

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第6799586号
(P6799586)

(45) 発行日 令和2年12月16日 (2020. 12. 16)

(24) 登録日 令和2年11月25日 (2020. 11. 25)

(51) Int. Cl.	F I
C 1 2 N 15/55 (2006. 01)	C 1 2 N 15/55 Z N A
C 1 2 N 9/16 (2006. 01)	C 1 2 N 9/16 Z
C 1 2 N 15/63 (2006. 01)	C 1 2 N 15/63 Z
C O 7 K 19/00 (2006. 01)	C O 7 K 19/00
C 1 2 N 15/62 (2006. 01)	C 1 2 N 15/62 Z

請求項の数 33 (全 88 頁) 最終頁に続く

(21) 出願番号	特願2018-510914 (P2018-510914)	(73) 特許権者	592017633
(86) (22) 出願日	平成28年8月26日 (2016. 8. 26)		ザ ジェネラル ホスピタル コーポレイ ション
(65) 公表番号	特表2018-525019 (P2018-525019A)		アメリカ合衆国 マサチューセッツ州 ボ ストン フルーツ ストリート 55
(43) 公表日	平成30年9月6日 (2018. 9. 6)	(74) 代理人	100092783
(86) 国際出願番号	PCT/US2016/049147		弁理士 小林 浩
(87) 国際公開番号	W02017/040348	(74) 代理人	100120134
(87) 国際公開日	平成29年3月9日 (2017. 3. 9)		弁理士 大森 規雄
審査請求日	令和1年8月26日 (2019. 8. 26)	(74) 代理人	100104282
(31) 優先権主張番号	62/211, 553		弁理士 鈴木 康仁
(32) 優先日	平成27年8月28日 (2015. 8. 28)	(72) 発明者	ジョン, ジェイ, キース
(33) 優先権主張国・地域又は機関	米国 (US)		アメリカ合衆国 マサチューセッツ州 O 1890, ウィンチェスター, マグノリア ウェイ 1
(31) 優先権主張番号	62/216, 033		最終頁に続く
(32) 優先日	平成27年9月9日 (2015. 9. 9)		
(33) 優先権主張国・地域又は機関	米国 (US)		

(54) 【発明の名称】 遺伝子操作CRISPR-Cas9ヌクレアーゼ

(57) 【特許請求の範囲】

【請求項 1】

配列番号2のアミノ酸配列と少なくとも95%同一であり、以下の位置：Y211、Y212、W229、Y230、R245、T392、またはY651のうちの1、2、3、4、5、6つまたはそれ以上における突然変異を有し、野生型SaCas9と比較して改善されたオフターゲット活性を有する、単離黄色ブドウ球菌 (*Staphylococcus aureus*) Cas9 (SaCas9) タンパク質。

【請求項 2】

以下の突然変異：Y211A、Y212A、W229A、Y230A、R245A、T392A、および/またはY651Aの1つ以上を含む、請求項1に記載の単離タンパク質。

【請求項 3】

N419における突然変異をさらに含む、請求項1に記載の単離タンパク質。

【請求項 4】

Y211、Y212、W229、Y230、R245、およびT392のうちの1、2、3、4、5または6つにおける突然変異、ならびに任意選択的にR654における突然変異を含む、請求項1に記載の単離タンパク質。

【請求項 5】

Y211A/R654A、Y211A/Y212A、Y211A/Y230A、Y211A/R245A、Y212A/Y230A、Y212A/R245A、Y230A/R

10

20

2 4 5 A、W 2 2 9 A / R 6 5 4 A、Y 2 1 1 A / Y 2 1 2 A / Y 2 3 0 A、Y 2 1 1 A / Y 2 1 2 A / R 2 4 5 A、Y 2 1 1 A / Y 2 1 2 A / Y 6 5 1 A、Y 2 1 1 A / Y 2 3 0 A / R 2 4 5 A、Y 2 1 1 A / Y 2 3 0 A / Y 6 5 1 A、Y 2 1 1 A / R 2 4 5 A / Y 6 5 1 A、Y 2 1 1 A / R 2 4 5 A / R 6 5 4 A、Y 2 1 1 A / R 2 4 5 A / N 4 1 9 A、Y 2 1 1 A / N 4 1 9 A / R 6 5 4 A、Y 2 1 2 A / Y 2 3 0 A / R 2 4 5 A、Y 2 1 2 A / Y 2 3 0 A / Y 6 5 1 A、Y 2 1 2 A / R 2 4 5 A / Y 6 5 1 A、Y 2 3 0 A / R 2 4 5 A / Y 6 5 1 A、R 2 4 5 A / N 4 1 9 A / R 6 5 4 A、T 3 9 2 A / N 4 1 9 A / R 6 5 4 A、R 2 4 5 A / T 3 9 2 A / N 4 1 9 A / R 6 5 4 A、Y 2 1 1 A / R 2 4 5 A / N 4 1 9 A / R 6 5 4 A、W 2 2 9 A / R 2 4 5 A / N 4 1 9 A / R 6 5 4 A、Y 2 1 1 A / R 2 4 5 A / T 3 9 2 A / N 4 1 9 A / R 6 5 4 A、または Y 2 1 1 A / W 2 2 9 A / R 2 4 5 A / N 4 1 9 A / R 6 5 4 A

10

における突然変異を含む、請求項 3 または 4 に記載の単離タンパク質。

【請求項 6】

N 4 4 ; R 4 5 ; R 5 1 ; R 5 5 ; K 5 7 ; R 5 9 ; R 6 0 ; R 6 1 ; H 1 1 1 ; K 1 1 4 ; R 1 1 6 ; V 1 6 4 ; R 1 6 5 ; N 1 6 9 ; R 2 0 8 ; R 2 0 9 ; T 2 3 8 ; Y 2 3 9 ; K 2 4 8 ; Y 2 5 6 ; R 3 1 4 ; N 3 9 4 ; Q 4 1 4 ; L 4 4 6 ; Q 4 8 8 ; N 4 9 2 ; Q 4 9 5 ; R 4 9 7 ; N 4 9 8 ; R 4 9 9 ; Q 5 0 0 ; K 5 1 8 ; K 5 2 3 ; K 5 2 5 ; H 5 5 7 ; R 5 6 1 ; K 5 7 2 ; R 6 3 4 ; G 6 5 5 ; N 6 5 8 ; S 6 6 2 ; N 6 6 7 ; R 6 8 6 ; K 6 9 2 ; R 6 9 4 ; H 7 0 0 ; K 7 5 1 ; D 7 8 6 ; T 7 8 7 ; L 7 8 8 ; Y 7 8 9 ; S 7 9 0 ; R 7 9 2 ; N 8 0 4 ; Y 8 6 8 ; K 8 7 0 ; K 8 7 8 ; K 8 7 9 ; K 8 8 1 ; T 8 8 2 ; K 8 8 6 ; Y 8 9 7 ; N 8 8 8 ; A 8 8 9 ; R 9 0 1 ; K 9 0 6 ; L 9 0 9 ; N 9 8 5 ; N 9 8 6 ; R 9 9 1 ; および / または R 1 0 1 5 における突然変異をさらに含む、請求項 1 ~ 5 のいずれか一項に記載の単離タンパク質。

20

【請求項 7】

以下の突然変異：E 7 8 2 K ; K 9 2 9 R ; N 9 6 8 K ; R 1 0 1 5 H ; E 7 8 2 K / N 9 6 8 K / R 1 0 1 5 H (K K H バリエント) ; E 7 8 2 K / K 9 2 9 R / R 1 0 1 5 H (K R H バリエント) ; または E 7 8 2 K / K 9 2 9 R / N 9 6 8 K / R 1 0 1 5 H (K R K H バリエント) のうちの 1 つ以上をさらに含む、請求項 1 ~ 5 のいずれか一項に記載の単離タンパク質。

【請求項 8】

30

ヌクレアーゼ活性を減少させる突然変異である、H 5 5 7 または N 5 8 0 における突然変異、および D 1 0、E 4 7 7、D 5 5 6、H 7 0 1 または D 7 0 4 における突然変異をさらに含む、請求項 1 ~ 5 のいずれか一項に記載の単離タンパク質。

【請求項 9】

前記突然変異が、
D 1 0 における突然変異が D 1 0 A もしくは D 1 0 N であるか、
D 5 5 6 における突然変異が D 5 5 6 A であるか、
H 5 5 7 における突然変異が H 5 5 7 A、H 5 5 7 N、もしくは H 5 5 7 Y であるか、または
N 5 8 0 における突然変異が N 5 8 0 A である、
請求項 8 に記載の単離タンパク質。

40

【請求項 10】

前記 S a C a s 9 タンパク質が、核局在化配列、細胞浸透ペプチド配列および / または親和性タグのうちの 1 つ以上と融合されている、請求項 1 ~ 9 のいずれか一項に記載の単離タンパク質。

【請求項 11】

任意選択の介在リンカーにより異種機能ドメインに融合されている請求項 1 ~ 10 のいずれか一項に記載の単離タンパク質を含む融合タンパク質であって、前記リンカーが前記融合タンパク質の活性を妨害しない、融合タンパク質。

【請求項 12】

50

前記異種機能ドメインが転写活性化ドメインである、請求項 1 1 に記載の融合タンパク質。

【請求項 1 3】

前記転写活性化ドメインが V P 6 4 または N F - B p 6 5 からのものである、請求項 1 2 に記載の融合タンパク質。

【請求項 1 4】

前記異種機能ドメインが転写サイレンサーまたは転写抑制ドメインである、請求項 1 1 に記載の融合タンパク質。

【請求項 1 5】

前記転写抑制ドメインが、クルップエル関連ボックス (K R A B) ドメイン、 E R F リプレッサードメイン (E R D)、または m S i n 3 A 相互作用ドメイン (S I D) である、請求項 1 4 に記載の融合タンパク質。

10

【請求項 1 6】

前記転写サイレンサーがヘテロクロマチンタンパク質 1 (H P 1) である、請求項 1 4 に記載の融合タンパク質。

【請求項 1 7】

前記 H P 1 が H P 1 または H P 1 である、請求項 1 6 に記載の融合タンパク質。

【請求項 1 8】

前記異種機能ドメインが、DNA のメチル化状態を改変する酵素である、請求項 1 1 に記載の融合タンパク質。

20

【請求項 1 9】

前記 DNA のメチル化状態を改変する前記酵素が DNA メチルトランスフェラーゼ (D N M T) またはテンイレブントランスロケーション (T E T) タンパク質である、請求項 1 8 に記載の融合タンパク質。

【請求項 2 0】

前記 T E T タンパク質が T E T 1 である、請求項 1 9 に記載の融合タンパク質。

【請求項 2 1】

前記異種機能ドメインが、ヒストンサブユニットを改変する酵素である、請求項 1 1 に記載の融合タンパク質。

【請求項 2 2】

30

ヒストンサブユニットを改変する前記酵素が、ヒストンアセチルトランスフェラーゼ (H A T)、ヒストンデアセチラーゼ (H D A C)、ヒストンメチルトランスフェラーゼ (H M T)、またはヒストンデメチラーゼである、請求項 2 1 に記載の融合タンパク質。

【請求項 2 3】

前記異種機能ドメインが生物学的係留物である、請求項 1 1 に記載の融合タンパク質。

【請求項 2 4】

前記生物学的係留物が M S 2、C s y 4 またはラムダ N タンパク質である、請求項 2 3 に記載の融合タンパク質。

【請求項 2 5】

前記異種機能ドメインが F o k I である、請求項 1 1 に記載の融合タンパク質。

40

【請求項 2 6】

請求項 1 ~ 1 0 のいずれか一項に記載のタンパク質をコードする単離核酸。

【請求項 2 7】

請求項 2 6 に記載の単離核酸を含み、任意選択的に前記単離核酸が 1 つ以上の調節ドメインに作動可能に結合されている、ベクター。

【請求項 2 8】

請求項 2 6 に記載の核酸を含む宿主細胞。

【請求項 2 9】

細胞のゲノムまたはエピゲノムを変更するインピトロの方法であって、前記細胞中で、請求項 1 ~ 1 0 のいずれか一項に記載の単離タンパク質および前記細胞の前記ゲノムの選

50

択部分に相補的な領域を有するガイドRNAを発現させるか、または前記単離タンパク質および前記ガイドRNAと前記細胞とを接触させることを含む方法。

【請求項30】

細胞のゲノムまたはエピゲノムを変更するインビトロの方法であって、前記細胞中で、請求項11～25のいずれか一項に記載の融合タンパク質および前記細胞の前記ゲノムの選択部分に相補的な領域を有するガイドRNAを発現させるか、または前記融合タンパク質および前記ガイドRNAと前記細胞とを接触させることを含む方法。

【請求項31】

二本鎖DNA(dsDNA)分子を変更するインビトロの方法であって、前記dsDNA分子を請求項1～10のいずれか一項に記載の単離タンパク質および前記dsDNA分子の選択部分に相補的な領域を有するガイドRNAと接触させることを含む方法。

10

【請求項32】

前記dsDNA分子がインビトロで存在する、請求項31に記載のインビトロの方法。

【請求項33】

二本鎖DNA(dsDNA)分子を変更するインビトロの方法であって、前記dsDNA分子を請求項11～25のいずれか一項に記載の融合タンパク質および前記dsDNA分子の選択部分に相補的な領域を有するガイドRNAと接触させることを含む方法。

【発明の詳細な説明】

【技術分野】

【0001】

20

優先権の主張

本出願は、35 USC § 119(e)のもと、2015年8月28日に出願された米国特許出願第62/211,553号明細書；2015年9月9日に出願された同第62/216,033号明細書；2015年11月20日に出願された同第62/258,280号明細書；2015年12月28日に出願された同第62/271,938号明細書；および2016年2月4日に出願された米国特許出願公開第15/015,947号明細書に対する優先権を主張する。上記の全体の内容は、参照により本明細書に組み込まれる。

【0002】

配列表

30

本出願は、ASCIIフォーマットで電子的に提出され、参照により全体として本明細書に組み込まれる配列表を含有する。前記ASCII複写物は、2016年8月26日に作成され、SEQUENCE LISTING.txtと命名され、サイズが129,955バイトである。

【0003】

連邦政府による資金提供を受けた研究または開発

本発明は、National Institutes of Healthにより授与された助成金番号DP1GM105378、およびR01GM088040のもとで政府支援により作製された。政府は、本発明における一定の権利を有する。

【0004】

40

本発明は、少なくとも部分的に、変更および改善された標的特異性を有する遺伝子操作クラスター化等間隔短鎖回文リピート(Clustered Regularly Interspaced Short Palindromic Repeats)(CRISPR)/CRISPR関連タンパク質9(Cas9)ヌクレアーゼならびにゲノム遺伝子操作、エピゲノミック遺伝子操作、ゲノム標的化、ゲノム編集、およびインビトロ診断におけるその使用に関する。

【背景技術】

【0005】

CRISPR-Cas9ヌクレアーゼは、広範な生物および細胞タイプにおいて効率的で効率的なゲノム編集を可能とする(Sander & Joung, Nat Biotech

50

h n o l 3 2 , 3 4 7 - 3 5 5 (2 0 1 4) ; H s u e t a l . , C e l l 1 5
7 , 1 2 6 2 - 1 2 7 8 (2 0 1 4) ; D o u d n a & C h a r p e n t i e r , S c i
e n c e 3 4 6 , 1 2 5 8 0 9 6 (2 0 1 4) ; B a r r a n g o u & M a y , E x p
e r t O p i n B i o l T h e r 1 5 , 3 1 1 - 3 1 4 (2 0 1 5)) 。 C a s
9 による標的部位認識は、標的プロトスペーサーに相補的な配列をコードするキメラ単一
ガイドRNA (s g RNA) によりプログラミングされるが (J i n e k e t a l .
, S c i e n c e 3 3 7 , 8 1 6 - 8 2 1 (2 0 1 2)) 、短鎖隣接PAMの認識も要
求する (M o j i c a e t a l . , M i c r o b i o l o g y 1 5 5 , 7 3 3 - 7
4 0 (2 0 0 9) ; S h a h e t a l . , R N A B i o l 1 0 , 8 9 1 - 8 9 9
(2 0 1 3) ; J i a n g e t a l . , N a t B i o t e c h n o l 3 1 , 2 3
3 - 2 3 9 (2 0 1 3) ; J i n e k e t a l . , S c i e n c e 3 3 7 , 8 1 6
- 8 2 1 (2 0 1 2) ; S t e r n b e r g e t a l . , N a t u r e 5 0 7 , 6
2 - 6 7 (2 0 1 4)) 。

10

【発明の概要】

【課題を解決するための手段】

【0006】

本明細書に記載のとおり、C a s 9 タンパク質は、理論的には、DNA についての C a
s 9 の結合親和性を低減させることにより、増加した特異性を示すように遺伝子操作する
ことができる。したがって、野生型タンパク質と比較して増加した特異性を有する (すな
わち、不完全マッチまたはミスマッチDNA 部位において実質的により少ないオフターゲ
ット効果を誘導する) 多数の C a s 9 バリエーションおよびそれらを使用する方法が本明細書
に記載される。

20

【0007】

第1の態様において、本発明は、以下の位置 : L 1 6 9 、 Y 4 5 0 、 N 4 9 7 、 R 6 6
1 、 Q 6 9 5 、 Q 9 2 6 、および / または D 1 1 3 5 E の 1 、 2 、 3 、 4 、 5 、 6 、または
7 つ全てにおける突然変異を有し、例えば、以下の位置 : L 1 6 9 A 、 Y 4 5 0 、 N 4
9 7 、 R 6 6 1 、 Q 6 9 5 、 Q 9 2 6 、 D 1 1 3 5 E の 1 、 2 、 3 、 4 、 5 、 6 または 7
つにおける突然変異を有する配列番号1のアミノ酸配列と少なくとも80% 同一である配
列ならびに任意選択的に核局在化配列、細胞浸透ペプチド配列、および / または親和性タ
グの1つ以上を含む単離化膿性連鎖球菌 (S t r e p t o c o c c u s p y o g e n e
s) C a s 9 (S p C a s 9) タンパク質を提供する。突然変異は、アミノ酸を天然アミ
ノ酸以外のアミノ酸に変更する (例えば、497 は、N 以外のいずれかである) 。好まし
い実施形態において、突然変異は、アミノ酸を天然のもの、アルギニンまたはリジン以外
の任意のアミノ酸に変化させ ; 一部の実施形態において、アミノ酸は、アラニンである。

30

【0008】

一部の実施形態において、バリエーション S p C a s 9 タンパク質は、以下 : N 4 9 7 、 R
6 6 1 、 Q 6 9 5 、および Q 9 2 6 の 1 、 2 、 3 、または 4 つ全てにおける突然変異、例
えば、以下の突然変異 : N 4 9 7 A 、 R 6 6 1 A 、 Q 6 9 5 A 、および Q 9 2 6 A の 1 、
2 、 3 、または 4 つ全てを含む。

【0009】

一部の実施形態において、バリエーション S p C a s 9 タンパク質は、Q 6 9 5 および / ま
たは Q 9 2 6 ならびに任意選択的に L 1 6 9 、 Y 4 5 0 、 N 4 9 7 、 R 6 6 1 、および D
1 1 3 5 E の 1 、 2 、 3 、 4 、または 5 つ全てにおける突然変異、例えば、例として、限
定されるものではないが、Y 4 5 0 A / Q 6 9 5 A 、 L 1 6 9 A / Q 6 9 5 A 、 Q 6 9 5
A / Q 9 2 6 A 、 Q 6 9 5 A / D 1 1 3 5 E 、 Q 9 2 6 A / D 1 1 3 5 E 、 Y 4 5 0 A /
D 1 1 3 5 E 、 L 1 6 9 A / Y 4 5 0 A / Q 6 9 5 A 、 L 1 6 9 A / Q 6 9 5 A / Q 9 2
6 A 、 Y 4 5 0 A / Q 6 9 5 A / Q 9 2 6 A 、 R 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A 、 N
4 9 7 A / Q 6 9 5 A / Q 9 2 6 A 、 Y 4 5 0 A / Q 6 9 5 A / D 1 1 3 5 E 、 Y 4 5 0
A / Q 9 2 6 A / D 1 1 3 5 E 、 Q 6 9 5 A / Q 9 2 6 A / D 1 1 3 5 E 、 L 1 6 9 A /
Y 4 5 0 A / Q 6 9 5 A / Q 9 2 6 A 、 L 1 6 9 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6

40

50

A, Y450A/R661A/Q695A/Q926A、N497A/Q695A/Q926A/D1135E、R661A/Q695A/Q926A/D1135E、およびY450A/Q695A/Q926A/D1135Eを含む。

【0010】

一部の実施形態において、バリエーションタンパク質は、N14;S15;S55;R63;R78;H160;K163;R165;L169;R403;N407;Y450;M495;N497;K510;Y515;W659;R661;M694;Q695;H698;A728;S730;K775;S777;R778;R780;K782;R783;K789;K797;Q805;N808;K810;R832;Q844;S845;K848;S851;K855;R859;K862;K890;Q920;Q926;K961;S964;K968;K974;R976;N980;H982;K1003;K1014;S1040;N1041;N1044;K1047;K1059;R1060;K1107;E1108;S1109;K1113;R1114;S1116;K1118;D1135;S1136;K1153;K1155;K1158;K1200;Q1221;H1241;Q1254;Q1256;K1289;K1296;K1297;R1298;K1300;H1311;K1325;K1334;T1337および/またはS1216における突然変異を含む。

【0011】

一部の実施形態において、バリエーションタンパク質は、以下の突然変異：N14A;S15A;S55A;R63A;R78A;R165A;R403A;N407A;N497A;Y450A;K510A;Y515A;R661A;Q695A;S730A;K775A;S777A;R778A;R780A;K782A;R783A;K789A;K797A;Q805A;N808A;K810A;R832A;Q844A;S845A;K848A;S851A;K855A;R859A;K862A;K890A;Q920A;Q926A;K961A;S964A;K968A;K974A;R976A;N980A;H982A;K1003A;K1014A;S1040A;N1041A;N1044A;K1047A;K1059A;R1060A;K1107A;E1108A;S1109A;K1113A;R1114A;S1116A;K1118A;D1135A;S1136A;K1153A;K1155A;K1158A;K1200A;Q1221A;H1241A;Q1254A;Q1256A;K1289A;K1296A;K1297A;R1298A;K1300A;H1311A;K1325A;K1334A;T1337Aおよび/またはS1216Aの1つ以上も含む。一部の実施形態において、バリエーションタンパク質は、HF1(N497A/R661A/Q695A/Q926A)+K810A、HF1+K848A、HF1+K855A、HF1+H982A、HF1+K848A/K1003A、HF1+K848A/R1060A、HF1+K855A/K1003A、HF1+K855A/R1060A、HF1+H982A/K1003A、HF1+H982A/R1060A、HF1+K1003A/R1060A、HF1+K810A/K1003A/R1060A、HF1+K848A/K1003A/R1060Aを含む。一部の実施形態において、バリエーションタンパク質は、HF1+K848A/K1003A、HF1+K848A/R1060A、HF1+K855A/K1003A、HF1+K855A/R1060A、HF1+K1003A/R1060A、HF1+K848A/K1003A/R1060Aを含む。一部の実施形態において、バリエーションタンパク質は、Q695A/Q926A/R780A、Q695A/Q926A/R976A、Q695A/Q926A/H982A、Q695A/Q926A/K855A、Q695A/Q926A/K848A/K1003A、Q695A/Q926A/K848A/K855A、Q695A/Q926A/K848A/H982A、Q695A/Q926A/K1003A/R1060A、Q695A/Q926A/K848A/R1060A、Q695A/Q926A/K855A/H982A、Q695A/Q926A/K855A/K1003A、Q695A/Q926A/K855A/R1060A、Q695A/Q926A/H982A/K1003A、Q695A/Q

10

20

30

40

50

9 2 6 A / H 9 8 2 A / R 1 0 6 0 A、Q 6 9 5 A / Q 9 2 6 A / K 1 0 0 3 A / R 1 0 6 0 A、Q 6 9 5 A / Q 9 2 6 A / K 8 1 0 A / K 1 0 0 3 A / R 1 0 6 0 A、Q 6 9 5 A / Q 9 2 6 A / K 8 4 8 A / K 1 0 0 3 A / R 1 0 6 0 Aを含む。一部の実施形態において、バリエーションは、N 4 9 7 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A / K 8 1 0 A、N 4 9 7 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A / K 8 4 8 A、N 4 9 7 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A / K 8 5 5 A、N 4 9 7 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A / R 7 8 0 A、N 4 9 7 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A / K 9 6 8 A、N 4 9 7 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A / H 9 8 2 A、N 4 9 7 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A / K 1 0 0 3 A、N 4 9 7 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A / K 1 0 1 4 A、N 4 9 7 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A / K 1 0 4 7 A、N 4 9 7 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A / R 1 0 6 0 A、N 4 9 7 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A / K 8 1 0 A / K 9 6 8 A、N 4 9 7 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A / K 8 1 0 A / K 8 4 8 A、N 4 9 7 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A / K 8 1 0 A / K 1 0 0 3 A、N 4 9 7 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A / K 8 1 0 A / R 1 0 6 0 A、N 4 9 7 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A / K 8 4 8 A / K 1 0 0 3 A、N 4 9 7 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A / K 8 5 5 A / K 1 0 0 3 A、N 4 9 7 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A / K 8 5 5 A / R 1 0 6 0 A、N 4 9 7 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A / K 9 6 8 A / K 1 0 0 3 A、N 4 9 7 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A / H 9 8 2 A / K 1 0 0 3 A、N 4 9 7 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A / H 9 8 2 A / R 1 0 6 0 A、N 4 9 7 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A / K 1 0 0 3 A / R 1 0 6 0 A、N 4 9 7 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A / K 8 1 0 A / K 1 0 0 3 A / R 1 0 6 0 A、N 4 9 7 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A / K 8 4 8 A / K 1 0 0 3 A / R 1 0 6 0 A、Q 6 9 5 A / Q 9 2 6 A / R 7 8 0 A、Q 6 9 5 A / Q 9 2 6 A / K 8 1 0 A、Q 6 9 5 A / Q 9 2 6 A / R 8 3 2 A、Q 6 9 5 A / Q 9 2 6 A / K 8 4 8 A、Q 6 9 5 A / Q 9 2 6 A / K 8 5 5 A、Q 6 9 5 A / Q 9 2 6 A / K 9 6 8 A、Q 6 9 5 A / Q 9 2 6 A / R 9 7 6 A、Q 6 9 5 A / Q 9 2 6 A / H 9 8 2 A、Q 6 9 5 A / Q 9 2 6 A / K 1 0 0 3 A、Q 6 9 5 A / Q 9 2 6 A / K 1 0 1 4 A、Q 6 9 5 A / Q 9 2 6 A / K 1 0 4 7 A、Q 6 9 5 A / Q 9 2 6 A / R 1 0 6 0 A、Q 6 9 5 A / Q 9 2 6 A / K 8 4 8 A / K 9 6 8 A、Q 6 9 5 A / Q 9 2 6 A / R 9 7 6 A、Q 6 9 5 A / Q 9 2 6 A / H 9 8 2 A、Q 6 9 5 A / Q 9 2 6 A / K 8 5 5 A、Q 6 9 5 A / Q 9 2 6 A / K 8 4 8 A / K 1 0 0 3 A、Q 6 9 5 A / Q 9 2 6 A / K 8 4 8 A / K 8 5 5 A、Q 6 9 5 A / Q 9 2 6 A / K 8 4 8 A / H 9 8 2 A、Q 6 9 5 A / Q 9 2 6 A / K 1 0 0 3 A / R 1 0 6 0 A、Q 6 9 5 A / Q 9 2 6 A / R 8 3 2 A / R 1 0 6 0 A、Q 6 9 5 A / Q 9 2 6 A / K 9 6 8 A / K 1 0 0 3 A、Q 6 9 5 A / Q 9 2 6 A / K 9 6 8 A / R 1 0 6 0 A、Q 6 9 5 A / Q 9 2 6 A / K 8 4 8 A / R 1 0 6 0 A、Q 6 9 5 A / Q 9 2 6 A / K 8 5 5 A / H 9 8 2 A、Q 6 9 5 A / Q 9 2 6 A / K 8 5 5 A / K 1 0 0 3 A、Q 6 9 5 A / Q 9 2 6 A / K 8 5 5 A / R 1 0 6 0 A、Q 6 9 5 A / Q 9 2 6 A / H 9 8 2 A / K 1 0 0 3 A、Q 6 9 5 A / Q 9 2 6 A / H 9 8 2 A / R 1 0 6 0 A、Q 6 9 5 A / Q 9 2 6 A / K 1 0 0 3 A / R 1 0 6 0 A、Q 6 9 5 A / Q 9 2 6 A / K 8 1 0 A / K 1 0 0 3 A / R 1 0 6 0 A、Q 6 9 5 A / Q 9 2 6 A / K 1 0 0 3 A / K 1 0 4 7 A / R 1 0 6 0 A、Q 6 9 5 A / Q 9 2 6 A / K 9 6 8 A / K 1 0 0 3 A / R 1 0 6 0 A、Q 6 9 5 A / Q 9 2 6 A / R 8 3 2 A / K 1 0 0 3 A / R 1 0 6 0 A、またはQ 6 9 5 A / Q 9 2 6 A / K 8 4 8 A / K 1 0 0 3 A / R 1 0 6 0 Aを含む。

【 0 0 1 2 】

アラニン以外のアミノ酸の突然変異も含まれ、本方法および組成物においてそれを作製および使用することができる。

【 0 0 1 3 】

一部の実施形態において、バリエーション S p C a s 9 タンパク質は、以下の追加の突然変

10

20

30

40

50

異：R 6 3 A、R 6 6 A、R 6 9 A、R 7 0 A、R 7 1 A、Y 7 2 A、R 7 4 A、R 7 5 A、K 7 6 A、N 7 7 A、R 7 8 A、R 1 1 5 A、H 1 6 0 A、K 1 6 3 A、R 1 6 5 A、L 1 6 9 A、R 4 0 3 A、T 4 0 4 A、F 4 0 5 A、N 4 0 7 A、R 4 4 7 A、N 4 9 7 A、I 4 4 8 A、Y 4 5 0 A、S 4 6 0 A、M 4 9 5 A、K 5 1 0 A、Y 5 1 5 A、R 6 6 1 A、M 6 9 4 A、Q 6 9 5 A、H 6 9 8 A、Y 1 0 1 3 A、V 1 0 1 5 A、R 1 1 2 2 A、K 1 1 2 3 A、K 1 1 2 4 A、K 1 1 5 8 A、K 1 1 8 5 A、K 1 2 0 0 A、S 1 2 1 6 A、Q 1 2 2 1 A、K 1 2 8 9 A、R 1 2 9 8 A、K 1 3 0 0 A、K 1 3 2 5 A、R 1 3 3 3 A、K 1 3 3 4 A、R 1 3 3 5 A、および T 1 3 3 7 A の 1 つ以上を含む。

【 0 0 1 4 】

一部の実施形態において、バリエーション SpCas9 タンパク質は、複数の置換突然変異：N 4 9 7 / R 6 6 1 / Q 6 9 5 / Q 9 2 6 (四重バリエーション突然変異体) ; Q 6 9 5 / Q 9 2 6 (二重突然変異体) ; R 6 6 1 / Q 6 9 5 / Q 9 2 6 および N 4 9 7 / Q 6 9 5 / Q 9 2 6 (三重突然変異体) を含む。一部の実施形態において、L 1 6 9、Y 4 5 0 および / または D 1 1 3 5 における追加の置換突然変異をそれらの二重、三重、および四重突然変異体に追加することができ、または Q 6 9 5 もしくは Q 9 2 6 における置換を担持する単一突然変異体に追加することができる。一部の実施形態において、突然変異体は、野生型アミノ酸に代えてアラニンを含む。一部の実施形態において、突然変異体は、アルギニンまたはリジン (または天然アミノ酸) 以外の任意のアミノ酸を含む。

【 0 0 1 5 】

一部の実施形態において、バリエーション SpCas9 タンパク質は、D 1 0、E 7 6 2、D 8 3 9、H 9 8 3、または D 9 8 6 ; および H 8 4 0 または N 8 6 3 における突然変異からなる群から選択される、ヌクレアーゼ活性を減少させる 1 つ以上の突然変異を含む。一部の実施形態において、突然変異は、(i) D 1 0 A または D 1 0 N、および (i i) H 8 4 0 A、H 8 4 0 N、または H 8 4 0 Y である。

【 0 0 1 6 】

一部の実施形態において、SpCas9 バリエーションは、以下の突然変異のセット：D 1 1 3 5 V / R 1 3 3 5 Q / T 1 3 3 7 R (VQR バリエーション) ; D 1 1 3 5 E / R 1 3 3 5 Q / T 1 3 3 7 R (EQR バリエーション) ; D 1 1 3 5 V / G 1 2 1 8 R / R 1 3 3 5 Q / T 1 3 3 7 R (VRQR バリエーション) ; または D 1 1 3 5 V / G 1 2 1 8 R / R 1 3 3 5 E / T 1 3 3 7 R (VREER バリエーション) の 1 つも含み得る。

【 0 0 1 7 】

本明細書において、以下の位置：Y 2 1 1、Y 2 1 2、W 2 2 9、Y 2 3 0、R 2 4 5、T 3 9 2、N 4 1 9、Y 6 5 1、または R 6 5 4 の 1、2、3、4、5、6 つ以上における突然変異を有し、例えば、以下の位置：Y 2 1 1、Y 2 1 2、W 2 2 9、Y 2 3 0、R 2 4 5、T 3 9 2、N 4 1 9、Y 6 5 1、または R 6 5 4 の 1、2、3、4、もしくは 5、または 6 つ以上における突然変異を有する配列番号 2 のアミノ酸配列と少なくとも 80 % 同一である配列ならびに任意選択的に核局在化配列、細胞浸透ペプチド配列、および / または親和性タグの 1 つ以上を含む単離黄色ブドウ球菌 (*Staphylococcus aureus*) Cas9 (SaCas9) タンパク質も提供される。一部の実施形態において、本明細書に記載の SaCas9 バリエーションは、以下の位置：Y 2 1 1、Y 2 1 2、W 2 2 9、Y 2 3 0、R 2 4 5、T 3 9 2、N 4 1 9、Y 6 5 1 および / または R 6 5 4 の 1、2、3、4、5、6 つ以上における突然変異を有する配列番号 2 のアミノ酸配列を含む。一部の実施形態において、バリエーションは、以下の突然変異：Y 2 1 1 A、Y 2 1 2 A、W 2 2 9、Y 2 3 0 A、R 2 4 5 A、T 3 9 2 A、N 4 1 9 A、Y 6 5 1、および / または R 6 5 4 A の 1 つ以上を含む。

【 0 0 1 8 】

一部の実施形態において、バリエーション SaCas9 タンパク質は、N 4 1 9 および / ま

10

20

30

40

50

たはR 6 5 4における突然変異ならびに任意選択的に追加の突然変異Y 2 1 1、Y 2 1 2、W 2 2 9、Y 2 3 0、R 2 4 5およびT 3 9 2の1、2、3、4つ以上、好ましくは、N 4 1 9 A / R 6 5 4 A、Y 2 1 1 A / R 6 5 4 A、Y 2 1 1 A / Y 2 1 2 A、Y 2 1 1 A / Y 2 3 0 A、Y 2 1 1 A / R 2 4 5 A、Y 2 1 2 A / Y 2 3 0 A、Y 2 1 2 A / R 2 4 5 A、Y 2 3 0 A / R 2 4 5 A、W 2 2 9 A / R 6 5 4 A、Y 2 1 1 A / Y 2 1 2 A / Y 2 3 0 A、Y 2 1 1 A / Y 2 1 2 A / R 2 4 5 A、Y 2 1 1 A / Y 2 1 2 A / Y 6 5 1 A、Y 2 1 1 A / Y 2 3 0 A / R 2 4 5 A、Y 2 1 1 A / Y 2 3 0 A / Y 6 5 1 A、Y 2 1 1 A / R 2 4 5 A / Y 6 5 1 A、Y 2 1 1 A / R 2 4 5 A / R 6 5 4 A、Y 2 1 1 A / R 2 4 5 A / N 4 1 9 A、Y 2 1 1 A / N 4 1 9 A / R 6 5 4 A、Y 2 1 2 A / Y 2 3 0 A / R 2 4 5 A、Y 2 1 2 A / Y 2 3 0 A / Y 6 5 1 A、Y 2 1 2 A / R 2 4 5 A / Y 6 5 1 A、Y 2 3 0 A / R 2 4 5 A / Y 6 5 1 A、R 2 4 5 A / N 4 1 9 A / R 6 5 4 A、T 3 9 2 A / N 4 1 9 A / R 6 5 4 A、R 2 4 5 A / T 3 9 2 A / N 4 1 9 A / R 6 5 4 A、Y 2 1 1 A / R 2 4 5 A / N 4 1 9 A / R 6 5 4 A、W 2 2 9 A / R 2 4 5 A / N 4 1 9 A / R 6 5 4 A、Y 2 1 1 A / R 2 4 5 A / T 3 9 2 A / N 4 1 9 A / R 6 5 4 A、またはY 2 1 1 A / W 2 2 9 A / R 2 4 5 A / N 4 1 9 A / R 6 5 4 Aを含む。

【0019】

一部の実施形態において、バリエーションS a C a s 9タンパク質は、Y 2 1 1 ; Y 2 1 2 ; W 2 2 9 ; Y 2 3 0 ; R 2 4 5 ; T 3 9 2 ; N 4 1 9 ; L 4 4 6 ; Q 4 8 8 ; N 4 9 2 ; Q 4 9 5 ; R 4 9 7 ; N 4 9 8 ; R 4 9 9 ; Q 5 0 0 ; K 5 1 8 ; K 5 2 3 ; K 5 2 5 ; H 5 5 7 ; R 5 6 1 ; K 5 7 2 ; R 6 3 4 ; Y 6 5 1 ; R 6 5 4 ; G 6 5 5 ; N 6 5 8 ; S 6 6 2 ; N 6 6 7 ; R 6 8 6 ; K 6 9 2 ; R 6 9 4 ; H 7 0 0 ; K 7 5 1 ; D 7 8 6 ; T 7 8 7 ; Y 7 8 9 ; T 8 8 2 ; K 8 8 6 ; N 8 8 8 ; 8 8 9 ; L 9 0 9 ; N 9 8 5 ; N 9 8 6 ; R 9 9 1 ; R 1 0 1 5 ; N 4 4 ; R 4 5 ; R 5 1 ; R 5 5 ; R 5 9 ; R 6 0 ; R 1 1 6 ; R 1 6 5 ; N 1 6 9 ; R 2 0 8 ; R 2 0 9 ; Y 2 1 1 ; T 2 3 8 ; Y 2 3 9 ; K 2 4 8 ; Y 2 5 6 ; R 3 1 4 ; N 3 9 4 ; Q 4 1 4 ; K 5 7 ; R 6 1 ; H 1 1 1 ; K 1 1 4 ; V 1 6 4 ; R 1 6 5 ; L 7 8 8 ; S 7 9 0 ; R 7 9 2 ; N 8 0 4 ; Y 8 6 8 ; K 8 7 0 ; K 8 7 8 ; K 8 7 9 ; K 8 8 1 ; Y 8 9 7 ; R 9 0 1 ; および/またはK 9 0 6 における突然変異を含む。

【0020】

一部の実施形態において、バリエーションS a C a s 9タンパク質は、以下の突然変異：Y 2 1 1 A ; Y 2 1 2 A ; W 2 2 9 A ; Y 2 3 0 A ; R 2 4 5 A ; T 3 9 2 A ; N 4 1 9 A ; L 4 4 6 A ; Q 4 8 8 A ; N 4 9 2 A ; Q 4 9 5 A ; R 4 9 7 A ; N 4 9 8 A ; R 4 9 9 A ; Q 5 0 0 A ; K 5 1 8 A ; K 5 2 3 A ; K 5 2 5 A ; H 5 5 7 A ; R 5 6 1 A ; K 5 7 2 A ; R 6 3 4 A ; Y 6 5 1 A ; R 6 5 4 A ; G 6 5 5 A ; N 6 5 8 A ; S 6 6 2 A ; N 6 6 7 A ; R 6 8 6 A ; K 6 9 2 A ; R 6 9 4 A ; H 7 0 0 A ; K 7 5 1 A ; D 7 8 6 A ; T 7 8 7 A ; Y 7 8 9 A ; T 8 8 2 A ; K 8 8 6 A ; N 8 8 8 A ; A 8 8 9 A ; L 9 0 9 A ; N 9 8 5 A ; N 9 8 6 A ; R 9 9 1 A ; R 1 0 1 5 A ; N 4 4 A ; R 4 5 A ; R 5 1 A ; R 5 5 A ; R 5 9 A ; R 6 0 A ; R 1 1 6 A ; R 1 6 5 A ; N 1 6 9 A ; R 2 0 8 A ; R 2 0 9 A ; T 2 3 8 A ; Y 2 3 9 A ; K 2 4 8 A ; Y 2 5 6 A ; R 3 1 4 A ; N 3 9 4 A ; Q 4 1 4 A ; K 5 7 A ; R 6 1 A ; H 1 1 1 A ; K 1 1 4 A ; V 1 6 4 A ; R 1 6 5 A ; L 7 8 8 A ; S 7 9 0 A ; R 7 9 2 A ; N 8 0 4 A ; Y 8 6 8 A ; K 8 7 0 A ; K 8 7 8 A ; K 8 7 9 A ; K 8 8 1 A ; Y 8 9 7 A ; R 9 0 1 A ; K 9 0 6 Aの1つ以上を含む。

【0021】

一部の実施形態において、バリエーションS a C a s 9タンパク質は、以下の追加の突然変異：Y 2 1 1 A、W 2 2 9 A、Y 2 3 0 A、R 2 4 5 A、T 3 9 2 A、N 4 1 9 A、L 4 4 6 A、Y 6 5 1 A、R 6 5 4 A、D 7 8 6 A、T 7 8 7 A、Y 7 8 9 A、T 8 8 2 A、K 8 8 6 A、N 8 8 8 A、A 8 8 9 A、L 9 0 9 A、N 9 8 5 A、N 9 8 6 A、R 9 9 1 A、R 1 0 1 5 A、N 4 4 A、R 4 5 A、R 5 1 A、R 5 5 A、R 5 9 A、R 6 0 A、R 1 1 6 A、R 1 6 5 A、N 1 6 9 A、R 2 0 8 A、R 2 0 9 A、T 2 3 8 A、Y 2 3 9 A

、K 2 4 8 A、Y 2 5 6 A、R 3 1 4 A、N 3 9 4 A、Q 4 1 4 A、K 5 7 A、R 6 1 A、H 1 1 1 A、K 1 1 4 A、V 1 6 4 A、R 1 6 5 A、L 7 8 8 A、S 7 9 0 A、R 7 9 2 A、N 8 0 4 A、Y 8 6 8 A、K 8 7 0 A、K 8 7 8 A、K 8 7 9 A、K 8 8 1 A、Y 8 9 7 A、R 9 0 1 A、K 9 0 6 Aの1つ以上を含む。

【 0 0 2 2 】

一部の実施形態において、バリエーション S a C a s 9 タンパク質は、複数の置換突然変異：R 2 4 5 / T 3 9 2 / N 4 1 9 / R 6 5 4 および Y 2 2 1 / R 2 4 5 / N 4 1 9 / R 6 5 4 (四重バリエーション突然変異体)；N 4 1 9 / R 6 5 4、R 2 4 5 / R 6 5 4、Y 2 2 1 / R 6 5 4、および Y 2 2 1 / N 4 1 9 (二重突然変異体)；R 2 4 5 / N 4 1 9 / R 6 5 4、Y 2 1 1 / N 4 1 9 / R 6 5 4、および T 3 9 2 / N 4 1 9 / R 6 5 4 (三重突然変異体)を含む。一部の実施形態において、突然変異体は、野生型アミノ酸に代えてアラニンを含む。

10

【 0 0 2 3 】

一部の実施形態において、バリエーション S a C a s 9 タンパク質は、D 1 0、E 4 7 7、D 5 5 6、H 7 0 1、または D 7 0 4；および H 5 5 7 または N 5 8 0 における突然変異からなる群から選択される、ヌクレアーゼ活性を減少させる1つ以上の突然変異も含む。一部の実施形態において、突然変異は、(i) D 1 0 A もしくは D 1 0 N、(ii) H 5 5 7 A、H 5 5 7 N、もしくは H 5 5 7 Y、(iii) N 5 8 0 A、および/または (iv) D 5 5 6 A である。

【 0 0 2 4 】

20

一部の実施形態において、バリエーション S a C a s 9 タンパク質は、以下の突然変異：E 7 8 2 K、K 9 2 9 R、N 9 6 8 K、または R 1 0 1 5 H の1つ以上を含む。具体的には、E 7 8 2 K / N 9 6 8 K / R 1 0 1 5 H (KKHバリエーション)；E 7 8 2 K / K 9 2 9 R / R 1 0 1 5 H (KRHバリエーション)；または E 7 8 2 K / K 9 2 9 R / N 9 6 8 K / R 1 0 1 5 H (KRKHバリエーション)である。

【 0 0 2 5 】

一部の実施形態において、バリエーション C a s 9 タンパク質は、特異性を増加させるための以下の領域の1つ以上の突然変異を含む。

【 0 0 2 6 】

【表 1】

機能的領域	SpCas9	SaCas9
スパーサー領域の DNA に接触する残基	L169; Y450; M495; N497; W659; R661; M694; Q695; H698; A728; Q926; E1108; V1015	Y211; W229; Y230; R245; T392; N419; L446; Y651; R654
非標的鎖の DNA に潜在的に接触する残基	N14; S15; S55; S730; K775; S777; R778; R780; K782; R783; K789; K797; Q805; N808; K810; R832; Q844; S845; K848; S851; K855; R859; K862; K890; Q920; K961; S964; K968; K974; R976; N980; H982; K1003; K1014; S1040; N1041; N1044; K1047; K1059; R1060; K1200; H1241; Q1254; Q1256; K1289; K1296; K1297; K1300; H1311; K1325	Q488A; N492A; Q495A; R497A; N498A; R499; Q500; K518; K523; K525; H557; R561; K572; R634; R654; G655; N658; S662; N667; R686; K692; R694; H700; K751
PAM 領域の DNA に接触する残基(直接的 PAM 接触を含む)	R71; Y72; R78; R165; R403; T404; F405; K1107; S1109; R1114; S1116; K1118; D1135; S1136; K1200; S1216; E1219; R1333; R1335; T1337	D786; T787; Y789; T882; K886; N888; A889; L909; N985; N986; R991; R1015
スパーサー領域の RNA に接触する残基	Y72; R75; K76; L101; S104; F105; R115; H116; I135; H160; K163; Y325; H328; R340; F351; D364; Q402; R403; I1110; K1113; R1122; Y1131	N44; R45; R51; R55; R59; R60; R116; R165; N169; R208; R209; Y211; T238; Y239; K248; Y256; R314; N394; Q414
リピート/アンチリピート領域の RNA に接触する残基	R63; R66; R70; R71; R74; R78; R403; T404; N407; R447; I448; Y450; K510; Y515; R661; V1009; Y1013	K57; R61; H111; K114; V164; R165; L788; S790; R792; N804; Y868; K870; K878; K879; K881; Y897; R901; K906
RNA ステムループに接触する残基	K30; K33; N46; R40; K44; E57; T62; R69; N77; L455; S460; R467; T472; I473; H721; K742; K1097; V1100; T1102; F1105; K1123; K1124; E1225; Q1272; H1349; S1351; Y1356	R47; K50; R54; R58; H62; R209; E213; S219; R452; K459; R774; N780; R781; L783

10

20

30

40

【 0 0 2 7 】

本明細書において、任意選択の介在リンカーにより異種機能ドメインに融合している本明細書に記載の単離バリエーション Cas9 タンパク質を含む融合タンパク質であって、リンカーが融合タンパク質の活性を妨害しない融合タンパク質も提供される。一部の実施形態において、異種機能ドメインは、DNA またはタンパク質、例えば、クロマチンに対して作用する。一部の実施形態において、異種機能ドメインは転写活性化ドメインである。一部の実施形態において、転写活性化ドメインは VP64 または NF- κ B p65 からのものである。一部の実施形態において、異種機能ドメインは転写サイレンサーまたは転写抑制ドメインである。一部の実施形態において、転写抑制ドメインは、クルップル関連ボックス (KRAB) ドメイン、ERF リプレッサードメイン (ERD)、または mSin

50

3 A相互作用ドメイン (SID) である。一部の実施形態において、転写サイレンサーはヘテロクロマチンタンパク質 1 (HP1)、例えば、HP1 または HP1 である。一部の実施形態において、異種機能ドメインは、DNA のメチル化状態を改変する酵素である。一部の実施形態において、DNA のメチル化状態を改変する酵素は、DNA メチルトランスフェラーゼ (DNMT) または TET タンパク質の全体もしくはジオキシゲナーゼドメイン、例えば、システインリッチ伸長部および 7 つの高度に保存されるエキソンによりコードされる 2 O G F e D O ドメイン、例えば、アミノ酸 1580 ~ 2052 を含む Tet 1 触媒ドメイン、アミノ酸 1290 ~ 1905 を含む Tet 2 およびアミノ酸 966 ~ 1678 を含む Tet 3 を含む触媒モジュールである。一部の実施形態において、TET タンパク質または TET 由来ジオキシゲナーゼドメインは、TET 1 からのものである。一部の実施形態において、異種機能ドメインは、ヒストンサブユニットを改変する酵素である。一部の実施形態において、ヒストンサブユニットを改変する酵素は、ヒストンアセチルトランスフェラーゼ (HAT)、ヒストンデアセチラーゼ (HDAC)、ヒストンメチルトランスフェラーゼ (HMT)、またはヒストンデメチラーゼである。一部の実施形態において、異種機能ドメインは生物学的係留物である。一部の実施形態において、生物学的係留物は MS2、Csy4 またはラムダ N タンパク質である。一部の実施形態において、異種機能ドメインは FokI である。

【0028】

本明細書において、本明細書に記載のバリエーション Cas9 タンパク質をコードする核酸、単離核酸、および任意選択的に本明細書に記載のバリエーション Cas9 タンパク質の発現のための 1 つ以上の調節ドメインに作動可能に結合されているその単離核酸を含むベクターも提供される。本明細書において、本明細書に記載の核酸を含み、かつ任意選択的に本明細書に記載のバリエーション Cas9 タンパク質を発現する宿主細胞、例えば細菌、酵母、昆虫もしくは哺乳動物宿主細胞またはトランスジェニック動物 (例えばマウス) も提供される。

【0029】

本明細書において、Cas9 バリエーションをコードする単離核酸および任意選択的にそのバリエーションの発現のための 1 つ以上の調節ドメインに作動可能に結合されているその単離核酸を含むベクターおよびその核酸を含み、かつ任意選択的にそのバリエーションタンパク質を発現する宿主細胞、例えば、哺乳動物宿主細胞も提供される。

【0030】

本明細書において、細胞中で、本明細書に記載のバリエーション Cas9 タンパク質または融合タンパク質、およびゲノム標的部位における最適なヌクレオチド間隔を有する、細胞のゲノムの選択部分に相補的な領域を有する少なくとも 1 つのガイド RNA を発現させることにより、またはそれらと細胞を接触させることにより、細胞のゲノムまたはエピゲノムを変更する方法も提供される。方法は、細胞を、例えば、単一ベクター中の Cas9 タンパク質およびガイド RNA をコードする核酸と接触させること；細胞を、例えば、複数のベクター中の Cas9 タンパク質コードする核酸およびガイド RNA をコードする核酸と接触させること；ならびに細胞を、精製 Cas9 タンパク質および合成または精製 gRNA の複合体と接触させることをとりわけ含み得る。一部の実施形態において、細胞は gRNA またはバリエーションタンパク質 / 融合タンパク質の一方または両方を安定的に発現し、他のエレメントは細胞中に形質移入または導入される。例えば、細胞は本明細書に記載のバリエーションタンパク質または融合タンパク質を安定的に発現し得、方法は、細胞を、合成 gRNA、精製組換え産生 gRNA、または gRNA をコードする核酸と接触させることを含み得る。一部の実施形態において、バリエーションタンパク質または融合タンパク質は、核局在化配列、細胞浸透ペプチド配列、および / または親和性タグの 1 つ以上を含む。

【0031】

本明細書において、dsDNA を、本明細書に記載の精製バリエーションタンパク質または融合タンパク質および dsDNA 分子の選択部分に相補的な領域を有するガイド RNA と接触させることにより、インビトロで単離 dsDNA 分子を変更する、例えば、選択的に

変更する方法も提供される。

【0032】

特に定義のない限り、本明細書において使用される全ての技術および科学用語は、本発明が属する技術分野の当業者により一般に理解されるものと同一の意味を有する。方法および材料は、本発明における使用のために本明細書に記載され、当技術分野において公知の他の好適な方法および材料を使用することもできる。材料、方法、および実施例は、説明のためのものにすぎず、限定するものではない。本明細書に挙げられる全ての刊行物、特許出願、特許、配列、データベースエントリー、および他の参考文献は、参照により全体として組み込まれる。矛盾する場合、本明細書が定義を含めて優先される。

【0033】

本発明の他の特徴部および利点は、以下の詳細な説明および図面、ならびに特許請求の範囲から明らかである。

【図面の簡単な説明】

【0034】

【図1-1】図1A：非特異的DNA接触を形成する残基中の突然変異を担持するSpCas9バリエーションの同定および特徴付けである。PDB 4OOGおよび4UN3（出典、それぞれ参考文献31および32）に基づく標的DNA：sgRNA二本鎖の野生型SpCas9認識を示す概略図である。

【図1-2】図1B：非特異的DNA接触を形成する残基中の突然変異を担持するSpCas9バリエーションの同定および特徴付けである。DNA骨格への水素結合を形成する位置におけるアラニン置換を含有するSpCas9バリエーションの特徴付けである。完全マッチsgRNAまたは標的部位に対するミスマッチをコードする4つの他のsgRNAを用いてプログラミングしたヒト細胞EGFP崩壊アッセイ使用して野生型SpCas9およびバリエーションを評価した。エラーバーは、 $n = 3$ についてのs.e.m.を表し；バックグラウンドEGFP損失の平均レベルを赤色点線により表す（このパネルおよびパネルCについて）。図1C：非特異的DNA接触を形成する残基中の突然変異を担持するSpCas9バリエーションの同定および特徴付けである。EGFP崩壊アッセイにより評価された24個の部位にわたる野生型SpCas9およびSpCas9-HF1のオンターゲット活性である。エラーバーは、 $n = 3$ についてのs.e.m.を表す。

【図1-3】図1D：非特異的DNA接触を形成する残基中の突然変異を担持するSpCas9バリエーションの同定および特徴付けである。T7E1アッセイにより評価された13個の内在性部位にわたる野生型SpCas9およびSpCas9-HF1のオンターゲット活性である。エラーバーは、 $n = 3$ についてのs.e.m.を表す。図1E：非特異的DNA接触を形成する残基中の突然変異を担持するSpCas9バリエーションの同定および特徴付けである。SpCas9-HF1のオンターゲット活性と、野生型SpCas9のオンターゲット活性との比（パネルCおよびDからのもの）である。

【図2-1】図2A：標準的な標的部位についてのsgRNAを用いた野生型SpCas9およびSpCas9-HF1のゲノムワイド特異性である。GUIDE-seqにより決定された、内在性ヒト遺伝子に標的化される8つのsgRNAを用いた野生型SpCas9およびSpCas9-HF1のオフターゲット部位である。リードカウントは、所与の部位における開裂頻度の尺度を表し；スパーサーまたはPAM内のミスマッチ位置を着色して強調する。

【図2-2】図2B：標準的な標的部位についてのsgRNAを用いた野生型SpCas9およびSpCas9-HF1のゲノムワイド特異性である。パネルAにおいて使用された8つのsgRNAからの野生型SpCas9およびSpCas9-HF1についてのGUIDE-seqにより同定されたゲノムワイドオフターゲット部位の総数のまとめである。図2C：標準的な標的部位についてのsgRNAを用いた野生型SpCas9およびSpCas9-HF1のゲノムワイド特異性である。オンターゲット活性に対するミスマッチ（プロトスパーサーおよびPAM内）の総数に従ってピン化された、8つのsgRNAについての野生型SpCas9およびSpCas9-HF1について同定されたオフ

10

20

30

40

50

ターゲット部位である。

【図3 - 1】図3 A : G U I D E - s e qにより同定されたオフターゲット部位の標的化ディープシーケンシングによる S p C a s 9 - H F 1 特異性改善のバリデーションである。図2からの6つの s g R N Aを用いた野生型 S p C a s 9 および S p C a s 9 - H F 1 についてのディープシーケンシングにより決定された平均オンターゲット改変パーセントである。エラーバーは、 $n = 3$ についての s . e . m . を表す。

【図3 - 2】図3 B : G U I D E - s e qにより同定されたオフターゲット部位の標的化ディープシーケンシングによる S p C a s 9 - H F 1 特異性改善のバリデーションである。インデル突然変異を含有するディープシーケンシングされたオンターゲット部位および G U I D E - s e q 検出されたオフターゲット部位の割合である。トリプリケート実験を、野生型 S p C a s 9、S p C a s 9 - H F 1、および対照条件についてプロットする。x軸下方の黒丸は、挿入突然変異も欠失突然変異も観察されなかったレプリケートを表す。P C Rにより増幅することができなかったオフターゲット部位を、アスタリスクとともに赤色文字で示す。プールリードカウントを用いる片側検定のフィッシャーの正確確率検定を使用する仮説検定は、E M X 1 - 1 オフターゲット1および F A N C F - 3 オフターゲット1のみにおける S p C a s 9 - H F 1 と対照条件との間の比較についての有意差を見出した (B e n j a m i n i - H o c h b e r g 法を使用する多重比較について調整後 $p < 0.05$)。有意差は、全てのオフターゲット部位における野生型 S p C a s 9 と S p C a s 9 - H F 1 との間、および R U N X 1 - 1 オフターゲット2を除き全てのオフターゲット部位における野生型 S p C a s 9 と対照条件との間でも見出された。

【図3 - 3】図3 C : G U I D E - s e qにより同定されたオフターゲット部位の標的化ディープシーケンシングによる S p C a s 9 - H F 1 特異性改善のバリデーションである。G U I D E - s e q リードカウント (図2 Aからのもの) と、野生型 S p C a s 9 を用いたオンおよびオフターゲット開裂部位におけるディープシーケンシングにより決定された平均改変パーセントとの間の相関の分散プロットである。

【図4 - 1】図4 A : 非標準的な反復部位についての s g R N Aを用いた野生型 S p C a s 9 および S p C a s 9 - H F 1 のゲノムワイド特異性である。多数のオフターゲット部位を開裂することが公知の2つの s g R N A (F u e t a l . , N a t B i o t e c h n o l 31, 822 - 826 (2013) ; T s a i e t a l . , N a t B i o t e c h n o l 33, 187 - 197 (2015)) を使用した野生型 S p C a s 9 および S p C a s 9 - H F 1 の G U I D E - s e q 特異性プロファイルである。G U I D E - s e q リードカウントは、所与の部位における開裂効率の尺度を表し ; スペースまたは P A M 内のミスマッチ位置を着色して強調し ; 赤色丸は、s g R N A - D N A 界面において、示されるバルジ (L i n e t a l . , N u c l e i c A c i d s R e s 42, 7473 - 7485 (2014)) を有する可能性が高い部位を示し ; 青色丸は、示されるものに対する代替ギャップありアラインメントを有し得る部位を示す (図8 参照)。

【図4 - 2】図4 B : 非標準的な反復部位についての s g R N Aを用いた野生型 S p C a s 9 および S p C a s 9 - H F 1 のゲノムワイド特異性である。パネル A において使用された2つの s g R N A からの野生型 S p C a s 9 および S p C a s 9 - H F 1 についての G U I D E - s e q により同定されたゲノムワイドオフターゲット部位の総数のまとめである。図4 C : 非標準的な反復部位についての s g R N Aを用いた野生型 S p C a s 9 および S p C a s 9 - H F 1 のゲノムワイド特異性である。オンターゲット部位に対するミスマッチ (プロトスペースおよび P A M 内) の総数に従ってピン化された、V E G F A 部位2および3についての野生型 S p C a s 9 または S p C a s 9 - H F 1 を用いて同定されたオフターゲット部位である。パネル A において赤色丸により標識されるオフターゲット部位は、これらのカウントに含めず ; パネル A において青色丸により標識される部位を、ギャップなしアラインメント中のミスマッチの数とともにカウントする。

【図5 - 1】図5 A : 追加の置換を担持する S p C a s 9 - H F 1 誘導体の活性である。8つの s g R N Aを用いた野生型 S p C a s 9、S p C a s 9 - H F 1、および S p C a

10

20

30

40

50

s 9 - H F 1 誘導体バリエーションのヒト細胞 E G F P 崩壊活性である。S p C a s 9 - H F 1 は、N 4 9 7 A、R 6 6 1 A、Q 6 9 5、および Q 9 2 6 A 突然変異を保有し；H F 2 = H F 1 + D 1 1 3 5 E；H F 3 = H F 1 + L 1 6 9 A；H F 4 = H F 1 + Y 4 5 0 A である。エラーバーは、n = 3 についての s . e . m . を表し；バックグラウンド E G F P 損失の平均レベルを赤色点線により表す。 図 5 B：追加の置換を担持する S p C a s 9 - H F 1 誘導体の活性である。パネル a からの 8 つの s g R N A を用いて野生型 S p C a s 9 と比較した S p C a s 9 - H F 1 バリエーションを使用した場合のオンターゲット活性のまとめである。中央値および四分位範囲を示し；野生型活性の > 70 % を示す間隔を緑色で強調する。 図 5 C：追加の置換を担持する S p C a s 9 - H F 1 誘導体の活性である。F A N C F 部位 2 および V E G F A 部位 3 オンターゲット部位、ならびに S p C a s 9 - H F 1 の効果に耐性の図 2 A および 4 A からのオフターゲット部位における S p C a s 9 および H F バリエーションによる平均改変パーセントである。改変パーセントは T 7 E 1 アッセイにより決定し；全ての実験についてバックグラウンドインデル割合を差し引いた。エラーバーは、n = 3 についての s . e . m . を表す。 図 5 D：追加の置換を担持する S p C a s 9 - H F 1 誘導体の活性である。オンターゲット活性とオフターゲット活性（パネル C からのもの）との比としてプロットされる、F A N C F 部位 2 または V E G F A 部位 3 s g R N A を用いた野生型 S p C a s 9 および H F バリエーションの特異性の比である。

10

【図 5 - 2】図 5 E：S p C a s 9 - H F 1 の効果に耐性のオフターゲット部位を有する s g R N A を用いた、S p C a s 9 - H F 1、- H F 2、および - H F 4 のゲノムワイド特異性である。パネル F における G U I D E - s e q 実験についての目的オンターゲット部位における平均 G U I D E - s e q タグインテグレーション。S p C a s 9 - H F 1 = N 4 9 7 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A；H F 2 = H F 1 + D 1 1 3 5 E；H F 4 = H F 1 + Y 4 5 0 A である。エラーバーは、n = 3 についての s . e . m . を表す。

20

【図 5 - 3】図 5 F：S p C a s 9 - H F 1 の効果に耐性のオフターゲット部位を有する s g R N A を用いた、S p C a s 9 - H F 1、- H F 2、および - H F 4 のゲノムワイド特異性である。F A N C F 部位 2 または V E G F A 部位 3 s g R N A のいずれかを用いた S p C a s 9 - H F 1、- H F 2、または - H F 4 の G U I D E - s e q 同定オフターゲット部位である。リードカウントは、所与の部位における開裂頻度の尺度を表し；スパーサーまたは P A M 内のミスマッチ位置を着色して強調する。オフターゲット区別の改善倍率は、S p C a s 9 - H F バリエーションについてのオフターゲットリードカウントを S p C a s 9 - H F バリエーション間の比較前にオンターゲット部位におけるリードカウントに正規化することにより計算した。

30

【図 6 - 1】図 6 A：s g R N A および標的 D N A との S p C a s 9 相互作用である。S p C a s 9：s g R N A 複合体を、s g R N A と標的 D N A との間の塩基対合とともに説明する概略図である。

【図 6 - 2】図 6 B：s g R N A および標的 D N A との S p C a s 9 相互作用である。P D B：4 U N 3（参照文献 32）からの、標的 D N A に結合している S p C a s 9：s g R N A 複合体の構造表示である。標的鎖 D N A 骨格への水素結合接触を形成する 4 つの残基を青色で強調し；可視目的のために H N H ドメインを非表示にする。

40

【図 7】図 7 A：G U I D E - s e q 実験に使用された種々の s g R N A を用いた野生型および S p C a s 9 - H F 1 のオンターゲット活性比較である。制限断片長多型アッセイにより定量された、図 2 A に示される G U I D E - s e q 実験についての目的オンターゲット部位における平均 G U I D E - s e q タグインテグレーションである。エラーバーは、n = 3 についての s . e . m . を表す。 図 7 B：G U I D E - s e q 実験に使用された種々の s g R N A を用いた野生型および S p C a s 9 - H F 1 のオンターゲット活性比較である。T 7 E 1 アッセイにより検出された、図 2 A に示される G U I D E - s e q 実験についての目的オンターゲット部位における平均改変パーセントである。エラーバーは、n = 3 についての s . e . m . を表す。 図 7 C：G U I D E - s e q 実験に使用され

50

た種々の sgRNA を用いた野生型および SpCas9 - HF1 のオンターゲット活性比較である。制限断片長多型アッセイにより定量された、図 4 A に示される GUIDE-seq 実験についての目的オンターゲット部位における平均 GUIDE-seq タグインテグレーションである。エラーバーは、 $n = 3$ についての s.e.m. を表す。 図 7 D : GUIDE-seq 実験に使用された種々の sgRNA を用いた野生型および SpCas9 - HF1 のオンターゲット活性比較である。T7E1 アッセイにより検出された、図 4 A に示される GUIDE-seq 実験についての目的オンターゲット部位における平均改変パーセントである。エラーバーは、 $n = 3$ についての s.e.m. を表す。

【図 8】VEGFA 部位 2 オフターゲット部位についての潜在的代替アラインメントである。Geneious (Kearse et al., Bioinformatics 28, 1647 - 1649 (2012)) バージョン 8.1.6 を使用してアラインされた、単一ヌクレオチドギャップを含有するオフターゲット部位として潜在的に認識することができる GUIDE-seq (左) により同定された 10 個の VEGFA 部位 2 オフターゲット部位 (Lin et al., Nucleic Acids Res 42, 7473 - 7485 (2014)) (右) である。

【図 9】トランケート sgRNA 14 を用いた野生型 SpCas9 および SpCas9 - HF1 の活性。EGFP 中の 4 つの部位に標的化される全長またはトランケート sgRNA を使用した野生型 SpCas9 および SpCas9 - HF1 の EGFP 崩壊活性である。エラーバーは、 $n = 3$ についての s.e.m. を表し；対照実験におけるバックグラウンド EGFP 損失の平均レベルを赤色点線により表す。

【図 10】5' - ミスマッチグアニン塩基を担持する sgRNA を用いた野生型 SpCas9 および SpCas9 - HF1 活性である。4 つの異なる部位に標的化される sgRNA を用いた野生型 SpCas9 および SpCas9 - HF1 の EGFP 崩壊活性である。それぞれの標的部位について、sgRNA は、マッチ非グアニン 5' - 塩基または意図的にミスマッチさせた 5' - グアニンのいずれかを含有する。

【図 11】野生型 SpCas9 および SpCas9 - HF1 発現プラスミドの量のタイトレーションである。変動量の野生型および SpCas9 - HF1 発現プラスミドによる形質移入からのヒト細胞 EGFP 崩壊活性。全ての形質移入について、sgRNA 含有プラスミドの量を 250 ng に固定した。別個の部位を標的化する 2 つの sgRNA を使用し；エラーバーは、 $n = 3$ についての s.e.m. を表し；陰性対照におけるバックグラウンド EGFP 損失の平均レベルを赤色点線により表す。

【図 12 - 1】図 12 A : SpCas9 - HF1 の PAM 認識特異性の変更である。T7E1 アッセイにより定量された、8 つの sgRNA を使用した SpCas9 - VQR (参照文献 15) および改善された SpCas9 - VRQR によるオンターゲット内在性ヒト部位の平均改変パーセントの比較である。両方のバリエーションは、NGAN PAM を認識するように遺伝子操作されている。エラーバーは、 $n = 2$ または 3 についての s.e.m. を表す。 図 12 B : SpCas9 - HF1 の PAM 認識特異性の変更である。8 つの sgRNA を使用した SpCas9 - VQR および SpCas9 - VRQR の、それらの - HF1 相当物と比較したオンターゲット EGFP 崩壊活性である。エラーバーは、 $n = 3$ についての s.e.m. を表し；陰性対照におけるバックグラウンド EGFP 損失の平均レベルを赤色点線により表す。

【図 12 - 2】図 12 C : SpCas9 - HF1 の PAM 認識特異性の変更である。T7E1 アッセイにより定量された 8 つの内在性ヒト遺伝子部位における SpCas9 - VQR および SpCas9 - VRQR による、それらの - HF1 バリエーションと比較した平均オンターゲット改変パーセントの比較である。エラーバーは、 $n = 3$ についての s.e.m. を表し；ND は、不検出である。 図 12 D : SpCas9 - HF1 の PAM 認識特異性の変更である。SpCas9 - VQR または SpCas9 - VRQR を使用し、それらの対応する - HF1 バリエーション (パネル B および C からのも) と比較した場合のオンターゲット活性の変化倍率のまとめである。中央値および四分位範囲を示し；野生型活性の > 70 % を示す間隔を緑色で強調する。

10

20

30

40

50

【図13】図13A：野生型SpCas9、SpCas9-HF1、および非標的DNA鎖に潜在的に接触し得る位置における1つ以上のアラニン置換を担持する野生型SpCas9誘導体の活性である。EGFP遺伝子中の部位に完全にマッチするsgRNAならびに11および12位において意図的にミスマッチさせたsgRNAを用いるEGFP崩壊アッセイを使用してヌクレアーゼを評価した。ミスマッチ位置に番号付与し、20位は最もPAMから遠位の位置であり；赤色点線は、EGFP崩壊のバックグラウンドレベルを表し；HF1 = N497A / R661A / Q695A / Q926A置換を有するSpCas9である。図13B：野生型SpCas9、SpCas9-HF1、および非標的DNA鎖に潜在的に接触し得る位置における1つ以上のアラニン置換を担持する野生型SpCas9誘導体の活性である。EGFP遺伝子中の部位に完全にマッチするsgRNAならびに9および10位（パネルB）において意図的にミスマッチさせたsgRNAを用いるEGFP崩壊アッセイを使用してヌクレアーゼを評価した。ミスマッチ位置に番号付与し、20位は最もPAMから遠位の位置であり；赤色点線は、EGFP崩壊のバックグラウンドレベルを表し；HF1 = N497A / R661A / Q695A / Q926A置換を有するSpCas9である。

10

【図14】図14A：野生型SpCas9、SpCas9-HF1、および非標的DNA鎖に潜在的に接触し得る位置における1つ以上のアラニン置換を担持するSpCas9-HF1誘導体の活性である。EGFP遺伝子中の部位に完全にマッチするsgRNAならびに11および12位において意図的にミスマッチさせたsgRNAを用いるEGFP崩壊アッセイを使用してヌクレアーゼを評価した。ミスマッチ位置に番号付与し、20位は最もPAMから遠位の位置であり；赤色点線は、EGFP崩壊のバックグラウンドレベルを表し；HF1 = N497A / R661A / Q695A / Q926A置換を有するSpCas9である。図14B：野生型SpCas9、SpCas9-HF1、および非標的DNA鎖に潜在的に接触し得る位置における1つ以上のアラニン置換を担持するSpCas9-HF1誘導体の活性である。EGFP遺伝子中の部位に完全にマッチするsgRNAならびに9および10位（パネルB）において意図的にミスマッチさせたsgRNAを用いるEGFP崩壊アッセイを使用してヌクレアーゼを評価した。ミスマッチ位置に番号付与し、20位は最もPAMから遠位の位置であり；赤色点線は、EGFP崩壊のバックグラウンドレベルを表し；HF1 = N497A / R661A / Q695A / Q926A置換を有するSpCas9である。

20

30

【図15】野生型SpCas9、SpCas9-HF1、および非標的DNA鎖に潜在的に接触し得る位置における1つ以上のアラニン置換を担持するSpCas9（Q695A / Q926A）誘導体の活性である。EGFP遺伝子中の部位に完全にマッチするsgRNAならびに11および12位において意図的にミスマッチさせたsgRNAを用いるEGFP崩壊アッセイを使用してヌクレアーゼを評価した。ミスマッチ位置に番号付与し、20位は最もPAMから遠位の位置であり；赤色点線は、EGFP崩壊のバックグラウンドレベルを表し；HF1 = N497A / R661A / Q695A / Q926A置換を有するSpCas9であり；Dbl = Q695A / Q926A置換を有するSpCas9である。

【図16】マッチsgRNAおよびスパーサー中のそれぞれ位置における単一ミスマッチを有するsgRNAを使用した野生型SpCas9、SpCas9-HF1、およびeSpCas9-1.1の活性である。EGFP遺伝子中の部位に完全にマッチするsgRNA（「マッチ」）ならびに示される位置において意図的にミスマッチさせたsgRNAを用いるEGFP崩壊アッセイを使用してヌクレアーゼを評価した。ミスマッチ位置に番号付与し、20位は最もPAMから遠位の位置である。SpCas9-HF1 = N497A / R661A / Q695A / Q926Aであり、eSP1.1 = K848A / K1003A / R1060Aである。

40

【図17】図17A：マッチsgRNAおよびスパーサー中の種々の位置における単一ミスマッチを有するsgRNAを使用した野生型SpCas9およびバリエーションの活性である。アラニン置換（標的または非標的DNA鎖に潜在的に接触し得る位置に指向される）

50

の組合せを含有する S p C a s 9ヌクレアーゼの活性を、E G F P 遺伝子中の部位に完全にマッチする s g R N A (「マッチ」)ならびに示されるスパー位置において意図的にミスマッチさせた s g R N A を用いる E G F P 崩壊アッセイを使用して評価した。ミスマッチ位置に番号付与し、20位は最も P A M から遠位の位置であり; m m = ミスマッチである。W T = 野生型であり、D b = Q 6 9 5 A / Q 9 2 6 A であり、H F 1 = N 4 9 7 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A であり、1 . 0 = K 8 1 0 A / K 1 0 0 3 A / R 1 0 6 0 A であり、1 . 1 = K 8 4 8 A / K 1 0 0 3 A / R 1 0 6 0 A である。図 1 7 B : マッチ s g R N A およびスパー中の種々の位置における単一ミスマッチを有する s g R N A を使用した野生型 S p C a s 9 およびバリエーションの活性である。マッチオンターゲット部位について残りの全ての考えられる単一ミスマッチの s g R N A を使用して、(a) からのこれらのヌクレアーゼのサブセットを試験した。ミスマッチ位置に番号付与し、20位は最も P A M から遠位の位置であり; m m = ミスマッチである。W T = 野生型であり、D b = Q 6 9 5 A / Q 9 2 6 A であり、H F 1 = N 4 9 7 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A であり、1 . 0 = K 8 1 0 A / K 1 0 0 3 A / R 1 0 6 0 A であり、1 . 1 = K 8 4 8 A / K 1 0 0 3 A / R 1 0 6 0 A である。

10

【図 1 8】マッチ s g R N A およびスパー中の種々の個々の位置におけるミスマッチを有する s g R N A を使用した野生型 S p C a s 9 およびバリエーションの活性である。アラニン置換 (標的または非標的 D N A 鎖に潜在的に接触し得る位置に指向される) の組合せを含有する S p C a s 9ヌクレアーゼの活性を、E G F P 遺伝子中の部位に完全にマッチする s g R N A (「マッチ」)ならびに示される位置において意図的にミスマッチさせた s g R N A を用いる E G F P 崩壊アッセイを使用して評価した。D b = Q 6 9 5 A / Q 9 2 6 A であり、H F 1 = N 4 9 7 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A である。

20

【図 1 9 - 1】図 1 9 A : マッチ s g R N A およびスパー中の種々の個々の位置におけるミスマッチを有する s g R N A を使用した野生型 S p C a s 9 およびバリエーションの活性である。アラニン置換 (標的または非標的 D N A 鎖に潜在的に接触し得る位置を対象とする) の組合せを含有する S p C a s 9ヌクレアーゼのオンターゲット活性を、E G F P 遺伝子中の部位に完全にマッチする 2 つの s g R N A を用いる E G F P 崩壊アッセイを使用して評価した。D b = Q 6 9 5 A / Q 9 2 6 A であり、H F 1 = N 4 9 7 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A である。

【図 1 9 - 2】図 1 9 B : マッチ s g R N A およびスパー中の種々の個々の位置におけるミスマッチを有する s g R N A を使用した野生型 S p C a s 9 およびバリエーションの活性である。スパー配列中の (s g R N A 「部位 1」) の 1 2、1 4、1 6、または 1 8 位におけるミスマッチを含有する s g R N A を用いて (a) からのこれらのヌクレアーゼのサブセットを試験して、それらの置換によりミスマッチの不寛容性が付与されたか否かを決定した。D b = Q 6 9 5 A / Q 9 2 6 A であり、H F 1 = N 4 9 7 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A である。

30

【図 2 0】四重突然変異体構築物中の突然変異の位置間の類似性を説明する S p C a s 9 (上段) および S a C a s 9 (下段) の構造比較 (黄色球状表示で示される) である。D N A 骨格に接触する他の残基も桃色球状表示で示す。

【図 2 1】図 2 1 A : 野生型 S a C a s 9 および 1 つ以上のアラニン置換を担持する S a C a s 9 誘導体の活性である。S a C a s 9 置換は、標的 D N A 鎖に潜在的に接触し得る位置を対象とした。E G F P 遺伝子中の部位に完全にマッチする s g R N A ならびに 1 1 および 1 2 位において意図的にミスマッチさせた s g R N A を用いる E G F P 崩壊アッセイを使用してヌクレアーゼを評価した。ミスマッチ位置に番号付与し、20位は最も P A M から遠位の位置であり; 赤色点線は、E G F P 崩壊のバックグラウンドレベルを表す。

40

図 2 1 B : 野生型 S a C a s 9 および 1 つ以上のアラニン置換を担持する S a C a s 9 誘導体の活性である。S a C a s 9 置換は、P A M 特異性に既に影響することが示されている位置を対象とした。E G F P 遺伝子中の部位に完全にマッチする s g R N A ならびに 1 1 および 1 2 位において意図的にミスマッチさせた s g R N A を用いる E G F P 崩壊アッセイを使用してヌクレアーゼを評価した。ミスマッチ位置に番号付与し、20位は最も

50

PAMから遠位の位置であり；赤色点線は、EGFP崩壊のバックグラウンドレベルを表す。

【図22-1】図22A：野生型(WT) SaCas9および標的DNA鎖に潜在的に接触し得る残基における1つ以上のアラニン置換を担持するSaCas9誘導体の活性である。EGFP遺伝子中の部位に完全にマッチするsgRNA(「マッチ」)ならびに19および20位において意図的にミスマッチさせたsgRNAを用いるEGFP崩壊アッセイを使用してヌクレアーゼを評価した。ミスマッチ位置に番号付与し、20位は最もPAMから遠位の位置である。

【図22-2】図22B：野生型(WT) SaCas9および標的DNA鎖に潜在的に接触し得る残基における1つ以上のアラニン置換を担持するSaCas9誘導体の活性である。EGFP遺伝子中の部位に完全にマッチするsgRNA(「マッチ」)ならびに19および20位において意図的にミスマッチさせたsgRNAを用いるEGFP崩壊アッセイを使用してヌクレアーゼを評価した。ミスマッチ位置に番号付与し、20位は最もPAMから遠位の位置である。

【図23】野生型(WT) SaCas9および標的DNA鎖に潜在的に接触し得る残基におけるアラニン置換の三重組合せを担持するSaCas9バリエーションの活性である。EGFP崩壊アッセイを使用してヌクレアーゼを評価した。4つの異なるsgRNAを使用し(マッチ#1~4)、さらに野生型SaCas9により効率的に使用されることが公知のミスマッチsgRNAを用いて4つの標的部位のそれぞれを試験した。それぞれの部位についてのミスマッチsgRNAを、それぞれのマッチsgRNAの右側に示す(例えば、マッチ部位3についての唯一のミスマッチsgRNAは、mm11および12である)。ミスマッチ位置に番号付与し、21位は最もPAMから遠位の位置であり；mmは、ミスマッチである。

【図24】図24A：野生型(WT) SaCas9および標的DNA鎖に潜在的に接触し得る残基における1つ以上のアラニン置換を担持するSaCas9誘導体の活性である。二重組合せ置換を担持するSaCas9バリエーションを、マッチおよび単一ミスマッチ内在性ヒト遺伝子標的的部位に対してT7E1アッセイを使用して評価した。マッチ「オンターゲット」部位を、Kleinstiver et al., Nature Biotechnology 2015からのそれらの遺伝子標的的部位のsgRNA番号に従って命名する。ミスマッチsgRNAに番号付与し、ミスマッチは最もPAMから遠位の位置の21位において生じ；ミスマッチsgRNAは、ミスマッチsgRNAの左側に列記されるマッチオンターゲット部位に由来する。図24B：野生型(WT) SaCas9および標的DNA鎖に潜在的に接触し得る残基における1つ以上のアラニン置換を担持するSaCas9誘導体の活性である。三重組合せ置換を担持するSaCas9バリエーションを、マッチおよび単一ミスマッチ内在性ヒト遺伝子標的的部位に対してT7E1アッセイを使用して評価した。マッチ「オンターゲット」部位を、Kleinstiver et al., Nature Biotechnology 2015からのそれらの遺伝子標的的部位のsgRNA番号に従って命名する。ミスマッチsgRNAに番号付与し、ミスマッチは最もPAMから遠位の位置の21位において生じ；ミスマッチsgRNAは、ミスマッチsgRNAの左側に列記されるマッチオンターゲット部位に由来する。

【発明を実施するための形態】

【0035】

CRISPR-Cas9ヌクレアーゼの制限は、不完全にマッチする標的的部位における不所望な「オフターゲット」突然変異を誘導するそれらの潜在性であり(例えば、Tsai et al., Nat Biotechnol. 2015参照)、ある場合には頻度が目的オンターゲット部位において観察されるものと同様である(Fu et al., Nat Biotechnol. 2013)。CRISPR-Cas9ヌクレアーゼを用いる従来の研究は、ガイドRNA(gRNA)と標的部位のスペーサー領域との間の配列特異的相互作用の数を低減させることにより、ヒト細胞中の開裂のオフターゲット部位における突然変異効果を低減させ得ることを示唆している(Fu et al., Nat

10

20

30

40

50

Biotechnol. 2014)。

【0036】

これは、gRNAをその5'末端において2または3 ntだけランケートすることにより早期に達成され、この増加した特異性の機序は、gRNA/Cas9複合体の相互作用エネルギーの減少であり、その結果、それはオンターゲット部位を開裂するためにちょうど十分なエネルギーとバランスを取り、それにより標的DNA部位中のミスマッチに起因するエネルギーペナルティーが存在すると想定されるオフターゲット部位を開裂するために十分なエネルギーを有する可能性が低くなることが仮定された(国際公開第2015/099850号パンフレット)。

【0037】

SpCas9のオフターゲット効果(ガイドRNAのための目的標的部位を有する不完全マッチまたはミスマッチのDNA部位におけるもの)は、その標的DNA部位との非特異的相互作用を減少させることにより最小化し得ることが仮定された。SpCas9-sgRNA複合体は、NGG PAM配列(SpCas9により認識される)(Deltcheva, E. et al. Nature 471, 602-607 (2011); Jinek, M. et al. Science 337, 816-821 (2012); Jiang, W., et al., Nat Biotechnol 31, 233-239 (2013); Sternberg, S. H., et al., Nature 507, 62-67 (2014))および隣接20 bp プロトSpacer配列(sgRNAの5'末端に相補的である)(Jinek, M. et al. Science 337, 816-821 (2012); Jinek, M. et al. Elife 2, e00471 (2013); Mali, P. et al., Science 339, 823-826 (2013); Cong, L. et al., Science 339, 819-823 (2013))から構成される標的部位を開裂する。SpCas9-sgRNA複合体は、その目的標的DNA部位の認識に必要とされるよりも大きいエネルギーを有し得、それにより、ミスマッチオフターゲット部位の開裂が可能となることが既に理論付けられた(Fu, Y., et al., Nat Biotechnol 32, 279-284 (2014))。この特性は、適応細菌免疫におけるCas9の目的の役割に有利であり得ることが想定することができ、突然変異し得る外来配列を開裂する能力をそれに付与する。この過剰なエネルギーモデルは、SpCas9濃度を減少させることにより(Hsu, P. D. et al. Nat Biotechnol 31, 827-832 (2013); Pattanayak, V. et al. Nat Biotechnol 31, 839-843 (2013))、またはsgRNAの相補性長さを低減させることにより(Fu, Y., et al., Nat Biotechnol 32, 279-284 (2014))、オフターゲット効果を低減させることができる(しかし、排除することはできない)ことを実証する従来の研究によっても支持されているが、この効果についての他の解釈も提案されている(Josephs, E. A. et al. Nucleic Acids Res 43, 8924-8941 (2015); Sternberg, S. H., et al. Nature 527, 110-113 (2015); Kiani, S. et al. Nat Methods 12, 1051-1054 (2015))。構造データは、SpCas9-sgRNA-標的DNA複合体を、4つのSpCas9残基(N497、R661、Q695、Q926)により作製される、標的DNA鎖のリン酸骨格への直接的な水素結合を含むいくつかのSpCas9媒介DNA接触により安定化し得ることを示唆する(Nishimasu, H. et al. Cell 156, 935-949 (2014); Anders, C., et al. Nature 513, 569-573 (2014))(図1aならびに図6aおよび6b)。本発明者らは、これらの接触の1つ以上の崩壊が、ロバストなオンターゲット活性を保持するためにちょうど十分であるが、ミスマッチオフターゲット部位を開裂する能力が縮小するレベルにおいてSpCas9-sgRNA複合体のエネルギー的なバランスを取り得ることを想定した。

【0038】

本明細書に記載のとおり、C a s 9 タンパク質は、理論的には、D N A についての C a s 9 の結合親和性を低減させることにより、増加した特異性を示すように遺伝子操作することができる。広く使用される化膿性連鎖球菌 (S t r e p t o c o c c u s p y o g e n e s) C a s 9 (S p C a s 9) のいくつかのバリエーションは、構造情報、細菌選択ベース指向進化、およびコンビナトリアル設計を使用して D N A 骨格上のリン酸と相互作用することが予測することができる S p C a s 9 中の種々の残基中に個々のアラニン置換を導入することにより遺伝子操作された。ロバストな大腸菌 (E . c o l i) ベーススクリーニングアッセイを使用してバリエーションを細胞活性についてさらに試験してそれらのバリエーションの細胞活性を評価し；この細菌系において、細胞生存率は、毒性ギラーゼ毒素 c c d B についての遺伝子ならびに g R N A および S p C a s 9 により標的化される 23 塩基対配列を含有する選択プラスミドの開裂および後続の崩壊に依存し、活性の保持または損失に関連する残基の同定をもたらした。さらに、ヒト細胞中で改善された標的特異性を示す別の S p C a s 9 バリエーションを同定および特徴付けした。

【0039】

さらに、細菌細胞ベース系において評価された S p C a s 9 の単一アラニン置換突然変異体の活性は、50 ~ 100 % の生存割合が通常ロバストな開裂を示す一方、0 % の生存率は、酵素が機能的に減衰したことを示すことを示した。次いで、R 6 3 A、R 6 6 A、R 6 9 A、R 7 0 A、R 7 1 A、Y 7 2 A、R 7 4 A、R 7 5 A、K 7 6 A、N 7 7 A、R 7 8 A、R 1 1 5 A、H 1 6 0 A、K 1 6 3 A、R 1 6 5 A、L 1 6 9 A、R 4 0 3 A、T 4 0 4 A、F 4 0 5 A、N 4 0 7 A、R 4 4 7 A、N 4 9 7 A、I 4 4 8 A、Y 4 5 0 A、S 4 6 0 A、M 4 9 5 A、K 5 1 0 A、Y 5 1 5 A、R 6 6 1 A、M 6 9 4 A、Q 6 9 5 A、H 6 9 8 A、Y 1 0 1 3 A、V 1 0 1 5 A、R 1 1 2 2 A、K 1 1 2 3 A、K 1 1 2 4 A、K 1 1 5 8 A、K 1 1 8 5 A、K 1 2 0 0 A、S 1 2 1 6 A、Q 1 2 2 1 A、K 1 2 8 9 A、R 1 2 9 8 A、K 1 3 0 0 A、K 1 3 2 5 A、R 1 3 3 3 A、K 1 3 3 4 A、R 1 3 3 5 A、および T 1 3 3 7 A を含む S p C a s 9 の追加の突然変異を細菌中でアッセイした。細菌中で < 5 % の生存率を有する 2 つの突然変異体 (R 6 9 A および F 4 0 5 A) を除き、それらの追加の単一突然変異の全てが S p C a s 9 のオンターゲット活性に対する効果をほとんど有さないと考えられた (細菌スクリーンにおいて > 70 % の生存率)。

【0040】

細菌スクリーンにおいて同定された C a s 9 のバリエーションがヒト細胞中で効率的に機能したか否かをさらに決定するため、ヒト U 2 O S 細胞ベース E G F P 崩壊アッセイを使用して種々のアラニン置換 C a s 9 突然変異体を試験した。このアッセイにおいて、単一のインテグレートされた構造的に発現される E G F P 遺伝子のコード配列中の標的部位の良好な開裂は、フローサイトメトリーにより定量的に評価されたインデル突然変異の誘導および E G F P 活性の崩壊をもたらした (例えば、Reyon et al., Nat Biotechnol. 2012 May; 30 (5): 460 - 5 参照)。

【0041】

これらの実験は、細菌細胞ベースアッセイにおいて得られた結果が、ヒト細胞中でヌクレアーゼ活性と十分に相関することを示し、それらの遺伝子操作方針を他の種および異なる細胞からの C a s 9 に拡張し得ることを示唆する。したがって、これらの知見は、本明細書において「バリエーション」または「そのバリエーション」と集合的に称される S p C a s 9 および S a C a s 9 バリエーションについての支持を提供する。

【0042】

本明細書に記載のバリエーションの全ては、例えば、単純な部位特異的突然変異誘発により既存の広く使用されるベクター中に急速に取り込むことができ、それらは少数の突然変異を要求するにすぎないため、バリエーションは、S p C a s 9 プラットフォームの既に記載された他の改善 (例えば、トランケート s g R N A (Tsai et al., Nat Biotechnol. 33, 187 - 197 (2015); Fu et al., Nat Biotechnol. 32, 279 - 284 (2014))、ニッカーゼ突然変異 (

Mali et al., Nat Biotechnol 31, 833 - 838 (2013); Ran et al., Cell 154, 1380 - 1389 (2013); FokI-dCas9融合 (Guilinger et al., Nat Biotechnol 32, 577 - 582 (2014); Tsai et al., Nat Biotechnol 32, 569 - 576 (2014); 国際公開第2014144288号パンフレット); および変更PAM特異性を有する遺伝子操作CRISPR-Cas9ヌクレアーゼ (Kleinstiver et al., Nature. 2015 Jul 23; 523 (7561): 481 - 5) とともに機能するはずである。

【0043】

したがって、本明細書において、Cas9バリエーション、例として、SpCas9バリエーションが提供される。SpCas9野生型配列は、以下のとおりである。

【化1】

10	20	30	40	50	60
MDKKYSIGLD	IGTNSVGWAV	ITDEYKVP SK	KFKVLGNTDR	HSIKKNLIGA	LLFDSGETAE
70	80	90	100	110	120
ATRLKRTARR	RYTRRKNRIC	YLQEIFSNE M	AKVDDSF FHR	LEESFLVEED	KKHERHPIFG
130	140	150	160	170	180
NIVDEVAYHE	KYPTIYHLRK	KLVDSTDKAD	LRLIYLALAH	MIKFRGHFLI	EGDLNPDNSD
190	200	210	220	230	240
VDKLFIQLVQ	TYNQLFEENP	INASGVDAKA	ILSARLSKSR	RLENLIAQLP	GEKKNGLFGN
250	260	270	280	290	300
LIALSLGLTP	NFKSNFDLAE	DAKLQLSKDT	YDDDLNLLA	QIGDQYADLF	LAAKNLSDAI
310	320	330	340	350	360
LLSDILRVNT	EITKAPLSAS	MIKRYDEHHQ	DLTLLKALVR	QQLPEKYKEI	FFDQSKNGYA
370	380	390	400	410	420
GYIDGGASQE	EFYKFIKPIL	EKMDGTEELL	VKLNREDLLR	KQRTFDNGSI	PHQIHLGELH
430	440	450	460	470	480
AILRRQEDFY	PFLKDNREKI	EKILTFRI PY	YVGPLARGNS	RFAWMTRKSE	ETITPWNFEE
490	500	510	520	530	540
VVDKGASAQ S	FIERMTNF DK	NLPNEKVL PK	HSLLYEYFTV	YNELTKVKYV	TEGMRKPAFL
550	560	570	580	590	600
SGEQKKAIVD	LLFKTNRKVT	VKQLKEDYFK	KIECFDSVEI	SGVEDRFNAS	LGTYHDLLKI
610	620	630	640	650	660
IKDKDFLDNE	ENEDILEDIV	LTLTLFEDRE	MIEERLKTYA	HLFDDKVMKQ	LKRRRYTGWG
670	680	690	700	710	720
RLSRKLINGI	RDQSGKTI L	DFLKSDGFAN	RNFMQLIHDD	SLTFKEDIQK	AQVSGQGD SL

20

30

【化 2】

```

      730      740      750      760      770      780
HEHIANLAGS PAIKKGILQT VKVVDLVKV MGRHKPENIV IEMARENQTT QKGQKNSRER

      790      800      810      820      830      840
MKRIEEGIKE LGSQILKEHP VENTQLQNEK LYLYYLQNGR DMYVDQELDI NRLSDYDVDH

      850      860      870      880      890      900
IVPQSFLKDD SIDNKVLTRS DKNRGKSDNV PSEEVVKMK NYWRQLLNAK LITQRKFDNL

      910      920      930      940      950      960
TKAERGGLSE LDKAGFIKRO LVETRQITKH VAQILDSRMN TKYDENDKLI REVKVITLKS

      970      980      990      1000      1010      1020
KLVSDFRKDF QFYKVVREINN YHHAHDAYLN AVVGITALIKK YPKLESEFVY GDYKVYDVRK

      1030      1040      1050      1060      1070      1080
MIAKSEQEIG KATAKYFFYS NIMNFFKTEI TLANGEIRKR PLIETNGETG EIVWDKGRDF

      1090      1100      1110      1120      1130      1140
ATVRKVLSPM QVNIVKKTEV QTGGFSKESI LPKRNSDKLI ARKKDWDPKK YGGFDSPTVA

      1150      1160      1170      1180      1190      1200
YSVLVVAKVE KGKSKKLKSV KELLGITIME RSSFEKNPID FLEAKGYKEV KKDLIKLPK

      1210      1220      1230      1240      1250      1260
YSLFELENGR KRMLASAGEL QKGNELALPS KYVNFLYLAS HYEKLKGSPE DNEQKQLFVE

      1270      1280      1290      1300      1310      1320
QHKHYLDEII EQISEFSKRV ILADANLDKV LSAYNKHHRDK PIREQAENII HLFTLTNLGA

      1330      1340      1350      1360
PAAFKYFDTT IDRKYRSTK EVLDATLIHQ SITGLYETRI DLSQLGGD (配列番号1)

```

10

20

【0044】

本明細書に記載の S p C a s 9 バリエーションは、以下の位置：N 4 9 7、R 6 6 1、Q 6 9 5、Q 9 2 6（またはそれらと類似する位置）の 1 つ以上における突然変異（すなわち、天然アミノ酸の、異なるアミノ酸、例えば、アラニン、グリシン、またはセリンとの置き換え）を有する配列番号 1 のアミノ酸配列を含み得る。一部の実施形態において、S p C a s 9 は、本明細書に記載の突然変位に加えて、配列番号 1 のアミノ酸配列と少なくとも 80%、例えば、少なくとも 85%、90%、または 95% 同一であり、例えば、保存的突然変異により置き換えられている、例えば、配列番号 1 の残基の最大 5%、10%、15%、または 20% における差異を有する。好ましい実施形態において、バリエーションは、親の所望の活性、例えば、ヌクレアーゼ活性（親がニッカーゼまたはデッド C a s 9 である場合を除外する）、ならびにノまたはガイド R N A および標的 D N A）と相互作用する能力を保持する。

30

【0045】

2 つの核酸配列の同一性パーセントを決定するため、最適な比較目的のために配列をアラインする（例えば、最適なアラインメントのため、ギャップを第 1 および第 2 のアミノ酸または核酸配列の一方または両方の中に導入することができ、比較目的のために、非同配列を無視することができる）。比較目的のためにアラインされる参照配列の長さは、参照配列の長さの少なくとも 80% であり、一部の実施形態において、少なくとも 90% または 100% である。次いで、対応するアミノ酸位置またはヌクレオチド位置におけるヌクレオチドを比較する。第 1 の配列中のある位置が、第 2 の配列中の対応する位置と同一のヌクレオチドにより占有される場合、その分子はその位置において同一である（本明細書において使用される核酸「同一性」は、核酸「相同性」に相当する）。2 つの配列間の同一性パーセントは、それらの配列により共有される同一位置の数の関数であり、それ

40

50

ら2つの配列の最適なアラインメントのために導入が必要とされるギャップの数、およびそれぞれのギャップの長さを考慮する。2つのポリペプチドまたは核酸配列間の同一性パーセントは、当技術分野の技能の範囲内の種々の手段で、例えば、公的に入手可能なコンピュータソフトウェア、例えば、Smith Waterman Alignment (Smith, T. F. and M. S. Waterman (1981) J Mol Biol 147:195-7); GeneMatcher Plus (商標) に組み込まれた“BestFit” (Smith and Waterman, Advances in Applied Mathematics, 482-489 (1981)), Schwarz and Dayhof (1979) Atlas of Protein Sequence and Structure, Dayhof, M. O., Ed, pp 353-358; BLASTプログラム (Basic Local Alignment Search Tool; (Altschul, S. F., W. Gish, et al. (1990) J Mol Biol 215:403-10)、BLAST-2、BLAST-P、BLAST-N、BLAST-X、WU-BLAST-2、ALIGN、ALIGN-2、CLUSTAL、またはMegalign (DNASTAR) ソフトウェアを使用して決定する。さらに、当業者は、アラインメントを計測するための適切なパラメータ、例として、比較される配列の長さにわたり最大アラインメントを達成するために必要される任意のアルゴリズムを決定することができる。一般に、タンパク質または核酸について、比較の長さは、全長以下の任意の長さ (例えば、5%、10%、20%、30%、40%、50%、60%、70%、80%、90%、95%、または100%) であり得る。本組成物および方法の目的のため、配列の全長の少なくとも80%がアラインされる。

【0046】

本発明の目的のため、2つの配列間の配列の比較および同一性パーセントの決定は、Blossum 62スコアリング行列をギャップペナルティ12、ギャップ伸長ペナルティ4、およびフレームシフトギャップペナルティ5でを使用して達成することができる。

【0047】

保存的置換としては、典型的には、以下の群：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；およびフェニルアラニン、チロシン内の置換が挙げられる。

【0048】

一部の実施形態において、SpCas9バリエーションは、以下の突然変異のセット：N497A/R661A/Q695/Q926A (四重アラニン突然変異体)；Q695A/Q926A (二重アラニン突然変異体)；R661A/Q695A/Q926AおよびN497A/Q695A/Q926A (三重アラニン突然変異体) の1つを含む。一部の実施形態において、L169および/またはY450における追加の置換突然変異をそれらの二重、三重、および四重突然変異体に追加することができ、またはQ695もしくはQ926における置換を担持する単一突然変異体に追加することができる。一部の実施形態において、突然変異体は、野生型アミノ酸に代えてアラニンを有する。一部の実施形態において、突然変異体は、アルギニンまたはリジン (または天然アミノ酸) 以外の任意のアミノ酸を有する。

【0049】

一部の実施形態において、SpCas9バリエーションは、タンパク質のヌクレアーゼ部分を触媒的に不活性にするため、Cas9のヌクレアーゼ活性を低減させ、または破壊する以下の突然変異：D10、E762、D839、H983、またはD986およびH840またはN863の1つ、例えば、D10A/D10NおよびH840A/H840N/H840Yも含み；それらの位置における置換は、アラニン (Nishimasu et al., Cell 156, 935-949 (2014) におけるとおり)、または他の残基、例えば、グルタミン、アスパラギン、チロシン、セリン、またはアスパラギン酸、例え

ば、E 7 6 2 Q、H 9 8 3 N、H 9 8 3 Y、D 9 8 6 N、N 8 6 3 D、N 8 6 3 S、またはN 8 6 3 Hであり得る（国際公開第2014/152432号パンフレット参照）。一部の実施形態において、バリエーションは、D 1 0 AまたはH 8 4 0 A（一本鎖ニッカーゼを作出する）における突然変異、またはD 1 0 AおよびH 8 4 0 A（ヌクレアーゼ活性を解消する；この突然変異体は、デッドC a s 9またはd C a s 9として公知である）における突然変異を含む。

【0050】

S p C a s 9 N 4 9 7 A / R 6 6 1 A / Q 6 9 5 A / R 9 2 6 A突然変異は、黄色ブドウ球菌（*Staphylococcus aureus*）C a s 9（S a C a s 9）中の類似残基を有し；例えば、図20を参照されたい。DNAまたはRNA骨格に接触する残基の突然変異は、S p C a s 9について観察されたとおり、S a C a s 9の特異性を増加させると予測される。したがって、本明細書においてS a C a s 9バリエーションも提供される。

10

【0051】

S a C a s 9野生型配列は、以下のとおりである。

【化 3】

10	20	30	40	50
MKRNYYILGLD	IGITSVGYGI	IDYETRDVID	AGVRLFKEAN	VENNEGRRSK
60	70	80	90	100
RGARRLKRRR	RHRIQRVKKL	LFDYNLLTDH	SELSGINPYE	ARVKGLSQKL
110	120	130	140	150
SEEEFSAALL	HLAKRRGVHN	VNEVEEDTGN	ELSTKEQISR	NSKALEEKYV
160	170	180	190	200
AELQLERLKK	DGEVRGSINR	EKTSYDVKEA	KQLLKVQKAY	HQLDQSFIDT
210	220	230	240	250
YIDLLETRRT	YYEGPGEKSP	FGWKDIKEWY	EMLMGHCTYF	PEELRSVKYA
260	270	280	290	300
YNADLYNALN	DLNNLVITRD	ENEKLEYEYK	FQIENVFQKQ	KKKPTLQKIA
310	320	330	340	350
KEILVNEEDI	KGYRVTSTGK	PEFTNLKVYH	DIKDITARKE	IIENAEELLDQ
360	370	380	390	400
IAKILTIYQS	SEDIQEELTN	LNSELTQEEI	EQISNLKGYT	GTHNLSLKAI
410	420	430	440	450
NLILDELWHT	NDNQIAIFNR	LKLVPKKVDL	SQQKEIPTTL	VDDFILSPVV
460	470	480	490	500
KRSFIQSIKV	INAIKKYGL	PNDIIIELAR	EKNSKDAQKM	INEMQKRNRR
510	520	530	540	550
TNERIBEIIR	TTGKENAKYL	IEKIKLHDMQ	EGKCLYSLEA	IPLEDLLNNP
560	570	580	590	600
FNYEVDHIIP	RSVSFDNSFN	NKVLVKQEEEN	SKKGNRTPFQ	YLSSSDSKIS
610	620	630	640	650
YETFKKHILN	LAKGKGRISK	TKKEYLLEER	DINRFSVQKD	FINRNLDVDR
660	670	680	690	700
YATRGLMNL	RSYFRVNNLD	VKVKSSINGGF	TSFLRRKWK	KKERNKGYKH
710	720	730	740	750
HAEDALIIAN	ADFIFKEWKK	LDKAKKVMEN	QMFEKQAESE	MPEIETEQEY
760	770	780	790	800
KEIFITPHQI	KHIKDFKDYK	YSHRVDKKPN	RELINDTLYS	TRKDDKGNTL
810	820	830	840	850
IVNNNLGLYD	KDNDKLKCLI	NKSPEKLLMY	HHDPQTYQKL	KLIMEQYGD
860	870	880	890	900
KNPLYKYEE	TGNYLTKYSK	KDNGPVIKKI	KYYGNKLNAH	LDITDDYPNS
910	920	930	940	950
RNKVVKLKSL	PYRFDVYLDN	GVYKFVTVKN	LDVIKKENYY	EVNSKCYEEA
960	970	980	990	1000
KKLKKISNQ	EFIASFYNN	LIKINGELYR	VIGVNNDLLN	RIEVNMIDIT
1010	1020	1030	1040	1050
YREYLENNMD	KRPRIIKTI	ASKTQSIKKY	STDILGNLYE	VKSKKHPQII

KKG (配列番号2)

【0052】

本明細書に記載の S a C a s 9 バリエーションは、以下の位置：Y 2 1 1、W 2 2 9、R 2 4 5、T 3 9 2、N 4 1 9、および/または R 6 5 4 の 1、2、3、4、5、または 6 つ全てにおける突然変異を有する配列番号 2 のアミノ酸配列を含み、例えば、以下の位置：Y 2 1 1、W 2 2 9、R 2 4 5、T 3 9 2、N 4 1 9、および/または R 6 5 4 の 1、2、3、4、5 または 6 つにおける突然変異を有する配列番号 2 のアミノ酸配列と少なくとも 80 % 同一である配列を含む。

【0053】

一部の実施形態において、バリエーション S a C a s 9 タンパク質は、以下の突然変異：Y 2 1 1 A；W 2 2 9 A；Y 2 3 0 A；R 2 4 5 A；T 3 9 2 A；N 4 1 9 A；L 4 4 6 A；Y 6 5 1 A；R 6 5 4 A；D 7 8 6 A；T 7 8 7 A；Y 7 8 9 A；T 8 8 2 A；K 8 8 6 A；N 8 8 8 A；A 8 8 9 A；L 9 0 9 A；N 9 8 5 A；N 9 8 6 A；R 9 9 1 A；R

10

20

30

40

50

1 0 1 5 A ; N 4 4 A ; R 4 5 A ; R 5 1 A ; R 5 5 A ; R 5 9 A ; R 6 0 A ; R 1 1 6 A ; R 1 6 5 A ; N 1 6 9 A ; R 2 0 8 A ; R 2 0 9 A ; Y 2 1 1 A ; T 2 3 8 A ; Y 2 3 9 A ; K 2 4 8 A ; Y 2 5 6 A ; R 3 1 4 A ; N 3 9 4 A ; Q 4 1 4 A ; K 5 7 A ; R 6 1 A ; H 1 1 1 A ; K 1 1 4 A ; V 1 6 4 A ; R 1 6 5 A ; L 7 8 8 A ; S 7 9 0 A ; R 7 9 2 A ; N 8 0 4 A ; Y 8 6 8 A ; K 8 7 0 A ; K 8 7 8 A ; K 8 7 9 A ; K 8 8 1 A ; Y 8 9 7 A ; R 9 0 1 A ; K 9 0 6 A の 1 つ以上も含む。

【 0 0 5 4 】

一部の実施形態において、バリエーション S a C a s 9 タンパク質は、以下の追加の突然変異：Y 2 1 1 A、W 2 2 9 A、Y 2 3 0 A、R 2 4 5 A、T 3 9 2 A、N 4 1 9 A、L 4 4 6 A、Y 6 5 1 A、R 6 5 4 A、D 7 8 6 A、T 7 8 7 A、Y 7 8 9 A、T 8 8 2 A、K 8 8 6 A、N 8 8 8 A、A 8 8 9 A、L 9 0 9 A、N 9 8 5 A、N 9 8 6 A、R 9 9 1 A、R 1 0 1 5 A、N 4 4 A、R 4 5 A、R 5 1 A、R 5 5 A、R 5 9 A、R 6 0 A、R 1 1 6 A、R 1 6 5 A、N 1 6 9 A、R 2 0 8 A、R 2 0 9 A、Y 2 1 1 A、T 2 3 8 A、Y 2 3 9 A、K 2 4 8 A、Y 2 5 6 A、R 3 1 4 A、N 3 9 4 A、Q 4 1 4 A、K 5 7 A、R 6 1 A、H 1 1 1 A、K 1 1 4 A、V 1 6 4 A、R 1 6 5 A、L 7 8 8 A、S 7 9 0 A、R 7 9 2 A、N 8 0 4 A、Y 8 6 8 A、K 8 7 0 A、K 8 7 8 A、K 8 7 9 A、K 8 8 1 A、Y 8 9 7 A、R 9 0 1 A、K 9 0 6 A の 1 つ以上を含む。

【 0 0 5 5 】

一部の実施形態において、バリエーション S a C a s 9 タンパク質は、複数の置換突然変異：R 2 4 5 / T 3 9 2 / N 4 1 9 / R 6 5 4 および Y 2 2 1 / R 2 4 5 / N 4 1 9 / R 6 5 4 (四重バリエーション突然変異体) ; N 4 1 9 / R 6 5 4、R 2 4 5 / R 6 5 4、Y 2 2 1 / R 6 5 4、および Y 2 2 1 / N 4 1 9 (二重突然変異体) ; R 2 4 5 / N 4 1 9 / R 6 5 4、Y 2 1 1 / N 4 1 9 / R 6 5 4、および T 3 9 2 / N 4 1 9 / R 6 5 4 (三重突然変異体) を含む。一部の実施形態において、突然変異体は、野生型アミノ酸に代えてアラニンを含む。

【 0 0 5 6 】

一部の実施形態において、バリエーション S a C a s 9 タンパク質は、E 7 8 2 K、K 9 2 9 R、N 9 6 8 K、および / または R 1 0 1 5 H における突然変異も含む。例えば、K K H バリエーション (E 7 8 2 K / N 9 6 8 K / R 1 0 1 5 H)、K R H バリエーション (E 7 8 2 K / K 9 2 9 R / R 1 0 1 5 H)、または K R K H バリエーション (E 7 8 2 K / K 9 2 9 R / N 9 6 8 K / R 1 0 1 5 H)] である。

【 0 0 5 7 】

一部の実施形態において、バリエーション S a C a s 9 タンパク質は、D 1 0、E 4 7 7、D 5 5 6、H 7 0 1、または D 7 0 4 ; および H 5 5 7 または N 5 8 0 における突然変異からなる群から選択される、ヌクレアーゼ活性を減少させる 1 つ以上の突然変異も含む。

【 0 0 5 8 】

一部の実施形態において、突然変異は、(i) D 1 0 A もしくは D 1 0 N、(i i) H 5 5 7 A、H 5 5 7 N、または H 5 5 7 Y、(i i i) N 5 8 0 A、および / または (i v) D 5 5 6 A である。

【 0 0 5 9 】

本明細書において、任意選択的に、バリエーションタンパク質の発現のための 1 つ以上の調節ドメインに作動可能に結合されている C a s 9 バリエーションをコードする単離核酸、単離核酸を含むベクター、ならびにその核酸を含み、および任意選択的にバリエーションタンパク質を発現する宿主細胞、例えば、哺乳動物宿主細胞も提供される。

【 0 0 6 0 】

本明細書に記載のバリエーションは、細胞のゲノムを変更するために使用することができ ; 本方法は、一般に、細胞中でバリエーションタンパク質を、細胞のゲノムの選択部分に相補的な領域を有するガイド RNA とともに発現させることを含む。細胞のゲノムを選択的に変更する方法は、当技術分野において公知であり、例えば、米国特許第 8, 9 9 3, 2 3 3 号明細書 ; 米国特許出願公開第 2 0 1 4 0 1 8 6 9 5 8 号明細書 ; 米国特許第 9, 0 2 3

10

20

30

40

50

, 649号明細書; 国際公開第2014/099744号パンフレット; 国際公開第2014/089290号パンフレット; 国際公開第2014/144592号パンフレット; 国際公開第144288号パンフレット; 国際公開第2014/204578号パンフレット; 国際公開第2014/152432号パンフレット; 国際公開第2115/099850号パンフレット; 米国特許第8,697,359号明細書; 米国特許出願公開第20160024529号明細書; 米国特許出願公開第20160024524号明細書; 米国特許出願公開第20160024523号明細書; 米国特許出願公開第20160024510号明細書; 米国特許出願公開第20160017366号明細書; 米国特許出願公開第20160017301号明細書; 米国特許出願公開第20150376652号明細書; 米国特許出願公開第20150356239号明細書; 米国特許出願公開第20150315576号明細書; 米国特許出願公開第20150291965号明細書; 米国特許出願公開第20150252358号明細書; 米国特許出願公開第20150247150号明細書; 米国特許出願公開第20150232883号明細書; 米国特許出願公開第20150232882号明細書; 米国特許出願公開第20150203872号明細書; 米国特許出願公開第20150191744号明細書; 米国特許出願公開第20150184139号明細書; 米国特許出願公開第20150176064号明細書; 米国特許出願公開第20150167000号明細書; 米国特許出願公開第20150166969号明細書; 米国特許出願公開第20150159175号明細書; 米国特許出願公開第20150159174号明細書; 米国特許出願公開第20150093473号明細書; 米国特許出願公開第20150079681号明細書; 米国特許出願公開第20150067922号明細書; 米国特許出願公開第20150056629号明細書; 米国特許出願公開第20150044772号明細書; 米国特許出願公開第20150024500号明細書; 米国特許出願公開第20150024499号明細書; 米国特許出願公開第20150020223号明細書; 米国特許出願公開第20140356867号明細書; 米国特許出願公開第20140295557号明細書; 米国特許出願公開第20140273235号明細書; 米国特許出願公開第20140273226号明細書; 米国特許出願公開第20140273037号明細書; 米国特許出願公開第20140189896号明細書; 米国特許出願公開第20140113376号明細書; 米国特許出願公開第20140093941号明細書; 米国特許出願公開第20130330778号明細書; 米国特許出願公開第20130288251号明細書; 米国特許出願公開第20120088676号明細書; 米国特許出願公開第20110300538号明細書; 米国特許出願公開第20110236530号明細書; 米国特許出願公開第20110217739号明細書; 米国特許出願公開第20110002889号明細書; 米国特許出願公開第20100076057号明細書; 米国特許出願公開第20110189776号明細書; 米国特許出願公開第20110223638号明細書; 米国特許出願公開第20130130248号明細書; 米国特許出願公開第20150050699号明細書; 米国特許出願公開第20150071899号明細書; 米国特許出願公開第20150050699号明細書; 米国特許出願公開第20150045546号明細書; 米国特許出願公開第20150031134号明細書; 米国特許出願公開第20150024500号明細書; 米国特許出願公開第20140377868号明細書; 米国特許出願公開第20140357530号明細書; 国際公開第2008/108989号パンフレット; 国際公開第2010/054108号パンフレット; 国際公開第2012/164565号パンフレット; 国際公開第2013/098244号パンフレット; 国際公開第2013/176772号パンフレット; 国際公開第20150071899号パンフレット; 米国特許出願公開第20140349400号明細書; 米国特許出願公開第20140335620号明細書; 米国特許出願公開第20140335063号明細書; 米国特許出願公開第20140315985号明細書; 米国特許出願公開第20140310830号明細書; 米国特許出願公開第20140310828号明細書; 米国特許出願公開第20140309487号明細書; 米国特許出願公開第20140304853号明細書; 米国特許出願公開第20140298547号明細書; 米国特許出願公開第201402

9 5 5 5 6 号明細書；米国特許出願公開第 2 0 1 4 0 2 9 4 7 7 3 号明細書；米国特許出願公開第 2 0 1 4 0 2 8 7 9 3 8 号明細書；米国特許出願公開第 2 0 1 4 0 2 7 3 2 3 4 号明細書；米国特許出願公開第 2 0 1 4 0 2 7 3 2 3 2 号明細書；米国特許出願公開第 2 0 1 4 0 2 7 3 2 3 1 号明細書；米国特許出願公開第 2 0 1 4 0 2 7 3 2 3 0 号明細書；米国特許出願公開第 2 0 1 4 0 2 7 1 9 8 7 号明細書；米国特許出願公開第 2 0 1 4 0 2 5 6 0 4 6 号明細書；米国特許出願公開第 2 0 1 4 0 2 4 8 7 0 2 号明細書；米国特許出願公開第 2 0 1 4 0 2 4 2 7 0 2 号明細書；米国特許出願公開第 2 0 1 4 0 2 4 2 7 0 0 号明細書；米国特許出願公開第 2 0 1 4 0 2 4 2 6 9 9 号明細書；米国特許出願公開第 2 0 1 4 0 2 4 2 6 6 4 号明細書；米国特許出願公開第 2 0 1 4 0 2 3 4 9 7 2 号明細書；米国特許出願公開第 2 0 1 4 0 2 2 7 7 8 7 号明細書；米国特許出願公開第 2 0 1 4 0 2 1 2 8 6 9 号明細書；米国特許出願公開第 2 0 1 4 0 2 0 1 8 5 7 号明細書；米国特許出願公開第 2 0 1 4 0 1 9 9 7 6 7 号明細書；米国特許出願公開第 2 0 1 4 0 1 8 9 8 9 6 号明細書；米国特許出願公開第 2 0 1 4 0 1 8 6 9 5 8 号明細書；米国特許出願公開第 2 0 1 4 0 1 8 6 9 1 9 号明細書；米国特許出願公開第 2 0 1 4 0 1 8 6 8 4 3 号明細書；米国特許出願公開第 2 0 1 4 0 1 7 9 7 7 0 号明細書；米国特許出願公開第 2 0 1 4 0 1 7 9 0 0 6 号明細書；米国特許出願公開第 2 0 1 4 0 1 7 0 7 5 3 号明細書；Makarova et al., "Evolution and classification of the CRISPR-Cas systems" 9(6) Nature Reviews Microbiology 467-477 (1-23) (Jun. 2011); Wiedenheft et al., "RNA-guided genetic silencing systems in bacteria and archaea" 482 Nature 331-338 (Feb. 16, 2012); Gasunas et al., "Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria" 109(39) Proceedings of the National Academy of Sciences USA E2579-E2586 (Sep. 4, 2012); Jin et al., "A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity" 337 Science 816-821 (Aug. 17, 2012); Carroll, "A CRISPR Approach to Gene Targeting" 20(9) Molecular Therapy 1658-1660 (Sep. 2012); 2012年5月25日に出願された米国特許出願第61/652,086号明細書；Al-Attar et al., Clustered Regularly Interspaced Short Palindromic Repeats (CRISPRs): The Hallmark of an Ingenious Antiviral Defense Mechanism in Prokaryotes, Biol Chem. (2011) vol. 392, Issue 4, pp. 277-289; Hale et al., Essential Features and Rational Design of CRISPR RNAs That Function With the Cas RAMP Module Complex to Cleave RNAs, Molecular Cell, (2012) vol. 45, Issue 3, 292-302を参照されたい。

【0061】

本明細書に記載のバリエーションタンパク質は、上記参照文献に記載のCas9タンパク質のいずれかに代えてもしくはそれらに加えて、またはそれらに記載の突然変異との組合せで使用することができる。さらに、本明細書に記載のバリエーションは、当技術分野において公知の野生型Cas9または他のCas9突然変異（例えば、上記のdCas9またはCas9ニッカーゼ）に代えて融合タンパク質、例えば、米国特許第8,993,233号明細書；米国特許出願公開第20140186958号明細書；米国特許第9,023,

10

20

30

40

50

649号明細書；国際公開第2014/099744号パンフレット；国際公開第2014/089290号パンフレット；国際公開第2014/144592号パンフレット；国際公開第144288号パンフレット；国際公開第2014/204578号パンフレット；国際公開第2014/152432号パンフレット；国際公開第2115/099850号パンフレット；米国特許第8,697,359号明細書；米国特許出願公開第2010/0076057号明細書；米国特許出願公開第2011/0189776号明細書；米国特許出願公開第2011/0223638号明細書；米国特許出願公開第2013/0130248号明細書；国際公開第2008/108989号パンフレット；国際公開第2010/054108号パンフレット；国際公開第2012/164565号パンフレット；国際公開第2013/098244号パンフレット；国際公開第2013/176772号パンフレット；米国特許出願公開第20150050699号明細書；米国特許出願公開第20150071899号明細書、および国際公開第2014/124284号パンフレットに記載の異種機能ドメインを有する融合タンパク質中で使用することができる。例えば、バリエーション、好ましくは、1つ以上のヌクレアーゼ低減、変更または殺傷突然変異を含むバリエーションをCas9のNまたはC末端上で、転写活性化ドメインまたは他の異種機能ドメイン（例えば、転写リプレッサー（例えば、KRAB、ERD、SIDなど、例えば、ets2リプレッサー因子（ERF）リプレッサードメイン（ERD）のアミノ酸473～530、KOX1のKRABドメインのアミノ酸1～97、またはMad mSIN3相互作用ドメイン（SID）のアミノ酸1～36；Beerli et al., PNAS USA 95:14628-14633 (1998) 参照）またはサイレンサー、例えば、ヘテロクロマチンタンパク質1（HP1、swi6としても公知）、例えば、HP1 またはHP1；固定RNA結合配列に融合している長鎖非コードRNA（lncRNA）をリクルートし得るタンパク質またはペプチド、例えば、MS2コートタンパク質、エンドリボヌクレアーゼCsy4、またはラムダNタンパク質により結合されるものに融合させることができ；当技術分野において公知のDNAのメチル化状態を改変する酵素（例えば、DNAメチルトランスフェラーゼ（DNMT）またはTETタンパク質）；またはヒストンサブユニットを改変する酵素（例えば、ヒストンアセチルトランスフェラーゼ（HAT）、ヒストンデアセチラーゼ（HDAC）、ヒストンメチルトランスフェラーゼ（例えば、リジンまたはアルギニン残基のメチル化のため）またはヒストンデメチラーゼ（例えば、リジンまたはアルギニン残基の脱メチル化のため））を使用することもできる。多数のこのようなドメインの配列、例えば、DNA中のメチル化シトシンのヒドロキシル化を触媒するドメインが当技術分野において公知である。例示的なタンパク質としては、テンイレブントランスロケーション（Ten-Eleven-Translocation）（TET）1～3ファミリー、DNA中の5-メチルシトシン（5-mC）を5-ヒドロキシメチルシトシン（5-hmC）に変換する酵素が挙げられる。

【0062】

ヒトTET1～3配列は当技術分野において公知であり、それを以下の表に示す。

【0063】

10

20

30

【表 2】

遺伝子	GenBank アクセッション番号	
	アミノ酸	核酸
TET1	NP_085128.2	NM_030625.2
TET2*	NP_001120680.1 (var 1)	NM_001127208.2
	NP_060098.3 (var 2)	NM_017628.4
TET3	NP_659430.1	NM_144993.1

* バリエント(1)は、より長い転写物を表し、より長いアイソフォームをコードする。

(a).バリエント(2)は、バリエント 1 と比較して 5' UTR および 3' UTR およびコード配列が異なる。得られるアイソフォーム(b)はより短く、アイソフォーム a と比較して区別される C 末端を有する。

10

【 0 0 6 4 】

一部の実施形態において、触媒ドメインの全長配列の全部または一部、例えば、システインリッチ伸長部および 7 つの高度に保存されるエキソンによりコードされる 2 O G F e D O ドメイン、例えば、アミノ酸 1 5 8 0 ~ 2 0 5 2 を含む T e t 1 触媒ドメイン、アミノ酸 1 2 9 0 ~ 1 9 0 5 を含む T e t 2、およびアミノ酸 9 6 6 ~ 1 6 7 8 を含む T e t 3 を含む触媒モジュールを含めることができる。3 つ全ての T e t タンパク質中のキー触媒残基を説明するアラインメント、および全長配列についてのその補助材料について (f t p サイト f t p . n c b i . n i h . g o v / p u b / a r a v i n d / D O N S / s u p p l e m e n t a r y _ m a t e r i a l _ D O N S . h t m l で入手可能) は、例えば、I y e r e t a l . , C e l l C y c l e . 2 0 0 9 J u n 1 ; 8 (1 1) : 1 6 9 8 - 7 1 0 . E p u b 2 0 0 9 J u n 2 7 の図 1 を参照されたい (例えば、s e q 2 c 参照) ; 一部の実施形態において、配列は、T e t 1 のアミノ酸 1 4 1 8 ~ 2 1 3 6 または T e t 2 / 3 中の対応領域を含む。

20

【 0 0 6 5 】

他の触媒モジュールは、I y e r e t a l . , 2 0 0 9 において同定されたタンパク質からのものであり得る。

30

【 0 0 6 6 】

一部の実施形態において、異種機能ドメインは、生物学的係留物であり、M S 2 コートタンパク質、エンドリボヌクレアーゼ C s y 4、またはラムダ N タンパク質の全部または一部 (例えば、それからの DNA 結合ドメイン) を含む。これらのタンパク質は、特異的ステム - ループ構造を含有する RNA 分子を d C a s 9 g R N A 標的化配列により規定される場所にリクルートするための使用することができる。例えば、M S 2 コートタンパク質、エンドリボヌクレアーゼ C s y 4、またはラムダ N に融合している d C a s 9 バリエントは、C s y 4、M S 2 またはラムダ N 結合配列に結合している長鎖非コード RNA (l n c R N A)、例えば、X I S T または H O T A I R をリクルートするために使用することができる ; 例えば、K e r y e r - B i b e n s e t a l . , B i o l . C e l l 1 0 0 : 1 2 5 - 1 3 8 (2 0 0 8) を参照されたい。あるいは、C s y 4、M S 2 またはラムダ N タンパク質結合配列は、例えば、K e r y e r - B i b e n s e t a l . , 前掲に記載のとおり、別のタンパク質に結合させることができ、本明細書に記載の方法および組成物を使用してタンパク質を d C a s 9 バリエント結合部位に標的化させることができる。一部の実施形態において、C s y 4 は、触媒的に不活性である。一部の実施形態において、C a s 9 バリエント、好ましくは、d C a s 9 バリエントは、米国特許第 8 , 9 9 3 , 2 3 3 号明細書 ; 米国特許出願公開第 2 0 1 4 0 1 8 6 9 5 8 号明細書 ; 米国特許第 9 , 0 2 3 , 6 4 9 号明細書 ; 国際公開第 2 0 1 4 / 0 9 9 7 4 4 号パンフレット ; 国際公開第 2 0 1 4 / 0 8 9 2 9 0 号パンフレット ; 国際公開第 2 0 1 4 / 1 4

40

50

4592号パンフレット；国際公開第144288号パンフレット；国際公開第2014/204578号パンフレット；国際公開第2014/152432号パンフレット；国際公開第2115/099850号パンフレット；米国特許第8,697,359号明細書；米国特許出願公開第2010/0076057号明細書；米国特許出願公開第2011/0189776号明細書；米国特許出願公開第2011/0223638号明細書；米国特許出願公開第2013/0130248号明細書；国際公開第2008/108989号パンフレット；国際公開第2010/054108号パンフレット；国際公開第2012/164565号パンフレット；国際公開第2013/098244号パンフレット；国際公開第2013/176772号パンフレット；米国特許出願公開第20150050699号明細書；米国特許出願公開第20150071899号明細書、および国際公開第2014/204578号パンフレットに記載のとおり、F o k Iに融合している。

10

【0067】

一部の実施形態において、融合タンパク質は、d C a s 9バリエーションおよび異種機能ドメイン間のリンカーを含む。これらの融合タンパク質中（または連結構造の融合タンパク質間）で 사용할 ことができるリンカーは、融合タンパク質の機能を妨害しない任意の配列を含み得る。好ましい実施形態において、リンカーは、短鎖であり、例えば、2～20アミノ酸であり、典型的には、フレキシブルである（すなわち、高い自由度を有するアミノ酸、例えば、グリシン、アラニン、およびセリンを含む）。一部の実施形態において、リンカーは、G G G S（配列番号3）またはG G G S（配列番号4）からなる1つ以上の単位、例えば、G G G S（配列番号5）またはG G G S（配列番号6）単位の2、3、4つ、またはそれより多いリピートを含む。他のリンカー配列を使用することもできる。

20

【0068】

一部の実施形態において、バリエーションタンパク質は、細胞内空間への送達を容易にする細胞浸透ペプチド配列、例えば、H I V由来T A Tペプチド、ペネトラチン、トランスポータータン、またはh C T由来細胞浸透ペプチドを含み、例えば、C a r o n e t a l . , (2 0 0 1) M o l T h e r . 3 (3) : 3 1 0 - 8 ; L a n g e l , C e l l - P e n e t r a t i n g P e p t i d e s : P r o c e s s e s a n d A p p l i c a t i o n s (C R C P r e s s , B o c a R a t o n F L 2 0 0 2) ; E l - A n d a l o u s s i e t a l . , (2 0 0 5) C u r r P h a r m D e s . 1 1 (2 8) : 3 5 9 7 - 6 1 1 ; およびD e s h a y e s e t a l . , (2 0 0 5) C e l l M o l L i f e S c i . 6 2 (1 6) : 1 8 3 9 - 4 9を参照されたい。

30

【0069】

細胞浸透ペプチド（C P P）は、細胞膜を通過する細胞質または他の細胞小器官、例えば、ミトコンドリアおよび核中への広範な生体分子の移動を容易にする短鎖ペプチドである。C P Pにより送達することができる分子の例としては、治療薬、プラスミドDNA、オリゴヌクレオチド、s i R N A、ペプチド核酸（P N A）、タンパク質、ペプチド、ナノ粒子、およびリポソームが挙げられる。C P Pは、一般に、30アミノ酸以下であり、天然または非天然存在タンパク質またはキメラ配列に由来し、高い相対存在量の正荷電アミノ酸、例えば、リジンもしくはアルギニン、または交互パターンの極性および非極性アミノ酸を含有する。当技術分野において一般に使用されるC P Pとしては、T a t (F r a n k e l e t a l . , (1 9 8 8) C e l l . 5 5 : 1 1 8 9 - 1 1 9 3 , V i v e s e t a l . , (1 9 9 7) J . B i o l . C h e m . 2 7 2 : 1 6 0 1 0 - 1 6 0 1 7)、ペネトラチン（D e r o s s i e t a l . , (1 9 9 4) J . B i o l . C h e m . 2 6 9 : 1 0 4 4 4 - 1 0 4 5 0）、ポリアルギニンペプチド配列（W e n d e r e t a l . , (2 0 0 0) P r o c . N a t l . A c a d . S c i . U S A 9 7 : 1 3 0 0 3 - 1 3 0 0 8 , F u t a k i e t a l . , (2 0 0 1) J . B i o l . C h e m . 2 7 6 : 5 8 3 6 - 5 8 4 0）、およびトランスポーター（P o o g a e t a l . , (1 9 9 8) N a t . B i o t e c h n o l . 1 6 : 8 5 7 - 8 6 1）が挙

40

50

げられる。

【0070】

CPPは、共有または非共有結合方針を介してその積み荷と結合させることができる。CPPおよびその積み荷を共有結合させる方法は当技術分野において公知であり、例えば、化学的架橋 (Stetsenko et al., (2000) J. Org. Chem. 65:4900-4909, Gait et al. (2003) Cell. Mol. Life. Sci. 60:844-853) または融合タンパク質のクローニング (Nagahara et al., (1998) Nat. Med. 4:1449-1453) である。積み荷と、極性および非極性ドメインを含む短鎖両親媒性CPPとの間の非共有カップリングは、静電および疎水性相互作用を介して確立される。

10

【0071】

CPPは、潜在的に治療的な生体分子を細胞中に送達するために当技術分野において利用されている。例としては、免疫抑制のためのポリアルギニンに結合しているシクロsporin (Rothbard et al., (2000) Nature Medicine 6(11):1253-1257)、腫瘍発生の阻害のためのMPGと呼ばれるCPPに結合しているサイクリンB1に対するsiRNA (Crombez et al., (2007) Biochem Soc. Trans. 35:44-46)、癌細胞成長を低減させるためのCPPに結合している腫瘍抑制因子p53ペプチド (Takenobu et al., (2002) Mol. Cancer Ther. 1(12):1043-1049, Snyder et al., (2004) PLoS Biol. 2:E36)、および喘息を治療するためのTatに融合しているRasまたはホスホイノシトール3キナーゼ (PI3K) のドミナントネガティブ型 (Myou et al., (2003) J. Immunol. 171:4399-4405) が挙げられる。

20

【0072】

CPPは、当技術分野において、造影およびバイオセンシング用途のために造影剤を細胞中に輸送するために利用されている。例えば、Tatに付着している緑色蛍光タンパク質 (GFP) が、癌細胞を標識するために使用されている (Shokolenko et al., (2005) DNA Repair 4(4):511-518)。量子ドットにコンジュゲートしているTatが、ラット脳の可視化のために血液脳関門を良好に通過させるために使用されている (Santra et al., (2005) Chem. Commun. 3144-3146)。CPPは、細胞画像化のために磁気共鳴画像化技術と組み合わせることもできる (Liu et al., (2006) Biochem. and Biophys. Res. Comm. 347(1):133-140)。Ramsey and Flynn, Pharmacol Ther. 2015 Jul 22. pii: S0163-7258(15)00141-2も参照されたい。

30

【0073】

代替的にまたはさらに、バリエーションタンパク質は、核局在化配列、例えば、SV40ラージT抗原NLS (PKKKRRV (配列番号7)) およびヌクレオプラスミンNLS (KRPAATKKAGQAKKKK (配列番号8)) を含む得る。他のNLSは、当技術分野において公知であり；例えば、Cokol et al., EMBO Rep. 2000 Nov 15; 1(5):411-415; Freitas and Cunha, Curr Genomics. 2009 Dec; 10(8):550-557を参照されたい。

40

【0074】

一部の実施形態において、バリエーションは、リガンドについての高い親和性を有する部分、例えば、GST、FLAGまたはヘキサヒスチジン配列を含む。このような親和性タグは、組換えバリエーションタンパク質の精製を容易にし得る。

【0075】

バリエーションタンパク質を細胞に送達する方法について、当技術分野において公知の任意の方法を使用して、例えば、バリエーションタンパク質コードする核酸からのインビトロ翻訳

50

、または好適な宿主細胞中の発現により、タンパク質を産生することができ；タンパク質産生についての当技術分野において多数の方法が公知である。例えば、タンパク質は、酵母、大腸菌（*E. coli*）、昆虫細胞系、植物、トランスジェニック動物、または培養哺乳動物細胞中で産生し、それから精製することができ；例えば、Palomares et al., "Production of Recombinant Proteins: Challenges and Solutions," *Methods Mol Biol.* 2004; 267: 15 - 52を参照されたい。さらに、バリエーションタンパク質は、任意選択的に、タンパク質が細胞内に存在すると開裂されるリンカーを用いて細胞中への移動を容易にする部分、例えば、脂質ナノ粒子に結合させることができる。例えば、LaFountaine et al., *Int J Pharm.* 2015 Aug 13; 494 (1): 180 - 194を参照されたい。

10

【0076】

発現系

本明細書に記載のCas9バリエーションを使用するため、それらをコードする核酸からそれらを発現させることが望ましいことがある。これは、種々の手段で実施することができる。例えば、Cas9バリエーションをコードする核酸を、複製および/または発現のための原核または真核細胞中への形質転換のための中間ベクター中にクローニングすることができる。中間ベクターは、典型的には、Cas9バリエーションの産生のためのCas9バリエーションをコードする核酸の貯蔵または操作のための原核生物ベクター、例えば、プラスミド、もしくはシャトルベクター、または昆虫ベクターである。Cas9バリエーションをコードする核酸は、植物細胞、動物細胞、好ましくは、哺乳動物細胞もしくはヒト細胞、真菌細胞、細菌細胞、または原虫細胞への投与のための発現ベクター中にクローニングすることもできる。

20

【0077】

発現を得るため、Cas9バリエーションをコードする配列は、典型的には、転写を指向するためのプロモーターを含有する発現ベクター中にサブクローニングする。好適な細菌および真核生物プロモーターは、当技術分野において周知であり、例えば、Sambrook et al., *Molecular Cloning, A Laboratory Manual* (3rd ed., 2001); Kriegler, *Gene Transfer and Expression: A Laboratory Manual* (1990); および *Current Protocols in Molecular Biology* (Ausubel et al., eds., 2010) に記載されている。遺伝子操作タンパク質を発現させるための細菌発現系は、例えば、大腸菌（*E. coli*）、バシラス属種（*Bacillus* sp.）、およびサルモネラ属（*Salmonella*）において利用可能である（Palva et al., 1983, *Gene* 22: 229 - 235）。このような発現系のためのキットは、市販されている。哺乳動物細胞、酵母、および昆虫細胞のための真核生物発現系も、当技術分野において周知であり、それらも市販されている。

30

【0078】

核酸の発現を指向するために使用されるプロモーターは、特定の用途に依存する。例えば、強力な構成的プロモーターは、典型的には、融合タンパク質の発現および精製に使用される。対照的に、Cas9バリエーションを遺伝子調節のためにインビボ投与すべき場合、Cas9バリエーションの特定の使用に応じて構成的または誘導性プロモーターのいずれかを使用することができる。さらに、Cas9バリエーションの投与に好ましいプロモーターは、弱いプロモーター、例えば、HSV TKまたは類似の活性を有するプロモーターであり得る。プロモーターは、トランス活性化に応答性であるエレメント、例えば、低酸素応答エレメント、Gal4応答エレメント、lacリプレッサー応答エレメント、ならびに小分子制御系、例えば、テトラサイクリン調節系およびRU-486系も含み得る（例えば、Gossen & Bujard, 1992, *Proc. Natl. Acad. Sci. USA*, 89: 5547; Oligino et al., 1998, *Gene* The

40

50

r. , 5 : 491 - 496 ; Wang et al. , 1997 , Gene Ther . , 4 : 432 - 441 ; Neering et al. , 1996 , Blood , 88 : 1147 - 55 ; および Rendahl et al. , 1998 , Nat . Biotechnol . , 16 : 757 - 761 参照) 。

【 0079 】

プロモーターに加え、発現ベクターは、典型的には、原核生物または真核生物のいずれであれ、宿主細胞中の核酸の発現に要求される全ての追加のエLEMENTを含有する転写単位または発現カセットを含有する。したがって、典型的な発現カセットは、例えば、Cas9バリエーションをコードする核酸配列に作動可能に結合されているプロモーター、および例えば、転写物の効率的なポリアダニル化、転写終結、リボソーム結合部位、または翻訳終結に要求される任意のシグナルを含有する。カセットの追加のエLEMENTとしては、例えば、エンハンサー、および異種スプライシングイントロンスイグナルを挙げることができる。

10

【 0080 】

遺伝子情報を細胞中に輸送するために使用される特定の発現ベクターは、Cas9バリエーションの目的の使用、例えば、植物、動物、細菌、真菌、原生動物中などでの発現に関して選択される。標準的な細菌発現ベクターとしては、pBR322ベースのプラスミド、pSKF、pET23Dなどのプラスミド、ならびに市販のタグ融合発現系、例えば、GSTおよびLacZが挙げられる。

【 0081 】

20

真核発現ベクターでは、真核生物ウイルスからの調節ELEMENTを含有する発現ベクター、例えば、SV40ベクター、パピローマウイルスベクター、およびエプスタイン・バーウイルス由来のベクターを使用することが多い。他の例示的な真核生物ベクターとしては、pMSG、pAV009/A+、pMTO10/A+、pMAMneo-5、バキュロウイルスpDSVE、およびSV40初期プロモーター、SV40後期プロモーター、メタロチオネインプロモーター、ネズミ乳腺腫瘍ウイルスプロモーター、ラウス肉腫ウイルスプロモーター、ポリヘドリンプロモーターまたは真核細胞中での発現に有効であることが示されている他のプロモーターの指向下でタンパク質の発現を可能とする他の任意のベクターが挙げられる。

【 0082 】

30

Cas9バリエーションを発現させるためのベクターは、ガイドRNAの発現を駆動するRNA PolIIIプロモーター、例えば、H1、U6または7SKプロモーターを含み得る。これらのヒトプロモーターは、プラスミド形質移入後に哺乳動物細胞中でのCas9バリエーションの発現を可能とする。

【 0083 】

一部の発現系は、安定的に形質移入された細胞系を選択するためのマーカー、例えばチミジンキナーゼ、ハイグロマイシンBホスホトランスフェラーゼ、およびジヒドロ葉酸レダクターゼを有する。高収率の発現系、例えば、バキュロウイルスベクターを昆虫細胞中で、ポリヘドリンプロモーターまたは他の強力なバキュロウイルスプロモーターの指向下でgRNAコード配列とともに使用するものも好適である。

40

【 0084 】

発現ベクター中に典型的に含まれるELEMENTとしては、大腸菌(E. coli)中で機能するレプリコン、組換えプラスミドを保有する細菌の選択を可能にする抗生物質耐性をコードする遺伝子、および組換え配列の挿入を可能にするためのプラスミドの非必須領域中のユニーク制限部位も挙げられる。

【 0085 】

標準的な形質移入法を使用して大量のタンパク質を発現する細菌、哺乳動物、酵母または昆虫の細胞系を産生し、次いで標準的な技術を用いてそのタンパク質を精製する(例えば、Colley et al. , 1989 , J . Biol . Chem . , 264 : 17619 - 22 ; Guide to Protein Purification , in

50

Methods in Enzymology, vol. 182 (Deutscher, ed., 1990) 参照)。真核および原核細胞の形質転換は、標準的な技術に従って実施する(例えば、Morrison, 1977, J. Bacteriol. 132:349-351; Clark-Curtiss & Curtiss, Methods in Enzymology 101:347-362 (Wu et al., eds, 1983 参照)。

【0086】

宿主細胞内に外来ヌクレオチド配列を導入する公知の手順のいずれかを使用することができる。このような方法としては、リン酸カルシウム形質移入、ポリブレン、プロトプラスト融合、エレクトロポレーション、ヌクレオフェクション、リボソーム、マイクロインジェクション、ネイキッドDNA、プラスミドベクター、ウイルスベクター(エピソード型および組込み型の両方)およびクローニングされたゲノムDNA、cDNA、合成DNAまたは他の外来遺伝子材料を宿主細胞内に導入する他の周知の方法のいずれかの使用が挙げられる(例えば、Sambrook et al., 前掲参照)。唯一必要なことは、使用される特定の遺伝子操作手順が、Cas9バリエーションを発現し得る宿主細胞内に少なくとも1つの遺伝子を良好に導入し得ることである。

【0087】

本方法は、精製Cas9タンパク質をgRNAと細胞中にリボ核タンパク質(RNP)複合体として導入することにより、およびgRNAとCas9タンパク質をコードするmRNAを導入することによりgDNAを改変することも含み得る。gRNAは、合成gRNAまたはガイドRNAをコードする核酸(例えば、発現ベクター中)であり得る。

【0088】

本発明は、ベクターおよびベクターを含む細胞も含む。

【実施例】

【0089】

本発明を以下の実施例においてさらに説明するが、それらの実施例は特許請求の範囲に記載の本発明の範囲を限定するものではない。

【0090】

方法

SpCas9バリエーションを進化させるための細菌ベース陽性選択アッセイ

陽性選択プラスミド(標的部位が埋め込まれた)を含有するコンピテント大腸菌(E. coli) BW25141(DE3)²³をCas9/sgRNAコードプラスミドにより形質転換した。SOB培地中の60分間のリカバリー後、クロラムフェニコール(非選択)またはクロラムフェニコール+10mMのアラビノース(選択)のいずれかを含有するLB培地上で形質転換物をプレーティングした。

【0091】

ゲノムワイド標的的特異性について重要であり得る追加の位置を同定するため、ホーミングエンドヌクレアーゼの特性を研究するために既に使用されている細菌選択系(以下、陽性選択と称される)(Chen & Zhao, Nucleic Acids Res 33, e154(2005); Doyon et al., J Am Chem Soc 128, 2477-2484(2006))を適合させた。

【0092】

この系の本適合において、誘導性毒性遺伝子をコードする陽性選択プラスミドのCas9媒介開裂は、線形化プラスミドの後続の分解および損失に起因する細胞生存を可能とする。SpCas9が陽性選択系中で機能し得ることを確立した後、野生型およびバリエーションの両方を、公知のヒトゲノムから選択される標的部位を保有する選択プラスミドを開裂するそれらの能力について試験した。標的部位を含有する陽性選択プラスミドを有する細菌中にこれらのバリエーションを導入し、選択培地上でプレーティングした。陽性選択プラスミドの開裂は、生存頻度: 選択プレート上のコロニー/非選択プレート上のコロニーを計算することにより推定した(図1, 5~6参照)。

【 0 0 9 3 】

【 表 3 】

この試験において使用されたプラスミドのサブセット(配列は以下に示す)

名称	Addgene ID	説明
JDS246	43861	CMV-T7-ヒト SpCas9-NLS-3xFLAG
VP12	係属中	CMV-T7-ヒト SpCas9-HF1(N497A, R661A, Q695A, Q926A)-NLS-3xFLAG
MSP2135	係属中	CMV-T7-ヒト SpCas9-HF2(N497A, R661A, Q695A, Q926A, D1135E)-NLS-3xFLAG
MSP2133	係属中	CMV-T7-ヒト SpCas9-HF4(Y450A, N497A, R661A, Q695A, Q926A)-NLS-3xFLAG
MSP469	65771	CMV-T7-ヒト SpCas9-VQR(D1135V, R1335Q, T1337R)-NLS-3xFLAG
MSP2440	係属中	CMV-T7-ヒト SpCas9-VQR-HF1(N497A, R661A, Q695A, Q926A, D1135V, R1335Q, T1337R)-NLS-3xFLAG
BPK2797	係属中	CMV-T7-ヒト SpCas9-VRQR(D1135V, G1218R, R1335Q, T1337R)-NLS-3xFLAG
MSP2443	係属中	CMV-T7-ヒト SpCas9-VRQR-HF1(N497A, R661A, Q695A, Q926A, D1135V, G1218R, R1335Q, T1337R)-NLS-3xFLAG
BPK1520	65777	U6-BsmBI カセット-Sp-sgRNA

10

20

【 0 0 9 4 】

ヒト細胞培養および形質移入

構成的に発現されるEGFP-PESTレポーター遺伝子¹⁵の単一インテグレートコピーを保有するU2OS・EGFP細胞を、10%のFBS、2mMのGlutaMax(Life Technologies)、ペニシリン/ストレプトマイシン、および400μg/mlのG418が補給されたAdvanced DMEM培地(Life Technologies)中で、37℃において5%CO₂を用いて培養した。Lonza 4D-nucleofectorのDN-100プログラムを製造業者のプロトコルに従って使用して、750ngのCas9プラスミドおよび250ngのsgRNAプラスミド(特に注釈のない限り)により、細胞を同時形質移入した。空のU6プロモータープラスミドと一緒に形質移入されるCas9プラスミドを全てのヒト細胞実験についての陰性対照として使用した(図2、7~10参照)。

30

【 0 0 9 5 】

ヒト細胞EGFP崩壊アッセイ

EGFP崩壊実験を既に記載のとおり実施した¹⁶。EGFP発現について、形質移入された細胞を形質移入の約52時間後にFortessaフローサイトメーター(BD Biosciences)を使用して分析した。バックグラウンドEGFP損失を全ての実験について約2.5%においてゲーティングした(図2、7参照)。

40

【 0 0 9 6 】

ヌクレアーゼ誘導突然変異率を定量するためのT7E1アッセイ、標的化ディープシーケンシング、およびGUIDE-seq

T7E1アッセイを、ヒト細胞について既に記載のとおり(Klein et al., Nature 523, 481-485 (2015))実施した。U2OS・EGFPヒト細胞について、Agencourt DNAdvance Genomic DNA Isolation Kit(Beckman Coulter Genomics)を使用して形質移入の約72時間後にゲノムDNAを形質移入細胞から抽出した。約200ngの精製PCR産物を変性させ、アニーリングし、T7E1(New England Biolabs)により消化した。ヒト細胞について既に記載

50

のとおり (Kleinstiver et al., Nature 523, 481 - 485 (2015); Reyon et al., Nat Biotechnol 30, 460 - 465 (2012)), Qiaxcel キャピラリー電気泳動機器 (QIAGEN) を使用して突然変異誘発頻度を定量した。

【0097】

GUIDE-seq 実験は、既に記載のとおり実施した (Tsai et al., Nat Biotechnol 33, 187 - 197 (2015))。簡潔に述べると、リン酸化ホスホロチオエート修飾二本鎖オリゴデオキシヌクレオチド (dsODN) を上記のとおり Cas9 および sgRNA 発現プラスミドとともに Cas9 スクレアーゼとともに U2OS 細胞中に形質移入した。dsODN 特異的増幅、ハイスループットシーケンシング、およびマッピングを実施して DSB 活性を含有するゲノム間隔を同定した。野生型対二重または四重変異体バリエーション実験のため、オフターゲットリードカウントをオンターゲットリードカウントに正規化して、試料間のシーケンシング深度の差を補正した。次いで、野生型およびバリエーション SpCas9 についての正規化比を比較してオフターゲット部位における活性の変化倍率を計算した。GUIDE-seq についての野生型およびバリエーション試料が、設定標的部位における類似のオリゴタグインテグレーション率を有するか否かを決定するため、100 ng のゲノム DNA (上記のとおり単離) から Phusion Hot-Start Flex により設定標的遺伝子座を増幅することにより制限断片長多型 (RFLP) アッセイを実施した。ほぼ 150 ng の PCR 産物を 20 U の NdeI (New England Biolabs) により 37 °C において 3 時間消化してから、Agencourt Ampure XP キットを使用してクリーンアップした。Qiaxcel キャピラリー電気泳動装置 (QIAGEN) を使用して RFLP 結果を定量してオリゴタグインテグレーション率を推定した。T7E1 アッセイを上記のとおり同様の目的のために実施した。

【0098】

実施例 1

CRISPR-Cas9 RNA ガイド遺伝子編集の標的化特異性に対処するための 1 つの潜在的な解決法は、新規突然変異を有する Cas9 バリエーションを遺伝子操作することである。

【0099】

これらの従来結果に基づき、CRISPR-Cas9 スクレアーゼの特異性は、DNA 上のリン酸基への結合により媒介される DNA についての Cas9 の非特異的結合親和性または DNA との疎水性もしくは塩基スタッキング相互作用を低減させることにより有意に増加させ得ることが仮定された (理論により拘束されるものではない)。このアプローチは、既に記載されたトランケート gRNA アプローチのような gRNA / Cas9 複合体により認識される標的部位の長さを減少させないという利点を有する。DNA についての Cas9 の非特異的結合親和性は、標的 DNA 上のリン酸基に接触するアミノ酸残基を突然変異させることにより低減させ得ることが推定された。

【0100】

非 Cas9 スクレアーゼ、例えば、TALEN のバリエーションを作出するための類似のアプローチが使用されている (例えば、Guillinger et al., Nat. Methods, 11: 429 (2014) 参照)。

【0101】

この仮定の最初の試験において、本発明者らは、DNA 骨格上のリン酸と相互作用することが予測することができる SpCas9 中の種々の残基中に個々のアラニン置換を導入することにより、広く使用される化膿性連鎖球菌 (S. pyogenes) Cas9 (SpCas9) の親和性低減バリエーションを遺伝子操作することを試行した。大腸菌 (E. coli) ベーススクリーニングアッセイを使用してこれらのバリエーションの活性を評価した (Kleinstiver et al., Nature, 2015 Jul 23; 523 (7561): 481 - 5)。この細菌系において、細胞生存率は、毒性ギラーゼ毒

素 c c d B についての遺伝子ならびに g R N A および S p C a s 9 により標的化される 2 3 塩基対配列を含有する選択プラスミドの開裂（および後続の崩壊）に依存した。この実験の結果は、活性を保持または損失する残基を同定した（表 1）。

【 0 1 0 2 】

【表 4】

表 1: 図 1 に示される細菌細胞ベース系において評価された Cas9 の単一アラニン置換突然変異体の活性

突然変異	%生存	突然変異	%生存	突然変異	%生存
R63A	84.2	Q926A	53.3	K1158A	46.5
R66A	0	K1107A	47.4	K1185A	19.3
R70A	0	E1108A	40.0	K1200A	24.5
R74A	0	S1109A	96.6	S1216A	100.4
R78A	56.4	K1113A	51.8	Q1221A	98.8
R165A	68.9	R1114A	47.3	K1289A	55.2
R403A	85.2	S1116A	73.8	R1298A	28.6
N407A	97.2	K1118A	48.7	K1300A	59.8
N497A	72.6	D1135A	67.2	K1325A	52.3
K510A	79.0	S1136A	69.2	R1333A	0
Y515A	34.1	K1151A	0	K1334A	87.5
R661A	75.0	K1153A	76.6	R1335A	0
Q695A	69.8	K1155A	44.6	T1337A	64.6

【 0 1 0 3 】

50 ~ 100 % の生存割合は通常ロバストな開裂を示す一方、0 % の生存率は酵素が機能的に減衰したことを示した。細菌中でアッセイした（しかし、上記表に示されない）追加の突然変異としては、R 6 9 A、R 7 1 A、Y 7 2 A、R 7 5 A、K 7 6 A、N 7 7 A、R 1 1 5 A、H 1 6 0 A、K 1 6 3 A、L 1 6 9 A、T 4 0 4 A、F 4 0 5 A、R 4 4 7 A、I 4 4 8 A、Y 4 5 0 A、S 4 6 0 A、M 4 9 5 A、M 6 9 4 A、H 6 9 8 A、Y 1 0 1 3 A、V 1 0 1 5 A、R 1 1 2 2 A、K 1 1 2 3 A、および K 1 1 2 4 A が挙げられる。R 6 9 A および F 4 0 5 A（細菌中の < 5 % の生存率を有した）を除き、それらの追加の単一突然変異の全ては、S p C a s 9 のオンターゲット活性に対する効果をほとんど有さないと考えられた（細菌スクリーンにおいて > 70 % の生存率）。

【 0 1 0 4 】

N 4 9 7 A、R 6 6 1 A、Q 6 9 5 A、および Q 9 2 6 A 突然変異の全ての考えられる単一、二重、三重および四重組合せを担持する 15 個の異なる S p C a s 9 バリエーションを構築して、それらの残基により作製される接触がオンターゲット活性について欠失可能であり得るか否かを試験した（図 1 b）。これらの実験について、単一インテグレート E G F P レポーター遺伝子内の開裂および非相同末端結合（N H E J）媒介修復による挿入または欠失突然変異（インデル）の誘導が細胞蛍光の損失をもたらす既に記載されたヒト細胞ベースアッセイを使用した（Reyon, D. et al., Nat Biotechnol. 30, 460 - 465, 2012）。野生型 S p C a s 9 と対合させた場合、ヒト細胞中で E G F P 発現を効率的に崩壊することが既に示されている E G F P 標的化 s g R N A を使用すると（Fu, Y. et al., Nat Biotechnol. 31, 822 - 826 (2013)、15 個全ての S p C a s 9 バリエーションは野生型 S p C a s 9 のものと同等の E G F P 崩壊活性を有した（図 1 b、灰色バー）。したがって、これらの残基の 1 つまたは全ての置換は、この E G F P 標的化 s g R N A を用いた S p C a s 9 のオンターゲット開裂効率を低減させなかった。

【 0 1 0 5 】

次に、ミスマッチ標的部位における 15 個全ての S p C a s 9 バリエーションの相対活性を

評価するための実験を実施した。これを行うため、13および14、15および16、17および18、ならびに18および19位（最もPAMから近位の塩基について1から出発し、最もPAMから遠位の塩基について20で終わる番号付与；図1b）における置換塩基のペアを含有する先の実験において使用されたEGFP標的化sgRNAの誘導体を用いてEGFP崩壊アッセイを繰り返した。この分析により、三重突然変異体の1つ（R661A/Q695A/Q926A）および四重突然変異体（N497A/R661A/Q695A/Q926A）の両方が、ミスマッチsgRNAの4つ全てを用いてバックグラウンドのものと同等のEGFP崩壊のレベルを示すことが明らかになった（図1b、着色バー）。特に、15個のバリエーションのうち、ミスマッチsgRNAを用いて最低活性を有するものは、全て、Q695AおよびQ926A突然変異を保有した。これらの結果および別のEGFP標的部位についてのsgRNAを使用した実験からの類似データに基づき、四重突然変異体（N497A/R661A/Q695A/Q926A）を追加の分析のために選択し、それをSpCas9-HF1（高フィデリティバリエーション#1について）と命名した。

10

【0106】

SpCas9-HF1のオンターゲット活性

より多数のオンターゲット部位においていかにロバストにSpCas9-HF1が機能するかを決定するため、追加のsgRNAを使用してこのバリエーションと野生型SpCas9との間の直接比較を実施した。合計で37個の異なるsgRNAを試験した：24個はEGFPに標的化され（EGFP崩壊アッセイを用いてアッセイする）13個は内在性ヒト遺伝子標的に標的化される（T7エンドヌクレアーゼI（T7EI）ミスマッチアッセイを使用してアッセイする）。EGFP崩壊アッセイを用いて試験した24個のsgRNAの20個（図1c）および内在性ヒト遺伝子部位に対して試験された13個のsgRNAの12個（図1d）は、SpCas9-HF1を用いて、同一のsgRNAを用いて野生型SpCas9の活性の少なくとも70%である活性を示した（図1e）。実際、SpCas9-HF1は、大部分のsgRNAを用いて野生型SpCas9と高度に同等の活性（90~140%）を示した（図1e）。試験した37個のsgRNAの3つは、SpCas9-HF1を用いて本質的に活性を示さず、それらの標的部位の試験は、高い活性が見られたものと比較してそれらの配列の特徴のいかなる明らかな差も示唆しなかった（表3）。概して、SpCas9-HF1は、試験したsgRNAの86%（32個/37個）について同等の活性を有した（野生型SpCas9活性の70%超）。

20

30

【0107】

【表 5】

表 3: sgRNA 標的のリスト

化膿性連鎖球菌(<i>S. pyogenes</i>) sgRNA						
EGFP						
予備 名称	名称	スパーサー 長さ(nt)	スパーサー配列	配列番号	拡張 PAM を有する配列	配列番号
FYF1 320	NGG 部位 1	20	GGGCACGGGC AGCTTGCCGG	9.	GGGCACGGGCAGCTTGC CGGTGGT	10.
FYF1 641	NGG 部位 1	18	GCACGGGCAG CTTGCCGG	11.	GCACGGGCAGCTTGCCG GTGGT	12.
CK10 12	NGG 部 位 1-13 および 14	20	GGGCACccGCA GCTTGCCGG	13.	GGGCACccGCAGCTTGCC GGTGGT	14.
FYF1 429	NGG 部 位 1-15 および 16	20	GGGctgGGGCA GCTTGCCGG	15.	GGGctgGGGCAGCTTGCC GGTGGT	16.
FYF1 430	NGG 部 位 1-17 および 18	20	GGcgACGGGCA GCTTGCCGG	17.	GGcgACGGGCAGCTTGCC GGTGGT	18.
FYF1 347	NGG 部 位 1-18 および 19	20	GccCACGGGCA GCTTGCCGG	19.	GccCACGGGCAGCTTGCC GGTGGT	20.
BPK1 345	NGG 部位 2	20	GTCGCCCTCGA ACTTCACCT	21.	GTCGCCCTCGAACTTCAC CTCGGC	22.
BPK1 350	NGG 部位 3	20	GTAGGTCAGGG TGGTCACGA	23.	GTAGGTCAGGGTGGTCA CGAGGGT	24.
BPK1 353	NGG 部位 4	20	GGCGAGGGCG ATGCCACCTA	25.	GGCGAGGGCGATGCCAC CTACGGC	26.
MSP7 92	NGG 部位 5	20	GGTCGCCACCA TGGTGAGCA	27.	GGTCGCCACCATGGTGA GCAAGGG	28.
MSP7 95	NGG 部位 6	20	GGTCAGGGTG GTCACGAGGG	29.	GGTCAGGGTGGTCACGA GGGTGGG	30.
FYF1 328	NGG 部位 7	20	GGTGGTGCAGA TGAACCTCA	31.	GGTGGTGCAGATGAACT TCAGGGT	32.
JAF1 001	NGG 部位 7	17	GGTGCAGATGA ACTTCA	33.	GGTGCAGATGAACTTCA GGGT	34.
BPK1 365	NGG 部位 8	20	GTTGGGGTCTT TGCTCAGGG	35.	GTTGGGGTCTTTGCTCAG GGCGGA	36.
MSP7 94	NGG 部位 9	20	GGTGGTCACGA GGGTGGGCC	37.	GGTGGTCACGAGGGTGG GCCAGGG	38.
FYF1 327	NGG 部位 10	20	GATGCCGTTCT TCTGCTTGT	39.	GATGCCGTTCTTCTGCTT GTCGGC	40.

10

20

30

【 0 1 0 8 】

【表 6】

JAF9 97	NGG 部位 10	17	GCCGTTCTTCT GCTTGT	41.	GCCGTTCTTCTGCTTGTC GGC	42.
BPK1 347	NGG 部位 11	20	GTCGCCACCAT GGTGAGCAA	43.	GTCGCCACCATGGTGAG CAAGGGC	44.
BPK1 369	NGG 部位 12	20	GCACTGCACGC CGTAGGTCA	45.	GCACTGCACGCCGTAGG TCAGGGT	46.
MSP2 545	NGG 部位 13	20	GTGAACCGCAT CGAGCTGAA	47.	GTGAACCGCATCGAGCT GAAGGGC	48.
MSP2 546	NGG 部位 14	20	GAAGGGCATCG ACTTCAAGG	49.	GAAGGGCATCGACTTCA AGGAGGA	50.
MSP2 547	NGG 部位 15	20	GCTTCATGTGG TCGGGGTAG	51.	GCTTCATGTGGTCGGGGT AGCGGC	52.
MSP2 548	NGG 部位 16	20	GCTGAAGCACT GCACGCCGT	53.	GCTGAAGCACTGCACGC CGTAGGT	54.
MSP2 549	NGG 部位 17	20	GCCGTCGTCCT TGAAGAAGA	55.	GCCGTCGTCCTTGAAGA AGATGGT	56.
MSP2 550	NGG 部位 18	20	GACCAGGATGG GCACCACCC	57.	GACCAGGATGGGCACCA CCCCGGT	58.
MSP2 551	NGG 部位 19	20	GACGTAGCCTT CGGGCATGG	59.	GACGTAGCCTTCGGGCAT GGCGGA	60.
MSP2 553	NGG 部位 20	20	GAAGTTCGAG GGCGACACCC	61.	GAAGTTCGAGGGCGACA CCCTGGT	62.
MSP2 554	NGG 部位 21	20	GAGCTGGACG GCGACGTAAA	63.	GAGCTGGACGGCGACGT AAACGGC	64.
MSP2 555	NGG 部位 22	20	GGCATCGCCCT CGCCCTCGC	65.	GGCATCGCCCTCGCCCTC GCCGGA	66.
MSP2 556	NGG 部位 23	20	GGCCACAAGTT CAGCGTGTC	67.	GGCCACAAGTTCAGCGT GTCCGGC	68.
FYF1 331	NGG 部位 24	20	GGGCGAGGAG CTGTTCACCG	69.	GGGCGAGGAGCTGTTCA CCGGGGT	70.
FYF1 560	NGG 部位 24	18	GCGAGGAGCT GTTACCG	71.	GCGAGGAGCTGTTACCC GGGGT	72.
BPK1 348	NGG 部位 25-非 5' G	20	CCTCGAACTTC ACCTCGGCG	73.	CCTCGAACTTCACCTCGG CGCGGG	74.
BPK1 349	NGG 部 位 25-mm 5' G	20	GCTCGAACTTC ACCTCGGCG	75.	GCTCGAACTTCACCTCGG CGCGGG	76.
BPK1 351	NGG 部位 26-非 5' G	20	CAACTACAAGA CCCGCGCCG	77.	CAACTACAAGACCCGCG CCGAGGT	78.
BPK1 352	NGG 部 位 26-mm 5' G	20	GAACTACAAGA CCCGCGCCG	79.	GAACTACAAGACCCGCG CCGAGGT	80.
BPK1 373	NGG 部位 27-非 5' G	20	CGCTCCTGGAC GTAGCCTTC	81.	CGCTCCTGGACGTAGCCT TCGGGC	82.
BPK1 375	NGG 部 位 27-mm 5' G	20	GGCTCCTGGAC GTAGCCTTC	83.	CGCTCCTGGACGTAGCCT TCGGGC	84.

10

20

30

【 0 1 0 9 】

【表 7】

BPK1 377	NGG 部位 28-非 5' G	20	AGGGCGAGGA GCTGTTTCACC	85.	AGGGCGAGGAGCTGTTC ACCGGGG	86.
BPK1 361	NGG 部 位 28-mm 5' G	20	GGGGCGAGGA GCTGTTTCACC	87.	GGGGCGAGGAGCTGTTC ACCGGGG	88.
BPK1 468	NGAA 部位 1	20	GTTCGAGGGCG ACACCCTGG	89.	GTTCGAGGGCGACACCC TGGTGAA	90.
MSP8 07	NGAA 部位 2	20	GTTCACCAGGG TGTCGCCCT	91.	GTTCACCAGGGTGTTCGC CCTCGAA	92.
MSP1 70	NGAC 部位 1	20	GCCCCACCCTCG TGACCACCC	93.	GCCCCACCCTCGTGACCA CCCTGAC	94.
MSP7 90	NGAC 部位 2	20	GCCCTTGCTCA CCATGGTGG	95.	GCCCTTGCTCACCATGGT GGCGAC	96.
MSP1 71	NGAT 部位 1	20	GTCGCCGTCCA GCTCGACCA	97.	GTCGCCGTCCAGCTCGA CCAGGAT	98.
MSP1 69	NGAT 部位 2	20	GTGTCCGGCGA GGGCGAGGG	99.	GTGTCCGGCGAGGGCGA GGGCGAT	100.
MSP1 68	NGAG 部位 1	20	GGGGTGGTGCC CATCCTGGT	101.	GGGGTGGTGCCCATCCTG GTCGAG	102.
MSP3 66	NGAG 部位 2	20	GCCACCATGGT GAGCAAGGG	103.	GCCACCATGGTGAGCAA GGGCGAG	104.

内在性遺伝子

EMXI

予備名称	名称	スパーサ ー 長 さ (nt)	スパーサー配列	配列 番号	拡張 PAM を有する配列	配列番号
FYF1548	NGG 部 位 1	20	GAGTCCGAG CAGAAGAAG AA	105.	GAGTCCGAGCAGAAGAAGAA GGGC	106.
MSP809	NGG 部 位 2	20	GTCACCTCCA ATGACTAGG G	107.	GTCACCTCCAATGACTAGGGT GGG	108.
VC475	NGG 部 位 3	20	GGGAAGACT GAGGCTACAT A	109.	GGGAAGACTGAGGCTACATAG GGT	110.
MSP814 *1	NGA 部 位 1	20	GCCACGAAG CAGGCCAAT GG	111.	GCCACGAAGCAGGCCAATGGG GAG	112.

FANCF

予備名称	名称	スパーサ ー 長 さ (nt)	スパーサー配列	配列 番号	拡張 PAM を有する配列	配列番号
DR348	NGG 部 位 1	20	GGAATCCCTT CTGCAGCACC	113.	GGAATCCCTTCTGCAGCACCT GGA	114.

【 0 1 1 0 】

【表 8】

MSP815	NGG 部位 2	20	GCTGCAGAA GGGATTCCAT G	115.	GCTGCAGAAAGGGATTCCATGA GGT	116.
MSP816	NGG 部位 3	20	GGCGGCTGC ACAACCAGT GG	117.	GGCGGCTGCACAACCAGTGGA GGC	118.
MSP817	NGG 部位 4	20	GCTCCAGAG CCGTGCGAAT G	119.	GCTCCAGAGCCGTGCGAATGG GGC	120.
MSP818 *2	NGA 部位 1	20	GAATCCCTTC TGCAGCACC T	121.	GAATCCCTTCTGCAGCACCTG GAT	122.
MSP820 *3	NGA 部位 2	20	GCGGCGGCT GCACAACCA GT	123.	GCGGCGGCTGCACAACCAGTG GAG	124.
MSP885 *4	NGA 部位 3	20	GGTTGTGCA GCCGCCGCT CC	125.	GGTTGTGCAGCCGCCGCTCCA GAG	126.

RUNXI

予備名称	名称	スパー ー 長さ (nt)	スパーサー配列	配列番 号	拡張 PAM を有する配列	配列番号
MSP822	NGG 部位 1	20	GCATTTTCAG GAGGAAGCG A	127.	GCATTTTCAGGAGGAAGCGAT GGC	128.
MSP825	NGG 部位 2	20	GGGAGAAGA AAGAGAGAT GT	129.	GGGAGAAGAAAGAGAGATGTA GGG	130.
MSP826 *5	NGA 部位 1	20	GGTGCATTTT CAGGAGGAA G	131.	GGTGCATTTTCAGGAGGAAGC GAT	132.
MSP828 *6	NGA 部位 2	20	GAGATGTAG GGCTAGAGG GG	133.	GAGATGTAGGGCTAGAGGGGT GAG	134.
MSP1725	NGAA 部位 1	20	GGTATCCAGC AGAGGGGAG A	135.	GGTATCCAGCAGAGGGGAGAA GAA	136.
MSP1726	NGAA 部位 2	20	GAGGCATCTC TGCACCGAG G	137.	GAGGCATCTCTGCACCGAGGT GAA	138.
MSP1728	NGAC 部位 1	20	GAGGGGTGA GGCTGAAAC AG	139.	GAGGGGTGAGGCTGAAACAGT GAC	140.

【 0 1 1 1 】

【表 9】

MSP1730	NGAC 部位 2	20	GAGCAAAAG TAGATATTAC A	141.	GAGCAAAAGTAGATATTACAA GAC	142.
MSP1732	NGAT 部位 1	20	GGAATTCAA ACTGAGGCAT A	143.	GGAATTCAAAGTCTGAGGCATATG AT	144.
MSP829	NGAT 部位 2	20	GCAGAGGGG AGAAGAAAG AG	145.	GCAGAGGGGAGAAGAAAGAG AGAT	146.
MSP1734	NGAG 部位 1	20	GCACCGAGG CATCTCTGCA C	147.	GCACCGAGGCATCTCTGCACC GAG	148.
MSP828	NGAG 部位 2	20	GAGATGTAG GGCTAGAGG GG	149.	GAGATGTAGGGCTAGAGGGGT GAG	150.

ZSCAN2

予備名称	名称	スパーサ ー 長さ (nt)	スパーサー配列	配列番 号	拡張 PAM を有する配列	配列番号
NN675	NGG 部 位	20	GTGCGGCAA GAGCTTCAG CC	151.	GTGCGGCAAGAGCTTCAGCCG GGG	152.

VEGFA

予備名称	名称	スパーサ ー 長さ (nt)	スパーサー配列	配列番 号	拡張 PAM を有する配列	配列番号
VC297	NGG 部 位 1	20	GGGTGGGGG GAGTTTGCTC C	153.	GGGTGGGGGGAGTTTGCTCCT GGA	154.
VC299	NGG 部 位 2	20	GACCCCTC CACCCCGCC TC	155.	GACCCCTCCACCCCGCCTCC GGG	156.
VC228	NGG 部 位 3	20	GGTGAGTGA GTGTGTGCGT G	157.	GGTGAGTGAGTGTGTGCGTGT GGG	158.
BPK1846 *7	NGA 部 位 1	20	GCGAGCAGC GTCTTCGAG AG	159.	GCGAGCAGCGTCTTCGAGAGT GAG	160.

【 0 1 1 2 】

【表 1 0】

ZNF629

予備名称	名称	スパーサ ー 長さ (nt)	スパーサー配列	配列番 号	拡張 PAM を有する配列	配列番号
NN675 *8	NGA 部 位	20	GTGCGGCAA GAGCTTCAG CC	161.	GTGCGGCAAGAGCTTCAGCCA GAG	162.

*1, Kleinstiver et al., Nature 2015 からの NGA EMX1 部位 4

*2, Kleinstiver et al., Nature 2015 からの NGA FANCF 部位 1

*3, Kleinstiver et al., Nature 2015 からの NGA FANCF 部位 3

*4, Kleinstiver et al., Nature 2015 からの NGA FANCF 部位 4

*5, Kleinstiver et al., Nature 2015 からの NGA RUNX1 部位 1

*6, Kleinstiver et al., Nature 2015 からの NGA RUNX1 部位 3

*7, Kleinstiver et al., Nature 2015 からの NGA VEGFA 部位 1

*8, Kleinstiver et al., Nature 2015 からの NGA ZNF629 部位

10

20

30

40

50

【0113】

S p C a s 9 - H F 1 のゲノムワイド特異性

S p C a s 9 - H F 1 がヒト細胞中で低減したオフターゲット効果を示したか否かを試験するため、シーケンシング (G U I D E - s e q) 方法により可能となる二本鎖分解のゲノムワイドの偏りのない同定を使用した。G U I D E - s e q は、隣接ゲノム配列の増幅およびシーケンシングを可能とするための二本鎖分解への短鎖二本鎖オリゴデオキシヌクレオチド (d s O D N) タグのインテグレーションを使用し、任意の所与の部位におけるタグインテグレーションの数は、開裂効率の定量的尺度を提供する (T s a i , S . Q . e t a l , N a t B i o t e c h n o l 33 , 187 - 197 (2015)) 。 G U I D E - s e q を使用して内在性ヒト E M X 1 、 F A N C F 、 R U N X 1 、 および Z S C A N 2 遺伝子中の種々の部位に標的化される 8 つの異なる s g R N A を使用して野生型 S p C a s 9 および S p C a s 9 - H F 1 により誘導されるオフターゲット効果の範囲を比較した。これらの s g R N A により標的化される配列はユニークであり、参照ヒトゲノム中の長さが様々な予測ミスマッチ部位を有する (表 2) 。 8 つの s g R N A についてのオンターゲット d s O D N タグインテグレーションの評価 (制限断片長多型 (R F L P) アッセイによる) およびインデル形成 (T 7 E I アッセイによる) により、野生型 S p C a s 9 および S p C a s 9 - H F 1 を用いた同等のオンターゲット活性が明らかになった (それぞれ図 7 a および 7 b) 。 G U I D E - s e q 実験は、 8 つの s g R N A の 7 つが、野生型 S p C a s 9 を用いて複数のゲノムワイドオフターゲット部位 (s g R N A 当たり 2 ~ 25 の範囲) における開裂を誘導する一方、 8 番目の s g R N A (F A N C F 部位 4 について) は、いかなる検出可能なオフターゲット部位も産生しないことを示した (図 2 a および 2 b) 。 しかしながら、野生型 S p C a s 9 を用いてインデルを誘導した 7 つの s g R N A の 6 つは、 S p C a s 9 - H F 1 を用いて G U I D E - s e q 検出可能オフターゲットイベントの顕著に完全な不存在を示し (図 2 a および 2 b) ; 残りの 7 番目の s g R N A (F A N C F 部位 2 について) は、プロトスペーサーシード配列内の 1 つのミスマッチを保有する部位における単一の検出可能なゲノムワイドオフターゲット開裂イベントのみを誘導した (図 2 a) 。 合計で、 S p C a s 9 - H F 1 を使用した場合に検出されなかったオフターゲット部位は、プロトスペーサーおよび / または P A M 配列中の 1 ~ 6 つのミスマッチを保有した (図 2 c) 。 野生型 S p C a s 9 に関して、 8 つの s g R N A (F A N C F 部位 4 について) は、 S p C a s 9 - H F 1 を用いて試験した場合にいかなる検出可能なオフターゲット開裂イベントも生じさせなかった (図 2 a) 。

【0114】

G U I D E - s e q の知見を確認するため、標的化アンプリコンシーケンシングを使用して野生型 S p C a s 9 および S p C a s 9 - H F 1 により誘導される N H E J 媒介インデル突然変異の頻度をより直接的に計測した。これらの実験のため、ヒト細胞に s g R N A - および C a s 9 コードプラスミドのみ (すなわち、 G U I D E - s e q タグを用いない) を形質移入した。次いで、次世代シーケンシングを使用して G U I D E - s e q 実験において 6 つの s g R N A について野生型 S p C a s 9 を用いて同定された 40 個のオフターゲット部位の 36 個を試験した (40 個の部位の 4 つは、それらをゲノム DNA から特異的に増幅させることができなかったため、試験することができなかった) 。 これらのディープシーケンシング実験は、 (1) 野生型 S p C a s 9 および S p C a s 9 - H F 1 が 6 つの s g R N A オンターゲット部位のそれぞれにおける同等の頻度のインデルを誘導すること (図 3 a および 3 b) ; (2) 野生型 S p C a s 9 が、予測されるとおり、同一部位についての G U I D E - s e q リードカウントと十分に相関する頻度における 36 個のオフターゲット部位 (図 3 b) の 35 個におけるインデル突然変異の統計的に有意な証拠を示すこと (図 3 c) ; および (3) 36 個のオフターゲット部位の 34 個における S p C a s 9 - H F 1 により誘導されるインデルの頻度は、対照形質移入からの試料中で観察されるインデルのバックグラウンドレベルと区別不能であることを示した (図 3 b) 。 陰性対照に対して S p C a s 9 - H F 1 を用いて統計的に有意な突然変異頻度を有することが考えられる 2 つのオフターゲット部位について、インデルの平均頻度は 0 . 049 %

および 0.037% であり、そのレベルにおいてそれらがシーケンシング / PCR エラーに起因するか、正規のヌクレアーゼ誘導インデルであるかを決定することは困難である。これらの結果に基づき、SpCas9-HF1 は、野生型 SpCas9 を用いて異なる頻度の範囲にわたり生じるオフターゲット突然変異を、検出不能なレベルに完全にまたはほぼ完全に低減させ得ることが結論付けられた。

【0115】

次に、非典型のホモポリマーまたは反復配列を標的化する sgRNA のゲノムワイドオフターゲット効果を低減させる SpCas9-HF1 の能力を評価した。目下のところ、多くがゲノムに対する直交性の相対的欠落に起因するの特徴を有するオンターゲット部位を回避しようと試行するが、SpCas9-HF1 がそれらの困難な標的についてもオフターゲットインデルを低減させ得るか否かを調査することが望ましいことがあった。したがって、ヒト VEGFA 遺伝子中のシトシンリッチホモポリマー配列または複数の TGRIP1 を含有する配列（それぞれ VEGFA 部位 2 および VEGFA 部位 3）のいずれかを標的化する、既に特徴付けされた sgRNA (Fu, Y. et al., Nat Biotechnol 31, Tsai, S. Q. et al., Nat Biotechnol 33, 187-197 (2015)) を使用した（表 2）。対照実験において、これらの sgRNA のそれぞれは、野生型 SpCas9 および SpCas9-HF1 の両方を用いて同等のレベルの GUIDE-seq sODN タグ取り込み（図 7c）およびインデル突然変異（図 7d）を誘導し、SpCas9-HF1 はそれらの sgRNA のいずれかを用いてオンターゲット活性が害されないことを実証した。重要なことに、GUIDE-seq 実験により、SpCas9-HF1 が、それらの sgRNA のオフターゲット部位の低減において高度に有効であることが明らかになり、VEGFA 部位 2 についての 123 個 / 144 個の部位および VEGFA 部位 3 についての 31 個 / 32 個の部位が検出されなかった（図 4a および 4b）。SpCas9-HF1 を用いて検出されなかったこれらのオフターゲット部位の試験は、それらがそれぞれ、それらのプロトスペーサーおよび PAM 配列内の一連の総ミスマッチ：VEGFA 部位 2 sgRNA についての 2 ~ 7 つのミスマッチおよび VEGFA 部位 3 sgRNA についての 1 ~ 4 つのミスマッチを有することを示し（図 4c）；さらに、VEGFA 部位 2 についてのそれらのオフターゲットの 9 つは、sgRNA-DNA 界面における潜在的なバルジ塩基 (Lin, Y. et al., Nucleic Acids Res 42, 7473-7485 (2014)) を有し得る（図 4a および 図 8）。SpCas9-HF1 を用いて検出されなかった部位は、VEGFA 部位 2 sgRNA についての 2 ~ 6 つのミスマッチおよび VEGFA 部位 3 sgRNA についての単一位点中の 2 つのミスマッチを有し（図 4c）、VEGFA 部位 2 sgRNA についての 3 つのオフターゲット部位はここでも潜在的バルジを有した（図 8）。まとめると、これらの結果は、SpCas9-HF1 が、単純リピート配列に標的化される sgRNA のオフターゲット効果の低減において高度に有効であり得、ホモポリマー配列に標的化される sgRNA へのかなりの影響も有し得ることを実証した。

【0116】

【表 1 1】

表 2| GUIDE-seq により試験された 10 個の sgRNA についての参照ヒトゲノム中の潜在的ミスマッチ部位のまとめ

部位	PAM を有するスパーサー	オンターゲット部位に対するミスマッチ*					
		1	2	3	4	5	6
EMX1-1	GAGTCCGAGCAGAAGAAGAAGGG (配列番号 163)	0	1	18	273	2318	15831
EMX1-2	GTCACCTCCAATGACTAGGGTGG (配列番号 164)	0	0	3	68	780	6102
FANCF-1	GGAATCCCTTCTGCAGCACCTGG (配列番号 165)	0	1	18	288	1475	9611
FANCF-2	GCTGCAGAAGGGATTCCATGAGG (配列番号 166)	1	1	29	235	2000	13047
FANCF-3	GGCGGCTGCACAACCACTGGAGG (配列番号 167)	0	0	11	79	874	6651
FANCF-4	GCTCCAGAGCCGTGCGAATGGGG (配列番号 168)	0	0	6	59	639	5078
RUNX1-1	GCATTTTCAGGAGGAAGCGATGG (配列番号 169)	0	2	6	189	1644	11546
ZSCAN2	GTGCGGCAAGAGCTTCAGCCGGG (配列番号 170)	0	3	12	127	1146	10687
VEGFA2	GACCCCTCCACCCCGCTCCGG (配列番号 171)	0	2	35	456	3905	17576
VEGFA3	GGTGAGTGAGTGTGTGCGTGTGG (配列番号 172)	1	17	383	6089	13536	35901

* Cas-OFFinder (Bae et al., Bioinformatics 30, 1473-1475 (2014))を使用して決定

【 0 1 1 7 】

【表 1 2】

表 4:試験において使用されたオリゴヌクレオチド

<u>T7E1 プライマーの説明</u>	<u>配列</u>	<u>配列番号</u>	
U2OS ヒト細胞中の EMX1 を増幅するためのフォワードプライマー	GGAGCAGCTGGTCAG AGGGG	173.	
U2OS ヒト細胞中の EMX1 を増幅するためのリバースプライマー	CCATAGGGAAGGGGG ACACTGG	174.	
U2OS ヒト細胞中の FANCF を増幅するためのフォワードプライマー	GGGCCGGGAAAGAGT TGCTG	175.	10
U2OS ヒト細胞中の FANCF を増幅するためのリバースプライマー	GCCCTACATCTGCTCT CCCTCC	176.	
U2OS ヒト細胞中の RUNX1 を増幅するためのフォワードプライマー	CCAGCACAACCTTACTC GCACTTGAC	177.	
U2OS ヒト細胞中の RUNX1 を増幅するためのリバースプライマー	CATCACCAACCCACAG CCAAGG	178.	
U2OS ヒト細胞中の VEGFA を増幅するためのフォワードプライマー	TCCAGATGGCACATTG TCAG	179.	
U2OS ヒト細胞中の VEGFA を増幅するためのリバースプライマー	AGGGAGCAGGAAAGT GAGGT	180.	20
U2OS ヒト細胞中の VEGFA (NGG 部位 2)を増幅するためのフォワードプライマー	CGAGGAAGAGAGAGA CGGGGTC	181.	
U2OS ヒト細胞中の VEGFA (NGG 部位 2)を増幅するためのリバースプライマー	CTCCAATGCACCCAAG ACAGCAG	182.	
U2OS ヒト細胞中の ZSCAN2 を増幅するためのフォワードプライマー	AGTGTGGGGTGTGTGG GAAG	183.	
U2OS ヒト細胞中の ZSCAN2 を増幅するためのリバースプライマー	GCAAGGGGAAGACTC TGGCA	184.	

【表 1 3】

U2OS ヒト細胞中の ZNF629 を増幅するためのフォワードプライマー	TACGAGTGCCTAGAGT GCG	185.	
U2OS ヒト細胞中の ZNF629 を増幅するためのリバープライマー	GCAGATGTAGGTCTTG GAGGAC	186.	
ディープシーケンシングプライマーの説明	配列	配列番号	
EMX1-1 オンターゲットを増幅するためのフォワードプライマー	GGAGCAGCTGGTCAG AGGGG	187.	
EMX1-1 オンターゲットを増幅するためのリバープライマー	CGATGTCCTCCCCATT GGCCTG	188.	10
EMX1-1-GUIDE_seq-OT#1 を増幅するためのフォワードプライマー	GTGGGGAGATTTGCAT CTGTGGAGG	189.	
EMX1-1-GUIDE_seq-OT#1 を増幅するためのリバープライマー	GCTTTTATACCATCTTG GGGTTACAG	190.	
EMX1-1-GUIDE_seq-OT#2 を増幅するためのフォワードプライマー	CAATGTGCTTCAACCC ATCACGGC	191.	
EMX1-1-GUIDE_seq-OT#2 を増幅するためのリバープライマー	CCATGAATTTGTGATG GATGCAGTCTG	192.	
EMX1-1-GUIDE_seq-OT#3 を増幅するためのフォワードプライマー	GAGAAGGAGGTGCAG GAGCTAGAC	193.	20
EMX1-1-GUIDE_seq-OT#3 を増幅するためのリバープライマー	CATCCCGACCTTCATC CCTCCTGG	194.	
EMX1-1-GUIDE_seq-OT#4 を増幅するためのフォワードプライマー	GTAGTTCTGACATTCC TCCTGAGGG	195.	
EMX1-1-GUIDE_seq-OT#4 を増幅するためのリバープライマー	TCAAACAAGGTGCAG ATACAGCA	196.	
EMX1-1-GUIDE_seq-OT#5 を増幅するためのフォワードプライマー	CAGGGTCGCTCAGTCT GTGTGG	197.	
EMX1-1-GUIDE_seq-OT#5 を増幅するためのリバープライマー	CCAGCGCACCATTAC TCCACCTG	198.	30
EMX1-1-GUIDE_seq-OT#6 を増幅するためのフォワードプライマー	GGCTGAAGAGGAAGA CCAGACTCAG	199.	
EMX1-1-GUIDE_seq-OT#6 を増幅するためのリバープライマー	GGCCCCTCTGAATTCA ATTCTCTGC	200.	
EMX1-1-GUIDE_seq-OT#7 を増幅するためのフォワードプライマー	CCACAGCGAGGAGTG ACAGCC	201.	
EMX1-1-GUIDE_seq-OT#7 を増幅するためのリバープライマー	CCAAGTCTTTCCTAAC TCGACCTTG	202.	
EMX1-1-GUIDE_seq-OT#8 を増幅するためのフォワードプライマー	CCCTAGGCCCCACACCA GCAATG	203.	
EMX1-1-GUIDE_seq-OT#8 を増幅するためのリバープライマー	GGGATGGGAATGGGAA TGTGAGGC	204.	40
EMX1-2 オンターゲットを増幅するためのフォワードプライマー	GCCCAGGTGAAGGTG TGGTTCC	205.	

【表 1 4】

EMX1-2 オンターゲットを増幅するためのリバー スプライマー	CCAAAGCCTGGCCAG	206.	
EMX1-2-GUIDE_seq-OT#1 を増幅するための フォワードプライマー	GGAGTG		
EMX1-2-GUIDE_seq-OT#1 を増幅するための リバープライマー	AGGCAAAGATCTAGGA	207.	
EMX1-2-GUIDE_seq-OT#2 を増幅するための フォワードプライマー	CCTGGATGG		
EMX1-2-GUIDE_seq-OT#2 を増幅するための リバープライマー	CCATCTGAGTCAGCCA	208.	
EMX1-2-GUIDE_seq-OT#3 を増幅するための フォワードプライマー	GCCTTGTC		
EMX1-2-GUIDE_seq-OT#3 を増幅するための リバープライマー	GGTTCCCTCCCTTCTG	209.	
EMX1-2-GUIDE_seq-OT#4 を増幅するための フォワードプライマー	AGCCC		10
EMX1-2-GUIDE_seq-OT#4 を増幅するための リバープライマー	GGATAGGAATGAAGAC	210.	
EMX1-2-GUIDE_seq-OT#5 を増幅するための フォワードプライマー	CCCCTCTCC		
EMX1-2-GUIDE_seq-OT#5 を増幅するための リバープライマー	GGACTGGCTGGCTGTG	211.	
EMX1-2-GUIDE_seq-OT#6 を増幅するための フォワードプライマー	TGTTTTGAG		
EMX1-2-GUIDE_seq-OT#6 を増幅するための リバープライマー	CTTATCCAGGGCTACC	212.	
EMX1-2-GUIDE_seq-OT#7 を増幅するための フォワードプライマー	TCATTGCC		
EMX1-2-GUIDE_seq-OT#7 を増幅するための リバープライマー	GCTGCTGCTGCTTTGA	213.	
EMX1-2-GUIDE_seq-OT#8 を増幅するための フォワードプライマー	TCACTCCTG		
EMX1-2-GUIDE_seq-OT#8 を増幅するための リバープライマー	CTCCTTAAACCCTCAG	214.	
EMX1-2-GUIDE_seq-OT#9 を増幅するための フォワードプライマー	AAGCTGGC		
EMX1-2-GUIDE_seq-OT#9 を増幅するための リバープライマー	GCACTGTCAGCTGATC	215.	20
EMX1-2-GUIDE_seq-OT#10 を増幅するための フォワードプライマー	CTACAGG		
EMX1-2-GUIDE_seq-OT#10 を増幅するための リバープライマー	ACGTTGGAACAGTCG	216.	
EMX1-2-GUIDE_seq-OT#11 を増幅するための フォワードプライマー	AGCTGTAGC		
EMX1-2-GUIDE_seq-OT#11 を増幅するための リバープライマー	TGTGCATAACTCATGTT	217.	
EMX1-2-GUIDE_seq-OT#12 を増幅するための フォワードプライマー	GGCAAACCT		
EMX1-2-GUIDE_seq-OT#12 を増幅するための リバープライマー	TCCACAACCTACCCTCA	218.	
EMX1-2-GUIDE_seq-OT#13 を増幅するための フォワードプライマー	GCTGGAG		
EMX1-2-GUIDE_seq-OT#13 を増幅するための リバープライマー	CCACTGACAATTCCT	219.	
EMX1-2-GUIDE_seq-OT#14 を増幅するための フォワードプライマー	CAACCCTGC		
EMX1-2-GUIDE_seq-OT#14 を増幅するための リバープライマー	AGGCAGACCAGTTATT	220.	
EMX1-2-GUIDE_seq-OT#15 を増幅するための フォワードプライマー	TGGCAGTC		30
EMX1-2-GUIDE_seq-OT#15 を増幅するための リバープライマー	ACAGGCGCAGTTCACT	221.	
EMX1-2-GUIDE_seq-OT#16 を増幅するための フォワードプライマー	GAGAAG		
EMX1-2-GUIDE_seq-OT#16 を増幅するための リバープライマー	GGGTAGGCTGACTTTG	222.	
FANCF-1 オンターゲットを増幅させるためのフォ ワードプライマー	GGCTCC		
FANCF-1 オンターゲットを増幅させるためのリバ ースプライマー	GCCCTCTTGCCTCCAC	223.	
FANCF-1-GUIDE_seq-OT#1 を増幅するため のフォワードプライマー	TGTTTG		
FANCF-1-GUIDE_seq-OT#1 を増幅するため のリバープライマー	CGCGGATGTTCCAATC	224.	
FANCF-1-GUIDE_seq-OT#2 を増幅するため のフォワードプライマー	AGTACGC		
FANCF-1-GUIDE_seq-OT#2 を増幅するため のリバープライマー	GCGGGCAGTGGCGTCT	225.	
FANCF-1-GUIDE_seq-OT#3 を増幅するため のフォワードプライマー	TAGTCG		
FANCF-1-GUIDE_seq-OT#3 を増幅するため のリバープライマー	CCCTGGGTTTGTTGG	226.	40
FANCF-1-GUIDE_seq-OT#4 を増幅するため のフォワードプライマー	CTGCTC		
FANCF-1-GUIDE_seq-OT#4 を増幅するため のリバープライマー	CTCCTTGCCGCCAGC	227.	
	CGGTC		

【表 15】

FANCF-1-GUIDE_seq-OT#2 を増幅するためのリバースプライマー	CACTGGGGAAGAGGC	228.	
FANCF-1-GUIDE_seq-OT#3 を増幅するためのフォワードプライマー	GAGGACAC		
FANCF-1-GUIDE_seq-OT#3 を増幅するためのリバースプライマー	CCAGTGTTTCCCATCC	229.	
FANCF-1-GUIDE_seq-OT#4 を増幅するためのフォワードプライマー	CCAACAC		
FANCF-1-GUIDE_seq-OT#4 を増幅するためのリバースプライマー	GAATGGATCCCCCCT	230.	
FANCF-1-GUIDE_seq-OT#4 を増幅するためのフォワードプライマー	AGAGCTC		
FANCF-1-GUIDE_seq-OT#4 を増幅するためのリバースプライマー	CAGGCCACAGGTCCT	231.	
FANCF-3 オンターゲットを増幅するためのフォワードプライマー	TCTGGA		
FANCF-3 オンターゲットを増幅するためのリバースプライマー	CCACACGGAAGGCTG	232.	10
FANCF-3 オンターゲットを増幅するためのフォワードプライマー	ACCACG		
FANCF-3 オンターゲットを増幅するためのリバースプライマー	GCGCAGAGAGAGCAG	233.	
FANCF-3 オンターゲットを増幅するためのフォワードプライマー	GACGTC		
FANCF-3 オンターゲットを増幅するためのリバースプライマー	GCACCTCATGGAATCC	234.	
FANCF-3-GUIDE_seq-OT#1 を増幅するためのフォワードプライマー	CTTCTGC		
FANCF-3-GUIDE_seq-OT#1 を増幅するためのリバースプライマー	CAAGTGATGCGACTTC	235.	
FANCF-3-GUIDE_seq-OT#1 を増幅するためのフォワードプライマー	CAACCTC		
FANCF-3-GUIDE_seq-OT#1 を増幅するためのリバースプライマー	CCCTCAGAGTTCAGCT	236.	
FANCF-3-GUIDE_seq-OT#2 を増幅するためのフォワードプライマー	TAAAAAGACC		
FANCF-3-GUIDE_seq-OT#2 を増幅するためのリバースプライマー	TGCTTCTCATCCACTCT	237.	20
FANCF-3-GUIDE_seq-OT#2 を増幅するためのフォワードプライマー	AGACTGCT		
FANCF-3-GUIDE_seq-OT#2 を増幅するためのリバースプライマー	CACCAACCAGCCATGT	238.	
FANCF-3-GUIDE_seq-OT#3 を増幅するためのフォワードプライマー	GCCATG		
FANCF-3-GUIDE_seq-OT#3 を増幅するためのリバースプライマー	CTGCCTGTGCTCCTCG	239.	
FANCF-3-GUIDE_seq-OT#3 を増幅するためのフォワードプライマー	ATGGTG		
FANCF-3-GUIDE_seq-OT#3 を増幅するためのリバースプライマー	GGGTTCAAAGCTCATC	240.	
FANCF-3-GUIDE_seq-OT#4 を増幅するためのフォワードプライマー	TGCCCC		
FANCF-3-GUIDE_seq-OT#4 を増幅するためのリバースプライマー	GCATGTGCCTTGAGAT	241.	
FANCF-3-GUIDE_seq-OT#4 を増幅するためのフォワードプライマー	TGCCTGG		
FANCF-3-GUIDE_seq-OT#4 を増幅するためのリバースプライマー	GACATTCAGAGAAGC	242.	
FANCF-3-GUIDE_seq-OT#5 を増幅するためのフォワードプライマー	GACCATGTGG		
FANCF-3-GUIDE_seq-OT#5 を増幅するためのリバースプライマー	CCATCTTCCCCTTTGG	243.	30
FANCF-3-GUIDE_seq-OT#5 を増幅するためのフォワードプライマー	CCCACAG		
FANCF-3-GUIDE_seq-OT#5 を増幅するためのリバースプライマー	CCCCAAAAGTGGCCA	244.	
FANCF-3-GUIDE_seq-OT#6 を増幅するためのフォワードプライマー	AGAGCCTGAG		
FANCF-3-GUIDE_seq-OT#6 を増幅するためのリバースプライマー	GTTCTCCAAAGGAAG	245.	
FANCF-3-GUIDE_seq-OT#6 を増幅するためのフォワードプライマー	AGAGGGGAATG		
FANCF-3-GUIDE_seq-OT#6 を増幅するためのリバースプライマー	GGTGCTGTGTCCTCAT	246.	
FANCF-3-GUIDE_seq-OT#7 を増幅するためのフォワードプライマー	GCATCC		
FANCF-3-GUIDE_seq-OT#7 を増幅するためのリバースプライマー	CGGCTTGCCTAGGGTC	247.	
FANCF-3-GUIDE_seq-OT#7 を増幅するためのフォワードプライマー	GTTGAG		
FANCF-3-GUIDE_seq-OT#7 を増幅するためのリバースプライマー	CCTTCAGGGGCTCTTC	248.	40
RUNX1-1 オンターゲットを増幅するためのフォワードプライマー	CAGGTC		
	GGGAACCTGGCAGGCA	249.	
	CCGAGG		

【表 16】

RUNX1-1 オンターゲットを増幅するためのリバースプライマー	GGGTGAGGCTGAAAC	250.	
RUNX1-1-GUIDE_seq-OT#1 を増幅するためのフォワードプライマー	AGTGACC		
RUNX1-1-GUIDE_seq-OT#1 を増幅するためのリバースプライマー	GGGAGGATGTTGGTTT	251.	
RUNX1-1-GUIDE_seq-OT#2 を増幅するためのフォワードプライマー	TAGGGAACCTG		
RUNX1-1-GUIDE_seq-OT#2 を増幅するためのリバースプライマー	TCCAATCACTACATGC	252.	
ZSCAN2 オンターゲットを増幅するためのフォワードプライマー	CATTTTGAAGA		
ZSCAN2 オンターゲットを増幅するためのリバースプライマー	CCACCCTCTTCCTTTG	253.	
ZSCAN2-GUIDE_seq-OT#1 を増幅するためのフォワードプライマー	ATCCTCCC		
ZSCAN2-GUIDE_seq-OT#1 を増幅するためのリバースプライマー	TCCTCCCTACTCCTTCA	254.	10
ZSCAN2-GUIDE_seq-OT#2 を増幅するためのフォワードプライマー	CCCAGG		
ZSCAN2-GUIDE_seq-OT#2 を増幅するためのリバースプライマー	GAGTGCCTGACATGTG	255.	
ZSCAN2-GUIDE_seq-OT#3 を増幅するためのフォワードプライマー	GGGAGAG		
ZSCAN2-GUIDE_seq-OT#3 を増幅するためのリバースプライマー	TCCAGCTAAAGCCTTT	256.	
ZSCAN2-GUIDE_seq-OT#4 を増幅するためのフォワードプライマー	CCCACAC		
ZSCAN2-GUIDE_seq-OT#4 を増幅するためのリバースプライマー	GAAGTCTCTGATGCAC	257.	
ZSCAN2-GUIDE_seq-OT#5 を増幅するためのフォワードプライマー	CTGAAGGCTG		
ZSCAN2-GUIDE_seq-OT#5 を増幅するためのリバースプライマー	ACCGTATCAGTGTGAT	258.	
ZSCAN2-GUIDE_seq-OT#6 を増幅するためのフォワードプライマー	GCATGTGGT		
ZSCAN2-GUIDE_seq-OT#6 を増幅するためのリバースプライマー	TGGGTTTAATCATGTGT	259.	20
ZSCAN2-GUIDE_seq-OT#7 を増幅するためのフォワードプライマー	TCTGCACTATG		
ZSCAN2-GUIDE_seq-OT#7 を増幅するためのリバースプライマー	CCCATCTTCCATTCTGC	260.	
ZSCAN2-GUIDE_seq-OT#8 を増幅するためのフォワードプライマー	CCTCCAC		
ZSCAN2-GUIDE_seq-OT#8 を増幅するためのリバースプライマー	CAGCTAGTCCATTTGT	261.	
ZSCAN2-GUIDE_seq-OT#9 を増幅するためのフォワードプライマー	TCTCAGACTGTG		
ZSCAN2-GUIDE_seq-OT#9 を増幅するためのリバースプライマー	GGCCAACATTGTGAAA	262.	
ZSCAN2-GUIDE_seq-OT#10 を増幅するためのフォワードプライマー	CCCTGTCTC		
ZSCAN2-GUIDE_seq-OT#10 を増幅するためのリバースプライマー	CCAGGGACCTGTGCTT	263.	
ZSCAN2-GUIDE_seq-OT#11 を増幅するためのフォワードプライマー	GGGTTC		
ZSCAN2-GUIDE_seq-OT#11 を増幅するためのリバースプライマー	CACCCCATGACCTGGC	264.	
ZSCAN2-GUIDE_seq-OT#12 を増幅するためのフォワードプライマー	ACAAGTG		
ZSCAN2-GUIDE_seq-OT#12 を増幅するためのリバースプライマー	AAGTGTTCTCAGAAT	265.	30
ZSCAN2-GUIDE_seq-OT#13 を増幅するためのフォワードプライマー	GCCAGCCC		
ZSCAN2-GUIDE_seq-OT#13 を増幅するためのリバースプライマー	CAGGAGTGCAGTTGTG	266.	
ZSCAN2-GUIDE_seq-OT#14 を増幅するためのフォワードプライマー	TTGGGAG		
ZSCAN2-GUIDE_seq-OT#14 を増幅するためのリバースプライマー	CTGATGAAGCACCAGA	267.	
ZSCAN2-GUIDE_seq-OT#15 を増幅するためのフォワードプライマー	GAACCCACC		
ZSCAN2-GUIDE_seq-OT#15 を増幅するためのリバースプライマー	CACACCTGGCACCCAT	268.	
ZSCAN2-GUIDE_seq-OT#16 を増幅するためのフォワードプライマー	ATGGC		
ZSCAN2-GUIDE_seq-OT#16 を増幅するためのリバースプライマー	GATCCACACTGGTGAG	269.	
ZSCAN2-GUIDE_seq-OT#17 を増幅するためのフォワードプライマー	AAGCCTTAC		
ZSCAN2-GUIDE_seq-OT#17 を増幅するためのリバースプライマー	CTTCCCACACTCACAG	270.	40
	CAGATGTAGG		

【0122】

SpCas9-HF1の特異性の改良

既に記載された方法、例えば、トランケートgRNA(Fu, Y. et al., Nat Biotechnol 32, 279-284 (2014))およびSpCas9-D1135Eバリエーション(Kleinstiver, B. P. et al., Nature 523, 481-485 (2015))は、SpCas9オフターゲット効果を部分的に低減させ得、本発明者らは、それらをSpCas9-HF1と組み合わせてそのゲノムワイド特異性をさらに改善することができるか否かを考慮した。ヒト細胞ベースEGF

P崩壊アッセイにおける4つの部位に標的化されるマッチ全長およびトランケートsgRNAを用いたSpCas9-HF1の試験により、sgRNA相補性長さの短縮が、オンターゲット活性を実質的に害することが明らかになった(図9)。対照的に、追加のD1135E突然変異を有するSpCas9-HF1(本明細書においてSpCas9-HF2と称されるバリエーション)は、ヒト細胞ベースEGFP崩壊アッセイを使用して試験した8つのsgRNAの6つを用いて野生型SpCas9の70%以上の活性を保持した(図5aおよび5b)。側鎖が標的DNAとのそのPAM近位末端上での疎水性非特異的相互作用を媒介する位置におけるそれぞれL169AまたはY450A突然変異を保有するSpCas9-HF3およびSpCas9-HF4バリエーションも作出した(Nishimasu, H. et al., Cell 156, 935-949 (2014); Jiang, F., et al., Science 348, 1477-1481 (2015))。SpCas9-HF3およびSpCas9-HF4は、8つのEGFP標的化sgRNAのうちの同一の6つを用いて野生型SpCas9を用いて観察された活性の70%以上を保持した(図5aおよび5b)。

【0123】

SpCas9-HF2、-HF3、および-HF4が、SpCas9-HF1に耐性である2つのオフターゲット部位(FANCF部位2およびVEGFA部位3sgRNAについて)におけるインデル頻度を低減させ得るか否かを決定するため、さらなる実験を実施した。プロトスペーサーのシード配列中の単一ミスマッチを担持するFANCF部位2オフターゲットについて、SpCas9-HF4は、T7E1アッセイにより判断されたとおり、インデル突然変異頻度をほぼバックグラウンドレベルに低減させた一方、オンターゲット活性も有益に増加させ(図5c)、3つのバリエーション間の特異性の最大増加をもたらした(図5d)。2つのプロトスペーサーミスマッチ(シード配列中に1つ、およびPAM配列から最も遠位のヌクレオチドにおいて1つ)を担持するVEGFA部位3オフターゲット部位について、SpCas9-HF2はインデル形成の最大低減を示した一方、オンターゲット突然変異頻度に対してわずかな効果のみを示し(図5c)、試験した3つのバリエーション間の特異性の最大増加をもたらした(図5d)。まとめると、これらの結果は、非特異的DNA接触を媒介し、またはPAM認識を変更し得る他の残基における追加の突然変異を導入することにより、SpCas9-HF1に耐性であるオフターゲット効果の低減についての潜在性を実証する。

【0124】

FANCF部位2およびVEGFA部位3sgRNAのオフターゲットに対してそれぞれ、SpCas9-HF4およびSpCas9-HF2がSpCas9-HF1に対して改善された区別を有することを示す上記のT7E1アッセイの知見を一般化するため、GUIDE-seqを使用してそれらのバリエーションのゲノムワイド特異性を試験した。RFLPアッセイを使用して、SpCas9-HF4およびSpCas9-HF2は、GUIDE-seqタグインテグレーション率によりアッセイされたとおり、SpCas9-HF1と類似のオンターゲット活性を有することが決定された(図5E)。GUIDE-seqデータを分析した場合、SpCas9-HF2またはSpCas9-HF4について新たなオフターゲット部位は同定されなかった(図5F)。SpCas9-HF1と比較して、全ての部位におけるオフターゲット活性はGUIDE-seqにより検出不能になり、または実質的に減少した。SpCas9-HF1に対して、SpCas9-HF4は、SpCas9-HF1の特異性の改善に抵抗性のままである単一のFANCF部位2オフターゲット部位に対してほぼ26倍良好な特異性を有した(図5F)。SpCas9-HF2は、高頻度VEGFA部位3オフターゲットについてSpCas9-HF1に対してほぼ4倍改善された特異性を有した一方、他の低頻度オフターゲット部位におけるGUIDE-seq検出可能イベントも大幅に低減し(>38倍)、または排除した。注目すべきことに、SpCas9-HF1について同定されたそれらの低頻度部位の3つのゲノム位置は、既に特徴付けされたバックグラウンドU2OS細胞切断点ホットスポットに隣接する。まとめると、これらの結果はSpCas9-HF2およびSpCas9-HF4

バリエーションが SpCas9 - HF1 のゲノムワイド特異性を改善し得ることを示唆する。

【0125】

標準的な非反復標的配列に対して設計された sgRNA を使用した場合、SpCas9 - HF1 はオフターゲット突然変異をロバストに一貫して低減させた。SpCas9 - HF1 に最も耐性であった2つのオフターゲット部位は、プロトスペーサー中の1および2つのミスマッチのみを有する。まとめると、これらの観察は、SpCas9 - HF1 を使用し、ゲノム中のいずれかの場所で1または2つのミスマッチを担持する密接に関連する部位を有さない非反復配列を標的化することにより（既存の公的に利用可能なソフトウェアプログラム（Baer, S., et al., Bioinformatics 30, 1473 - 1475 (2014)）を使用して容易に達成することができるもの）、オフターゲット突然変異を検出不能なレベルに最小化させ得ることを示唆する。使用者が留意すべき1つのパラメータは、SpCas9 - HF1 が、プロトスペーサー配列にミスマッチする sgRNA の5' 末端における G を使用する一般的手法と適合性であり得ないことである。標的部位に対してミスマッチする5' G を担持する4つの sgRNA の試験は、4つのうちの3つが野生型 SpCas9 と比較して SpCas9 - HF1 を用いて縮小した活性を示し（図10）、部分的にマッチする部位をより良好に区別する SpCas9 - HF1 の能力を反映する可能性がある。

【0126】

さらなる生化学的試験は、SpCas9 - HF1 その高いゲノムワイド特異性を達成する正確な機序を確認または明確化し得る。導入される4つの突然変異が細胞中の SpCas9 の安定性または定常状態発現レベルを変更することは明らかでない。それというもの、減少濃度の発現プラスミドを用いるタイトレーション実験は、野生型 SpCas9 および SpCas9 - HF1 が、それらの濃度が低下するにつれて同等に挙動することを示唆した（図11）。その代わり、最も簡易な機序的説明は、それらの突然変異が、オンターゲット活性を保持するためにちょうど十分であるがオフターゲット部位開裂を非効率または不存在にするほど低いレベルにおける複合体のエネルギーで Cas9 - sgRNA と標的 DNA との間の相互作用のエネルギーを減少させたことである。この機序は、構造データ中の突然変異残基と標的 DNA リン酸骨格との間で観察された非特異的相互作用と一致する（Nishimasu, H. et al., Cell 156, 935 - 949 (2014); Anders, C. et al., Nature 513, 569 - 573 (2014)）。いくらか類似の機序は、正荷電残基における置換を担持する転写アクチベーター様エフェクターヌクレアーゼの増加した特異性を説明するために提案されている（Guillinger, J. P. et al., Nat Methods 11, 429 - 435 (2014)）。

【0127】

SpCas9 - HF1 は、Cas9 機能を変更することが示された他の突然変異と組み合わせることもできることが可能であった。例えば、3つのアミノ酸置換を担持する SpCas9 突然変異体（D1135V / R1335Q / T1337R、SpCas9 - VQR バリエーションとしても公知）は、NGAN PAM（NGAG > NGAT = NGAA > NGAC についての相対効率を有する）を有する部位を認識し（Kleinstiver, B. P. et al., Nature 523, 481 - 485 (2015)）、近年同定された四重 SpCas9 突然変異体（D1135V / G1218R / R1335Q / T1337R、SpCas9 - VRQR バリエーションと称される）は、NGAH（H = A、C、または T）PAM を有する部位に対して VQR バリエーションに対して改善された活性を有する（図12a）。SpCas9 - HF1 から SpCas9 - VQR および SpCas9 - VRQR 中への4つの突然変異（N497A / R661A / Q695A / Q926A）の導入は、それぞれ SpCas9 - VQR - HF1 および SpCas9 - VRQR - HF1 を作出した。これらのヌクレアーゼの HF バージョンの両方は、EGFP レポーター遺伝子に標的化される8つの sgRNA の5つを用いて、および内在性ヒト遺伝子部位に標的化される8つの sgRNA の7つを用いてそれらの非 HF 相当物と同等（すなわち、70

%以上)のオンターゲット活性を示した(図12b~12d)。

【0128】

より広く考慮すると、これらの結果は、CRISPR関連ヌクレアーゼの追加の高フィデリティバリエーションの遺伝子操作のための一般的方針を説明する。非特異的DNA接触残基における追加の突然変異を追加することは、SpCas9-HF1を用いて存続する極めて少数の残留オフターゲット部位の一部をさらに低減させた。したがって、バリエーション、例えば、SpCas9-HF2、SpCas9-HF3、SpCas9-HF4などは、オフターゲット配列の性質に応じて特注様式で利用することができる。さらに、SpCas9の高フィデリティバリエーションの遺伝子操作を用いた成功は、非特異的DNA接触を突然変異させるアプローチが他の天然存在および遺伝子操作Cas9オルソログ(Ran, F. A. et al., Nature 520, 186-191 (2015), Esvelt, K. M. et al., Nat Methods 10, 1116-1121 (2013); Hou, Z. et al., Proc Natl Acad Sci USA (2013); Fonfara, I. et al., Nucleic Acids Res 42, 2577-2590 (2014); Kleinstiver, B. P. et al., Nat Biotechnol (2015)ならびに増加した頻度で発見および特徴付けされているより新たなCRISPR関連ヌクレアーゼ(Zetsche, B. et al., Cell 163, 759-771 (2015); Shmakov, S. et al., Molecular Cell 60, 385-397)に拡張し得ることを示唆する。

【0129】

実施例2

標的鎖DNAに接触する残基中のアラニン置換、例として、N497A、Q695A、R661A、およびQ926Aを有するSpCas9バリエーションが本明細書に記載される。これらの残基の他、本発明者らは、それらのバリエーション、例えば、SpCas9-HF1バリエーション(N497A/R661A/Q695A/Q926A)の特異性を、非標的DNA鎖に接触すると考えられる正荷電SpCas9残基: R780、K810、R832、K848、K855、K968、R976、H982、K1003、K1014、K1047、および/またはR1060(Slaymaker et al., Science. 2016 Jan 1; 351(6268): 84-8参照)中の置換を追加することにより、さらに改善することができるか否かを決定することを追求した。

【0130】

これらの位置における単一アラニン置換およびそれらの組合せを担持する野生型SpCas9誘導体の活性を、EGFP遺伝子中の部位に対して設計された完全マッチsgRNA(オンターゲット活性を評価するため)ならびに11および12位(1位は、最もPAMから近位の塩基である)における意図的なミスマッチを担持する同一のsgRNA(オフターゲット部位において見出されるミスマッチ部位における活性を評価するため)を用いるEGFP崩壊アッセイを使用して最初に試験した(図13A)。(三重置換K810A/K1003A/R1060AまたはK848A/K1003A/R1060Aを担持する誘導体は、それぞれeSpCas9(1.0)およびeSpCas9(1.1)として公知の近年記載されたバリエーションと同一であることに留意されたい; 参照文献1参照)。予測されるとおり、野生型SpCas9は、ロバストなオンターゲットおよびミスマッチ標的活性を有した。対照として、この実験においてSpCas9-HF1も試験し、それが予測されるとおりオンターゲット活性を維持する一方、ミスマッチ標的活性を低減させることを見出した(図13A)。非標的DNA鎖に潜在的に接触し得る位置における1つ以上のアラニン置換を担持する野生型SpCas9誘導体の全ては、野生型SpCas9と同等のオンターゲット活性を示した(図13A)。興味深いことに、これらの誘導体の一部は、野生型SpCas9を用いて観察された活性に対して、ミスマッチ11/12sgRNAを用いて低減した開裂も示し、それらの誘導体中の置換のサブセットが、野生型SpCas9に対してこのミスマッチ部位に対する向上した特異性を付与することを示

唆した(図13A)。しかしながら、これらの単一置換も置換の組合せも11/12 mismatches sgRNAを用いて観察された活性を完全に排除するために十分でなかった。9および10位における mismatches を担持する追加の sgRNA を使用して野生型 SpCas9、SpCas9-HF1、およびそれらの同一の野生型 SpCas9 誘導体を試験した場合(図13B)、ほとんどの誘導体について mismatches 標的活性の最小の変化のみが観察された。ここでも、これは、これらの潜在的な非標的鎖接触残基における単一、二重、またはさらには三重置換(既に記載された eSpCas9(1.0)および(1.1)バリエーションと等価である)が不完全 matches DNA 部位における活性を排除するには不十分であることを実証した。まとめると、これらのデータは、野生型 SpCas9 バリエーションが matches sgRNA を用いてオンターゲット活性を保持すること、および(野生型 SpCas9 に関して)それら自体でそれらの誘導体中に含有される置換が2つの異なる mismatches DNA 部位に対するヌクレアーゼ活性を排除するために十分でないことを実証する(図13Aおよび13B)。

【0131】

これらの結果を考慮すると、非標的 DNA 鎖に接触し得る残基における1つ以上の追加のアミノ酸置換を担持する SpCas9-HF1 誘導体が、親 SpCas9-HF1 タンパク質に対して特異性をさらに改善し得ることが仮定された。したがって、単一、二重、または三重アラニン置換の組合せを担持する種々の SpCas9-HF1 誘導体を、完全 matches sgRNA (オンターゲット活性を試験するため)ならびに11および12位における mismatches を担持する同一の sgRNA (オフターゲット部位について見出される mismatches 標的部位における活性を評価するため)を使用するヒト細胞ベース EGF 崩壊アッセイにおいて試験した。これらの sgRNA は、図13A~B に使用された同一のものである。この実験により、試験した SpCas9-HF1 誘導体バリエーションのほとんどが、野生型 SpCas9 および SpCas9-HF1 の両方を用いて観察されたものと同等のオンターゲット活性を示すことが明らかになった(図14A)。11/12 mismatches sgRNA を用いると、試験した SpCas9-HF1 誘導体の一部(例えば、SpCas9-HF1+R832A および SpCas9-HF1+K1014A)は、 mismatches sgRNA を用いて開裂の認識可能な変化を示さなかった。しかしながら、重要なことに SpCas9-HF1 誘導体のほとんどが、11/12 mismatches sgRNA を用いて SpCas9-HF1、eSpCas9(1.0)、または eSpCas9(1.1)を用いて観察されたものよりもかなり低い活性を有し、それらの新たなバリエーションのある組合せが、低減した mismatches 標的活性およびしたがって改善された特異性を有することを示唆した(図14A)。11/12 mismatches sgRNA を用いて mismatches 標的活性をほぼバックグラウンドレベルに低減させた16個の SpCas9-HF1 誘導体のうち、9つが、オンターゲット活性に対する最小の効果のみを有すると考えられた(完全 matches sgRNA を使用して評価した; 図14A)。9および10位において意図的に mismatches させた sgRNA を使用した EGF 崩壊アッセイにおけるこれらの SpCas9-HF1 誘導体のサブセットの追加の試験(図14B)により、それらのバリエーションがこの mismatches sgRNA を用いて、SpCas9-HF1(図14b)、eSpCas9(1.1)(図13A)、または野生型 SpCas9 ヌクレアーゼに追加された同一置換(図13B)のいずれかを用いて観察されたものよりも低い活性を有することも明らかになった。重要なことに、5つのバリエーションは、9/10 mismatches sgRNA を用いたこのアッセイにおいてバックグラウンドレベルのオフターゲット活性を示した。

【0132】

次に、非標的鎖のこれらのアラニン置換を、本発明者らの SpCas9-HF1 バリエーションからの Q695A および Q926A 置換のみを含有する SpCas9 バリエーション(ここでは、「二重」バリエーション)と組み合わせることができると否かを試験した。上記の試験 HF1 誘導体の多くがオンターゲット活性の観察可能(および不所望)な減少を示したため、SpCas9-HF1 からの2つの最も重要な置換のみ(Q695A および Q926A; 図1B 参照)を、1つ以上の非標的鎖接触置換と組み合わせることはオンターゲット

10

20

30

40

50

ト活性をレスキューし得るが、それらの置換を S p C a s 9 - H F 1 バリエーションに追加した場合に観察された特異性の増加を依然として維持することが仮定された。したがって、潜在的な非標的 D N A 鎖相互作用位置における単一、二重、または三重アラニン置換の組合せを担持する種々の S p C a s 9 (Q 6 9 5 A / Q 9 2 6 A) 誘導体を、図 1 3 A ~ B に使用された上記の E G F P に標的化される同一の完全マッチ s g R N A (オンターゲット活性を試験するため) ならびに 1 1 および 1 2 位におけるミスマッチを担持する同一の s g R N A (オフターゲット部位について見出されるミスマッチ標的部位における活性を評価するため) を使用するヒト細胞ベース E G F P 崩壊アッセイにおいて試験した。この実験により、試験した S p C a s 9 (Q 6 9 5 A / Q 9 2 6 A) 誘導体バリエーションのほとんどが野生型 S p C a s 9 および S p C a s 9 - H F 1 の両方を用いて観察されたものと同等のオンターゲット活性を示すことが明らかになった (図 1 5) 。重要なことに、S p C a s 9 - H F 1 誘導体の多くは、1 1 / 1 2 ミスマッチ s g R N A を用いて、S p C a s 9 - H F 1 、 e S p C a s 9 (1 . 0) 、または e S p C a s 9 (1 . 1) を用いて観察されたものと比較してかなり低い活性を有し、それらの新たなバリエーションのある組合せが、低減したミスマッチ標的活性およびしたがって改善された特異性を有することを示唆した (図 1 5) 。1 1 / 1 2 ミスマッチ s g R N A を用いてミスマッチ標的活性を、ほぼバックグラウンドレベルに低減させた 1 3 個の S p C a s 9 (Q 6 9 5 A / Q 9 2 6 A) 誘導体のうち、1 つのみがオンターゲット活性に対するかなりの効果を有すると考えられた (完全マッチ s g R N A を使用して評価した ; 図 1 5) 。

【 0 1 3 3 】

概して、これらのデータは、非標的 D N A 鎖に接触し得る位置における S p C a s 9 - H F 1 または S p C a s 9 (Q 6 9 5 A / Q 9 2 6 A) への 1 、 2 、または 3 つのアラニン置換の追加は、(それらの親クローンまたは近年記載された e S p C a s 9 (1 . 0) または (1 . 1) に対して) ミスマッチオフターゲット部位を区別する改善された能力を有する新たなバリエーションをもたらし得ることを実証する。重要なことに、野生型 S p C a s 9 に関するにそれらの同一の置換は、いかなる実質的な特異性利益も提供することが考えられない。

【 0 1 3 4 】

s g R N A - 標的 D N A 相補性界面におけるミスマッチに対する S p C a s 9 - H F 1 および e S p C a s 9 - 1 . 1 の寛容性をより良好に定義および比較するため、スパーサー相補性領域中の全ての考えられる位置における単一ミスマッチを含有する s g R N A を使用してそれらの活性を試験した。S p C a s 9 - H F 1 および e S p C a s 9 - 1 . 1 バリエーションの両方が野生型 S p C a s 9 と比較してほとんどの単一ミスマッチ s g R N A に対する類似の活性を有し、いくつかの例外で S p C a s 9 - H F 1 が e S p C a s 9 - 1 . 1 を上回った (図 1 6) 。

【 0 1 3 5 】

次に、二重突然変異体 (D b = Q 6 9 5 A / Q 9 2 6 A) 、 S p C a s 9 - H F 1 (N 4 9 7 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A) 、 e S p C a s 9 - 1 . 0 (1 . 0 = K 8 1 0 A / K 1 0 0 3 A / R 1 0 6 0 A) 、または e S p C a s 9 - 1 . 1 (1 . 1 = K 8 4 8 A / K 1 0 0 3 A / R 1 0 6 0 A) のいずれかからのアミノ酸置換と、標的鎖 D N A に接触し、または非標的鎖 D N A に潜在的に接触する残基中の追加のアラニン置換との組合せを含有する一部のバリエーションの単一ヌクレオチドミスマッチ寛容性を試験した (図 1 7 A ~ B) 。完全マッチ s g R N A を使用してオンターゲット活性を評価した一方、スパーサー配列中の 4 、 8 、 1 2 、または 1 6 位におけるそのようなミスマッチを担持する s g R N A を使用して単一ヌクレオチドミスマッチ寛容性を評価した (図 1 7 A) 。多数のこれらのバリエーションはオンターゲット活性を維持し、ミスマッチ s g R N A を用いて観察された活性はかなり低減した。これらのバリエーションの 3 つ (Q 6 9 5 A / K 8 4 8 A / Q 9 2 6 A / K 1 0 0 3 A / R 1 0 6 0 A 、 N 4 9 7 A / R 6 6 1 A / Q 6 9 5 A / K 8 5 5 A / Q 9 2 6 A / R 1 0 6 0 A 、 および N 4 9 7 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A / H 9 8 2 A / R 1 0 6 0 A) を、残りの単一ミスマッチ s g R N A (1 ~ 3 、

5～7、9～11、13～15、および17～20位におけるミスマッチを含有する)を用いてさらに試験した。これらのバリエーションは、eSpCas9-1.1と比較してsgRNA中の単一ヌクレオチド置換に対するロバストな不寛容性を実証し、それらの新たなバリエーションの改善された特異性プロファイルを実証した(図17B)。アミノ酸置換の代替組合せを含有する追加のバリエーションヌクレアーゼを、スパーサー中の5、7、および9位におけるミスマッチを含有するsgRNAを使用して試験した(従来のバリエーションがそれらの位置におけるミスマッチを寛容すると考えられたため、それらの特定のミスマッチsgRNAを使用した)(図18)。多数のこれらのヌクレアーゼはミスマッチ部位に対する改善された特異性を有し、オンターゲット活性はわずかに低減したにすぎなかった(図18)。

10

【0136】

突然変異の追加の組合せが、特異性の改善を付与し得るか否かをさらに決定するため、2つの追加のマッチsgRNAを用いるヌクレアーゼバリエーションの大幅に拡張されたパネルを試験して本発明者らのEGFP崩壊活性におけるオンターゲット活性を試験した(図19A)。多数のこれらのバリエーションはロバストなオンターゲット活性を維持し、それらが特異性に対するさらなる改善の生成に有用であり得ることを示唆した(図19B)。多数のこれらのバリエーションを、12、14、16、または18位における単一置換を含有するsgRNAを用いて試験して特異性の改善が観察されたか否かを決定し、それらの位置における単一ヌクレオチドミスマッチに対するより大きい不寛容性を示すことが見出された(図19B)。

20

【0137】

実施例3

SpCas9を用いて行ったものと類似する、黄色ブドウ球菌(*Staphylococcus aureus*) Cas9 (SaCas9)を用いる方針を採用することにより、標的DNA鎖に接触することが公知の残基(図20および21A)、非標的DNAに接触し得る残基(進行中の実験)、およびPAM特異性に影響し得ると本発明者らが既に示した残基(図21B)中のアラニン置換を導入することによりSaCas9の特異性を改善するための実験を実施した。標的鎖DNA骨格に接触し得る残基としては、Y211、Y212、W229、Y230、R245、T392、N419、L446、Y651、およびR654が挙げられ；非標的鎖DNAに接触し得る残基としては、Q848、N492、Q495、R497、N498、R499、Q500、K518、K523、K525、H557、R561、K572、R634、R654、G655、N658、S662、N668、R686、K692、R694、H700、K751が挙げられ；PAMに接触する残基としては、E782、D786、T787、Y789、T882、K886、N888、A889、L909、K929、N985、N986、R991、およびR1015が挙げられる。予備実験において、標的鎖DNA接触残基またはPAM接触残基(それぞれ図21AおよびB)中の単一アラニン置換(またはそれらの一部の組合せ)は、オンターゲットEGFP崩壊活性に対する可変効果を有し(完全マッチsgRNAを使用)、オフターゲット開裂を排除し得なかった(11および12位においてミスマッチするsgRNAを使用した場合)。興味深いことに、HF1中のSpCas9突然変異は、類似ミスマッチ標的/s gRNAペアを用いてオフターゲット活性を完全に停止させ得ず、(SpCas9を用いて観察されたとおり)標的鎖/非標的鎖置換の組合せを含有するバリエーションがそのような部位における特異性を改善するために必要であり得ることを示唆した。

30

40

【0138】

潜在的な標的鎖DNA接触を突然変異させてSaCas9特異性を改善する方針をさらに評価するため、sgRNA中の19および20位におけるミスマッチを寛容する突然変異の単一、二重、三重、および四重組合せの潜在性を試験した(図22AおよびB)。これらの組合せにより、Y230およびR245におけるアラニン置換が、他の置換と組み合わせた場合、ミスマッチ部位をより良好に区別する能力により判断されたとおり、特異

50

性を増加させ得ることが明らかになった。

【0139】

次に、これらの三重アラニン置換バリエーションの2つ(Y211A/Y230A/R245AおよびY212A/Y230A/R245A)のオンターゲット遺伝子崩壊活性を、EGFP中の4つのオンターゲット部位で試験した(マッチ部位#1~4;図23)。これらのバリエーションは、マッチ部位1および2についてロバストなオンターゲット活性を維持したが、部位3および4を用いてオンターゲット活性の約60~70%の損失を示した。これらの三重アラニン置換バリエーションの両方は、標的部位1~4のスペーサー中の種々の位置における二重ミスマッチを担持するsgRNAを使用することにより判断されたとおり、野生型SaCas9に対して特異性を大幅に改善した(図23)。

10

【0140】

これらのアラニン置換の二重および三重組合せ(それぞれ図24AおよびB)を担持するSaCas9バリエーションをオンターゲット活性について6つの内在性部位に対して試験し、21位(最もPAMから遠位の位置は、区別が困難なミスマッチであると予測される)における単一ミスマッチを含有するsgRNAを使用して特異性の改善を評価した。一部の例において、マッチsgRNAを用いるオンターゲット活性はバリエーションを用いて維持した一方、21位においてミスマッチするsgRNAを用いて「オフターゲット」活性を排除した(図24AおよびB)。他の場合、マッチsgRNAを用いて活性のわずかな損失から完全な損失が観察された。

20

【0141】

参考文献

1. Sander, J. D. & Joung, J. K. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat Biotechnol* 32, 347-355 (2014).
2. Hsu, P. D., Lander, E. S. & Zhang, F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell* 157, 1262-1278 (2014).
3. Doudna, J. A. & Charpentier, E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science* 346, 1258096 (2014).
4. Barrangou, R. & May, A. P. Unraveling the potential of CRISPR-Cas9 for gene therapy. *Expert Opin Biol Ther* 15, 311-314 (2015).
5. Jinek, M. et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816-821 (2012).
6. Sternberg, S. H., Redding, S., Jinek, M., Greene, E. C. & Doudna, J. A. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* 507, 62-67 (2014).
7. Hsu, P. D. et al. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol* 31, 827-832 (2013).
8. Tsai, S. Q. et al. GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat Biotechnol*

30

40

50

33, 187 - 197 (2015) .

9. Hou, Z. et al. Efficient genome engineering in human pluripotent stem cells using Cas9 from *Neisseria meningitidis*. *Proc Natl Acad Sci USA* (2013) .

10. Fonfara, I. et al. Phylogeny of Cas9 determines functional exchangeability of dual-RNA and Cas9 among orthologous type II CRISPR-Cas systems. *Nucleic Acids Res* 42, 2577 - 2590 (2014) .

10

11. Esvelt, K.M. et al. Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. *Nat Methods* 10, 1116 - 1121 (2013) .

12. Cong, L. et al. Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819 - 823 (2013) .

13. Horvath, P. et al. Diversity, activity, and evolution of CRISPR loci in *Streptococcus thermophilus*. *J Bacteriol* 190, 1401 - 1412 (2008) .

20

14. Anders, C., Niewoehner, O., Duerst, A. & Jinek, M. Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature* 513, 569 - 573 (2014) .

15. Reyon, D. et al. FLASH assembly of TALENs for high-throughput genome editing. *Nat Biotechnol* 30, 460 - 465 (2012) .

16. Fu, Y. et al. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol* 31, 822 - 826 (2013) .

30

17. Chen, Z. & Zhao, H. A highly sensitive selection method for directed evolution of homing endonucleases. *Nucleic Acids Res* 33, e154 (2005) .

18. Doyon, J.B., Pattanayak, V., Meyer, C.B. & Liu, D.R. Directed evolution and substrate specificity profile of homing endonuclease I-SceI. *J Am Chem Soc* 128, 2477 - 2484 (2006) .

40

19. Jiang, W., Bikard, D., Cox, D., Zhang, F. & Marraffini, L.A. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat Biotechnol* 31, 233 - 239 (2013) .

20. Mali, P. et al. RNA-guided human genome engineering via Cas9. *Science* 339, 823 - 826 (2013) .

21. Hwang, W.Y. et al. Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat Biotechnol* 31, 227 - 229 (2013) .

50

22. Chylinski, K., Le Rhun, A. & Charpentier, E. The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems. *RNA Biol* 10, 726-737 (2013).

23. Kleinstiver, B. P., Fernandes, A. D., Gloor, G. B. & Edgell, D. R. A unified genetic, computational and experimental framework identifies functionally relevant residues of the homing endonuclease I-BmoI. *Nucleic Acids Res* 38, 2411-2427 (2010).

10

24. Gagnon, J. A. et al. Efficient mutagenesis by Cas9 protein-mediated oligonucleotide insertion and large-scale assessment of single-guide RNAs. *PLoS One* 9, e98186 (2014).

【0142】

配列

配列番号271 - JDS246:CMV-T7-ヒトSpCas9-NLS-3xFLAG

ヒトコドン最適化化膿性連鎖球菌 (*S. pyogenes*) Cas9、通常フォント、NLS、二重下線、3xFLAGタグ、太字：

20

【化 4】

ATGGATAAAAAGTATTCTATTGGTTTAGACATCGGCACTAATTCGGTTGGATGGGCTGTCATAACCGAT
 GAATACAAAGTACCTTCAAAGAAATTTAAGGTGTTGGGGAACACAGACCGTCATTCGATTAAAAAGAAT
 CTTATCGGTGCCCTCCTATTTCGATAGTGGCGAAACGGCAGAGGGCACTCGCCTGAAACGAACCGCTCGG
 AGAAGGTATACACGTCGCAAGAACCGAATATGTTACTTTACAAGAAATTTTAGCAATGAGATGGCCAAA
 GTTGACGATTCTTTCTTTCCCGTTTGAAGAGTCCCTTCCTTGTGCAAGAGGACAAGAAACATGAACGG
 CACCCCATCTTTGGAAACATAGTAGATGAGGTGGCATAATCATGAAAAGTACCCAACGATTTATCACCTC
 AGAAAAAGCTAGTTGACTCAACTGATAAAGCGGACCTGAGGTAACTTACTTGGCTCTTGCCCATATG
 ATAAAGTTCGGTGGGCACTTTCTCATTTGAGGGTGATCTAAATCCGGACAACCTCGGATGTCGACAACTG
 TTCATCCAGTTAGTACAAACCTATAATCAGTTGTTTGAAGAGAACCCTATAAATGCAAGTGGCGTGGAT
 GCGAAGGCTATTCTTAGCGCCCGCTCTCTAAATCCCGACGGCTAGAAAACCTGATCGCACAAATTACCC
 GGAGAGAAGAAAAATGGGTGTTTCGGTAACCTTATAGCGCTCTCACTAGGCCCTGACACCAAATTTTAAG
 TCGAACTTCGACTTAGCTGAAGATGCCAAATTGCAGCTTAGTAAGGACACGTACGATGACGATCTCGAC
 AATCTACTGGCACAATTTGGAGATCAGTATGCGGACTTATTTTGGCTGCCAAAAACCTTAGCGATGCA
 ATCCTCCTATCTGACATACTGAGAGTTAATACTGAGATTACCAAGGCGCGCTTATCCGCTTCAATGATC
 AAAAGGTACGATGAACATCACCAAGACTTGACACTTCTCAAGGCCCTAGTCCGTCAGCAACTGCCTGAG
 AAATATAAGGAAATATTCTTTGATCAGTCGAAAAACGGGTACGCAGGTTATATTGACGGCGGAGCGAGT
 CAAGAGGAATTCTACAAGTTTATCAAACCCATATTAGAGAAGATGGATGGGACGGAAGAGTTGCTTGTA
 AAATCAATCGCGAAGATCTACTGCGAAAGCAGCGGACTTTTCGACAACGGTAGCATTCACATCAAATC
 CACTTAGGCGAATTGCATGCTATACTTAGAAGGCAGGAGGATTTTTATCCGTTCTCAAAGACAATCGT
 GAAAAGATTGAGAAAATCCTAACCTTTTCGCATACCTTACTATGTGGGACCCCTGGCCCGAGGGAACCTC
 CGGTTTCGCATGGATGACAAGAAAGTCCGAAGAACGATTACTCCATGGAATTTTGAGGAAGTTGTCGAT
 AAAGGTGCGTCAGCTCAATCGTTTCATCGAGAGGATGACCAACTTTGACAAGAAATTTACCGAACGAAAA
 GTATTGCGCTAAGCACAGTTTACTTTACGAGTATTTACAGTGTACAATGAACCTCACGAAAGTTAAGTAT
 GTCACTGAGGGCATGCGTAAACCCGCTTTCTAAGCGGAGAACAGAAGAAAGCAATAGTAGATCTGTTA
 TTCAAGACCAACCGCAAAGTGACAGTTAAGCAATTGAAAGAGGACTACTTTAAGAAAATTGAATGCTTC
 GATTCTGTCGAGATCTCCGGGTAGAAGATCGATTTAATGCGTCACTTGGTACGTATCATGACCTCCTA
 AAGATAATTAAAGATAAGGACTTCCTGGATAACGAAGAGAATGAAGATATCTTAGAAGATATAGTGTTG
 ACTCTTACCCTCTTTGAAGATCGGGAAATGATTGAGGAAAGACTAAAAACATACGCTCACCTGTTTCGAC
 GATAAGGTTATGAAACAGTTAAAGAGGCGTCGTATACGGGCTGGGGACGATTGTCGCGGAAACTTATC
 AACGGGATAAGAGACAAGCAAAGTGGTAAACCTATTCTCGATTTTCTAAAGAGCGACGGCTTCGCCAAT
 AGGAACCTTATGACAGCTGATCCATGATGACTCTTTAACCTTCAAAGAGGATATACAAAAGGCACAGGTT
 TCCGGACAAGGGGACTCATTTGCACGAACATATTGCGAATCTTGCTGGTTGCGCCAGCCATCAAAAAGGGC
 ATACTCCAGACAGTCAAAGTAGTGGATGAGCTAGTTAAGGTGATGGGACGTCACAAAACCGGAAAACATT
 GTAATCGAGATGGCAGCGGAAAATCAAACGACTCAGAAGGGGCAAAAAACAGTCGAGAGCGGATGAAG
 AGAATAGAAGAGGGTATTAAAGAACTGGGCAGCCAGATCTTAAAGGAGCATCTGTGAAAAATACCCAA
 TTGCAGAACGAGAACTTTACCTCTATTACCTACAAATGGAAGGGACATGTATGTTGATCAGGAACCTG
 GACATAAACCGTTTATCTGATTACGACGTCGATCACATTGTACCCCAATCCTTTTGAAGGACGATTCA
 ATCGACAATAAAGTGCTTACACGCTCGGATAAGAACCGAGGGAAAAGTGACAAATGTTCCAAGCGAGGAA
 GTCGTAAAGAAAATGAAGAACTATTGGCGGCAGCTCCTAAATGCGAACTGATAACGCAAAGAAAGTTC
 GATAACTTAACATAAAGCTGAGAGGGGTGGCTTGCTGAACTTGACAAGGCCGGATTATTAAACGTCAG
 CTCGTGGAAACCCGCCAAATCACAAAGCATGTTGCACAGATACTAGATTCCCGAATGAATACGAATAC
 GACGAGAACGATAAGCTGATTCGGGAAGTCAAAGTAATCACTTTAAAGTCAAATTTGGTGTGCGACTTC
 AGAAAGGATTTTCAATTCTATAAAGTTAGGGAGATAAATAACTACCACCATGCGCAGCAGCTTATCTT
 AATGCCGTCGTAGGGACCGCACTCATTAAGAAATACCCGAAGCTAGAAAGTGAGTTTGTGTATGGTGAT
 TACAAAGTTTATGACGTCGTAAGATGATCGCGAAAAGCGAACAGGAGATAGGCAAGGCTACAGCCAAA
 TACTTCTTTTATTCTAACATTATGAATTTCTTTAAGACGGAAATCACTCTGGCAAACGGAGAGATACGC
 AAACGACCTTTAATTGAAACCAATGGGGAGACAGGTGAAATCGTATGGGATAAGGGCCGGGACTTCGCG
 ACGGTGAGAAAGTTTGTCCATGCCCCAAGTCAACATAGTAAAGAAAACCTGAGGTGCAGACCGGAGGG
 TTTTCAAAGGAATCGATTCTTCCAAAAGGAATAGTGATAAGCTCATCGCTCGTAAAAAGGACTGGGAC
 CCGAAAAAGTACGGTGGCTTCGATAGCCCTACAGTTGCCTATTCTGTCTAGTAGTGCCAAAAGTTGAG
 AAGGGAAAATCCAAGAACTGAAGTCAGTCAAAGAATTATTGGGGATAACGATTATGGAGCGCTCGTCT
 TTTGAAAAGAACCCCATCGACTTCCTTGAGGCGAAAGGTTACAAGGAAGTAAAAAGGATCTCATAATT
 AAATACCAAAGTATAGTCTGTTTGAAGTTAGAAAATGGCCGAAAACGGATGTTGGCTAGCGCCGGAGAG
 CTTCAAAAGGGGAACGAACTCGCACTACCGTCTAAATACGTGAATTTCTGTATTTAGCGTCCCATTAC
 GAGAAGTTGAAAGGTTACCTGAAGATAACGAACAGAAGCAACTTTTGTGAGCAGCACAAACATTAT
 CTCGACGAAATCATAGAGCAAATTTCCGAATTGAGTAAGAGAGTCATCCTAGCTGATGCCAATCTGGAC
 AAAGTATTAAAGCGCATACAACAAGCACAGGGATAAACCATACGTGAGCAGGCGGAAAATATTATCCAT
 TTGTTTACTCTTACCAACCTCGGCGCTCCAGCGCAATTCAGTATTTTACACACACGATAGATCGCAAA
 CGATACACTTCTACCAAGGAGGTGCTAGACGCGACACTGATTACCAATCCATCACGGGATTATATGAA
 ACTCGGATAGATTTGTACAGCTTGGGGGTGACGGATCCCCAAGAAGAAGAGGAAAGTCTCGAGCGGAC
 TACAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATGACAAGTGA

10

20

30

40

【0143】

配列番号 272 - VP12:CMV-T7-ヒトSpCas9-HF1(N497A
 、R661A、Q695A、Q926A)-NLS-3xFLAG

50

ヒトコドン最適化化膿性連鎖球菌 (*S . p y o g e n e s*) C a s 9、通常フォント、改変コドン、小文字、N L S、二重下線、3 x F L A G タグ、太字：

【化5】

ATGGATAAAAAGTATTCTATTGGTTTAGACATCGGCCTAATTCCGTTGGATGGGCTGTCATAACCGAT
GAATACAAAGTACCTTCAAAGAAATTTAAGGTGTTGGGGAACACAGACCGTCATTTCGATTAAAAAGAAT
CTTATCGGTGCCCTCCTATTTCGATAGTGGCGAAACGGCAGAGGCGACTCGCCTGAAACGAACCGCTCGG
AGAAGGTATACACGTCGCAAGAACCGAATATGTTACTTACAAGAAATTTTAGCAATGAGATGGCCAAA
GTTGACGATTCTTTCTTTCACCGTTTGGGAAGAGTCCTTCCCTGTGCGAAGAGGACAAGAAACATGAACGG
CACCCCATCTTTGGAAACATAGTAGATGAGGTGGCATATCATGAAAAGTACCCAACGATTATCACCTC
AGAAAAAAGCTAGTTGACTCAACTGATAAAGCGGACCTGAGGTTAATCTACTTGGCTCTTGCCCATATG
ATAAAGTTCGGTGGGCACTTTCTCATTGAGGGTGATCTAAATCCGGACAACCTCGGATGTGACAAAACCTG
TTCATCCAGTTAGTACAAACCTATAATCAGTTGTTTGAAGAGAACCCTATAAATGCAAGTGGCGTGGAT
GCGAAGGCTATTCTTAGCGCCCGCTCTCTAAATCCCGACGGCTAGAAAACCTGATCGCACAAATTACCC
GGAGAGAAGAAAAATGGGTGTTTCGGTAACCTTATAGCGCTCTCACTAGGCTGACACCAAATTTTAAG
TCGAACCTTCGACTTAGCTGAAGATGCCAAATTGCAGCTTAGTAAGGACACGTACGATGACGATCTCGAC
AATCTACTGGCACAAATTGGAGATCAGTATGCGGACTTATTTTGGCTGCCAAAAACCTTAGCGATGCA
ATCCTCCTATCTGACATACTGAGAGTTAATACTGAGATTACCAAGGCGCGGTTATCCGCTTCAATGATC
AAAAGGTACGATGAACATACCAAGACTTGACACTTCTCAAGGCCCTAGTCCGTCAGCAACTGCCTGAG
AAATATAAGGAATATTCTTTGATCAGTCGAAAAACGGGTACGCAGGTTATATTGACGGCGGAGCGAGT
CAAGAGGAATTCTACAAGTTTATCAAACCCATATTAGAGAAGATGGATGGGACGGAAGAGTTGCTTGTA
AAACTCAATCGCGAAGATCTACTGCGAAAGCAGCGGACTTTGACACAACGGTAGCATTCCACATCAAATC
CACTTAGGCGAATTGCATGCTATACTTAGAAGGCAGGAGGATTTTATCCGTTCTCAAAGACAATCGT
GAAAAGATTGAGAAAATCCTAACCTTTTCGCATACCTTACTATGTGGGACCCCTGGCCCGAGGGAACCTCT
CGGTTTCGCATGGATGACAAGAAAGTCCGAAGAAACGATTACTCCCTGGAATTTTGAGGAAGTTGTCGAT
AAAGGTGCGTCAGCTCAATCGTTCATCGAGAGGATGACCg_{cc}TTTGACAAGAAATTTACCGAACGAAAAA
GTATTGCCTAAGCACAGTTTACTTTACGAGTATTTACAGTGTACAATGAACCTCACGAAAGTTAAGTAT
GTCATTGAGGACATGCGTAAACCCCGCTTTCTAAGCGGAGAACAGAAGAAAGCAATAGTAGATCTGTTA
TTCAAGACCAACCGCAAGTACAGTTAAGCAATTGAAAGAGGACTACTTTAAGAAAAATGAATGCTTTC
GATTCTGTGCGAGATCTCCGGGTAGAAGATCGATTAAATGCGTCACTTGGTACGTATCATGACCTCCTA
AAGATAATTAAAGATAAGGACTTCTTGATAACGAAGAGAATGAAGATATCTTAGAAGATATAGTGTTG
ACTCTTACCCTCTTTGAAGATCGGGAAATGATTGAGGAAAGACTAAAAACATACGCTCACCTGTTTCGAC
GATAAGGTTATGAAACAGTTAAAGAGGCGTCGTATACGGGCTGGGGAg_{cc}TTGTCGCGGAAACTTATC
AACGGGATAAGAGACAAGCAAAGTGGTAAACCTATTCTCGATTTTCTAAAGAGCGACGGCTTCGCCAAT
AGGAACCTTTATGg_{cc}CTGATCCATGATGACTCTTTAACCTTCAAAGAGGATATACAAAAGGCACAGGTT
TCCGGACAAGGGGACTCATTGCACGAACATATTGCGAATCTTGCTGGTTCGCCAGCCATCAAAAAGGGC
ATACTCCAGACAGTCAAAGTAGTGGATGAGCTAGTTAAGGTGATGGGACGTACAAAACCGGAAACATT
GTAATCGAGATGGCACGCGAAAAATCAAACGACTCAGAAGGGGCAAAAAACAGTCGAGAGCGGATGAAG
AGAATAGAAGAGGGTATTAAAGAACTGGGCAGCCAGATCTTAAAGGAGCATCCTGTGAAAAATACCCAA
TTGCAGAACGAGAACTTTACCTCTATTACCTACAAAATGGAAGGGACATGTATGTTGATCAGGAACCTG
GACATAAACCGTTTATCTGATTACGACGTCGATCAGATTGTACCCCAATCCTTTTGAAGGACGATTCA
ATCGACAATAAAGTGCTTACACGCTCGGATAAGAACCAGGGGAAAAGTGACAAATGTTCCAAGCGAGGAA
GTCGTAAAGAAAATGAAGAACTATTGGCGGCAGCTCCTAAATGCGAAACCTGATAACGCAAGAAAGTTC
GATAACTTAACTAAAGCTGAGAGGGGTGGCTTGTCTGAACCTTGACAAGGCCGGATTATTAAACGTCAG
CTCGTGGAAACCCGCG_{cc}ATCACAAAGCATGTTGCGCAGATACTAGATTCCCGAATGAATACGAAATAC
GACGAGAACGATAAGCTGATTGCGGAAGTCAAAGTAAATCACTTTAAAGTCAAATTTGGTGTGCGAATTC
AGAAAGGATTTTCAATTCTATAAAGTTAGGGAGATAAATAACTACCACCATGCGCACGACGCTTATCTT
AATGCCGTCGTAGGGACCGCACTCATTAAGAAATACCCGAAGCTAGAAAGTGAGTTTGTGTATGGTGAT
TACAAAGTTTATGACGTCGGTAAGATGATCGCGAAAAGCGAACAGGAGATAGGCAAGGCTACAGCCAAA
TACTTCTTTTATTTCTAACATTATGAATTTCTTTAAGACGGAAATCACTCTGGCAAACGGAGAGATACGC
AAACGACCTTTAATTGAAACCAATGGGGAGACAGGTGAAATCGTATGGGATAAGGGCCGGGACTTCGCG
ACGGTGAGAAAAGTTTGTCCATGCCCAAGTCAACATAGTAAAGAAAACCTGAGGTGACAGCCGGAGGG
TTTTCAAAGGAATCGATTCTTCCAAAAAGGAATAGTGATAAGCTCATCGCTCGTAAAAAGGACTGGGAC
CCGAAAAAGTACGGTGGCTTCGATAGCCCTACAGTTGCCTATTCTGTCTAGTAGTGGCAAAAGTTGAG
AAGGGAAAAATCCAAGAACTGAAGTCAGTCAAAGAATTATTTGGGGATAACGATTATGGAGCGCTCGTCT
TTTGAAAAGAACCCATCGACTTCCTTGAGGCGAAAGGTTACAAGGAAGTAAAAAAGGATCTCATAATT
AAACTACCAAAGTATAGTCTGTTTGAGTTAGAAAATGGCCGAAAACGGATGTTGGCTAGCGCCGGAGAG
CTTCAAAGGGGAACGAACTCGCACTACCGTCTAAATACGTGAATTTCTGTATTTAGCGTCCCATTAC
GAGAAGTTGAAAGGTTACCTGAAGATAACGAACAGAAGCAACTTTTTGTTGAGCAGCACAAACATTAT
CTCGACGAAATCATAGAGCAAATTTCCGAATTACGTAAGAGAGTCATCCTAGCTGATGCCAATCTGGAC
AAAGTATTAAGCGCATACAAACAGCACAGGGATAAACCCATACGTGAGCAGGCGGAAAAATATTATCCAT
TTGTTTACTCTTACCAACCTCGGCGCTCCAGCCGCAATCAAGTATTTTGACACAACGATAGATCGCAA
CGATACACTTCTACCAAGAGGTGCTAGACGCGCACTGATTCACCAATCCATCACGGGATTATATGAA
ACTCGGATAGATTTGTACAGCTTGGGGGTGACGGAATCCCCCAAGAAGAAGAGGAAAGTCTCGAGCGAC
TACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATGACAAAGTGA

10

20

30

40

配列番号 273 - MSP2135 : CMV - T7 - ヒト SpCas9 - HF2 (N497A、R661A、Q695A、Q926A、D1135E) - NLS - 3xFLAG
ヒトコドン最適化化膿性連鎖球菌 (S. pyogenes) Cas9、通常フォント、改変コドン、小文字、NLS、二重下線、3xFLAGタグ、太字：

【化 6】

ATGGATAAAAAGTATTTCTATTGGTTTACACATCGGCACTAATTCCGTTGGATGGGCTGTCTATAACCGAT
GAATACAAAGTACCTTCAAAGAAATTTAAGGTGTTGGGGAACACAGACCGTCATTTCGATTAAAAAGAAAT
CTTATCGGTGCCCTCCTATTTCGATAGTGGCGAAACGGCAGAGGCGACTCGCCTGAAACGAACCGCTCGG
AGAAGGTATACACGTCGCAAGAACCGAATATGTTACTTACAAGAAATTTTAGCAATGAGATGGCCAAA
GTTGACGATTCTTTCTTTACCGTTTGGGAAGAGTCCCTTCTTGTCTGAAGAGGACAAGAAACATGAACGG
CACCCCATCTTTGGAAACATAGTAGATGAGGTGGCATATCATGAAAAGTACCAACGATTTATCACCTC
AGAAAAAGCTAGTTGACTCAACTGATAAAGCGGACCTGAGGTTAATCTACTTGGGCTCTTGCCCATATG
ATAAAGTTCGGTGGGCACTTTCTCATTTGAGGGTGATCTAAATCCGGACAACCTCGGATGTCTGACAAACTG
TTCATCCAGTTAGTACAAACCTATAATCAGTTGTTTGAAGAGAACCCATATAATGCAAGTGGCGTGGAT
GCGAAGGCTATTCTTAGCGCCCGCCTCTCTAAATCCCGACGGCTAGAAAACCTGATCGCAAAATTTACCC
GGAGAGAAGAAAAATGGGTTTGTTCGGTAACCTTATAGCGCTCTCACTAGGCTGACACCAAAATTTAAG
TCGAACCTTCGACTTAGCTGAAGATGCCAAATTCGAGCTTAGTAAGGACACGTACGATGACGATCTCGAC
AATCTACTGGCACAAATTTGGAGATCAGTATGCGGACTTATTTTGGCTGCCAAAAACCTTAGCGATGCA
ATCCTCCTATCTGACATACTGAGAGTTAATACTGAGATTACCAAGGCGCGGTTATCCGCTTCAATGATC
AAAAGGTACGATGAACATCACCAAGACTTGACACTTCTCAAGGCCCTAGTCCGTGAGCAACTGCCTGAG
AAATATAAGGAAATATCTTTGATCAGTCGAAAAACGGGTACGCAGGTTATATTGACGGCGGAGCGAGT
CAAGAGGAATTTTACAAGTTTATCAAACCCATATTAGAGAAGATGGATGGGACGGAAGAGTTGCTTGTA
AAACTCAATCGCGAAGATCTACTGCGAAAGCAGCGGACTTTGACAAACGGTAGCATTCCACATCAAATC
CACTTAGGCGAATTGCATGCTATACTTAGAAGGCAGGAGGATTTTATCCGTTCTCAAAGACAATCGT
GAAAAGATTGAGAAAACTCTAACCTTTTCGCATACCTTACTATGTGGGACCCCTGGCCCGAGGGAACTCT
CGGTTTCGCATGGATGACAAGAAAGTCCGAAGAAACGATTACTCCCTGGAATTTTGAGGAAGTTGTCTGAT
AAAGGTGCGTCAGCTCAATCGTTTCATCGAGAGGATGACCgcccTTTGACAAGAATTTACCGAACGAAAAA
GTATTGCCTAAGCACAGTTTACTTTACGAGTATTTACAGTGTACAATGAACTCACGAAAGTTAAGTAT
GTCACTGAGGGCATGCGTAAACCCGCCCTTTCTAAGCGGAGAACAGAAGAAACCAATAGTAGATCTGTTA
TTCAAGACCAACCGCAAAGTGACAGTTAAGCAATTGAAAGAGGACTACTTTAAGAAAATTGAATGCTTC
GATTCTGTCTGAGATCTCCGGGGTAGAAGATCGATTTAATGCGTCACTTGGTACGTATCATGACCTCCTA
AAGATAATTAAAGATAAGGACTTCTTGATAACGAAGAGAATGAAGATATCTTAGAAGATATAGTGTG
ACTCTTACCCTCTTTGAAGATCGGGAATGATTGAGGAAAGACTAAAAACATACGCTCACCTGTTTCGAC
GATAAGGTTATGAAACAGTTAAAGAGGCGTCTGCTATACGGGCTGGGGAgcccTTGTCTCGCGAAACTTTATC
AACGGGATAAGAGACAAGCAAAGTGGTAAACTATTCTCGATTTTCTAAAGAGCGACGGCTTCGCCAAT
AGGAACCTTTATGgcccTGATCCATGATGACTCTTTAACCTTCAAAGAGGATATACAAAAGGCACAGGTT
TCCGGACAAGGGGACTCATTGCACGAACATATTGCGAATCTTGCTGGTTTCGCCAGCCATCAAAAAGGGC
ATACTCCAGACAGTCAAAGTAGTGGATGAGCTAGTTAAGGTCATGGGACGTACAAAACCGGAAAACATT
GTAATCGAGATGGCACGCGGAAAATCAAACGACTCAGAAGGGGCAAAAAACAGTCGAGAGCGGATGAAG
AGAATAGAAGAGGGTATTAAAGAACTGGGCAGCCAGATCTTAAAGGAGCATCTGTGGAAAATACCCAA
TTGCAGAACGAGAACTTTACCTCTATTACCTACAAAATGGAAGGGACATGTATGTTGATCAGGAACGTG
GACATAAACCGCTTTATCTGATTACGACGTCGATCACATTTGATCCCAATCCTTTTGAAGGACGATTCA
ATCGAATAAAGTGTCTTACACGCTCGGATAAGAACCGGAGGAAAAGTACCAATGTTCCAAAGCGAGGAA
GTCGTAAAGAAAATGAAGAACTATTGGCGGACGCTCCTAAATGCGAACTGATAACGCAAAGAAAGTTC
GATAACTTAACTAAAGCTGAGAGGGGTGGCTTGTCTGAACTTGACAAGGCCGATTATTAAACGTCAG
CTCGTGGAAACCCCGgcccATCACAAAGCATGTTGCGCAGATACTAGATTCCCGAATGAATACGAAATAC
GACGAGAACGATAAGCTGATTGCGGAAGTCAAAGTAATCACTTTAAAGTCAAATTTGGTGTCTGAGCTTC
AGAAAGGATTTTCAATTTCTATAAAGTTAGGGAGATAAATACTACCACCATGCGCACGACGCTTATCTT
AATGCCGTCTGAGGGACCGCACTCATTAAGAAATACCCGAAGCTAGAAAGTGAGTTTGTGTATGGTGAT
TACAAAGTTTATGACGTCGGTAAGATGATCGCGAAAAGCGAACAGGAGATAGGCAAGGCTACAGCCAAA
TACTTCTTTTATCTAACATTATGAATTTCTTTAAGACGGAATCACTCTGGCAAACGGAGAGATACGC
AAACGACCTTTAATTGAAACCAATGGGGAGACAGGTGAAATCGTATGGGATAAGGGCCGGGACTTCGCG
ACGGTGAGAAAAGTTTTGTCCATGCCCAAGTCAACATAGTAAAGAAAACCTGAGGTGCAGACCGGAGGG
TTTTCAAAGGAATCGATTCTTCCAAAAGGAATAGTGATAAGCTCATCGCTCGTAAAAAGGACTGGGAC
CCGAAAAAGTACGGTGGCTTCgagAGCCCTACAGTTGCCTATTCTGTCTAGTAGTGGCAAAAGTTGAG
AAGGGAAAATCCAAGAACTGAAGTCAGTCAAAGAATTATTGGGGATAACGATTATGGAGCGCTCGTCT
TTTGAAAAGAACCCATCGACTTCTTGAGGCGAAAAGGTTACAAGGAAGTAAAAAGGATCTCATAATT
AAACTACCAAAGTATAGTCTGTTTGAAGTTAGAAAATGGCCGAAAACGGATGTTGGCTAGCGCCGGAGAG
CTTCAAAGGGGAACGAACTCGCACTACCGTCTAAATACGTGAATTTCTGTATTTAGCGTCCCATTAC
GAGAAGTTGAAAGTTCACTGAAGATAACGAACAGAAGCAACTTTTTGTTGAGCAGCACAACATTAT
CTCGACGAATCATAGAGCAAAATTTGGAATTCAGTAAGAGAGTCATCTAGCTGATGCCAATCTGGAC
AAAGTATTAAGCGCATACAACAAGCACAGGGATAAACCCATACGTGAGCAGGCGGAAAATATTATCCAT
TTGTTTACTCTTACCAACCTCGGCGCTCCAGCGCATTTCAAGTATTTTGACACAACGATAGATCGCAA
CGATACACTTCTACCAAGGAGGTGCTAGACGCGCACTGATTCACCAATCCATCACGGGATTATATGAA
ACTCGGATAGATTTGTACAGCTTGGGGGTGACGGATCCCCCAAGAAGAAGAGGAAAGTCTCGAGCGAC
TACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATGACAAGTGA

10

20

30

40

【0145】

配列番号 274 - MSP2133:CMV-T7-ヒトSpCas9-HF4(Y4
50A、N497A、R661A、Q695A、Q926A)-NLS-3xFLAG

50

ヒトコドン最適化化膿性連鎖球菌 (*S . p y o g e n e s*) Cas 9、通常フォント、改変コドン、小文字、N L S、二重下線、3 x F L A G タグ、太字：

【化 7】

ATGGATAAAAAGTATTCTATTGGTTTAGACATCGGCACTAATTCGGTTGGATGGGCTGTCATAACCGAT
GAATACAAAGTACCTTCAAAGAAATTTAAGGTGTGGGGAACACAGACCGTCATTGATTAAAAAGAAT
CTTATCGGTGCCCTCCTATTTCGATAGTGGCGAAACGGCAGAGGCGACTCGCCTGAAACGAACCGCTCGG
AGAAGGTATACACGTCGCAAGAACC GAATATGTTACTTACAAGAAATTTTAGCAATGAGATGGCCAAA
GTTGACGATTCTTTCTTTCACCGTTTGGAAAGAGTCCCTTCTTGTGCGAAGAGGACAAGAAACATGAACGG
CACCCCATCTTTGGAAACATAGTAGATGAGGTGGCATATCATGAAAAGTACCCAACGATTTATCACCTC
AGAAAAAAGCTAGTTGACTCAACTGATAAAGCGGACCTGAGGTTAATCTACTTGGCTCTTGCCCATATG
ATAAAGTTCCGTGGGCACTTTCTCATTGAGGTGATCTAAATCCGGACAACCTCGGATGTCGACAAACTG
TTCATCCAGTTAGTACAAACCTATAATCAGTTGTTTGAAGAGAACCCTATAAATGCAAGTGGCGTGGAT
GCGAAGGCTATTCTTAGCGCCCGCTCTCTAAATCCCGACGGCTAGAAAACCTGATCGCACAAATTACCC
GGAGAGAAGAAAAATGGGTTGTTTCGGTAACCTTATAGCGCTCTCACTAGGCCTGACACCAAATTTTAAG
TCGAACCTTCGACTTAGCTGAAGATGCCAAATTGCAGCTTAGTAAGGACACGTACGATGACGATCTCGAC
AATCTACTGGCACAAATTGGAGATCAGTATGCGGACTTATTTTGGCTGCCAAAAACCTTAGCGATGCA
ATCCTCCTATCTGACATACTGAGAGTTAATACTGAGATTACCAAGGCGCGGTTATCCGCTTCAATGATC
AAAAGGTACGATGAACATACCAAGACTTGACACTTCTCAAGGCCCTAGTCCGTGAGCAACTGCCGTGAG
AAATATAAGGAAATATCTTTGATCAGTCGAAAAACGGGTACGCAGGTTATATTGACGGCGGAGCGAGT
CAAGAGGAATTCTACAAGTTTATCAAACCCATATTAGAGAAGATGGATGGGACGGAAGAGTTGCTTGTA
AAACTCAATCGCGAAGATCTACTGCGAAAGCAGCGGACTTTCGACAACGGTAGCATTCCACATCAAATC
CACTTAGGCGAATTGCGATGCTATACTTAGAAGGCAGGAGGATTTTATCCGTTCTCAAAGACAATCGT
GAAAAGATTGAGAAAACTTAACCTTTTCGCATACCTGccTATGTGGGACCCCTGGCCCGAGGGAACCTCT
CGGTTTCGCATGGATGACAAGAAAGTCCGAAGAACGATTACTCCCTGGAATTTTGAGGAAGTTGTCGAT
AAAGGTGCGTCAGCTCAATCGTTTCATCGAGAGGATGACCgcccTTTGACAAGAAATTTACCGAACGAAAAA
GTATTGCCCTAAGCACAGTTTACTTTACGAGTATTTACAGTGTACAATGAACCTCACGAAAGTTAAGTAT
GTCATGAGGGCATGCGTAAACCCGCTTTCTAAGCGGAGAACAGAAAGCAATAGTAGATCTGTTA
TTCAAGACCAACCGCAAGTGAAGTGAAGCAATTTGAAGAGGACTACTTTAAGAAAAATTTGAATGCTTTC
GATTCTGTGCGAGATCTCCGGGGTAGAAGATCGATTTAATGCGTCACTTGGTACGTATCATGACCTCCTA
AAGATAATTAAAGATAAGGACTTCTGGATAACGAAGAGAATGAAGATATCTTAGAAGATATAGTGTTG
ACTCTTACCCTCTTTGAAGATCGGGAAATGATTGAGGAAAGACTAAAAACATACGCTCACCTGTTTCGAC
GATAAGGTTATGAAACAGTTAAAGAGGCGTCGCTATACGGGCTGGGGAGcccTTGTCGCGGAAACCTTATC
AACGGGATAAGAGACAGCAAAAGTGGTAAACCTATTCTCGATTTTCTAAAGAGCGACGGCTTCGCCAAT
AGGAACCTTTATGgcccTGATCCATGATGACTCTTTAACCTTCAAAGAGGATATACAAAAGGCACAGGTT
TCCGGACAAGGGGACTCATTGCACGAACATATTGCGAATCTTGCTGGTTCCGACGCCATCAAAAAGGGC
ATACTCCAGACAGTCAAAGTAGTGGATGAGCTAGTTAAGGTCATGGGACGTCACAAACCGGAAACATT
GTAATCGAGATGGCAGCGGAAATCAAACGACTCAGAAGGGGCAAAAAACAGTCGAGAGCGGATGAAG
AGAATAGAAGAGGGTATTAAAGAACTGGGCAGCCAGATCTTAAAGGAGCATCCTGTGGAAATACCCAA
TTGCAGAACGAGAACTTTACCTCTATTACCTACAAAATGGAAGGGACATGTATGTTGATCAGGAACCTG
GACATAAACCGTTTATCTGATTACGACGTCGATCACATTGTACCCCAATCCTTTTGAAGGACGATTCA
ATCGACAATAAAGTGCTTACACGCTCGGATAAGAACCGAGGGAAAAGTGACAAATGTTCCAAGCGAGGAA
GTCGTAAAGAAAATGAAGAACTATTGGCGGCAGCTCCTAAATGCGAACTGATAACGCAAGAAAGTTC
GATAACTTAACCTAAAGCTGAGAGGGGTGGCTTGTCTGAACCTTGACAAGGCCGATTTATTAAACGTGAG
CTCGTGGAAACCCGCGcccATCACAAAGCATGTTGCGCAGATACTAGATTTCCGAATGAATACGAAATAC
GACGAGAACGATAAGCTGATTCGGGAAGTCAAAGTAATCACTTTAAAGTCAAAATTTGGTCTCGGACTTC
AGAAAGGATTTTCAATTTCTATAAAGTTAGGGAGATAAATAACTACCACATGCGCACGAGCTTATCTT
AATGCCGTCGTAGGGACCGCACTCATTAAGAAATACCCGAAGCTAGAAAGTGAGTTTGTGTATGGTGAT
TACAAAGTTTATGACGTCGTAAGATGATCGCGAAAAGCGAACAGGAGATAGGCAAGGCTACAGCCAAA
TACTTCTTTTATTTCTAACATTATGAATTTCTTTAAGACGGAAATCACTCTGGCAAACGGAGAGATACGC
AAACGACCTTTAATTGAAACCAATGGGGAGACAGGTGAAATCGTATGGGATAAGGGCCGGGACTTCGCG
ACGGTGAGAAAAGTTTTGTCCATGCCCAAGTCAACATAGTAAAGAAAACCTGAGGTGACAGCCGGAGGG
TTTTCAAAGGAATCGATTCTTCCAAAAGGAATAGTGATAAGCTCATCGCTCGTAAAAAGGACTGGGAC
CCGAAAAAGTACGGTGGCTTCGATAGCCCTACAGTTGCCTATTCTGTCTAGTAGTGCCAAAAGTTGAG
AAGGGAAAATCCAGAAACTGAAGTCAGTCAAAGAAATTTATGGGATAACGATTATGGAGCGCTCGTCT
TTTGAAAAGAACCCCTCGACTTCTTGGGCGAAAGGTTACAAGGAAGTAAAAAAGGATCTCATAAAT
AAACTACCAAAGTATAGTCTGTTGAGTTAGAAAATGGCCGAAAACGGATGTTGGCTAGCGCCGGAGAG
CTTCAAAGGGGAACGAACTCGCACTACCGCTCTAAATACGTGAATTTCTGTATTTAGCGTCCCATTAC
GAGAAGTTGAAAGGTTACCTGAAGATAACGAACAGAAGCAACTTTTTGTTGAGCAGCACAAACATTAT
CTCGACGAAATCATAGAGCAAATTTCCGAATTCAGTAAGAGAGTCATCCTAGCTGATGCCAATCTGGAC
AAAGTATTAAGCGCATACAACAAGCACAGGGATAAACCCATACGTGAGCAGGCGGAAAATATTATCCAT
TTGTTTACTCTTACCAACCTCGGCGCTCCAGCCGCATTCAAGTATTTTGACACAACGATAGATCGCAA
CGATACACTTCTACCAAGGAGGTGCTAGACGCGCACTGATTCACCAATCCATCACGGGATTATATGAA
ACTCGGATAGATTTGTACAGCTTGGGGGTGACGATCCCCAAGAAGAAGAGGAAAGTCTCGAGCGAC
TACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATGACAAGTGA

【 0 1 4 6 】

配列番号 275 - MSP469 : CMV - T7 - ヒト SpCas9 - VQR (D1135V、R1335Q、T1337R) - NLS - 3xFLAG

ヒトコドン最適化化膿性連鎖球菌 (S. pyogenes) Cas9、通常フォント、改変コドン、小文字、NLS、二重下線、3xFLAGタグ、太字：

【化 8】

ATGGATAAAAAGTATTCTATTGGTTTAGACATCGGCACTAATTCGGTTGGATGGGCTGTCATAACCGAT
 GAATACAAAGTACCTTCAAAGAAATTTAAGGTGTTGGGGAACACAGACCGTCATTTCGATTAAAAAGAAAT
 CTTATCGGTGCCCTCCTATTTCGATAGTGGCGAAACGGCAGAGGCGACTCGCCTGAAACGAACCGCTCGG
 AGAAGGTATACACGTGCAAGAACCGAATATGTTACTTACAAGAAATTTTAGCAATGAGATGGCCAAA
 GTTGACGATTCTTTCTTTACCGTTTGGAGAGTCTTCTTGTGCGAAGAGGACAAGAAACATGAACGG
 CACCCCATCTTTGGAACATAGTAGATGAGGTGGCATATCATGAAAAGTACCCAACGATTTATCACCTC
 AGAAAAAGCTAGTTGACTCAACTGATAAAGCGGACCTGAGGTTAATCTACTTGGCTCTTGCCCATATG
 ATAAAGTTCGGTGGGCACTTTCTCATTGAGGGTGATCTAAATCCGGACAACTCGGATGTCGACAACTG
 TTCATCCAGTTAGTACAAACCTATAATCAGTTGTTTGAAGAGAACCTATAAATGCAAGTGGCGTGGAT
 GCGAAGGCTATTCTTAGCGCCCGCCTCTCTAAATCCCGACGGCTAGAAAACCTGATCGCACAAATTACCC
 GGAGAGAAGAAAAATGGGTTGTTTCGGTAACCTTATAGCGCTCTCCTAGGCTTGACACCAAAATTTAAG
 TCGAAGTTCGACTTAGCTGAAGATGCCAAATTGCAGCTTAGTAAGGACACGTACGATGACGATCTCGAC
 AATCTACTGGCACAAATTGGAGATCAGTATGCGGACTTATTTTTGGCTGCCAAAACCTTAGCGATGCA
 ATCCTCCTATCTGACATACTGAGAGTTAATACTGAGATTACCAAGGCGCGTATCCGCTTCAATGATC
 AAAAGGTACGATGAACATCACCAAGACTTGACACTTCTCAAGGCCCTAGTCCGTGAGCAACTGCCTGAG
 AAATATAAGGAAATATTCTTTGATCAGTCGAAAAACGGGTACGCAGGTATATTGACGGCGGAGCGAGT
 CAAGAGGAATTCTACAAGTTTATCAAACCCATATTAGAGAAGATGGATGGGACGGAAGAGTTGCTTGTA
 AAACCTCAATCGGAAGATCTACTGCGAAAGCAGCGGACTTTGACAAACGGTAGCATTCCACATCAAATC
 CACTTAGGCGAATTGCATGCTATACTTAGAAGGCAGGAGGATTTTTATCCGTTCTCAAAGACAAATCGT
 GAAAAGATTGAGAAAAATCCTAACCTTTTCGCATACCTTACTATGTGGGACCCCTGGCCCGAGGGAACCTC
 CGGTTTCGCATGGATGACAAGAAAGTCCGAAGAAACGATTACTCCATGGAATTTTGAGGAAGTTGTCGAT
 AAAGGTGCGTCAGCTCAATCGTTCATCGAGAGGATGACCAACTTTGACAAGAATTTACCGAACGAAAAA
 GTATTGCCTAAGCACAGTTTACTTTACGAGTATTTTACAGTGTACAATGAACTCACGAAAGTTAAGTAT
 GTCAGTGGGGCATGCGTAAACCCGCTTTCTAAGCGGAGAACAGAAAGCAATAGTAGATCTGTTA
 TTCAAGACCAACCGCAAGTGACAGTTAAGCAATTGAAAGAGGACTACTTTAAGAAAATTGAATGCTTC
 GATTCTGTGAGATCTCCGGGGTAGAAGATCGATTAAATGCGTCACTTGGTACGTATCATGACCTCCTA
 AAGATAATTAAAGATAAGGACTTCTTGGATAACGAAGAGAATGAAGATATCTTAGAAGATATAGTGTG
 ACTCTTACCCTCTTTGAAGATCGGGAATGATTGAGGAAAGACTAAAAACATACGCTCACCTGCTTCGAC
 GATAAGGTTATGAAACAGTTAAAGAGGCGTCGCTATACGGGCTGGGACGATTGTGCGCGAACTTATC
 AACGGGATAAGAGACAAGCAAAGTGTTAAACTATTCTCGATTTTCTAAAGAGCGACGGCTTCGCCAAT
 AGGAACCTTTATGCAGCTGATCCATGATGACTCTTTAACCTTCAAAGAGGATATACAAAAGGCACAGGTT
 TCCGGACAAGGGGACTCATTGCACGAACATATTGCGAATCTTGCTGGTTGCCAGCCATCAAAAAGGGC
 ATACTCCAGACAGTCAAAGTAGTGGATGAGCTAGTTAAGGTCATGGGACGTCACAAACCGGAAAAACAT
 GTAATCGAGATGGCACGCGAAAAATCAAACGACTCAGAAGGGGCAAAAAACAGTCGAGAGCGGATGAAG
 AGAATAGAAGAGGGTATTAAAGAACTGGGCAGCCAGATCTTAAAGGAGCATCCTGTGGAATAACCCAA
 TTGCAGAACGAGAACTTTACCTCTATTACCTACAAAATGGAAGGGACATGTATGTTGATCAGGAAGT
 GACATAAACCGTTTATCTGATTACGACGTCGATCACATTTGACCCCAATCCTTTTGAAGGACGATTCA
 ATCGACAATAAAGTGCTTACACGCTCGGATAAGAACCGGAGGAAAAAGTGACAATGTTTCAAGCGAGGAA
 GTCGTAAAGAAAAATGAAGAATATTGGCGGACGCTCCTAAATGCGAACTGATAACGCAAAGAAAGTTC
 GATAACTTAACTAAAGCTGAGAGGGGTGGCTTGTCTGAAGTGTGACAAGGCCGATTTATTAAACGTCAG
 CTCGTGGAACCCGCCAAATCACAAAGCATGTTGCACAGATACTAGATTCCCGAATGAATACGAAATAC
 GACGAGAACGATAAGCTGATTCGGGAAGTCAAAGTAATCACTTTAAAGTCAAATTTGGTGTGCGGACTTC
 AGAAAGGATTTTCAATTCTATAAAGTTAGGGAGATAAATACTACCACCATGCGCACGACGCTTATCTT
 AATGCCGTGCTAGGGACCGCACTCATTAAGAAATACCCGAAGCTAGAAAGTGAGTTTGTGTATGGTGAT
 TACAAGGTTTATGACGTCGGTAAGATGATCGCGAAAAGCGAACAGGAGATAGGCAAGGCTACAGCCAAA
 TACTCTTTTATTCTAACATTATGAATTTCTTTAAGACGGAATCACTCTGGCAAACGGAGAGATACGC
 AAACGACCTTTTAAATTGAAACCAATGGGGAGACAGGTGAAATCGTATGGGATAAGGGCCGGGACTTCGCG
 ACGGTGAGAAAAGTTTGTCCATGCCCAAGTCAACATAGTAAAGAAAAGTGAAGTGCAGACCGGAGGG
 TTTTCAAAGGAATCGATTCTTCCAAAAGGAATAGTGATAAGCTCATCGCTCGTAAAAAGGACTGGGAC
 CCGAAAAAGTACGGTGGCTTCgtgAGCCCTACAGTTGCCTATTCTGTCTAGTAGTGGCAAAAGTTGAG
 AAGGGAAAAATCCAAGAACTGAAGTCAGTCAAAGAATTATTGGGGATAACGATTATGGAGCGCTCGTCT
 TTTGAAAAGAACCCCATCGACTTCCTTGAGGCGAAAGGTTACAAGGAAGTAAAAAAGGATCTCATAATT
 AAACCTACCAAAGTATAGTCTGTTTGAAGTGAAGAAATGGCCGAAAACGGATGTTGGCTAGCGCCGGAGAG
 CTTCAAAGGGGAACGAACCTCGCACTACCGTCTAAATACGTGAATTTCTGTATTTAGCGTCCCATTAC
 GAGAAGTTGAAAAGGTTTCACTGAAAGATAACGAAGCAAGCAACTTTTGTGAGCAGCACAAACATTAT
 CTCGACGAAATCATAGAGCAAATTTTCGGAATTCAGTAAGAGAGTCATCCTAGCTGATGCCAATCTGGAC
 AAAGTATTAAGCGCATACAACAAGCACAGGGATAAACCATACGTGAGCAGGCGGAAAAATATTATCCAT
 TTGTTTACTCTTACCAACCTCGGCGCTCCAGCCGCAATTCAAGTATTTTGACACAACGATAGATCGCAAA
 cagTACagTCTACCAAGGAGGTGCTAGACGCGACACTGATTCACCAATCCATCACGGGATTATATGAA
 ACTCGGATAGATTGTGTCAGCTTGGGGGTGACGGATCCCCCAAGAAGAGAGGAAGTCTCGAGCGAC
 TACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATGACAAGTGA

10

20

30

40

【0147】

配列番号 276 - MSP2440 : CMV - T7 - ヒト SpCas9 - VQR - HF
 1 (N497A、R661A、Q695A、Q926A、D1135V、R1335Q、

50

T 1 3 3 7 R) - N L S - 3 x F L A G

ヒトコドン最適化化膿性連鎖球菌 (*S . p y o g e n e s*) C a s 9、通常フォント、改変コドン、小文字、N L S、二重下線、3 x F L A G タグ、太字：

【化9】

ATGGATAAAAAGTATTCTATTGGTTTAGACATCGGCCTAATTCCGTTGGATGGGCTGTCATAACCGAT
GAATACAAAGTACCTTCAAAGAAATTTAAGGTGTTGGGGAACACAGACCGTCAATTCGATTAAAAAGAAT
CTTATCGGTGCCCTCCTATTGATAGTGGCGAAACGSCAGAGGCGACTCGCCTGAAACGAACCGCTCGG
AGAAGGTATACACGTCGCAAGAACCGAATATGTTACTTACAAGAAATTTTTAGCAATGAGATGGCCAAA
GTTGACGATTCTTTCTTCCACCGTTTGGAAAGAGTCCCTTCTTGTGCAAGAGGACAAGAAACATGAACGG
CACCCCATCTTTGGAAACATAGTAGATGAGGTGGCATATCATGAAAAGTACCCAACGATTTATCACCTC
AGAAAAAGCTAGTTGACTCAACTGATAAAGCGGACCTGAGGTAACTTACTTGGCTCTTGCCCATATG
ATAAAGTTCCGTGGGCACTTTCTCATTGAGGGTGATCTAAATCCGGACAACCTCGGATGTCGACAAATG
TTCATCCAGTTAGTACAAACCTATAATCAGTTGTTTGAAGAGAACCCTATAAAATGCAAGTGGCGTGGAT
GCGAAGGCTATTCTTAGCGCCCGCTCTCTAAATCCCGACGGCTAGAAAACCTGATCGCACAAATTACCC
GGAGAGAAGAAAAATGGGTGTTTCGGTAACCTTATAGCGCTCTCACTAGGCCTGACACCAAATTTTAAG
TCGAACCTCGACTTAGCTGAAGATGCCAAATTGCAGCTTAGTAAGGACACGTACGATGACGATCTCGAC
AATCTACTGGCACAAATTGGAGATCAGTATGCGGACTTATTTTTGGCTGCCAAAACCTTAGCGATGCA
ATCCTCCTATCTGACATACTGAGAGTTAATACTGAGATTACCAAGGCGCCGTTATCCGCTTCAATGATC
AAAAGGTACGATGAACATCACCAAGACTTGACACTTCTCAAGGCCCTAGTCCGTCAGCAACTGCCTGAG
AAATATAAGGAAATATTTCTTGATCAGTCGAAAAACGGGTACGCAGGTTATATTGACGGCGGAGCGAGT
CAAGAGGAATTTCTACAAGTTTATCAAACCCATATTAGAGAAGATGGATGGGACGGAAGAGATTGCTTGTA
AAACTCAATCGCGAAGATCTACTGCGAAAGCAGCGGACTTTTCGACAACGGTAGCATTTCCACATCAAATC
CACTTAGGCGAATTGCATGCTATACTTAGAAGGCAGGAGGATTTTTATCCGTTCTCAAAGACAATCGT
GAAAAGATTGAGAAAATCCTAACCTTTTCGCATACCTTACTATGTGGGACCCCTGGCCCGAGGGAACCTCT
CGGTTTCGCATGGATGACAAGAAAGTCCGAAGAACGATTACTCCCTGGAATTTTGAGGAAGTTGTCGAT
AAAGGTGCGTCAGCTCAATCGTTCATCGAGAGGATGACCgcccTTTGACAAGAAATTTACCGAACGAAAA
GTATTGCCTAAGCACAGTTTACTTTACGAGTATTTACAGTGTACAATGAACCTCACGAAAGTTAAGTAT
GTCACTGAGGGCATGCGTAAACCCGCTTTCTAAGCGGAGAACAGAAGAAAGCAATAGTAGATCTGTTA
TTCAAGACCAACCGCAAAGTGACAGTTAAGCAATTGAAAGAGGACTACTTTAAGAAAATTGAATGCTTC
GATTCTGTCGAGATCTCCGGGTTAGAAGATCGATTAAATGCGTCACTTGGTACGTATCATGACCTCCTA
AAGATAATTAAAGATAAAGGACTTCTGGATAACGAAAGAGAATGAAGATATCTTAGAAGATATAGTGTG
ACTCTTACCCTCTTTGAAGATCGGGAAATGATTGAGGAAAGACTAAAAACATACGCTCACCTGTTTCGAC
GATAAGGTTATGAAACAGTTAAAGAGGCGTCGCTATACGGGCTGGGGAgcccTTGTCGCGGAAACTTATC
AACGGGATAAGAGACAAGCAAAGTGGTAAACTATTCTCGATTTTCTAAAGAGCGACGGCTTCGCCAAT
AGGAACTTTATGgcccCTGATCCATGATGACTCTTTAACCTTCAAAGAGGATATACAAAAGGCACAGGTT
TCCGGACAAGGGGACTCAATGCACGAACATATTGCGAATCTTGCTGGTTTCGCCAGCCATCAAAAAGGGC
ATACTCCAGACAGTCAAAGTAGTGGATGAGCTAGTTAAGGTGATGGGACGTACAAAACCGGAAAACATTT
GTAATCGAGATGGCAGCGGAAAATCAAACGACTCAGAAGGGGCAAAAAACAGTTCGAGAGCGGATGAAG
AGAATAGAAGAGGGTATTAAAGAACTGGGCGAGCATCTTAAAGGAGCATCTGTGGAAAAATACCCAA
TTGCAGAAATGAGAACTTTTACCTCTATTACCTACAAAATGGAAGGGACATGTATGTTGATCAGGAATG
GACATAAACCGTTTTATCTGATTACGACGTCGATCACATTGTACCCCAATCCTTTTTGAAGGACGATTCA
ATCGACAATAAAGTGCTTACACGCTCGGATAAGAACCGAGGGAAGTGAATGTTCCAAGCGAGGAA
GTCGTAAAGAAAATGAAGAACTATTGGCGGCAGCTCCTAAATGCGAACTGATAACGCAAAGAAAGTTC
GATAACTTAACTAAAGCTGAGAGGGGTGGCTTGTCTGAACCTTGACAAGGCCGGATTTATTAAACGTCAG
CTCGTGGAACCCGCGcccATCACAAGCATGTTGCGCAGATACTAGATTCCCGAATGAATACGAAATAC
GACGAGAACGATAAGCTGATTCGGGAAGTCAAAGTAATCACTTTAAAGTCAAATTTGGTGTGCGACTTC
AGAAAGGATTTTCAATTTCTATAAAGTTAGGGAGATAAACTACCACCATGCGCAGCAGCCTTATCTT
AATGCCGTCGTAGGGACCGCACTATTAAAGAAATCCCGAAGCTAGAAAGTGAGTTTGTGTATGGTGAT
TACAAAGTTTATGACGTCGTAAGATGATCGCGAAAGCGAACAGGAGATAGGCAAGGCTACAGGCCAAA
TACTTCTTTTATTCTAACATTATGAATTTCTTTAAGACGGAAATCACTCTGGCAAACGGAGAGATACGC
AAACGACCTTTAATTGAAACCAATGGGGAGACAGGTGAAATCGTATGGGATAAGGGCCGGGACTTCGCG
ACGGTGAGAAAAGTTTTGTCCATGCCCCAAGTCAACATAGTAAAGAAAACCTGAGGTGCAGACCGGAGGG
TTTTCAAAGGAATCGATTCTTCCAAAAGGAATAGTGATAAGCTCATCGCTCGTAAAAAGGACTGGGAC
CCGAAAAAGTACGGTGGCTTCgtgAGCCCTACAGTTGCCTATTCTGTCTAGTAGTGCCAAAAGTTGAG
AAGGGAAAAATCCAAGAACTGAAGTCAGTCAAAGAAATTATTGGGGATAACGATTATGGAGCGCTCGTCT
TTTGAAAAGAACCCCATCGACTTCCTTGAGGCGAAAGGTTACAAGGAAGTAAAAAGGATCTCATAATT
AAACTACCAAGTATAGTCTGTTTGAGTTAGAAAATGGCCGAAAACGGATGTTGGCTAGCGCCGAGAG
CTTCAAAGGGGAACGAACTCGCACTACCGTCTAAATACGTGAATTTCTGTATTTAGCGTCCCATTAC
GAGAAGTTGAAAGGTTACCTGAAGATAACGAACAGAAGCAACTTTTTGTTGAGCAGCACAAACATTAT
CTCGACGAAATCATAGAGCAAATTTCCGAATTGAGTAAGAGAGTCATCCTAGCTGATGCCAATCTGGAC
AAAGTATTAAGCGCATACAACAAGCACAGGGATAAACCCTACGTGAGCAGGCGGAAAATATTATCCAT
TTGTTTACTCTTACCAACCTCGGCGCTCCAGCCGCTTCAAGTATTTTGACACAACGATAGATCGCAAA
cagTACagaTCTACCAAGGAGGTGCTAGACGCGACACTGATTACCAATCCATCACGGGATTATATGAA
ACTCGGATAGATTTGTACAGCTTGGGGGTGACGGATCCCCAAGAAGAAGAGGAAAGTCTCGAGCGAC
TACAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATGACAAGTGA

10

20

30

40

50

【 0 1 4 8 】

配列番号 2 7 7 - B P K 2 7 9 7 : C M V - T 7 - ヒト S p C a s 9 - V R Q R (D
1 1 3 5 V、G 1 2 1 8 R、R 1 3 3 5 Q、T 1 3 3 7 R) - N L S - 3 x F L A G
ヒトコドン最適化化膿性連鎖球菌 (S . p y o g e n e s) C a s 9、通常フォント、改
変コドン、小文字、N L S、二重下線、3 x F L A G タグ、太字：

【化 10】

ATGGATAAAAAGTATTCTATTGGTTTAGACATCGGCCTAATTCGGTTGGATGGGCTGTCATAACCGAT
 GAATACAAAGTACCTTCAAAGAAATTTAAGGTGTTGGGGAACACAGACCGTCATTTCGATTAAAAAGAAT
 CTTATCGGTGCCCTCCTATTTCGATAGTGGCGAAACGGCAGAGCGGACTCGCCTGAAACGAACCGCTCGG
 AGAAGGTATACACGTGCGCAAGAACCGAATATGTTACTTACAAGAAATTTTAGCAATGAGATGGCCAAA
 GTTGACGATTCTTTCTTTCCCGTTTGGGAAGAGTCTTCTTGTGCGAAGAGGACAAGAAACATGAACGG
 CACCCCATCTTTGGAAACATAGTAGATGAGGTGGCATATCATGAAAAGTACCCAACGATTTATCACCTC
 AGAAAAAGCTAGTTGACTCAACTGATAAAGCGGACCTGAGGTAACTTACTTGGCTCTTGCCCATATG
 ATAAAGTTCCGTGGGCACCTTTCTCATTGAGGGTGATCTAAATCCGGACAACCTCGGATGTCGACAAACTG
 TTCATCCAGTTAGTACAAACCTATAATCAGTTGTTTGAAGAGAACCCTATAAATGCAAGTGGCGTGGAT
 GCGAAGGCTATTCTTAGCGCCCGCTCTCTAAATCCCGACGGCTAGAAAACCTGATCGCACAAATTACCC
 GGAGAGAAGAAAAATGGGTTGTTTCGGTAACCTTATAGCGCTCTCACTAGGCCTGACACCAAATTTTAAG
 TCGAAGTTTCGACTTAGCTGAAGATGCCAAATTGCAGCTTAGTAAGGACACGTACGATGACGATCTCGAC
 AATCTACTGGGCACAAATTGGAGATCAGTATGCGGACTTATTTTGGCTGCCAAAAACCTTAGCGATGCA
 ATCCTCCTATCTGACATACTGAGAGTTAATACTGAGATTACCAAGGCGCGTTATCCGCTTCAATGATC
 AAAAGGTACGATGAACATCACCAGACTTGACACTTCTCAAGGCCCTAGTCCGTGACCAACTGCCTGAG
 AAATATAAGGAAATATTCTTTGATCAGTCGAAAAACGGGTACGCAGGTTATATTTGACGGCGGAGCGAGT
 CAAGAGGAATTCTACAAGTTTATCAAACCCATATTAGAGAAGATGGATGGGACGGAAGAGTTGCTTGTA
 AAATCAATCGCGAAGATCTACTGCGAAAGCAGCGGACTTTCGACAACGGTAGCATTCCACATCAAATC
 CACTTAGGCGAATTGCATGCTATACTTAGAAGGCAGGAGGATTTTTATCCGTTCTCAAAGACAATCGT
 GAAAAGATTGAGAAAATCCTAACCTTTGCGATACCTTACTATGTGGGACCCCTGGCCGAGGGAACTCT
 CGGTTGCGATGGATGACAAGAAAGTCCGAAGAAACGATTACTTCCATGGAATTTTGGAGAAAGTTGTCGAT
 AAAGGTGCGTCAGCTCAATCGTTTCATCGAGAGGATGACCAACTTTGACAAGAATTTACCGAACGAAAAA
 GTATTGCCTAAGCACAGTTTACTTTACGAGTATTTACAGTGTAACAATGAACTCACGAAAGTTAAGTAT
 GTCAGTGGGGCATGCGTAAACCCGCTTTCTAAGCGGAGAACAGAAGAAAGCAATAGTAGATCTGTTA
 TTCAAGACCAACCGCAAAGTGACAGTTAAGCAATTGAAAGAGGACTACTTTAAGAAAAATTGAATGCTTC
 GATTCTGTGCGAGATCTCCGGGTAGAAGATCGATTTAATGCGTCACTTGGTACGTATCATGACCTCCTA
 AAGATAATTAAAGATAAGGACTTCCTGGATAACGAAGAGAATGAAGATATCTTAGAAGATATAGTGTTG
 ACTCTTACCCTCTTTGAAGATCGGGAAATGATTGAGGAAAGACTAAAAACATACGCTCACCTGTTTCGAC
 GATAAGGTTATGAAACAGTTAAAGAGGGCTCGCTATACGGGCTGGGGACGATTGTCGCGGAAACTTATC
 AACGGGATAAGAGACAAGCAAGTGGTAAAGCTATCTCGATTTTCTAAAGAGCGACGGCTTCGCCAAT
 AGGAACCTTATGCGAGTATCCATGATGACTCTTTAACCTTCAAAGAGGATATACAAAAGGCCACAGGTT
 TCCGGACAAGGGGACTCATTGACAGCAACATATTGCGAATCTTGCTGGTTTCGCCAGCCATCAAAAAGGGC
 ATACTCCAGACAGTCAAAGTAGTGGATGAGCTAGTTAAGGTGATGGGACGTACAAAACCGGAAAAACATT
 GTAATCGAGATGGCAGCGCAAAATCAAACGACTCAGAAGGGGCAAAAAACAGTCGAGAGCGGATGAAG
 AGAATAGAAGAGGGTATTAAAGAACTGGGCAGCCAGATCTTAAAGGAGCATCCTGTGAAAAATACCCAA
 TTGCAGAACGAGAACTTTACCTCTATTACCTACAAAATGGAAGGGACATGTATGTTGATCAGGAACTG
 GACATAAACCGTTTATCTGATTACGACGTGATCATTGTACCCCAATCCTTTTTGAAGGACGATTCA
 ATCGACAATAAAGTGCTTACACGCTCGGATAAGAACCGGAGGAAAAAGTGACAATGTTCCAAGCGAGGAA
 GTCGTAAAGAAAAATGAAGAACTATTGGCGGCGAGTCTTAAATGCGAAAACGATAACGCAAAAGAAAGTTT
 GATAACTTAACTAAAGCTGAGAGGGGTGGCTTGTCTGAACTTGACAAGGCCGATTTATTAAACGTCAG
 CTCGTGGAACCCGCCAAATCACAAAGCATGTTGCACAGATACTAGATTCCCAGATGAATACGAAATAC
 GACGAGAACGATAAGCTGATTTCGGGAAGTCAAAGTAATCACTTTAAAGTCAAATTTGGTGTGCGACTTC
 AGAAAGGATTTTCAATTCTATAAAGTTAGGGAGATAAATACTACCACCATGCGCAGCAGCTTATCTT
 AATGCCGTGCTAGGGACCGCACTCATTAAGAAATACCCGAAGCTAGAAAGTGAGTTTGTGTATGGTGAT
 TACAAAGTTTATGACGTCCGTAAGATGATCGCGAAAAGCGAACAGGAGATAGGCAAGGCTACAGCCAAA
 TACTTCTTTTATTCTAACATTATGAATTTCTTTAAGACGGAAATCACTCTGGCAAACGGAGAGATACGC
 AAACGACCTTTAATTGAAACCAATGGGGAGACAGGTGAAATCGTATGGGATAAGGGCCGGGACTTCGCG
 ACGGTGAGAAAGTTTTGTCCATGCCCAAGTCAACATAGTAAAGAAAACTGAGGTGACAGACCGGAGGG
 TTTTCAAAGGAATCGATTCTTCCAAAAAGGAATAGTGATAAGCTCATCGCTCGTAAAGGAGGACTGGGAC
 CCGAAAAAGTACGGTGGCTTCgtgAGCCCTACAGTTGCCTATTCTGTCTAGTAGTGCAAAAGTTGAG
 AAGGAAAAATCCAAGAACTGAAGTCAGTCAAAGAATTATTTGGGATAACGATTATGGAGCGCTCGTCT
 TTTGAAAAGAACCCCATCGACTTCCTTGAGGCGAAAGGTTACAAGGAAGTAAAAAGGATCTCATAATT
 AAATACCAAAGTATAGTCTGTTGAGTTAGAAAATGGCCGAAAACGGATGTTGGCTAGCGCCGagAGAG
 CTTCAAAAGGGGAACGAACTCGCACTACCGTCTAAATACGTGAATTTCTGTATTTAGCGTCCCATTAC
 GAGAAGTTGAAAGGTTTCACTGAAGATAACGAACAGAAGCAACTTTTGTGAGCAGCACAAACATTAT
 CTCGACGAAATCATAGACAAATTCGGAATTAGTAAGAGAGTCATCCTAGCTGATGCCAATCTGGAC
 AAAGTATAAGCGCATACAACAGCACAGGAGATAAACCCATACGTGAGCAGGCGGAAAAATATTATCCAT
 TTGTTTACTCTTACCAACCTCGGCGCTCCAGCCGCATTCAAGTATTTTGACACAACGATAGATCGCAAA
 cagTACagTCTACCAAGGAGGTGCTAGACGCGACACTGATTCACCAATCCATCACGGGATTATATGAA
 ACTCGGATAGATTTGTACAGCTTGGGGGTGACGGATCCCCAAGAAGAAGAGGAAAGTCTCGAGCGAC
 TACAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATGACAAGTGA

10

20

30

40

【0149】

配列番号 278 - MSP2443:CMV-T7-ヒトSpCas9-VRQR-H
 F1(N497A、R661A、Q695A、Q926A、D1135V、G1218R

50

、 R 1 3 3 5 Q 、 T 1 3 3 7 R) - N L S - 3 x F L A G

ヒトコドン最適化膿性連鎖球菌 (*S . p y o g e n e s*) C a s 9 、 通常フォント、 改変コドン、 小文字、 N L S 、 二重下線、 3 x F L A G タグ、 太字 :

【化 1 1】

ATGGATAAAAAGTATTCTATTGGTTTAGACATCGGCACTAATTCCGTTGGATGGGCTGTCATAACCGAT
GAATACAAAGTACCTTCAAAGAAATTTAAGGTGTTGGGGAACACAGACCGTCATTCGATTAAAAAGAAT
CTTATCGGTGCCCTCCTATTGATAGTGGCGAAACGGCAGAGGCGACTCGCCTGAAACGAACCGCTCGG
AGAAGGTATACACGTCGCAAGAACCGAATATGTTACTTACAAGAAATTTTTAGCAATGAGATGGCCAAA
GTTGACGATTCTTTCTTTACCGTTTGGGAAGAGTCCCTTCTTGTGCGAAGAGGACAAGAAACATGAACGG
CACCCCATCTTTGGAAACATAGTAGATGAGGTGGCATATCATGAAAAGTACCCAACGATTTATCACCTC
AGAAAAAGCTAGTTGACTCAACTGATAAAGCGGACCTGAGGTAACTTACTTTGGCTCTTGCCCATATG
ATAAAGTTCCGTGGGCACTTTCTCATTGAGGGTGATCTAAATCCGGACAACCTCGGATGTCGACAAACTG
TTCATCCAGTTAGTACAAACCTATAATCAGTTGTTTGAAGAGAACCCTATAAATGCAAGTGGCGTGGAT
GCGAAGGCTATTCTTAGCGCCCGCTCTCTAAATCCCGACGGCTAGAAAACCTGATCGCACAAATTACCC
GGAGAGAAGAAAAATGGGTGTTTCGGTAACCTTATAGCGCTCTCACTAGGCCCTGACACCAAATTTTAAG
TCGAACCTTCGACTTAGCTGAAGATGCCAAATTGCAGCTTAGTAAGGACACGTACGATGACGATCTCGAC
AATCTACTGGCACAATTTGGAGATCAGTATGCGGACTTATTTTTGGCTGCCAAAACCTTAGCGATGCA
ATCCTCCTATCTGACATACTGAGAGTTAATACTGAGATTACCAAGGCGCCGTTATCCGCTTCAATGATC
AAAAGGTACGATGAACATCACCAAGACTTGACACTTCTCAAGGCCCTAGTCCGTCAGCAACTGCCTGAG
AAATATAAGGAAATATTCTTTGATCAGTCGAAAAACGGGTACGCAGGTTATATTGACGGCGGAGCGAGT
CAAGAGGAATTTCTACAAGTTTATCAAACCCATATTAGAGAAGATGGATGGGACGGAAGAGTTGCTTGTGA
AAACTCAATCGCGAAGATCTACTGCGAAAGCAGCGGACTTTTCGACAACGGTAGCATTCACATCAAATC
CACTTAGGCGAATTGCATGCTATACTTAGAAGGCAGGAGGATTTTTATCCGTTCTCAAAGACAATCGT
GAAAAGATTGAGAAAATCCTAACCTTTTCGCATACCTTACTATGTGGGACCCCTGGCCCGAGGGAACCTCT
CGGTTTCGCATGGATGACAAGAAAGTCCGAAGAAACGATTACTCCCTGGAATTTTGAGGAAGTTGTCGAT
AAAGGTGCGTCAGCTCAATCGTTTCATCGAGAGGATGACCgccTTTGACAAGAAATTACCGAACGAAAAA
GTATTGCCTAAGCACAGTTTACTTTACGAGTATTTTACAGTGTACAATGAACCTCACGAAAGTTAAGTAT
GTCACTGAGGGCATGCGTAAACCCCGCTTTCTAAGCGGAGAACAGAAGAAAGCAATAGTAGATCTGTTA
TTCAAGACCAACCGCAAAGTGACAGTTAAGCAATTGAAAGAGGACTACTTTAAGAAAATTGAATGCTTC
GATTCTGTGAGATCTCCGGGTAGAAGATCGATTAAATGCGTCACTTGGTACGTATCATGACTGCTCTTA
AAGATAATTAAAGATAAGGACTTCTGGATAACGGAAGAGAATGAAGATATCTTAGAAGATATAGTGTG
ACTCTTACCCTCTTTGAAGATCGGGAAATGATTGAGGAAAGACTAAAAACATACGCTCACCTGTTTCGAC
GATAAGGTTATGAAACAGTTAAAGAGGCGTCGCTATACGGGCTGGGGAgccTTGTCGCGGAAACTTATC
AACGGGATAAGAGACAAGCAAAGTGGTAAACTATTCTCGATTTTCTAAAGAGCGACGGCTTCGCCAAT
AGGAACCTTTATGgccCTGATCCATGATGACTCTTTAACCTTCAAAGAGGATATACAAAAGGCACAGGTT
TCCGGACAAGGGGACTCATTGCACGAACATATTTGCGAATCTTGCTGGTTTCGCCAGCCATCAAAAAGGGC
ATACTCCAGACAGTCAAAGTAGTGGATGAGCTAGTTAAGGTTCATGGGACGTCACAAAACCGGAAAACATT
GTAATCGAGATGGCAGCGCAAAATCAAACGACTCAGAAGGGGCAAAAAACAGTTCGAGAGCGGATGAAG
AGAATAGAAGAGGGTATTAAAGAACTGGGCGAGCAGATCTTAAAGGAGCATCCTGTGGAAAATACCCAA
TTGCAGAACGAGAACTTTACCTCTATTACCTACAAATGGAAGGGACATGTATGTTGATCAGGAAGT
GACATAAACCGTTTATCTGATTACGACGTCGATCACATTTGTACCCCAATCCTTTTGAAGGACGATTCA
ATCGACAATAAAGTGCTTACACGCTCGGATAAGAACCAGGGGAAAAGTGACAAATGTTCCAAGCGAGGAA
GTCGTAAAGAAAATGAAGAACTATTGGCGGCAGCTCTTAAATGCGAACTGATAACGCAAGAAAGTTTC
GATAACTTAATAAGCTGAGAGGGGTGGCTTGTCTGAACTTGACAAGGCCGGATTTATTAAACGTCAG
CTCGTGGAACCCCGgccATCACAAAGCATGTTGCGCAGATACTAGATTCCCGAATGAATACGAAATAC
GACGAGAACGATAAGCTGATTTCGGGAAGTCAAAGTAATCACTTTAAAGTCAAAATTTGGTGTGCGACTTC
AGAAAGGATTTTCAATTTCTATAAAGTTAGGGAGATAAATACTACCACCATGCGCAGCAGCTTATCTT
AATGCCGTCTGAGGACCGCACTCATTAAGAAATACCCGAAGCTAGAAAGTGAGTTTGTGTGATGGTAT
TACAAAGTTTATGACGTCCGTAAGATGATCGCGAAAAGCGAACAGGAGATAGGCAAGGCTACAGCCAAA
TACTTCTTTTATTCTAACATTATGAATTTCTTTAAGACGGAAATCACTCTGGCAAACGGAGAGATACGC
AAACGACCTTTAATTGAAACCAATGGGGAGACAGGTGAAATCGTATGGGATAAGGGCCGGGACTTCGCG
ACGGTGAGAAAAGTTTTGTCCATGCCCCAAGTCAACATAGTAAAGAAAACCTGAGGTGCAGACCGGAGGG
TTTTCAAAGGAATCGATTCTTCCAAAAGGAATAGTGATAAGCTCATCGCTCGTAAAAAGGACTGGGAC
CCGAAAAAGTACGGTGGCTTCgtgAGCCCTACAGTTGCCTATTCTGTCTAGTAGTGGCAAAAGTTGAG
AAGGGAAAAATCCAAGAACTGAAGTCAGTCAAAGAATTATTGGGGATAACGATTATGGAGCGCTCGTCT
TTTGAAAAGAACCCCATCGACTTCTTTGAGGCGAAAGGTTACAAGGAAGTAAAAAGGATCTCATAATT
AAACTACCAAGGTATGTTGTTGAGTTAGAAAATGGCCGAAAACGGATGTTGGCTAGCGCCAgAGAG
CTTCAAAGGGGAACGAACTCGCACTACCGTCTAAATACGTGAATTTCTGTATTATTAGCGTCCCATAC
GAGAAGTTGAAAGGTTACCTGAAGATAACGAACAGAAGCAACTTTTTGTTGAGCAGCACAAACATTAT
CTCGACGAAATCATAGAGCAAATTTTCGAATTTCAGTAAGAGAGTCATCCTAGCTGATGCCAATCTGGAC
AAAGTATTAAGCGCATACAACAAGCACAGGGATAAACCCTACGTGAGCAGGCGGAAAATATTATCCAT
TTGTTTACTCTTACCAACCTCGGCGCTCCAGCCGCATTCAGTATTTTACACACAGGATAGATCGCAA
cagTACAgTCTACCAAGGAGGTGCTAGACGCGCACTGATTACCAATCCATCACGGGATTATATGAA
ACTCGGATAGATTTGTCACAGCTTGGGGGTGACGGATCCCCAAGAAGAAGAGGAAAGTCTCGAGCGAC
TACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATGACAAGTGA

10

20

30

40

50

【 0 1 5 0 】

配列番号 279 - B P K 1 5 2 0 : U 6 - B s m B I カセット - S p - s g R N A
 U 6 プロモーター、通常フォント、B s m B I 部位、イタリック、化膿性連鎖球菌 (S .
 p y o g e n e s) s g R N A、小文字、U 6 ターミネーター、二重下線：
 【 化 1 2 】

TGTACAAAAAGCAGGCTTTAAAGGAACCAATTCAGTCGACTGGATCCGGTACCAAGGTCGGGCAGGAA
 GAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACAAGGCTGTTAGAGAGATAATTAGA
 ATTAATTTGACTGTAAACACAAAGATATTAGTACAAAATACGTGACGTAGAAAGTAATAATTTCTTGGG
 TAGTTTGCAGTTTTAAATTTATGTTTTAAATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTTC
 GATTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCGGAGACGATTAATGCGTCTCCgttttta
 gagctagaaatagcaagttaaataaggctagtcggttatcaacttgaaaaagtggcaccgagtcggtg
 cttttttttt

10

【 0 1 5 1 】

他の実施形態

本発明をその詳細な説明とともに記載した一方、上記の説明は、添付の特許請求の範囲
 の範囲により定義される本発明の範囲を説明するものであり、限定するものではないこと
 を理解すべきである。他の態様、利点、および改変は、以下の特許請求の範囲の範囲内
 である。

上記の開示に基づく発明の例として、以下のものが挙げられる。

20

[1] 以下の位置：L 1 6 9、Y 4 5 0、N 4 9 7、R 6 6 1、Q 6 9 5、Q 9 2 6、お
 よび／またはD 1 1 3 5 の 1、2、3、4、5、6、または7つ全てにおける突然変異を
 有し、好ましくは、以下の位置：L 1 6 9、Y 4 5 0、N 4 9 7、R 6 6 1、Q 6 9 5、
 Q 9 2 6、D 1 1 3 5 の 1、2、3、4、5、6、または7つにおける突然変異を有する
 配列番号 1 のアミノ酸配列と少なくとも 8 0 % 同一である配列ならびに任意選択的に核局
 在化配列、細胞浸透ペプチド配列、および／または親和性タグの 1 つ以上を含む単離化膿
 性連鎖球菌 (S t r e p t o c o c c u s p y o g e n e s) C a s 9 (S p C a s 9
) タンパク質。

[2] 以下：N 4 9 7、R 6 6 1、Q 6 9 5、およびQ 9 2 6 の 1、2、3、または4つ
 全てにおける突然変異、好ましくは、以下の突然変異：N 4 9 7 A、R 6 6 1 A、Q 6 9
 5 A、およびQ 9 2 6 A の 1、2、3、または4つ全てを含む、[1] に記載の単離タンパ
 ク質。

30

[3] Q 6 9 5 および／またはQ 9 2 6 の一方または両方ならびに任意選択的にL 1 6 9
 、Y 4 5 0、N 4 9 7、R 6 6 1、およびD 1 1 3 5 の 1、2、3、4、または5つ全て
 における突然変異、好ましくは、Y 4 5 0 A / Q 6 9 5 A、L 1 6 9 A / Q 6 9 5 A、Q
 6 9 5 A / Q 9 2 6 A、Q 6 9 5 A / D 1 1 3 5 E、Q 9 2 6 A / D 1 1 3 5 E、Y 4 5
 0 A / D 1 1 3 5 E、L 1 6 9 A / Y 4 5 0 A / Q 6 9 5 A、L 1 6 9 A / Q 6 9 5 A /
 Q 9 2 6 A、Y 4 5 0 A / Q 6 9 5 A / Q 9 2 6 A、R 6 6 1 A / Q 6 9 5 A / Q 9 2 6
 A、N 4 9 7 A / Q 6 9 5 A / Q 9 2 6 A、Y 4 5 0 A / Q 6 9 5 A / D 1 1 3 5 E、Y
 4 5 0 A / Q 9 2 6 A / D 1 1 3 5 E、Q 6 9 5 A / Q 9 2 6 A / D 1 1 3 5 E、L 1 6
 9 A / Y 4 5 0 A / Q 6 9 5 A / Q 9 2 6 A、L 1 6 9 A / R 6 6 1 A / Q 6 9 5 A / Q
 9 2 6 A、Y 4 5 0 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A、N 4 9 7 A / Q 6 9 5 A
 / Q 9 2 6 A / D 1 1 3 5 E、R 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A / D 1 1 3 5 E、ま
 たはY 4 5 0 A / Q 6 9 5 A / Q 9 2 6 A / D 1 1 3 5 Eを含む、[1] に記載の単離タン
 パク質。

40

[4] N 1 4 ; S 1 5 ; S 5 5 ; R 6 3 ; R 7 8 ; H 1 6 0 ; K 1 6 3 ; R 1 6 5 ; L 1
 6 9 ; R 4 0 3 ; N 4 0 7 ; Y 4 5 0 ; M 4 9 5 ; N 4 9 7 ; K 5 1 0 ; Y 5 1 5 ; W 6
 5 9 ; R 6 6 1 ; M 6 9 4 ; Q 6 9 5 ; H 6 9 8 ; A 7 2 8 ; S 7 3 0 ; K 7 7 5 ; S 7
 7 7 ; R 7 7 8 ; R 7 8 0 ; K 7 8 2 ; R 7 8 3 ; K 7 8 9 ; K 7 9 7 ; Q 8 0 5 ; N 8
 0 8 ; K 8 1 0 ; R 8 3 2 ; Q 8 4 4 ; S 8 4 5 ; K 8 4 8 ; S 8 5 1 ; K 8 5 5 ; R 8

50

5 9 ; K 8 6 2 ; K 8 9 0 ; Q 9 2 0 ; Q 9 2 6 ; K 9 6 1 ; S 9 6 4 ; K 9 6 8 ; K 9
 7 4 ; R 9 7 6 ; N 9 8 0 ; H 9 8 2 ; K 1 0 0 3 ; Y 1 0 1 3 ; K 1 0 1 4 ; V 1 0 1
 5 ; S 1 0 4 0 ; N 1 0 4 1 ; N 1 0 4 4 ; K 1 0 4 7 ; K 1 0 5 9 ; R 1 0 6 0 ; K 1
 1 0 7 ; E 1 1 0 8 ; S 1 1 0 9 ; K 1 1 1 3 ; R 1 1 1 4 ; S 1 1 1 6 ; K 1 1 1 8 ;
 R 1 1 2 2 ; K 1 1 2 3 ; K 1 1 2 4 ; D 1 1 3 5 ; S 1 1 3 6 ; K 1 1 5 3 ; K 1 1 5
 5 ; K 1 1 5 8 ; K 1 2 0 0 ; Q 1 2 2 1 ; H 1 2 4 1 ; Q 1 2 5 4 ; Q 1 2 5 6 ; K 1
 2 8 9 ; K 1 2 9 6 ; K 1 2 9 7 ; R 1 2 9 8 ; K 1 3 0 0 ; H 1 3 1 1 ; K 1 3 2 5 ;
 K 1 3 3 4 ; T 1 3 3 7 および / または S 1 2 1 6 における突然変異、好ましくは、N 4
 9 7 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A / K 8 1 0 A、N 4 9 7 A / R 6 6 1 A /
 Q 6 9 5 A / Q 9 2 6 A / K 8 4 8 A、N 4 9 7 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6
 A / K 8 5 5 A、N 4 9 7 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A / R 7 8 0 A、N 4
 9 7 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A / K 9 6 8 A、N 4 9 7 A / R 6 6 1 A /
 Q 6 9 5 A / Q 9 2 6 A / H 9 8 2 A、N 4 9 7 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6
 A / K 1 0 0 3 A、N 4 9 7 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A / K 1 0 1 4 A、
 N 4 9 7 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A / K 1 0 4 7 A、N 4 9 7 A / R 6 6
 1 A / Q 6 9 5 A / Q 9 2 6 A / R 1 0 6 0 A、N 4 9 7 A / R 6 6 1 A / Q 6 9 5 A /
 Q 9 2 6 A / K 8 1 0 A / K 9 6 8 A、N 4 9 7 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6
 A / K 8 1 0 A / K 8 4 8 A、N 4 9 7 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A / K 8
 1 0 A / K 1 0 0 3 A、N 4 9 7 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A / K 8 1 0 A
 / R 1 0 6 0 A、N 4 9 7 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A / K 8 4 8 A / K 1
 0 0 3 A、N 4 9 7 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A / K 8 4 8 A / R 1 0 6 0
 A、N 4 9 7 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A / K 8 5 5 A / K 1 0 0 3 A、N
 4 9 7 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A / K 8 5 5 A / R 1 0 6 0 A、N 4 9 7
 A / R 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A / K 9 6 8 A / K 1 0 0 3 A、N 4 9 7 A / R
 6 6 1 A / Q 6 9 5 A / Q 9 2 6 A / H 9 8 2 A / K 1 0 0 3 A、N 4 9 7 A / R 6 6 1
 A / Q 6 9 5 A / Q 9 2 6 A / H 9 8 2 A / R 1 0 6 0 A、N 4 9 7 A / R 6 6 1 A / Q
 6 9 5 A / Q 9 2 6 A / K 1 0 0 3 A / R 1 0 6 0 A、N 4 9 7 A / R 6 6 1 A / Q 6 9
 5 A / Q 9 2 6 A / K 8 1 0 A / K 1 0 0 3 A / R 1 0 6 0 A、N 4 9 7 A / R 6 6 1 A
 / Q 6 9 5 A / Q 9 2 6 A / K 8 4 8 A / K 1 0 0 3 A / R 1 0 6 0 A、Q 6 9 5 A / Q
 9 2 6 A / R 7 8 0 A、Q 6 9 5 A / Q 9 2 6 A / K 8 1 0 A、Q 6 9 5 A / Q 9 2 6 A
 / R 8 3 2 A、Q 6 9 5 A / Q 9 2 6 A / K 8 4 8 A、Q 6 9 5 A / Q 9 2 6 A / K 8 5
 5 A、Q 6 9 5 A / Q 9 2 6 A / K 9 6 8 A、Q 6 9 5 A / Q 9 2 6 A / R 9 7 6 A、Q
 6 9 5 A / Q 9 2 6 A / H 9 8 2 A、Q 6 9 5 A / Q 9 2 6 A / K 1 0 0 3 A、Q 6 9 5
 A / Q 9 2 6 A / K 1 0 1 4 A、Q 6 9 5 A / Q 9 2 6 A / K 1 0 4 7 A、Q 6 9 5 A /
 Q 9 2 6 A / R 1 0 6 0 A、Q 6 9 5 A / Q 9 2 6 A / K 8 4 8 A / K 9 6 8 A、Q 6 9
 5 A / Q 9 2 6 A / K 8 4 8 A / K 1 0 0 3 A、Q 6 9 5 A / Q 9 2 6 A / K 8 4 8 A /
 K 8 5 5 A、Q 6 9 5 A / Q 9 2 6 A / K 8 4 8 A / H 9 8 2 A、Q 6 9 5 A / Q 9 2 6
 A / K 1 0 0 3 A / R 1 0 6 0 A、Q 6 9 5 A / Q 9 2 6 A / R 8 3 2 A / R 1 0 6 0 A
 、Q 6 9 5 A / Q 9 2 6 A / K 9 6 8 A / K 1 0 0 3 A、Q 6 9 5 A / Q 9 2 6 A / K 9
 6 8 A / R 1 0 6 0 A、Q 6 9 5 A / Q 9 2 6 A / K 8 4 8 A / R 1 0 6 0 A、Q 6 9 5
 A / Q 9 2 6 A / K 8 5 5 A / H 9 8 2 A、Q 6 9 5 A / Q 9 2 6 A / K 8 5 5 A / K 1
 0 0 3 A、Q 6 9 5 A / Q 9 2 6 A / K 8 5 5 A / R 1 0 6 0 A、Q 6 9 5 A / Q 9 2 6
 A / H 9 8 2 A / K 1 0 0 3 A、Q 6 9 5 A / Q 9 2 6 A / H 9 8 2 A / R 1 0 6 0 A、
 Q 6 9 5 A / Q 9 2 6 A / K 1 0 0 3 A / R 1 0 6 0 A、Q 6 9 5 A / Q 9 2 6 A / K 8
 1 0 A / K 1 0 0 3 A / R 1 0 6 0 A、Q 6 9 5 A / Q 9 2 6 A / K 1 0 0 3 A / K 1 0
 4 7 A / R 1 0 6 0 A、Q 6 9 5 A / Q 9 2 6 A / K 9 6 8 A / K 1 0 0 3 A / R 1 0 6
 0 A、Q 6 9 5 A / Q 9 2 6 A / R 8 3 2 A / K 1 0 0 3 A / R 1 0 6 0 A、または Q 6
 9 5 A / Q 9 2 6 A / K 8 4 8 A / K 1 0 0 3 A / R 1 0 6 0 A をさらに含む、[1] に記
 載の単離タンパク質。
 [5] 以下の突然変異 : D 1 1 3 5 E ; D 1 1 3 5 V ; D 1 1 3 5 V / R 1 3 3 5 Q / T

10

20

30

40

50

1 3 3 7 R (V Q R バリアント) ; D 1 1 3 5 E / R 1 3 3 5 Q / T 1 3 3 7 R (E Q R バリアント) ; D 1 1 3 5 V / G 1 2 1 8 R / R 1 3 3 5 Q / T 1 3 3 7 R (V R Q R バリアント) ; または D 1 1 3 5 V / G 1 2 1 8 R / R 1 3 3 5 E / T 1 3 3 7 R (V R E R バリアント) の 1 つ以上をさらに含む、[1] に記載の単離タンパク質。

[6] D 1 0 、 E 7 6 2 、 D 8 3 9 、 H 9 8 3 、 または D 9 8 6 ; および H 8 4 0 または N 8 6 3 における突然変異からなる群から選択される、ヌクレアーゼ活性を減少させる 1 つ以上の突然変異をさらに含む、[1] に記載の単離タンパク質。

[7] ヌクレアーゼ活性を減少させる前記突然変異が、

(i) D 1 0 A または D 1 0 N 、 および

(i i) H 8 4 0 A 、 H 8 4 0 N 、 または H 8 4 0 Y

である、[6] に記載の単離タンパク質。

[8] 以下の位置 : Y 2 1 1 、 Y 2 1 2 、 W 2 2 9 、 Y 2 3 0 、 R 2 4 5 、 T 3 9 2 、 N 4 1 9 、 Y 6 5 1 、 R 6 5 4 の 1 、 2 、 3 、 4 、 5 、 6 つ以上における突然変異を有し、好ましくは、以下の位置 : Y 2 1 1 、 Y 2 1 2 、 W 2 2 9 、 Y 2 3 0 、 R 2 4 5 、 T 3 9 2 、 N 4 1 9 、 Y 6 5 1 、 R 6 5 4 の 1 、 2 、 3 、 4 、 または 5 、 6 つ以上における突然変異を有する配列番号 1 のアミノ酸配列と少なくとも 8 0 % 同一である配列ならびに任意選択的に核局在化配列、細胞浸透ペプチド配列、および / または親和性タグの 1 つ以上を含む単離黄色ブドウ球菌 (*Staphylococcus aureus*) Cas 9 (*SaCas9*) タンパク質。

[9] 以下の突然変異 : Y 2 1 1 A 、 Y 2 1 2 A 、 W 2 2 9 、 Y 2 3 0 A 、 R 2 4 5 A 、 T 3 9 2 A 、 N 4 1 9 A 、 Y 6 5 1 、 および / または R 6 5 4 A の 1 つ以上を含む、[8] に記載の単離タンパク質。

[1 0] N 4 1 9 および / または R 6 5 4 における突然変異ならびに任意選択的に追加の突然変異 Y 2 1 1 、 Y 2 1 2 、 W 2 2 9 、 Y 2 3 0 、 R 2 4 5 および T 3 9 2 の 1 、 2 、 3 、 4 つ以上、好ましくは、N 4 1 9 A / R 6 5 4 A 、 Y 2 1 1 A / R 6 5 4 A 、 Y 2 1 1 A / Y 2 1 2 A 、 Y 2 1 1 A / Y 2 3 0 A 、 Y 2 1 1 A / R 2 4 5 A 、 Y 2 1 2 A / Y 2 3 0 A 、 Y 2 1 2 A / R 2 4 5 A 、 Y 2 3 0 A / R 2 4 5 A 、 W 2 2 9 A / R 6 5 4 A 、 Y 2 1 1 A / Y 2 1 2 A / Y 2 3 0 A 、 Y 2 1 1 A / Y 2 1 2 A / R 2 4 5 A 、 Y 2 1 1 A / Y 2 1 2 A / Y 6 5 1 A 、 Y 2 1 1 A / Y 2 3 0 A / R 2 4 5 A 、 Y 2 1 1 A / Y 2 3 0 A / Y 6 5 1 A 、 Y 2 1 1 A / R 2 4 5 A / Y 6 5 1 A 、 Y 2 1 1 A / R 2 4 5 A / N 4 1 9 A 、 Y 2 1 1 A / N 4 1 9 A / R 6 5 4 A 、 Y 2 1 2 A / Y 2 3 0 A / R 2 4 5 A 、 Y 2 1 2 A / Y 2 3 0 A / Y 6 5 1 A 、 Y 2 1 2 A / R 2 4 5 A / Y 6 5 1 A 、 Y 2 3 0 A / R 2 4 5 A / Y 6 5 1 A 、 R 2 4 5 A / N 4 1 9 A / R 6 5 4 A 、 T 3 9 2 A / N 4 1 9 A / R 6 5 4 A 、 R 2 4 5 A / T 3 9 2 A / N 4 1 9 A / R 6 5 4 A 、 Y 2 1 1 A / R 2 4 5 A / N 4 1 9 A / R 6 5 4 A 、 W 2 2 9 A / R 2 4 5 A / N 4 1 9 A / R 6 5 4 A 、 Y 2 1 1 A / R 2 4 5 A / T 3 9 2 A / N 4 1 9 A / R 6 5 4 A 、 または Y 2 1 1 A / W 2 2 9 A / R 2 4 5 A / N 4 1 9 A / R 6 5 4 A を含む、[8] に記載の単離タンパク質。

[1 1] Y 2 1 1 ; Y 2 1 2 ; W 2 2 9 ; Y 2 3 0 ; R 2 4 5 ; T 3 9 2 ; N 4 1 9 ; L 4 4 6 ; Q 4 8 8 A ; N 4 9 2 A ; Q 4 9 5 A ; R 4 9 7 A ; N 4 9 8 A ; R 4 9 9 ; Q 5 0 0 ; K 5 1 8 ; K 5 2 3 ; K 5 2 5 ; H 5 5 7 ; R 5 6 1 ; K 5 7 2 ; R 6 3 4 ; Y 6 5 1 ; R 6 5 4 ; G 6 5 5 ; N 6 5 8 ; S 6 6 2 ; N 6 6 7 ; R 6 8 6 ; K 6 9 2 ; R 6 9 4 ; H 7 0 0 ; K 7 5 1 ; D 7 8 6 ; T 7 8 7 ; Y 7 8 9 ; T 8 8 2 ; K 8 8 6 ; N 8 8 8 ; 8 8 9 ; L 9 0 9 ; N 9 8 5 ; N 9 8 6 ; R 9 9 1 ; R 1 0 1 5 ; N 4 4 ; R 4 5 ; R 5 1 ; R 5 5 ; R 5 9 ; R 6 0 ; R 1 1 6 ; R 1 6 5 ; N 1 6 9 ; R 2 0 8 ; R 2 0 9 ; Y 2 1 1 ; T 2 3 8 ; Y 2 3 9 ; K 2 4 8 ; Y 2 5 6 ; R 3 1 4 ; N 3 9 4 ; Q 4 1 4 ; K 5 7 ; R 6 1 ; H 1 1 1 ; K 1 1 4 ; V 1 6 4 ; R 1 6 5 ; L 7 8 8 ; S 7 9 0 ; R 7 9 2 ; N 8 0 4 ; Y 8 6 8 ; K 8 7 0 ; K 8 7 8 ; K 8 7 9 ; K 8 8 1 ; Y 8 9 7 ; R 9 0 1 ; および / または K 9 0 6 における突然変異をさらに含む、[8] に記載の単離タンパク質。

10

20

30

40

50

[12] 以下の突然変異：E782K、K929R、N968K、またはR1015Hの1つ以上、具体的には、E782K/N968K/R1015H(KKHバリエーション)；E782K/K929R/R1015H(KRHバリエーション)；またはE782K/K929R/N968K/R1015H(KRKHバリエーション)をさらに含む、[8]に記載の単離タンパク質。

[13] D10、E477、D556、H701、またはD704；およびH557またはN580における突然変異からなる群から選択される、ヌクレアーゼ活性を減少させる1つ以上の突然変異をさらに含む、[8]に記載の単離タンパク質。

[14] 前記突然変異が、

(i) D10AもしくはD10N、および/または

(ii) H557A、H557N、もしくはH557Y、および/または

(iii) N580A、および/または

(iv) D556A

である、[13]に記載の単離タンパク質。

[15] 任意選択の介在リンカーにより異種機能ドメインに融合されている[1]～[14]のいずれか一項に記載の単離タンパク質を含む融合タンパク質であって、前記リンカーが前記融合タンパク質の活性を妨害しない、融合タンパク質。

[16] 前記異種機能ドメインが転写活性化ドメインである、[15]に記載の融合タンパク質。

[17] 前記転写活性化ドメインがVP64またはNF- κ B p65からのものである、[16]に記載の融合タンパク質。

[18] 前記異種機能ドメインが転写サイレンサーまたは転写抑制ドメインである、[15]に記載の融合タンパク質。

[19] 前記転写抑制ドメインが、クルッペル関連ボックス(KRAB)ドメイン、ERFリプレッサードメイン(ERD)、またはmSin3A相互作用ドメイン(SID)である、[18]に記載の融合タンパク質。

[20] 前記転写サイレンサーがヘテロクロマチンタンパク質1(HP1)、好ましくはHP1またはHP1である、[18]に記載の融合タンパク質。

[21] 前記異種機能ドメインが、DNAのメチル化状態を改変する酵素である、[15]に記載の融合タンパク質。

[22] 前記DNAのメチル化状態を改変する前記酵素がDNAメチルトランスフェラーゼ(DNMT)またはTETタンパク質である、[21]に記載の融合タンパク質。

[23] 前記TETタンパク質がTET1である、[22]に記載の融合タンパク質。

[24] 前記異種機能ドメインが、ヒストンサブユニットを改変する酵素である、[15]に記載の融合タンパク質。

[25] ヒストンサブユニットを改変する前記酵素が、ヒストンアセチルトランスフェラーゼ(HAT)、ヒストンデアセチラーゼ(HDAC)、ヒストンメチルトランスフェラーゼ(HMT)、またはヒストンデメチラーゼである、[15]に記載の融合タンパク質。

[26] 前記異種機能ドメインが生物学的係留物である、[15]に記載の融合タンパク質。

[27] 前記生物学的係留物がMS2、Csy4またはラムダNタンパク質である、[26]に記載の融合タンパク質。

[28] 前記異種機能ドメインがFokIである、[26]に記載の融合タンパク質。

[29] [1]～[14]のいずれか一項に記載のタンパク質をコードする単離核酸。

[30] 任意選択的に、[1]～[24]のいずれか一項に記載のタンパク質を発現させるための1つ以上の調節ドメインに作動可能に結合されている、[29]に記載の単離核酸を含むベクター。

[31] [29]に記載の核酸を含み、かつ任意選択的に[1]～[14]のいずれか一項に記載のタンパク質を発現する宿主細胞、好ましくは哺乳動物宿主細胞。

[32] 細胞のゲノムまたはエピゲノムを変更する方法であって、前記細胞中で、[1]～

10

20

30

40

50

[1 4]のいずれか一項に記載の単離タンパク質および前記細胞の前記ゲノムの選択部分に相補的な領域を有するガイドRNAを発現させるか、またはそれらと前記細胞を接触させることを含む方法。

[3 3] [1 5]に記載のタンパク質をコードする単離核酸。

[3 4] 任意選択的に、[1 5]に記載のタンパク質を発現させるための1つ以上の調節ドメインに作動可能に結合されている、[3 3]に記載の単離核酸を含むベクター。

[3 5] [3 3]に記載の核酸を含み、かつ任意選択的に[1 5]に記載のタンパク質を発現する宿主細胞、好ましくは哺乳動物宿主細胞。

[3 6] 細胞のゲノムまたはエピゲノムを変更する方法であって、前記細胞中で、[1]～[1 4]のいずれか一項に記載の単離タンパク質および前記細胞の前記ゲノムの選択部分に相補的な領域を有するガイドRNAを発現させるか、またはそれらと前記細胞を接触させることを含む方法。

10

[3 7] 細胞のゲノムまたはエピゲノムを変更する方法であって、前記細胞中で、[1 5]～[2 8]のいずれか一項に記載の単離融合タンパク質および前記細胞の前記ゲノムの選択部分に相補的な領域を有するガイドRNAを発現させるか、またはそれらと前記細胞を接触させることを含む方法。

[3 8] 前記単離タンパク質または融合タンパク質が、核局在化配列、細胞浸透ペプチド配列、および/または親和性タグの1つ以上を含む、[3 6]または[3 7]に記載の方法。

[3 9] 前記細胞が、幹細胞、好ましくは、胚性幹細胞、間葉系幹細胞、もしくは誘導多能性幹細胞であるか、生存動物中に存在するか、または胚中に存在する、[3 6]または[3 7]に記載の方法。

20

[4 0] 二本鎖DNA D (d s DNA) 分子を変更する方法であって、前記 d s DNA 分子を[1]～[1 4]のいずれか一項に記載の単離タンパク質および前記 d s DNA 分子の選択部分に相補的な領域を有するガイドRNAと接触させることを含む方法。

[4 1] 前記 d s DNA 分子がインビトロで存在する、[4 0]に記載の方法。

[4 2] 二本鎖DNA D (d s DNA) 分子を変更する方法であって、前記 d s DNA 分子を[1 5]に記載の融合タンパク質および前記 d s DNA 分子の選択部分に相補的な領域を有するガイドRNAと接触させることを含む方法。

[4 3] 前記 d s DNA 分子がインビトロで存在する、[4 2]に記載の方法。

30

【図 1 - 1】

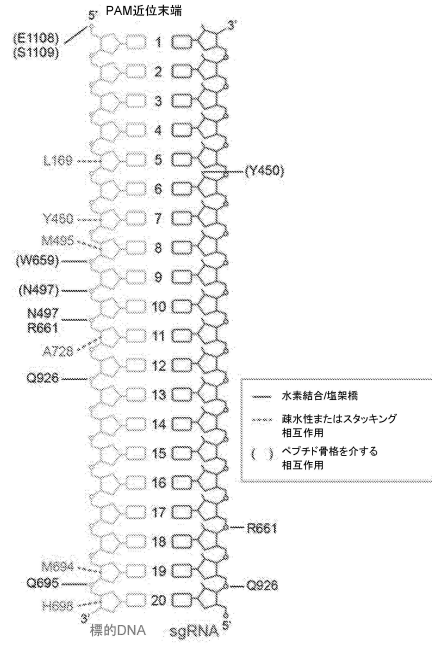
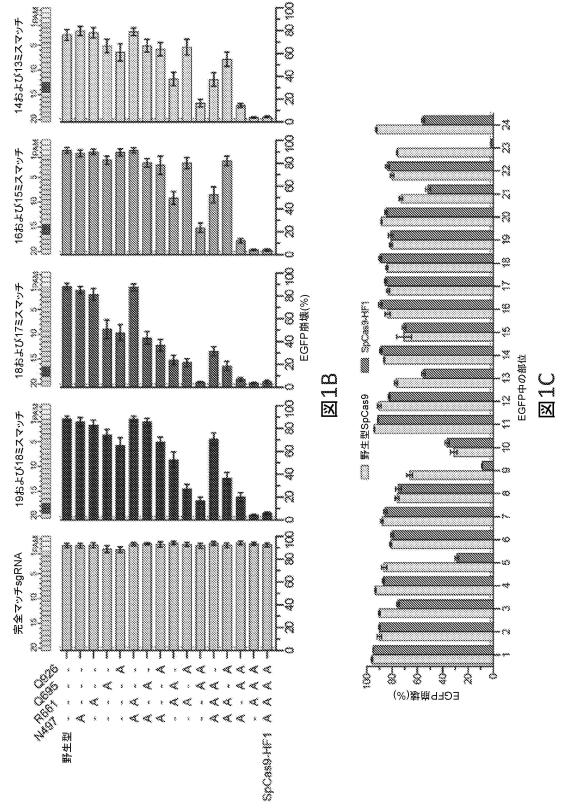
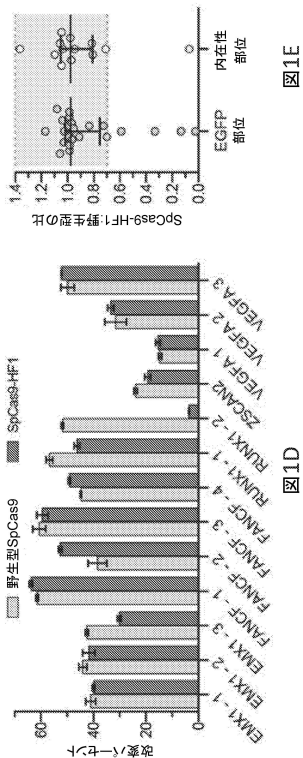


図1A

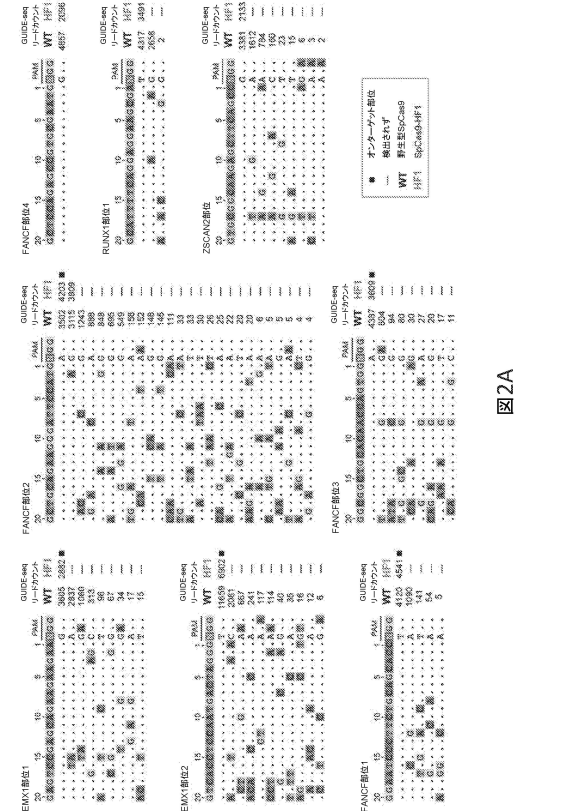
【図 1 - 2】



【図 1 - 3】



【図 2 - 1】



【図 2 - 2】

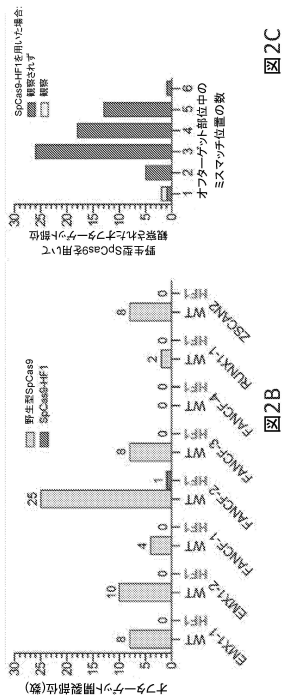


図2C

図2B

【図 3 - 1】

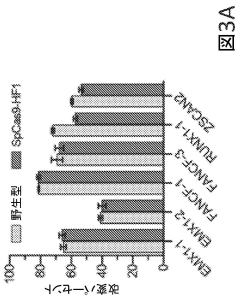


図3A

【図 3 - 2】

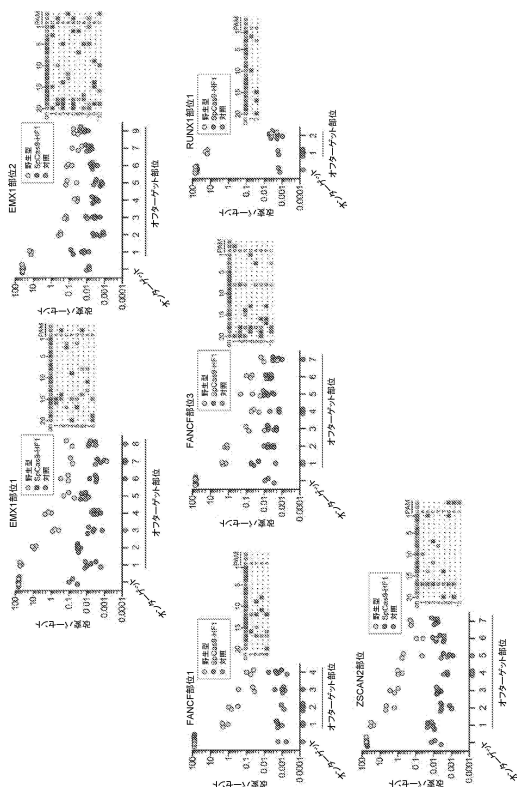


図3B

【図 3 - 3】

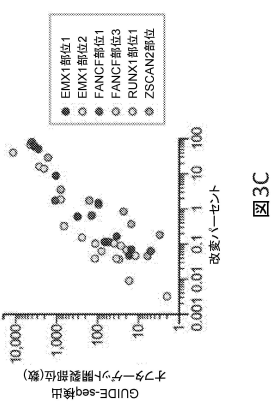
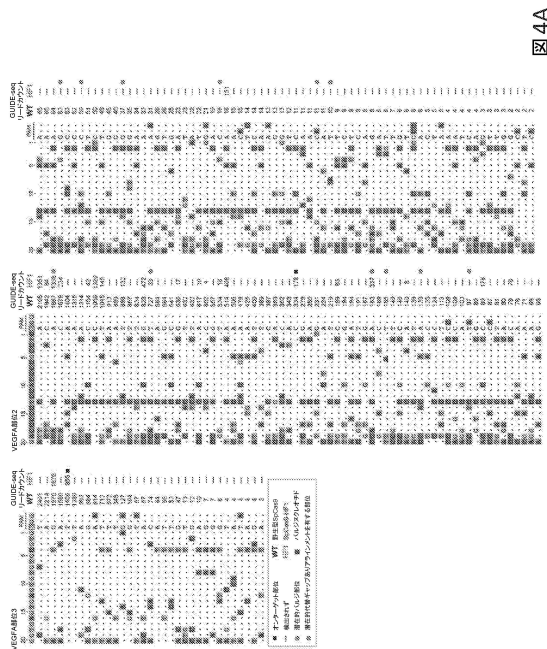


図3C

【図 4 - 1】



【図 4 - 2】

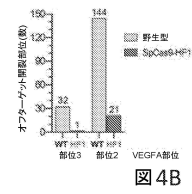


図 4B

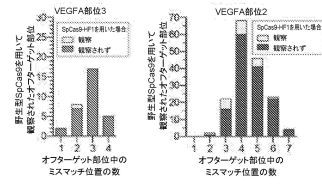


図 4C

【図 5 - 1】

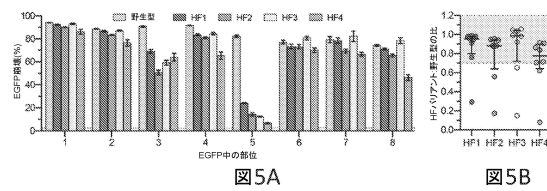


図 5A

【図 5 - 2】

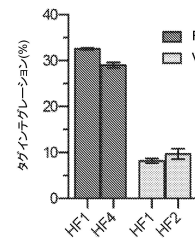


図 5B

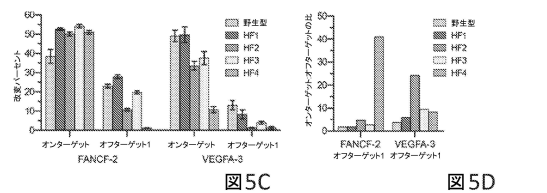


図 5C



図 5D



図 5E

【 図 6 - 1 】

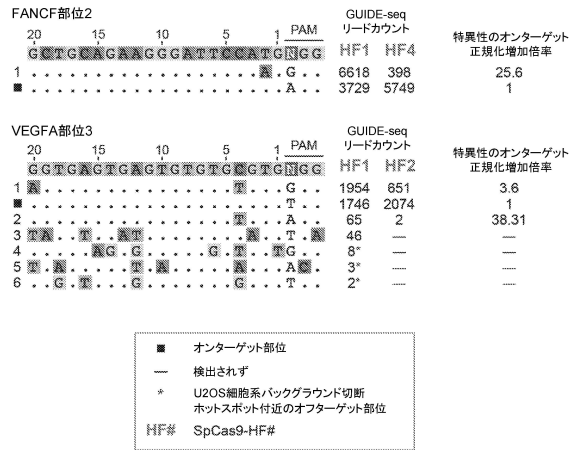


図5F

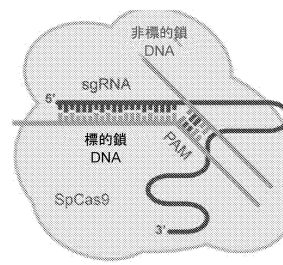


図6A

【圖 7】

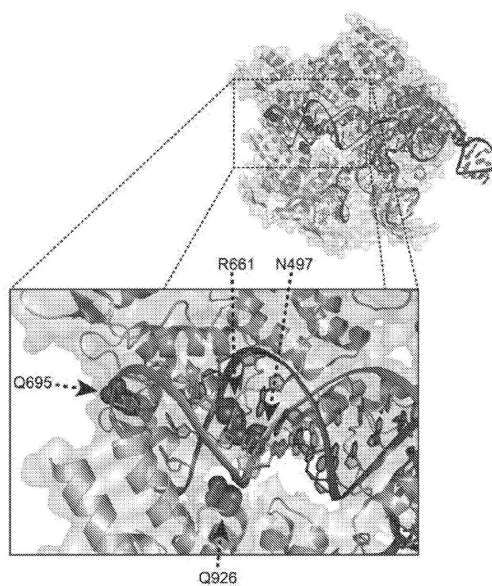
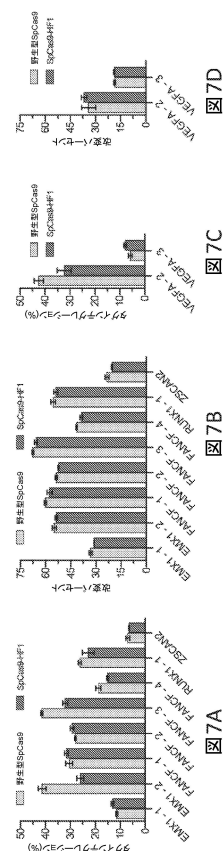


FIG. 6B



【図 8】

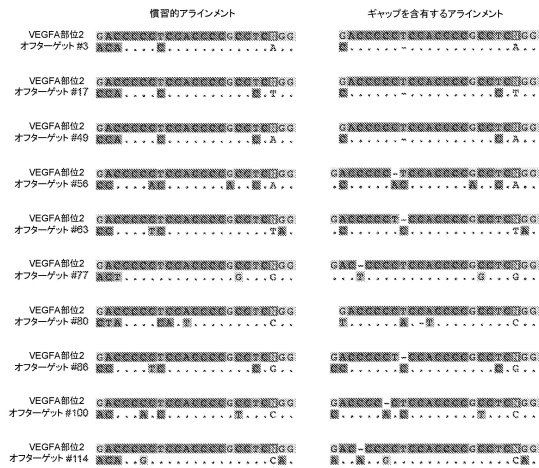


図 8

【図 9】

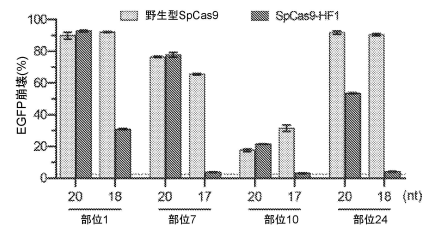


図 9

【図 10】

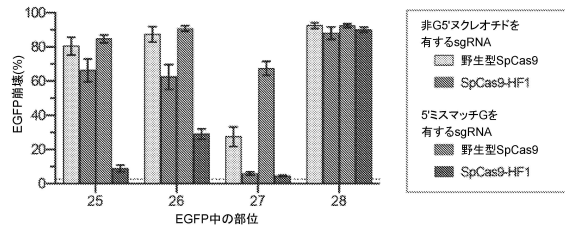


図 10

【図 12 - 1】

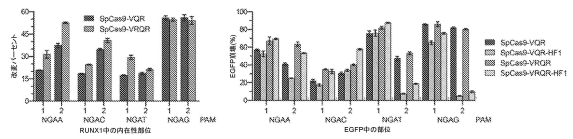


図 12A

図 12B

【図 11】

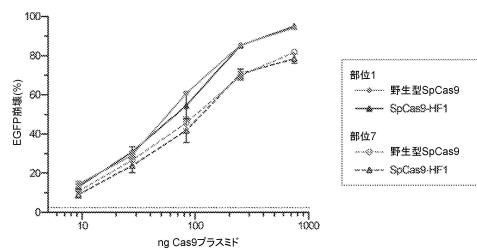


図 11

【図 12 - 2】

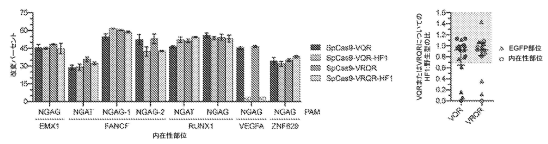


図 12C

図 12D

【 図 1 4 】

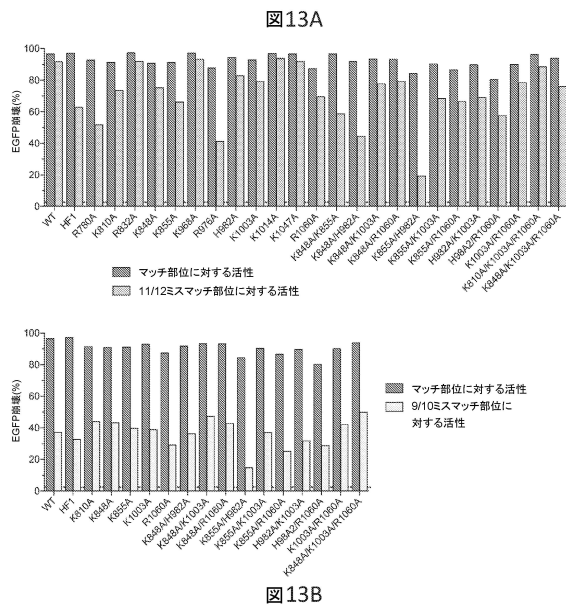


図13B

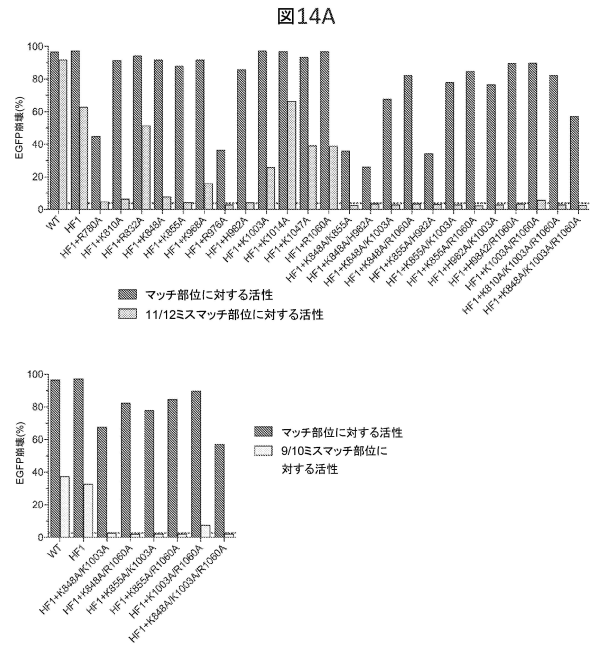


図14B

【 図 1 6 】

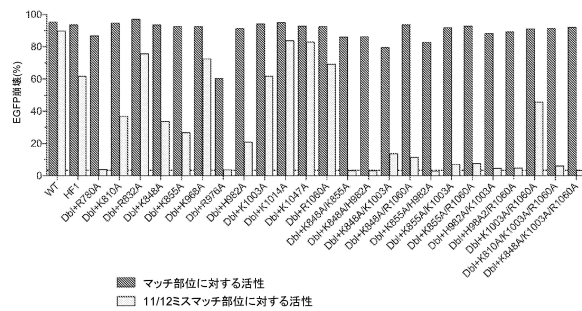


図15

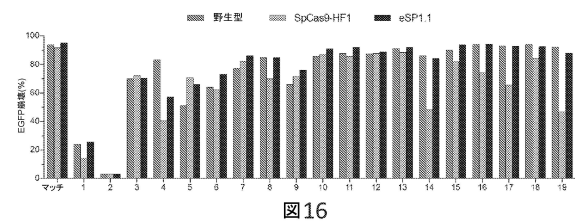


図16

【 図 1 7 】

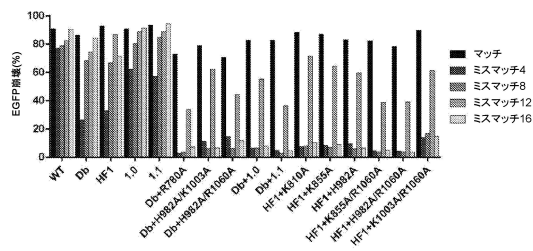


図17A

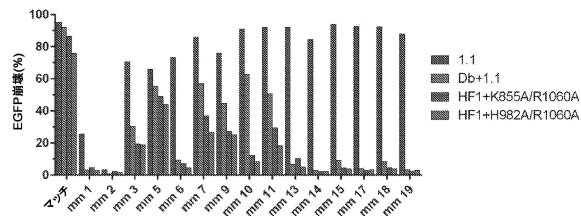


図17B

【 図 1 8 】

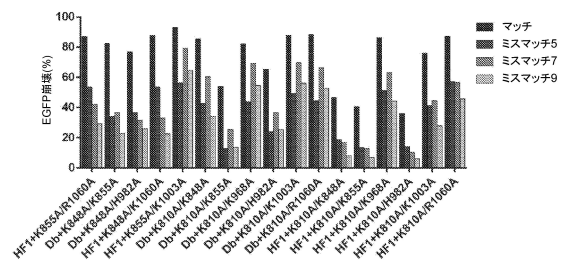


図18

【 図 1 9 - 1 】

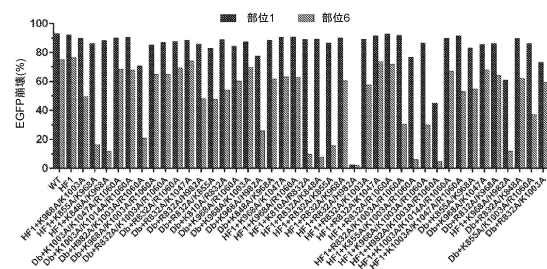


図19A

【 図 1 9 - 2 】

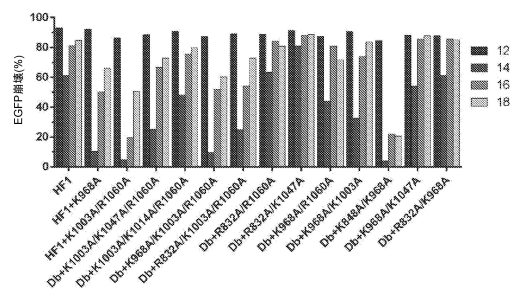
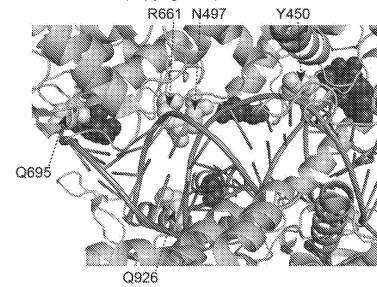


図19B

【 図 2 0 】

化膿性連鎖球菌(*S.pyogenes*)Cas9



Q926

黄色ブドウ球菌(*S.aureus*)Cas9

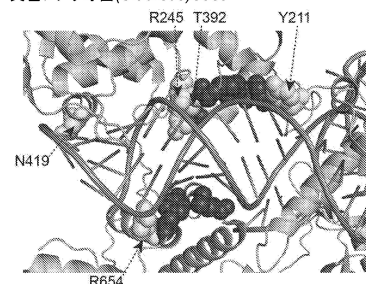


图 20

【 図 2 1 】

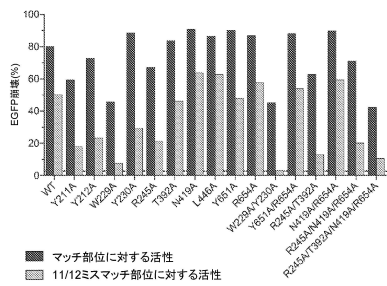


図21A

【 図 2 2 - 1 】

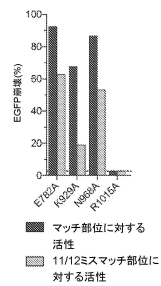


図21B

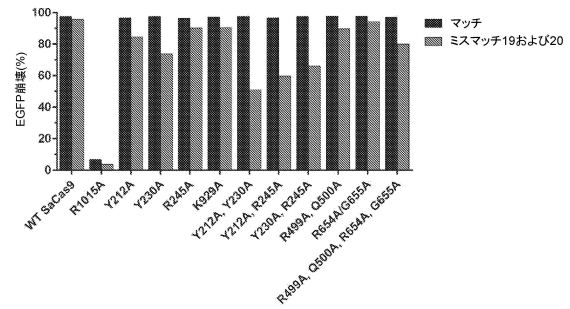


図22A

【 図 2 2 - 2 】

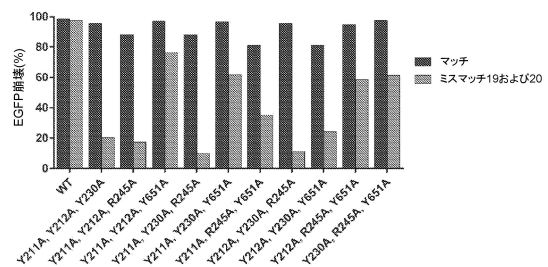


图 22B

【 ㄨ 2 3 】

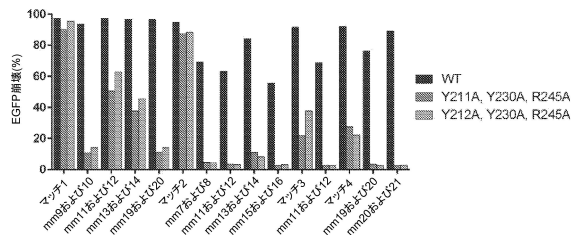


图 23

【 図 2 4 】

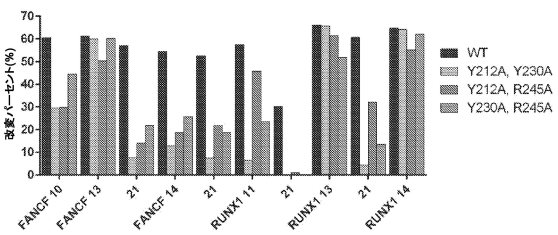


図24A

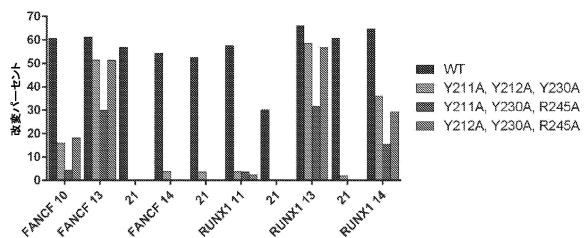


図24B

【配列表】

0006799586000001.app

フロントページの続き

(51)Int.Cl.		F I	
C 1 2 N	1/15	(2006.01)	C 1 2 N 1/15
C 1 2 N	1/19	(2006.01)	C 1 2 N 1/19
C 1 2 N	1/21	(2006.01)	C 1 2 N 1/21
C 1 2 N	5/10	(2006.01)	C 1 2 N 5/10
C 1 2 N	15/09	(2006.01)	C 1 2 N 15/09 1 1 0

- (31)優先権主張番号 62/258,280
 (32)優先日 平成27年11月20日(2015.11.20)
 (33)優先権主張国・地域又は機関
 米国(US)
- (31)優先権主張番号 62/271,938
 (32)優先日 平成27年12月28日(2015.12.28)
 (33)優先権主張国・地域又は機関
 米国(US)
- (31)優先権主張番号 15/015,947
 (32)優先日 平成28年2月4日(2016.2.4)
 (33)優先権主張国・地域又は機関
 米国(US)

早期審査対象出願

- (72)発明者 クラインスティーバー, ベンジャミン
 アメリカ合衆国 マサチューセッツ州 0 2 1 5 5, メッドフォード, ユニット 2 2 5 リバー
 ズ エッジ ドライブ 1 0 0
- (72)発明者 パターナヤック, ヴィクラム
 アメリカ合衆国 マサチューセッツ州 0 2 4 8 2, ウェルズレイ, リンデン ストリート 2 8
 0

審査官 西 賢二

- (56)参考文献 国際公開第 2 0 1 4 / 1 4 4 5 9 2 (W O , A 2)
 Kleinstiver, B. P. et al., "Engineered CRISPR-Cas9 nucleases with altered PAM specific
 ities", Nature, 2 0 1 5 年 6 月 2 2 日, Vol. 523, pp. 481-485
 Nishimasu, H. et al., "Crystal Structure of Staphylococcus aureus Cas9", Cell, 2 0 1
 5 年 8 月 2 7 日, Vol. 162, pp. 1113-1126, Supplemental Information

- (58)調査した分野(Int.Cl., D B 名)
 C 1 2 N 1 5 / 0 0 - 1 5 / 9 0
 C 1 2 N 9 / 0 0 - 9 / 9 9
 M E D L I N E / E M B A S E / B I O S I S / W P I D S (S T N)