

(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 I

C 1 2 N 9/16 (2006. 01)
 C O 7 K 19/00 (2006. 01)
 C 1 2 N 15/09 (2006. 01)
 C 1 2 N 1/15 (2006. 01)
 C 1 2 N 1/19 (2006. 01)

C 1 2 N 9/16 Z N A A
 C O 7 K 19/00
 C 1 2 N 15/09 1 1 O
 C 1 2 N 1/15
 C 1 2 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, ウィンチェスター, 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 F o k Iヌクレアーゼ (R F N) 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】

R N A 誘導型 F o k Iヌクレアーゼ (R F N) 融合タンパク質であって、触媒的に不活性である化膿性連鎖球菌 (S t r e p t o c o c c u s p y o g e n e s) の C R I S P R 関連 9 (d C a s 9) のアミノ末端に融合する F o k I 触媒ドメイン配列を含む、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。

【請求項 2】

介在リンカーを備えた前記 F o k 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 ~ 3 0 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、G l y₄ S e r を含む、請求項 3 に記載の融合タンパク質。

【請求項 5】

F o k I 触媒ドメインが、配列番号 4 のアミノ酸 3 8 8 ~ 5 8 3、または 4 0 8 ~ 5 8 3 を含む、請求項 1 に記載の融合タンパク質。

【請求項 6】

d C a s 9 が、D 1 0、E 7 6 2、H 9 8 3、または D 9 8 6 での変異、および H 8 4 0 または N 8 6 3 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 RNA 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 RNA 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を含む、

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、10 ~ 20 塩基対離れている、請求項 11 に記載の方法。

20

【請求項 13】

前記 2 つの標的ゲノム配列が、13 ~ 17 塩基対離れている、請求項 12 に記載の方法。

【請求項 14】

前記ガイド RNA が、

(a) 2 つの単一ガイド RNA であって、1つの単一ガイド RNA が第 1 の鎖を標的とし、他方のガイド RNA が相補鎖を標的とし、F o k I が各鎖を切断して、反対の DNA 鎖上に 1 対のニックをもたらす、それにより二重鎖を切断する、2つの単一ガイド RNA、または

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

30

である、請求項 11 に記載の方法。

【請求項 15】

前記 2 つのガイド RNA が、それぞれ、標的ゲノム配列の 17 ~ 20 個のヌクレオチドと相補的な相補領域を含む、請求項 11 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 11 ~ 15 のいずれか 1 項に記載の方法。

40

【請求項 17】

細胞における RNA 誘導型ゲノム編集の特異性が増大する、請求項 11 ~ 16 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 ~ 31 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 11 ~ 17 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 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、前記細胞を、請求項 1 ~ 7 のいずれかに記載の RNA 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。

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DP1 GM105378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】

RNA誘導型FokIヌクレアーゼ(RFN)、たとえばFokI-dCas9融合タンパク質を用いて、RNA誘導型ゲノム編集、例えばCRISPR/Cas9系を用いた編集の特異性を増大させる方法。

【背景技術】

【0004】

20

近年の研究では、クラスター化され等間隔にスパーサーが入った短い回文型の反復配列(CRISPR)/CRISPR関連(Cas)系(Wiedenheftら, Nature 482, 331-338(2012); Horvathら, Science 327, 167-170(2010); Ternsら, Curr Opin Microbiol 14, 321-327(2011))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in vivo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(Wangら, Cell 153, 910-918(2013); Shenら, Cell Res (2013); Dicarloら, Nucleic Acids Res (2013); Jiangら, Nat Biotechnol 31, 233-239(2013); Jinekら, Elife 2, e00471(2013); Hwangら, Nat Biotechnol 31, 227-229(2013); Congら, Science 339, 819-823(2013); Maliら, Science 339, 823-826(2013c); Choら, Nat Biotechnol 31, 230-232(2013); Gratzら, Genetics 194(4):1029-35(2013))。化膿性連鎖球菌(S. pyogenes)由来のCas9ヌクレアーゼ(以降、単にCas9と呼ぶ)は、人工的に設計されたgRNAの最初の20ヌクレオチドと、プロトスパーサー隣接モチーフ(PAM)、例えば、配列NGGまたはNAGにマッチするPAMに隣接する目的とする標的ゲノムDNA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(Shenら, Cell Res (2013); Dicarloら, Nucleic Acids Res (2013); Jiangら, Nat Biotechnol 31, 233-239(2013); Jinekら, Elife 2, e00471(2013); Hwangら, Nat Biotechnol 31, 227-229(2013); Congら, Science 339, 819-823(2013); Maliら, Science 339, 823-826(2013c); Choら, Nat Biotechnol 31, 230-232(2013); Jinekら, Science 337, 816-821(2012))。in vitro(Jinekら, Science 337, 816-821(2012))、細菌(Jiangら, Nat Biotechnol 31, 233-239(2013))およびヒト細胞(Congら, Science 339, 819-823(2013))で実施されたこれま

30

40

50

での研究では、C a s 9 を介した切断が、場合によっては、g R N A / 標的部 位接合部、特に 2 0 ヌクレオチド (n t) g R N A 相補性領域の 3 ' 末端にある最後の 1 0 ~ 1 2 のヌクレオチド (n t) における単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】

【 0 0 0 5 】

多くの研究で、C R I S P R - C a s 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、任意の単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、C R I S P R / C a s 系を用いるゲノム編集、例えば、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 - C a s 9 または F o k I - d C a s 9 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

【 0 0 0 6 】

第一の態様では、本発明は、d C a s 9 の末端、たとえば N 末端に融合する F o k I 触媒ドメイン配列を含む F o k I - d C a s 9 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 ~ 3 0 個、たとえば 4 ~ 1 2 個のアミノ酸、たとえば、G l y₄ S e r のリンカーを含む、F o k I - 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、F o k I 触媒ドメインは、配列番号 4 の 3 8 8 ~ 5 8 3、または 4 0 8 ~ 5 8 3 のアミノ酸を含む。いくつかの実施形態では、d C a s 9 は、D 1 0、E 7 6 2、H 9 8 3、または D 9 8 6 での変異、および H 8 4 0 または N 8 6 3 での変異、たとえば、(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。

20

【 0 0 0 7 】

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。

【 0 0 0 8 】

別の態様では、本発明は、細胞中の二重鎖 D N A 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 F o k I - d C a s 9 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 つの単一ガイド R N A であって、この 2 つの単一ガイド R N A 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 R N A 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 つの単一ガイド R N A は、第 1 の鎖を標的化し、他方のガイド R N A が、相補鎖を標的化する)、F o k I が、反対の D N A 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 つの単一ガイド R N A と、

(b) t r a c r R N A および 2 つの c r R N A であって、2 つの C r R N A のそれぞれが、標的配列の内の 1 つの鎖と相補的な配列を含むことにより、両方の c r R N A 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 c r R N A は第 1 の鎖を標的化し、他方は、相補鎖を標的化する)、F o k I が、反対の D N A 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、t r a c r R N A および 2 つの c r R N A とを含む。

40

【 0 0 0 9 】

別の態様では、本発明は、細胞の 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、本明細書中に記載される R N A 誘導型 F o k I ヌクレアーゼ (R F N) 融合タンパク質と細胞を接触させることを含む、方法を提供する。

【 0 0 1 0 】

50

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、標的配列の内の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)、FokIが、反対のDNA鎖上に1対のニックをもたらし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crRNAであって、2つのcrRNAのそれぞれが、標的配列の内の1つの鎖に相補的な配列を含むことにより、両方のcrRNA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crRNAが第1の鎖を標的化し、他方のcrRNA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しことにより、二重鎖が切断される、tracrRNAおよび2つのcrRNA

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、crRNA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。

【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方法および材料が記載されるが、ほかに、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。

【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。鉋はゲノム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。

【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-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NH2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

10

20

30

40

50

）。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 GGGCACGGGCAGCTTGCCGGTGG（配列番号1） EGFP部位2 GATGCCCGTTCTTCTGCTTGTCGG（配列番号2） EGFP部位3 GGTGGTGCAAGATGAACCTTCAGGG（配列番号3）

【図2D】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 GGGCACGGGCAGCTTGCCGGTGG（配列番号1）
EGFP部位2 GATGCCCGTTCTTCTGCTTGTCGG（配列番号2）
EGFP部位3 GGTGGTGCAAGATGAACCTTCAGGG（配列番号3）

【図2E】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 GGGCACGGGCAGCTTGCCGGTGG（配列番号1）
EGFP部位2 GATGCCCGTTCTTCTGCTTGTCGG（配列番号2）
EGFP部位3 GGTGGTGCAAGATGAACCTTCAGGG（配列番号3）

【図2F】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 GGGCACGGGCAGCTTGCCGGTGG（配列番号1）
EGFP部位2 GATGCCCGTTCTTCTGCTTGTCGG（配列番号2）
EGFP部位3 GGTGGTGCAAGATGAACCTTCAGGG（配列番号3）

【図2G】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2H】ヒト細胞ベースのU2OS EGFP崩壊アッセイにおいて15～25nt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（15nt、17nt、19nt、20nt、21nt、23ntおよび25nt）を保有するgRNAに限り評価することが可能であった。

【図3A】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。

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。Ctrl = 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TFEIAッセイによって測定したものである。Ctrl = 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。

【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。

【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。

【図4A】RNA誘導型FokIヌクレアーゼおよびCRISPR/CasサブタイプY pest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Y pest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。

【図4C】RNA誘導型FokIヌクレアーゼおよびCRISPR/CasサブタイプY pest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

10

20

30

40

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。

10

【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。

20

【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、配列の直右側の縦列中に示される。

30

【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、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、配列の直右側の縦列中に示される。

40

【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、配列の直右側の縦列中に示される。

【図6A】FokI-dCas9RFN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

50

び gRNA の 1 つまたは他方が、(VEGFA 遺伝子中の) 非 EGFP 配列を標的とする別の gRNA と置き換えられている対の存在下で評価される 2 つの REN 対の EGFP 崩壊の活性。EGFP 崩壊は、フローサイトメトリーにより定量化された。EGFP : 高感度緑色蛍光タンパク質、VEGFA : 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】FokI - dCas9 REN 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 EGFP 崩壊アッセイで使用されるのと同じ細胞由来のゲノム DNA で実施した T7E1 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】FokI - dCas9 REN の二量体化が、有効なゲノム編集の活性に必要である。(c) APC、MLH1 および VEGFA 遺伝子中の部位を標的とする REN の活性化。各標的で、本出願人は、「左の」片側部位では 1 つの gRNA のみ、または「右の」片側部位では 1 つの gRNA のみである、1 対の同族の gRNA と FokI - dCas9 を共発現させた。変異原性の比率は、T7E1 アッセイにより測定した。APC : 大腸腺腫症 ; MLH1 : mutL 相同体 1 ; VEGFA : 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】FokI - dCas9 REN の二量体化が、有効なゲノム編集の活性に必要である。(d) VEGFA 部位 1 を標的とするために使用される gRNA 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 (OT) 部位で VEGFA 部位 1 を標的とする REN 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 DNA 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 VEGFA 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 gRNA と共発現した Cas9 ニッカーゼまたは FokI - dCas9 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 gRNA の存在下での FokI - dCas9 (左のバー) または Cas9 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 gRNA 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 gRNA、または「右」の片側部位では他方の gRNA 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 DNA 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。VEGFA : 血管内皮増殖因子 A、DDB2 : 損傷特異的 DNA 結合タンパク質 2 (Damage - Specific DNA Binding Protein 2) ; FANCF : ファンコニ貧血相補群 F (Fanconi Anemia , Complementation Group F) ; FES : ネコ肉腫ウイルス癌遺伝子 ; RUNX 1 : Runt 関連転写因子 1 (Runt - Related Transcription Factor 1)。

【図 7 B】単一 gRNA と共発現した Cas9 ニッカーゼまたは FokI - dCas9 の変異原性活性。(b) 単一 gRNA の実験または対照の実験 (gRNA なし、および Cas9 ニッカーゼまたは FokI - dCas9 なし) のそれぞれに対して、gRNA 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、FokI - dCas9 (各対の左側のバー、薄灰色) および Cas9 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 Cas9 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。FokI - dCas9、Cas9 ニッカーゼ、または tdTomato 対照の存在下の (a) VEGFA、(b) FANCF、および (c) RUNX1 の遺伝子標的での、単一 gRNA 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に留意されたい。VEGFA：血管内皮増殖因子A；FANCF：ファンコニ貧血相補群F；RUNX1：Runx1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）

CRISPR 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。Marraffini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、様々な程度で単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タ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、単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での位置（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（Jiangら，Nat Biotechnol 31，233-239（2013）；Congら，Science 339，819-823（2013）；Jinekら，Science 337，816-821（2012））。しかしながら3'末端の単一および二重変異が十分に許容されら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トソン-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Marraffini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LiuらがZFN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、オンターゲット部位と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

10

20

30

40

50

って、真の R G N 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 R G N 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、T F E I 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、これまではるかに頻度の低い Z F N 誘発オフターゲット変化および T A L E N 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（P a t t a n a y a k ら, N a t M e t h o d s 8, 7 6 5 - 7 7 0 (2 0 1 1) ; P e r e z ら, N a t B i o t e c h n o l 2 6, 8 0 8 - 8 1 6 (2 0 0 8) ; G a b r i e l ら, N a t B i o t e c h n o l 2 9, 8 1 6 - 8 2 3 (2 0 1 1) ; H o c k e m e y e r ら, N a t B i o t e c h n o l 2 9, 7 3 1 - 7 3 4 (2 0 1 1) ）。このほか、ヒト細胞における R G N オフターゲット変異誘発の解析から、R G N 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。

【 0 0 2 0 】

いくつかの R G N 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 R G N では、ヒトゲノム内にある総数を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。T F E I 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ル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 R G N 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、細胞内での R G N 活性に影響を及ぼし得る R G N 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 R G N 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。

【 0 0 2 1 】

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、R G N 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 R G N によって効率的に変異され得るとすると、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 R G N には通常、2 0 b p の R N A : D N A 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、g R N A 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、G C 含有量が高いと R N A : D N A ハイブリッドが安定することが示されており（S u g i m o t o ら, B i o c h e m i s t r y 3 4, 1 1 2 1 1 - 1 1 2 1 6 (1 9 9 5) ）、したがって、g R N A / ゲノム D N A 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ことが考えられる。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 R N A : D N A ハイブリッドの安定性）がゲノム全域にわたる R G N 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、g R N A 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しても、このようなガイドラインの実施による影響が R G N 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】

RGN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1つに、細胞内で発現するgRNAおよびCas9ヌクレアーゼの濃度を低くすることが考えられる。U2OS・EGFP細胞のVEGFA標的部位2および3にRGNを用いて、この考えを検討した。トランスフェクトするgRNA発現およびCas9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(HEK293細胞およびK562細胞)でも、オンターゲット変異誘発の絶対的比率がU2OS・EGFP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。したがって、細胞内でのgRNAおよびCas9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gRNAおよび/またはCas9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。

10

【0023】

3種類の異なるヒト細胞型でRGN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。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DNA配列を標的とする複数のRGN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、RGN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。

20

【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. pyogenes*)Cas9タンパク質を土台とするCRISPR-Cas RNA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)。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、CRISPR-Cas RNA誘導型ヌクレアーゼ(RGN)の特異性を改善する方法が必要とされている。

30

【0025】

実施例1に記載されるように、Cas9 RGN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、Cradicら, 2013; Fuら, 2013; Hsura, 2013; Pattanayakra, 2013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、gRNA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Fuら, 2013; Hsura, 2013; Pattanayakra, 2013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。したがって、RNA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。

40

【0026】

二量体化は、Cas9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Cas9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、DNAのセグメント上に2つのCas9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、酵素活性に必要なではないため、単一のCas9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。

【0027】

したがって、RGNの特異性を改善する1つの手法として、D10AおよびH840Aの変異を有する触媒的に不活性な形態のCas9（dCas9としても知られる）に、FokIエンドヌクレアーゼドメインを融合することがある。FokI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DNA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9Aおよび実施例2）では、2つのFokI-dCas9融合物が、二重鎖の切断を得るために2つの異なるgRNA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、FokI-dCas9融合物は、単一のRGN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。

【0028】

よって、本明細書は、FokI-dCas9融合タンパク質であって、FokI配列が、dCas9（好ましくは、dCas9のアミノ酸末端、また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2～30個のアミノ酸、たとえば4～12個のアミノ酸、たとえばGly₄Ser（配列番号：23）または（Gly₄Ser）₃のリンカーを含む、FokI-dCas9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、dCas9およびFokI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2～20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GGGS（配列番号22）またはGGGGS（配列番号23）、たとえば、GGGS（配列番号22）またはGGGGS（配列番号23）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。

【0029】

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RNA誘導型FokI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gRNA対を発現するために有効な系であって、RFN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Cas9ニッカーゼは、一般的に、単一gRNA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Cas9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。

【0030】

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RNA誘導型FokIヌクレアーゼ（RFN）プラットフォームが記載される。RFNは、二量体としての活性と機能のために2つのgRNAを必要とする。驚くべきことに、活性RFNの遺伝子操作は、FokI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ー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ZenおよびTALENと異なる構築物である、dCas9タンパク質のアミノ末端へのFokI

ヌクレアーゼドメインの融合を必要とした。またRFNは、各FokI-dCas9/gRNA複合体により結合された片側部位が、14～17bpの長さの相対的に限定された介在スパーサーを備えた特定の相対配向(PAM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)。

【0031】

RFNの二量体の性質は、標準的な単量体Cas9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gRNAにより配向されるFokI-dCas9は、RFN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gRNA(6つのRFN標的部位に対する)を、共発現したFokI-dCas9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、gRNA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FokI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gRNAで観察されるいずれかの挿入欠失が、DNAへのFokI-dCas9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、FokI-dCas9を12のオンターゲット片側部位での単一gRNA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、VEGFAを標的とするRFNは、ディープシーケンシングにより判定されるように、gRNA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。

【0032】

RFNは真の二量体系であるので、RFNは、共局在化に依存するが二量体化を必要と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Cas9ニッカーゼは、一般的に、同一の個々のgRNAにより配向されるFokI-dCas9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Cas9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Cas9ニッカーゼが、同一の単一gRNAにより誘導されるFokI-dCas9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Cas9ニッカーゼは、二量体RFN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、PAM内部またはPAM外部、および、0～1000bpの範囲の長さのスパーサー配列を含んで配向できる(Ran et al., Cell 154, 1380-1389 (2013); Mali et al., Nat Biotechnol 31, 833-838 (2013); Cho et al., Genome Res (2013))。この無差別性は、Cas9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、RFN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FokI切断ドメインの必要条件のため、PAM外部を有していなければならない、14～17bp離れていなければならない。

【0033】

FokI

FokIは、DNA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、FokI全体、または触媒エンドヌクレアーゼドメイン、たとえば、GenBank アクセッション番号第AA24927.1のアミノ酸388～583または408～583(たとえばLi et al., Nucleic Acids Res. 39(1): 3

10

20

30

40

50

59-372 (2011); Cathomen and Joung, Mol. Ther. 16: 1200-1207 (2008)に記載)、またはMiller et al. Nat Biotechnol 25: 778-785 (2007); Szczepek et al., Nat Biotechnol 25: 786-793 (2007);もしくはBitinaite et al., Proc. Natl. Acad. Sci. USA. 95:10570-10575 (1998)に記載されるような、FokIの変異形態を含むことができる。

【0034】

FokIの例示的なアミノ酸配列は、以下の通りである。

【化1】

10	20	30	40	50	60
MFLSMVSKIR	TFGWVQNPGR	FENLKRVRVQV	FDRNSKVHNE	VKNIKIPTLV	KESKIQKELV
70	80	90	100	110	120
AIMNQHDLIY	TYKELVGTGT	SIRSEAPCDA	IIQATIADQG	NKKGYIDNWS	SDGFLRWAHA
130	140	150	160	170	180
LGFIEYINKS	DSFVITDVGL	AYSXSADGSA	IEKEILIEAI	SSYPPAIRIL	TLLEDGQHLL
190	200	210	220	230	240
KFDLGKNLGF	SGESGFTSLP	EGILLDTLAN	AMPKDKGEIR	NNWEGSSDKY	ARMIGGWLDK
250	260	270	280	290	300
LGLVKQGKKE	FIIPTLGKPD	NKEFISHAFK	ITGEGLKVLK	RAKGSTKFTR	VPKRVYWEML
310	320	330	340	350	360
ATNLTDKEYV	RTRRALILEI	LIKAGSLKIE	QIQDNLKKLG	FDEVIETIEN	DIKGLINTGI
370	380	390	400	410	420
FIEIKGRFYQ	LKDHILOFVI	PNRGVTKQLV	KSELEEKKSE	LRHKLKYVPH	EYIELIEIAR
430	440	450	460	470	480
NSTQDRILEM	KVMEFFMKVY	GYRGKHLGGS	RKPDGAIYTV	GSPIDYGVIV	DTKAYSGGYN
490	500	510	520	530	540
LPIGQADEMQ	RYVEENQTRN	KHINPNEWWK	VYPSSVTEFK	FLFVSGHFKG	NYKAQLTRLN
550	560	570	580		
HITNCNGAVL	SVEELLIGGE	MIKAGTLTLE	EVRRKFNNGE	INF (配列番号4)	

【0035】

FokIをコードする例示的な核酸配列は、以下の通りである。

10

20

30

【化 2】

ATGTTTTTGTAGTATGTTTTCTAAAAAAGAAGCTTTCGGTTGGGTTCAAAATCCAGGTAAA
 TTTGAGAATTTAAAACGAGTAGTTCAAGTATTTGATAGAAATTCTAAAGTACATAATGAA
 GTGAAAAATATAAAGATACCAACCCTAGTCAAAGAAAGTAAGATCCAAAAAGAAGCTAGTT
 GCTATTATGAATCAACATGATTTGATTTATACATATAAAGAGTTAGTAGGAACAGGAACT
 TCAATACGTTTCAAGAGCACCATGCGATGCAATTATTCAAGCAACAATAGCAGATCAAGGA
 AATAAAAAAGGCTATATCGATAATTGGTCATCTGACGGTTTTTTGCGTTGGGCACATGCT
 TTAGGATTTTATTGAATATATAAATAAAGTGATTTCTTTTGTAACTGATGTTGGACTT
 GCTTACTCTAAATCAGCTGACGGCAGCGCCATTGAAAAAGAGATTTTGATTGAAGCGATA
 TCATCTTATCCTCCAGCGATTTCGTATTTTAACTTTGCTAGAAGATGGACAACATTTGACA
 AAGTTTTGATCTTGGCAAGAATTTAGGTTTTAGTGGAGAAAGTGGATTTACTTCTCTACCG
 GAAGGAATTTCTTTTAGATACTCTAGCTAATGCTATGCCTAAAGATAAAGGCGAAATTCGT
 AATAATTGGGAAGGATCTTCAGATAAGTACGCAAGAATGATAGGTGGTTGGCTGGATAAA
 CTAGGATTAGTAAAGCAAGGAAAAAAGAATTTATCATTCTACTTTGGGTAAGCCGGAC
 AATAAAGAGTTTATATCCCACGCTTTTAAATTTACTGGAGAAGGTTTGAAAGTACTGCGT
 CGAGCAAAAGGCTCTACAAAATTTACACGTGTACCTAAAAGAGTATATTGGGAAATGCTT
 GCTACAAACCTAACCAGATAAAGAGTATGTAAGAACAAGAAGAGCTTTGATTTTAGAAATA
 TTAATCAAAGCTGGATCATTAAAAATAGAACAATAACAAGACAACCTGAAGAAATTAGGA
 TTTGATGAAGTTATAGAACTATTGAAATGATATCAAAGGCTTAATTAACACAGGTATA
 TTTATAGAAATCAAAGGGCGATTTTATCAATTGAAAGACCATATTCTTCAATTTGTAATA
 CCTAATCGTGGTGTGACTAAGCAACTAGTCAAAAGTGAAGTGGAGGAGAAGAAATCTGAA
 CTTTCGTCATAAATTGAAATATGTGCCTCATGAATATATTGAATTAATTGAAATGCGCAGA

10

AATTCCACTCAGGATAGAATTCTTGAAATGAAGGTAATGGAATTTTTTATGAAAGTTTAT
 GGATATAGAGGTAAACATTTGGGTGGATCAAGGAAACCGGACGGAGCAATTTATACTGTC
 GGATCTCCTATTGATTACGGTGTGATCGTGGATACTAAAGCTTATAGCGGAGGTTATAAT
 CTGCCAATTGGCCAAGCAGATGAAATGCAACGATATGTCGAAGAAAATCAAACACGAAAC
 AAACATATCAACCCTAATGAATGGTGGAAAGTCTATCCATCTTCTGTAACGGAATTTAAG
 TTTTTATTTGTGAGTGGTCACTTTAAAGGAAACTACAAAGCTCAGCTTACACGATTAAAT
 CATATCACTAATTGTAATGGAGCTGTTCTTAGTGTAGAAGAGCTTTTAATTGGTGGAGAA
 ATGATTAAGCCGGCACATTAACCTTAGAGGAAGTGAGACGGAAATTTAATAACGGCGAG
 ATAAACTTTTAA (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 F o k 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 とのいずれの差異も、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、これらは参照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、Esvelt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ら, 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】

10

20

30

【表 1】

代替のCas9タンパク質	
GenBank アクセッション番号	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トイ (<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	コプロコッカス・カツス (<i>Coprococcus catus</i>) GD-7
42525843	トレポネマ・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・インテステイニ (<i>Acidaminococcus intestini</i>) RYC-MR95
302336020	オルセンエラ・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>Fingoldia magna</i>) ATCC29328
47458868	マイコプラズマ・モービレ (<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・ポリトロプス (<i>Ilyobacter polytropus</i>) DSM2926
325677756	ルミノコッカス・アルプス (<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・セルロリテिकास (<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>Nitratifractor salsuginis</i>) DSM16511
325972003	スピロヘータ・グロバス (<i>Sphaerochaeta globus</i>) Buddy株
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・フラギリリス (<i>Bacteroides fragilis</i>) NCTC9343
256819408	カプノサイトファーガ・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
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	バーミンフロボクター・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
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 (Termite) 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・セルロリテカム (<i>Clostridium cellulolyticum</i>) H10
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	髄膜炎菌 (<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>)

10

20

30

40

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 excrementihominis</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

【 0 0 4 2 】

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 Cas 9 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 Cas 9 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. meningitides*) Cas 9 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. innocua* の Cas 9 オルソログ (異なるガイド RNA を利用すると思われる髄膜炎菌 (*N. meningitidis*) および *C. jejuni* の Cas 9 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ミド DNA を切断し得ることを *in vitro* で明らかにしている。

20

【 0 0 4 3 】

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*) Cas 9 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、(Nishimasuら, *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*) Cas 9 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。

30

【 0 0 4 4 】

【表 2】

10	20	30	40	50	60
MDKKYSIGLA	IGTNSVGWAV	ITDEYKVFSK	KFKVLGNTDR	HSIKKNLIGA	LLFDSGETAE
70	80	90	100	110	120
ATRLKRTARR	RYTRKKNRIC	YLQEIFSNEM	AKVDDSPFHR	LEESFLVEED	KKHERHPIFG
130	140	150	160	170	180
NIVDEVAYHE	KYPTIYHLRK	KLVDSTDKAD	LRLIYLALAH	MIKFRGHFLI	EGDLNPDNSD
190	200	210	220	230	240
VDKLFIOLVQ	TYNOLFEEENP	INASGVDAKA	ILSARLSKSR	RLENLIAQLP	GEKKNGLFGN
250	260	270	280	290	300
LIALSLGLTP	NFKSNFDLAE	DAKLQLSKDT	YDDDLNLLA	QIGDQYADLF	LAANKLSDAI
310	320	330	340	350	360
LLSDILRVNT	EITKAPLSAS	MIKRYDEHHQ	DLTLKALVR	QQLPEKYKEI	FFDQSKNGYA
370	380	390	400	410	420
GYIDGGASQE	EFYKFIKPIL	EKMDGTEELL	VKLNREDLLR	KQRTFDNGSI	PHQIHLGELH
430	440	450	460	470	480
AILRRQEDFY	PFLKDNREKI	EKILTFRIPY	YVGPLARGNS	RFAMNTRKSE	ETITPMNFEE
490	500	510	520	530	540
VVDKGASQAS	FIERMTNFDK	NLPNEKVLPR	HSLLEYFTV	YNELTKVKYV	TEGMRKPAFL
550	560	570	580	590	600
SGEQKKAIVD	LLFKTNRKVT	VKQLKEDYFK	KIECFDSVEI	SGVEDRFNAS	LGTYHOLLKI
610	620	630	640	650	660
IKDKDFLDNE	ENEDILEDIV	LTLTLFEDRE	MIEERLKTYA	HLFDDKVMKQ	LKRRRYTGWG
670	680	690	700	710	720
RLSRKLINGI	RDQSGKTL	DFLKSDGFAN	RNFMLIHDD	SLTFKEDIQK	AQVSGQGDLSL
730	740	750	760	770	780
HEHIANLAGS	PAIKKGILQT	VKVVDLVKV	MGRHKPENIV	IEMARENQTT	QKGQKNSRER
790	800	810	820	830	840
MKRIEEGIKE	LGSQILKEHP	VENTQLQNEK	LYLYYLQNGR	DMYVDQELDI	NRLSDYDVDA
850	860	870	880	890	900
IVPQSFLKDD	SIDNKVLTRS	DKNRGKSDNV	PSEEVVKKMK	NYWRQLLNAK	LITQRKFDNL
910	920	930	940	950	960
TKAERGGLSE	LKAGFIKRO	LVETRQITKH	VAQILDSPMN	TKYDENDKLI	REVKVITLKS
970	980	990	1000	1010	1020
KLVSDFRKDF	QFYKVRINN	YHHAHDAYLN	AVVGTALIKK	YPKLESEFVY	GDYKVYDVRK
1030	1040	1050	1060	1070	1080
MIKSEQZIG	KATAKYFFYS	NIMNFFKTEI	TLANGEIRKR	PLIETNGETG	EIVWDKGRDF
1090	1100	1110	1120	1130	1140
ATVRKVLSP	QVNIVKKTEV	Q7GGFSKESI	LPRNSDKLI	ARKKDWDPKK	YGGFDSPTVA

10

20

30

40

1150	1160	1170	1180	1190	1200
YSVLVVAKVE	KGKSKKLKSV	KELLGITIME	RSSFENPID	FLEAKGYKEV	KKDLIIKLPK
1210	1220	1230	1240	1250	1260
YSLFELENGR	KRMLASAGEL	QKGNELALPS	KYVNFYLLAS	HYEKLKGSPE	DNEQKQLFVE
1270	1280	1290	1300	1310	1320
QHKHYLDEII	EQISEFSKRV	ILADANLDKV	LSAYNKHRDK	PIREQAENII	HLFTLTNLGA
1330	1340	1350	1360		
PAAFKYFDTT	IDRKRYTSTK	EVLDAFLIHQ	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を単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1nt、2nt、3nt、4nt、5nt、6nt、7nt、8nt、9nt、10nt、15nt、20nt、25nt、30nt、35nt または 40nt 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1nt、2nt、3nt、4nt、5nt、6nt、7nt、8nt、9nt、10nt、15nt、20nt、25nt、30nt、35nt または 40nt 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1nt、2nt、3nt、4nt、5nt、6nt、7nt、8nt、9nt、10nt、15nt または 20nt、3'末端で少なくとも1nt、2nt、3nt、4nt、5nt、6nt、7nt、8nt、9nt、10nt、15nt、20nt、25nt、30nt、35nt または 40nt 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 お

30

40

50

よびFuら, Nat Biotechnol. 2013 Mar; 31(3): 227-9; Jinekら, 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Spacer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Spacer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)GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUC
CG(X_N) (配列番号6);
(X₁₇₋₂₀)GUUUUAGAGCUAUGCUGAAAAGCAUAGCAAGUUAAAAUAAGGCU
AGUCCGUUAUC(X_N) (配列番号7);
(X₁₇₋₂₀)GUUUUAGAGCUAUGCUGUUUUGGAAACAAAACAGCAUAGCAAGU
UAAAAUAAGGCUAGUCCGUUAUC(X_N) (配列番号8);
(X₁₇₋₂₀)GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUU
AUCAACUUGAAAAAGUGGCACCGAGUCGGUGC(X_N) (配列番号9);
(X₁₇₋₂₀)GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUU
AUCAACUUGAAAAAGUGGCACCGAGUCGGUGC (配列番号10);
(X₁₇₋₂₀)GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUUAAAAUAAGG
CUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC (配列番号11);
または

30

40

(X₁₇₋₂₀)GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUUAAAAUAAGG
CUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC (配列番号12)

であって、式中X₁₇₋₂₀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、RNA Pol III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、UUUUUU、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crRNA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crRNA、たとえば以下

(X₁₇ 20) GUUUUAGAGCUA (配列番号: 13);

(X₁₇ 20) GUUUUAGAGCUAUGCUGUUUUG (配列番号: 14);

または

(X₁₇ 20) GUUUUAGAGCUAAGCU (配列番号: 15)

およびtracrRNA配列と共に使用される。この場合、crRNA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】

GGAACCAUUCAAAACAGCAUAGCAAGUAAAAUAAGGCUAGUCCGUUA
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いてもよい。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】

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

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

【0051】

10

($X_{17 \quad 20}$) GUUUUAGAGCUAUGCUGUUUUG (配列番号: 14) を crRNA として使用するいくつかの実施形態では、以下の tracrRNA を使用する:

【化6】

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

($X_{17 \quad 20}$) GUUUUAGAGCUA (配列番号: 13) を crRNA として使用するいくつかの実施形態では、以下の tracrRNA を使用する:

20

【化7】

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

($X_{17 \quad 20}$) GUUUUAGAGCUAUGC (配列番号: 15) を crRNA として使用するいくつかの実施形態では、以下の tracrRNA を使用する:

【化8】

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

30

【0052】

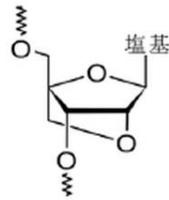
いくつかの実施形態では、gRNA は、オフターゲット作用を最小限にするために、ゲノムの残りのいずれかの配列と異なる少なくとも3つ以上のミスマッチである部位を標的とする。

【0053】

ロケット核酸 (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)。

40

【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-g RNA* 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-g RNA* 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 *cr RNA* および / または *tracr 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 *g RNA* - 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、*g RNA* の全部または一部、例えば *g RNA* 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 *g RNA* 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 *cr RNA* / *tracr RNA* 系の *cr RNA* および / または *tracr RNA* 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 *cr RNA* および *tracr RNA* 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、1つ以上の修飾 (たとえばロックト) 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ン化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ン化して植物細胞、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。

10

【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ン化する。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。

20

30

【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、lac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、Gossen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Rendahlら, 1998, Nat. Biotechnol., 16: 757-761を参照されたい)。

40

【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

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・バーウイ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+、pMTO10/A+、pMAMneo-5、バキュロウイルス pDSVE および SV40 初期プロモーター、SV40 後期プロモーター、メタロチオネインプロモーター、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 Polymerase 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 *in 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 Polymerase III または Polymerase II 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*) で機能するレプリコン、組換えプラスミドを保有する細菌の選択を可能にする抗生物質耐性をコードする遺伝子および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（エピソード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な方法で、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

CRISPR 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融合遺伝子の単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-ト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 LSRIIF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 1 M 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 QIA XCEL 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 誘発オンタ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) 250 ng、Cas9 発現プラスミド 750 ng および td-トマト発現プラスミド 30 ng をトランスフェクトした。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) 125 ng、Cas9 発現プラスミド 375 ng および td-トマト発現プラスミド 30 ng をトランスフェクトした。QIAamp DNA Blood Mini Kit (QIAGEN) を製造業者の説明書に従って用い、トランスフェクトした 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 回のヌクレオフェクション (U2OS、EGFP 細胞)、2 回のヌクレオフェクション (K562 細胞) または 2 回の Lipofectamine LTX トランスフ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、各トランスフェクションを計 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 (Agencourt) 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 (Reyonら, 2012, 上記)。

20

30

【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) にクローン化し、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' - GTAAAACGACGGCCAG - 3' (配列番号 19) を用いてサンガー法 (MGH DNA Sequencing Core) によりプラスミドの配列を決定した。

40

【0081】

実施例 1a . 単一ヌ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ヌ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 Reyon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、ヌ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

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ヌ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gRNA(sgRNA)を用いた:

EGFP部位1 GGGCACGGGCAGCTTTGCCGGTGG(配列番号1)

EGFP部位2 GATGCCCGTTCTTCTTGCTTGTCGG(配列番号2)

EGFP部位3 GGTGGTGCAGATGAACCTTCAGGG(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)。

【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単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)。

【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 Biotechnol 31, 233-239(2013); Congら, Science 339, 819-823(2013); Jinekら, Science 337, 816-821(2012))の通り、gRNAの位置1~10に単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の一部を除くいずれの位置に単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。

【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ヌ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(単一および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。

【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種類の標的部位いずれも、3つ以上の隣接するミ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

10

20

30

40

50

ッ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でなければならないことから、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った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った。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。

【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。

【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

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、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った。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125%(平均40%)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

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、TALEN¹⁵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ターゲット部位に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タ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(Maeder, M.L.ら, Mol Cell 31, 294-301(2008); Reynon, 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単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】

本発明者らのU2OS・EGFPレポーター細胞にトランスフェクトするCas9コードプラスミドの量を漸増させて(50ng~750ng)EGFP崩壊をアッセイした。図3F(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Cas9コードプラスミドの量を3分の1にしても、EGFP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RGNの活性は、トランスフェクトするCas9プラスミドの量を3分の1にすると直ちに低下した(図3F(上段))。以上の結果を踏まえて、実施例1a~1dに記載される実験には、EGFP標的部位#1、#2および#3に対してgRNA発現プラスミド/Cas9発現プラスミドをそれぞれ25ng/250ng、250ng/750ngおよび200ng/750ng用いた。

【0094】

一部のgRNA/Cas9の組合せが他の組合せよりもEGFP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sgRNA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sgRNA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1~6bp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RGNおよびEGFPレポーター遺伝子を標的とする3種類のRGNのヒトゲノムにおけるオフターゲット部位の数

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
EGFP 標的部位#1	0	0	3	10	156	1365	9755
EGFP 標的部位#2	0	0	0	11	96	974	7353
EGFP 標的部位#3	0	0	1	14	165	1439	10361

VEGFA、RNF2、FANCFおよびEMX1遺伝子を標的とする6種類のRGNならびにEGFP標的部位#1、#2および#3を標的とする3種類のRGNそれぞれのオフターゲット部位はヒトゲノム配列ビルドGRCh37で特定されたものである。ミスマッチはgRNAがアニールする20nt領域に対するもののみを認め、PAM配列に対するものは認めなかった。

【0096】

実施例2: FokI-dCas9融合タンパク質とのガイドRNA対の使用

単量体のCRISPR-Cas9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つの単一ガイドRNA(gRNA)に厳密に依存する新規の二量体RNA誘導型FokIヌクレアーゼ(RFN)を記載する。RFNは、内在性のヒト遺伝子のDNA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gRNAを発現する方法が記載され、これは、二量体RFN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、単量体Cas9ヌクレアーゼは、一般的に、マッチした単一gRNAにより導かれるRFN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。RFNは、二量体化の特異性の増強とCRISPR RNA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。

10

20

30

40

50

【 0 0 9 7 】

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 g R N A およびマルチプレックス g R N A 発現プラスミド

単一またはマルチプレックス g R N A をコードするプラスミドを、 B s m B I 消化型 C s y 4 - 隣接 g R N A 骨格 (p S Q T 1 3 1 3 ; A d d g e n e) を備える、アニールした標的部位のオリゴ二重鎖 (I n t e g r a t e d D N A T e c h n o l o g i e s) および定常領域のオリゴ二重鎖 (複数の g R N A に対して) の単一ステップの連結で構築した。

【 0 0 9 8 】

マルチプレックス g R N A 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) c r R N A 、 t r a c r R N A および C s y 4 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、 B s m B I 2 型制限酵素で消化した U 6 - C s y 4 部位 - g R N A プラスミド骨格の中に連結することにより構築した。 C s y 4 R N A 結合部位を、 g R N A 配列の 3 ' 末端および 5 ' 末端に付着させ、細胞において C a s 9 とともに発現させた。 C s y 4 R N A 結合部位の配列 ' G U U C A C U G C C G U A U A G G C A G C U A A G A A A ' (配列番号 : 2 0) を、標準的な g R N A 配列の 5 ' 末端および 3 ' 末端に融合した。

【 化 1 0 】

GUUCACUGCCGUAUAGGCAGNNNNNNNNNNNNNNNNNNNNNGUUUUAGAGCUA
GAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCA
CCGAGUCGGUGCGUUCACUGCCGUAUAGGCAGNNNNNNNNNNNNNNNNNNNNNN
GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUU
GAAAAAGUGGCACCGAGUCGGUGCGUUCACUGCCGUAUAGGCAG
 (配列番号 2 1)

【 0 0 9 9 】

この配列は、 C s y 4 部位に隣接するマルチプレックス g R N A 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。 C s y 4 隣接 s g R N A 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、 s g R N A は、 1 つの転写物上でコードされる C s y 4 部位により分離されるマルチプレックス s g R N A 、 および追加的な C s y 4 配列を有する個々の s g R N A でコードすることができる。この配列では、第 1 の N 2 0 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 N 2 0 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。

【 0 1 0 0 】

g R N A を含む C s y 4 認識部位をコードするプラスミドを、 「 2 A 」 ペプチド結合により分かれる C a s 9 および C s y 4 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スフェクトした。この結果から、 5 ' 末端および 3 ' 末端に融合した C s y 4 部位を備える g R N A は、上述した U 2 O S - E G F P 崩壊アッセイを使用して、ヒト細胞中での C a s 9 仲介性切断の誘導が可能なままであった。したがって、 C s y 4 R N A の結合部位は、 g R N A 配列の 3 ' 末端に付着することができ、 C a s 9 とこれら C s y 4 部位含有 g R N A の複合体は、細胞中で機能的なままである。

【 0 1 0 1 】

いくつかの実験では、 C s y 4 - T 2 A - F o k I - d C a s 9 をコードする構築物を使用した。 F o k I - d C a s 9 融合物の配列を以下に示す。この配列は、 F o k I および d C a s 9 および核局在配列の間に G G G G S (配列番号 : 2 3) 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)

MQLVKSELEKKSELRHKLKYPHEYIELIEIARNSTQDRILEMKVMEFFMKVYGYR
 GKHLGGSRKPDGAIYTVGSPIDYGVIVDTKAYS GGYNLP IQADEMQRYVEENQTRN
 KHINPNEWWKVYPSSVTEFKFLFVSGHFKGNYKAQLTRLNHITNCNGAVLSVEELLI
 GGEMIKAGTLTLEEVRRKFNNGEINFGGGGSDKKYSIGLAIGTNSVGWAVITDEYKV
 PSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQE
 IFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK
 LVDSTDKADLRILIYALAHMIKFRGHFLIEGDLNPDNSDVKLFILQVQTYNQLFEE
 NPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIASLGLTPNFKSN
 FDLAEDAKQLSKDITYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEI
 TKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQ
 EEFYKFIKPILEKMDGTEELLVKNLREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ
 EDFYFPLKDNREKIEKILTFRIPIYVGPLARGNSRFAMWTRKSEETITPWNFEVVVD
 KGASAQSFIERMTNFDKNLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFL
 SGEQKKAIVDILLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDL
 LKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMKQLKRR
 RYTGWGRLSRKLINGIRDQSGKTILDFLKSDFANRNFMLIHDDSLTFKEDIQKA
 QVSGQGDLSLHEHIANLAGSPAIIKKGILQTVKVVDELVKVMGRHKPENIVIAMARENQ
 TTQKGQKNSRERMKRIIEGKELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVD
 QELDINRLSDYDVAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNY
 WRQLLNAKLITQRKFDNLTAKERGGLSELDKAGFIKRLVETRQITKHVAQILDSRM
 NTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVRINNYHHAHDAYLNAVVGTA
 LIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLA
 NGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIKKTEVQTGGFSKESI
 LPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEKSKKLKSVKELLGIT
 IMERSSEFNPIDFLEAKGYKEVKKDLIIKLPKYSLEFLENKRKMLASAGELQKGN
 ELALPSKYVNFYLYLASHYEKLKGSPEDEQKQLFVEQHKHYLDEIIEQISEFSKRVI
 LADANLDKVL SAYNKHDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTS

10

20

TKEVL DATLIHQSI TGLYETRIDLSQLGGDGSPKKKRKVSSDYKDHDGDYKDHDIDY
 KDDDDK (配列番号 24)

30

【化 12】

FokI-dCas9 スクレオチド配列 (FokI-G4S-dCas9-nls-3XFLAG)

ATGCAACTAGTCAAAAGTGAACCTGGAGGAGAAGAAATCTGAACTTCGTCATAAATTG
 AAATATGTGCCTCATGAATATATTGAATTAATTGAAATTGCCAGAAATCCACTCAG
 GATAGAATTCTTGAAATGAAGGTAATGGAATTTTTTATGAAAGTTTATGGATATAGA
 GGTAACATTTGGGTGGATCAAGGAAACCGGACGGAGCAATTTATACTGTGCGATCT
 CCTATTGATTACGGTGTGATCGTGGATACTAAAGCTTATAGCGGAGGTTATAATCTG
 CCAATTGGCCAAGCAGATGAAATGCAACGATATGTGGAAGAAAATCAAACACGAAAC
 AAACATATCAACCCTAATGAATGGTGGAAAGTCTATCCATCTTCTGTAACGGAATTT
 AAGTTTTTATTTGTGAGTGGTCACTTTAAAGGAACTACAAAGCTCAGCTTACACGA
 TTAAATCATATCACTAATTGTAATGGAGCTGTTCTTAGTGTAGAAGAGCTTTTAATT
 GGTGGAGAAATGATTAAAGCCGGCACATTAACCTTAGAGGAAGTCAGACGGAATTT
 AATAACGGCGAGATAAACTTTGGTGGCGGTGGATCCGATAAAAAGTATTCTATTGGT
 TTAGCCATCGGCACTAATTCCGTTGGATGGGCTGTCATAACCGATGAATACAAAGTA
 CCTTCAAAGAAATTTAAGGTGTTGGGGAACACAGACCGTCATTTCGATTAAAAAGAAT
 CTTATCGGTGCCCTCCTATTTCGATAGTGGCGAAACGGCAGAGGCGACTCGCCTGAAA
 CGAACCGCTCGGAGAAGGTATACACGTCGCAAGAACCGAATATGTTACTTACAAGAA
 ATTTTTAGCAATGAGATGGCCAAAGTTGACGATTCTTTCTTTACCGTTTGGAAAGAG
 TCCTTCCTTGTCGAAGAGGACAAGAAACATGAACGGCACCCCATCTTTGGAAACATA
 GTAGATGAGGTGGCATATCATGAAAAGTACCAACGATTTATCACCTCAGAAAAAAG
 CTAGTTGACTCAACTGATAAAGCGGACCTGAGGTTAATCTACTTGGCTCTTGCCCAT
 ATGATAAAGTTCCGTGGGCACCTTTCTCATTGAGGGTGATCTAAATCCGGACAACCTCG
 GATGTCGACAAACTGTTTCATCCAGTTAGTACAAACCTATAATCAGTTGTTTGAAGAG
 AACCTTATAATGCAAGTGGCGTGGATGCGAAGGCTATTCTTAGCGCCCGCCTCTCT
 AAATCCCGACGGCTAGAAAACTGATCGCACAAATTACCCGGAGAGAGAGAAAAATGGG
 TTGTTTCGGTAACCTTATAGCGCTCTCACTAGGCCTGACACCAATTTTAAGTCGAAC
 TTCGACTTAGCTGAAGATGCCAAATTGCAGCTTAGTAAGGACACGTACGATGACGAT
 CTCGACAATCTACTGGCACAAATTGGAGATCAGTATGCGGACTTATTTTTGGCTGCC
 AAAACCTTAGCGATGCAATCCTCCTATCTGACATACTGAGAGTTAATACTGAGATT
 ACCAAGGCGCCGTTATCCGCTTCAATGATCAAAAGGTACGATGAACATCACCAAGAC
 TTGACACTTCTCAAGGCCCTAGTCCGTGAGCAACTGCCTGAGAAATATAAGGAAATA
 TTCTTTGATCAGTCGAAAAACGGGTACGCAGGTTATATTGACGGCGGAGCGAGTCAA
 GAGGAATTCTACAAGTTTATCAAACCCATATTAGAGAAGATGGATGGGACGGAAGAG
 TTGCTTGTAACCTCAATCGCGAAGATCTACTGCGAAAGCAGCGGACTTTTCGACAAC
 GGTAGCATTCACATCAAATCCACTTAGGCGAATTGCATGCTATACTTAGAAGGCAG
 GAGGATTTTTATCCGTTTCTCAAAGACAATCGTGAAAAGATTGAGAAAATCCTAACC
 TTTCGCATACCTTACTATGTGGGACCCCTGGCCCCGAGGGAATCTCGGTTTCGCATGG
 ATGACAAGAAAGTCCGAAGAAACGATTACTCCATGGAATTTTGAGGAAGTTGTGCGAT
 AAAGGTGCGTCAGCTCAATCGTTTCATCGAGAGGATGACCAACTTTGACAAGAATTTA
 CCGAACGAAAAAGTATTGCCAAGCACAGTTTACTTTACGAGTATTTACAGTGTAC
 AATGAATCAGCAAAGTTAAGTATGTCACTGAGGGCATGCGTAAACCCGCCTTTCTA
 AGCGGAGAACAGAAGAAAGCAATAGTAGATCTGTTATTCAAGACCAACCGCAAAGTG
 ACAGTTAAGCAATTGAAAGAGGACTACTTTAAGAAAATTGAATGCTTCGATTCTGTC
 GAGATCTCCGGGGTAGAAGATCGATTTAATGCGTCACTTGGTACGTATCATGACCTC
 CTAAAGATAATTAAAGATAAGGACTTCCTGGATAACGAAGAGAATGAAGATATCTTA
 GAAGATATAGTGTGACTCTTACCCTCTTTGAAGATCGGGAAATGATTGAGGAAAGA
 CTAAAAACATACGCTCACCTGTTTCGACGATAAGGTTATGAAACAGTTAAAGAGGCGT

10

20

30

40

CGCTATACGGGCTGGGGACGATTGTGCGGAACTTATCAACGGGATAAGAGACAAG
CAAAGTGGTAAACTATTCTCGATTTTCTAAAGAGCGACGGCTTCGCCAATAGGAAC
TTTATGCAGCTGATCCATGATGACTCTTTAACCTTCAAAGAGGATATACAAAAGGCA
CAGGTTTCCGGACAAGGGGACTCATTGCACGAACATATTGCGAATCTTGCTGGTTTCG
CCAGCCATCAAAAAGGGCATACTCCAGACAGTCAAAGTAGTGGATGAGCTAGTTAAG
GTCATGGGACGTCACAAACCGGAAAACATTGTAATCGAGATGGCACGCGAAAATCAA
ACGACTCAGAAGGGGCAAAAAACAGTCGAGAGCGGATGAAGAGAATAGAAGAGGGT
ATTAAAGAACTGGGCAGCCAGATCTTAAAGGAGCATCCTGTGGAAAATACCCAATTG
CAGAACGAGAACTTTACCTCTATTACCTACAAAATGGAAGGGACATGTATGTTGAT
CAGGAACCTGGACATAAACCGTTTATCTGATTACGACGTCGATGCCATTGTACCCCAA
TCCTTTTTTGAAGGACGATTCAATCGACAATAAAGTGCTTACACGCTCGGATAAGAAC
CGAGGGAAAAGTGACAATGTTCCAAGCGAGGAAGTCGTAAAGAAAATGAAGAACTAT
TGGCGGCAGCTCCTAAATGCGAACTGATAACGCAAAGAAAGTTCGATAACTTAACT
AAAGCTGAGAGGGGTGGCTTGTCTGAACTTGACAAGGCCGGATTTATTAACGTCAG
CTCGTGGAAACCCGCCAAATCACAAAGCATGTTGCACAGATACTAGATTCCCGAATG
AATACGAAATACGACGAGAACGATAAGCTGATTGCGGAAGTCAAAGTAATCACTTTA
AAGTCAAAATTGGTGTGCGACTTCAGAAAGGATTTTCAATTCTATAAAGTTAGGGAG
ATAAATAACTACCACCATGCGCACGACGCTTATCTTAATGCCGTCGTAGGGACCGCA
CTCATTAAGAAATACCCGAAGCTAGAAAGTGAGTTTGTGTATGGTGATTACAAAGTT
TATGACGTCCGTAAGATGATCGCGAAAAGCGAACAGGAGATAGGCAAGGCTACAGCC
AAATACTTCTTTTATTCTAACATTATGAATTTCTTTAAGACGGAAATCACTCTGGCA
AACGGAGAGATACGCAAACGACCTTTAATTGAAACCAATGGGGAGACAGGTGAAATC
GTATGGGATAAGGGCCGGGACTTCGCGACGGTGAGAAAAGTTTTGTCCATGCCCAA
GTCAACATAGTAAAGAAAACCTGAGGTGCAGACCGGAGGGTTTTCAAAGGAATCGATT
CTTCCAAAAGGAATAGTGATAAGCTCATCGCTCGTAAAAAGGACTGGGACCCGAAA
AAGTACGGTGGCTTCGATAGCCCTACAGTTGCCTATTCTGTCCTAGTAGTGCCAAA
GTTGAGAAGGGAAAATCCAAGAACTGAAGTCAGTCAAAGAATTATTGGGGATAACG
ATTATGGAGCGCTCGTCTTTTGAAGAAACCCCATCGACTTCCTTGAGGCGAAAGGT
TACAAGGAAGTAAAAAAGGATCTCATAATTAACTACCAAAGTATAGTCTGTTTGAG
TTAGAAAATGGCCGAAAACGGATGTTGGCTAGCGCCGGAGAGCTTCAAAGGGGAAC
GAACTCGCACTACCGTCTAAATACGTGAATTTCTGTATTTAGCGTCCCATTACGAG
AAGTTGAAAGGTTACCTGAAGATAACGAACAGAAGCAACTTTTTGTTGAGCAGCAC
AAACATTATCTCGACGAAATCATAGAGCAAATTTGGAATTCAGTAAGAGAGTCATC
CTAGCTGATGCCAATCTGGACAAAGTATTAAGCGCATACAACAAGCACAGGGATAAA
CCCATACGTGAGCAGGCGGAAAATATTATCCATTTGTTTACTCTTACCAACCTCGGC
GCTCCAGCCGATTCAAGTATTTTGACACAACGATAGATCGCAAACGATACACTTCT
ACCAAGGAGGTGCTAGACGCGACACTGATTACCAATCCATCACGGGATTATATGAA
ACTCGGATAGATTTGTACAGCTTGGGGGTGACGGATCCCCAAGAAGAAGAGGAAA
GTCTCGAGCGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTAC
AAGGATGACGATGACAAGTGA (配列番号 25)

10

20

30

【0102】

代替として、ヒトコドン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】

Nls-FokI-dCas9-nls アミノ酸配列

MDKKYSIGLAIGTNSVGWAVITDEYKVPSPKFKVLGNTDRHSIKKNLIGALLFDSGE
TAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHE
RHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIE

40

GDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQ
 LPGEKKNGLFGNLIASLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQ
 YADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQ
 LPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKNREDLL
 RKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPIYYVGPLA
 RGNSRFAMWTRKSEETITPWNFEVVDKGASASQSFIERMTNFDKNLPNEKVLPHSL
 LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFK
 KIECFDSVEISGVEDRFNASLGTYHDLKIIKDKDFLDNEENEDILEDIVLTTLTFE
 DREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLK
 SDGFANRNFQMLIHDDSLTFKEDIQKAQVSGQDLSLHEHIANLAGSPAICKGILQTV
 KVVDELVKVMGRHKPENIVIEEMARENQTTQKGQKNSRERMKRIEIEGKELGSQILKE
 HPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVAIVPQSFLKDDSIDNK
 VLTRSDKNRGKSDNVPSEEVVKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELD
 KAGFIKRLQVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKD
 FQFYKVRINNYYHHAHDAYLNAVVGTALEKYPKLESEFVYGDYKVDVRKMIKASE
 QEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATV
 RKVLSMPQVNIVKKEVQTTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVA
 YSVLVVAKVEKGSKKLKSVELLGITIMERSSEFKNPIDFLEAKGYKEVKKDLIIK
 LPKYSLEFELNGRKRMLASAGELQKGNELALPSKYVNFYLYLASHYEKLKGSPEDEQ
 KQLFVEQHKHYLDEIEQISEFSKRVLADANLDKVL SAYNKHDKPIREQAENIIH
 LFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSIETGLYETRIDLSQLGGD
 GSPKKKRKVSSDYKDHDGDYKDHDIDYKDDDDK (配列番号 26)

10

20

【化 14】

N1s-FokI-dCas9-n1s ヌクレオチド配列

ATGCCTAAGAAGAAGCGGAAGGTGAGCAGCCAACTTGTGAAGTCTGAACTCGAGGAG
 AAAAAATCAGAGTTGAGACACAAGTTGAAGTACGTGCCACACGAATACATCGAGCTT
 ATCGAGATCGCCAGAAACAGTACCCAGGATAGGATCCTTGAGATGAAAGTCATGGAG
 TTCTTTATGAAGGTCTACGGTTATAGAGGAAAGCACCTTGGCGGTAGCAGAAAGCCC
 GATGGCGCCATCTATACTGTCTGGATCTCCTATCGATTATGGGGTGATCGTGGATACC
 AAAGCTTACTCAGGCGGGTACAACCTTGCCCATAGGACAAGCCGACGAGATGCAGCGG
 TATGTCTGAAGAGAACCAGACGCGCAACAAGCACATCAACCCCAATGAATGGTGGAAA
 GTGTACCCAAGTAGTGTGACTGAGTTCAAGTTCCTGTTTGTCTCCGGCCACTTTAAG
 GGCAATTATAAAGCTCAGCTCACTAGACTCAATCACATCACAACTGCAACGGAGCT
 GTGTTGTCTAGTGGAGGAGCTCCTGATTGGAGGCGAGATGATCAAAGCCGGCACCCTT
 AACTGGAGGAGGTGCGGCGGAAGTTCAACAATGGAGAGATCAACTTCGGTGGCGGT
 GGATCCGATAAAAAGTATTCTATTGGTTTAGCCATCGGCACTAATTCGGTTGGATGG
 GCTGTCTATAACCGATGAATACAAAGTACCTTCAAAGAAATTTAAGGTGTTGGGGAAC
 ACAGACCGTCATTCGATTAAAAAGAATCTTATCGGTGCCCTCCTATTTCGATAGTGGC
 GAAACGGCAGAGGCGACTCGCCTGAAACGAACCGCTCGGAGAAGGTATACACGTCGC
 AAGAACCGAATATGTTACTTACAAGAAATTTTAGCAATGAGATGGCCAAAGTTGAC
 GATTCTTTCTTTTACCGTTTGGAAAGAGTCCTTTCCTTGTCTGAAGAGGACAAGAAACAT
 GAACGGCACCCCATCTTTGGAAACATAGTAGATGAGGTGGCATATCATGAAAAGTAC
 CCAACGATTTATCACCTCAGAAAAAGCTAGTTGACTCAACTGATAAAGCGGACCTG
 AGGTTAATCTACTTGGCTCTTGCCCATATGATAAAGTTCCGTGGGCACTTTCTCATT
 GAGGGTGATCTAAATCCGGACAACCTCGGATGTCTGACAACTGTTTCATCCAGTTAGTA
 CAAACCTATAATCAGTTGTTTGAAGAGAACCCTATAAATGCAAGTGGCGTGGATGCG
 AAGGCTATTCTTAGCGCCCGCCTCTCTAAATCCCGACGGCTAGAAAACCTGATCGCA
 CAATTACCCGGAGAGAAGAAAAATGGGTTGTTTCGGTAACCTTATAGCGCTCTCACTA
 GGCCTGACACCAAATTTTAAGTCGAACCTTCGACTTAGCTGAAGATGCCAAATTGCAG

30

40

CTTAGTAAGGACACGTACGATGACGATCTCGACAATCTACTGGCACAAATTGGAGAT
CAGTATGCGGACTTATTTTTGGCTGCCAAAACCTTAGCGATGCAATCCTCCTATCT
GACATACTGAGAGTTAATACTGAGATTACCAAGGCGCCGTTATCCGCTTCAATGATC
AAAAGGTACGATGAACATCACCAAGACTTGACACTTCTCAAGGCCCTAGTCCGTCAG
CAACTGCCTGAGAAATATAAGGAAATATTCTTTGATCAGTCGAAAAACGGGTACGCA
GGTTATATTGACGGCGGAGCGAGTCAAGAGGAATTCTACAAGTTTATCAAACCCATA
TTAGAGAAGATGGATGGGACGGAAGAGTTGCTTGTAACCACTCAATCGCGAAGATCTA
CTGCGAAAGCAGCGGACTTTTCGACAACGGTAGCATTCCACATCAATCCACTTAGGC
GAATTGCATGCTATACTTAGAAGGCAGGAGGATTTTTATCCGTTCTCAAAGACAAT
CGTGAAAAGATTGAGAAAATCCTAACCTTTTCGCATACCTTACTATGTGGGACCCCTG
GCCCCGAGGGAACCTCTCGGTTTCGCATGGATGACAAGAAAGTCCGAAGAAACGATTACT
CCATGGAATTTTGAGGAAGTTGTGCGATAAAGGTGCGTCAGCTCAATCGTTCATCGAG
AGGATGACCAACTTTGACAAGAATTTACCGAACGAAAAAGTATTGCCTAAGCACAGT
TTACTTTACGAGTATTTTACAGTGTACAATGAACTCACGAAAGTTAAGTATGTCACT
GAGGGCATGCGTAAACCCGCCTTTCTAAGCGGAGAACAGAAGAAAGCAATAGTAGAT
CTGTTATTCAAGACCAACCGCAAAGTGACAGTTAAGCAATTGAAAGAGGACTACTTT
AAGAAAATTGAATGCTTCGATTCTGTGCGAGATCTCCGGGGTAGAAGATCGATTTAAT
GCGTCACTTGGTACGTATCATGACCTCCTAAAGATAATTAAAGATAAGGACTTCCTG
GATAACGAAGAGAATGAAGATATCTTAGAAGATATAGTGTGACTCTTACCCTCTTT
GAAGATCGGGAAATGATTGAGGAAAGACTAAAAACATACGCTCACCTGTTTCGACGAT
AAGGTTATGAAACAGTTAAAGAGGCGTCGCTATACGGGCTGGGGACGATTGTCGCGG
AACTTATCAACGGGATAAGAGACAAGCAAAGTGGTAAACTATTCTCGATTTTCTA
AAGAGCGACGGCTTCGCCAATAGGAACCTTTATGCAGCTGATCCATGATGACTCTTTA
ACCTTCAAAGAGGATATACAAAAGGCACAGGTTTCCGGACAAGGGGACTCATTGCAC
GAACATATTGCGAATCTTGCTGGTTCCGACGCCATCAAAAAGGGCATACTCCAGACA
GTCAAAGTAGTGGATGAGCTAGTTAAGGTGATGGGACGTCACAAACCGGAAAACATT
GTAATCGAGATGGCAGCGGAAAATCAAACGACTCAGAAGGGGCAAAAAACAGTCGA
GAGCGGATGAAGAGAATAGAAGAGGGTATTAAAGAACTGGGCAGCCAGATCTTAAAG
GAGCATCCTGTGGAAAATACCCAATTGCAGAACGAGAACTTTACCTCTATTACCTA
CAAAATGGAAGGGACATGTATGTTGATCAGGAACTGGACATAAACCGTTTATCTGAT
TACGACGTCGATGCCATTGTACCCCAATCCTTTTTGAAGGACGATTCAATCGACAAT
AAAGTGCTTACACGCTCGGATAAGAACCGAGGGAAAAGTGACAATGTTCCAAGCGAG
GAAGTCGTAAAGAAAATGAAGAACTATTGGCGGCAGCTCCTAAATGCGAAACTGATA
ACGCAAAGAAAGTTTCGATAACTTAACTAAAGCTGAGAGGGGTGGCTTGTCTGAACCT
GACAAGGCCGATTTATTAACGTCAGCTCGTGGAAACCCGCCAAATCACAAAGCAT
GTTGCACAGATACTAGATTCCCGAATGAATACGAAATACGACGAGAACGATAAGCTG
ATTTCGGGAAGTCAAAGTAATCACTTTAAAGTCAAATTTGGTGTGCGACTTCAGAAAG
GATTTTCAATTCTATAAAGTTAGGGAGATAAATAACTACCACCATGCGCACGACGCT
TATCTTAATGCCGTCGTAGGGACCGCACTCATTAAGAAATACCCGAAGCTAGAAAGT
GAGTTTGTGTATGGTGATTACAAAGTTTATGACGTCCGTAAGATGATCGCGAAAAGC
GAACAGGAGATAGGCAAGGCTACAGCCAAATACTTCTTTTATTCTAACATTATGAAT
TTCTTTAAGACGGAAATCACTCTGGCAAACGGAGAGATACGCAAACGACCTTTAATT
GAAACCAATGGGGAGACAGGTGAAATCGTATGGGATAAGGGCCGGGACTTCGCGACG
GTGAGAAAAGTTTTGTCCATGCCCCAAGTCAACATAGTAAAGAAAACGAGGTGCAG
ACCGGAGGGTTTTCAAAGGAATCGATTCTTCCAAAAGGAATAGTGATAAGCTCATC
GCTCGTAAAAAGGACTGGGACCCGAAAAAGTACGGTGGCTTCGATAGCCCTACAGTT
GCCTATTCTGTCTAGTAGTGGCAAAAGTTGAGAAGGGAAAATCCAAGAACTGAAG
TCAGTCAAAGAATTATTGGGGATAACGATTATGGAGCGCTCGTCTTTTGAAAAGAAC
CCCATCGACTTCCTTGAGGCGAAAGGTTACAAGGAAGTAAAAAAGGATCTCATAATT
AAACTACCAAAGTATAGTCTGTTTGAGTTAGAAAATGGCCGAAAACGGATGTTGGCT

10

20

30

40

AGCGCCGGAGAGCTTCAAAAGGGGAACGAACTCGCACTACCGTCTAAATACGTGAAT
 TTCCTGTATTTAGCGTCCCATTACGAGAAGTTGAAAGGTTACCTGAAGATAACGAA
 CAGAAGCAACTTTTTGTTGAGCAGCACAAACATTATCTCGACGAAATCATAGAGCAA
 ATTTTCGGAATTCAGTAAGAGAGTCATCCTAGCTGATGCCAATCTGGACAAAGTATTA
 AGCGCATACAACAAGCACAGGGATAAACCCATACGTGAGCAGGCGGAAAATATTATC
 CATTTGTTTACTCTTACCAACCTCGGCGCTCCAGCCGCATTCAAGTATTTTGACACA
 ACGATAGATCGCAAACGATACACTTCTACCAAGGAGGTGCTAGACGCGGACACTGATT
 CACCAATCCATCACGGGATTATATGAAACTCGGATAGATTTGTACAGCTTGGGGGT
 GACGGATCCCCCAAGAAGAAGAGGAAAGTCTCGAGCGACTACAAAGACCATGACGGT
 GATTATAAAGATCATGACATCGATTACAAGGATGACGATGACAAGTGA

(配列番号 27)

10

【0103】

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、400 μg/mlのG418の存在下で培養した。

20

【0104】

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パーサー長解析の実験では、750 ngのpCAG-Csy4-FokI-dCas9-nlsヌクレアーゼプラスミドおよび250 ng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50 ng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、975 ng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、325 ngのgRNAベクターおよび10 ng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、750 ngのヌクレアーゼプラスミド、250 ngのgRNA発現プラスミド、および10 ng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

【0105】

単一のトランスフェクションを、最初のスパーサー活性スクリーニングのために実施し、焦点を当てたスパ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

【0106】

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 Biotechnol 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。

【0107】

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、Sciclone 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 DNAdvance 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/10

【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クロニングシ(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'-GTAAACGACGGCCAG-3';配列番号:19)を使用して配列決定した。20

【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、Sciclone 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ループッ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。30

【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。40

【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ィ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索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50

依存性 F o k I ヌクレアーゼドメインを、RNA 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 DNA 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、DNA 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のであると仮定された。

10

【 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 RNA 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 (l o w) 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 1 6 倍改善するであろう。

【 0 1 1 4 】

20

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 RNA 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 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 RNA 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。

30

【 0 1 1 5 】

実施例 2 c . 二量体 RNA 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、DNA 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 6 0 対の g R N A を設計した。これら g R N A 対のそれぞれが標的化とする 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

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 -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
1	EGFP 部位 1	74	GAGCTGGACGGCGACGTAAACGG	28.	CGCCGGACACGCTGAACCTGTGG	29.	0	内部
2	EGFP 部位 2	174	CCGGCAAGCTGCCCGTGCCCTGG	30.	GGTCAGGGTGGTCACGAGGGTGG	31.	1	内部
3	EGFP 部位 3	37	CGAGGAGCTGTTCAACGGGGTGG	32.	CCGTCCAGCTCGACCAGGATGGG	33.	2	内部
4	EGFP 部位 4	37	CGAGGAGCTGTTCAACGGGGTGG	34.	GCCGTCCAGCTCGACCAGGATGG	35.	3	内部
5	EGFP 部位 5	174	CCGGCAAGCTGCCCGTGCCCTGG	36.	GTAGGTCAAGGGTGGTCACGAGGG	37.	4	内部
6	EGFP 部位 6	34	GGGCGAGGAGCTGTTCAACGGGG	38.	CCGTCCAGCTCGACCAGGATGGG	39.	5	内部
7	EGFP 部位 7	33	AGGGCGAGGAGCTGTTCAACGGG	40.	CCGTCCAGCTCGACCAGGATGGG	41.	6	内部
8	EGFP 部位 8	32	AAGGCGAGGAGCTGTTCAACGG	42.	CCGTCCAGCTCGACCAGGATGGG	43.	7	内部
9	EGFP 部位 9	32	AAGGCGAGGAGCTGTTCAACGG	44.	GCCGTCCAGCTCGACCAGGATGG	45.	8	内部
10	EGFP 部位 10	106	CAGCGTGTCCGGCGAGGGCGAGG	46.	CTTCAGGGTCAGCTTGCCGTAGG	47.	9	内部

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
11	EGFP 部位 11	34	GGGCGAGGAGCTGTTCAACGGGG	48.	CGTCGCCGTCCAGCTCGACCAGG	49.	10	内部
12	EGFP 部位 12	33	AGGGCGAGGAGCTGTTCAACGGG	50.	CGTCGCCGTCCAGCTCGACCAGG	51.	11	内部
13	EGFP 部位 13	32	AAGGCGAGGAGCTGTTCAACGG	52.	CGTCGCCGTCCAGCTCGACCAGG	53.	12	内部
14	EGFP 部位 14	155	CTGAAGTTCATCTGCACCAACGG	54.	GTGGTCACGAGGGTGGGCCAGGG	55.	13	内部
15	EGFP 部位 15	101	AAGTTCAGCGTGTCCGGCGAGGG	56.	CTTCAGGGTCAGCTTGCCGTAGG	57.	14	内部
16	EGFP 部位 16	100	CAAGTTCAGCGTGTCCGGCGAGG	58.	CTTCAGGGTCAGCTTGCCGTAGG	59.	15	内部
17	EGFP 部位 17	58	GGTGGCCATCCTGGTCGAGCTGG	60.	CGCCGGACACGCTGAACCTGTGG	61.	16	内部
18	EGFP 部位 18	74	GAGCTGGACGGCGACGTAAACGG	62.	GGCATCGCCCTCGCCCTCGCCGG	63.	17	内部
19	EGFP 部位 19	307	GGAGCGCACCATCTTCTTCAGG	64.	CTCGAATTCACCTCGGCGCGGG	65.	18	内部
20	EGFP 部位 20	155	CTGAAGTTCATCTGCACCAACGG	66.	GTCAGGGTGGTCACGAGGGTGGG	67.	19	内部
21	EGFP 部位 21	95	GGCCACAAGTTCAGCGTGTCCGG	68.	CTTCAGGGTCAGCTTGCCGTAGG	69.	20	内部

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 「スパー サー」 距離	PAM
22	EGFP 部位 22	203	CTCGTGACCAACCTGACCTACGG	70.	CGTGCCTGCTTCATGTGGTCGGGG	71.	21	内部
23	EGFP 部位 23	174	CCGGCAAGCTGCGCGTGGCTGG	72.	GCTGAAGCACTGCACGCGTAGG	73.	22	内部
24	EGFP 部位 24	107	AGCGTGTCGGCGAGGCGGAGGG	74.	GGTGGTGCAGATGAACCTCAGGG	75.	23	内部
25	EGFP 部位 25	106	CAGCGTGTCGGCGAGGCGGAGGG	76.	GGTGGTGCAGATGAACCTCAGGG	77.	24	内部
26	EGFP 部位 26	49	CACCGGGGTGGTGGCCATCTGG	78.	CGCCGGACACGCTGAACCTGTGG	79.	25	内部
27	EGFP 部位 27	122	GGCGAGGCGGATGCCACTACGG	80.	GGGCGAGGCGAGCTTGGCGGTGG	81.	26	内部
28	EGFP 部位 28	203	CTCGTGACCAACCTGACCTACGG	82.	AGAAGTCGTGCTGCTTCATGTGG	83.	27	内部
29	EGFP 部位 29	337	CAACTACAAGACCGCGCGGAGG	84.	CGATGCCCTTCAGCTCGATGCGG	85.	28	内部
30	EGFP 部位 30	62	CCCATCTGGTGCAGCTGGACGG	86.	GGCATCGCCCTCGCCCTCGCCGG	87.	29	内部
31	EGFP 部位 31	100	CAAGTTCAGCGTGTCGGCGGAGG	88.	GGTGGTGCAGATGAACCTCAGGG	89.	30	内部
32	EGFP 部位 32	74	GAGCTGGACGGCGACGTAAACGG	90.	GACCAGGATGGGACACACCCCGG	91.	0	外部

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 「スパー サー」 距離	PAM
33	EGFP 部位 33	314	ACCATCTTCTTCAAGGACGACGG	92.	CGCTCCTGGACGTAGCCCTTCGGG	93.	1	外部
34	EGFP 部位 34	122	GGCGAGGCGGATGCCACTACGG	94.	CGCCGGACACGCTGAACCTGTGG	95.	2	外部
35	EGFP 部位 35	275	TTCAAGTCCGCCATGCCCGAAGG	96.	GTCGTGCTGCTTCATGTGGTCGG	97.	3	外部
36	EGFP 部位 36	275	TTCAAGTCCGCCATGCCCGAAGG	98.	TGCTGCTGCTTCATGTGGTCGGG	99.	4	外部
37	EGFP 部位 37	95	GGCCACAAGTTCAGCGTGTCGGG	100.	CGTCGCGTCCAGCTCGACCGAGG	101.	5	外部
38	EGFP 部位 38	203	CTCGTGACCAACCTGACCTACGG	102.	CCAGGGCACGGGCGAGCTTGGCGG	103.	6	外部
39	EGFP 部位 39	463	CAGCCACAAGCTCTATATCATGG	104.	TGTAATCCAGCTTGTGCCCCAGG	105.	7	外部
40	EGFP 部位 40	95	GGCCACAAGTTCAGCGTGTCGGG	106.	GCCGTCCAGCTCGACCGAGGATGG	107.	9	外部
41	EGFP 部位 41	95	GGCCACAAGTTCAGCGTGTCGGG	108.	CCGTCCAGCTCGACCGAGGATGGG	109.	10	外部
42	EGFP 部位 42	101	AAGTTCAGCGTGTCGGCGGAGGG	110.	CGTCGCGTCCAGCTCGACCGAGG	111.	11	外部
43	EGFP 部位 43	350	CGCGCCGAGGTGAAGTTCGAGGG	112.	GCCGTGCTCCTTGAAGAAGATGG	113.	12	外部

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 「スパー サー」 距離	PAM
44	EGFP 部位 44	174	CCGGCAAGCTGCCCCGTGCTGG	114.	CTTCAGGGTCAGCTTGCCGTAGG	115.	13	外部
45	EGFP 部位 45	100	CAAGTTCAGCGTGTCCGGCGAGG	116.	GCCGTCCAGCTCGACCCAGGATGG	117.	14	外部
46	EGFP 部位 46	100	CAAGTTCAGCGTGTCCGGCGAGG	118.	CCGTCCAGCTCGACCCAGGATGGG	119.	15	外部
47	EGFP 部位 47	101	AAGTTCAGCGTGTCCGGCGAGG	120.	CCGTCCAGCTCGACCCAGGATGGG	121.	16	外部
48	EGFP 部位 48	107	AGCGTGTCCGGCGAGGGCGAGG	122.	CGTCGCCGTCCAGCTCGACCCAGG	123.	17	外部
49	EGFP 部位 49	155	CTGAAGTTCATCTGCACCCCGG	124.	GGCATCGCCCTCGCCCTCGCCGG	125.	18	外部
50	EGFP 部位 50	106	CAGCGTGTCCGGCGAGGGCGAGG	126.	GCCGTCCAGCTCGACCCAGGATGG	127.	20	外部
51	EGFP 部位 51	95	GGCCACAAGTTCAGCGTGTCCGG	128.	GACCAGGATGGGCACCCACCCCGG	129.	21	外部
52	EGFP 部位 52	107	AGCGTGTCCGGCGAGGGCGAGG	130.	CCGTCCAGCTCGACCCAGGATGGG	131.	22	外部
53	EGFP 部位 53	337	CAACTACAAGACCCGCGCCGAGG	132.	GCGCTCCTGGACGTAGCCTTCGG	133.	23	外部
54	EGFP 部位 54	337	CAACTACAAGACCCGCGCCGAGG	134.	CGCTCCTGGACGTAGCCTTCGGG	135.	24	外部

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 「スパー サー」 距離	PAM
55	EGFP 部位 55	397	GCTGAAGGGCATCGACTTCAGG	136.	CCTCGAACTTCACCTCGGCGCGG	137.	25	外部
56	EGFP 部位 56	100	CAAGTTCAGCGTGTCCGGCGAGG	138.	GACCAGGATGGGCACCCACCCCGG	139.	26	外部
57	EGFP 部位 57	101	AAGTTCAGCGTGTCCGGCGAGG	140.	GACCAGGATGGGCACCCACCCCGG	141.	27	外部
58	EGFP 部位 58	400	GAAGGGCATCGACTTCAGGAGG	142.	CCTCGAACTTCACCTCGGCGCGG	143.	28	外部
59	EGFP 部位 59	337	CAACTACAAGACCCGCGCCGAGG	144.	CTGGACGTAGCCTTCGGGCGATGG	145.	29	外部
60	EGFP 部位 60	307	GGAGCGCACCATCTTCTTCAGG	146.	AGAAGTCGTGCTGCTTCATGTGG	147.	31	外部
61	EGFP 部位 61	100	CAAGTTCAGCGTGTCCGGCGAGG	148.	CGTCGCCGTCCAGCTCGACCCAGG	149.	10	外部
62	EGFP 部位 62	286	CATGCCCCAAGGCTACGTCCAGG	150.	AGAAGTCGTGCTGCTTCATGTGG	151.	10	外部
63	EGFP 部位 63	337	CAACTACAAGACCCGCGCCGAGG	152.	TGAAGAAGATGGTGCGCTCCTGG	153.	10	外部
64	EGFP 部位 64	382	GGTGAACCGCATCGAGCTGAGG	154.	CCTCGAACTTCACCTCGGCGCGG	155.	10	外部
65	EGFP 部位 65	275	TTCAGTCCGCCATGCCCGAAGG	156.	GCTTCATGTGCTCGGGGTAGCGG	157.	11	外部

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 「スパー サー」 距離	PAM
66	EGFP 部位 66	349	CCGCGCCGAGGTGAAGTTGAGG	158.	GCCGTCGTCTTGAAGAAGATGG	159	11	外部
67	EGFP 部位 67	382	GGTGAACCGCATCGAGCTGAAGG	160.	CTCGAACTTCACCTCGGCGCGGG	161	11	外部
68	EGFP 部位 68	383	GTAACCGCATCGAGCTGAAGGG	162.	CCTCGAACTTCACCTCGGCGCGGG	163	11	外部
69	EGFP 部位 69	520	CAAGATCCGCCACAACATCGAGG	164.	GATGCCGTTCTTCTGCTTGTGG	165	11	外部
70	EGFP 部位 70	383	GTGAACCGCATCGAGCTGAAGGG	166.	CTCGAACTTCACCTCGGCGCGGG	167	12	外部
71	EGFP 部位 71	415	CAAGGAGGACGGCAACATCCTGG	168.	TCAGCTCGATGCGGTTTACCAGG	169	13	外部
72	EGFP 部位 72	286	CATGCCCGAAGGCTACGTCCAGG	170.	GTCGTGCTGCTTCATGTGGTGG	171	14	外部
73	EGFP 部位 73	415	CAAGGAGGACGGCAACATCCTGG	172.	CAGCTCGATGCGGTTTACCAGG	173	14	外部
74	EGFP 部位 74	416	AAGGAGGACGGCAACATCCTGGG	174.	TCAGCTCGATGCGGTTTACCAGG	175	14	外部
75	EGFP 部位 75	101	AAGTTCAGCGTGTCCGCGAGGG	176.	GCCGTCCAGCTCGACAGGATGG	177	15	外部
76	EGFP 部位 76	286	CATGCCCGAAGGCTACGTCCAGG	178.	TCGTGCTGCTTCATGTGGTGGG	179	15	外部

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 「スパー サー」 距離	PAM
77	EGFP 部位 77	416	AAGGAGGACGGCAACATCCTGGG	180.	CAGCTCGATGCGGTTTACCAGG	181	15	外部
78	EGFP 部位 78	417	AGGAGGACGGCAACATCCTGGGG	182.	TCAGCTCGATGCGGTTTACCAGG	183	15	外部
79	EGFP 部位 79	524	ATCCGCCACAACATCGAGGACGG	184.	GATGCCGTTCTTCTGCTTGTGG	185	15	外部
80	EGFP 部位 80	106	CAGCGTGTCCGCGAGGGCGAGG	186.	CGTCGCGTCCAGCTCGACAGG	187	16	外部
81	EGFP 部位 81	174	CCGGCAAGCTGCCCGTGCCTGG	188.	CAGGCTCAGCTTGCCGTAGGTGG	189	16	外部
82	EGFP 部位 82	286	CATGCCCGAAGGCTACGTCCAGG	190.	CGTGTGCTTCATGTGGTGGGG	191	16	外部
83	EGFP 部位 83	417	AGGAGGACGGCAACATCCTGGGG	192.	CAGCTCGATGCGGTTTACCAGG	193	16	外部
84	EGFP 部位 84	427	CAACATCCTGGGGCACAGCTGG	194.	CGATGCCCTTCAGCTCGATGCGG	195	16	外部
85	EGFP 部位 85	397	GCTGAAGGGCATCGACTTCAAGG	196.	GTCGCCCTCGAACTTCACCTCGG	197	20	外部

【 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超のスパーサー長の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の範囲のスパーサー長を備え、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のスパ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のスパーサー長で試験し、13bpスパ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。T7EI解析およびサンガー配列決定

10

20

30

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、F o k I - d C a s 9 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g R N Aにより配向できる。単純性のため、2つのF o k I - d C a s 9 融合物および2つのg R N Aの複合体は、本明細書中でR N A誘導型F o k Iヌクレアーゼ(R F N)を指す。

【0117】

E G F Pレポーター遺伝子を用いて最初の知見を拡張し、R F N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、g R N A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R F Nのうち7つは、T 7 E Iにより判断されるように、ヒトU 2 O S . E G F P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、H E K 2 9 3細胞における同一の12 R F N対を用いて得た(表2)。U 2 O S . E G F P細胞からうまく標的化したアレルのサンガー配列決定から、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った(図5 F)。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R F Nの強さを例証するものである。

【0118】

実施例2 d . R F Nは、それらの切断部位に対して拡張した特異性を有する

R F N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g R N A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g R N Aは、F o k I - d C a s 9 誘導型挿入欠失を効率的に配向することが可能でないはずである。最初の試験を実施するために、ヒトU 2 O S . E G F P細胞における標的部位(E G F P部位47および81)に対してF o k I - d C a s 9 - 誘導型挿入欠失を効率的に配向することを示した、E G F Pにおける2つの標的部位に配向した2対のg R N Aを使用した(図5 C)。V E G F A中の関連しない部位を標的とするg R N Aでのこれら2つの対のそれぞれの1つまたは他のg R N Aの置換は、E G F P崩壊活性の低減をもたらした(図6 A)、T 7 E I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6 B)。同様に、2つのg R N Aのそれぞれの1つのみを使用した効果を、ヒトA P C、M L H 1、およびV E G F A遺伝子におけるF o k I - d C a s 9 - 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、T 7 E Iアッセイにより検出可能なR F N誘導型挿入欠失の損失を観察した(図6 C)。これらの結果は、R F N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g R N Aを必要とすることを例証する。

【0119】

本出願人のR F Nの活性が2つのg R N Aの発現に依存することを考慮すると、対の中の単一g R N A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R F Nを標的化するために必要とされる2つのg R N Aのうちの1つとしてそれ自体が作用できる単一g R N Aにより誘導される単量体C a s 9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C a s 9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V E G F A遺伝子の二量体R F N部位を標的とするために使用したg R N A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、V E G F A - 標的化R F N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6 Cに示されるT 7 E I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6 Dおよび表3)。これらの結果から、R F Nの使用することにより、C a s 9ヌクレアーゼおよび単一g R N A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、F o k I - d C a s 9と共に発現した単一g

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、追加の部位上での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 C a s 9 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	左の標的配列	配列番号	右の標的配列	配列番号	U2OS 細胞または 293 細胞中の RFN 標的配列の内在性配列 (小文字はスパー配列)	配列番号
APC1	CCAGAAAGTACGAGCGCGC CCGG	198.	TGGCAGGTGAGTGAGGCT GCAGG	199.	CCGGCGGCGCTCGTACTTCTGGCactggcgag cgtcTGGCAGGTGAGTGAGGCTGCAGG	200.
BRCA1	GAATACCCATCTGTCAGCT TCGG	201.	GGCGGAACCTGAGAGGCG TAAGG	202.	CCGAAGCTGACAGATGGGTATTTctttgacggggg tagggCGGGAACCTGAGAGGCGTAAGG	203.
DDB2	AATATTCAAGCAGCAGGCA CAGG	204.	CTCGGCAGGAGGCTGCA GCGGG	205.	CCTGTGCTGCTGCTTGAATATTTccgccttttag ggtgCTCGGCGAGGAGGCTGCAGCGGG	206.
EMX1	CCCAAAGCCTGGCCAGGGA GTGG	207.	GCCCCACAGGCTTGAAG CCCGG	208.	CCACTCCCTGGCCAGGCTTTGGGgagcctggagt catgCCCCACAGGCTTGAAGCCCGG	209.
FANCF - 部位 1	CCCTACTTCGCTTTCACC TTGG	210.	GGAATCCCTTCTGCAGCA CCTGG	211.	CCAAGGTGAAGCGGAGTAGGGccttgcgcacc tcatGGAATCCCTTCTGCAGCACCTGG	212.
FANCF - 部位 2	CGCTCCAGAGCCGTGCGAA TGGG	213.	TGGAGGCAGAGGCGCGC TTTGG	214.	CCCATTGCGACGGCTCTGAGCGGcggtgcacaa ccagTGGAGGCAGAGGCGGCTTTGG	215.
FES	CGAGGAGACTGGGACTGT AGGG	216.	CCAGCTGCTGCTTGCCT CCAGG	217.	CCCTACAGTCCCCAGCTCTCGtcccatgcctcc gtctCCAGCTGCTGCTTGCCTCCAGG	218.
GLI1	CATAGCTACTGATTGGTGG TGGG	219.	CGGGCCCCCTCCCAAGTCA GGGGG	220.	CCCACCACCAATCAGTAGCTATGcgagcctgct gtctCCGGCCCCCTCCCAAGTCCAGGCGG	221.
MLH1	GGAAACGCTAGATGCTCA ACGG	222.	CAAAATGCTGTTGCTGGC AGTGG	223.	CCGTTGAGCATCTAGACGTTTCTgtgtctcttg gcgcCAAAATGCTGTTCTGTCGACAGGGG	224.
RAR1	CTGTTGCTGGCCATGCCAA GCGG	225.	CCTGGGGCGGGCACCTC AATGG	226.	CCGCTTGGCATGGCCAGCACAGcagctcctgccc gacaCCTGGGGCGGGCACCTCAATGG	227.
RUNX	TTCCGAGCGAAACCAAGA CAGG	228.	GAGTCCCCCGCTTCAGA AGAGG	229.	CCTGTCTTGGTTTTCGCTCCGAAggtataaagaaat cattGAGTCCCCCGCTTCAGAAAGG	230.
SS18	GGCCCGGTCGACTCCGGGC CCGG	231.	TGCTGGGAATCAGCAGTG TTTGG	232.	CCGGCCCCGAGTCGACCGGGCGgagcgaggcg ggcctGCTGGGAATCAGCAGTGTGTTGG	233.
VEGFA - 部位 1	GGGTGGGGGAGTTTGCTC CTGG	234.	TCCTCTTTAGCCAGAGC CGGGG	235.	CCAGGAGCAACTCCCCCACCCctttccaaagc ccatTCCCTCTTTAGCCAGAGCCCGGG	236.
VEGFA - 部位 2	GCCGCCGCGGGGAGGAG GTGG	237.	GGCGAGCCGCGGCGAGGG GCCGG	238.	CCACTCTCTCCCGCGCGGCGgacagtgagcg cggcGGCGAGCCGCGGCGAGGCGCGG	239.
VEGFA - 部位 3	CCGTCTGCACACCCCGGCT CTGG	240.	CTCGGCCACACAGGGAA GCTGG	241.	CCAGAGCCGGGGTGTGACAGCGcagtcactaggg ggcgCTCGGCCACACAGGGAAAGCTGG	242.

10

20

30

40

表 3								
遺伝子名	T7E1 アッセイに使用されるプライマー 1	配列番号	T7E1 アッセイに使用されるプライマー 2	配列番号	OM50 の追加 (はい/いいえ)	サマールサイクロンのプロトコル	アンプルの寸法 (bp)	T7E1 の産物の生成切断の寸法 (bp)
APC1	GGCTGTGGGAAGCCAGCAA C	243.	AAGCCAGGGGCCA ACTGGAG	244.	いいえ	タッチダウン	634	447/187
BRCA1	GGCGGGAATTACAGATAA ATTAAA	245.	AGTCCCATCTCTC TCATACATACCA	246.	いいえ	タッチダウン	751	454/297
DDI2	ACCGCCCTTGGCACCAC	247.	CGGAGCTCATCTG CTTCCTGT	248.	いいえ	タッチダウン	627	456/171
EMX1	GGAGCAGCTGGTCAGAGGG G	249.	GGGAAGGGGGAC ACTGGGGA	250.	はい	2-ステップ	729	480/249
FANCF - 部位 1	GGCCTACATCTGCTCTCCCT CCA	251.	GGGCCGGGAAAG AGTTGCTG	252.	いいえ	タッチダウン	634	361/273
FANCF - 部位 2	GGCCTACATCTGCTCTCCCT CCA	253.	GGGCCGGGAAAG AGTTGCTG	254.	いいえ	タッチダウン	634	466/168
FES	GGGAGGGAGGCTCCAGGT T	255.	GGCACATGGGTC CCAAAGCA	256.	いいえ	タッチダウン	633	395/238
GLI1	CCTTACCCCTCCCTCACTC A	257.	AGAAGGGCGGGC CAGACAGT	258.	いいえ	タッチダウン	869	590/279
MLH1	ATATCCTTCTAGGTAGCGGG CAGTAGCC	259.	TCTCGGGGAGAG CGGTAAA	260.	いいえ	タッチダウン	610	332/278
RAR1	CCCAGGAAAAGTGCCAGCT CA	261.	TGATGGTCACTCC AACTGGA	262.	いいえ	タッチダウン	632	335/297
RUNX	AAGCGGCGCTGGCTTTT	263.	CCAGCACAACTTA CTCGCACTGA	264.	いいえ	タッチダウン	626	389/237
SS18	GGGATGCAGGGACGTCAC G	265.	GCCGCCCATCCC TAGAGAAA	266.	いいえ	タッチダウン	629	455/174
VEGFA - 部位 1	TCCAGATGGCACATTGTCAG	267.	AGGAGCAGGAA AGTGAGGT	268.	いいえ	タッチダウン	531	338/193
VEGFA - 部位 2	AGAGAAGTCGAGGAAGAGA GAG	269.	CAGCAGAAAGTTC ATGGTTTCG	270.	はい	タッチダウン	756	482/274
VEGFA - 部位 3	TCCAGATGGCACATTGTCAG	271.	AGGAGCAGGAA AGTGAGGT	272.	いいえ	タッチダウン	531	288/243

10

20

30

40

表 3				
遺伝子名	ディープシーケンシングに使用されるプライマー 1	配列番号	ディープシーケンシングに使用されるプライマー 2	配列番号
APC1				
BRCA1				
DDB2	CGATGGCTCCCAAGAAACGC	273.	GCAGGTAGAAATGCACAGCCG	274.
EMX1				
FANCF - 部位 1	GCCCAGAGTCAAGGAACAG	275.	AGGTAGTGTCTTGAGACCGCC	276.
FANCF - 部位 2	CATCCATCGGCGCTTTGGTC	277.	CCGGGAAAGAGTTGCTGCAC	278.
FES	CTCCCCGTCTGCAGTCCATC	279.	CCTGCAGGGACATGTGGTGA	280.
GLI1				
MLH1				
RAR1				
RUNX	TAGGGCTAGAGGGGTGAGGC	281.	CCGAGGTGAAACAAGCTGCC	282.
SS18				
VEGFA - 部位 1	ATGAGGGCTCCAGATGGCAC	283.	TTCACCCAGCTTCCCTGTGG	284.
VEGFA - 部位 2				
VEGFA - 部位 3				

【 0 1 2 0 】

実施例 2 e . 単量体 Cas 9 ニッカーゼは、単一 gRNA / Fok I - dCas 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 Cas 9 ニッカーゼの手法の重要となる脆弱性は、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 Ran et al. , Cell 154 , 1380 - 1389 (2013) ; Mali et al. , Nat Biotechnol 31 , 833 - 838 (2013) ; Cho et al. , Genome Res (2013) ; and Mali et al. , Science 339 , 823 - 826 (2013) 参照）。対形成される Cas 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。RFNは、二量体化依存性 Fok 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 Cas 9 ニッカーゼで観察されるものと比較して、1つの gRNA のみの存在下で

10

20

30

40

50

の望ましくない挿入欠失活性がより少ないことを示すと仮定される。

【 0 1 2 1 】

この仮定を試験するために、F o k I - d C a s 9 および C a s 9 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 g R N A の存在下で比較した（合計 1 2 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 g R N A により配向される単量体 C a s 9 ニッカーゼが、これらの標的で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、F o k I - d C a s 9 または C a s 9 ニッカーゼのゲノム編集活性を、1 つまたは他の g R N A のうち両方または 1 つのみが存在する中で評価した。F o k I - d C a s 9 および C a s 9 ニッカーゼの両方は、2 つの g R N A 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、1 2 個の単一 g R N A により配向した単量体 C a s 9 ニッカーゼは、0 . 0 0 4 8 % ~ 3 . 0 4 % の範囲の頻度で挿入欠失を誘導した（図 7 A および表 5）。対照的に、同一の 1 2 個の単一 g R N A により配向した F o k I - d C a s 9 は、0 . 0 0 4 5 ~ 0 . 4 7 3 % の範囲のより低い頻度で挿入欠失を誘導した（図 7 A および表 5）。これらのデータを直接比較することにより、F o k I - d C a s 9 は、1 2 個の単一 g R N A のうち 1 0 個で C a s 9 ニッカーゼよりも低い頻度で、挿入欠失を誘導した（図 7 A および表 5）。さらに、F o k I - d C a s 9 は、単一 g R N A 比率と、対形成した g R N A の比率を比較する際に、1 2 個の片側部位のうち 1 1 個で C a s 9 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 7 B）。

【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

【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

ス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	125912	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

10

20

30

40

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
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
2							
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
1							
tdTomato 対照 (-)	FANCF 部位	なし	11	22647331	18	578378	0.0031
1							
tdTomato 対照 (-)	FANCF 部位	なし	11	22647138	81	393821	0.0206
2							
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
(-)							

【0122】

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、特定の単量体の 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 標的の「右」の片側部位に対する単一 gRNA と共発現した 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 8A)。類似の結果が、FANCF 標的の部位 1 の「右」の片側部位 (位置 16 で 16.3% の変異頻度) (図 8B) または RUNX1 標的の部位の「右」の片側部位 (位置 17 で 2% の変異頻度) (図 8C) に配向される Cas9 ニッカーゼおよび単一 gRNA 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 gRNA が細胞で発現していない対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 8A ~ 8C)。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 鎖上の 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 8A ~ 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。

【0123】

実施例 2 f . 二量体 RFN は高い度合いの特異性を有する。

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をもたれる。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
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
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., Shivalila, 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 engineering in

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).

Ding, Q., Regan, S. N., Xia, Y., Ostrom, 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).

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).

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 RNA-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

10

20

30

40

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 nucleases. *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., Joungh, J. K. & Liu, D. R. Revealing off-target cleavage specificities of zinc-finger nucleases by in vitro selection. *Nat Methods* 8, 765-770 (2011).

10

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., Khabadi, A. M., Polstein, L. R., Thakore, P. I., Glass, K. A., Ousterout, D. G., Long, K. W., et al. RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nat Methods* 10, 973-976. (2013).

20

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).

30

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., Joungh, J. K., and Dobbs, D. ZiFiT (Zinc Finger Targeter): an updated zinc finger engineering tool. *Nucleic Acids Res* 38, W462-468. (2010).

40

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 nuclease cleavage profiles reveals an expanded landscape of off-target sites. *Nucleic Acids Res.* (2013).

Sander, J. D., Zaback, P., Joungh, 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).

10

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).

20

Yang, L., Guell, M., Byrne, S., Yang, J. L., De Los Angeles, A., Mali, P., Aach, J., Kim-Kisela, 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。

30

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。

40

[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]

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。 10
- [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 RNA 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 RNA 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。 20
- [1 2]
- 前記ガイド RNA が、
- (a) 2 つの単一ガイド RNA であって、1 つの単一ガイド RNA が第 1 の鎖を標的と
- し、他方のガイド RNA が相補鎖を標的とし、F o k I が各鎖を切断して、反対の DNA
- 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 つの単一ガイド RNA
- 、または
- (b) t r a c r RNA および 2 つの c r RNA であって、1 つの c r RNA が第 1 の
- 鎖を標的とし、他方の c r RNA が相補鎖を標的とし、F o k I が各鎖を切断して、反対
- の DNA 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 RN
- A および 2 つの c r RNA 30
- である、[1 0] に記載の方法。
- [1 3]
- 前記 2 つのガイド RNA 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 RNA 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。 40
- [1 6]
- 細胞における 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、前記細胞を
- 、[1] ~ [6] に記載の RNA 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。

【図 1】

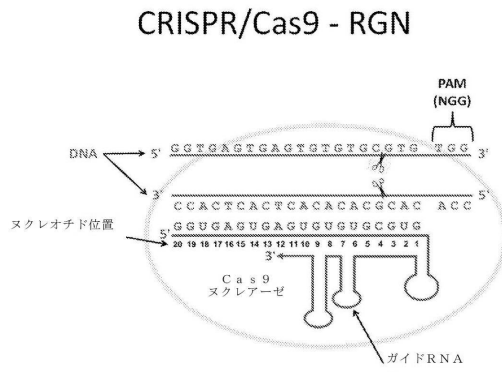


図 1

【図 2 A】

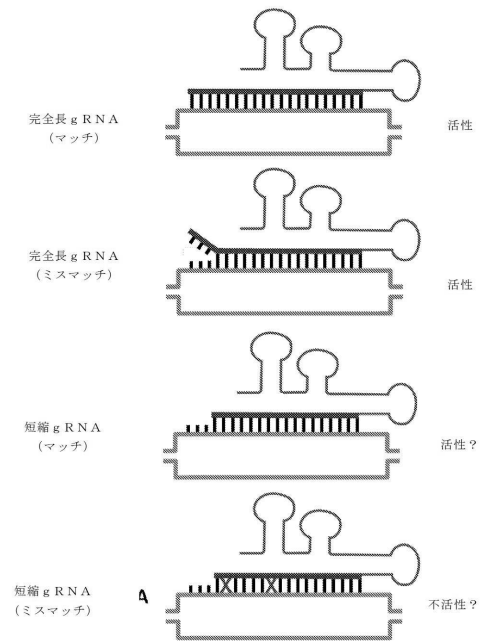


図 2 A

【図 2 B】

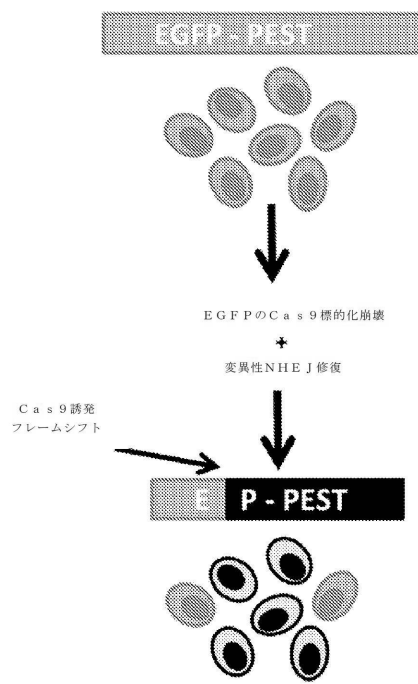


図 2 B

【図 2 C】

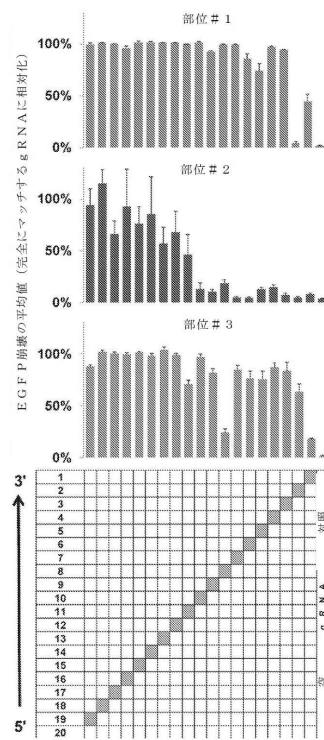


図 2 C

【図 2 D】

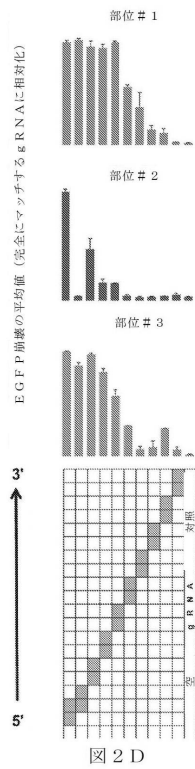


図 2 D

【図 2 E】

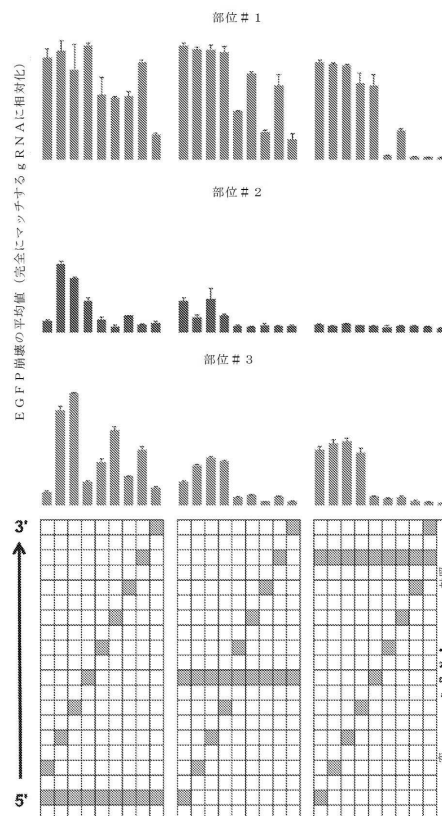


図 2 E

【図 2 F】

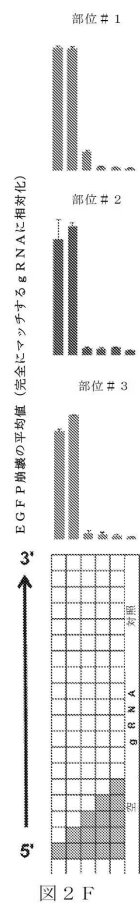


図 2 F

【図 2 G】

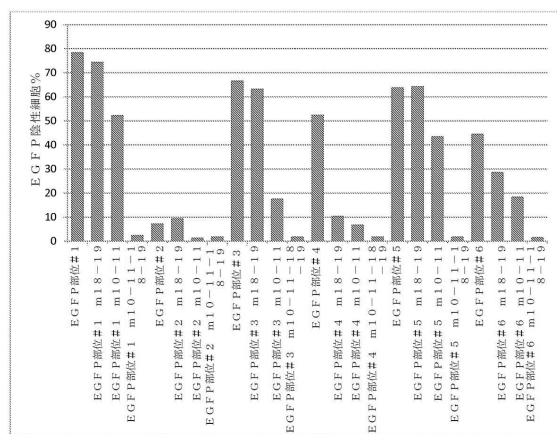


図 2 G

【 図 2 H 】

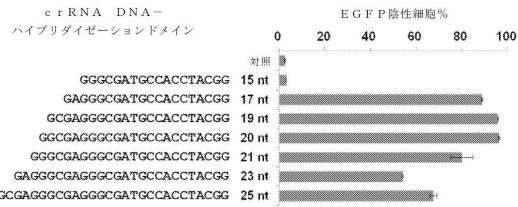


図 2 H

【 図 3 A 】

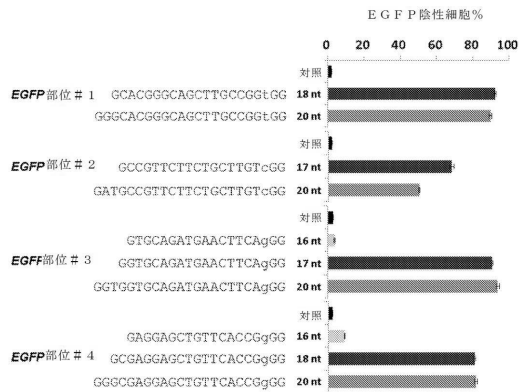


図 3 A

【 図 3 C 】

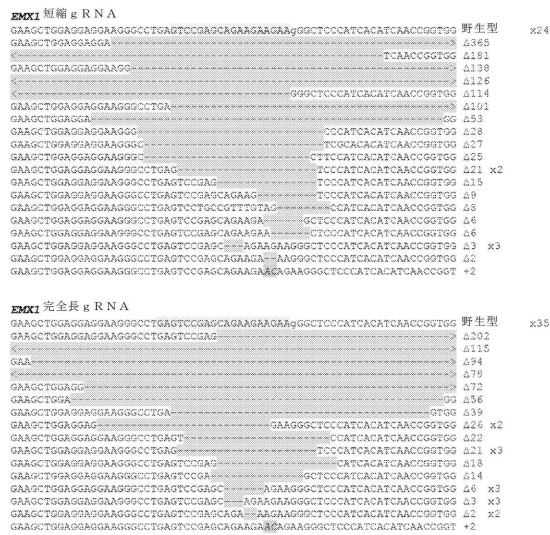


図 3 C

【 ㄨ 3 B 】

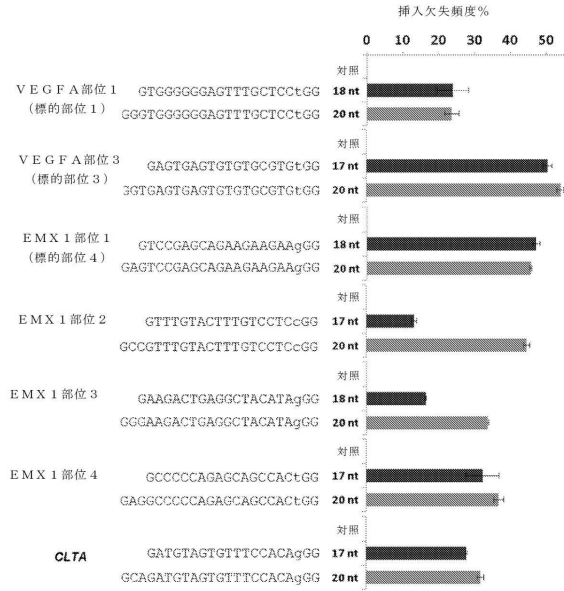


図 3 B

【 図 3 D 】

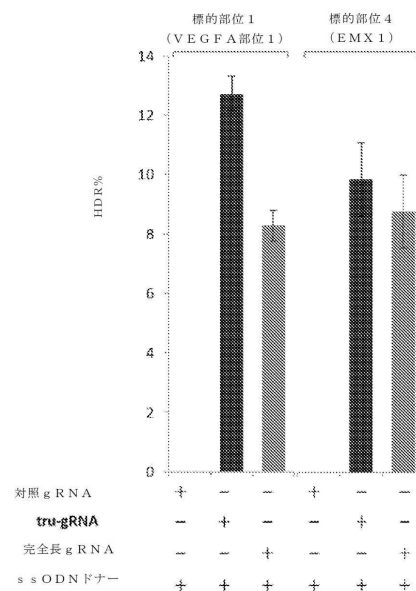


图 3 D

【図 3 E】

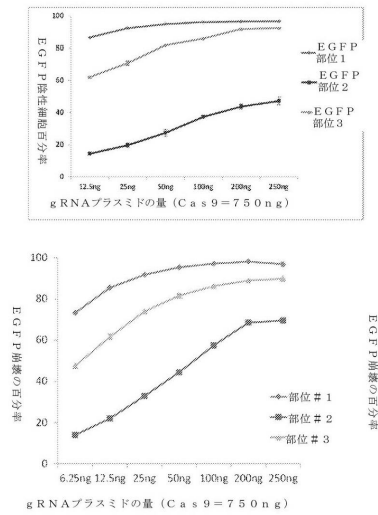


図 3 E

【図 3 F】

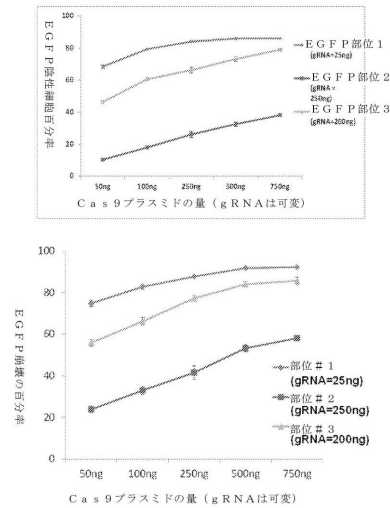


図 3 F

【図 4 A】

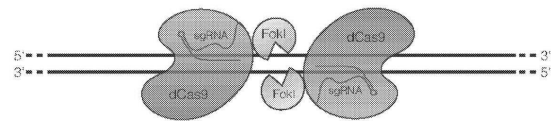


図 4 A

【図 4 B】

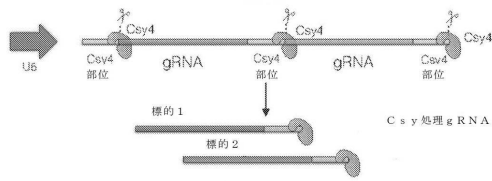


図 4 B

【図 4 C】

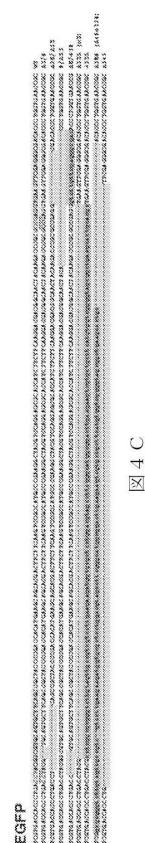


図 4 C

【図 5 A】

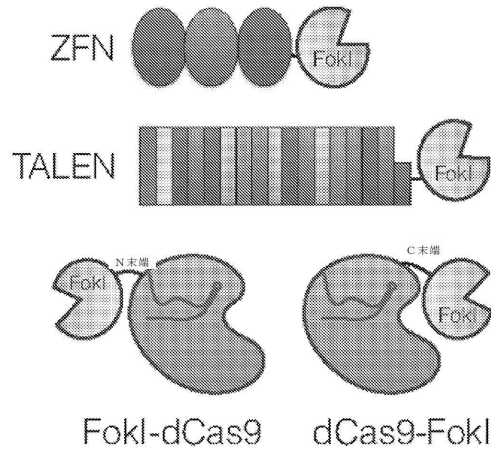


図 5 A

【図 5 B】

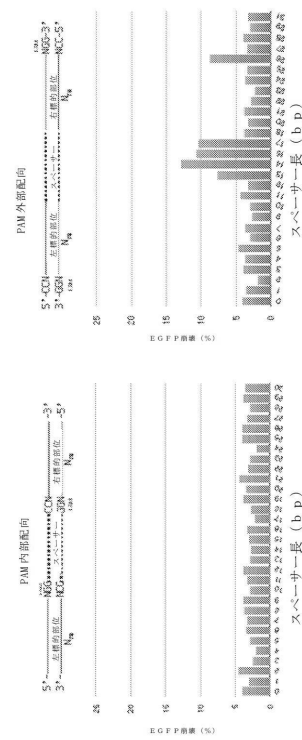


図 5 B

【図 5 C】

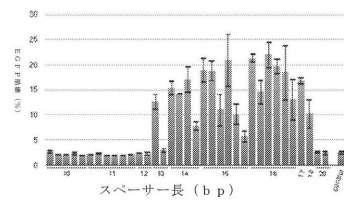


図 5 C

【図 5 D】

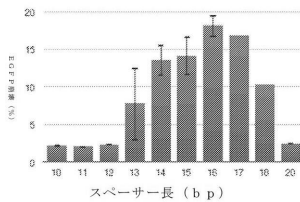


図 5 D

【図 5 E】

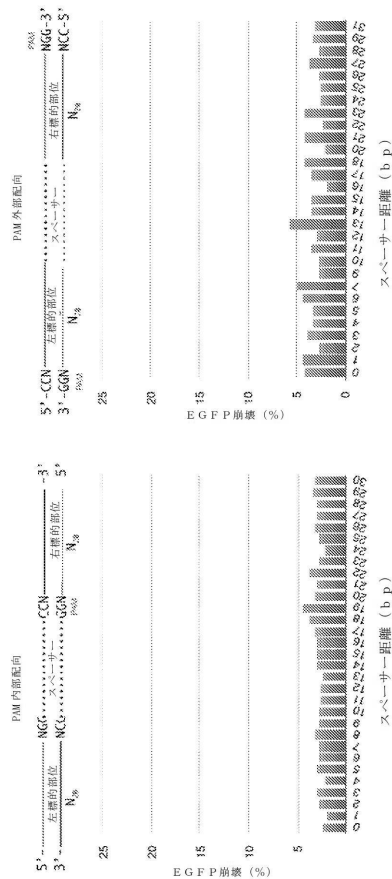
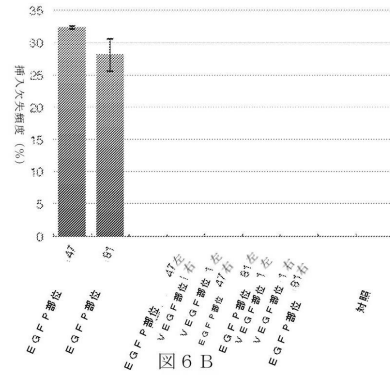


図 5 E

【 図 6 B 】



【 図 6 C 】

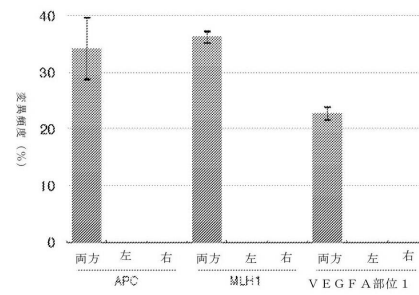
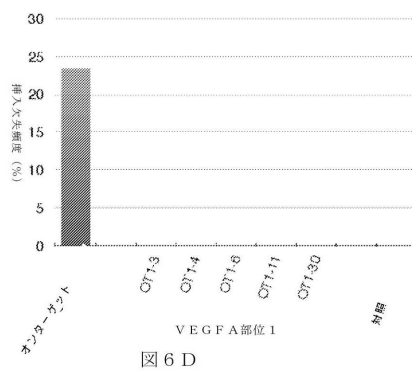
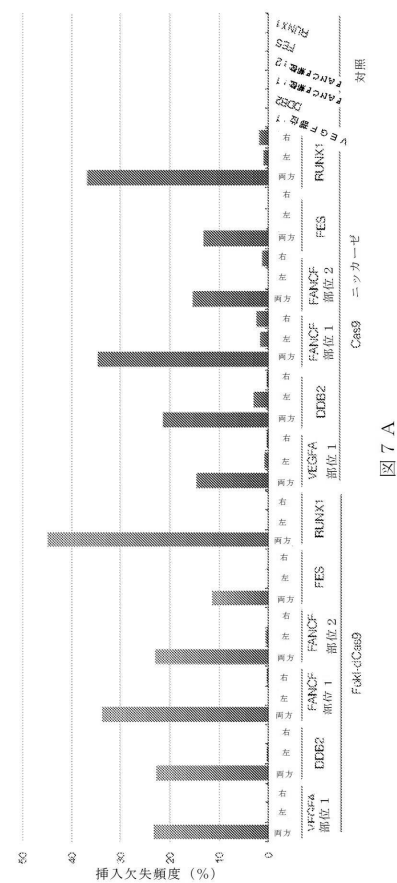


图 6 C

【 図 6 D 】



【 図 7 A 】



フロントページの続き

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

(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, シェンダー

アメリカ合衆国 0 2 1 2 9 マサチューセッツ州, チャールズタウン, プロスペクト ストリート 3 2, アpartment 1

審査官 西村 亜希子

(56)参考文献 国際公開第2 0 1 4 / 0 8 9 2 9 0 (W O , A 1)
 国際公開第2 0 1 4 / 0 9 9 7 4 4 (W O , A 1)
 米国特許出願公開第2 0 1 0 / 0 0 5 5 7 9 3 (U S , A 1)
 Cell , 2 0 1 3 年 2 月 2 8 日 , Vol.152 , pp.1173-1183
 Science , 2 0 1 2 年 8 月 1 7 日 , Vol.337 , pp.816-821
 PNAS , 1 9 9 6 年 , Vol.93 , pp.1156-1160 , FIG.1
 J. Biotechnol. , 2 0 0 9 年 , Vol.140 , pp.156-161 , Fig.1

(58)調査した分野(Int.Cl. , D B 名)

C 1 2 N 9 / 1 6
 C 0 7 K 1 9 / 0 0
 C 1 2 N 1 / 1 5
 C 1 2 N 1 / 1 9
 C 1 2 N 1 / 2 1
 C 1 2 N 5 / 1 0
 C 1 2 N 1 5 / 0 9
 J S T P l u s / J S T 7 5 8 0 (J D r e a m I I I)
 C A p l u s / M E D L I N E / W P I D S / B I O S I S (S T N)
 G e n B a n k / E M B L / D D B J / G e n e S e q
 U n i P r o t / G e n e S e q
 P u b M e d