

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7586601号
(P7586601)

(45)発行日 令和6年11月19日(2024.11.19)

(24)登録日 令和6年11月11日(2024.11.11)

(51)国際特許分類		F I	
A 6 1 K	31/7088(2006.01)	A 6 1 K	31/7088
C 1 2 N	15/09 (2006.01)	C 1 2 N	15/09 1 0 0
C 1 2 N	15/867(2006.01)	C 1 2 N	15/867 Z
C 1 2 N	15/864(2006.01)	C 1 2 N	15/864 1 0 0 Z
C 1 2 N	5/10 (2006.01)	C 1 2 N	5/10 Z N A
		請求項の数 36 (全492頁) 最終頁に続く	
(21)出願番号	特願2021-546888(P2021-546888)	(73)特許権者	520439645
(86)(22)出願日	令和2年2月13日(2020.2.13)		ビーム セラピューティクス インク.
(65)公表番号	特表2022-520080(P2022-520080 A)		アメリカ合衆国 0 2 1 4 2 マサチュー
(43)公表日	令和4年3月28日(2022.3.28)		セッツ州 ケンブリッジ メイン ストリ
(86)国際出願番号	PCT/US2020/018073	(74)代理人	100083806
(87)国際公開番号	WO2020/168051		弁理士 三好 秀和
(87)国際公開日	令和2年8月20日(2020.8.20)	(74)代理人	100095500
審査請求日	令和5年2月13日(2023.2.13)		弁理士 伊藤 正和
(31)優先権主張番号	62/805,271	(74)代理人	100111235
(32)優先日	平成31年2月13日(2019.2.13)		弁理士 原 裕子
(33)優先権主張国・地域又は機関	米国(US)	(74)代理人	100195257
(31)優先権主張番号	62/852,228		弁理士 大淵 一志
(32)優先日	令和1年5月23日(2019.5.23)	(72)発明者	スレイメイカー、 イアン
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】 遺伝的疾患の治療用を含めアデノシンデアミナーゼ塩基エディターを用いて疾患関連遺伝子を編集する方法

(57)【特許請求の範囲】

【請求項1】

対象における神経障害の治療用の医薬組成物であって、(i) アデノシン塩基エディターまたはそれをコードする核酸配列と、(ii) ガイドポリヌクレオチドまたはそれをコードする核酸配列とを含み、

前記アデノシン塩基エディターは、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み、

前記アデノシンデアミナーゼドメインは、配列番号2における番号付けでアミノ酸位置82またはそれに対応する位置でのV82S改変を含み、

前記ガイドポリヌクレオチドは、前記アデノシン塩基エディターを誘導して前記神経障害に関連する標的遺伝子またはその調節エレメントにおいてAからGへの核酸塩基改変をもたらし、それによって前記神経障害を治療し、ここで、

i) 前記標的遺伝子がアルファ-L-イズロニダーゼ (IDUA) 遺伝子であり、前記神経障害がハーラー症候群であるか；または

i i) 前記標的遺伝子がロイシンリッチリピートキナーゼ2 (LRRK2) 遺伝子であり、前記神経障害がパーキンソン病であるか；または

i i i) 前記標的遺伝子がメチルCpG結合タンパク質2 (MECP2) 遺伝子であり、前記神経障害がレット症候群であるか；または

i v) 前記標的遺伝子がATP結合カセットサブファミリーメンバー4 (ABCA4) 遺伝子であり、前記神経障害がシュタルガルト病である、

医薬組成物。

【請求項 2】

(i) 前記IDUA遺伝子またはその調節エレメントが、ハーラー症候群に関連するSNPを含むか；または

(i i) 前記LRRK2遺伝子またはその調節エレメントがパーキンソン病に関連するSNPを含むか；または

(i i i) 前記MECP2遺伝子またはその調節エレメントが、レット症候群に関連するSNPを含むか；または

(i v) 前記ABCA4遺伝子またはその調節エレメントが、シュタルガルト病に関連するSNPを含む、

請求項 1 に記載の医薬組成物。

【請求項 3】

前記AからGへの核酸塩基改変が、ハーラー症候群、パーキンソン病、レット症候群、またはシュタルガルト病に関連するSNPにおけるものである、請求項 2 に記載の医薬組成物。

【請求項 4】

(i) 前記ハーラー症候群に関連するSNPが、配列番号4における番号付けでIDUAポリペプチドにおけるW402XもしくはW401Xアミノ酸変異体またはIDUA遺伝子によってコードされるそのバリエーションをもたらし、ここでXは終止コドンであるか；または

(i i) 前記パーキンソン病に関連するSNPが、配列番号3における番号付けでLRRK2ポリペプチドにおけるA419V、R1441C、R1441H、もしくはG2019Sアミノ酸変異体またはLRRK2遺伝子によってコードされるそのバリエーションをもたらし、または

(i i i) 前記レット症候群に関連するSNPが、配列番号5における番号付けでMECP2ポリペプチドにおけるR106WもしくはT158Mアミノ酸変異体またはMECP2遺伝子によってコードされるそのバリエーションをもたらし、または

(i v) 前記シュタルガルト病に関連するSNPが、配列番号6における番号付けでABCA4ポリペプチドにおけるA1038VもしくはG1961Eアミノ酸変異体またはABCA4遺伝子によってコードされるそのバリエーションをもたらし、

請求項 2 または 3 に記載の医薬組成物。

【請求項 5】

前記AからGへの核酸塩基改変が、ハーラー症候群、パーキンソン病、レット症候群、またはシュタルガルト病に関連するSNPを野生型核酸塩基に変化させる、請求項 2 に記載の医薬組成物。

【請求項 6】

前記AからGへの核酸塩基改変が、ハーラー症候群に関連するSNPを、ハーラー症候群の1つ以上の症状の改善をもたらす非野生型核酸塩基に変化させる、請求項 2 に記載の医薬組成物。

【請求項 7】

ハーラー症候群に関連するSNPにおけるAからGへの改変が、IDUA遺伝子によってコードされる終止コドンをIDUAポリペプチドにおけるトリプトファンに変化させる、請求項 2 に記載の医薬組成物。

【請求項 8】

前記ガイドポリヌクレオチドが、

ハーラー症候群のための5' - GACUCUAGGCAGAGGUCUCAA -3'、5' - ACUCUAGGCAGAGGUCUCAA-3'、5' - CUCUAGGCCGAAGUGUCGC -3'、および5' -GCUCUAGGCCGAAGUGUCGC-3' からなる群；

パーキンソン病のための5' -AAGCGCAAGCCUGGAGGGAA -3'、5' -ACUACAGCAUUGCUCAGUAC-3' からなる群；

レット症候群のための5' - CUUUUCACUCCUGCCGGGG-3'、5' -AGCUUCCAUGUCCAGCCUUC-3'、5' - ACCAUGAAGUCAAAAUCAUU-3'、および5' - GCUUUCAGCCCCGUUUCUUG-3' からなる群；および

10

20

30

40

50

シュタルガルト病のための5'-CUCCAGGGCGAACUUCGACACACAGC-3'より選択される核酸配列を含むsgRNAである、請求項1～7のいずれか一項に記載の医薬組成物。

【請求項9】

a) 前記AからGへの核酸塩基の改変が、LRRK2遺伝子によってコードされるLRRK2ポリペプチド中のシステインまたはヒスチジンをアルギニンに変化させるか；または

b) 前記AからGへの核酸塩基の改変が、LRRK2遺伝子によってコードされるLRRK2ポリペプチドにおいてセリンをグリシンに変化させるか；または

c) 前記AからGへの核酸塩基の改変が、配列番号3における番号付けでのLRRK2ポリペプチドまたはLRRK2遺伝子によってコードされるそのバリエーションの位置144でシステインまたはヒスチジンをアルギニンに置換するか、または位置2019でセリンをグリシンに置換する、

10

請求項2に記載の医薬組成物。

【請求項10】

a) 前記レット症候群に関連するSNPが、MECP2遺伝子によってコードされるMECP2ポリペプチドにおけるR255XまたはR270Xアミノ酸変異体をもたらす、ここでXは終止コドンであるか；または

b) レット症候群に関連するSNPにおける前記AからGへの核酸塩基改変が、終止コドンをMECP2ポリペプチドにおけるトリプトファンに変化させる、

請求項2に記載の医薬組成物。

20

【請求項11】

神経障害に関連する標的遺伝子またはその調節エレメント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、前記標的遺伝子またはその調節エレメントを、(i) アデノシン塩基エディターと、(ii) ガイドポリヌクレオチドとに接触させることを含む、

前記アデノシン塩基エディターは、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、

前記アデノシンデアミナーゼドメインは、配列番号2における番号付けでアミノ酸位置82またはそれに対応する位置でのV82Sアミノ酸置換を含む、

前記ガイドポリヌクレオチドは、前記アデノシン塩基エディターを誘導して前記神経障害に関連する標的遺伝子またはその調節エレメントにおいてAからGへの核酸塩基改変をもたらす、ここで、

30

i) 前記標的遺伝子がアルファ-L-イソロニダーゼ (IDUA) 遺伝子であり、前記神経障害がハーパー症候群であるか；または

ii) 前記標的遺伝子がロイシンリッチリピートキナーゼ2 (LRRK2) 遺伝子であり、前記神経障害がパーキンソン病であるか；または

iii) 前記標的遺伝子がメチルCpG結合タンパク質2 (MECP2) 遺伝子であり、前記神経障害がレット症候群であるか；または

iv) 前記標的遺伝子がATP結合カセットサブファミリーメンバー4 (ABCA4) 遺伝子であり、前記神経障害がシュタルガルト病である、

方法。

40

【請求項12】

a) 前記AからGへの核酸塩基改変が、前記神経障害に関連するSNPを野生型核酸塩基に変化させるか；または

b) 前記AからGへの核酸塩基改変が、前記神経障害に関連するSNPを、前記神経障害の1つ以上の症状の改善をもたらす非野生型核酸塩基に変化させる、

請求項11に記載の方法。

【請求項13】

パーキンソン病に関連するロイシンリッチリピートキナーゼ2 (LRRK2) 遺伝子またはその調節エレメント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LRRK2遺伝子またはその調節エレメントを、(i) アデノシン塩基エディターまたはそれをコードす

50

る核酸配列と、(ii) ガイドポリヌクレオチドまたはそれをコードする核酸配列と接触させることを含み、

前記アデノシン塩基エディターは、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み、

前記アデノシンデアミナーゼドメインは、配列番号2における番号付けでアミノ酸位置82またはそれに対応する位置でのV82S改変を含み、

前記ガイドポリヌクレオチドが前記アデノシン塩基エディターを誘導してLRRK2遺伝子のSNPにおいてAからGへの核酸塩基改変をもたらす、前記SNPは、配列番号3における番号付けでLRRK2ポリペプチドにおけるG2019S変異体またはそのバリエーションをコードしない、方法。

【請求項14】

前記接触させることが細胞内で起こる、請求項13に記載の方法。

【請求項15】

前記アデノシンデアミナーゼドメインが、配列番号2における番号付けでアミノ酸位置166またはそれに対応する位置におけるアミノ酸置換をさらに含み；および／または

前記ガイドポリヌクレオチドは核酸配列5'-AAGCGCAAGCCUGGAGGGAA-3'または5'-ACUACAGCAUUGCUCAGUAC-3'を含むsgRNAである、

請求項13または14に記載の方法。

【請求項16】

(i) 前記SNPが、配列番号3における番号付けでLRRK2ポリペプチドにおけるA419V、R1441C、R1441H、もしくはG2019Sアミノ酸変異体またはLRRK2遺伝子によってコードされるそのバリエーションをもたらすか；または

(ii) 前記AからGへの核酸塩基の改変が、LRRK2遺伝子によってコードされるLRRK2ポリペプチド中のシステインまたはヒスチジンをアルギニンに変化させるか；または

(iii) 前記AからGへの核酸塩基の改変が、LRRK2遺伝子によってコードされるLRRK2ポリペプチドにおいてセリンをグリシンに変化させるか；または

(iv) 前記AからGへの核酸塩基の改変が、配列番号3における番号付けでのLRRK2ポリペプチドまたはLRRK2遺伝子によってコードされるそのバリエーションの位置144のシステインまたはヒスチジンをアルギニンに置換するか、または位置2019のセリンをグリシンに置換する、

請求項13～15のいずれか一項に記載の方法。

【請求項17】

ハーラー症候群に関連するアルファ-L-イズロニダーゼ (IDUA) 遺伝子またはその調節エレメント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、IDUA遺伝子またはその調節エレメントを、(i) アデノシン塩基エディターまたはそれをコードする核酸配列と、(ii) ガイドポリヌクレオチドまたはそれをコードする核酸配列と接触させることを含み、

前記アデノシン塩基エディターは、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み、

前記アデノシンデアミナーゼドメインは、配列番号2における番号付けでアミノ酸位置82またはそれに対応する位置でのV82Sアミノ酸置換を含み、

前記ガイドポリヌクレオチドは、前記アデノシン塩基エディターを誘導して前記IDUA遺伝子またはその調節エレメントにおいてAからGへの核酸塩基改変をもたらす、方法。

【請求項18】

前記接触させることが細胞内で起こる、請求項17に記載の方法。

【請求項19】

(i) 前記IDUA遺伝子またはその調節エレメントが、ハーラー症候群に関連するSNPを含むか；または

(ii) 前記AからGへの核酸塩基改変が、ハーラー症候群に関連するSNPにおけるものであるか；または

10

20

30

40

50

(i i i) 前記IDUA遺伝子またはその調節エレメントが、配列番号4における番号付けでIDUAポリペプチドにおけるW402XもしくはW401Xアミノ酸変異体またはIDUA遺伝子によってコードされるそのバリエーションをもたらすハーラー症候群に関連するSNPを含み、ここでXは終止コドンであるか；または

(i v) 前記IDUA遺伝子またはその調節エレメントが、ハーラー症候群に関連するSNPを含み、前記AからGへの核酸塩基改変が、ハーラー症候群に関連するSNPを野生型核酸塩基に変化させるか；または

(v) 前記IDUA遺伝子またはその調節エレメントが、ハーラー症候群に関連するSNPを含み、前記AからGへの核酸塩基改変が、ハーラー症候群に関連するSNPを、ハーラー症候群の1つ以上の症状の改善をもたらす非野生型核酸塩基に変化させるか；または

(v i) 前記IDUA遺伝子またはその調節エレメントが、ハーラー症候群に関連するSNPを含み、ハーラー症候群に関連するSNPにおける前記AからGへの改変が、IDUA遺伝子によってコードされる終止コドンをIDUAポリペプチドにおけるトリプトファンに変化させるか；または

(v i i) 前記IDUA遺伝子またはその調節エレメントが、ハーラー症候群に関連するSNPを含み、前記ガイドポリヌクレオチドが、ハーラー症候群に関連するSNPを含むIDUA遺伝子またはその調節エレメントに相補的な核酸配列を含むか；または

(v i i i) 前記IDUA遺伝子またはその調節エレメントが、ハーラー症候群に関連するSNPを含み、前記アデノシン塩基エディターが、ハーラー症候群に関連するSNPを含むIDUA遺伝子またはその調節エレメントに相補的な核酸配列を含む単一ガイドRNAと複合体をなすか；または

(i x) 前記ガイドポリヌクレオチドが、5' - GACUCUAGGCAGAGGUCUCAA -3'、5' - ACUCUAGGCAGAGGUCUCAA-3'、5' - CUCUAGGCCGAAGUGUCGC -3'、および5' - GCUCUAGGCCGAAGUGUCGC-3' からなる群より選択される核酸配列を含むsgRNAである、

請求項17または18に記載の方法。

【請求項20】

レット症候群に関連するメチルCpG結合タンパク質2 (MECP2) 遺伝子またはその調節エレメント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、(i) アデノシン塩基エディターまたはそれをコードする核酸配列と、(ii) ガイドポリヌクレオチドまたはそれをコードする核酸配列を、MECP2遺伝子またはその調節エレメントに接触させることを含み、

前記アデノシン塩基エディターは、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み、

前記アデノシンデアミナーゼドメインは、配列番号2における番号付けでアミノ酸位置82またはそれに対応する位置でのV82S改変を含み、

前記ガイドポリヌクレオチドは、前記アデノシン塩基エディターを誘導してMECP2遺伝子またはその調節エレメントにおけるAからGへの核酸塩基改変をもたらす、

方法。

【請求項21】

前記接触させることが細胞内で起こる、請求項20に記載の方法。

【請求項22】

(i) 前記MECP2遺伝子またはその調節エレメントが、レット症候群に関連するSNPを含み；および／または

(i i) 前記MECP2遺伝子またはその調節エレメントが、レット症候群に関連するSNPを含み、前記AからGへの核酸塩基改変が、レット症候群に関連するSNPにおけるものであり；および／または

(i i i) 前記MECP2遺伝子またはその調節エレメントが、レット症候群に関連するSNPを含み、前記レット症候群と関連するSNPが、配列番号5における番号付けでMECP2ポリペプチドにおけるR106WもしくはT158Mアミノ酸変異体またはMECP2遺伝子によ

10

20

30

40

50

ってコードされるそのバリエーションをもたらし；またはMECP2遺伝子によってコードされるMECP2ポリペプチドにおけるR255XまたはR270Xアミノ酸変異体をもたらし、ここでXは終止コドンであり；および／または

(i v) 前記MECP2遺伝子またはその調節エレメントが、レット症候群に関連するSNPを含み、前記AからGへの核酸塩基改変が、レット症候群に関連するSNPを野生型核酸塩基に変化させ；および／または

(v) 前記MECP2遺伝子またはその調節エレメントが、レット症候群に関連するSNPを含み、前記AからGへの核酸塩基改変が、レット症候群に関連するSNPを、レット症候群の1つ以上の症状の改善をもたらす非野生型核酸塩基に変化させ；および／または

(v i) 前記MECP2遺伝子またはその調節エレメントが、レット症候群に関連するSNPを含み、レット症候群に関連するSNPにおける前記AからGへの核酸塩基改変が、MECP2ポリペプチドにおいて終止コドンをトリプトファンに変化させ；および／または

(v i i) 前記MECP2遺伝子またはその調節エレメントが、レット症候群に関連するSNPを含み、前記ガイドポリヌクレオチドが、レット症候群に関連するSNPを含むMECP2遺伝子またはその調節エレメントに相補的な核酸配列を含み；および／または

(v i i i) 前記MECP2遺伝子またはその調節エレメントが、レット症候群に関連するSNPを含み、前記アデノシン塩基エディターが、レット症候群と関連するSNPを含むMECP2遺伝子またはその調節エレメントに相補的な核酸配列を含む単一ガイドRNAと複合体をなし；および／または

(i x) 前記ガイドポリヌクレオチドが、5' - CUUUUCACUCCUGCCGGGG-3'、5' - AGCUUCCAUGUCCAGCCUUC-3'、5' - ACCAUGAAGUCAAAAUCAUU-3'、および5' - GUUUCAGCCCCGUUUCUUG-3' からなる群より選択される核酸配列を含む、

請求項20または21に記載の方法。

【請求項23】

シュタルガルト病に関連するATP結合カセットサブファミリーメンバー4 (ABCA4) 遺伝子またはその調節エレメント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ABCA4遺伝子またはその調節エレメントを、(i) アデノシン塩基エディターまたはそれをコードする核酸配列と、(ii) ガイドポリヌクレオチドまたはそれをコードする核酸配列と接触させることを含み、

前記アデノシン塩基エディターは、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み、

前記アデノシンデアミナーゼドメインは、配列番号2における番号付けでアミノ酸位置82またはそれに対応する位置でのV82Sアミノ酸置換を含み、

前記ガイドポリヌクレオチドは、前記アデノシン塩基エディターを誘導してABCA4遺伝子またはその調節エレメントにおけるAからGへの核酸塩基改変をもたらす、

方法。

【請求項24】

前記接触させることが細胞内で起こる、請求項23に記載の方法。

【請求項25】

(i) ABCA4遺伝子が、シュタルガルト病に関連するSNPを含み；および／または

(i i) ABCA4遺伝子が、シュタルガルト病に関連するSNPを含み、前記AからGへの核酸塩基改変が、シュタルガルト病に関連するSNPにおけるものであり；および／または

(i i i) ABCA4遺伝子が、シュタルガルト病に関連するSNPを含み、シュタルガルト病に関連する前記SNPが、配列番号6における番号付けでABCA4ポリペプチドにおけるA1038VもしくはG1961Eアミノ酸変異体またはABCA4遺伝子によってコードされるそのバリエーションをもたらし；および／または

(i v) ABCA4遺伝子が、シュタルガルト病に関連するSNPを含み、前記AからGへの核酸塩基改変が、シュタルガルト病に関連するSNPを野生型核酸塩基に変化させ；および／または

(v) ABCA4遺伝子が、シュタルガルト病に関連するSNPを含み、前記AからGへの核

10

20

30

40

50

酸塩基の改変が、シュタルガルト病に関連するSNPを、シュタルガルト病の1つ以上の症状の改善をもたらす非野生型核酸塩基に変化させ；および／または

(v i) ABCA4遺伝子が、シュタルガルト病に関連するSNPを含み、前記ガイドポリヌクレオチドが、シュタルガルト病に関連するSNPを含むABCA4遺伝子またはその調節エレメントに相補的な核酸配列を含み；および／または

(v i i) ABCA4遺伝子が、シュタルガルト病に関連するSNPを含み、前記アデノシン塩基エディターが、シュタルガルト病に関連するSNPを含むABCA4遺伝子またはその調節エレメントに相補的な核酸配列を含む単一ガイドRNAと複合体をなし；および／または

(v i i i) 前記ガイドポリヌクレオチドが、配列5'-CUCCAGGGCGAACUUCGACACACAGC-3'を含むsgRNAである、

請求項 2 3 または 2 4 に記載の方法。

【請求項 2 6】

請求項 1 1 ～ 2 5 のいずれか一項に記載の方法によって産生される細胞。

【請求項 2 7】

(i) アデノシン塩基エディターまたはそれをコードする核酸配列と、(ii) ガイドポリヌクレオチドまたはそれをコードする核酸配列とを含む、塩基エディターシステムであって、

前記アデノシン塩基エディターが、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み、

前記アデノシンデアミナーゼドメインは、配列番号2における番号付けでアミノ酸位置82またはそれに対応する位置でのV82Sアミノ酸置換を含み、

前記ガイドポリヌクレオチドは、前記アデノシン塩基エディターを誘導して、LRRK2、IDUA、MECP2、およびABCA4からなる群から選択される神経障害に関連する標的遺伝子またはその調節エレメントにおけるAからGへの核酸塩基改変をもたらし、ここで、

i) 前記標的遺伝子がアルファ-L-イズロニダーゼ (IDUA) 遺伝子であり、前記神経障害がハーラー症候群であるか；または

i i) 前記標的遺伝子がロイシンリッチリピートキナーゼ2 (LRRK2) 遺伝子であり、前記神経障害がパーキンソン病であるか；または

i i i) 前記標的遺伝子がメチルCpG結合タンパク質2 (MECP2) 遺伝子であり、前記神経障害がレット症候群であるか；または

i v) 前記標的遺伝子がATP結合カセットサブファミリーメンバー4 (ABCA4) 遺伝子であり、前記神経障害がシュタルガルト病である、

塩基エディターシステム。

【請求項 2 8】

前記ポリヌクレオチドによりプログラム可能なDNA結合ドメインがCas9である、請求項 2 7 に記載の塩基エディターシステム。

【請求項 2 9】

前記Cas9がSpCas9またはSaCas9である、請求項 2 8 に記載の塩基エディターシステム。

【請求項 3 0】

前記ポリヌクレオチドによりプログラム可能なDNA結合ドメインが、NGG、NGA、NGCG、NGN、NNGRRT、NNNRRT、NGCG、NGCN、NGTN、およびNGCからなる群より選択されるヌクレオチド配列に対するプロトスペーサー隣接モチーフ特異性を有するSpCas9を含み、ここでNはA、G、CまたはTであり、RはAまたはGである、請求項 2 7 ～ 2 9 のいずれか一項に記載の塩基エディターシステム。

【請求項 3 1】

前記ポリヌクレオチドによりプログラム可能なDNA結合ドメインが、

(i) ヌクレアーゼ不活性であるか；または

(i i) ニックアーゼである、

請求項 2 7 ～ 3 0 のいずれか一項に記載の塩基エディターシステム。

【請求項 3 2】

10

20

30

40

50

(i) 前記アデノシンデアミナーゼが、T166R改変をさらに含むか；または
(ii) 前記アデノシンデアミナーゼが、Y147T、Y147R、Q154S、Y123H、Q154R、またはそれらの組み合わせのうちの1つ以上の改変をさらに含むか；または
(iii) 前記アデノシンデアミナーゼが、Y147R + Q154R + Y123H；Y147R + Q154R + I76Y；Y147R + Q154R + T166R；Y147T + Q154R；Y147T + Q154S；およびY123H + Y147R + Q154R + I76Yからなる群から選択される改変の組み合わせをさらに含むか；または
(iv) 前記アデノシン塩基エディタードメインがアデノシンデアミナーゼ単量体を含むか；または
(v) 前記アデノシン塩基エディターがアデノシンデアミナーゼ二量体を含む、

10

請求項 27～31 のいずれか一項記載の塩基エディターシステム。

【請求項 33】

請求項 27～32 のいずれか一項に記載の塩基エディターシステムを含む、インビトロまたはエクスピボの細胞。

【請求項 34】

請求項 27～32 のいずれか一項記載の塩基エディターシステム、または請求項 33 に記載の細胞と、薬学的に許容される担体とを含む、医薬組成物。

【請求項 35】

脂質をさらに含む、請求項 34 記載の医薬組成物。

【請求項 36】

請求項 27～32 のいずれか一項記載の塩基エディターシステム、または請求項 33 に記載の細胞を含む、キット。

20

【発明の詳細な説明】

【技術分野】

【0001】

関連出願への相互参照

この出願は、2019年2月13日に提出された米国仮出願第62 / 805,271号、2019年5月23日に提出された米国仮出願第62 / 852,228号、2019年5月23日に提出された米国仮出願第62 / 852,224号、2019年7月11日に提出された米国仮出願第62 / 873,138号、2019年8月19日に提出された米国仮出願第62 / 888,867号、2019年11月6日に提出された米国仮出願第62 / 931,722号、2019年11月27日に提出された米国仮出願第62 / 941,569号、2020年1月27日に提出された米国仮出願第62 / 966,526号の利益を主張し、それらの開示は参照によりその全体が本明細書に組み込まれる。

30

【0002】

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、個々の刊行物、特許、または特許出願が参照により組み込まれることが具体的かつ個別に示された場合と同程度に、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。

40

【背景技術】

【0003】

核酸配列の標的化された編集、例えば、ゲノムDNAの標的化された切断または標的化された改変は、遺伝子機能の研究にとって非常に有望なアプローチであり、ヒトの遺伝病に新しい治療法を提供する可能性も有する。現在利用可能な塩基エディターには、標的C・G塩基対をT・Aに変換するシチジン塩基エディター（BE4など）と、A・TをG・Cに変換するアデニン塩基エディター（ABE7.10など）が含まれる。より高い特異性および効率で標的配列内に改変を誘導することができる改良された塩基エディターが当技術分野で必要とされている。

【発明の概要】

50

【0004】

本発明は、効率が増加された新規のアデニン塩基エディター（例えばABE8）を含む組成物、および、標的配列を編集するためにアデノシンデアミナーゼバリエントを含む塩基エディターを使用する方法を提供する。

【0005】

ある態様において、本明細書では、対象における神経障害を治療する方法であって、(i) アデノシン塩基エディターまたはそれをコードする核酸配列と、(ii) ガイドポリヌクレオチドまたはそれをコードする核酸配列とを対象に投与することを含み、前記アデノシン塩基エディターは、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み、前記アデノシンデアミナーゼドメインは、配列番号2における番号付けでアミノ酸位置82もしくは166またはそれらに対応する位置でのアミノ酸置換を含み、前記ガイドポリヌクレオチドは、前記アデノシン塩基エディターを誘導して対象の前記神経障害に関連する標的遺伝子またはその調節エレメントにおいてAからGへの核酸塩基改変をもたらす、それによって対象の前記神経障害を治療する、方法が提供される。この態様のもう1つの実施形態では、前記標的遺伝子がアルファ-L-イズロニダーゼ（IDUA）遺伝子であり、神経疾患がハーラー症候群である。この態様の一実施形態では、前記標的遺伝子がロイシンリッチリピートキナーゼ2（LRRK2）遺伝子であり、神経疾患がパーキンソン病である。この態様の一実施形態では、前記標的遺伝子がメチルCpG結合タンパク質2（MECP2）遺伝子であり、神経疾患がレット症候群である。この態様の別の一実施形態では、前記標的遺伝子がATP結合カセットサブファミリーメンバー4（ABCA4）遺伝子であり、神経疾患がシュタルガルト病である。

【0006】

ある態様において、本明細書では、対象におけるハーラー症候群を治療する方法であって、(i) アデノシン塩基エディターまたはそれをコードする核酸配列と、(ii) ガイドポリヌクレオチドまたはそれをコードする核酸配列とを対象に投与することを含み、前記アデノシン塩基エディターは、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み、前記アデノシンデアミナーゼドメインは、配列番号2における番号付けでアミノ酸位置82もしくは166またはそれらに対応する位置でのアミノ酸置換を含み、前記ガイドポリヌクレオチドは、前記アデノシン塩基エディターを誘導して対象のアルファ-L-イズロニダーゼ（IDUA）遺伝子またはその調節エレメントにおいてAからGへの核酸塩基改変をもたらす、それによって対象のハーラー症候群を治療する、方法が提供される。

【0007】

ある実施形態では、前記投与が、ハーラー症候群に関連する少なくとも1つの症状を改善する。ある実施形態では、前記投与が、アデノシンデアミナーゼにおける前記アミノ酸置換を伴わない塩基エディターを用いた治療と比較して、ハーラー症候群に関連する少なくとも1つの症状のより迅速な改善をもたらす。

【0008】

ある実施形態では、前記IDUA遺伝子またはその調節エレメントが、ハーラー症候群に関連するSNPを含む。ある実施形態では、前記AからGへの核酸塩基改変は、ハーラー症候群に関連するSNPにおけるものである。ある実施形態では、ハーラー症候群に関連するSNPが、配列番号4における番号付けでIDUAポリペプチドにおけるW402XもしくはW401Xアミノ酸変異体またはIDUA遺伝子によってコードされるそのバリエーションをもたらす、ここでXは終止コドンである。ある実施形態では、前記AからGへの核酸塩基改変が、ハーラー症候群に関連するSNPを野生型核酸塩基に変化させる。ある実施形態では、前記AからGへの核酸塩基改変が、ハーラー症候群に関連するSNPを、ハーラー症候群の1つ以上の症状の改善をもたらす非野生型核酸塩基に変化させる。ある実施形態では、ハーラー症候群に関連するSNPにおけるAからGへの改変が、IDUA遺伝子によってコードされる終止コドンをIDUAポリペプチドにおけるトリプトファンに変化させる。

【0009】

ある実施形態では、前記ガイドポリヌクレオチドが、ハーラー症候群に関連するSNPを

含むIDUA遺伝子またはその調節エレメントに相補的な核酸配列を含む。ある実施形態では、前記アデノシン塩基エディターが、ハーラー症候群に関連するSNPを含むIDUA遺伝子またはその調節エレメントに相補的な核酸配列を含む単一ガイドRNA (sgRNA) と複合体をなす。ある実施形態では、sgRNAが、5' - GACUCUAGGCAGAGGUCUCA -3'、5' - ACUCUAGGCAGAGGUCUCA -3'、5' - CUCUAGGCCGAAGUGUCGC -3'、および 5' - GCUCUAGGCCGAAGUGUCGC -3' からなる群より選択される核酸配列を含む。

【0010】

ある態様において、本明細書では、対象におけるパーキンソン病を治療する方法であって、(i) アデノシン塩基エディターまたはそれをコードする核酸配列と、(ii) ガイドポリヌクレオチドまたはそれをコードする核酸配列とを対象に投与することを含み、前記アデノシン塩基エディターは、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み、前記アデノシンデアミナーゼドメインは、配列番号2における番号付けでアミノ酸位置82もしくは166またはそれらに対応する位置でのアミノ酸置換を含み、前記ガイドポリヌクレオチドは、前記アデノシン塩基エディターを誘導して対象のロイシンリッチリピートキナーゼ2 (LRRK2) 遺伝子またはその調節エレメントにおいてAからGへの核酸塩基改変をもたらす、それによって対象のパーキンソン病を治療する、方法が提供される。

【0011】

ある実施形態では、前記投与がパーキンソン病に関連する少なくとも一つの症状を改善する。ある実施形態では、前記投与が、アデノシンデアミナーゼにおける前記アミノ酸置換を伴わない塩基エディターを用いた治療と比較して、パーキンソン病に関連する少なくとも一つの症状のより迅速な改善をもたらす。

【0012】

ある実施形態では、LRRK2遺伝子またはその調節エレメントがパーキンソン病に関連するSNPを含む。ある実施形態では、前記AからGへの核酸塩基改変は、パーキンソン病に関連するSNPにおけるものである。ある実施形態では、パーキンソン病に関連するSNPが、配列番号3における番号付けでLRRK2ポリペプチドにおけるA419V、R1441C、R1441H、もしくはG2019Sアミノ酸変異体またはLRRK2遺伝子によってコードされるそのバリエーションをもたらす。

【0013】

ある実施形態では、前記AからGへの核酸塩基の改変が、パーキンソン病に関連するSNPを野生型核酸塩基に変化させる。ある実施形態では、前記AからGへの核酸塩基の改変が、パーキンソン病に関連するSNPを、パーキンソン病の1つ以上の症状の改善をもたらす非野生型核酸塩基に変化させる。ある実施形態では、前記AからGへの核酸塩基の改変が、LRRK2遺伝子によってコードされるLRRK2ポリペプチド中のシステインまたはヒスチジンをアルギニンに変化させる。ある実施形態では、前記AからGへの改変が、LRRK2遺伝子によってコードされるLRRK2ポリペプチドにおいてセリンをグリシンに変化させる。ある実施形態では、前記AからGへの改変が、配列番号3における番号付けでLRRK2ポリペプチドまたはLRRK2遺伝子によってコードされるそのバリエーションの位置144でシステイン (C) またはヒスチジン (H) をアルギニン (R) に置換するか、または位置2019でセリンをグリシン (G) に置換する。

【0014】

ある態様において、本明細書では、対象におけるパーキンソン病を治療する方法であって、(i) アデノシン塩基エディターまたはそれをコードする核酸配列と、(ii) ガイドポリヌクレオチドまたはそれをコードする核酸配列とを対象に投与することを含み、前記アデノシン塩基エディターは、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み、前記ガイドポリヌクレオチドは、前記アデノシン塩基エディターを誘導してパーキンソン病に関連するLRRK2遺伝子におけるSNPにおいてAからGへの核酸塩基改変をもたらす、前記SNPは、配列番号3における番号付けでLRRK2ポリペプチドにおけるG2019S変異体またはそのバリエーションをコードしない、方法が提供される。

【0015】

ある実施形態では、前記アデノシンデアミナーゼドメインが、配列番号2における番号付けでアミノ酸位置82もしくは166またはそれに対応する位置におけるアミノ酸置換を含む。ある実施形態では、前記ガイドポリヌクレオチドが、パーキンソン病に関連するSNPを含むLRRK2遺伝子またはその調節エレメントに相補的な核酸配列を含む。ある実施形態では、前記アデノシン塩基エディターが、パーキンソン病に関連するSNPを含むLRRK2遺伝子またはその調節エレメントに相補的な核酸配列を含む単一ガイドRNA (sgRNA) と複合体をなす。ある実施形態では、前記sgRNAが核酸配列5' -AAGCGCAAGCCUGGAGGGAA-3' ; または 5' -ACUACAGCAUUGCUCAGUAC-3' を含む。

【0016】

ある態様において、本明細書では、対象におけるレット症候群を治療する方法であって、(i) アデノシン塩基エディターまたはそれをコードする核酸配列と、(ii) ガイドポリヌクレオチドまたはそれをコードする核酸配列とを対象に投与することを含み、前記アデノシン塩基エディターは、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み、前記アデノシンデアミナーゼドメインは、配列番号2における番号付けでアミノ酸位置82もしくは166またはそれらに対応する位置でのアミノ酸置換を含み、前記ガイドポリヌクレオチドは、前記アデノシン塩基エディターを誘導して対象のメチルCpG結合タンパク質2 (MECP2) 遺伝子またはその調節エレメントにおいてAからGへの核酸塩基改変をもたらす、それによって対象のレット症候群を治療する、方法が提供される。

【0017】

ある実施形態では、前記投与が、レット症候群に関連する少なくとも一つの症状を改善する。ある実施形態では、前記投与が、アデノシンデアミナーゼにおける前記アミノ酸置換を伴わない塩基エディターを用いた治療と比較して、レット症候群に関連する少なくとも一つの症状のより迅速な改善をもたらす。ある実施形態では、前記MECP2遺伝子またはその調節エレメントが、レット症候群に関連するSNPを含む。ある実施形態では、前記AからGへの核酸塩基改変は、レット症候群に関連するSNPにおけるものである。ある実施形態では、前記レット症候群に関連するSNPが、配列番号5における番号付けでMECP2ポリペプチドにおけるR106WもしくはT158Mアミノ酸変異体またはMECP2遺伝子によってコードされるそのバリエーションをもたらす。ある実施形態では、前記レット症候群に関連するSNPが、MECP2遺伝子によってコードされるMECP2ポリペプチドにおけるR255XまたはR270Xアミノ酸変異体をもたらす、ここでXは終止コドンである。

【0018】

ある実施形態では、前記AからGへの核酸塩基改変が、レット症候群に関連するSNPを野生型核酸塩基に変化させる。ある実施形態では、前記AからGへの核酸塩基改変が、レット症候群に関連するSNPを、レット症候群症状の改善をもたらす非野生型核酸塩基に変化させる。ある実施形態では、レット症候群に関連するSNPにおける前記AからGへの核酸塩基改変が、終止コドンをMECP2ポリペプチドにおけるトリプトファンに変化させる。

【0019】

ある実施形態では、前記ガイドポリヌクレオチドが、レット症候群と関連するSNPを含むMECP2遺伝子またはその調節エレメントに対して相補的な核酸配列を含む。ある実施形態では、前記アデノシン塩基エディターが、レット症候群と関連するSNPを含むMECP2遺伝子またはその調節エレメントに相補的な核酸配列を含む単一ガイドRNA (sgRNA) と複合体をなす。ある実施形態では、前記ガイドポリヌクレオチドが、5' - CUUUUCACUUC CUGCCGGGG-3'、5' -AGCUUCCAUGUCCAGCCUUC-3'、5' - ACCAUGAAGUCAAAAUCAUU-3'、および5' - GCUUUCAGCCCCGUUUCUUG-3' からなる群より選択される核酸配列を含む。

【0020】

ある態様において、本明細書は、対象におけるシュタルガルト病を治療する方法であって、(i) アデノシン塩基エディターまたはそれをコードする核酸配列と、(ii) ガイドポリヌクレオチドまたはそれをコードする核酸配列とを対象に投与することを含み、前記アデ

10

20

30

40

50

ノシン塩基エディターは、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み、前記アデノシンデアミナーゼドメインは、配列番号2における番号付けでアミノ酸位置82もしくは166またはそれらに対応する位置でのアミノ酸置換を含み、前記ガイドポリヌクレオチドは、前記アデノシン塩基エディターを誘導して対象のATP結合カセットサブファミリーメンバー4 (ABCA4) 遺伝子またはその調節エレメントにおいてAからGへの核酸塩基改変をもたらす、それによって対象のシュタルガルト病を治療する、方法を提供する。

【0021】

ある実施形態では、前記投与が、シュタルガルト病に関連する少なくとも一つの症状を改善する。ある実施形態では、前記投与が、アデノシンデアミナーゼにおける前記アミノ酸置換を伴わない塩基エディターを用いた治療と比較して、シュタルガルト病に関連する少なくとも一つの症状のより迅速な改善をもたらす。

10

【0022】

ある実施形態では、ABCA4遺伝子が、シュタルガルト病に関連するSNPを含む。ある実施形態では、前記AからGへの核酸塩基改変は、シュタルガルト病に関連するSNPにおけるものである。ある実施形態では、シュタルガルト病に関連する前記SNPが、配列番号6における番号付けでABCA4ポリペプチドにおけるA1038VもしくはG1961Eアミノ酸変異体またはABCA4遺伝子によってコードされるそのバリエーションをもたらす。ある実施形態では、シュタルガルト病に関連する前記SNPが、配列番号6における番号付けでABCA4ポリペプチドにおけるG1961Eアミノ酸変異体またはそのバリエーションをもたらす。

20

【0023】

ある実施形態では、前記AからGへの核酸塩基改変が、シュタルガルト病に関連するSNPを野生型核酸塩基に変化させる。ある実施形態では、前記AからGへの核酸塩基改変が、シュタルガルト病に関連するSNPを、シュタルガルト病の1つ以上の症状の改善をもたらす非野生型核酸塩基に変化させる。ある実施形態では、前記ガイドポリヌクレオチドが、シュタルガルト病に関連するSNPを含むABCA4遺伝子またはその調節エレメントに相補的な核酸配列を含む。

【0024】

ある実施形態では、前記アデノシン塩基エディターは、シュタルガルト病に関連するSNPを含むABCA4遺伝子またはその調節エレメントに相補的な核酸配列を含む単一ガイドRNA (sgRNA) と複合体をなす。ある実施形態では、前記sgRNAが配列5' -CUCCAGGGCGAACUUCGACACACAGC-3' を含む。

30

【0025】

様々な態様において、本明細書に記述される治療は、前記アミノ酸置換を伴わない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用いた治療と比較して、神経障害の症状の改善をもたらす。

【0026】

ある態様において、本明細書では、神経障害に関連する標的遺伝子またはその調節エレメントを編集する方法であって、前記標的遺伝子またはその調節エレメントを、(i) アデノシン塩基エディターと、(ii) ガイドポリヌクレオチドとに接触させることを含み、前記アデノシン塩基エディターは、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み、前記アデノシンデアミナーゼドメインは、配列番号2における番号付けでアミノ酸位置82もしくは166またはそれらに対応する位置でのアミノ酸置換を含み、前記ガイドポリヌクレオチドは、前記アデノシン塩基エディターを誘導して前記神経障害に関連する標的遺伝子またはその調節エレメントにおいてAからGへの核酸塩基改変をもたらす、方法が提供される。この態様の一実施形態において、標的遺伝子はロイシンリッチリピートキナーゼ2 (LRRK2) 遺伝子であり、神経疾患がパーキンソン病である。この態様の別の一実施形態では、前記標的遺伝子がアルファ-L-イズロニダーゼ (IDUA) 遺伝子であり、神経疾患がハーラー症候群である。この態様の一実施形態では、前記標的遺伝子がメチルCpG結合タンパク質2 (MECP2) 遺伝子であり、神経疾患がレット症候群であ

40

50

る。この態様の別の一実施形態では、前記標的遺伝子がATP結合カセットサブファミリーメンバー4 (ABCA4) 遺伝子であり、神経疾患がシュタルガルト病である。

【0027】

ある態様において、本明細書では、ロイシンリッチリピートキナーゼ2 (LRRK2) 遺伝子またはその調節エレメントを編集する方法であって、LRRK2遺伝子またはその調節エレメントを、(i) アデノシン塩基エディターまたはそれをコードする核酸配列と、(ii) ガイドポリヌクレオチドまたはそれをコードする核酸配列とに接触させることを含み、前記アデノシン塩基エディターは、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み、前記アデノシンデアミナーゼドメインは、配列番号2における番号付けでアミノ酸位置82もしくは166またはそれらに対応する位置でのアミノ酸置換を含み、前記ガイドポリヌクレオチドは、前記アデノシン塩基エディターを誘導してLRRK2遺伝子またはその調節エレメントにおいてAからGへの核酸塩基改変をもたらす、方法が提供される。

10

【0028】

ある実施形態では、前記AからGへの核酸塩基改変は、パーキンソン病に関連するSNPにおけるものである。ある実施形態では、前記パーキンソン病に関連するSNPが、配列番号3における番号付けでLRRK2ポリペプチドにおけるA419V、R1441C、R1441H、もしくはG2019Sアミノ酸変異体またはLRRK2遺伝子によってコードされるそのバリエーションをもたらす。ある実施形態では、前記AからGへの核酸塩基改変が、パーキンソン病に関連するSNPを野生型核酸塩基に変化させる。ある実施形態では、前記AからGへの核酸塩基改変が、パーキンソン病に関連するSNPを、パーキンソン病の1つ以上の症状の改善をもたらす非野生型核酸塩基に変化させる。

20

【0029】

ある実施形態では、前記AからGへの核酸塩基の改変が、LRRK2遺伝子によってコードされるLRRK2ポリペプチド中のシステインまたはヒスチジンをアルギニンに変化させる。ある実施形態では、前記AからGへの改変が、LRRK2遺伝子によってコードされるLRRK2ポリペプチドにおいてセリンをグリシンに変化させる。ある実施形態では、前記AからGへの改変が、配列番号3における番号付けでLRRK2ポリペプチドまたはLRRK2遺伝子によってコードされるそのバリエーションの位置144のシステイン (C) またはヒスチジン (H) をアルギニン (R) に置換するか、または位置2019のセリンをグリシン (G) に置換する。

30

【0030】

ある態様において、本明細書では、ロイシンリッチリピートキナーゼ2 (LRRK2) 遺伝子またはその調節エレメントを編集する方法であって、前記LRRK2遺伝子またはその調節エレメントを、(i) アデノシン塩基エディターまたはそれをコードする核酸配列と、(ii) ガイドポリヌクレオチドまたはそれをコードする核酸配列とに接触させることを含み、前記アデノシン塩基エディターは、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み、前記ガイドポリヌクレオチドが前記アデノシン塩基エディターを誘導してLRRK2遺伝子のSNPにおいてAからGへの核酸塩基改変をもたらす、前記SNPは、配列番号3における番号付けでLRRK2ポリペプチドにおけるG2019S変異体またはそのバリエーションをコードしない、方法が提供される。

40

【0031】

ある実施形態では、前記アデノシンデアミナーゼドメインが、配列番号2における番号付けでアミノ酸位置82もしくは166またはそれに対応する位置におけるアミノ酸置換を含む。ある実施形態では、前記ガイドポリヌクレオチドが、パーキンソン病に関連するSNPを含むLRRK2遺伝子またはその調節エレメントに相補的な核酸配列を含む。

【0032】

ある実施形態では、前記アデノシン塩基エディターが、パーキンソン病に関連するSNPを含むLRRK2遺伝子またはその調節エレメントに相補的な核酸配列を含む単一ガイドRNA (sgRNA) と複合体をなす。ある実施形態では、前記sgRNAが核酸配列5' -AAGCGCAAGCCUGGAGGGAA -3' ; または5' -ACUACAGCAUUGCUCAGUAC -3' を含む。

50

【0033】

ある態様において、本明細書では、アルファ-L-イズロニダーゼ (IDUA) 遺伝子またはその調節エレメントを編集する方法であって、IDUA遺伝子またはその調節エレメントを、(i) アデノシン塩基エディターまたはそれをコードする核酸配列と、(ii) ガイドポリヌクレオチドまたはそれをコードする核酸配列と接触させることを含み、前記アデノシン塩基エディターは、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み、前記アデノシンデアミナーゼドメインは、配列番号2における番号付けでアミノ酸位置82もしくは166またはそれらに対応する位置でのアミノ酸置換を含み、前記ガイドポリヌクレオチドは、前記アデノシン塩基エディターを誘導して前記IDUA遺伝子またはその調節エレメントにおいてAからGへの核酸塩基改変をもたらす、方法が提供される。

10

【0034】

ある実施形態では、前記IDUA遺伝子またはその調節エレメントが、ハーラー症候群に関連するSNPを含む。ある実施形態では、前記AからGへの核酸塩基改変が、ハーラー症候群に関連するSNPにおけるものである。ある実施形態では、ハーラー症候群に関連する前記SNPが、配列番号4における番号付けでIDUAポリペプチドにおけるW402XもしくはW401Xアミノ酸変異体またはIDUA遺伝子によってコードされるそのバリエーションをもたらす、ここでXは終止コドンである。

【0035】

ある実施形態では、前記AからGへの核酸塩基改変が、ハーラー症候群に関連するSNPを野生型核酸塩基に変化させる。ある実施形態では、前記AからGへの核酸塩基改変が、ハーラー症候群に関連するSNPを、ハーラー症候群の1つ以上の症状の改善をもたらす非野生型核酸塩基に変化させる。ある実施形態では、ハーラー症候群に関連するSNPにおける前記AからGへの改変が、IDUA遺伝子によってコードされる終止コドンをIDUAポリペプチドにおけるトリプトファンに変化させる。

20

【0036】

ある実施形態では、前記ガイドポリヌクレオチドが、ハーラー症候群に関連するSNPを含むIDUA遺伝子またはその調節エレメントに相補的な核酸配列を含む。ある実施形態では、前記アデノシン塩基エディターが、ハーラー症候群に関連するSNPを含むIDUA遺伝子またはその調節エレメントに相補的な核酸配列を含む単一ガイドRNA (sgRNA)と複合体をなす。ある実施形態では、前記sgRNAが、5' - GACUCUAGGCAGAGGUCUCA - 3'、5' - ACUCUAGGCAGAGGUCUCA - 3'、5' - CUCUAGGCCGAAGUGUCGC - 3'、および5' - GCUCUAGGCCGAAGUGUCGC - 3' からなる群より選択される核酸配列を含む。

30

【0037】

ある態様において、本明細書では、メチルCpG結合タンパク質2 (MECP2) 遺伝子またはその調節エレメントを編集する方法であって、(i) アデノシン塩基エディターまたはそれをコードする核酸配列と、(ii) ガイドポリヌクレオチドまたはそれをコードする核酸配列を対象に投与することを含み、前記アデノシン塩基エディターは、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み、前記アデノシンデアミナーゼドメインは、配列番号2における番号付けでアミノ酸位置82もしくは166またはそれらに対応する位置でのアミノ酸置換を含み、前記ガイドポリヌクレオチドは、前記アデノシン塩基エディターを誘導してMECP2遺伝子またはその調節エレメントにおけるAからGへの核酸塩基改変をもたらす、方法が提供される。

40

【0038】

ある実施形態では、前記MECP2遺伝子またはその調節エレメントが、レット症候群に関連するSNPを含む。ある実施形態では、前記AからGへの核酸塩基改変が、レット症候群に関連するSNPにおけるものである。ある実施形態では、前記レット症候群と関連するSNPが、配列番号5における番号付けでMECP2ポリペプチドにおけるR106WもしくはT158Mアミノ酸変異体またはMECP2遺伝子によってコードされるそのバリエーションをもたらす。ある実施形態では、レット症候群に関連する前記SNPが、MECP2遺伝子によってコードされるMECP2ポリペプチドにおけるR255XまたはR270Xアミノ酸変異体をもたらす、ここで

50

Xは終止コドンである。

【0039】

ある実施形態では、前記AからGへの核酸塩基改変が、レット症候群に関連するSNPを野生型核酸塩基に変化させる。ある実施形態では、前記AからGへの核酸塩基改変が、レット症候群に関連するSNPを、レット症候群の1つ以上の症状の改善をもたらす非野生型核酸塩基に変化させる。ある実施形態では、レット症候群に関連するSNPにおける前記AからGへの核酸塩基改変が、MECP2ポリペプチドにおいて終止コドンをトリプトファンに変化させる。

【0040】

ある実施形態では、前記ガイドポリヌクレオチドが、レット症候群に関連するSNPを含むMECP2遺伝子またはその調節エレメントに相補的な核酸配列を含む。ある実施形態では、前記アデノシン塩基エディターが、レット症候群と関連するSNPを含むMECP2遺伝子またはその調節エレメントに相補的な核酸配列を含む単一ガイドRNA (sgRNA) と複合体をなす。ある実施形態では、前記ガイドポリヌクレオチドが、5' - CUUUUCACUCCUGCCGGGG-3'、5' -AGCUUCCAUGUCCAGCCUUC-3'、5' - ACCAUGAAGUCAAAAUCAUU-3'、および5' - GCUUUCAGCCCCGUUUCUUG-3' からなる群より選択される核酸配列を含む。

10

【0041】

ある態様において、本明細書は、ATP結合カセットサブファミリーメンバー4 (ABCA4) 遺伝子またはその調節エレメントを編集する方法であって、前記ABCA4遺伝子またはその調節エレメントを、(i) アデノシン塩基エディターまたはそれをコードする核酸配列と、(ii) ガイドポリヌクレオチドまたはそれをコードする核酸配列と接触させることを含み、前記アデノシン塩基エディターは、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み、前記アデノシンデアミナーゼドメインは、配列番号2における番号付けでアミノ酸位置82もしくは166、またはそれらに対応する位置でのアミノ酸置換を含み、前記ガイドポリヌクレオチドは、前記アデノシン塩基エディターを誘導してABCA4遺伝子またはその調節エレメントにおけるAからGへの核酸塩基改変をもたらす、方法が提供される。

20

【0042】

ある実施形態では、前記投与が、シュタルガルト病に関連する少なくとも一つの症状を改善する。ある実施形態では、前記投与が、アデノシンデアミナーゼにおける前記アミノ酸置換を伴わない塩基エディターを用いた治療と比較して、シュタルガルト病に関連する少なくとも一つの症状のより迅速な改善をもたらす。

30

【0043】

ある実施形態では、ABCA4遺伝子が、シュタルガルト病に関連するSNPを含む。ある実施形態では、前記AからGへの核酸塩基改変が、シュタルガルト病に関連するSNPにおけるものである。ある実施形態では、シュタルガルト病に関連する前記SNPが、配列番号6における番号付けでABCA4ポリペプチドにおけるA1038VもしくはG1961Eアミノ酸変異体またはABCA4遺伝子によってコードされるそのバリエーションをもたらす。ある実施形態では、シュタルガルト病に関連する前記SNPが、配列番号6における番号付けでABCA4ポリペプチドにおけるG1961Eアミノ酸変異体またはそのバリエーションをもたらす。

40

【0044】

ある実施形態では、前記AからGへの核酸塩基改変が、シュタルガルト病に関連するSNPを野生型核酸塩基に変化させる。ある実施形態では、前記AからGへの核酸塩基の改変が、シュタルガルト病に関連するSNPを、シュタルガルト病の1つ以上の症状の改善をもたらす非野生型核酸塩基に変化させる。ある実施形態では、前記ガイドポリヌクレオチドが、シュタルガルト病に関連するSNPを含むABCA4遺伝子またはその調節エレメントに相補的な核酸配列を含む。

【0045】

ある実施形態では、前記アデノシン塩基エディターが、シュタルガルト病に関連するSN

50

Pを含むABCA4遺伝子またはその調節エレメントに相補的な核酸配列を含む単一ガイドRNA (sgRNA) と複合体をなす。ある実施形態では、sgRNAが配列5'-CUCCAGGGCGAACUUCGACACACAGC-3'を含む。

【0046】

上記態様の様々な実施形態において、前記接触させることは、細胞内で起こる。ある実施形態では、前記接触させることが、細胞中のゲノムにおいて10%未満のインデルを生じ、ここでインデル率は、単一ヌクレオチド改変に隣接する配列と未改変配列との間のミスマッチ頻度によって測定される。ある実施形態では、前記接触させることが、細胞中のゲノムにおいて5%未満のインデルを生じ、ここでインデル率は、単一ヌクレオチド改変に隣接する配列と未改変配列との間のミスマッチ頻度によって測定される。ある実施形態では、前記接触させることが、細胞中のゲノムにおいて1%未満のインデルを生じ、ここでインデル率は、単一ヌクレオチド改変に隣接する配列と未改変配列との間のミスマッチ頻度によって測定される。

10

【0047】

上記態様の様々な実施形態において、前記細胞はニューロンである。ある実施形態では、前記接触させることが、細胞の集団内で起こる。ある実施形態では、前記接触させることが、前記接触の工程の後に前記細胞の集団の少なくとも40%においてAからGへの核酸塩基改変をもたらす。ある実施形態では、前記接触させることが、前記接触の工程の後に前記細胞の集団の少なくとも50%においてAからGへの核酸塩基改変をもたらす。ある実施形態では、前記接触させることが、前記接触の工程の後に前記細胞の集団の少なくとも70%においてAからGへの核酸塩基改変をもたらす。ある実施形態では、前記細胞の少なくとも90%が、前記接触の工程の後に生存性である。ある実施形態では、前記細胞の集団が、前記接触の工程の後に濃縮されなかった。ある実施形態では、前記細胞の集団はニューロンである。ある実施形態では、前記接触させることがインピボまたはエキスピボで起こる。

20

【0048】

上記様々な態様および実施形態において、前記ポリヌクレオチドによりプログラム可能なDNA結合ドメインはCas9である。ある実施形態では、前記Cas9がSpCas9、SaCas9、またはそれらのバリエーションである。ある実施形態では、前記ポリヌクレオチドによりプログラム可能なDNA結合ドメインが、改変されたプロトSpacer隣接モチーフ (PAM) 特異性を有する改変されたSpCas9を含む。ある実施形態では、前記Cas9が、NGG、NGA、NGCG、NGN、NNGRRT、NNRRT、NGCG、NGCN、NGTN、およびNGCからなる群より選択されるPAM配列に対する特異性を有し、ここでNはA、G、C、またはTでありRはAまたはGである。ある実施形態では、前記ポリヌクレオチドによりプログラム可能なDNA結合ドメインが、ヌクレアーゼ不活性バリエーションである。ある実施形態では、前記ポリヌクレオチドによりプログラム可能なDNA結合ドメインがニッカーゼバリエーションである。ある実施形態では、前記ニッカーゼバリエーションが、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。本明細書で提供される態様および実施形態において、前記アデノシンデアミナーゼドメインはTadAドメインを含む。ある実施形態では、前記アデノシンデアミナーゼが、V82S改変および/またはT166R改変を含むTadAデアミナーゼを含む。

30

【0049】

上記様々な態様および実施形態において、アデノシンデアミナーゼは、Y147T、Y147R、Q154S、Y123H、Q154R、またはそれらの組み合わせのうちの1つ以上の改変をさらに含む。本明細書で提供される態様および実施形態において、前記アデノシンデアミナーゼは、Y147R + Q154R + Y123H；Y147R + Q154R + I76Y；Y147R + Q154R + T166R；Y147T + Q154R；Y147T + Q154S；およびY123H + Y147R + Q154R + I76Yからなる群より選択される改変の組み合わせを含む。本明細書で提供される態様および実施形態において、前記アデノシン塩基エディタードメインはアデノシンデアミナーゼ単量体を含む。本明細書で提供される態様および実施形態において、前記アデノシン塩基エディターはアデノシンデアミナーゼ二量体を含む。ある実施形態では、前記TadAデアミナーゼがTadA*8バリエーションである。ある実施形態では、前記TadA*8バリエーションがTadA*8.

40

50

1、TadA*8.2、TadA*8.3、TadA*8.4、TadA*8.5、TadA*8.6、TadA*8.7、TadA*8.8、TadA*8.9、TadA*8.10、TadA*8.11、TadA*8.12、およびTadA*8.13からなる群から選択される。ある実施形態では、前記アデノシン塩基エディターが、ABE8.1、ABE8.2、ABE8.3、ABE8.4、ABE8.5、ABE8.6、ABE8.7、ABE8.8、ABE8.9、ABE8.10、ABE8.11、ABE8.12、およびABE8.13からなる群より選択されるABE8塩基エディターである。

【0050】

ある態様において、本明細書では、本明細書で開示される様々な態様および実施形態で記述される方法によって生成される細胞が提供される。ある態様において、本明細書では、本明細書で開示される様々な態様および実施形態で記述される方法によって生成される細胞の集団が提供される。

10

【0051】

ある態様において、本明細書では、(i) アデノシン塩基エディターまたはそれをコードする核酸配列と、(ii) ガイドポリヌクレオチドまたはそれをコードする核酸配列とを含む、塩基エディターシステムであって、前記アデノシン塩基エディターが、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み、前記アデノシンデアミナーゼドメインは、配列番号2における番号付けでアミノ酸位置82もしくは166またはそれらに対応する位置でのアミノ酸置換を含み、前記ガイドポリヌクレオチドは、前記アデノシン塩基エディターを誘導して神経障害に関連する標的遺伝子またはその調節エレメントにおけるAからGへの核酸塩基改変をもたらす、塩基エディターシステムが提供される。上記態様のある実施形態では、前記標的遺伝子がロイシンリッチリピートキナーゼ2 (LRRK2) 遺伝子であり、神経疾患がパーキンソン病である。この態様のもう1つの実施形態では、前記標的遺伝子がアルファ-L-イズロニダーゼ (IDUA) 遺伝子であり、神経疾患がハーラー症候群である。この態様の一実施形態では、前記標的遺伝子がメチルCpG結合タンパク質2 (MECP2) 遺伝子であり、神経疾患がレット症候群である。この態様の別の一実施形態では、前記標的遺伝子がATP結合カセットサブファミリーメンバー4 (ABCA4) 遺伝子であり、神経疾患がシュタルガルト病である。

20

【0052】

ある態様において、本明細書では、(i) アデノシン塩基エディターまたはそれをコードする核酸配列と、(ii) ガイドポリヌクレオチドまたはそれをコードする核酸配列を含む、塩基エディターシステムであって、前記アデノシン塩基エディターが、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み、前記アデノシンデアミナーゼドメインは、配列番号2における番号付けでアミノ酸位置82もしくは166またはそれらに対応する位置でのアミノ酸置換を含み、前記ガイドポリヌクレオチドが、前記アデノシン塩基エディターを誘導してLRRK2遺伝子またはその調節エレメントにおけるAからGへの核酸塩基改変をもたらす、塩基エディターシステムが提供される。

30

【0053】

ある実施形態では、前記AからGへの核酸塩基改変は、LRRK2遺伝子またはその調節エレメント中のパーキンソン病に関連するSNPにおけるものである。ある実施形態では、前記パーキンソン病に関連するSNPが、配列番号3における番号付けでLRRK2ポリペプチドにおけるA419V、R1441C、R1441H、もしくはG2019Sアミノ酸変異体またはLRRK2遺伝子によってコードされるそのバリエーションをもたらす。

40

【0054】

ある実施形態では、前記AからGへの核酸塩基改変が、前記パーキンソン病に関連するSNPを野生型核酸塩基に変化させる。ある実施形態では、前記AからGへの核酸塩基改変が、前記パーキンソン病に関連するSNPを、パーキンソン病の症状の改善をもたらす非野生型核酸塩基に変化させる。ある実施形態では、前記AからGへの核酸塩基の改変が、LRRK2遺伝子によってコードされるLRRK2ポリペプチド中のシステインまたはヒスチジンをアルギニンに変化させる。ある実施形態では、前記AからGへの改変が、LRRK2遺伝子によってコードされるLRRK2ポリペプチドにおいてセリンをグリシンに変化させる。ある実施

50

形態では、前記AからGへの改変が、配列番号3における番号付けでLRRK2ポリペプチドまたはLRRK2遺伝子によってコードされるそのバリエーションの位置144でシステイン (C) もしくはヒスチジン (H) をアルギニン (R) に置換するか、または位置2019でセリンをグリシン (G) に置換する。ある実施形態では、前記アデノシンデアミナーゼドメインが、配列番号2における番号付けでアミノ酸位置82もしくは166またはそれに対応する位置におけるアミノ酸置換を含む。

【0055】

ある実施形態では、前記ガイドポリヌクレオチドが、パーキンソン病に関連するSNPを含むLRRK2遺伝子またはその調節エレメントに相補的な核酸配列を含む。ある実施形態では、前記アデノシン塩基エディターが、パーキンソン病と関連する前記SNPを含むLRRK2遺伝子またはその調節エレメントに相補的な核酸配列を含む単一ガイドRNA (sgRNA) と複合体をなす。ある実施形態では、前記sgRNAが核酸配列5' -AAGCGCAAGCCUGGAGG GAA -3' ; または5' -ACUACAGCAUUGCUCAGUAC-3' を含む。

【0056】

ある態様において、本明細書では、(i) アデノシン塩基エディターまたはそれをコードする核酸配列と、(ii) ガイドポリヌクレオチドまたはそれをコードする核酸配列とを含む、塩基エディターシステムであって、前記アデノシン塩基エディターが、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み、前記アデノシンデアミナーゼドメインは、配列番号2における番号付けでアミノ酸位置82もしくは166またはそれに対応する位置でのアミノ酸置換を含み、前記ガイドポリヌクレオチドが、前記アデノシン塩基エディターを誘導してアルファ-L-イズロニダーゼ (IDUA) 遺伝子またはその調節エレメントにおけるAからGへの核酸塩基改変をもたらす、塩基エディターシステムが提供される。

【0057】

ある実施形態では、前記IDUA遺伝子またはその調節エレメントが、ハーラー症候群に関連するSNPを含む。ある実施形態では、前記AからGへの核酸塩基改変はハーラー症候群に関連するSNPにおけるものである。ある実施形態では、前記ハーラー症候群に関連するSNPが、配列番号4における番号付けでIDUAポリペプチドにおけるW402XもしくはW401Xアミノ酸変異体またはIDUA遺伝子によってコードされるそのバリエーションをもたらす、ここでXは終止コドンである。

【0058】

ある実施形態では、前記AからGへの核酸塩基改変が、ハーラー症候群に関連するSNPを野生型核酸塩基に変化させる。ある実施形態では、前記AからGへの核酸塩基の改変が、ハーラー症候群に関連するSNPを、ハーラー症候群の1つ以上の症状の改善をもたらす非野生型核酸塩基に変化させる。ある実施形態では、ハーラー症候群に関連するSNPにおけるAからGへの改変が、IDUA遺伝子によってコードされるIDUAポリペプチドにおいて終止コドントリプトファンに変化させる。

【0059】

ある実施形態では、前記ガイドポリヌクレオチドが、ハーラー症候群と関連するSNPを含むIDUA遺伝子またはその調節エレメントに相補的な核酸配列を含む。ある実施形態では、前記アデノシン塩基エディターが、ハーラー症候群と関連するSNPを含むIDUA遺伝子またはその調節エレメントに相補的な核酸配列を含む単一ガイドRNA (sgRNA) と複合体をなす。ある実施形態では、前記sgRNAが、5' - GACUCUAGGCAGAGGUCUCAA -3'、5' - ACUCUAGGCAGAGGUCUCAA-3'、5' - CUCUAGGCCGAAGUGUCGC -3'、および5' -GCUCUAGGCCGAAGUGUCGC-3' からなる群より選択される核酸配列を含む。

【0060】

いくつかの態様において、本明細書では、(i) アデノシン塩基エディターまたはそれをコードする核酸配列と、(ii) ガイドポリヌクレオチドまたはそれをコードする核酸配列とを含む、塩基エディターシステムであって、前記アデノシン塩基エディターは、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み、前記アデノシ

ンデアミナーゼドメインは、配列番号2における番号付けでアミノ酸位置82もしくは166またはそれらに対応する位置でのアミノ酸置換を含み、前記ガイドポリヌクレオチドは、前記アデノシン塩基エディターを誘導してメチルCpG結合タンパク質2 (MECP2) 遺伝子またはその調節エレメントにおけるAからGへの核酸塩基改変をもたらす、塩基エディターシステムが提供される。

【0061】

ある実施形態では、前記MECP2遺伝子またはその調節エレメントが、レット症候群に関連するSNPを含む。ある実施形態では、前記AからGへの核酸塩基改変は、レット症候群に関連するSNPにおけるものである。ある実施形態では、レット症候群に関連するSNPが、配列番号5における番号付けでMECP2ポリペプチドにおけるR106WまたはT158Mアミノ酸変異体またはMECP2遺伝子によってコードされるそのバリエーションをもたらす。ある実施形態では、レット症候群に関連するSNPが、MECP2遺伝子によってコードされるMECP2ポリペプチドにおけるR255XまたはR270Xアミノ酸変異体をもたらす、ここでXは終止コドンである。

【0062】

ある実施形態では、前記AからGへの核酸塩基改変が、レット症候群に関連するSNPを野生型核酸塩基に変化させる。ある実施形態では、前記AからGへの核酸塩基改変が、レット症候群に関連するSNPを、レット症候群の1つ以上の症状の改善をもたらす非野生型核酸塩基に変化させる。ある実施形態では、レット症候群に関連するSNPにおけるAからGへの核酸塩基改変が、MECP2ポリペプチドにおいて終止コドンをトリプトファンに変化させる。

【0063】

ある実施形態では、前記ガイドポリヌクレオチドが、レット症候群に関連するSNPを含むMECP2遺伝子またはその調節エレメントに相補的な核酸配列を含む。ある実施形態では、前記アデノシン塩基エディターが、レット症候群に関連するSNPを含むMECP2遺伝子またはその調節エレメントに相補的な核酸配列を含む単一ガイドRNA (sgRNA) と複合体をなす。ある実施形態では、前記ガイドポリヌクレオチドが、5' - CUUUUCACUCCUGC CGGGG-3'、5' - AGCUUCCAUGUCCAGCCUUC-3'、5' - ACCAUGAAGUCAAAAUC AUU-3'、および5' - GCUUUCAGCCCCGUUUCUUG-3' からなる群から選択される核酸配列を含む。

【0064】

ある態様において、本明細書では、(i) アデノシン塩基エディターまたはそれをコードする核酸配列と、(ii) ガイドポリヌクレオチドまたはそれをコードする核酸配列とを接合させることを含む、塩基エディターシステムであって、前記アデノシン塩基エディターが、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み、前記アデノシンデアミナーゼドメインは、配列番号2における番号付けでアミノ酸位置82もしくは166またはそれらに対応する位置でのアミノ酸置換を含み、前記ガイドポリヌクレオチドは、前記アデノシン塩基エディターを誘導してATP結合カセットサブファミリーメンバー4 (ABCA4) 遺伝子またはその調節エレメントにおけるAからGへの核酸塩基改変をもたらす、塩基エディターシステムが提供される。

【0065】

ある実施形態では、投与が、シュタルガルト病に関連する少なくとも一つの症状を改善する。ある実施形態では、投与が、アデノシンデアミナーゼにおける前記アミノ酸置換を伴わない塩基エディターを用いた治療と比較して、シュタルガルト病に関連する少なくとも一つの症状のより迅速な改善をもたらす。ある実施形態では、ABCA4遺伝子が、シュタルガルト病に関連するSNPを含む。ある実施形態では、前記AからGへの核酸塩基改変は、シュタルガルト病に関連するSNPにおけるものである。ある実施形態では、前記シュタルガルト病に関連するSNPが、配列番号6における番号付けでABCA4ポリペプチドにおけるA1038VもしくはG1961Eアミノ酸変異体またはABCA4遺伝子によってコードされるそのバリエーションをもたらす。ある実施形態では、前記シュタルガルト病に関連するSNPが、配

列番号6における番号付けでABCA4ポリペプチドにおけるG1961Eアミノ酸変異体またはそのバリエーションをもたらす。

【0066】

ある実施形態では、前記AからGへの核酸塩基改変が、シュタルガルト病に関連するSNPを野生型核酸塩基に変化させる。ある実施形態では、前記AからGへの核酸塩基改変が、シュタルガルト病に関連するSNPを、シュタルガルト病症状の改善をもたらす非野生型核酸塩基に変化させる。ある実施形態では、前記ガイドポリヌクレオチドは、シュタルガルト病に関連するSNPを含むABCA4遺伝子またはその調節エレメントに相補的な核酸配列を含む。

【0067】

ある実施形態では、前記アデノシン塩基エディターは、前記シュタルガルト病に関連するSNPを含むABCA4遺伝子またはその調節エレメントに相補的な核酸配列を含む単一ガイドRNA (sgRNA)と複合体をなす。ある実施形態では、前記sgRNAは配列5'-CUCCAGGCGAACUUCGACACACAGC-3'を含む。

【0068】

本明細書に提供される様々な態様および実施形態において、前記ポリヌクレオチドによりプログラム可能なDNA結合ドメインはCas9である。ある実施形態では、前記Cas9がSpCas9、SaCas9、またはそのバリエーションである。ある実施形態では、前記ポリヌクレオチドによりプログラム可能な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) 特異性を有する改変されたSpCas9を含む。ある実施形態では、前記Cas9が、NGG、NGA、NGCG、NGN、NNGRRT、NNNRRT、NGCG、NGCN、NGTN、およびNGCからなる群より選択されるPAM配列に対する特異性を有し、ここでNはA、G、CまたはTであり、RはAまたはGである。ある実施形態では、前記ポリヌクレオチドによりプログラム可能なDNA結合ドメインがヌクレアーゼ不活性バリエーションである。ある実施形態では、前記ポリヌクレオチドによりプログラム可能なDNA結合ドメインがニックアーゼバリエーションである。ある実施形態では、前記ニックアーゼバリエーションが、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。

【0069】

本明細書に提供される様々な態様および実施形態において、前記アデノシンデアミナーゼドメインはTadAドメインを含む。ある実施形態では、前記アデノシンデアミナーゼが、V82S改変および/またはT166R改変を含むTadAデアミナーゼを含む。

【0070】

本明細書に提供される様々な態様および実施形態において、前記アデノシンデアミナーゼは、Y147T、Y147R、Q154S、Y123H、Q154R、またはそれらの組み合わせのうちの1つ以上の改変をさらに含む。本明細書に提供される様々な態様および実施形態において、前記アデノシンデアミナーゼは、Y147R + Q154R + Y123H; Y147R + Q154R + I76Y; Y147R + Q154R + T166R; Y147T + Q154R; Y147T + Q154S; およびY123H + Y147R + Q154R + I76Yからなる群から選択される改変の組み合わせを含む。ある実施形態では、前記アデノシン塩基エディタードメインはアデノシンデアミナーゼ単量体を含む。ある実施形態では、前記アデノシン塩基エディターはアデノシンデアミナーゼ二量体を含む。

【0071】

本明細書に提供される様々な態様および実施形態において、前記TadAデアミナーゼはTadA*8バリエーションである。ある実施形態では、前記TadA*8バリエーションは、TadA*8.1、TadA*8.2、TadA*8.3、TadA*8.4、TadA*8.5、TadA*8.6、TadA*8.7、TadA*8.8、TadA*8.9、TadA*8.10、TadA*8.11、TadA*8.12、およびTadA*8.13からなる群から選択される。ある実施形態では、前記アデノシン塩基エディターが、ABE8.1、ABE8.2、ABE8.3、ABE8.4、ABE8.5、ABE8.6、ABE8.7、ABE8.8、ABE8.9、ABE8.10、ABE8.11、ABE8.12、およびABE8.13からなる群から選択されるABE8塩基エディターである。

【0072】

10

20

30

40

50

ある態様において、本明細書では、本明細書に記述されるアデノシン塩基エディターをコードする核酸配列を含むベクターが提供される。ある態様において、本明細書では、本明細書に記述されるアデノシン塩基エディターおよびガイドポリヌクレオチドをコードする核酸配列を含むベクターが提供される。ある実施形態では、前記ベクターはウイルスベクター、レンチウイルスベクター、またはAAVベクターである。

【0073】

ある態様において、本明細書では、本明細書に記述される塩基エディターシステムまたはベクターを含む、細胞が提供される。ある実施形態では、細胞は中枢神経系細胞である。ある実施形態では、細胞はニューロンである。ある実施形態では、細胞は光受容体である。ある実施形態では、細胞はインビトロ、インビボまたはエクスピボである。

10

【0074】

ある態様において、本明細書では、本明細書に記述される塩基エディター、ベクター、又は細胞と、薬学的に許容される担体とを含む、医薬組成物が提供される。一実施形態では、本明細書に記述される医薬組成物はさらに脂質を含む。別の一実施形態では、本明細書に記述される医薬組成物はさらにウイルスを含む。

【0075】

ある態様において、本明細書では、本明細書に記述される塩基エディターまたはベクターを含む、キットが提供される。

【0076】

本明細書に記述される方法の様々な実施形態において、前記ガイドポリヌクレオチドの少なくとも1つのヌクレオチドが非天然の修飾を含む。本明細書に記述される方法の様々な実施形態において、前記核酸配列の少なくとも一つのヌクレオチドが非天然の修飾を含む。様々な実施形態において、前記核酸配列の少なくとも一つのヌクレオチドが非天然の修飾を含む。ある実施形態では、前記非天然の修飾は化学修飾である。ある実施形態では、前記化学修飾は2'-O-メチル化である。ある実施形態では、前記核酸配列がホスホロチオエートを含む。

20

【0077】

本明細書における説明および例は、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、したがって変化し得ることが理解されるべきである。当業者であれば、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認識するであろう。

30

【0078】

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。

40

【0079】

本明細書で使用されているセクションの見出しは、まとめる目的のみのためのものであり、説明されている主題を限定するものとして解釈されるべきではない。

【0080】

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。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

50

のみのためであり、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。

【0081】

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。下記の詳細な説明（これは本開示の原理が利用される例示的实施形態を提示する）を参照することによって、そして下記において説明する添付図面を考慮して、本開示の特徴および利点のよりよい理解が得られる。

【0082】

定義

以下の定義は、当該技術分野の定義を補足するものであって本出願を対象としており、関連するまたは関連性のない案件、例えば共通の所有に係る特許または出願に帰するものではない。本明細書に記載されたものと同様または同等の任意の方法および材料を、本開示の試験の実施において使用することができるが、好ましい材料および方法を本明細書で説明する。従って、本明細書で使用される用語は、特定の实施形態を説明する目的のみのためのものであり、限定することを意図するものではない。

【0083】

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)。

【0084】

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。本明細書で使われるように、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しなければならない。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」を意味し、含める意味に理解される。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。

【0085】

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。

【0086】

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、例えば5倍以内、または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。

【0087】

本明細書で提供される範囲は、範囲内のすべて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。

【0088】

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徴、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。

【0089】

「脱塩基 (abasic) 塩基エディター」とは、核酸塩基を切り出してDNA核酸塩基(A、T、C、またはG)を挿入することができる薬剤を意味する。脱塩基塩基エディターは、核酸グリコシラーゼポリペプチドまたはその断片を含む。一実施形態において、核酸グリコシラーゼは、下記の配列のアミノ酸204またはウラシルDNAグリコシラーゼの対応位置においてAspを含み（例えば、アミノ酸204におけるAsnを置換して）、かつシトシン-DNAグリコシラーゼ活性を有する、変異型ヒトウラシルDNAグリコシラーゼまたはその活性断片である。一実施形態において、核酸グリコシラーゼは、下記の配列のアミノ酸147またはウラシルDNAグリコシラーゼの対応位置においてAla、Gly、Cys、またはSerを含み（例えば、アミノ酸147におけるTyrを置換して）、かつチミン-DNAグリコシラーゼ活性を有する、変異型ヒトウラシルDNAグリコシラーゼまたはその活性断片である。例示的なヒトウラシル-DNAグリコシラーゼ、アイソフォーム1の配列は以下の通りである：

```
1 mgvfclgpwg lgrklrtpgk gplqlslrlc gdhlqaipak kapagqeepeg tppssplsae
61 qldriqrnka aallrlaarn vpvfgeswk khlsgefkgp yfiklmgfva eerkhylvyp
121 ppqhvtwtq mcdikdvkvv ilgqdpvhgp nqahglcfsv qrpvppppsl eniykelstd
181 iedfvhpghg dlsgwakqgv llnavltvr ahqanshker gweqftdavu swlnqnsngl
241 vllwgsyaa kkgasidrkr hhvltahps plsvyrgffg crhfsktnel lqksgkkpid
301 wkel
```

【0090】

ヒトのウラシル-DNAグリコシラーゼ、アイソフォーム2の配列は以下の通りである。

```
1 migqktlysf fspsparkrh apseparvqg tgvagvpees gdaaapakk apagqeepeg
61 ppssplsaeq ldriqrnkaa allrlaarnv pvgfgeswkk hlsgefkgpy fiklmgfva
121 erkhytvypp phqvftwtqm cdikdvkvvi ilgqdpvhgpn qahglcfsvq rpvpvpppsle
181 niykelstdi edfvhpghgd lsgwakqgv llnavltvra hqanshker gweqftdavvs
241 wlnqnsnglv fllwgsyaqk kgsaidrkrh hvlqtahpsp lsvyrgffgc rhfsktnell
301 qksgkkpidw kel
```

【0091】

他の実施形態では、脱塩基エディターは、PCT/JP205/080958およびUS20170321210に記載されている脱塩基エディターのいずれかであり、これらは参照により本明細書に組み入れられる。特定の実施形態では、脱塩基エディターは、上記の配列に太字、下線で示された位置、または当該技術分野で知られる任意の他の脱塩基エディターまたはウラシルデグリコシラーゼにおける対応アミノ酸において突然変異を含む。一実施形態では、脱塩基エディターは、Y147、N204、L272、および/またはR276または対応位置に突然変異を含む。別の実施形態では、脱塩基エディターは、Y147AもしくはY147G突然変異または対応する突然変異を含む。別の実施形態では、脱塩基エディターは、N204D突然変異または対応する突然変異を含む。別の実施形態では、脱塩基エディターは、L272A突然変異または対応する突然変異を含む。別の実施形態では、脱塩基エディターは、R276EもしくはR276C突然変異または対応する突然変異を含む。

【0092】

「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。

【0093】

ある実施形態において、アデノシンデアミナーゼは、*TadA*デアミナーゼである。ある実施形態において、*TadA*デアミナーゼは*TadA*バリエーションである。ある実施形態において、*TadA*バリエーションは*TadA**8である。ある実施形態において、デアミナーゼまたはデアミナーゼドメインは、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 / 2007/045381 (WO2018 / 027078)およびPCT / US2016 / 058344(WO 2017/070632)に記載されており、これらのそれぞれは、その全体が参照により本明細書に組み込まれる。また、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) , 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（その全内容は参照により本明細書に組み込まれる）。

【0094】

野生型*TadA* (wt) アデノシンデアミナーゼには、次の配列を有する (*TadA*参照配列とも呼ばれる) :

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAH
AEIMALRQGGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAGS
LMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAKKAQSSTD (SEQ ID NO: 2).

【0095】

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、以下の配列における改変を含む :

MSEVEFSHEY WMRHALTLAK RARDEREVPV GAVLVLNRRV IGEGWNRAIG LHDPT
AHAEI MALRQGGGLVM QNYRLIDATL YVTFEPCVMC AGAMIHSRIG RVVFGVRNAK T
GAAGSLMDV LHYPGMNHRV EITEGILADE CAALLCYFFR MPRQVFNAQK KAQSSTD
(*TadA**7.10とも呼ばれる) 。

【0096】

いくつかの実施形態では、 $TadA^{*7.10}$ は、少なくとも1つの改変を含む。いくつかの実施形態において、 $TadA^{*7.10}$ は、アミノ酸82および/または166での改変を含む。特定の実施形態において、上記参照配列のバリエーションは、以下の改変のうちの1つ以上を含む： $Y147T$ 、 $Y147R$ 、 $Q154S$ 、 $Y123H$ 、 $V82S$ 、 $T166R$ 、および/または $Q154R$ 。改変 $Y123H$ は、本明細書では $H123H$ とも呼ばれる（ $TadA^{*7.10}$ における改変 $H123Y$ が $Y123H$ （wt）に戻ったもの）。他の実施形態では、 $TadA^{*7.10}$ 配列のバリエーションは、以下のものからなる群から選択される改変の組み合わせを含む： $Y147T + Q154R$ ； $Y147T + Q154S$ ； $Y147R + Q154S$ ； $V82S + Q154S$ ； $V82S + Y147R$ ； $V82S + Q154R$ ； $V82S + Y123H$ ； $I76Y + V82S$ ； $V82S + Y123H + Y147T$ ； $V82S + Y123H + Y147R$ ； $V82S + Y123H + Q154R$ ； $Y147R + Q154R + Y123H$ ； $Y147R + Q154R + I76Y$ ； $Y147R + Q154R + T166R$ ； $Y123H + Y147R + Q154R + I76Y$ ； $V82S + Y123H + Y147R + Q154R$ ；および $I76Y + V82S + Y123H + Y147R + Q154R$ 。

【0097】

他の実施形態において、本発明は、残基149、150、151、152、153、154、155、156、または157で始まるC末端の欠失を含む、欠失を含有するアデノシンデアミナーゼバリエーション、例えば $TadA^{*8}$ を提供する。他の実施形態において、アデノシンデアミナーゼバリエーションは、以下の改変のうちの1つ以上を含む $TadA$ （例えば $TadA^{*8}$ ）モノマーである： $Y147T$ 、 $Y147R$ 、 $Q154S$ 、 $Y123H$ 、 $V82S$ 、 $T166R$ 、および/または $Q154R$ 。他の実施形態において、アデノシンデアミナーゼバリエーションは、以下のものからなる群から選択される改変の組み合わせを含む $TadA$ （例えば $TadA^{*8}$ ）モノマーである： $Y147T + Q154R$ ； $Y147T + Q154S$ ； $Y147R + Q154S$ ； $V82S + Q154S$ ； $V82S + Y147R$ ； $V82S + Q154R$ ； $V82S + Y123H$ ； $I76Y + V82S$ ； $V82S + Y123H + Y147T$ ； $V82S + Y123H + Y147R$ ； $V82S + Y123H + Q154R$ ； $Y147R + Q154R + Y123H$ ； $Y147R + Q154R + I76Y$ ； $Y147R + Q154R + T166R$ ； $Y123H + Y147R + Q154R + I76Y$ ； $V82S + Y123H + Y147R + Q154R$ ；および $I76Y + V82S + Y123H + Y147R + Q154R$ 。

【0098】

さらに他の実施形態では、アデノシンデアミナーゼバリエーションは、それぞれが以下の改変 $Y147T$ 、 $Y147R$ 、 $Q154S$ 、 $Y123H$ 、 $V82S$ 、 $T166R$ 、および/または $Q154R$ のうちの1つ以上を有する2つのアデノシンデアミナーゼドメイン（例えば $TadA^{*8}$ ）を含むホモ二量体である。他の実施形態では、アデノシンデアミナーゼバリエーションは、それぞれが以下のものからなる群から選択される改変の組み合わせを有する2つのアデノシンデアミナーゼドメイン（例えば $TadA^{*8}$ ）を含むホモ二量体である： $Y147T + Q154R$ ； $Y147T + Q154S$ ； $Y147R + Q154S$ ； $V82S + Q154S$ ； $V82S + Y147R$ ； $V82S + Q154R$ ； $V82S + Y123H$ ； $I76Y + V82S$ ； $V82S + Y123H + Y147T$ ； $V82S + Y123H + Y147R$ ； $V82S + Y123H + Q154R$ ； $Y147R + Q154R + Y123H$ ； $Y147R + Q154R + I76Y$ ； $Y147R + Q154R + T166R$ ； $Y123H + Y147R + Q154R + I76Y$ ； $V82S + Y123H + Y147R + Q154R$ ；および $I76Y + V82S + Y123H + Y147R + Q154R$ 。

【0099】

他の実施形態において、アデノシンデアミナーゼバリエーションは、野生型 $TadA$ アデノシンデアミナーゼドメインと、以下の改変 $Y147T$ 、 $Y147R$ 、 $Q154S$ 、 $Y123H$ 、 $V82S$ 、 $T166R$ 、および/または $Q154R$ のうちの1つ以上を含むアデノシンデアミナーゼバリエーションドメイン（例えば $TadA^{*8}$ ）とを含むヘテロダイマーである。他の実施形態では、アデノシンデアミナーゼバリエーションは、野生型 $TadA$ アデノシンデアミナーゼドメインと、以下からなる群から選択される改変の組み合わせを含むアデノシンデアミナーゼバリエーションドメイン（例えば $TadA^{*8}$ ）とを含むヘテロダイマーである： $Y147T + Q154R$ ； $Y147T + Q154S$ ； $Y147R + Q154S$ ； $V82S + Q154S$ ； $V82S + Y147R$ ； $V82S + Q154R$ ； $V82S + Y123H$ ； $I76Y + V82S$ ； $V82S + Y123H + Y147T$ ； $V82S + Y123H + Y147R$ ； $V82S + Y123H + Q154R$ ； $Y147R + Q154R + Y123H$ ； $Y147R + Q154R + I76Y$ ； $Y147R + Q154R + T166R$ ； $Y123H + Y147R + Q154R + I76Y$ ； $V82S + Y123H + Y147R + Q154R$ ；および $I76Y + V82S + Y123H + Y147R + Q154R$ 。

【0100】

他の実施形態では、アデノシンデアミナーゼバリエーションは、TadA*7.10ドメインと、以下の改変Y147T、Y147R、Q154S、Y123H、V82S、T166R、および/またはQ154Rのうちの1つ以上を含むアデノシンデアミナーゼバリエーションドメイン（例えばTadA*8）とを含むヘテロダイマーである。他の実施形態において、アデノシンデアミナーゼバリエーションドメイン（例えばTadA*8）とを含むヘテロ二量体である：Y147T + Q154R; Y147T + Q154S; Y147R + Q154S; V82S + Q154S; V82S + Y147R; V82S + Q154R; V82S + Y123H; I76Y + V82S; V82S + Y123H + Y147T; V82S + Y123H + Y147R; V82S + Y123H + Q154R; Y147R + Q154R + Y123H; Y147R + Q154R + I76Y; Y147R + Q154R + T166R; Y123H + Y147R + Q154R + I76Y; V82S + Y123H + Y147R + Q154R; またはI76Y + V82S + Y123H + Y147R + Q154R。

10

【0101】

一実施形態では、アデノシンデアミナーゼは、以下の配列またはアデノシンデアミナーゼ活性を有するそのフラグメントを含むかまたは本質的にそれらからなるTadA*8である：MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVNNRVIGEGWNRAIGLHDPTAHA EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCTFFRMPRQVFNAQKKAQSSTD。

【0102】

いくつかの実施形態では、TadA*8は切り詰められている。いくつかの実施形態において、短縮型TadA*8は、全長TadA*8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いている。いくつかの実施形態において、短縮型TadA*8は、全長TadA*8と比較して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いている。いくつかの実施形態において、アデノシンデアミナーゼバリエーションは、完全長のTadA*8である。

20

【0103】

特定の実施形態において、アデノシンデアミナーゼヘテロダイマーは、TadA*8ドメインおよび以下のうちの1つから選択されるアデノシンデアミナーゼドメインを含む。

【0104】

特定の実施形態において、アデノシンデアミナーゼヘテロダイマーは、TadA*8ドメインおよび以下のうちの1つから選択されるアデノシンデアミナーゼドメインを含む。

30

【0105】

Escherichia coli TadA:

MRRAFITGVFFLSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAKKAQSS TD

【0106】

E. coli TadA (N末端切詰め):

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAKKAQSS TD

40

【0107】

Staphylococcus aureus (*S. aureus*) TadA:

MGSHMTNDIYFMTLAIIEAKKAAQLGEVPIGAIITKDDEVIARAHNLRETLQQPTAHAEHIAIERAAKVLGSRLEGCTLYVTLEPCVMCAGTIVMSRIPRVVYGADDPKGGCSGSLMNL LQQSNFNHRAIVDKGVLKEACSTLLTFFKNLRANKKSTN

【0108】

Bacillus subtilis (*B. subtilis*) TadA:

50

MTQDELYMKEAIKEAKKAEKGEVPIGAVLVINGEIIARAHNLRETEQRSIAHAEMLVIDE
ACKALGTWRLEGATLYVTLEPCPMCAGAVVLSRVEKVVFAGFDPKGGCSGTLMNLLQEE
RFNHQAEVVSGVLEEECGGMLSAFFRELRRKKKAARKNLSE

【0109】

Salmonella typhimurium (S. typhimurium) TadA:

MPPAFITGVTSLSDELDEHYWMRHALTLAKRAWDEREVPVGAVLVHNHRVIGEGWNR
PIGRHDPHTAHAEIMALRQGGLVLQNYRLLDTTLYVTLEPCVMCAGAMVHSRIGRVVFGA
RDAKTGAAGSLIDVLHHPGMNHRVEIEGVLRDECATLLSDFRMRRQEIKALKKADRAE
GAGPAV

【0110】

Shewanella putrefaciens (S. putrefaciens) TadA:

MDEYWMQVAMQMAEKAEAAAGEVPVGAVLVKDGQQTATGYNLSISQHDPTAHAEILCLR
SAGKKLENYRLLDATLYITLEPCAMCAGAMVHSRIARVVYGARDEKTGAAGTVVNLLQH
PAFNHQVEVTSGVLAEACSAQLSRFFKRRRDEKKALKLAQRAQQGIE

【0111】

Haemophilus influenzae F3031 (H. influenzae) TadA:

MDAAKVRSEFDEKMMRYALELADKAEALGEIPVGAVLVDDARNIIGEGWNLSIVQSDPT
AHAEIALRNGAKNIQNYRLLNSTLYVTLEPCTMCAGAILHSRIKRLVFGASDYKTGAIGS
RFHFFDDYKMNHTEITSGVLAEECSQLSTFFQKRREEKKIEKALLKSLSDK

【0112】

Caulobacter crescentus (C. crescentus) TadA:

MRTDESEDQDHRMMRLALDAARAAAEAGETPVGAVILDPSTGEVIATAGNGPIAAHDPT
AHAEIAAMRAAAAKLGNYRLTDLTLVVTLEPCAMCAGAISHARIGRVVFGADDPKGGAV
VHGPKFFAQPTCHWRPEVTGGVLADESADLLRGFFRARRKAKI

【0113】

Geobacter sulfurreducens (G. sulfurreducens) TadA:

MSSLKKTPIRDDAYWMGKAIREAAKAAARDEVPIGAVIVRDGAVIGRGNLREGSNDPSA
HAEMIAIRQAARRSANWRLTGATLYVTLEPCLMCMGAILARLERVVFGCYDPKGGGAAGS
LYDLSADPRLNHQVRLSPGVCQEECTMLSDFFRDLRRRKAKATPALFIDERKVPPEP

【0114】

TadA*7.10

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNRRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTD

【0115】

TadA*8を持つヘテロダイマーのコンポーネントとして企図される追加のTadA7.10ま
たはTadA7.10バリエーションとしては、次のものが挙げられる。

【0116】

GSSGSETPGTSESATPESSGSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNRRVI
GEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRI
RVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQ
KKAQSSTD

【0117】

TadA7.10 CP65

TAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGA
AGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTDGSSGS
ETPGTSESATPESSGSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNRRVIGEGW
NRAIGLHDP

【0118】

TadA7.10 CP83

10

20

30

40

50

YRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVE
ITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTDGSSGSETPGTSESATPESSGSEVEFS
HEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQ
GGLVMQN

【0119】

TadA7.10 CP136

MNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTDGSSGSETPGTSESATPESS
GSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAE
IMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLM
DVLHYPG

【0120】

TadA7.10 C-切り詰め

GSSGSETPGTSESATPESSGSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVI
GEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIG
RVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVFN

【0121】

TadA7.10 C-切り詰め2

GSSGSETPGTSESATPESSGSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVI
GEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIG
RVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQ

【0122】

TadA7.10 delta59-66+C-切り詰め

GSSGSETPGTSESATPESSGSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVI
GEGWNRAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVR
NAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVFN

【0123】

TadA7.10 delta 59-66

GSSGSETPGTSESATPESSGSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVI
GEGWNRAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVR
NAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSST
D.

【0124】

いくつかの実施形態において、アデノシンデアミナーゼバリエントは、TadA7.10にお
ける改変を含む。いくつかの実施形態において、TadA7.10は、アミノ酸82または166で
の改変を含む。特定の実施形態において、上記参照配列のバリエントは、以下の改変のう
ちの1つ以上を含む：Y147T、Y147R、Q154S、Y123H、V82S、T166R、およびQ1
54R。他の実施形態では、アデノシンデアミナーゼバリエントは、Y147R + Q154R + Y1
23H; Y147R + Q154R + I76Y; Y147R + Q154R + T166R; Y147T + Q154R; Y14
7T + Q154S; およびY123H + Y147R + Q154R + I76Yからなる群から選択される改変
の組み合わせを含む。

【0125】

他の実施形態において、本発明は、欠失を含むアデノシンデアミナーゼバリエント、例
えば、残基149、150、151、152、153、154、155、156、または157で始まるC末
端の欠失を含むTadA7.10を提供する。他の実施形態において、アデノシンデアミナーゼ
バリエントは、以下の改変のうちの1つ以上を含むTadAモノマーである：Y147T、Y147
R、Q154S、Y123H、V82S、T166R、Q154R。他の実施形態では、アデノシンデアミ
ナーゼバリエントは、以下の改変を含むモノマーである：Y147R + Q154R + Y123H; Y
147R + Q154R + I76Y; Y147R + Q154R + T166R; Y147T + Q154R; Y147T + Q1
54S; およびY123H + Y147R + Q154R + I76Y。さらに他の実施形態では、アデノシン
デアミナーゼバリエントは、それぞれが以下の改変Y147T、Y147R、Q154S、Y123H、

10

20

30

40

50

V82S、T166R、Q154Rのうちの1つ以上を有する2つのアデノシンデアミナーゼドメインを含むホモ二量体である。他の実施形態では、アデノシンデアミナーゼバリエントは、野生型アデノシンデアミナーゼドメインまたはTadA7.10ドメインと、以下の改変の1つまたは複数を含むアデノシンデアミナーゼバリエントドメインとを含むヘテロダイマーである：Y147T、Y147R、Q154S、Y123H、V82S、T166R、Q154R。他の実施形態では、アデノシンデアミナーゼバリエントは、TadA7.10ドメインと、以下の改変を含むTadA7.10のアデノシンデアミナーゼバリエントとを含むヘテロ二量体である。Y147R + Q154R + I76Y; Y147R + Q154R + T166R; Y147T + Q154R; Y147T + Q154S; および Y123H + Y147R + Q154R + I76Y。

【0126】

「投与すること」は、本明細書に記載される1以上の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、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。いくつかの実施形態において、非経口投与は、血管内、静脈内、筋肉内、動脈内、髄腔内、腫瘍内、皮内、腹腔内、経気管、皮下、角質下、関節内、莖膜内、くも膜下および胸骨内への注入または注射を含む。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。

【0127】

「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。

【0128】

「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構造、発現レベルまたは活性の変化（例えば増加または減少）を意味する。本明細書で使用される場合、改変は、ポリヌクレオチドもしくはポリペプチドの配列の変化、または発現レベルの変化、例えば10%の変化、25%の変化、40%の変化、50%の変化、またはそれ以上の発現レベルの変化を含む。

【0129】

「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、低減させる、停止させる、または安定させることを意味する。

【0130】

「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、ポリヌクレオチドまたはポリペプチドアナログは、対応する天然ポリヌクレオチドまたはポリペプチドの生物学的活性を保持しながら、天然ポリヌクレオチドまたは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修飾を有する。そのような修飾は、例えばリガンド結合を変化させることなく、アナログのDNA親和性、効率性、特異性、プロテアーゼもしくはヌクレアーゼ抵抗性、膜透過性、および／または半減期を増加させることができる。アナログは、非天然のポリヌクレオチドまたはアミノ酸を含み得る。

【0131】

「塩基エディター (BE)」あるいは「核酸塩基エディター (NBE)」とは、ポリヌクレオチドに結合して核酸塩基修飾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、核酸塩基修飾ポリペプチド（例えばデアミナーゼ）および核酸プログラミング可能なヌクレオチド結合ドメインを、ガイドポリヌクレオチド（例えばガイドRNA）と共に含む。様々な実施形態において、該薬剤は、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。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、デアミナーゼドメインに融合または連結される。1つの実施形態において、該薬剤は、塩基編集活性を有するドメ

10

20

30

40

50

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1つ以上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、アデノシン塩基エディター（ABE）である。

【0132】

「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。いくつかの実施形態において、シチジンデアミナーゼは、APOBECまたはAIDに対して少なくとも約85%の同一性を有する。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”）、または哺乳動物（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。

【0133】

いくつかの実施形態において、塩基エディターは、デアミナーゼ（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）に融合した再プログラム可能な塩基エディターである。いくつかの実施形態において、塩基エディターは、デアミナーゼ（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）に融合されたCas9である。いくつかの実施形態において、塩基エディターは、デアミナーゼ（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）に融合されたヌクレアーゼ不活Cas9（dCas9）である。いくつかの実施形態において、Cas9は、循環置換体（circular permutant）のCas9（例えばspCas9またはsaCas9）である。循環置換体Cas9は当該分野で公知であり、例えば、Oakes et al., Cell 176, 254-267, 2019に記載されている。いくつかの実施形態において、塩基エディターは、塩基除去修復の阻害因子、例えばUGIドメインまたはdISNドメインに融合される。ある実施形態において、融合タンパク質は、デアミナーゼと、UGIドメインまたはdISNドメインのような塩基除去修復の阻害因子に融合されたCas9ニックラーゼを含む。他の実施形態では、塩基エディターは脱塩基型（abasic）塩基エディターである。

【0134】

いくつかの実施形態では、塩基エディターは、アデノシン塩基エディター（ABE）である。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から進化されたものである。いくつかの実施形態において、本発明の塩基エディターは、内部的に融合された触媒（例えば、デアミナーゼ）ドメインを有するnapDNAbpドメインを含む。いくつかの実施形態において、napDNAbpは、内部的に融合されたデアミナーゼドメインを有するCas12a（Cpf1）である。いくつかの実施形態において、napDNAbpは、内部的に融合されたデアミナーゼドメインを有するCas12b（c2c1）である。いくつかの実施形態において、napDNAbpは、内部的に融合されたデアミナーゼドメインを有するCas12c（c2c3）である。いくつかの実施形態において、napDNAbpは、内部的に融合されたデアミナーゼドメインを有するCas12d（CasX）である。いくつかの実施形態において、napDNAbpは、内部的に融合されたデアミナーゼドメインを有するCas12e（CasY）である。いくつかの実施形態において、napDNAbpは、内部的に融合されたデアミナーゼドメインを有するCas12gである。いくつかの実施形態において、napDNAbpは、内部的に融合されたデアミナーゼドメインを有するCas12hである。いくつかの実施形態において、napDNAbpは、内部的に融合されたデアミナーゼドメインを有するCas12iである。いくつかの実施形態において、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死んだCas12（dCas12）である。いくつかの実施形態において、塩基エディターは、デアミナーゼドメイ

10

20

30

40

50

ンに融合されたCas12ニッカーゼ（nCas12）である。

【0135】

いくつかの実施形態において、塩基エディターは、アデノシンデアミナーゼバリエーション（例えばTadA*8）を、循環置換体Cas9（例えばspCAS9またはsaCAS9）および二部分核局在化配列を含む足場にクローニングすることによって生成される（例えばABE8）。循環置換体Cas9は当技術分野で公知であり、例えばOakes et al., Cell 176, 254-267, 2019に記載されている。例示的な循環置換体を以下に記載するが、ここで太字の配列はCas9に由来する配列を示し、斜体の配列はリンカー配列を示し、下線が引かれた配列は二部分核局在化配列を示す。

【0136】

CP5 (with MSP “NGC=Pam Variant with mutations Regular Cas9 likes NGG”
PID=Protein Interacting Domain and “D10A” nickase):

EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPMPQ
VNIVKKTEVQTTGGFSKESILPKRNSDKLIARKKDWDPPKKGGMQPTVAYSVLVVAKEVEKSKK
KLKSVKELLGITIMERSSEFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELNGRKRMLASAKF
LQKGNELALPSKYVNFYLYASHYEKLKGSPEDEQKQLFVEQHKHYLDEIIEQISEFSKRVILA
DANLDKVL SAYNKHRDKPIREQAENIIHLFTLTNLGAPRAFKYFDTTIARKEYRSTKEVLDATL
IHQSITGLYETRIDLSQLGGDGGSGGSGGSGGSGGSGGSGGMDKKYSIGLAIGTNSVGWAVITD
EYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRKNRICYLQEIFS
NEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADL
RLIYLAHAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEEENPINASGVDAKAILSAR
LSKSRLENLIAQLPGKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLL
AQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQL
PEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN
GSIPHQIHLGELHAILRRQEDFYFPFLKDNREKIEKILTFRIPIYYVGPLARGNSRFAWMTRKSEE
TITPWNFEVVDKGASAQSFIERMTNFDKNLPNEKVLPHSLLYEYFTVYNELTKVKYVTEGMR
KPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLK
IIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLS
RKLINGIRDKQSGKTILDFLKSDGFANRNFMLIHDDSLTFKEDIQAQVSGQGDSLHEHIANL
AGSPAIKKGILQTVKVVDLVKVMGRHKPENIVIEARENQTTQKGQKNSRERMKRIIEGKEL
GSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNK
VLTRSDKNRGKSDNVPSEEVVKMKKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKR
QLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHH
AHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVVKMIKSEQEGADKRTADGSEFESPKKKR
KV*

【0137】

いくつかの実施形態において、ABE8は、下記表6-9、13、または14からの塩基エディターから選択される。いくつかの実施形態において、ABE8は、TadAから進化されたアデ

ノシンデアミナーゼバリエーションを含む。いくつかの実施形態において、ABE8のアデノシンデアミナーゼバリエーションは、下記表7、9、13または14に記載されるようなTadA*8バリエーションである。いくつかの実施形態において、アデノシンデアミナーゼバリエーションは、Y147T、Y147R、Q154S、Y123H、V82S、T166R、および/またはQ154Rの群から選択される1つ以上の改変を含むTadA*7.10バリエーション（例えば、TadA*8）である。様々な実施形態において、ABE8は、Y147T + Q154R; Y147T + Q154S; Y147R + Q154S; V82S + Q154S; V82S + Y147R; V82S + Q154R; V82S + Y123H; I76Y + V82S; V82S + Y123H + Y147T; V82S + Y123H + Y147R; V82S + Y123H + Q154R; Y147R + Q154R + Y123H; Y147R + Q154R + I76Y; Y147R + Q154R + T166R; Y123H + Y147R + Q154R + I76Y; V82S + Y123H + Y147R + Q154R; およびI76Y + V82S + Y123H + Y147R + Q154Rからなる群から選択される改変の組み合わせを有するTadA*7.10（例えばTadA*8）を含む。いくつかの実施形態では、ABE8は単量体構築物である。いくつかの実施形態では、ABE8はヘテロ二量体構築物である。いくつかの実施形態では、ABE8塩基エディターは、以下の配列を含む：

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNRRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNRHVEITEGILADECAALLCTFFRMPRQVFNAQKKAQSSTD。

【0138】

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。

【0139】

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:eaa04774 (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）。

【0140】

例として、本明細書に記載の塩基編集組成物、システム、および方法において使用されるシチジン塩基エディターは、下記の核酸配列（8877塩基対）（Addgene, Watertown, MA.; Komor AC, et al., 2017, Sci Adv., 30;3(8):eaa04774. doi: 10.1126/sciadv.aao4774）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。

【0141】

1 atatgcccaag tacgccccct attgacgtca atgacggtaa atggcccgcc tggcattatg
61 cccagtacat gaccttatgg gactttccta cttggcagta catctacgta ttagtcatcg
121 ctattacat ggtgatgcgg ttttggcagt acatcaatgg gcgtggatag cggtttgact

181 cacggggatt tccaagtctc caccctattg acgtcaatgg gagtttgttt tggcaccaaaa
 241 atcaacggga ctttccaaaa tgtcgttaaca actccgcccc attgacgcaa atgggcggtg
 301 ggcgtgtacg gtgggaggtc tatataagca gagctggttt agtgaaccgt cagatccgct
 361 agagatccgc ggccgctaatac agactcact atagggagag ccgccaccat gagctcagag
 421 actggcccag tggctgtgga cccacattg agacggcgga tcgagcccca tgagtttgag
 481 gtattcttcg atccgagaga gctccgcaag gagacctgcc tgctttacga aattaattgg
 541 gggggccggc actccatttg gcgacataca tcacagaaca ctaacaagca cgtcgaagtc
 601 aacttcatcg agaagttcac gacagaaaga tatttctgtc cgaacacaag gtgcagcatt
 661 acctggtttc tcagctggag cccatgcggc gaatgtagta gggccatcac tgaattcctg
 721 tcaaggtatc cccacgtcac tctgtttatt tacatcgcaa ggctgtacca ccacgtgac
 781 ccccgaatc gacaaggcct gcgggatttg atctcttcag gtgtgactat ccaaattatg
 841 actgagcagg agtcaggata ctgctggaga aactttgtga attatagccc gagtaatgaa
 901 gccactggc ctaggtatcc ccatctgtgg gtacgactgt acgttcttga actgtactgc
 961 atcatactgg gcctgcctcc ttgtctcaac attctgagaa ggaagcagcc acagctgaca
 1021 tttcttacca tcgctcttca gtcttctcat taccagcgac tgccccaca cattctctgg
 1081 gccaccgggt tgaaatctgg tggttcttct ggtggttcta gcggcagcga gactcccggg
 1141 acctcagagt ccgccacacc cgaaagtctt ggtggttctt ctggtgggtc tgataaaaag
 1201 tattctattg gtttagccat cggcactaat tccgttggat gggctgtcat aaccgatgaa
 1261 tacaaggtac cttcaaagaa atttaagggtg ttggggaaca cagaccgtca ttcgattaaa
 1321 aagaatctta tcggtgccct cctattcgat agtggcgaaa cggcagaggc gactcgcctg
 1381 aaacgaaccg ctcggagaag gtatacacgt cgcaagaacc gaatatgtta cttacaagaa
 1441 atttttagca atgagatggc caaagttgac gattctttct ttcaccgttt ggaagagtcc
 1501 ttccttgtcg aagaggacaa gaaacatgaa cggcacccca tctttggaag catagtagat
 1561 gaggtggcat atcatgaaaa gtacccaacg atttatcacc tcagaaaaaa gctagttgac
 1621 tcaactgata aagcggacct gaggttaatc tacttggctc ttgccatat gataaagttc
 1681 cgtgggcact ttctcattga gggatgatcta aatccggaca actcggatgt cgacaaactg
 1741 ttcattccagt tagtacaac ctataatcag ttgtttgaag agaaccctat aaatgcaagt
 1801 ggcgtggatg cgaaggctat tcttagcgcc cgcctctcta aatcccgacg gctagaaaaa
 1861 ctgatcgcac aattaccgag agagaagaaa aatgggttgt tcggtaacct tatagcgctc
 1921 tctactaggc tgacacaaaa ttttaagtcg aacttcgact tagctgaaga tgccaaattg
 1981 cagcttagta aggacacgta cgatgacgat ctgcacaatc tactggcaca aattggagat
 2041 cagtatcgcg acttattttt ggctgcaaaa aaccttagcg atgcaatcct cctatctgac
 2101 atactgagag ttaatactga gattaccaag gcgccgttat ccgcttcaat gatcaaaagg
 2161 tacgatgaac atcaccaaga cttgacactt ctcaaggccc tagtccgtca gcaactgcct
 2221 gagaaatata aggaaatatt cttgatcag tcgaaaaacg ggtacgcagg ttatattgac
 2281 ggcggagcga gtcaagagga attctacaag ttatcaaac ccatattaga gaagatggat
 2341 gggacggaag agttgcttgt aaaactcaat cgcgaagatc tactgcgaaa gcagcggact
 2401 ttcgacaacg gtagcattcc acatcaaata cacttaggcg aattgcatgc tatacttaga
 2461 aggcaggagg atttttatcc gttctcaaaa gacaatcgtg aaaagattga gaaaatccta
 2521 acctttcgca tactttacta tgtgggaccc ctggcccgag ggaactctcg gttcgcattg
 2581 atgacaagaa agtccgaaga aacgattact ccatggaatt ttgaggaagt tgtcgataaa
 2641 ggtgcgtcag ctcaatcgtt catcgagagg atgaccaact ttgacaagaa ttaccgaac
 2701 gaaaaagtat tgcctaagca cagtttactt tacgagtatt tcacagtgt caatgaactc
 2761 acgaaaagta agtatgtcac tgagggcatt cgtaaaccgg cttttctaag cggagaacag
 2821 aagaaaagcaa tagtagatct gttattcaag accaaccgca aagtgcagat taagcaattg
 2881 aaagaggact actttaagaa aattgaatgc ttcgattctg tcgagatctc cggggtagaa
 2941 gatcgattta atgcgtcact tggtagctat catgacctcc taaagataat taaagataag
 3001 gacttcctgg ataacgaaga gaatgaagat atcttagaag atatagtgtt gactcttacc
 3061 ctctttgaag atcgggaaat gattgaggaa agactaaaaa catacgtca cctgttcgac
 3121 gataagggtta tgaaacagtt aaagaggcgt cgctatacgg gctggggacg attgtcgcgg

10

20

30

40

50

3181 aaacttatca acgggataag agacaagcaa agtggtaaaa ctattctcga ttttctaaag
 3241 agcgacggct tcgccaatag gaactttatg cagctgatcc atgatgactc tttaaccttc
 3301 aaagaggata taaaaaaggc acagggttcc ggacaagggg actcattgca cgaacatatt
 3361 gcgaatcttg ctggttcgcc agccatcaaa aagggcatac tccagacagt caaagtagtg
 3421 gatgagctag ttaagggtcat gggacgtcac aaaccggaaa acattgtaat cgagatggca
 3481 cgcgaaaatc aaacgactca gaaggggcaa aaaaacagtc gagagcggat gaagagaata
 3541 gaagagggta ttaaagaact gggcagccag atcttaaagg agcatcctgt ggaaaatacc
 3601 caattgcaga acgagaaact ttacctctat tacctacaaa atggaaggga catgtatgtt
 3661 gatcaggaac tggacataaa ccgtttatct gattacgacg tcgatcacat tgtaccccaa
 3721 tcctttttga aggacgattc aatcgacaat aaagtgttta cacgctcgga taagaaccga
 3781 gggaaaagtg acaatgttcc aagcgaggaa gtcgtaaaga aaatgaaga ctattggcgg
 3841 cagctcctaa atgcgaaact gataacgcaa agaaagtgcg ataacttaac taaagctgag
 3901 aggggtggct tgtctgaact tgacaaggcc ggatttatta aacgtcagct cgtggaaacc
 3961 cgccaaatca caaagcatgt tgcacagata ctagattccc gaatgaatac gaaatacgac
 4021 gagaacgata agctgattcg ggaagtcaaa gtaatcactt taaagtcaaa attggtgtcg
 4081 gacttcagaa aggattttca attctataaa gttagggaga taaataacta ccacatgcg
 4141 cacgacgctt atcttaatgc cgtcgtaggg accgcactca ttaagaaata cccgaagcta
 4201 gaaagtgagt ttgtgtatgg tgattacaaa gtttatgacg tccgtaagat gatcgcgaaa
 4261 agcgaacagg agataggcaa ggctacagcc aaatacttct tttattctaa cattatgaat
 4321 tttttaaga cggaaatcac tctggcaaac ggagagatac gcaaacgacc ttaattgaa
 4381 accaatgggg agacaggtga aatcgtatgg gataagggcc gggacttcgc gacggtgaga
 4441 aaagttttgt ccatgcccc aagtcaacata gtaaagaaaa ctgaggtgca gaccggaggg
 4501 ttttcaaagg aatcgattct tccaaaaagg aatagtgata agctcatcgc tcgtaaaaag
 4561 gactgggacc cgaaaaagta cgggtggcttc gatagcccta cagttgccta ttctgtccta
 4621 gtagtgagaa aagttgagaa gggaaaatcc aagaaactga agtcagtcaa agaattattg
 4681 gggataacga ttatggagcg ctctgttttt gaaaagaacc ccatcgactt ccttgaggcg
 4741 aaaggttaca aggaagtaaa aaaggatctc ataattaaac taccaaagta tagtctgttt
 4801 gagttagaaa atggccgaaa acggatgttg gctagcgccg gagagcttca aaaggggaac
 4861 gaactcgcac taccgtctaa atacgtgaat ttctgtatt tagcgtcca ttacgagaag
 4921 ttgaaagggt cacctgaaga taacgaacag aagcaacttt ttgttgagca gcacaaacat
 4981 tatctcgacg aaatcataga gcaaatttcg gaattcagta agagagtcac cctagctgat
 5041 gccaatctgg acaaagtatt aagcgcatat aacaagcaca gggataaacc catacgtgag
 5101 caggcggaag atattatcca ttgtttact ctaccaacc tcggcgctcc agccgcattc
 5161 aagtattttg acacaacgat agatcgcaaa cgatacactt ctaccaagga ggtgctagac
 5221 gcgacactga ttcaccaatc catcacggga ttatatgaaa ctcgataga tttgtcacag
 5281 cttgggggtg actctggttg ttctggagga tctggtggtt ctactaatct gtcagatatt
 5341 attgaaaagg agaccggtaa gcaactggtt atccaggaat ccatcctcat gctcccagag
 5401 gaggtggaag aagtcattgg gaacaagccg gaaagcgata tactcgtgca caccgcctac
 5461 gacgagagca ccgacgagaa tgtcatgctt ctgactagcg acgcccctga atacaagcct
 5521 tgggctctgg tcatagagga tagcaacggg gagaacaaga ttaagatgct ctctggtggt
 5581 tctggaggat ctggtggttc tactaatctg tcagatatta ttgaaaagga gaccggtgag
 5641 caactgggta tccaggaatc catcctcatg ctcccagagg aggtggaaga agtcattggg
 5701 aacaagccgg aaagcgatat actcgtgcac accgcctacg acgagagcac cgacgagaat
 5761 gtcattgctt tgactagcga cgcccctgaa tacaagcctt gggctctggt catacaggat
 5821 agcaacgggt agaacaagat taagatgctc tctggtggtt ctccaagaa gaagaggaaa
 5881 gtctaaccgg tcatcatcac catcaccatt gagtttaaac ccgctgatca gcctcgactg
 5941 tgccttctag ttgccagcca tctgtgttt gccctcccc cgtgccttcc ttgaccctgg
 6001 aaggtgccac tcccactgtc ctttctaat aaaatgagga aattgcatcg cattgtctga
 6061 gtaggtgtca ttctattctg ggggggtggg tggggcgagga cagcaagggg gaggattggg
 6121 aagacaatag caggcatgct ggggatgcgg tgggctctat ggcttctgag gcggaaagaa

10

20

30

40

50

6181 ccagctgggg ctgataccg tcgacctcta gctagagctt ggcgtaatca tggcatagc
 6241 tgtttcctgt gtgaaattgt tatccgctca caattccaca caacatacga gccggaagca
 6301 taaagtgtaa agcctagggt gcctaagag tgagctaact cacattaatt gcgttgcgct
 6361 cactgccccg tttcagtcg ggaaacctgt cgtgccagct gcattaatga atcgccaac
 6421 ggcgggggag aggcggtttg cgtattgggc gctctccgc ttcctcgctc actgactcg
 6481 tgcgctcggg cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacgg
 6541 tatccacaga atcaggggat aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg
 6601 ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca taggctccgc cccctgacg
 6661 agcatcaca aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat
 6721 accaggcgtt tccccctgga agctccctcg tgcgctctcc tgttcgacc ctgccgtta
 6781 ccgataacct gtccgccttt ctcccttcgg gaagcgtggc gctttctcat agctcacgct
 6841 gtaggtatct cagttcgggt taggtcgttc gctccaagct gggctgtgtg cacgaacccc
 6901 ccgttcagcc cgaccgtgc gccttatccg gtaactatcg tcttgagtc aaccggtaa
 6961 gacacgactt atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg
 7021 taggcgggtg tacagagttc ttgaagtggg ggcctaacta cggctacact agaagaacag
 7081 tatttggtat ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt
 7141 gatccggcaa acaaacacc gctggtagcg gtggttttt tgtttgcaag cagcagatta
 7201 cgcgagaaa aaaaggatct caagaagatc cttgatctt ttctacgggg tctgacgctc
 7261 agtggaaacg aaactcacgt taagggtatt tggcatgag attatcaaaa aggatcttca
 7321 cctagatcct tttaaattaa aaatgaagtt ttaaatcaat cttaaagtata tatgagtaaa
 7381 cttggtctga cagttacaa tgcttaatca gtgaggcacc tatctcagcg atctgtctat
 7441 ttcgttcac catagttgcc tgactccccg tcgtgtagat aactacgata cgggagggt
 7501 taccatctgg cccagtgct gcaatgatac cgcgagacc acgctcaccg gctccagatt
 7561 tatcagcaat aaaccagcca gccggaagg cggagcgag aagtggctc gcaactttat
 7621 ccgctccat ccagctctat aattgttgc gggaagctag agtaagtagt tcgccagtta
 7681 atagtttgc caacgttgtt gccattgcta caggcatcgt ggtgtcacgc tcgtcgtttg
 7741 gtatggcttc attcagctcc ggttcccaac gatcaaggcg agttacatga tccccatgt
 7801 tgtgcaaaaa agcggtagc tccttcggtc ctccgatcgt tgtcagaagt aagtggccg
 7861 cagtgttatc actcatggtt atggcagcac tgcataattc tcttactgtc atgccatccg
 7921 taagatgctt tctgtgact ggtgagtact caaccaagtc attctgagaa tagtgatgc
 7981 ggcgaccgag ttgctcttgc ccggcgtcaa tacgggataa taccgcgcca catagcagaa
 8041 cttaaaagt gctcatcatt ggaaaacgtt ctcggggcg aaaactctca aggatcttac
 8101 cgctgttgag atccagttcg atgtaacca ctctgcacc caactgatct tcagcatctt
 8161 ttactttcac cagcgtttct ggtgagcaa aaacaggaag gcaaatgcc gcaaaaaagg
 8221 gaataagggc gacacggaaa tgttgaatac tcatactctt ctttttcaa tattattgaa
 8281 gcatttatca ggttattgt ccatgagcg gatacatatt tgaatgtatt tagaaaaata
 8341 aacaaatagg ggttcgcgc acatttccc gaaaagtgc acctgacgtc gacggatcgg
 8401 gagatcgatc tccgatccc ctagggtcga ctctcagtac aatctgctt gatgccgat
 8461 agttaagcca gtatctgctc cctgcttggt tgttgagggt cgctgagtag tgcgagca
 8521 aaatttaagc tacaacaagg caaggcttga ccgacaattg catgaagaat ctgcttaggg
 8581 ttaggcgttt tgcgctgctt cgcatgtac gggccagata tacgcgttga cattgattat
 8641 tgactagtta ttaatagtaa tcaattacgg ggtcattagt tcatagccca tatatggagt
 8701 tccgcgttac ataactacg gtaaatggc cgctggctg accgccaac gacccccgc
 8761 cattgacgtc aataatgac tatgttcca tagtaacgcc aatagggact ttccattgac
 8821 gtcaatgggt ggagtattta cggtaaactg cccacttggc agtacatcaa gtgtatc

[0 1 4 2]

BE4アミノ酸配列：

MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELKTCCLLYEINWGGRHSIWRHTSQNTNK
 HVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSRAITEFLSRYPHVTFLFIYIARLYHH
 ADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFVNYSNEAHWPYPHLWVRLYVLE

10

20

30

40

50

LYCIILGLPPCLNILRRKQPQLTFFTIALQSCHYQRLPPHILWATGLKSGGSSGGSSGSETP
GTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIK
KNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESF
LVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYALAHMIKFRG
HFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRENLIA
QLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDITYDDDLNLLAQIGDQYA
DLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYK
EIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGS
IPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSE
ETITPWNFEEVVDKGASASQSFIERMTNFDKNLPNEKVLPHKSLLEYEFTVYNELTKVKYV
TEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASL
GTYHDLKKIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMKQLK
RRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQ
VSGQGDSLHEHIANLAGSPAIKKGILQTVKVDELVKVMGRHKPENIVIEMARENQTTQK
GQKNSRERMKRIEIGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINR
LSDYDVDHIVPQSFLKDDSIDNKVLTSDKNRGKSDNVPSEEVKKMKNYWRQLLNAKL
ITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIRE
VKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGD
YKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIV
WDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYG
GFDSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKD
LIIKLPKYSLELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLGSPEDNEQ
KQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLFTLT
NLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGDSGGSGGGSGGS
TNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLTSDAP
EYKPWALVIQDSNGENKIKMLSGSGGGSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEE
VIGNKPESDILVHTAYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSPKK
KRK

【0143】

例として、本明細書に記載の塩基編集組成物、システムおよび方法で使用されるアデニ
ン塩基エディター（ABE）は、下記に提供される核酸配列（8877塩基対）を有する（Add
gene, Watertown, MA.; Gaudelli NM, et al., Nature. 2017 Nov 23;551(7681):
464-471. doi: 10.1038/nature24644; Koblan LW, et al., Nat Biotechnol. 201
8 Oct;36(9):843-846. doi: 10.1038/nbt.4172.）。ABE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。

【0144】

ATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTAT
GCCCAGTACAT
GACCTTATGGGACTTTCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCA
TGGTGATGCGG
TTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCT
CCACCCCATTTG
ACGTCAATGGGAGTTTGTGTTTGGCACCAAATCAACGGGACTTTCCAAAATGTCGTAAC
AACTCCGCCCC
ATTGACGCAAATGGGCGGTAGGCGTGACGGTGGGAGGTCTATATAAGCAGAGCTGGTT
TAGTGAACCGT
CAGATCCGCTAGAGATCCGCGGCCGCTAATACGACTCACTATAGGGAGAGCCGCCACCA
TGAAACGGACA
GCCGACGGAAGCGAGTTCGAGTCACCAAAGAAGAAGCGGAAAGTCTCTGAAGTCGAGTT
TAGCCACGAGT

ATTGGATGAGGCACGCACTGACCCTGGCAAAGCGAGCATGGGATGAAAGAGAAGTCCCC
GTGGGCGCCGT
GCTGGTGCACAACAATAGAGTGATCGGAGAGGGATGGAACAGGCCAATCGGCCGCCACG
ACCTACCGCA
CACGCAGAGATCATGGCACTGAGGCAGGGAGGCCTGGTCATGCAGAATTACCGCCTGAT
CGATGCCACCC
TGTATGTGACACTGGAGCCATGCGTGATGTGCGCAGGAGCAATGATCCACAGCAGGATC
GGAAGAGTGGT
GTTCGGAGCACGGGACGCCAAGACCGGCGCAGCAGGCTCCCTGATGGATGTGCTGCACC
ACCCCGGCATG
AACCACCGGGTGGAGATCACAGAGGGAATCCTGGCAGACGAGTGCGCCGCCCTGCTGAG
CGATTTCTTTA
GAATGCGGAGACAGGAGATCAAGGCCCAGAAGAAGGCACAGAGCTCCACCGACTCTGGA
GGATCTAGCGG
AGGATCCTCTGGAAGCGAGACACCAGGCACAAGCGAGTCCGCCACACCAGAGAGCTCCG
GCGGCTCCTCC
GGAGGATCCTCTGAGGTGGAGTTTTCCACGAGTACTGGATGAGACATGCCCTGACCCT
GGCCAAGAGGG
CACGCGATGAGAGGGAGGTGCCTGTGGGAGCCGTGCTGGTGCTGAACAATAGAGTGATC
GGCGAGGGCTG
GAACAGAGCCATCGGCCTGCACGACCCAACAGCCCATGCCGAAATTATGGCCCTGAGAC
AGGGCGGCCTG
GTCATGCAGAACTACAGACTGATTGACGCCACCCTGTACGTGACATTCGAGCCTTGCGT
GATGTGCGCCG
GCGCCATGATCCACTCTAGGATCGGCCGCGTGGTGTTTGGCGTGAGGAACGCAAAAACC
GGCGCCGCAGG
CTCCCTGATGGACGTGCTGCACTACCCCGGCATGAATCACCGCGTCGAAATTACCGAGG
GAATCCTGGCA
GATGAATGTGCCGCCCTGCTGTGCTATTTCTTTCGGATGCCTAGACAGGTGTTCAATGC
TCAGAAGAAGG
CCCAGAGCTCCACCGACTCCGGAGGATCTAGCGGAGGCTCCTCTGGCTCTGAGACACCT
GGCACAAGCGA
GAGCGAACACCTGAAAGCAGCGGGGGCAGCAGCGGGGGGTCAGACAAGAAGTACAGC
ATCGGCCTGGCC
ATCGGCACCAACTCTGTGGGCTGGGCCGTGATCACCGACGAGTACAAGGTGCCCAGCAA
GAAATTCAAGG
TGCTGGGCAACACCGACCGGCACAGCATCAAGAAGAACCTGATCGGAGCCCTGCTGTTC
GACAGCGGCGA
AACAGCCGAGGCCACCCGGCTGAAGAGAACCGCCAGAAGAAGATACACCAGACGGAAG
AACCGGATCTGC
TATCTGCAAGAGATCTTCAGCAACGAGATGGCCAAGGTGGACGACAGCTTCTTCCACAG
ACTGGAAGAGT
CCTTCCTGGTGGAAAGAGGATAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTG
GACGAGGTGGC
CTACCACGAGAAGTACCCACCATCTACCACCTGAGAAAGAACTGGTGGACAGCACCG
ACAAGGCCGAC
CTGCGGCTGATCTATCTGGCCCTGGCCCACATGATCAAGTTCCGGGGCCACTTCTGAT
CGAGGGCGACC
TGAACCCCGACAACAGCGACGTGGACAAGCTGTTTCATCCAGCTGGTGCAGACCTACAAC
CAGCTGTTCTGA

10

20

30

40

50

GGAAAACCCCATCAACGCCAGCGGCGTGACGCCAAGGCCATCCTGTCTGCCAGACTGA
GCAAGAGCAGA
CGGCTGGAAAATCTGATCGCCCAGCTGCCCGGCGAGAAGAAGAATGGCCTGTTCGGAAA
CCTGATTGCC
TGAGCCTGGGCTGACCCCCAACTTCAAGAGCAACTTCGACCTGGCCGAGGATGCCAAA
CTGCAGCTGAG
CAAGGACACCTACGACGACGACCTGGACAACCTGCTGGCCCAGATCGGCGACCAGTACG
CCGACCTGTTT
CTGGCCGCCAAGAACCTGTCCGACGCCATCCTGCTGAGCGACATCCTGAGAGTGAACAC
CGAGATCACCA
AGGCCCCCTGAGCGCCTCTATGATCAAGAGATACGACGAGCACCACCAGGACCTGACC
CTGCTGAAAGC
TCTCGTGCGGCAGCAGCTGCCTGAGAAGTACAAAGAGATTTTCTTCGACCAGAGCAAGA
ACGGCTACGCC
GGCTACATTGACGGCGGAGCCAGCCAGGAAGAGTTCTACAAGTTCATCAAGCCCATCCT
GGAAAAGATGG
ACGGCACCGAGGAACTGCTCGTGAAGCTGAACAGAGAGGACCTGCTGCGGAAGCAGCGG
ACCTTCGACAA
CGGCAGCATCCCCACCAGATCCACCTGGGAGAGCTGCACGCCATTCTGCGGCGGCAGG
AAGATTTTTAC
CCATTCCTGAAGGACAACCGGGAAAAGATCGAGAAGATCCTGACCTTCGCATCCCCTA
CTACGTGGGCC
CTCTGGCCAGGGGAAACAGCAGATTGCTGCTGGATGACCAGAAAGAGCGAGGAAACCATC
ACCCCCTGGAA
CTTCGAGGAAGTGGTGGACAAGGGCGCTTCGCCCCAGAGCTTCATCGAGCGGATGACCA
ACTTCGATAAG
AACCTGCCCAACGAGAAGGTGCTGCCCAAGCACAGCCTGCTGTACGAGTACTTCACCGT
GTATAACGAGC
TGACCAAAGTGAAATACGTGACCGAGGGAATGAGAAAGCCCGCCTTCCTGAGCGGCGAG
CAGAAAAAGGC
CATCGTGGACCTGCTGTTCAAGACCAACCGGAAAGTGACCGTGAAGCAGCTGAAAGAGG
ACTACTTCAAG
AAAATCGAGTGCTTCGACTCCGTGGAAATCTCCGGCGTGGAAGATCGGTTCAACGCCTC
CCTGGGCACAT
ACCACGATCTGCTGAAAATTATCAAGGACAAGGACTTCCTGGACAATGAGGAAAACGAG
GACATTCTGGA
AGATATCGTGCTGACCCTGACACTGTTTGAGGACAGAGAGATGATCGAGGAACGGCTGA
AAACCTATGCC
CACCTGTTTCGACGACAAAGTGATGAAGCAGCTGAAGCGGCGGAGATACACCGGCTGGGG
CAGGCTGAGCC
GGAAGCTGATCAACGGCATCCGGGACAAGCAGTCCGGCAAGACAATCCTGGATTTCTG
AAGTCCGACGG
CTTCGCCAACAGAACTTCATGCAGCTGATCCACGACGACAGCCTGACCTTTAAAGAGG
ACATCCAGAAA
GCCCAGGTGTCCGGCCAGGGCGATAGCCTGCACGAGCACATTGCCAATCTGGCCGGCAG
CCCCGCCATTA
AGAAGGGCATCCTGCAGACAGTGAAGGTGGTGGACGAGCTCGTGAAAGTGATGGGCCGG
CACAAGCCCGA
GAACATCGTGATCGAAATGGCCAGAGAGAACCAGACCACCAGAAAGGGACAGAAGAAC
AGCCGCGAGAGA

10

20

30

40

50

ATGAAGCGGATCGAAGAGGGCATCAAAGAGCTGGGCAGCCAGATCCTGAAAGAACACCC
CGTGAAAAACA
CCCAGCTGCAGAACGAGAAGCTGTACCTGTACTACCTGCAGAATGGGCGGGATATGTAC
GTGGACCAGGA
ACTGGACATCAACCGGCTGTCCGACTACGATGTGGACCATATCGTGCCTCAGAGCTTTC
TGAAGGACGAC
TCCATCGACAACAAGGTGCTGACCAGAAGCGACAAGAACCGGGGCAAGAGCGACAACGT
GCCCTCCGAAG
AGGTCGTGAAGAAGATGAAGAACTACTGGCGGCAGCTGCTGAACGCCAAGCTGATTACC
CAGAGAAAGTT
CGACAATCTGACCAAGGCCGAGAGAGGGCGGCCTGAGCGAACTGGATAAGGCCGGCTTCA
TCAAGAGACAG
CTGGTGGAAACCCGGCAGATCACAAAGCACGTGGCACAGATCCTGGACTCCCGGATGAA
CACTAAGTACG
ACGAGAATGACAAGCTGATCCGGGAAGTGAAAGTGATCACCTGAAGTCCAAGCTGGTG
TCCGATTTCCG
GAAGGATTTCCAGTTTTTACAAAGTGCGCGAGATCAACAACTACCACCACGCCCACGACG
CCTACCTGAAC
GCCGTCGTGGGAACCGCCCTGATCAAAAAGTACCCTAAGCTGGAAAGCGAGTTCGTGTA
CGGCGACTACA
AGGTGTACGACGTGCGGAAGATGATCGCCAAGAGCGAGCAGGAAATCGGCAAGGCTACC
GCCAAGTACTT
CTTCTACAGCAACATCATGAACTTTTTCAAGACCGAGATTACCCTGGCCAACGGCGAGA
TCCGGAAGCGG
CCTCTGATCGAGACAAACGGCGAAACCGGGGAGATCGTGTGGGATAAGGGCCGGGATTT
TGCCACCGTGC
GGAAAGTGCTGAGCATGCCCCAAGTGAATATCGTGAAAAAGACCGAGGTGCAGACAGGC
GGCTTCAGCAA
AGAGTCTATCCTGCCCAAGAGGAACAGCGATAAGCTGATCGCCAGAAAGAAGGACTGGG
ACCCTAAGAAG
TACGGCGGCTTCGACAGCCCCACCGTGGCCTATTCTGTGCTGGTGGTGGCCAAAGTGGA
AAAGGGCAAGT
CCAAGAACTGAAGAGTGTGAAAGAGCTGCTGGGGATCACCATCATGGAAAGAAGCAGC
TTCGAGAAGAA
TCCCATCGACTTTCTGGAAGCCAAGGGCTACAAAGAAGTGAAAAAGGACCTGATCATCA
AGCTGCCTAAG
TACTCCCTGTTTCGAGCTGGAAAACGGCCGGAAGAGAATGCTGGCCTCTGCCGGCGAACT
GCAGAAGGGAA
ACGAACTGGCCCTGCCCTCCAAATATGTGAACTTCCTGTACCTGGCCAGCCACTATGAG
AAGCTGAAGGG
CTCCCCGAGGATAATGAGCAGAAACAGCTGTTTGTGGAACAGCACAAGCACTACCTGG
ACGAGATCATC
GAGCAGATCAGCGAGTTCTCCAAGAGAGTGATCCTGGCCGACGCTAATCTGGACAAAGT
GCTGTCCGCCT
ACAACAAGCACCGGGATAAGCCCATCAGAGAGCAGGCCGAGAATATCATCCACCTGTTT
ACCCTGACCAA
TCTGGGAGCCCCTGCCGCCTTCAAGTACTTTGACACCACCATCGACCGGAAGAGGTACA
CCAGCACCAA
GAGGTGCTGGACGCCACCCTGATCCACCAGAGCATCACCGGCCTGTACGAGACACGGAT
CGACCTGTCTC

10

20

30

40

50

AGCTGGGAGGTGACTCTGGCGGCTCAAAAAGAACCGCCGACGGCAGCGAATTCGAGCCC
AAGAAGAAGAG
GAAAGTCTAACCGGTCATCATCACCATCACCATTGAGTTTAAACCCGCTGATCAGCCTC
GACTGTGCCTT
CTAGTTGCCAGCCATCTGTTGTTTGCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTG
CCACTCCCAC
TGTCTTTCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTA
TTCTGGGGGGT
GGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGG
ATGCGGTGGGCT
CTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCGATACCGTCGACCTCTAGCTAG
AGCTTGGCGTA
ATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACA
TACGAGCCGGA
AGCATAAAGTGTAAGCCTAGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTT
GCGCTCACTGC
CCGCTTTCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCG
GGGAGAGGCGG
TTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTGTTTCG
GCTGCGGCGA
GCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGC
AGGAAAGAACA
TGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTT
TTCCATAGGCT
CCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGA
CAGGACTATAA
AGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCC
GCTTACCGGAT
ACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGG
TATCTCAGTTC
GGTGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTTCAGCCCGACC
GCTGCGCCTTA
TCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGC
AGCCACTGGTA
ACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCT
AACTACGGCTA
CACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAA
GAGTTGGTAGC
TCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTGTTGCAAGCAGCA
GATTACGCGCA
GAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACACTCAGTGG
AACGAAAACCTC
ACGTAAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAA
ATTAAAAATGA
AGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTT
AATCAGTGAGG
CACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCCTGACTCCCCGTCGTGT
AGATAACTAC
GATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCT
CACCGGCTCCA

10

20

30

40

50

GATTTATCAGCAATAAACCCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAAC
TTTATCCGCCT
CCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGT
TTGCGCAACGT
TGTTGCCATTGCTACAGGCATCGTGGTGTACGCTCGTCGTTTGGTATGGCTTCATTCA
GCTCCGGTTCC
CAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTT
CGGTCCTCCGA
TCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCAT
AATTCTCTTAC
TGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCT
GAGAATAGTGT
ATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAG
CAGAACTTTAA
AAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAACTCTCAAGGATCTTACCGCTG
TTGAGATCCAG
TTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTACCAGCGT
TTCTGGGTGA
GCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTT
GAATACTCATAC
TCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACA
TATTTGAATG
TATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTG
ACGTCGACGGA
TCGGGAGATCGATCTCCCGATCCCCTAGGGTCGACTCTCAGTACAATCTGCTCTGATGC
CGCATAGTTAA
GCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATT
TAAGCTACAAC
AAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGC
TGCTTCGCGAT
GTACGGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAAT
TACGGGGTCAT
TAGTTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCT
GGCTGACCGCC
CAACGACCCCCGCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAG
GGACTTTCCAT
TGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTA
TC

【0145】

「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させ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、シチジンデアミナーゼ活性、例えば標的C・GをT・Aに変換する活性である。別の実施形態では、塩基編集活性は、アデノシンまたはアデニンデアミナーゼ活性、例えばA・TをG・Cに変換する活性である。別の実施形態では、塩基編集活性は、シチジンデアミナーゼ活性、例えば標的C・GをT・Aに変換する活性であり、アデノシンまたはアデニンデアミナーゼ活性、例えばA・TをG・Cに変換する活性である。いくつかの実施形態では、塩基編集活性は、編集の効率によって評価される。塩基編集効率は、任意の適切な手段、例えば、サンガーシークエンシングまたは次世代シークエンシングによって測定することができる。いくつかの実施形態において、塩基編集効率は、塩基エディターによってもたらされる核酸塩基変換を伴う全配列決定リードのパーセンテージ、例え

10

20

30

40

50

ば、G・C塩基対に変換された標的A・T塩基対を伴う全配列決定リードのパーセンテージによって測定される。いくつかの実施形態において、塩基編集効率は、細胞の集団において塩基編集が行われる場合に、塩基エディターによってもたらされる核酸塩基変換を伴う全細胞のパーセンテージによって測定される。

【0146】

「塩基エディターシステム」という用語は、標的ヌクレオチド配列の核酸塩基を編集するためのシステムを指す。様々な実施形態において、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ム可能なヌクレオチド結合ドメイン（例えばCas9）；(2) 前記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（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）；および(3) 1つまたは複数の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、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。いくつかの実施形態において、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。いくつかの実施形態では、塩基エディターシステムはABE8である。

【0147】

いくつかの実施形態では、塩基エディターシステムは、2つ以上の塩基編集コンポーネントを含み得る。例えば、塩基エディターシステムは、複数のデアミナーゼを含み得る。いくつかの実施形態では、塩基エディターシステムは、1つまたは複数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。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単一の対を利用して、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。

【0148】

塩基エディター系のデアミナーゼドメイン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、またはその結合および相互作用の任意の組合せ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、デアミナーゼドメインは、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、ポリヌクレオチドと結合し、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、テロメラゼKu結合モチーフおよびKuタンパク質、テロメラゼ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。

【0149】

塩基エディターシステムは、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。

塩基エディター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。ある態様において、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、デアミナーゼドメインは、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、ポリヌクレオチドリンカー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。

【0150】

ある実施形態において、塩基エディターシステムは、塩基除去修復(BER)の阻害因子コンポーネントをさらに含むことができる。塩基エディター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、BER阻害因子を含み得る。ある実施形態において、BER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。ある実施形態において、BERの阻害因子は、イノシンBER阻害因子であり得る。ある実施形態において、BER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、BERの阻害因子に融合または連結され得る。ある実施形態において、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およびBER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、BERの阻害因子と非共有結合的に相互作用するか、またはBERの阻害因子と会合することによって、BER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、BER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異種部分またはドメインを含み得る。

【0151】

ある実施形態において、BER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、BER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、BER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

10

20

30

40

50

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、M S2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。

【0152】

用語「Cas9」または「Cas9ドメイン」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誘導ヌクレアーゼを指す。Cas9ヌクレアーゼは、casnl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呼ばれることもある。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) にプロセッシングされる。II型CRISPRシステムでは、pre-crRNAの正しいプロセッ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のプロセッ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的に切断する。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結合および切断はタンパク質と両方のRNAとを典型的に要する。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. 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あるい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。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., 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. et al., Nature 471:602-607(2011); および “A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity.” Jinek M. et al., 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オーソログは、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、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。

【0153】

例示的なCas9はStreptococcus pyogenes Cas9 (spCas9)であり、そのアミノ酸配列を下記に提供する。

10

20

30

40

50

MDKKYSIGLDIGTNSVGWAVITDDYKVPSKKFKVLGNTDRHSIKKNLIGALLFGSGETAETRL
 KRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAY
 HEKYPTIYHLRKKLADSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQIY
 NQLFEENPINASRVDAKAILSARLSKSRLENLIAQLPGEKRNGLFGNLIALSLGLTPNFKSNE
 DLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNSEITKAPLSAS
 MIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMD
 GTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRI
 PYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPHKS
 LLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFD
 SVEISGVEDRFNASLGAYHDLKIIKDKDFLDNEENEDILEDIVLTTLTLFEDRGMIEERLKYA
 HLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTF
 KEDIQKAQVSGQGHSLEQIANLAGSPAIIKKGILQTVKIVDELVKVMGHKPENIVIEMARENQT
 TQKGQKNSRERMKRIEEGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRL
 SDYDVDHIVPQSFIIKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKE
 DNLTKAERGGLSELDKAGFIKRLVETROITKHVAQIILDSRMNTKYDENDKLIREVKVITLKS
 LVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGITALIKKYPKLESEFVYGDYKVYDVRKMIAS
 EQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLMS
 PQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVEKKG
 SKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDLIIKLPKYSLELENGRKRMLASA
 GELQKGNELALPSKYVNFYLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIIEQISEFSKRV
 LADANLDKVL SAYNKHDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDA
 TLIHQSIITGLYETRIDLSQLGGD

10

20

30

(一重下線：HNHドメイン；二重下線：RuvCドメイン)

【0154】

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) または触媒的に不活なCas9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Jine
 k et al, Science. 337:816-821(2012); Qi et al, “Repurposing CRISPR as an RN
 A-Guided Platform for Sequence-Specific Control of Gene Expression” (2013) C
 ell. 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、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, Scien
 ce. 337:816-821(2012); Qi et al, Cell. 28;152(5): 1173-83 (2013))。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
 ッカーゼである。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

40

50

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%の同一性、少なくとも約80%の同一性、少なくとも約90%の同一性、少なくとも約95%の同一性、少なくとも約96%の同一性、少なくとも約97%の同一性、少なくとも約98%の同一性、少なくとも約99%の同一性、少なくとも約99.5%の同一性、または少なくとも約99.9%の同一性を有する。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%の同一性を有し、少なくとも約80%の同一性を有し、少なくとも約90%の同一性を有し、少なくとも約95%の同一性を有し、少なくとも約96%の同一性を有し、少なくとも約97%の同一性を有し、少なくとも約98%の同一性を有し、少なくとも約99%の同一性を有し、少なくとも約99.5%の同一性を有し、または少なくとも約99.9%の同一性を有する。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。

【0155】

ある実施形態において、断片は、長さが少なくとも100アミノ酸である。ある実施形態において、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。

【0156】

ある実施形態において、野生型Cas9は、*Streptococcus pyogenes*由来のCas9に対応する（NCBI参照配列：NC_017053.1、ヌクレオチド配列およびアミノ酸配列は以下の通り）。

```
ATGGATAAGAAATACTCAATAGGCTTAGATATCGGCACAAATAGCGTCGGATGGGCGGT
GATCACTGATGATTATAAGGTTCCGTCTAAAAAGTTCAAGGTTCTGGGAAATACAGACC
GCCACAGTATCAAAAAAAAAATCTTATAGGGGCTCTTTTATTTGGCAGTGGAGAGACAGCG
GAAGCGACTCGTCTCAAACGGACAGCTCGTAGAAGGTATACACGTCGGAAGAATCGTAT
TTGTTATCTACAGGAGATTTTTTCAAATGAGATGGCGAAAGTAGATGATAGTTTCTTTC
ATCGACTTGAAGAGTCTTTTTTGGTGGGAAGAAGACAAGAAGCATGAACGTCATCCTATT
TTTGGAAATATAGTAGATGAAGTTGCTTATCATGAGAAATATCCAATCTATCATCT
GCGAAAAAAATTGGCAGATTCTACTGATAAAGCGGATTTGCGCTTAATCTATTTGGCCT
TAGCGCATATGATTAAAGTTTCGTGGTCAATTTTTTGATTGAGGGAGATTTAAATCCTGAT
AATAGTGATGTGGACAACTATTTATCCAGTTGGTACAAATCTACAATCAATTATTTGA
AGAAAACCTATTAACGCAAGTAGAGTAGATGCTAAAGCGATTCTTTCTGCACGATTGA
GTAAATCAAGACGATTAGAAAATCTCATTGCTCAGCTCCCCGGTGAGAAGAGAAATGGC
TTGTTTGGGAATCTCATTGCTTTGTCATTGGGATTGACCCCTAATTTTAAATCAAATTTT
GATTTGGCAGAAGATGCTAAATTACAGCTTTCAAAGATACTTACGATGATGATTTAGA
TAATTTATTGGCGCAAATTGGAGATCAATATGCTGATTTGTTTTTGGCAGCTAAGAATT
TATCAGATGCTATTTTACTTTTCAAGATATCCTAAGAGTAAATAGTGAAATAACTAAGGCT
CCCCTATCAGCTTCAATGATTAAGCGCTACGATGAACATCATCAAGACTTGACTCTTTT
AAAAGCTTTAGTTTCGACAACAACCTCCAGAAAAGTATAAAGAAATCTTTTTTGATCAAT
CAAAAAACGGATATGCAGGTTATATTGATGGGGGAGCTAGCCAAGAAGAATTTTATAAA
```

TTTATCAAACCAATTTTAGAAAAAATGGATGGTACTGAGGAATTATTGGTGAAACTAAA
TCGTGAAGATTTGCTGCGCAAGCAACGGACCTTTGACAACGGCTCTATTCCCCATCAAA
TTCAC TTGGGTGAGCTGCATGCTATTTTGAGAAGACAAGAAGACTTTTATCCATTTTAA
AAAGACAATCGTGAGAAGATTGAAAAAATCTTGACTTTTCGAATTCCTTATTATGTTGG
TCCATTGGCGCGTGGAATAGTCGTTTTGCATGGATGACTCGGAAGTCTGAAGAAACAA
TTACCCCATGGAATTTTGAAGAAGTTGTCGATAAAGGTGCTTCAGCTCAATCATTTATT
GAACGCATGACAACTTTGATAAAAAATCTTCCAAATGAAAAAGTACTACCAAAACATAG
TTTGCTTTATGAGTATTTTACGGTTTATAACGAATTGACAAAGGTCAAATATGTTACTG
AGGGAATGCGAAAACCAGCATTTCTTTTCAGGTGAACAGAAGAAAGCCATTGTTGATTAA
CTCTTCAAAACAAATCGAAAAGTAACCGTTAAGCAATTAAGAAGATTATTTCAAAAA
AATAGAATGTTTTGATAGTGTTGAAATTTTCAGGAGTTGAAGATAGATTTAATGCTTCAT
TAGGCGCCTACCATGATTTGCTAAAAATTATTAAGATAAAGATTTTTTTGGATAATGAA
GAAATGAAGATATCTTAGAGGATATTGTTTTAACATTGACCTTATTTGAAGATAGGGG
GATGATTGAGGAAAGACTTAAACATATGCTCACCTCTTTGATGATAAGGTGATGAAAC
AGCTTAAACGTCGCCGTTATACTGGTTGGGGACGTTTGTCTCGAAAATTGATTAATGGT
ATTAGGGATAAGCAATCTGGCAAAACAATATTAGATTTTTTTGAAATCAGATGGTTTTGC
CAATCGCAATTTTATGCAGCTGATCCATGATGATAGTTTGACATTTAAAGAAGATATTC
AAAAAGCACAGGTGTCTGGACAAGGCCATAGTTTACATGAACAGATTGCTAACTTAGCT
GGCAGTCCTGCTATTAAAAAAGGTATTTTACAGACTGTAAAAATTGTTGATGAACTGGT
CAAAGTAATGGGGCATAAGCCAGAAAATATCGTTATTGAAATGGCACGTGAAAATCAGA
CAACTCAAAAGGGCCAGAAAAATTTCGCGAGAGCGTATGAAACGAATCGAAGAAGGTATC
AAAGAATTAGGAAGTCAGATTCTTAAAGAGCATCCTGTTGAAAATACTCAATTGCAAAA
TGAAAAGCTCTATCTCTATTATCTACAAAATGGAAGAGACATGTATGTGGACCAAGAAT
TAGATATTAATCGTTTAAAGTGATTATGATGTCGATCACATTGTTCCACAAAGTTTCATT
AAAGACGATTCAATAGACAATAAGGTACTAACGCGTTCTGATAAAAAATCGTGGTAAATC
GGATAACGTTCCAAGTGAAGAAGTAGTCAAAAAGATGAAAAACTATTGGAGACAACCTTC
TAAACGCCAAGTTAATCACTCAACGTAAGTTTGATAATTTAACGAAAGCTGAACGTGGA
GGTTTGAGTGAAC TTGATAAAGCTGGTTTTATCAAACGCCAATTGGTTGAAACTCGCCA
AATCACTAAGCATGTGGCACAATTTTGGATAGTCGCATGAATACTAAATACGATGAAA
ATGATAAACTTATTCGAGAGGTTAAAGTGATTACCTTAAATCTAAATTAGTTTCTGAC
TTCCGAAAAGATTTCCAATTCTATAAAGTACGTGAGATTAACAATTACCATCATGCCCA
TGATGCGTATCTAAATGCCGTCGTTGGAAC TGCTTTGATTAAGAAATATCCAAAAC TTG
AATCGGAGTTTGTCTATGGTGATTATAAAGTTTATGATGTTTCGTAAAATGATTGCTAAG
TCTGAGCAAGAAATAGGCCAAAGCAACCGCAAAATATTTCTTTTACTCTAATATCATGAA
CTTCTTCAAAACAGAAATTACACTTGCAAATGGAGAGATTTCGAAACGCCCTCTAATCG
AACTAATGGGGAACTGGAGAAATTGTCTGGGATAAAGGGCGAGATTTTGCCACAGTG
CGCAAAGTATTGTCCATGCCCCAAGTCAATATTGTCAAGAAAACAGAAGTACAGACAGG
CGGATTCTCCAAGGAGTCAATTTTACCAAAAAGAAATTCGGACAAGCTTATTGCTCGTA
AAAAAGACTGGGATCCAAAAAATATGGTGGTTTTGATAGTCCAACGGTAGCTTATTCA
GTCCTAGTGGTTGCTAAGGTGGAAAAAGGGAAATCGAAGAAGTTAAATCCGTTAAAGA
GTTACTAGGGATCACAATTATGGAAAGAAGTTCCTTTGAAAAAATCCGATTGACTTTT
TAGAAGCTAAAGGATATAAGGAAGTTAAAAAAGACTTAATCATTAACCTACCTAAATAT
AGTCTTTTTGAGTTAGAAAACGGTCGTAAACGGATGCTGGCTAGTGCCGGAGAATTACA
AAAAGGAAATGAGCTGGCTCTGCCAAGCAAATATGTGAATTTTTTATATTTAGCTAGTC
ATTATGAAAAGTTGAAGGGTAGTCCAGAAGATAACGAACAAAAACAATTGTTTGTGGAG
CAGCATAAGCATTATTTAGATGAGATTATTGAGCAAATCAGTGAATTTTCTAAGCGTGT
TATTTTAGCAGATGCCAATTTAGATAAAGTTCTTAGTG CATATAACAAACATAGAGACA
AACCAATACGTGAACAAGCAGAAAATATTATTCAATTTATTTACGTTGACGAATCTTGGA
GCTCCCGCTGCTTTTAAATATTTTGATACAACAATTGATCGTAAACGATATACGTCTAC
AAAAGAAGTTTTAGATGCCACTCTTATCCATCAATCCATCACTGGTCTTTATGAAACAC

10

20

30

40

50

GCATTGATTTGAGTCAGCTAGGAGGTGACTGA

MDKKYSIGLDIGTNSVGWAVITDDYKVPSKKFKVLGNTDRHSIKKNLIGALLFGSGETAETRL
 KRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVAY
 HEKYPTIYHLRKKLADSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQIY
 NQLFEENPINASRVDAKAILSARLSKSRLENLIAQLPGEKRNGLFGNLIALLSLGLTPNFKSNF
 DLAEDAKLQLSKDQYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNSEITKAPLSAS
 MIKRYDEHHQDLTLLKALVRQQLEPKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMD
 GTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRI
 PYYVGPLARGNSRFAWMTRKSEETITPWNFEFVVDKGASQSFIERMTNFDKNLPNEKVLPHKS
 LLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFD
 SVEISGVEDRFNASLGAYHDLKIIKDKDFLDNEENEDILEDIVLTLTLFEDRGMIEERLKYA
 HLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTF
 KEDIQKAQVSGQGHSLHEQIANLAGSPAIKKGILQTVKIVDELVKVMGHKPENIVIEMARENQT
 TQKGQKNSRERMKRIEELGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRL
 SDYDVDHIVPQSFIKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKF
 DNLTKAERGGLSELKAGFIKRQLVETROITKHVAQILDSRMNTKYDENDKLIREVKVITLKS
 LVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGITALIKKYPKLESEFVYGDYKVYDVRKMIKS
 EQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLMS
 PQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVEKKG
 SKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDLIIKLPKYSLFELNGRKRMLASA
 GELQKGNELALPSKYVNFYLYLASHYEKLKGSPEDEQKQLFVEQHKHYLDEIIIEQISEFSKRVI
 LADANLDKVL SAYNKHDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDA
 TLIHQSI TGLYETRIDLSQLGGD

10

20

30

(一重下線：HNHドメイン；二重下線：RuvCドメイン)

【0157】

ある実施形態において、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
 GAGGCGACTCGCCTGAAACGAACGCTCGGAGAAGGTATACACGTCGCAAGAACCGAAT
 ATGTTACTTACAAGAAATTTTTAGCAATGAGATGGCCAAAGTTGACGATTCTTTCTTTC
 ACCGTTTGGGAAGAGTCCTTCCTTGTCGAAGAGGACAAGAAACATGAACGGCACCCCATC
 TTTGGAAACATAGTAGATGAGGTGGCATATCATGAAAAGTACCCAACGATTTATCACCT
 CAGAAAAAAGCTAGTTGACTCAACTGATAAAGCGGACCTGAGGTTAATCTACTTGGCTC
 TTGCCCATATGATAAAGTTCCGTGGGCACTTTCTCATTGAGGGTGATCTAAATCCGGAC
 AACTCGGATGTCGACAACTGTTTCATCCAGTTAGTACAAACCTATAATCAGTTGTTTGA
 AGAGAACCCTATAAATGCAAGTGCGGTGGATGCGAAGGCTATTCTTAGCGCCCGCCTCT
 CTAATCCCGACGGCTAGAAAACCTGATCGCACAAATTACCCGGAGAGAAGAAAAATGGG

40

50

TTGTTTCGGTAACCTTATAGCGCTCTCACTAGGCCTGACACCAAATTTTAAAGTCGAACTT
CGACTTAGCTGAAGATGCCAAATTGCAGCTTAGTAAGGACACGTACGATGACGATCTCG
ACAATCTACTGGCACAAATTGGAGATCAGTATGCGGACTTATTTTTGGCTGCCAAAAAC
CTTAGCGATGCAATCCTCCTATCTGACATACTGAGAGTTAATACTGAGATTACCAAGGC
GCCGTTATCCGCTTCAATGATCAAAAGGTACGATGAACATCACCAAGACTTGACACTTC
TCAAGGCCCTAGTCCGTGAGCAACTGCCTGAGAAATATAAGGAAATATTCTTTGATCAG
TCGAAAAACGGGTACGCGAGGTTATATTGACGGCGGAGCGAGTCAAGAGGAATTCTACAA
GTTTATCAAACCCATATTAGAGAAGATGGATGGGACGGAAGAGTTGCTTGTA AAACTCA
ATCGCGAAGATCTACTGCGAAAGCAGCGGACTTTCGACAACGGTAGCATTCCACATCAA
ATCCACTTAGGCGAATTGCATGCTATACTTAGAAGGCAGGAGGATTTTTATCCGTTTCT
CAAAGACAATCGTGAAAAGATTGAGAAAATCCTAACCTTTCGCATACCTTACTATGTGG
GACCCCTGGCCCGAGGGA ACTCTCGGTTTCGCATGGATGACAAGAAAGTCCGAAGAAACG
ATTACTCCATGGAATTTTGAGGAAGTTGTGCGATAAAGGTGCGTCAGCTCAATCGTTCAT
CGAGAGGATGACCAACTTTGACAAGAATTTACCGAACGAAAAAGTATTGCCTAAGCACA
GTTTACTTTACGAGTATTTTACAGTGTACAATGAACTCACGAAAGTTAAGTATGTCACT
GAGGGCATGCGTAAACCCGCCTTTCTAAGCGGAGAACAGAAGAAAGCAATAGTAGATCT
GTTATTCAAGACCAACCGCAAAGTGACAGTTAAGCAATTGAAAGAGGACTACTTTAAGA
AAATTGAATGCTTCGATTCTGTGAGATCTCCGGGGTAGAAGATCGATTTAATGCGTCA
CTTGGTACGTATCATGACCTCCTAAAGATAATTAAGATAAGGACTTCCTGGATAACGA
AGAGAATGAAGATATCTTAGAAGATATAGTGTGACTCTTACCCTCTTTGAAGATCGGG
AAATGATTGAGGAAAGACTAAAAACATACGCTCACCTGTTTCGACGATAAGGTTATGAAA
CAGTTAAAGAGGCGTCGCTATACGGGCTGGGGACGATTGTGCGGAAACTTATCAACGG
GATAAGAGACAAGCAAAGTGGTAAAAC TATTCTCGATTTTCTAAAGAGCGACGGCTTCG
CCAATAGGAACTTTATGCAGCTGATCCATGATGACTCTTTAACCTTCAAAGAGGATATA
CAAAAGGCACAGGTTTCCGGACAAGGGGACTCATTGCACGAACATATTGCGAATCTTGC
TGGTTCGCCAGCCATCAAAAAGGGCATACTCCAGACAGTCAAAGTAGTGGATGAGCTAG
TTAAGGTCATGGGACGTCACAAACCGGAAAACATTGTAATCGAGATGGCACGCGAAAAT
CAAACGACTCAGAAGGGGGCAAAAAAACAGTCGAGAGCGGATGAAGAGAATAGAAGAGG
GTATTAAAGAACTGGGCAGCCAGATCTTAAAGGAGCATCCTGTGGAAAATACCCAATTG
CAGAACGAGAACTTTACCTCTATTACCTACAAAATGGAAGGGACATGTATGTTGATCA
GGA ACTGGACATAAACCGTTTATCTGATTACGACGTCGATCACATTGTACCCCAATCCT
TTTTGAAGGACGATTCAATCGACAATAAAGTGCTTACACGCTCGGATAAGAACCGAGGG
AAAAGTGACAATGTTCCAAGCGAGGAAGTCGTAAAGAAAATGAAGAACTATTGGCGGCA
GCTCCTAAATGCGAACTGATAACGCAAAGAAAGTTGATAACTTAACTAAAGCTGAGA
GGGGTGGCTTGTCTGAACTTGACAAGGCCGGATTTATTAAACGTCAGCTCGTGGAACCC
CGCCAAATCACAAAGCATGTTGCACAGATACTAGATTCCCGAATGAATACGAAATACGA
CGAGAACGATAAGCTGATTCGGGAAGTCAAAGTAATCACTTTAAAGTCAAATTGGTGT
CGGACTTCAGAAAGGATTTTCAATTCTATAAAGTTAGGGAGATAAATAACTACCACCAT
GCGCACGACGCTTATCTTAATGCCGTCGTAGGGACCGCACTCATTAAGAAATACCCGAA
GCTAGAAAGTGAGTTTGTGTATGGTGATTACAAAGTTTATGACGTCGTAAGATGATCG
CGAAAAGCGAACAGGAGATAGGCAAGGCTACAGCCAAATACTTCTTTTATTCTAACATT
ATGAATTTCTTTAAGACGGAAATCACTCTGGCAAACGGAGAGATACGCAAACGACCTTT
AATTGAAACCAATGGGGAGACAGGTGAAATCGTATGGGATAAGGGCCGGGACTTCGCGA
CGGTGAGAAAAGTTTTGTCCATGCCCAAGTCAACATAGTAAAGAAAAC TGAAGTGCAG
ACCGGAGGGTTTTCAAAGGAATCGATTCTTCAAAAAGGAATAGTGATAAGCTCATCGC
TCGTAAAAAGGACTGGGACCCGAAAAAGTACGGTGGCTTCGATAGCCCTACAGTTGCCT
ATTCTGTCTAGTAGTGGCAAAGTTGAGAAGGGGAAAATCCAAGAACTGAAGTCAGTC
AAAGAATTATTGGGGATAACGATTATGGAGCGCTCGTCTTTTGAAAAGAACCCCATCGA
CTTCCTTGAGGCGAAAGGTTACAAGGAAGTAAAAAAGGATCTCATAATTAACTACCAA
AGTATAGTCTGTTTGAGTTAGAAAATGGCCGAAAACGGATGTTGGCTAGCGCCGGAGAG

10

20

30

40

50

CTTCAAAGGGGAACGAACTCGCACTACCGTCTAAATACGTGAATTTCTGTATTTAGC
GTCCCATACGAGAAGTTGAAAGTTTACCTGAAGATAACGAACAGAAGCAACTTTTTG
TTGAGCAGCACAAACATTATCTCGACGAAATCATAGAGCAAATTTGGAATTCAGTAAG
AGAGTCATCCTAGCTGATGCCAATCTGGACAAAGTATTAAGCGCATAACAAGCACAG
GGATAAACCCATACGTGAGCAGGCGGAAATATTATCCATTTGTTTACTCTTACCAACC
TCGGCGCTCCAGCCGCATTCAAGTATTTTGACACAACGATAGATCGCAAACGATACACT
TCTACCAAGGAGGTGCTAGACGCGACACTGATTCACCAATCCATCACGGGATTATATGA
AACTCGGATAGATTTGTCACAGCTTGGGGGTGACGGATCCCCAAGAAGAAGAGGAAAG
TCTCGAGCGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAG
GATGACGATGACAAGGCTGCAGGA

10

MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRL
KRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLSESLVEEDKKHERHPIFGNIVDEVAY
HEKYPTIYHLRKKLVDSTDKADLRILIYLAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTY
NQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALLSLGLTPNFKSNF
DLAEDAKLQLSKDQYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSAS
MIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMD
GTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFLKDNREKIEKILTFRI
PYYVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGASQSFIERMTNFDKNLPNEKVLPHKS
LLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFD
SVEISGVEDRFNASLGTYHDLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYA
HLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTF
KEDIQKAQVSGQGSLSLHEHIANLAGSPAIKKGILQTVKVVDLVKVMGRHKPENIVIEMARENQ
TTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINR
LSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRK
FDNLTKAERGGLSELKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS
KLVSDFRKDFQFYKVBREINNYHHAHDAYLNAVVGITALIKKYPKLESEFVYGDYKVYDVRKMIAK
SEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLS
MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVEKG
KSKKLKSVKELLGITIMERSSEKPNIDFLEAKGYKEVKKDLIIKLPKYSLEFENGKRMLAS
AGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDEQKQLFVEQHKHYLDEIIIEQISEFSKR
VILADANLDKVL SAYNKHDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLD
ATLIHQSIITGLYETRIDLSQLGGD

20

30

40

(一重下線：HNHドメイン；二重下線：RuvCドメイン)

【0158】

いくつかの実施形態において、野生型Cas9は、*Streptococcus pyogenes*からのCas9
(NCBI参照配列：NC_002737.2 (下記のヌクレオチド配列) 及びUniprot参照配列：Q
99ZW2 (下記のアミノ酸配列) に対応する。

ATGGATAAGAAATACTCAATAGGCTTAGATATCGGCACAAATAGCGTCGGATGGGCGGT
GATCACTGATGAATATAAGGTTCCGTCTAAAAAGTTCAAGGTTCTGGGAAATACAGACC
GCCACAGTATCAAAAAAATCTTATAGGGGCTCTTTTATTTGACAGTGGAGAGACAGCG

50

GAAGCGACTCGTCTCAAACGGACAGCTCGTAGAAGGTATACACGTCGGAAGAATCGTAT
TTGTTATCTACAGGAGATTTTTTCAAATGAGATGGCGAAAGTAGATGATAGTTTCTTTC
ATCGACTTGAAGAGTCTTTTTTGGTGGGAAGAAGACAAGAAGCATGAACGTCATCCTATT
TTTGGAAATATAGTAGATGAAGTTGCTTATCATGAGAAATATCCAACCTATCTATCATCT
GCGAAAAAATTGGTAGATTCTACTGATAAAGCGGATTTGCGCTTAATCTATTTGGCCT
TAGCGCATATGATTAAGTTTCGTGGTCATTTTTTGGATTGAGGGAGATTTAAATCCTGAT
AATAGTGATGTGGACAAACTATTTATCCAGTTGGTACAAACCTACAATCAATTATTTGA
AGAAAACCCTATTAACGCAAGTGGAGTAGATGCTAAAGCGATTCTTTCTGCACGATTGA
GTAAATCAAGACGATTAGAAAATCTCATTGCTCAGCTCCCCGGTGAGAAGAAAAATGGC
TTATTTGGGAATCTCATTGCTTTGTCATTGGGTTTGACCCCTAATTTTAAATCAAATTTT
GATTTGGCAGAAGATGCTAAATTACAGCTTTCAAAGATACTTACGATGATGATTTAGA
TAATTTATTGGCGCAAATTGGAGATCAATATGCTGATTTGTTTTTGGCAGCTAAGAATT
TATCAGATGCTATTTTACTTTTACAGATATCCTAAGAGTAAATACTGAAATAACTAAGGCT
CCCCTATCAGCTTCAATGATTAAACGCTACGATGAACATCATCAAGACTTGACTCTTTT
AAAAGCTTTAGTTTCGACAACAACCTTCCAGAAAAGTATAAAGAAATCTTTTTTGATCAAT
CAAAAAACGGATATGCAGGTTATATTGATGGGGGAGCTAGCCAAGAAGAATTTTATAAA
TTTATCAAACCAATTTTAGAAAAAATGGATGGTACTGAGGAATTATTGGTGAAACTAAA
TCGTGAAGATTTGCTGCGCAAGCAACGGACCTTTGACAACGGCTCTATTCCCATCAAA
TTCACCTGGGTGAGCTGCATGCTATTTTGAGAAGACAAGAAGACTTTTATCCATTTTTA
AAAGACAATCGTGAGAAGATTGAAAAAATCTTGACTTTTCGAATTCCTTATTATGTTGG
TCCATTGGCGCGTGGAATAGTCGTTTTGCATGGATGACTCGGAAGTCTGAAGAAACAA
TTACCCCATGGAATTTTGAAGAAGTTGTCGATAAAGGTGCTTCAGCTCAATCATTTATT
GAACGCATGACAAACTTTGATAAAAAATCTTCAAATGAAAAAGTACTACCAAAACATAG
TTTGCTTTATGAGTATTTTACGGTTTATAACGAATTGACAAAGGTCAAATATGTTACTG
AAGGAATGCGAAAACCAGCATTTCTTTTACGGTGAACAGAAGAAAGCCATTGTTGATTTA
CTCTTCAAACAAATCGAAAAGTAACCGTTAAGCAATTAAGAAGATTATTTCAAAAA
AATAGAATGTTTTGATAGTGTTGAAATTTTACGGAGTTGAAGATAGATTTAATGCTTCAT
TAGGTACCTACCATGATTTGCTAAAAATTATTAAGATAAAGATTTTTTGGATAATGAA
GAAATGAAGATATCTTAGAGGATATTGTTTTAACATTGACCTTATTTGAAGATAGGGA
GATGATTGAGGAAAGACTTAAACATATGCTCACCTCTTTGATGATAAGGTGATGAAAC
AGCTTAAACGTCGCCGTTATACTGGTTGGGGACGTTTGTCTCGAAAATTGATTAATGGT
ATTAGGGATAAGCAATCTGGCAAAACAATATTAGATTTTTTGAATCAGATGGTTTTGC
CAATCGCAATTTTATGCAGCTGATCCATGATGATAGTTTGACATTTAAAGAAGACATTC
AAAAAGCACAAAGTGTCTGGACAAGGCGATAGTTTACATGAACATATTGCAAATTTAGCT
GGTAGCCCTGCTATTAAAAAAGGTATTTTACAGACTGTAAAAGTTGTTGATGAATTGGT
CAAAGTAATGGGGCGGCATAAGCCAGAAAATATCGTTATTGAAATGGCACGTGAAAATC
AGACAACTCAAAAGGGCCAGAAAAATTGCGGAGAGCGTATGAAACGAATCGAAGAAGGT
ATCAAAGAATTAGGAAGTCAGATTCTTAAAGAGCATCCTGTTGAAAATACTCAATTGCA
AAATGAAAAGCTCTATCTCTATTATCTCCAAAATGGAAGAGACATGTATGTGGACCAAG
AATTAGATATTAATCGTTTAAAGTGATTATGATGTCGATCACATTGTTCCACAAAGTTTC
CTTAAAGACGATTCAATAGACAATAAGGTCTTAACGCGTTCTGATAAAAAATCGTGGTAA
ATCGGATAACGTTCCAAGTGAAGAAGTAGTCAAAAAGATGAAAAACTATTGGAGACAAC
TTCTAAACGCCAAGTTAATCACTCAACGTAAGTTTGATAATTTAACGAAAGCTGAACGT
GGAGGTTTGAGTGAACCTTGATAAAGCTGGTTTTATCAAACGCCAATTGGTTGAAACTCG
CCAAATCACTAAGCATGTGGCACAAATTTTGATAGTCGCATGAATACTAAATACGATG
AAAATGATAAACTTATTCGAGAGGTTAAAGTGATTACCTTAAATCTAAATTAGTTTCT
GACTTCCGAAAAGATTTCCAATTCTATAAAGTACGTGAGATTAACAATTACCATCATGC
CCATGATGCGTATCTAAATGCCGTCGTTGGAAGTCTTTGATTAAGAAATATCCAAAAC
TTGAATCGGAGTTTGTCTATGGTGATTATAAAGTTTATGATGTTTCGTAAAATGATTGCT
AAGTCTGAGCAAGAAATAGGCAAAGCAACCGCAAAATATTTCTTTTACTCTAATATCAT

10

20

30

40

50

GAACCTTCTTCAAAACAGAAATTACACTTGCAAATGGAGAGATTGCGAAACGCCCTCTAA
TCGAAACTAATGGGGAACTGGAGAAATTGTCTGGGATAAAGGGCGAGATTTTGCCACA
GTGCGCAAAGTATTGTCCATGCCCCAAGTCAATATTGTCAAGAAAACAGAAGTACAGAC
AGGCGGATTCTCCAAGGAGTCAATTTTACCAAAAAGAAATTCGGACAAGCTTATTGCTC
GTA AAAAAGACTGGGATCCAAAAAATATGGTGGTTTTGATAGTCCAACGGTAGCTTAT
TCAGTCCTAGTGGTTGCTAAGGTGGAAAAAGGGAAATCGAAGAAGTTAAATCCGTTAA
AGAGTTACTAGGGATCACAATTATGGAAAGAAGTTCCTTTGAAAAAATCCGATTGACT
TTTTAGAAGCTAAAGGATATAAGGAAGTTAAAAAAGACTTAATCATTAAACTACCTAAA
TATAGTCTTTTTGAGTTAGAAAACGGTCGTAAACGGATGCTGGCTAGTGCCGGAGAATT
ACAAAAAGGAAATGAGCTGGCTCTGCCAAGCAAATATGTGAATTTTTTATATTTAGCTA
GTCATTATGAAAAGTTGAAGGGTAGTCCAGAAGATAACGAACAAAAACAATTGTTTGTG
GAGCAGCATAAGCATTATTTAGATGAGATTATTGAGCAAATCAGTGAATTTTCTAAGCG
TGTTATTTTAGCAGATGCCAATTTAGATAAAGTTCTTAGTG CATATAACAAACATAGAG
ACAAACCAATACGTGAACAAGCAGAAAATATTATTCATTTATTTACGTTGACGAATCTT
GGAGCTCCCGCTGCTTTTAAATATTTTGATACAACAATTGATCGTAAACGATATACGTC
TACAAAAGAAGTTTTAGATGCCACTCTTATCCATCAATCCATCACTGGTCTTTATGAAA
CACGCATTGATTTGAGTCAGCTAGGAGGTGACTGA

10

MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRL
KRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNIVDEVAY
HEKYPTIYHLRKKLVDSTDKADLRLLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTY
NQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALLSLGLTPNFKSNF
DLAEDAKLQLSKDITYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSAS
MIKRYDEHHQDLTLLKALVRQQLPKEYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMD
GTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRI
PYVVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNEKVLPHKS
LLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFD
SVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYA
HLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKS DGFANRNFMQLIHDDSLTF
KEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQ
TTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINR
LSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRK
FDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS
KLVSDFRKDFQFYKVBREINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIAK
SEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLS
MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEVG
KSKKLKSVKELLGITIMERSSEFEKNPIDFLEAKGYKEVKKDIIKLPKYSLFELENGRKRMLAS
AGELQKGNELALPSKYVNFLYLASHYEKLKGS PEDNEQKQLFVEQHKHYLDEIIEQISEFSKR
VILADANLDKVL SAYNKH RDKPIREQAENIIHLFTLTNLGAPAAF KYFDTTIDRKRYTSTKEVLD
ATLIHQ SITGLYETRIDLSQLGGD

20

30

40

50

(配列番号 1)

(一重下線：HNHドメイン；二重下線：RuvCドメイン)

【0159】

ある実施形態において、Cas9は、*Corynebacterium ulcerans* (NCBI Refs: NC_015683.1, NC_017317.1); *Corynebacterium diphtheria* (NCBI Refs: NC_016782.1, NC_016786.1); *Spiroplasma syrophidicola* (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 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、または任意の他の生物からのCas9を指す。

【0160】

いくつかの実施形態において、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、配列番号1における番号付けでD10AおよびH840A変異または別のCas9における対応する変異を含む。いくつかの実施形態において、dCas9は、dCas9 (D10AおよびH840A) のアミノ酸配列を含む：

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRL
KRTARRRYTRRKNRICYLQEIFSNEAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVAY
HEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTY
NQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALLSLGLTPNFKSNF
DLAEDAKLQLSKDITYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSAS
MIKRYDEHHQDLTLLKALVRQQLPKEYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMD
GTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRI
PYYVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGASQSFIERMTNFDKNLPNEKVLPHKS
LLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFD
SVEISGVEDRFNASLGTYHDLKIIKDKDFLDNEENEDILEDIVLTTLTLFEDREMIEERLKYA
HLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTF
KEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQ
TTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINR
LSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRK
FDNLTKAERGGLSELKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS
KLVSDFRKDFQFYKVINNYHHADAYLNAVVGITALIKKYPKLESEFVYGDYKVYDVRKMIAK
SEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL
MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEKG
KSKKLKSVKELLGITIMERSSEKPNIDFLEAKGYKEVKKDLIIKLPKYSLEFLENKRKMLAS
AGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDEQKQLFVEQHKHYLDEIIEQISEFSKR
VILADANLDKVL SAYNKHDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLD
ATLIHQSIITGLYETRIDLSQLGGD

10

20

30

40

50

(一重下線：HNHドメイン；二重下線：RuvCドメイン)

【0161】

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。

【0162】

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%の同一性、少なくとも約80%の同一性、少なくとも約90%の同一性、少なくとも約95%の同一性、少なくとも約98%の同一性、少なくとも約99%の同一性、少なくとも約99.5%の同一性、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。

【0163】

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。

【0164】

追加的なCas9タンパク質（例えば、ヌクレアーゼ不活 (dead) 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が理解されるべきである。例示的なCas9タンパク質は、限定されるものではないが、下記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、ヌクレアーゼ不活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。

【0165】

例示的な触媒不活Cas9 (dCas9):

DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEAT
RLKRTARRRYTRRKNNRICYLQEIFSNEMAKVDDSSFFHRLEESFLVEEDKKHERHPIFGNIV
DEVAYHEKYPTIYHLRKKLV DSTDKADLRLLIYLA LAHMIKFRGHFLIEGDLNPDNSDVK
LFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALS
LGLTPNFKSNFDLAEDAKLQLSKD TYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDIL
RVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQQLPEKYKEIFFDQSKNGYAGYIDGG
ASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQE
DFYPFLKDNREKIEKILTFRIPIYYVGPLARGNSRFAWMTRKSEETITPWNFEVVVDKGAS
AQSFIERMTNFDKNLPNEKVLPHKSHLLYEFYFTVYNELTKVKYVTEGMRKPAFLSGEQKKA
IVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLD
NEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLIN
GIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA
GSPAIIKKGILQTVKVVDELVKVMGRHKPENIVIMARENQTTQKGQKNSRERMKRIEEGI
KELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLK
DDSIDNKVLT RSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGG

LSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRK
DFQFYKVRINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQ
EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLS
MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAK
VEKGKSKKLKSVKELLGITIMERSSEFKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGR
KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKKHYLDEII
EQISEFSKRVLADANLDKVL SAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTID
RKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGD

【0166】

例示的な触媒のCas9ニッカーゼ (nCas9):

DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEAT
RLKRTARRRYTRRKNRICYLQEIFSNE MAKVDDSFHRLEESFLVEEDKKHERHPIFGNIV
DEVAYHEKYPTIYHLRKKLVDSTDKADRLIYLALAHMIKFRGHFLIEGDLNPDNSDVK
LFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIAS
LGLTPNFKSNFDLAEDAKLQLSKD TYDDDLDNLLAQIGDQYADFLAAKNLSDAILLSDIL
RVNTEITKAPLSASMIKRYDEHHQD LTLKALVRQQLP EKYKEIFFDQSKNGYAGYIDGG
ASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQE
DFYPFLKDNREKIEKILTFRIPIYYVG PLARGNSRFAWMTRKSEETITPWNFEVVVDKGAS
AQSFIERMTNFDKNLPNEKVL PKHSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKA
IVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLD
NEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLIN
GIRDKQSGKTILDFLKSDGFANRNF MQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA
GSPAIIKKGILQTVKVVDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGI
KELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK
DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGG
LSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRK
DFQFYKVRINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQ
EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLS
MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAK
VEKGKSKKLKSVKELLGITIMERSSEFKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGR
KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKKHYLDEII
EQISEFSKRVLADANLDKVL SAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTID
RKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGD

【0167】

例示的な触媒の活性Cas9:

DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEAT
RLKRTARRRYTRRKNRICYLQEIFSNE MAKVDDSFHRLEESFLVEEDKKHERHPIFGNIV
DEVAYHEKYPTIYHLRKKLVDSTDKADRLIYLALAHMIKFRGHFLIEGDLNPDNSDVK
LFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIAS
LGLTPNFKSNFDLAEDAKLQLSKD TYDDDLDNLLAQIGDQYADFLAAKNLSDAILLSDIL
RVNTEITKAPLSASMIKRYDEHHQD LTLKALVRQQLP EKYKEIFFDQSKNGYAGYIDGG
ASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQE
DFYPFLKDNREKIEKILTFRIPIYYVG PLARGNSRFAWMTRKSEETITPWNFEVVVDKGAS
AQSFIERMTNFDKNLPNEKVL PKHSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKA
IVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLD
NEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLIN
GIRDKQSGKTILDFLKSDGFANRNF MQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA
GSPAIIKKGILQTVKVVDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGI
KELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK
DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGG

LSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRK
DFQFYKVVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVDVRKMIKSEQ
EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLS
MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAK
VEKGKSKKLKSVKELLGITIMERSSEFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLELENGR
KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEII
EQISEFSKRVILADANLDKVL SAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTID
RKRYTSTKEVL DATLIHQSI TGLYETRIDLSQLGGD.

【0168】

ある実施形態において、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。ある実施形態において、Cas9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、Cas9は、CasXまたはCasXのバリエーションを表す。いくつかの実施形態において、Cas9は、CasYまたはCasYのバリエーションを表す。核酸プログラミング可能DNA結合タンパク質（napDNAbp: nucleic acid programmable DNA binding protein）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。

【0169】

特定の実施形態において、本発明の方法において有用なnapDNAbpは、循環置換体を含み、これは公知であって例えばOakes et al., Cell 176, 254-267, 2019に記載されている。下記は例示的な循環置換体であり、ここで太字の配列はCas9由来の配列を示し、斜体の配列はリンカー配列を表し、下線の配列は二部分核局在化配列を表す。

10

20

30

40

50

CP5 (with MSP “NGC=Pam Variant with mutations Regular Cas9 likes NGG” PID=Protein Interacting Domain and “D10A” nickase):

EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQ
VNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFMQPTVAYSVLVAKVEKGKSK
KLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDLIIKLPKYSLFELNGRKRMLASAKF
LQKGNELALPSKYVNFLYLASHYEKLKGSPEDEQKQLFVEQHKHYLDEIIEQISEFSKRVILA
DANLDKVL SAYNKH RD KPIREQAENIIHLFTLTNLGAPRAFKYFDTTIARKEYRSTKEVLDATL
IHQSITGLYETRIDLSQLGGDGGSGGSGGSGGSGGSGGMDKKYSIGLAIGTNSVGWAVITD
EYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS
NEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADL
RLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSAR
LSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLL
AQIGDQYADLFLAAKNLSAAILSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQOL
PEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN
GSIHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPIYYVGPLARGNSRFAWMTRKSEE
TITPWNFEVVDKGASAQSFIERMTNFDKNLPNEKVLPHSLLYEYFTVYNELTKVKYVTEGMR
KPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLK
IIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWRSL
RKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANL
AGSPAIKKGILQTVKVVDLVKVMGRHKPENIVIMARENQTTQKGQKNSRERMKRIEEGIKEL
GSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVHIVPQSFLKDDSIDNK
VLTRSDKNRGKSDNVPSEEVVKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKR
QLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHH
AHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIKSEQEGADKRTADGSEFESPKKKR
KV*

10

20

30

【0170】

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。

40

【0171】

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%

50

、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Cas12b/C2c1、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。

【0172】

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

MAVKSIVKVLRLDDMPEIRAGLWKLHKEVNAGVRYYTEWLSLLRQENLYRRSPNGDGEQ
ECDKTAECKAELLERLRARQVENGHARGPAGSDDELLQLARQLYELLVPQAIGAKGDAQ
QIARKFLSPLADKDAVGGLGIAKAGNKPRWVRMREAGEPGWEEKEKAETRKSADRTAD
VLRALADFGKPLMRVYTDSEMSSVEWKPLRKGQAVRTWDRDMFQQAIERMMSWESW
NQRVGQEYAKLVEQKNRFEQKNFVGQEHVHLVNQLQQDMKEASPGLESKEQTAHYVT
GRALRGSDKVFKEWGLAPDAPFDLYDAEIKNVQRRNTRRFGSHDLFAKLAEPEYQALW
REDASFLTRYAVYNSILRKLNHAKMFATFTLPDATAHPIWTRFDKLGGNLHQYTFLFNEF
GERRHAIRFHKLKLVENGVAEVDVTPISMSEQLDNLLPRDPNEPIALYFRDYGAEQH
FTGEFGGAKIQCRDQLAHMHRRRGARDVYLVNSVRVQSQSEARGERRPPYAAVFRLVG
DNHRAFVHFDKLSDYLAHPDDGKLGSEGLLSGLRVMSVDLGLRTSASISVFRVARKDEL
KPNSKGRVPFFFPKGNNDLVAVHERSQQLLKLPGETESKDLRAIREERQRTLRLQLRTQLAY
LRLLVRCGSEDVGRRERSWAKLIEQPVDAANHMTDPDWREAFENELQKLKSLHGICSDKE
WMDAVYESVRRVWRHMGKQVRDWRKDVRSGERPKIRGYAKDVVGGNSIEQIEYLERQY
KFLKSWSSFFGKVSQVIRAEKGSRAITLREHIDHAKEDRLKKLADRIIMEALGYVYALDE
RGKGKWWAKYPPCQLILLEELSEYQFNDRPPSENNQLMQWSHRGVFQELINQAQVHD
LLVGTMYAAFSSRFDARTGAPGIRCRRVPARCTQEHNPFPFWWLNKFVVEHTLDACPL
RADDLIPTGEGEIFVSPFSAEEGDFHQIHADLNAAQNLQQLRWSDFDISQIRLRCDWGEV
DGELVLIPRLTGKRTADSYSNKVFYTNVTGVTYERERGGKRRKVFAQEKLSEEEAELLVEA
DEAREKSVVLMRDPGSIINRGNWTRQKEFWSMV NQRIEGLVKQIRSRVPLQDSACENT
GDI

【0173】

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
KAKKKKGEEGETTTSNIILPLSGNDKNPWTETLKCYNFPTTVALSEVFKNFSQVKECEEVS
APSFVKPEFYEFGRSPGMVERTRRVKLEVEPHYLIIAAGWVLTRLGKAKVSEG DYVGVN
VFTPTRGILYSLIQNVNGIVPGIKPETAFLWIARKVVSSVTNPVSVVRIYTISDAVGQN
PTTINGGFSIDLTKLLEKRYLLSERLEAIARNALSISSNMREYIVLANIYIYELTG SKRLE
DLFYFANRDLIMNLSDDGKVRDLKLISAYVNGELIRGEG

【0174】

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

MEVPLYNIFGDNYIIQVATEAENSTIYNNKVEIDDEELRNVNLAYKIAKNNEDAAAERRG
KAKKKKGEEGETTTSNIILPLSGNDKNPWTETLKCYNFPTTVALSEVFKNFSQVKECEEVS
APSFVKPEFYKFGSPGMVERTRRVKLEVEPHYLIIMAAAGWVLTRLGKAKVSEG DYVGV
NVFTPTRGILYSLIQNVNGIVPGIKPETAFLWIARKVVSSVTNPVSVVSIYTISDAVGQ

10

20

30

40

50

NPTTINGGFSIDLTKLLEKRDLLSERLEAIARNALSISSNMRERYIVLANIYIYELTGSKRLE
 DLLYFANRDLIMNLSDDGKVRDLKLISAYVNGELIRGEG

[0175]

Deltaproteobacteria CasX

MEKRINKIRKKLSADNATKPVSRSGPMKTLLVRVMTDDLKKRLEKRRKKPEVMPQVISN
 NAANNLRMLDDYTKMKEAILQVYWQEFKDDHVGLMCKFAQPASKKIDQNKLPKPEMDE
 KGNLTAGFACSCGQPLFVYKLEQVSEK GKAYTNYFGRCNVAEHEKLILLAQLKPKV KDS
 DEAVTYSLGKFGQRALDFYSIHVTKESTHPVKPLAQIAGNRYASGPVGKALS DACMG TIA
 SFLSKYQDIIIEHQKVVKGNQKRLESLRELAKENLEYPSTLPPQPHTKEGV DAYNEVIA
 RVRMWVNLNLWQKLKLSRDDAKPLLRLKGFSPFVVERRENEVDWWNTINEVKKLIDA
 KRDMGRVFWSGVTAEK RNTILEGYNLPNENDHKKREGSLENPKKPAKRQFGDLLLYLE
 KKYAGDWGKVFDEAWERIDKKIAGLTSHIEREEARNAEDAQSKAVLTDWLRAKAS FVLE
 RLKEMDEKEFYACEIQLQKWYGD LRGNPF AVEAENRVVDISGFSIGSDGHSIQYRNLLAW
 KYLENGKREFYLLMNYGKKGRIRFTDGTDIKKS GKWQGLLYGGGKAKVIDLTFDPDDEQL
 IILPLAFGTRQGREFIWNDLLSLETGLIKLANGRVIEKTIYNKKIGRDEPALFVALTFERRE
 VVDPSNIKPVNLIGVARGENIPAVIALTDPEGCPLPEFKDSSGGPTDILRIGEGYKEKQRAI
 QAAKEVEQRRAGGYSRK FASKSRNLADDMVRNSARDLFYHAVTHDAVLVFANLSRGFGR
 QGKRTFMTERQYTKMEDWLTAKLAYEGLTSKTYLSKTLAQYTSKTC SNCGFTITYADMD
 VMLVRLKKTSDGWATT LNNKELKAEYQITYYNNRYKRQTVEKELSAELDR LSEESGNNDIS
 KWTGRRDEALFLLKKRFSHRPVQE QFVCLDCGHEVHAAEQ AALNIARSWLFLNSN STE
 FSKYKSGKQPFVGAWQAFYKRRLKEVWK PNA

10

20

[0176]

CasY (ncbi.nlm.nih.gov/protein/APG80656.1)

> APG80656.1 CRISPR-associated protein CasY [uncultured Parcubacteria gro
 up bacterium]

MSKRHPRISGVKGYRLHAQRLEYTGKSGAMRTIKYPLYSSPSGGRTVPREIVSAINDDYV
 GLYGLSNFDDLYNAEK RNEEKVYSVLDFWYDCVQYGAVFSYTA PGLLKNVAEVRGGSYE
 LTKTLKGSHLYDELQIDKVIKFLNKKEISRANGSLDKLKKDIIDCFKA EYRERHKDQC NKL
 ADDIKNAKKDAGASLGERQKKLFRDFFGISEQSENDKPSFTNPLNLTCCLLPFDTVNNNR
 NRGEVLFNKLKEYAQKLDKNEGSLEMWEYIGIGNSGTAFSNFLGEGFLGRLRENKITE LK
 KAMMDITDAWRGQE QEEEELEKRLRILAALTIKLREP KFDNHWGGYRSDINGKLSSWLQN
 YINQTVKIKEDLKGHKDLKKAKEMINRFGESDTKEEAVVSSLLESIEKIVPDDSADDEKP
 DIPAIAYRRFLSDGRLTLNR FVQREDVQEALIKERLEAEKKKKPKKRKKKS DA EDEKETID
 FKELFPHLAKPLKLVPNFYGDSKRELYKKYKNAAIYTDALWKAVEKIYKSAFSSSLKNSFF
 DTD FDKDFFIKRLQKIFS VYRRFNTDKWKPIVKN SFAPYCDIVSLAENEVLYKPKQSR SRK
 SAAIDKNRVRLPSTENIAKAGIALARELSVAGFDWKDLLKKEEHEEYIDLIELHK TALALL
 LAVTETQLDISALDFVENGT VKDFMKT RDGNLVLEGRFLEMFSQSIVFSEL RGLAGLMSR
 KEFITRS AIQTMNGKQAELLYIPHEFQSAKITTPKEMSRAFLDLAPAEFATSLEPESLSEKS
 LLKLKQMRYYPHYFGYELTRTGQGIDGGVAENALRLEKSPVKKREIKCKQYKTLGRGQNK
 IVLYVRSSYYQTQFLEWFLHRPKNVQTDVAVSGSFLIDEKKVKTRWNYDALTVALEPVS
 GSERVFVSQPFTIFPEKS AEEEGQRYLGIDIGEYGIAYTALEITGDSAKILDQNFISDPQLKT
 LREEVKGLKLDQRRGTFAMPSTKIARIRESLVHSLRNRIHHLALKHKAKIVYELEVSRFEE
 GKQKIKKVYATLKKADVSEIDADKNLQTTVWGKLAVASEISASYTSQFCGACKKLWRAE
 MQVDETITTQELIGTVRVIKGGTLIDAIKDFMRPPIFDENDTPFPKYRDFCDKHHISKMR
 GNSCLFICPFCRANADADIQASQTIALLRVKEEKKVEDYFERFRKLKNIKVLGQM KKI

30

40

[0177]

「Cas12」または「Cas12ドメイン」という用語は、Cas12タンパク質またはその断片
 (例えば、Cas12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Cas12のgRNA結合ドメインを含むタンパク質)を含むRNA誘導ヌクレアーゼを指す。
 Cas12はクラス2、タイプV CRISPR/Casシステムに属する。Cas12ヌクレアーゼは、CR

50

ISPR (clustered regularly interspaced short palindromic repeat) 関連ヌクレオチドと呼ばれることもある。例示的な *Bacillus hisashii* Cas 12b (BhCas12b) Cas 12 ドメインの配列を以下に提供する。

```
MAPKKKRKVGIHGVPAATRSLFKIEPNEEVKKGLWKTHEVLNHGIAYYMNILKLIRQE
AIYEHHEQDPKNPKKVSKEAIQAELWDFVLKMQKCNSTHEVDKDEVFNILRELYEELVP
SSVEKKGEANQLSNKFLYPLVDPNSQSGKGTASSGRKPRWYNLKIAGDPSWEEEEKKKWE
EDKKKDPLAKILGKLAEYGLIPLFIPYTDSEPIVKEIKWMEKSRNQSVRRLDKDMFIQAL
ERFLSWESWNLKVKEEYKVEKEYKTLEERIKEDIQALKALEQYKERQEQLLRDTLNTN
EYRLSKRGLRGWREIIQKWLKMDENEPSEKYLEVFQDYQRKHPREAGDYSVYEFLSKKE
NHFIWRNHPEYPYLYATFCEIDKKKKDAKQQTFTLADPINHPLWVRFEERSGSNLNKY
RILTEQLHTEKLKKKLTQVLDRLIPTESGGWEEKGKVDIVLLPSRQFYNQIFLDIEEKKGK
HAFTYKDESIKFLKGTLLGGARVQFDRDHLRRYPHKVESGNVGRIFYNMTVNIEPTESPV
SKSLKIHRDDFPKVVNFKPKELTEWIKDSKGKKLKSIGESLEIGLRVMSIDLGRQAAAASI
FEVVDQKPDIEGKLFPIKGTLEYAVHRASFNILKLPGETLVKSREVLRLKAREDNLKLNMNQ
KLNFLRNVLFHFQQFEDITEREKRVTKWISRQENSQVPLVYQDELIQIRELMYKPKDWV
AFLKQLHKKRLEVEIGKEVKHWRKSLSDGRKGLYGISLKNIDEIDRTRKFLLRWSLRPTEPG
EVRRLEPGQRFAIDQLNHLNALKEDRLKKMANTIIMHALGYCYDVRKKKKWQAKNPACQI
ILFEDLSNYPYEEERSRFENSKLMKWSRREIPRQVALQGEIYGLQVGEVGAQFSSRFHAK
TGSPGIRCSVVTKEKLQDNRFKLNQREGRLTLDKIAVLKEGDLYPDKGGEKFISLSKDRK
CVTTHADINAAQNLQKRFWTRTHGFYKVYCKAYQVDGQTVYIPESKDQKQKIIIEFEGEG
YFILKDGVEYEWVNAAGLKKIKKGSSKQSSSELVDSILKDSFDLASELKGEKLMLYRDPSTGN
VFPSDKWMAAGVFFGKLERILISKLTNQYSISTIEDSSKQSMKRPAATKKAGQAKKKK.
```

【0178】

BhCas12b アミノ酸配列に対して少なくとも85%以上の同一性を有するアミノ酸配列もまた、本発明の方法において有用である。

【0179】

「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")、または哺乳動物（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。

【0180】

「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。

【0181】

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

10

20

30

40

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。

【0182】

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、*Escherichia coli*、*Staphylococcus aureus*、*Salmonella typhimurium*、*Shewanella putrefaciens*、*Haemophilus influenzae*、または *Caulobacter crescentus* などの細菌に由来する。

【0183】

ある実施形態において、アデノシンデアミナーゼは、TadAデアミナーゼである。ある実施形態において、TadAデアミナーゼはTadAバリエーションである。ある実施形態において、TadAバリエーションはTadA*8である。ある実施形態において、デアミナーゼまたはデアミナーゼドメインは、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 (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); 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) 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 (その全内容が参照により本明細書に組み込まれる)。

【0184】

「検出」は、検出されるべき分析物の存在、不在または量を同定することを指す。1つの実施形態において、ポリヌクレオチドまたはポリペプチドにおける配列変化が検出される。別の実施形態では、インデル (indel) の存在が検出される。

【0185】

「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、後者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、酵素結合免疫吸着測定法（ELISA）で一般的に用いられるもの）、ビオチン、ジゴキシゲニン、またはハプテンが含まれる。

【0186】

「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。

【0187】

本明細書で使用される用語「有効量」は、所望の生物学的応答を誘導するために十分な生物学的活性剤の量を意味する。疾患の治療的処置のために本発明を実施するために使用される活性剤（複数可）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（例えば、プログラム可能DNA結合タンパク質、核酸塩基エディター、およびgRNA）の量である。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有効量、例えば、nCAs9ドメインおよび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）を含む核酸塩基エディターの有効量は、その核酸塩基エディターによって特異的に結合および編集される標的部位の編集を誘導するために十分な融合タンパク質の量を指し得る。一実施形態では、有効量は、治療効果（例えば、疾患またはその症状もしくは状態を低減または制御する）を達成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目的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目的遺伝子を変化させるだけでよい。

【0188】

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（例えばnCAs9ドメインおよびデアミナーゼドメイン（例えばアデノシンデアミナーゼ、シチジンデアミナーゼ）を含む核酸塩基エディターの有効量は、本明細書に記載される核酸塩基エディターによって特異的に結合され編集される標的部位の編集を誘導するのに十分な融合タンパク質の量を指す。当業者には理解されるように、作用物質（例えば融合タンパク質、ヌクレアーゼ、ハイブリッドタンパク質、タンパク質二量体、タンパク質（もしくはタンパク質二量体）とポリヌクレオチドとの複合体、またはポリヌクレオチド）の有効量は、例えば、望まれる生物学的応答、例えば編集されるべき特定のアルル、ゲノム、もしくは標的部位、標的化される細胞もしくは組織、および/または使用される作用物質などの種々の因子に応じて変化し得る。

【0189】

「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、参照核酸分子またはポリペプチドの全長の少なくとも約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。

【0190】

「ガイドRNA」あるいは「gRNA」は、標的配列に特異的となることができ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タンパク質（例えばCas9またはCpf1）と複合体を形成することができるポリヌクレオチドを意味する。一実施形態において、ガイドポリヌクレオチドはガイドRNA（gRNA）である。gRNAは2個以上のRNAの複合体として存在することもあれば、1個の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2つ以上の分子の複合体として存在するガイドRNA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、(

1) 標的核酸と相同性を共有するドメイン（例えば、標的へのCas9複合体の結合を指示する）；および(2) Cas9タンパク質に結合するドメインという二つのドメインを含む。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、いくつかの態様において、ドメイン(2)は、Jinek et al., Science 337:816-821(2012)（その全内容が参照により本明細書に組み入れられる）に提供されるようなtracrRNAと同一または相同である。gRNA（例えばドメイン2を含むもの）の他の例は、2013年9月6日出願された「Switchable Cas9 Nucleases and Uses Thereof」と題される米国仮特許出願U.S.S.N. 61/874,682および2013年9月6日出願された「Delivery System For Functional Nucleases」と題される米国仮特許出願U.S.S.N. 61/874,746に見出され得、それら各々の全内容が参照により本明細書に組み入れられる。いくつかの実施形態において、gRNAは、ドメイン(1)および(2)の二つ以上を含み、「伸長されたgRNA。」と称され得る。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における標的核酸と結合する。gRNAは、標的部位を相補するヌクレオチド配列を含み、これが標的部位へのヌクレアーゼ/RNA複合体の結合を媒介し、ヌクレアーゼ：RNA複合体の配列特異性を提供する。

【0191】

「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。

【0192】

用語「塩基修復の阻害因子 (inhibitor of base repair)」あるいは「IBR」は、核酸修復酵素、例えば塩基除去修復 (BER: base excision repair) 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 III、Endo IV、Endo V、Endo VIII、Fpg、hOGG1、hNEIL1、T7 Endo1、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Endo Vまたは触媒的に不活性なhAAGである。

【0193】

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ニングリコシラーゼ (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。

。

【0194】

「増加」とは、少なくとも10%、25%、50%、75%、または100%の正の変化を意味する。

【0195】

「インテイン (intein)」は、それ自身を切り出してそして残った断片 (エクステイン (extein)) をタンパク質スプライシングとして知られるプロセスにおいてペプチド結合で連結することができるタンパク質の断片である。インテインは「タンパク質イントロン」とも呼ばれる。インテインがそれ自身を切り出し、タンパク質の残りの部分を連結するプロセスは、本明細書において、「タンパク質スプライシング」または「インテイン媒介タンパク質スプライシング」と呼ばれる。ある態様において、前駆体タンパク質 (インテイン媒介タンパク質スプライシングの前のインテイン含有タンパク質) のインテインは、二つの遺伝子に由来する。そのようなインテインは、本明細書では分割 (split) インテインと呼ばれる (例えば、分割インテイン-Nおよび分割インテイン-C)。たとえば、シアノバクテリアでは、DNAポリメラーゼIIIの触媒サブユニットaであるDnaEは2つの別々の遺伝子dnaE-nおよびdnaE-cによってコードされている。dnaE-n遺伝子によってコードされるインテインは本明細書において「インテインN」と称され得る。dnaE-c遺伝子によってコードされるインテインは本明細書において「インテインC」と称され得る。

【0196】

他のインテイン系も使用され得る。例として、dnaEインテイン、すなわちCfa-N (例えば分割インテイン-N) およびCfa-C (例えば分割インテイン-C) のインテイン対に基づく合成インテインが記述されている (例えば、参照により本明細書に組み入れられるSteven et al., J Am Chem Soc. 2016 Feb. 24; 138(7):2162-5)。本開示に従って使用することができるインテイン対の非限定的な例としては、Cfa DnaEインテイン、Ssp GyrBインテイン、Ssp DnaXインテイン、Ter DnaE3インテイン、Ter ThyXインテイン、Rma DnaBインテイン、およびCne Prp8インテイン (例えば、参照により本明細書に組み込まれる米国特許第8,394,604号に記載のようなもの) が挙げられる。

【0197】

インテインの例示的なヌクレオチドおよびアミノ酸配列を提供する。

DnaE インテイン-N DNA: TGCCTGTCATACGAAACCGAGATACTGACAGTAGAATATGGCCTTCTGCCAATCGGGAAGATTGTGGAGAAACGGATAGAATGCACAGTTTACTCTGTGCGATAACAATGGTAACATTTATACTCAGCCAGTTGCCAGTGCGCACGACCGGGGAGAGCAGGAAGTATTCGAATACTGTCTGGAGGATGGAAGTCTCATTAGGGCCACTAAGGACCACAAATTTATGACAGTCGATGGCCAGATGCTGCCTATAGACGAAATCTTTGAGCGAGAGTTGGACCTCATGCGAGTTGACAACCTTCCTAAT

DnaE インテイン-N タンパク質: CLSYETEILTVEYGLLP I G K I V E K R I E C T V Y S V D N N G N I Y T Q P V A Q W H D R G E Q E V F E Y C L E D G S L I R A T K D H K F M T V D G Q M L P I D E I F E R E L D L M R V D N L P N

DnaE インテイン-C DNA: ATGATCAAGATAGCTACAAGGAAGTATCTTGGCAAACAAAA CGTTTATGA TATTGGAGTCGAAAGAGATCACAACCTTTGCTCTGAAGAACGGATTCATAGCTTCTAAT

Intein-C: MIKIATRKYLGKQNVYDIGVERDHN FALKNGFIASN

Cfa-N DNA: TGCCTGTCTTATGATACCGAGATACTTACCGTTGAATATGGCTTCTTGCC TATTGGAAAGATTGTGCGAAGAGAGAATTGAATGCACAGTATATACTGTAGACAAGAATG GTTTCGTTTACACACAGCCCATTTGCTCAATGGCACAATCGCGGCGAACAAGAAGTATTT GAGTACTGTCTCGAGGATGGAAGCATCATACGAGCAACTAAAGATCATAAATTCATGAC CACTGACGGGCAGATGTTGCCAATAGATGAGATATTCGAGCGGGGCTTGGATCTCAAAC AAGTGGATGGATTGCCA

Cfa-N タンパク質:

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

Cfa-C DNA: ATGAAGAGGACTGCCGATGGATCAGAGTTTGAATCTCCCAAGAAGAAGAG
GAAAGTAAAGATAATATCTCGAAAAAGTCTTGGTACCCAAAATGTCTATGATATTGGAG
TGGAGAAAGATCACAACTTCCTTCTCAAGAACGGTCTCGTAGCCAGCAAC
Cfa-C タンパク質: MKRTADGSEFESPCKKRKVKIISRKSLGTQNVYDIGVEKDHNFLLKN
GLVASN

【0198】

分割Cas9のN末端部と分割Cas9のC末端部とを接合するために、インティンNおよびインティンCをそれぞれ分割Cas9のN末端部および分割Cas9のC末端部に融合させ得る。例えば、いくつかの実施形態において、インティン-Nが、分割されたCas9のN-末端部分のC-末端に融合され、すなわち、N--[分割Cas9のN末端部分]-[インティン-N]--Cの構造を形成する。いくつかの実施形態において、インティン-Cが、分割されたCas9のC-末端部分のN-末端に融合され、すなわち、N-[インティン-C]--[分割Cas9のC末端部分]-Cの構造を形成する。インティンが融合されたところのタンパク質（例えば分割Cas9）を連結するためのインティン媒介性タンパク質スプライシングの機構は、例えば、参照により本明細書に組み入れられるShah et al., Chem Sci. 2014; 5(1):446-461に記載されているように、当該技術分野において公知である。インティンを設計および使用する方法は当該技術分野で公知であり、例えばWO2014004336、WO2017132580、US20150344549およびUS20180127780によって記載されており、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

【0199】

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。

【0200】

「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。

【0201】

「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。

【0202】

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（例えば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、シチジンデアミナーゼ、もしくはアデノシンデアミナーゼとシチジンデアミナーゼ）、またはnapDNABpドメイン（例えばCas12b）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もしく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。特定の実施形態では、リンカーは、Casタンパク質またはその断片内に挿入されたデアミナーゼドメインに隣接する。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。例えば、ある実施形態では、リンカーは、Cas12a/Cpf1、Cas12b/C2cl、Cas12c/C2c3、Cas12d/CasY、Cas12e/CasX、Cas12g、Cas12h、またはCas12iとデアミナーゼとを繋げ得る。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、塩基エディター系の脱アミノ化成分のRNA結合部分とnapDNABp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、塩基エディター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、またはプレクエオシン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、Sf Mu 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

10

20

30

40

50

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。

【0203】

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。

【0204】

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企図される。

【0205】

ある態様において、核酸塩基エディターのドメインは、SGGSSGSETPGTSESATPESSGGS, SGGSSGGSSGSETPGTSESATPESSGGSSGGS, またはGGSGGSPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGGSGGS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、核酸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、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、(GGGGS)_n、(G)_n、(EAAAK)_n、(GGG)_n、SGSETPGTSESATPES、もしくは(XP)_nモチーフ、またはこれらのいずれかの組合せを含み、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。いくつかの実施形態において、nは1、2、3、4、5、6、7、8、9、10、11、12、13、14、または15である。

【0206】

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、長さが40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SGGSSGGSSGSETPGTSESATPESSGGSSGGSSGGSSGGSSGSETPGTSESATPESSGGSSGGSSを含む。ある態様において、リンカーは、長さが64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SGGSSGGSSGSETPGTSESATPESSGGSSGGSSGGSSGGSSGSETPGTSESATPESSGGSSGGSS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。ある態様において、リンカーは、アミノ酸配列PGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSESを含む。

【0207】

「マーカー」とは、疾患または障害に関連する発現レベルまたは活性の変化を有する任

意のタンパク質またはポリヌクレオチドを意味する。

【0208】

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。

10

【0209】

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。

20

【0210】

「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に関し、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。

【0211】

「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日に出版され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、KKTELQTTNAENKTKKL、KRGINDRNFWRGENGRKTR、RKSGKIAAIVVKRPRK、PKKKRKV、またはMDSLLMNRKFLYQFKNVRWAKGRRETYLCを含む。

30

【0212】

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」、および「ポリヌクレオチド」は、ヌクレオチドのポリマー(例えば少なくとも3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

40

50

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'-フルオロリボース、リボース、2'-デオキシリボース、アラビノース、およびヘキソース)；および/または修飾リン酸基(例えばホスホロチオエートおよび5'-N-ホスホロアミダイト結合)であるか、またはそれらを含む。

【0213】

「核酸プログラミング可能DNA結合タンパク質」あるいは「napDNAbp」という用語は、「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」と互換的に使用され得、特異的な核酸配列にnapDNAbpをガイドするガイド核酸またはガイドポリヌクレオチド(例えばgRNA)のような核酸(例えばDNAまたはRNA)と会合するタンパク質を指す。ある態様において、ポリヌクレオチドによりプログラミング可能なヌクレオチド結合ドメインは、ポリヌクレオチドによりプログラミング可能なDNA結合ドメインである。ある態様において、ポリヌクレオチドによりプログラミング可能なヌクレオチド結合ドメインは、ポリヌクレオチドによりプログラミング可能なRNA結合ドメインである。ある態様において、ポリヌクレオチドによりプログラミング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に相補的な特異的DNA配列にCas9タンパク質を誘導するガイド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。核酸プログラミング可能DNA結合タンパク質の非限定的な例としては、Cas9(例えば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。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Csx12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、Cse2、Cse3、Cse4、Cse5e、Csc1、Csc2、Csa5、Csn1、Csn2、Csm1、Csm2、Csm3、Csm4、Csm5、Csm6、Cmr1、Cmr3、Cmr4、Cmr5、Cmr6、Csb1、Csb2、Csb3、Csx17、Csx14、Csx10、Csx16、CsaX、Csx3、Csx1、Csx15、Csx11、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、またはそれらの改変または操作さ

10

20

30

40

50

れたバージョンが挙げられる。他の核酸プログラミング可能なDNA結合タンパク質も、本開示に具体的に列記されていないかもしれないが、本開示の範囲内である。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を参照（それぞれの全内容が参照により本明細書に組み込まれる）。

【0214】

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。

【0215】

「核酸プログラム可能なDNA結合タンパク質」または「napDNAbp」という用語は、napDNAbpを特定の核酸配列に導くガイド核酸などの核酸 (例えば、DNAまたはRNA) と会合するタンパク質を指す。例えば、Cas12タンパク質は、Cas12タンパク質をガイドRNAに相補的な特異的DNA配列に導くガイドRNAと結合することができる。いくつかの実施形態において、napDNAbpは、Cas12ドメイン、例えば、ヌクレアーゼ活性Cas12ドメインである。napDNAbpsの例には、Cas12a/Cpf1、Cas12b/C2c1、Cas12c/C2c3、Cas12d/CasY、Cas12e/CasX、Cas12g、Cas12h、およびCas12iが含まれる。他のnapDNAbpも、本開示に具体的に列記されていないかもしれないが、本開示の範囲内である。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を参照（それぞれの全内容が参照により本明細書に組み込まれる）。

【0216】

「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、アデニンデアミナーゼもしくはアデノシンデアミナーゼ; またはシチジンデアミナ

ーゼもしくはシトシンデアミナーゼ)である。ある態様において、核酸塩基編集ドメインは、複数のデアミナーゼドメイン(例えば、アデニンデアミナーゼもしくはアデノシンデアミナーゼと、シチジンもしくはシトシンデアミナーゼ)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(WO 2018/027078)およびPCT/US2016/058344(WO 2017/070632)に記載されており、これらのそれぞれは、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:eaa04774 (2017)も参照(その全内容は、参照により本明細書に組み込まれる)。

10

【0217】

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、または他の方法で獲得することを含む。

20

【0218】

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。

【0219】

「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、疾患または障害と診断された、またはそれを有するリスクがある、またはそれを有することがあらかじめ決定された、またはそれを有することが疑われる患者として言及される。

30

【0220】

「病原性変異」、「病原性バリエーション」、「疾患原因変異」、「疾患原因バリエーション」、「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個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。

【0221】

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関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)。「賦形剤」、「担体」、「薬学的に許容される担体」、「ビヒクル」等のような用語は本明細書で互換的に使用される。

40

【0222】

50

「医薬組成物」という用語は、医薬的使用のために製剤化された組成物を表し得る。

【0223】

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。

【0224】

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。

【0225】

「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(napDNAbp)」という用語は、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を特定の核酸配列に導くガイドポリヌクレオチド(例えば、ガイドRNA)などの核酸(例えば、DNAまたはRNA)と結合するタンパク質を指す。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。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。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ム可能なヌクレオチド結合ドメインは、Cas12タンパク質である。

10

【0226】

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。

20

【0227】

「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。

【0228】

「参照(reference)」は、基準または対照の条件を意味する。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。他の実施形態において、限定されないが、参照は、試験条件に晒されていないか、またはプラセボもしくは通常の食塩水、培地、緩衝液、および/または、目的のポリヌクレオチドを保持しない対照ベクターに晒される、非処理細胞である。

30

【0229】

「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、少なくとも約20アミノ酸、少なくとも約25アミノ酸、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、少なくとも約60ヌクレオチド、少なくとも約75ヌクレオチド、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。いくつかの実施形態において、参照配列は、目的タンパク質の野生型配列である。他の実施形態では、参照配列は、野生型タンパク質をコードするポリヌクレオチド配列である。

40

【0230】

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。gRNAは2個以上のRNAの複合体として存在することもあれば、1個の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

50

(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2つ以上の分子の複合体として存在するガイドRNA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、(1) 標的核酸と相同性を共有するドメイン（例えば、標的へのCas9複合体の結合を指示する）；および(2) Cas9タンパク質に結合するドメインという二つのドメインを含む。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、いくつかの態様において、ドメイン(2)は、Jinek et al., Science 337:816-821(2012)（その全内容が参照により本明細書に組み入れられる）に提供されるようなtracrRNAと同一または相同である。gRNA（例えばドメイン2を含むもの）の他の例は、2013年9月6日出願された「Switchable Cas9 Nucleases and Uses Thereof」と題される米国仮特許出願U.S.S.N. 61/874,682および2013年9月6日出願された「Delivery System For Functional Nucleases」と題される米国仮特許出願U.S.S.N. 61/874,746に見出され得、それら各々の全内容が参照により本明細書に組み入れられる。いくつかの実施形態において、gRNAは、ドメイン(1) および (2) の二つ以上を含み、「伸長された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、例えば二つ以上のCas9タンパク質と結合し、二つ以上の異なる領域における標的核酸と結合する。gRNAは、標的部位を相補するヌクレオチド配列を含み、これが標的部位へのヌクレアーゼ/RNA複合体の結合を媒介し、ヌクレアーゼ：RNA複合体の配列特異性を提供する。

【0231】

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 (Cas9) である（例えば、"Complete genome sequence of an M1 strain of Streptococcus pyogenes." Ferretti J.J. et al., 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. et al., Nature 471:602-607(2011)参照）。

【0232】

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RNA: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)。

【0233】

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するが、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)に存在する。ある態様にお

10

20

30

40

50

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は一塩基改変とも呼ばれ得る。

【0234】

「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、核酸プログラミング可能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。

【0235】

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)。

【0236】

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。一実施形態では、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。

【0237】

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

10

20

30

40

50

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、好ましくは約30 mM未満NaClおよび3 mMクエン酸三ナトリウムであり、最も好ましくは約15 mM未満 NaClおよび1.5 mMクエン酸三ナトリウムである。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。一実施形態において、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。

【0238】

「分割 (split)」とは、2つ以上の断片に分割されることを意味する。

【0239】

「分割されたCas9タンパク質」あるいは「分割Cas9」とは、二つの別個のヌクレオチド配列によってコードされるN末端断片およびC末端断片として提供されるCas9タンパク質をいう。Cas9タンパク質のN末端部分およびC末端部分に対応するポリペプチドは、スプライスされて「再構成」Cas9タンパク質を形成することができる。特定の実施形態において、Cas9タンパク質は、例えば、Nishimasu et al., Cell, Volume 156, Issue 5, p. 935-949, 2014に記載されるように、またはJiang et al. (2016) Science 351: 867-871, PDB file: 5F9Rに記載されるように（それぞれ参照により本明細書に組み込まれる）、タンパク質の非秩序的な領域内で2つの断片に分割される。いくつかの実施形態において、タンパク質は、SpCas9のおよそアミノ酸A292-G364、F445-K483、もしくはE565-T637の間の領域内の任意のC、T、A、もしくはSにおいて、または任意の他のCas9、Cas9バリエーション（例えばnCas9、dCas9）、もしくは他のnapDNAbpにおける対応位置において、二つの断片に分割される。ある態様において、タンパク質は、SpCas9 T310、T313、A456、S469、またはC574において2つの断片に分割される。いくつかの実施形態において、タンパク質を2つの断片に分割するプロセスは、タンパク質を「分割（スプリッティング）する」と称される。

【0240】

他の実施形態において、Cas9タンパク質のN末端部分は、*S. pyogenes* Cas9野生型（SpCas9）（NCBI参照配列：NC_002737.2、Uniprot参照配列：Q99ZW2）のアミノ酸1～573または1～637を含み、Cas9タンパク質のC末端部分は、SpCas9野生型のアミノ酸574～1368または638～1368の部分を含む。

【0241】

分割されたCas9のC末端部分は、分割されたCas9のN末端部分と連結されて完全なCas9タンパク質を形成することができる。いくつかの実施形態において、Cas9タンパク質のC末端部分は、Cas9タンパク質のN末端部分が終わるところから始まる。このように、いくつかの実施態様において、分割Cas9のC末端部分は、spCas9のアミノ酸(551-651)-1368の部分を含む。「(551-651)-1368」とは、アミノ酸551～651（両端を含む）のあいだのアミノ酸から始まり、アミノ酸1368で終わることを意味する。例えば、分割Cas9のC末端部分は、spCas9のアミノ酸551-1368、552-1368、553-1368、554-1368

、555-1368、556-1368、557-1368、558-1368、559-1368、560-1368、561-1368、562-1368、563-1368、564-1368、565-1368、566-1368、567-1368、568-1368、569-1368、570-1368、571-1368、572-1368、573-1368、574-1368、575-1368、576-1368、577-1368、578-1368、579-1368、580-1368、581-1368、582-1368、583-1368、584-1368、585-1368、586-1368、587-1368、588-1368、589-1368、590-1368、591-1368、592-1368、593-1368、594-1368、595-1368、596-1368、597-1368、598-1368、599-1368、600-1368、601-1368、602-1368、603-1368、604-1368、605-1368、606-1368、607-1368、608-1368、609-1368、610-1368、611-1368、612-1368、613-1368、614-1368、615-1368、616-1368、617-1368、618-1368、619-1368、620-1368、621-1368、622-1368、623-1368、624-1368、625-1368、626-1368、627-1368、628-1368、629-1368、630-1368、631-1368、632-1368、633-1368、634-1368、635-1368、636-1368、637-1368、638-1368、639-1368、640-1368、641-1368、642-1368、643-1368、644-1368、645-1368、646-1368、647-1368、648-1368、649-1368、650-1368、または651-1368のいずれか1つの部分を含み得る。いくつかの実施形態において、分割されたCas9タンパク質のC末端部分は、SpCas9のアミノ酸574-1368または638-1368の部分を含む。

【0242】

「対象」とは、哺乳動物を意味し、ヒト、または、ウシ、ウマ、イヌ、ヒツジ、もしくはネコのような非ヒト哺乳動物を含むが、これらに限定されない。対象には、家畜、労働力を生み食料のような商品を提供するために飼育される飼育動物（ウシ、ヤギ、ニワトリ、ウマ、ブタ、ウサギ、およびヒツジを含むがこれらに限定されない）が含まれる。

【0243】

「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。一実施形態では、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%、80%もしくは85%、90%、95%または99%までもの同一性を有する。

【0244】

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、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.

【0245】

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（例えばシチジンまたはアデニンデアミナーゼ）またはそれを含む融合タンパク質によって脱アミノ化される。

【0246】

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、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、疾患および/またはそれ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、または治癒する。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。

【0247】

「ウラシルグリコシラーゼ阻害因子」、あるいは「UGI」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。一実施形態において、UGIは、ウラシル-DNAグリコシラーゼ塩基除去修復酵素を阻害することができるタンパク質、その断片、またはドメインである。いくつかの態様において、UGIドメインは、野生型UGIまたはその改変バージョンを含む。いくつかの実施形態において、UGI断片は、下記に提示される例示的アミノ酸配列の断片を含む。いくつかの実施形態において、UGI断片は、下記に提供される例示的UGI配列の少なくとも60%、少なくとも65%、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、または100%を含むアミノ酸配列を含む。いくつかの実施形態において、UGIは、以下に記載されるように、例示的UGIアミノ酸配列またはその断片に対して相等的なアミノ酸配列を含む。いくつかの実施態様において、UGIまたはその一部は、以下に記載されるように、野生型UGIもしくはUGI配列またはその一部に対して少なくとも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%、少なくとも99.5%、少なくとも99.9%または100%の同一性を有する。例示的なUGIは、以下のアミノ酸配列を含む：

```
> splP14739IUNGI_BPPB2 Uracil-DNA glycosylase inhibitor
MTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMMLTSD
APEYKPWALVIQDSNGENKIKML.
```

【0248】

「ベクター」という用語は、核酸配列を細胞に導入して形質転換された細胞をもたらす手段を指す。ベクターには、プラスミド、トランスポゾン、ファージ、ウイルス、リボソーム、およびエピソームが含まれる。「発現ベクター」は、レシピエント細胞で発現されるヌクレオチド配列を含む核酸配列である。発現ベクターは、開始配列、停止配列、エンハンサー配列、プロモーター配列、および分泌配列などの、導入された配列の発現を増進および/または促進するための追加の核酸配列を含み得る。

【0249】

本明細書で提供される組成物または方法は、本明細書で提供される他の組成物および方

10

20

30

40

50

法のいずれか1つまたは複数と組み合わせることができる。

【0250】

DNA編集は、病原性突然変異を遺伝子レベルで修正することによって病状を修正するための実行可能な手段として浮上してきた。最近まで、すべての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）を誘導し内因性DNA修復経路に依存して半確率的態様で産物帰結を決定することにより機能し、遺伝的産物の複雑な集団をもたらしていた。相同性誘導型修復（HDR）経路を介して正確でユーザー定義された修復結果を達成することはできるが、多くの課題により、治療に関連する細胞型でのHDRを使用した高効率修復が妨げられている。実際にはこの経路は、競合する、エラーが発生しやすい非相同末端結合経路に比べて非効率的である。さらに、HDRは細胞周期のG1期とS期とに厳密に制限されており、有糸分裂後の細胞におけるDSBの正確な修復を妨げている。その結果、これらの集団においてユーザー定義されかつプログラム可能な方法で高効率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が証明されている。

10

【0251】

本明細書で提供される組成物または方法は、本明細書で提供される他の組成物および方法のいずれか1つまたは複数と組み合わせることができる。

【0252】

DNA編集は、病原性突然変異を遺伝子レベルで修正することによって病状を修正するための実行可能な手段として浮上してきた。最近まで、すべての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）を誘導し内因性DNA修復経路に依存して半確率的態様で産物帰結を決定することにより機能し、遺伝的産物の複雑な集団をもたらしていた。相同性誘導型修復（HDR）経路を介して正確でユーザー定義された修復結果を達成することはできるが、多くの課題により、治療に関連する細胞型でのHDRを使用した高効率修復が妨げられている。実際にはこの経路は、競合する、エラーが発生しやすい非相同末端結合経路に比べて非効率的である。さらに、HDRは細胞周期のG1期とS期とに厳密に制限されており、有糸分裂後の細胞におけるDSBの正確な修復を妨げている。その結果、これらの集団においてユーザー定義されかつプログラム可能な方法で高効率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が証明されている。

20

【図面の簡単な説明】

【0253】

【図1】図1 A-1 Cはプラスミドを示す。図1 Aは、*TadA7.10-dCas9*塩基エディターをコードする発現ベクターである。図1 Bは、クロラムフェニコール耐性（*CamR*）およびスペクチノマイシン耐性（*SpectR*）を付与するタンパク質をコードする核酸分子を含むプラスミドである。このプラスミドはまた、2つの点突然変異によって不能にされたカナマイシン耐性遺伝子を含む。図1 Cは、クロラムフェニコール耐性（*CamR*）およびスペクチノマイシン耐性（*SpectR*）を付与するタンパク質をコードする核酸分子を含むプラスミドである。このプラスミドはまた、3つの点突然変異によって不能にされたカナマイシン耐性遺伝子を含む。

30

【図2】図2は、図1に示された発現ベクターで形質導入された細菌コロニーのイメージであり、これは非機能的カナマイシン耐性遺伝子を含む。ベクターは、エラープローンPCRを用いて生成されたABE7.10バリエーションを含んだ。これらの「進化された」ABE7.10バリエーションを発現する細菌細胞を、増加するカナマイシンの濃度を用いてカナマイシン耐性について選択した。アデノシンデアミナーゼ活性を有するABE7.10バリエーションを発現する細菌は、カナマイシン耐性遺伝子に導入された変異を修正し、カナマイシン耐性を回復させることができた。カナマイシン耐性細胞をさらなる分析のために選択した。

40

【図3】図3 Aおよび3 Bは、胎児ヘモグロビンの上方調節について治療的に関連する部位である、ヘモグロビンサブユニットガンマ（HGB1）遺伝子座の調節領域の編集を示す。図3 Aは、HGB1遺伝子の調節領域の一部の図である。図3 Bは、アデノシンデアミナーゼバリエーションの効率および特異性を定量化する。編集はHEK293T細胞のヘモグロビンサブユニットガンマ1（HGB1）遺伝子座でアッセイされ、それは胎児ヘモグロビンのアップレギ

50

レーションについて治療的に関連する部位である。上部パネルは、HGB1遺伝子の調節配列の標的領域におけるヌクレオチド残基を示す。A5、A8、A9、およびA11は、HGB1における編集されたアデノシン残基を示す。

【図4】図4は、非カノニカル（非正準）PAM配列を認識するdCas9を含むアデノシン塩基エディターの相対的有効性を示す。上のパネルはヘモグロビンのサブユニットのコード配列を示している。下のパネルは、さまざまな長さのガイドRNAを有するアデノシンデアミナーゼバリエーション塩基エディターの効率を示すグラフである。

【図5】図5は、ABE8塩基エディターの効率および特異性を示すグラフである。意図された標的（target）ヌクレオチドおよび意図されていない標的ヌクレオチド（バイスタンダー（bystander））における編集のパーセントが定量化されている。

【図6】図6は、ABE8塩基エディターの効率および特異性を示すグラフである。意図された標的ヌクレオチドおよび意図されていない標的ヌクレオチド（バイスタンダー）における編集のパーセントが定量化されている。

【図7】図7 A~7 Dは、第8世代アデニン塩基エディターがヒト細胞において、より優れたA・TからG・Cへの変換（conversion）を媒介することを示す。図7 Aは、アデニン塩基編集の概要を示す：i) ABE8は、ゲノム中のsgRNA標的化部位においてRループを創出する；ii) TadA*デアミナーゼは、Rループのss-DNA部分の加水分解脱アミノ化（Hydrolytic deamination）を介してアデニンをイノシンに化学的に変換する；iii) Cas9のD10Aニックアーゼはイノシン含有鎖の反対側の鎖にニックを入れる；iv) イノシン含有鎖は、DNA複製中に鋳型として使用され得る；v) DNAポリメラーゼ（DNA Polymerase）の場合、イノシンはシトシンと優先的に塩基対を形成する；vi) 複製後、イノシンはグアノシンに置き換えられる。図7 Bは、ABE8.x-mおよびABE8.x-dのアーキテクチャを示す。図7 Cは、tRNA Arg2 (PDB 2B3J) と複合体を形成した*S. aureus* TadA (図示せず) と整列した*E. coli* TadAデアミナーゼ (PDB 1Z3A) の3つの透視図を示す。8ラウンドの進化で同定された変異を強調している。図7 Dは、Hek293T細胞中の8つのゲノム部位にわたる、ABE7.10構築物に対するコアABE8構築物のA・TからG・C塩基編集効率を示すグラフである。値とエラーバーは、異なる日に行われた3つの独立した生物学的複製の平均とs.d.を反映している。

【0254】

【図8】図8 A~図8 Cは、ヒト細胞においてCas9 PAMバリエーションABE8と触媒的に不活性なCas9 ABE8バリエーションが、対応するABE7.10バリエーションよりも高いA・T→G・C変換（conversion）を媒介することを示す。値とエラーバーは、異なる日に行われた3つの独立した生物学的複製の平均とs.d.を反映している。図8 Aは、NG-Cas9 ABE8s (-NG PAM) を有するHek293T細胞におけるA・TからG・Cへの変換を示すグラフである。図8 Bは、Sa-Cas9 ABE8s (-NNGRRT PAM) を有するHek293T細胞におけるA・TからG・Cへの変換を示すグラフである。図8 Cは、触媒的に不活性化されたdCas9-ABE8s (*S. pyogenes* Cas9のD10A, H840A) を有するHek293T細胞におけるA・TからG・Cへの変換を示すグラフである。

【図9】図9 A-9 Eは、Hek293T細胞におけるABE7.10、ABEmaxおよび1つのBPNLSを有するABEmaxの間のオンターゲットおよびオフターゲット編集頻度の比較を示す。異なる日に実施したn=3の独立した生物学的反復試験について、個々のデータポイントが表示され、エラーバーはs.d.を表す。図9 Aおよび図9 Bは、標的上のDNA編集頻度を示すグラフである。図9 Bおよび図9 Cは、sgRNA誘導によるDNAオフターゲット編集の頻度を示すグラフである。図9 Eは、RNAオフターゲット編集頻度を示すグラフである。

【図10】図10 A-10 Bは、HEK293T細胞における、TadA、C末端αヘリックス切詰型ABE構築物のA・TからG・Cへの変換中央値および対応するインデル（INDEL）形成を示す。図10 Aは、8つのゲノム部位（Site）にわたるA・TからG・Cへの編集変換中央値を示すヒートマップである。図10 Bは、インデル（Indel）形成を示すヒートマップである。デルタ（delta）残基値はTadAにおける欠失位置に対応する。n=3生物学的複製から生じた中央値。

10

20

30

40

50

【図11】図11は、HEK293T細胞中の8つのゲノム部位にわたる、40個のABE8構築物のA・TからG・Cへの変換中央値を示すヒートマップである。中央値は2つ以上の生物学的反復から決定した。

【図12】図12は、HEK293T細胞中の8つのゲノム部位にわたる40個のABE8構築物のインデル%中央値を示すヒートマップである。中央値は2つ以上の生物学的反復から決定した。

【図13】図13は、編集における変化倍数 (Fold change in editing) を示すグラフである (ABE8:ABE7)。Hek293T細胞における8つの異なるゲノム部位の標的内の全てのA位置にわたる平均のABE8:ABE7 A・T→G・C編集の表示。位置 (position) 2~12は、20 ntプロトSpacer内の標的アデニンの位置を示し、位置20が-NGG PAMのすぐ5'にある。

10

【図14】図14は、ABE8のデンドログラムを示す。さらなる研究のために選択されたコアABE8構築物を黒でハイライトした。

【図15】図15は、HEK293T細胞における8つのゲノム部位 (site) にわたる、コアとなる8つのABE8構築物のA・TからG・Cへの変換中央値を示すヒートマップである。中央値は3つ以上の生物学的反復試験から決定した。

【0255】

【図16】図16は、HEK293T細胞中の8ゲノム部位 (Site) において試験されたコアとなる8つのABE8のインデル (Indel) 頻度中央値を示すヒートマップである。

【図17】図17は、HEK293T細胞における六つのゲノム部位 (Site) でのコアNG-ABE8構築物9 (-NG PAM) のA・TからG・Cへの変換中央値を示すヒートマップである。n=3生物学的複製から生じた中央値。

20

【図18】図18は、HEK293T細胞中の6つのゲノム部位 (Site) で試験されたコアNG-ABE8のインデル頻度中央値を示すヒートマップである。n=3生物学的複製から生じた中央値。

【図19】図19は、HEK293T細胞における六つのゲノム部位 (Site) でのコアSa-ABE8構築物 (-NNGRRT PAM) のA・TからG・Cへの変換中央値を示すヒートマップである。22 ntプロトSpacer内のサイト位置には、-2から20まで (5' から3') の番号が付けられている。位置20は、NNGRRT PAMの5'にある。n=3生物学的複製から生じた中央値。

【図20】図20は、HEK293T細胞中の8ゲノム部位 (Site) において試験されたコアsa-ABE8のインデル (Indel) 頻度の中央値を示すヒートマップである。n=3生物学的複製から生じた中央値。

30

【図21】図21は、HEK293T細胞における8つのゲノム部位 (Site) でのコアdC9-ABE8-m構築物の中央値A・TからG・Cへの変換を示すヒートマップである。Dead Cas9 (dC9) はS.pyogenes Cas9内のD10AおよびH840A変異と定義される。生物学的反復回数が3回以上の場合の中央値。

【図22】図22は、HEK293T細胞における8つのゲノム部位 (Site) でのコアdC9-ABE8-d構築物のA・TからG・Cへの変換中央値を示すヒートマップである。Dead (不活) Cas9 (dC9) はS. pyogenes Cas9内のD10AおよびH840A変異と定義される。n≥3の生物学的反復から生じた中央値。

40

【図23】図23 Aおよび23 Bは、HEK293T細胞中の8つのゲノム部位 (Site) で試験したコアdC9-ABE8のインデル (Indel) 頻度中央値を示す。n≥3の生物学的反復から生じた中央値。図23 Aは、ABE7.10と比較してdC9-ABE8-mバリエーションについて示されたインデルを示すヒートマップである。図23 Bは、ABE7.10と比較してdC9-ABE8-dバリエーションについて示されたインデルを示すヒートマップである。

【図24】図24は、ABE8およびABE7.10で処理されたHek293T細胞によるC・GからT・Aへの編集を示すグラフである。標的内のすべてのC位置にわたり平均化された各部位 (Site) の編集頻度。プロトSpacer内のシトシンを影付きで示している。

【図25】図25 A-25 Hは、ABE8構築物、およびDNAに対する特異性を改善するためのTadA変異を有するABE8構築物による、DNAオンターゲット編集およびsgRNA依存性DN

50

Aオフターゲット編集を示す。個々のデータポイントが表示され、エラーバーは異なる日に実施したn=3の独立した生物学的反復試験についてのs.d.を表す。図25 Aおよび25 Bは、ABE7と比較したコアABE8構築物についてのオンターゲットDNA編集頻度を示すグラフである。図25 Cおよび25 Dは、RNAオフターゲット編集を改善する突然変異を有するABE8についてのオンターゲットDNA編集頻度を示すグラフである。図25 Eおよび25 Fは、ABE7と比較したコアABE8構築物についての、sgRNA誘導によるDNAオフターゲット編集頻度をグラフである。図25 Gおよび25 Hは、RNAオフターゲット編集を改善する突然変異を有するABE8構築物についての、gRNA誘導によるDNAオフターゲット編集頻度を示すグラフである。

【0256】

【図26】図26は、ヒト細胞において、以前に同定された12個のsgRNA依存性Cas9オフターゲット遺伝子座におけるインデル頻度を示すグラフであり、個々のデータ点が示され、エラーバーは、異なる日に実施したn=3の独立した生物学的反復試験についてのs.d.を表す。

【図27】図27 Aおよび27 Bは、初代細胞におけるA・TからG・Cへの変換（conversion）および表現型の結果を示す。図27 Aは、二つの別々のドナー（donor）からのABE処理CD34+細胞における-198 HBG1/2部位でのA・TからG・Cへの変換を示すグラフである。処理後48および144時間で実施したNGS分析。-198 HBG1/2標的配列を、A7を強調表示して示す。A7についてプロットしたA・T→G・Cパーセント。図27 Bは、アルファ（alpha）グロビンに対する割合としての、形成されたガンマ（gamma）グロビンのパーセンテージを示すグラフである。ABE処理および赤血球分化の後の、異なる2ドナーから示された値。

【図28】図28 Aおよび28 Bは、ABE8で処理したCD34+細胞の、HBG1/2の上流の-198プロモーター部位でのA・T→G・C変換を示す。図28 Aは、二人のドナー（donor）からのCD34+細胞における、エディター処理後48時間および144時間でのABE8のA→G編集頻度を示すヒートマップであり、ここで、ドナー2は、鎌状赤血球症についてヘテロ接合である。図28 Bは、A7のみの編集または(A7+A8)組み合わせ編集のいずれかを含む全配列決定リード（reads）の分布のグラフ表示である。

【図29】図29は、ガンマグロビンプロモーターの-198部位における、ABE8処理CD34+細胞のインデル（Indel）頻度を示すヒートマップである。48時間および144時間時点における、二人のドナー（Donor）から示された頻度。

【図30】図30は、未処理の分化CD34+細胞（ドナー1）のUHPLC UV-Visトレース（220 nm）およびグロビン鎖レベルの積分を示す。

【図31】図31は、ABE7.10-mで処理した分化CD34+細胞（ドナー1）のUHPLC UV-Visトレース（220 nm）およびグロビン鎖レベルの積分を示す

【図32】図32は、ABE7.10-dで処理した分化CD34+細胞（ドナー1）のUHPLC UV-Visトレース（220 nm）およびグロビン鎖レベルの積分を示す。

【図33】図33は、ABE8.8-mで処理した分化CD34+細胞（ドナー1）のUHPLC UV-Visトレース（220 nm）およびグロビン鎖レベルの積分を示す

【0257】

【図34】図34は、ABE8.8-dで処理した分化CD34+細胞（ドナー1）のUHPLC UV-Visトレース（220 nm）およびグロビン鎖レベルの積分を示す。

【図35】図35は、ABE8.13-mで処理した分化CD34+細胞（ドナー1）のUHPLC UV-Visトレース（220 nm）およびグロビン鎖レベルの積分を示す。

【図36】図36は、ABE8.13-dで処理した分化CD34+細胞（ドナー1）のUHPLC UV-Visトレース（220 nm）およびグロビン鎖レベルの積分を示す。

【図37】図37は、ABE8.17-mで処理した分化CD34+細胞（ドナー1）のUHPLC UV-Visトレース（220 nm）およびグロビン鎖レベルの積分を示す。

【図38】図38は、ABE8.17-dで処理した分化CD34+細胞（ドナー1）のUHPLC UV-Visトレース（220 nm）およびグロビン鎖レベルの積分を示す。

【図39】図39は、ABE8.20-mで処理した分化CD34+細胞(ドナー1)のUHPLC UV-Visトレース(220 nm)およびグロビン鎖レベルの積分を示す。

【図40】図40は、ABE8.20-dで処理した分化CD34+細胞(ドナー1)のUHPLC UV-Visトレース(220 nm)およびグロビン鎖レベルの積分を示す。

【図41】図41は、未処理の分化CD34+細胞(ドナー2)のUHPLC UV-Visトレース(220 nm)およびグロビン鎖レベルの積分を示す。注：ドナー2は鎌状赤血球症についてヘテロ接合体である。

【図42】図42は、ABE7.10-mで処理した分化CD34+細胞(ドナー2)のUHPLC UV-Visトレース(220 nm)およびグロビン鎖レベルの積分を示す。注：ドナー2は鎌状赤血球症についてヘテロ接合体である。

【図43】図43は、ABE7.10-dで処理した分化CD34+細胞(ドナー2)のUHPLC UV-Visトレース(220 nm)およびグロビン鎖レベルの積分を示す。注：ドナー2は鎌状赤血球症についてヘテロ接合体である。

【図44】図44は、ABE8.8-mで処理した分化CD34+細胞(ドナー2)のUHPLC UV-Visトレース(220 nm)およびグロビン鎖レベルの積分を示す。注：ドナー2は鎌状赤血球症についてヘテロ接合体である。

【図45】図45は、ABE8.8-dで処理した分化CD34+細胞(ドナー2)のUHPLC UV-Visトレース(220 nm)およびグロビン鎖レベルの積分を示す。注：ドナー2は鎌状赤血球症についてヘテロ接合体である。

【図46】図46は、ABE8.13-mで処理した分化CD34+細胞(ドナー2)のUHPLC UV-Visトレース(220 nm)およびグロビン鎖レベルの積分を示す。注：ドナー2は鎌状赤血球症についてヘテロ接合体である。

【0258】

【図47】図47は、ABE8.13-dで処理した分化CD34+細胞(ドナー2)のUHPLC UV-Visトレース(220 nm)およびグロビン鎖レベルの積分を示す。注：ドナー2は鎌状赤血球症についてヘテロ接合体である。

【図48】図48は、ABE8.17-mで処理した分化CD34+細胞(ドナー2)のUHPLC UV-Visトレース(220 nm)およびグロビン鎖レベルの積分を示す。注：ドナー2は鎌状赤血球症についてヘテロ接合体である。

【図49】図49は、ABE8.17-dで処理した分化CD34+細胞(ドナー2)のUHPLC UV-Visトレース(220 nm)およびグロビン鎖レベルの積分を示す。注：ドナー2は鎌状赤血球症についてヘテロ接合体である。

【図50】図50は、ABE8.20-mで処理した分化CD34+細胞(ドナー2)のUHPLC UV-Visトレース(220 nm)およびグロビン鎖レベルの積分を示す。注：ドナー2は鎌状赤血球症についてヘテロ接合体である。

【図51】図51 A~51 Eは、2つの独立した部位でのABE8.8による編集が、赤血球分化後11日目の除核前には90%を超える編集に達し、赤血球分化後18日目には、アルファグロビンまたは全ベータファミリーグロビンに対して約60%のガンマグロビンに達したことを示す。図51 Aは、2つの独立した実験において、2人の健康なドナーにおけるABE8.8編集の平均を示すグラフである。編集効率はHBG1とHBG2を区別するプライマーで測定した。図51 Bは、2つの独立した実験における1人の健康なドナーの平均を示すグラフである。編集効率はHBG1とHBG2の両方を認識するプライマーで測定した。図51 Cは、ヘテロ接合E6V突然変異を有するドナーにおけるABE8.8の編集を示すグラフである。図51 Dおよび51 Eは、ABE8.8編集済み細胞におけるガンマグロビンの増加を示すグラフである。

【図52】図52 Aおよび52 Bは、鎌状赤血球突然変異を修正するためのABEバリエーションを用いた編集のパーセント(%Editing)を示す。図52 Aは、SCD患者線維芽細胞において約70%の編集を有する異なるエディターバリエーションのスクリーニングを示すグラフである。図52 Bは、リードABEバリエーションで編集された、健康ドナー由来のCD34細胞を示すグラフであり、編集ウィンドウ内に存在しSCD突然変異の編集についてのプロキシとして機能する隣接プロリンにおける同義突然変異A13を標的としている。ABE8バリエーションは

10

20

30

40

50

プロキシA13において約40%の平均編集頻度を示した。

【図53】図53 Aおよび53 Bは、ABE処理と関連するRNAにおける細胞内A→I編集を検出するためのRNAアンプリコン配列決定を示す。個々のデータポイントが表示され、エラーバーは、異なる日に実施したn=3の独立した生物学的反復試験についてのs.d.を表す。図53 Aは、ABE7およびCas9 (D10A) ニッカーゼコントロールと比較した、コアABE8構築物についての標的RNAアンプリコンにおけるA→I編集頻度を示すグラフである。図53 Bは、RNAオフターゲット編集を改善することが報告されている変異を有するABE8についての、標的RNAアンプリコンにおけるA→I編集頻度を示すグラフである。

【0259】

【図54】図54は、パーキンソン病におけるドーパミン作動性ニューロン喪失から生じるドーパミンの喪失を示す概略図である。

【図55】図55は、パーキンソン病に関連するLRRK2におけるR1441CおよびR1441H突然変異の補正のためのガイド (Guide) RNAおよび標的配列を示す概略図である。

【図56】図56は、パーキンソン病に関連するLRRK2におけるY1699C、G2019S、およびI2020突然変異の補正のための標的配列を示す概略図である。

【図57】図57 A～図57 Cは、グラフ、模式図及び表を提供する。図57 Aは、LRRK2標的配列の核酸位置7におけるAからGへの変換パーセントを定量する。使用されたエディターはPV1～PV14と呼ばれ、これについては以下に説明する。pCMVはCMVプロモーターを示す；bpNLSは二部分 (bipartite) 核局在化シグナルを示す；monoABE8.1は、モノマー型のABE8.1塩基エディターを示す。図57 Bは、パーキンソン病に関連するLRRK2におけるR1441C突然変異の修正のための標的配列およびガイド (Guide) RNAを示す。図57 Cは、LRRK2標的配列の核酸位置7におけるAからGへの変換パーセントを示す。エディターPV1～14を用いてLRRK2 R1441Cを編集した。エディター (15～28) を用いてG2109を編集した。LRRK2突然変異の修正に用いられたエディター (PV1～28) は以下の通りである。PV1 (PV15とも呼ばれる). pCMV_monoABE8.1_bpNLS + Y147TMSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNLRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCTFFRMPRQVFNAQKKAQSSTD PV2 (PV16とも呼ばれる). pCMV_monoABE8.1_bpNLS + Y147RMSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNLRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCRFFRMPRQVFNAQKKAQSSTD PV3 (PV17とも呼ばれる). pCMV_monoABE8.1_bpNLS + Q154SMSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNLRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRSVFNAQKKAQSSTD PV4 (PV18とも呼ばれる). pCMV_monoABE8.1_bpNLS + Y123HMSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNLRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTD PV5 (PV19とも呼ばれる). pCMV_monoABE8.1_bpNLS + V82SMSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNLRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYSTTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTD PV6 (PV20とも呼ばれる). pCMV_monoABE8.1_bpNLS + T166RMSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNLRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSRD PV7 (PV21とも呼ばれる). pCMV_monoABE8.1_bpNLS + Q154RMSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNLRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRRVFNAQKKAQSSTD PV8 (PV22とも呼ばれる). pCMV_m

10

20

30

40

50

onoABE8.1_bpNLS + Y147R_Q154R_Y123HMSEVEFSHEYWMRHALTLAKRARDER
 EVPVGAVLVNLRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFE
 PCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAAL
 LCRFFRMPRRVFNAQKKAQSSTD PV9 (PV23とも呼ばれる). pCMV_monoABE8.1_bp
 NLS + Y147R_Q154R_I76YMSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVNLRV
 RIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLYDLYVTFEPCVMCAGAMIH
 SRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCRFFRMPRRVF
 NAQKKAQSSTD PV10 (PV24とも呼ばれる). pCMV_monoABE8.1_bpNLS + Y147R_
 Q154R_T166RMSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVNLRVIGEGWNR
 AIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGV
 RNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCRFFRMPRRVFNAQKKAQSS
 RD PV11 (PV25とも呼ばれる). pCMV_monoABE8.1_bpNLS + Y147T_Q154RMSEV
EFSHEYWMRHALTLAKRARDEREVPVGAVLVNLRVIGEGWNRAIGLHDPTAHAEIMA
 LRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVL
 HYPGMNHRVEITEGILADECAALLCTFFRMPRRVFNAQKKAQSSTD PV12 (PV26とも
 呼ばれる). pCMV_monoABE8.1_bpNLS + Y147T_Q154SMSEVEFSHEYWMRHALTL
 AKRARDEREVPVGAVLVNLRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLID
 ATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGI
 LADECAALLCTFFRMPRRVFNAQKKAQSSTD PV13 (PV27とも呼ばれる). pCMV_mon
 oABE8.1_bpNLS + H123Y123H_Y147R_Q154R_I76YMSEVEFSHEYWMRHALTLA
 KRARDEREVPVGAVLVNLRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLYD
 ATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHHPGMNHRVEITEGI
 LADECAALLCRFFRMPRRVFNAQKKAQSSTD PV14 (PV28とも呼ばれる). pCMV_mon
 oABE8.1_bpNLS + V82S + Q154RMSEVEFSHEYWMRHALTLAKRARDEREVPVGAVL
 VLNLRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYSTFEPCVMCAGA
 MIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMP
 RRVFNAQKKAQSSTD

【0260】

【図58】図58 A～図58 Cは、グラフ、模式図および表を提供する。図58 AはLRRK2標
 的配列の核酸位置6におけるAからGへのパーセント変換を定量する。使用されたエディタ
 ーはPV15～PV28であり、これについての記述は上記に提供される。pCMVはCMVプロ
 モーターを示す；bpNLSは二部分核局在化シグナルを示す；monoABE8.1は、ABE8.1塩
 基エディターの一モノマー型を示す。図58 Bはパーキンソン病に関連するLRRK2におけ
 るG2019S突然変異の補正のための標的 (target) 配列およびガイド (Guide) RNAを示
 す。図58 Cは、LRRK2標的配列の核酸位置4および6におけるAからGへのパーセント変換
 を示す。4位のAからGへの転移はバイスタンダー効果である。

【図59】図59 A～59 Lは、R1441CをコードするLRRK2標的配列の位置7におけるAか
 らGへの遷移についての配列リードを示す(図57 A～57 C参照)。エディター表示されてい
 る (PV1～14)。PV1～28の説明は図56に提供される。

【図60】図60 A～60 Wは、G2019SをコードするLRRK2標的配列の位置4および6に
 おけるAからGへの遷移についての配列リードを示す(図58 A～58 Cを参照)。

【図61】図61 Aは、アンチセンス鎖G>A突然変異によってコードされる、LRRK2にお
 ける病原性突然変異A419Vの修正のための標的配列を示す概略図を提供する。変異は、T
 GG PAMに対する特異性を有するSpCas9バリエーションを用いて、位置12のAを標的とするA
 BEを用いて補正される。図61 Bは、パーキンソン病に関連するLRRK2における病原性
 突然変異L1114Lの補正のための標的配列を示す概略図である。この変異はアンチセンス
 鎖T>Cであり、シチジンデアミナーゼ活性を有する塩基エディター(CBE)を用いて修正さ
 れる。図61 Cは、パーキンソン病に関連するLRRK2における病原性突然変異I1122Vの
 補正のための標的配列を示す概略図である。この変異はアンチセンス鎖T>Cであり、シチ
 ジンデアミナーゼ活性を有する塩基エディター(CBE)を用いて修正される。図61 Dは、

パーキンソン病に関連するLRRK2における病原性突然変異M1869Vの補正のための標的配列を示す概略図を提供する。この変異はアンチセンス鎖T>Cであり、シチジンデアミナーゼ活性を有する塩基エディター(CBE)を用いて修正される。

【図62】図62 Aおよび62 Bは、HEK293T細胞におけるMus musculus IDUA W401X突然変異の正確な塩基編集補正を示す。図62 Aは、21ヌクレオチドガイドRNAを使用するABE8塩基エディターバリエーションを使用した、Mus musculus IDUA W401X突然変異の塩基編集(editing)のパーセンテージを示すグラフである。図62 Bは、21ヌクレオチドガイドRNAを使用するABE8塩基エディターバリエーションのインデル(Indel)のパーセンテージを示すグラフである。

【図63】図63は、20ヌクレオチドガイド(guide) RNAまたは21ヌクレオチドガイドRNAのいずれかを使用するABE8塩基エディターバリエーションを使用した、Mus musculus IDUA W401X突然変異の塩基編集(editing)の割合を示すグラフである。

【図64】図64は、W402X突然変異を修正するA→Gヌクレオチド塩基編集のための標的としてのホモ・サピエンスIDUAゲノム核酸およびアミノ酸配列の概略図を示す。図には、対応するガイド(Guide) RNA (gRNA) の核酸配列も示されている。図には、IDUA核酸配列中の標的(Target) アデノシン(A) 核酸塩基(ボックスで囲まれている) が示されている。

【図65】図65 Aおよび65 Bは、HEK293T細胞におけるホモ・サピエンスIDUA W402X突然変異の正確な塩基編集補正を示す。図65 Aは、20ヌクレオチドガイドRNAを使用するABE8塩基エディターバリエーションを使用した、ホモ・サピエンスIDUA W402X突然変異の塩基編集のパーセンテージを示すグラフである。図65 Bは、20ヌクレオチドガイドRNAを使用するABE8塩基エディターバリエーションのインデルのパーセンテージを示すグラフである。

【図66】図66 A～図66 Oは、ABE8塩基エディターバリエーションを用いたIDUA核酸配列におけるAからGへのヌクレオチド変化のパーセンテージの効率を示す表であり、塩基編集が行われた細胞におけるゲノムDNAのPCRに続くディープシーケンシング(MySeq)によって検出されている。図66 A～66 Mは、ABE8塩基エディターバリエーションABE8.1～ABE8.13のそれぞれ3つのサンプルを用いた、IDUA核酸標的部位の位置6におけるA→G塩基編集のパーセントを示す。図66 Nは、陽性対照塩基エディターABE7.10の3つのサンプルを用いた、IDUA核酸標的部位の位置6でのA→G塩基編集のパーセントを示す。図66 Oは、陰性対照の2つのサンプルを用いた、IDUA核酸標的部位における6位でのA→G塩基編集のパーセントを示す。

【図67】図67は、レット/MECP2: 変異補正を示している。MECP2の機能喪失は、多くの異なるde novo突然変異に起因し得る。X連鎖: XX患者はMECP2欠損についてモザイクである; XYは通常、乳児死亡をもたらす。

【図68】図68は、トップ3のガイド(Guide) 配列についてのレット症候群R106W突然変異補正を示す。

【図69】図69は、NGTT PAM最適化を有するエディターを用いたレット症候群R255X突然変異補正を示す。

【図70】図70 A～C: ハーラー/IDUA突然変異補正。図70 Aは、IDUA W402X突然変異補正の実験デザインを示す。図70 Bは、各エディター構築物のパーセント編集を示す。図70 Cは、編集された(edited) および編集されていない(unedited) 構築物についての比活性(nmol/mg/h)を示す。

【図71】図71は、ABE8.8によるインビボ塩基編集を示す。各サンプルについて左から順に、ガイド11(AAV9)、ガイド12(AAV9)、ガイド11(PHP.eB)、ガイド12(PHP.eB)、およびコントロール。

【0261】

【図72】図72 A～72 B。モデル細胞株における、ABCA4 G1961EアレルにおけるABE7.10およびABE8バリエーションによるA・TからG・Cへの変換。図72 A: 21 ntスパーサーsgRNAおよび塩基エディターバリエーションのプラスミドリポフェクション後の、ABCA4 G19

10

20

30

40

50

61Eコドンの統合された疾患アリル (Disease Allele) およびゆらぎ塩基 (Wobble Base) における、HEK293T細胞におけるA・TからG・Cへの変換 (Conversion)。リポフェクション後5日間インキュベートした細胞を編集について評価した。図72 B: ABCA4 G1961E疾患アリル、コドンのゆらぎ塩基、および21 ntスペーサーsgRNAによって使用される-NGG PAMを含む、対象部位におけるDNA配列。エラーバーは3つの複製のs.d.を表す。各データセットにおいて、疾患アリルは左側にあり、ゆらぎ塩基は右側にある。

【図73】モデル細胞株における、ABCA4 G1961EアリルにおけるsgRNAスペーサー長バリエーションによるA・TからG・Cへの変換。種々のスペーサー長のsgRNAおよびABE7.10のプラスミドリポフェクション後の、統合された疾患アリルおよびABCA4 G1961Eコドンのゆらぎ塩基における、HEK293T細胞におけるA・TからG・Cへの変換。リポフェクション後5日間インキュベートした細胞を編集について評価した。hRz=sgRNAの5'末端に自己開裂ハンマーヘッド型リボザイムを含む。エラーバーは3つの複製のs.d.を表す。各データセットにおいて、疾患アリルは左側にあり、ゆらぎ塩基は右側にある。

【図74】分割されたインティン (Intein) の再構成を使用した分割塩基エディターのデュアルAAV送達の概略図。2つのAAV粒子は、塩基編集に必要なコンポーネントで別々にパッケージ化される。1つのウイルスは、N末端分割インティン融合体を有する、塩基エディターのC末端領域をコードし、相補的なウイルスは、sgRNAと共に、C末端分割インティン融合体を有する塩基エディターのN末端領域をコードする。相補的なウイルスの同時形質導入により、sgRNAが転写され、塩基エディターの各半分が発現されて、分割インティンを介したタンパク質トランススプライシングを通じて再結合される。

【図75】図75 A~75 B。野生型細胞のABCA4 G1961における、分割 (Split) ABEバリエーションの二重AAV送達 (delivery) によるA・TからG・Cへの変換 (Conversion)。図75 A: 野生型ARPE-19細胞の野生型ABCA4 G1961標的部位におけるA・T→G・CおよびC・G→T・A転換であり、ここで、位置8Aにおける編集は、これらの細胞における編集についての代理 (surrogate) 標的として役立つ。細胞は、5E+4ウイルスゲノム/ウイルス/細胞のMOIで感染した。細胞を感染後2週間インキュベートし、編集について評価した。エラーバーは6回の繰り返しのs.d.を表す。各データポイントについて、位置8 (A>G) -代理サイトで処理されたサンプルを左側に、位置5 (C>T) を右に示す。図75 B: ABCA4 G1961アリル、および野生型配列を標的とする21 ntスペーサーsgRNAによって使用される-NGG PAMを含む、野生型標的部位におけるDNA配列。

【図76】図76 A~76 B。分割ABE7.10とABCA4 G1961Eの疾患アリルを標的とするsgRNAとを発現するAAV2で二重感染させた野生型ARPE-19細胞におけるオフターゲット塩基編集。図76 A: 未処理 (Untreated) 対照 (灰色) と比較した、二重AAV (青) での同時感染 (Treated) 2週間後の、ターゲット (On Targ) またはオフターゲット (OT) のプロトスペーサーにわたる最大A・T→G・C変換。図76 B: 未処理対照 (灰色) と比較した、二重AAV (青) での同時感染の2週間後の、ターゲット (On Targ) プロトスペーサーまたはオフターゲット (OT) プロトスペーサーにわたる最大の非A・T→G・C変換。各データポイントについて、野生型 (wt) ARPE-19細胞で処理 (Treated) した試料を左側に示し、未処理 (Untreated) のwt ARPE-19細胞を右側に示す。

【図77】分割ABE7.10と、ABCA4 G1961Eの疾患アリルを標的とするsgRNAとを発現するAAV2で二重感染させた野生型ARPE-19細胞における塩基編集によるインデル形成。未処理対照 (灰色) と比較した、二重AAV (青) での同時感染2週間後の、ターゲットまたはオフターゲットのプロトスペーサー内またはその近位に形成されたインデル (Indel) のパーセンテージ。各データポイントについて、野生型 (wt) ARPE-19細胞で処理 (Treated) した試料を左側に示し、未処理 (Untreated) のwt ARPE-19細胞を右側に示す。

【図78】培養後22日目における霊長類網膜完全性およびGFP発現。抗ロドプシン (Rhodopsin)、抗GFP、およびビオチン化ピーナッツ凝集素 (PNA) 抗体で4℃で一晩、切片を免疫標識した。Anc80L65.hGRK.eGFPは、GFPが光受容体含有外核層 (ONL) でのみ観察されることを示し、GRKプロモーターの光受容体特異的活性を確認した。一番上の行

10

20

30

40

50

は、形質導入されていない第0日である。2番目の行は形質導入されていない22日目である。第三列は22日目、GRKである。4番目の行は22日目、CMBである。カラムは、未染色(1stカラム)、DAPI(2ndカラム)、GFP(3rdカラム)、PNA(4thカラム)、ロドブシン(5thカラム)である。

【図79】NHPにおけるCas9発現。Cas9発現は霊長類網膜で培養後6日という早期に検出される。ABE7.10(列1および2)、ABE8.5(列2および3)、およびABE8.9(列3および4)の結果が示されている。上段：培養後6日目。下段：培養後17日目。結果は、AAV系がCas9を発現する分割インテインを送達することを実証している。スケールバー：100 μm。

【発明を実施するための形態】

【0262】

本発明は、標的核酸塩基配列の改変を生成するための、効率が増加した新規のアデニン塩基エディター(例えばABE8)を含む組成物およびそれらを使用する方法を提供する。

【0263】

〔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(例えば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)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(例えばCas9)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。

【0264】

〔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〕

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また、RNAに結合する核酸プログラム可能な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な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。

【0265】

塩基エディターの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、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ミング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ミング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、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ミング可能な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ミング可能な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。

10

20

30

40

50

【0266】

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。ある実施形態では、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を含むことができる。そのような実施形態では、残基H840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。

【0267】

例示的な触媒活性Cas9のアミノ酸配列は以下の通りである：

MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEA
TRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNI
VDEVAYHEKYPTIYHLRKKLV DSTDKADRLIYLALAHMIKFRGHFLIEGDLNPDNSDVD
KLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGKKNGLFGNLIAL
SLGLTPNFKSNFDLAEDAKLQLSKD TYDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDI
LRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDG
GASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ
EDFYPPFLKDNREKIEKILTFRIPIYVGPLARGNSRFWMTRKSEETITPWNFEEVVDKGAS
AQSFIERMTNFDKNLPNEKVLPKHSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKA
IVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLD
NEENEDILEDIVLTLT LFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLIN
GIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA
GSPAIAKKGILQTVKVVDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGI
KELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK
DDSIDNKVLTNRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGG
LSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRK
DFQFYKVRINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIKSEQ
EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLS
MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAK
VEKGKSKKLKSVKELLGITIMERSSEFKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELNGR
KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEII
EQISEFSKRVLADANLDKVL SAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTID
RKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGD.

【0268】

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。そのような実施形態では、非標的鎖は切断されない。

【0269】

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「ヌクレアーゼ不活」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RuvC1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。

【0270】

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。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))。

【0271】

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。いくつかの実施形態において、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。

【0272】

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. 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。

【0273】

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塩基スペーサーとからなる短い合成RNAである。したがって、当業者はCasタンパク質特異性のゲノム標的を変化させることができ、これは、ゲノムの他の部分と比較してgRNA標的化配列がゲノム標的に対していかに特異的であるかによって部分的に決定される。

【0274】

いくつかの実施形態において、gRNA足場配列は以下の通りである：GUUUUAGAGC U AGAAAUAGC AAGUAAAAU AAGGCUAGUC CGUUAUCAAC UUGAAAAAGU GGCACC GAGU CGGUGCUUUU。

【0275】

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。

【0276】

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、C

10

20

30

40

50

as5t、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、Cssx16、Cx16、Cx、Csx3、Csx1、Csx1S、Csf1、Csf2、Cs0、Csf4、Csd1、Csd2、Cst1、Cst2、Csh1、Csh2、Csa1、Csa2、Csa3、Csa4、Csa5、Cas12a/Cpf1、Cas12b/C2c1、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。

【0277】

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。

【0278】

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 syri* (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 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); *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。

【0279】

[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., 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. et al., Nature 471:602-607(2011); および “A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity.” Jinek M. et al., 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オーソログは、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、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。

【0280】

いくつかの実施形態において、核酸プログラミング可能DNA結合タンパク質 (napDNA bp) は、Cas9ドメインである。非限定的な例示的Cas9ドメインが本明細書で提供される。Cas9ドメインは、ヌクレアーゼ活性Cas9ドメイン、ヌクレアーゼ不活性Cas9ドメイン (dCas9)、またはCas9ニッカーゼ (nCas9) であり得る。ある態様において、Cas9ドメインは、ヌクレアーゼ活性ドメインである。例えば、Cas9ドメインは、二本鎖核酸の両方の鎖（例えば二本鎖DNA分子の両方の鎖）を切断するCas9ドメインであり得る。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1つ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1つと比較して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1つと比較して、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個の同一の一続きのアミノ酸残基を有するアミノ酸配列を含む。

【0281】

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またはその断片と同一性を共有する。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。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。

【0282】

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。

【0283】

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 (例えばdCas9およびnCas9)、CasX、CasY、Cpf1、Cas12b/C2C1、およびCas12c/C2C3が挙げられるが、これらに限定されない。ある実施形態において、野生型Cas9は、*Streptococcus pyogenes* 由来のCas9に対応する (NCBI参照配列: NC_017053.1、ヌクレオチドおよびアミノ酸の配列は以下の通り)。

【0284】

```
ATGGATAAGAAATACTCAATAGGCTTAGATATCGGCACAAATAGCGTCGGATGGGCGGT
GATCACTGATGATTATAAGGTTCCGTCTAAAAAGTTCAAGGTTCTGGGAAATACAGACC
GCCACAGTATCAAAAAAATCTTATAGGGGCTCTTTTATTTGGCAGTGGAGAGACAGCG
GAAGCGACTCGTCTCAAACGGACAGCTCGTAGAAGGTATACACGTCGGAAGAATCGTAT
TTGTTATCTACAGGAGATTTTTTCAAATGAGATGGCGAAAGTAGATGATAGTTTCTTTC
ATCGACTTGAAGAGTCTTTTTTGGTGGGAAGAAGACAAGAAGCATGAACGTCATCCTATT
TTTGGAAATATAGTAGATGAAGTTGCTTATCATGAGAAATATCCAACATCTATCATCT
GCGAAAAAATTGGCAGATTCTACTGATAAAGCGGATTTGCGCTTAATCTATTTGGCCT
TAGCGCATATGATTAAGTTTCGTGGTCATTTTTTGGATTGAGGGAGATTTAAATCCTGAT
AATAGTGATGTGGACAACTATTTATCCAGTTGGTACAAATCTACAATCAATTATTTGA
AGAAAACCCTATTAACGCAAGTAGAGTAGATGCTAAAGCGATTCTTTCTGCACGATTGA
GTAAATCAAGACGATTAGAAAATCTCATTGCTCAGCTCCCCGGTGAGAAGAGAAATGGC
TTGTTTGGGAATCTCATTGCTTTGTCATTGGGATTGACCCCTAATTTTAAATCAAATTTT
GATTTGGCAGAAGATGCTAAATTACAGCTTTCAAAGATACTTACGATGATGATTTAGA
TAATTTATTGGCGCAAATTGGAGATCAATATGCTGATTTGTTTTTGGCAGCTAAGAATT
TATCAGATGCTATTTTACTTTTCAAGATATCCTAAGAGTAAATAGTGAAATAACTAAGGCT
CCCCTATCAGCTTCAATGATTAAGCGCTACGATGAACATCATCAAGACTTGACTCTTTT
AAAAGCTTTAGTTTCGACAACAACCTTCCAGAAAAGTATAAAGAAATCTTTTTTGGATCAAT
CAAAAAACGGATATGCAGGTTATATTGATGGGGGAGCTAGCCAAGAAGAATTTTATAAA
TTTATCAAACCAATTTTAGAAAAAATGGATGGTACTGAGGAATTATTGGTGAAACTAAA
TCGTGAAGATTTGCTGCGCAAGCAACGGACCTTTGACAACGGCTCTATTCCCCATCAAA
TTCACTTGGGTGAGCTGCATGCTATTTTGAGAAGACAAGAAGACTTTTATCCATTTTTTA
AAAGACAATCGTGAGAAGATTGAAAAAATCTTGACTTTTCGAATTCCTTATTATGTTGG
TCCATTGGCGCGTGGCAATAGTCGTTTTGCATGGATGACTCGGAAGTCTGAAGAAACAA
TTACCCCATGGAATTTTGAAGAAGTTGTCGATAAAGGTGCTTCAGCTCAATCATTTATT
GAACGCATGACAACTTTGATAAAAAATCTTCAAATGAAAAAGTACTACCAAAACATAG
TTTGCTTTATGAGTATTTTACGGTTTATAACGAATTGACAAAGGTCAAATATGTTACTG
```

10

20

30

40

50

AGGGAATGCGAAAACCAGCATTTCTTTTCAGGTGAACAGAAGAAAGCCATTGTTGATTTA
CTCTTCAAAACAAATCGAAAAGTAACCGTTAAGCAATTAAGAAGATTATTTCAAAAA
AATAGAATGTTTTGATAGTGTTGAAATTTTCAGGAGTTGAAGATAGATTTAATGCTTCAT
TAGGCGCCTACCATGATTTGCTAAAAATTATTAAGATAAAGATTTTTTGGATAATGAA
GAAAATGAAGATATCTTAGAGGATATTGTTTTAACATTGACCTTATTTGAAGATAGGGG
GATGATTGAGGAAAGACTTAAACATATGCTCACCTCTTTGATGATAAGGTGATGAAAC
AGCTTAAACGTCGCCGTTATACTGGTTGGGGACGTTTGTCTCGAAAATTGATTAATGGT
ATTAGGGATAAGCAATCTGGCAAAACAATATTAGATTTTTTGAATCAGATGGTTTTGC
CAATCGCAATTTTATGCAGCTGATCCATGATGATAGTTTGACATTTAAAGAAGATATTC
AAAAAGCACAGGTGTCTGGACAAGGCCATAGTTTACATGAACAGATTGCTAACTTAGCT
GGCAGTCCTGCTATTAAAAAAGGTATTTTACAGACTGTAAAAATTGTTGATGAACTGGT
CAAAGTAATGGGGCATAAGCCAGAAAATATCGTTATTGAAATGGCACGTGAAAATCAGA
CAACTCAAAAGGGCCAGAAAAATTGCGGAGAGCGTATGAAACGAATCGAAGAAGGTATC
AAAGAATTAGGAAGTCAGATTCTTAAAGAGCATCCTGTTGAAAATACTCAATTGCAAAA
TGAAAAGCTCTATCTCTATTATCTACAAAATGGAAGAGACATGTATGTGGACCAAGAAT
TAGATATTAATCGTTTAAAGTGATTATGATGTCGATCACATTGTTCCACAAAGTTTCATT
AAAGACGATTCAATAGACAATAAGGTACTAACGCGTTCTGATAAAAATCGTGGTAAATC
GGATAACGTTCCAAGTGAAGAAGTAGTCAAAAAGATGAAAACTATTGGAGACAACTTC
TAAACGCCAAGTTAATCACTCAACGTAAGTTTGATAATTTAACGAAAGCTGAACGTGGA
GGTTTGAGTGAACCTTGATAAAGCTGGTTTTATCAAACGCCAATTGGTTGAACTCGCCA
AATCACTAAGCATGTGGCACAAATTTTGGATAGTCGCATGAATACTAAATACGATGAAA
ATGATAAACTTATTCGAGAGGTTAAAGTGATTACCTTAAAATCTAAATTAGTTTCTGAC
TTCCGAAAAGATTTCCAATTCTATAAAGTACGTGAGATTAACAATTACCATCATGCCCA
TGATGCGTATCTAAATGCCGTCGTTGGAACTGCTTTGATTAAGAAATATCCAAAACCTTG
AATCGGAGTTTGTCTATGGTGATTATAAAGTTTATGATGTTGTAATAATGATTGCTAAG
TCTGAGCAAGAAATAGGCAAAGCAACCGCAAAATATTTCTTTTACTCTAATATCATGAA
CTTCTTCAAAACAGAAATTACACTTGCAAATGGAGAGATTGCGAAACGCCCTCTAATCG
AACTAATGGGGAACTGGAGAAATTGTCTGGGATAAAGGGCGAGATTTTGCCACAGTG
CGCAAAGTATTGTCCATGCCCAAGTCAATATTGTCAAGAAAACAGAAGTACAGACAGG
CGGATTCTCCAAGGAGTCAATTTTACCAAAAAGAAATTCGGACAAGCTTATTGCTCGTA
AAAAAGACTGGGATCCAAAAAATATGGTGGTTTTGATAGTCCAACGGTAGCTTATTCA
GTCCTAGTGTTGCTAAGGTGGAAAAAGGGAAATCGAAGAAGTTAAAATCCGTTAAAGA
GTTACTAGGGATCACAATTATGGAAAGAAGTTCCTTTGAAAAAATCCGATTGACTTTT
TAGAAGCTAAAGGATATAAGGAAGTTAAAAAAGACTTAATCATTAACTACCTAAATAT
AGTCTTTTTGAGTTAGAAAACGGTCGTAAACGGATGCTGGCTAGTGCCGGAGAATTACA
AAAAGGAAATGAGCTGGCTCTGCCAAGCAAATATGTGAATTTTTTATATTTAGCTAGTC
ATTATGAAAAGTTGAAGGGTAGTCCAGAAGATAACGAACAAAAACAATTGTTTGTGGAG
CAGCATAAGCATTATTTAGATGAGATTATTGAGCAAATCAGTGAATTTTCTAAGCGTGT
TATTTTAGCAGATGCCAATTTAGATAAAGTTCTTAGTGATATAACAAACATAGAGACA
AACCAATACGTGAACAAGCAGAAAATATTATTCATTTATTTACGTTGACGAATCTTGA
GCTCCCGCTGCTTTTAAATATTTTGATACAACAATTGATCGTAAACGATATACGTCTAC
AAAAGAAGTTTGTAGATGCCACTCTTATCCATCAATCCATCACTGGTCTTTATGAAACAC
GCATTGATTTGAGTCAGCTAGGAGGTGACTGA

10

20

30

40

50

MDKKYSIGLDIGTNSVGWAVITDDYKVPSKKFKVLGNTDRHSIKKNLIGALLFGSGETAETRL
 KRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNIVDEVAY
 HEKYPTIYHLRKKLADSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQIY
 NQLFEENPINASRVDAKAILSARLSKSRLENLIAQLPGEKRNGLFGNLIALLSLGLTPNFKSNF
 DLAEDAKLQLSKDITYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNSEITKAPLSAS
 MIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMD
 GTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRI
 PYYVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHS
 LLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFD
 SVEISGVEDRFNASLGAYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDRGMIEERLKTYA
 HLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDFANRNFQMQLIHDDSLTF
 KEDIQKAQVSGQGHSLEQIANLAGSPAIKKGILQTVKIVDELVKVMGHKPENIVIEEMARENQT
 TQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRL
 SDYDVDHIVPQSFIKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKE
 DNLTKAERGGLELKDAGFIKRQLVETROITKHVAQILDSRMNTKYDENDKLIREVKVITLKS
 LVSDFRKDFQFYKVBREINNYHHAHDAYLNAVVGITALIKKYPKLESEFVYGDYKVYDVRKMIKS
 EQEIGKATAKYFFYSNIMNFEKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL
 SMPOQVNIKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEK
 GSKKLKSVKELLGITIMERSSEKPNIDFLEAKGYKEVKDLIIKLPKYSLEFLENKRKMLASA
 GELQKGNELALPSKYVNFYLLASHYEKLKGSPEQKQLFVEQHKHYLDEIIEQISEFSKRVI
 LADANLDKVL SAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDA
 TLIHQSIITGLYETRIDLSQLGGD

10

20

30

(一重下線：HNHドメイン；二重下線：RuvCドメイン)

【0285】

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
 GTCATTGATTAAAAAGAATCTTATCGGTGCCCTCCTATTGATAGTGGCGAAACGGCA
 GAGGCGACTCGCCTGAAACGAACCGCTCGGAGAAGGTATACACGTCGCAAGAACCGAAT
 ATGTTACTTACAAGAAATTTTAGCAATGAGATGGCCAAAGTTGACGATTCTTTCTTTC
 ACCGTTTGGGAAGAGTCCTTCCTTGTCGAAGAGGACAAGAAACATGAACGGCACCCCATC
 TTTGGAAACATAGTAGATGAGGTGGCATATCATGAAAAGTACCAACGATTTATCACCT
 CAGAAAAAAGCTAGTTGACTCAACTGATAAAGCGGACCTGAGGTTAATCTACTTGGCTC
 TTGCCCATATGATAAAGTTCCGTGGGCACTTTCTCATTGAGGGTGATCTAAATCCGGAC
 AACTCGGATGTCGACAACTGTTTCATCCAGTTAGTACAAACCTATAATCAGTTGTTTGA
 AGAGAACCCTATAAATGCAAGTGGCGTGGATGCGAAGGCTATTCTTAGCGCCCGCCTCT
 CTAATCCCGACGGCTAGAAAACCTGATCGCACAAATTACCGGAGAGAAGAAAAATGGG
 TTGTTGGTAACCTTATAGCGCTCTCACTAGGCCTGACACCAAATTTTAAGTCGAACTT

40

50

CGACTTAGCTGAAGATGCCAAATTGCAGCTTAGTAAGGACACGTACGATGACGATCTCG
ACAATCTACTGGCACAAATTGGAGATCAGTATGCGGACTTATTTTTGGCTGCCAAAAC
CTTAGCGATGCAATCCTCCTATCTGACATACTGAGAGTTAATACTGAGATTACCAAGGC
GCCGTTATCCGCTTCAATGATCAAAAGGTACGATGAACATCACCAAGACTTGACACTTC
TCAAGGCCCTAGTCCGTGAGCAACTGCCTGAGAAATATAAGGAAATATTCTTTGATCAG
TCGAAAAACGGGTACGCAGGTTATATTGACGGCGGAGCGAGTCAAGAGGAATTCTACAA
GTTTATCAAACCCATATTAGAGAAGATGGATGGGACGGAAGAGTTGCTTGTA AAACTCA
ATCGCGAAGATCTACTGCGAAAGCAGCGGACTTTCGACAACGGTAGCATTCCACATCAA
ATCCACTTAGGCGAATTGCATGCTATACTTAGAAGGCAGGAGGATTTTTATCCGTTCT
CAAAGACAATCGTGAAAAGATTGAGAAAATCCTAACCTTTCGCATACCTTACTATGTGG
GACCCCTGGCCCGAGGGAACCTCTCGGTTTCGCATGGATGACAAGAAAGTCCGAAGAAACG
ATTACTCCATGGAATTTTGAGGAAGTTGTGCGATAAAGGTGCGTCAGCTCAATCGTTCAT
CGAGAGGATGACCAACTTTGACAAGAATTTACCGAACGAAAAAGTATTGCCTAAGCACA
GTTTACTTTACGAGTATTTTACAGTGTACAATGAACTCACGAAAGTTAAGTATGTCACT
GAGGGCATGCGTAAACCCGCCTTTCTAAGCGGAGAACAGAAGAAAGCAATAGTAGATCT
GTTATTCAAGACCAACCGCAAAGTGACAGTTAAGCAATTGAAAGAGGACTACTTTAAGA
AAATTGAATGCTTCGATTCTGTGAGATCTCCGGGGTAGAAGATCGATTTAATGCGTCA
CTTGGTACGTATCATGACCTCCTAAAGATAATTAAAGATAAGGACTTCCTGGATAACGA
AGAGAATGAAGATATCTTAGAAGATATAGTGTTGACTCTTACCCTCTTTGAAGATCGGG
AAATGATTGAGGAAAGACTAAAAACATACGCTCACCTGTTTCGACGATAAGGTTATGAAA
CAGTTAAAGAGGCGTCGCTATACGGGCTGGGGACGATTGTGCGGAAACTTATCAACGG
GATAAGAGACAAGCAAAGTGGTAAAACCTATTCTCGATTTTCTAAAGAGCGACGGCTTCG
CCAATAGGAACTTTATGCAGCTGATCCATGATGACTCTTTAACCTTCAAAGAGGATATA
CAAAGGCACAGGTTTTCCGGACAAGGGGACTCATTGCACGAACATATTGCGAATCTTGC
TGGTTCGCCAGCCATCAAAAAGGGCATACTCCAGACAGTCAAAGTAGTGGATGAGCTAG
TTAAGGTCATGGGACGTCACAAACCGGAAAACATTGTAATCGAGATGGCACGCGAAAAT
CAAACGACTCAGAAGGGGCAAAAAAACAGTCGAGAGCGGATGAAGAGAATAGAAGAGG
GTATTAAGAAGTGGGCAGCCAGATCTTAAAGGAGCATCCTGTGGAAAATACCCAATTG
CAGAACGAGAACTTTACCTCTATTACCTACAAAATGGAAGGGACATGTATGTTGATCA
GGAACCTGGACATAAACCGTTTTATCTGATTACGACGTCGATCACATTGTACCCCAATCCT
TTTTGAAGGACGATTCAATCGACAATAAAGTGCTTACACGCTCGGATAAGAACCGAGGG
AAAAGTGACAATGTTCCAAGCGAGGAAGTCGTAAAGAAAATGAAGAACTATTGGCGGCA
GCTCCTAAATGCGAACTGATAACGCAAAGAAAGTTTCGATAACTTAACTAAAGCTGAGA
GGGGTGGCTTGTCTGAACTTGACAAGGCCGGATTTATTAAACGTCAGCTCGTGGAACCC
CGCCAAATCACAAAGCATGTTGCACAGATACTAGATTCCCGAATGAATACGAAATACGA
CGAGAACGATAAGCTGATTCGGGAAGTCAAAGTAATCACTTTAAAGTCAA AATTGGTGT
CGGACTTCAGAAAGGATTTTCAATTCTATAAAGTTAGGGAGATAAATAACTACCACCAT
GCGCACGACGCTTATCTTAATGCCGTCGTAGGGACCGCACTCATTAAGAAATACCCGAA
GCTAGAAAGTGAGTTTGTGTATGGTGATTACAAAGTTTATGACGTCGTAAGATGATCG
CGAAAAGCGAACAGGAGATAGGCAAGGCTACAGCCAAATACTTCTTTTATTCTAACATT
ATGAATTTCTTTAAGACGGAAATCACTCTGGCAAACGGAGAGATACGCAAACGACCTTT
AATTGAAACCAATGGGGAGACAGGTGAAATCGTATGGGATAAGGGCCGGGACTTCGCGA
CGGTGAGAAAAGTTTTGTCCATGCCCAAGTCAACATAGTAAAGAAAACCTGAGGTGCAG
ACCGGAGGGTTTTCAAAGGAATCGATTCTTCAAAAAGGAATAGTGATAAGCTCATCGC
TCGTAAAAAGGACTGGGACCCGAAAAAGTACGGTGGCTTCGATAGCCCTACAGTTGCCT
ATTCTGTCCTAGTAGTGGCAAAGTTGAGAAGGGGAAAATCCAAGAACTGAAGTCAGTC
AAAGAATTATTGGGGATAACGATTATGGAGCGCTCGTCTTTTGAAAAGAACCCCATCGA
CTTCCTTGAGGCGAAAGGTTACAAGGAAGTAAAAAAGGATCTCATAATTAACTACCAA
AGTATAGTCTGTTTGAGTTAGAAAATGGCCGAAAACGGATGTTGGCTAGCGCCGGAGAG
CTTCAAAAGGGGAACGAACTCGCACTACCGTCTAAATACGTGAATTCCTGTATTTAGC

10

20

30

40

50

GTCCATTACGAGAAGTTGAAAGGTTACCTGAAGATAACGAACAGAAGCAACTTTTTG
 TTGAGCAGCACAAACATTATCTCGACGAAATCATAGAGCAAATTTGGAATTCAGTAAG
 AGAGTCATCCTAGCTGATGCCAATCTGGACAAAGTATTAAGCGCATAACAAGCACAG
 GGATAAACCCATACGTGAGCAGGCGGAAATATTATCCATTTGTTTACTCTTACCAACC
 TCGGCGCTCCAGCCGCATTCAAGTATTTTGACACAACGATAGATCGCAAACGATACACT
 TCTACCAAGGAGGTGCTAGACGCGACACTGATTCACCAATCCATCACGGGATTATATGA
 AACTCGGATAGATTTGTACAGCTTGGGGGTGACGGATCCCCAAGAAGAAGAGGAAAG
 TCTCGAGCGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAG
 GATGACGATGACAAGGCTGCAGGA

MDKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRL
 KRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNIVDEVAY
 HEKYPTIYHLRKKLVDSTDKADLRILIYLAHAMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTY
 NQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALLSLGLTPNFKSNF
 DLAEDAKLQLSKDITYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSAS
 MIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMD
 GTEELLVKNLREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFLKDNREKIEKILTFRI
 PYYVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNEKVLPHKS
 LLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFD
 SVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYA
 HLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTF
 KEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQ
 TTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINR
 LSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRK
 FDNLTKAERGGELSELDKAGFIKRQLVETROITKHVAQILD SRMNTKYDENDKLIREVKVITLKS
 KLVSDFRKDFQFYKVBREINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIAK
 SEQEIGKATAKYFFYSNIMNFEKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL
 MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEKG
 KSKKLKSVKELLGITIMERSSEFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLEFENGRKRMLAS
 AGELQKGNELALPSKYVNFYLYLASHYEKLKGSPEDEQKQLFVEQHKHYLDEIIEQISEFSKRV
 ILADANLDKVL SAYNKH RDKPIREQAENI IHLFTLTNLGAPAAF KYFDTTIDRKRYTSTKEVLD
 ATLIHQ SITGLYETRIDLSQLGGD

10

20

30

40

(一重下線：HNHドメイン；二重下線：RuvCドメイン)

【0286】

ある態様において、野生型Cas9は、Streptococcus pyogenesからのCas9 (NCBI Reference Sequence: NC_002737.2 (ヌクレオチド配列は以下の通り) および Uniprot Reference Sequence: Q99ZW2 (アミノ酸配列は以下の通り) に対応する：

ATGGATAAGAAATACTCAATAGGCTTAGATATCGGCACAAATAGCGTCGGATGGGCGGT
 GATCACTGATGAATATAAGGTTCCGTCTAAAAAGTTCAAGGTTCTGGGAAATACAGACC
 GCCACAGTATCAAAAAAATCTTATAGGGGCTCTTTTATTTGACAGTGGAGAGACAGCG
 GAAGCGACTCGTCTCAAACGGACAGCTCGTAGAAGGTATACACGTCGGAAGAATCGTAT

50

TTGTTATCTACAGGAGATTTTTTCAAATGAGATGGCGAAAGTAGATGATAGTTTCTTTC
ATCGACTTGAAGAGTCTTTTTTGGTGGGAAGAAGACAAGAAGCATGAACGTCATCCTATT
TTTGAAATATAGTAGATGAAGTTGCTTATCATGAGAAATATCCAACCTATCTATCATCT
GCGAAAAAATTGGTAGATTCTACTGATAAAGCGGATTTGCGCTTAATCTATTTGGCCT
TAGCGCATATGATTAAGTTTCGTGGTCATTTTTTGGATTGAGGGAGATTTAAATCCTGAT
AATAGTGATGTGGACAACTATTTATCCAGTTGGTACAAACCTACAATCAATTATTTGA
AGAAAACCCTATTAACGCAAGTGAGTAGATGCTAAAGCGATTCTTTCTGCACGATTGA
GTAAATCAAGACGATTAGAAAATCTCATTGCTCAGCTCCCCGGTGAGAAGAAAAATGGC
TTATTTGGGAATCTCATTGCTTTGTCATTGGGTTTGACCCCTAATTTTAAATCAAATTT
GATTTGGCAGAAGATGCTAAATTACAGCTTTCAAAAGATACTTACGATGATGATTTAGA
TAATTTATTGGCGCAAATTGGAGATCAATATGCTGATTTGTTTTTGGCAGCTAAGAATT
TATCAGATGCTATTTTACTTTTCAAGATATCCTAAGAGTAAATACTGAAATAACTAAGGCT
CCCCTATCAGCTTCAATGATTAAACGCTACGATGAACATCATCAAGACTTGACTCTTTT
AAAAGCTTTAGTTTCGACAACAACCTTCCAGAAAAGTATAAAGAAATCTTTTTTGATCAAT
CAAAAAACGGATATGCAGGTTATATTGATGGGGGAGCTAGCCAAGAAGAATTTTATAAA
TTTATCAAACCAATTTTAGAAAAAATGGATGGTACTGAGGAATTATTGGTGAAACTAAA
TCGTGAAGATTTGCTGCGCAAGCAACGGACCTTTGACAACGGCTCTATTCCCCATCAAA
TTCATTGGGTGAGCTGCATGCTATTTTGAGAAGACAAGAAGACTTTTATCCATTTTTA
AAAGACAATCGTGAGAAGATTGAAAAAATCTTGACTTTTCGAATTCCTTATTATGTTGG
TCCATTGGCGCGTGGAATAGTCGTTTTGCATGGATGACTCGGAAGTCTGAAGAAACAA
TTACCCCATGGAATTTTGAAGAAGTTGTCGATAAAGGTGCTTCAGCTCAATCATTTATT
GAACGCATGACAACTTTGATAAAAAATCTTCAAATGAAAAAGTACTACCAAAACATAG
TTTGCTTTATGAGTATTTTACGGTTTATAACGAATTGACAAAGGTCAAATATGTTACTG
AAGGAATGCGAAAACCAGCATTTCTTTTCAAGGTGAACAGAAGAAAGCCATTGTTGATTTA
CTCTTCAAAACAAATCGAAAAGTAACCGTTAAGCAATTAAGAAGATTATTTCAAAAA
AATAGAATGTTTTGATAGTGTTGAAATTTCAAGGAGTTGAAGATAGATTTAATGCTTCAT
TAGGTACCTACCATGATTTGCTAAAAATTATTAAGATAAAGATTTTTTGGATAATGAA
GAAATGAAGATATCTTAGAGGATATTGTTTTAACATTGACCTTATTTGAAGATAGGGA
GATGATTGAGGAAAGACTTAAAACATATGCTCACCTCTTTGATGATAAGGTGATGAAAC
AGCTTAAACGTCGCCGTTATACTGGTTGGGGACGTTTGTCTCGAAAATTGATTAATGGT
ATTAGGGATAAGCAATCTGGCAAAACAATATTAGATTTTTTTGAAATCAGATGGTTTTGC
CAATCGCAATTTTATGCAGCTGATCCATGATGATAGTTTGACATTTAAAGAAGACATTC
AAAAAGCACAAGTGTCTGGACAAGGCGATAGTTTACATGAACATATTGCAAATTTAGCT
GGTAGCCCTGCTATTTAAAAAAGGTATTTTACAGACTGTAAAAGTTGTTGATGAATTGGT
CAAAGTAATGGGGCGGCATAAGCCAGAAAATATCGTTATTGAAATGGCACGTGAAAATC
AGACAACCTCAAAAGGGCCAGAAAAATTTCGCGAGAGCGTATGAAACGAATCGAAGAAGGT
ATCAAAGAATTAGGAAGTCAGATTCTTAAAGAGCATCCTGTTGAAAATACTCAATTGCA
AAATGAAAAGCTCTATCTCTATTATCTCCAAAATGGAAGAGACATGTATGTGGACCAAG
AATTAGATATTAATCGTTTAAAGTGATTATGATGTCGATCACATTGTTCCACAAAGTTTC
CTTAAAGACGATTCAATAGACAATAAGGTCTTAAACGCGTTCTGATAAAAAATCGTGGTAA
ATCGGATAACGTTCCAAGTGAAGAAGTAGTCAAAAAGATGAAAAACTATTGGAGACAAC
TTCTAAACGCCAAGTTAATCACTCAACGTAAGTTTGATAATTTAACGAAAGCTGAACGT
GGAGGTTTGAGTGAACCTTGATAAAGCTGGTTTTATCAAACGCCAATTGGTTGAAACTCG
CCAAATCACTAAGCATGTGGCACAATTTTGGATAGTCGCATGAATACTAAATACGATG
AAAATGATAAACTTATTCGAGAGGTTAAAGTGATTACCTTAAAATCTAAATTAGTTTCT
GACTTCCGAAAAGATTTCCAATTCTATAAAGTACGTGAGATTAACAATTACCATCATGC
CCATGATGCGTATCTAAATGCCGTCGTTGGAAGTCTTTGATTAAGAAATATCCAAAAC
TTGAATCGGAGTTTGTCTATGGTGATTATAAAGTTTATGATGTTTCGTAAAATGATTGCT
AAGTCTGAGCAAGAAATAGGCAAAGCAACCGCAAAATATTTCTTTTACTCTAATATCAT
GAACTTCTTCAAAACAGAAATTACACTTGCAAATGGAGAGATTTCGCAAACGCCCTCTAA

10

20

30

40

50

TCGAAACTAATGGGGAACTGGAGAAATTGTCTGGGATAAAGGGCGAGATTTTGCCACA
GTGCGCAAAGTATTGTCCATGCCCCAAGTCAATATTGTCAAGAAAACAGAAGTACAGAC
AGGCGGATTCTCCAAGGAGTCAATTTTACCAAAAAGAAATTCGGACAAGCTTATTGCTC
GTAAAAAAGACTGGGATCCAAAAAATATGGTGGTTTTGATAGTCCAACGGTAGCTTAT
TCAGTCCTAGTGGTTGCTAAGGTGGAAAAAGGGAAATCGAAGAAGTTAAAATCCGTTAA
AGAGTTACTAGGGATCACAATTATGGAAAGAAGTTCCTTTGAAAAAAATCCGATTGACT
TTTTAGAAGCTAAAGGATATAAGGAAGTTAAAAAAGACTTAATCATTAAACTACCTAAA
TATAGTCTTTTTGAGTTAGAAAACGGTCGTAAACGGATGCTGGCTAGTGCCGGAGAATT
ACAAAAAGGAAATGAGCTGGCTCTGCCAAGCAAATATGTGAATTTTTTATATTTAGCTA
GTCATTATGAAAAGTTGAAGGGTAGTCCAGAAGATAACGAACAAAAACAATTGTTTGTG
GAGCAGCATAAGCATTATTTAGATGAGATTATTGAGCAAATCAGTGAATTTTCTAAGCG
TGTTATTTTAGCAGATGCCAATTTAGATAAAGTTCTTAGTGCATATAACAAACATAGAG
ACAAACCAATACGTGAACAAGCAGAAAATATTATTCATTTATTTACGTTGACGAATCTT
GGAGCTCCCGCTGCTTTTAAATATTTTGATACAACAATTGATCGTAAACGATATACGTC
TACAAAAGAAGTTTTAGATGCCACTCTTATCCATCAATCCATCACTGGTCTTTATGAAA
CACGCATTGATTTGAGTCAGCTAGGAGGTGACTGA

10

MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRL
KRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNIVDEVAY
HEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTY
NQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNF
DLAEDAKLQLSKDQTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSAS
MIKRYDEHHQDLTLLKALVRQQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMD
GTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRI
PYYVGPLARGNSRFWMTRKSEETITPWNFEFVVDKGASQSFIERMTNFDKNLPNEKVLPHKS
LLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFD
SVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYA
HLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTF
KEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVEMARENQ
TTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINR
LSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRK
FDNLTKAERGGLSELDKAGFIKRQLVETROITKHVAQILDSRMNTKYDENDKLIREVKVITLKS
KLVSDFRKDFQFYKVINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAK
SEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLS
MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEKG
KSKKLKSVKELLGITIMERSSEFEKNPIDFLEAKGYKEVKKDIIKLPKYSLFEELENGRKRMLAS
AGELQKGNELALPSKYVNFYLYLASHYEKLKGSPEDEQKQLFVEQHKHYLDEIIIEQISEFSKR
VILADANLDKVL SAYNKH RD KPIREQAENI IHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLD
ATLIHQ SITGLYETRIDLSQLGGD

20

30

40

【0287】

いくつかの実施態様において、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 baltica* (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、または任意の他の生物からのCas9を表す。

10

【0288】

追加的なCas9タンパク質（例えば、ヌクレアーゼ不活（dead）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、ヌクレアーゼ不活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。

【0289】

20

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お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）中に提供される以下のアミノ酸配列を含む。

30

【0290】

例示的な、触媒的に不活性なCas9（dCas9）のアミノ酸配列は、以下の通りである：

MDKKYSIGLAIGTNSVGWAVITDEYKVPSPKFKVLGNTDRHSIKKNLIGALLFDSGETAETRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLVESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYLAHAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKILSARLSKSRLENLIAQLPGKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDITYDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPFLKDNREKIEKILTFRIPIYYVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDELVKVMGRHKPENIVIMARENQTTQKGQKNSRERMKRIEEGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVAIVPQSFLKDDSIDNKVLTNRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQ

40

50

EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLS
MPQVNIIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAK
VEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGR
KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEII
EQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLFTLTNLGAPAAFKYFDTTID
RKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGD

(例えば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(その内容全体が参照により本明細書に組み入れられる))。

【0291】

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、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))。

【0292】

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)または触媒的に不活性なCas9とも称される。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))。

【0293】

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

10

20

30

40

50

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。

【0294】

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。

【0295】

いくつかの実施形態において、dCas9は、dCas9 (D10AおよびH840A)のアミノ酸配列を含む：

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRL
KRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVAY
HEKYPTIYHLRKKLVDSTDKADLRLLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTY
NQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALLSLGLTPNFKSNF
DLAEDAKLQLSKDITYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSAS
MIKRYDEHHQDLTLLKALVRQQLPKEYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMD
GTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRI
PYYVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGASQSFIERMTNFDKNLPNEKVLPHKS
LLEYEFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFD
SVEISGVEDRFNASLGTYHDLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIERLKYA
HLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDFANRNFQMQLIHDDSLTF
KEDIQKAQVSGQGSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQ
TTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINR
LSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKMKNYWRQLLNAKLITQRK
FDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS
KLVSDFRKDFQFYKVRINNYHHADAYLNAVVGITALIKKYPKLESEFVYGDKYVDVRKMIAK
SEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLS
MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVEKG
KSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDLIIKLPKYSLELENGRKRMLAS
AGELQKGNELALPSKYVNFLYLASHYEKLGKSPEDNEQKQLFVEQHKHYLDEIIIEQISEFSKR
VILADANLDKVL SAYNKHDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLD
ATLIHQSI TGLYETRIDLSQLGGD

(一重下線：HNHドメイン；二重下線：RuvCドメイン)

【0296】

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。

【0297】

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) をもたらず、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。

【0298】

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。

【0299】

例示的な触媒的Cas9ニッカーゼ(nCas9)のアミノ酸配列は、以下の通りである：

MDKKYSIGLAIGTNSVGWAVITDEYKVPSPKFKVLGNTDRHSIKKNLIGALLFDSGETAEA
TRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNI
VDEVAYHEKYPTIYHLRKKLV DSTDKADRLRIYLALAHMIKFRGHFLIEGDLNPDNSDVD
KLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGKKNGLFGNLIAL
SLGLTPNFKSNFDLAEDAKLQLSKD TYDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDI
LRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDG
GASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ
EDFYFPFLKDNREKIEKILTFRIPIYVVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGS
AQSFIERMTNFDKNLPNEKVLPHKSHLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKA
IVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLD
NEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLIN
GIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA
GSPAIIKKGILQTVKVVDLVKVMGRHKPENIVIEAMARENQTTQKGQKNSRERMKRIEEGI
KELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK
DSDIDNKVLTTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNALKITQRKFDNLTKAERGG
LSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRK
DFQFYKVRINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQ
EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLS
MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAK
VEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLELENGR
KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLGKSPEDNEQKQLFVEQHKKHYLDEII
EQISEFSKRVILADANLDKVL SAYNKHDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTID
RKRYTSTKEVL DATLIHQSI TGLYETRIDLSQLGGD

【0300】

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、プログラミング可能なヌクレオチド結合タンパク質は、例えば、Burststein 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タンパク質であり得、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。

【0301】

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、プログラミング可能なヌクレオチド結合タンパク質は、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、プログラミング可能なヌクレオチド結合タンパク質は、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が理解されるべきである。

【0302】

いくつかの実施形態では、Cas9は、*Neisseria meningitidis* Cas9 (NmeCas9) またはそのバリエーションである。Edraki et al. Mol. Cell. (2019) 73(4): 714-726に記載されているNmeCas9の特徴とPAM配列は、その全体が参照により本明細書に組み込まれる。

【0303】

Nme1Cas9の例示的なアミノ酸配列を以下に提供する：
type II CRISPR RNA-guided endonuclease Cas9 [*Neisseria meningitidis*] WP_002235162.1

```
1  maafknpin yilgldigia svgwamveid edenpiclid lgvrverae vpktgdsiam
61  arllarsvrr ltrrrahrll rarrllkreg vlqaadfdn glikslpntp wqlraaaldr
121 kltplewsav llhlikhrgy lsqrkneget adkelgallk gvadnahalq tgdftrtpael
181 alnkfekesg hirnqrgdys htfsrkdlda elillfekqk efgnphvsgg lkegietllm
241 tqrpalsgda vqkmlghctf epaepkaakn tytaerfiwl tklnnlrile qgserpltdt
301 eratlmdepy rkskltyaqa rklglgledta ffkglyrgkd naeastlmem kayhaisral
361 ekeglkdkks plnlspelqd eigtafsllk tdeditgrlk driqpeilea llkhisfdkf
421 vqislkalrr ivplmeqgkr ydeacaeyg dhygkntee kiylppipad eirnpvvlra
```

481 lsqarkving vrrygspar ihietarevg ksfkdrkeie krqueenrkdr ekaaakfrey
 541 fpnfvgepks kdilklrlye qqhgkclysg keinlgrlne kgyveidhal pfsrtwddsf
 601 nnkvlvlgse nqnkgntqpy eyfngkdnsr ewqefkarve tsrfprskkq rillqkfdded
 661 gfkernlndt ryvnrflcqf vadrmlrtgk gkkrvfasng qitnllrgfw glrkvraend
 721 rhhaldavvv acstvamqqk itrfrvrykem nafdgktidk etgevlhqkt hfpqpweffa
 781 qevmirvfgk pdgkpefee dtpeklrtll aeklssrpea vheyvtplfv srpnrkmsg
 841 qghmetvksa krldegvsvl rvpltqlklk dlekmvnrer epklyealka rleahkddpa
 901 kafaepfyky dkagnrtqqv kavrveqvqk tgvwvrnhng iadnatmvrsv dvfekgd

kyy

961 lvpiyswqva kgilpdravv qgkdeedwql iddsfnfkfs lhpndlvevi tkkarmfgyf
 1021 aschrgtgni nirihldhk igkngilegi gvktalsfqk yqidalgkei rpcrlkkrrp
 1081 vr

10

【0304】

Nme2Cas9の例示的なアミノ酸配列を以下に提供する：

type II CRISPR RNA-guided endonuclease Cas9 [Neisseria meningitidis] WP_02230835.1

1 maafknpin yilgldigia svgwamveid eenpirlid lgvrverae vpktgdsalam
 61 arllarsvrr ltrrrahrll rarrllkreg vlqaadfdn glikslpntp wqlraaaldr
 121 kltplewsav llhlikhrgy lsqrkneget adkelgallk gvannahalq tgdftrtpael
 181 alnkfekesg hirnqrgdys htfsrkdldq elillfekqk efgnphvsgg lkegietllm
 241 tqrpalsgda vqkmlghctf epaepkaakn tytaerfiwl tklnnlrile qgserpltdt
 301 eratlmdepy rkskltyaqa rklglledta ffgklyrgkd naeastlmem kayhaisral
 361 ekeglkdkks plnlsselqd eigtafslfk tdeditgrlk drvqpeilea llkhisfdkf
 421 vqislkalrr ivplmeqgkr ydeacaeiyy dhygkntee kiylppipad einrpvvla
 481 lsqarkving vrrygspar ihietarevg ksfkdrkeie krqueenrkdr ekaaakfrey
 541 fpnfvgepks kdilklrlye qqhgkclysg keinlvrlne kgyveidhal pfsrtwddsf
 601 nnkvlvlgse nqnkgntqpy eyfngkdnsr ewqefkarve tsrfprskkq rillqkfdded
 661 gfkernlndt ryvnrflcqf vadhilltgk gkrrvfasng qitnllrgfw glrkvraend
 721 rhhaldavvv acstvamqqk itrfrvrykem nafdgktidk etgkvllhqkt hfpqpweffa
 781 qevmirvfgk pdgkpefee dtpeklrtll aeklssrpea vheyvtplfv srpnrkmsg
 841 ahkdtlrsak rfvkhnekis vkrvwlteik ladlenmvny kngreiele alkaleayg
 901 gnakqafdpk dnpfykkggq lvkavrvekt qesgvllnkk naytiadngd mvrvdvfckv
 961 dkkgknqyfi vpiyawqvae nilpdidckg yriddsytf fslhkydlia fqkdekskve
 1021 fayyincdss ngrfylawhd kgskeqqfri stqnlvliqk yqvnlgkei rpcrlkkrrp
 1081 vr

20

30

【0305】

ある実施形態では、Casタンパク質はCasXまたはCasYである。例示的なCasX ((uniprot.org/uniprot/F0NN87; uniprot.org/uniprot/F0NH53) tr | F0NN87 | F0NN87_SULIHCRISPR-associatedCasx protein OS = Sulfolobus islandicus (strain HVE10/4) GN = SiH_0402 PE=4 SV=1)アミノ酸配列は以下の通りである：

40

MEVPLYNIFGDNYIIQVATEAENSTIYNNKVEIDDEELRNVNLAYKIAKNNEDAAAERRG
 KAKKKKGEEGETTTSNIILPLSGNDKNPWTETLKCYNFPTTVALSEVFNFSQVKECEEVS
 APSFVKPEFYEFGRSPGMVERTRRVKLEVEPHYLIIAAGWVLTRLGKAKVSEG DYVGVN
 VFTPTRGILYSLIQNVNGIVPGIKPETAFLWIARKVVSSVTNPVSVVRIYTISDAVGQN
 PTTINGGFSIDLTKLLEKRYLLSERLEAIARNALSISSNMREYIVLANIYIYEYLTG SKRLE
 DLLYFANRDLIMNLNSDDGKVRDLKLISAYVNGELIRGEG.

【0306】

例示的なCasX (> tr | F0NH53 | F0NH53_SULIR CRISPR associated protein, Casx OS = Sulfolobus islandicus (strain REY15A) GN=SiRe_0771 PE=4 SV=1) アミノ酸配列は以下の通りである：

50

MEVPLYNIFGDNYIIQVATEAENSTIYNNKVEIDDEELRNVNLAYKIAKNNEDAAAERRG
KAKKKKGEEGETTTSNIILPLSGNDKNPWTETLKCYNFPTTVALSEVFNFSQVKECEEVS
APSFVKPEFYKFGRRSPGMVERTRRVKLEVEPHYLMIAAGWVLTRELKAKVSEG DYVGV
NVFTPTRGILYSLIQNVNGIVPGIKPETAFLWIARKVVSSVTNPVSVVSIYITSDAVGQ
NPTTINGGFSIDLTKLLEKRDLLSERLEAIARNALSISSNMRERYIVLANYIYEYLTGSKRLE
DLLYFANRDLIMNLNSDDGKVRDLKLISAYVNGELIRGEG.

【0307】

Deltaproteobacteria CasX

MEKRINKIRKKLSADNATKPVSRSGPMKTLLVRVMTDDLKKRLEKRRKKPEVMPQVISN
NAANNLRMLDDYTKMKEAILQVYWQEFKDDHVGLMCKFAQPASKKIDQNKLPKPEMDE
KGNLT TAGFACSQCQGPLFVYKLEQVSEK GKAYTNYFGRCNVAEHEKLILLAQLKPKVKDS
DEAVTYSLGKFGQRALDFYSIHVTKESTHPVKPLAQIAGNRYASGPVGKALS DACMG TIA
SFLSKYQDIIIEHQKVVKGNQKRLESLRELAKENLEYPSVTLPPQPHTKEGVDFAYNEVI
ARVRM WVNLNLWQKLKLSRDDAKPLLRLKGFPSPVVERRENEVDWWNTINEVKKLID
AKRDMGRVFWSGVTAEK RNTILEGYN YLPNENDHKKREGSLENPKKPAKRQFGDLLLYL
EKKYAGDWGKVFDEAWERIDKKIAGLTSHIEREEARNAEDAQSKAVLTDWLRAKASFVL
ERLKEMDEKEFYACEIQLQK WYGDLRGNPF AVEAENRVVDISGFSIGSDGHSIQYRNLLA
WKYLENGKREFYLLMNYGKKGRIRFTDGTDIKSGKWQGLLYGGGKAKVIDLTFDPDDE
QLIILPLAFGTRQGREFIWN DLLSLETGLIKLANGRVIEKTIYNKKIGRDEPALFVALT FER
REVVDPSNIKPVN LIGVARGENIPAVIALTDPEGCPLPEFKDSSGGPTDILRIGEGYKEKQR
AIQAAKEVEQRRAGGYSRK FASKSRNLADD MVRNSARDLFYHAVTHDAVLVFANLSRGF
GRQGKRTFMTERQYTKMEDWLTAKLAYEGLTSKTYLSKT LAQYTSKTC SNCGFTITYAD
MDVMLVRLKKTSDGWATT LNNKELKA EYQITYYNRYKRQTVEKELSAELDRLSEESGNN
DISKWT KGRRDEALFLLKKRFSHRPVQE QFVCLDCGHEVHAAEQ AALNIARSWFLNSN
STEFKSYKSGKQPFVGAWQAFYKRRLKEVWKPN A

【0308】

例示的なCasY ((ncbi.nlm.nih.gov/protein/APG80656.1) > APG80656.1 CRISP
R-associated protein CasY [uncultured Parcubacteria group bacterium]) アミノ
酸配列は以下の通りである：

MSKRHPRISGVKGYRLHAQRLEYTGKSGAMRTIKYPLYSSPSGGRTVPREIVSAINDDYV
GLYGLSNFDDLYNAEK RNEEKVYSVLD FWDYDCVQYGA VFSYTA PGLLKNVAEVRGGSYE
LTKTLKGSHLYDELQIDKVIKFLNKKEISRANGSLDKLKKDIIDCFKA EYRERHKDQC NKL
ADDIKNAKKDAGASLGERQKKLFRDFFGISEQSENDKPSFTNPLNLTCLLPFDTVNNNR
NRGEVLFNKLKEYAQKLDKNEGSLEMWEYIGIGNSGTAFSNFLGEGFLGRLRENKITE LK
KAMMDITDAWRGQE QEEEELEKRLRILAALT IKLREP KFDNH WGGYRSDINGKLSSWLQN
YINQTVKIKEDLKGHKKDLKKAKEMINRFGESDTKEEAVVSSLLESIEKIVPDD SADDEKP
DIPAIAIYRRFLSDGRLTLNRFVQREDVQEALIKERLEAEK KKKPKKRKKKS DA EDEKETID
FKELFPHLAKPLKLVPNFYGD SKRELYKKYKNA AIYTDALWKAVEKIYKSAFSSSLKNSFF
DTDFDKDFFIKRLQKIFSVYRRFNTDKWKPIVKNSFAPYCDIVSLAENEVLYKPKQSR SRK
SAAIDKNRVRLPSTENIAKAGIALARELSVAGFDWKDLLKKEEHEEYIDLIELHKTALALL
LAVTETQLDISALDFVENGTVKDFMKT RDGNLVLEGRFLEMFSQSIVFSELRGLAGLMSR
KEFITRSAIQTMNGKQAELLYIPHEFQSAKITTPKEMSRAFLDLAPAEFATSLEPE SLSEKS
LLKLKQMRYYPHYFGYELTRTGQGIDGGVAENALRLEKSPVKKREIKCKQYKTLGRGQNK
IVLYVRSSYYQTQFLEWFLHRPKNVQTDVAVSGSFLIDEKKVKTRWNYDALTVALEPVS
GSE RVFVSQPFTIFPEKS AEEEGQRYLGIDIGEYGIAYTALEITGDSAKILDQNFISDPQLKT
LREEVKGLKLDQRRGTFAMPSTKIARIRESLVHSLRNRIHHLALKHKAKIVYELEVSRFEE
GKQKIKKVYATLKKADVYSEIDADKNLQTTVWGKLAVASEISASYTSQFCGACKKLWRAE
MQVDETITTQELIGTVRVIKGGTLIDAIKDFMRPPIFDENDTPFPKYRDFCDKHHISKMR
GNSCLFICPFCRANADADIQASQTIALLRVKEEKKVEDYFERFRKLKNIKVLGQMKKI.

【0309】

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、Cas9は標的DNAに結合すると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) 効率は低いが高忠実度の高い相同性誘導修復 (HDR) 経路。

【0310】

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、ある実施形態では、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。)

【0311】

ある実施形態では、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 (insertion) や欠失 (deletion) (インデル: indel) 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)。

【0312】

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、重要な実証的意味を有する。ほとんどの実施形態で、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。

【0313】

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。

【0314】

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。修復タイプは、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。

【0315】

いくつかの実施態様において、Cas9は、修飾（改変）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て行うこともできる。

10

【0316】

ある実施形態では、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。ある実施形態では、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。

20

【0317】

ある実施形態では、バリエーションCas9タンパク質は低減されたヌクレアーゼ活性を有する。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。

【0318】

ある実施形態では、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）。

30

【0319】

ある実施形態では、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鎖の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。

40

50

【0320】

ある実施形態では、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、ある実施形態では、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。

【0321】

別の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、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。

10

【0322】

別の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、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。

【0323】

別の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、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、いくつかの実施形態において、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における触媒のHis残基が回復されている(A840H)。

20

【0324】

別の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、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、いくつかの実施形態において、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。ある実施形態において、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、このようないくつかの実施形態では、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、ある実施形態では、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

30

40

50

ができる。また、アラニン置換以外の変異も好適である。

【0325】

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）。

【0326】

いくつかの実施形態では、バリエーションCasタンパク質はspCas9、spCas9-VRQR、spCas9-VRER、xCas9 (sp)、saCas9、saCas9-KKH、spCas9-MQKSER、spCas9-LRKIQK、またはspCas9-LRVSQLであり得る。

【0327】

いくつかの実施形態において、アミノ酸置換D1135M、S1136Q、G1218K、E1219F、A1322R、D1332A、R1335E、およびT1337Rを含み（SpCas9-MQKFRAER）、改変されたPAM5'-NGC-3'に対する特異性を有する改変されたSpCas9が使用された。

【0328】

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。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り、CRISPR/Cas9系の制限のいくつかを克服する。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'に隣接するプロトスペーサー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。

【0329】

[核酸塩基エディターのCas12ドメイン]

通常、微生物CRISPR-Casシステムはクラス1システムとクラス2システムに分けられる。クラス1システムはマルチサブユニットエフェクター複合体を有する一方、クラス2システムは単一のタンパク質エフェクターを有する。たとえば、Cas9とCpf1は、タイプは異なるが（それぞれ、タイプIIとタイプV）クラス2エフェクターである。Cpf1に加えて、クラス2、タイプV CRISPR-Casシステムは、Cas12a/Cpf1、Cas12b/C2cl、Cas12c/C

10

20

30

40

50

2c3、Cas12d/CasY、Cas12e/CasX、Cas12g、Cas12h、およびCas12iも含む。たとえば、Shmakov et al., “Discovery and Functional Characterization of Diverse Class 2 CRISPR Cas Systems,” Mol. Cell, 2015 Nov. 5; 60(3): 385-397; Makarova et al., “Classification and Nomenclature of CRISPR-Cas Systems: Where from Here?” CRISPR Journal, 2018, 1(5): 325-336; およびYan et al., “Functionally Diverse Type V CRISPR-Cas Systems,” Science, 2019 Jan. 4; 363: 88-91参照; 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。タイプV Casタンパク質は、RuvC（またはRuvC様）エンドヌクレアーゼドメインを含有する。成熟CRISPR RNA (crRNA) の生成は一般にtracrRNAに依存しないが、たとえばCas12b/C2c1は、crRNAの生成にtracrRNAを必要とする。Cas12b/C2c1は、DNA切断のためにcrRNAとtracrRNAの両方に依存する。

10

【0330】

本発明で企図される核酸プログラム可能なDNA結合タンパク質には、クラス2、タイプVとして分類されるCasタンパク質（Cas12タンパク質）が含まれる。Casクラス2、タイプVタンパク質の非限定的な例には、Cas12a/Cpf1、Cas12b/C2c1、Cas12c/C2c3、Cas12d/CasY、Cas12e/CasX、Cas12g、Cas12h、およびCas12i、それらのホモログ、またはそれらの改変バージョンが含まれる。本明細書で使用される場合、Cas12タンパク質は、Cas12ヌクレアーゼ、Cas12ドメイン、またはCas12タンパク質ドメインと呼ばれることもある。いくつかの実施形態において、本発明のCas12タンパク質は、デアミナーゼドメインなどの内部的に融合されたタンパク質ドメインによって中断されたアミノ酸配列を含む。

20

【0331】

いくつかの実施形態において、Cas12ドメインは、ヌクレアーゼ不活性なCas12ドメインまたはCas12ニッカーゼである。いくつかの実施形態において、Cas12ドメインは、ヌクレアーゼ活性ドメインである。例えば、Cas12ドメインは、二本鎖核酸（例えば、二本鎖DNA分子）の一本の鎖にニックを入れるCas12ドメインであり得る。いくつかの実施形態では、Cas12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、Cas12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において、Cas12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、Cas12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の同一の連続するアミノ酸残基を有するアミノ酸配列を含む。

30

40

【0332】

いくつかの実施形態において、Cas12のフラグメント（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Cas12ドメインのうちの1つを含む：（1）Cas12のgRNA結合ドメイン、または（2）Cas12のDNA切断ドメイン。いくつかの実施形態において、Cas12またはそのフラグメントを含むタンパク質は、「Cas12バリエーション」と呼ばれる。Cas12バリエーションは、Cas12またはそのフラグメントと相同性を共有する。例えば、Cas12バリエーションは、野生型Cas12に対し

50

て少なくとも約70%の同一性、少なくとも約80%の同一性、少なくとも約90%の同一性、少なくとも約95%の同一性、少なくとも約96%の同一性、少なくとも約97%の同一性、少なくとも約98%の同一性、少なくとも約99%の同一性、少なくとも約99.5%の同一性、または少なくとも約99.9%の同一性である。いくつかの実施形態において、Cas12バリエーションは、野生型Cas12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、Cas12バリエーションは、Cas12のフラグメント（例えば、gRNA結合ドメインまたはDNA切断ドメイン）を含み、その結果、そのフラグメントは、野生型Cas12の対応するフラグメントに対して少なくとも約70%の同一性、少なくとも約80%の同一性、少なくとも約90%の同一性、少なくとも約95%の同一性、少なくとも約96%の同一性、少なくとも約97%の同一性、少なくとも約98%の同一性、少なくとも約99%の同一性、少なくとも約99.5%の同一性、または少なくとも約99.9%の同一性を有する。いくつかの実施形態において、フラグメントは、対応する野生型Cas12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%、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%の長さである。いくつかの実施形態において、フラグメントは、少なくとも100アミノ酸の長さである。いくつかの実施形態では、フラグメントは、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。

【0333】

いくつかの実施形態において、Cas12は、Cas12ヌクレアーゼ活性を変化させる1つ以上の突然変異を有するCas12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。そのような突然変異は、例として、Cas12のRuvCヌクレアーゼドメイン内のアミノ酸置換を含む。いくつかの実施形態において、野生型Cas12に対して少なくとも約70%の同一性、少なくとも約80%の同一性、少なくとも約90%の同一性、少なくとも約95%の同一性、少なくとも約98%の同一性、少なくとも約99%の同一性、少なくとも約99.5%の同一性、または少なくとも約99.9%の同一性であるCas12のバリエーションまたはホモログが提供される。いくつかの実施形態において、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、またはそれだけ長いアミノ酸配列を有するCas12のバリエーションが提供される。

【0334】

いくつかの実施形態では、本明細書で提供されるCas12融合タンパク質は、Cas12タンパク質の完全長アミノ酸配列、例えば、本明細書で提供されるCas12配列の1つを含む。しかしながら、他の実施形態では、本明細書で提供される融合タンパク質は、完全長のCas12配列を含まず、その1つまたは複数のフラグメントのみを含む。適切なCas12ドメインの例示的なアミノ酸配列が本明細書で提供され、Cas12ドメインおよびフラグメントの追加の適切な配列は、当業者には明らかであろう。

【0335】

一般に、クラス2、タイプV Casタンパク質は、単一の機能的なRuvCエンドヌクレアーゼドメインを有する（たとえば、Chen et al., “CRISPR-Cas12a target binding unleashes indiscriminate single-stranded DNase activity,” Science 360:436-439 (2018)参照）。場合によっては、Cas12タンパク質はバリエーションCas12bタンパク質である。（Strecker et al., Nature Communications, 2019, 10(1): Art. No.: 212を参照）。一実施形態では、バリエーションCas12ポリペプチドは、野生型Cas12タンパク質のアミノ酸配列と比較した場合、1、2、3、4、5またはそれ以上のアミノ酸が異なる（例えば、

10

20

30

40

50

欠失、挿入、置換、融合を有する) アミノ酸配列を有する。場合によっては、バリエーション Cas12 ポリペプチドは、Cas12 ポリペプチドの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、場合によっては、バリエーション Cas12 は、対応する野生型 Cas12b タンパク質のニッカーゼ活性の 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 Cas12b ポリペプチドである。場合によっては、バリエーション Cas12b タンパク質には実質的なニッカーゼ活性がない。

【0336】

場合によっては、バリエーション Cas12b タンパク質はニッカーゼ活性が低減されている。例えば、バリエーション Cas12b タンパク質は、野生型 Cas12b タンパク質のニッカーゼ活性の約 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。

【0337】

いくつかの実施形態において、Cas12 タンパク質は、哺乳動物細胞において活性を示す、Cas12a/Cpf1 ファミリーからの RNA 誘導エンドヌクレアーゼを含む。Prevotella および Francisella 1 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 システムの RNA ガイドエンドヌクレアーゼである。この獲得免疫メカニズムは、Prevotella および Francisella 細菌に見出される。Cpf1 遺伝子は CRISPR 遺伝子座に付随しており、ガイド RNA を使用してウイルス DNA を見つけて切断するエンドヌクレアーゼをコードしている。Cpf1 は Cas9 よりも小さくて単純なエンドヌクレアーゼであり、CRISPR/Cas9 システムの制約のいくつかを克服する。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' オーバーハングを伴う二本鎖切断である。Cpf1 のジグザグ状の切断パターンは、従来の制限酵素クローニングと同様に、遺伝子編集の効率を高めることができる、方向性のある遺伝子導入の可能性を開くことができる。上記の Cas9 バリエーションおよびオーソログと同様に、Cpf1 は、CRISPR によって標的にできるサイトの数を、SpCas9 が好む NGG PAM サイトを欠く AT リッチ領域または AT リッチゲノムに拡張することもできる。Cpf1 遺伝子座には、混合アルファ/ベータドメイン、RuvC-I とそれに続くヘリカル領域、RuvC-II、およびジンクフィンガー様ドメインが含まれている。Cpf1 タンパク質には、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がある。さらに、Cpf1 は Cas9 とは異なり、HNH エンドヌクレアーゼドメインを持たず、Cpf1 の N 末端には Cas9 のアルファヘリックス認識ローブがない。Cpf1 CRISPR-Cas ドメインアーキテクチャは、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ている。Cpf1 遺伝子座は Cas1、Cas2、および Cas4 タンパク質をコードし、タイプ II システムよりもタイプ I および III に類似している。機能的 Cpf1 はトランス活性化 CRISPR RNA (tracrRNA) を必要としないため、CRISPR (crRNA) のみが必要となる。Cpf1 は Cas9 よりも小さいだけでなく、sgRNA 分子もより小さい (Cas9 の約半分のヌクレオチド数) ため、これはゲノム編集において有利となる。Cpf1-crRNA 複合体は、Cas9 が標的とする G リッチ PAM とは対照的に、プロトスペーサー隣接モチーフ 5'-YTN-3' または 5'-TTTN-3' を同定することにより標的 DNA または RNA を切断する。PAM の同定後、Cpf1 は、4 または 5 ヌクレオチドのオーバーハングを持つスティッキーエンド様 DNA 二本鎖切断を導入する。

【0338】

本発明のいくつかの局面において、ベクターは、対応する野生型酵素に関して変異した CRISPR 酵素をコードし、その結果、変異した CRISPR 酵素は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く。Cas12 は、野生型の例示的な Cas12 ポリペプチド (例えば、Bacillus hisashii からの Cas12) に対して少なくとも、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。Cas12 は、野生型の例示的な Cas12 ポリペプチド (例えば、Bacillus hisashii (BhCas12b)、Bacillus sp. V3-13 (BvCas12b)、および Alicyclobacillus acidiphilus (AaCas12b) 由来のもの) に対して最大で、または最大でおよ

10

20

30

40

50

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/または配列相同性のポリペプチドを指すことができる。Cas12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、もしくはそれらの任意の組み合わせなどのアミノ酸変化を含み得るCas12タンパク質の改変形態を指すことができる。

【0339】

〔核酸によりプログラム可能なDNA結合タンパク質〕

本開示のある側面は、核酸プログラミング可能なDNA結合タンパク質として作用するドメインを含む融合タンパク質を提供し、これは、塩基エディターのようなタンパク質を特異的な核酸（例えばDNAまたはRNA）配列に誘導するために使用され得る。特定の実施形態において、融合タンパク質は、核酸プログラミング可能なDNA結合タンパク質ドメインとデアミナーゼドメインとを含む。核酸によりプログラム可能なDNA結合タンパク質の非限定的な例としては、Cas9（例えば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。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Csx12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、Cse2、Cse3、Cse4、Cse5e、Csc1、Csc2、Csa5、Csn1、Csn2、Csm1、Csm2、Csm3、Csm4、Csm5、Csm6、Cmr1、Cmr3、Cmr4、Cmr5、Cmr6、Csb1、Csb2、Csb3、Csx17、Csx14、Csx10、Csx16、CsaX、Csx3、Csx1、Csx1S、Csx11、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、またはそれらの改変または操作されたバージョンが挙げられる。他の核酸プログラミング可能なDNA結合タンパク質も、本開示に具体的に列記されていないかもしれないが、本開示の範囲内である。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を参照（それぞれの全内容が参照により本明細書に組み込まれる）。

【0340】

Cas9とは異なるPAM特異性を有する核酸プログラミング可能DNA結合タンパク質の一例はPrevotella および Francisella由来のClustered Regularly Interspaced Short Palindromic Repeats (Cpf1) である。Cas9と同様に、Cpf1もクラス2 CRISPRエフェクターである。Cpf1はCas9とは異なる特徴でロバストな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タンパク質は当技術分野で知られており、例えば過去に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。

【0341】

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) バリエーションは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のアルファ-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。

【0342】

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ミング可能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。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が理解されるべきである。

【0343】

野生型Francisella novicida Cpf1 (D917, E1006, およびD1255は太字で下線を引いている。)

MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYHQFFI
EEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSEKFKNLFN
QNLIDAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIISFKGWTTFYFKGFHEN
RKNVYSSNDIPTSIIRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIKKDLAEELTFDID
YKTSEVNQRVFSLDEVFEIANFNLYLNQSGITKFNTIIGGKFVNGENTKRKGINEYINLYS
QQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTMMQSFYEQIAAFKTVEEKS
IKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDYSVIGTAVLEYITQQIAPKNLD
NPSKKEQELIAKKTEKAKYLSLETIKLAALEEFNKHRDIDKQCRFEEILANFAAIPMIFDEIA
QNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIKDLLDQTNLLHKLKIFHISQSEDKAN
ILDKDEHFYLVFEECYFELANIVPLYNKIRNYITQKPYSDEKFKLNFENSTLANGWDKNKE
PDNTAILFIKDDKYYLGVMMNKKNNKIFDDKAIKENKGEYKKIVYKLLPGANKMLPKVFF
SAKSIKFYNPSEDILRIRNHSTHTKNGSPQKGYEKFEFNIEDCRKFIDFYKQSISKHPEWKD
FGFRFSDTQRYNSIDEFYREVENQGYKLTFENISESYIDSVVNQGKLYLFQIYNKDFSAYS
KGRPNLHTLYWKALFDERNLQDVVYKLNGEAEFYRKQSIPKKITHPAKEAIANKNKDNP
KKESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFNDEINLLLKEKANDVHILSIDRG
ERHLAYYTLVDGKGNIIKQDTFNIIGNDRMKTNYHDKLAAIEKDRDSARKDWKKINNIKE
MKEGYLSQVVHEIAKLVEYNAIVVFEDLNFGRFRGKVEKQVYQKLEKMLIEKLNLYLV
FKDNEFDKTGGVLRAYQLTAPFETFKKMGGKQTGIIYYVPAGFTSKICPVTGTVNQLYPKY
ESVSKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKAAKGKWTIASFGSRLINFRNSDKN
HNWDTREVPYPTKELEKLLKDYSIEYGHGECIKAAICGESDKKFFAKLTSVLNTILQMRNSK
TGTELDYLISPVADVNGNFFDSRQAPKNMPQDADANGAYHIGLKGLMLLGRIKNNQEGK
KLNLVIKNEEYFEFVQNRNN

【0344】

Francisella novicida Cpf1 D917A (A917, E1006, およびD1255は太字で下線を引いている。)

MSIQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYHQFFI
EEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSEKFNLFN
QNLIDAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIISFKGWTTFYFKGFHEN
RKNVYSSNDIPTSIIRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIKKDLAEELTFDID
YKTSEVNQRVFSLDEVFEIANFNLYLNQSGITKFNTIIGGKFVNGENTKRKGINEYINLYS
QQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTMMQSFYEQIAAFKTVEEKS
IKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDYSVIGTAVLEYITQQIAPKNLD
NPSKKEQELIAKKTEKAKYLSLETIKLAL EEFNKHRDIDKQCRFEEILANFAAIPMIFDEIA
QNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIKDLLDQTNLLHKLKIFHISQSEDKAN
ILDKDEHFYLVFEECYFELANIVPLYNKIRNYITQKPYSDEKFKLNFENSTLANGWDKNKE
PDNTAILFIKDDKYYLGVMNKKNNKIFDDKAIKENKGEYKKIVYKLLPGANKMLPKVFF
SAKSIKFYNPSEDILRIRNHSTHTKNGSPQKGYEKFENIEDCRKFIDFYKQSISKHPEWKD
FGFRFSDTQRYNSIDEFYREVENQGYKLTFENISESYIDSVVNQGKLYLFQIYNKDFSAYS
KGRPNLHTLYWKALFDERNLQDVVYKLNGEAEFYRKQSIPKKITHPAKEAIAKNKNDNP
KKESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFNDEINLLLKEKANDVHILSIARG
ERHLAYYTLVDGKGNIQKQDTFNIIGNDRMKTNYHDKLAAIEKDRDSARKDWKKINNIKE
MKEGYLSQVVHEIAKLVEYNAIVVFEDLNF¹GF²KRGRF³KEKQVYQKLEKMLIEKLNLYLV
FKDNEFDKTGGVLRAYQLTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGFVNQLYPKY
ESVSKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKAAKGKWTIASFGSRLINFRNSDKN
HNWDTREVVYPTKELEKLLKDYSIEYGHGECIKAAICGESDKKFFAKLTSVLNTILQMRNSK
TGTELDYLISPVADVNGNFFDSRQAPKNMPQDADANGAYHIGLKGLMLLGRIKNNQEGK
KLNLVIKNEEYFEFVQNRNN

10

20

【0345】

Francisella novicidaCpf1 E1006A (D917, A1006,および D1255は太字で下線を引いている。)

MSIQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYHQFFI
EEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSEKFNLFN
QNLIDAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIISFKGWTTFYFKGFHEN
RKNVYSSNDIPTSIIRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIKKDLAEELTFDID
YKTSEVNQRVFSLDEVFEIANFNLYLNQSGITKFNTIIGGKFVNGENTKRKGINEYINLYS
QQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTMMQSFYEQIAAFKTVEEKS
IKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDYSVIGTAVLEYITQQIAPKNLD
NPSKKEQELIAKKTEKAKYLSLETIKLAL EEFNKHRDIDKQCRFEEILANFAAIPMIFDEIA
QNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIKDLLDQTNLLHKLKIFHISQSEDKAN
ILDKDEHFYLVFEECYFELANIVPLYNKIRNYITQKPYSDEKFKLNFENSTLANGWDKNKE
PDNTAILFIKDDKYYLGVMNKKNNKIFDDKAIKENKGEYKKIVYKLLPGANKMLPKVFF
SAKSIKFYNPSEDILRIRNHSTHTKNGSPQKGYEKFENIEDCRKFIDFYKQSISKHPEWKD
FGFRFSDTQRYNSIDEFYREVENQGYKLTFENISESYIDSVVNQGKLYLFQIYNKDFSAYS
KGRPNLHTLYWKALFDERNLQDVVYKLNGEAEFYRKQSIPKKITHPAKEAIAKNKNDNP
KKESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFNDEINLLLKEKANDVHILSIDRG
ERHLAYYTLVDGKGNIQKQDTFNIIGNDRMKTNYHDKLAAIEKDRDSARKDWKKINNIKE
MKEGYLSQVVHEIAKLVEYNAIVVFADLNF¹GF²KRGRF³KEKQVYQKLEKMLIEKLNLYLV
FKDNEFDKTGGVLRAYQLTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGFVNQLYPKY
ESVSKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKAAKGKWTIASFGSRLINFRNSDKN
HNWDTREVVYPTKELEKLLKDYSIEYGHGECIKAAICGESDKKFFAKLTSVLNTILQMRNSK
TGTELDYLISPVADVNGNFFDSRQAPKNMPQDADANGAYHIGLKGLMLLGRIKNNQEGK
KLNLVIKNEEYFEFVQNRNN

30

40

50

【0346】

Francisella novicida Cpf1 D1255A (D917, E1006, および A1255は太字で下線を引いている。)

MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYHQFFI
EEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSEKFNLFN
QNLIDAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIISFKGWTTFYFKGFHEN
RKNVYSSNDIPTSIIRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIKKDLAEELTFDID
YKTSEVNQRVFSLDEVFEIANFNLYLNQSGITKFNTIIGGKFVNGENTKRKGINEYINLYS
QQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTMMQSFYEQIAAFKTVEEKS
IKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDYSVIGTAVLEYITQQIAPKNLD
NPSKKEQELIAKKTEKAKYLSLETIKLAL EEFNKHRDIDKQCRFEEILANFAAIPMIFDEIA
QNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIKDLLDQTNLLHKLKIFHISQSEDKAN
ILDKDEHFYLVFEECYFELANIVPLYNKIRNYITQKPYSDEKFKLNFENSTLANGWDKNKE
PDNTAILFIKDDKYYLGVMNKKNNKIFDDKAIKENKGEYKKIVYKLLPGANKMLPKVFF
SAKSIKFYNPSEDILRIRNHSTHTKNGSPQKGYEKFENIEDCRKFIDFYKQSISKHPEWKD
FGFRFSDTQRYNSIDEFYREVENQGYKLTFENISESYIDSVVNQGKLYLFQIYNKDFSAYS
KGRPNLHTLYWKALFDERNLQDVVYKLNGEAEFYRKQSIPKKITHPAKEAIANKNDNP
KKESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFNDEINLLLKEKANDVHILSIDRG
ERHLAYYTLVDGKGNIQKQDTFNIIGNDRMKTNYHDKLAAIEKDRDSARKDWKKINNIKE
MKEGYLSQVVHEIAKLVEYNAIVVFEDLNFGFKRGRFKVEKQVYQKLEKMLIEKLNLYLV
FKDNEFDKTGGVLRAYQLTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGFVNQLYPKY
ESVSKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKAAKGKWTIASFGSRLINFRNSDKN
HNWDTREVVYPTKELEKLLKDYSIEYGHGECIKAAICGESDKKFFAKLTSVLNTILQMRNSK
TGTELDYLISPVADVNGNFFDSRQAPKNMPQDAANGAYHIGLKGLMLLGRIKNNQEGK
KLNLVIKNEEYFEFVQNRNN

10

20

【0347】

Francisella novicida Cpf1 D917A/E1006A (A917, A1006, および D1255は太字で下線を引いている。)

MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYHQFFI
EEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSEKFNLFN
QNLIDAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIISFKGWTTFYFKGFHEN
RKNVYSSNDIPTSIIRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIKKDLAEELTFDID
YKTSEVNQRVFSLDEVFEIANFNLYLNQSGITKFNTIIGGKFVNGENTKRKGINEYINLYS
QQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTMMQSFYEQIAAFKTVEEKS
IKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDYSVIGTAVLEYITQQIAPKNLD
NPSKKEQELIAKKTEKAKYLSLETIKLAL EEFNKHRDIDKQCRFEEILANFAAIPMIFDEIA
QNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIKDLLDQTNLLHKLKIFHISQSEDKAN
ILDKDEHFYLVFEECYFELANIVPLYNKIRNYITQKPYSDEKFKLNFENSTLANGWDKNKE
PDNTAILFIKDDKYYLGVMNKKNNKIFDDKAIKENKGEYKKIVYKLLPGANKMLPKVFF
SAKSIKFYNPSEDILRIRNHSTHTKNGSPQKGYEKFENIEDCRKFIDFYKQSISKHPEWKD
FGFRFSDTQRYNSIDEFYREVENQGYKLTFENISESYIDSVVNQGKLYLFQIYNKDFSAYS
KGRPNLHTLYWKALFDERNLQDVVYKLNGEAEFYRKQSIPKKITHPAKEAIANKNDNP
KKESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFNDEINLLLKEKANDVHILSIARG
ERHLAYYTLVDGKGNIQKQDTFNIIGNDRMKTNYHDKLAAIEKDRDSARKDWKKINNIKE
MKEGYLSQVVHEIAKLVEYNAIVVFADLNFGFKRGRFKVEKQVYQKLEKMLIEKLNLYLV
FKDNEFDKTGGVLRAYQLTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGFVNQLYPKY
ESVSKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKAAKGKWTIASFGSRLINFRNSDKN
HNWDTREVVYPTKELEKLLKDYSIEYGHGECIKAAICGESDKKFFAKLTSVLNTILQMRNSK
TGTELDYLISPVADVNGNFFDSRQAPKNMPQDADANGAYHIGLKGLMLLGRIKNNQEGK
KLNLVIKNEEYFEFVQNRNN

30

40

50

【0348】

Francisella novicida Cpf1 D917A/D1255A (A917, E1006, および A1255は太字で下線を引いている。)

MSIQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYHQFFI
EEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSEKFKNLFN
QNLIDAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIISFKGWTTFYFKGFHEN
RKNVYSSNDIPTSIIRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIKKDLAEELTFDID
YKTSEVNQRVFSLDEVFEIANFNLYLNQSGITKFNTIIGGKFVNGENTKRKGINEYINLYS
QQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTMMQSFYEQIAAFKTVEEKS
IKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDYSVIGTAVLEYITQQIAPKNLD
NPSKKEQELIAKKTEKAKYLSLETIKLAL EEFNKHRDIDKQCRFEEILANFAAIPMIFDEIA
QNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIKDLLDQTNLLHKLKIFHISQSEDKAN
ILDKDEHFYLVFEECYFELANIVPLYNKIRNYITQKPYSDEKFKLNFENSTLANGWDKNKE
PDNTAILFIKDDKYYLGVMNKKNNKIFDDKAIKENKGEYKKIVYKLLPGANKMLPKVFF
SAKSIKFYNPSEDILRIRNHSTHTKNGSPQKGYEKFENIEDCRKFIDFYKQSISKHPEWKD
FGFRFSDTQRYNSIDEFYREVENQGYKLTFENISESYIDSVVNQGKLYLFQIYNKDFSAYS
KGRPNLHTLYWKALFDERNLQDVVYKLNGEAEFYRKQSIPKKITHPAKEAIANKNDNP
KKESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFNDEINLLLKEKANDVHILSIARG
ERHLAYYTLVDGKGNIQKQDTFNIIGNDRMKTNYHDKLAAIEKDRDSARKDWKKINNIKE
MKEGYLSQVVHEIAKLVEYNAIVVFEDLNFGFKRGRFKVEKQVYQKLEKMLIEKLNLYLV
FKDNEFDKTGGVLRAYQLTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGFVNQLYPKY
ESVSKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKAAKGKWTIASFGSRLINFRNSDKN
HNWDTREVVYPTKELEKLLKDYSIEYGHGECIKAAICGESDKKFFAKLTSVLNTILQMRNSK
TGTELDYLISPVADVNGNFFDSRQAPKNMPQDAANGAYHIGLKGLMLLGRIKNNQEGK
KLNLVIKNEEYFEFVQNRNN

10

20

【0349】

Francisella novicida Cpf1 E1006A/D1255A (D917, A1006, および A1255は太字で下線を引いている。)

MSIQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYHQFFI
EEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSEKFKNLFN
QNLIDAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIISFKGWTTFYFKGFHEN
RKNVYSSNDIPTSIIRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIKKDLAEELTFDID
YKTSEVNQRVFSLDEVFEIANFNLYLNQSGITKFNTIIGGKFVNGENTKRKGINEYINLYS
QQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTMMQSFYEQIAAFKTVEEKS
IKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDYSVIGTAVLEYITQQIAPKNLD
NPSKKEQELIAKKTEKAKYLSLETIKLAL EEFNKHRDIDKQCRFEEILANFAAIPMIFDEIA
QNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIKDLLDQTNLLHKLKIFHISQSEDKAN
ILDKDEHFYLVFEECYFELANIVPLYNKIRNYITQKPYSDEKFKLNFENSTLANGWDKNKE
PDNTAILFIKDDKYYLGVMNKKNNKIFDDKAIKENKGEYKKIVYKLLPGANKMLPKVFF
SAKSIKFYNPSEDILRIRNHSTHTKNGSPQKGYEKFENIEDCRKFIDFYKQSISKHPEWKD
FGFRFSDTQRYNSIDEFYREVENQGYKLTFENISESYIDSVVNQGKLYLFQIYNKDFSAYS
KGRPNLHTLYWKALFDERNLQDVVYKLNGEAEFYRKQSIPKKITHPAKEAIANKNDNP
KKESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFNDEINLLLKEKANDVHILSIDRG
ERHLAYYTLVDGKGNIQKQDTFNIIGNDRMKTNYHDKLAAIEKDRDSARKDWKKINNIKE
MKEGYLSQVVHEIAKLVEYNAIVVFADLNFGFKRGRFKVEKQVYQKLEKMLIEKLNLYLV
FKDNEFDKTGGVLRAYQLTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGFVNQLYPKY
ESVSKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKAAKGKWTIASFGSRLINFRNSDKN
HNWDTREVVYPTKELEKLLKDYSIEYGHGECIKAAICGESDKKFFAKLTSVLNTILQMRNSK
TGTELDYLISPVADVNGNFFDSRQAPKNMPQDAANGAYHIGLKGLMLLGRIKNNQEGK
KLNLVIKNEEYFEFVQNRNN

30

40

50

【0350】

Francisella novicida Cpf1 D917A/E1006A/D1255A (A917, A1006, および A1255は太字で下線を引いている。)

MSIQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYHQFFI
 EEILSSVCISEDLLQNYSDVYFKLKKSDDDLNQLKDFKSAKDTIKKQISEYIKDSEKFKNLFN
 QNLIDAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIISFKGWTTFYFKGFHEN
 RKNVYSSNDIPTSIYIRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIKKDLAEELTFDID
 YKTSEVNQRVFSLDEVFEIANFNLYLNQSGITKFNTIIGGKFVNGENTKRKGINEYINLYS
 QQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTMQSFYEQIAAFKTVEEKS
 IKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDYSVIGTAVLEYITQQIAPKNLD
 NPSKKEQELIAKKTEKAKYLSLETIKLALAEFNKHRDIDKQCRFEEILANFAAIPMIFDEIA
 QNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIKDLLDQTNLLHKLKIFHISQSEDKAN
 ILDKDEHFYLVFEECYFELANIVPLYNKIRNYITQKPYSDEKFKLNFENSTLANGWDKNKE
 PDNTAILFIKDDKYLLGVMNKKNNKIFDDKAIKENKGEYKKIVYKLLPGANKMLPKVFF
 SAKSIKFYNPSEDILRIRNHSTHTKNGSPQKGYEKFEFNIEDCRKFIDFYKQSISKHPEWKD
 FGFRFSDTQRYNSIDEFYREVENQGYKLTFENISESYIDSVVNQGKLYLFQIYNKDFSAYS
 GRPNLHTLYWKALFDERNLQDVVYKLNGEAEFYRKQSIPKKITHPAKEAIAKNKNDNP
 KKEVSFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFNDEINLLLKEKANDVHILSIARG
 ERHLAYYTLVDGKGNIIKQDTFNIIGNDRMKTNYHDKLAAIEKDRDSARKDWKKINNIKE
 MKEGYLSQVVHEIAKLVEYNAIVVFADLNF¹GFGRFRFKEKQVYQKLEKMLIEKLNLYLV
 FKDNEFDKTGGVLRAYQLTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGTVNQLYPKY
 ESVSKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKAAKGKWTIASFGSRLINFRNSDKN
 HNWDTREVVYPTKELEKLLKDYSIEYGHGECIKAAICGESDKKFFAKLTSVLNTILQMRNSK
 TGTELDYLISPVADVNGNFFDSRQAPKNMPQDAANGAYHIGLKGLMLLGRIKNNQEGK
 KNLNLIKNEEYFEFVQNRNN

【0351】

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、PAM配列を必要としないガイドヌクレオチド配列プログラミング可能DNA結合タンパク質ドメインで置換され得る。

【0352】

ある実施形態において、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、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。

【0353】

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、E781K、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。

【0354】

例示的なSaCas9のアミノ酸配列

KRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLLKRRRRH

RIQRVKKLLFDYNLLTDHSELGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGVHNVNE
VEEDTGNELSTKEQISRNSKALEEKYVAELQLERLKKDGEVRGSINRFKTS DYVKEAKQLL
KVQKAYHQLDQSFIDTYIDLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEE
LRSVKYAYNADLYNALNDLNNLVITRDENEKLEYEKFQIIENVFKQKKKPTLKQIAKEIL
VNEEDIKGYRVTSTGKPEFTNLKVYHDIKDITARKEIENAELLDQIAKILTIYQSSEDIQEE
LTNLNSELQTQEEIEQISNLKGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPKKV
DLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIKKYGLPNDIIELAREKNSKDAQMI
NEMQKRNRQTNERIEEIIRTTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNNPFN
YEVDHIIPRSVSFDNSFNNKVLVKQEENSKKGNRTPFQYLSSSDSKISYETFKKHILNLAK
GKGRISKTKKEYLLEERDINRFSVQKDFINRNLVDTRYATRGLMNLRLSYFRVNNLDVKV
KSINGGFTSFLRRKWKFKKERNKGYKHHAEADALIINANADFIFKEWKKLDKAKKVMENQM
FEEKQAESMPEIETE QEYKEIFITPHQIKHIKDFKDYKYSHRVDKKPNRELINDTLYSTRK
DDKGNTLIVNNLNGLYDKDNDKLLKLINKSPEKLLMYHHPQTYQKLKLIMEQYGDEKN
PLYKYYEETGNYLT KYSKKDN GPVIKKIKYYGNKLN AHL DITDDYPNSRNKVVKLSLKP
Y RFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSKCYEEAKKLLKISNQA EFIASFYNNDLIK
INGEL YRVIGVNNDLLNRIEVNMIDITYREYLENMNDKRPPRIIKTIASKTQSIKKYSTDIL
GNLYEVKSKKHPQIIKKG.

10

太字で下線を付している上記残基N579は、（例えばA579に）変異されてSaCas9ニッ
カーゼを生じ得る。

【0355】

20

例示的なSaCas9nのアミノ酸配列

KRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRRRRH
RIQRVKKLLFDYNLLTDHSELGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGVHNVNE
VEEDTGNELSTKEQISRNSKALEEKYVAELQLERLKKDGEVRGSINRFKTS DYVKEAKQLL
KVQKAYHQLDQSFIDTYIDLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEE
LRSVKYAYNADLYNALNDLNNLVITRDENEKLEYEKFQIIENVFKQKKKPTLKQIAKEIL
VNEEDIKGYRVTSTGKPEFTNLKVYHDIKDITARKEIENAELLDQIAKILTIYQSSEDIQEE
LTNLNSELQTQEEIEQISNLKGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPKKV
DLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIKKYGLPNDIIELAREKNSKDAQMI
NEMQKRNRQTNERIEEIIRTTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNNPFN
YEVDHIIPRSVSFDNSFNNKVLVKQEENSKKGNRTPFQYLSSSDSKISYETFKKHILNLAK
GKGRISKTKKEYLLEERDINRFSVQKDFINRNLVDTRYATRGLMNLRLSYFRVNNLDVKV
KSINGGFTSFLRRKWKFKKERNKGYKHHAEADALIINANADFIFKEWKKLDKAKKVMENQM
FEEKQAESMPEIETE QEYKEIFITPHQIKHIKDFKDYKYSHRVDKKPNRELINDTLYSTRK
DDKGNTLIVNNLNGLYDKDNDKLLKLINKSPEKLLMYHHPQTYQKLKLIMEQYGDEKN
PLYKYYEETGNYLT KYSKKDN GPVIKKIKYYGNKLN AHL DITDDYPNSRNKVVKLSLKP
Y RFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSKCYEEAKKLLKISNQA EFIASFYNNDLIK
INGEL YRVIGVNNDLLNRIEVNMIDITYREYLENMNDKRPPRIIKTIASKTQSIKKYSTDIL
GNLYEVKSKKHPQIIKKG.

30

N579から変異されてSaCas9ニッカーゼを生じることができる上記残基A579には太字
で下線を付している。

【0356】

40

例示的なSaKKH Cas9のアミノ酸配列

50

KRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRRRRHRIQ
RVKKLLFDYNLLTDHSELSGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGVHNVNEVEEDTG
NELSTKEQISRNSKALEEKYVAELQLERLKKDGEVRSINRFKTSYVKEAKQLLKVQKAYHQL
DQSFIDTYIDLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYN
ALNDLNNLVITRDENEKLEYEKFQIIENVFKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGKP
EFTNLKVYHDIKDITARKEIIEAELLDQIAKILTIYQSSEDIQEELTNLNSELTQEEIEQISN
LKGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILSPV
VKRSFIQSIKVINAIIKKYGLPNDIIIEELAREKNSKDAQKMINEMQKRNROTNERIEEII RTTG
KENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNPNFYEVVDHII PRSVSFDNSFNKVLVKQ
EEASKKGNRTPFQYLSSSDSKISYETFKKHILNLA KGGRISKTKKEYLLEERDINRFSVQKDF
INRNLDVTRYATRGLMNLLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKERNKGYKHAEDA
LIIANADFIKWKKLDKAKKVMENQMFEKQAESMPEIETE QEYKEIFITPHQIKHIKDFKDY
KYSHRVDKKPNRKLINDTLYSTRKDDKGNTLIVNNLNGLYDKDNDKLKKLINKSPEKLLMYHHD
PQTYQKLKLIMEQYGDEKNPLYKYEEETGNYLTKYSKKDNGPVIKKIKYYGNKLNAHLDITDDY
PNSRNKVVKLSLKPYRFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSKCYEEAKKLKKISNQAE
FIASFYKNDLIKINGELYRVIGVNNDLLNRIEVNMIDITYREYLENMNDKRPPHIIKTIASKTQ
SIKKYSTDILGNLYEVKSKKHPQIIKKG.

10

20

上記の残基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であり、太字で下線を付している。上記の残基K781、K967、およびH1014は、E781、N967、およびR1014から変異されてSaKKH Cas9を生じることができるものであり、斜体で下線を付している。

【0357】

30

ある実施形態では、napDNAbpは循環置換体 (circular permutant) である。下記配列において、ブレンテキストはアデノシンデアミナーゼ配列を表し、太字の配列はCas9由来の配列を示し、斜体の配列はリンカー配列を表し、下線の配列は二部分核局在化配列を表す。

40

50

CP5 (with MSP “NGC” PID and “D10A” nickase):

EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQ
 VNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFMQPTVAYSVLVAKVEKGKSK
 KLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDLIIKLPKYSLFELNGRKRMLASAKF
 LQKGNEALALPSKYVNFYLAHYEKLKGSPEDEQKQLFVEQHKHYLDEIIIEQISEFSKRVILA
 DANLDKVL SAYNKH RDKPIREQAENIIHLFTLTNLGAPRAFKYFDTTIARKEYRSTKEVLDATL
 IHQSITGLYETRIDLSQLGGDGGSGGSGGSGGSGGSGGSGGMDKKYSIGLAIGTNSVGWAVITD
 EYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS
 NEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADL
 RLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSAR
 LSKSRLENLIAQLPGKKNGLFGNLIASLGLTPNFKSNFDLAEDAKLQLSKD TYDDDLNLL
 AQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQL
 PEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN
 GSIHQIHLGELHAILRRQEDFYFPFLKDNREKIEKILTFRIPIYYVGPLARGNSRFAWMTRKSEE
 TITPWNFEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMR
 KPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGT YHDLK
 IIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLS
 RKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGD SLHEHIANL
 AGSPAIKKGILQTVKVVDLKVVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKEL
 GSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYD VDHIVPQSFLKDDSIDNK
 VLTRSDKNRGKSDNPSEEVVKMKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKR
 QLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQFYKVREINNYHH
 AHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEGADKRTADGSEFESP KKKR
 KV*

【 0 3 5 8 】

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/C2c3)がShmakov et al., “Discovery and Functional Characterization of Diverse Class 2 CRISPR Cas Systems”, Mol. Cell, 2015 Nov. 5; 60(3): 385-397によって記載されている(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。

【0359】

Alicyclobacillus acidoterrastris 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 acidoterrastris* 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。Cas 12b/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。

【0360】

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が理解されるべきである。

【0361】

Cas12b/C2c1 ((uniprot.org/uniprot/T0D7A2#2) sp | T0D7A2 | C2C1_ALIAG CRISPR-associated endonuclease C2c1 OS = *Alicyclobacillus acido-terrastris* (strain ATCC 49025 / DSM 3922 / CIP 106132 / NCIMB 13137/GD3B) GN=c2c1 PE=1 SV=1) のアミノ酸配列は以下の通りである：

MAVKSIVKLRLLDDMPEIRAGLWKLHKEVNAGVRYYTEWLSLLRQENLYRRSPNGDGEQ
ECDKTAEECKAELLERLRARQVENGHRGPAGSDDELLQLARQLYELLVPQAIGAKGDAQ
QIARKFLSPLADKDAVGGLGIAKAGNKPRWVRMREAGEPGWEEEEKAEATRSADRTAD
VLRALADFGKPLMRVYTDSEMSSVEWKPLRKGQAVRTWDRDMFQQAIERMMSWESW
NQRVQGQYAKLVEQKNRFEQKNFVGQEHVHLVNQLQQDMKEASPGLESKEQTAHYVT
GRALRGSDKVFKEKWGLAPDAFPDLYDAEIKNVQRRNTRRFGSHDLFAKLAEPEYQALW
REDASFLTRYAVYNSILRKLNHAKMFATFTLPDATAHPIWTRFDKLGGNLHQYTFLFNEF
GERRHAIRFHKLKLVENGVAEVDVTPISMSEQLDNLLPRDPNEPIALYFRDYGAEQH
FTGEFGGAKIQCRRDQLAHMHRRRGARDVYLVNSVRVQSQSEARGERRPPYAAVFRLVG
DNHRAFVHFDFKLSDYLAEHPPDDGKLGSEGLLSGLRVMSVDLGLRTSASISVFRVARKDEL
KPNSKGRVPFFFPKIGNDNLVAVHERSQLLKLPGETESKDLRAIREERQRTLRLQLRTQLAY
LRLLVRCGSEDVGRRERSWAKLIEQPVDAANHMTDPDWREAFENELQKLKSLHGICSDKE
WMDAVYESVRRVWRHMGKQVRDWRKDVRSGERPKIRGYAKDVVGGNSIEQIEYLERQY

KFLKSWSFSGKVSGQVIRAEKGSRAITLREHIDHAKEDRLKKLADRIIMEALGYVYALDE
RGKGKMWAKYPPCQLILLEELSEYQFNNDRPPSENNQLMQWSHRGVFQELINQAQVHD
LLVGTMYAAFSSRFDARTGAPGIRCRRVPARCTQEHNPFPWWLNKFVVEHTLDACPL
RADDLIPTGEGEIFVSPFSAEEGDFHQIHADLNAAQNLQQRLLWSDFDISQIRLRCDWGEV
DGELVLIPRLTGKRTADSYSNKVFYTTNTGVTYYERERGGKRRKVFAQEKLSEEEAELLVEA
DEAREKSVVLMRDPGSIINRGWTRQKEFSMV NQRIEGLVKQIRSRVPLQDSACENT
GDI.

[0362]

AacCas12b (*Alicyclobacillus acidiphilus*) - WP_067623834

MAVKSMKVKLRDLNMPEIRAGLWKLHTEVNAGVRYYTEWLSLLRQENLYRRSPNGDGE
QECYKTAECKAELLERLRARQVENGHCGPAGSDELLQLARQLYELLVPQAIGAKGDAQ
QIARKFLSPLADKDAVGGLGIAKAGNKPRWVRMREAGEPGWEEEEKAKAEARKSTDRTAD
VLRALADFGKPLMRVYTDSDMSSVQWKPLRKGQAVRTWDRDMFQQAIERMMSWES
WNQRVGEAYAKLVEQKSFEQKNFVGQEHVQLVNQLQQDMKEASHGLESKEQTAHYL
TGRALRGSDKVFKEKWEKLPDAPFDLYDTEIKNVQRRNTRRFGSHDLFAKLAEPKYQAL
WREDASFLTRYAVYNSIVRKLNHAKMFATFTLPDATAHPIWTRFDKLGGNLHQYTFLFN
EFGGRHAIRFQKLLTVEDGVAKEVDDVTVPISMSAQLDDLLPRDPHELVALYFQDYGAE
QHLAGEFGGAKIQYRRDQLNHLHARRGARDVYLNLSVRVQSQSEARGERRPPYAAVFRL
VGDNHRAVFHFDKLSDYLAHPDDGKLGSEGLLSGLRVMSVDLGLRTSASISVFRVARKD
ELKPNSEGRVPFCFPIEGNENLVAVHERSQQLLKLPGETESKDLRAIREERQRTLRLQRTQL
AYLRLLVRCGEDVGRRERSWAKLIEQPM DANQMTPDWREAFEDQLKLSLYGICGDR
EWTEAVYESVRRVWRHMGKQVRDWRKDVRSGERPKIRGYQKDVVGGNSIEQIEYLERQ
YKFLKSWSFSGKVSGQVIRAEKGSRAITLREHIDHAKEDRLKKLADRIIMEALGYVYALD
DERGKGKMWAKYPPCQLILLEELSEYQFNNDRPPSENNQLMQWSHRGVFQELINQAQV
HDLLVGTMYAAFSSRFDARTGAPGIRCRRVPARCAREQNPEPFPWWLNKFVAEHKLDGC
PLRADDLIPTGEGEFFVSPFSAEEGDFHQIHADLNAAQNLQRRLLWSDFDISQIRLRCDWG
EVDGEPVLIPRTTGKRTADSYGNKVFTTKGVTYYERERGGKRRKVFAQEELSEEEAELL
VEADEAREKSVVLMRDPGSIINRGDWTRQKEFSMV NQRIEGLVKQIRSRVRLQESAC
ENTGDI

[0363]

BhCas12b (*Bacillus hisashii*) NCBI Reference Sequence: WP_095142515

MAPKKKRKVGIGHGVPAAATRSFILKIEPNEEVKKGLWKTHEVLNHGIAYYMNILKLIRQE
AIYEHHEQDPKNPKKVSKAEIQAEWDFVLKMQKNSFTHEVDKDEVFNILRELYEELVP
SSVEKKGEANQLSNKFLYPLVDPNSQSGKGTASSGRKPRWYNLKIAGDPSWEEEEKKKWE
EDKKKDPLAKILGKLAEYGLIPLFIPYTDSEPIVKEIKWMEKSRNQSVRRLDKDMFIQAL
ERFLSWESWNLKVKEEYKVEKEYKTLEERIKEDIQALKALEQYEKERQEQLLRDTLNTN
EYRLSKRGLRGWREIIQKWLKMDENEPSEKYLEVFQDYQRKHPREAGDYSVYEFLSKKE
NHFIWRNHPEYPYLYATFCEIDKKKKDAKQQATFTLADPINHPLWVRFEERSGSNLNKY
RILTEQLHTEKLKKKLTVQLDRLIYPTESGGWEEKGKVDIVLLPSRQFYNQIFLDIEEKGK
HAFTYKDESIKFLKGTGGARVQFDRDHLRRYPHKVESGNVGRIFYNMTVNIEPTESPV
SKSLKIHRDDFPKVVNFKPKELTEWIKDSKGKKLKSIGESLEIGLRVMSIDLGQRQAAAASI
FEVVDQKPDIEGKLPPIKGTELYAVHRASFNKLPGETLVKSREVLARKAREDNLKL MNQ
KLNFLRNVLFHFQQFEDITEREKRVTKWISRQENS DVPLVYQDELIQIRELMYKPKDWV
AFLKQLHKKRLEVEIGKEVKHWRKSLSDGRKGLYGISLKNIDEIDTRKFLLRWSLRPTEPG
EVRRLPGQRFAIDQLNHLNALKEDRLKKMANTIIMHALGYCYDVRKKKWQAKNPACQI
ILFEDLSNYPYEERSRFENSKLMKWSRREIPRQVALQGEIYGLQVGEVGAQFSSRFHAK
TGSPGIRCSVVTKELQDNRFKKNLQREGRLTLDKIAVLKEGDLYPDKGGEKFISLSKDRK
CVTTHADINAAQNLQKRFWTRTHGFYKVYCKAYQVDGQTVYIPESKDQKQKIIIEFEGG
YFILKDGVEYEWVNAGKLKIKKGSSKQSSSELVDSILKDSFDLASELKGEKLMLYRDPGSGN
VFPSDKWMAAGVFFGKLERILISKLTNQYSISTIEDSSKQSMKRPAATKKAGQAKKKK

10

20

30

40

50

【0364】

BvCas12b V4（野生型に対してS893R/K846R/E837G変化）は以下のように発現される：5' mRNA Cap---5' UTR---bhCas12b---STOP配列---3' UTR --- 120polyAテール
5' UTR:

GGGAAATAAGAGAGAAAAGAAGAGTAAGAAGAAATATAAGAGCCACC

3' UTR (TriLink 標準UTR)

GCTGGAGCCTCGGTGGCCATGCTTCTTGCCCTTGGGCCTCCCCCAGCCCCTCCTCCCC
TTCCTGCACCCGTACCCCGTGGTCTTTGAATAAAGTCTGA

【0365】

bhCas12b (V4)の核酸配列

ATGGCCCCAAAGAAGAAGCGGAAGGTCGGTATCCACGGAGTCCCAGCAGCCGCCACCAG
ATCCTTCATCCTGAAGATCGAGCCCAACGAGGAAGTGAAGAAAGGCCTCTGGAAAACCC
ACGAGGTGCTGAACCACGGAATCGCCTACTACATGAATATCCTGAAGCTGATCCGGCAA
GAGGCCATCTACGAGCACCACGAGCAGGACCCCAAGAATCCCAAGAAGGTGTCCAAGGC
CGAGATCCAGGCCGAGCTGTGGGATTTCTGTGCTGAAGATGCAGAAGTGCAACAGCTTCA
CACACGAGGTGGACAAGGACGAGGTGTTCAACATCCTGAGAGAGCTGTACGAGGAACTG
GTGCCCAGCAGCGTGGAAAAGAAGGGCGAAGCCAACCAGCTGAGCAACAAGTTTCTGTGA
CCCTCTGGTGGACCCCAACAGCCAGTCTGGAAAGGGGAACAGCCAGCAGCGGCAGAAAGC
CCAGATGGTACAACCTGAAGATTGCCGGCGATCCCTCCTGGGAAGAAGAGAAGAAGAAG
TGGGAAGAAGATAAGAAAAAGGACCCGCTGGCCAAGATCCTGGGCAAGCTGGCTGAGTA
CGGACTGATCCCTCTGTTTCATCCCCTACACCGACAGCAACGAGCCCATCGTGAAAGAAA
TCAAGTGGATGGAAAAGTCCCGGAACCAGAGCGTGCGGCGGCTGGATAAGGACATGTTC
ATTCAGGCCCTGGAACGGTTCCTGAGCTGGGAGAGCTGGAACCTGAAAGTGAAAGAGGA
ATACGAGAAGGTCGAGAAAGAGTACAAGACCCTGGAAGAGAGGATCAAAGAGGACATC
CAGGCTCTGAAGGCTCTGGAACAGTATGAGAAAGAGCGGCAAGAACAGCTGCTGCGGGA
CACCTGAACACCAACGAGTACCGGCTGAGCAAGAGAGGCCTTAGAGGCTGGCGGGAAA
TCATCCAGAAATGGCTGAAAATGGACGAGAACGAGCCCTCCGAGAAGTACCTGGAAGTG
TTCAAGGACTACCAGCGGAAGCACCTAGAGAGGCCGGCGATTACAGCGTGTACGAGTT
CCTGTCCAAGAAAGAGAACCACTTTCATCTGGCGGAATCACCTGAGTACCCCTACCTGT
ACGCCACCTTCTGCGAGATCGACAAGAAAAAGAAGGACGCCAAGCAGCAGGCCACCTTC
ACACTGGCCGATCCTATCAATCACCTCTGTGGGTCCGATTTCGAGGAAAGAAGCGGCAG
CAACCTGAACAAGTACAGAATCCTGACCGAGCAGCTGCACACCGAGAAGCTGAAGAAAA
AGCTGACAGTGCAGCTGGACCGGCTGATCTACCCTACAGAATCTGGCGGCTGGGAAGAG
AAGGGCAAAGTGACATTGTGCTGCTGCCCAGCCGGCAGTTCTACAACCAGATCTTCCT
GGACATCGAGGAAAAGGGCAAGCACGCCTTACCTACAAGGATGAGAGCATCAAGTTCC
CTCTGAAGGGCACACTCGGCGGAGCCAGAGTGCAGTTCGACAGAGATCACCTGAGAAGA
TACCCTCACAAGGTGGAAAGCGGCAACGTGGGCAGAACTACTTCAACATGACCGTGAA
CATCGAGCCTACAGAGTCCCCAGTGTCCAAGTCTCTGAAGATCCACCGGGACGACTTCC
CCAAGGTGGTCAACTTCAAGCCCAAGAACTGACCGAGTGGATCAAGGACAGCAAGGGC
AAGAACTGAAGTCCGGCATCGAGTCCCTGGAAATCGGCCTGAGAGTGATGAGCATCGA
CCTGGGACAGAGACAGGCCGCTGCCGCCTCTATTTTCGAGGTGGTGGATCAGAAGCCCG
ACATCGAAGGCAAGCTGTTTTTCCCAATCAAGGGCACCGAGCTGTATGCCGTGCACAGA
GCCAGCTTCAACATCAAGCTGCCCAGGCGAGACACTGGTCAAGAGCAGAGAAGTGCTGCG
GAAGGCCAGAGAGGACAATCTGAACTGATGAACCAGAAGCTCAACTTCTGCGGAACG
TGCTGCACTTCCAGCAGTTCGAGGACATCACCGAGAGAGAGAAGCGGGTCACCAAGTGG
ATCAGCAGACAAGAGAACAGCGACGTGCCCTGGTGTACCAGGATGAGCTGATCCAGAT
CCGCGAGCTGATGTACAAGCCTTACAAGGACTGGGTGCGCTTCTGAAGCAGCTCCACA
AGAGACTGGAAGTCGAGATCGGCAAGAAAGTGAAGCACTGGCGGAAGTCCCTGAGCGAC
GGAAGAAAGGGCCTGTACGGCATCTCCCTGAAGAACATCGACGAGATCGATCGGACCCG
GAAGTTCCTGCTGAGATGGTCCCTGAGGCCTACCGAACCTGGCGAAGTGCGTAGACTGG

10

20

30

40

50

AACCCGGCCAGAGATTGCGCATCGACCAGCTGAATCACCTGAACGCCCTGAAAGAAGAT
CGGCTGAAGAAGATGGCCAACACCATCATCATGCACGCCCTGGGCTACTGCTACGACGT
GCGGAAGAAGAAATGGCAGGCTAAGAACCCCGCCTGCCAGATCATCCTGTTTCGAGGATC
TGAGCAACTACAACCCCTACGAGGAAAGGTCCCGCTTCGAGAACAGCAAGCTCATGAAG
TGGTCCAGACGCGAGATCCCCAGACAGGTTGCACTGCAGGGCGAGATCTATGGCCTGCA
AGTGGGAGAAGTGGGCGCTCAGTTCAGCAGCAGATTCCACGCCAAGACAGGCAGCCCTG
GCATCAGATGTAGCGTCGTGACCAAAGAGAAGCTGCAGGACAATCGGTTCTTCAAGAAT
CTGCAGAGAGAGGGGCAGACTGACCCTGGACAAAATCGCCGTGCTGAAAGAGGGCGATCT
GTACCCAGACAAAGGCGGCGAGAAGTTCATCAGCCTGAGCAAGGATCGGAAGTGCCTGA
CCACACACGCCGACATCAACGCCGCTCAGAACCTGCAGAAGCGGTTCTGGACAAGAACC
CACGGCTTCTACAAGGTGTACTGCAAGGCCTACCAGGTGGACGGCCAGACCGTGTACAT
CCCTGAGAGCAAGGACCAGAAGCAGAAGATCATCGAAGAGTTCGGCGAGGGGCTACTTCA
TTCTGAAGGACGGGGTGTACGAATGGGTCAACGCCGGAAGCTGAAAATCAAGAAGGGC
AGCTCCAAGCAGAGCAGCAGCGAGCTGGTGGATAGCGACATCCTGAAAGACAGCTTCGA
CCTGGCCTCCGAGCTGAAAGGCGAAAAGCTGATGCTGTACAGGGACCCCAGCGGCAATG
TGTTCCCCAGCGACAAATGGATGGCCGCTGGCGTGTTCTTCGGAAAGCTGGAACGCATC
CTGATCAGCAAGCTGACCAACCAGTACTCCATCAGCACCATCGAGGACGACAGCAGCAA
GCAGTCTATGAAAAGGCCGCGGCCACGAAAAGGCCGGCCAGGCAAAAAAGAAAAAG
【0366】

10

いくつかの実施形態では、Cas12bはBvCas12Bであり、これはBhCas12bのバリエーションであってBhCas12Bに対して以下の変化を含む：S893R、K846R、およびE837G。
BvCas12b (Bacillus sp. V3-13) NCBI Reference Sequence: WP_101661451.1
MAIRSIKLMKMTNSGTDISIYLRKALWRTHQLINEGIAYYMNLLTYRQEAIGDKTKEAYQ
AELINIIRNQQRNNGSSEEHGSDQEILALLRQLYELIIPSSIGESGDANQLGNKFLYPLVDP
NSQSGKGTSNAGRKPRWKRLKEEGNPDWELEKKKDEERKAKDPTVKIFDNLNKYGLLPL
FPLFTNIQKDIEWLPLGKRQSVRKWDKDMFIQAIERLLSWESWNRRVADEYKQLKEKTE
SYYKEHLTGGEEWIEKIRKFEKERNMELEKNAFAPNDGYFITSRQIRGWDRVYEKWSKLP
ESASPEELWKVVAEQQNKMSSEFGDPKVFSFLANRENARDIWRGHSEIRYHIAAYNGLQK
KLSRTKEQATFTLPDAIEHPLWIRYESPGGTNLNLFKLEEKQKKNYYVTLSKIIWPSEK
WIEKENIEIPLPSIQFNRQIKLKQHVKGKQEISFSDYSSRISLDGVLGGSRIQFNRKYIKN
HKELLGEGDIGPVFFNLVVDVAPLQETRNGRLQSPIGKALKVISSDFSKVIDYKPKELMD
WMNTGSASNSFGVASLLEGMRVMSIDMGQRTSASVSIFEVVKELPKDQEQKLFYSINDTE
LFAIHKRSFLLNLPGEVVTKNNKQQRQERRKKRQFVRSQIRMLANVLRLETKKTPDERKK
AIHKLMEIVQSYDSWTASQKEVWEKELNLLTNMAAFNDEIWKESLVELHHRIEYPVGGI
VSKWRKGLSEGRKNLAGISMWNIDELEDTRLLISWSKRSRTPGEANRIETDEPFGSSLL
QHIQNVKDDRLKQMANLIIMTALGFKYDKEEKDRYKRWKETYPACQIILFENLNRYLFNL
DRSRRENSRLMKWAHRSIPRTVSMQGEFGLQVGDRSEYSSRFHAKTGAPGIRCHALT
EEDLKAGSNTLKRLEDGFINESLAYLKKGDIIPSQGGELFVTLISKRYKKDSDNNELTVIH
ADINAAQNLQKRFWQQNSEVYRVPCQLARMGEDKLYIPKSQTETIKKYFGKGSFVKNNT
EQEVYKWEKSEKMKIKTDTTFDLQDLDFEDISKTIELAQQKKYLTMFRDPSGYFFNN
ETWRPQKEYWSIVNNIIKSCLKKKILSNKVEL
【0367】

20

30

40

ある実施形態では、Cas12bはBTCas12b。BTCas12b (Bacillus thermoamylovorans) NCBI Reference Sequence: WP_041902512である。
MATRSFILKIEPNEEVKKGLWKTHEVLNHGIAYYMNILKLIRQEAIYEHHEQDPKNPKKV
SKAEIQAELWDFVLKMQKCNSFTHEVDKDVVFNILRELYEELVPSSVEKKGEANQLSNKF
LYPLVDPNSQSGKGTASSGRKPRWYNLKIAGDPSWEEKKKKWEEDKKKDPLAKILGKLA
YGLIPLFIPFTDSNEPIVKEIKWMEKSRNQSVRRLDKDMFIQALERFLSWESWNLKVKEE
YEKVEKEHKTLEERIKEDIQAFKSLEQYEKERQEQLLRDTLNTNEYRLSKRGLRGWREII
QKWLKMDENEPSEKYLEVFKDYQRKHPREAGDYSVYEFLSKKENHFIWRNHPEYPYLYA

50

T

FCEIDKKKKDAKQQATFTLADPINHPLWVRFEERSGSNLNKYRILTEQLHTEKLKKKLTV
QLDRLIYPTESGGWEEKGKVDIVLLPSRQFYNQIFLDIEEKGKHAFTYKDESIKFPKGT
LGGARVQFDRDHLRRYPHKVESGNVGRIFYNMTVNIEPTESPVSKSLKIHRDDFPKFVNF
KPKELTEWIKDSKGKKLKSGIESLEIGLRVMSIDLGQRQAAAASIFEVVDQKPDIEGKLF
FPIKGTELYAVHRASFNIKLPGETLVKSREVLARKAREDNLKL MNQKLNFLRNVLFHFQQFE
DITEREKRVTKWISRQENS DVPLVYQDELIQIRELMYKPYKDWVAFLEKQLHHRLEVEIGK
EVKHWKSLSDGRKGLYGISLKNIDEIDRTRKFLLRWSLRPTEPGEVRRLEPGQRFAIDQ
LNHLNALKEDRLKKMANTIIMHALGYCYDVRKKKWQAKNPACQIILFEDLSNYPYEERS
RFENSKLMKWSRREIPRQVALQGEIYGLQVGEVGAQFSSRFHAKTGSPGIRCSVVTKEKL
QDNRRFFKNLQREGRLTLDKIAVLKEGDLYPDKGGEKFISLSKDRKLVTTHADINAAQNLQ
KRFWTRTHGIFYKYCKAYQVDGQTVYIPESKDQKQKIIEFGEYFILKDGVEYEWGNAGK
LKIKKGSSKQSSSELVDS DILKDSFDLASELKGEKLMLYRDP SGNVFSPDKWMAAGVFFG
KLERILISKLTNQYSISTIEDDSSKQSM

【0368】

いくつかの実施形態では、napDNAbpはCas12cを指す。いくつかの実施形態において、Cas12cタンパク質は、Cas12c1またはCas12c1のバリエーションである。いくつかの実施形態において、Cas12タンパク質は、Cas12c2またはCas12c2のバリエーションである。いくつかの実施形態において、Cas12タンパク質は、*Oleiphilus* sp. HI0009由来のCas12c (すなわちOspCas12c) またはOspCas12cのバリエーションである。これらのCas12c分子は、Yan et al., “Functionally Diverse Type V CRISPR-Cas Systems,” *Science*, 2019 Jan. 4; 363: 88-91に記載されており、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、napDNAbpは、天然に存在するCas12c1、Cas12c2、またはOspCas12c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Cas12c1、Cas12c2、またはOspCas12cタンパク質である。いくつかの実施形態において、napDNAbpは、本明細書に記載の任意のCas12c1、Cas12c2、またはOspCas12c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Cas12c1、Cas12c2、またはOspCas12cもまた、本開示に従って使用され得ることが理解されるべきである。

【0369】

Cas12c1 MQTKKTHLHLISAKASRKYRRTIACLSDTAKKDLERRKQSGAADPAQELSCCLK
TIKFKLEVPEGSKLPSFDRISQIYNALETIEKGSLSYLLFALILSGFRIFPNSSAAKTFASSSC
YKNDQFASQIKEIFGEMVKNFIPSELESILKKGRRKNNKDWTEENIKRVLNSEFGRKNSEG
SSALFDSFLSKFSQELFRKFD SWNEVNKKYLEAAELLD SMLASYGPFDSVCKMIGDSDSR
NSLPDKSTIAFTNNAEITVDIESSVMPYMAIAALLREYRQSKSKAAPVAVYVQSHLT TTNG
NGLSWFFKFGLDLIRKAPVSSKQSTSDGSKSLQELFSVPDDKLDGLKFIKEACEALPEASL
LCGEKGELLGYQDFRTSFAGHIDSWVANYVNRLFELIELVNQLPESIKLPSILTQKNHNLV
ASLGLQEAEVSHSLELFEGLVKNVRQT LKKLAGIDISSPNEQDIKEFYAFSDVLNRLGSIR
NQIENAVQTAKKDKIDLES AIEWKEWKLLKKLPKL NGLGGGV PKQ QELLDKALESVKQIR
HYQRIDFERVIQWAVNEHCLETVPKFLVDAEKKKINKESSTDFAAKENAVRFLLEGIGAA
ARGKTDSVSKAAYNWFV VNNFLAKKDLNRYFINCQGC IYKPPYSKRRSLAFALRSDNKDT
IEVVWEKFETFYKEISKEIEKFNFISQEFQTFLHLENLRMKLLLRRIQKPIPAEIAFFSLPQE
YYDSLPPNVAFLALNQEITPSEYITQFNLYSSFLNGNLILLRRSRSYLRAKFSWVGNSKLIY
AAKEARLWKIPNAYWKSDEWK MILDSNVLVFDKAGNVLPAPT LKKVCEREGDLRLFYPL
LRQLPHDWCYRNP FVKSVGREKNVIEVNKEGEPKVASALPGSLFRLIGPAPFKSLDDCF

FNPLDKDLRECMLIVDQEISQKVEAQKVEASLESC TY SIAVPIRYHLEEPKVS NQFENVLA
IDQGEAGLAYAVFSLKSIGEAETKPIAVGTIRIPSIRRLIHSVSTYRKKKQRLQNFKQNYDS
TAFIMRENV TGDVCAKIVGLMKEFNAPVLEYDVKNLESGSRQLSAVYKAVNSHFLYFKE
PGRDALRKQLWYGGDSWTIDGIEIVTRERKEDGKEGVEKIVPLKVFPGRSVSARFTSKTCS
CCGRNVFDWLFTEKKAKTNKKFNVNSKGELTTADGVIQLFEADRSKGPKFYARRKERTPL
TKPIAKGSYSLEEIERRVRTNLRAPKSKQSRDTSQSQYFCVYKDCALHFSGMQADENAAI
NIGRRFLTALRKNRRSDFPSNVKISDRLLDN

[0 3 7 0]

Cas12c2 MTKHSIPLHAFRNSGADARKWKGR IALLAKRGKETMRTLQFPLEMSEPEAAA I
NTTPFAVAYNAIEGTGKGT LFDYWAKLHLAGFRFFPSGGAATIFRQQAVFEDASWNAAF
CQQSGKDW P WLVPSKLYERFTKAPREVAKKDGSKK SIEFTQENVANESHVSLVGASITDK
TPEDQKEFFLKMAGALAEKFDSWKSANEDRIVAMKVIDEFLKSEGLHLPSLENI AVKCSV
ETKPDNATVAWHDAPMSGVQNLAIGVFATCASRIDNIYDLNGGKLSKLIQESATTPNVTA
LSWLFGKGLE YFR TTDIDTIMQDFNIPASAKESIKPLVES AQAIPTMTVLGKKNYAPFRPN
FGGKIDSWIANYASRLMLLNDILEQIEPGFELPQALLDNETLMSGIDMTGDELKELIEAVY
AWVDAAKQGLATLLGRGGNVDDAVQTFEQFSAMMDTLNGTLNTISARYVRVEMAGKD
EARLEK LIECKFDIPKWCKSVPKLVGISGGLPKVEEEIKVMNAAFKDVRARMFVRFE EIAA
YVASKGAGMDVYDALEKRELEQIKKLKSAVPERAHIQAYRAVLHRIGRAVQNCSEKTKQL
FSSKVIEMGVFKNP SHLNNFIFNQKGAIYRSPFDRSRHAPYQLHADKLLKNDWLELLAEI
SATLMASESTEQMEDALRLERTRLQLQLSGLPDWEYPASLAKPDIEVEIQ TALKMQLAKD
TVTSDVLQRAFNL YSSVLSGLTFKLLRRSFS LKMRF SVADTTQLIYVPK VCDWAIPKQYL
QAEGEIGIAARVVTESPAKMVTEVEMKEPKALGHFMQQAPHDWYFDASLGGTQVAGRI
VEKGKEVGKERKLVGYRMRGNSAYKTVL DKS LVGNTELSQC SMII EIPYTQTVDADFRAQ
VQAGLPKVSINLPVKETITASNKDEQMLFDRFVAIDLGERGLGYAVFDAKTLELQESGHR
PIKAITNLLNRTHHYEQRPNQ RQKFQAKFNVNLSELRENTVGDVCHQINRICAYYNAFPV
LEYMVPDRLDKQLKS VYESVTNRYIWSSTDAHKSARVQFWLGGETWEHPYLKSAKDKK
PLVLSPGRGASGKGT SQTCS CCGRNPFDLIKDMKPRAKIAVVDGKAKLENS ELKLFERNL
ESKDDMLARRHRNERAGMEQPLTPGNYTVDEIKALLRANLRRAPKNRRRTKDTTVSEYHC
VFSDCGKTMHADENAAVNIGGKFIADIEK

[0 3 7 1]

OspCas12c

MTKLRHRQKKLTHDWAGSKKREVLGSNGKLQNPLLMPVKKGQVTEFRKA FSAYARATK
GEMTDGRKNMFTHSFEPFKTKPSLHQCE LADKAYQSLHSYLPGLSLAHFLLSAHALGFRIF
SKSGEATAFQASSKIEAYESKLASELACVDLSIQNL TISTLFNALTTSVRGKGEETSADPLIA
RFYTLLTGKPLSRDTQGPERDLAEVISRKIASSFGTWKEMTANPLQSLQFFEEELHALDA
NVSLSPA FDLVIK MNDLQGD LKNRTIVFDPDAPVFEYNAEDPADIIIKLTARYAKEAVIKN
QNVGNYVKNAITTTNANGLGWLLNKGLSLLPVSTDDELLEFIGVERSHP SCHALIELIAQL
EAPELFEKNVFS DTRSEVQGMIDSAVSNHIARLSSSRNSLSMDSEELERLIKSFQIHTPHCS
LFIGAQSLSQQLES LPEALQSGVNSADILLGSTQYMLTNSLVEESIATYQRTLNRINYLSGV
AGQINGAIKRKAIDG EKIHLPAAWSELISLPFIGQPVIDVESDLAHLKNQYQTL SNEFDTLI
SALQKNFDLNFNKALLNRTQHFEAMCRSTKKNALSKPEIVSYRDLLARLT SCLYRGS LVL
RRAGIEVLKKHKIFESNSELREHVHERKH FVFSPLDRKAKKLLRLTDSRPDLLHVIDEIL
QHDNLENKDRESLWLVRSGYLLAGLPDQLSSSFINLPITQKGDRRLIDLIQYDQINRDAF
VMLVTSAFKSNLSGLQYRANKQS FVVTRTLSPYLGSKLVYVPKDKDWLVPSQMFEGRFA
DILQSDYMWVKDAGRLCVIDTAKHLSNIKKSVFSSEEVLAF LRELPHRTFIQTEVRGLGV
NVDGIAFNNGDIPSLKTF SNCVQVKVSRTNTSLVQTLNRWFEGGKVSPPSIQFERAYYKK
DDQIHEDA AKRKIRFQMPATELVHASDDAGWTPSYLLGIDPGEYGMGLSLVSINNGEVL
DSGFIHINSLINFASKKSNHQTKVVP RQQYKSPYANYLEQSKDSAAGDIAHILDRLIYKLN
ALPVFEALSGNSQSAADQVWTKVLSFYTWGDND AQNSIRKQHWFGASHWDIKGMLRQP
PTEKKPKPYIAFPGSQVSSYGNSQRCSCCGRNPIEQ LREMAKDTSIKELKIRNSEIQ LFDGT

IKLFNPDPSTVIERRRHNLGPSRIPVADRTFKNISPSSLEFKELITIVSR SIRHSPEFIAKKRG
IGSEYFCAYSDCNSSLNSEANAAANVAQKFQKQLFFEL

【0372】

いくつかの実施形態において、napDNAbpは、Cas12g、Cas12h、またはCas12iを指し、これらは、例えばYan et al., “Functionally Diverse Type V CRISPR-Cas Systems,” Science, 2019 Jan. 4; 363: 88-91に記載されており、その内容全体は、参照により本明細書に組み込まれる。10テラバイトを超える配列データを集約することにより、Cas12g、Cas12h、Cas12iを含め以前に特徴付けられたクラスVタンパク質との弱い類似性を示すタイプV Casタンパク質の新しいクラスが同定された。いくつかの実施形態において、Cas12タンパク質は、Cas12gまたはそのバリエーションである。いくつかの実施形態において、Cas12タンパク質は、Cas12hまたはそのバリエーションである。いくつかの実施形態において、Cas12タンパク質は、Cas12iまたはそのバリエーションである。他のRNA誘導DNA結合タンパク質がnapDNAbpとして使用され得、そして本開示の範囲内にあることを理解されたい。いくつかの実施形態において、napDNAbpは、天然に存在するCas12g、Cas12h、またはCas12i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Cas12g、Cas12h、またはCas12iタンパク質である。いくつかの実施形態において、napDNAbpは、本明細書に記載の任意のCas12g、Cas12h、またはCas12i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%の同一性であるアミノ酸配列を含む。他の細菌種由来のCas12g、Cas12h、またはCas12iもまた、本開示に従って使用され得ることが理解されるべきである。いくつかの実施形態では、Cas12iは、Cas12i1またはCas12i2である。

【0373】

Cas12g1

MAQASSTPAVSPRPRPRYREERTLVRKLLPRPGQSKQEFRENVKKLRKAFLQFNADVSGV
CQWAIQFRPRYGKPAEPTETFWKFFLEPETS LPPNDSRSPEFRRLQAFEAAAGINGAAAL
DDPAFTNELRDSILAVASRPKTKEAQRLFSRLKDYQPAHRMILAKVAAEWIESRYRRAHQ
NWERNYEEWKKEKQEWQNHPALTPEIREAFNQIFQQLEVKEKRVRICPAARLLQNKDN
CQYAGKNKHSVLCNQFNEFKKNHLQGKAIFFYKDAEKYLRCGLQSLKPNVQGPFPREDW
NKYLYRMNLKEETLRGKNGGRLPHCKNLGQCEFPNPTALCKYQQQLSSRPDLVQHD
ELYRKWRREYWREPRKPVFRYPSVKRHSIAKIFGENYFQADFKNSVVGRLRDSMPAGQY
LEFAFAPWPRNYRPQPGETEISSVHLHFVGTRPRIGFRFRVPHKRSRFDCTQEELDELRSR
TFPRKAQDQKFLEAARKRLLETFPGNAEQELRLLAVDLGTD SARAAFFIGKTFQQAFLKI
VKIEKLYEQWPNQKQAGDRRDASSKQPRPGLSRDHVGRHLQKMRAQASEIAQKRQELTG
TPAPETTTDQAAKKATLQPFDLRGLTVHTARMIRDWARLNARQIIQLAEENQVDLIVLES
LRGFRPPGYENLDQEKRRVAFFAHGRIRRVTEKAVERGMRVVTVPYLASSKVCAECCRK
KQKDNKQWEKNKKRGLFKCEGCGSQAQVDENAAARVLGRVFWGEIELPTAIP

【0374】

Cas12h1

MKVHEIPRSQLLKIKQYEGSFVEWYRDLQEDRKKFASLLFRWAAFGYAAREDDGATYISP
SQALLERRLLLGDAEDVAIKFLDVLFGGAPSSSCYSLFYEDFALRD KAKYSGAKREFIEGL
ATMPLDKIIRIRQDEQLSKIPAEEWLILGAEYSPEEIWEQVAPRIVNVDRSLGKQLRERL
GIKCRPHDAGYCKILMEVVARQLRSHNETYHEYLNQTHEMKT KVANNLTNEFDLVCEF
AEVLEEKNYGLGWYVLWQGVKQALKEQKKPTKIQIAVDQLRQPKFAGLLTAKWRALKG
AYDTWKLKKRLEKRKAFFPYMPNWDNDYQIPVGLTGLGVFTLEV KRTEVVVDLKEHGKL
FCSHSHYFGDLTAEKHPSRYHLKFRHKLKLRKRD SRVEPTIGPWIEAALREITIQKKPNGV

FYLGLPYALSHGIDNFQIAKRFFSAAKPDKEVINGLPSEMVVGAADLNLSNIVAPVKARIG
KGLEGPLHALDYGYGELIDGPKILTPDGPRCGELISLKR DIVEIKSAIKEFKACQREGLTMS
EETTTWLSEVESPSDSPRCMIQSRIADTSRRLNSFKYQMNKEGYQDLAEALRLLDAMDSY
NSLLESYQRMHLSPGEQSPKEAKFDTKRASFRDLLRRRVAHTIVEYFDDCDIVFFEDLDG
PSDSDSRNNALVKLLSPRTL LLYIRQALEKRGIGMVEVAKDGTSQNNPISGHVGVWRNKQ
NKSEIYFYEDKELLVMDADEVGAMNILCRGLNHSVCPYSFVTKAPEKKNDEKKEGDYGK
RVKRFLKDRYGSSNVRFLVASMGFVTVTTKRPKDALVGKRLYYHGGELVTHDLHNRMK
DEIKYLVEKEVLARRVSLSDSTIKSYKSFAHV

[0375]

Cas12i1

MSNKEKNASETRKAYTTKMIPRSHDRMKLLGNFMDYLMDGTPIFFELWNQFGGGIDRDI
ISGTANKDKISDDL LLA VNWFKVMPI NSKPQGVSPSNLANLFQQYSGSEPDIQAQEYFAS
NFDTEKHQWKDMRVEYERLLAELQLSRSDMHHD LKLMYKEKCIGLSLSTAHYITSVMFG
TGAKNNRQTKHQFYSKVIQLLEESTQINSVEQLASIILKAGDCDSYRKL RIRCSRKGATPSI
LKIVQDYELGTNH DDEVNVP SLIANLKEKLGRFEYECEWKCM EKIKAFLASKVGPYYLGS
YSAMLENALSPIKGMTTKNCKFVLKQIDAKNDIKYENEPFGKIVEGFFDSPYFESDTNVK
WVLHPHHIGESNIKTLWEDLNAIH SKYEEDIASLSEDKKEKRIKVYQGDVCQTINTYCEE
VGKEAKTPLVQLRLYLSRKDDIAVDKIIDGITFLSKKKHVEKQKINPVIQKYPSFNFNGN
SKLLGKIISPKDKLKHNLKCNRNQVDNYIWIEIKVLNTKTMRWEKH HYALSSTRFLEEYV
YPATSENPPDALAARFRTKTNGYEGKPALSAEQIEQIRSAPVGLRKVKKRQMRLEAARQQ
NLLPRYTWGKDFNINICKRGN NFEVTLATKVKKKKEKNYKVVLGYDANIVRKNTYAAIEA
HANGDGVIDYNDLPVKPIESGFVTVESQVRDKSYDQLSYNGVKLLYCKPHVESRRSFLEK
YRNGTMKDNRGNNIQIDFMKDFEAIADDETSLYYFNMKYCKLLQSSIRNHSSQAKEYREE
IFELLRDGKLSVLKLSSLSNLSFVMFKVAKSLIGTYFGHLLKKPKNSKSDVKAPPITDEDK
QKADPEMFALRLALEEKRLNKVKSKEVIANKIVAKALELRDKYGPVLIKGENISDTTKKG
KKSSTNSFLMDWLARGVANKVKEMVMMHQGLEFVEVNP NFTSHQDPFVHKNPENTFR
ARYSRCTPSELTEKNRKEILSFLSDKPSKRPTNAYYNEGAMAFLATYGLKKNDVLGVSL
KFKQIMANILHQRSEDQLLFPSRGGMFYLATYKLDADATSVNWNGKQFWVCNADLVAA
YNVGLVDIQKDFKKK

[0376]

Cas12i2

MSSAIKSYKSVLRPNERKNQLLKSTIQCLEDGSAFFFKMLQGLFGGITPEIVRFSTEQE KQ
QQDIALWCAVNWFRPVSQDSLTH TIASDNLVEKFEEYYGGTASDAIKQYFSASIGESYYW
NDCRQQYYDLCRELGVEVSDLTHDLEILCREKCLAVATESNQNN SIISVLFGTGEKEDRSV
KL RITKKILEAISNLKEIPKNVAPIQEIILNVAKATKETFRQVYAGNLGAPSTLEKFIADG
QKEFDLKKLQTDLKKVIRGKSKERDWCCQEELRSYVEQNTIQYDLWAWGEMFNKAHTA
LKIKSTRNYNFAKQRLEQFKEIQSLNNLLVVKLNDFFDSEFFSGEETY TICVHHLGGKDL
SKLYKAWEDDPADPENAI VVLCDDLKN NFKKEPIRNILRYIFTIRQECSAQDILAAAKYNQ
QLDRYKSQKANPSVLGNQGFTWTNAVILPEKAQRNDRPNSLDLRIWLYLKL RHPDGRW
KKHHIPFYDTRFFQEIYAAGNSPVDTCQFRTPRFGYHLPKLT DQT AIRVNKKHVKA AKTE
ARIRLAIQQGTLPVSNLKITEISATINSKGQVRIPVKFDVGRQKGT LQIGDRFCGYDQNQT
ASHAYSLWEVVKEGQYHKELGCFVRFISSGDIVSITENRG NQFDQLSYEGLAYPQYADWR
KKASKFVSLWQITKKNKKKEIVTVEAKEKFDAICKYQPRLYKFNKEYAYLLRDIVRGKSLV
ELQQIRQEIFRFIEQDCGVTRLGSLSLSTLETVKAVKGIIYSYFSTALNASKNNPISDEQRKE
FDP ELFALLEKLELIRTRKKKQKVERIANSLIQTCLENNIKFIRGEGDLSTTNNATKKKANS
RSMDWLARGVF NKIRQLAPMHNITLFGCGSLYTSHQDPLVHRNPDKAMKCRWAAIPVK
DIGDWVLRKLSQNLRAKNIGTGEYYHQGVKEFLSHYELQDLEEELLKWRSDRKSNI PCW
VLQNR LAEKLGNKEAVVYIPVRGGRIYFATHKVATGAVSIVFDQKQVWVCNADHVAAAN
IALTVKGIGEQQSSDEENPDGSR IKLQLTS

[0377]

10

20

30

40

50

塩基エディターの代表的な核酸配列およびタンパク質配列は以下の通りである。

【0378】

BhCas12b GGS GGS-ABE8-Xten20 at P153

BhCas12b GGS GGS-ABE8-Xten20 at P153

GCCACCATGGCCCCAAAGAAGAAGCGGAAGGTCGGTATCCACGGAGTCCCAGCAGCCGCCACCA
GATCCTTCATCCTGAAGATCGAGCCCAACGAGGAAGTGAAGAAAGGCCTCTGGAAAACCCACGA
GGTGTCTGAACCACGGAATCGCCTACTACATGAATATCCTGAAGCTGATCCGGCAAGAGGCCATC
TACGAGCACCACGAGCAGGACCCCAAGAATCCCAAGAAGGTGTCCAAGGCCGAGATCCAGGCCG
AGCTGTGGGATTTCTGTCTGAAGATGCAGAAGTGCAACAGCTTCACACACGAGGTGGACAAGGA
CGAGGTGTTCAACATCCTGAGAGAGCTGTACGAGGAAGTGGTGCCAGCAGCGTGGAAGAAG
GGCGAAGCCAACCAGCTGAGCAACAAGTTTCTGTACCCTCTGGTGGACCCCAACAGCCAGTCTG
GAAAGGGAACAGCCAGCAGCGGCAGAAAGCCCAGATGGTACAACCTGAAGATTGCCGGCGATCC
CggaggtctctggaggaagcTCCGAAGTCGAGTTTTCCCATGAGTACTGGATGAGACACGCATTG
ACTCTCGCAAAGAGGGCTCGAGATGAACGCGAGGTGCCCGTGGGGGCAGTACTCGTGCTCAACA
ATCGCGTAATCGGCGAAGGTTGGAATAGGGCAATCGGACTCCACGACCCCACTGCACATGCGGA
AATCATGGCCCTTCGACAGGGAGGGCTTGTGATGCAGAATTATCGACTTTATGATGCGACGCTG
TACGTACGTTTGAACCTTGCGTAATGTGCGCGGGAGCTATGATTCACTCCCGCATTGGACGAG
TTGTATTTCGGTGTTCGCAACGCCAAGACGGGTGCCGCGAGGTTCACTGATGGACGTGCTGCATCA
TCCAGGCATGAACCACCGGGTAGAAATCACAGAAGGCATATTGGCGGACGAATGTGCGGCGCTG
TTGTGTCTGTTTTTTTCGCATGCCAGGCGGGTCTTTAACGCCCAGAAAAAGCACAATCCTCTA
CTGACGGCTCTTCTGGATCTGAAACACCTGGCACAAGCGAGAGCGCCACCCCTGAGAGCTCTGG
CTCCTGGGAAGAAGAGAAGAAGAAGTGGGAAGAAGATAAGAAAAAGACCCGCTGGCCAAGATC
CTGGGCAAGCTGGCTGAGTACGGACTGATCCCTCTGTTCATCCCCTACACCGACAGCAACGAGC
CCATCGTGAAAGAAATCAAGTGGATGGAAAAGTCCCGGAACCAGAGCGTGCGGCGGCTGGATAA
GGACATGTTTCATTTCAGGCCCTGGAACGGTTCCTGAGCTGGGAGAGCTGGAACCTGAAAGTGAAA
GAGGAATACGAGAAGGTCGAGAAAAGGTACAAGACCCTGGAAGAGAGGATCAAAGAGGACATCC
AGGCTCTGAAGGCTCTGGAACAGTATGAGAAAGAGCGGCAAGAACAGCTGCTGCGGGACACCCT
GAACACCAACGAGTACCGGCTGAGCAAGAGAGGCCTTAGAGGCTGGCGGGAATCATCCAGAAA
TGGCTGAAAATGGACGAGAACGAGCCCTCCGAGAAGTACCTGGAAGTGTTCAGGACTACCAGC
GGAAGCACCCCTAGAGAGGCCGCGATTACAGCGTGTACGAGTTCCTGTCCAAGAAAGAGAACCA
CTTCATCTGGCGGAATCACCTGAGTACCCCTACCTGTACGCCACCTTCTGCGAGATCGACAAG
AAAAAGAAGGACGCCAAGCAGCAGGCCACCTTCACACTGGCCGATCCTATCAATCACCCCTCTGT
GGGTCCGATTGAGGAAAGAAGCGGCAGCAACCTGAACAAGTACAGAATCCTGACCGAGCAGCT
GCACACCGAGAAGCTGAAGAAAAAGCTGACAGTGCAGCTGGACCGGCTGATCTACCCTACAGAA
TCTGGCGGCTGGGAAGAGAAGGGCAAAGTGACATTGTGCTGCTGCCAGCCGGCAGTTCTACA
ACCAGATCTTCCTGGACATCGAGGAAAAGGGCAAGCACGCCTTCACCTACAAGGATGAGAGCAT
CAAGTTCCCTCTGAAGGGCACACTCGGCGGAGCCAGAGTGCAGTTCGACAGAGATCACCTGAGA

10

20

30

40

50

AGATACCCTCACAAGGTGGAAAGCGGCAACGTGGGCAGAATCTACTTCAACATGACCGTGAACA
TCGAGCCTACAGAGTCCCCAGTGTCCAAGTCTCTGAAGATCCACCGGGACGACTTCCCCAAGGT
GGTCAACTTCAAGCCCCAAGAACTGACCGAGTGGATCAAGGACAGCAAGGGCAAGAACTGAAG
TCCGGCATCGAGTCCCTGGAAATCGGCCTGAGAGTGATGAGCATCGACCTGGGACAGAGACAGG
CCGCTGCCGCTCTATTTTCGAGGTGGTGGATCAGAAGCCCGACATCGAAGGCAAGCTGTTTTT
CCCAATCAAGGGCACCGAGCTGTATGCCGTGCACAGAGCCAGCTTCAACATCAAGCTGCCCCGGC
GAGACACTGGTCAAGAGCAGAGAAGTGCTGCGGAAGGCCAGAGAGGACAATCTGAAACTGATGA
ACCAGAAGCTCAACTTCCCTGCGGAACGTGCTGCACTTCCAGCAGTTCGAGGACATCACCGAGAG
AGAGAAGCGGGTCACCAAGTGGATCAGCAGACAAGAGAACAGCGACGTGCCCCCTGGTGTACCAG
GATGAGCTGATCCAGATCCGCGAGCTGATGTACAAGCCTTACAAGGACTGGGTGCGCTTCCCTGA
AGCAGCTCCACAAGAGACTGGAAGTCGAGATCGGCCAAAGAAGTGAAGCACTGGCGGAAGTCCCT
GAGCGACGGAAGAAAGGGCCTGTACGGCATCTCCCTGAAGAACATCGACGAGATCGATCGGACC
CGGAAGTTCCTGCTGAGATGGTCCCTGAGGCCTACCGAACCTGGCGAAGTGCCTAGACTGGAAC
CCGGCCAGAGATTGCCCATCGACCAGCTGAATCACCTGAACGCCCTGAAAGAAGATCGGCTGAA
GAAGATGGCCAACACCATCATCATGCACGCCCTGGGCTACTGCTACGACGTGCGGAAGAAGAAA
TGGCAGGCTAAGAACCCCGCCTGCCAGATCATCCTGTTTCGAGGATCTGAGCAACTACAACCCCT
ACGAGGAAAGGTCCCGCTTCGAGAACAGCAAGCTCATGAAGTGGTCCAGACGCGAGATCCCCAG
ACAGGTTGCACTGCAGGGCGAGATCTATGGCCTGCAAGTGGGAGAAGTGGGCGCTCAGTTCAGC
AGCAGATTCCACGCCAAGACAGGCAGCCCTGGCATCAGATGTAGCGTCGTGACCAAAGAGAAGC
TGCAGGACAATCGGTTCTTCAAGAATCTGCAGAGAGAGGGCAGACTGACCCTGGACAAAATCGC
CGTGCTGAAAGAGGGCGATCTGTACCCAGACAAAGGCGGCGAGAAGTTCATCAGCCTGAGCAAG
GATCGGAAGTGCCTGACCACACACGCCGACATCAACGCCGCTCAGAACCTGCAGAAGCGGTTCT
GGACAAGAACCCACGGCTTCTACAAGGTGTACTGCAAGGCCTACCAGGTGGACGGCCAGACCG
TGTACATCCCTGAGAGCAAGGACCAGAAGCAGAAGATCATCGAAGAGTTCGGCGAGGGCTACT
TCATTCTGAAGGACGGGGTGTACGAATGGGTCAACGCCGGCAAGCTGAAAATCAAGAAGGGCAG
CTCCAAGCAGAGCAGCAGCGAGCTGGTGGATAGCGACATCCTGAAAGACAGCTTCGACCTGGCC
TCCGAGCTGAAAGGCGAAAAGCTGATGCTGTACAGGGACCCCAGCGGCAATGTGTTCCCCAGCG
ACAAATGGATGGCCGCTGGCGTGTCTTCGAAAGCTGGAACGCATCCTGATCAGCAAGCTGAC
CAACCAGTACTCCATCAGCACCATCGAGGACGACAGCAGCAAGCAGTCTATGAAAAGGCCGGCG
GCCACGAAAAAGGCCGGCCAGGCAAAAAAGAAAAAGGGATCC TACCCATACGATGTTCCAGATT
ACGCTTATCCCTACGACGTGCCTGATTATGCATACCCATATGATGTCCCCGACTATGCCTAA

10

20

30

40

MAPKKKRKVGIHGVPAATRSLFILKIEPNEEVKKGLWKTHEVLNHGIAYYMNILKLIRQE
AIYEHHEQDPKNPKKVSKAEIQAEWDFVLKMQKNSFTHEVDKDEVFNILRELYEELVP
SSVEKKGEANQLSNKFLYPLVDPNSQSGKGTASSGRKPRWYNLKIAGDPGSGGSSEVEF
SHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALR
QGGLVMQNYRLYDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLH
HPGMNHRVEITEGILADECAALLCRFFRMPRRVFNAQKKAQSSTDGSSGSETPGTSESAT
PESSGSWEEKKKKWEEDKKKDPLAKILGKLAEYGLIPLFIPYTDNEPIVKEIKWMEKSRN
QSVRRLDKDMFIQALERFLSWESWNLKVKEEYKVEKEYKTLEERIKEDIQALKALEQYE
KERQEQLLRDTLNTNEYRLSKRGLRGWREIIQKWLKMDENEPSEKYLEVFKDYQRKHPR

50

EAGDYSVYEFLSKKENHFIWRNHPEYPYLYATFCEIDKKKKDAKQQATFTLADPINHPLW
VRFEERSGSNLNKYRILTEQLHTEKLKKKLTVQLDRLIYPTESGGWEEKGKVDIVLLPSRQ
FYNQIFLDIEEKGKHAFTYKDESIKFPLKGTGGARVQFDRDHLRRYPHKVESGNVGRIFY
NMTVNIEPTESPVSKSLKIHRDDFPKVVNFKPKELTEWIKDSKGKKLKSGIESLEIGLRVM
SIDLGQRQAAAASIFEVVDQKPDIEGKLFFPIKGTELYAVHRASFNIKLPGETLVKSREVL
KAREDNLKLMNQKLNFLRNVLHFQQFEDITEREKRVTKWISRQENS DVPLVYQDELIQIR
ELMYKPYKDWVAFKQLHHRLEVEIGKEVKHWRKSLSDGRKGLYGISLKNIDEIDRTRKF
LLRWSLRPTEPGEVRRLEPGQRFAIDQLNHLNALKEDRLKKMANTIIMHALGYCYDVRK
KKWQAKNPACQIILFEDLSNYPYEERSRFENSKLMKWSRREIPRQVALQGEIYGLQVGE
VGAQFSSRFHAKTGSPGIRCSVVTKEKLQDNRRFFKNLQREGRLTDKIAVLKEGDLYPDK
GGEKFISLSKDRKCVTTHADINAAQNLQKRFWTRTHGFYKVYCKAYQVDGQTVYIPESK
DQKQKIIIEFGEFYFILKDGVEWVNAGKLIKKGSSKQSSSELVSDILKDSFDLASELKG
EKLMLYRDPGSGNVFPSDKWMAAGVFFGKLERILISKLTNQYSISTIEDDSSKQSMKRPA
TKKAGQAKKKKGSYPYDVPDYAYPYDVPDYAYPYDVPDYA
[0 3 7 9]

BhCas12b GSGGS-ABE8-Xten20 at K255

10

20

30

40

50

BhCas12b GGSIGGS-ABE8-Xten20 at K255

GCCACCATGGCCCCAAAGAAGAAGCGGAAGGTCCGTATCCACGGAGTCCCAGCAGCCGCCACCA
GATCCTTCATCCTGAAGATCGAGCCCAACGAGGAAGTGAAGAAAGGCCTCTGGAAAACCCACGA
GGTGTGAACACCGGAATCGCCTACTACATGAATATCCTGAAGCTGATCCGGCAAGAGGCCATC
TACGAGCACCACGAGCAGGACCCCAAGAATCCCAAGAAGGTGTCCAAGGCCGAGATCCAGGCCG
AGCTGTGGGATTTCTGTCTGAAGATGCAGAAGTGAACAGCTTCACACACGAGGTGGACAAGGA
CGAGGTGTTCAACATCCTGAGAGAGCTGTACGAGGAAGTGGTGCCACGAGCGTGGAAGAAG
GGCGAAGCCCAACCAGCTGAGCAACAAGTTTCTGTACCCTCTGGTGGACCCCAACAGCCAGTCTG
GAAAGGGAACAGCCAGCAGCGGCAGAAAGCCCAGATGGTACAACCTGAAGATTGCCGGCGATCC
CTCCTGGGAAGAAGAGAAGAAGAAGTGGGAAGAAGATAAGAAAAAGGACCCGCTGGCCAAGATC
CTGGGCAAGCTGGCTGAGTACGGACTGATCCCTCTGTTTCATCCCCTACACCGACAGCAACGAGC
CCATCGTGAAAGAAATCAAGTGGATGGAAGAGTCCCGGAACCAGAGCGTGCGGCGGCTGGATAA
GGACATGTTTATTAGGCCCTGGAACGGTTTCTGAGCTGGGAGAGCTGGAACCTGAAAGTGAAA
GAGGAATACGAGAAGGTGAGAAAGAGTACAAGACCCTGGAAGAGAGGATCAAAaggaggctctg
gagggaagcTCCGAAGTTCGAGTTTTCCCATGAGTACTGGATGAGACACGCATTGACTCTCGCAAA
GAGGGCTCGAGATGAACGCGAGGTGCCCGTGGGGCAGTACTCGTGCTCAACAATCGCGTAATC
GGCGAAGGTTGGAATAGGGCAATCGGACTCCACGACCCCACTGCACATGCGGAAATCATGGCCC
TTCGACAGGAGGGCTTGTGATGCAGAATTATCGACTTTATGATGCGACGCTGTACGTCACGTT
TGAACCTTGCGTAATGTGCGCGGGAGCTATGATTCACTCCCGCATTTGGACGAGTTGTATTCTGGT
GTTTCGCAACGCCAAGACGGGTGCCGCAGGTTCACTGATGGACGCTGCTGCATCATCCAGGCATGA
ACCACCGGTTAGAAATCACAGAAGGCATATTGGCGGACGAATGTGCGGCGCTGTGTGTGCTTT
TTTTCGATGCCCAGGCGGGTCTTTAACGCCCAGAAAAAGCACATCCTCTACTGACGGCTCT
TCTGGATCTGAAACACCTGGCACAAGCGAGAGCGCCACCCCTGAGAGCTCTGGCGAGGACATCC
AGGCTCTGAAGGCTCTGGAACAGTATGAGAAAGAGCGGCAAGAACAGCTGCTGCGGGACACCCT
GAACACCAACGAGTACCGGCTGAGCAAGAGAGGCCCTTAGAGGCTGGCGGGAAATCATCCAGAAA
TGGCTGAAAATGGACGAGAACGAGCCCTCCGAGAAGTACCTGGAAGTGTTCAGGACTACACAGC
GGAAGCACCCCTAGAGAGGCCGCGGATTACAGCGTGTACGAGTTCCTGTCCAAGAAAGAGAACCA
CTTCATCTGGCGGAATCACCCCTGAGTACCCCTACCTGTACGCCACCTTCTGCGAGATCGACAAG
AAAAAGAAGGACGCCAAGCAGCAGGCCACCTTCACACTGGCCGATCCTATCAATCACCCCTCTGT
GGGTCCGATTTCGAGGAAAGAAGCGGCAGCAACCTGAACAAGTACAGAATCCTGACCGAGCAGCT
GCACACCGAGAAGCTGAAGAAAAAGCTGACAGTGCAGCTGGACCGGCTGATCTACCCTACAGAA
TCTGGCGGCTGGGAAGAGAAGGGCAAAGTGGACATTGTGTGCTGCTGCCAGCCGGCAGTTCTACA
ACCAGATCTTCCTGGACATCGAGGAAAAGGGCAAGCACGCCTTCACCTACAAGGATGAGAGCAT
CAAGTTCCCTCTGAAGGGCACACTCGGCGGAGCCAGAGTGCAGTTCGACAGAGATCACCTGAGA

10

20

30

40

50

AGATACCCCTCACAAGGTGGAAAGCGGCAACGTGGGCAGAATCTACTTCAACATGACCGTGAACA
TCGAGCCTACAGAGTCCCCAGTGTCCAAGTCTCTGAAGATCCACCGGGACGACTTCCCCAAGGT
GGTCAACTTCAAGCCCCAAGAACTGACCGAGTGGATCAAGGACAGCAAGGGCAAGAACTGAAG
TCCGGCATCGAGTCCCTGGAAATCGGCCTGAGAGTGATGAGCATCGACCTGGGACAGAGACAGG
CCGCTGCCGCCTCTATTTTCGAGGTGGTGGATCAGAAGCCCCGACATCGAAGGCAAGCTGTTTTT
CCCAATCAAGGGCACCGAGCTGTATGCCGTGCACAGAGCCAGCTTCAACATCAAGCTGCCCCGC
GAGACACTGGTCAAGAGCAGAGAAGTGCTGCGGAAGGCCAGAGAGGACAATCTGAAACTGATGA
ACCAGAAGCTCAACTTCCTGCGGAACGTGCTGCACTTCCAGCAGTTCGAGGACATCACCGAGAG
AGAGAAGCGGGTCACCAAGTGGATCAGCAGACAAGAGAACAGCGACGTGCCCCCTGGTGTACCAG
GATGAGCTGATCCAGATCCGCGAGCTGATGTACAAGCCTTACAAGGACTGGGTGCGCTTCTCTGA
AGCAGCTCCACAAGAGACTGGAAGTCGAGATCGGCAAAGAAGTGAAGCACTGGCGGAAGTCCCT
GAGCGACGGAAGAAAGGGCCTGTACGGCATCTCCCTGAAGAACATCGACGAGATCGATCGGACC
CGGAAGTTCCTGCTGAGATGGTCCCTGAGGCCTACCGAACCTGGCGAAGTGCGTAGACTGGAAC
CCGGCCAGAGATTGCCCATCGACCAGCTGAATCACCTGAACGCCCTGAAAGAAGATCGGCTGAA
GAAGATGGCCAACACCATCATCATGCACGCCCTGGGCTACTGCTACGACGTGCGGAAGAAGAAA
TGCGAGGCTAAGAACCCCGCCTGCCAGATCATCTGTTTCGAGGATCTGAGCAACTACAACCCCT
ACGAGGAAAGGTCCCGCTTCGAGAACAGCAAGCTCATGAAGTGGTCCAGACGCGAGATCCCCAG
ACAGGTTGCACTGCAGGGCGAGATCTATGGCCTGCAAGTGGGAGAAGTGGGCGCTCAGTTCAGC
AGCAGATTCCACGCCAAGACAGGCAGCCCTGGCATCAGATGTAGCGTCGTGACCAAAGAGAAGC
TGCAGGACAATCGGTTCTTCAAGAATCTGCAGAGAGAGGGCAGACTGACCCTGGACAAAATCGC
CGTGCTGAAAGAGGGCGATCTGTACCCAGACAAAGGCGGCGAGAAGTTCATCAGCCTGAGCAAG
GATCGGAAGTGCGTGACCACACACGCCGACATCAACGCCGCTCAGAACCTGCAGAAGCGGTTCT
GGACAAGAACCACGGCTTCTACAAGGTGTACTGCAAGGCCTACCAGGTGGACGGCCAGACCGT
GTACATCCCTGAGAGCAAGGACCAGAAGCAGAAGATCATCGAAGAGTTCGGCGAGGGCTACTTC
ATTCTGAAGGACGGGTGTACGAATGGGTCAACGCCGCAAGCTGAAAATCAAGAAGGGCAGCT
CCAAGCAGAGCAGCAGCGAGCTGGTGGATAGCGACATCCTGAAAGACAGCTTCGACCTGGCCTC
CGAGCTGAAAGGCGAAAAGCTGATGCTGTACAGGGACCCCAGCGGCAATGTGTTCCCCAGCGAC
AAATGGATGGCCGCTGGCGTGTTCTTCGAAAGCTGGAACGCATCCTGATCAGCAAGCTGACCA
ACCAGTACTCCATCAGCACCATCGAGGACGACAGCAGCAAGCAGTCTATGAAAAGGCCGGCGGC
CACGAAAAAGGCCGGCCAGGCAAAAAAGAAAAAGGGATCC TACCCATACGATGTTCCAGATTAC
GCTTATCCCTACGACGTGCCTGATTATGCATACCCATATGATGTCCCCGACTATGCCTAA

10

20

30

40

MAPKKKRKVGIHGVPAATRSLKIEPNEEVKKGLWKTHEVLNHGIAYYMNLKLIRQE
AIYEHHEQDPKNPKKVSKEIQAEIQLWDFVLKMQKNSFTHEVDKDEVFNILRELYEELVP
SSVEKKGEANQLSNKFLYPLVDPNSQSGKGTASSGRKPRWYNLKIAGDPSWEEEEKKKWE
EDKKKDPLAKILGKLAEYGLIPLFIPYTDSEPIVKEIKWMEKSRNQSVRRLDKDMFIQAL
ERFLSWESWNLKVKEEYKVEKEYKTLERIKGGSGGSSEVEFSHEYWMRHALTLAKRA
RDEREVPVGAVLVLNRRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLYDRTL
YVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHHPGMNHRVEITEGILAD
ECAALLCRFFRMPRRVFNAQKKAQSSTDGSSGSETPGTSESATPESSGEDIQALKALEQYE
KERQEQLLRDTLNTNEYRLSKRGLRGWREIIQKWLKMDENEPSEKYLEVFKDYQRKHPR

50

EAGDYSVYEFLSKKENHFIWRNHPEYPYLYATFCEIDKKKKDAKQQATFTLADPINHPLW
VRFEERSGSNLNKYRILTEQLHTEKLKKKLTVQLDRLIYPTESGGWEEKGKVDIVLLPSRQ
FYNQIFLDIEEKGKHAFTYKDESIKFPLKGTGGARVQFDRDHLRRYPHKVESGNVGRIFY
NMTVNIEPTESPVSKSLKIHRDDFPKVVNFKPKELTEWIKDSKGKKLKSGIESLEIGLRVM
SIDLGQRQAAAASIFEVVDQKPDIEGKLFFPIKGTELYAVHRASFNIKLPGETLVKSREVL
KAREDNLKLMNQKLNFLRNVLHFQQFEDITEREKRVTKWISRQENS DVPLVYQDELIQIR
ELMYKPYKDWVAFKQLHHRLEVEIGKEVKHWRKSLSDGRKGLYGISLKNIDEIDRTRKF
LLRWSLRPTEPGEVRRLEPGQRFAIDQLNHLNALKEDRLKKMANTIIMHALGYCYDVRK
KKWQAKNPACQIILFEDLSNYPYEERSRFENSKLMKWSRREIPRQVALQGEIYGLQVGE
VGAQFSSRFHAKTGSPGIRCSVVTKEKLQDNRRFFKNLQREGRLTDKIAVLKEGDLYPDK
GGEKFISLSKDRKCVTTHADINAAQNLQKRFWTRTHGFYKVYCKAYQVDGQTVYIPESK
DQKQKIIIEFGEGYFILKDGVEWVNAGKLIKKGSSKQSSSELVSDILKDSFDLASELKG
EKLMLYRDPSGNVFP SDKWMAAGVFFGKLERILISKLTNQYSISTIEDDSSKQSMKRPA
TKKAGQAKKKKGSYPYDVPDYAYPYDVPDYAYPYDVPDYA
[0 3 8 0]

BhCas12b GSGGS-ABE8-Xten20 at D306

10

20

30

40

50

BhCas12b GSGGS-ABE8-Xten20 at D306

GCCACCATGGCCCCAAAGAAGAAGCGGAAGGTCGGTATCCACGGAGTCCCAGCAGCCGCCACCA
GATCCTTCATCCTGAAGATCGAGCCCCAACGAGGAAGTGAAGAAAGGCCTCTGGAAAACCCACGA
GGTGCTGAACCACGGAATCGCCTACTACATGAATATCCTGAAGCTGATCCGGCAAGAGGCCATC
TACGAGCACCACGAGCAGGACCCCCAAGAATCCCAAGAAGGTGTCCAAGGCCGAGATCCAGGCCG
AGCTGTGGGATTTCTGTGCTGAAGATGCAGAAGTGCAACAGCTTCACACACGAGGTGGACAAGGA
CGAGGTGTTCAACATCCTGAGAGAGCTGTACGAGGAAGTGGTGCCAGCAGCGTGAAAAAGAAG
GGCGAAGCCAACCAGCTGAGCAACAAGTTTCTGTACCCTCTGGTGACCCCCAACGCCAGTCTG
GAAAGGGAACAGCCAGCAGCGGCAGAAAGCCCAGATGGTACAACCTGAAGATTGCCGGCGATCC
CTCCTGGGAAGAAGAGAAGAAGAAGTGGGAAGAAGATAAGAAAAAGGACCCGCTGGCCAAGATC
CTGGGCAAGCTGGCTGAGTACGGACTGATCCCTCTGTTTCATCCCCTACACCGACAGCAACGAGC
CCATCGTGAAAGAAATCAAGTGGATGGAAAAGTCCCGGAACCAGAGCGTGCGGCGGCTGGATAA
GGACATGTTTCATTTCAGGCCCTGGAACGGTTCCTGAGCTGGGAGAGCTGGAACCTGAAAGTGAAA
GAGGAATACGAGAAGGTCGAGAAAGAGTACAAGACCCTGGAAGAGAGGATCAAAGAGGACATCC
AGGCTCTGAAGGCTCTGGAACAGTATGAGAAAGAGCGGCAAGAACAGCTGCTGCGGGACACCCCT
GAACACCAACGAGTACCGGCTGAGCAAGAGAGGCCTTAGAGGCTGGCGGGAAATCATCCAGAAA
TGGCTGAAAATGGACgagagctctggaggaagcTCCGAAGTCGAGTTTTCCCATGAGTACTGGA
TGAGACACGCATTGACTCTCGCAAAGAGGGCTCGAGATGAACGCGAGGTGCCCGTGGGGGCAGT
ACTCGTGCTCAACAATCGCGTAATCGGCGAAGGTTGGAATAGGGCAATCGGACTCCACGACCCC
ACTGCACATGCGGAAATCATGGCCCTTCGACAGGGAGGGCTTGTGATGCAGAAATATCGACTTT
ATGATGCGACGCTGTACGTCACGTTTGAACCTTGCGTAATGTGCGCGGGAGCTATGATTCACTC
CCGCATTGGACGAGTTGTATTTCGGTGTTTCGCAACGCCAAGACGGGTGCCGCAGGTTCACTGATG
GACGTGCTGCATCATCCAGGCATGAACACCGGGTAGAAATCACAGAAGGCATATTGGCGGACG
AATGTGCGGCGCTGTTGTGTCGTTTTTTTTTCGCATGCCAGGCGGGTCTTTAACGCCCAGAAAAA
AGCACAATCCTCTACTGACGGCTCTTCTGGATCTGAAACACCTGGCACAAGCGAGAGCGCCACC
CCTGAGAGCTCTGGCGAGAACGAGCCCTCCGAGAAGTACCTGGAAGTGTTCAGGACTACCAGC
GGAAGCACCTTAGAGAGCGCGGCGATTACAGCGTGTACGAGTTCCTGTCCAAGAAAGAGAACCA
CTTCATCTGGCGGAATCACCTGAGTACCCCTACCTGTACGCCACCTTCTGCGAGATCGACAAG
AAAAAGAAGGACGCCAAGCAGCAGGCCACCTTCACACTGGCCGATCCTATCAATCACCCCTCTGT
GGGTCCGATTTCGAGGAAAGAAGCGGCAGCAACCTGAACAAGTACAGAATCCTGACCGAGCAGCT
GCACACCGAGAAGCTGAAGAAAAAGCTGACAGTGCAGCTGGACCGGCTGATCTACCCTACAGAA
TCTGGCGGCTGGGAAGAGAAGGGCAAAGTGGACATTGTGCTGCTGCCAGCCGGCAGTTCTACA
ACCAGATCTTCCTGGACATCGAGGAAAAGGGCAAGCACGCCTTCACCTACAAGGATGAGAGCAT
CAAGTTCCCTCTGAAGGGCACACTCGGCGGAGCCAGAGTGCAGTTCGACAGAGATCACCTGAGA

10

20

30

40

50

AGATACCCTCACAAGGTGGAAAGCGGCAACGTGGGCAGAATCTACTTCAACATGACCGTGAACA
TCGAGCCTACAGAGTCCCCAGTGTCCAAGTCTCTGAAGATCCACCGGGACGACTTCCCCAAGGT
GGTCAACTTCAAGCCCAAAGAACTGACCGAGTGGATCAAGGACAGCAAGGGCAAGAACTGAAG
TCCGGCATCGAGTCCCTGGAAATCGGCCTGAGAGTGATGAGCATCGACCTGGGACAGAGACAGG
CCGCTGCCGCCTCTATTTTCGAGGTGGTGGATCAGAAGCCCCGACATCGAAGGCAAGCTGTTTTT
CCCAATCAAGGGCACCGAGCTGTATGCCGTGCACAGAGCCAGCTTCAACATCAAGCTGCCCCGC
GAGACACTGGTCAAGAGCAGAGAAGTGCTGCGGAAGGCCAGAGAGGACAATCTGAAACTGATGA
ACCAGAAGCTCAACTTCCTGCGGAACGTGCTGCACTTCCAGCAGTTCGAGGACATCACCGAGAG
AGAGAAGCGGGTCACCAAGTGGATCAGCAGACAAGAGAACAGCGACGTGCCCCCTGGTGTACCAG
GATGAGCTGATCCAGATCCGCGAGCTGATGTACAAGCCTTACAAGGACTGGGTGCGCTTCCTGA
AGCAGCTCCACAAGAGACTGGAAGTCGAGATCGGCAAAGAAGTGAAGCACTGGCGGAAGTCCCT
GAGCGACGGAAGAAAGGGCCTGTACGGCATCTCCCTGAAGAACATCGACGAGATCGATCGGACC
CGGAAGTTCCTGCTGAGATGGTCCCTGAGGCCTACCGAACCTGGCGAAGTGCGTAGACTGGAAC
CCGGCCAGAGATTCGCCATCGACCAGCTGAATCACCTGAACGCCCTGAAAGAAGATCGGCTGAA
GAAGATGGCCAACACCATCATCATGCACGCCCTGGGCTACTGCTACGACGTGCGGAAGAAGAAA
TGGCAGGCTAAGAACCCCGCCTGCCAGATCATCCTGTTTCGAGGATCTGAGCAACTACAACCCCT
ACGAGGAAAGGTCCCGCTTCGAGAACAGCAAGCTCATGAAGTGGTCCAGACGCGAGATCCCCAG
ACAGGTTGCACTGCAGGGCGAGATCTATGGCCTGCAAGTGGGAGAAGTGGGCGCTCAGTTCAGC
AGCAGATTCCACGCCAAGACAGGCAGCCCTGGCATCAGATGTAGCGTCGTGACCAAAGAGAAGC
TGCAGGACAATCGGTTCTTCAAGAATCTGCAGAGAGAGGGCAGACTGACCCTGGACAAAATCGC
CGTGCTGAAAGAGGGCGATCTGTACCCAGACAAAGGCGGCGAGAAGTTCATCAGCCTGAGCAAG
GATCGGAAGTGCGTGACCACACACGCCGACATCAACGCCGCTCAGAACCTGCAGAAGCGGTTCT
GGACAAGAACCCACGGCTTCTACAAGGTGTACTGCAAGGCCTACCAGGTGGACGGCCAGACCGT
GTACATCCCTGAGAGCAAGGACCAGAAGCAGAAGATCATCGAAGAGTTCGGCGAGGGCTACTTC
ATTCTGAAGGACGGGTGTACGAATGGGTCAACGCCGGAAGCTGAAAATCAAGAAGGGCAGCT
CCAAGCAGAGCAGCAGCGAGCTGGTGGATAGCGACATCCTGAAAGACAGCTTCGACCTGGCCTC
CGAGCTGAAAGGCGAAAAGCTGATGCTGTACAGGGACCCAGCGGCAATGTGTTCCCCAGCGAC
AAATGGATGGCCGCTGGCGTGTTCTTCGGAAAGCTGGAACGCATCCTGATCAGCAAGCTGACCA
ACCAGTACTCCATCAGCACCATCGAGGACGACAGCAGCAAGCAGTCTATGAAAAGGCCGGCGGC
CACGAAAAAGGCCGCGCCAGGCAAAAAAGAAAAAGGGATCC TACCCATACGATGTTCCAGATTAC
GCTTATCCCTACGACGTGCCTGATTATGCATACCCATATGATGTCCCCGACTATGCCTAA

10

20

30

40

MAPKKKRKVGIHGVPAATRSLKIEPNEEVKKGLWKTHEVLNHGIAYYMNILKLIRQE
AIYEHHEQDPKNPKKVSKAEIQAEWDFVLKMQKNSFTHEVDKDEVFNILRELYEELVP
SSVEKKGEANQLSNKFLYPLVDPNSQSGKGTASSGRKPRWYNLKIAGDPSWEEEEKKKWE
EDKKKDPLAKILGKLAEYGLIPLFIPYTDSEPIVKEIKWMEKSRNQSVRRLDKDMFIQAL
ERFLSWESWNLKVKEEYKVEKEYKTLERI KEDIQALKALEQYEKERQEQLLRDTLNTN
EYRLSKRGLRGWREIIQKWLKMDGGSGGSSEVEFSHEYWMRHALTLAKRARDEREVPVG
AVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLYDATLYVTFEPCVM
CAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLCRFF
RMPRRVFNAQKKAQSSTDGSSGSETPGTSESATPESSGENEPSEKYLEVFKDYQRKHPRE

50

AGDYSVYEFLSKKENHFIWRNHPEYPYLYATFCEIDKKKKDAKQQATFTLADPINHPLW
VRFEERSGSNLNKYRILTEQLHTEKLKKKLTVQLDRLIYPTESGGWEEKGKVDIVLLPSRQ
FYNQIFLDIEEKGKHAFTYKDESIKFPLKGTGGARVQFDRDHLRRYPHKVESGNVGRIFY
NMTVNIEPTESPVSKSLKIHRDDFPKVVNFKPKELTEWIKDSKGKKLKSGIESLEIGLRVM
SIDLGQRQAAAASIFEVVDQKPDIEGKLFFPIKGTELYAVHRASFNIKLPGETLVKSREVL
KAREDNLKLMNQKLNFLRNVLHFQQFEDITEREKRVTKWISRQENS DVPLVYQDELIQIR
ELMYKPYKDWVAFLKQLHHRLEVEIGKEVKHWRKSLSDGRKGLYGISLKNIDEIDRTRKF
LLRWSLRPTEPGEVRRLEPGQRFAIDQLNHLNALKEDRLKKMANTIIMHALGYCYDVRK
KKWQAKNPACQIILFEDLSNYPYEERSRFENSKLMKWSRREIPRQVALQGEIYGLQVGE
VGAQFSSRFHAKTGSPGIRCSVVTKEKLQDNRRFFKNLQREGRLTDKIAVLKEGDLYPDK
GGEKFISLSKDRKCVTTHADINAAQNLQKRFWTRTHGFYKVYCKAYQVDGQTVYIPESK
DQKQKIIIEFGEFYFILKDGVEWVNAGKLIKKGSSKQSSSELVSDILKDSFDLASELKG
EKLMLYRDPSGNVFPSDKWMAAGVFFGKLERILISKLTNQYSISTIEDDSSKQSMKRPA
TKKAGQAKKKKGSYPYDVPDYAYPYDVPDYAYPYDVPDYA
[0 3 8 1]

BhCas12b GSGGS-ABE8-Xten20 at D980

10

20

30

40

50

BhCas12b GSGGS-ABE8-Xten20 at D980

GCCACCATGGCCCCAAAGAAGAAGCGGAAGGTCGGTATCCACGGAGTCCCAGCAGCCGCCACCA
GATCCTTCATCCTGAAGATCGAGCCCCAACGAGGAAGTGAAGAAAGCCTCTGGAAAACCCACGA
GGTGTGAACCACGGAATCGCCTACTACATGAATATCCTGAAGCTGATCCGGCAAGAGGCCATC
TACGAGCACCACGAGCAGGACCCCCAAGAATCCCAAGAAGGTGTCCAAGGCCGAGATCCAGGCCG
AGCTGTGGGATTTCTGTCTGAAGATGCAGAAGTGCAACAGCTTCACACACGAGGTGGACAAGGA
CGAGGTGTTCAACATCCTGAGAGAGCTGTACGAGGAAGTGGTGCCAGCAGCGTGGAAGAAG
GGCGAAGCCAACCAGCTGAGCAACAAGTTTCTGTACCTCTGGTGGACCCCCAACAGCCAGTCTG
GAAAGGGAACAGCCAGCAGCGGCAGAAAGCCCAGATGGTACAACCTGAAGATTGCCGGCGATCC
CTCCTGGGAAGAAGAGAAGAAGAAGTGGGAAGAAGATAAGAAAAAGGACCCGCTGGCCAAGATC
CTGGGCAAGCTGGCTGAGTACGGACTGATCCCTCTGTTTCATCCCCTACACCGACAGCAACGAGC
CCATCGTGAAAGAAATCAAGTGGATGGAAAAGTCCCAGAACAGAGCGTGCGGCGGCTGGATAA
GGACATGTTTCATTTCAGGCCCTGGAACGGTTCCTGAGCTGGGAGAGCTGGAACCTGAAAGTGAA
GAGGAATACGAGAAGGTGAGAAAGAGTACAAGACCCTGGAAGAGAGGATCAAAGAGGACATCC
AGGCTCTGAAGGCTCTGGAACAGTATGAGAAAGAGCGGCAAGAAGAGCTGCTGCGGGACACCCT
GAACACCAACGAGTACCGGCTGAGCAAGAGAGGCCTTAGAGGCTGGCGGAAATCATCCAGAAA
TGGCTGAAAATGGACGAGAACGAGCCCTCCGAGAAGTACCTGGAAGTGTTCAGGACTACCAGC
GGAAGCACCTTAGAGAGGCCGCGATTACAGCGTGTACGAGTTCCTGTCCAAGAAAGAGAACCA
CTTCATCTGGCGGAATCACCTGAGTACCCCTACCTGTACGCCACCTTCTGCGAGATCGACAAG
AAAAAGAAGGACGCCAAGCAGCAGGCCACCTTCACACTGGCCGATCCTATCAATCACCTCTGT
GGGTCCGATTTCGAGGAAAGAAGCGGCAGCAACCTGAACAAGTACAGAATCTTGACCGAGCAGCT
GCACACCGAGAAGCTGAAGAAAAAGCTGACAGTGCAGCTGGACCGGCTGATCTACCCTACAGAA
TCTGGCGGCTGGGAAGAGAAGGGCAAAGTGGACATTGTGCTGCTGCCAGCCGGCAGTTCTACA
ACCAGATCTTCCTGGACATCGAGGAAAAGGGCAAGCACGCCTTCACCTACAAGGATGAGAGCAT
CAAGTTCCCTCTGAAGGGCACACTCGGCGGAGCCAGAGTGCAGTTCGACAGAGATCACCTGAGA
AGATACCCTCACAAGGTGGAAAGCGGCAACGTGGGCAGAATCTACTTCAACATGACCGTGAACA
TCGAGCCTACAGAGTCCCCAGTGTCCAAGTCTCTGAAGATCCACCGGGACGACTTCCCCAAGGT
GGTCAACTTCAAGCCCAAAGAACTGACCGAGTGGATCAAGGACAGCAAGGGCAAGAACTGAAG
TCCGGCATCGAGTCCCTGAAAATCGGCCTGAGAGTGTGAGCATCGACCTGGGACAGAGACAGG
CCGCTGCCGCTCTATTTTCGAGGTGGTGGATCAGAAGCCCGACATCGAAGGCAAGCTGTTTTT
CCCAATCAAGGGCACCGAGCTGTATGCCGTGCACAGAGCCAGCTTCAACATCAAGCTGCCCCGC
GAGACACTGGTCAAGAGCAGAGAAGTGTGCGGAAGGCCAGAGAGACAATCTGAAACTGATGA
ACCAGAAGCTCAACTTCCTGCGGAACGTGCTGCACTTCCAGCAGTTCGAGGACATCACCGAGAG
AGAGAAGCGGTCACCAAGTGGATCAGCAGACAAGAGAACAGCGACGTGCCCTGGTGTACCAG

10

20

30

40

50

GATGAGCTGATCCAGATCCGCGAGCTGATGTACAAGCCTTACAAGGACTGGGTCGCCTTCCTGA
 AGCAGCTCCACAAGAGACTGGAAGTCGAGATCGGCAAAGAAGTGAAGCACTGGCGGAAGTCCCT
 GAGCGACGGAAGAAAGGGCCTGTACGGCATCTCCCTGAAGAACATCGACGAGATCGATCGGACC
 CGGAAGTTCCTGCTGAGATGGTCCCTGAGGCCTACCGAACCTGGCGAAGTGCCTAGACTGGAAC
 CCGGCCAGAGATTGCCCATCGACCAGCTGAATCACCTGAACGCCCTGAAAGAAGATCGGCTGAA
 GAAGATGGCCAACACCATCATCATGCACGCCCTGGGCTACTGCTACGACGTGCGGAAGAAGAAA
 TGGCAGGCTAAGAACCCCGCCTGCCAGATCATCCTGTTTCGAGGATCTGAGCAACTACAACCCCT
 ACGAGGAAAGGTCCCGCTTCGAGAACAGCAAGCTCATGAAGTGGTCCAGACGCGAGATCCCCAG
 ACAGGTTGCACTGCAGGGCGAGATCTATGGCCTGCAAGTGGGAGAAGTGGGCGCTCAGTTCAGC
 AGCAGATTCCACGCCAAGACAGGCAGCCCTGGCATCAGATGTAGCGTCGTGACCAAAGAGAAGC
 TGCAGGACAATCGGTTCTTCAAGAATCTGCAGAGAGAGGGCAGACTGACCCTGGACAAAATCGC
 CGTGCTGAAAGAGGGCGATCTGTACCCAGACAAAGGCGGCGAGAAGTTCATCAGCCTGAGCAAG
 GATCGGAAGTGCCTGACCACACACGCCGACATCAACGCCGCTCAGAACCTGCAGAAGCGGTTCT
 GGACAAGAACCCACGGCTTCTACAAGGTGTACTGCAAGGCCTACCAGGTGGACgaggctctgg
 aggaagcTCCGAAGTCGAGTTTTTCCCATGAGTACTGGATGAGACACGCATTGACTCTCGCAAAG
AGGGCTCGAGATGAACGCGAGGTGCCCGTGGGGGCAGTACTCGTGCTCAACAATCGCGTAATCG
GCGAAGGTTGGAATAGGGCAATCGGACTCCACGACCCCACTGCACATGCGGAAATCATGGCCCT
TCGACAGGGAGGGCTTGTGATGCAGAATTATCGACTTTTATGATGCGACGCTGTACGTCACGTTT
GAACCTTGCGTAATGTGCGGGGAGCTATGATTCACTCCCGCATTGGACGAGTTGTATTCGGTG
TTTCGAACGCCAAGACGGGTGCCGAGGTTCACTGATGGACGTGCTGCATCATCCAGGCATGAA
CCACCGGGTAGAAATCACAGAAGGCATATTGGCGGACGAATGTGCGGCGCTGTTGTGTCGTTTT
TTTCGCATGCCCAGGCGGGTCTTTAACGCCCAGAAAAAGCACAAATCCTCTACTGACGGCTCTT
CTGGATCTGAAACACCTGGCACAAGCGAGAGCGCCACCCCTGAGAGCTCTGGCGGCCAGACCGT
 GTACATCCCTGAGAGCAAGGACCAGAAGCAGAAGATCATCGAAGAGTTTCGGCGAGGGCTACTTC
 ATTCTGAAGGACGGGTGTACGAATGGGTCAACGCCGGAAGCTGAAAATCAAGAAGGGCAGCT
 CCAAGCAGAGCAGCAGCGAGCTGGTGGATAGCGACATCCTGAAAGACAGCTTCGACCTGGCCTC
 CGAGCTGAAAGGCGAAAAGCTGATGCTGTACAGGGACCCAGCGGCAATGTGTTCCCCAGCGAC
 AAATGGATGGCCGCTGGCGTGTTCTTCGAAAGCTGGAACGCATCCTGATCAGCAAGCTGACCA
 ACCAGTACTCCATCAGCACCATCGAGGACGACAGCAGCAAGCAGTCTATGAAAAGGCCGGCGGC
CACGAAAAAGGCCGGCCAGGCAAAAAAGAAAAAGGGATCC TACCCATACGATGTTCCAGATTAC
 GCTTATCCCTACGACGTGCCTGATTATGCATACCCATATGATGTCCCCGACTATGCCTAA

10

20

30

40

MAPKKKRVGIHGVPAATRSLKIEPNEEVKKGLWKTHEVLNHGIAYYMNLKLIRQE
 AIYEHHEQDPKNPKKVSKEIQAEIQLWDFVLKMQKNSFTHEVDKDEVFNILRELYEELVP
 SSVEKKGEANQLSNKFLYPLVDPNSQSGKGTASSGRKPRWYNLKIAGDPSWEEEEKKKWE
 EDKKKDPLAKILGKLAEYGLIPLFIPYTDSEPIVKEIKWMEKSRNQSVRRLDKDMFIQAL
 ERFLSWESWNLKVKEEYKVEKEYKTLERI KEDIQALKALEQYEKERQEQLLRDTLNTN
 EYRLSKRGLRGWREIIQKWLKMDENEPSEKYLEVFQDYQRKHPREAGDYSVYEFLSKKE
 NHFIWRNHPEYPYLYATFCEIDKKKKDAKQQTFTLADPINHPLWVRFEERSGSNLNKY
 RILTEQLHTEKLKKKLTVQLDRLIYPTESGGWEEKGKVDIVLLPSRQFYNQIFLDIEEKGK
 HAFTYKDESIKFLKGTGGARVQFDRDHLRRYPHKVESGNVGRIYFNMTVNIEPTESPV

50

SKSLKIHRDDFPKVVNFKPKELTEWIKDSKGKKLKSGIESLEIGLRVMSIDLGQRQAAAASI
FEVVDQKPDIEGKLFFPIKGTELYAVHRASFNIKLPGETLVKSREVLRKAREDNLKLMNQ
KLNFLRNVLHFQQFEDITEREKRVTKWISRQENS DVPLVYQDELIQIRELMYKPYKDWV
AFLKQLHKRLEVEIGKEVKHWRKSLSDGRKGLYGISLKNIDEIDRTRKFLLRWSLRPTEPG
EVRRLPGQRFAIDQLNHLNALKEDRLKKMANTIIMHALGYCYDVRKKKWQAKNPACQI
ILFEDLSNYPYEERSRFENSKLMKWSRREIPRQVALQGEIYGLQVGEVGAQFSSRFHAK
TGSPGIRCSVVTKKLQDNRFFKNLQREGRLTLDKIAVLKEGDLYPDKGGEKFISLSKDRK
CVTTHADINAAQNLQKRFWTRTHGFYKVYCKAYQVDGGSGGSSEVEFSHEYWMRHALT
LAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRL
YDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHHPGMNHRVEITE
GILADECAALLCRFFRMPRRVFNAQKKAQSSTDGSSGSETPGTSESATPESSGGQTVYIPE
SKDQKQKIIIEFGEGYFILKDGVYEWVNAGKLKIKKGSSKQSSSELVDSILKDSFDLASEL
KGEKLMLYRDPSGNVFP SDKWMAAGVFFGKLERILISKLTNQYSISTIEDDSSKQSMKRP
AATKKAGQAKKKKGSPYDVPDYAYPYDVPDYAYPYDVPDYA
[0 3 8 2]

BhCas12b GGSGGS-ABE8-Xten20 at K1019

10

20

30

40

50

BhCas12b GSGGS-ABE8-Xten20 at K1019

GCCACCATGGCCCCAAAGAAGAAGCGGAAGGTCGGTATCCACGGAGTCCCAGCAGCCGCCACCA
GATCCTTCATCCTGAAGATCGAGCCCAACGAGGAAGTGAAGAAAGGCCTCTGGAAAACCCACGA
GGTGTGAACACGGAATCGCCTACTACATGAATATCCTGAAGCTGATCCGGCAAGAGGCCATC
TACGAGCACCACGAGCAGGACCCCAAGAATCCCAAGAAGGTGTCCAAGGCCGAGATCCAGGCCG
AGCTGTGGGATTTCTGTGCTGAAGATGCAGAAGTGCAACAGCTTCACACACGAGGTGGACAAGGA
CGAGGTGTTCAACATCCTGAGAGAGCTGTACGAGGAAGTGGTGCCAGCAGCGTGGAAAAGAA
GGCGAAGCCAACCAGCTGAGCAACAAGTTTCTGTACCCTCTGGTGGACCCCAACAGCCAGTCTG
GAAAGGGAACAGCCAGCAGCGGCAGAAAGCCCAGATGGTACAACCTGAAGATTGCCGGCGATCC
CTCCTGGGAAGAAGAGAAGAAGAAGTGGGAAGAAGATAAGAAAAAGGACCCGCTGGCCAAGATC
CTGGGCAAGCTGGCTGAGTACGGACTGATCCCTCTGTTTCATCCCCTACACCGACAGCAACGAGC
CCATCGTGAAAGAAATCAAGTGGATGGAAAAGTCCCGGAACCAGAGCGTGCGGGCGCTGGATAA
GGACATGTTTCATTTCAGGCCCTGGAACGGTTCTTGAGCTGGGAGAGCTGGAACCTGAAAGTGAAA
GAGGAATACGAGAAGGTCGAGAAAAGTACAAGACCCTGGAAGAGAGGATCAAAGAGGACATCC
AGGCTCTGAAGGCTCTGGAACAGTATGAGAAAGAGCGGCAAGAACAGCTGCTGCGGGACACCCT
GAACACCAACGAGTACCGGCTGAGCAAGAGAGGCCTTAGAGGCTGGCGGGAAATCATCCAGAAA
TGGCTGAAAATGGACGAGAACGAGCCCTCCGAGAAGTACCTGGAAGTGTTCAGGACTACCAGC
GGAAGCACCTAGAGAGGCCGGGATTACAGCGTGTACGAGTTCCTGTCCAAGAAAGAGAACCA
CTTCATCTGGCGGAATCACCTGAGTACCCCTACCTGTACGCCACCTTCTGCGAGATCGACAAG
AAAAAGAAGGACGCCAAGCAGCAGGCCACCTTCACACTGGCCGATCCTATCAATCACCTCTGT
GGGTCCGATTTCGAGGAAAGAAGCGGCAGCAACCTGAACAAGTACAGAATCCTGACCGAGCAGCT
GCACACCGAGAAGCTGAAGAAAAAGCTGACAGTGCAGCTGGACCGGCTGATCTACCTACAGAA
TCTGGCGGCTGGGAAGAGAAGGGCAAAGTGGACATTGTGCTGCTGCCCAGCCGGCAGTTCTACA
ACCAGATCTTCCTGGACATCGAGGAAAAGGGCAAGCACGCCTTCACCTACAAGGATGAGAGCAT
CAAGTTCCCTCTGAAGGGCACACTCGGCGGAGCCAGAGTGCAGTTCGACAGAGATCACCTGAGA
AGATACCCTCACAAGGTGGAAAGCGGCAACGTGGGCAGAATCTACTTCAACATGACCGTGAACA
TCGAGCCTACAGAGTCCCCAGTGTCCAAGTCTCTGAAGATCCACCGGGACGACTTCCCCAAGGT
GGTCAACTTCAAGCCCAAGAACTGACCGAGTGGATCAAGGACAGCAAGGGCAAGAACTGAAG
TCCGGCATCGAGTCCCTGGAAATCGGCCTGAGAGTGATGAGCATCGACCTGGGACAGAGACAGG
CCGCTGCCGCTCTATTTTCGAGGTGGTGGATCAGAAGCCCGACATCGAAGGCAAGCTGTTTTT
CCCAATCAAGGGCACCGAGCTGTATGCCGTGCACAGAGCCAGCTTCAACATCAAGCTGCCCCGC
GAGACACTGGTCAAGAGCAGAGAAGTGTGCGGAAGGCCAGAGAGGACAATCTGAAACTGATGA
ACCAGAAGCTCAACTTCCTGCGGAACGTGCTGCACCTCCAGCAGTTCGAGGACATCACCGAGAG
AGAGAAGCGGGTCACCAAGTGGATCAGCAGACAAGAGAACAGCGACGTGCCCCTGGTGTACCAG

10

20

30

40

50

GATGAGCTGATCCAGATCCGCGAGCTGATGTACAAGCCTTACAAGGACTGGGTGCGCTTCCTGA
AGCAGCTCCACAAGAGACTGGAAGTCGAGATCGGCAAAGAAGTGAAGCACTGGCGGAAGTCCCT
GAGCGACGGAAGAAAGGGCCTGTACGGCATCTCCCTGAAGAACATCGACGAGATCGATCGGACC
CGGAAGTTCCTGCTGAGATGGTCCCTGAGGCCTACCGAACCTGGCGAAGTGCGTAGACTGGAAC
CCGGCCAGAGATTGCCATCGACCAGCTGAATCACCTGAACGCCCTGAAAGAAGATCGGCTGAA
GAAGATGGCCAACACCATCATCATGCACGCCCTGGGCTACTGCTACGACGTGCGGAAGAAGAAA
TGGCAGGCTAAGAACCCCGCCTGCCAGATCATCCTGTTTCGAGGATCTGAGCAACTACAACCCCT
ACGAGGAAAGGTCCCGCTTCGAGAACAGCAAGCTCATGAAGTGGTCCAGACGCGAGATCCCCAG
ACAGGTTGCACTGCAGGGCGAGATCTATGGCCTGCAAGTGGGAGAAGTGGGCGCTCAGTTCAGC
AGCAGATTCCACGCCAAGACAGGCAGCCCTGGCATCAGATGTAGCGTCGTGACCAAAGAGAAGC
TGCAGGACAATCGGTTCTTCAAGAATCTGCAGAGAGAGGGCAGACTGACCCTGGACAAAATCGC
CGTGCTGAAAGAGGGCGATCTGTACCCAGACAAAGGCGGCGAGAAGTTCATCAGCCTGAGCAAG
GATCGGAAGTGCGTGACCACACACGCCGACATCAACGCCGCTCAGAACCTGCAGAAGCGGTTCT
GGACAAGAACCCACGGCTTCTACAAGGTGTACTGCAAGGCCCTACCAGGTGGACGGCCAGACCGT
GTACATCCCTGAGAGCAAGGACCAGAAGCAGAAGATCATCGAAGAGTTCGGCGAGGGCTACTTC
ATTCTGAAGGACGGGGTGACGAATGGGTCAACGCCGGCAAGGgaggctctggaggaagcTCCG
AAGTCGAGTTTTTCCCATGAGTACTGGATGAGACACGCATTGACTCTCGCAAAGAGGGCTCGAGA
TGAACGCGAGGTGCCCCGTGGGGGCAGTACTCGTGCTCAACAATCGCGTAATCGGCGAAGGTTGG
AATAGGGCAATCGGACTCCACGACCCCACTGCACATGCGGAAATCATGGCCCTTCGACAGGGAG
GGCTTGTGATGCAGAATTATCGACTTTATGATGCGACGCTGTACGTCACGTTTGAACCTTGCGT
AATGTGCGCGGGAGCTATGATTCACTCCCGCATTGGACGAGTTGTATTTCGGTGTTCGCAACGCC
AAGACGGGTGCCGACAGGTTCACTGATGGACGTGCTGCATCATCCAGGCATGAACCACCGGGTAG
AAATCACAGAAGGCATATTGGCGGACGAATGTGCGGCGCTGTTGTGTCGTTTTTTTCGCATGCC
CAGGCGGGTCTTTAACGCCCCAGAAAAAGCACAATCCTCTACTGACGGCTCTTCTGGATCTGAA
ACACCTGGCACAAGCGAGAGCGCCACCCCTGAGAGCTCTGGCCTGAAAATCAAGAAGGGCAGCT
CCAAGCAGAGCAGCAGCGAGCTGGTGGATAGCGACATCCTGAAAGACAGCTTCGACCTGGCCTC
CGAGCTGAAAGGCGAAAAGCTGATGCTGTACAGGGACCCAGCGGCAATGTGTTCCCCAGCGAC
AAATGGATGGCCGCTGGCGTGTCTTCGGAAAGCTGGAACGCATCCTGATCAGCAAGCTGACCA
ACCAGTACTCCATCAGCACCATCGAGGACGACAGCAGCAAGCAGTCTATGAAAAGGCCGGCGGC
CACGAAAAAGGCCGGCCAGGCAAAAAAGAAAAAGGGATCC TACCCATACGATGTTCCAGATTAC
GCTTATCCCTACGACGTGCCTGATTATGCATACCCATATGATGTCCCCGACTATGCCTAA

10

20

30

40

MAPKKKRKVGIHGVPAATRSLFILKIEPNEEVKKGLWKTHEVLNHGIAYYMNILKLIRQE
AIYEHHEQDPKNPKKVSKAEIQAEWDFVLKMQKCNSTHEVDKDEVFNILRELYEELVP
SSVEKKGEANQLSNKFLYPLVDPNSQSGKGTASSGRKPRWYNLKIAGDPSWEEEEKKKWE
EDKKKDPLAKILGKLAEYGLIPLFIPYTDSEPIVKEIKWMEKSRNQSVRRLDKDMFIQAL
ERFLSWESWNLKVKEEYKVEKEYKTLERI KEDIQALKALEQYEKERQEQLLRDTLNTN
EYRLSKRGLRGWREIIQKWLKMDENEPSEKYLEVFQDYQRKHPREAGDYSVYEFLSKKE
NHFIWRNHPEYPYLYATFCEIDKKKKDAKQQTFTLADPINHPLWVRFEERSGSNLNKY
RILTEQLHTEKLKKKLTVQLDRLIYPTESGGWEEKGKVDIVLLPSRQFYNQIFLDIEEKGK
HAFTYKDESIKFPLKGTLGGARVQFDRDHLRRYPHKVESGNVGRIFYNMTVNIEPTESPV

50

SKSLKIHRDDFPKVVNFKPKELTEWIKDSKGKKLKSGIESLEIGLRVMSIDLGRQAAAASI
FEVVDQKPDIEGKLFPIKGTELYAVHRASFNILKLPGETLVKSREVLARKAREDNKLKLMNQ
KLNFLRNVLHFQQFEDITEREKRVTKWISRQENS DVPLVYQDELIQIRELMYKPYKDWV
AFLKQLHKRLEVEIGKEVKHWRKSLSDGRKGLYGISLKNIDEIDRTRKFLLRWSLRPTEPG
EVRRLPEPGQRFAIDQLNHLNALKEDRLKKMANTIIMHALGYCYDVRKKKKWQAKNPACQI
ILFEDLSNYPYEERSRFENSKLMKWSRREIPRQVALQGEIYGLQVGEVGAQFSSRFHAK
TGSPGIRCSVVTKEKLQDNRFFKNLQREGRLTLDKIAVLKEGDLYPDKGGEKFISLSKDRK
CVTTHADINAAQNLQKRFWTRTHGFYKVYCKAYQVDGQTVYIPESKDQKQKIIEEFGE
YFILKDGVEYEWVNAGKGGSGGSSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVN
RVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLYDATLYVTFEPCVMCAGAMIH
SRIGRVVFGVRNAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLCRFFRMPRRV
FNAQKKAQSSTDGSSGSETPGTSESATPESSGLKIKKGSSKQSSSELVDSILKDSFDLASE
LKGEKLMLYRDP SGNVFPSDKWMAAGVFFGKLERILISKLTNQYSISTIEDDSSKQSMKR
PAATKKAGQAKKKKGSYPYDVPDYAYPYDVPDYAYPYDVPDYA

【0383】

上記の配列では、コザック配列は太字で下線が引かれている。N末端核局在化シグナル(NLS)をマークしている。小文字はGGGSGGSリンカーを示す。破線による下線はABE8をコードする配列をマークし、修飾されていない配列はBhCas12bをコードする。二重下線はXten20リンカーを示す。単一下線はC末端NLSを示す。点線による下線のGGATCCはGSリンカーを示す。イタリック体の文字は、3x血球凝集素(HA)タグのコード配列を表す。

【0384】

[ガイドポリヌクレオチド]

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。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.” Deltcheva 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, 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オーソログは、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はニックアーゼである。

【0385】

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(「s

gRNA」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。

【0386】

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。ある実施形態では、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。ある実施形態では、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。ある実施形態では、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。

【0387】

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。

【0388】

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(trRNA) 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。

【0389】

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。

【0390】

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。

【0391】

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、複合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。

【0392】

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。

【0393】

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。

【0394】

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。

【0395】

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。

【0396】

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。ある実施形態では、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。

【0397】

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。

【0398】

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。

【0399】

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。ある実施形態では、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;他の実施形態では、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、あるいはいくつかの実施形態では、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。

【0400】

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。

【0401】

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ペ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。ある実施形態では、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。

10

【0402】

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枝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。

20

【0403】

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。

30

40

【0404】

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

50

、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。

【0405】

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、HaloTag、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。

【0406】

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。

【0407】

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。

【0408】

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分子が環状であってもよい。

【0409】

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）。

【0410】

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。

【0411】

ある実施形態では、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。ある実施形態では、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。

【0412】

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。

【0413】

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スペーサー、C3スペーサー、C6スペーサー、dスペーサー、PCスペーサー、rスペーサー、スペーサー18、スペーサー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、DNP 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。

【0414】

ある実施形態では、修飾は永続的である。その他の実施形態では、修飾は一時的である

。ある実施形態では、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。

【0415】

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。

10

【0416】

修飾はホスホロチオエート置換でもあり得る。ある実施形態では、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。ある実施形態では、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。ある実施形態では、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

【0417】

異なるCas12bオルソログ（例えばBhCas12b、BvCas12b、AaCas12b）は、異なるスキファールド（足場）配列（tracrRNAとも呼ばれる）を使用する。いくつかの実施形態において、足場配列は、BhCas12bタンパク質と共に使用するために最適化されており、以下の配列を有する：（実際のgRNAにおいて、Tは、ウリジン（U）によって置き換えられる）。

BhCas12b sgRNAスキファールド（下線）+ 20nt〜23nt ガイド配列（Nによって示される）。

5 GTTCCTGTCTTTTGGTCAGGACAACCGTCTAGCTATAAGTGCTGCAGGGTGTGAGAAAC
TCCTATTGCTGGACGATGTCTCTTACGAGGCATTAGCACNNNNNNNNNNNNNNNNNNNN
 NN-3'

30

【0418】

ある実施形態では、スキファールド配列はBvCas12bタンパク質との使用のために最適化され、以下の配列を有する：（実際のgRNAにおいて、Tは、ウリジン（U）によって置き換えられる）。

BvCas12b sgRNAスキファールド（下線）+ 20nt〜23ntガイド配列（Nによって示される）。

5 GACCTATAGGGTCAATGAATCTGTGCGTGTGCCATAAGTAATTAATAAAATTACCCACCA
CAGGAGCACCTGAAAACAGGTGCTTGCCACNNNNNNNNNNNNNNNNNNNNNNNNNNNN-3'

40

【0419】

ある実施形態では、スキファールド配列はAaCas12bタンパク質との使用のために最適化され、以下の配列を有する：（実際のgRNAにおいて、Tは、ウリジン（U）によって置き換えられる）。

AaCas12b sgRNAスキファールド（下線）+ 20nt〜23ntガイド配列（Nによって示される）。

5 GTCTAAAGGACAGAATTTTTCAACGGGTGTGCCAATGGCCACTTTCCAGGTGGCAAAG
CCCGTTGAACCTTCTCAAAAAGAACGATCTGAGAAGTGGCACNNNNNNNNNNNNNNNNNNNN

50

NNNN-3'

【0420】

したがって、当業者はCasタンパク質特異性のゲノム標的を変化させることができ、これは、ゲノムの他の部分と比較してgRNA標的化配列がゲノム標的に対していかに特異的であるかによって部分的に決定される。

【0421】

〔プロトスペーサー隣接モチーフ〕

「プロトスペーサー隣接モチーフ(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 (すなわちプロトスペーサーの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 (すなわちプロトスペーサーの5' 末端の下流に位置する)であり得る。

【0422】

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。

【0423】

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ペーサー隣接モチーフ (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。

【0424】

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。いくつかのPAMバリエーションが下記表1に記載されている。

【0425】

表1. Cas9タンパク質および対応するPAM配列

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
spCas9-MQKFRAER	NGC
Cpf1	5' (TTTV)
SpyMac	5'-NAA-3'

【0426】

いくつかの実施形態では、PAMはNGCである。いくつかの実施形態では、NGC PAMは、Cas9バリエントによって認識される。いくつかの実施形態において、NGC PAMバリエントは、D1135M、S1136Q、G1218K、E1219F、A1322R、D1332A、R1335E、およびT1337R（総称して「MQKFRAER」と呼ばれる）から選択される1つ以上のアミノ酸置換を含む。

【0427】

いくつかの実施形態において、PAMはNGTである。いくつかの実施形態において、NGT PAMはバリエントCas9によって認識される。いくつかの実施形態において、NGT PAMバリエントは、残基1335、1337、1135、1136、1218、および/または1219の一つまたは複数における標的化突然変異を通じて作成される。いくつかの実施形態において、NGT

PAMバリエーションは、残基1219、1335、1337、1218の一つまたは複数における標的化突然変異を通じて作成される。いくつかの実施形態において、NGT PAMバリエーションは、残基1135、1136、1218、1219、および1335の一つまたは複数における標的化突然変異を通じて作成される。いくつかの実施形態において、NGT PAMバリエーションは、下記表2および表3に提供される標的化突然変異のセットから選択される。

【0428】

表2：残基1219、1335、1337、1218におけるNGT PAMバリエーション変異

【表2】

バリエーション	E1219V	R1335Q	T1337	G1218
1	F	V	T	
2	F	V	R	
3	F	V	Q	
4	F	V	L	
5	F	V	T	R
6	F	V	R	R
7	F	V	Q	R
8	F	V	L	R
9	L	L	T	
10	L	L	R	
11	L	L	Q	
12	L	L	L	
13	F	I	T	
14	F	I	R	
15	F	I	Q	
16	F	I	L	
17	F	G	C	
18	H	L	N	
19	F	G	C	A
20	H	L	N	V
21	L	A	W	
22	L	A	F	
23	L	A	Y	
24	I	A	W	
25	I	A	F	
26	I	A	Y	

【0429】

表3：残基1135、1136、1218、1219、および1335におけるNGT PAMバリエーション変異

10

20

30

40

50

【表 3】

バリエント	D1135L	S1136R	G1218S	E1219V	R1335Q
27	G				
28	V				
29	I				
30		A			
31		W			
32		H			
33		K			
34			K		
35			R		
36			Q		
37			T		
38			N		
39				I	
40				A	
41				N	
42				Q	
43				G	
44				L	
45				S	
46				T	
47					L
48					I
49					V
50					N
51					S
52					T
53					F
54					Y
55	N1286Q	I1331F			

【0430】

いくつかの実施形態において、NGT PAMバリエントは、表2および3のバリエント5、7、28、31、または36から選択される。いくつかの実施形態において、バリエントは、改善されたNGT PAM認識を有する。

【0431】

いくつかの実施形態では、NGT PAMバリエントは、残基1219、1335、1337および/または1218において突然変異を有する。いくつかの実施形態では、NGT PAMバリエントは、下記表4に提供されるバリエントから、認識を改善するための変異を伴って選択される。

【0432】

表4：残基1219、1335、1337、および1218におけるNGT PAMバリエント変異

【表4】

バリエント	E1219V	R1335Q	T1337	G1218
1	F	V	T	
2	F	V	R	
3	F	V	Q	
4	F	V	L	
5	F	V	T	R
6	F	V	R	R
7	F	V	Q	R
8	F	V	L	R

10

【0433】

いくつかの実施形態において、NGT PAMは、下記表5に提供されるバリエントから選択される。

【0434】

表5 A. NGT PAMバリエント

【表5-1】

	NGTNバ リエント	D1135	S1136	G1218	E1219	A1322R	R1335	T1337
バリエ ント1	LRKIQK	L	R	K	I	-	Q	K
バリエ ント2	LRSVQK	L	R	S	V	-	Q	K
バリエ ント3	LRSVQL	L	R	S	V	-	Q	L
バリエ ント4	LRKIRQK	L	R	K	I	R	Q	K
バリエ ント5	LRSVRQK	L	R	S	V	R	Q	K
バリエ ント6	LRSVRQL	L	R	S	V	R	Q	L

30

40

【0435】

いくつかの実施形態では、NGTNバリエントは、バリエント1である。いくつかの実施形

50

態では、NGTNバリエーションは、バリエーション2である。いくつかの実施形態では、NGTNバリエーションは、バリエーション3である。いくつかの実施形態では、NGTNバリエーションは、バリエーション4である。いくつかの実施形態では、NGTNバリエーションは、バリエーション5である。いくつかの実施形態では、NGTNバリエーションは、バリエーション6である。

【0436】

ある実施形態において、Cas9ドメインは、*Streptococcus pyogenes*由来のCas9ドメインである (SpCas9)。ある実施形態において、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、配列番号1における番号付けで、D10X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、配列番号1における番号付けで、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。ある実施形態において、SpCas9ドメイン、SpCas9dドメインまたはSpCas9nドメインは、非正準PAMを有する核酸配列に結合することができる。ある実施形態において、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、配列番号1における番号付けで、D1135X、R1335X、およびT1337X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対応する変異を含み、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、配列番号1における番号付けで、D1136E、R1335Q、およびT1337R変異のうちの1つ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変異を含む。いくつかの実施形態において、SpCas9ドメインは、配列番号1における番号付けで、D1135E、R1335Q、およびT1337R変異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変異を含む。いくつかの実施形態において、SpCas9ドメインは、配列番号1における番号付けで、D1135X、R1335X、およびT1337X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対応する変異を含み、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、配列番号1における番号付けで、D1135V、R1335Q、およびT1337R変異のうちの1つ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変異を含む。いくつかの実施形態において、SpCas9ドメインは、配列番号1における番号付けで、D1135V、R1335Q、およびT1337R変異、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変異を含む。いくつかの実施形態において、SpCas9ドメインは、配列番号1における番号付けで、D1135X、G1218X、R1335X、およびT1337X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対応する変異を含み、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、配列番号1における番号付けで、D1135V、G1218R、R1335Q、およびT1337R変異のうちの1つ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変異を含む。いくつかの実施形態において、SpCas9ドメインは、配列番号1における番号付けで、D1135V、G1218R、R1335Q、およびT1337R変異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変異を含む。

【0437】

ある実施形態では、Cas9は、改変されたPAM配列に対する特異性を有するCas9バリエーションである。ある実施形態では、追加的なCas9バリエーションおよびPAM配列はMiller et al., Continuous evolution of SpCas9 variants compatible with non-G PAMs. Nat Biotechnol (2020). <https://doi.org/10.1038/s41587-020-0412-8>に記述されており、その全体が参照により本明細書に組み入れられる。ある実施形態では、Cas9バリエーションは特定のPAM要求性を持たない。ある実施形態では、Cas9バリエーション、例えばSpCas9バリエーションは、NRNH PAMに対する特異性を有し、ここでRはAまたはGでありHはA、C、またはTである。ある実施形態では、SpCas9バリエーションは、PAM配列AAA, TAA, CA

A, GAA, TAT, GAT, またはCACに対する特異性を有する。ある実施形態では、SpCas9バリエーションは、配列番号1における番号付けで位置1114, 1134, 1135, 1137, 1139, 1151, 1180, 1188, 1211, 1218, 1219, 1221, 1249, 1256, 1264, 1290, 1318, 1317, 1320, 1321, 1323, 1332, 1333, 1335, 1337, もしくは1339またはそれに対応する位置におけるアミノ酸置換を含む。ある実施形態では、SpCas9バリエーションは、配列番号1における番号付けで位置1114, 1135, 1218, 1219, 1221, 1249, 1320, 1321, 1323, 1332, 1333, 1335, もしくは1337またはそれに対応する位置におけるアミノ酸置換を含む。ある実施形態では、SpCas9バリエーションは、配列番号1における番号付けで位置1114, 1134, 1135, 1137, 1139, 1151, 1180, 1188, 1211, 1219, 1221, 1256, 1264, 1290, 1318, 1317, 1320, 1323, 1333またはそれに対応する位置におけるアミノ酸置換を含む。ある実施形態では、SpCas9バリエーションは、配列番号1における番号付けで位置1114, 1131, 1135, 1150, 1156, 1180, 1191, 1218, 1219, 1221, 1227, 1249, 1253, 1286, 1293, 1320, 1321, 1332, 1335, 1339またはそれに対応する位置におけるアミノ酸置換を含む。ある実施形態では、SpCas9バリエーションは、配列番号1における番号付けで位置1114, 1127, 1135, 1180, 1207, 1219, 1234, 1286, 1301, 1332, 1335, 1337, 1338, 1349またはそれに対応する位置におけるアミノ酸置換を含む。SpCas9バリエーションの例示的なアミノ酸置換およびPAM特異性は下記表5B, 5C, 5D, および5Eに示される。

【0438】

表5B：追加のバリエーションの変異およびPAM

【表5-2】

SpCas9 /PAM	SpCas9 アミノ酸位置												
	1114	1135	1218	1219	1221	1249	1320	1321	1323	1332	1333	1335	1337
	R	D	G	E	Q	P	A	P	A	D	R	R	T
AAA		N		V	H						G		
AAA		N		V	H						G		
AAA				V							G		
TAA	G	N		V							I		
TAA		N		V							I		A
TAA	G	N		V							I		A
CAA				V							K		
CAA		N		V							K		
CAA		N		V							K		
GAA				V	H		V				K		
GAA		N		V			V				K		
GAA				V	H		V				K		
TAT			S	V	H	S		S				L	
TAT			S	V	H	S		S				L	
TAT			S	V	H	S		S				L	
GAT				V							I		
GAT				V					D			Q	
GAT				V					D			Q	
CAC				V						N		Q	N
CAC		N		V								Q	N
CAC				V						N		Q	N

【0439】

表5C：追加のバリエーションの変異およびPAM

10

20

30

40

50

【表 5 - 3】

SpCas9 /PAM	SpCas9 アミノ酸位置																		
	111 4	113 4	113 5	113 7	113 9	115 1	118 0	118 8	121 1	121 9	122 1	125 6	126 4	129 0	131 8	131 7	132 0	132 3	133 3
	R	F	D	P	V	K	D	K	K	E	Q	Q	H	V	L	N	A	A	R
GAA										V	H						V		K
GAA			N	S						V							V	D	K
GAA			N							V	H		Y				V		K
CAA			N							V	H		Y				V		K
CAA	G		N	S						V	H		Y				V		K
CAA			N					R		V	H						V		K
CAA			N				G		R	V	H		Y				V		K
CAA			N							V	H		Y				V		K
AAA			N				G			V	H	R	Y				V	D	K
CAA	G		N				G			V	H		Y				V	D	K
CAA		L	N				G			V	H		Y			T	V	D	K
TAA	G		N				G			V	H		Y	G	S		V	D	K
TAA	G		N			E	G			V	H		Y		S		V		K
TAA	G		N				G			V	H		Y		S		V	D	K
TAA	G		N				G		R	V	H						V		K
TAA			N				G		R	V	H		Y				V		K
TAA	G		N		A		G			V	H						V		K
TAA	G		N							V	H						V		K

【 0 4 4 0】

表 5 D：追加のバリエーションの変異およびPAM

【表 5 - 4】

SpCas9/ PAM	SpCas9 アミノ酸位置																			
	11 14	11 31	11 35	11 50	11 56	11 80	11 91	12 18	12 19	12 21	12 27	12 49	12 53	12 86	12 93	13 20	13 21	13 32	13 35	13 39
	R	Y	D	E	K	D	K	G	E	Q	A	P	E	N	A	A	P	D	R	T
SacB.T AT			N				N		V	H						V	S		L	
SacB.T AT			N					S	V	H		S					S	G	L	
AAT			N					S	V	H	V	S		K	T		S	G	L	I
TAT	G		N			G		S	V	H		S	K				S	G	L	
TAT	G		N			G		S	V	H		S					S	G	L	
TAT	G	C	N			G		S	V	H		S					S	G	L	
TAT	G	C	N			G		S	V	H		S					S	G	L	
TAT	G	C	N			G		S	V	H		S					S	G	L	
TAT	G	C	N		E	G		S	V	H		S					S	G	L	
TAT	G	C	N	V		G		S	V	H		S					S	G	L	
TAT		C	N			G		S	V	H		S					S	G	L	
TAT	G	C	N			G		S	V	H		S					S	G	L	

【 0 4 4 1】

表 5 E：追加のバリエーションの変異およびPAM

10

20

30

40

50

【表 5 - 5】

SpCas9	SpCas9 アミノ酸位置													
	111 4	112 7	113 5	118 0	120 7	121 9	123 4	128 6	130 1	133 2	133 5	133 7	133 8	134 9
	R	D	D	D	E	E	N	N	P	D	R	T	S	H
SacB.CA C			N			V				N	Q	N		
AAC	G		N			V				N	Q	N		
AAC	G		N			V				N	Q	N		
TAC	G		N			V				N	Q	N		
TAC	G		N			V		H		N	Q	N		
TAC	G		N		G	V	D	H		N	Q	N		
TAC	G		N			V				N	Q	N		
TAC	G	G	N	E		V		H		N	Q	N		
TAC	G		N			V		H		N	Q	N		
TAC	G		N			V				N	Q	N	T	R

【0 4 4 2】

ある実施形態では、Cas9は*Neisseria meningitidis* Cas9 (NmeCas9)またはそのバリエーションである。ある実施形態では、NmeCas9はNNNGAYW PAMに対する特異性を有し、ここでYはCまたはTでありWはAまたはTである。ある実施形態では、NmeCas9はNNNGYTT PAMに対する特異性を有し、ここでYはCまたはTである。ある実施形態では、NmeCas9はNNNGTCT PAMに対する特異性を有する。ある実施形態では、NmeCas9はNme1 Cas9である。ある実施形態では、NmeCas9はNNNGATT PAM、NNNCCTA PAM、NNNCCTC PAM、NNNCCTT PAM、NNNCCTG PAM、NNNNCCGT PAM、NNNNCCGGPAM、NNNNCCCA PAM、NNNNCCCT PAM、NNNNCCCC PAM、NNNNCCAT PAM、NNNNCCAG PAM、NNNNCCAT PAM、またはNNNGATT PAMに対する特異性を有する。ある実施形態では、Nme1Cas9はNNNGATT PAM、NNNCCTA PAM、NNNCCTC PAM、NNNCCTT PAM、またはNNNCCTG PAMに対する特異性を有する。ある実施形態ではNmeCas9はCAA PAM、CAAA PAM、またはCCA PAMに対する特異性を有する。ある実施形態ではNmeCas9はNme2 Cas9である。ある実施形態ではNmeCas9はNmeCas9はNNNNCC (N4CC) PAMに対する特異性を有し、ここでNはA, G, C,またはTのいずれかである。ある実施形態ではNmeCas9はNNNNCCGT PAM、NNNNCCGGPAM、NNNNCCCA PAM、NNNNCCCT PAM、NNNNCCCC PAM、NNNNCCAT PAM、NNNNCCAG PAM、NNNNCCAT PAM、またはNNNGATT PAMに対する特異性を有する。ある実施形態では、NmeCas9はNme3Cas9である。ある実施形態では、NmeCas9はNNNNCAAA PAM、NNNNCC PAM、またはNNNNCINN PAMに対する特異性を有する。Edraki et al. Mol. Cell. (2019) 73(4): 714-726に記述されたさらなるNmeCas9の特徴とPAM配列はその全体が参照により本明細書に組み入れられる。

【0 4 4 3】

Nme1Cas9の例示的なアミノ酸配列を以下に提供する：

type II CRISPR RNA誘導型エンドヌクレアーゼCas9 [*Neisseria meningitidis*] WP_002235162.1

```

1 maafknpin yilgldigia svgwamveid edenpiclid lgvrverae vpktgdsiam
61 arllarsvrr ltrrrahrll rarrllkreg vlqaadfdn glikslpntp wqlraaaldr
121 kltplewsav llhlikhrgy lsqrkneget adkelgallk gvadnahalq tgdftrpacl
181 alnkfekesg hirnqrgdys htfsrkdlda elillfekqk efgnphvsgg lkegiellm
241 tqrpalsgda vqkmlghctf epaepkaakn tytaerfiwl tklnnlrile qgserpltdt
301 eratlmdepy rskltyaqa rklglgledta ffgklyrgkd naeastlmem kayhaisral

```

361 ekeglkdkks plnlspelqd eigtafslfk tdeditgrlk driqpeilea llkhisfdkf
 421 vqislkalrr ivplmeqgkr ydeacaeiyg dhygkntee kiylppipad eirnpvvla
 481 lsqarkving vrrygspar ihietarevg ksfkdrkeie krqeenrkdr ekaaakfrey
 541 fpnfvgpeks kdilklrlye qqhgkclysg keinlgrlne kgyveidhal pfsrtwddsf
 601 nnkvlvlge nqnkgntqpy eyfngkdnsr ewqefkarve tsrfprskkq rillqkfdded
 661 gfkernlndt ryvnrflcqf vadmrltgk gkkrvfafng qitnllrgfw glrkvraend
 721 rhhaldavvv acstvamqqk itrfrvykem nafdgktidk etgevlhqkt hfpqpweffa
 781 qevmirvfgk pdgkpefeea dtpeklrtll aeklssrpea vheyvtplfv srpnrkmsg
 841 qghmetvksa krldegvsvl rvpltqlklk dlekmvnrer epklyealka rleahkddpa
 901 kafaepfyky dkagntqqv kavrveqvqk tgvwvrnhng iadnatmvrsv dvfekgd

kyy

961 lvpiyswqva kgilpdravv qgkdeedwql iddsfnfkfs lhpndlvevi tkkarmfgyf
 1021 aschrgtgni nirihldhk igkngilegi gvktalsfqk yqidalgkei rpcrlkkrrp
 1081 vr

【0444】

Nme2Cas9の例示的なアミノ酸配列を以下に提供する：

type II CRISPR RNA誘導型エンドヌクレアーゼCas9 [Neisseria meningitidis] WP_002230835.1

1 maafknpin yilgldigia svgwamveid eenpirlid lgvrvferae vpktgdslam
 61 arrlarsvrr ltrrrahrll rarrllkreg vlqaadfdn glikslntp wqlraaaldr
 121 kltplewsav llhlikhrgy lsqrkneget adkelgallk gvannahalq tgdftrpael
 181 alnkfekesg hirnqrgdys htfsrkdla elillfekqk efgnphvsgg lkegietllm
 241 tqrpalsgda vqkmlghctf epaepkaakn tytaerfiwl tklnnlrile qgserpltdt
 301 eratlmdepy rkskltyaqa rklglledta ffgklyrgkd naeastlmem kayhaisral
 361 ekeglkdkks plnlsselqd eigtafslfk tdeditgrlk drvqpeilea llkhisfdkf
 421 vqislkalrr ivplmeqgkr ydeacaeiyg dhygkntee kiylppipad eirnpvvla
 481 lsqarkving vrrygspar ihietarevg ksfkdrkeie krqeenrkdr ekaaakfrey
 541 fpnfvgpeks kdilklrlye qqhgkclysg keinlvrlne kgyveidhal pfsrtwddsf
 601 nnkvlvlge nqnkgntqpy eyfngkdnsr ewqefkarve tsrfprskkq rillqkfdded
 661 gfkecnlndt ryvnrflcqf vadhilltgk gkrrvfafng qitnllrgfw glrkvraend
 721 rhhaldavvv acstvamqqk itrfrvykem nafdgktidk etgkvhlqkt hfpqpweffa
 781 qevmirvfgk pdgkpefeea dtpeklrtll aeklssrpea vheyvtplfv srpnrkmsg
 841 ahkdtlrsak rfvkhnekis vkrvwlteik ladlenmvny kngreiele alkaleayg
 901 gnakqafdpk dnpfykkggq lkvavrvekt qesgvllnkk naytiadngd mvrvdvfckv
 961 dkkgknqyfi vpiyawqvae nilpdidckg yriddsytfc fslhkydlia fqkdekskve
 1021 fayyincdss ngrfylawhd kgskeqqfri stqnlvliqk yqvnelgkei rpcrlkkrrp
 1081 vr

【0445】

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。

【0446】

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A

10

20

30

40

50

AV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。そのような実施形態では、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。

【0447】

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。いくつかの実施形態では、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。いくつかの実施形態では、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キロベース短いため、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。ある実施形態では、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。いくつかの実施形態において、標的遺伝子は、Cas9 PAM、例えば、5'-NGGに隣接し得る。他の実施形態では、他のCas9オルソログは異なる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。

【0448】

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。ある実施形態では、隣接切断は、PAMの(約)3塩基対上流であり得る。ある実施形態では、隣接切断は、PAMの(約)10塩基対上流であり得る。ある実施形態では、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。PAM配列に結合することができる例示的なSpCas9タンパク質の配列は、以下の通りである：

【0449】

例示的なPAM結合SpCas9のアミノ酸配列は以下の通りである：

MDKKYSIGLDIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLVESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFQILVQTYNQLFEENPINASGVDAKILSARLSKSRRLLENLIAQLPGKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDITYDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGASQEEFYKFIKPILEKMDGTEELLVKLNRDRLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPFLKDNREKIEKILTFRIPIYVGPLARGNSRFAMTRKSEETITPWNFEEVVDKGAQAQSFIERMTNFDKNLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGG

LSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRK
DFQFYKVRINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQ
EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLS
MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAK
VEKGKSKKLKSVKELLGITIMERSSEFKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGR
KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDEQKQLFVEQHKHYLDEII
EQISEFSKRVLADANLDKVL SAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTID
RKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGD.

【0450】

例示的なPAM結合SpCas9nのアミノ酸配列は以下の通りである：

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEA
TRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNI
VDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYALAHMIKFRGHFLIEGDLNPDNSDVD
KLFIQLVQTYNQLFEENPINASGVDAKILSARLSKSRLENLIAQLPGEKKNGLFGNLIAL
SLGLTPNFKSNFDLAEDAKLQLSKD TYDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDI
LRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDG
GASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ
EDFYFPLKDNREKIEKILTFRIPIYVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGAS
AQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKA
IVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLD
NEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWRLSRKLIN
GIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA
GSPAIIKKGILQTVKVVDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGI
KELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK
DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGG
LSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRK
DFQFYKVRINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQ
EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLS
MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAK
VEKGKSKKLKSVKELLGITIMERSSEFKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGR
KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDEQKQLFVEQHKHYLDEII
EQISEFSKRVLADANLDKVL SAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTID
RKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGD.

【0451】

例示的なPAM結合SpEQR Cas9のアミノ酸配列は以下のとおりである：

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEA
TRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNI
VDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYALAHMIKFRGHFLIEGDLNPDNSDVD
KLFIQLVQTYNQLFEENPINASGVDAKILSARLSKSRLENLIAQLPGEKKNGLFGNLIAL
SLGLTPNFKSNFDLAEDAKLQLSKD TYDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDI
LRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDG
GASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ
EDFYFPLKDNREKIEKILTFRIPIYVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGAS
AQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKA
IVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLD
NEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWRLSRKLIN
GIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA
GSPAIIKKGILQTVKVVDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGI
KELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK
DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGG

LSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS~~KL~~VSDFRK
DFQFYKVREINNYHHAHDAYLNAVVG~~TAL~~IKKYPKLESEFVYGDYKVYDVRKMI~~AK~~SEQ
EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLS
MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWD~~PK~~KYGGFESPTVAYSVLVVAK
VEKGKSKKLKSVKELLGITIMERS~~SSF~~EKNPIDFLEAKGYKEVKKDLIIKLPKYS~~LF~~ELENGR
KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLK~~GS~~PEDNEQKQLFVEQHKKHYLDEII
EQISEFSKRVILADANLDKVL~~SAY~~NKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTID
RKQYRSTKEVL~~DAT~~LIHQ~~SIT~~GLYETRIDLSQLGGD

上記配列において、D1134、R1334およびT1336から変異されてSpEQR Cas9を生じることができる残基E1134、Q1334およびR1336には太字で下線を付している。

【0452】

例示的なPAM結合SpVQR Cas9のアミノ酸配列は以下の通りである：

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEA
TRLKRTARRRYTRRKNRICYLQEIFS~~NEM~~AKVDD~~SFF~~HRL~~EE~~SFLVEEDKKHERHPIFGNI
VDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVD
KLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIAL
SLGLTPNFKSNFDLAEDAKLQLSKD~~TY~~DDDLNLLAQIGDQYADLFLAAKNLSDAILLSDI
LRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDG
GASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ
EDFY~~PFL~~KDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGAS
AQSFIERMTNFDKNLPNEKVL~~PK~~HSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKA
IVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLD
NEENEDILEDIVLTLT~~LF~~FEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLIN
GIRDKQSGKTILDFLKS~~DGF~~ANRNF~~MQL~~IHDDSLTFKEDIQKAQVSGQGD~~SL~~HEHIANLA
GSPAIKKGILQTVKVVD~~EL~~VKVMGRHKPENIVIE~~MARE~~NQTTQKGQKNSRERMKRIEEGI
KELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYD~~VD~~HIVPQSFLK
DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGG
LSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS~~KL~~VSDFRK
DFQFYKVREINNYHHAHDAYLNAVVG~~TAL~~IKKYPKLESEFVYGDYKVYDVRKMI~~AK~~SEQ
EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLS
MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWD~~PK~~KYGGFVSPTVAYSVLVVAK
VEKGKSKKLKSVKELLGITIMERS~~SSF~~EKNPIDFLEAKGYKEVKKDLIIKLPKYS~~LF~~ELENGR
KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLK~~GS~~PEDNEQKQLFVEQHKKHYLDEII
EQISEFSKRVILADANLDKVL~~SAY~~NKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTID
RKQYRSTKEVL~~DAT~~LIHQ~~SIT~~GLYETRIDLSQLGGD.

上記配列において、D1134、R1334、およびT1336から変異されてSpVQR Cas9を生じることができる残基V1134、Q1334、およびR1336には太字で下線を付している。

【0453】

例示的なPAM結合SpVRER Cas9のアミノ酸配列は以下の通りである：

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEA
TRLKRTARRRYTRRKNRICYLQEIFS~~NEM~~AKVDD~~SFF~~HRL~~EE~~SFLVEEDKKHERHPIFGNI
VDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVD
KLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIAL
SLGLTPNFKSNFDLAEDAKLQLSKD~~TY~~DDDLNLLAQIGDQYADLFLAAKNLSDAILLSDI
LRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDG
GASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ
EDFY~~PFL~~KDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGAS
AQSFIERMTNFDKNLPNEKVL~~PK~~HSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKA
IVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLD
NEENEDILEDIVLTLT~~LF~~FEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLIN

GIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA
GSPAIKKGILQTVKVVDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGI
KELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK
DDSIDNKVLTTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGG
LSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRK
DFQFYKVRINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIKSEQ
EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLS
MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFVSPTVAYSVLVVAK
VEKGKSKKLKSVKELLGITIMERSSEFKNPIDFLEAKGYKEVKKDLIKLPKYSLELENGR
KRMLASARELQKGNELALPSKYVNFLYLASHYEKLGKSPEDNEQKQLFVEQHKKHYLDEII
EQISEFSKRVLADANLDKVL SAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTID
RKEYRSTKEVL DATLIHQ SITGLYETRIDLSQLGGD.

10

上記配列において、D1134、G1217、R1334、およびT1336から変異されてSpVRER Cas9を生じることができる残基V1134、R1217、Q1334、およびR1336には太字で下線を付している。

【0454】

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SpyMacCas9 (SpyMacCas9d)、またはSpyMacCas9ニッカーゼ(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

【0455】

天然の5'-NAAN-3' PAM特異性を有するStreptococcus macacaeにおけるSpy Cas9の例示的なCas9 Aホモログの配列は当技術分野で知られており、例えば、Jakimoらによって記載されており (www.biorxiv.org/content/biorxiv/early/2018/09/27/429654.full.pdf)、下記に提供する。

【0456】

30

SpyMacCas9

mdkkysigldigtnsvgwavitddykvpskkfkvlgntrhsikknligallfgsgetae
atrlkrtarrtryrrknricylqeifsnemakvddsffhrleesflveedkkherhpifg
nivdevayhekyptiylhrkkldstdkadrlriylalahmikfrghfliedlnpdnsd
vdklfiqlvqiynqlfeenpinasrvdakailsarlsksrrlenliaqlpgekrnglfgn
lialslgltpnfksnfdlaedakqlskdtyddldnllaqigdqyadflaaknlsdai
llsdilrvnseitkaplsasmikrydehhqdltlkalvrqqpkykeiffdqskngya
gyidggasqeefykfikpilekmdgteellvklrnredllrkqrtfdngsiphqihlgel
ailrrqedfypflkdnrekiekiltfripyyvgplargnsrfawmtrkseetitpwnfee
vvdkgasaqsfiermtndknlpnekvlpkhslllyeyftvyneltkvkyvtegmrkpafl
sgeqkkaivdillfktnrkvtkqlkedyfkkiecdsveisgvedrfnaslgayhdllki
ikdkdldneenediledivltltlfedrgmieerlktyahlfddkvmkqlkrrrytgw
rlsrklingirdkqsgktildflksdgifanrnfqmqlihddsltfkediqaqvsgqghsl
heqianlagspaikkgilqtvkivdelvkvmghkpeniviamarenqttqkgqknsrerm
krieegikelgsqilkehpventqlqneklylyylqngrdmyvdqeldinrlsdydvdi
vpqsfikddsidnkvltrsdknrgksdnvpseevvkkmknywrqllnaklitqrkfdnlt
kaergglseldkagfikrqlvetrqitkhvaqildsrmtkydendklirevkvitlksk
lvsdfrkdfqfykvreinnyhahdaylnavvg talikkypklesefvygdykvvdvrkm
iakseqeigkatakyffysnimnffkteitlangeirkrplietngetgeivwdkgrdfa
tvrkvlsmpqvnivkkteiqtvqngglfddnpksplevtpsklvplkkelnpkkyggyq

40

50

kpttaypvlitdtkqlipisvmnkkqfeqnpvkflrdrgyqqvgkndfiklpkytlvdi
gdgikrlwasskeihkgnglvvskksqillyhahhldsdlsndylqnhnqqfdvlfneii
sfskkcklgkehiqkienvysnkknsasieelaesfikllgftqlgatspfnflgvklnq
kqykgkkdyilpctegtltirqsitglyetrvdlskiged.

【0457】

ある実施形態では、バリエントCas9タンパク質はH840A、P475A、W476A、N477A、D1125A、W1126A、D1218Aの変異を有し、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、いくつかの実施形態では、バリエントCas9タンパク質は、D10A、H840A、P475A、W476A、N477A、D1125A、W1126A、およびD1218A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）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。いくつかの実施形態では、バリエントCas9タンパク質がW476AおよびW1126A変異を有する場合、またはバリエントCas9タンパク質がP475A、W476A、N477A、D1125A、W1126A、およびD1218A変異を有する場合、バリエント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、このようなバリエント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、ある実施形態では、このようなバリエント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。

【0458】

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。

【0459】

[PAM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)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。

【0460】

[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。特定の理論に拘束されることは望まないが、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。

【0461】

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。

【0462】

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/RuvC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Cas9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。HypaCas9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。

【0463】

例示的な高忠実度Cas9を以下に示す。

【0464】

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している

DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEAT
 RLKRTARRRYTRRRKNRICYLQEIFSNE**MA**KVDD**S**FFHRL**E**ESFLVEEDKKHERHPIFGNIV
 DEVAYHEKYPTIYHLRKKLV**D**STDKADLRLLIYALAHMIKFRGHFLIEGDLNP**D**NSDV**D**K
 LFIQLVQTYNQ**L**FEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALS
 LGLTPNFKS**N**FDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLS**D**AILLSDIL
 RVNTEITKAPLSASMIKRYDEHHQD**L**TLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGG
 ASQEEFYKFIKPILEKMDGTEELLV**K**LNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQE
 DFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWN**F**EEVVDKGAS
 AQSFIERMTA**F**DKNLPNEKVL**P**KHSLLYEYFTVYNELTKVKYVTEGM**R**KPAFLSGEQKKA
 IVDLLFKTNRKVT**V**KQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKD**K**DFLD
 NEENEDILEDIVLT**L**TLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTG**W**GA**L**SRKLIN
 GIRDKQSGKTILDFLKSDGFANR**N**FMA**L**IHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAG
 SPAIKKGILQTVKV**V**DELVKVMGRHKPENIV**I**EMARE**N**QTTQKGQKNSRERMKRIE**E**GK
 ELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDV**D**HIVPQSFLK
 DDSIDNKV**L**TRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLT**K**AERGG
 LSELDKAGFIKRQLVETRAITKHVAQILD**S**RMNTKYDENDKLIREVKVITLKS**K**LVSDFRK
 DFQFYKVREINNYHHAHDAYLNAVVG**T**ALIKKYPKLESEFVYGDYKVYDVRK**M**IAKSEQ
 EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLS
 MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWD**P**KKYGGFDSPTVAYSVLV**V**AK
 VEKGKSKKLKSVKELLGITIMERS**S**FEKNPIDFLEAKGYKEVKKDLI**K**LPKYSLFELENGR
 KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKKHYLDEII
 EQISEFSKR**V**ILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTID
 RKRYTSTKEVL**D**ATLIHQ**S**ITGLYETRIDLSQLGGD

10

20

【0465】

[Cas9ドメインと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とを含む融合タンパク質]

本開示のいくつかの態様は、napDNAbp（例えばCas9ドメイン）と、1つ以上のアデノシンデアミナーゼドメイン、シチジンデアミナーゼドメイン、および／またはDNAグリコシラーゼドメインとを含む融合タンパク質を提供する。ある実施形態では、融合タンパク質はCas9ドメインとアデノシンデアミナーゼドメイン（例えばTadA***A**）を含む。このCas9ドメインは、本明細書に提供されるCas9ドメインまたはCas9タンパク質（例えばdCas9またはnCas9）のいずれでもあり得ることが理解されるべきである。いくつかの実施形態において、本明細書に提供されるCas9ドメインまたはCas9タンパク質（例えばdCas9またはnCas9）のいずれかが、本明細書に提供される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のいずれか（例えばTadA***A**）と融合され得る。例えば、限定するものではないが、いくつかの実施形態において、融合タンパク質は、以下の構造を含む：

30

NH₂-[シチジンデアミナーゼ]-[Cas9ドメイン]-[アデノシンデアミナーゼ]-COOH；
 NH₂-[アデノシンデアミナーゼ]-[Cas9ドメイン]-[シチジンデアミナーゼ]-COOH；
 NH₂-[アデノシンデアミナーゼ]-[シチジンデアミナーゼ]-[Cas9ドメイン]-COOH；
 NH₂-[シチジンデアミナーゼ]-[アデノシンデアミナーゼ]-[Cas9ドメイン]-COOH；
 NH₂-[Cas9ドメイン]-[アデノシンデアミナーゼ]-[シチジンデアミナーゼ]-COOH；
 NH₂-[Cas9ドメイン]-[シチジンデアミナーゼ]-[アデノシンデアミナーゼ]-COOH；
 NH₂-[アデノシンデアミナーゼ]-[Cas9ドメイン]-COOH；
 NH₂-[Cas9ドメイン]-[アデノシンデアミナーゼ]-COOH；
 NH₂-[シチジンデアミナーゼ]-[Cas9ドメイン]-COOH；または
 NH₂-[Cas9 domain]-[シチジンデアミナーゼ]-COOH。

40

【0466】

50

ある実施形態において、シチジンデアミナーゼ、脱塩基エディター、およびアデノシンデアミナーゼ、ならびにnapDNAbp（例えばCas9ドメイン）を含む融合タンパク質は、リンカー配列を含まない。ある実施形態において、リンカーが、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ドメインならびにnapDNAbpの間に存在する。いくつかの実施形態において、上記の一般的な構築で使用される「-」は、任意のリンカーの存在を示す。いくつかの実施形態において、シチジンデアミナーゼおよびアデノシンデアミナーゼおよびnapDNAbpは、本明細書に提供される任意のリンカーを介して融合される。例えば、いくつかの実施形態において、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およびnapDNAbpは、本明細書に提供されるリンカーのいずれかを介して融合される。

10

【0467】

[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、二部分 (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PKKKRKVEGADKRTADGSEFESPKKKRKV, KRTADGSEFESPKKKRKV, KRPAATKKAGQAKKKK, KKTELQTTNAENKTKKL, KRGINDRNFWRGENGRKTR, RKSGKIAAIVVKRPRKPKKKRKV, またはMDSLLMNRKFLYQFKNVRWAKGRRETYLCを含む。

20

30

【0468】

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、二部分NLSである。二部分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、二部分と呼ばれ、一部分 (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、PKKKRKVEGADKRTADGSEFESPKKKRKVである。

40

【0469】

いくつかの実施形態において、本発明の融合タンパク質は、リンカー配列を含まない。いくつかの実施形態において、1つ以上のドメインまたはタンパク質間にリンカー配列が存在する。いくつかの実施形態において、アデノシンデアミナーゼまたはシチジンデアミナーゼとCas9ドメインを有する例示的なCas9融合タンパク質の一般的な構造は、以下の構造のいずれか1つを含み、ここで、NLSは核局在化配列（例えば、本明細書で提供されるいずれかのNLS）であり、NH₂は、融合タンパク質のN末端であり、COOHは融合タンパク質のC末端である：

50

NH₂-NLS-[アデノシンデアミナーゼ]-[Cas9ドメイン]-COOH;
NH₂-NLS [Cas9ドメイン]-[アデノシンデアミナーゼ]-COOH;
NH₂-[アデノシンデアミナーゼ]-[Cas9ドメイン]-NLS-COOH;
NH₂-[Cas9ドメイン]-[アデノシンデアミナーゼ]-NLS-COOH;
NH₂-NLS-[シチジンデアミナーゼ]-[Cas9ドメイン]-COOH;
NH₂-NLS [Cas9ドメイン]-[シチジンデアミナーゼ]-COOH;
NH₂-[シチジンデアミナーゼ]-[Cas9ドメイン]-NLS-COOH; または
NH₂-[Cas9ドメイン]-[シチジンデアミナーゼ]-NLS-COOH。

【0470】

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、核外輸送配列などの輸送配列、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、チオレドキシントグ、S-タグ、ソフトタグ (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。

【0471】

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。

【0472】

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。

【0473】

[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。

【0474】

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を含む。本明細書で特に説明されているように、デアミナーゼドメインは、シトシンデアミナーゼまたはアデノシンデアミナーゼを含む。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。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（その全内容が参照により本明細書に組み込まれる）。

【0475】

[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み得る。塩基エディターのそ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り、アデニン（A）核酸塩基からグアニン（G）核酸塩基への編集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（すなわち、アミン基を除去）することができる。

【0476】

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る。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。

【0477】

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、I156F、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。

【0478】

アデノシンデアミナーゼは、任意の適切な生物(例えば大腸菌)に由来することができる。ある態様において、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。

【0479】

〔アデノシンデアミナーゼ〕

いくつかの実施形態では、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有するデアミナーゼドメインを含み得る。そのような塩基エディターのアデノシンデアミナーゼドメインは、Aを脱アミノ化してイノシン(I)(Gの塩基対形成特性を示す)を形成することにより、アデニン(A)核酸塩基からグアニン(G)核酸塩基への編集を促進させることができる。アデノシンデアミナーゼは、デオキシリボ核酸(DNA)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(すなわちアミン基を除去する)ことができる。

【0480】

いくつかの実施形態において、本明細書に提供されるアデノシンデアミナーゼは、アデニンを脱アミノ化することができる。いくつかの実施形態において、本明細書で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、アデニンデアミナーゼは、本明細書に提供される突然変異のいずれか(例えばecTadAにおける突然変異)に対応する一つ以上の突然変異を含む、天然に存在するアデノシンデアミナーゼである。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、任意の天然に存在するアデノシンデアミナーゼ(例えばecTadAに対する相同性を有するもの)において、本明細書に記載された突然変異のいずれか(例えば、ecTadAにおいて同定される突然変異のいずれか)に対応する突然変異を生成することができる。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。

【0481】

本発明は、増加した効率(>50~60%)および特異性を有するアデノシンデアミナーゼバリエントを提供する。特に、本明細書に記載のアデノシンデアミナーゼバリエントは、ポリヌクレオチド内の所望の塩基を編集する可能性がより高く、改変が意図されていない塩基(すなわち「バイスタンダー」)を編集する可能性がより低い。

【0482】

特定の実施形態では、TadAは、PCT/US2017/045381 (WO2018/027078) に記載されているTadAのいずれか1つであり、これはその全体が参照により本明細書に組み込まれる。

【0483】

いくつかの実施形態において、本発明の核酸塩基エディターは、以下の配列における改変を含むアデノシンデアミナーゼバリエーションである：

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVNLRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTD (TadA*7.1
0とも呼ばれる)。

10

【0484】

特定の実施形態では、融合タンパク質は、単一の（例えば、単量体として提供される）TadA*8バリエーションを含む。いくつかの実施形態において、TadA*8は、Cas9ニッカーゼに連結されている。いくつかの実施形態において、本発明の融合タンパク質は、野生型TadA (TadA(wt)) がTadA*8バリエーションに連結されたものをヘテロ二量体として含む。他の実施形態では、本発明の融合タンパク質は、TadA*7.10がTadA*8バリエーションに連結されたものをヘテロ二量体として含む。いくつかの実施形態において、塩基エディターは、TadA*8バリエーションモノマーを含むABE8である。いくつかの実施形態において、塩基エディターは、TadA*8バリエーションおよびTadA(wt)のヘテロダイマーを含むABE8である。いくつかの実施形態において、塩基エディターは、TadA*8バリエーションおよびTadA*7.10のヘテロダイマーを含むABE8である。いくつかの実施形態において、塩基エディターは、TadA*8バリエーションのヘテロダイマーを含むABE8である。いくつかの実施形態において、TadA*8バリエーションは、表7から選択される。いくつかの実施形態において、ABE8は、表7から選択される。関連する配列は、以下の通りである。

20

【0485】

いくつかの実施形態において、本発明の融合タンパク質は、野生型TadA (TadA(wt)) がTadAバリエーション（例えばTadA*7.10バリエーション）に連結されたヘテロ二量体を含む。関連する配列は次のとおりである。

【0486】

野生型TadA (TadA(wt)) または「TadA参照配列」)

30

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAH
AEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAGS
LMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRREQEIKAKKAQSSTD (配列番号
2)

【0487】

TadA*7.10:

MSEVEFSHEYW MRHALTLAKR ARDEREVPVG AVLVLNLRVI GEGWNRAIGL HDPT
AHAEIM ALRQGGLVMQ NYRLIDATLY VTTFEPCVMCA GAMIHSRIGR VVFGVRNAKT
GAAGSLMDVL HYPGMNHRVE ITEGILADECA ALLCYFFRM PRQVFNAQKK AQSSD

【0488】

40

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

50

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。

10

【0489】

ある実施形態において、TadAデアミナーゼは、全長大腸菌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、以下のアミノ酸配列を含む：

MRRAFITGVFFLSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNR
PIGRHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGA
RDAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAKKAQSS
TD.

【0490】

20

しかしながら、本出願において有用なさらなるアデノシンデアミナーゼが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るべきである。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（ADAT）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ホモログのアミノ酸配列は以下のものを含む：

【0491】

Staphylococcus aureus TadA:

MGSHMTNDIYFMTLAIEEAKKAAQLGEVPIGAIITKDDEVIARAHNLRETLQQPTAHAEH
IAIERAAKVLGSRWLEGCTLYVTLEPCVMCAGTIVMSRIPRVVYGADDPKGGCSGS LMN
LLQQSNFNHRAIVDKGVLKEACSTLLTTFKKNLRANKKSTN

30

【0492】

Bacillus subtilis TadA:

MTQDELYMKEAIKEAKKAEKGEVPIGAVLVINGEIIARAHNLRETEQRSIAHAEMLVIDE
ACKALGTWRLEGATLYVTLEPCPMCAGAVVLSRVEKVVFGAFDPKGGCSGTLMNLLQEE
RFNHQAEVVSGLVEEECGGMLSAFFRELKKKKKAARKNLSE

【0493】

Salmonella typhimurium (S. typhimurium) TadA:

MPPAFITGVTSLSDVELDHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNR
PIGRHDPTAHAEIMALRQGGLVLQNYRLDITLYVTLEPCVMCAGAMVHSRIGRVVFGA
RDAKTGAAGSLIDVLHHPGMNHRVEIEGVLRDECATLLSDFFRMRRQEIKALKKADRAE
GAGPAV

40

【0494】

Shewanella putrefaciens (S. putrefaciens) TadA:

MDEYWMQVAMQMAEKAEAAAGEVPVGAVLVKDGQQIATGYNLSISQHDPTAHAEILCLR
SAGKKLENYRLLDATLYITLEPCAMCAGAMVHSRIARVVYGARDEKTGAAGTVVNLLQH
PAFNHQVEVTSGVLAEACSAQLSRFFKRRRDEKKALKLAQRAQQGIE

【0495】

Haemophilus influenzae F3031 (H. influenzae) TadA:

MDAAKVRSEFDEKMMRYALELADKAEALGEIPVGAVLVDDARNIIGEGWNLSIVQSDPT
AHAEIIALRNGAKNIQNYRLNSTLYVTLEPCTMCAGAILHSRIKRLVFGASDYKTGAIGS

50

RFHFFDDYKMNHTLEITSGVLAEECSQKLSTFFQKRREEKKIEKALLKSLSDK

【0496】

Caulobacter crescentus (C. crescentus) TadA:

MRTDESEDQDHRMMLRALDAARAAAEAGETPVGAVILDPSTGEVIATAGNGPIAAHDPT
AHAEIAAMRAAAAKLGNYRLTDLTLVVTLEPCAMCAGAISHARIGRVVFGADDPKGGAV
VHGPKFFAQPTCHWRPEVTGGVLADESADLLRGFFRARRKAKI

【0497】

Geobacter sulfurreducens (G. sulfurreducens) TadA:

MSSLKKTPIRDDAYWMGKAIREAAKAAARDEVPIGAVIVRDGAVIGRGNLREGSNDPSA
HAEMIAIRQAARRSANWRLTGATLYVTLEPCLMCMGAILARLERVVFGCYDPKGGGAAGS
LYDLSADPRLNHQVRLSPGVCQEECGTMLSDFFRDLRRRKAKATPALFIDERKVPPEP

【0498】

E. coli TadA (ecTadA) の実施形態は以下のものを含む：

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVNLRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNRHVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTD

【0499】

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。

【0500】

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。

【0501】

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

10

20

30

40

50

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。

【0502】

本明細書に提供される突然変異（例えばTadA参照配列に基づくもの）のいずれもが、*E. coli* TadA (ecTadA)、*S. aureus* TadA (saTadA)、または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）などの他のアデノシンデアミナーゼに導入され得ることが理解されるべきである。さらなるデアミナーゼを同様にアラインメントして、本明細書中に提供されるように変異させ得る相同的アミノ酸残基を同定できることは、当業者には明らかであろう。したがって、TadA参照配列において同定された突然変異のいずれも、相同的なアミノ酸残基を有する他の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。また、本明細書中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おいて、個々にまたは任意の組合せで作製できることが理解されるべきである。TadAバリエーションにおけるアミノ酸置換は、TadA参照配列（配列番号2）における番号付けによるものであり、相同的アミノ酸残基を有する他の任意のTadAバリエーションの対応するアミノ酸置換または位置であり得ることが理解されるべきである。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、使用される具体的な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べきである。TadA参照配列と同一性を共有するTadAバリエーションで番号付けが異なる可能性があり、種ごとの配列の違いが番号付けに影響を与え得る。当業者は、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、任意の相同タンパク質およびそれぞれのコード化核酸中の対応する残基を同定することができる。

【0503】

ある態様において、アデノシンデアミナーゼは、TadA参照配列における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。

【0504】

ある態様において、アデノシンデアミナーゼは、TadA参照配列におけるA106X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、TadA参照配列におけるA106V突然変異、または別のアデノシンデアミナーゼ（例えば野生型TadAまたはecTadA）における対応する突然変異を含む。

【0505】

ある態様において、アデノシンデアミナーゼは、TadA参照配列におけるE155X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参照配列における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（例えばecTadA）における対応する突然変異を含む。

【0506】

ある態様において、アデノシンデアミナーゼは、TadA参照配列におけるD147X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参照配列におけるD147Y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む。

【0507】

ある態様において、アデノシンデアミナーゼは、TadA参照配列におけるA106X、E15

5X、またはD147X、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、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。

【0508】

例えば、アデノシンデアミナーゼは、TadA参照配列における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（例えばecTadA）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参照配列におけ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（例えばecTadA）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；および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。

【0509】

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参照配列における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（例えばecTadA）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（例えばecTadA）における1つ以上の対応する突然変異を含む。

【0510】

ある態様において、アデノシンデアミナーゼは、TadA参照配列における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（例えばecTadA）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参照配列における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（例えばecTadA）における1以上の対応する突然変異を含む。

【0511】

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参照配列における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（例えばecTadA）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参照配列における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（例えばecTadA）における1つ以上の対応する突然変異を含む。

【0512】

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*参照配列における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（例えば*ecTadA*）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*参照配列における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（例えば*ecTadA*）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*参照配列における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（例えば*ecTadA*）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

【0513】

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、R26X、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*参照配列における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（例えば*ecTadA*）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

30

【0514】

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*参照配列における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（例えば*ecTadA*）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*参照配列における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（例えば*ecTadA*）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*参照配列における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（例えば*ecTadA*）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*参照配列における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（例えば*ecTadA*）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*参照配列におけるH8Y、R26W、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（例えば*ecTadA*）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*参照配列にお

40

50

る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（例えばecTadA）における対応する突然変異（複数可）を含む。

【0515】

本明細書に提供される変異のいずれか、および任意のさらなる変異（例えばecTadAアミノ酸配列に基づくもの）を、任意の他のアデノシンデアミナーゼに導入することができる。本明細書に提供される変異のいずれも、TadA参照配列または別のアデノシンデアミナーゼ（例えばecTadA）において、個々に、または任意の組合せで作製することができる。

【0516】

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。

【0517】

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（例えばecTadA）において一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参照配列においてD108N、D108G、もしくはD108V突然変異、または別のアデノシンデアミナーゼ（例えばecTadA）において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参照配列における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（例えばecTadA）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参照配列における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（例えばecTadA）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参照配列における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（例えばecTadA）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参照配列におけるH8Y、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（例えばecTadA）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参照配列における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（例えばecTadA）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参照配列における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（例えばecTadA）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参照配列における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（例えばecTadA）における対応する突然変異を含む。

【0518】

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参照配列における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参照配列における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（例えばecTadA）における1つ以上の対応する突然変異を含む。

【0519】

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参照配列におけるL84F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む。

【0520】

ある態様において、アデノシンデアミナーゼは、TadA参照配列におけるH123X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、TadA参照配列におけるH123Y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む。

【0521】

ある態様において、アデノシンデアミナーゼは、TadA参照配列におけるI156X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、TadA参照配列におけるI156F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む。

【0522】

ある態様において、アデノシンデアミナーゼは、TadA参照配列において、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（例えばecTadA）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参照配列において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（例えばecTadA）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参照配列において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（例えばecTadA）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。

【0523】

ある態様において、アデノシンデアミナーゼは、TadA参照配列において、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（例えばecTadA）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参照配列において、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。

【0524】

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参照配列において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（例えばecTadA）における対応する突然変異もしくは突然変異を含む。

【0525】

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参照配列において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（例えばecTadA）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基準配列においてE25M、E25D、E25A、E25R、E25V、E25S、E25Y、R26G、R26N、R26Q、R26C

、R26L、R26K、R107P、R107K、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（例えばecTadA）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（例えばecTadA）における対応する突然変異のうちの1つ以上を含む。

【0526】

ある態様において、アデノシンデアミナーゼは、TadA参照配列におけるE25X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、TadA参照配列における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（例えばecTadA）における対応する突然変異を含む。

10

【0527】

ある態様において、アデノシンデアミナーゼは、TadA参照配列におけるR26X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。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参照配列における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（例えばecTadA）における対応する突然変異を含む。

20

【0528】

ある態様において、アデノシンデアミナーゼは、TadA参照配列におけるR107X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、TadA参照配列におけるR107P、R107K、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（例えばecTadA）における対応する突然変異を含む。

【0529】

ある態様において、アデノシンデアミナーゼは、TadA参照配列におけるA142X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、TadA参照配列におけるA142N、A142D、A142G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む。

30

【0530】

ある態様において、アデノシンデアミナーゼは、TadA参照配列におけるA143X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、TadA参照配列における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（例えばecTadA）における対応する突然変異を含む。

40

【0531】

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参照配列における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（例えばecTadA）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参照配列におけるH

50

36L、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（例えばecTadA）における1以上の対応する突然変異を含む。

【0532】

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参照配列におけるH36X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、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参照配列におけるH36L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む。

10

【0533】

ある態様において、アデノシンデアミナーゼは、TadA参照配列におけるN37X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、TadA参照配列における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（例えばecTadA）における対応する突然変異を含む。

【0534】

ある態様において、アデノシンデアミナーゼは、TadA参照配列におけるP48X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、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参照配列における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（例えばecTadA）における対応する突然変異を含む。

20

【0535】

ある態様において、アデノシンデアミナーゼは、TadA参照配列における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参照配列における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（例えばecTadA）における対応する突然変異を含む。

30

【0536】

ある態様において、アデノシンデアミナーゼは、TadA参照配列における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参照配列における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（例えばecTadA）における対応する突然変異を含む。

【0537】

ある態様において、アデノシンデアミナーゼは、TadA参照配列における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参照配列におけるK157N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む。

40

【0538】

ある態様において、アデノシンデアミナーゼは、TadA参照配列におけるP48X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、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参照配列における

50

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（例えばecTadA）における対応する突然変異を含む。

【0539】

ある態様において、アデノシンデアミナーゼは、TadA参照配列におけるA142X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、TadA参照配列におけるA142N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む。

【0540】

ある態様において、アデノシンデアミナーゼは、TadA参照配列におけるW23X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、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参照配列における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（例えばecTadA）における対応する突然変異を含む。

【0541】

ある態様において、アデノシンデアミナーゼは、TadA参照配列におけるR152X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、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参照配列における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（例えばecTadA）における対応する突然変異を含む。

【0542】

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、突然変異の各組合せは括弧内にある：

(A106V_D108N),
 (R107C_D108N),
 (H8Y_D108N_N127S_D147Y_Q154H),
 (H8Y_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),
 (R26G_L84F_A106V_R107H_D108N_H123Y_A142N_A143D_D147Y_E155V_I1

10

20

30

40

50

56F),
(E25G_R26G_L84F_A106V_R107H_D108N_H123Y_A142N_A143D_D147Y_E15
5V
_I156F), (E25D_R26G_L84F_A106V_R107K_D108N_H123Y_A142N_A143G_D1
47Y_E155V_
I156F),
(R26Q_L84F_A106V_D108N_H123Y_A142N_D147Y_E155V_I156F),
(E25M_R26G_L84F_A106V_R107P_D108N_H123Y_A142N_A143D_D147Y_E15
5V
_I156F),
(R26C_L84F_A106V_R107H_D108N_H123Y_A142N_D147Y_E155V_I156F), (L
84F_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_E15
5V
_I156F),
(R26G_L84F_A106V_R107H_D108N_H123Y_A142N_A143D_D147Y_E155V_I1
56F),
(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_K15
7N),
(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),
(H36L_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F_K157N), (H36L_L8
4F_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),
(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),

10

20

30

40

50

(H36L_R51L_L84F_A106V_D108N_H123Y_A142N_S146C_D147Y_E155V_I156F_K157N), (N37S_L84F_A106V_D108N_H123Y_A142N_D147Y_E155V_I156F_K161T),
(L84F_A106V_D108N_D147Y_E155V_I156F),
(R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N_K161T),
(L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K161T),
(L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N_K160E_K161T),
(L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N_K160E),
(R74Q_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_A142N),
(P48T_I49V_L84F_A106V_D108N_H123Y_A142N_D147Y_E155V_I156F_L157N),
(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_E155V_I156F_K157N),
(H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_A142N_D147Y_E155V_I156F_K157N),
(W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_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_E155V_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),

10

20

30

40

50

(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_R152P_E155V_I156F_K157N).

【0543】

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させる1つまたは複数の特徴を含む。例えば、本明細書で提供される融合タンパク質のいずれかは、ヌクレアーゼ活性が低減された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のいずれかは、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分子の一つの鎖を切断するCas9ドメインを有し得る。

【0544】

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*7.10である。いくつかの実施形態では、TadA*7.10は、少なくとも1つの改変を含む。特定の実施形態では、TadA*7.10は、以下の改変のうちの1つもしくは複数またはTadA*7.10へのさらなる改変を含む：Y147T、Y147R、Q154S、Y123H、V82S、T166R、およびQ154R。改変Y123Hは、本明細書ではH123Hとも呼ばれる（TadA*7.10における改変H123YがY123H（wt）に戻されたもの）。他の実施形態では、TadA*7.10は、以下の群から選択される改変の組み合わせを含む：Y147T + Q154R；Y147T + Q154S；Y147R + Q154S；V82S + Q154S；V82S + Y147R；V82S + Q154R；V82S + Y123H；I76Y + V82S；V82S + Y123H + Y147T；V82S + Y123H + Y147R；V82S + Y123H + Q154R；Y147R + Q154R + Y123H；Y147R + Q154R + I76Y；Y147R + Q154R + T166R；Y123H + Y147R + Q154R + I76Y；V82S + Y123H + Y147R + Q154R；および I76Y + V82S + Y123H + Y147R + Q154R。特定の実施形態において、アデノシンデアミナーゼバリエーションは、残基149、150、151、152、153、154、155、156、および157で始まるC末端の欠失を含む。

【0545】

いくつかの実施形態では、TadAバリエーションは、TadA7.10と比較して少なくとも1つの改変を含む。いくつかの実施形態において、TadAバリエーションは、野生型TadAと比較して少なくとも1つの改変を含む。TadAバリエーションにおけるアミノ酸改変は、TadA7.10または野生型TadAと比較して、本明細書に記載されるアミノ酸置換のいずれかが1つであり得る。いくつかの実施形態では、TadAバリエーション、例えば、TadA8は、アミノ酸位置23、26、36、37、48、49、51、72、84、87、105、108、123、125、142、145、147、152、155、16、157、161、またはそれらの任意の組み合わせにおいてアミノ酸改変を含む。いくつかの実施形態において、TadAバリエーションは、TadA7.10に対するアミノ酸改変V82Xを含み、ここで、Xは、V以外の任意のアミノ酸である。いくつかの実施形態において、TadAバリエーションは、TadA7.10に対するアミノ酸改変T166Xを含み、ここで、Xは、T以外の任意のアミノ酸である。いくつかの実施形態において、アミノ酸Xは、酸性アミノ酸、塩基性アミノ酸、または中性アミノ酸である。いくつかの実施形態において、TadAバリエーションは、TadA7.10に対するアミノ酸改変V82X、Y147X、Q154X、I76X、Y123X、R23X、L36X、A48X、L51X、F84X、V106X、N108X、Y123X、C146X、Y147X、P152X、Q154X、V155X、F156X、N157X、T166Xまたはそれらの任意の組み合わせを含み、ここで、Xは、TadA7.10におけるアミノ酸以

10

20

30

40

50

外の任意のアミノ酸である。いくつかの実施形態において、Xは、酸性アミノ酸、塩基性アミノ酸、または中性アミノ酸である。いくつかの実施形態において、Xは、アミノ酸を、TadA参照配列における野生型アミノ酸に戻す。

【0546】

他の実施形態では、本発明の塩基エディターは、以下の改変のうちの1つまたは複数を含むアデノシンデアミナーゼバリエーション（例えばTadA*8）を含むモノマーである：TadA*7.10または参照TadA配列と比較してY147T、Y147R、Q154S、Y123H、V82S、T166R、および／またはQ154R。他の実施形態において、アデノシンデアミナーゼバリエーション（TadA*8）は、以下の群から選択される改変の組み合わせを含む単量体である：Y147T + Q154R; Y147T + Q154S; Y147R + Q154S; V82S + Q154S; V82S + Y147R; V82S + Q154R; V82S + Y123H; I76Y + V82S; V82S + Y123H + Y147T; V82S + Y123H + Y147R; V82S + Y123H + Q154R; Y147R + Q154R + Y123H; Y147R + Q154R + I76Y; Y147R + Q154R + T166R; Y123H + Y147R + Q154R + I76Y; V82S + Y123H + Y147R + Q154R; およびI76Y + V82S + Y123H + Y147R + Q154R。他の実施形態では、塩基エディターは、野生型アデノシンデアミナーゼと、以下の改変のうちの1つまたは複数を含むアデノシンデアミナーゼバリエーション（例えばTadA*8）とを含むヘテロダイマーである：Y147T, Y147R, Q154S, Y123H, V82S, T166R, および／またはQ154R。他の実施形態では、塩基エディターは、TadA*7.10ドメインと、以下の群から選択される改変の組み合わせを含むアデノシンデアミナーゼバリエーションドメイン（例えばTadA*8）とを含むヘテロダイマーである：Y147T + Q154R; Y147T + Q154S; Y147R + Q154S; V82S + Q154S; V82S + Y147R; V82S + Q154R; V82S + Y123H; I76Y + V82S; V82S + Y123H + Y147T; V82S + Y123H + Y147R; V82S + Y123H + Q154R; Y147R + Q154R + Y123H; Y147R + Q154R + I76Y; Y147R + Q154R + T166R; Y123H + Y147R + Q154R + I76Y; V82S + Y123H + Y147R + Q154R; およびI76Y + V82S + Y123H + Y147R + Q154R。

【0547】

一実施形態では、アデノシンデアミナーゼは、アデノシンデアミナーゼ活性を有する以下の配列またはその断片を含むか、または本質的にそれからなる、TadA*8である：

MSEVEFSHEYWMRHALTLAKRARDEREVPVGVAVLVLNNRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNHRVEITEGILADECAALLCTFFRMPRQVFNAQKKAQSSTD

【0548】

いくつかの実施形態では、TadA*8は切り詰められている。いくつかの実施形態において、切り詰め型TadA*8は、全長TadA*8と比較して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いている。いくつかの実施形態において、切り詰め型TadA*8は、全長TadA*8と比較して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いている。いくつかの実施形態において、アデノシンデアミナーゼバリエーションは、完全長のTadA*8である。

【0549】

いくつかの実施形態において、TadA*8は、TadA*8.1, TadA*8.2, TadA*8.3, TadA*8.4, TadA*8.5, TadA*8.6, TadA*8.7, TadA*8.8, TadA*8.9, TadA*8.10, TadA*8.11, TadA*8.12, TadA*8.13, TadA*8.14, TadA*8.15, TadA*8.16, TadA*8.17, TadA*8.18, TadA*8.19, TadA*8.20, TadA*8.21, TadA*8.22, TadA*8.23, TadA*8.24である。

【0550】

一実施形態では、本発明の融合タンパク質は、野生型TadAが本明細書に記載のアデノシンデアミナーゼバリエーション（例えばTadA*8）に連結されそしてそれがCas9ニッカーゼに連結されているものを含む。特定の実施形態において、融合タンパク質は単一のTadA*8ドメインを（すなわちモノマーとして提供されて）含む。別の実施形態では、塩基エディ

ターはTadA*8とTadA(wt)とを含み、これらはヘテロダイマーを形成することができる。典型的な配列は以下の通りである。

TadA(wt):

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAH
AEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAGS
LMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAKKAQSSTD

TadA*7.10:

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNRRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTD

TadA*8:

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNRRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNHRVEITEGILADECAALLCTFFRMPRQVFNAQKKAQSSTD.

【0551】

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。

【0552】

特定の実施形態では、TadA*8は、太字で示されている以下の位置のいずれかに1つまたは複数の変異を含む。他の実施形態では、TadA*8は、下線を引いて示された位置のいずれかに1つまたは複数の変異を含む：

MSEVEFSHEY WMRHALTLAK RARDEREVPV GAVLVLNRRV IGEGWNRAIG⁵⁰
LHDPTAHAEI MALRQGGLVM QNYRLIDATL Y¹⁰⁰VTFEPCVMC AGAMIHSRIG¹⁰⁰
RVVFGVRNAK TGAAGSLMDV LH¹⁵⁰YPGMNHRV EITEGILADE CAALLCY¹⁵⁰FFR¹⁵⁰
MPR¹⁵⁰QVFNAQK KAQSS¹⁵⁰TD

【0553】

例えば、TadA*8は、アミノ酸位置82および/または166における改変(例えばV82S、T166R)を、単独で、または以下のうちのいずれか1つ以上と組み合わせて含む：Y147T,

Y147R, Q154S, Y123H, および/またはQ154R。特定の実施形態では、改変の組み合わせが、以下のものからなる群から選択される：Y147T + Q154R; Y147T + Q154S; Y147R + Q154S; V82S + Q154S; V82S + Y147R; V82S + Q154R; V82S + Y123H; I76Y + V82S; V82S + Y123H + Y147T; V82S + Y123H + Y147R; V82S + Y123H + Q154R; Y147R + Q154R + Y123H; Y147R + Q154R + I76Y; Y147R + Q154R + T166R; Y123H + Y147R + Q154R + I76Y; V82S + Y123H + Y147R + Q154R; および I76Y + V82S + Y123H + Y147R + Q154R。

【0554】

ある実施形態では、アデノシンデアミナーゼは、アデノシンデアミナーゼ活性を有する以下の配列またはその断片を含むか、または本質的にそれからなる、TadA*8である：

MSEVEFSHEY WMRHALTLAK RARDEREVPV GAVLVNLRV IGEGWNRAIG LHDPT
AHAEI MALRQGGLVM QNYRLIDATL YVTFEPCVMC AGAMIHSRIG
RVVFGVRNAK TGAAGSLMDV LHYPGMNHRV EITEGILADE CAALLCTFFR
MPRQVFNAQK KAQSSTD

【0555】

いくつかの実施形態では、TadA*8は切り詰められている (truncated)。いくつかの実施形態において、切り詰め型TadA*8は、全長TadA*8と比較して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いている。いくつかの実施形態において、切り詰め型TadA*8は、全長TadA*8と比較して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いている。いくつかの実施形態において、アデノシンデアミナーゼバリエーションは、完全長のTadA*8である。

【0556】

一実施形態では、本発明の融合タンパク質は、野生型TadAが本明細書に記載のアデノシンデアミナーゼバリエーション (例えばTadA*8) に連結されそしてそれがCas9ニッカーゼに連結されているものを含む。特定の実施形態において、融合タンパク質は単一のTadA*8ドメインを (すなわちモノマーとして提供されて) 含む。別の実施形態では、塩基エディターはTadA*8とTadA(wt)とを含み、これらはヘテロダイマーを形成することができる。

【0557】

いくつかの実施形態において、アデノシンデアミナーゼアレル (例えばTadAアレル) の合成ライブラリーを利用して、改変された塩基編集効率および/または特異性を有するアデノシン塩基エディターを生成することができる。いくつかの実施形態において、合成ライブラリーから生成されたアデノシン塩基エディターは、より高い塩基編集効率および/または特異性を含む。いくつかの実施形態において、合成ライブラリーから生成されたアデノシン塩基エディターは、野生型TadAを有するアデノシン塩基エディターと比較して、増加した塩基編集効率、増加した塩基編集特異性、低減されたオフターゲット編集、低減されたバイズタンダー編集、低減されたインデル形成、および/または低減された目的外編集を示す。いくつかの実施形態において、合成ライブラリーから生成されたアデノシン塩基エディターは、TadA*7.10を有するアデノシン塩基エディターと比較して、増加した塩基編集効率、増加した塩基編集特異性、低減されたオフターゲット編集、低減されたバイズタンダー編集、低減されたインデル形成、および/または低減された目的外編集を示す。いくつかの実施形態において、合成ライブラリーは、ABEのランダム化されたTadA部分を含む。いくつかの実施形態では、合成ライブラリーは、TadAの各位置に20個すべての正準的なアミノ酸の置換を含む。いくつかの実施形態において、合成ライブラリーは、ライブラリー構成員あたり1〜2ヌクレオチド置換突然変異の平均頻度を含む。いくつかの実施形態において、合成ライブラリーは、TadA*7.10に見られるバックグラウンド突然変異を含む。

【0558】

いくつかの実施形態では、本明細書に記載の塩基編集システムは、TadAがCas9に挿入された (inserted) ABEを含む。TadAがCas9に挿入された関連ABEの配列を提供する。

【 0 5 5 9 】

101 Cas9 TadAins 1015

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENP
INASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
NFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAI
LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI
FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLR
KQRTFDNGSIPHQIHLGELHAILRRQEDFYPPFLKDNREKIEKILTFRIPY
YVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDK
NLPNEKVLPHKSHLLYEFYTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD
LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQ
LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDD
SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIAKKGILQTVKVVDELVKV
MGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEELGELGSQILKEHP
VENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDD
SIDNKVLTRSDKNRGKSDNVPSEEVVKMKMKNYWRQLLNAKLITQRKFDNL
TKAERGGLSELDKAGFIKRLVETRQITKHVAQILDSRMNTKYDENDKLI
REVKVITLKSCLVSDFRKDFQFYKVINNYHHAHDAYLNAVVGTAIIKK
YPKLESEFVYGDYKVGSSGSETPGTSESATPESSGSEVEFSHEYWMRHAL
TLAKRARDEREVPVGAVLVNLRVIGEGWNRAIGLHDPTAHAEIMALRQG
GLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGS
LMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSS
DYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIE
TNGETGEIVWDKGRDFATVRKVLSPQVNVIVKKTEVQTGGFSKESILPKR
NSDKLIARKKDWDPPKYGGFDSPTVAYSVLVVAKEKGSKKLKS VKELL
GITIMERSSEFKNPIDFLEAKGYKEVKKDLIILPKYSLFELENGRKRML
ASAGELQKGNELALPSKYVNFYLYLASHYEKLKGS PEDNEQKQLFVEQHKH
YLDEIIEQISEFSKRVILADANLDKVL SAYNKH RDKPIREQAENIIHLFT
LTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQ SITGLYETRIDL SQ
LGGD

10

20

30

【 0 5 6 0 】

102 Cas9 TadAins 1022

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENP
INASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
NFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAI
LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI
FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLR
KQRTFDNGSIPHQIHLGELHAILRRQEDFYPPFLKDNREKIEKILTFRIPY
YVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDK
NLPNEKVLPHKSHLLYEFYTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD
LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQ
LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDD

40

50

SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIAKKGILQTVKKVDELVKV
MGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEELGSGILKEHP
VENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVHIVPQSFLKDD
SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL
TKAERGGELSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI
REVKVITLKSCLVSDFRKDFQFYKVRINNYHHAHDAYLNAVVGTAIIKK
YPKLESEFVYGDYKVYDVRKMIGSSGSETPGTSESATPESSGSEVEFSHE
YWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAE
IMALRQGGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNA
KTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQ
KKAQSSTDAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIE
TNGETGEIVWDKGRDFATVRKVLSPQVNIKKTEVQTGGFSKESILPKR
NSDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVEKGKSKKLKSVKELL
GITIMERSSEFKNPIDFLEAKGYKEVKKDLIIKLPKYSLELENGRKRML
ASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDEQKQLFVEQHKH
YLDEIIEQISEFSKRVILADANLDKVL SAYNKH RDKPIREQAENIIHLFT
LTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDL SQ
LGGD

[0561]

103 Cas9 TadAins 1029

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDD SFFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDFIQLVQTYNQLFEENP
INASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
NFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAI
LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI
FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLR
KQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPY
YVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGASAQSFIERM TNFDK
NLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD
LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMKQ
LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNF MQLIHDD
SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIAKKGILQTVKKVDELVKV
MGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEELGSGILKEHP
VENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVHIVPQSFLKDD
SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL
TKAERGGELSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI
REVKVITLKSCLVSDFRKDFQFYKVRINNYHHAHDAYLNAVVGTAIIKK
YPKLESEFVYGDYKVYDVRKMIKSEQEIGSSGSETPGTSESATPESSGS
EVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLH
DPTAHAEIMALRQGGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRV
VFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRM P
RQVFNAQKKAQSSTDGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIE
TNGETGEIVWDKGRDFATVRKVLSPQVNIKKTEVQTGGFSKESILPKR
NSDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVEKGKSKKLKSVKELL
GITIMERSSEFKNPIDFLEAKGYKEVKKDLIIKLPKYSLELENGRKRML
ASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDEQKQLFVEQHKH
YLDEIIEQISEFSKRVILADANLDKVL SAYNKH RDKPIREQAENIIHLFT

10

20

30

40

50

LTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDL SQ
LGGD

[0 5 6 2]

103 Cas9 TadAins 1040

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDDSSFFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
LRLIYLALAHMIKFRGHFLIEGDLNPDNSDV DKLFIQLVQTYNQLFEENP
INASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
NFKSNFDLAEDAKLQLSKD TYDDDLDNLLAQIGDQYADLFLAAKNLSDAI
LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI
FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLR
KQRTFDNGSIPHQIHLGELHAILRRQEDFY PFLKDNREKIEKILTFRIPY
YVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDK
NLPNEKVLPKHSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD
LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQ
LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFL KSDGFANRNFMQLIHDD
SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPA IKKGILQTVKVVDDELVKV
MGRHKPENIVIAMARENQTTQKGQKNSRERM KRIE EGikelGSQILKEHP
VENTQLQNEKLYLYYLQNGRDMYVDQELDIN RLSDYDVDHIVPQSFLKDD
SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNY WRQLLNAKLITQRKFDNL
TKAERGGELSELDKAGFIKRQLVETRQITKHVA QILDSRMNTKYDENDKLI
REVKVITLKSCLVSDFRKDFQFYK VREINNYHHAHDAYLNAVVG TALIKK
YPKLESEFVYGDYKVYDVRKMIAKSEQEIGKAT AKYFFYSGSSGSETPGT
SESATPESSGSEVEFSHEYWMRHALTLAKRARD ERVPVGAVLVLNRRVI
GEGWNRAIGLHDPTAHAEIMALRQGGLVMQNY RLIDATLYVTFEPCVMCA
GAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYP GNMNRVEITEGILADEC
AALLCYFFRMPRQVFNAQKKAQSSTDNIMNFFK TEITLANGEIRKRPLIE
TNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKR
NSDKLIARKKDWDPKKYGGFDSPTVAYSVLV VAKVEKGKSKKLKSVKELL
GITIMERSSF EKNPIDFLEAKGYKEVKKDLI IKLPKYSLFELNGRKRML
ASAGELQKGNELALPSKYVNFLYLASHYEKLKG SPEDNEQKQLFVEQHKH
YLDEIIEQISEFSKRVILADANLDKVL SAYNKH RDKPIREQAENIIHLFT
LTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDL SQ
LGGD

10

20

30

[0 5 6 3]

105 Cas9 TadAins 1068

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDDSSFFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
LRLIYLALAHMIKFRGHFLIEGDLNPDNSDV DKLFIQLVQTYNQLFEENP
INASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
NFKSNFDLAEDAKLQLSKD TYDDDLDNLLAQIGDQYADLFLAAKNLSDAI
LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI
FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLR
KQRTFDNGSIPHQIHLGELHAILRRQEDFY PFLKDNREKIEKILTFRIPY
YVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDK
NLPNEKVLPKHSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD
LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI

40

50

IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQ
LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDD
SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDELVKV
MGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEDIKELGSQILKEHP
VENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDD
SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL
TKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI
REVKVITLKSCLVSDFRKDFQFYKVRINNYHHAHDAYLNAVVGTAIIKK
YPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEI
TLANGEIRKRPLIETNGEGSSGSETPGTSESATPESSGSEVEFSHEYWMR
HALTLAKRARDEREVPVGAVLVNNRVIGEGWNRAIGLHDPTAHAEIMAL
RQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGA
AGSLMDVLHYPGMNHHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQ
SSTDTEIVWDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKR
NSDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVEKGKSKKLKSVKELL
GITIMERSSEFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLEENGRKRML
ASAGELQKGNELALPSKYVNFLYLASHYEKLKGGSPEDNEQKQLFVEQHKH
YLDEIIEQISEFSKRVILADANLDKVL SAYNKH RDKPIREQAENIIHLFT
LTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDL SQ
LGGD

10

20

[0 5 6 4]

106 Cas9 TadAins 1247

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDD SFFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
LRLIYLALAHMIKFRGHFLIEGDLNPDNSDV DKLFIQLVQTYNQLFEENP
INASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
NFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAI
LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI
FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLR
KQRTFDNGSIPHQIHLGELHAILRRQEDFY PFLKDNREKIEKILTFRIPY
YVGPLARGNSRFAWMTRKSEETITPWNFE EVVDKGASAQSFIERMTNFDK
NLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD
LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQ
LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDD
SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDELVKV
MGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEDIKELGSQILKEHP
VENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDD
SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL
TKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI
REVKVITLKSCLVSDFRKDFQFYKVRINNYHHAHDAYLNAVVGTAIIKK
YPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEI
TLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIVKKTEV
QTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVE
KGKSKKLKSVKELLGITIMERSSEFEKNPIDFLEAKGYKEVKKDLIIKLPK
YSLFEENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGGSS
GSETPGTSESATPESSGSEVEFSHEYWMRHALTLAKRARDEREVPVGAVL
VLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTF
EPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHHRVEITE

30

40

50

GILADECAALLCYFFRMPRQVFNAQKKAQSSTDSPEDNEQKQLFVEQHKH
YLDEIIEQISEFSKRVLADANLDKVL SAYNKH RDKPIREQAENIIHLFT
LTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDL SQ
LGGD

[0 5 6 5]

107 Cas9 TadAins 1054

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDD SFFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEEENP
INASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
NFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAI
LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI
FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLR
KQRTFDNGSIPHQIHLGELHAILRRQEDFY PFLKDNREKIEKILTFRIPY
YVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDK
NLPNEKVLPKHSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD
LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQ
LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMLIHDD
SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKV
MGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHP
VENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDD
SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL
TKAERGGSELDAKAGFIKRLVETRQITKHVAQILDSRMNTKYDENDKLI
REVKVITLKSCLVSDFRKDFQFYKVINNYHHAHDAYLNAVVG TALIKK
YPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEI
TLANGSSGSETPGTSESATPESSGSEVEFSHEYWMRHALTLAKRARDERE
VPVGAVLV LNNRVIGEGWNRAIGLHDPTAHAEIMALRQGG LVMQNYRLID
ATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMN
HRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTDGEIRKRPLIE
TNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKR
NSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEKGKSKKLKSVKELL
GITIMERSSSFENPIDFLEAKGYKEVKKDLI IKLPKYSLFELENGRKRML
ASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKH
YLDEIIEQISEFSKRVLADANLDKVL SAYNKH RDKPIREQAENIIHLFT
LTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDL SQ
LGGD

[0 5 6 6]

108 Cas9 TadAins 1026

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDD SFFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEEENP
INASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
NFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAI
LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI
FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLR
KQRTFDNGSIPHQIHLGELHAILRRQEDFY PFLKDNREKIEKILTFRIPY
YVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDK

NLPNEKVLPKHSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD
LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQ
LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMLIHDD
SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIAKKGILQTVKVVDELVKV
MGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEELGKELGSQILKEHP
VENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDD
SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL
TKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI
REVKVITLKSCLVSDFRKDFQFYKVRINNYYHHAHDAYLNAVVGTAIIKK
YPKLESEFVYGDYKVDYVRKMIKSEGSSGSETPGTSESATPESSGSEVE
FSHEYWMRHALTLAKRARDEREVPVGAVLVNLRVIGEGWNRAIGLHDPT
AHAEIMALRQGGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFG
VRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQV
FNAQKKAQSSTDQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIE
TNGETGEIVWDKGRDFATVRKVLSPQVNIIVKKTEVQTGGFSKESILPKR
NSDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVEKGKSKKLKSVKELL
GITIMERSSEFKNPIDFLEAKGYKEVKKDLIKLPKYSLELENGRKRML
ASAGELQKGNELALPSKYVNFYLYLASHYEKLKGPEDNEQKQLFVEQHKH
YLDEIIEQISEFSKRVILADANLDKVL SAYNKHDKPIREQAENIIHLFT
LTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDL SQ
LGGD

[0 5 6 7]

109 Cas9 TadAins 768

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENP
INASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
NFKSNFDLAEDAKLQLSKD TYDDDLNLLAQIGDQYADLFLAAKNLSDAI
LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI
FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLR
KQRTFDNGSIPHQIHLGELHAILRRQEDFYFPFLKDNREKIEKILTFRIPY
YVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDK
NLPNEKVLPKHSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD
LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQ
LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMLIHDD
SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIAKKGILQTVKVVDELVKV
MGRHKPENIVIAMARENQSSGSETPGTSESATPESSGSEVEFSHEYWMR
HALTLAKRARDEREVPVGAVLVNLRVIGEGWNRAIGLHDPTAHAEIMAL
RQGGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGA
AGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRTTQKGQKNSR
ERMKRIEELGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL
DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKK
MKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQIT
KHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVR
NNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVDYVRKMIKSE
IGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGR
DFATVRKVLSPQVNIIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDP

10

20

30

40

50

KKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSSFENP
IDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELAL
PSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSK
RVILADANLDKVL SAYNKH RD KPIREQAENIIHLFTLTNLGAPAAFKYFD
TTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGD

[0 5 6 8]

110.1 Cas9 TadAins 1250

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDD SFFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENP
INASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
NFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAI
LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI
FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLR
KQRTFDNGSIPHQIHLGELHAILRRQEDFY PFLKDNREKIEKILTFRIPY
YVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDK
NLPNEKVL PKHSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD
LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQ
LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNF MQLIHDD
SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKV
MGRHKPENIVIEMARENQTTQKGQKNSRERMKRIE EG IKELGSQILKEHP
VENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYD VDHIVPQSFLKDD
SIDNKVLTRSDKNRGKSDNVPSEEVVKMKMKNYWRQLLNAKLITQRKFDNL
TKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI
REVKVITLKS KLVSDFRKDFQFYK VREINNYHHAHDAYLNAVVG TALIKK
YPKLESEFVYGDYKVYDVRKMI AKSEQEIGKATAKYFFYSNIMNFFKTEI
TLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEV
QTGGFSKESILPKRNSDKLIARKKDWD PKKYGGFDSPTVAYSVLVVAKVE
KGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDLIIKLPK
YSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPG
SSGSETPGTSESATPESSGSEVEFSHEYWMRHALTLAKRARDEREVPVGA
VLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYV
TFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEI
TEGILADECAALLCYFFRMPREDNEQKQLFVEQHKHYLDEIIEQISEFSK
RVILADANLDKVL SAYNKH RD KPIREQAENIIHLFTLTNLGAPAAFKYFD
TTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGD

[0 5 6 9]

110.2 Cas9 TadAins 1250

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDD SFFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENP
INASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
NFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAI
LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI
FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLR
KQRTFDNGSIPHQIHLGELHAILRRQEDFY PFLKDNREKIEKILTFRIPY
YVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDK

NLPNEKVLPKHSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD
LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQ
LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDD
SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIAKKGILQTVKVVDELVKV
MGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEDIKELGSQILKEHP
VENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDD
SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL
TKAERGGELSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI
REVKVITLKSCLVSDFRKDFQFYKVRINNYHHAHDAYLNAVVGTAIIKK
YPKLESEFVYGDYKVDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEI
TLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIKKTEV
QTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVE
KGKSKKLKSVKELLGITIMERSSEFEKNPIDFLEAKGYKEVKKDLIIKLPK
YSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPG
SSGSSGSETPGTSESATPESGSEVEFSHEYWMRHALTLAKRARDEREVP
VGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGGLVMQNYRLIDAT
LYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHR
VEITEGILADECAALLCYFFRMPREDNEQKQLFVEQHKKHYLDEIEEISE
FSKRVLADANLDKVL SAYNKHDKPIREQAENIIHLFTLTNLGAPAAFK
YFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGD
[0 5 7 0]

10

20

110.3 Cas9 TadAins 1250

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENP
INASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
NFKSNFDLAEDAKLQLSKD TYDDDLDNLLAQIGDQYADLFLAAKNLSDAI
LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI
FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLR
KQRTFDNGSIPHQIHLGELHAILRRQEDFYPPFLKDNREKIEKILTFRIPY
YVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERM TNFDK
NLPNEKVLPKHSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD
LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQ
LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDD
SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIAKKGILQTVKVVDELVKV
MGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEDIKELGSQILKEHP
VENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDD
SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL
TKAERGGELSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI
REVKVITLKSCLVSDFRKDFQFYKVRINNYHHAHDAYLNAVVGTAIIKK
YPKLESEFVYGDYKVDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEI
TLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIKKTEV
QTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVE
KGKSKKLKSVKELLGITIMERSSEFEKNPIDFLEAKGYKEVKKDLIIKLPK
YSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPG
SSGSSGSETPGTSESATPESGSSSGSEVEFSHEYWMRHALTLAKRARDER
EVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGGLVMQNYRLI

30

40

50

DATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGM
NHRVEITEGILADECAALLCYFFRMPREDNEQKQLFVEQHKHYLDEIEQ
ISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLFTLTNLGAPA
AFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGD
[0 5 7 1]

110.4 Cas9 TadAins 1250

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDD SFFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
LRLIYLALAHMIKFRGHFLIEGDLNPDNSDV DKLFIQLVQTYNQLFEENP
INASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
NFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAI
LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI
FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVVLNREDLLR
KQRTFDNGSIPHQIHLGELHAILRRQEDFY PFLKDNREKIEKILTFRIPY
YVGPLARGNSRFAWMTRKSEETITPWNFE EVVDKGASAQSFIERMTNFDK
NLPNEKVLPKHSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD
LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQ
LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFL KSDGFANRNFMLIHDD
SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPA IKKGILQTVKVVDELVKV
MGRHKPENIVIEMARENQTTQKGQKNSRERM KRIE EG I KELGSQILKEHP
VENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDV DHIVPQSFLKDD
SIDNKVLTRSDKNRGKSDNVPSEEVVKMKMKNYWRQLLNAKLITQRKFDNL
TKAERGGSEL D KAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI
REVKVITLKS KLVSDFRKDFQFYK VREINNYHHAH DAYLNAVVG TALIKK
YPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEI
TLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEV
QTGGFSKESILPKRNSDKLIARKKDWD PKKYGGFDSPTVAYSVLVAKVE
KGKSKKLKSVKELLGITIMERS SFEKNPIDFLEAKGYKEVKKD LIIKLPK
YSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPG
SSGSSGSETPGTSESATPESGSSSGSEVEFSHEYW MRHALTLAKRARDER
EVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLI
DATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGM
NHRVEITEGILADECAALLCYFFRMRREDNEQKQLFVEQHKHYLDEIEQ
ISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLFTLTNLGAPA
AFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGD
[0 5 7 2]

110.5 Cas9 TadAins 1249

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDD SFFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
LRLIYLALAHMIKFRGHFLIEGDLNPDNSDV DKLFIQLVQTYNQLFEENP
INASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
NFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAI
LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI
FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVVLNREDLLR
KQRTFDNGSIPHQIHLGELHAILRRQEDFY PFLKDNREKIEKILTFRIPY
YVGPLARGNSRFAWMTRKSEETITPWNFE EVVDKGASAQSFIERMTNFDK
NLPNEKVLPKHSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD

10

20

30

40

50

LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQ
LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDD
SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKV
MGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEDIKELGSQILKEHP
VENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDD
SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL
TKAERGGELSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI
REVKVITLKSCLVSDFRKDFQFYKVRINNYYHHAHDAYLNAVVGTAIIKK
YPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEI
TLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIKKTEV
QTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVE
KGKSKKLKSVKELLGITIMERSSEFEKNPIDFLEAGYKEVKKDLIIKLPK
YSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKSGS
SGSSGSETPGTSESATPESGSSSGSEVEFSHEYWMRHALTLAKRARDERE
VPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLID
ATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMN
HRVEITEGILADECAALLCYFFRMRRPEDNEQKQLFVEQHKHYLDEIIEQ
ISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLFTLTNLGAPA
AFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGD

10

20

[0 5 7 3]

110.5 Cas9 TadAins delta 59-66 1250

MDKKYSIGLAIGTNSVGWAVITDEYKVPSSKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENP
INASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTP
NFKSNFDLAEDAKLQLSKD TYDDDLNLLAQIGDQYADLFLAAKNLSDAI
LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPKEYKEI
FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLR
KQRTFDNGSIPHQIHLGELHAILRRQEDFYFPFLKDNREKIEKILTFRIPY
YVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDK
NLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD
LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQ
LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDD
SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKV
MGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEDIKELGSQILKEHP
VENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDD
SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL
TKAERGGELSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI
REVKVITLKSCLVSDFRKDFQFYKVRINNYYHHAHDAYLNAVVGTAIIKK
YPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEI
TLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIKKTEV
QTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVE
KGKSKKLKSVKELLGITIMERSSEFEKNPIDFLEAGYKEVKKDLIIKLPK
YSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKSGPG
SSGSSGSETPGTSESATPESGSSSGSEVEFSHEYWMRHALTLAKRARDERE
VPVGAVLVLNNRVIGEGWNRAHAEIMALRQGGLVMQNYRLIDATLYVTFE
PCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEG

30

40

50

ILADECAALLCYFFRMPRQVFNAQKKAQSSTDEDNEQKQLFVEQHKHYLD
EIEQISEFSKRVLADANLDKVL SAYNKHRDKPIREQAENIIHLFTLTN
LGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGG
D

【0 5 7 4】

110.6 Cas9 TadAins 1251

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDD SFFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
LRLIYLALAHMIKFRGHFLIEGDLNPDNSDV DKLFIQLVQTYNQLFEENP
INASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
NFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAI
LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI
FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLR
KQRTFDNGSIPHQIHLGELHAILRRQEDFY PFLKDNREKIEKILTFRIPY
YVGPLARGNSRFAWMTRKSEETITPWNFE EVVDKGASAQSFIERMTNFDK
NLPNEKVLPHKSHLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD
LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQ
LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFL KSDGFANRNF MQLIHDD
SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPA IKKGILQTVKVVDELVKV
MGRHKPENIVIEMARENQTTQKGQKNSRERM KRIEEGIKELGSQILKEHP
VENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYD VDHIVPQSFLKDD
SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL
TKAERGGSEL DKA GFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI
REVKVITLKS KLVSDFRKDFQFYK VREINNYHHAH DAYLNAVVG TALIKK
YPKLESEFVYGDYKVYDVRKMIKSEQEIGKAT AKYFFYSNIMNFFKTEI
TLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEV
QTGGFSKESILPKRNSDKLIARKKDWD PKKYGGFDSPTVAYSVLVAKVE
KGKSKKLKSVKELLGITIMERS SFEKNPIDFLEAKGYKEVKKD LIIKLPK
YSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPE
GSSGSSGSETPGTSESATPESGSSSGSEVEFSHEYW MRHALTLAKRARDE
REVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGG LVMQNYRL
IDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPG
MNRHVEITEGILADECAALLCYFFRMRRDNEQKQLFVEQHKHYLDEIEQ
ISEFSKRVLADANLDKVL SAYNKHRDKPIREQAENIIHLFTLTNLGAPA
AFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGD

【0 5 7 5】

110.7 Cas9 TadAins 1252

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDD SFFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
LRLIYLALAHMIKFRGHFLIEGDLNPDNSDV DKLFIQLVQTYNQLFEENP
INASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
NFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAI
LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI
FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLR
KQRTFDNGSIPHQIHLGELHAILRRQEDFY PFLKDNREKIEKILTFRIPY
YVGPLARGNSRFAWMTRKSEETITPWNFE EVVDKGASAQSFIERMTNFDK
NLPNEKVLPHKSHLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD

LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQ
LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDD
SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIAKKGILQTVKVVDELVKV
MGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEELGELGSQILKEHP
VENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDD
SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL
TKAERGGELSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI
REVKVITLKSCLVSDFRKDFQFYKVRINNYYHHAHDAYLNAVVGTAIIKK
YPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEI
TLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIKKTEV
QTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVE
KGKSKKLKSVKELLGITIMERSSEFEKNPIDFLEAKGYKEVKKDLIIKLPK
YSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPE
DGSSGSSGSETPGTSESATPESGSSSGSEVEFSHEYWMRHALTLAKRARD
EREVPVGAVLVNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGGLVMQNYR
LIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYP
GMNHRVEITEGILADECAALLCYFFRMRRNEQKQLFVEQHKHYLDEIEEQ
ISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLFTLTNLGAPA
AFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGD

10

20

[0576]

110.8 Cas9 TadAins delta 59-66 C-truncate 1250

MDKKYSIGLAIGTNSVGWAVITDEYKVPSSKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENP
INASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTP
NFKSNFDLAEDAKLQLSKD TYDDDLNLLAQIGDQYADLFLAAKNLSDAI
LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPKEYKEI
FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLR
KQRTFDNGSIPHQIHLGELHAILRRQEDFYFPFLKDNREKIEKILTFRIPY
YVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDK
NLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD
LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQ
LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDD
SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIAKKGILQTVKVVDELVKV
MGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEELGELGSQILKEHP
VENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDD
SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL
TKAERGGELSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI
REVKVITLKSCLVSDFRKDFQFYKVRINNYYHHAHDAYLNAVVGTAIIKK
YPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEI
TLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIKKTEV
QTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVE
KGKSKKLKSVKELLGITIMERSSEFEKNPIDFLEAKGYKEVKKDLIIKLPK
YSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPE
SSGSETPGTSESATPESGSSSGSEVEFSHEYWMRHALTLAKRARDEREVPVGA
VLVLNNRVIGEGWNRAHAEIMALRQGGGLVMQNYRLIDATLYVTFEPCVMC
AGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADE

30

40

50

CAALLCYFFRMPRQEDNEQKQLFVEQHKHYLDEIEQISEFSKRVLADA
NLDKVL SAYNKH RD KPIREQAENIIHLFTLTNLGAPAAF KYFDTTIDRK R
YTSTKEVL DATLIHQ SITGLYETRIDLSQLGGD

[0 5 7 7]

111.1 Cas9 TadAins 997

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDD SFFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
LRLIYLALAHMIKFRGHFLIEGDLNPDNSDV DKLFIQLVQTYNQLFEENP
INASGVDAKAILSARLSKSRREN LIAQLPGEKKNGLFGNLIALSLGLTP
NFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAI
LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI
FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLV KLNREDLLR
KQRTFDNGSIPHQIHLGELHAILRRQEDFY PFLKDNREKIEKILTFRIPY
YVGPLARGNSRFAWMTRKSEETITPWNFE EVVDKGASAQSFIERMTNFDK
NLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD
LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQ
LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKS DGFANRNF MQLIHDD
SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPA IKKGILQTVKVVDELVKV
MGRHKPENIVIEMARENQTTQKGQKNSRERM KRIE EGikelGSQILKEHP
VENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDV DHIVPQSFLKDD
SIDNKVLTRSDKNRGKSDNVPSEEVVKMKMKNYWRQLLNAKLITQRKFDNL
TKAERGGLSELDKAGFIK RQLVETRQITKHVAQILDSRMNTKYDENDKLI
REVKVITLKS KLVSDFRKDFQFYK VREINNYHHAH DAYLNAVVG TALSHE
YWMRHALTLAKRARDEREVPVGAVLV LNNRVIGEGWNRAIGLHDPTAHAE
IMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNA
KTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQ
KKAQSSTDGSSGSETPGTSESATPESSGIKKYPKLESEFVYGDYKVYDVR
KMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET
GEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKL
IARKKDWDPKKYGGFDSPTVAYSVLV VAKVEKGSKKLKSVKELLGITIM
ERSSFEKNPIDFLEAKGYKEVKKDLI IKLPKYSLFEL ENGRKRMLASAGE
LQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEI
IEQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLFTLTNLG
APAAF KYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGD

[0 5 7 8]

111.2 Cas9 TadAins 997

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDD SFFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
LRLIYLALAHMIKFRGHFLIEGDLNPDNSDV DKLFIQLVQTYNQLFEENP
INASGVDAKAILSARLSKSRREN LIAQLPGEKKNGLFGNLIALSLGLTP
NFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAI
LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI
FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLV KLNREDLLR
KQRTFDNGSIPHQIHLGELHAILRRQEDFY PFLKDNREKIEKILTFRIPY
YVGPLARGNSRFAWMTRKSEETITPWNFE EVVDKGASAQSFIERMTNFDK
NLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD
LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI

IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQ
LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMLIHDD
SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDELVKV
MGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEELGSGILKEHP
VENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDD
SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL
TKAERGGELSELDKAGFIKRLVETRQITKHVAQILDSRMNTKYDENDKLI
REVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAHSHE
YWMRHALTLAKRARDEREVPVGAVLVLNRRVIGEGWNRAIGLHDPTAHAE
IMALRQGGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNA
KTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQ
KKAQSSTDGSSGSSGSETPGTSESATPESSGSSIKKYPKLESEFVYGDY
KVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLI
ETNGETGEIVWDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPK
RNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEVGKSKKLKSVKEL
LGITIMERSSEKPNIDFLEAKGYKEVKKDLIIKLPKYSLELENGRKRML
LASAGELQKGNELALPSKYVNFLYLASHYEKLGSPEDNEQKQLFVEQHK
HYLDEIIEQISEFSKRVILADANLDKVL SAYNKH RDKPIREQAENIIHLF
TLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLS
QLGGD

10

20

[0579]

112 delta HNH TadA

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDFLFIQLVQTYNQLFEENP
INASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
NFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAI
LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPKEYKEI
FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLR
KQRTFDNGSIPHQIHLGELHAILRRQEDFYFPFLKDNREKIEKILTFRIPY
YVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDK
NLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD
LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQ
LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMLIHDD
SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDELVKV
MGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEELGSGEVEFSHE
YWMRHALTLAKRARDEREVPVGAVLVLNRRVIGEGWNRAIGLHDPTAHAE
IMALRQGGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNA
KTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQ
KKAQSSTDGGLSELKAGFIKRLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAH
IKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFK
TEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIVKK
TEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVA
KVEKGKSKKLKSVKELLGITIMERSSEKPNIDFLEAKGYKEVKKDLIIK
LPKYSLELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLG
SPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVL SAYNKH
RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATL

30

40

50

IHQSITGLYETRIDLSQLGGD

[0580]

113 N-term single TadA helix trunc 165-end

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIG
 LHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTTFEPCVMCAGAMIHSRIG
 RVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFR
 MPRSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSV
 GWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKR
 TARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERH
 PIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYALAHMIKFRG
 HFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARL
 SKSRRLLENLIAQLPGEKKNGLFGNLIASLGLTPNFKS NFDAEDAKLQL
 SKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAP
 LSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGG
 ASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHL
 GELHAILRRQEDFYFPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMT
 RKSEETITPWNFEEVVDKGASAQSFIERM TNFDKNLPNEKVLPHKSHLLYE
 YFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE
 DYFKKIECFDSVEISGVEDRFNASLGTYHDLKIIKDKDFLDNEENEDIL
 EDIVLTTLTFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKL
 INGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQ
 GDSLHEHIANLAGSPAIIKKGILQTVKVVDELVKVMGRHKPENIVIAMARE
 NQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYL
 QNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGK
 SDNVPSEEVVKMKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGF
 IKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDF
 RKDFQFYKVINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVY
 DVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETN
 GETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNS
 DKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEKGKSKLKS VKELLGI
 TIMERSSEKNPIDFLEAKGYKEVKKDLI IKLPKYSLFELENGRKRMLAS
 AGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYL
 DEIEQISEFSKRVLADANLDKVL SAYNKH RDKPIREQAENIIHLFTLT
 NLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLG
 GD

[0581]

114 N-term single TadA helix trunc 165-end delta 59-65

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRTAH
 AEIMALRQGGLVMQNYRLIDATLYVTTFEPCVMCAGAMIHSRIGRVVFGVR
 NAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRSGGS
 SGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSV GWAVITD
 EYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKR TARRRYT
 RRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNIV
 DEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYALAHMIKFRGHFLIEGD
 LNPDSNDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRL
 NLIAQLPGEKKNGLFGNLIASLGLTPNFKS NFDAEDAKLQLSKDTYDD
 DDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIK
 RYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFY
 KFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAIL
 RRQEDFYFPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETI

TPWNFEVVDK GASAQSFIERMTNFDKNLPNEKVL PKHSLLEYFTVYNE
LTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIE
CFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTL
TLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDK
QSGKTILDFLKSDGFANRNF MQLIHDDSLTFKEDIQKAQVSGQGDSLHEH
IANLAGSPAIIKKGILQTVKVVDELVKVMGRHKPENIVIAMARENQTTQKG
QKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMY
VDQELDINRLSDYD VDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSE
EVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVE
TRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFY
KVREINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIA
KSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIV
WDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARK
KDWDPPKKYGGFDSPTVAYSVLV VAKVEKGSKKLKSVKELLGITIMERS
FEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENG RKRMLASAGELQKG
NELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKKHYLDEIIEQI
SEFSKRVLADANLDKVL SAYNKH RDKPIREQAENIIHLFTLTNLGAPAA
FKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGD

[0582]

115.1 Cas9 TadAins1004

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDDSFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
LRLIYLALAHMIKFRGHFLIEGDLNPDNSDV DKLFIQLVQTYNQLFEENP
INASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
NFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAI
LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI
FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLR
KQRTFDNGSIPHQIHLGELHAILRRQEDFY PFLKDNREKIEKILTFRIPY
YVGPLARGNSRFAWMTRKSEETITPWNFEVVDK GASAQSFIERMTNFDK
NLPNEKVL PKHSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD
LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
IKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQ
LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNF MQLIHDD
SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDELVKV
MGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHP
VENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYD VDHIVPQSFLKDD
SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL
TKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI
REVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG TALIKK
YPKGSSGSETPGTSESATPESSGSEVEFSHEYWMRHALTLAKRARDEREV
PVGAVLV LNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDA
TLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNH
RVEITEGILADECAALLCYFFRMPRQLESEFVYGDYKVYDVRKMIAKSEQ
EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKG
RDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWD
PKKYGGFDSPTVAYSVLV VAKVEKGSKKLKSVKELLGITIMERS
SFEPIDFLEAKGYKEVKKDLIIKLPKYSLFELENG RKRMLASAGELQKGNELA
LPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKKHYLDEIIEQISEFS
KRVILADANLDKVL SAYNKH RDKPIREQAENIIHLFTLTNLGAPAAF KYF

10

20

30

40

50

DTTIDRKRYTSTKEVLDATLIHQSI TGLYETRIDLSQLGGD

[0 5 8 3]

115.2 Cas9 TadAins1005

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
 LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDDSFHR
 LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
 LRLIYLALAHMIKFRGHFLIEGDLNPDNSDV DKLFIQLVQTYNQLFEENP
 INASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
 NFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAI
 LLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI
 FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLV KLNREDLLR
 KQRTFDNGSIPHQIHLGELHAILRRQEDFY PFLKDNREKIEKILTFRIPY
 YVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDK
 NLPNEKVLPHKSHLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD
 LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
 IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQ
 LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFL KSDGFANRNF MQLIHDD
 SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPA IKKGILQTVKVVDELVKV
 MGRHKPENIVIEMARENQTTQKGQKNSRERM KRIE EGikelGSQILKEHP
 VENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDV DHIVPQSFLKDD
 SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQL LNAKLITQRKFDNL
 TKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI
 REVKVITLKS KLVSDFRKDFQFYK VREINNYHHAH DAYLNAVVG TALIKK
 YPKLGSSGSETPGTSESATPESSGSEVEFSHEYW MRHALTLAKRARDERE
 VPGAVLV LNNRVIGEGWNRAIGLHDPTAHAEIM ALRQGGLVMQNYRLID
 ATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMN
 HRVEITEGILADECAALLCYFFRM PRQESEFVYGDYKVYDVRKMIKSEQ
 EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKG
 RDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDW
 PKKYGGFDSPTVAYSVLVAKVEKGSKKLKSVKELLGITIMERS SFEKN
 PIDFLEAKGYKEVKKDLIILPKYSLFELENGRKRMLASAGELQKGNELA
 LPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQH KHYLDEIIEQISEFS
 KRVILADANLDKVLSAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYF
 DTTIDRKRYTSTKEVLDATLIHQSI TGLYETRIDLSQLGGD

[0 5 8 4]

115.3 Cas9 TadAins1006

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
 LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDDSFHR
 LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
 LRLIYLALAHMIKFRGHFLIEGDLNPDNSDV DKLFIQLVQTYNQLFEENP
 INASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
 NFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAI
 LLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI
 FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLV KLNREDLLR
 KQRTFDNGSIPHQIHLGELHAILRRQEDFY PFLKDNREKIEKILTFRIPY
 YVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDK
 NLPNEKVLPHKSHLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD
 LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
 IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQ
 LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFL KSDGFANRNF MQLIHDD

SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKKVDELVKV
MGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEDIKELGSQILKEHP
VENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDD
SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL
TKAERGGLSELDKAGFIKRLVETRQITKHVAQILDSRMNTKYDENDKLI
REVKVITLKSCLVSDFRKDFQFYKVRINNYHHAHDAYLNAVVGITALIKK
YPKLEGGSSGSETPGTSESATPESSGSEVEFSHEYWMRHALTLAKRARDER
EVPVGAVLVLNRRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLI
DATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGM
NHRVEITEGILADECAALLCYFFRMPRQSEFVYGDYKVVYDVRKMIKSEQ
EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKG
RDFATVRKVLSTMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWD
PKKYGGFDSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSSFEN
PIDFLEAKGYKEVKKDLIKLPKYSLEFENGRKRMLASAGELQKGNELA
LPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFS
KRVILADANLDKVL SAYNKHDKPIREQAENIIHLFTLTNLGAPAAFKYF
DTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGD

[0 5 8 5]

115.4 Cas9 TadAins1007

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENP
INASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
NFKSNFDLAEDAKLQLSKD TYDDDLNLLAQIGDQYADLFLAAKNLSDAI
LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI
FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKNREDLLR
KQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPY
YVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDK
NLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD
LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMKQ
LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDFANRNF MQLIHDD
SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKKVDELVKV
MGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEDIKELGSQILKEHP
VENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDD
SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL
TKAERGGLSELDKAGFIKRLVETRQITKHVAQILDSRMNTKYDENDKLI
REVKVITLKSCLVSDFRKDFQFYKVRINNYHHAHDAYLNAVVGITALIKK
YPKLESGSSGSETPGTSESATPESSGSEVEFSHEYWMRHALTLAKRARDE
REVPVGAVLVLNRRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRL
IDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPG
MNHRVEITEGILADECAALLCYFFRMPRQSEFVYGDYKVVYDVRKMIKSEQ
EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKG
RDFATVRKVLSTMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWD
PKKYGGFDSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSSFEN
PIDFLEAKGYKEVKKDLIKLPKYSLEFENGRKRMLASAGELQKGNELA
LPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFS
KRVILADANLDKVL SAYNKHDKPIREQAENIIHLFTLTNLGAPAAFKYF
DTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGD

10

20

30

40

50

【0586】

116.1 Cas9 TadAins C-term truncate2 792

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
 LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFHR
 LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
 LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENP
 INASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
 NFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAI
 LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI
 FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLR
 KQRTFDNGSIPHQIHLGELHAILRRQEDFYPPFLKDNREKIEKILTFRIPY
 YVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDK
 NLPNEKVLPHKSHLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD
 LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
 IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQ
 LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDFANRNFQMQLIHDD
 SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKV
 MGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEELGEGSSGSETP
 GTSESATPESSGSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVNLR
 VIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVM
 CAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILAD
 ECAALLCYFFRMPRQSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQE
 LDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVK
 KMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRLVETRQI
 TKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVRE
 INNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQ
 EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKG
 RDFATVRKVLSPQVNVIVKKTEVQTGGFSKESILPKRNSDKLIARKKDW
 PKKYGGFDSPTVAYSVLVAKVEKGSKKLKSVKELLGITIMERSSSFEN
 PIDFLEAKGYKEVKKDLIIKLPKYSLFELNGRKRMLASAGELQKGNELA
 LPSKYVNFYLYLASHYEKLKGSPEDEQKQLFVEQHKKHYLDEIIEQISEFS
 KRVILADANLDKVL SAYNKHDKPIREQAENIIHLFTLTNLGAPAAFKYF
 DTTIDRKRYTSTKEVLDATLIHQSTGLYETRIDLSQLGGD

10

20

30

【0587】

116.2 Cas9 TadAins C-term truncate2 791

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
 LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFHR
 LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
 LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENP
 INASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
 NFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAI
 LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI
 FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLR
 KQRTFDNGSIPHQIHLGELHAILRRQEDFYPPFLKDNREKIEKILTFRIPY
 YVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDK
 NLPNEKVLPHKSHLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD
 LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
 IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQ
 LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDFANRNFQMQLIHDD
 SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKV

40

50

MGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGIKELGSSGSETPG
TSESATPESSGSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRV
IGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMC
AGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADE
CAALLCYFFRMPRQGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQE
LDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVK
KMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI
TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVRE
INNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQ
EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKG
RDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWD
PKKYGGFDSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSSFEN
PIDFLEAKGYKEVKKDLIIKLPKYSLFELNGRKRMLASAGELQKGNELA
LPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFS
KRVLADANLDKVL SAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYF
DTTIDRKRYTSTKEVLDATLIHQ SITGLYETRIDLSQLGGD

[0 5 8 8]

116.3 Cas9 TadAins C-term truncate2 790

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEFSNEMAKVDDSFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENP
INASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
NFKSNFDLAEDAKLQLSKD TYDDDLNLLAQIGDQYADLFLAAKNLSDAI
LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI
FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLR
KQRTFDNGSIPHQIHLGELHAILRRQEDFY PFLKDNREKIEKILTFRIPY
YVGPLARGNSRFAWMTRKSEETITPWNFE EVVDKGASAQSFIERM TNFDK
NLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD
LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQ
LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKS DGFANRNF MQLIHDD
SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPA IKKGILQTVKVVDELVKV
MGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGIKEGSSGSETPGT
SESATPESSGSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVI
GEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCA
GAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADE C
AALLCYFFRMPRQLGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQE
LDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVK
KMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI
TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVRE
INNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQ
EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKG
RDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWD
PKKYGGFDSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSSFEN
PIDFLEAKGYKEVKKDLIIKLPKYSLFELNGRKRMLASAGELQKGNELA
LPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFS
KRVLADANLDKVL SAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYF
DTTIDRKRYTSTKEVLDATLIHQ SITGLYETRIDLSQLGGD

[0 5 8 9]

10

20

30

40

50

117 Cas9 delta 1017-1069

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENP
INASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
NFKSNFDLAEDAKLQLSKD TYDDDLNLLAQIGDQYADLFLAAKNLSDAI
LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI
FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLR
KQRTFDNGSIPHQIHLGELHAILRRQEDFYPPFLKDNREKIEKILTFRIPY
YVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDK
NLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD
LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQ
LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDD
SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDELVKV
MGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHP
VENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDD
SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL
TKAERGGELSELDKAGFIKQQLVETRQITKHVAQILDSRMNTKYDENDKLI
REVKVITLKSCLVSDFRKDFQFYKVINNYHHAHDAYLNAVVGTAIIKK
YPKLESEFVYGDYKVYSSGSEVEFSHEYWMRHALTLAKRARDEREVPVGA
VLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGGLVMQNYRLIDATLYV
TFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEI
TEGILADECAALLCYFFRMPRQVFNAQKKAQSSTDGEIVWDKGRDFATVR
KVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGF
DSPTVAYSVLVVAKEGKSKKLKSVKELLGITIMERSSEKKNPIDFLEA
KGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVN
FLYLASHYEKLGKSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLAD
ANLDKVL SAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRK
RYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGD

[0590]

118 Cas9 TadA-CP116ins 1067

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENP
INASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
NFKSNFDLAEDAKLQLSKD TYDDDLNLLAQIGDQYADLFLAAKNLSDAI
LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI
FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLR
KQRTFDNGSIPHQIHLGELHAILRRQEDFYPPFLKDNREKIEKILTFRIPY
YVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDK
NLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD
LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQ
LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDD
SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDELVKV
MGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHP
VENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDD

SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL
TKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI
REVKVITLKSCLVSDFRKDFQFYKVINNYHHAHDAYLNAVVGITALIKK
YPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEI
TLANGEIRKRPLIETNMNHRVEITEGILADECAALLCYFFRMPRQVFNAQ
KKAQSSTDGSSGSETPGTSESATPESSGSEVEFSHEYWMRHALTLAKRAR
DEREVPVGAVLVLNRRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNY
RLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHY
PGGETGEIVWDKGRDFATVRKVLSPQVNIKKTEVQTGGFSKESILPKR
NSDKLIARKKDWDPPKYGGFDSPTVAYSVLVAKVEKGKSKKLKSVKELL
GITIMERSSEFKNPIDFLEAKGYKEVKKDLIKLPKYSLELENGRKRML
ASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDEQKQLFVEQHKH
YLDEIIEQISEFSKRVILADANLDKVL SAYNKH RDKPIREQAENIIHLFT
LTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDL SQ
LGGD

10

[0 5 9 1]

119 Cas9 TadAins 701

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDD SFFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENP
INASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
NFKSNFDLAEDAKLQLSKD TYDDDLNLLAQIGDQYADLFLAAKNLSDAI
LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI
FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLR
KQRTFDNGSIPHQIHLGELHAILRRQEDFY PFLKDNREKIEKILTFRIPY
YVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDK
NLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD
LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQ
LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKS DGFANRNF MQLIHDD
SGSSGSETPGTSESATPESSGSEVEFSHEYWMRHALTLAKRARDEREVPV
GAVLVLNRRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATL
YVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRV
EITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTD LTFKEDIQKAQVS
GQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMA
RENQTTQKGQKNSRERMKRIE EGikelGSQILKEHPVENTQLQNEKLYLY
YLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNR
GKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKA
GFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVS
DFRKDFQFYKVINNYHHAHDAYLNAVVGITALIKKYPKLESEFVYGDYK
VYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIE
TNGETGEIVWDKGRDFATVRKVLSPQVNIKKTEVQTGGFSKESILPKR
NSDKLIARKKDWDPPKYGGFDSPTVAYSVLVAKVEKGKSKKLKSVKELL
GITIMERSSEFKNPIDFLEAKGYKEVKKDLIKLPKYSLELENGRKRML
ASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDEQKQLFVEQHKH
YLDEIIEQISEFSKRVILADANLDKVL SAYNKH RDKPIREQAENIIHLFT
LTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDL SQ
LGGD

20

30

40

[0 5 9 2]

50

120 Cas9 TadACP136ins 1248

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENP
INASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
NFKSNFDLAEDAKLQLSKDITYDDDLNLLAQIGDQYADLFLAAKNLSDAI
LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI
FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLR
KQRTFDNGSIPHQIHLGELHAILRRQEDFYPPFLKDNREKIEKILTFRIPIY
YVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMNTNFDK
NLPNEKVLPHKSHLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD
LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQ
LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDD
SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDELVKV
MGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHP
VENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDD
SIDNKVLTRSDKNRGKSDNVPSEEVVKMKNYWRQLLNAKLITQRKFDNL
TKAERGGSELDAKAGFIKRLVETRQITKHVAQILDSRMNTKYDENDKLI
REVKVITLKSCLVSDFRKDFQFYKVINNYHHAHDAYLNAVVGTAIIKK
YPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEI
TLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNVKKTEV
QTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVE
KGKSKKLKSVKELLGITIMERSSEKNPIDFLEAKGYKEVKKDLIIKLPK
YSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSMD
HRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTDGSSGSETPGT
SESATPESSGSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVI
GEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCA
GAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGPEDNEQKQLFVEQHKH
YLDEIIEQISEFSKRVLADANLDKVL SAYNKH RDKPIREQAENIIHLFT
LTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDL SQ
LGGD

[0593]

121 Cas9 TadACP136ins 1052

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENP
INASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
NFKSNFDLAEDAKLQLSKDITYDDDLNLLAQIGDQYADLFLAAKNLSDAI
LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI
FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLR
KQRTFDNGSIPHQIHLGELHAILRRQEDFYPPFLKDNREKIEKILTFRIPIY
YVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMNTNFDK
NLPNEKVLPHKSHLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD
LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQ
LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDD
SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDELVKV

MGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEDIKELGSQILKEHP
VENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYVDHIVPQSFLKDD
SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL
TKAERGGELSELDKAGFIKRLVETRQITKHVAQILDSRMNTKYDENDKLI
REVKVITLKSCLVSDFRKDFQFYKVRINNYYHHAHDAYLNAVVGTAIIKK
YPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEI
TLAMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTDGSSGS
ETPGTSESATPESSGSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVL
NNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEP
CVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGNGEIRKRPLIE
TNGETGEIVWDKGRDFATVRKVLSPQVNVKKTEVQTGGFSKESILPKR
NSDKLIARKKDWDPPKYGGFDSPTVAYSVLVVAKEKGKSKKLKSVKELL
GITIMERSSEKPNIDFLEAKGYKEVKKDLIKLPKYSLELENGRKRML
ASAGELQKGNELALPSKYVNFLYLASHYEKLGSPEDNEQKQLFVEQHKH
YLDEIIEQISEFSKRVILADANLDKVL SAYNKH RD KPIREQAENIIHLFT
LTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDL SQ
LGGD

[0 5 9 4]

122 Cas9 TadACP136ins 1041

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENP
INASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
NFKSNFDLAEDAKLQLSKD TYDDDLNLLAQIGDQYADLFLAAKNLSDAI
LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI
FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLR
KQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRI PY
YVGPLARGNSRFAWMTRKSEETITPWNFEVVVDKGASAQSFIERM TNFDK
NLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD
LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQ
LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKS DGFANRNF MQLIHDD
SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDDELVKV
MGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEDIKELGSQILKEHP
VENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYVDHIVPQSFLKDD
SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL
TKAERGGELSELDKAGFIKRLVETRQITKHVAQILDSRMNTKYDENDKLI
REVKVITLKSCLVSDFRKDFQFYKVRINNYYHHAHDAYLNAVVGTAIIKK
YPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNHRVEITEG
ILADECAALLCYFFRMPRQVFNAQKKAQSSTDGSSGSETPGTSESATPES
SGSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAI
GLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEP CVMCAGAMIHSRI
GRVVFGVRNAKTGAAGSLMDVLHYPGNIMNFFKTEITLANGEIRKRPLIE
TNGETGEIVWDKGRDFATVRKVLSPQVNVKKTEVQTGGFSKESILPKR
NSDKLIARKKDWDPPKYGGFDSPTVAYSVLVVAKEKGKSKKLKSVKELL
GITIMERSSEKPNIDFLEAKGYKEVKKDLIKLPKYSLELENGRKRML
ASAGELQKGNELALPSKYVNFLYLASHYEKLGSPEDNEQKQLFVEQHKH
YLDEIIEQISEFSKRVILADANLDKVL SAYNKH RD KPIREQAENIIHLFT
LTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDL SQ

10

20

30

40

50

LGGD

【0 5 9 5】

123 Cas9 TadACP139ins 1299

MDKKYSIGLAIGTNSVGWAVITDEYKVPSSKFKVLGNTDRHSIKKNLIGA
 LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHR
 LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
 LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENP
 INASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
 NFKSNFDLAEDAKLQLSKDITYDDDLNLLAQIGDQYADLFLAAKNLSDAI
 LLSLILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI
 FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLR
 KQRTFDNGSIPHQIHLGELHAILRRQEDFYPPFLKDNREKIEKILTFRIPY
 YVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDK
 NLPNEKVLPHKSHLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD
 LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
 IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQ
 LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDD
 SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDELVKV
 MGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEELGSGILKEHP
 VENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDD
 SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL
 TKAERGGLSELDKAGFIKRLVETRQITKHVAQILDSRMNTKYDENDKLI
 REVKVITLKSCLVSDFRKDFQFYKVRINNYHHAHDAYLNAVVGITALIKK
 YPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEI
 TLANGAIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIVKKTEV
 QTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVE
 KGKSKKLKSVKELLGITIMERSSEFEKNPIDFLEAKGYKEVKKDLIIKLPK
 YSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPE
 DNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVL SAYNKH RMN
 HRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTDGSSGSETPGT
 SESATPESSGSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNRRVI
 GEGWNRAIGLHDPTAHAEIMALRQGGMLVMQNYRLIDATLYVTFEPCVMCA
 GAMIHRSRIGRVVFGVRNAKTGAAGSLMDVLHYPGDKPIREQAENIIHLFT
 LTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDL SQ
 LGGD

10

20

30

【0 5 9 6】

124 Cas9 delta 792-872 TadAins

MDKKYSIGLAIGTNSVGWAVITDEYKVPSSKFKVLGNTDRHSIKKNLIGA
 LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHR
 LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
 LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENP
 INASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
 NFKSNFDLAEDAKLQLSKDITYDDDLNLLAQIGDQYADLFLAAKNLSDAI
 LLSLILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI
 FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLR
 KQRTFDNGSIPHQIHLGELHAILRRQEDFYPPFLKDNREKIEKILTFRIPY
 YVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDK
 NLPNEKVLPHKSHLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD
 LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
 IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQ

40

50

LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDD
 SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKV
 MGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEDIKELGSEVEFSHE
 YWMRHALTLAKRARDEREVPVGAVLVNLRVIGEGWNRAIGLHDPTAHAE
 IMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNA
 KTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQ
 KKAQSSTDEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKA
 GFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVS
 DFRKDFQFYKVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYK
 VYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIE
 TNGETGEIVWDKGRDFATVRKVLSPQVNVIVKKTEVQTGGFSKESILPKR
 NSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEGKSKKLKSVKELL
 GITIMERSSEKPNIDFLEAKGYKEVKKDLIKLPKYSLELENGRKRML
 ASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDEQKQLFVEQHKH
 YLDEIIEQISEFSKRVILADANLDKVL SAYNKH RD KPIREQAENIIHLFT
 LTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDL SQ
 LGGD

[0 5 9 7]

125 Cas9 delta 792-906 TadAins

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
 LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHR
 LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
 LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENP
 INASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
 NFKSNFDLAEDAKLQLSKD TYDDDLNLLAQIGDQYADLFLAAKNLSDAI
 LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI
 FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKNREDLLR
 KQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPY
 YVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDK
 NLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD
 LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
 IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQ
 LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDD
 SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKV
 MGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEDIKELGSEVEFSHE
 YWMRHALTLAKRARDEREVPVGAVLVNLRVIGEGWNRAIGLHDPTAHAE
 IMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNA
 KTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQ
 KKAQSSTDGLSELKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDK
 LIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAII
 KKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKT
 EITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNVIVKKT
 EVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKE
 GKGSKKLKSVKELLGITIMERSSEKPNIDFLEAKGYKEVKKDLIKLPKYSLE
 LENGKRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGS
 PEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVL SAYNKH R
 DKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLI
 HQSITGLYETRIDL SQ LGGD

[0 5 9 8]

126 TadA CP65ins 1003

10

20

30

40

50

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
 LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHR
 LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
 LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENP
 INASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
 NFKSNFDLAEDAKLQLSKD TYDDDLNLLAQIGDQYADLFLAAKNLSDAI
 LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI
 FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKNLNREDLLR
 KQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPY
 YVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDK
 NLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD
 LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
 IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQ
 LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDD
 SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKV
 MGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEDIKELGSQILKEHP
 VENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVHDHVPQSFLKDD
 SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL
 TKAERGGLSELDKAGFIKRLVETRQITKHVAQILDSRMNTKYDENDKLI
 REVKVITLKSCLVSDFRKDFQFYKVINNYHHAHDAYLNAVVGITALIKK
 YPKTAHAIEIMALRQGGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGR
 VVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRM
 PRQVFNAQKKAQSSTDGSSGSETPGTSESATPESSGSEVEFSHEYWMRHA
 LTLAKRARDEREVPVGAVLVLNRRVIGEGWNRAIGLHDPLESEFVYGDYK
 VYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIE
 TNGETGEIVWDKGRDFATVRKVLSPQVNVIVKKTEVQTGGFSKESILPKR
 NSDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVEKGKSKKLKSVKELL
 GITIMERSSEFKNPIDFLEAKGYKEVKKDLIKLPKYSLEFLENKRKML
 ASAGELQKGNELALPSKYVNFLYLASHYEKLGKSPEDNEQKQLFVEQHKH
 YLDEIIEQISEFSKRVILADANLDKVL SAYNKHDKPIREQAENIIHLFT
 LTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDL SQ
 LGGD

[0599]

127 TadA CP65ins 1016

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
 LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHR
 LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
 LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENP
 INASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
 NFKSNFDLAEDAKLQLSKD TYDDDLNLLAQIGDQYADLFLAAKNLSDAI
 LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI
 FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKNLNREDLLR
 KQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPY
 YVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDK
 NLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD
 LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
 IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQ
 LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDD
 SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKV
 MGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEDIKELGSQILKEHP

VENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDD
SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL
TKAERGGELSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI
REVKVITLKSCLVSDFRKDFQFYKVINNYHHAHDAYLNAVVGTAIIKK
YPKLESEFVYGDYKVTAAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVM
CAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILAD
ECAALLCYFFRMPRQVFNAQKKAQSSTDGSSGSETPGTSESATPESSGSE
VEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHD
PYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIE
TNGETGEIVWDKGRDFATVRKVLSPQVNIKKTEVQTGGFSKESILPKR
NSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEKGKSKKLKSVKELL
GITIMERSSEKKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRML
ASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKH
YLDEIIEQISEFSKRVILADANLDKVL SAYNKH RDKPIREQAENIIHLFT
LTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDL SQ
LGGD

[0600]

128 TadA CP65ins 1022

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEFSNEMAKVDDSFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENP
INASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
NFKSNFDLAEDAKLQLSKD TYDDDLDNLLAQIGDQYADLFLAAKNLSDAI
LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI
FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLR
KQRTFDNGSIPHQIHLGELHAILRRQEDFYPLKDNREKIEKILTFRIPY
YVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGASAQSFIERM TNFDK
NLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD
LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQ
LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKS DGFANRNF MQLIHDD
SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDELVKV
MGRHKPENIVIEMARENQTTQKGQKNSRERMKRIE EGIKELGSQILKEHP
VENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDD
SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL
TKAERGGELSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI
REVKVITLKSCLVSDFRKDFQFYKVINNYHHAHDAYLNAVVGTAIIKK
YPKLESEFVYGDYKVDVRKMITAAHAEIMALRQGGLVMQNYRLIDATLYV
TFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEI
TEGILADECAALLCYFFRMPRQVFNAQKKAQSSTDGSSGSETPGTSESAT
PESSGSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWN
RAIGLHDPKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIE
TNGETGEIVWDKGRDFATVRKVLSPQVNIKKTEVQTGGFSKESILPKR
NSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEKGKSKKLKSVKELL
GITIMERSSEKKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRML
ASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKH
YLDEIIEQISEFSKRVILADANLDKVL SAYNKH RDKPIREQAENIIHLFT
LTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDL SQ
LGGD

10

20

30

40

50

【0601】

129 TadA CP65ins 1029

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENP
INASGVDAKAILSARLSKSRRENLIAQLPGEKKNGLFGNLIALSLGLTP
NFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAI
LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPKEYKEI
FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLR
KQRTFDNGSIPHQIHLGELHAILRRQEDFYPPFLKDNREKIEKILTFRIPY
YVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDK
NLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD
LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQ
LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDD
SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIAKKGILQTVKVVDELVKV
MGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEELGSGILKEHP
VENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDD
SIDNKVLTRSDKNRGKSDNVPSEEVVKMKKNYWRQLLNAKLITQRKFDNL
TKAERGGLSELDKAGFIKRLVETRQITKHVAQILDSRMNTKYDENDKLI
REVKVITLKSCLVSDFRKDFQFYKVINNYHHAHDAYLNAVVGTAIIKK
YPKLESEFVYGDYKVYDVRKMIKSEQEITAHAEIMALRQGGGLVMQNYRL
IDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPG
MNRHVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTDGSSGSETP
GTSESATPESSGSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNRR
VIGEGWNRAIGLHDPGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIE
TNGETGEIVWDKGRDFATVRKVLSPQVNVIVKKTEVQTGGFSKESILPKR
NSDKLIARKKDWDPPKYGGFDSPTVAYSVLVVAKEKGSKKLKS VKELL
GITIMERSSEFKNPIDFLEAKGYKEVKKDLIKLPKYSLELENGRKRML
ASAGELQKGNELALPSKYVNFYLYLASHYEKLGSPEDNEQKQLFVEQHKH
YLDEIIEQISEFSKRVILADANLDKVL SAYNKH RDKPIREQAENIIHLFT
LTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQ SITGLYETRIDL SQ
LGGD

10

20

30

【0602】

130 TadA CP65ins 1041

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENP
INASGVDAKAILSARLSKSRRENLIAQLPGEKKNGLFGNLIALSLGLTP
NFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAI
LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPKEYKEI
FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLR
KQRTFDNGSIPHQIHLGELHAILRRQEDFYPPFLKDNREKIEKILTFRIPY
YVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDK
NLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD
LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQ
LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDD

40

50

SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKKVDELVKV
MGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEDIKELGSQILKEHP
VENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDD
SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL
TKAERGGELSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI
REVKVITLKSCLVSDFRKDFQFYKVRINNYHHAHDAYLNAVVGTAIIKK
YPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSTAHAEIMALR
QGGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAA
GSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQS
STDGSSGSETPGTSESATPESSGSEVEFSHEYWMRHALTLAKRARDEREV
PVGAVLVLNNRVIGEGWNRAIGLHDPNIMNFFKTEITLANGEIRKRPLIE
TNGETGEIVWDKGRDFATVRKVLSPQVNVKKTEVQTGGFSKESILPKR
NSDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVEKGKSKKLKSVKELL
GITIMERSSEFKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRML
ASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDEQKQLFVEQHKH
YLDEIIEQISEFSKRVILADANLDKVL SAYNKHDKPIREQAENIIHLFT
LTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQSI TGLYETRIDL SQ
LGGD

10

[0 6 0 3]

131 TadA CP65ins 1054

20

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDDSFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENP
INASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALLSLGLTP
NFKSNFDLAEDAKLQLSKD TYDDDLNLLAQIGDQYADLFLAAKNLSDAI
LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI
FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLR
KQRTFDNGSIPHQIHLGELHAILRRQEDFY PFLKDNREKIEKILTFRI PY
YVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDK
NLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD
LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMKQ
LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNF MQLIHDD
SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKKVDELVKV
MGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEDIKELGSQILKEHP
VENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDD
SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL
TKAERGGELSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI
REVKVITLKSCLVSDFRKDFQFYKVRINNYHHAHDAYLNAVVGTAIIKK
YPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSTAHAEIMALR
QGGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAA
GSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQS
STDGSSGSETPGTSESATPESSGSEVEFSHEYWMRHALTLAKRARDEREV
PVGAVLVLNNRVIGEGWNRAIGLHDPNIMNFFKTEITLANGEIRKRPLIE
TNGETGEIVWDKGRDFATVRKVLSPQVNVKKTEVQTGGFSKESILPKR
NSDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVEKGKSKKLKSVKELL
GITIMERSSEFKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRML
ASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDEQKQLFVEQHKH
YLDEIIEQISEFSKRVILADANLDKVL SAYNKHDKPIREQAENIIHLFT

30

40

50

LTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSI TGLYETRIDL SQ
LGGD

【0604】

132 TadA CP65ins 1246

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDD SFFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENP
INASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
NFKSNFDLAEDAKLQLSKD TYDDDLDNLLAQIGDQYADLFLAAKNLSDAI
LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI
FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLR
KQRTFDNGSIPHQIHLGELHAILRRQEDFY PFLKDNREKIEKILTFRI PY
YVGPLARGNSRFAWMTRKSEETITPWNFE EVVDKGASAQSFIERMTNFDK
NLPNEKVLPHKSHLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD
LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
IKDKDFLDNEENEDILEDIVLTLT LFEDREMIEERLKTYAHLFDDKVMKQ
LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKS DGFANRNF MQLIHDD
SLTFKEDIQKAQVSGQGD SLHEHIANLAGSPA IKKGILQTVKVVD ELVKV
MGRHKPENIVIE MARENQTTQKGQKNSRERMKRIEEG IKELGSQILKEHP
VENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYD VDHIVPQSFLKDD
SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL
TKAERGG LSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI
REVKVITLKS KLVSDFRKDFQFYK VREINNYHHAH DAYLNAVVG TALIKK
YPKLESEFVYGDYKVYDVRKMI AKSEQEIGKATAKYFFYSNIMNFFKTEI
TLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEV
QTGGFSKESILPKRNSDKLIARKKD WDPKKYGGFDSPTVAYSVLV VAKVE
KGKSKKLKSVKELLGITIMERS SFEKNPIDFLEAKGYKEVKKDLI IKLPK
YSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGTAH
AEIMALRQGG LVMQNYRLIDATLYVT FEPCVMCAGAMIHSRIGRVVFGVR
NAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVFN
AQKKAQSSTDGSSGSETPGTSESATPESSGSEVEFSHEYWMRHALTLAKR
ARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPSPEDNEQKQLFVEQHKKH
YLDEIIEQISEFSKRVLADANLDKVL SAYNKH RDKPIREQAENIIHLFT
LTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSI TGLYETRIDL SQ
LGGD

【0605】

したがって、いくつかの実施形態において、Cas9ポリペプチドの同定された位置においてTadAまたはそのバリエーションを挿入するアデノシンデアミナーゼ塩基エディターが生成された。

【0606】

いくつかの実施形態において、アデノシンデアミナーゼアリル（例えばTadAアリル）の合成ライブラリーを利用して、改変された塩基編集効率および／または特異性を有するアデノシン塩基エディターを生成することができる。いくつかの実施形態において、合成ライブラリーから生成されたアデノシン塩基エディターは、より高い塩基編集効率および／または特異性を含む。いくつかの実施形態において、合成ライブラリーから生成されたアデノシン塩基エディターは、野生型TadAを有するアデノシン塩基エディターと比較して、増加した塩基編集効率、増加した塩基編集特異性、低減されたオフターゲット編集、低減されたバイスタンダー編集、低減されたインデル形成、および／または低減された目的外編集を示す。いくつかの実施形態において、合成ライブラリーから生成されたアデノシン

10

20

30

40

50

塩基エディターは、*TadA**7.10を有するアデノシン塩基エディターと比較して、増加した塩基編集効率、増加した塩基編集特異性、低減されたオフターゲット編集、低減されたバイズタンダー編集、低減されたインデル形成、および／または低減された目的外編集を示す。いくつかの実施形態において、合成ライブラリーは、ABEのランダム化された*TadA*部分を含む。いくつかの実施形態では、合成ライブラリーは、*TadA*の各位置に20個すべての正準的なアミノ酸の置換を含む。いくつかの実施形態において、合成ライブラリーは、ライブラリー構成員あたり1～2ヌクレオチド置換突然変異の平均頻度を含む。いくつかの実施形態において、合成ライブラリーは、*TadA**7.10に見られるバックグラウンド突然変異を含む。

【0607】

〔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（例えば細胞修復機構によって）チミジン塩基で置換されて、C:GからT:Aへの遷移を生じさせる。他の実施形態では、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。

【0608】

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、シチジン（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。

【0609】

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あるいはデアミナーゼドメイン（例えばシチジンデアミナーゼドメイン）を含む。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。いくつかの実施形態では、追加のドメインが、デアミナーゼドメインと共に内部的に融合されている。

【0610】

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

10

20

30

40

50

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、napDNAbpドメインがCas12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、Cas12-gRNA-標的DNA複合体の形成の際に対合しないままにしておくことができ、その結果、Cas12「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。

【0611】

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ファミリーは進化的に保存されたシチジンデアミナーゼを含む。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」と呼ばれる)、APOBEC3F、APOBEC3G、APOBEC3H、APOBEC4、および活性化誘導 (シチジン) デアミナーゼ (AID) が含まれる。SaBE3、SaKKH-BE3、VQR-BE3、EQR-BE3、VRER-BE3、YE1-BE3、EE-BE3、YE2-BE3、およびYEE-BE3を含むがこれらに限定されない多くの改変シチジンデアミナーゼを含む塩基エディターが市販されており、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。

【0612】

ある実施形態において、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。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が理解されるであろう。いくつかの実施形態において、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、またはマウスに由来する。いくつかの実施形態において、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)。いくつかの実施形態に

10

20

30

40

50

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。いくつかの実施形態において、塩基エディターのデアミナーゼドメインは、pmCDA1である。

【0613】

PmCDA1の塩基配列とアミノ酸配列、およびヒトAIDのCDSの塩基配列とアミノ酸配列を以下に示す。

【0614】

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

MTDAEYVRIHEKLDIYTFKKQFFNNKKS VSHRCYVLFELKRRGERRACFWGYAVNKPQS
GTERGIHAEIFSIRKVEEYLRDNPQGFTINWYSSWSPCADCAEKILEWYNQELRGNGHTL
KIWACKLYYEKNARNQIGLWNLRDNGVGLNVMVSEHYQCCRKIFIQSSHNQLNENRWL
EKTCLKRAEKRRSELSIMIQVKILHTTKSPAV

10

【0615】

> EF094822.1 Petromyzon marinus isolate PmCDA.21 cytosine deaminase mRNA, complete cds

TGACACGACACAGCCGTGTATATGAGGAAGGGTAGCTGGATGGGGGGGGGGGAATACG
TTCAGAGAGGACATTAGCGAGCGTCTTGTGGTGGCCTTGAGTCTAGACACCTGCAGAC
ATGACCGACGCTGAGTACGTGAGAATCCATGAGAAGTTGGACATCTACACGTTTAAAGAA
ACAGTTTTTCAACAACAAAAAATCCGTGTCGCATAGATGCTACGTTCTCTTTGAATTAA
AACGACGGGGTGAACGTAGAGCGTGT TTTTGGGGCTATGCTGTGAATAAACACAGAGC
GGGACAGAACGTGGAATTCACGCCGAAATCTTTAGCATTAGAAAAGTCGAAGAATACCT
GCGCGACAACCCCGGACAATTCACGATAAATTGGTACTCATCCTGGAGTCCTTGTGCAG
ATTGCGCTGAAAAGATCTTAGAATGGTATAACCAGGAGCTGCGGGGGGAACGGCCACACT
TTGAAAATCTGGGCTTGCAAACTCTATTACGAGAAAAATGCGAGGAATCAAATTGGGCT
GTGGAACCTCAGAGATAACGGGGTTGGGTTGAATGTAATGGTAAGTGAACACTACCAAT
GTTGCAGGAAAATATTCATCCAATCGTCGCACAATCAATTGAATGAGAATAGATGGCTT
GAGAAGACTTTGAAGCGAGCTGAAAAACGACGGAGCGAGTTGTCCATTATGATTTCAGGT
AAAAATACTCCACACCACTAAGAGTCCTGCTGTTTAAAGAGGCTATGCGGATGGTTTTTC

20

【0616】

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

MDSLLMNRKFLYQFKNVRWAKGRRETYLCYVVKRRDSATSFSLDFGYLRNKNKGCHVEL
LFLRYISDWDLDPGRCYRVTWFTSWSPCYDCARHVADFLRGPNPNSLRIFTARLYFCEDR
KAEPEGLRRLHRAGVQIAIMTFKAPV

30

【0617】

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

AGAGAACCATCATTAATTGAAGTGAGATTTTCTGGCCTGAGACTTGCAGGGAGGCAAG
AAGACACTCTGGACACCACTATGGACAGGTAAAGAGGCAGTCTTCTCGTGGGTGATTGC
ACTGGCCTTCTCTCAGAGCAAATCTGAGTAATGAGACTGGTAGCTATCCCTTTCTCTC
ATGTAACCTGTCTGACTGATAAGATCAGCTTGATCAATATGCATATATATTTTTTGTATCT
GTCTCCTTTTCTTCTATTTCAGATCTTATACGCTGTCAGCCCAATTCTTTCTGTTTCAGAC
TTCTCTTGATTTCCCTCTTTTTTCATGTGGCAAAAGAAGTAGTGCGTACAATGTACTGAT
TCGTCCTGAGATTTGTACCATGGTTGAACTAATTTATGGTAATAATATTAACATAGCA
AATCTTTAGAGACTCAAATCATGAAAAGGTAATAGCAGTACTGTACTAAAAACGGTAGT
GCTAATTTTCGTAATAATTTTGTAAATATTCAACAGTAAACAACTTGAAGACACACTT
TCCTAGGGAGGCGTTACTGAAATAATTTAGCTATAGTAAGAAAATTTGTAATTTTAGAA
ATGCCAAGCATTCTAAATTAATTGCTTGAAAGTCACTATGATTGTGTCCATTATAAGGA
GACAAATTCATTCAAGCAAGTTATTTAATGTTAAAGGCCCAATTGTTAGGCAGTTAATG
GCACTTTTACTATTAATAATCTTTCCATTTGTTTCAGACGTAGCTTAACTTACCTCTTAG

40

50

GTGTGAATTTGGTTAAGGTCCTCATAATGTCTTTATGTGCAGTTTTTGTATAGGTTATTG
TCATAGAACTTATTCTATTCTACATTTATGATTACTATGGATGTATGAGAATAACACC
TAATCCTTATACTTTACCTCAATTTAACTCCTTTATAAAGAACTTACATTACAGAATAA
AGATTTTTTTAAAAATATATTTTTTTGTAGAGACAGGGTCTTAGCCCAGCCGAGGCTGGT
CTCTAAGTCCTGGCCCAAGCGATCCTCCTGCCTGGGCCTCCTAAAGTGCTGGAATTATA
GACATGAGCCATCACATCCAATATACAGAATAAAGATTTTTTAATGGAGGATTTAATGTT
CTTCAGAAAATTTTCTTGAGGTCAGACAATGTCAAATGTCTCCTCAGTTTACACTGAGA
TTTTGAAAACAAGTCTGAGCTATAGGTCCTTGTGAAGGGTCCATTGGAAATACTTGTTT
AAAGTAAAATGGAAAGCAAAGGTAAAATCAGCAGTTGAAATTCAGAGAAAGACAGAAA
AGGAGAAAAGATGAAATTC AACAGGACAGAAGGGGAAATATATTATCATTAAAGGAGGAC
AGTATCTGTAGAGCTCATTAGTGATGGCAAAATGACTTGGTCAGGATTATTTTTTAACCC
GCTTGTTTTCTGGTTTGCACGGCTGGGGATGCAGCTAGGGTTCTGCCTCAGGGAGCACAG
CTGTCCAGAGCAGCTGTCAGCCTGCAAGCCTGAAACACTCCCTCGGTAAAGTCCTTCCT
ACTCAGGACAGAAATGACGAGAACAGGGAGCTGGAAACAGGGCCCCTAACCAGAGAAGG
GAAGTAATGGATCAACAAAGTTAACTAGCAGGTCAGGATCACGCAATTCATTTCACTCT
GACTGGTAACATGTGACAGAAACAGTG TAGGCTTATTGTATTTTCATGTAGAGTAGGAC
CCAAAAATCCACCCAAAGTCCTTTATCTATGCCACATCCTTCTTATCTATACTTCCAGG
ACACTTTTTCTTCCTTATGATAAGGCTCTCTCTCTCCACACACACACACACACACA
CACACACACACACACACACACACAAACACACACCCCGCCAACCAAGGTGCATGTAAAAA
GATGTAGATTCTCTGCCTTTCTCATCTACACAGCCCAGGAGGGTAAGTTAATATAAGA
GGGATTTATTGGTAAGAGATGATGCTTAATCTGTTTAACTGGGCCTCAAAGAGAGAA
TTTCTTTTCTTCTGTACTTATTAAGCACCTATTATGTGTTGAGCTTATATATACAAAGGG
TTATTATATGCTAATATAGTAATAGTAATGGTGGTTGGTACTATGGTAATTACCATAAA
AATTATTATCCTTTTAAAATAAAGCTAATTATTATTGGATCTTTTTTTAGTATTCATTTTA
TGTTTTTTATGTTTTTGATTTTTTAAAAGACAATCTCACCTGTTACCCAGGCTGGAGTG
CAGTGGTGCAATCATAGCTTTCTGCAGTCTTGAACCTCCTGGGCTCAAGCAATCCTCTG
CCTTGGCCTCCCAAAGTGTTGGGATACAGTCATGAGCCACTGCATCTGGCCTAGGATCC
ATTTAGATTAAAATATGCATTTTAAATTTTAAAATAATATGGCTAATTTTACCTTATG
TAATGTGTATACTGGCAATAAATCTAGTTTGCTGCCTAAAGTTTAAAGTGCTTTCCAGT
AAGCTTCATGTACGTGAGGGGAGACATTTAAAGTGAAACAGACAGCCAGGTGTGGTGGC
TCACGCCTGTAATCCCAGCACTCTGGGAGGCTGAGGTGGGTGGATCGCTTGAGCCCTGG
AGTTCAAGACCAGCCTGAGCAACATGGCAAAACGCTGTTTCTATAACAAAAATTAGCCG
GGCATGGTGGCATGTGCCTGTGGTCCCAGCTACTAGGGGGCTGAGGCAGGAGAATCGTT
GGAGCCCAGGAGGTCAAGGCTGCACTGAGCAGTGCTTGCGCCACTGCACTCCAGCCTGG
GTGACAGGACCAGACCTTGCCTCAAAAAAATAAGAAGAAAAATTA AAAATAAATGGAAA
CAACTACAAAGAGCTGTTGTCCTAGATGAGCTACTTAGTTAGGCTGATATTTTGGTATT
TAACTTTTAAAGTCAGGGTCTGTCACCTGCACTACATTATTA AAATATCAATTCTCAAT
GTATATCCACACAAAGACTGGTACGTGAATGTT CATAGTACCTTTATTCACAAAACCCC
AAAGTAGAGACTATCCAATATCCATCAACAAGTGAACAAATAAACAAAATGTGCTATA
TCCATGCAATGGAATACCACCCTGCAGTACAAAGAAGCTACTTGGGGATGAATCCCAA
GTCATGACGCTAAATGAAAGAGTCAGACATGAAGGAGGAGATAATGTATGCCATACGAA
ATTCTAGAAAATGAAAGTAACTTATAGTTACAGAAAGCAAATCAGGGCAGGCATAGAGG
CTCACACCTGTAATCCCAGCACTTTGAGAGGCCACGTGGGAAGATTGCTAGAACTCAGG
AGTTCAAGACCAGCCTGGGCAACACAGTGAAACTCCATTCTCCACAAAAATGGGAAAAA
AAGAAAGCAAATCAGTGGTTGTCCTGTGGGGAGGGGAAGGACTGCAAAGAGGGGAAGAA
GCTCTGGTGGGGTGAGGGTGGTGATTCAGGTTCTGTATCCTGACTGTGGTAGCAGTTTG
GGGTGTTTACATCCAAAAATATTCGTAGAATTATGCATCTTAAATGGGTGGAGTTTACT
GTATGTAAATTATACCTCAATGTAAGAAAAAATAATGTGTAAGAAAACTTTCAATTCTC
TTGCCAGCAAACGTTATTCAAATTCCTGAGCCCTTTACTTCGCAAATTCTCTGCACTTCT
GCCCCGTACCATTAGGTGACAGCACTAGCTCCACAAATTGGATAAATGCATTTCTGGAA

10

20

30

40

50

AAGACTAGGGACAAAATCCAGGCATCACTTGTGCTTTTCATATCAACCATGCTGTACAGC
TTGTGTTGCTGTCTGCAGCTGCAATGGGGACTCTTGATTTCTTTAAGGAACTTGGGTT
ACCAGAGTATTTCCACAAATGCTATTCAAATTAGTGCTTATGATATGCAAGACACTGTG
CTAGGAGCCAGAAAACAAAGAGGAGGAGAAATCAGTCATTATGTGGGAACAACATAGCA
AGATATTTAGATCATTTTACTAGTTAAAAAAGCAGCAGAGTACAAAATCACACATGCA
ATCAGTATAATCCAAATCATGTAAATATGTGCCTGTAGAAAGACTAGAGGAATAAACAC
AAGAATCTTAACAGTCATTGTCTATTAGACACTAAGTCTAATTATTATTATTAGACACTA
TGATATTTGAGATTTAAAAAATCTTTAATATTTTAAAATTTAGAGCTCTTCTATTTTTCC
ATAGTATTCAAGTTTGACAATGATCAAGTATTACTCTTTCTTTTTTTTTTTTTTTTTTT
TTTTTGAGATGGAGTTTTTGGTCTTGTTGCCCATGCTGGAGTGGAATGGCATGACCATAG
CTCACTGCAACCTCCACCTCCTGGGTTCAAGCAAAGCTGTCGCCTCAGCCTCCCGGGTA
GATGGGATTACAGGCGCCACCACCACACTCGGCTAATGTTTGTATTTTTTAGTAGAGAT
GGGGTTTCACCATGTTGGCCAGGCTGGTCTCAAACCTCCTGACCTCAGAGGATCCACCTG
CCTCAGCCTCCCAAAGTGCTGGGATTACAGATGTAGGCCACTGCGCCCGGCCAAGTATT
GCTCTTATACATTAAAAAACAGGTGTGAGCCACTGCGCCAGCCAGGTATTGCTCTTAT
ACATTAAAAAATAGGCCGGTGCAGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAAGC
CAAGGCGGGCAGAACACCCGAGGTGAGGAGTCCAAGGCCAGCCTGGCCAAGATGGTGAA
ACCCCGTCTCTATTAAAAATACAAACATTACCTGGGCATGATGGTGGGCGCCTGTAATC
CCAGCTACTCAGGAGGCTGAGGCAGGAGGATCCGCGGAGCCTGGCAGATCTGCCTGAGC
CTGGGAGGTTGAGGCTACAGTAAGCCAAGATCATGCCAGTATACTTCAGCCTGGGCGAC
AAAGTGAGACCGTAACAAAAAATAAATTTAAAAAAGAAATTTAGATCAAGATCCA
ACTGTAAAAAGTGGCCTAAACACCACATTAAAGAGTTTGGAGTTTATTCTGCAGGCAGA
AGAGAACCATCAGGGGGTCTTCAGCATGGGAATGGCATGGTGCACCTGGTTTTTGTGAG
ATCATGGTGGTGACAGTGTGGGGAATGTTATTTTGGAGGGACTGGAGGCAGACAGACCG
GTTAAAAGGCCAGCACACAGATAAGGAGGAAGAAGATGAGGGCTTGGACCGAAGCAG
AGAAGAGCAAACAGGGAAGGTACAAATTCAAGAAATATTGGGGGGTTTGAATCAACACA
TTTAGATGATTAATTAAATATGAGGACTGAGGAATAAGAAATGAGTCAAGGATGGTTCC
AGGCTGCTAGGCTGCTTACCTGAGGTGGCAAAGTCGGGAGGAGTGGCAGTTTAGGACAG
GGGGCAGTTGAGGAATATTGTTTTGATCATTTTGAAGTTTGAAGTACAAGTTGGACACTT
AGGTAAAGACTGGAGGGGAAATCTGAATATACAATTATGGGACTGAGGAACAAGTTTAT
TTATTTTTTTGTTTCGTTTTCTTGTTGAAGAACAATTTAATTGTAATCCCAAGTCATCA
GCATCTAGAAGACAGTGGCAGGAGGTGACTGTCTTGTTGGGTAAGGGTTTGGGGTCCTTG
ATGAGTATCTCTCAATTGGCCTTAAATATAAGCAGGAAAAGGAGTTTATGATGGATTCC
AGGCTCAGCAGGGCTCAGGAGGGCTCAGGCAGCCAGCAGAGGAAGTCAGAGCATCTTCT
TTGGTTTAGCCCAAGTAATGACTTCCTTAAAAAGCTGAAGGAAAATCCAGAGTGACCAG
ATTATAAACTGTACTCTTGCATTTTCTCTCCCTCCTCTCACCCACAGCCTCTTGATGAAC
CGGAGGAAGTTTCTTTACCAATTCAAAAATGTCCGCTGGGCTAAGGGTCGGCGTGAGAC
CTACCTGTGCTACGTAGTGAAGAGGCGTGACAGTGCTACATCCTTTTCACTGGACTTTG
GTTATCTTCGCAATAAGGTATCAATTAAAGTCGGCTTTGCAAGCAGTTTAATGGTCAAC
TGTGAGTGCTTTTAGAGCCACCTGCTGATGGTATTACTTCCATCCTTTTTTGGCATTGT
GTCTCTATCACATTCTCAAATCCTTTTTTTTATTTCTTTTTCCATGTCCATGCACCCAT
ATTAGACATGGCCCAAATATGTGATTTAATTCTCCTCCCAAGTAATGCTGGGCACCCTAA
TACCACTCCTTCTTTCAGTGCCAAGAACAACCTGCTCCCAAACCTGTTTACCAGCTTTCCTC
AGCATCTGAATTGCCTTTGAGATTAATTAAGCTAAAAGCATTTTTATATGGGAGAATAT
TATCAGCTTGTCCAAGCAAAAATTTTAAATGTGAAAAACAATTGTGTCTTAAGCATTT
TTGAAAATTAAGGAAGAAGAATTTGGGAAAAAATTAACGGTGGCTCAATTCTGTCTTCC
AAATGATTTCTTTTCCCTCCTACTCACATGGGTCGTAGGCCAGTGAATACATTCAACAT
GGTGATCCCCAGAAAACCTCAGAGAAGCCTCGGCTGATGATTAATTAATGATCTTTTCG
GCTACCCGAGAGAATTACATTTCCAAGAGACTTCTTCACCAAAATCCAGATGGGTTTAC
ATAAACTTCTGCCCACGGGTATCTCCTCTCTCCTAACACGCTGTGACGTCTGGGCTTGG

10

20

30

40

50

TGGAATCTCAGGGAAGCATCCGTGGGGTGGAAAGGTCATCGTCTGGCTCGTTGTTTGATG
GTTATATTACCATGCAATTTTCTTTGCCTACATTTGTATTGAATACATCCCAATCTCCTT
CCTATTCGGTGACATGACACATTCTATTTT CAGAAGGCTTTGATTTTATCAAGCACTTTC
ATTTACTTCTCATGGCAGTGCCTATTACTTCTCTTACAATACCCATCTGTCTGCTTTACC
AAAATCTATTTCCCTTTTTCAGATCCTCCCAAATGGTCCTCATAAACTGTCCTGCCTCCA
CCTAGTGGTCCAGGTATATTTCCACAATGTTACATCAACAGGCACTTCTAGCCATTTTC
CTTCTCAAAAGGTGCAAAAAGCAACTTCATAAACACAAATTAATCTTCGGTGAGGTAG
TGTGATGCTGCTTCCTCCCAACTCAGCGCACTTCGTCTTCCTCATTCCACAAAAACCCAT
AGCCTTCCTTCACTCTGCAGGACTAGTGCTGCCAAGGGTTCAGCTCTACCTACTGGTGT
GCTCTTTT GAGCAAGTTGCTTAGCCTCTCTGTAACACAAGGACAATAGCTGCAAGCATC
CCCAAAGATCATTGCAGGAGACAATGACTAAGGCTACCAGAGCCGCAATAAAAGTCAGT
GAATTTTAGCGTGGTCTCTCTGTCTCTCCAGAACGGCTGCCACGTGGAATTGCTCTTC
CTCCGCTACATCTCGGACTGGGACCTAGACCCTGGCCGCTGCTACCGCGTCACCTGGTT
CACCTCCTGGAGCCCCTGCTACGACTGTGCCCCGACATGTGGCCGACTTTCTGCGAGGGA
ACCCCAACCTCAGTCTGAGGATCTTCACCGCGCGCCTCTACTTCTGTGAGGACCGCAAG
GCTGAGCCCGAGGGGCTGCGGCGGCTGCACCGCGCCGGGGTGCAAATAGCCATCATGAC
CTTCAAAGGTGCGAAAGGGCCTTCCGCGCAGGCGCAGTGCAGCAGCCCGCATTTCGGGAT
TGCGATGCGGAATGAATGAGTTAGTGGGGAAGCTCGAGGGGAAGAAGTGGGCGGGGATT
CTGGTTACCTCTGGAGCCGAAATTAAGATTAGAAGCAGAGAAAAGAGTGAATGGCTC
AGAGACAAGGCCCGAGGAAATGAGAAAATGGGGCCAGGGTTGCTTCTTTCCCTCGAT
TTGGAACCTGAACTGTCTTCTACCCCCATATCCCCGCCTTTTTTTCTTTTTTTTTTTT
GAAGATTATTTTTACTGCTGGAATACTTTTGTAGAAAACCACGAAAGAACTTTCAAAGC
CTGGGAAGGGCTGCATGAAAATTCAGTTTCGTCTCTCCAGACAGCTTCGGCGCATCCTTT
TGGTAAGGGGCTTCCTCGCTTTTTTAAATTTTCTTTCTTTCTCTACAGTCTTTTTTGGAGT
TTCGTATATTTCTTATATTTTCTTATTGTTCAATCACTCTCAGTTTTTCATCTGATGAAAA
CTTTATTTCTCCTCCACATCAGCTTTTTCTTCTGCTGTTTCACCATT CAGAGCCCTCTGC
TAAGGTTCTTTTTCCCTCCCTTTTTCTTTCTTTTGTGTTTCACATCTTTAAATTTCTGTC
TCTCCCCAGGGTTGCGTTTCCTTCCTGGTCAGAAATCTTTTCTCCTTTTTTTTTTTTTT
TTTTTTTTTTTTTAAACAAACAAACAAAAAACCCAAAAAACTCTTTCCCAATTTACTTT
CTTCCAACATGTTACAAAGCCATCCACTCAGTTTAGAAGACTCTCCGGCCCCACCGACC
CCCAACCTCGTTTTGAAGCCATTCACTCAATTTGCTTCTCTCTTTCTCTACAGCCCCTGT
ATGAGGTTGATGACTTACGAGACGCATTTTCGTACTTTGGGACTTTGATAGCAACTTCCA
GGAATGTCACACACGATGAAATATCTCTGCTGAAGACAGTGGATAAAAAACAGTCCTTC
AAGTCTTCTCTGTTTTTATTCTTCAACTCTCACTTTCTTAGAGTTTACAGAAAAAATATT
TATATACGACTCTTTAAAAAGATCTATGTCTTGAAAATAGAGAAGGAACACAGGTCTGG
CCAGGGACGTGCTGCAATTGGTGCAGTTTTGAATGCAACATTGTCCCCTACTGGGAATA
ACAGAACTGCAGGACCTGGGAGCATCTAAAGTGTCAACGTTTTTCTATGACTTTTAGG
TAGGATGAGAGCAGAAGGTAGATCCTAAAAAGCATGGTGAGAGGATCAAATGTTTTTAT
ATCAACATCCTTTATTATTTGATTCATTTGAGTTAACAGTGGTGTAGTGATAGATTTTT
CTATTCTTTTCCCTTGACGTTTACTTTCAAGTAACACAACTCTTCCATCAGGCCATGAT
CTATAGGACCTCCTAATGAGAGTATCTGGGTGATTGTGACCCCAAACCATCTCTCCAAA
GCATTAATATCCAATCATGCGCTGTATGTTTTAATCAGCAGAAGCATGTTTTTATGTTT
GTACAAAAGAAGATTGTTATGGGTGGGGATGGAGGTATAGACCATGCATGGTCACTTC
AAGCTACTTTAATAAAGGATCTTAAATGGGCAGGAGGACTGTGAACAAGACACCCTAA
TAATGGGTTGATGTCTGAAGTAGCAAATCTTCTGGAAACGCAAACCTTTTTAAGGAAGT
CCCTAATTTAGAAACACCCACAACTTCACATATCATAATTAGCAAACAATTGGAAGGA
AGTTGCTTGAATGTTGGGGAGAGGAAAATCTATTGGCTCTCGTGGGTCTCTTCATCTCA
GAAATGCCAATCAGGTCAAGGTTTGCTACATTTTGTATGTGTGTGATGCTTCTCCAAA
GGTATATTAATACTATATAAGAGAGTTGTGACAAAACAGAATGATAAAGCTGCGAACCGTG
GCACACGCTCATAGTTCTAGCTGCTTGGGAGGTTGAGGAGGGAGGATGGCTTGAACACA

10

20

30

40

50

GGTGTTC AAGGCCAGCCTGGGCAACATAACAAGATCCTGTCTCTCAAAAAAAAAAAAAAAAAA
AAAAAAGAAAGAGAGAGGGCCGGGCGTGGTGGCTCACGCCTGTAATCCCAGCACTTTGG
GAGGCCGAGCCGGGCGGATCACCTGTGGTCAGGAGTTTGAGACCAGCCTGGCCAACATG
GCAAAACCCCGTCTGTACTCAAAATGCAAAATTAGCCAGGCGTGGTAGCAGGCACCTG
TAATCCCAGCTACTTGGGAGGCTGAGGCAGGAGAATCGCTTGAACCCAGGAGGTGGAGG
TTGCAGTAAGCTGAGATCGTGCCGTTGCACTCCAGCCTGGGCGACAAGAGCAAGACTCT
GTCTCAGAAAAAAAAAAAAAAAAAGAGAGAGAGAGAGAAAGAGAACAAATATTTGGGAGA
GAAGGATGGGGAAGCATTGCAAGGAAATTGTGCTTTATCCAACAAAATGTAAGGAGCCA
ATAAGGGATCCCTATTTGTCTCTTTTGGTGTCTATTTGTCCCTAACAACTGTCTTTGACA
GTGAGAAAAATATTCAGAATAACCATATCCCTGTGCCGTTATTACCTAGCAACCCTTGC
AATGAAGATGA

10

GCAGATCCACAGGAAAACCTTGAATGCACAACTGTCTTATTTTAATCTTATTGTACATAA
GTTTGTAAGAGATTAAAAATTGTTACTTCATGTATTCAATTTATATTTTATATTATTTTG
CGTCTAATGATTTTTTTTATTAACATGATTTCTTTTCTGATATATTGAAATGGAGTCTCAA
AGCTTCATAAATTTATACTTTAGAAATGATTCTAATAACAACGTATGTAATTGTAACA
TTGCAGTAATGGTGTCTACGAAGCCATTTCTCTTGATTTTTTAGTAACTTTTATGACAGC
AAATTTGCTTCTGGCTCACTTTCAATCAGTTAAATAAATGATAAATAATTTTGGGAAGCT
GTGAAGATAAAATACCAAATAAAATAATATAAAAGTGATTTATATGAAGTTAAATAAA
AAATCAGTATGATGGAATAAACTTG

20

【0618】

本発明の方法において有用な他のシチジンデアミナーゼを以下に提供する。

【0619】

rAPOBEC-1 *Rattus norvegicus*

MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELKTCCLLYEINWGGRHSIWRHTSQNTNK
HVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSRAITEFLSRYPHVTLFIYIARLYHH
ADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFVNYSNEAHWPRYPHLWVRLYVLE
LYCIILGLPPCLNILRRKQPQLTFFTIALQSCHYQRLPPHILWATGLK

【0620】

mAPOBEC-1 *Mus musculus*

MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELKTCCLLYEINWGGRHSVWRHTSQNTSN
HVEVNFLEKFTTERYFRPNTRCSITWFLSWSPCGECSRAITEFLSRHPYVTLFIYIARLY
HHTDQRNRQGLRDLISSGVTIQIMTEQEYCYCWRNFVNYPSPNEAYWPRYPHLWVKLY
VLELYCIILGLPPCLKILRRKQPQLTFFTITLQTCHYQRIPPHLLWATGLK

30

【0621】

maAPOBEC-1 *Mesocricetus auratus*

MSSETGPVVVDPTLRRRIEPHEFDAFFDQGELRKTCCLLYEIRWGGRHNIWRHTGQNTSR
HVEINFIEKFTSERYFYPSTRCSIVWFLSWSPCGECSKAITEFLSGHPNVTLFIYAARLY
HHTDQRNRQGLRDLISRGVTIRIMTEQEYCYCWRNFVNYPSPNEVYWPYPNLWMRLY
ALELYCIHLGLPPCLKIKRRHQYPLTFFRLNLQSCHYQRIPPHILWATGFI

40

【0622】

hAPOBEC-1 *Homo sapiens*

MTSEKGPSTGDPTLRRRIEPWEFDVFYDPRELKTCCLLYEIKWGMSRKIWRSSGKNTTN
HVEVNFIEKFTSERDFHPSMCSITWFLSWSPCWECSQAIREFLSRHPGVTLVIYVARLF
WHMDQQNRQGLRDLVNSGVTIQIMRASEYYHCWRNFVNYPSPGDEAHWPQYPPLWMM
LYALELHCIILSLPPCLKISRRWQNHLTFFRLHLQNCHYQTIPPHILLATGLIHPSVAWR

【0623】

ppAPOBEC-1 *Pongo pygmaeus*

MTSEKGPSTGDPTLRRRIESWEFDVFYDPRELKTCCLLYEIKWGMSRKIWRSSGKNTTN
HVEVNFIEKFTSERRFHSSISCSITWFLSWSPCWECSQAIREFLSQHPGVTLVIYVARLF

50

WHMDQRNRQGLRDLVNSGVTIQIMRASEYYHCWRNFVNYPGDEAHWPQYPPLWMM
LYALELHCCIISLPPCLKISRROWQHHLAFFRLHLQNCHYQTIPPHILLATGLIHPSVTWR
[0 6 2 4]

ocAPOBEC1 *Oryctolagus cuniculus*

MASEKGPSNKDYTLRRRIEPWEFEVFFDPQELRKEACLLYEIKWGASSKTWRSSGKNTTN
HVEVNFLEKLTSEGR LGPSTCCSITWFLSWSPCWECSMAIREFLSQHPGVTLIIFVARLF
QHMDRRNRQGLKDLVTSGVTVRVMSVSEYCYCWENFVNYPGKAAQWPRYPPRWML
MYALELYCII LGLPPCLKISR RHQKQLTFFSLTPQYCHYKMIPPYILLATGLLQPSVPWR
[0 6 2 5]

mdAPOBEC-1 *Monodelphis domestica*

MNSKTGPSVG DATLRRRIKPWEFVAFFNPQELRKETCLLYEIKWGNQNIWRHSNQNTSQ
HAEINFMEKFTAERHFNSSVRCSITWFLSWSPCWECSKAIRKFLDHYPNVT LAIFISRLY
WHMDQQHRQGLKELVHSGVTIQIMSYSEYHYCWRNFVDYPQGEEDYWPKY PYLWIML
YVLELHCCI LGLPPCLKISGSHSNQLALFSLDLQDCHYQKIPYNVLVATGLVQPFVTWR
[0 6 2 6]

mAPOBEC-2 *Mus musculus*

MAQKEEAAEAAAPASQNGDDLENLEDPEKLKELIDLP PFEIVTGVRLPVNFFKFQFRNVE
YSSGRNKTF L CYVVEVQSKGGQAQATQGYLEDEHAGAHAE EAFNTILPAFDPALKYNVT
WYVSSSPCAACADRILKTL SKTKNLRLLILVSRLFMWEEPEVQAALKKLKEAGCKLRIMK
PQDFEYIWQNFVEQE EGESKAFEPWEDIQENFLY YEEKLADILK
[0 6 2 7]

hAPOBEC-2 *Homo sapiens*

MAQKEEAAVATEAASQNGEDLENLDDPEKLKELIELPPFEIVTGERLPANFFKFQFRNVE
YSSGRNKTF L CYVVEAQGKGGQVQASRGYLEDEHAAAHAE EAFNTILPAFDPALRYNVT
WYVSSSPCAACADRIIKTL SKTKNLRLLILVGR LFMWEEPEIQ AALKKLKEAGCKLRIMKP
QDFEYVWQNFVEQE EGESKAFQPWEDIQENFLY YEEKLADILK
[0 6 2 8]

ppAPOBEC-2 *Pongo pygmaeus*

MAQKEEAAAATEAASQNGEDLENLDDPEKLKELIELPPFEIVTGERLPANFFKFQFRNVE
YSSGRNKTF L CYVVEAQGKGGQVQASRGYLEDEHAAAHAE EAFNTILPAFDPALRYNVT
WYVSSSPCAACADRIIKTL SKTKNLRLLILVGR LFMWEELEIQ DALKKLKEAGCKLRIMKP
QDFEYVWQNFVEQE EGESKAFQPWEDIQENFLY YEEKLADILK
[0 6 2 9]

btAPOBEC-2 *Bos Taurus*

MAQKEEAAAAAEPASQNGEEVENLEDPEKLKELIELPPFEIVTGERLP AHYFKFQFRNVE
YSSGRNKTF L CYVVEAQSKGGQVQASRGYLEDEHATNHAE EAFNSIMPTFDPALRYMV
TWYVSSSPCAACADRIVKTLNKTKNLRLLILVGR LFMWEEPEIQ AALRKLKEAGCRLRIM
KPQDFEYIWQNFVEQE EGESKAFEPWEDIQENFLY YEEKLADILK
[0 6 3 0]

mAPOBEC-3 *Mus musculus*

MQPQRLGPRAGMGPFC LGCSHRKCYSPIRNLISQETFKFHKNLGYAKGRKDTFLCYEVT
RKDCDSPVSLHHGV FKNKDNIHAEICFLYWFHDKVLKVLSPREEFKITWYMSWSPCFEC
AEQIVRFLATHHNLSLDIFSSRLYNVQDPETQQNL CRLVQEGAQVAAMDLYEFKKCWKK
FVDNGGRRFRPWKRLLTNFRYQDSKLQEILRPCYISVPSSSSSTLSNICLTGLPETRFWV
EGRRMDPLSEEEFY SQFYNQRVKHL CYYHRMKPYLCYQLEQFNGQAPLKGCLLSEK GKQ
HAEILFLDKIRSMELSQVTITCYLTWSPCPNCAWQLAAFKRDRPD LILHIYTSRLYFHWK
RPFQKGLCSLWQSGILVDVMDLPQFTDCWTNFVNPKRPFWPWKGLEISRR TQRRLRRI
KESWGLQDLVND FGNLQLGPPMS
[0 6 3 1]

hAPOBEC-3A *Homo sapiens*

MEASPASGPRHLMDPHIFTSNFNNGIGRHKTYLCYEVERLDNGTSVKMDQHRGFLHNQA
KNLLCGFYGRHAELRFLDLVPSLQLDPAQIYRVTFWISWSPCFSWGCAGEVRAFLQENT
HVRLRIFAARIYDYDPLYKEALQMLRDAGAQVSIMTYDEFKHCWDTFVDHQQGCPFQPW
DGLDEHSQALSGLRLAILQNQGN

[0 6 3 2]

hAPOBEC-3B

Homo sapiens

MNPQIRNPMERMRYRDTFYDNFENEPILYGRSYTWLCYEVKIKRGRSNLLWDTGVFRGQV
YFKPQYHAEMCFLSWFCGNQLPAYKCFQITWVSWTPCPDCVAKLAEFLSEHPNVTLT
SAARLYYYWERDYRRALCRLSQAGARVTIMDYEEFAYCWENFVYNEGQQFMPWYKFDE
NYAFLHRTLKEILRYLMDPDTFTFNFNNDPLVLRRTQTYLCYEVERLDNGTWWLMDQH
MGFLCNEAKNLLCGFYGRHAELRFLDLVPSLQLDPAQIYRVTFWISWSPCFSWGCAGEV
RAFLQENTHVRLRIFAARIYDYDPLYKEALQMLRDAGAQVSIMTYDEFYCWDTFVYRQ
GCPFQPWDGLEEHSQALSGLRLAILQNQGN

[0 6 3 3]

hAPOBEC-3C

Homo sapiens

MNPQIRNPMKAMYPGTFFYFQFKNLWEANDRNETWLCFTVEGIKRRSVVSWKTGVFRNQ
VDSETHCHAERCFLSWFCDDILSPNTKYQVTWYTSWSPCPDCAGEVAEFLARHSNVNLT
IFTARLYYFQYPCYQEGRLSLSQEGVAVEIMDYEDFKYCWENFVYNDNEPFKPWKGLKT
NFRLLKRRLRESLQ

[0 6 3 4]

hAPOBEC-3D

Homo sapiens

MNPQIRNPMERMRYRDTFYDNFENEPILYGRSYTWLCYEVKIKRGRSNLLWDTGVFRGPV
LPKRQSNHRQEVYFRFENHAEMCFLSWFCGNRLPANRRFQITWVSWNPCLPCVVKVT
KFLAEHPNVTLTISAARLYYYRDRDWRWVLLRLHKAGARVKIMDYEDFAYCWENFVCN
EGQPFMPWYKFDDNYASLHRTLKEILRNPMEAMYPHIFYFHFNLLKACGRNESWLCFT
MEVTKHHSVAVFRKRGVFRNQVDPETHCHAERCFLSWFCDDILSPNTNYEVTWYTSWSP
CPECAGEVAEFLARHSNVNLTIFTARLCYFWDTDYQEGLCSLSQEGASVKIMGYKDFVSC
WKNFVYSDDEPFKPWKGLQTNFRLLKRRLREILQ

[0 6 3 5]

hAPOBEC-3F

Homo sapiens

MKPHFRNTVERMYRDTFSYNFYNRPILSRRNTVWLCYEVKTKGSPRPRLDAKIFRGQVY
SQPEHHAEMCFLSWFCGNQLPAYKCFQITWVSWTPCPDCVAKLAEFLAEHPNVTLTIS
AARLYYYWERDYRRALCRLSQAGARVKIMDDEEFAYCWENFVYSEGQPFMPWYKFDDN
YAFLHRTLKEILRNPMEAMYPHIFYFHFNLRKAYGRNESWLCFTMEVVKHHSPVSWK
RGVFRNQVDPETHCHAERCFLSWFCDDILSPNTNYEVTWYTSWSPCPECAGEVAEFLAR
HSNVNLTIFTARLYYFWDTDYQEGRLSLSQEGASVEIMGYKDFKYCWENFVYNDDEPFK
PWKGLKYNFLFLDSKLQEILE

[0 6 3 6]

hAPOBEC-3G

Homo sapiens

MKPHFRNTVERMYRDTFSYNFYNRPILSRRNTVWLCYEVKTKGSPRPPLDAKIFRGQVY
SELKYHPEMRFFHWFSKWRKLHRDQEYEVWYISWSPCTKCTRDMATFLAEDPKVTLT
IFVARLYYFWDPDYQEALRSLCQKRDGPRATMKIMNYDEFQHCWSKFVYSQRELFE
PWNNLPKYIILLHIMLGEILRHSMDPPTFTFNFNNEPWVRGRHETLYCYEVERMHNDTWV
LLNQRRGFLCNQAPHKHGFLEGRHAELCFLDVIPFWKLDLDQDYRVTCFTSWSPCFSCA
QEMAKFISKNHVSLCIFTARIYDDQGRCQEGRLTLAEAGAKISIMTYSEFKHCWDTFVD
HQQGCPFQPWDGLDEHSQDLSGLRLAILQNQEN

[0 6 3 7]

hAPOBEC-4

Homo sapiens

MEPIYEEYLANHGTIVKPYYWLSFSLDCSNCPYHIRTGEEARVSLTEFCQIFGFPYGTTF
PQTKHLTFYELKTSSGSLVQKGHASSCTGNYIHPEMMLFEMNGYLDIAIYNNDISIRHIL

10

20

30

40

50

YSNNSPCNEANHCCISKMYNFLTYPGITLSIYFSQLYHTEMDFPASAWNREALRSLASL
WPRVVLSPISGGIWHSVLHSFISGVSGSHVFQPILTGRALADRHNAYEINAITGVKPYFT
DVLLQTKRNPNTKAQEALSYPLNNAFPQGQFFQMPSGQLQPNLPPDLRAPVVFVLVPLR
DLPPMHMGQNPKNPRNIVRHLNMPQMSFQETKDLGRLPTGRSVEIVEITEQFASSKEAD
EKKKKKGKK

[0638]

mAPOBEC-4

Mus musculus

MDSLLMKQKKFLYHFKNVRWAKGRHETLYCYVVKRRDSATSCSLDFGHLRNKSGCHVE
LLFLRYISDWDLDPGRCYRVTWFTSWSPCYDCARHVAEFLRWNPNLRLIFTARLYFCE
DRKAEPEGLRRLHRAGVQIGIMTFKDYFYCWNTFVENRERTFKAWEGLHENSVRRLTRQL
RRILLPLYEVDDLRLDAFRMLGF

10

rAPOBEC-4

Rattus norvegicus

MEPLYEEYLTHSGTIVKPYWLSVSLNCTNCPYHIRTGEEARVPYTEFHQTFGFPWSTYP
QTKHLTFYELRSSGNLIQKGLASNCTGSHTHPESMLFERDGYLDSLIFHDSNIRHIILY
SNNSPCDEANHCCISKMYNFLMNYPEVTLSVFFSPLYHTENQFPTSAWNREALRGLASL
WPQVTLAISGGIWQSILETFVSGISEGLTAVRPFTAGRTLTDRYNAYEINCITEVKPYFT
DALHSWQKENQDQKVWAASENQPLHNTTPAQWQPDMSQDCRTPAVFMLVPYRDLPP
HVNPSPPQKPRTVVRHLNLTQLSASKVKALRKSPSGRPVKKEEARKGSTRSQEANETNKS
WKKQTLFIKSNICHLLEREQKKIGILSSWSV

[0639]

20

mfAPOBEC-4

Macaca fascicularis

MEPTYEEYLANHGTIVKPYWLSFSLDCSNCPYHIRTGEEARVSLTEFCQIFGFPYGT
PQTKHLTFYELKTSSGSLVQKGHASSCTGNYIHPEMMLFEMNGYLDIAIYNND
SIRHIILYCNNSPCNEANHCCISKVYNFLTYPGITLSIYFSQLYHTEMDFPASAWNREALRSLASL
WPRVVLSPISGGIWHSVLHSFVSGVSGSHVFQPILTGRALTDRYNAYEINAITGVKPF
FTDVLLHTKRNPNTKAQMALESYPLNNAFPGQSFQMTSGIPDLRAPVVFVLLPLRDLPP
MHMGQDPNKPRNIIRHLNMPQMSFQETKDLERLPTRRSVETVEITERFASSKQAEKTKKK
KGKK

[0640]

hAID *Homo sapiens*

30

MDSLLMNRKFLYQFKNVRWAKGRRETYLCYVVKRRDSATSFSLDFGYLRNKNKGCHVEL
LLFLRYISDWDLDPGRCYRVTWFTSWSPCYDCARHVADFLRGPNLRLIFTARLYFCEDR
KAEPEGLRRLHRAGVQIAIMTFKDYFYCWNTFVENHERTFKAWEGLHENSVRLSRQLRR
ILLPLYEVDDLRLDAFRTLGL

[0641]

clAID *Canis lupus familiaris*

MDSLLMKQRKFLYHFKNVRWAKGRHETLYCYVVKRRDSATSFSLDFGHLRNKSGCHVE
LLFLRYISDWDLDPGRCYRVTWFTSWSPCYDCARHVADFLRGYPNLRLIFAARLYFCED
RKAPEPEGLRRLHRAGVQIAIMTFKDYFYCWNTFVENREKTFKAWEGLHENSVRLSRQLR
RILLPLYEVDDLRLDAFRTLGL

40

[0642]

btAID *Bos Taurus*

MDSLLKKQRQFLYQFKNVRWAKGRHETLYCYVVKRRDSPTSFSLDFGHLRNKAGCHVEL
LLFLRYISDWDLDPGRCYRVTWFTSWSPCYDCARHVADFLRGYPNLRLIFTARLYFCDKE
RKAPEPEGLRRLHRAGVQIAIMTFKDYFYCWNTFVENHERTFKAWEGLHENSVRLSRQLR
RILLPLYEVDDLRLDAFRTLGL

[0643]

mAID *Mus musculus*

MDSLLMNRKFLYQFKNVRWAKGRRETYLCYVVKRRDSATSFSLDFGYLRNKNKGCHVEL
LLFLRYISDWDLDPGRCYRVTWFTSWSPCYDCARHVADFLRGPNLRLIFTARLYFCEDR

50

KAEPEGLRRLHRAGVQIAIMTFKDYFCWNTFVENHERTFKAWEGLHENSVRLSRQLRR
ILLPLYEVDDLRDAFRTLGL

[0644]

pmCDA-1 *Petromyzon marinus*

MAGYECVRVSEKLDFTFEFQFENLHYATERHRTYVIFDVKPKQSAGGRSRLWGYIINNP
NVCHAEILMSMIDRHLESNPGVYAMTWYMSWSPCANCSSKLNPNWLKNLLEEQGHRTL
MHFSRIYDRDREGDHRGLRGLKHVSNSFRMGVVGRAEVKECLAHEYVEASRRTLTLWLDTT
ESMAAKMRRKLCILVRCAGMRESGIPLHLFTLQTPLLSGRVVWWRV

[0645]

pmCDA-2 *Petromyzon marinus*

MELREVVDALASCVRHEPLSRVAFLRCFAAPSQKPRGTVILFYVEGAGRGVTGGHAVNY
NKQGTSIHAEVLLLSAVRAALLRRRRCEDGEEATRGTCTLHCYSTYSPCRDCVEYIQEFGAS
TGVRVVIHCCRLYELDVNRRRSEAEGVLRSLRGRDFRLMGPRDAIALLLGGRLANTAD
GESGASNAWVTETNVVEPLVDMTGFGDEDLHAQVQRNKQIREAYANYASAVSLMLGE
LHVDPAKFPFLAEFLAQTSEPSGTPRETRGRPRGASSRGPEIGRQRPADFERALGAYGLF
LHPRIVSREADREEIKRDLIVVMRKHNYQGP

[0646]

pmCDA-5 *Petromyzon marinus*

MAGDENVRVSEKLDFTFEFQFENLHYATERHRTYVIFDVKPKQSAGGRSRLWGYIINNP
NVCHAEILMSMIDRHLESNPGVYAMTWYMSWSPCANCSSKLNPNWLKNLLEEQGHRTL
MMHFSRIYDRDREGDHRGLRGLKHVSNSFRMGVVGRAEVKECLAHEYVEASRRTLTLWLD
TTESMAAKMRRKLCILVRCAGMRESGMPLHLFT

[0647]

yCD *Saccharomyces cerevisiae*

MVTGGMASKWDQKGMEDIAEYEEAALGYKEGGVPIGGCLINNKDGSVLGRGHNMRFQKGS
ATLHGEISTLENCGRLEGKVYKDTTLYTTLSPCDMCTGAIIMYGIPRCVVGENVNFKSKGE
KYLQTRGHEVVVDDERCKKIMKQFIDERPQDWFEDIGE

[0648]

rAPOBEC-1 (delta 177-186)

MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELKTCCLYEINWGGRHSIWRHTSQNTNK
HVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSRAITEFLSRYPHVTLFIYIARLYHH
ADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFVNYSNPSNEAHWPRYPHLWVRGLPPC
LNILRRKQPQLTFFTIALQSCHYQRLPPHILWATGLK

[0649]

rAPOBEC-1 (delta 202-213)

MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELKTCCLYEINWGGRHSIWRHTSQNTNK
HVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSRAITEFLSRYPHVTLFIYIARLYHH
ADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFVNYSNPSNEAHWPRYPHLWVRLYVLE
LYCIILGLPPCLNILRRKQPQHYQRLPPHILWATGLK

[0650]

Human AID:

MDSLLMNRRKFLYQFKNVRWAKGRRETYLCYVVKRRDSATSFSLDFGYLRNKNKGCHVELLFLRY
ISDWDLDPGRCYRVTWFTSWSPCYDCARHVADFLRGNPNLSLRIFTARLYFCEDRKAAEPEGLRR
LHRAGVQIAIMTFKDYFCWNTFVENHERTFKAWEGLHENSVRLSRQLRRILLPLYEVDDLRDA
FRTLGL

(下線：核局在化配列；二重下線：核外搬出シグナル)

[0651]

10

20

30

40

50

Mouse AID:

MDSLLMKQKKFLYHFKNVRWAKGRHETYLCYVVKRRDSATSCSLDFGHLRNKSGCHVELLFLRY
ISDWDLDPGRCYRVTWFTSWSPCYDCARHVAEFLRWNPNSLRIFTARLYFCEDRKAEP EGLRR
LHRAGVQIGIMTFKDYFYCWNTFVENRERTFKAWEGLHENSVR LTRQLRRILLPLYEVDDLRD
AFRMLGF

(下線：核局在化配列；二重下線：核外搬出シグナル)

【0652】

10

Canine AID:

MDSLLMKQKQKFLYHFKNVRWAKGRHETYLCYVVKRRDSATSFSLDFGHLRNKSGCHVELLFLRY
ISDWDLDPGRCYRVTWFTSWSPCYDCARHVADFLRGYPNLSLRIFAARLYFCEDRKAEP EGLRR
LHRAGVQIAIMTFKDYFYCWNTFVENREKTFKAWEGLHENSVR LSRQLRRILLPLYEVDDLRD
AFRTLGL

(下線：核局在化配列；二重下線：核外搬出シグナル)

【0653】

20

Bovine AID:

MDSLLKKQKQKFLYQFKNVRWAKGRHETYLCYVVKRRDSPTSFSLDFGHLRNKAGCHVELLFLRY
ISDWDLDPGRCYRVTWFTSWSPCYDCARHVADFLRGYPNLSLRIFTARLYFCDKERKAEP EGLR
RLHRAGVQIAIMTFKDYFYCWNTFVENHERTFKAWEGLHENSVR LSRQLRRILLPLYEVDDLRD
AFRTLGL

(下線：核局在化配列；二重下線：核外搬出シグナル)

【0654】

30

Rat AID:

MAVGSKPKAALVGPHWERERIWCFLCSTGLGTQQTGQTSRWLRPAATQDPVSPPRSLLMKQKRF
LYHFKNVRWAKGRHETYL CYVVKRRDSATSFSLDFGYLRNKSGCHVELLFLRYISDWDLDPGRC
YRVTWFTSWSPCYDCARHVADFLRGNPNSLRIFTARLTGWGALPAGLMSPARPSDYFYCWNTF
VENHERTFKAWEGLHENSVR LSRRLRRILLPLYEVDDLRDAFRTLGL

(下線：核局在化配列；二重下線：核外搬出シグナル)

【0655】

40

Mouse APOBEC-3

MGPFCLGCSHRKCYSPIRNLISQETFKFHFKNLGYAKGRKDTFLCYEVTRKDCDSPVSLHHGV
KNKDNIIHAEICFLYWFHDKVLKVLSPREEFKITWYMSWSPCFECAEQIVRFLATHHNLSLDIFS
SRLYNVQDPETQQNLCLRVQEGAQVAAMDLYEFKKCWKKFVDNGGRRFRPWKRLLTNFRYQDSK
LQEILRPCYIPVPSSSSSTLSNICLTGKLPETRFCVEGRRMDPLSEEEFYFYNQRVKHLCTY
HRMKPYLCYQLEQFNGQAPLKGCLLSEKKGQHAEILFLDKIRSMELSQVTITCYLTWSPCPNCA
WQLAAFKRDRPDILILHIYTSRLYFHWKRPFQKGLCSLWQSGILVDVMDLPQFTDCWTNFEVNPGR
PFWPWKGLEIISRRTQRRRLRIKESWGLQDLVNDFGNLQLGPPMS

10

(斜体：核酸編集ドメイン)

【0656】

Rat APOBEC-3:

MGPFCLGCSHRKCYSPIRNLISQETFKFHFKNRLRYAIDRKDTFLCYEVTRKDCDSPVSLHHGV
FKNKDNIIHAEICFLYWFHDKVLKVLSPREEFKITWYMSWSPCFECAEQVLRFLATHHNLSLDIF
SSRLYNIRDPENQQNLCLRVQEGAQVAAMDLYEFKKCWKKFVDNGGRRFRPWKKLLTNFRYQDS
KLQEILRPCYIPVPSSSSSTLSNICLTGKLPETRFCVERRRVHLLSEEEFYFYNQRVKHLCTY
YHGKPYLCYQLEQFNGQAPLKGCLLSEKKGQHAEILFLDKIRSMELSQVIITCYLTWSPCPNC
AWQLAAFKRDRPDILILHIYTSRLYFHWKRPFQKGLCSLWQSGILVDVMDLPQFTDCWTNFEVNPGR
PFWPWKGLEIISRRTQRRRLHRIKESWGLQDLVNDFGNLQLGPPMS

20

(斜体：核酸編集ドメイン)

【0657】

Rhesus macaque APOBEC-3G:

MVEPMDPRTFVSNFNNRPILSGLNTVWLCCEVKTDPSPGPPLDAKIFQGVYSKAKYHPEMRFL
RWFHKWRQLHHDQEYKVTWYVSWSPCTRCANSVATFLAKDPKVTLTIFVARLYYFWKPDYQQAL
RILCQKRGGPHATMKIMNYNEFQDCWNKFVDGRGKPFKPRNNLPKHYTLLQATLGELLRHLMDP
GTFTSNFNNKPWVSGQHETYLCKVERLHNDTWVPLNQHRGFLRNQAPNIHGFPKGRHAELCFL
DLIPFWKLDGQQYRVTCFTSWSPCFSCAQEMAKFISNNEHVSLCIFAARIYDDQGRYQEGRLAL
HRDGAKIAMMNYSEFEYCWDTFVDRQGRPFQPWDGLDEHSQALSGRLRAI

30

(斜体：核酸編集ドメイン；下線：細胞質局在化シグナル)

【0658】

Chimpanzee APOBEC-3G:

40

MKPHFRNPVERMYQDTFSDNFYNRPILSHRNTVWLCYEVKTKGPSRPPLDAKIFRGQVYSKLKY
HPEMRFFHWFSKWRKLHRDQEYEV TWYISWSPCTKCTRDVATFLAEDPKVTLTIFVARLYYFWD
PDYQEALRSLCQKRDGPRATMKIMNYDEFQHCWSKFVYSQRELFEPWNNLPKYYILLHIMLGEI
LRHSMDPPTFTSNFNNELWVRGRHETYLCEVERLHNDTWVLLNQRRGFLCNQAPHKHGFLEGR
HAELCFLDVIPFWKLDLHQDYRVTCFTSWSPCFSCAQEMAKFISNNKHVSLCIFAARIYDDQGR
CQEGLRTLAKAGAKISIMTYSEFKHCWDTFVDHQGCPFPQPDGLEEHSQALSGRLRAILQNQGN

10

(斜体：核酸編集ドメイン；下線：細胞質局在化シグナル)

【0659】

Green monkey APOBEC-3G:

MNPQIRNMVEQMEPDIFVYYFNRPILSGRNTVWLCYEVKTKDPSGPPLDANIFQGKLYPEAKD
HPEMKFLHWFRKWRQLHRDQEYEV TWYVSWSPCTRCANSVATFLAEDPKVTLTIFVARLYYFWK
PDYQQALRILCQERGGPHATMKIMNYNEFQHCWNEFVDGQGKPKPRKNLPKHYTLLHATLGEL
LRHVMDPGTFTSNFNKPVWSGQRETYLCYKVERSHNDTWVLLNQHRGFLRNQAPDRHGFPKGR
HAELCFLDLIPFWKLDDQYRVTCFTSWSPCFSCAQKMAKFISSNNKHVSLCIFAARIYDDQGR
QEGLRTLHRDGAKIAVMNYSEFEYCWDTFVDRQGRPFQPDGGLDEHSQALSGRLRAI

20

(斜体：核酸編集ドメイン；下線：細胞質局在化シグナル)

【0660】

Human APOBEC-3G:

MKPHFRNTVERMYRDTFSYNFYNRPILSRNTVWLCYEVKTKGPSRPPLDAKIFRGQVYSELKY
HPEMRFFHWFSKWRKLHRDQEYEV TWYISWSPCTKCTRDMATFLAEDPKVTLTIFVARLYYFWD
PDYQEALRSLCQKRDGPRATMKIMNYDEFQHCWSKFVYSQRELFEPWNNLPKYYILLHIMLGEI
LRHSMDPPTFTFNFNNEPWWVRGRHETYLCEVERMHNDTWVLLNQRRGFLCNQAPHKHGFLEGR
HAELCFLDVIPFWKLDLDQDYRVTCFTSWSPCFSCAQEMAKFISKNKHVSLCIFTARIYDDQGR
CQEGLRTLAEAGAKISIMTYSEFKHCWDTFVDHQGCPFPQPDGGLDEHSQDLSGRLRAILQNQEN

30

(斜体：核酸編集ドメイン；下線：細胞質局在化シグナル)

【0661】

Human APOBEC-3F:

MKPHFRNTVERMYRDTFSYNFYNRPILSRNTVWLCYEVKTKGPSRPRLDAKIFRGQVYSQPEH
HAEMCFLSWFCGNQLPAYKCFQITWVFSWTPCPDCVAKLAEF LAEHPNVTLTISAARLYYWER
DYRRALCRLSQAGARVKIMDDEEFAYCWENFVYSEGQPFMPWYKFDDNYAFLHRTLKEILRNPM
EAMYPHIFYFHFKNLKAYGRNESWLCFTMEVVKKHSPVSWKRGVFRNQVDPETHCHAERCFLS
WFCDDILSPNTNYEV TWYTSWSPCECAGEVAEFLARHSNVNLTIFTARLYYFWD TDYQEGLRS
LSQEGASVEIMGYKDFKYCWENFVYNDDEPFKPKWGLKYNFLFLDSKLQEILE

40

(斜体：核酸編集ドメイン)

50

【0662】

Human APOBEC-3B:

MNPQIRNPMERMYRDTFYDNFENEPILYGRSYTWLCYEVKIKRGRSNLLWDTGVFRGQVYFKPQ
YHAEMCFLSWFCGNQLPAYKCFQITWVFSWTPCPDCVAKLAEFLSEHPNVTLTISAARLYYYWE
RDYRRALCRLSQAGARVTIMDYEEFAYCWENFVYNEGQQFMPWYKFDENYAFHLRTLKEILRYL
MDPDTFTFNFNNDPLVLRRTYLYCYEVERLDNGTWVLMQHMGLCNEAKNLLCGFYGRHAEL
RFLDLVPSLQLDPAQIYRVTFISWSPCFSWGCAEVRAFLQENTHVRLRIFAARIYDYDPLYK
EALQMLRDAGAQVSIMTYDEFEYCWDTFVYRQGCPFPQPDGEEHSQALSGRLRAILQNQGN

10

(斜体：核酸編集ドメイン)

【0663】

Rat APOBEC-3B:

MQPQGLGNAGMGPVCLGCSHRRPYSPIRNPLKKLYQQTFYFHFKNVRYAWGRKNNFL
CYEVNGMDCALPVPLRQGVFRKQGHHAELCFIYWFHDKVLRVLSPMEEFKVTWYMSW
SPCSKCAEQVARFLAAHRNLSLAIFSSRLYYYLRNPNYQQKLCRLIQEGVHVAAMDLPF
KKCWNFVDNDGQPFRPWMRLRINFSFYDCKLQEISRMNLLREDVFYLQFNNSHRVKP
VQNRYYRRKSYLCYQLERANGQEPLKGYLLYKKGEQHVEILFLEKMRSMEELSQVRITCYL
TWSPCPNCARQLAAFKKDHDPDLILRIYTSRLYFWRKKFQKGLCTLWRSIHDVMDLPQ
FADCWTNFVNPRPFRPWNELEKNSWRIQRRRLRIKESWGL

20

【0664】

Bovine APOBEC-3B:

DGWEVAFRSGTVLKAGVLGVSMTEGWAGSGHPGQACVWTPGTRNTMNLLREVLFKQ
QFGNQPRVPAPYYRRKTYLCYQLKQRNDLTLDRCFRNKKQRHAERFIDKINSLDLNPS
QSYKIICYITWSPCPNCANELVNFITRNNHLKLEIFASRLYFHWIKSFKMGLQDLQNAGIS
VAVMTHTEFEDCWEQFVDNQSRFPQPDWKLEQYSASIRRLRILTAI

【0665】

Chimpanzee APOBEC-3B:

MNPQIRNPMEMYQRTFYYNFENEPILYGRSYTWLCYEVKIRRGHSNLLWDTGVFRGQ
MYSQPEHHAEMCFLSWFCGNQLSAYKCFQITWVFSWTPCPDCVAKLAKFLAHPNVTL
TISAARLYYYWERDYRRALCRLSQAGARVKIMDDEEFAYCWENFVYNEGQPFMPWYKF
DDNYAFHLRTLKEIIRHLMDDPDTFTFNFNNDPLVLRHQTLYCYEVERLDNGTWVLMQ
QHMGLCNEAKNLLCGFYGRHAELRFLDLVPSLQLDPAQIYRVTFISWSPCFSWGCA
QVRAFLQENTHVRLRIFAARIYDYDPLYKEALQMLRDAGAQVSIMTYDEFEYCWDTFVY
RQGCPFPQPDGEEHSQALSGRLRAILQVRASSLCMVPHRPPPPQSPGPCPLPCSEPPL
GSLLPTGRPAPSLPFLLTASFSPPPASLPPLPSLSLSPGHLVPSPFHSLSLSCSIQPPCSSRI
RETEGWASVSKEGRDLG

30

【0666】

Human APOBEC-3C:

MNPQIRNPMKAMYPGTFYFQFKNLWEANDRNETWLCFTVEGIKRRSVVSWKTGVFRNQVDSETH
CHAERCFLSWFCDDILSPNTKYQVTWYTSWSPCPDCAGEVAEFLARHSNVNLTIFTARLYYFQY
PCYQEGRLSLSQEGVAVEIMDYEDFKYCWENFVYNDNEPFKPKWGLKTNFRLLKRLRESLQ

40

(斜体：核酸編集ドメイン)

【0667】

Gorilla APOBEC-3C3C

50

MNPQIRNPMKAMYPGTFYFQFKNLWEANDRNETWLCFTVEGIKRRSVVSWKTGVFRNQVDSETH
CHAERCFLSWECDDILSPNTNYQVTWYTSWSPCECAGEVAEFLARHSNVNLTIFTARLYYFQD
TDYQEGRLSLSQEGVAVKIMDYKDFKYCWENFVYNDDEPFKPKWGLKYNFRFLKRRLQEILE

【0668】

Human APOBEC-3A:

MEASPASGPRHLMDPHIFTSNFNNGIGRHKTYLCYEVERLDNGTSVKMDQHRGFLHNQAKNLLC
GFYGRHAELRFLDLVPSLQLDPAQIYRVTWFIWSWPCFSWGCAGEVRAFLQENTHVRLRIFAAR
IYDYDPLYKEALQMLRDAGAQVSIMTYDEFKHCWDTFVDHQGCPFQPWDGLDEHSQALSGRLRA
ILQNQGN

10

(斜体：核酸編集ドメイン)

【0669】

Rhesus macaque APOBEC-3A:

MDGSPASRPRHLMDPNTFTFNFNNDLSVRGRHQTYLCYEVERLDNGTWVPMDERRGFLCNKAKN
VPCGDYGCHEVLRFLCEVPSWQLDPAQTYRVTWFIWSWPCFRRGCAGQVRVFLQENKHVRLRIF
AARIYDYDPLYQEALRTL RDAGAQVSIMTYEEFKHCWDTFVDRQGRPFQPWDGLDEHSQALSGR
LRAILQNQGN

20

(斜体：核酸編集ドメイン)

【0670】

Bovine APOBEC-3A3A:

MDEYTFTENFNQGWPSKTYLCYEMERLDGDATIPLDEYKGFVRNKGLDQPEKPCHAEIYFLGK
IHSWNLDNRNQHYRLTCFISWSPCYDCAQKLTTFLKENHHISLHILASRIYTHNRFQCHQSGLCE
LQAAGARITIMTFEDFKHCWETFVDHKGKPFQPWEGLVKSQALCTELQAILKTQON

30

(斜体：核酸編集ドメイン)

【0671】

Human APOBEC-3H:

MALLTAETFRLQFNNKRRLRRPYYPRKALLCYQLTPQNGSTPTRGYFENKKKCHAEICFINEIK
SMGLDETQCYQVTCYLTWSPCSCAWELVDFIKAHDHLNLGIFASRLYYHWCKPQQKGLRLLCG
SQVPVEVMGFPKFADCWENFVDHEKPLSFNPYKMLEELDKNSRAIKRRLERIKIPGVRAQGRYM
DILCDAEV

40

(斜体：核酸編集ドメイン)

【0672】

Rhesus macaque APOBEC-3H:

MALLTAKTFSLQFNNKRRVNKPYYPRKALLCYQLTPQNGSTPTRGHLKNKKKDHAEIFI
NLIKSMGLDETQCYQVTCYLTWSPCPCAGELVDFIKAHRHLNLRFASRLYYHWRPNY
QEGLLLLCGSQVPVEVMGLPEFTDCWENFVDHKEPPSFNPSEKLEELDKNSQAIKRRLER

50

IKSRSDVLENGLRSLQLGPVTPSSSIRNSR

【0673】

Human APOBEC-3D:

MNPQIRNPMERMYRDTFYDNFENEPILYGRSYTWLCYEVKIKRGRSNLLWDTGVFRGPVLPKRQ
SNHRQEVYFRFENHAEMCFLSWFCGNRLPANRRFQITWFSWNPCLPCVVKVTKFLAEHPNVTL
TISAARLYYYRDRDWRWVLLRLHKAGARVKIMDYEDFAYCWENFVCNEGQPFMPWYKFDDNYAS
LHRTLKEILRNPMEMYPHIFYFHFKNLLKACGRNESWLCFTMEVTKHHSVFRKRGVFRNQVD
PETHCHAERCFLSWFCDDILSPNTNYEVTWYTSWSPCECAGEVAEFLARHSNVNLTIFTARLC
YFWDTDYQEGLCSLSQEGASVKIMGYKDFVSCWKNFVYSDDEPFKPKWGLQTNFRLLKRRLREI
LQ

10

(斜体：核酸編集ドメイン)

【0674】

Human APOBEC-1:

MTSEKGPSTGDPTLRRRIEPWEFDVFDPRELRKEACLLYEIKWGMSRKIWRSSGKNNTN
HVEVNFIIKFTSERDFHPSMSCSITWFLSWSPCWECSQAIREFLSRHPGVTLVIYVARLF
WHMDQQNRQGLRDLVNSGVTIQIMRASEYYHCWRNFVNYPGDEAHWPQYPPLWMM
LYALELHCCIISLPPCLKISRRWQNHLTFFRLHLQNCHYQTIPPHILLATGLIHPSVAWR
【0675】

20

Mouse APOBEC-1:

MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELRKETCLLYEINWGGRHSVWRHTSQNTSN
HVEVNFLEKFTTERYFRPNTRCSITWFLSWSPCGECSRAITEFLSRHPYVTLFIYIARLYH
HTDQRNRQGLRDLISSGVTIQIMTEQEYCYCWRNFVNYPSPNEAYWPRYPHLWVKLYV
LELYCIILGLPPCLKILRRKQPQLTFFTITLQTCHYQRIPPHLLWATGLK
【0676】

Rat APOBEC-1:

30

MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELRKETCLLYEINWGGRHSIWRHTSQNTNK
HVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSRAITEFLSRYPHVTLFIYIARLYHH
ADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFVNYSNEAHWPYPHLWVRLYVLE
LYCIILGLPPCLNILRRKQPQLTFFTIALQSCHYQRLPPHILWATGLK
【0677】

Human APOBEC-2:

MAQKEEAAVATEAASQNGEDLENLDDPEKLKELIPLPFEIVTGERLPANFFKFQFRNVE
YSSGRNKTFCLCYVVEAQKGGQVQASRGYLEDEHAAAHAEAAFFNTILPAFDPALRYNVT
WYVSSSPCAACADRIIKTLTKNLRLILVGRLFMWEEPEIQAAALKKLKEAGCKLRIMKP
QDFEYVWQNFVEQEEGESKAFQPWEDIQENFLYYEEKLADILK
【0678】

40

Mouse APOBEC-2:

MAQKEEAAEAAAPASQNGDDLENLEDPEKLKELIDLPPFEIVTGVRLPVNFFKFQFRNVE
YSSGRNKTFCLCYVVEVQSKGGQAQATQGYLEDEHAGAHAEAAFFNTILPAFDPALKYNVT
WYVSSSPCAACADRILKTLTKNLRLILVSRLFMWEEPEVQAALKKLKEAGCKLRIMK
PQDFEYIWQNFVEQEEGESKAFEPWEDIQENFLYYEEKLADILK
【0679】

Rat APOBEC-2:

MAQKEEAAEAAAPASQNGDDLENLEDPEKLKELIDLPPFEIVTGVRLPVNFFKFQFRNVE
YSSGRNKTFCLCYVVEAQKGGQVQATQGYLEDEHAGAHAEAAFFNTILPAFDPALKYNVT

50

WYVSSSPCAACADRILKTLSTKNLRLLLILVSRLFMWEEPEVQAALKKLKEAGCKLRIMK
PQDFEYLWQNFVEQEEGESKAFEPWEDIQENFLYYEEKLADILK

【0680】

Bovine APOBEC-2:

MAQKEEAAAAAEPASQNGEEVENLEDPEKLEKELIELPPFEIVTGERLPAHYFKFQFRNVE
YSSGRNKTFCLCYVVEAQSKGGQVQASRGYLEDEHATNHAEAAFFNSIMPTFDPALRYMV
TWYVSSSPCAACADRIVKTLNKTKNLRLLLILVGRLFMWEEPEIQAALRKLKEAGCRLRIM
KPQDFEYIWQNFVEQEEGESKAFEPWEDIQENFLYYEEKLADILK

【0681】

Petromyzon marinus CDA1 (pmCDAL):

MTDAEYVRIHEKLDIYTFKKQFFNNKKSVSHRCYVLFELKRRGERRACFWGYAVNKPQS
GTERGIHAEIFSIRKVEEYLRDNPQGFTINWYSSWSPCADCAEKILEWYNQELRGNGHTL
KIWACKLYYEKNARNQIGLWNLRDNGVGLNVMVSEHYQCCRKIFIQSSHNQLNENRWL
EKTLKRAEKRRSELSFMIQVKILHTTKSPAV

【0682】

Human APOBEC3G chain A:

MDPPTFTFNFNNEPWWGRHETLYCYEVERMHNDTWVLLNQRRGFLCNQAPHKHGFLE
GRHAELCFLDVIPFWKLDLDQDYRVTCFTSWSPCFSCAQEMAKFISKKNKHVSLCIFTARI
YDDQGRQCQGLRTLAEAGAKISFTYSEFKHCWDTFVDHQGCPFPWDGLDEHSQDLSGR
LRAILQ

【0683】

本開示の態様によるCas12のアミノ酸配列内で内部的に融合することができる他の例示的なデアミナーゼを以下に提供する。いくつかの実施形態では、それぞれの配列の活性ドメイン、例えば、局在化シグナル（例えば、核局在化配列、核外輸送シグナル、または細胞質局在化シグナル）を伴わないドメインを使用できることを理解されたい。

【0684】

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。

【0685】

[シチジンデアミナーゼ]

一実施形態では、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。いくつかの実施形態では、本明細書で提供されるシチジンデアミナーゼは、シトシンまたは5-メチルシトシンをウラシルまたはチミンに脱アミノ化することができる。いくつかの実施形態では、本明細書で提供されるシチジンデアミナーゼは、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し得る。いくつかの実施形態において、シチジンデアミナーゼは、原核生物に由来する。いくつかの実施形態において、シチジンデアミナーゼは、細菌に由来する。いくつかの実施形態において、シチジンデアミナーゼは、哺乳動物（例えば、ヒト）に由来する。

【0686】

いくつかの実施形態において、シチジンデアミナーゼは、本明細書に記載のシチジンデアミナーゼアミノ酸配列のいずれか1つ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。

【0687】

本発明の融合タンパク質は、核酸編集ドメインを含む。いくつかの実施形態において、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。いくつかの実施形態

10

20

30

40

50

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。いくつかの実施形態において、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。いくつかの実施形態において、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である。いくつかの実施形態において、デアミナーゼは、APOBEC1デアミナーゼである。いくつかの実施形態において、デアミナーゼは、APOBEC2デアミナーゼである。いくつかの実施形態において、デアミナーゼは、APOBEC3デアミナーゼである。いくつかの実施形態において、デアミナーゼは、APOBEC3Aデアミナーゼである。いくつかの実施形態において、デアミナーゼは、APOBEC3Bデアミナーゼである。いくつかの実施形態において、デアミナーゼは、APOBEC3Cデアミナーゼである。いくつかの実施形態において、デアミナーゼは、APOBEC3Dデアミナーゼである。いくつかの実施形態において、デアミナーゼは、APOBEC3Eデアミナーゼである。いくつかの実施形態において、デアミナーゼは、APOBEC3Fデアミナーゼである。いくつかの実施形態において、デアミナーゼは、APOBEC3Gデアミナーゼである。いくつかの実施形態において、デアミナーゼは、APOBEC3Hデアミナーゼである。いくつかの実施形態において、デアミナーゼは、APOBEC4デアミナーゼである。いくつかの実施形態において、デアミナーゼは、活性化誘導デアミナーゼ（AID）である。いくつかの実施形態において、デアミナーゼは、脊椎動物のデアミナーゼである。いくつかの実施形態において、デアミナーゼは、無脊椎動物のデアミナーゼである。いくつかの実施形態において、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、またはマウスのデアミナーゼである。いくつかの実施形態において、デアミナーゼは、ヒトデアミナーゼである。いくつかの実施形態において、デアミナーゼは、ラットデアミナーゼ、例えば、rAPOBEC1である。いくつかの実施形態において、デアミナーゼは、*Petromyzon marinus*シチジンデアミナーゼ1（pmCDA1）である。いくつかの実施形態では、デアミナーゼはヒトAPOBEC3Gである。いくつかの実施形態において、デアミナーゼは、ヒトAPOBEC3Gのフラグメントである。いくつかの実施形態において、核酸編集ドメインは、本明細書に記載の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。

【0688】

〔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。ある実施形態では、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。

【0689】

ある実施形態において、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。ある実施形態では、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。ある実施形態では、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。このような実施形態では、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。

【0690】

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:eaao4774 (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）。

10

【0691】

さらに、いくつかの実施形態では、Gamタンパク質が、塩基エディターのN末端に融合され得る。いくつかの実施形態において、Gamタンパク質は、塩基エディターのC末端に融合され得る。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では、Gamを使用してDSBの自由端を結合することにより、塩基編集プロセスの際のインデル形成を減らすことができる。いくつかの実施形態において、174残基のGamタンパク質が、塩基エディターのN末端に融合される。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:eaao4774 (2017)参照。いくつかの実施形態では、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エディタードメインの長さを変え得る。たとえば、少なくとも1つのドメインの少なくとも1つのアミノ酸の削除が、塩基エディターの長さを短くし得る。別のケースでは、1つまたは複数の突然変異は、野生型ドメインと比較してドメインの長さを変更しない。たとえば、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。

20

【0692】

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。ある実施形態では、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アルファ、ベータ、ガンマ、デルタ、イプシロン、ガンマ、イータ、イオタ、カッパ、ラムダ、ミュー、またはニュー成分である。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。

30

40

【0693】

【その他の核酸塩基エディター】

本発明は、モジュラー式マルチエフェクター核酸塩基エディターを提供し、ここでは、実質的にあらゆる当技術分野で知られる核酸塩基エディターが、本明細書中に記載された融合タンパク質に挿入され得、または、シチジンデアミナーゼもしくはアデノシンデアミナーゼの代わりに交換され得る。1つの実施形態において、本発明は、脱塩基型核酸塩基エディタードメインを含むマルチエフェクター核酸塩基エディターを特徴とする。脱塩基

50

型核酸塩基エディターは当技術分野で公知であり、例えば、Kavli et al., EMBO J. 15:3442-3447, 1996によって記述されており、この文献は参照により本明細書に組み込まれる。

【0694】

一実施形態では、マルチエフェクター核酸塩基エディターは、以下のドメインA～C、A～D、またはA～Eを含み：

$\text{NH}_2\text{-[A-B-C]-COOH}$ 、

$\text{NH}_2\text{-[A-B-C-D]-COOH}$ 、または

$\text{NH}_2\text{-[A-B-C-D-E]-COOH}$

ここで、AおよびC、またはA、C、およびEは、それぞれ、以下のうちの1つまたは複数を含む：アデノシンデアミナーゼドメインまたはその活性断片、シチジンデアミナーゼドメインまたはその活性断片、DNAグリコシラーゼドメインまたはその活性断片；ここで、BまたはBとDは、それぞれ、核酸配列特異的結合活性を有する1つまたは複数のドメインを含む。

【0695】

一実施形態では、マルチエフェクター核酸塩基エディターは、 $\text{NH}_2\text{-[A}_n\text{-B}_0\text{-C}_n\text{]-COOH}$

、 $\text{NH}_2\text{-[A}_n\text{-B}_0\text{-C}_n\text{-D}_0\text{]-COOH}$ 、または

$\text{NH}_2\text{-[A}_n\text{-B}_0\text{-C}_p\text{-D}_0\text{-E}_q\text{]-COOH}$

を含み、ここで、AおよびC、またはA、C、およびEは、それぞれ、以下のうちの1つまたは複数を含む：アデノシンデアミナーゼドメインまたはその活性断片、シチジンデアミナーゼドメインまたはその活性断片、およびDNAグリコシラーゼドメインまたはその活性断片；ここで、nは整数1、2、3、4、または5であり、pは整数0、1、2、3、4、または5であり、BまたはBとDはそれぞれ、核酸配列特異的結合活性を有するドメインを含み、oは整数1、2、3、4、または5である。

【0696】

[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、二本鎖または一本鎖の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）とガイドポリ核酸（例えばgRNA）とを含む塩基エディターシステムと接触させ、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。

【0697】

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1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。

【0698】

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。

【0699】

本明細書で提供されるのは、塩基エディターシステムを使用して核酸塩基を編集するためのシステム、組成物、および方法である。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) 該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組み合わせられたガイドポリヌクレオチド(例えば、ガイドRNA)とを含む。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。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。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。いくつかの実施形態において、核酸塩基編集ドメインは、デアミナーゼドメインである。いくつかの実施形態において、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し得る。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。

【0700】

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:eaa04774 (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)。

【0701】

いくつかの実施形態では、単一のガイドポリヌクレオチドを利用して、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。ある実施形態では、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。

【0702】

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、ポリヌクレオチドと結合し、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。

10

【0703】

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。ある態様において、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。

20

30

【0704】

ある態様において、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)であり得る。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り得る。ある態様において、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、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

40

50

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。

【0705】

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(BER)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カーゼ活性を含む。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。

【0706】

いくつかの実施形態では、本方法は、正準的(例えばNGG)PAMサイトを必要としない。いくつかの実施形態において、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。

【0707】

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。ある実施形態では、標的は4塩基領域内にあり得る。ある実施形態では、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。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:eaao4774 (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。

【0708】

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。いくつかの実施形態において、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。いくつかの実施形態において、標的ウィンドウは、脱アミノ化ウィンドウである。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定義された領域であり得る。いくつかの実施形態において、脱アミノ化ウィンドウは、2、3、4、5、6、7、8、9、または10塩基の領域内にある。いくつかの実施形態において、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、または25塩基上流である。

【0709】

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。

【0710】

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处配列、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、ソフトタグ (例ソフトタグ1、ソフトタグ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。

【0711】

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、アデノシンデアミナーゼ)、ウラ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、およびレポーター遺伝子配列が挙げられる。

【0712】

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トグ (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。

【0713】

塩基エディターまたは塩基エディターシステムは、本明細書に記載されている任意の塩基エディターであり得る。いくつかの実施形態では、塩基エディターは、本明細書に記載されるように、アデノシンデアミナーゼモノマーまたはアデノシンデアミナーゼダイマーを含む。いくつかの実施形態では、塩基エディターは、デアミナーゼを有するCas9、例えばCas9ポリペプチドの可撓性ループに挿入されたアデノシンデアミナーゼを含む。いくつかの実施形態では、塩基エディターは、それぞれがCas9分割インティン融合タンパク質のN末端フラグメントおよびC末端フラグメントをコードする2つのポリヌクレオチドの発現によって生成される。塩基エディターは、RNP、ベクター、ウイルスベクター、またはmRNA送達などの核酸を介して宿主細胞に送達され得る。いくつかの実施形態において、塩基エディターをコードするポリヌクレオチド構築物は、構造pMRNA-Trilink-ISLAY3-mo noTadA-ABE7.10(V82S)-MQKFRAER 120A BbsI、pMRNA-Trilink-ISLAY3-ABE7.10(V82S, Y147T, Q154S)-MQKFRAER 120A BbsI、またはpMRNA-Trilink-ISLAY3-ABE7.10(V82T, Y147T, Q154S)-MQKFRAER 120A BbsIを含む。ガイドRNAは化学的に修飾され得る。いくつかの実施形態において、ガイドRNAは、2'-O-メチル (2'-OMe)、2'-デオキシ (2'-H)、2'-O-C1-3アルキル-O-C1-3アルキル、例えば2'-メトキシエチル (「2'-MOE」)、2'-フルオロ (「2'-F」)、2'-アミノ (「2'-NH₂」)、2'-アラビノシル (「2'-アラビノ」) ヌクレオチド、2'-F-アラビノシル (「2'-F-アラビノ」) ヌクレオチド、2'-ロックド核酸 (「LNA」) ヌクレオチド、2'-アンロックド核酸 (「ULNA」) ヌクレオチド、L型の糖 (「L-糖」)、4'-チオリボシルヌクレオチド、または本明細書に記載の任意の化学修飾などの、1つ以上の化学修飾核酸塩基を含む。いくつかの実施形態において、ガイドRNAは、ホスホロチオエート「P(S)」(P(S))、ホスホノアセテート「PACE」(P(CH₂COO-))などのホスホノカルボキシレート (P(CH₂)_nCOOR)、例えばチオホスホノアセテート「thioPACE」((S)P(CH₂)_nCOO-)などのチオホスホノカルボキシレート ((S)P(CH₂)_nCOOR)、例えばメチルホスホネート-P(CH₃)などのアルキルホスホネート (P(C1-3アルキル)、ボラノホスホネート (P(BH₃))、およびホスホロジチオエート(P(S)₂)などの、ヌクレオチド間結合修飾を含む。いくつかの実施形態において、ガイドRNAは、2-チオウラシル (「2-チオU」)、2-チオシトシン (「2-チオC」)、4-チオウラシル (「4-thioU」)、6-チオグアニン (「6-thioG」)、2-アミノアデニン (「2-aminoA」)、2-アミノプリン、シュードウラシル、ヒポキサンチン、7-デアザグアニン、7-デアザ-8-アザグアニン、7-デアザアデニン、7-デアザ-8-アザアデニン、5-メチルシトシン (「5-メチルC」)、5-メチルウラシル (「5-メチルU」)、5-ヒドロキシメチルシトシン、5-ヒドロキシメチルウラシル、5,6-デヒドロウラシル、5-プロピニルシトシン、5-プロピニルウラシル、5-エチニルシトシン、5-エチニルウラシル、5-アリルウラシル (「5-アリルU」)、5-アリルシトシン (「5-アリルC」)、5-アミノアリルウラシル (「5-アミノアリルU」)、5-アミノアリル-シトシン (「5-アミノアリルC」)、脱塩基ヌクレオチド、Z塩基、P塩基、非構造化核酸 (「UNA」)、イソグアニン (「isoG」)、イソシトシン (「isoC」) などの核酸塩基化学修飾を含む。いくつかの実施形態において、ガイドRNAは、ヌクレオチド糖、核酸塩基、ホスホジエステル結合および／またはヌクレオチドリノ酸に対する1つ以上の同位体修飾を含む。このような修飾には、1つまたは複数の¹⁵N、¹³C、¹⁴C、重水素、³H、³²P、¹²⁵I、¹³¹I原子もしくは他の原子またはそれらの元素を含むヌクレオチドが含まれる。

【0714】

一部の実施形態において、アデノシン塩基エディター（ABE）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、ABEは、BE3のAPOBEC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。

【0715】

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。

【0716】

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I156F)を伴うABE 2.3である。

【0717】

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ABE4.3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。

【0718】

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スセット(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6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は第六世代ABEである。一部の実施形態では、ABEは、以下の表6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6に示されるように、ABE7.1、ABE7.2、ABE7.3、ABE7.4、ABE7.5、ABE7.6、ABE7.7、ABE7.8、ABE7.9、またはABE7.10である。

【0719】

表6．ABEの遺伝子型

10

20

30

40

50

【表 6】

	23	26	36	37	48	49	51	72	84	87	105	108	123	125	142	145	147	152	155	156	157	161
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
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

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	161
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
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

20

30

【0720】

いくつかの実施形態において、塩基エディターは、アデノシン塩基エディターである。

いくつかの実施形態において、アデノシン塩基エディターは、第8世代ABE（ABE8）である。いくつかの実施形態では、ABE8は、TadA*8バリエントを含む。いくつかの実施形態において、ABE8は、TadA*8バリエントを含有する単量体構築物（「ABE8.x-m」）を有する。いくつかの実施形態において、ABE8は、Y147T変異を有するTadA*7.10（TadA*8.1）を含有する単量体構築物を有するABE8.1-mである。いくつかの実施形態において、ABE8は、Y147R変異を有するTadA*7.10（TadA*8.2）を含有する単量体構築物を有するABE8.2-mである。いくつかの実施形態において、ABE8は、ABE8.3-mであり、これは、Q154S変異を有するTadA*7.10（TadA*8.3）を含有する単量体構築物を有する。いくつかの実施形態において、ABE8は、ABE8.4-mであり、これは、Y123H変異を有す

40

50

るTadA*7.10 (TadA*8.4) を含有する単量体構築物を有する。いくつかの実施形態において、ABE8は、V82S変異を有するTadA*7.10 (TadA*8.5) を含有する単量体構築物を有するABE8.5-mである。いくつかの実施形態において、ABE8は、ABE8.6-mであり、これは、T166R変異を有するTadA*7.10 (TadA*8.6) を含む単量体構築物を有する。いくつかの実施形態において、ABE8は、ABE8.7-mであり、これは、Q154R突然変異を有するTadA*7.10 (TadA*8.7) を含む単量体構築物を有する。いくつかの実施形態において、ABE8は、Y147R、Q154R、およびY123H突然変異を有するTadA*7.10 (TadA*8.8) を含む単量体構築物を有するABE8.8-mである。いくつかの実施形態において、ABE8はABE8.9-mであり、これは、Y147R、Q154RおよびI76Y突然変異を有するTadA*7.10 (TadA*8.9) を含む単量体構築物を有する。いくつかの実施形態において、ABE8は、ABE8.10-mであり、これは、Y147R、Q154R、およびT166R突然変異を有するTadA*7.10 (TadA*8.10) を含む単量体構築物を有する。いくつかの実施形態において、ABE8は、ABE8.11-mであり、これは、Y147TおよびQ154R突然変異を有するTadA*7.10 (TadA*8.11) を含む単量体構築物を有する。いくつかの実施形態において、ABE8は、ABE8.12-mであり、これは、Y147TおよびQ154S突然変異を有するTadA*7.10 (TadA*8.12) を含む単量体構築物を有する。いくつかの実施形態において、ABE8は、ABE8.13-mであり、これは、Y123H (H123Yから戻されたY123H)、Y147R、Q154RおよびI76Y突然変異を有するTadA*7.10 (TadA*8.13) を含む単量体構築物を有する。いくつかの実施形態において、ABE8は、I76YおよびV82S突然変異を有するTadA*7.10 (TadA*8.14) を含む単量体構築物を有するABE8.14-mである。いくつかの実施形態において、ABE8は、V82SおよびY147R突然変異を有するTadA*7.10 (TadA*8.15) を含む単量体構築物を有するABE8.15-mである。いくつかの実施形態において、ABE8は、ABE8.16-mであり、これは、V82S、Y123H (H123Yから戻されたY123H) およびY147R突然変異を有するTadA*7.10 (TadA*8.16) を含む単量体構築物を有する。いくつかの実施形態では、ABE8はABE8.17-mであり、これは、V82SおよびQ154R突然変異を有するTadA*7.10 (TadA*8.17) を含む単量体構築物を有する。いくつかの実施形態では、ABE8はABE8.18-mであり、これは、V82S、Y123H (H123Yから戻されたY123H) およびQ154R突然変異を有するTadA*7.10 (TadA*8.18) を含む単量体構築物を有する。いくつかの実施形態では、ABE8はABE8.19-mであり、これは、V82S、Y123H (H123Yから戻されたY123H)、Y147RおよびQ154R突然変異を有するTadA*7.10 (TadA*8.19) を含む単量体構築物を有する。いくつかの実施形態において、ABE8は、ABE8.20-mであり、これは、I76Y、V82S、Y123H (H123Yから戻されたY123H)、Y147RおよびQ154R突然変異を有するTadA*7.10 (TadA*8.20) を含む単量体構築物を有する。いくつかの実施形態において、ABE8は、ABE8.21-mであり、これは、Y147RおよびQ154S突然変異を有するTadA*7.10 (TadA*8.21) を含む単量体構築物を有する。いくつかの実施形態において、ABE8は、V82SおよびQ154S突然変異を有するTadA*7.10 (TadA*8.22) を含む単量体構築物を有するABE8.22-mである。いくつかの実施形態において、ABE8は、ABE8.23-mであり、これは、V82SおよびY123H (H123Yから戻されたY123H) 突然変異を伴うTadA*7.10 (TadA*8.23) を含む単量体構築物を有する。いくつかの実施形態において、ABE8は、ABE8.24-mであり、これは、V82S、Y123H (H123Yから戻されたY123H)、およびY147T突然変異を有するTadA*7.10 (TadA*8.24) を含む単量体構築物を有する。

【0721】

いくつかの実施形態において、ABE8は、TadA*8バリエーションに融合された野生型E.coli TadAを含むヘテロ二量体構築物 (「ABE8.x-d」) を有する。いくつかの実施形態において、ABE8は、Y147T突然変異を有するTadA*7.10 (TadA*8.1) に融合された野生型E.coli TadAを含むヘテロ二量体構築物を有するABE8.1-dである。いくつかの実施形態において、ABE8は、Y147R突然変異を有するTadA*7.10 (TadA*8.2) に融合された野生型E.coli TadAを含むヘテロ二量体構築物を有するABE8.2-dである。いくつかの実施形態において、ABE8は、ABE8.3-dであり、これは、Q154S突然変異を有するTadA*7.10 (Ta

10

20

30

40

50

dA*8.3) に融合された野生型E.coli TadAを含むヘテロ二量体構築物を有する。いくつかの実施形態において、ABE8は、ABE8.4-dであり、これは、Y123H突然変異を有するTadA*7.10 (TadA*8.4) に融合された野生型E.coli TadAを含むヘテロ二量体構築物を有する。いくつかの実施形態において、ABE8は、V82S突然変異を有するTadA*7.10 (TadA*8.5) に融合された野生型E.coli TadAを含むヘテロ二量体構築物を有するABE8.5-dである。いくつかの実施形態において、ABE8は、T166R突然変異を有するTadA*7.10 (TadA*8.6) に融合された野生型E.coli TadAを含むヘテロ二量体構築物を有するdABE8.6-dである。いくつかの実施形態において、ABE8は、ABE8.7-dであり、これは、Q154R突然変異を有するTadA*7.10 (TadA*8.7) に融合された野生型E.coli TadAを含むヘテロ二量体構築物を有する。いくつかの実施形態において、ABE8は、Y147R、Q154R、およびY123H突然変異を有するTadA*7.10 (TadA*8.8) に融合された野生型E.coli TadAを含むヘテロ二量体構築物を有するABE8.8-dである。いくつかの実施形態において、ABE8は、ABE8.9-dであり、これは、Y147R、Q154RおよびI76Y突然変異を有するTadA*7.1 (TadA*8.9) 0に融合された野生型E.coli TadAを含むヘテロ二量体構築物を有する。いくつかの実施形態において、ABE8は、ABE8.10-dであり、これは、Y147R、Q154R、およびT166R突然変異を有するTadA*7.10 (TadA*8.10) に融合された野生型E.coli TadAを含むヘテロ二量体構築物を有する。いくつかの実施形態において、ABE8は、ABE8.11-dであり、これは、Y147TおよびQ154R突然変異を有するTadA*7.10 (TadA*8.11) に融合された野生型E.coli TadAを含むヘテロ二量体構築物を有する。いくつかの実施形態において、ABE8は、Y147TおよびQ154S突然変異を有するTadA*7.10 (TadA*8.12) に融合された野生型E.coli TadAを含むヘテロ二量体構築物を有するABE8.12-dである。いくつかの実施形態において、ABE8は、ABE8.13-dであり、Y123H (H123Yから戻されたY123H)、Y147R、Q154RおよびI76Y突然変異を有するTadA*7.10 (TadA*8.13) に融合された野生型E.coli TadAを含むヘテロ二量体構築物を有する。いくつかの実施形態において、ABE8は、I76YおよびV82S突然変異を有するTadA*7.10 (TadA*8.14) に融合された野生型E.coli TadAを含むヘテロ二量体構築物を有するABE8.14-dである。いくつかの実施形態において、ABE8は、V82SおよびY147R突然変異を有するTadA*7.10 (TadA*8.15) に融合された野生型E.coli TadAを含むヘテロ二量体構築物を有するABE8.15-dである。いくつかの実施形態では、ABE8はABE8.16-dであり、これは、V82S、Y123H (H123Yから戻されたY123H) およびY147R突然変異を有するTadA*7.10 (TadA*8.16) に融合された野生型E.coli TadAを含むヘテロ二量体構築物を有する。いくつかの実施形態において、ABE8は、V82SおよびQ154R突然変異を有するTadA*7.10 (TadA*8.17) に融合された野生型E.coli TadAを含むヘテロ二量体構築物を有するABE8.17-dである。いくつかの実施形態では、ABE8はABE8.18-dであり、これは、V82S、Y123H (H123Yから戻されたY123H) およびQ154R突然変異を有するTadA*7.10 (TadA*8.18) 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。いくつかの実施形態では、ABE8はABE8.19-dであり、これは、V82S、Y123H (YH123Yから戻されたY123H)、Y147RおよびQ154R突然変異を有するTadA*7.10 (TadA*8.19) に融合された野生型E.coli TadAを含むヘテロ二量体構築物を有する。いくつかの実施形態において、ABE8は、ABE8.20-dであり、これはI76Y、V82S、Y123H (H123Yから戻されたY123H)、Y147RおよびQ154R突然変異を有するTadA*7.10 (TadA*8.20) 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。いくつかの実施形態において、ABE8は、ABE8.21-dであり、これは、Y147RおよびQ154S突然変異を有するTadA*7.10 (TadA*8.21) に融合された野生型E.coli TadAを含むヘテロ二量体構築物を有する。いくつかの実施形態において、ABE8は、V82SおよびQ154S突然変異を有するTadA*7.10 (TadA*8.22) に融合された野生型E.coli TadAを含むヘテロ二量体構築物を有するABE8.22-dである。いくつかの実施形態では、ABE8はABE8.23-dであり、これは、V82SおよびY123H (H123Yから戻されたY123H) 突然変異を有するTadA*7.10 (TadA*8.23) に融合された野生型E.coli TadAを含むヘテロ二量体構築物を有する。いくつかの実施形態では、ABE8はABE8.24-dであり、これは、V82S、Y123H (H123Y

10

20

30

40

50

から戻されたY123H)、およびY147T突然変異を有するTadA*7.10 (TadA*8.24) に融合された野生型E.coli TadAを含むヘテロ二量体構築物を有する。

【0722】

いくつかの実施形態において、ABE8は、TadA*8バリエーションに融合されたTadA*7.10 (「ABE8.x-7」) を含むヘテロ二量体構築物を有する。いくつかの実施形態において、ABE8はY147T突然変異を有するTadA*7.10 (TadA*8.1) に融合されたTadA*7.10を含むヘテロ二量体構築物を有するABE8.1-7である。いくつかの実施形態において、ABE8はY147R突然変異を有するTadA*7.10 (TadA*8.2) に融合されたTadA*7.10を含むヘテロ二量体構築物を有するABE8.2-7である。いくつかの実施形態において、ABE8はQ154S突然変異を有するTadA*7.10に融合されたTadA*7.10 (TadA*8.3) を含むヘテロ二量体構築物を有するABE8.3-7である。いくつかの実施形態において、ABE8はY123H突然変異を有するTadA*7.10 (TadA*8.4) に融合されたTadA*7.10を含むヘテロ二量体構築物を有するABE8.4-7である。いくつかの実施形態において、ABE8はV82S突然変異を有するTadA*7.10 (TadA*8.5) に融合されたTadA*7.10を含むヘテロ二量体構築物を有するABE8.5-7である。いくつかの実施形態において、ABE8はT166R突然変異を有するTadA*7.10 (TadA*8.6) に融合されたTadA*7.10を含むヘテロ二量体構築物を有するABE8.6-7である。いくつかの実施形態において、ABE8はABE8.7-7であり、これは、Q154R突然変異を有するTadA*7.10 (TadA*8.7) に融合されたTadA*7.10を含むヘテロ二量体構築物を有する。いくつかの実施形態において、ABE8はY147R、Q154R、およびY123H突然変異を有するTadA*7.10 (TadA*8.8) に融合されたTadA*7.10を含むヘテロ二量体構築物を有するABE8.8-7である。いくつかの実施形態において、ABE8はY147R、Q154RおよびI76Y突然変異を有するTadA*7.10 (TadA*8.9) に融合されたTadA*7.10を含むヘテロ二量体構築物を有するABE8.9-7である。いくつかの実施形態において、ABE8はY147R、Q154R、およびT166R突然変異を有するTadA*7.10 (TadA*8.10) に融合されたTadA*7.10を含むヘテロ二量体構築物を有するABE8.10-7である。いくつかの実施形態において、ABE8はABE8.11-7であり、これは、Y147TおよびQ154R突然変異を有するTadA*7.10 (TadA*8.11) に融合されたTadA*7.10を含むヘテロ二量体構築物を有する。いくつかの実施形態において、ABE8はABE8.12-7であり、これは、Y147TおよびQ154S突然変異を有するTadA*7.10 (TadA*8.12) に融合されたTadA*7.10を含むヘテロ二量体構築物を有する。いくつかの実施形態において、ABE8はABE8.13-7であり、これは、Y123H (H123Yから戻されたY123H)、Y147R、Q154RおよびI76Y突然変異を有するTadA*7.10 (TadA*8.13) に融合されたTadA*7.10を含むヘテロ二量体構築物を有する。いくつかの実施形態において、ABE8はI76YおよびV82S突然変異を有するTadA*7.10 (TadA*8.14) に融合されたTadA*7.10を含むヘテロ二量体構築物を有するABE8.14-7である。いくつかの実施形態において、ABE8はV82SおよびY147R突然変異を有するTadA*7.10 (TadA*8.15) に融合されたTadA*7.10を含むヘテロ二量体構築物を有するABE8.15-7である。いくつかの実施形態では、ABE8はABE8.16-7であり、これは、V82S、Y123H (H123Yから戻されたY123H) およびY147R突然変異を有するTadA*7.10 (TadA*8.16) に融合されたTadA*7.10を含むヘテロ二量体構築物を有する。いくつかの実施形態において、ABE8はV82SおよびQ154R突然変異を有するTadA*7.10 (TadA*8.17) に融合されたTadA*7.10を含むヘテロ二量体構築物を有するABE8.17-7である。いくつかの実施形態では、ABE8はABE8.18-7であり、これは、V82S、Y123H (H123Yから戻されたY123H) およびQ154R突然変異を有するTadA*7.10 (TadA*8.18) に融合されたTadA*7.10を含むヘテロ二量体構築物を有する。いくつかの実施形態では、ABE8はABE8.19-7であり、これは、V82S、Y123H (H123Yから戻されたY123H)、Y147RおよびQ154R突然変異を有するTadA*7.10 (TadA*8.19) に融合されたTadA*7.10を含むヘテロ二量体構築物を有する。いくつかの実施形態において、ABE8はABE8.20-7であり、I76Y、V82S、Y123H (H123Yから戻されたY123H)、Y147RおよびQ154R突然変異を有するTadA*7.10 (TadA*8.20) に融合されたTadA*7.10を含むヘテロ二量体構築物を有する。いくつかの実施形態において、ABE8はABE8.21-

10

20

30

40

50

7であり、これは、Y147RおよびQ154S突然変異を有するTadA*7.10 (TadA*8.21) に融合されたTadA*7.10を含むヘテロ二量体構築物を有する。いくつかの実施形態において、ABE8はV82SおよびQ154S突然変異を有するTadA*7.10 (TadA*8.22) に融合されたTadA*7.10を含むヘテロ二量体構築物を有するABE8.22-7である。いくつかの実施形態では、ABE8はABE8.23-7であり、これは、V82SおよびY123H突然変異 (TadA*8.23) を有するTadA*7.10 (H123Yから戻されたY123H) に融合されたTadA*7.10を含むヘテロ二量体構築物を有する。いくつかの実施形態では、ABE8はABE8.24-7であり、これは、V82S、Y123H (H123Yから戻されたY123H)、およびY147T突然変異を有するTadA*7.10 (TadA*8.24) に融合されたTadA*7.10を含むヘテロ二量体構築物を有する。
【0723】

10

いくつかの実施形態において、ABEは、下記表7に示されているようにABE8.1-m、ABE8.2-m、ABE8.3-m、ABE8.4-m、ABE8.5-m、ABE8.6-m、ABE8.7-m、ABE8.8-m、ABE8.9-m、ABE8.10-m、ABE8.11-m、ABE8.12-m、ABE8.13-m、ABE8.14-m、ABE8.15-m、ABE8.16-m、ABE8.17-m、ABE8.18-m、ABE8.19-m、ABE8.20-m、ABE8.21-m、ABE8.22-m、ABE8.23-m、ABE8.24-m、ABE8.1-d、ABE8.2-d、ABE8.3-d、ABE8.4-d、ABE8.5-d、ABE8.6-d、ABE8.7-d、ABE8.8-d、ABE8.9-d、ABE8.10-d、ABE8.11-d、ABE8.12-d、ABE8.13-d、ABE8.14-d、ABE8.15-d、ABE8.16-d、ABE8.17-d、ABE8.18-d、ABE8.19-d、ABE8.20-d、ABE8.21-d、ABE8.22-d、ABE8.23-d、またはABE8.24-dである。

【0724】

20

表7：ABE塩基エディター

30

40

50

【表 7】

ABE8 塩基 エディタ ー	アデノシ ンデアミ ナーゼ	アデノシンデアミナーゼの説明
ABE8.1-m	TadA*8.1	Monomer_TadA*7.10 + Y147T
ABE8.2-m	TadA*8.2	Monomer_TadA*7.10 + Y147R
ABE8.3-m	TadA*8.3	Monomer_TadA*7.10 + Q154S
ABE8.4-m	TadA*8.4	Monomer_TadA*7.10 + Y123H
ABE8.5-m	TadA*8.5	Monomer_TadA*7.10 + V82S
ABE8.6-m	TadA*8.6	Monomer_TadA*7.10 + T166R
ABE8.7-m	TadA*8.7	Monomer_TadA*7.10 + Q154R
ABE8.8-m	TadA*8.8	Monomer_TadA*7.10 + Y147R_Q154R_Y123H
ABE8.9-m	TadA*8.9	Monomer_TadA*7.10 + Y147R_Q154R_I76Y
ABE8.10-m	TadA*8.10	Monomer_TadA*7.10 + Y147R_Q154R_T166R
ABE8.11-m	TadA*8.11	Monomer_TadA*7.10 + Y147T_Q154R
ABE8.12-m	TadA*8.12	Monomer_TadA*7.10 + Y147T_Q154S
ABE8.13-m	TadA*8.13	Monomer_TadA*7.10 + Y123H_Y147R_Q154R_I76Y
ABE8.14-m	TadA*8.14	Monomer_TadA*7.10 + I76Y_V82S
ABE8.15-m	TadA*8.15	Monomer_TadA*7.10 + V82S_Y147R
ABE8.16-m	TadA*8.16	Monomer_TadA*7.10 + V82S_Y123H_Y147R
ABE8.17-m	TadA*8.17	Monomer_TadA*7.10 + V82S_Q154R
ABE8.18-m	TadA*8.18	Monomer_TadA*7.10 + V82S_Y123H_Q154R
ABE8.19-m	TadA*8.19	Monomer_TadA*7.10 + V82S_Y123H_Y147R_Q154R
ABE8.20-m	TadA*8.20	Monomer_TadA*7.10 + I76Y_V82S_Y123H_Y147R_Q154R
ABE8.21-m	TadA*8.21	Monomer_TadA*7.10 + Y147R_Q154S
ABE8.22-m	TadA*8.22	Monomer_TadA*7.10 + V82S_Q154S
ABE8.23-m	TadA*8.23	Monomer_TadA*7.10 + V82S_Y123H
ABE8.24-m	TadA*8.24	Monomer_TadA*7.10 + V82S_Y123H_Y147T
ABE8.1-d	TadA*8.1	Heterodimer_(WT) + (TadA*7.10 + Y147T)
ABE8.2-d	TadA*8.2	Heterodimer_(WT) + (TadA*7.10 + Y147R)
ABE8.3-d	TadA*8.3	Heterodimer_(WT) + (TadA*7.10 + Q154S)

10

20

30

40

50

ABE8 塩基 エディタ ー	アデノシ ンデアミ ナーゼ	アデノシンデアミナーゼの説明
ABE8.4-d	TadA*8.4	Heterodimer_(WT) + (TadA*7.10 + Y123H)
ABE8.5-d	TadA*8.5	Heterodimer_(WT) + (TadA*7.10 + V82S)
ABE8.6-d	TadA*8.6	Heterodimer_(WT) + (TadA*7.10 + T166R)
ABE8.7-d	TadA*8.7	Heterodimer_(WT) + (TadA*7.10 + Q154R)
ABE8.8-d	TadA*8.8	Heterodimer_(WT) + (TadA*7.10 + Y147R_Q154R_Y123H)
ABE8.9-d	TadA*8.9	Heterodimer_(WT) + (TadA*7.10 + Y147R_Q154R_I76Y)
ABE8.10-d	TadA*8.10	Heterodimer_(WT) + (TadA*7.10 + Y147R_Q154R_T166R)
ABE8.11-d	TadA*8.11	Heterodimer_(WT) + (TadA*7.10 + Y147T_Q154R)
ABE8.12-d	TadA*8.12	Heterodimer_(WT) + (TadA*7.10 + Y147T_Q154S)
ABE8.13-d	TadA*8.13	Heterodimer_(WT) + (TadA*7.10 + Y123H_Y147T_Q154R_I76Y)
ABE8.14-d	TadA*8.14	Heterodimer_(WT) + (TadA*7.10 + I76Y_V82S)
ABE8.15-d	TadA*8.15	Heterodimer_(WT) + (TadA*7.10 + V82S_Y147R)
ABE8.16-d	TadA*8.16	Heterodimer_(WT) + (TadA*7.10 + V82S_Y123H_Y147R)
ABE8.17-d	TadA*8.17	Heterodimer_(WT) + (TadA*7.10 + V82S_Q154R)
ABE8.18-d	TadA*8.18	Heterodimer_(WT) + (TadA*7.10 + V82S_Y123H_Q154R)
ABE8.19-d	TadA*8.19	Heterodimer_(WT) + (TadA*7.10 + V82S_Y123H_Y147R_Q154R)
ABE8.20-d	TadA*8.20	Heterodimer_(WT) + (TadA*7.10 + I76Y_V82S_Y123H_Y147R_Q154R)
ABE8.21-d	TadA*8.21	Heterodimer_(WT) + (TadA*7.10 + Y147R_Q154S)
ABE8.22-d	TadA*8.22	Heterodimer_(WT) + (TadA*7.10 + V82S_Q154S)
ABE8.23-d	TadA*8.23	Heterodimer_(WT) + (TadA*7.10 + V82S_Y123H)
ABE8.24-d	TadA*8.24	Heterodimer_(WT) + (TadA*7.10 + V82S_Y123H_Y147T)

【 0 7 2 5 】

いくつかの実施形態において、塩基エディター（例えば、ABE8）は、アデノシンデアミナーゼバリエーション（例えば、TadA*8）を、循環置換体（circular permutant）Cas9（例えば、CP5またはCP6）および二部分核局在化配列を含む足場にクローニングすることによって生成される。いくつかの実施形態では、塩基エディター（例えば、ABE7.9、ABE7.10、またはABE8）は、NGC PAM CP5バリエーション（*S. pyrogenes* Cas9またはspVRQR Cas9）である。いくつかの実施形態では、塩基エディター（例えば、ABE7.9、ABE7.10、またはABE8）は、AGA PAM CP5バリエーション（*S. pyrogenes* Cas9またはspVRQR Cas9）である。いくつかの実施形態では、塩基エディター（例えば、ABE7.9、AB

E7.10、またはABE8)は、NGC PAM CP6バリエーション (S.pyrogenes Cas9またはspVR QR Cas9)である。いくつかの実施形態では、塩基エディター (例えば、ABE7.9、ABE 7.10、またはABE8)は、AGA PAM CP6バリエーション (S.pyrogenes Cas9またはspVRQR Cas9)である。

【0726】

いくつかの実施形態において、ABEは、以下の表8に示されるような遺伝子型を有する。

【0727】

表8．ABEの遺伝子型

【表8】

	23	26	36	37	48	49	51	72	84	87	105	108	123	125	142	145	147	152	155	156	157	161
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

【0728】

以下の表9に示すように、40個のABE8の遺伝子型が記述されている。ABEの進化されたE.coli TadA部分の残基位置が示されている。ABE8における変異変化は、ABE7.10の変異とは異なる場合に示されている。いくつかの実施形態では、ABEは、以下の表9に示されるABEのうちの1つの遺伝子型を有する。

20

【0729】

表9．進化されたTadAの残基アイデンティティ

30

40

50

【表 9】

	23	36	48	51	76	82	84	106	108	123	146	147	152	154	155	156	157	166
ABE7.10	R	L	A	L	I	V	F	V	N	Y	C	Y	P	Q	V	F	N	T
ABE8.1-m												T						
ABE8.2-m												R						
ABE8.3-m														S				
ABE8.4-m										H								
ABE8.5-m						S												
ABE8.6-m																		R
ABE8.7-m														R				
ABE8.8-m										H		R		R				
ABE8.9-m					Y							R		R				
ABE8.10-m												R		R				R
ABE8.11-m												T		R				
ABE8.12-m												T		S				
ABE8.13-m					Y					H		R		R				
ABE8.14-m					Y	S												
ABE8.15-m						S						R						
ABE8.16-m						S				H		R						
ABE8.17-m						S								R				
ABE8.18-m						S				H				R				
ABE8.19-m						S				H		R		R				
ABE8.20-m					Y	S				H		R		R				
ABE8.21-m												R		S				
ABE8.22-m						S								S				
ABE8.23-m						S				H								
ABE8.24-m						S				H		T						
ABE8.1-d												T						
ABE8.2-d												R						
ABE8.3-d														S				
ABE8.4-d										H								
ABE8.5-d						S												
ABE8.6-d																		R
ABE8.7-d														R				
ABE8.8-d										H		R		R				
ABE8.9-d					Y							R		R				
ABE8.10-d												R		R				R
ABE8.11-d												T		R				
ABE8.12-d												T		S				
ABE8.13-d					Y					H		R		R				
ABE8.14-d					Y	S												
ABE8.15-d						S						R						
ABE8.16-d						S				H		R						

10

20

30

40

50

ABE8.17-d					S								R				
ABE8.18-d					S				H				R				
ABE8.19-d					S				H		R		R				
ABE8.20-d				Y	S				H		R		R				
ABE8.21-d											R		S				
ABE8.22-d					S								S				
ABE8.23-d					S				H								
ABE8.24-d					S				H		T						

10

【 0 7 3 0 】

いくつかの実施形態では、塩基エディターはABE8.1であり、これは、アデノシンデアミナーゼ活性を有する以下の配列またはそのフラグメントを含むか、または本質的にそれらからなる。

【 0 7 3 1 】

ABE8.1_Y147T_CP5_NGC PAM_monomer

20

30

40

50

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVI GEGWNRAIGLHDPTAHAEIMALR
QGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNH
RVEITEGILADECAALLCTFFRMPRQVFNAQKKAQSSTD **SGGSSGGSSGSETPGTSESATP****ESS**
GGSSGGSEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVR
KVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFMQPTVAYSVLVVAK
VEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKR
MLASAKFLQKGNELALPSKYVNFLYLASHYEKLKGSPEDEQKQLFVEQHKHYLDEIIEQISEF 10
SKRVILADANLDKVL SAYNKH RD KPIREQAENIIHLFTLTNLGAPRAFKYFDTTIARKEYRSTK
EVLDATLIHQ SITGLYETRIDLSQLGGDGGSSGGSSGGSSGGSSGGMDKKYSIGLAIGTNSV
GWAVITDEYKVPSKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRIC
YLQEIFSNEMAKVDDSFHRL EESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVD
STDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEEENPINASGVDA
KAILSARLSKSRLENLIAQLPGEKKNGLFGNLI ALSGLTPNFKSNEFLAEDAKLQLSKDTYD
DDL DNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLK 20
ALVRQQLP EKYKEIFFDQSKNGYAGYIDGGASQEEFYKF I KPILEKMDGTEELLVKLNREDLLR
KQRTFDNGSIPHQIHLGELHAILRRQEDFY PFLKDNREKIEKILTFRIPIYVGPLARGNSRFAW
MTRKSEETITPWNFE EVVDKGASAQSFIERMTNFDKNLPNEKVLPHSLLYEYFTVYNELTKVK
YVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLG
TYHDLLKIIKDKDFLDNEENEDILEDIVLTLTTLFEDREMIEERLKYAHLFDDKVMKQLKRRRY
TGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQGD SL
HEHIANLAGSPA IKKGILQTVKVVDLVKVMGRHKPENIV IEMARENQTTQKGQKNSRERMKRI
EEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHI VQSF LK 30
DDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKKNYWRQLLNAKLITQRKFDNLTKAERGGLSELD
KAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQFYKVR
EINNYHHAH DAYLNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMI AKSEQEGADKRTADGSEF
ESPKKKRKV*

上記の配列において、プレーンテキストはアデノシンデアミナーゼ配列を示し、太字の配列はCas9に由来する配列を示し、斜体の配列はリンカー配列を示し、下線が引かれた配列は二部分核局在化配列を示す。

【0732】

いくつかの実施形態では、塩基エディターはABE8.1であり、これは、アデノシンデアミナーゼ活性を有する以下の配列またはそのフラグメントを含むか、または本質的にそれらからなる。

【0733】

pNMG-B335 ABE8.1_Y147T_CP5_NGC PAM_monomer

10

20

30

40

50

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVNLRVIGEGWNRAIGLHDPTAHAEIMALR
 QGGLVMQNYRLIDATLYVTTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNH
 RVEITEGILADECAALLCTFFRMPRQVFNAQKKAQSSTDSGGSSGGSSGSETPGTSESATPES
 GGSSGGSEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVR
 KVLSPQVNIIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFMQPTVAYSVLVVAK
 VEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDLIIKLPKYSLELENGRKR
 MLASAKFLQKGNELALPSKYVNFLYLASHYEKLKGSPEDEQKQLFVEQHKHYLDEIIEQISEF
 SKRVILADANLDKVL SAYNKHDKPIREQAENIIHLFTLTNLGAPRAFKYFDTTIARKEYRSTK
 EVLDATLIHQSIITGLYETRIDLSQLGGDGGSSGGSSGGSSGGSSGGMDKKYSIGLAIGTNSV
 GWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRIC
 YLQEIFSNEMAKVDDSFHRL EESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVD
 STDKADLRILIYALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDA
 KAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKS NFDLAEDAKLQLSKD TYD
 DDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLK
 ALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLR
 KQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRFAW
 MTRKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNEKVL PKHSLLEYFTVYNELTKVK
 YVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLG
 TYHDL LKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRY
 TGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSL
 HEHIANLAGSPAIKKGILQTVKVDELVKVMGRHKPENIVIVEMARENQTTQKGQKNSRERMKRI
 EEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK
 DDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKMKNYWRQLLNAKLITQRFKDNLTKAERGGLSELD
 KAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQFYKVR
 EINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEGADKRTADGSEF
ESPKKKRKV*

10

20

30

上記の配列において、プレーンテキストはアデノシンデアミナーゼ配列を示し、太字の配列はCas9に由来する配列を示し、斜体の配列はリンカー配列を示し、下線が引かれた配列は二部分核局在化配列を示す。

40

【0734】

いくつかの実施形態では、塩基エディターはABE8.14であり、これは、アデノシンデアミナーゼ活性を有する以下の配列またはそのフラグメントを含むか、または本質的にそれらからなる。

【0735】

pNMG-357_ABE8.14 with NGC PAM CP5

50

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVI GEGWNRPIGRHDPTAHAEIMALR
QGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAGSLMDVLHHPGMNH
RVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDGSSGGSSGSETPGTSESATPESSG
GSSGGSMSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVI GEGWNRRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLH
YPGMNRHVEITEGILADECAALLCTFFRMPRQVFNAQKKAQSSTDGSSGGSSGSETPGTSES
ATPESSGGSSGGSEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGR
DFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFMQPTVAYS
VLVVAKEKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKDLIIKLPKYSLFEL
ENGRKRMLASAKFLQKGNELALPSKYVNFLYLASHYEKLKGS PEDNEQKQLFVEQHKHYLDEII
EQISEFSKRVLADANLDKVL SAYNKH RDKPIREQAENIIHLFTLTNLGAPRAFKYFDTTIARK
EYRSTKEVL DATLIHQ SITGLYETRIDL SQLGGDGGSSGGSSGGSSGGSSGGMDKKYSIGLA
IGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTR
RKNRICYLQEIFSNE MAKVDD SFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHL
RKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLF EENPIN
ASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQL
SKD TYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQ
DLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLN
REDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPIYYVGPLARG
NSRFAWMTRKSEETITPWNFEEVVDKGASASQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYN
ELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDR
FNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFFDDKVMKQ
LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVS
GQGDSLHEHIANLAGSPAIKKGILQTVKVVDLVKVMGRHKPENIVIEMARENQTTQKGQKNSR
ERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIV
PQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKMKNYWRQLLNAKLITQRKFDNLTKAERG
GLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDF
QFYKVR EINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIKSEQEGADKRT
ADGSEFESPKKKRKV*

10

20

30

40

上記の配列において、プレーンテキストはアデノシンデアミナーゼ配列を示し、太字の配列はCas9に由来する配列を示し、斜体の配列はリンカー配列を示し、下線が引かれた配列は二部分核局在化配列を示す。

【0736】

いくつかの実施形態では、塩基エディターはABE8.8-mであり、これは、アデノシンデアミナーゼ活性を有する以下の配列またはそのフラグメントを含むか、または本質的にそれらからなる。

【0737】

ABE8.8-m

50

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVNNRVIGEGWNRAIGLHDPTAHAEIMALR
 QGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHHPGMNH
 RVEITEGILADECAALLCRFFRMPRRVFNAQKKAQSSTDSSGSSGGSSGSETPGTSESATPESS
 GGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTRHSIKKNLIGALLFDSGET
 AEATRLKRTARRRYTRKKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNI
 VDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLEFI
 QLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
 NFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITK
 APLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKP
 ILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEK
 ILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEK
 VLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFK
 KIECFDSVEISGVEDRFNASLGTYHDLKIIKDKDFLDNEENEDILEDIVLTTLTLFEDREMIEE
 RLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIH
 DDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIE
 MARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQ
 ELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKKNYWRQLLNAK
 LITQRKFDNLTKAERGGSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVK
 VITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDV
 RKMIASEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFAT
 VRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVV
 AKVEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKDLIIKLPKYSLFELENGR
 KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGGSPEDNEQQLFVEQHKHYLDEIIEQIS
 EFSKRVILADANLDKVL SAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTS
 TKEVL DATLIHQ SITGLYETRIDLSQLGGDEGADKRTADGSEFESPKKKRKV*

10

20

30

上記の配列において、プレーンテキストはアデノシンデアミナーゼ配列を示し、太字の配列はCas9に由来する配列を示し、斜体の配列はリンカー配列を示し、下線が引かれた配列は二部分核局在化配列を示し、二重下線が引かれた配列は変異を示す。

40

【0738】

いくつかの実施形態では、塩基エディターはABE8.8-dであり、これは、アデノシンデアミナーゼ活性を有する以下の配列またはそのフラグメントを含むか、または本質的にそれらからなる。

【0739】

ABE8.8-d

50

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHAEIMALR
QGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAGSLMDVLHHPGMNH
RVEITEGILADECAALLSDFFRMRRQEIKAKKAQSSTDSGGSSGGSSGSETPGTSESATPES
GGSSGGSSSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVHNNRVIGEGWNRRAIGLHDPTA
EIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLH
HPGMNRVEITEGILADECAALLCRFFRMPRRVFNAQKAQSSTDSGGSSGGSSGSETPGTSES
ATPESGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALL
FDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERH
PIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSD
VDKLFIQLVQTYNQLFEEPNINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIAL
SLGLTPNFKSNFDLAEDAKLQLSKDITYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRV
NTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEF
YKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDN
REKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDK
NLPNEKVLPHKSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQL
KEDYFKKIECFDSVEISGVEDRFNASLGTYHDLKIIKDKDFLDNEENEDILEDIVLTLTLFED
REMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRN
FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDLVKVMGRHKP
ENIVIEMARENQTTQKGQKNSRERMKRIE EG IKELGSQILKEHPVENTQLQNEKLYLYYLQNGR
DMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWR
QLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDK
LIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGITALIKKYPKLESEFVYGD
YKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDK
GRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVA
YSVLVVAKEGKSKKLKSVKELLGITIMERSSEFKNPIDFLEAKGYKEVKKDLIIKLPKYSLF
ELENKRKRLASAGELQKGNELALPSKYVNFLYLASHYEKLGSPEDNEQKQLFVEQHKHYLDE
IIEQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLFTLTNLGAPAAFKYFDTTID
RKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGDE EGADKRTADGSEFESPKKKRKV*

10

20

30

40

上記の配列において、プレーンテキストはアデノシンデアミナーゼ配列を示し、太字の配列はCas9に由来する配列を示し、斜体の配列はリンカー配列を示し、下線が引かれた配列は二部分核局在化配列を示し、二重下線が引かれた配列は変異を示す。

【0740】

いくつかの実施形態では、塩基エディターはABE8.13-mであり、これは、アデノシンデアミナーゼ活性を有する以下の配列またはそのフラグメントを含むか、または本質的にそれらからなる。

【0741】

ABE8.13-m

50

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVNNRVIGEGWNRAIGLHDPTAHAEIMALR
 QGGLVMQNYRLYDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHHPGMNH
 RVEITEGILADECAALLCRFFRMPRRVFNAQKKAQSSTDSGGSSGGSSGSETPGTSESATPSESS
 GGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGET
 AEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNI
 VDEVAYHEKYPTIYHLRKKLV DSTDKADRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFI
 QLVQTYNQLFEEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
 NFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLF LAAKNLSDAILLSDILRVNTEITK
 APLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKP
 ILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPFLKDNREKIEK
 ILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEVVVDKGASAQSFIERMTNFDKNLPNEK
 VLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFK
 KIECFDSVEISGVEDRFNASLGTYHDLKKI IKDKDFLDNEENEDILEDIVLTLTLFEDREMIEE
 RLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIH
 DDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDDELVKVMGRHKPENIVIE
 MARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQ
 ELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKKNYWRQLLNAK
 LITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILD SRMNTKYDENDKLIREVK
 VITLKSCLVSDFRKDFQFYKVRINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDV
 RKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFAT
 VRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVV
 AKVEKGKSKKLKSVKELLGITIMERSSSF EKNPIDFLEAKGYKEVKKDII IKLPKYSLFELNGR
 KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGS PEDNEQKQLFVEQHKKHYLDEIEQIS
 EFSKRVLADANLDKVL SAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTS
 TKEVL DATLIHQ SITGLYETRIDLSQLGGDEGADKRTADGSEFESP KKKRKV*

10

20

30

上記の配列において、プレーンテキストはアデノシンデアミナーゼ配列を示し、太字の配列はCas9に由来する配列を示し、斜体の配列はリンカー配列を示し、下線が引かれた配列は二部分核局在化配列を示し、二重下線が引かれた配列は変異を示す。

40

【0742】

いくつかの実施形態では、塩基エディターはABE8.13-dであり、これは、アデノシンデアミナーゼ活性を有する以下の配列またはそのフラグメントを含むか、または本質的にそれらからなる。

【0743】

ABE8.13-d

50

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHAEIMALR
 QGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAGSLMDVLHHPGMNH
 RVEITEGILADECAALLSDFFRMRRQEIKAKKAQSSTDSSGGSSGGSSGSETPGTSESATPESS
 GGSSGSSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNRRVIGEGWNRRAIGLHDPTAHA
 EIMALRQGGLVMQNYRLYDATLYVTLEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLH
 HPGMNRVEITEGILADECAALLCRFFRMRRVFNAQKKAQSSTDSSGGSSGGSSGSETPGTSES
 ATPESSGGSSGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALL
 FDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHRLEESFLVEEDKKHERH
 PIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADRLIYLALAHMIKFRGHFLIEGDLNPDNSD
 VDKLFIQLVQTYNQLFEEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIAL
 SLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRV
 NTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEF
 YKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDN
 REKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDK
 NLPNEKVLPHKSHLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQL
 KEDYFKKIECFDSVEISGVEDRFNASLGTYHDLKIIKDKDFLDNEENEDILEDIVLTLTLFED
 REMIEERLKTYAHLEFDDKVMKQLKRRRYTGWRLSRKLINGIRDKQSGKTILDFLKSDGFANRN
 FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDLVKVMGRHKP
 ENIVIEMARENQTTQKGQKNSRERMKRIEIGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGR
 DMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKMKNYWR
 QLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDK
 LIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGD
 YKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDK
 GRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVA
 YSVLVVAKVEKGKSKKLKSVKELLGITIMERSSEFKNPIDFLEAKGYKEVKKDLIIKLPKYSLF
 ELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGS PEDNEQKQLFVEQHKHYLDE
 IIEQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLFTLTNLGAPAAFKYFDTTID
 RKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGDEGADKRTADGSEFESP KKKRKV*

10

20

30

40

上記の配列において、プレーンテキストはアデノシンデアミナーゼ配列を示し、太字の配列はCas9に由来する配列を示し、斜体の配列はリンカー配列を示し、下線が引かれた配列は二部分核局在化配列を示し、二重下線が引かれた配列は変異を示す。

【0744】

いくつかの実施形態では、塩基エディターはABE8.17-mであり、これは、アデノシンデアミナーゼ活性を有する以下の配列またはそのフラグメントを含むか、または本質的にそれらからなる。

【0745】

ABE8.17-m

50

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVNNRVIGEGWNRAIGLHDPTAHAEIMALR
QGGLVMQNYRLIDATLYSTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNH
RVEITEGILADECAALLCYFFRM^{PRR}VFNAQKKAQSSTDSSGSSGGSSGSETPGTSESATPES
GGSSGGSDKKYSIGL^{AI}GTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLEDSGET
AEATRLKRTARRRYTRKKNRICYLQE^{IFS}NEMAKVDDSFHRL^{EES}FLVEEDKKHERHPIFGNI
VDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFI
QLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
NFKSNFDLAEDAKLQLSKD^{TYDDDL}NLLAQIGDQYADLEFLAAKNLSAILLSDILRVNTEITK
APLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKP
ILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEK
ILTFRIPIYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEK
VLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFK
KIECFDSVEISGVEDRFNASLGT^{YHDL}LKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEE
RLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMOLIH
DDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDDELVKVMGRHKPENIVIE
MARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQ
ELDINRLSDYD^{VDHIV}QSF^{LKDD}SIDNKVLTRSDKNRGKSDNVPSEEVVKMKKNYWRQLLNAK
LITQRKFDNLTKAERGGLSELDKAGFIKRLVETRQITKHVAQILDSRMNTKYDENDKLIREV
VITLKS^{KLVS}DFRKDFQFYKVREINNYHHAHDAYLNAVVG^{TALIK}KYPKLESEFVYGDYKVYDV
RKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFAT
VRKVL^{SMPQ}VNIVKKTEVQTTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLV
AKVEKGKSKKLKSVKELLGITIMERS^{SFEKN}PIDFLEAKGYKEVKKD^{LI}IKLPKYSLFELENGR
KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLK^{GSPED}NEQKQLFVEQHKHYLDEIIEQIS
EFSKRVILADANLDKVL^{SAYN}KHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTS
TKEVL^{DATLI}HQSITGLYETRIDLSQLGGDEGADKRTADGSEFESP^{KKKRKV}*

10

20

30

上記の配列において、プレーンテキストはアデノシンデアミナーゼ配列を示し、太字の配列はCas9に由来する配列を示し、斜体の配列はリンカー配列を示し、下線が引かれた配列は二部分核局在化配列を示し、二重下線が引かれた配列は変異を示す。

40

【0746】

いくつかの実施形態では、塩基エディターはABE8.17-dであり、これは、アデノシンデアミナーゼ活性を有する以下の配列またはそのフラグメントを含むか、または本質的にそれらからなる。

【0747】

ABE8.17-d

50

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHAEIMALR
QGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAGSLMDVLHHPGMNH
RVEITEGILADECAALLSDFFRMRREQEIKAQKKAQSSTDSSGSSGGSSGSETPGTSESATPESS
GGSSGGSSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVNLRVIGEGWNRRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYSTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLH
YPGMNRHVEITEGILADECAALLCYFFRMPRRVFNAQKKAQSSTDSSGSSGGSSGSETPGTSES
ATPESSGGSSGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALL
FDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERH
PIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSD
VDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIAL
SLGLTPNFKSNEFDLAEDAKLQLSKDITYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRV
NTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEF
YKFIKPILEKMDGTEELLVKNLREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDN
REKIEKILTFRIPYVGPLARGNSRFAWMTRKSEETITPWNFEVVVDKGASAQSFIERMTNFDK
NLPNEKVLPKHSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQL
KEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFED
REMIEERLKTYAHLFDDKVMKQLKRRRYTGWRLSRKLINGIRDKQSGKTILDFLKSDGFANRN
FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDLKVVMGRHKP
ENIVIAMARENQTTQKGQKNSRERMKRIEIEGKELGSQILKEHPVENTQLQNEKLYLYYLQNGR
DMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVKKMKNYWR
QLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDK
LIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGD
YKVYDVRKMIKSEQEI GKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDK
GRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVA
YSVLVVAKEVGKSKKLKSVKELLGITIMERSSEFKNPIDFLEAKGYKEVKKDLIIKLPKYSLF
ELENKRKRLASAGELQKGNELALPSKYVNFLYLASHYEKLGSPEDNEQKQLEFVEQHKHYLDE
IIEQISEFSKRVLADANLDKVL SAYNKHDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTID
RKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGDEGADKRTADGSEFESPKKRKY*

10

20

30

40

上記の配列において、プレーンテキストはアデノシンデアミナーゼ配列を示し、太字の配列はCas9に由来する配列を示し、斜体の配列はリンカー配列を示し、下線が引かれた配列は二部分核局在化配列を示し、二重下線が引かれた配列は変異を示す。

【0748】

いくつかの実施形態では、塩基エディターはABE8.20-mであり、これは、アデノシンデアミナーゼ活性を有する以下の配列またはそのフラグメントを含むか、または本質的にそれらからなる。

【0749】

ABE8.20-m

50

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVNNRVIGEGWNRAIGLHDPTAHAEIMALR
QGGLVMQNYRLYDATLYSTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHHPGMNH
RVEITEGILADECAALLCRFFRMPRRVFNAQKKAQSSTDSGGSSGGSSGSETPGTSESATPS
GGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGET
AEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNI
VDEVAYHEKYPTIYHLRKKLVDDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFI
QLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTP
NFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLF~~L~~AAKNLSDAILLSDILRVNTEITK
APLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKP
ILEKMDGTEELLVKNLREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEK
ILTFRIPIYYVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNEK
VLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFK
KIECFDSVEISGVEDRFNASLGTYHDLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEE
RLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMLIH
DDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVDELVKVMGRHKPENIVIE
MARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQ
ELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKKNYWRQLLNAK
LITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVK
VITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDV
RKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFAT
VRKVLSPQVNIIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVV
AKVEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDLIIKLPKYSLFELNGR
KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDEQKQLFVEQHKKHYLDEIEQIS
EFSKRVLADANLDKVL SAYNKHRDKPIREQAENIIHLETLTNLGAPAAFKYFDTTIDRKRYTS
TKEVLDTLIIHQSIITGLYETRIDLSQLGGDEGADKRTADGSEFESPKKKRKV*

10

20

30

上記の配列において、プレーンテキストはアデノシンデアミナーゼ配列を示し、太字の配列はCas9に由来する配列を示し、斜体の配列はリンカー配列を示し、下線が引かれた配列は二部分核局在化配列を示し、二重下線が引かれた配列は変異を示す。

【0750】

40

いくつかの実施形態では、塩基エディターはABE8.20-dであり、これは、アデノシンデアミナーゼ活性を有する以下の配列またはそのフラグメントを含むか、または本質的にそれらからなる。

【0751】

ABE8.20-d

50

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHAEIMALR
QGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAGSLMDVLHHPGMNH
RVEITEGILADECAALLSDFFRMRRQEIKAKKAQSSTDSSGSSGGSSGSETPGTSESATPES
GGSSGSSSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVHNNRVIGEGWNRRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYSTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLH
HPGMNHRVEITEGILADECAALLCRFFRMPRRVFNAQKKAQSSTDSSGSSGGSSGSETPGTSES
ATPESGGSSGSSDKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALL
FDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHRLEESFLVEEDKKHERH
PIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSD
VDKLFIQLVQTYNQLFEEPNINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIAL
SLGLTPNEFSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADFLAAKNLSDAILLSDILRV
NTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEF
YKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDN
REKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDK
NLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQL
KEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFED
REMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRN
FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDLVKVMGRHKP
ENIVIEARENQTTQKGQKNSRERMKRIEIEGKELGSQILKEHPVENTQLQNEKLYLYYLQNGR
DMYVDQELDINRLSDYDVHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKKNYWR
QLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDK
LIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGD
YKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDK
GRDFATVRKVLSPQVNIIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVA
YSVLVVAKEVGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDIIKLPKYSLE
ELENKRKMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDEQKQLFVEQHKHYLDE
IIEQISEFSKRVLADANLDKVL SAYNKHDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTID
RKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGDEGADKRTADGSEFESPKKRKV*

10

20

30

40

上記の配列において、プレーンテキストはアデノシンデアミナーゼ配列を示し、太字の配列はCas9に由来する配列を示し、斜体の配列はリンカー配列を示し、下線が引かれた配列は二部分核局在化配列を示し、二重下線が引かれた配列は変異を示す。

【0752】

いくつかの実施形態では、本発明のABE8は以下の配列から選択される。

【0753】

01. monoABE8.1_bpNLS + Y147T

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVHNNRVIGEGWNRRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHHPGMNHRVEITEGILADECAALLCTFFRMPRQVFNAQKKAQSSTDSSGSSGGSS

50

GSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTD
RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYALAHM
IKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRL
NLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNDLAEDAKLQLSKDITYDDDLNLLAQIGD
QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLP
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
NGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR
KSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN
ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMK
QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDFANRNFQMQLIHDDSLTFKEDIQ
KAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDELVKVMGRHKPENIVIAMARENQT
TQKGQKNSRERMKRIEELGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL
DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLL
NAKLITQRKFDNLTKAERGGLSELDKAGFIKQRLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKSCLVSDFRKDFQFYKVVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEF
VYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET
GEIVWDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK
KYGGFVSPTVAYSVLVAKVEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEV
KKDLIIKLPKYSLELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLKGPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKHDKPIREQAENIIHLF
TLTNLGAPAAFKYFDTTIDRKQYRSTKEVLDTLIHQ SITGLYETRIDLSQLGGDEGADKR
TADGSEFESPKKKRKV

[0 7 5 4]

02. monoABE8.1_bpNLS + Y147R

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNRRVIGEGWNRAIGLHDPTAHA
EIMALRQGGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNHRVEITEGILADECAALLCRFFRMPRQVFNAQKKAQSSTDSGGSSGGSS
GSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTD
RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYALAHM
IKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRL
NLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNDLAEDAKLQLSKDITYDDDLNLLAQIGD
QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLP
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
NGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR
KSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN
ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMK
QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDFANRNFQMQLIHDDSLTFKEDIQ
KAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDELVKVMGRHKPENIVIAMARENQT
TQKGQKNSRERMKRIEELGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL
DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLL
NAKLITQRKFDNLTKAERGGLSELDKAGFIKQRLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKSCLVSDFRKDFQFYKVVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEF
VYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET
GEIVWDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK
KYGGFVSPTVAYSVLVAKVEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEV
KKDLIIKLPKYSLELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLKGPED

NEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVLSAYNKHrdKPIREQAENIIHLF
TLTNLGAPAAFKYFDDTTIDRKQYRSTKEVLDATLIHQSI TGLYETRIDLSQLGGDEGADKR
TADGSEFESPKKKRKV

[0 7 5 5]

03. monoABE8.1_bpNLS + Q154S

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNRRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRSVFNAQKKAQSSTDSSGGSSGGSSG
SETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDR
HSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRL
EESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYLAHAM
KFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRL
NLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNDLAEDAKLQLSKDTYDDDLNLLAQIGD
QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPE
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
NGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR
KSEETITPWNFEVVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN
ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMK
QLKRRRYTGWRSLRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQ
KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDDELVKVMGRHKPENIVIAMARENQT
TQKGQKNSRERMKRIEELGSGILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL
DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVKKMKNYWRQLL
NAKLITQRKFDNLTKAERGGSELKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG TALIKKYPKLESEF
VYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET
GEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK
KYGGFVSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSEKNPIDFLEAGYKEV
KKDLIIKLPKYSLFELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLKGSPE
NEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVLSAYNKHrdKPIREQAENIIHLF
TLTNLGAPAAFKYFDDTTIDRKQYRSTKEVLDATLIHQSI TGLYETRIDLSQLGGDEGADKR
TADGSEFESPKKKRKV

[0 7 5 6]

04. monoABE8.1_bpNLS + Y123H

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNRRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHHPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTDSSGGSSGGSS
GSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTD
RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRL
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYLAHAM
IKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRL
NLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNDLAEDAKLQLSKDTYDDDLNLLAQIGD
QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPE
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
NGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR
KSEETITPWNFEVVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN
ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMK
QLKRRRYTGWRSLRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQ
KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDDELVKVMGRHKPENIVIAMARENQT

TQKGQKNSRERMKRIEELGSGILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL
DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVKKMKNYWRQLL
NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEF
VYGDYKVVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET
GEIVWDKGRDFATVRKVLSPQVNVVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK
KYGGFVSPTVAYSVLVVAKEKGSKKLKSVELLGITIMERSSEKKNPIDFLEAKGYKEV
KKDLIIKLPKYSLFELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLGSPED
NEQKQLFVEQHKHYLDEIEQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLF
TLTNLGAPAAFKYFDTTIDRKQYRSTKEVL DATLIHQ SITGLYETRIDLSQLGGDEGADKR
TADGSEFESPKKKRKV

10

[0 7 5 7]

05. monoABE8.1_bpNLS + V82S

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVNLRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYSTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTDSGGSSGGSS
GSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTD
RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDD SFFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLR LIYLALAHM
IKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRL
NLIAQLPGKKNGLFGNLIALSLGLTPNFKS NFDLAEDAKLQLSKD TYDDDLNLLAQIGD
QYADFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPE
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
NGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR
KSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN
ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMK
QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKS DGFANRNF MQLIHDDSLTFKEDIQ
KAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDDELVKVMGRHKPENIVIAMARENQT
TQKGQKNSRERMKRIEELGSGILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL
DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVKKMKNYWRQLL
NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEF
VYGDYKVVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET
GEIVWDKGRDFATVRKVLSPQVNVVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK
KYGGFVSPTVAYSVLVVAKEKGSKKLKSVELLGITIMERSSEKKNPIDFLEAKGYKEV
KKDLIIKLPKYSLFELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLGSPED
NEQKQLFVEQHKHYLDEIEQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLF
TLTNLGAPAAFKYFDTTIDRKQYRSTKEVL DATLIHQ SITGLYETRIDLSQLGGDEGADKR
TADGSEFESPKKKRKV

20

30

40

[0 7 5 8]

06. monoABE8.1_bpNLS + T166R

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVNLRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSRDSGGSSGGSS
GSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTD
RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDD SFFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLR LIYLALAHM
IKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRL
NLIAQLPGKKNGLFGNLIALSLGLTPNFKS NFDLAEDAKLQLSKD TYDDDLNLLAQIGD

50

QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPE
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
NGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR
KSEETITPWNFEVVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN
ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMK
QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQ
KAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDELVKVMGRHKPENIVIAMARENQT
TQKGQKNSRERMKRIEEGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL
DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLL
NAKLITQRKFDNLTAKAERGGLSELDKAGFIKQRLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKSCLVSDFRKDFQFYKVVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEF
VYGDYKVVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET
GEIVWDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK
KYGGFVSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSSFENPIDFLEAGKYKEV
KKDLIIKLPKYSLELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLKGSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLF
TLTNLGAPAAFKYFDTTIDRKQYRSTKEVL DATLIHQ SITGLYETRIDLSQLGGDEGADKR
TADGSEFESPKKKRKV

10

【0759】

20

07. monoABE8.1_bpNLS + Q154R

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNRRVIGEGWNRAIGLHDPTAHA
EIMALRQGGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRRVFNAQKKAQSSTDSGGSSGGSS
GSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTD
RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDD SFFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLR LIYLALAHM
IKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRL
NLIAQLPGEKKNGLFGNLIALSLGLTPNFKS NFDLAEDAKLQLSKD TYDDDLNLLAQIGD
QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPE
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
NGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR
KSEETITPWNFEVVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN
ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMK
QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQ
KAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDELVKVMGRHKPENIVIAMARENQT
TQKGQKNSRERMKRIEEGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL
DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLL
NAKLITQRKFDNLTAKAERGGLSELDKAGFIKQRLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKSCLVSDFRKDFQFYKVVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEF
VYGDYKVVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET
GEIVWDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK
KYGGFVSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSSFENPIDFLEAGKYKEV
KKDLIIKLPKYSLELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLKGSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLF
TLTNLGAPAAFKYFDTTIDRKQYRSTKEVL DATLIHQ SITGLYETRIDLSQLGGDEGADKR
TADGSEFESPKKKRKV

30

40

【0760】

08. monoABE8.1_bpNLS + Y147R_Q154R_Y123H

50

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHHPGMNHRVEITEGILADECAALLCRFFRMPRRVFNAQKKAQSSTDSSGGSSGGSS
GSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTD
RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYALAHM
IKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRL
NLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGD
QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLP
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
NGSIPHQIHLGELHAILRRQEDFYFPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR
KSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN
ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMK
QLKRRRYTGWGRLSRKLINGIRDQKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQ
KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDLKVVMGRHKPENIVIAMARENQT
TQKGQKNSRERMKRIEELGSGILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL
DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLL
NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG TALIKKYPKLESEF
VYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET
GEIVWDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK
KYGGFVSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSEFNPIDFLEAGYKEV
KKDLIIKLPKYSLFELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLGSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLF
TLTNLGAPAAFKYFDTTIDRKQYRSTKEVLDATLIHQ SITGLYETRIDLSQLGGDEGADKR
TADGSEFESPKKKRKV

[0761]

09. monoABE8.1_bpNLS + Y147R_Q154R_I76Y

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLYDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNHRVEITEGILADECAALLCRFFRMPRRVFNAQKKAQSSTDSSGGSSGGSSG
SETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTDR
HSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRL
EESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYALAHMI
KFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRL
NLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGD
QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLP
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
NGSIPHQIHLGELHAILRRQEDFYFPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR
KSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN
ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMK
QLKRRRYTGWGRLSRKLINGIRDQKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQ
KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDLKVVMGRHKPENIVIAMARENQT
TQKGQKNSRERMKRIEELGSGILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL
DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLL
NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG TALIKKYPKLESEF
VYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET

GEIVWDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK
KYGGFVSPTVAYSVLVVAKEKGSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEV
KKDLIIKLPKYSLFELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLKGSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLF
TLTNLGAPAAFKYFDTTIDRKQYRSTKEVL DATLIHQ SITGLYETRIDLSQLGGDEGADKR
TADGSEFESP KKKRKV

【0 7 6 2】

10. monoABE8.1_bpNLS + Y147R_Q154R_T166R

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLV LNNRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNHRVEITEGILADECAALLCRFFRMPRRVFNAQKKAQSSRDSGGSSGGSS
GSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTD
RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDD SFFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLR LIYLALAHM
IKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRL E
NLIAQLPG EKKNGLFGNLIALSLGLTPNFKS NFDLAEDAKLQLSKD TYDDDLNLLAQIGD
QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPE
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
NGSIPHQIHLGELHAILRRQEDFYFPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR
KSEETITPWNFEVVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN
ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMK
QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKS DGFANRNF MQLIHDDSLTFKEDIQ
KAQVSGQGDSLHEHIANLAGSPA IKKGILQTVKVVD ELVKVMGRHKPENIVIEMARENQT
TQKGQKNSRERMKRIE EGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL
DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKMKNYWRQLL
NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKS KLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG TALIKKYPKLESEF
VYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET
GEIVWDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK
KYGGFVSPTVAYSVLVVAKEKGSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEV
KKDLIIKLPKYSLFELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLKGSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLF
TLTNLGAPAAFKYFDTTIDRKQYRSTKEVL DATLIHQ SITGLYETRIDLSQLGGDEGADKR
TADGSEFESP KKKRKV

【0 7 6 3】

11. monoABE8.1_bpNLS + Y147T_Q154R

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLV LNNRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNHRVEITEGILADECAALLCTFFRMPRRVFNAQKKAQSSTDSSGGSSGGSSG
SETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDR
HSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDD SFFHRL
EESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLR LIYLALAHMI
KFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRL E
NLIAQLPG EKKNGLFGNLIALSLGLTPNFKS NFDLAEDAKLQLSKD TYDDDLNLLAQIGD
QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPE
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
NGSIPHQIHLGELHAILRRQEDFYFPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR
KSEETITPWNFEVVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN

ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTLFEDREMIEERLKTYAHLFDDKVMK
QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQ
KAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDLVKVMGRHKPENIVIAMARENQT
TQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL
DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLL
NAKLITQRKFDNLTAKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG TALIKKYPKLESEF
VYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET
GEIVWDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK
KYGGFVSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEV
KKDLIIKLPKYSLELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLGKSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLF
TLTNLGAPAAFKYFDTTIDRKQYRSTKEVL DATLIHQ SITGLYETRIDLSQLGGDEGADKR
TADGSEFESPKKKRKV

[0 7 6 4]

12. monoABE8.1_bpNLS + Y147T_Q154S

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNRRVIGEGWNRAIGLHDPTAHA
EIMALRQGGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNHRVEITEGILADECAALLCTFFRMPRSVFNAQKKAQSSTDSGGSSGGSSG
SETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDR
HSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDD SFFHRL
EESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLR LIYLALAHMI
KFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRL
NLIAQLPGEKKNGLFGNLIALSLGLTPNFKS NFDLAEDAKLQLSKD TYDDDLNLLAQIGD
QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPE
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
NGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPIYYVGPLARGNSRFAWMTR
KSEETITPWNFEVVVDKGASAQSFIERM TNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN
ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTLFEDREMIEERLKTYAHLFDDKVMK
QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQ
KAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDLVKVMGRHKPENIVIAMARENQT
TQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL
DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLL
NAKLITQRKFDNLTAKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG TALIKKYPKLESEF
VYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET
GEIVWDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK
KYGGFVSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEV
KKDLIIKLPKYSLELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLGKSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLF
TLTNLGAPAAFKYFDTTIDRKQYRSTKEVL DATLIHQ SITGLYETRIDLSQLGGDEGADKR
TADGSEFESPKKKRKV

[0 7 6 5]

13. monoABE8.1_bpNLS + H123Y123H_Y147R_Q154R_I76Y

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNRRVIGEGWNRAIGLHDPTAHA
EIMALRQGGGLVMQNYRLYDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHHPGMNHRVEITEGILADECAALLCRFFRMPRRVFNAQKKAQSSTDSGGSSGGSS
GSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTD
RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDD SFFHR

LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYALAHM
IKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRE
NLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDITYDDDLNLLAQIGD
QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLP
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
NGSIPHQIHLGELHAILRRQEDFYFPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR
KSEETITPWNFEEVVDKGASAQSFIERMNTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN
ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMK
QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQ
KAQVSGQGDSLHEHIANLAGSPAIAKKGILQTVKVVDLKVVMGRHKPENIVIAMARENQT
TQKGQKNSRERMKRIEELGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL
DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLL
NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEF
VYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET
GEIVWDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK
KYGGFVSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSEKKNPIDFLEAKGYKEV
KKDLIIKLPKYSLELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLGSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLF
TLTNLGAPAAFKYFDTTIDRKQYRSTKEVLDATLIHQ SITGLYETRIDLSQLGGDEGADKR
TADGSEFESPKKKRKV

[0766]

14. monoABE8.1_bpNLS + V82S + Q154R

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNRRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYSTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNRHVEITEGILADECAALLCYFFRMPRRVFNAQKKAQSSTDSGGSSGGSS
GSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVP SKFKVLGNTD
RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDDSFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYALAHM
IKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRE
NLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDITYDDDLNLLAQIGD
QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLP
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
NGSIPHQIHLGELHAILRRQEDFYFPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR
KSEETITPWNFEEVVDKGASAQSFIERMNTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN
ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMK
QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQ
KAQVSGQGDSLHEHIANLAGSPAIAKKGILQTVKVVDLKVVMGRHKPENIVIAMARENQT
TQKGQKNSRERMKRIEELGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL
DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLL
NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEF
VYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET
GEIVWDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK
KYGGFVSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSEKKNPIDFLEAKGYKEV
KKDLIIKLPKYSLELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLGSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLF
TLTNLGAPAAFKYFDTTIDRKQYRSTKEVLDATLIHQ SITGLYETRIDLSQLGGDEGADKR

10

20

30

40

50

TADGSEFESPKKKRKV

【0767】

いくつかの実施形態において、塩基エディターは、ポリヌクレオチドにより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）が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）に融合されたものを含む融合タンパク質である。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させる1つまたは複数の特徴を含む。例えば、本明細書で提供される融合タンパク質のいずれかは、ヌクレアーゼ活性が低減された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のいずれかは、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分子の一つの鎖を切断するCas9ドメインを有し得る。

10

【0768】

いくつかの実施形態において、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。いくつかの実施形態において、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。いくつかの実施形態において、塩基エディターは、核酸ポリメラーゼの全部または一部を含むドメインを含む。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る核酸ポリメラーゼまたはその一部は、損傷乗り越え型（translesion）DNAポリメラーゼである。

【0769】

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。

20

30

【0770】

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。いくつかの実施形態では、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

40

50

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 (例えばUGI等) とを連結する。

10

【0771】

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。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 Commun. 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 (rigid)」リンカーとも呼ばれる。

20

30

40

【0772】

本発明の融合タンパク質は、核酸編集ドメインを含む。いくつかの実施形態において、デアミナーゼは、アデノシンデアミナーゼである。いくつかの実施形態において、デアミナーゼは、脊椎動物のデアミナーゼである。いくつかの実施形態において、デアミナーゼは、無脊椎動物のデアミナーゼである。いくつかの実施形態において、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、またはマウスのデアミナーゼである。いくつかの実施形態において、デアミナーゼは、ヒトデアミナーゼである。いく

50

つかの実施形態において、デアミナーゼは、ラットデアミナーゼである。

【0773】

本明細書で使用される場合、「ヘテロ二量体（ヘテロダイマー）」は、野生型TadAドメインとTadA * 7.10ドメインのバリエーション、または2つのバリエーションTadAドメイン（例えば、TadA7.10と、Y147TおよびQ154S変換を伴うTadA7.10）を含む融合タンパク質を指すことができる。

【0774】

いくつかの実施形態では、塩基エディターは、napDNABpに内部で融合または挿入された異種ポリペプチドを含む融合タンパク質を含む。異種ポリペプチド（例えば、デアミナーゼ）は、例えばnapDNABpが標的ポリヌクレオチドおよびガイド核酸に結合するその能力を保持するように、適切な位置でnapDNABp（例えば、Cas9）に挿入され得る。デアミナーゼは、デアミナーゼの機能（例えば、塩基編集活性）またはnapDNABpの機能（例えば、標的核酸およびガイド核酸に結合する能力）を損なうことなく、napDNABpに挿入することができる。デアミナーゼは、例えば、結晶学的研究によって示される無秩序な領域または高温因子もしくはB因子を含む領域でnapDNABpに挿入することができる。秩序化されていない、無秩序である、または構造化されていないタンパク質の領域、例えば、溶媒に曝される領域およびループは、構造または機能を損なうことなく挿入に使用することができる。デアミナーゼは、可撓性ループ領域または溶媒に曝される領域においてnapDNABpに挿入することができる。いくつかの実施形態において、デアミナーゼは、Cas9ポリペプチドの可撓性ループに挿入される。

【0775】

いくつかの実施形態において、デアミナーゼの挿入位置は、Cas9ポリペプチドの結晶構造のB因子分析によって決定される。いくつかの実施形態において、デアミナーゼは、平均よりも高いB因子（例えば、総タンパク質または無秩序領域を含むタンパク質ドメインと比較してより高いB因子）を含むCas9ポリペプチドの領域に挿入される。Bファクター（因子）または温度ファクターは、平均位置からの原子の変動を示すことができる（たとえば、温度に依存する原子の振動または結晶格子の静的無秩序の結果として）。骨格原子の高いBファクター（たとえば、平均よりも高いBファクター）は、比較的高い局所移動度を持つ領域を示している可能性がある。このような領域は、構造または機能を損なうことなくデアミナーゼを挿入するために使用することができる。デアミナーゼは、総タンパク質の平均Bファクターより50%、60%、70%、80%、90%、100%、110%、120%、130%、140%、150%、160%、170%、180%、190%、200%、またはそれ以上高いBファクターのC α 原子を有する残基の位置に挿入することができる。デアミナーゼは、その残基を含むCas9タンパク質ドメインの平均B因子よりも50%、60%、70%、80%、90%、100%、110%、120%、130%、140%、150%、160%、170%、180%、190%、200%、またはそれ以上高いBファクターのC α 原子を有する残基の位置に挿入することができる。平均よりも高いB因子を含むCas9ポリペプチド位置は、例えば、配列番号1における番号付けで残基768、792、1052、1015、1022、1026、1029、1067、1040、1054、1068、1246、1247、および1248を含み得る。平均よりも高いB因子を含むCas9ポリペプチド領域は、例えば、配列番号1における番号付けで残基792-872、792-906、および2-791を含み得る。

【0776】

異種ポリペプチド（例えばデアミナーゼ）は、配列番号1における番号付けで768、791、792、1015、1016、1022、1023、1026、1029、1040、1052、1054、1067、1068、1069、1246、1247、および1248からなる群より選択されるアミノ酸残基、または別のCas9ポリペプチドにおける対応するアミノ酸残基においてnapDNABpに挿入され得る。いくつかの実施形態において、異種ポリペプチドは、配列番号1における番号付けでアミノ酸位置768-769、791-792、792-793、1015-1016、1022-1023、1026-1027、1029-1030、1040-1041、1052-1053、1054-1055、1067-1068、1068-1069、1247-1248、または1248-1249の間、またはそれに対応するアミノ酸位置

の間に挿入される。いくつかの実施形態において、異種ポリペプチドは、配列番号1における番号付けでアミノ酸位置769-770、792-793、793-794、1016-1017、1023-1024、1027-1028、1030-1031、1041-1042、1053-1054、1055-1056、1068-1069、1069-1070、1248-1249、または1249-1250の間、またはそれに対応するアミノ酸位置の間に挿入される。いくつかの実施形態において、異種ポリペプチドは、配列番号1における番号付けで768、791、792、1015、1016、1022、1023、1026、1029、1040、1052、1054、1067、1068、1069、1246、1247、および1248からなる群より選択されるアミノ酸残基、または別のCas9ポリペプチドにおける対応するアミノ酸残基を置き換える。挿入位置に関して配列番号1を参照することは、例示目的であることを理解されたい。ここで論じられる挿入は、配列番号1のCas9ポリペプチド配列に限定されず、例えばCas9ニッカーゼ (nCas9)、ヌクレアーゼ不活Cas9 (dCas9)、ヌクレアーゼドメインを欠くCas9バリエーション、切り詰め型Cas9、または部分的もしくは完全なHNHドメインを欠くCas9ドメインなどの、バリエーションCas9ポリペプチドにおける対応する位置への挿入も含む。

【0777】

異種ポリペプチド（例えばデアミナーゼ）は、配列番号1における番号付けで、768、792、1022、1026、1040、1068、および1247からなる群より選択されるアミノ酸残基、または別のCas9ポリペプチド中の対応するアミノ酸残基においてnapDNAbp中に挿入され得る。いくつかの実施形態において、異種ポリペプチドは、配列番号1における番号付けでアミノ酸位置768-769、792-793、1022-1023、1026-1027、1029-1030、1040-1041、1068-1069、または1247-1248の間、またはそれに対応するアミノ酸位置の間に挿入される。いくつかの実施形態において、異種ポリペプチドは、配列番号1における番号付けでアミノ酸位置769-770、793-794、1023-1024、1027-1028、1030-1031、1041-1042、1069-1070、または1248-1249の間、またはそれに対応するアミノ酸位置の間に挿入される。いくつかの実施形態において、異種ポリペプチドは、配列番号1における番号付けで768、792、1022、1026、1040、1068、および1247からなる群より選択されるアミノ酸残基、または別のCas9ポリペプチドにおける対応するアミノ酸残基を置き換える。

【0778】

いくつかの実施形態において、ABE（例えばTadA）は、配列番号1における番号付けで1015、1022、1029、1040、1068、1247、1054、1026、768、1067、1248、1052、および1246からなる群から選択されるアミノ酸残基、または別のCas9ポリペプチドにおける対応するアミノ酸残基において挿入される。いくつかの実施形態において、ABE（例えばTadA）は、配列番号1における番号付けで残基792-872、792-906、または2-791に置き換わって、または別のCas9ポリペプチドにおける対応するアミノ酸残基に置き換わって挿入される。いくつかの実施形態において、ABEは、配列番号1における番号付けで1015、1022、1029、1040、1068、1247、1054、1026、768、1067、1248、1052、および1246からなる群から選択されるアミノ酸、または別のCas9ポリペプチドにおける対応するアミノ酸残基のN末端に挿入される。いくつかの実施形態において、ABEは、配列番号1における番号付けで1015、1022、1029、1040、1068、1247、1054、1026、768、1067、1248、1052、および1246からなる群から選択されるアミノ酸、または別のCas9ポリペプチドの対応するアミノ酸残基のC末端に挿入される。いくつかの実施形態において、ABEは、配列番号1における番号付けで1015、1022、1029、1040、1068、1247、1054、1026、768、1067、1248、1052、および1246からなる群から選択されるアミノ酸、または別のCas9ポリペプチドにおける対応するアミノ酸残基を置き換えて挿入される。

【0779】

いくつかの実施形態において、CBE（例えばAPOBEC1）は、配列番号1における番号付けで1016、1023、1029、1040、1069、および1247からなる群から選択されるアミノ酸残基、または別のCas9ポリペプチドにおける対応するアミノ酸残基において挿入され

10

20

30

40

50

【 0 7 8 5 】

10

【 0 7 8 6 】

20

【 0 7 8 7 】

30

【 0 7 8 8 】

40

【 0 7 8 9 】

50

【 0 7 9 0 】

【 0 7 9 1 】

10

【 0 7 9 2 】

20

【 0 7 9 3 】

30

【 0 7 9 4 】

40

【 0 7 9 5 】

50

別のCas9ポリペプチドにおける対応するアミノ酸残基のC末端に挿入される。いくつかの実施形態において、ABEは、配列番号1における番号付けでアミノ酸1247、または別のCas9ポリペプチドの対応するアミノ酸残基を置き換えて挿入される。

【0796】

いくつかの実施形態において、デアミナーゼは、配列番号1における番号付けでアミノ酸残基1248、または別のCas9ポリペプチドの対応するアミノ酸残基において挿入される。いくつかの実施形態において、ABEは、配列番号1における番号付けでアミノ酸1248、または別のCas9ポリペプチドにおける対応するアミノ酸残基のN末端に挿入される。いくつかの実施形態において、ABEは、配列番号1における番号付けでアミノ酸1248、または別のCas9ポリペプチドにおける対応するアミノ酸残基のC末端に挿入される。いくつかの実施形態において、ABEは、配列番号1における番号付けでアミノ酸1248、または別のCas9ポリペプチドの対応するアミノ酸残基を置き換えて挿入される。

10

【0797】

ある実施形態において、異種ポリペプチド（例えばデアミナーゼ）は、Cas9ポリペプチドの可撓性ループに挿入される。可撓性ループ部分は、配列番号1における番号付けで530-537、569-570、686-691、943-947、1002-1025、1052-1077、1232-1247、もしくは1298-1300、または別のCas9ポリペプチド中の対応するアミノ酸残基からなる群から選択することができる。可撓性ループ部分は、配列番号1における番号付けで1-529、538-568、580-685、692-942、948-1001、1026-1051、1078-1231、もしくは1248-1297、または別のCas9ポリペプチド中の対応するアミノ酸残基からなる群から選択することができる。

20

【0798】

異種ポリペプチド（例えばデアミナーゼ）は、配列番号1における番号付けでアミノ酸残基1017-1069、1242-1247、1052-1056、1060-1077、1002-1003、943-947、530-537、568-579、686-691、1242-1247、1298-1300、1066-1077、1052-1056、もしくは1060-1077、または別のCas9ポリペプチドの対応するアミノ酸残基に対応する、Cas9ポリペプチド領域に挿入することができる。

【0799】

異種ポリペプチド（例えばデアミナーゼ）はCas9ポリペプチドの欠失領域の代わりに挿入され得る。欠失領域は、Cas9ポリペプチドのN末端またはC末端部分に相当し得る。いくつかの実施形態において、欠失領域は、配列番号1における番号付けで残基792~872、または別のCas9ポリペプチドにおける対応するアミノ酸残基に相当する。いくつかの実施形態において、欠失領域は、配列番号1における番号付けで残基792~906、または別のCas9ポリペプチドにおける対応するアミノ酸残基に相当する。いくつかの実施形態において、欠失領域は、配列番号1における番号付けで残基2~791、または別のCas9ポリペプチドにおける対応するアミノ酸残基に相当する。

30

【0800】

異種ポリペプチド（例えばデアミナーゼ）は、Cas9ポリペプチドの構造的または機能的ドメイン内に挿入され得る。異種ポリペプチド（例えばデアミナーゼ）は、Cas9ポリペプチドの2つの構造的または機能的ドメインの間に挿入することができる。異種ポリペプチド（例えばデアミナーゼ）は、Cas9ポリペプチドの構造的または機能的ドメインの代わりに、例えば、Cas9ポリペプチドからそのドメインを削除した後に、挿入され得る。Cas9ポリペプチドの構造的または機能的ドメインは、例えば、RuvC I、RuvC II、RuvC III、Rec1、Rec2、PI、またはHNHを含み得る。

40

【0801】

いくつかの実施形態において、Cas9ポリペプチドは、RuvC I、RuvC II、RuvC III、Rec1、Rec2、PI、またはHNHドメインからなる群より選択される1以上のドメインを欠く。ある実施形態において、Cas9ポリペプチドは、ヌクレアーゼドメインを欠く。ある実施形態において、Cas9ポリペプチドは、HNHドメインを欠く。いくつかの実施形態において、Cas9ポリペプチドは、そのCas9ポリペプチドがHNH活性を低下させるかまたは消失

50

させるように、HNHドメインの一部を欠く。

【0802】

ある実施形態において、Cas9ポリペプチドは、ヌクレアーゼドメインの欠失を含み、デアミナーゼは、ヌクレアーゼドメインを置き換えて挿入される。いくつかの実施形態において、HNHドメインは削除され、デアミナーゼはその場所に挿入される。いくつかの実施形態において、RuvCドメインの1つ以上が欠失され、デアミナーゼがその場所に挿入される。

【0803】

異種ポリペプチドを含む融合タンパク質は、napDNAbpのN末端およびC末端断片によって隣接され得る。ある実施形態において、融合タンパク質は、Cas9ポリペプチドのN末端断片およびC末端断片によって隣接されるデアミナーゼを含む。該N末端断片またはC末端断片は、標的ポリヌクレオチド配列に結合することができる。N末端断片のC末端またはC末端断片のN末端は、Cas9ポリペプチドの可撓性ループの一部を含み得る。N末端断片のC末端またはC末端断片のN末端は、Cas9ポリペプチドの α ヘリックス構造の一部を含み得る。N末端断片またはC末端断片は、DNA結合ドメインを含み得る。N末端断片またはC末端断片は、RuvCドメインを含み得る。N末端断片またはC末端断片は、HNHドメインを含み得る。いくつかの実施形態において、N末端断片およびC末端断片のいずれもHNHドメインを含まない。

【0804】

いくつかの実施形態において、N末端Cas9断片のC末端は、融合タンパク質が標的核酸塩基を脱アミノ化する際に標的核酸塩基に近接するアミノ酸を含む。いくつかの実施形態において、C末端Cas9断片のN末端は、融合タンパク質が標的核酸塩基を脱アミノ化する際に標的核酸塩基に近接するアミノ酸を含む。標的核酸塩基とN末端Cas9断片のC末端またはC末端Cas9断片のN末端におけるアミノ酸との間の近接性を有するためには、異なるデアミナーゼは挿入位置が異なり得る。例えば、ABEの挿入位置は、配列番号1における番号付けで1015、1022、1029、1040、1068、1247、1054、1026、768、1067、1248、1052、および1246からなる群から選択されるアミノ酸残基、または別のCas9ポリペプチドにおける対応するアミノ酸残基におけるものであり得る。CBEの適切な挿入位置は、配列番号1における番号付けで1016、1023、1029、1040、1069、および1247からなる群より選択されるアミノ酸残基、または別のCas9ポリペプチドの対応するアミノ酸残基であり得る。ある実施形態において、ABEの挿入は、上記のアミノ酸残基のいずれか1つのN末端またはC末端に挿入され得る。いくつかの実施形態において、ABEの挿入は、上記のアミノ酸残基のいずれか1つを置き換えて挿入され得る。

【0805】

融合タンパク質のC末端Cas9断片（すなわち、融合タンパク質中のデアミナーゼに隣接するC末端Cas9断片）は、Cas9ポリペプチドのC末端を含み得る。融合タンパク質のC末端Cas9断片は、少なくとも約100、200、300、400、500、600、700、800、900、1000、1100、1200、または1300アミノ酸の長さを含むことができる。融合タンパク質のC末端Cas9断片は、配列番号1における番号付けでアミノ酸残基1-56、1-95、1-200、1-300、1-400、1-500、1-600、1-700、1-718、1-765、1-780、1-906、1-918、もしくは1-1100、または別のCas9ポリペプチドの対応するアミノ酸残基に相当する配列を含み得る。N末端Cas9断片は、配列番号1における番号付けでアミノ酸残基1-56、1-95、1-200、1-300、1-400、1-500、1-600、1-700、1-718、1-765、1-780、1-906、1-918、もしくは1-1100または別のCas9ポリペプチド中の対応するアミノ酸残基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%の配列同一性を含む配列を含み得る。

【0806】

融合タンパク質のN末端Cas9断片（すなわち、融合タンパク質中のデアミナーゼに隣接

10

20

30

40

50

するN末端Cas9断片)は、Cas9ポリペプチドのN末端を含み得る。融合タンパク質のN末端Cas9断片は、少なくとも約100、200、300、400、500、600、700、800、900、1000、1100、1200、または1300アミノ酸の長さを含むことができる。融合タンパク質のN末端Cas9断片は、配列番号1における番号付けでアミノ酸残基1099-1368、918-1368、906-1368、780-1368、765-1368、718-1368、94-1368、もしくは56-1368、または別のCas9ポリペプチドの対応するアミノ酸残基に相当する配列を含み得る。N末端Cas9断片は、配列番号1における番号付けでアミノ酸残基1099-1368、918-1368、906-1368、780-1368、765-1368、718-1368、94-1368、もしくは56-1368または別のCas9ポリペプチド中の対応するアミノ酸残基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%の配列同一性を含む配列を含み得る。

10

【0807】

融合タンパク質のN末端Cas9断片およびC末端Cas9断片を合わせたものは、例えば、配列番号1に記載されているもののような、全長の天然Cas9ポリペプチド配列に相当していなくてもよい。

【0808】

本明細書に記載された融合タンパク質は、非標的部位(例えばオフターゲット部位)での脱アミノ化が低減された、例えばゲノム全体に渡り目的外脱アミノ化が低減された、標的化脱アミノ化をもたらすことができる。本明細書に記載される融合タンパク質は、非標的部位におけるバスターゲット脱アミノ化が低減された標的化脱アミノ化をもたらすことができる。望まれない脱アミノ化、またはオフターゲット脱アミノ化は、例えばCas9ポリペプチドのN末端またはC末端に融合されたデアミナーゼを含むエンド末端融合タンパク質と比較して、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95%、または少なくとも99%低減され得る。望まれない脱アミノ化、またはオフターゲット脱アミノ化は、例えばCas9ポリペプチドのN末端またはC末端に融合されたデアミナーゼを含むエンド末端融合タンパク質と比較して、少なくとも1倍(one-fold)、少なくとも2倍、少なくとも3倍、少なくとも4倍、少なくとも5倍、少なくとも10倍、少なくとも15倍、少なくとも20倍、少なくとも30倍、少なくとも40倍、少なくとも50倍、少なくとも60倍、少なくとも70倍、少なくとも80倍、少なくとも90倍、または少なくとも100倍低下させることができる。

20

30

【0809】

いくつかの実施形態において、融合タンパク質のデアミナーゼは、Rループの範囲内で2つを超える核酸塩基を脱アミノ化しない。いくつかの実施形態において、融合タンパク質のデアミナーゼは、Rループの範囲内で3つを超える核酸塩基を脱アミノ化しない。いくつかの実施形態において、融合タンパク質のデアミナーゼは、Rループの範囲内で2、3、4、5、6、7、8、9、または10を超える核酸塩基を脱アミノ化しない。Rループは、DNA:RNAハイブリッド、DNA:DNAまたはRNA:RNA相補的構造を含み、一本鎖DNAと会合した三本鎖核酸構造である。本明細書中で用いられる場合、R-ループは、標的ポリヌクレオチドがCRISPR複合体または塩基編集複合体と接触された場合に形成され得、ここで、ガイドポリヌクレオチド(例えばガイドRNA)の一部が、標的ポリヌクレオチド(例えば標的DNA)の一部とハイブリダイズし、それと取って代わる。ある実施形態において、Rループは、スパーサー配列および標的DNA相補配列のハイブリダイズした領域を含む。Rループ領域の長さは約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核酸塩基対であり得る。いくつかの実施形態において、Rループ領域は長さが約20核酸塩基対である。本明細書で使用される場合、Rループ領域は、ガイドポリヌクレオチドとハイブリダイズする標的DNA鎖には限定されないことが理解されるべきである。例えば、Rループ領域内の標的核酸塩基の編集は、ガイドRNAと相補的な鎖を含むDNA鎖に対してであってもよく、

40

50

またはガイドRNAと相補的な鎖の反対側の鎖であるDNA鎖に対してであってもよい。ある実施形態において、Rループの領域における編集は、標的DNA配列中のうち、ガイドRNAに対して非相補的な鎖（プロトスペーサー鎖）上の核酸塩基の編集を含む。

【0810】

本明細書に記載される融合タンパク質は、正準的（canonical）な塩基編集とは異なる編集ウィンドウにおいて標的脱アミノ化をもたらし得る。ある実施形態において、標的核酸塩基は、標的ポリヌクレオチド配列におけるPAM配列の約1～約20塩基上流にある。ある実施形態において、標的核酸塩基は、標的ポリヌクレオチド配列におけるPAM配列の約2～約12塩基上流にある。いくつかの実施形態において、標的核酸塩基は、PAM配列から約1～9塩基対、約2～10塩基対、約3～11塩基対、約4～12塩基対、約5～13塩基対、約6～14塩基対、約7～15塩基対、約8～16塩基対、約9～17塩基対、約10～18塩基対、約11～19塩基対、約12～20塩基対、約1～7塩基対、約2～8塩基対、約3～9塩基対、約4～10塩基対、約5～11塩基対、約6～12塩基対、約7～13塩基対、約8～14塩基対、約9～15塩基対、約10～16塩基対、約11～17塩基対、約12～18塩基対、約13～19塩基対、約14～20塩基対、約1～5塩基対、約2～6塩基対、約3～7塩基対、約4～8塩基対、約5～9塩基対、約6～10塩基対、約7～11塩基対、約8～12塩基対、約9～13塩基対、約10～14塩基対、約11～15塩基対、約12～16塩基対、約13～17塩基対、約14～18塩基対、約15～19塩基対、約16～20塩基対、約1～3塩基対、約2～4塩基対、約3～5塩基対、約4～6塩基対、約5～7塩基対、約6～8塩基対、約7～9塩基対、約8～10塩基対、約9～11塩基対、約10～12塩基対、約11～13塩基対、約12～14塩基対、約13～15塩基対、約14～16塩基対、約15～17塩基対、約16～18塩基対、約17～19塩基対、約18～20塩基対だけ離れているか、またはそれだけ上流にある。いくつかの実施形態において、標的核酸塩基は、PAM配列から約1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20もしくはそれより多くの塩基対だけ離れているか、またはそれだけ上流にある。いくつかの実施形態では、標的核酸塩基は、PAM配列の約1、2、3、4、5、6、7、8、または9塩基対上流にある。いくつかの実施形態において、標的核酸塩基は、PAM配列の約2、3、4または6塩基対上流にある。

【0811】

従って、例えばBE4のような正準的な塩基エディターと比較して別の塩基編集ウィンドウ嗜好性を可能にする、塩基編集を最適化するために融合タンパク質ライブラリーおよびその使用方法も本明細書で提供される。いくつかの実施形態において、本開示は、複数の融合タンパク質を含む、最適化された塩基編集のためのタンパク質ライブラリーを提供し、ここで、該複数の融合タンパク質の各々は、Cas9ポリペプチドのN末端断片およびC末端断片によって隣接されたデアミナーゼを含み、融合タンパク質の各々のN末端断片が、該複数の融合タンパク質の残りのもののN末端断片と異なるか、または、融合タンパク質の各々のC末端断片が、該複数の融合タンパク質の残りのもののC末端断片と異なり、該融合タンパク質の各々のデアミナーゼは、標的ポリヌクレオチド配列中のプロトスペーサー隣接モチーフ（Protospacer Adjacent Motif：PAM）配列に近接した標的核酸塩基を脱アミノ化し、該N末端断片またはC末端断片は標的ポリヌクレオチド配列に結合する。ある実施形態において、CRISPR Rループ内の各核酸塩基について、該複数の融合タンパク質のうちの少なくとも1つが、核酸塩基を脱アミノ化する。いくつかの実施形態において、標的ポリヌクレオチドのうちPAM配列から1～20塩基対離れた各核酸塩基について、該複数の融合タンパク質のうちの少なくとも一つが、核酸塩基を脱アミノ化する。いくつかの実施形態において、本明細書では、最適化された塩基編集を可能にする融合タンパク質ライブラリーを含むキットが提供される。

【0812】

融合タンパク質は、2つ以上の異種ポリペプチドを含むことができる。例えば、融合タンパク質は、1つ以上のUGIドメインおよび/または1つ以上の核局在化シグナルをさらに含み得る。2つ以上の異種ドメインは、タンデムに挿入され得る。2つ以上の異種ドメインは、それらがNapDNAbp中でタンデムにならないような位置に挿入され得る。

【0813】

いくつかの実施形態において、塩基エディターは、内部に融合された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）を有するnapDNAbpドメイン（例えばCas12由来ドメイン）を含む融合タンパク質を含む。いくつかの実施形態では、napDNAbpはCas12bである。いくつかの実施形態において、塩基エディターは、以下の表Aに提供される遺伝子座に挿入された内部融合TadA*8ドメインを有するBhCas12bドメインを含む。

【0814】

表A：Cas12bタンパク質における挿入遺伝子座

BhCas12b	挿入部位	挿入された両隣のアミノ酸
position 1	153	PS
position 2	255	KE
position 3	306	DE
position 4	980	DG
position 5	1019	KL
position 6	534	FP
position 7	604	KG
position 8	344	HF
BvCas12b	挿入部位	挿入された両隣のアミノ酸
position 1	147	PD
position 2	248	GG
position 3	299	PE
position 4	991	GE
position 5	1031	KM
AaCas12b	挿入部位	挿入された両隣のアミノ酸
position 1	157	PG
position 2	258	VG
position 3	310	DP
position 4	1008	GE
position 5	1044	GK

【0815】

いくつかの実施形態では、塩基エディターは、複数のドメインを含むことができる。例えば、Cas12タンパク質に由来するnapDNAbpドメインを含む塩基エディターは、野生型あるいは天然のCas12のRECローブ（lobe）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、1つまたは複数のRuvCドメインWEDドメインを含むことができる。いくつかの実施形態では、塩基エディターの1つまたは複数のドメインは、そのドメインを含むポリペプチドの野生型バージョンと比較して突然変異（例えば、置換、挿入、欠失）を含む。

【0816】

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。

10

【0817】

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。

【0818】

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および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および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。

20

30

【0819】

[リンカー]

ある実施形態において、本発明のペプチドまたはペプチドドメインのいずれかを連結するためにリンカーが使用され得る。リンカーは、共有結合のように単純であり得、またはそれは、多原子の長さであるポリマーリンカーであり得る。ある実施形態において、リンカーは、ポリペプチドであるか、またはアミノ酸に基づくものである。他の実施形態において、リンカーはペプチド様ではない。ある実施形態において、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、等）である。ある実施形態において、リンカーは、アミド連結の炭素-窒素結合である。特定の実施形態では、リンカーは、環状又は非環状、置換又は非置換、分枝又は非分枝の脂肪族又はヘテロ脂肪族リンカーである。ある実施形態において、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、その他）である。特定の実施形態では、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。ある実施形態において、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、ベータ-アラニン、3-アミノプロパン酸、4-アミノブタン酸、5-ペンタン酸、等）を含む。特定の実施形態では、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

40

50

タン、シクロヘキサン)に基づく。他の実施形態において、リンカーは、ポリエチレングリコール部分 (PEG) を含む。他の実施形態において、リンカーはアミノ酸を含む。ある態様において、リンカーはペプチドを含む。ある実施形態において、リンカーは、アリールまたはヘテロアリール部分を含む。ある実施形態において、リンカーは、フェニル環に基づく。リンカーは、ペプチドからの求核剤 (例えばチオール、アミノ) がリンカーに結合することを促進するための官能化部分を含み得る。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。

【0820】

ある実施形態において、リンカーは、一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。いくつかの実施形態において、リンカーは、結合 (例えば共有結合)、有機分子、基、ポリマー、または化学的部分である。ある実施形態において、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、または100) アミノ酸である。

【0821】

いくつかの実施形態において、アデノシンデアミナーゼおよびnapDNAbpは、長さが4、16、32、または104アミノ酸である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、長さが約3から約104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、リンカーを介して互いに融合されたアデノシンデアミナーゼおよびCas9ドメインを含む。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ために、デアミナーゼドメイン (例えば、工学操作されたecTadA) およびCas9ドメインの間で種々のリンカーの長さおよび柔軟性を利用することができる (例えば、(GGGS)_n、(GGGGS)_n、および(G)_nの形態の非常に柔軟なリンカーから、(EAAAK)_n、(SGGS)_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)、および(XP)_n)。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。ある実施形態において、リンカーは、(GGGS)_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いずれかのアデノシンデアミナーゼおよびCas9ドメインは、アミノ酸配列SGSETPGTSESATPESを含むリンカー (例えばXTENリンカー) を介して融合される。

【0822】

[ガイドRNAを伴うCas9複合体]

本開示のいくつかの側面は、本明細書で提供された融合タンパク質のいずれかと、融合タンパク質のCas9ドメイン (例えばdCas9、ヌクレアーゼ活性Cas9、またはCas9ニックアーゼ) に結合されたガイドRNA (例えば、A\突然変異を標的とするガイド) とを含む複合体を提供する。これらの複合体はリボヌクレオタンパク質 (RNP) とも呼ばれる。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ために、融合タンパク質のドメインを連結する任意の方法が利用され得る (例えば、(GGGS)_n、(GGGGS)_n、および(G)_nの形態の非常に柔軟なリンカーから、(EAAAK)_n、(SGGS)_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)、および(XP)_n)。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。ある実施形態におい

10

20

30

40

50

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。

【0823】

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1に列挙されている配列または5'-NAA-3')にすぐ隣接している。いくつかの実施形態において、ガイド核酸(例えばガイドRNA)は、目的の遺伝子(例えば、疾患または障害に関連する遺伝子)の配列に相補的である。

【0824】

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、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、標準PAM配列(NGG)にすぐ隣接していない。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、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。

【0825】

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と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。

【0826】

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。

【0827】

[アデノシンデアミナーゼバリエーションおよびCas9ドメインを含む融合タンパク質を使用する方法]

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、タンパク質の突然変異形態をコードする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、正準PAM配列(NGG)に直接隣接する。いくつかの実施形態において、標的配列の3'末端は、正準PAM配列(NGG)に直接隣接していない。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、NGCG、NGN、NNGRRT、NNNRRT、NGCG、NGCN、NGTN、NGTN、または5' (TTTV) 配列にすぐ隣接している。

【0828】

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と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。

【0829】

本明細書に開示された、Cas9ドメインとアデノシンデアミナーゼバリエーション(例えばABE8)を含む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、例えばsgRNA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。

【0830】

[ガイドRNAを伴うCas12複合体]

本開示のいくつかの側面は、本明細書で提供された融合タンパク質のいずれかと、ガイドRNA(例えば、標的ポリヌクレオチドを編集のための標的とするガイド)とを含む複合体を提供する。

【0831】

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配列にすぐ隣接している。

【0832】

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、例えばTTN、DTTN、GTTN、ATTN、ATTC、DTTNT、WTTN、HATY、TTTN、TTTV、TTTC、TG、RTR、またはYTN PAMサイトにすぐ隣接している。

【0833】

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と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。

【0834】

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、Cas12結合を可能にするtracrRNAフレームワークと、Cas12: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。Cas12: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。

【0835】

本明細書に開示される塩基エディターのドメインは、デアミナーゼドメインがCas12タンパク質に内在化されている限り、任意の順序で配置することができる。例えばCas12ドメインと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-[Cas12ドメイン]-リンカー1-[ABE8]-リンカー2-[Cas12ドメイン]-COOH;

NH2-[Cas12ドメイン]-リンカー1-[ABE8]-[Cas12ドメイン]-COOH;

NH2-[Cas12ドメイン]-[ABE8]-リンカー2-[Cas12ドメイン]-COOH;

NH2-[Cas12ドメイン]-[ABE8]-[Cas12ドメイン]-COOH;

NH2-[Cas12ドメイン]-リンカー1-[ABE8]-リンカー2-[Cas12ドメイン]-[イノシンBER

10

20

30

40

50

阻害因子]-COOH;

NH2-[Cas12ドメイン]-リンカー1-[ABE8]-[Cas12ドメイン]-[イノシンBER阻害因子]-COOH;

NH2-[Cas12ドメイン]-[ABE8]-リンカー2-[Cas12ドメイン]-[イノシンBER阻害因子]-COOH;;

NH2-[Cas12ドメイン]-[ABE8]-[Cas12ドメイン]-[イノシンBER阻害因子]-COOH;

NH2-[イノシンBER阻害因子]-[Cas12ドメイン]-リンカー1-[ABE8]-リンカー2-[Cas12ドメイン]-COOH;

NH2-[イノシンBER阻害因子]-[Cas12ドメイン]-リンカー1-[ABE8]-[Cas12ドメイン]-COOH;

NH2-[イノシンBER阻害因子]-[Cas12ドメイン]-[ABE8]-リンカー2-[Cas12ドメイン]-COOH;

NH2-[イノシンBER阻害因子]NH2-[Cas12ドメイン]-[ABE8]-[Cas12ドメイン]-COOH;
【0836】

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。Komor, A.C., et al., “Improved base excision repair in 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。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。すべてのドメインの長さが野生型ドメインと同じであるそのような塩基エディターの非限定的な例としては、以下のものが挙げられる：

NH2-[Cas12ドメイン]-リンカー1-[APOBEC1]-リンカー2-[Cas12ドメイン]-COOH;

NH2-[Cas12ドメイン]-リンカー1-[APOBEC1]-[Cas12ドメイン]-COOH;

NH2-[Cas12ドメイン]-[APOBEC1]-リンカー2-[Cas12ドメイン]-COOH;

NH2-[Cas12ドメイン]-[APOBEC1]-[Cas12ドメイン]-COOH;

NH2-[Cas12ドメイン]-リンカー1-[APOBEC1]-リンカー2-[Cas12ドメイン]-[UGI]-COOH;

NH2-[Cas12ドメイン]-リンカー1-[APOBEC1]-[Cas12ドメイン]-[UGI]-COOH;

NH2-[Cas12ドメイン]-[APOBEC1]-リンカー2-[Cas12ドメイン]-[UGI]-COOH;

NH2-[Cas12ドメイン]-[APOBEC1]-[Cas12ドメイン]-[UGI]-COOH;

NH2-[UGI]-[Cas12ドメイン]-リンカー1-[APOBEC1]-リンカー2-[Cas12ドメイン]-COOH;

NH2-[UGI]-[Cas12ドメイン]-リンカー1-[APOBEC1]-[Cas12ドメイン]-COOH;

NH2-[UGI]-[Cas12ドメイン]-[APOBEC1]-リンカー2-[Cas12ドメイン]-COOH;

NH2-[UGI]-[Cas12ドメイン]-[APOBEC1]-[Cas12ドメイン]-COOH;

【0837】

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。Komor, A.C., et al., “Programmable editing of a target base in genomic

10

20

30

40

50

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。

【0838】

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。

【0839】

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napDNAbp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、napDNAbpドメインに対してC末端に位置する。

【0840】

融合タンパク質に含まれるタンパク質ドメインは、異種機能ドメインであり得る。融合タンパク質に含めることができるタンパク質ドメインの非限定的な例には、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）、ウラ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、およびレポーター遺伝子配列が含まれる。タンパク質ドメインは、例えば、以下の活性のうちの1つ以上を有する異種機能ドメインであり得る：転写活性化活性、転写抑制活性、転写放出因子活性、遺伝子サイレンシング活性、クロマチン修飾活性、エピジェネティック修飾活性、ヒストン修飾活性、RNA切断活性、および核酸結合活性。そのような異種機能ドメインは、標的DNAに付随する標的ポリペプチド（例えば、ヒストン、DNA結合タンパク質など）の修飾などの、機能活性を付与し得、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をもたらすことができる。付与される他の機能および／または活性には、トランスポゼイズ活性、インテグラーゼ活性、リコンビナーゼ活性、リガーゼ活性、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、脱デニル化活性、SUMO化活性、脱SUMO化活性、または上記の任意の組み合わせが含まれ得る。

【0841】

ドメインは、エピトープタグ、レポータータンパク質、その他の結合ドメインで検出または標識され得る。エピトープタグの非限定的な例には、ヒスチジン (His) タグ、V5タグ、FLAGタグ、インフルエンザ血球凝集素 (HA) タグ、Mycタグ、VSV-Gタグ、およびチオレドキシン (Trx) タグが含まれる。レポーター遺伝子の例には、グルタチオン-S-トランスフェラーゼ (GST)、ホースラディッシュペルオキシダーゼ (HRP)、クロラムフェニコールアセチルトランスフェラーゼ (CAT)、ベータガラクトシダーゼ、ベータグルクロニダーゼ、ルシフェラーゼ、緑色蛍光タンパク質 (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含まれますが、これらに限定されない。追加のタンパク質配列としては、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、GAL4 DNA結合ドメイン融合、および単純ヘルペスウイルス (HSV) BP16タンパク質融合などを含むがこれらに限定されない、DNA分子または他の細胞分子に結合するアミノ酸配列が挙げられる。

【0842】

いくつかの実施形態において、BhCas12bガイドポリヌクレオチドは、以下の配列を有する：

BhCas12b sgRNAスキファールド（下線）+ 20nt～23ntガイド配列（N_nによって示される）

5 GTTCTGTCTTTTGGTCAGGACAACCGTCTAGCTATAAGTGCTGCAGGGTGTGAGAAAC
TCCTATTGCTGGACGATGTCTCTTACGAGGCATTAGCACNNNNNNNNNNNNNNNNNNNN
NN-3'

【0843】

いくつかの実施形態において、BvCas12bおよびAaCas12bガイドポリヌクレオチドは、以下の配列を有する：

BvCas12b sgRNAスキファールド（下線）+ 20nt～23ntガイド配列（N_nによって示される）

5 GACCTATAGGGTCAATGAATCTGTGCGTGTGCCATAAGTAATTAATAAAATTACCCACCA
CAGGAGCACCTGAAAACAGGTGCTTGGCACNNNNNNNNNNNNNNNNNNNNNNNNNNNN-3'

AaCas12b sgRNAスキファールド（下線）+ 20nt～23ntガイド配列（N_nによって示される）

5 GTCTAAAGGACAGAATTTTTCAACGGGTGTGCCAATGGCCACTTTCCAGGTGGCAAAG
CCCGTTGAACCTTCTCAAAAAGAACGATCTGAGAAGTGGCACNNNNNNNNNNNNNNNNNNNN
NNNN-3'

【0844】

[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、ゲノム編集を提供するための新しい方法を提供する。

【0845】

本発明の融合タンパク質は、著しい割合のインデルを生成することなく、変異を含むタンパク質をコードする特定のヌクレオチド塩基を有利に改変させる。本明細書において使用されるところの「インデル(indel)」は、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、核酸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かが、インデルと比べて意図された改変(例えば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。

【0846】

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かは、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。

【0847】

いくつかの実施形態において、本明細書に記載されるABE8塩基エディターバリエーションの1つを含む塩基エディター系のいずれかは、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。いくつかの実施形態において、本明細書に記載のABE8塩基エディターバリエーションの1つを含む塩基エディターシステムのうちいずれかは、標的ポリヌクレオチド配列において0.8%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、本明細書に記載のABE8塩基エディターバリエーションの1つを含む塩基エディターシステムのうちいずれかは、標的ポリヌクレオチド配列において最大で0.8%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、本明細書に記載のABE8塩基エディターバリエーションの1つを含む塩基エディターシステムのうちいずれかは、標的ポリヌクレオチド配列において0.3%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、記載されたABE8塩基エディターバリエーションの1つを含む塩基エディターシステムのうちいずれかは、ABE7塩基エディターの1つを含む塩基エディターシステムと比較して、標的ポリヌクレオチド配列におけるより低いインデル形成をもたらす。いくつかの実施形態において、本明細書に記載のABE8塩基エディターバリエーションの1つを含むエディターシステムのうちいずれかは、ABE7.10を含む塩基エディターシステムと比較して、標的ポリヌクレオチド配列におけるより低いインデル形成をもたらす。

【0848】

いくつかの実施形態では、本明細書に記載のABE8塩基エディターバリエーションの1つを含む塩基エディターシステムのうちいずれかは、ABE7塩基エディターの1つを含む塩基エディターシステムと比較してインデル頻度が減少している。いくつかの実施形態では、本明細書に記載のABE8塩基エディターバリエーションの1つを含む塩基エディターシステムのうちいずれかは、ABE7塩基エディターの1つを含む塩基エディターシステムと比較してインデル頻度が少なくとも0.01%、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%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、または少なくとも95%減少している。いくつかの実施形態では、本明細書に記載のABE8塩基エディターバリエーションの1つを含む塩基エディターシステムは、ABE7.10を含む塩基エディターシステムと比較してインデル頻度が少なくとも0.01%、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%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、または少なくとも95%減少している。

【0849】

本発明は、効率および特異性が増加したアデノシンデアミナーゼバリエーション（例えば、ABE8バリエーション）を提供する。特に、本明細書に記載のアデノシンデアミナーゼバリエーションは、ポリヌクレオチド内の所望の塩基を編集する可能性がより高く、塩基編集ウィン

10

20

30

40

50

ドウ中で改変されることが意図されていない塩基（例えば「バイスタンダー」）を編集する可能性がより低い。

【0850】

いくつかの実施形態では、本明細書に記載のABE8塩基エディターバリエーションの1つを含む塩基編集システムのいずれかは、バイスタンダーの編集または変異が低減されている。いくつかの実施形態において、意図しない編集または変異は、バイスタンダー変異またはバイスタンダー編集、例えば、標的ヌクレオチド配列の標的ウィンドウ中の意図しない位置または非標的位置における標的塩基（例えばAまたはC）の塩基編集である。いくつかの実施形態では、本明細書に記載のABE8塩基エディターバリエーションの1つを含む塩基編集システムのいずれかは、ABE7塩基エディター、例えば、ABE7.10を含む塩基エディターシステムと比較して、バイスタンダーの編集または変異が減少している。いくつかの実施形態において、本明細書に記載のABE8塩基エディターバリエーションの1つを含む塩基編集システムのいずれかは、ABE7塩基エディター、例えばABE7.10を含む塩基エディターシステムと比べてバイスタンダーの編集または変異が少なくとも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%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも95%、または少なくとも99%低減されている。いくつかの実施形態において、本明細書に記載のABE8塩基エディターバリエーションの1つを含む塩基編集システムのいずれかは、ABE7塩基エディター、例えばABE7.10を含む塩基エディターシステムと比べてバイスタンダーの編集または変異が少なくとも1.1倍、少なくとも1.2倍、少なくとも1.3倍、少なくとも1.4倍、少なくとも1.5倍、少なくとも1.6倍、少なくとも1.7倍、少なくとも1.8倍、少なくとも1.9倍、少なくとも2.0倍、少なくとも2.1倍、少なくとも2.2倍、少なくとも2.3倍、少なくとも2.4倍、少なくとも2.5倍、少なくとも2.6倍、少なくとも2.7倍、少なくとも2.8倍、少なくとも2.9倍、または少なくとも3.0倍低減されている。

【0851】

いくつかの実施形態では、本明細書に記載のABE8塩基エディターバリエーションの1つを含む塩基編集システムのいずれかは、目的外（spurious）の編集が低減されている。いくつかの実施形態において、意図しない編集または変異は、目的外の変異または目的外の編集、例えば、ゲノムの意図しない領域または非標的領域における標的塩基（例えばAまたはC）の非特異的編集またはガイド非依存性編集である。いくつかの実施形態では、本明細書に記載のABE8塩基エディターバリエーションの1つを含む塩基編集システムのいずれかは、ABE7塩基エディター、例えば、ABE7.10を含む塩基エディターシステムと比較して、目的外編集が低減されている。いくつかの実施形態では、本明細書に記載のABE8塩基エディターバリエーションの1つを含む塩基編集システムのいずれかは、ABE7塩基エディター、例えば、ABE7.10を含む塩基エディターシステムと比較して、目的外編集が少なくとも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%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも95%、または少なくとも99%低減されている。いくつかの実施形態では、本明細書に記載のABE8塩基エディターバリエーションの1つを含む塩基編集システムのいずれかは、ABE7塩基エディター、例えば、ABE7.10を含む塩基エディターシステムと比較して、目的外編集が少なくとも1.1倍、少なくとも1.2倍、少なくとも1.3倍、少なくとも1.4倍、少なくとも1.5倍、少なくとも1.6倍、少なくとも1.7倍、少なくとも1.8倍、少なくとも1.9倍、少なくとも2.0倍、少なくとも2.1倍、少なくとも2.2倍、少なくとも2.3倍、少なくとも2.4倍、少なくとも2.5倍、少なくとも2.6倍、少なくとも2.7倍、少なくとも2.8倍、少なくとも2.9倍、または少なくとも3.0倍低減さ

れている。

【0852】

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、例えば非意図的な点突然変異など、有意な数の非意図的な突然変異（すなわちバースタンドの変異）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。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。

10

【0853】

いくつかの実施形態において、本明細書に記載のABE8塩基エディターバリエーションのいずれかは、少なくとも0.01%、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%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも95%、または少なくとも99%の塩基編集効率を有する。いくつかの実施形態において、塩基編集効率は、細胞の集団における編集された核酸塩基のパーセンテージを計算することによって測定され得る。いくつかの実施形態では、本明細書に記載のABE8塩基エディターバリエーションのいずれかは、細胞の集団における編集された核酸塩基により測定される場所で少なくとも0.01%、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%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも95%、または少なくとも99%の塩基編集効率を有する。

20

【0854】

いくつかの実施形態では、本明細書に記載のABE8塩基エディターバリエーションのいずれも、ABE7塩基エディターと比較してより高い塩基編集効率を有する。いくつかの実施形態では、本明細書に記載のABE8塩基エディターバリエーションのいずれかは、ABE7塩基エディター（例えばABE7.10）と比較して、少なくとも1%、少なくとも2%、少なくとも3%、少なくとも4%、少なくとも5%、少なくとも10%、少なくとも15%、少なくとも20%、少なくとも25%、少なくとも30%、少なくとも35%、少なくとも40%、少なくとも45%、少なくとも50%、少なくとも55%、少なくとも60%、少なくとも65%、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも95%、少なくとも99%、少なくとも100%、少なくとも105%、少なくとも110%、少なくとも115%、少なくとも120%、少なくとも125%、少なくとも130%、少なくとも135%、少なくとも140%、少なくとも145%、少なくとも150%、少なくとも155%、少なくとも160%、少なくとも165%、少なくとも170%、少なくとも175%、少なくとも180%、少なくとも185%、少なくとも190%、少なくとも195%、少なくとも200%、少なくとも210%、少なくとも220%、少なくとも230%、少なくとも240%、少なくとも250%、少なくとも260%、少なくとも270%、少なくとも280%、少なくとも290%、少なくとも300%、少なくとも310%、少なくとも320%、少なくとも330%、少なくとも340%、少なくとも350%、少なくとも360%、少なくとも370%、少なくとも380%、少なくとも390%、少なくとも400%、少なくとも450%、または少なくとも500%高い塩基編集効率を有する。

30

40

【0855】

いくつかの実施形態において、本明細書に記載のABE8塩基エディターバリエーションのい

50

ずれかは、ABE7塩基エディター（例えばABE7.10）と比較して、少なくとも1.1倍、少なくとも1.2倍、少なくとも1.3倍、少なくとも1.4倍、少なくとも1.5倍、少なくとも1.6倍、少なくとも1.7倍を有する。少なくとも1.8倍、少なくとも1.9倍、少なくとも2.0倍、少なくとも2.1倍、少なくとも2.2倍、少なくとも2.3倍、少なくとも2.4倍、少なくとも2.5倍、少なくとも2.6倍、少なくとも2.7倍、少なくとも2.8倍、少なくとも2.9倍、少なくとも3.0倍、少なくとも3.1倍、少なくとも3.2倍、少なくとも3.3倍、少なくとも3.4倍、少なくとも3.5倍、少なくとも3.6倍、少なくとも3.7倍、少なくとも3.8倍、少なくとも3.9倍、少なくとも4.0倍、少なくとも4.1倍、少なくとも4.2倍、少なくとも4.3倍、少なくとも4.4倍、少なくとも4.5倍、少なくとも4.6倍、少なくとも4.7倍、少なくとも4.8倍、少なくとも4.9倍、または少なくとも5.0倍高い塩基編集効率を有する。

10

【0856】

いくつかの実施形態において、本明細書に記載のABE8塩基エディターバリエーションのいずれかは、少なくとも0.01%、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%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも95%、または少なくとも99%のオンターゲット塩基編集効率を有する。いくつかの実施形態では、本明細書に記載のABE8塩基エディターバリエーションのいずれかは、細胞の集団における編集された標的核酸塩基により測定されるところで少なくとも0.01%、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%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも95%、または少なくとも99%のオンターゲット塩基編集効率を有する。

20

【0857】

いくつかの実施形態では、本明細書に記載のABE8塩基エディターバリエーションのいずれかは、ABE7塩基エディターと比較して、より高いオンターゲット塩基編集効率を有する。いくつかの実施形態では、本明細書に記載のABE8塩基エディターバリエーションのいずれかは、ABE7塩基エディタ（例えばABE7.10）と比較して、少なくとも1%、少なくとも2%、少なくとも3%、少なくとも4%、少なくとも5%、少なくとも10%、少なくとも15%、少なくとも20%、少なくとも25%、少なくとも30%、少なくとも35%、少なくとも40%、少なくとも45%、少なくとも50%、少なくとも55%、少なくとも60%、少なくとも65%、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも95%、少なくとも99%、少なくとも100%、少なくとも105%、少なくとも110%、少なくとも115%、少なくとも120%、少なくとも125%、少なくとも130%、少なくとも135%、少なくとも140%、少なくとも145%、少なくとも150%、少なくとも155%、少なくとも160%、少なくとも165%、少なくとも170%、少なくとも175%、少なくとも180%、少なくとも185%、少なくとも190%、少なくとも195%、少なくとも200%、少なくとも210%、少なくとも220%、少なくとも230%、少なくとも240%、少なくとも250%、少なくとも260%、少なくとも270%、少なくとも280%、少なくとも290%、少なくとも300%、少なくとも310%、少なくとも320%、少なくとも330%、少なくとも340%、少なくとも350%、少なくとも360%、少なくとも370%、少なくとも380%、少なくとも390%、少なくとも400%、少なくとも450%、または少なくとも500%高いオンターゲット塩基編集効率を有する。

30

40

【0858】

いくつかの実施形態では、本明細書に記載のABE8塩基エディターバリエーションのいずれかは、ABE7塩基エディタ（例えばABE7.10）と比較して、少なくとも1.1倍、少なくとも1.2倍、少なくとも1.3倍、少なくとも1.4倍、少なくとも1.5倍、少なくとも1.6倍、少

50

なくとも1.7倍、少なくとも1.8倍、少なくとも1.9倍、少なくとも2.0倍、少なくとも2.1倍、少なくとも2.2倍、少なくとも2.3倍、少なくとも2.4倍、少なくとも2.5倍、少なくとも2.6倍、少なくとも2.7倍、少なくとも2.8倍、少なくとも2.9倍、少なくとも3.0倍、少なくとも3.1倍、少なくとも3.2倍、少なくとも3.3倍、少なくとも3.4倍、少なくとも3.5倍、少なくとも3.6倍、少なくとも3.7倍、少なくとも3.8倍、少なくとも3.9倍、少なくとも4.0倍、少なくとも4.1倍、少なくとも4.2倍、少なくとも4.3倍、少なくとも4.4倍、少なくとも4.5倍、少なくとも4.6倍、少なくとも4.7倍、少なくとも4.8倍、少なくとも4.9倍、または少なくとも5.0倍高いオンターゲット塩基編集効率を有する。

【0859】

本明細書に記載のABE8塩基エディターバリエーションは、プラスミド、ベクター、LNP複合体、またはmRNAを介して宿主細胞に送達され得る。いくつかの実施形態において、本明細書に記載されるABE8塩基エディターバリエーションのいずれかは、mRNAとして宿主細胞に送達される。いくつかの実施形態では、核酸ベースの送達系（例えばmRNA）により送達されるABE8塩基エディターは、編集された核酸塩基により測定されるところで少なくとも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%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも95%、または少なくとも99%のオンターゲット塩基編集効率を有する。いくつかの実施形態において、mRNAシステムによって送達されるABE8塩基エディターは、プラスミドまたはベクターシステムによって送達されるABE8塩基エディターと比較して、より高い塩基編集効率を有する。いくつかの実施形態において、本明細書に記載のABE8塩基エディターバリエーションのいずれかは、プラスミドまたはベクター系によって送達される場合と比較して、mRNA系で送達された場合に少なくとも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%、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99%、少なくとも100%、少なくとも105%、少なくとも110%、少なくとも115%、少なくとも120%、少なくとも125%、少なくとも130%、少なくとも135%、少なくとも140%、少なくとも145%、少なくとも150%、少なくとも155%、少なくとも160%、少なくとも165%、少なくとも170%、少なくとも175%、少なくとも180%、少なくとも185%、少なくとも190%、少なくとも195%、少なくとも200%、少なくとも210%、少なくとも220%、少なくとも230%、少なくとも240%、少なくとも250%、少なくとも260%、少なくとも270%、少なくとも280%、少なくとも290%、少なくとも300%高い、少なくとも310%、少なくとも320%、少なくとも330%、少なくとも340%、少なくとも350%、少なくとも360%、少なくとも370%、少なくとも380%、少なくとも390%、少なくとも400%、少なくとも450%、または少なくとも500%のオンターゲット塩基編集効率を有する。ある実施形態において、本明細書に記載のABE8塩基エディターバリエーションのいずれかは、プラスミドまたはベクター系によって送達される場合と比較して、mRNA系で送達された場合に少なくとも1.1倍、少なくとも1.2倍、少なくとも1.3倍、少なくとも1.4倍、少なくとも1.5倍、少なくとも1.6倍、少なくとも1.7倍、少なくとも1.8倍、少なくとも1.9倍、少なくとも2.0倍、少なくとも2.1倍、少なくとも2.2倍、少なくとも2.3倍、少なくとも2.4倍、少なくとも2.5倍、少なくとも2.6倍、少なくとも2.7倍、少なくとも2.8倍、少なくとも2.9倍、少なくとも3.0倍、少なくとも3.1倍、少なくとも3.2倍、少なくとも3.3倍、少なくとも3.4倍、少なくとも3.5倍、少なくとも3.6倍、少なくとも3.7倍、少なくとも3.8倍、少なくとも3.9倍、少なくとも4.0倍、少なくとも4.1倍、少なくとも4.2倍、少なくとも4.3倍、少なくとも4.4倍、少なくとも4.5倍、少なくとも4.6倍、少なくとも4.7倍、少なくとも4.8倍、少なくとも4.9倍、または少なくとも5.0倍高いオンターゲット塩基編集効率を有する。

10

20

30

40

50

【0860】

ある実施形態において、本明細書に記載のABE8塩基エディターバリエーションの1つを含む塩基エディターシステムのいずれかは、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、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、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%未満のオフターゲット編集をもたらす。

【0861】

ある実施形態において、本明細書に記載のABE8塩基エディターバリエーションのいずれかは、プラスミドまたはベクター系によって送達される場合と比較して、mRNA系で送達された場合に、より低いガイド依存（guided）オフターゲット編集効率を有する。ある実施形態において、本明細書に記載のABE8塩基エディターバリエーションのいずれかは、プラスミドまたはベクター系によって送達される場合と比較して、mRNA系で送達された場合に、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%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも95%、または少なくとも99%低いガイド依存オフターゲット編集効率を有する。ある実施形態において、本明細書に記載のABE8塩基エディターバリエーションのいずれかは、プラスミドまたはベクター系によって送達される場合と比較して、mRNA系で送達された場合に、少なくとも1.1倍、少なくとも1.2倍、少なくとも1.3倍、少なくとも1.4倍、少なくとも1.5倍、少なくとも1.6倍、少なくとも1.7倍、少なくとも1.8倍、少なくとも1.9倍、少なくとも2.0倍、少なくとも2.1倍、少なくとも2.2倍、少なくとも2.3倍、少なくとも2.4倍、少なくとも2.5倍、少なくとも2.6倍、少なくとも2.7倍、少なくとも2.8倍、少なくとも2.9倍、または少なくとも3.0倍低いガイド依存オフターゲット編集効率を有する。ある実施形態において、本明細書に記載のABE8塩基エディターバリエーションのいずれかは、プラスミドまたはベクター系によって送達される場合と比較して、mRNA系で送達された場合に、ガイド依存オフターゲット編集効率における少なくとも約2.2倍の減少を有する。

【0862】

ある実施形態において、本明細書に記載のABE8塩基エディターバリエーションのいずれかは、プラスミドまたはベクター系によって送達される場合と比較して、mRNA系で送達された場合に、より低いガイド非依存性オフターゲット編集効率を有する。ある実施形態において、本明細書に記載のABE8塩基エディターバリエーションのいずれかは、プラスミドまたはベクター系によって送達される場合と比較して、mRNA系で送達された場合に、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%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも95%、または少なくとも99%低いガイド非依存性オフターゲット編集効率を有する。ある実施形態において、本明細書に記載のABE8塩基エディターバリエーションのいずれかは、プラスミドまたはベクター系によって送達される場合と比較して、mRNA系で送達された場合に、少なくとも1.1倍、少なくとも1.2倍、少なくとも1.3倍、少なくとも1.4倍、少なくとも1.5倍、少なくとも1.6倍、少なくとも1.7倍、少なくとも1.8倍、少なくとも1.9倍、少なくとも2.0倍、少なくとも2.1倍、少なくとも2.2倍、少なくとも2.3倍、少なくとも2.4倍、少なくとも2.5倍、少なくとも2.6倍、少なくとも2.7倍、少なくとも2.8倍、少なくとも2.9倍、少なくとも3.0倍、少なくとも5.0倍、少なくとも10.0倍、少なくとも20.0倍、少なくとも50.0倍、少なくとも70.0倍、少なく

とも100.0倍、少なくとも120.0倍、少なくとも130.0倍、または少なくとも150.0倍低いガイド非依存性オフターゲット編集効率を有する。ある実施形態において、本明細書に記載のABE8塩基エディターバリエーションは、プラスミドまたはベクター系によって送達される場合と比較して、mRNA系で送達された場合に、ガイド非依存性オフターゲット編集効率（例えば目的外RNA脱アミノ化）における134.0倍の減少を有する。ある実施形態において、本明細書に記載のABE8塩基エディターバリエーションは、ゲノムに渡るガイド非依存性突然変異率を上昇させない。

【0863】

ある実施形態において、本明細書で提供される塩基エディターは、1:1超の意図される変異対インデルの比を生ずることができる。ある実施形態において、本明細書で提供される塩基エディターは、少なくとも1.5:1、少なくとも2:1、少なくとも2.5:1、少なくとも3:1、少なくとも3.5:1、少なくとも4:1、少なくとも4.5:1、少なくとも5:1、少なくとも5.5:1、少なくとも6:1、少なくとも6.5:1、少なくとも7:1、少なくとも7.5:1、少なくとも8:1、少なくとも10:1、少なくとも12:1、少なくとも15:1、少なくとも20:1、少なくとも25:1、少なくとも30:1、少なくとも40:1、少なくとも50:1、少なくとも100:1、少なくとも200:1、少なくとも300:1、少なくとも400:1、少なくとも500:1、少なくとも600:1、少なくとも700:1、少なくとも800:1、少なくとも900:1、もしくは少なくとも1000:1、またはそれ以上の、意図される変異対インデルの比を生ずることができる。

【0864】

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:eaa04774 (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。

【0865】

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塩基以上長いかまたは短い場合、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。

【0866】

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

10

20

30

40

50

とが理解されるべきである。

【0867】

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塩基エディターは、核酸の領域におけるインデルの形成を制限することができる。ある実施形態において、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。いくつかの実施形態では、本明細書で提供される塩基エディターのいずれかは、核酸の領域においてインデルの形成を1%未満、1.5%未満、2%未満、2.5%未満、3%未満、3.5%未満、4%未満、4.5%未満、5%未満、6%未満、7%未満、8%未満、9%未満、10%未満、12%未満、15%未満、または20%未満に制限することができる。核酸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、核酸（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。

【0868】

本開示のいくつかの態様は、本明細書で提供されるいずれかの塩基エディターが、有意な数の非意図的突然変異（例えば目的外のオフターゲット編集またはバイスタンダー編集）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突然変異を効率的に生成することができるという認識に基づく。ある実施形態において、意図される突然変異は、標的遺伝子における突然変異を改変または補正するように特に設計されたgRNAに結合した特異的塩基エディターによって生成される突然変異である。ある実施形態において、意図される突然変異は、HBG変異を改変または補正するように特に設計されたgRNAに結合した特異的塩基エディターによって生成される突然変異である。

【0869】

いくつかの実施形態において、本明細書で提供されるいずれかの塩基エディターは、意図された突然変異対意図されていない突然変異の比（例えば、意図した突然変異：意図しない突然変異）を1:1より大きくすることができるいくつかの実施形態において、本明細書で提供されるいずれかの塩基エディターは、意図された突然変異対意図されていない突然変異の比を、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にすることができる。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。

【0870】

【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（マルチプレックス）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。

10

【0871】

いくつかの実施形態において、複数の核酸塩基対は、1つ以上の遺伝子内にある。ある実施形態において、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。

【0872】

20

いくつかの実施形態において、編集は、少なくとも1つのタンパク質コード領域における複数の核酸塩基対の編集である。いくつかの実施形態において、編集は、少なくとも1つのタンパク質非コード領域における複数の核酸塩基対の編集である。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である。

【0873】

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。

30

40

【0874】

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複数の核酸塩基対の多重編集が可能な塩基エディターシステムは、本明細書に記載のABE8塩基エディターバリエーションのうちの1つを含む。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複数の核酸塩基対の多重編集が可能な塩基エディターシステムは、ABE7塩基エディターのうちの1つを含む。いくつかの実施形態では、本明細書に記載のABE8塩基エディターバリエーション

50

トの1つを含む、多重編集が可能な塩基エディターシステムは、ABE7塩基エディターの1つを含む多重編集が可能な塩基エディターシステムと比較して、より高い多重編集効率を有する。いくつかの実施形態では、本明細書に記載のABE8塩基エディターバリエーションの1つを含む、多重編集が可能な塩基エディターシステムは、ABE7塩基エディターの1つを含む多重編集が可能な塩基エディターシステムと比較して、少なくとも1%、少なくとも2%、少なくとも3%、少なくとも4%、少なくとも5%、少なくとも10%、少なくとも15%、少なくとも20%、少なくとも25%、少なくとも30%、少なくとも35%、少なくとも40%、少なくとも45%、少なくとも50%、少なくとも55%、少なくとも60%、少なくとも65%、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも95%、少なくとも99%、少なくとも100%、少なくとも105%、少なくとも110%、少なくとも115%、少なくとも120%、少なくとも125%、少なくとも130%、少なくとも135%、少なくとも140%、少なくとも145%、少なくとも150%、少なくとも155%、少なくとも160%、少なくとも165%、少なくとも170%、少なくとも175%、少なくとも180%、少なくとも185%、少なくとも190%、少なくとも195%、少なくとも200%、少なくとも210%、少なくとも220%、少なくとも230%、少なくとも240%、少なくとも250%、少なくとも260%、少なくとも270%、少なくとも280%、少なくとも290%、少なくとも300%高い、少なくとも310%、少なくとも320%、少なくとも330%、少なくとも340%、少なくとも350%、少なくとも360%、少なくとも370%、少なくとも380%、少なくとも390%、少なくとも400%、少なくとも450%、または少なくとも500%高い多重編集効率を有する。いくつかの実施形態では、本明細書に記載のABE8塩基エディターバリエーションの1つを含む、多重編集が可能な塩基エディターシステムは、ABE7塩基エディターの1つを含む多重編集が可能な塩基エディターシステムと比較して、少なくとも1.1倍、少なくとも1.2倍、少なくとも1.3倍、少なくとも1.4倍、少なくとも1.5倍、少なくとも1.6倍、少なくとも1.7倍、少なくとも1.8倍、少なくとも1.9倍、少なくとも2.0倍、少なくとも2.1倍、少なくとも2.2倍、少なくとも2.3倍、少なくとも2.4倍、少なくとも2.5倍、少なくとも2.6倍、少なくとも2.7倍、少なくとも2.8倍、少なくとも2.9倍、少なくとも3.0倍、少なくとも3.1倍、少なくとも3.2倍、少なくとも3.3倍、少なくとも3.4倍、少なくとも3.5倍、少なくとも4.0倍、少なくとも4.5倍、少なくとも5.0倍、少なくとも5.5倍、または少なくとも6.0倍高い多重編集効率を有する。

【0875】

〔核酸を編集する方法〕

本開示のいくつかの態様は、核酸を編集するための方法を提供する。いくつかの実施形態において、本方法は、タンパク質をコードする核酸分子の核酸塩基（例えば二本鎖DNA配列の塩基対）を編集するための方法である。いくつかの実施形態において、本方法は、a) 核酸（例えば二本鎖DNA配列）の標的領域を、塩基エディター（例えばアデノシンデアミナーゼに融合されたCas9ドメイン）およびガイド核酸（例えばgRNA）を含む複合体と接触させる、工程と、b) 前記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前記標的核酸塩基対の第一の核酸塩基を第二の核酸塩基に変換する工程と、d) n Cas9を使用して、前記標的領域の、一本を超えない数の鎖を切断する工程とを含み、ここで、第一の核酸塩基に相補的な第三の核酸塩基が、第二の核酸塩基に相補的な第四の核酸塩基によって置き換えられる。ある実施形態において、本方法は、核酸において20%未満のインデル形成をもたらす。一部の実施形態では、工程bが省略されることが理解されるべきである。いくつかの実施形態において、本方法は、19%未満、18%未満、16%未満、14%未満、12%未満、10%未満、8%未満、6%未満、4%未満、2%未満、1%未満、0.5%未満、0.2%未満、または0.1%未満のインデル形成をもたらす。いくつかの実施形態において、本方法は、第二の核酸塩基を、第四の核酸塩基に相補的な第五の核酸塩基で置き換え、それによって意図された編集塩基対（例えばG・CからA・T）を生成することをさらに含む。いくつかの実施形態では、意図された塩基対の少なくとも5%が編集される。いくつかの実施形態では、意図された塩基対の少なくとも10%、15%、20%、25%、30%、35%、40%、45%、または50%が編集される。

【0876】

ある実施形態において、標的ヌクレオチドにおける意図された生成物対意図されない生成物の比は、少なくとも2:1、少なくとも5:1、少なくとも10:1、少なくとも20:1、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60:1、少なくとも70:1、少なくとも80:1、少なくとも90:1、少なくとも100:1、もしくは少なくとも200:1、またはそれ以上である。いくつかの実施形態において、意図された突然変異対インデル形成の比は、1:1超、10:1超、50:1超、100:1超、500:1超、もしくは1000:1超、またはそれ以上である。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）が、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は反対の鎖である。一部の実施形態では、塩基エディターはdCas9ドメインを含む。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。いくつかの実施形態において、意図される編集塩基対は、PAM部位の上流にある。ある実施形態において、意図される編集塩基対は、PAM部位の1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、または20ヌクレオチド上流である。いくつかの実施形態において、意図される編集塩基対は、PAM部位の下流にある。いくつかの実施形態において、意図される編集塩基対は、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。いくつかの実施形態において、本方法は、正準（例えばNGG）PAMサイトを必要としない。一部の実施形態では、核酸塩基エディターはリンカーを含む。ある態様において、リンカーは、長さが1〜25アミノ酸である。ある態様において、リンカーは、長さが5〜20アミノ酸である。ある実施形態において、リンカーは、長さが10、11、12、13、14、15、16、17、18、19、または20アミノ酸である。一実施形態において、リンカーは、長さが32アミノ酸である。別の実施形態では、「長いリンカー（long linker）」は、長さが少なくとも約60アミノ酸である。他の実施形態において、リンカーは、長さが約3〜100アミノ酸である。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。ある実施形態において、標的ウィンドウは、1〜10ヌクレオチドを含む。いくつかの実施形態において、標的ウィンドウは、長さが1〜9、1〜8、1〜7、1〜6、1〜5、1〜4、1〜3、1〜2、または1ヌクレオチドである。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、意図される編集塩基対は、標的ウィンドウ内にある。いくつかの実施形態において、標的ウィンドウは、意図される編集塩基対を含む。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。ある実施形態では、標的ウィンドウはメチル化ウィンドウである。

【0877】

いくつかの実施形態において、本開示は、ヌクレオチド（例えばタンパク質をコードする遺伝子におけるSNP）を編集するための方法を提供する。いくつかの実施形態において、本開示は、二本鎖DNA配列の核酸塩基対を編集するための方法を提供する。いくつかの実施形態において、本方法は、a) 二本鎖DNA配列の標的領域を、塩基エディターおよびガイド核酸（例えばgRNA）を含む複合体と接触させ、ここで、標的領域が標的核酸塩基対を含む、工程と、b) 上記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前記標的核酸塩基対の第一の核酸塩基を第二の核酸塩基に変換する工程と、d) 前記標的領域の一本鎖を超えない数の鎖を切断する工程と、を含み、ここで、第一の核酸塩基に相補的な第三の核酸塩基が、第二の核酸塩基に相補的な第四の核酸塩基によって置き換えられ、第二の核酸塩基が、第四の核酸塩基に相補的な第五の核酸塩基によって置き換えられ、それによって、意図された編集塩基対を生成し、ここで意図された塩基対を生成する効率は少なくとも5%である。一部の実施形態では、工程bは省略されることが理解されるべきである。いくつかの実施形態では、意図された塩基対の少なくとも5%が編集される。いくつかの実施形態では、意図された塩基対の少なくとも10%、15%、20%、25%、30%、35%、40%、45%、または50%が編集される。いくつかの態様において、本方法は、19%

10

20

30

40

50

未満、18%未満、16%未満、14%未満、12%未満、10%未満、8%未満、6%未満、4%未満、2%未満、1%未満、0.5%未満、0.2%未満、または0.1%未満のインデル形成を引き起こす。ある実施形態において、標的ヌクレオチドにおける意図された生成物対意図されない生成物の比は、少なくとも2:1、少なくとも5:1、少なくとも10:1、少なくとも20:1、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60:1、少なくとも70:1、少なくとも80:1、少なくとも90:1、少なくとも100:1、もしくは少なくとも200:1、またはそれ以上である。いくつかの実施形態において、意図された突然変異対インデル形成の比は、1:1超、10:1超、50:1超、100:1超、500:1超、もしくは1000:1超、またはそれ以上である。いくつかの実施形態において、切断された一本鎖が、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は反対の鎖である。いくつかの実施形態において、意図される編集塩基対は、PAM部位の上流にある。ある実施形態において、意図される編集塩基対は、PAM部位の1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、または20ヌクレオチド上流である。いくつかの実施形態において、意図される編集塩基対は、PAM部位の下流にある。いくつかの実施形態において、意図される編集塩基対は、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。いくつかの実施形態において、本方法は、正準（例えばNGG）PAMサイトを必要としない。ある態様において、リンカーは、長さが1~25アミノ酸である。ある態様において、リンカーは、長さが5~20アミノ酸である。ある実施形態において、リンカーは、長さが10、11、12、13、14、15、16、17、18、19、または20アミノ酸である。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。ある実施形態において、標的ウィンドウは、1~10ヌクレオチドを含む。いくつかの実施形態において、標的ウィンドウは、長さが1~9、1~8、1~7、1~6、1~5、1~4、1~3、1~2、または1ヌクレオチドである。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、意図される編集塩基対は、標的ウィンドウ内で起こる。いくつかの実施形態において、標的ウィンドウは、意図される編集塩基対を含む。いくつかの実施形態において、核酸塩基エディターは、本明細書に提供される塩基エディターのいずれかである。

【0878】

【宿主細胞における融合タンパク質の発現】

アデノシンデアミナーゼバリエーションを含む本発明の融合タンパク質は、細菌、酵母、真菌、昆虫、植物、および動物細胞を含むがこれらに限定されない実質的にあらゆる関心対象の宿主細胞において、当業者に知られている通常の方法を用いて発現され得る。例えば、本発明のアデノシンデアミナーゼをコードするDNAは、cDNA配列に基づいてCDSの上流および下流のための適切なプライマーを設計することによってクローニングすることができる。クローン化されたDNAは、直接に、または必要に応じて制限酵素で消化した後に、または適切なリンカーおよび/または核局在化シグナルを追加した後で、塩基編集システムの1つ以上のさらなる成分をコードするDNAとライゲーションされ得る。塩基編集システムは宿主細胞内で翻訳されて複合体を形成する。

【0879】

核酸塩基修飾活性を有する一つ以上のドメインをコードする一つ以上のポリヌクレオチド（例えばアデノシンデアミナーゼ）を、napDNAbpをコードするポリヌクレオチドに作動可能に連結して、本発明の融合タンパク質をコードするポリヌクレオチドを調製することにより、融合タンパク質が生成される。いくつかの実施形態において、napDNAbpをコードするポリヌクレオチド、および核酸塩基修飾活性を有するドメインをコードするDNAは、それぞれ、結合ドメインまたはその結合パートナーをコードするDNAと融合されてもよく、または、両方のDNAが、分離インテインをコードするDNAと融合されてもよく、それによって、核酸配列認識変換モジュールおよび核酸塩基変換酵素は、宿主細胞中で翻訳されて複合体を形成する。これらの場合、所望であれば、リンカーおよび/または核局在化

10

20

30

40

50

シグナルがDNAの一方または両方の適切な位置に連結され得る。

【0880】

本明細書に記載のタンパク質ドメインをコードするDNAは、DNAを化学的に合成することによって、または合成された部分的に重複するオリゴDNA短鎖をPCR法およびギブソンアセンブリ法を利用して連結し、その全長をコードするDNAを構築することによって、得ることができる。化学合成またはPCR法もしくはGibsonアセンブリ法の組み合わせによって完全長DNAを構築する利点は、使用されるコドン、DNAが導入される宿主に応じてCDS完全長で設計できることである。異種DNAの発現では、そのDNA配列を宿主生物で高頻度で使用されるコドンに変換することにより、タンパク質発現レベルが増加することが期待される。使用する宿主におけるコドン使用頻度のデータとしては、例えば、上総DNA研究所ホームページに開示されている遺伝コード使用頻度データベース (<http://www.kazusa.or.jp/codon/index.html>) を使用することができ、または各宿主におけるコドン使用頻度を示す文献を参照してもよい。得られたデータおよび導入されるDNA配列を参照して、そのDNA配列に用いられるコドンのうち宿主での使用頻度の低いコドンを、同じアミノ酸をコードしかつ使用頻度の高いコドンに変換し得る。

10

【0881】

核酸配列認識モジュールおよび/または核酸塩基変換酵素をコードするDNAを含む発現ベクターは、例えば、適切な発現ベクター中のプロモーターの下流にそのDNAを連結することによって作製することができる。

【0882】

発現ベクターとしては、Escherichia coli由来プラスミド（例えばpBR322、pBR325、pUC12、pUC13）；Bacillus subtilis由来プラスミド（例えば、pUB110、pTP5、pC194）；酵母由来プラスミド（例えばpSH19、pSH15）；昆虫細胞発現プラスミド（例えばpFast-Bac）；動物細胞発現プラスミド（例えばpA1-11、pXT1、pRc/CMV、pRc/RSV、pcDNA1/Neo）；λファージ等のバクテリオファージ；バキュロウイルス等の昆虫ウイルスベクター（例えばBmNPV、AcNPV）；レトロウイルス、ワクシニアウイルス、アデノウイルス等の動物ウイルスベクター等が使用される。

20

【0883】

プロモーターとしては、遺伝子発現に用いる宿主に適した任意のプロモーターを用いることができる。DSBを用いた従来の方法では、毒性のために宿主細胞の生存率が著しく低下することがあるため、誘導性プロモーターを用いることにより誘導開始までに細胞数を増やすことが望ましい。しかし、本発明の核酸修飾酵素複合体を発現させることにより十分な細胞増殖も得られるので、構成的プロモーターを用いることもできるが、これに限定されるものではない。

30

【0884】

例えば、宿主が動物細胞である場合には、SRアルファプロモーター、SV40プロモーター、LTRプロモーター、CMV（サイトメガロウイルス）プロモーター、RSV（ラウス肉腫ウイルス）プロモーター、MoMuLV（モロニーマウス白血病ウイルス）LTR、HSV-TK（単純ヘルペスウイルスのチミジンキナーゼ）プロモーター等が用いられる。これらのうち、CMVプロモーター、SRアルファプロモーター等が好ましい。宿主が大腸菌である場合、trpプロモーター、lacプロモーター、recAプロモーター、ラムダP.sub.Lプロモーター、lppプロモーター、T7プロモーター等が好ましい。宿主がバチルス属である場合、SP01プロモーター、SP02プロモーター、penPプロモーター等が好ましい。宿主が酵母の場合には、Gal1/10プロモーター、PHO5プロモーター、PGKプロモーター、GAPプロモーター、ADHプロモーター等が好ましい。宿主が昆虫細胞である場合には、ポリヘドリンプロモーター、P10プロモーター等が好ましい。宿主が植物細胞である場合、CaMV35Sプロモーター、CaMV19Sプロモーター、NOSプロモーター等が好ましい。

40

【0885】

ある実施形態において、発現ベクターは、必要に応じて、エンハンサー、スプライシングシグナル、ターミネーター、ポリA付加シグナル、選択マーカー（例えば薬物耐性遺伝

50

子、栄養要求性相補遺伝子など）、複製起点などを含むことができる。

【0886】

本明細書に記載されるタンパク質ドメインをコードするRNAは、例えば、上記核酸配列認識モジュールおよび/または核酸塩基変換酵素をコードするDNAをコードするベクターをテンプレートとして使用することにより、それ自体公知のインビトロ転写系においてmRNAに転写することにより調製することができる。

【0887】

本発明の融合タンパク質は、核酸配列認識モジュールおよび/または核酸塩基変換酵素をコードするDNAを含む発現ベクターを宿主細胞に導入し、宿主細胞を培養することにより、細胞内で発現させることができる。

【0888】

本発明において有用な宿主細胞としては、細菌細胞、酵母、昆虫細胞、動物細胞などが挙げられる。

【0889】

Escherichia属には、Escherichia coli K12.cndot.DH1 (Proc. Natl. Acad. Sci. U SA, 60, 160 (1968))、Escherichia coli JM103 (Nucleic Acids Research, 9, 309 (1981))、Escherichia coli JA221 (Journal of Molecular Biology, 120, 517 (1978))、Escherichia coli HB101 (Journal of Molecular Biology, 41, 459 (1969))、Escherichia coli C600 (Genetics, 39, 440 (1954)) 等が含まれる。

【0890】

Bacillus属には、Bacillus subtilis M1114 (Gene, 24, 255 (1983))、Bacillus subtilis 207-21 (Journal of Biochemistry, 95, 87 (1984)) 等が含まれる。

【0891】

本発明の融合タンパク質の発現のために有用な酵母としては、Saccharomyces cerevisiae AH22、AH22R-、NA87-11A、DKD-5D、20B-12、Schizosaccharomyces pombe NCYC1913、NCYC2036、Pichia pastoris KM71等が挙げられる。

【0892】

融合タンパク質は、例えばAcNPVのようなウイルスベクターを用いて昆虫細胞で発現される。昆虫宿主細胞としては、ヨトウガ幼虫由来樹立株 (Spodoptera frugiperda細胞; SF細胞)、Trichoplusia niの中腸由来MG1細胞、Trichoplusia niの卵由来のHigh Five (商標)細胞、Mamestra brassicae由来細胞、Estigmena acrea由来細胞等のいずれかの細胞株が用いられる。ウイルスがBmNPVの場合には、Bombyx mori由来樹立株の細胞 (Bombyx mori N細胞; BmN細胞) 等を昆虫細胞として用いる。Sf細胞としては、例えば、Sf9細胞 (ATCC CRL1711)、Sf21細胞 (上記すべて、In Vivo, 13, 213-217 (1977)) 等が挙げられる。

【0893】

昆虫としては、例えば、Bombyx moriの幼虫、Drosophila、コオロギなどを用いて本発明の融合タンパク質が発現される (Nature, 315, 592 (1985))。

【0894】

融合タンパク質を発現するために哺乳動物細胞株が使用され得る。このような細胞株としては、サルCOS-7細胞、サルVero細胞、チャイニーズハムスター卵巣 (CHO) 細胞、dhfr遺伝子欠損CHO細胞、マウスL細胞、マウスAtT-20細胞、マウスミエローマ細胞、ラットGH3細胞、ヒトFL細胞など、ヒトその他の哺乳動物のiPS細胞、ES細胞などの多能性幹細胞、および種々の組織から調製された初代培養細胞が挙げられる。また、ゼブラフィッシュ胚、Xenopus卵母細胞などを用いることもできる。

【0895】

当業者によく知られる方法を用いて植物細胞が培養物中に維持され得る。植物細胞培養には、様々な植物 (例えば、イネ、コムギ、トウモロコシ等の穀物、トマト、キュウリ、ナス等の農作物、カーネーション、Eustoma russellianum、タバコ、Arabidopsis thaliana) から調製される、懸濁培養細胞、カルス、プロトプラスト、葉断片、根断片等が関

10

20

30

40

50

わる。

【0896】

上記の宿主細胞は全て、ハプロイド（一倍体）または倍数体（例えば二倍体、三倍体、四倍体など）であり得る。従来の変異導入法では、変異は原則として1つの相同染色体にだけ導入されてヘテロ遺伝子型を作製する。したがって、優性変異が起こらない限り、望まれる表現型は発現せず、ホモ接合性は不都合にも労力と時間を必要とする。これに対して、本発明によれば、ゲノム中の相同染色体上のどの対立遺伝子にも変異を導入することができるので、劣性変異の場合でも一世代で所望の表現型を発現させることができ、従来の方法の問題点を解決することができるので極めて有用である。

【0897】

本発明の融合タンパク質をコードする発現ベクターは、任意のトランスフェクション法を用いて宿主細胞に導入される（例えばリゾチーム法、コンピテント法、PEG法、CaCl₂共沈法、エレクトロポレーション法、マイクロインジェクション法、粒子銃法、リポフェクション法、アグロバクテリウム法）。トランスフェクション方法は、トランスフェクションされるべき宿主細胞に基づいて選択される。

【0898】

*Escherichia coli*は、例えばProc. Natl. Acad. Sci. USA, 69, 2110 (1972)、Gene, 17, 107 (1982)等に記載された方法に従って形質転換され得る。

【0899】

*Bacillus*属は、例えば、Molecular & General Genetics, 168, 111 (1979) 等に記載された方法によってベクター導入され得る。

【0900】

酵母細胞は、例えばMethods in Enzymology, 194, 182-187 (1991)、Proc. Natl. Acad. Sci. USA, 75, 1929 (1978) 等に記載された方法によってベクター導入され得る。

【0901】

昆虫細胞は、例えばBio/Technology, 6, 47-55 (1988)等に記載された方法によってベクター導入され得る。

【0902】

哺乳動物細胞は、例えば、Cell Engineering additional volume 8, New Cell Engineering Experiment Protocol, 263-267 (1995) (Shujunsha刊行)、およびVirology, 52, 456 (1973) に記載されている方法によってベクター導入され得る。

【0903】

本発明の発現ベクターを含む細胞は、宿主に応じて変わる公知の方法に従って培養される。

【0904】

例えば、*Escherichia coli*または*Bacillus*属を培養する場合、その培養のために使用される培地として液体培地が好ましい。培地は、好ましくは、形質転換体の成長に必要な炭素源、窒素源、無機物質等を含む。炭素源の例としては、グルコース、デキストリン、可溶性デンプン、スクロースなどが挙げられる。窒素源の例としては、アンモニウム塩、硝酸塩、コーンステーパーリカー、ペプトン、カゼイン、肉エキス、大豆ケーキ、ジャガイモエキス等のような無機又は有機物質が挙げられ、無機物質としては、塩化カルシウム、リン酸二水素ナトリウム、塩化マグネシウム等が挙げられる。培地は、酵母エキス、ビタミン、成長促進因子などを含有してもよい。培地のpHは好ましくは約5～約8である。

【0905】

*Escherichia coli*を培養するための培地としては、例えば、グルコース、カザミノ酸を含むM 9培地 (Journal of Experiments in Molecular Genetics, 431-433, Cold Spring Harbor Laboratory, New York 1972) が好ましい。必要な場合には、例えば3-ベータ-インドリルアクリル酸のような薬剤が、プロモーターの効率的な機能を保証するために培地に添加され得る。*Escherichia coli*は、一般に約15～約43℃で培養される。必要ならば、曝気および攪拌を行ってもよい。

10

20

30

40

50

【0906】

Bacillus属は一般に約30～約40℃で培養する。必要に応じて曝気及び攪拌が行われ得る。

【0907】

酵母の培養のための培地の例としては、パークホルダー最少培地 (Proc. Natl. Acad. Sci. USA, 77, 4505 (1980))、0.5%カザミノ酸含有SD培地 (Proc. Natl. Acad. Sci. USA, 81, 5330 (1984)) 等が挙げられる。培地のpHは好ましくは約5～約8である。培養は、一般に約20℃～約35℃で行う。必要に応じて曝気及び攪拌を行ってもよい。

【0908】

昆虫細胞や昆虫を培養するための培地としては、例えば、不活化10%ウシ血清等の添加剤を適宜含有するグレース昆虫培地 (Nature, 195, 788 (1962)) 等が用いられる。培地のpHは好ましくは約6.2～約6.4である。培養は一般に約27℃で行われる。必要に応じて曝気および攪拌を行ってもよい。

10

【0909】

動物細胞を培養するための培地としては、例えば、約5～約20%のウシ胎児血清を含有する最小必須培地 (MEM) (Science, 122, 501 (1952))、Dulbecco改変イーグル培地 (DMEM) (Virology, 8, 396 (1959))、RPMI 1640培地 (The Journal of the American Medical Association, 199, 519 (1967))、199培地 (Proceeding of the Society for the Biological Medicine, 73, 1 (1950)) 等が使用される。培地のpHは好ましくは約6～約8である。培養は、一般に約30℃～約40℃で行われる。必要に応じて曝気および攪拌を行ってもよい。

20

【0910】

植物細胞を培養するための培地としては、例えば、MS培地、LS培地、B5培地等が用いられる。培地のpHは好ましくは約5～約8である。培養は、一般に約20℃～約30℃で行われる。必要に応じて曝気及び攪拌を行ってもよい。

【0911】

動物細胞、昆虫細胞、植物細胞等の高等真核細胞を宿主細胞として用いる場合には、本発明の塩基編集系 (例えばアデノシンデアミナーゼバリエーションを含むもの) をコードするDNAを、誘導性プロモーター (例えばメタロチオネインプロモーター (重金属イオンによって誘導される)、熱ショックタンパク質プロモーター (熱ショックによって誘導される)、Tet-ON/Tet-OFF系プロモーター (テトラサイクリンまたはその誘導体の追加または除去によって誘導される)、ステロイド応答性プロモーター (ステロイドホルモンまたはその誘導体によって誘導される) 等) の制御下において宿主細胞に導入し、誘導物質が適切な段階で培地に添加されて (または培地から除去されて) 核酸修飾酵素複合体の発現を誘導し、一定期間培養を行って塩基編集および標的遺伝子への変異の導入を実行し、塩基編集系の一過性の発現を実現することができる。

30

【0912】

大腸菌などの原核細胞は誘導性プロモーターを利用することができる。誘導性プロモーターの例としては、lacプロモーター (IPTGによって誘導される)、cspAプロモーター (寒冷ショックによって誘導される)、araBADプロモーター (アラビノースによって誘導される) などが挙げられるが、これらに限定されない。

40

【0913】

あるいは、動物細胞、昆虫細胞、植物細胞等の高等真核細胞を宿主細胞として用いる場合には、上述の誘導性プロモーターをベクター除去機序として利用することもできる。すなわち、宿主細胞で機能する複製起点をベクターに搭載し、複製に必要なタンパク質 (例えば動物細胞についてはSV40 onおよびラージT抗原、oriP、およびEBNA-1等) をコードする核酸の発現を上記誘導性プロモーターにより制御する。その結果、誘導物質の存在下ではベクターは自律的に複製可能であるが、誘導物質を除去すると、自律的な複製はできず、ベクターは細胞分裂とともに自然に脱落する (Tet-OFF系ベクターではテトラサイクリンやドキシサイクリンの添加により自律的な複製ができない)。

50

【0914】

[塩基エディターを使用する方法]

疾患関連の遺伝子およびアリの点突然変異の補正は、治療および基礎研究への応用を伴う遺伝子補正のための新しい戦略を提供する。

【0915】

本開示は、本明細書で提供される塩基エディターシステムによって修正することができる点突然変異に関連するかまたはそれによって引き起こされる疾患と診断された対象の、治療のための方法を提供する。例えば、いくつかの実施形態において、そのような疾患（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、疾患関連遺伝子の点突然変異を修正する有効量の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）を投与することを含む方法が提供される。本開示は、デアミナーゼ媒介遺伝子編集によって修正することができる点突然変異に関連するかまたはそれによって引き起こされる疾患の治療のための方法を提供する。本明細書で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適切な疾患は、本開示に基づいて当業者には明らかであろう。本明細書で提供されるのは、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である。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Cas12ドメインを含むもの）の活性が、点突然変異の修正をもたらす。いくつかの実施形態では、標的DNA配列は、疾患または障害に関連するG→A点突然変異を含み、変異A塩基の脱アミノ化が、疾患または障害に関連しない配列をもたらす。いくつかの実施形態において、標的DNA配列は、疾患または障害に関連するT→C点突然変異を含み、変異C塩基の脱アミノ化が、疾患または障害に関連しない配列をもたらす。

【0916】

いくつかの実施形態では、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。いくつかの実施形態では、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。

【0917】

いくつかの実施形態では、本明細書で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を脱アミノ化することができる。本開示の他の態様は、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）と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Cas12）を含む融合タンパク質とを提供する。例えば、アデノシンはイノシン残基に変換されることができ、これは通常、シトシン残基と塩基対を形成する。そのような融合タンパク質は、とりわけ、核酸配列の標的化された編集に有用である。このような融合タンパク質は、インビトロでのDNAの標的化編集のために、例えば、変異細胞または変異動物の生成のために、標的突然変異の導入のために、例えば、エクスピボでの細胞における（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ける）遺伝的欠陥の補正のために、および、インビボでの標的化変異の導入、例えば、遺伝的欠陥の補正のために使用され得、または、GからAもしくはTからCへの変異における疾患関連遺伝子への不活性化突然変異の導入が、本明細書で提供される核酸塩基エディターを使用して治療され得る。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、システムなどを提供する。

【0918】

[意図する変異の生成]

いくつかの実施形態では、本明細書で提供される方法の目的は、遺伝子編集を介して機能不全の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、機能不全の遺伝子の機能は、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では、本明細書で提供される方法を使用して、遺伝子産物の正常な機能を破壊することができる。本明細書で提供される核酸塩基編集タンパク質は、例えば、ヒト細胞培養において疾患関連突然変異を修正することによって、インビトロでの遺伝子編集ベースのヒト治療について検証することができる。本明細書で提供される核酸塩基編集タンパク質、例えば、napDNAbpドメイン（例えば、Cas12）と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）を含む融合タンパク質は、AからGまたはCからTへのあらゆる単一点変異を修正するために使用され得ることが当業者によって理解されるであろう。前者の場合は変異体AからIへの脱アミノ化、後者の場合は変異体Tと塩基対を形成するAの脱アミノ化、およびそれに続く1ラウンドの複製により、変異が修正される。

10

【0919】

いくつかの実施形態では、本開示は、有意な数の意図しない突然変異（例えば意図しない点突然変異）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、点突然変異などの意図する突然変異を効率的に生成す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、意図される突然変異は、その意図される突然変異を生成するように特別に設計されたガイドポリヌクレオチド（例えばgRNA）に結合した特異的塩基エディター（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）によって生成される突然変異である。いくつかの実施形態において、意図される突然変異は、疾患または障害に関連する突然変異である。いくつかの実施形態において、意図される突然変異は、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。いくつかの実施形態において、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）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。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シトシン（C）からチミン（T）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内の停止コドン、例えば、中途（premature）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、終止コドンを除去する突然変異である。

20

30

【0920】

いくつかの実施形態において、本明細書で提供されるいずれかの塩基エディターは、意図された突然変異対意図されていない突然変異の比（例えば、意図した点突然変異：意図しない点突然変異）を1:1より大きくすることができるいくつかの実施形態において、本明細書で提供されるいずれかの塩基エディターは、意図された突然変異対意図されていない突然変異（例えば、意図した点突然変異：意図しない点突然変異）の比を、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、もしくは少なくとも1000:1またはそれ以上にすることができる。

40

【0921】

塩基エディターの効率の詳細は、国際PCT出願番号PCT/2017/045381（WO 2018/027078）およびPCT/US2016/058344（WO 2017/070632）に記載されており、それぞれが参照によりその全体が本明細書に組み込まれる。Komor, A.C., et al., “Progra

50

mmable 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）。

【0922】

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を使用した1つ以上の遺伝子における複数の核酸塩基対の編集が、少なくとも1つの意図された突然変異の形成をもたらす。いくつかの実施形態において、前記少なくとも1つの意図された突然変異の形成が、疾患を引き起こす突然変異の正確な修正をもたらす。マルチプレックス編集は、本明細書で提供される任意の方法または方法の組み合わせを使用して達成できることが理解されるべきである。

【0923】

〔送達システム〕

核酸塩基エディターおよびgRNAの、核酸に基づく送達

本開示による塩基編集システムをコードする核酸は、公知の方法によって、または本明細書に記載するように、対象に投与されるか、またはインビトロもしくはインビボで細胞に送達され得る。一実施形態において、核酸塩基エディターは、例えばベクター（例えばウイルスベクター又は非ウイルスベクター）、ベクターに基づかない方法（例えば裸のDNA、DNA複合体、mRNA、脂質ナノ粒子を用いて）、またはそれらの組み合わせによって送達され得る。

【0924】

核酸塩基エディターをコードする核酸は、例えばトランスフェクションまたはエレクトロポレーションによって、裸のDNAまたはRNA、例えばmRNAとして細胞（例えば、造血細胞もしくはその前駆細胞、造血幹細胞、および/または人工多能性幹細胞）に直接送達されるか、または標的細胞による取り込みを促進させる分子（例えばN-アセチルガラクトサミン）にコンジュゲート化され得る。本明細書に記載されるベクターのような核酸ベクターも使用され得る。

【0925】

核酸ベクターは、本明細書中に記載される融合タンパク質のドメインをコードする1以上の配列を含むことができる。ベクターはまた、タンパク質をコードする配列に付随する（例えば、挿入されているか、融合されている）シグナルペプチド（例えば、核局在化、核小体局在化、またはミトコンドリア局在化のためのもの）をコードする配列も含むことができる。一例として、核酸ベクターは、一つ以上の核局在化配列（例えばSV40からの核局在化配列）を含むCas9コード配列と、アデノシンデアミナーゼバリエーション（ABE8）を含むことができる。

【0926】

核酸ベクターはまた、任意の適切な数の調節/制御エレメント、例えばプロモーター、エンハンサー、イントロン、ポリアデニル化シグナル、Kozakコンセンサス配列、または内部リボソームエンタリー部位（IRES）を含むことができる。これらのエレメントは当技術分野でよく知られている。造血細胞については、適切なプロモーターは、IFNベータまたはCD45を含み得る。

【0927】

本開示による核酸ベクターは、組換えウイルスベクターを含む。例示的なウイルスベクターは本明細書中に記載される。当技術分野で知られる他のウイルスベクターも使用することができる。さらに、ウイルス粒子を用いて、核酸および/またはペプチドの形態の塩基編集システム成分を送達することができる。例えば、「空の」ウイルス粒子は、任意の適切なカーゴを含有するようにアセンブルされ得る。ウイルスベクターおよびウイルス粒子

はまた、標的化リガンドを組み込んで標的組織特異性を変化させるように操作することができる。

【0928】

ウイルス性ベクターに加えて、本開示によるゲノム編集システムをコードする核酸を送達するために非ウイルス性ベクターを使用することもできる。非ウイルス性核酸ベクターの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のものである。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）ナノ粒子が、本開示の特定の実施形態において送達ビヒクルとしての使用のために適切となり得る。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10（下記）に示す。

【0929】

10

20

30

40

50

【表 10】

遺伝子移入のために使用される脂質			
Lipid	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	10
1,2-Dioleoyloxy-3-trimethylammonium-propane	DOTAP	Cationic	
Diioctadecylamidoglycylspermine	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(sperminocarboxamido-ethyl)-N,N-dimethyl-1-propanaminium trifluoroacetate	DOSPA	Cationic	20
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)-carbonyl]cholesterol	DC-Chol	Cationic	
Bis-guanidium-tren-cholesterol	BGTC	Cationic	
1,3-Dioleoyl-2-(6-carboxy-spermyl)-propylamide	DOSPER	Cationic	
Dimethyloctadecylammonium bromide	DDAB	Cationic	30
Diioctadecylamidoglycylspermidin	DSL	Cationic	
rac-[(2,3-Dioctadecyloxypropyl)(2-hydroxyethyl)]-dimethylammonium chloride	CLIP-1	Cationic	
rac-[2(2,3-Dihexadecyloxypropyl)-oxymethyloxy)ethyl]trimethylammonium bromide	CLIP-6	Cationic	
Ethylidimyristoylphosphatidylcholine	EDMPC	Cationic	
1,2-Distearyloxy-N,N-dimethyl-3-aminopropane	DSDMA	Cationic	
1,2-Dimyristoyl-trimethylammonium propane	DMTAP	Cationic	40

遺伝子移入のために使用される脂質

Lipid	略語	特徴	
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	10
N-t-Butyl-N0-tetradecyl-3-tetradecylaminopropionamidinium	diC14-amidinium	Cationic	
Octadecenyl-oxy[ethyl-2-heptadecenyl-3-hydroxyethyl]imidazolium chloride	DOTIM	Cationic	
N1 -Cholesteryl-oxy-carbonyl-3,7-diazanonane-1,9-diamine	CDAN	Cationic	
2-(3-[Bis(3-amino-propyl)-amino]propylamino)-N-ditetradecylcarbamoyl-methyl-ethyl-acetamide	RPR209120	Cationic	
1,2-dilinoleyl-oxy-3-dimethylaminopropane	DLinDMA	Cationic	
2,2-dilinoleyl-4-dimethylaminoethyl-[1,3]-dioxolane	DLin-KC2-DMA	Cationic	20
dilinoleyl-methyl-4-dimethylaminobutyrate	DLin-MC3-DMA	Cationic	

【 0 9 3 0 】

表 1 1 は、遺伝子移入および／またはナノ粒子製剤における使用のための例示的なポリマーを列記する。

30

40

50

【表 1 1】

遺伝子移入のために使用されるポリマー		
ポリマー	略語	
Poly(ethylene)glycol	PEG	
Polyethylenimine	PEI	
Dithiobis (succinimidylpropionate)	DSP	
Dimethyl-3,3'-dithiobispropionimide	DTBP	10
Poly(ethylene imine)biscarbamate	PEIC	
Poly(L-lysine)	PLL	
Histidine modified PLL		
Poly(N-vinylpyrrolidone)	PVP	
Poly(propylenimine)	PPI	
Poly(amidoamine)	PAMAM	
Poly(amidoethylenimine)	SS-PAEI	
Triethylenetetramine	TETA	
Poly(β -aminoester)		20
Poly(4-hydroxy-L-proline ester)	PHP	
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	
Poly(phosphoramidate)s	PPA	
Poly(N-2-hydroxypropylmethacrylamide)	pHPMA	30
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	40

【0 9 3 1】

表 1 2 は、本明細書に記載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方法を要約する。

【表 1 2】

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
生物学的非ウイルス性送達媒体	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
	弱毒化細菌	YES	一過性	NO	核酸
	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

【0 9 3 2】

別の局面において、例えばCas9もしく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。

【0933】

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。

【0934】

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などのプロモーターが挙げられ得る。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。骨芽細胞については、適切なプロモーターにはOG-2が含まれ得る。

【0935】

いくつかの実施形態において、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小ささである。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。

【0936】

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。

【0937】

[ウイルスベクター]

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得る。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、ウイルスベクターに含まれる核酸上にコードされ得る。いくつかの実施形態において、塩基エディターシステムの1つ以上の構成要素は、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は、単一のウイルスベクター上にコードされ得る。他の実施形態において、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。いずれの場合も、塩基エディターおよびガイド核酸はそれぞれ、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。ウイルスベクター上にコードされる成分の組合せは、選択されたウイルスベクターのカーゴサイズ制約により決定され得る。

10

20

30

40

50

【0938】

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。

10

【0939】

ウイルスベクターは、レンチウイルス（例えばHIVおよびHIVベースのベクター）、アデノウイルス（例えばAD100）、レトロウイルス（例えばマロニーマウス白血病ウイルス、MML-V）、ヘルペスウイルス（例えばHSV-2）、およびアデノ随伴ウイルス（例えばAAV）、またはその他のプラスミドもしくはウイルスベクタータイプを含み、特に、例えば米国特許第8,454,972号（製剤、アデノウイルスの用量）、米国特許第8,404,658号（製剤、AAVの用量）および米国特許第5,846,946号（製剤、DNAプラスミドの用量）、ならびにレンチウイルス、AAVおよびアデノウイルスを含む臨床試験に関する臨床試験および刊行物からの製剤および用量を用いるものが挙げられる。例えば、AAVについては、投与の経路、製剤および用量は、米国特許第8,454,972号におけるものおよびAAVを含む臨床試験におけるものとして行うことができる。アデノウイルスについては、投与の経路、製剤および用量は、米国特許第8,404,658号におけるものおよびアデノウイルスを含む臨床試験におけるものとして行うことができる。プラスミド送達については、投与の経路、製剤および用量は、米国特許第5,846,946号におけるものおよびプラスミドを含む臨床研究におけるものとして行うことができる。用量は平均的70 kgの個体（例えば成人男性）に基づくか外挿され得、体重および種の異なる患者、被験者、哺乳動物について調整することができる。投与の頻度は、医療従事者または獣医（例えば医師、獣医）の領域の範囲内であり、年齢、性別、全般的な健康状態、患者または対象の他の状態および対処される特定の状態または症状を含む通常の因子に依存する。ウイルスベクターを目的の組織に注入され得る。細胞型特異的塩基編集のためには、塩基エディターおよび任意のガイド核酸の発現は、細胞型特異的プロモーターによって駆動され得る。

20

30

【0940】

レトロウイルスの指向性 (tropism) は、外来のエンベローブ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); Sommerfelt 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)。

40

【0941】

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

50

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し得る。ある態様において、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。いくつかの実施形態では、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。

【0942】

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、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。アデノ随伴ウイルス（「AAV」）ベクターもまた、例えば核酸およびペプチドのインビトロ生成において、ならびにインビボおよびエクスビボの遺伝子治療手順のために、標的核酸で細胞を形質導入することに使用され得る（例えば、West et al., *Virology* 160:38-47 (1987); 米国特許第4,797,368号; WO 93/24641; Kotin, *Human Gene Therapy* 5:793-801 (1994); Muzyczka, *J. Clin. Invest.* 94:1351 (1994)を参照）。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。

【0943】

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。

【0944】

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。

【0945】

ウイルスベクターは、アプリケーションに基づいて選択することができる。例えば、インビボ遺伝子送達については、AAVが他のウイルスベクターよりも有利であり得る。ある実施形態において、AAVは低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。ある実施形態において、AAVは、宿主ゲノムに組み込まれないため、挿入変異誘発を引き起こす可能性の低さを可能にする。アデノウイルスは強い免疫応答を誘発するため、ワクチンとしてよく用いられる。ウイルスベクターのパッケージング能力が、ベクターにパッケージングすることができる塩基エディターのサイズを制限し得る。

【0946】

AAVは約4.5 Kbまたは二つの145塩基逆末端反復配列 (ITR) を含め4.75 Kbのパッケ

ージング能を有する。これは、開示された塩基エディター、ならびにプロモーターおよび転写ターミネーターが、単一のウイルスベクター中に収まり得ることを意味する。4.5 Kb または4.75 Kbより大きい構築物は、ウイルス産生の著しい低下をもたらし得る。例えば、SpCas9は非常に大きく、遺伝子自体が4.1 Kbを超えているため、AAVに詰め込むのは困難である。したがって、本開示の実施形態は、従来の塩基エディターよりも短い、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。開示された塩基エディターは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。いくつかの実施形態において、開示された塩基エディターの長さは4.5 kb以下である。

10

【0947】

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)に見出される。

【0948】

20

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。

【0949】

レンチウイルスは以下のように調製できる。pCasES10 (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 (pCasES10) 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する。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に変更する。これらの方法は細胞培養中に血清を用いるが、無血清法が好ましい。

30

【0950】

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°Cで一晩再懸濁する。次いでそれらを等分し、直ちに-80°Cで凍結する。

40

【0951】

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 (登録商標) 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。

【0952】

当該システムのあらゆるRNA、例えばガイドRNAまたは塩基エディターをコードするm

50

RNAを、RNAの形態で細胞に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、インビトロ転写を使用し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使用され得る。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。

【0953】

発現を増強し、毒性の可能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、1以上の修飾ヌクレオチドを含むように修飾され得る（例えば擬似Uまたは5-メチル-Cを用いて）。

【0954】

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る。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。

【0955】

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。

【0956】

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。

【0957】

[インテイン]

ある実施形態において、ヌクレアーゼ（例えばCas9）の一部または断片が、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

10

20

30

40

50

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。

【0958】

インテイン (intervening protein) は、多種多様な生物に見出される自己プロセッシングドメインであり、タンパク質スプライシングとして知られるプロセスを行うものである。タンパク質スプライシングは、ペプチド結合の切断と形成の両方からなる多段階の生化学的反応である。タンパク質スプライシングの内因性基質は、インテインを含む生物に見出されるタンパク質であるが、インテインはまた、実質的にあらゆるポリペプチド骨格を化学的に操作するために使用することもできる。

【0959】

タンパク質スプライシングでは、インテインは、二つのペプチド結合を切断することによって一前駆体ポリペプチドから自身を切り出し、それによって、隣接するエクステイン (外部タンパク質) 配列を、新しいペプチド結合の形成を介して連結する。この転位は翻訳後に起こる (翻訳と同時に起こる可能性もある)。インテイン媒介性タンパク質スプライシングは自発的に起こり、インテインドメインの折りたたみだけを必要とする。

【0960】

インテインの約5%が分割インテインであり、これらはN-インテインとC-インテインという二つの別々のポリペプチドとして転写され翻訳され、その各々が一つのエクステインに融合している。翻訳の際に、インテイン断片は、自発的にかつ非共有結合的に標準インテイン構造へとアSEMBルし、トランスにタンパク質スプライシングを行う。タンパク質スプライシングの機序には一連のアシル転移反応が関わっており、これが、インテイン-エクステイン接合部における2つのペプチド結合の切断と、N-エクステインとC-エクステインの間の新しいペプチド結合の形成とをもたらす。このプロセスは、N-エクステインとインテインのN-末端とを繋げるペプチド結合の活性化によって開始される。事実上すべてのインテインは、N末端にシステインまたはセリンを有し、これがN-エクステインのC末端残基のカルボニル炭素を攻撃する。このNからO/Sへのアシル基の移動は、保存されたトレオニンとヒスチジン (TXXHモチーフと呼ばれる)、および一般に見出されるアスパラギン酸によって促進され、直鎖状 (チオ) エステル中間体の形成をもたらす。次に、この中間体は、システイン、セリン、またはトレオニンであるC-エクステインの最初の残基 (+1) の求核攻撃によってトランス-(チオ) エステル化される。生成した分枝 (チオ) エステル中間体は、インテインの高度に保存されたC末端アスパラギンの環化というユニークな変換によって、解消される。この過程は、ヒスチジン (高度に保存されたHNFモチーフに見出されるもの) と最後から2番目のヒスチジンによって促進され、アスパラギン酸も関与し得る。このスクシンイミド生成反応は、反応複合体からインテインを切除し、非ペプチド結合を介して結合されたエクステインを残す。この構造は、インテイン非依存的態様で迅速に転位し、安定なペプチド結合になる。

【0961】

いくつかの実施形態では、塩基エディター (例えばABE、CBE) のN末端断片が分割インテイン-Nに融合され、C末端断片が分割インテイン-Cに融合される。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。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。

【0962】

いくつかの実施形態において、ABEは、SpCas9の選択された領域内のAla、Ser、Thr、またはCys残基においてN末端断片およびC末端断片に分割された。これらの領域は、Cas9結晶構造解析により同定されたループ領域に対応する。各断片のN-末端をインテイン-Nに融合させ、各断片のC-末端を、アミノ酸位置S303、T310、T313、S355、A456、S460、A463、T466、S469、T472、T474、C574、S577、A589、およびS590（下記配列で太字の大文字で示されている）においてインテインCに融合させる。

1 mdkkysigld igtnsvgwav itdeykvpsk kfkvlgntr hsikknliga llfdsgetae

61 atrlkrtarr rytrrknic ylqeifsnem akvddsffhr leesflveed kkherhpifg
 121 nivdevayhe kyptiyhlrk klvdstdkad lrliylalah mikfrghfli egdlndpnsd
 181 vdklfiqlvq tynqlfeenp inasgvdaka ilsarlsksr rlenliaqlp gekknlgfng
 241 lialslgltp nfksnfdlae daklqlskdt yddldnlla qigdqyadlf laaknlsdai
 301 llSdilrvnT eiTkapsas mikrydehhq dltllkalvr qqlpekykei ffdqSkngya
 361 gyidggasqe efykfikpil ekmdgteell vclnredllr kqrtfdngsi phqihlgelg
 421 ailrrqedfy pflkdnreki ekiltfripy yvgplArgnS rfAwmTrkSe eTiTpwnfee
 481 vvdkgasaqs fiermtnfkd nlpnekvlpk hsllyeyftv yneltkvkv tegmrkpafl
 541 sgeqkkaivd llfktnrkvt vkqlkedyfk kieCfdSvei sgvedrfnAS lgtyhdllki
 601 ikdkdfldne enedilediv ltltlfedre mieerlktya hlfdkvmkq lkrrrytgwg
 661 rlsrklingi rdkqsgktil dflksdgfan rnfmqlihdd sltfkediql aqvsgqgdsi
 721 hehianlags paikkgilqt vkvvdclkv mgrhkpeniv iemarenqtt kqgqknsrer
 781 mkrieegike lgsqilkehp ventqlqnek lylyylqngr dmyvdqeldi nrlsdydvdh
 841 ivpqslkdd sidnkvtrs dknrgksdnv pseevvkkmk nywrqlnak litqrkfdnl
 901 tkaergglse ldkagfikrq lvetrqitkh vaqildsrnm tkydendkli revkvitlks
 961 klvsdfrkdf qfykvreinn yhhahdayln avvgtalikk ypklesefyv gdykvydvrvk
 1021 miaksegeig katakyffys nimnffktei tlangeirkr plietngetg eivwdkgrdf
 1081 atvrkvlsmv qvnivkktev qtggfskesi lpkrnsdkli arkkdwdpkk yggfdsptva
 1141 ysvlvvakve kgkskklksv kllgitime rssfeknpid fleakgykev kkdliiklpk
 1201 yslfelengr krmlasagel qkgnelalps kyvnflylas hyeklkgspe dneqkqlfve
 1261 qhkhyldcii eqisefskrv iladanldkv lsaynkhrdk pireqaenii hlftltnlga
 1321 paafkyfddt idrkrytstk evldatlihq sitglyetri dlsqllggd

【0963】

〔核酸塩基エディターを用いた変異の標的化〕

変異を標的とする核酸塩基エディターの適合性は、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、対象の単一の細胞が、レポーター（例えばGFP）をコードする少量のベクターと一緒に、塩基編集システムで形質導入される。これらの細胞は、293T、K562またはU2OSなどの不死化ヒト細胞株を含め、当技術分野で知られる任意の細胞株であり得る。あるいは、（例えばヒトの）一次細胞を使用してもよい。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。

【0964】

送達は、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。

【0965】

核酸塩基エディターの活性は、本明細書に記載のように、すなわち細胞のゲノムの配列決定を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）で使用するのために）。

【0966】

最初の試験で最大レベルの標的特異的変化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。

【0967】

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、細胞のゲノム内の目的の変異を標的化するために使用されるガイドRNAと共に、細胞（例えば造血細胞もしくはその前駆細胞、造血幹細胞、および/または人工多能性幹細胞）に送達され、それによって変異を変化させる。いくつかの実施形態において、塩基エディターは、ガイドRNAによって標的指向化され、対象遺伝子の配列に1つ以上の編集を導入する。

【0968】

一実施形態では、核酸塩基エディターを用いて、スプライス部位、エンハンサー、および転写調節エレメントを含むがこれらに限定されない調節配列を標的化する。次いで、その調節エレメントにより制御される遺伝子の発現に対する該変化の影響を、当該技術分野で知られる任意の方法を用いてアッセイする。

【0969】

さらに他の実施形態では、本発明の核酸塩基エディターを用いて、生物のゲノム内の目的のポリヌクレオチドを標的化する。一実施形態において、本発明の核酸塩基エディターは、細胞のゲノム内の様々な配列をタイル状にカバーするために使用されるガイドRNAのライブラリと共に細胞に送達され、それによってゲノム全体に渡って配列を系統的に改変する。

【0970】

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、所望の細胞型における発現のためにコドン最適化され、それは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である。

【0971】

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

10

20

30

40

50

ンに対応する。

【0972】

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。場合によっては、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。

10

【0973】

【マルチエフェクター核酸塩基エディターのアプリケーション】

マルチエフェクター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とし、タンパク質発現を改変する変化を作り出すことができる。一実施形態では、マルチエフェクター核酸塩基エディターは、スプライス部位、エンハンサー、および転写調節エレメントを含むがこれらに限定されない非コード配列または調節配列を改変するために使用される。次いで、当該技術分野で公知の任意の方法を用いて、その調節エレメントによって制御される遺伝子の発現に対する当該改変の影響をアッセイする。特定の実施形態において、マルチエフェクター核酸塩基エディターは、調節配列を実質的に変化させて、それによって、遺伝子発現を調節するその能力を消失させることができる。有利なことに、これは、他のRNAプログラミング可能ヌクレアーゼとは対照的に、ゲノム標的配列に二本鎖切断を生じることなく行うことができる。

20

30

【0974】

マルチエフェクター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とし、タンパク質活性を改変する変化を作り出すことができる。例えば、突然変異誘発の文脈において、マルチエフェクター核酸塩基エディターは、エラープローンPCRおよび他のポリメラーゼベースの方法と比べて多くの利点を有する。本発明のマルチエフェクター核酸塩基エディターは、標的領域内の複数の塩基に変化を生じるので、そのような変異は、エラープローンPCRによって導入される変異と比較して、タンパク質レベルで発現される可能性がより高い。エラープローンPCRによって導入される変異は、コドン中の単一ヌクレオチド変化が依然として同じアミノ酸をコードし得ること（コドンの縮重）を考慮すると、タンパク質レベルで発現される可能性がより低い。ポリヌクレオチド全体にわたってランダムな変化を誘導するエラープローンPCRとは異なり、本発明のマルチエフェクター核酸塩基エディターは、対象タンパク質の小さな領域または限定された領域内の特定のアミノ酸を標的化するために使用できる。

40

【0975】

他の実施形態において、本発明のマルチエフェクター核酸塩基エディターは、生物のゲノム内の目的のポリヌクレオチドを標的化するために使用される。一実施形態において、生物は、微生物叢の細菌（例えばBacteroidetes、Verrucomicrobia、Firmicutes；ガンマプロテオバクテリア、アルファプロテオバクテリア、Bacteroidetes、Clostridia、Erysipelotrichia、Bacilli；Enterobacteriales、Bacteriales、Verrucomicrobia

50

les、Clostridiales、Erysipelotrichales、Lactobacillales ; Enterobacteriaceae、Bacteroidaceae、Erysipelotrichaceae、Prevotellaceae、Coriobacteriaceae、およびAlcaligenaceae、Escherichia、Bacteroides、Alistipes、Akkermansia、Clostridium、Lactobacillus) である。別の実施形態では、生物は農業上重要な動物（例えば牛、羊、山羊、馬、鶏、七面鳥）または植物（例えば大豆、小麦、トウモロコシ、米、タバコ、リンゴ、ブドウ、モモ、プラム、チェリー）である。1つの実施形態において、本発明のマルチエフェクター核酸塩基エディターは、細胞のゲノム内の種々の配列を標的化するために使用されるガイドRNAのライブラリーと共に細胞に送達され、それによってゲノム全体の配列を系統的に変化させる。1つの実施形態において、本発明のマルチエフェクター核酸塩基エディターは、細胞のゲノム内の種々の配列をタイル状にカバーするために使用されるガイドRNAのライブラリーと共に細胞に送達され、それによってゲノム全体の配列を系統的に変化させる。

【0976】

突然変異を、種々のタンパク質のいずれかにおいて作製して、構造-機能分析を促進させ、またはタンパク質の内因性活性を変化させ得る。突然変異は、例えば、酵素（例えばキナーゼ、ホスファターゼ、カルボキシラーゼ、ホスホジエステラーゼ）または酵素基質、受容体またはそのリガンド、ならびに抗体およびその抗原において作製され得る。一実施形態では、マルチエフェクター核酸塩基エディターは、酵素の活性部位、受容体のリガンド結合部位、または抗体の相補性決定領域 (CDR) をコードする核酸分子を標的とする。酵素の場合、活性部位に変異を導入すると、酵素活性が増加、減少、または消失し得る。酵素に対する突然変異の影響は、当技術分野で知られるおよび/または当業者には明らかである多くのアッセイのいずれかを含む、酵素活性アッセイにおいて特徴付けされる。受容体の場合、リガンド結合部位に生成された変異は、リガンドに対する受容体親和性を増加、減少、または消失させ得る。このような突然変異の効果は、当技術分野で知られるおよび/または当業者には明らかである多くのアッセイのいずれかを含む、受容体/リガンド結合アッセイにおいてアッセイされる。

【0977】

[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。いくつかの態様において、医薬組成物は、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、医薬組成物は、さらなる薬剤（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。

【0978】

適切な薬学的に許容される担体は、一般に、対象に医薬組成物を投与することを助け、医薬組成物を送達可能な調製物に加工することを助け、または投与前に医薬組成物を保存することを助ける、不活性物質を含む。薬学的に許容される担体は、製剤の形態、質感、粘度、pH、薬物動態、溶解性を安定化、最適化、または他の態様で変化させることができる。

【0979】

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非限定的な例は、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、ペニバナ油、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。緩衝剤、湿潤剤、乳化剤、希釈剤、封入剤、皮膚浸透増強剤、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。例えば、担体は、限定されるものではないが、生理食塩水、緩衝生理食塩水、デキストロース、アルギニン、スクロース、水、グリセロール、エタノール、ソルビトール、デキストラン、カルボキシメチルセルロースナトリウム、およびそれらの組み合わせを含むことができる。

10

【0980】

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緩衝化合物は、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。

【0981】

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)をレシピエント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。

20

30

【0982】

いくつかの実施形態において、薬学的組成物は、対象への送達のために、例えば遺伝子編集のために製剤化される。いくつかの実施形態において、本明細書で企図される薬学的組成物の投与は、限定されるものではないが、注入、輸血、または非経口的を含む従来の技術を用いて実施され得る。いくつかの実施形態において、非経口投与は、血管内、静脈内、筋肉内、動脈内、髄腔内、腫瘍内、皮内、腹腔内、経気管、皮下、角質下、関節内、莖膜内、くも膜下および胸骨内への注入または注射を含む。いくつかの実施形態において、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。

40

【0983】

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。

【0984】

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

50

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。

10

【0985】

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。

20

【0986】

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）。

30

【0987】

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。

40

【0988】

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することができる。

50

【0989】

別の態様では、上記疾患の治療に有用な材料を含む製品が含まれる。いくつかの実施態様において、製品は、容器及びラベルを含む。適切な容器は、例えばボトル、バイアル、シリンジ、および試験管を含む。容器は、ガラスまたはプラスチックなどの様々な材料から形成することができる。いくつかの実施形態において、容器は、本明細書に記載される疾患を治療するために有効である組成物を保持し、無菌アクセスポートを有し得る。例えば、容器は、静脈内溶液バッグ、または皮下注射針によって穿刺可能なストッパーを有するバイアルであり得る。組成物中の活性剤は、本発明の化合物である。いくつかの態様において、容器上のまたは容器に付随するラベルは、選択される疾患を治療するために組成物が使用されることを示す。製品は、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添付文書を含め、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

【0990】

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。

20

【0991】

いくつかの実施形態において、本明細書で提供される組成物は、対象内で標的化されたゲノム改変を行うために、対象、例えばヒト対象に投与される。いくつかの実施形態において、対象から細胞が取得され、本明細書中に提供される医薬組成物のいずれかと接触される。いくつかの実施形態において、対象から取り出されて医薬組成物とex vivoで接触された細胞は、任意で、所望のゲノム改変が細胞内で行われたか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医薬組成物を送達する方法は公知であり、例えば米国特許第6,453,242号、6,503,717号、6,534,261号、6,599,692号、6,607,882号、6,689,558号、6,824,978号、6,933,113号、6,979,539号、7,013,219号、および7,163,824号に記載されており、それらの全ての開示は、その全体が参照により本明細書に組み込まれる。本明細書で提供される医薬組成物の説明は、主として、ヒトへの投与に適した医薬組成物に向けられているが、当業者には、そのような組成物が、あらゆる種類の動物または生物への投与に一般的に適していることが理解される（例えば獣医学的用途）。

30

【0992】

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。

40

【0993】

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

50

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WO2011/053982 A8)も参照のこと。

10

【0994】

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。

【0995】

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

【0996】

いくつかの実施形態において、本開示による組成物は、様々な疾患、障害、および/または状態のいずれかの治療に使用することができる。

【0997】

[治療方法]

疾患もしくは状態、および/または疾患もしくは状態を引き起こす遺伝子変異を治療する方法も提供される。これらの方法は、対象(例えば、ヒトなどの哺乳動物)に、本明細書に記載の塩基エディターシステム(例えば、塩基エディターおよびgRNA)をコードするポリヌクレオチドを含む治療有効量の医薬組成物を投与することを含む。いくつかの実施形態において、塩基エディターは、napDNAbpドメインと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とを含む融合タンパク質である。対象の細胞が、塩基エディターと、その塩基エディターを標的指向化させる1つまたは複数のガイドポリヌクレオチドとで形質導入され、目的の遺伝子に変異を含む核酸配列のA・TからG・Cへの改変(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ている場合)またはC・GからU・Aへの改変(細胞がシチジンデアミナーゼドメインで形質導入されている場合)がもたらされる。

30

【0998】

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。

40

【0999】

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)の目的遺伝子を標的とする、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、その疾患または状態を患っている、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。

50

【1000】

一実施形態では、本発明は、治療の進行を監視する方法を提供する。本方法は、疾患、障害、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）（例えば疾患または状態に関連する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。

10

【1001】

ある実施形態において、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、所望のゲノム変化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; および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。

20

【1002】

本明細書に提供されるのは、本明細書に記載の組成物またはシステム、例えば、塩基エディターシステムまたは塩基エディタータンパク質を用いて疾患または障害を治療する方法である。任意で、本明細書に記載の塩基エディターシステムは、1つまたは複数の他の処置と組み合わせることができる。一態様では、本明細書に提供されるのは、本明細書に記載の塩基エディターを投与することによって、それを必要とする対象におけるシュタルガルト病（SD）を治療する方法である。一態様では、本明細書で提供されるのは、本明細書で説明される塩基エディターを投与することによって、それを必要とする対象のパーキンソン病（PD）を治療する方法である。一態様では、本明細書で提供されるのは、本明細書で説明される塩基エディターを投与することによって、それを必要とする対象のレット症候群（RTT）を治療する方法である。一態様では、本明細書に提供されるのは、本明細書に記載の塩基エディターを投与することによって、それを必要とする対象においてハーラー症候群（HS）を治療する方法である。

30

【1003】

個々の対象の応答は、完全な応答、部分的な応答、または安定した疾患として特徴付けられ得る。いくつかの実施形態では、応答は部分的応答（PR）である。いくつかの実施形態では、応答は完全応答（CR）である。いくつかの実施形態において、応答は、対象の無増悪生存期間（例えば、安定した疾患）をもたらす。

40

【1004】

いくつかの実施形態において、治療は、そのヒト対象がその化合物（例えば塩基エディターシステム）で処置されなかった場合のヒト対象の予想される生存時間と比較して、ヒト対象の生存時間の増加をもたらす。いくつかの実施形態では、生存時間の増加は、ABE7塩基エディター（例えばABE7.10）を含む塩基エディターで処置された対象と比較して、疾患の進行がより遅いことを含む。

【1005】

いくつかの実施形態では、記載された方法で治療されるヒト対象は、子供（例えば、0

50

～18歳)である。他の実施形態では、記載された方法で治療されるヒト対象は、成人（例えば、18歳超）である。

【1006】

〔シュタルガルト病 (SD) 〕

シュタルガルト病 (シュタルガルト黄斑ジストロフィー、若年性黄斑変性症、または黄色斑眼底としても知られる) は、網膜 (すなわち、光を感知する眼の後ろの組織) の遺伝性障害である。シュタルガルト病は、黄斑変性症を引き起こすいくつかの遺伝性疾患の1つである。この病気は一般的に、小児期または青年期に視力喪失を引き起こす。ただし、場合によっては成人期の後半まで視力低下に気付かないこともあり得る。病気が進行して完全に失明することは稀である。一般に、黄斑の進行性の損傷 (変性) が時間の経過とともに起こるにつれて、視力喪失が時間の経過とともにゆっくりと進行し、20/200以下になる。一例では、本明細書に記載の方法で治療されるシュタルガルト病は、若年性シュタルガルト病を含む。別の例では、本明細書に記載の方法で治療されるシュタルガルト病は、遅発性シュタルガルト病を含む。別の例では、本明細書に記載の方法で治療されるシュタルガルト病は、シュタルガルト型優性黄斑ジストロフィーを含む。別の例では、本明細書に記載の方法で治療されるシュタルガルト病は、優性シュタルガルト様黄斑ジストロフィーを含む。

【1007】

シュタルガルト病の症状の進行は、患者ごとに異なり得る。病気の発症が早い患者は、一般的に、より急速な視力喪失を起こす傾向がある。視力低下は、最初はゆっくりと減少し、その後、急速に悪化してから横ばいになることもあり得る。シュタルガルト病のほとんどの患者は、20/200以下の視力に陥る。シュタルガルト病の人は、年をとるにつれ周辺 (横) の視力の一部を失い始めることもある。

【1008】

いくつかの実施形態では、病原性SNPがシュタルガルト病に関連している。場合により、病原性SNPはABCA4遺伝子にある。そして、場合によって、病原性突然変異は、A1038V、L541P、G1961E、またはそれらの組み合わせを含む。いくつかの実施形態において、病原性SNPは、弾性線維性仮性黄色腫に関連している。場合によって、病原性SNPはABCC6遺伝子にある。場合によって、病原性突然変異はR1141* (ナンセンス変異) を含む。いくつかの実施形態では、病原性SNPは中鎖アシル-CoAデヒドロゲナーゼ欠損症に関連している。場合によって、病原性SNPはACADM遺伝子にある。そして場合によっては、病原性突然変異はK329Eを含む。いくつかの実施形態において、病原性SNPは、重症複合免疫不全症に関連している。場合によって、病原性SNPはADA遺伝子にある。そして場合によって、病原性突然変異は、G216R、Q3*、またはそれらの組み合わせを含む。

【1009】

ABCA4ポリペプチドの例示的なアミノ酸配列を以下に提供する：

```
> sp | P78363 | ABCA4_HUMAN Retinal-specific phospholipid-transporting ATPase
ABCA4 OS=Homo sapiens OX=9606 GN=ABCA4 PE=1 SV=3
MGFVRQIQLLLWKNWTLRKRQKIRFVVELVWPLSLFLVLIWLRNANPLYSHHECHFPN
KAMPSAGMLPWLQGIFCNVNNPCFQSPTPGESPGIVSNYNNNSILARVYRDFQELLMNAP
ESQHLGRIWTELHILSQFMDTLRTHPERIAGRIRIRDILKDEETLTLFLIKNIGLSDSVVY
LLINSQVRPEQFAHGVPDLALKDIACSEALLERFIIFSQRRGAKTVRYALCSLSQGTQWI
EDTLYANVDFFKLFRVLPTLLDSRSQGINLRSWGILSDMSPRIQEFHRPSMQDLLWVT
RPLMQNGGPETFTKLMGILSDLLCGYPEGGGSRVLSFNWYEDNNYKAFLGIDSTRKDPY
SYDRRTTSFCNALIQSLESNPLTKIAWRAAKPLLGMKILYTPDSPAARRILKNANSTFEEL
EHVRKLVKAWEEVGPQIWYFFDNSTQMNMI RDTLGNPTVKDFLNRQLGEEGITA EAILN
FLYKGPRESQADDMANFDWRDIFNITDRTLRLVNQYLECLVLDKFESYNDETQLTQRALS
LLEENMFWAGVVFPDMYPWTSSLP PHVKYKIRMDIDVVEKTNKIKDRYWD SGPRADPV
EDFRYIWGGFAYLQDMVEQGITRSQVQAEAPVGIYLQQMPYPFCFVDDSFMIILNRCFPIF
MVLAWIYSVSM TVKSIVLEKELRLKETLKNQGV SNAVIWCTWFLDSFSIMSMSIFLLTIFI
```

MHGRILHYSDFILFLFLLAFSTATIMLCFLLSTFFSKASLAAACSGVIYFTLYLPHILCFA
WQDRMTAELKKAVSLLSPVAFGFGTEYLVRFEEQGLGLQWSNIGNSPTEGDEFSFLLSM
QMMLLDAAVYGLLAWYLDQVFPGDYGTPLPWYFLLQESYWLGGEGCSTREERALEKTE
PLTEETEDPEHPEGIHDSFFEREHPGWVPGVCVKNLVKIFPCGRPAVDRLNITFYENQIT
AFLGHNGAGKTTTSLILTGLLPPTSGTVLVGGRDIETSLDAVRQSLGMCPQHNLFFHHLT
VAEHMLFYAQLKGKSQEEAQLMEAMLEDTG LHHKRNEEAQDLSGGMQRKLSVAIAFV
GDAKVVLDEPTSGVDPYSRRSIWDLLLKYRSGRTIIMSTHHMDEADLLGDRIAIQAQRL
YCSGTPLFLKNCFGTGLYLT LVRKMKNIQSQRKGSEGTCSKSGFSTTCPAHVDDLTP
QVLDDGVNELMDVVLHHVPEAKLVEICIGQELIFLLPNKNFKHRAVASLFRELEETLADLG
LSSFGISDTPLEEIFLKVTEDSDSGLFLFAGGAQQKRENVNPRHPCLGPREKAGQTPQDSNV
CSPGAPAAHPEGQPPPEPECPGPQLNTGTQLVLQHVQALLVKRFQHTIRSHKDFLAQIVL
PATFVFLALMLSIVIPPFGEYPALTLHPWIYGGQYTFFSMDEPGSEQFTVLADVLLNKP
FGNRCLKEGWLPPEYPCGNSTPWKTPSVSPNITQLFQKQKWTQVNPSPSCRCSTREKLT
LPECPEGAGGLPPPQRTQRSTEILQDLTDNRNISDFLVKTYPALIRSSLKSKFWVNEQRYGG
ISIGGKLPVVPITGEALVGFLSDLGRIMNVSGGPITREASKEIPDFLKHLETEDNIKVWFNN
KGWHALVSFLNVAHNAILRASLPKDRSPEEYGITVISQPLNLTKEQLSEITVLTTSVDAVV
AICVIFSMSFVPASFVLYLIQERVNKS KHLQFISGVSPTTYWVTNFLWDIMNYSVSAGLV
VGIFIGFQKKAYTSPENLPALVALLLLYGWAVIPMMYPASFLFDVPSTAYVALSCANLFIG
INSSAITFILELFENNRTLLRFNAVLRKLLIVFPFHCLGRGLIDLALSQAVTDVYARFGEH
SANPFHWDLIGKNLFAMVVEGVVYFLLTLLVQRHFFLSQWIAEPTKEPIVDEDDDDVAEE
RQRIITGGNKTDILRLHELTKIYPGTSSPAVDRLCVGVRPGECFGLLVNGAGKTTTTFKML
TGDTTTSVSGDATVAGKSILTNISEVHQNMGYCPQFDAIDELLTGREHLYLYARLRGVPAE
EIEKVANWSIKSLGLTVYADCLAGTYSGGNKRKLSTAIALIGCPPLVLLDEPTTGMDPQAR
RMLWNVIVSIIREGRAVVLTSHSMEECEALCTRLAIMVKGAFCRMGTIQHLKSKFGDGYI
VTMKIKSPKDDLLPDLNPVEQFFQGNFPGSVQRRHYNMLQFQVSSSSSLARIFQLLLSHK
DSSLIEEYSVTQTTLQVFNFAKQQTESHDLPLHPRAAGASRQAQD (配列番号6)。

ガイドRNA配列は、配列GCTGTGTGTCGAAGTTCGCCCTGGAGAGGGTGまたはGCTGTG
TGTCGGAGTTCGCCCTGGAGAGGGTGにおいてABCA4遺伝子を標的化し、ここでPAM配
列には下線を付している。ガイドRNAは配列CACCUCUCCAGGGCGAACUUCGACACACA
GCまたはCACCUCUCCAGGGCGAACUCCGACACACAGCを含む。

【1010】

シュタルガルト病の1つまたは複数の症状には、視力の中央における両眼の灰色、黒色
、または曇ったスポットにおける中心的視力の、変動的で遅い喪失；明るい環境から暗い
環境に移動するときに、目が調整するのに通常よりも時間がかかること；目が明るい光に
対してより敏感になり得る；病気の後期における色盲、網膜色素上皮（RPE）の細胞にお
けるA2Eなどの有毒なリポフスチン色素の蓄積、光受容体の死、11-cis-レチナルデヒド
（11cRALまたはレチナール）の合成の増加、ロドプシンの再生の増加、リポフスチンの
蓄積、リポフスチン色素の形成、網膜変性、老廃物の生成、A2E（およびA2E関連分子）
の形成、A2E（およびA2E関連分子）の蓄積、脈絡膜血管新生、脈絡網膜萎縮、またはそ
れらの組み合わせが含まれるが、これらに限定されない。対象は、シュタルガルト病の
1つまたは複数の症状の改善を示すことができる。一実施形態では、症状の一つ以上にお
ける改善は少なくとも5%である。別の実施形態では、症状の一つ以上における改善は少な
くとも10%である。別の実施形態では、症状の一つ以上における改善は少なくとも15%で
ある。別の実施形態では、症状の一つ以上における改善は少なくとも20%である。別の実
施形態では、症状の一つ以上における改善は少なくとも25%である。別の実施形態では、
症状の一つ以上における改善は少なくとも30%である。別の実施形態では、症状の一つ以
上における改善は少なくとも35%である。別の実施形態では、症状の一つ以上における改
善は少なくとも40%である。別の実施形態では、症状の一つ以上における改善は少なくと
も50%である。別の実施形態では、症状の一つ以上における改善は少なくとも60%である
。別の実施形態では、症状の一つ以上における改善は少なくとも70%である。別の実施形

態では、症状の一つ以上における改善は少なくとも75%である。別の実施形態では、症状の一つ以上における改善は少なくとも80%である。別の実施形態では、症状の一つ以上における改善は少なくとも85%である。別の実施形態では、症状の一つ以上における改善は少なくとも90%である。別の実施形態では、症状の一つ以上における改善は少なくとも95%である。

【1011】

〔パーキンソン病（PD）〕

パーキンソン病（PD）は最も一般的な運動障害であり、世界中で600万人以上が罹患している。PDは若年性または早期発症を呈する可能性があるが、主には55歳以上の人々を苦しみ、65歳以降は疾患の発生率が急激に上昇する。PDの臨床的特徴には、動作緩慢、姿勢の不安定性、安静時振戦、および硬直が含まれ、これらは主に、黒質（SN）緻密部におけるドーパミン作動性ニューロンの進行性喪失に関連している。通常の老化の間、この領域のドーパミン作動性ニューロンの約0.1～0.2%が毎年失われると考えられているが、この率はPDの患者では大幅に加速され、これらのニューロンの約70～80%が失われると症状が現れる。PDのもう1つの病理学的特徴は、残ったドーパミン作動性ニューロンにおけるレビー小体（LB）として知られる α -シヌクレインタンパク性封入体の存在である。

【1012】

PDの症例の大部分は特発性であるが、症例の約10%が家族歴を報告しており、家族性および散発性の疾患に関連づけられる突然変異の数が増えています。ロイシンリッチリピートキナーゼ2（LRRK2）の常染色体優性G2019S変異は、家族性および散発性のPD患者の最も一般的な既知原因である。

【1013】

LRRK2ポリペプチドの例示的なアミノ酸配列を以下に提供する：

```
> sp | Q5S007 | LRRK2_HUMAN Leucine-rich repeat serine/threonine-protein kinase 2 OS=Homo sapiens OX=9606 GN=LRRK2 PE=1 SV=2
MASGSCQGCEEDEETLKKLIVRLNNVQEGKQIETLVQILEDLLVFTYSERASKLFQGKNIH
VPLLIVLDSYMRVASVQQVGWSLLCKLIEVCPGTMQSLMGPQDVGNDWEVLGVHQLIL
KMLTVHNASVNLVIGLKTLDLLLTSGKITLLILDEESDIFMLIFDAMHSFPANDEVQKLG
CKALHVLFFERVSEEQLTEFVENKDYMILLSALTNFKDEEEIVLHVHCLHSLAIPCNNVEV
LMSGNVRCYNIVVEAMKAFPMSERIQEVSCCLLHRLTLGNFFNILVLNEVHEFVVKAVQ
QYPENAALQISALSCALLTETIFLNQDLEEKNNENQENDDEGEEDKLFWLEACYKALTWH
RKNKHVQEAACWALNNLLMYQNSLHEKIGDEDGHFPAHREVMLSMLMHSSSKEVFQAS
ANALSTLLEQNVNFRKILLSKGIHLNVLELMQKHIHSPEVAESGCKMLNHLFEQSNTSLDI
MAAVVPKILTVMKRHETSLPVQLEALRAILHFIVPGMPEESREDTEFHKLNMVKKQCF
KNDIHKLVLAALNRFIGNPGIQKCGLKVISSIVHFPDALEMLSLEGAMDSVLHTLQMYPD
DQEIQCGLSLIGYLITKKNVFIGTGHLAKILVSSLYRFKDVAEIQTGKFQTLAILKLSAS
FSKLLVHHSFDLVIFHQMSSNIMEQKDQQFLNLCKCKFAKVMDDYLKNVMLERACDQ
NNSIMVECLLLLGADANQAKEGSSLICQVCEKESSPKLVLLLLNSGSREQDVRKALTISIG
KGDSQIISLLRLALDVANNSICLGFGCIGKVEPSWLGLPLFPDKTSNLRKQTNIASTLAR
MVIRYQMKSAVEEGTASGSDGNFSEDLVSKFDEWTFIPDSSMDSVFAQSDDLDEGSEGS
FLVKKKSNSISVGEFYRDAVLQRCSPNLQRHSNSLGPFDHEDLLKRKRKILSSDDSLRSSK
LQSHMRHSDSISSLASEREYITSLDLSANELRDIDALSQKCCISVHLEHLEKLELHQNALTS
FPQQLCETLKSLLTHLDLHSNKFTSFPSYLLKMSCIANLDVSRNDIGPSVVLDPVTKCPTLK
QFNLSYNQLSFVPENLTDVVEKLEQLILEGNKISGICSPRLKELKILNLSKNHISLSENFL
EACPKVESFSARMNFLAAMPFLPPSMTILKLSQNKFCIPEAILNPLHLRSLDMSSNDIQY
LPGPAHWKSLNLRRELLFSHNQISILDSEKAYLWSRVEKLHLSHNKLKEIPPEIGCLENLT
SLDVSYNLELRSFPNEMGKLSKIWDLPDELHLNFDKFHIGCKAKDIIRFLQQRLLKAVPY
NRMKLMIVGNTGSGKTTLLQQLMKTKKSDLGMQSATVGIDVKDWPIQIRDKRKRDLVL
NVWDFAGREEFYSTHPHFMFTQRALYLAVYDLKSGQAEVDAMKPWLFNIKARASSSPVIL
VGTHLDVSDEKQRKACMSKITKELLNKRGFPAIRDYHFNATEESDALAKLRKTIINESLN
```

FKIRDQLVVGQLIPDCYVELEKIILSERKNVPIEFVIDRKRLQLLVRENQLQLDENELPHA
VHFLNESGVLLHFQDPALQLSDLYFVEPKWLCIKMAQILTVKVEGCPKHPKGIISRRDVE
KFLSKKRKFPPKNYMSQYFKLLEKFQIALPIGEEYLLVPSSLSDHRPVIELPHCENSEIIIRLY
EMPYFPMGFWSRLINRLLEISPYMLSGRERLRPNRMWYWRQGIYLNWSPEAYCLVGSEV
LDNHPESEFLKITVPSCRKGCILLGQVVDHIDSLMEEWFPGLLEIDICGEGETLLKKWALYS
FNDGEEHQKILLDDLMMKKAEEGDLLVNPDPRLTIPISQIAPDLILADLPRNIMLNNDLE
FEQAPEFLLGDGSFGSVYRAAYEGEEVAVKIFNKHTSLRLLRQELVVLCHLHHPSLISLLA
AGIRPRMLVMELASKGSLDRLLQQDKASLTRTLQHRIALHVADGLRYLHSAMIIYRDLKP
HNVLLFTLYPNAIIAKIADYGIAQYCCRMGIKTSEGTGPFRAPEVARGNVIYNQQADVYS
FGLLLYDILTGTGRIVEGLKFPNEFDELEIQGKLPDPVKEYGCAPWPMVEKLIKQCLKENP
QERPTSAQVFDILNSAELVCLTRRILLPKNVIVECMVATHHNSRNASIWLGCIGHTDRGQL
SFLDLNTEGYTSEEVADSRILCLALVHLPVEKESWIVSGTQSGTLLVINTEDGKKRHTLEK
MTDSVTCLYCNSFSKQSKQKNFLLVGTADGKLAIFEDKTVKLKGAAPLKILNIGNVSTPL
MCLSESTNSTERNVMWGGCGTKIFSFSNDFTIQKLIETRISQLFSYAAFSDSNIITVVVD
ALYIAKQNSPVVEVWDKKTEKLCGLIDCVHFLREVMVKENKESKHKMSYSGRVKTLCLQ
KNTALWIGTGGGHILLDLSTRRLIRVIYNFCNSVRVMMTAQLGSLKNVMLVLGYNRKN
TEGTQKQKEIQSCLTVWDINLPHEVQNLEKHIEVRKELAEKMRRTSVE (配列番号3)

パーキンソン病は、運動に影響を与える進行性神経系障害です。症状は通常徐々に始まり、片方の手だけにかろうじて気づく震えから始まることもある。振戦が一般的であり得るが、障害はこわばりや動きの減速も一般的に引き起こす。パーキンソン病の初期段階では、被験者の顔はほとんどまたはまったく表情を示さない場合がある。歩行中に患者の腕が振れない場合がある。話し方が小声になったり不明瞭になったりすることがある。パーキンソン病の症状は、一般的に時間の経過とともに悪化する。

【1014】

パーキンソン病の兆候と症状は患者によって異なる。初期の兆候は軽度で見過ごされ得る。多くの場合、症状は体の片側から始まり、症状が両側に影響を及ぼし始めた後でも、通常はその側の方が悪くあり続ける。パーキンソン病の兆候と症状には、振戦、運動緩慢（動作緩慢）、筋肉硬直、姿勢とバランスの障害、自動動作の喪失（例えば、まばたき、微笑み、歩行中の腕の振りなど、無意識の動作を行う能力の低下）、話し方の変化（例えば、小声で話す、早口で話す、ろれつが回らない、または話す前にためらう、話し方が通常の抑揚ではなくより単調であり得る）、書き方の変化（たとえば、書くことが難しくなり得、書いたものが小さく見え得る）、血圧の変化（例えば、患者は、立つときに血圧の突然の低下（起立性低血圧）のためにめまいや立ちくらみを感じることもある）、嗅覚の機能不全（例えば、患者は匂いの感覚に困難を経験し得る）、倦怠感（例えば、パーキンソン病の多くの患者はエネルギーを失い、特に一日の後半に倦怠感を経験する）、痛み（例えば、パーキンソン病の一部の患者は、体の特定の領域または体全体のいずれかで痛みを経験する）、性的機能不全（例えば、パーキンソン病の一部の患者は性的欲求または性的能力の低下に気づく）、またはそれらの組み合わせが含まれるが、これらに限定されない。

【1015】

対象は、パーキンソン病の1つまたは複数の症状の改善を示すことができる。一実施形態では、症状の一つ以上における改善は少なくとも5%である。別の実施形態では、症状の一つ以上における改善は少なくとも10%である。別の実施形態では、症状の一つ以上における改善は少なくとも15%である。別の実施形態では、症状の一つ以上における改善は少なくとも20%である。別の実施形態では、症状の一つ以上における改善は少なくとも25%である。別の実施形態では、症状の一つ以上における改善は少なくとも30%である。別の実施形態では、症状の一つ以上における改善は少なくとも35%である。別の実施形態では、症状の一つ以上における改善は少なくとも40%である。別の実施形態では、症状の一つ以上における改善は少なくとも50%である。別の実施形態では、症状の一つ以上における改善は少なくとも60%である。別の実施形態では、症状の一つ以上における改善は少なく

とも70%である。別の実施形態では、症状の一つ以上における改善は少なくとも75%である。別の実施形態では、症状の一つ以上における改善は少なくとも80%である。別の実施形態では、症状の一つ以上における改善は少なくとも85%である。別の実施形態では、症状の一つ以上における改善は少なくとも90%である。別の実施形態では、症状の一つ以上における改善は少なくとも95%である。

【 1 0 1 6 】

対象（例えば、ヒトなどの哺乳動物）に、本明細書に記載の塩基エディターシステム（例えば塩基エディターおよびgRNA）をコードするポリヌクレオチドを含む治療有効量の医薬組成物を投与することを含む、PDおよび/またはそれを引き起こすLRRK2の遺伝子変異を治療する方法も提供される。いくつかの実施形態において、塩基エディターは、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またはシチジンデアミナーゼドメインとを含む融合タンパク質である。対象の細胞は、塩基エディターと、その塩基エディターを標的指向化させる1つまたは複数のガイドポリヌクレオチドとで形質導入され、LRRK2遺伝子に変異を含む核酸配列のA・TからG・Cへの改変（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への改変（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）がもたらされる。

【 1 0 1 7 】

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。

【 1 0 1 8 】

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LRRK2遺伝子を標的とする、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、PD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。本明細書の組成物はまた、PDが関係している可能性がある他の任意の障害の治療にも使用され得る。

【 1 0 1 9 】

いくつかの実施形態において、ガイドRNAは、標的配列GCTCGCCCTTCTTCTTCCCCTGTGA、GTCTTTCCCTCCAGGCTCGCCCTTCTTCTTCCCCTGTGA、TCACAGGGGAAGAAGAAGGGCGAGCCTGGAGGGAAAGACにおいてLRRK遺伝子を標的とする。いくつかの実施形態において、ガイドRNAは、配列UCGCCCUUCUUCUUCUUCCCCUGUGA、GUCUUUCCCUCCAGGCUCGCCCUUCUUCUUCUUCCCCUGUGA、UCACAGGGGAAGAAGAAGGGCGAGC、またはUCACAGGGGAAGAAGAAGGGCGAGCCUGGAGGGAAAGACを含む。いくつかの実施形態において、ガイドRNAは、配列UCCGACUAUAUGAAAUGCCUUAUUUCCA AUGGGAUUUUGGまたはUUGCAAAGAUUGCUGACUAGGGCAUUGCUCAGUACUGCUGUAGAAUGGを含む。

【 1 0 2 0 】

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、PDに関連する障害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PD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。

【1021】

いくつかの実施形態において、本明細書で提供される組成物は、対象内で標的化されたゲノム修飾をもたらすために、対象、例えば、ヒト対象に投与される。ある実施形態において、細胞が対象から取得され、本明細書に提供される医薬組成物のいずれかと接触される。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公知であり、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; および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。

【1022】

〔ハーラー症候群 (HS) 〕

ハーラー症候群 (Hurler Syndrome) は、ムコ多糖症1型 (MPS1) の最も重症な形態であり、約20万人に1人で起こるするまれな常染色体劣性リソソーム蓄積症である。MPS1は、骨格異常、認知障害、心臓病、呼吸器系の問題、肝臓と脾臓の肥大、および予測寿命の減少を特徴とする。MPS1は、アルファ-L-イズロニダーゼ (IDUA) 遺伝子の変異によって引き起こされ、リソソーム内のグリコサミノグリカンの分解に不可欠なアルファ-L-イズロニダーゼの欠損をもたらす。ハーラー症候群の患者の標準治療は骨髄移植であるが、この治療法では、特に中枢神経系ですでに発生している損傷を修正することはできず、移植に伴う死亡のリスクもある。したがって、ハーラー症候群の患者を治療するための新規の組成物および方法が必要である。

【1023】

本開示は、塩基エディター媒介遺伝子編集によって修正することができる、および／または症状を治療もしくは改善することができる点突然変異に、関連するかまたはそれによって引き起こされる、ハーラー症候群の治療のための方法を提供する。

【1024】

対象 (例えば、ヒトなどの哺乳動物) に、本明細書に記載の塩基エディターシステム (例えばABE8塩基エディターおよびgRNA) をコードするポリヌクレオチドを含む治療有効量の医薬組成物を投与することを含む、ハーラー症候群および/またはそれを引き起こすIDUA遺伝子の変異を治療する方法も提供される。いくつかの実施形態において、塩基エディターは、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を含む融合タンパク質である。対象の細胞は、塩基エディターと、その塩基エディターを標的指向化させる1つまたは複数のガイドポリヌクレオチドとで形質導入され、IDUA遺伝子に変異を含む核酸配列のA・TからG・Cへの改変がもたらされる。

【1025】

IDUAポリペプチドの例示的なアミノ酸配列を以下に提供する：

```
> sp | P35475 | IDUA_HUMAN Alpha-L-iduronidase OS=Homo sapiens OX=9606
GN=IDUA PE=1 SV=2
MRPLRPRAALLALLASLLAAPPVAPAEAPHLVHVDAARALWPLRRFWRSTGFCPLPHS
QADQYVLSWDQQLNLAYVGAVPHRGIKQVRTHWLLELVTTTRGSTGRGLSYNFTHL DGY
LDLLRENQLLPGFELMGSA SGHFTDFEDKQQVF EWKDLVSSLARRYIGRYGLAHVSKWN
FETWNEPDHHD FDNVSM TMQGFLNYYDACSEGLRAAS PALRLGGPGDSFHTPPRSPLS
WGLLRHCHDGTNFFTGEAGVRLDYISLHRKGARSSISILEQEKVVAQQIRQLFPKFADTPI
YNDEADPLVGWSLPQPWRADV TYAAMVVKVIAQHQNLLANTTSAPFYALLSNDNAFL
```

SYHPHPFAQRTLTARFQVNNTRPPHVQLLRKPVLTAMGLLALLDEEQLWAEVSQAGTV
LDSNHTVGVLASAHRPQGPADAWRAAVLIYASDDTRAHPNRSVAVTLRLRGVPPGPGLV
YVTRYLDNGLCSPDGEWRRLLGRPVFPTAEQFRRMRAAEDPVAAAPRPLPAGGRLTLRPA
LRLPSLLLHVHVCARPEKPPGQVTRLRALPLTQGQLVLVWSDEHVGSKCLWTYEIQFSQD
GKAYTPVSRKPSTFNLVVFSPDTGAVSGSYRVRALDYWARPGPFSDPVPYLEVPVPRGPP
SPGNP (配列番号4).

【1026】

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。

10

【1027】

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IDUA遺伝子を標的とする、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、ハーラー症候群を患っている、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。本明細書の組成物はまた、ハーラー症候群が関係している可能性がある他の任意の障害の治療にも使用され得る。ガイドRNA配列は、スパーサー配列CTTTTC ACTTTTCCTGCCGGGG (R255X)、AGCTTCCATGTCCAGCCTTC (R106W)、ACCATG AAGTCAAAATCATT (T158M)、またはGCTTTCAGCCCCGTTTCTTG (R270X)を含み得る。

20

【1028】

一実施形態では、本発明は、治療の進行を監視する方法を提供する。本方法は、ハーラー症候群に関連する障害、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)(例えばハーラー症候群に関連する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。

30

【1029】

ある実施形態において、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; および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。

40

【1030】

治療されるハーラー症候群の1つまたは複数の症状には、粗い顔の特徴、角膜薄濁、肝肥大、後弯/突背、ヘルニア、睡眠障害/いびきなどの気道関連症状、脾腫、心臓弁異常、

50

認知障害、多発性ジストーシス、舌の肥大、関節収縮、扁桃腺の肥大、またはそれらの組み合わせが含まれるが、これらに限定されない。

【1031】

対象は、ハーラー症候群の1つまたは複数の症状の改善を示すことができる。一実施形態では、症状の一つ以上における改善は少なくとも5%である。別の実施形態では、症状の一つ以上における改善は少なくとも10%である。別の実施形態では、症状の一つ以上における改善は少なくとも15%である。別の実施形態では、症状の一つ以上における改善は少なくとも20%である。別の実施形態では、症状の一つ以上における改善は少なくとも25%である。別の実施形態では、症状の一つ以上における改善は少なくとも30%である。別の実施形態では、症状の一つ以上における改善は少なくとも35%である。別の実施形態では、症状の一つ以上における改善は少なくとも40%である。別の実施形態では、症状の一つ以上における改善は少なくとも50%である。別の実施形態では、症状の一つ以上における改善は少なくとも60%である。別の実施形態では、症状の一つ以上における改善は少なくとも70%である。別の実施形態では、症状の一つ以上における改善は少なくとも75%である。別の実施形態では、症状の一つ以上における改善は少なくとも80%である。別の実施形態では、症状の一つ以上における改善は少なくとも85%である。別の実施形態では、症状の一つ以上における改善は少なくとも90%である。別の実施形態では、症状の一つ以上における改善は少なくとも95%である。

10

【1032】

〔レット症候群 (RTT) 〕

対象（例えば、ヒトなどの哺乳動物）に、本明細書に記載の塩基エディターシステム（例えば塩基エディターおよびgRNA）をコードするポリヌクレオチドを含む治療有効量の医薬組成物を投与することを含む、レット症候群（RTT）および/またはそれを引き起こすMecp2の遺伝子変異を治療する方法も提供される。いくつかの実施形態において、塩基エディターは、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またはシチジンデアミナーゼドメインとを含む融合タンパク質である。対象の細胞は、塩基エディターと、その塩基エディターを標的指向化させる1つまたは複数のガイドポリヌクレオチドとで形質導入され、Mecp2遺伝子に変異を含む核酸配列のA・TからG・Cへの改変（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への改変（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）がもたらされる。

20

30

【1033】

MECP2ポリペプチドの例示的なアミノ酸配列を以下に提供する：

```
> sp | P51608 | MECP2_HUMAN Methyl-CpG-binding protein 2 OS=Homo sapiens
OX=9606 GN=MECP2 PE=1 SV=1
MVAGMLGLREEKSEDQDLQGLKDKPLKFKKVKKDKKKEKEGKHEPVQPSAHHSAEPAEA
GKAETSESGSGSAPAVPEASASPKQRRSIIRDRGPMYDDPTLPEGWTRKLKQRKSGRSAGK
YDVYLINPQGKAFRSKVELIAYFEKVGDTSLDPNDFDFTVTGRGSPSRREQPKPKPKSPK
APGTGRGRGRPKGSGTTRPKAATSEGVQVKRVLEKSPGKLLVKMPFQTSPGGKAEGGGA
TTSTQVMVIKRPGRKRKAADPQAIPKKRGRKPGSVVAAAAAEAKKKAVKESSIRSVQET
VLPIKKRKTRETVSIEVKEVVKPLLVLSTLGEKSGKGLKTKCKSPGRKSKESSPKGRSSSASSP
PKKEHHHHHHHSESPKAPVPLLPLPPPPPEPESSEDPTSPPEPQDLSSSVCKEEKMPRG
GSLESDGCPKEPAKTQPAVATAATAAEKYKHRGEGERKDIVSSSMRPNREEPVDSRTPV
TERVS (配列番号5).
```

40

【1034】

ガイドRNA配列は、スパーサーCTTTTCACTTTTCCTGCCGGGG (R255X)、AGCTTCCATGTCCAGCCTTC (R106W)、ACCATGAAGTCAAAATCATT (T158M)、またはGCTTT CAGCCCCGTTTCTTG (R270X)を含み得る。

【1035】

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

50

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。

【1036】

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Mecp2遺伝子を標的とする、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、RTT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。本明細書の組成物はまた、RTTが関係している可能性がある他の任意の障害の治療にも使用され得る。

10

【1037】

一実施形態では、本発明は、治療の進行を監視する方法を提供する。本方法は、RTTに関連する障害、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RTT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。

20

【1038】

ある実施形態において、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; および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。

30

【1039】

治療されるレット症候群の1つまたは複数の症状には、就寝抵抗、入眠遅延、睡眠時間、睡眠不安、夜の目覚め、睡眠時随伴症、睡眠呼吸障害、日中の眠気、手の機能、歩行、言語的および非言語的コミュニケーション、理解、注意、行動の問題、気分、発作活動、挙動のおよび感情的な特徴、またはそれらの組み合わせが含まれるが、これらに限定されない。

40

【1040】

対象は、レット症候群の1つまたは複数の症状の改善を示すことができる。一実施形態では、症状の一つ以上における改善は少なくとも5%である。別の実施形態では、症状の一つ以上における改善は少なくとも10%である。別の実施形態では、症状の一つ以上における改善は少なくとも15%である。別の実施形態では、症状の一つ以上における改善は少なくとも20%である。別の実施形態では、症状の一つ以上における改善は少なくとも25%である。別の実施形態では、症状の一つ以上における改善は少なくとも30%である。別の実施形態では、症状の一つ以上における改善は少なくとも35%である。別の実施形態では、症状の一つ以上における改善は少なくとも40%である。別の実施形態では、症状の一つ以

50

上における改善は少なくとも50%である。別の実施形態では、症状の一つ以上における改善は少なくとも60%である。別の実施形態では、症状の一つ以上における改善は少なくとも70%である。別の実施形態では、症状の一つ以上における改善は少なくとも75%である。別の実施形態では、症状の一つ以上における改善は少なくとも80%である。別の実施形態では、症状の一つ以上における改善は少なくとも85%である。別の実施形態では、症状の一つ以上における改善は少なくとも90%である。別の実施形態では、症状の一つ以上における改善は少なくとも95%である。

【1041】

[キット]

本開示の種々の態様は、塩基エディターシステムを含むキットを提供する。1つの実施形態において、キットは、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。融合タンパク質は、デアミナーゼドメイン（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）と核酸プログラミング可能なDNA結合タンパク質（napDNAbp）とを含む。ある実施形態において、キットは、対象核酸分子を標的化することができる少なくとも1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。

【1042】

キットは、いくつかの実施形態において、1つ以上の変異を編集するためにキットを使用するための指示書を提供する。指示書は、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、指示書は、以下のうちの少なくとも1つを含む：注意書き；警告；臨床試験；および/または参考文献。指示書は、容器上に直接印刷してもよいし（それが存在する場合）、容器に貼付するラベルとして印刷されてもよいし、容器内または容器と一緒に供給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してもよい。さらなる実施形態では、キットは、適切な操作パラメータのためのラベルまたは別個の挿入物（パッケージインサート）の形態で指示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（複数可）として使用される、適切な陽性および陰性対照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。キットはさらに、（無菌の）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の容器を含むことができる。それはさらに、他の緩衝液、希釈剤、フィルター、針、シリンジ、および使用説明書付きの添付文書を含む、商業的および使用者の観点から望ましい他の材料を含むことができる。

【1043】

本発明の実施は、別段の表示がない限り、分子生物学（組換え技術を含む）、微生物学、細胞生物学、生化学および免疫学の従来技術を利用し、これらは当業者の技量の範囲内である。そのような技術は、“Molecular Cloning: A Laboratory Manual”, second edition (Sambrook, 1989); “Oligonucleotide Synthesis” (Gait, 1984); “Animal Cell Culture” (Freshney, 1987); “Methods in Enzymology” “Handbook of Experimental Immunology” (Weir, 1996); “Gene Transfer Vectors for Mammalian Cells” (Miller and Calos, 1987); “Current Protocols in Molecular Biology” (Ausubel, 1987); “PCR: The Polymerase Chain Reaction”, (Mullis, 1994); “Current Protocols in Immunology” (Coligan, 1991)などの文献で詳しく説明されている。これらの技術は、本発明のポリヌクレオチドおよびポリペプチドの製造に適用可能であり、したがって、本発明の製造および実施において考慮され得る。特定の実施形態のために特に有用な技術は、以下のセクションで論じられる。

【1044】

以下の実施例は、当業者に本発明のアッセイ、スクリーニング、および治療方法をいかに作って使用するかの完全な開示および説明を提供するために記載されており、本発明者らがその発明とみなすものの範囲を限定することを意図するものではない。

【実施例】

【1045】

【実施例1：編集効率が向上したアデノシン塩基エディター】

Tad7.10-dCas9融合タンパク質を含む塩基編集システムは、約10～20%の効率で標的ポリヌクレオチドを編集することができるが、より高い効率を必要とする用途ではその使用が制限され得る。効率および特異性が増大したアデニン塩基エディターを同定するための取り組みにおいて、アデノシンデアミナーゼTadA7.10を含む構築物を、エラープローンPCRによって変異誘発させ、続いて、核酸プログラム可能なDNA結合タンパク質であるdCas9をコードする核酸配列に隣接させて発現ベクター中にクローン化した（図1A）。アデノシンデアミナーゼバリエーションを含む発現ベクターを、クロラムフェニコール耐性（CamR）およびスペクチノマイシン耐性（SpectR）をコードしかつ2つの点突然変異によって機能しなくなったカナマイシン耐性遺伝子を有する選択プラスミドと共に、コンピテントな細菌細胞に共形質転換した（進化ラウンド7戦略）（図1B）。カナマイシン耐性の回復について細胞を選択し、これはアデノシンデアミナーゼ活性の指標であった。その後の選択ラウンドでは、クロラムフェニコール耐性（CamR）およびスペクチノマイシン耐性（SpectR）をコードしかつ3つの点突然変異によって非機能性とされたカナマイシン耐性遺伝子を有するプラスミドと共に、該発現ベクターをコンピテント細胞に共形質転換した（進化ラウンド8戦略）（図1C）。不活性化されたカナマイシン耐性遺伝子の核酸配列を以下に提供する。

```
ccggaattgccagctggggcgccctctggttaagggttgggaagccctgcaaagtaaactggatgg
ctttcttgccgccaaggatctgatggcgaggggatcaagatctgatcaagagacaggatgagg
atccttttcgcATGATCGAATAAGATGGATTGCACGCAGGTTCTCCGGCCGCTTAGGTGGAGCGC
CTATTCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGT
CAGCGCAGGGGCGCCCGGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAAGTGA
GGACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGCGCAGCTGTGCTCGAC
GTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGT
CATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATAC
GCTTGATCCGGCTACCTGCCCATTTCGACCACCAAGCGAAACATCGCATCGAGCGAGCACGTACT
CGGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAG
CCGAAGTGTTCGCCAGGCTCAAGGCGCGCATGCCCCGACGGCGAGGATCTCGTCGTGACCCATGG
CGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCAATTAAGTGGGC
CGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGC
TTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTTCGACGCG
CATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTAA
```

【1046】

上記配列において、小文字はカナマイシン耐性プロモーター領域を示し、太字の配列は標的化された不活性化部分（Q4*及びW15*）を示し、斜体の配列はカナマイシン耐性遺伝子の標的化されあ不活性部位（D208N）を示し、下線の配列はPAM配列を示す。

【1047】

再び、カナマイシン濃度を増加させていった一連のアガロースプレート上で細胞を平板培養した。図2に示されるように、効率的な塩基編集活性を有するアデノシンデアミナーゼバリエーションは、カナマイシン耐性遺伝子に存在する突然変異を修正することができ、さらなる分析のために選択された。細菌細胞において効率的な塩基編集を示すアデノシンデ

アミナーゼバリエーション塩基エディターを表13に記載する。選択されたアデノシンデアミナーゼバリエーションを含む塩基エディターをコードする哺乳類発現ベクターを作製した。

【1048】

E6V (E7Vとも呼ばれる) 変異を含む鎌状赤血球症に関連するβグロビン蛋白質を発現するHek293T細胞を用いて、アデノシンデアミナーゼバリエーションの編集効率を試験した(図3Aおよび3B)。「Hek293T/HBBE6V」細胞と称されるこれらの細胞は、表13に列挙されるABE8塩基エディターを含む融合タンパク質を含む塩基編集システムを発現するレンチウイルスベクターを用いて形質導入された。ABE8塩基エディターは、循環置換体(circular permutant) Cas9および二部分核局在配列を含む足場にアデノシンデアミナーゼバリエーションをクローニングすることによって作製した。循環置換体Cas9は当技術分野で公知であり、例えばOakes et al., Cell 176,254-267, 2019に記載されている。これらの配列は下記に提供される。

【1049】

胎児ヘモグロビンのアップレギュレーションは鎌状赤血球症を克服するための治療アプローチである。図3Aは、胎児ヘモグロビンの上方調節のための治療的に関連する部位を示す。残基5および8のアデノシンを編集することにより、BCL11A結合を有意に低下させることができ、それにより胎児ヘモグロビンの発現を増加させることができる。図3Aを参照すると、ABE8塩基エディターは、ABE7.10塩基エディターよりも約2~3倍多い塩基編集活性を示した。

【1050】

表13：新規アデニン塩基エディターABE8

10

20

30

40

50

【表 13】

プラスミド ID	説明	機能
280	ABE8.1	monomer_TadA*7.10 + Y147T
281	ABE8.2	monomer_TadA*7.10 + Y147R
282	ABE8.3	monomer_TadA*7.10 + Q154S
283	ABE8.4	monomer_TadA*7.10 + Y123H
284	ABE8.5	monomer_TadA*7.10 + V82S
285	ABE8.6	monomer_TadA*7.10 + T166R
286	ABE8.7	monomer_TadA*7.10 + Q154R
287	ABE8.8	monomer_Y147R_Q154R_Y123H
288	ABE8.9	monomer_Y147R_Q154R_I76Y
289	ABE8.10	monomer_Y147R_Q154R_T166R
290	ABE8.11	monomer_Y147T_Q154R
291	ABE8.12	monomer_Y147T_Q154S
292	ABE8.13	monomer_H123Y123H_ Y147R_Q154R_I76Y
293	ABE8.14	heterodimer_TadA*7.10 + Y147T
294	ABE8.15	heterodimer_TadA*7.10 + Y147R
295	ABE8.16	heterodimer_TadA*7.10 + Q154S
296	ABE8.17	heterodimer_TadA*7.10 + Y123H
297	ABE8.18	heterodimer_TadA*7.10 + V82S
298	ABE8.19	heterodimer_TadA*7.10 + T166R
299	ABE8.20	heterodimer_TadA*7.10 + Q154R
300	ABE8.21	heterodimer_Y147R_Q154R_Y123H
301	ABE8.22	heterodimer_Y147R_Q154R_I76Y
302	ABE8.23	heterodimer_Y147R_Q154R_T166R
303	ABE8.24	heterodimer_Y147T_Q154R
304	ABE8.25	heterodimer_Y147T_Q154S

【1051】

図4を参照して、ABE8塩基エディターを、HBB E6Vをコードするポリヌクレオチドを標的とする18、19、20、21、または22ヌクレオチドのガイドRNAと共に、Hek293T/HBBE6V細胞に導入した。ABE8エディターは、循環置換体 (Cp)-Cas9に融合された場合、編集効率の増加を示した。合計で、40の異なるABE8構築物 (表14) および3つのABE7.10構築物を、Hek293T/HBBE6V細胞における編集活性について試験した。例示的な構築物の配列は以下の通りである。編集の特異性を評価するために、標的および意図しないまたはバイスタンダーの変異をモニターした (図5)。コドン5におけるアデノシンの意図しない編集はサイレントであった。しかし、コドン9の意図しない編集はセリンからプロリン

への変異をもたらした。再び図5を参照すると、複数のABE8塩基エディターが、ABE7.10エディターと比較して編集効率および特異性の増加を示し、いずれのエディターもセリンからプロリンへのミスセンス突然変異につながる有意なバースタンダー編集を有していなかった。

【1052】

鎌状赤血球変異を含む線維芽細胞において、選択されたABE8塩基エディターおよびABE7.10塩基エディター対照のさらなる分析を行った。図6に示されるように、ABE8エディターは、ABE7.10と比較して、塩基編集活性が増加していた。ABE8.18は約70%の効率を示した。選択されたABE8エディターは前例のない特異性も示した。重要なことに、全てのABE8エディターについての平均インデル形成は0.1%未満であった。

【1053】

10

20

30

40

50

【表 1 4】

プラスミド ID	説明	機能
335	NGC PAM CP5 variant (<i>S. pyogenes</i> Cas9)_ABE8.1	monomer_TadA*7.10 + Y147T
336	NGC PAM CP5 variant (<i>S. pyogrouns</i> Cas9)_ABE8.2	monomer_TadA*7.10 + Y147R
337	NGC PAM CP5 variant (<i>S. pyogenes</i> Cas9)_ABE8.3	monomer_TadA*7.10 + Q154S
338	NGC PAM CP5 variant (<i>S. pyogenes</i> Cas9)_ABE8.4	monomer_TadA*7.10 + Y123H
339	NGC PAM CP5 variant (<i>S. pyogenes</i> Cas9)_ABE8.5	monomer_TadA*7.10 + V82S
340	NGC PAM CP5 variant (<i>S. pyogenes</i> Cas9)_ABE8.6	monomer_TadA*7.10 + T166R
341	NGC PAM CP5 variant (<i>S. pyogenes</i> Cas9)_ABE8.7	monomer_TadA*7.10 + Q154R
342	NGC PAM CP5 variant (<i>S. pyogenes</i> Cas9)_ABE8.8	monomer_Y147R_Q154R_Y123H
343	NGC PAM CP5 variant (<i>S. pyogenes</i> Cas9)_ABE8.9	monomer_Y147R_Q154R_I76Y
344	NGC PAM CP5 variant (<i>S. pyogenes</i> Cas9)_ABE8.10	monomer_Y147R_Q154R_T166R
345	NGC PAM CP5 variant (<i>S. pyogenes</i> Cas9)_ABE8.11	monomer_Y147T_Q154R
346	NGC PAM CP5 variant (<i>S. pyogenes</i> Cas9)_ABE8.12	monomer_Y147T_Q154S
347	NGC PAM CP5 variant (<i>S. pyogenes</i> Cas9)_ABE8.13	monomer_H123Y123H_Y147R_Q154R_I76Y
348	NGC PAM CP5 variant (<i>S. pyogenes</i> Cas9)_ABE7.10	monomer_deletion at TadA7.10* residue 149
349	NGC PAM CP5 variant (<i>S. pyogenes</i> Cas9)_ABE7.10	monomer_deletion at TadA7.10* residue 150
350	NGC PAM CP5 variant (<i>S. pyogenes</i> Cas9)_ABE7.10	monomer_deletion at TadA7.10* residue 151
351	NGC PAM CP5 variant (<i>S. pyogenes</i> Cas9)_ABE7.10	monomer_deletion at TadA7.10* residue 152
352	NGC PAM CP5 variant (<i>S. pyogenes</i> Cas9)_ABE7.10	monomer_deletion at TadA7.10* residue 153

10

20

30

40

50

353	NGC PAM CP5 variant (<i>S. pyogenes</i> Cas9)_ABE7.10	monomer_deletion at TadA7.10* residue 154
354	NGC PAM CP5 variant (<i>S. pyogenes</i> Cas9)_ABE7.10	monomer_deletion at TadA7.10* residue 155
355	NGC PAM CP5 variant (<i>S. pyogenes</i> Cas9)_ABE7.10	monomer_deletion at TadA7.10* residue 156
356	NGC PAM CP5 variant (<i>S. pyogenes</i> Cas9)_ABE7.10	monomer_deletion at TadA7.10* residue 157
357	NGC PAM CP5 variant (<i>S. pyogenes</i> Cas9)_ABE8.14	heterodimer_TadA*7.10 + Y147T
358	NGC PAM CP5 variant (<i>S. pyogenes</i> Cas9)_ABE8.15	heterodimer_TadA*7.10 + Y147R
359	NGC PAM CP5 variant (<i>S. pyogenes</i> Cas9)_ABE8.16	heterodimer_TadA*7.10 + Q154S
360	NGC PAM CP5 variant (<i>S. pyogenes</i> Cas9)_ABE8.17	heterodimer_TadA*7.10 + Y123H
361	NGC PAM C P5 variant (<i>S. pyogenes</i> Cas9)_ABE8.18	heterodimer_TadA*7.10 + V82S
362	NGC PAM CP5 variant (<i>S. pyogenes</i> Cas9)_ABE8.19	heterodimer_TadA*7.10 + T166R
363	NGC PAM CP5 variant (<i>S. pyogenes</i> Cas9)_ABE8.20	heterodimer_TadA*7.10 + Q154R
364	NGC PAM CP5 variant (<i>S. pyogenes</i> Cas9)_ABE8.21	heterodimer_Y147R_Q154R_Y123H
365	NGC PAM CP5 variant (<i>S. pyogenes</i> Cas9)_ABE8.22	heterodimer_Y147R_Q154R_I76Y
366	NGC PAM CP5 variant (<i>S. pyogenes</i> Cas9)_ABE8.23	heterodimer_Y147R_Q154R_T166R
367	NGC PAM CP5 variant (<i>S. pyogenes</i> Cas9)_ABE8.24	heterodimer_Y147T_Q154R
368	NGC PAM CP5 variant (<i>S. pyogenes</i> Cas9)_ABE8.25	heterodimer_Y147T_Q154S
369	NGC PAM CP5 variant (<i>S. pyogenes</i> Cas9)_ABE8.26	heterodimer_H123Y123H_Y147R_Q154R_I76Y
370	NGC PAM CP5 variant (<i>S. pyogenes</i> Cas9)_ABE7.10	heterodimer_deletion at TadA7.10* residue 149
371	NGC PAM CP5 variant (<i>S. pyogenes</i> Cas9)_ABE7.10	heterodimer_deletion at TadA7.10* residue 150

10

20

30

40

50

372	NGC PAM CP5 variant (<i>S. pyogenes</i> Cas9)_ABE7.10	heterodimer_deletion at TadA7.10* residue 151
373	NGC PAM CP5 variant (<i>S. pyogenes</i> Cas9)_ABE7.10	heterodimer_deletion at TadA7.10* residue 152
374	NGC PAM CP5 variant (<i>S. pyogenes</i> Cas9)_ABE7.10	heterodimer_deletion at TadA7.10* residue 153
375	NGC PAM CP5 variant (<i>S. pyogenes</i> Cas9)_ABE7.10	heterodimer_deletion at TadA7.10* residue 154
376	NGC PAM CP5 variant (<i>S. pyogenes</i> Cas9)_ABE7.10	heterodimer_deletion at TadA7.10* residue 155
377	NGC PAM CP5 variant (<i>S. pyogenes</i> Cas9)_ABE7.10	heterodimer_deletion at TadA7.10* residue 156
378	NGC PAM CP5 variant (<i>S. pyogenes</i> Cas9)_ABE7.10	heterodimer_deletion at TadA7.10* residue 157

10

20

【1054】

【実施例2：ABE8設計のためのコドン最適化とNLS選択】

Cas9のコドン用法と核局在化配列が真核生物のゲノム編集効率を劇的に変化させ得ることが立証されている（例えば、Kim, S. et al., Rescue of high-specificity Cas9 variants using sgRNAs with matched 5' nucleotides. *Genome Biol* 18, 218, doi:10.1186/s13059-017-1355-3 (2017); Mikami, M. et al., Comparison of CRISPR/Cas9 expression constructs for efficient targeted mutagenesis in rice. *Plant Mol Biol* 88, 561-572, doi:10.1007/s11103-015-0342-x (2015); Jinek, M. et al., RNA-programmed genome editing in human cells. *Elife* 2, e00471, doi:10.7554/eLife.00471 (2013)参照）。塩基エディターの元のCas9n成分は、6つの潜在的なポリアデニル化部位を含み、真核生物での発現を低下させる（例えば、Kim, S. et al., Rescue of high-specificity Cas9 variants using sgRNAs with matched 5' nucleotides. *Genome Biol* 18, 218, doi:10.1186/s13059-017-1355-3 (2017); Komor, A. C. et al., Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533, 420-424, doi:10.1038/nature17946 (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, doi:10.1038/nature24644 (2017)参照）。これを高度に最適化されたコドン配列で置き換えると、塩基編集効率が向上する（例えばCong, L. et al. Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819-823, doi:10.1126/science.1231143 (2013); Koblan, L. W. et al. Improving cytidine and adenine base editors by expression optimization and ancestral reconstruction. *Nat Biotechnol*, doi:10.1038/nbt.4172 (2018); Zafra, M. P. et al. Optimized base editors enable efficient editing in cells, organoids and mice. *Nat Biotechnol*, doi:10.1038/nbt.4194 (2018) 参照）。

30

40

【1055】

4つのABE構築物に関連して、DNAオンターゲット(図9 A、9 B)、DNAオフターゲット(図9 C、9 D)およびRNAオフターゲット(図9 E)の塩基編集頻度を評価したが、これらは全てコドン最適化したCas9(D10A)を含む：i) 単一C末端BP-SV40 NLSを有するABE

50

7.10 ; ii) ABE7.10の5' TadA野生型部分を欠くmonoABE7.10 ; iii) コドン最適化したTadA領域及び2つのBPNLS配列を含むABE_{max} ; およびiv) ABE_{max}のようにTadAコドン最適化を有するが単一のC末端BP-SV40 NLSを含むABE_{max}(-BPNLS)。

【1056】

4つ全ての構築物は、著しく類似したオンターゲット編集効率を示し、NLSアーキテクチャとTadAコドン最適化はオンターゲット編集効率について決定的役割は有さないことを示した(図9 A、9 B)。オフターゲットプロファイルも非常に類似していたが、ABE_{max}は、ABE7.10と比較した場合、1部位で有意に大きなDNAオフターゲット編集を示した($p=0.00027$ 、ステューデントの両側T検定)(図9 C、9 D)。ABE_{max}(-NBPNLS)は、ABE7.10より1.6倍大きなRNAオフターゲット編集の平均頻度を示した(図9 E)。

10

【1057】

【実施例3：標的化範囲を拡大した優れたアデニン塩基エディター】

ABEは、進化された(Evolved)大腸菌tRNAARG修飾酵素であるTadAが、触媒的に減損されたCas9蛋白質(D10AニッカーゼCas9、nCas9)に共有結合的に融合されたものを含む分子機械である(図7 Aおよび7 B)。従来のアデニン塩基エディターの制限を克服するために、抗生物質選択を生き残るために3つ同時のA・T→G・C復帰編集を行わなければならないABEを設計することによって、細菌選択システムのストリンジェンシーを増加させた。以前のABE進化では、TadAライブラリーは、エラーブローンPCRを介して作成された。対照的に、TadAの各位置に20個全ての正準的アミノ酸の置換を含むTadAアリの合成ライブラリーを利用し、ライブラリー構成員当たり1~2ヌクレオチド置換変異の平均頻度とした。この化学ライブラリーは、エラーブローンPCR技術で達成できるよりも大きな配列空間へのアクセスを可能にした。

20

【1058】

約300クローンを単離し、配列を決定した。得られた配列データから、高頻度に濃縮(enriched)されたTadA*内の8つの変異を同定した(表7および9)。同定された8つのアミノ酸変異のうち6つはコドン当たり少なくとも2つの核酸塩基変化を必要としたが、これは以前のTadAエラーブローンライブラリーでは観察されなかったことである。濃縮された突然変異のうちの2つはアデニン脱アミノ化の活性部位の近位残基を変化させる(I76およびV82)(図7 C)。TadA*7.10のC末端 α ヘリックスにおける以前に報告された四つの変異に加えて、二つの新しい変異が同じ α ヘリックス内に観察された(Y147RとQ154R)(図7 C)。この高度に変異された α ヘリックスは、切り詰められると塩基編集効率が実質的に低下したため、ロバストな産物形成のために必要なものである(図10 Aおよび図10 B)。

30

【1059】

哺乳類細胞におけるTadA*バリエーションの活性を試験するために、BEコドン最適化およびNLS配向を、最も有利なオンターゲットおよびオフターゲットプロファイルで利用した(実施例2 ; 図9A~9E参照)。8つの濃縮されたTadA*変異を種々の組み合わせでABE7.10に組み込み、40の新しいABE8バリエーションを得た(表7と9)。ABEのTadA領域が不活性(野生型)および活性(進化型)TadA*プロトマーのヘテロ二量体融合体であるか、または工学操作されたTadA*の単一プロトマーであるABE8構築物を作製し、これは約500塩基対小さいエディターを生じた。これらの構築物バリエーションは、それぞれABE8.x-dおよびABE8.x-mと呼ばれる(表7と9)。

40

【1060】

最初に、これら40の構築物について、正準的な20 nt *S. pyogenes* プロトスペーサー内の2~20の範囲の位置(NGG PAM=位置21、22、23)に標的A塩基を含む8つのゲノム部位に渡り、ABE7.10と比較したオンターゲットDNA編集効率を評価した(図11)。N末端野生型TadA構築物は、ABE8を用いたロバストなDNA編集にとって必要でなかった。実際、N末端野生型TadAを含む構築物(ABE8.x-d)は、その経路化された構造(ABE8.x-m)と比較して、編集ウィンドウ選好性、全DNA編集結果、およびインデル頻度のいずれに関しても同様に機能する(図7 D、図11、図12)。構築物内TadA(wt):TadA*8二量体化はABE8活性に必要でないかもしれないが、それはトランス(in trans) TadA*8:TadA*8二

50

量体化がABE8発現塩基エディター間で起こる可能性を排除しない。

【1061】

試験されたすべてのサイトで、ABE8はABE7.10と比較して、プロトスペーサにおける正準位置 (A 5~A 7) で約1.5倍高く、非正準位置 (A 3~A 4、A 8~A 10) で約3.2倍高い編集をもたらした(図13)。倍数の違いは、標的の配列、標的ウィンドウ内の「A」の位置、およびABE8構築物のアイデンティティに応じて異なる(図7 D、図11、図13)。全体として、試験した全てのサイトの全ての位置にわたる編集の中央値変化は、ABE7.10と比較して1.94倍であった (範囲1.34~4.49)。

【1062】

次に、40の構築物の大きなABE8プールから、より詳細に評価するためにABE8構築物のサブセット (ABE8.8-m、ABE8.13-m、ABE8.17-m、ABE8.20-m、ABE8.8-d、ABE8.13-d、ABE8.17-dおよびABE8.20-d) を選択した。これらの構築物は、階層的クラスタリング分析を介して決定された、8ゲノム部位間の編集性能における明確な差異を有するABE8を表す(図14)。これらのABE8は全て、試験した全てのゲノム部位でABE7.10より有意に優れており (p値=0.0006871、両側Wilcoxon順位和検定)、ABE8誘導進化戦略から同定された変異の様々な組み合わせを含んでいる(図15および図16)。

【1063】

非NGG PAMを認識するABEバリエーションが報告されているが、これら構築物の編集効率は、NGG PAM配列を標的とする*S. pyogenes* Cas9で観察された結果と比較した場合、多くの例で低下している (例えば、Huang, T. P. et al. Circularly permuted and PAM-modified Cas9 variants broaden the targeting scope of base editors. *Nat Biotechnol* 37, 626-631, doi:10.1038/s41587-019-0134-y (2019); Hua, K. et al., Expanding the base editing scope in rice by using Cas9 variants. *Plant Biotechnol J*, doi:10.1111/pbi.12993 (2018); Yang, L. et al., Increasing targeting scope of adenosine base editors in mouse and rat embryos through fusion of TadA deaminase with Cas9 variants. *Protein Cell* 9, 814-819, doi:10.1007/s13238-018-0568-x (2018)参照)。進化されたデアミナーゼが、非NGG PAMを有する標的部位での編集効率も増加させるかどうかを決定するために、*S. pyogenes* Cas9を、遺伝子操作された*S. py.*バリエーションNG-Cas9 (PAM :NG) (Nishimasu, H. et al. Engineered CRISPR-Cas9 nuclease with expanded targeting space. *Science* (2018)) または*Staphylococcus aureus* Cas9 (SaCas9、PAM: NNGRRT) (Ran, F. A. et al. In vivo genome editing using *Staphylococcus aureus* Cas9. *Nature* 520, 186-191, doi:10.1038/nature14299 (2015)) で置き換えたABE8エディターを作成した。ABE8バリエーションをABE7.10と比較した場合、SpCas9-NG (NG-ABE8.x-m/d) および SaCas9 (Sa-ABE8.x-m/d) の両方について、A・T→G・C編集頻度においてそれぞれ1.6倍と2.0倍の中央値増加が観察された (図8A、8B、および17~20)。SpCas9-ABE8と同様に、非NGG PAMバリエーションについてABE7.10とABE8構築物の間の編集効率における最大の違いは、編集ウィンドウにおける好まれる位置の周辺に位置する標的A位置で観察される (*S. pyogenes*: 位置4~8; *S. aureus*: 位置6~13; Rees, H. A. & Liu, D. R., Base editing: precision chemistry on the genome and transcriptome of living cells. *Nat Rev Genet* 19, 770-788, doi:10.1038/s41576-018-0059-1 (2018)も参照)。非NGG PAMを利用するABE8オーソログは、効率的なA塩基編集のための標的範囲を拡大する。

【1064】

インデル形成を最小にすることが必要となる用途のために、Cas9の触媒的に減損されたD10Aニッカーゼバリエーションを触媒的に「死んだ」バージョンのCas9 (D10A+H840A) (Jinek, M. et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816-821, doi:10.1126/science.1225829 (2012)) で置き換えることの効果を、コアとなる8つのABE8構築物 (「dC9-ABE8.x-m/d」) において調査した。ABEにおいてニッカーゼを死んだCas9で置き換えることによ

10

20

30

40

50

り、ABE7.10と比較してdC9-ABE8バリエーションでは、有意に高い (2.1倍) オンターゲットDNA編集効率を維持しながらインデル頻度の90%低下が観察された (図8 C、21、22、23 A、および23 B)。バックグラウンドを超えるインデルは観察されるが、試験した部位での頻度はわずか0.3~0.8%であった。幸いなことに、dC9-ABE8バリエーションは、標準的ABE8と比較して、オンターゲットDNA編集効率の中央値14%の低下しか有さない。

【1065】

オンターゲット遺伝子座における望ましくないABE媒介性ゲノム編集のもう1つの種類は、ABE依存性シトシン→ウラシル (C・G→T・A) 変換である (Grunewald, J. et al., CRISPR DNA base editors with reduced RNA off-target and self-editing activities. *Nat Biotechnol* 37, 1041-1048, doi:10.1038/s41587-019-0236-6 (2019); Lee, C. et al. CRISPR-Pass: Gene Rescue of Nonsense Mutations Using Adenine Base Editors. *Mol Ther* 27, 1364-1371, doi:10.1016/j.ymthe.2019.05.013 (2019)参照)。試験された8つの標的部位において、C→Tへ編集の95パーセンタイルは、ABE8バリエーションで0.45%、ABE7.10-dまたは-mで0.15%と測定され、ABEによるオンターゲットシトシン脱アミノ化は起こりうるがその頻度は一般的に非常に低いことが示された (図24)。まとめると、これらのデータは、ABE8が、他のしばしば望ましくない副産物と比較してA→G変換に対する高い特異性を保持していることを示す。

【1066】

[実施例4: ABE8構築物によるDNAオンターゲット及びsgRNA依存性DNAオフターゲット編集はDNAに対する特異性を向上させる]

すべての塩基エディターと同様に、ABE8は、ゲノムやトランスクリプトームのオフターゲット遺伝子座に作用する可能性がある (例えば、Gaudelli, N. M. et al. Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage. *Nature* 551, 464-471, doi:10.1038/nature24644 (2017); Komor, A. C., et al., Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533, 420-424, doi:10.1038/nature17946 (2016); Grunewald, J. et al. CRISPR DNA base editors with reduced RNA off-target and self-editing activities. *Nat Biotechnol* 37, 1041-1048, doi:10.1038/s41587-019-0236-6 (2019); Rees, H. A., et al., Analysis and minimization of cellular RNA editing by DNA adenine base editors. *Sci Adv* 5, eaax5717, doi:10.1126/sciadv.aax5717 (2019); Rees, H. A. et al. Improving the DNA specificity and applicability of base editing through protein engineering and protein delivery. *Nat Commun* 8, 15790, doi:10.1038/ncomms15790 (2017); Jin, S. et al. Cytosine, but not adenine, base editors induce genome-wide off-target mutations in rice. *Science* 364, 292-295, doi:10.1126/science.aaw7166 (2019); Zuo, E. et al. Cytosine base editor generates substantial off-target single-nucleotide variants in mouse embryos. *Science* 364, 289-292, doi:10.1126/science.aav9973 (2019); Lee, H. K., et al., Cytosine but not adenine base editor generates mutations in mice. *Biorxiv*, doi:https://doi.org/10.1101/731927 (2019); Grunewald, J. et al. Transcriptome-wide off-target RNA editing induced by CRISPR-guided DNA base editors. *Nature* 569, 433-437, doi:10.1038/s41586-019-1161-z (2019); Zhou, C. et al. Off-target RNA mutation induced by DNA base editing and its elimination by mutagenesis. *Nature* 571, 275-278, doi:10.1038/s41586-019-1314-0 (2019)参照)。

【1067】

4つのオンターゲット遺伝子座 (図25 Aおよび25 B)、およびゲノムDNAにおける12の以前に同定されたsgRNA関連オフターゲット遺伝子座 (Tsai, S. Q. et al. GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat Biotechnol* 33, 187-197, doi:10.1038/nbt.3117 (2015)) (図25 Eおよび25 F) における塩基編集を測定し、その全てがHEK293T細胞における真のCas9オフター

ゲット遺伝子座であることが確認された（図26）。オンターゲット遺伝子座での増加された活性から予想されるように、ABE8構築物は、ABE7.10よりも3～6倍大きいDNAオフターゲット編集頻度を示す。これは、ABE8構築物の使用についての注意事項であるが、sgRNAの慎重な選択および分析によりsgRNA依存性オフターゲット編集を実質的に減少させることができる（Tsai, S. Q. et al. GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat Biotechnol* 33, 187-197, doi:10.1038/nbt.3117 (2015); Yeh, W. H., et al., In vivo base editing of post-mitotic sensory cells. *Nat Commun* 9, 2184, doi:10.1038/s41467-018-04580-3 (2018)参照）。無差別的なsgRNAの使用を必要とするアプリケーションについては、ABE8.17mのTadAドメインへのDNAおよびRNA特異性増強V106W変異（Rees, H. A., Wilson, C., Doman, J. L. & Liu, D. R. Analysis and minimization of cellular RNA editing by DNA adenine base editors. *Sci Adv* 5, eaax5717, doi:10.1126/sciadv.aax5717 (2019)）の設置が、ABE7.10のものを超えるオンターゲット編集のレベルを維持しながらDNAオフターゲット編集を2.6倍減少させることができる（図25 C、25 D、25 Gおよび25 H）。

【1068】

ABE8のsgRNA非依存性オフターゲット活性を測定するために、ABEで処理されたHEK 293T細胞において細胞RNAの標的化増幅とハイスループットシーケンシングを行った（Gaudelli, N. M. et al. Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage. *Nature* 551, 464-471, doi:10.1038/nature24644 (2017); Rees, H. A., et al., Analysis and minimization of cellular RNA editing by DNA adenine base editors. *Sci Adv* 5, eaax5717, doi:10.1126/sciadv.aax5717 (2019)参照）。このアッセイでは、ABE8は、ABE7.10と比較して、細胞RNAアデノシン脱アミノ化の平均頻度が2.3～5.3倍高かった（図25 A）。

【1069】

目的外RNAオフターゲット編集を緩和するために、以前に発表された変異（Grunewald, J. et al. CRISPR DNA base editors with reduced RNA off-target and self-editing activities. *Nat Biotechnol* 37, 1041-1048, doi:10.1038/s41587-019-0236-6 (2019); Rees, H. A., et al., Analysis and minimization of cellular RNA editing by DNA adenine base editors. *Sci Adv* 5, eaax5717, doi:10.1126/sciadv.aax5717 (2019); Grunewald, J. et al. Transcriptome-wide off-target RNA editing induced by CRISPR-guided DNA base editors. *Nature* 569, 433-437, doi:10.1038/s41586-019-1161-z (2019); Zhou, C. et al. Off-target RNA mutation induced by DNA base editing and its elimination by mutagenesis. *Nature* 571, 275-278, doi:10.1038/s41586-019-1314-0 (2019)）をABE8.17-m中のデアミナーゼ酵素のTadA部分に導入し、オフターゲット編集頻度の減少を評価した。これらの変異は全て、異なる程度においてABE8.17-mのオンターゲット編集頻度を低減させ、V106WとF148AはABE8の減損が最も少なかった（図25 Cおよび25 D）。これらのうち、V106WのみがオフターゲットのRNAとDNA編集のレベルを実質的に低下させることができた（図25 B）。このように、ABE8へのV106W変異の組み込みが、細胞トランスクリプトームの一時的攪乱が回避されなければならない場合に、または無差別的sgRNAを用いた使用のために、適用可能である。

【1070】

【実施例5：血液学的障害の治療のためのアデニン塩基エディター】

ABE8構築物をヒト造血幹細胞（HSC）において評価した。細胞療法として患者に投与する前のHSCのex vivo操作および/または編集は、血液学的障害の治療のための有望なアプローチである。ABEはHBG1/2のプロモーター領域の-198位置にTからCへの置換を導入することが以前に示された（Gaudelli, N. M. et al. Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage. *Nature* 551, 464-471, doi:10.1038/nature24644 (2017)）。この天然に存在するアリルは、胎児ヘモグロビンの

遺伝的持続性 (HPFH: Hereditary Persistence of Fetal Hemoglobin) を生じ、成人期まで γ グロビンレベルの増加をもたらし、これは鎌状赤血球症および β サラセミアで見られる β グロビンの欠損を緩和することができる (Wienert, B. et al. KLF1 drives the expression of fetal hemoglobin in British HPFH. *Blood* 130, 803-807, doi:10.1182/blood-2017-02-767400 (2017))。HPFH表現型を再現し ABE8 の臨床的関連性を評価する目的で、CD34+造血幹細胞を2人のドナーから単離し、ABE8エディターをコードする mRNA と、標的 A をプロトスペーサー内の7位に配置する末端修飾 sgRNA とでトランスフェクトした。

【1071】

-198 HBG1/2 プロモーター標的部位での平均 ABE8 編集効率は、初期時点 (48時間) ではいずれの ABE7.10 構築物よりも2~3倍高く、後期時点 (144時間) ではいずれの ABE7.10 よりも1.3~2倍高かった (図27 A、図28、図29)。これらの速度論的区別は、細胞療法の投与前の細胞培養を最小限に抑えなければならない *ex vivo* 療法にとって臨床的に重要である。

【1072】

次に、ABE処理と赤血球分化後に産生される γ (Gamma) グロビン蛋白質の量を UPLC により定量した (図30~50)。模擬 (Mock) 処理細胞と比較した場合、ABE8処理群由来の赤血球における γ (Gamma) グロビン/ α (Alpha) グロビン発現の平均3.5倍の増加が観察され、ABE7.10 m/d で達成されたレベルと ABE8.13-d を比較した場合には、約1.4倍の増加が観察された (図27 B)。鎌状赤血球症の症状を改善するには $\geq 20\%$ の HbF が必要であると予測されており、 β サラセミア患者ではさらに高い最低濃度を要する可能性が高い (例えば Canver, M. C. & Orkin, S. H. Customizing the genome as therapy for the beta-hemoglobinopathies. *Blood* 127, 2536-2545, doi:10.1182/blood-2016-01-678128 (2016); Fitzhugh, C. D. et al. At least 20% donor myeloid chimerism is necessary to reverse the sickle phenotype after allogeneic HSCT. *Blood* 130, 1946-1948, doi:10.1182/blood-2017-03-772392 (2017) 参照)。ABE8処理後に観察された γ グロビンレベルはこの閾値を超えていた。

【1073】

全体として、ABE8は γ グロビン遺伝子 HBG 1 と HBG 2 のプロモーターにおける胎児ヘモグロビン (HPFH) アリルの自然発生遺伝的持続性を再現し、ヒト CD34+細胞培養における最大60%の編集効率と、分化赤血球における γ グロビン発現の対応するアップレギュレーションを達成した。

【1074】

[実施例6: 鎌状赤血球症およびベータサラセミアの治療のための相補的塩基編集アプローチ]

鎌状赤血球症 (SCD) およびベータサラセミアは、 β グロビンの産生および機能の障害であり、重度の貧血および多数の臓器系にわたる重大な疾患合併症を引き起こす。胎児ヘモグロビン (HbF) のアップレギュレーションまたは β グロビン遺伝子の補正により操作された造血幹細胞の自己移植は、 β 異常ヘモグロビン症患者の疾患負荷を減少させる可能性がある。塩基編集は最近開発された技術であり、二本鎖 DNA 切断を導入することなくゲノムの正確な改変を可能にする。

【1075】

HbF抑制を解除する改変を同定するために、シトシンおよびアデニン塩基エディター (ABE) を用いて、 γ グロビン遺伝子プロモーターを包括的にスクリーニングした。HbFを有意にアップレギュレートする3領域を同定したが、最も効果的なヌクレオチド残基変換は、胎児ヘモグロビンの遺伝的持続性 (HPFH) 患者で見られる天然のバリエーションにより支持された。HBG 1 および HBG 2 プロモーター内の重要な調節モチーフにおけるヌクレオチド変換後に HbF レベルを有意に増加させる ABE が開発された。CD34+造血幹細胞および前駆細胞 (HSPC) を臨床スケールで精製し、自己再生能を保持するように設計されたプロセスを用いて編集した。異なる ABE を用いて、2つの独立した部位での編集は94%に達し

10

20

30

40

50

、UPLCによれば63%までのガンマグロビンが得られた(図51 A~51 E)。この観察されたHbFレベルは、HPFHの臨床的観察と、より高いHbF用量をより軽度の疾患と結びつける非介入療法とに基づけば、SCDとベータサラセミア患者の大部分に保護を与えるはずである(Ngo et al., 2011 Brit J Hem; Musallam et al., 2012 Blood)。

【1076】

SCDのGlu6Val変異を直接修正することが、SCD集団のためにデザインされた遺伝子治療の最近の目標となってきた。現在の塩基編集技術では、鎌状βグロビンにおけるA-T転換で生じるような変異はまだ変換できない；しかしながら、バリンの反対側の鎖のアデニン残基を認識し編集するようなABEバリエーションは設計されている。この結果、バリンがアラニンに変換され、Hb G-Makassarとして知られる天然の変異体が産生される。この位置にアラニンを有するβグロビンはポリマー形成に寄与せず、Hb G-Makassarを有する患者は正常な血液学的パラメータおよび赤血球形態を呈する。これらのABEバリエーションで編集されたSCD患者線維芽細胞は、標的アデニンの70%に至る変換を達成した(図52 A)。次いで、健康なドナーからのCD34細胞をリードABEバリエーションで編集し、これは、編集ウィンドウ内に存在しSCD変異編集についてのプロキシとして働く隣接プロリンにおける同義突然変異を標的とした。平均編集頻度は40%であった(図52 B)。同種移植の場面でこれらのレベルで記録されたドナー骨髄キメラ現象は、鎌状赤血球表現型を逆転させるのに必要な20%を超えている(Fitzhugh et al, 2017 Blood)。

【1077】

【実施例7：材料と方法】

一般的なメソッド：

全てのクローニングは、USER酵素(New England Biolabs) クローニング法(Geu-Floures et al., USER fusion: a rapid and efficient method for simultaneous fusion and cloning of multiple PCR products. Nucleic Acids Res 35, e55, doi:10.1093/nar/gkm106 (2007)参照)を介して行い、PCR増幅のための鋳型は、細菌または哺乳動物のコドン最適化遺伝子断片として購入した(GeneArt)。作成したベクターをMach T1^Rコンピテント細胞(Thermo Fisher Scientific)に形質転換し、長期保存のために-80℃で維持した。この研究で使用されたすべてのプライマーは、Integrated DNA Technologiesから購入され、PCRSは、Phusion U DNA Polymerase Green MultiPlex PCR Master Mix (ThermoFisher) またはQ5 Hot Start High-Fidelity 2x Master Mix (New England Biolabs) のいずれかを用いて実施された。本研究で用いたすべてのプラスミドは、エンドトキシン除去手順を含むZymoPURE Plasmid Midiprep (Zymo Research Corporation)を用いて、50 mLのMach 1培養物から新たに調製された。分子生物学グレードのHyclone水(GE Healthcare Life Sciences)をすべてのアッセイ、トランスフェクション、およびPCR反応に用いて、DNase活性の排除を確実にした。

【1078】

Hek293T哺乳類細胞トランスフェクションに使用されるsgRNAのアミノ酸配列は、以下の表15に提供される。20 nt標的プロトスペーサーは太字で示してある。ヒトU6プロモーターは転写開始部位において「G」を好むことが確立されているので、標的DNA配列が「G」で始まっていない場合にはプライマーの5'末端に「G」を加えた(Cong, L. et al., Multiplex genome engineering using CRISPR/Cas systems. Science 339, 819-823, doi:10.1126/science.1231143 (2013)参照)。前に記述されたpFYF sgRNAプラスミドをPCR増幅のための鋳型として使用した。

【1079】

表15：Hek293T哺乳類細胞トランスフェクションに用いられるsgRNAの配列。

10

20

30

40

50

【表 15】

サイト	RNA プロトスペーサー配列	Cas9 足場	PAM
1	GAACACAAAGCAUAGACUGC	<i>S. pyogenes</i>	NGG
2	GGGAAAGACCCAGCAUCCGU	<i>S. pyogenes</i>	NGG
3	GCUCCCAUCACAUAACCGG	<i>S. pyogenes</i>	NGG
4	GGUGAGUGAGUGUGUGCGUG	<i>S. pyogenes</i>	NGG
5	GGCUUCAGGUUCUAAAUGAG	<i>S. pyogenes</i>	NGG
6	GCAGAGAGUCGCCGUCUCCA	<i>S. pyogenes</i>	NGG
7	GUGUAAGACCUCAAAAGCAC	<i>S. pyogenes</i>	NGG
8	GAUGAGAAGGAGAAGUUCUU	<i>S. pyogenes</i>	NGG
9	GAGGACAAAGUACAAACGGC	<i>S. pyogenes</i>	AGA
10	GCCACCACAGGGAAGCUGGG	<i>S. pyogenes</i>	TGA
11	GCUCUCAGGCCUGUCCGCA	<i>S. pyogenes</i>	CGT
12	GAGCAAAUACCAGAGAUAAAG	<i>S. pyogenes</i>	AGA
13	GAUCAGGAAAUAAGAGCCACA	<i>S. pyogenes</i>	GGC
14	GCCCAUCCCUGAGUCCAGCG	<i>S. pyogenes</i>	AGC
15	GAACACGAAGACAUCUGAAGGUA	<i>S. aureus</i>	TTGAAT
16	GAUUUACAGCCUGGCCUUUGGGG	<i>S. aureus</i>	TCGGGT
17	GGAGAGAAAGAGAAGUUGAUUG	<i>S. aureus</i>	ATGGGT
18	GAGGGUGAGGGAUGAGAUAAUG	<i>S. aureus</i>	ATGAGT
19	GGUGGAGGAGGGUGCAUGGGGU	<i>S. aureus</i>	CAGAAT
20	GCUGUUGCAUGAGGAAAGGGAC	<i>S. aureus</i>	TAGAGT
HEK2	GAACACAAAGCAUAGACUGC	<i>S. pyogenes</i>	CGG
HEK3	GGCCCAGACUGAGCACGUGA	<i>S. pyogenes</i>	TGG
HEK4	GGCACUGCGGCUGGAGGUGG	<i>S. pyogenes</i>	GGG
LDLR	GCAGAGCACUGGAAUUCGUCA	<i>S. pyogenes</i>	GGG

【1080】

sgRNA足場配列は以下の通りである。

S. pyogenes:

GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAA
AGUGGCACCGAGUCGGUGC

S. aureus: GUUUUAGUACUCUGUAAUGAAAAUUACAGAAUCUACUAAAACAAGGCAA
AAUGCCGUGUUUAUCUCGUCAACUUGUUGGCGAGA

【1081】

【誘導進化のためのインプット細菌TadA*ライブラリーの生成】

TadA*8.0ライブラリーは、TadA*7.10オープンリーディングフレームの各アミノ酸位置において20アミノ酸すべてをコードするように設計された (Gaudelli, N. M. et al., Programmable base editing of A*T to G*C in genomic DNA without DNA cleavag

10

20

30

40

50

e. Nature 551, 464-471, doi:10.1038/nature24644 (2017))。各TadA*8.0ライブラリー構成員は、約1~2個の新しいコード変異を含み、化学的に合成されたものをRanomics Inc. (カナダ国トロント) から購入した。TadA*8.0ライブラリーをPhusion U Green MultiPlex PCR Master MixでPCR増幅し、ABE誘導進化のために最適化された細菌ベクターにUSER法によりアセンブルした (Gaudelli, N. M. et al., Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage. Nature 551, 464-471, doi:10.1038/nature24644 (2017))。

【1082】

[TadAバリエーションの細菌進化]

TadA*8ライブラリーを含むABEの誘導進化は、以下の変更を加えて、以前に記述されたように行った (Gaudelli, N. M. et al., Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage. Nature 551, 464-471, doi:10.1038/nature24644 (2017)) : i) 大腸菌10ベータ (ニューイングランド・バイオラボ) を進化宿主として使用した ; ii) カナマイシン上の生存を3つの遺伝子不活性化成分の補正に依存させた (例えば、生存には、カナマイシンにおける2つの停止変異と1つの活性部位変異の復帰変異を必要とした)。カナマイシン耐性遺伝子配列は、ABE8進化のための選択変異を含む。10ベータ宿主細胞中で選択プラスミドおよびエディターを一晩共培養した後、ライブラリー培養物を、プラスミド維持抗生物質および増加する濃度の選択抗生物質カナマイシン (64~512 $\mu\text{g/mL}$) を添加した2xYT-寒天培地上に平板培養した。細菌を1日増殖させ、生存クローンのTadA*8部分を濃縮後にSanger配列決定した。次いで、目的の同定されたTadA*8突然変異を、USERアセンブリーを介して哺乳動物発現ベクターに組み込んだ。

【1083】

[一般的なHEK293TおよびRPMI-8226哺乳類培養条件]

細胞を37°Cで5% CO₂とともに培養した。HEK293T細胞 [CLBTx 013, American T type Cell Culture Collection (ATCC)] は、10% (v/v) 胎児ウシ血清 (A31606-02、Thermo Fisher Scientific社) を有するダルベッコ改変イーグル培地+グルタマックス (10566-016、Thermo Fisher Scientific社) 中で培養した。RPMI-8226 (CCL-155、ATCC) 細胞は、10% (v/v) 胎児ウシ血清 (Gibco) を含むRPMI-1640培地 (Gibco) で培養した。供給者からの受領後、細胞はマイコプラズマについて陰性であった。

【1084】

[Hek293TプラスミドトランスフェクションおよびgDNA抽出]

HEK293T細胞を、48ウェルのPoly-D-Lysine処理したBioCoatプレート (Corning) 上に35,000細胞/ウェルの密度で播種し、プレーティングの18~24時間後にトランスフェクトした。NucleoCounter NC-200 (ケモメテック) を使用して細胞を計数した。これらの細胞に、OptiMEM血清低減培地 (ThermoFisher Scientific) において総容積12.5 μL に希釈した750 ngの塩基エディターまたはヌクレアーゼ対照、250 ngのsgRNA、および10 ngのGFP-maxプラスミド (Lonza) を添加した。この溶液を11 μL のOptiMEM血清低減培地中1.5 μL のリポフェクタミン2000 (ThermoFisher) と混合し、15分間室温で放置した。次いで、全25 μL 混合物を予め播種しておいたHek293T細胞に移し、約120時間インキュベートした。インキュベーション後、培地を吸引し、細胞を250 μL の1xPBS溶液 (ThermoFisher Scientific) で2回洗浄し、新しく調製された100 μL の溶解緩衝液を添加した (100 mM Tris-HCl、pH 7.0、0.05% SDS、25 g/mL Proteinase K (Thermo Fisher Scientific))。溶解緩衝液を含むトランスフェクションプレートを37°Cで1時間インキュベートし、混合物を96ウェルPCRプレートに移し、80°Cで30分間加熱した。

【1085】

[ABE構造およびABE8構築物についてのDNAおよびRNAオフターゲット編集の分析]

HEK293T細胞を、48ウェルのポリ-D-リジンコーティングプレート (Corning) 上に16~20時間平板培養した後、抗生物質を含まないDMEM+Glutamax培地 (Thermo Fisher

10

20

30

40

50

Scientific) 中で3万細胞/ウェルの密度でリポフェクションを行った。750 ngのニッカーゼまたは塩基エディター発現プラスミドDNAを、15 µl OPTIMEM+Glutamax中で250 ngのsgRNA発現プラスミドDNAと組み合わせた。これを、ウェル当たり1.5 µlのリポフェクタミン2000と8.5 µlのOPTIMEM+Glutamaxを含む10 µlの脂質混合物と混合した。トランスフェクションの3日後に細胞を回収し、DNAまたはRNAのいずれかを回収した。DNA分析のためには、細胞を1X PBS中で一回洗浄し、次いで製造業者の指示に従って100 µlのQuickExtract (商標) 緩衝液 (Lucigen) 中で溶解した。RNAの回収には、製造業者の指示に従ってMagMAX (商標) mirVana (商標) Total RNA Isolation Kit (Thermo Fisher Scientific) をKingFisher (商標) Flex Purification Systemと共に使用した。

10

【1086】

標的化されたRNA配列決定は、ほぼ以前に記述された通りに行われた (Rees, H. A. et al., Analysis and minimization of cellular RNA editing by DNA adenine base editors. Sci Adv 5, eaax5717, doi:10.1126/sciadv.aax5717 (2019)参照)。cDNAは、製造業者の指示に従ってEZDnase (Thermo Fisher Scientific) と共にSuperScript IV One-Step RT-PCRシステムを用いて、単離されたRNAから調製した。以下のプログラムを使用した: 58°C、12分間; 98°C、2分間; その後、アンプリコンに応じて変化したPCRサイクル: CTNNB 1およびIP90のためには32サイクルの[98°C、10秒; 60°C、10秒; 72°C、30秒]、およびRSL1D1のためには35サイクルの[98°C、10秒; 58°C、10秒; 72°C、30秒]。サンプルと同時にRT無し対照を実施した。複合RT-PCRに続いて、アンプリコンをバーコード化し、上述のようにIllumina Miseqを用いて配列決定した。各アンプリコン中のフォワードプライマーの末端後の最初の塩基から始まる、各アンプリコン中の最初の125 ntを、参照配列に整列させ、各アンプリコン中の平均および最大A→I頻度の分析のために使用した (図53 Aおよび53 B)。

20

【1087】

オフターゲットDNA配列決定は、上記のようにIllumina Miseqシーケンサーを用いた配列決定のためのサンプルを調製するための二工程PCRおよびバーコード化法を用いて、下記表16に列挙される以前に公表されたプライマー (Komor, A. C. et al., Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. Nature 533, 420-424, doi:10.1038/nature17946 (2016); Rees, H. A. et al., Analysis and minimization of cellular RNA editing by DNA adenine base editors. Sci Adv 5, eaax5717, doi:10.1126/sciadv.aax5717 (2019)) を使用して行った。

30

【1088】

表16: ゲノム部位の増幅に用いたHTSプライマー:

40

50

【表 1 6】

プライマー名	配列
fwd_site_1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNCCAGCCCCATCTGTCAAAC
rev_site_1	TGGAGTTCAGACGTGTGCTCTTCCGATCTTGAATGGATTCCTTGGAACAATGA
fwd_site_2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNTGAGGGAGAGCCGTGTAGTT
rev_site_2	TGGAGTTCAGACGTGTGCTCTTCCGATCTGCCTCTCAAAGTGCTGGGAT
fwd_site_3	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNCCATCAGGCTCTCAGCTCAG
rev_site_3	TGGAGTTCAGACGTGTGCTCTTCCGATCTCTCGTGGGTTTGTGGTTGC
fwd_site_4	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNGCCCATTCCTCTTTAGCCA
rev_site_4	TGGAGTTCAGACGTGTGCTCTTCCGATCTGAGCCGTTCCCTCTTTGCTA
fwd_site_5	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNAACCTGTGTGACACTTGGCA
rev_site_5	TGGAGTTCAGACGTGTGCTCTTCCGATCTGTCTGGCCCAAGATCACACA
fwd_site_6	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNCACGGATAAAGACGCTGGGA
rev_site_6	TGGAGTTCAGACGTGTGCTCTTCCGATCTGGGGTCCCAGGTGCTGAC
fwd_site_7	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNTTGATTGTCTCCTTTGCCGC
rev_site_7	TGGAGTTCAGACGTGTGCTCTTCCGATCTTGACCCAGTGTGATAGATCAGT
fwd_site_8	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNCACCCCTTCAGTCCATGCTT
rev_site_8	TGGAGTTCAGACGTGTGCTCTTCCGATCTTCTGATGGGGAGGAACGAGT
fwd_site_9	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNCAGCTCAGCCTGAGTGTGA
rev_site_9	TGGAGTTCAGACGTGTGCTCTTCCGATCTGCCCACCCTAGTCATTGGAG
fwd_site_10	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNGTCAGAGGGACACACTGTGG
rev_site_10	TGGAGTTCAGACGTGTGCTCTTCCGATCTCACACTCACTCACCCACACA
fwd_site_11	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNTGTGTGGGTGAGTGAGTGTG
rev_site_11	TGGAGTTCAGACGTGTGCTCTTCCGATCTCACCAAGGTTACAGCCTGA
fwd_site_12	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNTGTCTCTGCCTGTAGCTGC
rev_site_12	TGGAGTTCAGACGTGTGCTCTTCCGATCTCGCTCTGGGCTTCATCTTCA
fwd_site_13	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNTGGGATTATGGGTGTGAGCC

10

20

30

40

50

プライマー名	配列	
rev_site_13	TGGAGTTCAGACGTGTGCTCTTCCGATCTTGCCTTCCTCCTCTCTCTCC	
fwd_site_14	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNTCAGACCAGATTTCGGAGAA	
rev_site_14	TGGAGTTCAGACGTGTGCTCTTCCGATCTGTTTCCAGGGGGTCC	
fwd_site_15	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNTCGCACAGCCTTAGTTCAA	
rev_site_15	TGGAGTTCAGACGTGTGCTCTTCCGATCTAACTTGAAGAGACGGCAGCA	10
fwd_site_16	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNCCCCCAGCTACAGAAAGGTC	
rev_site_16	TGGAGTTCAGACGTGTGCTCTTCCGATCTATTTCACCCGCAAAATGGCC	
fwd_site_17	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNTCAGTTCAGCCCAGGAGTAT	
rev_site_17	TGGAGTTCAGACGTGTGCTCTTCCGATCTTGTGTATGGTGAGAGGTAGGGA	
fwd_site_18	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNGTCTGAGGTCACACAGTGGG	
rev_site_18	TGGAGTTCAGACGTGTGCTCTTCCGATCTCTGAGAGCAGGGACCACATC	
fwd_site_19	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNGGAGGTGGAGAGAGGATGT	20
rev_site_19	TGGAGTTCAGACGTGTGCTCTTCCGATCTACTCTTCTGAGGTCTAGGAACCCG	
fwd_site_20	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNCCCTGTTCTTAAAGCCCACC	
rev_site_20	TGGAGTTCAGACGTGTGCTCTTCCGATCTACTCTCTGGTCTGTTTGTGGCCA	
fwd_CTNNB1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNATTGATGGAGTTGGACATG GCC	
rev_CTNNB1	TGGAGTTCAGACGTGTGCTCTCCAGCTACTTGTCTTCTGAGTGAAGG	
fwd_RSL1D1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNTTGGCTTTCCAAATCAGTGGG TC	
rev_RSL1D1	TGGAGTTCAGACGTGTGCTCTTCCGATCTCTCATAAGCTTAGACCAACAAGC	30
fwd_IP90	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNCTGGTTGACCAATCTGTGGT G	
rev_IP90	TGGAGTTCAGACGTGTGCTCTCTGCGTCTGGATCAGGTACG	
fwd_HEK293_site2_off1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNGTGTGGAGAGTGAGTAAGCC A	
rev_HEK293_site2_off1	TGGAGTTCAGACGTGTGCTCTTCCGATCTACGGTAGGATGATTTTCAGGCA	
fwd_HEK293_site2_off2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNCACAAGCAGTGTAGCTCAG G	
rev_HEK293_site2_off2	TGGAGTTCAGACGTGTGCTCTTCCGATCTTTTTTGGTACTCGAGTGTATTTCAG	
fwd_HEK293_site3_off1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNTCCCCTGTTGACCTGGAGAA	40

プライマー名	配列
rev_HEK293_site3_off1	TGGAGTTCAGACGTGTGCTCTTCCGATCTCACTGTACTTGCCCTGACCA
fwd_HEK293_site3_off2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNTTGGTGTGACAGGGAGCAA
rev_HEK293_site3_off2	TGGAGTTCAGACGTGTGCTCTTCCGATCTCTGAGATGTGGGCAGAAGGG
fwd_HEK293_site3_off3	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNTGAGAGGGAACAGAAGGGCT
rev_HEK293_site3_off3	TGGAGTTCAGACGTGTGCTCTTCCGATCTGTCCAAAGGCCCAAGAACCT
fwd_HEK293_site3_off4	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNCTTAGCACTTTGGAAGGTCG
rev_HEK293_site3_off4	TGGAGTTCAGACGTGTGCTCTTCCGATCTGCTCATCTTAATCTGCTCAGCC
fwd_HEK293_site3_off5	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNAAGGAGCAGCTCTTCCTGG
rev_HEK293_site3_off5	TGGAGTTCAGACGTGTGCTCTTCCGATCTGTCTGCACCATCTCCACAA
fwd_HEK293_site4_off1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNGGCATGGCTTCTGAGACTCA
rev_HEK293_site4_off1	TGGAGTTCAGACGTGTGCTCTTCCGATCTGTCTCCCTTGCACTCCCTGTCTTT
fwd_HEK293_site4_off2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNTTTGGCAATGGAGGCATTGG
rev_HEK293_site4_off2	TGGAGTTCAGACGTGTGCTCTTCCGATCTGAAGAGGCTGCCCATGAGAG
fwd_HEK293_site4_off3	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNGGCTGAGGCTCGAATCCTG
rev_HEK293_site4_off3	TGGAGTTCAGACGTGTGCTCTTCCGATCTCTGTGGCCTCCATATCCCTG
fwd_HEK293_site4_off4	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNTTCCACCAGAACTCAGCCC
rev_HEK293_site4_off4	TGGAGTTCAGACGTGTGCTCTTCCGATCTCCTCGGTTCCCTCCACAACAC
fwd_HEK293_site4_off5	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNCACGGGAAGGACAGGAGAAG
rev_HEK293_site4_off5	TGGAGTTCAGACGTGTGCTCTTCCGATCTGCAGGGGAGGGATAAAGCAG
fwd_HEK_site3	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNGGAAACGCCCATGCAATTAGTC
rev_HEK_site3	TGGAGTTCAGACGTGTGCTCTTCCGATCTCTTGTCACCAGTATCCCGGTG
fwd_HEK_site2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNTGAATGGATTCCCTTGAAACAATG
rev_HEK_site2	TGGAGTTCAGACGTGTGCTCTTCCGATCTCCAGCCCCATCTGTCAAACCT
fwd_HEK_site4	TGGAGTTCAGACGTGTGCTCTTCCGATCTTCCTTTCAACCCGAACGGAG
rev_HEK_site4	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNGCTGGTCTTCTTTCCCTCC
fwd_LDLR	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNGCCCTGCTTCTTTTCTCTGTGT

10

20

30

40

50

プライマー名	配列
rev_LDLR	TGGAGTTCAGACGTGTGCTCTTCCGATCTACCATTAACGCAGCCAACTTCA
fwd_TRAC	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCATGAGGTCTATGGACTTCAAGAGCAA
Rev_TRAC	TGGAGTTCAGACGTGTGCTCTTCCGATCTCATCATTGACCAGAGCTCTGGGCAGAA
fwd_CBLB	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCACTTACCAGCATTACTTCCTAAACC
Rev_CBLB	TGGAGTTCAGACGTGTGCTCTTCCGATCTATGGGCTCCACTTTTCAGCTCTGTAA
fwd_CD7	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCAGTTCAGGCACATGTAGGAGGGA
Rev_CD7	TGGAGTTCAGACGTGTGCTCTTCCGATCTACCGCCTGCAGCTGTCGGACACTGGCA
fwd_B2M	ACACTCTTTCCCTACACGACGCTCTTCCGATCTAAAAGATGAGTATGCCTGCCGTG
Rev_B2M	TGGAGTTCAGACGTGTGCTCTTCCGATCTCAGATTGTTTATATCAGATGGGATGGG
fwd_CIITA	ACACTCTTTCCCTACACGACGCTCTTCCGATCTATGCAAGTTTGGTCCTGAGCCCTCCC
Rev_CIITA	TGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTGGGTTCCTGCGCTCTGCA
fwd_PDCD1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCCAGGGACTGAGGGTGGAAGGTCC
Rev_PDCD1	TGGAGTTCAGACGTGTGCTCTTCCGATCTACCTCCGCCTGAGCAGTGGAGAA

【1089】

〔CD34+細胞において使用されるABEエディターのためのmRNA産生〕

dT7プロモーターに続いて5' UTR、Kozak配列、ORF、および3' UTRをコードするプラスミド中に、エディターをクローニングした。dT7プロモーターは、環状プラスミドからの転写を妨げる、T7プロモーター内の不活性化点突然変異を有する。このプラスミドをPCR反応（Q5ホットスタート2Xマスターミックス）の鋳型にし、そこでは、フォワードプライマーがT7プロモーター内のSNPを修正し、リバースプライマーは3' UTRに120Aターミナルを付加した。得られたPCR産物をZymo Research 25 µg DCCカラムで精製し、その後のin vitro転写においてmRNAテンプレートとして使用した。使用マニュアルに従ってNEB HiScribe High-Yield Kitを使用した。ただしウリジンをN1-メチル-プソイドウリジンに完全置換し、CleanCap AG (Trilink) で転写同時キャッピングを行った。反応物クリーニングアップは塩化リチウム沈殿により行った。増幅に使用されたプライマーは表17に見出すことができる。

【1090】

表17：ABE8 T7のin vitro転写反応に用いられたプライマー

10

20

30

40

50

【表 17】

名称	配列
fwd_IVT	TCGAGCTCGGTACCTAATACGACTCAC
rev_IVT	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT CTTCCTACTCAGGCTTTATTCAAAGACCA

10

【 1 0 9 1 】

[CD34+細胞調製]

動員化 (mobilized) 末梢血を得て、ヒトCD34+HSPCを濃縮した (HemaCare、M001F-GCSF/MOZ-2)。CD34+細胞を解凍し、エレクトロポレーションの48時間前に、1% Glutamax (Gibco)、100 ng/mLのTPO (Peprotech)、SCF (Peprotech) およびFlt-3 (Peprotech) を含有するX-VIVO 10 (Lonza) に入れた。

【1092】

[CD34+細胞の電口ポレーション]

解凍48時間後、細胞をスピンドウンしてX-VIVO 10培地を除去し、0.1% HSA (Akron Biotechnologies) を有するMaxCyte緩衝液(HyClone) で洗浄した。次いで、細胞を冷MaxCyte緩衝液中で125万細胞/mLに再懸濁し、複数の20 μ Lアリコートに分割した。次いで、ABE mRNA (0.15 μ M) および-198HBG1/2 sgRNA (4.05 μ M) を、実験条件に従って分注し、MaxCyte緩衝液中で総量5 μ Lまでに上昇させた。20 μ Lの細胞を3つのグループに分けて5 μ LのRNA混合物に添加し、エレクトロポレーションのためにOC25x3 MaxCyteキュベットの各チャンバーに充填した。電荷を受けた後、チャンバーから25 μ Lを回収し、24ウェル未処理培養プレート中のウェルの中央に置いた。インキュベーター (37°C、5% CO₂) 内で20分間細胞を回復させた。20分の回復後、1% Glutamax、100 ng/mLのTPO、SCFおよびFlt-3を含有するX-VIVO 10を、100万細胞/mLの濃度で細胞に添加した。次いで細胞を、インキュベーター (37°C、5% CO₂) 内で48時間さらに回復させた。

20

【 1 0 9 3 】

「ABEエレクトロポレーション後の赤血球分化」

48時間のエレクトロポレーション後休止に続いて(培養0日目)、細胞をスピンドウンし、5%ヒト血清、330 µg/mLトランスフェリン(Sigma)、10 µg/mLヒトインスリン(Sigma)、2 U/mLヘパリンナトリウム(Sigma)、3 U/mL EPO (Peprotech)、100 ng/mL SCF (Peprotech)、5 µg/mL IL 3および50 µMヒドロコルチゾン(Sigma)を含有する「第1相」IMDM培地(ATCC)に2万細胞/mLで移した。培養4日目に、細胞に4倍容量の同じ培地を与えた。7日目に、細胞をスピンドウンし、5%ヒト血清(Sigma)、330 µg/mLトランスフェリン、10 µg/mLヒトインスリン、2 U/mLヘパリンナトリウム、3 U/mL EPO、および100 ng/mL SCFを含有する「第2相」IMDM培地に20万細胞/mLで移した。11日目に、細胞をスピンドウンし、5%ヒト血清、330 µg/mLトランスフェリン、10 µg/mLヒトインスリン、2 U/mLヘパリンナトリウム、および3 U/mL EPOを含む「第3相」IMDM培地に100万細胞/mLで移した。14日目に、細胞をスピンドウンし、11日目と同じ培地中に再懸濁したが、ただし500万細胞/mLとした。18日目に、分化した赤血球を50万個の細胞アリコートで収集し、500 µLのDPBS(Gibco)中で一回洗浄し、UHP-LC処理の前に-80°Cで24時間凍結した。

30

【 1 0 9 4 】

「UHPLC分析用赤血球サンプルの調製」

凍結赤血球ペレットを室温で解凍した。ペレットをACK溶解緩衝液で最終濃度 5×10^4

40

50

細胞/ μL まで希釈した。試料をピペットで混合し、室温で5分間インキュベートした。次いで、試料を -80°C で5分間凍結し、解凍し、ピペットで混合した後に $6,700\text{ g}$ で10分間遠心分離した。上清を注意深く除去し（細胞デブリペレットを乱すことなく）、新しいプレートに移し、UHPLC分析のために超純水中で 5×10^3 細胞/ μL まで希釈した。

【1095】

〔超高速液体クロマトグラフィー (UHPLC) 分析〕

二元ポンプとUV検出器と共に構成されたUHPLCシステム上で（サーモ・フィッシャー・サイエンティフィック、ヴァンキッシュ・ホライズン）、グロビン鎖の逆相分離を行った。固定相は、AQUITYペプチドBEH C18 VanGuardプレカラム（ $2.1 \times 5\text{ mm}$ 、 $1.7\text{ }\mu\text{m}$ ビーズ、 $300\text{ }\text{\AA}$ 細孔）の後のAQUITYペプチドBEH C18カラム（ $2.1 \times 150\text{ mm}$ 、 $1.7\text{ }\mu\text{m}$ ビーズ、 $300\text{ }\text{\AA}$ 細孔）（どちらもWaters社）からなり、カラム温度は 60°C とした。溶離は、水中の 0.1% トリフルオロ酢酸（TFA）（A）とアセトニトリル中の 0.08% TFA（B）を用い、流速 0.25 mL/min で行った。グロビン鎖の分離は、 $0\sim 10$ 分で $40\sim 52\%$ B； $10\sim 10.5$ 分で $52\sim 40\%$ B； 12 分まで 40% Bという線形勾配を用いて達成された。試料注入体積を $10\text{ }\mu\text{L}$ とし、データレート 5 Hz で波長 220 nm のUVスペクトルを分析中に収集した。グロビン鎖のアイデンティティは、ヘモグロビン標準のLC/MS分析により確認した。

【1096】

〔CD34+細胞のゲノムDNA抽出〕

ABEエレクトロポレーション後（48時間後）、細胞のアリコートに 1% Glutamax（Gibco）、 100 ng/mL のTPO（Peprtech）、SCF（Peprtech）及びFlt-3（Peprtech）を含むX-VIVO 10培地（Lonza）中で培養した。48時間および144時間の後培養に続き、10万個の細胞を収集し、スピンドウンした。 $50\text{ }\mu\text{L}$ のQuick Extract（Lucigen）を細胞ペレットに添加し、細胞混合物を96ウェルPCRプレート（Bio-Rad）に移した。溶解物を 65°C で15分間加熱し、次いで 98°C で10分間加熱した。細胞溶解物は -20°C で保存した。

【1097】

〔配列〕

以下の配列において、小文字はカナマイシン耐性プロモーター領域を示し、太字の配列は標的化された不活性化部分（Q4*及びW15*）を示し、斜体の配列はカナマイシン耐性遺伝子の標的化された不活性化部位（D208N）を示し、下線の配列はPAM配列を示す。

不活化されたカナマイシン耐性遺伝子：

10

20

30

40

50

ccggaattgccagctggggcgccctctggtaagggttgggaagccctgcaaagtaaactggatgg
ctttcttgccgccaaggatctgatggcgaggggatcaagatctgatcaagagacaggatgagg
atcctttcgc**ATGATCGAATAAGAT**GGATTGCACGCAGGTTCTCCGGCC**GCTTAGGTGGAGCGC**
CTATTCGGCTTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGT
CAGCGCAGGGGCGCCCGGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTGCA
GGACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGCGCAGCTGTGCTCGAC
GTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGT
CATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATAC
GCTTGATCCGGCTACCTGCCCATTGACCACCAAGCGAAACATCGCATCGAGCGAGCACGTACT
CGGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAG
CCGAACTGTTGCCAGGCTCAAGGCGCGCATGCCCCGACGGCGAGGATCTCGTCGTGACCCATGG
CGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATTA**ACTGTGGC**
CGGCTGGGTTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGC
TTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTGCAGCG
CATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTAA

10

20

【1098】

以下の配列において、プレーンテキストはアデノシンデアミナーゼ配列を示し、太字の配列はCas9由来の配列を示し、斜体の配列はリンカー配列を示し、下線付きの配列は二部分核局在化配列を示し、二重下線付きの配列は突然変異を示す。

【1099】

30

40

50

EIGKATAYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQ
VNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFMQPTVAYSVLVAKVEKGKSK
KLKSVKELLGITIMERSSEFKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAKF
LQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLEVEQHKHYLDEIIEQISEFSKRVILA
DANLDKVL SAYNKHRDKPIREQAENIIHLFTLTNLGAPRAFKYFDTTIARKEYRSTKEVLDTL
IHQSITGLYETRIDLSQLGGDGGSGGSGGSGGSGGSGGSGGMDKKYSIGLAIGTNSVGWAVITD
EYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS
NEMAKVDDSFHRL EESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADL
RLIYLA LAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSAR
LSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKS NFDLAEDAKLQLSKDTYDDDLNLL
AQIGDQYADLFLAAKNLS DAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQOL
PEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN
GSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEE
TITPWNFEVVDKGASAQSFIERMTNFDKNLPNEKVL PKHSLLEYFTVYNELTKVKYVTEGMR
KPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLK
IIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLS
RKLINGIRDKQSGKTILDFLKS DGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANL
AGSPAIKKGILQTVKVVDLVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKEL
GSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNK
VLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKR
QLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQFYKVREINNYHH
AHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEGADKRTADGSEFESP KKKR
KV*

10

20

30

【 1 1 0 0 】

40

50

MSEVEFSHEYWMRHALTLAKRARDEREVPGAVLVLNRRVIGEGWNRAIGLHDPTAHAEIMALR
QGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNH
RVEITEGILADECAALLCTFFRMPRQVFNAQKKAQSSTD **SGGSSGGSSGSETPGTSESATPESS**
GGSSGGSEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVR
KVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFMQPTVAYSVLVVAK
VEKGSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKR
MLASAKFLQKGNELALPSKYVNFLYLASHYEKLKGS PEDNEQKQLFVEQHKKHYLDEIIIEQISEF
SKRVILADANLDKVL SAYNKH RDKPIREQAENIIHLFTLTNLGAPRAFKYFDTTIARKEYRSTK
EVL DATLIHQ SITGLYETRIDLSQLGGDGGSGGSGGSGGSGGSGGSGGMDKKYSIGLAIGTNSV
GWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRIC
YLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVD
STDKADLR LIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDA
KAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKD TYD
DDL DNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLK
ALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLR
KQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRFAW
MTRKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNEKVL PKHSLLEYEFTVYNELTKVK
YVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLG
TYHDL LKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRY
TGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDSL
HEHIANLAGSPAIKKGILQTVKVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRI
EEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK
DDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKKNYWRQLLNAKLITQRKFDNLTKAERGGLSELD
KAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQFYKVR
EINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQ EGADKRTADGSEF
ESP KKKRKV*

10

20

30

【 1 1 0 1 】

40

50

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVNNRVIGEGWNRAIGLHDPTAHAEIMALR
QGGLVMQNYRLIDATLYVTTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNH
RVEITEGILADECAALLCTFFRMPRQVFNAQKKAQSSTDSSGGSSGGSSGSETPGTSESATPSS
GGSSGGSEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVR
KVLSPQVNIIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFMQPTVAYSVLVVAK
VEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDIIKLPKYSLFELENGRKR
MLASAKFLQKGNELALPSKYVNFLYLASHYEKLKGSPEDEQKQLFVEQHKKHYLDEIIEQISEF
SKRVILADANLDKVL SAYNKHDKPIREQAENIIHLFTLTNLGAPRAFKYFDTTIARKEYRSTK
EVL DATLIHQSI TGLYETRIDLSQLGGDGGSSGGSSGGSSGGSSGGMDKKYSIGLAIGTNSV
GWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRKNRIC
YLQEIFS NEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVD
STDKADRLRIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLEENPINASGVDA
KAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKD TYD
DDLNDLLAQIGDQYADFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLK
ALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLR
KQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPIYYVGPLARGNSRFAW
MTRKSEETITPWNFEVVDKGASQSFIERMTNFDKNLPNEKVLPHSLLYEYFTVYNELTKVK
YVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLG
TYHDLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQLKRRRY
TGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGD SL
HEHIANLAGSPAIIKGILOTVKVVDLVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRI
EEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVHIVPQSFLK
DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELD
KAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVR
EINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIKSEQEGADKRTADGSEF
ESPKKRKV*

10

20

30

【 1 1 0 2 】

40

50

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHAEIMALR
QGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAGSLMDVLHHPGMNH
RVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDGGSSGGSSGSETPGTSESATPESSE
GSSGSSMSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVHNNRVIGEGWNRRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLH
YPGMNRVEITEGILADECAALLCTFFRMPRQVFNAQKKAQSSTDSSGSSGGSSGSETPGTSES
ATPESSEGGSSGS

10

EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQ
VNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPPKYGGFMQPTVAYSVLVAKVEKGKSK
KLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDLIIKLPKYSLEFENGRKRMLASAKF
LQKGNELALPSKYVNFYLASHYEKLKGSPEDEQKQLFVEQHKHYLDEIEQISEFSKRVILA
DANLDKVL SAYNKHDKPIREQAENIIHLFTLTNLGAPRAFKYFDTTARKEYRSTKEVLDATL
IHQSITGLYETRIDLSQLGGDGGSSGGSSGGSSGGSSGGMDKKYSIGLAIGTNSVGWAVITD
EYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS
NEMAKVDDSSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADL
RLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSAR
LSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKD TYDDDLNLL
AQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQL
PEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN
GSIPHQIHLGELHAILRRQEDFYPPFLKDNREKIEKILTFRIPIYVGPLARGNSRFAWMTRKSEE
TITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMR
KPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLK
IIKDKDFLDNEENEDILEDIVLTLTLEFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLS
RKLINGIRDKQSGKTILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQDLSLHEHIANL
AGSPAIKKGILQTVKVVDLVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIE EG I K E L
GSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNK
VLTRSDKNRGKSDNVPSEEVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKR
QLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVINNYHH
AHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIKSEQEGADKRTADGSEFESPKKKR

20

30

KV*

40

[1 1 0 3]

50

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVNNRVIGEGWNRAIGLHDPTAHAEIMALR
QGGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHHPGMNH
RVEITEGILADECAALLCRFFRMPRRVFNAQKKAQSSTDSSGSSGGSSGSETPGTSESATPESS
GGSSGSSDKKYSIGLAIGTNSVGWAVITDEYKVP SKKFV LGNTDRHSIKKNLIGALLFDSGET
AEATRLKRTARRRYTRKNRICYLQEIFS NEMAKVDD SFFHRLEESFLVEEDKKHERHPIFGNI
VDEVAYHEKYPTIYHLRKKLV DSTDKADLR LIYLALAHMIKFRGHFLIEGDLNPDNSDVKLEFI
QLVQTYNQLFEE NPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
NFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITK
APLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKP
ILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPFLKDNREKIEK
ILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGASAQSFIERM TNFDKNLPNEK
VLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFK
KIECFDSVEISGVEDRFNASLGTYHDLKKI IKDKDFLDNEENEDI EDIVLTTLTFEDREMIEE
RLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKS DGFANRNF MQLIH
DDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPA IKKGILQTVKVVD ELVKVMGRHKPENIVIE
MARENQTTQKGQNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQ
ELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKKNYWRQLLNAK
LITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILD SRMNTKYDENDKLIREVK
VITLKS KLVSDFRKDFQFYK VREINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDV
RKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFAT
VRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWD PKKYGGFDSPTVAYSVLVV
AKVEKGKSKKLKSVKELLGITIMERS SFEKNPIDFLEAKGYKEVKDLII IKLPKYSLFEL ENGR
KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLGSPEDNEQKQLFVEQH KHYLDEIIEQIS
EFSKR VILADANLDKVL SAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTS
TKEVL DATLIHQ SITGLYETRIDLSQLGGDEGADKRTADGSEFESP KKKRKV*

10

20

30

[1 1 0 4]

40

50

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHAEIMALR
QGGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAGSLMDVLHHPGMNH
RVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSSGSSGGSSGSETPGTSESATPESS
GGSSGGSSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVNLRVIGEGWNRRAIGLHDPTAHA
EIMALRQGGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLH
HPGMNRVEITEGILADECAALLCRFFRMPRRVFNAQKKAQSSTDSSGSSGGSSGSETPGTSES
ATPESSGGSSGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALL
FDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERH
PIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSD
VDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLF'GNLIAL
SLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSAAILSDILRV
NTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEF
YKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDN
REKIEKILTFRIPIYYVGPLARGNSRFawMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDK
NLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQL
KEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFED
REMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRN
FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDLVKVMGRHKP
ENIVIEMARENQTTQKGQKNSRERMKRIE EGIKELGSQILKEHPVENTQLQNEKLYLYLQNGR
DMYVDQELDINRLSDYDVHDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKMKNYWR
QLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILD SRMNTKYDENDK
LIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGD
YKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDK
GRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWD PKKYGGFDSPTVA
YSVLVVAKVEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDLIIKLPKYSLF
ELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGS PEDNEQKQLFVEQHKHYLDE
II EQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLFTLTNLGAPAAFKYFDTTID
RKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGDEGADKRTADGSEFESP KKKRKV*

10

20

30

40

【 1 1 0 5 】

50

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVNNRVIGEGWNRAIGLHDPTAHAEIMALR
QGGVLVMQNYRLYDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHHPGMNH
RVEITEGILADECAALLCRFFRMPRRVFNAQKKAQSSTDSGGSSGGSSGSETPGTSESATPESS
GGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGET
AEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNI
VDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFI
QLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
NFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLLAAKNLSDAILLSDILRVNTEITK
APLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKP
ILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPFLKDNREKIEK
ILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEFVVDKGASAQSFIERMTNFDKNLPNEK
VLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFK
KIECFDSVEISGVEDRFNASLGTYHDLKKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEE
RLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMLIH
DDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVDELVKVMGRHKPENIVIE
MARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQ
ELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKKNYWRQLLNAK
LITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVK
VITLKSCLVSDFRKDFQFYKVRREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDV
RKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFAT
VRKVLSPQVNIIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVV
AKVEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDLIIKLPKYSLEFELNGR
KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDEQKQLFVEQHKKHYLDEIIEQIS
EFSKRVILADANLDKVL SAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTS
TKEVL DATLIHQ SITGLYETRIDLSQLGGDEGADKRTADGSEFESPKKKRKV*

10

20

30

【 1 1 0 6 】

40

50

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHAEIMALR
QGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAGSLMDVLHHPGMNH
RVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSSGSSGGSSGSETPGTSESATPESS
GGSSGGSSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRRAIGLHDPTAHA
EIMALRQGGLVMQNYRLYDATLYVTLEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLH
HPGMNRVEITEGILADECAALLCRFFRMPRRVFNAQKKAQSSTDSSGSSGGSSGSETPGTSES
ATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALL
FDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERH
PIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSD
VDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIAL
SLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRV
NTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEF
YKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDN
REKIEKILTRIPYVVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDK
NLPNEKVLPKHSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQL
KEDYFKKIECFDSVEISGVEDRFNASLGTYHDLKIIKDKDFLDNEENEDILEDIVLTTLTFED
REMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRN
FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVDELVKVMGRHKP
ENIVIAMARENQTTQKGQKNSRERMKRIE EG IKELGSQILKEHPVENTQLQNEKLYLYYLQNGR
DMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKMKNYWR
QLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDK
LIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGD
YKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDK
GRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVA
YSVLVVAKEKGKSKKLKSVKELLGITIMERSSEKNPIDFLEAKGYKEVKKDIIKLPKYSLE
ELENKRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDE
IIEQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLFTLTNLGAPAAFKYFDTTID
RKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGDEGADKRTADGSEFESPKKKRKV*

10

20

30

40

[1 1 0 7]

50

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALR
QGGGLVMQNYRLIDATLYSTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNH
RVEITEGILADECAALLCYFFRMPRRVFNAQKKAQSSTDSGGSSGGSSGSETPGTSESATPESS
GGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGET
AEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNI
VDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFI
QLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTP
NFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLLAAKNLSDAILLSDILRVNTEITK
APLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKP
ILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEK
ILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNEK
VLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFK
KIECFDSVEISGVEDRFNASLGTYHDLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEE
RLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMLIH
DDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVDELVKVMGRHKPENIVIE
MARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQ
ELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKKNYWRQLLNAK
LITQRKFDNLTKAERGGLSELDKAGFIKRLVETRQITKHVAQILDSRMNTKYDENDKLIREVK
VITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDV
RKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFAT
VRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVV
AKVEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDLIIKLPKYSLELENGR
KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGS PEDNEQKQLFVEQHKKHYLDEIIEQIS
EFSKRVLADANLDKVL SAYNKHRDKPIREQAENIIHLEFTLTNLGAPAAFKYFDTTIDRKRYTS
TKEVL DATLIHQ SITGLYETRIDLSQLGGDEGADKRTADGSEFESP KKKRKV*

[1 1 0 8]

10

20

30

40

50

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHAEIMALR
QGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAGSLMDVLHHPGMNH
RVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSSGSSGGSSGSETPGTSESATPESS
GGSSGGSSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVNNRVIGEGWNRRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYSTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLH
YPGMNRVEITEGILADECAALLCYFFRMPRRVFNAQKKAQSSTDSSGSSGGSSGSETPGTSES
ATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALL
FDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERH
PIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSD
VDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIAL
SLGLTPNFKSNFDLAEDAKLQLSKD TYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRV
NTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEF
YKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDN
REKIEKILTRIPYVVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDK
NLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQL
KEDYFKKIECFDSVEISGVEDRFNASLGTYHDLKIIKDKDFLDNEENEDILEDIVLTLTLFED
REMIEERLKTYAHLFDDKVMKQLKRRRYTGWRLSRKLINGIRDKQSGKTILDFLKSDGFANRN
FMQLIHDDSLTFKEDIQKAQVSGQDSLHEHIANLAGSPAIKKGILQTVKVVDLVKVMGRHKP
ENIVIAMARENQTTQKGQKNSRERMKRIE EG IKELGSQILKEHPVENTQLQNEKLYLYYLQNGR
DMYVDQELDINRLSDYVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKNYWR
QLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDK
LIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGD
YKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDK
GRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVA
YSVLVVAKEKGKSKKLKSVKELLGITIMERSSEKNPIDFLEAKGYKEVKKDIIKLPKYSLE
ELENKRKRLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDEQKQLFVEQHKHYLDE
IIEQISEFSKRVLADANLDKVL SAYNKHDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTID
RKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGDEGADKRTADGSEFESPKKKRKV*

10

20

30

40

【 1 1 0 9 】

50

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVNNRVIGEGWNRAIGLHDPTAHAEIMALR
QGGLVMQNYRLYDATLYSTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHHPGMNH
RVEITEGILADECAALLCRFFRMPRVFNAQKKAQSSTDSGGSSGGSSGSETPGTSESATPESS
GGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGET
AEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNI
VDEVAYHEKYPTIYHLRKKLV DSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFI
QLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLLENLIAQLPGEKKNGLFGNLIALSLGLTP
NFKSNFDLAEDAKLQLSKD TYDDDLNLLAQIGDQYADLF LAAKNLSDAILLSDILRVNTEITK
APLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKP
ILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPLKDNREKIEK
ILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEK
VLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFK
KIECFDSVEISGVEDRFNASLGTYHDLKIIKDKDFLDNEENEDILEDIVLTTLTLFEDREMIEE
RLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMOLIH
DDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVDELVKVMGRHKPENIVIE
MARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQ
ELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKKNYWRQLLNAK
LITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVK
VITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDV
RKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFAT
VRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWD PKKYGGFDSP TVAYSVLVV
AKVEKGKSKKLKSVKELLGITIMERSSSF EKNPIDFLEAKGYKEVKDLIIKLPKYSLFEL ENGR
KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQIS
EFSKRVLADANLDKVL SAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTS
TKEVL DATLIHQ SITGLYETRIDLSQLGGDEGADKRTADGSEFESP KKKRKV*

10

20

30

【 1 1 1 0 】

40

50

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHAEIMALR
 QGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAGSLMDVLHHPGMNH
 RVEITEGILADECAALLSDFFRMRRQEIKAKKAQSSTDSSGGSSGGSSGSETPGTSESATPESS
 GGSSGSSSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVHNNRVIGEGWNRRAIGLHDPTAHA
 EIMALRQGGLVMQNYRLYDATLYSTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLH
 HPGMNRVEITEGILADECAALLCRFFRMPRRVFNAQKAQSSTDSSGGSSGGSSGSETPGTSES
 ATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALL
 FDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHRLEESFLVEEDKKHERH
 PIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYALAHMIKFRGHFLIEGDLNPDNSD
 VDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIAL
 SLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRV
 NTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEF
 YKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDN
 REKIEKILTFRIPIYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDK
 NLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQL
 KEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFED
 REMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRN
 FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIAKKGILQTVKVVDLKVVMGRHKP
 ENIVIEMARENQTTQKGQKNSRERMKRIEIGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGR
 DMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKKNYWR
 QLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDK
 LIREVKVITLKSCLVSDFRKDFQFYKVRINNYHHAHDAYLNAVVGITALIKKYPKLESEFVYGD
 YKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDK
 GRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVA
 YSVLVVAKVEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKDLIIKLPKYSLF
 ELENGKRKMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEQKQLEFVEQHKHYLDE
 IIEQISEFSKRVLADANLDKVL SAYNKHDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTID
 RKRYTSTKEVLDTLHQSIITGLYETRIDLSQLGGDEGADKRTADGSEFESPKKKRKV*

10

20

30

40

[1 1 1 1]

01. monoABE8.1_bpNLS + Y147T

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVHNNRVIGEGWNRRAIGLHDPTAHA
 EIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
 MDVLHYPGMNHRVEITEGILADECAALLCTFFRMPRQVFNAQKAQSSTDSSGGSSGGSS
 GSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTD
 RSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHR
 LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYALAHM
 IKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLE
 NLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGD

50

QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLP
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
NGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR
KSEETITPWNFEVVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN
ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMK
QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQ
KAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDELVKVMGRHKPENIVIAMARENQT
TQKGQKNSRERMKRIEELGELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL
DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLL
NAKLITQRKFDNLTKAERGGLSELDKAGFIKQRLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKSCLVSDFRKDFQFYKVVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEF
VYGDYKVVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET
GEIVWDKGRDFATVRKVLSPQVNVIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK
KYGGFVSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEV
KKDLIIKLPKYSLELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLKGSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLF
TLTNLGAPAAFKYFDTTIDRKQYRSTKEVL DATLIHQ SITGLYETRIDLSQLGGDEGADKR
TADGSEFESPKKKRKV

[1 1 1 2]

02. monoABE8.1_bpNLS + Y147R

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNRRVIGEGWNRAIGLHDPTAHA
EIMALRQGGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNHRVEITEGILADECAALLCRFFRMPRQVFNAQKKAQSSTDSGGSSGGSS
GSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTD
RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYALAHM
IKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRE
NLIAQLPGEKKNGLFGNLIALSLGLTPNFKS NFD LAEDAKLQLSKD TYDDDLNLLAQIGD
QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLP
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
NGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR
KSEETITPWNFEVVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN
ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMK
QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQ
KAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDELVKVMGRHKPENIVIAMARENQT
TQKGQKNSRERMKRIEELGELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL
DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLL
NAKLITQRKFDNLTKAERGGLSELDKAGFIKQRLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKSCLVSDFRKDFQFYKVVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEF
VYGDYKVVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET
GEIVWDKGRDFATVRKVLSPQVNVIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK
KYGGFVSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEV
KKDLIIKLPKYSLELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLKGSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLF
TLTNLGAPAAFKYFDTTIDRKQYRSTKEVL DATLIHQ SITGLYETRIDLSQLGGDEGADKR
TADGSEFESPKKKRKV

[1 1 1 3]

03. monoABE8.1_bpNLS + Q154S

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRSVFNAQKKAQSSTDSSGGSSGGSSG
SETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDR
HSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHRL
EESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYALAHMI
KFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRL
NLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGD
QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLP
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
NGSIPHQIHLGELHAILRRQEDFYFPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR
KSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN
ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMK
QLKRRRYTGWGRLSRKLINGIRDQKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQ
KAQVSGQGDSLHEHIANLAGSPAIAKKGILQTVKVDELVKVMGRHKPENIVIAMARENQT
TQKGQKNSRERMKRIEELGSGILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL
DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVKKMKNYWRQLL
NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG TALIKKYPKLESEF
VYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET
GEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK
KYGGFVSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSEKNPIDFLEAKGYKEV
KKDLIIKLPKYSLFELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLKGSPE
NEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RDKPIREQAENIIHLF
TLTNLGAPAAFKYFDTTIDRKQYRSTKEVL DATLIHQ SITGLYETRIDLSQLGGDEGADKR
TADGSEFESPKKKRKV

[1 1 1 4]

04. monoABE8.1_bpNLS + Y123H

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHHPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTDSSGGSSGGSS
GSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTD
RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYALAHM
IKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRL
NLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGD
QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLP
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
NGSIPHQIHLGELHAILRRQEDFYFPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR
KSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN
ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMK
QLKRRRYTGWGRLSRKLINGIRDQKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQ
KAQVSGQGDSLHEHIANLAGSPAIAKKGILQTVKVDELVKVMGRHKPENIVIAMARENQT
TQKGQKNSRERMKRIEELGSGILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL
DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVKKMKNYWRQLL
NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG TALIKKYPKLESEF
VYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET

GEIVWDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK
KYGGFVSPTVAYSVLVVAKEKGSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEV
KKDLIIKLPKYSLFELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLKGSPED
NEQKQLFVEQHKHYLDEIIQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLF
TLTNLGAPAAFKYFDTTIDRKQYRSTKEVL DATLIHQ SITGLYETRIDLSQLGGDEGADKR
TADGSEFESP KKKRKV

【1 1 1 5】

05. monoABE8.1_bpNLS + V82S

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYSTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNRHVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTDSGGSSGGSS
GSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTD
RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDD SFFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLR LIYLALAHM
IKFRGHFLIEGDLNPDNSDV DKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRE
NLIAQLPG EKKNGLFGNLIALSLGLTPNFKS NFDLAEDAKLQLSKD TYDDDLNLLAQIGD
QYADLFLAAKNLS DAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPE
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
NGSIPHQIHLGELHAILRRQEDFYFPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR
KSEETITPWNFEVVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN
ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMK
QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKS DGFANRNF MQLIHDDSLTFKEDIQ
KAQVSGQGDSLHEHIANLAGSPA IKKGILQTVKVVD ELVKVMGRHKPENIVIEMARENQT
TQKGQKNSRERMKRIE EGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL
DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKMKNYWRQLL
NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKS KLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG TALIKKYPKLESEF
VYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET
GEIVWDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK
KYGGFVSPTVAYSVLVVAKEKGSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEV
KKDLIIKLPKYSLFELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLKGSPED
NEQKQLFVEQHKHYLDEIIQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLF
TLTNLGAPAAFKYFDTTIDRKQYRSTKEVL DATLIHQ SITGLYETRIDLSQLGGDEGADKR
TADGSEFESP KKKRKV

【1 1 1 6】

06. monoABE8.1_bpNLS + T166R

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNRHVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSRDSGGSSGGSS
GSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTD
RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDD SFFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLR LIYLALAHM
IKFRGHFLIEGDLNPDNSDV DKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRE
NLIAQLPG EKKNGLFGNLIALSLGLTPNFKS NFDLAEDAKLQLSKD TYDDDLNLLAQIGD
QYADLFLAAKNLS DAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPE
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
NGSIPHQIHLGELHAILRRQEDFYFPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR
KSEETITPWNFEVVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN

ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTLFEDREMIEERLKTYAHLFDDKVMK
QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQ
KAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDLVKVMGRHKPENIVIAMARENQT
TQKGQKNSRERMKRIEEGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL
DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLL
NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG TALIKKYPKLESEF
VYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET
GEIVWDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK
KYGGFVSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEV
KKDLIIKLPKYSLFELNGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLGKSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RDKPIREQAENIIHLF
TLTNLGAPAAFKYFDTTIDRKQYRSTKEVLDATLIHQ SITGLYETRIDLSQLGGDEGADKR
TADGSEFESPKKKRKV

[1 1 1 7]

07. monoABE8.1_bpNLS + Q154R

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHA
EIMALRQGGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRRVFNAQKKAQSSTDSGGSSGGSS
GSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTD
RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDDSFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLR LIYLALAHM
IKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRL
NLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKD TYDDDLNLLAQIGD
QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPE
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
NGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPIYYVGPLARGNSRFAWMTR
KSEETITPWNFEVVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN
ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTLFEDREMIEERLKTYAHLFDDKVMK
QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQ
KAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDLVKVMGRHKPENIVIAMARENQT
TQKGQKNSRERMKRIEEGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL
DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLL
NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG TALIKKYPKLESEF
VYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET
GEIVWDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK
KYGGFVSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEV
KKDLIIKLPKYSLFELNGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLGKSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RDKPIREQAENIIHLF
TLTNLGAPAAFKYFDTTIDRKQYRSTKEVLDATLIHQ SITGLYETRIDLSQLGGDEGADKR
TADGSEFESPKKKRKV

[1 1 1 8]

08. monoABE8.1_bpNLS + Y147R_Q154R_Y123H

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHA
EIMALRQGGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHHPGMNHRVEITEGILADECAALLCRFFRMPRRVFNAQKKAQSSTDSGGSSGGSS
GSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTD
RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDDSFHR

LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYALAHM
IKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRL
NLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDQTYDDDLNLLAQIGD
QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLP
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
NGSIPHQIHLGELHAILRRQEDFYFPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR
KSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN
ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMK
QLKRRRYTGWGRLSRKLINGIRDQKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQ
KAQVSGQGDSLHEHIANLAGSPAIAKKGILQTVKVVDELVKVMGRHKPENIVIAMARENQT
TQKGQKNSRERMKRIEELGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL
DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLL
NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEF
VYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET
GEIVWDKGRDFATVRKVLSPQVNVIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK
KYGGFVSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSEKKNPIDFLEAKGYKEV
KKDLIIKLPKYSLELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLGKSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLF
TLTNLGAPAAFKYFDTTIDRKQYRSTKEVLDATLIHQ SITGLYETRIDLSQLGGDEGADKR
TADGSEFESPKKKRKV

[1 1 1 9]

09. monoABE8.1_bpNLS + Y147R_Q154R_I76Y

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVNNRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLYDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNHRVEITEGILADECAALLCRFFRMPRRVFNAQKKAQSSTDSGGSSGGSSG
SETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDR
HSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRL
EESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYALAHMI
KFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRL
NLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDQTYDDDLNLLAQIGD
QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLP
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
NGSIPHQIHLGELHAILRRQEDFYFPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR
KSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN
ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMK
QLKRRRYTGWGRLSRKLINGIRDQKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQ
KAQVSGQGDSLHEHIANLAGSPAIAKKGILQTVKVVDELVKVMGRHKPENIVIAMARENQT
TQKGQKNSRERMKRIEELGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL
DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLL
NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEF
VYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET
GEIVWDKGRDFATVRKVLSPQVNVIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK
KYGGFVSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSEKKNPIDFLEAKGYKEV
KKDLIIKLPKYSLELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLGKSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLF
TLTNLGAPAAFKYFDTTIDRKQYRSTKEVLDATLIHQ SITGLYETRIDLSQLGGDEGADKR

10

20

30

40

50

TADGSEFESPKKKRKV

【1 1 2 0】

10. monoABE8.1_bpNLS + Y147R_Q154R_T166R

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVNNRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNHRVEITEGILADECAALLCRFFRMPRRVFNAQKKAQSSRDSGGSSGGSS
GSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTD
RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRL
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHM
IKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRL
NLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNDLAEDAKLQLSKDQTYDDDLNLLAQIGD
QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPE
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
NGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTRIPYYVGPLARGNSRFAWMTR
KSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN
ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMK
QLKRRRYTGWGRLSRKLINGIRDQKQSGKTILDFLKSDGDFANRNFQMQLIHDDSLTFKEDIQ
KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDDELVKVMGRHKPENIVIAMARENQT
TQKGQKNSRERMKRIEELGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL
DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVKKMKNYWRQLL
NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKSCLVSDFRKDFQFYKVVREINNYHHAHDAYLNAVVG TALIKKYPKLESEF
VYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET
GEIVWDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK
KYGGFVSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSEKKNPIDFLEAKGYKEV
KKDLIIKLPKYSLFELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLGSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RDKPIREQAENIIHLF
TLTNLGAPAAFKYFDTTIDRKQYRSTKEVL DATLIHQ SITGLYETRIDLSQLGGDEGADKR
TADGSEFESPKKKRKV

【1 1 2 1】

11. monoABE8.1_bpNLS + Y147T_Q154R

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVNNRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNHRVEITEGILADECAALLCTFFRMPRRVFNAQKKAQSSDSSGGSSGGSSG
SETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDR
HSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRL
EESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMI
KFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRL
NLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNDLAEDAKLQLSKDQTYDDDLNLLAQIGD
QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPE
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
NGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTRIPYYVGPLARGNSRFAWMTR
KSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN
ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMK
QLKRRRYTGWGRLSRKLINGIRDQKQSGKTILDFLKSDGDFANRNFQMQLIHDDSLTFKEDIQ
KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDDELVKVMGRHKPENIVIAMARENQT
TQKGQKNSRERMKRIEELGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL
DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVKKMKNYWRQLL

NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG TALIKKYPKLESEF
VYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET
GEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK
KYGGFVSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSSF EKNPIDFLEAKGYKEV
KKDLIIKLPKYSLFELENGKRKRLASARELQKGNELALPSKYVNFLYLASHYEKLKGSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLF
TLTNLGAPAAFKYFDTTIDRKQYRSTKEVLDATLIHQ SITGLYETRIDLSQLGGDEGADKR
TADGSEFESP KKKRKV

[1 1 2 2]

10

12. monoABE8.1_bpNLS + Y147T_Q154S

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHA
EIMALRQGG LVMQNYRLIDATLYVT FEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNHRVEITEGILADECAALLCTFFRMPRSVFNAQKKAQSSTD SGGSSGGSSG
SETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDR
HSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDD SFFHRL
EESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLR LIYLALAHMI
KFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRL E
NLIAQLPGEKKNGLFGNLIALSLGLTPNFKS NFDLAEDAKLQLSKD TYDDDLNLLAQIGD
QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPE
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
NGSIPHQIHLGELHAILRRQEDFY PFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR
KSEETITPWNFEVVDKGASAQSFIERM TNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN
ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLT LFEDREMIEERLKYAHLFDDKVMK
QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKS DGFANRNF MQLIHDDSLTFKEDIQ
KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVD ELVKVMGRHKPENIV IEMARENQT
TQKGQKNSRERMKRIE EGikelGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL
DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNPVSEEVVKMKMKNYWRQLL
NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG TALIKKYPKLESEF
VYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET
GEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK
KYGGFVSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSSF EKNPIDFLEAKGYKEV
KKDLIIKLPKYSLFELENGKRKRLASARELQKGNELALPSKYVNFLYLASHYEKLKGSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLF
TLTNLGAPAAFKYFDTTIDRKQYRSTKEVLDATLIHQ SITGLYETRIDLSQLGGDEGADKR
TADGSEFESP KKKRKV

20

30

[1 1 2 3]

13. monoABE8.1_bpNLS + H123Y123H_Y147R_Q154R_I76Y

40

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHA
EIMALRQGG LVMQNYRLYDATLYVT FEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHHPGMNHRVEITEGILADECAALLCRFFRMPRRVFNAQKKAQSSTD SGGSSGGSSG
GSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTD
RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDD SFFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLR LIYLALAHM
IKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRL E
NLIAQLPGEKKNGLFGNLIALSLGLTPNFKS NFDLAEDAKLQLSKD TYDDDLNLLAQIGD
QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPE
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD

50

NGSIPHQIHLGELHAILRRQEDFYPPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR
KSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN
ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMK
QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQ
KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDLVKVMGRHKPENIVIAMARENQT
TQKGQKNSRERMKRIEIEGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL
DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLL
NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEF
VYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET
GEIVWDKGRDFATVRKVLSPQVNVIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK
KYGGFVSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSEKNPIDFLEAGYKEV
KKDLIIKLPKYSLELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLGSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLF
TLTNLGAPAAFKYFDTTIDRKQYRSTKEVL DATLIHQ SITGLYETRIDLSQLGGDEGADKR
TADGSEFESPKKKRKV

【1 1 2 4】

14. monoABE8.1_bpNLS + V82S + Q154R

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNRRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYSTFEPVCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRRVFNAQKKAQSSTDSGGSSGGSS
GSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVP SKKFKVLGNTD
RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDDSFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLR LIYLALAHM
IKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRL
NLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGD
QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLP
E KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
NGSIPHQIHLGELHAILRRQEDFYPPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR
KSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN
ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMK
QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQ
KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDLVKVMGRHKPENIVIAMARENQT
TQKGQKNSRERMKRIEIEGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL
DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLL
NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEF
VYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET
GEIVWDKGRDFATVRKVLSPQVNVIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK
KYGGFVSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSEKNPIDFLEAGYKEV
KKDLIIKLPKYSLELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLGSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLF
TLTNLGAPAAFKYFDTTIDRKQYRSTKEVL DATLIHQ SITGLYETRIDLSQLGGDEGADKR
TADGSEFESPKKKRKV

【1 1 2 5】

【実施例8：パーキンソン病】

材料と方法

本明細書に記載される実施例で提供される結果は、以下の材料および方法を用いて得ら

10

20

30

40

50

れた。

【1 1 2 6】

実施例で使用されたABEの配列は次のとおり：

【1 1 2 7】

01. monoABE8.1_bpNLS + Y147T

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNRRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNRHVEITEGILADECAALLCTFFRMPRQVFNAQKKAQSSTDSSGGSSGGSS
GSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTD
RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHHR
LEESFLVEEDKKHERHPHIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYALAHM
IKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRL
NLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNDLAEDAKLQLSKDTYDDDLNLLAQIGD
QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLP
E
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
NGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR
KSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN
ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMK
QLKRRRYTGWGRLSRKLINGIRDQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQ
KAQVSGQGDSLHEHIANLAGSPAIAKKGILQTVKVVDLKVVMGRHKPENIVIAMARENQT
TQKGQKNSRERMKRIEELGSGILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL
DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVKKMKNYWRQLL
NAKLITQRKFDNLTKAERGGSELKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG TALIKKYPKLESEF
VYGDKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET
GEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK
KYGGFVSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSEKKNPIDFLEAKGYKEV
KKDLIIKLPKYSLFELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLKGSPE
D
NEQKQLFVEQHKKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RDKPIREQAENIIHLF
TLTNLGAPAAFKYFDTTIDRKQYRSTKEVL DATLIHQ SITGLYETRIDLSQLGGDEGADKR
TADGSEFESPKKKRKV

【1 1 2 8】

02. monoABE8.1_bpNLS + Y147R

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNRRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNRHVEITEGILADECAALLCRFFRMPRQVFNAQKKAQSSTDSSGGSSGGSS
GSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTD
RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHHR
LEESFLVEEDKKHERHPHIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYALAHM
IKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRL
NLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNDLAEDAKLQLSKDTYDDDLNLLAQIGD
QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLP
E
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
NGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR
KSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN
ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMK
QLKRRRYTGWGRLSRKLINGIRDQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQ
KAQVSGQGDSLHEHIANLAGSPAIAKKGILQTVKVVDLKVVMGRHKPENIVIAMARENQT

TQKGQKNSRERMKRIEELGSGILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL
DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVKKMKNYWRQLL
NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEF
VYGDYKVVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET
GEIVWDKGRDFATVRKVLSPQVNVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK
KYGGFVSPTVAYSVLVVAKEKGSKKLSVKELLGITIMERSSEKKNPIDFLEAKGYKEV
KKDLIIKLPKYSLELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLGSPED
NEQKQLFVEQHKHYLDEIEQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLF
TLTNLGAPAAFKYFDTTIDRKQYRSTKEVL DATLIHQ SITGLYETRIDLSQLGGDEGADKR
TADGSEFESPKKKRKV

10

[1 1 2 9]

03. monoABE8.1_bpNLS + Q154S

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRSVFNAQKKAQSSTDSSGSSGGSSG
SETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDR
HSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDD SFFHRL
EESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLR LIYLALAHMI
KFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRL
NLIAQLPGKKNGLFGNLIALSLGLTPNFKS NFDLAEDAKLQLSKD TYDDDLNLLAQIGD
QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPE
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
NGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR
KSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN
ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMK
QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKS DGFANRNF MQLIHDDSLTFKEDIQ
KAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDDELVKVMGRHKPENIVIAMARENQT
TQKGQKNSRERMKRIEELGSGILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL
DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVKKMKNYWRQLL
NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEF
VYGDYKVVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET
GEIVWDKGRDFATVRKVLSPQVNVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK
KYGGFVSPTVAYSVLVVAKEKGSKKLSVKELLGITIMERSSEKKNPIDFLEAKGYKEV
KKDLIIKLPKYSLELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLGSPED
NEQKQLFVEQHKHYLDEIEQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLF
TLTNLGAPAAFKYFDTTIDRKQYRSTKEVL DATLIHQ SITGLYETRIDLSQLGGDEGADKR
TADGSEFESPKKKRKV

20

30

40

[1 1 3 0]

04. monoABE8.1_bpNLS + Y123H

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHHPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTDSSGSSGGSSG
GSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTD
RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDD SFFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLR LIYLALAHM
IKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRL
NLIAQLPGKKNGLFGNLIALSLGLTPNFKS NFDLAEDAKLQLSKD TYDDDLNLLAQIGD

50

QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPE
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
NGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR
KSEETITPWNFEVVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN
ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMK
QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQ
KAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDELVKVMGRHKPENIVIAMARENQT
TQKGQKNSRERMKRIEELGELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL
DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLL
NAKLITQRKFDNLTAKAERGGLSELDKAGFIKQRLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKSCLVSDFRKDFQFYKVVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEF
VYGDYKVVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET
GEIVWDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK
KYGGFVSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSEKKNPIDFLEAKGYKEV
KKDLIIKLPKYSLELNGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLKGSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLF
TLTNLGAPAAFKYFDTTIDRKQYRSTKEVL DATLIHQ SITGLYETRIDLSQLGGDEGADKR
TADGSEFESPKKKRKV

[1 1 3 1]

05. monoABE8.1_bpNLS + V82S

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNRRVIGEGWNRAIGLHDPTAHA
EIMALRQGGGLVMQNYRLIDATLYSTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTDSGGSSGGSS
GSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTD
RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYALAHM
IKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRE
NLIAQLPGEKKNGLFGNLIALSLGLTPNFKS NFDLAEDAKLQLSKD TYDDDLNLLAQIGD
QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPE
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
NGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR
KSEETITPWNFEVVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN
ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMK
QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQ
KAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDELVKVMGRHKPENIVIAMARENQT
TQKGQKNSRERMKRIEELGELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL
DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLL
NAKLITQRKFDNLTAKAERGGLSELDKAGFIKQRLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKSCLVSDFRKDFQFYKVVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEF
VYGDYKVVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET
GEIVWDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK
KYGGFVSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSEKKNPIDFLEAKGYKEV
KKDLIIKLPKYSLELNGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLKGSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLF
TLTNLGAPAAFKYFDTTIDRKQYRSTKEVL DATLIHQ SITGLYETRIDLSQLGGDEGADKR
TADGSEFESPKKKRKV

[1 1 3 2]

06. monoABE8.1_bpNLS + T166R

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNRRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNRHVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSRDSSGGSSGGSS
GSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSSKFKVLGNTD
RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADRLRIYLALAHM
IKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRE
NLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGD
QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLP
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
NGSIPHQIHLGELHAILRRQEDFYFPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR
KSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN
ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMK
QLKRRRYTGWGRLSRKLINGIRDQKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQ
KAQVSGQGDSLHEHIANLAGSPAIAKKGILQTVKVDELVKVMGRHKPENIVIAMARENQT
TQKGQKNSRERMKRIEELGSGILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL
DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLL
NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG TALIKKYPKLESEF
VYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET
GEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK
KYGGFVSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSEKNPIDFLEAGYKEV
KKDLIIKLPKYSLFELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLKGSPE
NEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RDKPIREQAENIIHLF
TLTNLGAPAAFKYFDTTIDRKQYRSTKEVLDATLIHQ SITGLYETRIDLSQLGGDEGADKR
TADGSEFESPKKKRKV

[1 1 3 3]

07. monoABE8.1_bpNLS + Q154R

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNRRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNRHVEITEGILADECAALLCYFFRMPRRVFNAQKKAQSSTDSSGGSSGGSS
GSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSSKFKVLGNTD
RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADRLRIYLALAHM
IKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRE
NLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGD
QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLP
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
NGSIPHQIHLGELHAILRRQEDFYFPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR
KSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN
ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMK
QLKRRRYTGWGRLSRKLINGIRDQKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQ
KAQVSGQGDSLHEHIANLAGSPAIAKKGILQTVKVDELVKVMGRHKPENIVIAMARENQT
TQKGQKNSRERMKRIEELGSGILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL
DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLL
NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG TALIKKYPKLESEF
VYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET

GEIVWDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK
KYGGFVSPTVAYSVLVVAKEKGSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEV
KKDLIIKLPKYSLFELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLKGSPED
NEQKQLFVEQHKHYLDEIIQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLF
TLTNLGAPAAFKYFDTTIDRKQYRSTKEVL DATLIHQ SITGLYETRIDLSQLGGDEGADKR
TADGSEFESPKKKRKV

【1 1 3 4】

08. monoABE8.1_bpNLS + Y147R_Q154R_Y123H

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHHPGMNHRVEITEGILADECAALLCRFFRMPRRVFNAQKKAQSSTDSGGSSGGSS
GSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTD
RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDD SFFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLR LIYLALAHM
IKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRL
NLIAQLPGEKKNGLFGNLIALSLGLTPNFKS NFDLAEDAKLQLSKD TYDDDLNLLAQIGD
QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPE
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
NGSIPHQIHLGELHAILRRQEDFYFPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR
KSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN
ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMK
QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKS DGFANRNF MQLIHDDSLTFKEDIQ
KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDDELVKVMGRHKPENIVIEMARENQT
TQKGQKNSRERMKRIEELGSGILKEHPVENTQLQNEKLYLYLQNGRDMYVDQEL
DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKMKNYWRQLL
NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKS KLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG TALIKKYPKLESEF
VYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET
GEIVWDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK
KYGGFVSPTVAYSVLVVAKEKGSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEV
KKDLIIKLPKYSLFELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLKGSPED
NEQKQLFVEQHKHYLDEIIQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLF
TLTNLGAPAAFKYFDTTIDRKQYRSTKEVL DATLIHQ SITGLYETRIDLSQLGGDEGADKR
TADGSEFESPKKKRKV

【1 1 3 5】

09. monoABE8.1_bpNLS + Y147R_Q154R_I76Y

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLYDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNHRVEITEGILADECAALLCRFFRMPRRVFNAQKKAQSSTDSGGSSGGSSG
SETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDR
HSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDD SFFHRL
EESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLR LIYLALAHMI
KFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRL
NLIAQLPGEKKNGLFGNLIALSLGLTPNFKS NFDLAEDAKLQLSKD TYDDDLNLLAQIGD
QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPE
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
NGSIPHQIHLGELHAILRRQEDFYFPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR
KSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN

ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTLFEDREMIEERLKTYAHLFDDKVMK
QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQ
KAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDLVKVMGRHKPENIVIAMARENQT
TQKGQKNSRERMKRIEELGSGILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL
DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLL
NAKLITQRKFDNLTKAERGGSELKAGFIKRLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG TALIKKYPKLESEF
VYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET
GEIVWDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK
KYGGFVSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSEKKNPIDFLEAKGYKEV
KKDLIIKLPKYSLELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLGKSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLF
TLTNLGAPAAFKYFDTTIDRKQYRSTKEVL DATLIHQ SITGLYETRIDLSQLGGDEGADKR
TADGSEFESPKKKRKV

[1 1 3 6]

10. monoABE8.1_bpNLS + Y147R_Q154R_T166R

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNRRVIGEGWNRAIGLHDPTAHA
EIMALRQGGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNHRVEITEGILADECAALLCRFFRMPRRVFNAQKKAQSSRDSGGSSGGSS
GSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTD
RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDDSFH R
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLR LIYLALAHM
IKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRL E
NLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKD TYDDDLDNLLAQIGD
QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPE
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
NGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPIYVVGPLARGNSRFAWMTR
KSEETITPWNFEVVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN
ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTLFEDREMIEERLKTYAHLFDDKVMK
QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQ
KAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDLVKVMGRHKPENIVIAMARENQT
TQKGQKNSRERMKRIEELGSGILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL
DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLL
NAKLITQRKFDNLTKAERGGSELKAGFIKRLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG TALIKKYPKLESEF
VYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET
GEIVWDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK
KYGGFVSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSEKKNPIDFLEAKGYKEV
KKDLIIKLPKYSLELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLGKSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLF
TLTNLGAPAAFKYFDTTIDRKQYRSTKEVL DATLIHQ SITGLYETRIDLSQLGGDEGADKR
TADGSEFESPKKKRKV

[1 1 3 7]

11. monoABE8.1_bpNLS + Y147T_Q154R

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNRRVIGEGWNRAIGLHDPTAHA
EIMALRQGGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNHRVEITEGILADECAALLCTFFRMPRRVFNAQKKAQSSTDSGGSSGGSSG
SETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDR
HSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDDSFHRL

EESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYLAHAHMI
KFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRL
NLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNDLAEDAKLQLSKDQTYDDDLNLLAQIGD
QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLP
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
NGSIPHQIHLGELHAILRRQEDFYFPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR
KSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN
ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMK
QLKRRRYTGWGRLSRKLINGIRDQKQSGKTILDFLKSDFANRNFQMQLIHDDSLTFKEDIQ
KAQVSGQGDSLHEHIANLAGSPAIAKKGILQTVKVDELVKVMGRHKPENIVIAMARENQT
TQKGQKNSRERMKRIEELGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL
DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLL
NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEF
VYGDKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET
GEIVWDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK
KYGGFVSPTVAYSVLVVAKEGKSKKLKSVKELLGITIMERSSEKKNPIDFLEAKGYKEV
KKDLIIKLPKYSLELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLGKSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RDKPIREQAENIIHLF
TLTNLGAPAAFKYFDTTIDRKQYRSTKEVLDATLIHQ SITGLYETRIDLSQLGGDEGADKR
TADGSEFESPKKKRKV

[1 1 3 8]

12. monoABE8.1_bpNLS + Y147T_Q154S

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNRRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNHRVEITEGILADECAALLCTFFRMPRSVFNAQKKAQSSTDSSGGSSGGSSG
SETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDR
HSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRL
EESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYLAHAHMI
KFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRL
NLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNDLAEDAKLQLSKDQTYDDDLNLLAQIGD
QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLP
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
NGSIPHQIHLGELHAILRRQEDFYFPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR
KSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN
ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMK
QLKRRRYTGWGRLSRKLINGIRDQKQSGKTILDFLKSDFANRNFQMQLIHDDSLTFKEDIQ
KAQVSGQGDSLHEHIANLAGSPAIAKKGILQTVKVDELVKVMGRHKPENIVIAMARENQT
TQKGQKNSRERMKRIEELGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL
DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLL
NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEF
VYGDKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET
GEIVWDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK
KYGGFVSPTVAYSVLVVAKEGKSKKLKSVKELLGITIMERSSEKKNPIDFLEAKGYKEV
KKDLIIKLPKYSLELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLGKSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RDKPIREQAENIIHLF
TLTNLGAPAAFKYFDTTIDRKQYRSTKEVLDATLIHQ SITGLYETRIDLSQLGGDEGADKR

10

20

30

40

50

TADGSEFESPKKKRKV

[1 1 3 9]

13. monoABE8.1_bpNLS + H123Y123H_Y147R_Q154R_I76Y

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVNNRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLYDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHHPGMNHRVEITEGILADECAALLCRFFRMPRRVFNAQKKAQSSTDSSGSSGGSS
GSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTD
RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHR
LEESFLVEEDKKHERHPFIGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYALAHM
IKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRE
NLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNDLAEDAKLQLSKDITYDDDLNLLAQIGD
QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPE
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
NGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR
KSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN
ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMK
QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQ
KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDDELVKVMGRHKPENIVIAMARENQT
TQKGQKNSRERMKRIEELGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL
DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVKKMKNYWRQLL
NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG TALIKKYPKLESEF
VYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET
GEIVWDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK
KYGGFVSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSEKKNPIDFLEAKGYKEV
KKDLIIKLPKYSLFELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLGSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLF
TLTNLGAPAAFKYFDTTIDRKQYRSTKEVL DATLIHQ SITGLYETRIDLSQLGGDEGADKR
TADGSEFESPKKKRKV

[1 1 4 0]

14. monoABE8.1_bpNLS + V82S + Q154R

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVNNRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYSTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRRVFNAQKKAQSSTDSSGSSGGSS
GSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTD
RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHR
LEESFLVEEDKKHERHPFIGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYALAHM
IKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRE
NLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNDLAEDAKLQLSKDITYDDDLNLLAQIGD
QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPE
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
NGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR
KSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN
ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMK
QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQ
KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDDELVKVMGRHKPENIVIAMARENQT
TQKGQKNSRERMKRIEELGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL
DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVKKMKNYWRQLL

NAKLITQRKFDNLTKAERGGSELKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND
 KLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG TALIKKYPKLESEF
 VYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET
 GEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK
 KYGGFVSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSSF EKNPIDFLEAKGYKEV
 KKDLIIKL PKYSLFELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLKGSPED
 NEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLF
 TLTNLGAPAAFKYFDTTIDRKQYRSTKEVLDATLIHQ SITGLYETRIDLSQLGGDEGADKR
 TADGSEFESPKKKRKV

【1141】

アミノ酸置換D1135M、S1136Q、G1218K、E1219F、A1322R、D1332A、R1335E、およびT1337R (MQKFRAER) を含み、改変されたPAM 5'-NGC-3'に対する特異性を有する改変SpCas9を、G2019Sの補正のために使用した。改変されたPAM 5'-NGA-3'に対する特異性を有する改変SpCas9-VRQRをR1441Cの補正のために使用した。

【1142】

〔クローニング〕

標的ポリヌクレオチドならびに使用されるgRNAおよびプライマーのDNA配列は、本明細書中に記載される。gRNAについては、以下のスカフォールド（足場）配列が提示される：GUUUUAGAGC UAGAAAUAGC AAGUUAUUU AAGGCUAGUC CGUUAUCAAC UUGAAAAAGU GGCACCGAGU CGGUGCUUUU。この足場は、図57 A～図57 Cおよび図58 A～図58 Cで記述されるPAM（例えば、それぞれNGAおよびNGC PAM）のために使用された。gRNAは、足場配列と、本明細書に記載されるかまたは当業者の知識に基づいて決定され当業者に理解される病原性変異を含むLRRK2遺伝子のためのスペーサー配列（標的配列）を包含する（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: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)。

【1143】

PCRはVeraSeq ULtra DNAポリメラーゼ (Enzymatics) またはQ 5ホットスタート高忠実度DNAポリメラーゼ (New England Biolabs) を用いて行った。塩基エディター (BE) プラスミドはUSERクローニング (New England Biolabs) を用いて構築した。デアミナーゼ遺伝子は、gBlocks遺伝子断片 (Integrated DNA Technologies) として合成した。使用したCas9遺伝子を下記に示す。Cas9遺伝子は以前に報告されたプラスミドから得られた。デアミナーゼおよび融合遺伝子をpCMV（哺乳類コドン最適化）またはpET28b (E. coliコドン最適化) バックボーン中にクローン化した。部位特異的変異誘発を用いてsgRNA発現プラスミドを構築した。

【1144】

簡単に説明すると、プライマーは、製造業者の説明書に従ってT4ポリヌクレオチドキナーゼ (New England Biolabs) を用いて5'リン酸化された。ガイド1および2の領域を増幅するために、以下のプライマーを使用した：

ガイド1プライマー (oAM129 ; 5' -AAGCGCAAGCCTGGAGGGAA-3' について) :

5' -GAAGCGCAAGCCTGGAGGGAAAGTTTTAGAGCTAGAAATAGCA-3' ;

ガイド2プライマー (oAM130; 5' -ACTACAGCATTGCTCAGTAC-3') :

5' -GACTACAGCATTGCTCAGTACGTTTTAGAGCTAGAAATAGCA-3' ;

共通プライマー (oAM95):

10

20

30

40

50

5' -GGTGTTCGTCCTTTCCACAAG-3'

【1145】

次に、製造業者の指示に従って、リン酸化プライマーおよびテンプレートとして目的遺伝子をコードするプラスミド共にQ5ホットスタート高忠実度ポリメラーゼ (New England Biolabs) を用いてPCRを行った。PCR産物をDpnI (20 U、New England Biolabs)とともに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。

【1146】

[ssDNAに対するin vitro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。全てのCy3標識基質は、Integrated DNA Technologies (IDT) から得た。デアミナーゼは、製造業者の指示に従ってTNT T7 Quick Coupled Transcription/Translation Kit (Promega) を使用して1 µgのプラスミドを用いてin vitroで発現させた。タンパク質発現後、5 µlの溶解物を、CutSmart緩衝液 (New England Biolabs) (50 mM酢酸カリウム、29 mM酢酸トリス、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。

【1147】

[His6-ABE8/PV1-28-リンカー-dCas9融合体の発現および精製]

E. coli BL21 STAR (DE3) コンピテント細胞(ThermoFisher Scientific) をプラスミド (例えばpET28b-His6-ABE8/PV1-28-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、100 µg ml⁻¹のカナマイシンを含むLuria-Bertani (LB) ブロス中で37°Cにて一晚増殖させた。細胞を同じ培地に1:100に希釈し、OD₆₀₀=~0.6になるまで37°Cで増殖させた。培養物を2時間かけて4°Cまで冷却し、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で添加して蛋白質発現を誘導した。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のトリス(ヒドロキシメチル)-アミノメタン (Tris) -HCl (pH 7.5)、1 M 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社)) 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、溶解物上清を単離した。溶解物をHis-Purニッケル-ニトロ酢酸 (ニッケル-NTA) 樹脂 (ThermoFisher Scientific) と4°Cで1時間インキュベートして、Hisタグ付き融合タンパク質を捕捉した。樹脂をカラムに移し、40 mlの溶解緩衝液で洗浄した。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させ、限外濾過(Amicon-Millipore、100 kDaの分子量カットオフ)によって全容量1 mlまで濃縮した。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を含有する低塩精製緩衝液中で20 mlに希釈し、SPセファロース高速流動樹脂 (GEライフサイエンス) 上にロードした。樹脂をこの低塩緩衝液40 mlで洗浄し、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)、0.5 M NaCl、20%グリセロール、10 mM TCEPを含有する活性緩衝液5 mlで溶出した。溶出したタンパク質をSDS-PAGEで定量した。

【1148】

[sgRNAのインビトロ転写]

T7プロモーターに続いて20 bpのsgRNA標的配列を含む線状DNA断片を、製造業者の指示に従ってトランスクリプトエイドT7高収率転写キット (ThermoFisher Scientific) を用いてインビトロで転写した。sgRNA産物を、製造業者の指示に従ってMEGAclear Kit (ThermoFisher Scientific) を用いて精製し、UV吸光度によって定量した。

【1149】

10

20

30

40

50

[Cy3コンジュゲート化dsDNA基質の調製]

典型的には、非標識配列鎖（例えば80 ntの非標識鎖の配列）は、PAGE精製オリゴヌクレオチドとしてIDTに注文した。各80 nt基質の3'末端に相補的な25 nt Cy3標識プライマーは、HPLC精製オリゴヌクレオチドとしてIDTに注文した。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、dNTP（0.75 µlの100 mM溶液）を伴うNEBuffer2（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Cy3標識プライマー（5 µlの100 µM溶液）と組み合わせ、95℃に5分間加熱し、続いて45℃まで0.1℃/秒の率で徐々に冷却した。このアニリング期間の後、Klenow exo-（5 U、New England Biolabs）を添加し、反応物を37℃で1時間インキュベートした。この溶液を緩衝液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のデアミナーゼアッセイ。精製された融合タンパク質（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の最終濃度まで添加し、得られた溶液を37℃で2時間インキュベートした。dsDNAを緩衝液PB（100 µl、Qiagen）およびイソプロパノール（25 µl）の添加によって融合体から分離し、EconoSpinマイクロスピナラム（エポックライフサイエンス）で精製し、20 µlのCutSmart緩衝液（ニューイングランド・バイオラプス）で溶出した。USER酵素（1 U、ニューイングランド・バイオラプス）を、精製され編集されたdsDNAに加え、37℃で1時間インキュベートした。Cy3標識鎖は、5 µlの反応溶液と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）とを組み合わせることによって、その相補鎖から完全に変性分離された。完全長C含有基質を、10% TBE-尿素ゲル（Bio-Rad）上で、切断されたU含有編集化基質から分離し、GE Amersham Typhoonイメージャ上で画像化した。

【1150】

[ハイスループット配列決定のためのin vitro編集dsDNAの調製]

オリゴヌクレオチドはIDTから得た。相補配列をトリス緩衝液中で組み合わせ（100 µM溶液5 µl）、95℃に5分間加熱することによりアニールし、続いて45℃まで0.1℃/秒の速度で徐冷して60 bpのdsDNA基質を生成した。精製された融合タンパク質（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。60量体dsDNA基質を125 nMの最終濃度で添加し、得られた溶液を37℃で2時間インキュベートした。dsDNAを緩衝液PB（100 µl、Qiagen）およびイソプロパノール（25 µl）の添加によって融合物から分離し、EconoSpinマイクロスピナラム（エポックライフサイエンス）上で20 µlのトリス緩衝液で溶出して精製した。得られた編集されたDNA（1 µlを鋳型として使用した）を、高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、製造業者の説明書に従って、13サイクルの増幅でPCRにより増幅した。PCR反応生成物は、RapidTips（Diffinity Genomics）を用いて精製され、精製されたDNAは、配列決定アダプターを含有するプライマーを用いたPCRによって増幅され、精製され、先に記載されたようにMiSeqハイスループットDNAシーケンサー（Illumina）上で配列決定された。

【1151】

[細胞培養]

HEK293T（ATCC CRL-3216）およびU2OS（ATCC HTB-96）は、10%（v/v）ウシ胎児血清（FBS）を補充したダルベッコ改変イーグル培地+GlutaMax（ThermoFisher）中で、37℃、5% CO₂で維持した。HCC1954細胞（ATCC CRL-2338）は、上述のように補充されたRPMI-1640培地（ThermoFisher Scientific）中に維持した。LRRK2を含む不死化細胞（Taconic Biosciences社）は、10%（v/v）ウシ胎児血清（FBS）および200 µg ml⁻¹ Geneticin（ThermoFisher Scientific社）を添加したダルベッコ改変イーグル培地+GlutaMax（ThermoFisher Scientific社）中で培養した。

【1152】

[トランスフェクション]

HEK293T細胞を48ウェルのコラーゲンコートBioCoatプレート (Corning) 上に播種し、約85%の集密度においてトランスフェクトした。簡単に述べれば、製造業者のプロトコールに従って、ウェル当たり1.5 µlのリポフェクタミン2000 (ThermoFisher Scientific) を用いて、750 ngのBEおよび250 ngのsgRNA発現プラスミドをトランスフェクトした。HEK293T細胞は、製造業者の説明書に従って、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 (HEK293T細胞用のプログラムQ-001を用いたVキット)。

【1153】

[ゲノムDNA試料のハイスループットDNA配列決定]

トランスフェクトされた細胞を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。関心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用いて、製造業者の指示に従ってPhusion high-fidelity DNAポリメラーゼ(ThermoFisher) を用いて行った。反応が増幅のリニアレンジで止まることを確実にするために、各プライマー対についてサイクル数を別々に決定した。PCR産物はRapidTips (Diffinity Genomics) を用いて精製した。精製された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アッセイキット (ThermoFisher) およびKAPA Library Quantification Kit-Illumina (KAPA Biosystems) を用いて定量した。サンプルは、前述のようにイルミナMiSeq上で配列決定した (Pattanayak, Nature Biotechnol. 31, 839-843 (2013))。

【1154】

[データ分析]

配列決定リード(reads)をMiSeq Reporter (Illumina) を用いて自動的に脱多重化し、個々のFASTQファイルをカスタムMatlabで解析した。各リードは、Smith-Watermanアルゴリズムを用いて適切な参照配列にペアワイズに整列させた。Qスコアが31未満の塩基コールはNに置き換えられ、従ってヌクレオチド頻度の計算から除外された。この処理により、MiSeq塩基呼び出しエラー率の予測値はおおよそ1,000個に1個になる。リードおよび参照配列がギャップを含まないアラインメント配列をアラインメント表に保存し、そこから各遺伝子座について塩基頻度を一覧にすることができた。インデル頻度は、以前に記載された基準 (Zuris, et al., Nature Biotechnol. 33, 73-80 (2015)) を使用してカスタムMatlabスクリプトにより定量した。配列決定リードを走査して、インデル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配列との完全一致を探索した。完全一致が見つからなかった場合、そのリードは分析から除外された。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する場合、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。

【1155】

[塩基エディターでのPAMバリエーション検証]

新規のCRISPRシステムおよびPAMバリエーションは、塩基エディター (例えばPV1-PV28) が、LRRK2ポリヌクレオチド中に存在する標的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バリエーションが評価され、検証された。PAM評価及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) ; PCT/US2016/058344 (WO 2017/070632) ; 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

10

20

30

40

50

Biotechnology 33, 1293-1298 (2015)に記述されており、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); 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)も参照（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）。

【1156】

【遺伝子編集によるパーキンソン病突然変異の補正】

LRRK2における病原性突然変異R1441CおよびR1441Hはパーキンソン病に関連している。図55に示されるように、LRRK2遺伝子のアンチセンス鎖におけるG>A突然変異と関連するR1441C突然変異は、アデノシンデアミナーゼ活性およびAGA PAM特異性を有する塩基エディターを用いて補正される。LRRK2におけるR1441Hは、LRRK2遺伝子におけるG>A突然変異によってコードされるが、アデノシンデアミナーゼ活性および図55に示される標的配列の3位におけるTGA PAM特異性を有する、または5位におけるTGT特異性を有する塩基エディターを用いて補正される。

【1157】

図56は、パーキンソン病に関連するLRRK2におけるY1699C、G2019S、およびI2020T突然変異の補正のための標的配列を示す概略図である。Y1699C変異はLRRK2遺伝子のアンチセンス鎖上のT>C変異と関連し、これはシチジンデアミナーゼ活性を有する塩基エディターを用いて補正される。G2019S変異は、LRRK2遺伝子のアンチセンス鎖上のG>A変異に関連し、これはアデノシンデアミナーゼ活性を有する塩基エディターを用いて補正される。I2020Tは、LRRK2遺伝子におけるT>C突然変異によってコードされ、これはシチジンデアミナーゼ活性およびTGC PAM特異性を有する塩基エディターを用いて補正される。

【1158】

図57 A～図57 Cに示すように、エディターPV 1～14を使用して、図57 Bに示す配列を有するガイドRNAを用いてLRRK2 R1441Cを編集したが、ただし標的配列中の全てのチミジン (T) はウリジン(U) によって置換した (ガイドRNA1: 5' -AAGCGCAAGCCUGGAGGGAA-3')。AからGへの変換パーセントを図57 Aに示す。例示的な配列リードが図57 Cに示される。

【1159】

図58 A～図58 Cに示すように、エディターPV 15～28を使用して、図58 Bに示す配列を有するガイドRNAを用いてLRRK2 G2019Sを編集したが、ただし標的配列中の全てのチミジン (T) は、ウリジン(U) によって置換した (ガイドRNA2: 5' -ACUACAGCAUUGCUCAGUAC-3')。標的部位およびオフターゲット部位におけるAからGへの変換率を図58 Aに示す。例示的な配列リードが図58 Cに示されている。エディター (PV 15-28) を用いてG2019Sを編集した。

【1160】

LRRK2突然変異の補正に使用されたエディター (PV 15～28) の記述は以下のとおりである。

PV1 (also PV15). pCMV_monoABE8.1_bpNLS + Y147T

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNHRVEITEGILADECAALLCTFFRMPRQVFNAQKKAQSSTD

PV2 (also PV16). pCMV_monoABE8.1_bpNLS + Y147R

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL

MDVLHYPGMNHHRVEITEGILADECAALLCRFFRMPRQVFNAQKKAQSSTD
PV3 (also PV17). pCMV_monoABE8.1_bpNLS + Q154S
MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVNLRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNHHRVEITEGILADECAALLCYFFRMPRSVFNAQKKAQSSTD
PV4 (also PV18). pCMV_monoABE8.1_bpNLS + Y123H
MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVNLRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHHPGMNHHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTD
PV5 (also PV19). pCMV_monoABE8.1_bpNLS + V82S
MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVNLRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYSTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNHHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTD
PV6 (also PV20). pCMV_monoABE8.1_bpNLS + T166R
MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVNLRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNHHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSRD
PV7 (also PV21). pCMV_monoABE8.1_bpNLS + Q154R
MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVNLRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNHHRVEITEGILADECAALLCYFFRMPRRVFNAQKKAQSSTD
PV8 (also PV22). pCMV_monoABE8.1_bpNLS + Y147R_Q154R_Y123H
MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVNLRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHHPGMNHHRVEITEGILADECAALLCRFFRMPRRVFNAQKKAQSSTD
PV9 (also PV23). pCMV_monoABE8.1_bpNLS + Y147R_Q154R_I76Y
MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVNLRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLYDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNHHRVEITEGILADECAALLCRFFRMPRRVFNAQKKAQSSTD
PV10 (also PV24). pCMV_monoABE8.1_bpNLS + Y147R_Q154R_T166R
MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVNLRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNHHRVEITEGILADECAALLCRFFRMPRRVFNAQKKAQSSRD
PV11 (also PV25). pCMV_monoABE8.1_bpNLS + Y147T_Q154R
MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVNLRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNHHRVEITEGILADECAALLCTFFRMPRRVFNAQKKAQSSTD
PV12 (also PV26). pCMV_monoABE8.1_bpNLS + Y147T_Q154S
MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVNLRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNHHRVEITEGILADECAALLCTFFRMPRSVFNAQKKAQSSTD
PV13 (also PV27). pCMV_monoABE8.1_bpNLS + H123Y123H_Y147R_Q154R_I
76Y
MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVNLRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLYDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHHPGMNHHRVEITEGILADECAALLCRFFRMPRRVFNAQKKAQSSTD
PV14 (also PV28). pCMV_monoABE8.1_bpNLS + V82S + Q154R
MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVNLRVIGEGWNRAIGLHDPTAHA
EIMALRQGGLVMQNYRLIDATLYSTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSL
MDVLHYPGMNHHRVEITEGILADECAALLCYFFRMPRRVFNAQKKAQSSTD

10

20

30

40

50

【1161】

図59 A～59 Lは、R1441Cをコードする、LRRK2標的配列の位置7におけるAからGへの遷移についての例示的な配列リードを提供する。エディターが表示されている (PV 1～14)。

【1162】

図60 A～60 Wは、G2019Sをコードする、LRRK2標的配列の位置4および6におけるAからGへの遷移についての配列のリードを示す。エディターが表示されている (PV 15～28)。

【1163】

パーキンソン病に関連するLRRK2における他の病原性突然変異が、同様の戦略を用いて補正される(図61 A～61 D)。

【1164】

[実施例9：マウスアルファ-L-イズロニダーゼ (IDUA) 遺伝子におけるW401X変異の補正のためのHurler症候群塩基編集]

ハーラー (Hurler) 症候群は、ムコ多糖症1型 (MPS1) の最も重度の病型である。MPS1はアルファ-L-イズロニダーゼ (IDUA) 遺伝子の変異によって引き起こされる。現在、ヒトIDUA遺伝子を含むトランスジェニックマウスモデルは存在しない。しかし、Mus musculusのIDUAタンパク質アミノ酸配列とHomo sapiensのIDUAタンパク質アミノ酸配列との間には高い保存性がある。ヒトにおいて、重篤なHurler症候群の表現型と関連するMPS1の一般的な変異はW402Xである。この変異は、IDUA蛋白質の402番目に終止コドン (W402X) を導入する一塩基置換であり、ホモ接合体では極めて重度の臨床表現型と関連している。マウスでは、IDUAタンパク質の同等の突然変異はW401Xである。標的部位でA>Gを効率的に変換することによってW401X変異を補正することにより、ABEが、Mus musculus IDUA遺伝子の補正に使用され得る。SNPにおけるA>G補正は、位置401における終止コドン (W401X) をIDUAポリペプチドにおけるトリプトファンに変化させる。

【1165】

検証されたプロトスペーサー隣接モチーフ (PAM) 配列選択性を有するCas9部分を用いるA・TからG・CへのDNA塩基エディター (ABE) を用いて、W401X変異を野生型配列への復帰のために標的化した。どのガイドRNA (gRNA) とABE8-Cas9プラットフォームが標的化IDUA変異を最も効率的かつ正確に補正できるかを決定するために、Hurler症候群W401X標的化変異を有するMus musculus IDUAアリルをレンチウイルス形質導入によりHEK293T細胞のゲノムに組み入れた。上記のようにOptiMEM培地およびリポフェクタミン2000を用いて、250 ngのgRNAおよび750 ngのABE8バリエーション塩基エディター発現プラスミドでHEK293T細胞をトランスフェクトし、48ウェルプレートのウェル当たり3万細胞を平板培養化した。ABE8塩基エディターバリエーションはNGG PAM配列を含んだ (つまりSpCas9)。トランスフェクションの5日後 (トランスフェクション3日後に培地交換を伴う)、細胞を溶解し配列決定のために前処理し、miSeq分析によって、所望の部位における塩基編集について分析した。

【1166】

Mus musculus IDUAのDNA標的化/挿入配列は以下に示され、NCBI参照配列番号NM_008325.4の下で見出される代表的なMus musculus IDUA遺伝子配列の核酸1077～1358に対応する。

```
CTCCCAGCGCACACTTACTGCTCGATTCCAGGTCAACAATACTCACCCACCCACGTGC
AGTTGCTGCGAAAGCCAGTACTCACAGTCATGGGGCTCATGGCCCTGTTGGATGGAGAA
CAACTCTAGGCAGAGGTCTCAAAGGCTGGGGCTGTGTTGGACAGCAATCATACAGTGGG
TGTCCTGGCCAGCACCCATCACCCCTGAAGGCTCCGCGAGCGGCCTGGAGTACCACAGTCC
TCATCTACACTAGTGATGACACCCACGACACCCCAACCACAGTAT
```

【1167】

W401X突然変異についての上記Mus musculus DNA標的/挿入配列は、「A」核酸塩基

10

20

30

40

50

(太字および下線で示す)を含む一方、Mus musculus IDUA遺伝子配列は位置1202に「G」核酸塩基を含む。

【1168】

2つのガイドRNAを、Mus musculus IDUA W401X変異の塩基編集について試験した。gRNAは、足場配列と、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ところの疾患関連遺伝子のためのスパーサー配列（標的配列）を含む（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 p protein yields C:G-to-T:A base editors with higher efficiency and product purity” Science Advances 3:eaa04774 (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/s41576-018-0059-1参照)。

【1169】

ABE8塩基エディターバリエーションを用いて、IDUA W401X変異を標的として21ヌクレオチドガイドRNA (gRNA) をテストした（図62 Aおよび62 B）。上記DNA標的配列の相補鎖とハイブリダイズする21ヌクレオチドgRNA配列を以下に示す：gACTCTAGGCAGAG GTCTCAAAGG。NGG PAM配列（つまりSpCas9）には下線が引かれている。gRNA配列中の小文字の「g」は、ポリメラーゼ（例えばPol III）が転写を開始しなければならない配列におけるミスマッチを示す。ガイドRNAは、配列UUGAGACCUCUGCCUAGAGUを含む。

【1170】

上記gRNA配列について、足場配列は以下の通りである：
GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAG TGGCACCGAGTCGGTGCTTTTTTTT。

【1171】

使用されたABE8塩基エディターは、ABE8モノマー（monomer）バリエーション：ABE8.1、ABE8.12、ABE8.13、ABE8.4、ABE8.5、ABE8.6、ABE8.7、ABE8.8、ABE8.9、ABE8.10、ABE8.11、ABE8.12、およびABE8.13を含む。陽性対照塩基エディターABE7.10および陰性対照も比較のために使用した。

【1172】

ABE8塩基エディターバリエーションは、W401Xの約40%塩基編集補正で、ABE7.10陽性対照と比較して同等または増加した塩基編集活性を示した（図62 A）。図62 Bに示されるように、インデル形成のパーセントは約0.4~0.6%であり、ABE7.10陽性対照と同等であった。

【1173】

ABE8塩基エディターバリエーションを用いてIDUA W401X突然変異を標的として、20ヌクレオチドのガイドRNA (gRNA) も試験し、上述の21ヌクレオチドのgRNAと比較した（図63）。上記DNA標的配列の相補鎖とハイブリダイズする20ヌクレオチドgRNA配列を以下に示す：ACTCTAGGCAGAGGTCTCAA AGG。NGG PAM配列（つまりSpCas9）に下線を付している。

【1174】

上記gRNA配列について、足場配列は以下の通りである：
GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAG TGGCACCGAGTCGGTGCTTTTTTTT。

【1175】

使用されたABE8塩基エディターは、ABE8モノマー変異体：ABE8.1、ABE8.12、ABE8.13、ABE8.4、ABE8.5、ABE8.6、ABE8.7、ABE8.8、ABE8.9、ABE8.10、ABE8.11

10

20

30

40

50

、ABE8.12、およびABE8.13を含む。陽性対照塩基エディターABE7.10および陰性対照も比較のために使用した。

【1176】

ABE8塩基エディターバリエーションは、20ヌクレオチドgRNAまたは21ヌクレオチドgRNAのいずれを用いても、W401Xの約40%塩基編集補正で、同等の塩基編集活性を有していた(図63)。

【1177】

[実施例10：ヒトアルファ-L-イズロニダーゼ (IDUA) 遺伝子におけるW402X変異の補正のためのHurler症候群塩基編集]

ハーラー (Hurler) 症候群は、ムコ多糖症1型 (MPS1) の最も重度の病型である。MPS1はアルファ-L-イズロニダーゼ (IDUA) 遺伝子の変異によって引き起こされる。ヒトにおいて、重篤なHurler症候群の表現型と関連するMPS1の一般的な変異はW402Xである。この変異は、IDUA蛋白質の402番目に終止コドン (W402X) を導入する一塩基置換であり、ホモ接合体では極めて重度の臨床表現型と関連している。標的部位でA>Gを効率的に変換することによってW402X変異を補正することにより、ABEが、ヒトIDUA遺伝子の補正に使用され得る。

【1178】

検証されたプロトスペーサー隣接モチーフ (PAM) 配列選択性を有するCas9部分を用いるA・TからG・CへのDNA塩基エディター (ABE) を用いて、W402X変異を野生型配列への復帰のために標的化した。図64に示すように、ABE塩基エディターおよびガイドRNA (gRNA) を使用して、Homo sapiens IDUA核酸配列中のアデノシン (A) 核酸塩基 (囲み) を標的とし、W402X突然変異を補正することができる。SNPにおけるA>G補正は、位置402における終止コドン(W402X)をIDUAポリペプチドにおけるトリプトファンに変化させる。

【1179】

どのABE8-Cas9プラットフォームが標的化IDUA変異を最も効率的かつ正確に補正できるかを決定するために、Hurler症候群W402X標的化変異を有するHomo sapiens IDUAアレルをレンチウイルス形質導入によりHEK293T細胞のゲノムに組み入れた。上記のようにOptiMEM培地およびリポフェクタミン2000を用いて、250 ngのgRNAおよび750 ngのABE8バリエーション塩基エディター発現プラスミドでHEK293T細胞をトランスフェクトし、48ウェルプレートのウェル当たり3万細胞を平板培養化した。ABE8塩基エディターバリエーションはNGG PAM配列を含んだ (つまりSpCas9)。トランスフェクションの5日後 (トランスフェクション3日後に培地交換を伴う)、細胞を溶解し配列決定のために前処理し、miSeq分析によって、所望の部位における塩基編集について分析した。

【1180】

Homo sapiens IDUAポリヌクレオチド配列におけるDNA標的化/挿入配列は、以下に示されており、NCBI参照配列番号NM_000203.5の下で見出される代表的なHomo sapiens IDUA遺伝子配列の核酸1076~1358に対応する。

```
CCTTCGCGCAGCGCACGCTCACCGCGCGCTTCCAGGTCAACAACACCCGCCGCGGCAC  
GTGCAGCTGTTGCGCAAGCCGGTGCTCACGGCCATGGGGCTGCTGGCGCTGCTGGATGA  
GGAGCAGCTCTAGGCCGAAGTGTCGCAGGCCGGGACCGTCTGGACAGCAACCACACGG  
TGGGCGTCTTGGCCAGCGCCACCGCCCCAGGGCCCGGCGACGCCTGGCGCGCCGCG  
GTGCTGATCTACGCGAGCGACGACACCCGCGCCACCCCAACCGCAG
```

【1181】

上記のHomo sapiens標的/挿入配列は「A」核酸塩基 (太字および下線で示す) を含む一方、Homo sapiens IDUA遺伝子配列は、IDUA配列の位置1205に「G」核酸塩基を含む。

【1182】

20ヌクレオチドガイドRNA (gRNA) を、ABE8塩基エディターバリエーションによるW402X変異を標的としてテストした(図65 A)。gRNAは、足場配列と、本明細書に提供される

10

20

30

40

50

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ところの疾患関連遺伝子のためのスペーサー配列（標的配列）とを含む（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: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/s41576-018-0059-1参照）。

10

【1183】

上記のDNA標的配列の相補鎖とハイブリダイズするgRNA配列は：GCTCTAGGCCGAAGTGTCGC AGGである。NGG PAM配列(つまりSpCas9)には下線を付している。

【1184】

上記gRNA配列について、足場配列は以下の通りである。

GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTT

【1185】

使用されたABE塩基エディターはABE8モノマーバリエーション：ABE8.1、ABE8.2、ABE8.3、ABE8.4、ABE8.5、ABE8.6、ABE8.7、ABE8.8、ABE8.9、ABE8.10、ABE8.11、ABE8.12、およびABE8.13を含む。陽性対照塩基エディターABE7.10および陰性対照も比較のために使用した。

20

【1186】

ABE8塩基エディターバリエーションは、W402Xの約30～40%の塩基編集補正で、ABE7.10陽性対照と比較して塩基編集活性が同等または増加していた(図65 A)。図65 Bに示されるように、インデル形成のパーセントは、約0.2～0.5%で、ABE7.10陽性対照と同等またはそれ未満であった。

【1187】

PCR産物のディープシーケンシングによって検出される、IDUA酸配列の6位における標的A核酸塩基のA→G塩基編集の効率を図66 A～66 Oに示す。下記表18は、ABE7.10陽性対照(図66 N)および陰性対照(図66 O)と比較して、ABE8塩基エディターバリエーション(図66 A～66 M)によって達成されるIDUA核酸標的部位の位置6におけるAからGへの塩基編集のパーセントを要約している。

30

【1188】

表18：塩基エディターバリエーションを用いたIDUA標的部位塩基編集のパーセンテージ。

40

50

【表 18】

サンプル	ABE	A→G 塩基編集 (%)
NMG280-1	ABE8.1	31.17%
NMG280-2	ABE8.1	31.65%
NMG280-3	ABE8.1	33.61%
NMG281-1	ABE8.2	41.89%
NMG281-2	ABE8.2	40.96%
NMG281-3	ABE8.2	35.02%
NMG282-1	ABE8.3	34.07%
NMG282-2	ABE8.3	31.86%
NMG282-3	ABE8.3	30.78%
NMG283-1	ABE8.4	32.72%
NMG283-2	ABE8.4	31.87%
NMG283-3	ABE8.4	27.74%
NMG284-1	ABE8.5	8.11%
NMG284-2	ABE8.5	10.02%
NMG284-3	ABE8.5	9.5%
NMG285-1	ABE8.6	39%
NMG285-2	ABE8.6	31.94%
NMG285-3	ABE8.6	37.55%
NMG286-1	ABE8.7	39.84%
NMG286-2	ABE8.7	32.14%
NMG286-3	ABE8.7	40.85%
NMG287-1	ABE8.8	38.92%
NMG287-2	ABE8.8	38.87%
NMG287-3	ABE8.8	35.53%
NMG288-1	ABE8.9	36.73%
NMG288-2	ABE8.9	38.97%
NMG288-3	ABE8.9	28.23%
NMG289-1	ABE8.10	34.1%
NMG289-2	ABE8.10	33.7%
NMG289-3	ABE8.10	39.92%
NMG290-1	ABE8.11	41.41%

10

20

30

40

50

サンプル	ABE	A→G 塩基編集 (%)
NMG290-2	ABE8.11	36.78%
NMG290-3	ABE8.11	39.83%
NMG291-1	ABE8.12	37.87%
NMG291-2	ABE8.12	45.05%
NMG291-3	ABE8.12	35.94%
NMG292-1	ABE8.13	34.44%
NMG292-2	ABE8.13	33.73%
NMG292-3	ABE8.13	39.7%
NP20-1	ABE7.10	33.73%
NP20-2	ABE7.10	33.93%
NP20-3	ABE7.10	26.89%
Negct-1	Negative Control	0.04%
Negct-2	Negative Control	0.03%

【1189】

表18および図66 A～66 Mに示されるように、ABE8塩基エディターバリエーションを使用し、IDUA核酸配列の位置6すなわち標的「A」核酸塩基部位において、平均約33.9%のA→G塩基編集が達成された。これは陽性対照ABE7.10(図66 N)と同等であった。

【1190】

[実施例11：細胞培養とトランスフェクション]

HEK293T (293T) 細胞株はAmerican Tissue Culture Collection (ATCC) から入手した。293T細胞は、10%ウシ胎児血清および1%ペニシリン/ストレプトマイシンを添加したDMEM中、37℃、5% CO₂で維持した。全ての細胞株を、製造業者の指示に従ってリポフェクタミン2000 (Invitrogen) を用いて24ウェルプレート中でトランスフェクトした。リポフェクションに用いたDNA量は1 μg/ウェルであった。トランスフェクション効率は、対照GFP発現プラスミドの送達後に蛍光顕微鏡によって測定したところ、293T細胞について80%よりも常に高かった。

【1191】

プラスミドトランスフェクションのために、HEK293T細胞を平板培養し、OptiMEM培地およびリポフェクタミン2000を用いて、U6プロモーターを含みgRNAをコードする250 ngの発現プラスミドと、Cas9/ABE8バリエーション塩基エディターをコードする750 ngの発現プラスミドとでトランスフェクトした。使用したABE8塩基エディターバリエーションにはN GG PAM配列が含まれていた。細胞を37℃、5% CO₂で5日間維持し、トランスフェクション後3日目には培地を交換した。その後、細胞を溶解し、ゲノムDNAを単離し、典型的には20～100 ngの鋳型DNAを用いて標準的な手順を用いてPCRを行った。アダプター (Illumina) を加えた後、DNAをディープシーケンシングした。所望の部位における塩基編集をMiSeq分析により分析した。

【1192】

293T細胞の複製トランスフェクションから採取したゲノムDNAまたはRNAからのPCRアンプリコンについて、ディープシーケンシングを行った。ゲル電気泳動によってPCR

産物の質を検証した後、PCR産物を、例えばZymo Research社のZymo clean Gel DNA Recovery Kitを用いたゲル抽出によって単離した。剪断せずにショットガンライブラリーを調製した。ライブラリーをqPCRによって定量し、MiSeq 500サイクル配列決定キットバージョン2を用いて、断片の各末端から、251サイクルに渡り一つのMiSeq Nanoフローセル上で配列決定した。Fastqファイルが生成され、bcl2fastq v2.17.1.14変換ソフトウェア (Illumina) によって脱多重化された。

【1193】

[実施例12：レット症候群突然変異の塩基編集補正]

R106WおよびR255Xを含む、6つのRett突然変異を有するMECP2の単一コピーを含むレンチ-HEK株を作製し、ガイドRNA配列およびABEバリエーションのスクリーニングに用いた。ガイドおよびABE発現のためのプラスミドをレンチ-HEK株にトランスフェクトし、ゲノムDNAを収集し、標的部位における突然変異についてNextGen配列決定により分析した。図67および図68を参照すると、ガイド1は全体的に最高の編集を示し、ABE8.8、ABE8.9、およびABE8.13はガイド1および2の両方と良好に協働した。ガイドRNA配列はスペーサーCTTTTCACTTTTCTGCGGGG (R255X)、AGCTTCCATGTCCAGCCTTC (R106W)、ACCATGAAGTCAAAATCATT (T158M)、またはGCTTTCAGCCCCGTTTCTTG (R270X)を含む。

【1194】

異なるPAM認識特異性を有する塩基エディターバリエーション (variant) を試験する；PAMの変化に対応するCas9の置換を表19に示す。*バリエーション5は、R255Xでの25%編集に相当する。この方法は、上述のR106Wの方法に類似しており、ABE上の変異は他のグループによって生成されている。図69および下記表19を参照すると、バリエーション5に寄与する変異が、この特定の遺伝子座において最も編集量が多いことを示した。

【1195】

10

20

30

40

50

【表 19】

variant	E1219V	R1335Q	T1337	G1218
1	F	V	T	
2	F	V	R	
3	F	V	Q	
4	F	V	L	
5	F	V	T	R
6	F	V	R	R
7	F	V	Q	R
8	F	V	L	R
9	L	L	T	
10	L	L	R	
11	L	L	Q	
12	L	L	L	
13	F	I	T	
14	F	I	R	
15	F	I	Q	
16	F	I	L	
17	F	G	C	
18	H	L	N	
19	F	G	C	A
20	H	L	N	V
21	L	A	W	
22	L	A	F	
23	L	A	Y	
24	I	A	W	
25	I	A	F	
26	I	A	Y	

【1196】

【実施例13：Hurler/IDUA変異補正】

Hurler突然変異を有するMECP2の単一コピーを含有するレンチ-HEK株を作製し、上記のようにガイドRNA配列およびバリエーションをスクリーニングするために使用した。図70 A～70 Cを参照して、遺伝性IDUA機能喪失 (W402X) 変異の補正を、塩基エディターバリエーションを用いて検討した。2人のW402Xホモ接合性Hurler患者(GM06214およびGM00798)ならびに非罹患ヘテロ接合性親 (GM00799) からの線維芽細胞をCoriellから得た。患者由来およびBJ線維芽細胞を、ABE8.8のためのmRNAおよびヒトW402Xガイドでエレクトロポレーションした。ゲノムDNAをQuickExtract溶解溶液で抽出し、AからGへの編集を評価するためにNextGen配列決定で配列決定した。イズロニダーゼ活性は分光光度法により評価した。溶解物を酸性緩衝液中で4-メチルウンベリフェロン-イズロニドと共に2時間インキュベートし、次いでアルカリ溶液でクエンチした。4-メチルウンベリフェロン開裂を蛍光(365 nm励起、445 nm発光)により測定した。患者の線維芽細胞では高い編集が観察され、これは非罹患ヘテロ接合個体であるGM00799で観察されたものに匹敵する酵素活性の増加をもたらした。

【1197】

【実施例14：ABE8.8によるin vivo塩基編集】

図70を参照すると、分割インテインABE8.8およびマウスROSA26遺伝子座のためのgRNAをコードするウイルスゲノムをAAV9およびPHP.eBカプシドにパッケージした。C57BL/6マウスの側脳室に9e11総vgのAAV9 (4.5e11の各インテイン分割体)と7.5e11総vg PHP.eB (3.75e11の各インテイン分割体)を注入した。6週間後、脳と脊髄を採取し、ゲノムDNAの抽出と配列決定のために解剖した。AAV9形質導入は海馬で最も高く、それは

側脳室に最も近い構造であり、13%編集に達した。

【1198】

〔実施例15：ABE8バリエーション塩基編集〕

ABCA4におけるG1961E変異を復帰させるための最適な塩基エディターを決定するために、我々が以前にこの標的部位上の最適なスパーサー長として示した21 ntスパーサー配列を有するsgRNAを用いて(図73)、レンチウイルスノックインモデル細胞株において、疾患アリルでのA→G塩基変換を行うために、40のユニークなABE8バリエーションをABE7.10と比較した(図72 A~72 B)。全てのバリエーションは、疾患アリルとゆらぎ(wobble)塩基で測定可能なA→G編集を提供した。ゆらぎ塩基における編集はサイレント変異をもたらし、有害ではない。最も性能の良いバリエーションのうちの6つをその後コドン最適化し、さらなる検証のために分割AAVシステムに組み込んだ。塩基エディター、sgRNA、および発現調節エレメントをコードするDNA配列のサイズは、単一のAAV粒子のパッケージング限界を超えるので、必要な部分は、2つのAAV粒子の間で分割され同時送達され得る。この送達方法では、塩基エディターをコードする遺伝子を2つのウイルスに分割し、分割インテインを用いて、同時感染後に全長タンパク質を再構成する(図74)。野生型TadAドメインを欠く分割ABE8バリエーション(ABE8-m)をAAV2ベクターの対にパッケージ化し、そのうちの1つは、関心のある野生型ABCA4部位を標的とするsgRNAの単一コピーもコードする。この実験では、野生型細胞には存在しない疾患アリルの代理(surrogate)として、ABCA4 G1961コドンのゆらぎ塩基において編集を評価した。野生型ARPE-19細胞を二重AAVで共形質導入し、塩基編集率を対象の21 nt標的配列(図75 A~75 B)上で評価した。ABEバリエーション7.9、7.10、8.5-m、8.8-m、8.9-m及び8.18-mは代理部位8Aの変換において同等に効率的であったが、バリエーション8.8-m、8.9-m及び8.18-mは位置5Cで望ましくないC→T変換も触媒した。野生型TadAドメインが除去されたABE7.10のバリエーション(ABE7.10-m)は、親のABE7.10バリエーションと比較して、見かけ上50%の活性損失を有した。これらの結果は、バリエーションABE8.5-mが、野生型TadAドメインも欠如して関心部位で最も効率的なエディターであることを示し、これは、塩基エディターの総サイズをDNAで594 bpあるいは198アミノ酸残基も減少させる。

【1199】

ガイドRNA配列は、配列GCTGTGTGTCTGAAGTTTCGCCCTGGAGAGGTTG またはGCTGTGTGTCTGGAGTTTCGCCCTGGAGAGGTTGにてABCA4遺伝子を標的化し、ここでPAM配列には下線が引かれている。ガイドRNAは、配列CACCUCUCCAGGGCGAACUUCGACACACAGC またはCACCUCUCCAGGGCGAACUCCGACACACAGCを含む。

【1200】

ABCA4 G1961E遺伝子座を標的とする21 ntスパーサー長sgRNAを用いたヒトゲノム内でのオフターゲット塩基編集の可能性を評価した。SpCas9 NGG PAMとマッチする3'配列に先立つABCA4 G1961E gRNAプロトスパーサー配列に対する不完全マッチングの全てについてヒト参照ゲノム(GRCh38)をコンピュータ上でスキャンすることによって、ゲノム内の潜在的なオフターゲット部位のin silico予測を行った。最大5個のミスマッチおよび単一のRNAまたはDNAバルジを含む全配列を評価した。(a) ミスマッチの数が少ないこと、および(b) コーディングエクソン(GENCODE転写産物アノテーションによって決定)およびがん関連遺伝子(COSMICがん遺伝子センサスで報告されたもの)との重複に基づいて、潜在的なオフターゲット部位を実験的評価のために優先順位付けした。21 ntスパーサー配列に対するミスマッチが3個以下である予測オフターゲットは、in silico解析ではされるゲノム中に見つからなかった。野生型ARPE-19細胞に、ABE7.10とABCA4 G1961E疾患アリルを標的とする21 ntスパーサーsgRNAとを同時送達するためにデュアルAAVシステムを使用し、28のin silico予測オフターゲット部位の標的化アンプリコン配列決定によって編集を評価した。予測されたオフターゲット部位のいずれも、未処理細胞と比較して処理細胞において有意に塩基編集されてはならず(図76 A~76 B)、オフターゲット部位でもオンターゲット部位でも有意なインデルは見出されなかった(図77)。観察された唯一の有意な編集は、処理細胞におけるABCA4 G1961ゆらぎ塩基において生じ

、このことは、ABCA4 G1961E疾患アリルを標的とするsgRNAはこれらの細胞に存在する野生型アリルと1つのミスマッチ塩基対しか含まないため予測されたことだった。これらの結果は、このsgRNAが、評価されたin silico予測オフターゲット部位のいずれにおいてもオフターゲットDNA編集を促進しないことを示す。

【1201】

【実施例20：霊長類網膜の実施例】

非ヒト霊長類の眼を死後1～2時間で採取し、死後4～8時間の間に培養した。6 mmの生検パンチを用いて神経網膜全体からパンチを採取した。光受容体側を下にした網膜を6ウェル組織培養プレート中のヌクレオポア膜の上に置いた。ベクター (10 ul、1.26E+12 v g/ml) を神経網膜と上記膜の間にピペットで注入し、網膜組織下に液胞 (bleb) を形成させた。培地を3日毎に交換し、組織を0～22日間インキュベートした。異なる時点で組織を採取し、10%中性緩衝ホルマリン中で固定し、組織学的検査のために処理した。

【1202】

(霊長類網膜の完全性)

切片を、抗ロドプシン、抗GFP、およびビオチン化ピーナッツ凝集素抗体で4℃で一晩免疫標識した。PBSで洗浄した後、サンプルを二次抗体と室温で1時間インキュベートした。スライドをPBS中で洗浄し、DAPIを含有するグリセロール系液体マウント剤でマウントした。

【1203】

非ヒト霊長類からの網膜外植片を0日目 (D0) および22日目に採取した。D0とD22の非形質導入網膜外植片を比較した組織学的染色は、22日間まで培養された場合でも網膜の細胞型が保存されることを示した。D22では、光受容体特異的GFP発現ベクター (Anc80 L65.hGRK.eGFP) と比較して、Anc80L65.CMV.eGFPに曝露された網膜培養ではGFP発現が定性的に明かった。Anc80L65.hGRK1.eGFPによる網膜外植片の形質導入は、GFPが光受容体含有外核層 (ONL) にのみ存在することを示し、hGRK1プロモーターの光受容体特異的活性を確認した。図78参照。

【1204】

(NHPにおけるCas9発現)

切片を、マウスおよびウサギモノクローナルCas9抗体で4℃で一晩免疫標識した。PBSで洗浄した後、サンプルを二次抗体と室温で1時間インキュベートした。スライドをPBS中で洗浄し、DAPIを含有するグリセロール系液体マウント剤でマウントした。

【1205】

最適化された分割ABE (ABE7.10、8.5、8.9) 塩基エディターをコードするデュアルAAV2粒子を、非ヒト霊長類網膜外植片で試験した。全長塩基エディターの再構成を試験するために、各塩基エディター-分割インテイン半分を発現するAAV粒子による同時感染後に、異なる時点で組織を収集し、Cas9のNおよびC末端について染色した。Cas9 N末端 (緑染色) とC末端 (赤染色) の両方の発現が早くも感染後6日目に観察し、17日目まで維持されたことから、塩基エディターの編集活性ウィンドウの可能性が示唆された。これらの結果は、二重AAV分割インテイン塩基エディターが非ヒト霊長類網膜外移植片においてCas9を発現することを立証する。図79参照。

【1206】

本開示は以下の実施形態を含む。

実施形態1

対象における神経障害を治療する方法であって、(i) アデノシン塩基エディターまたはそれをコードする核酸配列と、(ii) ガイドポリヌクレオチドまたはそれをコードする核酸配列とを対象に投与することを含み、

前記アデノシン塩基エディターは、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み、

前記アデノシンデアミナーゼドメインは、配列番号2における番号付けでアミノ酸位置82もしくは166またはそれらに対応する位置でのアミノ酸置換を含み、

10

20

30

40

50

前記ガイドポリヌクレオチドは、前記アデノシン塩基エディターを誘導して対象の前記神経障害に関連する標的遺伝子またはその調節エレメントにおいてAからGへの核酸塩基改変をもたらし、それによって対象の前記神経障害を治療する、

方法。

実施形態 2

前記標的遺伝子がアルファ-L-イゾロニダーゼ (IDUA) 遺伝子であり、前記神経疾患がハーラー症候群である、実施形態1記載の方法。

実施形態 3

対象におけるハーラー症候群を治療する方法であって、(i) アデノシン塩基エディターまたはそれをコードする核酸配列と、(ii) ガイドポリヌクレオチドまたはそれをコードする核酸配列とを対象に投与することを含み、

前記アデノシン塩基エディターは、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み、

前記アデノシンデアミナーゼドメインは、配列番号2における番号付けでアミノ酸位置82もしくは166またはそれらに対応する位置でのアミノ酸置換を含み、

前記ガイドポリヌクレオチドは、前記アデノシン塩基エディターを誘導して対象のアルファ-L-イゾロニダーゼ (IDUA) 遺伝子またはその調節エレメントにおいてAからGへの核酸塩基改変をもたらし、それによって対象のハーラー症候群を治療する、

方法。

実施形態 4

前記投与が、ハーラー症候群に関連する少なくとも1つの症状を改善する、実施形態2または3に記載の方法。

実施形態 5

前記投与が、アデノシンデアミナーゼにおける前記アミノ酸置換を伴わない塩基エディターを用いた治療と比較して、ハーラー症候群に関連する少なくとも1つの症状のより迅速な改善をもたらす、実施形態4に記載の方法。

実施形態 6

前記IDUA遺伝子またはその調節エレメントが、ハーラー症候群に関連するSNPを含む、実施形態2～5のいずれか一項記載の方法。

実施形態 7

前記AからGへの核酸塩基改変が、ハーラー症候群に関連するSNPにおけるものである、実施形態2～6のいずれか一項記載の方法。

実施形態 8

ハーラー症候群に関連するSNPが、配列番号4における番号付けでIDUAポリペプチドにおけるW402XもしくはW401Xアミノ酸変異体またはIDUA遺伝子によってコードされるそのバリエーションをもたらす、ここでXは終止コドンである、実施形態6または7記載の方法。

実施形態 9

前記AからGへの核酸塩基改変が、ハーラー症候群に関連するSNPを野生型核酸塩基に変化させる、実施形態6～8のいずれか一項に記載の方法。

実施形態 10

前記AからGへの核酸塩基改変が、ハーラー症候群に関連するSNPを、ハーラー症候群の1つ以上の症状の改善をもたらす非野生型核酸塩基に変化させる、実施形態6～8のいずれか一項に記載の方法。

実施形態 11

ハーラー症候群に関連するSNPにおけるAからGへの改変が、IDUA遺伝子によってコードされる終止コドンをIDUAポリペプチドにおけるトリプトファンに変化させる、実施形態6～10のいずれか一項に記載の方法。

実施形態 12

前記ガイドポリヌクレオチドが、ハーラー症候群に関連するSNPを含むIDUA遺伝子また

10

20

30

40

50

はその調節エレメントに相補的な核酸配列を含む、実施形態6～10のいずれか一項に記載の方法。

実施形態13

前記アデノシン塩基エディターが、ハーラー症候群に関連するSNPを含むIDUA遺伝子またはその調節エレメントに相補的な核酸配列を含む単一ガイドRNA (sgRNA) と複合体をなす、実施形態6～12のいずれか一項に記載の方法。

実施形態14

前記sgRNAが、5' - GACUCUAGGCAGAGGUCUCAA -3'、5' - ACUCUAGGCAGAGGUCUCAA-3'、5' - CUCUAGGCCGAAGUGUCGC -3'、および5' -GCUCUAGGCCGAAGUGUCGC-3' からなる群より選択される核酸配列を含む、実施形態13記載の方法。

10

実施形態15

前記標的遺伝子がロイシンリッチリピートキナーゼ2 (LRRK2) 遺伝子であり、神経疾患がパーキンソン病である、実施形態1記載の方法。

実施形態16

対象におけるパーキンソン病を治療する方法であって、(i) アデノシン塩基エディターまたはそれをコードする核酸配列と、(ii) ガイドポリヌクレオチドまたはそれをコードする核酸配列とを対象に投与することを含み、

前記アデノシン塩基エディターは、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み、

前記アデノシンデアミナーゼドメインは、配列番号2における番号付けでアミノ酸位置82もしくは166またはそれらに対応する位置でのアミノ酸置換を含み、

20

前記ガイドポリヌクレオチドは、前記アデノシン塩基エディターを誘導して対象のロイシンリッチリピートキナーゼ2 (LRRK2) 遺伝子またはその調節エレメントにおいてAからGへの核酸塩基改変をもたらす、それによって対象のパーキンソン病を治療する、

方法。

実施形態17

前記投与がパーキンソン病に関連する少なくとも一つの症状を改善する、実施形態1、15または16のいずれか一項に記載の方法。

実施形態18

前記投与が、アデノシンデアミナーゼにおける前記アミノ酸置換を伴わない塩基エディターを用いた治療と比較して、パーキンソン病に関連する少なくとも一つの症状のより迅速な改善をもたらす、実施形態17に記載の方法。

30

実施形態19

LRRK2遺伝子またはその調節エレメントがパーキンソン病に関連するSNPを含む、実施形態15～18のいずれか一項記載の方法。

実施形態20

前記AからGへの核酸塩基改変が、パーキンソン病に関連するSNPにおけるものである、実施形態19記載の方法。

実施形態21

パーキンソン病に関連するSNPが、配列番号3における番号付けでLRRK2ポリペプチドにおけるA419V、R1441C、R1441H、もしくはG2019Sアミノ酸変異体またはLRRK2遺伝子によってコードされるそのバリエーションをもたらす、実施形態19または20に記載の方法。

40

実施形態22

前記AからGへの核酸塩基の改変が、パーキンソン病に関連するSNPを野生型核酸塩基に変化させる、実施形態19～21のいずれか一項記載の方法。

実施形態23

前記AからGへの核酸塩基の改変が、パーキンソン病に関連するSNPを、パーキンソン病の1つ以上の症状の改善をもたらす非野生型核酸塩基に変化させる、実施形態19～21のいずれか一項に記載の方法。

50

実施形態 2 4

前記AからGへの核酸塩基の改変が、LRRK2遺伝子によってコードされるLRRK2ポリペプチド中のシステインまたはヒスチジンをアルギニンに変化させる、実施形態19～22のいずれか一項記載の方法。

実施形態 2 5

前記AからGへの改変が、LRRK2遺伝子によってコードされるLRRK2ポリペプチドにおいてセリンをグリシンに変化させる、実施形態19～22のいずれか一項記載の方法。

実施形態 2 6

前記AからGへの改変が、配列番号3における番号付けでLRRK2ポリペプチドまたはLRRK2遺伝子によってコードされるそのバリエーションの位置144でシステイン (C) またはヒスチジン (H) をアルギニン (R) に置換するか、または位置2019でセリンをグリシン (G) に置換する、実施形態19～22のいずれか一項記載の方法。

10

実施形態 2 7

対象におけるパーキンソン病を治療する方法であって、(i) アデノシン塩基エディターまたはそれをコードする核酸配列と、(ii) ガイドポリヌクレオチドまたはそれをコードする核酸配列とを対象に投与することを含み、

前記アデノシン塩基エディターは、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み、

前記ガイドポリヌクレオチドは、前記アデノシン塩基エディターを誘導してパーキンソン病に関連するLRRK2遺伝子におけるSNPにおいてAからGへの核酸塩基改変をもたらし、

20

前記SNPは、配列番号3における番号付けでLRRK2ポリペプチドにおけるG2019S変異体またはそのバリエーションをコードしない、

方法。

実施形態 2 8

前記アデノシンデアミナーゼドメインが、配列番号2における番号付けでアミノ酸位置82もしくは166またはそれに対応する位置におけるアミノ酸置換を含む、実施形態27記載の方法。

実施形態 2 9

前記ガイドポリヌクレオチドが、パーキンソン病に関連するSNPを含むLRRK2遺伝子またはその調節エレメントに相補的な核酸配列を含む、実施形態15～28のいずれか一項に記載の方法。

30

実施形態 3 0

前記アデノシン塩基エディターが、パーキンソン病に関連するSNPを含むLRRK2遺伝子またはその調節エレメントに相補的な核酸配列を含む単一ガイドRNA (sgRNA) と複合体をなす、実施形態15～28のいずれか一項に記載の方法。

実施形態 3 0

前記sgRNAが核酸配列5' -AAGCGCAAGCCUGGAGGGAA -3' ; または 5' -ACUACAGCAUUGCUCAGUAC -3' を含む、実施形態30記載の方法。

実施形態 3 1

前記標的遺伝子がメチルCpG結合タンパク質2 (MECP2) 遺伝子であり、神経疾患がレット症候群である、実施形態1記載の方法。

40

実施形態 3 3

対象におけるレット症候群を治療する方法であって、(i) アデノシン塩基エディターまたはそれをコードする核酸配列と、(ii) ガイドポリヌクレオチドまたはそれをコードする核酸配列とを対象に投与することを含み、

前記アデノシン塩基エディターは、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み、

前記アデノシンデアミナーゼドメインは、配列番号2における番号付けでアミノ酸位置82もしくは166またはそれらに対応する位置でのアミノ酸置換を含み、

前記ガイドポリヌクレオチドは、前記アデノシン塩基エディターを誘導して対象のメチ

50

ルCpG結合タンパク質2 (MECP2) 遺伝子またはその調節エレメントにおいてAからGへの核酸塩基改変をもたらし、それによって対象のレット症候群を治療する、

方法。

実施形態34

前記投与が、レット症候群に関連する少なくとも一つの症状を改善する、実施形態32または33に記載の方法。

実施形態35

前記投与が、アデノシンデアミナーゼにおける前記アミノ酸置換を伴わない塩基エディターを用いた治療と比較して、レット症候群に関連する少なくとも一つの症状のより迅速な改善をもたらす、実施形態34に記載の方法。

実施形態36

前記MECP2遺伝子またはその調節エレメントが、レット症候群に関連するSNPを含む、実施形態32～35のいずれか一項記載の方法。

実施形態37

前記AからGへの核酸塩基改変が、レット症候群に関連するSNPにおけるものである、実施形態36記載の方法。

実施形態38

前記レット症候群に関連するSNPが、配列番号5における番号付けでMECP2ポリペプチドにおけるR106WもしくはT158Mアミノ酸変異体またはMECP2遺伝子によってコードされるそのバリエーションをもたらす、実施形態36または37に記載の方法。

実施形態39

前記レット症候群に関連するSNPが、MECP2遺伝子によってコードされるMECP2ポリペプチドにおけるR255XまたはR270Xアミノ酸変異体をもたらす、ここでXは終止コドンである、実施形態36または37記載の方法。

実施形態40

前記AからGへの核酸塩基改変が、レット症候群に関連するSNPを野生型核酸塩基に変化させる、実施形態36～39のいずれか一項記載の方法。

実施形態41

前記AからGへの核酸塩基改変が、レット症候群に関連するSNPを、レット症候群症状の改善をもたらす非野生型核酸塩基に変化させる、実施形態36～39のいずれか一項記載の方法。

実施形態42

レット症候群に関連するSNPにおける前記AからGへの核酸塩基改変が、終止コドンをMECP2ポリペプチドにおけるトリプトファンに変化させる、実施形態36～39のいずれか一項に記載の方法。

実施形態43

前記ガイドポリヌクレオチドが、レット症候群と関連するSNPを含むMECP2遺伝子またはその調節エレメントに対して相補的な核酸配列を含む、実施形態36～42のいずれか一項に記載の方法。

実施形態44

前記アデノシン塩基エディターが、レット症候群と関連するSNPを含むMECP2遺伝子またはその調節エレメントに相補的な核酸配列を含む単一ガイドRNA (sgRNA) と複合体をなす、実施形態36～42のいずれか一項に記載の方法。

実施形態45

前記ガイドポリヌクレオチドが、5' - CUUUUCACUUCUGCCGGGG-3'、5' - AGCUUCCAUGUCCAGCCUUC-3'、5' - ACCAUGAAGUCAAAAUCAUU-3'、および5' - GCUUUCAGCCCCGUUUCUUG-3' からなる群より選択される核酸配列を含む、実施形態44に記載の方法。

実施形態46

前記標的遺伝子がATP結合カセットサブファミリーメンバー4 (ABCA4) 遺伝子であり、

10

20

30

40

50

神経疾患がシュタルガルト病である、実施形態1記載の方法。

実施形態 4 7

対象におけるシュタルガルト病を治療する方法であって、(i) アデノシン塩基エディターまたはそれをコードする核酸配列と、(ii) ガイドポリヌクレオチドまたはそれをコードする核酸配列とを対象に投与することを含み、

前記アデノシン塩基エディターは、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み、

前記アデノシンデアミナーゼドメインは、配列番号2における番号付けでアミノ酸位置82もしくは166またはそれらに対応する位置でのアミノ酸置換を含み、

前記ガイドポリヌクレオチドは、前記アデノシン塩基エディターを誘導して対象のATP結合カセットサブファミリーメンバー4 (ABCA4) 遺伝子またはその調節エレメントにおいてAからGへの核酸塩基改変をもたらし、それによって対象のシュタルガルト病を治療する、

方法。

実施形態 4 8

前記投与が、シュタルガルト病に関連する少なくとも一つの症状を改善する、実施形態46または47に記載の方法。

実施形態 4 9

前記投与が、アデノシンデアミナーゼにおける前記アミノ酸置換を伴わない塩基エディターを用いた治療と比較して、シュタルガルト病に関連する少なくとも一つの症状のより迅速な改善をもたらす、実施形態48に記載の方法。

実施形態 5 0

ABCA4遺伝子が、シュタルガルト病に関連するSNPを含む、実施形態46～49のいずれか一項記載の方法。

実施形態 5 1

前記AからGへの核酸塩基改変が、シュタルガルト病に関連するSNPにおけるものである、実施形態50記載の方法。

実施形態 5 2

シュタルガルト病に関連する前記SNPが、配列番号6における番号付けでABCA4ポリペプチドにおけるA1038VもしくはG1961Eアミノ酸変異体またはABCA4遺伝子によってコードされるそのバリエーションをもたらす、実施形態50または51に記載の方法。

実施形態 5 3

シュタルガルト病に関連する前記SNPが、配列番号6における番号付けでABCA4ポリペプチドにおけるG1961Eアミノ酸変異体またはそのバリエーションをもたらす、実施形態52記載の方法。

実施形態 5 4

前記AからGへの核酸塩基改変が、シュタルガルト病に関連するSNPを野生型核酸塩基に変化させる、実施形態51～53のいずれか一項記載の方法。

実施形態 5 5

前記AからGへの核酸塩基改変が、シュタルガルト病に関連するSNPを、シュタルガルト病の1つ以上の症状の改善をもたらす非野生型核酸塩基に変化させる、実施形態51～53のいずれか一項に記載の方法。

実施形態 5 6

前記ガイドポリヌクレオチドが、シュタルガルト病に関連するSNPを含むABCA4遺伝子またはその調節エレメントに相補的な核酸配列を含む、実施形態51～55のいずれか一項記載の方法。

実施形態 5 7

前記アデノシン塩基エディターが、シュタルガルト病に関連するSNPを含むABCA4遺伝子またはその調節エレメントに相補的な核酸配列を含む単一ガイドRNA (sgRNA) と複合体をなす、実施形態51～56のいずれか一項に記載の方法。

10

20

30

40

50

実施形態 5 8

前記sgRNAが配列5'-CUCCAGGGCGAACUUCGACACACAGC-3'を含む、実施形態57記載の方法。

実施形態 5 9

前記治療が、前記アミノ酸置換を伴わない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用いた治療と比較して、前記神経障害の症状の改善をもたらす、実施形態1～58のいずれか一項に記載の方法。

実施形態 6 0

神経障害に関連する標的遺伝子またはその調節エレメントを編集する方法であって、前記標的遺伝子またはその調節エレメントを、(i) アデノシン塩基エディターと、(ii) ガイドポリヌクレオチドとに接触させることを含み、

前記アデノシン塩基エディターは、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み、

前記アデノシンデアミナーゼドメインは、配列番号2における番号付けでアミノ酸位置8もしくは166またはそれらに対応する位置でのアミノ酸置換を含み、

前記ガイドポリヌクレオチドは、前記アデノシン塩基エディターを誘導して前記神経障害に関連する標的遺伝子またはその調節エレメントにおいてAからGへの核酸塩基改変をもたらす、

方法。

実施形態 6 1

標的遺伝子がロイシンリッチリピートキナーゼ2 (LRRK2) 遺伝子であり、神経疾患がパーキンソン病である、実施形態60記載の方法。

実施形態 6 2

ロイシンリッチリピートキナーゼ2 (LRRK2) 遺伝子またはその調節エレメントを編集する方法であって、LRRK2遺伝子またはその調節エレメントを、(i) アデノシン塩基エディターまたはそれをコードする核酸配列と、(ii) ガイドポリヌクレオチドまたはそれをコードする核酸配列とに接触させることを含み、

前記アデノシン塩基エディターは、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み、

前記アデノシンデアミナーゼドメインは、配列番号2における番号付けでアミノ酸位置8もしくは166またはそれらに対応する位置でのアミノ酸置換を含み、

前記ガイドポリヌクレオチドは、前記アデノシン塩基エディターを誘導してLRRK2遺伝子またはその調節エレメントにおいてAからGへの核酸塩基改変をもたらす、

方法。

実施形態 6 3

前記AからGへの核酸塩基改変が、パーキンソン病に関連するSNPにおけるものである、実施形態62記載の方法。

実施形態 6 4

前記パーキンソン病に関連するSNPが、配列番号3における番号付けでLRRK2ポリペプチドにおけるA419V、R1441C、R1441H、もしくはG2019Sアミノ酸変異体またはLRRK2遺伝子によってコードされるそのバリエーションをもたらす、実施形態62または63に記載の方法。

実施形態 6 5

前記AからGへの核酸塩基改変が、パーキンソン病に関連するSNPを野生型核酸塩基に変化させる、実施形態62～64のいずれか一項に記載の方法。

実施形態 6 6

前記AからGへの核酸塩基改変が、パーキンソン病に関連するSNPを、パーキンソン病の1つ以上の症状の改善をもたらす非野生型核酸塩基に変化させる、実施形態62～64のいずれか一項に記載の方法。

実施形態 6 7

10

20

30

40

50

前記AからGへの核酸塩基の改変が、LRRK2遺伝子によってコードされるLRRK2ポリペプチド中のシステインまたはヒスチジンをアルギニンに変化させる、実施形態62～66のいずれか一項記載の方法。

実施形態68

前記AからGへの改変が、LRRK2遺伝子によってコードされるLRRK2ポリペプチドにおいてセリンをグリシンに変化させる、実施形態62～66のいずれか一項に記載の方法。

実施形態69

前記AからGへの改変が、配列番号3における番号付けでLRRK2ポリペプチドまたはLRRK2遺伝子によってコードされるそのバリエーションの位置144のシステイン (C) またはヒスチジン (H) をアルギニン (R) に置換するか、または位置2019のセリンをグリシン (G) に置換する、実施形態62～66のいずれか一項記載の方法。

10

実施形態70

ロイシンリッチリピートキナーゼ2 (LRRK2) 遺伝子またはその調節エレメントを編集する方法であって、前記LRRK2遺伝子またはその調節エレメントを、(i) アデノシン塩基エディターまたはそれをコードする核酸配列と、(ii) ガイドポリヌクレオチドまたはそれをコードする核酸配列と接触させることを含み、前記アデノシン塩基エディターは、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み、前記ガイドポリヌクレオチドが前記アデノシン塩基エディターを誘導してLRRK2遺伝子のSNPにおいてAからGへの核酸塩基改変をもたらす、前記SNPは、配列番号3における番号付けでLRRK2ポリペプチドにおけるG2019S変異体またはそのバリエーションをコードしない、方法。

20

実施形態71

前記アデノシンデアミナーゼドメインが、配列番号2における番号付けでアミノ酸位置82もしくは166またはそれに対応する位置におけるアミノ酸置換を含む、実施形態70記載の方法。

実施形態72

前記ガイドポリヌクレオチドが、パーキンソン病に関連するSNPを含むLRRK2遺伝子またはその調節エレメントに相補的な核酸配列を含む、実施形態61～71のいずれか一項に記載の方法。

実施形態73

前記アデノシン塩基エディターが、パーキンソン病に関連するSNPを含むLRRK2遺伝子またはその調節エレメントに相補的な核酸配列を含む単一ガイドRNA (sgRNA) と複合体をなす、実施形態61～71のいずれか一項に記載の方法。

30

実施形態74

前記sgRNAが核酸配列5' -AAGCGCAAGCCUGGAGGGAA -3' ; または5' -ACUACAGCAUUGCUCAGUAC-3' を含む、実施形態73記載の方法。

実施形態75

前記標的遺伝子がアルファ-L-イズロニダーゼ (IDUA) 遺伝子であり、神経疾患がハーレー症候群である、実施形態60記載の方法。

実施形態76

アルファ-L-イズロニダーゼ (IDUA) 遺伝子またはその調節エレメントを編集する方法であって、IDUA遺伝子またはその調節エレメントを、(i) アデノシン塩基エディターまたはそれをコードする核酸配列と、(ii) ガイドポリヌクレオチドまたはそれをコードする核酸配列と接触させることを含み、前記アデノシン塩基エディターは、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み、

40

前記アデノシンデアミナーゼドメインは、配列番号2における番号付けでアミノ酸位置82もしくは166またはそれらに対応する位置でのアミノ酸置換を含む、

前記ガイドポリヌクレオチドは、前記アデノシン塩基エディターを誘導して前記IDUA遺伝子またはその調節エレメントにおいてAからGへの核酸塩基改変をもたらす、

方法。

実施形態77

50

前記IDUA遺伝子またはその調節エレメントが、ハーラー症候群に関連するSNPを含む、実施形態75または76のいずれか一項に記載の方法。

実施形態78

前記AからGへの核酸塩基改変が、ハーラー症候群に関連するSNPにおけるものである、実施形態75～77のいずれか一項に記載の方法。

実施形態79

ハーラー症候群に関連する前記SNPが、配列番号4における番号付けでIDUAポリペプチドにおけるW402XもしくはW401Xアミノ酸変異体またはIDUA遺伝子によってコードされるそのバリエーションをもたらす、ここでXは終止コドンである、実施形態77または78記載の方法。

実施形態80

前記AからGへの核酸塩基改変が、ハーラー症候群に関連するSNPを野生型核酸塩基に変化させる、実施形態77～79のいずれか一項に記載の方法。

実施形態81

前記AからGへの核酸塩基改変が、ハーラー症候群に関連するSNPを、ハーラー症候群の1つ以上の症状の改善をもたらす非野生型核酸塩基に変化させる、実施形態77～79のいずれか一項に記載の方法。

実施形態82

ハーラー症候群に関連するSNPにおける前記AからGへの改変が、IDUA遺伝子によってコードされる終止コドンをIDUAポリペプチドにおけるトリプトファンに変化させる、実施形態77～81のいずれか一項に記載の方法。

実施形態83

前記ガイドポリヌクレオチドが、ハーラー症候群に関連するSNPを含むIDUA遺伝子またはその調節エレメントに相補的な核酸配列を含む、実施形態77～81のいずれか一項に記載の方法。

実施形態84

前記アデノシン塩基エディターが、ハーラー症候群に関連するSNPを含むIDUA遺伝子またはその調節エレメントに相補的な核酸配列を含む単一ガイドRNA (sgRNA)と複合体をなす、実施形態77～83のいずれか一項に記載の方法。

実施形態85

前記sgRNAが、5'-GACUCUAGGCAGAGGUCUCAA-3'、5'-ACUCUAGGCAGAGGUCUCAA-3'、5'-CUCUAGGCCGAAGUGUCGC-3'、および5'-GCUCUAGGCCGAAGUGUCGC-3'からなる群より選択される核酸配列を含む、実施形態84記載の方法。

実施形態86

前記標的遺伝子がメチルCpG結合タンパク質2 (MECP2) 遺伝子であり、神経疾患がレット症候群である、実施形態60記載の方法。

実施形態87

メチルCpG結合タンパク質2 (MECP2) 遺伝子またはその調節エレメントを編集する方法であって、(i) アデノシン塩基エディターまたはそれをコードする核酸配列と、(ii) ガイドポリヌクレオチドまたはそれをコードする核酸配列を対象に投与することを含み、前記アデノシン塩基エディターは、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み、

前記アデノシンデアミナーゼドメインは、配列番号2における番号付けでアミノ酸位置82もしくは166またはそれらに対応する位置でのアミノ酸置換を含み、

前記ガイドポリヌクレオチドは、前記アデノシン塩基エディターを誘導してMECP2遺伝子またはその調節エレメントにおけるAからGへの核酸塩基改変をもたらす、

方法。

実施形態88

前記MECP2遺伝子またはその調節エレメントが、レット症候群に関連するSNPを含む、実施形態86または87に記載の方法。

10

20

30

40

50

実施形態 8 9

前記AからGへの核酸塩基改変が、レット症候群に関連するSNPにおけるものである、実施形態88記載の方法。

実施形態 9 0

前記レット症候群と関連するSNPが、配列番号5における番号付けでMECP2ポリペプチドにおけるR106WもしくはT158Mアミノ酸変異体またはMECP2遺伝子によってコードされるそのバリエーションをもたらす、実施形態88または89に記載の方法。

実施形態 9 1

レット症候群に関連する前記SNPが、MECP2遺伝子によってコードされるMECP2ポリペプチドにおけるR255XまたはR270Xアミノ酸変異体をもたらす、ここでXは終止コドンである、実施形態88または89記載の方法。

10

実施形態 9 2

前記AからGへの核酸塩基改変が、レット症候群に関連するSNPを野生型核酸塩基に変化させる、実施形態88～81のいずれか一項記載の方法。

実施形態 9 3

前記AからGへの核酸塩基改変が、レット症候群に関連するSNPを、レット症候群の1つ以上の症状の改善をもたらす非野生型核酸塩基に変化させる、実施形態88～91のいずれか一項記載の方法。

実施形態 9 4

レット症候群に関連するSNPにおける前記AからGへの核酸塩基改変が、MECP2ポリペプチドにおいて終止コドントリプトファンに変化させる、実施形態88～91のいずれか一項に記載の方法。

20

実施形態 9 5

前記ガイドポリヌクレオチドが、レット症候群に関連するSNPを含むMECP2遺伝子またはその調節エレメントに相補的な核酸配列を含む、実施形態88～94のいずれか一項に記載の方法。

実施形態 9 6

前記アデノシン塩基エディターが、レット症候群と関連するSNPを含むMECP2遺伝子またはその調節エレメントに相補的な核酸配列を含む単一ガイドRNA (sgRNA) と複合体をなす、実施形態88～94のいずれか一項に記載の方法。

30

実施形態 9 7

前記ガイドポリヌクレオチドが、5' - CUUUUCACUUCUGCCGGGG-3'、5' - AGCUUCC AUGUCCAGCCUUC-3'、5' - ACCAUGAAGUCAAAAUCAUU-3'、および5' - GCUUUCAG CCCCUGUUUCUUG-3' からなる群より選択される核酸配列を含む、実施形態96に記載の方法。

実施形態 9 8

前記標的遺伝子がATP結合カセットサブファミリーメンバー4 (ABCA4) 遺伝子であり、神経疾患がシュタルガルト病である、実施形態60記載の方法。

実施形態 9 9

ATP結合カセットサブファミリーメンバー4 (ABCA4) 遺伝子またはその調節エレメントを編集する方法であって、前記ABCA4遺伝子またはその調節エレメントを、(i) アデノシン塩基エディターまたはそれをコードする核酸配列と、(ii) ガイドポリヌクレオチドまたはそれをコードする核酸配列と接触させることを含み、前記アデノシン塩基エディターは、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み、前記アデノシンデアミナーゼドメインは、配列番号2における番号付けでアミノ酸位置82もしくは166、またはそれらに対応する位置でのアミノ酸置換を含み、

40

前記ガイドポリヌクレオチドは、前記アデノシン塩基エディターを誘導してABCA4遺伝子またはその調節エレメントにおけるAからGへの核酸塩基改変をもたらす、方法。

実施形態 1 0 0

50

前記投与が、シュタルガルト病に関連する少なくとも一つの症状を改善する、実施形態98または99に記載の方法。

実施形態101

前記投与が、アデノシンデアミナーゼにおける前記アミノ酸置換を伴わない塩基エディターを用いた治療と比較して、シュタルガルト病に関連する少なくとも一つの症状のより迅速な改善をもたらす、実施形態100に記載の方法。

実施形態102

ABCA4遺伝子が、シュタルガルト病に関連するSNPを含む、実施形態98~101のいずれか一項記載の方法。

実施形態103

前記AからGへの核酸塩基改変が、シュタルガルト病に関連するSNPにおけるものである、実施形態102記載の方法。

実施形態104

シュタルガルト病に関連する前記SNPが、配列番号6における番号付けでABCA4ポリペプチドにおけるA1038VもしくはG1961Eアミノ酸変異体またはABCA4遺伝子によってコードされるそのバリエーションをもたらす、実施形態102または103に記載の方法。

実施形態105

シュタルガルト病に関連する前記SNPが、配列番号6における番号付けでABCA4ポリペプチドにおけるG1961Eアミノ酸変異体またはそのバリエーションをもたらす、実施形態104記載の方法。

実施形態106

前記AからGへの核酸塩基改変が、シュタルガルト病に関連するSNPを野生型核酸塩基に変化させる、実施形態103~105のいずれか一項記載の方法。

実施形態107

前記AからGへの核酸塩基の改変が、シュタルガルト病に関連するSNPを、シュタルガルト病の1つ以上の症状の改善をもたらす非野生型核酸塩基に変化させる、実施形態103~105のいずれか一項記載の方法。

実施形態108

前記ガイドポリヌクレオチドが、シュタルガルト病に関連するSNPを含むABCA4遺伝子またはその調節エレメントに相補的な核酸配列を含む、実施形態103~107のいずれか一項記載の方法。

実施形態109

前記アデノシン塩基エディターが、シュタルガルト病に関連するSNPを含むABCA4遺伝子またはその調節エレメントに相補的な核酸配列を含む単一ガイドRNA (sgRNA) と複合体をなす、実施形態103~108のいずれか一項に記載の方法。

実施形態110

sgRNAが配列5' -CUCCAGGGCGAACUUCGACACACAGC-3' を含む、実施形態109記載の方法。

実施形態111

前記接触させることは細胞内で起こる、実施形態60~110のいずれか一項記載の方法。

実施形態112

前記接触させることが、細胞中のゲノムにおいて10%未満のインデルを生じ、ここでインデル率は、単一ヌクレオチド改変に隣接する配列と未改変配列との間のミスマッチ頻度によって測定される、実施形態111記載の方法。

実施形態113

前記接触させることが、細胞中のゲノムにおいて5%未満のインデルを生じ、ここでインデル率は、単一ヌクレオチド改変に隣接する配列と未改変配列との間のミスマッチ頻度によって測定される、実施形態111記載の方法。

実施形態114

前記接触させることが、細胞中のゲノムにおいて1%未満のインデルを生じ、ここでイン

10

20

30

40

50

デル率は、単一ヌクレオチド改変に隣接する配列と未改変配列との間のミスマッチ頻度によって測定される、実施形態111記載の方法。

実施形態115

前記細胞がニューロンである、実施形態111～114のいずれか一項記載の方法。

実施形態116

前記接触させることが、細胞の集団内で起こる、実施形態60～110のいずれか一項記載の方法。

実施形態117

前記接触させることが、前記接触の工程の後に前記細胞の集団の少なくとも40%においてAからGへの核酸塩基改変をもたらす、実施形態116に記載の方法。

実施形態118

前記接触させることが、前記接触の工程の後に前記細胞の集団の少なくとも50%においてAからGへの核酸塩基改変をもたらす、実施形態116に記載の方法。

実施形態119

前記接触させることが、前記接触の工程の後に前記細胞の集団の少なくとも70%においてAからGへの核酸塩基改変をもたらす、実施形態116に記載の方法。

実施形態120

前記細胞の少なくとも90%が、前記接触の工程の後に生存性である、実施形態116～119のいずれか一項記載の方法。

実施形態121

前記細胞の集団が、前記接触の工程の後に濃縮されなかった、実施形態116～120のいずれか一項記載の方法。

実施形態122

前記細胞の集団がニューロンである、実施形態116～121のいずれか一項記載の方法。

実施形態123

前記接触させることがインビボまたはエクスピボで起こる、実施形態111～122のいずれか一項記載の方法。

実施形態124

前記ポリヌクレオチドによりプログラム可能なDNA結合ドメインがCas9である、実施形態1～123のいずれか一項に記載の方法。

実施形態125

前記Cas9がSpCas9、SaCas9、またはそれらのバリエーションである、実施形態124に記載の方法。

実施形態126

前記ポリヌクレオチドによりプログラム可能なDNA結合ドメインが、改変されたプロトSpacer隣接モチーフ (PAM) 特異性を有する改変されたSpCas9を含む、実施形態124または125に記載の方法。

実施形態127

前記Cas9が、NGG、NGA、NGCG、NGN、NNGRRT、NNNRRT、NGCG、NGCN、NGTN、およびNGCからなる群より選択されるPAM配列に対する特異性を有し、ここでNはA、G、C、またはTでありRはAまたはGである、実施形態126に記載の方法。

実施形態128

前記ポリヌクレオチドによりプログラム可能なDNA結合ドメインが、ヌクレアーゼ不活性バリエーションである、実施形態124～127のいずれか一項に記載の方法。

実施形態129

前記ポリヌクレオチドによりプログラム可能なDNA結合ドメインがニッカーゼバリエーションである、実施形態124～127のいずれか一項に記載の方法。

実施形態130

前記ニッカーゼバリエーションが、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、実施形態129記載の方法。

10

20

30

40

50

実施形態 1 3 1

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1～130のいずれか一項記載の方法。

実施形態 1 3 2

前記アデノシンデアミナーゼが、V82S改変および/またはT166R改変を含むTadAデアミナーゼを含む、実施形態131記載の方法。

実施形態 1 3 3

前記アデノシンデアミナーゼが、Y147T、Y147R、Q154S、Y123H、Q154R、またはそれらの組み合わせのうちの1つ以上の改変をさらに含む、実施形態1～132のいずれか一項記載の方法。

10

実施形態 1 3 4

前記アデノシンデアミナーゼが、Y147R + Q154R + Y123H；Y147R + Q154R + I76Y；Y147R + Q154R + T166R；Y147T + Q154R；Y147T + Q154S；およびY123H + Y147R + Q154R + I76Yからなる群より選択される改変の組み合わせを含む、実施形態1～133のいずれか一項記載の方法。

実施形態 1 3 5

前記アデノシン塩基エディタードメインがアデノシンデアミナーゼ単量体を含む、実施形態1～134のいずれか一項記載の方法。

実施形態 1 3 6

前記アデノシン塩基エディターがアデノシンデアミナーゼ二量体を含む、実施形態1～135のいずれか一項記載の方法。

20

実施形態 1 3 7

前記TadAデアミナーゼがTadA*8バリエーションである、実施形態131～136のいずれか一項に記載の方法。

実施形態 1 3 8

前記TadA*8バリエーションがTadA*8.1、TadA*8.2、TadA*8.3、TadA*8.4、TadA*8.5、TadA*8.6、TadA*8.7、TadA*8.8、TadA*8.9、TadA*8.10、TadA*8.11、TadA*8.12、およびTadA*8.13からなる群から選択される、実施形態137に記載の方法。

実施形態 1 3 9

前記アデノシン塩基エディターが、ABE8.1、ABE8.2、ABE8.3、ABE8.4、ABE8.5、ABE8.6、ABE8.7、ABE8.8、ABE8.9、ABE8.10、ABE8.11、ABE8.12、およびABE8.13からなる群より選択されるABE8塩基エディターである、実施形態138に記載の方法。

30

実施形態 1 4 0

実施形態111～115のいずれか一項に記載の方法によって産生される細胞。

実施形態 1 4 1

実施形態116～122のいずれか一項に記載の方法によって産生される細胞の集団。

実施形態 1 4 2

(i) アデノシン塩基エディターまたはそれをコードする核酸配列と、(ii) ガイドポリヌクレオチドまたはそれをコードする核酸配列とを含む、塩基エディターシステムであって、前記アデノシン塩基エディターが、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み、前記アデノシンデアミナーゼドメインは、配列番号2における番号付けでアミノ酸位置82もしくは166またはそれらに対応する位置でのアミノ酸置換を含み、

40

前記ガイドポリヌクレオチドは、前記アデノシン塩基エディターを誘導して神経障害に関連する標的遺伝子またはその調節エレメントにおけるAからGへの核酸塩基改変をもたらす、

塩基エディターシステム。

実施形態 1 4 3

前記標的遺伝子がロイシンリッチリピートキナーゼ2 (LRRK2) 遺伝子であり、神経疾患がパーキンソン病である、実施形態142記載の塩基エディターシステム。

50

実施形態 1 4 4

(i) アデノシン塩基エディターまたはそれをコードする核酸配列と、(ii) ガイドポリヌクレオチドまたはそれをコードする核酸配列を含む、塩基エディターシステムであって、前記アデノシン塩基エディターが、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み、前記アデノシンデアミナーゼドメインは、配列番号2における番号付けでアミノ酸位置82もしくは166またはそれらに対応する位置でのアミノ酸置換を含み、

前記ガイドポリヌクレオチドは、前記アデノシン塩基エディターを誘導してLRRK2遺伝子またはその調節エレメントにおけるAからGへの核酸塩基改変をもたらす、

塩基エディターシステム。

10

実施形態 1 4 5

前記AからGへの核酸塩基改変が、LRRK2遺伝子またはその調節エレメント中のパーキンソン病に関連するSNPにおけるものである、実施形態144に記載の塩基エディターシステム。

実施形態 1 4 6

前記パーキンソン病に関連するSNPが、配列番号3における番号付けでLRRK2ポリペプチドにおけるA419V、R1441C、R1441H、もしくはG2019Sアミノ酸変異体またはLRRK2遺伝子によってコードされるそのバリエーションをもたらす、実施形態144または145に記載の塩基エディターシステム。

実施形態 1 4 7

前記AからGへの核酸塩基改変が、前記パーキンソン病に関連するSNPを野生型核酸塩基に変化させる、実施形態145または146のいずれか一項に記載の塩基エディターシステム。

20

実施形態 1 4 8

前記AからGへの核酸塩基改変が、前記パーキンソン病に関連するSNPを、パーキンソン病の症状の改善をもたらす非野生型核酸塩基に変化させる、実施形態145または146のいずれか一項に記載の塩基エディターシステム。

実施形態 1 4 9

前記AからGへの核酸塩基の改変が、LRRK2遺伝子によってコードされるLRRK2ポリペプチド中のシステインまたはヒスチジンをアルギニンに変化させる、実施形態145～148のいずれか一項記載の塩基エディターシステム。

30

実施形態 1 5 0

前記AからGへの改変が、LRRK2遺伝子によってコードされるLRRK2ポリペプチドにおいてセリンをグリシンに変化させる、実施形態145～148のいずれか一項に記載の塩基エディターシステム。

実施形態 1 5 1

前記AからGへの改変が、配列番号3における番号付けでLRRK2ポリペプチドまたはLRRK2遺伝子によってコードされるそのバリエーションの位置144でシステイン (C) もしくはヒスチジン (H) をアルギニン (R) に置換するか、または位置2019でセリンをグリシン (G) に置換する、実施形態145～148のいずれか一項に記載の塩基エディターシステム。

40

実施形態 1 5 2

前記アデノシンデアミナーゼドメインが、配列番号2における番号付けでアミノ酸位置82もしくは166またはそれに対応する位置におけるアミノ酸置換を含む、実施形態151に記載の塩基エディターシステム。

実施形態 1 5 3

前記ガイドポリヌクレオチドが、パーキンソン病に関連するSNPを含むLRRK2遺伝子またはその調節エレメントに相補的な核酸配列を含む、実施形態143～152のいずれか一項に記載の塩基エディターシステム。

実施形態 1 5 4

前記アデノシン塩基エディターが、パーキンソン病と関連する前記SNPを含むLRRK2遺

50

伝子またはその調節エレメントに相補的な核酸配列を含む単一ガイドRNA (sgRNA) と複合体をなす、実施形態143～152のいずれか一項に記載の塩基エディターシステム。

実施形態 1 5 5

前記sgRNAが核酸配列5' -AAGCGCAAGCCUGGAGGGAA -3' ; または5' -ACUACAG CAUUGCUCAGUAC-3' を含む、実施形態154に記載の塩基エディターシステム。

実施形態 1 5 6

前記標的遺伝子がアルファ-L-イズロニダーゼ (IDUA) 遺伝子であり、神経疾患がハーラー症候群である、実施形態142に記載の塩基エディターシステム。

実施形態 1 5 7

(i) アデノシン塩基エディターまたはそれをコードする核酸配列と、 (ii) ガイドポリヌクレオチドまたはそれをコードする核酸配列とを含む、塩基エディターシステムであって、前記アデノシン塩基エディターが、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み、前記アデノシンデアミナーゼドメインは、配列番号2における番号付けでアミノ酸位置82もしくは166またはそれに対応する位置でのアミノ酸置換を含み、

前記ガイドポリヌクレオチドが、前記アデノシン塩基エディターを誘導してアルファ-L-イズロニダーゼ (IDUA) 遺伝子またはその調節エレメントにおけるAからGへの核酸塩基改変をもたらす、

塩基エディターシステム。

実施形態 1 5 8

前記IDUA遺伝子またはその調節エレメントが、ハーラー症候群に関連するSNPを含む、実施形態156または157に記載の塩基エディターシステム。

実施形態 1 5 9

前記AからGへの核酸塩基改変がハーラー症候群に関連するSNPにおけるものである、実施形態156～158のいずれか一項記載の塩基エディターシステム。

実施形態 1 6 0

前記ハーラー症候群に関連するSNPが、配列番号4における番号付けでIDUAポリペプチドにおけるW402XもしくはW401Xアミノ酸変異体またはIDUA遺伝子によってコードされるそのバリエーションをもたらす、ここでXは終止コドンである、実施形態158または159に記載の塩基エディターシステム。

実施形態 1 6 1

前記AからGへの核酸塩基改変が、ハーラー症候群に関連するSNPを野生型核酸塩基に変化させる、実施形態158～160のいずれか一項記載の塩基エディターシステム。

実施形態 1 6 2

前記AからGへの核酸塩基の改変が、ハーラー症候群に関連するSNPを、ハーラー症候群の1つ以上の症状の改善をもたらす非野生型核酸塩基に変化させる、実施形態158～160のいずれか一項に記載の塩基エディターシステム。

実施形態 1 6 3

ハーラー症候群に関連するSNPにおけるAからGへの改変が、IDUA遺伝子によってコードされるIDUAポリペプチドにおいて終止コドンをトリプトファンに変化させる、実施形態158～162のいずれか一項記載の塩基エディターシステム。

実施形態 1 6 4

前記ガイドポリヌクレオチドが、ハーラー症候群と関連するSNPを含むIDUA遺伝子またはその調節エレメントに相補的な核酸配列を含む、実施形態158～162のいずれか一項記載の塩基エディターシステム。

実施形態 1 6 5

前記アデノシン塩基エディターが、ハーラー症候群と関連するSNPを含むIDUA遺伝子またはその調節エレメントに相補的な核酸配列を含む単一ガイドRNA (sgRNA) と複合体をなす、実施形態158～164のいずれか一項に記載の塩基エディターシステム。

実施形態 1 6 6

前記sgRNAが、5'-GACUCUAGGCAGAGGUCUCAA-3'、5'-ACUCUAGGCAGAGGUCUCAA-3'、5'-CUCUAGGCCGAAGUGUCGC-3'、および5'-GCUCUAGGCCGAAGUGUCGC-3'からなる群より選択される核酸配列を含む、実施形態165に記載の塩基エディターシステム。

実施形態167

前記標的遺伝子がメチルCpG結合タンパク質2 (MECP2) 遺伝子であり、神経疾患がレット症候群である、実施形態142に記載の塩基エディターシステム。

実施形態168

(i) アデノシン塩基エディターまたはそれをコードする核酸配列と、(ii) ガイドポリヌクレオチドまたはそれをコードする核酸配列とを含む、塩基エディターシステムであって、前記アデノシン塩基エディターは、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み、前記アデノシンデアミナーゼドメインは、配列番号2における番号付けでアミノ酸位置82もしくは166またはそれらに対応する位置でのアミノ酸置換を含み、

10

前記ガイドポリヌクレオチドは、前記アデノシン塩基エディターを誘導してメチルCpG結合タンパク質2 (MECP2) 遺伝子またはその調節エレメントにおけるAからGへの核酸塩基改変をもたらす、

塩基エディターシステム。

実施形態169

前記MECP2遺伝子またはその調節エレメントが、レット症候群に関連するSNPを含む、実施形態167または168に記載の塩基エディターシステム。

20

実施形態170

前記AからGへの核酸塩基改変が、レット症候群に関連するSNPにおけるものである、実施形態169に記載の塩基エディターシステム。

実施形態171

レット症候群に関連するSNPが、配列番号5における番号付けでMECP2ポリペプチドにおけるR106WまたはT158Mアミノ酸変異体またはMECP2遺伝子によってコードされるそのバリエーションをもたらす、実施形態169または170に記載の塩基エディターシステム。

実施形態172

レット症候群に関連するSNPが、MECP2遺伝子によってコードされるMECP2ポリペプチドにおけるR255XまたはR270Xアミノ酸変異体をもたらす、ここでXは終止コドンである、実施形態169または170に記載の塩基エディターシステム。

30

実施形態173

前記AからGへの核酸塩基改変が、レット症候群に関連するSNPを野生型核酸塩基に変化させる、実施形態169～172のいずれか一項に記載の塩基エディターシステム。

実施形態174

前記AからGへの核酸塩基改変が、レット症候群に関連するSNPを、レット症候群の1つ以上の症状の改善をもたらす非野生型核酸塩基に変化させる、実施形態169～172のいずれか一項に記載の塩基エディターシステム。

実施形態175

レット症候群に関連するSNPにおけるAからGへの核酸塩基改変が、MECP2ポリペプチドにおいて終止コドントリプトファンに変化させる、実施形態169～172のいずれか一項に記載の塩基エディターシステム。

40

実施形態176

前記ガイドポリヌクレオチドが、レット症候群に関連するSNPを含むMECP2遺伝子またはその調節エレメントに相補的な核酸配列を含む、実施形態169～175のいずれか一項に記載の塩基エディターシステム。

実施形態177

前記アデノシン塩基エディターが、レット症候群に関連するSNPを含むMECP2遺伝子またはその調節エレメントに相補的な核酸配列を含む単一ガイドRNA (sgRNA) と複合体を

50

なす、実施形態169～175のいずれか一項に記載の塩基エディターシステム。

実施形態178

前記ガイドポリヌクレオチドが、5'-CUUUUCACUUCCUGCCGGGG-3'、5'-AGCUUCCAUGUCCAGCCUUC-3'、5'-ACCAUGAAGUCAAAAUCAUU-3'、および5'-GCUUUCAGCCCCGUUUCUUG-3'からなる群から選択される核酸配列を含む、実施形態177に記載の塩基エディターシステム。

実施形態179

前記標的遺伝子がATP結合カセットサブファミリーメンバー4 (ABCA4) 遺伝子であり、神経疾患がシュタルガルト病である、実施形態142に記載の塩基エディターシステム。

実施形態180

(i) アデノシン塩基エディターまたはそれをコードする核酸配列と、(ii) ガイドポリヌクレオチドまたはそれをコードする核酸配列とを接触させることを含む、塩基エディターシステムであって、前記アデノシン塩基エディターが、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み、前記アデノシンデアミナーゼドメインは、配列番号2における番号付けでアミノ酸位置82もしくは166またはそれに対応する位置でのアミノ酸置換を含み、

前記ガイドポリヌクレオチドは、前記アデノシン塩基エディターを誘導してATP結合カセットサブファミリーメンバー4 (ABCA4) 遺伝子またはその調節エレメントにおけるAからGへの核酸塩基改変をもたらす、

塩基エディターシステム。

実施形態181

投与が、シュタルガルト病に関連する少なくとも一つの症状を改善する、実施形態179または180に記載の塩基エディターシステム。

実施形態182

投与が、アデノシンデアミナーゼにおける前記アミノ酸置換を伴わない塩基エディターを用いた治療と比較して、シュタルガルト病に関連する少なくとも一つの症状のより迅速な改善をもたらす、実施形態181に記載の塩基エディターシステム。

実施形態183

ABCA4遺伝子が、シュタルガルト病に関連するSNPを含む、実施形態179～182のいずれか一項記載の塩基エディターシステム。

実施形態184

前記AからGへの核酸塩基改変が、シュタルガルト病に関連するSNPにおけるものである、実施形態183に記載の塩基エディターシステム。

実施形態185

前記シュタルガルト病に関連するSNPが、配列番号6における番号付けでABCA4ポリペプチドにおけるA1038VもしくはG1961Eアミノ酸変異体またはABCA4遺伝子によってコードされるそのバリエーションをもたらす、実施形態183または184に記載の塩基エディターシステム。

実施形態186

前記シュタルガルト病に関連するSNPが、配列番号6における番号付けでABCA4ポリペプチドにおけるG1961Eアミノ酸変異体またはそのバリエーションをもたらす、実施形態185に記載の塩基エディターシステム。

実施形態187

前記AからGへの核酸塩基改変が、シュタルガルト病に関連するSNPを野生型核酸塩基に変化させる、実施形態184～186のいずれか一項記載の塩基エディターシステム。

実施形態188

前記AからGへの核酸塩基改変が、シュタルガルト病に関連するSNPを、シュタルガルト病症状の改善をもたらす非野生型核酸塩基に変化させる、実施形態184～186のいずれか一項に記載の塩基エディターシステム。

実施形態189

前記ガイドポリヌクレオチドが、シュタルガルト病に関連するSNPを含むABCA4遺伝子またはその調節エレメントに相補的な核酸配列を含む、実施形態184～188のいずれか一項に記載の塩基エディターシステム。

実施形態190

前記アデノシン塩基エディターが、前記シュタルガルト病に関連するSNPを含むABCA4遺伝子またはその調節エレメントに相補的な核酸配列を含む単一ガイドRNA (sgRNA)と複合体をなす、実施形態184～189のいずれか一項に記載の塩基エディターシステム。

実施形態191

前記sgRNAが配列5'-CUCCAGGGCGAACUUCGACACACAGC-3'を含む、実施形態190に記載の塩基エディターシステム。

10

実施形態192

前記ポリヌクレオチドによりプログラム可能なDNA結合ドメインがCas9である、実施形態142～191のいずれか一項に記載の塩基エディターシステム。

実施形態193

前記Cas9がSpCas9、SaCas9、またはそのバリエーションである、実施形態192に記載の塩基エディターシステム。

実施形態194

前記ポリヌクレオチドによりプログラム可能なDNA結合ドメインが、改変されたプロトSpacer隣接モチーフ (PAM) 特異性を有する改変されたSpCas9を含む、実施形態192または193に記載の塩基エディターシステム。

20

実施形態195

前記Cas9が、NGG、NGA、NGCG、NGN、NNGRRT、NNNRRT、NGCG、NGCN、NGTN、およびNGCからなる群より選択されるPAM配列に対する特異性を有し、ここでNはA、G、CまたはTであり、RはAまたはGである、実施形態194に記載の塩基エディターシステム。

実施形態196

前記ポリヌクレオチドによりプログラム可能なDNA結合ドメインがヌクレアーゼ不活性バリエーションである、実施形態192～195のいずれか一項に記載の塩基エディターシステム。

実施形態197

前記ポリヌクレオチドによりプログラム可能なDNA結合ドメインがニックアーゼバリエーションである、実施形態192～195のいずれか一項に記載の塩基エディターシステム。

30

実施形態198

前記ニックアーゼバリエーションが、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、実施形態197に記載の塩基エディターシステム。

実施形態199

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142～198のいずれか一項に記載の塩基エディターシステム。

実施形態200

前記アデノシンデアミナーゼが、V82S改変および/またはT166R改変を含むTadAデアミナーゼを含む、実施形態199に記載の塩基エディターシステム。

40

実施形態201

前記アデノシンデアミナーゼが、Y147T、Y147R、Q154S、Y123H、Q154R、またはそれらの組み合わせのうちの1つ以上の改変をさらに含む、実施形態142～200のいずれか一項に記載の塩基エディターシステム。

実施形態202

前記アデノシンデアミナーゼが、Y147R + Q154R + Y123H; Y147R + Q154R + I76Y; Y147R + Q154R + T166R; Y147T + Q154R; Y147T + Q154S; およびY123H + Y147R + Q154R + I76Yからなる群から選択される改変の組み合わせを含む、実施形態142～201のいずれか一項に記載の塩基エディターシステム。

50

実施形態 203

前記アデノシン塩基エディタードメインがアデノシンデアミナーゼ単量体を含む、実施形態142～202のいずれか一項記載の塩基編集システム。

実施形態 204

前記アデノシン塩基エディタがアデノシンデアミナーゼ二量体を含む、実施形態142～203のいずれか一項記載の塩基エディターシステム。

実施形態 205

前記TadAデアミナーゼがTadA*8バリエーションである、実施形態199～204のいずれか一項に記載の塩基エディターシステム。

実施形態 206

前記TadA*8バリエーションは、TadA*8.1、TadA*8.2、TadA*8.3、TadA*8.4、TadA*8.5、TadA*8.6、TadA*8.7、TadA*8.8、TadA*8.9、TadA*8.10、TadA*8.11、TadA*8.12、およびTadA*8.13からなる群から選択される、実施形態205に記載の塩基エディターシステム。

実施形態 207

前記アデノシン塩基エディターが、ABE8.1、ABE8.2、ABE8.3、ABE8.4、ABE8.5、ABE8.6、ABE8.7、ABE8.8、ABE8.9、ABE8.10、ABE8.11、ABE8.12、およびABE8.13からなる群から選択されるABE8塩基エディターである、実施形態206に記載の塩基エディターシステム。

実施形態 208

実施形態142～207のいずれか一項記載のアデノシン塩基エディターをコードする核酸配列を含むベクター。

実施形態 209

実施形態142～207のいずれか一項に記載のアデノシン塩基エディターおよびガイドポリヌクレオチドをコードする核酸配列を含むベクター。

実施形態 210

前記ベクターがウイルスベクター、レンチウイルスベクター、またはAAVベクターである、実施形態208または209に記載のベクター。

実施形態 211

実施形態142～207のいずれか一項記載の塩基エディターシステムまたは実施形態208～210のいずれか一項記載のベクターを含む、細胞。

実施形態 212

中枢神経系細胞である、実施形態211記載の細胞。

実施形態 213

ニューロンである、実施形態211に記載の細胞。

実施形態 214

光受容体である、実施形態211記載の細胞。

実施形態 215

細胞がインビトロ、インビボまたはエクスピボである、実施形態211～214のいずれか一項記載の細胞。

実施形態 216

実施形態142～207のいずれか一項記載の塩基エディター、実施形態208～210のいずれか一項記載のベクター、又は実施形態211～215のいずれか一項記載の細胞と、薬学的に許容される担体とを含む、医薬組成物。

実施形態 217

脂質をさらに含む、実施形態216記載の医薬組成物。

実施形態 218

ウイルスをさらに含む、実施形態216記載の医薬組成物。

実施形態 219

実施形態142～207のいずれか一項記載の塩基エディター、または実施形態208～210

10

20

30

40

50

のいずれか一項記載のベクターを含む、キット。

実施形態 2 2 0

前記ガイドポリヌクレオチドの少なくとも1つのヌクレオチドが非天然の修飾を含む、実施形態1～139のいずれか一項記載の方法。

実施形態 2 2 1

前記核酸配列の少なくとも一つのヌクレオチドが非天然の修飾を含む、実施形態14、31、45、58、74、85、97または110のいずれか一項記載の方法。

実施形態 2 2 2

前記核酸配列の少なくとも一つのヌクレオチドが非天然の修飾を含む、実施形態155、166、178または191のいずれか一項に記載の塩基エディターシステム。

10

実施形態 2 2 3

前記非天然の修飾が化学修飾である、実施形態222に記載の塩基エディターシステム。

実施形態 2 2 4

前記化学修飾が2'-O-メチル化である、実施形態223に記載の塩基エディターシステム。

実施形態 2 2 5

前記核酸配列がホスホロチオエートを含む、実施形態223に記載の塩基エディターシステム。

〔他の実施形態〕

以上の説明から、本明細書に記載された発明を様々な用途および条件に適合させるために、それに変形および修正を加えることができることは明らかである。そのような実施形態も、下記の特許請求の範囲の範囲内である。

20

【1 2 0 7】

本明細書における可変要素の定義における要素のリストの記載は、リストされた要素のいずれかの単一の要素または組み合わせ（またはサブ組み合わせ）としてのその可変要素の定義を含む。本明細書における一実施形態の記載は、その実施形態を単一実施形態として、または任意の他の実施形態もしくはその一部との組み合わせとして含む。

【1 2 0 8】

本明細書に記載されるすべての刊行物、特許、および特許出願は、それぞれの個々の刊行物、特許、または特許出願が参照により組み込まれるように具体的かつ個別に指示された場合と同じ範囲で、参照により本明細書に組み込まれる。別段の表示がない場合、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、それらの全体が参照により本明細書に組み込まれる。

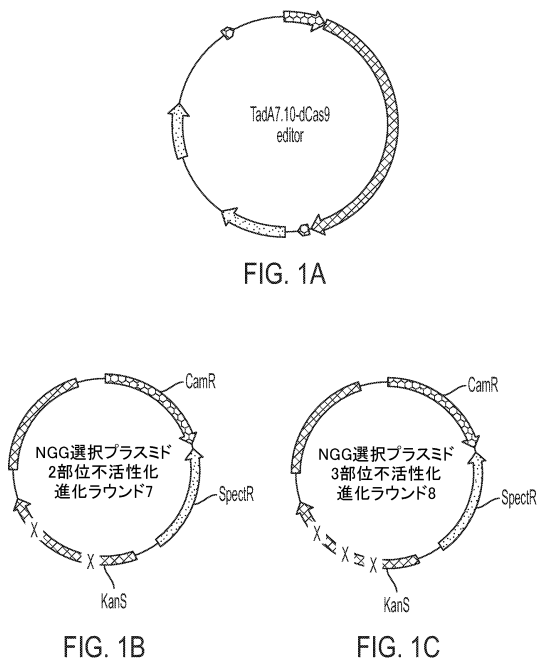
30

40

50

【図面】

【図 1】



【図 2】

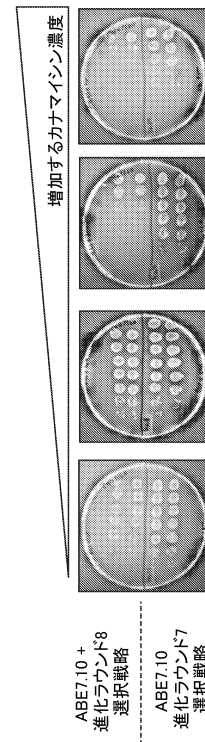
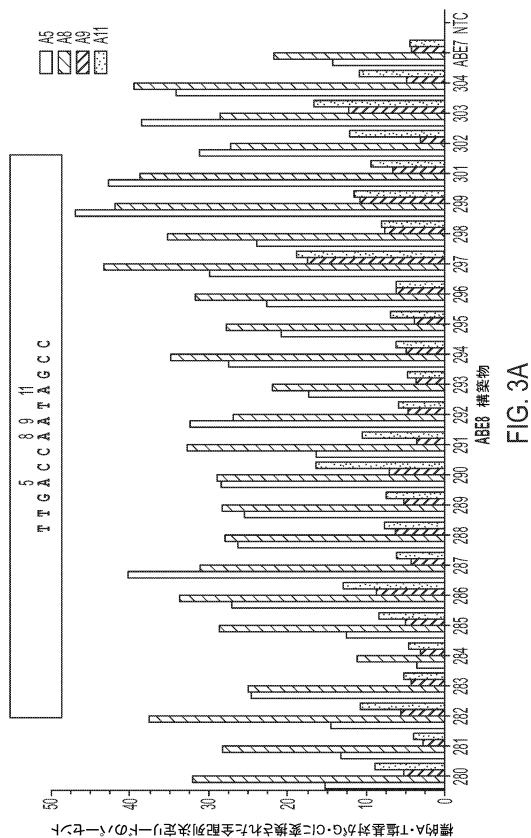
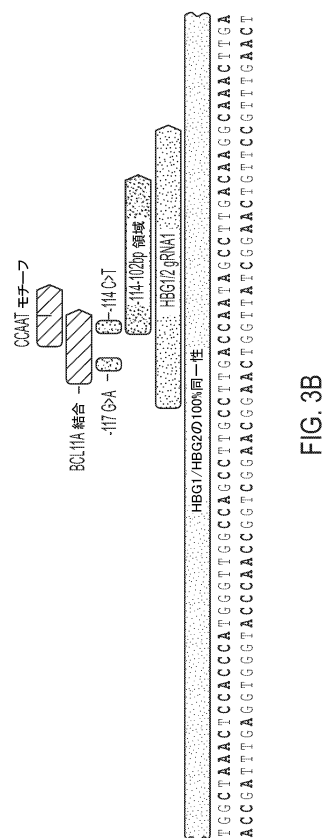


FIG. 2

【図 3 - 1】



【図 3 - 2】



10

20

30

40

50

【図 5 - 1】

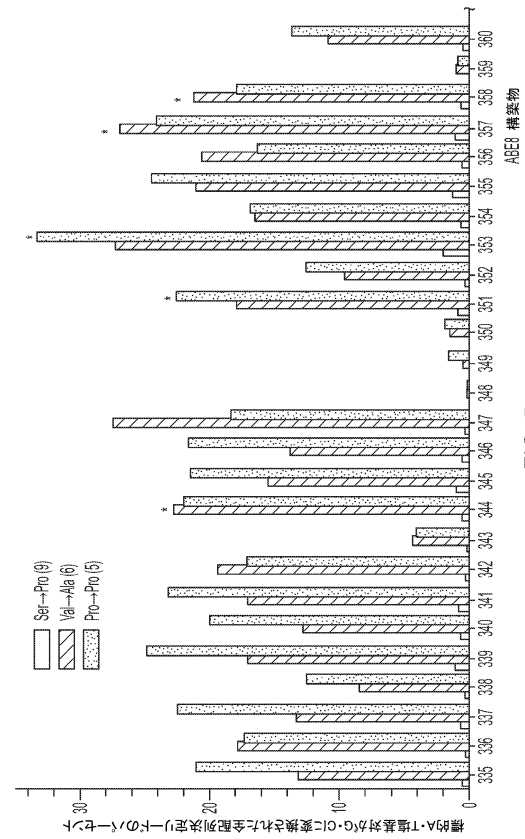
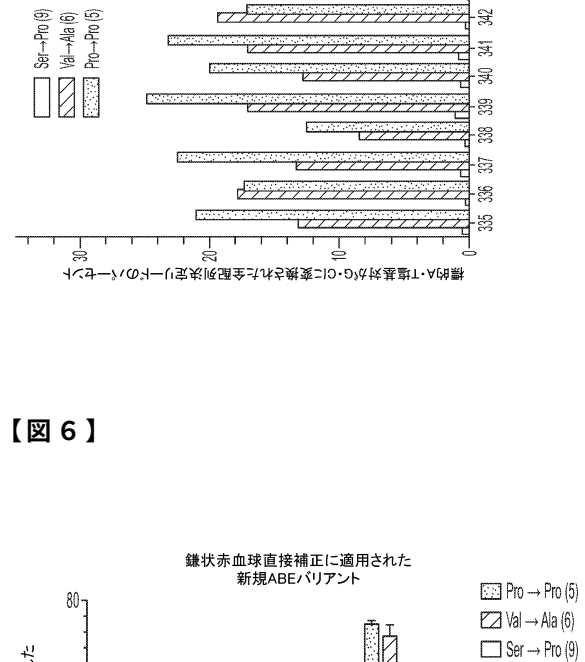


FIG. 5



【図 6】

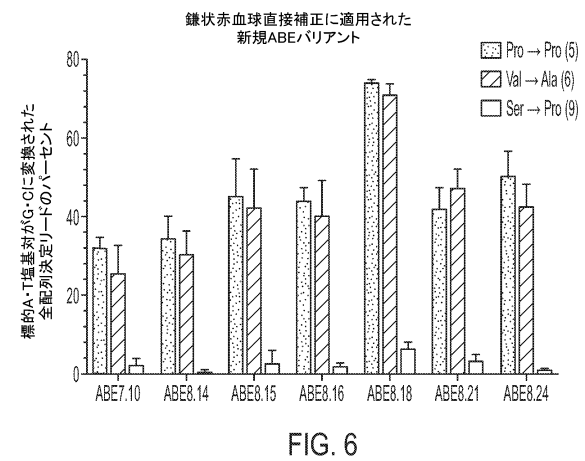


FIG. 6

【図 7 A】

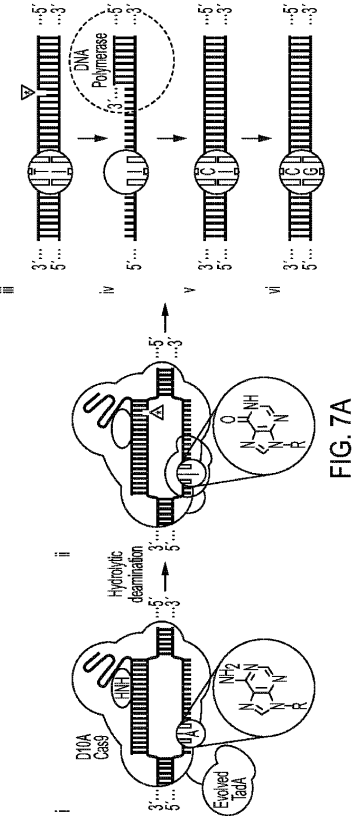


FIG. 7A

【図 7 B】

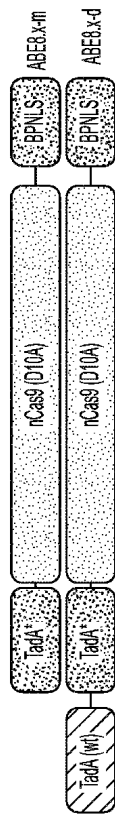


FIG. 7B

【図 7 C】

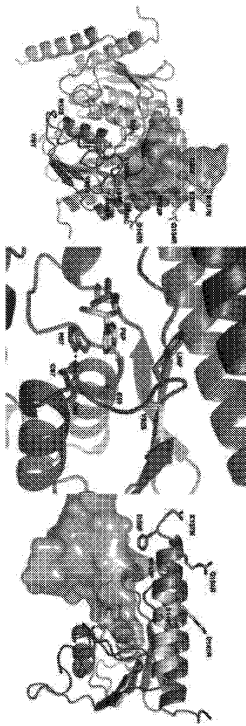


FIG. 7C

【図 7 D】

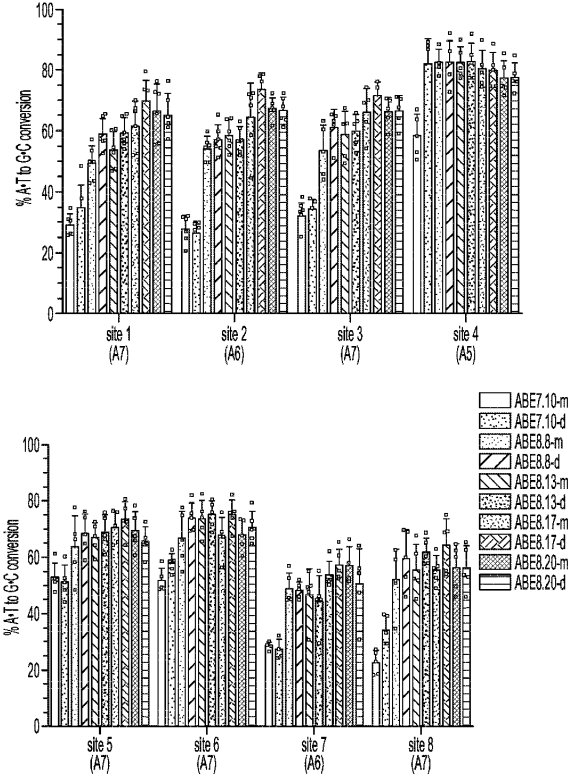


FIG. 7D

【図 8 A】

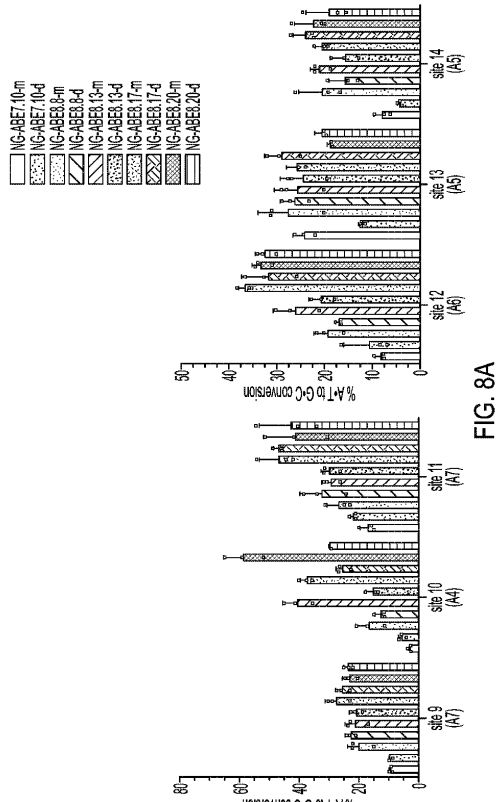


FIG. 8A

【図 8 B】

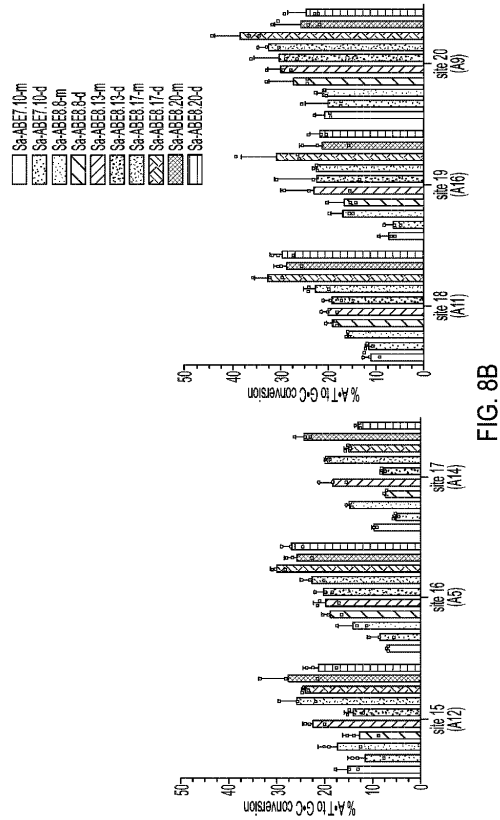


FIG. 8B

【図 8 C】

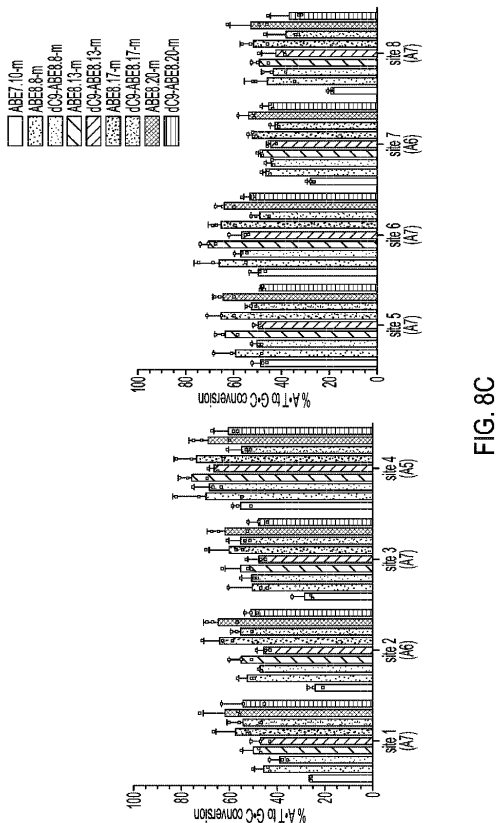


FIG. 8C

【図 9 - 1】

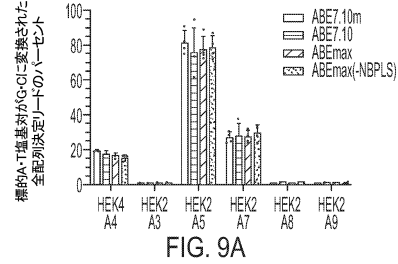


FIG. 9A

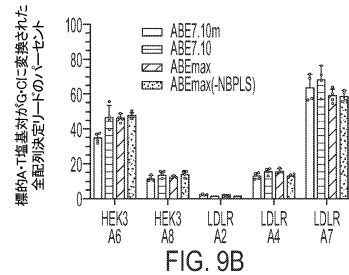


FIG. 9B

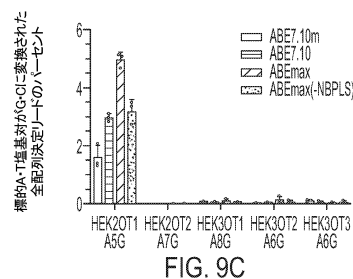


FIG. 9C

10

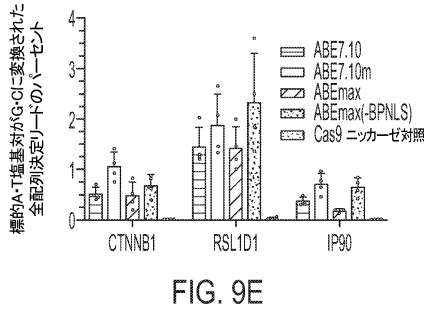
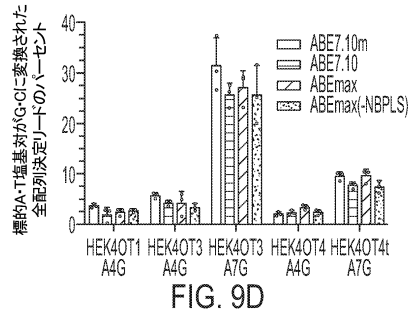
20

30

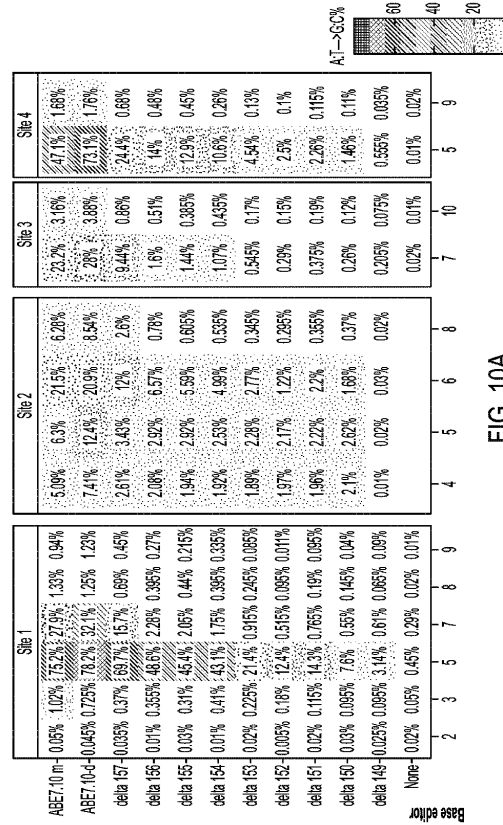
40

50

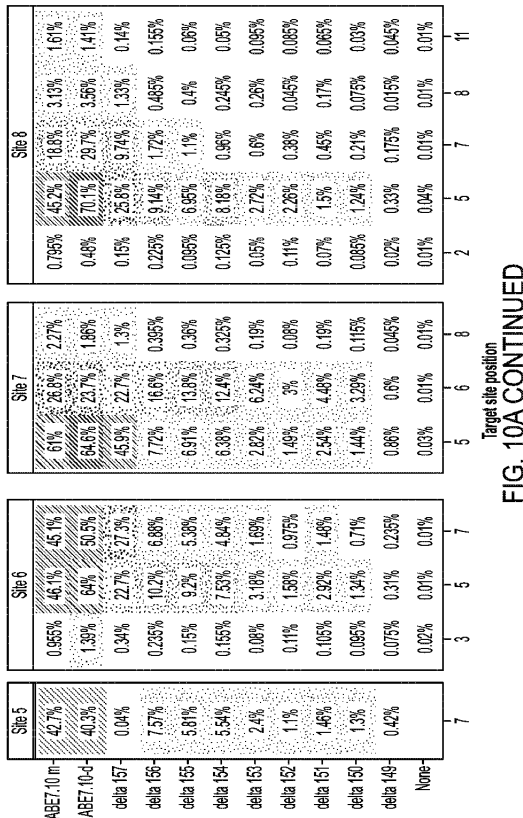
【図 9 - 2】



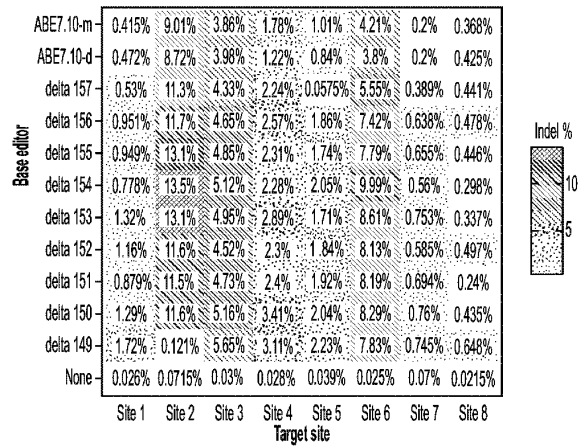
【図 10 A - 1】



【図 10 A - 2】



【図 10 B】



10

20

30

40

50

【図 1 1】

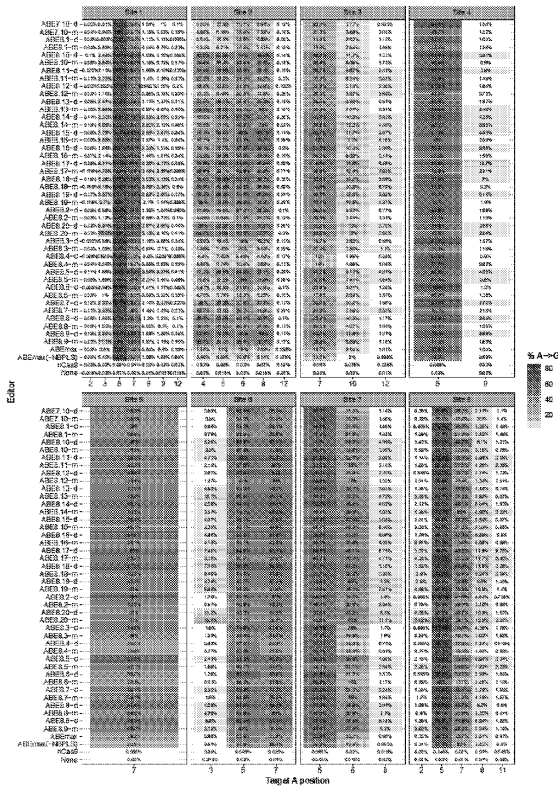


FIG. 11

【図 12-1】

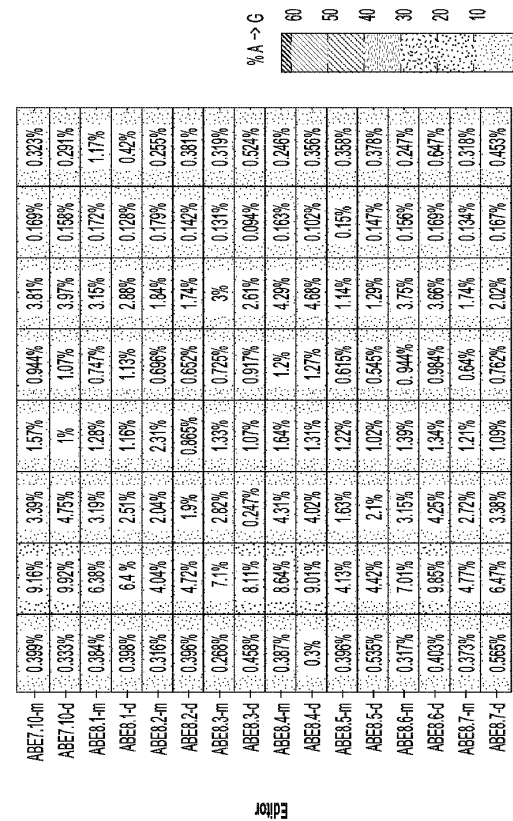


FIG. 12

【図 12-2】

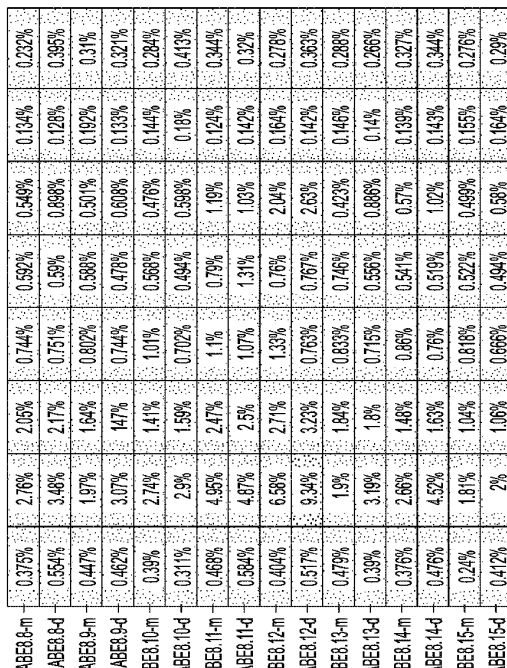


FIG. 12 CONTINUED

【圖 12-3】

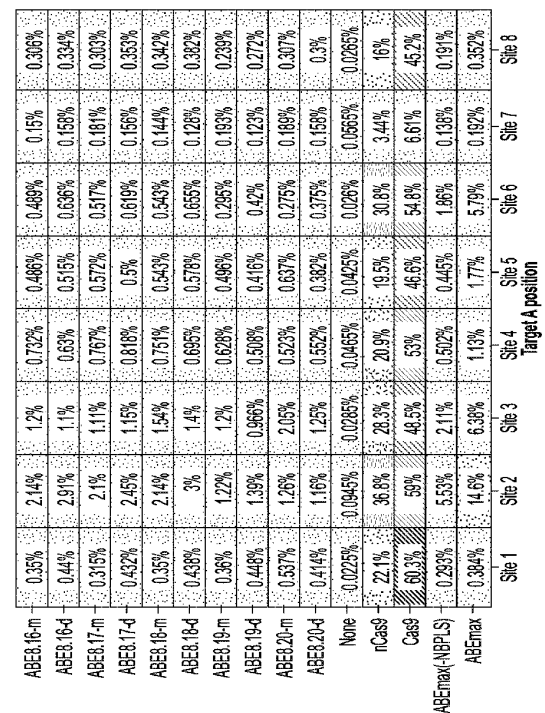
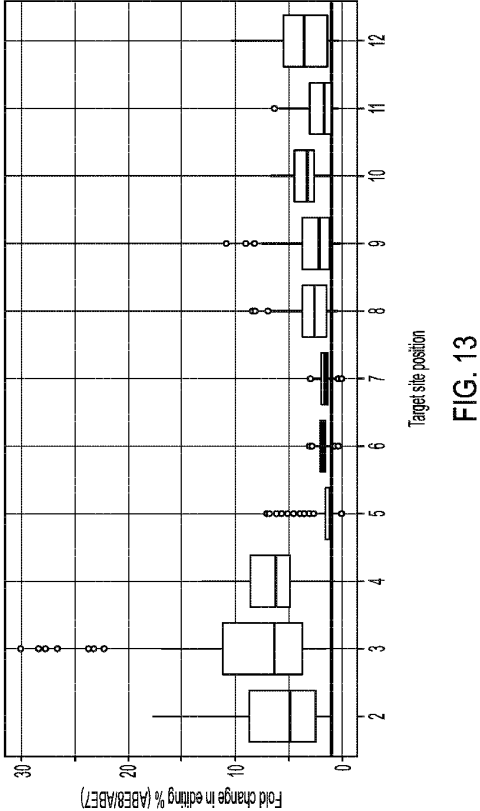
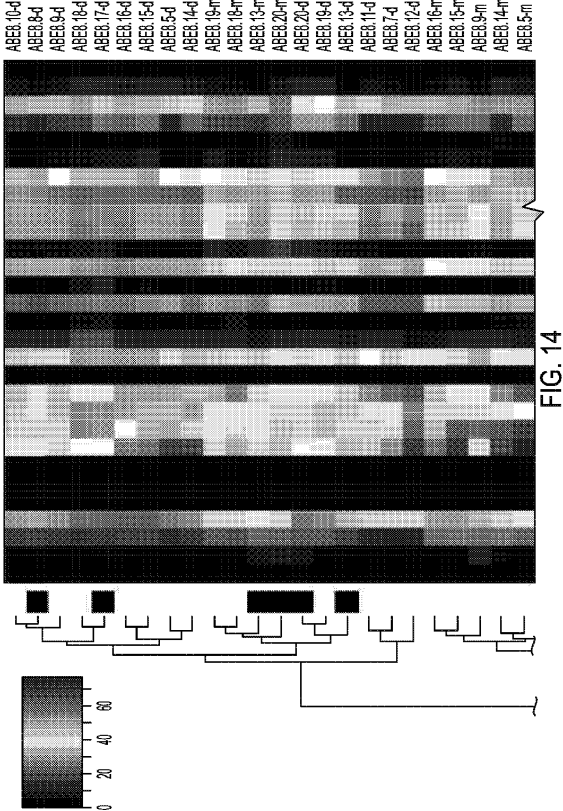


FIG. 12 CONTINUED

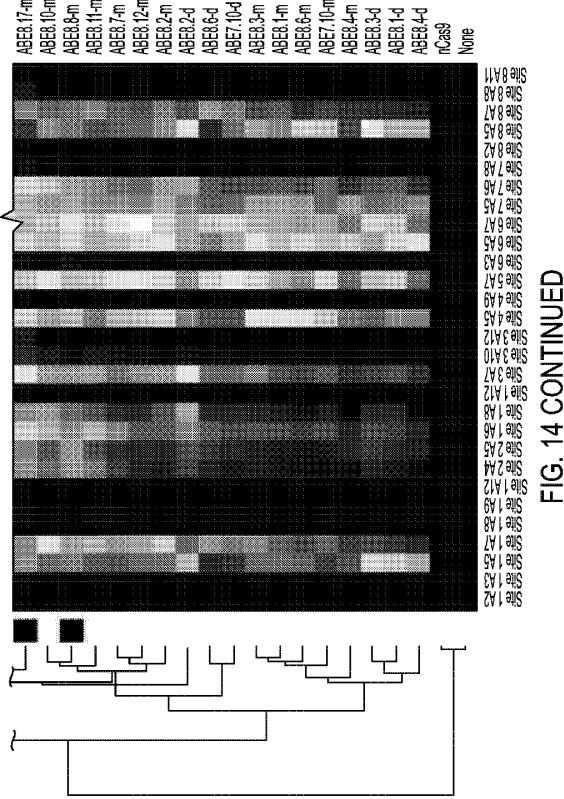
【図 1 3】



【図 1 4 - 1】



【図 1 4 - 2】



【図 1 5 - 1】

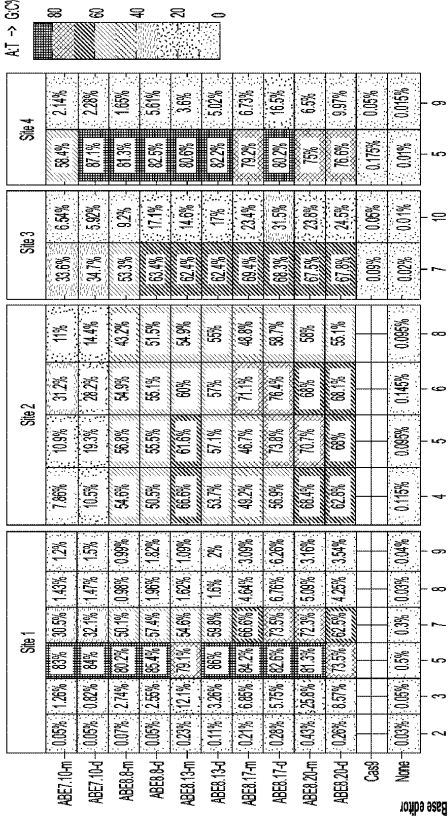


FIG. 15

10

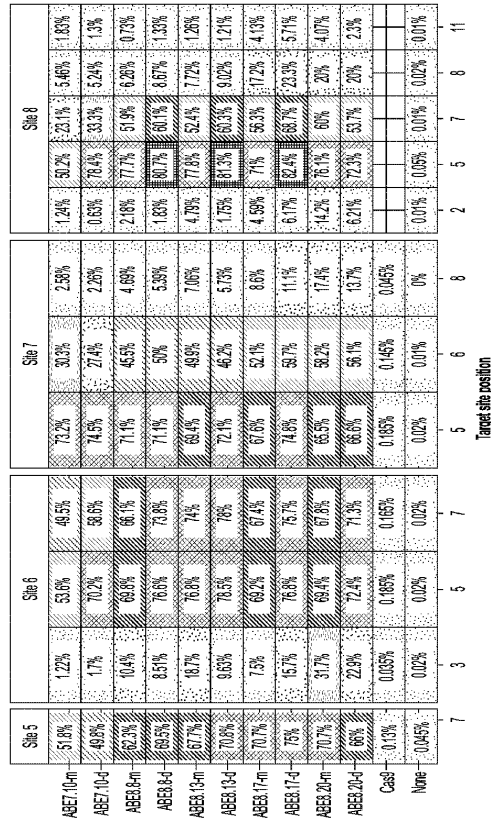
20

30

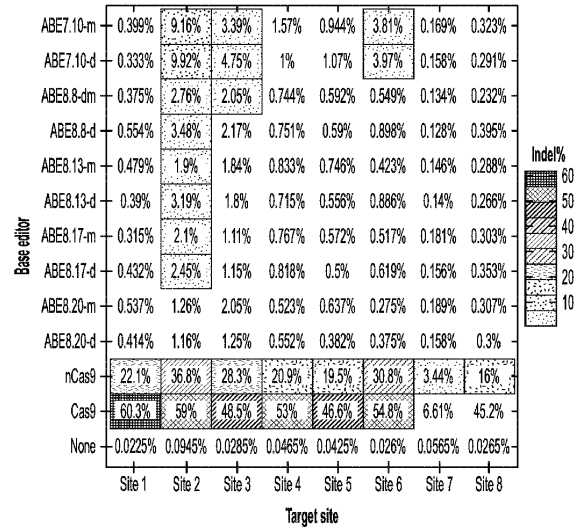
40

50

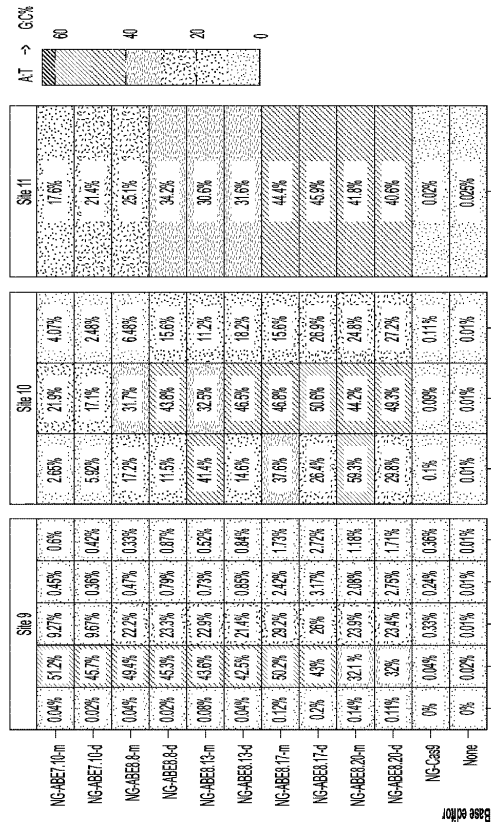
【図 15 - 2】



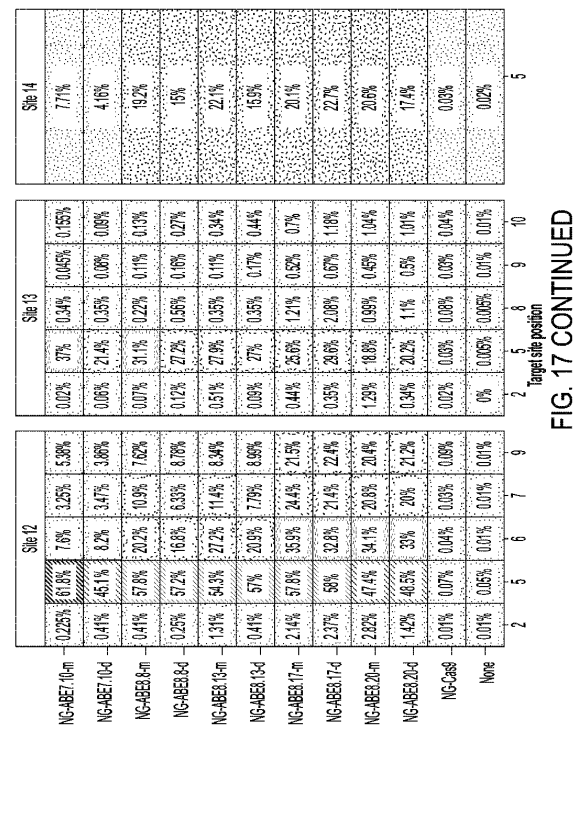
【図 16】



【図 17 - 1】



【図 17 - 2】



10

20

30

40

50

【図 18】

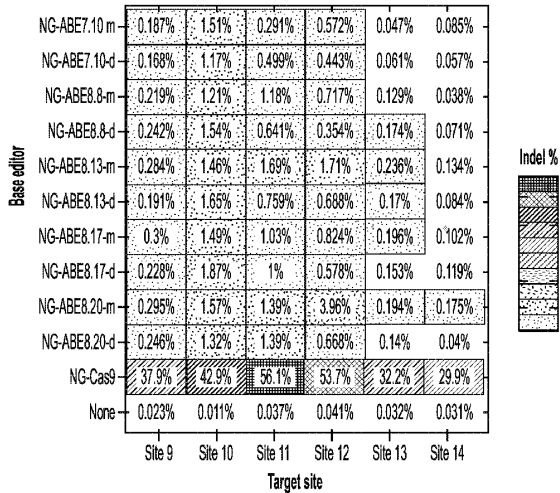


FIG. 18

【図 19 - 1】

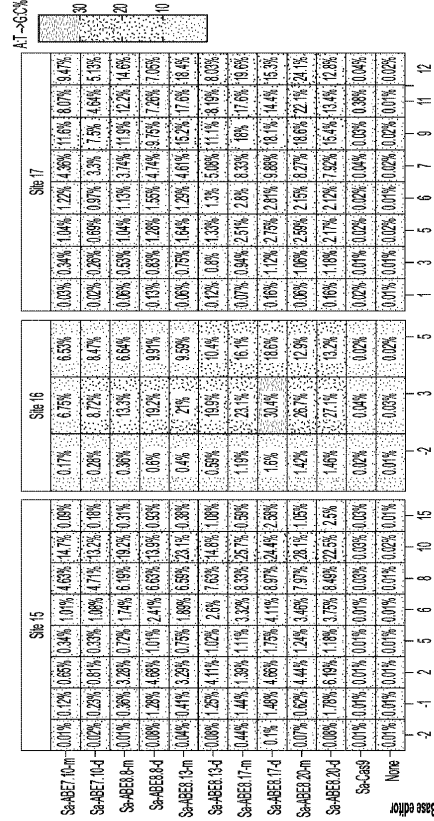


FIG. 19

【図 19 - 2】

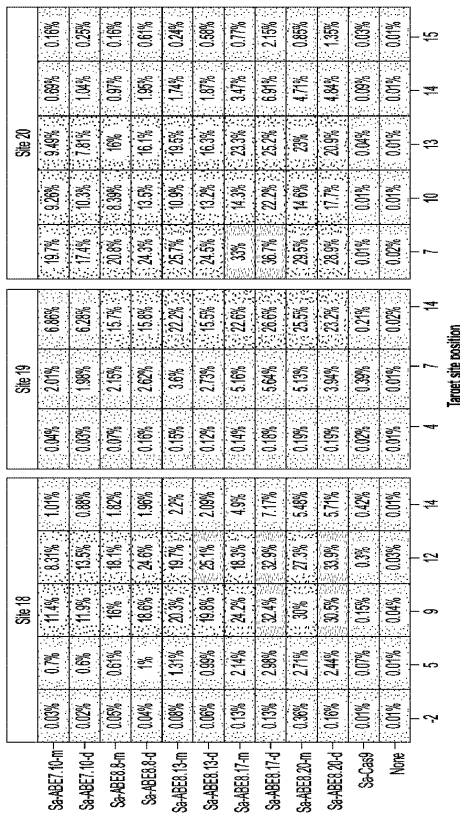


FIG. 19 CONTINUED

【図 20】

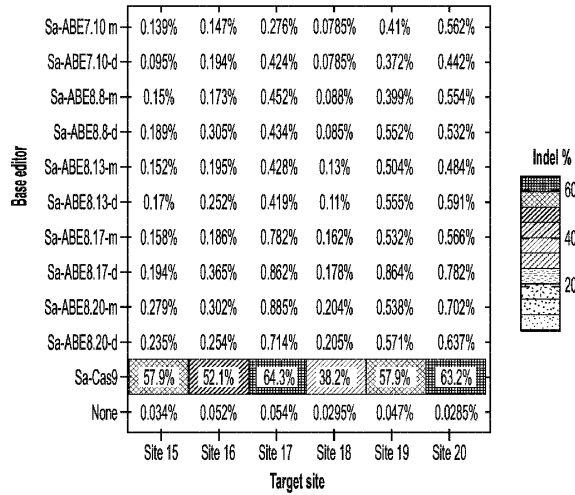


FIG. 20

【図 2 1 - 1】

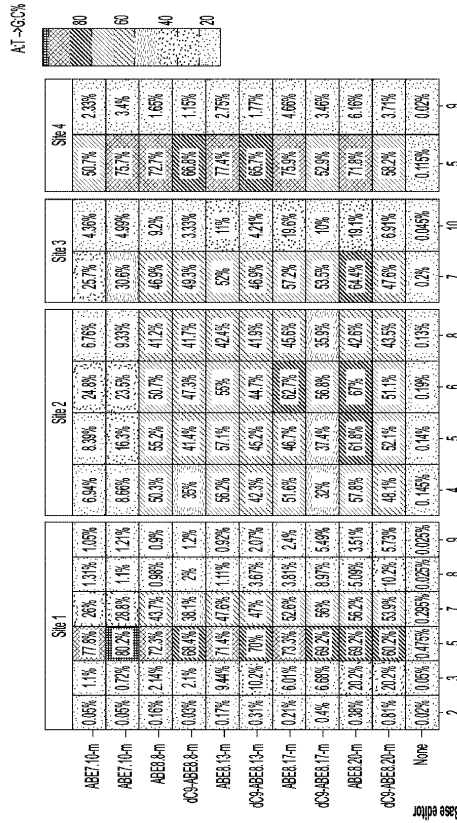


FIG. 21

【図 2 1 - 2】

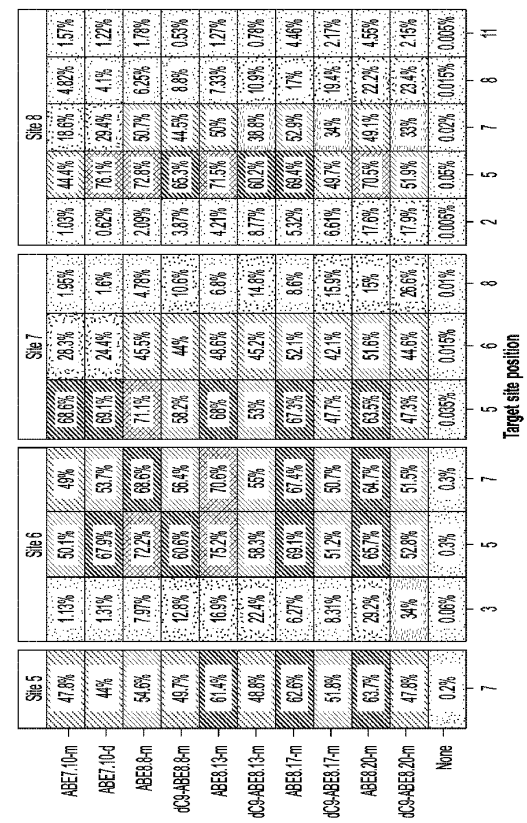


FIG. 21 CONTINUED

【図 2 2 - 1】

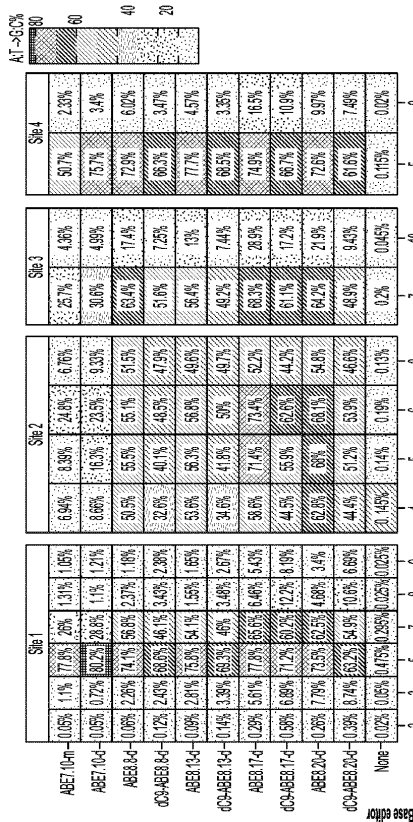


FIG. 22

【図 2 2 - 2】

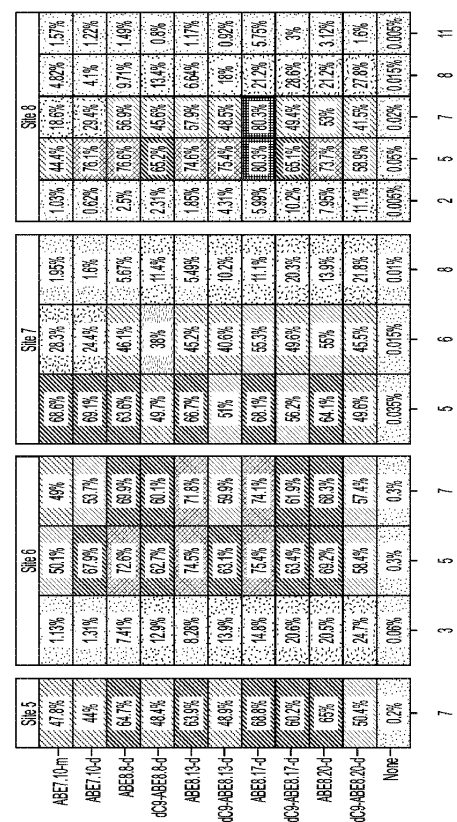


FIG. 22 CONTINUED

【図 2 3 A】

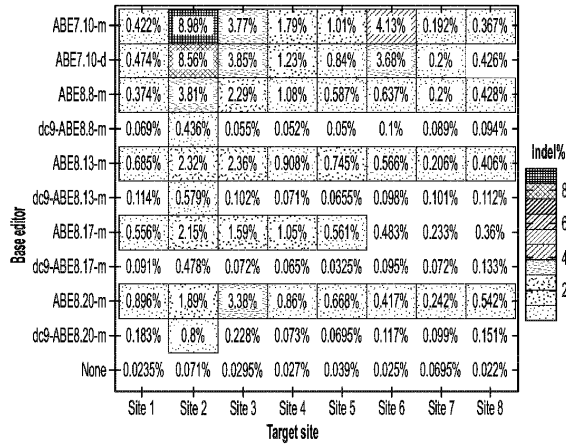


FIG. 23A

【図 2 3 B】

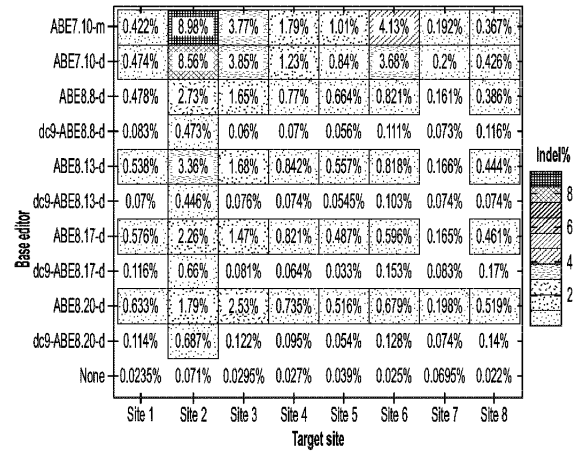
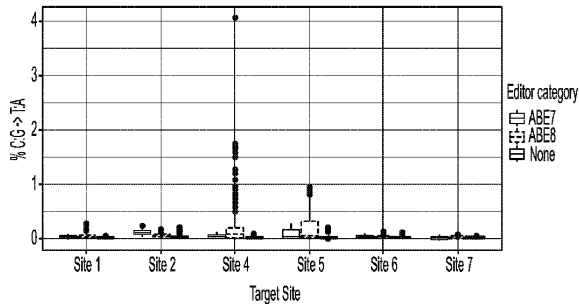


FIG. 23B

【図 2 4】



Site	target DNA sequence
1	GAACACAAAGCATAGACTGC
2	GGGAAGAGCCAGCATCCGT
3	GCTCCCATCACATCAACCGG
5	GGCTTCAGGTTCTAAATGAG
6	GCAGAGAGTCCCGTCTCCA
7	GTGTAAGACCTCAAAAGCAC

FIG. 24

【図 2 5 A】

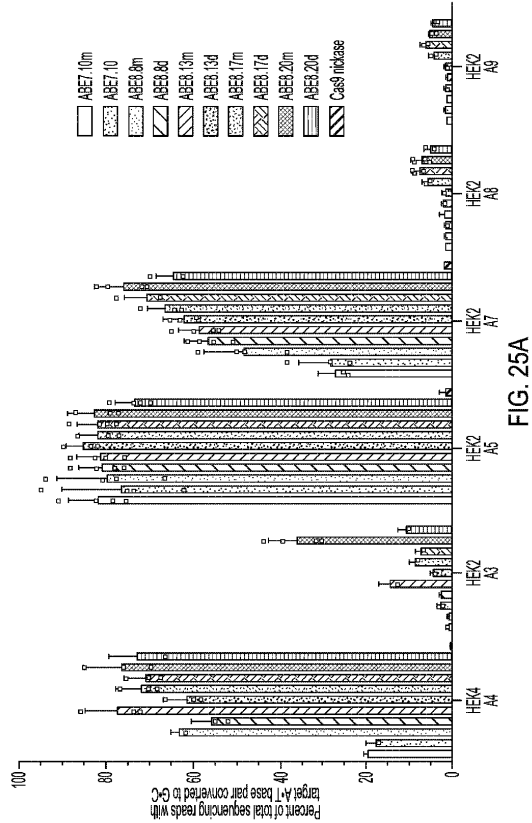


FIG. 25A

10

20

30

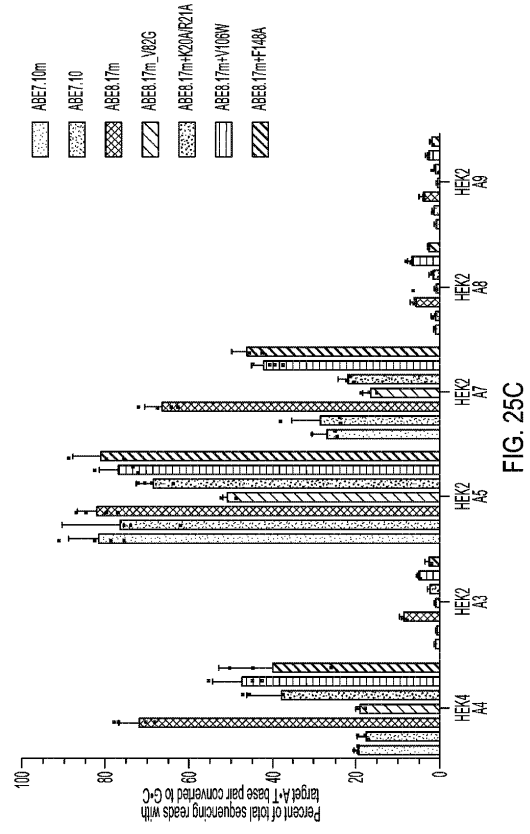
40

50

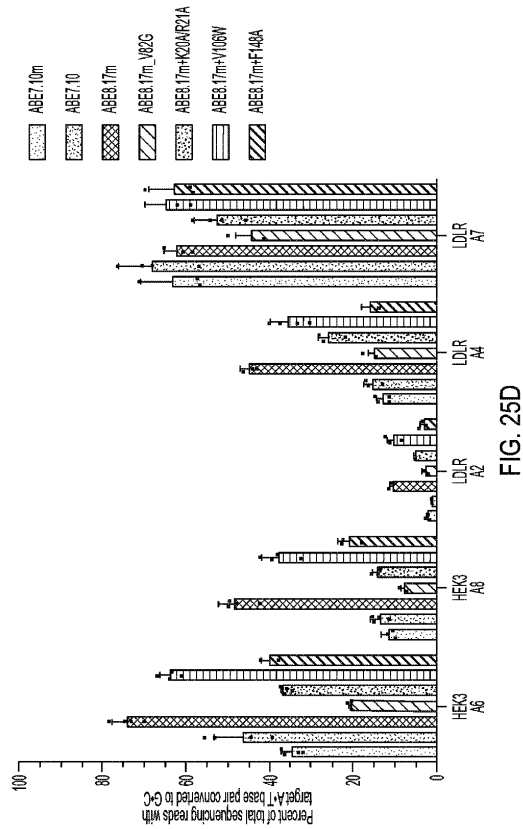
【図 2 5 B】



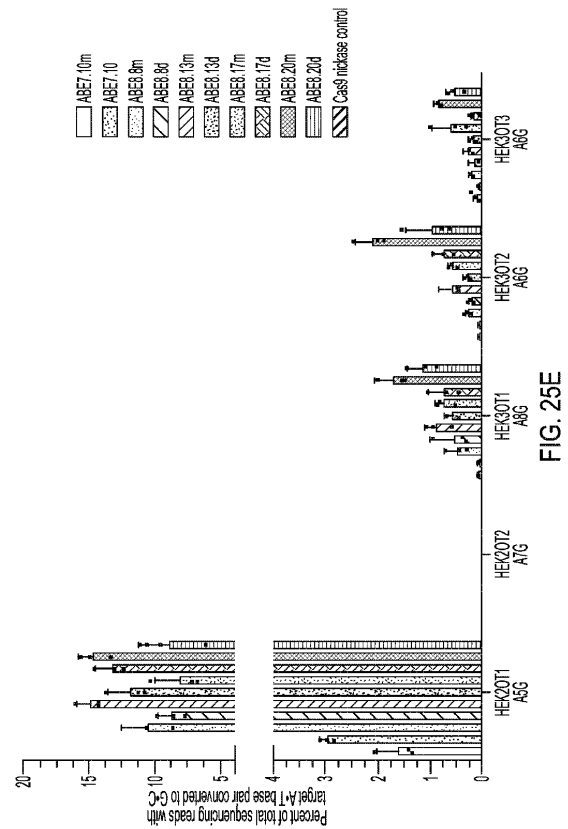
【図 2 5 C】



【図 2 5 D】



【図 2 5 E】



10

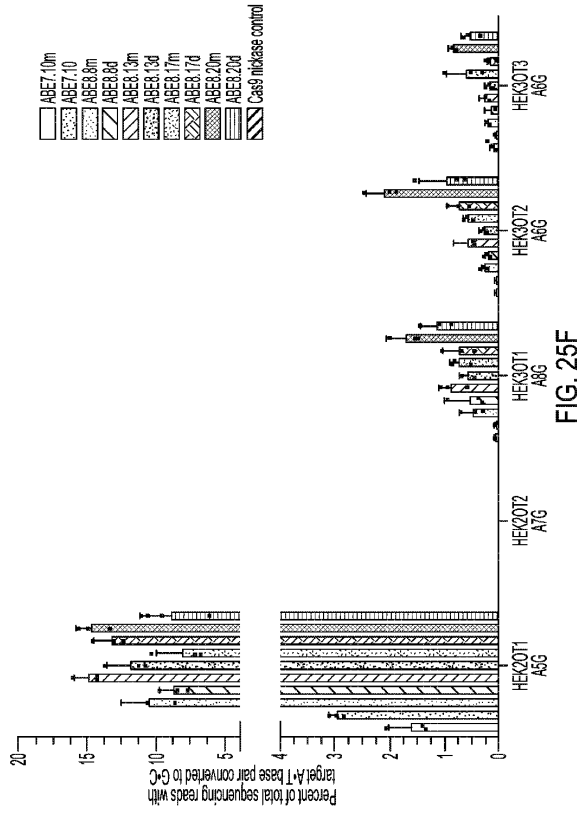
20

30

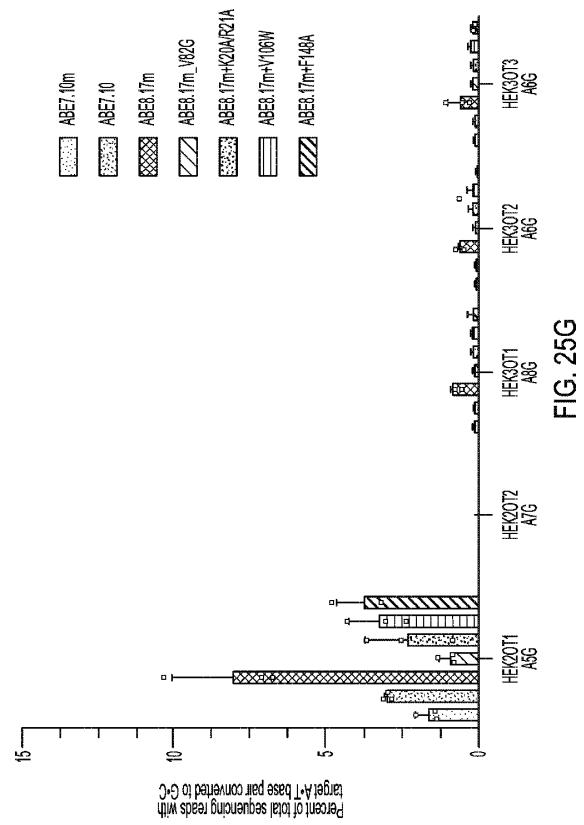
40

50

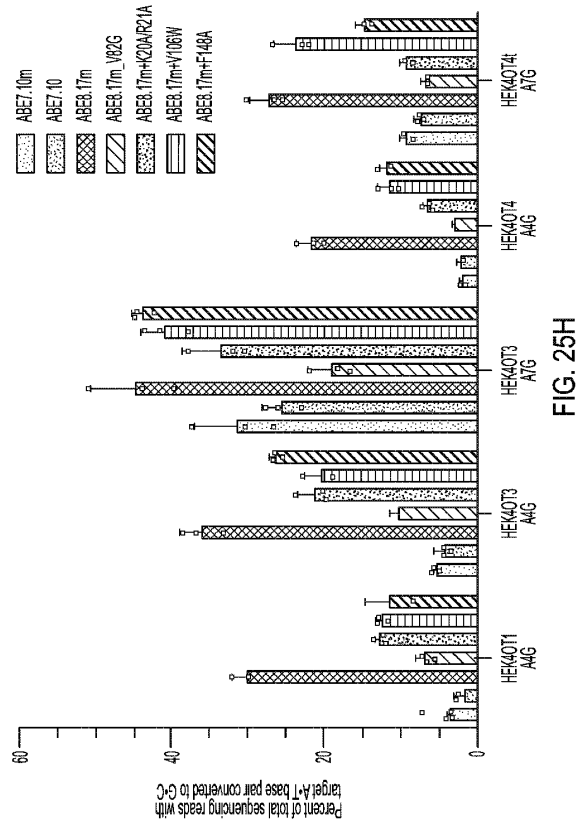
【図 25 F】



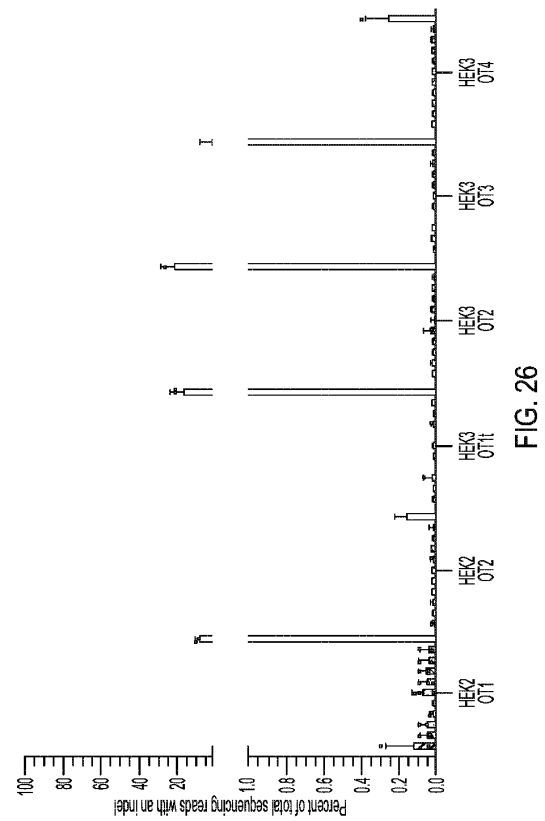
【図 25 G】



【図 25 H】



【図 26 - 1】



10

20

30

40

50

【図 26 - 2】

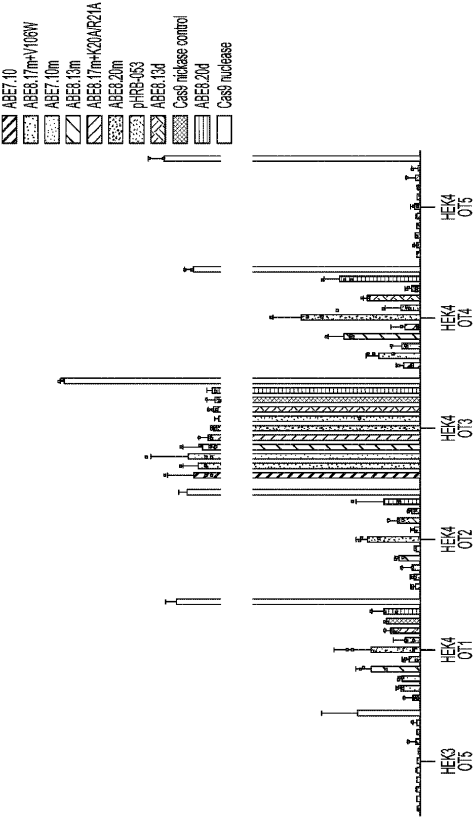


FIG. 26 CONTINUED

【図 27 A】

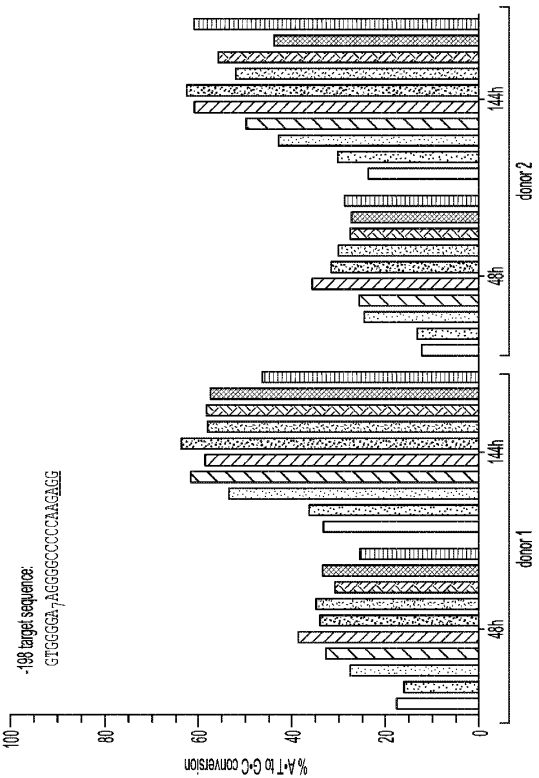


FIG. 27A

【図 27 B】

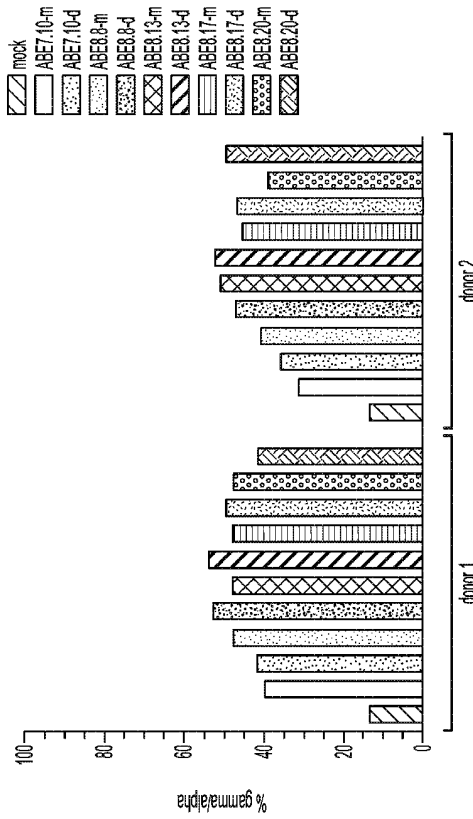


FIG. 27B

【図 28 A】

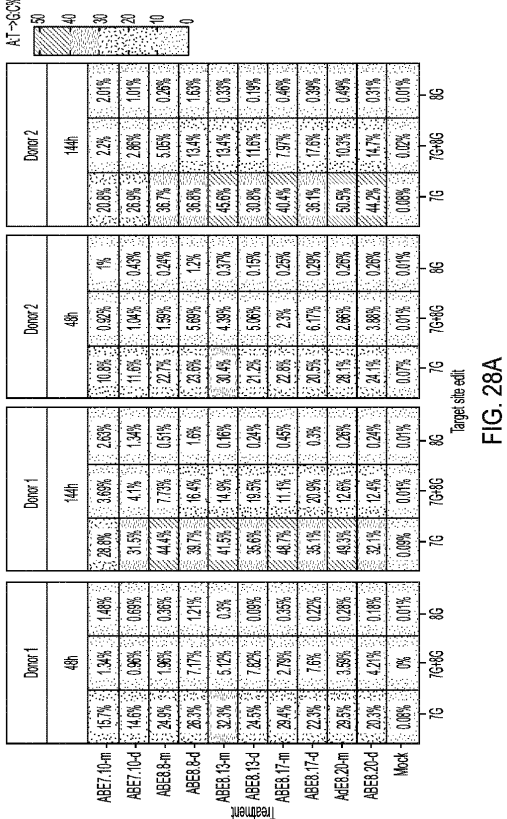


FIG. 28A

10

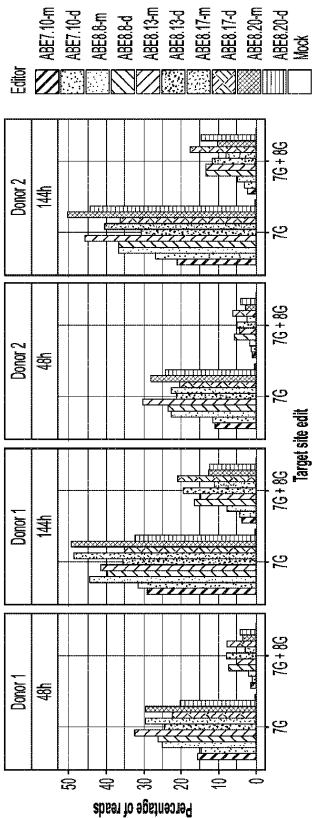
20

30

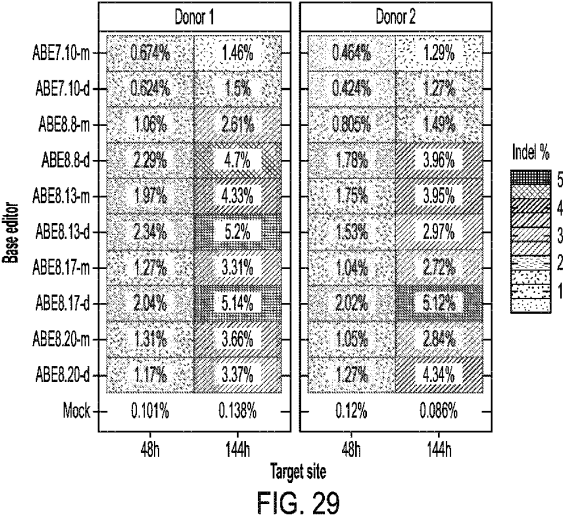
40

50

【図 28 B】



【図 29】



【図 30】

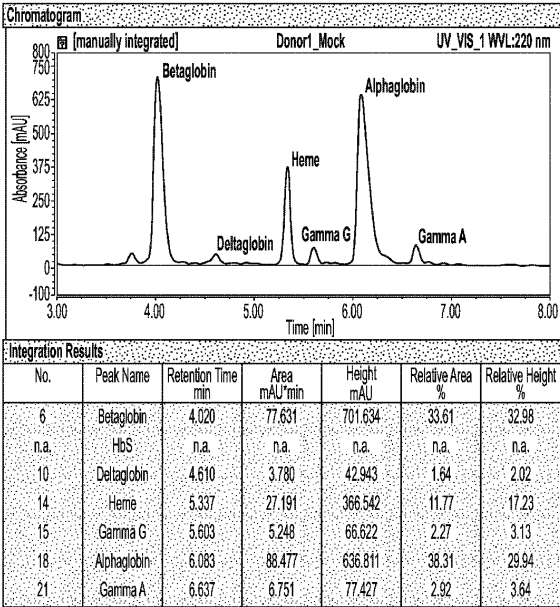


FIG. 30

【図 31】

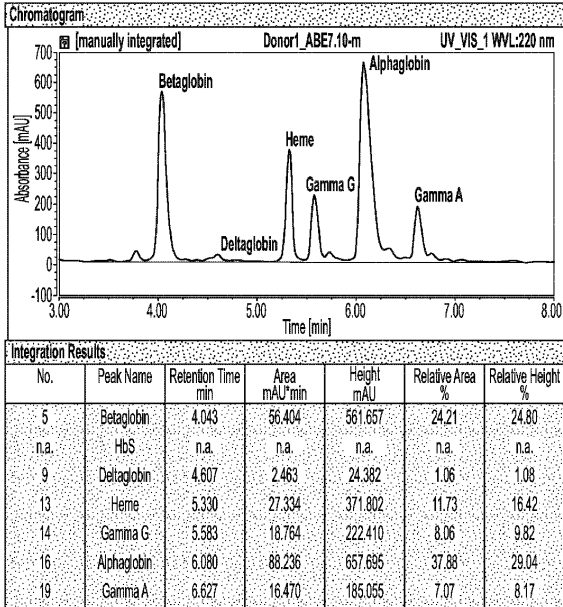


FIG. 31

【図 3 2】

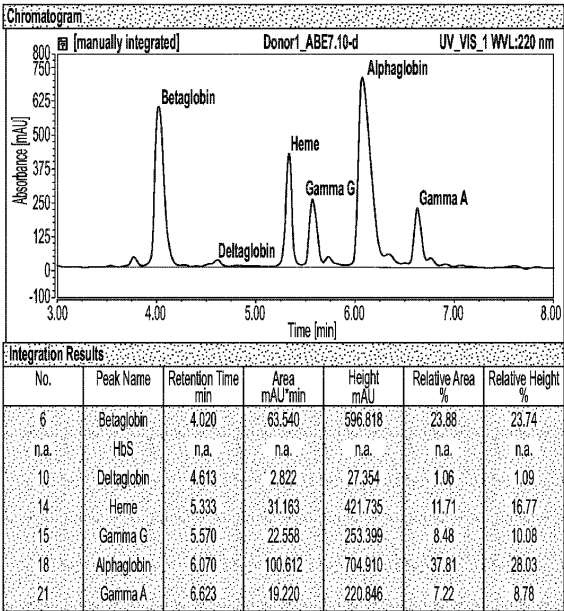


FIG. 32

【図 3 3】

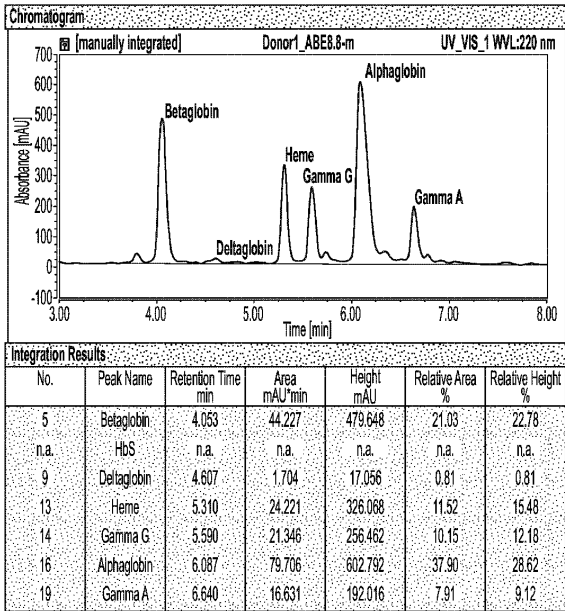


FIG. 33

【図 3 4】

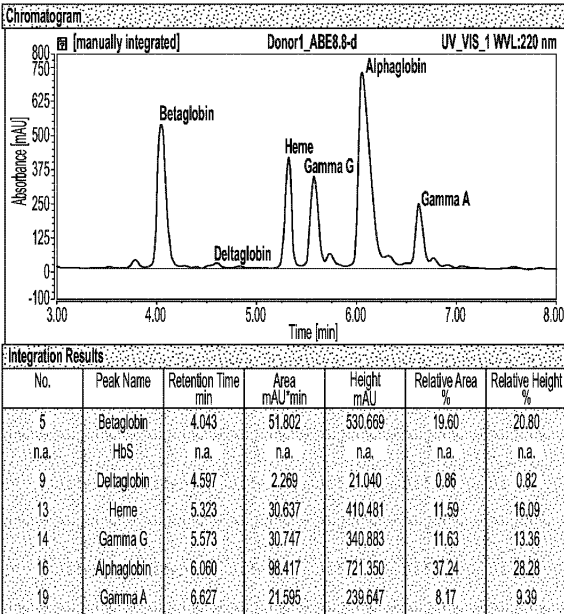


FIG. 34

【図 3 5】

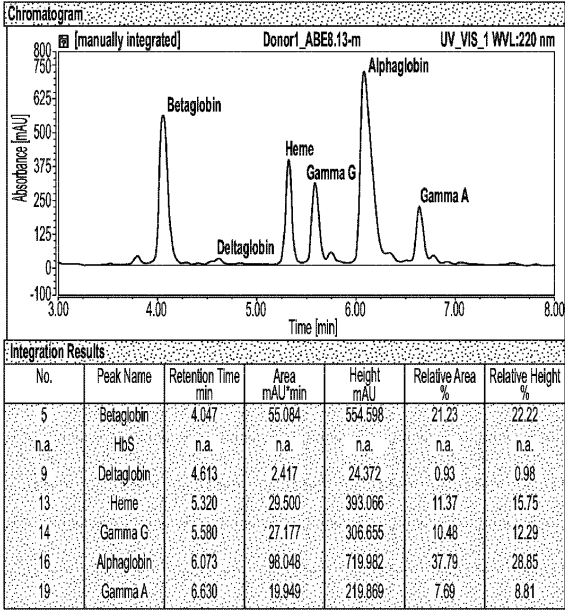


FIG. 35

【図 3 6】

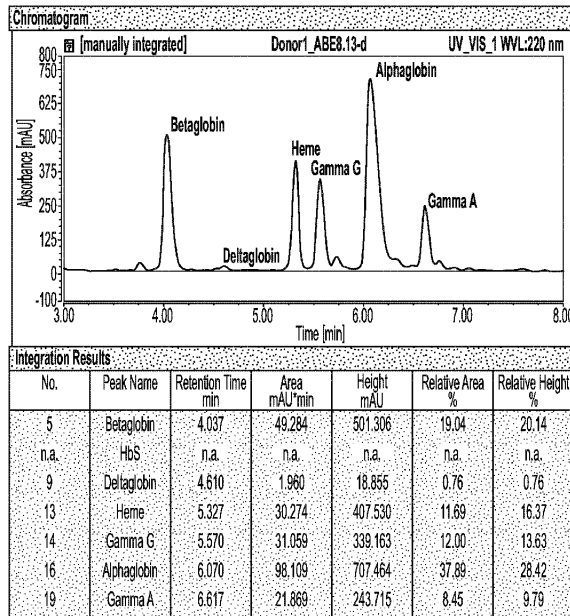


FIG. 36

【図 3 7】

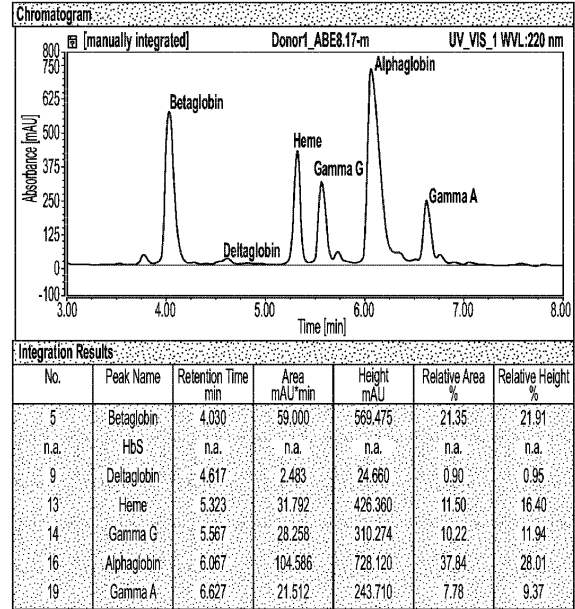


FIG. 37

【図 3 8】

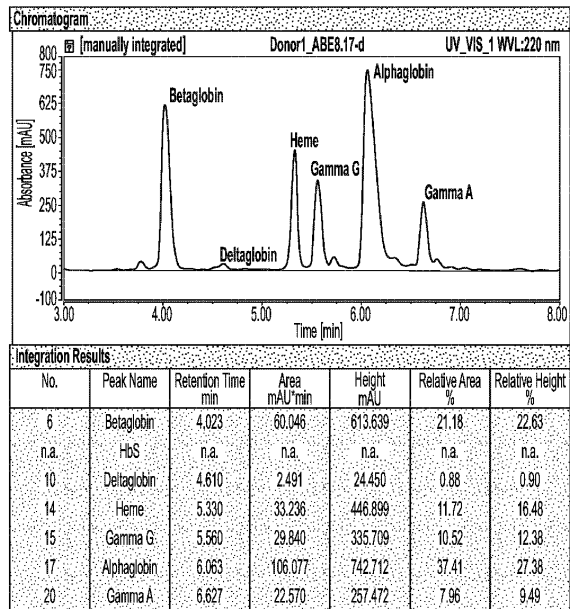


FIG. 38

【図 3 9】

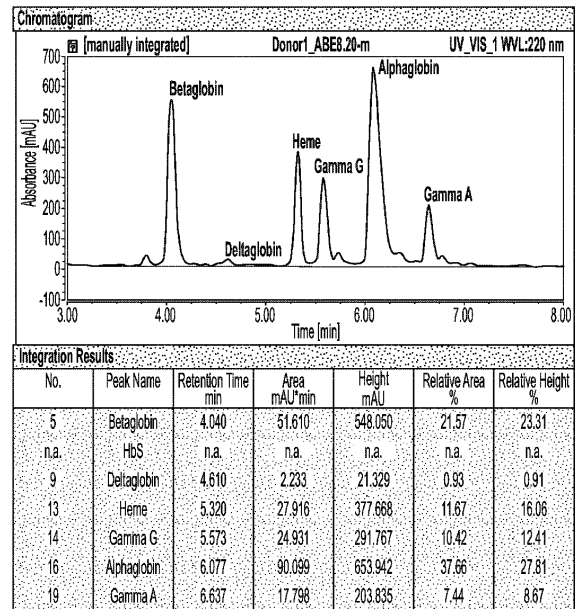


FIG. 39

10

20

30

40

50

【図 40】

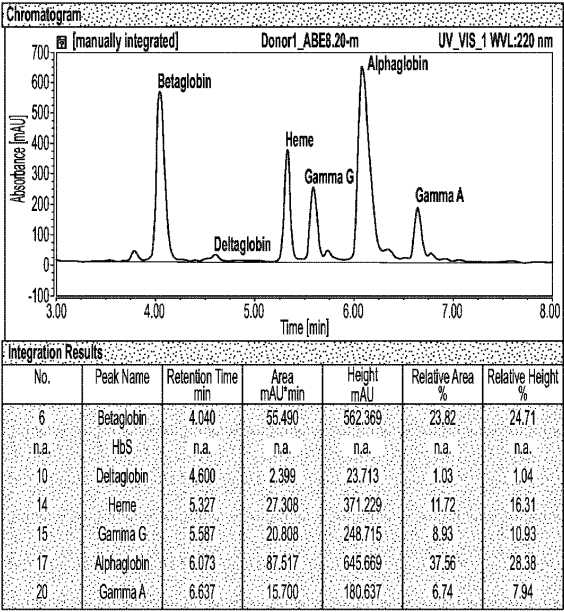


FIG. 40

【図 41】

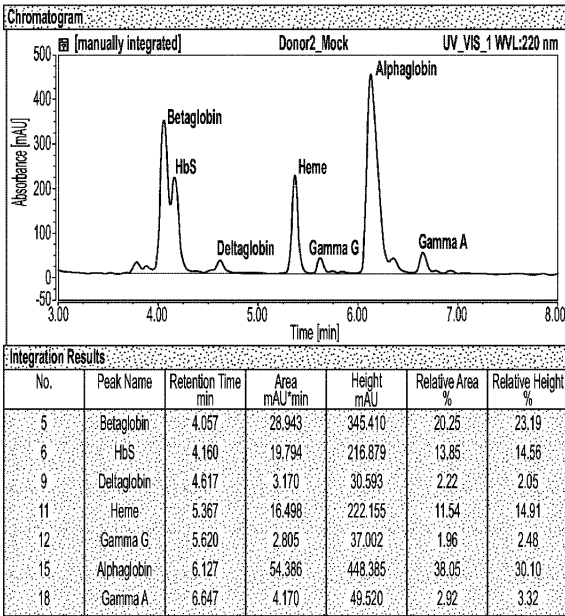


FIG. 41

【図 42】

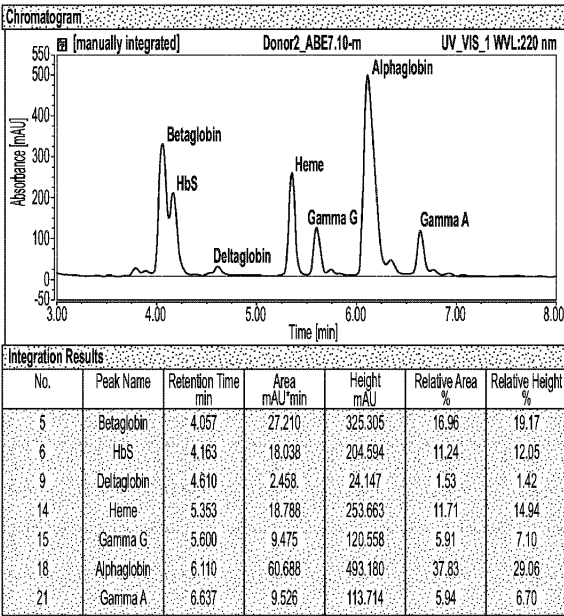


FIG. 42

【図 43】

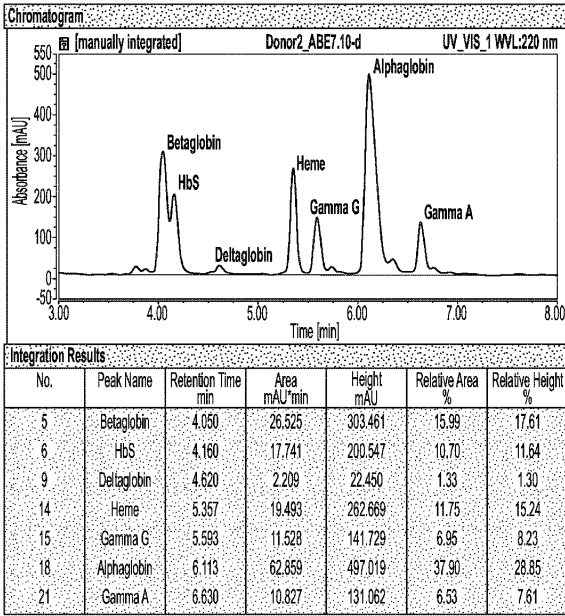


FIG. 43

10

20

30

40

50

【 図 4 4 】

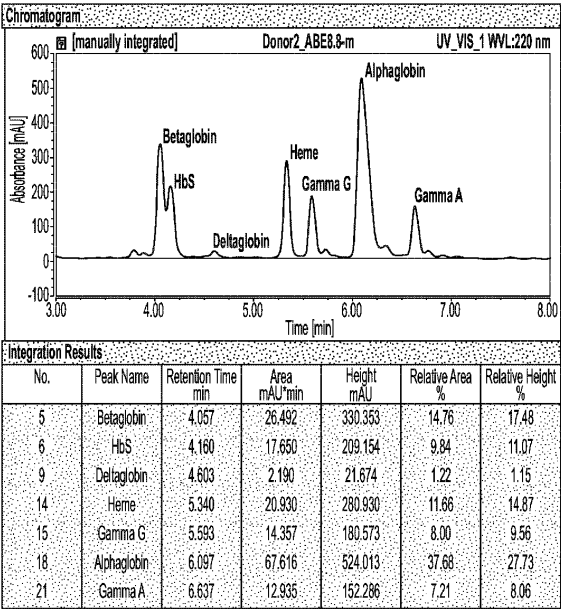


FIG. 44

【 図 4 5 】

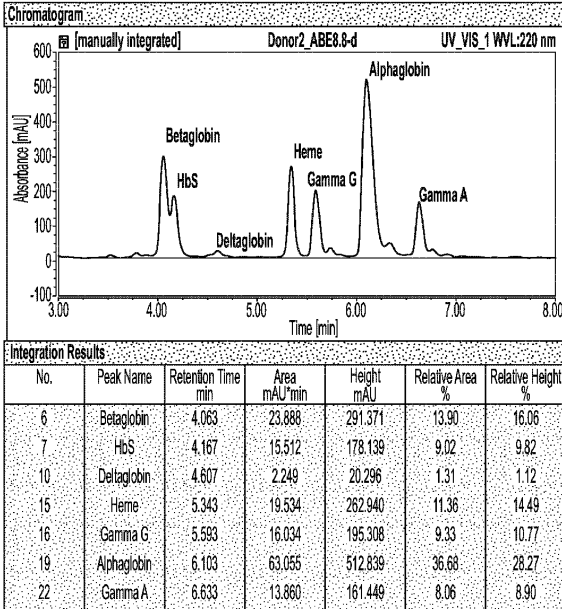


FIG. 45

【 図 4 6 】

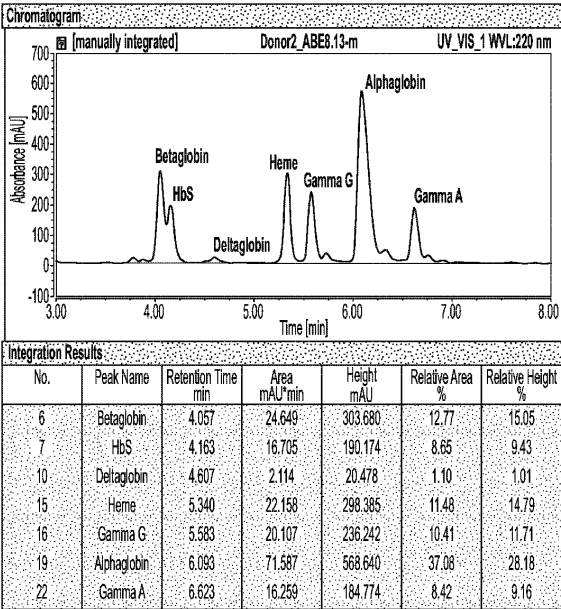


FIG. 46

【 図 4 7 】

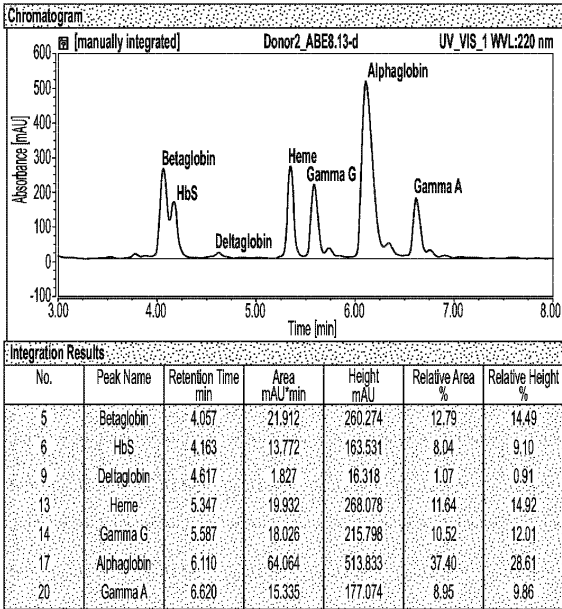


FIG. 47

10

20

30

40

50

【図 48】

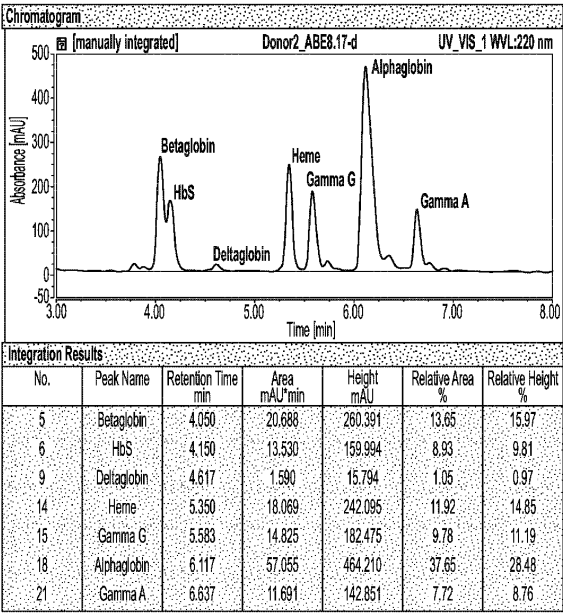


FIG. 48

【図 49】

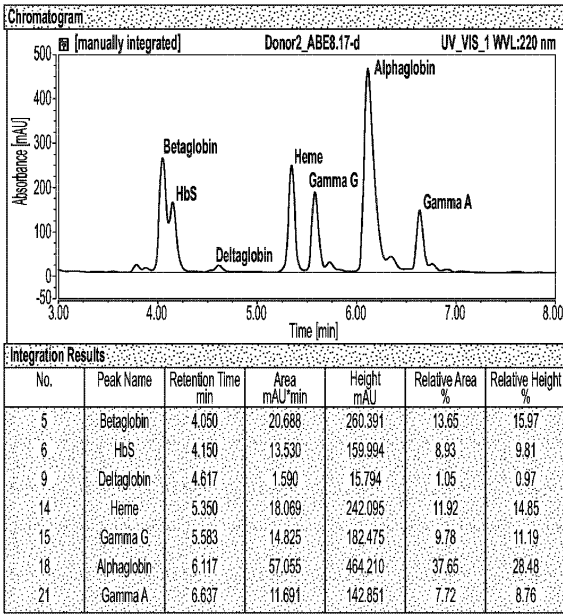


FIG. 49

【図 50】

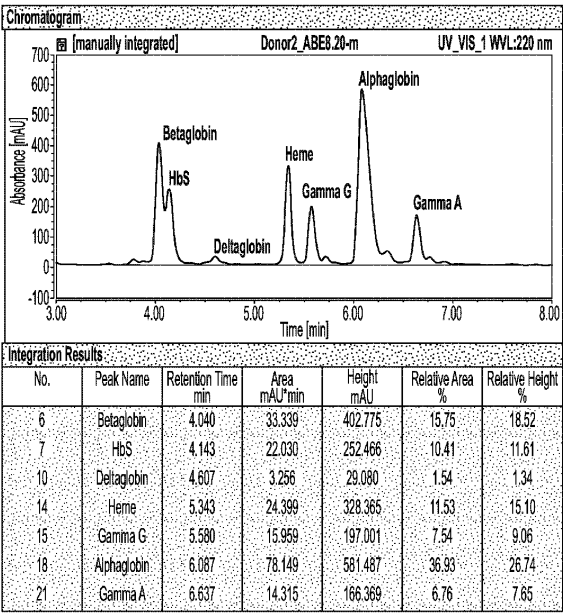


FIG. 50

【図 51 - 1】

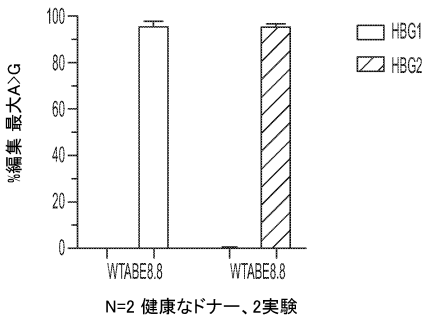


FIG. 51A

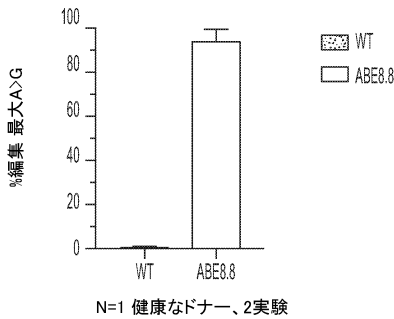


FIG. 51B

10

20

30

40

50

【図 5 1 - 2】

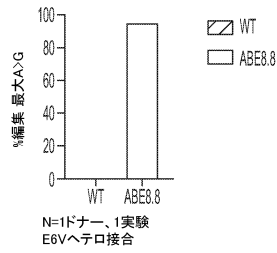


FIG. 51C

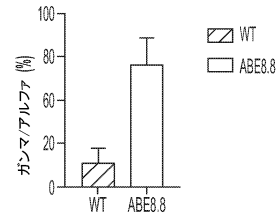


FIG. 51D

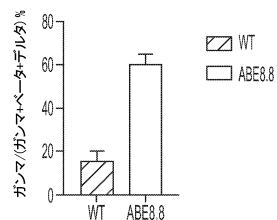


FIG. 51E

【図 5 2 - 1】

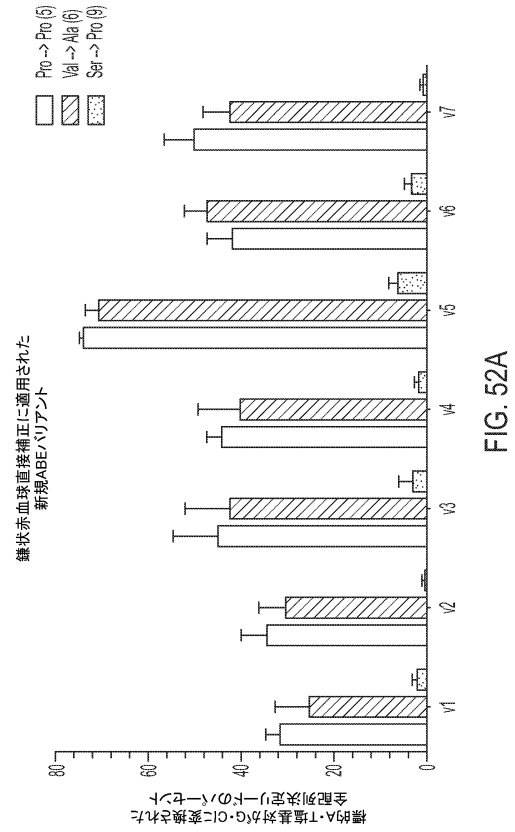


FIG. 52A

【図 5 2 - 2】

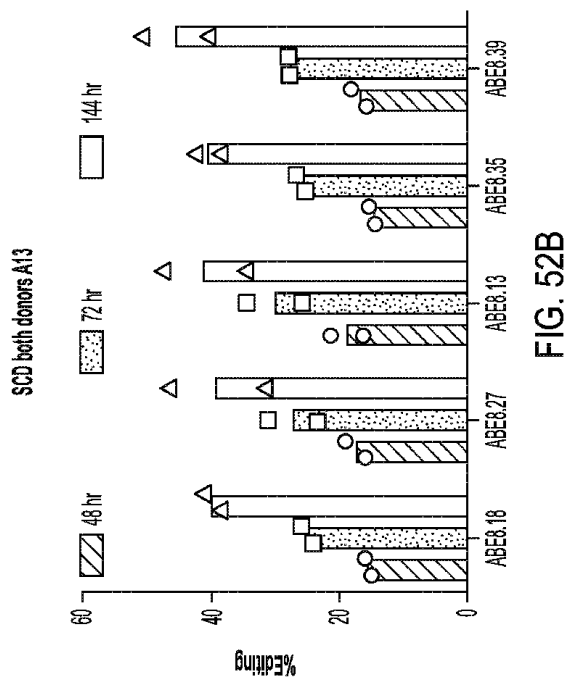


FIG. 52B

【図 5 3 - 1】

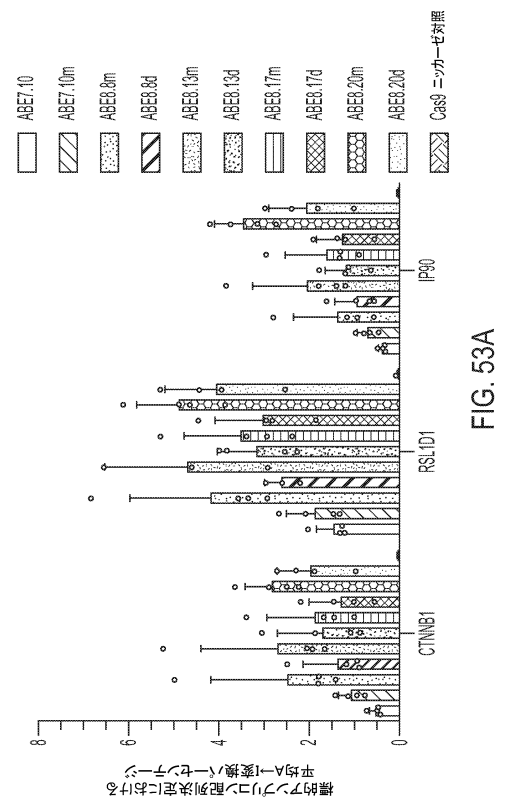


FIG. 53A

10

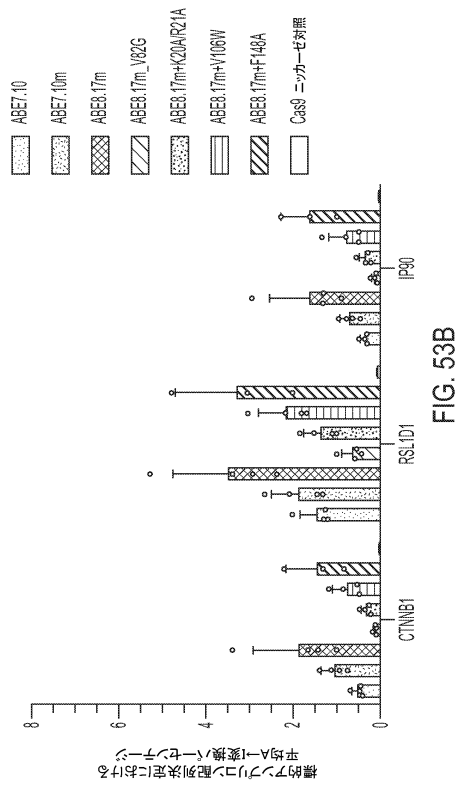
20

30

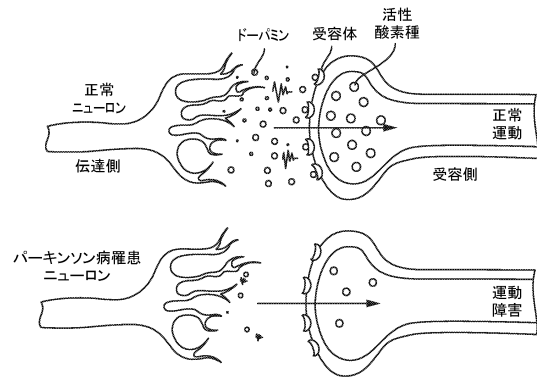
40

50

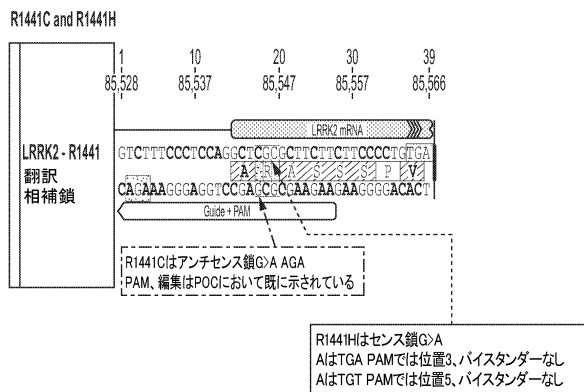
【図 5 3 - 2】



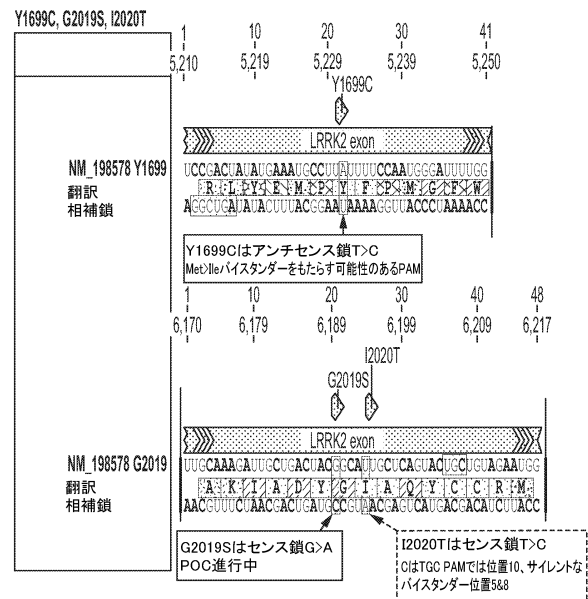
【図 5 4】



【図 5 5】



【図 5 6】



10

20

30

40

50

【図 5 7 A】

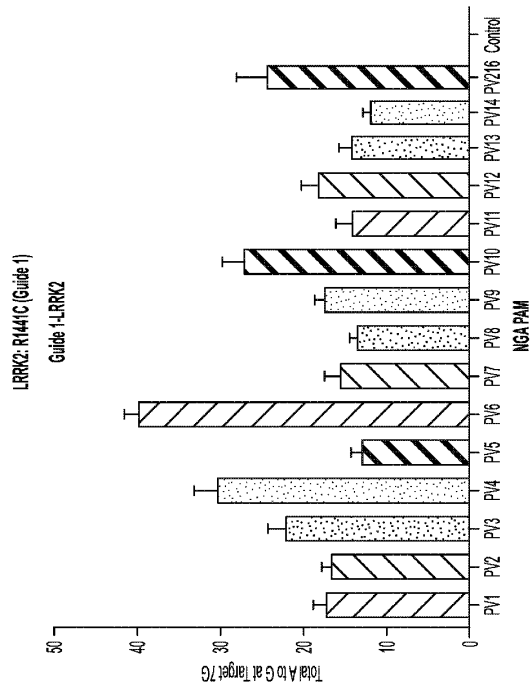


FIG. 57A

【図 5 7 B】

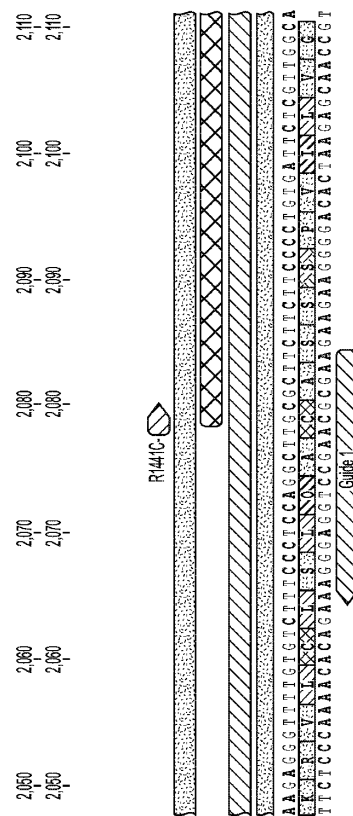


FIG. 57B

【図 5 7 C】

[illegible]

FIG. 57C

【図 5 8 A - 1】

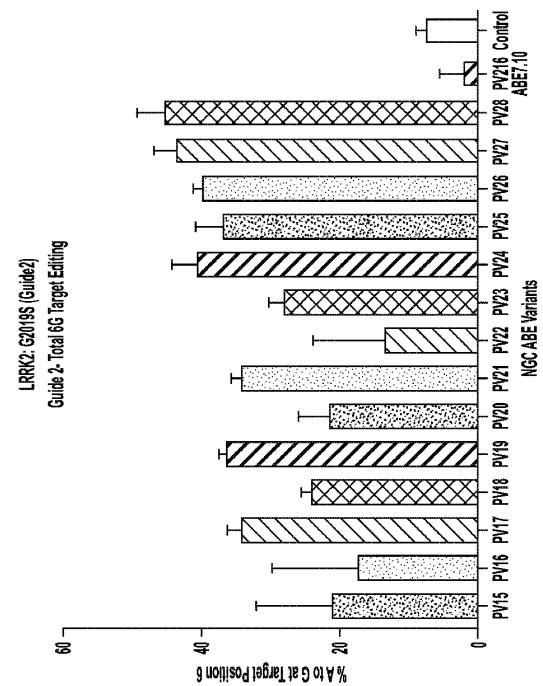


FIG. 58A

【図 5 9 C - 2】

PCR BT248																							
S02PPI1-G1-4CM42-2638																							
T	0.01%	0%	0.02%	0.02%	0.01%	0.01%	0.03%	0%	0.01%	0%	0.01%	0.03%	0.01%	0%	0.01%	0.01%	0.01%	0%	0.01%	0.01%	0.01%	0.01%	0.01%
G	0.01%	0.02%	0.02%	0.02%	0%	0.01%	0%	0.03%	0%	0%	0%	0.03%	0.01%	0%	0.01%	0.01%	0.01%	0.02%	0.01%	0.02%	0.01%	0.01%	0%
C	0.01%	0%	0%	0%	0.03%	0%	0.03%	0%	0%	0.01%	0%	0.03%	0.01%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
A	0.01%	0.02%	0.02%	0.02%	0.01%	0.01%	0.03%	0%	0.01%	0%	0.01%	0.03%	0.01%	0%	0.01%	0.01%	0.01%	0.02%	0.01%	0.02%	0.01%	0.01%	0.01%
#	2565	3441	3352	2417	3508	3442	3353	3365	3332	2406	3332	3365	3332	2406	3332	3365	3332	2406	3332	3365	3332	2406	3332
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	

PCR BT248																							
S02PPI1-G1-4CM42-2638																							
T	0%	0%	0.01%	0.01%	0.01%	0%	0%	0.01%	0.01%	0%	0.01%	0.01%	0.01%	0%	0.01%	0.01%	0.02%	0.01%	0%	0.01%	0.01%	0%	0%
G	0%	0.06%	0.06%	0.01%	0.06%	0%	0.07%	0.02%	0.01%	0.01%	0.01%	0.06%	0.06%	0.01%	0.01%	0.06%	0.06%	0.02%	0.01%	0%	0.06%	0.01%	0%
C	0%	0%	0%	0%	0.06%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
A	0.06%	0.06%	0%	0.04%	0.02%	0.02%	0.06%	0.01%	0.01%	0.02%	0%	0.06%	0.06%	0.02%	0.06%	0.01%	0%	0.02%	0.01%	0.02%	0.06%	0.01%	0.06%
#	3503	3405	3505	3500	3572	3516	3530	3601	3624	3524	3504	3594	3626	3616	3601	3615	3616	3566	3614	3606	3606	3636	3636

FIG. 59C CONTINUED

【図 5 9 D - 1】

PCR BT248																							
S02-PIG-G1-4CM17-2616																							
T	0.01%	0%	0.01%	0%	0%	0%	0%	0%	0.01%	0%	0%	0.01%	0.05%	0.02%	0.01%	0.02%	0.01%	0.02%	0%	0.01%	0%	0.01%	0%
G	0.11%	0.05%	0.05%	0%	0.05%	0%	0.05%	0%	0.05%	0%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.02%	0.02%	0.02%	0.02%
C	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0.05%	0%	
A	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	
#	3075	3012	3049	2938	2971	2967	3034	3035	3046	2945	3012	2966	3028	3046	3035	3046	3035	3046	3035	3046	3035	3046	
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	

PCR BT248																							
S02-PIG-G1-4CM17-2617																							
T	0.01%	0%	0.05%	0.01%	0.01%	0.01%	0%	0%	0%	0.01%	0.01%	0.02%	0.02%	0%	0%	0.01%	0.01%	0.01%	0%	0%	0.02%	0%	
G	0.01%	0.05%	0.05%	0.01%	0.01%	0.02%	17.19%	1.61%	0.01%	0%	0.05%	0.02%	0.02%	0.01%	0%	0.05%	0.05%	0.01%	0.01%	0.01%	0.02%	0.02%	
C	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	
A	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	
#	2935	2901	2936	2936	2906	2914	2910	2938	2936	2934	2932	2935	2915	2931	2936	2934	2933	2945	2946	2937	2935	2934	
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	

FIG. 59D

【図 5 9 D - 2】

PCR BT248																							
S02PPI1-G1-4CM17-2616																							
T	0%	0.03%	0.02%	0.1%	0.03%	0.01%	0%	0%	0.03%	0.01%	0%	0.03%	0.01%	0%	0.02%	0.02%	0%	0%	0%	0%	0.01%	0.01%	
G	0%	0.1%	0.01%	0.01%	0.01%	0.01%	0%	0.03%	0.01%	0%	0.03%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.02%	0%	0.02%	0.01%	0.01%	
C	0.01%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0.03%	0.01%	0%	
A	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0%	0.03%	0.01%	0%	0.03%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.02%	0.01%	0.02%	0.01%	0.01%	
#	2202	2178	2178	2130	2157	2161	2164	2130	2164	2101	2154	2130	2164	2101	2154	2130	2164	2101	2154	2130	2164	2101	
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	

PCR BT248																							
S02PPI1-G1-4CM15-2640																							
T	0%	0%	0.03%	0.02%	0.01%	0%	0%	0%	0.03%	0.01%	0.02%	0.02%	0.01%	0.01%	0.01%	0%	0%	0%	0%	0%	0.01%	0.02%	
G	0.05%	0.1%	0.03%	0%	0.01%	0%	0.05%	0.03%	0%	0%	0%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.02%	0.01%	0%	0%	
C	0%	0.01%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0.01%	0.01%	0%	
A	0.04%	0.05%	0.01%	0.02%	0.01%	0%	0.05%	0.02%	0%	0.05%	0.03%	0.02%	0%	0%	0%	0%	0%	0%	0%	0.01%	0.01%	0%	
#	2211	2162	2178	2100	2154	2157	2147	2167	2153	2141	2162	2173	2167	2204	2172	2143	2162	2130	2146	2181	2149	2251	2163

FIG. 59D CONTINUED

【図 5 9 E - 1】

PCR BT248																							
S02PPI1-G1-4CM17-261																							
T	0.01%	0%	0.01%	0.05%	0.04%	0.01%	0%	0%	0.01%	0%	0.01%	0%	0.02%	0.01%	0.02%	0.02%	0%	0.01%	0%	0%	0.01%	0%	
G	0.05%	0.01%	0.01%	0.01%	0.01%	0.01%	0%	0.01%	0.01%	0%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	
C	0%	0%	0%	0%	0%	0%	0%	0%	0.01%	0.01%	0%	0.01%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	
A	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0%	0.01%	0.01%	0%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	
#	2552	2547	2552	2553	2550	2553	2547	2552	2553	2547	2552	2553	2547	2552	2553	2547	2552	2553	2547	2552	2553	2547	
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	

PCR BT248																							
S02PPI1-G1-4CM17-262																							
T	0%	0%	0.02%	0.02%	0.02%	0.05%	0%	0.05%	0.05%	0.02%	0%	0.05%	0.05%	0.04%	0%	0.01%	0%	0%	0.01%	0.02%	0.01%	0%	
G	0.02%	0.07%	0.01%	0.01%	0%	0%	0.01%	0.05%	0.04%	0%	0.01%	0.05%	0.05%	0.05%	0%	0%	0.08%	0.02%	0.02%	0.02%	0.01%	0%	
C	0%	0%	0%	0%	0%	0%	0%	0%	0.01%	0.01%	0%	0.05%	0%	0%	0%	0%	0%	0%	0%	0.12%	0%	0%	
A	0.05%	0.01%	0.01%	0.01%	0.01%	0.01%	0%	0.01%	0.01%	0%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	
#	2552	2547	2552	2553	2550	2553	2547	2552	2553	2547	2552	2553	2547	2552	2553	2547	2552	2553	2547	2552	2553	2547	
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	

FIG. 59E

10

20

30

40

50

【図 5 9 E - 2】

PCR B1248																								
S03-PH-G-AQM16-2619																								
T	0%	0%	0.01%	0%	0.02%	0.01%	0.01%	0%	0.01%	0%	0%	0.01%	0.01%	0.01%	0.02%	0.01%	0%	0.01%	0.01%	0%	0.01%	0.01%	0.02%	
G	0.01%	0.01%	0.01%	0%	0.01%	0.01%	0.47%	1.5%	0.01%	0%	0%	0.02%	0.01%	0.01%	0%	0.01%	0.01%	0.01%	0.01%	0%	0.01%	0.01%	0.01%	
C	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0.05%	0%	0%	
A	0.01%	0.01%	0%	0.05%	0.02%	0.15%	0.53%	0%	0.02%	0.03%	0%	0.03%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	
#	39059/39057/39055/39021/39098/39010/39090/39007/39060/39042/39090/39003/40024/39070/40156/40232/40085/40178/40101/40034/40060/39087/	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23

PCR B1248																								
S03-PH-G-AQM16-2620																								
T	0.01%	0%	0.01%	0.01%	0%	0%	0.01%	0%	0.01%	0.01%	0.01%	0.01%	0.01%	0%	0%	0.01%	0.01%	0%	0%	0.01%	0.01%	0%	0%	
G	0%	0.02%	0.01%	0.01%	0%	0%	0.01%	0%	0.01%	0.01%	0%	0%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.02%	0.01%	0.01%	0.01%	
C	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0.01%	0%	0%	
A	0.01%	0.01%	0%	0.02%	0.01%	0.23%	0.61%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.02%	0%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	
#	39059/39057/39055/39021/39098/39010/39090/39007/39060/39042/39090/39003/40024/39070/40156/40232/40085/40178/40101/40034/40060/39087/	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23

FIG. 59E CONTINUED

【図 5 9 F - 1】

PCR B1248																								
S03-PH-G-AQM16-2621																								
T	0%	0%	0.01%	0.02%	0.01%	0.02%	0%	0%	0.01%	0.01%	0%	0.01%	0.03%	0%	0.01%	0.03%	0.02%	0%	0%	0.02%	0%	0.01%		
G	0.01%	0.05%	0.02%	0%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0%	0%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0%	0.02%	0.01%	0.01%		
C	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		
A	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0%	0.01%	0.01%	0.01%		
#	39059/39017/39055/39021/39098/39010/39090/39007/39060/39042/39090/39003/40024/39070/40156/40232/40085/40178/40101/40034/40060/39087/	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23

PCR B1248																								
S04-PH-G-AQM16-2640																								
T	0%	0%	0.03%	0.01%	0%	0%	0.07%	0%	0.01%	0%	0%	0.01%	0.01%	0%	0.02%	0.03%	0.01%	0%	0%	0.01%	0.01%	0%	0%	
G	0.4%	0.05%	0.08%	0%	0.02%	1.45%	0.05%	0%	0%	0.4%	0.04%	0.04%	0.01%	0.01%	0.02%	0.03%	0.01%	0.02%	0.01%	0.01%	0.01%	0.01%	0%	
C	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	
A	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0%	0.01%	0.01%	0.01%	0%	
#	4252/42024/41917/41640/41788/42190/41944/42161/41226/41994/41803/42310/42485/42035/42642/42638/4231/42523/42129/42514/42331/42103/	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23

FIG. 59F

【図 5 9 F - 2】

PCR B1248																								
S04-PM2.5-AQM16-264																								
T	0%	0%	0.01%	0.03%	0.05%	0.06%	0%	0%	0%	0.01%	0.04%	0.01%	0.01%	0.01%	0.01%	0%	0.01%	0.01%	0%	0.01%	0.01%	0%		
G	0.04%	0.1%	0.01%	0.02%	0%	0.02%	0%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0%	0.01%	0.01%	0%		
C	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		
A	0.01%	0.01%	0.01%	0.02%	0.03%	0.01%	0.01%	0.01%	0%	0.01%	0.01%	0.01%	0.01%	0.01%	0.02%	0.01%	0.01%	0.01%	0%	0%	0%	0%		
#	39045/30175/39027/39007/39065/39041/39098/39046/30196/39040/39023/39043/39095/39007/39060/39042/39090/39003/40024/39070/40156/40232/40085/40178/40101/40034/40060/39087/	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23

PCR B1248																								
S04-PM2.5-AQM16-2645																								
T	0%	0%	0.01%	0.02%	0.02%	0.01%	0%	0%	0.01%	0.01%	0.01%	0.01%	0.02%	0.02%	0.01%	0%	0.02%	0.01%	0%	0%	0.01%	0.01%		
G	0.01%	0.04%	0.01%	0%	0.02%	0%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0%	0.01%	0.01%	0%		
C	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		
A	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0%	0%	0.01%	0.01%	0.01%	0.01%	0.02%	0.02%	0.01%	0%	0.01%	0.01%	0%	0.01%	0.01%	0.01%		
#	39045/30175/39027/39007/39065/39041/39098/39046/30196/39040/39023/39043/39095/39007/39060/39042/39090/39003/40024/39070/40156/40232/40085/40178/40101/40034/40060/39087/	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23

FIG. 59F CONTINUED

【図 5 9 G - 1】

PCR B1248																								
S04-P16-G-AQM19-2622																								
T	0%	0%	0.01%	0.01%	0.01%	0%	0%	0%	0%	0%	0%	0%	0.02%	0.02%	0.01%	0.01%	0.02%	0%	0%	0%	0%	0%	0%	
G	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	
C	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	
A	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	
#	3536/3729/3723/3682/3698/3704/3745/3776/3742/3689/3763/3700/3746/3704/3747/3728/3729/3750/3769/3737/3759/3756/3749/	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23

PCR B1248																								
S04-P16-G-AQM19-2623																								
T	0%	0%	0%	0.01%	0.01%	0.02%	0%	0%	0.01%	0.01%	0%	0%	0%	0.02%	0.01%	0.01%	0.02%	0.01%	0.01%	0%	0%	0%	0%	
G	0.01%	0.2%	0%	0.01%	0%	0.01%	0%	0.01%	0%	0%	0%	0%	0%	0.02%	0.02%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0%	
C	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0.02%	0.02%	0.02%	0%	0%	0%	0%	0%	0.2%	0%	
A	0.01%	0.01%	0%	0.02%	0.02%	0.02%	0.01%	0.01%	0.01%	0.01%	0%	0%	0%	0.02%	0.02%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.02%	
#	3649/3624/3616/3594/3590/3653/3594/3608/3571/3659/3626/3612/3642/3650/3647/3682/3676/3655/3671/3632/3653/3657/3651/	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23

FIG. 59G

【図 5 9 I - 2】

PCR BT248																							
S86P14-G14QM2-2620																							
T	0.01%	0%	0.01%	0.15%	0.01%	0.01%	0.01%	0.11%	0.12%	0%	0%	0.01%	0.02%	0.01%	0%	0.01%	0.01%	0.01%	0%	0%	0.12%	0%	0.01%
G	0.11%	0.14%	0.03%	0%	0.05%	0.01%	0.01%	0.05%	5.10%	0.01%	0%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0%	0.01%	0.01%	
C	0%	0%	0.11%	0.05%	0%	0.05%	0%	0%	0%	0.01%	0.01%	0.01%	0.01%	0%	0%	0.01%	0%	0.01%	0%	0%	0.05%	0.12%	0%
A	0.05%	0.05%	0.01%	0.01%	0.03%	0%	0.03%	0.41%	0.41%	0%	0.04%	0.04%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	
#	39033	3822	3948	3852	3816	3822	3855	3829	3846	3830	3862	3858	3846	4015	3876	4035	4050	4015	4025	4057	4016	4002	
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	

PCR BT248																						
S86P14-G14QM2-2621																						
T	0.01%	0%	0.01%	0.01%	0.01%	0%	0.01%	0.01%	0.02%	0.01%	0%	0%	0.01%	0.01%	0%	0%	0.01%	0.01%	0%	0%	0.01%	0%
G	0.02%	0.08%	0.01%	0%	0.01%	0%	0.01%	0%	0%	0%	0%	0%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.02%
C	0%	0%	0%	0%	0.01%	0%	0.01%	0.05%	0%	0.01%	0.01%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
A	0.01%	0.01%	0.01%	0.02%	0.02%	0.02%	0.02%	0.02%	0.01%	0.01%	0%	0%	0.01%	0.01%	0%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%
#	39031	39055	39030	3855	3830	3827	3858	3824	3803	3805	3816	3874	3830	4016	3897	4021	4023	4037	4022	4004	4017	4016
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23

FIG. 59I CONTINUED

【図 5 9 J - 1】

PCR_BT248																							
S87-P17-G14QM2-2628																							
T	0.01%	0%	0.08%	0.01%	0.02%	0.01%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	
G	0%	0.01%	0.01%	0%	0.01%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	
C	0%	0%	0%	0%	0.01%	0%	0.04%	0%	0.01%	0.01%	0.01%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	
A	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0%	0%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0%	
#	4101	4101	4084	4055	4050	4073	4121	4083	4133	4043	4015	4013	4112	4130	4181	4153	4152	4145	4158	4134	4131	4135	4189
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	

PCR_BT248																						
S87P17-G14QM2-2629																						
T	0.01%	0%	0.01%	0.02%	0%	0.02%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
G	0.4%	0.07%	0.01%	0%	0%	0.01%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
C	0.01%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
A	0.01%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
#	37189	36549	36584	36381	36548	36553	37040	36520	37055	36520	37055	37102	37250	37148	37470	37311	37428	37226	37326	37342	37342	37342
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23

FIG. 59J

【図 5 9 J - 2】

PCR BT248																							
S87P17-G14QM2-2630																							
T	0%	0.01%	0%	0%	0.04%	0.02%	0%	0%	0%	0%	0%	0%	0%	0%	0.01%	0.02%	0.01%	0%	0.01%	0.02%	0.01%	0%	0%
G	0.05%	0.15%	0.01%	0%	0.05%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0%	0%
C	0%	0%	0%	0%	0.01%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
A	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0%	0%	0%	0%	0%	0%	0%	0%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0%	0%
#	30061	2888	2801	2836	2826	2807	2800	2804	2838	28182	2878	2844	3009	3021	3005	3025	3035	3046	3031	3014	3016	3002	3007
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	

PCR BT248																						
S87P17-G14QM2-2632																						
T	0.08%	0%	0.01%	0.04%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
G	0.01%	0.01%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
C	0%	0.07%	0%	0.01%	0%	0.01%	0%	0.01%	0%	0.01%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
A	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
#	41382	41382	41382	41382	41382	41382	41382	41382	41382	41382	41382	41382	41382	41382	41382	41382	41382	41382	41382	41382	41382	41382
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23

FIG. 59J CONTINUED

【図 5 9 K - 1】

PCR BT248																						
S87P17-G14QM2-2633																						
T	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
G	0.01%	0.01%	0%	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%
A	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
#	41021	40780	40655	40327	40447	40853	40741	40910	38765	40633	40461	40815	41087	40832	41234	41278	41112	41344	41037	41087	41057	40914
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23

PCR BT248																						
S88P1216-G14QM2-2634																						
T	0.01%	0%	0%	0%	0.06%	0.01%	0.01%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
G	0%	0.01%	0%	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%
A	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
#	41327	40855	39889	38419	38588	38321	40391	40855	40768	39350	40066	39813	40235	40455	40625	40637	40843	41449	40555	40530	40534	40534
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23

【図 5 9 K - 2】

PCR BT248																					
S08-PI6-G-AQM2-2631																					
T	0.01%	0%	0.01%	0.02%	0.03%	0.01%	0%	0%	0.01%	0.01%	0%	0%	0.01%	0.01%	0.02%	0.02%	0.01%	0.01%	0%	0%	0.01%
G	0.01%	0.02%	0.02%	0.02%	0%	0%	0%	0.01%	0.02%	0%	0%	0.01%	0.02%	0.01%	0.01%	0.01%	0.01%	0%	0%	0.02%	0.01%
C	0%	0%	0%	0%	0%	0%	0%	0.02%	0.01%	0.02%	0%	0%	0.01%	0.01%	0.01%	0%	0%	0%	0%	0%	0%
A	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0%	0%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%
#	3335	3716	3804	3726	3784	3750	3740	3795	3812	3761	3802	3811	3859	3846	3845	3835	3836	38476	3845	3837	3821
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22

PCR BT248																					
S08-PI6-G-AQM2-2632																					
T	0.01%	0%	0.01%	0.01%	0.02%	0%	0%	0.01%	0.02%	0%	0%	0.01%	0.01%	0%	0.01%	0.01%	0.01%	0%	0%	0.01%	0.01%
G	0.01%	0.01%	0.01%	0%	0%	0%	0.02%	0.01%	0.01%	0%	0%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0%	0%	0.01%	0.01%
C	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
A	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0%	0%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%
#	3717	3786	3760	3742	3759	3740	3795	3812	3761	3802	3811	3859	3846	3845	3835	3836	38476	3845	3837	3821	3813
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22

FIG. 59K CONTINUED

【図 5 9 L - 1】

PCR BT248																					
S08-PI6-G-AQM2-2633																					
T	0.02%	0%	0.01%	0.03%	0.02%	0.01%	0%	0.01%	0.01%	0.01%	0%	0%	0.02%	0.03%	0.01%	0.01%	0%	0%	0%	0%	0%
G	0.01%	0.01%	0.01%	0.01%	0%	0%	0%	0.01%	0.01%	0%	0%	0%	0.01%	0.01%	0.01%	0.01%	0.01%	0.02%	0.01%	0%	0%
C	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
A	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0%	0%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%
#	2016	2040	2075	2091	2047	2035	2030	2070	2095	2099	2078	2065	2055	2070	2092	2093	2071	2093	2091	2093	2093
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22

PCR BT248																					
S08-PI6-G-AQM2-2635																					
T	0%	0.01%	0.01%	0.02%	0.01%	0.01%	0%	0%	0.02%	0.01%	0.01%	0%	0%	0.01%	0.01%	0.02%	0.01%	0%	0%	0%	0%
G	0%	0%	0%	0%	0%	0%	0%	0%	0.01%	0%	0%	0%	0%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0%
C	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
A	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0%	0%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0%
#	2009	2076	2057	2060	2039	2019	2009	2019	2009	2009	2019	2009	2009	2019	2009	2009	2019	2009	2009	2009	2009
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22

FIG. 59L

【図 5 9 L - 2】

PCR BT248																					
S08-PI6-G-AQM2-2635																					
T	0%	0%	0.01%	0.04%	0.02%	0.01%	0%	0%	0.02%	0.01%	0%	0%	0.01%	0.01%	0.02%	0%	0%	0%	0%	0.01%	0.01%
G	0.01%	0.01%	0.01%	0.01%	0.01%	0%	0%	0%	0.01%	0.01%	0%	0%	0.01%	0.01%	0.01%	0.01%	0.01%	0%	0%	0.01%	0.01%
C	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
A	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0%	0%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%
#	2015	2015	2015	2015	2015	2015	2015	2015	2015	2015	2015	2015	2015	2015	2015	2015	2015	2015	2015	2015	2015
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22

PCR BT248																					
S08-PI6-G-AQM2-2637																					
T	0%	0%	0%	0%	0%	0%	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%
C	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
A	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0%	0%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%
#	4035	4014	3935	3920	3962	3960	4040	4041	4024	3916	4005	3945	4014	4045	4024	4034	4074	4048	4065	4047	4049
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22

FIG. 59L CONTINUED

【図 6 0 A - 1】

PCR BT139																					
S01-AQM-G-PI6-AQM-PI6-BT139																					
T	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
G	0.03%	0%	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%
A	0.04%	0.04%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
#	165	173	174	100	167	107	165	170	165	174	170	165	169	169	173	165	172	173	170	174	173
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22

PCR BT139																					
S02-AQM-G-PI6-AQM-PI6-BT139																					
T	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
G	4.39%	0%	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%
A	0.01%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
#	12	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22

FIG. 60A

10

20

30

40

50

【図 60A - 2】

PCR-BTX139																							
S103-A0M-G2-P17-A0M-Pbe-H-PCR-BTX139																							
T	0.02%	0.13%	0.03%	0.03%	0.03%	0.02%	0.03%	0.01%	0.01%	0.01%	0.04%	0.02%	0.02%	0.01%	0.02%	0.06%	0.01%	0.03%	0.01%	0.03%	0.04%	0.01%	
G	0.48%	0%	0%	0%	0%	0%	0%	3.5%	0.1%	0%	0%	0%	0%	0%	0.02%	0.03%	0%	0.04%	0%	0%	0.05%	0%	
C	0.03%	0.03%	0.02%	0.02%	0.03%	0.08%	0.01%	0.03%	0.02%	0.01%	0.01%	0.01%	0.03%	0.03%	0.04%	0.11%	0%	0.01%	0%	0.02%	0%	0.03%	
A	0.45%	0.13%	0%	0%	0%	0%	0.02%	0.02%	0.02%	0.03%	0.08%	0.04%	0.01%	0.15%	0.04%	0.01%	0%	0.02%	0.01%	0.07%	0.01%	0%	
#	5365	5372	5379	5386	5393	5400	5407	5414	5421	5428	5435	5442	5449	5456	5463	5470	5477	5484	5491	5498	5505	5512	
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	

PCR-BTX139																							
S104-A0M-G2-P16-A0M-Pbe-I-PCR-BTX139																							
T	0.03%	0.01%	0.03%	0.04%	0.04%	0.02%	0.01%	0.02%	0.01%	0.02%	0.06%	0.02%	0.02%	0.01%	0.03%	0.03%	0.01%	0.05%	0.01%	0.03%	0.01%	0.01%	
G	0.35%	0.01%	0%	0%	0%	0%	2.02%	0.09%	0%	0.05%	0%	0%	0%	0%	0.01%	0.03%	0%	0.03%	0%	0%	0.03%	0%	
C	0.04%	0.03%	0.02%	0.01%	0.02%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.02%	0.02%	0.03%	0.12%	0.02%	0.03%	0%	0.01%	0%	0.01%	
A	0.58%	0.13%	0%	0%	0.02%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	
#	5246	5253	5260	5267	5274	5281	5288	5295	5302	5309	5316	5323	5330	5337	5344	5351	5358	5365	5372	5379	5386	5393	
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	

FIG. 60A CONTINUED

【図 60B - 1】

PCR-BTX139																							
S103-A0M-G2-P17-A0M-Pbe-H-PCR-BTX139																							
T	0.04%	0.02%	0.04%	0.03%	0.05%	0.03%	0.01%	0.04%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	
G	0.95%	0.01%	0%	0.45%	0%	33.52%	99.95%	0%	6.66%	0.07%	0.01%	99.91%	0%	0%	0.02%	99.91%	0%	0.02%	0%	0%	99.95%	0%	
C	0.03%	0.03%	0.02%	0.01%	99.92%	0%	0.01%	99.99%	0.02%	0.01%	0.01%	99.98%	0.04%	99.83%	0.4%	0.01%	0.01%	0.01%	0.01%	99.96%	0.04%	0%	
A	99.97%	0.1%	0.01%	47.51%	0.01%	60.47%	0.01%	0.01%	0.01%	0.01%	0.02%	0.01%	0%	0.16%	97.7%	0.02%	0%	99.77%	0.01%	0%	0.03%	0.02%	
#	6087	6094	6101	6108	6115	6122	6129	6136	6143	6150	6157	6164	6171	6178	6185	6192	6199	6206	6213	6220	6227	6234	
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	

PCR-BTX139																							
S103-A0M-G2-P17-A0M-Pbe-H-PCR-BTX139																							
T	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%	0%	33.33%	0%	24.44%	100%	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%	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	
A	100%	0%	0%	66.67%	0%	75.56%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	
#	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	

FIG. 60B

【図 60B - 2】

PCR-BTX139																							
S107-A0M-G2-P17-A0M-Pbe-H-PCR-BTX139																							
T	0.05%	0.05%	0.05%	0.05%	0.04%	0.04%	0.02%	0.01%	0.05%	0.05%	0.02%	0.02%	0.02%	0.02%	0.03%	0.04%	0.05%	0.01%	0.05%	0.05%	0.05%	0.05%	0.01%
G	1.19%	0.03%	0%	0.13%	0%	36.79%	95.95%	0%	5.36%	0.04%	0%	99.95%	0%	0.01%	0%	0.01%	0.01%	0%	0.03%	0%	0%	99.97%	0%
C	0.05%	0.05%	0.02%	0.01%	0.02%	0.01%	0.02%	0%	99.97%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.04%	0.02%	0.01%	0%	99.95%	0.02%	0%	99.93%
A	0.87%	0.12%	0%	48.02%	0.02%	55.17%	0.01%	0.02%	54.85%	0.01%	0.01%	0.1%	0.02%	0%	0.44%	99.81%	0.01%	0%	99.95%	0.01%	0%	0.04%	0.01%
#	7017	7024	7031	7038	7045	7052	7059	7066	7073	7080	7087	7094	7101	7108	7115	7122	7129	7136	7143	7150	7157	7164	7171
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	

PCR-BTX139																							
S108-A0M-G2-P17-A0M-Pbe-H-PCR-BTX139																							
T	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%	0%	33.33%	0%	8.33%	100%	0%	0%	0%	0%	100%	0%	0%	0%	100%	0%	0%	0%	0%	0%	100%	0%
C	0%	0%	0%	0%	0%	0%	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
A	100%	0%	0%	66.67%	0%	91.67%	0%	0%	100%	0%	0%	0%	0%	0%	0%	0%	0%	100%	0%	0%	0%	0%	0%
#	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23

FIG. 60B CONTINUED

【図 60C - 1】

PCR_BTX139																							
S103-A0M-G2-P17-A0M-Pbe-H-PCR-BTX139																							
T	0.02%	0.06%	0.02%	0.06%	0.04%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.04%	0.01%	0.01%	0.04%	0.03%	0.01%	0.02%	0.01%	0.02%	0.04%	0.01%	
G	0.55%	0%	0%	40.37%	0.01%	27.87%	0.92%	0%	2.88%	0.06%	0%	0.32%	0.01%	0%	0%	0.02%	0%	0.02%	0%	0%	0.35%	0%	
C	0.04%	0.03%	0.02%	0.03%	0.03%	0.02%	0.04%	0.04%	0%	0.04%	0.02%	0.01%	0.03%	0.03%	0.03%	0.04%	0.01%	0.01%	0.01%	0.02%	0%	0.35%	
A	0.93%	0.12%	0%	59.35%	0.02%	7.16%	0.02%	0.02%	0.02%	0.02%	0.01%	0.02%	0.01%	0.01%	0.15%	0.03%	0.03%	0.01%	0.01%	0.05%	0.02%		
#	7451	7452	7453	7454	7455	7456	7457	7458	7459	7460	7461	7462	7463	7464	7465	7466	7467	7468	7469	7470	7471		
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	

PCR_BTX139																							
S103-A0M-G2-P17-A0M-Pbe-H-PCR-BTX139																							
T	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	22%	0%	0%	0%	0%	0%	0%	
G	0%	0%	0%	25.67%	0%	18.52%	100%	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%	
A	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	
#	7472	7473	7474	7475	7476	7477	7478	7479	7480	7481	7482	7483	7484	7485	7486	7487	7488	7489	7490	7491	7492	7493	
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	

FIG. 60C

10

20

30

40

50

【図 60C - 2】

PCR B1x139																					
S117-ACM-G2-P117-ACM-PdBe-E2-PCR-B1x139																					
T	0.02%	0.13%	0.07%	0.03%	0.07%	0.09%	0.05%	0.07%	0.04%	0.02%	0.04%	0.01%	0.04%	0.02%	0.02%	0.04%	0.02%	0.04%	0.02%	0.02%	0.01%
G	0.08%	0%	0%	42.45%	0%	35.57%	0.02%	0%	3.73%	0.11%	0%	0.07%	0%	0.01%	0%	0.02%	0.04%	0%	0.02%	0%	0%
C	0.02%	0.07%	0.02%	0.02%	0.02%	0.08%	0.05%	0.07%	0.07%	0.01%	0.07%	0.01%	0.07%	0.01%	0.01%	0.03%	0.01%	0.01%	0.01%	0.01%	0%
A	0.04%	0.12%	0%	57.52%	0.02%	63.33%	0.04%	0.02%	0.04%	0.01%	0%	0.08%	0.02%	0.01%	0.02%	0.04%	0.01%	0%	0.09%	0.01%	0%
#	4537	4538	4539	4540	4541	4542	4543	4544	4545	4546	4547	4548	4549	4550	4551	4552	4553	4554	4555	4556	4557
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22

FIG. 60C CONTINUED

【図 60D - 1】

PCR B1x139																					
S117-ACM-G2-P117-ACM-PdBe-E2-PCR-B1x139																					
T	0.02%	0.05%	0.07%	0.03%	0.07%	0.09%	0.05%	0.07%	0.04%	0.02%	0.04%	0.01%	0.04%	0.02%	0.02%	0.04%	0.02%	0.04%	0.02%	0.02%	0.01%
G	0.08%	0.01%	0%	51.19%	0.01%	59.25%	0.01%	0%	6.31%	0.08%	0%	0.07%	0%	0.01%	0%	0.02%	0.04%	0%	0.02%	0%	0%
C	0.02%	0.07%	0.02%	0.02%	0.02%	0.08%	0.05%	0.07%	0.07%	0.01%	0.07%	0.01%	0.07%	0.01%	0.01%	0.03%	0.01%	0.01%	0.01%	0.01%	0%
A	0.04%	0.12%	0%	57.52%	0.02%	63.33%	0.04%	0.02%	0.04%	0.01%	0%	0.08%	0.02%	0.01%	0.02%	0.04%	0.01%	0%	0.09%	0.01%	0%
#	4558	4559	4560	4561	4562	4563	4564	4565	4566	4567	4568	4569	4570	4571	4572	4573	4574	4575	4576	4577	4578
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22

FIG. 60D

【図 60D - 2】

PCR B1x139																					
S116-ACM-G2-P116-ACM-PdBe-E2-PCR-B1x139																					
T	0.04%	0.07%	0.05%	0.05%	0.06%	0.02%	0.01%	0.04%	0.04%	0.02%	0.04%	0.01%	0.04%	0.02%	0.02%	0.04%	0.02%	0.04%	0.02%	0.02%	0.01%
G	0.08%	0.01%	0%	54.18%	0%	58.07%	0.02%	0%	5.67%	0.12%	0%	0.07%	0%	0.01%	0%	0.02%	0.04%	0%	0.02%	0%	0%
C	0.02%	0.07%	0.02%	0.02%	0.02%	0.08%	0.05%	0.07%	0.07%	0.01%	0.07%	0.01%	0.07%	0.01%	0.01%	0.03%	0.01%	0.01%	0.01%	0.01%	0%
A	0.04%	0.12%	0%	57.52%	0.02%	63.33%	0.04%	0.02%	0.04%	0.01%	0%	0.08%	0.02%	0.01%	0.02%	0.04%	0.01%	0%	0.09%	0.01%	0%
#	4579	4580	4581	4582	4583	4584	4585	4586	4587	4588	4589	4590	4591	4592	4593	4594	4595	4596	4597	4598	4599
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22

FIG. 60D CONTINUED

【図 60E - 1】

PCR B1x139																					
S117-ACM-G2-P117-ACM-PdBe-E2-PCR-B1x139																					
T	0.03%	0.08%	0.05%	0.02%	0.02%	0.02%	0.01%	0.04%	0.04%	0.02%	0.04%	0.01%	0.04%	0.02%	0.02%	0.04%	0.02%	0.04%	0.02%	0.02%	0.01%
G	0.05%	0.01%	0%	44.04%	0%	51.23%	0.01%	0%	2.53%	0.05%	0%	0.02%	0%	0.01%	0%	0.02%	0.04%	0%	0.02%	0%	0%
C	0.02%	0.07%	0.02%	0.02%	0.02%	0.08%	0.05%	0.07%	0.07%	0.01%	0.07%	0.01%	0.07%	0.01%	0.01%	0.03%	0.01%	0.01%	0.01%	0.01%	0%
A	0.04%	0.12%	0%	57.52%	0.02%	63.33%	0.04%	0.02%	0.04%	0.01%	0%	0.08%	0.02%	0.01%	0.02%	0.04%	0.01%	0%	0.09%	0.01%	0%
#	4600	4601	4602	4603	4604	4605	4606	4607	4608	4609	4610	4611	4612	4613	4614	4615	4616	4617	4618	4619	4620
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22

FIG. 60E

10

20

30

40

50

【図 60E - 2】

PCR-BTX39																							
S19-AQM-G2-P17-AQM-Pae1-C3-PCR-BTX39																							
T	0.01%	0.04%	99.96%	0.02%	0.07%	0.05%	0.03%	0.01%	0.02%	99.96%	99.99%	0.03%	0.02%	99.96%	0.01%	0.06%	0.04%	99.96%	0.01%	0.01%	99.96%	0.03%	0.01%
G	0.48%	0%	0%	46.29%	0%	39.13%	99.97%	0%	4.47%	0.11%	0%	99.97%	0.01%	0%	0.01%	0.03%	99.97%	0%	0.02%	0.02%	99.97%	0%	0%
C	0.03%	99.75%	0.01%	0.03%	99.97%	0.03%	0.01%	99.96%	0%	0.01%	0.02%	0.03%	99.95%	0.03%	99.97%	0.11%	0.02%	0.02%	0.01%	99.96%	0.01%	0%	99.96%
A	99.40%	0.16%	0%	93.67%	0.02%	90.79%	0.03%	0.01%	95.51%	0.01%	0%	0.12%	0.03%	0.01%	0.16%	99.75%	0.02%	96.95%	0.01%	0%	0.03%	0.01%	
#	2785	2785	2785	2785	2785	2785	2785	2785	2785	2785	2785	2785	2785	2785	2785	2785	2785	2785	2785	2785	2785	2785	

PCR-BTX39																							
S20-AQM-G2-P18-AQM-Pae1-C3-PCR-BTX39																							
T	0.01%	0.03%	99.97%	0.02%	0.04%	0.08%	0.02%	0.01%	0.02%	99.97%	99.99%	0.03%	0.01%	99.99%	0.01%	0.05%	98.99%	0.02%	0.01%	99.97%	0.03%	0.01%	
G	0.23%	0.02%	0%	99.97%	0%	26.89%	99.97%	0%	2.11%	0.1%	0%	99.97%	0%	0%	0.01%	0.04%	99.94%	0%	0.02%	0%	99.97%	0%	
C	0.05%	99.97%	0.03%	0.04%	99.93%	0.01%	0.01%	99.97%	0%	0.02%	0.02%	0.01%	99.97%	0.04%	98.97%	0.13%	0%	0.01%	99.96%	0.03%	0%	99.99%	
A	99.65%	0.14%	0%	90.65%	0.02%	73.02%	0.02%	97.07%	0.01%	0.03%	0.03%	0.02%	0.01%	0.17%	99.75%	0.01%	99.95%	0.03%	0%	0.1%	0%		
#	21602	21602	21602	21602	21602	21602	21602	21602	21602	21602	21602	21602	21602	21602	21602	21602	21602	21602	21602	21602	21602	21602	

FIG. 60E CONTINUED

【図 60F - 1】

PCR-BTX39																							
S21-AQM-G2-P19-AQM-Pae1-E3-PCR-BTX39																							
T	0.04%	0.06%	99.94%	0.02%	0.08%	0.07%	0.01%	0.04%	0.05%	99.95%	99.95%	0.05%	0.04%	99.94%	0.01%	0.05%	0.03%	99.94%	0.01%	0.02%	99.94%	0.03%	0.02%
G	0.93%	0.02%	0%	93.94%	0%	40.41%	99.97%	0%	7.44%	0.1%	0%	99.97%	0%	0%	0.01%	0.01%	99.93%	0%	0.09%	0%	99.94%	0%	
C	0.04%	99.97%	0.05%	0.02%	99.99%	0.01%	0.03%	99.94%	0.04%	0.03%	0.01%	0.01%	99.95%	0.04%	99.92%	0.08%	0%	0.01%	99.97%	0.03%	0%	99.97%	
A	99.99%	0.14%	0%	95.99%	0.02%	99.51%	0.03%	0.02%	92.04%	0.02%	0.03%	0.04%	0.01%	0%	0.05%	99.95%	0.03%	0%	99.94%	0.01%	0.04%	0.02%	
#	2162	2162	2162	2162	2162	2162	2162	2162	2162	2162	2162	2162	2162	2162	2162	2162	2162	2162	2162	2162	2162	2162	

PCR_BT139																							
S22-AQM-G2-P20-AQM-Pae1-F3-PCR_BT139																							
T	0.01%	0%	99.99%	0%	0.08%	0.03%	0.02%	0.01%	0.04%	0.04%	99.97%	0.03%	0.03%	99.97%	0%	0.04%	0.05%	99.94%	0%	0.03%	99.94%	0.04%	0%
G	0.02%	0%	0%	21.05%	0%	25.04%	99.97%	0%	2.65%	0.1%	0%	99.97%	0%	0%	0.01%	0.01%	99.94%	0%	0.03%	0.01%	99.92%	0%	
C	0.03%	99.99%	0.06%	0.01%	99.95%	0%	99.97%	0%	0.02%	0.02%	0.01%	99.95%	0.01%	99.95%	0.13%	0%	99.94%	0.01%	0%	99.92%	0.01%	99.99%	
A	99.97%	0.03%	0%	97.95%	0.02%	74.13%	0%	0.02%	97.22%	0.01%	0%	0.05%	0.02%	0.01%	0.05%	99.92%	0.01%	99.95%	0.04%	0.02%	0.04%	0.01%	
#	2169	2169	2169	2169	2169	2169	2169	2169	2169	2169	2169	2169	2169	2169	2169	2169	2169	2169	2169	2169	2169	2169	

FIG. 60F

【図 60F - 2】

PCR-BTX39																							
S23-AQM-G2-P21-AQM-Pae1-H3-PCR-BTX39																							
T	0.06%	0.1%	99.97%	0.03%	0.06%	0.08%	0.01%	0.01%	0.03%	99.97%	99.97%	0.03%	0.02%	99.97%	0%	0.04%	0.04%	99.97%	0.01%	0.05%	99.97%	0.05%	0.01%
G	0.37%	0%	45.99%	0%	55.07%	99.97%	0.01%	4.3%	0.09%	0%	99.92%	0.01%	0%	99.92%	0%	0.01%	99.93%	0%	0.02%	0%	99.98%	0%	99.98%
C	0.02%	99.77%	0.02%	0.03%	99.93%	0%	0.01%	99.97%	0.01%	0.03%	0.01%	99.97%	0.02%	99.94%	0.01%	0.01%	99.94%	0.02%	0%	99.98%	0%	99.98%	
A	99.95%	0.15%	0%	90.55%	0.02%	93.77%	0.03%	0.02%	95.05%	0.01%	0.03%	0.03%	0.01%	99.96%	0.01%	0.01%	99.96%	0.01%	0.07%	0.01%	0%	0%	
#	2740	2740	2740	2740	2740	2740	2740	2740	2740	2740	2740	2740	2740	2740	2740	2740	2740	2740	2740	2740	2740	2740	

PCR-BTX39																							
S24-AQM-G2-P22-AQM-Pae1-H3-PCR-BTX39																							
T	0%	0%	100%	2%	0%	0%	0%	0%	0%	100%	100%	0%	0%	100%	0%	0%	94.47%	0%	100%	100%	0%	0%	0%
G	0%	0%	21.40%	0%	15.38%	100%	0%	5.88%	0%	0%	100%	0%	0%	100%	5.88%	0%	100%	5.88%	0%	100%	0%	100%	0%
C	0%	100%	0%	100%	0%	0%	0%	0%	0%	100%	100%	0%	0%	100%	0%	0%	100%	0%	100%	0%	0%	100%	0%
A	100%	0%	75.5%	0%	84.62%	0%	84.62%	0%	0%	0%	0%	0%	0%	0%	100%	0%	100%	0%	100%	0%	0%	100%	0%
#	17	18	16	14	17	15	18	17	18	17	17	17	17	17	18	18	18	18	18	18	18	17	17

FIG. 60F CONTINUED

【図 60G - 1】

PCR-BTX39																							
S25-AQM-G2-P23-AQM-Pae1-H4-PCR-BTX39																							
T	0.01%	99.97%	0.03%	0.07%	0.07%	0.01%	0.02%	0.03%	99.97%	99.97%	0.03%	0.02%	99.99%	0.01%	0.03%	0.05%	99.97%	0.01%	0.01%	99.99%	0.03%	0.01%	
G	0.01%	0%	42.07%	0%	32.77%	99.95%	0%	4.57%	0.09%	0%	99.95%	0%	0.01%	0%	0.03%	99.93%	0%	0.02%	0%	99.94%	0%	0%	
C	99.93%	0.02%	0.04%	99.91%	0%	0%	99.95%	0.01%	0.01%	0.02%	0%	99.97%	0.01%	0.03%	0.01%	0.02%	0.01%	99.97%	0.02%	0%	99.97%	0%	
A	99.99%	0.15%	0%	92.85%	0.02%	97.7%	0.02%	95.29%	0.01%	0%	0.02%	0.01%	0%	0.45%	99.81%	0.01%	0.01%	99.99%	0.02%	0%	99.97%	0%	
#	6884	6884	6884	6884	6884	6884	6884	6884	6884	6884	6884	6884	6884	6884	6884	6884	6884	6884	6884	6884	6884	6884	

PCR-BTX39																							
S26-AQM-G2-P24-AQM-Pae1-H4-PCR-BTX39																							
T	0.05%	0.04%	99.99%	0.04%	0.05%	0.02%	0.06%	0.05%	99.99%	99.97%	0.04%	0.01%	99.99%	0.01%	0.05%	0.05%	99.96%	0.01%	0.03%	99.96%	0.03%	0%	
G	0.98%	0%	61.44%	0%	47.3%	99.93%	0%	5.58%	0.08%	0.01%	99.99%	0%	0%	0.01%	99.9%	0%	0.02%	0%	99.99%	0%	99.99%	0%	
C	0.03%	99.96%	0.02%	0.02%	99.92%	0.03%	0%	99.94%	0.02%	0.02%	99.98%	0.02%	99.98%	0.14%	0.02%	0.02%	0.01%	99.99%	0.02%	0%	99.99%	0%	
A	99.93%	0.16%	0%	93.9%	0.02%	92.42%	0.03%	0%	94.39%	0%	0%	0.04%	0.03%	0%	99.9%	0.01%	99.98%	0%	0%	99.97%	0.01%	99.97%	
#	4405	4404	4507	4550	4490	4363	4484	4501	4488	4493	4493	4493	4493	4493	4493	4493	4493	4493	4493	4493	4493	4493	

FIG. 60G

【図 60G - 2】

PCR 5Tx139																					
S02-4CM-G2-PV2-4CM-PH-E4-PCR-5Tx139																					
T	0.04%	0.05%	0.08%	0.08%	0.05%	0.08%	0.02%	0.02%	0.03%	0.05%	0.05%	0.05%	0.02%	0.03%	0.04%	0.04%	0.04%	0.04%	0.05%	0.01%	0.01%
G	0.04%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
C	0.03%	0.03%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%
A	0.03%	0.03%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%
#	5020	5021	5022	5023	5024	5025	5026	5027	5028	5029	5030	5031	5032	5033	5034	5035	5036	5037	5038	5039	5040
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22

FIG. 60G CONTINUED

【図 60H - 1】

PCR 5Tx139																					
S02-4CM-G2-PV2-4CM-PH-E4-PCR-5Tx139																					
T	0.04%	0.05%	0.08%	0.08%	0.05%	0.08%	0.02%	0.02%	0.03%	0.05%	0.05%	0.05%	0.02%	0.03%	0.04%	0.04%	0.04%	0.04%	0.05%	0.01%	0.01%
G	0.04%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
C	0.03%	0.03%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%
A	0.03%	0.03%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%
#	5020	5021	5022	5023	5024	5025	5026	5027	5028	5029	5030	5031	5032	5033	5034	5035	5036	5037	5038	5039	5040
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22

FIG. 60H

【図 60H - 2】

PCR 5Tx139																					
S03-4CM-G2-PV2-4CM-PH-E4-PCR-5Tx139																					
T	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%
G	0.02%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
C	0.03%	0.03%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%
A	0.04%	0.04%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%
#	5041	5042	5043	5044	5045	5046	5047	5048	5049	5050	5051	5052	5053	5054	5055	5056	5057	5058	5059	5060	5061
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22

FIG. 60H CONTINUED

【図 60I - 1】

PCR 5Tx139																					
S03-4CM-G2-PV2-4CM-PH-E4-PCR-5Tx139																					
T	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%
G	0.05%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
C	0.04%	0.04%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%
A	0.05%	0.05%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%
#	5062	5063	5064	5065	5066	5067	5068	5069	5070	5071	5072	5073	5074	5075	5076	5077	5078	5079	5080	5081	5082
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22

FIG. 60I

10

20

30

40

50

【図 60 I - 2】

PCR BTx139																							
S303-ACM-G2-PV25-ACM-Phe-I-3S-PCR-BTx139																							
T	0.03%	0.05%	59.95%	0.02%	0.08%	0.05%	0.02%	0.02%	0.05%	59.95%	59.95%	0.05%	0.03%	59.97%	0.01%	0.05%	0.02%	59.99%	0%	0.02%	59.99%	0.01%	0%
G	0.94%	0.01%	56.33%	0.01%	55.91%	59.97%	0.01%	4.39%	0.08%	0%	59.97%	0%	0%	0%	0%	0%	0.02%	59.91%	0%	0.05%	0%	59.97%	0%
C	0.05%	59.97%	0.02%	0.02%	59.98%	0.02%	0.01%	59.95%	0.01%	0.01%	0.02%	0.01%	59.94%	0.00%	0.02%	0.01%	0.02%	59.95%	0.01%	0%	59.95%	0.01%	0%
A	59.97%	0.14%	0%	51.53%	0.03%	51.01%	0.03%	0.01%	55.55%	0.03%	0.01%	0.03%	0%	0.55%	59.97%	0.02%	0.01%	59.93%	0.02%	0%	0.03%	0.01%	0%
#	47911	48023	48166	47444	47476	47635	47683	47687	47696	47705	47706	47759	47784	47811	48084	47320	47644	47622	48025	47593	48000	48467	47903
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	

PCR BTx139																							
S303-ACM-G2-PV25-ACM-Phe-I-3S-PCR-BTx139																							
T	0.01%	0.01%	59.99%	0.02%	0.04%	0.05%	0.01%	0.02%	0.00%	59.97%	59.95%	0.03%	0.03%	59.95%	0.01%	0.04%	0.04%	59.97%	0.01%	0.03%	59.99%	0.04%	0.01%
G	0.46%	0.01%	56.44%	0.01%	52.21%	59.94%	0%	4.41%	0.07%	0.01%	59.94%	0%	0.01%	0%	0%	0%	0.04%	59.93%	0%	0.01%	0%	59.97%	0%
C	0.02%	59.98%	0.02%	0.02%	59.94%	0.07%	0.01%	59.95%	0.02%	0.01%	59.95%	0.00%	59.91%	0.12%	0.02%	0.03%	0.04%	59.95%	0.02%	0%	59.95%	0.02%	0%
A	59.97%	0.12%	0%	51.51%	0.02%	51.55%	0.02%	0.03%	55.22%	0%	0%	0.02%	0.03%	0.01%	0.48%	59.93%	0.01%	0%	59.97%	0.01%	0%	0.04%	0%
#	45394	45394	45394	45394	45394	45394	45394	45394	45394	45394	45394	45394	45394	45394	45394	45394	45394	45394	45394	45394	45394	45394	
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	

FIG. 60I CONTINUED

【図 60 J - 1】

PCR BTx139																							
S307-ACM-G2-PV27-ACM-Phe-I-5S-PCR-BTx139																							
T	0.04%	0.01%	59.99%	0.03%	0.08%	0.09%	0.03%	0.01%	0.04%	59.96%	59.95%	0.04%	0.01%	59.97%	0%	0.04%	0.04%	59.99%	0.02%	0.05%	59.95%	0.02%	0%
G	1%	0.01%	0%	59.07%	0.95%	49.23%	59.97%	0%	4.52%	0.11%	0%	59.95%	0%	0%	0%	0.01%	0.02%	59.93%	0%	0.01%	59.93%	0.01%	0%
C	0.03%	59.98%	0.01%	0.05%	59.97%	0%	0%	59.98%	0.01%	0.01%	0.02%	0.03%	59.95%	0.03%	59.95%	0.01%	0.02%	59.94%	0.03%	0%	59.95%	0.03%	0%
A	59.97%	0.15%	0%	51.94%	0.08%	51.67%	0.04%	0.01%	55.43%	0.03%	0.01%	0.08%	0%	0%	0.17%	59.93%	0.01%	0%	59.95%	0%	0%	0.05%	0.01%
#	47495	47694	47695	48015	47447	48401	47322	47909	48674	47752	47989	48986	47729	48619	47946	47614	48095	47725	47446	47989	47939	47917	
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	

PCR BTx139																							
S308-ACM-G2-PV28-ACM-Phe-I-5S-PCR-BTx139																							
T	0.05%	0.04%	59.93%	0.04%	0.08%	0.07%	0.04%	0.06%	0.00%	59.97%	59.97%	0.03%	0.01%	59.99%	0.01%	0.05%	0.06%	59.97%	0.01%	0.04%	59.99%	0.03%	0.01%
G	1.35%	0.01%	0%	52.35%	0.91%	52.21%	59.91%	0%	11.8%	0.08%	0.01%	59.94%	0%	0%	0%	0%	0.02%	59.92%	0%	0.06%	0%	59.91%	0%
C	0.03%	59.97%	0.04%	0.04%	59.98%	0.03%	0.01%	59.99%	0.07%	0.01%	0.02%	0.01%	59.99%	0.03%	59.91%	0.12%	0%	0.02%	0%	59.95%	0.02%	0%	
A	59.93%	0.16%	0%	51.55%	0.05%	51.67%	0.02%	0.02%	58.05%	0.02%	0%	0.03%	0.01%	0.15%	59.91%	0.01%	0.01%	59.93%	0.01%	0%	0.08%	0.02%	
#	48531	48530	48744	47739	48507	48531	48531	48531	48531	48531	48531	48531	48531	48531	48531	48531	48531	48531	48531	48531	48531	48531	
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	

FIG. 60J

【図 60 J - 2】

PCR BTx139																							
S308-ACM-G2-PV28-ACM-Phe-I-5S-PCR-BTx139																							
T	0.01%	0.05%	59.97%	0.01%	0.03%	0.04%	0.01%	0.01%	0.02%	59.98%	59.97%	0.04%	0.02%	59.97%	0.01%	0.04%	0.05%	59.97%	0%	0.03%	59.99%	0.04%	0.01%
G	0.02%	0%	5%	0.53%	0%	0.66%	59.97%	0%	0.12%	0.11%	0%	59.95%	0%	0%	0%	0.01%	59.94%	0%	0%	59.91%	0%	59.91%	
C	0.03%	59.97%	0.02%	0.01%	59.96%	0%	0%	59.97%	0%	0.01%	0.02%	0.02%	59.98%	0.03%	59.94%	0.12%	0%	0.02%	0%	59.95%	0.02%	0%	
A	59.97%	0.14%	0%	59.95%	0.01%	59.94%	0.01%	0.02%	59.97%	0%	0.05%	59.93%	0.01%	0%	59.98%	0.02%	0%	59.98%	0.02%	0%	0.05%	0%	
#	55982	55982	56106	56106	55904	56177	55790	56295	56295	56194	55797	55445	55795	56165	56165	56110	56542	56172	55789	56442	56459	55759	
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	

PCR BTx139																							
S304-ACM-G2-Phe-I-ACM-Phe-I-5S-PCR-BTx139																							
T	0%	0%	100%	0%	0%	0%	0%	0%	0%	100%	100%	0%	0%	100%	0%	0%	0%	100%	0%	0	100%	0%	0%
G	0%	0%	0%	19.8%	0%	10%	100%	0%	0%	0%	0%	100%	0%	0%	100%	0%	0%	0%	0%	0%	100%	0%	0%
C	0%	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	100%	0%	0%	0%	0%	0%	0%	0%	0%	100%	0%
A	100%	0%	0%	81.82%	0%	90%	0%	0%	100%	0%	0%	0%	0%	0%	100%	0%	0%	0%	0%	0%	0%	0%	0%
#	28	29	29	22	30	20	29	28	27	28	29	29	29	29	26	30	30	30	29	29	30	30	30
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	

FIG. 60J CONTINUED

【図 60 K - 1】

PCR BTx139																							
S304-ACM-G2-PV28-ACM-Phe-I-5S-PCR-BTx139																							
T	0.04%	0.26%	59.95%	0.02%	0.08%	0.09%	0.02%	0.01%	0.04%	59.97%	59.97%	0.05%	0.01%	59.94%	0.01%	0.04%	0.05%	59.99%	0.01%	0.04%	59.95%	0.06%	0.01%
G	0.95%	0%	43.67%	0%	23.27%	59.95%	0%	4.08%	0.12%	0%	59.95%	0%	0%	5%	0%	0.03%	59.93%	0%	0%	5%	59.93%	0%	0%
C	0.04%	59.99%	0.02%	0.01%	59.9%	0%	0%	59.98%	0.04%	0.01%	0.02%	0.02%	59.97%	0.03%	59.97%	0.15%	0%	0.01%	59.94%	0.03%	0%	59.98%	0%
A	59.9%	0.14%	0%	55.55%	0.01%	57.07%	0.02%	0.01%	55.83%	0%	0.01%	0.04%	0.01%	0%	0.2%	59.97%	0.01%	0%	59.98%	0.01%	0.01%	0.02%	0.01%
#	71801	71775	71901	68871	71501	69255	71491	72171	70951	72201	72031	71051	71191	71005	71911	70419	72057	72501	72021	71591	72401	72312	
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	

PCR BTx139																							
S302-ACM-G2-PV24-ACM-Phe-I-5S-PCR-BTx139																							
T	0.03%	0.08%	59.94%	0.04%	0.08%	0.03%	0.02%	0.02%	0.03%	59.99%	59.97%	0.04%	0.01%	59.94%	0.02%	0.04%	0.04%	59.97%	0%	0.03%	59.99%	0.03%	0.01%
G	0.84%	0%	51.8%	0%	40.85%	59.94%	0%	4.29%	0.05%	0.01%	59.95%	0%	0%	5%	0%	0.02%	59.94%	0%	0.02%	0%	59.91%	0%	0%
C	0.04%	59.95%	0.04%	0.02%	59.94%	0.01%	0%	59.94%	0%	59.94%	0%	0.02%	0.02%	59.98%	0.05%	59.94%	0.13%	0.01%	0.02%	0%	59.95%	0.03%	0%
A	59.07%	0.13%	0%	48.12%	0.02%	59.04%	0.03%	0.03%	55.66%	0.01%	0%	0.01%	0.01%	0%	0.4%	59.81%	0.02%	0%	59.98%	0.01%	0%	0.05%	0.02%
#	72471	45543	45645	44295	45351	45789	45271	45752	45013	45743	44980	45067	45208	45046	45593	44467	45955	46201	45955	45440	45989	45817	45907
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	

【図 60K-2】

PCR-BT139																							
SNA-AON-G2-P76-AONPae-EG-PCR-BT139																							
T	0.03%	0.1%	0.03%	0.04%	0.07%	0.08%	0.04%	0.02%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%
G	0.95%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
C	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%
A	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%
#	5845	4402	4402	4402	4402	4402	4402	4402	4402	4402	4402	4402	4402	4402	4402	4402	4402	4402	4402	4402	4402	4402	4402
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	

PCR-BT139																							
SNA-AON-G2-P76-AONPae-EG-PCR-BT139																							
T	0.02%	0.06%	0.02%	0.07%	0.04%	0.02%	0.04%	0.02%	0.01%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%
G	0.43%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
C	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%
A	0.51%	0.1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
#	5232	5336	5348	5367	5367	5370	5388	5371	5387	5385	5385	5385	5385	5385	5385	5385	5385	5385	5385	5385	5385	5385	5385
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	

FIG. 60K CONTINUED

【図 60L-1】

PCR-BT139																							
SNA-AON-G2-P76-AONPae-EG-PCR-BT139																							
T	0.03%	0.05%	0.03%	0.07%	0.03%	0.04%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%
G	1.1%	0.03%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
C	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%
A	0.04%	0.1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
#	4785	4884	4877	4877	4877	4877	4885	4875	4885	4885	4885	4885	4885	4885	4885	4885	4885	4885	4885	4885	4885	4885	4885
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	

PCR-BT139																							
SNA-AON-G2-P76-AONPae-FG-PCR-BT139																							
T	0.08%	0.02%	0.03%	0.02%	0.02%	0.04%	0.04%	0.01%	0.04%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%
G	1.46%	0.01%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
C	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%
A	0.41%	0.1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
#	4438	4457	4476	4476	4476	4476	4476	4476	4476	4476	4476	4476	4476	4476	4476	4476	4476	4476	4476	4476	4476	4476	4476
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	

FIG. 60L

【図 60L-2】

PCR-BT139																							
SNA-AON-G2-P76-AONPae-FG-PCR-BT139																							
T	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%	0%	0%	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%	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%
#	112	112	112	112	112	112	112	112	112	112	112	112	112	112	112	112	112	112	112	112	112	112	112
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	

PCR-BT139																							
SNA-AON-G2-P76-AONPae-FG-PCR-BT139																							
T	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%	0%	0%	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%	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%
#	11	10	11	7	11	9	10	10	11	11	10	11	10	11	11	10	11	11	10	11	10	11	11
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	

FIG. 60L CONTINUED

【図 60M-1】

PCR-BT139																							
S302-NMP-ae-002-PCR-BT139																							
T	99.99%	0.04%	0.01%	99.99%	0.03%	0.34%	0.1%	99.97%	0.01%	0.04%	0.02%	0.03%	0.01%	0.02%	99.99%	99.97%	0.04%	0.01%	99.99%	0.01%	0.01%	0.04%	99.99%
G	0%	99.99%	0%	0.05%	99.97%	0.03%	0.08%	0%	12.67%	0%	0.01%	99.99%	0%	0.01%	0.08%	0.01%	99.99%	0%	0%	0.02%	99.95%	0%	
C	0%	0.02%	99.99%	0.02%	0.01%	0.03%	99.97%	0.02%	0.02%	99.97%	0%	99.97%	0%	0.05%	0.02%	0.02%	99.99%	0.01%	99.99%	0.04%	0%	0%	
A	0%	0.02%	0.19%	0%	0%	99.99%	0.2%	0%	99.99%	0.01%	0.02%	99.97%	0.01%	0%	0.02%	0.01%	0%	0.15%	99.99%	0.01%	0%	0%	
#	6071	5976	5964	5976	5967	5949	5976	5976	5976	5976	5976	5976	5976	5976	5976	5976	5976	5976	5976	5976	5976	5976	
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	

PCR-BT139																							
S302-NMP-ae-002-PCR-BT139																							
T	99.99%	0.05%	0.01%	99.99%	0.1%	0.39%	0.28%	99.97%	0.02%	0.01%	0.05%	0.01%	0%	0.01%	99.99%	99.99%	0.03%	0.01%	99.99%	0.01%	0.02%	0.04%	99.99%
G	0%	99.94%	0%	0.05%	99.99%	0.15%	0%	0.01%	13.83%	0%	0.31%	99.99%	0%	0.01%	0.08%	0.01%	99.97%	0%	0%	0.01%	99.94%	0%	
C	0.01%	0%	99.97%	0.01%	0%	0.03%	99.99%	0.02%	0.01%	99.97%	0%	0%	0%	0.02%	0.01%	0%	99.97%	0.01%	99.99%	0.04%	0%	0.01%	
A	0%	0.01%	0.08%	0.04%	0%	99.99%	0.1%	0%	99.94%	0%	0.02%	99.99%	0.01%	0%	0%	0.02%	0.01%	0.16%	99.99%	0.02%	0%	0%	
#	5137	5041	5055	5102	5055	4985	5037	5017	4957	4957	4957	4955	5041	5042	5033	5017	4955	5004	4965	5065	5044	5065	5065

【図 6 0 0 - 2】

PCR BTx139																						
S312-NP-neo3-002-PCR-BT139																						
T	99.95%	0.04%	0.01%	99.95%	0.32%	0.67%	0%	99.97%	0.02%	0%	0.04%	0.01%	0.01%	0.02%	99.94%	99.97%	0.05%	0.02%	99.97%	0.05%	0.03%	99.98%
G	0%	99.95%	0%	0.05%	99.99%	0%	0.04%	0%	0.04%	0%	0.04%	99.99%	0%	0.03%	0.13%	0.01%	99.97%	0%	0%	0.01%	0.01%	99.94%
C	0.02%	0%	99.93%	0.01%	0%	0.03%	99.97%	0%	0%	99.97%	0%	0%	99.97%	0%	0.03%	0.01%	0.01%	99.97%	0.02%	99.97%	0.15%	0%
A	0%	0%	0.05%	0%	0%	99.9%	0.04%	0%	78.41%	0.02%	99.97%	0.01%	0.02%	99.95%	0%	0.01%	0.01%	0.01%	0.01%	0.25%	99.91%	0%
#	5399/	5302/	5294/	5294/	5247/	5111/	5204/	5216/	5107/	5265/	5081/	5205/	5245/	5192/	5297/	5294/	5198/	5170/	5162/	5290/	5188/	52915/

PCR BTx139																						
S313-NP-neo3-002-PCR-BT139																						
T	100%	0%	0%	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	100%	100%	0%	0%	100%	0%	0%	100%
G	0%	100%	0%	0%	100%	5.58%	0%	0%	0%	0%	100%	0%	0%	0%	0%	0%	0%	0%	100%	0%	0%	100%
C	0%	0%	100%	0%	0%	0%	100%	0%	0%	100%	0%	0%	100%	0%	0%	0%	0%	0%	100%	0%	100%	0%
A	0%	0%	0%	0%	0%	94.44%	0%	0%	100%	0%	100%	0%	100%	0%	0%	0%	0%	0%	0%	0%	100%	0%
#	23	22	20	23	19	16	21	18	19	24	23	20	23	22	23	20	22	22	21	20	20	23

FIG. 600 CONTINUED

【図 6 0 P - 1】

PCR BTx139																						
S314-NP-neo3-001-PCR-BT139																						
T	99.98%	0.05%	0%	99.95%	0.23%	0.44%	0.05%	99.97%	0.03%	0.01%	0.04%	0.02%	0.01%	0.01%	99.97%	99.99%	0.03%	0.01%	99.99%	0.01%	0.04%	99.98%
G	0%	99.94%	0%	0.06%	99.7%	11.22%	0.05%	0%	14.53%	0%	0.23%	99.97%	0.01%	0.01%	0%	99.99%	0%	0%	0%	0%	0.01%	99.94%
C	0.02%	0.01%	99.97%	0.01%	0%	0.03%	99.76%	0.02%	0.01%	99.97%	0%	0%	99.95%	0%	0.01%	0.01%	0.03%	99.95%	0.01%	99.95%	0.03%	0%
A	0%	0%	0.05%	0%	0%	99.25%	0.13%	0.01%	15.43%	0.01%	99.72%	0.01%	0.02%	99.97%	0%	0.04%	0.01%	0%	0.16%	99.92%	0.01%	0%
#	5522/	5457/	5409/	5409/	5307/	5370/	5434/	5447/	5330/	5430/	5307/	5478/	5173/	5392/	5474/	5482/	5385/	5408/	5344/	5457/	5392/	5505/

PCR BTx139																						
S315-NP-neo3-002-PCR-BT139																						
T	99.98%	0.04%	0%	99.9%	0.38%	0.5%	0.01%	99.98%	0.02%	0.01%	0.05%	0.02%	0.01%	0.03%	99.95%	99.99%	0.04%	0.01%	99.99%	0%	0.03%	99.98%
G	0%	99.95%	0%	0.07%	99.64%	15.64%	0.14%	0%	16.53%	0%	0.19%	99.97%	0%	0.01%	0.12%	0%	99.94%	0%	0%	0%	0.01%	99.98%
C	0.02%	0%	99.97%	0.03%	0%	0.02%	99.73%	0.02%	0.02%	99.98%	0.01%	0%	99.95%	0%	0.01%	0.03%	99.95%	0.01%	99.95%	0.01%	0.03%	0%
A	0%	0.01%	0.03%	0%	0%	99.84%	0.14%	0%	13.41%	0.01%	99.73%	0%	0.01%	99.95%	0%	0.01%	0.03%	0%	0.19%	99.94%	0.01%	0%
#	4733/	4629/	4619/	4706/	4639/	4055/	4024/	4074/	4090/	4578/	4646/	4950/	4447/	4497/	4024/	4036/	4090/	4090/	4036/	4090/	4036/	4134/
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22

FIG. 60P

【図 6 0 P - 2】

PCR BTx139																						
S316-NP-neo3-003-PCR-BT139																						
T	99.9%	0.05%	0.01%	99.9%	0.41%	0.49%	0%	99.97%	0.02%	0%	0.04%	0.01%	0%	0.02%	99.98%	99.99%	0.03%	0.02%	100%	0%	0.03%	99.9%
G	0%	99.94%	0%	0.06%	99.58%	17.1%	0%	0%	17.92%	0%	0.57%	99.99%	0%	0.03%	0.1%	0%	99.99%	0%	0%	0.01%	99.94%	0%
C	0.01%	0%	99.94%	0%	0.01%	99.94%	0.02%	0.01%	99.97%	0%	0%	99.97%	0%	0.01%	0.01%	0%	99.97%	0%	99.97%	0.17%	0%	0%
A	0.01%	0%	0.05%	0%	92.23%	0.16%	0%	92.05%	0.02%	99.9%	0%	0.03%	99.95%	0.01%	0%	0.01%	0.01%	0%	0.4%	99.76%	0.01%	0%
#	2517/	2557/	2548/	2317/	2317/	2247/	2257/	2262/	2295/	2355/	2295/	2350/	2371/	2347/	2373/	2347/	2349/	2349/	2349/	2349/	2349/	2349/

PCR BTx139																						
S317-NP-neo3-001-PCR-BT139																						
T	99.9%	0.05%	0.01%	99.92%	0.4%	0.01%	0.01%	99.99%	0.01%	0.01%	0.02%	0.01%	0.01%	0.03%	99.97%	99.99%	0.03%	0.01%	99.99%	0.01%	0.03%	99.9%
G	0%	99.94%	0%	0.06%	99.65%	0.02%	0%	0%	0.05%	0%	0.18%	99.99%	0.01%	0.01%	0.11%	0%	99.99%	0%	0%	0.02%	99.95%	0%
C	0.01%	0%	99.96%	0.02%	0%	0.03%	99.95%	0.01%	0.01%	99.99%	0%	0%	99.95%	0%	0.01%	0.01%	99.97%	0.02%	99.94%	0.13%	0.01%	0%
A	0.01%	0%	0.03%	0%	99.94%	0.14%	0%	99.94%	0.01%	99.9%	0.01%	0.03%	99.97%	0.01%	0.01%	0%	0.01%	0%	0.15%	99.82%	0.01%	0%
#	5559/	5451/	5462/	5530/	5576/	5438/	5479/	5522/	5403/	5473/	5350/	5478/	5549/	5537/	5501/	5461/	5439/	5404/	5407/	5389/	5490/	5559/

FIG. 60P CONTINUED

【図 6 0 Q - 1】

PCR_BT139																						
S318-NP-neo1-002-PCR-BT139																						
T	99.99%	0.00%	0%	99.94%	0.22%	0.01%	0%	99.99%	0.01%	0.05%	0.02%	0.01%	0.03%	99.98%	99.99%	0.03%	0.01%	99.99%	0%	0.03%	0.04%	99.98%
G	0%	99.97%	0%	0.05%	99.67%	0.02%	0%	0%	0.03%	0%	0.11%	99.97%	0%	0.01%	0.08%	0%	99.99%	0%	0%	0%	99.94%	0%
C	0.01%	0%	99.93%	0.01%	0%	0.02%	99.95%	0.01%	0.02%	99.99%	0%	0%	99.98%	0%	0.01%	0.01%	99.97%	0.01%	99.94%	0.13%	0%	0.01%
A	0.05%	0.02%	0.04%	0%	99.95%	0.14%	0%	99.94%	0.01%	99.94%	0.01%	0.01%	99.95%	0.01%	0%	0.02%	0%	0.16%	99.83%	0.01%	0%	
#	5829/	5577/	5574/	5942/	5835/	5574/	5557/	5572/	5573/	5495/	5579/	5525/	5544/	5524/	5519/	5549/	5534/	5501/	5508/	5613/	5509/	
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23

PCR_BT139																						
S319-NP-neo3-003-PCR-BT139																						
T	99.99%	0.04%	0.05%	99.92%	0.35%	0.01%	0.1%	99.99%	0%	0.01%	0.02%	0.02%	0.01%	0.03%	99.97%	99.99%	0.03%	0.01%	99.99%	0%	0.04%	99.99%
G	0%	99.95%	0%	0.06%	99.65%	0.02%	0%	0%	0.03%	0%	0.05%	99.99%	0%	0.01%	0.12%	0%	99.99%	0%	0%	0.01%	99.95%	0%
C	0.01%	0%	99.97%	0.02%	0%	0.03%	99.75%	0.02%	0.02%	99.97%	0%	0%	99.97%	0%	0.01%	0.01%	99.97%	0.01%	99.98%	0.14%	0%	0.01%
A	0%	0.01%	0.03%	0%	99.94%	0.12%	0%	99.95%	0.02%	99.95%	0%	0.02%	99.95%	0.01%	0%	0%	0%	0.13%	99.81%	0.02%	0%	0%
#	5245/	5125/	5147/	5174/	5165/	5174/	5132/	5157/	5080/	5126/	5203/	5130/	5164/	5087/	5162/	5169/	5089/	5086/	5145/	5145/	5171/	

FIG. 60Q

10

20

30

40

50

【図 60Q-2】

PCR-BTX139																							
S32-NP-eval-001-PCR-BTX139																							
T	99.9%	0.05%	0.01%	99.9%	0.3%	0%	0.02%	99.9%	0.01%	0%	0.05%	0.01%	0.01%	0.03%	99.9%	0.05%	0.03%	99.9%	0.01%	0.02%	0.05%	99.9%	0%
G	0%	99.9%	0%	0.04%	99.9%	0.01%	0%	0%	0.02%	0%	0.01%	99.9%	0%	0.01%	99.9%	0%	0.01%	99.9%	0%	0%	0.01%	99.9%	0%
C	0.02%	0.01%	99.9%	0.02%	0%	0.03%	99.9%	0.02%	0.01%	99.9%	0%	0%	99.9%	0%	99.9%	0.02%	0.03%	99.9%	0.02%	99.9%	0.01%	0.01%	0.02%
A	0.8%	0.01%	0.04%	0%	0.01%	99.9%	0.1%	0%	99.9%	0.01%	99.9%	0.01%	0.02%	99.9%	0%	0%	0.02%	0.01%	0%	0.16%	99.9%	0.01%	0%
#	5499/	4855/	4950/	4916/	4855/	4845/	4815/	4883/	4860/	4861/	4734/	4853/	4837/	4851/	4810/	4833/	4823/	4810/	4793/	4801/	4801/	4811/	4841/
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	

PCR-BTX139																							
S32-NP-eval-002-PCR-BTX139																							
T	99.9%	0.04%	0.02%	99.9%	0.02%	0.02%	0.01%	99.9%	0.01%	0%	0.05%	0.02%	0.02%	0.02%	99.9%	0.04%	0.01%	99.9%	0.01%	0.05%	0.02%	99.9%	0%
G	0%	99.9%	0%	0.05%	99.9%	0.02%	0%	0%	0.01%	99.9%	0%	0.01%	99.9%	0%	0.08%	0%	99.9%	0%	0%	0.02%	99.9%	0%	0%
C	0.02%	0.01%	99.9%	0.01%	0%	0.03%	99.9%	0.01%	0.02%	99.9%	0%	99.9%	0%	0.01%	0.02%	0.02%	99.9%	0.02%	99.9%	0.01%	0%	0.01%	0.01%
A	0%	0%	0.03%	0%	0%	99.9%	0.1%	0%	99.9%	0.02%	99.9%	0%	0.02%	99.9%	0%	0.01%	0%	0.07%	0.02%	0%	0.05%	99.9%	0%
#	5706/	5654/	5686/	5741/	5723/	5635/	5714/	5613/	5631/	5591/	5591/	5723/	5554/	5713/	5760/	5533/	5532/	5544/	5591/	5599/	5718/	5776/	0%
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	

FIG. 60Q CONTINUED

【図 60R-1】

PCR-BTX139																							
S32-NP-eval-003-PCR-BTX139																							
T	100%	0%	0%	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	100%	100%	0%	0%	100%	0%	0%	0%	100%
G	0%	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
C	0%	0%	100%	0%	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%
#	32	29	32	32	32	33	33	33	27	31	29	32	32	32	32	32	32	32	32	32	32	31	33
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	

PCR-BTX139																							
S32-NP-eval-001-PCR-BTX139																							
T	99.9%	0.06%	0.01%	99.9%	0.2%	0.01%	0%	99.9%	0.01%	0%	0.04%	0.05%	0.01%	0.03%	99.9%	0.05%	0.01%	99.9%	0.01%	0.04%	0.01%	99.9%	0%
G	0%	99.9%	0%	0.04%	99.9%	0.03%	0%	0%	0.04%	0%	0.03%	99.9%	0%	0.01%	99.9%	0%	0.01%	99.9%	0%	0%	0.01%	99.9%	0%
C	0.01%	0%	99.9%	0.01%	0%	0.04%	99.9%	0.01%	0.01%	99.9%	0%	0%	99.9%	0%	99.9%	0.01%	0.01%	99.9%	0.01%	0.01%	0.01%	99.9%	0%
A	0%	0.01%	0.04%	0%	0.01%	99.9%	0.1%	0%	99.9%	0.01%	99.9%	0%	0.02%	99.9%	0%	0.02%	0.03%	0%	0.44%	99.9%	0.01%	0.01%	0%
#	5834/	4756/	4782/	4855/	4794/	4738/	4765/	4770/	4745/	4659/	4736/	4784/	4705/	4749/	4765/	4735/	4732/	4751/	4635/	4772/	4633/	4633/	4633/
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	

FIG. 60R

【図 60R-2】

PCR-BTX139																							
S32-NP-eval-002-PCR-BTX139																							
T	99.9%	0.04%	0.01%	99.9%	0.31%	0.00%	0%	99.9%	0.01%	0%	0.05%	0.01%	0%	0.01%	99.9%	0.04%	0.01%	99.9%	0.01%	0.04%	0.05%	99.9%	0%
G	0%	99.9%	0%	0.05%	99.9%	0.02%	0%	0%	0.01%	99.9%	0%	0.01%	99.9%	0%	0%	99.9%	0%	0%	99.9%	0%	0.02%	99.9%	0%
C	0.02%	0.01%	99.9%	0.03%	0%	0.03%	99.9%	0.01%	0.01%	99.9%	0%	0%	99.9%	0%	99.9%	0.01%	0.01%	99.9%	0.01%	0.03%	0.05%	99.9%	0%
A	0%	0%	0.03%	0%	0.01%	99.9%	0.1%	0%	99.9%	0.01%	99.9%	0.01%	0.02%	99.9%	0%	0%	0.01%	0%	0.15%	99.9%	0.01%	0%	0%
#	5512/	5437/	5415/	5478/	5461/	5403/	5420/	5445/	5551/	5421/	5295/	5425/	5436/	5391/	5457/	5459/	5370/	5389/	5380/	5435/	5330/	5443/	5395/
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	

PCR-BTX139																							
S32-NP-eval-003-PCR-BTX139																							
T	99.9%	0.05%	0.02%	99.9%	0.3%	0%	0%	99.9%	0.02%	0.01%	0.05%	0.01%	0.05%	0.01%	99.9%	0.04%	0.01%	99.9%	0.01%	0.04%	0.04%	99.9%	0%
G	0%	99.94%	0%	0%	0.05%	99.9%	0%	0%	0.03%	0%	0.06%	99.9%	0%	0.01%	99.9%	0%	0%	99.94%	0%	0.01%	99.95%	0%	0%
C	0.01%	0%	99.9%	0.02%	0%	0.03%	99.9%	0.01%	0.02%	99.9%	0%	0%	99.9%	0%	0.01%	0.02%	0%	99.9%	0.01%	99.9%	0.01%	0.01%	0%
A	0%	0.01%	0.05%	0%	0.01%	99.9%	0.1%	0%	99.9%	0.01%	99.9%	0.01%	0%	0.01%	0%	0.01%	0%	0%	0.15%	99.9%	0.01%	0%	0%
#	4901/	4832/	4914/	4837/	4820/	4841/	4746/	4827/	4719/	4824/	4840/	4793/	4853/	4816/	4735/	4787/	4837/	4759/	4842/	4814/	4812/	4814/	4814/
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	

FIG. 60R CONTINUED

【図 60R-3】

PCR-BTX139																							
S32-NP-eval-001-PCR-BTX139																							
T	99.9%	0.05%	0.3%	0.01%	0.05%	0.01%	0.01%	0.04%	0.01%	0.01%	0.02%	0.02%	0.01%	0.02%	99.9%	0.05%	0.01%	99.9%	0.01%	0.03%	0.04%	99.9%	0%
G	0%	99.9%	0%	0.05%	99.9%	0%	0%	0.02%	0.01%	0.05%	99.9%	0%	0.02%	0.1%	0%	99.9%	0%	0%	0%	0%	99.9%	0%	0%
C	0.02%	0%	99.9%	0.01%	0%	0.04%	99.9%	0.02%	0.02%	99.9%	0.01%	0%	99.9%	0%	99.9%	0.01%	0.03%	99.9%	0.01%	0.03%	0.05%	99.9%	0%
A	0.06%	0.01%	0.04%	0%	0%	99.9%	0.1%	0%	99.9%	0.01%	99.9%	0.01%	0%	0.01%	99.9%	0%	0.01%	0%	0.15%	99.9%	0.01%	0.01%	0%
#	5215/	5102/	5153/	5229/	5211/	5169/	5176/	5209/	5106/	5169/	5049/	5176/	5215/	5239/	5209/	5210/	5143/	5151/	5191/	5101/	5205/	5259/	5259/
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	

PCR-BTX139																							
S32-NP-eval-002-PCR-BTX139																							
T	99.9%	0.05%	0.01%	99.9%	0.2%	0%	0%	99.9%	0%	0.01%	0.04%	0.01%	0.02%	0.01%	99.9%	0.05%	0.01%	99.9%	0.01%	0.03%	0.04%	99.9%	0%
G	0%	99.9%	0%	0.05%	99.9%	0.01%	0%	0%	0.03%	0%	0.01%	99.9%	0%	0.01%	99.9%	0%	0%	99.9%	0%	0%	0.02%	99.9%	0%
A	0.01%	0.01%	99.9%	0%	0%	0.04%	99.9%	0%	0.01%	99.9%	0%	99.9%	0%	0%	99.9%	0.02%	0%	99.9%	0.02%	99.9%	0.01%	0%	0%
C	0.07%	0%	0.02%	0%	0.01%	99.9%	0.04%	0%	99.9%	0.02%	99.9%	0.01%	0.02%	99.9%	0.01%	0%	0.01%	99.9%	0.01%	0.03%	99.9%	0.01%	0.01%
52.05 51.91 51.75 51.63 51.70 51.53 51.94 51.55 50.65 50.62 50.42 51.04 51.66 50.86 51.67 51.55 50.53 50.01 50.80 51.16 50.47 55.51 52.12																							
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23																							

【図 60 R - 4】

PCR BTx139																							
S329-NP-rem-003-PCR-BT139																							
T	99.9%	0.02%	0.02%	99.94%	0.34%	0%	0%	99.99%	0.01%	0.01%	0.04%	0.01%	0.01%	0.02%	99.94%	99.99%	0.02%	0.02%	99.99%	0.05%	0.05%	0.08%	100%
G	0%	0%	0%	99.97%	0.04%	0.02%	0%	0%	0.02%	0%	0.02%	99.99%	0%	0.01%	0.4%	0%	99.75%	0%	0%	0%	0.07%	99.93%	0%
C	0.01%	0%	99.92%	0.04%	0.01%	0.03%	99.88%	0.02%	0%	99.97%	0.01%	0.01%	0.01%	0.03%	99.97%	0.01%	99.97%	0.01%	99.97%	0.05%	0.05%	0.16%	0%
A	0%	0.01%	0.06%	0%	0.02%	99.95%	0.11%	0%	99.94%	0.01%	99.93%	0.01%	0.02%	99.95%	0.01%	0.01%	0.2%	0.02%	0%	0.15%	99.79%	0.01%	0%
#	3392	33435	3394	33997	33615	3300	3303	3351	32974	33346	33927	33557	33969	33195	33599	33517	33007	33317	33069	33457	33949	33302	33917
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	

PCR BTx139																							
S329-NP-rem-004-PCR-BT139																							
T	99.9%	0.05%	0.01%	99.94%	0.33%	0%	0%	99.99%	0.01%	0%	0.03%	0.01%	0.01%	0.03%	99.9%	99.99%	0.04%	0.01%	99.99%	0.01%	0.04%	0.08%	99.98%
G	0%	0%	0%	99.94%	0.05%	0.01%	0%	0%	0.02%	0%	0.02%	99.97%	0%	0.02%	0.08%	0%	99.89%	0%	0%	0.07%	99.93%	0%	
C	0.01%	0%	99.94%	0.02%	0%	0.05%	99.87%	0.01%	0.02%	99.98%	0%	0.01%	99.95%	0%	0.02%	0.02%	99.97%	0.02%	99.97%	0.05%	0.05%	0.15%	0%
A	0%	0%	0.04%	0%	0.01%	99.94%	0.13%	0%	99.95%	0.02%	99.95%	0.01%	0.03%	99.95%	0.01%	0%	0.04%	0.01%	0%	0.14%	99.93%	0.02%	0%
#	5944	5909	5913	5912	5962	5799	5805	5945	5790	5907	5911	5912	5955	57945	5946	5943	5792	5717	5726	5932	5724	59445	59107
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	

FIG. 60R CONTINUED

【図 60 S - 1】

PCR BTx139																								
S334-NP-rem-003-PCR-BT139																								
T	99.97%	0.04%	0.01%	99.95%	0.35%	0.01%	0.01%	99.99%	0.01%	0.03%	0.03%	0.01%	0.03%	0.01%	0.02%	99.92%	99.99%	0.4%	0.01%	99.97%	0.01%	0.03%	0.05%	99.99%
G	0%	99.94%	0%	99.95%	99.85%	0.01%	0%	0%	0.02%	0%	0.01%	99.97%	0%	0.02%	0.07%	0.01%	99.95%	0%	0%	0%	0.01%	99.94%	0%	
C	0.01%	0%	99.95%	0.01%	0%	0.04%	99.86%	0.01%	0.01%	99.97%	0%	0%	99.97%	0.01%	0%	99.96%	0.03%	99.86%	0.03%	99.86%	0.17%	0%	0%	
A	0.02%	0.01%	0.06%	0%	0.01%	99.94%	0.13%	0%	99.95%	0.01%	99.95%	0%	0.02%	99.95%	0.01%	0.01%	0.01%	0.01%	0%	0.13%	99.79%	0.01%	0%	
#	5798	59455	5937	57059	5900	5929	5949	5975	55948	59487	59305	59495	59336	59145	59539	5973	59955	59117	59985	5951	55994	59731	5791	
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23		

PCR BTx139																							
S335-NP-rem-004-PCR-BT139																							
T	99.9%	0.06%	0.02%	99.95%	0.38%	0.01%	0%	99.99%	0%	0.04%	0.04%	0.01%	0%	0.02%	99.95%	99.99%	0.05%	0.02%	99.97%	0.01%	0.03%	0.05%	99.98%
G	0%	0%	0.03%	99.93%	0%	0%	0%	0%	0.03%	0%	0.01%	99.98%	0%	0.01%	0.12%	0%	99.93%	0%	0%	0%	0.02%	99.93%	0%
C	0.02%	0.01%	99.99%	0.01%	0%	0.05%	99.85%	0.01%	0.02%	99.98%	0.01%	0.01%	99.97%	0.02%	0.01%	99.97%	0.03%	99.97%	0.03%	99.85%	0.15%	0%	0.02%
A	0%	0%	0.06%	0%	0%	99.94%	0.15%	0%	99.95%	0.01%	99.95%	0%	0.02%	99.95%	0.01%	0.01%	0%	0.01%	0%	0.44%	99.81%	0.02%	0.01%
#	5323	5226	5278	5296	5246	5744	5246	5295	5175	5243	5192	5290	5278	5271	5279	5297	5198	5204	5185	5243	5153	5292	5240
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	

FIG. 60S

【図 60 S - 2】

PCR BTx139																							
S336-NP-rem-002-PCR-BT139																							
T	100%	0%	0%	100%	0%	0%	0%	100%	0%	0%	0%	0%	0%	0%	100%	100%	0%	0%	100%	0%	0%	0%	100%
G	0%	100%	0%	0%	0%	0%	0%	0%	0%	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
C	0%	0%	100%	0%	0%	0%	0%	0%	0%	0%	100%	0%	0%	0%	0%	0%	100%	0%	100%	0%	0%	0%	0%
A	0%	0%	0%	0%	100%	0%	0%	0%	0%	0%	0%	100%	0%	0%	0%	0%	0%	0%	0%	0%	100%	0%	0%
#	60	59	59	57	59	60	52	59	57	57	59	57	61	59	56	59	59	59	57	59	57	59	62
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	

PCR BTx139																							
S337-NP-rem-003-PCR-BT139																							
T	100%	0%	0%	0%	0%	0%	0%	0%	100%	0%	0%	0%	0%	0%	100%	100%	0%	0%	100%	0%	0%	0%	100%
G	0%	100%	0%	0%	0%	0%	0%	0%	0%	0%	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
C	0%	0%	100%	0%	0%	0%	0%	0%	0%	0%	0%	100%	0%	0%	0%	0%	100%	0%	100%	0%	0%	0%	0%
A	0%	0%	0%	0%	100%	0%	0%	0%	0%	0%	0%	0%	100%	0%	0%	0%	0%	0%	0%	0%	100%	0%	0%
#	16	14	15	16	16	15	13	12	14	15	13	12	16	13	12	13	14	15	14	15	14	12	16
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	

FIG. 60S CONTINUED

【図 60 S - 3】

PCR BTx139																								
S338-NP-rem-002-PCR-BT139																								
T	99.95%	0.05%	0.01%	99.92%	0.44%	0.01%	0.01%	99.97%	0.02%	0.02%	0.02%	0.01%	0.03%	0.01%	0.02%	99.98%	99.97%	0.04%	0.01%	99.95%	0.01%	0.05%	0.04%	99.95%
G	0%	0%	0.05%	99.95%	0.01%	0%	0%	0%	0.02%	0.01%	0.05%	99.95%	0%	0.02%	0.1%	0%	99.95%	0%	0%	0%	0.01%	99.94%	0%	
C	0.02%	0.01%	99.95%	0.01%	0%	0.04%	99.85%	0.02%	0.01%	99.98%	0%	0%	99.95%	0%	0.01%	0.02%	99.98%	0.02%	99.84%	0.44%	0%	0%	0%	
A	0.03%	0.01%	0.04%	0%	0.01%	99.94%	0.16%	0%	99.95%	0.01%	99.95%	0.01%	0.03%	99.95%	0.01%	0%	0.01%	0.01%	0%	0.15%	99.83%	0.01%	0%	
#	5955	5705	5715	5950	5834	5779	5709	5912	5701	5794	5943	5792	5432	5935	5935	5757	5705	5735	5935	5703	5701	5909	5941	
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23		

PCR_BT139																							
S331-NP-rem-003-PCR-BT139																							
T	99.9%	0.05%	0.01%	99.94%	0.35%	0.01%	0%	99.99%	0.02%	0.01%	0.04%	0.03%	0.01%	0.03%	99.95%	99.99%	0.04%	0.01%	99.95%	0%	0.05%	0.04%	99.95%
G	0%	99.94%	0%	0.05%	99.94%	0.01%	0%	0%	0.03%	0%	0.03%	99.95%	0%	0.02%	0.09%	0%	99.87%	0%	0%	0%	0.01%	99.95%	0.01%
C	0.01%	0%	0.01%	99.97%	0.02%	0%	0.04%	99.95%	0.01%	0.02%	99.98%	0.01%	0%	99.97%	0%	0.01%	0.01%	0.02%	99.98%	0.02%	99.92%	0.63%	0%
A	0.01%	0%	0.02%	99.97%	0.02%	0%	0.04%	99.95%	0.01%	0%	99.94%	0.01%	0.02%	99.95%	0.02%	0%	0.08%	0%	0%	0.17%	99.77%	0.01%	0%
del	99.95%	99.94%	99.98%	99.99%	99.92%	99.99%	99.99%	99.99%	99.99%	99.99%	99.99%	99.99%	99.99%	99.99%	99.99%	99.99%	99.99%	99.99%	99.99%	99.99%	99.99%	99.99%	99.99%

【図 60 S - 4】

PCR BTx139																						
S332NP-neo-001-PCR-BT139																						
T	99.9%	0.03%	0.01%	99.9%	0.33%	0%	0.05%	99.9%	0.02%	0.01%	0.02%	0.05%	0.01%	0.02%	99.9%	0.03%	0.01%	99.9%	0.03%	0.05%	99.9%	
G	0%	99.9%	0%	0.05%	99.0%	0.01%	0%	0%	0%	0.01%	0.08%	99.9%	0%	0.01%	0.08%	0%	0.01%	0%	99.9%	0.33%	0%	
C	0.01%	0%	99.9%	0.01%	0%	0.05%	99.9%	0.01%	0.02%	99.9%	0%	0%	0.01%	0.01%	0.01%	0.01%	0.01%	99.0%	0.01%	0.01%	0%	
A	0%	0.01%	0.03%	0%	0%	99.9%	0.12%	0%	99.9%	0%	0.02%	99.9%	0.02%	0.01%	0.03%	0.02%	0%	0.14%	99.9%	0.01%	0%	
#	57045/5807/5917/5707/5834/5812/5102/5945/5500/5512/5941/5768/5540/5551/5642/5513/5573/5573/5551/5504/5532/5539/5707																					
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23

PCR BTx139																						
S333NP-neo-002-PCR-BT139																						
T	99.90%	0.05%	0.01%	99.93%	0.33%	0.01%	0.1%	99.99%	0.02%	0%	0.05%	0.02%	0.01%	0.02%	99.99%	0.05%	0.01%	99.99%	0%	0.04%	0.04%	99.99%
G	0%	99.94%	0%	99.95%	0.01%	0%	0%	99.99%	0%	0.02%	0%	0.02%	99.97%	0%	0.01%	0.1%	99.91%	0%	0.01%	0.01%	99.95%	0%
C	0.01%	0.01%	99.95%	0.02%	0%	0.03%	99.70%	0.01%	0.01%	99.99%	0%	0%	99.98%	0%	0.01%	0.01%	0.01%	99.99%	0.02%	99.95%	0.01%	0%
A	0.01%	0%	0.04%	0.01%	0.01%	99.95%	0.14%	0%	99.95%	0.01%	99.99%	0.01%	99.97%	0%	0.01%	0%	0.02%	0.01%	0%	99.95%	0.01%	0%
#	5556/5753/5722/5917/5504/5741/5752/5783/5807/5735/5535/5795/5904/5704/5704/5795/5745/5705/5724/5700/5778/5553/5722/5545																					
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23

FIG. 60S CONTINUED

【図 60 T - 1】

PCR BTx139																					
S342NP-neo-002-PCR-BT139																					
T	55.9%	0.03%	0.01%	53.9%	0.33%	0.01%	0.01%	50.0%	0.03%	0.01%	0.01%	50.0%	0.03%	0.01%	50.0%	0.03%	0.01%	50.0%	0.03%	0.05%	0.05%
G	0%	50.9%	0%	0.05%	50.0%	0.02%	0%	0%	0.03%	0%	0.17%	50.0%	0%	0.01%	0.08%	0%	50.0%	0%	0%	50.0%	0%
C	0.01%	0%	50.0%	0.01%	0%	0.03%	50.0%	0.03%	0.02%	50.0%	0%	0%	0.01%	0.02%	0.01%	50.0%	0.01%	0.01%	50.0%	0%	
A	0.01%	0%	0.03%	0%	0%	50.0%	0.11%	0%	50.0%	0.02%	50.0%	0.01%	0.03%	50.0%	0.01%	0.01%	0.01%	0.2%	50.0%	0.01%	0%
#	4276/4400/4465/4884/4885/4438/4438/4450/4450/4057/4530/4038/4450/4486/4724/4495/4724/4111/4221/4171/4181/4187/4276																				
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22

PCR BTx139																						
S343NP-neo-002-PCR-BT139																						
T	55.9%	0.04%	0.01%	53.9%	0.29%	0.01%	0.05%	50.0%	0.02%	0.05%	0.02%	0.02%	0.02%	50.0%	0.03%	0.01%	50.0%	0.03%	0.01%	50.0%	0.03%	0.05%
G	0%	50.9%	0%	0.05%	50.0%	0.01%	0%	0%	0.02%	0%	0.03%	50.0%	0%	0.01%	0.08%	0%	50.0%	0%	0%	50.0%	0%	0%
C	0.01%	0%	50.0%	0.02%	0%	0.03%	50.0%	0.03%	0.02%	50.0%	0%	0%	0.01%	0.02%	0.02%	50.0%	0.01%	0.01%	50.0%	0.01%	50.0%	0%
A	0.01%	0%	0.05%	0%	0%	50.9%	0.15%	0%	50.0%	0%	50.0%	0%	0.01%	50.0%	0%	0.05%	0.01%	0.14%	50.0%	0.01%	0%	0%
#	4588/4496/4496/4234/4234/4185/4185/4271/4143/4153/4108/4195/4259/4190/4277/4215/4190/4134/4189/4719/4190/4718/4291																					
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	

FIG. 60T

【図 60 T - 2】

PCR BTx139																						
S344NP-neo-001-PCR-BT139																						
T	50.0%	0.04%	0.02%	50.0%	0.01%	0%	50.0%	0.03%	0.01%	0.02%	50.0%	0.03%	0.01%	50.0%	0%	0.04%	0.03%	50.0%	0%	0.04%	0.03%	50.0%
G	0%	50.9%	0%	0.07%	50.0%	0.01%	0%	0.01%	0%	0.01%	50.0%	0%	0%	0.08%	0%	50.0%	0%	0%	50.0%	0%	50.0%	0%
C	0.02%	0%	50.0%	0.02%	0.01%	0.03%	50.0%	0.02%	0.03%	50.0%	0%	0.01%	0.02%	0%	50.0%	0.01%	0.01%	50.0%	0.01%	50.0%	0.01%	0.04%
A	0%	0.01%	0.05%	0%	0.01%	50.0%	0.12%	0.01%	50.0%	0.01%	0.02%	50.0%	0%	0.01%	0.01%	0%	0.15%	50.0%	0.02%	0%	0%	0%
#	4011/59405/3936/3972/5975/3917/3947/3959/3904/3950/3946/3949/3970/3937/3949/3917/3922/3916/3951/3916/3959/4041																					
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23

PCR BTx139																						
S345NP-neo-002-PCR-BT139																						
T	50.0%	0.07%	0.01%	50.0%	0.33%	0.01%	0.04%	50.0%	0%	0.01%	0.03%	0.01%	0.01%	0.02%	50.0%	0.03%	0.01%	50.0%	0%	0.03%	0.05%	50.0%
G	0%	50.9%	0%	0.04%	50.0%	0.01%	0%	0%	0.01%	0%	0.02%	50.0%	0%	0.01%	0.1%	0%	50.0%	0%	0%	50.0%	0%	
C	0.03%	0%	50.0%	0.03%	0%	0.05%	50.0%	0.02%	0.02%	50.0%	0%	0%	0.01%	0.01%	0%	50.0%	0.01%	0.01%	50.0%	0.01%	50.0%	
A	0.08%	0.01%	0.05%	0%	0.01%	50.0%	0.13%	0%	50.0%	0.02%	50.0%	0%	0.01%	50.0%	0%	0.12%	50.0%	0.01%	0%	50.0%	0.01%	
#	4280/4247/4200/4239/4241/4191/4262/4262/4198/4263/4273/4216/4247/4246/4245/4286/4801/4801/4181/4237/4181/4243/4275																					
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	

FIG. 60T CONTINUED

【図 60 T - 3】

PCR BTx139																						
S338NP-neo-001-PCR-BT139																						
T	50.0%	0.06%	0.01%	50.0%	0.3%	0.01%	0%	50.0%	0.01%	0.01%	0.06%	0.02%	0.01%	0.03%	50.0%	0.04%	0.01%	50.0%	0%	0.04%	0.05%	50.0%
G	0%	50.9%	0%	0.07%	50.0%	0.01%	0%	0%	0.02%	50.0%	0%	0.02%	50.0%	0%	0.05%	0.01%	50.0%	0%	0%	50.0%	0%	
C	0.01%	0%	50.0%	0.01%	0%	0.04%	50.0%	0.02%	0.01%	50.0%	0%	0%	50.0%	0%	0.01%	0%	50.0%	0%	0.12%	50.0%	0%	
A	0%	0.01%	0.04%	0%	0%	50.9%	0.15%	0%	50.0%	0.01%	50.0%	0.01%	0.01%	0%	50.0%	0%	0.14%	50.0%	0.01%	0%	0%	
#	3271/3707/3162/3200/3185/3165/3192/3519/3130/3175/3103/3170/3190/3191/3195/3165/3159/3145/3196/3147/3181/3223																					
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	

PCR BTx139																						
S339NP-neo-002-PCR-BT139																						
	50.0%	0.04%	0.01%	50.0%	0.33%	0.01%	0.08%	50.0%	0%	0.02%	0.04%	0.02%	0.02%	0.01%	50.0%	0.04%	0.01%	50.0%	0.02%	0.05%	0.02%	50.0%
	0%	50.0%	0%	50.0%	0.02%	0%	0.02%	0%	0.02%	0%	0.02%	0.02%	0%	0%	50.0%	0%	50.0%	0%	0%	0.01%	50.0%	
	0.02%	0%	50.0%	0.03%	0.03%	0%	0.04%	50.0%	0.01%	50.0%	0%	0%	0.01%	0.01%	50.0%	0.02%	0%	50.0%	0.02%	0.01%	0.01%	
	0.05%	0%	0.04%	0.01%	50.0%	0.02%	0%	50.0%	0.01%	50.0%	0.01%	0.01%	0%	0.01%	50.0%	0%	0.01%	0.01%	0.01%	0.02%	0.01%	
	50.0%	50.0%	50.0%	50.0%	50.0%	50.0%	50.0%	50.0%	50.0%	50.0%	50.0%	50.0%	50.0%	50.0%	50.0%	50.0%	50.0%	50.0%	50.0%	50.0%	50.0%	

【図 60U-4】

PCR B5x139																							
Sis4MP-new140C-PCR-B5x139																							
T	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
G	0%	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
C	0%	0%	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
A	0%	0%	0%	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
#	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
PCR B5x139																							
Sis4MP-new140C-PCR-B5x139																							
T	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
G	0%	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
C	0%	0%	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
A	0%	0%	0%	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
#	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24

【図 60V-1】

PCR BTx139																								
S35-WP-nm14-00-PCRBTx139																								
T	99.95%	0.00%	0.01%	99.95%	0.25%	0.01%	0.05%	99.99%	0%	0.01%	0.04%	0.05%	0%	0.02%	99.99%	99.99%	0.04%	0.02%	99.99%	0%	0.05%	0.05%	99.99%	0%
G	0%	99.97%	0%	0.04%	99.97%	0.01%	0%	0%	0.01%	0.01%	0.05%	99.95%	0%	0%	99.95%	99.95%	0%	0%	99.95%	0%	0.01%	99.95%	0%	0%
C	0.04%	0%	99.95%	0.01%	0%	0.04%	99.97%	0.01%	0.01%	99.99%	0%	0%	99.97%	0%	0.01%	0.01%	0.01%	0.01%	99.95%	0.01%	99.95%	0.15%	0%	0.02%
A	0%	0%	0.03%	0%	99.94%	0%	0%	99.97%	0%	99.97%	0%	0.02%	99.99%	0%	0%	0.03%	0.02%	0%	0.44%	99.92%	0.01%	0%	0%	0%
#	4240	4170	4394	4294	4304	4157	4165	4125	4166	4078	4160	4195	4148	4197	4179	4177	4140	4139	4175	4105	4155	4234	423	23
PCR BTx139																								
S35-WP-NP237-00-PCRBTx139																								
T	99.95%	0.05%	0%	99.94%	0.33%	0.24%	0%	99.97%	0.01%	0.01%	0.03%	0.02%	0.01%	0.01%	99.99%	99.99%	0.04%	0%	99.99%	0.01%	0.05%	0.05%	99.99%	0%
G	0%	99.94%	0%	0.05%	99.95%	0.80%	0%	0%	4.01%	0.01%	0.05%	99.97%	0%	0.02%	0.07%	0%	99.95%	0%	0%	0.01%	0.02%	99.95%	0%	0%
C	0.01%	0%	99.95%	0.02%	0.01%	0.03%	99.98%	0.02%	0.02%	99.97%	0%	0.01%	99.97%	0%	0.01%	0.01%	0%	99.99%	0.01%	99.92%	0.44%	0%	0.01%	0%
A	0%	0.01%	0.05%	0%	0.01%	92.05%	0.11%	0.01%	95.95%	0.02%	99.98%	0.01%	0.02%	99.97%	0.01%	0.01%	0%	0.01%	0%	0.17%	99.91%	0.01%	0%	0%
#	4077	9349	9349	95629	95153	91065	93407	95557	95694	94693	93807	93369	93617	93022	96655	94946	94939	91733	93045	94945	93947	93005	90077	4077

FIG. 60V

【図 60V-2】

PCR BT139																								
S361-AP-AP231-403-PCR-BT139																								
100%	12%	0%	0%	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	100%	0%	0%	0%	100%
0%	30.0%	0%	0%	100%	5.88%	0%	0%	0%	0%	0%	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	100%
0%	0%	0%	0%	100%	0%	0%	0%	0%	0%	0%	100%	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%	0%	0%	0%	0%	0%	0%	0%	100%
5%	5%	5%	5%	5%	5%	5%	5%	5%	5%	5%	5%	5%	5%	5%	5%	5%	5%	5%	5%	5%	5%	5%	5%	5%
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

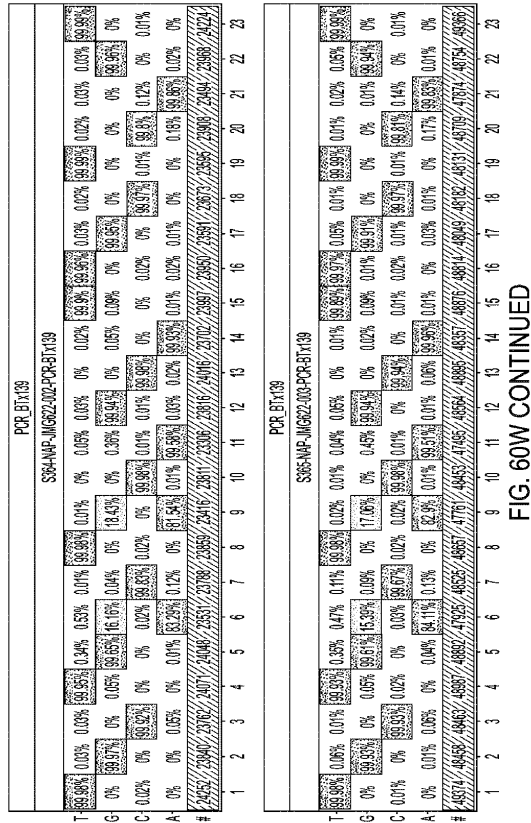
PCR BT139																								
S361-AP-AP231-403-PCR-BT139																								
100%	0%	0%	0%	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	100%	0%	0%	0%	100%
0%	100%	0%	0%	100%	0%	0%	0%	0%	0%	0%	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	100%
0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	100%
0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	100%
13	12	11	10	12	11	9	10	11	12	11	11	12	10	12	12	10	10	10	11	12	12	13	14	15

【図 60V-3】

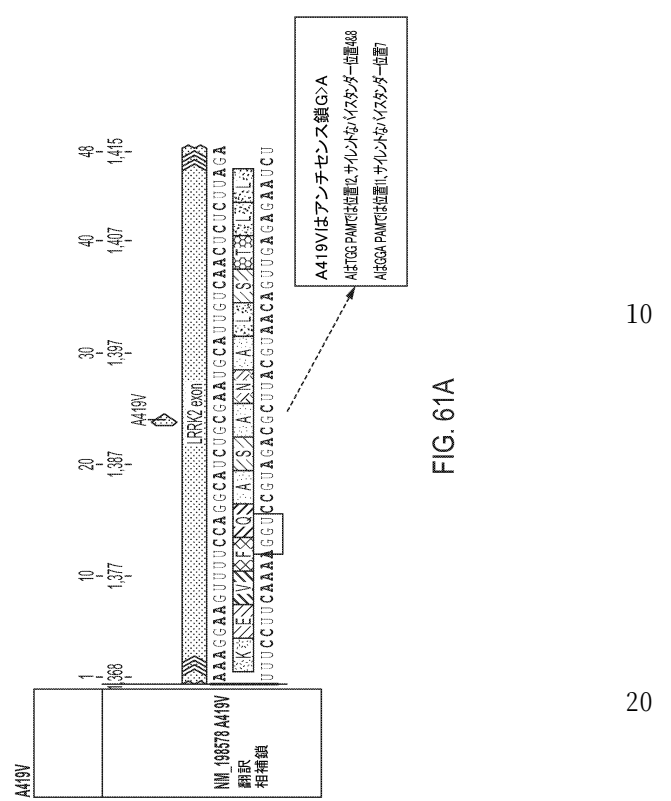
POB BT-139																								
SS54-MP-neat3-002-PCB-BT-139																								
99.95%	0.05%	0.01%	99.95%	0.04%	0.01%	99.97%	0.01%	0.04%	0.02%	0.01%	0.05%	0.03%	0.01%	99.97%	0.01%	0.04%	0.04%	99.95%	0%					
0%	99.92%	0%	0.07%	99.95%	0.05%	0%	0.05%	0%	0.04%	99.95%	0%	0.01%	0.1%	99.93%	0%	0.01%	0.02%	99.95%	0%					
0.01%	0%	99.96%	0.04%	0%	0.04%	0.03%	0.03%	99.96%	0%	0%	0%	0.01%	0%	99.96%	0.03%	0.03%	0%	0.01%	0.01%					
0%	0.01%	0.03%	0%	0.02%	99.89%	0.05%	0%	99.92%	0%	99.97%	0%	0.01%	0.02%	99.96%	0.02%	0.03%	0.02%	0%	0.03%	0.01%				
99.97%	0.03%	99.97%	0.02%	99.95%	0.03%	99.97%	0.01%	99.95%	0.03%	99.97%	0.01%	99.95%	0.03%	99.97%	0.01%	99.95%	0.03%	99.97%	0.01%	99.95%				
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23		
POB BT-139																								
SS55-MP-neat3-003-PCB-BT-139																								
99.95%	0.04%	0.01%	99.95%	0.25%	0%	0.04%	99.99%	0.01%	0%	0.03%	0.01%	0.03%	0.03%	99.91%	99.98%	0.01%	0.05%	0.04%	99.99%					
0%	99.95%	0%	0.04%	99.95%	0.01%	0%	0%	0.02%	0%	0.1%	99.95%	0%	0.08%	0%	99.95%	0%	0%	0.01%	99.94%					
0.01%	0%	99.95%	0.01%	0.01%	0.04%	99.94%	0.01%	0.01%	99.96%	0%	0%	0.02%	0%	99.97%	0.01%	99.96%	0.12%	0%	0.01%					
0.03%	0.01%	0.04%	0%	0%	99.95%	0.12%	0%	99.95%	0.01%	99.96%	0%	0.01%	0.01%	99.97%	0%	0.13%	99.95%	0.01%	0%					
99.97%	0.03%	99.97%	0.02%	99.95%	0.03%	99.97%	0.01%	99.95%	0.03%	99.97%	0.01%	99.95%	0.03%	99.97%	0.01%	99.95%	0.03%	99.97%	0.01%	99.95%				
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23		
POB BT-139																								
SS56-MP-neat3-004-PCB-BT-139																								
99.95%	0.04%	0.01%	99.95%	0.25%	0%	0.04%	99.99%	0.01%	0%	0.03%	0.01%	0.03%	0.03%	99.91%	99.98%	0.01%	0.05%	0.04%	99.99%					
0%	99.95%	0%	0.04%	99.95%	0.01%	0%	0%	0.02%	0%	0.1%	99.95%	0%	0.08%	0%	99.95%	0%	0%	0.01%	99.94%					
0.01%	0%	99.95%	0.01%	0.01%	0.04%	99.94%	0.01%	0.01%	99.96%	0%	0%	0.02%	0%	99.97%	0.01%	99.96%	0.12%	0%	0.01%					
0.03%	0.01%	0.04%	0%	0%	99.95%	0.12%	0%	99.95%	0.01%	99.96%	0%	0.01%	0.01%	99.97%	0%	0.13%	99.95%	0.01%	0%					
99.97%	0.03%	99.97%	0.02%	99.95%	0.03%	99.97%	0.01%	99.95%	0.03%	99.97%	0.01%	99.95%	0.03%	99.97%	0.01%	99.95%	0.03%	99.97%	0.01%	99.95%				
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23		

FIG. 60V CONTINUED

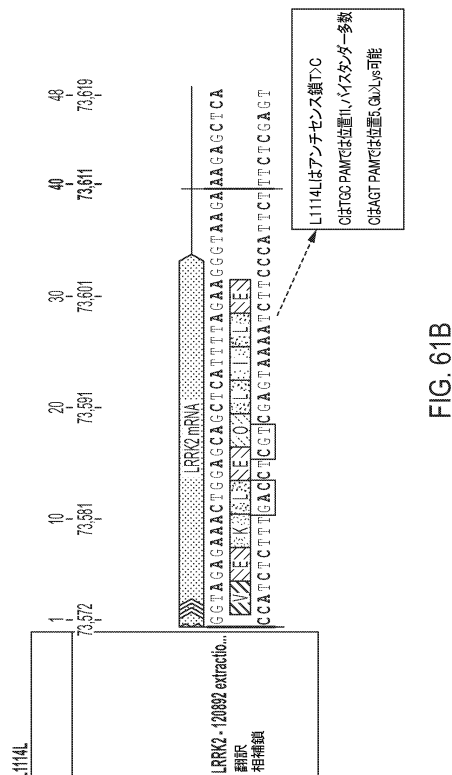
【図 60W-4】



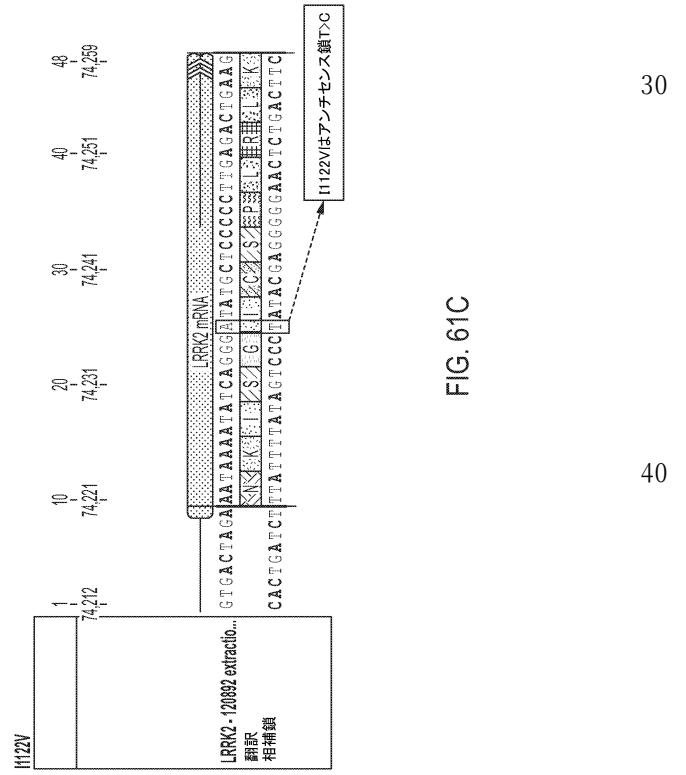
【図 6 1 - 1】



【圖 6 1 - 2】



【図 6 1 - 3】



【図 6 1 - 4】

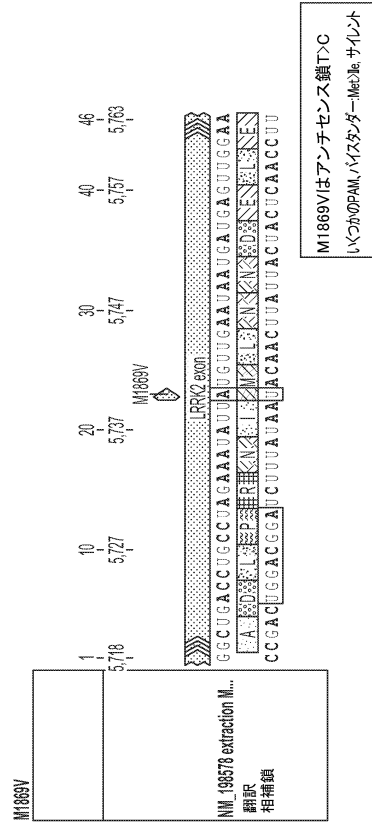


FIG. 61D

【図 6 2 A】

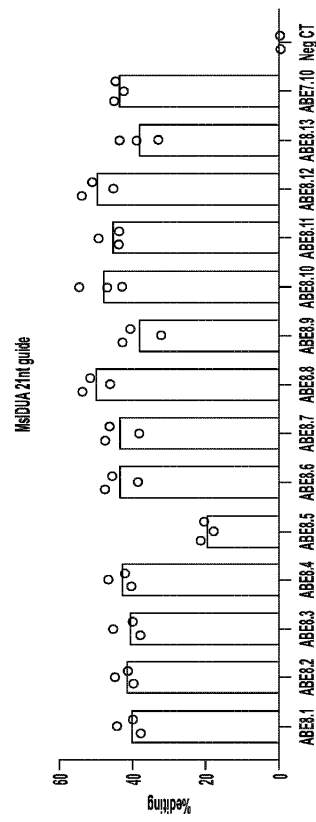


FIG. 62A

【図 6 2 B】

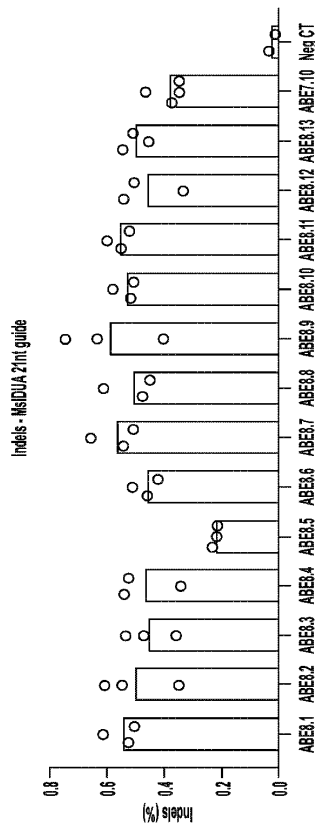


FIG. 62B

【図 6 3】

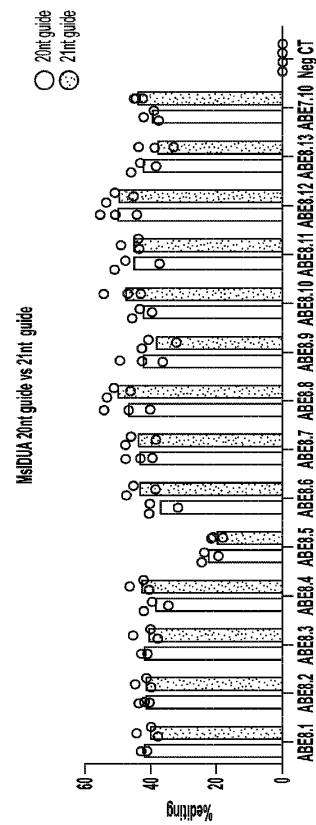


FIG. 63

【図 6 6 B】

		hudsonium																				
		S128-hudsonium28-1																				
T	0.04%	0%	100%	0.01%	99.99%	0.02%	0%	0.02%	0.01%	0.01%	0%	0%	0.01%	0.01%	0%	100%	0%	100%	0%	0%	0%	0.01%
G	99.99%	0%	0%	0%	0%	41.65%	99.99%	99.99%	0%	99.99%	0%	0%	99.99%	0%	0%	99.99%	0%	0%	0%	0%	99.97%	0%
C	0%	0%	99.99%	0.01%	99.99%	0%	0%	0%	99.99%	99.99%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	99.97%	0%
A	0%	0%	0%	0.01%	0%	99.02%	0.01%	0%	0.01%	0.01%	0%	99.99%	99.99%	0%	0%	0%	0%	0%	0%	0%	0.02%	0.01%
#	2138	134	2124	142	7485	2137	2159	2159	2159	2159	2159	2159	2159	2159	2159	2159	2159	2159	2159	2159	2159	2159
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20			

		hudsonium																				
		S127-hudsonium28-2																				
T	0.02%	0.05%	99.99%	0.01%	99.99%	0.01%	0.02%	0%	0.02%	0.01%	0%	0%	0.01%	99.99%	0%	99.99%	0%	100%	0%	0%	0%	0.01%
G	99.95%	0%	0%	0.05%	0%	40.68%	99.99%	99.99%	0%	99.99%	0.01%	0.01%	99.99%	0%	0%	100%	0%	0%	0%	0%	99.99%	0%
C	0%	0%	99.94%	0.01%	99.97%	0%	0%	0%	99.99%	99.99%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	99.99%	0%
A	0.01%	0%	0%	0%	0%	99.91%	0.01%	0.01%	0%	0.02%	99.99%	99.99%	0%	0%	0%	0%	0%	0%	0%	0%	0.03%	0.01%
#	2201	773	2794	130	7980	2776	2741	2798	2792	2791	2795	2795	2795	2795	2795	2795	2795	2795	2795	2795	2795	2795
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20			

		hudsonium																				
		S128-hudsonium28-3																				
T	0.02%	0.05%	99.99%	0.02%	99.99%	0.04%	0%	0.01%	0.01%	0%	0%	0%	0%	0.01%	100%	0%	99.99%	0%	0%	0%	0.01%	0.01%
G	99.99%	0%	0%	0.01%	0%	35.02%	99.99%	99.99%	0%	99.99%	0.02%	0.02%	99.99%	0%	99.97%	0%	99.99%	0%	0%	0%	99.99%	0%
C	0%	0%	99.95%	0.01%	99.97%	0.01%	0%	0%	99.97%	99.99%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	99.97%	0%
A	0%	0%	0%	0%	0%	99.99%	0.04%	0.01%	0.02%	0%	99.97%	99.99%	0%	0%	0%	0%	0%	0%	0%	0%	0.04%	0.01%
#	9274	9206	2903	2531	2945	2526	2841	2885	2507	2514	2944	2937	2842	2936	2939	2940	2938	2939	2939	2939	2939	2939
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20			

		hudsonium																				
		S128-hudsonium28-4																				
T	0.02%	0.05%	99.99%	0.02%	99.99%	0.04%	0%	0.01%	0.01%	0%	0%	0%	0%	0.01%	100%	0%	99.99%	0%	0%	0%	0.01%	0.01%
G	99.99%	0%	0%	0.01%	0%	35.02%	99.99%	99.99%	0%	99.99%	0.02%	0.02%	99.99%	0%	99.97%	0%	99.99%	0%	0%	0%	99.99%	0%
C	0%	0%	99.95%	0.01%	99.97%	0.01%	0%	0%	99.97%	99.99%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	99.97%	0%
A	0%	0%	0%	0%	0%	99.99%	0.04%	0.01%	0.02%	0%	99.97%	99.99%	0%	0%	0%	0%	0%	0%	0%	0%	0.04%	0.01%
#	9274	9206	2903	2531	2945	2526	2841	2885	2507	2514	2944	2937	2842	2936	2939	2940	2938	2939	2939	2939	2939	2939
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20			

Position along large site (bp)

FIG. 66b

【図 6 6 C】

	huda:samplem																			
	SI02-huda:rmq2d2-1																			
T	0.01%	0.05%	39.97%	0.01%	39.99%	0.02%	0.02%	0.01%	0.01%	0.01%	0%	0%	0.01%	39.99%	0.02%	39.99%	0.02%	0.01%	0.01%	0.01%
G	39.97%	0%	0.01%	0.01%	0%	34.01%	39.95%	39.97%	0.01%	0%	39.97%	0.02%	0.02%	39.98%	0%	39.97%	0%	0%	39.95%	0%
C	0%	39.93%	0.02%	39.97%	0.01%	0.03%	0%	0%	39.99%	39.95%	0%	0.01%	0%	0%	0.01%	0%	0.01%	39.97%	0%	39.97%
A	0.01%	0.01%	0%	0%	39.95%	0.01%	0.01%	0.01%	0%	0%	0.01%	0%	0%	0%	0%	0%	0.02%	0.02%	0.02%	0.02%
#	2545	3948	2398	2537	6715	2537	2510	2300	2377	3398	2587	3323	2017	2593	347	3653	355	2572	19	2945

	huda:samplem																			
	SI03-huda:rmq2d2-2																			
T	0.04%	0.02%	39.95%	0.01%	39.99%	0.04%	0.01%	0%	0.01%	0.02%	0.01%	0.01%	0%	0%	39.99%	0.01%	39.99%	0.02%	0.01%	0.01%
G	39.95%	0%	0.01%	0%	0%	31.98%	39.95%	39.99%	0%	0%	39.99%	0.01%	0.01%	39.99%	0%	39.99%	0%	0%	39.99%	0%
C	0%	39.97%	0%	39.97%	0.01%	0.03%	0%	0%	39.99%	39.97%	0%	0%	0.01%	0%	0.01%	0%	0.01%	39.95%	0%	39.99%
A	0%	0.01%	0%	0.01%	0.01%	0%	39.94%	0%	0%	0%	0.01%	0.01%	39.97%	39.99%	0%	0.01%	0%	0.02%	0.02%	0.01%
#	2305	3217	2919	2323	3445	2345	2367	2995	2367	2915	244	2340	2671	2390	3239	2371	2387	2329	2594	2321

	huda:samplem																			
	SI01-huda:rmq2d2-3																			
T	0.02%	0.04%	39.98%	0.03%	39.99%	0.01%	0.01%	0.02%	0%	0%	0.02%	0.01%	0%	0%	39.99%	0.02%	39.99%	0.02%	0.01%	0.01%
G	39.98%	0%	0%	0.03%	0%	30.01%	39.97%	39.99%	0%	0%	39.97%	0.01%	0.01%	39.99%	0%	39.97%	0%	0%	39.97%	0%
C	0%	39.94%	0%	39.93%	0.01%	0%	39.99%	39.99%	0%	0%	39.99%	0.01%	0%	0%	0.01%	0%	0.01%	39.97%	0%	39.99%
A	0%	0.01%	0.01%	0.01%	0.01%	39.97%	0.02%	0.01%	0.01%	0%	0%	0.01%	39.99%	0.01%	0%	0.01%	0%	0.01%	0.02%	0.01%
#	2612	2011	2985	629	6548	2617	2562	2535	294	2325	294	2545	2613	2590	2627	2688	2597	2615	19	2945

Position along target site (bp)

FIG. 66C

【図 6 6 D】

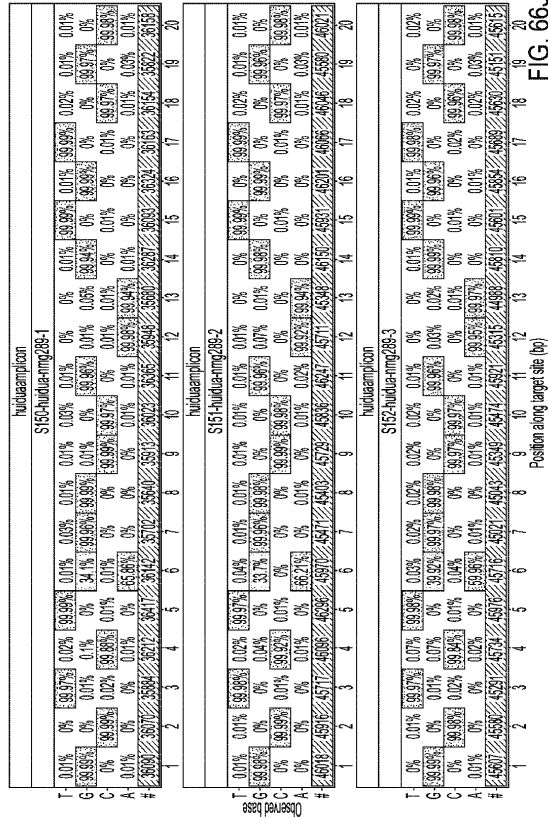
[illegible]

【図 6 6 E】

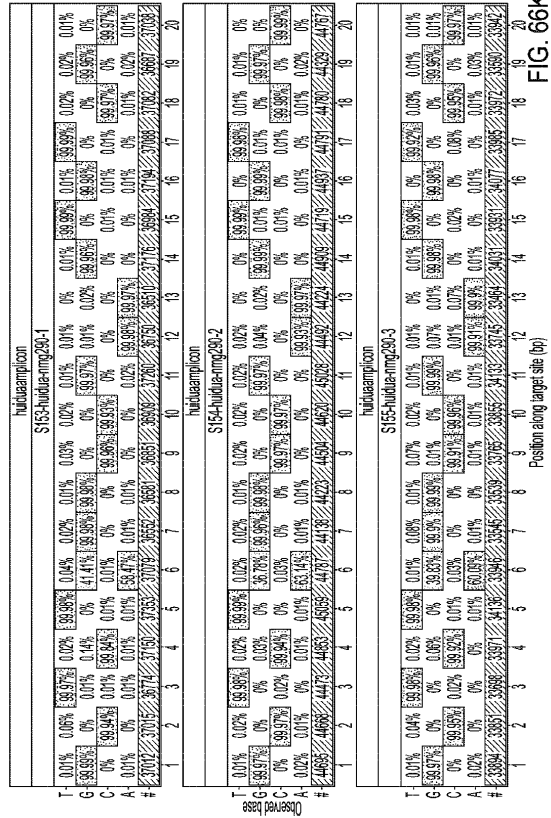
		huda:sampleon																			
		S165-huda:rmc2d-4																			
T	0.01%	0.01%	0%	0%	0.01%	0%	0.01%	0%	0.01%	0%	0.01%	0%	0.01%	0%	0.01%	0%	0.01%	0%	0.01%	0.01%	
G	99.99%	0%	0%	0%	8.14%	99.99%	99.99%	0%	0%	0%	99.99%	0.01%	0%	0%	99.99%	0%	99.99%	0%	99.97%	0.01%	
C	0%	99.99%	0.02%	99.99%	0.01%	0.01%	0%	0%	0%	99.99%	99.99%	0%	0%	0%	0%	0.02%	99.99%	0.01%	99.99%	99.99%	
A	0%	0%	0%	0%	99.99%	0.01%	0.01%	0%	0%	0%	99.99%	99.99%	0.01%	0%	0%	0%	0.02%	0.02%	0.01%	0.01%	
#	43242	0.05%	43168	4330	9732	43495	43665	43815	43907	4371	43747	4322	4307	43616	4443	43546	43513	43516	44055	44112	
		huda:sampleon																			
		S165-huda:rmc2d-2																			
T	0%	0.01%	99.97%	0.01%	99.99%	0%	0.05%	0.01%	0.03%	0.01%	0%	0.01%	0.01%	0%	99.98%	0%	99.98%	0.02%	0.02%	0.01%	
G	99.99%	0%	0%	0%	0%	0%	0.02%	99.99%	99.99%	0%	0%	99.97%	0.01%	0%	99.99%	0%	99.97%	0.03%	0%	99.99%	
C	0%	99.97%	0.02%	99.99%	0.01%	0%	0%	0%	99.99%	99.99%	0%	0%	0.03%	0%	0.01%	0.02%	0.01%	99.97%	0%	99.99%	
A	0.02%	0.01%	0%	0%	99.97%	0.01%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0.02%	0.01%	0.01%	
#	33007	4001	94601	4461	9513	94922	9481	9441	94737	9378	95714	94762	94495	93951	94611	93930	94943	94401	94949	94883	
		huda:sampleon																			
		S167-huda:rmc2d-3																			
T	0.01%	0%	99.99%	0.01%	99.99%	0%	0.04%	0.02%	0.01%	0.02%	0.01%	0%	0%	0%	99.99%	0%	99.99%	0.02%	0.02%	0.01%	
G	99.99%	0%	0%	0%	0%	0%	9.3%	99.97%	99.97%	0%	0%	99.99%	0.01%	0%	99.99%	0%	99.99%	0%	99.99%	0%	
C	0%	99.99%	0.01%	99.99%	0%	0%	0%	0%	99.99%	99.97%	0%	0%	0%	0%	0%	0%	0%	0.02%	99.99%	99.97%	
A	0.01%	0.01%	0%	0%	99.96%	0.01%	0.01%	0.01%	0%	0%	0%	99.97%	99.99%	0%	0%	0%	0%	0.01%	0.02%	0.01%	
#	31978	3192	3175	3204	3208	31942	3167	3155	3175	31675	32176	3197	3192	32126	31970	32441	32024	32465	31634	3164	
		Position along data site (bp)																			
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20

FIG 66E

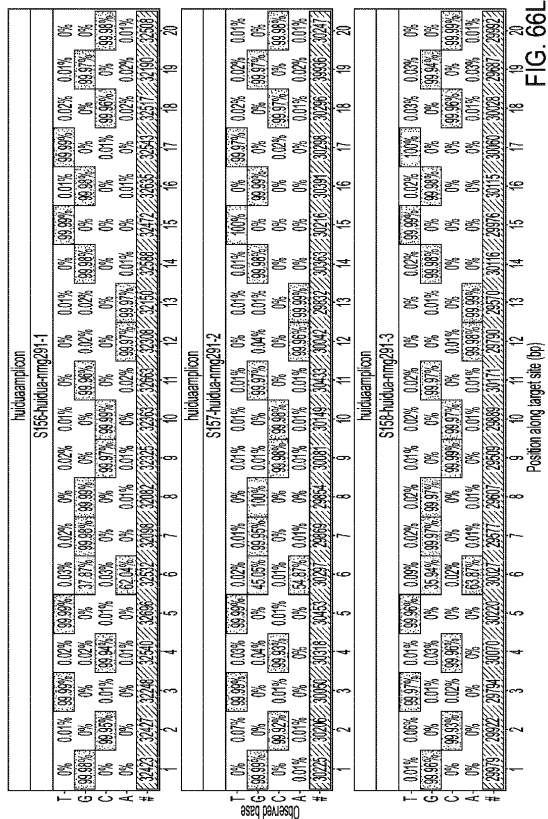
【図 6 6 J】



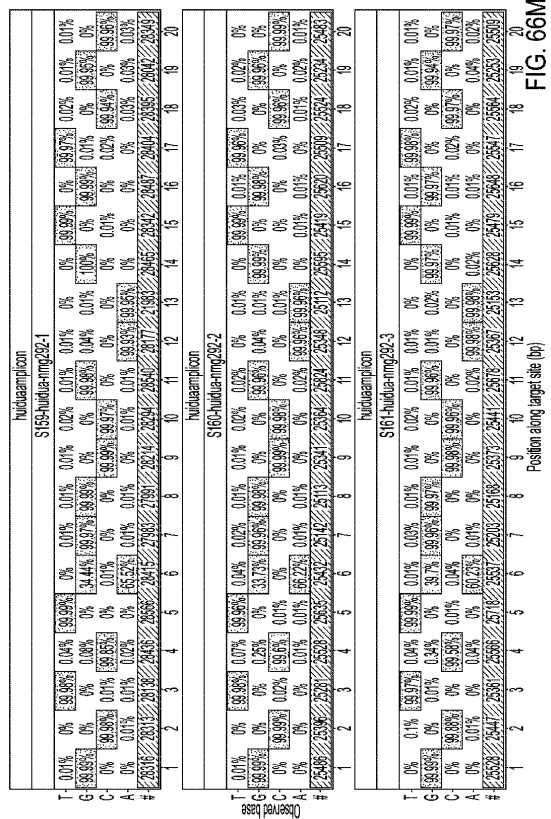
【図 6 6 K】



【図 6 6 L】



【図 6 6 M】



10

20

30

40

50

【図 6 6 N】

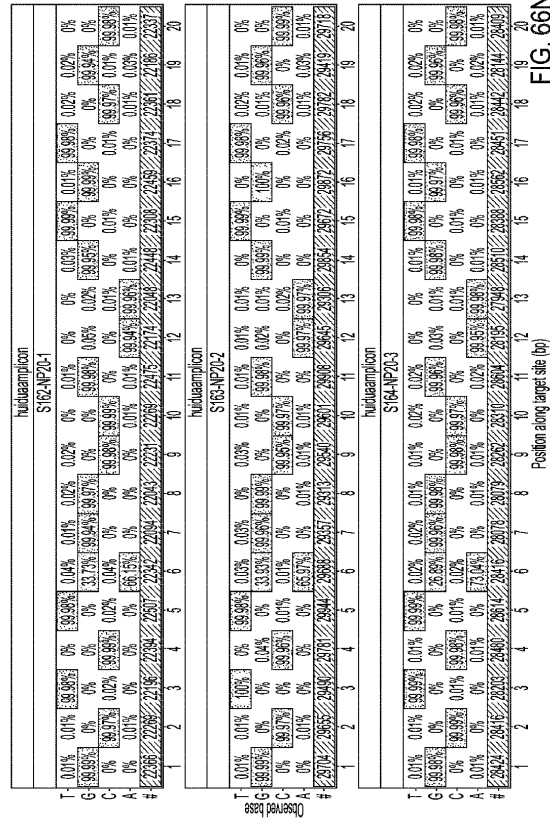


FIG. 66N

【図 6 6 O】

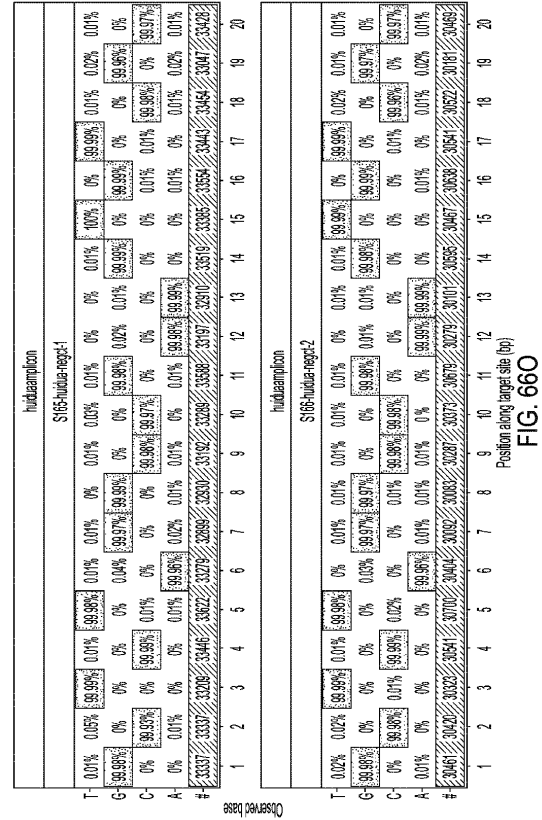


FIG. 66O

【図 6 7】

変異	スパーサー	PAM	レンチHEKスクリーニング
R255X	CITTTCACTTCTCGCGGGG	NGTT	第1:低い編集 第2(最適化):新規 バリエーションで35%に至る
R106W	AGCTTCCATGTCCAGCCTTC	NGG	35%に至る編集、ABE8.9
T158M	ACCATGAAGTCAAAATCATT	NGG	低い標的、高いバスターンダー
R270X	GCTTTTCAGCCCGCTTTCTTG	NGG	編集なし、設計問題?

FIG. 67

【図 6 8 - 1】

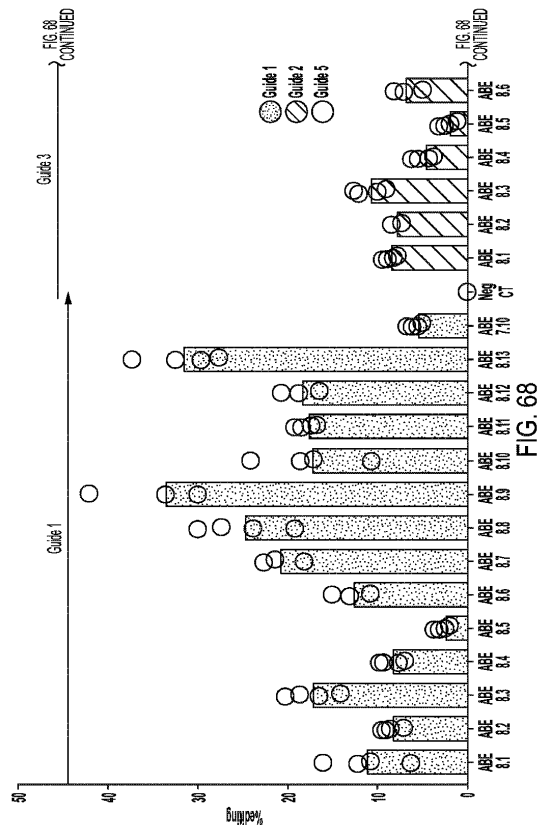


FIG. 68

10

20

30

40

50

【図 7 2 A - 1】

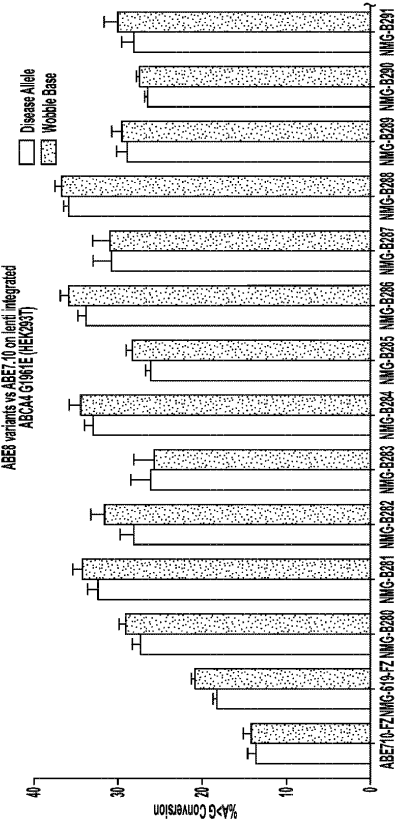


FIG. 72A

【図 7 2 A - 2】

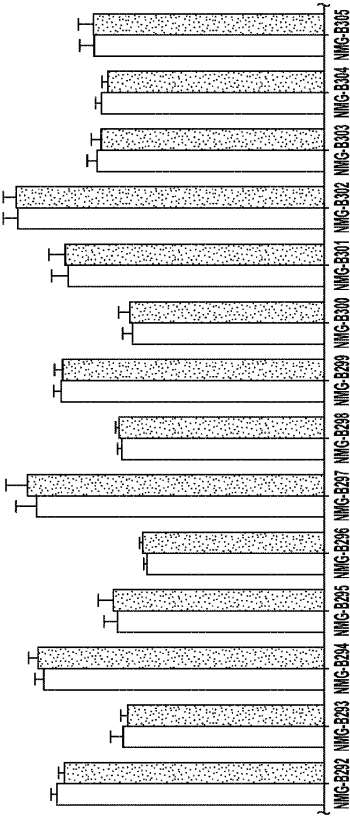


FIG. 72A CONTINUED

【図 7 2 A - 3】

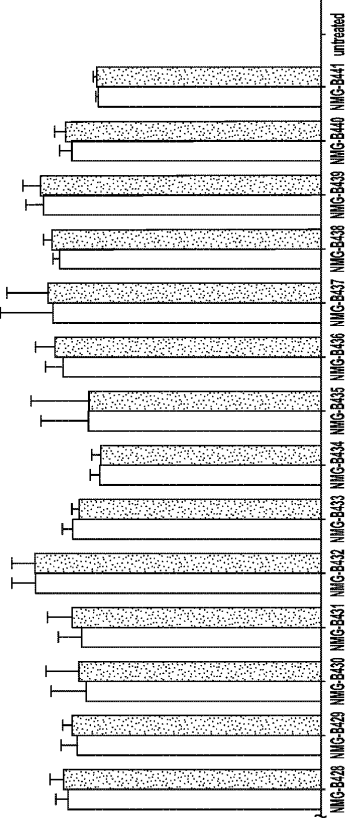


FIG. 72A CONTINUED

【図 7 2 B】

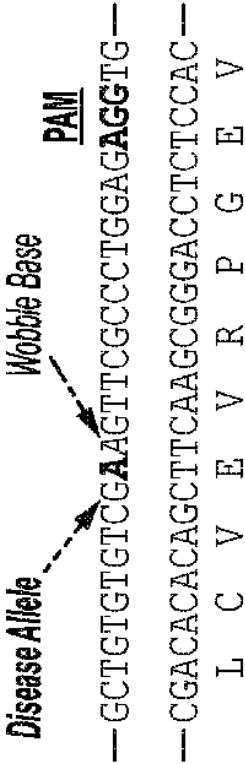


FIG. 72B

10

20

30

40

50

【図 7 3】

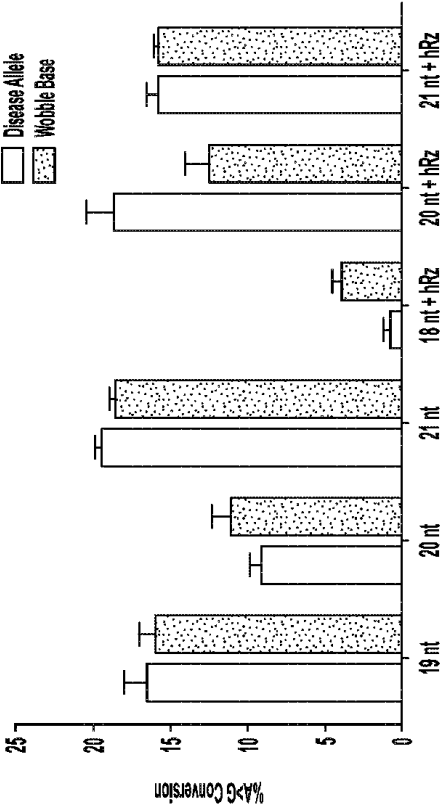


FIG. 73

【図 7 4】

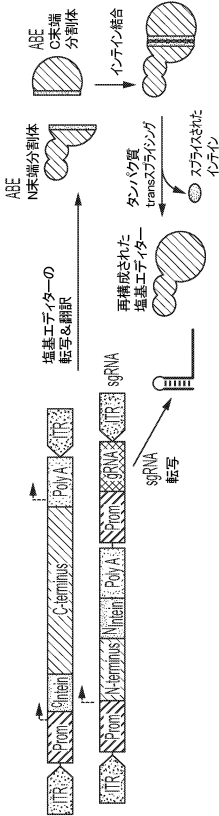


FIG. 74

【図 7 5 A】

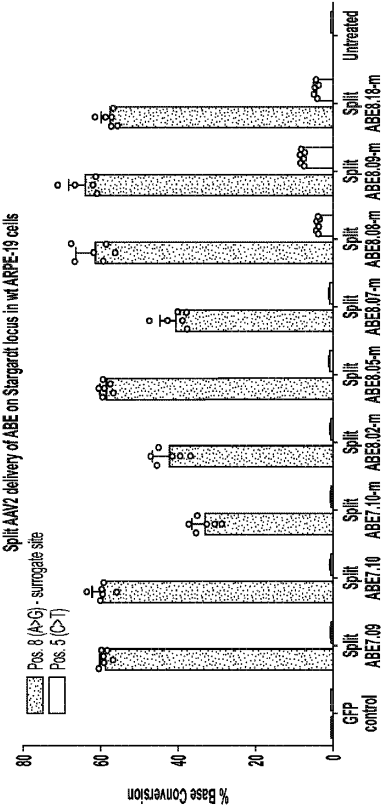


FIG. 75A

【図 7 5 B】

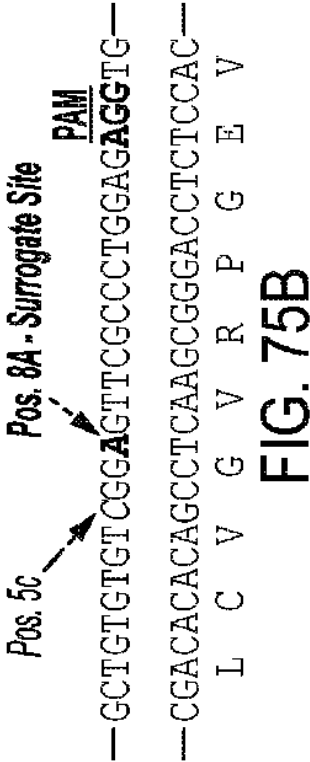
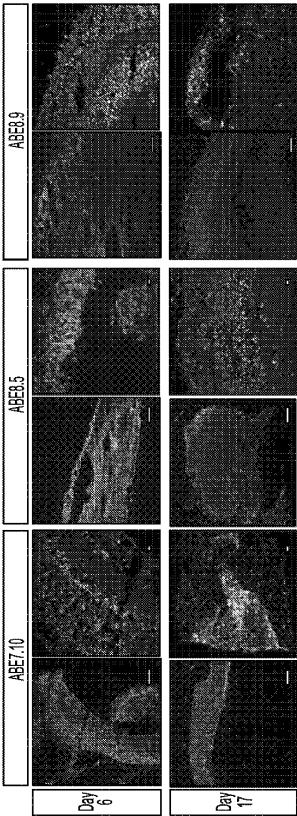


FIG. 75B

【図 79】



Our dual AAV system that delivers split-inteins expresses Cas9

FIG. 79

【配列表】

0007586601000001.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	5/0793(2010.01)	C 1 2 N	5/0793
A 6 1 P	25/00 (2006.01)	A 6 1 P	25/00
A 6 1 P	25/16 (2006.01)	A 6 1 P	25/16
A 6 1 K	48/00 (2006.01)	A 6 1 K	48/00
A 6 1 K	35/76 (2015.01)	A 6 1 K	35/76
A 6 1 K	35/12 (2015.01)	A 6 1 K	35/12
A 6 1 K	35/30 (2015.01)	A 6 1 K	35/30
A 6 1 K	31/712(2006.01)	A 6 1 K	31/712
A 6 1 K	31/7125(2006.01)	A 6 1 K	31/7125
C 1 2 N	15/56 (2006.01)	C 1 2 N	15/56

(33)優先権主張国・地域又は機関

米国(US)

(31)優先権主張番号 62/852,224

(32)優先日 令和1年5月23日(2019.5.23)

(33)優先権主張国・地域又は機関

米国(US)

(31)優先権主張番号 62/873,138

(32)優先日 令和1年7月11日(2019.7.11)

(33)優先権主張国・地域又は機関

米国(US)

(31)優先権主張番号 62/888,867

(32)優先日 令和1年8月19日(2019.8.19)

(33)優先権主張国・地域又は機関

米国(US)

(31)優先権主張番号 62/931,722

(32)優先日 令和1年11月6日(2019.11.6)

(33)優先権主張国・地域又は機関

米国(US)

(31)優先権主張番号 62/941,569

(32)優先日 令和1年11月27日(2019.11.27)

(33)優先権主張国・地域又は機関

米国(US)

(31)優先権主張番号 62/966,526

(32)優先日 令和2年1月27日(2020.1.27)

(33)優先権主張国・地域又は機関

米国(US)

(31)優先権主張番号 62/850,919

(32)優先日 令和1年5月21日(2019.5.21)

(33)優先権主張国・地域又は機関

米国(US)

セッツ州 ケンブリッジ ランズダウン ストリート 26 2nd フロア

(72)発明者 ゴーデッリ、ニコル

アメリカ合衆国 02139 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 26
2nd フロア

(72)発明者 ユ、イ

アメリカ合衆国 02139 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 26
2nd フロア

(72)発明者 ゼッツェ、ベルンド

アメリカ合衆国 02139 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 26

-
- 2nd フロア
- (72)発明者 ボーン、 デイビッド エー.
アメリカ合衆国 02139 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 26
2nd フロア
- (72)発明者 イ、 スノージュ
アメリカ合衆国 02139 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 26
2nd フロア
- (72)発明者 パッカー、 マイケル
アメリカ合衆国 02139 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 26
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 02139 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 26
2nd フロア
- (72)発明者 ペトロシャン、 ナタリー
アメリカ合衆国 02139 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 26
2nd フロア
- (72)発明者 メッサーナ、 アンジェリカ
アメリカ合衆国 02139 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 26
2nd フロア
- (72)発明者 バーコヴィッチ、 シャウナ
アメリカ合衆国 02139 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 26
2nd フロア
- (72)発明者 ブライソン、 デイビッド
アメリカ合衆国 02139 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 26
2nd フロア
- 審査官 榎本 佳予子
- (56)参考文献 国際公開第2018/027078 (WO, A1)
国際公開第2019/005884 (WO, A1)
国際公開第2018/213708 (WO, A1)
- (58)調査した分野 (Int.Cl., DB名)
A61K 31/00-33/44
A61K 48/00
A61P 1/00-43/00
C12N 15/00-15/90
C12N 5/00- 5/28
A61K 35/00-35/768
C A p l u s / R E G I S T R Y / M E D L I N E / E M B A S E / B I O S I S (S T
N)