

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第6984096号
(P6984096)

(45) 発行日 令和3年12月17日(2021.12.17)

(24) 登録日 令和3年11月29日(2021.11.29)

(51) Int. Cl. F I
C 1 2 N 9/16 (2006.01) C 1 2 N 9/16 Z N A Z
 C 1 2 N 15/55 (2006.01) C 1 2 N 15/55

請求項の数 11 (全 34 頁)

(21) 出願番号	特願2016-532542 (P2016-532542)	(73) 特許権者	507044516
(86) (22) 出願日	平成26年11月19日 (2014.11.19)		プレジデント アンド フェローズ オブ ハーバード カレッジ
(65) 公表番号	特表2017-501688 (P2017-501688A)		アメリカ合衆国 マサチューセッツ O2 138, ケンブリッジ, クインシー ストリート 17
(43) 公表日	平成29年1月19日 (2017.1.19)	(74) 代理人	100079049
(86) 国際出願番号	PCT/US2014/066375		弁理士 中島 淳
(87) 国際公開番号	W02015/077318	(74) 代理人	100084995
(87) 国際公開日	平成27年5月28日 (2015.5.28)		弁理士 加藤 和詳
審査請求日	平成29年11月17日 (2017.11.17)	(72) 発明者	チャベス、アレハンドロ
審査番号	不服2020-925 (P2020-925/J1)		アメリカ合衆国 O2446 マサチュー セッツ州 ブルックリン スーワル アベ ニュー 108 ユニット 4
審査請求日	令和2年1月23日 (2020.1.23)		
(31) 優先権主張番号	61/906,374		
(32) 優先日	平成25年11月19日 (2013.11.19)		
(33) 優先権主張国・地域又は機関	米国 (US)		

最終頁に続く

(54) 【発明の名称】 変異C a s 9タンパク質

(57) 【特許請求の範囲】

【請求項1】

C a s 9タンパク質のファミリー内の低度保存された1または複数のアミノ酸配列を同定すること、

前記1または複数の同定されたアミノ酸配列に対応する核酸配列を同定すること、

前記1または複数の同定されたアミノ酸配列に対応する前記核酸配列を欠いた、変異C a s 9タンパク質をコードする核酸を作製すること、および

前記作製された核酸から変異C a s 9タンパク質を作製することを含み、

前記変異C a s 9タンパク質は、R N A誘導型D N A結合活性を依然として有している

10

変異C a s 9タンパク質を作製する方法。

【請求項2】

前記変異C a s 9タンパク質がヌクレアーゼ活性を有する、請求項1に記載の方法。

【請求項3】

前記変異C a s 9タンパク質がニッカーゼである、請求項1に記載の方法。

【請求項4】

前記変異C a s 9タンパク質がヌクレアーゼ欠損である、請求項1に記載の方法。

【請求項5】

前記1または複数の同定されたアミノ酸配列に対応する前記核酸配列からなる核酸を、

20

前記変異 C a s 9 タンパク質をコードする核酸から欠失させる、請求項 1 に記載の方法。

【請求項 6】

前記 1 または複数の同定されたアミノ酸配列に対応する前記核酸配列からなる核酸を、前記変異 C a s 9 タンパク質をコードする核酸から欠失させてリンカーで置換する、請求項 1 に記載の方法。

【請求項 7】

前記 1 または複数の同定されたアミノ酸配列に対応する前記核酸配列からなる核酸を、前記変異 C a s 9 タンパク質をコードする核酸から欠失させて S G G G S リンカーで置換する、請求項 1 に記載の方法。

【請求項 8】

前記 1 または複数の同定されたアミノ酸配列が、前記 1 または複数の同定されたアミノ酸配列を隣接アミノ酸配列から区別する一対のエッジアミノ酸配列を有する、請求項 1 に記載の方法。

【請求項 9】

C a s 9 タンパク質のファミリー内の、高度保存されたか、または低度保存された 1 または複数のアミノ酸配列を同定することをさらに含み、前記高度保存されたか、または低度保存された 1 または複数のアミノ酸配列のそれぞれには、低度保存された 2 つの隣接アミノ酸配列か、または高度保存された 2 つの隣接アミノ酸配列が隣接している、請求項 1 に記載の方法。

【請求項 10】

配列番号 1 の核酸配列によってコードされる変異 C a s 9 タンパク質。

【請求項 11】

配列番号 2 の核酸配列によってコードされる変異 C a s 9 タンパク質。

【発明の詳細な説明】

【技術分野】

【0001】

関連出願

本願は、2013年11月19日に提出された米国仮特許出願第61/906,374号に基づく優先権を主張するものであり、あらゆる目的のために、その全体が参照により本明細書に取り込まれる。

【0002】

政府利益の陳述

本発明は、米国国立衛生研究所助成番号第P50HG005550号の政府助成を受けてなされた。米国政府は、本発明における一定の権利を有する。

【背景技術】

【0003】

細菌および古細菌のCRISPR-Casシステムは、Casタンパク質と複合体を形成した短鎖ガイドRNAに依存して、侵入する外来核酸中に存在する相補配列を分解する。Deltcheva, E. et al. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* 471, 602-607 (2011); Gasiunas, G., Barrangou, R., Horvath, P. & Siksnys, V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proceedings of the National Academy of Sciences of the United States of America* 109, E2579-2586 (2012); Jinek, M. et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816-821 (2012); Sapranaukas, R. et al. The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. *Nucleic acids research* 39, 9275-9282 (2011); および Bhaya, D., Davison, M. & Barrangou, R. CRISPR-Cas systems in bacteria and archaea: versatile small RNAs for adaptive defense and regulation. *Annual review of genetics* 45, 273-297 (2011) を参照されたい。近年、*Streptococcus pyogenes* (S. pyogenes) のII型CRISPRシステム

10

20

30

40

50

のインビトロ再構成により、通常はトランスにコードされる *tracrRNA* (「トランス活性化型 *CRISPR RNA*」) と融合した *crRNA* (「*CRISPR RNA*」) が、その *crRNA* に一致する標的 *DNA* 配列を *Cas9* タンパク質に配列特異的に切断させるのに十分であることが実証された。標的部位に相同な *gRNA* の発現により、*Cas9* が動員され、標的 *DNA* が分解される。H. Deveau et al, Phage response to *CRISPR*-encoded resistance in *Streptococcus thermophilus*. *Journal of Bacteriology* 190, 1390 (Feb, 2008) を参照されたい。

【発明の概要】

【発明が解決しようとする課題】

【0004】

Cas9 は、標的化が望まれるゲノム領域に相補的な 20 塩基対と *Cas9* を動員するモチーフとを含むガイド *RNA* (*gRNA*) を発現させることにより、ゲノムのほぼ全ての領域を標的化するようにプログラムし得る *DNA* ヌクレアーゼである。特徴解析されている *Cas9* ファミリーのメンバーおよび推定上の *Cas9* ファミリーのメンバーの全ては、サイズが数千ベース (> 3000 塩基対) であり、機能的に検証されている最小のメンバーである *NM-Cas9* (ナイセリア・メニンジティディス (*Neisseria meningitidis*) の *Cas9*) のサイズは 3,249 塩基対である。このタンパク質はサイズが大きいため、送達および操作が難しく、生命工学および治療的応用への可能性が制限される。

【課題を解決するための手段】

【0005】

本開示の態様は、タンパク質の一部が削除 (omit) されるように改変されているが、標的 *DNA* に結合して標的 *DNA* 中に二本鎖切断を形成することができる *RNA* 誘導型 *DNA* 結合ヌクレアーゼとして機能する、*DNA* に結合し且つ 1 種類または複数種類の *RNA* によってガイドされる *II* 型 *CRISPR* システムの *RNA* 誘導型 *DNA* 結合タンパク質に関する。ある態様によれば、*II* 型 *CRISPR* システムの *RNA* 誘導型 *DNA* 結合タンパク質は *Cas9* タンパク質である。

【0006】

本開示の態様は、タンパク質の一部が削除されるように改変されているが、標的 *DNA* に結合して標的 *DNA* 中に一本鎖切断またはニックを形成することができる *RNA* 誘導型 *DNA* 結合ニッカーゼとして機能する、*II* 型 *CRISPR* システムの *RNA* 誘導型 *DNA* 結合タンパク質に関する。ある態様によれば、*II* 型 *CRISPR* システムの *RNA* 誘導型 *DNA* 結合タンパク質は *Cas9* タンパク質である。

【0007】

本開示の態様は、タンパク質の一部が削除されるように改変されているが、ヌクレアーゼ欠損 (nuclease null) である、すなわち *RNA* 誘導型 *DNA* 結合タンパク質がヌクレアーゼ活性を欠いている、*RNA* 誘導型 *DNA* 結合タンパク質として機能する、*II* 型 *CRISPR* システムの *RNA* 誘導型 *DNA* 結合タンパク質に関する。ある態様によれば、*II* 型 *CRISPR* システムの *RNA* 誘導型 *DNA* 結合タンパク質は *Cas9* タンパク質である。

【0008】

ある態様によれば、特定の *RNA* 誘導型 *DNA* 結合タンパク質ファミリー内であまり保存されていないかまたは高度に分岐した *RNA* 誘導型 *DNA* 結合タンパク質配列、および/または特定の *RNA* 誘導型 *DNA* 結合タンパク質ファミリー内の、本明細書中で「保存エッジ (conservation edge)」と呼ばれる、低度保存と高度保存との境界の間のタンパク質配列を、種の集団内で同定することによって、*RNA* 誘導型 *DNA* 結合タンパク質の部分を欠失のために同定する。この態様によれば、本明細書に記載され、当業者に知られているような方法を用いて、*Cas9* などの *RNA* 誘導型 *DNA* 結合タンパク質などの *DNA* 結合タンパク質内のアミノ酸配列が、高度保存されているか、または低度保存されていると特定される。ある態様によれば、高度保存されたアミノ酸配列および低度保存されたアミノ酸配列は、全体として、*DNA* 結合タンパク質のタンパク質配列内で互いに隣接

10

20

30

40

50

、例えば直接隣接している。高度保存されたアミノ酸配列および低度保存されたアミノ酸配列は、高度保存されたアミノ酸配列を低度保存されたアミノ酸配列から切り離すアミノ酸によって区別される。このように、高度保存されたアミノ酸配列を低度保存されたアミノ酸配列から切り離すアミノ酸は、高度保存されたアミノ酸配列または低度保存されたアミノ酸配列のいずれかのエッジまたは末端部分である限りは、本明細書では「エッジアミノ酸」または「保存エッジ」と呼ばれる。したがって、本開示の方法は、高度保存されたアミノ酸配列を低度保存されたアミノ酸配列と切り離す、あるいは高度保存されたアミノ酸配列と低度保存されたアミノ酸配列を区別する、アミノ酸の同定を検討する。そのようなアミノ酸を本明細書では「エッジアミノ酸」と呼ぶ。この態様によれば、一对のエッジアミノ酸が、高度保存されたアミノ酸配列のいずれかの端部に隣接または接していてもよい。同様に、一对のエッジアミノ酸が、低度保存されたアミノ酸配列のいずれかの端部に隣接または接していてもよい。さらにこの態様によれば、ある例示的实施形態は、DNA結合タンパク質全体のタンパク質配列内での、高度保存された隣接アミノ酸配列の同定、および低度保存されたアミノ酸配列の同定に関する。特に、ある例示的实施形態は、DNA結合タンパク質全体のタンパク質配列内での、低度保存された配列とタンデムまたは直列に存在する高度保存された配列の同定、特に、いずれかの末端が低度保存(LC)された配列に隣接している高度保存(HC)配列、またはいずれかの末端が高度保存(HC)された配列に隣接している低度保存(LC)配列の同定に関する。このように、本明細書に記載の方法によって同定される例示的なタンデム配列または直列配列は、模式的に LC HC LC または HC LC HC と表され得る。この態様によれば、高度保存または低度保存された中央配列には、それぞれ低度保存または高度保存された隣接配列が接している。例示的なタンデム配列 LC HC LC では、一对のエッジアミノ酸により、高度保存(HC)されたアミノ酸配列が、高濃度アミノ酸配列の両端にあるかまたは高濃度アミノ酸配列に隣接している2つの隣接する低度保存(LC)されたアミノ酸配列から、区別されるかまたは切り離されている。例示的なタンデム配列 HC LC HC では、一对のエッジアミノ酸により、低度保存(LC)されたアミノ酸配列が、低度保存されたアミノ酸配列の両端にあるかまたは低度保存されたアミノ酸配列に隣接している2つの隣接する高度保存(HC)されたアミノ酸配列から、区別されるかまたは切り離されている。中央配列が高度保存されたアミノ酸配列または低度保存されたアミノ酸配列のいずれであっても、本明細書に記載の方法を用いてこのような例示的なタンデム配列が同定された場合、本明細書に記載の方法に従って中央配列を除去することにより、DNA結合活性が維持されており、且つ野生型DNA結合タンパク質よりサイズが小さい、変異DNA結合タンパク質を作成する。エッジアミノ酸は、中央配列に隣接することにより、または直列する隣接配列から中央配列を切り離すことにより、変異DNA結合タンパク質を作成するために除去される中央配列を規定する。本開示の特定の態様によれば、変異DNA結合タンパク質が有用なDNA結合タンパク質活性を保持している限り、中央配列が高度保存されたアミノ酸配列であるか低度保存されたアミノ酸配列であるかに関係なく、タンデム配列中の中央配列が除去され得る。

【0009】

本開示におけるあまり保存されていない配列、または保存エッジ間のタンパク質配列を同定するために、アラインメントを、PFAMから得るか、またはCas9ホモログのデータベースサーチにより得られた配列のコレクションから作成した。このアラインメントを計算的に調整し、位置ごとのアミノ酸頻度の(相対)エントロピーとして保存度を計算した。次に、RNA誘導型DNA結合タンパク質から配列を除去して変異体を作製する。

【0010】

ある態様によれば、欠失に適していると考えられるマルチドメインCas9タンパク質内の領域を同定するために、バイオインフォマティクスによるアプローチを用いて、Cas9タンパク質における潜在的なドメイン境界を同定する。MUSCLEを用いてPFAMデータベース中のCas9配列(PF13395)をリアライメントすることにより多重配列アラインメントを作成し、多様性および全長配列についてアラインメントを計算的

10

20

30

40

50

に調整する。大腸菌の全遺伝子の平均頻度を基準として、観察されたアミノ酸頻度の相対エントロピーとして配列保存度を計算する。保存プロファイルにマルチスケールエッジフィルタ（ガウス差分（DoG）バンドパスフィルタ）を適用して、本明細書で保存エッジと呼ばれる潜在的なタンパク質ドメイン境界を帰属させる。保存エッジ間の領域を1回目の欠失変異体の欠失のために選択する。

【0011】

ある態様によれば、本開示は、サイズがより小さいが、野生型のタンパク質に近い活性を保持している、合成NM Cas9欠失変異体を記載する。合成NM Cas9欠失変異体は、細胞中でガイドRNAとの共局在複合体としてDNAに結合させて二本鎖切断もしくは一本鎖切断を生じさせるため、または目的の標的DNAの近傍にエフェクター基を配置させて所望の機能を発揮させるために用いることができる。

10

【0012】

ある態様によれば、DNAへの結合に不必要であり、野生型DNA結合タンパク質よりサイズの小さい変異DNA結合タンパク質を形成するために除去可能である、Cas9などのDNA結合タンパク質の領域を同定するための、アラインメントに基づくドメイン検出方法が提供される。本明細書に記載の方法によれば、細菌細胞およびヒト細胞において強力な活性を示す最小化されたCas9バリエーションが作製される。本明細書に記載の態様によれば、野生型DNA結合タンパク質より小さい、変異機能性Cas9バリエーションなどの変異機能性DNA結合タンパク質バリエーションが提供される。

20

【0013】

ある態様によれば、例示的なDNA結合タンパク質としては、原核生物および高等真核生物の両方で機能することが示されている、ナイセリア・メニンジティディス (*Neisseria meningitidis*) のCas9 (NM、GI: 218767588) およびストレプトコッカス・サーモフィラス (*Streptococcus thermophilus*) のCas9 (ST1、GI: 116627542) などのCas9オルソログが挙げられる。その全体が参照により援用されるHou, Z. et al. Efficient genome engineering in human pluripotent stem cells using Cas9 from *Neisseria meningitidis*. *Proceedings of the National Academy of Sciences of the United States of America* 110, 15644-15649, doi:10.1073/pnas.1313587110 (2013) を参照されたい。

30

【0014】

これらの例示的Cas9オルソログは、ストレプトコッカス・ピオゲネス (*Streptococcus pyogenes*) のCas9 (SP、GI: 13622193) より遺伝子サイズが小さく、すなわち、4100塩基対に対して約3200塩基対である。したがって、本開示の態様は、特に遺伝子長がウイルス価に大きく影響し得るウイルスパッケージング技術を用いて、DNA結合タンパク質Cas9を送達可能な効率が上昇するように、DNA結合タンパク質Cas9のサイズを小さくすることに関する。それぞれの全体が参照により援用されるKumar, M., Keller, B., Makalou, N. & Sutton, R. E. Systematic determination of the packaging limit of lentiviral vectors. *Human gene therapy* 12, 1893-1905, doi:10.1089/104303401753153947 (2001); Wu, Z., Yang, H. & Colosi, P. Effect of genome size on AAV vector packaging. *Molecular therapy: the journal of the American Society of Gene Therapy* 18, 80-86, doi:10.1038/mt.2009.255 (2010); およびGelin as, C. & Temin, H. M. Nondefective spleen necrosis virus-derived vectors define the upper size limit for packaging reticuloendotheliosis viruses. *Proceedings of the National Academy of Sciences of the United States of America* 83, 9211-9215 (1986) を参照されたい。

40

【0015】

Cas9遺伝子のサイズを合成的に小さくすることで、より複雑な制御系および機能的ドメインを単一ベクター内にパッケージングすることができる。別の態様によれば、標的化能が向上したより小さなCas9バリエーションの作製を可能にする、PAM特異性を合成的に変化させる方法が提供される。

50

【0016】

ある態様によれば、第1の種類のCas9のC末端ドメインを第2の種類のCas9のC末端ドメインと交換することによりCas9キメラを作製する方法が提供される。ある態様によれば、本開示は、NMのC末端ドメインをST1と交換することにより、機能的NM-ST1-Cas9キメラなどのドメイン交換Cas9キメラを提供する。キメラCas9タンパク質はST1のガイドRNAおよびPAMに対する特異性を示す。

【0017】

ある態様によれば、細胞は、原核細胞または真核細胞である。ある態様によれば、細胞は、細菌細胞、酵母細胞、植物細胞、または動物細胞である。ある態様によれば、細胞は、哺乳動物細胞である。

10

【0018】

ある態様によれば、RNAは約10~約500ヌクレオチドである。ある態様によれば、RNAは約20~約100ヌクレオチドである。

【0019】

ある態様によれば、1種類または複数種類のRNAはガイドRNAである。ある態様によれば、1種類または複数種類のRNAはcrRNAである。ある態様によれば、1種類または複数種類のRNAはtracrRNAである。ある態様によれば、1種類または複数種類のRNAはtracrRNA-crRNA融合体である。

【0020】

ある態様によれば、標的DNAは、ゲノムDNA、ミトコンドリアDNA、ウイルスDNA、結合可能なエレメント、または外来性DNAである。

20

【0021】

ある態様によれば、RNA誘導型DNA結合タンパク質は、DNAに結合し、且つ1種類または複数種類のRNAによってガイドされる、I型CRISPRシステムのRNA誘導型DNA結合タンパク質である。ある態様によれば、RNA誘導型DNA結合タンパク質は、DNAに結合し、且つ1種類または複数種類のRNAによってガイドされる、Cas9タンパク質である。

【0022】

本発明の特定の実施形態のその他の特徴および利点は、以下の実施形態およびそれらの図面の説明において、および特許請求の範囲から、さらに十分に明らかになるであろう。

30

【0023】

本特許または出願ファイルは、カラーで作製された少なくとも1枚の図面を含む。カラー図面を含む本特許または特許出願公開のコピーは、特許庁へ申請し、必要な料金を支払うことで提供される。本発明の、前述した特徴およびその他の特徴、ならびに利点は、添付の図面と共に、具体的実施形態についての以下の詳細な説明からより十分に理解されるであろう。

【図面の簡単な説明】

【0024】

【図1】NM-Cas9活性に対するYFPレポーターを含み、種々のNM-Cas9ヌクレアーゼ欠損遺伝子で形質転換された大腸菌細胞の画像を示す図である。NM-Cas9の非存在下では細胞は蛍光性であり(右上の陰性対照)、全長ヌクレアーゼ欠損NM-Cas9存在下では細胞は非蛍光性である(左上の陽性対照)。作製したNM-Cas9欠失体のうちの2つが示されており、NM-Cas9²⁵⁵⁻⁴⁴⁹は、野生型レベルに近い抑制を示し(左下)、NM-Cas9⁸⁷⁴⁻⁹²²は、DNA結合能の大部分の欠如を示している(右下)。

40

【図2】その全体が参照により援用されるFonfara I, et al., (2014) Phylogeny of Cas9 determines functional exchangeability of dual-RNA and Cas9 among orthologous type II CRISPR-Cas systems, Nucleic Acids Res. 42, 2577-90に記載されるような系統樹を示す図である。赤色で印が付けられているのが、PFA Mリアライメントの初期シードとして用いた配列である。

50

【図3】NM Cas9の位置まで短縮 (truncate) した後のCas9アラインメントの保存プロファイルを示す図である。

【図4】SP Cas9における位置まで短縮 (truncate) した保存プロファイルを示す図である。

【図5A】Cas9タンパク質内の一次アミノ酸保存プロットを示す図である。大腸菌の全遺伝子の平均アミノ酸頻度に対する相対エントロピーが計算されている。プロットの上の縦線は、アラインメントに基づく境界検出アルゴリズムにより決定された境界を表し、太線は検出された最も有意な6つの境界を表す。

【図5B】その全体が参照により援用されるFonfara, I. et al. Phylogeny of Cas9 determines functional exchangeability of dual-RNA and Cas9 among orthologous type I CRISPR-Cas systems. *Nucleic acids research* 42, 2577-2590, doi:10.1093/nar/gkt1074 (2014) に基づくドメイン帰属を示す図である。RuvCI~IIIは、RuvCヌクレアーゼドメインへと折り畳まれる部分である。HNHはHNHヌクレアーゼドメインである。RRRはアルギニンリッチヘリックス領域である。網掛けは、140位周辺のアルギニンリッチな区間に基づくこの領域の伸長部である。

【図6A】NM Cas9転写レポーターのデザインを示す模式図である。プロトSpacerおよびNM特異的PAMの位置を示す。

【図6B】大腸菌において試験した種々のヌクレアーゼ欠損NM変異体を用いた、NM Cas9転写抑制アッセイデータのグラフを示す図である。データは平均値 ± 標準偏差を表す (n = 4)。

【図6C】NM Cas9のドメイン構造を示す模式図である。輪郭が破線である白い四角は、DNA結合活性にほとんど変化を起こさないNM変異体からの最も大きな切出し領域の範囲を示す。

【図7】NM Cas9欠失解析に関し、NM Cas9転写レポーターのデザインを示す図である。プロトSpacerおよびNM特異的PAMの位置を示す。種々のヌクレアーゼ欠損NM変異体を用いたNM Cas9転写抑制アッセイを大腸菌において試験した。データは平均値 ± 標準偏差を表す (n = 5)。

【図8A】大腸菌およびヒト細胞におけるST1 Cas9欠失解析および機能検証ならびに特にST1 Cas9転写レポーターのデザインを示す図である。プロトSpacerおよびST1特異的PAMの位置を示す。大腸菌において試験したST1ヌクレアーゼ欠損の欠失変異体を用いた、ST1 Cas9転写抑制因子アッセイ。データは平均値 ± 標準偏差を表す (n = 4)。

【図8B】tdTomatoレポーターの上流に最小CMVプロモーター (min CMV) を含む、ST1活性化試験用のレポーターコンストラクトを示す模式図である。最小CMVプロモーターの上流へのST1ヌクレアーゼ欠損Cas9 VP64融合タンパク質の結合により、ヒト細胞内での転写活性化および蛍光が引き起こされる。

【図8C】欠失変異体を含むST1活性化因子をsgRNAおよびtdTomatoレポーターと共にトランスフェクトして蛍光顕微鏡で可視化した細胞の画像を示す図である。

【図8D】フローサイトメトリーによる図8CのST1活性化の定量化を示すグラフである。データは平均値 ± 標準偏差を表す (n = 3)。

【図9】大腸菌におけるTD Cas9欠失解析および機能検証に関する図である。転写抑制因子アッセイを用いてTDヌクレアーゼ欠損欠失変異体を試験した。データは平均値 ± 標準偏差を表す (n = 4)。

【図10A】転写抑制アッセイにより測定されるNM ST1ドメイン交換解析、および特に図示されているNM特異的PAMまたはST1特異的PAMの配列を有するNM転写レポーターおよびST1転写レポーターのデザインに関する図である。

【図10B】NMおよびST1のCas9の概略をアミノ酸交換ポイントの位置と共に示す模式図である。

【図10C】NMのCas9に特異的なPAM配列を有するレポーターと共に、NMのCas9に特有なガイドRNAと同時発現されたNM ST1ヌクレアーゼ欠損ドメイン交

10

20

30

40

50

換変異体の蛍光のグラフを示す図である。データは平均値 ± 標準偏差を表す (n = 4)。

【図10D】ST1のCas9に特異的なPAM配列を有するレポーターと共に、NMのCas9に特異的なガイドRNAと同時発現されたNM-ST1ヌクレアーゼ欠損ドメイン交換変異体の蛍光のグラフを示す図である。データは平均値 ± 標準偏差を表す (n = 4)。

【図10E】ST1のCas9に特異的なPAM配列を有するレポーターと共に、ST1のCas9に特異的なガイドRNAと同時発現されたNM-ST1ヌクレアーゼ欠損ドメイン交換変異体の蛍光のグラフを示す図である。データは平均値 ± 標準偏差を表す (n = 4)。

【図10F】NMのCas9に特異的なPAM配列を有するレポーターと共に、ST1のCas9に特異的なガイドRNAと同時発現されたNM-ST1ヌクレアーゼ欠損ドメイン交換変異体の蛍光のグラフを示す図である。データは平均値 ± 標準偏差を表す (n = 4)。

【図11A】転写抑制アッセイにより測定されるNM-ST1ドメイン交換解析、および特に図示されているST1特異的PAMの配列を有するST1転写レポーターのデザインに関する図である。

【図11B】NMおよびST1-Cas9の概略を、アミノ酸交換ポイントの位置と共に示す模式図である。

【図11C】ST1のCas9に特異的なPAM配列を有するレポーターと共に、ST1のCas9に特異的なガイドRNAと同時発現されたNM-ST1ヌクレアーゼ欠損ドメイン交換変異体の蛍光のグラフを示す図である。データは平均値 ± 標準偏差を表す (n = 4)。

【図11D】ST1のCas9に特異的なPAM配列を有するレポーターと共に、NMのCas9に特異的なガイドRNAと同時発現されたNM-ST1ヌクレアーゼ欠損ドメイン交換変異体の蛍光のグラフを示す図である。データは平均値 ± 標準偏差を表す (n = 4)。

【発明を実施するための形態】

【0025】

本発明の実施形態は、II型CRISPRシステムの変異RNA誘導型DNA結合タンパク質に関する。そのような変異体は、II型CRISPRシステムのRNA誘導型DNA結合タンパク質属内の種間であまり保存されていないか、または高度に分岐した (divergent) 配列を除去することにより作成される。ある態様によれば、RNA誘導型DNA結合タンパク質のファミリー内の種の配列をアライメントし、低度保存配列または保存エッジ間配列を決定する。次に、これらの配列を特定のRNA誘導型DNA結合タンパク質から欠失させる。例示的なRNA誘導型DNA結合タンパク質としては、例えばII型CRISPRシステム中に存在するCas9タンパク質が挙げられる。そのようなCas9タンパク質およびII型CRISPRシステムは、当該技術分野で十分に裏付けられている。Makarova et al., Nature Reviews, Microbiology, Vol. 9, June 2011, pp. 467-477を参照されたい。本明細書に記載の変異DNA結合タンパク質は、標的DNAにおける二本鎖切断もしくは標的DNAにおける一本鎖切断を生じさせるため、または標的DNAの近くにエフェクター基が配置されてエフェクター基が標的DNAと相互作用できるように標的DNAに結合するために用いることができる。そのようなエフェクター基としては、当業者に公知の活性化因子、抑制因子、またはエピジェネティック修飾因子が挙げられる。

【0026】

ヌクレアーゼ活性を有する例示的なDNA結合タンパク質は、二本鎖DNAにニックを形成するか、または二本鎖DNAを切断するように機能する。そのようなヌクレアーゼ活性は、ヌクレアーゼ活性を示す1つまたは複数のポリペプチド配列を有するDNA結合タンパク質から生じ得る。そのような例示的なDNA結合タンパク質は、それぞれが二本鎖DNAの特定のストランドを切断すること、またはニックを形成することに関与する、2

10

20

30

40

50

つの別個のヌクレアーゼドメインを有していてもよい。当業者に知られているヌクレアーゼ活性を有する例示的なポリペプチド配列としては、McrA HNHヌクレアーゼ関連ドメインおよびRuvC様ヌクレアーゼドメインが挙げられる。したがって、例示的なDNA結合タンパク質は、McrA HNHヌクレアーゼ関連ドメインおよびRuvC様ヌクレアーゼドメインのうちの1つまたは複数を含む天然タンパク質である。

【0027】

ある態様によれば、2つ以上のヌクレアーゼドメインを有するDNA結合タンパク質は、ヌクレアーゼドメインのうちの1つを除いた全てが不活性化されるように修飾または改変されてもよい。そのような修飾または改変されたDNA結合タンパク質は、DNA結合タンパク質が二本鎖DNAの一方のストランドのみを切断する、またはニックを形成するものである限り、DNA結合タンパク質ニッカーゼと呼ばれる。RNAによってDNAにガイドされる場合、DNA結合タンパク質ニッカーゼは、RNA誘導型DNA結合タンパク質ニッカーゼと呼ばれる。

10

【0028】

例示的なDNA結合タンパク質は、II型CRISPRシステムのRNA誘導型DNA結合タンパク質である。例示的なDNA結合タンパク質は、Cas9タンパク質である。

【0029】

ストレプトコッカス・ピオゲネス (*S. pyogenes*) では、Cas9は、タンパク質中の2つの触媒ドメイン、すなわちDNAの相補鎖を切断するHNHドメインおよび非相補鎖を切断するRuvC様ドメインが介在するプロセスによって、プロトスペーサー隣接モチーフ (PAM) の3bp上流に平滑末端二本鎖切断を形成する。その全体が参照により援用されるJinke et al., *Science* 337, 816-821 (2012) を参照されたい。Cas9タンパク質は、Makarova et al., *Nature Reviews, Microbiology*, Vol. 9, June 2011, pp. 467-477の補足情報で確認されている以下のものを含む、多数のII型CRISPRシステム中に存在することが知られている。メタノコッカス・マリパルディス (*Methanococcus maripaludis*) C7株; コリネバクテリウム・ジフテリアエ (*Corynebacterium diphtheriae*); コリネバクテリウム・エフィシエンス (*Corynebacterium efficiens*) YS 314株; コリネバクテリウム・グルタニカム (*Corynebacterium glutamicum*) ATCC 13032 Kitasato株; コリネバクテリウム・グルタニカム ATCC 13032 Bielefeld株; コリネバクテリウム・グルタニカム R株; コリネバクテリウム・クロッペンステッティイ (*Corynebacterium kroppenstedtii*) DSM 44385株; マイコバクテリウム・アブセサス (*Mycobacterium abscessus*) ATCC 19977株; ノカルディア・ファルシニカ (*Nocardia farcinica*) IFM 10152株; ロドコッカス・エリスロポリス (*Rhodococcus erythropolis*) PR 4株; ロドコッカス・ジョスティイ (*Rhodococcus jostii*) RHA 1株; ロドコッカス・オパカス (*Rhodococcus opacus*) B4 uid 36573株; アシドサーマス・セルロリティカス (*Acidothermus cellulolyticus*) 11B株; アルスロバクター・クロロフェノリカス (*Arthrobacter chlorophenolicus*) A6株; クリベラ・フラビダ (*Kribbella flavida*) DSM 17836 uid 43465株; サーマモノスポラ・カーバタ (*Thermomonospora curvata*) DSM 43183株; ビフィドバクテリウム・デンティウム (*Bifidobacterium dentium*) Bd 1株; ビフィドバクテリウム・ロングム (*Bifidobacterium longum*) DJO 10A株; スラッキア・ヘリオトリニレデューセンス (*Slackia heliotrinireducens*) DSM 20476株; パーセフォネラ・マリナ (*Persephonella marina*) EX H 1株; バクテロイデス・フラギリス (*Bacteroides fragilis*) NCTC 9434株; カプノサイトファガ・オクラセア (*Capnocytophaga ochracea*) DSM 7271株; フラボバクテリウム・サイクロフィルム (*Flavobacterium psychrophilum*) JIP 02 86株; アッカーマンシア・ムシニフィラ (*Akkermansia muciniphila*) ATCC BAA 835株; ロゼイフレクサス・キャステンホルツィイ (*Roseiflexus castenholzii*) DSM 13941株; ロゼイフレクサス (*Roseiflexus*) RS 1株; シネコシスティス (*Synechocystis*) PCC 6803株; エルシマイクロビウム・ミヌトゥム (*Elusimicrobium minutum*) Pei 191株; 未培養細菌 Te

20

30

40

50

r m i t e g r o u p 1 b a c t e r i u m p h y l o t y p e R s D 1 7
 ; フィブロバクター・サクシノゲネス (*Fibrobacter succinogenes*) S 8 5 株 ; パチルス
 ・セレウス (*Bacillus cereus*) A T C C 1 0 9 8 7 株 ; リステリア・イノキュア (*Liste*
ria innocua) ; ラクトバチルス・カゼイ (*Lactobacillus casei*) ; ラクトバチルス・ラ
 ムノーサス (*Lactobacillus rhamnosus*) G G 株 ; ラクトバチルス・サリバリウス (*Lacto*
bacillus salivarius) U C C 1 1 8 株 ; ストレプトコッカス・アガラクティアエ (*Strept*
ococcus agalactiae) A 9 0 9 株 ; ストレプトコッカス・アガラクティアエ N E M 3
 1 6 株 ; ストレプトコッカス・アガラクティアエ 2 6 0 3 株 ; ストレプトコッカス・デ
 イスガテクティアエ亜種エクイシミリス (*Streptococcus dysgalactiae equisimilis*) G
 G S 1 2 4 株 ; ストレプトコッカス・エクイ亜種ズーエピデミカス (*Streptococcus equi*
zooepidemicus) M G C S 1 0 5 6 5 株 ; ストレプトコッカス・ガロリティカス (*Strept*
ococcus gallolyticus) U C N 3 4 u i d 4 6 0 6 1 株 ; ストレプトコッカス・ゴルド
 ニイ (*Streptococcus gordonii*) チャリス (*Challis*) 株 C H 1 亜株 ; ストレプトコッカ
 ス・ミュータンス (*Streptococcus mutans*) N N 2 0 2 5 u i d 4 6 3 5 3 株 ; ストレ
 プトコッカス・ミュータンス ; ストレプトコッカス・ピオゲネス (*Streptococcus pyogen*
es) M 1 G A S 株 ; ストレプトコッカス・ピオゲネス M G A S 5 0 0 5 株 ; ストレプト
 コッカス・ピオゲネス M G A S 2 0 9 6 株 ; ストレプトコッカス・ピオゲネス M G A S
 9 4 2 9 株 ; ストレプトコッカス・ピオゲネス M G A S 1 0 2 7 0 株 ; ストレプトコッ
 カス・ピオゲネス M G A S 6 1 8 0 株 ; ストレプトコッカス・ピオゲネス M G A S 3
 1 5 株 ; ストレプトコッカス・ピオゲネス S S I 1 株 ; ストレプトコッカス・ピオゲ
 ネス M G A S 1 0 7 5 0 株 ; ストレプトコッカス・ピオゲネス N Z 1 3 1 株 ; ストレ
 プトコッカス・サーモフィラス (*Streptococcus thermophiles*) C N R Z 1 0 6 6 株 ; ス
 トレプトコッカス・サーモフィラス L M D 9 株 ; ストレプトコッカス・サーモフィラ
 ス L M G 1 8 3 1 1 株 ; クロストリジウム・ボツリヌム (*Clostridium botulinum*) A
 3 L o c h M a r e e 株 ; クロストリジウム・ボツリヌム B E k l u n d 1 7
 B 株 ; クロストリジウム・ボツリヌム B a 4 6 5 7 株 ; クロストリジウム・ボツリヌ
 ム F L a n g e l a n d 株 ; クロストリジウム・セルロリティカム (*Clostridium ce*
llulolyticum) H 1 0 株 ; フィネゴルディア・マグナ (*Finegoldia magna*) A T C C 2 9
 3 2 8 株 ; ユウバクテリウム・レクターレ (*Eubacterium rectale*) A T C C 3 3 6 5 6
 株 ; マイコプラズマ・ガリセプティカム (*Mycoplasma gallisepticum*) ; マイコプラズマ
 ・モービレ (*Mycoplasma mobile*) 1 6 3 K 株 ; マイコプラズマ・ペネトランス (*Mycopla*
sma penetrans) ; マイコプラズマ・シノビアエ (*Mycoplasma synoviae*) 5 3 株 ; ストレ
 プトバチルス・モニリフォルミス (*Streptobacillus moniliformis*) D S M 1 2 1 1 2 株
 ; ブラジリゾビウム (*Bradyrhizobium*) B T A i 1 株 ; ニトロバクター・ハンブルゲンシ
 ス (*Nitrobacter hamburgensis*) X 1 4 株 ; ロドシュードモナス・パルストリス (*Rhodop*
seudomonas palustris) B i s B 1 8 ; ロドシュードモナス・パルストリス B i s B 5
 株 ; パルビバクラム・ラバメンティボランス (*Parvibaculum lavamentivorans*) D S 1
 株 ; ディノロセオバクター・シバエ (*Dinoroseobacter shibae*) D F L 1 2 株 ; グルコン
 アセトバクター・ジアゾトロフィクス (*Gluconacetobacter diazotrophicus*) P a l 5
 F A P E R J 株 ; グルコンアセトバクター・ジアゾトロフィクス P a l 5 J G I
 株 ; アゾスピリillum (*Azospirillum*) B 5 1 0 u i d 4 6 0 8 5 株 ; ロドスピリillum・
 ルブラム (*Rhodospirillum rubrum*) A T C C 1 1 1 7 0 株 ; ディアフォロバクター (*Dia*
phorobacter) T P S Y u i d 2 9 9 7 5 株 ; フェルミネフォロバクター・エイセニア
 エ (*Verminephrobacter eiseniae*) E F 0 1 2 株 ; ナイセリア・メニンジティディス (*N*
eisseria meningitidis) 0 5 3 4 4 2 株 ; ナイセリア・メニンジティディス a l p h
 a 1 4 ; ナイセリア・メニンジティディス Z 2 4 9 1 株 ; デスルホビブリオ・サレキシ
 ゲンス (*Desulfovibrio salexigens*) D S M 2 6 3 8 株 ; キャンピロバクク
 ー・ジェジュニ亜種ドイレイ (*Campylobacter jejuni doylei*) 2 6 9 9 7 株 ; キャンピロバクク
 ー・ジェジュニ (*Campylobacter jejuni*) 8 1 1 1 6 株 ; キャンピロバクク
 ー・ラリ (*Campylobacter lari*) R M 2 1 0 0 株 ; ヘリコバクター

10

20

30

40

50

・ヘパティカス (*H Helicobacter hepaticus*) ; ウォリネラ・サクシノゲネス (*Wolinella succinogenes*) ; トルモナス・アウエンシス (*Tolomonas auensis*) D S M 9 1 8 7 株 ; シュードアルテロモナス・アトランティカ (*Pseudoalteromonas atlantica*) T 6 c 株 ; シューワネラ・ペアレアナ (*Shewanella pealeana*) A T C C 7 0 0 3 4 5 株 ; レジオネラ・ニューモフィラ (*Legionella pneumophila*) P a r i s 株 ; アクチノバチルス・サクシノゲネス (*Actinobacillus succinogenes*) 1 3 0 Z 株 ; パスツレラ・ムルトシダ (*Pasteurella multocida*) ; フランシセラ・ツラレンシス亜種ノビシダ (*Francisella tularensis novicida*) U 1 1 2 株 ; フランシセラ・ツラレンシス亜種ハラークティカ (*Francisella tularensis holarctica*) ; フランシセラ・ツラレンシス F S C 1 9 8 株 ; フランシセラ・ツラレンシス亜種ツラレンシス (*Francisella tularensis tularensis*) ; フランシセラ・ツラレンシス W Y 9 6 3 4 1 8 株 ; およびトレポネーマ・デンティコラ (*Treponema denticola*) A T C C 3 5 4 0 5 株。したがって、本開示の態様は、上記で特定される種のいずれか1つなど、I I 型 C R I S P R システム中に存在する C a s 9 タンパク質の変異体に関する。例示的な C a s 9 タンパク質は、ナイセリア・メニンジティディス (*Neisseria meningitidis*) 0 5 3 4 4 2 株、ナイセリア・メニンジティディス 1 4 株、ナイセリア・メニンジティディス Z 2 4 9 1 株などのナイセリア・メニンジティディスにおいて見出される C a s 9 タンパク質である。

【 0 0 3 0 】

本開示に係る細胞には、本明細書に記載されるように外来核酸の導入でき、発現させることができるあらゆる細胞が含まれる。本明細書に記載される本開示の基本的概念は細胞の種類によって限定されないと理解されるべきである。本開示に係る細胞には、真核細胞、原核細胞、動物細胞、植物細胞、真菌細胞、古細菌細胞、真正細菌細胞などが含まれる。細胞には、酵母細胞、植物細胞、および動物細胞などの真核細胞が含まれる。具体的な細胞としては、哺乳動物細胞が挙げられる。具体的な細胞としては、ヒト誘導多能性幹細胞など、多能性幹細胞などの幹細胞が挙げられる。

【 0 0 3 1 】

標的核酸としては、変異 R N A 誘導型 D N A 結合タンパク質ヌクレアーゼがニック形成または切断に有用であり得る任意の核酸配列が挙げられる。標的核酸としては遺伝子が挙げられる。本開示の目的のために、二本鎖 D N A などの D N A は標的核酸を含んでいてもよく、共局在複合体は、共局在複合体が標的核酸に対する所望の効果を与えられ得よう、その標的核酸において、またはその標的核酸に隣接して、またはその標的核酸の近傍で、D N A に結合、または D N A と共局在することができる。そのような標的核酸は、内在性(または天然)の核酸および外来性(または外来)の核酸を含んでいてもよい。当業者は、本開示に基づき、標的核酸を含む D N A に共局在するガイド R N A および変異 C a s 9 タンパク質を容易に特定またはデザインすることができる。D N A には、ゲノム D N A 、ミトコンドリア D N A 、ウイルス D N A 、結合可能なエレメント、または外来性 D N A が含まれる。

【 0 0 3 2 】

外来核酸(すなわち、細胞の天然核酸組成物の一部でないもの)は、当業者に公知の任意の導入方法を用いて細胞に導入されてもよい。そのような方法には、遺伝子導入法、形質導入法、ウイルス形質導入法、マイクロインジェクション法、リポフェクション法、ヌクレオフェクション法、ナノ粒子銃(nanoparticle bombardment)法、形質転換法、結合法(conjugation)などが含まれる。当業者は、容易に特定可能な文献資料を用いてそのような方法を容易に理解し、適用するであろう。

【 0 0 3 3 】

ある態様によれば、C a s 9 のファミリー内の種内であまり保存されていないか、もしくは分岐した C a s 9 タンパク質の一部分、または C a s 9 のファミリー内の種内の保存エッジ間にある C a s 9 タンパク質の一部分を欠失させることにより、C a s 9 タンパク質をコードするために必要な遺伝物質を小さくする。機能的 C a s 9 をコードするために必要な核酸のサイズを小さくすることにより、C a s 9 を送達するようにデザインされた

ベクターに、ガイドRNAまたは調節エレメントまたはエフェクタードメインをコードする核酸などの付加的な核酸を含めることができる。最も小さな特徴解析されているCas9ファミリーメンバーを使用する場合、Cas9を所望のゲノム座位に適切に局在させるために必要な遺伝因子(Cas9タンパク質およびgRNA)をコードするのに約4,500キロベースのDNAが必要となる。約4,500塩基対のCas9は、アデノ随伴ウイルス(AAV)系ウイルスベクター(欧州で規制承認されているウイルスベクター)内へのパッケージングのサイズ限度に近い。さらに、プログラム可能なDNA結合タンパク質に融合される第1の転写ドメインおよびエピジェネティックエフェクタードメインのいくつかは2,000塩基対より大きいため、Cas9に融合させた場合、AAVベクターのパッケージング限度を遙かに超え、レンチウイルスのパッケージングシステムの限度(約8,000塩基対)に近づく。

10

【0034】

以下の実施例は本開示の代表例として記載するものである。本開示、図面、および添付の特許請求の範囲を考慮して、これらの実施例および他の同等の実施形態が明らかになるので、これらの実施例は本開示の範囲を限定するものと解釈されるべきではない。

【実施例】

【0035】

実施例 I

Cas9変異体

Cas9をコードする遺伝子サイズが大きいという問題を克服するために、最も小さい特徴解析されているCas9ファミリーメンバーの1つであるNM Cas9(ナイセリア・メニンジティディスのCas9)内で種々の領域の標的化欠失を行う。NM Cas9のサイズは3,249bpである。ゲノム標的化のための必要条件およびヌクレアーゼ活性に関わる残基を決定する。よりサイズの小さいNM Cas9を作製するために、Cas9タンパク質のアラインメントを作製し、低度保存された隣接区間または保存エッジ間の区間を欠失のために同定した。複数の目的領域を同定し、NM Cas9から選択的に除去した後、Cas9抑制因子アッセイを用いて機能を評価した。アッセイでは、ヌクレアーゼ活性を欠いているレポーター遺伝子の5'領域を標的化可能なNM Cas9変異体を用いた。NM Cas9は、レポーター遺伝子に結合できる場合は転写を抑制し、蛍光レポーターの場合、細胞が非蛍光性となる。

20

30

【0036】

Cas9のアラインメントおよび欠失予測:PFAMデータベース、またはjackHMMER(その全体が参照により援用されるR.D. Finn, J. Clements, S.R. Eddy, Nucleic Acids Research (2011) Web Server Issue 39:W29-W37)などのデータベースから、Cas9ホモログの全長配列を得た。配列のコレクションがアライメントされない場合、CLUSTALW(その全体が参照により援用されるSievers F, Wilm A, Dineen DG, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Soding J, Thompson JD, Higgins DG (2011))などのアラインメントアルゴリズムまたは同等物を用いてアラインメントを作成する。アラインメントを目的配列の位置まで計算的にカットし、アラインメントの偏りが少なくなるように調整した(例えば、ペアワイズ同一性が95%を超える配列を除去した)。その位置でのアミノ酸およびギャップの量を考慮して、位置当たりのアミノ酸頻度のエントロピーまたは相対エントロピーとして保存度を計算する。低度保存領域または保存エッジ間領域を欠失の標的とする。実験検証を反復して、欠失を拡大するか、またはシフトさせる。

40

【0037】

欠失体の構築および特徴解析:その全体が参照により援用されるEsvelt, K. M., Mali, P., Braff, J. L., Moosburner, M., Yaung, S. J., and Church, G. M. (2013) Nat Methods 10, 1116-1121に記載されるように、ヌクレアーゼ欠損NM Cas9を発現する細菌プラスミドをあらかじめ作製した。NM Cas9内に標的欠失を生じさせるために、その全体が参照により援用されるGibson, D. G., Young, L., Chuang, R. Y., Venter,

50

J. C., Hutchison, C. A., 3rd, and Smith, H. O. (2009) Nat Methods 6, 343-345に記載されるように、ギブソン・アセンブリを用いた。重複する相補的部分を含み、NM Cas 9内の標的領域を除去してSGGG Sリンカーを挿入するようにデザインされたプライマーを購入し、PCR反応に用いた。PCR断片をゲル精製し、ギブソン・アセンブリを用いてインビトロでアセンブルし、大腸菌に形質転換した。クローンの配列を検証し、(2つではなく)単一のプラスミドがNM Cas 9のスペーサー、標的プロトスペーサー、およびNM Cas 9活性に対するYFPレポーターを含む、あらかじめ作製されたNM Cas 9レポータープラスミド((2013) Nat Methods 10, 1116-1121参照)の改変型を用いてクローンを試験した。簡潔に述べると、このアッセイでは、合成NM Cas 9バリエーションおよびレポータープラスミドで細胞を同時形質転換する。次に、二重に形質転換された細胞を37 °Cで生育し、蛍光プレートリーダーを用いてYFP蛍光の量を測定し、野生型ヌクレアーゼ欠損NM Cas 9を有する対照プラスミドおよびレポータープラスミドで形質転換された細胞と比較する。

10

【0038】

以下の配列は、YFPレポーターアッセイにより測定されるように、ほぼ野生型レベルに近い活性を保持していた2つの最も大きなNM Cas 9単一欠失変異体の配列である。NM Cas 9内の欠失領域を置換しているSGGG Sリンカーの配列は、大文字で示される。

【0039】

【化 1】

NM-Cas9-Δ255-449

atggccgccttcaagcccaacccatcaactacatcctgggectggccatcggcatcgcagcgtggctggccatggtgagatcga
cgaggacgagaacccatctgctgatcgaacctgggtgtgctgctgttcgagcgcctgaggtgcccagactggtgacagtctggcta
tggctcggcgttctgctcgtctgttcggccttactcggcggcgcctcaccgcttctgcgcctcgcggcctgtgaagcgcgagg
gtgtctgcaggtgccgacttcgacgagaacggcctgatcaagagcctgcccaactccttggcagctgcgcgtgccctctggac
cgcaagctgactcctctggagtggagcggctgtctgtcacctgatcaagcaccggcctacctgagccagcgcaagaacgagggc
gagaccgcccagcaaggagctgggtgtctgtgtaaggcgtggccgacaaccccacgcccctgcagactggtgactccgcactcctg
ctgagctggccctgaaagttcgagaaggagagcggccacatccgcaaccagcgcggcgactacagccacaccttcagccgcaagg
acctgcagggcagctgatcctgtgttcgagaagcagaaggagttcggcaacccccacgtgagcggcgccctgaaaggaggatcgc
agacctgtctgatgccagcggcccggcctgagcggcgacccgtgcagaagatgTCCGGCGGCGGTTTCGggcgacc
actacggcaagaacaccgaggagaagatctactgcctctatccccggcagagatccgcaaccccggtgtgtctgcgcgacct
gagccagcccgcaagggtgatcaacggcgtgtgtgcccgtacggcagcccccccgcatccacatcgagaccgcccgcgagggtg
gcaagagcttcaaggaccgcaaggagatcgagaagcggcaggagagaaccgcaaggaccgcgagaaggccggcccaagttcgc
cgagtacttcccactctggtggcgagcccaagagcaaggacatctgaagctgcgcctgtacgagcagcagcacggcaagtgcctgt
acagcggcaaggagatcaacctgggcccctgaacgagaaggcctacgtggagatcgcctgcccctgccttcagccgcacctggg
acgacagcttcaacaacaaggtgtgtgtgtggcagcgggctcagaacaaggcaaccagacccccctacgagtacttcaacggca
ggacaacagccgcgagtgagcaggagttcaaggcccggctggagaccagccgttccccgcagcaagaagcagcgcctctgtctc
agaagttgacgaggacggcttcaaggagcgaaccgaaacgacaccgctacgtgaaccgttctgtgcccagttctgtgcccagccg
atgcgcctgaccggcaaggcgaagaagcgcgtgttcgcagcaacggccagatccaacctgtctgcgcggttctgtgggctgtgc
aaggtgcgcgccgagaacgaccgcaccacgcctggacgccgtggtgtggcctgcagcacctggccatgcagcagaagatac
ccgtctgtgcgtacaaggagatgaacgcttcgacggtaaacaccatcgacaaggagaccggcgaggtgtgcaccagaagaccac
ttccccagccctgggagttcttgcggaggagggtatgatccgctgttcggcaagcccagccgagttcgaggaggccg
acaccccgcgagaagctgcgaccctgtgcccgagaagctgagcagcggcctgaggccgtgcagagttactctctgttctgt
gagccgcgcccccaaccgcaagatgagcgggtcagggcctacgtgagaccgtgaagagcgccaaagcgcctggacgagggcgtgagc
gtgtgcgcgtgccccgaccagctgaagctgaaggacctggagaagatgtgaaccgcgagcgcgagcccagctgtacgagcc
ctgaagccccgctggaggcccacaaggacgaccccccaaggccttgcggagcccttctaacgtacgacaaggccggcaaccgc
accagcagggtgaagccgtgcgcgtgagcaggtgcagaagaccggcgtgtgggtgcgaaccacaacggcatcggcagacaacgc
caccatggtgcgcgtggacgtttcgagaaggcgacaagtactacctgggtcccctacagctggcaggtggccaaaggcctcctgc
ccgaccgcgccgtggtgcagggcaaggacgaggaggactggcagctgatcgacgacagcttcaacttcaagttcagcctgcacccaa
cgacctggtgaggtgatccaagaagcccgcctgttcggctacttccagctgcccaccgcccaccggcaacatcaacatccgc
atccacgacctggaccacaagatcggcaagaacggcctcctggaaggcctcggcgtgaagaccgcccctgagcttccagaagtaccaga
tcgacgagctggcaaggagatccgcccctgcgctgaagaagcgcctcctgtgcgctaa [配列番号 1]

10

20

30

40

【 0 0 4 0 】

【化 2】

NM-Cas9-Δ567-654

atggccgcttcaagcccaaccccatcaactacatcctgggctggccatcggcatcgcagcgtgggctgggcatggtgagatcga
cgaggacgagaaccccatctgctgacgacctgggtgtgctgctgagcgcctgaggtgcccagactggtgacagtctggcta
tggctcggcggcttctctctcttctggccttactcggcggcgcctcacgccttctgctcgcctcggcctgctgaagcgcgagg
gtgtctgacaggtgccactctgacgagaacggcctgatcaagagcctgcccaacactccttggcagctgctcgcctcctggac
cgcaagctgactcctctggagtggagcggcgtgctgctgacctgatcaagcaccgctgctacctgagccagcgcaagaacgaggc
gagaccgccgacaagagctgggtgctctgctgaagggcgtggccgacaacgcccacgcccctgacagctggtgactccgactcctg
ctgagctggccctgaacaagttcagaaggagagcggccacatccgcaaccagcggcggactacagccacacttcagccgcaagg
acctgacggccgagctgactctgctgttcgagaagcagaaggagttcgcaacccccacgtgagcggcggcctgaaggaggcgcacg
agacctgctgatgaccagcggccccctgagcggcgcacggctgacagaagatgctggccactgcaccttcgagccagccgagc
ccaaggccgcaagaacactacaccggcagcgttctctgctgaccaagctgaacaacctgctcctgagcagggcagcgc
agcggccccctgaccgacaccgagcggcgcaccctgatggagcagccctaccgcaagagcaagctgacctagcccaggccccgaag
ctgctgggtctggaaggacaccgcttctcaaggccctgctgctacggcaaggacaacggcggagccagcaccctgatggagatgaagg
cctacacgccatcagccgccccctggagaaggaggccctgaaggacaagaagatcctgaacctgagccccgagctgcaaggacg
agatcggcaccgcttcagcctgttcaagaccgacgaggacatcaccggcggcctgaaggaccgcatccagcccagatcctggaggc
cctgctgaagcacatcagcttcgacaagttctgctgagatcagcctgaaggccctgctcgcctcagctgccccctgatggagcagggcaagc
gctacgacgagcctgctgccccgagatctacggcaccactacggcaagaagaacacggaggagaagatctacctgctcctatccccgc
cgacgagatccgcaacccccctggtgctgctgccccctgagccagccccgaaggtgatcaacggcgtggtgcccgtctacggcagccc
cgccccgatccacatcgagaccggccgaggtgggcaagagctcaaggaccgcaaggagatcagaagcggcagggagagaac
cgcaaggaccgagagaaggccggccaagtccgagtagtctcccaactctggtggcggagcccgaagagcaaggacatctgaag
ctgctgctgacgagcagcagcagcggcaagtgcTCCGGCGGCGGTTCGcagaagttcagcaggacggcttcaaggagc
gcaacctgaacgacaccgctacgtgaaccgcttctgtgccaagtctgtggccgaccgcatgctgctgaccggcaagggcaagaagc
cgtgtctgcccagcaacggccagatcaccacctgtgctgctgctgctggcctgctgcaaggtgctgcccggagaaagaccgcccacca
cgccccctgacgcccgtgtgtgctgctgacaccgtggccatcagcagaagatcaccgcttctgctgcttacaaggagatgaacgcc
ttcagcggtaaaaccatcgacaaggagaccggcgaaggctgctgaccagaagaccacttccccagccctgggagttcttgcgccagg
aggatgatgaccgctgtgtgcaagcccagcggcaagcccgagttcgaggaggccgacacccccgagaagctgctgacacctgctgg
ccgagaagctgagcagccgcccctgaggccgtgacgagtagctcctctgtctgtgagccgcccccaaccgcaagatgagcgg
tcagggtcacatggagaccgtgaagagcggcaagcctgacgagggcgtgagcgtgtgctgctgcccctgaccagctgaagct
gaaggacctggaagaatggtgaaccgagcggcagcccaagctgtacgagccctgaaggcccctggaggcccacaaggac
gacccccgcaaggccttctgcccagcccttctacaagtacgacaagccggcaaccgaccagcaggtgaaggccgtgctgctgag
caggtgcagaagaccggcgtgtgggtgcaaccacaacggcctcgcgacaacgccaccatggtgctgctgagcgtgttcgagaag
ggcgacaagtactactgtgcccctctacagctggcaggtggccaaggccatcctgcccagccgcccgtgtgctgagggcaaggac
gaggaggactggcagctgacgacagcttcaactcaagttcagcctgcaccccacacactggagggtgatcacaagaaggc
ccgcatgttctgctacttctgagctgcccagcggcaccggcaacatcaacatccgatccacgacctggaccacaagatggcaaga
acggcatcctgagggcctggcgtgaagaccggcctgagcttcagaagtagcagatcagcagctgggcaaggagatccgcccct
gccgctgaagaagcggcctcctgtgctgctaa [配列番号 2]

10

20

30

40

【 0 0 4 1 】

【化3 - 1】

YFPレポータープラスミド

agctctcgaacccagagtccegcfcagaagaactcgtcaagaaggcgcgtagaaggcgcgctgcgaatcgggagcggcgatacc
 gtaaagcagagggaagcggcagccattcgcgcgaagctctcagcaatcaccggtagccaacgctatgtctctgatagcggcgcgc
 cacaccagccggccacagtcgatgaatccagaaaaggcgcattttccaccatgatattcggcaagcaggcatcgccatgggtcacga
 cgatcctcgcgcggcagtcggcgccttgagcctggcgaacagctcggcctggcgcgagcccctgatgctcttctccagatcatcct
 gatcgacaagaccggcttccatccgagtacgtgctcgcgcgatgcgagtcttcgcttgggtggcgaatgggcaggtagccggatcaagcgt
 atgcagccggccattgcatcagccatgatggatactttcggcaggagcaaggtgagatgacaggagatcctgccccggcacttcgcc
 caatagcagccagtccttcccgcttcagtgacacgctcagcagcagctgcgaaggaaacgcccgtcgtggccagccacgatagccgc
 gctgcctcgtctcagttcattcagggcaccggacaggtcggcttgcacaaaaaagaccggcggccccgcgctgacagccggaacac
 ggcgcgatcagagcagccgattgtctgttgcctcagtcataccgaatagcctctccaccaagcggccggagaacctcgtgcaatc
 catcttgtcaatcatgcgaaacgatcctcatcctgtctcttgcagatcttgatccccgcgcacatcagatccttggcggcaagaaagccat
 ccagttactttgcagggttcccaacctaccagagggcggccccagctggcaattccgacgtctaagaaaccattattatcatgacattaac
 ctataaaaataggcgtatcacagggcccttctgtctcaccctcaggggacaatgaaaacgttagtcatggcgcgccttgacggctagctca
 gtcttaggtacagtgctagcttaatgctcgcacatagcagaactttaaagttatcgcctatggtgtagctccccttctcatttcgagtgctaca
 tccggccgtatggctccacgtagagcatacggaaaaaaagtcaaaagcctccgaccggaggtcggccttacttctagcagagtttgta
 gaaacgcaaaaaggccatccgctcaggatggccttctgttaatttgatgcttgcagtttatgggggcgtccctgcccggccacctccggg
 ccgttgcctcgaacgttcaaatccgctcccggcggatttctctactcaggagagcgttaccgacaacaacagataaaacgaaaggc
 ccagctttcagactgagcctttcgttttattgatgcttggcagttccctactctcgcagtggggagaccccacactaccafcggcgcctacggcg
 tttcacttctgagttcggcatggggctcaggtgggaccaccgcgctactcgcggcaggcaaatctgttttatcagacccttctcgttctgat
 ttaatctgtatcaggtgaaaatcttctcctccgcaaaacagccccgtagaaaaaggagcgttgatcggcacgtaagaggtccacgat
 aatatctaacaccgtcgttggactatfttaccctcggcgtgataatggttgcagtactagaattctttaaacttaagaaggagataacata
 tgaatcccctatggcgaatacttttaagctcgtaaaggagaagaactttcactggagttgtcccattcttggtaattagatgggtgatfta
 atgggcacaaatttctgctagtgagaggggtgaaggtgatcacaatacggaaaacttacccttaaatttatttgcactactggaaaactac
 ctgttccatggccaacacttgcactacttccggftatggctctaaaatgctttgctagataccagatcatatgaaacggcatgactttttcaaga
 gtgccatgcccgaaggttatgtacagaaaagaactatattttcagaatgacgggaactacaagacacgtgctgaagtcaagtttgaaggt
 gatacccttftaataagaatcaggttaaaaggtattgattttaaagagaaggaaacattcttgacacaattggaatacaactataactaca
 caatgtatacatcatggcagacaaacaaaagaatggaatcaagtttaacttcaaaatttagacacaacattgaagatggaagcgttcaactag
 cagaccattatacaaaaaactcctaattggcagtgcccctgtccttttaccagacaaccattaccctgcctatcaatctgcccttccgaaagat
 cccaacgaaaagagagaccacatggctccttctgagtttgtaacagctgctgggattacacatggcatggtgaaactatacaaaatagctta
 accgaagcgttgatagttgatacttttgcctcggcgcgaactagaggcatcaataaaacgaaaggctcagtcgaaagactgggccttt
 cgttttatctgttgttgcggtaacgctctcctgagtaggacaaatccggcgccttagacctagggtacgggttttctgcccgaacgg

10

20

30

【化3 - 2】

gctgttctggtgttgctagttgttatcagaatcgcagatccggctcaggttgcggctgaaagcgtatttctccagaattgccatgatttt
tccccacgggagggcgtcactggctcccgtgtgtcggcagcttgattcgataagcagcctgcctgttccaggctgtctatgtgtgactgttg
agctgtaacaagttgtctcagggttcaatttcatgttctagttgctttgtttactggtttcacctgttctattagggtttacatgctgttcacgtgta
cattgtcgtatctgttcattggtgaacagctttaaattgcacaaaaactcgtaaaagctctgatgtatctatctttttacaccgttttcatctgtgcat
atggacagttttccctttgatatacagggtgaacagttgttctacttttgtttgttagctttgatcttccactgafagatacaagaccataagaac
ctcagatccttccgatttagccagtagttctctagtggtggttctgttttgcgtgagccatgagaacgaaccattgagatcatgcttactttg
catgtcactcaaaaatttgcctcaaaaactgggtgagctgaattttgcagttaaagcatcgtgtagtgttttcttagtccgttacgtaggtaggaa
tctgatgtaattggtgtgtgattttgtcaccattcattttatctgggtgttctcaagttcgggttacgagatcatttgtctatctagttcaacttgaa
aatcaacgtatcagtcggcggcctcgttatcaaccaccaatttcatattgtctgaagtgtttaaacttttacttattggtttcaaaaaccattgg
ttaagcctttaaactcatgtagttattttcaagcattaacatgaacttaaattcatcaaggcctaactctatatttgcctgtgagtttctttgtgtt
agttcttttaataaccactcataaactcctagatgattgttttcaaaagacttaacatgttccagattatattttatgaattttttaaactggaaaag
ataaggcaatctcttactaaaaactaatttcaattttctcgtgagaacttggcagatgttgcactggaaaactcaaaagcctttaaaccaaa
ggattcctgatttccacagttctcgtcctcagctctctgtggttcttagctaaacaccataagcatttccctactgagttcatcatctgagcgtgta
ttggttataagtgaaacgataccgtccgttcttccctttaggtggtttcaatcgtgggttgagtagtggccacacagcataaaatagcttggtttc
atgctccgtaagtcatagcgcactaatcgtagttcatttgccttggaaaaaactaattcagacatacatctcaattgtctaggtgattttaaataca
ctataccaattgagatgggctagtcattgataaattactagctcttttccctttaggtgtggtatctgttaaattctgctagaccttctgfgaaaact
tgtaaattctgctagacctctgtaaattccgctagaccttgggtgtttttttgtttatattcaagtggttataattatagaataaagaagaataa
aaaaagataaaaagaaatagatccagccctgtgtataactcactacttttagctcagttccgcagattacaaaaggatgtcgcaaaacgctgttt
gctcctctacaaaaacagaccttaaaacctaaaggcttaagtagcaccctcgaagctcgggcaaatcgtgaatattcttttctcctcagc
catcaggcacctgagtcgtcttttctgtacattcagttcgtcgtcgcctcagcgtctgagtgaaatgggggtaaaaggcactacagggc
gccttttatggattcatgcaaggaaactaccataatacaagaaaagcccgacagggccttctcagggcgttttatggcgggtctgtctatgtg
gtgctatctgacttttctgttccagcagttctcctctgattttccagctgaccacttccgattatcccgtgacaggtcattcagactggccta
atgcaccagtaaggcagcgggtatcatcaacagcttacccttactgtccctagtgcttggattctaccaataaaaaaacgcccggcgg
caaccgagcgttctgaacaaatccagatggagttctgaggtcattactggatctatcaacaggagccaagcg [配列番号3]

10

20

30

【0042】

本明細書に記載の方法に従って、NM Cas 9内の複数の欠失が同定された。最大の NM Cas 9 255 449は595塩基対が除去されており、レポーターアッセイにより測定されるように、活性低下はわずか16%である。ある態様によれば、NM Cas 9などの野生型Cas 9より1000塩基対または900塩基対少なく、野生型レベルに近い活性を保持している、変異Cas 9タンパク質が提供される。

【0043】

実施例II

欠失のためのCas 9ヌクレアーゼドメインの標的化

Cas 9のニックーゼ欠損アレルまたはヌクレアーゼ欠損アレルが望まれる場合、低度配列保存領域または配列保存エッジ間領域を標的化すると共に、Cas 9ヌクレアーゼドメインを周囲のヌクレオチドと共に欠失のための標的としてもよい。このようなアプローチを用いて、HNHモチーフおよび周囲のヌクレオチドを欠いた機能的NM Cas 9アレルであるNM Cas 9 567 654を作製した。このアレルは、YFPレポーターアッセイで測定されるように、野生型とほぼ同じDNA結合能を保持していた。

40

【0044】

実施例III

偏りのないNM Cas 9欠失ライブラリーの構築方法

Cas 9欠失体を作製するために標的アプローチに加えて、本開示の態様には、ランダム欠失生成および機能的変異体のスクリーニングのためのハイスループットアプローチが

50

含まれる。例示的な方法によれば、無差別 (promiscuous)ヌクレアーゼ、超音波処理、サンプルの反復ピペティング、または他の化学的、酵素的、もしくは環境的手段を用いて、所望の Cas9 アレルを含むプラスミド DNA を切断してもよい。断片化後、プラスミド DNA をエキソヌクレアーゼで処理して、Cas9 遺伝子からヌクレオチドを除去してもよい。エキソヌクレアーゼ処理後、断片化末端をマングベーンヌクレアーゼまたはクレノーポリメラーゼなどの酵素で平滑末端化し、互いにライゲーションさせて、ランダム欠失を含む Cas9 プラスミドを再生する。Cas9 の欠失部分内にリンカーまたはエフェクターモチーフなどの外来性ドメインを挿入するために、平滑末端化された断片 DNA にそのようなドメインをライゲーションさせてもよく、次に、プラスミドの環状化により、Cas9 の欠失部分内に外来性ドメインが挿入された Cas9 コード配列が得られる。

10

【0045】

この時点で、ライブラリーを Cas9 活性のレポーターアッセイを含む細胞に形質転換してもよく、機能的活性を維持しているライブラリーのメンバーを同定してもよい。あるいは、スクリーニングするライブラリーのサイズを小さくするために、新たに作製されたライブラリー由来の Cas9 のコード配列を消化または PCR により単離してもよく、最初の野生型 Cas9 遺伝子より短くなるように断片をサイズで選択してもよい。次に、これらのより小さいメンバーを最初のベクター中に再度ライゲーションし、Cas9 活性に対するレポーターを含む細胞に形質転換してもよい。

【0046】

20

プラスミドの切断に加えて、3 側に Cas9 遺伝子に対する相同性を有するが 5 側には互いに対する相同性を有し、各オリゴヌクレオチドの 3 末端が Cas9 内の約 30 塩基対の異なる区間に結合する、オリゴヌクレオチドのライブラリーを作製してもよい。これらのオリゴヌクレオチドは Cas9 コード配列のセンス鎖およびアンチセンス鎖の両方をカバーする。次に、これらのオリゴヌクレオチドを用いて PCR を行い、所与のセンス PCR 反応産物のそれぞれが他のアンチセンス PCR 産物の全てに対する相補性を有し、逆も同じである、一連の Cas9 断片を作製することができる。次に、ギブソン・アセンブリまたはオーバーラップ・エクステンション PCR などの方法を用いてこれらの断片同士をアニーリングさせた後、ベクター骨格中へのライゲーションおよび細胞への形質転換を行い、Cas9 遺伝子のランダムな区間が除去された Cas9 バリエーションのライブラリーを作製することができる。より長いリンカーのため、または欠失領域内にエフェクタードメインを挿入するためには、オリゴヌクレオチドはその 5 末端に、より長いリンカーまたはエフェクタードメインに対する相補性を有するべきであり、次に、このドメインは、ギブソン・アセンブリ反応中またはオーバーラップ・エクステンション PCR 中に導入されるべきである。ライブラリー作製されると、本明細書に記載の YFP レポーターシステムなどのレポーターアッセイを用いて機能的バリエーションを同定することができる。

30

【0047】

実施例 IV

ベクター構築

Cas9ヌクレアーゼ欠損プラスミドは、ST1 (Addgene #48659)であるか、またはNMプラスミドおよびTDプラスミド(それぞれ、Addgene #48646および48648)から、以下の点変異(NM: D16A D587A H588A N611A、およびTD: D13A D878A H879A N902A)を導入することにより構築した。ギブソン・アセンブリを用いてCas9欠失体を作製した。内部欠失体を作製する場合、連結断片間のリンカーを欠くNM 566 620以外は、内部欠失体を5アミノ酸のSer Gly Gly Gly Serリンカーで連結した。N末端ドメイン交換により、ST1の残基1~117をNMの残基118~1082に融合させた。C末端ドメイン交換により、NMの残基1~727をST1の残基743~1121に融合させた。

40

【0048】

50

実施例 V

細菌レポーターコンストラクト

欠失変異体の解析に用いるレポーターコンストラクトは、単一の SC101 kanR プラスミド骨格中にスペーサーエレメントおよび YFP レポーターを併せ持つこと以外は、以前に発表されているものと同様である。ドメイン交換解析のためのレポーターコンストラクトは以前に使用されているものと同じである。その全体が参照により援用される Esvelt, K. M. et al. Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. Nature methods 10, 1116-1121, doi:10.1038/nmeth.2681 (2013) を参照されたい。

【0049】

10

実施例 VI

哺乳動物レポーターコンストラクト

MST1n VP64 コンストラクト、ST1 ガイド RNA プラスミド、および ST1 特異的哺乳動物転写レポーターは、Esvelt, K. M. et al. Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. Nature methods 10, 1116-1121, doi:10.1038/nmeth.2681 (2013) において以前に発表されている (それぞれ、Addgene # 48675、48672、および 48678)。細菌のコンストラクトの場合と同様に欠失変異体を作製した。

【0050】

20

実施例 VII

抑制アッセイ

試験対象の適切なスペーサー/レポーターコンストラクトおよび Cas9 ベクターで NEB 10 beta 細胞 (New England Biolabs 社) を同時形質転換することにより、Cas9 抑制アッセイを行った。形質転換によって生じたコロニーをピックアップし、96 ウェルプレート中で連続振盪しながら 37 °C で生育した。翌日、Synergy Neo マイクロプレートリーダー (BioTek 社) を用いてプレートを読み取り、495 ~ 528 nm の蛍光および 600 nm の吸光度を測定した。交換実験では、以前に公開されている 2 つの異なったスペーサー/プロトスペーサーの組合せ (A および B) (Esvelt, K. M. et al. Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. Nature methods 10, 1116-1121, doi:10.1038/nmeth.2681 (2013) 参照) を試験した。他の全ての実験では、スペーサー/プロトスペーサーの組合せ B のみを調べた。

30

【0051】

実施例 VIII

細胞培養およびトランスフェクション

10% FBS (Invitrogen 社) およびペニシリン/ストレプトマイシン (Invitrogen 社) を添加した、高グルコースダルベッコ変法イーグル培地 (Invitrogen 社) 中で HEK 293T 細胞を維持した。加湿インキュベーター中に細胞を 37 °C、5% CO₂ で維持した。1 ウェル当たり 50,000 細胞を播種した 24 ウェルプレート中で細胞に遺伝子導入した。2.5 μl のリポフェクタミン 2000 を用いて、400 ng の Cas9 活性化因子、100 ng の gRNA、および 60 ng のレポータープラスミドを各ウェルに送達した。細胞をさらに 36 ~ 48 時間生育した後、免疫蛍光または FACS を用いてアッセイした。

40

【0052】

実施例 IX

多重配列アラインメントおよびエッジフィルタ

MUSCLE で PFAM データベースの Cas9 配列 (PF13395、798 個の配列) をリアライメントし、MATLAB スクリプトを用いて多様性および全長配列についてアラインメントを調整することにより多重配列アラインメントを作成した。以下に更に具体的に説明するこの方法により、217 個の配列が得られた。

50

【 0 0 5 3 】

P F A Mデータベース中のC a s 9配列（P F 1 3 3 9 5、7 9 8個の配列）の最終的なリアライメントを得るために、以下のステップ（プログラムコードは全てM A T L A B）を行った。Fonfara I, et al., (2014) Phylogeny of Cas9 determines functional exchangeability of dual-RNA and Cas9 among orthologous type II CRISPR-Cas systems, *Nucleic Acids Res.* 42, 2577-90からのリアライメントを得、グループI I AおよびI I Cからの配列のみを含め、後者からは、I I AとI I Cを分離する分岐までの配列のみ（図2中で赤色の四角で印付けられている）の合計49個の配列を含めた。

【 0 0 5 4 】

次に、配列を2つのグループ、すなわち150位付近に大きな挿入を有するグループおよび有さないグループに分けた（これは、例えば一方でN M C a s 9とS T C a s 9を区別し、他方でS P C a s 9とT D C a s 9を区別する）。M U S C L E (Edgar, RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput, *Nucleic Acids Res* 32, 1792-97参照)を用いてこれらのグループを別個にリアライメントし、（配列の長さのため）ウィンドウアプローチ (windowed approach) を用いてリアライメントした後、1つのシードリアライメントへとプロファイル間リアライメントを行った。

【 0 0 5 5 】

P F 1 3 3 9 5中の配列を、M U S C L Eでシードリアライメントを用いてリアライメントした。シードを用いた全てのリアライメントは、各標的配列を1つずつシードに対してリアライメントすることにより行われる。このリアライメントを用いて、シードと標的配列の間でトップヒット同一性を決定し、トップヒット同一性の降順に再度整理する。次に、今度は同一性の降順に、標的配列をシードに対して1つずつリアライメントする。この2ステップのアプローチをとることでリアライメントのロバスト性を確保する。さらに、これらの配列を、挿入を含むか否かによって分け、2つの別個のグループをプロファイルとしてシードとリアライメントする。短い配列および大きな短縮を有する配列を手動で除く。90%を超えるペアワイズ類似性を有する配列を除く。

【 0 0 5 6 】

得られた2つのリアライメントをプロファイル間で互いにリアライメントし、合計217個の配列を得た。得られたリアライメントを、目的のC a s 9オルソログの位置まで短縮する。

【 0 0 5 7 】

大腸菌O 1 5 7の全遺伝子の平均アミノ酸バックグラウンド頻度に対する相対エントロピーとして配列保存度を計算した。得られた保存プロファイルにガウス差分 (D o G) エッジフィルタ（その全体が参照により援用されるMarr, D. & Hildreth, E. Theory of edge detection. *Proceedings of the Royal Society of London. Series B, Containing papers of a Biological character.* Royal Society 207, 187-217 (1980) 参照) を適用し、パラメータ選択に対するロバスト性を実現するために複数の長さスケールで平均化し、種々の長さスケールで検出することにより、ドメイン境界検出を行った。

【 0 0 5 8 】

具体的には、リアライメントの保存度を、大腸菌O 1 5 7の全遺伝子の平均頻度に対するアミノ酸頻度の相対エントロピーとして計算した (Cover, TM and Thomas, JT (2006) *Elements of Information Theory*; 2nd edition, Wiley-Interscience参照)。

【 0 0 5 9 】

10

20

30

40

【数1】

$$D_i = \sum_{a=1}^{20} p_i^a \log \frac{p_i^a}{q^a}$$

【0060】

式中、 p_i^a は、位置 i におけるアミノ酸 a の頻度であり、 q^a は、アミノ酸 a の平均頻度である。20アミノ酸の全体の総和である。 \log の底は2であり、エントロピーは

10

ビットで与えられる。平均頻度 q^a は以下の通りである。

	A	C	D	E	F	G	H	I	K	L	M	N	P	Q	R	S	T	V	W	Y
	0.094	0.012	0.052	0.058	0.038	0.073	0.022	0.059	0.045	0.104	0.027	0.041	0.044	0.044	0.057	0.060	0.055	0.070	0.015	0.029

NM Cas 9 の位置まで短縮した後、Cas 9のアラインメントにより、図3の保存プロファイルが得られる。

【0061】

保存プロファイルを、(IIA型サブファミリーの、より大きなCas 9タンパク質の一例である)SP Cas 9における位置までプロット短縮する。N末端のおよそNM 145 (SP 170) の位置まで、およびさらにNM 200 (SP 400) の位置の後に、同様な特徴が観察され、これは前述の大きな挿入である。図4を参照されたい。

20

【0062】

マルチスケールエッジフィルタを保存プロファイルに適用することにより潜在的ドメイン境界を同定した。このフィルタにより、様々なスケールでガウス差分(DoG)(その全体が参照により援用されるMarr, D and Hildreth, E (1980) Theory of Edge Detection, Proc R Soc Lond B Biol Sci 207, 187-217参照)が計算され、得られたグラフが総計される。この曲線の極値は、タンパク質中の低度保存領域と高度保存領域の境界であると解釈される。これらのドメインは、進化的分岐速度の差につながる潜在的に異なる機能的な重要性のため、異なる保存レベルを示し得る。異なる保存度は、進化的時間軸にわたるドメイン挿入により生じたマルチドメインタンパク質の特徴であり得る。極値の値は、重要性の点から境界を順位付けるために用いられる。

30

【0063】

NMの位置に限定したCas 9アラインメントでは、本明細書に記載の方法を用いて以下の境界または保存エッジまたはエッジアミノ酸が同定された。

【0064】

【表 1】

順位	位置	極値の Abs 値
1	736	0.0225
2	620	0.0148
3	554	0.0148
4	472	0.0144
5	288	0.0136
6	144	0.0131
7	87	0.0106
8	661	0.0101
9	825	0.0100
10	205	0.0099
11	512	0.0092
12	414	0.0092
13	108	0.0092
14	853	0.0067
15	438	0.0050
16	936	0.0041
17	1011	0.0038

10

20

【0065】

実施例 X

Cas9ファミリーメンバーの計算解析

欠失に適していると考えられるマルチドメイン Cas9 タンパク質内の領域を同定するために、バイオインフォマティクスのアプローチを用いてドメイン間の潜在的境界を同定した。よく精選されたシードアラインメントを用いて、前述したように、PFAM から得た全長 Cas9 配列 (PF13395) を再度アラインメントし、高いペアワイズ同一性を有する配列を除去した。大腸菌の全遺伝子の平均頻度に対する観察されたアミノ酸頻度の相対エントロピー (Cover, T. M., Thomas, J.T. Elements of Information Theory, 2nd edition. (Wiley-Interscience, 2006) 参照) として一次配列保存度を計算した。マルチドメインタンパク質中のドメインは様々なレベルの配列保存度を示すと予想できるので、ドメイン境界の位置を同定するためには、保存プロファイルにマルチスケールエッジフィルタを適用することが好適であると考えられる。

30

【0066】

狭い範囲の空間周波数 (spatial frequency) に対して感受性があるガウス差分 (DOG) バンドパスフィルタを用いて保存プロファイルについてエッジ検出を行った。Marr, D. & Hildreth, E. Theory of edge detection. Proceedings of the Royal Society of London. Series B, Containing papers of a Biological character. Royal Society 207, 187-217 (1980) を参照されたい。種々の長さスケールでの検出を可能にするため、および特定のパラメータ選択に対してフィルタを非感受性にするため、複数のスケール (5 ~ 50 アミノ酸) にわたる平均化を行った。次に、Cas9 アラインメントの保存プロファイルにバンドパスフィルタを適用した。NM Cas9 についての同定された潜在的境界位置を図 5A に示す。上位 6 つの境界を赤色の長い太線で示す。図に示すように、それぞれの全体が参照により援用される Sapranauskas, R. et al. The Streptococcus thermo

40

50

philus CRISPR/Cas system provides immunity in Escherichia coli. Nucleic acids research 39, 9275-9282, doi:10.1093/nar/gkr606 (2011)およびFonfara, I. et al. Phylogeny of Cas9 determines functional exchangeability of dual-RNA and Cas9 among orthologous type II CRISPR-Cas systems. Nucleic acids research 42, 2577-2590, doi:10.1093/nar/gkt1074 (2014)において以前に帰属された500位の上流から始まり750位まで及び既知のHNHおよびRuvCドメインの配置をフィルタは正確に同定している(図5B参照)。560位周辺および620位周辺のドメイン間境界も正確に予想されている。88位の第1のアルギニンリッチヘリックスの境界も予想されている。144位にある、タンパク質のN末端側の上位の境界の1つは、過去にFonfara, I. et al. Phylogeny of Cas9 determines functional exchangeability of dual-RNA and Cas9 among orthologous type II CRISPR-Cas systems. Nucleic acids research 42, 2577-2590, doi:10.1093/nar/gkt1074 (2014)においてドメイン境界として同定されていないが、アルギニン(Arg)リッチヘリックス領域をより詳細に描写しており、今回、第2の保存されたアルギニンリッチヘリックスを含んでいた。

【0067】

実施例XI

Cas9の短縮体および欠失解析

Cas9内の機能的ドメインおよび潜在的非機能的ドメインを実験的に探索するために、N末端およびC末端の短縮体、並びに、本明細書に記載のドメイン検出解析に基づく中程度の一連の内部欠失体を作製した。機能的ヌクレアーゼ欠損Cas9タンパク質がYFPレポーターの5'末端に結合し、それによりレポーターの発現レベルが低下する、大腸菌における転写抑制因子アッセイを用いて欠失の影響を解析した(図6A参照)。それぞれの全体が参照により援用されるQi, L. S. et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell 152, 1173-1183, doi:10.1016/j.cell.2013.02.022 (2013); Esvelt, K. M. et al. Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. Nature methods 10, 1116-1121, doi:10.1038/nmeth.2681 (2013); およびBikard, D. et al. Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. Nucleic acids research 41, 7429-7437, doi:10.1093/nar/gkt520 (2013)を参照されたい。N末端またはC末端の短縮体はいずれもレポーターを抑制できなかったが、288位の境界の下流および上流の2つの内部欠失体であるNM 255 289およびNM 330 389、ならびにNM 566 620欠失体は野生型レベルに近い抑制を示した(図6Bおよび図7)。欠失を反復的に拡大した複数回の解析をさらに行い、レポーターを抑制する能力をアッセイした(図6B)。NM活性をほとんど低下させずに除去可能であった2つの大きな非重複領域254 449および567 654(それぞれ、タンパク質の全長の18%および8%を構成する)が同定された(図6C)。アラインメントに示されるように、NMの254 449位は、Cas9タンパク質に特異的なタンパク質領域内において、比較的低度保存された区間である。567から654位はHNHドメインであり、これはCas9のDNA触媒に重要であることが知られているが、DNA結合については不必要であることが分かっているドメインである。

【0068】

NM Cas9から除去された領域がNMに固有でなく、他のCas9ファミリーメンバーから除去可能な一般的領域であることを実証するために、ストレプトコッカス・サーモフィラス(*Streptococcus thermophilus*)のCas9(ST1)およびトレポネマ・デンティコラ(*Treponema denticola*)のCas9(TD、GI:42525843)のヌクレアーゼ欠損バリエーション内で対応する欠失体を作製し、転写抑制アッセイを用いてこれらの機能を測定した(図8Aおよび図9)。ST1およびTDの両方における対応する欠失変異体がそれらの対応する野生型と同様な活性を示したことから、除去された領域が、Cas9系統樹全体において、さらにTDなどのタイプII Aサブファミリー内のより離れたメンバー間においても、Cas9のDNA結合に不必要であることが示唆される

【 0 0 6 9 】

本明細書に記載の2つの最も大きい機能的欠失体の活性を、転写活性化因子アッセイを用いてST1 Cas9内で試験した(図8B)。大腸菌内での解析と一致して、VP64活性化ドメインに融合させてヒト細胞において蛍光レポーターを標的化した場合に、両方のST1欠失変異体が野生型タンパク質に匹敵する活性を保持していた(図8C~図8D)。2つの欠失変異体の大きい方である255450(ST1のナンバリング)は、サイズが2,793塩基対であるCas9遺伝子を生じる。

【 0 0 7 0 】

実施例XII

Cas9ドメイン交換

Cas9のN末端ドメインおよびC末端ドメインは、crRNA:tracrRNA結合および/またはPAM選択性に重要な役割を果たし得る。活性を解析するために、NMのN末端および/またはC末端をST1の相同な領域と置換した、NMとST1との間の一連のドメイン交換変異体を作製した。次に、レポーター内のガイドRNAおよび/またはCas9特異的PAMを変更して、本明細書に記載の転写レポーターアッセイを用いてキメラタンパク質を試験して、タンパク質特異性に対するドメイン交換の影響を決定した(図10A)。ドメイン交換の正確な位置はドメイン境界解析に基づいて決定された。すなわち、同定された最も有意なN末端境界およびC末端境界(図5A)にできるだけ近い位置であると同時にアラインメント内でほぼ完全に保存されている位置を選択した(図10B)。NMとST1との間のN末端ドメイン交換体はいずれもNMに新規な特性を付与しなかったことから、ST1のN末端はモジュール式ではなく、導入されなかったST1の他の領域と関連して機能することが示唆される(図10C~10F)。C末端交換により、ST1のcrRNA:tracrRNA複合体と相互作用可能であり、さらにST1特異的PAMを有するレポーターを抑制可能である、NM-ST1ハイブリッドが作製された(図10E)。この結果は、図11(A)~11(D)に示されているように、他のST1特異的レポーターを用いてさらに検証された。

本発明の態様は以下を含む。

付記1

DNA結合タンパク質のファミリー内で、低度保存された2つの隣接アミノ酸配列と隣接する高度保存されたアミノ酸配列であるか、または高度保存された2つの隣接アミノ酸配列と隣接する低度保存されたアミノ酸配列である、第1のアミノ酸配列を同定すること

前記第1のアミノ酸配列に対応する核酸配列を同定すること、

前記第1のアミノ酸配列に対応する核酸配列を欠いた、標的DNA結合タンパク質のための核酸配列を作製すること、および

前記作製された核酸配列から変異DNA結合タンパク質を作製することを含む、変異DNA結合タンパク質を作製する方法。

付記2

前記標的DNA結合タンパク質がヌクレアーゼ活性を有する、付記1に記載の方法。

付記3

前記標的DNA結合タンパク質がニッカーゼである、付記1に記載の方法。

付記4

前記標的DNA結合タンパク質がヌクレアーゼ欠損である、付記1に記載の方法。

付記5

前記第1のアミノ酸配列に対応する核酸配列を、前記標的DNA結合タンパク質のための核酸配列から欠失させる、付記1に記載の方法。

付記6

前記第1のアミノ酸配列に対応する核酸配列を、前記標的DNA結合タンパク質のための核酸配列から欠失させてリンカーで置換する、付記1に記載の方法。

10

20

30

40

50

付記 7

前記第 1 のアミノ酸配列に対応する核酸配列を、前記標的 DNA 結合タンパク質のための核酸配列から欠失させて S G G G S リンカーで置換する、付記 1 に記載の方法。

付記 8

前記 DNA 結合タンパク質が Cas 9 タンパク質である、付記 1 に記載の方法。

付記 9

前記第 1 のアミノ酸配列が、前記第 1 のアミノ酸配列を隣接アミノ酸配列から区別する一対のエッジアミノ酸配列を有する、付記 1 に記載の方法。

付記 1 0

配列番号 1 の変異 Cas 9 タンパク質。

10

付記 1 1

配列番号 2 の変異 Cas 9 タンパク質。

付記 1 2

高度には保存されていないか、または低度保存された Cas 9 タンパク質のファミリー内の 1 または複数のアミノ酸配列を同定すること、

前記 1 または複数の同定されたアミノ酸配列に対応する核酸配列を同定すること、

前記 1 または複数の同定されたアミノ酸配列に対応する核酸配列を欠いた、標的 Cas 9 タンパク質のための核酸配列を作製すること、および

前記作製された核酸配列から変異 Cas 9 タンパク質を作製することを含む、変異 Cas 9 タンパク質を作製する方法。

20

付記 1 3

前記標的 Cas 9 タンパク質がヌクレアーゼ活性を有する、付記 1 2 に記載の方法。

付記 1 4

前記標的 Cas 9 タンパク質がニッカーゼである、付記 1 2 に記載の方法。

付記 1 5

前記標的 Cas 9 タンパク質がヌクレアーゼ欠損である、付記 1 2 に記載の方法。

付記 1 6

前記 1 または複数の同定されたアミノ酸配列に対応する核酸配列を、前記標的 Cas 9 タンパク質のための核酸配列から欠失させる、付記 1 2 に記載の方法。

付記 1 7

前記 1 または複数の同定されたアミノ酸配列に対応する核酸配列を、前記標的 Cas 9 タンパク質のための核酸配列から欠失させてリンカーで置換する、付記 1 2 に記載の方法。

30

付記 1 8

前記 1 または複数の同定されたアミノ酸配列に対応する核酸配列を、前記標的 Cas 9 タンパク質のための核酸配列から欠失させて S G G G S リンカーで置換する、付記 1 2 に記載の方法。

付記 1 9

第 1 の Cas 9 タンパク質種の C 末端ドメインを第 2 の Cas 9 タンパク質種の C 末端ドメインで置換することを含む、変異 Cas 9 タンパク質を作製する方法。

40

付記 2 0

前記第 1 の Cas 9 タンパク質種が NM Cas 9 であり、前記第 2 の Cas 9 タンパク質種が ST 1 Cas 9 である、付記 1 9 に記載の方法。

付記 2 1

第 1 の Cas 9 タンパク質種の N 末端ドメインを第 2 の Cas 9 タンパク質種の N 末端ドメインで置換することを含む、キメラ Cas 9 タンパク質を作製する方法。

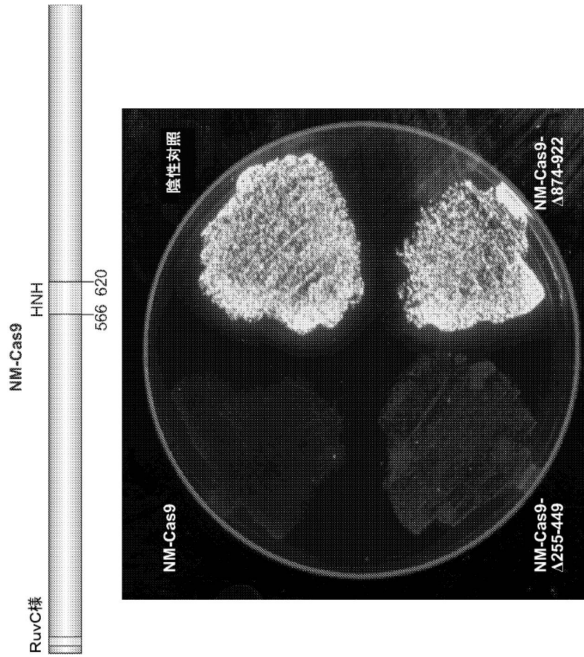
付記 2 2

前記第 1 の Cas 9 タンパク質種が NM Cas 9 であり、前記第 2 の Cas 9 タンバ

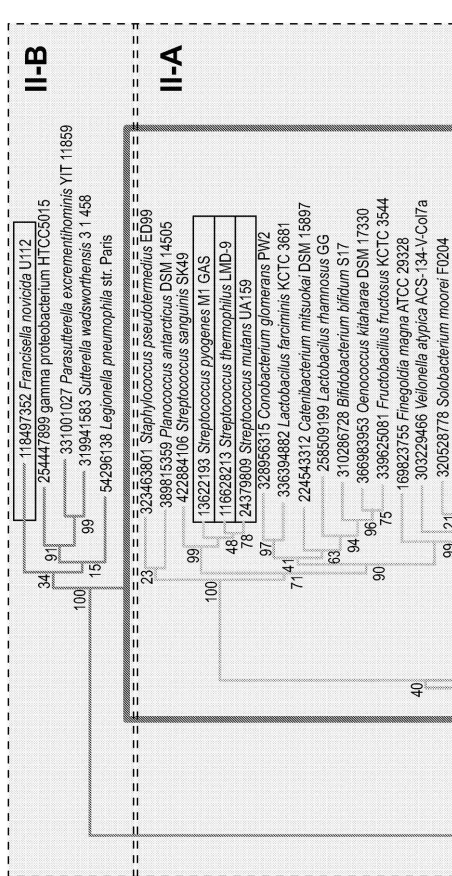
50

ク質種が S T 1 C a s 9 である、付記 2 1 に記載の方法。

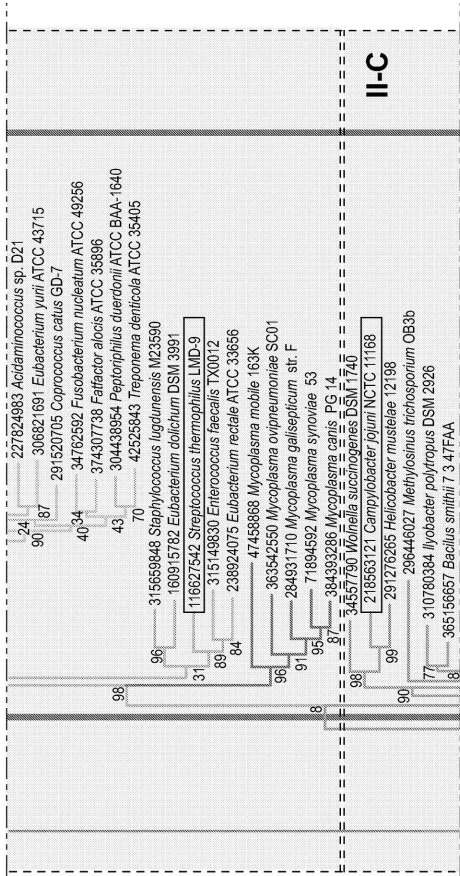
【 図 1 】



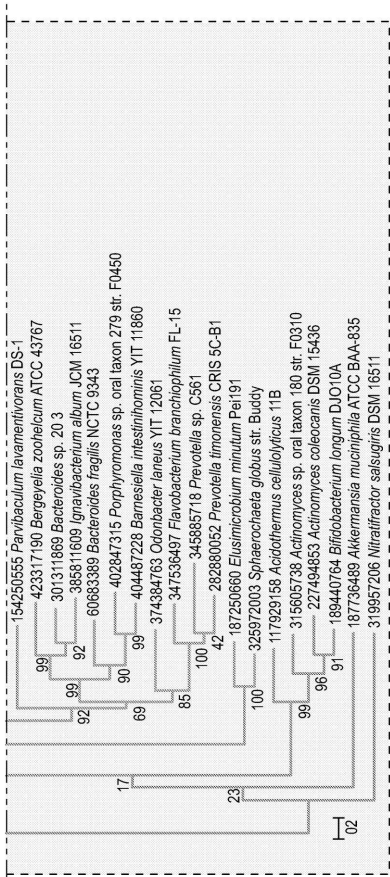
【 図 2 - 1 】



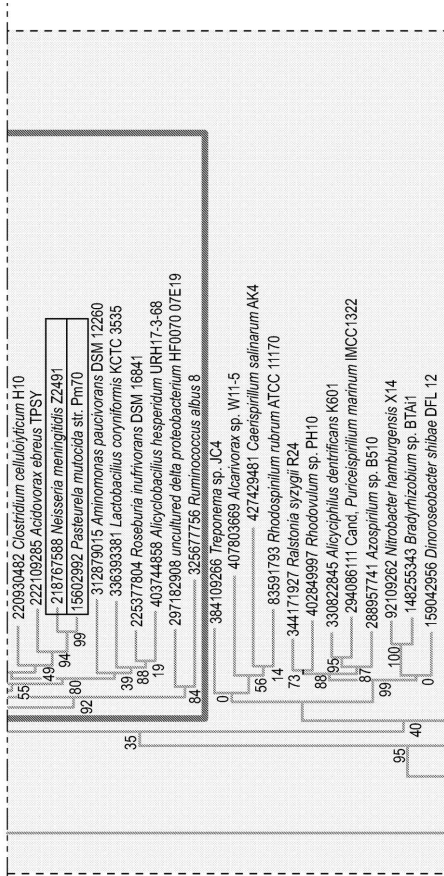
【 2 - 2 】



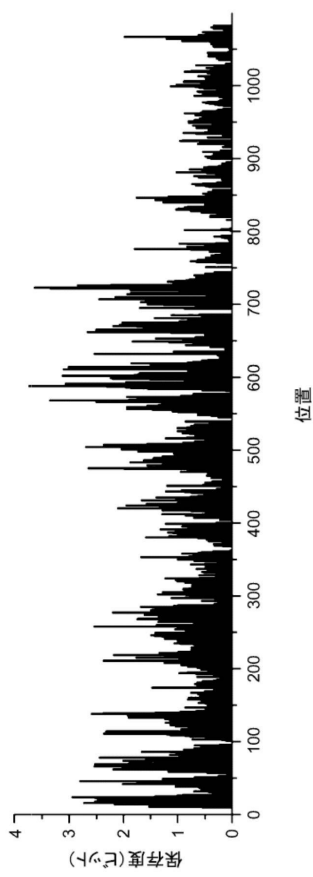
【 2 - 4 】



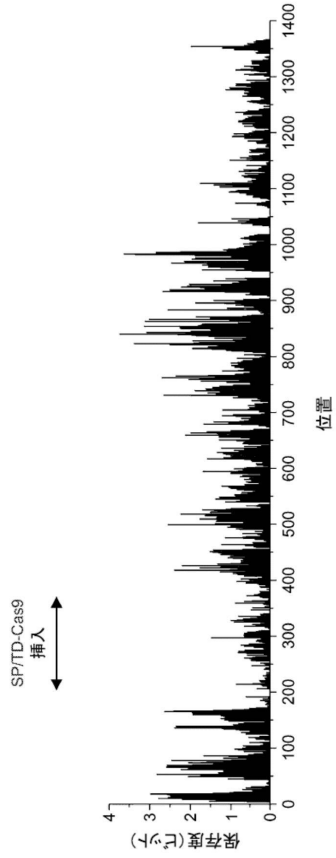
【 2 - 3 】



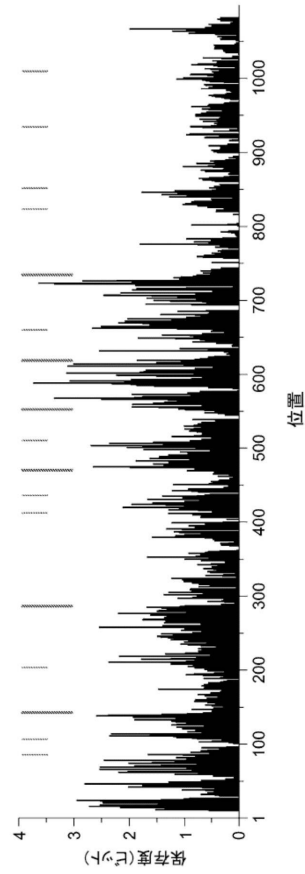
【 3 】



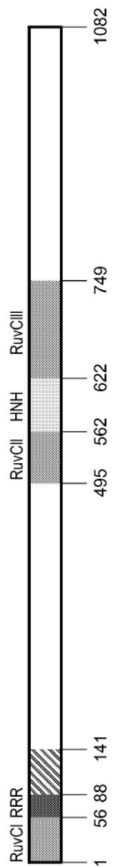
【 図 4 】



【 図 5 A 】



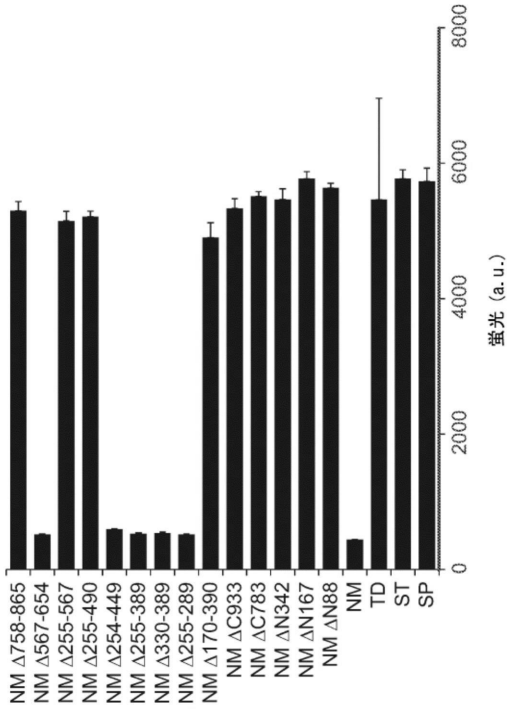
【 図 5 B 】



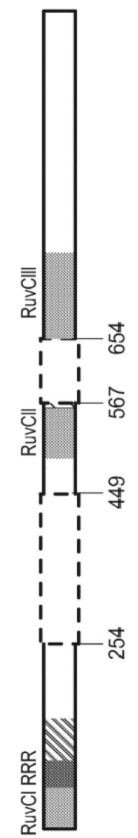
【 図 6 A 】



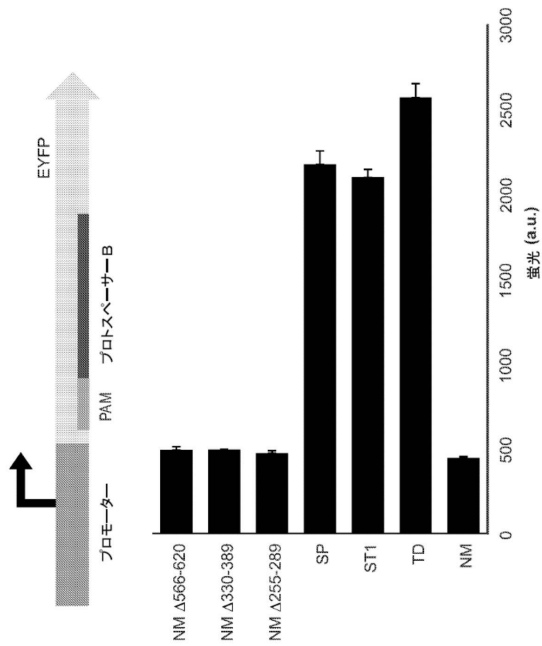
【 図 6 B 】



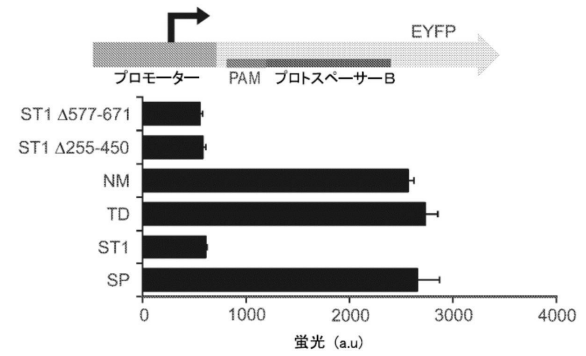
【 図 6 C 】



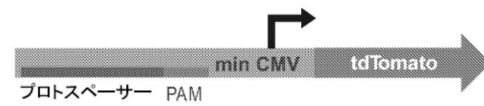
【 図 7 】



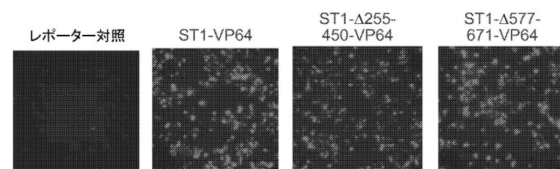
【 図 8 A 】



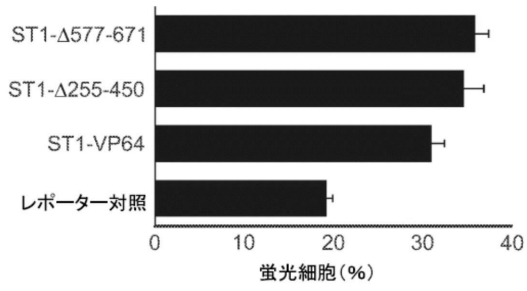
【 図 8 B 】



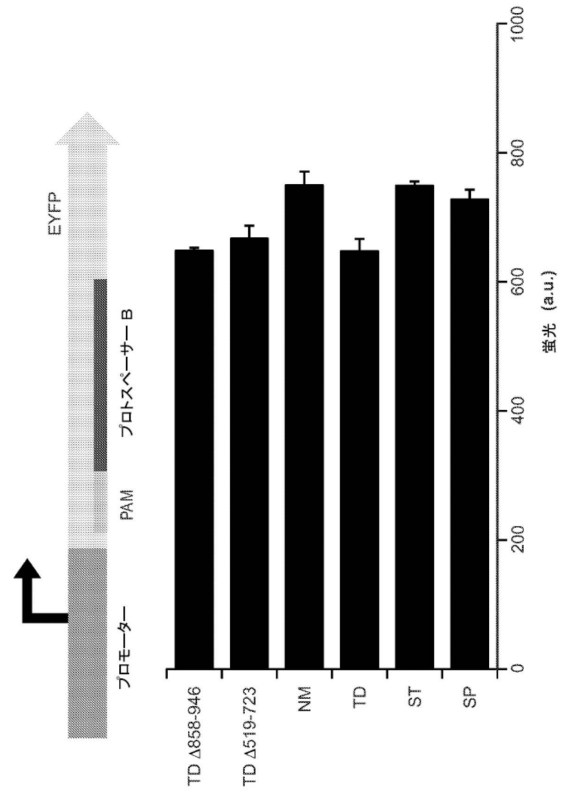
【 図 8 C 】



【図 8 D】



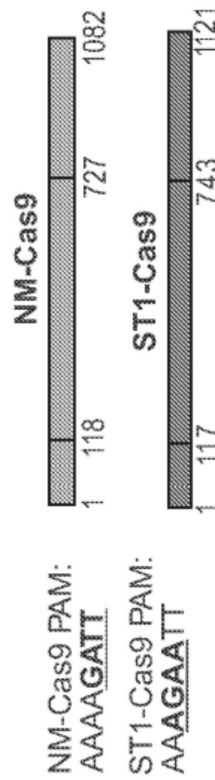
【図 9】



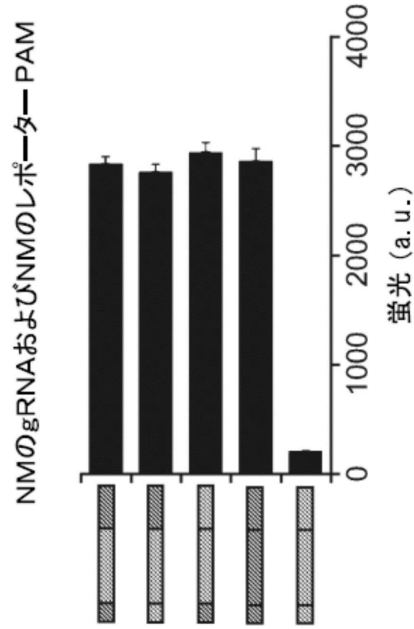
【図 10 A】



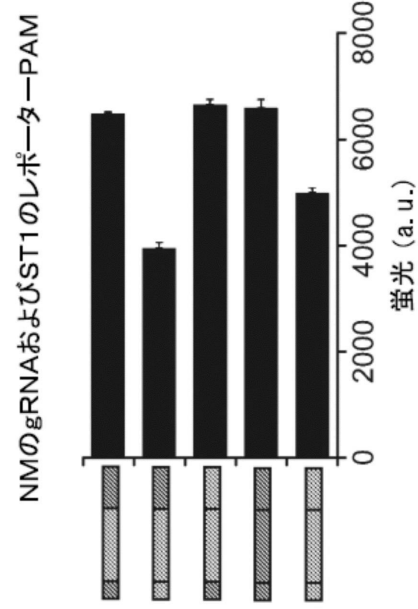
【図 10 B】



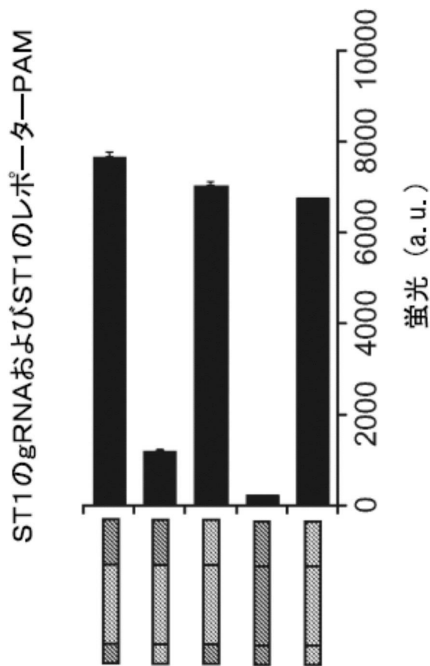
【 図 1 0 C 】



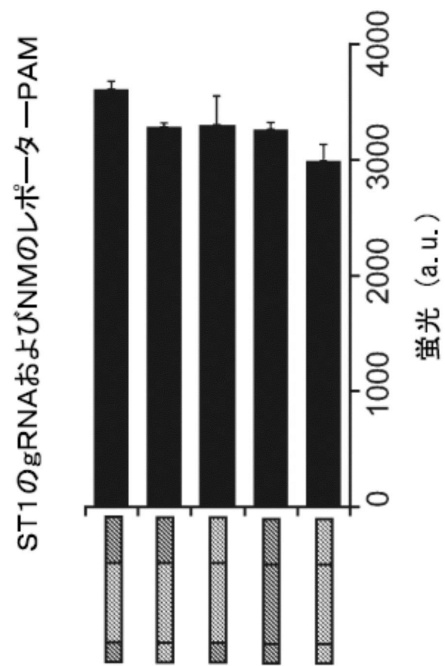
【 図 1 0 D 】



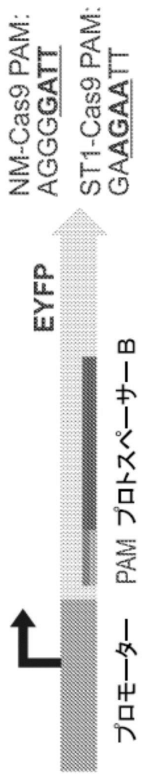
【 図 1 0 E 】



【 図 1 0 F 】



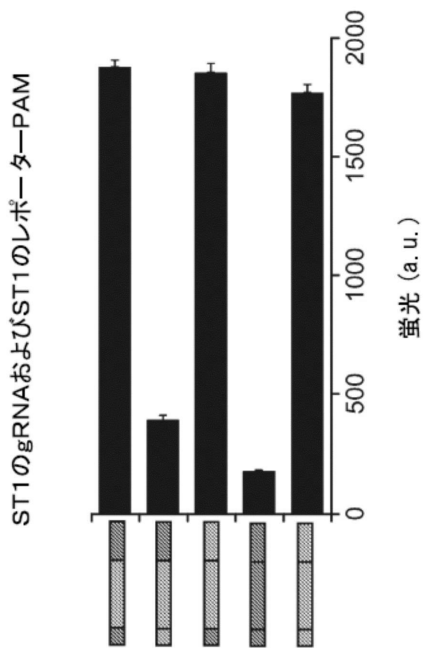
【 図 1 1 A 】



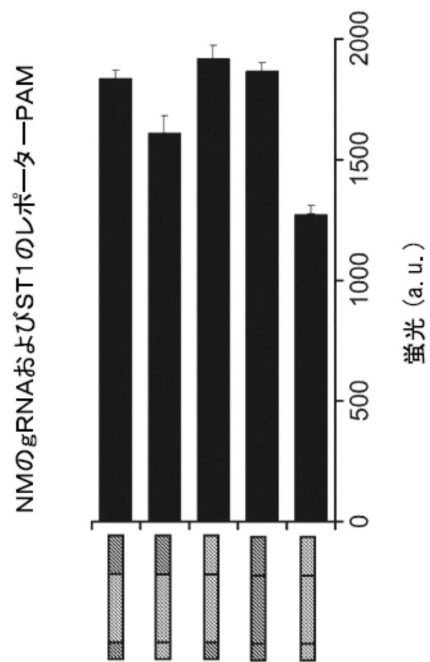
【 図 1 1 B 】



【 図 1 1 C 】



【 図 1 1 D 】



【配列表】

0006984096000001.app

フロントページの続き

- (72)発明者 プールヴィク、 フランク
アメリカ合衆国 75231 テキサス州 ダラス ウォーリング レーン 7482
- (72)発明者 チャーチ、 ジョージ エム.
アメリカ合衆国 02446 マサチューセッツ州 ブルックリン ケント ストリート 218

合議体

審判長 長井 啓子

審判官 松野 広一

審判官 中島 庸子

- (56)参考文献 国際公開第2013/169802(WO, A1)
特表2005-529837(JP, A)
国際公開第2013/141680(WO, A1)
米国特許出願公開第2011/0286980(US, A1)
国際公開第2014/150624(WO, A1)
J. Virol., 1990, 64(2), pp. 613-620
Nucleic Acids Res., 2009, 37(21), pp. 7110-7123
Mol. Cell. Biol., 1996, 16(7), pp. 3866-3877
RNA Biol., 2013, 10(5), pp. 726-737, Published online 2013 Apr 5
Nat. Biotechnol., 2013, 31(3), pp. 230-232, Published online 2013 Jan 29
Nat. Biotechnol., 2013, 31(9), pp. 833-838, Published online 2013 Aug 1

(58)調査した分野(Int.Cl., DB名)

C12N 15/00-15/90

C07K 1/00-19/00

JSTPlus/JMEDPlus/JST7580(JDreamIII)

CAplus/MEDLINE/BIOSIS(STN)

GenBank/EMBL/DDBJ/GeneSeq

UniProt/GeneSeq