

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第6622183号
(P6622183)

(45) 発行日 令和1年12月18日(2019.12.18)

(24) 登録日 令和1年11月29日(2019.11.29)

(51) Int.Cl.

F 1

C 12 N	9/16	(2006.01)	C 12 N	9/16	Z N A A
C 07 K	19/00	(2006.01)	C 07 K	19/00	
C 12 N	15/09	(2006.01)	C 12 N	15/09	1 1 O
C 12 N	1/15	(2006.01)	C 12 N	1/15	
C 12 N	1/19	(2006.01)	C 12 N	1/19	

請求項の数 19 (全 70 頁) 最終頁に続く

(21) 出願番号	特願2016-502853 (P2016-502853)
(86) (22) 出願日	平成26年3月14日 (2014.3.14)
(65) 公表番号	特表2016-517276 (P2016-517276A)
(43) 公表日	平成28年6月16日 (2016.6.16)
(86) 国際出願番号	PCT/US2014/028630
(87) 国際公開番号	W02014/144288
(87) 国際公開日	平成26年9月18日 (2014.9.18)
審査請求日	平成29年3月7日 (2017.3.7)
(31) 優先権主張番号	61/799,647
(32) 優先日	平成25年3月15日 (2013.3.15)
(33) 優先権主張国・地域又は機関	米国(US)
(31) 優先権主張番号	61/838,178
(32) 優先日	平成25年6月21日 (2013.6.21)
(33) 優先権主張国・地域又は機関	米国(US)

(73) 特許権者	592017633 ザ ジェネラル ホスピタル コーポレイ ション 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
(74) 代理人	100092783 弁理士 小林 浩
(74) 代理人	100120134 弁理士 大森 規雄
(74) 代理人	100104282 弁理士 鈴木 康仁
(72) 発明者	ジョン, ジェー. キース アメリカ合衆国 01890 マサチュー セッツ州, ウィンチエスター, マグノリア ウェイ 1

最終頁に続く

(54) 【発明の名称】 RNA誘導型 Fok I ヌクレアーゼ (RFN) を用いた RNA誘導型ゲノム編集の特異性の増大

(57) 【特許請求の範囲】

【請求項 1】

RNA誘導型 Fok I ヌクレアーゼ (RFN) 融合タンパク質であって、触媒的に不活性である化膿性連鎖球菌 (*Streptococcus pyogenes*) の C R I S P R 関連 9 (dCas9) のアミノ末端に融合する Fok I 触媒ドメイン配列を含む、RNA誘導型 Fok I ヌクレアーゼ (RFN) 融合タンパク質。

【請求項 2】

介在リンカーを備えた前記 Fok I 触媒ドメイン配列を含む、請求項 1 に記載の融合タンパク質。

【請求項 3】

2 ~ 30 個のアミノ酸のリンカーを含む、請求項 1 に記載の融合タンパク質。

【請求項 4】

前記リンカーが、 Gly₄ Ser を含む、請求項 3 に記載の融合タンパク質。

【請求項 5】

Fok I 触媒ドメインが、配列番号 4 のアミノ酸 388 ~ 583、または 408 ~ 583 を含む、請求項 1 に記載の融合タンパク質。

【請求項 6】

dCas9 が、D10、E762、H983、または D986 での変異、および H840 または N863 での変異を含む、請求項 1 に記載の融合タンパク質。

【請求項 7】

10

20

d C a s 9 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 Y、または H 8 4 0 N

での変異を含む、請求項 6 に記載の融合タンパク質。

【請求項 8】

請求項 1 ~ 7 のいずれかに記載の融合タンパク質をコードする核酸。

【請求項 9】

請求項 8 に記載の核酸を含むベクター。

【請求項 10】

請求項 1 ~ 7 のいずれかに記載の融合タンパク質を発現する宿主細胞。 10

【請求項 11】

生体外で、細胞において、ゲノム配列の配列特異的な崩壊を誘導する方法であって、前記方法が、請求項 1 ~ 7 のいずれかに記載の R N A 誘導型 F o k I ヌクレアーゼ (R F N) 融合タンパク質、および、2つの標的ゲノム配列に前記 R F N を配向するガイド R N A を、前記細胞に発現すること、または前記細胞と接触させることとを含み、

前記 2 つの標的ゲノム配列の外側の境界に、P A M 配列が位置している、方法。

【請求項 12】

前記 2 つの標的ゲノム配列が、1 0 ~ 2 0 塩基対離れている、請求項 1 1 に記載の方法。 20

【請求項 13】

前記 2 つの標的ゲノム配列が、1 3 ~ 1 7 塩基対離れている、請求項 1 2 に記載の方法。 20

【請求項 14】

前記ガイド R N A が、

(a) 2 つの單一ガイド R N A であって、1 つの單一ガイド R N A が第 1 の鎖を標的とし、他方のガイド R N A が相補鎖を標的とし、F o k I が各鎖を切断して、反対の D N A 鎖上に 1 対のニックをもたらし、それにより二重鎖を切断する、2 つの單一ガイド R N A 、または

(b) t r a c r R N A および 2 つの c r R N A であって、1 つの c r R N A が第 1 の鎖を標的とし、他方の c r R N A が相補鎖を標的とし、F o k I が各鎖を切断して、反対の D N A 鎖上に 1 対のニックをもたらし、それにより二重鎖を切断する、t r a c r R N A および 2 つの c r R N A 30

である、請求項 1 1 に記載の方法。

【請求項 15】

前記 2 つのガイド R N A が、それぞれ、標的ゲノム配列の 1 7 ~ 2 0 個のヌクレオチドと相補的な相補領域を含む、請求項 1 1 に記載の方法。

【請求項 16】

挿入欠失変異が、前記 2 つの標的配列の間に誘導される、請求項 1 1 ~ 1 5 のいずれか 1 項に記載の方法。 40

【請求項 17】

細胞における R N A 誘導型ゲノム編集の特異性が増大する、請求項 1 1 ~ 1 6 のいずれか 1 項に記載の方法。

【請求項 18】

前記 2 つの標的ゲノム配列が、0 ~ 3 1 ヌクレオチド離れており、かつ、それぞれ 3 ' 末端で P A M 配列を有する、請求項 1 1 ~ 1 7 のいずれか 1 項に記載の方法。

【請求項 19】

生体外で、細胞における R N A 誘導型ゲノム編集の特異性を増大させる方法であって、前記細胞を、請求項 1 ~ 7 のいずれかに記載の R N A 誘導型 F o k I ヌクレアーゼ (R F N) 融合タンパク質と接触させることとを含む、方法。 50

【発明の詳細な説明】**【技術分野】****【0001】****(優先権の主張)**

本願は、2013年3月15日に出願された米国仮特許出願第61/799,647号；2013年6月21日に出願された米国仮特許出願第61/838,178号；2013年6月21日に出願された米国仮特許出願第61/838,148号および2013年12月26日に出願された米国仮特許出願第61/921,007号に対する米国特許法119条(e)に基づく優先権を主張するものである。上記出願の内容全体が参照により本明細書に組み込まれる。

10

【0002】**(連邦支援による研究または開発)**

本発明は、米国国立衛生研究所により授与された助成番号D P 1 G M 1 0 5 3 7 8の下、政府の支援を受けてなされたものである。政府は本発明に一定の権利を有する。

【0003】

R N A 誘導型 F o k I ヌクレアーゼ (R F N) 、たとえば F o k I - d C a s 9 融合タンパク質を用いて、R N A 誘導型ゲノム編集、例えば C R I S P R / C a s 9 系を用いた編集の特異性を増大させる方法。

【背景技術】**【0004】**

近年の研究では、クラスター化され等間隔にスペーサーが入った短い回文型の反復配列 (C R I S P R) / C R I S P R 関連 (C a s) 系 (W i e d e n h e f t ら , N a t u r e 4 8 2 , 3 3 1 - 3 3 8 (2 0 1 2) ; H o r v a t h ら , S c i e n c e 3 2 7 , 1 6 7 - 1 7 0 (2 0 1 0) ; T e r n s ら , C u r r O p i n M i c r o b i o l 1 1 4 , 3 2 1 - 3 2 7 (2 0 1 1)) が、細菌、酵母およびヒト細胞のほか、ショウジョウバエ、ゼブラフィッシュおよびマウスなどのそのままの生物体において *i n v i t r o* で基板としてのゲノム編集の役割を果たし得ることが明らかにされている (W a n g ら , C e l l 1 5 3 , 9 1 0 - 9 1 8 (2 0 1 3) ; S h e n ら , C e l l R e s (2 0 1 3) ; D i c a r l o ら , N u c l e i c A c i d s R e s (2 0 1 3) ; J i a n g ら , 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 ら , E l i f e 2 , e 0 0 4 7 1 (2 0 1 3) ; H w a n g ら , N a t B i o t e c h n o l 3 1 , 2 2 7 - 2 2 9 (2 0 1 3) ; C o n g ら , S c i e n c e 3 3 9 , 8 1 9 - 8 2 3 (2 0 1 3) ; M a l i ら , S c i e n c e 3 3 9 , 8 2 3 - 8 2 6 (2 0 1 3 c) ; C h o ら , N a t B i o t e c h n o l 3 1 , 2 3 0 - 2 3 2 (2 0 1 3) ; G r a t z ら , G e n e t i c s 1 9 4 (4) : 1 0 2 9 - 3 5 (2 0 1 3)) 。化膿性連鎖球菌 (S . p y o g e n e s) 由来の C a s 9 ヌクレアーゼ (以降、単に C a s 9 と呼ぶ) は、人工的に設計された g R N A の最初の 2 0 ヌクレオチドと、プロトスペーサー隣接モチーフ (P A M) 、例えば、配列 N G G または N A G にマッチする P A M に隣接する目的とする標的ゲノム D N A 配列の相補鎖との間の塩基対相補性を介して誘導され得る (S h e n ら , C e l l R e s (2 0 1 3) ; D i c a r l o ら , N u c l e i c A c i d s R e s (2 0 1 3) ; J i a n g ら , 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 ら , E l i f e 2 , e 0 0 4 7 1 (2 0 1 3) ; H w a n g ら , N a t B i o t e c h n o l 3 1 , 2 2 7 - 2 2 9 (2 0 1 3) ; C o n g ら , S c i e n c e 3 3 9 , 8 1 9 - 8 2 3 (2 0 1 3) ; M a l i ら , S c i e n c e 3 3 9 , 8 2 3 - 8 2 6 (2 0 1 3 c) ; C h o ら , N a t B i o t e c h n o l 3 1 , 2 3 0 - 2 3 2 (2 0 1 3) ; J i n e k ら , S c i e n c e 3 3 7 , 8 1 6 - 8 2 1 (2 0 1 2)) 。 *i n v i t r o* (J i n e k ら , S c i e n c e 3 3 7 , 8 1 6 - 8 2 1 (2 0 1 2)) 、細菌 (J i a n g ら , 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)) およびヒト細胞 (C o n g ら , S c i e n c e 3 3 9 , 8 1 9 - 8 2 3 (2 0 1 3)) で実施されたこれま

20

30

40

50

での研究では、Cas9を介した切断が、場合によっては、gRNA / 標的部位接合部、特に20ヌクレオチド(nt)gRNA相補性領域の3'末端にある最後の10~12のヌクレオチド(nt)における单一のミスマッチによって、無効になり得ることが示されている。

【発明の概要】

【0005】

多くの研究で、CRISPR-Casヌクレアーゼが、最大5つのミスマッチを許容し、なおも切断することが可能であるが、任意の单一のミスマッチまたはミスマッチの組合せが活性に及ぼす影響を予測するのは困難であることが示されている。まとめると、これらのヌクレアーゼは著しいオフターゲット効果を示し得るが、その部位を予測するのが困難であり得る。本明細書には、CRISPR/Cas系を用いるゲノム編集、例えば、RNA誘導型FokIヌクレアーゼ(RFN)、たとえばFokI-Cas9またはFokI-dCas9ベースの融合タンパク質を用いるゲノム編集の特異性を増大させる方法が記載される。10

【0006】

第一の態様では、本発明は、dCas9の末端、たとえばN末端に融合するFokI触媒ドメイン配列を含むFokI-dCas9融合タンパク質であって、任意に、介在リンカー、たとえば、2~30個、たとえば4~12個のアミノ酸、たとえば、Gly₄Serのリンカーを含む、FokI-Cas9融合タンパク質を提供する。いくつかの実施形態では、FokI触媒ドメインは、配列番号4の388~583、または408~583のアミノ酸を含む。いくつかの実施形態では、dCas9は、D10、E762、H983、またはD986での変異、およびH840またはN863での変異、たとえば、(i) D10AまたはD10N；および(ii) H840A、H840YまたはH840Nでの変異を含む。20

【0007】

別の態様では、本発明は、これらの融合タンパク質をコードする核酸、この核酸を含むベクター、およびこの核酸、ベクター、または融合タンパク質を保有または発現する宿主細胞を提供する。

【0008】

別の態様では、本発明は、細胞中の二重鎖DNA分子の中、たとえば、ゲノム配列の中で配列特異的な切断を誘導する方法であって、本明細書に記載されるFokI-dCas9融合タンパク質を、細胞中で発現させるか、または細胞と接触させることと。30

(a) 2つの单一切断RNAであって、この2つの单一切断RNAのそれぞれが、標的配列の1本鎖にそれぞれ相補的な配列を含むことにより、両方のガイドRNAを使用することにより、両方の鎖の標的化をもたらし、(すなわち、1つの单一切断RNAは、第1の鎖を標的化し、他方のガイドRNAが、相補鎖を標的化する)、FokIが、反対のDNA鎖上に1対のニックをもたらす各鎖を切断することにより、二重鎖が切断される、2つの单一切断RNAと。

(b) tracrRNAおよび2つのcrrRNAであって、2つのcrrRNAのそれぞれが、標的配列内の1つの鎖と相補的な配列を含むことにより、両方のcrrRNAを使用することにより、両方の鎖の標的化をもたらし(すなわち、1つのcrrRNAは第1の鎖を標的化し、他方は、相補鎖を標的化する)、FokIが、反対のDNA鎖上に1対のニックをもたらす各鎖を切断することにより、二重鎖が切断される、tracrRNAおよび2つのcrrRNAとを含む。40

【0009】

別の態様では、本発明は、細胞のRNA誘導型ゲノム編集の特異性を増大させる方法であって、本明細書中に記載されるRNA誘導型FokIヌクレアーゼ(RFN)融合タンパク質と細胞を接触させることを含む、方法を提供する。

【0010】

本方法は、

(a) 2つの単一ガイドRNAであって、この2つの単一ガイドRNAのそれぞれが、標的配列内の1つの鎖とそれぞれ相補的である配列を含むことにより、両方のガイドRNAを使用することにより、両鎖の標的化をもたらし（すなわち1つの単一ガイドRNAが第1の鎖を標的化し、他方のガイドRNAが相補鎖を標的化する）、Fok Iが、反対のDNA鎖上に1対のニックをもたらす各鎖を切断することにより、二本鎖が切断される、2つの単一ガイド鎖と、

(b) tracrRNAおよび2つのcrrRNAであって、2つのcrrRNAのそれぞれが、標的配列内の1つの鎖に相補的な配列を含むことにより、両方のcrrRNAを使用することにより、両方の鎖の標的化をもたらし（すなわち、1つのcrrRNAが第1の鎖を標的化し、他方のcrrRNAが、相補鎖を標的化する）、Fok Iが、反対のDNA鎖上に1対のニックをもたらすことにより、二重鎖が切断される、tracrRNAおよび2つのcrrRNA

を細胞で発現させるか、この細胞と接触させることをさらに含んでもよい。

【0011】

いくつかの実施形態では、2つの標的ゲノム配列（すなわち、crrRNAまたは単一ガイドRNAの標的相補領域に相補的な配列）は、10～20塩基対、好ましくは13～17塩基対離れている。

【0012】

いくつかの実施形態では、挿入欠失変異が、2つの標的配列の間に誘導される。

20

【0013】

いくつかの実施形態では、細胞のRNA誘導型ゲノム編集の特異性が増大する。

【0014】

特に明記されない限り、本明細書で使用される技術用語および科学用語はいずれも、本発明が属する技術分野の当業者に一般に理解されるものと同じ意味を有する。本明細書には本発明で使用する方法および材料が記載されるが、ほかにも、当該技術分野で公知の他の適切な方法および材料を使用することができる。材料、方法および具体例は単に例示的なものであって、限定することを意図するものではない。本明細書で言及される刊行物、特許出願、特許、配列、データベースエントリをはじめとする参照物はいずれも、その全体が参照により組み込まれる。不一致が生じた場合、定義を含めた本明細書が優先される。

30

【0015】

本発明のその他の特徴と利点は、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になるであろう。

【図面の簡単な説明】

【0016】

【図1】標的DNA部位と結合したgRNA/Cas9ヌクレアーゼ複合体を示す模式図である。鉄はゲノムDNA標的部位のCas9ヌクレアーゼのおよその切断点を示す。ガイドRNAのヌクレオチドの番号が5'から3'に向かって逆向きに進むことに留意されたい。

40

【図2A】gRNAの5'相補性領域を短縮する原理を示す模式図である。灰色の太線=標的DNA部位、灰色の細線の構造=gRNA、灰色の楕円=Cas9ヌクレアーゼであり、黒線はgRNAと標的DNA部位との間の塩基対形成を示している。

【図2B】EGFP崩壊アッセイの模式的概観である。単一の組み込まれたEGFP-P-PESTレポーター遺伝子の標的化Cas9仲介性二本鎖切断が、誤りがちなNHEJ仲介性修復によって修復されると、細胞にコード配列を崩壊させるフレームシフト変異およびそれに付随する蛍光の消失が起こる。

【図2C】EGFPレポーター遺伝子配列の3つの異なる標的部位でアッセイした、(C)単一ミスマッチを有するsgRNAを保有するRGNの活性。反復試験の活性を完全にマッチするgRNAの活性に正規化した平均値が示されている（オンラインの方法を参照

50

)。エラーバーは平均値の標準誤差を表す。各 g RNA のミスマッチの位置が下の格子に灰色で強調されている。3つのEGFP 標的部位の配列は以下の通りであった： EGFP 部位 1 GGGCACGGGCAGCTTGCAGGTGG (配列番号 1) EGFP 部位 2 GATGCCGTTCTTCAGCTTGTCTGG (配列番号 2) EGFP 部位 3 GGTGGTGCAAGATGAACCTTCAGGG (配列番号 3)

【図 2 D】EGFP レポーター遺伝子配列の3つの異なる標的部位でアッセイした、(D)隣接する二重ミスマッチを有するsgRNA を保有する RGN の活性。反復試験の活性を完全にマッチする gRNA の活性に正規化した平均値が示されている（オンラインの方法を参照）。エラーバーは平均値の標準誤差を表す。各 gRNA のミスマッチの位置が下の格子に灰色で強調されている。3つのEGFP 標的部位の配列は以下の通りであった：

EGFP 部位 1 GGGCACGGGCAGCTTGCAGGTGG (配列番号 1)
EGFP 部位 2 GATGCCGTTCTTCAGCTTGTCTGG (配列番号 2) EGFP 部位 3 GGTGGTGCAAGATGAACCTTCAGGG (配列番号 3)

【図 2 E】EGFP レポーター遺伝子配列の3つの異なる標的部位でアッセイした、(E)様々な間隔の二重ミスマッチを有するsgRNA を保有する RGN の活性。反復試験の活性を完全にマッチする gRNA の活性に正規化した平均値が示されている（オンラインの方法を参照）。エラーバーは平均値の標準誤差を表す。各 gRNA のミスマッチの位置が下の格子に灰色で強調されている。3つのEGFP 標的部位の配列は以下の通りであった： EGFP 部位 1 GGGCACGGGCAGCTTGCAGGTGG (配列番号 1)

) EGFP 部位 2 GATGCCGTTCTTCAGCTTGTCTGG (配列番号 2) EGFP 部位 3 GGTGGTGCAAGATGAACCTTCAGGG (配列番号 3)

【図 2 F】EGFP レポーター遺伝子配列の3つの異なる標的部位でアッセイした、(F)数の増加する隣接するミスマッチを有するsgRNA を保有する RGN の活性。反復試験の活性を完全にマッチする gRNA の活性に正規化した平均値が示されている（オンラインの方法を参照）。エラーバーは平均値の標準誤差を表す。各 gRNA のミスマッチの位置が下の格子に灰色で強調されている。3つのEGFP 標的部位の配列は以下の通りであった： EGFP 部位 1 GGGCACGGGCAGCTTGCAGGTGG (配列番号 1) EGFP 部位 2 GATGCCGTTCTTCAGCTTGTCTGG (配列番号 2) EGFP 部位 3 GGTGGTGCAAGATGAACCTTCAGGG (配列番号 3)

【図 2 G】gRNA の5'末端のミスマッチの方が3'のミスマッチよりも CRISPR / Cas の感度が高くなる。gRNA は、ワトソン - クリックトランスバージョンを用いてミスマッチさせた「m」で表される位置を除いて、RNA と DNA との間でワトソン - クリック塩基対を形成する（すなわち、EGFP 部位 # 2 M18 - 19 は、位置 18 および 19 で gRNA をそのワトソン - クリックパートナーに変化させることによってミスマッチさせたものである）。gRNA の5'付近の位置は一般に許容性が極めて高いが、他の残基がミスマッチである場合、この位置でのマッチがヌクレアーゼ活性には重要である。4つの位置すべてがミスマッチである場合、ヌクレアーゼ活性はもはやは検出されなくなる。このことはさらに、この5'の位置でのマッチがほかにも3'の位置に生じたミスマッチを相殺し得ることを示している。これらの実験はコドン最適化型よりも低い絶対レベルのヌクレアーゼ活性を示し得る非コドン最適化型の Cas9 を用いて実施したものであることに留意されたい。

【図 2 H】ヒト細胞ベースの U2OS EGFP 崩壊アッセイにおいて 15 ~ 25 nt の範囲の様々な長さの相補性領域を有する gRNA によって誘導された Cas9 ヌクレアーゼ活性の効率。U6 プロモーターからの gRNA 発現には 5' G の存在が必要であるため、標的 DNA 部位に対して特定の長さの相補性 (15 nt, 17 nt, 19 nt, 20 nt, 21 nt, 23 nt および 25 nt) を保有する gRNA に限り評価することが可能であった。

【図 3 A】Cas9 および EGFP レポーター遺伝子の4か所の標的部位に対する完全長 gRNA または短縮 gRNA によって仲介される EGFP 崩壊のヒト細胞における効率。

10

20

30

40

50

相補性領域の長さおよび対応する標的DNA部位が示されている。Ctr1 = 相補性領域を欠く対照gRNA。

【図3B】マッチする標準的なRGNおよびtru-RGNによって7つの異なるヒト内在遺伝子標的に導入された標的化挿入欠失変異の効率。相補性領域の長さおよび対応する標的DNA部位が示されている。挿入欠失頻度はT7EIアッセイによって測定したものである。Ctr1 = 相補性領域を欠く対照gRNA。

【図3C】EMX1部位を標的とするtru-gRNAまたはマッチする完全長gRNAを用いたRGNによって誘発された挿入欠失変異のDNA配列。標的DNA部位のgRNA相補性領域と相互作用する部分が灰色で強調され、PAM配列の最初の塩基が小文字で示されている。欠失が灰色で強調された破線で表され、挿入が灰色で強調されたイタリック体の文字で表されている。欠失または挿入塩基の正味の数および各配列が単離された回数が右側に示されている。10

【図3D】マッチする標準的なRGNおよびtru-RGNによって2つの内在ヒト遺伝子に導入された正確なHDR/ssODN仲介性変化の効率。%HDRはBamHI制限消化アッセイを用いて測定したものである(実施例2の実験手順を参照されたい)。対照gRNA = 空のU6プロモーターベクター。

【図3E】U2OS.EGFP細胞に可変量の完全長gRNA発現プラスミド(上段)またはtru-gRNA発現プラスミド(下段)を一定量のCas9発現プラスミドとともにトランスフェクトした後、EGFP発現が減少した細胞の百分率をアッセイした。二重反復実験の平均値が平均値の標準誤差とともに示されている。この3か所のEGFP標的部位では、tru-gRNAで得られたデータがtru-gRNAプラスミドの代わりに完全長gRNA発現プラスミドで実施した実験のデータと緊密に一致していることに留意されたい。20

【図3F】U2OS.EGFP細胞に可変量のCas9発現プラスミドを可変量の完全長gRNA発現プラスミド(上段)またはtru-gRNA発現プラスミド(下段)(図3Eの実験から各tru-gRNAの量を決定した)とともにトランスフェクトした。二重反復実験の平均値が平均値の標準誤差とともに示されている。この3か所のEGFP標的部位では、tru-gRNAで得られたデータがtru-gRNAプラスミドの代わりに完全長gRNA発現プラスミドで実施した実験のデータと緊密に一致していることに留意されたい。これらの漸増法の結果から、実施例1および2で実施したEGFP崩壊アッセイに使用するプラスミドの濃度を決定した。30

【図4A】RNA誘導型FokIヌクレアーゼおよびCRISPR/CasサブタイプYpestタンパク質4(Csy4)ベースのマルチプレックスgRNA発現系である。

(a) RNA誘導型FokIヌクレアーゼの図式的概要。2つのFokI-dCas9融合タンパク質が、FokIの二量体化およびDNAの切断を促進するために、2つの異なるgRNAより隣接した標的部位に動員される。

【図4B】RNA誘導型FokIヌクレアーゼおよびCRISPR/CasサブタイプYpestタンパク質4(Csy4)ベースのマルチプレックスgRNA発現系である。

(b) Csy4ベースのマルチプレックスgRNA発現系の図式的概要。2つのgRNA(いずれかの5'末端ヌクレオチドを備える)が、Csy4認識部位に隣接するそれぞれのgRNAと、UGプロモーターから単一の転写物中で共発現される。Csy4は、転写物からgRNAを切断し、放出する。Csy4認識領域は、gRNAの3'末端に維持され、Csy4ヌクレアーゼがこの部位に結合する。40

【図4C】RNA誘導型FokIヌクレアーゼおよびCRISPR/CasサブタイプYpestタンパク質4(Csy4)ベースのマルチプレックスgRNA発現系である。

(c) 多重Csy4ベースの系の検証。EGFPの隣接部位を標的とした2つのgRNAを、Csy4およびCas9のヌクレアーゼと共に、ヒトU2OS.EGFP細胞のCsy4ベースの系を使用して単一のRNA転写物中に発現した。これらの細胞に誘導された挿入欠失変異の配列が示される。野生型の配列が上部に示され、両方の標的部位が灰色で強調されて、PAM配列が下線を施した文字として示される。欠失は、灰色の背景に対し50

て破線により示され、挿入は、灰色の背景に対して小文字により示される。各配列の右側には、挿入(+)または欠失()の大きさが特定される。

【図5A】RNA誘導型FokIヌクレアーゼの設計および最適化。(a)ZFN、TALEN、FokI-dCas9融合物およびdCas9-FokI融合物の概略的な例示。

【図5B】RNA誘導型FokIヌクレアーゼの設計および最適化。(b)2つの配向:PAM内部(PAM_{in})(左側のパネル)およびPAM外部(PAM_{out})(右側のパネル)のうちの1つの片側部位を標的化とするgRNA対でのFokI-dCas9融合物のEGFP崩壊の活性化のスクリーニング。片側部位は、0~31bpの範囲の可変的な長さのスペーサー配列により分離した。EGFP崩壊は、フローサイトメトリーにより定量した(n=1)。dCas9-FokI融合物および同一のgRNA対の対応するデータを図5Eに示す。

【図5C】RNA誘導型FokIヌクレアーゼの設計および最適化。(c)PAM外部と共に配向され、10~20bpの範囲の長さのスペーサー長と共に配向される片側部位を備える標的部位のFokI-dCas9仲介性EGFP崩壊の追加的な評価。EGFP崩壊は、フローサイトメトリーにより定量化された。エラーバーは、平均値の標準誤差(s.e.m)を示す(n=2)。

【図5D】RNA誘導型FokIヌクレアーゼの設計および最適化。(d)スペーサー長により分類した(c)からのデータの、EGFP崩壊の平均値。エラーバーは、s.e.mを示す。

【図5E】RNA誘導型FokIヌクレアーゼの設計および最適化。(e)これらのプロットは、0~31bpのスペーシングおよびPAM内部およびPAM外部の配向を伴う60個のgRNA対でのU2OS.EGFP細胞のEGFP崩壊アッセイにおけるdCas9-FokI活性のスクリーニング結果を示す。

【図5F-1】RNA誘導型FokIヌクレアーゼの設計および最適化。(f)U2OS細胞のFokI-dCas9誘導型変異の配列を示す。Cas9またはFokI-dCas9により結合される23nt標的配列を灰色で示す。プロトスペーサー隣接モチーフ、すなわちPAM配列を、下線を付して太字で表す。欠失を、薄い灰色の背景上に破線を付して表す。挿入は灰色で強調する。挿入または欠失した塩基の正味の数は、配列の直右側の縦列中に示される。

【図5F-2】RNA誘導型FokIヌクレアーゼの設計および最適化。(f)U2OS細胞のFokI-dCas9誘導型変異の配列を示す。Cas9またはFokI-dCas9により結合される23nt標的配列を灰色で示す。プロトスペーサー隣接モチーフ、すなわちPAM配列を、下線を付して太字で表す。欠失を、薄い灰色の背景上に破線を付して表す。挿入は灰色で強調する。挿入または欠失した塩基の正味の数は、配列の直右側の縦列中に示される。

【図5F-3】RNA誘導型FokIヌクレアーゼの設計および最適化。(f)U2OS細胞のFokI-dCas9誘導型変異の配列を示す。Cas9またはFokI-dCas9により結合される23nt標的配列を灰色で示す。プロトスペーサー隣接モチーフ、すなわちPAM配列を、下線を付して太字で表す。欠失を、薄い灰色の背景上に破線を付して表す。挿入は灰色で強調する。挿入または欠失した塩基の正味の数は、配列の直右側の縦列中に示される。

【図5F-4】RNA誘導型FokIヌクレアーゼの設計および最適化。(f)U2OS細胞のFokI-dCas9誘導型変異の配列を示す。Cas9またはFokI-dCas9により結合される23nt標的配列を灰色で示す。プロトスペーサー隣接モチーフ、すなわちPAM配列を、下線を付して太字で表す。欠失を、薄い灰色の背景上に破線を付して表す。挿入は灰色で強調する。挿入または欠失した塩基の正味の数は、配列の直右側の縦列中に示される。

【図6A】FokI-dCas9-RFNの二量体化が、有効なゲノム編集の活性に必要である。(a)(EGFP部位47および81を)正確に標的とするgRNA対、およ

10

20

30

40

50

び g RNA の 1 つまたは他方が、(V E G F A 遺伝子中の) 非 E G F P 配列を標的とする別の g RNA と置き換えられている対の存在下で評価される 2 つの R E N 対の E G F P 崩壊の活性。 E G F P 崩壊は、フローサイトメトリーにより定量化された。 E G F P : 高感度緑色蛍光タンパク質、 V E G F A : 血管内皮増殖因子 A 。エラーバーは、平均値の標準誤差の (s . e . m) を表す (n = 3) 。

【図 6 B】 F o k I - d C a s 9 R F N の二量体化が、有効なゲノム編集の活性に必要である。(b) (a) の E G F P 崩壊アッセイで使用されるのと同一の細胞由来のゲノム D N A で実施した T 7 E I アッセイによる変異誘発の頻度の定量化。エラーバーは、 s . e . m を表す (n = 3) 。

【図 6 C】 F o k I - d C a s 9 R F N の二量体化が、有効なゲノム編集の活性に必要である。(c) A P C 、 M L H 1 および V E G F A 遺伝子中の部位を標的とする R F N の活性化。各標的で、本出願人は、「左の」片側部位では 1 つの g RNA のみ、または「右の」片側部位では 1 つの g RNA のみである、 1 対の同族の g RNA と F o k I - d C a s 9 を共発現させた。変異原性の比率は、 T 7 E 1 アッセイにより測定した。 A P C : 大腸腺腫症 ; M L H 1 : m u t L 相同体 1 ; V E G F A : 血管内皮細胞増殖因子。エラーバーは s . e . m を表す (n = 3) 。

【図 6 D】 F o k I - d C a s 9 R F N の二量体化が、有効なゲノム編集の活性に必要である。(d) V E G F A 部位 1 を標的とするために使用される g RNA の 1 つでのオンターゲット部位および 5 つの従来より知られているオフターゲット (O T) 部位で V E G F A 部位 1 を標的とする R F N の変異誘発頻度。変異の頻度は、ディープシーケンシングにより決定した。報告される各値は、 3 つの独立したトランスフェクションの実験から集められたゲノム D N A から調製した単一のディープシーケンシングライブラリーから決定した。オンターゲット V E G F A 部位 1 について示される値 (アスタリスクが付される) は、以下の図 4 a の値と同一であり、この図面に表される値との比較を簡単にすることのみを目的として示される。

【図 7 A】 単一 g RNA と共に発現した C a s 9 ニッカーゼまたは F o k I - d C a s 9 の変異原性活性。(a) 6 つの異なるヒト遺伝子部位を標的とする 1 つまたは 2 つの g RNA の存在下での F o k I - d C a s 9 (左のバー) または C a s 9 ニッカーゼ (中央のバー) により誘導される挿入欠失変異の頻度。各遺伝子標的について、両方の g RNA での挿入欠失の頻度を評価した。「左」の片側部位では 1 つのみの g RNA 、または「右」の片側部位では他方の g RNA のみを用いる。変異の頻度を、ディープシーケンシングにより決定した。報告するそれぞれの挿入欠失の頻度は、 3 つの独立したトランスフェクションの実験から集めたゲノム D N A から調製した単一のディープシーケンシングライブラリーから決定した。 V E G F A : 血管内皮増殖因子 A 、 D D B 2 : 損傷特異的 D N A 結合タンパク質 2 (D a m a g e - S p e c i f i c D N A B i n d i n g P r o t e i n 2) ; F A N C F : ファンコニ貧血相補群 F (F a n c o n i A n e m i a , C o m p l e m e n t a t i o n G r o u p F) ; F E S : ネコ肉腫ウイルス癌遺伝子 ; R U N X 1 : R u n t 関連転写因子 1 (R u n t - R e l a t e d T r a n s c r i p t i o n F a c t o r 1) 。

【図 7 B】 単一 g RNA と共に発現した C a s 9 ニッカーゼまたは F o k I - d C a s 9 の変異原性活性。(b) 単一 g RNA の実験または対照の実験 (g RNA なし、および C a s 9 ニッカーゼまたは F o k I - d C a s 9 なし) のそれぞれに対して、 g RNA 対を備える各標的部位で得られた値と比較した挿入欠失の頻度における倍数減少として表される (a) 由来のデータ。この倍数減少は、 F o k I - d C a s 9 (各対の左側のバー、薄灰色) および C a s 9 ニッカーゼ (各対の右側のバー、濃灰色) の両方について計算された。

【図 8】 単一の C a s 9 ニッカーゼは、それぞれの標的部位に点変異を高い効率で導入することができる。 F o k I - d C a s 9 、 C a s 9 ニッカーゼ、または t d T o m a t o 対照の存在下の (a) V E G F A 、 (b) F A N C F 、および (c) R U N X 1 の遺伝子標的での、単一 g RNA が標的化とする片側部位の各位置で見出される異なる点変異の

10

20

30

40

50

頻度。変異の頻度を、ディープシーケンシングにより決定した。報告する各点変異の値は、3つの独立したトランスフェクションの実験から集めたゲノムDNAから調製した単一ディープシーケンシングから決定した。これらの実験に使用するゲノムDNAは、図7A～Bの挿入欠失変異で分析した同一の細胞から単離したこと留意されたい。VEGFA：血管内皮増殖因子A；FANCF：ファンコニ貧血相補群F；RUNX1：Runt関連転写因子1。

【発明を実施するための形態】

【0017】

(詳細な説明)

C R I S P R RNA誘導型ヌクレアーゼ(RGN)は、簡便で効率的なゲノム編集のプラットフォームとして急速に登場した。Marraffinら(Jiangら,Nat Biotechnol 31, 233-239(2013))は近年、細菌でのCas9 RGNの特異性を体系的に研究したが、ヒト細胞でのRGN特異性については十分に明らかにされていない。これらのヌクレアーゼを研究および治療応用に広く用いるのであれば、ヒトをはじめとする真核細胞においてRGNによるオフターゲット効果が及ぶ範囲を理解することが極めて重要になる。本発明者らは、ヒト細胞ベースのレポーターアッセイを用いてCas9ベースのRGNによるオフターゲット切断の特徴を明らかにした。ガイドRNA(gRNA)-DNA接合部に沿った位置に応じて、様々な程度で單一および二重のミスマッチが許容された。一部ミスマッチのある部位を調べることによって、ヒト細胞の内在遺伝子座を標的とした6種類のRGNのうちの4種類によって誘発されるオフターゲット変化が迅速に検出された。特定されたオフターゲット部位は最大5つのミスマッチを保有しており、その多くが目的とするオントーゲット部位で観察された頻度と同等(またはそれ以上)の頻度で変異していた。したがって、RGNはヒト細胞において、不完全にマッチしたRNA-DNAに対しても高い活性を示し、この観察結果は研究および治療応用での使用を複雑なものにしかねないものである。

【0018】

本明細書に記載される結果は、任意のRGNの特異性プロファイルを予測することが容易なものでもなく単純なものでもないことを示している。EGFPレポーターアッセイの実験は、單一および二重のミスマッチがヒト細胞でのRGN活性に様々な影響を及ぼし得るものであり、その影響はミスマッチの標的部位内の位置(1つまたは複数)に厳密に依存するものではないことを示している。例えば、これまでに公開されている報告の通り、一般にgRNA/DNA接合部の3'側半分にみられる変化の方が、5'側半分にみられる変化よりも影響が大きいが(Jiangら,Nat Biotechnol 31, 233-239(2013); Congら, Science 339, 819-823(2013); Jinneklら, Science 337, 816-821(2012))。しかしながら3'末端の單一および二重変異が十分に許容されると思われる場合もあるのに対して、5'末端の二重変異は活性を大幅に低下させ得る。さらに、任意の位置(1つまたは複数)のミスマッチの影響の大きさは部位依存的であると思われる。可能なあらゆるヌクレオチド置換(本発明者らのEGFPレポーター実験で用いられるワツソン-クリックトランスバージョン以外にも)を試験し、大きな一連のRGNに関して広範囲にわたるプロファイリングを実施すれば、オフターゲットの及び得る範囲についてさらに洞察が得られると考えられる。この点に関して、近年記載されたMarraffinらの細菌細胞ベースの方法(Jiangら,Nat Biotechnol 31, 233-239(2013))または以前にLiulagZFNに適用したin vitroのコンビナトリアルライブラリーベースの切断部位選択手法(Pattanayakら,Nat Methods 8, 765-770(2011))が、さらに大きなRGN特異性プロファイルの作製に有用であると考えられる。

【0019】

RGN特異性を広範囲にわたって予測するには上に挙げたような困難が伴うが、オントーゲット部位と1～5つのミスマッチの分だけ異なる一部のゲノム部位を調べることによ

10

20

30

40

50

つて、真の RGN によるオフターゲットを特定することができた。注目すべきことに、これらの実験の条件下では、多くのこれらのオフターゲット部位における RGN 誘発変異の頻度は、目的とするオンターゲット部位で観察された頻度とほぼ同じ（またはそれ以上）であり、T7E1 アッセイ（本発明者らの実験室で実施した場合、信頼できる変異頻度の検出限界が約 2 ~ 5 % である）を用いてこれらの部位を検出することが可能であった。これらの変異率は極めて高いため、これまでるかに頻度の低い ZFN 誘発オフターゲット変化および TALEN 誘発オフターゲット変化の検出に必要とされたディープシーケンシング法の使用を回避することができた（Pattanayakula, Nat Methods 8, 765 - 770 (2011); Perezら, Nat Biotechnol 26, 808 - 816 (2008); Gabrielら, Nat Biotechnol 29, 816 - 823 (2011); Hockemeyerら, Nat Biotechnol 29, 731 - 734 (2011)）。このほか、ヒト細胞における RGN オフターゲット変異誘発の解析から、RGN 特異性を予測することが困難であることが確認され、単一および二重のミスマッチのあるオフターゲット部位がすべて変異の証拠を示すわけではないが、ミスマッチが 5 つに及ぶ一部の部位でも変化がみられた。さらに、特定された真のオフターゲット部位には、目的とする標的配列と比較して、転移または転換による差への明らかな偏りは全くみられない。

【0020】

いくつかの RGN にオフターゲット部位がみられたが、このような部位の特定は広範囲の規模でもゲノム全域にわたる規模でもなかった。研究対象にした 6 種類の RGN では、ヒトゲノム内にある総数がはるかに多い潜在的オフターゲット配列のごく一部のものだけを検討した。T7E1 アッセイによってこれだけ多数のオフターゲット変異の遺伝子座を検討するのは実用的な戦略でも費用効果に優れた戦略でもないが、今後の研究でハイスクープットシーケンシングを使用することになれば、多数のオフターゲット部位の候補を調べることが可能になり、より高感度に真のオフターゲット変異を検出する方法が得られるものと考えられる。例えば、そのような方法を用いれば、本発明者らがいかなるオフターゲット変異も明らかにできなかった 2 種類の RGN について、さらなるオフターゲット部位を明らかにすることが可能になると考えられる。さらに、細胞内の RGN 活性に影響を及ぼし得る RGN 特異性とエピゲノミックな要因（例えば、DNA メチル化およびクロマチン状態）の両方の理解が進展すればほかにも、検討する必要のある潜在的部位の数が減少し、これによりゲノム全域にわたる RGN オフターゲットの評価の実用性が高まり、価格も手頃なものになり得ると考えられる。

【0021】

ゲノムオフターゲット変異の頻度を最小限に抑えるのにいくつかの戦略を用いることができる。例えば、RGN 標的部位の特異的選択を最適化することができる。目的とする標的部位と最大 5 つの位置において異なるオフターゲット部位が RGN によって効率的に変異されるとすると、ミスマッチの計数によって判定される最小数のオフターゲット部位を有する標的部位を選択するのが効果的であるとは思われない。ヒトゲノム内の配列を標的とする任意の RGN には通常、20 bp の RNA : DNA 相補性領域内の 4 つまたは 5 つの位置が異なる何千もの潜在的オフターゲット部位が存在することになる。このほか、gRNA 相補性領域のヌクレオチド含有量が潜在的オフターゲット効果の範囲に影響を及ぼす可能性がある。例えば、GC 含有量が高いと RNA : DNA ハイブリッドが安定することが示されており（Sugimotoら, Biochemistry 34, 11211 - 11216 (1995)）、したがって、gRNA / ゲノム DNA のハイブリダイゼーションの安定性およびミスマッチの許容性も高まることが予想されると考えられる。この 2 つのパラメータ（ゲノム内のミスマッチ部位の数および RNA : DNA ハイブリッドの安定性）がゲノム全域にわたる RGN 特異性に影響を及ぼす可能性およびその機序を評価するには、gRNA の数を増やした実験がさらに必要となる。しかし、このような予測パラメータを定めることができるとても、このようなガイドラインの実施による影響が RGN の標的化の範囲をさらに制限することになる可能性があることに留意することが重

要である。

【0022】

R G N 誘発オフターゲット効果を低減する一般的な戦略で有望なものの中の1つに、細胞内で発現するg R N A およびC a s 9 ヌクレアーゼの濃度を低くすることが考えられる。U 2 O S . E G F P 細胞のV E G F A 標的部位2および3にR G N を用いて、この考えを検討した。トランسفェクトするg R N A 発現およびC a s 9 発現プラスミドを減らしたところ、オンターゲット部位での変異率が低下したが、オフターゲット変異の相対的比率はあまり変化しなかった。これと同じように、他の2種類のヒト細胞型（H E K 2 9 3 細胞およびK 5 6 2 細胞）でも、オンターゲット変異誘発の絶対的比率がU 2 O S . E G F P 細胞より低いものの、高レベルのオフターゲット変異誘発率が観察された。10 したがって、細胞内でのg R N A およびC a s 9 の発現レベルを低下させてもオフターゲット効果を低減する解決策にはならないと思われる。さらに、上の結果はほかにも、ヒト細胞に観察される高率のオフターゲット変異誘発がg R N A および/またはC a s 9 の過剰発現に起因するものではないことを示唆している。

【0023】

3種類の異なるヒト細胞型でR G N によって大幅なオフターゲット変異誘発が引き起こされ得るという観察結果は、このゲノム編集プラットフォームの使用に重要な意味を持つ。研究に適用する場合、特に望ましくない変化の他配が課題となる世代時間の長い培養細胞または生物体を用いる実験では、高頻度のオフターゲット変異の潜在的に複雑な作用を考慮に入れる必要がある。20 オフターゲット効果はランダムなものではなく、標的とする部位と関係があるため、このような作用を制御する方法の1つとして、異なるD N A 配列を標的とする複数のR G N を用いて同じゲノム変化を誘発することが考えられる。しかし、治療に適用する場合、ここに挙げた観察結果は、これらのヌクレアーゼをヒト疾患の治療により長期間にわたって安全に使用するのであれば、R G N 特異性を慎重に定め、かつ/または改善する必要があることを明確に示している。

【0024】

特異性を改善する方法

本明細書に示されるように、化膿性連鎖球菌（S . p y o g e n e s ）C a s 9 タンパク質を土台とするC R I S P R - C a s R N A 誘導型ヌクレアーゼは、目的とするオンターゲット活性と同等以上の著しいオフターゲット変異誘発効果を有し得る（実施例1）。30 このようなオフターゲット効果は、研究の適用する際に、また特に将来治療に適用する際に問題となり得る。したがって、C R I S P R - C a s R N A 誘導型ヌクレアーゼ（R G N ）の特異性を改善する方法が必要とされている。

【0025】

実施例1に記載されるように、C a s 9 R G N はヒト細胞のオフターゲット部位に高頻度の挿入欠失変異を誘発し得る（このほか、C r a d i c k l a , 2 0 1 3 ; F u l a , 2 0 1 3 ; H s u l a , 2 0 1 3 ; P a t t a n a y a k l a , 2 0 1 3 を参照されたい）。このような望ましくない変化は、目的とするオンターゲット部位と5つものミスマッチによって異なるゲノム配列に起こり得る（実施例1を参照されたい）。さらに、g R N A 相補性領域の5'末端のミスマッチは一般に、3'末端のミスマッチよりも許容されやすいが、このような関係は絶対的なものではなく、部位依存性を示す（実施例1およびF u l a , 2 0 1 3 ; H s u l a , 2 0 1 3 ; P a t t a n a y a k l a , 2 0 1 3 を参照されたい）。その結果、現在、ミスマッチの数および/または位置に依存するコンピュータを用いる方法は、真のオフターゲット部位を特定するために予測に用いる価値が低下している。40 したがって、R N A 誘導型ヌクレアーゼを研究および治療応用に使用するのであれば、オフターゲット変異の頻度を低下させる方法が依然として重要な優先事項である。

【0026】

二量体化は、C a s 9 ヌクレアーゼの特異性を改善する魅力的な戦略となる可能性がある。これは、真の二量体系ではない、対形成されたC a s 9 ニッカーゼの手法とは異なるものである。対形成されたニッカーゼは、D N A のセグメント上に2つのC a s 9 ニッカ50

ーゼを同時に局在することにより働き、これにより、定義されていない機構を介して効率の高いゲノム編集を誘導する。二量体化は、酵素活性に必要ではないため、単一の C a s 9 ニッカーゼもまた、特定の部位で（知られていない機構を介して）挿入欠失を高い効率で誘導することもでき、したがって、ゲノム中の望ましくないオフターゲット変異を潜在的に引き起こす可能性がある。

【 0 0 2 7 】

したがって、R G N の特異性を改善する 1 つの手法として、D 1 0 A および H 8 4 0 A の変異を有する触媒的に不活性な形態の C a s 9 (d C a s 9 としても知られる) に、F o k I エンドヌクレアーゼドメインを融合することがある。F o k I ヌクレアーゼドメインは、二量体として機能し、したがって、二重鎖の切断を誘導するために、2 つのサブユニットが D N A の同一の局所部分に動員されなければならない。この構成（図 9 A および実施例 2 ）では、2 つの F o k I - d C a s 9 融合物が、二重鎖の切断を得るために 2 つの異なる g R N A を使用した適切な構成で動員される。したがって、この系では、F o k I - d C a s 9 融合物は、単一の R G N の部位より 2 倍長い部位に結合し、それにより、この系はより特異的になると予測される。

【 0 0 2 8 】

よって、本明細書は、F o k I - d C a s 9 融合タンパク質であって、F o k I 配列が、d C a s 9 (好ましくは、d C a s 9 のアミノ酸末端、または任意にカルボキシ末端) に融合し、任意に、介在リンカー、たとえば、任意に 2 ~ 3 0 個のアミノ酸、たとえば 4 ~ 1 2 個のアミノ酸、たとえば G l y 4 S e r (配列番号： 2 3) または (G l y 4 S e r) 3 のリンカーを含む、F o k I - d C a s 9 融合タンパク質を提供する。いくつかの実施形態では、融合タンパク質は、d C a s 9 および F o k I のドメインの間にリンカーを含む。これらの融合タンパク質（または連結構造の融合タンパク質の間）に使用し得るリンカーは、融合タンパク質の機能に干渉しない任意の配列を含み得る。好ましい実施形態では、リンカーは短く、例えば 2 ~ 2 0 アミノ酸であり、通常は柔軟である（すなわち、グリシン、アラニンおよびセリンなどの自由度の高いアミノ酸を含む）。いくつかの実施形態では、リンカーは G G G S (配列番号 2 2) または G G G G S (配列番号 2 3) 、たとえば、G G G S (配列番号 2 2) または G G G G S (配列番号 2 3) ユニットのうちの 2 、 3 、 4 つ、またはそれ以上の反復ユニットからなる 1 つ以上のユニットを含む。同様に他のリンカー配列も使用できる。

【 0 0 2 9 】

また本明細書は、単なる共局在化ではなく二量体化が効率的なゲノム編集の活性に必要である R N A 誘導型 F o k I ヌクレアーゼプラットフォームを記載する。これらのヌクレアーゼは、ヒトの細胞で効率の高いゲノム編集を強く仲介することができ、高感度ディープシーケンシング法により判定して、検出できないレベルまでオフターゲット変異を減少させることができる。また、5' 末端のヌクレオチドを備える g R N A 対を発現するために有効な系であって、R F N プラットフォーム上のより広い標的化範囲を与える方法を記載する。最後に、単量体の C a s 9 ニッカーゼは、一般的に、単一 g R N A が存在する中、本明細書中に記載されるヌクレアーゼよりもよりも多く望ましくない挿入欠失および点変異を導入する。これらの結果は、良好に特徴付けられた二量体構築物および対形成した C a s 9 ニッカーゼと比較した変異誘発プロファイルの改善、最も可能性のあるゲノム編集の正確さを必要とする研究または治療の適用に重要な特徴といった、特異性の利点を備えた、強力であり、取扱いが簡単なヌクレアーゼプラットフォームを定義する。

【 0 0 3 0 】

したがって、本明細書中で、ヒトの細胞で強力かつ特異性の高いゲノム編集を行うための新規の R N A 誘導型 F o k I ヌクレアーゼ (R F N) プラットフォームが記載される。R F N は、二量体としての活性と機能のために二つの g R N A を必要とする。驚くべきことに、活性 R F N の遺伝子操作は、F o k I ドメインが遺伝子操作した亜鉛フィンガーまたは転写アクチベータ 様エフェクター反復アレイのカルボキシ末端に融合した Z e n および T A L E N と異なる構築物である、d C a s 9 タンパク質のアミノ末端への F o k I

10

20

30

40

50

ヌクレアーゼドメインの融合を必要とした。また R F N は、各 F o k - d C a s 9 / g R N A 複合体により結合された片側部位が、14 ~ 17 b p の長さの相対的に限定された介在スペーサーを備えた特定の相対配向 (P A M 外部) を有することを必要とする。(活性は、追加的なスペーシングで可能であり得るが、ほとんど一貫して成功しない)。

【 0 0 3 1 】

R F N の二量体の性質は、標準的な単量体 C a s 9 ヌクレアーゼに対して重要な特異性に関する利点を提供する。理想的な二量体系では、活性は、片側部位で、モノマーではほとんど観察されない。提示されるデータから、単一 g R N A により配向される F o k I - d C a s 9 は、R F N 片側部位で変異誘発をほとんど誘導しないか、または全く誘導しないことが例証される。12つの単一 g R N A (6つのR F N 標的部位に対する)を、共発現した F o k I - d C a s 9 を用いて試験し、挿入欠失が、非常に低い頻度 (0.0045 % ~ 0.47 % の範囲)、場合によっては、g R N A またはヌクレアーゼの発現がない対照細胞で観察されるバックグラウンドの比率と同程度の低いレベルで観察された。二量体として F o k I ヌクレアーゼドメインが機能することを考慮すると、単一 g R N A で観察されるいずれかの挿入欠失が、D N A への F o k I - d C a s 9 ダイマーの動員によるものである可能性が高いことが推定される。機構に関わらず、非常に低いレベルの変異誘発のみが、F o k I - d C a s 9 を 12 のオフターゲット片側部位での単一 g R N A を用いて試験する際に観察されることを考慮すると、何等かの変異誘発が、部分的にミスマッチのオフターゲット片側部位で誘導される可能性が非常に少ない。実際に、V E G F A を標的とする R F N は、ディープシーケンシングにより判定されるように、g R N A のうちの 1 つの知られているオフターゲット部位で検出可能な変異を誘導しなかった。10

【 0 0 3 2 】

R F N は真の二量体系であるので、R F N は、共局在化に依存するが二量体化を必要するものではない対形成するニッカーゼの技術に勝って、多くの重要な利点を有している。第 1 に、本明細書中の直接的な比較から、単一の C a s 9 ニッカーゼは、一般的に、同一の個々の g R N A により配向される F o k I - d C a s 9 融合タンパク質よりも挿入欠失変異を効率良く導入することが示される。第 2 に、単量体の C a s 9 ニッカーゼはまた、標的片側部位に塩基対の置換を、この試験で包有されていない従来では知られていない変異原性の副作用を伴って、高い効率で誘導する場合がある。ここでも、直接的な比較から、単量体の C a s 9 ニッカーゼが、同一の単一 g R N A により誘導される F o k I - d C a s 9 融合物よりも実質的に高い比率で点変異を誘導することが示される。第 3 に、対形成した C a s 9 ニッカーゼは、二量体 R F N よりも標的片側部位の配向およびスペーシングにおいて高い無差別性を示し、それゆえオフターゲット変異が誘導され得る可能性がより高くなる範囲の部位を有する。対形成したニッカーゼの片側部位は、P A M 内部または P A M 外部、および、0 ~ 1 0 0 0 b p の範囲の長さのスペーサー配列を含んで配向できる (R a n e t a l . , C e l l 1 5 4 , 1 3 8 0 - 1 3 8 9 (2 0 1 3) ; M a l i e t a l . , N a t B i o t e c h n o l 3 1 , 8 3 3 - 8 3 8 (2 0 1 3) ; C h o e t a l . , G e n o m e R e s (2 0 1 3)) 。この無差別性は、C a s 9 ニッカーゼのゲノム編集活性が、酵素の二量体化に依存するものではなく、むしろ 2 つのニックの単なる共局在に依存するために存在する。対照的に、R F N は、特異性において、より厳密性が高い。片側部位は、有効な切断のための 2 つのおおよそ配置された F o k I 切断ドメインの必要条件のため、P A M 外部を有していないければならず、14 ~ 17 b p 離れてはいけなければならない。30

【 0 0 3 3 】

F o k I

F o k I は、D N A 認識ドメインおよび触媒 (エンドヌクレアーゼ) ドメインを含む 2 型制限エンドヌクレアーゼである。本明細書中に記載される融合タンパク質は、F o k I 全体、または触媒エンドヌクレアーゼドメイン、たとえば、G e n B a n k アクセシジョン番号第 A A A 2 4 9 2 7 . 1 のアミノ酸 3 8 8 ~ 5 8 3 または 4 0 8 ~ 5 8 3 (たとえば L i e t a l . , N u c l e i c A c i d s R e s . 3 9 (1) : 3)40

5 9 - 3 7 2 (2 0 1 1) ; Cathomen and Joun g , Mol . Ther . 16 : 1 2 0 0 - 1 2 0 7 (2 0 0 8) に記載) 、または Miller et al . Nat Biotechnol 25 : 778 - 785 (2 0 0 7) ; Szczeppek et al . , Nat Biotechnol 25 : 786 - 793 (2 0 0 7) ; もしくは Bitinaite et al . , Proc . Natl . Acad . Sci . USA . 95 : 1 0 5 7 0 - 1 0 5 7 5 (1 9 9 8) に記載されるような、Fok I の変異形態を含むことができる。

【 0 0 3 4 】

Fok I の例示的なアミノ酸配列は、以下の通りである。

【 化 1 】

10

10	20	30	40	50	60
MFLSMVSKIR	TFGWVQNP GK	FENLKR VVQ V	FDRNSKVHNE	VKNIKIPTLV	KESKIQKELV
70	80	90	100	110	120
AIMNQHD LIY	TYKELVGT GT	SIRSEAPCDA	IIQATIADQG	NKKGYIDNWS	SDGFLRWAHA
130	140	150	160	170	180
LGFIEYINKS	DSFVITDVGL	AYSKSADGSA	IEKEILIEAI	SSYPPAIRIL	TLLEDGQH LT
190	200	210	220	230	240
KFDLGKNL GF	SGESGFTSL P	EGLLDT LAN	AMPDKGEIR	NNWEGSSDKY	ARMIGGWLDK
250	260	270	280	290	300
LGLVKQGKKE	FIIPTLGKPD	NKEFISHAFK	ITGEGLKV LR	RAKGSTKFTR	VPKRVYWEML
310	320	330	340	350	360
ATNLTDKEYV	RTRRALILEI	LIKAGSLKIE	QI QDNLKKLG	FDEV ETIEN	DIKGLINTGI
370	380	390	400	410	420
FIEIKGRFYQ	LKDHLQFVI	PNRGVT KQLV	KSELEEK KSE	LRHKLKYVPH	EYIELIEIAR
430	440	450	460	470	480
NSTQDRILEM	KVMEFFMKVY	GYRGKHLGG S	RKPDGAIYTV	GSPIDYGVIV	DTKAYSGGY N
490	500	510	520	530	540
LPIGQA DEMQ	RYVEENQTRN	KHINPNEW WK	VYPSSVTEFK	FLFVSGHF KG	NYKAQLTRLN
550	560	570	580		
HITNCNGAVL	SVEELLIG GE	MIKAGTLTLE	EVRRKFNN GE	INF	(配列番号 4)

20

【 0 0 3 5 】

Fok I をコードする例示的な核酸配列は、以下の通りである。

30

【化2】

ATGTTTTGAGTATGGTTCTAAAATAAGAACCTTCGGTGGGTCAAAATCCAGGTAAA
 TTTGAGAATTAAACGAGTAGTCAAGTATTGATAGAAATTCAAAGTACATAATGAA
 GTGAAAAATATAAAGATACCAACCCAGTCAAAGAAAGTAAGATCCAAAAGAACTAGTT
 GCTATTATGAATCAACATGATTGATTTACATATAAAGAGTTAGTAGGAACAGGAACCT
 TCAATACGTTCAGAACGACCAGCGATGCAATTATTCAAGCAACAATAGCAGATCAAGGA
 AATAAAAAGGCTATATCGATAATTGGTCATCTGACGGTTTTGCGTTGGGCACATGCT
 TTAGGATTATTGAATATATAAAGTGAATTCTTGTAATAACTGATGTTGGACTT
 GCTTACTCTAAATCAGCTGACGCCAGCGCATTGAAAAAGAGATTGATTGAAGCGATA
 TCATCTTATCCTCCAGCGATTCGTTACCTTGCTAGAAGATGGACACATTTGACA
 AAGTTGATCTTGGCAAGAATTAGGTTAGTGGAGAAAGTGGATTACTTCTACCG
 GAAGGAATTCTTTAGATACTCTAGCTATGCTATGCCTAAAGATAAAGGCAGAAATTGCT
 AATAATTGGGAAGGATCTCAGATAAGTACGCAAGAATGATAGGTGGCTGGATAAA
 CTAGGATTAGTAAAGCAAGGAAAAAGAATTATCATTCCCTACTTGGGTAAGCCGGAC
 AATAAAAGAGTTATATCCCACGCTTTAAAATTACTGGAGAAGGTTGAAAGTACTGCGT
 CGAGCAAAAGGCTCTACAAAATTACACGTGTACCTAAAGAGTATATTGGGAAATGCTT
 GCTACAAACCTAACCGATAAAGAGTATGTAAGAACAAAGAGCTTGATTTAGAAATA
 TTAATCAAAGCTGGATCATTTAAAGAACAAATACAAGACAACCTGAAAGAAATTAGGA
 TTGATGAAGTTATAGAAACTATTGAAAATGATATCAAAGGCTTAATTACACAGGTATA
 TTATAGAAATCAAAGGGCGATTTATCAATTGAAAGACCATACTTCAATTGTAATA
 CCTAATCGTGGTGTGACTAAGCAACTAGTCAAAAGTGAACCTGGAGGAGAAGAAATCTGAA
 CTTCGTCATAAATTGAAATATGTGCCTCATGAATATATTGAAATTGAAATTGCCAGA

10

AATTCCACTCAGGATAGAATTCTTGAAATGAAGGTAATGGAATTGGGTTATGAAAGTTAT
 GGATATAGAGGTAACACATTGGGTTGATCAAGGAAACCGGACGGAGCAATTATACTGTC
 GGATCTCCTATTGATTACGGTGTGATCGGGATACTAAAGCTTATAGCGGAGGTTATAAT
 CTGCCAATTGCCAACGCAGATGAAATGCAACGATATGTCGAAGAAAATCAAACACGAAAC
 AAACATATCAACCTAATGAATGGTGGAAAGTCTATCCATCTCTGTAAACGGAATTAAAG
 TTTTATTGAGTGGTCACTTAAAGGAAACTACAAAGCTCAGCTACACGATTAAT
 CATATCACTAATTGTAATGGAGCTGTTCTAGTGTAGAAGAGCTTTAATTGGTGGAGAA
 ATGATTAAGCCGGCACATTAACCTTAGAGGAAGTGAACCGGAAATTAAACGGCGAG
 ATAAACTTTAA (配列番号5)

20

【0036】

いくつかの実施形態では、本明細書中で使用されるFok Iヌクレアーゼは、配列番号4、たとえば、配列番号4のアミノ酸388～583、または408～583と少なくとも約50%同一である。これらの変異体ヌクレアーゼは、DNAを切断する特性を保持したままでなければならない。いくつかの実施形態では、ヌクレオチド配列は、配列番号4のアミノ酸383～583、または408～583と約50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、99%または100%同一である。いくつかの実施形態では、配列番号4のアミノ酸388～583、または408～583とのいづの差異も、非保存領域にある。

30

【0037】

2つの配列のパーセント同一性を決定するには、最適比較目的で配列を整列させる（最適な整列に必要であれば第一および第二のアミノ酸または核酸配列の一方または両方にギャップを導入し、また非相同配列を比較目的のため無視することができる）。比較目的で整列させる参照配列の長さは少なくとも50%である（いくつかの実施形態では、参照配列の長さの約50%、55%、60%、65%、70%、75%、85%、90%、95%または100%を整列させる）。次いで、対応する位置のヌクレオチドまたは残基を比較する。第一の配列のある位置が、第二の配列の対応する位置と同じヌクレオチドまたは残基によって占められていれば、2つの分子はその位置において同一である。2つの配列間のパーセント同一性は、2つの配列の最適な整列のために導入する必要があるギャップの数および各ギャップの長さを考慮に入れた、2つの配列に共通する同一の位置の数の関数となる。

40

【0038】

50

配列の比較および2つの配列間のパーセント同一性の決定は、数学的アルゴリズムを用いて実施することができる。本願の目的には、GCGソフトウェアパッケージのGAPプログラムに組み込まれているNeedlemanおよびWunsch((1970)J.Mol.Biol.48:444-453)のアルゴリズムを使用し、ギャップペナルティが12、ギャップ伸長ペナルティが4、フレームシフトギャップペナルティが5のBlossum62スコア行列を用いて、2つのアミノ酸配列間のパーセント同一性を決定する。

【0039】

Cas9

いくつかの細菌がCas9タンパク質変異体を発現する。現在、化膿性連鎖球菌(*Streptococcus pyogenes*)Cas9が最もよく用いられているが、他のCas9タンパク質にも化膿性連鎖球菌(*S.pyogenes*)Cas9と高レベルの配列同一性を有し、同じガイドRNAを利用するものがある。それ以外のものはさらに多様であり、利用するgRNAが異なり、認識するPAM配列(RNAによって定められる配列に隣接するタンパク質によって定められる、2~5のヌクレオチド配列)も異なる。*Chylinski*らは多数の細菌群のCas9タンパク質を分類し(RNA Biology 10:5, 1-12; 2013)、その付図1および付表1に多数のCas9タンパク質を列記しており、これらは参考により本明細書に組み込まれる。ほかのCas9タンパク質については、*Esveldt*ら, *Nat Methods*. 2013 Nov; 10(11):1116-21および*Fonfara*ら, "Phylogeny of Cas9 determines functional exchangeability of dual-RNA and Cas9 among orthologous type II CRISPR-Cas systems." *Nucleic Acids Res.* 2013 Nov 22. [印刷前電子出版] doi:10.1093/nar/gkt1074に記載されている。

【0040】

本明細書に記載の方法および組成物には様々な種のCas9分子を用いることができる。化膿性連鎖球菌(*S.pyogenes*)およびサーモフィルス菌(*S. thermophilus*)のCas9分子が本開示の大部分の対象となるが、ほかにも、本明細書に列記する他の種のCas9タンパク質に由来するか、これに基づくCas9分子を用いることができる。換言すれば、本記載の大部分が化膿性連鎖球菌(*S.pyogenes*)およびサーモフィルス菌(*S. thermophilus*)のCas9分子を用いるものであるが、他の種のCas9分子をその代わりに用いることができる。このような種としては、*Chylinski*, 2013の付図に基づいて作成した以下の表に記載される種が挙げられる。

【0041】

【表1】

代替のCas9タンパク質	
GenBank アクセション番号	細菌
303229466	ベイロネラ・アティピカ (<i>Veillonella atypica</i>) ACS-134-V-Col7a
34762592	フソバクテリウム・ヌクレアタム亜種ビンセンティ (<i>Fusobacterium nucleatum</i> subsp. <i>vincentii</i>)
374307738	フィリファクター・アロシス (<i>Filifactor alocis</i>) ATCC35896
320528778	ソロバクテリウム・ムーレイ (<i>Solobacterium moorei</i>) F0204
291520705	コブロコッカス・カツス (<i>Coprococcus catus</i>) GD-7
42525843	トレボネーマ・デンティコラ (<i>Treponema denticola</i>) ATCC35405
304438954	ペプトニフィラス・デュアデニイ (<i>Peptoniphilus duerdenii</i>) ATCC BAA-1640
224543312	カテニバクテリウム・ミツオカイ (<i>Catenibacterium mitsuokai</i>) DSM15897
24379809	ストレプトコッカス・ミュータンス (<i>Streptococcus mutans</i>) UA159
15675041	化膿性連鎖球菌 (<i>Streptococcus pyogenes</i>) SF370
16801805	リステリア・イノキュア (<i>Listeria innocua</i>) Clip11262
116628213	サーモフィルス菌 (<i>Streptococcus thermophilus</i>) LMD-9
323463801	スタフィロコッカス・シュードインターメディウス (<i>Staphylococcus pseudintermedius</i>) ED99
352684361	アシダミノコッカス・インテスティニ (<i>Acidaminococcus intestini</i>) RyC-MR95
302336020	オルゼナ・ウリ (<i>Olsenella uli</i>) DSM7084
366983953	オエノコッカス・キタハラエ (<i>Oenococcus kitaharae</i>) DSM17330
310286728	ビフィドバクテリウム・ビフィダム (<i>Bifidobacterium bifidum</i>) S17
258509199	ラクトバチルス・ラムノサス (<i>Lactobacillus rhamnosus</i>) GG
300361537	ラクトバチルス・ガセリ (<i>Lactobacillus gasseri</i>) JV-V03
169823755	フィネゴルディア・マグナ (<i>Finegoldia magna</i>) ATCC29328
47458868	マイコプラズマ・モーピレ (<i>Mycoplasma mobile</i>) 163K
284931710	マイコプラズマ・ガリセプチカム (<i>Mycoplasma gallisepticum</i>) F株
363542550	マイコプラズマ・オビニュー・モニエ (<i>Mycoplasma ovipneumoniae</i>) SC01
384393286	マイコプラズマ・カニス (<i>Mycoplasma canis</i>) PG14
71894592	マイコプラズマ・シノビエ (<i>Mycoplasma synoviae</i>) 53
238924075	ユーバクテリウム・レクターレ (<i>Eubacterium rectale</i>) ATCC33656
116627542	サーモフィルス菌 (<i>Streptococcus thermophilus</i>) LMD-9
315149830	フェカリス菌 (<i>Enterococcus faecalis</i>) TX0012
315659848	スタフィロコッカス・ルグドゥネンシス (<i>Staphylococcus lugdunensis</i>) M23590
160915782	ユーバクテリウム・ドリカム (<i>Eubacterium dolichum</i>) DSM3991
336393381	ラクトバチルス・コリニフォルミス亜種トルケンス (<i>Lactobacillus coryniformis</i> subsp. <i>torquens</i>)
310780384	イリオバクター・ポリトロpus (<i>Ilyobacter polytropus</i>) DSM2926
325677756	ルミノコッカス・アルブス (<i>Ruminococcus albus</i>) 8
187736489	アッカーマンシア・ムシニフィラ (<i>Akkermansia muciniphila</i>) ATCCBAA-835

10

20

30

40

117929158	アシドサーマス・セルロリティカス (<i>Acidothermus cellulolyticus</i>) 11B	
189440764	ビフィドバクテリウム・ロンガム (<i>Bifidobacterium longum</i>) DJ010A	
283456135	ビフィドバクテリウム・デンティウム (<i>Bifidobacterium dentium</i>) Bd1	
38232678	ジフテリア菌 (<i>Corynebacterium diphtheriae</i>) NCTC13129	
187250660	エルシミクロビウム・ミヌトゥム (<i>Elusimicrobium minutum</i>) Pei191	
319957206	ニトラティフラクター・サルスギニス (<i>Nitratirfractor salsuginis</i>) DSM16511	
325972003	スピロヘータ・グロバス (<i>Sphaerochaeta globus</i>) Buddy株	10
261414553	フィブロバクター・サクシノゲネス亜種サクシノゲネス (<i>Fibrobacter succinogenes</i> subsp. <i>succinogenes</i>)	
60683389	バクテロイデス・フラギリス (<i>Bacteroides fragilis</i>) NCTC9343	
256819408	カブノサイトファーガ・オクラセア (<i>Capnocytophaga ochracea</i>) DSM7271	
90425961	ロドシュードモナス・パルストリス (<i>Rhodopseudomonas palustris</i>) BisB18	
373501184	プレボテラ・ミカンス (<i>Prevotella micans</i>) F0438	
294674019	プレボテラ・ルミニコラ (<i>Prevotella ruminicola</i>) 23	
365959402	フラボバクテリウム・カラムナリ (<i>Flavobacterium columnare</i>) ATCC49512	20
312879015	アミノモナス・パウシボランス (<i>Aminomonas paucivorans</i>) DSM12260	
83591793	ロドスピリルム・ルブルム (<i>Rhodospirillum rubrum</i>) ATCC11170	
294086111	Candidatus プニセイスピリルム・マリヌム (<i>Puniceispirillum marinum</i>) IMCC1322	
121608211	バーミンフロバクター・エイセニア (<i>Verminephrobacter eiseniae</i>) EF01-2	
344171927	ラルストニア・シジギイ (<i>Ralstonia syzygii</i>) R24	
159042956	ディノロセオバクター・シバエ (<i>Dinoroseobacter shibae</i>) DFL12	
288957741	アゾスピリルム (<i>Azospirillum</i>) 菌種-B510	
92109262	ニトロバクター・ハンブルゲンシス (<i>Nitrobacter hamburgensis</i>) X14	30
148255343	ブラディリゾビウム (<i>Bradyrhizobium</i>) 菌種-BTAi1	
34557790	ウォリネラ・サクシノゲネス (<i>Wolinella succinogenes</i>) DSM1740	
218563121	カンピロバクター・ジェジュニ亜種ジェジュニ (<i>Campylobacter jejuni</i> subsp. <i>jejuni</i>)	
291276265	ヘリコバクター・ムステラエ (<i>Helicobacter mustelae</i>) 12198	
229113166	セレウス菌 (<i>Bacillus cereus</i>) Rock1-15	
222109285	アシドボラックス・エブレウス (<i>Acidovorax ebreus</i>) TPSY	
189485225	未培養テルミテ (<i>Termite</i>) グループ1	
182624245	ウェルシュ菌 (<i>Clostridium perfringens</i>) D株	
220930482	クロストリジウム・セルロリティカム (<i>Clostridium cellulolyticum</i>) H10	40
154250555	パルビバクラム・ラバメンティボランス (<i>Parvibaculum lavamentivorans</i>) DS-1	
257413184	ロゼブリア・インテスティナリス (<i>Roseburia intestinalis</i>) L1-82	
218767588	膿膜炎菌 (<i>Neisseria meningitidis</i>) Z2491	
15602992	パスツレラ・ムルトシダ亜種ムルトシダ (<i>Pasteurella multocida</i> subsp. <i>multocida</i>)	

319941583	ステレラ・ワズワーセンシス(<i>Sutterella wadsworthensis</i>)31
254447899	ガンマプロテオバクテリアHTCC5015
54296138	レジオネラ・ニューモフィラ(<i>Legionella pneumophila</i>)Paris株
331001027	パラステラ・エクスクレメンティホミニス(<i>Parasutterella exrementihominis</i>)YIT11859
34557932	ウォリネラ・サクシノゲネス(<i>Wolinella succinogenes</i>)DSM1740
118497352	フランシセラ・ノビシダ(<i>Francisella novicida</i>)U112

【0042】

10

本明細書に記載の構築物および方法は上に挙げたいずれかの Cas9 タンパク質およびその対応するガイド RNA またはこれに相当する他のガイド RNA の使用を含み得る。このほか、サーモフィルス菌 (*Streptococcus thermophilus*) LMD-9 の Cas9 CRISPR1 系がヒト細胞で機能することが Cong らに示されている (Science 339, 819 (2013))。髄膜炎菌 (*N. meningitidis*) Cas9 オルソログが Houら, Proc Natl Acad Sci U S A. 2013 Sep 24; 110(39): 15644-9 および Esveltら, Nat Methods. 2013 Nov; 10(11): 1116-21 に記載されている。さらに、Jinek らは、サーモフィルス菌 (*S. thermophilus*) および L. イノキュア (*L. innocua*) の Cas9 オルソログ (異なるガイド RNA を利用するとと思われる髄膜炎菌 (*N. meningitidis*) および ジェジュニ菌 (*C. jejuni*) の Cas9 オルソログではない) が、わずかに効率が低下するものの、化膿性連鎖球菌 (*S. pyogenes*) 二重 gRNA によって誘導されて標的プラスマド DNA を切断し得ることを *in vitro* で明らかにしている。

【0043】

いくつかの実施形態では、本発明の系は、細菌にコードされるか、哺乳動物細胞での発現にコドン最適化され、D10、E762、H983 または D986 および H840 または N863 の変異、例えば、D10A / D10N および H840A / H840N / H840Y を含む化膿性連鎖球菌 (*S. pyogenes*) Cas9 タンパク質を用いてタンパク質のヌクレアーゼ部分の触媒能を不活性化し；これらの位置における置換は、(Nishimatsu, Cell 156, 935-949 (2014)) に記載されているように) アラニンであってよく、また他の残基、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 E762Q、H983N、H983Y、D986N、N863D、N863S または N863H であってよい (図 1C)。本明細書に記載の方法および組成物に使用することができる触媒能が不活性化された化膿性連鎖球菌 (*S. pyogenes*) Cas9 は以下の通りであり、例示的な D10A および H840A の変異は太字で表され、下線が施されている。

20

【0044】

30

【表2】

10	20	30	40	50	60	
NDKKYSIGLA	IGTNSVGWAV	ITDEYKVPSK	KFKVLGNTDR	HSIKKNLIGA	LLFDSEGETAE	
70	80	90	100	110	120	
ATRLKRTARR	RYTRRKNRIC	YLQEIFSNEM	AKVDDSFHHR	LEESFLVEED	KKHERHPIFG	
130	140	150	160	170	180	
NIVDEVAYME	KYPTIYHLRK	KLVDSTDKAD	LRLIYLALAH	MIKFRGMFLI	EGDLNPDNSD	
190	200	210	220	230	240	10
VDKLFQOLVQ	TYNQLFEEENP	INASGVDAKA	IISARLSKSR	RLENLIAOLP	GEKKNGLFGN	
250	260	270	280	290	300	
LIALSLGLTP	NFKSNFQLDAAE	DAKLQLSKDT	YDDDDLNLLA	QIGDQYADLF	LAAKNLSDAI	
310	320	330	340	350	360	
LLSDILRVNT	EITKAPLSAS	MIKRYDEHHQ	DLTLLKALVR	QQLPPEKYKEI	FFDQSKNGYA	
370	380	390	400	410	420	
GYIDGGASQE	EFYKFIKPIL	EKMDGTEELL	VKLNREDLLR	KQRTFDNGSI	PHQIHLGELH	
430	440	450	460	470	480	20
AILRRQEDFY	PFLKDNREKI	EKILTFRIPY	YVGPLARGNS	RFANMTRKSE	ETITPWNFEE	
490	500	510	520	530	540	
VVDKGASAQS	FIERMTNFDK	NLPNEKVLPK	HSLLYEYFTV	YNELTKVKYV	TEGMRKPAFL	
550	560	570	580	590	600	
SGEQKKAIVD	LLFKTNRKT	VKQLKEDYFK	KIECFDSVEI	SGVEDRFNAS	LGTYHDLKII	
610	620	630	640	650	660	
IKDKDFLDNE	ENEDILEDIV	LTTLTFEDRE	MIEERLKTYA	HLFDDKVMKQ	LKRRRYTGWG	
670	680	690	700	710	720	
RLSRKLINGI	RDKQSGKTTIL	DFLKSDGFAN	RNFMQLIHDD	SLTFKEDIQK	AQVSGQGDSL	30
730	740	750	760	770	780	
HEHIANLAGS	PAIKKGILQT	VKVVDDELVKV	MGRHKPENIV	IEMARENQTT	QKGQKNSRER	
790	800	810	820	830	840	
MKRIEEGIKE	LGSQILKEHP	VENTQLQNEK	LYLYYLQNGR	DMYVDQELDI	NRLSDYDVDA	
850	860	870	880	890	900	
IVPQSFLKDD	SIDNKVLTRS	DKNRGKSDNV	PSSEVVKKMK	NYWRQLLNAK	LITQRKFDSL	
910	920	930	940	950	960	
TKAERGGLSE	LDKAGFIKRQ	LVETRQITKH	VAQILDLSRMN	TKYDENDKLI	REVKVITLKS	
970	980	990	1000	1010	1020	
KLVSDFRKDF	QFYKVREINN	YHHAHDAYLN	AVVGTALIKK	YPKLESEFVY	GDYKVYDVRK	
1030	1040	1050	1060	1070	1080	
MIAKSEQEIG	KATAKYFFYS	NIMNFFKTEI	TLANGEIRKR	PLIETNGETG	EIVWDKGRDF	
1090	1100	1110	1120	1130	1140	
ATVRKVL SMP	QVNIVKKTEV	Q7GGFSKESI	LPKRNSDKLI	ARKKDWDPKK	YGGFDSPTVA	

1150	1160	1170	1180	1190	1200
YSVLVVAKVE	KGKSKKLKV	KELLGITIME	RSSFEKNPID	FLEAKGYKEV	KKDLIIKLPK
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	IDRKRYTSTK	EVLDATLIHQ	SITGLYETRI	DLSQLGGD	(配列番号5)

10

【0045】

いくつかの実施形態では、本明細書で使用される Cas 9 ヌクレアーゼは、化膿性連鎖球菌 (S. pyogenes) Cas 9 の配列と少なくとも約 50 % 同一である、すなわち、配列番号 5 と少なくとも 50 % 同一である。いくつかの実施形態では、ヌクレオチド配列は配列番号 5 と約 50 %、55 %、60 %、65 %、70 %、75 %、80 %、85 %、90 %、95 %、99 % または 100 % 同一である。いくつかの実施形態では、配列番号 5 との差はいずれも非保存領域内にあり、これは Chylinski ら, RNA Biology 10 : 5, 1 - 12; 2013 (例えば、その付図 1 および付表 1); Esvelt ら, Nat Methods. 2013 Nov; 10(11): 1116 - 21 および Fonfara ら, Nucl. Acids Res. (2014) 42(4): 2577 - 2590. [印刷前電子出版 2013 Nov 22] doi: 10.1093/nar/gkt1074 に記載されている配列の配列アライメントによって確認される。同一性を上述のように決定する。

20

【0046】

ガイド RNA (gRNA)

一般的に言えば、ガイド RNA には 2 つの系、すなわち、一緒に機能して Cas 9 による切断を誘導する別個の crRNA と tracrRNA を用いる系 1 および 2 つの別個のガイド RNA を单一の系に組み合わせるキメラ crRNA - tracrRNA ハイブリッドを用いる系 2 (単一ガイド RNA または sgRNA と呼ばれる。このほか、Jinek ら, Science 2012; 337: 816 - 821 を参照されたい) がある。tracrRNA は様々な長さに短縮することが可能であり、様々な長さのものが別個の系 (系 1) およびキメラ gRNA 系 (系 2) の両方で機能することが示されている。例えば、いくつかの実施形態では、tracrRNA は、その 3' 末端から少なくとも 1 nt、2 nt、3 nt、4 nt、5 nt、6 nt、7 nt、8 nt、9 nt、10 nt、15 nt、20 nt、25 nt、30 nt、35 nt または 40 nt 短縮されていてよい。いくつかの実施形態では、tracrRNA 分子は、その 5' 末端から少なくとも 1 nt、2 nt、3 nt、4 nt、5 nt、6 nt、7 nt、8 nt、9 nt、10 nt、15 nt または 20 nt、25 nt、30 nt、35 nt または 40 nt 短縮されていてよい。あるいは、tracrRNA 分子は、5' 末端および 3' 末端の両方から、例えば、5' 末端で少なくとも 1 nt、2 nt、3 nt、4 nt、5 nt、6 nt、7 nt、8 nt、9 nt、10 nt、15 nt または 20 nt、3' 末端で少なくとも 1 nt、2 nt、3 nt、4 nt、5 nt、6 nt、7 nt、8 nt、9 nt、10 nt、15 nt または 20 nt、30 nt、35 nt または 40 nt 短縮されていてよい。例えば、Jinek ら, Science 2012; 337: 816 - 821; Mali ら, Science. 2013 Feb 15; 339(6121): 823 - 6; Cong ら, Science. 2013 Feb 15; 339(6121): 819 - 23; ならびに Hwang ら, Nature 2013; 500(7467): 472 - 5; ならびに Doudna ら, Cell 2014; 157(7): 1344 - 56; ならびに Charpentier ら, Cell 2014; 157(7): 1346 - 57; ならびに Zhang ら, Cell 2014; 157(7): 1347 - 58; ならびに Zhou ら, Cell 2014; 157(7): 1349 - 60; ならびに Gootz ら, Cell 2014; 157(7): 1361 - 71; ならびに Liu ら, Cell 2014; 157(7): 1372 - 82; ならびに Li ら, Cell 2014; 157(7): 1383 - 93; ならびに Li ら, Cell 2014; 157(7): 1394 - 404; ならびに Li ら, Cell 2014; 157(7): 1405 - 16; ならびに Li ら, Cell 2014; 157(7): 1417 - 28; ならびに Li ら, Cell 2014; 157(7): 1429 - 40; ならびに Li ら, Cell 2014; 157(7): 1441 - 51; ならびに Li ら, Cell 2014; 157(7): 1453 - 63; ならびに Li ら, Cell 2014; 157(7): 1465 - 75; ならびに Li ら, Cell 2014; 157(7): 1477 - 87; ならびに Li ら, Cell 2014; 157(7): 1489 - 99; ならびに Li ら, Cell 2014; 157(7): 1491 - 501; ならびに Li ら, Cell 2014; 157(7): 1493 - 511; ならびに Li ら, Cell 2014; 157(7): 1495 - 513; ならびに Li ら, Cell 2014; 157(7): 1497 - 515; ならびに Li ら, Cell 2014; 157(7): 1499 - 517; ならびに Li ら, Cell 2014; 157(7): 1501 - 519; ならびに Li ら, Cell 2014; 157(7): 1503 - 521; ならびに Li ら, Cell 2014; 157(7): 1505 - 523; ならびに Li ら, Cell 2014; 157(7): 1507 - 525; ならびに Li ら, Cell 2014; 157(7): 1509 - 527; ならびに Li ら, Cell 2014; 157(7): 1511 - 529; ならびに Li ら, Cell 2014; 157(7): 1513 - 531; ならびに Li ら, Cell 2014; 157(7): 1515 - 533; ならびに Li ら, Cell 2014; 157(7): 1517 - 535; ならびに Li ら, Cell 2014; 157(7): 1519 - 537; ならびに Li ら, Cell 2014; 157(7): 1521 - 539; ならびに Li ら, Cell 2014; 157(7): 1523 - 541; ならびに Li ら, Cell 2014; 157(7): 1525 - 543; ならびに Li ら, Cell 2014; 157(7): 1527 - 545; ならびに Li ら, Cell 2014; 157(7): 1529 - 547; ならびに Li ら, Cell 2014; 157(7): 1531 - 549; ならびに Li ら, Cell 2014; 157(7): 1533 - 551; ならびに Li ら, Cell 2014; 157(7): 1535 - 553; ならびに Li ら, Cell 2014; 157(7): 1537 - 555; ならびに Li ら, Cell 2014; 157(7): 1539 - 557; ならびに Li ら, Cell 2014; 157(7): 1541 - 559; ならびに Li ら, Cell 2014; 157(7): 1543 - 561; ならびに Li ら, Cell 2014; 157(7): 1545 - 563; ならびに Li ら, Cell 2014; 157(7): 1547 - 565; ならびに Li ら, Cell 2014; 157(7): 1549 - 567; ならびに Li ら, Cell 2014; 157(7): 1551 - 569; ならびに Li ら, Cell 2014; 157(7): 1553 - 571; ならびに Li ら, Cell 2014; 157(7): 1555 - 573; ならびに Li ら, Cell 2014; 157(7): 1557 - 575; ならびに Li ら, Cell 2014; 157(7): 1559 - 577; ならびに Li ら, Cell 2014; 157(7): 1561 - 579; ならびに Li ら, Cell 2014; 157(7): 1563 - 581; ならびに Li ら, Cell 2014; 157(7): 1565 - 583; ならびに Li ら, Cell 2014; 157(7): 1567 - 585; ならびに Li ら, Cell 2014; 157(7): 1569 - 587; ならびに Li ら, Cell 2014; 157(7): 1571 - 589; ならびに Li ら, Cell 2014; 157(7): 1573 - 591; ならびに Li ら, Cell 2014; 157(7): 1575 - 593; ならびに Li ら, Cell 2014; 157(7): 1577 - 595; ならびに Li ら, Cell 2014; 157(7): 1579 - 597; ならびに Li ら, Cell 2014; 157(7): 1581 - 599; ならびに Li ら, Cell 2014; 157(7): 1583 - 601; ならびに Li ら, Cell 2014; 157(7): 1585 - 603; ならびに Li ら, Cell 2014; 157(7): 1587 - 605; ならびに Li ら, Cell 2014; 157(7): 1589 - 607; ならびに Li ら, Cell 2014; 157(7): 1591 - 609; ならびに Li ら, Cell 2014; 157(7): 1593 - 611; ならびに Li ら, Cell 2014; 157(7): 1595 - 613; ならびに Li ら, Cell 2014; 157(7): 1597 - 615; ならびに Li ら, Cell 2014; 157(7): 1599 - 617; ならびに Li ら, Cell 2014; 157(7): 1601 - 619; ならびに Li ら, Cell 2014; 157(7): 1603 - 621; ならびに Li ら, Cell 2014; 157(7): 1605 - 623; ならびに Li ら, Cell 2014; 157(7): 1607 - 625; ならびに Li ら, Cell 2014; 157(7): 1609 - 627; ならびに Li ら, Cell 2014; 157(7): 1611 - 629; ならびに Li ら, Cell 2014; 157(7): 1613 - 631; ならびに Li ら, Cell 2014; 157(7): 1615 - 633; ならびに Li ら, Cell 2014; 157(7): 1617 - 635; ならびに Li ら, Cell 2014; 157(7): 1619 - 637; ならびに Li ら, Cell 2014; 157(7): 1621 - 639; ならびに Li ら, Cell 2014; 157(7): 1623 - 641; ならびに Li ら, Cell 2014; 157(7): 1625 - 643; ならびに Li ら, Cell 2014; 157(7): 1627 - 645; ならびに Li ら, Cell 2014; 157(7): 1629 - 647; ならびに Li ら, Cell 2014; 157(7): 1631 - 649; ならびに Li ら, Cell 2014; 157(7): 1633 - 651; ならびに Li ら, Cell 2014; 157(7): 1635 - 653; ならびに Li ら, Cell 2014; 157(7): 1637 - 655; ならびに Li ら, Cell 2014; 157(7): 1639 - 657; ならびに Li ら, Cell 2014; 157(7): 1641 - 659; ならびに Li ら, Cell 2014; 157(7): 1643 - 661; ならびに Li ら, Cell 2014; 157(7): 1645 - 663; ならびに Li ら, Cell 2014; 157(7): 1647 - 665; ならびに Li ら, Cell 2014; 157(7): 1649 - 667; ならびに Li ら, Cell 2014; 157(7): 1651 - 669; ならびに Li ら, Cell 2014; 157(7): 1653 - 671; ならびに Li ら, Cell 2014; 157(7): 1655 - 673; ならびに Li ら, Cell 2014; 157(7): 1657 - 675; ならびに Li ら, Cell 2014; 157(7): 1659 - 677; ならびに Li ら, Cell 2014; 157(7): 1661 - 679; ならびに Li ら, Cell 2014; 157(7): 1663 - 681; ならびに Li ら, Cell 2014; 157(7): 1665 - 683; ならびに Li ら, Cell 2014; 157(7): 1667 - 685; ならびに Li ら, Cell 2014; 157(7): 1669 - 687; ならびに Li ら, Cell 2014; 157(7): 1671 - 689; ならびに Li ら, Cell 2014; 157(7): 1673 - 691; ならびに Li ら, Cell 2014; 157(7): 1675 - 693; ならびに Li ら, Cell 2014; 157(7): 1677 - 695; ならびに Li ら, Cell 2014; 157(7): 1679 - 697; ならびに Li ら, Cell 2014; 157(7): 1681 - 699; ならびに Li ら, Cell 2014; 157(7): 1683 - 701; ならびに Li ら, Cell 2014; 157(7): 1685 - 703; ならびに Li ら, Cell 2014; 157(7): 1687 - 705; ならびに Li ら, Cell 2014; 157(7): 1689 - 707; ならびに Li ら, Cell 2014; 157(7): 1691 - 709; ならびに Li ら, Cell 2014; 157(7): 1693 - 711; ならびに Li ら, Cell 2014; 157(7): 1695 - 713; ならびに Li ら, Cell 2014; 157(7): 1697 - 715; ならびに Li ら, Cell 2014; 157(7): 1699 - 717; ならびに Li ら, Cell 2014; 157(7): 1701 - 719; ならびに Li ら, Cell 2014; 157(7): 1703 - 721; ならびに Li ら, Cell 2014; 157(7): 1705 - 723; ならびに Li ら, Cell 2014; 157(7): 1707 - 725; ならびに Li ら, Cell 2014; 157(7): 1709 - 727; ならびに Li ら, Cell 2014; 157(7): 1711 - 729; ならびに Li ら, Cell 2014; 157(7): 1713 - 731; ならびに Li ら, Cell 2014; 157(7): 1715 - 733; ならびに Li ら, Cell 2014; 157(7): 1717 - 735; ならびに Li ら, Cell 2014; 157(7): 1719 - 737; ならびに Li ら, Cell 2014; 157(7): 1721 - 739; ならびに Li ら, Cell 2014; 157(7): 1723 - 741; ならびに Li ら, Cell 2014; 157(7): 1725 - 743; ならびに Li ら, Cell 2014; 157(7): 1727 - 745; ならびに Li ら, Cell 2014; 157(7): 1729 - 747; ならびに Li ら, Cell 2014; 157(7): 1731 - 749; ならびに Li ら, Cell 2014; 157(7): 1733 - 751; ならびに Li ら, Cell 2014; 157(7): 1735 - 753; ならびに Li ら, Cell 2014; 157(7): 1737 - 755; ならびに Li ら, Cell 2014; 157(7): 1739 - 757; ならびに Li ら, Cell 2014; 157(7): 1741 - 759; ならびに Li ら, Cell 2014; 157(7): 1743 - 761; ならびに Li ら, Cell 2014; 157(7): 1745 - 763; ならびに Li ら, Cell 2014; 157(7): 1747 - 765; ならびに Li ら, Cell 2014; 157(7): 1749 - 767; ならびに Li ら, Cell 2014; 157(7): 1751 - 769; ならびに Li ら, Cell 2014; 157(7): 1753 - 771; ならびに Li ら, Cell 2014; 157(7): 1755 - 773; ならびに Li ら, Cell 2014; 157(7): 1757 - 775; ならびに Li ら, Cell 2014; 157(7): 1759 - 777; ならびに Li ら, Cell 2014; 157(7): 1761 - 779; ならびに Li ら, Cell 2014; 157(7): 1763 - 781; ならびに Li ら, Cell 2014; 157(7): 1765 - 783; ならびに Li ら, Cell 2014; 157(7): 1767 - 785; ならびに Li ら, Cell 2014; 157(7): 1769 - 787; ならびに Li ら, Cell 2014; 157(7): 1771 - 789; ならびに Li ら, Cell 2014; 157(7): 1773 - 791; ならびに Li ら, Cell 2014; 157(7): 1775 - 793; ならびに Li ら, Cell 2014; 157(7): 1777 - 795; ならびに Li ら, Cell 2014; 157(7): 1779 - 797; ならびに Li ら, Cell 2014; 157(7): 1781 - 799; ならびに Li ら, Cell 2014; 157(7): 1783 - 801; ならびに Li ら, Cell 2014; 157(7): 1785 - 803; ならびに Li ら, Cell 2014; 157(7): 1787 - 805; ならびに Li ら, Cell 2014; 157(7): 1789 - 807; ならびに Li ら, Cell 2014; 157(7): 1791 - 809; ならびに Li ら, Cell 2014; 157(7): 1793 - 811; ならびに Li ら, Cell 2014; 157(7): 1795 - 813; ならびに Li ら, Cell 2014; 157(7): 1797 - 815; ならびに Li ら, Cell 2014; 157(7): 1799 - 817; ならびに Li ら, Cell 2014; 157(7): 1801 - 819; ならびに Li ら, Cell 2014; 157(7): 1803 - 821; ならびに Li ら, Cell 2014; 157(7): 1805 - 823; ならびに Li ら, Cell 2014; 157(7): 1807 - 825; ならびに Li ら, Cell 2014; 157(7): 1809 - 827; ならびに Li ら, Cell 2014; 157(7): 1811 - 829; ならびに Li ら, Cell 2014; 157(7): 1813 - 831; ならびに Li ら, Cell 2014; 157(7): 1815 - 833; ならびに Li ら, Cell 2014; 157(7): 1817 - 835; ならびに Li ら, Cell 2014; 157(7): 1819 - 837; ならびに Li ら, Cell 2014; 157(7): 1821 - 839; ならびに Li ら, Cell 2014; 157(7): 1823 - 841; ならびに Li ら, Cell 2014; 157(7): 1825 - 843; ならびに Li ら, Cell 2014; 157(7): 1827 - 845; ならびに Li ら, Cell 2014; 157(7): 1829 - 847; ならびに Li ら, Cell 2014; 157(7): 1831 - 849; ならびに Li ら, Cell 2014; 157(7): 1833 - 851; ならびに Li ら, Cell 2014; 157(7): 1835 - 853; ならびに Li ら, Cell 2014; 157(7): 1837 - 855; ならびに Li ら, Cell 2014; 157(7): 1839 - 857; ならびに Li ら, Cell 2014; 157(7): 1841 - 859; ならびに Li ら, Cell 2014; 157(7): 1843 - 861; ならびに Li ら, Cell 2014; 157(7): 1845 - 863; ならびに Li ら, Cell 2014; 157(7): 1847 - 865; ならびに Li ら, Cell 2014; 157(7): 1849 - 867; ならびに Li ら, Cell 2014; 157(7): 1851 - 869; ならびに Li ら, Cell 2014; 157(7): 1853 - 871; ならびに Li ら, Cell 2014; 157(7): 1855 - 873; ならびに Li ら, Cell 2014; 157(7): 1857 - 875; ならびに Li ら, Cell 2014; 157(7): 1859 - 877; ならびに Li ら, Cell 2014; 157(7): 1861 - 879; ならびに Li ら, Cell 2014; 157(7): 1863 - 881; ならびに Li ら, Cell 2014; 157(7): 1865 - 883; ならびに Li ら, Cell 2014; 157(7): 1867 - 885; ならびに Li ら, Cell 2014; 157(7): 1869 - 887; ならびに Li ら, Cell 2014; 157(7): 1871 - 889; ならびに Li ら, Cell 2014; 157(7): 1873 - 891; ならびに Li ら, Cell 2014; 157(7): 1875 - 893; ならびに Li ら, Cell 2014; 157(7): 1877 - 895; ならびに Li ら, Cell 2014; 157(7): 1879 - 897; ならびに Li ら, Cell 2014; 157(7): 1881 - 899; ならびに Li ら, Cell 2014; 157(7): 1883 - 901; ならびに Li ら, Cell 2014; 157(7): 1885 - 903; ならびに Li ら, Cell 2014; 157(7): 1887 - 905; ならびに Li ら, Cell 2014; 157(7): 1889 - 907; ならびに Li ら, Cell 2014; 157(7): 1891 - 909; ならびに Li ら, Cell 2014; 157(7): 1893 - 911; ならびに Li ら, Cell 2014; 157(7): 1895 - 913; ならびに Li ら, Cell 2014; 157(7): 1897 - 915; ならびに Li ら, Cell 2014; 157(7): 1899 - 917; ならびに Li ら, Cell 2014; 157(7): 1901 - 919; ならびに Li ら, Cell 2014; 157(7): 1903 - 921; ならびに Li ら, Cell 2014; 157(7): 1905 - 923; ならびに Li ら, Cell 2014; 157(7): 1907 - 925; ならびに Li ら, Cell 2014; 157(7): 1909 - 927; ならびに Li ら, Cell 2014; 157(7): 1911 - 929; ならびに Li ら, Cell 2014; 157(7): 1913 - 931; ならびに Li ら, Cell 2014; 157(7): 1915 - 933; ならびに Li ら, Cell 2014; 157(7): 1917 - 935; ならびに Li ら, Cell 2014; 157(7): 1919 - 937; ならびに Li ら, Cell 2014; 157(7): 1921 - 939; ならびに Li ら, Cell 2014; 157(7): 1923 - 941; ならびに Li ら, Cell 2014; 157(7): 1925 - 943; ならびに Li ら, Cell 2014; 157(7): 1927 - 945; ならびに Li ら, Cell 2014; 157(7): 1929 - 947; ならびに Li ら, Cell 2014; 157(7): 1931 - 949; ならびに Li ら, Cell 2014; 157(7): 1933 - 951; ならびに Li ら, Cell 2014; 157(7): 1935 - 953; ならびに Li ら, Cell 2014; 157(7): 1937 - 955; ならびに Li ら, Cell 2014; 157(7): 1939 - 957; ならびに Li ら, Cell 2014; 157(7): 1941 - 959; ならびに Li ら, Cell 2014; 157(7): 1943 - 961; ならびに Li ら, Cell 2014; 157(7): 1945 - 963; ならびに Li ら, Cell 2014; 157(7): 1947 - 965; ならびに Li ら, Cell 2014; 157(7): 1949 - 967; ならびに Li ら, Cell 2014; 157(7): 1951 - 969; ならびに Li ら, Cell 2014; 157(7): 1953 - 971; ならびに Li ら, Cell 2014; 157(7): 1955 - 973; ならびに Li ら, Cell 2014; 157(7): 1957 - 975; ならびに Li ら, Cell 2014; 157(7): 1959 - 977; ならびに Li ら, Cell 2014; 157(7): 1961 - 979; ならびに Li ら, Cell 2014; 157(7): 1963 - 981; ならびに Li ら, Cell 2014; 157(7): 1965 - 983; ならびに Li ら, Cell 2014; 157(7): 1967 - 985; ならびに Li ら, Cell 2014; 157(7): 1969 - 987; ならびに Li ら, Cell 2014; 157(7): 1971 - 989; ならびに Li ら, Cell 2014; 157(7): 1973 - 991; ならびに Li ら, Cell 2014; 157(7): 1975 - 993; ならびに Li ら, Cell 2014; 157(7): 1977 - 995; ならびに Li ら, Cell 2014; 157(7): 1979 - 997; ならびに Li ら, Cell 2014; 157(7): 1981 - 999; ならびに Li ら, Cell 2014; 157(7): 1983 - 1001; ならびに Li ら, Cell 2014; 157(7): 1985 - 1003; ならびに Li ら, Cell 2014; 157(7): 1987 - 1005; ならびに Li ら, Cell 2014; 157(7): 1989 - 1007; ならびに Li ら, Cell 2014; 157(7): 1991 - 1009; ならびに Li ら, Cell 2014; 157(7): 1993 - 1011; ならびに Li ら, Cell 2014; 157(7): 1995 - 1013; ならびに Li ら, Cell 2014; 157(7): 1997 - 1015; ならびに Li ら, Cell 2014; 157(7): 1999 - 1017; ならびに Li ら, Cell 2014; 157(7): 2001 - 1019; ならびに Li ら, Cell 2014; 157(7): 2003 - 1021; ならびに Li ら, Cell 2014; 157(7): 2005 - 1023; ならびに Li ら, Cell 2014; 157(7): 2007 - 1025; ならびに Li ら, Cell 2014; 157(7): 2009 - 1027; ならびに Li ら, Cell 2014; 157(7): 2011 - 1029; ならびに Li ら, Cell 2014; 157(7): 2013 - 1031; ならびに Li ら, Cell 2014; 157(7): 2015 - 1033; ならびに Li ら, Cell 2014; 157(7): 2017 - 1035; ならびに Li ら, Cell 2014; 157

および Furukawa, Nat Biotechnol. 2013 Mar; 31(3):227-9; Jinekura, Elife 2, e00471 (2013) を参照されたい。系2では一般に、キメラgRNAの長さが長いほどオンターゲット活性が高いことがわかっているが、様々な長さのgRNAの相対的特異性は現時点では明らかにされておらず、したがって、場合によっては短いgRNAを用いる方が望ましいことがある。いくつかの実施形態では、gRNAは、転写開始部位の上流約100~800bp以内、例えば、転写開始部位の上流約500bp以内、または転写開始部位の下流約100~800bp以内、例えば、約500bp以内にある領域に相補的である。いくつかの実施形態では、2種類以上のgRNAをコードするベクター（例えば、プラスミド）、例えば、標的遺伝子の同じ領域の異なる部位を対象とする2種類、3種類、4種類、5種類またはそれ以上のgRNAをコードするプラスミドを用いる。
10

【0047】

Cas9ヌクレアーゼは、たとえば、ゲノムDNA標的部位の相補鎖に相補的な5'末端で17~20ntを有する単一gRNAまたはtracrRNA/crRNAといった、ガイドRNAを使用して、配列NGGといった追加的なプロトスペーサー隣接モチーフ(PAM)を有する、特異的な17~20ntのゲノム標的に誘導することができる。したがって、本発明の方法は、たとえば、Mali et al., Science 2013 Feb 15; 339(6121):823-6に記載されるような単一のCas9ガイドRNAといった、標準的なトランスコード化tracrRNAに融合されるcrRNAを含む単一のガイドRNAの使用を含むことができる。これは、たとえば、NGG、NAG、またはNNGGといったプロトスペーサー隣接モチーフ(PAM)の5'末端標的配列に相補的な鎖のうち、20nt、19nt、18nt、または17nt、好ましくは17ntまたは18ntといった、25~17、任意に20以下のヌクレオチド(nt)の標的配列に相補的な5'末端での配列を備える。いくつかの実施形態では、単一のCas9誘導型RNAは、配列
20

【化3】

(X₁₇₋₂₀)GUUUUAGAGCUAGAAAAGCAAGUUAAAAUAAGGCUAGUC
CG(X_N) (配列番号6);

(X₁₇₋₂₀)GUUUUAGAGCUAUGCUGAAAAGCAUAGCAAGUUAAAAUAAGGC
GUCCGUUAUC(X_N) (配列番号7);

(X₁₇₋₂₀)GUUUUAGAGCUAUGCUGUUUUGGAAACAAAACAGCAUAGCAAGU
UAAAAUAAGGCUAGUCGUUAUC(X_N) (配列番号8);

(X₁₇₋₂₀)GUUUUAGAGCUAGAAAAGCAAGUUAAAAUAAGGCUAGUCGUU
AUCAACUUGAAAAAGUGGCACCGAGUCGGUGC(X_N) (配列番号9);

(X₁₇₋₂₀)GUUUUAGAGCUAGAAAAGCAAGUUAAAAUAAGGCUAGUCGUU
AUCAACUUGAAAAAGUGGCACCGAGUCGGUGC (配列番号10);
40

(X₁₇₋₂₀)GUUUUAGAGCUAUGCUGAAAAGCAUAGCAAGUUAAAAUAAGG
CUAGUCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC (配列番号11);
または

(X₁₇₋₂₀)GUUUUAGAGCUAUGCUGAAAAGCAUAGCAAGUUAAAAUAAGG
CUAGUCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC (配列番号12)

であって、式中X_{17~20}が、標的配列の17~20個の連続したヌクレオチドに相補
50

的なヌクレオチド配列である、配列からなる。単一ガイドRNAをコードするDNAは、以前に文献で記載されている（Jinek et al., Science. 337 (6096) : 816 - 21 (2012) および Jinek et al., Elife. 2 : e00471 (2013)）。

【0048】

ガイドRNAは、リボ核酸とCas9との結合に干渉しない任意の配列であり得るX_Nを含んでよく、(RNA中の)Nは0～200、例えば、0～100、0～50または0～20であり得る。

【0049】

いくつかの実施形態では、ガイドRNAは、3'末端に1つまたは複数のアデニン(A)またはウラシル(U)ヌクレオチドを含む。いくつかの実施形態では、RNAPol II転写を停止させる終止シグナルとして用いられる1つまたは複数のTが任意選択で存在する結果として、RNAは、分子の3'末端に1つまたは複数のU、例えば、1～8つまたはそれ以上のU(例えば、U、UU、UUU、UUUU、UUUUU、UUUUUUU、UUUUUUUU、UUUUUUUUU)を含む。

【0050】

本明細書中に記載されるいくつかの例は単一gRNAを利用するが、本方法は、二重のgRNA(たとえば天然に存在する系で見出されるcrrRNAおよびtracrRNA)を用いても使用できる。この場合、単一のtracrRNA、本発明の系を使用して発現する複数の異なるcrrRNA、たとえば以下

(X₁₇ 20) GUUUUAGAGCUA(配列番号：13)；
(X₁₇ 20) GUUUUAGAGCUAUGCUGUUUG(配列番号：14)；
または

(X₁₇ 20) GUUUUAGAGCUAUGCUC(配列番号：15)
およびtracrRNA配列と共に使用される。この場合、crrRNAは、本明細書に記載される方法および分子におけるガイドRNAとして使用され、tracrRNAは、同一または異なるRNA分子から発現できる。いくつかの実施形態では、本方法は、配列【化4】

GGAACCAUUCAAAACAGCAUAGCAAGUUAAAAUAAGGUAGUCCGUUA
UCAACUUGAAAAAGUGGCACCGAGUCGGUGC(配列番号16)

またはその活性部位(活性部位は、Cas9またはdCas9との複合体を形成する特性を保持する部位である)を含む、またはからなるtracrRNAと細胞を接触させることを含む。いくつかの実施形態では、tracrRNA分子は、少なくとも1nt、2nt、3nt、4nt、5nt、6nt、7nt、8nt、9nt、10nt、15nt、20nt、25nt、30nt、35nt、または40ntだけ、3'末端から短縮されてもよい。別の実施形態では、tracrRNA分子は、少なくとも1nt、2nt、3nt、4nt、5nt、6nt、7nt、8nt、9nt、10nt、15nt、20nt、25nt、30nt、35nt、または40ntだけ、5'末端から短縮されてもよい。あるいは、tracrRNA分子は、たとえば5'末端上で少なくとも1nt、2nt、3nt、4nt、5nt、6nt、7nt、8nt、9nt、10nt、15nt、または20nt分、および3'末端上で少なくとも1nt、2nt、3nt、4nt、5nt、6nt、7nt、8nt、9nt、10nt、15nt、20nt、25nt、30nt、35nt、または40nt分、5'末端および3'末端の両方から短縮されてもよい。配列番号8に加えて例示的なtracrRNA配列は、以下を含む：

【化5】

UAGCAAGUUAAAAUAAGGUAGCUAGUCCGUUAUCAACUUGAAAAAGUGGCA
CCGAGUCGGUGC (配列番号17) もしくはその活性部分、または

AGCAUAGCAAGUUAAAAUAAGGUAGCUAGUCCGUUAUCAACUUGAAAAAGU
GGCACCGAGUCGGUGC (配列番号18) もしくはその活性部分。

【0051】

(X₁₇ - ₂₀) G U U U U A G A G C U A U G C U G U U U U G (配列番号：14)
をcRNAとして使用するいくつかの実施形態では、以下のt r a c r RNAを使用する
：

【化6】

GGAACCAUUCAAAACAGCAUAGCAAGUUAAAAUAAGGUAGCUAGUCCGUUA
UCAACUUGAAAAAGUGGCACCGAGUCGGUGC (配列番号16) またはその活性部分。

(X₁₇ - ₂₀) G U U U U A G A G C U A (配列番号：13) をc r RNAとして使用
するいくつかの実施形態では、以下のt r a c r RNAを使用する：

【化7】

UAGCAAGUUAAAAUAAGGUAGCUAGUCCGUUAUCAACUUGAAAAAGUGGCA
CCGAGUCGGUGC (配列番号17) またはその活性部分。

(X₁₇ - ₂₀) G U U U U A G A G C U A U G C U (配列番号：15) をc r RNAと
して使用するいくつかの実施形態では、以下のt r a c r RNAを使用する：

【化8】

AGCAUAGCAAGUUAAAAUAAGGUAGCUAGUCCGUUAUCAACUUGAAAAAGU
GGCACCGAGUCGGUGC (配列番号18) またはその活性部分。

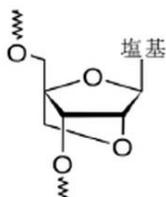
【0052】

いくつかの実施形態では、gRNAは、オフターゲット作用を最小限にするために、ゲノムの残りのいずれかの配列と異なる少なくとも3つ以上のミスマッチである部位を標的とする。

【0053】

ロックト核酸(LNA)などの修飾RNAオリゴヌクレオチドが、修飾オリゴヌクレオチドをより好ましい(安定な)コンホメーションにロックすることによってRNA-DNAハイブリダイゼーションの特異性を増大させることが示されている。例えば、2'-O-メチルRNAは2'酸素と4'炭素の間に追加の共有結合が存在する修飾塩基であり、これをオリゴヌクレオチド内に組み込むことによって全体の熱安定性および選択性を改善することができる(式I)。

【化9】



式I ロックト核酸

【0054】

10

したがって、いくつかの実施形態では、本明細書に開示される tru-gRNA は、1つまたは複数の修飾 RNA オリゴヌクレオチドを含み得る。例えば、本明細書に記載の短縮ガイド RNA 分子は、標的配列に相補的なガイド RNA の 17 ~ 18 nt または 17 ~ 19 nt の 5' 領域のうちの1つ、一部または全部が修飾されていてよく、例えば、ロックされていてよく (2' - O - 4' - C メチレン架橋) 、5' - メチルシチジンであってよく、2' - O - メチル - プソイドウリジンであってよく、あるいはリボースリン酸骨格がポリアミド鎖に置き換わっていてよく (ペプチド核酸) 、例えば、合成リボ核酸であってよい。

【0055】

20

他の実施形態では、tru-gRNA 配列の1つ、一部または全部のヌクレオチドが修飾されていてよく、例えば、ロックされていてよく (2' - O - 4' - C メチレン架橋) 、5' - メチルシチジンであってよく、2' - O - メチル - プソイドウリジンであってよく、あるいはリボースリン酸骨格がポリアミド鎖に置き換わっていてよく (ペプチド核酸) 、例えば、合成リボ核酸であってよい。

【0056】

いくつかの実施形態では、単一ガイド RNA および / または crrNA および / または tracrRNA は、3' 末端に1つまたは複数のアデニン (A) またはウラシル (U) ヌクレオチドを含み得る。

【0057】

30

既存の Cas9 ベースの RGN は、目的とするゲノム部位への標的化を誘導するのに gRNA - DNA ヘテロ二本鎖の形成を利用するものである。しかし、RNA - DNA ヘテロ二本鎖ではその DNA - DNA 対応物よりも無差別な範囲の構造が形成され得る。実際、DNA - DNA 二本鎖の方がミスマッチに対する感度が高く、このことは、DNA 誘導型ヌクレアーゼがオフターゲット配列と容易には結合せず、RNA 誘導型ヌクレアーゼに比して特異性が高いことを示唆している。したがって、本明細書に記載の方法に使用可能なガイド RNA はハイブリッド、すなわち、1つまたは複数のデオキシリボヌクレオチド、例えば短い DNA オリゴヌクレオチドが、gRNA の全部または一部、例えば gRNA の相補性領域の全部または一部と置き換わったハイブリッドであり得る。この DNA ベースの分子は、単一 gRNA 系の全部または一部と置き換わってもよく、あるいは二重 crrRNA / tracrRNA 系の crrRNA および / または tracrRNA の全部または一部と置き換わってもよい。一般に DNA - DNA 二本鎖の方が RNA - DNA 二本鎖よりもミスマッチに対する許容性が低いことから、相補性領域内に DNA を組み込むこのような系の方がより高い信頼性で目的とするゲノム DNA 配列を標的とするはずである。このような二本鎖を作製する方法は当該技術分野で公知であり、例えば、Barkerら, BMC Genomics. 2005 Apr 22; 6: 57; および Sugimotoら, Biochemistry. 2000 Sep 19; 39 (37): 11270 - 81 を参照されたい。

40

【0058】

さらに、別個の crrRNA および tracrRNA を使用する系では、1つまたは両方を合成とすることができます、1つ以上の修飾 (たとえばロック) ヌクレオチドまたはデオ

50

キシリボヌクレオチドを含むことができる。

【0059】

細胞との関連において、Cas9と合成gRNAの複合体を使用して、CRISPR/Cas9ヌクレアーゼ系のゲノム全域にわたる特異性を改善することが可能である。

【0060】

記載される方法は、本明細書中に記載されるCas9 gRNA + 融合タンパク質を、細胞中で発現すること、または細胞と接触させることを含むことができる。

【0061】

発現系

記載される融合タンパク質を使用するためには、それをコードする核酸から発現させるのが望ましいであろう。これは様々な方法で実施することができる。例えば、ガイドRNAをコードする核酸を中間ベクターにクローニングして原核細胞または真核細胞を形質転換し、複製および／または発現させる。中間ベクターは通常、融合タンパク質をコードする核酸を保管または操作して融合タンパク質を産生する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。このほか、融合タンパク質をコードする核酸を発現ベクターにクローニングして植物細胞、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、細菌細胞または原生動物細胞に投与してもよい。

【0062】

発現させるためには、融合タンパク質をコードする配列を通常、転写を指令するプロモーターを含む発現ベクターにサブクローニングする。適切な細菌プロモーターおよび真核プロモーターは当該技術分野で周知であり、例えば、Sambrookら, Molecular Cloning, A Laboratory Manual(第3版, 2001); Kriegler, Gene Transfer and Expression: A Laboratory Manual(1990); およびCurrent Protocols in Molecular Biology(Ausubelら編, 2010)に記載されている。設計したタンパク質を発現させるための細菌発現系を例えば、大腸菌(*E. coli*)、バチルス(*Bacillus*)菌種およびサルモネラ(*Salmonella*)で入手することができる(Palvaら, 1983, Gene 22: 229-235)。そのような発現系のキットが市販されている。哺乳動物細胞、酵母および昆虫細胞用の真核発現系は当該技術分野で周知であり、同じく市販されている。

【0063】

核酸の発現を指令するために使用するプロモーターは、具体的な用途によって決まる。例えば、融合タンパク質の発現および精製には通常、強力な構成的プロモーターを使用する。これに対して、遺伝子調節のためにガイドRNAをin vivoで投与する場合、ガイドRNAの具体的な用途に応じて構成的プロモーターまたは誘導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、ガイドRNAの投与に好ましいプロモーターは、HSV-TKなどの弱いプロモーターまたはこれと類似する活性を有するプロモーターであり得る。プロモーターはほかにも、トランス活性化に応答性の要素、例えば、低酸素応答要素、Gal4応答要素、laczリプレッサー応答要素ならびにテトラサイクリン調節系およびRU-486系などの小分子制御系を含み得る(例えば、GosseんおよびBujard, 1992, Proc. Natl. Acad. Sci. USA, 89: 5547; Oliginoら, 1998, Gene Ther., 5: 491-496; Wangら, 1997, Gene Ther., 4: 432-441; Neeringら, 1996, Blood, 88: 1147-55;ならびにRendahilら, 1998, Nat. Biotechnol., 16: 757-761を参照されたい)。

【0064】

発現ベクターは通常、プロモーターに加えて、原核または真核宿主細胞で核酸を発現するのに必要なほかのあらゆる要素を含む転写単位または発現力セットを含む。したがって、典型的な発現力セットは、例えばgRNAをコードする核酸配列と作動可能に連結され

10

20

30

40

50

たプロモーターと、例えば効率的な転写産物のポリアデニル化、転写停止、リボソーム結合部位または翻訳停止に必要な任意のシグナルとを含む。カセットのほかの要素としては、例えば、エンハンサーおよび異種スプライスイントロンシグナルを挙げ得る。

【0065】

細胞内に遺伝情報を輸送するのに使用する具体的な発現ベクターは、意図するgRNAの用途、例えば、植物、動物、細菌、真菌、原生動物などでの発現を考慮して選択する。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GSTおよびLacZなどの市販のタグ融合発現系が挙げられる。

【0066】

真核発現ベクターには真核ウイルスの調節要素を含む発現ベクター、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・バーウィルス由来のベクターを用いることが多い。その他の例示的な真核ベクターとしては、pMSG、pAV009/A+、pMT010/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、マウス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をはじめとする真核細胞での発現に効果を示すプロモーターの指令を受けてタンパク質を発現させる他の任意のベクターが挙げられる。

【0067】

ガイドRNA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 Pol III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、プラスミドトランスフェクション後に哺乳動物細胞にgRNAを発現させる。あるいは、例えばvitro転写にT7プロモーターを使用してもよく、RNAをin vitroで転写させてから精製することができる。短いRNA、例えば、siRNA、shRNAをはじめとする低分子RNAの発現に適したベクターを使用することができる。図4Bに記載のCys4ベースのマルチプレックス系を用いて、複数のgRNAを、単一の転写物に発現でき(RNA Pol IIIまたはPol IIIプロモーターにより駆動)、次いでより大きな転写物から切断され得る。

【0068】

一部の発現系は、安定にトランスフェクト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、ヒグロマイシン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などを有する。このほか、昆虫細胞にバキュロウイルスベクターを用いてポリヘドリンプロモーターをはじめとする強力なバキュロウイルスプロモーターの指令の下にgRNAコード配列を置くものなど、高収率の発現系が適している。

【0069】

発現ベクターに通常含まれる要素としてはほかにも、大腸菌(E.coli)で機能するレブリコン、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プラスミドの非必須領域にあり組換え配列の挿入を可能にする固有の制限部位が挙げられる。

【0070】

標準的なトランスフェクション法を用いて、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作製し、次いで標準的な技術を用いてこのタンパク質を精製する(例えば、Colleyら, 1989, J.Biol.Chem., 264:17619-22; Guide to Protein Purification, in Methods in Enzymology, vol. 182 (Deutscherら編, 1990)を参照されたい)。真核および原核細胞の形質転換を標準的な技術により実施する(例えば、Morrison, 1977, J.Bacteriol. 132:349-351; Clark-CurtissおよびCurtiss, Methods in Enzymology 101:347-362 (Wuら編, 1983)を参照されたい)。

10

20

30

40

50

【0071】

宿主細胞内に外来ヌクレオチド配列を導入するあらゆる既知の方法を用い得る。このような方法としては、リン酸カルシウムトランスフェクション、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、リポソーム、マイクロインジェクション、裸のDNA、プラスミドベクター、ウイルスベクター（エピソーム型および組込み型の両方）およびクローン化したゲノムDNA、cDNA、合成DNAをはじめとする外来遺伝物質を宿主細胞内に導入する他のあらゆる周知の方法の使用が挙げられる（例えば、Sambrookら、上記を参照されたい）。唯一必要なのは、用いる具体的な遺伝子工学的方法で、宿主細胞内にgRNAの発現が可能な遺伝子を少なくとも1つ良好に導入することができるることである。

10

【0072】

本発明は、ベクターおよびベクターを含む細胞を含む。

【実施例】

【0073】

本発明は以下の実施例でさらに説明されるが、これらの実施例は特許請求の範囲に記載される本発明の範囲を限定するものではない。

【0074】

実施例1.RNA誘導型エンドヌクレアーゼの特異性の評価

C R I S P R RNA誘導型ヌクレアーゼ（RGN）は、簡便で効率的なゲノム編集のプラットフォームとして急速に登場したものである。この実施例では、ヒト細胞ベースのレポーターアッセイを用いてCas9ベースのRGNのオフターゲット切断の特徴を明らかにすることについて記載する。

20

【0075】

材料および方法

実施例1では以下の材料および方法を用いた。

【0076】

ガイドRNAの構築

Cas9標的化のための可変20nt配列を保有するDNAオリゴヌクレオチドをアニールさせて、4bpオーバーハングを有しBsmBI消化プラスミドpMLM3636へのライゲーションに適合した短い二本鎖DNAフラグメントを作製した。このアニールしたオリゴヌクレオチドのクローン化により、U6プロモーターの発現下に20の可変5'ヌクレオチドを有するキメラ+103一本鎖ガイドRNAをコードするプラスミドが得られる（Hwangら、Nat Biotechnol 31, 227-229 (2013)；Maliら、Science 339, 823-826 (2013)）。この研究に使用するpMLM3636および発現プラスミドpJDS246（コドン最適化型のCas9をコードする）はともに非営利プラスミド配布サービスAddgene（addgene.org/crispr-cas）から入手可能である。

30

【0077】

EGFP活性アッセイ

EGFP-PEST融合遺伝子の單一コピーが組み込まれたU2OS. EGFP細胞を既に記載されている通りに培養した（Reyonら、Nat Biotech 30, 460-465 (2012)）。トランスフェクションでは、SE Cell Line 4D-Nucleofector（商標）Xキット（Lonza）を製造業者のプロトコルに従って用い、示される量のgRNA発現プラスミドおよびpJDS246をTd-Tマトコードプラスミド30ngとともに200,000個の細胞にヌクレオフェクトした。トランスフェクションの2日後、BD LSR IIフローサイトメータを用いて細胞を解析した。gRNA/Cas9プラスミドの濃度を最適化するトランスフェクションを三重反復で実施し、他のトランスフェクションをいずれも二重反復で実施した。

40

【0078】

内在ヒトゲノム部位のPCR増幅および配列検証

50

Phusion Hot Start II 高忠実度DNAポリメラーゼ(NEB)を用いてPCR反応を実施した。ほとんどの遺伝子座がタッチダウンPCR([98、10秒；72～62、-1／サイクル、15秒；72、30秒]10サイクル、[98、10秒；62、15秒；72、30秒]25サイクル)を用いて良好に増幅された。必要に応じて、68または72の一定のアニーリング温度および3%DMSOまたは1Mベタインを用いて残りの標的のPCRを35サイクル実施した。PCR産物をQIAxcelキャピラリー電気泳動系で分析して、その大きさおよび純度を検証した。妥当性が確認された産物をExoSap-IT(Affymetrix)で処理し、サンガーフラッシュ法(MGH DNA Sequencing Core)により配列決定して各標的部位を検証した。

10

【0079】

ヒト細胞におけるRGN誘発オントーゲットおよびオフターゲット変異の頻度の決定
U2OS.EGFP細胞およびK562細胞では、4D Nucleofector System(Lonza)を製造業者の説明書に従って用い、 2×10^5 個の細胞にgRNA発現プラスミドまたは空のU6プロモータープラスミド(陰性対照)250ng、Cas9発現プラスミド750ngおよびtd-Tomat発現プラスミド30ngをトランسفェクトした。HEK293細胞では、Lipofectamine LTX試薬(Life Technologies)を製造業者の指示に従って用い、 1.65×10^5 個の細胞にgRNA発現プラスミドまたは空のU6プロモータープラスミド(陰性対照)125ng、Cas9発現プラスミド375ngおよびtd-Tomat発現プラスミド30ngをトランسفェクトした。QIAamp DNA Blood Mini Kit(QIAGEN)を製造業者の説明書に従って用い、トランسفェクトしたU2OS.EGFP細胞、HEK293細胞またはK562細胞からゲノムDNAを回収した。オフターゲット候補部位を増幅するのに十分なゲノムDNAが得られるように、3回のスクレオフェクション(U2OS.EGFP細胞)、2回のスクレオフェクション(K562細胞)または2回のLipofectamine LTXトランسفェクションで得られたDNAをプールしてからT7EIを実施した。試験した各条件に対してこの操作を2回実施することにより同じゲノムDNAのプールを2つ作製し、各トランسفェクションを計4回または6回分得た。次いで、これらのゲノムDNAを鋳型に用いてPCRを上記の通りに実施し、Ampure XPビーズ(Agenceourt)を製造業者の説明書に従って用い精製した。T7EIアッセイを既に記載されている通りに実施した(Reyonaら, 2012, 上記)。

20

【0080】

NHEJ仲介性挿入欠失変異のDNA配列決定

T7EIアッセイに使用した精製PCR産物をZero Blunt TOPOベクター-(Life Technologies)にクローニングし、MGH DNA Automation Coreによりアルカリ溶解ミニプレップ法を用いてプラスミドDNAを単離した。M13順方向プライマー(5'-GTAAAAACGACGGCCAG-3'(配列番号19)を用いてサンガーフラッシュ法(MGH DNA Sequencing Core)によりプラスミドの配列を決定した。

30

【0081】

実施例1a. 単一スクレオチドミスマッチ

ヒト細胞におけるRGN特異性決定因子を明らかにすることを始めるにあたり、複数のgRNA/標的DNA接合部内の様々な位置に系統的にミスマッチを生じさせることの影響を評価するために大規模な試験を実施した。これを実施するため、既に記載されている標的スクレアーゼ活性の迅速の定量化が可能な定量的なヒト細胞ベースの高感度緑色蛍光タンパク質(EGFP)崩壊アッセイ(上の「方法」およびReyonaら, 2012, 上記を参照されたい)(図2B)を用いた。このアッセイでは、スクレアーゼ誘発二本鎖切断(DSB)の誤りがちな非相同末端結合(NHEJ)修復によって導入されるフレームシフト挿入/欠失(挿入欠失)変異を不活性化することによって起こるヒトU2OS.E

40

50

GFP細胞内の蛍光シグナルを評価することによって、単一の組み込まれたEGFPレポーター遺伝子を標的とするスクレアーゼの活性を定量化することができる（図2B）。ここに記載される研究では、EGFP内の異なる配列を標的とする以下のような3種類の約100ntの単一sgRNA（sgRNA）を用いた：

EGFP部位1 GGGCACGGCAGCTTGCCTGGTGG（配列番号1）

EGFP部位2 GATGCCGTTCTCTGCTTGTCTGG（配列番号2）

EGFP部位3 GGTGGTGCAGATGAACATTCAAGGG（配列番号3）。

上記sgRNAはそれぞれ、Cas9仲介性のEGFP発現崩壊を効率的に誘導することができる（実施例1eおよび2aならびに図3E（最上段）および3F（最上段）を参照されたい）。

10

【0082】

最初の実験では、3種類のEGFP標的化sgRNAの相補的標的化領域の20ヌクレオチドのうち19ヌクレオチドにおける單一ヌクレオチドミスマッチの影響を試験した。これを実施するため、3種類の標的部位のそれぞれについて、位置1～19（3'から5'の方向に1～20の番号を付した；図1を参照されたい）にワトソン-クリックトランスバージョンミスマッチを保有する変異体sgRNAを作製し、これらの各種sgRNAがヒト細胞においてCas9仲介性EGFP崩壊を誘導する能力を試験した（位置20のヌクレオチドはU6プロモーター配列の一部であり、発現への影響を避けるためグアニンでなければならないことから、この位置に置換のある変異体sgRNAは作製しなかった）。

20

【0083】

EGFP標的部位#2では、gRNAの3'末端よりも5'末端のミスマッチの方が許容性に高いことを示唆するこれまでの研究結果（Jiangら,Nat Biotech no1 31, 233-239(2013); Congら, Science 339, 819-823(2013); Jinsekら, Science 337, 816-821(2012)）の通り、gRNAの位置1～10に单一ミスマッチがあると、付随するCas9の活性に著しい影響がみられた（図2C、中央パネル）。しかし、EGFP標的部位#1および#3では、gRNAの一部を除くいすれの位置に单一ミスマッチがあっても、それが配列の3'末端内でも、高い許容性がみられた。さらに、上記2種類の標的にはミスマッチに対する感度が高い具体的な位置に差がみられた（図2C、最上段と最下段のパネルを比較されたい）。例えば、標的部位#1は特に位置2のミスマッチに対して高い感度を示したのに対して、標的部位#3は位置1および8のミスマッチに対して最も高い感度を示した。

30

【0084】

実施例1b. 複数のミスマッチ

gRNA/DNA接合部の2つ以上のミスマッチによる影響を試験するため、隣接する位置および離れた位置にワトソン-クリックトランスバージョンミスマッチを2つ有する一連の変異体sgRNAを作製し、EGFP崩壊アッセイを用いて、ヒト細胞でこれらのsgRNAがCas9スクレアーゼ活性を誘導する能力を試験した。全般的に標的部位は3種類とも、一方または両方のミスマッチがgRNA標的化領域の3'側半分に起こる2つの変化に対して感度が高くなることが分かった。しかし、この影響の大きさには部位による差がみられ、標的部位#2がこの2つのミスマッチに対して最も高い感度を示し、標的部位#1が全般的に最も低い感度を示した。許容され得る隣接したミスマッチの数を試験するため、gRNA標的化領域の5'末端の位置19～15（单一および2つのミスマッチの許容性が高いと思われる位置）の範囲でミスマッチの位置の数が漸増する変異体sgRNAを構築した。

40

【0085】

このようにミスマッチを漸増させたsgRNAを試験したところ、3種類の標的部位でいずれも、3つ以上の隣接するスマッチを導入することによってRGN活性が大幅に消失することが明らかになった。5'末端の位置19から開始し、3'末端に向かってミスマ

50

ツチを追加して漸増させていくと、3種類の異なるEGFP標的化gRNAの活性が突然低下した。具体的には、位置19および19+18にミスマッチを含むgRNAが実質的に完全な活性を示すのに対して、位置19+18+17、19+18+17+16および19+18+17+16+15にミスマッチのあるgRNAには陰性対照に比して実質的に差がみられなかった(図2F)。(位置20はgRNAの発現を駆動するU6プロモーターの一部であるためGでなければならぬことから、本発明者らは上記変異体gRNAの位置20にはミスマッチを生じさせなかつことに留意されたい。)

【0086】

gRNA相補性を短縮することにより特異性が増大したRGNを得ることができる根拠がほかにも以下の実験で得られた:4種類の異なるEGFP標的化gRNA(図2H)では、位置18および19に2つのミスマッチを導入しても活性にあまり影響を及ぼさなかつた。しかし、上記gRNAの位置10および11にさらに2つのミスマッチを導入すると活性がほぼ完全に消失する。10/11の2つのミスマッチのみを導入しても、全般的に活性にそれほど大きな影響を及ぼさないのは興味深い。

10

【0087】

以上をまとめると、ヒト細胞で得られたこれらの結果は、RGNの活性がgRNA標的化配列の3'側半分のミスマッチに対する方が高い感度を示し得ることを裏付けるものである。しかし、以上のデータはほかにも、RGN特異性が複雑で標的部位依存性であり、単一および2つのミスマッチであれば、RNA標的化領域の3'側半分に1つまたは複数のミスマッチが生じても高い許容性を示すことが多いことを明確に示している。さらに、以上のデータはほかにも、gRNA/DNA接合部の5'側半分のあらゆるミスマッチが必ずしも許容性が高いわけではないことを示唆している。

20

【0088】

さらに、以上の結果は、より短い領域の相補性(具体的には約17nt)を有するgRNAの方が活性の特性が高くなることを強く示唆している。本発明者らは、17ntの特異性と、PAM配列によって付与される2ntの特異性とを組み合わせることによって、ヒト細胞にみられるゲノムのような大型で複雑なゲノム内で固有なものとなるのに十分な長さの1つである19bp配列の仕様が得られることに注目する。

【0089】

実施例1c. オフターゲット変異

30

内在ヒト遺伝子を標的とするRGNのオフターゲット変異を特定することができるかどうかを明らかにするため、VEGFA遺伝子の3つの異なる部位、EMX1遺伝子の1つの部位、RNF2遺伝子の1つの部位およびFANCF遺伝子の1つの部位を標的とする6種類のsgRNAを用いた。上記6種類のsgRNAは、T7エンドヌクレアーゼI(T7EI)アッセイによって検出されたように、ヒトU2OS.EGFP細胞のそれぞれの内在遺伝子座におけるCas9仲介性挿入欠失を効率的に誘導するものであった(上の「方法」)。次いで本発明者らは、この6種のRGNそれぞれについて、U2OS.EGFP細胞におけるヌクレアーゼ誘発NHEJ仲介性挿入欠失変異の証拠を得るために、候補となるオフターゲット部位を数十か所(46から64に及ぶ箇所数)検討した。評価した遺伝子座には、ヌクレオチドが1つまたは2つ異なる全ゲノム部位のほかにも、ヌクレオチドが3~6つ異なるゲノム部位のサブセットを含め、gRNA標的化配列の5'側半分にこのようなミスマッチを1つまたは複数有するものを重視した。T7EIアッセイを用いて、VEGFA部位1では(検討した53か所の候補部位うち)4か所のオフターゲット部位、VEGFA部位2では(検討した46か所のうち)12か所、VEGFA部位3では(検討した64か所のうち)7か所、EMX1部位では(検討した46か所のうち)1か所が容易に特定された。RNF2またはFANCF遺伝子について検討したそれぞれ43か所および50か所の候補部位にはオフターゲット変異は検出されなかつた。実証されたオフターゲット部位の変異率は極めて高く、目的とする標的部位に観察された変異率の5.6%から12.5%(平均4.0%)に及ぶものであった。このような真のオフターゲットには、標的部位の3'末端にミスマッチを有し、合計5つに及ぶミスマッチを有する

40

50

配列が含まれ、ほとんどのオフターゲット部位がタンパク質をコードする遺伝子内にみられた。一部のオフターゲット部位のDNAシーケンシングから、予測されるRGN切断部位に挿入欠失変異が起こることを示す分子的な裏付けがさらに得られた(図8A~8C)。

【0090】

実施例1d. その他の細胞型のオフターゲット変異

RGNがU2OS・EGFP細胞に高頻度でオフターゲット変異を誘発し得ることが確認されたため、次に、これらのヌクレアーゼが他のタイプのヒト細胞にもこのような影響を及ぼすかどうかを明らかにしようとした。これらの細胞は、以前、TALEN^{1~5}の活性を評価するのにU2OS・EGFP細胞を用いたため、最初の実験にはU2OS・EGFP細胞を選択したが、標的化ヌクレアーゼの活性の試験にはヒトHEK293細胞およびK562細胞の方が広く用いられている。したがって、HEK293細胞およびK562細胞についてもVEGFA部位1、2および3ならびにEMX1部位を標的とする4種類のRGNの活性をも評価した。この4種類のそれぞれのRGNが、変異頻度はU2OS・EGFP細胞に観察された頻度よりもいくぶん低いものの、上記のさらなる2種類のヒト細胞系でもその目的とするオントーゲット部位にNHEJ仲介性挿入欠失変異を効率的に誘発した(T7E1アッセイによる評価)。最初にU2OS・EGFP細胞で特定された上記4種類のRGNの24か所のオフターゲット部位を評価したところ、多くの部位が、HEK293細胞およびK562細胞でも同様にその対応するオントーゲット部位と同程度の頻度で変異することが明らかになった。予想された通り、HEK293細胞のこれらのオフターゲット部位の一部のDNAシーケンシングにより、予測されたゲノム遺伝子座に変化が生じることを示す分子的な根拠がさらに得られた。U2OS・EGFP細胞で特定されたオフターゲット部位のうち、HEK293細胞の4か所、K562細胞の11か所が検出可能な変異を示さなかった理由は正確にはわからない。しかし、これらのオフターゲット部位の多くがU2OS・EGFP細胞でも比較的低い変異頻度を示したことが注目される。したがって、本発明者らの実験ではU2OS・EGFP細胞に比してHEK293細胞およびK562細胞の方が全般的にRGNの活性が低いと思われるため、HEK293細胞およびK562細胞のこれらの部位における変異率がT7E1アッセイの信頼できる検出限界(約2~5%)未満になり得る。以上をまとめると、HEK293細胞およびK562細胞で得た結果は、今回RGNに観察される高頻度のオフターゲット変異が複数のヒト細胞型にみられる一般的な現象である根拠を示すものである。

【0091】

実施例1e. EGFP崩壊アッセイに使用するgRNA発現およびCas9発現プラスミドの量の漸増

非相同末端結合を介したフレームシフト変異の誘発がEGFPの発現を確実に崩壊させ得る位置であるEGFPヌクレオチド502の上流に位置する3つの異なる配列(上に示したEGFP部位1~3)に対して単一ガイドRNA(sgRNA)を作製した(Mae der, M. L. ら, Mol Cell 31, 294-301 (2008); Rey o n, D. ら, Nat Biotech 30, 460-465 (2012))。

【0092】

3つの標的部位について、最初に様々な量のgRNA発現プラスミド(12.5~250ng)を、構成的に発現するEGFP-PESTレポーター遺伝子の單一コピーを有する本発明者らのU2OS・EGFPレポーター細胞に、コドン最適化型のCas9ヌクレアーゼ発現プラスミド750ngとともにトランスフェクトした。最高濃度のgRNAプラスミド(250ng)ではRGNが3種類とも効率的にEGFP発現を崩壊させた(図3E(上段))。しかし、これより少ない量のgRNA発現プラスミドをトランスフェクトした場合、標的部位#1および#3に対するRGNが同レベルの崩壊を示したのに対して、標的部位#2のRGN活性は、トランスフェクトするgRNA発現プラスミドの量を減らすと直ちに低下した(図3E(上段))。

【0093】

10

20

30

40

50

本発明者らの U 2 O S . E G F P レポーター細胞にトランスフェクトする Cas 9 コードプラスミドの量を漸増させて (50 ng ~ 750 ng) E G F P 崩壊をアッセイした。図 3 F (上段) に示されるように、標的部位 # 1 では、トランスフェクトする Cas 9 コードプラスミドの量を 3 分の 1 にしても、E G F P 崩壊活性が実質的に低下せずに許容された。しかし、標的部位 # 2 および # 3 を標的とする R G N の活性は、トランスフェクトする Cas 9 プラスミドの量を 3 分の 1 にすると直ちに低下した (図 3 F (上段))。以上の結果を踏まえて、実施例 1 a ~ 1 d に記載される実験には、E G F P 標的部位 # 1 、 # 2 および # 3 に対して g R N A 発現プラスミド / Cas 9 発現プラスミドをそれぞれ 25 ng / 250 ng 、 250 ng / 750 ng および 200 ng / 750 ng 用いた。

【 0094 】

10

一部の g R N A / Cas 9 の組合せが他の組合せよりも E G F P 発現の崩壊に高い効果を示す理由も、これらの組合せの一部がトランスフェクションに用いるプラスミドの量に多かれ少なかれ感受性を示す理由も理解されていない。上記 3 種類の s g R N A ゲノム内に存在するオフターゲット部位の範囲がそれぞれの活性に影響を及ぼしている可能性があるが、3 種類の s g R N A の挙動の差の原因となり得るこれらの特定の標的部位の 1 ~ 6 bp だけ異なるゲノム部位の数に差はみられなかった (表 1)。

【 0095 】

【表 3 】

表 1. 内在ヒト遺伝子を標的とする 6 種類の R G N および E G F P レポーター遺伝子を標的とする 3 種類の R G N のヒトゲノムにおけるオフターゲット部位の数

20

標的部位	オンターゲット部位に対するミスマッチの数						
	0	1	2	3	4	5	6
標的 1(VEGFA 部位 1)	1	1	4	32	280	2175	13873
標的 2(VEGFA 部位 2)	1	0	2	35	443	3889	17398
標的 3(VEGFA 部位 3)	1	1	17	377	6028	13398	35517
標的 4(EMX)	1	0	1	18	276	2309	15731
標的 5(RNF2)	1	0	0	6	116	976	7443
標的 6(FANCF)	1	0	1	18	271	1467	9551
E GFP 標的部位 #1	0	0	3	10	156	1365	9755
E GFP 標的部位 #2	0	0	0	11	96	974	7353
E GFP 標的部位 #3	0	0	1	14	165	1439	10361

30

VEGFA 、 RNF2 、 FANCF および EMX1 遺伝子を標的とする 6 種類の R G N ならびに E G F P 標的部位 # 1 、 # 2 および # 3 を標的とする 3 種類の R G N それぞれのオフターゲット部位はヒトゲノム配列ビルト G R C h 3 7 で特定されたものである。ミスマッチは g R N A がアニールする 20 nt 領域に対するものののみを認め、 P A M 配列に対するものは認めなかった。

【 0096 】

実施例 2 : F o k I - d Cas 9 融合タンパク質とのガイド R N A 対の使用
単量体の C R I S P R - Cas 9 ヌクレアーゼは、標的化ゲノム編集に広く使用されているが、望まれていないオフターゲット変異を高い頻度で誘導し得る。本実施例は、伸長した、二本鎖の配列を認識し、切断活性のための 2 つの單一ガイド R N A (g R N A) に厳密に依存する新規の二量体 R N A 誘導型 F o k I ヌクレアーゼ (R F N) を記載する。 R F N は、内在性のヒト遺伝子の D N A 配列を効率良く強力に編集できる。さらに、いずれかの 5' 末端のヌクレオチドを保有する g R N A を発現する方法が記載され、これは、二量体 R F N に有益な標的化範囲を与える重要な利点を有する。直接の比較では、单量体 Cas 9 ニックアーゼは、一般的に、マッチした單一 g R N A により導かれる R F N よりも高い頻度で、望ましくない挿入欠失および予期していない限局的点変異を誘導する。 R F N は、二量体化の特異性の増強と C R I S P R R N A ベースの標的化の簡便性を組み合わせ、非常に正確なゲノム編集を必要とする研究および治療上の適用のための重要な新規プラットフォームを提供する。

40

50

【0097】

材料および方法

以下の材料および方法を実施例2に使用した。

単一gRNAおよびマルチプレックスgRNA発現プラスミド

単一またはマルチプレックスgRNAをコードするプラスミドを、BsmBI消化型Csy4-隣接gRNA骨格(pSQT1313; Addgene)を備える、アニールした標的部位のオリゴ二重鎖(Integrated DNA Technologies)および定常領域のオリゴ二重鎖(複数のgRNAに対して)の単一ステップの連結で構築した。

【0098】

10

マルチプレックスgRNAコードプラスミドを、1)第1の標的部位をコードするアニーリングしたオリゴ、2)crrRNA、tcrRNAおよびCsy4結合部位をコードするリン酸化しアニールしたオリゴ、および3)第2の標的部位をコードするアニール化オリゴを、BsmBI 2型制限酵素で消化したU6-Csy4部位-gRNAプラスミド骨格の中に連結することにより構築した。Csy4 RNA結合部位を、gRNA配列の3'末端および5'末端に付着させ、細胞においてCas9とともに発現させた。Csy4 RNA結合部位の配列‘GUUCACUGCCGUAUAGGCAGNNNNNNNNNNNNNNGUUUUAGAGCUA
GAAAUAGCAAGUUAAAUAAGGCUAGUCGUUAUCUACUUGAAAAAGUGGCA
CCGAGUCGGUGCGUUCACUGCCGUAUAGGCAGNNNNNNNNNNNNNNNN
GUUUUAGAGCUAGAAAUAAGGCUAGUCCGUUAUCUACUU
GAAAAAGUGGCACCGAGUCGGUGCGUUCACUGCCGUAUAGGCAG’(配列番号：20)を、標準的なgRNA配列の5'末端および3'末端に融合した。

【化10】

20

GUUCACUGCCGUAUAGGCAGNNNNNNNNNNNNNNNNGUUUUAGAGCUA
GAAAUAGCAAGUUAAAUAAGGCUAGUCGUUAUCUACUUGAAAAAGUGGCA
CCGAGUCGGUGCGUUCACUGCCGUAUAGGCAGNNNNNNNNNNNNNNNN
GUUUUAGAGCUAGAAAUAAGGCUAGUCCGUUAUCUACUU
GAAAAAGUGGCACCGAGUCGGUGCGUUCACUGCCGUAUAGGCAG

(配列番号21)

【0099】

30

この配列は、Csy4部位に隣接するマルチプレックスgRNA配列である(下線)。機能的に、1つの転写物上のマルチプレックスにこれらをコードすることは、別々にこれらをコードしたものと同一の結果を有する。Csy4隣接sgRNAの全ての部分は、本明細書に記載される実験の多重部分で発現したが、sgRNAは、1つの転写物上でコードされるCsy4部位により分離されるマルチプレックスsgRNA、および追加的なCsy4配列を有する個々のsgRNAでコードすることができる。この配列では、第1のN20配列は、標的ゲノム配列の1つの鎖に相補的な配列を表し、第2のN20配列は、標的ゲノム配列の他方の鎖に相補的な配列を表す。

【0100】

40

gRNAを含むCsy4認識部位をコードするプラスミドを、「2A」ペプチド結合により分かれるCas9およびCsy4タンパク質をコードするプラスミドと共に同時にトランسفエクトした。この結果から、5'末端および3'末端に融合したCsy4部位を備えるgRNAは、上述したU2OS-EGFP崩壊アッセイを使用して、ヒト細胞中のCas9介性切断の誘導が可能なままであった。したがって、Csy4 RNAの結合部位は、gRNA配列の3'末端に付着することができ、Cas9とこれらCsy4部位含有gRNAの複合体は、細胞中で機能的なままである。

【0101】

50

いくつかの実験では、Csy4-T2A-FokI-dCas9をコードする構築物を使用した。FokI-dCas9融合物の配列を以下に示す。この配列は、FokIおよびdCas9および核局在配列の間にGGGGS(配列番号：23)リンカー(下線)を含む。

【化11】

F o k I - d C a s 9 アミノ酸配列 (F o k I - G 4 S - d C a s 9 - n l s - 3 X F L
A G)

MQLVKSELEKKSELRHKLKYVPHEYIELIEIARNSTQDRILEMKVMEFFMKVYGYR
GKHLGGSRKPDGAIYTVGSPIDYGVIVDTKAYSGGYNLPIGQADEMQRYVEENQTRN
KHINPNEWWKVYPSSVTEFKFLVSGHFKGNYKAQLTRLNHTNCNGAVLSVEELLI
GGEMIKAGTTLTLEEVRRKFNNGEINFGGGGSDKKYSIGLAIGTNSVGWAVITDEYKV
PSKKFKVLGNTDRHSIKKNLIGALLFDSETAEATRLKRTARRRYTRRKNRICYLQE
IFSNEMAKVDDSFHRLLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK
LVDSTDKAIDLRIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEE
NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSN
FDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEI
TKAPLSASMIKRYDEHHQDLTLLKALVRQQLPKEYKEIFFDQSNGYAGYIDGGASQ
EEFYKFIKPILEKMDGTEELLVQLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ
EDFYPFLKDNRKIEKILTFRIPYYVGPLARGNSRFAMTRKSEETITPWNFEVV
KGASAQSFIERMNTFDKLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKP AFL
SGEQKKAIVDILLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDL
LKI IKDKDFLDNEENEDILEDIVLTTLFEDREMIEERLKTYAHLFDDKVMKQLKRR
RYTGWGRRLSRKLINGIRDQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKA
QVSGQGDLSLHEHIANLAGSPAIIKGILQTVKVVDELVKVMGRHKPENIVIEMARENQ
TTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVD
QELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKNY
WRQLLNAKLTQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILD
NTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVR
REINNYHHAHDAYLNAV
VTA LIKKYPKLESEFVYGDYKVDVRKMI
AKSEQEIGKATAKYFFYSNIMNFFKTEITLA
NGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL
SMPQVNIVKKTEVQTGGFSKESI
LPKRNSDKLIARKKDWP
KPYGGFDSPTVAYS
VUVAKVEKGKSKKL
KSVKELLGIT
IMERSSFEKN
PIDFLEAKGYKEV
KKDLI
IKLPKYS
LFELENGR
KMLASAG
ELQGN
ELALPSKYVN
FLYASHYE
KLKGSPED
NEQKQLF
VEQH
KHYLDE
IIEQ
ISEFS
KRVI
LADANLD
KVLS
SAYNK
HRDK
PIREQ
AENI
IHLFT
LTNL
GAPA
AFKY
FDTT
IDRK
KRYTS

TKEVLDATL
IHQ SIT GLY ETR
IDLSQLGGDGSPKKRK
VSSDYKDHDG
DYKDHD
DIDY
KDDDDK (配列番号 24)

10

20

30

【化12】

F o k I - d C a s 9 ヌクレオチド配列 (F o k I - G 4 S - d C a s 9 - n l s - 3
X F L A G)

ATGCAACTAGTCAAAAGTGAACCTGGAGGAGAAGAAATCTGAACCTCGTCATAAATTG
AAATATGTGCCCTCATGAATATATGAAATTAAATTGAAATTGCCAGAAATTCCACTCAG
GATAGAATTCTGAAATGAAGGTAAATGGAATTTTATGAAAGTTATGGATATAGA
GGTAAACATTGGGTGGATCAAGGAAACCGGACGGAGCAATTATACTGTCGGATCT
CCTATTGATTACGGTGTGATCGGGATACTAAAGCTATAGCGGAGGTTATAATCTG
CCAATTGGCCAAGCAGATGAAATGCAACGATATGTCGAAGAAAATCAAACACGAAAC
AAACATATCAACCCTAATGAATGGTGGAAAGTCTATCCATCTTGTAAACGGAATT
AAGTTTATTGTGAGTGGTCACTTAAAGGAAACTACAAAGCTCAGCTTACACGA
TTAAATCATATCACTAATTGAAATGGAGCTGTTAGTGTAGAAGAGCTTTAATT
GGTGGAGAAATGATTAAAGCCGGCACATTAACCTTAGAGGAAGTCAGACGGAAATT
AATAACGGCGAGATAAACCTTGGTGGCGGTGGATCGATAAAAAGTATTCTATTGGT
TTAGGCATCGGCACTAATTCCGTGGATGGCTGTCAATAACCGATGAATAACAGTA
CCTTCAAAGAAATTAAAGGTGTTGGGAACACAGACCGTCATTGATTAAAAAGAAT
CTTATCGGTGCCCTCCTATTGATAGTGGCGAAACGGCAGAGGGCACTCGCCTGAAA
CGAACCGCTCGGAGAAGGTATACACGTCGAAGAACCGAATATGTTACTTACAAGAA
ATTTTAGCAATGAGATGGCAAAGTTGACGATTCTTCTTCACCCTTGAAAGAG
TCCTCCTTGTGAAAGAGGACAAGAAACATGAACGGCACCCATCTTGAAACATA
GTAGATGAGGGGGCATATCATGAAAAGTACCCAACGATTATCACCTCAGAAAAAG
CTAGTTGACTCAACTGATAAAGCGGACCTGAGGTTAATCTACTTGGCTTGCCT
ATGATAAAAGTCCGTGGCACTTCTATTGAGGGTGTCTAAATCCGACAACCTCG
GATGTCGACAAACTGTTCATCCAGTTAGTACAAACCTATAATCAGTTGTTGAAGAG
AACCTATAAATGCAAGTGGCGTGGATGCGAAGGCTATTCTAGCGCCCGCTCT
AAATCCCACGGCTAGAAAACCTGATCGCACAATTACCCGGAGAGAAGAAAAATGG
TTGTTCGGTAAACCTTATAGCGCTCTCACTAGGCCTGACACCAAATTAAAGTCGAAC
TTCGACTTAGCTGAAGATGCAAATTGCAAGCTAGTAAAGGACACGTACGATGACGAT
CTCGACAATCTACTGGCACAAATTGGAGATCAGTATGCGGACTTATTGGCTGCC
AAAAACCTTAGCGATGCAATCCTCTATCTGACATACTGAGAGTTAAACTGAGATT
ACCAAGGCGCCGTTATCCGCTTCAATGATCAAAGGTACGATGAACATCACCAAGAC
TTGACACTTCTCAAGGCCCTAGTCGTCACTGCCTGAGAAATATAAGGAAATA
TTCTTGATCGAAAAACGGGTACGCAGGTTATATTGACGGCGGAGCGAGTCAA
GAGGAATTCTACAAGTTATCAAACCCATATTAGAGAAGATGGATGGACGGAAGAG
TTGCTTGTAAAACCTCAATCGCGAAGATCTACTGCGAAAGCAGCGGACTTCGACAAC
GGTAGCATTCCACATCAAATCCACTTAGCGAATTGCTATGCTATACTTAGAAGGCAG
GAGGATTTTATCCGTTCTCAAAGACAATCGTAAAAGATTGAGAAAATCTTAACC
TTTCGATACCTTACTATGTGGACCCCTGGCCCGAGGGAACTCTCGGTTCGCATGG
ATGACAAGAAAGTCCGAAGAAACGATTACTCCATGGAATTGGAGGAAGTTGTCGAT
AAAGGTGCGTCAGCTCAATCGTTCATCGAGAGGATGACCAACTTGACAGAATTAA
CCGAACGAAAAGTATTGCTAAGCACAGTTACTTACGAGTATTACAGTGTAC
AATGAACCTACGAAAGTTAAGTATGTCACTGAGGGCATGCGTAAACCGCCTTCTA
AGCGGAGAACAGAAGAAAGCAATAGTAGATCTGTTATTCAAGACCAACCGCAAAGTG
ACAGTTAAGCAATTGAAAGAGGACTACTTTAAGAAAATTGAATGCTTCGATTCTGTC
GAGATCTCGGGGTAGAAGATCGATTTAATGCGTCACTGGTACGTATCATGACCTC
CTAAAGATAATTAAAGATAAGGACTTCCTGGATAACGAAGAGAATGAAGATATCTA
GAAGATATAGTGTGACTCTTACCCCTTTGAAGATCGGGAAATGATTGAGGAAGA
CTAAAAACATACGCTCACCTGTTGACGATAAGGTTATGAAACAGTTAAAGAGGCGT

10

20

30

40

CGCTATACTGGCTGGGACGATTGTCGCGAAACTTATCAACGGGATAAGAGACAAG
CAAAGTGGTAAAACATTCTGATTTCTAAAGAGCGACGGCTCGCCAATAGGAAC
TTTATGCAGCTGATCCATGATGACTCTTAACCTTAAAGAGGATATACAAAAGGCA
CAGGTTCCGGACAAGGGACTCATTCGACGAAACATATTGCGAATCTTGCTGGTCG
CCAGGCCATAAAAAGGGCATACTCCAGACAGTCAGTAAAGTAGTGGATGAGCTAGTTAAG
GTCATGGGACGTACAAACCGGAAAACATTGTAATCGAGATGGCACGCGAAAATCAA
ACGACTCAGAAGGGCAAAAAACAGTCGAGAGCGGATGAAGAGAATAGAAGAGGGT
ATTAAAGAACTGGGAGCCAGATCTTAAAGGAGCATTCTGTGGAAAATACCCAATTG
CAGAACGAGAAACTTACCTTACAAATGGAAGGGACATGTATGTTGAT
CAGGAACCTGGACATAAACGTTATCTGATTACGACGTCGATGCCATTGTACCCCAA
TCCTTTGAAGGGACGATTCAATCGACAATAAAGTGTCTACACGCTCGGATAAGAAC
CGAGGGAAAAGTGACAATGTTCCAAGCGAGGAAGTCGTAAGAAAATGAAGAACTAT
TGGCGGCAGCTCCTAAATGCGAAACTGATAACGCAAAGAAAAGTTCGATAACTTAAC
AAAGCTGAGAGGGTGGCTGTGACTTGACAAGGCCGGATTATTAAACGTCAG
CTCGTGGAAACCCGCCAATCACAAAGCATGTTGACAGATACTAGATTCCGAATG
AATACGAAATACGACGAGAACGATAAGCTGATTGGGAAGTCAAAGTAATCACTTA
AAGTCAAAATTGGTGTGGACTTCAGAAAGGATTTCATTCTATAAAAGTTAGGGAG
ATAAAATAACTACCACCATGCGCACGACGCTTATCTTAATGCCGCGTAGGGACCGCA
CTCATTAAAGAAATACCGAAGCTAGAAAGTGAGTTGTATGGTATTACAAAGTT
TATGACGTCCGTAAGATGATCGCAGAACAGGAGATAGGCAAGGCTACAGCC
AAATACTTCTTTATTCTAACATTATGAATTCTTAAGACGGAAATCAGCTGGCA
AACGGAGAGATA CGCAAACGACCTTAATTGAAACCAATGGGGAGACAGGTGAAATC
GTATGGGATAAGGGCCGGACTTCGCGACGGTGAGAAAAGTTGTCCATGCCCAA
GTCAACATAGTAAAGAAAAGTGGCTCGAGACCGGAGGGTTTCAAAGGAATCGATT
CTTCCAAAAAGGAATAGTGATAAGCTCATCGCTCGTAAAAGGACTGGGACCCGAAA
AAGTACGGTGGCTCGATAGCCCTACAGTTGCCTATTCTGCTCTAGTAGTGGCAAAA
GTTGAGAAGGGAAAATCCAAGAAACTGAAAGTCAGTCAAAGAATTATTGGGATAACG
ATTATGGAGCGCTCGTCTTGAAAAGAACCCCATCGACTTCCTGAGGCGAAAGGT
TACAAGGAAGTAAAAAAGGATCTCATAATTAAACTACCAAAGTATAGTCTGTTGAG
TTAGAAAATGGCGAAAACGGATGTTGGCTAGCGCCGGAGAGCTCAAAAGGGAAAC
GAACTCGCACTACCGCTAAATACGTGAATTCCCTGTATTAGCGTCCCATTACGAG
AAGTGAAAGGTTACCTGAAGATAACGAACAGAACGAACTTTGTTGAGCAGCAC
AAACATTATCTCGACGAAATCATAGAGCAAATTTCGGAATTCACTGAGAGTCATC
CTAGCTGATGCCAATCTGGACAAAGTATTAAGCGCATAACAACAGCACAGGGATAAA
CCCATACGTGAGCAGCGGAAAATATTATCCATTGTTACTCTTACCAACCTCGGC
GCTCCAGCCGCATTCAAGTATTTGACACAACGATAGATCGCAAACGATACTTCT
ACCAAGGAGGTGCTAGACGCGACACTGATTACCAATCCATCACGGATTATATGAA
ACTCGGATAGATTGTCACAGCTGGGGTGACGGATCCCCAAGAAGAAGAGGAAA
GTCTCGAGCGACTACAAAGACCATGACGGTGATTATAAGATCATGACATCGATTAC
AAGGATGACGATGACAAGTGA (配列番号 25)

【0102】

40

代替として、ヒトコドンを最適化した構築物を使用した。これは、NおよびC末端核局在シグナルを含むものであった。配列を以下に示す。

【化13】

N 1 s - F o k I - d C a s 9 - n 1 s アミノ酸配列

MDKKYSIGLAIGTNSVGWAVITDEYKVP SKFKVLGNTDRHSIKKNLIGALLFDSGE
TAEATRLKRTARRRYTRRKNRICYLQEIFSNE MAKVDDSFHRLEESFLVEEDKKHE
RHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTD KADLRLIYLALAHMIKFRGHFLIE

GDLNPNDNSDVKLFQQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQ
 LPGEKKNGLFGNLIALSGLTPNFSNFDAEDAKLQLSKDTYDDLDNLLAQIGDQ
 YADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQ
 LPEKYKEIFFDQSNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKNLNREDLL
 RKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNRKREKIEKILTFRIPYYVGPLA
 RGNNSRFAMTRKSEETITPWNFEVVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL
 LYEFYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFK
 KIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLFE
 DREMIEERLKTYAHLFDDKVMKQLKRRRTGWRGLSRKLINGIRDQSGKTILDFLK
 SDGFANRNFNMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKGKILQTV
 KVVDLVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEELGKELGSQILKE
 HPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFQFLKDDSIDNK
 VLTRSDKNRGKSDNVPSEEVVKMKNYWRQLLNAKLITQRKFQDNLTKAERGGLSELD
 KAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKD
 FQFYKVREINNYHHAHDAYLNAVGTALIKKYPKLESEFVYGDYKVDVRKMIAKSE
 QEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPRIETNGETGEIVWDKGRDFATV
 RKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWPCKYGGFDSPTVA
 YSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLI
 LPKYSLFELENGKRMLASAGELQKGNELALPSKYVNFLYLAHYEKLKGSPEDNEQ
 KQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVL SAYNHRDKPIREQAENI
 LFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQ SITGLYETRIDSQLGGD
 GSPKKKRKVSSDYKDHDGDYKDHDIDYKDDDDK (配列番号 26) 20

【化14】

N1s-FokI-dCas9-n1s ヌクレオチド配列

ATGCCCTAAGAAGAACCGGAAGGGAGCAGCCAACCTGTGAAGTCTGAACTCGAGGAG
 AAAAAATCAGAGTTGAGACACAAGTTGAAGTACGTGCCACACGAATACATCGAGCTT
 ATCGAGATCGCCAGAAACAGTACCCAGGATAGGATCCTTGAGATGAAAGTCATGGAG
 TTCTTATGAAGGTCTACGGTTATAGAGGAAAGCACCTGGCGGTAGCAGAAAGCCC
 GATGGCGCCATCTATACTGTCGGATCTCTATCGATTATGGGGTAGCTGGATACC
 AAAGCTTACTCAGGCAGGTACAACCTGCCATAGGACAAGCCGACGAGATGCAGCGG
 TATGTCGAAGAGAACCAAGACCGCAACAAAGCACATCAACCCCATGAATGGGGAAA
 GTGTACCCAAGTAGTGTGACTGAGTTCAAGTCTGTTGTCCTCCGCCACTTAAG
 GGCAATTATAAAGCTCAGCTCACTAGACTCAATCACATCACAAACTGCAACGGAGCT
 GTGTTGTCAGTGGAGGGACTCCTGATTGGAGGCGAGATGATCAAAGCCGGACCC
 ACACGGAGGGAGGTGCGGCGGAAGTTCAACAATGGAGAGATCAACTCGGTGGCGGT
 GGATCCGATAAAAAGTATTCTATTGGTTAGCCATCGGACTAATCCGTTGGATGG
 GCTGTCATAACCGATGAATACAAGTACCTCAAAGAAATTAGGTGTTGGGAAC
 ACAGACCGTCATTGATTAAGAATCTATCGGTGCCCTCCTATCGATAGTGGC
 GAAACGGCAGAGGCGACTCGCCTGAAACGAACCGCTCGGAGAAGGTATACACGTC
 AAGAACCGAATATGTTACTTACAAGAAATTAGCAATGAGATGGCCAAAGTTGAC
 GATTCTTCTTCACCGTTGGAAAGAGTCCTCCTGTCGAAGAGGACAAGAACAT
 GAACGGCACCCATCTTGGAAACATAGTAGATGAGGTGGCATATCATGAAAGTAC
 CCAACGATTATCACCTCAGAAAAAGCTAGTTGACTCAACTGATAAAGCGGACCTG
 AGGTAAATCTACTTGGCTCTGCCATATGATAAAGTCCGTGGGCACTTCCTCATT
 GAGGGTGTAAATCCGGACAACCGGATGTCGACAAACTGTTCATCCAGTAGTA
 CAAACCTATAATCAGTTGTTGAAGAGAACCCATATAATGCAAGTGGCGTGGATGCG
 AAGGCTATTCTAGCGCCCGCTCTAAATCCGACGGCTAGAAAACCTGATCGCA
 CAATTACCCGGAGAGAAGAAAAATGGGTTGTTGGTAACCTTATAGCGCTCTCACTA
 GGCCTGACACCAAATTAAAGTCGAACTCGACTAGCTGAAGATGCCAAATTGCAG 30
 40

CTTAGTAAGGACACGTACGATGACGATCTGACAATCTACTGGCACAAATTGGAGAT
 CAGTATGCGGACTTATTTGGCTGCCAAAAACCTTAGCGATGCAATCCTCTATCT
 GACATACTGAGAGTTAATACTGAGATTACCAAGGCAGCTTATCCGCTTCATGATC
 AAAAGGTACGATGAACATCACCAAGACTGACACTCTCAAGGCCAGTCCGTCAG
 CAACTGCCTGAGAAATATAAGGAAATATTCTTGATCAGTCGAAAAACGGGTACGCA
 GGTTATATTGACGGCGAGCGAGTCAAGAGGAATTCTACAAGTTATCAAACCCATA
 TTAGAGAAGATGGATGGACGGAAGAGTGCTTGTAAAACCTCAATCGGAAGATCTA
 CTGCGAAAGCAGCGACTTCGACAACGGTAGCATCCACATCAAACCACTTAGGC
 GAATTGCATGCTATACTTAGAAGGCAGGAGGATTTTATCCGTTCTCAAAGACAAT
 CGTAAAAAGATTGAGAAAATCTAACCTTCGATACCTTACTATGTGGGACCCCTG
 GCCCGAGGGAACCTCGGTTCGATGGATGACAAGAAAGTCCGAAGAAACGATTACT
 CCATGGAATTTGAGGAAGTTGTCGATAAAGGTGCGTCAGCTCAATCGTTATCGAG
 AGGATGACCAACTTGACAAGAATTACCGAACGAAAAAGTATTGCTTAAGCACAGT
 TTACTTTACGAGTATTCACAGTGTACAATGAACCTACGAAAGTTAAGTATGTCACT
 GAGGGCAGCGTAAACCCGCTTCTAACCGGAGAACAGAACAGAAAGCAATAGTAGAT
 CTGTTATTCAAGACCAACCGCAAAGTGCAGTTAACAGTAAAGAATTGAAAGAGGACTACTTT
 AAGAAAATTGAATGCTCGATTCTGTCGAGATCTCCGGGGTAGAAGATCGATTTAAT
 GCGTCACTGGTACGTATCATGACCTCTAAAGATAATTAAAGATAAGGACTTCCTG
 GATAACGAAGAGAATGAAGATACTTAGAAGATATAGTGTGACTCTTACCCCTTT
 GAAGATCGGGAAATGATTGAGGAAAGACTAAAAACATACGCTCACCTGTTGACGAT
 AAGGTTATGAAACAGTTAAAGAGGCAGCTACGGGCTGGGACGATTGTCGCGG
 AAACCTTATCAACGGGATAAGAGACAAGCAAAGTGGTAAAACCTATTCTGATTTCTA
 AAGAGCGACGGCTTCGCCAATAGGAACCTTATGCAGCTGATCCATGATGACTCTTA
 ACCTCAAAGAGGATATACAAAAGGCACAGGTTCCGGACAAGGGACTCATTGCAC
 GAACATATTGCGAATCTGCTGGTCGCCAGCCATCAAAAAGGGCATACTCCAGACA
 GTCAAAGTAGTGGATGAGCTAGTTAACGGTATGGGACGTACAAACCGGAAACATT
 GTAATCGAGATGGCACCGAAAATCAAACGACTCAGAACGGGCAAAAAACAGTCGA
 GAGCGGATGAAGAGAATAGAACGGGTATTAAGAACACTGGCAGCCAGATCTTAAAG
 GAGCATCCTGTGGAAAATACCCATTGAGAACGAGAAACTTACCTCTATTACCTA
 CAAAATGGAAGGGACATGTATGTTGATCAGGAACCTGGACATAACCGTTATCTGAT
 TACGACGTCGATGCCATTGTACCCATTCTTTGAAGGACGATTCAATCGACAAT
 AAAGTGTACACGCTCGATAAGAACGGAGGAAAAGTGACAATGTTCAAGCGAG
 GAAGTCGTAAGAAAATGAAGAAACTATTGGCGGAGCTCTAAATGCGAAACTGATA
 ACGCAAAGAAAGTTCGATAACTAACTAAAGCTGAGAGGGTGGCTGTGACTT
 GACAAGGCCGGATTATTAAACGTCAGCTCGTGGAAACCCGCCAAATCACAAAGCAT
 GTTGCACAGATACTAGATTCCCGAATGAATACGAAATACGACGAGAACGATAAGCTG
 ATTGGGAAGTCAAAGTAATCACTTAAAGTCAAAATTGGTGTGGACTTCAGAAAG
 GATTTCAATTCTATAAAAGTTAGGGAGATAAAATAACTACCACCATGCGCACGACGCT
 TATCTTAATGCCGTCGTAGGGACCGCACTCTAACGAAATACCGAAGCTAGAAAAGT
 GAGTTGTGATGGTATTACAAAGTTATGACGTCGTAAGATGATCGCAGAACG
 GAACAGGAGATAGGCAAGGCTACGCCAAACTCTTTATTCTAACATTATGAAT
 TTCTTAAAGACGGAAATCACTCTGGCAACGGAGAGATACGAAACGACCTTAATT
 GAAACCAATGGGGAGACAGGTGAAATCGTATGGATAAGGGCCGGACTTCGCGACG
 GTGAGAAAAGTTTGTCCATGCCCAAGTCACACATAGTAAAGAAAACGTGAGGTGAG
 ACCGGAGGGTTTCAAAGGAATCGATTCTCCAAAAGGAATAGTGATAAGCTCAGT
 GCTCGTAAAAGGACTGGGACCGAAAAAGTACGGTGGCTCGATAGCCCTACAGTT
 GCCTATTCTGCTCTAGTAGTGGCAAAAGTTGAGAACGGAAAATCCAAGAAACTGAAAG
 TCAGTCAAAGAATTATTGGGATAACGATTATGGAGCGCTCGTCTTGAAAAGAAC
 CCCATCGACTTCCTTGAGGCAGAGGTTACAAGGAAGTAAAAAGGATCTCATAATT
 AAACTACCAAAGTATAGTGTGTTGAGTTAGAAAATGGCCGAAAACGGATGTTGGCT

10

20

30

40

AGGCCGGAGAGCTCAAAAGGGAACGAACTCGCACTACCGTCTAAATACGTGAAT
 TTCCTGTATTAGCGTCCCATTACGAGAAGTTGAAAGGTTCACCTGAAGATAACGAA
 CAGAAGCAACTTTTGTTGAGCAGCACAAACATTATCTCGACGAAATCATAGAGCAA
 ATTCGGAATTCAAGTAGAGTCATCCTAGCTGATGCCAATCTGGACAAAGTATTA
 AGCGCATAACAACAAGCACAGGGATAAACCCATACGTGAGCAGGCGAAAATATTATC
 CATTGTTACTCTTACCAACCTCGGCCTCCAGCCGCATTCAAGTATTTGACACA
 ACGATAGATCGCAAACGATACTTCTACCAAGGGAGGTGCTAGACGCGACACTGATT
 CACCAATCCATCACGGGATTATATGAAACTCGGATAGATTGTCACAGCTGGGGGT
 GACGGATCCCCAAGAAGAAGAGGAAAGTCTCGAGCGACTACAAAGACCATGACGGT
 GATTATAAAGATCATGACATCGATTACAAGGATGACGATGACAAGTGA

10

(配列番号 27)

【 0 1 0 3 】**組織培養およびトランスフェクション**

すべての細胞培養実験を、HEK293細胞、U2OS細胞、または安定して統合された単一複製の不安定化EGFP遺伝子を保有するU2OS細胞(U2OS·EGFP細胞)で行った。細胞株を、10%のFBS、2mM Glutamax(Life Technologies)、およびペニシリン/ストレプトマイシンを補充したAdvanced DMEM(Life Technologies)の中、5%のCO₂、37°Cで培養した。さらに、U2OS·EGFP細胞を、400μg/mlのG418の存在下で培養した。

20

【 0 1 0 4 】

U2OS細胞およびU2OS·EGFP細胞を、製造元の説明書にしたがって、Lonza 4D-NucleofectorのDN-100プログラムを使用してトランスフェクトした。最初のFokI-dCas9活性のスクリーニングおよび融合したスペーサー長解析の実験では、750ngのpCAG-Csy4-FokI-dCas9-nlsヌクレアーゼプラスミドおよび250ngのgRNAコードプラスミドを、トランスフェクションの対照としての50ngのtdTomato発現プラスミド(Clontech)と共にトランスフェクトした。U2OS細胞およびU2OS·EGFP細胞の他のすべての実験では、975ngのヒトコドン最適化pCAG-Csy4-T2A-nls-hFokI-dCas9-nls(SQT1601)またはpCAG-Cas9-D10Aニッカーゼ(NW3)を、325ngのgRNAベクターおよび10ngのTd_tomato発現プラスミドと共にトランスフェクトし、トランスフェクションから3日後に解析した。HEK293細胞を、製造元の説明書にしたがってリポフェクタミン(Life Technologies)を使用し、750ngのヌクレアーゼプラスミド、250ngのgRNA発現プラスミド、および10ngのTd_tomatoでトランスフェクトし、トランスフェクションから3日後に、NHEJ媒介性変異原性について解析した。

30

【 0 1 0 5 】

単一のトランスフェクションを、最初のスペーサー活性スクリーニングのために実施し、焦点を当てたスペーサー長解析のためにトランスフェクションを二重に行った。他のすべてのトランスフェクションを3重に行った。

40

【 0 1 0 6 】**EGFP崩壊アッセイ**

EGFP崩壊アッセイを、U2OS·EGFPレポーター細胞を使用して、以前に記載されているように実施した(実施例1およびReyon et al., Nat Biotech 30, 460-465(2012)参照)。細胞を、BD Biosciences LSR IIまたはFortessa FACSのアナライザーを使用してEGFPおよびtdTomatoの発現についてアッセイした。

【 0 1 0 7 】**T7E1アッセイによるヌクレアーゼまたはニッカーゼ誘導型変異比率の定量化**

50

T7E1アッセイを、以前に記載されているように実施した(Reyon et al., Nat Biotech 30, 460-465 (2012))。簡潔に述べると、Sciclon G3 リキッドハンドリングワークステーション(Caliper)と共に製造元の説明書にしたがってAgencourt DNA advance ゲノムDNA単離キット(Beckman Coulter Genomics)を使用して、ゲノムDNAをトランスフェクションから72時間後に単離した。ゲノムの座位を増幅するPCR反応を、Phusion Hot-start Flex DNAポリメラーゼ(New England Biolabs)を使用して実施した。2つのステッププロトコル(98、30秒；(98、7秒；72、30秒)×35、72、5分)、またはタッチダウンプロトコル((98°C、10秒；72~62°C、?1°C/周期、15秒；72°C、30秒)×10周期(98°C、10秒；62°C、15秒；72°C、30秒)×25周期)を使用して試料を増幅した。200ngの精製したPCRアンプリコンを、変性、ハイブリダイズし、T7エンドヌクレアーゼ(I)(New England Biolabs)を用いて処理した。変異の頻度を、以前に記載されているようにQiaxcel キャピラリー電気泳動器具(Qiagen)を用いて定量した(Reyon et al., Nat Biotech 30, 460-465 (2012))。

【0108】

変異誘発ゲノムDNAのサンガー配列決定

T7E1アッセイで使用した同一の精製したPCR産物を、Topoクローニングし(Life Technologies)、個々のクローンのプラスミドDNAを単離し、M13リバースプライマー(5'-GTAATACGACTGGCCAG-3'；配列番号：19)を使用して配列決定した。

【0109】

照度ライブラリーの調製および解析

200~350bpの短いPCR産物を、Phusion Hot-start FLEX DNAポリメラーゼを使用して増幅した。PCR産物を、製造元の説明書にしたがってAmpure XPビーズ(Beckman Coulter Genomics)を使用して生成した。Dual-indexed Truseq Illuminaディープシーケンシングライブラリーを、Sciclon G3リキッドハンドリングワークステーション上のハイスクロープトライトライブラリー調製系(Kapa Biosystems)を使用して調製した。最終的なアダプター-連結ライブラリーを、Qiaxcel キャピラリー電気泳動器具(Qiagen)を用いて定量した。150bpの対形成した末端の配列決定を、Dana-Farber Cancer Institute Molecular Biology CoreによるIllumina MiSeq シーケンサー上で実施した。

【0110】

MiSeqの対形成した末端の読み取り値を、bwaを使用してヒトゲノム参照GChr37に対してマッピングした。30超の平均定量スコアを有する読み取り値を、統合した標的または候補となるオフターゲットヌクレアーゼの結合部位に重複する挿入または欠失の変異について解析した。変異解析を、Genome Analysis Toolkit(GATK)およびPythonを用いて行った。

【0111】

オフターゲット探索アルゴリズム

ヒトゲノムを通してスライディングウインドウ中に特定数未満のミスマッチを有するマッチを探査する標的部位のマッチングアルゴリズムを実施した。

【0112】

実施例2a：二量体RNA誘導型ヌクレアーゼを設計する根拠

Cas9の標的化の簡便性と二量体の特異性の利点を組み合わせる単一のプラットフォームを開発し得ると仮定した。それを実施するために、良好に特徴付けられた、二量体化

10

20

30

40

50

依存性 F o k I ヌクレアーゼドメインを、 R N A 誘導型触媒不活性 C a s 9 (d C a s 9) タンパク質に融合した。 F o k I 含有 Z F N および T A L E N のように、これら融合体の二量体は、これらの間に特定の長さの「スペーサー」配列を備える 2 つの「片側部位」から構成される部位を標的化するために結合する際、配列特異的な D N A 切断を媒介し得ることが期待された(図 4 A)。そのような融合体は、活性のために 2 つの g R N A を必要とし、かつ単一 g R N A が、 D N A 切断に必要な 2 つの F o k I 含有融合タンパク質を動員するには恐らく不十分であるか、動員することができないので、高い特異性を有すると仮定された(図 4 A)。そのような二量体系は、標準的な単量体 C a s 9 ヌクレアーゼと比較して改善した特異性を示し、また、単一のニッカーゼが望ましくない変異原性作用を発揮する場合のある対形成されたニッカーゼ系に勝って、重要な特性の利点を有する可能性があるであろうと仮定された。

【 0 1 1 3 】

実施例 2 b : 5' 末端のヌクレオチドの制限なしの g R N A のマルチプレックス発現

二量体の R N A 誘導型ヌクレアーゼのための標的化の範囲は、現存する g R N A 発現方法を使用すると、狭い(low)ものである。 2 つの配列の要件、すなわち d C a s 9 により特定される 5' N G G の P A M 配列の要件、およびほとんどの発現ベクターにおける U 6 プロモーターの使用により課される g R N A の 5' 末端での G ヌクレオチドの要件は、概して、 d C a s 9 モノマーの標的化範囲を制限する。しかしながら、 g R N A の 5' G の要件が軽減されるなら、標的化の範囲は 16 倍改善するであろう。

【 0 1 1 4 】

5' ヌクレオチドを備える g R N A の発現を可能にするマルチプレックス系を開発するために、プラスミドを構築したが、そのプラスミドから、それぞれが C s y 4 リボヌクレアーゼの切断部位に隣接する(H a u r w i t z e t a l . , S c i e n c e 3 2 9 , 1 3 5 5 - 1 3 5 8 (2 0 1 0)) 2 つの g R N A が、 U 6 プロモーターから転写した単一の R N A の中に発現できる(図 4 B)。 C s y 4 は、この転写物を処理し、これにより 2 つの g R N A を放出すると予期される。知られている C s y 4 媒介性切断の機構((H a u r w i t z e t a l . , S c i e n c e 3 2 9 , 1 3 5 5 - 1 3 5 8 (2 0 1 0) ; S t e r n b e r g e t a l . , R N A 1 8 , 6 6 1 - 6 7 2 (2 0 1 2)) に基づき、それぞれ処理された g R N A は、 3' 末端上に C s y 4 認識領域を保持し、 C s y 4 タンパク質がこの部位に結合される(図 4 B)。この構成では、いずれの 5' ヌクレオチドを備える g R N A をも発現することが可能であるはずである。この系を、 E G F P レポーター遺伝子内の部位を標的とする 2 つの g R N A を発現するために使用することにより、試験した。ヒト細胞において、 C s y 4 および C a s 9 ヌクレアーゼと共にこの転写物を共発現することにより、両方の E G F P 標的部位での挿入欠失変異、およびこれら部位の間の配列の欠失が導入される(図 4 C)。これらの実験から、両方の g R N A が、単一の親 R N A 転写物から処理され、両方が、ヒト細胞での C a s 9 ヌクレアーゼの活性を誘導できることが示唆される。

【 0 1 1 5 】

実施例 2 c . 二量体 R N A 誘導型ヌクレアーゼの構築および最適化

F o k I ヌクレアーゼドメインおよび d C a s タンパク質を保有する 2 つの異なるハイブリッドタンパク質を構築した。このうち 1 つでは、 F o k I ヌクレアーゼドメインが、 d C a s 9 のカルボキシ末端に融合しており(d C a s 9 - F o k I)、もう一つでは、アミノ末端に融合している(F o k I - d C a s 9)(図 5 A)。 d C a s 9 - F o k I タンパク質は、構造上 Z F N および T A L E N に類似する(図 5 A)。これらの融合物のいずれかまたは両方が、 D N A の部位特異性切断を媒介し得るかどうかを確かめるために、 N H E J 媒介性挿入欠失の E G F P レポーター遺伝子への導入を迅速かつ簡便に定量化できる良好に確立したヒト細胞系アッセイを使用した(実施例 1 で上述の E G F P 崩壊アッセイ)。効率的な切断に必要とされる片側部位の幾何学形状は知られていないので、 E G F P の様々な部位を標的とする 60 対の g R N A を設計した。これら g R N A 対のそれが標的化とする 2 つの片側部位は、 P A M 配列の両方が、スペーサー配列に直接隣接

10

20

30

40

50

しているか（PAM内部配向）、または完全長の標的部位の外の境界に位置している（「PAM外部」配向）（図5B）ように配向された。さらに、スペーサー配列は、0～31bpの長さでも変動させた（図5Bおよび表2）。

【表4】

表2								
FokI-dCas9 EGFP 対 #	名称	標的開始位置 (+)	配列(+)部位	配列番号	配列(-)部位	配列番号	端から端までの 「スペーサー」 距離	PAM
1	EGFP 部位 1	74	GAGCTGGACGGCGACGTAACCGG	28.	CGCCGGACACGCTGAACTTGTGG	29.	0	内部
	EGFP 部位 2	174	CCGGCAAGCTGCCGTGCCCTGG	30.	GGTCAGGGTGGTCACGAGGGTGG	31.	1	内部
	EGFP 部位 3	37	CGAGGAGCTGTTCACCGGGGTGG	32.	CCGTCCAGCTCGACCAGGATGGG	33.	2	内部
	EGFP 部位 4	37	CGAGGAGCTGTTCACCGGGGTGG	34.	GCGCTCCAGCTCGACCAGGATGGG	35.	3	内部
	EGFP 部位 5	174	CCGGCAAGCTGCCGTGCCCTGG	36.	GTAGGTCAGGGTGGTCACGAGGG	37.	4	内部
	EGFP 部位 6	34	GGGCGAGGAGCTGTTCACCGGGG	38.	CCGTCCAGCTCGACCAGGATGGG	39.	5	内部
	EGFP 部位 7	33	AGGGCGAGGAGCTGTTCACCGGG	40.	CCGTCCAGCTCGACCAGGATGGG	41.	6	内部
	EGFP 部位 8	32	AAGGGCGAGGAGCTGTTCACCGG	42.	CCGTCCAGCTCGACCAGGATGGG	43.	7	内部
	EGFP 部位 9	32	AAGGGCGAGGAGCTGTTCACCGG	44.	GCGTCCAGCTCGACCAGGATGG	45.	8	内部
	EGFP 部位 10	106	CAGCGTGTCCGGCGAGGGCGAGG	46.	CTTCAGGGTCAGCTTGCCTAGG	47.	9	内部

10

表2								
FokI-dCas9 EGFP 対 #	名称	標的開始位置 (+)	配列(+)部位	配列番号	配列(-)部位	配列番号	端から端までの 「スペーサー」 距離	PAM
2	EGFP 部位 11	34	GGGCGAGGAGCTGTTCACCGGGG	48.	CGTCGCCGTCCAGCTCGACCAGG	49.	10	内部
	EGFP 部位 12	33	AGGGCGAGGAGCTGTTCACCGGG	50.	CGTCGCCGTCCAGCTCGACCAGG	51.	11	内部
	EGFP 部位 13	32	AAGGGCGAGGAGCTGTTCACCGG	52.	CGTCGCCGTCCAGCTCGACCAGG	53.	12	内部
	EGFP 部位 14	155	CTGAAGTTCATCTGCAACCACCGG	54.	GTGGTCACGAGGGTGCGCAGGG	55.	13	内部
	EGFP 部位 15	101	AAGTTCAAGCGTGTCCGGCGAGGG	56.	CTTCAGGGTCAGCTTGCCTAGG	57.	14	内部
	EGFP 部位 16	100	CAAGTTCAAGCGTGTCCGGCGAGG	58.	CTTCAGGGTCAGCTTGCCTAGG	59.	15	内部
	EGFP 部位 17	58	GGTGCCTCATCTGGTCAAGCTGG	60.	CGCCGGACACGCTGAACCTGTGG	61.	16	内部
	EGFP 部位 18	74	GAGCTGGACGGCGACGTAACCGG	62.	GGCATCGCCCTCGCCCTCGCCGG	63.	17	内部
	EGFP 部位 19	307	GGAGCGCACCATCTTCTTCAGG	64.	CTCGAACCTCACCTCGCCGGCGGG	65.	18	内部
	EGFP 部位 20	155	CTGAAGTTCATCTGCAACCACCGG	66.	GTCAGGGTGGTCACGAGGGTGGG	67.	19	内部
	EGFP 部位 21	95	GGCCACAAAGTTCAAGCGTGTCCGG	68.	CTTCAGGGTCAGCTTGCCTAGG	69.	20	内部

20

30

40

表 2								
FokI -dCa s9 EGFP 対 #	名称	標的開 始位置 (+)	配列 (+) 部位	配 列 番 号	配列 (-) 部位	配 列 番 号	端から端 までの 「スペー サー」 距離	PAM
22	EGFP 部位 22	203	CTCGTGACCACCCCTGACCTACGG	70.	CGTGCTGCTTCATGTGGTGGGG	71.	21	内部
23	EGFP 部位 23	174	CCGGCAAGCTGCCGTGCCCTGG	72.	GCTGAAAGCACTGCACGCCGTAGG	73.	22	内部
24	EGFP 部位 24	107	AGCGTGTCCGGCGAGGGCGAGGG	74.	GGTGGTGCAGATGAACCTTCAGGG	75.	23	内部
25	EGFP 部位 25	106	CAGCGTGTCCGGCGAGGGCGAGG	76.	GGTGGTGCAGATGAACCTTCAGGG	77.	24	内部
26	EGFP 部位 26	49	CACCGGGTGGTGCCTACGG	78.	CGCCGGACACGCTGAACTTGTGG	79.	25	内部
27	EGFP 部位 27	122	GGCGAGGGCGATGCCACCTACGG	80.	GGGCACGGCAGCTTGCCTGGTGG	81.	26	内部
28	EGFP 部位 28	203	CTCGTGACCACCCCTGACCTACGG	82.	AGAAGTCGTGCTGCTTCATGTGG	83.	27	内部
29	EGFP 部位 29	337	CAACTACAAGACCCGCCCGAGG	84.	CGATGCCCTTCAGCTCGATGCGG	85.	28	内部
30	EGFP 部位 30	62	CCCATCTGGTCGAGCTGGACGG	86.	GGCATGCCCTCGGCCCTCGCCGG	87.	29	内部
31	EGFP 部位 31	100	CAAGTCAGCGTGTCCGGCGAGG	88.	GGTGGTGCAGATGAACCTTCAGGG	89.	30	内部
32	EGFP 部位 32	74	GAGCTGGACGGCGACGTAAACGG	90.	GACCAGGAATGGGCACCAACCCGG	91.	0	外部

10

20

表 2								
FokI -dCa s9 EGFP 対 #	名称	標的開 始位置 (+)	配列 (+) 部位	配 列 番 号	配列 (-) 部位	配 列 番 号	端から端 までの 「スペー サー」 距離	PAM
33	EGFP 部位 33	314	ACCATCTTCTTCAGGACGACGG	92.	CGCTCTGGACGTAGCCTTCGGG	93.	1	外部
34	EGFP 部位 34	122	GGCGAGGGCGATGCCACCTACGG	94.	CGCCGGACACGCTGAACTTGTGG	95.	2	外部
35	EGFP 部位 35	275	TTCAAGTCGCCATGCCGAAGG	96.	GTCGTGCTGCTTCATGTGGTGGG	97.	3	外部
36	EGFP 部位 36	275	TTCAAGTCGCCATGCCGAAGG	98.	TCGTGCTGCTTCATGTGGTGGG	99.	4	外部
37	EGFP 部位 37	95	GGCCACAAAGTTAGCGTGTCCGG	100.	CGTCGCCGTCCAGCTCGACCAAGG	101.	5	外部
38	EGFP 部位 38	203	CTCGTGACCACCCCTGACCTACGG	102.	CCAGGGCACGGCAGCTTGCCTGG	103.	6	外部
39	EGFP 部位 39	463	CAGCCACAAACGTCTATATCATGG	104.	TGTACTCCAGCTTGTGCCCAAGG	105.	7	外部
40	EGFP 部位 40	95	GGCCACAAAGTTAGCGTGTCCGG	106.	GCCGTCCAGCTCGACCAAGGATGG	107.	9	外部
41	EGFP 部位 41	95	GGCCACAAAGTTAGCGTGTCCGG	108.	CCGTCCAGCTCGACCAAGGATGG	109.	10	外部
42	EGFP 部位 42	101	AAGTCAGCGTGTCCGGCGAGGG	110.	CGTCGCCGTCCAGCTCGACCAAGG	111.	11	外部
43	EGFP 部位 43	350	CGCGCCGAGGTGAAGTTCGAGGG	112.	GCCGTGCTTGAAGAAGATGG	113.	12	外部

30

40

表 2								
FokI-dCa s9 EGFP 対 #	名称	標的開始位置 (+)	配列 (+) 部位	配列番号	配列 (-) 部位	配列番号	端から端までの「スペーサー」距離	PAM
44	EGFP 部位 44	174	CGGGCAAGCTGCCGTGCGCTGG	114.	CTTCAGGGTCAGCTTGCCTAGG	115.	13	外部
45	EGFP 部位 45	100	CAAGTTCAAGCGTGTCCGGCGAGG	116.	GCGTCCAGCTCGACCAAGGATGG	117.	14	外部
46	EGFP 部位 46	100	CAAGTTCAAGCGTGTCCGGCGAGG	118.	CCGTCCAGCTCGACCAAGGATGG	119.	15	外部
47	EGFP 部位 47	101	AAGTTCAAGCGTGTCCGGCGAGGG	120.	CCGTCCAGCTCGACCAAGGATGG	121.	16	外部
48	EGFP 部位 48	107	AGCGTGTCCGGCGAGGGCGAGGG	122.	CGTCGCGTCCAGCTCGACCAAGG	123.	17	外部
49	EGFP 部位 49	155	CTGAAGTTCATCTGCACCAACGG	124.	GGCATCGCCCTCGCCCTCGCCGG	125.	18	外部
50	EGFP 部位 50	106	CAGCGTGTCCGGCGAGGGCGAGG	126.	GCGTCCAGCTCGACCAAGGATGG	127.	20	外部
51	EGFP 部位 51	95	GGCCACAAGTTCAAGCGTGTCCGG	128.	GACCAGGATGGGCACCAACCCCGG	129.	21	外部
52	EGFP 部位 52	107	AGCGTGTCCGGCGAGGGCGAGGG	130.	CCGTCCAGCTCGACCAAGGATGG	131.	22	外部
53	EGFP 部位 53	337	CAACTACAAGACCCGCGCGAGG	132.	GCGCTCCTGGACGTAGCCTTCGG	133.	23	外部
54	EGFP 部位 54	337	CAACTACAAGACCCGCGCGAGG	134.	CGCTCCTGGACGTAGCCTTCGG	135.	24	外部

10

20

表 2								
FokI-dCa s9 EGFP 対 #	名称	標的開始位置 (+)	配列 (+) 部位	配列番号	配列 (-) 部位	配列番号	端から端までの「スペーサー」距離	PAM
55	EGFP 部位 55	397	GCTGAAGGCATCGACTTCAGG	136.	CCTCGAACTTCACCTCGCGCGG	137.	25	外部
56	EGFP 部位 56	100	CAAGTTCAAGCGTGTCCGGCGAGG	138.	GACCAGGATGGGCACCAACCCCGG	139.	26	外部
57	EGFP 部位 57	101	AAGTTCAAGCGTGTCCGGCGAGGG	140.	GACCAGGATGGGCACCAACCCCGG	141.	27	外部
58	EGFP 部位 58	400	GAAGGGCATCGACTTCAGGGAGG	142.	CCTCGAACTTCACCTCGCGCGG	143.	28	外部
59	EGFP 部位 59	337	CAACTACAAGACCCGCGCGAGG	144.	CTGGACGTAGCCTTCGGCATGG	145.	29	外部
60	EGFP 部位 60	307	GGAGCGCACCATCTTCAAGG	146.	AGAAGTCGTGCTGCTTCATGTGG	147.	31	外部
61	EGFP 部位 61	100	CAAGTTCAAGCGTGTCCGGCGAGG	148.	CGTCGCGTCCAGCTCGACCAAGG	149.	10	外部
62	EGFP 部位 62	286	CATGCCGAAGGCTACGTCCAGG	150.	AGAACGCGTGTGCTTCATGTGG	151.	10	外部
63	EGFP 部位 63	337	CAACTACAAGACCCGCGCGAGG	152.	TGAAGAAGATGGTGCGCTCTGG	153.	10	外部
64	EGFP 部位 64	382	GGTGAACCGCATCGAGCTGAAGG	154.	CCTCGAACTTCACCTCGCGCGG	155.	10	外部
65	EGFP 部位 65	275	TTCAAGTCGCCATGCCGAAGG	156.	GCTTCATGTGGTCGGGGTAGCGG	157.	11	外部

30

40

FokI -dCa s9 EGFP 対 #	名称	標的開始位置 (+)	配列 (+) 部位	配 列 番 号	配列 (-) 部位	配 列 番 号	端から端 までの 「スペー サー」 距離	PAM
66	EGFP 部位 66	349	CCGGCGCCGAGGTGAAGTTCGAGG	158.	GCGCTCGTCCTTGAAGAAGATGG	159	11	外部
67	EGFP 部位 67	382	GGTGAAACCGCATCGAGCTGAAGG	160.	CTCGAACITCACCTCGCGCGGG	161	11	外部
68	EGFP 部位 68	383	GTGAACCGCATCGAGCTGAAGG	162.	CCTCGAACITCAACCTCGCGCGGG	163	11	外部
69	EGFP 部位 69	520	CAAGATCCGCCACAACATCGAGG	164.	GATGCCGTTCTCTGCTTGTGG	165	11	外部
70	EGFP 部位 70	383	GTGAACCGCATCGAGCTGAAGG	166.	CTCGAACITCACCTCGCGCGGG	167	12	外部
71	EGFP 部位 71	415	CAAGGAGGACGGCARACATCTGG	168.	TCAGCTCGATGCCTTCACCAGG	169	13	外部
72	EGFP 部位 72	286	CATGCCGAAGGCTACGTCCAGG	170.	GTCGTGCTGCTTCATGTGGTGG	171	14	外部
73	EGFP 部位 73	415	CAAGGAGGACGGCAACATCTGG	172.	CAGCTCGATGCCTTCACCAGG	173	14	外部
74	EGFP 部位 74	416	AAGGAGGACGGCAACATCTGG	174.	TCAGCTCGATGCCTTCACCAGG	175	14	外部
75	EGFP 部位 75	101	AAGTTCAGCGTGTCCGGCGAGG	176.	GCGCTCAGCTCGACCAGGATGG	177	15	外部
76	EGFP 部位 76	286	CATGCCGAAGGCTACGTCCAGG	178.	TCTGCTGCTTCATGTGGTGG	179	15	外部

10

20

FokI -dCa s9 EGFP 対 #	名称	標的開始位置 (+)	配列 (+) 部位	配 列 番 号	配列 (-) 部位	配 列 番 号	端から端 までの 「スペー サー」 距離	PAM
77	EGFP 部位 77	416	AAGGAGGACGGCAACATCTGG	180.	CAGCTCGATGCCTTCACCAGG	181	15	外部
78	EGFP 部位 78	417	AGGAGGACGGCAACATCTGG	182.	TCAGCTCGATGCCTTCACCAGG	183	15	外部
79	EGFP 部位 79	524	ATCCGCCAACATCGAGGACGG	184.	GATGCCGTTCTCTGCTTGTGG	185	15	外部
80	EGFP 部位 80	106	CAGCGTGTCCGGCGAGGGCGAGG	186.	CGTCGCCGTCAGCTCGACCAGG	187	16	外部
81	EGFP 部位 81	174	CGGGCAAGCTGCCGTGCCCTGG	188.	CAGGGTCAGCTTGCCTAGGTGG	189	16	外部
82	EGFP 部位 82	286	CATGCCGAAGGCTACGTCCAGG	190.	CGTGCTGCTTCATGTGGTGG	191	16	外部
83	EGFP 部位 83	417	AGGAGGACGGCAACATCTGG	192.	CAGCTCGATGCCTTCACCAGG	193	16	外部
84	EGFP 部位 84	427	CAACATCTGGGCACAGCTGG	194.	CGATGCCCTTCAGCTCGATGCCG	195	16	外部
85	EGFP 部位 85	397	GCTGAAGGGCATCGACTTCAGG	196.	GTCGCCCTCGAACITCACCTGG	197	20	外部

30

【 0 1 1 6 】

驚くべきことに、dCas9-FokIタンパク質は、ヒトU2OS. EGFP細胞において60個のgRNA対のいずれと共に発現する際にも、検出可能なEGFP崩壊活性を示さなかった(図5E)。しかしながら、同一の60のgRNA対でのFokI-dCas9タンパク質のスクリーニングは、PAM外部の配向の片側部位から構成され、13~17bpおよび26bp(およそ1回転で13~17bp超のスペーサー長のDNAヘリックス)を備える標的部位のEGFP崩壊活性を明らかにした(図5B)。10~20bpの範囲のスペーサー長を備え、PAM外部の配向の片側部位を備える追加の25個の標的DNA部位上でのFokI-dCas9の試験は、13~18bpのスペーサー長を備える標的での効率的な切断を例証した(図5C~D)。これらの実験では、1つの部位を、それぞれ17bpまたは18bpのスペーサー長で試験し、13bpスペーサー長を備える全ての部位が活性を示したわけではなかった。TEI解析およびサンガーパターン決定

40

50

により成功して標的とされた部位のサブセット解析は、さらに、意図する位置に挿入欠失の存在を確認した。したがって、Fok I - dCas9は、対象となる完全長の標的部位を効率的に切断するために、2つの適切に配置されたgRNAにより配向できる。単純性のため、2つのFok I - dCas9融合物および2つのgRNAの複合体は、本明細書中でRNA誘導型Fok Iヌクレアーゼ(RFN)を指す。

【0117】

EGFPレポーター遺伝子を用いて最初の知見を拡張し、RFNが、内在性ヒト遺伝子の定常化したゲノム編集を実施するために使用できるかどうかを確認するために、gRNA対を、9つの異なるヒト遺伝子の12の異なる標的部位のために設計した(表2)。試験した12のRFNのうち7つは、T7EIにより判断されるように、ヒトU2OS・EGFP細胞において意図される標的部位で、高い効率の(3~40%の範囲)挿入欠失を誘導した(表2)。類似の結果を、HEK293細胞における同一の12RFN対を用いて得た(表2)。U2OS・EGFP細胞からうまく標的化したアレルのサンガーパターン決定から、予期した切断部位で、ある範囲の挿入欠失(おもに欠損)の導入が明らかとなつた(図5F)。2つの異なるヒト細胞株で観察される修飾の高い成功率および高い効率は、内在性ヒト遺伝子を修飾するRFNの強さを例証するものである。

10

【0118】

実施例2d. RFNは、それらの切断部位に対して拡張した特異性を有する

RFNが、二量体化に関連する高い認識特異性を有するかどうかを試験するために、これらヌクレアーゼが、対の中の両方のgRNAの存在に密に依存するかどうかを試験した。理想的な二量体系では、単一gRNAは、Fok I - dCas9誘導型挿入欠失を効率的に配向することが可能ではないはずである。最初の試験を実施するために、ヒトU2OS・EGFP細胞における標的部位(EGFP部位47および81)に対してFok I - dCas9 - 誘導型挿入欠失を効率的に配向することを示した、EGFPにおける2つの標的部位に配向した2対のgRNAを使用した(図5C)。VEGFA中の関連しない部位を標的とするgRNAでのこれら2つの対のそれぞれの1つまたは他のgRNAの置換は、EGFP崩壊活性の低減をもたらし(図6A)、T7EIアッセイにより判定して、検出できないレベルまで標的化した変異の低減をもたらした(図6B)。同様に、2つのgRNAのそれぞれの1つのみを使用した効果を、ヒトAPC、MLH1、およびVEGFA遺伝子におけるFok I - dCas9 - 仲介性挿入欠失を効率的に誘導する対を使用して試験し(表2)、ここでも、T7EIアッセイにより検出可能なRFN誘導型挿入欠失の損失を観察した(図6C)。これらの結果は、RFNによるゲノム編集の効率的な誘導が、完全長の標的部位に対する適切な相補性を備える2つのgRNAを必要とすることを例証する。

20

【0119】

本出願人のRFNの活性が2つのgRNAの発現に依存することを考慮すると、対中の単一gRNAのうち1つの知られているオフターゲット部位上の変異原性作用は無視できるべきものであろうと推定される。これら直接的な比較を実施することは、二量体RFNを標的化するために必要とされる2つのgRNAのうちの1つとしてそれ自体が作用できる単一gRNAにより誘導される単量体Cas9ヌクレアーゼのオフターゲット部位を知ることが必要である。単量体のCas9ヌクレアーゼオフターゲット部位が、文献でほとんど定義されていないが、本出願人がヒトVEGFA遺伝子の二量体RFN部位を標的とするために使用したgRNAのうちの1つについて5つのオフターゲット部位が以前に特定されている(実施例1)。ディープシーケンシングを使用して、これら5つのオフターゲット部位が、VEGFA - 標的化RFNが発現した細胞での変異の根拠を示したかどうかを確認した。(これらは、図6Cに示されるT7EIアッセイで使用したものと同一の細胞である)。5つすべてのオフターゲット部位での挿入欠失変異の頻度は、背景と区別できないものであった(図6Dおよび表3)。これらの結果から、RFNの使用することにより、Cas9ヌクレアーゼおよび単一gRNAにより本来誘導されたオフターゲット作用を本質的に除去できることを例証し、Fok I - dCas9と共に発現した単一g

30

40

50

R N A が、挿入欠失を効率的に誘導しないとの本出願人の観察と一致した。しかしながら、現在のところ、追加の部位上でこれらの直接的な比較を実施することは可能ではなく、このような実験は、二量体 R F N の片側部位をも標的化できる、より多くの単一 g R N A 部位のオフターゲット部位の特定を待つ必要があり、二量体 R F N は、標準的な単量体 Cas9 ヌクレアーゼと比較して特異性を高めたとの結論となった。

【表5】

表3				
遺伝子名	左の標準配列	配列番号	右の標準配列	配列番号
CCAGAAGTACGAGGCCGC APC1 CGGG	198.	TGGCAGGTGAGTGAGGCT GCAGG	199.	CGGGGGGGCTGCTGTACTTCTGAGTGGGactgtggcgaa C9tctGGAGGTGAGTTGAGTGAAGCTTCATTCTtttagccccgggg 200.
GAATACCCATCTGTCAGCT BRCA1 TCGG	201.	GGCGAACCTGTAGAGGGCG TAAGG	202.	CCGAAGCTGACAGATGGGTATTCTtttagccccgggg tagggGGGAACCTGTAGAGGGCTAAAG 203.
AATATTCAAGCAGGGCA DDB2 CAGG	204.	CTCGGCCAGGGCTGCA GCGGG	205.	CTCTGGCTGTGCTGAATATTTCGcccttttag ggtgtCTCGGGCAGGAGGGCTGCAAGCGGG 206.
CCCRAAGCCTGGCCAGGGA EMX1 GTGG	207.	GCCCCACAGGGCTTGAG CCCGG	208.	CCACTCCCTGGCAGGGCTTGGggggctggagt catggcccacaggcctTGAGCCCGG 209.
CCCTACTTCCGCTTTCACC #号1 TTGG	210.	GGAAATCCCTCTGCAECA CCTGG	211.	CCAAGGTAAAGGGAAAGTAGGGGccttcgcgcac tcatGGAAATCCCTCTGAGCACCTGG 212.
CGCTCCAGAGCCGTGCGAA FANCF - #号2 TGGG	213.	TGGAGCAAGAGGGGGC TTTTG	214.	CCCATTCGGCACGGCTCTGGAGGGGcggtgcacaa ccagtGGGGGGAAAGGGGGGGCTTTGG 215.
CGAGGAGACTGGGACTGT FES AGGG	216.	CCAGCTGCTGCGCT CCAGG	217.	CCCTACAGTCCCAGCCTCCCTGtcctatgcctc gtctCCAGTGTGCTGCCTTGCTCAGG 218.
CATAGCTACTGATTTGG GLII TGGG	219.	CCGGCCCTCCCCAGTCA GGGG	220.	CCCACACCAATCAGTGTATGGcgagccctgtct gtctCCGGCCCTCCCCAGTCAAGCGGGGG 221.
GGAAACGCTTATAGTGTCA MLH1 ACGG	222.	CAAATGTCGTTCTGGC AGTGG	223.	CCGTTGAGCATCTAGACGTTCTCTtttggtctttctgt gcggccaaaaATGTCGTTCTGTCAGGGGG 224.
CTGTTGCTGGCCATGCCAA RARAI GGGG	225.	CCTGGGGGGGGCACCTC AATGG	226.	CCGCTTGGCATGGCCAGCAACAGcgtctctggcc gacacCTGGGGGGGGCACCTCAATGG 227.
TTCGGAGCGAAAACCAAGA RUNX CAGG	228.	GAGTCCCCGGCTTCAAGA AGGG	229.	CCTGTTGGTTCTGCTCCGAAAgtaaaaat attGAGTCCCCGGCTTCAAGAAGGG 230.
GGCCCGGGTGCAGCTGGGGC SS18 CGGG	231.	TGCTGGAAATCAGCACTG TTTGG	232.	CGGGGGGGAGTCGACCCGGGGCgaggcgggggg qgcctGCTGGGGAAATCAGCAGTGTGTTGG 233.
GGGGGGGGAGTTGCTC VEGFA - #号1 CTGG	234.	TCCCTCTTATGCCAGAGC CGGG	235.	CCAGGAGCAAACCCCCCCCCtttccaaagc ccattCCCTCTTTAGCCAGGGGGGG 236.
GCGCGGGGGGGGGAG VEGFA - #号2 GTGG	237.	GGCGAGGGGGGGGG GCCGG	238.	CCACCTCTCCCCGGGGGGGGGGGG cggcggGGGGGGGGGGGGGGGGGG 239.
CCGTCTGGCACACCCGGCT VEGFA - #号3 CTGG	240.	CTCGGCCACCACAGGGAA GCTGG	241.	CCAGAGGCCACCACAGGGAAAGCTGG ggcgctCGGGCACACAGGGAA 242.

表 3

遺伝子名	TTEI アッセイに使用されるプラットフォーム	配列番号	TTEI アッセイに使用されるプライマー 1	配列番号	DNA の重畠 (はさみ) / リバース	DNA の重畠 (はさみ) / リバースのプロトコル	アンブリッジのプロトコル	アンブリッジのプロトコル (b p)	TTEI の薬物の生産切断物の寸法 (b p)
APCI	GGCTGTGGAAAGCCAGCAA C	243.	AAGCAGGGGCCA ACTGGAG	244.	ハサミ	ハサミ	ハサミ	ハサミ	447/187
BRCA1	GCGCGGAATTACAGATAA ATTAaaa	245.	AGGTCCCATTCTCTC TCATACATACCA	246.	ハサミ	ハサミ	ハサミ	ハサミ	454/297
DBB2	ACCGCCCCCTGGCACCAAC GGGAGCAGCTGGTAGAGGG G	247.	CGGAGCTCATCTG CTTCTCTGT	248.	ハサミ	ハサミ	ハサミ	ハサミ	456/171
EMX1	GCCTCTACATCTGCTCTCCCT CCA	249.	GGGAAGGGGGAC ACTGGGGA	250.	ハサミ	ハサミ	ハサミ	ハサミ	634 447/187
FANCF- 部位 1	GCCTCTACATCTGCTCTCCCT CCA	251.	GGGCCGGGAAAG AGTTGCTG	252.	ハサミ	ハサミ	ハサミ	ハサミ	456/171
FANCF- 部位 2	GGGGAGGGAGGCTCCAGGT T	253.	GGGCCGGGAAAG AGTTGCTG	254.	ハサミ	ハサミ	ハサミ	ハサミ	480/249
FES	CCTTACCCCTCCCCTCACTC A	255.	GGCACAAATGGCTC OCAAGCA	256.	ハサミ	ハサミ	ハサミ	ハサミ	634 361/273
GLJ1	ATATCCTCTAGGTAGCGGG CAGTAGCC	257.	AGAAGGGGGGGC CAGACAGT	258.	ハサミ	ハサミ	ハサミ	ハサミ	466/168
MLH1	CCCAGGAAAAGTGCCAGCT A	259.	TCTCGGGGGAGAG CGCTAAA	260.	ハサミ	ハサミ	ハサミ	ハサミ	634 361/273
RARA1	ACTGGAA CA	261.	TGATGGTCACCCC AACTGGAA	262.	ハサミ	ハサミ	ハサミ	ハサミ	634 361/273
RUNX	AAGGGCGCGCTGGCTTTT GGGATGCAGGGACGGTCAA G	263.	CCAGCACAACTTA CTGGCACTTGA	264.	ハサミ	ハサミ	ハサミ	ハサミ	395/238
SS18		265.	GGGGCCCCATCCCC TAGAGAAA	266.	ハサミ	ハサミ	ハサミ	ハサミ	395/238
VEGFA- 部位 1	TCCAGATGGCACATTTGTCAG G	267.	AGGGAGCAGGAA AGTGAGGT	268.	ハサミ	ハサミ	ハサミ	ハサミ	332/278
VEGFA- 部位 2	AGAGAAGTICGAGGAAGAGA GAG	269.	CAGCAGAAAGTTTC ATGGTTTGG	270.	ハサミ	ハサミ	ハサミ	ハサミ	338/193
VEGFA- 部位 3	TCCAGATGGCACATTTGTCAG	271.	AGGGAGCAGGAA AGTGAGGT	272.	ハサミ	ハサミ	ハサミ	ハサミ	482/274

表 3			
遺伝子名	プライマー番号	プライマー番号	アリーブシーケンシングに使用されるプライマー番号
APC1 マサード			
BRCAl			
DDB2	CGATGGCTCCAAAGAACGC	273.	GCAGGGTAGAATGCCACAGCCG
EMX1			274.
FANCF - 部位 1	GCCCCAGACTCAAGGAACACG	275.	AGGTAGTGCCTTGAGACCGCC
FANCF - 部位 2	CATCCCATCGGGCTTTGGTC	277.	CCGGGAAAGAGAGTGTGCTGCAC
FES	CTCCCCGTCAGTCCATC	279.	CCTGCAGGGACATGTGGTGA
GLI1			280.
MLH1			
RAPAI			
RUNX	TAGGGCTAGAGGGGTGAGGC	281.	CCGAGGGTAGAAACAAGCTGCCC
SS18			282.
VEGFA - 部位 1	ATGAGGGCTCCAGATGGCAC	283.	TTCACCCAGCTTCCCTGTGG
VEGFA - 部位 2			
VEGFA - 部位 3			

【 0 1 2 0 】

実施例 2 e . 単量体 C a s 9 ニッカーゼは、単一 g R N A / F o k I - d C a s 9 複合体よりも高い比率の変異原性を誘導する。

上述のように、対形成した C a s 9 ニッカーゼの手法の重要な脆弱性は、単一の単量体ニッカーゼが、ある特定の標的部位で、挿入欠失変異を高頻度で誘導する可能性があることである（実施例 1 および R a n e t a l . , C e l l 1 5 4 , 1 3 8 0 - 1 3 8 9 (2 0 1 3) ; M a l i e t a l . , N a t B i o t e c h n o l 1 3 1 , 8 3 3 - 8 3 8 (2 0 1 3) ; C h o e t a l . , G e n o m e R e s (2 0 1 3) ; a n d M a l i e t a l . , S c i e n c e 3 3 9 , 8 2 3 - 8 2 6 (2 0 1 3) 参照）。対形成される C a s 9 ニッカーゼ系の二量体化 - 依存性の欠損は、2つの単量体ニッカーゼがゲノムの他の場所に望ましくない挿入欠失変異をそれぞれ作製し得るため、オフターゲットの潜在的な原因である。 R F N は、二量体化依存性 F o k I ヌクレアーゼを使用して変質を誘導するため、これらの融合物は、単量体 C a s 9 ニッカーゼで観察されるものと比較して、1つの g R N A のみの存在下で

の望ましくない挿入欠失活性がより少ないと仮定される。

【0121】

この仮定を試験するために、FokI-dCas9およびCas9ニッカーゼの活性を、6つの二量体ヒト遺伝子標的部位での単一gRNAの存在下で比較した（合計12の片側部位、表4）。これら特定の部位は、対の中の1つのみおよび／または他のgRNAにより配向される単量体Cas9ニッカーゼが、これらの標的に挿入欠失変異を誘導し得るため、選択された。ディープシーケンシングを使用して、FokI-dCas9またはCas9ニッカーゼのゲノム編集活性を、1つまたは他のgRNAのうち両方または1つのみが存在する中で評価した。FokI-dCas9およびCas9ニッカーゼの両方は、2つのgRNAの存在する中、高い効率で、6つすべての標的部位で挿入欠失を誘導した（表5）。仮説として、12個の単一gRNAにより配向した単量体Cas9ニッカーゼは、0.0048%～3.04%の範囲の頻度で挿入欠失を誘導した（図7Aおよび表5）。対照的に、同一の12個の単一gRNAにより配向したFokI-dCas9は、0.0045～0.473%の範囲のより低い頻度で挿入欠失を誘導した（図7Aおよび表5）。これらのデータを直接比較することにより、FokI-dCas9は、12個の単一gRNAのうち10個でCas9ニッカーゼよりも低い頻度で、挿入欠失を誘導した（図7Aおよび表5）。さらに、FokI-dCas9は、単一gRNA比率と、対形成したgRNAの比率を比較する際に、12個の片側部位のうち11個でCas9ニッカーゼよりも多くの挿入欠失の頻度の倍数減少を示した（図7B）。

【表6】

表4

染色体	位置	部位	FokI-dC as9 挿入欠失	FokI-dC as9 の合計	FokI-dC as9 挿入欠失の頻度(%)	tdToma to 対照 挿入欠失	tdTo ma合 計	tdToma to 挿入欠失の頻度(%)
6	4373729 0	VEGFA 部位1	35000	150158	23.3087 8	10	2581 08	0.0038 7
15	6563753 1	OT1-3	1	169681	0.00058	1	1398 47	0.0007 1
12	1316901 82	OT1-4	4	190111	0.00210	5	1397 62	0.0035 7
12	1988060	OT1-6	3	258976	0.00115	2	1781 62	0.0011 2
1	9934764 5	OT1-11	4	235853	0.00169	4	1862 87	0.0021 4
17	3979632 2	OT1-30	1	261605	0.00038	1	2868 50	0.0003 4

10

20

30

【表7】

表5. 単一および対のgRNA(図7に提示されるものと同一のデータ)を用いた、6つの部位でのFokI-dCas9、Cas9n、およびtdTomato対照のディープシーケンシング

ヌクレアーゼ の種類または 対照	部位	ガイド RNA	染色体	位置	挿入 欠失	合計	パーセン テージ
FokI-dCas9	VEGFA 部位 1	両方	6	43737290	35000	150158	23.3088
FokI-dCas9	VEGFA 部位 1	左	6	43737290	5	95476	0.0052
FokI-dCas9	VEGFA 部位 1	右	6	43737290	9	91962	0.0098
FokI-dCas9	DDB2	両方	11	47236820	11303	50062	22.5780
FokI-dCas9	DDB2	左	11	47236820	311	85726	0.3628
FokI-dCas9	DDB2	右	11	47236820	153	95050	0.1610
FokI-dCas9	FANCF 部位 1	両方	11	22647331	65846	195311	33.7134
FokI-dCas9	FANCF 部位 1	左	11	22647331	19	27487	0.0691
FokI-dCas9	FANCF 部位 1	右	11	22647331	845	225154	0.3753
FokI-dCas9	FANCF 部位 2	両方	11	22647138	27743	120314	23.0588
FokI-dCas9	FANCF 部位 2	左	11	22647138	989	205832	0.4805
FokI-dCas9	FANCF 部位 2	右	11	22647138	142	165130	0.0860
FokI-dCas9	FES	両方	15	91428181	14260	126912	11.3254
FokI-dCas9	FES	左	15	91428181	4	143877	0.0028
FokI-dCas9	FES	右	15	91428181	7	145495	0.0048
FokI-dCas9	RUNX1	両方	21	36421217	61057	136164	44.8408
FokI-dCas9	RUNX1	左	21	36421217	222	162636	0.1365
FokI-dCas9	RUNX1	右	21	36421217	109	169122	0.0645
Cas9n	VEGFA 部位 1	両方	6	43737290	14294	99036	14.4331
Cas9n	VEGFA 部位 1	左	6	43737290	573	82316	0.6961
Cas9n	VEGFA 部位 1	右	6	43737290	315	101957	0.3090
Cas9n	DDB2	両方	11	47236820	6673	31168	21.4098
Cas9n	DDB2	左	11	47236820	1680	56019	2.9990
Cas9n	DDB2	右	11	47236820	172	42424	0.4054
Cas9n	FANCF 部位 1	両方	11	22647331	66827	193111	34.6055
Cas9n	FANCF 部位 1	左	11	22647331	1565	109029	1.4354
Cas9n	FANCF 部位 1	右	11	22647331	2457	109289	2.2482

表 5. 単一および対の g RNA (図 7 に提示されるものと同一のデータ) を用いた、6 つの部位での FokI-dCas9、Cas9n、および tdTomato 対照のディープシーケンシング

ヌクレアーゼ の種類または 対照	部位	ガイド RNA	染色体	位置	挿入 欠失	合計	パーセン テージ
1							
Cas9n	FANCF 部位	両方	11	22647138	17007	111468	15.2573
2							
Cas9n	FANCF 部位	左	11	22647138	120	100591	0.1193
Cas9n	FANCF 部位	右	11	22647138	1063	93162	1.1410
2							
Cas9n	FES	両方	15	91428181	16529	126597	13.0564
Cas9n	FES	左	15	91428181	6	125196	0.0048
Cas9n	FES	右	15	91428181	23	46102	0.0499
Cas9n	RUNX1	両方	21	36421217	80029	216800	36.9137
Cas9n	RUNX1	左	21	36421217	1106	108670	1.0178
Cas9n	RUNX1	右	21	36421217	2169	121413	1.7865
tdTomato 対 照 (-)	VEGF 部位	なし	6	43737290	29	313517	0.0092
tdTomato 対照 (-)	FANCF 部位	なし	11	22647331	18	578378	0.0031
tdTomato 対照 (-)	FANCF 部位	なし	11	22647138	81	393821	0.0206
tdTomato 対照 (-)	FES	なし	15	91428181	21	410620	0.0051
tdTomato 対照 (-)	DDB2	なし	11	47236820	14	165314	0.0085
tdTomato 対照 (-)	RUNX1	なし	21	36421217	13	511977	0.0025

【 0 1 2 2 】

また、ディープシーケンシングの実験は、標的部位の中の特定の位置の点変異の導入といった、特定の单量体の Cas9 ニッカーゼの従来説明されておらず、予期していない副作用を明らかにした。VEGFA 標的部位の「右」の片側部位に対する単一 g RNA と共に発現した Cas9 ニッカーゼは、10.5% の頻度で、認識部位の位置 15 に塩基置換を誘導した (図 8 A)。類似の結果が、FACF 標的部位 1 の「右」の片側部位 (位置 16 で 16.3% の変異頻度) (図 8 B) または RUNX1 標的部位の「右」の片側部位 (位置 17 で 2% の変異頻度) (図 8 C) に配向される Cas9 ニッカーゼおよび単一 g RNA で観察された。これらの位置での点変異は、Cas9 ニッカーゼまたは g RNA が細胞で発現していない对照試料のバックグラウンドレベルを超えて観察されるものではなかった (図 8 A ~ 8 C)。興味深いことに、この高頻度変異が観察された 3 つの部位のうち 2 つでは、観察された置換基の大部分が、非標的 DNA 鎮上の C の G へ塩基転換である。これら点変異が観察された位置は、dCas9 / gRNA / 標的 DNA 複合体において in vitro で P1 ヌクレアーゼに感受性があると観察された標的部位の鎖 - 分離領域内にある。重要なことに、これらの点変異は、FokI - dCas9 タンパク質および同一の gRNA を発現する細胞において、より低い頻度 (5 ~ 100 倍低い) で起こる (図 8 A ~ C)。全体的に見て、単一 gRNA により配向される FokI - dCas9 ヌクレアーゼは、概して、マッチした単一の Cas9 ニッカーゼよりも低い頻度で変異原性の挿入欠失および点変異を誘導する。

【 0 1 2 3 】

実施例 2 f . 二量体 RFN は高い度合いの特異性を有する。

30

40

50

2つのgRNAにより配向される二量体RFNは、ヒトの細胞中で明らかなオフターゲット変異を誘導すると予期されるものではない。2つの片側部位から構成される完全長の配列を切断する1対のgRNAにより配向されるRFNは、標的部位のDNAのうち最大44bpを特定すると予測される。この長さの配列は、偶然であるが、ほとんど常に固有である（標的が複製したゲノム配列にある特定の状況を除く）。さらに、この完全長部位に対して、ゲノムで最も近接してマッチした部位は、ほとんどの場合、多数のミスマッチを保有しており、それが次いで、RFN二量体による切断活性を最小限にするか、または消失させると予期される。実際に、この試験でのRFNでの標的化が成功した15の完全長の配列に0～16個のミスマッチ（12～17bpの長さのスペーサーを可能にする）を有するヒトゲノムの全ての部位を特定した。この解析から、全ての15個の完全長配列は固有であり、ゲノム中のほとんど近接して一致した部位は、7～12のミスマッチの範囲であることが示された（表6）。この数のミスマッチを含む部位は、RFNにより効率的に変異誘発されるものではなく、この仮説を確認するためのさらなる試験に興味がもたれる。全体的にみて、二量体RFNは、ヒト細胞において高い度合いの特異性を有するが、特異性の最終的な特徴付けは、ゲノム全体を通してRFNの特異性を包括的に定義できる公平な方法の開発が待たれる。

【表8】

表6：定義される数のミスマッチを有するヒトゲノムにおいて候補となるFokI-dCas9オフターゲット部位の頻度

遺伝子	0	7	8	9	10	11	12	13	14	15	16	
APC	1					1	2	16	74	414	2254	10
BRCA1	1	1						5	20	164	983	
DDB2	1						2	7	58	267	1335	
EMX1	1			1		2	8	40	175	828	3494	
FANCF	1						2	4	44	298	1639	
FANCF	1						2	12	79	358	1718	
FES	1					3	8	32	191	939	4505	
GLI1	1					2	1	7	69	343	1711	
MLH1	1						2	5	22	96	643	
RARA	1				1	2	8	39	187	698	2849	20
RUNX1	1							3	25	145	800	
SS18	1						1	2	6	39	280	1207
VEGFA-1	1					1	2	3	22	103	543	2676
VEGFA-2	1					4	9	99	447	1675	5608	18599
VEGFA-3	1							3	20	120	623	2783

【0124】

参考文献

- Cheng, A. W., Wang, H., Yang, H., Shi, L., Katz, Y., Theunissen, T. W., Rangarajan, S., Shivalilaa, C. S., Dadon, D. B., and Jaenisch, R. Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. *Cell Res.* 23, 1163-1171. (2013).
 Cho, S. W., Kim, S., Kim, J. M. & Kim, J. S. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat Biotechnol* 31, 230-232 (2013).
 Cong, L. et al. Multiplex genome engineerin 50

- g using CRISPR/Cas systems. *Science* 339, 819-823 (2013).
- Cradick, T. J., Fine, E. J., Antico, C. J., and Bao, G. CRISPR/Cas9 systems targeting beta-globin and CCR5 genes have substantial off-target activity. *Nucleic Acids Res.* (2013).
- Dicarlo, J. E. et al. Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. *Nucleic Acids Res.* (2013). 10
- Ding, Q., Regan, S. N., Xia, Y., Oostrom, L. A., Cowan, C. A., and Musunuru, K. Enhanced efficiency of human pluripotent stem cell genome editing through replacing TALENs with CRISPRs. *Cell Stem Cell* 12, 393-394. (2013).
- Fisher, S., Barry, A., Abreu, J., Minie, B., Nolan, J., Delorey, T. M., Young, G., Fennell, T. J., Allen, A., Ambrogio, L., et al. A scalable, fully automated process for construction of sequence-ready human exome targeted capture libraries. *Genome Biol* 12, R1. (2011). 20
- Friedland, A. E., Tzur, Y. B., Esvelt, K. M., Colaiacovo, M. P., Church, G. M., and Calarco, J. A. Heritable genome editing in *C. elegans* via a CRISPR-Cas9 system. *Nat Methods* 10, 741-743. (2013).
- Fu, Y., Foden, J. A., Khayter, C., Maeder, M. L., Reyon, D., Joung, J. K., and Sander, J. D. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol* 31, 822-826. (2013). 30
- Gabriel, R. et al. An unbiased genome-wide analysis of zinc-finger nuclease specificity. *Nat Biotechnol* 29, 816-823 (2011).
- Gilbert, L. A., Larson, M. H., Morsut, L., Liu, Z., Brar, G. A., Torres, S. E., Stern-Ginossar, N., Brandman, O., Whitehead, E. H., Doudna, J. A., et al. (2013). CRISPR-Mediated Modular RN 40 A-Guided Regulation of Transcription in Eukaryotes. *Cell* 154, 442-451.
- Gratz, S. J. et al. Genome engineering of *Drosophila* with the CRISPR RNA-guided Cas9 nuclease. *Genetics* (2013).
- Hockemeyer, D. et al. Genetic engineering of human pluripotent cells using TALE nucleases. *Nat Biotechnol* 29, 731-734 (2011).
- Horvath, P. & Barrangou, R. CRISPR/Cas, the immune system of bacteria and archaea. *Sci* 50

ence 327, 167 - 170 (2010).
Hsu, P. D., Scott, D. A., Weinstein, J. A., Ran, F. A., Konermann, S., Agarwala, V., Li, Y., Fine, E. J., Wu, X., Shalem, O., et al. DNA targeting specificity of RNA-guided Cas9 nuclease. *Nat Biotechnol* 31, 827 - 832. (2013).
Hwang, W. Y. et al. Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat Biotechnol* 31, 227 - 229 (2013).
Hwang, W. Y., Fu, Y., Reyon, D., Maeder, M. L., Kaini, P., Sander, J. D., Joung, J. K., Peterson, R. T., and Yeh, J. R. Heritable and Precise Zebrafish Genome Editing Using a CRISPR-Cas System. *PLoS One* 8, e68708. (2013a).
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).
Jinek, M. et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816 - 821 (2012).
Jinek, M. et al. RNA-programmed genome editing in human cells. *Elife* 2, e00471 (2013).
Li, D., Qiu, Z., Shao, Y., Chen, Y., Guan, Y., Liu, M., Li, Y., Gao, N., Wang, L., Lu, X., et al. Heritable gene targeting in the mouse and rat using a CRISPR-Cas system. *Nat Biotechnol* 31, 681 - 683. (2013a).
Li, W., Teng, F., Li, T., and Zhou, Q. Simultaneous generation and germline transmission of multiple gene mutations in rat using CRISPR-Cas systems. *Nat Biotechnol* 31, 684 - 686. (2013b).
Maeder, M. L., Linder, S. J., Cascio, V. M., Fu, Y., Ho, Q. H., and Joung, J. K. CRISPR RNA-guided activation of endogenous human genes. *Nat Methods* 10, 977 - 979. (2013).
Mali, P., Aach, J., Stranges, P. B., Esvelt, K. M., Moosburner, M., Kosuri, S., Yang, L., and Church, G. M. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat Biotechnol* 31, 833 - 838. (2013a).
Mali, P., Esvelt, K. M., and Church, G. M. Cas9 as a versatile tool for engineering biology. *Nat Methods* 10, 957 - 963. (2013b).
Mali, P. et al. RNA-guided human genome engineering via Cas9. *Science* 339, 823 - 826 (2013c).

10

20

30

40

50

- Pattanayak, V., Lin, S., Guilinger, J.P., Ma, E., Doudna, J.A., and Liu, D.R. High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat Biotechnol* 31, 839-843. (2013).
- Pattanayak, V., Ramirez, C.L., Joung, J.K. & Liu, D.R. Revealing off-target cleavage specificities of zinc-finger nucleases by *in vitro* selection. *Nat Methods* 8, 765-770 (2011).
- Perez, E.E. et al. Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. *Nat Biotechnol* 26, 808-816 (2008).
- Perez-Pinera, P., Kocak, D.D., Vockley, C.M., Adler, A.F., Kabadi, A.M., Polstein, L.R., Thakore, P.I., Glass, K.A., Ousterout, D.G., Leong, K.W., et al. RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nat Methods* 10, 973-976. (2013).
- Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., and Lim, W.A. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* 152, 1173-1183. (2013).
- Ran, F.A., Hsu, P.D., Lin, C.Y., Gootenberg, J.S., Konermann, S., Trevino, A.E., Scott, D.A., Inoue, A., Matoba, S., Zhang, Y., et al. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* 154, 1380-1389. (2013).
- Reyon, D. et al. FLASH assembly of TALENs for high-throughput genome editing. *Nat Biotech* 30, 460-465 (2012).
- Sander, J.D., Maeder, M.L., Reyon, D., Voytas, D.F., Joung, J.K., and Dobbs, D. ZiFiT (Zinc Finger Targeter): an updated zinc finger engineering tool. *Nucleic Acids Res* 38, W462-468. (2010).
- Sander, J.D., Ramirez, C.L., Linder, S.J., Pattanayak, V., Shores, N., Ku, M., Foden, J.A., Reyon, D., Bernstein, B.E., Liu, D.R., et al. *In silico* abstraction of zinc finger nucleic acid cleavage profiles reveals an expanded landscape of off-target sites. *Nucleic Acids Res.* (2013).
- Sander, J.D., Zaback, P., Joung, J.K., Voytas, D.F., and Dobbs, D. Zinc Finger Targeter (Zi) 50

FiT) : an engineered zinc finger / target site design tool. *Nucleic Acids Res* 35, W599 - 605. (2007).

Shen, B. et al. Generation of gene-modified mice via Cas9 / RNA-mediated gene targeting. *Cell Res* (2013).

Sugimoto, N. et al. Thermodynamic parameters to predict stability of RNA / DNA hybrid duplexes. *Biochemistry* 34, 11211 - 11216 (1995).

Terns, M. P. & Terns, R. M. CRISPR-based adaptive immune systems. *Curr Opin Microbiol* 14, 321 - 327 (2011).

Wang, H. et al. One-Step Generation of Mice Carrying Mutations in Multiple Genes by CRISPR/Cas-Mediated Genome Engineering. *Cell* 153, 910 - 918 (2013).

Wiedenheft, B., Sternberg, S. H. & Doudna, J. A. RNA-guided genetic silencing systems in bacteria and archaea. *Nature* 482, 331 - 338 (2012).

Yang, L., Guel, M., Byrne, S., Yang, J. L., De Los Angeles, A., Mali, P., Aach, J., Kim-Kiselak, C., Briggs, A. W., Rios, X., et al. (2013). Optimization of scarless human stem cell genome editing. *Nucleic Acids Res* 41, 9049 - 9061.

【0125】

その他の実施形態

ここまで本発明をその詳細な説明と関連させて記載してきたが、上記説明は例示を目的とするものであり、添付の「特許請求の範囲」の範囲によって定められる本発明の範囲を限定するものではないことを理解するべきである。その他の態様、利点および改変は以下の特許請求の範囲内にある。

本発明は以下を提供する。

[1]

RNA誘導型FokIヌクレアーゼ(RFN)融合タンパク質であって、触媒的に不活性であるCRISPR関連9(dCas9)のアミノ末端に融合するFokI触媒ドメイン配列を含み、任意に、介在リンカーを備える、RNA誘導型FokIヌクレアーゼ(RFN)融合タンパク質。

[2]

2~30個のアミノ酸のリンカーを含む、[1]に記載の融合タンパク質。

[3]

前記リンカーが、Gly₄Serを含む、[2]に記載の融合タンパク質。

[4]

FokI触媒ドメインが、配列番号4のアミノ酸388~583、または408~583を含む、[1]に記載の融合タンパク質。

[5]

dCas9が、D10、E762、H983、またはD986での変異、およびH840またはN863での変異を含む、[1]に記載の融合タンパク質。

[6]

10

20

30

40

50

d C a s 9 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 Y、または H 8 4 0 N

での変異を含む、[5] に記載の融合タンパク質。

[7]

[1] ~ [6] に記載の融合タンパク質をコードする核酸。

[8]

[7] に記載の核酸を含むベクター。

[9]

[1] ~ [6] に記載の融合タンパク質を発現する宿主細胞。

[1 0]

細胞において、ゲノム配列の配列特異的な崩壊を誘導する方法であって、前記方法が、[1] ~ [6] に記載の R N A 誘導型 F o k I ヌクレアーゼ (R F N) 融合タンパク質、および好ましくは 0 ~ 3 1 ヌクレオチド離れた 2 つの標的ゲノム配列に前記 R F N を配向するガイド R N A を、前記細胞に発現すること、または前記細胞と接触させることとを含み、好ましくは、前記 2 つの標的配列が、それぞれ 3' 末端で P A M 配列を有する、方法。

[1 1]

前記 2 つの標的ゲノム配列が、1 0 ~ 2 0 塩基対、好ましくは、1 3 ~ 1 7 塩基対離れている、[1 0] に記載の方法。

[1 2]

前記ガイド R N A が、

(a) 2 つの單一ガイド R N A であって、1 つの單一ガイド R N A が第 1 の鎖を標的とし、他方のガイド R N A が相補鎖を標的とし、F o k I が各鎖を切断して、反対の D N A 鎖上に 1 対のニックをもたらし、それにより二重鎖を切断する、2 つの單一ガイド R N A 、または

(b) t r a c r R N A および 2 つの c r R N A であって、1 つの c r R N A が第 1 の鎖を標的とし、他方の c r R N A が相補鎖を標的とし、F o k I が各鎖を切断して、反対の D N A 鎖上に 1 対のニックをもたらし、それにより二重鎖を切断する、t r a c r R N A および 2 つの c r R N A

である、[1 0] に記載の方法。

[1 3]

前記 2 つのガイド R N A が、それぞれ、標的ゲノム配列の 1 7 ~ 2 0 個のヌクレオチドと相補的な相補領域を含む、[1 0] に記載の方法。

[1 4]

挿入欠失変異が、前記 2 つの標的配列の間に誘導される、[1 0] ~ [1 3] のいずれか 1 項に記載の方法。

[1 5]

細胞における R N A 誘導型ゲノム編集の特異性が増大する、[1 0] ~ [1 4] のいずれか 1 項に記載の方法。

[1 6]

細胞における R N A 誘導型ゲノム編集の特異性を増大させる方法であって、前記細胞を、[1] ~ [6] に記載の R N A 誘導型 F o k I ヌクレアーゼ (R F N) 融合タンパク質と接触させることを含む、方法。

10

20

30

40

【図1】

CRISPR/Cas9 - RGN

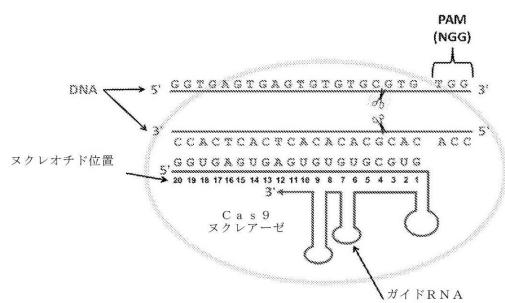


図1

【図2 A】

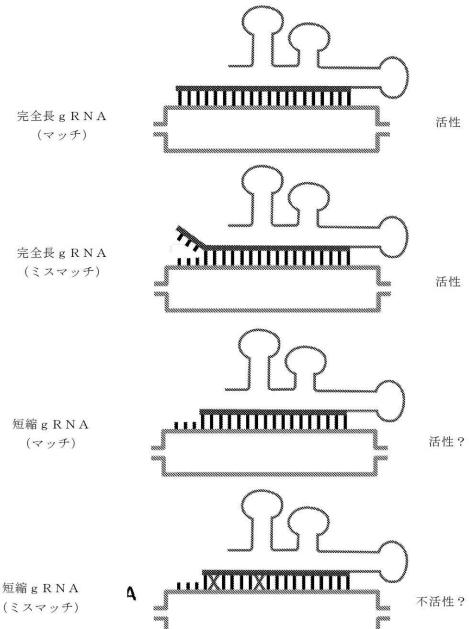


図2 A

【図2 B】

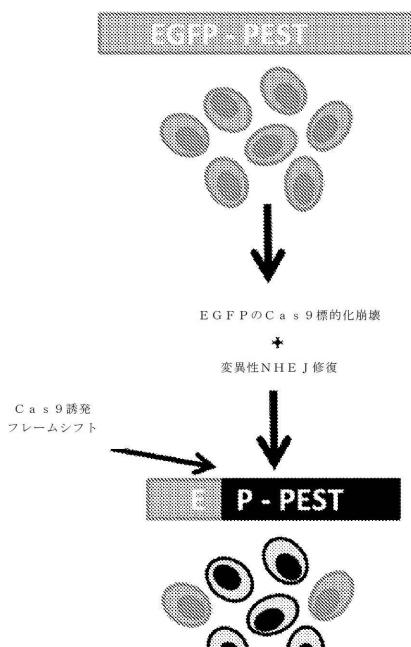


図2 B

【図2 C】

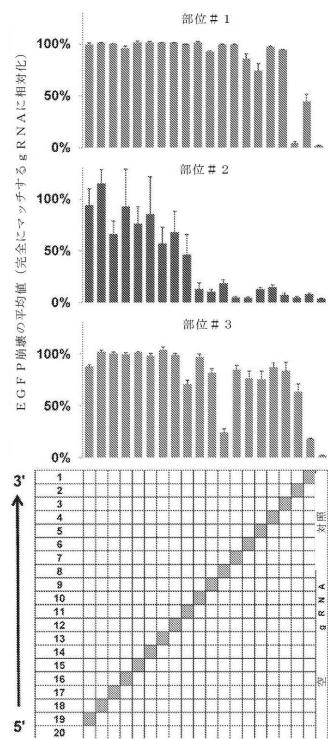


図2 C

【図 2 D】

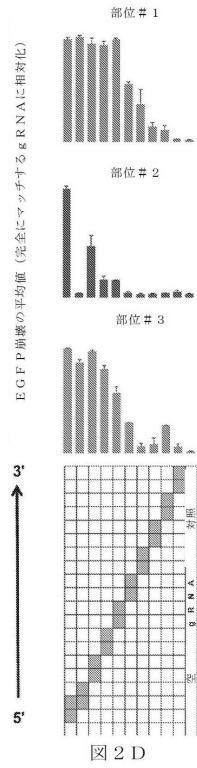


図 2 D

【図 2 E】

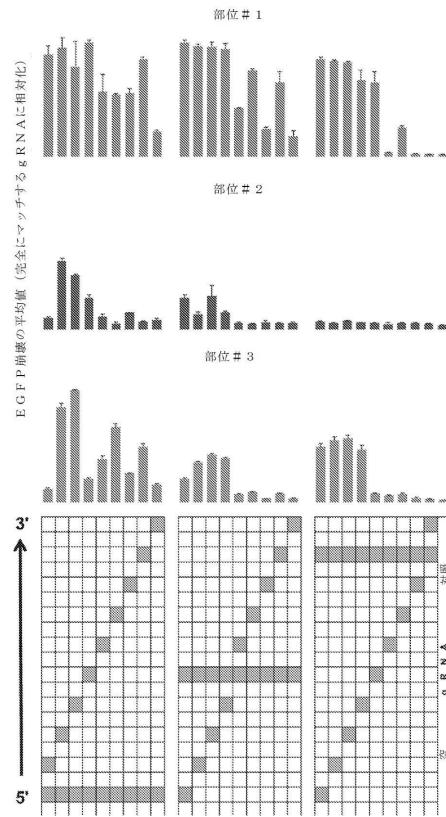


図 2 E

【図 2 F】

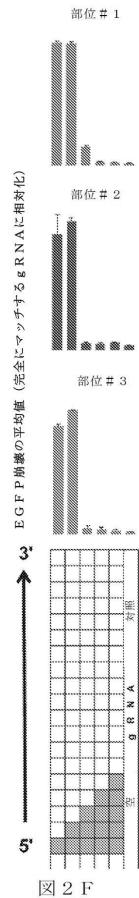


図 2 F

【図 2 G】

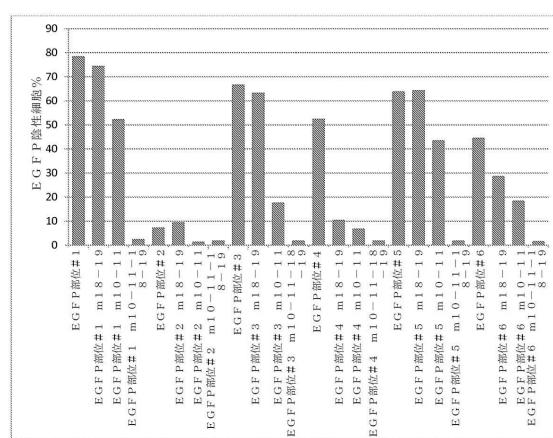


図 2 G

【 図 2 H 】

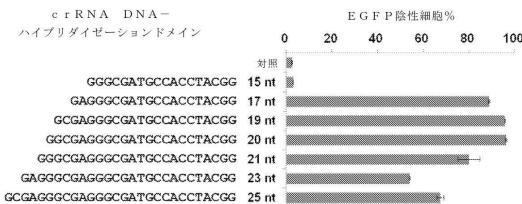


図 2 H

【図3A】

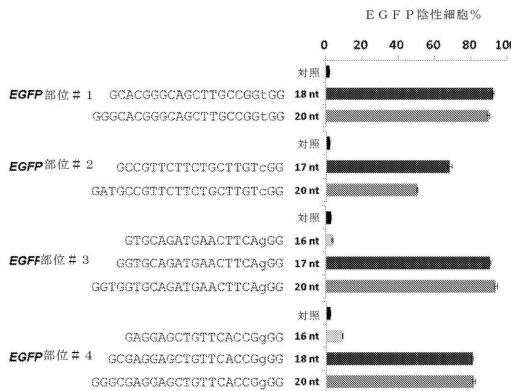


図3A

【図3C】

EMK1	短縮 RNA	野生型	x2
GAACTGGGAGGAGGAGGCTGAGTCGGAGCAGAAGGGTCCCGATGCTAACCGGTG92			
GAAGCTGGGAGGAGG		A345	
GAAGCTGGAGGGAGG		TCACCGGGTG	A181
GAAGCTGGAGGGAGG			A138
GAAGCTGGAGGGAGG			A126
GAAGCTGGAGGGAGG	GGGCTCCCATCACATCACCGGTG		A114
GAAGCTGGAGGGAGG			A191
GAAGCTGGAGGA		CGG	A53
GAAGCTGGAGGA	CCCATCACATCACCGGTG		A28
GAAGCTGGAGGA	TGGCACACATCACACCGGTG		A27
GAAGCTGGAGGA	CITTCACATCACACCGGTG		A25
GAAGCTGGAGGA	TCCGATCACATCACACCGGTG		A21 x2
GAAGCTGGAGGA	TCCGATCACATCACACCGGTG		A13
GAAGCTGGAGGA	TCCGATCACATCACACCGGTG		A9
GAAGCTGGAGGA	CCCATCACATCACACCGGTG		A8
GAAGCTGGAGGA	GCTCCCATCACATCACACCGGTG		A6
GAAGCTGGAGGA	TCCTGCACATCACACCGGTG		A6
GAAGCTGGAGGA	AGAGAGGGCTCATCACATCACACCGGTG		A3 x3
GAAGCTGGAGGA	AGAGAGGGCTCATCACATCACACCGGTG		A2
GAAGCTGGAGGA	AGAGAGGGCTCATCACATCACACCGGTG		A2
GAAGCTGGAGGA	AGAGAGGGCTCATCACATCACACCGGTG		A2
GAAGCTGGAGGA	AGAGAGGGCTCATCACATCACACCGGTG		+2

图 3 C

【図3B】

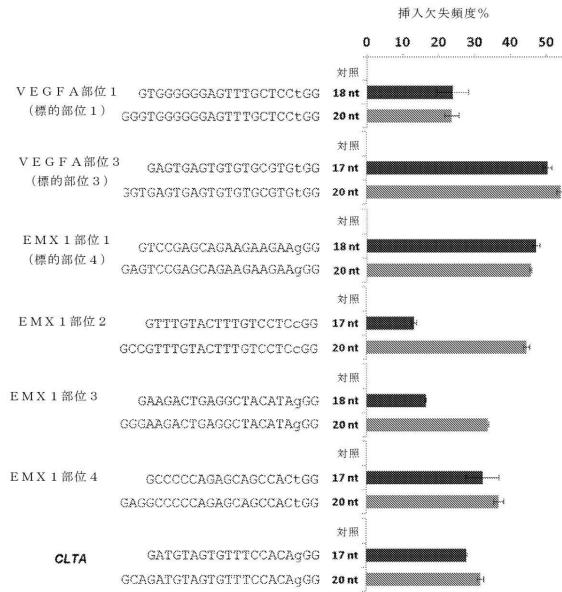


図 3 B

【図3D】

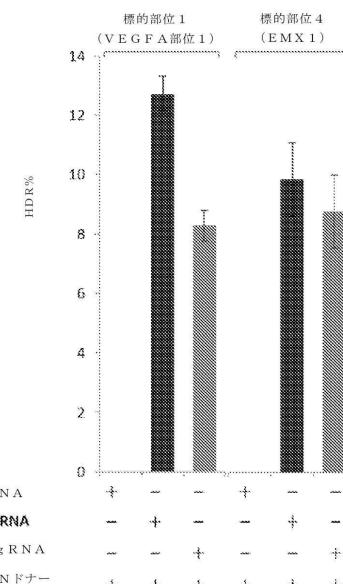


图 3 D

【図3 E】

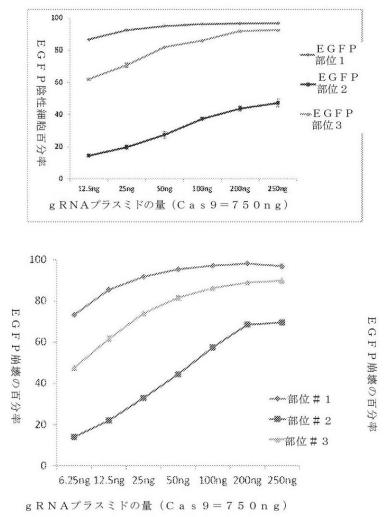


図3 E

【図3 F】

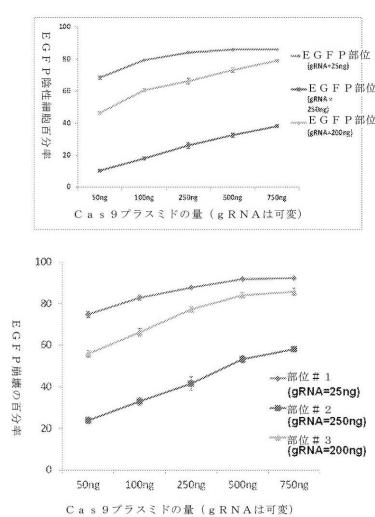


図3 F

【図4 A】

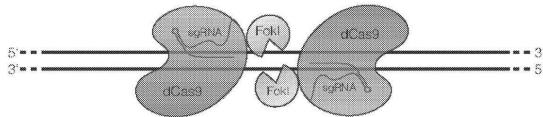


図4 A

【図4 B】

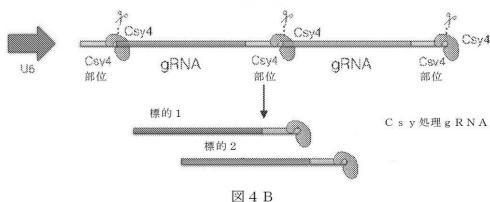
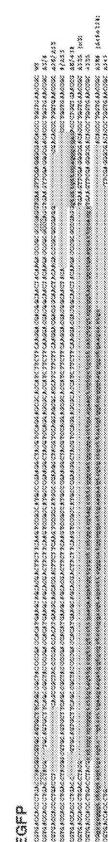


図4 B

【図4 C】



【図 5 A】

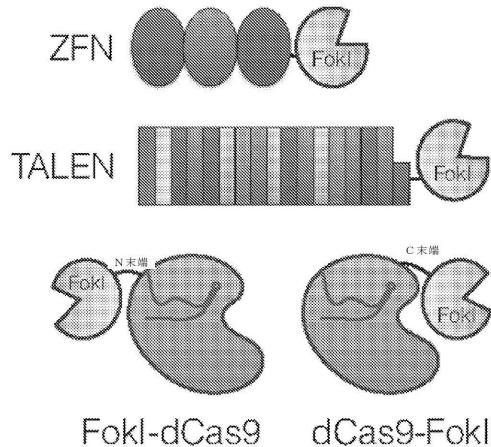


図 5 A

【図 5 B】

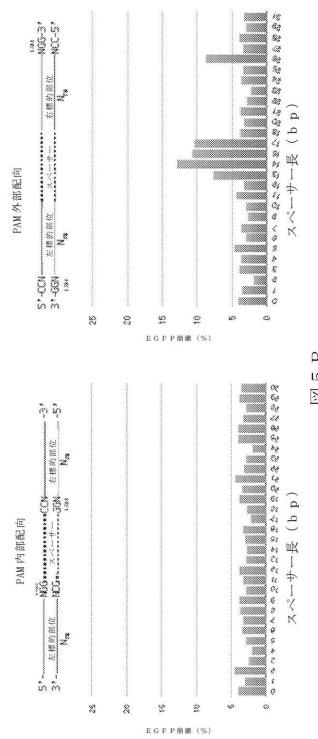


図 5 B

【図 5 C】

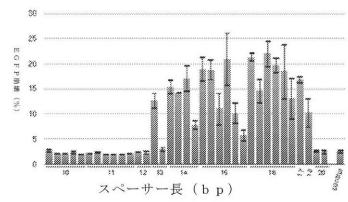


図 5 C

【図 5 E】

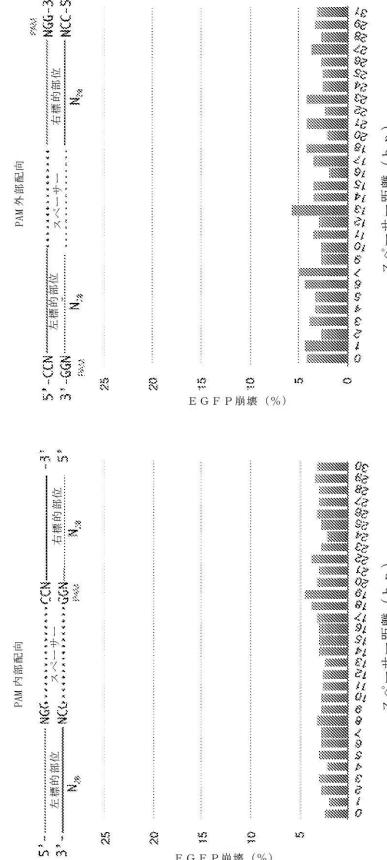


図 5 E

【図 5 D】

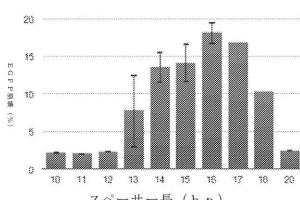


図 5 D

【図 6 A】

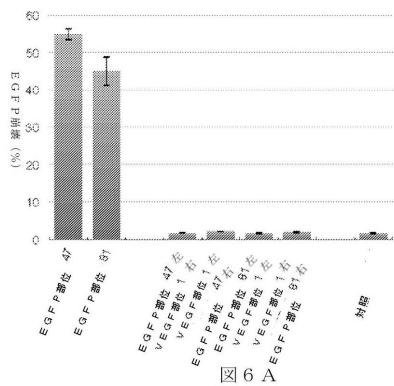


図 6 A

【図 6 B】

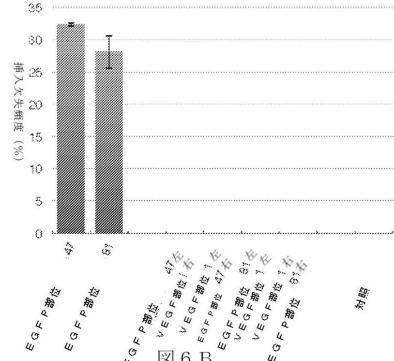


図 6 B

【図 6 C】

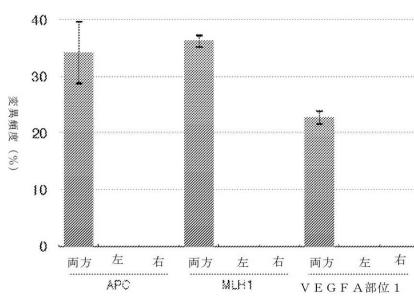


図 6 C

【図 6 D】

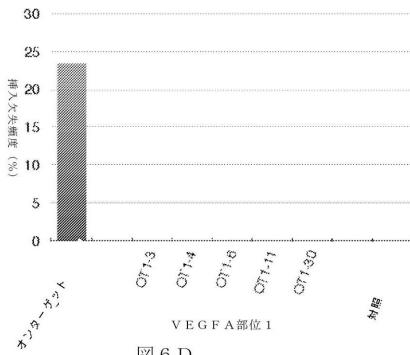


図 6 D

【図 7 A】

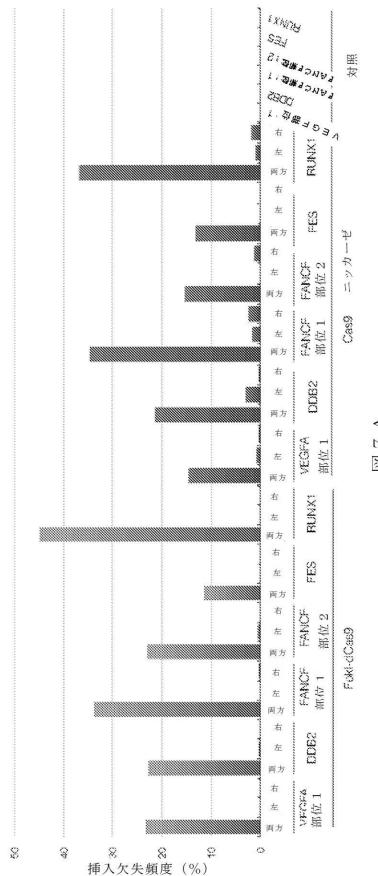
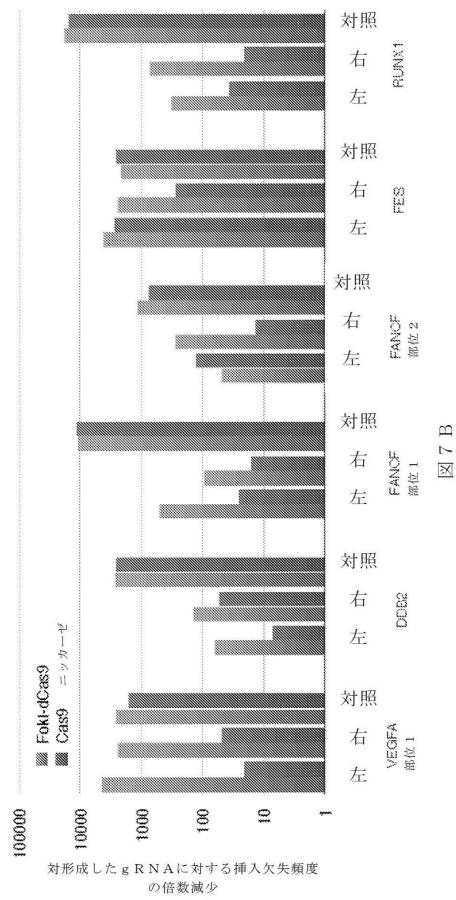
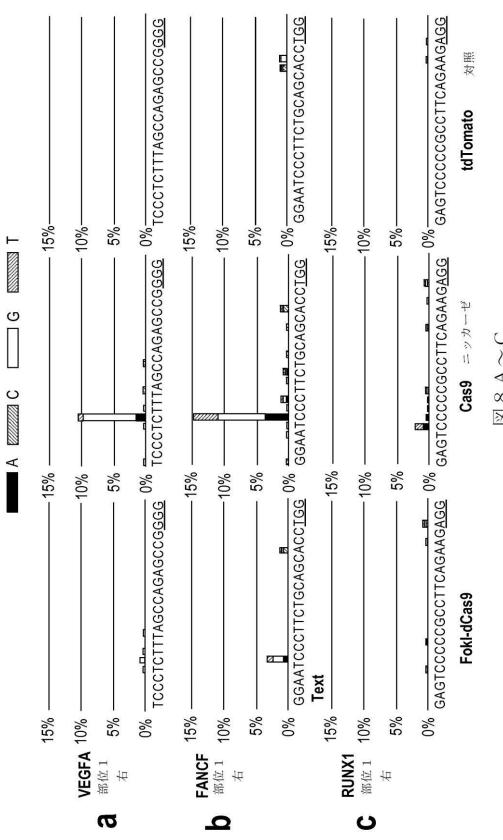


図 7 A

【図7B】



【図8】



【配列表】

0006622183000001.app

フロントページの続き

(51)Int.Cl.	F I
C 12 N 1/21 (2006.01)	C 12 N 1/21
C 12 N 5/10 (2006.01)	C 12 N 5/10

(31)優先権主張番号 61/838,148

(32)優先日 平成25年6月21日(2013.6.21)

(33)優先権主張国・地域又は機関

米国(US)

(31)優先権主張番号 61/921,007

(32)優先日 平成25年12月26日(2013.12.26)

(33)優先権主張国・地域又は機関

米国(US)

(72)発明者 ツァイ , シエンダー

アメリカ合衆国 02129 マサチューセッツ州 , チャールズタウン , プロスペクト ストリート 32 , アパートメント 1

審査官 西村 亜希子

(56)参考文献 國際公開第2014/089290 (WO , A1)

國際公開第2014/099744 (WO , A1)

米国特許出願公開第2010/0055793 (US , A1)

Cell , 2013年 2月28日 , Vol.152 , pp.1173-1183

Science , 2012年 8月17日 , Vol.337 , pp.816-821

PNAS , 1996年 , Vol.93 , pp.1156-1160 , FIG.1

J. Biotechnol. , 2009年 , Vol.140 , pp.156-161 , Fig.1

(58)調査した分野(Int.Cl. , DB名)

C 12 N 9 / 16

C 07 K 19 / 00

C 12 N 1 / 15

C 12 N 1 / 19

C 12 N 1 / 21

C 12 N 5 / 10

C 12 N 15 / 09

JST Plus / JST 580 (JDreamIII)

Caplus / MEDLINE / WPI-IDS / BIOSIS (STN)

GenBank / EMBL / DDBJ / GeneSeq

UniProt / GeneSeq

PubMed