRQAPKNMPQDADANGAYHIGLKGLMLLGRIKNNQEGK
KLNLVIKNEEYFEFVQNRNN

10

【0236】

Francisella novicida Cpf1 D917A/D1255A (A917、E1006、A1255を太字下線
で示す。)

MSIQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYHQFFI
EEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSEKFNLFN
QNLIDAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIISFKGWTTYFKGFHEN
RKNVYSSNDIPTSIIRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIKKDLAEELTFDID
YKTSEVNQRVFSLDEVFEIANFNLYLNQSGITKFNTIIGGKFVNGENTKRKGINEYINLYS
QQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTMMQSFYEQIAAFKTVEEKS
IKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLQQVFDVSVIGTAVLEYITQQIAPKNLD
NPSKKEQELIAKKTEKAKYLSLETIKLAL~~EE~~FNKHRDIDKQCRFEEILANFAAIPMIFDEIA
QNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIKDLLDQTNNLLHKLKIFHISQSEDKAN
ILDKDEHFYLVFEECYFELANIVPLYNKIRNYITQKPYSDEKFKLNFENSTLANGWDKNKE
PDNTAILFIKDDKYYLGVMNKKNNKIFDDKA~~IK~~ENKGE~~GY~~KKIVYKLLPGANKMLPKVFF
SAKSIKFYNPSEDILRIRNHSTHTKNGSPQKG~~YE~~KFEFNIEDCRKFIDFYKQSISKHPEWKD
FGFRFSDTQRYNSIDEFYREVENQGYKLTFENISESYIDSVVNQGKLYLFQIYNKDFSAYS
KGRPNLHTLYWKALFDERNLQDVVYKLNGEAEFYRKQSIPKKITHPAKEAIANKNDNP
KKESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFNDEINLLLKEKANDVHILSIARG
ERHLAYYTLVDGKGNIQKQDTFNIIGNDRMKTNYHDKLAAIEKDRDSARKDWKKINNIKE
MKEGYLSQVVHEIAKLVEIYNAIVVFEDLNF~~FG~~FKRGRFKVEKQVYQKLEKMLIEKLNLYLV
FKDNEFDKTGGVLRAYQLTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGTVNQLYPKY
ESVSKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKAAKGKWTIASFGSRLINFRNSDKN
HNWDTREVVYPTKELEKLLKDYSIEYGHGECIKAAICGESDKKFFAKLTSVLNTILQMRNSK
TGTELDYLISPVADVNGNFFDSRQAPKNMPQDAAANGAYHIGLKGLMLLGRIKNNQEGK
KLNLVIKNEEYFEFVQNRNN

20

30

【0237】

Francisella novicida Cpf1 E1006A/D1255A (D917、A1006、A1255を太字下線
で示す。)

MSIQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYHQFFI
EEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSEKFNLFN
QNLIDAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIISFKGWTTYFKGFHEN
RKNVYSSNDIPTSIIRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIKKDLAEELTFDID
YKTSEVNQRVFSLDEVFEIANFNLYLNQSGITKFNTIIGGKFVNGENTKRKGINEYINLYS
QQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTMMQSFYEQIAAFKTVEEKS
IKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLQQVFDVSVIGTAVLEYITQQIAPKNLD
NPSKKEQELIAKKTEKAKYLSLETIKLAL~~EE~~FNKHRDIDKQCRFEEILANFAAIPMIFDEIA
QNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIKDLLDQTNNLLHKLKIFHISQSEDKAN
ILDKDEHFYLVFEECYFELANIVPLYNKIRNYITQKPYSDEKFKLNFENSTLANGWDKNKE
PDNTAILFIKDDKYYLGVMNKKNNKIFDDKA~~IK~~ENKGE~~GY~~KKIVYKLLPGANKMLPKVFF
SAKSIKFYNPSEDILRIRNHSTHTKNGSPQKG~~YE~~KFEFNIEDCRKFIDFYKQSISKHPEWKD

40

50

FGFRFSDTQRYNSIDEFYREVENQGYKLTFENISESYIDSVVNQGKLYLFQIYNKDFSAYS
KGRPNLHTLYWKALFDERNLQDVVYKLNGEAEFYRKQSIPKKITHPAKEAIANKNDNP
KKESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFNDEINLLLKEKANDVHILSIDRG
ERHLAYYTLVDGKGNIQKQDTFNIIGNDRMKTNYHDKLAAIEKDRDSARKDWKKINNIKE
MKEGYLSQVVHEIAKLVEIYNAIVVFADLNFSGFRGRFVKVEKQVYQKLEKMLIEKLNLYLV
FKDNEFDKTGGVLRAYQLTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGTVNQLYPKY
ESVSKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKAAKGKWTIASFGSRLINFRNSDKN
HNWDTREVVYPTKELEKLLKDYSIEYGHGECIKAAICGESDKKFFAKLTSVLNTILQMRNSK
TGTELDYLISPVADVNGNFFDSRQAPKNMPQDAANGAYHIGLKGLMLLGRIKNNQEGK
KLNLVIKNEEYFEFVQNRNN

10

【0238】

Francisella novicida Cpf1 D917A/E1006A/D1255A (A917、A1006、A1255を太字下線で示す。)

MSIQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYHQFFI
EEILSSVCISEDLLQNYSDVYFKLKSSDDNLQKDFKSAKDTIKKQISEYIKDSEKFNLFN
QNLIDAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIISFKGWTTFYFKGFHEN
RKNVYSSNDIPTSIIRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIKKDLAEELTFDID
YKTSEVNQRVFSLDEVFEIANFNLYLNQSGITKFNTIIGGKFVNGENTKRKGINEYINLYS
QQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTMMQSFYEQIAAFKTVEEKS
IKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSSQVFDYVIGTAVLEYITQQIAPKNLD
NPSKKEQELIAKKTEKAKYLSLETIKLALAEFNKHRDIDKQCRFEEILANFAAIPMIFDEIA
QNKDNLAQISIKYQNGGKDLLQASAEDDVKAIKDLLDQTNNLLHKLKIFHISQSEDKAN
ILDKDEHFYLVFEECYFELANIVPLYNKIRNYITQKPYSDEKFKLNFENSTLANGWDKNKE
PDNTAILFIKDDKYYLGVMNKKNNKIFDDKAIKENKGEGYKKIVYKLLPGANKMLPKVFF
SAKSIKFYNPSEDILRIRNHSTHTKNGSPQKGYEKFEFNIEDCRKFIDFYKQSISKHPEWKD
FGFRFSDTQRYNSIDEFYREVENQGYKLTFENISESYIDSVVNQGKLYLFQIYNKDFSAYS
KGRPNLHTLYWKALFDERNLQDVVYKLNGEAEFYRKQSIPKKITHPAKEAIANKNDNP
KKESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFNDEINLLLKEKANDVHILSIARG
ERHLAYYTLVDGKGNIQKQDTFNIIGNDRMKTNYHDKLAAIEKDRDSARKDWKKINNIKE
MKEGYLSQVVHEIAKLVEIYNAIVVFADLNFSGFRGRFVKVEKQVYQKLEKMLIEKLNLYLV
FKDNEFDKTGGVLRAYQLTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGTVNQLYPKY
ESVSKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKAAKGKWTIASFGSRLINFRNSDKN
HNWDTREVVYPTKELEKLLKDYSIEYGHGECIKAAICGESDKKFFAKLTSVLNTILQMRNSK
TGTELDYLISPVADVNGNFFDSRQAPKNMPQDAANGAYHIGLKGLMLLGRIKNNQEGK
KLNLVIKNEEYFEFVQNRNN

20

30

【0239】

塩基エディターのポリヌクレオチドプログラム可能ヌクレオチド結合ドメインは、それ自体が1つ以上のドメインを含むことができる。例えば、ポリヌクレオチドプログラム可能なヌクレオチド結合ドメインは、1つ以上のヌクレアーゼドメインを含むことができる。ある態様において、ポリヌクレオチドプログラム可能ヌクレオチド結合ドメインのヌクレアーゼドメインは、エンドヌクレアーゼまたはエキソヌクレアーゼを含むことができる。本明細書において、用語「エキソヌクレアーゼ」は、核酸(例えばRNAまたはDNA)を遊離末端から消化することができるタンパク質またはポリペプチドを指し、用語「エンドヌクレアーゼ」は、核酸(例えばDNAまたはRNA)の内部領域を触媒(例えば劈開)することができるタンパク質またはポリペプチドを指す。ある態様において、エンドヌクレアーゼは、二本鎖核酸の一本鎖を切断することができる。ある態様において、エンドヌクレアーゼは、二本鎖核酸分子の両方の鎖を切断することができる。ある態様において、ポリヌクレオチドプログラム可能ヌクレオチド結合ドメインは、デオキシリボヌクレアーゼであり得る。ある態様において、ポリヌクレオチドプログラム可能ヌクレオチド結合ドメインは、リボヌクレアーゼであり得る。

40

50

【0240】

ある態様において、ポリヌクレオチドプログラム可能なヌクレオチド結合ドメインのヌクレアーゼドメインは、標的ポリヌクレオチドの0本、1本または2本の鎖を切断することができる。ある場合には、ポリヌクレオチドプログラム可能なヌクレオチド結合ドメインは、ニッカーゼドメインを含むことができる。本明細書において、用語「ニッカーゼ」は、二本鎖核酸分子(例えばDNA)中の二本鎖の一方の鎖のみを切断することができるヌクレアーゼドメインを含むポリヌクレオチドプログラム可能なヌクレオチド結合ドメインを指す。ある態様において、ニッカーゼは、活性なポリヌクレオチドプログラム可能なヌクレオチド結合ドメインに一つ以上の突然変異を導入することによって、ポリヌクレオチドプログラム可能なヌクレオチド結合ドメインの完全に触媒活性な(例えば天然の)形態から誘導することができる。例えば、ポリヌクレオチドプログラム可能なヌクレオチド結合ドメインがCas9に由来するニッカーゼドメインを含む場合、Cas9に由来するニッカーゼドメインは、D10A突然変異および位置840におけるヒスチジン(H)を含むことができる。そのような場合、残基H 840は触媒活性を保持し、それによって核酸二本鎖の一本鎖を切断することができる。別の例において、Cas9由来ニッカーゼドメインは、H840A突然変異を含むことができ、一方、位置10におけるアミノ酸残基は、Dのままである。いくつかの実施形態において、ニッカーゼは、ニッカーゼ活性に必要ではないヌクレアーゼドメインの全部または一部を除去することによって、ポリヌクレオチドプログラム可能なヌクレオチド結合ドメインの完全に触媒活性な(例えば天然の)形態から誘導することができる。例えば、ポリヌクレオチドプログラム可能なヌクレオチド結合ドメインがCas9に由来するニッカーゼドメインを含む場合、Cas9に由来するニッカーゼドメインは、RuvCドメインまたはHNHドメインの全部または一部の欠失を含むことができる。

【0241】

したがって、ニッカーゼドメインを含むポリヌクレオチドプログラム可能なヌクレオチド結合ドメインを含む塩基エディターは、特定のポリヌクレオチド標的配列(例えば結合したガイド核酸の相補的配列によって決定される)において一本鎖DNA切断(ニック)を生成することができる。ある態様において、ニッカーゼドメイン(例えばCas9由来のニッカーゼドメイン)を含む塩基エディターによって切断される核酸二本鎖標的ポリヌクレオチド配列の鎖は、塩基エディターによって編集されない鎖である(すなわち、塩基エディターによって切断される鎖は、編集される塩基を含む鎖とは反対の鎖である)。他の態様において、ニッカーゼドメイン(例えばCas9由来のニッカーゼドメイン)を含む塩基エディターは、編集のために標的とされるDNA分子の鎖を切断することができる。このような場合、非標的鎖は切断されない。

【0242】

触媒的に死んだ(すなわち標的ポリヌクレオチド配列を切断することができない)ポリヌクレオチドプログラム可能なヌクレオチド結合ドメインを含む塩基エディターも本明細書中に提供される。本明細書において、用語「触媒的に死んだ」および「死んだヌクレアーゼ」は、核酸の鎖を切断することができない結果となる一つ以上の突然変異および/または欠失を有するポリヌクレオチドプログラム可能なヌクレオチド結合ドメインを指すために交換可能に使用される。いくつかの実施形態において、触媒的に死んだポリヌクレオチドプログラム可能なヌクレオチド結合ドメイン塩基エディターは、1つ以上のヌクレアーゼドメインにおける特定の点突然変異の結果としてヌクレアーゼ活性を欠くことができる。例えば、Cas9ドメインを含む塩基エディターの場合、Cas9は、D10A突然変異およびH840A突然変異の両方を含むことができる。このような変異は両方のヌクレアーゼドメインを不活性化し、その結果ヌクレアーゼ活性を失う。他の実施形態において、触媒的に死んだポリヌクレオチドプログラム可能なヌクレオチド結合ドメインは、触媒ドメイン(例えばRuvC 1および/またはHNHドメイン)の全部又は一部の一つ以上の欠失を含むことができる。さらなる態様において、触媒的に死んだポリヌクレオチドプログラム可能なヌクレオチド結合ドメインは、点突然変異(例えばD10AまたはH840A)ならびにヌクレアーゼドメインの全部または一部の欠失を含む。

【0243】

また、本明細書では、ポリヌクレオチドプログラム可能ヌクレオチド結合ドメインの以前に機能していたバージョンから、触媒的に死んだポリヌクレオチドプログラム可能ヌクレオチド結合ドメインを生成することができる突然変異も企図される。例えば、触媒的に死んだCas9 (「dCas9」) の場合、D10AおよびH840A以外の突然変異を有してヌクレアーゼ不活性化Cas9をもたらすバリエーションが提供される。このような突然変異は、例えば、D10およびH840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サブドメインにおける置換)を含む。

【0244】

さらなる適切なヌクレアーゼ不活性dCas9ドメインは、本開示および当該分野における知識に基づいて当業者に明らかとなり得、本開示の範囲内である。このようなさらなる例示的な適切なヌクレアーゼ不活性Cas9ドメインには、限定されるものではないが、D10A/H840A、D10A/D839A/H840A、およびD10A/D839A/H840A/N863A突然変異体ドメインが含まれる(例えば、Prashant et al., CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. Nature Biotechnology. 2013; 31(9): 833-838を参照されたい(その全内容は参照により本明細書に組み込まれる))。いくつかの実施形態において、dCas9ドメインは、本明細書に提供されるdCas9ドメインのいずれかに対して少なくとも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%の同一性を有するアミノ酸配列を含む。いくつかの実施形態において、Cas9ドメインは、本明細書に記載されるアミノ酸配列のいずれかと比較して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またはそれ以上の突然変異を有するアミノ酸配列を含む。いくつかの実施形態において、Cas9ドメインは、本明細書に記載されたアミノ酸配列のいずれかと比較して、少なくとも10、少なくとも15、少なくとも20、少なくとも30、少なくとも40、少なくとも50、少なくとも60、少なくとも70、少なくとも80、少なくとも90、少なくとも100、少なくとも150、少なくとも200、少なくとも250、少なくとも300、少なくとも350、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1100、または少なくとも1200の同一の連続したアミノ酸残基を有するアミノ酸配列を含む。

【0245】

塩基エディターに組み込むことができるポリヌクレオチドプログラム可能ヌクレオチド結合ドメインの非限定的な例としては、CRISPRタンパク質由来ドメイン、制限ヌクレアーゼ、メガヌクレアーゼ、TALヌクレアーゼ(TALEN)、およびジンクフィンガーヌクレアーゼ(ZFN)が挙げられる。いくつかの場合において、塩基エディターは、核酸のCRISPR(すなわちClustered Regularly Interspaced Short Palindromic Repeats)媒介修飾の際に、結合ガイド核酸を介して核酸配列に結合することができる天然もしくは修飾されたタンパク質またはその一部を含むポリヌクレオチドプログラム可能ヌクレオチド結合ドメインを含む。そのようなタンパク質は、本明細書において「CRISPRタンパク質」と呼ばれる。従って、本明細書に開示されているのは、CRISPRタンパク質の全部または一部を含むポリヌクレオチドプログラム可能ヌクレオチド結合ドメインを含む塩基エディター(すなわち、CRISPRタンパク質の全部または一部をドメインとして含む塩基エディター(それは塩基エディターの「CRISPRタンパク質由来ドメイン」とも呼ばれる))である。塩基エディターに組み込まれたCRISPRタンパク質由来ドメインは、野生型または天然型のCRISPRタンパク質と比較して改変され得る。例えば、以下に記載するように、CRISPRタンパク質由来ドメインは、野生型または天然型のCRISPRタンパク質と比較して、1

10

20

30

40

50

以上の突然変異、挿入、欠失、再配列および/または組換えを含み得る。

【0246】

いくつかの実施形態において、塩基エディターに組み込まれたCRISPRタンパク質由来ドメインは、結合ガイド核酸と組み合わせられた場合に標的ポリヌクレオチドに結合することができるエンドヌクレアーゼ(例えばデオキシリボヌクレアーゼまたはリボヌクレアーゼ)である。いくつかの実施形態において、塩基エディターに組み込まれたCRISPRタンパク質由来ドメインは、結合ガイド核酸と組み合わせられた場合に標的ポリヌクレオチドに結合することができるニッカーゼである。いくつかの実施形態において、塩基エディターに組み込まれたCRISPRタンパク質由来ドメインは、結合ガイド核酸と組み合わせられた場合に標的ポリヌクレオチドに結合することができる触媒的に死んだドメインである。或る実施態様では、塩基エディターのCRISPRタンパク質由来ドメインに結合する標的ポリヌクレオチドはDNAであり、或る実施態様では、塩基エディターのCRISPRタンパク質由来ドメインに結合する標的ポリヌクレオチドはRNAである。

10

【0247】

ある態様において、塩基エディターのCRISPRタンパク質由来ドメインは、*Corynebacterium ulcerans* (NCBI Refs: NC_015683.1, NC_017317.1); *Corynebacterium diphtheria* (NCBI Refs: NC_016782.1, NC_016786.1); *Spiroplasma syrphidicola* (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); *Neisseria meningitidis* (NCBI Ref: YP_002342100.1), *Streptococcus pyogenes*, または *Staphylococcus aureus* 由来のCas9の全部または一部を含むことができる。

20

【0248】

ある態様において、塩基エディターのCas9由来ドメインは、*Staphylococcus aureus* 由来のCas9ドメイン(SaCas9)である。ある態様において、SaCas9ドメインは、ヌクレアーゼ活性SaCas9、ヌクレアーゼ不活性SaCas9 (SaCas9d)、またはSaCas9ニッカーゼ(SaCas9n)である。ある態様において、SaCas9ドメインは、N579X突然変異を含む。ある態様において、SaCas9ドメインは、N579A突然変異を含む。いくつかの実施形態において、SaCas9ドメイン、SaCas9dドメイン、またはSaCas9nドメインは、非標準PAMを有する核酸配列に結合することができ、いくつかの実施形態において、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することができる。いくつかの実施形態において、SaCas9ドメインは、E781X、N967X、およびR1014X突然変異のうちの1つ以上を含む。

30

【0249】

ある態様において、Cas9ドメインは、*Staphylococcus aureus* 由来のCas9ドメイン(SaCas9)である。ある態様において、SaCas9ドメインは、ヌクレアーゼ活性SaCas9、ヌクレアーゼ不活性SaCas9 (SaCas9d)、またはSaCas9ニッカーゼ (SaCas9n) である。いくつかの実施形態において、SaCas9は、N579A突然変異、または本明細書に提供されるアミノ酸配列のいずれかにおける対応する突然変異を含む。

40

【0250】

いくつかの実施形態において、SaCas9ドメイン、SaCas9dドメイン、またはSaCas9nドメインは、非標準PAMを有する核酸配列に結合することができ、いくつかの実施形態において、SaCas9ドメイン、SaCas9dドメイン、またはSaCas9nドメインは、NNGRRTまたはNNNRRT PAM配列を有する核酸配列に結合することができる。いくつかの実施形態において、SaCas9ドメインは、E781X、N967X、およびR1014X突然変異の1以上、または本明細書に提供されるアミノ酸配列のいずれかにおける対応する突然変異を含み、ここでXは任意のアミノ酸である。いくつかの実施形態において、SaCas9ドメインは、E78

50

1K、N967K、およびR1014H突然変異のうちの1つ以上、または本明細書に提供されるアミノ酸配列のいずれかにおける1つ以上の対応する突然変異を含む。いくつかの実施形態において、SaCas9ドメインは、E781K、N967K、またはR1014H突然変異、または本明細書に提供されるアミノ酸配列のいずれかにおける対応する突然変異を含む。

【0251】

塩基エディターは、高忠実度Cas9であるCas9の全部又は一部に由来するドメインを含むことができる。いくつかの実施形態において、塩基エディターの高忠実度Cas9ドメインは、対応する野生型Cas9ドメインと比較して、Cas9ドメインとDNAの糖-リン酸骨格との間の静電相互作用を減少させる1つ以上の突然変異を含む、操作されたCas9ドメインである。DNAの糖-リン酸骨格との静電相互作用を減少させた高忠実度Cas9ドメインは、より少ないオフターゲット効果を有することができる。いくつかの実施形態において、Cas9ドメイン(例えば、野生型Cas9ドメイン)は、Cas9ドメインとDNAの糖-リン酸主鎖との間の結合を減少させる一つ以上の突然変異を含む。いくつかの実施形態において、Cas9ドメインは、Cas9ドメインとDNAの糖-リン酸主鎖との間の結合を少なくとも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%、少なくとも10%、少なくとも15%、少なくとも20%、少なくとも25%、少なくとも30%、少なくとも35%、少なくとも40%、少なくとも45%、少なくとも50%、少なくとも55%、少なくとも60%、少なくとも65%、少なくとも70%以上減少させる一つ以上の突然変異を含む。

【0252】

いくつかの実施形態において、変異型Casタンパク質は、spCas9、spCas9-VRQR、spCas9-VRER、xCas9(sp)、saCas9、saCas9-KKKH、spCas9-MQKSER、spCas9-LRKIQK、またはspCas9-LRVSQLであり得る。例示的なsaCas9配列を以下に提供する：

KRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRRRRH
RIQRVKKLLFDYNLLTDHSELGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGVHNVNE
VEEDTGNELSTKEQISRNSKALEEKYVAELQLERLKKDGEVRGSINRFKTSYVKEAKQLL
KVQKAYHQLDQSFIDTYIDLLETRRTYYEGPGEKSPFGWKDIKEWYEMLMGHCTYFPEE
LRSVKYAYNADLYNALNDLNNLVITRDENEKLEYEYEFQIIENVFKQKKKPTLKQIAKEIL
VNEEDIKGYRVTSTGKPEFTNLKVYHDIKDITARKEIENAELLDQIAKILTIYQSSEDIQEE
LTNLNSELQTQEEIEQISNLKGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPKKV
DLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIKKYGLPNDIIELAREKNSKDAQKMI
NEMQKRNRQTNERIEEIIRTTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNPNFN
YEVDHIIPRSVSFDNSFNKVLVKQEENS~~SK~~GNRTPFQYLSSSDSKISYETFKKHILNLAK
GKGRISKTKKEYLLEERDINRFSVQKDFINRNLVDTRYATRGLMNLRSYFRVNNLDVKV
KSINGGFTSFLRRKWKFKERNKGYKHHAEALIANADFIKWKLDKAKKVMENQM
FEEKQAESMPEIETEQEYKEIFITPHQIKHIKDFKDYKYSHRVDKKPNRELINDTLYSTRK
DDKGNTLIVNNLNGLYDKDNDKLKKLINKSPEKLLMYHHDPPQTYQKLKLIMEQYGDEN
PLYKYYEETGNYLTKYSKKDNGPVIKKIKYYGNKLNALHDITDDYPNSRNKVVKLSLKP
YRFDVYLDNGVYKFVTVKNLDVIKKENYEVNSKCYEEAKKLKISNQAEFIASFYNNDLIK
INGELYRVIGVNNDLLNRIEVMIDITYREYLENMNDKRPPRIKTIASKTQSIKKYSTDIL
GNLYEVKSKKHPQIIKKG。

上記saCas9配列において、太字で下線を引いた残基N579は(例えばA579へと)変異されてSaCas9ニッカーゼを生じることができる。

【0253】

例示的SaCas9n配列を以下に提供する：

KRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRRRRH
RIQRVKKLLFDYNLLTDHSELGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGVHNVNE
VEEDTGNELSTKEQISRNSKALEEKYVAELQLERLKKDGEVRGSINRFKTSYVKEAKQLL
KVQKAYHQLDQSFIDTYIDLLETRRTYYEGPGEKSPFGWKDIKEWYEMLMGHCTYFPEE
LRSVKYAYNADLYNALNDLNNLVITRDENEKLEYEYEFQIIENVFKQKKKPTLKQIAKEIL
VNEEDIKGYRVTSTGKPEFTNLKVYHDIKDITARKEIENAELLDQIAKILTIYQSSEDIQEE

LTNLNSELQTQEEIEQISNLKGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPKKV
 DLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIIKKYGLPNDIIELAREKNSKDAQKMI
 NEMQKRNRQTNERIEEIIRTTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNNPFI
 YEVDHIIPRSVSFDNSFNKVLVKQEEASKKGNRTPFQYLSSSDSKISYETFKKHILNLAK
 GKGRISKTKKEYLLEERDINRFSVQKDFINRNLVDTRYATRGLMNLRLSYFRVNNLDVKV
 KSINGGFTSFLRRKWKFKKERNKGYKHHAEADALIINANADFIFKEWKKLDKAKKVMENQM
 FEEKQAESMPEIETEQEYKEIFITPHQIKHIKDFKDYKYSHRVDKKPNRELINDTLYSTRK
 DDKGNTLIVNNLNGLYDKDNDKLKKLINKSPEKLLMYHHDPPQTYQKLKLIMEQYGDEKN
 PLYKYYEETGNYLTKEYSKKDNGPVIKKIKYYGNKLNALHDITDDYPNSRNKVVKLSLKP
 YRFDVYLDNGVYKFVTVKNLVDVIKKENYYEVNSKCYEEAKKLKKISNQAEFIASFYNNDLIK
 INGELYRVIGVNNDLLNRIEVMIDITYREYLENMNDKRPPRIKTIASKTQSIKKYSTDIL
 GNLYEVKSKKHPQIIKKG。

上記saCas9n配列において、N579から変異されてSaCas9ニッカーゼを生じることが
 できる残基A579に太字で下線を引いている。

【0254】

例示的SaKKH Cas9の配列を以下に提供する：

【0255】

KRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRRRRH
 RIQRVKKLLFDYNLLTDHSELGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGVHNVNE
 VEEDTGNELSTKEQISRNSKALEEKYVAELQLERLKKDGEVRGSINRFKTSYVKEAKQLL
 KVQKAYHQLDQSFIDTYIDLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEE
 LRSVKYAYNADLYNALNDLNNLVITRDENEKLEYEYKFQIIENVFKQKKKPTLKQIAKEIL
 VNEEDIKGYRVTSTGKPEFTNLKVYHDIKDITARKEIENAELLDQIAKILTIYQSSEDIQEE
 LTNLNSELQTQEEIEQISNLKGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPKKV
 DLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIIKKYGLPNDIIELAREKNSKDAQKMI
 NEMQKRNRQTNERIEEIIRTTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNNPFI
 YEVDHIIPRSVSFDNSFNKVLVKQEEASKKGNRTPFQYLSSSDSKISYETFKKHILNLAK
 GKGRISKTKKEYLLEERDINRFSVQKDFINRNLVDTRYATRGLMNLRLSYFRVNNLDVKV
 KSINGGFTSFLRRKWKFKKERNKGYKHHAEADALIINANADFIFKEWKKLDKAKKVMENQM
 FEEKQAESMPEIETEQEYKEIFITPHQIKHIKDFKDYKYSHRVDKKPNRKLINDTLYSTRK
 DDKGNTLIVNNLNGLYDKDNDKLKKLINKSPEKLLMYHHDPPQTYQKLKLIMEQYGDEKN
 PLYKYYEETGNYLTKEYSKKDNGPVIKKIKYYGNKLNALHDITDDYPNSRNKVVKLSLKP
 YRFDVYLDNGVYKFVTVKNLVDVIKKENYYEVNSKCYEEAKKLKKISNQAEFIASFYNNDLIK
 INGELYRVIGVNNDLLNRIEVMIDITYREYLENMNDKRPPHIKTIASKTQSIKKYSTDIL
 GNLYEVKSKKHPQIIKKG。

上記残基A579は、N579から変異されてSaCas9ニッカーゼを生じることができるもの
 だが、太字で下線を引いている。上記の残基K781、K967、およびH1014は、E781、N
 967、およびR1014から変異されて、SaKKH Cas9を生じることができるが、下線および
 イタリック体で示している。

【0256】

ある態様において、改変されたCas9は、高忠実度Cas9酵素である。いくつかの実施形
 態において、高忠実度Cas9酵素は、SpCas9(K855A)、eSpCas9(1.1)、SpCas9-HF1、
 または高精度Cas9バリエーション(HypaCas9)である。修飾Cas9eSpCas9(1.1)は、HNH/R
 uvCグループと非標的DNA鎖との間の相互作用を弱めるアラニン置換を含み、オフターゲ
 ット部位での鎖分離と切断を防止する。同様に、SpCas9-HF1は、DNAリン酸骨格とのCa
 s9の相互作用を破壊するアラニン置換を介して、オフターゲット編集を低下させる。Hyp
 aCas9は、Cas9の校正と標的識別を増加させる突然変異をREC3ドメインに含む(SpCas9
 N692A/M694A/Q695A/H698A)。3つの高忠実度酵素はすべて、野生型Cas9よりもオ
 フターゲット編集が少ない。以下に、高忠実度Cas9の一例を示す。Cas9に対する高忠実
 度Cas9ドメイン突然変異を太字および下線で示す。

【0257】

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEA
 TRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNI
 VDEVAYHEKYPTIYHLRKKLVDSTDKADRLIYLALAHMIKFRGHFLIEGDLNPDNSDVD
 KLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLLENLIAQLPGEKKNGLFGNLIAL
 SLGLTPNFKSNFDLAEDAKLQLSKDITYDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDI
 LRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDG
 GASQEEFYKFIKPILEKMDGTEELLVKLNRDILLRKQRTFDNGSIPHQIHLGELHAILRRQ
 EDFYPFLKDNREKIEKILTFRIPIYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGAS
 AQSFIERMATAFDKNLPNEKVLPKHSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKA
 IVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLD
 NEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGALSRLKIN
 GIRDKQSGKTILDFLKSDGAFANRNFMA¹LIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAG
 SPAIKKGILQTVKVVDDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGK
 ELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYVDHIVPQSFLK
 DDSIDNKVLT²RS³DKNRGKSDNVPSEEVKKMKNYWRQLLNAKLITQRKFDNLTKAERGG
 LSELDKAGFIKRQLVETRAITKHVAQILDSRMNTKYDENDKLIREVKVITLKS⁴KLVSDFRK
 DFQFYKVREINNYHHAHDAYLNAVVG⁵TALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQ
 EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLS
 MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAK
 VEKGKSKKLKSVKELLGITIMERS⁶SFEKNPIDFLEAKGYKEVKKDLIKLPKYS⁷LFELNGR
 KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKKHYLDEII
 EQISEFSKR⁸VILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTID
 RKRYTSTKEVLDATLIHQ⁹SITGLYETRIDL¹⁰SQLGGD

10

20

【0258】

[ガイドポリヌクレオチド]

本明細書中で使用される、用語「ガイドポリヌクレオチド」とは、標的配列に特異的で
 あり得かつポリヌクレオチドプログラム可能ヌクレオチド結合ドメインタンパク質(例えば
 Cas9またはCpf1)と複合体を形成し得るポリヌクレオチドをいう。一実施形態では、ガイ
 ドポリヌクレオチドはガイドRNAである。本明細書において使用される用語「ガイドRNA
 (gRNA)」およびその文法的等価物は、標的DNAに対して特異的であり得かつCasタンパク
 質と複合体を形成し得るRNAを表し得る。RNA/Cas複合体は、Casタンパク質を標的DNA
 に「ガイド」するのを補助することができる。Cas9/crRNA/tracrRNAは、スパーサーに
 相補的な線状または環状dsDNA標的をエンドヌクレアーゼ的に切断する。crRNAに相補的
 でない標的鎖が最初にエンドヌクレアーゼで切断され、次いでエキソヌクレアーゼ的に3'-
 5'方向にトリムされる。自然界では、DNA結合と切断にはタンパク質と両方のRNAが通常
 必要とされる。しかしながら、crRNAおよびtracrRNAの両方の側面を単一のRNA種に組
 み込むように、単一ガイドRNA(「sgRNA」、または単に「gRNA」)を作製することがで
 きる。例えば、Jinek M. et al., Science 337:816-821(2012)を参照されたい(その
 内容全体が参照により本明細書に組み入れられる)。Cas9は、CRISPR反復配列中の短い
 モチーフ(PAMまたはプロトスパーサー隣接モチーフ)を認識して、「自己」と「非自己」
 を区別するのを助ける。Cas9ヌクレアーゼの配列および構造は、当業者によく知られてい
 る(例えば“Complete genome sequence of an M1 strain of Streptococcus pyogenes.”
 Ferretti, J.J. et al., Natl. Acad. Sci. U.S.A. 98:4658-4663(2001); “CRISPR
 RNA maturation by trans-encoded small RNA and host factor RNase III.” De
 ltcheva E. et al., Nature 471:602-607(2011); および “Programmable dual-RNA-
 guided DNA endonuclease in adaptive bacterial immunity.” Jinek M. et al, Scie
 nce 337:816-821(2012)参照。その内容全体が参照により本明細書に組み入れられる)
 。Cas9オーソログは、限定されるものではないが、S. pyogenes および S. thermophil
 usを含む様々な種において記述されている。さらなる適切なCas9ヌクレアーゼおよび配列

30

40

50

は、本開示に基づいて当業者に明らかとなり得、そのようなCas9ヌクレアーゼおよび配列には、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遺伝子座からのCas9配列が含まれる（その全内容は参照により本明細書に組み込まれる）。ある態様において、Cas9ヌクレアーゼは、不活性(例えば不活化) DNA切断ドメインを有し、すなわち、Cas9はニッカーゼである。

【0259】

ある態様において、ガイドポリヌクレオチドは、少なくとも一つの単一ガイドRNA (「sgRNA」または「gRNA」) である。ある態様において、ガイドポリヌクレオチドは、少なくとも一つのtracrRNAであり、ある態様において、ガイドポリヌクレオチドは、ポリヌクレオチドプログラム可能なDNA結合ドメイン(例:Cas9またはCpf1)を標的ヌクレオチド配列に導くためにPAM配列を必要としない。

【0260】

本願明細書に開示される塩基エディターのポリヌクレオチドプログラム可能ヌクレオチド結合ドメイン(例えばCRISPR由来ドメイン)は、ガイドポリヌクレオチドと会合することによって標的ポリヌクレオチド配列を認識することができる。ガイドポリヌクレオチド(例えばgRNA)は、典型的には一本鎖であり、ポリヌクレオチドの標的配列に部位特異的に結合(すなわち相補的塩基対形成を介して)するようにプログラムすることができ、それによって、ガイド核酸を伴った塩基エディターを標的配列に導く。ガイドポリヌクレオチドはDNAであり得る。ガイドポリヌクレオチドはRNAであり得る。場合によっては、ガイドポリヌクレオチドは天然ヌクレオチド(例えばアデノシン)を含む。場合によっては、ガイドポリヌクレオチドは、非天然(または不自然)ヌクレオチド(例えばペプチド核酸またはヌクレオチドアナログ)を含む。ある場合には、ガイド核酸配列の標的化領域は、長さが少なくとも15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、または30ヌクレオチドであり得る。ガイド核酸の標的化領域は、長さが10~30ヌクレオチドの間、または長さが15~25ヌクレオチドの間、または長さが15~20ヌクレオチドの間であり得る。

【0261】

いくつかの実施形態において、ガイドポリヌクレオチドは、例えば相補的塩基対形成を介して互いに相互作用できる、二つ以上の個別のポリヌクレオチドを含む(例えば二重ガイドポリヌクレオチド)。例えば、ガイドポリヌクレオチドは、CRISPR RNA (crRNA) およびトランス活性化CRISPR RNA (tracrRNA) を含むことができる。例えば、ガイドポリヌクレオチドは、1以上のトランス活性化CRISPR RNA (tracrRNA) を含むことができる。

【0262】

II型CRISPRシステムにおいて、CRISPRタンパク質(例Cas9)による核酸の標的化は、典型的には、標的配列を認識する配列を含む第一のRNA分子(crRNA) と、ガイドRNA-CRISPRタンパク質複合体を安定化させる足場領域を形成する反復配列を含む第二のRNA分子(tracrRNA) との間の相補的塩基対形成を必要とする。このような二重ガイドRNAシステムは、本明細書に開示された塩基エディターを標的ポリヌクレオチド配列に導くためのガイドポリヌクレオチドとして利用され得る。

【0263】

いくつかの実施形態において、本明細書で提供される塩基エディターは、単一ガイドポリヌクレオチド(例えばgRNA)を利用する。いくつかの実施形態において、本明細書で提供される塩基エディターは、デュアル(二重)ガイドポリヌクレオチド(例えばデュアルgRNA)を利用する。いくつかの実施形態において、本明細書で提供される塩基エディターは、一つ以上のガイドポリヌクレオチド(例えば複数のgRNA)を利用する。いくつかの実施形態において、単一ガイドポリヌクレオチドが、本明細書に記載される異なる塩基エディターのために利用される。例えば、単一ガイドポリヌクレオチドを、シチジン塩基エディターおよびアデノシン塩基エディターのために利用することができる。

【0264】

10

20

30

40

50

他の実施形態では、ガイドポリヌクレオチドは、核酸のポリヌクレオチド標的化部分および核酸の足場部分の両方を単一分子(すなわち単一分子ガイド核酸)中に含むことができる。例えば、単一分子ガイドポリヌクレオチドは、単一ガイドRNA (sgRNAまたはgRNA)であり得る。本明細書において、「ガイドポリヌクレオチド配列」という用語は、塩基エディターと相互作用し標的ポリヌクレオチド配列に塩基エディターを導くことができる任意の単一、二重または複数分子核酸を企図する。

【0265】

典型的には、ガイドポリヌクレオチド(例えばcrRNA/trRNA複合体またはgRNA)は、標的ポリヌクレオチド配列を認識しそれに結合することができる配列を含む「ポリヌクレオチド標的セグメント」と、塩基エディターのポリヌクレオチドプログラム可能ヌクレオチド結合ドメイン成分内でガイドポリヌクレオチドを安定化させる「タンパク質結合セグメント」とを含む。或る実施態様では、ガイドポリヌクレオチドのポリヌクレオチド標的セグメントは、DNAポリヌクレオチドを認識してそれに結合し、それによりDNA中の塩基の編集を促進する。他の実施態様では、ガイドポリヌクレオチドのポリヌクレオチド標的セグメントは、RNAポリヌクレオチドを認識しそれに結合し、それによりRNA中の塩基の編集を促進する。ここで、「セグメント」とは、分子の一部分または領域、例えばガイドポリヌクレオチド中の連続したヌクレオチドのストレッチをいう。また、セグメントは、複合体の領域/セクションも表し得、従ってセグメントは2つ以上の分子の領域を含むことができる。例えば、ガイドポリヌクレオチドが複数の核酸分子を含む場合、タンパク質結合セグメントは、例えば相補性の領域に沿ってハイブリダイズした複数の別個の分子の全てまたは一部を含むことができる。ある態様において、二つの別個の分子を含むDNA標的化RNAのタンパク質結合セグメントは、(i) 長さが100塩基対である第一のRNA分子の塩基対40~75; および(ii) 長さが50塩基対である第二のRNA分子の10~25塩基対を含み得る。「セグメント」の定義は、特定の文脈において特に定義されない限り、全塩基対のうちの特定の数に限定されず、所与のRNA分子からの塩基対の特定の数に限定されず、複体内の別個の分子の特定の数に限定されず、任意の全長のRNA分子の領域を含み得、他の分子に対する相補性を有する領域を含み得る。

【0266】

ガイドRNAまたはガイドポリヌクレオチドは、2つ以上のRNA、例えばCRISPR RNA (crRNA) およびトランス活性化crRNA (tracrRNA) を含むことができる。ガイドRNAまたはガイドポリヌクレオチドは、時に、crRNAおよびtracrRNAの一部(例えば機能的部分)の融合によって形成される単一鎖RNA、あるいは単一ガイドRNA (sgRNA) を含み得る。ガイドRNAまたはガイドポリヌクレオチドは、crRNAおよびtracrRNAを含む二重RNAであってもよい。さらに、crRNAは標的DNAとハイブリダイズし得る。

【0267】

上述のように、ガイドRNAまたはガイドポリヌクレオチドは、発現産物であり得る。例えば、ガイドRNAをコードするDNAは、ガイドRNAをコードする配列を含むベクターであり得る。ガイドRNAまたはガイドポリヌクレオチドは、単離されたガイドRNA、またはガイドRNAをコードする配列およびプロモーターを含むプラスミドDNAを細胞にトランスフェクトすることによって、細胞に導入することができる。ガイドRNAまたはガイドポリヌクレオチドは、ウイルス媒介遺伝子送達の使用など、他の方法で細胞に導入することもできる。

【0268】

ガイドRNAまたはガイドポリヌクレオチドを単離することができる。例えば、ガイドRNAは、単離されたRNAの形態で細胞または生物にトランスフェクトされ得る。ガイドRNAは、当技術分野で公知の任意のインビトロ転写系を用いたインビトロ転写によって調製することができる。ガイドRNAは、ガイドRNAをコードする配列を含むプラスミドの形態ではなく、単離されたRNAの形態で細胞に移入され得る。

【0269】

ガイドRNAまたはガイドポリヌクレオチドは、以下の3つの領域を含み得る：染色体配

列中の標的部位に相補的であり得る5'末端における第1の領域、ステムループ構造を形成し得る第2の内部領域、および一本鎖であり得る第3の3'領域。各ガイドRNAが融合タンパク質を特異的な標的部位にガイドするように、各ガイドRNAの第1の領域を異ならせることもできる。さらに、各ガイドRNAの第2および第3の領域は、すべてのガイドRNAにおいて同一であり得る。

【0270】

ガイドRNAまたはガイドポリヌクレオチドの第1の領域は、そのガイドRNAの第1の領域が標的部位と塩基対を形成できるように、染色体配列の標的部位における配列に相補的であり得る。場合によっては、ガイドRNAの第一の領域は、約10ヌクレオチド~25ヌクレオチド(すなわち、10ヌクレオチドからヌクレオチドまで;または約10ヌクレオチドから約25ヌクレオチドまで;または10ヌクレオチドから約25ヌクレオチドまで;または約10ヌクレオチドから25ヌクレオチドまで)またはそれ以上を含むことができる。例えば、ガイドRNAの第一の領域と染色体配列中の標的部位との間の塩基対形成の領域は、長さが約10、11、12、13、14、15、16、17、18、19、20、22、23、24、25またはそれ以上のヌクレオチドであり得るか、またはおよそその長さであり得る。ときに、ガイドRNAの第一の領域は、長さが19、20、もしくは21ヌクレオチド、または約19、20、もしくは21ヌクレオチドであり得る。

【0271】

ガイドRNAまたはガイドポリヌクレオチドは、二次構造を形成する第2の領域を含むこともできる。例えば、ガイドRNAによって形成される二次構造は、ステム(またはヘアピン)およびループを含むことができる。ループとステムの長さはさまざまであり得る。例えば、ループの長さは(約)3から10ヌクレオチドの範囲であり得、ステムの長さは(約)6から20塩基対の範囲であり得る。ステムは、1から10または約10ヌクレオチドの一つ以上のバルジを含むことができる。第二の領域の全長は、約16から60ヌクレオチド長の範囲であり得る。例えば、ループは長さが(約)4ヌクレオチドであり得、ステムは(約)12塩基対であり得る。

【0272】

ガイドRNAまたはガイドポリヌクレオチドは、本質的に一本鎖状態であり得る3'末端の第3の領域も含むこともできる。たとえば、第3の領域は、対象とする細胞のいずれの染色体配列とも相補的でないこともあれば、ガイドRNAの残りの部分と相補的でないこともある。さらに第3の領域の長さはさまざまであり得る。第3の領域は、長さが約4ヌクレオチド以上であり得る。例えば、第3の領域の長さは、(約)5から60ヌクレオチド長の範囲であり得る。

【0273】

ガイドRNAまたはガイドポリヌクレオチドは、遺伝子標的の任意のエクソンまたはイントロンを標的とすることができる。場合によっては、ガイドは遺伝子のエキソン1または2を標的にすることができる;他の場合には、ガイドは遺伝子のエキソン3または4を標的にすることができる。組成物は、全てが同じエクソンを標的とする複数のガイドRNA、または場合によっては、異なるエクソンを標的とする複数のガイドRNAを含むことができる。遺伝子のエキソンとイントロンが標的とされ得る。

【0274】

ガイドRNAまたはガイドポリヌクレオチドは、(約)20ヌクレオチドの核酸配列を標的とし得る。標的核酸は、(約)20ヌクレオチド未満であり得る。標的核酸は、長さが少なくとも(約)5、10、15、16、17、18、19、20、21、22、23、24、25、30、または1~100ヌクレオチドの間の任意の長さであり得る。標的核酸は、長さが多くとも約5、10、15、16、17、18、19、20、21、22、23、24、25、30、40、50ヌクレオチド、または1~100ヌクレオチドの間の任意の長さであり得る。標的核酸配列は、PAMの最初のヌクレオチドの5'直近の(約)20塩基であり得る。ガイドRNAは、核酸配列を標的化し得る。標的核酸は、少なくとも(約)1~10、1~20、1~30、1~40、1~50、1~60、1~70、1~80、1~90、または1~100ヌクレオチドであり得る。

【0275】

ガイドポリヌクレオチド、例えば、ガイドRNAは、別の核酸、例えば、細胞のゲノム中の標的核酸またはプロトスパーサーにハイブリダイズし得る核酸を指すことができる。ガイドポリヌクレオチドはRNAであり得る。ガイドポリヌクレオチドはDNAであり得る。ガイドポリヌクレオチドは、核酸の配列に部位特異的に結合するようにプログラムまたは設計することができる。ガイドポリヌクレオチドは、一ポリヌクレオチド鎖を含み得、単一ガイドポリヌクレオチドと呼ばれ得る。ガイドポリヌクレオチドは、2つのポリヌクレオチド鎖を含み得、二重ガイドポリヌクレオチドと呼ばれ得る。ガイドRNAはRNA分子として細胞や胚に導入され得る。例えば、RNA分子は、インビトロで転写され得、および/または化学的に合成され得る。RNAは、合成DNA分子、例えばgBlocks（登録商標）遺伝子断片から転写され得る。ガイドRNAはRNA分子として細胞や胚に導入され得る。ガイドRNAはまた、非RNA核酸分子、例えばDNA分子の形態で細胞または胚に導入され得る。例えば、ガイドRNAをコードするDNAが、対象の細胞または胚におけるガイドRNAの発現のためにプロモーター制御配列に作動可能に連結され得る。RNAコード配列は、RNAポリメラーゼIII (Pol III)によって認識されるプロモーター配列に作動可能に連結され得る。ガイドRNAを発現するために使用することができるプラスミドベクターは、px330ベクターおよびpx333ベクターを含むが、これらに限定されない。場合によっては、プラスミドベクター(例えばpx333ベクター)は、少なくとも二つのガイドRNAをコードするDNA配列を含むことができる。

【0276】

ガイドポリヌクレオチド（例えばガイドRNA）およびターゲティング配列を選択、設計および検証するための方法は、本明細書中に記載され、当業者に知られている。例えば、核酸塩基エディター系におけるデアミナーゼドメイン(例えばAIDドメイン)の潜在的基質混合性 (promiscuity) の影響を最小限にするために、意図せず脱アミノ化の標的となり得る残基(例えば、標的核酸遺伝子座内のssDNA上に潜在的に位置し得るオフターゲットC残基)の数を最小限にすることができる。さらに、ソフトウェアツールを使用して、標的核酸配列に対応するgRNAを最適化することができ、例えば、ゲノム全体の総オフターゲット活性を最小限化することができる。例えば、*S. pyogenes* Cas9を用いた標的化ドメイン選択肢の各可能性について、(例えばNAGまたはNGGなどの選択されたPAMに先行する)ある数(例えば1、2、3、4、5、6、7、8、9、または10)までのミスマッチ塩基対を含むすべてのオフターゲット配列をゲノムにわたって同定することができる。標的部位に相補的なgRNAの第一の領域を同定することができ、すべての第一の領域(例:crRNA)をその総予測オフターゲットスコアに従ってランク付けすることができ;上位にランクされた標的化ドメインは、最大のオンターゲット活性と最小のオフターゲット活性を持つ可能性が高いものを表す。候補標的化gRNAは、当技術分野で公知の方法および/または本明細書に記載の方法を用いて機能的に評価することができる。

【0277】

非限定的な例として、Cas9と共に使用するためのガイドRNAのcrRNAにおける標的DNAハイブリダイズ配列は、DNA配列探索アルゴリズムを用いて同定することができる。gRNA設計は、Bae S., Park J., & Kim J.-S. Cas-OFFinder: A fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics* 30, 1473-1475 (2014)に記載されているような公開ツールcas-offinderに基づくカスタムgRNA設計ソフトウェアを用いて実施することができる。このソフトウェアは、ガイドのゲノム全体的オフターゲット傾向を計算した後にガイドにスコアを付ける。通常、完全マッチから7個のミスマッチまでの範囲の一致が、17から24までの長さの範囲のガイドについて考慮される。いったんオフターゲットサイトが計算的に決定されると、各ガイドについて集約スコアが計算され、ウェブインターフェースを使用した表形式の出力に要約される。PAM配列に隣接する潜在的な標的部位を同定することに加えて、ソフトウェアは、選択された標的部位から1、2、3または3超のヌクレオチドだけ異なるすべてのPAM隣接配列も同定する。標的核酸配列、例えば標的遺伝子についての

ゲノムDNA配列を得て、反復エレメントを、公に入手可能なツール、例えば、RepeatMaskerプログラムを用いてスクリーニングすることができる。RepeatMaskerは、入力されたDNA配列から、反復エレメントや低複雑性領域を探し出す。所与のクエリー配列に存在する反復の詳細な注釈が出力となる。

【0278】

同定後、ガイドRNA（例えばcrRNA）の第1の領域が、標的部位へのそれらの距離、それらの直交性（orthogonality）、および関連するPAM配列との密接な一致のための5'ヌクレオチドの存在（例えば、関連するPAM（例えば、化膿レンサ球菌についてのNGG PAM、黄色ブドウ球菌についてのNNGRRTまたはNNGRRV PAM）を含むヒトゲノムにおける密接な一致の同定に基づく5'G）に基づいて、階層にランク付けされ得る。本明細書において使用される場合、直交性とは、標的配列に対して最小数のミスマッチを含む、ヒトゲノムにおける配列の数を表す。「高レベルの直交性」または「良好な直交性」は、例えば、意図された標的以外にヒトゲノムにおいて同一の配列を有さず、標的配列において一または二のミスマッチを含有する配列も有さない20マー標的化ドメインを指し得る。良好な直交性を有する標的化ドメインは、オフターゲットDNA切断を最小化するように選択され得る。

【0279】

いくつかの実施形態において、塩基編集活性を検出すること、および候補ガイドポリヌクレオチドを試験することのためにレポーターシステムが使用され得る。いくつかの態様において、レポーターシステムは、塩基編集活性がレポーター遺伝子の発現をもたらす、レポーター遺伝子ベースのアッセイを含むことができる。例えば、レポーターシステムは、不活性化開始コドン、例えば、鋳型鎖上の3'-TAC-5'から3'-CAC-5'への突然変異を含むレポーター遺伝子を含み得る。標的Cの脱アミノ化に成功すると、対応するmRNAは5'-GUG-3'ではなく5'-AUG-3'として転写され、レポーター遺伝子の翻訳を可能にする。適切なレポーター遺伝子は、当業者には明らかであろう。レポーター遺伝子の非限定的な例としては、緑色蛍光タンパク質（GFP）、赤色蛍光タンパク質（RFP）、ルシフェラーゼ、分泌型アルカリホスファターゼ（SEAP）、または、発現が検出可能であり当業者には明らかである他の任意の遺伝子をコードする遺伝子が挙げられる。レポーター系を用いて、多くの異なるgRNAを試験することができ、例えば、標的DNA配列に関してどの残基を、それぞれの対応デアミナーゼが標的とするかを決定することができる。非鋳型鎖を標的とするsgRNAも、特定の塩基編集タンパク質（例えばCas9デアミナーゼ融合タンパク質）のオフターゲット効果を評価するために試験することができる。いくつかの実施形態において、そのようなgRNAは、変異開始コドンがgRNAと塩基対を形成しないように設計することができる。ガイドポリヌクレオチドは、標準リボヌクレオチド、修飾リボヌクレオチド（例えば、プソイドウリジン）、リボヌクレオチド異性体、および/またはリボヌクレオチド類似体を含むことができる。いくつかの実施形態において、ガイドポリヌクレオチドは、少なくとも1つの検出可能な標識を含むことができる。検出可能な標識は、フルオロフォア（例えば、FAM、TMR、Cy3、Cy5、テキサスレッド、オレゴングリーン、Alexa Fluor、Haloタグ、または適切な蛍光色素）、検出タグ（例えば、ビオチン、ジゴキシゲニン等）、量子ドット、または金粒子であり得る。

【0280】

ガイドポリヌクレオチドは、化学的に合成され得るか、酵素的に合成され得るか、またはそれらの組み合わせであり得る。例えば、ガイドRNAは、標準的なホスホロアミダイトベースの固相合成法を用いて合成することができる。あるいは、ガイドRNAをコードするDNAを、ファージRNAポリメラーゼにより認識されるプロモーター制御配列に作動可能に連結することによって、ガイドRNAをインビトロで合成することができる。適切なファージプロモーター配列の例は、T7、T3、SP6プロモーター配列、またはそれらのバリエーションを含む。ガイドRNAが二つの別々の分子（例えば、crRNAとtracrRNA）を含む実施形態において、crRNAは化学的に合成することができ、tracrRNAは酵素的に合成することができる。

【0281】

ある態様において、塩基エディター系は、複数のガイドポリヌクレオチド、例えばgRNAを含み得る。例えば、gRNAは、塩基エディター系に含まれる一つ以上の標的遺伝子座(例えば少なくとも1 gRNA、少なくとも2 gRNA、少なくとも5 gRNA、少なくとも10 gRNA、少なくとも20 gRNA、少なくとも30 gRNA、少なくとも50 gRNA)を標的とすることができる。前記複数のgRNA配列はタンデムに配置され得、好ましくは直接反復によって分離される。

【0282】

ガイドRNAまたはガイドポリヌクレオチドをコードするDNA配列がベクターの一部であってもよい。さらに、ベクターは、追加の発現制御配列(例えばエンハンサー配列、Kozak配列、ポリアデニル化配列、転写終結配列など。)、選択可能なマーカーまたはレポーター配列(例えばGFPまたはピューロマイシンなどの抗生物質耐性遺伝子)、複製起点などを含むことができる。ガイドRNAをコードするDNA分子は線状であってもよい。ガイドRNAまたはガイドポリヌクレオチドをコードするDNA分子も環状であってもよい。

10

【0283】

いくつかの態様において、塩基エディター系の1つ以上の構成要素は、DNA配列によってコードされ得る。このようなDNA配列は、発現系(例えば細胞)と一緒に、または別々に導入することができる。例えば、ポリヌクレオチドプログラム可能なヌクレオチド結合ドメインおよびガイドRNAをコードするDNA配列を細胞に導入することができ、各DNA配列は別個の分子の一部であることができ(例えばポリヌクレオチドプログラム可能なヌクレオチド結合ドメインのコード配列を含む1つのベクターおよびガイドRNAコード配列を含む第2のベクター)、または両方が同じ分子の一部であることができる(例えばポリヌクレオチドプログラム可能なヌクレオチド結合ドメイン及びガイドRNAの両方のためのコード(および調節)配列を含む一つのベクター)。

20

【0284】

RNAガイド型エンドヌクレアーゼおよびガイドRNAをコードするDNA配列を細胞に導入する場合、各DNA配列は、別々の分子の一部であってもよく(例えばRNAガイド型エンドヌクレアーゼのコード配列を含む1つのベクターおよびガイドRNAのコード配列を含む第2のベクター)、または両方が同じ分子の一部であってもよい(例えばRNAガイド型エンドヌクレアーゼおよびガイドRNAの両方のコード(および調節)配列を含む一つのベクター)。

30

【0285】

ガイドポリヌクレオチドは、新しいまたは増強された特徴を有する核酸を提供するための1以上の改変を含むことができる。ガイドポリヌクレオチドは、核酸親和性タグを含むことができる。ガイドポリヌクレオチドは、合成ヌクレオチド、合成ヌクレオチドアナログ、ヌクレオチド誘導体、および/または修飾ヌクレオチドを含むことができる。

【0286】

場合によっては、gRNAまたはガイドポリヌクレオチドは修飾(改変)を含むことができる。修飾は、gRNAまたはガイドポリヌクレオチドの任意の位置に施され得る。単一のgRNAまたはガイドポリヌクレオチドに対して複数の修飾を行うことができる。gRNAまたはガイドポリヌクレオチドは、修飾後に品質管理を受けることができる。場合によっては、品質管理は、PAGE、HPLC、MS、またはそれらの任意の組み合わせを含むことができる。

40

【0287】

gRNAまたはガイドポリヌクレオチドの修飾は、置換、挿入、欠失、化学的修飾、物理的修飾、安定化、精製、またはそれらの任意の組合せであり得る。

【0288】

gRNAまたはガイドポリヌクレオチドはまた、5' アデニル酸、5' グアノシン三リン酸キャップ、5' N7-メチルグアノシン三リン酸キャップ、5' 三リン酸キャップ、3' リン酸、3' チオリン酸、5' リン酸、5' チオリン酸、Cis-Synチミジン二量体、三量体、C12スパーサー、C3スパーサー、C6スパーサー、dスパーサー、PCスパーサー、rスパーサー、スパー

50

サー18、スパーサー9、3'-3'修飾、5'-5'修飾、塩基脱落、アクリジン、アゾベンゼン、
ビオチン、ビオチンBB、ビオチンTEG、コレステロールTEG、デスチオビオチンTEG、D
NP TEG、DNP-X、DOTA、dT-ビオチン、二重ビオチン、ソラレン、ソラレンC2、ソラ
レンC6、TINA、3' DABCYL、ブラックホールクエンチャー1、ブラックホールクエンチャー
2、DABCYL SE、dT-DABCYL、IRDye QC-1、QSY-21、QSY-35、QSY-7、QSY-9
、カルボキシリリンカー、チオールリンカー、2'-デオキシリボヌクレオシド類似体プリン
、2'-デオキシリボヌクレオシド類似体ピリミジン、リボヌクレオシド類似体、2'-O-メチ
ルリボヌクレオシド類似体、糖修飾類似体、ウォブル/ユニバーサル塩基、蛍光色素標識、
2'-フルオロRNA、2'-O-メチルRNA、メチルホスホネート、ホスホジエステルDNA、ホス
ホジエステルRNA、ホスホチオエステルDNA、ホスホロチオネートRNA、UNA、プソイド
リジン-5'-トリホスフェート、5'-メチルシジン-5'-トリホスフェート、またはそれらの任
意の組合せでも修飾され得る。

10

【0289】

場合によっては、修飾は永続的である。その他の場合、修飾は一時的である。場合によ
っては、gRNAまたはガイドポリヌクレオチドに対して複数の修飾が行われる。gRNAまた
はガイドポリヌクレオチドの修飾は、それらのコンホメーション、極性、疎水性、化学反
応性、塩基対形成相互作用、またはそれらの組み合わせなどの、ヌクレオチドの物理化学
的特性を変化させることができる。

【0290】

修飾はホスホロチオエート置換でもあり得る。ある場合には、天然のホスホジエステル
結合は細胞のヌクレアーゼによって急速に分解されやすく、ホスホロチオエート (PS) 結
合置換体を用いたヌクレオチド間結合の修飾は、細胞分解による加水分解に対してより安
定である。修飾は、gRNAまたはガイドポリヌクレオチドの安定性を増大させることがで
きる。修飾は生物学的活性を増強させることもできる。ある場合には、ホスホロチオエート
強化RNA gRNAは、RNアーゼA、RNアーゼT1、子牛血清ヌクレアーゼ、またはそれら
の組み合わせを阻害することができる。これらの特性は、in vivoまたはin vitroでヌクレ
アーゼへの暴露がある可能性が高いアプリケーションにおいて、PS-RNA gRNAの使用を
可能にする。例えば、ホスホロチオエート (PS) 結合をgRNAの5'末端または3'末端におけ
る最後の3~5ヌクレオチドの間に導入することができ、それはエキソヌクレアーゼ分解を
阻害し得る。ある場合には、ホスホロチオエート結合をgRNA全体に加えてエンドヌクレ
アーゼによる攻撃を減らすことができる。

20

30

【0291】

【プロトスパーサ隣接モチーフ】

「プロトスパーサ隣接モチーフ(PAM)」またはPAM様モチーフは、CRISPR細菌適応免疫
系においてCas9ヌクレアーゼによって標的化されるDNA配列の直後の2~6塩基対DNA
配列を指す。いくつかの実施形態では、PAMは5' PAM (すなわちプロトスパーサの5' 末
端の上流に位置する)であり得る。他の実施形態では、PAMは3' PAM (すなわちプロトス
パーサの5' 末端の下流に位置する)であり得る。

【0292】

プロトスパーサ隣接モチーフ(PAM)またはPAM様モチーフは、CRISPR細菌適応免疫
系においてCas9ヌクレアーゼによって標的化されるDNA配列の直後の2~6塩基対DNA配
列を指す。いくつかの実施形態では、PAMは5' PAM (すなわちプロトスパーサの5' 末端
の上流に位置する)であり得る。他の実施形態では、PAMは3' PAM (すなわちプロトス
パーサの5' 末端の下流に位置する)であり得る。PAM配列は標的結合に必須であるが、正確な
配列はCasタンパク質の種類に依存する。

40

【0293】

本明細書で提供される塩基エディターは、標準的または非標準的プロトスパーサ隣接
モチーフ (PAM) 配列を含むヌクレオチド配列に結合することができるCRISPRタンパク質
由来ドメインを含むことができる。PAM部位は、標的ポリヌクレオチド配列に近接するヌ
クレオチド配列である。本開示のいくつかの側面は、異なるPAM特異性を有するCRISPR

50

タンパク質の全部または一部を含む塩基エディターを提供する。例えば、*S. pyogenes*由来のCas9 (spCas9) などのCas9タンパク質は、典型的に、特定の核酸領域に結合するために標準的なNGG PAM配列を必要とし、ここで「NGG」中の「N」はアデニン (A)、チミン(T)、グアニン (G)、またはシトシン (C) であり、Gはグアニンである。PAMはCRISPRタンパク質特異的であり得、異なるCRISPRタンパク質由来ドメインを含む異なる塩基エディター間で異なり得る。PAMは標的配列の5' または3' にあり得る。PAMは、標的配列の上流または下流にあり得る。PAMは、1、2、3、4、5、6、7、8、9、10またはそれ以上のヌクレオチドの長さであり得る。多くの場合、PAMは2~6ヌクレオチドの長さである。

【0294】

[Cas9タンパク質配列]

ある態様において、Cas9ドメインは、*Streptococcus pyogenes*由来のCas9ドメインである (SpCas9)。ある態様において、SpCas9ドメインは、ヌクレアーゼ活性SpCas9、ヌクレアーゼ不活性SpCas9 (SpCas9d)、またはSpCas9ニッカーゼ (SpCas9n) である。いくつかの実施形態において、SpCas9は、D9XD9X突然変異、または本明細書に提供されるアミノ酸配列のいずれかにおける対応する突然変異を含み、ここでXはD以外のアミノ酸である。いくつかの実施形態において、SpCas9は、D9AD9A突然変異、または本明細書に提供されるアミノ酸配列のいずれかにおける対応する突然変異を含む。ある態様において、SpCas9ドメイン、SpCas9dドメインまたはSpCas9nドメインは、非標準PAMを有する核酸配列に結合することができる。ある態様において、SpCas9ドメイン、SpCas9dドメインまたはSpCas9nドメインは、NGG、NGAまたはNGCG PAM配列を有する核酸配列に結合することができる。いくつかの実施形態において、SpCas9ドメインは、D1135X、R1335X、およびT1337X突然変異の1つ以上、または本明細書に提供されるアミノ酸配列のいずれかにおける対応する突然変異突然変異を含み、ここでXは任意のアミノ酸である。いくつかの実施形態において、SpCas9ドメインは、D1135E、R1335Q、およびT1337R突然変異の1つ以上、または本明細書に提供されるアミノ酸配列のいずれかにおける対応する突然変異を含む。いくつかの実施形態において、SpCas9ドメインは、D1135E、R1335Q、およびT1337R突然変異、または本明細書に提供されるアミノ酸配列のいずれかにおける対応する突然変異を含む。いくつかの実施形態において、SpCas9ドメインは、D1135X、R1335X、およびT1337X突然変異のうちの1つ以上、または本明細書において提供されるアミノ酸配列のいずれかにおける対応する突然変異を含み、ここでXは任意のアミノ酸である。いくつかの実施形態において、SpCas9ドメインは、D1135V、R1335Q、およびT1337R突然変異の1以上、または本明細書に提供されるアミノ酸配列のいずれかにおける対応する突然変異を含む。いくつかの実施形態において、SpCas9ドメインは、D1135V、R1335Q、およびT1337R突然変異、または本明細書に提供されるアミノ酸配列のいずれかにおける対応する突然変異を含む。いくつかの実施形態において、SpCas9SpCas9ドメインは、D1135X、G1218X、R1335X、およびT1337X突然変異の1以上、または本明細書に提供されるアミノ酸配列のいずれかにおける対応する突然変異を含み、ここでXは任意のアミノ酸である。いくつかの実施形態において、SpCas9ドメインは、D1135V、G1218R、R1335Q、およびT1337R突然変異の1つ以上、または本明細書に提供されるアミノ酸配列のいずれかにおける対応する突然変異を含む。いくつかの実施形態において、SpCas9いくつかの実施形態において、SpCas9ドメインは、D1135V、G1218R、R1335Q、およびT1337R突然変異、または本明細書において提供されるアミノ酸配列のいずれかにおける対応する突然変異を含む。

【0295】

いくつかの実施形態において、本明細書に提供される融合タンパク質のいずれかのCas9ドメインは、本明細書に記載されるCas9ポリペプチドに対して少なくとも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%の同一性であるアミノ酸配列を含む。いくつかの実施形態

10

20

30

40

50

において、本明細書に提供される融合タンパク質のいずれかのCas9ドメインは、本明細書に記載される任意のCas9ポリペプチドのアミノ酸配列を含む。いくつかの実施形態において、本明細書に提供される融合タンパク質のいずれかのCas9ドメインは、本明細書に記載される任意のCas9ポリペプチドのアミノ酸配列からなる。

【0296】

以下は、例示的なSpCas9配列を提供する：

MDKKYSIGLDIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEA
TRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNI
VDEVAYHEKYPTIYHLRKKLVDSTDKADRLIYLALAHMIKFRGHFLIEGDLNPDNSDVD
KLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIAL
SLGLTPNFKSNFDLAEDAKLQLSKDITYDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDI
LRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDG
GASQEEFYKFIKPILEKMDGTEELLVKLNRDILLRKQRTFDNGSIPHQIHLGELHAILRRQ
EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGAS
AQSFIERMTNFDKNLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKA
IVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLD
NEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLIN
GIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA
GSPAIIKKGILQTVKVVDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGI
KELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK
DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGG
LSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRK
DFQFYKVVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVDVRKMIKSEQ
EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLS
MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAK
VEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDLIIKLPKYSLELENGR
KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKKHYLDEII
EQISEFSKRVLADANLDKVL SAYNKHRRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTID
RKRYTSTKEVLDATLIHQSIITGLYETRIDLSQLGGD

【0297】

以下は、例示的なSpCas9n配列を提供する：

MDKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEA
TRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNI
VDEVAYHEKYPTIYHLRKKLVDSTDKADRLIYLALAHMIKFRGHFLIEGDLNPDNSDVD
KLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIAL
SLGLTPNFKSNFDLAEDAKLQLSKDITYDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDI
LRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDG
GASQEEFYKFIKPILEKMDGTEELLVKLNRDILLRKQRTFDNGSIPHQIHLGELHAILRRQ
EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGAS
AQSFIERMTNFDKNLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKA
IVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLD
NEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLIN
GIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA
GSPAIIKKGILQTVKVVDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGI
KELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK
DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGG
LSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRK
DFQFYKVVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVDVRKMIKSEQ
EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLS
MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAK

VEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDLIIKLPKYSLFELNGR
KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKKHYLDEII
EQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLFTLTNLGAPAAFKYFDTTID
RKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGD

【0298】

以下は、例示的なSpEQR Cas9配列を提供する：

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEA
TRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNI
VDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYLAHAMIKFRGHFLIEGDLNPDNSDVD
KLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIAL
SLGLTPNFKSNFDLAEDAKLQLSKD TYDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDI
LRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDG
GASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ
EDFYFPLKDNREKIEKILTFRIPIYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGAS
AQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKA
IVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLD
NEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLIN
GIRDKQSGKTILDFLKSDGFANRNF MQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA
GSPAIIKKGILQTVKVVD ELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGI
KELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK
DDSIDNKVLT RSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGG
LSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRK
DFQFYK VREINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIKSEQ
EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLS
MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWD PKKYGGFESPTVAYSVLVVAK
VEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDLIIKLPKYSLFELNGR
KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKKHYLDEII
EQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLFTLTNLGAPAAFKYFDTTID
RKQYRSTKEVL DATLIHQ SITGLYETRIDLSQLGGD。

上記SpEQR Cas9配列において、残基E1135、Q1335およびR1337は、D1135、R1335、およびT1337から変異されてSpEQR Cas9を生じることができるものであり、下線、太字が付されている。

【0299】

以下は、例示的なSpVQR Cas9配列を提供する：

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEA
TRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNI
VDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYLAHAMIKFRGHFLIEGDLNPDNSDVD
KLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIAL
SLGLTPNFKSNFDLAEDAKLQLSKD TYDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDI
LRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDG
GASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ
EDFYFPLKDNREKIEKILTFRIPIYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGAS
AQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKA
IVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLD
NEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLIN
GIRDKQSGKTILDFLKSDGFANRNF MQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA
GSPAIIKKGILQTVKVVD ELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGI
KELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK
DDSIDNKVLT RSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGG
LSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRK

DFQFYKVVREINNYHHAHDAYLNAVVGITALIKKYPKLESEFVYGDYKVYDVRKMIKSEQ
EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLS
MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFVSPTVAYSVLVVAK
VEKGKSKKLKSVKELLGITIMERSSEFKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGR
KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEII
EQISEFSKRVLADANLDKVL SAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTID
RKQYRSTKEVL DATLIHQ SITGLYETRIDLSQLGGD。

上記SpVQR Cas9配列において、残基V1135、Q1335、およびR1337は、D1135、R1335、およびT1337から変異されてSpVQR Cas9を生じることができるものであり、下線、太字が付されている。

【0300】

以下は、例示的なSpVRER Cas9配列を提供する：

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEA
TRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNI
VDEVAYHEKYPTIYHLRKKLVDSTDKADRLIYLALAHMIKFRGHFLIEGDLNPDNSDVD
KLFIQLVQTYNQLFEENPINASGVDAKILSARLSKSRLENLIAQLPGEKKNGLFGNLIAL
SLGLTPNFKSNFDLAEDAKLQLSKD TYDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDI
LRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDG
GASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ
EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGAS
AQSFIERMTNFDKNLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKA
IVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLD
NEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLIN
GIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA
GSPAIKKGILQTVKVVDLVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGI
KELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK
DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGG
LSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRK
DFQFYKVVREINNYHHAHDAYLNAVVGITALIKKYPKLESEFVYGDYKVYDVRKMIKSEQ
EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLS
MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFVSPTVAYSVLVVAK
VEKGKSKKLKSVKELLGITIMERSSEFKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGR
KRMLASARELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEII
EQISEFSKRVLADANLDKVL SAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTID
RKEYRSTKEVL DATLIHQ SITGLYETRIDLSQLGGD。

【0301】

以下は、例示的なSpVRQR Cas9配列を提供する：

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEA
TRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNI
VDEVAYHEKYPTIYHLRKKLVDSTDKADRLIYLALAHMIKFRGHFLIEGDLNPDNSDVD
KLFIQLVQTYNQLFEENPINASGVDAKILSARLSKSRLENLIAQLPGEKKNGLFGNLIAL
SLGLTPNFKSNFDLAEDAKLQLSKD TYDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDI
LRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDG
GASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ
EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGAS
AQSFIERMTNFDKNLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKA
IVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLD
NEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLIN
GIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA
GSPAIKKGILQTVKVVDLVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGI

KELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK
DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGG
LSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRK
DFQFYKVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQ
EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLS
MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFVSPTVAYSVLVVAK
VEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDLIIKLPKYSLELENGR
KRMLASARELQKGNELALPSKYVNFLYLASHYEKLKGSPEDEQKQLFVEQHKHYLDEII
EQISEFSKRVLADANLDKVLSAYNKHRRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTID
RKQYRSTKEVLDTLIHQSTGLYETRIDLSQLGGD。

10

上記の残基V1135、R1218、Q1335、およびR1337は、D1135、G1218、R1335、
およびT1337から変異されてSpVRQR Cas9を生じることができ、下線および太字で示
されている。

【0302】

ある態様において、Cas9ドメインは、組換えCas9ドメインである。ある態様において
、組換えCas9ドメインは、SpyMacCas9ドメインである。いくつかの実施形態において
、SpyMacCas9ドメインは、ヌクレアーゼ活性SpyMacCas9、ヌクレアーゼ不活性SpyM
acCas9 (SpyMacCas9d)、またはSpyMacCas9ニックアーゼ(SpyMacCas9n) である。
いくつかの実施形態において、SaCas9ドメイン、SaCas9dドメイン、またはSaCas9nド
メインは、非標準PAMを有する核酸配列に結合することができる。いくつかの実施形態に
おいて、SpyMacCas9ドメイン、SpCas9dドメイン、またはSpCas9nドメインは、NAA
PAM配列を有する核酸配列に結合することができる。

20

例示的SpyMacCas9

MDKKYSIGLDIGTNSVGWAVITDDYKVPSKKFKVLGNTDRHSIKKNLIGALLFGSGETAEA
TRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNI
VDEVAYHEKYPTIYHLRKKLADSTDKADRLIYLALAHMIKFRGHFLIEGDLNPDNSDVD
KLFIQLVQIYNQLFEENPINASRVDAKAILSARLSKSRLENLIAQLPGEKRNGLFGNLI
SLGLTPNFKSNFDLAEDAKLQLSKDITYDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDI
LRVNSEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGG
ASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQE
DFYPFLKDNREKIEKILTFRIPIYVVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGGAS
AQSFIERMTNFDKNLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKA
IVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGAYHDLLKIIKDKDFLD
NEENEDILEDIVLTTLTFEDRGMIEERLKYAHLFDDKVMKQLKRRRYTGWRSLRKLIN
GIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGHSLEHQIANLA
GSPAIIKKGILQTVKIVDELVKVMGHKPNIVIMARENQTTQKGQKNSRERMKRIEEGIKE
LGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFIKDD
SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLS
ELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDF
QFYKVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIG
KATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQ
VNIVKKTEIQTVGQNGGLFDDNPKSPLEVTSPKLVPLKKELNPKKYGGYQKPTTAYPVLL
ITDTKQLIPISVMNKKQFEQNPVKFLRDRGYQQVGKNDFIKLPKYTLVDIGDGIKRLWAS
SKEIHKGNQLVVSCKSQILLYHAHHLSDLSNDYLQNHNNQFQDVLNFNEIISFSKKCKLGK
EHIQKIENVYSNKKNSASIEELAESFIKLLGFTQLGATSPFNFLGVKLNQKQYKGGKDYILP
CTEGTLIRQSITGLYETRVDSLKIGED。

30

40

【0303】

〔高忠実度Cas9ドメイン〕

本開示のいくつかの態様は、高忠実度Cas9ドメインを提供する。いくつかの実施形態に
おいて、高忠実度Cas9ドメインは、対応する野生型Cas9ドメインと比較して、Cas9ドメ

50

インとDNAの糖・リン酸骨格との間の静電相互作用を減少させる1つ以上の突然変異を含む人工Cas9ドメインである。特定の理論に縛られることを望まないが、DNAの糖・リン酸骨格との静電相互作用を減少させた高忠実度Cas9ドメインは、より少ないオフターゲット効果を有し得る。ある態様において、Cas9ドメイン(例えば、野生型Cas9ドメイン)は、Cas9ドメインとDNAの糖・リン酸骨格との間の結合を低減させる一つ以上の突然変異を含む。ある態様において、Cas9ドメインは、Cas9ドメインとDNAの糖・リン酸骨格との間の結合を少なくとも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%、少なくとも10%、少なくとも15%、少なくとも20%、少なくとも25%、少なくとも30%、少なくとも35%、少なくとも40%、少なくとも45%、少なくとも50%、少なくとも55%、少なくとも60%、少なくとも65%、または少なくとも70%だけ低減させる一つ以上の突然変異を含む。

10

【0304】

いくつかの実施形態において、本明細書で提供されるCas9融合タンパク質のいずれかは、N497X、R661X、Q695X、および/またはQ926X突然変異の1以上、または本明細書で提供されるアミノ酸配列のいずれかにおける対応する突然変異を含み、ここでXは任意のアミノ酸である。いくつかの実施形態において、本明細書で提供されるCas9融合タンパク質のいずれかは、N497A、R661A、Q695A、および/またはQ926A突然変異の1以上、または本明細書で提供されるアミノ酸配列のいずれかにおける対応する突然変異を含む。いくつかの実施形態において、Cas9ドメインは、D10A突然変異、または本明細書に提供されるアミノ酸配列のいずれかにおける対応する突然変異を含む。高い忠実度を有するCas9ドメインは当技術分野で公知であり、当業者には明らかであろう。例えば、高い忠実度を有するCas9ドメインは、Kleinstiver, B.P., et al. “High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects.” *Nature* 529, 490-495 (2016); および Slaymaker, I.M., et al. “Rationally engineered Cas9 nucleases with improved specificity.” *Science* 351, 84-88 (2015)に記述されており、それぞれの全内容は参照により本明細書に組み入れられる。以下の高忠実度Cas9ドメインでは、Cas9と比した突然変異が太字および下線で示されている。

20

【0305】

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEA
TRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNI
VDEVAYHEKYPTIYHLRKKLVDSTDKADRLIYLALAHMIKFRGHFLIEGDLNPDNSDVD
KLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIAL
SLGLTPNFKSNFDLAEDAKLQLSKDITYDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDI
LRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDG
GASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ
EDFYFPFLKDNREKIEKILTFRIPIYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGAS
AQSFIERMTAFDKNLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKA
IVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLD
NEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGALSRKLIN
GIRDKQSGKTILDFLKSDGFANRNFMALIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAG
SPAIKKGILQTVKVVDELVKVMGRHKPENIVIMARENQTTQKGQKNSRERMKRIEEGIK
ELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVHDIVPQSFLK
DDSIDNKVLTTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGG
LSELDKAGFIKRQLVETRAITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRK
DFQFYKVRINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQ
EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLS
MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAK
VEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDLIIKLPKYSLELENGR
KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGPEDNEQKQLFVEQHKKHYLDEII
EQISEFSKRVLADANLDKVLSAYNKHDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTID

30

40

50

RKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGD

【0306】

ある場合には、バリエーションCas9タンパク質はH840A、P475A、W476A、N477A、D1125A、W1126A、D1127Aの突然変異を有し、その結果、標的DNAまたはRNAを切断する能力が低下している。このようなCas9タンパク質は、標的DNA(例えば一本鎖標的DNA)を切断する能力は低下しているが、標的DNA(例えば一本鎖標的DNA)に結合する能力は保持している。別の非限定的な例として、いくつかの場合において、バリエーションCas9タンパク質は、D10A、H840A、P475A、W476A、N477A、D1125A、W1126A、およびD1127A突然変異を有し、その結果、ポリペプチドは、標的DNA(例えば一本鎖標的DNA)を切断する能力が低下している。そのようなCas9タンパク質は、標的DNA(例えば一本鎖標的DNA)に結合する能力を低下しているが、標的DNA(例えば一本鎖標的DNA)に結合する能力を保持している。バリエーションCas9タンパク質がW476AおよびW1126A変異を有する場合、またはバリエーションCas9タンパク質がP475A、W476A、N477A、D1125A、W1126A、およびD1127A変異を有する場合、バリエーションCas9タンパク質はPAM配列に効率的に結合しない。したがって、このような場合には、このようなバリエーションCas9タンパク質を結合の方法に用いると、この方法はPAM配列を必要としない。換言すれば、ある場合には、このようなバリエーションCas9タンパク質を結合の方法に用いる場合、この方法はガイドRNAを含み得るが、この方法は、PAM配列の非存在下で行うことができる(したがって、結合の特異性はガイドRNAの標的セグメントによってもたらされる)。上記の効果を達成するために、他の残基が変異され得る(すなわち一方または他方のヌクレアーゼ部分を不活性化する)。非限定的な例として、残基D10、G12、G17、E762、H840、N854、N863、H982、H983、A984、D986、および/またはA987を変更(すなわち置換)することができる。また、アラニン置換以外の変異も好適である。

【0307】

ある態様において、塩基エディターのCRISPRタンパク質由来ドメインは、標準的PAM配列(NGG)を有するCas9タンパク質の全部または一部を含み得る。他の実施形態では、塩基エディターのCas9由来ドメインは、非標準的PAM配列を用いることができる。そのような配列は本技術分野で記述されており当業者には明らかであろう。例えば、非標準的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 aureus* CRISPR-Cas9 by modifying PAM recognition" *Nature Biotechnology*, 33, 1293-1298 (2015)に記述されており、それぞれの全内容を参照によりここに組み込む。

【0308】

いくつかの例において、本明細書に開示される塩基エディターのCRISPRタンパク質由来ドメインによって認識されるPAMは、塩基エディターをコードするインサート(例えばAAVインサート)とは別個のオリゴヌクレオチド上で細胞に提供され得る。そのような場合、別個のオリゴヌクレオチド上にPAMを提供することは、さもなくば標的配列と同じポリヌクレオチド上に隣接するPAMが存在しないために切断することができない標的配列の切断を可能にする。

【0309】

一実施形態において、*S. pyogenes* Cas9 (SpCas9) を、ゲノム工学のためのCRISPRエンドヌクレアーゼとして使用することができる。ただし、他のものも使用され得る。場合によっては、異なるエンドヌクレアーゼを用いて特定のゲノム標的を標的化することができる。場合によっては、非NGG PAM配列を有する合成SpCas9由来バリエーションを使用することができる。さらに、様々な種からの他のCas9オルソログが同定されており、これらの「非SpCas9」は、本開示でも有用になり得る種々のPAM配列に結合し得る。例えば、比較的大きなサイズのSpCas9(約4 kbのコード配列)は、細胞内で効率的に発現することができないSpCas9 cDNAプラスミドをもたらすこともあり得る。逆に、*Staphylococcus*

aureus Cas9 (SaCas9) のコード配列は、SpCas9よりも約1キロ塩基短いため、細胞内で効率的に発現させ得る。SpCas9と同様に、SaCas9エンドヌクレアーゼは、in vitroの哺乳類細胞およびin vivoのマウスにおいて標的遺伝子を修飾する能力がある。ある場合には、Casタンパク質は異なるPAM配列を標的とすることができる。いくつかの場合において、標的遺伝子は、Cas9 PAM、例えば、5' -NGGに隣接し得る。他の場合には、他のCas9オーソログは異なるPAM要件を有し得る。例えば、*S. thermophilus* のもののようなPAM (CRISPR1の場合は5' -NNAGAA、CRISPR3の場合は5' -NGGNG) および *Neisseria meningitidis* のもの (5' -NNNNGATT) のような他のPAMも標的遺伝子に隣接して見出され得る。

【0310】

いくつかの実施形態において、*S. pyogenes*系について、標的遺伝子配列は、5' -NGG PAMの前(すなわち、その5'側)にあり得、20 ntガイドRNA配列が、反対側の鎖と塩基対を形成して、PAMに隣接するCas9切断を媒介することができる。ある場合には、隣接切断は、PAMの(約)3塩基対上流であり得る。ある場合には、隣接切断は、PAMの(約)10塩基対上流であり得る。ある場合には、隣接切断は、PAMの(約)0~20塩基対上流であり得る。例えば、隣接切断は、PAM上流の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、または30塩基対の隣であり得る。隣接切断はPAMの1~30塩基対下流でもあり得る。

【0311】

【核局在化配列 (NLS) を含む融合タンパク質】

いくつかの態様において、本明細書で提供される融合タンパク質は、一つ以上(例:2, 3, 4, 5)の核ターゲティング配列、例えば、核局在化配列(NLS)をさらに含む。一実施形態では、二部(bipartite) NLSが使用される。いくつかの態様において、NLSは、NLSを含むタンパク質の細胞核中への輸入(例えば核輸送によるもの)を促進するアミノ酸配列を含む。いくつかの実施形態において、本明細書において提供される融合タンパク質のいずれかは、核局在化配列(NLS)をさらに含む。いくつかの実施形態において、NLSは融合タンパク質のN末端に融合される。いくつかの実施形態において、NLSは融合タンパク質のC末端に融合される。いくつかの実施形態において、NLSはCas9ドメインのN末端に融合される。いくつかの実施形態において、NLSはnCas9ドメインまたはdCas9ドメインのC末端に融合される。いくつかの実施形態において、NLSはデアミナーゼのN末端に融合される。いくつかの実施形態において、NLSはデアミナーゼのC末端に融合される。いくつかの実施形態において、NLSは、1つ以上のリンカーを介して融合タンパク質に融合される。ある態様において、NLSは、リンカーなしで融合タンパク質に融合される。いくつかの実施形態において、NLSは、本明細書において提供または参照されるNLS配列のいずれか1つのアミノ酸配列を含む。さらなる核局在化配列は当技術分野で公知であり、当業者には明らかであろう。例えば、NLS配列は、Plank et al., PCT/EP2000/011690に記載されており、その内容は、例示的な核局在化配列の開示について参照により本明細書に組み込まれる。ある態様において、NLSは、アミノ酸配列PKKKRKVEGADKRTADGSEFES PKKKRKV, KRTADGSEFESPKKKRKV, KRPAATKKAGQAKKKK, KKTELQTTNAENKTKKL, KRGINDRNFWRGENGRKTR, RKSGKIAAIVVKRPRKPKKKRKV, またはMDSLLMNRRKFLYQFKNVRWAKGRRETYLCを含む。いくつかの実施形態において、NLSはリンカー中に存在するか、またはNLSはリンカー、例えば本明細書に記載されるリンカーによって隣接される。いくつかの実施形態において、N末端またはC末端NLSは、二部NLSである。二部NLSは、比較的短いスペーサー配列によって分離される二つの塩基性アミノ酸クラスターを含む(それゆえにbipartite、2部分と呼ばれ、一部(monopartite) NLSは異なる)。ヌクレオプラスミンのNLSであるKR[PAATKKAGQA]KKKKは遍在的な二部シグナルのプロトタイプであり、塩基性アミノ酸の二つのクラスターが約10アミノ酸のスペーサーによって隔てられたものである。例示的な二部NLSの配列は、PKKKRKVEGADKRTADGSEFES PKKKRKVである。

【0312】

いくつかの態様において、本発明の融合タンパク質は、リンカー配列を含まない。ある態様において、1以上のドメインまたはタンパク質の間にリンカー配列が存在する。

【0313】

本開示の融合タンパク質は、1つ以上のさらなる特徴を含み得ることが理解されるべきである。例えば、いくつかの態様において、融合タンパク質は、阻害因子、細胞質局在化配列、核輸出配列などの輸出配列、または他の局在化配列、ならびに融合タンパク質の可溶化、精製、または検出に有用な配列タグを含むことができる。本明細書において提供される適切なタンパク質タグには、限定されるものではないが、ビオチンカルボキシラーゼキャリアータグ (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、ポリヒスチジンタグ (ヒスチジンタグまたはHis-タグとも呼ばれる)、マルトース結合タンパク質 (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、緑色蛍光タンパク質 (GFP) -タグ、チオレドキシンタグ、S-タグ、ソフタグ(例えばSoftag1、Softag3)、ストレプトタグ、ビオチンリガーゼタグ、FLAsHタグ、V5タグ、およびSBP-タグが含まれる。さらなる適切な配列は、当業者には明らかであろう。いくつかの実施形態において、融合タンパク質は、1つ以上のHisタグを含む。

【0314】

1以上の核局在化配列 (NLS) を含むCRISPR酵素をコードするベクターを使用することができる。例えば、約1、2、3、4、5、6、7、8、9、10個のNLSを使用することができる。CRISPR酵素は、アミノ末端またはその近傍のNLS、カルボキシ末端またはその近傍の約1個、2個、3個、4個、5個、6個、7個、8個、9個、10個またはそれ以上のNLS、またはこれらの任意の組み合わせ(例えばアミノ末端における1以上のNLSおよびカルボキシ末端における1以上のNLS)を含むことができる。複数のNLSが存在する場合、各NLSは他から独立して選択することができ、したがって1つのNLSが1つ以上のコピーで存在することができ、および/または1つ以上のコピーで存在する1つ以上の他のNLSと組み合わせて存在することができる。

【0315】

法において使用されるCRISPR酵素は、約6個のNLSを含むことができる。NLSに最も近いアミノ酸がN-またはC-末端からポリペプチド鎖に沿って約50アミノ酸内(例えば1、2、3、4、5、10、15、20、25、30、40、または50アミノ酸内)にあるとき、そのNLSはN-またはC-末端の近傍にあると考えられる。

【0316】

いくつかの実施形態において、NLSはリンカー中に存在するか、またはNLSはリンカー、例えば本明細書に記載されるリンカーによって隣接される。いくつかの実施形態において、N末端またはC末端のNLSは、二部NLSである。二部NLSは、比較的短いスペーサー配列によって分離される二つの塩基性アミノ酸クラスターを含む(それゆえにbipartite、2部分と呼ばれ、一部(monopartite) NLSは異なる)。ヌクレオプラスミンのNLSであるKR[PAATKKAGQA]KKKKは遍在的な二部シグナルのプロトタイプであり、塩基性アミノ酸の二つのクラスターが約10アミノ酸のスペーサーによって隔てられたものである。例示的な二部NLSの配列は、PKKKRKVEGADKRTADGSEFES PKKKRKVである。

【0317】

いくつかの実施形態において、NLSはリンカー中に存在するか、またはNLSはリンカー、例えば本明細書に記載されるリンカーによって隣接される。いくつかの実施形態において、N末端またはC末端のNLSは、二部NLSである。二部NLSは、比較的短いスペーサー配列によって分離される二つの塩基性アミノ酸クラスターを含む(それゆえにbipartite、2部分と呼ばれ、一部(monopartite) NLSは異なる)。ヌクレオプラスミンのNLSであるKR[PAATKKAGQA]KKKKは遍在的な二部シグナルのプロトタイプであり、塩基性アミノ酸の二つのクラスターが約10アミノ酸のスペーサーによって隔てられたものである。例示的な二部NLSの配列は、PKKKRKVEGADKRTADGSEFES PKKKRKVである。

【0318】

PAM配列は、当該技術分野で公知の任意のPAM配列であり得る。適切なPAM配列には、以下が含まれるが、これらに限定されない：NGG, NGA, NGC, NGN, NGT, NGCG, NGAG, NGAN, NGNG, NGCN, NGCG, NGTN, NNGRRT, NNNRRT, NNGRR(N), TTTV, TYCV, TYCV, TATV, NNNNGATT, NNAGAAW, または NAAAAC。Yはピリミジンであり、Nは任意のヌクレオチド塩基であり、WはAまたはTである。

【0319】

〔排他性が低減されたCas9ドメイン〕

典型的には、*S. pyogenes*由来のCas9 (spCas9) などのCas9タンパク質は、特定の核酸領域に結合するために標準的なNGG PAM配列を必要とし、ここで「NGG」の「N」はアデノシン (A)、チミジン (T) またはシトシン (C) であり、Gはグアノシンである。これは、ゲノム内の所望の塩基を編集する能力を制限し得る。いくつかの実施形態において、本明細書に提供される塩基編集融合タンパク質は、正確な位置、例えばPAMの上流にある標的塩基を含む領域に配置することが必要になり得る。例えばKomor, A.C., et al., “Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage” *Nature* 533, 420-424 (2016)参照（これらの内容全体は、参照により本明細書に組み込まれる）。従って、いくつかの実施形態において、本明細書で提供される融合タンパク質のいずれかは、標準的（例えばNGG）PAM配列を含まないヌクレオチド配列に結合することができるCas9ドメインを含み得る。非標準的PAM配列に結合するCas9ドメインは本技術分野において記述されており当業者には明らかであろう。例えば、非標準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 aureus* CRISPR-Cas9 by modifying PAM recognition” *Nature Biotechnology* 33, 1293-1298 (2015); Nishimasu, H., et al., “Engineered CRISPR-Cas9 nuclease with expanded targeting space” *Science*. 2018 Sep 21;361(6408):1259-1262, Chatterjee, P., et al., “Minimal PAM specificity of a highly similar *S. pyogenes* Cas9 ortholog” *Sci Adv*. 2018 Oct 24;4(10):eaau0766. doi: 10.1126/sciadv.aau0766に記述されており、それぞれの全内容を参照によりここに組み込む。次の表に、いくつかのPAMバリエーションを示す。

10

20

30

40

50

【表 1】

バリエント	PAM
spCas9	NGG
spCas9-VRQR	NGA
spCas9-VRER	NGCG
xCas9 (sp)	NGN
saCas9	NNGRRT
saCas9-KKH	NNNRRT
spCas9-MQKSER	NGCG
spCas9-MQKSER	NGCN
spCas9-LRKIQK	NGTN
spCas9-LRVSQK	NGTN
spCas9-LRVSQL	NGTN
SpyMacCas9	NAA
Cpf1	5' (TTTV)

【0320】

〔核酸塩基編集ドメイン〕

ポリヌクレオチドプログラム可能なヌクレオチド結合ドメインおよび核酸塩基（塩基）編集ドメイン（例えばデアミナーゼドメイン）を含む融合タンパク質を含む塩基エディターを本明細書に記載する。塩基エディターは、標的配列を認識することができるガイドポリヌクレオチドと相互作用することによって、標的ポリヌクレオチド配列中の1以上の塩基を編集するようにプログラムすることができる。標的配列がいったん認識されると、編集が行われるポリヌクレオチド上に塩基エディターが固定され、次いで、塩基エディターのデアミナーゼドメイン成分が標的塩基を編集することができる。

【0321】

ある態様において、核酸塩基編集ドメインは、デアミナーゼドメインである。場合によっては、デアミナーゼドメインは、シトシンデアミナーゼまたはシチジンデアミナーゼであり得る。いくつかの実施形態では、「シトシンデアミナーゼ」および「シチジンデアミナーゼ」という用語は、交換可能に使用され得る。場合によっては、デアミナーゼドメインは、アデニンデアミナーゼまたはアデノシンデアミナーゼであり得る。いくつかの実施形態では、「アデニンデアミナーゼ」および「アデノシンデアミナーゼ」という用語は、交換可能に使用され得る。核酸塩基編集タンパク質の詳細は、国際PCT出願番号PCT/2017/045381 (WO2018/027078) およびPCT/US2016/058344 (WO2017/070632) に記載されており、これらの各々は、その全体が参照により本明細書に組み込まれる。Komor, A.C., et al., “Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage” Nature 533, 420-424 (2016); Gaudelli, N. M., et al., “Programmable base editing of A・T to G・C in genomic DNA without DNA cleavage” Nature 551, 464-471 (2017); および Komor, A.C., et al., “Improved base excision repair inhibition and bacteriophage Mu Gam protein yie

lds C:G-to-T:A base editors with higher efficiency and product purity” Science Advances 3:eaa04774 (2017)も参照（その全内容が参照により本明細書に組み込まれる）。

【0322】

【CからTへの編集】

いくつかの実施形態において、本明細書に開示される塩基エディターは、ポリヌクレオチドの標的シチジン (C) 塩基を脱アミノ化してチミンの塩基対形成特性を有するウリジン (U) を生成することができる、シチジンデアミナーゼを含む融合タンパク質を含む。いくつかの実施形態において、例えば、ポリヌクレオチドが二本鎖(例えばDNA)である場合、ウリジン塩基は次いで、チミジン塩基で置換されて(例えば細胞修復機構により)、C:GからT:Aへの転移を生じさせることができる。他の実施形態において、塩基エディターによる核酸中のCからUへの脱アミノ化は、UからTへの置換を伴うことができない。

【0323】

ポリヌクレオチド中の標的Cを脱アミノ化してUを生じさせることは、本明細書に記載の塩基エディターによって実行することができるタイプの塩基編集の非限定的な例である。別の例において、シチジンデアミナーゼドメインを含む塩基エディターは、シトシン (C) 塩基からグアニン (G) 塩基への変換を媒介することができる。例えば、塩基エディターのシチジンデアミナーゼドメインによるシチジンの脱アミノ化によって生成されたポリヌクレオチドのUは、塩基除去修復機構(例えばウラシルDNAグリコシラーゼ (UDG) ドメインによるもの)によってポリヌクレオチドから切り出され、脱塩基部位を生成することができる。次に、脱塩基部位の反対側の核酸塩基が、Cのような別の塩基で置換され得(例えば塩基修復機構により)、例えばこれは損傷乗り越えポリメラーゼによって行われる。脱塩基部位の反対側の核酸塩基がCで置換されるのが典型的であるが、他の置換(例えばA、G、T)も起こりうる。

【0324】

従って、いくつかの実施形態において、本明細書に記載される塩基エディターは、ポリヌクレオチド中の標的Cを脱アミノ化してUにすることができる脱アミノ化ドメイン(例えばシチジンデアミナーゼドメイン)を含む。さらに、以下に記載するように、塩基エディターは、脱アミノ化から生じるUの、ある態様においてはTまたはGへの、変換を促進させるさらなるドメインを含むことができる。例えば、シチジンデアミナーゼドメインを含む塩基エディターは、さらにウラシルグリコシラーゼ阻害因子 (UGI) ドメインを含み、TによるUの置換を媒介し、CからTへの塩基編集事象を完了することができる。損傷乗り越えポリメラーゼは脱塩基部位の反対側でCの取り込みを促進することができるので(すなわち脱塩基部位でのGの取り込みをもたらし、CからGへの塩基編集事象を完了する)、別の例では塩基エディターは損傷乗り越えポリメラーゼを組み込んで、CからGへの塩基編集の効率を改善させることができる。

【0325】

ドメインとしてシチジンデアミナーゼを含む塩基エディターは、DNA、RNAおよびDNA-RNAハイブリッドを含む任意のポリヌクレオチドにおいて標的Cを脱アミノ化することができる。典型的には、シチジンデアミナーゼは、ポリヌクレオチドの一本鎖部分の文脈に配置されるC核酸塩基を触媒する。ある態様において、標的Cを含むポリヌクレオチド全体が一本鎖であり得る。例えば、塩基エディターに組み込まれたシチジンデアミナーゼは、一本鎖RNAポリヌクレオチド中の標的Cを脱アミノ化することができる。他の実施形態において、シチジンデアミナーゼドメインを含む塩基エディターは、二本鎖ポリヌクレオチドに作用することができるが、標的Cは、脱アミノ化反応の時点で一本鎖状態になるポリヌクレオチドの一部に位置し得る。例えば、NAGPBドメインがCas9ドメインを含む実施形態において、いくつかのヌクレオチドは、Cas9-gRNA-標的DNA複合体の形成の際に適合しないままにしておくことができ、その結果、Cas9「Rループ複合体」が形成される。これらの不對ヌクレオチドは、一本鎖特異的ヌクレオチドデアミナーゼ酵素(例えばシチジンデアミナーゼ)の基質として作用し得る一本鎖DNAのバブルを形成することができる。

【0326】

ある態様において、塩基エディターのシチジンデアミナーゼは、アポリポタンパク質B mRNA編集複合体 (APOBEC) ファミリーデアミナーゼの全部または一部を含むことができる。APOBECは進化的に保存されたシチジンデアミナーゼファミリーである。このファミリーのメンバーはCからUへの編集酵素である。APOBEC様タンパク質のN末端ドメインは触媒ドメインであり、C末端ドメインは偽触媒ドメインである。より具体的には、触媒ドメインは亜鉛依存性のシチジンデアミナーゼドメインであり、シチジンの脱アミノ化にとって重要である。APOBECファミリーメンバーには、APOBEC1、APOBEC2、APOBEC3A、APOBEC3B、APOBEC3C、APOBEC3D(今では「APOBEC3E」はこれを指す)、APOBEC3F、APOBEC3G、APOBEC3H、APOBEC4、および活性化誘導 (シチジン) デアミナーゼが含まれる。多数の改変シチジンデアミナーゼが市販されており、限定されるものではないが、SaBE3、SaKKH-BE3、VQR-BE3、EQR-BE3、VRER-BE3、YE1-BE3、EE-BE3、YE2-BE3、およびYE-BE3が挙げられ、これらはAddgeneから入手可能である(プラスミド85169、85170、85171、85172、85173、85174、85175、85176、85177)。一部の実施形態では、塩基エディターに組み込まれたデアミナーゼは、APOBEC1デアミナーゼの全部または一部を含む。一部の実施形態では、塩基エディターに組み込まれたデアミナーゼは、APOBEC2デアミナーゼの全部または一部を含む。一部の実施形態では、塩基エディターに組み込まれたデアミナーゼは、APOBEC3デアミナーゼの全部または一部を含む。一部の実施形態では、塩基エディターに組み込まれたデアミナーゼは、APOBEC3Aデアミナーゼの全部または一部を含む。一部の実施形態では、塩基エディターに組み込まれたデアミナーゼは、APOBEC3Bデアミナーゼの全部または一部を含む。一部の実施形態では、塩基エディターに組み込まれたデアミナーゼは、APOBEC3Cデアミナーゼの全部または一部を含む。一部の実施形態では、塩基エディターに組み込まれたデアミナーゼは、APOBEC3Dデアミナーゼの全部または一部を含む。一部の実施形態では、塩基エディターに組み込まれたデアミナーゼは、APOBEC3Eデアミナーゼの全部または一部を含む。一部の実施形態では、塩基エディターに組み込まれたデアミナーゼは、APOBEC3Fデアミナーゼの全部または一部を含む。一部の実施形態では、塩基エディターに組み込まれたデアミナーゼは、APOBEC3Gデアミナーゼの全部または一部を含む。一部の実施形態では、塩基エディターに組み込まれたデアミナーゼは、APOBEC3Hデアミナーゼの全部または一部を含む。一部の実施形態では、塩基エディターに組み込まれたデアミナーゼは、APOBEC4デアミナーゼの全部または一部を含む。ある態様において、塩基エディターに組み込まれたデアミナーゼは、活性化誘導デアミナーゼ (AID) の全部または一部を含む。

【0327】

ある態様において、塩基エディターに組み込まれたデアミナーゼは、シチジンデアミナーゼ1 (CDA1) の全部または一部を含む。塩基エディターは、任意の適切な生物(例えばヒトまたはラット)からのデアミナーゼを含むことができることを理解されるべきである。ある態様において、塩基エディターの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、またはマウス由来である。ある態様において、塩基エディターのデアミナーゼドメインは、ラットに由来する(例えばラットAPOBEC1)。ある態様において、塩基エディターのデアミナーゼドメインは、ヒトAPOBEC1である。ある態様において、塩基エディターのデアミナーゼドメインは、PmCDA1である。

【0328】

PmCDA1の塩基配列およびアミノ酸配列、ならびにヒトAIDのCDSの塩基配列およびアミノ酸配列を以下に示す。

> tr | A5H718 | A5H718_PETMA Cytosine deaminase OS=Petromyzon marinus OX=7757 PE=2 SV=1:

MTDAEYVRIHEKLDIYTFKKQFFNNKKS VSHRCYVLFELKRRGERRACFWGYAVNKPQSG
TERGIHAEIFSIRKVEEYLRDNPQGFTINWYSSWSPCADCAEKILEWYNQELRGNGHTLK
IWACKLYYEKNARNQIGLWNL RDNGVGLNVMVSEHYQCCRKIFIQSSHNQLNENRWLE

KT

LKRAEKRRSELSIMIQVKILHTTKSPAV

核酸配列: > EF094822.1 *Petromyzon marinus* isolate PmCDA.21 cytosine deaminase mRNA, complete cds:

TGACACGACACAGCCGTGTATATGAGGAAGGGTAGCTGGATGGGGGGGGGGGAATACG
TTCAGAGAGGA
CATTAGCGAGCGTCTTGTTGGTGGCCTTGAGTCTAGACACCTGCAGACATGACCGACGC
TGAGTACGTGA
GAATCCATGAGAAGTTGGACATCTACACGTTTAAGAAACAGTTTTTCAACAACAAAAA
TCCGTGTCGCA
TAGATGCTACGTTCTCTTTGAATTAAAACGACGGGGTGAACGTAGAGCGTGTTTTTGGG
GCTATGCTGTG
AATAAACCACAGAGCGGGACAGAACGTGGAATTCACGCCGAAATCTTTAGCATTAGAAA
AGTCGAAGAAT
ACCTGCGCGACAACCCCGGACAATTCACGATAAATTGGTACTCATCCTGGAGTCCTTGT
GCAGATTGCGC
TGAAAAGATCTTAGAATGGTATAACCAGGAGCTGCGGGGGAACGGCCACACTTTGAAAA
TCTGGGCTTGC
AAACTCTATTACGAGAAAAATGCGAGGAATCAAATTGGGCTGTGGAACCTCAGAGATAA
CGGGGTTGGGT
TGAATGTAATGGTAAGTGAACACTACCAATGTTGCAGGAAAATATTCATCCAATCGTCG
CACAATCAATT
GAATGAGAATAGATGGCTTGAGAAGACTTTGAAGCGAGCTGAAAAACGACGGAGCGAGT
TGTCATTATG
ATTCAGGTAAAAATACTCCACACCACTAAGAGTCCTGCTGTTTAAGAGGCTATGCGGAT
GTTTTTC

ヒト活性化誘導シチジンデアミナーゼ(AID)のコード配列(CDS)のアミノ酸配列および核酸配列を下記に示す:

> tr | Q6QJ80 | Q6QJ80_HUMAN Activation-induced cytidine deaminase OS=Homo sapiens OX=9606 GN=AICDA PE=2 SV=1

MDSLLMNRKFLYQFKNVRWAKGRRETYLCYVVKRRDSATSFSLDFGYLRNKNKGCHVEL
L
FLRYISDWDLDPGRCYRVTWFTSWSPCYDCARHVADFLRGNPNLSLRIFTARLYFCEDRK
AEPEGLRRLHRAGVQIAIMTFKAPV

核酸配列: > NG_011588.1:5001-15681 Homo sapiens activation induced cytidine deaminase (AICDA), RefSeqGene (LRG_17) on chromosome 12:

AGAGAACCATCATTAATTGAAGTGAGATTTTTCTGGCCTGAGACTTGCAGGGAGGCAAG
AAGACACTCTG
GACACCACTATGGACAGGTAAAGAGGCAGTCTTCTCGTGGGTGATTGCACTGGCCTTCC
TCTCAGAGCAA
ATCTGAGTAATGAGACTGGTAGCTATCCCTTTCTCTCATGTAAGTGTCTGACTGATAAG
ATCAGCTTGAT
CAATATGCATATATATTTTTTGATCTGTCTCCTTTTCTTCTATTTCAGATCTTATACGCTG
TCAGCCCAAT
TCTTTCTGTTTCAGACTTCTCTTGATTTCCCTCTTTTTCATGTGGCAAAGAAGTAGTGC
GTACAATGTA
CTGATTCGTCCTGAGATTTGTACCATGGTTGAACTAATTTATGGTAATAATATTAACA
TAGCAAATCTT
TAGAGACTCAAATCATGAAAAGGTAATAGCAGTACTGTACTAAAAACGGTAGTGCTAAT
TTTCGTAATAA

10

20

30

40

50

TTTTGTAAATATTCAACAGTAAACAACCTTGAAGACACACTTTCCTAGGGAGGCGTTAC
TGAAATAATT
AGCTATAGTAAGAAAATTTGTAATTTTAGAAATGCCAAGCATTCTAAATTAATTGCTTG
AAAGTCACTAT
GATTGTGTCCATTATAAGGAGACAAATTCATTCAAGCAAGTTATTTAATGTTAAAGGCC
CAATTGTTAGG
CAGTTAATGGCACTTTTACTATTAATAATCTTCCATTTGTTTCAGACGTAGCTTAACTT
ACCTCTTAGG
TGTGAATTTGGTTAAGGTCCTCATAATGTCTTTATGTGCAGTTTTTGATAGGTTATTGTC
ATAGAACTTA
TTCTATTCCTACATTTATGATTACTATGGATGTATGAGAATAACACCTAATCCTTATAC
TTTACCTCAAT
TTAACTCCTTTATAAAGAACTTACATTACAGAATAAAGATTTTTTAAAAATATATTTTT
TTGTAGAGACA
GGGTCTTAGCCAGCCGAGGCTGGTCTCTAAGTCCTGGCCCAAGCGATCCTCCTGCCTG
GGCCTCCTAAA
GTGCTGGAATTATAGACATGAGCCATCACATCCAATATACAGAATAAAGATTTTTTAATG
GAGGATTTAAT
GTTCTTCAGAAAATTTTCTTGAGGTCAGACAATGTCAAATGTCTCCTCAGTTTACTG
AGATTTTGAAA
ACAAGTCTGAGCTATAGGTCCTTGTGAAGGGTCCATTGGAAATACTTGTTCAAAGTAAA
ATGGAAAGCAA
AGGTAAAATCAGCAGTTGAAATTCAGAGAAAGACAGAAAAGGAGAAAAGATGAAATTC
AACAGGACAGAA
GGGAAATATATTATCATTAAAGGAGGACAGTATCTGTAGAGCTCATTAGTGATGGCAAAA
TGACTTGGTCA
GGATTATTTTTTAACCCGCTTGTTTCTGGTTTGCACGGCTGGGGATGCAGCTAGGGTTCT
GCCTCAGGGAG
CACAGCTGTCCAGAGCAGCTGTCAGCCTGCAAGCCTGAAACACTCCCTCGGTAAAGTCC
TTCCTACTCAG
GACAGAAATGACGAGAACAGGGAGCTGGAAACAGGCCCTAACCAGAGAAGGGAAGTA
ATGGATCAACAA
AGTTAACTAGCAGGTCAGGATCACGCAATTCATTTCACTCTGACTGGTAACATGTGACA
GAAACAGTGTA
GGCTTATTGTATTTTCATGTAGAGTAGGACCCAAAAATCCACCCAAAGTCCTTTATCTA
TGCCACATCCT
TCTTATCTATACTTCCAGGACACTTTTTCTTCTTATGATAAGGCTCTCTCTCTCTCCAC
ACACACACAC
ACACACACACACACACACACACACACACACACAAACACACACCCCGCCAACCAAGGTGC
ATGTAAAAAGA
TGTAGATTCCTCTGCCTTTCTCATCTACACAGCCCAGGAGGGTAAGTTAATATAAGAGG
GATTTATTGGT
AAGAGATGATGCTTAATCTGTTTAACACTGGGCCTCAAAGAGAGAATTTCTTTTCTTCT
GTACTTATTAA
GCACCTATTATGTGTTGAGCTTATATATACAAAGGGTTATTATATGCTAATATAGTAAT
AGTAATGGTGG
TTGGTACTATGGTAATTACCATAAAAATTATTATCCTTTTAAAATAAAGCTAATTATTA
TTGGATCTTTT
TTAGTATTCATTTTATGTTTTTTATGTTTTTTGATTTTTTAAAAGACAATCTCACCTGTT
ACCCAGGCTG

10

20

30

40

50

GAGTGCAGTGGTGCAATCATAGCTTTCTGCAGTCTTGAACTCCTGGGCTCAAGCAATCC
TCCTGCCTTGG
CCTCCCAAAGTGTTGGGATACAGTCATGAGCCACTGCATCTGGCCTAGGATCCATTTAG
ATTAAAATATG
CATTTTAAATTTTAAAATAATATGGCTAATTTTTACCTTATGTAATGTGTATACTGGCA
ATAAATCTAGT
TTGCTGCCTAAAGTTTAAAGTGCTTTCCAGTAAGCTTCATGTACGTGAGGGGAGACATT
TAAAGTGAAAC
AGACAGCCAGGTGTGGTGGCTCACGCCTGTAATCCCAGCACTCTGGGAGGCTGAGGTGG
GTGGATCGCTT
GAGCCCTGGAGTTCAAGACCAGCCTGAGCAACATGGCAAAACGCTGTTTCTATAACAAA
AATTAGCCGGG
CATGGTGGCATGTGCCTGTGGTCCCAGCTACTAGGGGGCTGAGGCAGGAGAATCGTTGG
AGCCCAGGAGG
TCAAGGCTGCACTGAGCAGTGCTTGCGCCACTGCACTCCAGCCTGGGTGACAGGACCAG
ACCTTGCCTCA
AAAAAATAAGAAGAAAAATTAAAAATAAATGGAAACAACTACAAAGAGCTGTTGTCCTA
GATGAGCTACT
TAGTTAGGCTGATATTTTGGTATTTAACTTTTAAAGTCAGGGTCTGTCACCTGCACTAC
ATTATTAATAAT
ATCAATTCTCAATGTATATCCACACAAAGACTGGTACGTGAATGTTCATAGTACCTTTA
TTCACAAAACC
CCAAAGTAGAGACTATCCAAATATCCATCAACAAGTGAACAAATAAACAAAATGTGCTA
TATCCATGCAA
TGGAATACCACCCTGCAGTACAAAGAAGCTACTTGGGGATGAATCCCAAAGTCATGACG
CTAAATGAAAG
AGTCAGACATGAAGGAGGAGATAATGTATGCCATACGAAATTCTAGAAAATGAAAGTAA
CTTATAGTTAC
AGAAAGCAAATCAGGGCAGGCATAGAGGCTCACACCTGTAATCCCAGCACTTTGAGAGG
CCACGTGGGAA
GATTGCTAGAACTCAGGAGTTCAAGACCAGCCTGGGCAACACAGTGAAACTCCATTCTC
CACAAAAATGG
GAAAAAAGAAAGCAAATCAGTGGTTGTCCTGTGGGGAGGGGAAGGACTGCAAAGAGG
GAAGAAGCTCTG
GTGGGGTGAGGGTGGTGATTGAGTTCTGTATCCTGACTGTGGTAGCAGTTTGGGGTGT
TTACATCCAAA
AATATTCGTAGAATTATGCATCTTAAATGGGTGGAGTTTACTGTATGTAAATTATACCT
CAATGTAAGAA
AAAATAATGTGTAAGAAAACTTTCAATTCTCTTGCCAGCAAACGTTATTCAAATTCCTG
AGCCCTTTACT
TCGCAAATTCTCTGCACTTCTGCCCCGTACCATTAGGTGACAGCACTAGCTCCACAAAT
TGGATAAATGC
ATTTCTGGAAAAGACTAGGGACAAAATCCAGGCATCACTTGTGCTTTCATATCAACCAT
GCTGTACAGCT
TGTGTTGCTGTCTGCAGCTGCAATGGGGACTCTTGATTTCTTTAAGGAACTTGGGTTA
CCAGAGTATTT
CCACAAATGCTATTCAAATTAGTGCTTATGATATGCAAGACACTGTGCTAGGAGCCAGA
AAACAAAGAGG
AGGAGAAATCAGTCATTATGTGGGAACAACATAGCAAGATATTTAGATCATTTTGACTA
GTTAAAAAAGC

10

20

30

40

50

AGCAGAGTACAAAATCACACATGCAATCAGTATAATCCAAATCATGTAAATATGTGCCT
GTAGAAAGACT
AGAGGAATAAACACAAGAATCTTAACAGTCATTGTCATTAGACACTAAGTCTAATTATT
ATTATTAGACA
CTATGATATTTGAGATTTAAAAAATCTTTAATATTTTAAAATTTAGAGCTCTTCTATTTT
TCCATAGTAT
TCAAGTTTGACAATGATCAAGTATTACTCTTTCTTTTTTTTTTTTTTTTTTTTTTTGA
GATGGAGTTT
TGGTCTTGTTGCCCATGCTGGAGTGGAATGGCATGACCATAGCTCACTGCAACCTCCAC
CTCCTGGGTTC
AAGCAAAGCTGTCGCCTCAGCCTCCCGGGTAGATGGGATTACAGGCGCCACCACCACA
CTCGGCTAATG
TTTGTATTTTGTAGTAGAGATGGGGTTTCACCATGTTGGCCAGGCTGGTCTCAAACCTCCT
GACCTCAGAGG
ATCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGATGTAGGCCACTGCGCCCGG
CCAAGTATTGC
TCTTATACATTAAAAAACAGGTGTGAGCCACTGCGCCCAGCCAGGTATTGCTCTTATAC
ATAAAAAATA
GGCCGGTGCAAGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAAGCCAAGGCGGGCAGA
ACACCCGAGGT
CAGGAGTCCAAGGCCAGCCTGGCCAAGATGGTGAAACCCGTCTCTATTAATAAATACAA
ACATTACCTGG
GCATGATGGTGGGCGCCTGTAATCCCAGCTACTCAGGAGGCTGAGGCAGGAGGATCCGC
GGAGCCTGGCA
GATCTGCCTGAGCCTGGGAGGTTGAGGCTACAGTAAGCCAAGATCATGCCAGTATACTT
CAGCCTGGGCG
ACAAAGTGAGACCGTAACAAAAAAAAAAAAAAAAATTTAAAAAAGAAATTTAGATCAAGATC
CAACTGTAAAA
AGTGGCCTAAACACCACATTAAAGAGTTTGGAGTTTATTCTGCAGGCAGAAGAGAACCA
TCAGGGGGTCT
TCAGCATGGGAATGGCATGGTGCACCTGGTTTTTGTGAGATCATGGTGGTGACAGTGTG
GGGAATGTTAT
TTTGGAGGGACTGGAGGCAGACAGACCGGTTAAAAGGCCAGCACAAACAGATAAGGAGG
AAGAAGATGAGG
GCTTGGACCGAAGCAGAGAAGAGCAAACAGGGAAGGTACAAATTCAAGAAATATTGGG
GGGTTTGAATCA
ACACATTTAGATGATTAATTAATATGAGGACTGAGGAATAAGAAATGAGTCAAGGATG
GTTCCAGGCTG
CTAGGCTGCTTACCTGAGGTGGCAAAGTCGGGAGGAGTGGCAGTTTAGGACAGGGGGCA
GTTGAGGAATA
TTGTTTTGATCATTTTGTAGTTTGTAGGTACAAGTTGGACACTTAGGTAAAGACTGGAGGG
GAAATCTGAAT
ATACAATTATGGGACTGAGGAACAAGTTTATTTTATTTTTTGTTCGTTTTCTTGTTGAA
GAACAAATTT
AATTGTAATCCCAAGTCATCAGCATCTAGAAGACAGTGGCAGGAGGTGACTGTCTTGTG
GGTAAGGGTTT
GGGGTCCTTGATGAGTATCTCTCAATTGGCCTTAAATATAAGCAGGAAAAGGAGTTTAT
GATGGATTCCA
GGCTCAGCAGGGCTCAGGAGGGCTCAGGCAGCCAGCAGAGGAAGTCAGAGCATCTTCTT
TGGTTTAGCCC

10

20

30

40

50

AAGTAATGACTTCCTTAAAAAGCTGAAGGAAAATCCAGAGTGACCAGATTATAAACTGT
ACTCTTGCAAT
TTCTCTCCCTCCTCTCACCCACAGCCTCTTGATGAACCGGAGGAAGTTTCTTTACCAATT
CAAAAATGTC
CGCTGGGCTAAGGGTCGGCGTGAGACCTACCTGTGCTACGTAGTGAAGAGGCGTGACAG
TGCTACATCCT
TTTCACTGGACTTTGGTTATCTTCGCAATAAGGTATCAATTAAAGTCGGCTTTGCAAGC
AGTTTAATGGT
CAACTGTGAGTGCTTTTAGAGCCACCTGCTGATGGTATTACTTCCATCCTTTTTTGGCAT
TTGTGTCTCT
ATCACATTCTCAAATCCTTTTTTTTTATTTCTTTTTTCCATGTCCATGCACCCATATTAGA
CATGGCCCAA
AATATGTGATTTAATTCCTCCCCAGTAATGCTGGGCACCCTAATACCACTCCTTCCTTC
AGTGCCAAGAA
CAACTGCTCCCAAACCTGTTTACCAGCTTTCCTCAGCATCTGAATTGCCTTTGAGATTAA
TTAAGCTAAAA
GCATTTTTATATGGGAGAATATTATCAGCTTGTCCAAGCAAAAATTTTAAATGTGAAAA
ACAAATTGTGT
CTTAAGCATTTTTGAAAATTAAGGAAGAAGAATTTGGGAAAAAATTAACGGTGGCTCAA
TTCTGTCTTCC
AAATGATTTCTTTTCCCTCCTACTCACATGGGTGCTAGGCCAGTGAATACATTCAACAT
GGTGATCCCCA
GAAAACTCAGAGAAGCCTCGGCTGATGATTAATTAATTGATCTTTCGGCTACCCGAGA
GAATTACATT
CCAAGAGACTTCTTCACCAAATCCAGATGGGTTTACATAAACTTCTGCCCACGGGTAT
CTCCTCTCTCC
TAACACGCTGTGACGTCTGGGCTTGGTGGAAATCTCAGGGAAGCATCCGTGGGGTGGAAG
GTCATCGTCTG
GCTCGTTGTTTGATGGTTATATTACCATGCAATTTTCTTTGCCTACATTTGTATTGAATA
CATCCCAATC
TCCTTCCTATTTCGGTGACATGACACATTCTATTTCAGAAGGCTTTGATTTTATCAAGCA
CTTTCATTTAC
TTCTCATGGCAGTGCCTATTACTTCTCTTACAATACCCATCTGTCTGCTTTACCAAATC
TATTTCCCT
TTTCAGATCCTCCCAAATGGTCCTCATAAACTGTCTGCCTCCACCTAGTGGTCCAGGT
ATATTTCCACA
ATGTTACATCAACAGGCACTTCTAGCCATTTTCTTCTCAAAGGTGCAAAAAGCAACT
TCATAAACACA
AATTAAATCTTCGGTGAGGTAGTGTGATGCTGCTTCCTCCCAAACCTCAGCGCACTTCGTC
TTCCTCATTCC
ACAAAAACCCATAGCCTTCCTTCACTCTGCAGGACTAGTGCTGCCAAGGGTTTCAGCTCT
ACCTACTGGTG
TGCTCTTTTGAGCAAGTTGCTTAGCCTCTCTGTAACACAAGGACAATAGCTGCAAGCAT
CCCCAAAGATC
ATTGCAGGAGACAATGACTAAGGCTACCAGAGCCGCAATAAAAGTCAGTGAATTTTAGC
GTGGTCCTCTC
TGTCTCTCCAGAACGGCTGCCACGTGGAATTGCTCTTCTCCGCTACATCTCGGACTGG
GACCTAGACCC
TGGCCGCTGCTACCGCGTCACCTGGTTCACCTCCTGGAGCCCCTGCTACGACTGTGCC
GACATGTGGCC

10

20

30

40

50

GACTTTCTGCGAGGGAACCCCAACCTCAGTCTGAGGATCTTCACCGCGCGCCTCTACTT
CTGTGAGGACC
GCAAGGCTGAGCCCGAGGGGCTGCGGCGGCTGCACCGCGCCGGGGTGCAAATAGCCATC
ATGACCTTCAA
AGGTGCGAAAGGGCCTTCCGCGCAGGCGCAGTGCAGCAGCCCGCATTCGGGATTGCGAT
GCGGAATGAAT
GAGTTAGTGGGGAAGCTCGAGGGGAAGAAGTGGGCGGGGATTCTGGTTCACCTCTGGAG
CCGAAATTAAA
GATTAGAAGCAGAGAAAAGAGTGAATGGCTCAGAGACAAGGCCCCGAGGAAATGAGAA
AATGGGGCCAGG
GTTGCTTCTTTCCCTCGATTTGGAACCTGAACTGTCTTCTACCCCCATATCCCCGCCTT
TTTTTCCTT
TTTTTTTTTTGAAGATTATTTTTACTGCTGGAATACTTTTGTAGAAAACCACGAAAGAA
CTTCAAAGCC
TGGGAAGGGCTGCATGAAAATTCAGTTCGTCTCTCCAGACAGCTTCGGCGCATCCTTTT
GGTAAGGGGCT
TCCTCGCTTTTTTAAATTTTCTTCTTCTCTACAGTCTTTTTTGGAGTTTCGTATATTTT
TTATATTTT
TTATTGTTCAATCACTCTCAGTTTTTCATCTGATGAAAACCTTTATTTCTCCTCCACATCAG
CTTTTTCTT
TGCTGTTTCACCATTGAGAGCCCTCTGCTAAGGTTCTTTTCCCTCCCTTTTCTTCTT
TGTTGTTTCA
CATCTTTAAATTTCTGTCTCTCCCCAGGGTTGCGTTTCTTCTGTCAGAAATCTTTTC
TCCTTTTTTT
TTTTTTTTTTTTTTTTTTTTTAAACAAACAAACAAAAAACCCAAAAAACTCTTTCCCAAT
TTACTTTCTT
CCAACATGTTACAAAGCCATCCACTCAGTTTAGAAGACTCTCCGGCCCCACCGACCCCC
AACCTCGTTTT
GAAGCCATTCACTCAATTTGCTTCTCTCTTTCTCTACAGCCCCTGTATGAGGTTGATGAC
TTACGAGACG
CATTTCTGACTTTGGGACTTTGATAGCAACTTCCAGGAATGTCACACACGATGAAATAT
CTCTGCTGAAG
ACAGTGGATAAAAAACAGTCCTTCAAGTCTTCTCTGTTTTTATTCTTCAACTCTCACTTT
CTTAGAGTTT
ACAGAAAAAATATTTATATACGACTCTTTAAAAAGATCTATGTCTTGAAAATAGAGAAG
GAACACAGGTC
TGGCCAGGGACGTGCTGCAATTGGTGCAGTTTTGAATGCAACATTGTCCCCTACTGGGA
ATAACAGAACT
GCAGGACCTGGGAGCATCCTAAAGTGTCACGTTTTTCTATGACTTTTAGGTAGGATGA
GAGCAGAAGGT
AGATCCTAAAAAGCATGGTGAGAGGATCAAATGTTTTTATATCAACATCCTTTATTATT
TGATTCATTTG
AGTTAACAGTGGTGTAGTGATAGATTTTTCTATTCTTTTCCCTTGACGTTTACTTTCAA
GTAACACAAA
CTCTCCATCAGGCCATGATCTATAGGACCTCCTAATGAGAGTATCTGGGTGATTGTGA
CCCCAAACCAT
CTCTCCAAAGCATTAAATATCCAATCATGCGCTGTATGTTTTAATCAGCAGAAGCATGTT
TTTATGTTTGT
ACAAAAGAAGATTGTTATGGGTGGGGATGGAGGTATAGACCATGCATGGTCACCTTCAA
GCTACTTTAAT

10

20

30

40

50

AAAGGATCTTAAATGGGCAGGAGGACTGTGAACAAGACACCCTAATAATGGGTTGATG
TCTGAAGTAGC
AAATCTTCTGGAAACGCAAACCTCTTTTAAGGAAGTCCCTAATTTAGAAACACCCACAAA
CTTCACATATC
ATAATTAGCAAACAATTGGAAGGAAGTTGCTTGAATGTTGGGGAGAGGAAAATCTATTG
GCTCTCGTGGG
TCTCTTCATCTCAGAAATGCCAATCAGGTCAAGGTTTGCTACATTTTGTATGTGTGTGA
TGCTTCTCCCA
AAGGTATATTAACCTATATAAGAGAGTTGTGACAAAACAGAATGATAAAGCTGCGAACCG
TGGCACACGCT
CATAGTTCTAGCTGCTTGGGAGGTTGAGGAGGGAGGATGGCTTGAACACAGGTGTTCAA
GGCCAGCCTGG
GCAACATAACAAGATCCTGTCTCTCAAAAAAAAAAAAAAAAAAAGAAAGAGAGAGGG
CCGGGCGTGGTG
GCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGCCGGGCGGATCACCTGTGGTCA
GGAGTTTGAGA
CCAGCCTGGCCAACATGGCAAACCCCGTCTGTACTCAAATGCAAAAATTAGCCAGGC
GTGGTAGCAGG
CACCTGTAATCCCAGCTACTTGGGAGGCTGAGGCAGGAGAATCGCTTGAACCCAGGAGG
TGGAGGTTGCA
GTAAGCTGAGATCGTGCCGTTGCACTCCAGCCTGGGCGACAAGAGCAAGACTCTGTCTC
AGAAAAAAAAA
AAAAAAGAGAGAGAGAGAGAAAGAGAACAATATTTGGGAGAGAAGGATGGGGAAGCA
TTGCAAGGAAAT
TGTGCTTTATCCAACAAAATGTAAGGAGCCAATAAGGGATCCCTATTTGTCTCTTTTGG
TGTCTATTTGT
CCCTAACAACTGTCTTTGACAGTGAGAAAAATATTCAGAATAACCATATCCCTGTGCCG
TTATTACCTAG
CAACCCTTGCAATGAAGATGAGCAGATCCACAGGAAAACCTTGAATGCACAACTGTCTTA
TTTTAATCTTA
TTGTACATAAGTTTGTAAGAGAGTTAAAAATTGTTACTTCATGTATTCATTTATATTTT
ATATTATTTTG
CGTCTAATGATTTTTTTATTAACATGATTTCTTTTCTGATATATTGAAATGGAGTCTCAA
AGCTTCATAA
ATTTATAACTTTAGAAATGATTCTAATAACAACGTATGTAATTGTAACATTGCAGTAAT
GGTGCTACGAA
GCCATTTCTCTTGATTTTTTAGTAACTTTTATGACAGCAAATTTGCTTCTGGCTCACTTT
CAATCAGTTA
AATAAATGATAAATAATTTTGGAAAGCTGTGAAGATAAAATACCAAATAAAATAATATAA
AAGTGATTTAT
ATGAAGTTAAATAAAAAATCAGTATGATGGAATAAACTTG

【 0 3 2 9 】

本開示の態様に従ってCas9に融合され得る他の例示的デアミナーゼが、以下に提供される。いくつかの実施形態において、それぞれの配列の活性ドメイン、例えば、局在化シグナル(核局在化配列、核外搬出シグナル、細胞質局在化シグナル)を伴わないドメインが使用されることが理解されるべきである。

Human AID:

10

20

30

40

50

MDSLLMNRRKFLYQFKNVRWAKGRRETYLCYVVKRRDSATSFSLDFGYLRNKNGCH
VELLFLRYISDWDLDPGRCYRVTWFTSWSPCYDCARHVADFLRGPNLRLRIFTARLYF
CEDRKAEPGLRRLHRAGVQIAIMTFKDYFYCWNTFVENHERTFKAWEGLHENSVRLS
RQLRRILLPLYEVDDLRLDAFRTLGL (SEQ ID NO:)

(下線: 核局在化配列; 二重下線: 核外搬出シグナル)

Mouse AID:

MDSLLMKQKKFLYHFKNVRWAKGRHETLYCYVVKRRDSATSCSLDFGHLRNKSGCH
VELLFLRYISDWDLDPGRCYRVTWFTSWSPCYDCARHVAEFLRWPNLRLRIFTARLYF
CEDRKAEPGLRRLHRAGVQIGIMTFKDYFYCWNTFVENRERTFKAWEGLHENSVRLT
RQLRRILLPLYEVDDLRLDAFRMLGE (SEQ ID NO:)

10

(下線: 核局在化配列; 二重下線: 核外搬出シグナル)

Canine AID:

MDSLLMKQKFLYHFKNVRWAKGRHETLYCYVVKRRDSATSFSLDFGHLRNKSGCHV
ELLFLRYISDWDLDPGRCYRVTWFTSWSPCYDCARHVADFLRGYPNLSLRIFAARLYFC
EDRKAEPGLRRLHRAGVQIAIMTFKDYFYCWNTFVENREKTFKAWEGLHENSVRLSR
QLRRILLPLYEVDDLRLDAFRTLGL (SEQ ID NO:)

20

(下線: 核局在化配列; 二重下線: 核外搬出シグナル)

Bovine AID:

MDSLLKKQKQFLYQFKNVRWAKGRHETLYCYVVKRRDSPTSFSLDFGHLRNKAGCHV
ELLFLRYISDWDLDPGRCYRVTWFTSWSPCYDCARHVADFLRGYPNLSLRIFTARLYFC
DKERKAEPGLRRLHRAGVQIAIMTFKDYFYCWNTFVENHERTFKAWEGLHENSVRLS
RQLRRILLPLYEVDDLRLDAFRTLGL (SEQ ID NO:)

30

(下線: 核局在化配列; 二重下線: 核外搬出シグナル)

Rat AID

MAVGSKPKAALVGPHWERERIWCFLCSTGLGTQQTGQTSRWLRPAATQDPVSPPRSLL
MKQKFLYHFKNVRWAKGRHETLYCYVVKRRDSATSFSLDFGYLRNKSGCHVELLFL
RYISDWDLDPGRCYRVTWFTSWSPCYDCARHVADFLRGPNLRLRIFTARLTGWGALP
AGLMSPARPSDYFYCWNTFVENHERTFKAWEGLHENSVRLSRRLRRILLPLYEVDDLRL
DAFRTLGL (SEQ ID NO:)

40

(下線: 核局在化配列; 二重下線: 核外搬出シグナル)

Mouse APOBEC-3

50

MGPFCLGCSHRKCYSPIRNLISQETFKFHFKNLGYAKGRKDTFLCYEVTRKDCDSPVSL
HHGVFKNKDNIAEICFLYWFDKVLKVLSPREEFKITWYMSWSPCFECAEQIVRFLATHH
NLSLDIFSSRLYNVQDPETQQNLCLRLVQEGAQVAAMDLYEFKKCWKKFVDNGGRRFR
PWKRLLTNFRYQDSKLQEILRPCYIPVSSSSSTLSNICLTKGLPETRFCVEGRRMDPLSE
EEFYSQFYNNQVRVKHLCYYHRMKPYLCYQLEQFNGQAPLKGCLLSEKGGKQHAELFLDKI
RSMELSQVTITCYLTWSPCPNCAWQLAAFKRDRPDILHIYTSRLYFHWKRPFQKGLCSL
WQSGILVDVMDLPQFTDCWTNFNPKRPFWPWKGLEIISRRTQRRLLRIKESWGLQDL
VNDFGNLQLGPPMS (SEQ ID NO:)

10

(斜体: 核酸編集ドメイン)

Rat APOBEC-3:

MGPFCLGCSHRKCYSPIRNLISQETFKFHFKNRLRYAIDRKDTFLCYEVTRKDCDSPVSL
HHGVFKNKDNIAEICFLYWFDKVLKVLSPREEFKITWYMSWSPCFECAEQVLRFLATH
HNLSLDIFSSRLYNIRDPENQQNLCLRLVQEGAQVAAMDLYEFKKCWKKFVDNGGRRFR
PWKKLLTNFRYQDSKLQEILRPCYIPVSSSSSTLSNICLTKGLPETRFCVERRRVHLLSE
EEFYSQFYNNQVRVKHLCYYHGVKPYLCYQLEQFNGQAPLKGCLLSEKGGKQHAELFLDKI
RSMELSQVTITCYLTWSPCPNCAWQLAAFKRDRPDILHIYTSRLYFHWKRPFQKGLCSL
WQSGILVDVMDLPQFTDCWTNFNPKRPFWPWKGLEIISRRTQRRLLHRIKESWGLQDL
VNDFGNLQLGPPMS (SEQ ID NO:)

20

(斜体: 核酸編集ドメイン)

Rhesus macaque APOBEC-3G:

MVEPMDPRTFVSNNFNNRPILSGLNTVWLCCEVKT KDPSGPPLDAKIFQGVYSKAKYH
PEMRFLRW FHKWRQLHHDQEYKVTWYVSWSPCTRCANSVATFLAKDPKVTLTIFVARLY
YFWKPDYQQALRILCQKRGGPHATMKIMNYNEFQDCWNKFVDGRGKPFKPRNNLPKH
YTLLQATLGELLRHLMDPGTFTSNFNKPKWVSGQHETLYLCYKVERLHNDTWVPLNQH
RGFLRNQAPNIHGFPGKGRHAELCFDLIPFWKLDGQQYRVTCFTSWSPCFSCAQEMAKFIS
NNEHVSLCIFAARIYDDQGRYQEGLRALHRDGAKIAMMNYSEFEYCWDTFVDRQGRPF
QPWDGLDEHSQALSGRLRAI (SEQ ID NO:)

30

(斜体: 核酸編集ドメイン; 下線: 細胞質局在化シグナル)

Chimpanzee APOBEC-3G:

40

50

MKPHFRNPVERMYQDTFSDNFYNRPILSHRNTVWLCYEVKTKGPSRPPLDAKIFRGQV
YSKLKYHPEMRFFHWFSKWRKLHRDQEYEV TWYISWSPCTKCTRDVATFLAEDPKVTLTI
FVARLYYFWDPDYQEALRSLCQKRDGPRA TMKIMNYDEFQHCWSKFVYSQRELFEPW
NNLPKYYILLHIMLGEILRHSM DPPTFTSNFN NELWVRGRHET YLCYEVERLHNDTWVL
LNQRRGFLCNQAPHKHGFLEGRHAELCFLDVIPFWKLDLHQDYRVTCFTSWSPCFSCAQE
MAKFISNNKHVSLCIFAARIYDDQGRCQEGLRTLAKAGAKISIMTYSEFKHCWDTFVDH
QGCPFPQWDGLEEHSQALS GRLRAILQNQGN (SEQ ID NO:)

10

(斜体: 核酸編集ドメイン; 下線: 細胞質局在化シグナル)

Green monkey APOBEC-3G:

MNPQIRNMVEQMEPDIFVYYFNNRPILSGRNTVWLCYEVKTKDPSGPPLDANIFQGKLY
PEAKDHPEMKFLHWFRKWRQLHRDQEYEV TWYVSWSPCTRCANSVATFLAEDPKVTLTI
VARLYYFWKPDYQQALRILCQERGGPHATMKIMNYNEFQHCWNEFVDGQGKPKPRK
NLPKHYTLLHATLGELLRHVMDPGTFTSNFNK PWVSGQRETYLCYKVERSHNDTWV
LLNQHRGFLRNQAPDRHGFPKGRHAELCFLDLIPFWKLDLDDQQYRVTCFTSWSPCFSCAQK
MAKFISNNKHVSLCIFAARIYDDQGRCQEGLRTLHRDGAKIAVMNYSEFEYCWDTFVD
RQGRPFQPWDGLDEHSQALS GRLRAI (SEQ ID NO:)

20

(斜体: 核酸編集ドメイン; 下線: 細胞質局在化シグナル)

Human APOBEC-3G:

MKPHFRNTVERMYRDTFSYNFYNRPI LSRNTVWLCYEVKTKGPSRPPLDAKIFRGQV
YSELKYHPEMRFFHWFSKWRKLHRDQEYEV TWYISWSPCTKCTRDMATFLAEDPKVTLTI
FVARLYYFWDPDYQEALRSLCQKRDGPRA TMKIMNYDEFQHCWSKFVYSQRELFEPW
NNLPKYYILLHIMLGEILRHSM DPPTFTFNFN NEPWVRGRHET YLCYEVERMHNDTWV
LLNQRRGFLCNQAPHKHGFLEGRHAELCFLDVIPFWKLDLDDQDYRVTCFTSWSPCFSCAQ
EMAKFISKNKHVSLCIFTARIYDDQGRCQEGLRTLA EAGAKISIMTYSEFKHCWDTFVD
HQGCPFPQWDGLDEHSQDLS GRLRAILQNQEN (SEQ ID NO:)

30

(斜体: 核酸編集ドメイン; 下線: 細胞質局在化シグナル)

Human APOBEC-3F:

40

MKPHFRNTVERMYRDTFSYNFYNRPILSRRNTVWLCYEYVKTGSPSRPRLDAKIFRGQV
YSQPEHHAEMCFLSWFCGNQLPAYKCFQITWVSWTPCPDCVAKLAEFLAEHPNVTLTIS
AARLYYYWERDYRRALCRLSQAGARVKIMDDEEFAYCWENFVYSEGQPFMPWYKFD
DNYAFLHRTLKEILRNPMEAMYPHIFYHFKNLRKAYGRNESWLCFTMEVVKHHSPVS
WKRGVFRNQVDPETHCHAERCFLSWFCDDILSPNTNYEVTWYTSWSPCPECAGEVAEF
LARHSNVNLTIFTARLYYFWDTDYQEGLRSLSQEGASVEIMGYKDFKYCWENFVYND
DEPFKPWKGLKYNFLFLDSKLQEILE (SEQ ID NO:)

10

(斜体: 核酸編集ドメイン)

Human APOBEC-3B:

MNPQIRNPMERMYRDTFYDNFENEPILYGRSYTWLCYEYVKKRGRSNLLWDTGVFRGQ
VYFKPQYHAEMCFLSWFCGNQLPAYKCFQITWVSWTPCPDCVAKLAEFLSEHPNVTLTI
SAARLYYYWERDYRRALCRLSQAGARVTIMDYEEFAYCWENFVYNegQQFMPWYKF
DENYAFLHRTLKEILRYLMDPDTFTFNFNNDPLVLRRRQTYLCYEYVERLDNGTWVLM
QHMGLFCNEAKNLLCGFYGRHAELRFLDLVPSLQLDPAQIYRVTWFIWSPCFSWGCAGE
VRAFLQENTHVRLRIFAARIYDYDPLYKEALQMLRDAGAQVSIMTYDEFEYCWDTFVY
RQGCPFPQWDGLEEHSQALSGRLRAILQNQGN (SEQ ID NO:)

20

(斜体: 核酸編集ドメイン)

Rat APOBEC-3B:

MQPQGLGNAGMGPVCLGCSHRRPYSPIRNPLKKLYQQTFYHFKNVRYAWGRKNNFL
CYEVNGMDALPVPLRQGVFRKQGHIAELCFIYWFHDKVLRVLSPMEEFKVTWYMSW
SPCSKCAEQVARFLAAHRNLSLAIFSSRLYYYLRNPYQQKLCRLIQEGVHVAAMDLP
KKCWNKFVDNDGQPFRPWRMLRINFSFYDCKLQEIFSRMNLLREDVFYLQFNNSHRVKP
VQNRYYRRKSYLCYQLERANGQEPLKGYLLYKKGEQHVEILFLEKMRSMELSQVRITCYL
TWSPCPNCARQLAAFKKDHDPDLILRIYTSRLYFWRKKFQKGLCTLWRSGLHVDVMDLPQ
FADCWTNFVNPQRPFPRPWNELKNSWRIQRRLRIKESWGL (SEQ ID NO:)

30

Bovine APOBEC-3B: DGWEVAFRSGTVLKAGVLGVSMTEGWAGSGHPGQGACVWTPG
TRNTMNLLREVLFKQQFGNQPRVPAPYYRRKTYLCYQLKQRNDLTLDRCFRNKKQRH
AERFIDKINSLDLNPSQSYKIICYITWSPCPNCANELVNFITRNNHLKLEIFASRLYFHWIK
SFKMGLQDLQNAgisVAVMTHTEFEDCWEQFVDNQSRPFQPWDKLEQYSASIRRLQRI
LTAPI (SEQ ID NO:)

Chimpanzee APOBEC-3B:

MNPQIRNPMEMYQRTFYYNFENEPILYGRSYTWLCYEYKIRRGHSNLLWDTGVFRGQ
MYSQPEHHAEMCFLSWFCGNQLSAYKCFQITWVSWTPCPDCVAKLAKFLAEHPNVTL
TISAARLYYYWERDYRRALCRLSQAGARVKIMDDEEFAYCWENFVYNegQPFMPWYKF
DDNYAFLHRTLKEIIRHLMDDPDTFTFNFNNDPLVLRRHQTYLCYEYVERLDNGTWVLM
QHMGLFCNEAKNLLCGFYGRHAELRFLDLVPSLQLDPAQIYRVTWFIWSPCFSWGCAG
QVRAFLQENTHVRLRIFAARIYDYDPLYKEALQMLRDAGAQVSIMTYDEFEYCWDTFVY
RQGCPFPQWDGLEEHSQALSGRLRAILQVRASSLCMVPHRPPPPQSPGPCPLPCSEPPL
GSLLPTGRPAPSLPFLLTASFSFPPASLPPLPSLSLSPGHLVPVPSFHSLSLSCSIQPPCSSRI
RETEGWASVSKEGRDLG (SEQ ID NO:)

40

Human APOBEC-3C:

50

MNPQIRNPMKAMYPGTFYFQFKNLWEANDRNETWLCFTVEGIKRRSVVSWKTGVFRN
QVDSETH*CHAERCFLSWFCDDILSPNTKYQVTWYTSWSPCPDC*AGEVAEFLARHSNVNLT
IFTARLYYFQYPCYQEGLRSLSQEGVAVEIMDYEDFKYCWENFVYNDNEPFPWKGLK
TNFRLLKRRLRESLQ (SEQ ID NO:)

(斜体: 核酸編集ドメイン)

Gorilla APOBEC-3C

MNPQIRNPMKAMYPGTFYFQFKNLWEANDRNETWLCFTVEGIKRRSVVSWKTGVFRN
QVDSETH*CHAERCFLSWECDDILSPNTNYQVTWYTSWSPCPEC*AGEVAEFLARHSNVNLT
FTARLYYFQDQDYQEGLRSLSQEGVAVKIMDYKDFKYCWENFVYNDDEPFPWKGLK
YNFRFLKRRLQEILE (SEQ ID NO:)

10

Human APOBEC-3A:

MEASPASGPRHLMDPHIFTSNFNNGIGRHKTYLCYEVERLDNGTSVKMDQHRGFLHNQ
AKNLLCGFYGR*HAELRFLDLVPSLQLDPAQIYRVTWFI*SWSPCFSWGCAGEVRAFLQENT
HVRLRIFAARIYDYDPLYKEALQMLRDAGAQVSIMTYDEFKHCWDTFVDHQGCPFQP
WDGLDEHSQALSGRLRAILQNQGN (SEQ ID NO:)

20

(斜体: 核酸編集ドメイン)

Rhesus macaque APOBEC-3A:

MDGSPASRPRHLMDPNTFTFNFNNDLSVRGRHQTYLCYEVERLDNGTWVPMDERRGF
LCNKAKNVPCGDYGCH*VELRFLCEVPSWQLDPAQTYRVTWFI*SWSPCFRRGCAGQVRVF
LQENKHVRLRIFAARIYDYDPLYQEALRTL RDAGAQVSIMTYEEFKHCWDTFVDRQGR
PFQPWDGLDEHSQALSGRLRAILQNQGN (SEQ ID NO:)

30

(斜体: 核酸編集ドメイン)

Bovine APOBEC-3A:

MDEYTFTENFNQGWPSKTYLCYEMERLDGDATIPLDEYKGFVRNKGLDQPEKPCHAE
*LYFLGKIHSWNLDNRNQHYRLTCFI*SWSPCYDCAQKLTTFLKENHHISLHILASRIYTHNRFG
CHQSGLCCELQAAGARITIMTFEDFKHCWETFVDHKGKPFQPWEGLNVKSQALCTELQA
ILKTQQN (SEQ ID NO:)

40

(斜体: 核酸編集ドメイン)

Human APOBEC-3H:

MALLTAETFRLQFNNKRRLRRPYYPKALLCYQLTPQNGSTPTRGYFENKKKCHAEICF
*INEIKSMGLDETQCYQVTCYLTW*SPCSSCAWELVDFIKAHDHLNLGIFASRLYYHWCKPQ
QKGLRLLCGSQVPVEVMGFPKFADCWENFVDHEKPLSFNPYKMLEELDKNSRAIKRRL
ERIKIPGVRAQGRYMDILCDAEV (SEQ ID NO:)

50

(斜体: 核酸編集ドメイン)

Rhesus macaque APOBEC-3H:

MALLTAKTFSLQFNNKRRVKNKPYPRKALLCYQLTPQNGSTPTRGHLKNKKKDHAEIFI
NLIKSMGLDETQCYQVTCYLTWSPCPCAGELVDFIKAHHRHLNLRIFASRLYYHWRPNY
QEGLLLLCGSQVPVEVMGLPEFTDCWENFVDHKEPPSFNPSEKLEELDKNNSQAIKRRLER
IKSRSDVLENGLRSLQLGPVTPSSSIRNSR (SEQ ID NO:)

Human APOBEC-3D:

MNPQIRNPMERMYRDTFYDNFENEPILYGRSYTWLCYEVKIKRGRSNLLWDTGVFRGP
VLPKRQSNHRQEVYFRFEN*HAEMCFLSWFCGNRLPANRRFQITWFWVSWNPCLPCVVKVT*
KFLAEHPNVTLTISAARLYYYRDRDWRWVLLRLHKAGARVKIMDYEDFAYCWENFVC
NEGQPFMPWYKFDDNYASLHRTLKEILRNPMEAMYPHIFYHFKNLLKACGRNESWLC
FTMEVTKHHSVFRKRGVFRNQVDPETH*CHAERCFLSWFCDDILSPNTNYEVTWYTSWSP*
CPECAGEVAEFLARHSNVNLTIFTARLCYFWDTDYQEGLCSLSQEGASVKIMGYKDFV
SCWKNFVYSDDEPFKPKWGLQTNFRLKRLREILQ (SEQ ID NO:)

10

(斜体: 核酸編集ドメイン)

Human APOBEC-1:

MTSEKGPSTGDPTLRRRIEPWEFDVFYDPRELRKEACLLYEIKWGMSRKIWRSSGKNTTN
HVEVNFICKFTSERDFHPSMSCSITWFLSWSPCWECSQAIREFLSRHPGVTLVIYVARLF
WHMDQQNRQGLRDLVNSGVTIQIMRASEYYHCWRNFVNYPGDEAHWPQYPPLWMM
LYALELHCCIILSLPPCLKISRWRQNHLTFFRLHLQNCHYQTIPPHILLATGLIHPSVAWR (SEQ ID NO:)

Mouse APOBEC-1:

MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELRKETCLLYEINWGGRHSVWRHTSQNTSN
HVEVNFLEKFTTERYFRPNTRCSITWFLSWSPCGECSRAITEFLSRHPYVTLFIYIARLYH
HTDQRNRQGLRDLISSGVTIQIMTEQEYCYCWRNFVNYPSPNEAYWPRYPHLWVKLYV
LELYCIILGLPPCLKILRRKQPQLTFFTTITLQTCHYQRIPPHLLWATGLK (SEQ ID NO:)

30

Rat APOBEC-1:

MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELRKETCLLYEINWGGRHSIWRHTSQNTNK
HVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSRAITEFLSRYPHVTLFIYIARLYHH
ADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFVNYSNPSNEAHWPRYPHLWVRLYVLE
LYCIILGLPPCLNILRRKQPQLTFFTTIALQSCHYQRLPPHILWATGLK (SEQ ID NO:)

Human APOBEC-2:

MAQKEEA AVATEAASQNGEDLENLDDPEKLKELIELPPFEIVTGERLPANFFKFQFRNVE
YSSGRNKTFLCYVVEAQGGGQVQASRGYLEDEHAAAHAEAAFFNTILPAFDPALRYNVT
WYVSSSPCAACADRIIKTLSTKNLRLILVGRLFMWEEPEIQAALKKLKEAGCKLRIMKP
QDFEYVWQNFVEQEEGESKAFQPWEDIQENFLYEEKLADILK (SEQ ID NO:)

40

Mouse APOBEC-2:

MAQKEEA AEAAPASQNGDDLENLEDPEKLKELIDLPPFEIVTGVRLPVNFFKFQFRNVE
YSSGRNKTFLCYVVEVQSKGGQAQATQGYLEDEHAGAHAEAAFFNTILPAFDPALKYNVT
WYVSSSPCAACADRILKTLSTKNLRLILVSRLFMWEEPEVQAALKKLKEAGCKLRIMK
PQDFEYIWQNFVEQEEGESKAFEPWEDIQENFLYEEKLADILK (SEQ ID NO:)

Rat APOBEC-2:

MAQKEEA AEAAPASQNGDDLENLEDPEKLKELIDLPPFEIVTGVRLPVNFFKFQFRNVE
YSSGRNKTFLCYVVEAQSKGGQVQATQGYLEDEHAGAHAEAAFFNTILPAFDPALKYNVT
WYVSSSPCAACADRILKTLSTKNLRLILVSRLFMWEEPEVQAALKKLKEAGCKLRIMK

50

PQDFEYLWQNFVEQEEGESKAFEPWEDIQENFLYYEEKLADILK (SEQ ID NO:)

Bovine APOBEC-2:

MAQKEEAAAAAEPASQNGEEVENLEDPEKLKELIELPPFEIVTGERLPAHYFKFQFRNVE
YSSGRNKTFLCYVVEAQSKGGQVQASRGYLEDEHATNHAEAAFFNSIMPTFDPALRYMV
TWYVSSSPCAACADRVKTLNKTKNLRLLLVGRLFMWEEPEIQAALRKLKEAGCRLRIM
KPQDFEYIWQNFVEQEEGESKAFEPWEDIQENFLYYEEKLADILK (SEQ ID NO:)

Petromyzon marinus CDA1 (pmCDAL)

MTDAEYVRIHEKLDIYTFKKQFFNNKKSVSHERCYVLFELKRRGERRACFWGYAVNK PQS
GTERGIHAEIFSIRKVEEYLRDNPQGFTINWYSSWSPCADCAEKILEWYNQELRG NGHTL
KIWACKLYYEKNARNQIGLWNLRDNGVGLNVMVSEHYQCCRKIFIQSSHNQ LNENRWL
EKTLLKRAEKRRSELSFMIQVKILHTTKSPAV (SEQ ID NO:)

Human APOBEC3G D316R D317R

MKPHFRNTVERMYRDTFSYNFYNRPILSRRNTVWLCYEVKTKGSPRPPLDAKIFRGQ VY
SELKYHPEMRFFHWFSKWRKLHRDQEYEVTVYISWSPCTKCTRDMAFLAEDP KVTLT
IFVARLYYFWDPDYQEALRSLCQKRDGPRATMKF NYDEFQHCWSKFVYSQ RELFEPW
NNLPKYIILLHFMGLGEILRHSMDPPTFTFNFNNEPWVRGRHETLYLCYEVEVER MHNDTWV
LLNQRRGFLCNQAPHKHGFLGRHAELCFLDVIPFWKLDLDQDYRVTC FTSWSPCFSCA
QEMAKFISK KHVSLCIFTARIYRRQGRCEGLRTLAEAGAKISF T YSEFKHCWDTFVDH
QGCPFPQWDGLDEHSQDLSGRLRAILQNQEN

(SEQ ID NO:)

Human APOBEC3G chain A

MDPPTFTFNFNNEPWVRGRHETLYLCYEVEVERMHNDTWVLLNQRRGFLCNQAPHKHG FL
EGRHAELCFLDVIPFWKLDLDQDYRVTCFTSWSPCFSCAQEMAKFISKNKHVSLCI FTAR
IYDDQGRCEGLRTLAEAGAKISF TYSEFKHCWDTFVDH QGCPFPQWDGLD EHSQDLS
GRLRAILQ (SEQ ID NO:)

Human APOBEC3G chain A D120R D121R

MDPPTFTFNFNNEPWVRGRHETLYLCYEVEVERMHNDTWVLLNQRRGFLCNQAPHKHG FL
EGRHAELCFLDVIPFWKLDLDQDYRVTCFTSWSPCFSCAQEMAKFISKNKHVSLCI FTAR
IYRRQGRCEGLRTLAEAGAKISFMTYSEFKHCWDTFVDH QGCPFPQWDGLD HSQDL
SGRLRAILQ (SEQ ID NO:)

【0330】

本開示のいくつかの態様は、例えばデアミナーゼドメインに点突然変異を生じさせることによって、本明細書に記載される融合タンパク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、融合タンパク質(例えば塩基エディター)のプロセス特性に影響を及ぼすという認識に基づく。例えば、塩基編集融合タンパク質内のデアミナーゼドメインの触媒活性を減少させるが除去はしない突然変異は、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可能性を低くし、それによって脱アミノ化のウィンドウを狭めることができる。脱アミノ化のウィンドウを狭める能力は、特定の標的残基に隣接する残基の望まれない脱アミノ化を防ぐことができ、それは、オフターゲット効果を減少させるか、または防ぐことができる。

【0331】

一例として、いくつかの実施形態において、塩基エディターに組み込まれたAPOBECデアミナーゼは、rAPOBEC1のH121X、H122X、R126X、R126X、R118X、W90X、W90X、およびR132Xからなる群より選択される1以上の突然変異、または別のAPOBECデアミナーゼにおける1以上の対応する突然変異を含むことができ、ここでXは任意のアミノ酸である。ある態様において、塩基エディターに組み込まれたAPOBECデアミナーゼは、rAPOBEC1のH121R、H122R、R126A、R126E、R118A、W90A、W90Y、およびR132Eからなる群より選択される1以上の突然変異、または別のAPOBECデアミナーゼにおける1以上の対応する突然変異を含むことができる。

【0332】

10

20

30

40

50

ある態様において、塩基エディターに組み込まれたAPOBECデアミナーゼは、hAPOBEC3GのD316X、D317X、R320X、R320X、R313X、W285X、W285X、R326Xからなる群より選択される1以上の突然変異、または別のAPOBECデアミナーゼにおける1以上の対応する突然変異を含むことができ、ここでXは任意のアミノ酸である。いくつかの実施形態において、本明細書で提供される融合タンパク質のいずれかは、hAPOBEC3GのD316R、D317R、R320A、R320E、R313A、W285A、W285Y、R326Eからなる群から選択される1以上の突然変異、または別のAPOBECデアミナーゼにおける1以上の対応する突然変異を含むAPOBECデアミナーゼを含む。

【0333】

ある態様において、塩基エディターに組み込まれたAPOBECデアミナーゼは、rAPOBEC1のH121RおよびH122R突然変異、または別のAPOBECデアミナーゼにおける1つ以上の対応する突然変異を含むことができる。いくつかの実施形態において、塩基エディターに組み込まれたAPOBECデアミナーゼは、rAPOBEC1のR126A突然変異、または別のAPOBECデアミナーゼにおける1つ以上の対応する突然変異を含むAPOBECデアミナーゼを含むことができる。いくつかの実施形態において、塩基エディターに組み込まれたAPOBECデアミナーゼは、rAPOBEC1のR126E突然変異、または別のAPOBECデアミナーゼにおける1つ以上の対応する突然変異を含むAPOBECデアミナーゼを含むことができる。いくつかの実施形態において、塩基エディターに組み込まれたAPOBECデアミナーゼは、rAPOBEC1のR118A突然変異、または別のAPOBECデアミナーゼにおける1つ以上の対応する突然変異を含むAPOBECデアミナーゼを含むことができる。いくつかの実施形態において、塩基エディターに組み込まれたAPOBECデアミナーゼは、rAPOBEC1のW90A突然変異、または別のAPOBECデアミナーゼにおける1つ以上の対応する突然変異を含むAPOBECデアミナーゼを含むことができる。いくつかの実施形態において、塩基エディターに組み込まれたAPOBECデアミナーゼは、rAPOBEC1のW90Y突然変異、または別のAPOBECデアミナーゼにおける1つ以上の対応する突然変異を含むAPOBECデアミナーゼを含むことができる。いくつかの実施形態において、塩基エディターに組み込まれたAPOBECデアミナーゼは、rAPOBEC1のR132E突然変異、または別のAPOBECデアミナーゼにおける1つ以上の対応する突然変異を含むAPOBECデアミナーゼを含むことができる。いくつかの実施形態において、塩基エディターに組み込まれたAPOBECデアミナーゼは、rAPOBEC1のW90YおよびR126E突然変異、または別のAPOBECデアミナーゼにおける1つ以上の対応する突然変異を含むAPOBECデアミナーゼを含むことができる。いくつかの実施形態において、塩基エディターに組み込まれたAPOBECデアミナーゼは、rAPOBEC1のR126EおよびR132E突然変異、または別のAPOBECデアミナーゼにおける1つ以上の対応する突然変異を含むAPOBECデアミナーゼを含むことができる。いくつかの実施形態において、塩基エディターに組み込まれたAPOBECデアミナーゼは、rAPOBEC1のW90Y、R126E、およびR132E突然変異、または別のAPOBECデアミナーゼにおける1つ以上の対応する突然変異を含むAPOBECデアミナーゼを含むことができる。

【0334】

いくつかの実施形態において、塩基エディターに組み込まれたAPOBECデアミナーゼは、hAPOBEC3GのD316RおよびD317R突然変異、または別のAPOBECデアミナーゼにおける1つ以上の対応する突然変異を含むAPOBECデアミナーゼを含むことができる。いくつかの実施形態において、本明細書で提供される融合タンパク質のいずれかは、hAPOBEC3GのR320A突然変異、または別のAPOBECデアミナーゼにおける1つ以上の対応する突然変異を含むAPOBECデアミナーゼを含む。いくつかの実施形態において、塩基エディターに組み込まれたAPOBECデアミナーゼは、hAPOBEC3GのR320E突然変異、または別のAPOBECデアミナーゼにおける1つ以上の対応する突然変異を含むAPOBECデアミナーゼを含むことができる。いくつかの実施形態において、塩基エディターに組み込まれたAPO

10

20

30

40

50

BECデアミナーゼは、hAPOBEC3GのR313A突然変異、または別のAPOBECデアミナーゼにおける1つ以上の対応する突然変異を含むAPOBECデアミナーゼを含むことができる。いくつかの実施形態において、塩基エディターに組み込まれたAPOBECデアミナーゼは、hAPOBEC3GのW285A突然変異、または別のAPOBECデアミナーゼにおける1つ以上の対応する突然変異を含むAPOBECデアミナーゼを含むことができる。いくつかの実施形態において、塩基エディターに組み込まれたAPOBECデアミナーゼは、hAPOBEC3GのW285Y突然変異、または別のAPOBECデアミナーゼにおける1つ以上の対応する突然変異を含むAPOBECデアミナーゼを含むことができる。いくつかの実施形態において、塩基エディターに組み込まれたAPOBECデアミナーゼは、hAPOBEC3GのR326E突然変異、または別のAPOBECデアミナーゼにおける1つ以上の対応する突然変異を含むAPOBECデアミナーゼを含むことができる。いくつかの実施形態において、塩基エディターに組み込まれたAPOBECデアミナーゼは、hAPOBEC3GのW285YおよびR320E突然変異、または別のAPOBECデアミナーゼにおける1つ以上の対応する突然変異を含むAPOBECデアミナーゼを含むことができる。いくつかの実施形態において、塩基エディターに組み込まれたAPOBECデアミナーゼは、hAPOBEC3GのR320EおよびR326E突然変異、または別のAPOBECデアミナーゼにおける1つ以上の対応する突然変異を含むAPOBECデアミナーゼを含むことができる。いくつかの実施形態において、塩基エディターに組み込まれたAPOBECデアミナーゼは、hAPOBEC3GのW285YおよびR326E突然変異、または別のAPOBECデアミナーゼにおける1つ以上の対応する突然変異を含むAPOBECデアミナーゼを含むことができる。いくつかの実施形態において、塩基エディターに組み込まれたAPOBECデアミナーゼは、hAPOBEC3GのW285Y、R320E、およびR326E突然変異、または別のAPOBECデアミナーゼにおける1つ以上の対応する突然変異を含むAPOBECデアミナーゼを含むことができる。

【0335】

限定されるものではないが、SaBE3、SaKKH-BE3、VQR-BE3、EQR-BE3、VRER-BE3、YE1-BE3、EE-BE3、YE2-BE3、およびYE-BE3を含む、多数の改変シチジンデアミナーゼがAddgeneから市販されている(プラスミド85169、85170、85171、85172、85173、85174、85175、85176、85177)。

【0336】

CからTへの核酸塩基編集タンパク質の詳細は、国際PCT出願番号PCT/US2016/058344 (WO2017/070632) およびKomor, A.C., et al., “Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage” Nature 533, 420-424 (2016) に記載されており、これらの全内容は参照により本明細書に組み入れられる。

【0337】

[AからGへの編集]

いくつかの実施形態において、本明細書に記載される塩基エディターは、アデノシンデアミナーゼを含むデアミナーゼドメインを含むことができる。塩基エディターのこのようなアデノシンデアミナーゼドメインは、Aを脱アミノ化して、Gの塩基対形成特性を示すイノシン (I) を形成することによって、アデニン (A) 核酸塩基からグアニン (G) 核酸塩基への編集することを促進することができる。アデノシンデアミナーゼは、デオキシリボ核酸 (DNA) 中のデオキシアデノシン残基のアデニンを脱アミノ化すること(すなわちアミン基を除去すること)ができる。

【0338】

いくつかの実施形態において、本明細書において提供される核酸塩基エディターは、1つ以上のタンパク質ドメインを融合させ、それによって融合タンパク質を生成することによって作製することができる。ある態様において、本明細書において提供される融合タンパク質は、融合タンパク質の塩基編集活性(例えば効率、選択性、特異性)を改善する一つ以上の特徴を含む。例えば、本明細書中に提供される融合タンパク質は、ヌクレアーゼ活性が低下したCas9ドメインを含むことができる。いくつかの実施形態において、本明細書

中で提供される融合タンパク質は、ヌクレアーゼ活性を有さないCas9ドメイン (dCas9)、または、Cas9ニッカーゼ (nCas9) と呼ばれる、二本鎖DNA分子の一鎖を切断するCas9ドメインを有することができる。いかなる特定の理論にも拘束されることを望まないが、触媒残基(例えばH840)の存在が、標的Aと反対のTを含有する非編集(例えば非脱アミノ化)鎖を切断するCas9の活性を維持する。Cas9の触媒残基の突然変異(例えばD10からA10)は、標的A残基を含有する編集鎖の切断を妨げる。このようなCas9バリエーションは、gRNAによって規定される標的配列に基づく特定の位置で単一鎖DNA切断(ニック)を生じさせることができ、非編集鎖の修復を導き、最終的には編集鎖においてTからCへの変化をもたらす。ある態様において、AからGへの塩基エディターは、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をさらに含む。いかなる特定の理論にも束縛されることを望まないが、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することができ、これが、塩基エディターの活性または効率を改善させることができる。

【0339】

アデノシンデアミナーゼを含む塩基エディターは、DNA、RNAおよびDNA-RNAハイブリッドを含む任意のポリヌクレオチドに作用することができる。特定の実施形態では、アデノシンデアミナーゼを含む塩基エディターは、RNAを含むポリヌクレオチドの標的Aを脱アミノ化することができる。例えば、塩基エディターは、RNAポリヌクレオチドおよび/またはDNA-RNAハイブリッドポリヌクレオチドの標的Aを脱アミノ化することができるアデノシンデアミナーゼドメインを含むことができる。一実施形態によると、塩基エディターに組み込まれたアデノシンデアミナーゼは、RNAに作用するアデノシンデアミナーゼ(ADAR、例えばADAR1またはADAR2)の全部または一部を含む。別の実施形態では、塩基エディターに組み込まれたアデノシンデアミナーゼは、tRNAに作用するアデノシンデアミナーゼ(ADAT)の全部または一部を含む。アデノシンデアミナーゼドメインを含む塩基エディターはまた、DNAポリヌクレオチドのA核酸塩基を脱アミノ化する能力も有し得る。一実施形態では、塩基エディターのアデノシンデアミナーゼドメインは、1以上の突然変異を含むADATの全部または一部を含み、これがDNA中の標的AをADATが脱アミノ化することを可能にする。例えば、塩基エディターは、D108N、A106V、D147Y、E155V、L84F、H123Y、I157F、または別のアデノシンデアミナーゼにおける対応する突然変異の1以上を含む大腸菌由来のADAT (EcTadA)の全部または一部を含むことができる。いくつかの実施形態において、TadAデアミナーゼは、大腸菌TadAデアミナーゼ(ecTadA)またはその断片である。例えば、切り詰められたecTadAは、全長ecTadAに対して一つ以上のN末端アミノ酸を欠失していてもよい。いくつかの実施形態において、切り詰められたecTadAは、全長ecTadAに対して1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、6、17、18、19または20個のN末端アミノ酸残基を欠失していてもよい。いくつかの実施形態において、切り詰められたecTadAは、全長ecTadAに対して1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、6、17、18、19または20個のC末端アミノ酸残基を欠失していてもよい。いくつかの実施形態において、ecTadAデアミナーゼは、N末端メチオニンを含まない。いくつかの実施形態において、TadAデアミナーゼは、N末端が切り詰められたTadAである。特定の実施形態において、TadAは、その全体が参照により本明細書に組み込まれるPCT/US2017/045381に記載されているTadAのいずれかである。

【0340】

アデノシンデアミナーゼは、任意の適切な生物に由来することができる。ある態様において、アデノシンデアミナーゼは、原核生物由来である。ある態様において、アデノシンデアミナーゼは、細菌由来である。ある態様において、アデノシンデアミナーゼは、*Escherichia coli*、*Staphylococcus aureus*、*Salmonella typhi*、*Shewanella putrefaciens*、*Haemophilus influenzae*、*Caulobacter crescentus*、または*Bacillus subtilis*に由来する。ある態様において、アデノシンデアミナーゼは、大腸菌由来である。ある態

10

20

30

40

50

様において、アデニンデアミナーゼは、天然に存在するアデノシンデアミナーゼが本明細書に提供される突然変異のいずれかに対応する一つ以上の突然変異を含むものである(例えば、ecTadAにおける突然変異)。任意の相同タンパク質中の対応する残基は、例えば、配列アラインメントおよび相同残基の決定によって同定することができる。本明細書中に記載された突然変異のいずれか(例えば、ecTadAで同定された突然変異のいずれか)に対応する、任意の天然アデノシンデアミナーゼ(例えば、ecTadAに対して相同性を有するもの)における突然変異を、それに応じて生成することができる。

【0341】

[TadA (tRNAアデノシンデアミナーゼA)]

特定の実施形態では、TadAは、その全体が参照により本明細書に組み込まれるPCT/US 2017/045381(国際公開第2018/027078号)に記載されているTadAのいずれかである。

【0342】

1つの実施形態において、本発明の融合タンパク質は、TadA7.10に連結された野生型TadAがCas9ニッカーゼに連結されたものを含む。特定の実施形態では、融合タンパク質は、単一のTadA7.10ドメイン(例えば、モノマーとして提供される)を含む。他の実施形態では、ABE7.10エディターは、TadA7.10およびTadA(wt)を含み、これらはヘテロ二量体を形成することができる。関連するアミノ酸配列は以下のとおりである。

(M)SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAG
SLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAKKAQSSTD

これは「TadA参照配列」または野生型TadA (TadA(wt)) と呼ばれる。

TadA7.10アミノ酸配列

(M)SEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNRRVIGEGWNRAIGLHDPTAH
AEIMALRQGGLVMQNYRLIDATLYVTTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGS
LMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTD

【0343】

ある態様において、アデノシンデアミナーゼは、本明細書に提供されるアデノシンデアミナーゼのいずれかに記載のアミノ酸配列のいずれかに対して少なくとも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%の同一性であるアミノ酸配列を含む。本明細書において提供されるアデノシンデアミナーゼは、一つ以上の突然変異(例えば、本明細書に提供される突然変異のいずれか)を含み得ることが理解されるべきである。本開示は、特定のパーセント同一性を有するデアミナーゼドメインが加えて本明細書に記載される突然変異のいずれかまたはその組合せを含むものを提供する。いくつかの実施形態において、アデノシンデアミナーゼは、参照配列または本明細書に提供されるアデノシンデアミナーゼのいずれかと比較して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、またはそれ以上の変異を有するアミノ酸配列を含む。いくつかの実施形態において、アデノシンデアミナーゼは、当該技術分野において既知であるかまたは本明細書に記載されたアミノ酸配列のいずれかと比較して、少なくとも5個、少なくとも10個、少なくとも15個、少なくとも20個、少なくとも25個、少なくとも30個、少なくとも35個、少なくとも40個、少なくとも45個、少なくとも50個、少なくとも60個、少なくとも70個、少なくとも80個、少なくとも90個、少なくとも100個、少なくとも110個、少なくとも120個、少なくとも130個、少なくとも140個、少なくとも150個、少なくとも160個、または少なくとも170個の同一の連続するアミノ酸残基を有するアミノ酸配列を含む。

【0344】

いくつかの実施形態において、TadAデアミナーゼは、全長E.coli TadAデアミナーゼである。例えば、特定の実施形態において、アデノシンデアミナーゼは、以下のアミノ酸配

10

20

30

40

50

列を含む:

MRRAFITGVFFLSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNR
PIGRHDPTAHAEIMALRQGGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGA
RDAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIAQKKAQSST
D。

【0345】

本出願において有用な追加のアデノシンデアミナーゼは当業者には明らかであり、本開示の範囲内であることが理解されよう。例えば、アデノシンデアミナーゼは、tRNAに作用するアデノシンデアミナーゼ(AD AT)のホモログであり得る。限定されるものではないが、例示的なAD AT相同体のアミノ酸配列は、以下を含む:

【0346】

Staphylococcus aureus TadA:

MGSHMTNDIYFMTLAIEEAKKAAQLGEVPIGAIITKDDEVIARAHNLRETLQQPTAHAEH
IAIERAAKVLGSRLEGCTLYVTLEPCVMCAGTIVMSRIPRVVYGADDPKGGCS GS LMN
LLQQS NFNHRAIVDKG VLKE AC S TLLTTFFKNLRANKKS TN

【0347】

Bacillus subtilis TadA:

MTQDELYMKEAIKEAKKAAEEKGEVPIGAVLVINGEIIARAHNLRETEQRSIAHAEML VID
EACKALGTWRLEGATLYVTLEPCPM CAGAVVLSRVEKVVFGAFDPKGGC S GTLMN LL
QEERFNHQAEVVSGVLEEECGGMLSAFFREL RKKKKAARKNLSE

【0348】

Salmonella typhimurium (S. typhimurium) TadA:

MPPAFITGVTSLSDELDEHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNR
PIGRHDPTAHAEIMALRQGGGLVLQNYRLDDTTLYVTLEPCVMCAGAMVHSRIGRVVFGA
RDAKTGAAGSLIDVLHHPGMNHRVEIEGVLRDECATLLSDFFRMRRQEIK ALKKADRA
EGAGPAV

【0349】

Shewanella putrefaciens (S. putrefaciens) TadA:

MDE YWMQVAMQM AEKAEAGE VPVGA VLVKDGQQIATGYNLS IS QHDPT AHAEI
LCLRSAGKKLENYRLLDATLYITLPCAMCAGAMVHSRIARVVYGARDEKTGAAGTVVN
LLQHAPFNHQVEVTSGVLAEACSAQLSRFFKRRRDEKKALKLAQRAQQGIE

【0350】

Haemophilus influenzae F3031 (H. influenzae) TadA:

MDAAKVRSEFDEKMMRYALELADKAEALGEIPVGAVLVDDARNIIGEGWNLSIVQSDPT
A HAEIIALRNGAKNIQNYRLNSTLYVTLEPCTMCAGAILHS RIKRLVFG ASDYKTGAI
GSRFHFFDDYKMNHTLEITSGVLAEECSQKLSTFFQKRREEKKIEKALLKSLSDK

【0351】

Caulobacter crescentus (C. crescentus) TadA:

MRTDESEDQDHRMMRLALDAARAAAEAGETPVGAVILDPSTGEVIATAGNGPIAAHDPT
AHAEIAAMRAAAAKLGNYRLTDLTLLVVTLEPCAMCAGAISHARIGRVVFGADD PKGGAV
VHGPKFFAQPTCHWRPEVTGGVLADESADLLRGFFRARRKAKI

【0352】

Geobacter sulfurreducens (G. sulfurreducens) TadA:

MSSLKKTPIRDDAYWMGKAIREAAKAAARDEVPIGAVIVRDGAVIGRGHNLREGSNDPSA
HAEMIAIRQAARRSANWRLTGATLYVTLEPCLMCMGAILARLERVVFGCYDPKGGGAAGS
LYDLSADPRLNHQVRLSPGVCQEECGTMLSDFFRDLRRRK KAKATPALF IDERKVPPEP.

【0353】

E. coli TadA (ecTadA):

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVNNRVIGEGWNRAIGLHDPTAHA
EIMALRQGGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL

10

20

30

40

50

MDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTD.

【0354】

ある態様において、アデノシンデアミナーゼは、TadA参照配列に対するD108X突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含み、Xは野生型アデノシンデアミナーゼにおける対応するアミノ酸以外の任意のアミノ酸を示す。いくつかの実施形態において、アデノシンデアミナーゼは、TadA参照配列に対するD108G、D108N、D108V、D108A、またはD108Yの、突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含む。しかしながら、さらなるデアミナーゼを同様に位置合わせさせて、本明細書中に提供されるように変異され得る相同的アミノ酸残基を同定し得ることが理解されるべきである。

10

【0355】

ある態様において、アデノシンデアミナーゼは、TadA参照配列に対するA106X突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含み、Xは野生型アデノシンデアミナーゼにおける対応するアミノ酸以外の任意のアミノ酸を示す。ある態様において、アデノシンデアミナーゼは、TadA参照配列に対するA106V突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含む。

【0356】

ある態様において、アデノシンデアミナーゼは、TadA参照配列に対するE155X突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含み、Xの存在は、野生型アデノシンデアミナーゼにおける対応するアミノ酸以外の任意のアミノ酸を示す。ある態様において、アデノシンデアミナーゼは、TadA参照配列に対するE155D、E155G、またはE155V突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含む。

20

【0357】

ある態様において、アデノシンデアミナーゼは、TadA参照配列に対するD147X突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含み、Xの存在は、野生型アデノシンデアミナーゼにおける対応するアミノ酸以外の任意のアミノ酸を示す。ある態様において、アデノシンデアミナーゼは、TadA参照配列に対するD147Y突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含む。

【0358】

本明細書に提供される突然変異(例えば、TadA参照配列のアミノ酸配列に基づくもの)のいずれも、黄色ブドウ球菌TadA (saTadA) などの他のアデノシンデアミナーゼ、またはその他のアデノシンデアミナーゼ(例えば、細菌のアデノシンデアミナーゼ)に導入することができることが理解されるべきである。TadA参照配列中の変異残基に相同的な配列をいかに同定するかは当業者に明らかであろう。従って、TadA基準配列に関連して同定された突然変異のいずれも、相同的アミノ酸残基を有する他のアデノシンデアミナーゼにおいて作製され得る。また、本明細書に提供される突然変異のいずれも、TadAまたは別のアデノシンデアミナーゼにおいて、個々にまたは任意の組合せで作製され得ることも理解されるべきである。例えば、アデノシンデアミナーゼは、TadA参照配列に対するD108N、A106V、E155V、および/またはD147Y突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含み得る。ある態様において、アデノシンデアミナーゼは、TadA参照配列に関連する以下の突然変異群(変異のグループは「;」で分けられる)、または別のアデノシンデアミナーゼにおける対応する突然変異を含む: D108NおよびA106V; D108NおよびE155V; D108NおよびD147Y; A106VおよびE155V; A106VおよびD147Y; E155VおよびD147Y; D108N、A106V、およびE155V; D108N、A106V、およびD147Y; D108N、E155V、およびD147Y; A106V、E155V、およびD147Y; ならびにD108N、A106V、E155V、およびD147Y; ただし、ここで提供される対応する突然変異の任意の組合せがアデノシンデアミナーゼ(例えばecTadA)において作製され得ることを理解されたい。

30

40

【0359】

50

ある態様において、アデノシンデアミナーゼは、H8X、T17X、L18X、W23X、L34X、W45X、R51X、A56X、E59X、E85X、M94X、I95X、V102X、F104X、A106X、R107X、D108X、K110X、M118X、N127X、A138X、F149X、M151X、R153X、Q154X、I156X、および/またはK157X突然変異のうちの1以上、または他のアデノシンデアミナーゼにおける1以上の対応する突然変異を含み、Xの存在は、アデノシン野生型デアミナーゼにおける対応するアミノ酸以外のいずれかのアミノ酸を示す。いくつかの実施形態において、アデノシンデアミナーゼは、TadA基準配列に対するH8Y、T17S、L18E、W23L、L34S、W45L、R51H、A56E、またはA56S、E59G、E85K、またはE85G、M94L、I95I、V102A、F104L、A106V、R107C、またはR107H、またはR107P、D108G、またはD108N、またはD108A、またはD108Y、K110I、M118K、N127S、A138V、F149Y、M151V、R153C、Q154L、I156D、および/またはK157R突然変異のうちの1つ以上、または他のアデノシンデアミナーゼにおける1つ以上の対応する突然変異を含む。

【0360】

いくつかの実施形態において、アデノシンデアミナーゼは、tadA基準配列に関連するH8X、T17X、L18X、W23X、L34X、W45X、R51X、A56X、E59X、E85X、M94X、I95X、V102X、F104X、A106X、R107X、D108X、K110X、M118X、N127X、A138X、F149X、M151X、R153X、Q154X、I156X、および/またはK157X突然変異のうちの1つ以上、または別のアデノシンデアミナーゼにおける1つ以上の対応する突然変異を含み、Xの存在は、野生型アデノシンデアミナーゼにおける対応するアミノ酸以外のいずれかのアミノ酸を示す。いくつかの実施形態において、アデノシンデアミナーゼは、tadA配列に関連するH8Y、T17S、L18E、W23L、L34S、W45L、R51H、A56E、またはA56S、E59G、E85K、またはE85G、M94L、I95I、V102A、F104L、A106V、R107C、またはR107H、またはR107P、D108G、またはD108N、またはD108A、またはD108Y、K110I、M118K、N127S、A138V、F149Y、M151V、R153C、Q154L、I156D、および/またはK157R突然変異のうちの1つ以上、または他のアデノシンデアミナーゼにおける1つ以上の対応する突然変異を含む。ある態様において、アデノシンデアミナーゼは、TadA参照配列に関連するH8X、D108X、および/またはN127X突然変異のうちの1つ以上、または別のアデノシンデアミナーゼにおける1つ以上の対応する突然変異を含み、Xは任意のアミノ酸の存在を示す。いくつかの実施形態において、アデノシンデアミナーゼは、TadA参照配列に関連するH8Y、D108N、および/またはN127S突然変異の1以上、または別のアデノシンデアミナーゼにおける1以上の対応する突然変異を含む。

【0361】

いくつかの実施形態において、アデノシンデアミナーゼは、TadA参照配列に対するH8X、D108X、および/またはN127X突然変異の1以上、または別のアデノシンデアミナーゼにおける1以上の対応する突然変異を含み、ここでXは任意のアミノ酸の存在を示す。ある態様において、アデノシンデアミナーゼは、TadA参照配列に対するH8Y、D108N、および/またはN127S突然変異のうちの1つ以上、または別のアデノシンデアミナーゼにおける1つ以上の対応する突然変異を含む。

【0362】

いくつかの実施形態において、アデノシンデアミナーゼは、TadA参照配列に対するH8X、D108X、および/またはN127X突然変異の1以上、または別のアデノシンデアミナーゼにおける1以上の対応する突然変異を含み、ここでXは任意のアミノ酸の存在を示す。ある態様において、アデノシンデアミナーゼは、TadA参照配列に対するH8Y、D108N、および/またはN127S突然変異のうちの1つ以上、または別のアデノシンデアミナーゼにおける1つ以上の対応する突然変異を含む。

【0363】

いくつかの実施形態において、アデノシンデアミナーゼは、TadA参照配列に対するH8X、D108X、および/またはN127X突然変異の1以上、または別のアデノシンデアミナーゼにおける1以上の対応する突然変異を含み、ここでXは任意のアミノ酸の存在を示す。あ

る態様において、アデノシンデアミナーゼは、*TadA*参照配列に対するH8Y、D108N、および/またはN127S突然変異のうちの1つ以上、または別のアデノシンデアミナーゼにおける1つ以上の対応する突然変異を含む。

【0364】

いくつかの実施形態において、アデノシンデアミナーゼは、H8X、D108X、N127X、D147X、R152X、およびQ154Xからなる群より選択される1つ、2つ、3つ、4つ、5つ、または6つの突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異（複数可）を含み、ここでXは野生型アデノシンデアミナーゼにおける対応するアミノ酸以外のいずれかのアミノ酸の存在を示す。いくつかの実施形態において、アデノシンデアミナーゼは、H8X、M61X、M70X、D108X、N127X、Q154X、E155X、およびQ163Xからなる群より選択される1つ、2つ、3つ、4つ、5つ、6つ、7つ、または8つの突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異（複数可）を含み、ここでXは野生型アデノシンデアミナーゼにおける対応するアミノ酸以外のいずれかのアミノ酸の存在を示す。いくつかの実施形態において、アデノシンデアミナーゼは、*TadA*参照配列に対してH8X、D108X、N127X、E155X、およびT166Xからなる群より選択される1つ、2つ、3つ、4つ、または5つの突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異（複数可）を含み、Xは、参照または野生型アデノシンデアミナーゼにおける対応するアミノ酸以外のいずれかのアミノ酸の存在を示す。いくつかの実施形態において、アデノシンデアミナーゼは、*Tad*参照配列に対するH8X、A106X、D108X突然変異からなる群から選択される1つ、2つ、3つ、4つ、5つ、または6つの突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含み、Xは、参照または野生型アデノシンデアミナーゼにおける対応するアミノ酸以外の任意のアミノ酸の存在を示す。いくつかの実施形態において、アデノシンデアミナーゼは、*Tad*参照配列に対するH8X、R126X、L68X、D108X、N127X、D147X、およびE155Xからなる群より選択される1つ、2つ、3つ、4つ、5つ、6つ、7つ、または8つの突然変異、または他のアデノシンデアミナーゼにおける対応する突然変異（複数可）を含み、ここでXは野生型アデノシンデアミナーゼにおける対応するアミノ酸以外のいずれかのアミノ酸の存在を示す。ある態様において、アデノシンデアミナーゼは、*TadA*参照配列に対してH8X、D108X、A109X、N127X、およびE155Xからなる群より選択される1つ、2つ、3つ、4つ、または5つの突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異（複数可）を含み、Xは野生型アデノシンデアミナーゼにおける対応するアミノ酸以外のいずれかのアミノ酸の存在を示す。

【0365】

いくつかの実施形態において、アデノシンデアミナーゼは、*TadA*参照配列に対してH8Y、D108N、N127S、D147Y、R152C、およびQ154Hからなる群より選択される1つ、2つ、3つ、4つ、5つ、または6つの突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含む。いくつかの実施形態において、アデノシンデアミナーゼは、*TadA*参照配列に対してH8Y、M61I、M70V、D108N、N127S、Q154R、E155GおよびQ163Hからなる群より選択される1つ、2つ、3つ、4つ、5つ、6つ、7つまたは8つの突然変異、または他のアデノシンデアミナーゼにおける対応する突然変異を含む。いくつかの実施形態において、アデノシンデアミナーゼは、*TadA*参照配列に対してH8Y、D108N、N127S、E155V、およびT166Pからなる群より選択される1つ、2つ、3つ、4つ、または5つの突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異（複数可）を含む。いくつかの実施形態において、アデノシンデアミナーゼは、*TadA*参照配列に対してH8Y、A106T、D108N、N127S、E155D、およびK161Qからなる群から選択される1つ、2つ、3つ、4つ、5つ、または6つの突然変異、または他のアデノシンデアミナーゼにおける対応する突然変異を含む。いくつかの実施形態において、アデノシンデアミナーゼは、*TadA*参照配列に対してH8Y、R126W、L68Q、D108N、N127S、D147Y、およびE155Vからなる群より選択される1つ、2つ、3つ、4つ、5つ、6つ、7つ、または8つの突然変異、または他のアデノシンデアミナーゼにおける対応する突然変異を含む

。いくつかの実施形態において、アデノシンデアミナーゼは、TadA参照配列に対してH8Y、D108N、A109T、N127S、およびE155Gからなる群より選択される1つ、2つ、3つ、4つ、または5つの突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異（複数可）を含む。

【0366】

ある態様において、アデノシンデアミナーゼは、別のアデノシンデアミナーゼにおける1以上のまたは1以上の対応する突然変異を含む。ある態様において、アデノシンデアミナーゼは、TadA参照配列に関連するD108N、D108G、またはD108V突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含む。ある態様において、アデノシンデアミナーゼは、TadA参照配列に対してA106VおよびD108N突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含む。ある態様において、アデノシンデアミナーゼは、TadA参照配列に関連するR107CおよびD108N突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含む。ある態様において、アデノシンデアミナーゼは、TadA参照配列に対してH8Y、D108N、N127S、D147Y、およびQ154H突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含む。いくつかの実施形態において、アデノシンデアミナーゼは、TadA参照配列に対してH8Y、R24W、D108N、N127S、D147Y、およびE155V突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含む。ある態様において、アデノシンデアミナーゼは、TadA参照配列に対するD108N、D147Y、およびE155V突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含む。ある態様において、アデノシンデアミナーゼは、TadA参照配列に関連するH8Y、D108N、およびN127S突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含む。ある態様において、アデノシンデアミナーゼは、TadA参照配列に対するA106V、D108N、D147YおよびE155V突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含む。

【0367】

いくつかの実施形態において、アデノシンデアミナーゼは、tadA参照配列に対してS2X、H8X、I49X、L84X、H123X、N127X、I156Xおよび/またはK160X突然変異の1つ以上、または別のアデノシンデアミナーゼにおける1つ以上の対応する突然変異を含み、Xの存在は、野生型アデノシンデアミナーゼにおける対応するアミノ酸以外のいずれかのアミノ酸を示す。いくつかの実施形態において、アデノシンデアミナーゼは、TadA参照配列に対してS2A、H8Y、I49F、L84F、H123Y、N127S、I156Fおよび/またはK160S突然変異のうちの1つ以上、または別のアデノシンデアミナーゼにおける1つ以上の対応する突然変異を含む。

【0368】

ある態様において、アデノシンデアミナーゼは、L84X突然変異アデノシンデアミナーゼを含み、Xは、野生型アデノシンデアミナーゼにおける対応するアミノ酸以外の任意のアミノ酸を示す。ある態様において、アデノシンデアミナーゼは、TadA参照配列に対するL84F突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含む。

【0369】

ある態様において、アデノシンデアミナーゼは、TadA参照配列に関連するH123X突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含み、Xは、野生型アデノシンデアミナーゼにおける対応するアミノ酸以外の任意のアミノ酸を示す。ある態様において、アデノシンデアミナーゼは、TadA参照配列に対するH123Y突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含む。

【0370】

ある態様において、アデノシンデアミナーゼは、TadA参照配列に関連するI157X突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含み、Xは、野生型アデノシンデアミナーゼにおける対応するアミノ酸以外の任意のアミノ酸を示す。ある態様において、アデノシンデアミナーゼは、TadA参照配列に関連するI157F突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含む。

【0371】

ある態様において、アデノシンデアミナーゼは、*TadA*参照配列に対して、L84X、A106X、D108X、H123X、D147X、E155X、およびI156Xからなる群より選択される1つ、2つ、3つ、4つ、5つ、6つ、または7つの突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異（複数可）を含み、ここでXは、野生型アデノシンデアミナーゼにおける対応するアミノ酸以外のいずれかのアミノ酸の存在を示す。ある態様において、アデノシンデアミナーゼは、*tadA*参照配列に対してS2X、I49X、A106X、D108X、D147X、およびE155Xからなる群より選択される1つ、2つ、3つ、4つ、5つ、または6つの突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異（複数可）を含み、ここでXは野生型アデノシンデアミナーゼにおける対応するアミノ酸以外のいずれかのアミノ酸の存在を示す。ある態様において、アデノシンデアミナーゼは、*TadA*参照配列に対してH8X、A106X、D108X、N127X、およびK160Xからなる群より選択される1、2、3、4、または5個の突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異（複数可）を含み、ここでXは野生型アデノシンデアミナーゼにおける対応するアミノ酸以外のいずれかのアミノ酸の存在を示す。

10

【0372】

ある態様において、アデノシンデアミナーゼは、*TadA*参照配列に対して、L84F、A106V、D108N、H123Y、D147Y、E155V、およびI156Fからなる群より選択される1つ、2つ、3つ、4つ、5つ、6つ、または7つの突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含む。ある態様において、アデノシンデアミナーゼは、*TadA*参照配列に対して、S2A、I49F、A106V、D108N、D147Y、およびE155Vからなる群より選択される1つ、2つ、3つ、4つ、5つ、または6つの突然変異を含む。

20

【0373】

いくつかの実施形態において、アデノシンデアミナーゼは、*TadA*参照配列に対してH8Y、A106T、D108N、N127S、およびK160Sからなる群より選択される1つ、2つ、3つ、4つ、または5つの突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異もしくは突然変異を含む。

【0374】

いくつかの実施形態において、アデノシンデアミナーゼは、*TadA*参照配列に対してE25X、R26X、R107X、A142X、および/またはA143X突然変異のうちの1以上、または別のアデノシンデアミナーゼにおける1以上の対応する突然変異を含み、Xの存在は、野生型アデノシンデアミナーゼにおける対応するアミノ酸以外のいずれかのアミノ酸を示す。ある態様において、アデノシンデアミナーゼは、*TadA*基準配列に関連するE25M、E25D、E25A、E25R、E25V、E25S、E25Y、R26G、R26N、R26Q、R26C、R26L、R26K、R107P、R07K、R107A、R107N、R107W、R107H、R107S、A142N、A142D、A142G、A143D、A143G、A143E、A143L、A143W、A143M、A143S、A143Q および/またはA143R突然変異のうちの1以上、または他のアデノシンデアミナーゼにおける1以上の対応する突然変異を含む。いくつかの実施形態において、アデノシンデアミナーゼは、*TadA*参照配列に対応する本明細書に記載の突然変異のうちの1つ以上、または別のアデノシンデアミナーゼにおける対応する突然変異のうちの1つ以上を含む。

30

40

【0375】

ある態様において、アデノシンデアミナーゼは、*TadA*参照配列に関連するE25X突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含み、Xは、野生型アデノシンデアミナーゼにおける対応するアミノ酸以外の任意のアミノ酸を示す。ある態様において、アデノシンデアミナーゼは、*TadA*参照配列に対してE25M、E25D、E25A、E25R、E25V、E25S、またはE25Y突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含む。

【0376】

ある態様において、アデノシンデアミナーゼは、*TadA*参照配列に関連するR26X突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含み、Xは、野生型

50

アデノシンデアミナーゼにおける対応するアミノ酸以外の任意のアミノ酸を示す。いくつかの実施形態において、アデノシンデアミナーゼは、*TadA*参照配列に対してR26G、R26N、R26Q、R26C、R26L、またはR26K突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含む。

【0377】

ある態様において、アデノシンデアミナーゼは、*TadA*参照配列に対するR107X突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含み、Xは、野生型アデノシンデアミナーゼにおける対応するアミノ酸以外の任意のアミノ酸を示す。ある態様において、アデノシンデアミナーゼは、*TadA*参照配列に対してR107P、R07K、R107A、R107N、R107W、R107H、またはR107S突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含む。

10

【0378】

ある態様において、アデノシンデアミナーゼは、*TadA*参照配列に対するA142X突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含み、Xは、野生型アデノシンデアミナーゼにおける対応するアミノ酸以外の任意のアミノ酸を示す。ある態様において、アデノシンデアミナーゼは、*TadA*参照配列に対するA142N、A142D、A142G突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含む。

【0379】

ある態様において、アデノシンデアミナーゼは、*TadA*参照配列に対してA143X突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含み、Xは、野生型アデノシンデアミナーゼにおける対応するアミノ酸以外の任意のアミノ酸を示す。ある態様において、アデノシンデアミナーゼは、*TadA*参照配列に対するA143D、A143G、A143E、A143L、A143W、A143M、A143S、A143Qおよび/またはA143Rの突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含む。

20

【0380】

いくつかの実施形態において、アデノシンデアミナーゼは、*TadA*参照配列に対するH36X、N37X、P48X、I49X、R51X、M70X、N72X、D77X、E134X、S146X、Q154X、K157X、および/またはK161X突然変異の1つ以上、または別のアデノシンデアミナーゼにおける1つ以上の対応する突然変異を含み、ここでXの存在は、野生型アデノシンデアミナーゼにおける対応するアミノ酸以外のいずれかのアミノ酸を示す。いくつかの実施形態において、アデノシンデアミナーゼは、*TadA*参照配列に関連するH36L、N37T、N37S、P48T、P48L、I49V、R51H、R51L、M70L、N72S、D77G、E134G、S146R、S146C、Q154H、K157N、および/またはK161T突然変異の1以上、または別のアデノシンデアミナーゼにおける1以上の対応する突然変異を含む。

30

【0381】

いくつかの実施形態において、アデノシンデアミナーゼは、*TadA*参照配列に対するH36X突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含み、ここでXは、野生型アデノシンデアミナーゼにおける対応するアミノ酸以外の任意のアミノ酸を示す。いくつかの実施形態において、アデノシンデアミナーゼは、*TadA*参照配列に対するH36L突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含む。

40

【0382】

ある態様において、アデノシンデアミナーゼは、*TadA*参照配列に対するN37X突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含み、Xは野生型アデノシンデアミナーゼにおける対応するアミノ酸以外の任意のアミノ酸を示す。ある態様において、アデノシンデアミナーゼは、*TadA*参照配列に対するN37TまたはN37S突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含む。

【0383】

ある態様において、アデノシンデアミナーゼは、*TadA*参照配列に対するP48X突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含み、ここでXは、野生型アデノシンデアミナーゼにおける対応するアミノ酸以外の任意のアミノ酸を示す。あ

50

る態様において、アデノシンデアミナーゼは、TadA参照配列に対するP48TまたはP48L突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含む。

【0384】

ある態様において、アデノシンデアミナーゼは、TadA参照配列に対するR51X突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含み、Xは野生型アデノシンデアミナーゼにおける対応するアミノ酸以外の任意のアミノ酸を示す。ある態様において、アデノシンデアミナーゼは、TadA参照配列に対するR51HまたはR51L突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含む。

【0385】

ある態様において、アデノシンデアミナーゼは、TadA参照配列に対するS146X突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含み、Xは野生型アデノシンデアミナーゼにおける対応するアミノ酸以外の任意のアミノ酸を示す。ある態様において、アデノシンデアミナーゼは、TadA参照配列に対するS146RまたはS146C突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含む。

【0386】

ある態様において、アデノシンデアミナーゼは、TadA参照配列に対するK157X突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含み、Xは野生型アデノシンデアミナーゼにおける対応するアミノ酸以外の任意のアミノ酸を示す。ある態様において、アデノシンデアミナーゼは、TadA参照配列に対するK157N突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含む。

【0387】

ある態様において、アデノシンデアミナーゼは、TadA参照配列に対するP48X突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含み、ここでXは、野生型アデノシンデアミナーゼにおける対応するアミノ酸以外の任意のアミノ酸を示す。ある態様において、アデノシンデアミナーゼは、TadA参照配列に対するP48S、P48T、またはP48A突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含む。

【0388】

ある態様において、アデノシンデアミナーゼは、TadA参照配列に対するA142X突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含み、Xは野生型アデノシンデアミナーゼにおける対応するアミノ酸以外の任意のアミノ酸を示す。ある態様において、アデノシンデアミナーゼは、TadA参照配列に対するA142N突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含む。

【0389】

ある態様において、アデノシンデアミナーゼは、TadA参照配列に対するW23X突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含み、ここでXは、野生型アデノシンデアミナーゼにおける対応するアミノ酸以外の任意のアミノ酸を示す。ある態様において、アデノシンデアミナーゼは、TadA参照配列に対するW23RまたはW23L突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含む。

【0390】

ある態様において、アデノシンデアミナーゼは、TadA参照配列に対するR152X突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含み、ここでXは、野生型アデノシンデアミナーゼにおける対応するアミノ酸以外の任意のアミノ酸を示す。ある態様において、アデノシンデアミナーゼは、TadA参照配列に対するR152PまたはR52H突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含む。

【0391】

一実施形態では、アデノシンデアミナーゼは、突然変異H36L、R51L、L84F、A106V、D108N、H123Y、S146C、D147Y、E155V、I156F、およびK157Nを含むことができる。一部の実施形態では、アデノシンデアミナーゼは、tadA基準配列に関連する突然変異の以下の組合せを含み、ここで、組合せの各突然変異は「_」によって分離され、突然変

10

20

30

40

50

異の各組合せは括弧内にある：

(A106V_D108N), (R107C_D108N), (H8Y_D108N_N127S_D 147Y_Q154H), (H8Y_R24W_D108N_N127S_D147Y_E155V), (D108N_D147Y_E155V), (H8Y_D108N_N127S), (H8Y_D108N_N127S_D147Y_Q154H), (A106V_D108N_D147Y_E155V) (D108Q_D147Y_E155V) (D108M_D147Y_E155V), (D108L_D147Y_E155V), (D108K_D147Y_E155V), (D108I_D147Y_E155V), (D108F_D147Y_E155V), (A106V_D108N_D147Y), (A106V_D108M_D147Y_E155V), (E59A_A106V_D108N_D147Y_E155V), (E59A cat dead_A106V_D108N_D147Y_E155V), (L84F_A106V_D108N_H123Y_D147Y_E155V_I156Y), (L84F_A106V_D108N_H123Y_D147Y_E155V_I156F), (D103A_D104N), (G22P_D103A_D104N), (G22P_D103A_D104N_S138 A), (D103 A_D104N_S138 A), (R26G_L84F_A106V_R107H_D108N_H123Y_A142N_A143D_D147Y_E155V_I156F), (E25G_R26G_L84F_A106V_R107H_D108N_H123Y_A142N_A143D_D147Y_E155V_I156F), (E25D_R26G_L84F_A106V_R107K_D108N_H123Y_A142N_A143G_D147Y_E155V_I156F), (R26Q_L84F_A106V_D108N_H123Y_A142N_D147Y_E155V_I156F), (E25M_R26G_L84F_A106V_R107P_D108N_H123Y_A142N_A143D_D147Y_E155V_I156F), (R26C_L84F_A106V_R107H_D108N_H123Y_A142N_D147Y_E155V_I156F), (L84F_A106V_D108N_H123Y_A142N_A143L_D147Y_E155V_I156F), (R26G_L84F_A106V_D108N_H123Y_A142N_D147Y_E155V_I156F), (E25A_R26G_L84F_A106V_R107N_D108N_H123Y_A142N_A143E_D147Y_E155V_I156F), (R26G_L84F_A106V_R107H_D108N_H123Y_A142N_A143D_D147Y_E155V_I156F), (A106V_D108N_A142N_D147Y_E155V), (R26G_A106V_D108N_A142N_D147Y_E155V), (E25D_R26G_A106V_R107K_D108N_A142N_A143G_D147Y_E155V), (R26G_A106V_D108N_R107H_A142N_A143D_D147Y_E155V), (E25D_R26G_A106V_D108N_A142N_D147Y_E155V), (A106V_R107K_D108N_A142N_D147Y_E155V), (A106V_D108N_A142N_A143G_D147Y_E155V), (A106V_D108N_A142N_A143L_D147Y_E155V), (H36L_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N), (N37T_P48T_M70L_L84F_A106V_D108N_H123Y_D147Y_I49V_E155V_I156F), (N37S_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F_K161T), (H36L_L84F_A106V_D108N_H123Y_D147Y_Q154H_E155V_I156F), (N72S_L84F_A106V_D108N_H123Y_S146R_D147Y_E155V_I156F), (H36L_P48L_L84F_A106V_D108N_H123Y_E134G_D147Y_E155V_I156F_K157N), (H36L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F), (L84F_A106V_D108N_H123Y_S146R_D147Y_E155V_I156F_K161T), (N37S_R51H_D77G_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F),

10

20

30

40

50

(R51L_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F_K157N),
(D24G_Q71R_L84F_H96L_A106V_D108N_H123Y_D147Y_E155V_I156F_K160E
) ,
(H36L_G67V_L84F_A106V_D108N_H123Y_S146T_D147Y_E155V_I156F),
(Q71L_L84F_A106V_D108N_H123Y_L137M_A143E_D147Y_E155V_I156F),
(E25G_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F_Q159L),
(L84F_A91T_F104I_A106V_D108N_H123Y_D147Y_E155V_I156F),
(N72D_L84F_A106V_D108N_H123Y_G125A_D147Y_E155V_I156F),
(P48S_L84F_S97C_A106V_D108N_H123Y_D147Y_E155V_I156F),
(W23G_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F),
(D24G_P48L_Q71R_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F_Q159L
) ,
(L84F_A106V_D108N_H123Y_A142N_D147Y_E155V_I156F),
(H36L_R51L_L84F_A106V_D108N_H123Y_A142N_S146C_D147Y_E155V_I156
F
_K157N), (N37S_L84F_A106V_D108N_H123Y_A142N_D147Y_E155V_I156F_K
161T),
(L84F_A106V_D108N_D147Y_E155V_I156F),
(R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N_K16
1T),
(L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K161T),
(L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N_K160E_K1
61T),
(L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N_K160E), (R
74Q
L84F_A106V_D108N_H123Y_D147Y_E155V_I156F),
(R74A_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F),
(L84F_A106V_D108N_H123Y_D147Y_E155V_I156F),
(R74Q_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F),
(L84F_R98Q_A106V_D108N_H123Y_D147Y_E155V_I156F),
(L84F_A106V_D108N_H123Y_R129Q_D147Y_E155V_I156F),
(P48S_L84F_A106V_D108N_H123Y_A142N_D147Y_E155V_I156F), (P48S_A14
2N),
(P48T_I49V_L84F_A106V_D108N_H123Y_A142N_D147Y_E155V_I156F_L157
N),
(P48T_I49V_A142N),
(H36L_P48S_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F
_K157N),
(H36L_P48S_R51L_L84F_A106V_D108N_H123Y_S146C_A142N_D147Y_E155V
_I156F (H36L_P48T_I49V_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_
E155V_I156F_K157N),
(H36L_P48T_I49V_R51L_L84F_A106V_D108N_H123Y_A142N_S146C_D147Y_
E155V_I156F_K157N),
(H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F
_K157N),
(H36L_P48A_R51L_L84F_A106V_D108N_H123Y_A142N_S146C_D147Y_E155
V_I156F_K157N),
(H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_A142N_D147Y_E155
V_I156F_K157N),
(W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V

10

20

30

40

50

_I156F_K157N),
(W23R_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V
_I156F_K157N),
(W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146R_D147Y_E155V
_I156F_K161T),
(H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152H_E155
V_I156F_K157N),
(H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152P_E155V
_I156F_K157N),
(W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152P
_E155V_I156F_K157N),
(W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_A142A_S146C_D147Y
_E155V
_I156F_K157N),
(W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_A142A_S146C_D147Y
_R152P_E155V_I156F_K157N),
(W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146R_D147Y_E155V
_I156F_K161T),
(W23R_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152P
_E155V_I156F_K157N),
(H36L_P48A_R51L_L84F_A106V_D108N_H123Y_A142N_S146C_D147Y_R152
P_E155V
_I156F_K157N).

【0392】

特定の実施形態では、本明細書で提供される融合タンパク質は、融合タンパク質の塩基編集活性を改善する1つ以上の特徴を含む。例えば、本明細書に提供される融合タンパク質は、ヌクレアーゼ活性が低下したCas9ドメインを含み得る。いくつかの実施形態において、本明細書で提供される融合タンパク質は、ヌクレアーゼ活性を有さないCas9ドメイン(dCas9)、またはCas9ニッカーゼ(nCas9)と呼ばれる、二本鎖DNA分子の1鎖を切断するCas9ドメインを有し得る。

【0393】

ある態様において、アデノシンデアミナーゼは、TadA参照配列または野生型配列に対するD108X突然変異、または別のアデノシンデアミナーゼ(例えばecTadA)における対応する突然変異を含み、Xは野生型アデノシンデアミナーゼにおける対応するアミノ酸以外の任意のアミノ酸を示す。ある態様において、アデノシンデアミナーゼは、D108G、D108N、D108V、D108A、またはD108Y突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含む。

【0394】

ある態様において、アデノシンデアミナーゼは、TadA参照配列または野生型配列に関連するA106X、E155X、またはD147X、突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含み、Xは野生型アデノシンデアミナーゼにおける対応するアミノ酸以外の任意のアミノ酸を示す。ある態様において、アデノシンデアミナーゼは、E155D、E155G、またはE155V突然変異を含む。ある態様において、アデノシンデアミナーゼは、D147Yを含む。

【0395】

本明細書に提供される突然変異(例えばTadA基準アミノ酸配列に基づくもの)のいずれも、*S. aureus* TadA (saTadA)などの他のアデノシンデアミナーゼ、またはその他のアデノシンデアミナーゼ(例えば細菌アデノシンデアミナーゼ)に導入され得ることを理解されたい。TadA基準配列に基づいて同定された突然変異のいずれも、相同的アミノ酸残基を有する他のアデノシンデアミナーゼにおいて作製することができる。また、本明細書中に提

10

20

30

40

50

供される突然変異のいずれも、*TadA*または別のアデノシンデアミナーゼに関連して、個々にまたは任意の組合せで作製することができることを理解されたい。

【0396】

例えば、アデノシンデアミナーゼは、*TadA*参照配列に対するD108N、A106V、E155V、および/またはD147Y突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含み得る。ある態様において、アデノシンデアミナーゼは、*TadA*参照配列に関連する以下の突然変異群(突然変異の群は「;」によって分離されている)、または別のアデノシンデアミナーゼにおける対応する突然変異を含む：D108NおよびA106V；D108NおよびE155V；D108NおよびD147Y；A106VおよびE155V；A106VおよびD147Y；E155VおよびD147Y；D108N, A106V, およびE55V；D108N, A106V, およびD147Y；D108N, E55V, およびD147Y；A106V, E55V, およびD147Y；およびD108N, A106V, E55V, およびD147Y。ただし、ここで提供される対応する突然変異の任意の組合せを、アデノシンデアミナーゼ(例えば*ecTadA*)において作製することができることを理解されたい。

10

【0397】

いくつかの実施形態において、アデノシンデアミナーゼは、*TadA*参照配列に対するH8X、T17X、L18X、W23X、L34X、W45X、R51X、A56X、E59X、E85X、M94X、I95X、V102X、F104X、A106X、R107X、D108X、K110X、M118X、N127X、A138X、F149X、M151X、R153X、Q154X、I156X、および/またはK157X突然変異のうちの1以上、または他のアデノシンデアミナーゼにおける1以上の対応する突然変異を含み、ここで、Xの存在は、野生型アデノシンデアミナーゼにおける対応するアミノ酸以外のいずれかのアミノ酸を示す。いくつかの実施形態において、アデノシンデアミナーゼは、*TadA*参照配列に関連するH8Y、T17S、L18E、W23L、L34S、W45L、R51H、A56E、またはA56S、E59G、E85K、またはE85G、M94L、I95L、V102A、F104L、A106V、R107C、またはR107H、またはR107P、D108G、またはD108N、またはD108A、D108Y、K110I、M118K、N127S、A138V、F149Y、M151V、R153C、Q154L、I156D、および/またはK157R突然変異のうちの1つ以上、または他のアデノシンデアミナーゼにおける1つ以上の対応する突然変異を含む。いくつかの実施形態において、アデノシンデアミナーゼは、*TadA*参照配列に対するH8X、D108X、および/またはN127X突然変異のうちの1つ以上、または別のアデノシンデアミナーゼにおける1つ以上の対応する突然変異を含み、Xは任意のアミノ酸の存在を示す。ある態様において、アデノシンデアミナーゼは、*TadA*参照配列に対するH8Y、D108N、および/またはN127S突然変異のうちの1つ以上、または別のアデノシンデアミナーゼにおける1つ以上の対応する突然変異を含む。

20

30

【0398】

いくつかの実施形態において、アデノシンデアミナーゼは、*TadA*参照配列に対するH8X、R26X、M61X、L68X、M70X、A106X、D108X、A109X、N127X、D147X、R152X、Q154X、E155X、K161X、Q163X、および/またはT166X突然変異の1以上、または別のアデノシンデアミナーゼにおける1以上の対応する突然変異を含み、ここでXは、野生型アデノシンデアミナーゼにおける対応するアミノ酸以外のいずれかのアミノ酸の存在を示す。いくつかの実施形態において、アデノシンデアミナーゼは、*TadA*参照配列に対するH8Y、R26W、M61I、L68Q、M70V、A106T、D108N、A109T、N127S、D147Y、R152C、Q154HもしくはQ154R、E155GもしくはE155VもしくはE155D、K161Q、Q163H、および/もしくはT166P突然変異の1つ以上、または他のアデノシンデアミナーゼにおける1つ以上の対応する突然変異を含む。

40

【0399】

いくつかの実施形態において、アデノシンデアミナーゼは、*TadA*参照配列に対するH8X、D108X、N127X、D147X、R152X、およびQ154Xからなる群より選択される1つ、2つ、3つ、4つ、5つ、または6つの突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異(複数可)を含み、ここでXは野生型アデノシンデアミナーゼにお

50

ける対応するアミノ酸以外のいずれかのアミノ酸の存在を示す。いくつかの実施形態において、アデノシンデアミナーゼは、TadA参照配列に対するH8X、M61X、M70X、D108X、N127X、Q154X、E155X、およびQ163Xからなる群より選択される1つ、2つ、3つ、4つ、5つ、6つ、7つ、または8つの突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異（複数可）を含み、ここでXは野生型アデノシンデアミナーゼにおける対応するアミノ酸以外のいずれかのアミノ酸の存在を示す。ある態様において、アデノシンデアミナーゼは、TadA参照配列に対してH8X、D108X、N127X、E155X、およびT166Xからなる群より選択される1つ、2つ、3つ、4つ、または5つの突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異（複数可）を含み、ここでXは野生型アデノシンデアミナーゼにおける対応するアミノ酸以外のいずれかのアミノ酸の存在を示す。

10

【0400】

ある態様において、アデノシンデアミナーゼは、H8X、A106X、D108X、別のアデノシンデアミナーゼにおける突然変異（複数可）からなる群から選択される1、2、3、4、5、または6個の突然変異を含み、ここでXは野生型アデノシンデアミナーゼにおける対応するアミノ酸以外の任意のアミノ酸の存在を示す。ある態様において、アデノシンデアミナーゼは、H8X、R126X、L68X、D108X、N127X、D147X、およびE155Xからなる群より選択される1つ、2つ、3つ、4つ、5つ、6つ、7つ、または8つの突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異（複数可）を含み、Xは野生型アデノシンデアミナーゼにおける対応するアミノ酸以外のいずれかのアミノ酸の存在を示す。ある態様において、アデノシンデアミナーゼは、TadA参照配列に対してH8X、D108X、A109X、N127X、およびE155Xからなる群より選択される1つ、2つ、3つ、4つ、または5つの突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異（複数可）を含み、Xは野生型アデノシンデアミナーゼにおける対応するアミノ酸以外のいずれかのアミノ酸の存在を示す。

20

【0401】

いくつかの実施形態において、アデノシンデアミナーゼは、TadA参照配列に対してH8Y、D108N、N127S、D147Y、R152C、およびQ154Hからなる群より選択される1つ、2つ、3つ、4つ、5つ、または6つの突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異を含む。いくつかの実施形態において、アデノシンデアミナーゼは、TadA参照配列に対してH8Y、M61I、M70V、D108N、N127S、Q154R、E155GおよびQ163Hからなる群より選択される1つ、2つ、3つ、4つ、5つ、6つ、7つまたは8つの突然変異、または他のアデノシンデアミナーゼにおける対応する突然変異を含む。いくつかの実施形態において、アデノシンデアミナーゼは、TadA参照配列に対してH8Y、D108N、N127S、E155V、およびT166Pからなる群より選択される1つ、2つ、3つ、4つ、または5つの突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異（複数可）を含む。いくつかの実施形態において、アデノシンデアミナーゼは、TadA参照配列に対してH8Y、A106T、D108N、N127S、E155D、およびK161Qからなる群から選択される1つ、2つ、3つ、4つ、5つ、または6つの突然変異、または他のアデノシンデアミナーゼにおける対応する突然変異を含む。いくつかの実施形態において、アデノシンデアミナーゼは、TadA参照配列に対してH8Y、R126W、L68Q、D108N、N127S、D147Y、およびE155Vからなる群より選択される1つ、2つ、3つ、4つ、5つ、6つ、7つ、または8つの突然変異、または他のアデノシンデアミナーゼにおける対応する突然変異を含む。いくつかの実施形態において、アデノシンデアミナーゼは、TadA参照配列に対してH8Y、D108N、A109T、N127S、およびE155Gからなる群より選択される1つ、2つ、3つ、4つ、または5つの突然変異、または別のアデノシンデアミナーゼにおける対応する突然変異（複数可）を含む。

30

40

【0402】

本明細書に提供される突然変異のいずれか、および任意のさらなる突然変異(例えばecTadAアミノ酸配列に基づくもの)は、任意の他のアデノシンデアミナーゼに導入することが

50

できる。本明細書に提供される突然変異のいずれも、TadA参照配列または別のアデノシンデアミナーゼに対して、個々に、または任意の組合せで作製することができる。

【0403】

AからGへの核酸塩基編集タンパク質の詳細は、国際PCT出願第PCT/2017/045381号(国際公開第2018/027078号)およびGaudelli, N.M., et al., “Programmable base editing of A・T to G・C in genomic DNA without DNA cleavage” Nature, 551, 464-471 (2017)に記載されており、その内容全体が参照により本明細書に組み込まれる。

【0404】

[シチジンデアミナーゼ]

1つの実施形態において、本発明の融合タンパク質は、シチジンデアミナーゼを含む。ある態様において、本明細書において提供されるシチジンデアミナーゼは、シトシンまたは5-メチルシトシンを脱アミノ化してウラシルまたはチミンにすることができる。いくつかの実施形態において、本明細書に提供されるシトシンデアミナーゼは、DNA中のシトシンを脱アミノ化することができる。シチジンデアミナーゼは、任意の適切な生物に由来することができる。いくつかの実施形態において、シチジンデアミナーゼは、天然に存在するシチジンデアミナーゼが本明細書に提供される突然変異のいずれかに対応する1つ以上の突然変異を含むところのものである。当業者は、例えば、配列アラインメントおよび相同的残基の決定によって、任意の相同的タンパク質中の対応する残基を同定することができる。従って、当業者は、本明細書に記載された突然変異のいずれかに対応する突然変異を、任意の天然に存在するシチジンデアミナーゼにおいて生じさせることができる。ある態様において、シチジンデアミナーゼは、原核生物由来である。ある態様において、シチジンデアミナーゼは、細菌由来である。ある態様において、シチジンデアミナーゼは、哺乳動物(例えば、ヒト)由来である。

【0405】

いくつかの実施形態において、シチジンデアミナーゼは、本明細書に記載されるシチジンデアミナーゼアミノ酸配列のいずれかに対して少なくとも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%の同一性であるアミノ酸配列を含む。本明細書において提供されるシチジンデアミナーゼは、一つ以上の突然変異(例えば、本明細書に提供される突然変異のいずれか)を含み得ることを理解されたい。本開示は、特定のパーセント同一性を有する任意のデアミナーゼドメインに、本明細書に記載される突然変異またはその組合せのいずれかが加えられたものを提供する。いくつかの実施形態において、シチジンデアミナーゼは、参照配列または本明細書に提供されるいずれかのシチジンデアミナーゼと比較して、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個またはそれ以上の変異を有するアミノ酸配列を含む。いくつかの実施形態において、シチジンデアミナーゼは、当技術分野において既知であるかまたは本明細書に記載されたアミノ酸配列のいずれかと比較して、少なくとも5、少なくとも10、少なくとも15、少なくとも20、少なくとも25、少なくとも30、少なくとも35、少なくとも40、少なくとも45、少なくとも50、少なくとも60、少なくとも70、少なくとも80、少なくとも90、少なくとも100、少なくとも110、少なくとも120、少なくとも130、少なくとも140、少なくとも150、少なくとも160、または少なくとも170個の同一の連続するアミノ酸残基を有するアミノ酸配列を含む。

【0406】

[追加的ドメイン]

本明細書に記載される塩基エディターは、ポリヌクレオチドの核酸塩基の編集、修飾、または改変を促進させるのを助ける任意のドメインを含むことができる。ある態様において、塩基エディターは、ポリヌクレオチドプログラム可能なヌクレオチド結合ドメイン(例えばCas9)、核酸塩基編集ドメイン(例えばデアミナーゼドメイン)、および一つ以上のさ

らなるドメインを含む。場合によっては、追加のドメインは、塩基エディターの酵素的または触媒的機能、塩基エディターの結合機能を促進することができ、または所望の塩基編集結果に干渉し得る細胞機構(例えば酵素)の阻害因子であり得る。ある態様において、塩基エディターは、ヌクレアーゼ、ニッカーゼ、リコンビナーゼ、デアミナーゼ、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、転写アクチベーター、または転写リプレッサードメインを含むことができる。

【0407】

ある態様において、塩基エディターは、ウラシルグリコシラーゼ阻害因子 (UGI) ドメインを含むことができる。UGIドメインは、例えば、Cの脱アミノ化によって形成されるUがC核酸塩基に再変換されることを阻害することによって、シチジンデアミナーゼドメインを含む塩基エディターの効率を改善することができる。ある場合には、U:Gヘテロ二本鎖DNAの存在に対する細胞DNA修復応答が、細胞における核酸塩基編集効率の低下の原因となりうる。ある場合には、ウラシルDNAグリコシラーゼ (UDG) が、細胞内のDNAからのUの除去を触媒し得、それが塩基除去修復 (BER) を開始し得、大部分がU:G対からC:G対への復帰をもたらす。このような場合、一本鎖に結合する、編集された塩基をブロックする、UGIを阻害する、BERを阻害する、編集された塩基を保護する、および/または編集されていない鎖の修復を促進する、1以上のドメインを含む塩基エディターにおいてBERは阻害され得る。したがって、本開示は、UGIドメインを含む塩基編集融合タンパク質を企図する。

【0408】

ある態様において、塩基エディターは、二本鎖切断 (DSB) 結合タンパク質の全部または一部をドメインとして含む。例えば、DSB結合タンパク質は、バクテリオファージMuのGamタンパク質を含み得、これはDSBの末端に結合してそれらを分解から保護することができる。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” Science Advances 3:eaa04774 (2017)を参照されたい(その全内容が参照により本明細書に組み入れられる)。

【0409】

ある態様において、塩基エディターは、核酸ポリメラーゼ (NAP) の全てまたは一部をドメインとして含むことができる。例えば、塩基エディターは、真核生物NAPの全部または一部を含むことができる。いくつかの実施形態において、塩基エディターに組み込まれるNAPまたはその一部は、DNAポリメラーゼである。ある態様において、塩基エディターに組み込まれるNAPまたはその一部は、損傷乗り越えポリメラーゼ活性を有する。場合によっては、塩基エディターに組み込まれるNAPまたはその一部は損傷乗り越えDNAポリメラーゼである。いくつかの実施形態において、塩基エディターに組み込まれるNAPまたはその一部は、Rev7、Rev1複合体、ポリメラーゼイオタ、ポリメラーゼカッパ、またはポリメラーゼイタである。いくつかの実施形態において、塩基エディターに組み込まれるNAPまたはその一部は、真核生物ポリメラーゼ α 、 β 、 γ 、 δ 、 ϵ 、 γ 、 η 、iota、 κ 、 λ 、mu、またはnu成分である。ある態様において、塩基エディターに組み込まれるNAPまたはその一部は、核酸ポリメラーゼ(例えば損傷乗り越えDNAポリメラーゼ)に対して少なくとも75%、80%、85%、90%、95%、96%、97%、98%、99%、または99.5%の同一性であるアミノ酸配列を含む。

【0410】

[塩基エディターシステム]

本明細書で提供される塩基エディターシステムの使用は、(a) 対象のポリヌクレオチド(例えば、二本鎖DNAまたはRNA、一本鎖DNAまたはRNA)の標的ヌクレオチド配列を、核酸塩基エディター(例えば、アデノシン塩基エディターまたはシチジン塩基エディター)およびガイドポリ核酸(例えばgRNA)を含む塩基エディターシステムと接触させる工程であって、標的ヌクレオチド配列は標的核酸塩基対を含む、工程と；(b) 標的領域の鎖分離を誘導する工程と；(c) 標的領域の一本鎖における標的核酸塩基対の第1の核酸塩基を第2の核酸

塩基に変換する工程と；(d) 標的領域の鎖を2本以上は切断することなく、第1の核酸塩基に相補的な第3の核酸塩基が、第2の核酸塩基に相補的な第4の核酸塩基に置換される工程とを含む。一部の実施形態では、ステップ (b) は省略されることを理解されたい。ある態様において、標的核酸塩基対は、1以上の遺伝子における複数の核酸塩基対である。いくつかの実施形態において、本明細書に提供される塩基エディターシステムは、1以上の遺伝子における複数の核酸塩基対の多重編集を可能にする。ある態様において、複数の核酸塩基対は、同一遺伝子内に位置する。いくつかの実施形態において、複数の核酸塩基対は、1またはそれより多い遺伝子に位置し、ここで、少なくとも1つの遺伝子は、異なる遺伝子座に位置する。

【0411】

いくつかの実施形態において、切断された一本鎖(ニック鎖)は、ガイド核酸にハイブリダイズされる。いくつかの実施形態において、切断された一本鎖は、第一の核酸塩基を含む鎖と反対である。一部の実施形態では、塩基エディターはCas9ドメインを含む。いくつかの実施形態において、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。いくつかの実施形態において、第2の塩基はイノシンである。

【0412】

本明細書に提供される塩基編集システムは、触媒的に欠陥のあるStreptococcus pyogenes Cas9と、シチジンデアミナーゼと、塩基除去修復の阻害因子とを含む融合タンパク質を用いて、二本鎖DNA切断を生じることなく、ドナーDNA鋳型を必要とせず、かつ過剰な確率的挿入および欠失を誘導することなく、DNAにおけるプログラム可能な一塩基(C→TまたはA→G)変化を誘導する、ゲノム編集の新規なアプローチを提供する。

【0413】

本明細書では、塩基エディター系を用いて核酸塩基を編集するためのシステム、組成物、および方法を提供する。いくつかの実施形態において、塩基エディターシステムは、(1) 核酸塩基を編集するための、ポリヌクレオチドプログラム可能なヌクレオチド結合ドメインおよび核酸塩基編集ドメイン(例えばデアミナーゼドメイン)を含む塩基エディター (BE) と、(2) ポリヌクレオチドプログラム可能なヌクレオチド結合ドメインに付随するガイドポリヌクレオチド(例えばガイドRNA)とを含む。ある態様において、塩基エディターシステムは、シトシン塩基エディター (CBE) を含む。いくつかの実施形態では、塩基エディターシステムは、アデノシン塩基エディター (ABE) を含む。ある態様において、ポリヌクレオチドプログラム可能なヌクレオチド結合ドメインは、ポリヌクレオチドプログラム可能なDNA結合ドメインである。ある態様において、ポリヌクレオチドプログラム可能なヌクレオチド結合ドメインは、ポリヌクレオチドプログラム可能なRNA結合ドメインである。ある態様において、核酸塩基編集ドメインは、デアミナーゼドメインである。場合によっては、デアミナーゼドメインは、シトシンデアミナーゼまたはシチジンデアミナーゼであり得る。いくつかの実施形態では「シトシンデアミナーゼ」および「シチジンデアミナーゼ」という用語は交換可能に使用され得る。場合によっては、デアミナーゼドメインは、アデニンデアミナーゼまたはアデノシンデアミナーゼであり得る。いくつかの実施形態では「アデニンデアミナーゼ」および「アデノシンデアミナーゼ」という用語は交換可能に使用され得る。核酸塩基編集タンパク質の詳細は、国際PCT出願番号PCT/2017/045381 (WO2018/027078) およびPCT/US2016/058344 (WO2017/070632) に記載されており、これらの各々は、その全体が参照により本明細書に組み込まれる。Komor, A.C., et al., “Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage” Nature 533, 420-424 (2016); Gaudelli, N.M., et al., “Programmable base editing of A·T to G·C in genomic DNA without DNA cleavage” Nature 551, 464-471 (2017); および 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” Science Advances 3: eaao4774 (2017)も参照のこと(その内容全体が参照により本明細書に組み入れられる)。

10

20

30

40

50

【0414】

ある態様において、塩基エディターは、編集された鎖の塩基除去修復を阻害する。いくつかの実施形態において、塩基エディターは、編集されていない鎖を保護または結合する。一部の実施形態では、塩基エディターはUGI活性を含む。ある態様において、塩基エディターは、触媒的に不活性なイノシン特異的ヌクレアーゼを含む。一部の実施形態では、塩基エディターはニックアーゼ活性を含む。ある態様において、塩基対の意図された編集は、PAM部位の上流にある。ある態様において、塩基対の意図された編集は、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。いくつかの実施形態において、塩基対の意図された編集は、PAM部位の下流にある。ある態様において、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ヌクレオチド下流にある。

10

【0415】

いくつかの実施形態では、本方法は、標準(例えばNGG) PAMサイトを必要としない。いくつかの実施態様において、核酸塩基エディターは、リンカー又はスペーサーを含む。ある態様において、リンカーまたはスペーサーは、長さが1~25アミノ酸である。ある態様において、リンカーまたはスペーサーは、長さが5~20アミノ酸である。ある態様において、リンカーまたはスペーサーは、長さが10、11、12、13、14、15、16、17、18、19、または20アミノ酸である。

20

【0416】

一部の実施形態では、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。ある態様において、標的ウィンドウは、1~10ヌクレオチドを含む。ある態様において、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。いくつかの実施形態において、塩基対の意図された編集は、標的ウィンドウ内にある。いくつかの実施態様において、標的ウィンドウは、塩基対の意図された編集を含む。いくつかの実施形態において、本方法は、本明細書に提供される塩基エディターのいずれかを用いて実施される。いくつかの実施態様において、標的ウィンドウは、脱アミノ化ウィンドウである。

【0417】

いくつかの実施形態では、塩基エディターはシチジン塩基エディター (CBE) である。いくつかの実施形態において、非限定的な例示的CBEは、BE1 (APOBEC1-XTEN-dCas9)、BE2 (APOBEC1-XTEN-dCas9-UGI)、BE3 (APOBEC1-XTEN-dCas9(A840H)-UGI)、BE3-Gam、saBE3、saBE4-Gam、BE4、BE4-Gam、saBE4、またはsaBE4E-Gamである。BE4はAPOBEC1 - Cas9n(D10A)リンカーを32アミノ酸まで、Cas9n - UGIリンカーを9アミノ酸まで拡張し、単一塩基エディター構築物のC末端に別の9アミノ酸リンカーと共にUGIの第二コピーを付加する。塩基性エディタsaBE3およびsaBE4では、*S. pyogenes* Cas9n(D10A)をより小さい*S. aureus* Cas9n(D10A)で置き換えた。BE3-Gam、saBE3-Gam、BE4-Gam、およびsaBE4-Gamは、174残基のGamタンパク質が、16アミノ酸のXTENリンカーを介してBE3、saBE3、BE4、およびsaBE4のN末端に融合されたものを有する。

30

40

【0418】

一部の実施形態では、塩基エディターはアデノシン塩基エディター (ABE) である。ある態様において、アデノシン塩基エディターは、DNA中のアデニンを脱アミノ化することができる。ある態様において、アデノシン塩基エディターは、RNA中のアデニンを脱アミノ化することができる。ある態様において、ABEは、BE3のAPOBEC1成分を、天然または操作された大腸菌TadA、ヒトADAR2、マウスADA、またはヒトADAT2で置換することによって生成される。いくつかの実施形態において、ABEは、進化されたTadAバリエーションを含む。いくつかの実施態様では、ABEはABE 1.2 (TadA*-XTEN-nCas9-NLS) である。ある態様において、TadA*はA106VおよびD108N突然変異を含む。

【0419】

50

ある態様において、ABEは第二世代ABEである。ある態様において、ABEはABE 2.1であり、これは、*TadA** (*TadA*2.1*) において追加の突然変異D147YおよびE155Vを含むものである。いくつかの実施形態において、ABEはABE 2.2であり、これはABE 2.1が触媒的に不活性化されたバージョンのヒトアルキルアデニンDNAグリコシラーゼ(E125Q変異を伴うAAG)に融合されたものである。いくつかの実施形態において、ABEはABE 2.3であり、ここではABE 2.1が触媒的に不活性化された大腸菌Endo V (D35A変異で不活性化される)に融合されたものである。いくつかの態様において、ABEはABE 2.6であり、これはABE 2.1におけるリンカーの二倍の長さ(32アミノ酸、(SGGS)₂-XTEN-(SGGS)₂)のリンカーを有する。いくつかの実施態様において、ABEはABE 2.7であり、これはABE 2.1が追加の野生型*TadA*モノマーと連結されたものである。いくつかの実施態様では、ABEはABE 2.8であり、これはABE 2.1が追加の*TadA*2.1*モノマーでつながれたものである。いくつかの実施形態において、ABEはABE 2.9であり、これは進化された*TadA* (*TadA*2.1*) とABE 2.1のN末端との直接融合体である。いくつかの実施形態において、ABEはABE 2.10であり、これは野生型*TadA*とABE 2.1のN末端との直接融合体である。いくつかの実施形態において、ABEはABE 2.11であり、これは、*TadA**モノマーのN末端において不活性化E59A突然変異を有するABE 2.9である。いくつかの実施形態において、ABEはABE 2.12であり、これは、内部*TadA**モノマーにおいて不活性化E59A突然変異を有するABE 2.9である。

【0420】

ある態様において、ABEは、第三世代ABEである。ある態様において、ABEはABE 3.1であり、これは、三つのさらなる*TadA*突然変異(L84F、H123Y、およびI157F)を伴うABE 2.3である。

【0421】

いくつかの実施形態において、ABEは第四世代ABEである。いくつかの実施形態において、ABEはABE 4.3であり、これはさらなる*TadA*突然変異A142N (*TadA*4.3*) を伴うABE 3.1である。

【0422】

いくつかの実施形態において、ABEは第五世代ABEである。いくつかの実施形態において、ABEはABE 5.1であり、これは生存クローンからの突然変異のコンセンサセット(H36L、R51L、S146C、およびK157N) をABE 3.1に輸入することによって生成される。いくつかの実施形態において、ABEはABE 5.3であり、これは内部の進化された*TadA**に融合された野生型大腸菌*TadA*を含むヘテロ二量体構築物を有する。ある態様において、ABEは、以下の表2に示されるように、ABE 5.2、ABE 5.4、ABE 5.5、ABE 5.6、ABE 5.7、ABE 5.8、ABE 5.9、ABE 5.10、ABE 5.11、ABE 5.12、ABE 5.13、またはABE 5.14である。一部の実施形態では、ABEは第6世代ABEである。一部の実施形態では、ABEは、以下の表2に示されるように、ABE 6.1、ABE 6.2、ABE 6.3、ABE 6.4、ABE 6.5、またはABE 6.6である。いくつかの実施形態において、ABEは、第七世代ABEである。いくつかの実施形態において、ABEは、以下の表2に示されるように、ABE 7.1、ABE 7.2、ABE 7.3、ABE 7.4、ABE 7.5、ABE 7.6、ABE 7.7、ABE 7.8、ABE 7.9、またはABE 7.10である。

表2：ABEの遺伝子型

10

20

30

40

50

【表 2】

	23	26	36	37	48	49	51	72	84	87	105	108	123	125	142	145	147	152	155	156	157	16
ABE0.1	W	R	H	N	P		R	N	L	S	A	D	H	G	A	S	D	R	E	I	K	K
ABE0.2	W	R	H	N	P		R	N	L	S	A	D	H	G	A	S	D	R	E	I	K	K
ABE1.1	W	R	H	N	P		R	N	L	S	A	N	H	G	A	S	D	R	E	I	K	K
ABE1.2	W	R	H	N	P		R	N	L	S	V	N	H	G	A	S	D	R	E	I	K	K
ABE2.1	W	R	H	N	P		R	N	L	S	V	N	H	G	A	S	Y	R	V	I	K	K
ABE2.2	W	R	H	N	P		R	N	L	S	V	N	H	G	A	S	Y	R	V	I	K	K
ABE2.3	W	R	H	N	P		R	N	L	S	V	N	H	G	A	S	Y	R	V	I	K	K
ABE2.4	W	R	H	N	P		R	N	L	S	V	N	H	G	A	S	Y	R	V	I	K	K
ABE2.5	W	R	H	N	P		R	N	L	S	V	N	H	G	A	S	Y	R	V	I	K	K
ABE2.6	W	R	H	N	P		R	N	L	S	V	N	H	G	A	S	Y	R	V	I	K	K
ABE2.7	W	R	H	N	P		R	N	L	S	V	N	H	G	A	S	Y	R	V	I	K	K
ABE2.8	W	R	H	N	P		R	N	L	S	V	N	H	G	A	S	Y	R	V	I	K	K
ABE2.9	W	R	H	N	P		R	N	L	S	V	N	H	G	A	S	Y	R	V	I	K	K
ABE2.10	W	R	H	N	P		R	N	L	S	V	N	H	G	A	S	Y	R	V	I	K	K
ABE2.11	W	R	H	N	P		R	N	L	S	V	N	H	G	A	S	Y	R	V	I	K	K
ABE2.12	W	R	H	N	P		R	N	L	S	V	N	H	G	A	S	Y	R	V	I	K	K
ABE3.1	W	R	H	N	P		R	N	F	S	V	N	Y	G	A	S	Y	R	V	F	K	K
ABE3.2	W	R	H	N	P		R	N	F	S	V	N	Y	G	A	S	Y	R	V	F	K	K
ABE3.3	W	R	H	N	P		R	N	F	S	V	N	Y	G	A	S	Y	R	V	F	K	K
ABE3.4	W	R	H	N	P		R	N	F	S	V	N	Y	G	A	S	Y	R	V	F	K	K
ABE3.5	W	R	H	N	P		R	N	F	S	V	N	Y	G	A	S	Y	R	V	F	K	K
ABE3.6	W	R	H	N	P		R	N	F	S	V	N	Y	G	A	S	Y	R	V	F	K	K
ABE3.7	W	R	H	N	P		R	N	F	S	V	N	Y	G	A	S	Y	R	V	F	K	K
ABE3.8	W	R	H	N	P		R	N	F	S	V	N	Y	G	A	S	Y	R	V	F	K	K
ABE4.1	W	R	H	N	P		R	N	L	S	V	N	H	G	N	S	Y	R	V	I	K	K

10

20

30

40

50

	23	26	36	37	48	49	51	72	84	87	105	108	123	125	142	145	147	152	155	156	157	16
ABE4.2	W	G	H	N	P		R	N	L	S	V	N	H	G	N	S	Y	R	V	I	K	K
ABE4.3	W	R	H	N	P		R	N	F	S	V	N	Y	G	N	S	Y	R	V	F	K	K
ABE5.1	W	R	L	N	P		L	N	F	S	V	N	Y	G	A	C	Y	R	V	F	N	K
ABE5.2	W	R	H	S	P		R	N	F	S	V	N	Y	G	A	S	Y	R	V	F	K	T
ABE5.3	W	R	L	N	P		L	N	I	S	V	N	Y	G	A	C	Y	R	V	I	N	K
ABE5.4	W	R	H	S	P		R	N	F	S	V	N	Y	G	A	S	Y	R	V	F	K	T
ABE5.5	W	R	L	N	P		L	N	F	S	V	N	Y	G	A	C	Y	R	V	F	N	K
ABE5.6	W	R	L	N	P		L	N	F	S	V	N	Y	G	A	C	Y	R	V	F	N	K
ABE5.7	W	R	L	N	P		L	N	F	S	V	N	Y	G	A	C	Y	R	V	F	N	K
ABE5.8	W	R	L	N	P		L	N	F	S	V	N	Y	G	A	C	Y	R	V	F	N	K
ABE5.9	W	R	L	N	P		L	N	F	S	V	N	Y	G	A	C	Y	R	V	F	N	K
ABE5.10	W	R	L	N	P		L	N	F	S	V	N	Y	G	A	C	Y	R	V	F	N	K
ABE5.11	W	R	L	N	P		L	N	F	S	V	N	Y	G	A	C	Y	R	V	F	N	K
ABE5.12	W	R	L	N	P		L	N	F	S	V	N	Y	G	A	C	Y	R	V	F	N	K
ABE5.13	W	R	H	N	P		L	D	F	S	V	N	Y	A	A	S	Y	R	V	F	K	K
ABE5.14	W	R	H	N	S		L	N	F	C	V	N	Y	G	A	S	Y	R	V	F	K	K
ABE6.1	W	R	H	N	S		L	N	F	S	V	N	Y	G	N	S	Y	R	V	F	K	K
ABE6.2	W	R	H	N	T	V	L	N	F	S	V	N	Y	G	N	S	Y	R	V	F	N	K
ABE6.3	W	R	L	N	S		L	N	F	S	V	N	Y	G	A	C	Y	R	V	F	N	K
ABE6.4	W	R	L	N	S		L	N	F	S	V	N	Y	G	N	C	Y	R	V	F	N	K
ABE6.5	W	R	L	N	I	V	L	N	F	S	V	N	Y	G	A	C	Y	R	V	F	N	K
ABE6.6	W	R	L	N	T	V	L	N	F	S	V	N	Y	G	N	C	Y	R	V	F	N	K
ABE7.1	W	R	L	N	A		L	N	F	S	V	N	Y	G	A	C	Y	R	V	F	N	K
ABE7.2	W	R	L	N	A		L	N	F	S	V	N	Y	G	N	C	Y	R	V	F	N	K
ABE7.3	I	R	L	N	A		L	N	F	S	V	N	Y	G	A	C	Y	R	V	F	N	K
ABE7.4	R	R	L	N	A		L	N	F	S	V	N	Y	G	A	C	Y	R	V	F	N	K
ABE7.5	W	R	L	N	A		L	N	F	S	V	N	Y	G	A	C	Y	H	V	F	N	K
ABE7.6	W	R	L	N	A		L	N	I	S	V	N	Y	G	A	C	Y	P	V	I	N	K

10

20

30

40

50

	23	26	36	37	48	49	51	72	84	87	105	108	123	125	142	145	147	152	155	156	157	16
ABE7.7	L	R	L	N	A		L	N	F	S	V	N	Y	G	A	C	Y	P	V	F	N	K
ABE7.8	I	R	L	N	A		L	N	F	S	V	N	Y	G	N	C	Y	R	V	F	N	K
ABE7.9	L	R	L	N	A		L	N	F	S	V	N	Y	G	N	C	Y	P	V	F	N	K
ABE7.10	R	R	L	N	A		L	N	F	S	V	N	Y	G	A	C	Y	P	V	F	N	K

10

【0423】

いくつかの実施形態では、塩基エディターは、ポリヌクレオチドプログラム可能DNA結合ドメインおよびシチジン核酸塩基を脱アミノ化するためのシチジンデアミナーゼドメインを含み、ここで、ガイドポリヌクレオチドが、病原性突然変異に関連する遺伝子、例えば、ACADM、HBB、PDS、SNCA、もしくはSERPINA1のコード領域、または本明細書の表4に列挙されているもののような遺伝子の調節領域に位置する標的ヌクレオチド配列に塩基エディターをターゲティングする。

【0424】

ある態様において、塩基エディターは、核酸塩基編集ドメイン(例えばデアミナーゼドメインの全部または一部)に融合されたポリヌクレオチドプログラム可能なヌクレオチド結合ドメイン(例えばCas9由来ドメイン)を含む融合タンパク質である。ある態様において、塩基エディターは、ウラシルグリコシラーゼ阻害因子(UGI)の全部または一部を含むドメインをさらに含む。ある態様において、塩基エディターは、ウラシルDNAグリコシラーゼ(UDG)などの、ウラシル結合タンパク質(UBP)の全部または一部を含むドメインを含む。ある態様において、塩基エディターは、核酸ポリメラーゼの全てまたは一部を含むドメインを含む。ある態様において、塩基エディターに組み込まれる核酸ポリメラーゼまたはその一部は、損傷乗り越えDNAポリメラーゼである。

20

【0425】

いくつかの実施形態では、塩基エディターのドメインは、複数のドメインを含むことができる。例えば、Cas9に由来するポリヌクレオチドプログラム可能なヌクレオチド結合ドメインを含む塩基エディターは、野生型または天然のCas9のRECローブおよびNUCローブに対応するRECローブおよびNUCローブを含むことができる。別の例では、塩基エディターは、RuvCIドメイン、BHドメイン、REC1ドメイン、REC2ドメイン、RuvCIIドメイン、L1ドメイン、HNHドメイン、L2ドメイン、RuvCIIIドメイン、WEDドメイン、TOPOドメインまたはCTDドメインのうちの1つ以上を含むことができる。ある態様において、塩基エディターの一つ以上のドメインは、そのドメインを含むポリペプチドの野生型バージョンに対する突然変異(例えば置換、挿入、削除)を含む。例えば、ポリヌクレオチドプログラム可能DNA結合ドメインのHNHドメインは、H840A置換を含むことができる。別の例では、ポリヌクレオチドプログラム可能なDNA結合ドメインのRuvCIドメインは、D10A置換を含むことができる。

30

40

【0426】

本明細書に開示される塩基エディターの異なるドメイン(例えば隣接ドメイン)は、一つ以上のリンカードメイン(例えばXTENリンカードメイン)を使用して、または使用せずに、互いに接続することができる。場合によっては、リンカードメインは、二つの分子もしくは部分(例えば、第一のドメイン(例えばCas9由来ドメイン)および第二のドメイン(例えばシチジンデアミナーゼドメインまたはアデノシンデアミナーゼドメイン)のような融合タンパク質の二つのドメイン)をつなげる結合(例えば共有結合)、化学基、または分子であり得る。ある態様において、リンカーは、共有結合(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。ある態様において、リンカーは、アミド結合

50

における炭素窒素結合である。特定の実施形態では、リンカーは、環状もしくは非環状、置換もしくは非置換、分枝もしくは非分枝の脂肪族もしくはヘテロ脂肪族リンカーである。ある態様において、リンカーは、ポリマー(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)である。特定の実施態様では、リンカーは、アミノアルカン酸のモノマー、ダイマー又はポリマーを含む。ある態様において、リンカーは、アミノアルカン酸(例えばグリシン、エタン酸、アラニン、 β -アラニン、3-アミノプロパン酸、4-アミノブタン酸、5-ペンタン酸等)を含む。いくつかの実施態様において、リンカーは、アミノヘキサン酸(Ahx)のモノマー、ダイマー又はポリマーを含む。ある態様において、リンカーは炭素環部分(例えばシクロペンタン、シクロヘキサン)に基づく。他の態様において、リンカーはポリエチレングリコール部分(PEG)を含む。ある態様において、リンカーは、アリールまたはヘテロアリール部分を含む。ある態様において、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)の結合を促進させる官能化部分を含むことができる。リンカーの一部として任意の求電子剤を使用することができる。例示的な求電子剤としては、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、ハロゲン化アリール、ハロゲン化アシル、およびイソチオシアナートが挙げられるが、これらに限定されない。ある態様において、リンカーは、Cas9ヌクレアーゼドメインを含むRNAプログラム可能ヌクレアーゼのgRNA結合ドメインと、核酸編集タンパク質の触媒ドメインとを結合する。ある態様において、リンカーは、dCas9と第二のドメイン(例、シチジンデアミナーゼ、UGI等)とを連結する。【0427】

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配置されるか、または2つの基、分子、または他の部分によって隣接され、共有結合を介してそれぞれに連結され、したがって2つを連結する。典型的には、リンカーは、このように2つを連結する。ある態様において、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タンパク質)である。ある態様において、リンカーは、有機分子、基、ポリマー、または化学的部分である。ある態様において、リンカーは、長さが2~100アミノ酸であり、例えば、長さが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、30~35、35~40、40~45、45~50、50~60、60~70、70~80、80~90、90~100、100~150~200アミノ酸である。ある態様において、リンカーは、長さが約3~約104(例えば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、55、60、65、70、75、80、85、90、95、100)アミノ酸である。より長いまたはより短いリンカーも考えられる。ある態様において、リンカードメインは、XTENリンカーとも呼ばれ得るアミノ酸配列SGSETPGTSESATPESを含む。融合タンパク質ドメインを連結するための任意の方法を用いて(例えば、(SGGS)n、(GGGS)n、(GGGG)n、および(G)nの形態の非常に柔軟なリンカーから、(EAAAK)n、(GGS)n、SGSETPGTSESATPESの形態のより剛性の高いリンカーまで(例えばGuilinger JP, Thompson DB, Liu DR. Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. Nat. Biotechnol. 2014; 32(6): 577-82参照。その全内容は参照によりここに組み込まれる)、または(XP)nモチーフ)、核酸塩基エディターの活性のための最適な長さを達成することができる。いくつかの実施態様において、nは、1、2、3、4、5、6、7、8、9、10、11、12、13、14又は15である。ある態様において、リンカーは、(GGS)nモチーフを含み、ここで、nは、1、3、または7である。いくつかの実施形態において、本明細書で提供される融合タンパク質のCas9ドメインは、アミノ酸配列SGSETPGTSESATPESを含むリンカーを介して融合される。いくつかの実施形態において、リンカーは、複数のプロリン残基を含み長さが5~21、5~14、5~9、5~7アミノ酸であり、例えば、PAPAP、PAPAPA、PAPAPAP、PAPAPAPA、P(AP)₄、P(AP)₇、P(AP)₁₀である(例えば、Tan J, Zhang F, Karcher D, Bock R. Engineering of high-precision base editors for site-specific single nucleotide replacement. Nat Comm

10

20

30

40

50

un. 2019 Jan 25;10(1):439を参照；その全内容は参照によりここに組み込まれる)。このようなプロリンに富むリンカーはまた、「硬質」リンカーとも呼ばれる。

【0428】

本発明の融合タンパク質は、核酸編集ドメインを含む。ある態様において、核酸編集ドメインは、CからUへの塩基変化を触媒することができる。ある態様において、核酸編集ドメインは、デアミナーゼドメインである。ある態様において、デアミナーゼは、シチジンデアミナーゼまたはアデノシンデアミナーゼである。ある態様において、デアミナーゼは、アポリボタンパク質B mRNA編集複合体 (APOBEC) ファミリーデアミナーゼである。ある態様において、デアミナーゼは、APOBEC1デアミナーゼである。ある態様において、デアミナーゼは、APOBEC 2デアミナーゼである。ある態様において、デアミナーゼは、APOBEC3デアミナーゼである。ある態様において、デアミナーゼは、APOBEC3Aデアミナーゼである。ある態様において、デアミナーゼは、APOBEC3Bデアミナーゼである。ある態様において、デアミナーゼは、APOBEC3Cデアミナーゼである。ある態様において、デアミナーゼは、APOBEC3Dデアミナーゼである。ある態様において、デアミナーゼは、APOBEC3Eデアミナーゼである。ある態様において、デアミナーゼは、APOBEC3Fデアミナーゼである。ある態様において、デアミナーゼは、APOBEC3Gデアミナーゼである。ある態様において、デアミナーゼは、APOBEC3Hデアミナーゼである。ある態様において、デアミナーゼは、APOBEC4デアミナーゼである。ある態様において、デアミナーゼは、活性化誘導デアミナーゼ (AID) である。ある態様において、デアミナーゼは、脊椎動物デアミナーゼである。ある態様において、デアミナーゼは、無脊椎動物デアミナーゼである。ある態様において、デアミナーゼは、ヒト、チンパンジー、ゴリラ、サル、ウシ、イヌ、ラット、またはマウスデアミナーゼである。ある態様において、デアミナーゼは、ヒトデアミナーゼである。ある態様において、デアミナーゼは、ラットデアミナーゼ、例えば、rAPOBEC1である。ある態様において、デアミナーゼは、*Petromyzon marinus*シチジンデアミナーゼ1 (pmCDA1) である。ある態様において、デアミナーゼは、ヒトAPOBEC3Gである。ある態様において、デアミナーゼは、ヒトAPOBEC3Gの断片である。ある態様において、デアミナーゼは、D316R D317R突然変異を含むヒトAPOBEC3Gバリエーションである。ある態様において、デアミナーゼは、ヒトAPOBEC3Gの断片であり、D316R D317R突然変異に対応する突然変異を含む。ある態様において、核酸編集ドメインは、本明細書に記載される任意のデアミナーゼのデアミナーゼドメインに対して少なくとも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%の同一性である。

【0429】

[ガイドRNAを伴うCas9複合体]

本開示のいくつかの態様は、本明細書に提供される融合タンパク質のいずれか、およびガイドRNA (例えば、表3A及び3Bに記載されている遺伝子のような目的の遺伝子、又は表4に記載されている遺伝子の調節配列を標的とするガイド)を含む複合体を提供する。

【0430】

ある態様において、ガイド核酸(例えばガイドRNA)は、15~100ヌクレオチド長であり、標的配列に相補的である少なくとも10個の連続するヌクレオチドの配列を含む。いくつかの実施形態において、ガイドRNAは、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、または50ヌクレオチドの長さである。いくつかの実施形態において、ガイドRNAは、標的配列に相補的な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個の連続したヌクレオチドの配列を含む。ある態様において、標的配列はDNA配列である。ある態様において、標的配列は、細菌、酵母、真菌、昆虫、植物または動物のゲノムにおける配列である。ある態様において、標的配列は、ヒトのゲノムにおける配列である。いくつかの実施形態において、標的配列の3'末端は、標準PAM配列

(NGG) にすぐ隣接している。いくつかの実施形態において、標的配列の3'末端は、非標準PAM配列(例えば、表3A若しくは3Bに列挙されている配列又は5'-NAA-3')にすぐ隣接している。ある態様において、ガイド核酸(例えばガイドRNA)は、疾患標的化可能突然変異を有する遺伝子における配列に相補的である。

【0431】

本開示のいくつかの態様は、本明細書に提供される融合タンパク質または複合体を使用する方法を提供する。例えば、本開示のいくつかの局面は、DNA分子を、本明細書中に提供される融合タンパク質のいずれか、および少なくとも一つのガイドRNAと接触させることを含む方法を提供し、ここで、ガイドRNAは、約15~100ヌクレオチド長であり、標的配列に相補的である少なくとも10個の連続したヌクレオチドの配列を含む。いくつかの実施形態において、標的配列の3'末端は、AGC、GAG、TTT、GTG、またはCAA配列にすぐ隣接している。いくつかの実施形態において、標的配列の3'末端は、NGA、NAA、NGCG、NGN、NNGRRT、NNNRRT、NGCG、NGCN、NGTN、NGTN、NGTN、または5'(TTTV)配列にすぐ隣接している。

【0432】

それぞれの配列における特定の位置または残基の番号付けは、使用される特定のタンパク質および番号付けスキームに依存することが理解されるであろう。たとえば、成熟タンパク質の前駆体と成熟タンパク質そのものでは番号付けが異なることがあり、種ごとの配列の違いが番号付けに影響することがある。当業者は、当業者に周知の方法、例えば、配列アラインメントおよび相同的残基の決定によって、任意の相同的タンパク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。

【0433】

本明細書に開示された融合タンパク質のいずれかを標的部位、例えば、編集される突然変異を含む部位にターゲティングするためには、融合タンパク質をガイドRNAと共に共発現させることが典型的に必要であることは当業者には明らかであろう。本明細書の他の箇所より詳細に説明されるように、ガイドRNAは、典型的には、Cas9結合を可能にするtracrRNAフレームワークと、Cas9:核酸編集酵素/ドメイン融合タンパク質に配列特異性を付与するガイド配列とを含む。あるいは、ガイドRNAおよびtracrRNAは、2つの核酸分子として別々に提供され得る。いくつかの態様において、ガイドRNAは、ガイド配列が標的配列に相補的な配列を含むという構造を含む。ガイド配列は典型的には20ヌクレオチド長である。Cas9:核酸編集酵素/ドメイン融合タンパク質を特定のゲノム標的部位にターゲティングするための適切なガイドRNAの配列は、本開示に基づいて当業者に明らかであろう。そのような適切なガイドRNA配列は、典型的には、編集される標的ヌクレオチドの50ヌクレオチド以内の上流または下流内の核酸配列に相補的なガイド配列を含む。提供された融合タンパク質のいずれかを特定の標的配列に標的化するのに適したいくつかの例示的なガイドRNA配列が本明細書に提供される。

【0434】

本明細書に開示される塩基エディターのドメインは、任意の順序で配置することができる。例えばポリヌクレオチドプログラム可能なヌクレオチド結合ドメインおよびデアミナーゼドメインを含む融合タンパク質を含む塩基エディターの非限定的な例は、以下のように配置され得る。

NH₂-[核酸塩基編集ドメイン]-リンカー1-[例えば Cas9 由来ドメイン]-COOH;
NH₂-[例えば シチジンデアミナーゼ]-リンカー1-[例えば Cas9 由来ドメイン]-COOH;
NH₂-[例えば シチジンデアミナーゼ]-リンカー1-[例えば Cas9 由来ドメイン]-リンカー2-[UGI]-COOH;
NH₂-[例えば APOBEC]-リンカー1-[例えば Cas9 由来ドメイン]-COOH;
NH₂-[例えば シチジンデアミナーゼ]-リンカー1-[例えば Cas9 由来ドメイン]-COOH;
NH₂-[例えば APOBEC]-リンカー1-[例えば Cas9 由来ドメイン]-COOH;
NH₂-[例えば APOBEC]-リンカー1-[例えば Cas9 由来ドメイン]-リンカー2-[UGI]-COOH

OH

NH₂-[例えば アデノシンデアミナーゼ]-[例えば Cas9 由来ドメイン]-COOH;
 NH₂-[例えば Cas9 由来ドメイン]-[例えば アデノシンデアミナーゼ]-COOH;
 NH₂-[例えば アデノシンデアミナーゼ]-[例えば Cas9 由来ドメイン]-[イノシンBER阻害因子]-COOH;
 NH₂-[例えば アデノシンデアミナーゼ]-[イノシンBER阻害因子]-[例えば Cas9 由来ドメイン]-COOH;
 NH₂-[イノシンBER阻害因子]-[例えば アデノシンデアミナーゼ]-[例えば Cas9 由来ドメイン]-COOH;
 NH₂-[例えば Cas9 由来ドメイン]-[例えば アデノシンデアミナーゼ]-[イノシンBER阻害因子]-COOH;
 NH₂-[例えば Cas9 由来ドメイン]-[イノシンBER阻害因子]-[例えば アデノシンデアミナーゼ]-COOH; または
 NH₂-[イノシンBER阻害因子]-[例えば Cas9 由来ドメイン]-[例えば アデノシンデアミナーゼ]-COOH
 【0435】

さらに、場合によっては、Gamタンパク質を塩基エディターのN末端に融合させることができる。場合によっては、Gamタンパク質を塩基エディターのC末端に融合させることができる。バクテリオファージMuのGamタンパク質は二本鎖切断 (DSB) の末端に結合し、それらを分解から保護することができる。いくつかの実施形態において、Gamを使用してDSBの遊離末端を結合することにより、塩基編集のプロセス中のインデル形成を低減することができる。いくつかの実施形態において、174残基のGamタンパク質は、塩基エディターのN末端に融合される。例えば、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” Science Advances 3:eaao4774 (2017)を参照されたい。場合によっては、1つまたは複数の突然変異が、野生型ドメインと比較して、塩基編集ドメインの長さを変化させることができる。例えば、少なくとも1つのドメインにおける少なくとも1つのアミノ酸の欠失は、塩基エディターの長さを短くすることができる。別の例では、1つまたは複数の突然変異は、野生型ドメインに対するドメインの長さを変化させない。例えば、いずれかのドメインにおける置換（複数可）は塩基エディターの長さを変えない。全てのドメインの長さが野生型ドメインと同じであるこのような塩基エディターの非限定的な例は、以下を含むことができる：

NH₂-[APOBEC1]-リンカー1-[Cas9(D10A)]-リンカー2-[UGI]-COOH;
 NH₂-[CDA1]-リンカー1-[Cas9(D10A)]-リンカー2-[UGI]-COOH;
 NH₂-[AID]-リンカー1-[Cas9(D10A)]-リンカー2-[UGI]-COOH;
 NH₂-[APOBEC1]-リンカー1-[Cas9(D10A)]-リンカー2-[SSB]-COOH;
 NH₂-[UGI]-リンカー1-[APOBEC1]-リンカー2-[Cas9(D10A)]-COOH;
 NH₂-[APOBEC1]-リンカー1-[Cas9(D10A)]-リンカー2-[UGI]-リンカー3-[UGI]-COOH;
 NH₂-[Cas9(D10A)]-リンカー1-[CDA1]-リンカー2-[UGI]-COOH;
 NH₂-[Gam]-リンカー1-[APOBEC1]-リンカー2-[Cas9(D10A)]-リンカー3-[UGI]-COOH;
 NH₂-[Gam]-リンカー1-[APOBEC1]-リンカー2-[Cas9(D10A)]-リンカー3-[UGI]-リンカー4-[UGI]-COOH;
 NH₂-[APOBEC1]-リンカー1-[dCas9(D10A, H840A)]-リンカー2-[UGI]-COOH; または
 NH₂-[APOBEC1]-リンカー1-[dCas9(D10A, H840A)]-COOH
 【0436】

いくつかの実施形態において、本明細書中で提供される塩基編集融合タンパク質は、正確な位置、例えば、標的塩基が規定された領域(例えば「脱アミノ化ウィンドウ」)内に配置されるような位置に配置される必要がある。場合によっては、標的は4塩基領域内にあ

り得る。場合によっては、そのような規定された標的領域は、PAMの約15塩基上流にあり得る。例えば、Komor, A.C., et al., “Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage” Nature 533, 420-424 (2016); Gaudelli, N.M., et al., “Programmable base editing of A·T to G·C in genomic DNA without DNA cleavage” Nature 551, 464-471 (2017); and 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” Science Advances 3:eaa04774 (2017) 参照；その全内容が参照により本明細書に組み入れられる。

【0437】

定義された標的領域は、脱アミノ化ウィンドウであり得る。脱アミノ化ウィンドウは、塩基エディターが標的ヌクレオチドに作用して脱アミノ化する、規定された領域であり得る。いくつかの実施態様において、脱アミノ化ウィンドウは、2、3、4、5、6、7、8、9又は10塩基領域内にある。いくつかの実施態様において、脱アミノ化ウィンドウは、PAMの5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24又は25塩基上流にある。

【0438】

本開示の塩基エディターは、標的ポリヌクレオチド配列の編集を促進させる任意のドメイン、特徴またはアミノ酸配列を含むことができる。例えば、いくつかの実施形態において、塩基エディターは、核局在化配列 (NLS) を含む。いくつかの実施形態において、塩基エディターのNLSは、デアミナーゼドメインとポリヌクレオチドプログラム可能なヌクレオチド結合ドメインとの間に位置する。いくつかの実施形態において、塩基エディターのNLSは、ポリヌクレオチドプログラム可能なヌクレオチド結合ドメインに対してC末端に位置する。

【0439】

本明細書に開示される塩基エディターに存在し得る他の例示的特徴は、細胞質局在化配列、核輸出配列などの輸出配列、または他の局在化配列などの局在化配列、ならびに融合タンパク質の可溶化、精製、または検出に有用な配列タグである。本明細書において提供される適切なタンパク質タグには、限定されるものではないが、ビオチンカルボキシラーゼキャリアータグ (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、ヒスチジンタグまたはHis-タグとも呼ばれるポリヒスチジンタグ、マルトース結合タンパク質 (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、緑色蛍光タンパク質 (GFP) -タグ、チオレドキシントグ、S-タグ、ソフタグ (例ソフタグ1、ソフタグ3)、ストレプトタグ、ビオチンリガーゼタグ、FLASHタグ、V5タグ、およびSBP-タグが含まれる。さらなる適切な配列は、当業者には明らかであろう。いくつかの実施形態において、融合タンパク質は、1つ以上のHisタグを含む。

【0440】

融合タンパク質に含まれ得るタンパク質ドメインの非限定的な例としては、デアミナーゼドメイン (例えばシチジンデアミナーゼおよび/またはアデノシンデアミナーゼ)、ウリシルグリコシラーゼ阻害因子 (UGI) ドメイン、エピトープタグ、レポーター遺伝子配列、および/または以下の活性の一つ以上を有するタンパク質ドメインが挙げられる: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、転写放出因子活性、ヒストン修飾活性、RNA切断活性、および核酸結合活性。追加のドメインは、異種の機能的ドメインであり得る。そのような異種機能ドメインは、DNAメチル化、DNA損傷、DNA修復、標的DNAと結合した標的ポリペプチド (例えばヒストン、DNA結合タンパク質など) の修飾などの機能活性を付与し、例えばヒストンメチル化、ヒストンアセチル化、ヒストンユビキチン化などをもたらすことができる。

【0441】

付与される他の機能には、メチルトランスフェラーゼ活性、脱メチル化活性、脱アミノ化活性、ジスミターゼ活性、アルキル化活性、脱プリン化活性、酸化活性、ピリミジンニ

10

20

30

40

50

量体形成活性、インテグラーゼ活性、トランスポザーゼ活性、リコンビナーゼ活性、ポリメラーゼ活性、リガーゼ活性、ヘリカーゼ活性、フォトリアーゼ活性またはグリコシラーゼ活性、アセチルトランスフェラーゼ活性、デアセチラーゼ活性、キナーゼ活性、ホスファターゼ活性、ユビキチンリガーゼ活性、脱ユビキチン化活性、アデニル化活性、脱アデニル化活性、SUMO化活性、脱SUMO化活性、リボシル化活性、脱リボシル化活性、ミリスチル化活性、リモデリング活性、プロテアーゼ活性、酸化還元酵素活性、トランスフェラーゼ活性、ヒドロラーゼ活性、イソメラーゼ活性、シンターゼ活性、シンテターゼ活性、および脱ミリスチル化活性、またはこれらの任意の組合せが含まれる。

【0442】

エピトープタグの非限定的な例は、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ン (Trx) タグを含む。レポーター遺伝子の例としては、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)、 β -ガラクトシダーゼ、 β -グルクロニダーゼ、ルシフェラーゼ、緑色蛍光タンパク質 (GFP)、HcRed、DsRed、シアン蛍光タンパク質 (CFP)、黄色蛍光タンパク質 (YFP)、および青色蛍光タンパク質 (BFP) を含む自己蛍光タンパク質が挙げられるが、これらに限定されない。さらなるタンパク質配列には、DNA分子に結合するかまたは他の細胞分子に結合するアミノ酸配列が含まれ得、これらには、マルトース結合タンパク質 (MBP)、Sタグ、Lex A 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タンパク質融合体が挙げられるが、これらに限定されない。

【0443】

〔塩基エディターの効率性〕

CRISPR-Cas9ヌクレアーゼは、標的ゲノム編集を媒介するために広く使用されている。ほとんどのゲノム編集応用において、Cas9はガイドポリヌクレオチド(例えば単一ガイドRNA (sgRNA))と複合体を形成し、sgRNA配列により指定される標的部位で二本鎖DNA切断 (DSB) を誘導する。細胞は主に非相同末端結合 (NHEJ) 修復経路を介してこのDSBに応答し、遺伝子を破壊するフレームシフト変異を生じる確率的挿入または欠失 (インデル) を生じる。DSBに隣接する配列と高度の相同性を有するドナーDNA鋳型の存在下で、相同性指向性修復 (HDR) として知られる代替経路を介して遺伝子補正が達成され得る。残念ながら、ほとんどの非侵襲的条件下では、HDRは非効率であり、細胞状態および細胞型に依存し、より高いインデルの頻度によって支配される。ヒトの疾患に関連する既知の遺伝的変異の大部分は点突然変異であるため、より効率的かつクリーンに正確な点突然変異を作製する方法が必要である。本明細書に提供される塩基編集システムは、二本鎖DNA切断を生じることなく、ドナーDNA鋳型を必要とせず、かつ過剰な確率的挿入および欠失を誘導することなく、ゲノム編集を編集するための新しい方法を提供する。

【0444】

本明細書に提供される塩基エディターは、著しい割合のインデルを生成することなく、特定のヌクレオチド塩基を改変することができる。「インデル (indel)」という用語は、本明細書において使用される場合、核酸内のヌクレオチド塩基の挿入または欠失を指す。このような挿入または欠失は、遺伝子のコード領域内でフレームシフト突然変異を引き起こす可能性がある。いくつかの実施形態において、標的ヌクレオチド配列において多数の挿入または欠失(すなわちインデル)を生じることなく、核酸内の特定のヌクレオチドを効率的に改変(例えば変異または脱アミノ化)する塩基エディターを生成することが望ましい。特定の実施形態では、本明細書に提供される塩基エディターのいずれも、インデルに対して意図された改変(例えば点突然変異または脱アミノ化)のより大きな割合を生成することができる。

【0445】

いくつかの実施形態において、本明細書で提供される塩基エディター系のいずれも、標的ポリヌクレオチド配列において、50%未満、40%未満、30%未満、20%未満、19%未満

、18%未満、17%未満、16%未満、15%未満、14%未満、13%未満、12%未満、11%未満、10%未満、9%未満、8%未満、7%未満、6%未満、5%未満、4%未満、3%未満、2%未満、1%未満、0.9%未満、0.8%未満、未満、0.7%未満、0.6%未満、0.5%未満、0.4%未満、0.3%未満、0.2%未満、0.1%未満、0.09%未満、0.08%未満、0.07%未満、未満、0.06%未満、0.05%未満、0.04%未満、0.03%未満、0.02%未満、または0.01%未満のインデル形成をもたらす。

【0446】

本開示のいくつかの態様は、本明細書に提供される塩基エディターが、有意な数の非意図的な突然変異（例えば非意図的な点突然変異など）を生成することなく、核酸（例えば対象のゲノム内の核酸）において、意図する突然変異（例えば点突然変異など）を効率的に生成することができるという認識に基づく。

10

【0447】

いくつかの実施形態において、本明細書に提供される塩基エディターは、少なくとも0.01%の意図された突然変異（すなわち少なくとも0.01%の編集効率）を生じることができる。いくつかの実施形態において、本明細書に提供される塩基エディターは、少なくとも0.01%、1%、2%、3%、4%、5%、10%、15%、20%、25%、30%、40%、45%、50%、60%、70%、80%、90%、95%、または99%の意図された突然変異を生成することができる。

【0448】

いくつかの実施形態において、本明細書に提供される塩基エディターは、1:1より大きい、意図された点突然変異対インデルの比を生成することができる。いくつかの実施形態において、本明細書に提供される塩基エディターは、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8.5:1、少なくとも9:1、少なくとも10:1、少なくとも11:1、少なくとも12:1、少なくとも13:1、少なくとも14:1、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、またはそれ以上の、意図される点突然変異対インデルの比を生じることができる。

20

30

【0449】

意図された突然変異およびインデルの数は、例えば、国際PCT出願番号PCT/2017/045381 (WO2030/027078) およびPCT/US2016/058344 (WO2017/070632); Komor, A.C., et al., “Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage” Nature 533, 420-424 (2016); Gaudelli, N. M., et al., “Programmable base editing of A·T to G·C in genomic DNA without DNA cleavage” Nature 551, 464-471 (2017); および 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,” Science Advances 3:eaao4774 (2017)（その全内容は参照によりここに組み込まれる）に記載されているもののような、任意の適切な方法を用いて決定することができる。

40

【0450】

いくつかの実施形態において、インデル頻度を計算するために、インデルが生じ得るウィンドウの両側に隣接する二つの10 bp配列に対する正確なマッチについて配列決定リードがスキャンされる。完全マッチが見つからない場合、そのリードは解析から除外される。このインデルウィンドウの長さが参照配列と完全にマッチする場合、リードはインデルを含まないものとして分類される。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、配列決定リードは、それぞれ挿入または欠失として分類される。いくつかの態様において、本明細書において提供される塩基エディターは、核酸の領域におけ

50

るインデルの形成を制限することができる。ある態様において、その領域は、塩基エディターによって標的化されるヌクレオチドのところにあるか、または塩基エディターによって標的化されるヌクレオチドの2、3、4、5、6、7、8、9または10ヌクレオチド以内の領域である。

【0451】

標的ヌクレオチド領域で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)が塩基エディターに曝される時間の量に依存し得る。いくつかの実施形態において、インデルの数または割合は、標的ヌクレオチド配列(例えば細胞のゲノム内の核酸)を塩基エディターに曝してから少なくとも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日後に決定される。本明細書に記載される塩基エディターの特徴は、本明細書に提供される融合タンパク質、または融合タンパク質を使用する方法のいずれにも適用され得ることが理解されるべきである。

【0452】

[多重編集]

いくつかの実施形態において、本明細書に提供される塩基エディターシステムは、1以上の遺伝子における複数の核酸塩基対の多重編集を可能にする。ある態様において、複数の核酸塩基対は、同一遺伝子内に位置する。いくつかの実施形態において、複数の核酸塩基対は、1つまたはそれより多い遺伝子に位置し、ここで、少なくとも1つの遺伝子は、異なる遺伝子座に位置する。ある態様において、多重編集は、1以上のガイドポリヌクレオチドを含むことができる。いくつかの実施形態では、多重編集は、1つ以上の塩基エディターシステムを含むことができる。いくつかの実施形態において、多重編集は、単一のガイドポリヌクレオチドを有する1つ以上の塩基エディターシステムを含むことができる。いくつかの実施形態において、多重編集は、複数のガイドポリヌクレオチドを有する1つ以上の塩基エディターシステムを含むことができる。いくつかの実施形態において、多重編集は、単一の塩基エディター系を有する1以上のガイドポリヌクレオチドを含むことができる。いくつかの実施形態において、多重編集は、標的ポリヌクレオチド配列への結合を標的化するためにPAM配列を必要としない少なくとも1つのガイドポリヌクレオチドを含むことができる。いくつかの実施形態において、多重編集は、標的ポリヌクレオチド配列への結合を標的化するためにPAM配列を必要とする少なくとも1つのガイドポリヌクレオチドを含むことができる。いくつかの実施形態において、多重編集は、標的ポリヌクレオチド配列への結合を標的化するためにPAM配列を必要としない少なくとも1つのガイドポリヌクレオチドと、標的ポリヌクレオチド配列への結合を標的化するためにPAM配列を必要とする少なくとも1つのガイドポリヌクレオチドとの混合物を含むことができる。本明細書に記載される塩基エディターのいずれかを使用する多重編集の特徴は、本明細書に提供される塩基エディターのいずれかを使用する方法の任意の組み合わせに適用され得ることを理解されたい。また、本明細書に記載される塩基エディターのいずれかを使用する多重編集は、複数の核酸塩基対の順次的編集を含むことができることを理解されたい。

【0453】

本明細書で提供される方法は、(a) 対象のポリヌクレオチドの標的ヌクレオチド配列(例えば二本鎖DNA配列)を、核酸塩基エディター(例えばアデノシン塩基エディターまたはシチジン塩基エディター)およびガイドポリヌクレオチド酸(例えばgRNA)を含む塩基エディターシステムと接触させる工程であって、ここで、標的ヌクレオチド配列は標的核酸塩基対を含む、工程と；(b) 標的領域の鎖分離を誘導する工程と；(c) 標的領域の一本鎖における標的核酸塩基対の第一の核酸塩基を第二の核酸塩基に編集する工程と；(d) 標的領域の鎖を1本より多くは切断せず、第1の核酸塩基に相補的な第3の核酸塩基が第2の核酸塩基に相補的な第4の核酸塩基に置換される工程とを含む。

【0454】

いくつかの実施形態において、複数の核酸塩基対は、1つ以上の遺伝子に存在する。あ

10

20

30

40

50

る態様において、複数の核酸塩基対は、同じ遺伝子内にある。いくつかの実施形態において、1つ以上の遺伝子における少なくとも1つの遺伝子は、異なる遺伝子座に位置する。

【0455】

ある態様において、塩基編集は、少なくとも1つのタンパク質コード領域における複数の核酸塩基対の編集を含む。いくつかの実施形態において、塩基編集は、少なくとも1つのタンパク質非コード領域における複数の核酸塩基対の編集を含む。ある態様において、塩基編集は、少なくとも1つのタンパク質コード領域および少なくとも1つのタンパク質非コード領域における複数の核酸塩基対の編集を含む。

【0456】

いくつかの態様において、編集は、1以上のガイドポリヌクレオチドを伴う。いくつかの実施形態では、塩基エディターシステムは、1つ以上の塩基エディターシステムを含むことができる。いくつかの実施形態では、塩基エディターシステムは、単一のガイドポリヌクレオチドとともに1つ以上の塩基エディターシステムを含むことができる。いくつかの実施形態において、塩基エディターシステムは、複数のガイドポリヌクレオチドとともに1つ以上の塩基エディターシステムを含むことができる。いくつかの実施形態において、編集は、単一の塩基エディター系を有する1以上のガイドポリヌクレオチドを伴う。いくつかの実施形態において、編集は、標的ポリヌクレオチド配列への結合を標的化するためにPAM配列を必要としない少なくとも1つのガイドポリヌクレオチドを伴う。いくつかの実施形態において、編集は、標的ポリヌクレオチド配列への結合を標的化するためにPAM配列を必要とする少なくとも1つのガイドポリヌクレオチドを伴う。いくつかの実施形態において、編集は、標的ポリヌクレオチド配列への結合を標的化するためにPAM配列を必要としない少なくとも1つのガイドポリヌクレオチドと、標的ポリヌクレオチド配列への結合を標的化するためにPAM配列を必要とする少なくとも1つのガイドポリヌクレオチドとの混合物を伴う。本明細書に記載される塩基エディターのいずれかを使用する多重編集の特徴は、本明細書に提供される塩基エディターのいずれかを使用する方法の任意の組み合わせに適用され得ることを理解されたい。また、編集は、複数の核酸塩基対の順次的編集を含むことができることを理解されたい。

【0457】

〔塩基エディターを使用する方法〕

疾患関連遺伝子およびアリルにおける点突然変異の補正は、治療および基礎研究への応用を有する遺伝子補正のための新しい戦略を提示し提供する。

【0458】

本開示は、本明細書に提供される塩基エディターシステムによって補正することができる点突然変異に関連するまたはその点突然変異によって引き起こされる疾患と診断された対象の治療方法を提供する。例えば、いくつかの態様において、疾患、例えば遺伝子突然変異によって引き起こされる疾患を有する対象に、疾患表現型を軽減する良性アミノ酸に病原性アミノ酸を置換する核酸塩基エディター(例えばアデノシンデアミナーゼ塩基エディターまたはシチジンデアミナーゼ塩基エディター)の有効量を投与する工程を含む方法が提供される。様々な実施形態において、疾患は、増殖性疾患、遺伝性疾患、腫瘍性疾患、または代謝性疾患である。このような疾患および障害の非限定的な例としては、ヘモグロビン疾患または障害、鎌状赤血球症、 β -サラセミア、 α -1アンチトリプシン欠損症(A1AD)、肝性ポルフィリン症、ACADM欠損症、ペンドレッド症候群、または家族性パーキンソン病が挙げられる。非限定的な例として、鎌状赤血球症の表現型を軽減する良性アミノ酸(Ala)で病原性アミノ酸(Val)を置換する有効量のA $\rightarrow$ G核酸塩基エディター(例えばアデノシンデアミナーゼ塩基エディター)を、鎌状赤血球症を有する対象に投与することを含む方法が提供される。

【0459】

点突然変異を補正するかまたは疾患関連遺伝子に不活性化突然変異を導入することによって治療することができる他の疾患は、当業者に知られており、本開示はこの点において限定されない。本開示は、追加の疾患または障害の治療方法、例えばデアミナーゼ媒介遺

伝子編集によって補正することができる点突然変異に関連する、またはそれによって引き起こされる疾患または障害の治療方法を提供する。いくつかのそのような疾患は本明細書に記載されており、そして本明細書に提供される戦略および融合タンパク質で治療することができるさらなる適切な疾患は、本開示に基づいて当業者に明らかであろう。それぞれの配列における特定の位置または残基の番号付けは、使用される特定のタンパク質および番号付けスキームに依存することが理解されるであろう。例えば成熟タンパク質の前駆体と成熟タンパク質そのものの番号付けは異なることがあり、種間の配列の違いが番号付けに影響することがある。当業者は、任意の相同的タンパク質およびそれぞれのコード核酸配列中のそれぞれの残基を、当技術分野で周知の方法、例えば配列アラインメントおよび相同的残基の決定によって同定することができる。

10

【0460】

本明細書では、疾患または障害に関連する標的ヌクレオチド配列中の核酸塩基を編集するために、塩基エディターまたは塩基エディターシステムを使用する方法が提供される。いくつかの実施形態において、塩基エディター(例えばアデノシンデアミナーゼおよびCas9ドメインを含むもの)の活性は、点突然変異の補正をもたらす。ある態様において、標的DNA配列は、疾患または障害と関連するG→A点突然変異を含み、ここで、突然変異体A塩基の脱アミノ化は、疾患または障害と関連しない配列をもたらす。ある態様において、標的DNA配列は、疾患または障害と関連するT→C点突然変異を含み、ここで、突然変異体C塩基の脱アミノ化は、疾患または障害と関連しない配列をもたらす。

【0461】

20

いくつかの実施形態において、標的DNA配列はタンパク質をコードし、点突然変異はコドン中にあり、野生型コドンと比較して突然変異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。いくつかの実施形態において、突然変異体Aの脱アミノ化は、突然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。いくつかの実施形態において、突然変異体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。いくつかの実施形態において、突然変異体Cの脱アミノ化は、突然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。いくつかの実施形態において、突然変異体C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。ある態様において、対象は、疾患もしくは障害を有するか、または疾患もしくは障害と診断されている。

【0462】

30

ある態様において、本明細書に提供されるアデノシンデアミナーゼは、DNAのデオキシアデノシン残基のアデニンを脱アミノ化することができる。開示の他の態様は、アデノシンデアミナーゼ(例えば、本明細書に記載のDNA中のデオキシアデノシンを脱アミノ化するアデノシンデアミナーゼ)および特定のヌクレオチド配列に結合することができるドメイン(例えばCas9やCpf1タンパク質)を含む融合タンパク質を提供する。例えば、アデノシンは、典型的にはシトシン残基と塩基対を形成するイノシン残基に変換されることができる。このような融合タンパク質は、特に核酸配列の標的化編集に有用である。そのような融合タンパク質は、インビトロでのDNAの標的化された編集のために、例えば、突然変異体の細胞または動物の生成のために；標的化された突然変異の導入のために、例えば、ex vivoの細胞における遺伝的欠陥の補正のために、例えば、対象から得られその後同じまたは別の対象に再導入される細胞において；およびin vivoでの標的突然変異の導入のために使用することができ、例えば、GからAへ、またはTからCへの突然変異による疾患関連遺伝子における遺伝子欠陥の補正または不活性化突然変異の導入は、本明細書に提供される核酸塩基エディターを用いて処理することができる。本開示は、デアミナーゼおよび核酸塩基エディターを利用するデアミナーゼ、融合タンパク質、核酸、ベクター、細胞、組成物、方法、キット、システム等を提供する。

40

【0463】

[HBG1/2の調節領域におけるヌクレオチドを標的とするための核酸塩基エディターの使用]

HBG1/2遺伝子の調節領域内のヌクレオチドを標的とする核酸塩基エディターの適性は

50

、本明細書に記載されるように評価される。一実施形態では、関心対象の単一細胞を、レポーター(例えばGFP)をコードする少量のベクターとともに、本明細書に記載される核酸塩基エディターをコードする核酸分子(複数可)を用いてトランスフェクト、形質導入、またはその他の方法で改変する。これらの細胞は、例えば293T細胞、K562細胞、またはU20S細胞などの不死化ヒト細胞株であり得る。あるいは、初代ヒト細胞、例えばCD34+細胞を使用してもよい。細胞は、対象または個体から、例えば組織生検、外科手術、血液、血漿、血清、または他の生物学的流体などから得ることもできる。そのような細胞は、最終的な細胞標的に関連したものであり得る。

【0464】

送達は、以下にさらに記載されるように、ウイルスベクターを用いて実施され得る。一実施形態において、トランスフェクションは、脂質トランスフェクション(LipofectamineやFugeneなど)を用いて、またはエレクトロポレーションによって実施することができる。トランスフェクション後、GFPの発現を蛍光顕微鏡またはフローサイトメトリーのいずれかによって測定して、一貫した高レベルのトランスフェクションを確認することができる。最も高い活性を与えるエディターの組み合わせを決定するために、これらの予備的なトランスフェクションは異なる複数の核酸塩基エディターを含むことができる。

【0465】

核酸塩基エディターの活性は、本明細書に記載のように、すなわち標的遺伝子を配列決定して標的配列の変化を検出することによって評価される。サンガー配列決定のためには、精製されたPCRアンプリコンをプラスミド骨格中にクローン化し、形質転換し、ミニプレップし、単一のプライマーで配列決定する。配列決定はまた、次世代配列決定技術を用いて実施されてもよい。次世代配列決定を使用する場合、アンプリコンは、意図された切断部位が非対称に配置された300~500 bpであり得る。PCRに続いて、次世代配列決定のアダプターおよびバーコード(たとえば、Illumina multiplexアダプターとインデックス)をアンプリコンの末端に付加し得る(例えば高スループット配列決定(例えばIllumina MiSeq上でのもの)で使用するために)。

【0466】

最初の試験で最大レベルの標的特異的改変を誘導する融合タンパク質を、さらなる評価のために選択し得る。

【0467】

特定の実施形態では、核酸塩基エディターを用いて、目的のポリヌクレオチドを標的化する。一実施形態では、本発明の核酸塩基エディターは、核酸配列、例えば、HBG1/2遺伝子に関連する調節領域の標的核酸配列を標的化するために使用されるガイドRNAと一緒に、適切な細胞(例えば、肝細胞、CD34+細胞などの造血細胞、またはその前駆細胞)に送達され、それによって、遺伝子の非正常または異常な機能または活性を補正または減少させる。

【0468】

いくつかの態様において、塩基エディターは、ガイドRNAによってターゲティングされ、関心のある調節配列に1以上の編集を導入する。いくつかの実施形態において、Mecp2遺伝子に導入される1以上の改変は、下記表4に提示される通りである。

【0469】

[意図する変異の生成]

いくつかの態様において、本明細書において提供される方法の目的は、遺伝子編集を介して機能不全遺伝子の機能を回復させることである。いくつかの実施形態において、機能不全遺伝子の機能は、意図された突然変異を導入することによって回復される。本明細書において提供される核酸塩基編集タンパク質は、例えばヒト細胞培養物中の疾患関連突然変異を補正することによって、インビトロでの遺伝子編集に基づくヒト治療のために検証され得る。本明細書において提供される核酸塩基編集タンパク質、例えば、ポリヌクレオチドプログラム可能なヌクレオチド結合ドメイン(例:Cas9)および核酸塩基編集ドメイン(例えば、アデノシンデアミナーゼドメインまたはシチジンデアミナーゼドメイン)を含む融

10

20

30

40

50

合タンパク質は、任意のAからGまたはCからTへの単一点突然変異を補正するために使用することができる。このことが当業者によって理解されよう。最初のケースでは、変異体AからIへの脱アミノ化が変異を修正し、後者のケースでは、変異体Tと塩基対を形成しているAの脱アミノ化とそれに続く一回の複製が変異を修正する。

【0470】

いくつかの態様において、本開示は、有意な数の非意図的な突然変異（例えば非意図的な点突然変異）を生成することなく、核酸（例えば対象のゲノム内の核酸）において、意図する突然変異（例えば点突然変異）を効率的に生成することができる塩基エディターを提供する。いくつかの実施形態において、意図された突然変異は、意図された突然変異を生じさせるように特に設計されたガイドポリヌクレオチド（例えばgRNA）に結合した特定の塩基エディター（例えばシチジン塩基エディターまたはアデノシン塩基エディター）によって生じる突然変異である。ある態様において、意図される突然変異は、疾患または障害に関連する突然変異である。ある態様において、意図される突然変異は、疾患または障害に関連するアデニン（A）からグアニン（G）への点突然変異である。ある態様において、意図される突然変異は、疾患または障害に関連するシトシン（C）からチミン（T）への点突然変異である。いくつかの実施形態において、意図される突然変異は、遺伝子のコード領域または非コード領域内のアデニン（A）からグアニン（G）への点突然変異である。ある態様において、意図される突然変異は、遺伝子のコード領域または非コード領域内のシトシン（C）からチミン（T）への点突然変異である。いくつかの実施形態において、意図される突然変異は、停止コドン、例えば遺伝子のコード領域内の未成熟停止コドンを生成する点突然変異である。いくつかの実施形態において、意図される突然変異は、停止コドンを除去する突然変異である。

【0471】

いくつかの実施形態において、本明細書に提供され記載される塩基エディターのいずれかは、意図される突然変異対意図されない突然変異の比（例えば、意図される点突然変異：意図されない点突然変異）を1:1よりも大きくすることができる。いくつかの実施形態において、本明細書に提供される塩基エディターのいずれかは、意図される突然変異対意図されない突然変異の比（例えば意図される点突然変異：意図されない点突然変異）を、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、または少なくとも1000:1以上とすることができる。

【0472】

塩基エディター効率の詳細は、国際PCT出願番号PCT/2017/045381 (WO2030/027078) およびPCT/US2016/058344 (WO2017/070632) に記載されており、これらの各々は、その全体が参照により本明細書に組み込まれる。Komor, A.C., et al., “Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage” Nature 533, 420-424 (2016); Gaudelli, N.M., et al., “Programmable base editing of A·T to G·C in genomic DNA without DNA cleavage” Nature 551, 464-471 (2017); および 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” Science Advances 3:eaao4774 (2017)も参照（その全体が参照により本明細書に組み込まれる）。

【0473】

いくつかの実施形態において、少なくとも1つの意図された突然変異の形成は、病原性タンパク質の病原性アミノ酸を、野生型非病原性タンパク質とは異なる良性アミノ酸で置換し、それによって、病原性アミノ酸を良性アミノ酸で置換することにより遺伝的障害を

治療する。本明細書に記載される塩基エディターの多重編集の特徴は、本明細書に提供される塩基エディターを使用する方法の任意の組み合わせに適用され得るを理解されたい。

【0474】

〔良性代替アリルへの病原性アミノ酸置換〕

ある態様において、意図される突然変異は、病原性突然変異または疾患を引き起こす突然変異を良性突然変異に変換することができる突然変異である。病原性突然変異から良性代替アリルへの非限定的な例示的変換は、下記の表3Aおよび3Bに列挙されている。表3Aおよび3Bに示されている良性編集は、野生型に戻るための正確な補正を行う代わりに病原性突然変異を補正する可能性を有する代替的变化を表す。突然変異および他の配列バリエーションを記述する命名法の詳細は、den Dunnen, J.T. and Antonarakis, S.E., “Mutation Nomenclature Extensions and Suggestions to Describe Complex Mutations: A Discussion.” Human Mutation 15:712 (2000)に記載されており、その内容全体が参照により本明細書に組み入れられる。ある態様において、疾患または障害は、 α -1アンチトリプシン欠損症 (A1AD) であり、病原性突然変異は、A1ATをコードするSERPINA1遺伝子にある。一態様において、SERPINA1の病原性突然変異は、E342K (PiZアリル)である。別の実施形態では、SERPINA1の病原性突然変異は、E264V (PiSアリル)である。

【0475】

表3Aは、代表的な疾患遺伝子を示し、病原性アミノ酸から良性代替アリルへの置換が示されている。例として、中鎖アシルCoAデヒドロゲナーゼ (ACADM) 遺伝子における一つ以上の突然変異は、中鎖アシルCoAデヒドロゲナーゼ欠損症に関連しおよび/またはその原因となる。代表的なヒトACADMアミノ酸配列は、UniProtKB参照番号P11310に見出される。SLC26A4；ペンドリンタンパク質をコードする溶質輸送体ファミリー26メンバー4 (PDS) 遺伝子における1以上の突然変異は、ペンドレッド症候群に関連しおよび/またはその原因となる。代表的なヒトペンドリンのアミノ酸配列は、UniProtKB参照番号O43511-1に見出される。 α -シヌクレイン (SNCA) 遺伝子における一つ以上の突然変異は、常染色体優性パーキンソン病に関連し、および/またはその原因となる。代表的なヒト α シヌクレイン (SCNA) アミノ酸配列は、UniProtKB参照番号P37840の下に見出される。

【0476】

特定の実施形態において、SERPINA1遺伝子の5位および7位のA核酸塩基を脱アミノ化して、D341Gアリルを得た。本明細書に記載されるSERPINA1遺伝子配列の塩基編集は、コードされたA1ATタンパク質においてD341G、E342G、E342R、K343E、またはK343G置換をもたらすことができる。一実施形態によると、SERPINA1遺伝子の7位および8位のA核酸塩基が脱アミノ化され、E342Gアリルが得られた。一実施形態によると、E342K病原性突然変異の塩基編集は、E342G良性アリルをもたらした(図3および図4)。一部の実施形態では、塩基編集は、オフターゲット編集をもたらす得る。一実施形態によると、オフターゲット編集は、SERPINA1 E342K (PiZ) アリルのD341Gである(図3および図4)。一実施形態では、塩基エディターによる病原性アミノ酸の置換は、HBB遺伝子によってコードされるHbタンパク質においてE7VからE7Iへの変化をもたらす。一実施形態によると、塩基エディターで病原性アミノ酸を置換すると、HBB遺伝子によってコードされる成熟型 β -グロビンにおけるE6VからE6A (E7VからE7A)の変化がもたらされる(表3B)。

10

20

30

40

50

【表 3 A】

表 3 A. 疾患遺伝子の病原性アミノ酸置換の、良性代替アリルへの変換					
	遺伝子	病原性>代替アリル	塩基エディター	gRNA 標的配列	PAM
1.	<i>ACADM</i>	K329E>K329G	ABE	CAAUGGAAGUUGAACUAGCU	NGA
2.	<i>PDS</i>	T416P>T416F	CBE	GAGCCCUGGAGGAAAGACAC	NGG
3.	<i>SNCA</i>	A30P>A30L	CBE	AGCACCAGGAAAGACAAAAG	NGG
4.	<i>SERPINA1</i>	E342K>E342G	ABE	GACAAGAAAGGGACUGAAGC	NGC

10

【表 3 B】

表 3 B. HBB 遺伝子における病原性アミノ酸置換の、良性代替アリルへの変換					
	遺伝子	病原性>代替アリル	塩基エディター	gRNA 標的配列	PAM
1.	<i>HBB</i>	E6V>E6A	ABE	UCCACAGGAGUCAGAUACAC	NNNRRT* (CATGGT)
2.	<i>HBB</i>	E6V>E6A	ABE	UGAAGAGGUGUCCUCAGUCUA	NNACCA (CGTGGT)
3.	<i>HBB</i>	E6V>E6A	ABE	UCUGAAGAGGUGUCCUCAGUCU	NGC (ACG)
4.	<i>HBB</i>	E26K>E26G	ABE	UGGUAAGGCCUGGGCAGGU	NGG

20

30

【0477】

[遺伝子調節性編集の導入]

いくつかの態様において、本明細書において提供される方法の目的は、ゲノム編集を介して機能不全遺伝子の機能を回復させることである。いくつかの実施形態において、機能不全遺伝子の機能は、意図された突然変異を導入することによって回復される。いくつかの実施形態において、意図される突然変異は、遺伝子の調節配列(例えば遺伝子プロモーターまたは遺伝子リプレッサー)を変化させる突然変異である。ある態様において、意図された突然変異は、遺伝子調節因子編集を導入する突然変異である。特定の遺伝子に関連する遺伝子調節因子編集の非限定的な例示的導入、例えば、胎児ヘモグロビンの遺伝性持続性(HPFH)を伴う知的発達障害および胎児ヘモグロビン定量形質5に関連するBAFクロマチンリモデリング複合体成分(BCL11A)遺伝子；γグロビン遺伝子HBG1およびHBG2；哺乳類ヘム合成経路の律速酵素である赤血球型5-アミノレブリン酸シンターゼ1(ALAS1)；血中でコレステロールを運ぶ低密度リポタンパク質(LDL)粒子に結合し特定のリガンドの受容体媒介エンドサイトーシスに関与する低密度リポタンパク質受容体(LDLR)は、。低密度リポタンパク質受容体が、下記の表4に示されている。表4に示した突然変異および他の配列バリエーションを記述する命名法の詳細は、den Dunnen, J.T. and Antonarakis, S.E., "Mutation Nomenclature Extensions and Suggestions to Describe Complex Mutations: A Discussion." Human Mutation 15:712 (2000)に記述されてお

40

50

り、その全体の内容は参照により本明細書に組み込まれる。非限定的に、ヒトBCL11Aアミノ酸配列はGenBank受託番号ADL14508に見出され、その核酸コード配列はGenBank受託番号GU324937.1に見出される。ヒトHBG1およびHBG2のアミノ酸配列および核酸配列は、それぞれNCBI参照配列番号NM_000559.2およびNB_000184.3の下に見出される。ヒトALAS1アミノ酸および核酸配列は、UniProtKB受付番号Q5JAM2およびNCBI参照配列番号NM_000688.6の下に見出される。ヒトLDLRアミノ酸配列はNCBI受託番号NP_000518.1に見出され、その核酸コード配列はNCBI受託番号NP_000527.4に見出される。

【0478】

いくつかの実施形態において、本明細書に提供される塩基エディターは、遠隔部位に意図された突然変異を導入することができる。遠隔部位は、限定されるものではないが、遺伝子プロモーターおよび/またはエンハンサーおよびエクソンまたはイントロンを含む。ある態様において、意図される突然変異は、遺伝子のスプライシングを変化させる突然変異である。いくつかの実施形態において、遺伝子のスプライシングを変化させる意図された突然変異は、エクソンまたはイントロン内にある。ある態様において、遺伝子のスプライシングを変化させる意図された突然変異は、スプライシング率を減少させる。ある態様において、遺伝子のスプライシングを変化させる意図された突然変異は、スプライシング率を増加させる。ある態様において、遺伝子のプロモーターおよび/またはエンハンサー内の意図された突然変異は、転写を増加させる。ある態様において、転写の増加は、リプレッサータンパク質（単数または複数）による結合を減少/除去する、遺伝子のプロモーターおよび/またはエンハンサー内の意図された突然変異に起因する。ある態様において、転写の増加は、新規の転写活性化タンパク質（複数可）の結合を可能にする、遺伝子のプロモーターおよび/またはエンハンサー内の意図された突然変異に起因する。ある態様において、転写を低下させる。ある態様において、転写の減少は、新規なリプレッサータンパク質の結合を可能にする、遺伝子のプロモーターおよび/またはエンハンサー内の意図された突然変異に起因する。ある態様において、転写の減少は、転写活性化タンパク質による結合を減少/除去する、遺伝子のプロモーターおよび/またはエンハンサー内の意図された突然変異に起因する。

。

10

20

30

40

50

【表 4】

表 4. 遺伝子調節編集の導入					
	遺伝子	ヌクレオチド 変化	塩基エ ディタ ー	gRNA 標的配列	PAM
1.	<i>BCL11A</i>	c.386- 24278G>A	ABE	UGAAAGAAAUUAACACAAA	NGA
2.	<i>BCL11A</i>	c.386-24983T>C	CBE	UUCCUGCACCGAAGCUUUGC	NGT
3.	<i>HBG1</i>	c.-167C>T	CBE	CUUGACCAAUAGCCUUGACA	NGG
4.	<i>HBG1</i>	c.-170G>A	CBE	GCUAUUGGUCAAGGCAAGGC	NGG
5.	<i>HBG1</i>	c.-249C>T	CBE	CUUCCCCACACUAUCUCAAU	NNRR T
6.	<i>HBG2</i>	c.-211C>T	CBE		
7.	<i>HBG2</i>	c.-228T>C	ABE	AUAUUUGCAUUGAGAUAGUG	NGG
8.	<i>ALAS1</i>	c.3G>A	CBE	CUCUCCAUGUUCAGGAAGUA	TGCT
9.	<i>ALAS1</i>	c.2T>C	ABE	CUCUCCAUGUUCAGGAAGUA	TGCT
10.	<i>ALAS1</i>	c.46C>T	CBE	AGUCCLDCCCAGGCCUUUCUG C	AGAA
11.	<i>ALAS1</i>	c.91C>T	CBE	CUAUGCCCCAAAACUGCCCCA	AGAT
12.	<i>ALAS1</i>	c.91C>T	CBE	UGCCCCAAAACUGCCCCAAGA	TGAT
13.	<i>ALAS1</i>	c.226C>T	CBE	AAGGUCCAACAGACUCCUGA	TGG
14.	<i>ALAS1</i>	c.229C>T	CBE	AAGGUCCAACAGACUCCUGA	TGG
15.	<i>ALAS1</i>	c.226C>T	CBE	AGGUCCAACAGACUCCUGAU	GGAT
16.	<i>ALAS1</i>	c.229C>T	CBE	AGGUCCAACAGACUCCUGAU	GGAT
17.	<i>ALAS1</i>	c.247C>T	CBE	GGAUCCCAGCAGAGUCCAGA	TGG
18.	<i>ALAS1</i>	c.250C>T	CBE	GGAUCCCAGCAGAGUCCAGA	TGG
19.	<i>ALAS1</i>	c.247C>T	CBE	GAUCCCAGCAGAGUCCAGAU	GGCA
20.	<i>ALAS1</i>	c.250C>T	CBE	GAUCCCAGCAGAGUCCAGAU	GGCA
21.	<i>ALAS1</i>	c.340C>T	CBE	GCAGCACAGAUGAAUCAGAG	AGG
22.	<i>ALAS1</i>	c.340C>T	CBE	CAGCACAGAUGAAUCAGAGA	GGCA
23.	<i>ALAS1</i>	c.349C>T	CBE	AUGAAUCAGAGAGGCAGCAG	TGTC
24.	<i>ALAS1</i>	c.391C>T	CBE	UGAGCUUCAGGAGGAUGUGC	AGG

10

20

30

40

50

表 4. 遺伝子調節編集の導入					
	遺伝子	ヌクレオチド 変化	塩基エ ディタ ー	gRNA 標的配列	PAM
25.	<i>ALAS1</i>	c.391C>T	CBE	GAGCUUCAGGAGGAUGUGCA	GGAA
26.	<i>ALAS1</i>	c.403C>T	CBE	GAUGUGCAGGAAAUGAAUGC	CGTG
27.	<i>ALAS1</i>	c.403C>T	CBE	UGUGCAGGAAAUGAAUGCCG	TGAG
28.	<i>ALAS1</i>	c.199+1G>A	CBE	CUUACUCUCACUGGCCGGAG	GGG
29.	<i>ALAS1</i>	c.199+2T>C	ABE	CUUACUCUCACUGGCCGGAG	GGG
30.	<i>ALAS1</i>	c.199+1G>A	CBE	ACUUACUCUCACUGGCCGGA	GGG
31.	<i>ALAS1</i>	c.199+2T>C	ABE	ACUUACUCUCACUGGCCGGA	GGG
32.	<i>ALAS1</i>	c.199+1G>A	CBE	CACUUACUCUCACUGGCCGG	AGG
33.	<i>ALAS1</i>	c.199+2T>C	ABE	CACUUACUCUCACUGGCCGG	AGG
34.	<i>ALAS1</i>	c.199+1G>A	CBE	UGACACUUACUCUCACUGGC	CGG
35.	<i>ALAS1</i>	c.199+2T>C	ABE	UGACACUUACUCUCACUGGC	CGG
36.	<i>ALAS1</i>	c.200-1G>A	CBE	UUGUCUUCUGAGGGAGGAAA	TGG
37.	<i>ALAS1</i>	c.200-2A>G	ABE	CUCAGAAGACAAAACUGCUA	AGG
38.	<i>ALAS1</i>	c.427+1G>A	CBE	UCAUCUCUUACCUUCCUCA	CGG
39.	<i>ALAS1</i>	c.427+2T>C	ABE	UCAUCUCUUACCUUCCUCA	CGG
40.	<i>ALAS1</i>	c.1165+1G>A	CBE	CACACUUACCAUCCAUGAA	TGG
41.	<i>ALAS1</i>	c.1165+2T>C	ABE	CACACUUACCAUCCAUGAA	TGG
42.	<i>ALAS1</i>	c.1166-1A>G	ABE	CUCAGGGGCGGUGUGCCCAC	TGG
43.	<i>ALAS1</i>	c.1331-2A>G	ABE	CUCCUCCCAGGCAAAGCCUU	TGG
44.	<i>HBG1/2</i>	c. -198 T>C	ABE	GUGGGGAAGGGGCCCCCAAG	AGG
45.	<i>HBG1/2</i>	c. -198 T>C	ABE	AUUGAGAUAGUGUGGGGAAG	GGG
46.	<i>HBG1/2</i>	c. -198 T>C	ABE	CAUUGAGAUAGUGUGGGGA	GGG
47.	<i>HBG1/2</i>	c. -198 T>C	ABE	GCAUUGAGAUAGUGUGGGGA	AGG
48.	<i>HBG1/2</i>	c. -114 ~ -102 欠 損	CBE and/or ABE	GCUAUUGGUCAAGGCAAGGC	TGG
49.	<i>HBG1/2</i>	c. -114 ~ -102 欠 損	CBE and/or ABE	CAAGGCUAUUGGUCAAGGCA	AGG
50.	<i>HBG1/2</i>	c. -114 ~ -102 欠 損	CBE and/or ABE	CUUGUCAAGGCUAUUGGUCA	AGG

10

20

30

40

50

表 4. 遺伝子調節編集の導入					
	遺伝子	ヌクレオチド 変化	塩基エ ディタ ー	gRNA 標的配列	PAM
51.	<i>HBG1/2</i>	c. -114 ~ -102 欠 損	CBE and/or ABE	CUUGACCAAUAGCCUUGACA	AGG
52.	<i>HBG1/2</i>	c. -114 ~ -102 欠 損	CBE and/or ABE	GUUUGCCUUGUCAAGGCUAU	TGG
53.	<i>HBG1/2</i>	c. -114 ~ -102 欠 損	CBE and/or ABE	UGGUCAAGUUUGCCUUGUCA	AGG
54.	<i>HBG1/2</i>	c. -198 T>C	ABE	UGGGGAAGGGGCCCCCAAGA	GGA
55.	<i>HBG1/2</i>	c. -198 T>C	ABE	GUGUGGGGAAGGGGCCCCCA	AGA
56.	<i>HBG1/2</i>	c. -175 T>C	ABE	UCAGACAGAUUUUGCAUUG	AGA
57.	<i>HBG1/2</i>	c. -175 T>C	ABE	UUUCAGACAGAUUUUGCAU	TGA
58.	<i>HBG1/2</i>	c. -114 ~ -102 欠 損	CBE and/or ABE	CUUGCCUUGACCAAUAGCCU	TGA
59.	<i>HBG1/2</i>	c. -114 ~ -102 欠 損	CBE and/or ABE	UAGCCUUGACAAGGCAAACU	TGA
60.	<i>HBG1/2</i>	c. -90 BCL11A 結合	CBE and/or ABE	CAAACUUGACCAAUAGUCUU	AGA
61.	<i>HBG1/2</i>	c. -198 T>C	ABE	UGUGGGGAAGGGGCCCCCAA	GAGGA T
62.	<i>HBG1/2</i>	c. -202 C>T, - 201 C>T, -198 T>C, -197 C>T, -196 C>T, -195 C>G	CBE and/or ABE	GGGCCCCUCCCCACACUAU	CTCAA T
63.	<i>HBG1/2</i>	c. -197 C>T, - 196 C>T, -195 C>G	CBE	CUUCCCCACACUAUCUCAAU	GCAAA T
64.	<i>HBG1/2</i>	c. -175 T>C	ABE	CAGACAGAUUUUGCAUUGA	GATAG T
65.	<i>HBG1/2</i>	c. -175 T>C	ABE	UUUCAGACAGAUUUUGCAU	TGAGA T
66.	<i>HBG1/2</i>	c. -114 ~ -102 欠 損	CBE	AAGUUUGCCUUGUCAAGGCU	ATTGG T
67.	<i>HBG1/2</i>	c. -114 ~ -102 欠 損	CBE and/or ABE	GCCUUGACAAGGCAAACUUG	ACCAA T
68.	<i>HBG1/2</i>	c. -114 ~ -102 欠 損	CBE and/or	UUGACAAGGCAAACUUGACC	AATAG T

10

20

30

40

50

表 4. 遺伝子調節編集の導入					
	遺伝子	ヌクレオチド 変化	塩基エ ディタ ー	gRNA 標的配列	PAM
		損	ABE		
69.	<i>HBG1/2</i>	c. -90 BCL11A 結合	CBE and/or ABE	UGACCAAUAGUCUUAGAGUA	TCCAG T
70.	<i>HBG1/2</i>	c. -175 T>C	ABE	AGACAGAUUUUGCAUUGAGA UA	TTT
71.	<i>LDLR</i>	c. 81C > T	CBE	CAGAUGCGAAAGAAACGAGU	NNRR T

【0479】

[送達システム]

本明細書に開示される塩基エディターは、ウイルスベクターに含まれる核酸上にコードされ得る。例示的なウイルスベクターとしては、レトロウイルスベクター(例えばMaloney murine leukemia virus, MML-V)、アデノウイルスベクター(例えばAD 100)、レンチウイルスベクター(HIVおよびFIVベースのベクター)、ヘルペスウイルス(例えばHSV-2)、およびアデノ関連ウイルスベクターが挙げられる。

【0480】

[アデノ随伴ウイルスベクター (AAV)]

AAVアデノ随伴ウイルス(「AAV」)ベクターは、標的核酸で細胞を形質導入するために使用することもでき、例えば、核酸およびペプチドのインビトロでの産生において、ならびにインビボおよびエクスピボでの遺伝子治療手順のために使用され得る(例えば、West et al., Virology 160:38-47 (1987); U.S. Patent No. 4,797,368; WO 93/24641; Kotin, Human Gene Therapy 5:793-801 (1994); Muzyczka, J. Clin. Invest. 94:1351 (1994))。組換えAAVベクターの構築は、米国特許第5,173,414号; 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)を含む多くの刊行物に記載されている。

【0481】

in vivo送達に関して、AAVは他のウイルスベクターよりも有利であり得る。場合によっては、AAVは低い毒性を可能にするが、これは、免疫応答を活性化させ得る細胞粒子の超遠心分離を必要としない精製方法に起因し得る。ある場合には、AAVは宿主ゲノムに組み込まれないため、挿入突然変異を引き起こす確率の低さを可能にする。

【0482】

AAVはパルボウイルスファミリーに属する小さな一本鎖DNA依存性ウイルスである。4.7 kbの野生型 (wt) AAVゲノムは、それぞれ四つの複製タンパク質および三つのキャプシドタンパク質をコードする二つの遺伝子からなり、両側に145 bpの逆方向末端反復配列 (ITR) がある。ビリオンは三つのキャプシドタンパク質Vp1、Vp2およびVp3から成り、これらは同じオープンリーディングフレームから1:1:10比で産生されるが、異なるスプライシング (Vp1) および選択的翻訳開始部位(Vp2とVp3のそれぞれ)から産生される。Vp3はビリオン中に最も豊富に存在するサブユニットであり、ウイルスの指向性を規定する細胞表面での受容体認識に関与する。ウイルス感染性において機能するホスホリパーゼドメインがVp1のユニークなN末端に同定されている。

【0483】

AAVのパッケージング限界は4.5または4.75 Kbである。従って、開示された塩基エディター並びにプロモーター及び転写ターミネーターは、単一のウイルスベクターに収容することができる。4.5 Kbまたは4.75 Kbを超える構築物は、ウイルス産生を有意に減少させることができる。例えば、SpCas9は非常に大きく、遺伝子自体が4.1 Kbを超え、AAVに詰め込むことが困難であるため、本開示の実施形態は、従来の塩基エディターよりも短い長さの開示された塩基エディターを利用することを含む。いくつかの例では、塩基エディターは4 kb未満である。開示される塩基エディターは、4.5 kb、4.4 kb、4.3 kb、4.2 kb、4.1 kb、4 kb、3.9 kb、3.8 kb、3.7 kb、3.6 kb、3.5 kb、3.4 kb、3.3 kb、3.2 kb、3.1 kb、3 kb、2.9 kb、2.8 kb、2.7 kb、2.6 kb、2.5 kb、2 kb、または1.5 kb未満であり得る。いくつかの場合において、開示された塩基エディターは、長さが4.5 kb以下である。

10

【0484】

AAVは、AAV1、AAV2、AAV5、またはそれらの任意の組み合わせであり得る。標的とする細胞に関してAAVのタイプを選択することができる；例えば、AAV血清型1、2、5またはハイブリッドキャプシドAAV1、AAV2、AAV5またはそれらの任意の組合せを選択して、脳または神経細胞を標的化することができる；また、心臓組織を標的とするAAV4を選択することができる。AAV8は肝臓への送達に有用である。これらの細胞に関する特定のAAV血清型の表は、Grimm, D. et al, J. Virol. 82: 5887-5911 (2008)に見出され得る。

【0485】

20

組換えAAV (rAAV) は、wt AAVと同様に、ベクター導入遺伝子カセットを挟むシス作用性145 bp ITRを利用し、外来DNAのパッケージングのために最大4.5 kbを提供する。感染後、rAAVは本発明の融合タンパク質を発現することができ、環状の頭・尾コンカテマーの状態のエピソームとして存在することにより、宿主ゲノムに組み込まれることなく存続することができる。このシステムを用いたrAAVの成功例は数多くあるが、in vitroおよびin vivoでは、パッケージング能力が限られているため、遺伝子のコード配列の長さがwt AAVゲノムの長さ以上である場合にAAV媒介遺伝子送達の使用が制限されている。

【0486】

AAVベクターの小さいパッケージング能力は、このサイズを超える多数の遺伝子の送達および/または大きな生理学的調節エレメントの使用を困難にする。これらの課題は、例えば、送達されるタンパク質を2つ以上の断片に分割することによって対処することができ、ここで、N末端断片を分割インテイン-Nに融合し、C末端断片を分割インテイン-Cに融合する。そしてこれらの断片を2つ以上のAAVベクターにパッケージングする。本明細書中で使用される場合、「インテイン」とは、隣接するN末端エクステインおよびC末端エクステイン(例えば、連結されるべき断片)をライゲートする自己スプライシングタンパク質イントロン(例えばペプチド)を指す。異種タンパク質断片を連結するための特定のインテインの使用は、例えば、Wood et al., J. Biol. Chem. 289(21); 14512-9 (2014)に記載されている。例えば、インテインIntNおよびIntCは、別々のタンパク質断片に融合された場合、互いを認識して、自身をスプライスして排出し、それと同時に、融合している上記タンパク質断片の隣接するN-およびC-末端エクステインをライゲートして、それによってそれによって2つのタンパク質断片から完全長タンパク質を再構成する。他の適切なインテインは当業者に明らかであろう。

30

40

【0487】

本発明の融合タンパク質の断片は、長さを変えることができる。ある態様において、タンパク質断片は、長さが2アミノ酸～約1000アミノ酸の範囲である。ある態様において、タンパク質断片は、長さが約5アミノ酸～約500アミノ酸の範囲である。ある態様において、タンパク質断片は、長さが約20アミノ酸～約200アミノ酸の範囲である。ある態様において、タンパク質断片は、長さが約10アミノ酸～約100アミノ酸の範囲である。他の長さの適切なタンパク質断片は、当業者には明らかであろう。

【0488】

50

ある態様において、ヌクレアーゼ(例えばCas9)の一部または断片は、インテインに融合される。ヌクレアーゼは、インテインのN末端またはC末端に融合され得る。ある態様において、融合タンパク質の一部または断片は、インテインに融合され、AAVキャプシドタンパク質に融合される。インテイン、ヌクレアーゼおよびキャプシドタンパク質は、任意の配置(例えば、ヌクレアーゼ-インテイン-キャプシド、インテイン-ヌクレアーゼ-キャプシド、キャプシド-インテイン-ヌクレアーゼなど)と一緒に融合され得る。いくつかの実施形態において、インテインのN末端は融合タンパク質のC末端に融合され、インテインのC末端はAAVキャプシドタンパク質のN末端に融合される。

【0489】

一実施形態では、大きな導入遺伝子発現カセットを二つの別々の半分(5' および3' 末端、またはヘッドおよびテール)に分割することによって二重AAVベクターが生成され、カセットの各半分が単一のAAVベクター(5 kb未満)内にパッケージングされる。次いで、両方の二重AAVベクターによる同一細胞の同時感染に続いて、(1) 5' および3' ゲノム(二重AAV重複ベクター)間の相同組換え(HR); (2) ITRを介した5' 及び3' ゲノム(二重AAVトランスプライシングベクター)の尾・頭コンカテマー化; または (3) それら二つのメカニズムの組み合わせ(二重AAVハイブリッドベクター)により、完全長トランスジーン発現カセットの再構築が達成される。in vivoでの二重AAVベクターの使用は完全長タンパク質の発現をもたらす。二重AAVベクタープラットフォームの使用は、 > 4.7 kbの大きさの導入遺伝子のための効率的で実行可能な遺伝子導入戦略を表す。

【0490】

塩基エディターの送達のためのRNAまたはDNAウイルスベースのシステムの使用は、培養中または宿主中の特定の細胞にウイルスをターゲティングし、ウイルスの積荷を核または宿主細胞ゲノムに輸送する、高度に進化したプロセスを利用する。ウイルスベクターは、培養中の細胞、患者に直接投与することができ(in vivo)、またはそれらを用いて細胞をin vitroで処理し得、改変された細胞を任意で患者に投与することができる(ex vivo)。従来のウイルスベースのシステムは、遺伝子導入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、アデノ随伴及び単純ヘルペスウイルスベクターを含み得る。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の遺伝子導入法では、宿主ゲノムへの組み込みが可能であり、しばしば挿入された導入遺伝子の長期発現をもたらす。さらに、多くの異なる細胞型および標的組織において高い形質導入効率が観察されている。

【0491】

レトロウイルスの指向性は、外来のエンベロープタンパク質を組み込み、標的細胞の潜在的な標的集団を拡大することによって変化させることができる。レンチウイルスベクターは、非分裂細胞を形質導入または感染させることができ、典型的には高いウイルス力価を産生するレトロウイルスベクターである。したがって、レトロウイルス遺伝子導入系の選択は標的組織に依存する。レトロウイルスベクターは、6~10 kbまでの外来配列のパッケージング能力を有するシス作用性の長い末端反復から構成される。最小シス作用性LTRは、ベクターの複製およびパッケージングのために十分であり、次いでそれを用いて治療遺伝子が標的細胞に組み込まれ、永続的な導入遺伝子発現が提供される。広く使用されているレトロウイルスベクターには、マウス白血病ウイルス (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が含まれる(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/US94/05700参照)。

【0492】

レトロウイルスベクター、特にレンチウイルスベクターは、標的細胞への効率的な組み込みのために所定の長さより小さいポリヌクレオチド配列を必要とし得る。例えば、9 kbを超える長さのレトロウイルスベクターは、より小さいサイズのものと比較して低いウイルス力価をもたらす得る。いくつかの態様において、本開示の塩基エディターは、レトロ

ウイルスベクターを介して効率的なパッケージングおよび標的細胞への送達を可能にするのに十分なサイズである。場合によっては、塩基エディターは、ガイド核酸および/またはターゲティング可能なヌクレアーゼ系の他の成分と共に発現された場合でも、効率的なパッケージングおよび送達を可能にするサイズである。

【0493】

一過性発現が好ましい用途では、アデノウイルスベースのシステムを使用することができる。アデノウイルスベクターは、多くの細胞型において非常に高い形質導入効率が可能であり、細胞分裂を必要としない。このようなベクターでは、高い力価および発現レベルが得られている。このベクターは比較的簡単な系で大量に生成できる。

【0494】

従って、本明細書に記載される塩基エディターは、ウイルスベクターを用いて送達される。塩基エディターシステムの1つ以上のコンポーネントは、1つ以上のウイルスベクター上にコードされ得る。例えば、塩基エディターおよびガイド核酸が、単一のウイルスベクター上にコードされ得る。他の場合には、塩基エディターおよびガイド核酸は、異なるウイルスベクター上にコードされる。いずれの場合にも、塩基エディターおよびガイド核酸は、各々、プロモーターおよびターミネーターに作動可能に連結され得る。

【0495】

ウイルスベクター上にコードされる成分の組み合わせは、選択されたウイルスベクターのカーゴサイズ制約によって決定され得る。

【0496】

任意の適切なプロモーターを使用して、塩基エディターおよび、適当な場合には、ガイド核酸の発現を誘導することができる。遍在性発現のために使用することができるプロモーターとしては、CMV、CAG、CBh、PGK、SV40、フェリチン重鎖または軽鎖などが挙げられる。脳または他のCNS細胞での発現については、適切なプロモーターとしては、全てのニューロンについてのシナプシンI、興奮性ニューロンについてのCaMKII α 、GABA作動性ニューロンについてのGAD67もしくはGAD65またはVGATなどが挙げられ得る。肝細胞での発現については、適切なプロモーターとしては、アルブミンプロモーターが挙げられる。肺細胞の発現については、適切なプロモーターはSP-Bを含むことができ、内皮細胞については、適切なプロモーターはICAMを含むことができ、造血細胞については、適切なプロモーターはIFN β またはCD45を含むことができる。骨芽細胞用。適当なプロモーターにはOG-2が含まれる。

【0497】

核酸分子発現をコードする塩基エディターを駆動するために使用されるプロモーターは、AAV ITRを含み得る。これは、ベクター中のスペースを占領してしまい得るさらなるプロモーター要素の必要性を排除するために有利であり得る。解放された追加のスペースは、ガイド核酸または選択マーカーなどの追加の要素の発現を駆動するために使用することができる。ITR活性は比較的弱いので、これは選択したヌクレアーゼの過剰発現による潜在的毒性を低減するために使用できる。

【0498】

いくつかの場合において、本開示の塩基エディターは、同じ核酸分子内で別々のプロモーターが塩基エディターおよび適合性ガイド核酸の発現を駆動することを可能にするのに十分な大きさである。例えば、ベクターまたはウイルスベクターは、塩基エディターをコードする核酸に作動可能に連結された第1のプロモーター、およびガイド核酸に作動可能に連結された第2のプロモーターを含むことができる。

【0499】

ガイド核酸の発現を駆動するために使用されるプロモーターには、U6またはH1などのPol IIIプロモーター、ならびにgRNAアデノ随伴ウイルス (AAV) を発現するためのPol IIプロモーターとイントロンのカセットの使用が含まれ得る。

【0500】

一以上のガイド核酸を伴うまたは伴わない本明細書に記載される塩基エディターは、ア

10

20

30

40

50

デノ随伴ウイルス (AAV)、レンチウイルス、アデノウイルス、または他のプラスミドもしくはウイルスベクター型を用いて、特に、例えば、米国特許第8,454,972号(アデノウイルスの製剤、用量)、米国特許第8,404,658号(AAVの製剤、用量)および米国特許第5,846,946号(DNAプラスミドの製剤、用量)からの製剤および用量、ならびにレンチウイルス、AAVおよびアデノウイルスが関わる臨床試験およびそれに関する出版物からの製剤および用量を用いて送達することができる。例えば、AAVについては、投与経路、製剤化および用量は、米国特許第8,454,972号およびAAVが関わる臨床試験におけるようにすることができる。アデノウイルスについては、投与経路、製剤化および用量は、米国特許第8,404,658号およびアデノウイルスが関わる臨床試験におけるようにすることができる。プラスミド送達については、投与経路、処方および用量は、米国特許第5,846,946号およびプラスミドが関わる臨床試験におけるようにすることができる。投与量は平均的な70 kgの個体(例えば成人男性)に基づいて算出するか、またはそれに外挿することができ、患者、被験者、体重および動物種の異なる哺乳動物について調整することができる。投与の頻度は、年齢、性別、全般的な健康状態、患者又は対象の他の状態及び対処される特定の状態又は症状を含む通常の要因に依存して、医学又は獣医学の実施者(例えば医師、獣医師)の技量の範囲内である。ウイルスベクターを対象組織に注射することができる。細胞型特異的塩基編集のために、塩基エディターおよび任意のガイド核酸の発現を、細胞型特異的プロモーターによって駆動し得る。

【0501】

レンチウイルスは複雑なレトロウイルスであり、有糸分裂細胞と有糸分裂後細胞の両方に感染してその遺伝子を発現する能力をもつ。最も一般的に知られているレンチウイルスはヒト免疫不全ウイルス (HIV) であり、これは広範囲の細胞型を標的とするために他のウイルスのエンベロープ糖タンパク質を使用する。

【0502】

レンチウイルスは以下のように調製できる。pCasES 10 (これはレンチウイルス伝達プラスミド骨格を含む)をクローニングした後、低継代(p=5)のHEK293FT細胞を、T-75フラスコ中に播種し、抗生物質なしで、10%ウシ胎児血清を含むDMEM中でトランスフェクションの前日に50%コンフルエンスにした。20時間後、培地をOptiMEM (無血清) 培地に変え、トランスフェクションを4時間後に行った。10 µgのレンチウイルストランスファープラスミド (pCasES 10) および以下のパッケージングプラスミド：5 µgのpMD2.G (VSV-g偽型)および7.5 µgのpsPAX2 (gag/pol/rev/tat) で細胞をトランスフェクトした。トランスフェクションは4 mLのOptiMEM中でカチオン性脂質送達剤(50 µlのリポフェクタミン2000および100 µlのプラス試薬)を用いて行うことができる。6時間後、培地を10%ウシ胎児血清を含む抗生物質不含DMEMに変更した。これらの方法は細胞培養中に血清を用いたが、無血清法が最適である。

【0503】

レンチウイルスは以下のように精製できる。ウイルス上清を48時間後に回収する。上清をまずデブリから除去し、0.45 µm低タンパク質結合 (PVDF) フィルターを通して濾過する。次いで、これらを24,000 rpmで2時間超遠心機で遠心する。ウイルスペレットを50 µlのDMEM中に4℃で一晩再懸濁し、次いで等分し、直ちに-80℃で凍結する。

【0504】

別の実施形態では、ウマ伝染性貧血ウイルス (EIAV) に基づく最小の非霊長類レンチウイルスベクターも考えられる。別の実施形態では、血管静止タンパク質エンドスタチンおよびアンジオスタチンを発現するウマ伝染性貧血ウイルスに基づくレンチウイルス遺伝子治療ベクター、RetinoStat RTMを、網膜下注射を介して送達することが考えられる。別の実施形態では、自己不活化レンチウイルスベクターの使用が考えられる。

【0505】

この系の任意のRNA、例えば、ガイドRNAまたは塩基エディターをコードするmRNAは、RNAの形態で送達することができる。塩基エディターをコードするmRNAは、インビトロ転写を用いて生成することができる。例えば、ヌクレアーゼmRNAは、以下の要素を含

むPCRカセットを用いて合成することができる：T7プロモーター、任意でコザック配列（GCCACC）、ヌクレアーゼ配列、およびβグロビン-ポリAテールからの3' UTRのような3' UTR。このカセットは、T7ポリメラーゼによる転写のために用いることができる。ガイドポリヌクレオチド（例えばgRNA）もまた、インビトロ転写を用いて、T7プロモーター、次いで配列「GG」、およびガイドポリヌクレオチド配列を含有するカセットから転写され得る。

【0506】

発現を増強し、可能性のある毒性を低減するために、塩基エディターコード配列および/またはガイド核酸は、例えば擬似Uまたは5-メチル-Cを用いて、1以上の修飾ヌクレオチドを含むように修飾され得る。

【0507】

本開示はいくつかの態様において、細胞または生物を改変する方法を含む。細胞は、原核細胞または真核細胞であり得る。細胞は哺乳動物細胞であり得る。哺乳動物細胞は、ヒト、非ヒト霊長類、ウシ、ブタ、齧歯類またはマウス細胞であり得る。本開示の塩基エディター、組成物および方法によって細胞に導入される改変は、細胞および細胞の子孫が、タンパク質、抗体、澱粉、アルコールまたは他の所望の細胞アウトプットなどの生物学的産物の生産を改善するために改変されるようなものであり得る。本開示の方法によって細胞に導入される改変は、細胞およびその子孫が、産生される生物学的産物、例えば疾患関連タンパク質産物を変化させる改変を含むようなものであり得る。

【0508】

システムは、1つまたは複数の異なるベクターを含むことができる。一態様では、塩基エディターは、所望の細胞型、好ましくは真核細胞、好ましくは哺乳動物細胞またはヒト細胞の発現のためにコドン最適化される。

【0509】

一般に、コドン最適化とは、目的の宿主細胞における発現を増強するために、天然配列の少なくとも一つのコドン（例えば約1個、2個、3個、4個、5個、10個、15個、20個、25個、50個以上のコドン）を、天然アミノ酸配列を維持しながら、その宿主細胞の遺伝子においてより頻繁にまたは最も頻繁に使用されるコドンで置換することによって、核酸配列を改変するプロセスをいう。様々な種は特定のアミノ酸の特定のコドンについて特定の偏りを示す。コドンバイアス（生物間のコドン利用の違い）は、メッセンジャーRNA（mRNA）の翻訳効率としばしば相関し、それは特に、翻訳されるコドンの性質および特定のトランスファーRNA（tRNA）分子の利用可能性に依存すると考えられている。細胞内の選択されたtRNAの優位性は、一般にペプチド合成で最も頻繁に使われるコドンを反映している。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に合わせて調整することができる。コドン使用表は、例えば、「コドン使用データベース」www.kazusa.or.jp/codon/（2002年7月9日に訪問した）で容易に入手可能であり、これらの表は、多くの方法で適合させることができる。Nakamura, Y., et al. "Codon usage tabulated from the international DNA sequence databases: status for the year 2000" Nucl. Acids Res. 28:292 (2000)参照。例えばGene Forge (Aptagen; Jacobus, Pa.)のような、特定の宿主細胞における発現のために特定の配列をコドン最適化するためのコンピュータアルゴリズムも利用可能である。ある態様において、操作されたヌクレアーゼをコードする配列中の一つ以上のコドン（例えば1、2、3、4、5、10、15、20、25、50以上、または全てのコドン）は、特定のアミノ酸について最も頻繁に使用されるコドンに対応する。

【0510】

パッケージング細胞は、典型的には、宿主細胞に感染し得るウイルス粒子を形成するために使用される。そのような細胞は、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psi.2細胞またはPA317細胞を含む。遺伝子治療に使用されるウイルスベクターは、通常、核酸ベクターをウイルス粒子にパッケージする細胞株を作製することによって生成される。ベクターは、典型的には、パッケージング

10

20

30

40

50

およびその後の宿主への組み込みに必要な最小のウイルス配列を含み、他のウイルス配列は、発現されるべきポリヌクレオチドのための発現カセットによって置き換えられる。欠けているウイルス機能は、典型的にはパッケージング細胞株によってトランスに供給される。例えば、遺伝子治療に使用されるAAVベクターは、典型的には、パッケージングおよび宿主ゲノムへの組み込みに必要なAAVゲノムからのITR配列のみを有する。ウイルスDNA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む細胞株においてパッケージングされ得る。細胞株は、ヘルパーとしてのアデノウイルスによっても感染され得る。ヘルパーウイルスは、AAVベクターの複製およびヘルパープラスミドからのAAV遺伝子の発現を促進することができる。ある場合には、ヘルパープラスミドは、ITR配列の欠如のために、有意な量ではパッケージングされない。アデノウイルスによる汚染は、例えばAAVよりもアデノウイルスの方が感受性である熱処理によって減少させることができる。

【0511】

〔塩基エディターの非ウイルス性送達〕

塩基エディターのための非ウイルス性送達アプローチも利用可能である。非ウイルス核酸ベクターの1つの重要なカテゴリーは、有機または無機であり得るナノ粒子である。ナノ粒子は当技術分野でよく知られており、任意の適切なナノ粒子設計を用いて、ゲノム編集システム構成要素またはそのような構成要素をコードする核酸を送達することができる。例えば、有機(例えば、脂質および/またはポリマー)ナノ粒子は、本開示の特定の実施形態における送達ビヒクルとしての使用に好適であり得る。ナノ粒子製剤、および/または遺伝子導入において使用するための例示的な脂質を、表5に示す(下記)。

10

20

30

40

50

【表 5】

遺伝子移入のために用いられる脂質		
脂質	略語	特徴
1,2-Dioleoyl-sn-glycero-3-phosphatidylcholine	DOPC	Helper
1,2-Dioleoyl-sn-glycero-3-phosphatidylethanolamine	DOPE	Helper
Cholesterol		Helper
N-[1-(2,3-Dioleoyloxy)propyl]N,N,N-trimethylammonium chloride	DOTMA	Cationic
1,2-Dioleoyloxy-3-trimethylammonium-propane	DOTAP	Cationic
Diocetadecylamidoglycylspermine	DOGS	Cationic
N-(3-Aminopropyl)-N,N-dimethyl-2,3-bis(dodecyloxy)-1-propanaminium bromide	GAP-DLRIE	Cationic
Cetyltrimethylammonium bromide	CTAB	Cationic
6-Lauroxyhexyl ornithinate	LHON	Cationic
1-(2,3-Dioleoyloxypropyl)-2,4,6-trimethylpyridinium	2Oc	Cationic
2,3-Dioleoyloxy-N-[2(sperminecarboxamido-ethyl)-N,N-dimethyl-1-propanaminium trifluoroacetate	DOSPA	Cationic
1,2-Dioleoyl-3-trimethylammonium-propane	DOPA	Cationic
N-(2-Hydroxyethyl)-N,N-dimethyl-2,3-bis(tetradecyloxy)-1-propanaminium bromide	MDRIE	Cationic
Dimyristoxypropyl dimethyl hydroxyethyl ammonium bromide	DMRI	Cationic
3β-[N-(N',N'-Dimethylaminoethane)-carbamoyl]cholesterol	DC-Chol	Cationic
Bis-guanidium-tren-cholesterol	BGTC	Cationic
1,3-Dioleoyl-2-(6-carboxy-spermyl)-propylamide	DOSPER	Cationic
Dimethyloctadecylammonium bromide	DDAB	Cationic
Diocetadecylamidoglycylspermidin	DSL	Cationic
rac-[1-(2,3-Dioctadecyloxypropyl)(2-hydroxyethyl)]-dimethylammonium chloride	CLIP-1	Cationic
rac-[2(2,3-Dihexadecyloxypropyl-oxy-methyloxy)ethyl]trimethylammonium bromide	CLIP-6	Cationic
Ethyl-dimyristoylphosphatidylcholine	EDMPC	Cationic
1,2-Distearoyloxy-N,N-dimethyl-3-aminopropane	DSDMA	Cationic
1,2-Dimyristoyl-trimethylammonium propane	DMTAP	Cationic
O,O'-Dimyristyl-N-lysyl aspartate	DMKE	Cationic
1,2-Distearoyl-sn-glycero-3-ethylphosphocholine	DSEPC	Cationic
N-Palmitoyl D-erythro-sphingosyl carbamoyl-spermine	CCS	Cationic
N-t-Butyl-N0-tetradecyl-3-tetradecylaminopropionamidin	diC14-amidin	Cationic
Octadecenoloxo[ethyl-2-heptadecenyl-3 hydroxyethyl]imidazolinium chloride	DOTIM	Cationic
N1 -Cholesterylloxycarbonyl-3,7-diazanonane-1,9-diamine	CDAN	Cationic
2-(3-[Bis(3-amino-propyl)-amino]propylamino)-N-ditetradecylcarbamoylme-ethyl-acetamide	RPR209120	Cationic
1,2-dilinoleoyloxy-3-dimethylaminopropane	DLinDMA	Cationic
2,2-dilinoleyl-4-dimethylaminoethyl-[1,3]-dioxolane	DLin-KC2-	Cationic

遺伝子移入のために用いられる脂質

脂質	略語	特徴
dilinoleyl-methyl-4-dimethylaminobutyrate	DMA DLin-MC3- DMA	Cationic

表6は、遺伝子導入および/またはナノ粒子製剤における使用のための例示的ポリマーを
 列挙する。

【表 6】

遺伝子移入のために用いられるポリマー		
ポリマー	略語	
Poly(ethylene)glycol	PEG	10
Polyethylenimine	PEI	
Dithiobis (succinimidylpropionate)	DSP	
Dimethyl-3,3'-dithiobispropionimide	DTBP	
Poly(ethylene imine)biscarbamate	PEIC	
Poly(L-lysine)	PLL	
Histidine modified PLL		
Poly(N-vinylpyrrolidone)	PVP	20
Poly(propylenimine)	PPI	
Poly(amidoamine)	PAMAM	
Poly(amidoethylenimine)	SS-PAEI	
Triethylenetetramine	TETA	
Poly(β -aminoester)		
Poly(4-hydroxy-L-proline ester)	PHP	30
Poly(allylamine)		
Poly(α -[4-aminobutyl]-L-glycolic acid)	PAGA	
Poly(D,L-lactic-co-glycolic acid)	PLGA	
Poly(N-ethyl-4-vinylpyridinium bromide)		
Poly(phosphazene)s	PPZ	
Poly(phosphoester)s	PPE	50
Poly(phosphoramidate)s	PPA	
Poly(N-2-hydroxypropylmethacrylamide)	pHPMA	
Poly (2-(dimethylamino)ethyl methacrylate)	pDMAEMA	
Poly(2-aminoethyl propylene phosphate)	PPE-EA	
Chitosan		
Galactosylated chitosan		
N-Dodacylated chitosan		
Histone		
Collagen		
Dextran-spermine	D-SPM	

表7は、本明細書で記述される融合タンパク質をコードするポリヌクレオチドのための
 送達法を要約する。

40

50

【表 7】

送達	ベクター／モード	非分裂細胞への送達	発現の期間	ゲノム挿入	送達される分子の種類
物理的	例えば電気穿孔、パーティクルガン、リン酸カルシウムトランスフェクション	YES	一過性	NO	核酸およびタンパク質
ウイルス性	レトロウイルス	NO	安定	YES	RNA
	レンチウイルス	YES	安定	YES／NOだが改変を伴う	RNA
	アデノウイルス	YES	一過性	NO	DNA
	アデノ随伴ウイルス (AAV)	YES	安定	NO	DNA
	ワクシニアウイルス	YES	高度に一過性	NO	DNA
非ウイルス性	単純ヘルペスウイルス	YES	安定	NO	DNA
	カチオン性リポソーム	YES	一過性	送達されるものによる	核酸およびタンパク質
	ポリマーナノ粒子	YES	一過性	送達されるものによる	核酸およびタンパク質
生物学的非ウイルス性	弱毒化細菌	YES	一過性	NO	核酸
送達媒体	遺伝子操作バクテリオファージ	YES	一過性	NO	核酸
	哺乳類ウイルス様粒子	YES	一過性	NO	核酸
	生物学的リポソーム：赤血球ゴーストおよびエキソソーム	YES	一過性	NO	核酸

【0512】

別の局面において、例えばCas9またはそのバリエーションなどの核酸結合タンパク質および目的のゲノム核酸配列を標的とするgRNAのような、ゲノム編集システム構成要素またはそのような構成要素をコードする核酸の送達は、リボ核酸タンパク質 (RNP) を細胞に送達することによって達成され得る。RNPは、標的化gRNAと複合体を形成した核酸結合タンパク質、例えば、Cas9を含む。RNPは、エレクトロポレーション、ヌクレオフェクション、またはカチオン性脂質媒介方法、例えば、Zuris, J.A. et al., 2015, Nat. Biotechnology, 33(1):73-80によって報告されているもののような、公知の方法を用いて細胞に送達することができる。RNPは、CRISPR塩基編集システムにおける使用のために有利であ

り、特に一次細胞のようなトランスフェクトが困難な細胞のために有利である。さらに、RNPは、細胞におけるタンパク質発現で起こり得る困難を軽減することもでき、特に、CRISPRプラスミドにおいて使用され得るCMVまたはEF1Aなどの真核生物プロモーターがよく発現されない場合にはそうである。有利なことに、RNPの使用は、細胞への外来DNAの送達を必要としない。さらに、核酸結合タンパク質およびgRNA複合体を含むRNPは、経時的に分解されるので、RNPの使用は、オフターゲット効果を制限する可能性を有する。プラスミドベースの技術の場合と同様の方法で、RNPは、結合タンパク質(例えばCas9バリエーション)を送達するために、そして相同性依存修復(HDR)を導くために使用され得る。一実施形態において、RNP送達は、本明細書に記載されるように、鎌状赤血球症(SCD)などの血液疾患の治療に関連する塩基編集のための、細胞への核酸結合タンパク質およびgRNAの送達に適している。

10

【0513】

〔医薬組成物〕

本開示の他の態様は、本明細書に記載される塩基エディター、融合タンパク質、または融合タンパク質-ガイドポリヌクレオチド複合体のいずれかを含む医薬組成物に関する。用語「医薬組成物」は、本明細書中で使用される場合、薬学的使用のために処方される組成物を指す。いくつかの態様において、薬学的組成物は、薬学的に許容される担体をさらに含む。ある態様において、薬学的組成物は、さらなる剤(例えば特異的送達、半減期の延長のためのもの、または他の治療化合物)を含む。

【0514】

20

ここで使用されるところの用語「薬学的に許容される担体」は、液体または固体の充填剤、希釈剤、賦形剤、製造助剤(例えば潤滑剤、タルクマグネシウム、ステアリン酸カルシウムまたはステアリン酸亜鉛、あるいはステアリン酸)、または溶媒カプセル化材料などの、身体のある部位(例えば送達部位)から別の部位(例えば器官、組織または身体の一部)への化合物の運搬または輸送に参与する薬学的に許容される材料、組成物またはビヒクルを意味する。薬学的に許容される担体は、製剤の他の成分と適合し、対象の組織に有害でないという意味で許容される(例えば生理的適合性、無菌性、生理的pH等)。

【0515】

薬学的に許容される担体として役立つことができる物質のいくつかの例は、以下を含む:
(1) ラクトース、グルコースおよびスクロースのような糖;(2) コーンスターチ、ジャガイモでん粉等のでん粉;(3) セルロース及びその誘導体(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、酢酸セルロース等);(4) トラガント末;(5) モルト;(6) ゼラチン;(7) ステアリン酸マグネシウム、ラウリル硫酸ナトリウム、タルク等の潤滑剤;(8) ココアバター、坐剤用ワックス等の添加剤;(9) 落花生油、綿実油、ベニバナ油、ゴマ油、オリーブ油、コーン油、大豆油等の油;(10) プロピレングリコール等のグリコール;(11) ポリオール、例えばグリセリン、ソルビトール、マンニトール及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など;(13) 寒天;(14) 水酸化マグネシウム、水酸化アルミニウム等の緩衝剤;(15) アルギン酸;(16) パイロジェンフリー水;(17) 生理食塩液;(18) リンガー液;(19) エチルアルコール;(20) pH緩衝液;(21) ポリエステル、ポリカーボネート及び／又はポリ無水物;(22) ポリペプチドおよびアミノ酸のような増量剤(23) エタノールのような血清アルコール;及び(23) 製剤に使用される他の非毒性適合性物質。湿潤剤、着色剤、離型剤、コーティング剤、甘味剤、風味料、香料、保存剤および抗酸化剤も製剤中に存在させることができる。「賦形剤」、「担体」、「薬学的に許容される担体」などの用語は、本明細書では交換可能に使用される。

30

40

【0516】

薬学的組成物は、約5.0~約8.0の範囲などの生理学的pHを反映する所定のレベルに製剤のpHを維持するために、一つ以上のpH緩衝化合物を含むことができる。水性液体製剤で使用されるpH緩衝化合物は、アミノ酸またはヒスチジンなどのアミノ酸の混合物、またはヒスチジンおよびグリシンなどのアミノ酸の混合物であり得る。あるいは、pH緩衝化

50

物は、製剤のpHを所定のレベル、例えば約5.0～約8.0の範囲に維持し、カルシウムイオンをキレートしない剤であることが好ましい。このようなpH緩衝化合物の典型的な例としては、イミダゾールおよび酢酸イオンが挙げられるが、これらに限定されない。pH緩衝化合物は、製剤のpHを所定のレベルに維持するのに適した任意の量で存在し得る。

【0517】

薬学的組成物はまた、一つ以上の浸透圧調節剤、すなわち、処方物の浸透圧特性(例えば、等張性、オスモラリティ、および／または浸透圧)をレシipient個体の血流および血液細胞にとって許容可能なレベルに調節する化合物を含有することができる。浸透圧調節剤は、カルシウムイオンをキレートしない薬剤であり得る。浸透圧調節剤は、製剤の浸透圧特性を調節する当業者に公知または入手可能な任意の化合物であり得る。当業者は、本発明の処方における使用のための所定の浸透圧調節剤の適合性を経験的に決定することができる。適切なタイプの浸透圧調節剤の例示的な例としては、塩化ナトリウムおよび酢酸ナトリウムのような塩；スクロース、デキストロース、マンニトールなどの糖；グリシンなどのアミノ酸；これらの薬剤および／または薬剤タイプの1つ以上の混合物が挙げられるが、これらに限定されない。浸透圧調節剤は、製剤の浸透圧特性を調節するのに十分な任意の濃度で存在し得る。

【0518】

いくつかの態様において、薬学的組成物は、対象への送達のために、例えば遺伝子編集または塩基編集のために製剤化される。本明細書に記載の医薬組成物を投与する適切な経路は、限定されるものではないが、局所、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、経鼓膜、器官内、硬膜外、髄腔内、筋肉内、静脈内、血管内、骨内、眼周囲、腫瘍内、脳内、および脳室内投与を含む。

【0519】

いくつかの態様において、本明細書に記載される薬学的組成物は、患部(例えば腫瘍部位)に局所的に投与される。いくつかの実施形態において、本明細書に記載される薬学的組成物は、注射、カテーテル、坐剤、またはインプラントによって、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、例えば、シアラスティック膜、または繊維などの膜を含む。

【0520】

他の態様において、本明細書に記載される薬学的組成物は、制御放出システムにおいて送達される。一実施形態では、ポンプを使用することができる(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: 574参照)。別の実施形態では、ポリマー材料を使用することができる。(例えばMedical Applications of Controlled Release (Langer and Wise eds., CRC Press, Boca Raton, Fla., 1974); Controlled Drug Bioavailability, Drug Product Design and Performance (Smolen and Ball eds., Wiley, New York, 1984); Ranger and Peppas, 1983, Macromol. Sci. Rev. Macromol. Chem. 23:61. See also Levy et al., 1985, Science 228: 190; During et al., 1989, Ann. Neurol. 25:351; Howard et al., 1989, J. Neurosurg. 71: 105.)。他の制御放出システムは、例えば、上記Langerに記載されている。

【0521】

いくつかの態様において、薬学的組成物は、対象、例えばヒトへの静脈内または皮下投与に適合された組成物として、ルーチンの手順に従って製剤化される。いくつかの態様において、注射による投与のための医薬組成物は、可溶化剤としての無菌等張使用の溶液および注射部位の痛みを緩和するためのリグノカインなどの局所麻酔薬である。一般に、成分は、単位投与形態、例えば、活性剤の量を示すアンプルまたはサシエットのような密閉容器内の乾燥凍結乾燥粉末または水を含まない濃縮物として、別々にまたは一緒に供給される。薬剤が注入によって投与される場合には、無菌の医薬グレードの水または生理食塩水を含む注入ボトルを用いてそれを分注することができる。医薬組成物が注射によって投

与される場合、投与前に成分を混合することができるように、注射用滅菌水または生理食塩水のアンプルを提供することができる。

【0522】

全身投与のための医薬組成物は、液体、例えば滅菌生理食塩水、乳酸化リンゲル液またはハंक液であり得る。さらに、医薬組成物は、固体形態であって、使用の直前に再溶解または懸濁され得る。凍結乾燥形態もまた考えられる。医薬組成物は、非経口投与にも適した、リポソームまたは微結晶などの脂質粒子または小胞内に含有することができる。粒子は、組成物がその中に含まれる限り、単ラメラまたは複数ラメラのような任意の適切な構造であり得る。化合物は、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)、低量(5-10モル%)のカチオン性脂質を含み、ポリエチレングリコール(PEG)コーティングにより安定化された、「安定化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)。このような粒子および小胞には、N-[1-(2,3-ジオレオイルキシ)プロピル]-N,N,N-トリメチル-アモニウムメチル硫酸塩、あるいは「DOTAP」のような正電荷脂質が特に好ましい。このような脂質粒子の調製はよく知られている。例えば、米国特許第4,880,635; 4,906,477; 4,911,928; 4,917,951; 4,920,016; および4,921,757号参照(その各々は、参照により本明細書に組み込まれる)。

10

【0523】

本明細書に記載の医薬組成物は、例えば、単位用量として投与または包装することができる。「単位用量」という用語は、本開示の薬学的組成物に関して使用される場合、対象のための単一用量として適した物理的に個別の単位を指し、各単位は、必要な希釈剤(例えば担体、またはビヒクル)と合わせて所望の治療効果を生じるように計算された所定量の活性物質を含有する。

20

【0524】

さらに、この薬学的組成物は、(a) 凍結乾燥形態の本発明の化合物を含有する容器、および(b) 本発明の凍結乾燥化合物の再構成または希釈のために使用される、薬学的に許容される希釈剤(例えば無菌のもの)を含有する第2の容器を含む薬学的キットとして提供することができる。必要に応じて、そのような容器には、医薬品又は生物学的製剤の製造、使用又は販売を規制する政府機関によって規定された様式の通知であって、人に投与するための製造、使用又は販売の機関による承認を反映するものとすることができる。

30

【0525】

いくつかの態様において、本明細書に記載される融合タンパク質、gRNA、および/または複合体のいずれかは、薬学的組成物の一部として提供される。いくつかの態様において、薬学的組成物は、本明細書に提供される融合タンパク質のいずれかを含む。いくつかの態様において、薬学的組成物は、本明細書に提供される複合体のいずれかを含む。いくつかの態様において、薬学的組成物は、gRNAおよびカチオン性脂質と複合体を形成するRNA-ガイドヌクレアーゼ(例えばCas9)を含むリボ核タンパク質複合体を含む。ある態様において、薬学的組成物は、gRNA、核酸プログラム可能DNA結合タンパク質、カチオン性脂質、および薬学的に許容される賦形剤を含む。薬学的組成物は、場合により、1つ以上の追加の治療活性物質を含むことができる。

40

【0526】

種々の動物への投与に適した組成物を与えるための、ヒトへの投与に適した医薬組成物の改変は十分に理解されており、通常の熟練した獣医薬理学者は、もし必要だとしても単に通常の実験で、そのような改変を設計および/または実施することができる。薬学的組成物の投与が意図される対象には、限定されるものではないが、ヒトおよび/または他の霊長類; 哺乳動物、家畜、ペット、およびウシ、ブタ、ウマ、ヒツジ、ネコ、イヌ、マウス、および/またはラットなどの商業的に関連のある哺乳動物; ニワトリ、カモ、ガチョウ及び/又は七面鳥のような商業的に関連のある鳥類が含まれる。

【0527】

本明細書に記載される薬学的組成物の製剤は、薬学の分野において公知の、または今後

50

開発される任意の方法によって調製することができる。一般に、このような調製方法は、活性成分を賦形剤および/または1つ以上の他の補助成分と会合させ、次いで、必要および/または所望であれば、製品を所望の単回または複数回投与単位に成形および/または包装する工程を含む。薬学的製剤はさらに、薬学的に許容される賦形剤を含むことができ、それは、本明細書で使用される場合、所望の特定の剤形に適した、溶媒、分散媒体、希釈剤、または他の液体ビヒクル、分散または懸濁助剤、界面活性剤、等張剤、増粘剤または乳化剤、保存剤、固体結合剤、潤滑剤などのいずれかおよび全てを含む。Remington's The Science and Practice of Pharmacy, 21st Edition, A. R. Gennaro (Lippincott, Williams & Wilkins, Baltimore, MD, 2006 (参照によりその全体が本明細書に組み込まれる))は、薬学的組成物を製剤化する際に使用される種々の賦形剤およびその調製のための公知の技術を開示する。ヌクレアーゼを含む医薬組成物を製造するためのさらなる適切な方法、試薬、賦形剤および溶媒については、参照によりその全体が本明細書に組み込まれるPCT出願PCT/US2010/055131(2010年11月2日出願、公開番号WO2011053982 A8)も参照のこと。

10

【0528】

あらゆる従来の賦形剤媒体は、望ましくない生物学的効果を生じさせること、または他のかたちで医薬組成物の何らかの他の成分と有害な様式で相互作用することなどによって、物質またはその誘導体と不適合である場合を除き、その使用が本開示の範囲内にあると考えられる。

【0529】

上記の組成物は、有効量で投与することができる。有効量は、投与方法、治療される特定の状態、および所望の結果に依存する。それはまた、状態のステージ、対象の年齢および身体的状態、もしあれば併用療法の性質、および医師によく知られた同様の因子に依存し得る。治療用途のためには、それは医学的に望ましい結果を達成するのに十分な量である。

20

【0530】

[病理学的突然変異に関連する疾患の治療方法]

本明細書(例えば上記表3Aおよび3B)に記載される疾患関連または疾患原因遺伝子、または、例えば上記表4に列挙されるような、突然変異を有する遺伝子に関連する調節配列(例えば遺伝子プロモーター、エンハンサー、リプレッサーなど)への塩基編集の導入が関わる、疾患または障害を治療する方法も提供される。

30

【0531】

本方法は、対象(例えばヒト等の哺乳動物)に、本明細書に記載の塩基エディター系(例えば、塩基エディターとgRNA)をコードするポリヌクレオチドを含む薬学的組成物の治療有効量を投与することを含む。ある態様において、塩基エディターは、ポリヌクレオチドプログラム可能なDNA結合ドメインおよびアデノシンデアミナーゼドメインまたはシチジンデアミナーゼドメインを含む融合タンパク質である。対象の細胞は、疾患関連遺伝子、疾患原因遺伝子、または疾患原因遺伝子に関連する調節核酸配列のA・TからG・Cへの変化(細胞がアデノシンデアミナーゼドメインで形質導入される場合)またはC・GからU・Aへの変化(細胞がシチジンデアミナーゼドメインで形質導入される場合)をもたらすために、塩基エディターおよびそれをターゲティングする一つ以上のガイドポリヌクレオチドとともに形質導入される。

40

【0532】

本明細書に記載の方法は、本明細書に記載の組成物の有効量を対象(そのような治療を必要としていると特定された対象、又は疾患のリスクがありそのような治療を必要としていると疑われる対象を含む)に投与することを含む。このような治療を必要とする対象を同定することは、対象または医療専門家の判断にあり得、主観的(例えば所見)または客観的(例えば、検査または診断方法によって測定可能)であり得る。

【0533】

治療方法は、一般的に、例えば、それを必要とする対象(例えば、ヒト患者)の疾患原因

50

遺伝子、疾患関連遺伝子(例えば、上記の表3Aおよび3Bに示すもの)、または例えば表4に列挙された疾患遺伝子に関連する調節配列(遺伝子プロモーター、エンハンサー、またはリプレッサーなど)を標的とする塩基エディターおよびgRNAをコードするベクターを含む薬学的組成物の治療有効量の投与を含む。このような治療は、疾患または障害を患っている、それを有している、それに罹患しやすい、またはその罹患のリスクがある対象、特にヒト対象に適切に投与される。本明細書中の組成物はまた、記載された遺伝子または記載された遺伝子の調節配列が関与し得る任意の他の障害の治療にも使用され得る。

【0534】

一実施形態では、治療の進行を監視する方法が提供される。本方法は、疾患に関連する特定の遺伝子に関連する疾患またはその症状に罹患しているかまたは罹患しやすい対象において、診断マーカー(マーカー)(例えば、本明細書中に記載される疾患関連遺伝子に関連するSNP)または診断測定(例えば、スクリーニング、アッセイ)のレベルを決定する工程を含み、ここでその対象は、疾患またはその症状を治療するのに十分な本明細書中の組成物の治療量を投与された対象である。この方法で測定されたマーカーのレベルを、健康な正常対照または他の罹患患者のいずれかにおけるマーカーの既知のレベルと比較して、対象の疾患状態を確立することができる。好ましい態様において、第1のレベルの決定よりも後の時点で、対象におけるマーカーの第2のレベルが決定され、疾患の経過または治療の有効性をモニターするためにこの2つのレベルが比較される。特定の好ましい実施形態では、対象における治療前のマーカーのレベルが、本発明による治療を開始する前に決定される。このマーカーの治療前レベルを治療開始後の対象のマーカーレベルと比較して、治療の有効性を判定することができる。

【0535】

ある態様において、細胞が対象から取得され、本明細書に提供される薬学的組成物と接触される。いくつかの態様において、対象から取り出されてex vivoで薬学的組成物と接触された細胞は、場合により、所望のゲノム改変が細胞において実行または検出された後に、対象に再導入される。ヌクレアーゼを含む薬学的組成物を送達する方法は、例えば、米国特許第6,453,242; 6,503,717; 6,534,261; 6,599,692; 6,607,882; 6,689,558; 6,824,978; 6,933,113; 6,979,539; 7,013,219; and 7,163,824号に記載されており、その開示全体が参照により本明細書に組み込まれる。本明細書に提供される薬学的組成物の説明は、主に、ヒトへの投与に適した薬学的組成物に向けられているが、そのような組成物は、一般に、あらゆる種類の動物または生物への投与、例えば、獣医学的使用に適していることが、当業者によって理解されるであろう。

【0536】

[キット]

本開示の様々な態様は、塩基エディターシステムを含むキットまたは製品を提供する。一実施形態では、キットまたは製品は、核酸塩基エディター融合タンパク質をコードするヌクレオチド配列を含む核酸構築物を含む。この融合タンパク質は、デアミナーゼ(例えば、シチジンデアミナーゼまたはアデニンデアミナーゼ)および核酸プログラム可能DNA結合タンパク質(napDNAbp)を含む。ある態様において、キットは、対象核酸分子、例えば疾患原因遺伝子、疾患関連遺伝子(表3Aおよび表3Bに記載されているようなもの)、または例えば表4に列挙されたような疾患関連遺伝子に関連する調節配列(遺伝子プロモーター、エンハンサー、またはリプレッサーなど)を標的とすることができる、少なくとも一つのガイドRNAを含む。いくつかの実施形態において、キットは、少なくとも1つのガイドRNAをコードするヌクレオチド配列を含む核酸構築物を含む。

【0537】

いくつかの実施形態において、キットは、それを使用して一つ以上の疾患関連遺伝子もしくは疾患原因遺伝子、または疾患関連遺伝子に関連する一つ以上の調節配列(例えば表3A、3Bまたは4に列挙されている遺伝子、例えばSERPINA1、鎌状赤血球遺伝子(HBB)またはHBG1/2遺伝子)を編集するための使用説明書を提供する。説明書は、一般に、核酸分子を編集するためのキットの使用に関する情報を含む。他の実施形態では、説明書は

、以下のうちの少なくとも1つを含む：注意事項；警告；臨床試験；および/または参考資料。使用説明書は、容器（もしあれば）に直接印刷するか、容器に貼付するラベルとして印刷するか、または容器内もしくは容器と共に提供される独立したシート、パンフレット、カードまたはフォルダーとして印刷され得る。さらなる実施形態では、キットは、適切な動作パラメータのためのラベルまたは別個のインサート（添付文書）の形態で説明書を含むことができる。さらに別の実施形態では、キットは、検出、較正、または正規化のための標準として使用される、適切なポジティブおよびネガティブコントロールまたは対照サンプルを有する1つ以上の容器を含むことができる。キットは、（滅菌）リン酸緩衝生理食塩水、リンゲル液、またはデキストロース溶液などの薬学的に許容される緩衝液を含む第2の容器をさらに含むことができる。さらに、他の緩衝剤、希釈剤、フィルター、針、注射器、および使用説明書付きの添付文書を含む、商業的観点および使用者の観点から望ましい他の物質を含むことができる。

10

【0538】

ある態様において、キットは、鎌状赤血球症またはAIADのような表3Aおよび3Bに記載される遺伝子、または表4に記載される遺伝子の調節配列と関連する疾患または障害を有する対象の治療に有用である。

【0539】

本明細書に開示される実施形態の実施は、別段の指示がない限り、当業者の技量の範囲内にある、免疫学、生化学、化学、分子生物学、微生物学、細胞生物学、ゲノミクスおよび組換えDNAの従来技術を使用する。例えば、Sambrook and Green, *Molecular Cloning: A Laboratory Manual*, 4th Edition (2012); the series *Current Protocols in Molecular Biology* (F. M. Ausubel, et al. eds.); the series *Methods In Enzymology* (Academic Press, Inc.), PCR 2: A Practical Approach (M.J. MacPherson, B.D. Hames and G.R. Taylor eds. (1995)), Harlow and Lane, eds. (1988) *Antibodies, A Laboratory Manual*, and *Culture of Animal Cells: A Manual of Basic Technique and Specialized Applications*, 6th Edition (R.I. Freshney, ed. (2010)) 参照。

20

【0540】

塩基エディターシステムの方法および組成物ならびに使用を包含する、以下の番号付けされた追加の実施形態が本明細書で企図される：

30

1. 必要とする対象における疾患を治療する方法であって、

ガイドポリヌクレオチドまたはそれをコードする核酸と；

ポリヌクレオチドプログラム可能なDNA結合ドメインまたはそれをコードする核酸と；

アデノシンデアミナーゼドメインまたはそれをコードする核酸と

を含む塩基エディターシステムを前記対象に投与することを含み、

該ガイドポリヌクレオチドは、該対象中の細胞のHBBポリヌクレオチドにおいてA・TからG・Cへの改変を生じさせるように該塩基エディターシステムをターゲティングすることができ、それによって該疾患を治療する、

方法。

2. 必要とする対象における疾患を治療する方法であって、

40

(a) 細胞に、

ガイドポリヌクレオチドまたはそれをコードする核酸と；

ポリヌクレオチドプログラム可能なDNA結合ドメインまたはそれをコードする核酸と；

アデノシンデアミナーゼドメインまたはそれをコードする核酸と

を含む塩基エディターシステムを導入すること；および

(b) 前記細胞を前記対象に投与すること

を含み、前記ガイドポリヌクレオチドは、前記細胞におけるHBBポリヌクレオチド中の核酸塩基にA・TからG・Cへの改変を生じさせるように前記塩基エディターシステムをターゲティングすることができ、それによって該疾患を治療する、

方法。

50

3. 前記細胞が前駆細胞である実施形態2の方法。
4. 前記細胞が造血幹細胞、骨髓系共通前駆細胞、前赤芽球、赤芽球、網状赤血球、または赤血球である、実施形態2または3の方法。
5. 前記造血幹細胞がCD34⁺細胞である、実施形態4の方法。
6. 前記細胞が対象に対して自家性である、実施形態2～5のいずれか一記載の方法。
7. 前記細胞が対象に対して同種異系である、実施形態2～5のいずれか1つの方法。
8. 前記細胞が対象に対して異種性である、実施形態2～5のいずれか1つの方法。
9. 前記対象が哺乳動物である、先の実施形態のいずれか1つの方法。
10. HBBポリヌクレオチドを編集する方法であって、
HBBポリヌクレオチドを、
ガイドポリヌクレオチドと；
ポリヌクレオチドプログラム可能なDNA結合ドメインと；
アデノシンデアミナーゼドメインと
を含む塩基エディターシステムと接触させることを含み、
ここで、ガイドポリヌクレオチドは、HBBポリヌクレオチド中の核酸塩基においてA・TからG・Cへの改変をもたらすために前記塩基エディターシステムをターゲティングすることができるものである、
方法。
11. 疾患の治療のための改変された細胞を産生する方法であって、
細胞に、
ガイドポリヌクレオチドまたはそれをコードする核酸と；
ポリヌクレオチドプログラム可能なDNA結合ドメインまたはそれをコードする核酸と；
アデノシンデアミナーゼドメインまたはそれをコードする核酸と
を含む塩基エディターシステムを導入すること
を含み、
ここで、ガイドポリヌクレオチドは、前記細胞におけるHBBポリヌクレオチド中の核酸塩基においてA・TからG・Cへの改変をもたらすために前記塩基エディターシステムをターゲティングすることができるものである、
方法。
12. 前記導入がインビボである、実施形態11の方法。
13. 前記導入がex vivoである実施形態11の方法。
14. 前記細胞は前記疾患を有する対象から得られたものである、実施形態13の方法。
15. 前記細胞が哺乳動物細胞である、実施形態11～14のいずれか一記載の方法。
16. 前記細胞が前駆細胞である実施形態15の方法。
17. 前記細胞が造血幹細胞、骨髓系共通前駆細胞、前赤芽球、赤芽球、網状赤血球、または赤血球である、実施形態15または16の方法。
18. 前記造血幹細胞がCD34⁺細胞である、実施形態17の方法。
19. 前記ポリヌクレオチドプログラム可能なDNA結合ドメインがCas9ドメインである、前記実施形態のいずれか1つの方法。
20. Cas9ドメインがヌクレアーゼ不活性Cas9ドメインである、実施形態19の方法。
21. Cas9ドメインがCas9ニッカーゼドメインである、実施形態20の方法。
22. Cas9ドメインがSpCas9ドメインを含む、実施形態19～21のいずれか1つの方法。
23. SpCas9ドメインがD10Aおよび/またはH840Aアミノ酸置換またはその対応するアミノ酸置換を含む、実施形態22の方法。
24. SpCas9ドメインがNGG PAMに対する特異性を有する、実施形態22または23の方法。
25. SpCas9ドメインが、NGA PAM、NGT PAM、またはNGC PAMに対する特異性を有する、実施形態22～24のいずれか1つの方法。
26. SpCas9ドメインがアミノ酸置換L1111R、D1135V、G1218R、E1219F、A1322R、R1335V、T1337R、ならびにL1111、D1135L、S1136R、G1218S、E1219V

10

20

30

40

50

、D1332A、R1335Q、T1337I、T1337V、T1337F、およびT1337Mのうちの1つ以上、またはそれらの対応するアミノ酸置換を含む、実施態様22～25のいずれか1つの方法。

27. SpCas9ドメインがアミノ酸置換L1111R、D1135V、G1218R、E1219F、A1322R、R1335V、T1337R、ならびにL1111、D1135L、S1136R、G1218S、E1219V、D1332A、D1332S、D1332T、D1332V、D1332L、D1332K、D1332R、R1335Q、T1337I、T1337V、T1337F、T1337S、T1337N、T1337K、T1337R、T1337H、T1337Q、およびT1337Mのうちの1つ以上、またはそれらの対応するアミノ酸置換を含む、実施態様22～25のいずれか1つの方法。

28. SpCas9ドメインがアミノ酸置換D1135L、S1136R、G1218S、E1219V、A1322R、R1335Q、T1337、およびA1322R、ならびにL1111、D1135L、S1136R、G1218S、E1219V、D1332A、D1332S、D1332T、D1332V、D1332L、D1332K、D1332R、R1335Q、T1337I、T1337V、T1337F、T1337S、T1337N、T1337K、T1337R、T1337H、T1337Q、およびT1337Mのうちの1つ以上、またはそれらの対応するアミノ酸置換を含む、実施態様22～25のいずれか1つの方法。

29. SpCas9ドメインがアミノ酸置換D1135M、S1136Q、G1218K、E1219F、A1322R、D1332A、R1335E、およびT1337R、またはそれらの対応するアミノ酸置換を含む、実施態様22～25のいずれか1つの方法。

30. SpCas9ドメインが、NG PAM、NNG PAM、GAA PAM、GAT PAM、またはCAA PAMに対する特異性を有する、実施形態22～24のいずれか1つの方法。

31. Cas9ドメインがアミノ酸置換E480K、E543K、およびE1219V、またはそれらの対応するアミノ酸置換を含む、実施形態30の方法。

32. Cas9ドメインがSaCas9ドメインを含む、実施態様19～21のいずれか1つの方法。

33. SaCas9ドメインは、NNNRRT PAMに対する特異性を有する、実施形態32の方法。

34. SaCas9ドメインは、NNGRRT PAMに対する特異性を有する、実施形態33の方法。

35. SaCas9ドメインがアミノ酸置換N579Aまたはその対応するアミノ酸置換を含む、実施態様32～34のいずれか1つの方法。

36. SaCas9ドメインがアミノ酸置換E782K、N968K、およびR1015H、またはそれらの対応するアミノ酸置換を含む、実施態様32～35のいずれか1つの方法。

37. Cas9ドメインがSt1Cas9ドメインを含む、実施態様19～21のいずれか1つの方法。

38. St1Cas9ドメインがNNACCA PAMに対する特異性を有する実施形態37の方法。

39. 前記アデノシンデアミナーゼドメインが、天然には存在しない改変アデノシンデアミナーゼドメインである、前記実施形態のいずれか1つの方法。

40. アデノシンデアミナーゼドメインがTadAドメインを含む、実施形態39の方法。

41. TadAドメインは、TadA 7.10のアミノ酸配列を含む、実施形態36の方法。

42. 前記塩基エディターシステムがジンクフィンガードメインをさらに含む、前記実施形態のいずれか1つの方法。

43. ジンクフィンガードメインが、認識ヘリックス配列RNEHLEV、QSTTLKR、およびRTEHLAR、または認識ヘリックス配列RGEHLRQ、QSGTLKR、およびRNDKLVLPを含む、実施形態42の方法。

44. 実施形態42または43の方法であって、ジンクフィンガードメインがzf1raまたはzf1rbである、方法。

45. 前記塩基エディターシステムが核局在化シグナル (NLS) をさらに含む、前記実施形態のいずれか一つの方法。

46. 前記塩基エディターシステムが1つ以上のリンカーをさらに含む、前記実施形態のいずれか1つの方法。

47. ポリヌクレオチドプログラム可能なDNA結合ドメイン、アデノシンデアミナーゼドメイン、ジンクフィンガードメイン、およびNLSのうちの二つ以上がリンカーを介して連結されている、実施形態46の方法。

48. リンカーがペプチドリンカーであり、それによって塩基エディター融合タンパク質

10

20

30

40

50

を形成する、実施形態47の方法。

49. ペプチドリンカーが、SGGSSGSETPGTSESATPESSGGS, SGGSSGGSSGSETPGTSESATPESSGGSSGGS, GSGGSPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGGSSGGS, SGGSSGGSSGSETPGTSESATPES, SGGSSGGSSGSETPGTSESATPESSGGSSGGSSGGSSGGS, SGGSSGGSSGSETPGTSESATPESSGGSSGGSSGGSSGGSSGSETPGTSESATPESSGGSSGGS, PGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATS, (SGGS)_n, (GGGS)_n, (GGGGS)_n, (G)_n, (EAAK)_n, (GGS)_n, SGSETPGTSESATPES, および (XP)_nからなる群より選択されるアミノ酸配列を含む、実施形態48の方法。

10

50. 塩基エディター融合タンパク質が、

MPKKKRKVSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSSGGSSGSETPGTSESATPESSGGSSGGSSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVHNNRVIGEGWNRRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTDSSGGSSGSSGSETPGTSESATPESDLVLGLAIGISVGVGILNKVTGEIIHKNSRIFPAAQAENNLVRRRTNRQGRRLARRKKHRRVRLNRLFEESSLITDFTKISINLNPYQLRVKGLTDELSNEELFIALKNMVKHGHSYLDASDDGNSSVGDYAQIVKENSQKLETKTGQIQQLERYQTYGQLRGDFTVEKDGGKHRLINVFPTSAYRSEALRILQTQQEFNPQITDEFINRYLEILTGKRKYHGPNEKSRTDYGRYRTSGETLDNIFGILIGKCTFYDEFRAAKASYTAQEFNLLNDLNNLTVPPTETKLSKEQKNQIINYVKNEKAMGPAKLFKYIAKLSCDVADIKGYRIDKSGKAEIHTFEAYRKMKTLETLDIEQMDRETLDKLAYVLTNTEREQIQEALHEFADGSFSQKQVDELVQFRKANSSIFGKGWHNFSVKLMMELIPELYETSEEQMTILTRLGKQKTTSSSNKTKYIDEKLLTEIYNPVAKSVRQAIKIVNAAIKEYGDFDNIVIEMARETNEDDEKKAIQKIQKANKDEKDAAMLKAANQYNGKAELPHSVFHGHKQLATKIRLWHQQGERCLYTGTISIHDLINNSNQFEVDHILPLSITFDDSLANKVLVYATANQEKQRTPYQALDSMDDAWSFRELKAFVRESKTLNKKKEYLLTEEDISKFDVRKKFIERNLVDITYASRVVLNALQEHFRAHKIDTKVSVVRGQFTSQLRRHWGIEKTRDITYHHHAVDALIIAASQLNLWKKQKNTLVSYSEDQLLDIETGELISDDEYKESVFKAPYQHFDVTLKSKEFEDSILFSYQVDSKFNKISDATIYATRQAKVGKDKADETYVLGKIKDIYTQDGYDAFMKIYKKDKSKFLMYRHDPQTFEKVIENYPNKQINDKGKEVPCNPFLKYKEEHGYIRKYSKKGNGPEIKSLKYYDSKLGHNHIDITPKDSNNKVVLQSVSPWRADVFNKTTGKYEILGLKYADLQFDKGTGTYKISQEKYNDIKKKEGVDSSEFKFTLYKNDLLLVKDTETKEQQLFRFLSRTMPKQKHVELKPYDKQKFEGGEALIKVLGNVANSGQCKKGLGKSNISYKVRTDVLGNQHIIKNEGDKPKLDFPKKKRKEGADKRTADGSEFESPKKKRKV, M SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSSGGSSGSSGSETPGTSESATPESSGGSSGGSSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVHNNRVIGEGWNRRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTDSSGGSSGSSGSETPGTSESATPESSGGSSGSKRNYILGLAIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGAARLKRHHRIQRVKLLFDYNLLTDHSELGINPYEARVKGLSQKLS EEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNSKALEEKYVAELQLERLKKDGEVRGSINRFKTS DYVKEAKQLLKVKAYHQLDQSFDITYIDLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALNDLNNLVITRDENEKLEYEKFQIENVFQKQKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEFTNLKVYHDIKDITARKEIIE NAELLDQIAKILTIYQSSEDIQEELTNLNSLTQEEIEQISNLKGYTGTHNLSLKAINLILDE

20

30

40

50

LWHTNDNQIAIFNRLKLVPPKKVDLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIKKY
GLPNDIIIELAREKNSKDAQKMINEMQKRNRTNERIEEIIRTTGKENAKYLIEKIKLHDM
QEGKCLYSLEAIPLEDLLNPNFNYEVDHIIPRSVSFDNSFNKVLVKQEENSKKGNRTPFQ
YLSSSDSKISYETFKKHILNLAAGKGRISKTKEYLLEERDINRFSVQKDFINRNLVDTRYA
TRGLMNLRLSYFRVNNLDVKKVKSINGGFTSFLRRKWKFKKERNKGYKHHAE DALIIANAD
FIFKEWKKLDKAKKVMENQMFEKQAESMPEIETE QEYKEIFITPHQIKHIKDFKDYKYS
HRVDKKPNRKLINDTLYSTRKDDKGNTLIVNNLNGLYDKDNDKLLKLINKSPEKLLMYH
HDPQTYQKLKLIMEQYGDEKNPLYKYEEETGNYLTKYSSKDNNGPVIKKIKYYGNKLNAH
LDITDDYPNSRNKVVKLSLKPYRFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSKCYEEA
KKLKKISNQAEFIASFYKNDLIKINGELYRVIGVNNDLLNRIEVNMIDITYREYLENMNDK
RPPHIIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKGEGADKRTADGSEFESPKKKR
KV、

10

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAH
AEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAGS
LMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAKKAQSSTDSSGSSGGSS
GSETPGTSESATPESSGGSSGSSSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVN
RVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHS
RIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVF
NAQKKAQSSTDSSGSSGGSSGSETPGTSESATPESSGGSSGGSKRNYILGLAIGITSVGYGII
DYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRRRRHRIQRVKLLFDYNLLTDHS
ELSGINPYEARVKGLSQKLSEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNS
KALEEKYVAELQLERLKKDGEVRGSINRFKTSYVKEAKQLLKVKQAYHQLDQSFIDTYI
DLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALND
LNNLVITRDENEKLEYEKFQIENVFQKQKKPTLQKIAKEILVNEEDIKGYRVTSTGKPEF
TNLKVYHDIKDITARKEIENAELLDQIAKILTIIYQSSEDIQEELTNLNSLTQEEIEQISNL
KGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPPKKVDLSQQKEIPTTLVDDFILS
PVVKRSFIQSIKVINAIKKYGLPNDIIIELAREKNSKDAQKMINEMQKRNRTNERIEEII
RTTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNPNFNYEVDHIIPRSVSFDNSFN
KVLVKQEENSKKGNRTPFQYLSSSDSKISYETFKKHILNLAAGKGRISKTKEYLLEERDIN
RFSVQKDFINRNLVDTRYATRGLMNLRLSYFRVNNLDVKKVKSINGGFTSFLRRKWKFKKE
RNKGYKHHAE DALIIANADFIFKEWKKLDKAKKVMENQMFEKQAESMPEIETE QEYKEI
FITPHQIKHIKDFKDYKYSHRVDKKPNRKLINDTLYSTRKDDKGNTLIVNNLNGLYDKD
DKLKKLINKSPEKLLMYHHDPQTYQKLKLIMEQYGDEKNPLYKYEEETGNYLTKYSSK
DNNGPVIKKIKYYGNKLNAHLDITDDYPNSRNKVVKLSLKPYRFDVYLDNGVYKFVTVKN
LDVIKKENYYEVNSKCYEEAKKLKKISNQAEFIASFYKNDLIKINGELYRVIGVNNDLLN
RIEVNMIDITYREYLENMNDKRPPHIIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQII
KKGEGADKRTADGSEFESPKKKRKVSSGNSNANSRGPSSGLVPLSLRGSHSRPGERPFQ
CRICMRNFSRNEHLEVHTRTHTGEKPFQCRICMRNFSQSTTLKRHLRTHTGEKPFQCRIC
MRNFSRTEHLARHLKTHLRGSSAQ、または

20

30

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAH
AEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAGS
LMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAKKAQSSTDSSGSSGGSS
GSETPGTSESATPESSGGSSGSSSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVN
RVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHS
RIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVF
NAQKKAQSSTDSSGSSGGSSGSETPGTSESATPESSGGSSGGSKRNYILGLAIGITSVGYGII
DYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRRRRHRIQRVKLLFDYNLLTDHS
ELSGINPYEARVKGLSQKLSEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNS
KALEEKYVAELQLERLKKDGEVRGSINRFKTSYVKEAKQLLKVKQAYHQLDQSFIDTYI
DLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALND

40

50

AGUCAGAUGUUUUUGUACUCUCAAGAUUUUAAGUAAACUGUACAACGAAACUUACACAG
UUACUUAAAUCUUGCAGAAGCUACAAAGAUAAAGGCUUCAUGCCGAAAUCAACACCCUG
UCAUUUUUAUGGCAGGGUGからなる群より選択される配列を含む、実施形態62の方法。

64. 必要とする対象において鎌状赤血球症を治療する方法であって、

単一ガイドRNA (sgRNA)と；

MPKKKRKVSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGR
HDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAK
TGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDG
GSSGGSSGSETPGTSESATPESSGGSSGSSEVEFSHEYWMRHALTLAKRARDEREVPVGA
VLVLNNRVIGEGWNRRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCA
GAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFR
MPRQVFNAQKKAQSSTDGSSGGSSGSETPGTSESATPESDLVLGLAIGISVGVGILNK
VTGEIIHKNSRIFPAAQAENNLVRRTNRQGRRLARRKKHRRVRLNRLFEEGLITDFTKISI
NLNPYQLRVKGLTDELSNEELFIALKNMVKHGRISYLLDASDDGNSSVGDYAQIVKENS
KQLETKTPGQIQLERYQTYGQLRGDFTVEKDGGKHRLINVFPTSAYRSEALRILQTQQEFN
PQITDEFINRYLEILTGKRKYHGPNEKSRTDYGRYRTSGETLDNIFGILIGKCTFYPDEF
RAAKASYTAQEFNLLNDLNNLTVPPTETKKLSKEQKNQIINYVKNEKAMGPAKLFKYIAKL
LSCDVADIKGYRIDKSGKAEIHTFEAYRKMKTLETLDIEQMDRETLDKLAYVLTNTEREG
IQEALEHEFADGSFSQKQVDELVQFRKANSSIFGKGWHNFSVKLMMELIPELYETSEEQ
MTILRLGKQKTTSSSNKTKYIDEKLLTEEIYNPVVAKSVRQAIKIVNAAIKEYGDFDNIVIE
MARETNEDDEKKAIQKIQKANKDEKDAAMLKAANQYNGKAELPHSVFHGHKQLATKIRL
WHQQGERCLYTGKTISIHDLINNSNQFEVDHILPLSITFDDSLANKVLVYATANQEKQR
TPYQALDSMDDAWSFRELKAFVRESKTLNKKKEYLLTEEDISKFDVRKKFIERNLVDL
YASRVVLNALQEHFRAHKIDTKVSVVRGQFTSQLRRHWGIEKTRDTYHHHAVDALIIAAS
SQLNLWKKQKNTLVSYSEDQLLDIETGELISDDEYKESVFKAPYQHFDVTLKSKEFEDSIL
FSYQVDSKFNKRISDATIYATRQAKVGKDKADETYVLGKIKDIYTQDGYDAFMKIYKKDK
SKFLMYRHDPQTFEKVIEPIENYPNKQINDKGKEVPCNPFLKYKEEHGYIRKYSKKGNG
PEIKSLKYYDSKLGHNHIDITPKDSNNKVVLQSVSPWRADVFNKTTGKYEILGLKYADLQ
FDKGTGTYKISQEKYNDIKKKEGVDSSEFKFTLYKNDLLLVKDTETKEQQFLRFLSRTMP
KQKHVELKPYDKQKFEGGEALIKVLGNVANSGQCKKGLGKSNISYKVRTDVLGNQHII
KNEGDKPKLDFPKKKRKVEGADKRTADGSEFESPKKKRKV、
MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAH
AEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAGS
LMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDGSSGGSSG
GSETPGTSESATPESSGGSSGSSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNN
RVIGEGWNRRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHS
RIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVF
NAQKKAQSSTDGSSGGSSGSETPGTSESATPESSGGSSGSKRNYILGLAIGITSVGYGII
DYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLRKRRRRHRIQRVKLLFDYNLLTDHS
ELSGINPYEARVKGLSQKLSEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNS
KALEEKYVAELQLERLKKDGEVRGSINRFKTSYVKEAKQLLKVKAYHQLDQSFIDTYI
DLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALND
LNNLVITRDENEKLEYEYEFQIENVFQKQKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEF
TNLKVYHDIKDITARKEIENAELLDQIAKILTIIYQSSEDIQEELTNLNSLTQEEIEQISNL
KGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPPKKVDLSQQKEIPTTLVDDFILS
PVVKRSFIQSIKVINAIIKKYGLPNDIIELAREKNSKDAQKMINEMQKRNRQTNERIEEII
RTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNNPFNYEVDHIIPRSVSFDNSFNN
KVLVKQEENSKKGNRTPFQYLSSSDSKISYETFKKHILNLAGKGRISKTKKEYLLEERDIN
RFSVQKDFINRLVDTRYATRGLMNLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKE
RNKGKXHAEDALIIANADFIKFWKKLDKAKKVMENQMFEKQAESMPEIETEQEYKEI

10

20

30

40

50

FITPHQIKHIKDFKDYKYSHRVDKKPNRKLINDTLYSTRKDDKGNTLIVNNLNGLYDKDN
DKLKKLINKSPEKLLMYHHDPTQYQKLKLIMEQYGDEKNPLYKYYEETGNYLTKYSKKD
NGPVIKKIKYYGNKLNAHLDDITDDYPNSRNKVVKLSLKPYPYFDVYLDNGVYKFVTVKNLD
VIKKENYYEVNSKCYEEAKKLKISNQAEFIASFYKNDLIKINGELYRVIGVNNDLLNRIEV
NMIDITYREYLENMNDKRPPHIIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKGEGA
DKRTADGSEFESPKKKRKV、

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAH
AEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAGS
LMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAKKAQSSTDSSGSSGGSS
GSETPGTSESATPESSGGSSGSSGSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNN
RVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHS
RIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVF
NAQKKAQSSTDSSGSSGGSSGSETPGTSESATPESSGGSSGGSKRNYILGLAIGITSVGYGII
DYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRRRRHRIQRVKKLLFDYNLLTDHS
ELSGINPYEARVKGLSQKLSEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNS
KALEEKYVAELQLERLKKDGEVRGSINRFKTSYVKEAKQLLKVKQKAYHQLDQSFIDTYI
DLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALND
LNNLVITRDENEKLEYEYEFQIIENVFKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEF
TNLKVYHDIKDITARKEIENAELLDQIAKILTIYQSSEDIQEELTNLNSLTQEEIEQISNL
KGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILS
PVVKRSFIQSIKVINAIKKYGLPNDIIIELAREKNSKDAQKMINEMQKRNRQTNERIEEIIIR
TTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNPNFNYEVDHIIPRSVSFDNSFNN
KVLVKQEENSCKGNRTPFQYLSSSDSKISYETFKKHILNLAAGKGRISKTKEYLLEERDIN
RFSVQKDFINRNLVDTRYATRGLMNLRSYFRVNNLDVVKVKSINGGFTSFLRRKWKFKKE
RNKGYKHAEDALIINANADFIFKEWKKLDKAKKVMENQMFEKQAESMPEIETEYQYKEI
FITPHQIKHIKDFKDYKYSHRVDKKPNRKLINDTLYSTRKDDKGNTLIVNNLNGLYDKDN
DKLKKLINKSPEKLLMYHHDPTQYQKLKLIMEQYGDEKNPLYKYYEETGNYLTKYSKKD
NGPVIKKIKYYGNKLNAHLDDITDDYPNSRNKVVKLSLKPYPYFDVYLDNGVYKFVTVKNLD
VIKKENYYEVNSKCYEEAKKLKISNQAEFIASFYKNDLIKINGELYRVIGVNNDLLNRIEV
NMIDITYREYLENMNDKRPPHIIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKGEGA
DKRTADGSEFESPKKKRKVSSGNSNANSRGPSFSSGLVPLSLRGSHSRPGERPFQCRICMR
NFSRNEHLEVHTRTHTGEKPFQCRICMRNFSQSTTLKRHLRTHTGEKPFQCRICMRNFSR
TEHLARHLKTHLRGSSAQ、または

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAH
AEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAGS
LMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAKKAQSSTDSSGSSGGSS
GSETPGTSESATPESSGGSSGSSGSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNN
RVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHS
RIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVF
NAQKKAQSSTDSSGSSGGSSGSETPGTSESATPESSGGSSGGSKRNYILGLAIGITSVGYGII
DYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRRRRHRIQRVKKLLFDYNLLTDHS
ELSGINPYEARVKGLSQKLSEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNS
KALEEKYVAELQLERLKKDGEVRGSINRFKTSYVKEAKQLLKVKQKAYHQLDQSFIDTYI
DLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALND
LNNLVITRDENEKLEYEYEFQIIENVFKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEF
TNLKVYHDIKDITARKEIENAELLDQIAKILTIYQSSEDIQEELTNLNSLTQEEIEQISNL
KGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILS
PVVKRSFIQSIKVINAIKKYGLPNDIIIELAREKNSKDAQKMINEMQKRNRQTNERIEEIIIR
TTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNPNFNYEVDHIIPRSVSFDNSFNN
KVLVKQEENSCKGNRTPFQYLSSSDSKISYETFKKHILNLAAGKGRISKTKEYLLEERDIN

10

20

30

40

50

RFSVQKDFINRNLVDTRYATRGLMNLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKE
 RNKGKHHHAEDALIIANADFIFKEWKKLDKAKKVMENQMFEKQAESMPEIETEQEYKEI
 FITPHQIKHIKDFKDYKYSHRVDKKPNRKLINDTLYSTRKDDKGNTLIVNNLNGLYDKDN
 DKLKKLINKSPEKLLMYHHPQTYQKLKLIMEQYGDEKNPLYKYEEETGNLYLTKYSKKD
 NGPVIKKIKYGGNKLNAHLDDITDDYPNSRNKVVKLSLKPYRFDVYLDNGVYKFVTVKNLD
 VIKKENYYEVNSKCYEEAKKLKISNQAEFISFYKNDLIKINGELYRVIGVNNDLLNRIEV
 NMIDITYREYLENMNDKRPPHIIKTISKYSTDILGNLYEVKSKKHPQIIKKGEGA
 DKRTADGSEFESPKKKRKVSSGNSNANSRGPSSFSSGLVPLSLRGSHSRPGERPFQCRICMR
 NFSRGEHLRQHTRTHTGEKPFQCRICMRNFSQSGTLKRHLRTHHTGEKPFQCRICMRNFSR
 NDKLVPHLKTHLRGSSAQ

10

からなる群から選択されるアミノ酸配列を含む融合タンパク質と

を含む核酸塩基エディターシステムを前記対象に投与することを含み、

前記sgRNAは、前記対象の細胞中のHBBポリヌクレオチドの位置17においてA・Tから
 G・Cへの改変を生じさせるように前記塩基エディターシステムをターゲティングすること
 ができ、それによって鎌状赤血球症を治療し、ここでsgRNAは、CUUCUCCACAGGAGUC
 AGAU, ACUUCUCCACAGGAGUCAGAU, GACUUCUCCACAGGAGUCAGAU, GUUUUUG
 UACUCUCAAGAUUUAAGUAAACUGUACAACGAAACUACACAGUACUAAAUCUUGC
 AGAAGCUACAAAGAUAAAGGCUUCAUGCCGAAAUCAACACCCUGUCAUUUUUAUGGCAGG
 GUG, CUUCUCCACAGGAGUCAGAUUUUUUGUACUCUCAAGAUUUAAGUAAACUGUAC
 AACGAAACUACACAGUACUAAAUCUUGCAGAAGCUACAAAGAUAAAGGCUUCAUG
 CCGAAAUCAACACCCUGUCAUUUUUAUGGCAGGGUG, ACUUCUCCACAGGAGUCAGAU
 UUUUUGUACUCUCAAGAUUUAAGUAAACUGUACAACGAAACUACACAGUACUAAA
 UCUUGCAGAAGCUACAAAGAUAAAGGCUUCAUGCCGAAAUCAACACCCUGUCAUUUUUA
 UGGCAGGGUG; および GACUUCUCCACAGGAGUCAGAUUUUUUGUACUCUCAAGAU
 UUAAGUAAACUGUACAACGAAACUACACAGUACUAAAUCUUGCAGAAGCUACAAA
 GAUAAGGCUUCAUGCCGAAAUCAACACCCUGUCAUUUUUAUGGCAGGGUGからなる群よ
 り選択される配列を含む、方法。

20

65. 必要とする対象において鎌状赤血球症を治療する方法であって、

(a) 前記対象から得られた細胞に、

単一ガイドRNA (sgRNA)と;

30

MPKKKRKVSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGR
 HDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAK
 TGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAKKAQSSTDSG
 GSSGGSSGSETPGTSESATPESSGGSSGSSEVEFSHEYWMRHALTLAKRARDEREVPVGA
 VLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCA
 GAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFR
 MPRQVFNAQKKAQSSTDSGGSSGGSSGSETPGTSESATPESDLVLGLAIGIGSVGVGILNK
 VTGEIIHKNSRIFPAAQAENNLVRRNTNRQGRRLARRKKHRRVRLNRLFEEGLITDFTKISI
 NLNPYQLRVKGLTDELSNEELFIALKNMVKHGRISYLDASDDGNSSVGDYAQIVKENS
 KQLETKTGQIQQLERYQTYGQLRGDFTVEKDGGKHRLINVFPTSAYRSEALRILQTQQEFN
 PQITDEFINRYLEILTGKRKYHGPNEKSRTDYGRYRTSGETLDNIFGILIGKCTFYDEF
 RAAKASYTAQEFNLLNDLNNLTVPETETKLSKEQKNQIINYVKNEKAMGPAKLFKYIAKL
 LSCDVADIKGYRIDKSGKAEIHTFEAYRKMKTLETLDIEQMDRETLDKLAYVLTNTEREG
 IQEALHEFADGSFSQKQVDELVQFRKANSSIFGKGWHNFSVKLMMELIPELYETSEEQ
 MTILRLGKQKTSSSNKTKYIDEKLLTEIYNPVVAKSVRQAIKIVNAAIKEYGDFDNIVIE
 MARETNEDDEKKAIQKIQKANKDEKDAAMLKAANQYNGKAELPHSVFHGHKQLATKIRL
 WHQQGERCLYTGTISIHDLINNSNQFEVDHILPLSITFDDSLANKVLVYATANQEKQQR
 TPYQALDSMDDAWSFRELKAFVRESKTLNKKKEYLLTEEDISKFDVRKKFIERNLVDL
 YASRVVLNALQEHFRAHKIDTKVSVVRGQFTSQLRRHWGIEKTRDTYHHHAVDALIIAAS
 SQLNLWKKQKNTLVSYSEDQLLDIETGELISDDEYKESVFKAPYQHFVDTLKSKEFEDSIL

40

50

FSYQVDSKFNRKISDATIYATRQAKVGKDKADETYVLGKIKDIYTQDGYDAFMKIYKKDK
SKFLMYRHDPQTFEKVIPILENYPNKQINDKGKEVPCNPFLKYKEEHGYIRKYSKKGNG
PEIKSLKYYDSKLGNHIDITPKDSNNKVVLQSVSPWRADVFNKTTGKYEILGLKYADLQ
FDKGTGTYKISQEKYNDIKKKEGVDSSEFKFTLYKNDLLLVKDTETKEQQFLRFLSRTMP
KQKHVELKPYDKQKFEGGEALIKVLGNVANSGQCKKGLGKSNISYKVRTDVLGNQHII
KNEGDKPKLDFPKKKRKVEGADKRTADGSEFESPKKRKV、

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAH
AEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAGS
LMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSSGSSGGSS
GSETPGTSESATPESSGGSSGSSGSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNN
RVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHS
RIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVF
NAQKKAQSSTDSSGSSGGSSGSETPGTSESATPESSGGSSGGSKRNYILGLAIGITSVGYGII
DYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRRRRHRIQRVKLLFDYNLLTDHS
ELSGINPYEARVKGLSQKLSEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNS
KALEEKYVAELQLERLKKDGEVRGSINRFKTSYVKEAKQLLKVKQAYHQLDQSFIDTYI
DLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALND
LNNLVITRDENEKLEYEYEFQIENVFQKQKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEF
TNLKVYHDIKDITARKEIENAELLDQIAKILTIIYQSSEDIQEELTNLNSLTQEEIEQISNL
KGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFELS
PVVKRSFIQSIKVINAIKKYGLPNDIIELAREKNSKDAQKMINEMQKRNRQTNERIEEIIIR
TTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNNPFNYEVDHIIPRSVSFDNSFNN
KVLVKQEENSCKGNRTPFQYLSSSDSKISYETFKKHILNLAAGKGRISKTKEYLLEERDIN
RFSVQKDFINRNLVDTRYATRGLMNLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKE
RNKGYKHAEDALIINANADFIFKEWKKLDKAKKVMENQMFEKQAESMPEIETEYKEYKEI
FITPHQIKHIKDFDKYKYSHRVDKKPNRKLINDTLYSTRKDDKGNTLIVNNLNGLYDKDN
DKLKKLINKSPEKLLMYHHPQTYQKLKLIMEQYGDEKNPLYKYEEETGNYLTKYSKKD
NGPVIKKIKYYGNKLNAHLDDITDDYPNSRNKVVKLSLKPYPYFDVYLDNGVYKFVTVKNLD
VIKKENYYEVNSKCYEEAKKLKISNQAEFIASFYKNDLIKINGELYRVIGVNNDLLNRIEV
NMIDITYREYLENMNDKRPPHIIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKGEGA
DKRTADGSEFESPKKRKV、

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAH
AEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAGS
LMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSSGSSGGSS
GSETPGTSESATPESSGGSSGSSGSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNN
RVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHS
RIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVF
NAQKKAQSSTDSSGSSGGSSGSETPGTSESATPESSGGSSGGSKRNYILGLAIGITSVGYGII
DYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRRRRHRIQRVKLLFDYNLLTDHS
ELSGINPYEARVKGLSQKLSEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNS
KALEEKYVAELQLERLKKDGEVRGSINRFKTSYVKEAKQLLKVKQAYHQLDQSFIDTYI
DLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALND
LNNLVITRDENEKLEYEYEFQIENVFQKQKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEF
TNLKVYHDIKDITARKEIENAELLDQIAKILTIIYQSSEDIQEELTNLNSLTQEEIEQISNL
KGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFELS
PVVKRSFIQSIKVINAIKKYGLPNDIIELAREKNSKDAQKMINEMQKRNRQTNERIEEIIIR
TTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNNPFNYEVDHIIPRSVSFDNSFNN
KVLVKQEENSCKGNRTPFQYLSSSDSKISYETFKKHILNLAAGKGRISKTKEYLLEERDIN
RFSVQKDFINRNLVDTRYATRGLMNLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKE
RNKGYKHAEDALIINANADFIFKEWKKLDKAKKVMENQMFEKQAESMPEIETEYKEYKEI

10

20

30

40

50

FITPHQIKHIKDFKDYKYSHRVDKKPNRKLINDTLYSTRKDDKGNTLIVNNLNGLYDKDN
DKLKKLINKSPEKLLMYHHDPTQTYQKLKLIMEQYGDEKNPLYKYYEETGNLYLTYSKKD
NGPVIKKIKYYYGNKLNAHLDDITDDYPNSRNKVVKLSLKPYPYRFDVYLDNGVYKFVTVKNLD
VIKKENYYEVNSKCYEEAKKLKISNQAEFIASFYKNDLIKINGELYRVIGVNNDLLNRIEV
NMIDITYREYLENMNDKRPPHIIKTASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKGEGA
DKRTADGSEFESPKKKRKVSSGNSNANSRGPSSFSSGLVPLSLRGSHSRPGERPFQCRICMR
NFSRNEHLEVHTRTHTGEKPFQCRICMRNFSQSSTTLKRHLRTHHTGEKPFQCRICMRNFSR
TEHLARHLKTHLRGSSAQ、または

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAH
AEIMALRQGGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAGS
LMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAKKAQSSTDSSGSSGGSS
GSETPGTSESATPESSGGSSGSSGSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNN
RVIGEGWNRAIGLHDPTAHAEIMALRQGGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHS
RIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVF
NAQKKAQSSTDSSGSSGGSSGSETPGTSESATPESSGGSSGSSGSKRNYILGLAIGITSVGYGII
DYETRDVIDAGVRLFKEANVENNEGRRSKRGRRLKRRRRHRIQRVKKLLFDYNLLTDHS
ELSGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNS
KALEEKYVAELQLERLKKDGEVRGSINRFKTSYVKEAKQLLKVKQKAYHQLDQSFIDTYI
DLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALND
LNNLVITRDENEKLEYEKFQIIENVFKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEF
TNLKVYHDIKDITARKEIENAELLDQIAKILTIYQSSEDIQEELTNLNSLTQEEIEQISNL
KGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILS
PVVKRSFIQSIKVINAIKKYGLPNDIIIELAREKNSKDAQKMINEMQKRNRQTNERIEEIIIR
TTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNPNFNYEVDHIIIPRSVSFDNSFNN
KVLVKQEENSKKGNRTPFQYLSSSDSKISYETFKKHILNLAGKGRISKTKEYLLEERDIN
RFSVQKDFINRNLVDTRYATRGLMNLRLSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKE
RNKGKYYHAEDALIINANADFIFKEWKKLDKAKKVMENQMFEKQAESMPEIETEYKEYKEI
FITPHQIKHIKDFKDYKYSHRVDKKPNRKLINDTLYSTRKDDKGNTLIVNNLNGLYDKDN
DKLKKLINKSPEKLLMYHHDPTQTYQKLKLIMEQYGDEKNPLYKYYEETGNLYLTYSKKD
NGPVIKKIKYYYGNKLNAHLDDITDDYPNSRNKVVKLSLKPYPYRFDVYLDNGVYKFVTVKNLD
VIKKENYYEVNSKCYEEAKKLKISNQAEFIASFYKNDLIKINGELYRVIGVNNDLLNRIEV
NMIDITYREYLENMNDKRPPHIIKTASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKGEGA
DKRTADGSEFESPKKKRKVSSGNSNANSRGPSSFSSGLVPLSLRGSHSRPGERPFQCRICMR
NFSRGEHLRQHTRTHTGEKPFQCRICMRNFSQSSTTLKRHLRTHHTGEKPFQCRICMRNFSR
NDKLVPHLKTHLRGSSAQ

からなる群から選択されるアミノ酸配列を含む融合タンパク質と
を含む塩基エディターシステムを導入すること、および

(b) 前記細胞を前記対象に投与することを含み、

前記sgRNAは、前記対象の細胞のHBBポリヌクレオチドの位置17においてA・TからG
・Cへの改変を生じさせるように前記塩基エディターシステムをターゲティングすること
ができ、それによって鎌状赤血球症を治療し、ここでsgRNAは、CUUCUCCACAGGAGUC
AGAU, ACUUCUCCACAGGAGUCAGAU, GACUUCUCCACAGGAGUCAGAU, GUUUUUG
UACUCUCAAGAUUUAAGUAACUGUACAACGAAACUACACAGUACUUAUAAUCUUGC
AGAAGCUACAAAGAUAAAGGCUUCAUGCCGAAAUCAACACCCUGUCAUUUUUAUGGCAGG
GUG, CUUCUCCACAGGAGUCAGAUUUUUUGUACUCUCAAGAUUUAAGUAACUGUAC
AACGAAACUACACAGUACUUAUAAUCUUGCAGAAAGCUACAAAGAUAAAGGCUUCAUG
CCGAAAUCAACACCCUGUCAUUUUUAUGGCAGGGUG, ACUUCUCCACAGGAGUCAGAU
UUUUUGUACUCUCAAGAUUUAAGUAACUGUACAACGAAACUACACAGUACUUAUAA
UCUUGCAGAAAGCUACAAAGAUAAAGGCUUCAUGCCGAAAUCAACACCCUGUCAUUUUUA
UGGCAGGGUG; および GACUUCUCCACAGGAGUCAGAUUUUUUGUACUCUCAAGAU

UUAAGUAAACUGUACAACGAAACUUACACAGUUACUAAAAUCUUGCAGAAGCUACAAA
GAUAAGGCUUCAUGCCGAAAUCAACACCCUGUCAUUUUAUGGCAGGGUGからなる群より
選択される配列を含む、

方法。

66．塩基エディターシステムを含む、改変された細胞であって、

前記塩基エディターシステムは、

ガイドポリヌクレオチドまたはそれをコードする核酸と；

ポリヌクレオチドプログラム可能なDNA結合ドメインまたはそれをコードする核酸と；

アデノシンデアミナーゼドメインまたはそれをコードする核酸と

を含み、

ここで、ガイドポリヌクレオチドは、前記細胞におけるHBBポリヌクレオチド中の核酸
塩基においてA・TからG・Cへの改変をもたらすために前記塩基エディターシステムをタ
ーゲティングすることができるものである、

細胞。

67．前記細胞は疾患を有する対象から得られたものである、実施形態66の改変された細胞。

68．前記細胞は哺乳動物細胞である、実施形態66または67の改変された細胞。

69．前記細胞が前駆細胞である、実施形態68の改変された細胞。

70．前記細胞が造血幹細胞、骨髄系共通前駆細胞、前赤芽球、赤芽球、網状赤血球、ま
たは赤血球である、実施形態68または69の改変された細胞。

71．前記造血幹細胞がCD34⁺細胞である、実施形態70の改変された細胞。

72．前記ポリヌクレオチドプログラム可能なDNA結合ドメインがCas9ドメインである、
実施形態68～71のいずれか1つの改変された細胞。

73．前記Cas9ドメインがヌクレアーゼ不活性Cas9ドメインである、実施形態72の改変
された細胞。

74．前記Cas9ドメインがCas9ニッカーゼドメインである、実施形態72の方法。

75．前記Cas9ドメインがSpCas9ドメインを含む、実施形態72～74のいずれか1つの改
変された細胞。

76．SpCas9ドメインがD10Aおよび/またはH840Aアミノ酸置換、またはその対応する
アミノ酸置換を含む、実施形態75の改変された細胞。

77．SpCas9ドメインがNGG PAMに対する特異性を有する、実施形態75または76の改
変された細胞。

78．SpCas9ドメインがNGA PAM、NGT PAM、またはNGC PAMに対する特異性を有
する、実施形態75～77のいずれか1つの改変された細胞。

79．SpCas9ドメインがアミノ酸置換L1111R, D1135V, G1218R, E1219F, A1322
R, R1335V, T1337R、ならびに L1111, D1135L, S1136R, G1218S, E1219V, D13
32A, R1335Q, T1337I, T1337V, T1337F, および T1337Mのうちの1つ以上、また
はその対応するアミノ酸置換を含む、実施形態75～78のいずれか1つの改変された細胞。

80．SpCas9ドメインがアミノ酸置換L1111R, D1135V, G1218R, E1219F, A1322
R, R1335V, T1337R、ならびにL1111, D1135L, S1136R, G1218S, E1219V, D13
32A, D1332S, D1332T, D1332V, D1332L, D1332K, D1332R, R1335Q, T1337I,
T1337V, T1337F, T1337S, T1337N, T1337K, T1337R, T1337H, T1337Q, およ
びT1337Mのうちの1つ以上、またはその対応するアミノ酸置換を含む、実施形態75～7
8のいずれか1つの改変された細胞。

81．SpCas9ドメインがアミノ酸置換D1135L, S1136R, G1218S, E1219V, A1322R
, R1335Q, T1337, および A1322R、ならびにL1111, D1135L, S1136R, G1218S,
E1219V, D1332A, D1332S, D1332T, D1332V, D1332L, D1332K, D1332R, R13
35Q, T1337I, T1337V, T1337F, T1337S, T1337N, T1337K, T1337R, T1337H,
T1337Q, および T1337Mのうちの1つ以上、またはその対応するアミノ酸置換を含む、
実施形態75～78のいずれか1つの改変された細胞。

10

20

30

40

50

82. SpCas9ドメインが、アミノ酸置換D 1135 M、S 1136 Q、G 1218 K、E 1219 F、A 1322 R、D 1332 A、R 1335 E、およびT 1337 R、またはそれらの対応するアミノ酸置換を含む、実施態様75～78のいずれか1つの改変された細胞。

83. SpCas9ドメインが、NG PAM、NNG PAM、GAA PAM、GAT PAM、またはCAA PAMに対する特異性を有する実施形態75～76のいずれか1つの改変された細胞。

84. Cas9ドメインがアミノ酸置換E 480 K、E 543 K、およびE 1219 V、またはそれらの対応するアミノ酸置換を含む、実施形態83の改変された細胞。

85. Cas9ドメインがSaCas9ドメインを含む、実施態様72～74のいずれか1つの改変された細胞。

86. SaCas9ドメインがNNNRRT PAMに対する特異性を有する、実施形態85の改変された細胞。

87. SaCas9ドメインがNNGRRT PAMに対する特異性を有する、実施形態86の改変された細胞。

88. SaCas9ドメインがアミノ酸置換N 579 Aまたはその対応するアミノ酸置換を含む、実施態様85～88のいずれか1つの改変された細胞。

89. SaCas9ドメインがアミノ酸置換E 782 K、N 968 K、およびR 1015 H、またはそれらの対応するアミノ酸置換を含む、実施態様85～89のいずれか1つの改変された細胞。

。

90. Cas9ドメインがSt1Cas9ドメインを含む、実施態様72～74のいずれか1つの改変された細胞。

91. St1Cas9ドメインがNNACCA PAMに対する特異性を有する、実施形態90の改変された細胞。

92. 前記アデノシンデアミナーゼドメインが、天然には存在しない改変アデノシンデアミナーゼドメインである、実施形態66～91のいずれか1つの改変された細胞。

93. 前記アデノシンデアミナーゼドメインがTadAドメインを含む、実施形態92の改変された細胞。

94. 前記TadAドメインがTad A7.10のアミノ酸配列を含む、実施形態93の改変された細胞。

95. 前記塩基エディターシステムがジンクフィンガードメインをさらに含む、実施態様66～94のいずれか1つの改変された細胞。

96. 前記ジンクフィンガードメインが、認識ヘリックス配列RNEHLEV、QSTTLKR、およびRTEHLAR、または認識ヘリックス配列RGEHLRQ、QSGTLKR、およびRNDKLVpを含む、実施形態95の改変された細胞。

97. 前記ジンクフィンガードメインがzf1raまたはzf1rbである、実施形態95または96の改変された細胞。

98. 前記塩基エディターシステムが核局在化シグナル (NLS) をさらに含む、実施形態66～97のいずれか一つの改変された細胞。

99. 前記塩基エディターシステムが1以上のリンカーをさらに含む、実施形態66～98のいずれか1つの改変された細胞。

100. ポリヌクレオチドプログラム可能なDNA結合ドメイン、アデノシンデアミナーゼドメイン、ジンクフィンガードメイン、およびNLSのうちの二つ以上がリンカーを介して連結されている、実施形態99の改変された細胞。

101. 前記リンカーがペプチドリンカーであり、それによって塩基エディター融合タンパク質を形成する、実施形態100の改変された細胞。

102. 前記ペプチドリンカーが、SGGSSGSETPGTSESATPESSGGS, SGGSSGGSSGSETPGTSESATPESSGGSSGGS, GSGSGSPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSA PGSPAGSPTSTEEGTSTEPSEGSA PGSTSTEPSEGSA PGSTSESATPESGPGSEPATSGGSSGGS, SGGSSGGSSGSETPGTSESATPES, SGGSSGGSSGSETPGTSESATPESSGGSSGGSSGGSSGGS, SGGSSGGSSGSETPGTSESATPESSGGSSGGSSGGSSGGSSGSETPGTSESATPESSGGS SGGs, PGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSA PGSPAGSPTSTEEGTSTEPSE

10

20

30

40

50

GSAP GTSTEPSEGSAPGTSESATPESGPGSEPATS, (SGGS)n, (GGGS)n, (GGGGS)n, (G)n, (EAAAK)n, (GGG)n, SGSETPGTSESATPES, および (XP)nからなる群より選択されるアミノ酸配列を含む、実施形態101の改変された細胞。

103. 前記塩基エディター融合タンパク質が、

MPKKKRKVSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGR
HDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAK
TGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDG
GSSGGSSGSETPGTSESATPESSGGSSGSSGSEVEFSHEYWMRHALTLAKRARDEREVPVGA
VLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCA
GAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFR
MPRQVFNAQKKAQSSTDGSSGGSSGSETPGTSESATPESDLVLGLAIGISVGVGILNK
VTGEIIHKNSRIFPAAQAENNLVRRRTNRQGRRLARRKKHRRVRLNRLFEEESGLITDFTKISI
NLNPYQLRVKGLTDELSNEELFIALKNMVKHGRGISYLLDDASDDGNSSVGDYAQIVKENS
KQLETKTPGQIQRLERYQTYGQLRGDFTVEKDGGKHRLINVFPTSAYRSEALRILQTQQEFN
PQITDEFINRYLEILTGKRKYHGPNEKSRTDYGRYRTSGETLDNIFGILIGKCTFYPDEF
RAAKASYTAQEFNLLNDLNNLTVPPTETKLSKEQKNQIINYVKNEKAMGPAKLFKYIAKL
LSCDVADIKGYRIDKSGKAEIHTFEAYRKMKTLETLDIEQMDRETLDKLAYVLTNTEREG
IQEALHEFADGSFSQKQVDELVQFRKANSSIFGKGWHNFSVKLMMELIPELYETSEEQ
MTILRLGKQKTTSSSNKTKYIDEKLLTEEIYNPVVAKSVRQAIKIVNAAIKEYGDFDNIVIE
MARETNEDDEKKAIQKIQKANKDEKDAAMLKAANQYNGKAELPHSVFHGHKQLATKIRL
WHQQGERCLYTGTISIHDLINNSNQFEVDHILPLSITFDDSLANKVLVYATANQEKQQR
TPYQALDSMDDAWSFRELKAFVRESKTLNKKKEYLLTEEDISKFDVRKKFIERNLVDL
YASRVVLNALQEHFRAHKIDTKVSVVRGQFTSQLRRHWGIEKTRDTYHHHAVDALIIAAS
SQLNLWKKQKNTLVSYSEDQLLDIETGELISDDEYKESVFKAPYQHFDVTLKSKEFEDSIL
FSYQVDSKFNKISDATIYATRQAKVGKDKADETYVLGKIKDIYTQDGYDAFMKIYKKDK
SKFLMYRHDPQTFEKVIEPILENYPNKQINDKGKEVPCNPFLKYKEEHGYIRKYSKKGNG
PEIKSLKYYDSKLGNHIDITPKDSNNKVVLQSVSPWRADVFNKTTGKYEILGLKYADLQ
FDKGTGTYKISQEKYNDIKKKEGVDSSEFKFTLYKNDLLLVKDTETKEQQLFRFLSRTMP
KQKHVELKPYDKQKFEGGEALIKVLGNVANSGQCKKGLGKSNISYKVRTDVLGNQHII
KNEGDKPKLDFPKKKRKVEGADKRTADGSEFESPKKKRKV、

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAH
AEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAGS
LMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDGSSGGSS
GSETPGTSESATPESSGGSSGSSGSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNN
RVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHS
RIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVF
NAQKKAQSSTDGSSGGSSGSETPGTSESATPESSGGSSGSSGSKRNYILGLAIGITSVGYGII
DYETRVIDAGVRLFKEANVENNEGRRSKRGARRLKRRRRHRIQRVKLLFDYNLLTDHS
ELSGINPYEARVKGLSQKLSEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNS
KALEEKYVAELQLERLKKDGEVRGSINRFKTSYVKEAKQLLVQKAYHQLDQSFIDTYI
DLLETRRTYYEGPGEKSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALND
LNNLVITRDENEKLEYEKFQIENVFQKQKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEF
TNLKVYHDIKDITARKEIENAELLDQIAKILTIIYQSSEDIQEELTNLNSLTQEEIEQISNL
KGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPPKKVDLSQQKEIPTTLVDDFILS
PVVKRSFIQSIKVINAIKKYGLPNDIIIELAREKNSKDAQKMINEMQKRNRQTNERIEEIIIR
TTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNNPFNYEVDHIIIPRSVSFDNSFNN
KVLVKQEENSCKGNRTPFQYLSSSDSKISYETFKKHILNLAAGKGRISKTKEYLLEERDIN
RFSVQKDFINRNLVDTRYATRGLMNLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKE
RNKGKHHHAEDALIIANADFIFKEWKKLDKAKKVMENQMFEKQAESMPEIETEYQYKEI
FITPHQIKHIKDFDKYKSHRVDKKPNRKLINDTLSTRKDDKGNTLIVNNLNGLYDKDN

10

20

30

40

50

DKLKKLINKSPEKLLMYHHDPPQTYQKLKLIMEQYGDEKNPLYKYYEETGNYLTKYSKKD
NGPVIKKIKYYGNKLNALHDITDDYPNSRNKVVKLSLKPYPYFDVYLDNGVYKFVTVKNLD
VIKKENYYEVNSKCYEEAKKLKKISNQAEFIASFYKNDLIKINGELYRVIGVNNDLLNRIEV
NMIDITYREYLENMNDKRPPHIIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKGEKA
DKRTADGSEFESPKKKRKV、

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAH
AEIMALRQGGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAGS
LMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAKKAQSSTDSSGSSGGSS
GSETPGTSESATPESSGGSSGSSSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNN
RVIGEGWNRRAIGLHDPTAHAEIMALRQGGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHS
RIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVF
NAQKKAQSSTDSSGSSGGSSGSETPGTSESATPESSGGSSGSSGSKRNYILGLAIGITSVGYGII
DYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRRRRRHRIQRVKKLLFDYNLLTDHS
ELSGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNS
KALEEKYVAELQLERLKKDGEVRGSINRFKTSYVKEAKQLLKVKQAYHQLDQSFIDTYI
DLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALND
LNNLVITRDENEKLEYEYEFQIIENVFKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEF
TNLKVYHDIKDITARKEIENAELLQIAKILTIIYQSSEDIQEELTNLNSLTQEEIEQISNL
KGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPPKKVDLSQQKEIPTTLVDDFILS
PVVKRSFIQSIKVINAIKKYGLPNDIIIELAREKNSKDAQKMINEMQKRNRTNERIEEIIIR
TTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNPNFNYEVDHIIPRSVSFDNSFNN
KVLVKQEENSCKGNRTPFQYLSSSDSKISYETFKKHILNLAGKGRISKTKEYLLEERDIN
RFSVQKDFINRNLVDTRYATRGLMNLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKE
RNKGYKHAEDALIANADFIKWKLDKAKKVMENQMFEKKAESMPEIETEQEYKEI
FITPHQIKHIKDFDKYKSHRVDKPKNRKLINDTLSTRKDDKGNTLIVNNLNGLYDKDN
DKLKKLINKSPEKLLMYHHDPPQTYQKLKLIMEQYGDEKNPLYKYYEETGNYLTKYSKKD
NGPVIKKIKYYGNKLNALHDITDDYPNSRNKVVKLSLKPYPYFDVYLDNGVYKFVTVKNLD
VIKKENYYEVNSKCYEEAKKLKKISNQAEFIASFYKNDLIKINGELYRVIGVNNDLLNRIEV
NMIDITYREYLENMNDKRPPHIIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKGEKA
DKRTADGSEFESPKKKRKVSSGNSNANSRGPSSFSSGLVPLSLRGSHSRPGERPFQCRICMR
NFSRNEHLEVHTRTHTGEKPFQCRICMRNFSQSTTLKRHLRTHHTGEKPFQCRICMRNFSR
TEHLARHLKTHLRGSSAQ、または

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAH
AEIMALRQGGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAGS
LMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAKKAQSSTDSSGSSGGSS
GSETPGTSESATPESSGGSSGSSSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNN
RVIGEGWNRRAIGLHDPTAHAEIMALRQGGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHS
RIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVF
NAQKKAQSSTDSSGSSGGSSGSETPGTSESATPESSGGSSGSSGSKRNYILGLAIGITSVGYGII
DYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRRRRRHRIQRVKKLLFDYNLLTDHS
ELSGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNS
KALEEKYVAELQLERLKKDGEVRGSINRFKTSYVKEAKQLLKVKQAYHQLDQSFIDTYI
DLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALND
LNNLVITRDENEKLEYEYEFQIIENVFKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEF
TNLKVYHDIKDITARKEIENAELLQIAKILTIIYQSSEDIQEELTNLNSLTQEEIEQISNL
KGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPPKKVDLSQQKEIPTTLVDDFILS
PVVKRSFIQSIKVINAIKKYGLPNDIIIELAREKNSKDAQKMINEMQKRNRTNERIEEIIIR
TTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNPNFNYEVDHIIPRSVSFDNSFNN
KVLVKQEENSCKGNRTPFQYLSSSDSKISYETFKKHILNLAGKGRISKTKEYLLEERDIN
RFSVQKDFINRNLVDTRYATRGLMNLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKE

10

20

30

40

50

RNKG YKHHAEDALI IANADFIFKEWKKLDKAKKVMENQMFE EKQAESMPEIETE QEYKEI
 FITPHQIKHIKDFKDYKYSHRVDKKPNRKLINDTL YSTRKDDKGNTLIVNNLNGLYDKDN
 DKLKKLINKSPEKLLMYHHDPQTYQKLKLIMEQY GDEKNPLYKY YEETGNYLT KYSKKD
 NGPVIKKIKYYGNKLNAHL DITDDYPNSRNKVVKLSLKP YRFDVYLDNGVYKFVTVKNLD
 VIKKENYYEVNSKCYEEAKKLK KISNQAEFIASFYKNDLIKINGELYRVIGVNNDLLNRIEV
 NMIDITYREYLENMNDKRPPHIIKT IASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKGEGA
 DKRTADGSEFESPKKKRKVSSGNSNANSRGP SFSSGLVPLSLRGSHSRPGERPFQCRICMR
 NFSRGEHLRQHTRTHTGEKPFQCRICMRNFSQSGTLKRHLRTHTGEKPFQCRICMRNFSR
 NDKLVP HLKTHLRGSSAQ

からなる群より選択されるアミノ酸配列を含む、実施形態101または102の改変された細胞。

10

104. 前記HBBポリヌクレオチドは、疾患の原因となる病原性一塩基多型 (SNP) を含む、実施形態66~103のいずれか一つの改変された細胞。

105. 疾患が鎌状赤血球症である、実施形態104の改変された細胞。

106. 前記HBBポリヌクレオチドは、病原性SNPから生じたアミノ酸変異を含むヘモグロビンのβサブユニット (HbB) をコードする、実施形態105の改変された細胞。

107. 前記脱アミノ化の結果、前記アミノ酸変異が良性アミノ酸で置換され、ここで前記良性アミノ酸はHbBの野生型アミノ酸とは異なる、実施形態106の改変された細胞。

108. 前記アミノ酸変異が位置6またはその対応する位置にある、実施形態106または107の改変された細胞。

20

109. 前記アミノ酸変異が、グルタミン酸 (E) からバリン (V) への、位置6 (E6V) またはその対応する位置での変異である、実施形態108の改変された細胞。

110. 前記脱アミノ化が位置6またはその対応する位置におけるE6V変異のアラニンによる置換をもたらす、実施形態109の改変された細胞。

111. 前記脱アミノ化が位置17またはその対応する位置である、実施形態110の改変された細胞。

112. 前記ガイドポリヌクレオチドが2つの個別のポリヌクレオチドを含み、前記2つの個別のポリヌクレオチドは2つのDNA、2つのRNA、またはDNAとRNAである、実施形態66~111のいずれか1つの改変された細胞。

113. 前記ガイドポリヌクレオチドがcrRNAおよびtracrRNAを含み、前記crRNAは前記HBBポリヌクレオチド中の標的配列に相補的な核酸配列を含む、実施形態66~112のいずれか1つの改変された細胞。

30

114. 前記標的配列が位置17またはその対応する位置を含む、実施形態113の改変された細胞。

115. 前記塩基エディターシステムが単一ガイドRNA (sgRNA) を含む、実施形態113または114の改変された細胞。

116. 前記sgRNAが、CUUCUCCACAGGAGUCAGAU, ACUUCUCCACAGGAGUCAGAU, GACUUCUCCACAGGAGUCAGAU, GUUUUUGUACUCUCAAGAUUUAAGUAACUGUACAACGAAACUACACAGUUACUUAUAAUUCUUGCAGAAAGCUACAAAGAUAAAGGCUUCAUGCCGAAAUCAACACCCUGUCAUUUUUAUGGCAGGGUG, CUUCUCCACAGGAGUCAGAUUUUUUGUACUCUCAAGAUUUAAGUAACUGUACAACGAAACUACACAGUUACUUAUAAUUCUUGCAGAAAGCUACAAAGAUAAAGGCUUCAUGCCGAAAUCAACACCCUGUCAUUUUUAUGGCAGGGUG; およびGACUUCUCCACAGGAGUCAGAUUUUUUGUACUCUCAAGAUUUAAGUAACUGUACAACGAAACUACACAGUUACUUAUAAUUCUUGCAGAAAGCUACAAAGAUAAAGGCUUCAUGCCGAAAUCAACACCCUGUCAUUUUUAUGGCAGGGUGからなる群より選択される配列を含む、実施形態115の改変された細胞。

40

117. 単一ガイドRNA (sgRNA) と;

50

MPKKKRKVSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGR
HDPTAHAEIMALRQGGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAK
TGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDG
GSSGGSSGSETPGTSESATPESSGGSSGSSEVEFSHEYWMRHALTLAKRARDEREVPVGA
VLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGGLVMQNYRLIDATLYVTLEPCVMCA
GAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFR
MPRQVFNAQKKAQSSTDGSSGGSSGSETPGTSESATPESDLVLGLAIGIGSVGVGILNK
VTGEIIHKNSRIFPAAQAENNLVRRTNRQGRRLARRKKHRRVRLNRLFEESSLITDFTKISI
NLNPYQLRVKGLTDELSNEELFIALKNMVKHGRISYLDASDDGNSSVGDYAQIVKENS
QLETKTPGQIQLERYQTYGQLRGDFTVEKDGGKHRLINVFPTSAYRSEALRILQTQQEFN
PQITDEFINRYLEILTGKRKYHGPNGEKSRTDYGRYRTSGETLDNIFGILIGKCTFYPDEF
RAAKASYTAQEFNLLNDLNNLTVPETETKLSKEQKNQIINYVKNEKAMGPAKLFKYIAKL
LSCDVADIKGYRIDKSGKAEIHTFEAYRKMKTLETLDIEQMDRETLDKLAYVLTNTEREG
IQEALEHEFADGSFSQKQVDELVQFRKANSSIFGKGWHNFSVKLMMELIPELYETSEEQ
M TILTRLGKQKTSSSNKTKYIDEKLLTEEIYNPVVAKSVRQAIKIVNAAIKEYGDFDNIVIE
MARETNEDDEKKAIQKIQKANKDEKDAAMLKAANQYNGKAELPHSVFHGHKQLATKIRL
WHQQGERCLYTGTISIHDLINNSNQFEVDHILPLSITFDDSLANKVLVYATANQEKQQR
TPYQALDSMDDAWSFRELKAFVRESKTLNKKKEYLLTEEDISKFDVRKKFIERNLVDL
YASRVVLNALQEHFRAHKIDTKVSVVRGQFTSQLRRHWGIEKTRDTHHHAVDALIIAAS
SQLNLWKKQKNTLVSYSEDQLLDIETGELISDDEYKESVFKAPYQHFDVTLKSKEFEDSIL
FSYQVDSKFNKRISDATIYATRQAKVGKDKADETYVLGKIKDIYTQDGYDAFMKIYKKDK
SKFLMYRHDPQTFEKVIEPILENYPNKQINDKGKEVPCNPFLKYKEEHGYIRKYSKKGNG
PEIKSLKYYDSKLGHNHIDITPKDSNNKVVLQSVSPWRADVFNKTTGKYEILGLKYADLQ
FDKGTGTYKISQEKYNDIKKKEGVDSSEFKFTLYKNDLLLVDKTETKEQQFLRFLSRTMP
KQKHVELKPYDKQKFEGGEALIKVLGNVANSGQCKKGLGKSNISYKVRTDVLGNQHII
KNEGDKPKLDFPKKKRKVEGADKRTADGSEFESPKKKRKV、

10

20

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAH
AEIMALRQGGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAGS
LMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDGSSGGSS
GSETPGTSESATPESSGGSSGSSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNN
RVIGEGWNRAIGLHDPTAHAEIMALRQGGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHS
RIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVF
NAQKKAQSSTDGSSGGSSGSETPGTSESATPESSGGSSGSKRNYILGLAIGITSVGYGII
DYETRVIDAGVRLFKEANVENNEGRRSKRGARRLRKRRRRHRIQRVKKLLFDYNLLTDHS
ELSGINPYEARVKGLSQKLSEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNS
KALEEKYVAELQLERLKKDGEVRGSINRFKTSYVKEAKQLLKVKAYHQLDQSFIDTYI
DLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALND
LNNLVITRDENEKLEYEYEFQIENVFKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEF
TNLKVYHDIKDITARKEIENAELLDQIAKILTIIYQSSEDIQEELTNLNSLTQEEIEQISNL
KGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILS
PVVKRSFIQSIKVINAIIKKYGLPNDIIELAREKNSKDAQKMINEMQKRNRQTNERIEEII
RTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNNPFNYEVDHIIPRSVSFDNSFNN
KVLVKQEENSCKGNRTPFQYLSSSDSKISYETFKKHILNLAGKGRISKTKEYLLEERDIN
RFSVQKDFINRNLVDTRYATRGLMNLRSYFRVNNLDVKVKSINGGFTSFLRRKWFKKE
RNKGKYYHAEDALIIANADFIFKEWKKLDKAKKVMENQMFEKQAESMPEIETEYQYKEI
FITPHQIKHIKDFKDYKYSHRVDKKPNRKLINDTLYSTRKDDKGNTLIVNNLNGLYDKDN
DKLKKLINKSPEKLLMYHHDPTQYQKLKLIMEQYGDEKNPLYKYYEETGNLYTKYSKKD
NGPVIKKIKYYGNKLNAHLDDYDYPNSRNKVVKLSLKPYRFDVYLDNGVYKFVTVKNLD
VIKKENYYEVNSKCYEEAKKLKISNQAEFIASFYKNDLIKINGELYRVIGVNNDLLNRIEV
NMIDITYREYLENMNDKRPPHIIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKGEA

30

40

50

DKRTADGSEFESP KKKRKV、

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAH
AEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAGS
LMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSSGSSGGSS
GSETPGTSESATPESSGGSSGSSGSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNN
RVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHS
RIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVF
NAQKKAQSSTDSSGSSGGSSGSETPGTSESATPESSGGSSGSSGSKRNYILGLAIGITSVGYGII
DYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRRIIRHRIQRVKKLLFDYNLLTDHS
ELSGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNS
KALEEKYVAELQLERLKKDGEVRGSINRFKTSYVKEAKQLLKVKQKAYHQLDQSFIDTYI
DLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALND
LNNLVITRDENEKLEYEKFQIENVFVKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEF
TNLKVYHDIKDITARKEIENAELLDDQIAKILTIYQSSEDIQEELTNLNLSELTQEEIEQISNL
KGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILS
PVVKRSFIQSIKVINAIKKYGLPNDIIIELAREKNSKDAQKMINEMQKRNRQTNERIEEIIIR
TTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNPNFNYEVDHIIPRSVSFDNSFNN
KVLVKQEENSCKGNRTPFQYLSSSDSKISYETFKKHILNLAAGKGRISKTKEYLLEERDIN
RFSVQKDFINRNLVDTRYATRGLMNNLLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKE
RNKGYKHHAEADALIANADFIFKEWKKLDKAKKVMENQMFEKKAESMPEIETE QEYKEI
FITPHQIKHIKDFKDYKYSHRVDKKPNRKLINDTLYSTRKDDKGNTLIVNNLNGLYDKDN
DKLKKLINKSPEKLLMYHHPQTYQKLKLIMEQYGDEKNPLYKYYEETGNYLT KYSKKD
NGPVIKKIKYYGNKLNAHLDDITDDYPNSRNKVVKLSLKPYPYRFDVYLDNGVYKFVTVKNLD
VIKKENYYEVNSKCYEEAKKLKISNQAEFIASFYKNDLIKINGELYRVIGVNNDLLNRIEV
NMIDITYREYLENMNDKRPPHIIKTIASTQSIKKYSTDILGNLYEVKSKKHPQIIKKGEGA
DKRTADGSEFESP KKKRKVSSGNSNANSRGPSFSSGLVPLSLRGSHSRPGERPFQCRICMR
NFSRNEHLEVHTRTHTGEKPFQCRICMRNFSQSTTLKRHLRTHHTGEKPFQCRICMRNFSR
TEHLARHLKTHLRGSSAQ、または

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAH
AEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAGS
LMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSSGSSGGSS
GSETPGTSESATPESSGGSSGSSGSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNN
RVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHS
RIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVF
NAQKKAQSSTDSSGSSGGSSGSETPGTSESATPESSGGSSGSSGSKRNYILGLAIGITSVGYGII
DYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRRIIRHRIQRVKKLLFDYNLLTDHS
ELSGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNS
KALEEKYVAELQLERLKKDGEVRGSINRFKTSYVKEAKQLLKVKQKAYHQLDQSFIDTYI
DLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALND
LNNLVITRDENEKLEYEKFQIENVFVKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEF
TNLKVYHDIKDITARKEIENAELLDDQIAKILTIYQSSEDIQEELTNLNLSELTQEEIEQISNL
KGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILS
PVVKRSFIQSIKVINAIKKYGLPNDIIIELAREKNSKDAQKMINEMQKRNRQTNERIEEIIIR
TTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNPNFNYEVDHIIPRSVSFDNSFNN
KVLVKQEENSCKGNRTPFQYLSSSDSKISYETFKKHILNLAAGKGRISKTKEYLLEERDIN
RFSVQKDFINRNLVDTRYATRGLMNNLLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKE
RNKGYKHHAEADALIANADFIFKEWKKLDKAKKVMENQMFEKKAESMPEIETE QEYKEI
FITPHQIKHIKDFKDYKYSHRVDKKPNRKLINDTLYSTRKDDKGNTLIVNNLNGLYDKDN
DKLKKLINKSPEKLLMYHHPQTYQKLKLIMEQYGDEKNPLYKYYEETGNYLT KYSKKD
NGPVIKKIKYYGNKLNAHLDDITDDYPNSRNKVVKLSLKPYPYRFDVYLDNGVYKFVTVKNLD

10

20

30

40

50

VIKKENYYEVNSKCYEEAKKLKKISNQAEFIASFYKNDLIKINGELYRVIGVNNDLLNRIEV
NMIDITYREYLENMNDKRPPHIIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKGEGA
DKRTADGSEFESPKKKRKVSSGNSNANSRGPSSFSSGLVPLSLRGSHSRPGERPFQCRICMR
NFSRGEHLRQHTRTHTGEKPFQCRICMRNFSQSGTLKRHLRTHHTGEKPFQCRICMRNFSR
NDKLVPHLKTHLRGSSAQ

からなる群から選択されるアミノ酸配列を含む融合タンパク質と
を含む塩基エディターシステムを含む、改変された細胞であって、

前記sgRNAは、前記細胞中のHBBポリヌクレオチドの位置17においてA・TからG・Cへの改変をもたらすために前記塩基エディターシステムをターゲティングすることができ、
前記sgRNAは、CUUCUCCACAGGAGUCAGAU, ACUUCUCCACAGGAGUCAGAU, GACU
UCUCCACAGGAGUCAGAU, GUUUUUGUACUCUCAAGAUUUAAGUAACUGUACAACGA
AACUUACACAGUUACUAAAUCUUGCAGAAGCUACAAAGAUAAAGGCUUCAUGCCGAA
AUCAACACCCUGUCAUUUUUAUGGCAGGGUG, CUUCUCCACAGGAGUCAGAUUUUUU
GUACUCUCAAGAUUUAAGUAACUGUACAACGAAACUUACACAGUUACUAAAUCUUG
CAGAAGCUACAAAGAUAAAGGCUUCAUGCCGAAAUCAACACCCUGUCAUUUUUAUGGCAG
GGUG, ACUUCUCCACAGGAGUCAGAUUUUUUGUACUCUCAAGAUUUAAGUAACUGU
ACAACGAAACUUACACAGUUACUAAAUCUUGCAGAAGCUACAAAGAUAAAGGCUUCA
UGCCGAAAUCAACACCCUGUCAUUUUUAUGGCAGGGUG; および GACUUCUCCACAGGA
GUCAGAUUUUUUGUACUCUCAAGAUUUAAGUAACUGUACAACGAAACUUACACAGU
UACUAAAUCUUGCAGAAGCUACAAAGAUAAAGGCUUCAUGCCGAAAUCAACACCCUGU
CAUUUUUAUGGCAGGGUGからなる群より選択される配列を含み、前記細胞は赤血球である、

改変された細胞。

118. ガイドポリヌクレオチドまたはそれをコードする核酸と;

ポリヌクレオチドプログラム可能なDNA結合ドメインまたはそれをコードする核酸と;
アデノシンデアミナーゼドメインまたはそれをコードする核酸と

を含む、塩基エディターシステムであって、

前記ガイドポリヌクレオチドは、HBBポリヌクレオチド中の核酸塩基においてA・TからG・Cへの改変をもたらすために前記塩基エディターシステムをターゲティングすることができるものである、

塩基エディターシステム。

119. 前記ポリヌクレオチドプログラム可能なDNA結合ドメインがCas9ドメインである、実施形態118の塩基エディターシステム。

120. 前記Cas9ドメインがヌクレアーゼ不活性Cas9ドメインである、実施形態119の改変された細胞。

121. 前記Cas9ドメインがCas9ニッカーゼドメインである、実施形態119の塩基エディターシステム。

122. 前記Cas9ドメインがSpCas9ドメインを含む、実施形態119~121のいずれか1つの塩基エディターシステム。

123. 前記SpCas9ドメインがD10Aおよび/またはH840Aアミノ酸置換またはそれらの対応するアミノ酸置換を含む、実施形態122の塩基エディターシステム。

124. 前記SpCas9ドメインがNGG PAMに対する特異性を有する、実施形態122または123の塩基エディターシステム。

125. 前記SpCas9ドメインがNGA PAM、NGT PAM、またはNGC PAMに対する特異性を有する、実施形態122~124のいずれか1つの塩基エディターシステム。

126. 前記SpCas9ドメインが、アミノ酸置換L1111R, D1135V, G1218R, E1219F, A1322R, R1335V, T1337R、ならびに L1111, D1135L, S1136R, G1218S, E1219V, D1332A, R1335Q, T1337I, T1337V, T1337F, および T1337Mのうちの1つ以上、またはその対応するアミノ酸置換を含む、実施形態122~125のいずれか1つの塩基編集システム。

127. 前記SpCas9ドメインが、アミノ酸置換L1111R, D1135V, G1218R, E1219F, A1322R, R1335V, T1337R、ならびにL1111, D1135L, S1136R, G1218S, E1219V, D1332A, D1332S, D1332T, D1332V, D1332L, D1332K, D1332R, R1335Q, T1337I, T1337V, T1337F, T1337S, T1337N, T1337K, T1337R, T1337H, T1337QおよびT1337Mのうちの1つ以上、またはその対応するアミノ酸置換を含む、実施形態122~125のいずれか1つの塩基エディターシステム。

128. 前記SpCas9ドメインが、アミノ酸置換D1135L, S1136R, G1218S, E1219V, A1322R, R1335Q, T1337, および A1322R、ならびにL1111, D1135L, S1136R, G1218S, E1219V, D1332A, D1332S, D1332T, D1332V, D1332L, D1332K, D1332R, R1335Q, T1337I, T1337V, T1337F, T1337S, T1337N, T1337K, T1337R, T1337H, T1337Q, および T1337Mのうちの1つ以上、またはその対応するアミノ酸置換を含む、実施形態122~125のいずれか1つの塩基エディターシステム。

10

129. 前記SpCas9ドメインが、アミノ酸置換D1135M, S1136Q, G1218K, E1219F, A1322R, D1332A, R1335E、およびT1337R、またはそれらの対応するアミノ酸置換を含む、実施形態122~125のいずれか1つの塩基エディターシステム。

130. 前記SpCas9ドメインがNG PAM、NNG PAM、GAA PAM、GAT PAM、またはCAA PAMに対する特異性を有する、実施形態122~123のいずれか1つの塩基エディターシステム。

131. 前記Cas9ドメインがアミノ酸置換E480K, E543K、およびE1219V、またはそれらの対応するアミノ酸置換を含む、実施形態130の塩基エディターシステム。

20

132. 前記Cas9ドメインがSaCas9ドメインを含む、実施形態119~121のいずれか1つの塩基エディターシステム。

133. 前記SaCas9ドメインは、NNNRRT PAMに対する特異性を有する、実施形態132の塩基エディターシステム。

134. 前記SaCas9ドメインは、NNGRRT PAMに対する特異性を有する、実施形態133の塩基エディターシステム。

135. 前記SaCas9ドメインは、アミノ酸置換N579Aまたはその対応するアミノ酸置換を含む、実施形態132~134のいずれか1つの塩基エディターシステム。

136. 前記SaCas9ドメインがアミノ酸置換E782K, N968K、およびR1015H、またはそれらの対応するアミノ酸置換を含む、実施形態134~135のいずれか1つの塩基エディターシステム。

30

137. 前記Cas9ドメインがSt1Cas9ドメインを含む、実施形態119~121のいずれか1つの塩基エディターシステム。

138. 前記St1Cas9ドメインがNNACCA PAMに対する特異性を有する、実施形態137の塩基エディターシステム。

139. 前記アデノシンデアミナーゼドメインが、天然には存在しない改変されたアデノシンデアミナーゼドメインである、実施形態118~138のいずれか1つの塩基エディターシステム。

140. 前記アデノシンデアミナーゼドメインがTadAドメインを含む、実施形態139の塩基エディターシステム。

40

141. 前記TadAドメインは、TadA 7.10のアミノ酸配列を含む、実施形態140の塩基エディターシステム。

142. 前記塩基エディターシステムは、ジンクフィンガードメインをさらに含む、実施形態118~138のいずれか1つの塩基エディターシステム。

143. 前記ジンクフィンガードメインが、認識ヘリックス配列RNEHLEV, QSTTLKR、およびRTEHLAR、または認識ヘリックス配列RGEHLRQ, QSGTLKR、およびRNDKLVPを含む、実施形態142の塩基エディターシステム。

144. 前記ジンクフィンガードメインがzf1raまたはzf1rbである、実施形態142または143の塩基エディターシステム。

145. 前記塩基エディターシステムは、核局在化シグナル (NLS) をさらに含む、実施形

50

態118～138のいずれか1つの塩基エディターシステム。

146. 前記塩基エディターシステムは、1つ以上のリンカーをさらに含む、実施形態118～145のいずれか1つの塩基エディターシステム。

147. ポリヌクレオチドプログラム可能なDNA結合ドメイン、アデノシンデアミナーゼドメイン、ジンクフィンガードメイン、およびNLSのうちの二つ以上がリンカーを介して連結されている、実施形態146の塩基エディターシステム。

148. 前記リンカーがペプチドリンカーであり、それによって塩基編集融合タンパク質を形成する、実施形態147の塩基エディターシステム。

149. 前記ペプチドリンカーが、SGGSSGSETPGTSESATPESSGGS, SGGSSGGSSGSETPGTSESATPESSGGSSGGS, GSGSGSPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGGSGGS, SGGSSGGSSGSETPGTSESATPES, SGGSSGGSSGSETPGTSESATPESSGGSSGGSSGGSSGGS, SGGSSGGSSGSETPGTSESATPESSGGSSGGSSGGSSGGSSGSETPGTSESATPESSGGS, PGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGGSGGS, (SGGS)_n, (GGGS)_n, (GGGS)_n, (G)_n, (EAAAK)_n, (GGS)_n, SGSETPGTSESATPES, および (XP)_nからなる群より選択されるアミノ酸配列を含む、実施形態148の塩基エディターシステム。

150. 前記塩基編集融合タンパク質が、

MPKKKRKVSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAKKAQSSTDSSGGSSGSETPGTSESATPESSGGSSGGSSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVHNNRVIGEGWNRRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTDSSGGSSGSSGSETPGTSESATPESDLVLGLAIGIGSVGVGILNKVTGEIIHKNSRIFPAAQAENNLVRRNTNRQGRRLARRKKHRRVRLNRLFEEGLITDFTKISINLNPYQLRVKGLTDELSNEELFIALKNMVKHGHSYLDASDDGNSSVGDYAQIVKENSQLETKTPGQIQLERYQTYGQLRGDFTVEKDGKKHRLINVFPTSAYRSEALRILQTQQEFNPQITDEFINRYLEILTGKRKYHGPNGNEKSRTDYGRYRTSGETLDNIFGILIGKCTFYPDEFRAAKASYTAQEFNLLNDLNNLTVPPTETKKLSKEQKNQIINYVKNEKAMGPAKLFKYIAKLSCDVADIKGYRIDKSGKAEIHTFEAYRKMKLTLETLDIEQMDRETLDKLAYVLTNTEREGIQEALHEFADGSFSQKQVDELVQFRKANSSIFGKGWHNFSVKLMMELIPELYETSEEQMTILTRLGKQKTTSSSNKTKYIDEKLLTEEIYNPVVAKSVRQAIKIVNAAIKEYGDFDNIVIEMARETNEDDEKKAIQKIQKANKDEKDAAMLKAANQYNGKAELPHSVFHHGKQLATKIRLWHQQGERCLYTGTISIHDLINNSNQFEVDHILPLSITFDDSLANKVLVYATANQEKQRTPYQALDSMDDAWSFRELKAFVRESKTLNKKKEYLLTEEDISKFDVRKKFIERNLVDLYASRVVLNALQEHFRAHKIDTKVSVVRGQFTSQLRRHWGIEKTRDTYHHHAVDALIIAASQLNLWKKQKNTLVSYSEDQLLDIETGELISDDEYKESVFKAPYQHFDVTLKSKEFEDSILFSYQVDSKFNKRKISDATIYATRQAKVGKDKADETYVLGKIKDIYTQDGYDAFMKIYKKDKSKFLMYRHDPQTFEKVIENYPNPKQINDKGKEVPCNPFLKYKEEHGYIRKYSKKGNGPEIKSLKYYDSKLGNHIDITPKDSNNKVVLQSVSPWRADVFNKTTGKYEILGLKYADLQFDKGTGTYKISQEKYNDIKKKEGVDSSEFKFTLYKNDLLLVDKTETKEQQFLRFLSRTMPKQKHVELKPYDKQKFEGGEALIKVLGNVANSGQCKKGLGKSNISYKVRTDVLGNQHIIKNEGDKPKLDFPKKKRKVEGADKRTADGSEFESPKKKRKV、

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAKKAQSSTDSSGGSSGSETPGTSESATPESSGGSSGGSSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVHNNRVIGEGWNRRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVF

10

20

30

40

50

NAQKKAQSSTDSSGSSGGSSGSETPGTSESATPESSGGSSGGSKRNYILGLAIGITSVGYGII
DYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRRRRHRIQRVKKLLFDYNLLTDHS
ELSGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNS
KALEEKYVAELQLERLKKDGEVRGSINRFKTSYVKEAKQLLKVKQAYHQLDQSFIDTYI
DLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALND
LNNLVITRDENEKLEYEKFQIIENVFKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEF
TNLKVYHDIKDITARKEIENAELLDQIAKILTIIYQSSEDIQEELTNLNSELTQEEIEQISNL
KGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFELS
PVVKRSFIQSIKVINAIKKYGLPNDIIIELAREKNSKDAQKMINEMQKRNRQTNERIEEIIIR
TTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNPNFNYEVDHIIPRSVSFDNSFNN
KVLVKQEENSCKGNRTPFQYLSSSDSKISYETFKKHILNLAGKGRISKTKEYLLEERDIN
RFSVQKDFINRNLVDTRYATRGLMNLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKE
RNKGYKHAEDALIIANADFIFKEWKKLDKAKKVMENQMFEKQAESMPEIETE QEYKEI
FITPHQIKHIKDFKDYKYSHRVDKKPNRKLINDTLYSTRKDDKGNTLIVNNLNGLYDKDN
DKLKKLINKSPEKLLMYHHDPPQTYQKLKLIMEQYGDEKNPLYKYYEETGNYLTKEYSKKD
NGPVIKKIKYYGNKLNALHDITDDYPNSRNKVVKLSLKPYPYFDVYLDNGVYKFVTVKNLD
VIKENYYEVNSKCYEEAKKLKISNQAEFIASFYKNDLIKINGELYRVIGVNNDLLNRIEV
NMIDITYREYLENMNDKRPPHIIKTASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKGEA
DKRTADGSEFESPKKKRKV、

10

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAH
AEIMALRQGGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAGS
LMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAKKAQSSTDSSGSSGGSS
GSETPGTSESATPESSGGSSGGSSSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNN
RVIGEGWNRAIGLHDPTAHAEIMALRQGGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHS
RIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVF
NAQKKAQSSTDSSGSSGGSSGSETPGTSESATPESSGGSSGGSKRNYILGLAIGITSVGYGII
DYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRRRRHRIQRVKKLLFDYNLLTDHS
ELSGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNS
KALEEKYVAELQLERLKKDGEVRGSINRFKTSYVKEAKQLLKVKQAYHQLDQSFIDTYI
DLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALND
LNNLVITRDENEKLEYEKFQIIENVFKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEF
TNLKVYHDIKDITARKEIENAELLDQIAKILTIIYQSSEDIQEELTNLNSELTQEEIEQISNL
KGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFELS
PVVKRSFIQSIKVINAIKKYGLPNDIIIELAREKNSKDAQKMINEMQKRNRQTNERIEEIIIR
TTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNPNFNYEVDHIIPRSVSFDNSFNN
KVLVKQEENSCKGNRTPFQYLSSSDSKISYETFKKHILNLAGKGRISKTKEYLLEERDIN
RFSVQKDFINRNLVDTRYATRGLMNLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKE
RNKGYKHAEDALIIANADFIFKEWKKLDKAKKVMENQMFEKQAESMPEIETE QEYKEI
FITPHQIKHIKDFKDYKYSHRVDKKPNRKLINDTLYSTRKDDKGNTLIVNNLNGLYDKDN
DKLKKLINKSPEKLLMYHHDPPQTYQKLKLIMEQYGDEKNPLYKYYEETGNYLTKEYSKKD
NGPVIKKIKYYGNKLNALHDITDDYPNSRNKVVKLSLKPYPYFDVYLDNGVYKFVTVKNLD
VIKENYYEVNSKCYEEAKKLKISNQAEFIASFYKNDLIKINGELYRVIGVNNDLLNRIEV
NMIDITYREYLENMNDKRPPHIIKTASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKGEA
DKRTADGSEFESPKKKRKVSSGNSNANSRGPSSFSSGLVPLSLRGSHSRPGERPFQCRICMR
NFSRNEHLEVHTRTHTGEKPFQCRICMRNFSQSTTLKRHLRTHTGEKPFQCRICMRNFSR
TEHLARHLKTHLRGSSAQ、または

20

30

40

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAH
AEIMALRQGGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAGS
LMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAKKAQSSTDSSGSSGGSS
GSETPGTSESATPESSGGSSGGSSSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNN

50

RVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHS
RIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVF
NAQKKAQSSTDSSGSSGGSSGSETPGTSESATPESSGGSSGGSKRNYILGLAIGITSVGYGII
DYETRVIDAGVRLFKEANVENNEGRRSKRGARRLKRHHRIQRVKKLLFDYNLLTDHS
ELSGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNS
KALEEKYVAELQLERLKKDGEVRGSINRFKTSYVKEAKQLLKVKQAYHQLDQSFIDTYI
DLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALND
LNNLVITRDENEKLEYEKFQIIENVFKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEF
TNLKVYHDIKDITARKEIENAELLDQIAKILTIYQSSEDIQEELTNLNLSELTQEEIEQISNL
KGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILS
PVVKRSFIQSIKVINAIKKYGLPNDIIIELAREKNSKDAQKMINEMQKRNRQTNERIEEIIIR
TTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNPNFNYEVDHIIPRSVSFDNSFNN
KVLVKQEENSCKGNRTPFQYLSSSDSKISYETFKKHILNLAAGKGRISKTKEYLLEERDIN
RFSVQKDFINRNLVDTRYATRGLMNLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKE
RNKGKYYHAEDALIANADFIFKEWKKLDKAKKVMENQMFEKQAESMPEIETEQEYKEI
FITPHQIKHIKDFDKYKYSHRVDKKPNRKLINDTLSTRKDDKGNTLIVNNLNGLYDKDN
DKLKKLINKSPEKLLMYHHPQTYQKLKLIMEQYGDEKNPLYKYEEETGNYLTKYSKKD
NGPVIKKIKYYGNKLNAHLDDITDDYPNSRNKVVKLSLKPYPYRFDVYLDNGVYKFVTVKNLD
VIKKENYYEVNSKCYEEAKKLKISNQAEFIASFYKNDLIKINGELYRVIGVNNDLLNRIEV
NMIDITYREYLENMNDKRPPHIIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKGEA
DKRTADGSEFESPKKRKVSSGNSNANSRGPSSFSSGLVPLSLRGSHSRPGERPFQCRICMR
NFSRGEHLRQHTRTHTGEKPFQCRICMRNFSQSGTLKRHLRTHHTGEKPFQCRICMRNFSR
NDKLVPHLKTHLRGSSAQ

からなる群より選択されるアミノ酸配列を含む、実施形態148または149の塩基エディターシステム。

151. 前記HBBポリヌクレオチドが、疾患の原因となる病原性一塩基多型 (SNP) を含む、実施形態118~150のいずれか一項の塩基エディターシステム。

152. 前記疾患が鎌状赤血球症である、実施形態151の塩基エディターシステム。

153. 前記HBBポリヌクレオチドは、病原性SNPから生じたアミノ酸変異を含むヘモグロビンのβサブユニット (HbB) をコードする、実施形態152の塩基エディターシステム。

154. 脱アミノ化の結果、前記アミノ酸変異が良性アミノ酸で置換され、ここで前記良性アミノ酸はHbBの野生型アミノ酸とは異なる、実施形態153の塩基エディターシステム。

155. 前記アミノ酸変異は位置6またはその対応する位置にある、実施形態153または154の塩基エディターシステム。

156. 前記アミノ酸変異が、グルタミン酸 (E) からバリン (V) への位置6 (E6V) またはその対応する位置での変異である、実施形態155に記載の塩基エディターシステム。

157. 脱アミノ化が、E6V変異の位置またはその対応する位置でのアラニンによる置換をもたらす、実施形態156の塩基エディターシステム。

158. 脱アミノ化が位置17またはその対応する位置である、実施形態157の塩基エディターシステム。

159. 前記ガイドポリヌクレオチドが2つの個別のポリヌクレオチドを含み、前記2つの個別のポリヌクレオチドは2つのDNA、2つのRNA、またはDNAとRNAである、実施形態118~158のいずれか1つの塩基エディターシステム。

160. 前記ガイドポリヌクレオチドがcrRNAおよびtracrRNAを含み、前記crRNAはHBBポリヌクレオチド中の標的配列に相補的な核酸配列を含む、実施形態118~159のいずれか一項記載の塩基エディターシステム。

161. 前記標的配列が位置17またはその対応する位置を含む、実施形態160の塩基エディターシステム。

162. 前記塩基エディターシステムが単一ガイドRNA (sgRNA) を含む、実施形態160

10

20

30

40

50

または161の塩基エディターシステム。

163. 前記sgRNAがCUUCUCCACAGGAGUCAGAU, ACUUCUCCACAGGAGUCAGAU, GACUUCUCCACAGGAGUCAGAU, GUUUUUGUACUCUCAAGAUUUAAGUAACUGUACA ACGAAACUUACACAGUUACUUAUAUCUUGCAGAAGCUACAAAGAUAAAGGCUUCAUGC CGAAAUCAACACCCUGUCAUUUUUAUGGCAGGGUG, CUUCUCCACAGGAGUCAGAUGU UUUUGUACUCUCAAGAUUUAAGUAACUGUACAACGAAACUUACACAGUUACUUAUAU CUUGCAGAAGCUACAAAGAUAAAGGCUUCAUGCCGAAAUCAACACCCUGUCAUUUUUAU GGCAGGGUG, ACUUCUCCACAGGAGUCAGAUGUUUUUGUACUCUCAAGAUUUAAGUA ACUGUACAACGAAACUUACACAGUUACUUAUAUCUUGCAGAAGCUACAAAGAUAAAGG CUUCAUGCCGAAAUCAACACCCUGUCAUUUUUAUGGCAGGGUG; およびGACUUCUCCA CAGGAGUCAGAUGUUUUUGUACUCUCAAGAUUUAAGUAACUGUACAACGAAACUUAC ACAGUUACUUAUAUCUUGCAGAAGCUACAAAGAUAAAGGCUUCAUGCCGAAAUCAACAC CCUGUCAUUUUUAUGGCAGGGUGからなる群から選択される配列を含む、実施形態162の塩基エディターシステム。

10

164. 単一ガイドRNA (sgRNA)と、MPKKKRKVSEVEFSHEYWMRHALTLAKRAWDE REVPGAVLVHNNRVIGEGWNRPIGRHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTL EPCVMCAGAMIHSRIGRVVFGARDAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAA LLSDFFRMRRQEIKAKKAQSSTDSSGSSGGSSGSETPGTSESATPESSGGSSGSSSEVEFS HEYWMRHALTLAKRARDEREVPVGAVLVHNNRVIGEGWNRRAIGLHDPTAHAEIMALRQ GGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYP GMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTDSSGSSGGSSGSETPGTS ESATPESDLVLGLAIGIGSVGVGILNKVTGEIIHKNSRIFPAAQAENNLVRRRTNRQGRRLAR RKKHRRVRLNRLFEESSGLITDFTKISINLNPYQLRVKGLTDELSNEELFIALKNMVKHGRI SYLDDASDDGNSSVG DYAQIVKENSQKLETKTPGQIQQLERYQTYGQLRGDFTVEKDGGKH RLINVFPTSAYRSEALRILQTQQEFNPQITDEFINRYLEILTGKRKYHGPNGEKSRTDYG RYRTSGETLDNIFGILIGKCTFYPDEFRAAKASYTAQEFNLLNDLNNLTVPETETKLSKEQ KNQIINYVKNEKAMGPAKLFKYIAKLLSCDVADIKGYRIDKSGKAEIHTFEAYRKMKTLET LDIEQMDRETLDKLAYVLTNTEREGIQEALHEFADGSFSQKQVDELVQFRKANSSIFGK GWHNFSVKLMMELIPELYETSEEQMTILTRLGKQKTSSSNKTKYIDEKLLTEEIYNPVV AKSVRQAIKIVNAAIKEYGDFDNIVIEMARETNEDDEKKAIQKIQKANKDEKDAAMLKAA NQYNGKAELPHSVFHGHKQLATKIRLWHQQGERCLYTGKTISIHDLINNSNQFEVDHILP LSITFDDSLANKVLVYATANQEKQRTPYQALDSMDDAWSFRELKAFVRESKTLNKKK EYLLTEEDISKFDVRKKFIERNLVDTLYASRVVLNALQEHFRAHKIDTKVSVVRGQFTSQL RRHWGIEKTRDTYHHHAVDALIIAASSQLNLWKKQKNTLVSYSEDQLLDIETGELISDDE YKESVFKAPYQHFDVTLKSKEFEDSILFSYQVDSKFNRKISDATIYATRQAKVGKDKADET YVLGKIKDIYTQDGYDAFMKIYKKDKSKFLMYRHPQTFEKVIEPILENYPNKQINDKGK EVPCNPFLKYKEEHGYIRKYSKKGNGPEIKSLKYDSDKLGNHIDITPKDSNNKVVLQSVSP WRADVYFNKTTGKYEILGLKYADLQFDKGTGTYKISQEKYNDIKKKEGVDSDSEFKFTLY KNDLLLVKDTETKEQQLFRFLSRTMPKQKHVELKPYDKQKFEGGEALIKVLGNVANSG QCKKGLGKSNIYKVRTDVLGNQHIIKNEGDKPKLDFPKKKRKVEGADKRTADGSEFESP KKKRKV、

20

30

40

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAH AEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAGS LMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAKKAQSSTDSSGSSGGSS GSETPGTSESATPESSGGSSGSSSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVHNN RVIGEGWNRRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHS RIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVF NAQKKAQSSTDSSGSSGGSSGSETPGTSESATPESSGGSSGSSSKRNYILGLAIGITSVGYGII DYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLRRRRRHRIQRVKKLLFDYNLLTDHS ELSGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNS

50

KALEEKYVAELQLERLKKDGEVRGSINRFKTSYVKEAKQLLKVQKAYHQLDQSFIDTYI
DLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALND
LNNLVITRDENEKLEYEKFQIENVFKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEF
TNLKVYHDIKDITARKEIENAELLDQIAKILTIIYQSSEDIQEELTNLNSELTQEEIEQISNL
KGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPPKKVDLSQQKEIPTTLVDDFELS
PVVKRSFIQSIKVINAIKKYGLPNDIIIELAREKNSKDAQKMINEMQKRNRQTNERIEEIIIR
TTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNPNFNYEVDHIIIPRSVSFDNSFNN
KVLVKQEENSCKGNRTPFQYLSSSDSKISYETFKKHILNLAGKGRISKTKEYLLEERDIN
RFSVQKDFINRNLVDTRYATRGLMNLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKE
RNKGYKHHAEADALIANADFIFKEWKKLDKAKKVMENQMFEKKAESMPEIETEQEYKEI
FITPHQIKHIKDFKDYKYSHRVDKKPNRKLINDTLSTRKDDKGNTLIVNNLNGLYDKDN
DKLKKLINKSPEKLLMYHHPQTYQKLKLIMEQYGDEKNPLYKYYEETGNLYLTYSKKD
NGPVIKKIKYYGNKLNAHLDDITDDYPNSRNKVVKLSLKPYPYFDVYLDNGVYKFVTVKNLD
VIKKENYYEVNSKCYEEAKKLKISNQAEFIASFYKNDLIKINGELYRVIGVNNDLLNRIEV
NMIDITYREYLENMNDKRPPHIIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKGEGA
DKRTADGSEFESPKKKRKV、

10

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAH
AEIMALRQGGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAGS
LMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAKKAQSSTDSSGGSSGGSS
GSETPGTSESATPESSGGSSSGSSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNN
RVIGEGWNRAIGLHDPTAHAEIMALRQGGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHS
RIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVF
NAQKKAQSSTDSSGGSSGGSSGSETPGTSESATPESSGGSSGGSSKRNILGLAIGITSVGYGII
DYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLRRRRRHRIQRVKKLLFDYNLLTDHS
ELSGINPYEARVKGLSQKLSEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNS
KALEEKYVAELQLERLKKDGEVRGSINRFKTSYVKEAKQLLKVQKAYHQLDQSFIDTYI
DLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALND
LNNLVITRDENEKLEYEKFQIENVFKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEF
TNLKVYHDIKDITARKEIENAELLDQIAKILTIIYQSSEDIQEELTNLNSELTQEEIEQISNL
KGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPPKKVDLSQQKEIPTTLVDDFELS
PVVKRSFIQSIKVINAIKKYGLPNDIIIELAREKNSKDAQKMINEMQKRNRQTNERIEEIIIR
TTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNPNFNYEVDHIIIPRSVSFDNSFNN
KVLVKQEENSCKGNRTPFQYLSSSDSKISYETFKKHILNLAGKGRISKTKEYLLEERDIN
RFSVQKDFINRNLVDTRYATRGLMNLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKE
RNKGYKHHAEADALIANADFIFKEWKKLDKAKKVMENQMFEKKAESMPEIETEQEYKEI
FITPHQIKHIKDFKDYKYSHRVDKKPNRKLINDTLSTRKDDKGNTLIVNNLNGLYDKDN
DKLKKLINKSPEKLLMYHHPQTYQKLKLIMEQYGDEKNPLYKYYEETGNLYLTYSKKD
NGPVIKKIKYYGNKLNAHLDDITDDYPNSRNKVVKLSLKPYPYFDVYLDNGVYKFVTVKNLD
VIKKENYYEVNSKCYEEAKKLKISNQAEFIASFYKNDLIKINGELYRVIGVNNDLLNRIEV
NMIDITYREYLENMNDKRPPHIIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKGEGA
DKRTADGSEFESPKKKRKVSSGNSNANSRGPSSFSSGLVPLSLRGSHSRPGERPFQCRICMR
NFSRNEHLEVHTRTHTGEKPFQCRICMRNFSQSTTLKRHLRTHHTGEKPFQCRICMRNFSR
TEHLARHLKTHLRGSSAQ、または

20

30

40

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAH
AEIMALRQGGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAGS
LMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAKKAQSSTDSSGGSSGGSS
GSETPGTSESATPESSGGSSSGSSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNN
RVIGEGWNRAIGLHDPTAHAEIMALRQGGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHS
RIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVF
NAQKKAQSSTDSSGGSSGGSSGSETPGTSESATPESSGGSSGGSSKRNILGLAIGITSVGYGII

50

DYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRRRRRHRIQRVKKLLFDYNLLTDHS
ELSGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNS
KALEEKYVAELQLERLKKDGEVRGSINRFKTS DYVKEAKQLLKVQKAYHQLDQSFIDTYI
DLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALND
LNNLVITRDENEKLEYEKFQIIENVFKQKKKPTLQKQIAKEILVNEEDIKGYRVTSTGKPEF
TNLKVYHDIKDITARKEIENAELLDQIAKILTIIYQSSEDIQEELTNLNLSELTQEEIEQISNL
KGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPPKKVDLSQQKEIPTTLVDDFILS
PVVKRSFIQSIKVINAIKKYGLPNDIIIELAREKNSKDAQKMINEMQKRNRQTNERIEEIIIR
TTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNPNFNYEVDHIIPRSVSFDNSFNN
KVLVKQEENSKKGNRTPFQYLSSSDSKISYETFKKHILNLA KGKGRISKTKKEYLLEERDIN
RFSVQKDFINRNLVDTRYATRGLMNLRLSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKE
RNKGKHHHAEDALIINANADFIKFWKKLDKAKKVMENQMFEKQAESMPEIETE QEYKEI
FITPHQIKHIKDFKDYKYSHRVDKKPNRKLINDTLYSTRKDDKGNTLIVNNLNGLYDKDN
DKLKKLINKSPEKLLMYHHPQTYQKLKLIMEQYGDEKNPLYKYYEETGNLYLTKYSKKD
NGPVIKKIKYGGNKLNAHLDDYDPSNRNKVVKLSLKPYRFDVYLDNGVYKFVTVKNLD
VIKKENYYEVNSKCYEEAKKLKISNQAEFIASFYKNDLIKINGELYRVIGVNNDLLNRIEV
NMIDITYREYLENMNDKRPPHIIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKGE GA
DKRTADGSEFESPKKKRKVSSGNSNANSRGPSSFSSGLVPLSLRGSHSRPGERPFQCRICMR
NFSRGEHLRQHTRTHTGEKPFQCRICMRNFSQSGTLKRHLRTHTGEKPFQCRICMRNFSR
NDKLVPHLKTHLRGSSAQ

10

20

からなる群から選択されるアミノ酸配列を含む融合タンパク質と
を含む、塩基エディターシステムであって、

前記sgRNAは、細胞におけるHBBポリヌクレオチドの位置17においてA・TからG・Cへ
の改変をもたらすために前記塩基エディターシステムをターゲティングすることができ、
ここで前記sgRNAは、CUUCUCCACAGGAGUCAGAU, ACUUCUCCACAGGAGUCAGAU,
GACUUCUCCACAGGAGUCAGAU, GUUUUUGUACUCUCAAGAUUUAAGUAACUGUACA
ACGAAACUUACACAGUUACUAAAUCUUGCAGAAGCUACAAAGAUAAAGGCUUCAUGC
CGAAAUCAACACCCUGUCAUUUUUAUGGCAGGGUG, CUUCUCCACAGGAGUCAGAUGU
UUUUGUACUCUCAAGAUUUAAGUAACUGUACAACGAAACUUACACAGUUACUAAA
CUUGCAGAAGCUACAAAGAUAAAGGCUUCAUGCCGAAAUCAACACCCUGUCAUUUUUAU
GGCAGGGUG, ACUUCUCCACAGGAGUCAGAUGUUUUUGUACUCUCAAGAUUUAAGUA
ACUGUACAACGAAACUUACACAGUUACUAAAUCUUGCAGAAGCUACAAAGAUAAAGG
CUUCAUGCCGAAAUCAACACCCUGUCAUUUUUAUGGCAGGGUG; およびGACUUCUCCA
CAGGAGUCAGAUGUUUUUGUACUCUCAAGAUUUAAGUAACUGUACAACGAAACUUAC
ACAGUUACUAAAUCUUGCAGAAGCUACAAAGAUAAAGGCUUCAUGCCGAAAUCAACAC
CCUGUCAUUUUUAUGGCAGGGUGからなる群から選択される配列を含む、

30

塩基エディターシステム。

165. 必要とする対象において疾患を治療する方法であって、

ガイドポリヌクレオチドまたはそれをコードする核酸と；

ポリヌクレオチドプログラム可能なDNA結合ドメインまたはそれをコードする核酸と；

デアミナーゼドメインまたはそれをコードする核酸と

40

を含む塩基エディターシステムを前記対象に投与することを含み、

ここで、ガイドポリヌクレオチドは、対象中の細胞の標的ポリヌクレオチド中の核酸塩
基の脱アミノ化をもたらすために塩基エディターシステムをターゲティングすることがで
きるものであり、標的ポリヌクレオチドは、疾患の原因となる病原性アミノ酸を含むタン
パク質をコードし、脱アミノ化が、病原性アミノ酸から良性アミノ酸への置換をもたらし
、それにより疾患を治療し、前記良性アミノ酸は、前記タンパク質の野生型アミノ酸とは
異なる、

方法。

166. 必要とする対象において疾患を治療する方法であって、

50

(a) 細胞に、
ガイドポリヌクレオチドまたはそれをコードする核酸と；
ポリヌクレオチドプログラム可能なDNA結合ドメインまたはそれをコードする核酸と；
デアミナーゼドメインまたはそれをコードする核酸と
を含む塩基エディターシステムを導入すること、および

(b) 前記細胞を前記対象に投与すること
を含み、

ここで、ガイドポリヌクレオチドは、対象中の細胞の標的ポリヌクレオチド中の核酸塩基の脱アミノ化をもたらすために塩基エディターシステムをターゲティングすることができるものであり、

標的ポリヌクレオチドが、疾患の原因となる病原性アミノ酸を含むタンパク質をコードし、脱アミノ化が、病原性アミノ酸から良性アミノ酸への置換をもたらし、それにより疾患を治療し、前記良性アミノ酸は、前記タンパク質の野生型アミノ酸とは異なる、

方法。

167. 疾患の治療のための改変された細胞を産生する方法であって、

ガイドポリヌクレオチドまたはそれをコードする核酸と；
ポリヌクレオチドプログラム可能なDNA結合ドメインまたはそれをコードする核酸と；
デアミナーゼドメインまたはそれをコードする核酸と
を含む塩基エディターシステムを細胞に導入することを含み、

ここで、ガイドポリヌクレオチドは、細胞中の標的ポリヌクレオチド中の核酸塩基の脱アミノ化をもたらすために、塩基エディターシステムをターゲティングすることができるものであり、

前記標的ポリヌクレオチドが、疾患の原因となる病原性アミノ酸を含むタンパク質をコードし、前記脱アミノ化が、前記病原性アミノ酸から良性アミノ酸への置換をもたらし、前記良性アミノ酸は、前記タンパク質の野生型アミノ酸とは異なる、

方法。

168. 導入がインビボまたはエクスピボである、実施形態167の方法。

169. 細胞が哺乳動物細胞である、実施形態167または168の方法。

170. 標的ポリヌクレオチドが病原性アミノ酸の原因となる病原性一塩基多型 (SNP) を含む、実施形態164~169のいずれか一項記載の方法。

171. 疾患が、MCAD欠損症、鎌状赤血球症、ヘモグロビン病、 β -サラセミア、ペンドレッド症候群、家族性パーキンソン病、または α -1アンチトリプシン欠損症 (A1AD) である、実施形態170の方法。

172. 核酸塩基の脱アミノ化がタンパク質の発現、活性、または安定性を改変する、実施形態164~171のいずれか1つの方法。

173. 核酸塩基の脱アミノ化がタンパク質の発現、活性、または安定性を増加させる、実施形態172の方法。

174. 標的ポリヌクレオチドがACADMポリヌクレオチドであり、タンパク質がMCADタンパク質である、実施形態164~173のいずれか1つの方法。

175. 脱アミノ化が、MCADタンパク質におけるアミノ酸置換K329E> K329Gのアミノ酸変化をもたらす、実施形態174の方法。

176. 標的ポリヌクレオチドがHBBポリヌクレオチドであり、タンパク質がヘモグロビンの β サブユニット (HbB) である、実施形態164~173のいずれか1つの方法。

177. 脱アミノ化がHbBにおけるアミノ酸置換E26K> E26Gをもたらす、実施形態176の方法。

178. 標的ポリヌクレオチドがPDSポリヌクレオチドであり、タンパク質がペンドリンである、実施形態164~173のいずれか1つの方法。

179. 脱アミノ化が、ペンドリンにおけるアミノ酸置換T416P> T416Fをもたらす、実施形態178の方法。

180. 標的ポリヌクレオチドがSCNAポリヌクレオチドであり、タンパク質がアルファ

10

20

30

40

50

シヌクレインタンパク質である、実施形態164~173のいずれか1つの方法。

181. 脱アミノ化が、アルファシヌクレインタンパク質におけるアミノ酸置換A30P> A30Lをもたらす、実施形態180の方法。

182. 遺伝子がSERPINA1ポリヌクレオチドであり、タンパク質がA1ATタンパク質である、実施形態164~173のいずれか1つの方法。

183. 脱アミノ化が、A1ATタンパク質におけるアミノ酸置換E342K> E342Gをもたらす、実施形態182の方法。

184. ポリヌクレオチドプログラム可能なDNA結合ドメインがCas9ドメインである、実施形態164~183のいずれか1つの方法。

185. Cas9ドメインがヌクレアーゼ不活性Cas9ドメインまたはCas9ニッカーゼドメインである、実施形態184の方法。

186. Cas9ドメインがSpCas9ドメインを含む、実施形態184または185の方法。

187. SpCas9ドメインがD10Aおよび/またはH840Aアミノ酸置換またはその対応するアミノ酸置換を含む、実施形態186の方法。

188. SpCas9ドメインがNGN PAMに対する特異性を有する、実施形態186または187の方法。

189. SpCas9ドメインがアミノ酸置換D 1135 M、S 1136 Q、G 1218 K、E 1219 F、A 1322 R、D 1332 A、R 1335 E、およびT 1337 R、またはそれらの対応するアミノ酸置換を含む、実施形態186~188のいずれか1つの方法。

190. Cas9ドメインがSaCas9ドメインを含む、実施形態184または185の方法。

191. SaCas9ドメインは、NNNRRT PAMに対する特異性を有する、実施形態190の方法。

192. SaCas9ドメインがアミノ酸置換N 579 Aまたはその対応するアミノ酸置換を含む、実施形態190または191の方法。

193. SaCas9ドメインがアミノ酸置換E 782 K、N 968 K、およびR 1015 H、またはそれらの対応するアミノ酸置換を含む、実施形態190~192のいずれか1つの方法。

194. Cas9ドメインがSt1Cas9ドメインを含む、実施形態184または185の方法。

195. Cas9ドメインがNNACCA PAMに対する特異性を有する、実施形態194の方法。

196. デアミナーゼドメインがシチジンデアミナーゼドメインを含む、実施形態164~195のいずれか1つの方法。

197. シチジンデアミナーゼドメインがAPOBEC1ドメインを含む、実施形態196の方法。

198. デアミナーゼドメインがアデノシンデアミナーゼドメインを含む、実施形態164~195のいずれか1つの方法。

199. アデノシンデアミナーゼドメインは、TADA 7.10のアミノ酸配列を含むTadAドメインを含む、実施形態198の方法。

200. 塩基編集システムがUGIドメインをさらに含む、実施形態164~199のいずれか1つの方法。

201. 塩基編集システムが1つ以上のリンカーをさらに含む、実施形態164~200のいずれか1つの方法。

202. ポリヌクレオチドプログラム可能なDNA結合ドメイン、デアミナーゼドメイン、およびUGIドメインのうちの二つ以上がリンカーを介して連結されている、実施形態201の方法。

203. リンカーがペプチドリンカーであり、それによって塩基編集融合タンパク質を形成する、実施形態202の方法。

204. 塩基編集融合タンパク質がBE4のアミノ酸配列を含む、実施形態203の方法。

205. 塩基編集融合タンパク質は、TadA 7.10のアミノ酸配列を含む、実施形態203の方法。

206. 疾患の治療のための塩基エディターシステムであって、

ガイドポリヌクレオチドまたはそれをコードする核酸と；

10

20

30

40

50

ポリヌクレオチドプログラム可能なDNA結合ドメインまたはそれをコードする核酸と；
デアミナーゼドメインまたはそれをコードする核酸と
を含み、

前記ガイドポリヌクレオチドは、標的ポリヌクレオチド中の核酸塩基の脱アミノ化をもたらすために前記塩基エディターシステムをターゲティングすることができるものであり、
標的ポリヌクレオチドは表2に記載された標的配列を含み、

前記標的ポリヌクレオチドは、疾患の原因となる病原性アミノ酸を含むタンパク質をコードし、前記脱アミノ化が、前記病原性アミノ酸から良性アミノ酸への置換をもたらし、
ここで前記良性アミノ酸は、前記タンパク質の野生型アミノ酸とは異なる。

207. 必要とする対象において疾患を治療する方法であって、

ガイドポリヌクレオチドまたはそれをコードする核酸と；

ポリヌクレオチドプログラム可能なDNA結合ドメインまたはそれをコードする核酸と；
デアミナーゼドメインまたはそれをコードする核酸と

を含む塩基エディターシステムを、前記対象に投与することを含み、

ここで、ガイドポリヌクレオチドは、対象における細胞の標的ポリヌクレオチド中の核酸塩基の脱アミノ化をもたらすために前記塩基エディターシステムをターゲティングすることができ、それによって疾患を治療し、

ここで、前記核酸塩基は、遺伝子の調節エレメント中にある、
方法。

209. 必要とする対象における疾患を治療する方法であって、

(a) 細胞に、

ガイドポリヌクレオチドまたはそれをコードする核酸と；

ポリヌクレオチドプログラム可能なDNA結合ドメインまたはそれをコードする核酸と；
デアミナーゼドメインまたはそれをコードする核酸と

を含む塩基エディターシステムを導入すること、および

(b) 前記細胞を前記対象に投与すること

を含み、

ここで、ガイドポリヌクレオチドは、前記細胞の標的ポリヌクレオチド中の核酸塩基の脱アミノ化をもたらすために前記塩基エディターシステムをターゲティングすることができるものであり、それによって疾患を治療し、ここで、前記核酸塩基は遺伝子の調節エレメント中にある、

方法。

209. 疾患の治療のための改変された細胞を産生する方法であって、

ガイドポリヌクレオチドまたはそれをコードする核酸と；

ポリヌクレオチドプログラム可能なDNA結合ドメインまたはそれをコードする核酸と；
デアミナーゼドメインまたはそれをコードする核酸と

を含む塩基エディターシステムを細胞に導入することを含み、

ここで、ガイドポリヌクレオチドは、前記細胞の標的ポリヌクレオチド中の核酸塩基の脱アミノ化を行うために前記塩基エディターシステムをターゲティングすることができるものであり、ここで、前記核酸塩基は遺伝子の調節エレメント中にある、

方法。

210. 導入がインビボまたはエクスピボである、実施形態209の方法。

211. 細胞が哺乳動物細胞である方法の実施形態209または210の方法。

212. 細胞がCD34+細胞である、実施形態211の方法。

213. 前記遺伝子が前記疾患に関連するものであり、実施形態207～212のいずれか1つの方法。

214. 前記遺伝子の発現、活性、またはプロセッシングが疾患の原因である、実施形態213の方法。

215. 脱アミノ化が前記遺伝子の発現、活性、またはプロセッシングを改変する、実施形態207～214のいずれか1つの方法。

10

20

30

40

50

216. 脱アミノ化が、調節エレメントに対する少なくとも1つのタンパク質の結合パターンを改変する、実施形態207~215のいずれか1つの方法。

217. 前記調節エレメントが、遺伝子のプロモーター、エンハンサー、リプレッサー、サイレンサー、インスレーター、開始コドン、終止コドン、Kozakコンセンサス配列、スプライスアクセプター、スプライスドナー、スプライスサイト、3'非翻訳領域(UTR)、5'非翻訳領域(UTR)、または遺伝子間領域である、実施形態207~216のいずれか一項記載の方法。

218. 脱アミノ化が、スプライス部位、開始コドン、終止コドン、またはKozakコンセンサス配列の除去をもたらす、実施形態207~218のいずれか1つの方法。

219. 脱アミノ化が、スプライス部位、開始コドン、終止コドン、またはKozakコンセンサス配列の付加をもたらす、実施形態207~218のいずれか1つの方法。

220. 脱アミノ化がイントロン含有物を生じる、実施形態207~218のいずれか1つの方法。

221. 脱アミノ化がエキソンスキッピングをもたらす、実施形態207~218のいずれか1つの方法。

222. 脱アミノ化が、表2に記載された遺伝子のいずれか1つから選択される遺伝子の調節エレメント内である、実施形態207~221のいずれか1つの方法。

223. 前記疾患が胎児ヘモグロビン遺伝性持続症(HPFH)である、実施形態222の方法。

224. 前記遺伝子はHBG1またはHBG2である、実施形態223の方法。

225. 標的ポリヌクレオチドが、TCCACAGGAGTCAGATGCAC, TGAAGAGGTGTCCTCAGTCTA, TCTGAAGAGGTGTCCTCAGTCT,および TGGTAAGGCCCTGGGCAGGTからなる群より選択される配列、またはその任意の相補鎖配列を含む、実施形態224の方法。

226. 前記ポリヌクレオチドプログラム可能なDNA結合ドメインがCas9ドメインである、実施形態207~225のいずれか1つの方法。

227. 前記Cas9ドメインがヌクレアーゼ不活性Cas9ドメインまたはCas9ニッカーゼドメインである、実施形態226の方法。

228. 前記Cas9ドメインがSpCas9ドメインを含む、実施形態226または227の方法。

229. 前記SpCas9ドメインがD10Aおよび/またはH840Aアミノ酸置換またはその対応するアミノ酸置換を含む、実施形態228の方法。

230. 前記SpCas9ドメインが、NGN PAMに対する特異性を有する、実施形態228または229の方法。

231. 前記SpCas9ドメインが、アミノ酸置換D 1135 M、S 1136 Q、G 1218 K、E 1219 F、A 1322 R、D 1332 A、R 1335 E、およびT 1337 R、またはそれらの対応するアミノ酸置換を含む、実施形態228~230のいずれか1つの方法。

232. 前記Cas9ドメインがSaCas9ドメインを含む、実施形態226または227の方法。

233. 前記SaCas9ドメインは、NNRRT PAMに対する特異性を有する、実施形態232の方法。

234. 前記SaCas9ドメインが、アミノ酸置換N 579 Aまたはその対応するアミノ酸置換を含む、実施形態232または233の方法。

235. 前記SaCas9ドメインが、アミノ酸置換E 782 K、N 968 K、およびR 1015 H、またはそれらの対応するアミノ酸置換を含む、実施形態232~234のいずれか1つの方法。

236. 前記Cas9ドメインがSt1Cas9ドメインを含む、実施形態226または227の方法。

237. 前記Cas9ドメインがNNACCA PAMに対する特異性を有する、実施形態236の方法。

238. 前記デアミナーゼドメインがシチジンデアミナーゼドメインを含む、実施形態207~237のいずれか1つの方法。

239. 前記シチジンデアミナーゼドメインがAPOBEC1ドメインを含む、実施形態238の方法。

10

20

30

40

50

240. 前記デアミナーゼドメインがアデノシンデアミナーゼドメインを含む、実施形態207～237のいずれか1つの方法。

241. 前記アデノシンデアミナーゼドメインが、*TadA 7.10*のアミノ酸配列を含む*TadA*ドメインを含む、実施形態240の方法。

242. 塩基エディターシステムがUGIドメインをさらに含む、実施形態207～41のいずれか1つの方法。

243. 塩基エディターシステムが1つ以上のリンカーをさらに含む、実施形態207～242のいずれか1つの方法。

244. ポリヌクレオチドプログラム可能なDNA結合ドメイン、デアミナーゼドメイン、およびUGIドメインのうちの二つ以上がリンカーを介して連結されている、実施形態244の方法。

10

245. 前記リンカーがペプチドリンカーであり、それによって塩基編集融合タンパク質を形成する、実施形態244の方法。

246. 前記塩基編集融合タンパク質がBE4のアミノ酸配列を含む、実施形態245の方法。

247. 前記塩基編集融合タンパク質がBE4のアミノ酸配列を含む、実施形態245の方法。

248. ガイドポリヌクレオチドまたはそれをコードする核酸と；

ポリヌクレオチドプログラム可能なDNA結合ドメインまたはそれをコードする核酸と；
デアミナーゼドメインまたはそれをコードする核酸と
を含む、塩基エディターシステムであって、

前記ガイドポリヌクレオチドは、標的ポリヌクレオチド中の核酸塩基の脱アミノ化をもたらすために塩基エディターをターゲティングすることができるものであり、

20

前記標的ポリヌクレオチドが、表3Aまたは表3Bに記載された標的配列を含み、

前記核酸塩基は、表3Aまたは表3Bに記載された遺伝子の調節エレメント内にある、
塩基エディターシステム。

【実施例】

【0541】

以下の実施例は、例示のみを目的として提供され、本願で提供される特許請求の範囲を限定することは意図しない。

【0542】

【実施例1. 塩基エディターにおけるPAMバリエーションの検証】

30

新しいCRISPRシステムとPAMバリエーションは、塩基エディターが標的SNPにおいて正確な補正を行うことを可能にする。いくつかの新規PAMバリエーションを評価し、検証した。PAM評価および塩基エディターの詳細は、例えば、国際PCT出願番号PCT/2017/045381 (WO2018/027078) およびPCT/US2016/058344 (WO2017/070632) に記載されており、これらの各々は、その全体が参照により本明細書に組み込まれる。Komor, A.C., et al., “Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage” *Nature* 533, 420-424 (2016); Gaudelli, N.M., et al., “Programmable base editing of A·T to G·C in genomic DNA without DNA cleavage” *Nature* 551, 464-471 (2017); および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” *Science Advances* 3 :eaao4774 (2017)も参照（その全体が参照により本明細書に組み込まれる）。

40

【0543】

【実施例2. アルファ1アンチトリプシン欠損症 (A1AD) を補正する遺伝子編集

$\alpha 1$ アンチトリプシン (A1A) は、14番染色体上のSERPINA1遺伝子によってコードされるプロテアーゼインヒビターである。この糖タンパク質は主に肝臓で合成され、血中に分泌され、血清中濃度は健康な成人で1.5～3.0 g/L (20～52 μ mol/L)である(図1)。肺間質および肺胞上皮液中に拡散し、そこで好中球エラスターゼを不活性化することにより、プロテアーゼを介した損傷から肺組織を保護する。 $\alpha 1$ アンチトリプシン欠損症は常染色体共優性遺伝する。100を超えるSERPINA1遺伝子の遺伝的バリエーションが記述されてい

50

るが、すべてが疾患と関連しているわけではない。これらのバリエーションのアルファベット表記は、ゲル電気泳動での移動速度に基づいている。最もよくみられるバリエーションはM (中程度の移動度)アリルであり、最も頻度の高い2つの欠損症アリルはPiSおよびPiZ (後者は移動速度が最も遅い)である。測定可能な血清蛋白を産生しないいくつかの突然変異が報告されている;これらは「ヌル」アリルと呼ばれる。最も一般的な遺伝子型はMMであり、それは $\alpha 1$ アンチトリプシンの正常な血清値を生じる。重度の欠損症の人のほとんどは、Zアリルのホモ接合体(ZZ)である。Zタンパク質は、肝細胞の小胞体で産生される際にミスフォールドし重合する;これらの異常なポリマーは肝臓において捕捉され、 $\alpha 1$ アンチトリプシンの血清中レベルを大幅に低下させる。 $\alpha - 1$ アンチトリプシン欠損症患者に見られる肝疾患は、肝細胞における異常な $\alpha - 1$ アンチトリプシンタンパク質の蓄積とその結果としての細胞応答 (オートファジー、小胞体ストレス応答およびアポトーシスを含む)によって引き起こされる。図2は、最も一般的な遺伝子型および対応する $\alpha - 1$ アンチトリプシンの血清レベルを示す。 $\alpha - 1$ アンチトリプシンの循環レベルが低下すると、肺の好中球エラスターゼ活性が上昇する;プロテアーゼとアンチプロテアーゼのこの不均衡が、この状態に関連する肺疾患をもたらす(図1)。

【0544】

アルファ-1アンチトリプシン欠損症 (「A1AD」) は白人に最も多くみられ、この障害は罹患者の肺および肝臓に最も高頻度に影響を及ぼす。肺で最もよくみられる症状は、早期発症型(30歳代と40歳代の患者)の汎細葉性肺気腫であり、肺底部で最も著明である。しかしながら、びまん性または上葉性の気腫が生じることがあり、気管支拡張症も起こり得る。最も頻繁に報告される症状は、呼吸困難、喘鳴および咳である。肺機能検査はCOPDと一致する所見を示す;しかし、気管支拡張薬の反応性がみられ、喘息と診断され得る。

【0545】

ZZ遺伝子型に起因する肝疾患は様々な形で発現する。罹患した乳児は、新生児期に胆汁うっ滞性黄疸を呈することがあり、ときに無胆汁性便(青白いか粘土色)および肝腫大を伴う。血液中の抱合型ビリルビン、トランスアミナーゼおよび γ -グルタミルトランスフェラーゼの濃度が上昇する。より年長の小児および成人における肝疾患は、トランスアミナーゼ上昇の偶発的所見、または静脈瘤出血や腹水などの確立された肝硬変の徴候を呈することがある。 $\alpha 1$ アンチトリプシン欠損症は肝細胞癌の素因にもなる。ホモ接合ZZ遺伝子型は肝疾患の発達に必要なであるが、ヘテロ接合Z変異は、C型肝炎感染や嚢胞性線維症肝疾患などのより重度の肝疾患のリスクを高めることにより、他の疾患に対する遺伝的修飾因子として作用しうる。

【0546】

A1ADの最も一般的な臨床的バリエーションはE264V (PiS) とE342K (PiZ) アリルである。A1AD患者の半数以上は、少なくとも1コピーのE342Kアリル変異を持っていた。HDRを介するヌクレアーゼゲノム編集は非効率的であり、豊富なインデルが循環レベルを低下させ肺症状を悪化させ得る。AAVベクターを用いた肝細胞の形質導入が関わる遺伝子治療は、更なるミスフォールドタンパク質の産生により肝病理を悪化させる。野生型A1ATとE342K A1ATをノックダウンするsiRNAとの両方をコードするAAVは、両方の病理に対処する有望性を示す。

【0547】

本明細書に記載するような塩基編集を用いて、A1ATをコードするSERPINA1遺伝子における突然変異を補正するための戦略を、図3に示す。NGC PAMを受け容れるように進化され操作されたSpCas9 DNA結合ドメインを有するABEのmRNA+gRNAトランスフェクションを伴う本明細書中に記載される塩基編集システム(ngcABE)を用いて、SERPINA1核酸配列中の核酸塩基A5/A7を編集し、D341Gを導入した。SERPINA1核酸配列中の核酸塩基A7/A8を編集し、E342Gを導入した。ABEまたはREPAIR (RNA編集)を用いたE342Kバリエーションについての異なる補正戦略は、D341G、E342G、E343E、およびK343G生成物の混合物を生じた。図4に示されるように、A1ATタンパク質機能についてD341Gオフターゲット編集はそれ自体は良性であったが、E342Kとの組合せでは有害であった。E

342GはA1AT機能を復元させた。E342Rは機能しなかった。

【0548】

下記表8は、野生型およびバリエント (E342K) SERPINA1にコードされるアミノ酸配列、ならびに野生型およびバリエント (E342K) SERPINA 1のポリヌクレオチドのオープンリーディングフレーム (ORF) 核酸配列の代表的なリストを示す。

【表8】

	配列 番号	配列	
<i>SERPINA1</i> アミノ酸	1	MPSSVSWGILLLAGLCCCLVPVSLAEDPQGDAQAQKTDTSHHQDQDHPTFN KITPNLAEFASFSLYRQLAHQSNSTNIFFSPVSIATAFAMLSLGTKADTHD EILEGLNFNLTEIPEAQIHEGFQELLRTLNPDSQLQLTTGNGLFLSEGLK LVDKFLVDVKKLYHSEAFVNFVGDTEEAQKQINDYVEKGTQGKIVDLV KELDRDRTVFALVNYIFFKGKWERPFVVDTEEDFHVDQVTTVKVPM MKRLGMFNIQHCKKLSSWVLLMKYLGNTAIFFLPDEGKLQHLLENELT HDIITKFLNEDRRSASLHLPKLSITGTLDKSVLGQLGITKVFNSGADL SGVTEEAPLKLSKAVHKAVLTIDEKGTEAAGAMFLEAIPMSIPPEVKFN KPFVFLMIEQNTKSPLFMGKVVNPQK	10
<i>SERPINA1</i> ORF	2	ATGCCGTCTTCTGTCTCGTGGGGCATCCTCCTGCTGGCAGGCCTGTG CTGCCTGGTCCCTGTCTCCCTGGCTGAGGATCCCCAGGGAGATGCTG CCCAGAAGACAGATACATCCCACCATGATCAGGATCACCAACCTT CAACAAGATCACCCCCAACCTGGCTGAGTTCGCCTTCAGCCTATACC GCCAGCTGGCACACCAGTCCAACAGCACCAATATCTTCTTCTCCCA GTGAGCATCGCTACAGCCTTTGCAATGCTCTCCCTGGGGACCAAGGC TGACACTCACGATGAAATCCTGGAGGGCCTGAATTTCAACCTCACG GAGATTCCGGAGGCTCAGATCCATGAAGGCTTCCAGGAACCTCCTCC GTACCTCAACCAGCCAGACAGCCAGCTCCAGCTGACCACCGGCAA TGGCCTGTTCCTCAGCGAGGGCCTGAAGCTAGTGGATAAGTTTTTG AGGATGTTAAAAAGTTGTACCACTCAGAAGCCTTCACTGTCAACTTC GGGGACACCGAAGAGGGCCAAGAAACAGATCAACGATTACGTGGAG AAGGGTACTCAAGGGAAAATTGTGGATTGGTCAAGGAGCTTGACA GAGACACAGTTTTTGTCTGGTGAATTACATCTTCTTTAAAGGCAAA TGGGAGAGACCCTTTGAAGTCAAGGACACCGAGGAAGAGGACTTCC ACGTGGACCAGGTGACCACCGTGAAGGTGCCTATGATGAAGCGTTT AGGCATGTTTAACATCCAGCACTGTAAGAAGCTGTCCAGCTGGGTG CTGCTGATGAAATACCTGGGCAATGCCACCGCCATCTTCTTCTGCTG TGATGAGGGGAACTACAGCACCTGGAAAATGAACTCACCCACGAT ATCATCACCAGTTCTTGGAAAATGAAGACAGAAGGTCTGCCAGCT TACATTTACCCAACTGTCCATTACTGGAACCTATGATCTGAAGAGC GTCCTGGGTCAACTGGGCATCACTAAGGTCTTCAGCAATGGGGCTG ACCTCTCCGGGGTCACAGAGGAGGCACCCCTGAAGCTCTCCAAGCG CGTGCATAAGGCTGTGCTGACCATCGACGAGAAAGGGACTGAAGCT GCTGGGGCCATGTTTTTAGAGGCCATACCCATGTCTATCCCCCCCCGA GGTCAAGTTCAACAAACCCTTTGTCTTCTTAATGATTGAACAAAATA CCAAGTCTCCCCTCTTCATGGGAAAAGTGGTGAATCCCACCCAAAA A	20 30
<i>SERPINA1</i> E342K アミノ酸	3	MPSSVSWGILLLAGLCCCLVPVSLAEDPQGDAQAQKTDTSHHQDQDHPTFN KITPNLAEFASFSLYRQLAHQSNSTNIFFSPVSIATAFAMLSLGTKADTHD EILEGLNFNLTEIPEAQIHEGFQELLRTLNPDSQLQLTTGNGLFLSEGLK LVDKFLVDVKKLYHSEAFVNFVGDTEEAQKQINDYVEKGTQGKIVDLV KELDRDRTVFALVNYIFFKGKWERPFVVDTEEDFHVDQVTTVKVPM MKRLGMFNIQHCKKLSSWVLLMKYLGNTAIFFLPDEGKLQHLLENELT HDIITKFLNEDRRSASLHLPKLSITGTLDKSVLGQLGITKVFNSGADL SGVTEEAPLKLSKAVHKAVLTIDKKGTEAAGAMFLEAIPMSIPPEVKFN	40

		KPFVFLMIEQNTKSPLFMGKVVNPTQK
<i>SERPI</i> <i>NAI</i> E342K ORF	4	ATGCCGTCTTCTGTCTCGTGGGGCATCCTCCTGCTGGCAGGCCTGTG CTGCCTGGTCCCTGTCTCCCTGGCTGAGGATCCCCAGGGAGATGCTG CCCAGAAGACAGATACATCCCACCATGATCAGGATCACCCAACCTT CAACAAGATCACCCCCAACCTGGCTGAGTTTCGCCTTCAGCCTATACC GCCAGCTGGCACACCAGTCCAACAGCACCAATATCTTCTTCTCCCA GTGAGCATCGCTACAGCCTTTGCAATGCTCTCCCTGGGGACCAAGGC TGACACTCACGATGAAATCCTGGAGGGCCTGAATTTCAACCTCACG GAGATTCCGGAGGCTCAGATCCATGAAGGCTTCCAGGAACCTCCTCC GTACCCTCAACCAGCCAGACAGCCAGCTCCAGCTGACCACCGGCAA TGGCCTGTTCTCAGCGAGGGCCTGAAGCTAGTGGATAAGTTTTTGG AGGATGTTAAAAAGTTGTACCACTCAGAAGCCTTCACTGTCAACTTC GGGGACACCGAAGAGGCCAAGAAACAGATCAACGATTACGTGGAG AAGGGTACTCAAGGGAAAATTGTGGATTTGGTCAAGGAGCTTGACA GAGACACAGTTTTTGTCTCTGGTGAATTACATCTTCTTTAAAGGCAAA TGGGAGAGACCCCTTTGAAGTCAAGGACACCGAGGAAGAGGACTTCC ACGTGGACCAGGTGACCACCGTGAAGGTGCCTATGATGAAGCGTTT AGGCATGTTTAACATCCAGCACTGTAAGAAGCTGTCCAGCTGGGTG CTGCTGATGAAATACCTGGGCAATGCCACCGCCATCTTCTTCCTGCC TGATGAGGGGAAACTACAGCACCTGGAAAATGAACTCACCCACGAT ATCATCACCAAGTTCCTGGAAAATGAAGACAGAAGGTCTGCCAGCT TACATTTACCCAAACTGTCCATTACTGGAACCTATGATCTGAAGAGC GTCCTGGGTCAACTGGGCATCACTAAGGTCTTCAGCAATGGGGCTG ACCTCTCCGGGGTCACAGAGGAGGCACCCCTGAAGCTCTCCAAGGC CGTGCATAAGGCTGTGCTGACCATCGACaAGAAAGGGACTGAAGCT GCTGGGGCCATGTTTTTAGAGGCCATAACCCATGTCTATCCCCCCCCGA GGTCAAGTTCAACAAACCCTTTGTCTTCTTAATGATTGAACAAAATA CCAAGTCTCCCCTCTTCATGGGAAAAGTGGTGAATCCCACCCAAAA A

10

20

【0549】

[実施例3. HBB遺伝子の鎌状赤血球症部位における遺伝子編集]

AからGへの核酸塩基エディター (BE) が構築され、これは、NNNRRT PAMを受け容れる改変 *Staphylococcus aureus* Cas9 (SaCas9 KKH) 核酸プログラム可能DNA結合ドメインおよび野生型 TadA と TadA7.10 のヘテロ二量体を含んだ (ABE SaCas9 KKH)。野生型 SaCas9 (Kleinstiver et al. Nat Biotechnol., 2015; 33(12): 1293-1298) に対するアミノ酸置換 E782K/N968K/R1015H を含む SaCas9 KKH は、部位枯渇アッセイに基づいて、NNNRRT PAM においてヌクレアーゼ活性を有するものとして同定された。TadA7.10 ドメインは DNA 中のアデニンに対してアデノシンデアミナーゼ活性を有し、*S. aureus* nCas9 (D10A) ドメインはニッカーゼ活性を有する。真核生物のゲノム編集に使用するために、ABE SaCas9 KKH は、その C 末端に核局在化シグナルを含む。ABE SaCas9 KKH のアミノ酸配列を以下に示す (wtTadA 下線—a.a. リンカー 斜体—TadA*7.10 下線—アミノ酸リンカー 斜体—SaCas9 KKH—a.a. リンカー 斜体—NLS 太字斜体)。

30

40

50

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGA
AGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSGGSS
GGSSGSETPGTSESATPESSGGSSGSSSEVEFSHEYWMRHALTLAKRARDEREVPVGAVL
VLNNRVIGEGWNRRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCA
GAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFF
RMPRQVFNAQKKAQSSTDSGGSSGGSSGSETPGTSESATPESSGGSSGSKRNYILGLAIGI
 TSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRRRRRHRIQRVKLLF
 DYNLLTDHSELGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGVHNVNEVEEDTGNE
 LSTKEQISRNSKALEEKYVAELQLERLKKDGEVRGSINRFKTSYVKEAKQLLKVQKA
 YHQLDQSFIDTYIDLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVK
 YAYNADLYNALNDLNNLVITRDENEKLEYEKFQIENVFKQKKKPTLKQIAKEILVNE
 EDIKGYRVTSTGKPEFTNLKVYHDIKDITARKEIENAELLDQIAKILTIYQSSEDIQEELT
 NLNSELTQEEIEQISNLKGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPPKVD
 LSQKKEIPTTLVDDFILSPVVKRSFIQSIKVINAIIKKYGLPNDIIIELAREKNSKDAQKMIN
 EMQKRNRQTNERIEEIIRTTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNNPFN
 YEVDHIIPRSVSFDNSFNKVLVKQEENSKKGNRTPFQYLSSSDSKISYETFKKHILNLAK
 GKGRISKTKKEYLLEERDINRFVQKDFINRNLVDTRYATRGLMNLRLSYFRVNNLDVK
 VKSINGGFTSFLRRKWFKKERNKGYKHAEDALIINANADFIFKEWKLDKAKKVMEN
 QMFEEKQAESMPEIETEQEYKEIFITPHQIKHIKDFKDYKYSHRVDKKPNRKLINDTLYST
 RKDDKGNTLIVNNLNGLYDKDNDKLKLINKSPEKLLMYHHDHPQTYQKLKLIMEQYG
 DEKNPLYKYEEETGNYLTKYSKKDNGPVIKKIKYYGNKLNAHLDITDDYPNSRNKVVK
 LSLKPYRFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSKCYEEAKKLKKISNQAEFIAS
 FYKNDLIKINGELYRVIGVNNDLLNRIEVNMIDITYREYLENMNDKRPPHIIKTIASKTQSI
 KKYSTDILGNLYEVKSKKHPQIIKKGEGADKRTADGSEFESPKKKKRV.

10

20

30

【0550】

鎌状赤血球症部位での塩基編集を検討するために、鎌状赤血球症SNPを有するHBBポリ
 ヌクレオチドのレンチウイルス挿入型コピーを含むHEK293T細胞を、ABE SaCas9 KKH
 、ABE SpCas9 MQKSER、またはABE SpCas9 VRQRをコードするプラスミド、および
 鎌状細胞部位をそれぞれのPAMに隣接させて標的化するスペーサーを含む対応するガイド
 RNAで、共トランスフェクトした。特に、ABE7.10 SaCas9 KKHは、NNNRRT PAM (
 CATGGT、図7A)に隣接する鎌状細胞部位を標的化するスペーサーを含むガイドRNAとの
 組合せで、鎌状赤血球症SNP (A4) において活性を示した：UCCACAGGAGUCAGAUGCA
 C (20 nt)。塩基エディターABE SaCas9 KKHは、高スループットシーケンシング (H
 TS) によって決定されるように、鎌状赤血球症標的部位SNPにおいてA・TをG・Cに変換
 した。編集のパーセントは、ヌクレオチド位置A4、A6、A9およびA13におけるA・Tから
 G・Cへの塩基編集によって測定した。

40

【0551】

50

鎌状赤血球症部位での塩基編集をさらに調べるために、ジンクフィンガータンパク質融合体をABE SaCas9 KKH (SAKKH ABE7.10)に融合した(図8A)。ABE SaCas9 KKH/saKKH ABE7.10へのジンクフィンガーの追加は、塩基編集を増強させる可能性があるという仮説を立てた。HBB遺伝子座に結合するジンクフィンガーヌクレアーゼが生成され、ヌクレアーゼに基づく遺伝子補正戦略の一部として使用されている (Sebastiano et al. *Stem Cells*. 2011; 29(11): 1717-1726)。二つの構築物を作成し、saKKH ABE7.10 zf1ra (認識ヘリックス配列RNEHLEV、QSTTLKR、およびRTEHLAR)およびsaKKH ABE7.10 zf1rb (認識ヘリックス配列RGEHLRQ、QSGTL、RNDKLVP)と命名した(図8A)。

【0552】

鎌状赤血球症SNPを有するHBBポリヌクレオチドのレンチウイルス組み込みコピーを含むHEK293T細胞を、saKKH ABE7.10、saKKH ABE7.10 zf1ra、またはsaKKH ABE7.10 zf1rbをコードするプラスミド、およびCATGGT PAMに隣接する鎌状赤血球部位を標的化するスパーサーを含むガイドRNAで、共トランスフェクトした(図8B)。高スループットシーケンシング (HTS) によって決定されるように、すべての塩基エディターは、鎌状赤血球症SNP (図8Bの位置2に対応する)において活性を示した(図8C、図8D、図8Eおよび図8F)。特に、ABE7.10 zf1raおよびsaKKH ABE7.10 zf1rb塩基エディターは、20 ntガイドと組み合わせられて使用された場合、saKKH ABE7.10と比較して鎌状赤血球症SNPにおいて活性の増加を示した(図8C)。従って、鎌状赤血球症SNPの塩基編集は、鎌状赤血球症を治療するための組成物および方法において有用である。

【0553】

[実施例4. *Streptococcus thermophilus* 1 Cas9ポリヌクレオチドプログラム可能DNA結合ドメインを備えたAからGへの塩基エディター(ABE St1Cas9)は鎌状赤血球症に関連するSNPに対する塩基編集活性を有する]

Streptococcus thermophilus 1 Cas9核酸プログラム可能DNA結合ドメインおよび野生型TadAとTadA 7.10のヘテロ二量体を含む、AからGへの核酸塩基エディターを構築した(ABE St1Cas9)。TadA 7.10ドメインはDNA中のアデニンに対してアデノシンデアミナーゼ活性を有し、*S. thermophilus* nCas9 (D9A) ドメインはニッカーゼ活性を有する。真核生物ゲノム編集で使用するために、ABE St1Cas9はそのN末端とC末端に核局在化シグナルを含む。ABE St1Cas9のアミノ酸配列を以下に示す。(NLS 太字—wtTadA 下線—a.a.リンカー 斜体—TadA*7.10 下線—a.a.リンカー 斜体—nSt1Cas9—NLS-BP-NLS 太字斜体)：

10

20

30

40

50

MPKKKRKVSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPI
GRHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGAR
DAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSS
TDSSGSSGGSSGSETPGTSESATPESGGSSGGSSEVEFSHEYWMRHALTLAKRARDEREV
PVGAVLVNNRVIGEGWNRRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEP
CVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAA
LLCYFFRMPROVFNAQKKAQSSTDSSGSSGGSSGSETPGTSESATPESDLVLGLAIGISV
 GVGILNKVTGEIHKNSRIFPAAQAENNLVRRRTNRQGRRLARRKKHRRVRLNRLFEESEG
 LITDFTKISINLPYQLRVKGLTDELSNEELFIALKNMVKHRGISYLLDDASDDGNSSVGD
 YAQIVKENSQKLETCTPGQIQLERYQTYGQLRGDFTVEKDGGKHRLINVFPTSAYRSEA
 LRILQTQQEFNPQITDEFINRYLEILTGKRKYHGPNGNEKSRTDYGRYRTSGETLDNIFGI
 LIGKCTFYPDEFRAAKASYTAQEFNLLNDLNNLTVPPTETKKLSKEQKNQIINYVKNEKA
 MGPALFKYIAKLLSCDVADIKGYRIDKSGKAEIHTFEAYRKMKTLETLDIEQMDRETL
 DKLAYVLTNTEREGIQEALHEFADGSFSQKQVDELVQFRKANSSIFGKGWHNFSVKL
 MMELIPELYETSEEQMTILTRLGKQKTTSSSNKTKYIDEKLLTEEIYNPVVAKSVRQAIKI
 VNAAIKEYGDFDNIVIAMARETNEDDEKKAIQKIQKANKDEKDAAMLKAANQYNGKA
 ELPHSVFHGHKQLATKIRLWHQQGERCLYTGKTISIHDLINNSNQFEVDHILPLSITFDDS
 LANKVLVYATANQEKGQRTPYQALDSMDDAWSFRELKAFVRESKTLNKKKEYLLTE
 EDISKFDVRKKFIERNLVDTLYASRVVLNALQEHFRAHKIDTKVSVVRGQFTSQLRRHW
 GIEKTRDITYHHHAVDALIAASSQLNLWKKQKNTLVSYSEDQLLDIETGELISDDEYKES
 VFKAPYQHFVDTLKSKEFEDSILFSYQVDSKFNRKISDATIYATRQAKVGKDKADETYV
 LGKIKDIYTQDGYDAFMKIYKKDKSKFLMYRHDPQTFEKVIEPILENYPNKQINDKGKE
 VPCNPFLKYKEEHGYIRKYSKKGNGPEIKSLKYYDSKLGNHIDITPKDSNNKVVLQSVSP
 WRADVYFNKTTGKYEILGLKYADLQFDKGTGTYKISQEKYNDIKKKEGVDSSEFKFT
 LYKNDLLLVKDTETKEQQLFRFLSRTMPKQKHVELKPYDKQKFEGGEALIKVLGNVA
 NSGQCKKGLGKSNISIYKVRTDVLGNQHIIKNEGDKPKLDF**PKKKRKVEGADKRTADGS**
EFESPKKKRKV

10

20

30

【0554】

40

5'-NNAGAA-3'および5'-NNACCA-3'のPAM配列に隣接する部位におけるABE St1Cas9の塩基編集活性を、HEK293T細胞において試験した。これらのPAMは、部位枯渇アッセイに基づいてSt1Cas9ヌクレアーゼの代替PAMとして同定された(Kleistiver et. al. Nature (2015) 523: 481-485)、ABE St1Cas9の推定PAMを表す。ABE St1Cas9がこれらのPAMを受容することを確認するために、HEK293T細胞を、ABE St1Cas9をコードするプラスミド、およびNNAGAAまたはNNACCAのPAM配列の5'の20ヌクレオチド配列を標的化する適切なsgRNAで、同時トランスフェクトした。核酸塩基編集を五日間行った後、細胞からゲノムDNAを抽出し、ハイスループット配列決定(HTS)により遺伝子座を分析した。

【0555】

50

ABE St1Cas9は、NNAGAA PAM配列に隣接する標的部位においてAをGに変換した(図9Aおよび9B、上)。これはたとえば、PAMの反対側の鎖でTがCに変換されることによって示されている(図9A参照；CTAGAAの逆総補鎖であるTTCTAGが示されている)。同様の条件下で維持されたがただし塩基エディターを存在させなかった未処理細胞は、そのような改変は示さなかった(図9Aおよび図9B、下)。St1Cas9による塩基編集は、例えば同じ配列を標的とするSt1Cas9ヌクレアーゼと比較して、インデル形成のレベルが低かった(図9Aおよび9B、挿入表)。

【0556】

ABE St1Cas9はNNACCA PAM配列に隣接する標的部位においてAをGに変換した(図9Cおよび9D、上)。これはPAMの反対側の鎖におけるTからCへの変換によって示された(NNACCAの逆総補鎖であるTGGTNNが示されている)。同様の条件下で維持されたがただし塩基エディターを存在させなかった未処理細胞は、そのような改変を示さなかった(図9Cおよび9D、下)。St1Cas9による塩基編集は、例えば同じ配列を標的とするSt1Cas9ヌクレアーゼと比較して、インデル形成のレベルが低かった(図9Cおよび9D、挿入表)。

【0557】

HBB遺伝子の鎌状赤血球部位にはNNACCAが存在し、疾患関連「A」(鎌状赤血球症SNP「T」の反対)を、ガイドRNAの21 ntプロトスペーサー：ACUUCUCCACAGGAGUCAGAUの9位に配置する。疾患関連「A」を編集するとTからCへのコード変異が起こり、βグロビンのVal→Ala置換が起こる。スペーサー領域の位置11および14におけるTからCへのコード突然変異もまた可能であることがわかる。しかし、これらはそれぞれのコドンのゆらぎ(wobble)位置で起こり、従ってこれらの位置のアミノ酸(T4およびP5)は変化しないので、サイレントである。鎌状赤血球症SNPを標的とするガイドRNAは、20 ntスペーサー：CUUCUCCACAGGAGUCAGAUまたは22 ntスペーサー：CUUCUCCACAGGAGUCAGAUを用いて作製することもできる。

【0558】

鎌状赤血球症部位での塩基編集を検討するために、鎌状赤血球症SNPを有するHBBポリヌクレオチドのレンチウイルス組み込みコピーを含むHEK293T細胞を、ABE St1Cas9をコードするプラスミドおよびNNACCA PAMに隣接する鎌状赤血球症部位を標的化するスペーサーを含むガイドRNAで、共トランスフェクトした。ABE St1Cas9は鎌状赤血球症標的部位SNPにおいてA・TをG・Cに変換した(図9E上)。このことはNNACCA PAMの反対鎖でTがCに変換されることによって示された(GCACCAの逆総補鎖であるTGGTGCが図9Eに示されている)。同様の条件下で維持されたがただし塩基エディターを存在させなかった未処理細胞は、そのような改変を示さなかった(図9E下)。St1Cas9による塩基編集は、例えば同じ配列を標的とするSt1Cas9ヌクレアーゼと比較して、インデル形成のレベルが低かった(図9E、挿入表)。従って、鎌状赤血球症SNPを塩基編集するこの能力は、鎌状赤血球症を治療するための組成物および方法において有用である。

【0559】

[実施例5. 改変Streptococcus pyogenes Cas9 (SpCas9) ポリヌクレオチドプログラム可能DNA結合ドメインを有するAからGへの塩基エディター(ngc-ABE)は、鎌状赤血球症に関連するSNPに対する塩基編集活性を有する]

野生型SpCas9アミノ酸配列に対して、D1135M、S1136Q、G12128K、E1219F、A1322R、R1335E、およびT1337Rのうちの1以上と組み合わせた置換D1332Aを含む、改変されたStreptococcus pyogenes Cas9が開発された。その結果、D1135M、S1136Q、G12128K、E1219F、A1322R、D1332A、R1335EおよびT1337R置換を有する改変SpCas9がNGC PAMを受容できることが見出された。NGC PAMを受容する改変S. pyogenes Cas9核酸プログラム可能DNA結合ドメインおよび野生型TadAとTadA 7.10のヘテロ二量体を含む、AからGへの核酸塩基エディターを構築した(ngc-Cas9)。TadA 7.10ドメインはDNA中のアデニンに対してアデノシンデアミナーゼ活性を有し、S. pyogenes nCas9 (D10A) ドメインはニッカーゼ活性を有する。真核生物のゲノム編集に用いるために、ngc-ABEはそのC末端に核局在化シグナルを含む。ngc-ABEのアミノ酸配列を以

10

20

30

40

50

下に示す (wtTadA 下線—a.a.リンカー 斜体—TadA*7.10 下線—a.a.リンカー 斜体—n gc-Cas9—a.a.リンカー 斜体—NLS 太字斜体 ; a.a.置換 小文字) :

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGA
AGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSGGSS
GGSSGSETPGTSESATPESSGGSSGGSSSEVEFSHEYWMRHALTLAKRARDEREVPVGAVL
VLNNRVIGEGWNRRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCA
GAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFF
RMPRQVFNAQKKAQSSTDSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGLAIG
 TNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRR
 YTRRKNRICYLQEIFSNEMAKVDDSSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEK
 YPTIYHLRKKLVSDTKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFQILVQT
 YNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIASLGLTPNF
 KSNFDLAEDAKLQLSKDITYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTE
 ITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQE
 EFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYF
 FLKDNREKIEKILTRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEVVVDKGASAQSF
 IERMTNFDKNLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIV
 DLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIKDKDFLDN
 EENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLIN
 GIRDKQSGKTILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANL
 AGSPAIAKKGILQTVKVVDDELVKVMGRHKKPENIVIMARENQTTQKGQKNSRERMKRIE
 EGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYVDHIVPQ
 SFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTK
 AERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS
 LVSDFRKDFQFYKVRINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDV
 RKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGR
 DFATVRKVLSPQVNVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGF~~mg~~
 PTVAYSVLVVAKEKKGSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLII
 KLPKYSLELENGRKRMLASA~~kl~~LQKGNELALPSKYVNFLYLASHYEKLKGSPEDENEQK
 QLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKHARDKPIREQAENIIHLFTLT
 NLGAP~~r~~AFKYFDTTI~~a~~RK~~e~~Y~~r~~STKEVLDATLIHQSTGLYETRIDLSQLGGDEGADKRTAD
 GSEFES~~PKKKRKV~~.

10

20

30

40

【0560】

鎌状赤血球症部位での塩基編集を調べるために、鎌状赤血球症SNPを有するHBBポリヌクレオチドのレンチウイルス組み込みコピーを含むHEK293T細胞に、ngc-ABEをコードするプラスミドと、以下から選択される、NGC PAM (GCAの逆相補鎖であるTGCを図10に示す)に隣接する鎌状赤血球症部位を標的化するスペーサーを含むガイドRNAとを共トランスフェクトした: UUCUCCACAGGAGUCAGA (18-nt); CUUCUCCACAGGAGUCAGA (19-nt); ACUUCUCCACAGGAGUCAGA (20-nt); GACUUCUCCACAGGAGUCAGA (21-nt)

50

t); および AGACUUCUCCACAGGAGUCAGA (22-nt)。高スループットシーケンシング (HTS) によって決定されるように、塩基エディター-ngc-ABEは、鎌状赤血球症標的部位 SNPにおいてA・TをG・Cに変換した。編集のパーセントは、開始コドンによってコードされるメチオニンを含む全長β-グロビンポリペプチドに対する、アミノ酸位置P6、E7およびS10に対応するコドンにおけるA・TからG・Cへの塩基編集によって測定した。これらの位置での塩基編集は、P6P (サイレント)、E7A、及びS10P置換をもたらした。GFPをトランスフェクトし塩基エディターを含まなかった細胞は、塩基編集を示さなかった。従って、鎌状赤血球症SNPを塩基編集するこの能力は、鎌状赤血球症を治療するための組成物および方法において有用である。図10には、NGC PAMを受け入れるSpCas9 DNA結合ドメインを有するABE (ngcABE)を用いた、鎌状赤血球症標的部位における塩基編集のパーセントが示されている。

10

【0561】

gRNAは、以下の核酸配列：5'-GACUUCUCCACAGGAGUCAGA GUUUUAGAGC UAGA AAUAGC AAGUAAAAU AAGGCUAGUC CGUUAUCAAC UUGAAAAAGU GGCACCGAG U CGGUGCUUUU-3'を含んでおり、ここで、最初と最後の3つの塩基がホスホロチオエートと20-MEで修飾されていた。使用されたngcABE塩基エディターの配列は次のとおりである。

ATGAGCGAGGTGGAATTCAGCCACGAGTACTGGATGCGGCACGCCCTGACACTGGCCAA
AAGAGCTTGGGACGAGAGGGAAGTGCCTGTGGGAGCTGTGCTGGTGCACAACAACAGAG
TGATCGGCGAAGGCTGGAACAGACCCATCGGCAGACACGATCCTACAGCTCACGCCGAG
ATCATGGCCCTGAGACAAGGCGGACTGGTTCATGCAGAACTACCGGCTGATCGACGCCAC
ACTGTACGTGACCCTGGAACCTTTCGTGATGTGTGCCGGCGCTATGATCCACAGCAGAA
TCGGCAGAGTGGTGTTCGGCGCCAGAGATGCCAAAACAGGCGCTGCCGGAAGCCTGATG
GATGTGCTGCATCACCCCGGCATGAACCACAGAGTGGAAATCACCGAGGGGCATCCTGGC
CGATGAATGTGCCGCTCTGCTGAGCGACTTCTTCCGGATGCGGCGGCAAGAGATCAAGG
CCCAGAAGAAGGCCAGTCCAGCACAGATAGCGGCGGATCTAGCGGAGGCAGCTCTGGA
TCTGAGACACCTGGCACAAGCGAGAGCGCCACACCTGAAAGTTCTGGCGGTTCTTCTGG
CGGCAGCAGCGAGGTTCGAGTTCTCTACGAATATTGGATGAGACACGCTCTCACCCCTGG
CTAAGAGAGCCAGGGACGAAAGAGAGGTTGCCAGTTGGCGCTGTCCTGGTGTGTAACAAT
CGCGTCATCGGAGAAGGATGGAATCGCGCCATTGGCCTGCACGATCCAACCGCACATGC
CGAAATTATGGCTCTGCGGCAAGGCGGCCTCGTGATGCAAATTACAGACTGATCGATG
CTACCCTCTACGTCACCTTCGAGCCCTGTGTTCATGTGTGCTGGGGCAATGATTCACTCC
CGGATTGGCCGCGTGGTGTGTTGGAGTGCGGAATGCCAAGACTGGCGCCGCTGGATCTCT
GATGGACGTCCTGCACTATCCTGGGATGAACCACCGGGTCGAGATCACAGAGGGAATTC
TGGCTGACGAGTGCGCTGCCCTGCTGTGCTACTTCTTTAGAATGCCCAGACAGGTGTTC
AACGCCCAGAAAAAGCTCAGAGCAGCACCGATTCCGGCGGAAGCAGCGGAGGATCTTC
TGGAAGCGAAACCCAGGCACCAGCGAGTCTGCCACACCAGAATCATCTGGCGGTAGCT
CCGGCGGCAGCGACAAGAAGTATTCTATCGGACTGGCCATCGGCACCAACTCTGTTGGA
TGGGCCGTGATCACCGACGAGTACAAGGTGCCCAGCAAGAAATTCAAGGTGCTGGGCAA
CACCGACAGGCACAGCATCAAGAAGAACCTGATCGGCGCACTGCTGTTGACTCTGGCG
AAACAGCCGAGGCCACCAGACTGAAGAGAACAGCCCGCAGACGGTACACCAGAAGAAA
GAACCGGATCTGCTACCTCCAAGAGATCTTCAGCAACGAGATGGCCAAGGTGGACGACA
GCTTCTTCCACAGACTGGAAGAGTCCTTCTGGTGGAGAGGACAAGAAGCACGAGAGA
CACCCCATCTTCGGCAACATCGTGGACGAGGTGGCCTACCACGAGAAGTACCCACCAT
CTACCACCTGAGAAAGAACTGGTGGACAGCACCGACAAGGCCGACCTGAGACTGATCT
ATCTGGCCCTGGCTCACATGATCAAGTTCCGGGGCCACTTCTCTGATCGAGGGCGACCTG
AATCCTGACAACAGCGACGTGGACAAGCTGTTTCATCCAGCTGGTGCAGACCTACAACCA
GCTGTTTCGAGGAAAACCCCATCAACGCCAGCGGAGTGGATGCCAAGGCCATCCTGTCTG
CCAGACTGAGCAAGAGCAGACGGCTGGAAAATCTGATCGCCAGCTGCCTGGCGAGAAG
AAGAATGGCCTGTTTCGGCAACCTGATTGCCCTGAGCCTGGGCCTGACACCTAACTTCAA

20

30

40

50

GAGCAACTTCGACCTGGCCGAGGACGCCAAACTGCAGCTGAGCAAGGACACCTACGACG
ACGACCTGGACAATCTGCTGGCCCAGATCGGCGATCAGTACGCCGACTTGTTTCTGGCC
GCCAAGAATCTGAGCGACGCCATCCTGCTGTCCGACATCCTGAGAGTGAACACCGAGAT
CACCAAGGCACCTCTGAGCGCCTCTATGATCAAGAGATACGACGAGCACCACCAGGATC
TGACCCTGCTGAAGGCCCTCGTTAGACAGCAGCTGCCAGAGAAGTACAAAGAGATTTTC
TTCGACCAGAGCAAGAACGGCTACGCCGGCTACATTGATGGCGGAGCCAGCCAAGAGGA
ATTCTACAAGTTCATCAAGCCCATCCTCGAGAAGATGGACGGCACCGAGGAACTGCTGG
TCAAGCTGAACAGAGAGGACCTGCTGAGAAAAGCAGAGAACCTTCGACAACGGCAGCATC
CCTCACCAGATCCACCTGGGAGAACTGCACGCCATTCTGCGGAGACAAGAGGACTTTTA
CCCATTCTGAAGGACAACCGGGAAAAGATCGAGAAAATCCTGACCTTCAGGATCCCCT
ACTACGTGGGACCACTGGCCAGAGGCAATAGCAGATTTCGCTGGATGACCAGAAAGAGC
GAGGAAACCATCACTCCCTGGAACCTTCGAGGAAGTGGTGGACAAGGGCGCCAGCGCTCA
GTCCTTCATCGAGCGGATGACCAACTTCGATAAGAACCTGCCTAACGAGAAGGTGCTGC
CCAAGCACAGCCTGCTGTACGAGTACTTCACCGTGTACAACGAGCTGACCAAAGTGAAA
TACGTGACCGAGGGAATGAGAAAGCCCGCCTTTCTGAGCGGCGAGCAGAAAAAGGCCAT
CGTGGATCTGCTGTTCAAGACCAACCGGAAAGTGACCGTGAAGCAGCTGAAAGAGGACT
ACTTCAAGAAAATCGAGTGCTTCGACAGCGTCGAGATCTCCGGCGTGGAAGATCGGTTC
AATGCCAGCCTGGGCACATACCACGATCTGCTGAAAATTATCAAGGACAAGGACTTCCT
GGACAACGAAGAGAACGAGGACATCCTTGAGGACATCGTGCTGACACTGACCCTGTTTG
AGGACAGAGAGATGATCGAGGAACGGCTGAAAACATACGCCACCTGTTTCGACGACAAA
GTGATGAAGCAACTGAAGCGGCGGAGATACACCGGCTGGGGCAGACTGTCTCGGAAGCT
GATCAACGGCATCCGGGATAAGCAGTCCGGCAAGACCATCCTGGACTTTCTGAAGTCCG
ACGGCTTCGCCAACAGAACTTCATGCAGCTGATTACGACGACAGCCTCACCTTCAAA
GAGGATATCCAGAAAGCCCAGGTGTCCGGCCAGGGCGATTCTCTGCATGAGCACATTGC
CAACCTGGCCGGCTCTCCCGCCATTAAGAAAGGCATCCTGCAGACAGTGAAGGTGGTGG
ACGAGCTTGTGAAAGTGATGGGCAGACACAAGCCCGAGAACATCGTGATCGAAATGGCC
AGAGAGAACCAGACCACACAGAAGGGACAGAAGAACAGCCGCGAGAGAATGAAGCGGA
TCGAAGAGGGCATCAAAGAGCTGGGCAGCCAGATCCTGAAAGAACACCCCGTGGA AAAAC
ACCCAGCTGCAGAACGAGAAGCTGTACCTGTACTACCTGCAGAATGGACGGGATATGTA
CGTGGACCAAGAGCTGGACATCAACAGACTGTCCGACTACGATGTGGACCATATCGTGC
CCCAGTCTTTTCTGAAGGACGACTCCATCGACAACAAGGTCCTGACCAGATCCGACAAG
AATCGGGGCAAGAGCGACAACGTGCCCTCCGAAGAGGTGGTCAAGAAGATGAAGAACTA
CTGGCGACAGCTGCTGAACGCCAAGCTGATTACCCAGCGGAAGTTCGACAATCTGACCA
AGGCCGAAAGAGGGCGGCCTGAGCGAACTGGATAAGGCCGGCTTCATCAAGAGACAGCTG
GTGGA AACCCGGCAGATCACAAGCACGTGGCACAGATTCTGGACTCTCGGATGAACAC
TAAGTACGACGAGAACGACAACTGATCCGCGAAGTGAAAGTCATCACCTGAAGTCCA
AGCTGGTGTCCGATTTCCGGAAGGATTTCCAGTTCTACAAAGTGCGCGAGATCAACAAC
TACCATCACGCCACGACGCCTACCTGAATGCCGTTGTTGGAACAGCCCTGATCAAAAA
GTACCCTAAGCTGGAAAGCGAGTTCGTGTACGGCGACTACAAGGTGTACGACGTGCGGA
AGATGATCGCCAAGAGCGAGCAAGAGATTGGCAAGGCAACCGCCAAGTACTTCTTCTAC
AGCAACATCATGAACTTTTTCAAGACAGAGATCACCTCGCCAACGGCGAGATCAGAAA
GCGGCCTCTGATCGAGACAAACGGCGAAACCGGGCAGATTGTGTGGGATAAGGGCAGAG
ACTTTGCCACAGTGCGGAAAGTGCTGAGCATGCCCAAGTGAATATCGTGAGAAAACC
GAGGTGCAGACAGGCGGCTTCAGCAAAGAGTCTATCCTGCCTAAGCGGAACTCCGACAA
GCTGATCGCCAGAAAGAAGGACTGGGACCCCAAGAAATACGGCGGCTTTATGCAGCCCA
CCGTGGCCTATtctGtTCTGGTGGTGGCCAAAGTGGA AAAAGGGCAAGTCCAAGAACTCA
AGAGCGTGAAAGAGCTGCTGGGGATCACCATCATGGAAAGAAGCAGCTTCGAGAAGAAT
CCGATCGATTTCTCGAGGCCAAGGGTTACAAAGAAGTGAAAAAGGACCTGATCATCAA
GCTCCCCAAGTACTCCCTGTTTCGAGCTGGAAAACGGCCGGAAGAGAATGCTGGCCTCTG
CcAAGttcCTGCAGAAGGGAAACGAACTGGCCCTGCCTAGCAAATATGTGAACTTCCTGT

10

20

30

40

50

ACCTGGCCAGCCACTATGAGAAGCTGAAGGGCAGCCCCGAGGACAATGAGCAAAAGCAG
CTGTTTGTGGAACAGCACAAAGCACTACCTGGACGAGATCATCGAGCAGATCAGCGAGTT
TAGCAAGAGAGTGATTCTGGCCGACGCCAATCTGGACAAAGTGCTGTCCGCCTACAACA
AGCACCGGGACAAGCCTATCAGAGAGCAGGGCCGAGAATATCATCCACCTGTTTACCCTG
ACCAACCTGGGAGCCCCCTagaGCCTTCAAGTACTTTGACACCACCATCgcccGGAAGGAG
TACcggTCCACCAAAGAGGTGCTGGACGCCACTCTGATCCACCAGTCTATCACCGGCCTG
TACGAGACACGGATCGACCTGTCTCAACTCGGAGGCGACGAAGGCGCCGATAAGAGAAC
CGCCGATGGCTCTGAGTTCGAGAGCCCTAAGAAAAAGCGCAAAGTGTag

【0562】

【実施例6．鎌状赤血球症 (SCD) / 遺伝性高胎児性ヘモグロビン症 (HPFH) に関連するHBG1/2遺伝子の調節領域の編集】

10

HPFHは、生涯を通じて高レベルHbF発現を起こすヘテロ接合欠失またはヌクレオチド置換を伴う良性遺伝的状態である。胎児ヘモグロビン (HbF) レベルの増加は、βサラセミアや鎌状赤血球貧血のようなβ-グロビン遺伝子発現の遺伝性障害の臨床経過を改善することができる。遺伝性胎児ヘモグロビン遺残症 (HPFH) と呼ばれる障害の群では、HbFのγグロビン遺伝子 (HBG1/2) の一方または両方の発現が成人赤血球細胞において高レベルで持続する。

【0563】

γ-グロビン遺伝子発現の正常なパターンのためのいくつかの調節エレメントならびにHPFHの欠失および非欠失型がForget B.G. “Molecular basis of hereditary persistence of fetal hemoglobin.” Ann NY Acad Sci. 1998 Jun 30;850:38-44に記載されており、その内容全体が参照により本明細書に組み込まれる。

20

【0564】

HPFHの非欠失型は、HBG1またはHBG2 γ-グロビン遺伝子のいずれかのプロモーター領域における点突然変異の存在によって特徴付けられる。このような点突然変異は、種々の転写因子と遺伝子のプロモーターとの間の相互作用を変化させると考えられている。HPFHの欠失型は、通常であれば離れたところにあるシス作用性因子がγ-グロビン遺伝子の近傍に並置されるため、γ-グロビン遺伝子発現の正常な発生的パターンを脱制御し得る。上記表4は、本明細書に記載の塩基エディターおよび方法によってそのような遺伝子の調節領域において影響を受け得るヌクレオチド変化および関連する遺伝子（例えばHBG1/2遺伝子）を提示する。残存的胎児ヘモグロビンを誘導するために、HBGに関連する遺伝子は、表4に示す遺伝領域について標的化される。これらの遺伝子標的部位は、遺伝性持続胎児ヘモグロビン (HPFH) 患者で見出される自然発生の変異/SNPを再現するために有用である。

30

【0565】

HBG1/2プロモーターの調節領域の非調節領域を標的化するためにガイドRNA (gRNA) を設計した(図11)。示されたgRNAおよびCas9塩基エディターでトランスフェクトされた293T細胞における標的塩基編集率を図12に示す。塩基編集効果のパーセンテージはMiseqにより決定した。市販のトランスフェクション試薬 (Mirus) を用いて、C塩基エディターまたはA塩基エディターのいずれかを含有するプラスミドDNAを、対応するgRNAと共にHEK293Tにトランスフェクトした。図12および図13において、gRNAの標的配列は以下の通りである：

40

gRNA1: CTTGACCAATAGCCTTGACAAGG;
gRNA3: CAAGGCTATTGGTCAAGGCA;
gRNA4: GCTATTGGTCAAGGCAAGGC;
gRNA11: CTTGCCTTGACCAATAGCCTTGA;
gRNA42: TCAGACAGATATTTGCATTGAGA;
gRNA 43: TTTGAGACAGATATTTGCATTGA;
gRNA44: TAGCCTTGACAAGGCAAACCTTGA;
gRNA45: AAGTTTGCCTTGTCAAGGCTATTGGT

50

【0566】

上述し図12および13にも記載されたgRNA1、gRNA4およびgRNA45が、「A」塩基エディター (ABE) および「C」塩基エディター (CBE) とともに使用され、残りのgRNAは「C」塩基エディター (CBE) と共に使用された。一実施形態では、gRNA足場配列は以下の通りである：GUUUUAGAGC UAGAAAUAGC AAGUUAUUU AAGGCUAGUC CGUUAU CAAC UUGAAAAAGU GGCACCGAGU CGGUGCUUUU。

【0567】

CD34+細胞へのRNAトランスフェクションのために、ヒト骨髄CD34+細胞をHemaCareから購入した。ガイドRNA核酸配列の最初と最後の3塩基対が2'-OMeで修飾された化学修飾gRNAがSynthegoにより合成され、塩基エディターmRNAはTriLinkにより転写された。例として、mRNA編集のためには、mRNAはN1MePseudoU修飾mRNAであり得る。当業者には理解されるように、2'-O-メチル化はRNAの一般的なヌクレオシド修飾であり、ヌクレオシドのリボース部分の2'ヒドロキシルにメチル基が付加されてメトキシ基を生成する。CD34+細胞を解凍し、2〜3日間培養して増殖させた後、エレクトロポレーションに供した。500 ngのgRNAおよび2500 ngの塩基エディターmRNAを、製造業者の指示に従って20万のCD34+細胞にエレクトロポレーションした。RNAエレクトロポレーションの3日後にゲノムDNAを採取した。QuickExtract™ DNA抽出溶液 (Epicentre) を用いて65°Cで30分間ゲノムDNAを抽出した後、65°Cで5分間加熱活性化した。図13は、示されたgRNAおよび塩基エディターでトランスフェクトされた初代ヒト骨髄CD34+細胞における標的化塩基編集率を示す。HBG1およびHBG2遺伝子座の両方における塩基編集のパーセンテージを、Miseqにより評価した。ゲノム部位をPCR増幅し、MiSeq上で配列決定した。そのような評価のために当該技術分野で典型的に使用される分析方法および手順を使用して、各位置における塩基頻度およびインデルのパーセントについて結果を分析した。

【0568】

[実施例7. 材料と方法]

本明細書に記載される実施例で提供された結果は、以下の材料および方法を用いて得られた。

[クローニング/トランスフェクション]

PCRはVeraSeq Ultra DNAポリメラーゼ (Enzymatics) またはQ5 Hot Start High-Fidelity DNAポリメラーゼ(New England Biolabs)を用いて行った。USERクローニング(New England Biolabs)を用いて塩基エディター (BE) プラスミドを構築した。デアミナーゼ遺伝子はgBlocks Gene Fragments (Integrated DNA Technologies)として合成された。用いたCas9遺伝子は下記に示す。Cas9遺伝子は以前に報告されたプラスミドから得られた。デアミナーゼと融合遺伝子をpCMV (哺乳類コドン最適化)またはpET28b (大腸菌コドン最適化)の骨格にクローニングした。sgRNA発現プラスミドは、部位特異的変異誘発を用いて構築した。例えば、gRNAプラスミドは、gRNAの発現を駆動する、当技術分野で認識されるU6プロモーターを含んだ。BEプラスミド、例えば、ABEプラスミドについては、CMVプロモーターを用いて、塩基エディターのオープンリーディングフレームの発現を駆動した。

【0569】

簡単に説明すると、プライマーは、製造業者の指示に従ってT4ポリヌクレオチドキナーゼ(New England Biolabs)を用いて5'リン酸化した。次に、リン酸化プライマー、およびテンプレートとしての、目的の遺伝子をコードする発現プラスミドと共に、Q5 Hot Start High-Fidelity Polymerase (New England Biolabs)を用いて、製造業者の指示に従ってPCRを実施した。PCR産物をDpnI (20 U、New England Biolabs)と共に37°Cで1時間インキュベートし、QIAprepスピンカラム (Qiagen) で精製し、製造業者の指示に従ってQuickLigase (New England Biolabs)を用いてライゲーションした。DNAベクター増幅はMach1コンピテント細胞(ThermoFisher Scientific)を用いて行った。

gRNAについて、以下の足場配列を示す：GUUUUAGAGC UAGAAAUAGC AAGUUAUUU

AU AAGGCUAGUC CGUUAUCAAC UUGAAAAAGU GGCACCGAGU CGGUGCUUUU。この足場は、本明細書の表に示されるPAM、例えばNGG、NGA、NGC、NGT PAMのために使用された；gRNAは、このスキャフォールド配列と、本明細書に提供されるか、または当業者の知識に基づいて決定されかつ当業者に理解されるような、疾患関連遺伝子(例えば表3A、3Bおよび4)のためのおよびスパーサー配列(標的配列)とを包含する。(例えば、Komor, A.C., et al., “Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage” *Nature* 533, 420-424 (2016); Gaudelli, N.M., et al., “Programmable base editing of A・T to G・C in genomic DNA without DNA cleavage” *Nature* 551, 464-471 (2017); 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” *Science Advances* 3:eaa04774 (2017), and Rees, H.A., et al., “Base editing: precision chemistry on the genome and transcriptome of living cells.” *Nat Rev Genet.* 2018 Dec;19(12):770-788. doi: 10.1038/s41576-018-0059-1参照。)

【0570】

[ssDNAに対するインビトロのデアミナーゼアッセイ]

すべてのssDNA基質の配列を下記に示す。Cy3標識基質はすべてIntegrated DNA Technologies (IDT) から入手した。デアミナーゼは、1 μgのプラスミドを使用して、製造業者の指示に従ってTNT T7 Quick Coupled Transcription/Translation Kit (Promega) を使用してin vitroで発現させた。タンパク質発現に続いて、5 μlの溶解物を、CutSmart緩衝液(New England Biolabs) (50 mM酢酸カリウム、29 mMトリス酢酸、10 mM酢酸マグネシウム、100 μg ml⁻¹ BSA、pH 7.9)中で35 μlのssDNA (1.8 μM)およびUSER酵素(1ユニット)と組み合わせ、37°Cで2時間インキュベートした。切断されたU含有基質を、10% TBE - 尿素ゲル (Bio - Rad) 上で完全長未改変基質から分離した。

【0571】

[BE-リンカー-結合タンパク質融合体の発現および精製]

大腸菌BL21 STAR (DE 3) コンピテント細胞(ThermoFisher Scientific)などのコンピテント細胞を、リンカーを介して核酸結合タンパク質に融合された塩基エディター (BE) (例えばBE-リンカー-dCas9) をコードするプラスミドで形質転換した。得られた発現株を、カナマイシン100 μg ml⁻¹を含むLuria-Bertani (LB) ブロス中で37°Cで一晩増殖させた。細胞を同じ増殖培地に1:100希釈し、OD600=〜0.6になるまで37°Cで増殖させた。培養物を4°Cで2時間冷却し、イソプロピル-β-d-1-チオガラクトピラノシド (IPTG) を0.5 mMで添加してタンパク質発現を誘導した。約16時間後、4,000 gで遠心分離して細胞を集め、溶解緩衝液(50 mMトリス(ヒドロキシメチル)アミノメタン(Tris) -HCl (pH 7.5)、1M NaCl、20%グリセロール、10 mMトリス(2-カルボキシエチル)ホスフィン(TCEP, Soltec Ventures))に再懸濁した。細胞を超音波処理(20秒パルス・オン、20秒パルス・オフ、合計8分間、6 W出力)によって溶解し、25,000 gで15分間遠心分離した後に溶解物上清を単離した。溶解物をHis-Purニッケル-ニトリロ酢酸 (ニッケル-NTA) 樹脂(ThermoFisher Scientific)と共に4°Cで1時間インキュベートし、His標識融合タンパク質を捕捉した。樹脂をカラムに移し、溶解緩衝液40 mlで洗浄した。Hisタグ融合タンパク質を、285 mMイミダゾールを補充した溶解緩衝液中で溶出し、限外濾過(100 kDa分子量カットオフのAmicon-Millipore)によって全容量1 mlに濃縮した。タンパク質を、50 mMトリス(ヒドロキシメチル)-アミノメタン (Tris) -HCl (pH 7.0)、0.1 M NaCl、20%グリセロール、10 mM TCEPを含む低塩精製緩衝液中で20mlに希釈し、SPセファローズ Fast Flow樹脂(GE Life Sciences)上にロードした。樹脂をこの低塩緩衝液40 mlで洗浄し、50 mMトリス(ヒドロキシメチル)アミノメタン (Tris) -HCl (pH 7.0)、0.5 M NaCl、20%グリセロール、10 mM TCEPを含む活性緩衝液5 mlでタンパク質を溶出した。溶出したタンパク質をSDS-PAGEで定量した。

【0572】

[sgRNAのin vitro転写]

T7プロモーターに続いてsgRNA標的配列を含む直鎖DNA断片を、製造業者の指示に従ってTranscriptAid T 7 High Yield Transcription Kit (ThermoFisher Scientific) と共にプライマー配列を用いて、in vitroで転写した。製造業者の指示に従ってMEGAClear キット(ThermoFisher Scientific)を用いてsgRNA産物を精製し、UV吸収により定量した。

【0573】

【Cy3共役dsDNA基質の調製】

典型的には、ラベルされていない鎖、例えば長さ80 ntの配列を、PAGE精製オリゴヌクレオチドとしてIntegrated DNA Technologies (IDT) にオーダーした。標識プライマーは、各基質、例えば80 nt基質の3'末端に相補的である。このプライマーは、HPLC精製オリゴヌクレオチドとしてIDTにオーダーした。Cy3標識dsDNA基質を生成するために、80 nt鎖(5 µlの100 µM溶液)を、Cy3標識プライマー(5 µlの100 µM溶液)とNEBuffer 2 (38.25 µlの50 mM NaCl、10 mM Tris-HCl、10 mM MgCl₂、1 mM DTT、pH 7.9 溶液、New England Biolabs)中でdNTP (0.75 µlの100 mM溶液)と組み合わせ、95°Cで5分間加熱し、その後45°Cまで0.1°C/秒の速度で徐々に冷却した。このアニーリング期間の後、Klenow exo-(5U、New England Biolabs)を加え、反応物を37°Cで1時間インキュベートした。溶液を緩衝液PB (250 µl、Qiagen)およびイソプロパノール(50 µl)で希釈し、QIAprepスピンカラム(Qiagen)で精製し、50 µlのTris緩衝液で溶出した。dsDNAに対するデアミナーゼアッセイ。精製した融合タンパク質(活性緩衝液中1.9 µMを20 µl)を1当量の適切なsgRNAと組み合わせ、周囲温度で5分間インキュベートした。Cy3標識dsDNA基質を125 nMの最終濃度で添加し、得られた溶液を37°Cで2時間インキュベートした。緩衝液PB (100 µl、Qiagen)およびイソプロパノール(25 µl)の添加により融合物からdsDNAを分離し、EconoSpinマイクロスピンカラム(EPOCH Life Science)で精製し、20 µlのCutSmart緩衝液(New England Biolabs)で溶出した。精製された編集dsDNAにUSER酵素(1 U、New England Biolabs)を添加し、37°Cで1時間インキュベートした。5 µlの反応溶液を15 µlのDMSOベースのローディング緩衝液(5 mM Tris、0.5 mM EDTA、12.5%グリセロール、0.02%ブロモフェノールブルー、0.02%キシレンシアン、80% DMSO)と組み合わせることによって、Cy3標識鎖をその相補鎖から完全に変性させた。全長C含有基質を、10% TBE-尿素ゲル (Bio-Rad) 上で、切断されたU含有編集基質から分離し、GE Amersham Typhoonイメージャー上で画像化した。

【0574】

【ハイスループットシーケンシングのためのin vitro編集dsDNAの調製】

オリゴヌクレオチドはIntegrated DNA Technologies (IDT) から入手した。相補的配列をTris緩衝液中で組み合わせ(5 µlの100 µM溶液)、95°Cに5分間加熱し、続いて45°Cまで0.1°C/秒の速度で徐々に冷却することによってアニールし、60 bpのdsDNA基質を生成した。精製融合タンパク質(活性緩衝液中1.9 µMを20 µl)を1当量の適切なsgRNAと組み合わせ、周囲温度で5分間インキュベートした。60マーdsDNA基質を125 nMの最終濃度で添加し、得られた溶液を37°Cで2時間インキュベートした。緩衝液PB (100 µl、Qiagen)およびイソプロパノール(25 µl)の添加により融合物からdsDNAを分離し、EconoSpinマイクロスピンカラム(EPOCH Life Science)で精製し、20 µlのTris緩衝液で溶出した。得られた編集されたDNA (1 µlをテンプレートとして使用した)を、製造業者の指示に従ってハイスループット配列決定プライマー対およびVeraSeq Ultra (Enzymatics) を用いて13サイクルの増幅を伴うPCRによって増幅した。RapidTips (Diffinity Genomics)を用いてPCR反応生成物を精製し、精製したDNAを、シーケンシングアダプターを含むプライマーを用いてPCRにより増幅し、精製し、先に記載したようにMiSeqハイスループットDNAシーケンサー (Illumina) 上で配列決定した。

【0575】

【細胞培養】

HEK293T (ATCC CRL-3216)、U2OS (ATCC HTB-96)、CD34+細胞、または他の関連する細胞型を、10% (v/v) ウシ胎児血清 (FBS) を補充したダルベッコの改変イーグル

培地プラスGlutaMax (ThermoFisher) 中で、37°C、5% CO₂で維持した。HCC1954細胞(ATCC CRL-2338)を、上記のように補充したRPMI-1640培地(ThermoFisher Scientific)中で維持した。標的配列を含有する遺伝子を含有する不死化細胞(Taconn Biosciences)を、10% (v/v) ウシ胎児血清 (FBS) および200 µg ml⁻¹ゲネチシン(ThermoFisher Scientific)を補足したダルベッコの改変イーグル培地プラスGlutaMax (ThermoFisher Scientific)中で培養した。

【0576】

[トランスフェクション]

HEK293Tまたは他の宿主細胞を、48ウェルのコラーゲン被覆BioCoatプレート (Corning) 上に播種し、約85%コンフルエントの状態で作製した。簡単に述べれば、750 ngのBEおよび250 ngのsgRNA発現プラスミドを、1.5 µlのリポフェクタミン2000 (ThermoFisher Scientific)/ウェルを用いて、製造業者のプロトコルに従ってトランスフェクトした。HEK293T細胞を、メーカーの指示に従って、適切なAmaxa Nucleofector IIプログラムを用いてトランスフェクトした(HEK 293T細胞のためのプログラムQ-001を用いるVキット)。

【0577】

[ゲノムDNA試料のハイスループットDNA配列決定]

トランスフェクトした細胞を3日後に収穫し、Agencourt DNAdvance Genomic DNA Isolation Kit (Beckman Coulter)を製造業者の指示に従って用いてゲノムDNAを単離した。対象のオン・ターゲットおよびオフ・ターゲットゲノム領域を、隣接するハイスループット配列決定プライマー対を用いたPCRによって増幅した。PCR増幅は、5 ngのゲノムDNAをテンプレートとして使用し、Phusion高忠実度DNAポリメラーゼ (ThermoFisher)を用いて、製造業者の指示に従って実施した。反応が直線的増幅範囲内で停止することを確認するために、各プライマー対について別々にサイクル数を決定した。RapidTips (Diffinity Genomics)を用いてPCR産物を精製した。精製DNAを、配列決定アダプターを含むプライマーを用いたPCRにより増幅した。生成物をゲル精製し、Quant-iT PicoGreen dsDNA Assay Kit (ThermoFisher) およびKAPA Library Quantification Kit-Illumina (KAPA Biosystems)を用いて定量した。サンプルは、前述のようにIllumina MiSeq上で配列決定された(Pattanayak, Nature Biotechnol. 31, 839-843 (2013))。

【0578】

[データ分析]

配列決定リード (read) は、MiSeq Reporter (Illumina) を使用して自動的に脱マルチプレックス化され、個々のFASTQファイルを、カスタムMatlabを使用して分析した。各リードは、Smith-Watermanアルゴリズムを使用して、適切な参照配列にペアワイズでアラインメントされた。Qスコアが31未満の塩基コールはNに置き換え、ヌクレオチド頻度の計算から除外した。この処理により、予測されるMiSeq塩基呼び出しエラー率は約1,000分の1になる。リードと参照配列がギャップを含まないアラインメント配列は、アラインメントテーブルに保存し、そこから各遺伝子座について塩基頻度を表にすることができた。インデル頻度は、前述の基準 (Zuris, et al., Nature Biotechnol. 33, 73-80 (2015)) を用いて、カスタムのMatlabスクリプトで定量した。配列決定リードをスキャンして、インデル (indel) が生じる可能性のあるウィンドウの両側に隣接する二つの10 bp配列との正確なマッチを探した。完全なマッチが見つからなかった場合、そのリードは分析から除外した。このインデルウィンドウの長さが参照配列と正確に一致した場合、そのリードはインデルを含まないものとして分類された。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、その配列決定リードは、それぞれ挿入または欠失として分類された。

【0579】

[他の実施形態]

上述の説明から、種々の用途および条件に適用するために、本明細書に記載する本発明に変形および修正を加えることができることが明らかであろう。そのような実施形態もま

10

20

30

40

50

た、特許請求の範囲内である。

【0580】

本明細書における変数の定義における要素のリストの記載は、単一要素としての、またはリストされた要素の組合せ(またはサブコンビネーション)としてのその変数の定義を含む。本明細書における実施形態の説明は、単一の実施形態として、または任意の他の実施形態またはその一部と組み合わせての、その実施形態を含む。

【0581】

本明細書で言及されているすべての特許および刊行物は、それぞれの独立した特許および刊行物が、参照により組み入れられることが具体的かつ個別に示されている場合と同じ程度に、参照により本明細書に組み入れられる。

10

20

30

40

50

【図面】

【図 1】

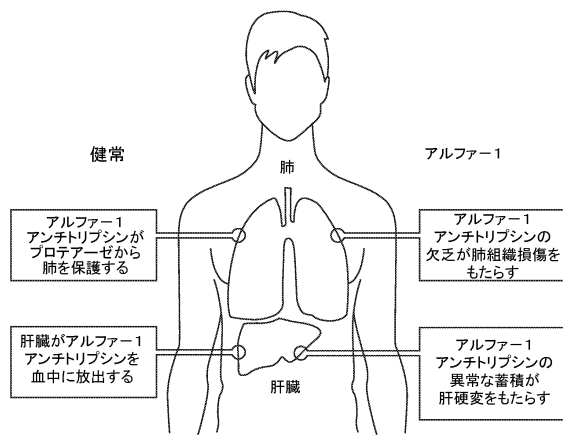


FIG. 1

【図 2】

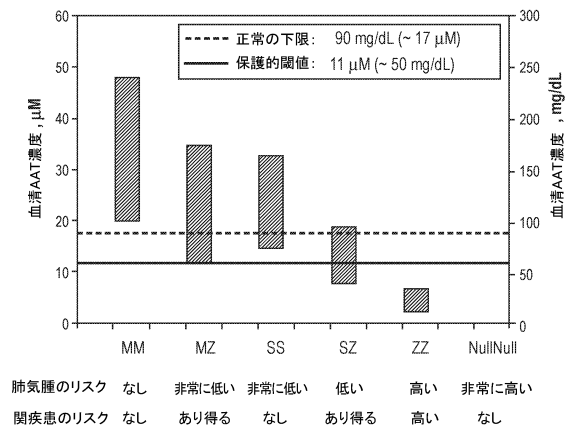


FIG. 2

【図 3】



FIG. 3

【図 4】

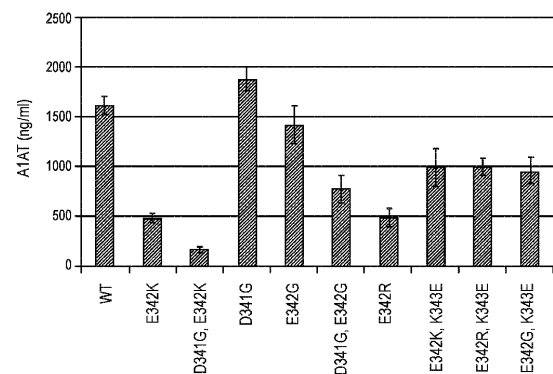


FIG. 4

【図 8 A】

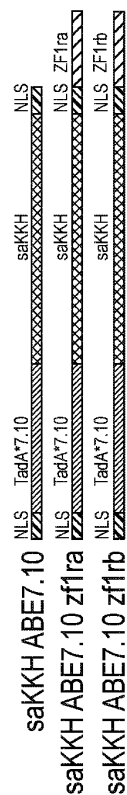


FIG. 8A

【図 8 B】



FIG. 8B

【図 8 C】

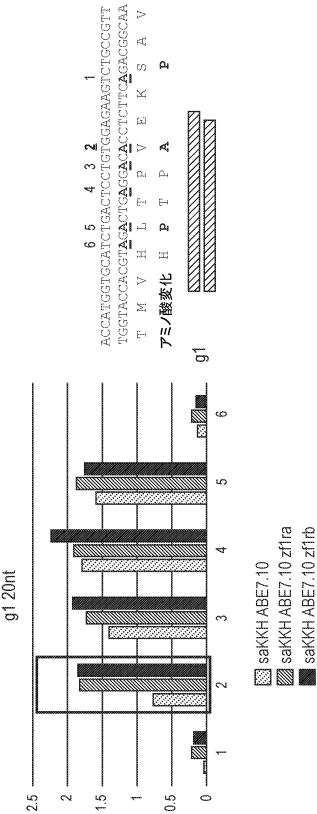


FIG. 8C

【図 8 D】

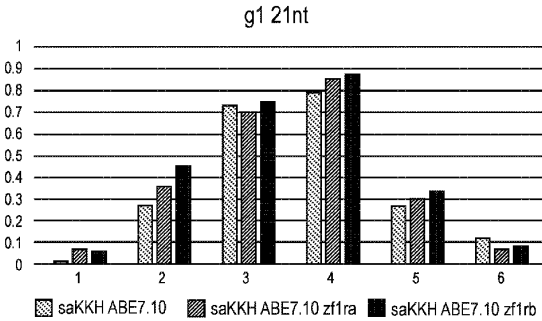


FIG. 8D

10

20

30

40

50

【図 8 F】

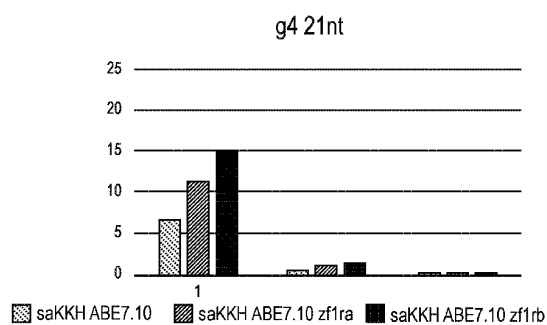


FIG. 8F

【図 9 A】

[illegible]

FIG. 9A

【図 9 B】

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	L	P	A	M	
	G	C	T	G	A	A	T	G	A	A	G	T	T	A	T	G	A	T	C	G	N	N	A	G	A
ABC STI-CAS9																									
A	0.0	0.0	0.0	0.1	98.8	96.3	0.0	0.0	96.1	95.0	0.0	0.0	0.0	95.5	0.0	0.0	100.0	0.0	0.0	0.0	100.0	100.0	100.0	0.0	100.0
C	0.0	99.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	99.9	0.0	0.0	0.0	0.0	0.0	0.0	
G	99.9	0.0	0.0	99.9	1.2	3.7	0.0	99.9	9.9	5.0	99.9	0.0	0.0	4.5	0.0	100.0	0.0	0.0	100.0	0.0	0.0	99.9	0.0	0.0	
T	0.0	0.0	100.0	0.0	0.0	0.0	100.0	0.1	0.0	0.0	100.0	100.0	0.0	100.0	0.0	100.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	
非处理																									
A	0.0	0.0	0.0	0.0	100.0	100.0	0.0	0.0	100.0	100.0	0.0	0.0	0.0	100.0	0.0	0.1	99.7	0.1	0.1	0.0	100.0	99.7	100.0	0.0	100.0
C	0.0	99.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	99.8	0.0	0.0	0.3	0.0	0.0	0.0	
G	0.0	0.0	0.0	99.9	0.0	99.9	0.0	99.9	0.0	99.9	0.0	0.0	0.0	0.0	0.0	99.7	1.0	0.0	99.9	0.0	0.0	100.0	0.0	100.0	
T	0.0	0.1	100.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	100.0	100.0	0.0	100.0	0.0	100.0	0.1	0.2	99.9	0.1	0.0	0.0	0.0	0.0	

FIG. 9B

	インデル%
ABE-ST1 CAS9	0.07%
ST1-CAS9 スクレアーゼ	25.5%
非処理	0.04%

【図 9 C】

	P	A	M		20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1
	T	G	G	T	N	N	T	T	T	G	A	C	T	T	A	C	A	G	T	G	G	G	G	C
ABE ST1-CAS9	A	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	100.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
A	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
C	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	77.1	0.0	100.0	0.0	0.0	5.3	0.0	0.0	0.0	0.0	0.0
G	0.0	99.9	100.0	0.0	100.0	0.0	0.0	100.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	99.9	100.0	100.0	100.0	100.0	0.0
T	0.0	0.0	0.0	100.0	0.0	0.0	100.0	0.0	99.9	0.0	100.0	0.0	0.0	92.3	0.0	0.0	0.0	94.1	0.1	0.0	0.0	0.0	0.0	0.0
非処理	A	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	100.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
C	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
A	0.0	100.0	100.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
T	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
G	0.0	100.0	100.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
T	0.0	0.0	0.0	100.0	0.0	0.0	100.0	0.0	100.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0

FIG. 9C

	インデル%
ABE-ST1 CAS9	0.16%
ST1-CAS9 スクレアーゼ	2.8%
非処理	0.12%

【図 9 D】

	T	G	G	T	N	C	A	G	G	A	T	T	G	T	C	T	T	A	C	T	T	G	T	C
ABE ST-CAS9																								
0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	99.9	100.0	0.0	0.0	0.0	0.1	0.3	0.0	0.7	100.0	0.5	0.4	4.2	0.0	100.0	0.4	0.2	0.0	0.2	99.9	0.0
0.0	99.9	99.9	0.0	99.9	0.0	0.0	99.9	100.0	0.0	0.0	0.0	99.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0
100.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	99.9	99.9	99.7	0.0	99.3	99.5	99.9	99.8	0.0	0.0	99.8	99.8	0.0	99.8	0.0
非处理																								
0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	100.0	99.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	100.0
0.0	99.9	99.9	0.0	99.9	0.0	0.0	100.0	99.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	99.9	0.0
100.0	0.0	0.0	100.0	0.0	0.0	0.1	0.0	0.0	0.0	100.0	100.0	100.0	0.0	100.0	0.0	0.0	0.0	100.0	100.0	100.0	0.0	100.0	100.0	0.0

FIG. 9D

	インデル%
ABE-ST1 CAS9	0.10%
ST1-CAS9 スクレアーゼ	8.73%
非処理	0.08%

【図 9 E】

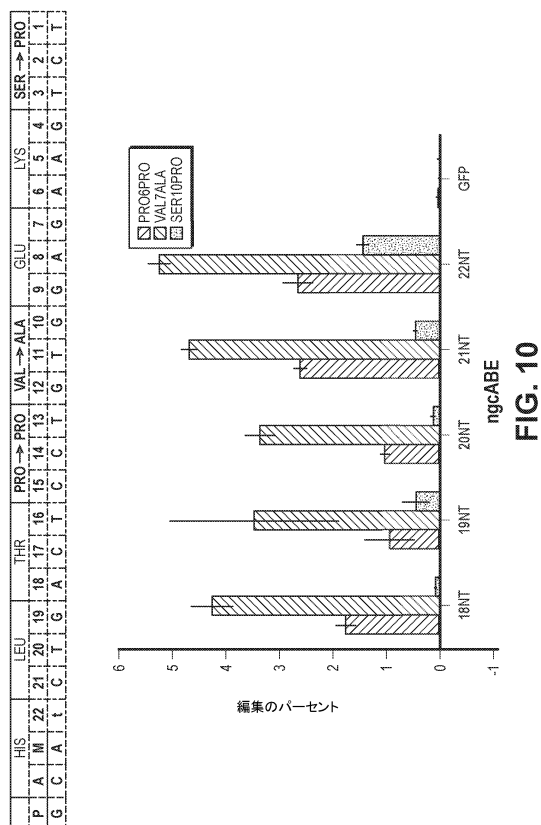
VAL				HIS				LEU				THR				PRO→PRO				VAL→ALA				GLU				LYS			
P	A	M	T	21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1							
T	G	A	T	G	C	A	T	C	T	G	A	C	T	C	T	T	G	T	G	A	G	A	A	G	T						
ABE ST1-CAS9																															
C	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0							
A	0.0	0.0	0.0	0.0	0.0	0.0	99.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0							
M	0.0	99.9	99.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	99.9	99.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0							
T	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	99.3	99.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0							
G	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	99.0	99.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0							
A	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	99.3	99.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0							
M	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	99.0	99.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0							
T	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	99.3	99.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0							
G	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	99.0	99.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0							
非处理																															
C	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0							
A	0.0	0.0	0.0	0.0	0.0	0.0	99.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	99.9	99.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0							
M	0.0	99.9	100.0	0.0	99.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	99.9	99.9	100.0	0.0	0.0	0.0	0.0	0.0	0.0							
T	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	99.9	99.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0							
G	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0							
A	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0							
M	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0							
T	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0							
G	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0							

FIG. 9E

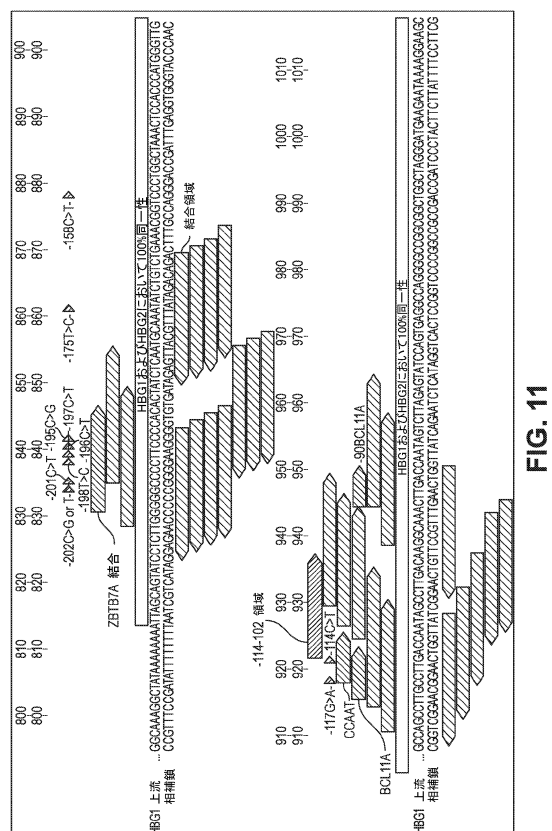
インデル%	
ABE-ST1 CAS9	0.07%
ST1-CAS9 スクレアーゼ	2.30%
非処理	0.04%

鑑状赤血球症標的部位においてABE-STICAS9により誘導されたA・TからG・Cへの変異！

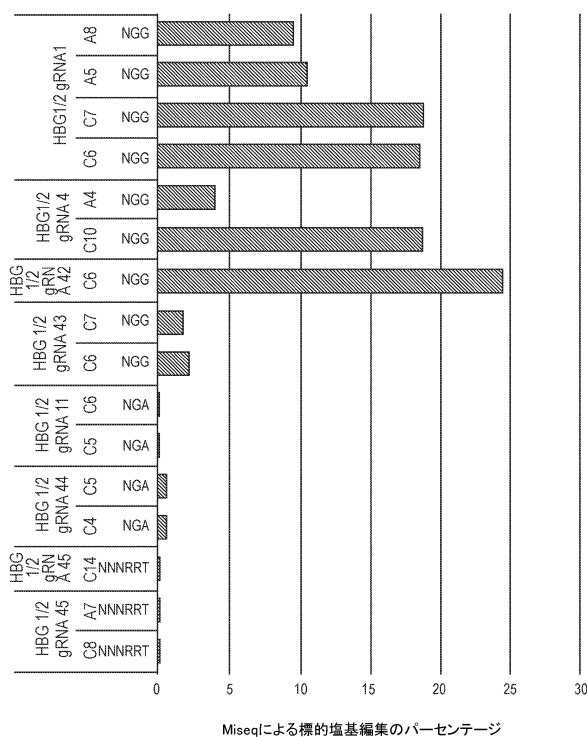
【図 10】



【図 11】



【図 12】



【図 13】

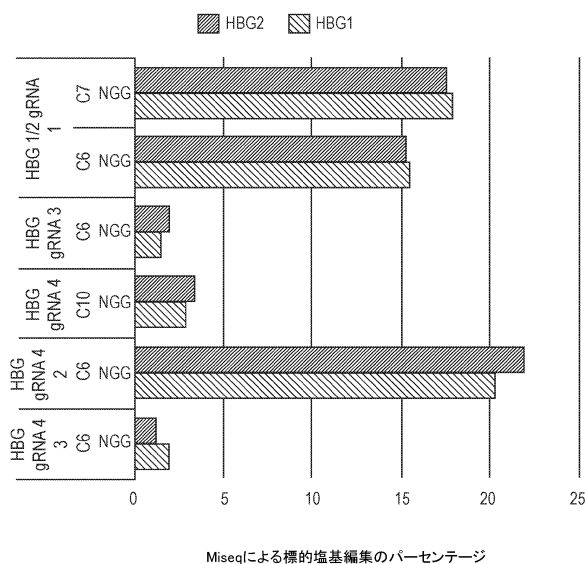


FIG. 13

【配列表】
0007642531000001.app

10

20

30

40

50

フロントページの続き

(51)国際特許分類

F I

C 1 2 N	15/62	(2006.01)	C 1 2 N	15/62	Z
C 1 2 N	15/55	(2006.01)	C 1 2 N	15/55	
C 1 2 N	9/16	(2006.01)	C 1 2 N	9/16	Z
A 6 1 P	7/00	(2006.01)	A 6 1 P	7/00	
A 6 1 K	35/18	(2015.01)	A 6 1 K	35/18	
A 6 1 K	35/28	(2015.01)	A 6 1 K	35/28	
A 6 1 K	48/00	(2006.01)	A 6 1 K	48/00	
A 6 1 P	25/28	(2006.01)	A 6 1 P	25/28	
A 6 1 K	38/50	(2006.01)	A 6 1 K	38/50	
A 6 1 P	1/16	(2006.01)	A 6 1 P	1/16	
A 6 1 P	11/00	(2006.01)	A 6 1 P	11/00	
A 6 1 K	38/55	(2006.01)	A 6 1 K	38/55	

(33)優先権主張国・地域又は機関

米国(US)

(31)優先権主張番号 62/780,890

(32)優先日 平成30年12月17日(2018.12.17)

(33)優先権主張国・地域又は機関

米国(US)

前置審査

セッツ州 ケンブリッジ ランズダウン ストリート 2 6 2nd フロア

(72)発明者 エヴァンス、 ジョン

アメリカ合衆国 0 2 1 3 9 マサチューセッツ州 ケンブリッジ ランズダウン ストリート 2 6 2nd フロア

(72)発明者 バッカー、 マイケル

アメリカ合衆国 0 2 1 3 9 マサチューセッツ州 ケンブリッジ ランズダウン ストリート 2 6 2nd フロア

(72)発明者 フ、 ヤンファン

アメリカ合衆国 0 2 1 3 9 マサチューセッツ州 ケンブリッジ ランズダウン ストリート 2 6 2nd フロア

(72)発明者 ゴーデッリ、 ニコル

アメリカ合衆国 0 2 1 3 9 マサチューセッツ州 ケンブリッジ ランズダウン ストリート 2 6 2nd フロア

(72)発明者 ゲールケ、 ジェイソン マイケル

アメリカ合衆国 0 2 1 3 9 マサチューセッツ州 ケンブリッジ ランズダウン ストリート 2 6 2nd フロア

(72)発明者 ジョン、 ジェイ、 キース

アメリカ合衆国 0 2 1 2 9 マサチューセッツ州 チャールズタウン 1 3 ス ストリート 1 4 9 6 ス フロア マサチューセッツ ジェネラル ホスピタル

審査官 松井 一泰

(56)参考文献 特表 2 0 1 6 - 5 2 8 8 9 0 (J P, A)

(58)調査した分野 (Int.Cl., D B名)

C 1 2 N 1 5 / 0 0 - 1 5 / 9 0

C 1 2 N 1 / 0 0 - 7 / 0 8

C A p l u s / M E D L I N E / E M B A S E / B I O S I S (S T N)