

(19)日本国特許庁(JP)

(12)公開特許公報(A)

(11)公開番号

特開2023-30067

(P2023-30067A)

(43)公開日 令和5年3月7日(2023.3.7)

(51)国際特許分類

C 1 2 N 15/09 (2006.01)

F I

C 1 2 N 15/09 1 1 0

C 1 2 N 15/09 Z N A

審査請求 有 請求項の数 1 O L (全138頁)

(21)出願番号 特願2022-202406(P2022-202406)
 (22)出願日 令和4年12月19日(2022.12.19)
 (62)分割の表示 特願2019-517246(P2019-517246)の分割
 原出願日 平成29年9月28日(2017.9.28)
 (31)優先権主張番号 62/402,849
 (32)優先日 平成28年9月30日(2016.9.30)
 (33)優先権主張国・地域又は機関 米国(US)

(71)出願人 506115514
 ザ リージェンツ オブ ザ ユニバーシテ
 イ オブ カリフォルニア
 The Regents of the U
 niversity of Califo
 rnia
 アメリカ合衆国, カリフォルニア州 9
 4 6 0 7 - 5 2 0 0 , オークランド, フ
 ランクリン ストリート 1 1 1 1 , 1 2
 番 フロア
 (74)代理人 100114557
 弁理士 河野 英仁
 (74)代理人 100078868
 弁理士 河野 登夫
 (72)発明者 ダウドナ, ジェニファー エー.
 最終頁に続く

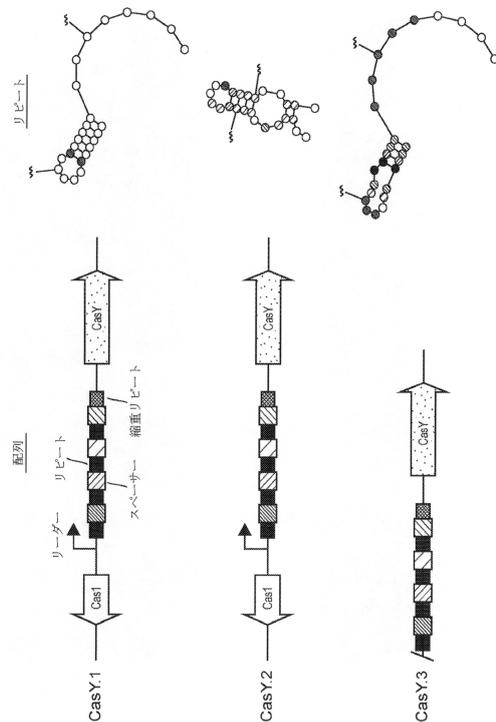
(54)【発明の名称】 RNA誘導型核酸修飾酵素及びその使用方法

(57)【要約】 (修正有)

【課題】クラス2のCRISPR/Casシステムを提供する。

【解決手段】本開示は、CasYタンパク質、CasYタンパク質をコードする核酸、ならびにCasYタンパク質及び/またはそれをコードする核酸を含む改変宿主細胞を提供する。CasYタンパク質は、提供される多様な用途に有用である。本開示は、CasYタンパク質に結合して、CasYタンパク質に配列特異性を与えるCasYガイドRNA、CasYガイドRNAをコードする核酸、ならびにCasYガイドRNA及び/またはそれをコードする核酸を含む改変宿主細胞を提供する。CasYガイドRNAは、提供される多様な用途に有用である。本開示は、CRISPR RNA誘導型エンドヌクレアーゼを同定する方法を提供する。

【選択図】図4 - 1



【特許請求の範囲】

【請求項 1】

組成物であって、
図面に実質的に示されており、明細書に実質的に記載されている、
組成物。

【発明の詳細な説明】

【背景技術】

【0001】

CRISPR-Casシステムは、DNAシーケンシングの時代以前の科学では未知であった経路の一例であるが、今ではファージ及びウイルスに対する獲得免疫を細菌及び古細菌に付与するものであることがわかっている。過去10年間の集中的研究により、このシステムの生化学が明らかになった。

【0002】

CRISPR-Casシステムは、外来DNAまたはRNAの獲得、標的化、及び切断に関与するCasタンパク質と、Casタンパク質をその標的に誘導する短いスペーサー配列に隣接する直列反復配列を含むCRISPR配列からなる。クラス2のCRISPR-Casは簡素化型であり、RNAに結合された単一のCasタンパク質が、標的配列への結合とその切断を担う。

【0003】

このようなプログラム可能な性質をもつ最小限のシステムにより、汎用技術としての使用が可能になったことで、ゲノム操作の分野に変革をもたらされつつある。

【0004】

現在のCRISPR-Cas技術は、培養細菌に由来するシステムに基づいているため、単離されていない大部分の生物は未開発のままである。今のところ、クラス2のCRISPR/Casシステムは、ごくわずかしが発見されていない。さらなるクラス2のCRISPR/Casシステム（例えば、Casタンパク質とガイドRNAとの組み合わせ）が当技術分野で必要とされている。

【発明の概要】

【0005】

本開示は、本明細書で「CasY」ポリペプチド（別称「CasYタンパク質」）と称するRNA誘導型エンドヌクレアーゼポリペプチド；CasYポリペプチドをコードする核酸；ならびにCasYポリペプチド及び/またはそれをコードする核酸を含む改変宿主細胞を提供する。CasYポリペプチドは、提供される多様な用途に有用である。

【0006】

本開示は、CasYタンパク質に結合して、CasYタンパク質に配列特異性を与えるガイドRNA（本明細書で「CasYガイドRNA」とも称する）；CasYガイドRNAをコードする核酸；ならびにCasYガイドRNA及び/またはそれをコードする核酸を含む改変宿主細胞を提供する。CasYガイドRNAは、提供される多様な用途に有用である。

【0007】

本開示は、CRISPR RNA誘導型エンドヌクレアーゼを同定する方法を提供する。

【図面の簡単な説明】

【0008】

【図1-1】天然CasYタンパク質配列の例を示す。

【図1-2】天然CasYタンパク質配列の例を示す。

【図1-3】天然CasYタンパク質配列の例を示す。

【図1-4】天然CasYタンパク質配列の例を示す。

【図2-1】天然CasYタンパク質配列のアライメントを示す。

【図2-2】天然CasYタンパク質配列のアライメントを示す。

【図2-3】天然CasYタンパク質配列のアライメントを示す。

10

20

30

40

50

- 【図 2 - 4】天然 CasY タンパク質配列のアライメントを示す。
- 【図 2 - 5】天然 CasY タンパク質配列のアライメントを示す。
- 【図 2 - 6】天然 CasY タンパク質配列のアライメントを示す。
- 【図 2 - 7】天然 CasY タンパク質配列のアライメントを示す。
- 【図 2 - 8】天然 CasY タンパク質配列のアライメントを示す。
- 【図 2 - 9】天然 CasY タンパク質配列のアライメントを示す。
- 【図 2 - 10】天然 CasY タンパク質配列のアライメントを示す。
- 【図 2 - 11】天然 CasY タンパク質配列のアライメントを示す。
- 【図 2 - 12】天然 CasY タンパク質配列のアライメントを示す。
- 【図 2 - 13】天然 CasY タンパク質配列のアライメントを示す。 10
- 【図 2 - 14】天然 CasY タンパク質配列のアライメントを示す。
- 【図 3 - 1】(パネル A 及び B) CasY ドメインの模式図を示す。CasY のホモログの同定を試行する様々な検索から得た結果も示している。また、同定された CasY を含む CRISPR 遺伝子座の部分も図示している。
- 【図 3 - 2】(パネル A 及び B) CasY ドメインの模式図を示す。CasY のホモログの同定を試行する様々な検索から得た結果も示している。また、同定された CasY を含む CRISPR 遺伝子座の部分も図示している。
- 【図 4 - 1】CasY 及び C2c3 遺伝子座の模式図を示す。干渉タンパク質を緑で、取り込みタンパク質を赤で示している。RNA 構造を利用して折り畳まれるリピートを右側に示す。5' 末端に強度のヘアピンが表れており、CasY による CRISPR 配列の自己プロセッシングを示唆している。 20
- 【図 4 - 2】CasY 及び C2c3 遺伝子座の模式図を示す。干渉タンパク質を緑で、取り込みタンパク質を赤で示している。RNA 構造を利用して折り畳まれるリピートを右側に示す。5' 末端に強度のヘアピンが表れており、CasY による CRISPR 配列の自己プロセッシングを示唆している。
- 【図 4 - 3】CasY 及び C2c3 遺伝子座の模式図を示す。干渉タンパク質を緑で、取り込みタンパク質を赤で示している。RNA 構造を利用して折り畳まれるリピートを右側に示す。5' 末端に強度のヘアピンが表れており、CasY による CRISPR 配列の自己プロセッシングを示唆している。
- 【図 5 - 1】(パネル A ~ D) CasY の PAM 配列を決定するために実施した実験 (CasY による PAM 依存性プラスミド干渉) を示す。 30
- 【図 5 - 2】(パネル A ~ D) CasY の PAM 配列を決定するために実施した実験 (CasY による PAM 依存性プラスミド干渉) を示す。
- 【図 6】(パネル A 及び B) 天然 CasY ガイド RNA の「リピート」配列、及び標的 DNA にハイブリダイズする CasY ガイド RNA の例を表す。(上から下へ、配列番号 11 ~ 15、及び 20)
- 【図 7】(パネル A 及び B) 未培養生物から新規同定された CRISPR - Cas システムを表す。A は、Huguet al. 32 のデータに基づいた、すべての細菌及び古細菌のうち、単離された標本のある主要系統または単離された標本のない主要系統が占める比率である。この結果は、これらのドメインで大部分の生物学がほとんど未研究であることを明らかにしている。古細菌 Cas9 及び新規 CRISPR - CasY は、単離された標本のない系統でのみ見出された。B は、新たに発見された CRISPR - Cas システムの遺伝子座構成である。 40
- 【図 8 - 1】(パネル A 及び B) ARMAN - 1 CRISPR 配列の多様性、及び ARMAN - 1 Cas9 PAM 配列の同定を示す。A は、15 種類の AMD 試料から再構成された CRISPR 配列である。白いボックスはリピートを示し、カラーのひし形はスペーサーを示す(同一のスペーサーは同色であり、一意なスペーサーは黒色である)。配列の保存領域を(右側に)ハイライト表示している。最近獲得されたスペーサー(左側)の多様性は、システムが活動性であることを示す。リードデータからの CRISPR 断片も含む分析を図 14 に示す。B は、AMD メタゲノムデータから再構成した単一の推定ウ 50

イルスコンティグが、ARMAN-1 CRISPR配列からの56のプロトスペーサー（赤い垂直バー）を含むことを示す。

【図8-2】（パネルC）非標的鎖上のプロトスペーサー下流の保存された「NGG」PAMモチーフを明らかにした配列分析を示す。

【図9-1】（パネルA及びB）CasXがE.coliにおいてプログラム可能なDNA干渉を媒介することを示すデータを表す。Aは、CasXプラスミド干渉アッセイの図である。最小CasX遺伝子座を発現するE.coliを、CRISPR配列のスペーサーが一致する配列を含むプラスミド（標的）、または一致しないスペーサーを含むプラスミド（非標的）で形質転換する。形質転換後、培養物をプレーティングし、コロニー形成単位（cfu）を定量する。Bは、スペーサー1（sX.1）を標的とするPlancetomycesのCasX遺伝子座を発現し、指定の標的で形質転換されたE.coliの段階希釈である（sX1、CasXスペーサー1；sX2、CasXスペーサー2；NT、非標的）。

【図9-2】（パネルC及びD）CasXがE.coliにおいてプログラム可能なDNA干渉を媒介することを示すデータを表す。Cは、Deltaproteobacteria CasXによるプラスミド干渉である。実験は3連で実施し、平均±s.d.を示す。Dは、E.coliで発現したPlancetomyces CasX遺伝子座についてのPAM欠失アッセイである。対照ライブラリと比較して30倍超、欠失しているPAM配列を使用して、WebLogoを作成した。

【図10】（パネルA～C）CasXが二本鎖誘導型CRISPR複合体であることを示すデータを表す。Aの下部に、環境RNA配列（メタトランスクリプトームデータ）をCasX CRISPR遺伝子座とマッピングしたものを図示した（赤の矢印、推定tracrRNA；白のボックス、リピート配列；緑のひし形、スペーサー配列）。挿入図は、最初のリピートとスペーサーの詳細図を示す。Bは、CasX二本鎖DNA干渉の図である。RNAプロセシング部位は黒の矢印で示す。Cは、推定tracrRNAをCasX遺伝子座からロックアウトしたプラスミド干渉アッセイの結果である（T、標的；NT、非標的）。実験は3連で実施し、平均±s.d.を示す。

【図11-1】（パネルA）E.coliのCasY遺伝子座の発現が、DNA干渉に十分であることを示すデータを表す。Aは、CasY遺伝子座及び隣接タンパク質の図である。

【図11-2】（パネルB及びC）E.coliのCasY遺伝子座の発現が、DNA干渉に十分であることを示すデータを表す。Bは、対照ライブラリと比較してCasYによって3倍超、欠失している5' PAM配列のWebLogoである。Cは、CasY.1を発現し、指定のPAMを含む標的で形質転換されたE.coliによるプラスミド干渉である。実験は3連で実施し、平均±s.d.を示す。

【図12】（パネルA及びB）既知のシステムに関連して新たに同定されたCRISPR-Casを表す。Aは、一般的なCas1タンパク質の簡略化した系統発生樹である。CRISPR型の既知のシステムはくさび形及び分岐で示し、新たに記載されるシステムは太字で示す。詳細なCas1の系統は、補足データ2に示す。Bは、II-B型遺伝子座とII-C型遺伝子座との組換えの結果、古細菌II型システムに生じた進化シナリオを提案したものである。

【図13-1】ARMAN-4からの古細菌Cas9が、縮重CRISPR配列をもつ多数のコンティグに見出されることを示す。16の異なるコンティグに関して、ARMAN-4からのCas9を暗赤色でハイライト表示している。推定ドメインまたは機能をもつタンパク質は標識されており、仮想タンパク質は非標識である。15のコンティグは、2つの縮重直列リピート（1bpのミスマッチ）と、単一の保存スペーサーを含む。残りのコンティグは、直列リピートを1つのみ含む。ARMAN-1とは異なり、ARMAN-4ではCas9に隣接する追加のCasタンパク質は見られない。

【図13-2】ARMAN-4からの古細菌Cas9が、縮重CRISPR配列をもつ多数のコンティグに見出されることを示す。16の異なるコンティグに関して、ARMAN

10

20

30

40

50

- 4からのCas9を暗赤色でハイライト表示している。推定ドメインまたは機能をもつタンパク質は標識されており、仮想タンパク質は非標識である。15のコンティグは、2つの縮重直列リピート(1bpのミスマッチ)と、単一の保存スパーサーを含む。残りのコンティグは、直列リピートを1つのみ含む。ARMAN-1とは異なり、ARMAN-4ではCas9に隣接する追加のCasタンパク質は見られない。

【図13-3】ARMAN-4からの古細菌Cas9が、縮重CRISPR配列をもつ多数のコンティグに見出されることを示す。16の異なるコンティグに関して、ARMAN-4からのCas9を暗赤色でハイライト表示している。推定ドメインまたは機能をもつタンパク質は標識されており、仮想タンパク質は非標識である。15のコンティグは、2つの縮重直列リピート(1bpのミスマッチ)と、単一の保存スパーサーを含む。残りのコンティグは、直列リピートを1つのみ含む。ARMAN-1とは異なり、ARMAN-4ではCas9に隣接する追加のCasタンパク質は見られない。

10

【図13-4】ARMAN-4からの古細菌Cas9が、縮重CRISPR配列をもつ多数のコンティグに見出されることを示す。16の異なるコンティグに関して、ARMAN-4からのCas9を暗赤色でハイライト表示している。推定ドメインまたは機能をもつタンパク質は標識されており、仮想タンパク質は非標識である。15のコンティグは、2つの縮重直列リピート(1bpのミスマッチ)と、単一の保存スパーサーを含む。残りのコンティグは、直列リピートを1つのみ含む。ARMAN-1とは異なり、ARMAN-4ではCas9に隣接する追加のCasタンパク質は見られない。

【図14-1】ARMAN-1 CRISPR配列の再構成全体を表す。CRISPR配列の再構成には、参照アセンブル配列、ならびに短いDNAリードから再構成された配列セグメントを含む。緑の矢印はリピートを示し、カラーの矢印はCRISPRスパーサーを示す(同一のスパーサーは同色であり、一意なスパーサーは黒色である)。CRISPRシステムでは、スパーサーは通常、一方向に追加されるので、左側のスパーサーの不一致が大きいのは、最近獲得されたことに起因する。

20

【図14-2】ARMAN-1 CRISPR配列の再構成全体を表す。CRISPR配列の再構成には、参照アセンブル配列、ならびに短いDNAリードから再構成された配列セグメントを含む。緑の矢印はリピートを示し、カラーの矢印はCRISPRスパーサーを示す(同一のスパーサーは同色であり、一意なスパーサーは黒色である)。CRISPRシステムでは、スパーサーは通常、一方向に追加されるので、左側のスパーサーの不一致が大きいのは、最近獲得されたことに起因する。

30

【図14-3】ARMAN-1 CRISPR配列の再構成全体を表す。CRISPR配列の再構成には、参照アセンブル配列、ならびに短いDNAリードから再構成された配列セグメントを含む。緑の矢印はリピートを示し、カラーの矢印はCRISPRスパーサーを示す(同一のスパーサーは同色であり、一意なスパーサーは黒色である)。CRISPRシステムでは、スパーサーは通常、一方向に追加されるので、左側のスパーサーの不一致が大きいのは、最近獲得されたことに起因する。

【図14-4】ARMAN-1 CRISPR配列の再構成全体を表す。CRISPR配列の再構成には、参照アセンブル配列、ならびに短いDNAリードから再構成された配列セグメントを含む。緑の矢印はリピートを示し、カラーの矢印はCRISPRスパーサーを示す(同一のスパーサーは同色であり、一意なスパーサーは黒色である)。CRISPRシステムでは、スパーサーは通常、一方向に追加されるので、左側のスパーサーの不一致が大きいのは、最近獲得されたことに起因する。

40

【図14-5】ARMAN-1 CRISPR配列の再構成全体を表す。CRISPR配列の再構成には、参照アセンブル配列、ならびに短いDNAリードから再構成された配列セグメントを含む。緑の矢印はリピートを示し、カラーの矢印はCRISPRスパーサーを示す(同一のスパーサーは同色であり、一意なスパーサーは黒色である)。CRISPRシステムでは、スパーサーは通常、一方向に追加されるので、左側のスパーサーの不一致が大きいのは、最近獲得されたことに起因する。

【図14-6】ARMAN-1 CRISPR配列の再構成全体を表す。CRISPR配

50

列の再構成には、参照アセンブル配列、ならびに短いDNAリードから再構成された配列セグメントを含む。緑の矢印はリピートを示し、カラーの矢印はCRISPRスペーサーを示す（同一のスペーサーは同色であり、一意なスペーサーは黒色である）。CRISPRシステムでは、スペーサーは通常、一方向に追加されるので、左側のスペーサーの不一致が大きいのは、最近獲得されたことに起因する。

【図15】（パネルA及びB）ARMAN-1スペーサーと、古細菌群集のメンバーのゲノムとのマッピングを示す。Aは、ARMAN-1からのプロトスペーサー（赤の矢印）と、同じ環境からのナノ古細菌ARMAN-2のゲノムとのマッピングである。6つのプロトスペーサーは、2つの長末端反復配列（LTR）が隣接するゲノムの一部に一意にマッピングされ、さらに2つの別のプロトスペーサーがLTR内で完全に一致している（青及び緑）。この領域はトランスポゾンである可能性が高く、これは、ARMAN-1のCRISPR-Casシステムがこの要素の動員を抑制する役割を果たしていることを示唆する。Bでは、ARMAN生物と同じ試料に見られる、リッチモンド鉱山生態系のもう1つのメンバーであるThermoplasmales古細菌（I-プラズマ）ともプロトスペーサーをマッピングしている。プロトスペーサーは、短い仮想タンパク質をコードするゲノムの領域内に集合しており、これが流動要素を表し得ることも示唆している。

10

【図16-1】（パネルA）ARMAN-1 crRNA及びtracrRNAの予測される二次構造を表す。Aは、CRISPRのリピート及びtracrRNAアンチリピートを黒で図示し、スペーサー由来の配列を一連の緑のNで示している。遺伝子座からは明確な終結シグナルを予測できないため、二次構造に基づいて、3つの異なるtracrRNA長を試験した（69、104、及び179、それぞれ赤、青、及びピンク）。

20

【図16-2】（パネルB）ARMAN-1 crRNA及びtracrRNAの予測される二次構造を表す。Bは、Aにおける二本鎖ガイドに対応する遺伝子操作された一本鎖ガイドRNAである。

【図16-3】（パネルC～E）ARMAN-1 crRNA及びtracrRNAの予測される二次構造を表す。Cは、tracrRNAの3'末端に2つの異なるヘアピン（75及び122）をもつARMAN-4 Cas9の二本鎖ガイドである。Dは、Cにおける二本鎖ガイドに対応する遺伝子操作された一本鎖ガイドRNAである。Eは、E. coliのin vivo標的アッセイの試験条件である。

30

【図17】（パネルA及びB）in vitroでの生化学研究のための精製スキーマを表す。Aは、補足資料に概説されている様々な条件下で、ARMAN-1（AR1）及びARMAN-4（AR4）Cas9を発現し、精製したことを示す。青いボックスで囲まれたタンパク質のin vitroでの切断活性について試験した。Bは、AR1-Cas9及びAR4-Cas9精製物の画分を10% SDS-PAGEゲル上で分離したものを示す。

【図18】新たに同定されたCRISPR-Casシステムと既知のタンパク質との比較を表す。（1）NCBIの非重複（NR）タンパク質に対するBLAST検索、（2）すべての既知のタンパク質のHMMデータベースに対する、隠れマルコフモデル（HMM）検索、及び（3）HHpred³⁰を使用した遠隔相同性検索に基づいた、既知のタンパク質に対するCas X及びCas Yの類似性を示す。

40

【図19-1】（パネルA及びB）Cas XによってプログラムされたDNA干渉に関連するデータを表す。Aは、図9のパネルCから続く、Cas X2（Planctomyces）及びCas X1（Deltaproteobacteria）についてのプラスミド干渉アッセイである（sX1、Cas Xスペーサー1；sX2、Cas Xスペーサー2；NT、非標的）。実験は3連で実施し、平均±s.d.を示す。Bは、図9のパネルBから続く、Cas X遺伝子座を発現し、指定された標的で形質転換された、E. coliの段階希釈である。

【図19-2】（パネルC）Cas XによってプログラムされたDNA干渉に関連するデータを表す。Cは、E. coliで発現したDeltaproteobacteria

50

Cas X についての PAM 欠失アッセイである。対照ライブラリと比較して、示された PAM 欠失閾値 (PDVT) を超えて欠失している PAM 配列を使用して、Web Logo を作成した。

【図 19 - 3】(パネル D) Cas X によってプログラムされた DNA 干渉に関連するデータを表す。D は、E. coli で発現した *Planctomyces Cas X* についての PAM 欠失アッセイである。対照ライブラリと比較して、示された PAM 欠失閾値 (PDVT) を超えて欠失している PAM 配列を使用して、Web Logo を作成した。

【図 20】Cas 9 ホモログの進化樹を表す。以前に記載されているシステムを示す、Cas 9 タンパク質の最尤系統樹。各型に基づいて、II - A は青、II - B は緑、及び II - C は紫に色分けされている。II - C 型 CRISPR - Cas システムに、新たに記載された未培養細菌からの 2 つの細菌 Cas 9 を加えて群化した古細菌 Cas 9 である。

【図 21】ARMAN - 1 及び ARMAN - 4 からの Cas 9 についてアッセイした、切断条件の表を表す。

【発明を実施するための形態】

【0009】

本明細書で使用される場合、「異種」とは、ヌクレオチドまたはポリペプチドの配列が、それぞれ天然の核酸またはタンパク質に見られないものであることを意味する。例えば、Cas Y ポリペプチドに関して、異種ポリペプチドは、Cas Y ポリペプチド以外のタンパク質からのアミノ酸配列を含む。場合によって、1 つの種由来の Cas Y タンパク質の一部が、異なる種由来の Cas Y タンパク質の一部に融合されている。そのため、それぞれの種由来の Cas Y 配列は、互いに対して異種であるとみなすことができる。別の例として、Cas Y タンパク質 (例えば、dCas Y タンパク質) を非 Cas Y タンパク質 (例えば、ヒストンデアセチラーゼ) 由来の活性ドメインに融合させることができ、その活性ドメインの配列は、異種ポリペプチド (Cas Y タンパク質に対して異種である) とみなすことができる。

【0010】

本明細書で同義に使用される用語「ポリヌクレオチド」及び「核酸」は、重合形態の、長さを問わないヌクレオチド、すなわちリボヌクレオチドまたはデオキシヌクレオチドを指す。したがって、この用語には、一本鎖、二本鎖、もしくは多重鎖の DNA もしくは RNA、ゲノム DNA、cDNA、DNA - RNA ハイブリッド、またはプリン及びピリミジン塩基、もしくは他の天然ヌクレオチド塩基、化学的もしくは生化学的に修飾されたヌクレオチド塩基、非天然のヌクレオチド塩基、もしくは誘導体化ヌクレオチド塩基を含む重合体が含まれるが、それらに限定されない。用語「ポリヌクレオチド」及び「核酸」は、記載される実施形態に適用可能である場合、一本鎖 (例えば、センスまたはアンチセンス) ポリヌクレオチド及び二本鎖ポリヌクレオチドを含むと理解されるべきである。

【0011】

本明細書で同義に使用される用語「ポリペプチド」、「ペプチド」、及び「タンパク質」は、任意の長さのアミノ酸の重合形態を指し、これには、遺伝的にコードされた及び非遺伝的にコードされたアミノ酸、化学的にまたは生化学的に修飾または誘導体化されたアミノ酸、ならびに修飾ペプチド骨格を有するポリペプチドを含み得る。本用語には、限定されるものではないが、異種アミノ酸配列を有する融合タンパク質、異種及び同種のリーダー配列を有する融合体、N 末端メチオニン残基を有するまたは有しない融合体、免疫標識されたタンパク質などを含む、融合タンパク質が含まれる。

【0012】

核酸、タンパク質、細胞、または生物に適用される場合、本明細書で使用される用語「天然の」は、天然に存在する核酸、細胞、タンパク質、または生物を指す。

【0013】

本明細書で使用される場合、用語「単離された」とは、ポリヌクレオチド、ポリペプチド、または細胞が、そのポリヌクレオチド、ポリペプチド、または細胞が天然に存在する

10

20

30

40

50

場合とは異なる環境に存在することを意味する。単離された遺伝子改変宿主細胞は、遺伝子改変宿主細胞の混合集団に存在する場合がある。

【0014】

本明細書で使用される場合、用語「外因性核酸」とは、自然界で得られる細菌、生物、もしくは細胞に通常は見られないもしくは天然で見られない核酸、及び/またはそれらによって産生されない核酸を指す。本明細書で使用される場合、用語「内因性核酸」とは、自然界で得られる細菌、生物、もしくは細胞に通常見られる核酸、及び/またはそれらによって産生される核酸を指す。「内因性核酸」は「天然の核酸」とも呼ばれる、所与の細菌、生物、または細胞に「天然の」核酸である。

【0015】

本明細書で使用される場合、「組換え」とは、特定の核酸(DNAまたはRNA)が、天然の系に見られる内因性核酸とは区別できる構造コード配列または非コード配列を有する構築物を生じさせるクローニング、制限酵素による切断、及び/またはライゲーションといった工程の様々な組み合わせの産物であることを意味する。一般に、構造コード配列をコードするDNA配列を、cDNA断片及び短いオリゴヌクレオチドリンカーから、または一連の合成オリゴヌクレオチドからアセンブルして合成核酸を提供し、これを細胞内または無細胞の転写系及び翻訳系に含まれる組換え転写単位から発現させることができる。そのような配列は、真核生物の遺伝子に通常存在する内部非翻訳配列、すなわちイントロンによって中断されないオープンリーディングフレームの形態で提供することができる。関連配列を含むゲノムDNAはまた、組換え遺伝子または転写単位の形成に使用することもできる。非翻訳DNAの配列は、オープンリーディングフレームの5'側に存在しても、または3'側に存在してもよく、そのような配列は、コード領域の操作または発現を妨害せず、かつ、事実上、様々な機構によって望ましい産物の産生を調節するように作用することができる(以下の「DNA制御配列」を参照)。

【0016】

したがって、例えば、用語「組換え」ポリヌクレオチドまたは「組換え」核酸とは、天然に存在しないもの、例えば、本来は2つに分離されている配列セグメントを、人為的介入を経て人工的に組み合わせることによって作製されたものを指す。この人工的な組み合わせは、化学合成手段によって、または核酸の単離セグメントの人工的な操作、例えば遺伝子工学技術のいずれかによって行われる場合が多い。それらは通常、配列認識部位を導入または除去すると同時に、同一アミノ酸または保存的アミノ酸をコードする縮重コドンでコドン置換して行われる。あるいは、所望する機能の核酸セグメントを一つに連結し、所望する機能の組み合わせを生じさせることで実行される。この人工的な組み合わせは、化学合成手段によって、または核酸の単離セグメントの人工的な操作、例えば遺伝子工学技術のいずれかによって行われる場合が多い。

【0017】

同様に、用語「組換え」ポリペプチドとは、天然に存在しないポリペプチド、例えば、本来は2つに分離されているアミノ配列セグメントを、人為的介入を経て人工的に組み合わせることによって作製されたポリペプチドを指す。したがって、例えば、異種アミノ酸配列を含むポリペプチドは組換えである。

【0018】

「構築物」または「ベクター」とは、特定のヌクレオチド配列(複数可)の発現及び/または伝播を目的として生成された、または他の組換えヌクレオチド配列の構築に使用することを意図する組換え核酸、一般には組換えDNAを意味する。

【0019】

本明細書で同義に使用される用語「DNA制御配列」、「制御要素」、及び「調節要素」は、宿主細胞において、コード配列の発現及び/またはコードされたポリペプチドの産生をもたらす及び/または調節するようなプロモーター、エンハンサー、ポリアダニル化シグナル、ターミネーター、タンパク質分解シグナルなどの転写及び翻訳制御配列を指す。

10

20

30

40

50

【0020】

用語「形質転換」は、本明細書において「遺伝子改変」と同義に使用され、細胞への新たな核酸（例えば、細胞に対して外因性のDNA）の導入後に誘導される恒久的または一過性の遺伝的变化を指す。遺伝的变化（「改変」）は、宿主細胞のゲノムへの新たな核酸の取り込みによって、または新たな核酸のエピソーム要素としての一過性の保持もしくは安定保持のいずれかによってなされ得る。細胞が真核細胞である場合、恒久的な遺伝的变化は、一般に、細胞のゲノムへの新たなDNAの導入によってなされる。原核細胞の場合、恒久的な変化が染色体に導入されることもあれば、またはプラスミド及び発現ベクターなどの染色体外要素を介して導入されることもあり、染色体外要素には、組換え宿主細胞での保持に利用される1つ以上の選択マーカを含む場合がある。遺伝子改変に適した方法として、ウイルス感染、トランスフェクション、コンジュゲート、プロトプラスト融合、エレクトロポレーション、パーティクルガン技術、リン酸カルシウム沈降、直接マイクロインジェクションなどが挙げられる。方法の選択は一般に、形質転換される細胞の種類及び形質転換が行われる状況（すなわち、*in vitro*、*ex vivo*、または*in vivo*）に応じて異なる。これらの方法の一般的な考察は、Ausubel, et al, Short Protocols in Molecular Biology, 3rd ed., Wiley & Sons, 1995で参照することができる。

10

【0021】

「機能的に連結された」とは、そのように記載される構成要素が意図された方法で機能できるような関係にある並置を指す。例えば、プロモーターがコード配列の転写または発現に影響を及ぼす場合、プロモーターはコード配列と機能的に連結されている。本明細書で使用される場合、用語「異種プロモーター」及び「異種制御領域」は、自然界において特定の核酸と通常は関連しないプロモーター及び他の制御領域を指す。例えば、「コード領域と異種である転写制御領域」とは、自然界においてコード領域と通常は関連しない転写制御領域である。

20

【0022】

本明細書で使用される場合、「宿主細胞」とは、*in vivo*もしくは*in vitro*の真核細胞、原核細胞、または真核細胞でも原核細胞でもよい単細胞実体として培養された多細胞生物由来の細胞（例えば、細胞株）、もしくは核酸（例えば、発現ベクター）のレシピエントとして使用され、核酸によって遺伝子改変された元の細胞の子孫を含んでいる多細胞生物由来の細胞を意味する。単一細胞の子孫は、天然の、偶発的な、または意図的な変異のため、モルホロジーまたはゲノムもしくは全DNA相補性が、必ずしも元の親と完全に同一でなくてもよいものと理解される。「組換え宿主細胞」（別称「遺伝子改変宿主細胞」）は、異種核酸、例えば発現ベクターを導入された宿主細胞である。例えば、本発明の原核宿主細胞は、例えば原核宿主細胞に対して外来性である（自然界では通常見られない）外因性核酸、または原核宿主細胞には通常見られない組換え核酸といった異種核酸を適切な原核宿主細胞に導入することによって遺伝子改変された原核宿主細胞（例えば、細菌）であり、本発明の真核宿主細胞は、例えば真核宿主細胞に対して外来性である外因性核酸、または真核宿主細胞には通常見られない組換え核酸といった異種核酸を適切な真核宿主細胞に導入することによって遺伝子改変された真核宿主細胞である。

30

40

【0023】

用語「保存的アミノ酸置換」とは、類似する側鎖を有するアミノ酸残基のタンパク質と互換性であることを意味する。例えば、脂肪族側鎖を有するアミノ酸の群は、グリシン、アラニン、バリン、ロイシン、及びイソロイシンからなり；脂肪族ヒドロキシル側鎖を有するアミノ酸の群は、セリン及びスレオニンからなり；アミドを含む側鎖を有するアミノ酸の群は、アスパラギン及びグルタミンからなり；芳香族側鎖を有するアミノ酸の群は、フェニルアラニン、チロシン、及びトリプトファンからなり；塩基性側鎖を有するアミノ酸の群は、リジン、アルギニン、及びヒスチジンからなり；ならびに硫黄を含む側鎖を有するアミノ酸の群は、システイン及びメチオニンからなる。例示的な保存的アミノ酸置換群は、バリン - ロイシン - イソロイシン、フェニルアラニン - チロシン、リジン - アルギ

50

ニン、アラニン - バリン、及びアスパラギン - グルタミンである。

【0024】

ポリヌクレオチドまたはポリペプチドは、別のポリヌクレオチドまたはポリペプチドに対して一定の比率の「配列同一性」を有する。これは、アライメントしたとき、塩基またはアミノ酸の比率が同じであり、2つの配列を比較すると、同一の相対位置にあることを意味する。配列類似性はいくつかの異なる方法で決定することができる。配列同一性を決定するには、ncbi.nlm.nih.gov/BLASTでワールドワイドウェブ経由で利用可能であるBLASTを含む方法及びコンピュータープログラムを使用して配列をアライメントすることができる。例えば、Altschul et al. (1990), J. Mol. Biol. 215: 403 - 10を参照のこと。別のアライメントアルゴリズムは、Oxford Molecular Group, Incの完全子会社であるGenetics Computing Group (GCG) (所在地Madison, Wisconsin, USA)のパッケージで利用可能なFASTAである。アライメントのための他の技術については、Methods in Enzymology, vol. 266: Computer Methods for Macromolecular Sequence Analysis (1996), ed. Doolittle, Academic Press, Inc. (Harcourt Brace & Co., San Diego, California, USAの一部門)に記載されている。特に興味深いのは、配列中のギャップを許容するアライメントプログラムである。Smith-Watermanは、配列アライメント中のギャップを許容するアルゴリズムの一種である。Meth. Mol. Biol. 70: 173 - 187 (1997)を参照のこと。また、Needleman及びWunschのアライメント法を使用するGAPプログラムを、配列のアライメントに利用することができる。J. Mol. Biol. 48: 443 - 453 (1970)を参照のこと。

【0025】

本明細書で使用される場合、「治療」、「治療する」などの用語は、望ましい薬学的効果及び/または生理学的効果を得ることを指す。効果は、疾患もしくはその症状を完全にもしくは部分的に予防する点では予防的であり得、及び/または、疾患及び/または疾患に起因する有害作用の部分的もしくは完全な治癒という点では治療的であり得る。本明細書で使用される場合、「治療」は、哺乳動物、例えばヒトの疾患のいずれかの治療を包含し、それには、(a)疾患の素因を有し得るが、まだそれを有すると診断されていない対象に疾患が生じることを予防すること；(b)疾患を抑制する、すなわちその発生を抑止すること；及び(c)疾患を緩和する、すなわち疾患の退縮を引き起こすことを含む。

【0026】

本明細書で同義に使用される用語「個体」、「対象」、「宿主」、及び「患者」は、個々の生物、例えば、マウス、サル、ヒト、哺乳動物の家畜、哺乳動物の競技用動物、及び哺乳動物のペットを含むが、それらに限定されない哺乳動物を指す。

【0027】

本発明をさらに記載するにあたり、本発明は記載される特定の実施形態に限定されず、当然ながら、これらを変更できるものと理解されるべきである。また、本発明の範囲は添付の特許請求の範囲によってのみ限定されるため、本明細書で使用される用語は特定の実施形態の記載のみを目的としており、限定を意図しないものと理解されるべきである。

【0028】

ある範囲の数値が与えられるとき、文脈から特に明示されない限り、この範囲の上限と下限との間の各中間値、ならびにこの記載された範囲内の他のいずれの記載値または中間値も下限の単位の10分の1まで本発明に包含されるものと理解されるべきである。これよりも狭い範囲の上限と下限は、その狭い範囲に独立して包含されてもよく、また記載された範囲に何らかの具体的な除外制限があることを条件として本発明内に包含される。記載された範囲が上限と下限の一方または両方を含む場合、その含まれる上限下限の一方または両方を除外した範囲もまた本発明の範囲内に包含される。

10

20

30

40

50

【 0 0 2 9 】

特に定義しない限り、本明細書で使用される技術用語及び科学用語はすべて、本発明が属する技術分野の当業者に共通して理解されているものと同じ意味を有する。好ましい方法及び材料を後述するが、本明細書に記載するものと同様または同等であるいかなる方法及び材料も本発明の実施または試験に使用することができる。本明細書で言及する全ての刊行物は、刊行物の引用箇所と関連する方法及び/または材料の開示及び記載について参照することにより本明細書に組み込まれる。

【 0 0 3 0 】

本明細書及び添付の特許請求の範囲で使用される場合、単数形「a」、「an」、及び「the」は、文脈によって特に明示されない限り、複数の指示対象物を包含することに留意されたい。したがって、例えば「CasYポリペプチド」に関する言及は、複数のこのようなポリペプチドを含み、「ガイドRNA」に関する言及は、当業者に公知である1つ以上のガイドRNA及びその等価物に関する言及を含むなどである。さらに、任意の構成要素を除外するように特許請求の範囲を起草することにも留意されたい。そのため、本記載は、特許請求の範囲の構成要素の列挙に関連して、「単独で」、「のみ」などのような排他的用語を使用する際、または「否定的」限定を使用する際の前提基準としての役割を果たすことを意図とする。

【 0 0 3 1 】

明確性を目的として個々の実施形態に関連して記載される本発明の特定の特徴は、組み合わせることで単一の実施形態としても提供することができる。逆に、簡潔性を目的として単一の実施形態に関連して記載される本発明の種々の特徴は、個別に提供することも、または任意の好適なサブコンビネーションで提供することもできる。本発明と関連する実施形態のあらゆる組み合わせは、本発明に明確に包含され、どの組み合わせもすべて個別に及び明示的に開示されているかのように本明細書に開示される。加えて、様々な実施形態及びその要素のサブコンビネーションもすべて本発明に明確に包含され、そのような、どのサブコンビネーションもすべて個別に及び明示的に本明細書に開示されているかのように本明細書に開示される。

【 0 0 3 2 】

本明細書で考察する刊行物は、本出願の出願日に先立つその開示のためにのみ提供される。本明細書のいかなる内容も、本発明が、先行発明を理由としてこのような刊行物に先行する権利を与えられないことを容認するものとはみなされない。さらに、記載された刊行物の日付は実際の公開日と異なる場合があり、それぞれ確認を要する場合がある。

【 0 0 3 3 】

本開示は、本明細書で「CasY」ポリペプチド（別称「CasYタンパク質」）と称するRNA誘導型エンドヌクレアーゼポリペプチド；CasYポリペプチドをコードする核酸；ならびにCasYポリペプチド及び/またはそれをコードする核酸を含む改変宿主細胞を提供する。CasYポリペプチドは、提供される多様な用途に有用である。

【 0 0 3 4 】

本開示は、CasYタンパク質に結合して、CasYタンパク質に配列特異性を与えるガイドRNA（本明細書で「CasYガイドRNA」とも称する）；CasYガイドRNAをコードする核酸；ならびにCasYガイドRNA及び/またはそれをコードする核酸を含む改変宿主細胞を提供する。CasYガイドRNAは、提供される多様な用途に有用である。

【 0 0 3 5 】

本開示は、CRISPR RNA誘導型エンドヌクレアーゼを同定する方法を提供する。

【 0 0 3 6 】

組成物

CRISPR / CasYタンパク質及びガイドRNA

CRISPR / CasYエンドヌクレアーゼ（例えば、CasYタンパク質）は、対応するガイドRNA（例えば、CasYガイドRNA）と相互作用（結合）し、ガイドRNA

10

20

30

40

50

と標的核酸分子内の標的配列との間の塩基対形成を介して標的核酸内の特定部位を標的化するリボ核タンパク質（RNP）複合体を形成する。ガイドRNAは、標的核酸の配列（標的部位）に相補的なヌクレオチド配列（ガイド配列）を含む。したがって、CasYタンパク質は、CasYガイドRNAと複合体を形成し、このガイドRNAはガイド配列によってRNP複合体に配列特異性を与える。複合体のCasYタンパク質は、部位特異的な活性を与える。換言すれば、CasYタンパク質は、ガイドRNAとの会合によって、標的核酸配列（例えば、染色体配列または染色体外配列、例えばエピソーム配列、ミニサークル配列、ミトコンドリア配列、葉緑体配列など）内の標的部位に誘導される（例えば、標的部位で安定である）。

【0037】

本開示は、CasYポリペプチド（及び/またはCasYポリペプチドをコードする核酸）を含む組成物を提供する（例えば、この場合のCasYポリペプチドは、天然タンパク質、ニッカーゼCasYタンパク質、dCasYタンパク質、キメラCasYタンパク質などであり得る）。本開示は、CasYガイドRNA（及び/またはCasYガイドRNAをコードする核酸）を含む組成物を提供する。本開示は、（a）CasYポリペプチド（及び/またはCasYポリペプチドをコードする核酸）（例えば、この場合のCasYポリペプチドは、天然タンパク質、ニッカーゼCasYタンパク質、dCasYタンパク質、キメラCasYタンパク質などであり得る）、及び（b）CasYガイドRNA（及び/またはCasYガイドRNAをコードする核酸）を含む組成物を提供する。本開示は、（a）本開示のCasYポリペプチド（例えば、この場合のCasYポリペプチドは、天然タンパク質、ニッカーゼCasYタンパク質、dCasYタンパク質、キメラCasYタンパク質などであり得る）；及び（b）CasYガイドRNAを含む、核酸/タンパク質複合体（RNP複合体）を提供する。

【0038】

CasYタンパク質

CasYポリペプチド（この用語は用語「CasYタンパク質」と同義に使用される）は、標的核酸及び/または標的核酸と会合したポリペプチドに結合する、及び/またはそれらを修飾（例えば、切断、ニック、メチル化、脱メチル化など）することができる（例えば、ヒストンテールのメチル化またはアセチル化）（例えば、CasYタンパク質は活性を有する融合パートナーを含む場合もあれば、CasYタンパク質がヌクレアーゼ活性を与える場合もある）。場合によって、CasYタンパク質は天然のタンパク質である（例えば、原核細胞に天然に存在する）。他の場合には、CasYタンパク質は、天然のポリペプチドではない（例えば、CasYタンパク質は変異体CasYタンパク質、キメラタンパク質などである）。

【0039】

所与のタンパク質がCasYガイドRNAと相互作用するかどうかを決定するアッセイは、タンパク質と核酸との間の結合を試験する任意の利便な結合アッセイであってよい。好適な結合アッセイ（例えば、ゲルシフトアッセイ）を当業者は把握しているであろう（例えば、CasYガイドRNA及びタンパク質を標的核酸に添加することを含むアッセイ）。タンパク質が活性を有するかどうかを決定する（例えば、タンパク質が、標的核酸を切断するヌクレアーゼ活性及び/または何らかの異種活性を有するかどうかを決定する）アッセイは、任意の利便なアッセイであってよい（例えば、核酸の切断について試験する任意の利便な核酸切断アッセイ）。好適なアッセイ（例えば切断アッセイ）を当業者は把握しているであろう。

【0040】

天然のCasYタンパク質は、標的とする二本鎖DNA（dsDNA）の特異的配列での二本鎖切断を触媒するエンドヌクレアーゼとして機能する。配列特異性は、会合するガイドRNAが標的DNA内の標的配列にハイブリダイズすることによって与えられる。天然のCasYガイドRNAはcrRNAであり、crRNAは、（i）標的DNA内の標的配列にハイブリダイズするガイド配列、及び（ii）CasYタンパク質に結合するス

10

20

30

40

50

テムループ（ヘアピン - dsRNA 二重鎖）を含むタンパク質結合セグメントを含む。

【0041】

いくつかの実施形態では、本発明の方法及び/または組成物の CasY タンパク質は天然（野生型）タンパク質である（またはそれらに由来する）。天然 CasY タンパク質の例を図 1 に示し、配列番号 1 ~ 7 として記載する。天然 CasY タンパク質の例を図 1 に示し、配列番号 1 ~ 8 として記載する。例示的な天然 CasY タンパク質のアライメントを図 2 に表す（タンパク質には「Y1」、「Y2」、「Y3」などの名前がつけられている）。（配列決定データからアセンブルした）7 つの天然 CasY CRISPR 遺伝子座の部分 DNA 骨格を配列番号 21 ~ 27 として記載する。重要な点として、この新たに発見されたタンパク質（CasY）は、以前に同定された CRISPR - Cas エンドヌクレアーゼと比較して短く、したがって、このタンパク質を代わりに使用することで、タンパク質をコードするヌクレオチド配列が比較的短くなるという利点をもたらす。これは、例えば、CasY タンパク質をコードする核酸が望ましい場合、例えば、研究及び/または臨床用途のために真核細胞（例えば、哺乳動物細胞、ヒト細胞、マウス細胞、in vitro、ex vivo、in vivo）などの細胞への送達にウイルスベクター（例えば、AAV ベクター）を用いる場合に有用である。CasY CRISPR 遺伝子座を保有する細菌が、低温（例えば、10 ~ 17）で採取された環境試料中に存在したことも本明細書で指摘されている。したがって、CasY は、低温（例えば、10 ~ 14、10 ~ 17、10 ~ 20）で良好に機能できると期待される（例えば、これまでに発見された他の Cas エンドヌクレアーゼ（endocnuclease）よりも優れている）。

10

20

【0042】

場合によって、（本発明の組成物及び/または方法の）CasY タンパク質は、配列番号 1 として記載される CasY タンパク質配列と 20% 以上の配列同一性（例えば、30% 以上、40% 以上、50% 以上、60% 以上、70% 以上、80% 以上、85% 以上、90% 以上、95% 以上、97% 以上、98% 以上、99% 以上、または 100% の配列同一性）を有するアミノ酸配列を含む。例えば、場合によって、CasY タンパク質は、配列番号 1 として記載される CasY タンパク質配列と 50% 以上の配列同一性（例えば、60% 以上、70% 以上、80% 以上、85% 以上、90% 以上、95% 以上、97% 以上、98% 以上、99% 以上、または 100% の配列同一性）を有するアミノ酸配列を含む。場合によって、CasY タンパク質は、配列番号 1 として記載される CasY タンパク質配列と 80% 以上の配列同一性（例えば、85% 以上、90% 以上、95% 以上、97% 以上、98% 以上、99% 以上、または 100% の配列同一性）を有するアミノ酸配列を含む。場合によって、CasY タンパク質は、配列番号 1 として記載される CasY タンパク質配列と 90% 以上の配列同一性（例えば、95% 以上、97% 以上、98% 以上、99% 以上、または 100% の配列同一性）を有するアミノ酸配列を含む。場合によって、CasY タンパク質は、配列番号 1 として記載される CasY タンパク質配列を有するアミノ酸配列を含む。場合によって、CasY タンパク質は、その配列が、（例えば、後述するアミノ酸位置などに）タンパク質の天然の触媒活性を減少させるアミノ酸置換（例えば、1、2、または 3 つのアミノ酸置換）を含むという点を除いて、配列番号 1 として記載される CasY タンパク質配列を有するアミノ酸配列を含む。

30

40

【0043】

場合によって、CasY タンパク質は、配列番号 2 として記載される CasY タンパク質配列と 20% 以上の配列同一性（例えば、30% 以上、40% 以上、50% 以上、60% 以上、70% 以上、80% 以上、85% 以上、90% 以上、95% 以上、97% 以上、98% 以上、99% 以上、または 100% の配列同一性）を有するアミノ酸配列を含む。場合によって、CasY タンパク質は、配列番号 2 として記載される CasY タンパク質配列と 50% 以上の配列同一性（例えば、60% 以上、70% 以上、80% 以上、85% 以上、90% 以上、95% 以上、97% 以上、98% 以上、99% 以上、または 100% の配列同一性）を有するアミノ酸配列を含む。場合によって、CasY タンパク質は、配

50

列番号2として記載されるCasYタンパク質配列と80%以上の配列同一性(例えば、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性)を有するアミノ酸配列を含む。場合によって、CasYタンパク質は、配列番号2として記載されるCasYタンパク質配列と90%以上の配列同一性(例えば、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性)を有するアミノ酸配列を含む。場合によって、CasYタンパク質は、その配列が、(例えば、後述するアミノ酸位置などに)タンパク質の天然の触媒活性を減少させるアミノ酸置換(例えば、1、2、または3つのアミノ酸置換)を含むという点を除いて、配列番号2として記載されるCasYタンパク質配列を有するアミノ酸配列を含む。

10

【0044】

場合によって、CasYタンパク質は、配列番号3として記載されるCasYタンパク質配列と20%以上の配列同一性(例えば、30%以上、40%以上、50%以上、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性)を有するアミノ酸配列を含む。場合によって、CasYタンパク質は、配列番号3として記載されるCasYタンパク質配列と50%以上の配列同一性(例えば、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性)を有するアミノ酸配列を含む。場合によって、CasYタンパク質は、配列番号3として記載されるCasYタンパク質配列と80%以上の配列同一性(例えば、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性)を有するアミノ酸配列を含む。場合によって、CasYタンパク質は、配列番号3として記載されるCasYタンパク質配列と90%以上の配列同一性(例えば、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性)を有するアミノ酸配列を含む。場合によって、CasYタンパク質は、その配列が、(例えば、後述するアミノ酸位置などに)タンパク質の天然の触媒活性を減少させるアミノ酸置換(例えば、1、2、または3つのアミノ酸置換)を含むという点を除いて、配列番号3として記載されるCasYタンパク質配列を有するアミノ酸配列を含む。

20

30

【0045】

場合によって、CasYタンパク質は、配列番号4として記載されるCasYタンパク質配列と20%以上の配列同一性(例えば、30%以上、40%以上、50%以上、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性)を有するアミノ酸配列を含む。場合によって、CasYタンパク質は、配列番号4として記載されるCasYタンパク質配列と50%以上の配列同一性(例えば、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性)を有するアミノ酸配列を含む。場合によって、CasYタンパク質は、配列番号4として記載されるCasYタンパク質配列と80%以上の配列同一性(例えば、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性)を有するアミノ酸配列を含む。場合によって、CasYタンパク質は、配列番号4として記載されるCasYタンパク質配列と90%以上の配列同一性(例えば、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性)を有するアミノ酸配列を含む。場合によって、CasYタンパク質は、その配列が、(例えば、後述するアミノ酸位置などに)タンパク質の天然の触媒活性を減少させるアミノ酸置換(例えば、1、2、または3つのアミノ酸置換)を含むという点を除いて、配列番号4として記載されるCasYタンパク質配列を有するアミノ酸配列を含む。

40

50

るアミノ酸配列を含む。

【0046】

場合によって、C a s Yタンパク質は、配列番号5として記載されるC a s Yタンパク質配列と20%以上の配列同一性（例えば、30%以上、40%以上、50%以上、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有するアミノ酸配列を含む。場合によって、C a s Yタンパク質は、配列番号5として記載されるC a s Yタンパク質配列と50%以上の配列同一性（例えば、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有するアミノ酸配列を含む。場合によって、C a s Yタンパク質は、配列番号5として記載されるC a s Yタンパク質配列と80%以上の配列同一性（例えば、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有するアミノ酸配列を含む。場合によって、C a s Yタンパク質は、配列番号5として記載されるC a s Yタンパク質配列と90%以上の配列同一性（例えば、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有するアミノ酸配列を含む。場合によって、C a s Yタンパク質は、その配列が、（例えば、後述するアミノ酸位置などに）タンパク質の天然の触媒活性を減少させるアミノ酸置換（例えば、1、2、または3つのアミノ酸置換）を含むという点を除いて、配列番号5として記載されるC a s Yタンパク質配列を有するアミノ酸配列を含む。

【0047】

場合によって、C a s Yタンパク質は、配列番号6として記載されるC a s Yタンパク質配列と20%以上の配列同一性（例えば、30%以上、40%以上、50%以上、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有するアミノ酸配列を含む。場合によって、C a s Yタンパク質は、配列番号6として記載されるC a s Yタンパク質配列と50%以上の配列同一性（例えば、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有するアミノ酸配列を含む。場合によって、C a s Yタンパク質は、配列番号6として記載されるC a s Yタンパク質配列と80%以上の配列同一性（例えば、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有するアミノ酸配列を含む。場合によって、C a s Yタンパク質は、配列番号6として記載されるC a s Yタンパク質配列と90%以上の配列同一性（例えば、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有するアミノ酸配列を含む。場合によって、C a s Yタンパク質は、その配列が、（例えば、後述するアミノ酸位置などに）タンパク質の天然の触媒活性を減少させるアミノ酸置換（例えば、1、2、または3つのアミノ酸置換）を含むという点を除いて、配列番号6として記載されるC a s Yタンパク質配列を有するアミノ酸配列を含む。

【0048】

場合によって、C a s Yタンパク質は、配列番号7として記載されるC a s Yタンパク質配列と20%以上の配列同一性（例えば、30%以上、40%以上、50%以上、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有するアミノ酸配列を含む。場合によって、C a s Yタンパク質は、配列番号7として記載されるC a s Yタンパク質配列と50%以上の配列同一性（例えば、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有するアミノ酸配列を含む。場合によって、C a s Yタンパク質は、配

列番号7として記載されるC a s Yタンパク質配列と80%以上の配列同一性（例えば、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有するアミノ酸配列を含む。場合によって、C a s Yタンパク質は、配列番号7として記載されるC a s Yタンパク質配列と90%以上の配列同一性（例えば、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有するアミノ酸配列を含む。場合によって、C a s Yタンパク質は、その配列が、（例えば、後述するアミノ酸位置などに）タンパク質の天然の触媒活性を減少させるアミノ酸置換（例えば、1、2、または3つのアミノ酸置換）を含むという点を除いて、配列番号7として記載されるC a s Yタンパク質配列を有するアミノ酸配列を含む。

10

【0049】

場合によって、C a s Yタンパク質は、配列番号8として記載されるC a s Yタンパク質配列と20%以上の配列同一性（例えば、30%以上、40%以上、50%以上、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有するアミノ酸配列を含む。場合によって、C a s Yタンパク質は、配列番号8として記載されるC a s Yタンパク質配列と50%以上の配列同一性（例えば、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有するアミノ酸配列を含む。場合によって、C a s Yタンパク質は、配列番号8として記載されるC a s Yタンパク質配列と80%以上の配列同一性（例えば、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有するアミノ酸配列を含む。場合によって、C a s Yタンパク質は、配列番号8として記載されるC a s Yタンパク質配列と90%以上の配列同一性（例えば、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有するアミノ酸配列を含む。場合によって、C a s Yタンパク質は、その配列が、（例えば、後述するアミノ酸位置などに）タンパク質の天然の触媒活性を減少させるアミノ酸置換（例えば、1、2、または3つのアミノ酸置換）を含むという点を除いて、配列番号8として記載されるC a s Yタンパク質配列を有するアミノ酸配列を含む。

20

30

【0050】

場合によって、C a s Yタンパク質は、配列番号9として記載されるC a s Yタンパク質配列と20%以上の配列同一性（例えば、30%以上、40%以上、50%以上、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有するアミノ酸配列を含む。場合によって、C a s Yタンパク質は、配列番号9として記載されるC a s Yタンパク質配列と50%以上の配列同一性（例えば、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有するアミノ酸配列を含む。場合によって、C a s Yタンパク質は、配列番号9として記載されるC a s Yタンパク質配列と80%以上の配列同一性（例えば、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有するアミノ酸配列を含む。場合によって、C a s Yタンパク質は、配列番号9として記載されるC a s Yタンパク質配列と90%以上の配列同一性（例えば、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有するアミノ酸配列を含む。場合によって、C a s Yタンパク質は、その配列が、（例えば、後述するアミノ酸位置などに）タンパク質の天然の触媒活性を減少させるアミノ酸置換（例えば、1、2、または3つのアミノ酸置換）を含むという点を除いて、配列番号9として記載されるC a s Yタンパク質配列を有するアミノ酸配列を含む。

40

50

るアミノ酸配列を含む。

【0051】

場合によって、C a s Yタンパク質は、配列番号1～4として記載されるC a s Yタンパク質配列のいずれか1つと20%以上の配列同一性（例えば、30%以上、40%以上、50%以上、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有するアミノ酸配列を含む。場合によって、C a s Yタンパク質は、配列番号1～4として記載されるC a s Yタンパク質配列のいずれか1つと50%以上の配列同一性（例えば、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有するアミノ酸配列を含む。場合によって、C a s Yタンパク質は、配列番号1～4として記載されるC a s Yタンパク質配列のいずれか1つと80%以上の配列同一性（例えば、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有するアミノ酸配列を含む。場合によって、C a s Yタンパク質は、配列番号1～4として記載されるC a s Yタンパク質配列のいずれか1つと90%以上の配列同一性（例えば、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有するアミノ酸配列を含む。場合によって、C a s Yタンパク質は、配列番号1～4のいずれか1つに記載されるC a s Yタンパク質配列を有するアミノ酸配列を含む。場合によって、C a s Yタンパク質は、その配列が、（例えば、後述するアミノ酸位置などに）タンパク質の天然の触媒活性を減少させるアミノ酸置換（例えば、1、2、または3つのアミノ酸置換）を含むという点を除いて、配列番号1～4のいずれか1つに記載されるC a s Yタンパク質配列を有するアミノ酸配列を含む。

【0052】

場合によって、C a s Yタンパク質は、配列番号1～5として記載されるC a s Yタンパク質配列のいずれか1つと20%以上の配列同一性（例えば、30%以上、40%以上、50%以上、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有するアミノ酸配列を含む。場合によって、C a s Yタンパク質は、配列番号1～5として記載されるC a s Yタンパク質配列のいずれか1つと50%以上の配列同一性（例えば、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有するアミノ酸配列を含む。場合によって、C a s Yタンパク質は、配列番号1～5として記載されるC a s Yタンパク質配列のいずれか1つと80%以上の配列同一性（例えば、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有するアミノ酸配列を含む。場合によって、C a s Yタンパク質は、配列番号1～5として記載されるC a s Yタンパク質配列のいずれか1つと90%以上の配列同一性（例えば、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有するアミノ酸配列を含む。場合によって、C a s Yタンパク質は、その配列が、（例えば、後述するアミノ酸位置などに）タンパク質の天然の触媒活性を減少させるアミノ酸置換（例えば、1、2、または3つのアミノ酸置換）を含むという点を除いて、配列番号1～5のいずれか1つに記載されるC a s Yタンパク質配列を有するアミノ酸配列を含む。

【0053】

場合によって、C a s Yタンパク質は、配列番号1～7として記載されるC a s Yタンパク質配列のいずれか1つと20%以上の配列同一性（例えば、30%以上、40%以上、50%以上、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有するアミノ酸配列を含む。場合によって、C a s Yタンパク質は、配列番号1～7として記載されるC a s Yタンパク質配列のいずれか1つと50%以上の配列同一性（例えば、60%

以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性)を有するアミノ酸配列を含む。場合によって、CasYタンパク質は、配列番号1~7として記載されるCasYタンパク質配列のいずれか1つと80%以上の配列同一性(例えば、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性)を有するアミノ酸配列を含む。場合によって、CasYタンパク質は、配列番号1~7として記載されるCasYタンパク質配列のいずれか1つと90%以上の配列同一性(例えば、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性)を有するアミノ酸配列を含む。場合によって、CasYタンパク質は、その配列が、(例えば、後述するアミノ酸位置などに)タンパク質の天然の触媒活性を減少させるアミノ酸置換(例えば、1、2、または3つのアミノ酸置換)を含むという点を除いて、配列番号1~7のいずれか1つに記載されるCasYタンパク質配列を有するアミノ酸配列を含む。

10

【0054】

場合によって、CasYタンパク質は、配列番号1~8として記載されるCasYタンパク質配列のいずれか1つと20%以上の配列同一性(例えば、30%以上、40%以上、50%以上、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性)を有するアミノ酸配列を含む。場合によって、CasYタンパク質は、配列番号1~8として記載されるCasYタンパク質配列のいずれか1つと50%以上の配列同一性(例えば、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性)を有するアミノ酸配列を含む。場合によって、CasYタンパク質は、配列番号1~8として記載されるCasYタンパク質配列のいずれか1つと80%以上の配列同一性(例えば、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性)を有するアミノ酸配列を含む。場合によって、CasYタンパク質は、配列番号1~8として記載されるCasYタンパク質配列のいずれか1つと90%以上の配列同一性(例えば、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性)を有するアミノ酸配列を含む。場合によって、CasYタンパク質は、その配列が、(例えば、後述するアミノ酸位置などに)タンパク質の天然の触媒活性を減少させるアミノ酸置換(例えば、1、2、または3つのアミノ酸置換)を含むという点を除いて、配列番号1~8のいずれか1つに記載されるCasYタンパク質配列を有するアミノ酸配列を含む。

20

30

【0055】

CasYタンパク質ドメイン

CasYタンパク質のドメインを図3に示す。図3の模式図(アミノ酸は、CasY1タンパク質(配列番号1)に基づいて付番されている)に見られるように、CasYタンパク質には、おおよそ800~1000アミノ酸長(例えば、CasY1は約815、及びCasY5は約980)のN末端ドメインと、3つの部分RuvCドメイン(RuvC-I、RuvC-II、及びRuvC-III。本明細書でサブドメインとも称する)を含むC末端ドメインとが含まれ、RuvCドメインは、CasYタンパク質の一次アミノ酸配列では隣接していないが、タンパク質が産生され、折り畳まれるとRuvCドメインを形成する。したがって、場合によって、(本発明の組成物及び/または方法の)CasYタンパク質は、750~1050アミノ酸(例えば、750~1025、750~1000、750~950、775~1050、775~1025、775~1000、775~950、800~1050、800~1025、800~1000、または800~950アミノ酸)の範囲の長さを有するN末端ドメインをもつアミノ酸配列を含む(例えば、NLS及び/または触媒活性を有するドメインのような何らかの融合された異種配列

40

50

は含まない)。場合によって、(本発明の組成物及び/または方法の) Cas Yタンパク質は、分割された Ruv Cドメイン(例えば、3つの部分 Ruv Cドメイン - Ruv C - I、Ruv C - II、及び Ruv C - III)に対してN末端側である、750~1050アミノ酸(例えば、750~1025、750~1000、750~950、775~1050、775~1025、775~1000、775~950、800~1050、800~1025、800~1000、または800~950アミノ酸)の範囲の長さを有するアミノ酸配列を含む(例えば、NLS及び/または触媒活性を有するドメインのような何らかの融合された異種配列は含まない)。

【0056】

場合によって、(本発明の組成物及び/または方法の) Cas Yタンパク質は、配列番号1として記載される Cas Yタンパク質配列のN末端ドメイン(例えば、図3、パネルAの Cas Y1においてアミノ酸1~812として示されるドメイン)と20%以上の配列同一性(例えば、30%以上、40%以上、50%以上、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性)を有するアミノ酸配列を含む。例えば、場合によって、Cas Yタンパク質は、配列番号1として記載される Cas Yタンパク質配列のN末端ドメイン(例えば、図3、パネルAの Cas Y1においてアミノ酸1~812として示されるドメイン)と50%以上の配列同一性(例えば、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性)を有するアミノ酸配列を含む。場合によって、Cas Yタンパク質は、配列番号1として記載される Cas Yタンパク質配列のN末端ドメイン(例えば、図3、パネルAの Cas Y1においてアミノ酸1~812として示されるドメイン)と80%以上の配列同一性(例えば、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性)を有するアミノ酸配列を含む。場合によって、Cas Yタンパク質は、配列番号1として記載される Cas Yタンパク質配列のN末端ドメイン(例えば、図3、パネルAの Cas Y1においてアミノ酸1~812として示されるドメイン)と90%以上の配列同一性(例えば、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性)を有するアミノ酸配列を含む。場合によって、Cas Yタンパク質は、配列番号1として記載される Cas Yタンパク質配列のアミノ酸1~812を有するアミノ酸配列を含む。

【0057】

場合によって、(本発明の組成物及び/または方法の) Cas Yタンパク質は、配列番号1~4のいずれか1つに記載される Cas Yタンパク質配列のN末端ドメイン(例えば、図3、パネルAの Cas Y1においてアミノ酸1~812として示されるドメイン)と20%以上の配列同一性(例えば、30%以上、40%以上、50%以上、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性)を有するアミノ酸配列を含む。例えば、場合によって、Cas Yタンパク質は、配列番号1~4のいずれか1つに記載される Cas Yタンパク質配列のN末端ドメイン(例えば、図3、パネルAの Cas Y1においてアミノ酸1~812として示されるドメイン)と50%以上の配列同一性(例えば、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性)を有するアミノ酸配列を含む。場合によって、Cas Yタンパク質は、配列番号1~4のいずれか1つに記載される Cas Yタンパク質配列のN末端ドメイン(例えば、図3、パネルAの Cas Y1においてアミノ酸1~812として示されるドメイン)と80%以上の配列同一性(例えば、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性)を有するアミノ酸配列を含む。場合によって、Cas Yタンパク質は、配列番号1~4のいずれか1つに記載される Cas Yタンパク質配列のN末端ドメイン(例えば、図3、パネルAの Cas Y1においてアミノ酸1~812として示されるドメイン)と90%以上の配列同一性(例えば、95%以上、97%以上、98%以上、99%以上

10

20

30

40

50

、または100%の配列同一性)を有するアミノ酸配列を含む。場合によって、CasYタンパク質は、配列番号1として記載されるCasYタンパク質配列のアミノ酸1~812に対応する、配列番号1~4のいずれか1つのアミノ酸配列の断片を含む。

【0058】

場合によって、(本発明の組成物及び/または方法の)CasYタンパク質は、配列番号1~5のいずれか1つに記載されるCasYタンパク質配列のN末端ドメイン(例えば、図3、パネルAのCasY1においてアミノ酸1~812として示されるドメイン)と20%以上の配列同一性(例えば、30%以上、40%以上、50%以上、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性)を有するアミノ酸配列を含む。例えば、
場合によって、CasYタンパク質は、配列番号1~5のいずれか1つに記載されるCasYタンパク質配列のN末端ドメイン(例えば、図3、パネルAのCasY1においてアミノ酸1~812として示されるドメイン)と50%以上の配列同一性(例えば、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性)を有するアミノ酸配列を含む。場合によって、CasYタンパク質は、配列番号1~5のいずれか1つに記載されるCasYタンパク質配列のN末端ドメイン(例えば、図3、パネルAのCasY1においてアミノ酸1~812として示されるドメイン)と80%以上の配列同一性(例えば、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性)を有するアミノ酸配列を含む。場合によって、CasYタンパク質は、配列番号1~5のいずれか1つに記載されるCasYタンパク質配列のN末端ドメイン(例えば、図3、パネルAのCasY1においてアミノ酸1~812として示されるドメイン)と90%以上の配列同一性(例えば、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性)を有するアミノ酸配列を含む。場合によって、CasYタンパク質は、配列番号1として記載されるCasYタンパク質配列のアミノ酸1~812に対応する、配列番号1~5のいずれか1つのアミノ酸配列の断片を含む。

【0059】

場合によって、(本発明の組成物及び/または方法の)CasYタンパク質は、配列番号1~7のいずれか1つに記載されるCasYタンパク質配列のN末端ドメイン(例えば、図3、パネルAのCasY1においてアミノ酸1~812として示されるドメイン)と20%以上の配列同一性(例えば、30%以上、40%以上、50%以上、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性)を有するアミノ酸配列を含む。例えば、
場合によって、CasYタンパク質は、配列番号1~7のいずれか1つに記載されるCasYタンパク質配列のN末端ドメイン(例えば、図3、パネルAのCasY1においてアミノ酸1~812として示されるドメイン)と50%以上の配列同一性(例えば、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性)を有するアミノ酸配列を含む。場合によって、CasYタンパク質は、配列番号1~7のいずれか1つに記載されるCasYタンパク質配列のN末端ドメイン(例えば、図3、パネルAのCasY1においてアミノ酸1~812として示されるドメイン)と80%以上の配列同一性(例えば、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性)を有するアミノ酸配列を含む。場合によって、CasYタンパク質は、配列番号1~7のいずれか1つに記載されるCasYタンパク質配列のN末端ドメイン(例えば、図3、パネルAのCasY1においてアミノ酸1~812として示されるドメイン)と90%以上の配列同一性(例えば、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性)を有するアミノ酸配列を含む。場合によって、CasYタンパク質は、配列番号1として記載されるCasYタンパク質配列のアミノ酸1~812に対応する、配列番号1~7のいずれか1つのアミノ酸配列の断片を含む。

【0060】

50

場合によって、（本発明の組成物及び/または方法の）C a s Yタンパク質は、配列番号1～8のいずれか1つに記載されるC a s Yタンパク質配列のN末端ドメイン（例えば、図3、パネルAのC a s Y 1においてアミノ酸1～812として示されるドメイン）と20%以上の配列同一性（例えば、30%以上、40%以上、50%以上、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有するアミノ酸配列を含む。例えば、場合によって、C a s Yタンパク質は、配列番号1～8のいずれか1つに記載されるC a s Yタンパク質配列のN末端ドメイン（例えば、図3、パネルAのC a s Y 1においてアミノ酸1～812として示されるドメイン）と50%以上の配列同一性（例えば、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有するアミノ酸配列を含む。場合によって、C a s Yタンパク質は、配列番号1～8のいずれか1つに記載されるC a s Yタンパク質配列のN末端ドメイン（例えば、図3、パネルAのC a s Y 1においてアミノ酸1～812として示されるドメイン）と80%以上の配列同一性（例えば、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有するアミノ酸配列を含む。場合によって、C a s Yタンパク質は、配列番号1～8のいずれか1つに記載されるC a s Yタンパク質配列のN末端ドメイン（例えば、図3、パネルAのC a s Y 1においてアミノ酸1～812として示されるドメイン）と90%以上の配列同一性（例えば、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有するアミノ酸配列を含む。場合によって、C a s Yタンパク質は、配列番号1として記載されるC a s Yタンパク質配列のアミノ酸1～812に対応する、配列番号1～8のいずれか1つのアミノ酸配列の断片を含む。

10

20

【0061】

場合によって、（本発明の組成物及び/または方法の）C a s Yタンパク質は、配列番号1～4として記載されるC a s Yタンパク質配列のいずれか1つのN末端ドメイン（例えば、図3、パネルAのC a s Y 1においてアミノ酸1～812として示されるドメイン）と20%以上の配列同一性（例えば、30%以上、40%以上、50%以上、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有する第1のアミノ酸配列と、分割R u v Cドメイン（例えば、3つの部分R u v Cドメイン - R u v C - I、R u v C - I I、及びR u v C - I I I）を含む、第1のアミノ酸配列に対してC末端側にある第2のアミノ酸配列とを含む。例えば、場合によって、C a s Yタンパク質は、配列番号1～4として記載されるC a s Yタンパク質配列のいずれか1つのN末端ドメイン（例えば、図3、パネルAのC a s Y 1においてアミノ酸1～812として示されるドメイン）と50%以上の配列同一性（例えば、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有する第1のアミノ酸配列と、分割R u v Cドメイン（例えば、3つの部分R u v Cドメイン - R u v C - I、R u v C - I I、及びR u v C - I I I）を含む、第1のアミノ酸配列に対してC末端側にある第2のアミノ酸配列とを含む。場合によって、C a s Yタンパク質は、配列番号1～4として記載されるC a s Yタンパク質配列のいずれか1つのN末端ドメイン（例えば、図3、パネルAのC a s Y 1においてアミノ酸1～812として示されるドメイン）と80%以上の配列同一性（例えば、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有する第1のアミノ酸配列と、分割R u v Cドメイン（例えば、3つの部分R u v Cドメイン - R u v C - I、R u v C - I I、及びR u v C - I I I）を含む、第1のアミノ酸配列に対してC末端側にある第2のアミノ酸配列とを含む。場合によって、C a s Yタンパク質は、配列番号1～4として記載されるC a s Yタンパク質配列のいずれか1つのN末端ドメイン（例えば、図3、パネルAのC a s Y 1においてアミノ酸1～812として示されるドメイン）と90%以上の配列同一性（例えば、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有する第1のアミノ酸配

30

40

50

列と、分割 R u v C ドメイン（例えば、3つの部分 R u v C ドメイン - R u v C - I、R u v C - I I、及び R u v C - I I I）を含む、第1のアミノ酸配列に対してC末端側にある第2のアミノ酸配列とを含む。場合によって、C a s Y タンパク質は、配列番号1として記載される C a s Y タンパク質配列のアミノ酸1～812に対応するアミノ酸配列と、分割 R u v C ドメイン（例えば、3つの部分 R u v C ドメイン - R u v C - I、R u v C - I I、及び R u v C - I I I）を含む、第1のアミノ酸配列に対してC末端側にある第2のアミノ酸配列とを含む。場合によって、（本発明の組成物及び/または方法の）C a s Y タンパク質は、配列番号1～5として記載される C a s Y タンパク質配列のいずれか1つのN末端ドメイン（例えば、図3、パネルAの C a s Y 1においてアミノ酸1～812として示されるドメイン）と20%以上の配列同一性（例えば、30%以上、40%以上、50%以上、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有する第1のアミノ酸配列と、分割 R u v C ドメイン（例えば、3つの部分 R u v C ドメイン - R u v C - I、R u v C - I I、及び R u v C - I I I）を含む、第1のアミノ酸配列に対してC末端側にある第2のアミノ酸配列とを含む。例えば、場合によって、C a s Y タンパク質は、配列番号1～5として記載される C a s Y タンパク質配列のいずれか1つのN末端ドメイン（例えば、図3、パネルAの C a s Y 1においてアミノ酸1～812として示されるドメイン）と50%以上の配列同一性（例えば、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有する第1のアミノ酸配列と、分割 R u v C ドメイン（例えば、3つの部分 R u v C ドメイン - R u v C - I、R u v C - I I、及び R u v C - I I I）を含む、第1のアミノ酸配列に対してC末端側にある第2のアミノ酸配列とを含む。場合によって、C a s Y タンパク質は、配列番号1～5として記載される C a s Y タンパク質配列のいずれか1つのN末端ドメイン（例えば、図3、パネルAの C a s Y 1においてアミノ酸1～812として示されるドメイン）と80%以上の配列同一性（例えば、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有する第1のアミノ酸配列と、分割 R u v C ドメイン（例えば、3つの部分 R u v C ドメイン - R u v C - I、R u v C - I I、及び R u v C - I I I）を含む、第1のアミノ酸配列に対してC末端側にある第2のアミノ酸配列とを含む。場合によって、C a s Y タンパク質は、配列番号1～5として記載される C a s Y タンパク質配列のいずれか1つのN末端ドメイン（例えば、図3、パネルAの C a s Y 1においてアミノ酸1～812として示されるドメイン）と90%以上の配列同一性（例えば、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有する第1のアミノ酸配列と、分割 R u v C ドメイン（例えば、3つの部分 R u v C ドメイン - R u v C - I、R u v C - I I、及び R u v C - I I I）を含む、第1のアミノ酸配列に対してC末端側にある第2のアミノ酸配列とを含む。場合によって、C a s Y タンパク質は、配列番号1として記載される C a s Y タンパク質配列のアミノ酸1～812に対応するアミノ酸配列と、分割 R u v C ドメイン（例えば、3つの部分 R u v C ドメイン - R u v C - I、R u v C - I I、及び R u v C - I I I）を含む、第1のアミノ酸配列に対してC末端側にある第2のアミノ酸配列とを含む。

【0062】

場合によって、（本発明の組成物及び/または方法の）C a s Y タンパク質は、配列番号1～7として記載される C a s Y タンパク質配列のいずれか1つのN末端ドメイン（例えば、図3、パネルAの C a s Y 1においてアミノ酸1～812として示されるドメイン）と20%以上の配列同一性（例えば、30%以上、40%以上、50%以上、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有する第1のアミノ酸配列と、分割 R u v C ドメイン（例えば、3つの部分 R u v C ドメイン - R u v C - I、R u v C - I I、及び R u v C - I I I）を含む、第1のアミノ酸配列に対してC末端側にある第2のアミノ酸配列とを含む。例えば、場合によって、C a s Y タンパク質は、配列番号1

~ 7として記載されるCasYタンパク質配列のいずれか1つのN末端ドメイン（例えば、図3、パネルAのCasY1においてアミノ酸1~812として示されるドメイン）と50%以上の配列同一性（例えば、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有する第1のアミノ酸配列と、分割RuvCドメイン（例えば、3つの部分RuvCドメイン-RuvC-I、RuvC-II、及びRuvC-III）を含む、第1のアミノ酸配列に対してC末端側にある第2のアミノ酸配列とを含む。場合によって、CasYタンパク質は、配列番号1~7として記載されるCasYタンパク質配列のいずれか1つのN末端ドメイン（例えば、図3、パネルAのCasY1においてアミノ酸1~812として示されるドメイン）と80%以上の配列同一性（例えば、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有する第1のアミノ酸配列と、分割RuvCドメイン（例えば、3つの部分RuvCドメイン-RuvC-I、RuvC-II、及びRuvC-III）を含む、第1のアミノ酸配列に対してC末端側にある第2のアミノ酸配列とを含む。場合によって、CasYタンパク質は、配列番号1~7として記載されるCasYタンパク質配列のいずれか1つのN末端ドメイン（例えば、図3、パネルAのCasY1においてアミノ酸1~812として示されるドメイン）と90%以上の配列同一性（例えば、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有する第1のアミノ酸配列と、分割RuvCドメイン（例えば、3つの部分RuvCドメイン-RuvC-I、RuvC-II、及びRuvC-III）を含む、第1のアミノ酸配列に対してC末端側にある第2のアミノ酸配列とを含む。

【0063】

場合によって、（本発明の組成物及び/または方法の）CasYタンパク質は、配列番号1~8として記載されるCasYタンパク質配列のいずれか1つのN末端ドメイン（例えば、図3、パネルAのCasY1においてアミノ酸1~812として示されるドメイン）と20%以上の配列同一性（例えば、30%以上、40%以上、50%以上、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有する第1のアミノ酸配列と、分割RuvCドメイン（例えば、3つの部分RuvCドメイン-RuvC-I、RuvC-II、及びRuvC-III）を含む、第1のアミノ酸配列に対してC末端側にある第2のアミノ酸配列とを含む。例えば、場合によって、CasYタンパク質は、配列番号1~8として記載されるCasYタンパク質配列のいずれか1つのN末端ドメイン（例えば、図3、パネルAのCasY1においてアミノ酸1~812として示されるドメイン）と50%以上の配列同一性（例えば、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有する第1のアミノ酸配列と、分割RuvCドメイン（例えば、3つの部分RuvCドメイン-RuvC-I、RuvC-II、及びRuvC-III）を含む、第1のアミノ酸配列に対してC末端側にある第2のアミノ酸配列とを含む。場合によって、CasYタンパク質は、配列番号1~8として記載されるCasYタンパク質配列のいずれか1つのN末端ドメイン（例えば、図3、パネルAのCasY1においてアミノ酸1~812として示されるドメイン）と80%以上の配列同一性（例えば、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有する第1のアミノ酸配列と、分割RuvCドメイン（例えば、3つの部分RuvCドメイン-RuvC-I、RuvC-II、及びRuvC-III）を含む、第1のアミノ酸配列に対してC末端側にある第2のアミノ酸配列とを含む。場合によって、CasYタンパク質は、配列番号1~8として記載されるCasYタンパク質配列のいずれか

1つのN末端ドメイン（例えば、図3、パネルAのCasY1においてアミノ酸1～812として示されるドメイン）と90%以上の配列同一性（例えば、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有する第1のアミノ酸配列と、分割RuvCドメイン（例えば、3つの部分RuvCドメイン-RuvC-I、RuvC-II、及びRuvC-III）を含む、第1のアミノ酸配列に対してC末端側にある第2のアミノ酸配列とを含む。場合によって、CasYタンパク質は、配列番号1として記載されるCasYタンパク質配列のアミノ酸1～812に対応するアミノ酸配列と、分割RuvCドメイン（例えば、3つの部分RuvCドメイン-RuvC-I、RuvC-II、及びRuvC-III）を含む、第1のアミノ酸配列に対してC末端側にある第2のアミノ酸配列とを含む。

10

【0064】

いくつかの実施形態では、（本発明の組成物及び/または方法の）CasYタンパク質の分割RuvCドメインは、RuvC-IIサブドメインとRuvC-IIIサブドメインとの間にRuvC-IIIサブドメインより大きい領域を含む。例えば、場合によっては、RuvC-IIIサブドメインの長さに対する、RuvC-IIサブドメインとRuvC-IIIサブドメインとの間の領域の長さの比が1.1以上である（例えば、1.2）。場合によって、RuvC-IIIサブドメインの長さに対する、RuvC-IIサブドメインとRuvC-IIIサブドメインとの間の領域の長さの比が1より大きい。場合によって、RuvC-IIIサブドメインの長さに対する、RuvC-IIサブドメインとRuvC-IIIサブドメインとの間の領域の長さの比は、1より大きく、かつ1～1.5（例えば、1～1.4、1～1.3、または1～1.2）である。

20

【0065】

（本発明の組成物及び/または方法のCasYタンパク質に関する）いくつかの実施形態では、RuvC-IIIサブドメインの長さに対する、RuvC-IIサブドメインの長さの比は2以下（例えば、1.8以下、1.7以下、1.6以下、1.5以下、または1.4以下）である。例えば、場合によっては、RuvC-IIIサブドメインの長さに対する、RuvC-IIサブドメインの長さの比は1.5以下（例えば、1.4以下）である。いくつかの実施形態では、RuvC-IIIサブドメインの長さに対する、RuvC-IIサブドメインの長さの比は、1～2（例えば、1.1～2、1.2～2、1～1.8、1.1～1.8、1.2～1.8、1～1.6、1.1～1.6、1.2～1.6、1～1.4、1.1～1.4、または1.2～1.4）の範囲内である。

30

【0066】

（本発明の組成物及び/または方法のCasYタンパク質に関して）場合によって、RuvC-IIIサブドメインの長さに対する、RuvC-IIサブドメインとRuvC-IIIサブドメインとの間の領域の長さの比は1より大きい。場合によって、RuvC-IIIサブドメインの長さに対する、RuvC-IIサブドメインとRuvC-IIIサブドメインとの間の領域の長さの比は、1より大きく、かつ1～1.3（例えば、1～1.2）である。

【0067】

（本発明の組成物及び/または方法のCasYタンパク質に関して）場合によって、RuvC-IIサブドメインとRuvC-IIIサブドメインとの間の領域は、少なくとも60アミノ酸長（例えば、少なくとも65、68、または70アミノ酸長）である。場合によって、RuvC-IIサブドメインとRuvC-IIIサブドメインとの間の領域は、60～110アミノ酸の範囲（例えば、60～105、60～100、60～95、60～90、65～110、65～105、65～100、65～95、または65～90アミノ酸の範囲）の長さを有する。

40

【0068】

場合によって、（本発明の組成物及び/または方法の）CasYタンパク質は、配列番号1～4として記載されるCasYタンパク質配列のいずれか1つのN末端ドメイン（例えば、図3、パネルAのCasY1においてアミノ酸1～812として示されるドメイン

50

)と50%以上の配列同一性(例えば、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性)を有する第1のアミノ酸配列と、3つの部分RuvCドメイン-RuvC-I、RuvC-II、及びRuvC-IIIを含む、第1のアミノ酸配列に対してC末端側にある第2のアミノ酸配列とを含み、ここで、(i)RuvC-IIIサブドメインの長さに対する、RuvC-IIサブドメインとRuvC-IIIサブドメインとの間の領域の長さの比は1.1以上(例えば、1.2)であるか；(ii)RuvC-IIIサブドメインの長さに対する、RuvC-IIサブドメインとRuvC-IIIサブドメインとの間の領域の長さの比は1より大きいか；(iii)RuvC-IIIサブドメインの長さに対する、RuvC-IIサブドメインとRuvC-IIIサブドメインとの間の領域の長さの比は、1より大きく、かつ1~1.5(例えば、1~1.4、1~1.3、または1~1.2)であるか；(iv)RuvC-IIIサブドメインの長さに対する、RuvC-IIサブドメインの長さの比は2以下(例えば、1.8以下、1.7以下、1.6以下、1.5以下、または1.4以下)であるか；(v)RuvC-IIIサブドメインの長さに対する、RuvC-IIサブドメインの長さの比は1.5以下(例えば、1.4以下)であるか；(vi)RuvC-IIIサブドメインの長さに対する、RuvC-IIサブドメインの長さの比は、1~2(例えば、1.1~2、1.2~2、1~1.8、1.1~1.8、1.2~1.8、1~1.6、1.1~1.6、1.2~1.6、1~1.4、1.1~1.4、または1.2~1.4)の範囲内であるか；(vii)RuvC-IIIサブドメインの長さに対する、RuvC-IIサブドメインとRuvC-IIサブドメインとの間の領域の長さの比は1より大きいか；(viii)RuvC-IIIサブドメインの長さに対する、RuvC-IIサブドメインとRuvC-IIIサブドメインとの間の領域の長さの比は、1より大きく、かつ1~1.5(例えば、1~1.2)であるか；(ix)RuvC-IIサブドメインとRuvC-IIIサブドメインとの間の領域は、少なくとも60アミノ酸長(例えば、少なくとも65または少なくとも70アミノ酸長)であるか；(x)RuvC-IIサブドメインとRuvC-IIIサブドメインとの間の領域は、少なくとも65アミノ酸長であるか；(xi)RuvC-IIサブドメインとRuvC-IIIサブドメインとの間の領域は、60~110アミノ酸の範囲(例えば、60~105、60~100、60~95、60~90、65~110、65~105、65~100、65~95、または65~90アミノ酸の範囲)の長さを有するか；または(xii)RuvC-IIサブドメインとRuvC-IIIサブドメインとの間の領域は、65~95アミノ酸の範囲の長さを有する。

【0069】

場合によって、(本発明の組成物及び/または方法の)CasYタンパク質は、配列番号1~4として記載されるCasYタンパク質配列のいずれか1つのN末端ドメイン(例えば、図3、パネルAのCasY1においてアミノ酸1~812として示されるドメイン)と75%以上の配列同一性(例えば、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性)を有する第1のアミノ酸配列と、3つの部分RuvCドメイン-RuvC-I、RuvC-II、及びRuvC-IIIを含む、第1のアミノ酸配列に対してC末端側にある第2のアミノ酸配列とを含み、ここで、(i)RuvC-IIIサブドメインの長さに対する、RuvC-IIサブドメインとRuvC-IIIサブドメインとの間の領域の長さの比は1.1以上(例えば、1.2)であるか；(ii)RuvC-IIIサブドメインの長さに対する、RuvC-IIサブドメインとRuvC-IIIサブドメインとの間の領域の長さの比は1より大きいか；(iii)RuvC-IIIサブドメインの長さに対する、RuvC-IIサブドメインとRuvC-IIIサブドメインとの間の領域の長さの比は、1より大きく、かつ1~1.5(例えば、1~1.4、1~1.3、または1~1.2)であるか；(iv)RuvC-IIIサブドメインの長さに対する、RuvC-IIサブドメインの長さの比は2以下(例えば、1.8以下、1.7以下、1.6以下、1.5以下、または1.4以下)であるか；(v)RuvC-IIIサブドメインの長さに対する、Ruv

C - I Iサブドメインの長さの比は1.5以下(例えば、1.4以下)であるか；(v i) R u v C - I I Iサブドメインの長さに対する、R u v C - I Iサブドメインの長さの比は、1~2(例えば、1.1~2、1.2~2、1~1.8、1.1~1.8、1.2~1.8、1~1.6、1.1~1.6、1.2~1.6、1~1.4、1.1~1.4、または1.2~1.4)の範囲内であるか；(v i i) R u v C - I I Iサブドメインの長さに対する、R u v C - I IサブドメインとR u v C - I I Iサブドメインとの間の領域の長さの比は1より大きいか；(v i i i) R u v C - I I Iサブドメインの長さに対する、R u v C - I IサブドメインとR u v C - I I Iサブドメインとの間の領域の長さの比は、1より大きく、かつ1~1.5(例えば、1~1.2)であるか；(i x) R u v C - I IサブドメインとR u v C - I I Iサブドメインとの間の領域は、少なくとも60アミノ酸長(例えば、少なくとも65または少なくとも70アミノ酸長)であるか；(x) R u v C - I IサブドメインとR u v C - I I Iサブドメインとの間の領域は、少なくとも65アミノ酸長であるか；(x i) R u v C - I IサブドメインとR u v C - I I Iサブドメインとの間の領域は、60~110アミノ酸の範囲(例えば、60~105、60~100、60~95、60~90、65~110、65~105、65~100、65~95、または65~90アミノ酸の範囲)の長さを有するか；または(x i i) R u v C - I IサブドメインとR u v C - I I Iサブドメインとの間の領域は、65~95アミノ酸の範囲の長さを有する。

10

【0070】

場合によって、(本発明の組成物及び/または方法の)C a s Yタンパク質は、配列番号1~4として記載されるC a s Yタンパク質配列のいずれか1つのN末端ドメイン(例えば、図3、パネルAのC a s Y1においてアミノ酸1~812として示されるドメイン)と85%以上の配列同一性(例えば、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性)を有する第1のアミノ酸配列と、3つの部分R u v Cドメイン - R u v C - I、R u v C - I I、及びR u v C - I I Iを含む、第1のアミノ酸配列に対してC末端側にある第2のアミノ酸配列とを含み、ここで、(i) R u v C - I I Iサブドメインの長さに対する、R u v C - I IサブドメインとR u v C - I I Iサブドメインとの間の領域の長さの比は1.1以上(例えば、1.2)であるか；(i i) R u v C - I I Iサブドメインの長さに対する、R u v C - I IサブドメインとR u v C - I I Iサブドメインとの間の領域の長さの比は1より大きいか；(i i i) R u v C - I I Iサブドメインの長さに対する、R u v C - I IサブドメインとR u v C - I I Iサブドメインとの間の領域の長さの比は、1より大きく、かつ1~1.5(例えば、1~1.4、1~1.3、または1~1.2)であるか；(i v) R u v C - I I Iサブドメインの長さに対する、R u v C - I Iサブドメインの長さの比は2以下(例えば、1.8以下、1.7以下、1.6以下、1.5以下、または1.4以下)であるか；(v) R u v C - I I Iサブドメインの長さに対する、R u v C - I Iサブドメインの長さの比は1.5以下(例えば、1.4以下)であるか；(v i) R u v C - I I Iサブドメインの長さに対する、R u v C - I Iサブドメインの長さの比は、1~2(例えば、1.1~2、1.2~2、1~1.8、1.1~1.8、1.2~1.8、1~1.6、1.1~1.6、1.2~1.6、1~1.4、1.1~1.4、または1.2~1.4)の範囲内であるか；(v i i) R u v C - I I Iサブドメインの長さに対する、R u v C - I IサブドメインとR u v C - I I Iサブドメインとの間の領域の長さの比は1より大きいか；(v i i i) R u v C - I I Iサブドメインの長さに対する、R u v C - I IサブドメインとR u v C - I I Iサブドメインとの間の領域の長さの比は、1より大きく、かつ1~1.5(例えば、1~1.2)であるか；(i x) R u v C - I IサブドメインとR u v C - I I Iサブドメインとの間の領域は、少なくとも60アミノ酸長(例えば、少なくとも65または少なくとも70アミノ酸長)であるか；(x) R u v C - I IサブドメインとR u v C - I I Iサブドメインとの間の領域は、少なくとも65アミノ酸長であるか；(x i) R u v C - I IサブドメインとR u v C - I I Iサブドメインとの間の領域は、60~110アミノ酸の範囲(例えば、60~105、60~100、60~95

20

30

40

50

、60～90、65～110、65～105、65～100、65～95、または65～90アミノ酸の範囲)の長さを有するか；または(xii)RuvC-IIサブドメインとRuvC-IIISubドメインとの間の領域は、65～95アミノ酸の範囲の長さを有する。

【0071】

場合によって、(本発明の組成物及び/または方法の)CasYタンパク質は、配列番号1～5として記載されるCasYタンパク質配列のいずれか1つのN末端ドメイン(例えば、図3、パネルAのCasY1においてアミノ酸1～812として示されるドメイン)と50%以上の配列同一性(例えば、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性)を有する第1のアミノ酸配列と、3つの部分RuvCドメイン-RuvC-I、RuvC-II、及びRuvC-IIISubドメインを含む、第1のアミノ酸配列に対してC末端側にある第2のアミノ酸配列とを含み、ここで、(i)RuvC-IIISubドメインの長さに対する、RuvC-IIサブドメインとRuvC-IIISubドメインとの間の領域の長さの比は1.1以上(例えば、1.2)であるか；(ii)RuvC-IIISubドメインの長さに対する、RuvC-IIサブドメインとRuvC-IIISubドメインとの間の領域の長さの比は1より大きいか；(iii)RuvC-IIISubドメインの長さに対する、RuvC-IIサブドメインとRuvC-IIISubドメインとの間の領域の長さの比は、1より大きく、かつ1～1.5(例えば、1～1.4、1～1.3、または1～1.2)であるか；(iv)RuvC-IIISubドメインの長さに対する、RuvC-IIサブドメインの長さの比は2以下(例えば、1.8以下、1.7以下、1.6以下、1.5以下、または1.4以下)であるか；(v)RuvC-IIISubドメインの長さに対する、RuvC-IIサブドメインの長さの比は1.5以下(例えば、1.4以下)であるか；(vi)RuvC-IIISubドメインの長さに対する、RuvC-IIサブドメインの長さの比は、1～2(例えば、1.1～2、1.2～2、1～1.8、1.1～1.8、1.2～1.8、1～1.6、1.1～1.6、1.2～1.6、1～1.4、1.1～1.4、または1.2～1.4)の範囲内であるか；(vii)RuvC-IIISubドメインの長さに対する、RuvC-IIサブドメインとRuvC-IIISubドメインとの間の領域の長さの比は1より大きいか；(viii)RuvC-IIISubドメインの長さに対する、RuvC-IIサブドメインとRuvC-IIISubドメインとの間の領域の長さの比は、1より大きく、かつ1～1.5(例えば、1～1.2)であるか；(ix)RuvC-IIサブドメインとRuvC-IIISubドメインとの間の領域は、少なくとも60アミノ酸長(例えば、少なくとも65または少なくとも70アミノ酸長)であるか；(x)RuvC-IIサブドメインとRuvC-IIISubドメインとの間の領域は、少なくとも65アミノ酸長であるか；(xi)RuvC-IIサブドメインとRuvC-IIISubドメインとの間の領域は、60～110アミノ酸の範囲(例えば、60～105、60～100、60～95、60～90、65～110、65～105、65～100、65～95、または65～90アミノ酸の範囲)の長さを有するか；または(xii)RuvC-IIサブドメインとRuvC-IIISubドメインとの間の領域は、65～95アミノ酸の範囲の長さを有する。

【0072】

場合によって、(本発明の組成物及び/または方法の)CasYタンパク質は、配列番号1～5として記載されるCasYタンパク質配列のいずれか1つのN末端ドメイン(例えば、図3、パネルAのCasY1においてアミノ酸1～812として示されるドメイン)と75%以上の配列同一性(例えば、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性)を有する第1のアミノ酸配列と、3つの部分RuvCドメイン-RuvC-I、RuvC-II、及びRuvC-IIISubドメインを含む、第1のアミノ酸配列に対してC末端側にある第2のアミノ酸配列とを含み、ここで、(i)RuvC-IIISubドメインの長さに対する、RuvC-IIサブドメインとRuvC-IIISubドメインとの間の領域の長さの比は1.1以上

10

20

30

40

50

(例えば、1.2)であるか；(ii) RuvC - IIIサブドメインの長さに対する、RuvC - IIサブドメインとRuvC - IIIサブドメインとの間の領域の長さの比は1より大きいか；(iii) RuvC - IIIサブドメインの長さに対する、RuvC - IIサブドメインとRuvC - IIIサブドメインとの間の領域の長さの比は、1より大きく、かつ1~1.5(例えば、1~1.4、1~1.3、または1~1.2)であるか；(iv) RuvC - IIIサブドメインの長さに対する、RuvC - IIサブドメインの長さの比は2以下(例えば、1.8以下、1.7以下、1.6以下、1.5以下、または1.4以下)であるか；(v) RuvC - IIIサブドメインの長さに対する、RuvC - IIサブドメインの長さの比は1.5以下(例えば、1.4以下)であるか；(vi) RuvC - IIIサブドメインの長さに対する、RuvC - IIサブドメインの長さの比は、1~2(例えば、1.1~2、1.2~2、1~1.8、1.1~1.8、1.2~1.8、1~1.6、1.1~1.6、1.2~1.6、1~1.4、1.1~1.4、または1.2~1.4)の範囲内であるか；(vii) RuvC - IIIサブドメインの長さに対する、RuvC - IIサブドメインとRuvC - IIIサブドメインとの間の領域の長さの比は1より大きいか；(viii) RuvC - IIIサブドメインの長さに対する、RuvC - IIサブドメインとRuvC - IIIサブドメインとの間の領域の長さの比は、1より大きく、かつ1~1.5(例えば、1~1.2)であるか；(ix) RuvC - IIサブドメインとRuvC - IIIサブドメインとの間の領域は、少なくとも60アミノ酸長(例えば、少なくとも65または少なくとも70アミノ酸長)であるか；(x) RuvC - IIサブドメインとRuvC - IIIサブドメインとの間の領域は、少なくとも65アミノ酸長であるか；(xi) RuvC - IIサブドメインとRuvC - IIIサブドメインとの間の領域は、60~110アミノ酸の範囲(例えば、60~105、60~100、60~95、60~90、65~110、65~105、65~100、65~95、または65~90アミノ酸の範囲)の長さを有するか；または(xii) RuvC - IIサブドメインとRuvC - IIIサブドメインとの間の領域は、65~95アミノ酸の範囲の長さを有する。

【0073】

場合によって、(本発明の組成物及び/または方法の) CasYタンパク質は、配列番号1~5として記載されるCasYタンパク質配列のいずれか1つのN末端ドメイン(例えば、図3、パネルAのCasY1においてアミノ酸1~812として示されるドメイン)と85%以上の配列同一性(例えば、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性)を有する第1のアミノ酸配列と、3つの部分RuvCドメイン - RuvC - I、RuvC - II、及びRuvC - IIIを含む、第1のアミノ酸配列に対してC末端側にある第2のアミノ酸配列とを含み、ここで、(i) RuvC - IIIサブドメインの長さに対する、RuvC - IIサブドメインとRuvC - IIIサブドメインとの間の領域の長さの比は1.1以上(例えば、1.2)であるか；(ii) RuvC - IIIサブドメインの長さに対する、RuvC - IIサブドメインとRuvC - IIIサブドメインとの間の領域の長さの比は1より大きいか；(iii) RuvC - IIIサブドメインの長さに対する、RuvC - IIサブドメインとRuvC - IIIサブドメインとの間の領域の長さの比は、1より大きく、かつ1~1.5(例えば、1~1.4、1~1.3、または1~1.2)であるか；(iv) RuvC - IIIサブドメインの長さに対する、RuvC - IIサブドメインの長さの比は2以下(例えば、1.8以下、1.7以下、1.6以下、1.5以下、または1.4以下)であるか；(v) RuvC - IIIサブドメインの長さに対する、RuvC - IIサブドメインの長さの比は1.5以下(例えば、1.4以下)であるか；(vi) RuvC - IIIサブドメインの長さに対する、RuvC - IIサブドメインの長さの比は、1~2(例えば、1.1~2、1.2~2、1~1.8、1.1~1.8、1.2~1.8、1~1.6、1.1~1.6、1.2~1.6、1~1.4、1.1~1.4、または1.2~1.4)の範囲内であるか；(vii) RuvC - IIIサブドメインの長さに対する、RuvC - IIサブドメインとRuvC - IIIサブドメインとの間の領域の長さの比は1より大き

いか；(viii) RuvC - IIIサブドメインの長さに対する、RuvC - IIサブドメインとRuvC - IIIサブドメインとの間の領域の長さの比は、1より大きく、かつ1~1.5(例えば、1~1.2)であるか；(ix) RuvC - IIサブドメインとRuvC - IIIサブドメインとの間の領域は、少なくとも60アミノ酸長(例えば、少なくとも65または少なくとも70アミノ酸長)であるか；(x) RuvC - IIサブドメインとRuvC - IIIサブドメインとの間の領域は、少なくとも65アミノ酸長であるか；(xi) RuvC - IIサブドメインとRuvC - IIIサブドメインとの間の領域は、60~110アミノ酸の範囲(例えば、60~105、60~100、60~95、60~90、65~110、65~105、65~100、65~95、または65~90アミノ酸の範囲)の長さを有するか；または(xii) RuvC - IIサブドメインとRuvC - IIIサブドメインとの間の領域は、65~95アミノ酸の範囲の長さを有する。

10

【0074】

場合によって、(本発明の組成物及び/または方法の) CasYタンパク質は、配列番号1~7として記載されるCasYタンパク質配列のいずれか1つのN末端ドメイン(例えば、図3、パネルAのCasY1においてアミノ酸1~812として示されるドメイン)と50%以上の配列同一性(例えば、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性)を有する第1のアミノ酸配列と、3つの部分RuvCドメイン - RuvC - I、RuvC - II、及びRuvC - IIIを含む、第1のアミノ酸配列に対してC末端側にある第2のアミノ酸配列とを含み、ここで、(i) RuvC - IIIサブドメインの長さに対する、RuvC - IIサブドメインとRuvC - IIIサブドメインとの間の領域の長さの比は1.1以上(例えば、1.2)であるか；(ii) RuvC - IIIサブドメインの長さに対する、RuvC - IIサブドメインとRuvC - IIIサブドメインとの間の領域の長さの比は1より大きいか；(iii) RuvC - IIIサブドメインの長さに対する、RuvC - IIサブドメインとRuvC - IIIサブドメインとの間の領域の長さの比は、1より大きく、かつ1~1.5(例えば、1~1.4、1~1.3、または1~1.2)であるか；(iv) RuvC - IIIサブドメインの長さに対する、RuvC - IIサブドメインの長さの比は2以下(例えば、1.8以下、1.7以下、1.6以下、1.5以下、または1.4以下)であるか；(v) RuvC - IIIサブドメインの長さに対する、RuvC - IIサブドメインの長さの比は1.5以下(例えば、1.4以下)であるか；(vi) RuvC - IIIサブドメインの長さに対する、RuvC - IIサブドメインの長さの比は、1~2(例えば、1.1~2、1.2~2、1~1.8、1.1~1.8、1.2~1.8、1~1.6、1.1~1.6、1.2~1.6、1~1.4、1.1~1.4、または1.2~1.4)の範囲内であるか；(vii) RuvC - IIIサブドメインの長さに対する、RuvC - IIサブドメインとRuvC - IIIサブドメインとの間の領域の長さの比は1より大きいか；(viii) RuvC - IIIサブドメインの長さに対する、RuvC - IIサブドメインとRuvC - IIIサブドメインとの間の領域の長さの比は、1より大きく、かつ1~1.5(例えば、1~1.2)であるか；(ix) RuvC - IIサブドメインとRuvC - IIIサブドメインとの間の領域は、少なくとも60アミノ酸長(例えば、少なくとも65または少なくとも70アミノ酸長)であるか；(x) RuvC - IIサブドメインとRuvC - IIIサブドメインとの間の領域は、少なくとも65アミノ酸長であるか；(xi) RuvC - IIサブドメインとRuvC - IIIサブドメインとの間の領域は、60~110アミノ酸の範囲(例えば、60~105、60~100、60~95、60~90、65~110、65~105、65~100、65~95、または65~90アミノ酸の範囲)の長さを有するか；または(xii) RuvC - IIサブドメインとRuvC - IIIサブドメインとの間の領域は、65~95アミノ酸の範囲の長さを有する。

20

30

40

【0075】

場合によって、(本発明の組成物及び/または方法の) CasYタンパク質は、配列番

50

号1~7として記載されるCasYタンパク質配列のいずれか1つのN末端ドメイン(例えば、図3、パネルAのCasY1においてアミノ酸1~812として示されるドメイン)と75%以上の配列同一性(例えば、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性)を有する第1のアミノ酸配列と、3つの部分RuvCドメイン-RuvC-I、RuvC-II、及びRuvC-IIIを含む、第1のアミノ酸配列に対してC末端側にある第2のアミノ酸配列とを含み、ここで、(i)RuvC-IIIサブドメインの長さに対する、RuvC-IIサブドメインとRuvC-IIIサブドメインとの間の領域の長さの比は1.1以上(例えば、1.2)であるか；(ii)RuvC-IIIサブドメインの長さに対する、RuvC-IIサブドメインとRuvC-IIIサブドメインとの間の領域の長さの比は1より大きいか；(iii)RuvC-IIIサブドメインの長さに対する、RuvC-IIサブドメインとRuvC-IIIサブドメインとの間の領域の長さの比は、1より大きく、かつ1~1.5(例えば、1~1.4、1~1.3、または1~1.2)であるか；(iv)RuvC-IIIサブドメインの長さに対する、RuvC-IIサブドメインの長さの比は2以下(例えば、1.8以下、1.7以下、1.6以下、1.5以下、または1.4以下)であるか；(v)RuvC-IIIサブドメインの長さに対する、RuvC-IIサブドメインの長さの比は1.5以下(例えば、1.4以下)であるか；(vi)RuvC-IIIサブドメインの長さに対する、RuvC-IIサブドメインの長さの比は、1~2(例えば、1.1~2、1.2~2、1~1.8、1.1~1.8、1.2~1.8、1~1.6、1.1~1.6、1.2~1.6、1~1.4、1.1~1.4、または1.2~1.4)の範囲内であるか；(vii)RuvC-IIIサブドメインの長さに対する、RuvC-IIサブドメインとRuvC-IIIサブドメインとの間の領域の長さの比は1より大きいか；(viii)RuvC-IIIサブドメインの長さに対する、RuvC-IIサブドメインとRuvC-IIIサブドメインとの間の領域の長さの比は、1より大きく、かつ1~1.5(例えば、1~1.2)であるか；(ix)RuvC-IIサブドメインとRuvC-IIIサブドメインとの間の領域は、少なくとも60アミノ酸長(例えば、少なくとも65または少なくとも70アミノ酸長)であるか；(x)RuvC-IIサブドメインとRuvC-IIIサブドメインとの間の領域は、少なくとも65アミノ酸長であるか；(xi)RuvC-IIサブドメインとRuvC-IIIサブドメインとの間の領域は、60~110アミノ酸の範囲(例えば、60~105、60~100、60~95、60~90、65~110、65~105、65~100、65~95、または65~90アミノ酸の範囲)の長さを有するか；または(xii)RuvC-IIサブドメインとRuvC-IIIサブドメインとの間の領域は、65~95アミノ酸の範囲の長さを有する。

【0076】

場合によって、(本発明の組成物及び/または方法の)CasYタンパク質は、配列番号1~7として記載されるCasYタンパク質配列のいずれか1つのN末端ドメイン(例えば、図3、パネルAのCasY1においてアミノ酸1~812として示されるドメイン)と85%以上の配列同一性(例えば、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性)を有する第1のアミノ酸配列と、3つの部分RuvCドメイン-RuvC-I、RuvC-II、及びRuvC-IIIを含む、第1のアミノ酸配列に対してC末端側にある第2のアミノ酸配列とを含み、ここで、(i)RuvC-IIIサブドメインの長さに対する、RuvC-IIサブドメインとRuvC-IIIサブドメインとの間の領域の長さの比は1.1以上(例えば、1.2)であるか；(ii)RuvC-IIIサブドメインの長さに対する、RuvC-IIサブドメインとRuvC-IIIサブドメインとの間の領域の長さの比は1より大きいか；(iii)RuvC-IIIサブドメインの長さに対する、RuvC-IIサブドメインとRuvC-IIIサブドメインとの間の領域の長さの比は、1より大きく、かつ1~1.5(例えば、1~1.4、1~1.3、または1~1.2)であるか；(iv)RuvC-IIIサブドメインの長さに対する、RuvC-IIサブドメインの長さの比は2以下(例え

ば、1.8以下、1.7以下、1.6以下、1.5以下、または1.4以下)であるか；
 (v) R u v C - I I Iサブドメインの長さに対する、R u v C - I Iサブドメインの長さの比は1.5以下(例えば、1.4以下)であるか；(vi) R u v C - I I Iサブドメインの長さに対する、R u v C - I Iサブドメインの長さの比は、1~2(例えば、1.1~2、1.2~2、1~1.8、1.1~1.8、1.2~1.8、1~1.6、1.1~1.6、1.2~1.6、1~1.4、1.1~1.4、または1.2~1.4)の範囲内であるか；(vii) R u v C - I I Iサブドメインの長さに対する、R u v C - I IサブドメインとR u v C - I I Iサブドメインとの間の領域の長さの比は1より大きいか；(viii) R u v C - I I Iサブドメインの長さに対する、R u v C - I IサブドメインとR u v C - I I Iサブドメインとの間の領域の長さの比は、1より大きく、かつ1~1.5(例えば、1~1.2)であるか；(ix) R u v C - I IサブドメインとR u v C - I I Iサブドメインとの間の領域は、少なくとも60アミノ酸長(例えば、少なくとも65または少なくとも70アミノ酸長)であるか；(x) R u v C - I IサブドメインとR u v C - I I Iサブドメインとの間の領域は、少なくとも65アミノ酸長であるか；(xi) R u v C - I IサブドメインとR u v C - I I Iサブドメインとの間の領域は、60~110アミノ酸の範囲(例えば、60~105、60~100、60~95、60~90、65~110、65~105、65~100、65~95、または65~90アミノ酸の範囲)の長さを有するか；または(xii) R u v C - I IサブドメインとR u v C - I I Iサブドメインとの間の領域は、65~95アミノ酸の範囲の長さを有する。

10

20

【0077】

場合によって、(本発明の組成物及び/または方法の)C a s Yタンパク質は、配列番号1~8として記載されるC a s Yタンパク質配列のいずれか1つのN末端ドメイン(例えば、図3、パネルAのC a s Y1においてアミノ酸1~812として示されるドメイン)と50%以上の配列同一性(例えば、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性)を有する第1のアミノ酸配列と、3つの部分R u v Cドメイン - R u v C - I、R u v C - I I、及びR u v C - I I Iを含む、第1のアミノ酸配列に対してC末端側にある第2のアミノ酸配列とを含み、ここで、(i) R u v C - I I Iサブドメインの長さに対する、R u v C - I IサブドメインとR u v C - I I Iサブドメインとの間の領域の長さの比は1.1以上(例えば、1.2)であるか；(ii) R u v C - I I Iサブドメインの長さに対する、R u v C - I IサブドメインとR u v C - I I Iサブドメインとの間の領域の長さの比は1より大きいか；(iii) R u v C - I I Iサブドメインの長さに対する、R u v C - I IサブドメインとR u v C - I I Iサブドメインとの間の領域の長さの比は、1より大きく、かつ1~1.5(例えば、1~1.4、1~1.3、または1~1.2)であるか；(iv) R u v C - I I Iサブドメインの長さに対する、R u v C - I Iサブドメインの長さの比は2以下(例えば、1.8以下、1.7以下、1.6以下、1.5以下、または1.4以下)であるか；(v) R u v C - I I Iサブドメインの長さに対する、R u v C - I Iサブドメインの長さの比は1.5以下(例えば、1.4以下)であるか；(vi) R u v C - I I Iサブドメインの長さに対する、R u v C - I Iサブドメインの長さの比は、1~2(例えば、1.1~2、1.2~2、1~1.8、1.1~1.8、1.2~1.8、1~1.6、1.1~1.6、1.2~1.6、1~1.4、1.1~1.4、または1.2~1.4)の範囲内であるか；(vii) R u v C - I I Iサブドメインの長さに対する、R u v C - I IサブドメインとR u v C - I I Iサブドメインとの間の領域の長さの比は1より大きいか；(viii) R u v C - I I Iサブドメインの長さに対する、R u v C - I IサブドメインとR u v C - I I Iサブドメインとの間の領域の長さの比は、1より大きく、かつ1~1.5(例えば、1~1.2)であるか；(ix) R u v C - I IサブドメインとR u v C - I I Iサブドメインとの間の領域は、少なくとも60アミノ酸長(例えば、少なくとも65または少なくとも70アミノ酸長)であるか；(x) R u v C - I IサブドメインとR u v C - I I Iサブドメ

30

40

50

インとの間の領域は、少なくとも65アミノ酸長であるか；(xi) RuvC - IIサブドメインとRuvC - IIIサブドメインとの間の領域は、60～110アミノ酸の範囲（例えば、60～105、60～100、60～95、60～90、65～110、65～105、65～100、65～95、または65～90アミノ酸の範囲）の長さを有するか；または(xii) RuvC - IIサブドメインとRuvC - IIIサブドメインとの間の領域は、65～95アミノ酸の範囲の長さを有する。

【0078】

場合によって、（本発明の組成物及び/または方法の）CasYタンパク質は、配列番号1～8として記載されるCasYタンパク質配列のいずれか1つのN末端ドメイン（例えば、図3、パネルAのCasY1においてアミノ酸1～812として示されるドメイン）と75%以上の配列同一性（例えば、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有する第1のアミノ酸配列と、3つの部分RuvCドメイン - RuvC - I、RuvC - II、及びRuvC - IIIを含む、第1のアミノ酸配列に対してC末端側にある第2のアミノ酸配列とを含み、ここで、(i) RuvC - IIIサブドメインの長さに対する、RuvC - IIサブドメインとRuvC - IIIサブドメインとの間の領域の長さの比は1.1以上（例えば、1.2）であるか；(ii) RuvC - IIIサブドメインの長さに対する、RuvC - IIサブドメインとRuvC - IIIサブドメインとの間の領域の長さの比は1より大きいか；(iii) RuvC - IIIサブドメインの長さに対する、RuvC - IIサブドメインとRuvC - IIIサブドメインとの間の領域の長さの比は、1より大きく、かつ1～1.5（例えば、1～1.4、1～1.3、または1～1.2）であるか；(iv) RuvC - IIIサブドメインの長さに対する、RuvC - IIサブドメインの長さの比は2以下（例えば、1.8以下、1.7以下、1.6以下、1.5以下、または1.4以下）であるか；(v) RuvC - IIIサブドメインの長さに対する、RuvC - IIサブドメインの長さの比は1.5以下（例えば、1.4以下）であるか；(vi) RuvC - IIIサブドメインの長さに対する、RuvC - IIサブドメインの長さの比は、1～2（例えば、1.1～2、1.2～2、1～1.8、1.1～1.8、1.2～1.8、1～1.6、1.1～1.6、1.2～1.6、1～1.4、1.1～1.4、または1.2～1.4）の範囲内であるか；(vii) RuvC - IIIサブドメインの長さに対する、RuvC - IIサブドメインとRuvC - IIIサブドメインとの間の領域の長さの比は1より大きいか；(viii) RuvC - IIIサブドメインの長さに対する、RuvC - IIサブドメインとRuvC - IIIサブドメインとの間の領域の長さの比は、1より大きく、かつ1～1.5（例えば、1～1.2）であるか；(ix) RuvC - IIサブドメインとRuvC - IIIサブドメインとの間の領域は、少なくとも60アミノ酸長（例えば、少なくとも65または少なくとも70アミノ酸長）であるか；(x) RuvC - IIサブドメインとRuvC - IIIサブドメインとの間の領域は、少なくとも65アミノ酸長であるか；(xi) RuvC - IIサブドメインとRuvC - IIIサブドメインとの間の領域は、60～110アミノ酸の範囲（例えば、60～105、60～100、60～95、60～90、65～110、65～105、65～100、65～95、または65～90アミノ酸の範囲）の長さを有するか；または(xii) RuvC - IIサブドメインとRuvC - IIIサブドメインとの間の領域は、65～95アミノ酸の範囲の長さを有する。

【0079】

場合によって、（本発明の組成物及び/または方法の）CasYタンパク質は、配列番号1～8として記載されるCasYタンパク質配列のいずれか1つのN末端ドメイン（例えば、図3、パネルAのCasY1においてアミノ酸1～812として示されるドメイン）と85%以上の配列同一性（例えば、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有する第1のアミノ酸配列と、3つの部分RuvCドメイン - RuvC - I、RuvC - II、及びRuvC - IIIを含む、第1のアミノ酸配列に対してC末端側にある第2のアミノ酸配列とを含み、ここで、(i

10

20

30

40

50

) R u v C - I I Iサブドメインの長さに対する、R u v C - I IサブドメインとR u v C - I I Iサブドメインとの間の領域の長さの比は1.1以上(例えば、1.2)であるか；(i i) R u v C - I I Iサブドメインの長さに対する、R u v C - I IサブドメインとR u v C - I I Iサブドメインとの間の領域の長さの比は1より大きいか；(i i i) R u v C - I I Iサブドメインの長さに対する、R u v C - I IサブドメインとR u v C - I I Iサブドメインとの間の領域の長さの比は、1より大きく、かつ1~1.5(例えば、1~1.4、1~1.3、または1~1.2)であるか；(i v) R u v C - I I Iサブドメインの長さに対する、R u v C - I Iサブドメインの長さの比は2以下(例えば、1.8以下、1.7以下、1.6以下、1.5以下、または1.4以下)であるか；(v) R u v C - I I Iサブドメインの長さに対する、R u v C - I Iサブドメインの長さの比は1.5以下(例えば、1.4以下)であるか；(v i) R u v C - I I Iサブドメインの長さに対する、R u v C - I Iサブドメインの長さの比は、1~2(例えば、1.1~2、1.2~2、1~1.8、1.1~1.8、1.2~1.8、1~1.6、1.1~1.6、1.2~1.6、1~1.4、1.1~1.4、または1.2~1.4)の範囲内であるか；(v i i) R u v C - I I Iサブドメインの長さに対する、R u v C - I IサブドメインとR u v C - I I Iサブドメインとの間の領域の長さの比は1より大きいか；(v i i i) R u v C - I I Iサブドメインの長さに対する、R u v C - I IサブドメインとR u v C - I I Iサブドメインとの間の領域の長さの比は、1より大きく、かつ1~1.5(例えば、1~1.2)であるか；(i x) R u v C - I IサブドメインとR u v C - I I Iサブドメインとの間の領域は、少なくとも60アミノ酸長(例えば、少なくとも65または少なくとも70アミノ酸長)であるか；(x) R u v C - I IサブドメインとR u v C - I I Iサブドメインとの間の領域は、少なくとも65アミノ酸長であるか；(x i) R u v C - I IサブドメインとR u v C - I I Iサブドメインとの間の領域は、60~110アミノ酸の範囲(例えば、60~105、60~100、60~95、60~90、65~110、65~105、65~100、65~95、または65~90アミノ酸の範囲)の長さを有するか；または(x i i) R u v C - I IサブドメインとR u v C - I I Iサブドメインとの間の領域は、65~95アミノ酸の範囲の長さを有する。

10

20

【0080】

場合によって、(本発明の組成物及び/または方法の) C a s Yタンパク質は、750~1050アミノ酸(例えば、750~1025、750~1000、750~950、775~1050、775~1025、775~1000、775~950、800~1050、800~1025、800~1000、または800~950アミノ酸)の範囲の長さを有するN末端ドメインをもつ第1のアミノ酸配列(例えば、NLS及び/または触媒活性を有するドメインのような何らかの融合された異種配列は含まない)；ならびに3つの部分R u v Cドメイン - R u v C - I、R u v C - I I、及びR u v C - I I Iを有する分割R u v Cドメインをもつ(第1のアミノ酸配列に対してC末端側にある)第2のアミノ酸配列を含み、ここで、(i) R u v C - I I Iサブドメインの長さに対する、R u v C - I IサブドメインとR u v C - I I Iサブドメインとの間の領域の長さの比は1.1以上(例えば、1.2)であるか；(i i) R u v C - I I Iサブドメインの長さに対する、R u v C - I IサブドメインとR u v C - I I Iサブドメインとの間の領域の長さの比は1より大きいか；(i i i) R u v C - I I Iサブドメインの長さに対する、R u v C - I IサブドメインとR u v C - I I Iサブドメインとの間の領域の長さの比は、1より大きく、かつ1~1.5(例えば、1~1.4、1~1.3、または1~1.2)であるか；(i v) R u v C - I I Iサブドメインの長さに対する、R u v C - I Iサブドメインの長さの比は2以下(例えば、1.8以下、1.7以下、1.6以下、1.5以下、または1.4以下)であるか；(v) R u v C - I I Iサブドメインの長さに対する、R u v C - I Iサブドメインの長さの比は1.5以下(例えば、1.4以下)であるか；(v i) R u v C - I I Iサブドメインの長さに対する、R u v C - I Iサブドメインの長さの比は、1~2(例えば、1.1~2、1.2~2、1~1.8、1.1~1

30

40

50

． 8、 1． 2 ~ 1． 8、 1 ~ 1． 6、 1． 1 ~ 1． 6、 1． 2 ~ 1． 6、 1 ~ 1 4、 1． 1 ~ 1． 4、 または 1． 2 ~ 1． 4) の範囲内であるか ; (v i i) R u v C - I I I サブドメインの長さに対する、 R u v C - I I サブドメインと R u v C - I I I サブドメインとの間の領域の長さの比は 1 より大きいか ; (v i i i) R u v C - I I I サブドメインの長さに対する、 R u v C - I I サブドメインと R u v C - I I I サブドメインとの間の領域の長さの比は、 1 より大きく、かつ 1 ~ 1． 5 (例えば、 1 ~ 1． 2) であるか ; (i x) R u v C - I I サブドメインと R u v C - I I I サブドメインとの間の領域は、少なくとも 6 0 アミノ酸長 (例えば、少なくとも 6 5 または少なくとも 7 0 アミノ酸長) であるか ; (x) R u v C - I I サブドメインと R u v C - I I I サブドメインとの間の領域は、少なくとも 6 5 アミノ酸長であるか ; (x i) R u v C - I I サブドメインと R u v C - I I I サブドメインとの間の領域は、 6 0 ~ 1 1 0 アミノ酸の範囲 (例えば、 6 0 ~ 1 0 5、 6 0 ~ 1 0 0、 6 0 ~ 9 5、 6 0 ~ 9 0、 6 5 ~ 1 1 0、 6 5 ~ 1 0 5、 6 5 ~ 1 0 0、 6 5 ~ 9 5、 または 6 5 ~ 9 0 アミノ酸の範囲) の長さを有するか ; または (x i i) R u v C - I I サブドメインと R u v C - I I I サブドメインとの間の領域は、 6 5 ~ 9 5 アミノ酸の範囲の長さを有する。

10

【 0 0 8 1 】

場合によって、 (本発明の組成物及び / または方法の) C a s Y タンパク質は、配列番号 1 として記載される C a s Y タンパク質配列の C 末端ドメイン (例えば、図 3、パネル A の C a s Y 1 においてアミノ酸 8 1 2 ~ 1 1 2 5 として示されるドメイン) と 2 0 % 以上の配列同一性 (例えば、 3 0 % 以上、 4 0 % 以上、 5 0 % 以上、 6 0 % 以上、 7 0 % 以上、 8 0 % 以上、 8 5 % 以上、 9 0 % 以上、 9 5 % 以上、 9 7 % 以上、 9 8 % 以上、 9 9 % 以上、または 1 0 0 % の配列同一性) を有するアミノ酸配列を含む。例えば、場合によって、 C a s Y タンパク質は、配列番号 1 として記載される C a s Y タンパク質配列の C 末端ドメイン (例えば、図 3、パネル A の C a s Y 1 においてアミノ酸 8 1 2 ~ 1 1 2 5 として示されるドメイン) と 5 0 % 以上の配列同一性 (例えば、 6 0 % 以上、 7 0 % 以上、 8 0 % 以上、 8 5 % 以上、 9 0 % 以上、 9 5 % 以上、 9 7 % 以上、 9 8 % 以上、 9 9 % 以上、または 1 0 0 % の配列同一性) を有するアミノ酸配列を含む。場合によって、 C a s Y タンパク質は、配列番号 1 として記載される C a s Y タンパク質配列の C 末端ドメイン (例えば、図 3、パネル A の C a s Y 1 においてアミノ酸 8 1 2 ~ 1 1 2 5 として示されるドメイン) と 8 0 % 以上の配列同一性 (例えば、 8 5 % 以上、 9 0 % 以上、 9 5 % 以上、 9 7 % 以上、 9 8 % 以上、 9 9 % 以上、または 1 0 0 % の配列同一性) を有するアミノ酸配列を含む。場合によって、 C a s Y タンパク質は、配列番号 1 として記載される C a s Y タンパク質配列の C 末端ドメイン (例えば、図 3、パネル A の C a s Y 1 においてアミノ酸 8 1 2 ~ 1 1 2 5 として示されるドメイン) と 9 0 % 以上の配列同一性 (例えば、 9 5 % 以上、 9 7 % 以上、 9 8 % 以上、 9 9 % 以上、または 1 0 0 % の配列同一性) を有するアミノ酸配列を含む。場合によって、 C a s Y タンパク質は、配列番号 1 として記載される C a s Y タンパク質配列のアミノ酸 8 1 2 ~ 1 1 2 5 を有するアミノ酸配列を含む。

20

30

【 0 0 8 2 】

場合によって、 (本発明の組成物及び / または方法の) C a s Y タンパク質は、配列番号 1 ~ 4 として記載される C a s Y タンパク質配列のいずれか 1 つの C 末端ドメイン (例えば、図 3、パネル A の C a s Y 1 においてアミノ酸 8 1 2 ~ 1 1 2 5 として示されるドメイン) と 2 0 % 以上の配列同一性 (例えば、 3 0 % 以上、 4 0 % 以上、 5 0 % 以上、 6 0 % 以上、 7 0 % 以上、 8 0 % 以上、 8 5 % 以上、 9 0 % 以上、 9 5 % 以上、 9 7 % 以上、 9 8 % 以上、 9 9 % 以上、または 1 0 0 % の配列同一性) を有するアミノ酸配列を含む。例えば、場合によって、 C a s Y タンパク質は、配列番号 1 ~ 4 として記載される C a s Y タンパク質配列のいずれか 1 つの C 末端ドメイン (例えば、図 3、パネル A の C a s Y 1 においてアミノ酸 8 1 2 ~ 1 1 2 5 として示されるドメイン) と 5 0 % 以上の配列同一性 (例えば、 6 0 % 以上、 7 0 % 以上、 8 0 % 以上、 8 5 % 以上、 9 0 % 以上、 9 5 % 以上、 9 7 % 以上、 9 8 % 以上、 9 9 % 以上、または 1 0 0 % の配列同一性) を有するア

40

50

ミノ酸配列を含む。場合によって、C a s Yタンパク質は、配列番号1～4として記載されるC a s Yタンパク質配列のいずれか1つのC末端ドメイン（例えば、図3、パネルAのC a s Y 1においてアミノ酸812～1125として示されるドメイン）と80%以上の配列同一性（例えば、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有するアミノ酸配列を含む。場合によって、C a s Yタンパク質は、配列番号1～4として記載されるC a s Yタンパク質配列のいずれか1つのC末端ドメイン（例えば、図3、パネルAのC a s Y 1においてアミノ酸812～1125として示されるドメイン）と90%以上の配列同一性（例えば、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有するアミノ酸配列を含む。場合によって、C a s Yタンパク質は、配列番号1として記載されるC a s Yタンパク質配列のアミノ酸812～1125に対応する、配列番号1～4として記載されるC a s Yタンパク質配列のいずれか1つのアミノ酸配列の断片を含む。

10

【0083】

場合によって、（本発明の組成物及び/または方法の）C a s Yタンパク質は、配列番号1～5として記載されるC a s Yタンパク質配列のいずれか1つのC末端ドメイン（例えば、図3、パネルAのC a s Y 1においてアミノ酸812～1125として示されるドメイン）と20%以上の配列同一性（例えば、30%以上、40%以上、50%以上、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有するアミノ酸配列を含む。例えば、場合によって、C a s Yタンパク質は、配列番号1～5として記載されるC a s Yタンパク質配列のいずれか1つのC末端ドメイン（例えば、図3、パネルAのC a s Y 1においてアミノ酸812～1125として示されるドメイン）と50%以上の配列同一性（例えば、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有するアミノ酸配列を含む。場合によって、C a s Yタンパク質は、配列番号1～5として記載されるC a s Yタンパク質配列のいずれか1つのC末端ドメイン（例えば、図3、パネルAのC a s Y 1においてアミノ酸812～1125として示されるドメイン）と80%以上の配列同一性（例えば、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有するアミノ酸配列を含む。場合によって、C a s Yタンパク質は、配列番号1～5として記載されるC a s Yタンパク質配列のいずれか1つのC末端ドメイン（例えば、図3、パネルAのC a s Y 1においてアミノ酸812～1125として示されるドメイン）と90%以上の配列同一性（例えば、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有するアミノ酸配列を含む。場合によって、C a s Yタンパク質は、配列番号1として記載されるC a s Yタンパク質配列のアミノ酸812～1125に対応する、配列番号1～5として記載されるC a s Yタンパク質配列のいずれか1つのアミノ酸配列の断片を含む。

20

30

【0084】

場合によって、（本発明の組成物及び/または方法の）C a s Yタンパク質は、配列番号1～7として記載されるC a s Yタンパク質配列のいずれか1つのC末端ドメイン（例えば、図3、パネルAのC a s Y 1においてアミノ酸812～1125として示されるドメイン）と20%以上の配列同一性（例えば、30%以上、40%以上、50%以上、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有するアミノ酸配列を含む。例えば、場合によって、C a s Yタンパク質は、配列番号1～7として記載されるC a s Yタンパク質配列のいずれか1つのC末端ドメイン（例えば、図3、パネルAのC a s Y 1においてアミノ酸812～1125として示されるドメイン）と50%以上の配列同一性（例えば、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有するアミノ酸配列を含む。場合によって、C a s Yタンパク質は、配列番号1～7として記載されるC a s Yタンパク質配列のいずれか1つのC末端ドメイン（例えば、図3、パネルA

40

50

の C a s Y 1 においてアミノ酸 8 1 2 ~ 1 1 2 5 として示されるドメイン) と 8 0 % 以上の配列同一性 (例えば、8 5 % 以上、9 0 % 以上、9 5 % 以上、9 7 % 以上、9 8 % 以上、9 9 % 以上、または 1 0 0 % の配列同一性) を有するアミノ酸配列を含む。場合によって、C a s Y タンパク質は、配列番号 1 ~ 7 として記載される C a s Y タンパク質配列のいずれか 1 つの C 末端ドメイン (例えば、図 3、パネル A の C a s Y 1 においてアミノ酸 8 1 2 ~ 1 1 2 5 として示されるドメイン) と 9 0 % 以上の配列同一性 (例えば、9 5 % 以上、9 7 % 以上、9 8 % 以上、9 9 % 以上、または 1 0 0 % の配列同一性) を有するアミノ酸配列を含む。場合によって、C a s Y タンパク質は、配列番号 1 として記載される C a s Y タンパク質配列のアミノ酸 8 1 2 ~ 1 1 2 5 に対応する、配列番号 1 ~ 7 として記載される C a s Y タンパク質配列のいずれか 1 つのアミノ酸配列の断片を含む。

10

【 0 0 8 5 】

場合によって、(本発明の組成物及び/または方法の) C a s Y タンパク質は、配列番号 1 ~ 8 として記載される C a s Y タンパク質配列のいずれか 1 つの C 末端ドメイン (例えば、図 3、パネル A の C a s Y 1 においてアミノ酸 8 1 2 ~ 1 1 2 5 として示されるドメイン) と 2 0 % 以上の配列同一性 (例えば、3 0 % 以上、4 0 % 以上、5 0 % 以上、6 0 % 以上、7 0 % 以上、8 0 % 以上、8 5 % 以上、9 0 % 以上、9 5 % 以上、9 7 % 以上、9 8 % 以上、9 9 % 以上、または 1 0 0 % の配列同一性) を有するアミノ酸配列を含む。例えば、場合によって、C a s Y タンパク質は、配列番号 1 ~ 8 として記載される C a s Y タンパク質配列のいずれか 1 つの C 末端ドメイン (例えば、図 3、パネル A の C a s Y 1 においてアミノ酸 8 1 2 ~ 1 1 2 5 として示されるドメイン) と 5 0 % 以上の配列同一性 (例えば、6 0 % 以上、7 0 % 以上、8 0 % 以上、8 5 % 以上、9 0 % 以上、9 5 % 以上、9 7 % 以上、9 8 % 以上、9 9 % 以上、または 1 0 0 % の配列同一性) を有するアミノ酸配列を含む。場合によって、C a s Y タンパク質は、配列番号 1 ~ 8 として記載される C a s Y タンパク質配列のいずれか 1 つの C 末端ドメイン (例えば、図 3、パネル A の C a s Y 1 においてアミノ酸 8 1 2 ~ 1 1 2 5 として示されるドメイン) と 8 0 % 以上の配列同一性 (例えば、8 5 % 以上、9 0 % 以上、9 5 % 以上、9 7 % 以上、9 8 % 以上、9 9 % 以上、または 1 0 0 % の配列同一性) を有するアミノ酸配列を含む。場合によって、C a s Y タンパク質は、配列番号 1 ~ 8 として記載される C a s Y タンパク質配列のいずれか 1 つの C 末端ドメイン (例えば、図 3、パネル A の C a s Y 1 においてアミノ酸 8 1 2 ~ 1 1 2 5 として示されるドメイン) と 9 0 % 以上の配列同一性 (例えば、9 5 % 以上、9 7 % 以上、9 8 % 以上、9 9 % 以上、または 1 0 0 % の配列同一性) を有するアミノ酸配列を含む。場合によって、C a s Y タンパク質は、配列番号 1 として記載される C a s Y タンパク質配列のアミノ酸 8 1 2 ~ 1 1 2 5 に対応する、配列番号 1 ~ 8 として記載される C a s Y タンパク質配列のいずれか 1 つのアミノ酸配列の断片を含む。

20

30

【 0 0 8 6 】

場合によって、(本発明の組成物及び/または方法の) C a s Y タンパク質は、7 5 0 ~ 1 0 5 0 アミノ酸 (例えば、7 5 0 ~ 1 0 2 5、7 5 0 ~ 1 0 0 0、7 5 0 ~ 9 5 0、7 7 5 ~ 1 0 5 0、7 7 5 ~ 1 0 2 5、7 7 5 ~ 1 0 0 0、7 7 5 ~ 9 5 0、8 0 0 ~ 1 0 5 0、8 0 0 ~ 1 0 2 5、8 0 0 ~ 1 0 0 0、または 8 0 0 ~ 9 5 0 アミノ酸) の範囲の長さを有する第 1 のアミノ酸配列 (N 末端ドメイン) (例えば、N L S 及び/または触媒活性を有するドメインのような何らかの融合された異種配列は含まない) ; 及び配列番号 1 ~ 4 として記載される C a s Y タンパク質配列のいずれか 1 つの C 末端ドメイン (例えば、図 3、パネル A の C a s Y 1 においてアミノ酸 8 1 2 ~ 1 1 2 5 として示されるドメイン) と 2 0 % 以上の配列同一性 (例えば、3 0 % 以上、4 0 % 以上、5 0 % 以上、6 0 % 以上、7 0 % 以上、8 0 % 以上、8 5 % 以上、9 0 % 以上、9 5 % 以上、9 7 % 以上、9 8 % 以上、9 9 % 以上、または 1 0 0 % の配列同一性) を有し、第 1 のアミノ酸配列に対して C 末端側に位置する第 2 のアミノ酸配列を含む。例えば、場合によって、C a s Y タンパク質は、7 5 0 ~ 1 0 5 0 アミノ酸 (例えば、7 5 0 ~ 1 0 2 5、7 5 0 ~ 1 0 0 0、7 5 0 ~ 9 5 0、7 7 5 ~ 1 0 5 0、7 7 5 ~ 1 0 2 5、7 7 5 ~ 1 0 0 0、7 7 5 ~ 9 5 0、8 0 0 ~ 1 0 5 0、8 0 0 ~ 1 0 2 5、8 0 0 ~ 1 0 0 0、または 8 0 0 ~

40

50

950アミノ酸)の範囲の長さを有する第1のアミノ酸配列(N末端ドメイン)(例えば、NLS及び/または触媒活性を有するドメインのような何らかの融合された異種配列は含まない);及び配列番号1~4として記載されるCasYタンパク質配列のいずれか1つのC末端ドメイン(例えば、図3、パネルAのCasY1においてアミノ酸812~1125として示されるドメイン)と50%以上の配列同一性(例えば、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性)を有し、第1のアミノ酸配列に対してC末端側に位置する第2のアミノ酸配列を含む。場合によって、CasYタンパク質は、750~1050アミノ酸(例えば、750~1025、750~1000、750~950、775~1050、775~1025、775~1000、775~950、800~1050、800~1025、800~1000、または800~950アミノ酸)の範囲の長さを有する第1のアミノ酸配列(N末端ドメイン)(例えば、NLS及び/または触媒活性を有するドメインのような何らかの融合された異種配列は含まない);及び配列番号1~4として記載されるCasYタンパク質配列のいずれか1つのC末端ドメイン(例えば、図3、パネルAのCasY1においてアミノ酸812~1125として示されるドメイン)と80%以上の配列同一性(例えば、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性)を有し、第1のアミノ酸配列に対してC末端側に位置する第2のアミノ酸配列を含む。場合によって、CasYタンパク質は、750~1050アミノ酸(例えば、750~1025、750~1000、750~950、775~1050、775~1025、775~1000、775~950、800~1050、800~1025、800~1000、または800~950アミノ酸)の範囲の長さを有する第1のアミノ酸配列(N末端ドメイン)(例えば、NLS及び/または触媒活性を有するドメインのような何らかの融合された異種配列は含まない);及び配列番号1~4として記載されるCasYタンパク質配列のいずれか1つのC末端ドメイン(例えば、図3、パネルAのCasY1においてアミノ酸812~1125として示されるドメイン)と90%以上の配列同一性(例えば、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性)を有し、第1のアミノ酸配列に対してC末端側に位置する第2のアミノ酸配列を含む。場合によって、CasYタンパク質は、750~1050アミノ酸(例えば、750~1025、750~1000、750~950、775~1050、775~1025、775~1000、775~950、800~1050、800~1025、800~1000、または800~950アミノ酸)の範囲の長さを有する第1のアミノ酸配列(N末端ドメイン)(例えば、NLS及び/または触媒活性を有するドメインのような何らかの融合された異種配列は含まない);及び配列番号1として記載されるCasYタンパク質配列のアミノ酸812~1125に対応する、配列番号1~4として記載されるCasYタンパク質配列のいずれか1つのアミノ酸配列の断片を有し、第1のアミノ酸配列に対してC末端側に位置する第2のアミノ酸配列を含む。

【0087】

場合によって、(本発明の組成物及び/または方法の)CasYタンパク質は、750~1050アミノ酸(例えば、750~1025、750~1000、750~950、775~1050、775~1025、775~1000、775~950、800~1050、800~1025、800~1000、または800~950アミノ酸)の範囲の長さを有する第1のアミノ酸配列(N末端ドメイン)(例えば、NLS及び/または触媒活性を有するドメインのような何らかの融合された異種配列は含まない);及び配列番号1~5として記載されるCasYタンパク質配列のいずれか1つのC末端ドメイン(例えば、図3、パネルAのCasY1においてアミノ酸812~1125として示されるドメイン)と20%以上の配列同一性(例えば、30%以上、40%以上、50%以上、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性)を有し、第1のアミノ酸配列に対してC末端側に位置する第2のアミノ酸配列を含む。例えば、場合によって、Cas

Yタンパク質は、750～1050アミノ酸（例えば、750～1025、750～1000、750～950、775～1050、775～1025、775～1000、775～950、800～1050、800～1025、800～1000、または800～950アミノ酸）の範囲の長さを有する第1のアミノ酸配列（N末端ドメイン）（例えば、NLS及び/または触媒活性を有するドメインのような何らかの融合された異種配列は含まない）；及び配列番号1～5として記載されるCasYタンパク質配列のいずれか1つのC末端ドメイン（例えば、図3、パネルAのCasY1においてアミノ酸812～1125として示されるドメイン）と50%以上の配列同一性（例えば、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有し、第1のアミノ酸配列に対してC末端側に位置する第2のアミノ酸配列を含む。場合によって、CasYタンパク質は、750～1050アミノ酸（例えば、750～1025、750～1000、750～950、775～1050、775～1025、775～1000、775～950、800～1050、800～1025、800～1000、または800～950アミノ酸）の範囲の長さを有する第1のアミノ酸配列（N末端ドメイン）（例えば、NLS及び/または触媒活性を有するドメインのような何らかの融合された異種配列は含まない）；及び配列番号1～5として記載されるCasYタンパク質配列のいずれか1つのC末端ドメイン（例えば、図3、パネルAのCasY1においてアミノ酸812～1125として示されるドメイン）と80%以上の配列同一性（例えば、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有し、第1のアミノ酸配列に対してC末端側に位置する第2のアミノ酸配列を含む。場合によって、CasYタンパク質は、750～1050アミノ酸（例えば、750～1025、750～1000、750～950、775～1050、775～1025、775～1000、775～950、800～1050、800～1025、800～1000、または800～950アミノ酸）の範囲の長さを有する第1のアミノ酸配列（N末端ドメイン）（例えば、NLS及び/または触媒活性を有するドメインのような何らかの融合された異種配列は含まない）；及び配列番号1～5として記載されるCasYタンパク質配列のいずれか1つのC末端ドメイン（例えば、図3、パネルAのCasY1においてアミノ酸812～1125として示されるドメイン）と90%以上の配列同一性（例えば、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有し、第1のアミノ酸配列に対してC末端側に位置する第2のアミノ酸配列を含む。場合によって、CasYタンパク質は、750～1050アミノ酸（例えば、750～1025、750～1000、750～950、775～1050、775～1025、775～1000、775～950、800～1050、800～1025、800～1000、または800～950アミノ酸）の範囲の長さを有する第1のアミノ酸配列（N末端ドメイン）（例えば、NLS及び/または触媒活性を有するドメインのような何らかの融合された異種配列は含まない）；及び配列番号1として記載されるCasYタンパク質配列のアミノ酸812～1125に対応する、配列番号1～5として記載されるCasYタンパク質配列のいずれか1つのアミノ酸配列の断片を有し、第1のアミノ酸配列に対してC末端側に位置する第2のアミノ酸配列を含む。

【0088】

場合によって、（本発明の組成物及び/または方法の）CasYタンパク質は、750～1050アミノ酸（例えば、750～1025、750～1000、750～950、775～1050、775～1025、775～1000、775～950、800～1050、800～1025、800～1000、または800～950アミノ酸）の範囲の長さを有する第1のアミノ酸配列（N末端ドメイン）（例えば、NLS及び/または触媒活性を有するドメインのような何らかの融合された異種配列は含まない）；及び配列番号1～7として記載されるCasYタンパク質配列のいずれか1つのC末端ドメイン（例えば、図3、パネルAのCasY1においてアミノ酸812～1125として示されるドメイン）と20%以上の配列同一性（例えば、30%以上、40%以上、50%以上、6

0%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性を有し、第1のアミノ酸配列に対してC末端側に位置する第2のアミノ酸配列を含む。例えば、場合によって、CasYタンパク質は、750~1050アミノ酸（例えば、750~1025、750~1000、750~950、775~1050、775~1025、775~1000、775~950、800~1050、800~1025、800~1000、または800~950アミノ酸）の範囲の長さを有する第1のアミノ酸配列（N末端ドメイン）（例えば、NLS及び/または触媒活性を有するドメインのような何らかの融合された異種配列は含まない）；及び配列番号1~7として記載されるCasYタンパク質配列のいずれか1つのC末端ドメイン（例えば、図3、パネルAのCasY1においてアミノ酸812~1125として示されるドメイン）と50%以上の配列同一性（例えば、60%以上、70%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有し、第1のアミノ酸配列に対してC末端側に位置する第2のアミノ酸配列を含む。場合によって、CasYタンパク質は、750~1050アミノ酸（例えば、750~1025、750~1000、750~950、775~1050、775~1025、775~1000、775~950、800~1050、800~1025、800~1000、または800~950アミノ酸）の範囲の長さを有する第1のアミノ酸配列（N末端ドメイン）（例えば、NLS及び/または触媒活性を有するドメインのような何らかの融合された異種配列は含まない）；及び配列番号1~7として記載されるCasYタンパク質配列のいずれか1つのC末端ドメイン（例えば、図3、パネルAのCasY1においてアミノ酸812~1125として示されるドメイン）と80%以上の配列同一性（例えば、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有し、第1のアミノ酸配列に対してC末端側に位置する第2のアミノ酸配列を含む。場合によって、CasYタンパク質は、750~1050アミノ酸（例えば、750~1025、750~1000、750~950、775~1050、775~1025、775~1000、775~950、800~1050、800~1025、800~1000、または800~950アミノ酸）の範囲の長さを有する第1のアミノ酸配列（N末端ドメイン）（例えば、NLS及び/または触媒活性を有するドメインのような何らかの融合された異種配列は含まない）；及び配列番号1~7として記載されるCasYタンパク質配列のいずれか1つのC末端ドメイン（例えば、図3、パネルAのCasY1においてアミノ酸812~1125として示されるドメイン）と90%以上の配列同一性（例えば、95%以上、97%以上、98%以上、99%以上、または100%の配列同一性）を有し、第1のアミノ酸配列に対してC末端側に位置する第2のアミノ酸配列を含む。場合によって、CasYタンパク質は、750~1050アミノ酸（例えば、750~1025、750~1000、750~950、775~1050、775~1025、775~1000、775~950、800~1050、800~1025、800~1000、または800~950アミノ酸）の範囲の長さを有する第1のアミノ酸配列（N末端ドメイン）（例えば、NLS及び/または触媒活性を有するドメインのような何らかの融合された異種配列は含まない）；及び配列番号1として記載されるCasYタンパク質配列のアミノ酸812~1125に対応する、配列番号1~7として記載されるCasYタンパク質配列のいずれか1つのアミノ酸配列の断片を有し、第1のアミノ酸配列に対してC末端側に位置する第2のアミノ酸配列を含む。

【0089】

場合によって、（本発明の組成物及び/または方法の）CasYタンパク質は、750~1050アミノ酸（例えば、750~1025、750~1000、750~950、775~1050、775~1025、775~1000、775~950、800~1050、800~1025、800~1000、または800~950アミノ酸）の範囲の長さを有する第1のアミノ酸配列（N末端ドメイン）（例えば、NLS及び/または触媒活性を有するドメインのような何らかの融合された異種配列は含まない）；及び配列番

号 1 ~ 8 として記載される C a s Y タンパク質配列のいずれか 1 つの C 末端ドメイン (例えば、図 3、パネル A の C a s Y 1 においてアミノ酸 8 1 2 ~ 1 1 2 5 として示されるドメイン) と 2 0 % 以上の配列同一性 (例えば、3 0 % 以上、4 0 % 以上、5 0 % 以上、6 0 % 以上、7 0 % 以上、8 0 % 以上、8 5 % 以上、9 0 % 以上、9 5 % 以上、9 7 % 以上、9 8 % 以上、9 9 % 以上、または 1 0 0 % の配列同一性) を有し、第 1 のアミノ酸配列に対して C 末端側に位置する第 2 のアミノ酸配列を含む。例えば、場合によって、C a s Y タンパク質は、7 5 0 ~ 1 0 5 0 アミノ酸 (例えば、7 5 0 ~ 1 0 2 5、7 5 0 ~ 1 0 0 0、7 5 0 ~ 9 5 0、7 7 5 ~ 1 0 5 0、7 7 5 ~ 1 0 2 5、7 7 5 ~ 1 0 0 0、7 7 5 ~ 9 5 0、8 0 0 ~ 1 0 5 0、8 0 0 ~ 1 0 2 5、8 0 0 ~ 1 0 0 0、または 8 0 0 ~ 9 5 0 アミノ酸) の範囲の長さを有する第 1 のアミノ酸配列 (N 末端ドメイン) (例えば、N L S 及び / または触媒活性を有するドメインのような何らかの融合された異種配列は含まない) ; 及び配列番号 1 ~ 8 として記載される C a s Y タンパク質配列のいずれか 1 つの C 末端ドメイン (例えば、図 3、パネル A の C a s Y 1 においてアミノ酸 8 1 2 ~ 1 1 2 5 として示されるドメイン) と 5 0 % 以上の配列同一性 (例えば、6 0 % 以上、7 0 % 以上、8 0 % 以上、8 5 % 以上、9 0 % 以上、9 5 % 以上、9 7 % 以上、9 8 % 以上、9 9 % 以上、または 1 0 0 % の配列同一性) を有し、第 1 のアミノ酸配列に対して C 末端側に位置する第 2 のアミノ酸配列を含む。場合によって、C a s Y タンパク質は、7 5 0 ~ 1 0 5 0 アミノ酸 (例えば、7 5 0 ~ 1 0 2 5、7 5 0 ~ 1 0 0 0、7 5 0 ~ 9 5 0、7 7 5 ~ 1 0 5 0、7 7 5 ~ 1 0 2 5、7 7 5 ~ 1 0 0 0、7 7 5 ~ 9 5 0、8 0 0 ~ 1 0 5 0、8 0 0 ~ 1 0 2 5、8 0 0 ~ 1 0 0 0、または 8 0 0 ~ 9 5 0 アミノ酸) の範囲の長さを有する第 1 のアミノ酸配列 (N 末端ドメイン) (例えば、N L S 及び / または触媒活性を有するドメインのような何らかの融合された異種配列は含まない) ; 及び配列番号 1 ~ 8 として記載される C a s Y タンパク質配列のいずれか 1 つの C 末端ドメイン (例えば、図 3、パネル A の C a s Y 1 においてアミノ酸 8 1 2 ~ 1 1 2 5 として示されるドメイン) と 8 0 % 以上の配列同一性 (例えば、8 5 % 以上、9 0 % 以上、9 5 % 以上、9 7 % 以上、9 8 % 以上、9 9 % 以上、または 1 0 0 % の配列同一性) を有し、第 1 のアミノ酸配列に対して C 末端側に位置する第 2 のアミノ酸配列を含む。場合によって、C a s Y タンパク質は、7 5 0 ~ 1 0 5 0 アミノ酸 (例えば、7 5 0 ~ 1 0 2 5、7 5 0 ~ 1 0 0 0、7 5 0 ~ 9 5 0、7 7 5 ~ 1 0 5 0、7 7 5 ~ 1 0 2 5、7 7 5 ~ 1 0 0 0、7 7 5 ~ 9 5 0、8 0 0 ~ 1 0 5 0、8 0 0 ~ 1 0 2 5、8 0 0 ~ 1 0 0 0、または 8 0 0 ~ 9 5 0 アミノ酸) の範囲の長さを有する第 1 のアミノ酸配列 (N 末端ドメイン) (例えば、N L S 及び / または触媒活性を有するドメインのような何らかの融合された異種配列は含まない) ; 及び配列番号 1 ~ 8 として記載される C a s Y タンパク質配列のいずれか 1 つの C 末端ドメイン (例えば、図 3、パネル A の C a s Y 1 においてアミノ酸 8 1 2 ~ 1 1 2 5 として示されるドメイン) と 9 0 % 以上の配列同一性 (例えば、9 5 % 以上、9 7 % 以上、9 8 % 以上、9 9 % 以上、または 1 0 0 % の配列同一性) を有し、第 1 のアミノ酸配列に対して C 末端側に位置する第 2 のアミノ酸配列を含む。場合によって、C a s Y タンパク質は、7 5 0 ~ 1 0 5 0 アミノ酸 (例えば、7 5 0 ~ 1 0 2 5、7 5 0 ~ 1 0 0 0、7 5 0 ~ 9 5 0、7 7 5 ~ 1 0 5 0、7 7 5 ~ 1 0 2 5、7 7 5 ~ 1 0 0 0、7 7 5 ~ 9 5 0、8 0 0 ~ 1 0 5 0、8 0 0 ~ 1 0 2 5、8 0 0 ~ 1 0 0 0、または 8 0 0 ~ 9 5 0 アミノ酸) の範囲の長さを有する第 1 のアミノ酸配列 (N 末端ドメイン) (例えば、N L S 及び / または触媒活性を有するドメインのような何らかの融合された異種配列は含まない) ; 及び配列番号 1 として記載される C a s Y タンパク質配列のアミノ酸 8 1 2 ~ 1 1 2 5 に対応する、配列番号 1 ~ 8 として記載される C a s Y タンパク質配列のいずれか 1 つのアミノ酸配列の断片を有し、第 1 のアミノ酸配列に対して C 末端側に位置する第 2 のアミノ酸配列を含む。

【 0 0 9 0 】

C a s Y 変異体

変異体 C a s Y タンパク質は、対応する野生型 C a s Y タンパク質のアミノ酸配列と比較した場合に、少なくとも 1 つのアミノ酸が異なっている (例えば、欠失、挿入、置換、

融合を有する)アミノ酸配列を有する。二本鎖標的核酸の一方の鎖を切断するが、他方の鎖は切断しないCasYタンパク質を、本明細書で「ニッカーゼ」(例えば、「ニッカーゼCasY」と称する。実質的にヌクレアーゼ活性をもたないCasYタンパク質を、本明細書で不活性CasYタンパク質(「dCasY」と称する(ただし、以下で詳細に説明するキメラCasYタンパク質の場合、融合パートナーである異種ポリペプチドによってヌクレアーゼ活性を得ることができる)。本明細書に記載するCasY変異体タンパク質のいずれでも(例えば、ニッカーゼCasY、dCasY、キメラCasY)、CasY変異体は、上記と同じパラメーター(例えば、存在するドメイン、同一性パーセントなど)をもつCasYタンパク質配列を含み得る。

【0091】

変異体 - 触媒活性

場合によって、CasYタンパク質は、例えば、天然の触媒活性配列と比較して変異された変異体CasYタンパク質であり、対応する天然配列と比較した場合に、減少した切断活性を示す(例えば、90%以下、80%以下、70%以下、60%以下、50%以下、40%以下、または30%以下の切断活性を示す)。場合によって、そのような変異体CasYタンパク質は、触媒「不活性」タンパク質であり(実質的に切断活性をもたない)、「dCasY」と称されることがある。場合によって、変異体CasYタンパク質はニッカーゼである(二本鎖標的核酸、例えば二本鎖標的DNAの一方の鎖のみを切断する)。本明細書で詳細に記載されるように、場合によって、CasYタンパク質(野生型切断活性をもつCasYタンパク質の場合もあれば、減少した切断活性をもつ変異体CasY、例えばdCasYまたはニッカーゼCasYの場合もある)は、目的とする活性(例えば、目的とする触媒活性)をもつ異種ポリペプチドと融合(複合体化)され、融合タンパク質(キメラCasYタンパク質)を形成する。

【0092】

CasY1(配列番号1)に従う付番によれば、CasYの触媒残基は、D828、E914、D1074を含む(配列番号1を示す図1では、これらの残基に下線が引かれている)。(例えば、図2のパネルA及びBのアライメントを参照)。

【0093】

このように、場合によって、CasYタンパク質は、減少した活性をもち、かつ、上記アミノ酸の1つ以上(または任意CasYタンパク質の1つ以上の対応するアミノ酸)が変異している(例えば、アラニンで置換されている)。場合によって、変異体CasYタンパク質は触媒「不活性」タンパク質であり(触媒不活性であり)、「dCasY」と称される。dCasYタンパク質は、活性を与える融合パートナーに融合することができ、場合によって、dCasY(例えば、触媒活性を与える融合パートナーをもたないが、真核細胞に発現するとNLSを有し得るもの)は標的DNAに結合することができ、RNAポリメラーゼを遮断して標的DNAからの翻訳を阻害することができる。場合によって、変異体CasYタンパク質はニッカーゼである(二本鎖標的核酸、例えば二本鎖標的DNAの一方の鎖のみを切断する)。

【0094】

変異体 - キメラCasY(すなわち融合タンパク質)

上記で述べたように、場合によって、CasYタンパク質(野生型切断活性をもつCasYタンパク質の場合もあれば、減少した切断活性をもつ変異体CasY、例えばdCasYまたはニッカーゼCasYの場合もある)は、目的とする活性(例えば、目的とする触媒活性)をもつ異種ポリペプチドと融合(複合体化)され、融合タンパク質(キメラCasYタンパク質)を形成する。CasYタンパク質が融合することができる異種ポリペプチドを本明細書で「融合パートナー」と称する。

【0095】

場合によって、融合パートナーは、標的DNAの転写を調節(例えば、転写を阻害、転写を増加)することができる。例えば、場合によって、融合パートナーは、転写を阻害するタンパク質(またはタンパク質由来のドメイン)である(例えば、転写抑制因子、なら

10

20

30

40

50

びに転写阻害タンパク質の動員、メチル化などの標的DNAの修飾、DNA修飾因子の動員、標的DNAと会合するヒストンの調節、ヒストンのアセチル化及び/またはメチル化を変更するもののようなヒストン修飾因子の動員などを介して機能するタンパク質)。場合によって、融合パートナーは、転写を増加させるタンパク質(またはタンパク質由来のドメイン)である(例えば、転写活性化因子、ならびに転写活性化タンパク質の動員、脱メチル化などの標的DNAの修飾、DNA修飾因子の動員、標的DNAと会合するヒストンの調節、ヒストンのアセチル化及び/またはメチル化を変更するもののようなヒストン修飾因子の動員などを介して作用するタンパク質)。

【0096】

場合によって、キメラCasYタンパク質は、標的核酸を修飾する酵素活性(例えば、ヌクレアーゼ活性、メチルトランスフェラーゼ活性、デメチラーゼ活性、DNA修復活性、DNA損傷活性、脱アミノ化活性、ジスムターゼ活性、アルキル化活性、脱プリン活性、酸化活性、ピリミジン二量体を形成する活性、インテグラーゼ活性、トランスポザーゼ活性、リコンビナーゼ活性、ポリメラーゼ活性、リガーゼ活性、ヘリカーゼ活性、フォトリナーゼ活性、またはグリコシラーゼ活性)をもつ異種ポリペプチドを含む。

10

【0097】

場合によって、キメラCasYタンパク質は、標的核酸と会合するポリペプチド(例えばヒストン)を修飾する酵素活性(例えば、メチルトランスフェラーゼ活性、デメチラーゼ活性、アセチルトランスフェラーゼ活性、デアセチラーゼ活性、キナーゼ活性、ホスファターゼ活性、ユビキチンリガーゼ活性、脱ユビキチン化活性、アデニル化活性、脱アデニル化活性、SUMO化活性、脱SUMO化活性、リボシル化活性、脱リボシル化活性、ミリストイル化活性、または脱ミリストイル化活性)をもつ異種ポリペプチドを含む。

20

【0098】

転写の増加に使用することができるタンパク質(またはその断片)の例として、VP16、VP64、VP48、VP160、p65サブドメイン(例えば、NF-B由来)、ならびにEDLLの活性化ドメイン及び/またはTAL活性化ドメイン(例えば、植物での活性のため)などの転写活性化因子;SET1A、SET1B、MLL1~5、ASH1、SYM2、NSD1などのヒストンリジンメチルトランスフェラーゼ;JHDM2a/b、UTX、JMJD3などのようなヒストンリジンデメチラーゼ;GCN5、PCAF、CBP、p300、TAF1、TIP60/PLIP、MOZ/MYST3、MORF/MYST4、SRC1、ACTR、P160、CLOCKなどのようなヒストンアセチルトランスフェラーゼ;ならびにTen-Eleven Translocation(TET)ジオキシゲナーゼ1(TET1CD)、TET1、DME、DML1、DML2、ROS1などのようなDNAデメチラーゼが挙げられるが、これらに限定されない。

30

【0099】

転写の減少に使用することができるタンパク質(またはその断片)の例として、Kruppel関連ボックス(KRABまたはSKD)などの転写抑制因子;KOX1抑制ドメイン;MadmSIN3相互作用ドメイン(SID);ERF抑制因子ドメイン(ERD)、SRDX抑制ドメイン(例えば、植物での抑制のため)など;Pr-SET7/8、SUV4-20H1、RIZ1などのようなヒストンリジンメチルトランスフェラーゼ;JMJD2A/JHDM3A、JMJD2B、JMJD2C/GASC1、JMJD2D、JARID1A/RBP2、JARID1B/PLU-1、JARID1C/SMCX、JARID1D/SMCYなどのようなヒストンリジンデメチラーゼ;HDAC1、HDAC2、HDAC3、HDAC8、HDAC4、HDAC5、HDAC7、HDAC9、SIRT1、SIRT2、HDAC11などのようなヒストンリジンデアセチラーゼ;HhaI DNA m5cメチルトランスフェラーゼ(M.HhaI)、DNAメチルトランスフェラーゼ1(DNMT1)、DNAメチルトランスフェラーゼ3a(DNMT3a)、DNAメチルトランスフェラーゼ3b(DNMT3b)、METI、DRM3(植物)、ZMET2、CMT1、CMT2(植物)などのようなDNAメチラーゼ;なら

40

50

びにラミンA、ラミンBなどのような周辺動員要素が挙げられるが、これらに限定されない。

【0100】

場合によって、融合パートナーは、標的核酸（例えば、ssRNA、dsRNA、ssDNA、dsDNA）を修飾する酵素活性をもつ。融合パートナーが与えることができる酵素活性の例としては、制限酵素（例えば、FokIヌクレアーゼ）によって得られるようなヌクレアーゼ活性、メチルトランスフェラーゼ（例えば、HhaI DNA m5cメチルトランスフェラーゼ（M.HhaI）、DNAメチルトランスフェラーゼ1（DNMT1）、DNAメチルトランスフェラーゼ3a（DNMT3a）、DNAメチルトランスフェラーゼ3b（DNMT3b）、METI、DRM3（植物）、ZMET2、CMT1、CMT2（植物）など）によって得られるようなメチルトランスフェラーゼ活性；デメチラーゼ（例えば、Ten-Eleven Translocation（TET）ジオキシゲナーゼ1（TET1CD）、TET1、DME、DML1、DML2、ROS1など）によって得られるようなデメチラーゼ活性、DNA修復活性、DNA損傷活性、デアミナーゼ（例えば、ラットAPOBEC1のようなシトシンデアミナーゼ酵素）によって得られるような脱アミノ化活性、ジスムターゼ活性、アルキル化活性、脱プリン活性、酸化活性、ピリミジン二量体を形成する活性、インテグラーゼ及び/またはレゾルパーゼ（例えば、Ginインベルターゼの機能亢進変異体、GinH106YなどのGinインベルターゼ；ヒト免疫不全ウイルス1型インテグラーゼ（IN）；Tn3レゾルパーゼなど）によって得られるようなインテグラーゼ活性、トランスポザーゼ活性、リコンビナーゼ（例えば、Ginリコンビナーゼの触媒ドメイン）によって得られるようなリコンビナーゼ活性、ポリメラーゼ活性、リガーゼ活性、ヘリカーゼ活性、フォトリアーゼ活性、ならびにグリコシラーゼ活性が挙げられるが、これらに限定されない。

10

20

【0101】

場合によって、融合パートナーは、標的核酸（例えば、ssRNA、dsRNA、ssDNA、dsDNA）と会合するタンパク質（例えば、ヒストン、RNA結合タンパク質、DNA結合タンパク質など）を修飾する酵素活性をもつ。融合パートナーが与えることができる（標的核酸と会合するタンパク質を修飾する）酵素活性の例としては、ヒストンメチルトランスフェラーゼ（HMT）（例えば、suppressor of variegation 3-9 homolog 1（SUV39H1、別称KMT1A）、ユークロマチンヒストンリジンメチルトランスフェラーゼ2（G9A、別称KMT1C及びEHMT2）、SUV39H2、ESET/SETDB1など、SET1A、SET1B、MLL1~5、ASH1、SYMD2、NSD1、DOT1L、Pr-SET7/8、SUV4-20H1、EZH2、RIZ1）によって得られるようなメチルトランスフェラーゼ活性、ヒストンデメチラーゼ（例えば、リジンデメチラーゼ1A（KDM1A、別称LSD1）、JHDM2a/b、JMJD2A/JHDM3A、JMJD2B、JMJD2C/GASC1、JMJD2D、JARID1A/RBP2、JARID1B/PLU-1、JARID1C/SMCX、JARID1D/SMCY、UTX、JMJD3など）によって得られるようなデメチラーゼ活性、ヒストンアセチラーゼトランスフェラーゼ（例えば、ヒトアセチルトランスフェラーゼp300、GCN5、PCAF、CBP、TAF1、TIP60/PLIP、MOZ/MYST3、MORF/MYST4、HBO1/MYST2、HMOF/MYST1、SRC1、ACTR、P160、CLOCKなどの触媒コア/断片）によって得られるようなアセチルトランスフェラーゼ活性、ヒストンデアセチラーゼ（例えば、HDAC1、HDAC2、HDAC3、HDAC8、HDAC4、HDAC5、HDAC7、HDAC9、SIRT1、SIRT2、HDAC11など）によって得られるようなデアセチラーゼ活性、キナーゼ活性、ホスファターゼ活性、ユビキチンリガーゼ活性、脱ユビキチン化活性、アデニル化活性、脱アデニル化活性、SUMO化活性、脱SUMO化活性、リボシル化活性、脱リボシル化活性、ミリストイル化活性、ならびに脱ミリストイル化活性が挙げられるが、これらに限定されない。

30

40

【0102】

50

好適な融合パートナーのさらなる例は、(例えば、化学的に制御可能なキメラ Cas Y タンパク質を生成するための)ジヒドロ葉酸レダクターゼ(DHFR)不安定化ドメイン、及び葉緑体輸送ペプチドである。好適な葉緑体輸送ペプチドとして、以下が挙げられるが、これらに限定されない。

【0103】

M A S M I S S S A V T T V S R A S R G Q S A A M A P F G G L K S M T G F P V R
 K V N T D I T S I T S N G G R V K C M Q V W P P I G K K K F E T L S Y L P P L T
 R D S R A (配列番号83); M A S M I S S S A V T T V S R A S R G Q S A A M A P
 F G G L K S M T G F P V R K V N T D I T S I T S N G G R V K S (配列番号84);
 M A S S M L S S A T M V A S P A Q A T M V A P F N G L K S S A A F P A T R K A N 10
 N D I T S I T S N G G R V N C M Q V W P P I E K K K F E T L S Y L P D L T D S G
 G R V N C (配列番号85); M A Q V S R I C N G V Q N P S L I S N L S K S S Q R
 K S P L S V S L K T Q Q H P R A Y P I S S S W G L K K S G M T L I G S E L R P L
 K V M S S V S T A C (配列番号86); M A Q V S R I C N G V W N P S L I S N L S
 K S S Q R K S P L S V S L K T Q Q H P R A Y P I S S S W G L K K S G M T L I G S
 E L R P L K V M S S V S T A C (配列番号87); M A Q I N N M A Q G I Q T L N P
 N S N F H K P Q V P K S S S F L V F G S K K L K N S A N S M L V L K K D S I F M
 Q L F C S F R I S A S V A T A C (配列番号88); M A A L V T S Q L A T S G T V
 L S V T D R F R R P G F Q G L R P R N P A D A A L G M R T V G A S A A P K Q S R
 K P H R F D R R C L S M V V (配列番号89); M A A L T T S Q L A T S A T G F G 20
 I A D R S A P S S L L R H G F Q G L K P R S P A G G D A T S L S V T T S A R A T
 P K Q Q R S V Q R G S R R F P S V V V C (配列番号90); M A S S V L S S A A V
 A T R S N V A Q A N M V A P F T G L K S A A S F P V S R K Q N L D I T S I A S N
 G G R V Q C (配列番号91); M E S L A A T S V F A P S R V A V P A A R A L V R
 A G T V V P T R R T S S T S G T S G V K C S A A V T P Q A S P V I S R S A A A A
 (配列番号92); 及び M G A A A T S M Q S L K F S N R L V P P S R R L S P V P N
 N V T C N N L P K S A A P V R T V K C C A S S W N S T I N G A A A T T N G A S A
 A S S (配列番号93)。

【0104】

場合によって、本開示の Cas Y 融合ポリペプチドは、a) 本開示の Cas Y ポリペプチド; 及び b) 葉緑体輸送ペプチドを含む。したがって、例えば、CRISPR-Cas Y 複合体は、葉緑体を標的とすることができる。場合によって、この標的化は、葉緑体輸送ペプチド(CTP)または色素体輸送ペプチドと呼ばれるN末端伸長の存在によってなされ得る。発現されるポリペプチドが、植物色素体(例えば、葉緑体)において区画化されるようにする場合、細菌源からの染色体導入遺伝子は、発現されるポリペプチドをコードする配列に融合されたCTP配列をコードする配列を有していなければならない。したがって、葉緑体への外因性ポリペプチドの移行は、外因性ポリペプチドをコードするポリヌクレオチドの5'領域に機能可能に連結されているCTP配列をコードするポリヌクレオチド配列によって達成される場合が多い。CTPは、色素体への移行時の処理段階で除去される。しかしながら、処理効率は、ペプチドのNH₂末端のCTP及び近傍配列の 40
 アミノ酸配列によって影響され得る。記載されている葉緑体を標的化するための他の選択肢は、トウモロコシ cab-m7シグナル配列(米国特許第7,022,896号、WO97/41228)、エンドウグルタチオンレダクターゼシグナル配列(WO97/41228)、及びUS2009029861に記載のCTPである。

【0105】

場合によって、本開示の Cas Y 融合ポリペプチドは、a) 本開示の Cas Y ポリペプチド; 及び b) エンドソーム放出ペプチドを含み得る。場合によって、エンドソーム放出ポリペプチドは、アミノ酸配列 G L F X A L L X L L X S L W X L L L X A (配列番号94)を含み、ここで、各Xは独立して、リジン、ヒスチジン、及びアルギニンから選択される。場合によって、エンドソーム放出ポリペプチドは、アミノ酸配列 G L F H A L L H 50

LLHSLWHL L LHA (配列番号95)を含む。

【0106】

(部位特異的標的核酸修飾、転写の調節、及び/または標的タンパク質修飾、例えばヒストン修飾のための) Cas9、ジンクフィンガー、及び/またはTALEタンパク質との融合に関連して使用される上記の融合パートナーのいくつか(及びそれより多く)の例については、例えば、Nomura et al., J Am Chem Soc. 2007 Jul 18; 129(28): 8676-7; Rivenbark et al., Epigenetics. 2012 Apr; 7(4): 350-60; Nucleic Acids Res. 2016 Jul 8; 44(12): 5615-28; Gilbert et al., Cell. 2013 Jul 18; 154(2): 442-51; Kearns et al., Nat Methods. 2015 May; 12(5): 401-3; Mendenhall et al., Nat Biotechnol. 2013 Dec; 31(12): 1133-6; Hilton et al., Nat Biotechnol. 2015 May; 33(5): 510-7; Gordley et al., Proc Natl Acad Sci USA. 2009 Mar 31; 106(13): 5053-8; Akopian et al., Proc Natl Acad Sci USA. 2003 Jul 22; 100(15): 8688-91; Tan et al., J Virol. 2006 Feb; 80(4): 1939-48; Tan et al., Proc Natl Acad Sci USA. 2003 Oct 14; 100(21): 11997-2002; Papworth et al., Proc Natl Acad Sci USA. 2003 Feb 18; 100(4): 1621-6; Sanjana et al., Nat Protoc. 2012 Jan 5; 7(1): 171-92; Beerli et al., Proc Natl Acad Sci USA. 1998 Dec 8; 95(25): 14628-33; Snowden et al., Curr Biol. 2002 Dec 23; 12(24): 2159-66; Xu et al., Xu et al., Cell Discov. 2016 May 3; 2: 16009; Komore et al., Nature. 2016 Apr 20; 533(7603): 420-4; Chaikind et al., Nucleic Acids Res. 2016 Aug 11; Choudhury et al., Oncotarget. 2016 Jun 23; Du et al., Cold Spring Harb Protoc. 2016 Jan 4; Pham et al., Methods Mol Biol. 2016; 1358: 43-57; Balboa et al., Stem Cell Reports. 2015 Sep 8; 5(3): 448-59; Hara et al., Sci Rep. 2015 Jun 9; 5: 11221; Piatek et al., Plant Biotechnol J. 2015 May; 13(4): 578-89; Hu et al., Nucleic Acids Res. 2014 Apr; 42(7): 4375-90; Cheng et al., Cell Res. 2013 Oct; 23(10): 1163-71; 及び Maeder et al., Nat Methods. 2013 Oct; 10(10): 977-9を参照のこと。

【0107】

その他の好適な異種ポリペプチドとして、標的核酸の転写及び/または翻訳を直接的及び/または間接的に増加させるポリペプチド(例えば、転写活性化因子またはその断片、転写活性化因子を動員するタンパク質またはその断片、小分子/薬物応答性転写制御因子及び/または翻訳制御因子、翻訳制御タンパク質など)が挙げられるが、これらに限定されない。転写の増加または減少を果たす異種ポリペプチドの非限定的な例として、転写活性化因子ドメイン及び転写抑制因子ドメインが挙げられる。いくつかのそのような場合には、キメラCasYポリペプチドは、ガイド核酸(ガイドRNA)によって、標的核酸内の特定の場所(すなわち、配列)に対して標的化され、例えば、プロモーター(転写活性化因子の機能を選択的に阻害する)へのRNAポリメラーゼの結合を遮断する、及び/ま

たはクロマチンの局所状態を変化させるなどの遺伝子座特異的制御を及ぼす（例えば、標的核酸の修飾、または標的核酸に会合するポリペプチドの修飾に融合配列が使用される場合）。場合によって、変更（例えば、転写抑制または活性化）は一過性である。場合によって、変更は継承される（例えば、エピジェネティック修飾が、標的核酸または標的核酸に会合するタンパク質、例えばヌクレオソームヒストンに加えられた場合）。

【0108】

ssRNA 標的核酸を標的とする場合に使用される異種ポリペプチドの非限定的な例として、（限定はされないが）スプライシング因子（例えば、RSドメイン）；タンパク質翻訳構成要素（例えば、翻訳開始因子、伸長因子、及び/または放出因子；例えばeIF4G）；RNAメチラーゼ；RNA編集酵素（例えばRNAデアミナーゼ、例えばRNAに作用するアデノシンデアミナーゼ（ADAR）。AからI及び/またはCからUに編集する酵素を含む）；ヘリカーゼ；RNA結合タンパク質などが挙げられる。異種ポリペプチドは、タンパク質全体を含むこともあれば、場合によってタンパク質の断片（例えば、機能ドメイン）を含むこともあると理解される。

10

【0109】

本発明のキメラCasYポリペプチドの異種ポリペプチドは、一過性または不可逆性、直接的または間接的を問わず、ssRNA（本開示の目的では、分子内及び/または分子間二次構造、例えば、ヘアピン、ステムループなどの二本鎖RNA二重鎖を含む）と相互作用できる任意のドメインであってよく、限定はされないが、以下を含む群から選択されるエフェクタードメインを含む：エンドヌクレアーゼ（例えば、SMG5及びSMG6などのタンパク質由来のRNaseIII、CRR22 DYWドメイン、ダイサー、及びPIN（PilT N末端）ドメイン）；RNA切断の刺激を担うタンパク質及びタンパク質ドメイン（例えば、CPSF、CstF、CFIm、及びCFIIm）；エキソヌクレアーゼ（例えば、XRN-1またはエキソヌクレアーゼT）；デアデニラーゼ（例えばHNT3）；ナンセンス介在性のRNA分解を担うタンパク質及びタンパク質ドメイン（例えば、UPF1、UPF2、UPF3、UPF3b、RNP S1、Y14、DEK、REF2、及びSRm160）；RNAの安定化を担うタンパク質及びタンパク質ドメイン（例えばPABP）；翻訳の抑制を担うタンパク質及びタンパク質ドメイン（例えば、Ago2及びAgo4）；翻訳刺激を担うタンパク質及びタンパク質ドメイン（例えばStaufen）；翻訳の調節を担う（例えば、可能である）タンパク質及びタンパク質ドメイン（例えば、開始因子、伸長因子、放出因子などの翻訳因子、例えばeIF4G）；RNAのポリアデニル化を担うタンパク質及びタンパク質ドメイン（例えば、PAP1、GLD-2、及びStar-PAP）；RNAのポリウリジニル化を担うタンパク質及びタンパク質ドメイン（例えばCID1及び末端ウリジル酸トランスフェラーゼ）；RNAの移行を担うタンパク質及びタンパク質ドメイン（例えば、IMP1、ZBP1、She2p、She3p、及びBicaudal-Dに由来）；RNAの核内保持を担うタンパク質及びタンパク質ドメイン（例えばRrp6）；RNAの核外輸送を担うタンパク質及びタンパク質ドメイン（例えば、TAP、NXF1、THO、TREX、REF、及びAlly）；RNAスプライシングの抑制を担うタンパク質及びタンパク質ドメイン（例えば、PTB、Sam68、及びhnRNP A1）；RNAスプライシングの刺激を担うタンパク質及びタンパク質ドメイン（例えばセリン/アルギニンリッチ（SR）ドメイン）；転写効率の低下を担うタンパク質及びタンパク質ドメイン（例えばFUS（TLS））；転写の刺激を担うタンパク質及びタンパク質ドメイン（例えばCDK7及びHIV Tat）。あるいは、エフェクタードメインは、エンドヌクレアーゼ；RNA切断の刺激が可能であるタンパク質及びタンパク質ドメイン；エキソヌクレアーゼ；デアデニラーゼ；ナンセンス介在性のRNA分解活性をもつタンパク質及びタンパク質ドメイン；RNAの安定化が可能であるタンパク質及びタンパク質ドメイン；翻訳の抑制が可能であるタンパク質及びタンパク質ドメイン；翻訳の刺激が可能であるタンパク質及びタンパク質ドメイン；翻訳の調節が可能であるタンパク質及びタンパク質ドメイン（例えば、開始因子、伸長因子、放出因子などの翻訳因子、例えばeIF4G）；RNAのポリアデニル化が可

20

30

40

50

能であるタンパク質及びタンパク質ドメイン；RNAのポリアデニル化が可能であるタンパク質及びタンパク質ドメイン；RNAの移行活性をもつタンパク質及びタンパク質ドメイン；RNAの核内保持が可能であるタンパク質及びタンパク質ドメイン；RNA核外輸送活性をもつタンパク質及びタンパク質ドメイン；RNAスプライシングの抑制が可能であるタンパク質及びタンパク質ドメイン；RNAスプライシングの刺激が可能であるタンパク質及びタンパク質ドメイン；転写効率の低下が可能であるタンパク質及びタンパク質ドメイン；ならびに転写の刺激が可能であるタンパク質及びタンパク質ドメインを含む群から選択してもよい。別の好適な異種ポリペプチドは、WO2012068627（その全体が参照により本明細書に組み込まれる）に詳細に記載されているPUF RNA結合ドメインである。

10

【0110】

キメラCasYポリペプチドにおいて異種ポリペプチドとして（全体またはその断片として）使用できるRNAスプライシング因子のいくつかは、個別の配列特異的RNA結合モジュール及びスプライシングエフェクタードメインをもつモジュール構造を有する。例えば、セリン/アルギニンリッチ（SR）タンパク質ファミリーのメンバーは、mRNA前駆体においてエキソンスプライシングエンハンサー（ESE）に結合するN末端RNA認識モチーフ（RRM）、及びエキソンの取り込みを促進するC末端RSドメインで構成される。別の例として、hnRNPタンパク質hnRNP A1は、RRMドメインを介してエキソンスプライシングサイレンサー（ESS）に結合し、C末端グリシンリッチドメインを介してエキソンの取り込みを阻害する。いくつかのスプライシング因子は、2つの別の部位間の制御配列に結合することにより、スプライス部位（SS）の選択的使用を制御することができる。例えば、ASF/SF2は、ESEを認識し、イントロン近位の部位の使用を促進することができ、一方のhnRNP A1はESSに結合し、イントロン遠位の部位を使用するようにスプライシングをシフトすることができる。このような因子の一用途は、内因性遺伝子、特に疾患関連遺伝子の選択的スプライシングを調節するESFを生成することである。例えば、Bcl-x mRNA前駆体は、相対する機能のタンパク質をコードする2つの選択的5'スプライス部位を有する2つのスプライシングアイソフォームを生成する。長鎖スプライシングアイソフォームBcl-xLは、長寿命の有糸分裂後細胞に発現する強力なアポトーシス阻害因子であり、多くのがん細胞において上方制御され、アポトーシスシグナルから細胞を保護する。短鎖アイソフォームBcl-xSは、アポトーシス促進性アイソフォームであり、代謝回転率の高い（例えば、リンパ球を発生する）細胞に高レベルで発現する。2つのBcl-xスプライシングアイソフォームの比は、エキソン中核領域またはエキソン伸長領域のいずれか（すなわち、2つの選択的5'スプライス部位間）に配置された複数のシスエレメントによって制御される。その他の例については、その全体が参照により本明細書に組み込まれるWO2010075303を参照のこと。

20

30

【0111】

さらに好適な融合パートナーとして、境界要素であるタンパク質（またはその断片）（例えば、CTCF）、周辺動員をもたらすタンパク質及びその断片（例えば、ラミンA、ラミンBなど）、タンパク質ドッキング要素（例えば、FKBP/FRB、Pill1/Aby1など）が挙げられるが、これらに限定されない。

40

【0112】

本発明のキメラCasYポリペプチドに適した様々な追加の異種ポリペプチド（またはその断片）の例としては、以下の出願：PCT特許出願：WO2010075303、WO2012068627、及びWO2013155555に記載されるものが挙げられるが、これらに限定されない（公報は、Cas9などの他のCRISPRエンドヌクレアーゼに関するものであるが、記載される融合パートナーを代わりにCasYと共に使用することができる）。また、例えば、米国特許及び特許出願：第8,906,616号；第8,895,308号；第8,889,418号；第8,889,356号；第8,871,445号；第8,865,406号；第8,795,965号；第8,771,945

50

号；第 8, 697, 359 号；第 20140068797 号；第 20140170753 号；第 20140179006 号；第 20140179770 号；第 20140186843 号；第 20140186919 号；第 20140186958 号；第 20140189896 号；第 20140227787 号；第 20140234972 号；第 20140242664 号；第 20140242699 号；第 20140242700 号；第 20140242702 号；第 20140248702 号；第 20140256046 号；第 20140273037 号；第 20140273226 号；第 20140273230 号；第 20140273231 号；第 20140273232 号；第 20140273233 号；第 20140273234 号；第 20140273235 号；第 20140287938 号；第 20140295556 号；第 20140295557 号；第 20140298547 号；第 20140304853 号；第 20140309487 号；第 20140310828 号；第 20140310830 号；第 20140315985 号；第 20140335063 号；第 20140335620 号；第 20140342456 号；第 20140342457 号；第 20140342458 号；第 20140349400 号；第 20140349405 号；第 20140356867 号；第 20140356956 号；第 20140356958 号；第 20140356959 号；第 20140357523 号；第 20140357530 号；第 20140364333 号；第 20140377868 号で参照することができ、すべての文献はその全体が参照により本明細書に組み込まれる。

10

【0113】

20

場合によって、異種ポリペプチド（融合パートナー）は細胞内移行をもたらす。すなわち、この異種ポリペプチドは、細胞内移行配列（例えば、核を標的とする核移行シグナル（NLS）、融合タンパク質を核に入れないようにする配列、例えば、核外輸送配列（NES）、融合タンパク質を細胞質内に保持したままにする配列、ミトコンドリアを標的とするミトコンドリア移行シグナル、葉緑体を標的とする葉緑体移行シグナル、ER 保持シグナルなど）を含んでいる。いくつかの実施形態では、CasY 融合ポリペプチドは、タンパク質が核を標的にしないように NLS を含まない（例えば、標的核酸が、サイトゾル内に存在する RNA である場合に有利になる可能性がある）。いくつかの実施形態では、異種ポリペプチドは、追跡及び/または精製を容易にするためのタグ（例えば、蛍光タンパク質、例えば、緑色蛍光タンパク質（GFP）、YFP、RFP、CFP、mCherry、tdTomato など；ヒスチジンタグ、例えば、6XHis タグ；ヘマグルチニン（HA）タグ；FLAG タグ；Myc タグなど）を提供することができる（すなわち、異種ポリペプチドは検出可能な標識である）。

30

【0114】

場合によって、CasY タンパク質（例えば、野生型 CasY タンパク質、変異体 CasY タンパク質、キメラ CasY タンパク質、dCasY タンパク質、減少したヌクレアーゼ活性を CasY 部分をもつキメラ CasY タンパク質（例えば、融合パートナーに融合された dCasY タンパク質）など）は、核移行シグナル（NLS）（例えば、場合によって 2 つ以上、3 つ以上、4 つ以上、または 5 つ以上の NLS）を含む（に融合されている）。したがって、場合によって、CasY ポリペプチドは、1 つ以上の NLS（例えば、2 つ以上、3 つ以上、4 つ以上、または 5 つ以上の NLS）を含む。場合によって、1 つ以上の NLS（2 つ以上、3 つ以上、4 つ以上、または 5 つ以上の NLS）は、N 末端及び/または C 末端に、またはその付近（例えば 50 アミノ酸以内）に位置する。場合によって、1 つ以上の NLS（2 つ以上、3 つ以上、4 つ以上、または 5 つ以上の NLS）は、N 末端に、またはその付近（例えば 50 アミノ酸以内）に位置する。場合によって、1 つ以上の NLS（2 つ以上、3 つ以上、4 つ以上、または 5 つ以上の NLS）は、C 末端に、またはその付近（例えば 50 アミノ酸以内）に位置する。場合によって、1 つ以上の NLS（3 つ以上、4 つ以上、または 5 つ以上の NLS）は、N 末端と C 末端の両方に、または両方の付近（例えば 50 アミノ酸以内）に位置する。場合によって、1 つの NLS が N 末端に位置し、1 つの NLS が C 末端に位置している。

40

50

【0115】

場合によって、CasYタンパク質（例えば、野生型CasYタンパク質、変異体CasYタンパク質、キメラCasYタンパク質、dCasYタンパク質、減少したヌクレアーゼ活性をCasY部分をもつキメラCasYタンパク質（例えば、融合パートナーに融合されたdCasYタンパク質）など）は、1～10個のNLS（例えば、1～9、1～8、1～7、1～6、1～5、2～10、2～9、2～8、2～7、2～6、または2～5個のNLS）を含む（に融合されている）。場合によって、CasYタンパク質（例えば、野生型CasYタンパク質、変異体CasYタンパク質、キメラCasYタンパク質、dCasYタンパク質、減少したヌクレアーゼ活性をCasY部分をもつキメラCasYタンパク質（例えば、融合パートナーに融合されたdCasYタンパク質）など）は、2～5個のNLS（例えば、2～4または2～3個のNLS）を含む（に融合されている）。

10

【0116】

NLSの非限定的な例としては、アミノ酸配列PKKKRKV（配列番号96）を有するSV40ウイルスラージT抗原のNLS；ヌクレオプラスミン由来のNLS（例えば、配列KRPAATKKAQAKKKK（配列番号97）を有するヌクレオプラスミン二分型NLS）；アミノ酸配列PAAKRVKLD（配列番号98）またはRQRRNELKRSP（配列番号99）を有するc-myc NLS；配列NQSSNFGPMMKGGNFGGRSSGPYGGGGQYFAKPRNQGGY（配列番号100）を有するhRNPA1 M9 NLS；インポーチン-アルファ由来のIBBDメインの配列RMRI ZFKNKGKDTAELRRRVEVSVELRKAkkDEQILKRRNV（配列番号101）；筋腫タンパク質の配列VSRKRPRP（配列番号102）及びPKKARED（配列番号103）；ヒトp53の配列PQPKKKPL（配列番号104）；マウスc-abl IVの配列SALIKKKKMAP（配列番号105）；インフルエンザウイルスNS1の配列DRLRR（配列番号106）及びPKQKKRK（配列番号107）；肝炎ウイルスデルタ抗原の配列RKLKKKIKKL（配列番号108）；マウスMx1タンパク質の配列REKKKFLKRR（配列番号109）；ヒトポリ（ADPリボース）ポリメラーゼの配列KRKGDEVDGVDEVAKKKSKK（配列番号110）；ステロイドホルモン受容体（ヒト）グルココルチコイドの配列RKLQAGMNL EARKTKK（配列番号111）に由来するNLS配列が挙げられる。一般に、NLS（または複数のNLS）は、真核細胞の核内への検出可能な量のCasYタンパク質の蓄積を誘導するのに十分な強度のものである。核内の蓄積の検出は、任意の好適な技術によって実施することができる。例えば、細胞内での位置を可視化できるように、検出可能なマーカーをCasYタンパク質に融合してもよい。細胞核は細胞から単離することができ、その後、内容物を免疫組織化学法、ウェスタンブロット、または酵素活性アッセイなどのタンパク質の検出に適した任意の方法によって分析することができる。核内の蓄積はまた、間接的に決定することもできる。

20

30

【0117】

場合によって、CasY融合ポリペプチドは、脂質二重層、ミセル、細胞膜、オルガネラ膜、または小胞膜の透過を促進するポリペプチド、ポリヌクレオチド、炭水化物、または有機もしくは無機化合物を指す、「タンパク質導入ドメイン」すなわちPTD（細胞膜透過性ペプチド、CPPとも呼ばれる）を含む。極性小分子から大きな巨大分子及び/またはナノ粒子までを範囲とし得る別の分子に結合したPTDは、例えば、細胞外空間から細胞内空間への、またはサイトゾルからオルガネラ内への移行といった、分子の膜透過を促進する。いくつかの実施形態では、PTDはポリペプチドのアミノ末端に共有結合している（例えば、野生型CasYに結合して融合タンパク質を形成する、またはdCasY、ニッカーゼCasY、もしくはキメラCasYタンパク質などの変異体CasYタンパク質に結合して融合タンパク質を形成する）。いくつかの実施形態では、PTDはポリペプチドのカルボキシル末端に共有結合している（例えば、野生型CasYに結合して融合タンパク質を形成する、またはdCasY、ニッカーゼCasY、もしくはキメラCas

40

50

Yタンパク質などの変異体CasYタンパク質に結合して融合タンパク質を形成する)。場合によって、PTDは、適切な挿入部位でCasY融合ポリペプチド内に内部挿入されている(すなわち、CasY融合ポリペプチドのN末端またはC末端にはない)。場合によって、本発明のCasY融合ポリペプチドは、1つ以上のPTD(例えば、2つ以上、3つ以上、4つ以上のPTD)を含む(PTDと複合体化される、融合される)。場合によって、PTDは、核移行シグナル(NLS)(例えば、場合によって2つ以上、3つ以上、4つ以上、または5つ以上のNLS)を含む。したがって、場合によって、CasY融合ポリペプチドは、1つ以上のNLS(例えば、2つ以上、3つ以上、4つ以上、または5つ以上のNLS)を含む。いくつかの実施形態では、PTDは核酸(例えば、CasYガイド核酸、CasYガイド核酸をコードするポリヌクレオチド、CasY融合ポリペプチドをコードするポリヌクレオチド、ドナーポリヌクレオチドなど)に共有結合している。PTDの例としては、最小のウンデカペプチドタンパク質導入ドメイン(YGRKKRRQRRR: 配列番号112を含むHIV-1TATの47~57残基に対応); 細胞への直接導入に十分な複数のアルギニン(例えば、3、4、5、6、7、8、9、10、または10~50のアルギニン)を含むポリアルギニン配列; VP22ドメイン(Zender et al. (2002) Cancer Gene Ther. 9(6): 489-96); Drosophila Antennapediaタンパク質導入ドメイン(Noguchi et al. (2003) Diabetes 52(7): 1732-1737); 短縮ヒトカルシトニンペプチド(Trehin et al. (2004) Pharm. Research 21: 1248-1256); ポリリジン(Wender et al. (2000) Proc. Natl. Acad. Sci. USA 97: 13003-13008); RRQRRTSKLMKR(配列番号113); トランスポータンGWTLNSAGYLLGKINLKALAAALAKKIL(配列番号114); KALAWEAKLAKALAKALAKHLAKALAKALKCEA(配列番号115); 及びRQIKIWFQNRMRMKWKK(配列番号116)が挙げられるが、これらに限定されない。例示的なPTDとして、限定はされないが、YGRKKRRQRRR(配列番号117)、RKKRRQRRR(配列番号118); 3アルギニン残基から50アルギニン残基までからなるアルギニンホモポリマーが挙げられ、例示的なPTDドメインアミノ酸配列には、限定はされないが、以下のいずれかが含まれる: YGRKKRRQRRR(配列番号119); RKKRRQRRR(配列番号120); YARAAARQARA(配列番号121); THRLPRRRRRR(配列番号122); 及びGGRRARRRRRRR(配列番号123)。いくつかの実施形態では、PTDは、活性化可能なCPP(ACPP)である(Aguilera et al. (2009) Integr Biol (Camb) June; 1(5-6): 371-381)。ACPPには、対となるポリアニオン(例えば、Glu9すなわち「E9」)に切断可能なリンカーを介して接続されたポリカチオン性CPP(例えば、ARG9すなわち「R9」)が含まれ、これが実効電荷をほぼゼロに低下させ、それによって細胞への接着及び取り込みを阻害する。リンカーの切断時に、ポリアニオンが遊離して、ポリアルギニン及びそれに備わる接着性が局所的に露出され、それによってACPPの膜透過が「活性化」する。

【0118】

リンカー(例えば、融合パートナー用)

いくつかの実施形態では、本発明のCasYタンパク質は、リンカーポリペプチド(例えば、1つ以上のリンカーポリペプチド)を介して融合パートナーに融合することができる。リンカーポリペプチドは、種々のアミノ酸配列のいずれかを有し得る。タンパク質は、一般に柔軟な性質のスペーサーペプチドによって連結することができるが、他の化学的結合を排除するものではない。好適なリンカーには、4アミノ酸長~40アミノ酸長、または4アミノ酸長~25アミノ酸長のポリペプチドを含む。これらのリンカーは、タンパク質を結合するようなリンカーをコードする合成オリゴヌクレオチドを用いて作製することも、または融合タンパク質をコードする核酸配列によってコードすることもできる。ある程度の柔軟性をもつペプチドリンカーを使用することができる。連結ペプチドは、実質

10

20

30

40

50

的にいずれのアミノ酸配列を有していてもよいが、好ましいリンカーは、一般に柔軟なペプチドが得られるような配列を有することに留意されたい。グリシン及びアラニンなどの小さいアミノ酸の使用は、柔軟なペプチドの作製に有益なものである。そのような配列の作製は当業者にとって日常的である。多種多様なリンカーが市販されており、使用に適していると考えられる。

【0119】

リンカーポリペプチドの例として、グリシンポリマー (G)_n、グリシン-セリンポリマー (例えば、(GS)_n、GSGGS_n (配列番号124)、GGSGGS_n (配列番号125)、及びGGGS_n (配列番号126)を含む。ここで、nは少なくとも1つの整数である)、グリシン-アラニンポリマー、アラニン-セリンポリマーが挙げられる。例示的なリンカーには、GSGG (配列番号127)、GSGGG (配列番号128)、GSGSG (配列番号129)、GSGGG (配列番号130)、GGGSG (配列番号131)、GSSSG (配列番号132)などを含むが、それに限定されないアミノ酸配列を含み得る。任意の望ましい要素と複合体化されたペプチドの設計は、柔軟なリンカーに加え、柔軟性の低い構造を与える1つ以上の部分をリンカーが含み得るように、全体的にまたは部分的に柔軟であるリンカーを含んでもよいことを当業者は理解しているであろう。

10

【0120】

検出可能な標識

場合によって、本開示のCasYポリペプチドは、検出可能な標識を含む。検出可能なシグナルをもたらすことができる好適な検出可能な標識及び/または部分には、酵素、放射性同位体、特異的結合対のメンバー；フルオロフォア；蛍光タンパク質；量子ドットなどを含み得るが、これらに限定されない。

20

【0121】

好適な蛍光タンパク質として、緑色蛍光タンパク質 (GFP) またはその変異体、GFPの青色蛍光変異体 (BFP)、GFPのシアン蛍光変異体 (CFP)、GFPの黄色蛍光変異体 (YFP)、改良型GFP (EGFP)、改良型CFP (ECFP)、改良型YFP (EYFP)、GFPS65T、Emerald、Topaz (TYFP)、Venus、Citrine、mCitrine、GFPUV、不安定化EGFP (dEGFP)、不安定化ECFP (dECFP)、不安定化EYFP (dEYFP)、mCFPm、Cerulean、T-Sapphire、Cypet、YPet、mKO、HcRed、t-HcRed、DsRed、DsRed2、DsRedモノマー、J-Red、dimer2、t-dimer2 (12)、mRFP1、ポシロポリン、ウミシイタケGFP、Monster GFP、paGFP、カエデタンパク質及びキンドリングタンパク質、フィコピリタンパク質及びフィコピリタンパク質コンジュゲート、例えばB-フィコエリスリン、R-フィコエリスリン、及びアロフィコシアニンが挙げられるが、これらに限定されない。蛍光タンパク質のその他の例として、mHoneydew、mBanana、mOrange、dTomato、tdTomato、mTangerine、mStrawberry、mCherry、mGrape1、mRaspberry、mGrape2、mPlum (Shaner et al. (2005) Nat. Methods 2: 905-909)などが挙げられる。Matz et al. (1999) Nature Biotechnol. 17: 969-973に記載されるような花虫綱種由来の種々の蛍光タンパク質及び有色タンパク質はいずれも使用に適している。

30

40

【0122】

好適な酵素として、西洋ワサビペルオキシダーゼ (HRP)、アルカリホスファターゼ (AP)、ベータ-ガラクトシダーゼ (GAL)、グルコース-6-リン酸デヒドロゲナーゼ、ベータ-N-アセチルグルコサミニダーゼ、-グルクロニダーゼ、インペルターゼ、キサンチンオキシダーゼ、ホタルルシフェラーゼ、グルコースオキシダーゼ (GO)などが挙げられるが、これらに限定されない。

【0123】

50

プロトスペーサー隣接モチーフ (PAM)

CasYタンパク質は、DNAを標的とするRNAと標的DNAとの相補性領域によって規定される標的配列の位置で標的DNAに結合する。多くのCRISPRエンドヌクレアーゼについていえることだが、二本鎖標的DNAの部位特異的結合（及び/または切断）は、(i)ガイドRNAと標的DNAとの間の塩基対相補性；かつ(ii)標的DNAの短鎖モチーフ〔プロトスペーサー隣接モチーフ(PAM)と呼ぶ〕の両方によって決定される位置で発生する。

【0124】

いくつかの実施形態では、CasYタンパク質におけるPAMは、標的DNAの非相補鎖の標的配列のすぐ5'側である（相補鎖はガイドRNAのガイド配列にハイブリダイズするが、非相補鎖は、非相補鎖の逆相補鎖であるガイドRNAとは直接ハイブリダイズしない）。いくつかの実施形態では（例えば、本明細書に記載のCasY1が使用される場合）、非相補鎖のPAM配列は、5'-TA-3'である（場合によってはXTA、ここで、XはC、A、またはTである）。一例として、図5及び図7を参照のこと（PAMはTA、またはPAMがXTA（ここで、XはC、A、またはT）であると考えられる場合はCTAである）。いくつかの実施形態では（例えば、本明細書に記載のCasY1が使用される場合）、非相補鎖のPAM配列は、5'-TA-3'である（場合によってはHTA、ここで、HはC、A、またはTである）。一例として、図5及び図7を参照のこと（PAMはTA、またはPAMがHTA（ここで、HはC、A、またはT）であると考えられる場合はCTAである）。場合によっては（例えば、本明細書に記載のCasY2が使用される場合）、非相補鎖のPAM配列は、標的の5'隣接配列5'-YR-3'である（ここで、YはTまたはCであり、RはAまたはGである）。場合によっては（例えば、本明細書に記載のCasY2が使用される場合）、非相補鎖のPAM配列は、5'-TR-3'（例えば、5'-DTR-3'）である（ここで、RはAまたはGであり、DはA、G、またはTである）。一例として、図5Dを参照のこと。

【0125】

場合によって、種々の提供される方法で、異なるCasYタンパク質（すなわち、様々な種由来のCasYタンパク質）を使用し、それによって異なるCasYタンパク質の様々な酵素特性を利用すると有利な場合がある（例えば、異なるPAM配列選択性；酵素活性の増加または減少；細胞傷害性レベルの増加または減少；NHEJ、相同組換え修復、一本鎖切断、二本鎖切断など同士の均衡変化；短鎖全配列の利用などのため）。異なる種由来のCasYタンパク質は、標的DNAに異なるPAM配列を必要とする場合がある。したがって、選択した特定のCasYタンパク質について、PAM配列条件が、上記の5'-TA-3'（またはXTA、HTA）配列と異なってもよい。適切なPAM配列を同定するための(in silico及び/またはウェットラボ方法を含む)様々な方法が当技術分野において公知かつ日常的であり、任意の利便な方法を使用することができる。本明細書に記載のTA(XTA、HTA)PAM配列は、PAM欠失アッセイを使用して同定された（例えば、下記の実施例の図5を参照）。

【0126】

CasYガイドRNA

CasYタンパク質に結合してリボ核タンパク質複合体(RNP)を形成し、標的核酸（例えば、標的DNA）内の特定の場所に複合体を標的化する核酸分子を、本明細書で「CasYガイドRNA」または単に「ガイドRNA」と称する。場合によって、CasYガイドRNAにRNA塩基に加えてDNA塩基を含むようなハイブリッドDNA/RNAを作製できるが、用語「CasYガイドRNA」は、本明細書でそのような分子を包含する場合にも使用されると理解すべきである。

【0127】

CasYガイドRNAは、標的化セグメント及びタンパク質結合セグメントという2つのセグメントを含むと言ってよい。CasYガイドRNAの標的化セグメントには、標的核酸（例えば、標的ssRNA、標的ssDNA、二本鎖標的DNAの相補鎖など）内の

特定の配列（標的部位）に相補的な（それによってその配列とハイブリダイズされる）ヌクレオチド配列（ガイド配列）を含む。タンパク質結合セグメント（または「タンパク質結合配列」）は、C a s Yポリペプチドと相互作用（結合）する。本発明のC a s YガイドRNAのタンパク質結合セグメントは、互いにハイブリダイズして二本鎖RNA二重鎖（d s RNA二重鎖）を形成する、ヌクレオチドの2つの相補ストレッチを含む。標的核酸（例えば、ゲノムDNA）の部位特異的結合及び/または切断は、C a s YガイドRNA（C a s YガイドRNAのガイド配列）と標的核酸との間の塩基対相補性によって決定される位置（例えば、標的遺伝子座の標的配列）で発生し得る。

【0128】

C a s YガイドRNAとC a s Yタンパク質、例えば融合C a s Yポリペプチドは、複合体を形成する（例えば、非共有結合性相互作用による結合）。C a s YガイドRNAは、ガイド配列（標的核酸の配列に相補的であるヌクレオチド配列）を含む標的化セグメントを含むことによって、複合体に標的特異性を与える。複合体のC a s Yタンパク質は部位特異的活性（C a s Yタンパク質によって得られる切断活性及び/またはキメラC a s Yタンパク質の場合には、融合パートナーによって得られる活性）を与える。換言すれば、C a s Yタンパク質は、C a s YガイドRNAとの会合によって、標的核酸配列（例えば、標的配列）に誘導される。

【0129】

C a s YガイドRNAの「標的化配列」とも称する「ガイド配列」は、（例えば、本明細書に記載されるように）P A M配列が考慮される可能性があることを除いて、C a s YガイドRNAが、C a s Yタンパク質（例えば、天然C a s Yタンパク質、融合C a s Yポリペプチド（キメラC a s Y）など）を、任意の所望する標的核酸の任意の所望する配列に標的化できるように変更することができる。したがって、例えば、C a s YガイドRNAは、真核細胞の核酸、例えば、ウイルス核酸、真核生物の核酸など（例えば、真核生物の染色体、染色体配列、真核生物のRNAなど）の配列に相補的である（例えば、ハイブリダイズすることができる）ガイド配列を有してもよい。

【0130】

C a s YガイドRNAのガイド配列

本発明のC a s YガイドRNAは、標的核酸の配列（標的部位）に相補的なヌクレオチド配列であるガイド配列（すなわち、標的化配列）を含む。換言すれば、C a s YガイドRNAのガイド配列は、標的核酸（例えば、二本鎖DNA（d s DNA）、一本鎖DNA（s s DNA）、一本鎖RNA（s s RNA）、または二本鎖RNA（d s RNA））と、ハイブリダイゼーション（すなわち、塩基対形成）を介して配列特異的に相互作用することができる。C a s YガイドRNAのガイド配列は、標的核酸（例えば、真核生物の標的核酸、ゲノムDNAなど）内の任意の所望する標的配列にハイブリダイズするように（例えば、d s DNA標的を標的とする場合には、P A Mを考慮して）修飾（例えば、遺伝子工学によって）/設計することができる。

【0131】

いくつかの実施形態では、ガイド配列と標的核酸の標的部位との相補性パーセントは60%以上（例えば、65%以上、70%以上、75%以上、80%以上、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%）である。場合によって、ガイド配列と標的核酸の標的部位との相補性パーセントは80%以上（例えば、85%以上、90%以上、95%以上、97%以上、98%以上、99%以上、または100%）である。場合によって、ガイド配列と標的核酸の標的部位との相補性パーセントは90%以上（例えば、95%以上、97%以上、98%以上、99%以上、または100%）である。場合によって、ガイド配列と標的核酸の標的部位との相補性パーセントは100%である。

【0132】

場合によって、ガイド配列と標的核酸の標的部位との相補性パーセントは、標的核酸の標的部位の7つの連続する最も3'側のヌクレオチドにわたって100%である。

上、98%以上、99%以上、または100%)である。場合によって、ガイド配列と標的核酸の標的部位との相補性パーセントは、19~25の連続するヌクレオチドにわたって100%である。

【0137】

場合によって、ガイド配列は17~30ヌクレオチド(nt)(例えば、17~25、17~22、17~20、19~30、19~25、19~22、19~20、20~30、20~25、または20~22nt)の範囲の長さを有する。場合によって、ガイド配列は17~25ヌクレオチド(nt)(例えば、17~22、17~20、19~25、19~22、19~20、20~25、または20~22nt)の範囲の長さを有する。場合によって、ガイド配列は、17nt以上(例えば、18nt以上、19nt以上、20nt以上、21nt以上、または22nt以上;19nt、20nt、21nt、22nt、23nt、24nt、25ntなど)の長さを有する。場合によって、ガイド配列は、19nt以上(例えば、20nt以上、21nt以上、または22nt以上;19nt、20nt、21nt、22nt、23nt、24nt、25ntなど)の長さを有する。場合によって、ガイド配列は17ntの長さを有する。場合によって、ガイド配列は18ntの長さを有する。場合によって、ガイド配列は19ntの長さを有する。場合によって、ガイド配列は20ntの長さを有する。場合によって、ガイド配列は21ntの長さを有する。場合によって、ガイド配列は22ntの長さを有する。場合によって、ガイド配列は23ntの長さを有する。

10

【0138】

20

CasYガイドRNAのタンパク質結合セグメント

本発明のCasYガイドRNAのタンパク質結合セグメントはCasYタンパク質と相互作用する。CasYガイドRNAは結合したCasYタンパク質を、上述したガイド配列を介して標的核酸内の特異的ヌクレオチド配列に誘導する。CasYガイドRNAのタンパク質結合セグメントには、互いに相補的であり、ハイブリダイズして二本鎖RNA二重鎖(dsRNA二重鎖)を形成する、ヌクレオチドの2つのストレッチを含む。したがって、タンパク質結合セグメントは、dsRNA二重鎖を含む。

【0139】

場合によって、dsRNA二重鎖領域は、5~25塩基対(bp)(例えば、5~22、5~20、5~18、5~15、5~12、5~10、5~8、8~25、8~22、8~18、8~15、8~12、12~25、12~22、12~18、12~15、13~25、13~22、13~18、13~15、14~25、14~22、14~18、14~15、15~25、15~22、15~18、17~25、17~22、または17~18bp、例えば、5bp、6bp、7bp、8bp、9bp、10bpなど)の範囲を含む。場合によって、dsRNA二重鎖領域は、6~15塩基対(bp)(例えば、6~12、6~10、または6~8bp、例えば、6bp、7bp、8bp、9bp、10bpなど)の範囲を含む。場合によって、二重鎖領域は5bp以上(例えば、6bp以上、7bp以上、または8bp以上)を含む。場合によって、二重鎖領域は6bp以上(例えば、7bp以上または8bp以上)を含む。場合によって、二重鎖領域のヌクレオチドすべてが対を形成するとは限らないため、二重鎖形成領域はバルジを含むことがある。本明細書の用語「バルジ」は、二本鎖二重鎖に関与しないが、関与するヌクレオチドによって5'及び3'が取り囲まれているヌクレオチドのストレッチ(ヌクレオチドは1つの可能性がある)を意味する際に使用され、そのようなバルジは二重鎖領域の一部とみなされる。場合によって、dsRNAは1つ以上のバルジ(例えば、2つ以上、3つ以上、4つ以上のバルジ)を含む。場合によって、dsRNA二重鎖は2つ以上のバルジ(例えば、3つ以上、4つ以上のバルジ)を含む。場合によって、dsRNA二重鎖は、1~5個のバルジ(例えば、1~4、1~3、2~5、2~4、または2~3個のバルジ)を含む。

30

40

【0140】

したがって、場合によって、互いにハイブリダイズしてdsRNA二重鎖を形成するヌ

50

クレオチドのストレッチは、互いに70%~100%の相補性(例えば、75%~100%、80%~100%、85%~100%、90%~100%、95%~100%の相補性)を有する。場合によって、互いにハイブリダイズしてdsRNA二重鎖を形成するヌクレオチドのストレッチは、互いに70%~100%の相補性(例えば、75%~100%、80%~100%、85%~100%、90%~100%、95%~100%の相補性)を有する。場合によって、互いにハイブリダイズしてdsRNA二重鎖を形成するヌクレオチドのストレッチは、互いに85%~100%の相補性(例えば、90%~100%、95%~100%の相補性)を有する。場合によって、互いにハイブリダイズしてdsRNA二重鎖を形成するヌクレオチドのストレッチは、互いに70%~95%の相補性(例えば、75%~95%、80%~95%、85%~95%、90%~95%の相補性)を有する。

10

【0141】

換言すれば、いくつかの実施形態では、dsRNA二重鎖は、互いに70%~100%の相補性(例えば、75%~100%、80%~100%、85%~100%、90%~100%、95%~100%の相補性)を有するヌクレオチドの2つのストレッチを含む。場合によって、dsRNA二重鎖は、互いに85%~100%の相補性(例えば、90%~100%、95%~100%の相補性)を有するヌクレオチドの2つのストレッチを含む。場合によって、dsRNA二重鎖は、互いに70%~95%の相補性(例えば、75%~95%、80%~95%、85%~95%、90%~95%の相補性)を有するヌクレオチドの2つのストレッチを含む。

20

【0142】

本発明のCasYガイドRNAの二重鎖領域は、天然の二重鎖領域と比較して1つ以上(1つ、2つ、3つ、4つ、5つなど)の変異を含み得る。例えば、場合によって、各セグメントからの塩基対に関するヌクレオチドが異なっても、塩基対を維持することができる。場合によって、本発明のCasYガイドRNAの二重鎖領域は、(天然のCasYガイドRNAの)天然の二重鎖領域と比較して、それよりも多くの塩基対、少ない塩基対、小さいバルジ、大きいバルジ、少ないバルジ、多くのバルジ、またはそれらの任意の好都合な組み合わせを含む。

【0143】

様々なCas9ガイドRNAの例を当技術分野において見出すことができ、場合によって、Cas9ガイドRNAに導入されたものと類似する変形例も本開示のCasYガイドRNAに導入することができる(例えば、dsRNA二重鎖領域に対する変異、安定性を付加して別のタンパク質との相互作用をもたらすための5'末端または3'末端の伸長など)。例えば、Jinek et al., *Science*. 2012 Aug 17; 337(6096): 816-21; Chylinski et al., *RNA Biol*. 2013 May; 10(5): 726-37; Ma et al., *Biomed Res Int*. 2013; 2013: 270805; Hou et al., *Proc Natl Acad Sci U S A*. 2013 Sep 24; 110(39): 15644-9; Jinek et al., *Elife*. 2013; 2: e00471; Pattanayak et al., *Nat Biotechnol*. 2013 Sep; 31(9): 839-43; Qi et al., *Cell*. 2013 Feb 28; 152(5): 1173-83; Wang et al., *Cell*. 2013 May 9; 153(4): 910-8; Auer et al., *Genome Res*. 2013 Oct 31; Chen et al., *Nucleic Acids Res*. 2013 Nov 1; 41(20): e19; Cheng et al., *Cell Res*. 2013 Oct; 23(10): 1163-71; Cho et al., *Genetics*. 2013 Nov; 195(3): 1177-80; DiCarlo et al., *Nucleic Acids Res*. 2013 Apr; 41(7): 4336-43; Dickinson et al., *Nat Methods*. 2013 Oct; 10(10): 1028-34; Ebin et al., *Sci Re*

30

40

50

p. 2013; 3: 2510; Fujii et al., *Nucleic Acids Res.* 2013 Nov 1; 41(20): e187; Hu et al., *Cell Res.* 2013 Nov; 23(11): 1322-5; Jiang et al., *Nucleic Acids Res.* 2013 Nov 1; 41(20): e188; Larson et al., *Nat Protoc.* 2013 Nov; 8(11): 2180-96; Mali et al., *Nat Methods.* 2013 Oct; 10(10): 957-63; Nakayama et al., *Genesis.* 2013 Dec; 51(12): 835-43; Ran et al., *Nat Protoc.* 2013 Nov; 8(11): 2281-308; Ran et al., *Cell.* 2013 Sep 12; 154(6): 1380-9; Upadhyaya et al., *G3 (Bethesda).* 2013 Dec 9; 3(12): 2233-8; Walsh et al., *Proc Natl Acad Sci US A.* 2013 Sep 24; 110(39): 15514-5; Xie et al., *Mol Plant.* 2013 Oct 9; Yang et al., *Cell.* 2013 Sep 12; 154(6): 1370-9; Briner et al., *Mol Cell.* 2014 Oct 23; 56(2): 333-9; ならびに米国特許及び出願：第8,906,616号；第8,895,308号；第8,889,418号；第8,889,356号；第8,871,445号；第8,865,406号；第8,795,965号；第8,771,945号；第8,697,359号；第20140068797号；第20140170753号；第20140179006号；第20140179770号；第20140186843号；第20140186919号；第20140186958号；第20140189896号；第20140227787号；第20140234972号；第20140242664号；第20140242699号；第20140242700号；第20140242702号；第20140248702号；第20140256046号；第20140273037号；第20140273226号；第20140273230号；第20140273231号；第20140273232号；第20140273233号；第20140273234号；第20140273235号；第20140287938号；第20140295556号；第20140295557号；第20140298547号；第20140304853号；第20140309487号；第20140310828号；第20140310830号；第20140315985号；第20140335063号；第20140335620号；第20140342456号；第20140342457号；第20140342458号；第20140349400号；第20140349405号；第20140356867号；第20140356956号；第20140356958号；第20140356959号；第20140357523号；第20140357530号；第20140364333号；及び第20140377868号を参照のこと（すべての文献はその全体が参照により本明細書に組み込まれる）。

【0144】

CasYガイドRNAは、ガイド配列、及びハイブリダイズしてタンパク質結合セグメントのdsRNA二重鎖を形成するヌクレオチドの2つのストレッチ（「二重鎖形成セグメント」）の両方を含む。所与のCasYガイドRNAの特定の配列は、crRNAが存在する種の特徴であり得る。好適なCasYガイドRNAの例は、本明細書に記載されている。

【0145】

例示的なガイドRNA配列

図6（パネルA及びB）に示されているリピート配列（例示的なCasYガイドRNAの非ガイド配列部分）は、CasY1～Y5では天然遺伝子座由来である。場合によって、本発明のCasYガイドRNAは、（例えばガイド配列に加えて）crRNA配列CTCCGAAGTATCGGGGATAAAGGC（配列番号31）[RNAは、CUC CGAAAGUAUCGGGGAUAAAGGC（配列番号11）である]を含む（例え

ば、図6を参照)。場合によって、本発明のCasYガイドRNAは、crRNA配列CTCCGAAGTATCGGGGATAAAGGC(配列番号31)[RNAは、CUCCGAAAGUAUCGGGAUAAAGGC(配列番号11)]である]と80%以上の同一性(例えば、85%以上、90%以上、93%以上、95%以上、97%以上、98%以上、または100%の同一性)を有するヌクレオチド配列を含む。場合によって、本発明のCasYガイドRNAは、crRNA配列CTCCGAAGTATCGGGGATAAAGGC(配列番号31)[RNAは、CUCCGAAAGUAUCGGGAUAAAGGC(配列番号11)]である]と90%以上の同一性(例えば、93%以上、95%以上、97%以上、98%以上、または100%の同一性)を有するヌクレオチド配列を含む。

10

【0146】

場合によって、本発明のCasYガイドRNAは、(例えばガイド配列に加えて)crRNA配列CACCGAAATTTGGAGAGGATAAGGC(配列番号32)[RNAは、CACCGAAAUUGGAGAGGAUAAAGGC(配列番号12)]を含む(例えば、図6を参照)。場合によって、本発明のCasYガイドRNAは、crRNA配列CACCGAAATTTGGAGAGGATAAGGC(配列番号32)[RNAは、CACCGAAAUUGGAGAGGAUAAAGGC(配列番号12)]である]と80%以上の同一性(例えば、85%以上、90%以上、93%以上、95%以上、97%以上、98%以上、または100%の同一性)を有するヌクレオチド配列を含む。場合によって、本発明のCasYガイドRNAは、crRNA配列CACCGAAATTTGGAGAGGATAAGGC(配列番号32)[RNAは、CACCGAAAUUGGAGAGGAUAAAGGC(配列番号12)]である]と90%以上の同一性(例えば、93%以上、95%以上、97%以上、98%以上、または100%の同一性)を有するヌクレオチド配列を含む。

20

【0147】

場合によって、本発明のCasYガイドRNAは、(例えばガイド配列に加えて)crRNA配列CTCCGAATTAATCGGGAGGATAAGGC(配列番号33)[RNAは、CUCCGAAUAUCGGGAGGAUAAAGGC(配列番号13)]である]を含む(例えば、図6を参照)。場合によって、本発明のCasYガイドRNAは、crRNA配列CTCCGAATTAATCGGGAGGATAAGGC(配列番号33)[RNAは、CUCCGAAUAUCGGGAGGAUAAAGGC(配列番号13)]である]と80%以上の同一性(例えば、85%以上、90%以上、93%以上、95%以上、97%以上、98%以上、または100%の同一性)を有するヌクレオチド配列を含む。場合によって、本発明のCasYガイドRNAは、crRNA配列CTCCGAATTAATCGGGAGGATAAGGC(配列番号33)[RNAは、CUCCGAAUAUCGGGAGGAUAAAGGC(配列番号13)]である]と90%以上の同一性(例えば、93%以上、95%以上、97%以上、98%以上、または100%の同一性)を有するヌクレオチド配列を含む。

30

【0148】

場合によって、本発明のCasYガイドRNAは、(例えばガイド配列に加えて)crRNA配列CCCCGAATATAGGGGACAAAAAGGC(配列番号34)[RNAは、CCCCGAUAUAGGGGACAAAAAGGC(配列番号14)]である]を含む(例えば、図6を参照)。場合によって、本発明のCasYガイドRNAは、crRNA配列CCCCGAATATAGGGGACAAAAAGGC(配列番号34)[RNAは、CCCCGAUAUAGGGGACAAAAAGGC(配列番号14)]である]と80%以上の同一性(例えば、85%以上、90%以上、93%以上、95%以上、97%以上、98%以上、または100%の同一性)を有するヌクレオチド配列を含む。場合によって、本発明のCasYガイドRNAは、crRNA配列CCCCGAATATAGGGGACAAAAAGGC(配列番号34)[RNAは、CCCCGAUAUAGGGGACAAAAAGGC(配列番号14)]である]と90%以上の同一性(例え

40

50

ば、93%以上、95%以上、97%以上、98%以上、または100%の同一性)を有するヌクレオチド配列を含む。

【0149】

場合によって、本発明のCasYガイドRNAは、(例えばガイド配列に加えて) crRNA配列GTCCTAGACATACAGGTGGAAAGGTGAGAGTAAAGAC(配列番号35)[RNAは、GUCUAGACAUAACAGGUGGAAAGGUGAGAGUAAGAC(配列番号15)]を含む(例えば、図6を参照)。場合によって、本発明のCasYガイドRNAは、crRNA配列GTCCTAGACATACAGGTGGAAAGGTGAGAGTAAAGAC(配列番号35)[RNAは、GUCUAGACAUAACAGGUGGAAAGGUGAGAGUAAGAC(配列番号15)]と80%以上の同一性(例えば、85%以上、90%以上、93%以上、95%以上、97%以上、98%以上、または100%の同一性)を有するヌクレオチド配列を含む。場合によって、本発明のCasYガイドRNAは、crRNA配列GTCCTAGACATACAGGTGGAAAGGTGAGAGTAAAGAC(配列番号35)[RNAは、GUCUAGACAUAACAGGUGGAAAGGUGAGAGUAAGAC(配列番号15)]と90%以上の同一性(例えば、93%以上、95%以上、97%以上、98%以上、または100%の同一性)を有するヌクレオチド配列を含む。

10

【0150】

場合によって、本発明のCasYガイドRNAは、(例えばガイド配列に加えて)配列番号11~15のいずれか1つに記載されるcrRNA配列を含む。場合によって、本発明のCasYガイドRNAは、配列番号11~15のいずれか1つに記載されるcrRNA配列と80%以上の同一性(例えば、85%以上、90%以上、93%以上、95%以上、97%以上、98%以上、または100%の同一性)を有するヌクレオチド配列を含む。場合によって、本発明のCasYガイドRNAは、配列番号11~15のいずれか1つに記載されるcrRNA配列と90%以上の同一性(例えば、93%以上、95%以上、97%以上、98%以上、または100%の同一性)を有するヌクレオチド配列を含む。

20

【0151】

場合によって、本発明のCasYガイドRNAは、(例えばガイド配列に加えて)配列番号11~14のいずれか1つに記載されるcrRNA配列を含む。場合によって、本発明のCasYガイドRNAは、配列番号11~14のいずれか1つに記載されるcrRNA配列と80%以上の同一性(例えば、85%以上、90%以上、93%以上、95%以上、97%以上、98%以上、または100%の同一性)を有するヌクレオチド配列を含む。場合によって、本発明のCasYガイドRNAは、配列番号11~14のいずれか1つに記載されるcrRNA配列と90%以上の同一性(例えば、93%以上、95%以上、97%以上、98%以上、または100%の同一性)を有するヌクレオチド配列を含む。

30

【0152】

CasY18の天然遺伝子由来のリピート配列(例示的なCasYガイドRNAの非ガイド配列部分)は、CTCCGTGAATACGTGGGGTAAAGGC(配列番号36)である[RNAは、CUCUGUAUAACGUGGGGUAAGGC(配列番号16)]である。場合によって、本発明のCasYガイドRNAは、(例えばガイド配列に加えて)crRNA配列CTCCGTGAATACGTGGGGTAAAGGC(配列番号36)[RNAは、CUCUGUAUAACGUGGGGUAAGGC(配列番号16)]を含む。場合によって、本発明のCasYガイドRNAは、crRNA配列CTCCGTGAATACGTGGGGTAAAGGC(配列番号36)[RNAは、CUCUGUAUAACGUGGGGUAAGGC(配列番号16)]と80%以上の同一性(例えば、85%以上、90%以上、93%以上、95%以上、97%以上、98%以上、または100%の同一性)を有するヌクレオチド配列を含む。場合によって、本発明のCasYガイドRNAは、crRNA配列CTCCGTGAATACGTGGGGTAAAGGC(配列番号36)[RNAは、CUCUGUAUAACGUGGGGUAAGGC(配列番号16)]と80%以上の同一性(例えば、85%以上、90%以上、93%以上、95%以上、97%以上、98%以上、または100%の同一性)を有するヌクレオチド配列を含む。

40

50

C G T G G G G T A A A G G C (配列番号 36) [RNAは、C U C C G U G A A U A C G U G G G U A A A G G C (配列番号 16)である]と90%以上の同一性(例えば、93%以上、95%以上、97%以上、98%以上、または100%の同一性)を有するヌクレオチド配列を含む。

【0153】

場合によって、本発明のCasYガイドRNAは、(例えばガイド配列に加えて)配列番号11~16のいずれか1つに記載されるcrRNA配列を含む。場合によって、本発明のCasYガイドRNAは、(例えばガイド配列に加えて)配列番号11~16のいずれか1つに記載されるcrRNA配列と80%以上の同一性(例えば、85%以上、90%以上、93%以上、95%以上、97%以上、98%以上、または100%の同一性)を有するヌクレオチド配列を含む。場合によって、本発明のCasYガイドRNAは、配列番号11~16のいずれか1つに記載されるcrRNA配列と90%以上の同一性(例えば、93%以上、95%以上、97%以上、98%以上、または100%の同一性)を有するヌクレオチド配列を含む。

10

【0154】

CasYシステム

本開示は、CasYシステムを提供する。本開示のCasYシステムは以下を含み得る：
 a) 本開示のCasYポリペプチド及びCasYガイドRNA；
 b) 本開示のCasYポリペプチド、CasYガイドRNA、及びドナー鑄型核酸；
 c) 本開示のCasY融合ポリペプチド及びCasYガイドRNA；
 d) 本開示のCasY融合ポリペプチド、CasYガイドRNA、及びドナー鑄型核酸；
 e) 本開示のCasYポリペプチドをコードするmRNA及びCasYガイドRNA；
 f) 本開示のCasYポリペプチドをコードするmRNA、CasYガイドRNA、及びドナー鑄型核酸；
 g) 本開示のCasY融合ポリペプチドをコードするmRNA及びCasYガイドRNA；
 h) 本開示のCasY融合ポリペプチドをコードするmRNA、CasYガイドRNA、及びドナー鑄型核酸；
 i) 本開示のCasYポリペプチドをコードするヌクレオチド配列及びCasYガイドRNAをコードするヌクレオチド配列を含む組換え発現ベクター；
 j) 本開示のCasYポリペプチドをコードするヌクレオチド配列、CasYガイドRNAをコードするヌクレオチド配列、及びドナー鑄型核酸をコードするヌクレオチド配列を含む組換え発現ベクター；
 k) 本開示のCasY融合ポリペプチドをコードするヌクレオチド配列及びCasYガイドRNAをコードするヌクレオチド配列を含む組換え発現ベクター；
 l) 本開示のCasY融合ポリペプチドをコードするヌクレオチド配列、CasYガイドRNAをコードするヌクレオチド配列、及びドナー鑄型核酸をコードするヌクレオチド配列を含む組換え発現ベクター；
 m) 本開示のCasYポリペプチドをコードするヌクレオチド配列を含む第1の組換え発現ベクター、及びCasYガイドRNAをコードするヌクレオチド配列を含む第2の組換え発現ベクター；
 n) 本開示のCasYポリペプチドをコードするヌクレオチド配列を含む第1の組換え発現ベクター、ならびにCasYガイドRNAをコードするヌクレオチド配列及びドナー鑄型核酸を含む第2の組換え発現ベクター；
 o) 本開示のCasY融合ポリペプチドをコードするヌクレオチド配列を含む第1の組換え発現ベクター、及びCasYガイドRNAをコードするヌクレオチド配列を含む第2の組換え発現ベクター；
 p) 本開示のCasY融合ポリペプチドをコードするヌクレオチド配列を含む第1の組換え発現ベクター、ならびにCasYガイドRNAをコードするヌクレオチド配列及びドナー鑄型核酸を含む第2の組換え発現ベクター；
 q) 本開示のCasYポリペプチドをコードするヌクレオチド配列、第1のCasYガイドRNAをコードするヌクレオチド配列、及び第2のCasYガイドRNAをコードするヌクレオチド配列を含む組換え発現ベクター；
 もしくは
 r) 本開示のCasY融合ポリペプチドをコードするヌクレオチド配列、第1のCasYガイドRNAをコードするヌクレオチド配列、及び第2のCasYガイドRNAをコードするヌクレオチド配列を含む組換え発現ベクター；
 または(a)~(r)のいずれか1つの何らかの変形例。

20

30

40

【0155】

50

核酸

本開示は、ドナーポリヌクレオチド配列、C a s Yポリペプチド（例えば、野生型C a s Yタンパク質、ニックーゼC a s Yタンパク質、d C a s Yタンパク質、キメラC a s Yタンパク質など）をコードするヌクレオチド配列、C a s YガイドRNA、及びC a s YガイドRNAをコードするヌクレオチド配列のうち1つ以上を含む1つ以上の核酸を提供する。本開示は、C a s Y融合ポリペプチドをコードするヌクレオチド配列を含む核酸を提供する。本開示は、C a s Yポリペプチドをコードするヌクレオチド配列を含む組換え発現ベクターを提供する。本開示は、C a s Y融合ポリペプチドをコードするヌクレオチド配列を含む組換え発現ベクターを提供する。本開示は、a) C a s Yポリペプチドをコードするヌクレオチド配列；及びb) C a s YガイドRNA（複数可）をコードするヌクレオチド配列を含む組換え発現ベクターを提供する。本開示は、a) C a s Y融合ポリペプチドをコードするヌクレオチド配列；及びb) C a s YガイドRNA（複数可）をコードするヌクレオチド配列を含む組換え発現ベクターを提供する。場合によって、C a s Yタンパク質をコードするヌクレオチド配列及び/またはC a s YガイドRNAをコードするヌクレオチド配列は、選択した細胞型（例えば、原核細胞、真核細胞、植物細胞、動物細胞、哺乳動物細胞、霊長類細胞、齧歯類細胞、ヒト細胞など）において機能できるプロモーターに機能的に連結される。

10

【0156】

場合によって、本開示のC a s Yポリペプチドをコードするヌクレオチド配列は、コドン最適化されている。この種類の最適化は、同じタンパク質をコードしながら、意図する宿主生物または細胞のコドン選択性を模倣する、C a s Yをコードするヌクレオチド配列の変異を引き起こすことができる。したがって、コドンを変更することはできるが、コードされたタンパク質は変更されないままである。例えば、意図する標的細胞がヒト細胞であった場合、ヒトコドンに最適化された、C a s Yをコードするヌクレオチド配列を使用することができる。別の非限定的な例として、意図する宿主細胞がマウス細胞であった場合、マウスコドンに最適化された、C a s Yをコードするヌクレオチド配列を生成することができる。別の非限定的な例として、意図する宿主細胞が植物細胞であった場合、植物コドンに最適化された、C a s Yをコードするヌクレオチド配列を生成することができる。別の非限定的な例として、意図する宿主細胞が昆虫細胞であった場合、昆虫コドンに最適化された、C a s Yをコードするヌクレオチド配列を生成することができる。

20

30

【0157】

本開示は、(i)ドナー鋳型核酸のヌクレオチド配列（ドナー鋳型は、標的核酸（例えば標的ゲノム）の標的配列と相同性を有するヌクレオチド配列を含む）；(ii)標的ゲノムの標的遺伝子座の標的配列にハイブリダイズするC a s YガイドRNAをコードするヌクレオチド配列（例えば、真核細胞などの標的細胞において機能できるプロモーターに機能的に連結される）；及び(iii)C a s Yタンパク質をコードするヌクレオチド配列（例えば、真核細胞などの標的細胞において機能できるプロモーターに機能的に連結される）を含む1つ以上の組換え発現ベクターを提供する（異なる組換え発現ベクターの場合もあれば、同一の組換え発現ベクターの場合もある）。本開示は、(i)ドナー鋳型核酸のヌクレオチド配列（ドナー鋳型は、標的核酸（例えば標的ゲノム）の標的配列と相同性を有するヌクレオチド配列を含む）；及び(ii)標的ゲノムの標的遺伝子座の標的配列にハイブリダイズするC a s YガイドRNAをコードするヌクレオチド配列（例えば、真核細胞などの標的細胞において機能できるプロモーターに機能的に連結される）を含む1つ以上の組換え発現ベクターを提供する（異なる組換え発現ベクターの場合もあれば、同一の組換え発現ベクターの場合もある）。本開示は、(i)標的ゲノムの標的遺伝子座の標的配列にハイブリダイズするC a s YガイドRNAをコードするヌクレオチド配列（例えば、真核細胞などの標的細胞において機能できるプロモーターに機能的に連結される）；及び(ii)C a s Yタンパク質をコードするヌクレオチド配列（例えば、真核細胞などの標的細胞において機能できるプロモーターに機能的に連結される）を含む1つ以上の組換え発現ベクターを提供する（異なる組換え発現ベクターの場合もあれば、同一の組

40

50

換え発現ベクターの場合もある)。

【0158】

好適な発現ベクターとして、ウイルス発現ベクター(例えば、ワクシニアウイルス; ポリオウイルス; アデノウイルス(例えば、Li et al., Invest Ophthalmol Vis Sci 35:2543-2549, 1994; Borrás et al., Gene Ther 6:515-524, 1999; Li and Davidson, PNAS 92:7700-7704, 1995; Sakamoto et al., Hum Gene Ther 5:1088-1097, 1999; WO94/12649、WO93/03769; WO93/19191; WO94/28938; WO95/11984、及びWO95/00655を参照); アデノ随伴ウイルス(AAV) (例えば、Ali et al., Hum Gene Ther 9:81-86, 1998; Flannery et al., PNAS 94:6916-6921, 1997; Bennett et al., Invest Ophthalmol Vis Sci 38:2857-2863, 1997; Jomary et al., Gene Ther 4:683-690, 1997; Rolling et al., Hum Gene Ther 10:641-648, 1999; Ali et al., Hum Mol Genet 5:591-594, 1996; WO93/09239のSrivastava, Samulski et al., J. Vir. (1989) 63:3822-3828; Mendelson et al., Virol. (1988) 166:154-165; 及びFlotte et al., PNAS (1993) 90:10613-10617を参照); SV40; 単純ヘルペスウイルス; ヒト免疫不全ウイルス(例えば、Miyoshi et al., PNAS 94:10319-23, 1997; Takahashi et al., J. Virol 73:7812-7816, 1999を参照)系のウイルスベクター; レトロウイルスベクター(例えば、マウス白血病ウイルス、脾臓壊死ウイルス、ならびにラウス肉腫ウイルス、ハーベイ肉腫ウイルス、トリ白血病ウイルス、レンチウイルス、ヒト免疫不全ウイルス、骨髄増殖性肉腫ウイルス、及び乳癌ウイルスなどのレトロウイルス由来のベクター)などが挙げられる。場合によって、本開示の組換え発現ベクターは、組換えアデノ随伴ウイルス(AAV)ベクターである。場合によって、本開示の組換え発現ベクターは、組換えレンチウイルスベクターである。場合によって、本開示の組換え発現ベクターは、組換えレトロウイルスベクターである。

【0159】

利用する宿主/ベクター系に応じて、構成的プロモーター及び誘導性プロモーター、転写エンハンサー要素、転写ターミネーターなどを含む、いくつかの好適な転写及び翻訳制御要素のいずれかを発現ベクターに使用することができる。

【0160】

いくつかの実施形態では、CasYガイドRNAをコードするヌクレオチド配列は、制御要素、例えば、プロモーターなどの転写制御要素に機能的に連結される。いくつかの実施形態では、CasYタンパク質またはCasY融合ポリペプチドをコードするヌクレオチド配列は、制御要素、例えば、プロモーターなどの転写制御要素に機能的に連結される。

【0161】

転写制御要素はプロモーターであり得る。場合によって、プロモーターは構成的に活性化プロモーターである。場合によって、プロモーターは制御可能なプロモーターである。場合によって、プロモーターは誘導性プロモーターである。場合によって、プロモーターは組織特異的プロモーターである。場合によって、プロモーターは細胞型特異的プロモーターである。場合によって、転写制御要素(例えば、プロモーター)は、標的細胞型または標的細胞集団において機能的である。例えば、場合によって、転写制御要素は、真核細胞、例えば造血幹細胞(例えば、動員された末梢血(mPB)CD34(+)細胞、骨髄(BM)CD34(+)細胞など)において機能的であり得る。

【0162】

真核生物プロモーター（真核細胞において機能するプロモーター）の非限定的な例として、EF1、サイトメガロウイルス（CMV）最初期由来のもの、単純ヘルペスウイルス（HSV）チミジンキナーゼ、初期及び後期SV40、レトロウイルス由来の長末端反復配列（LTR）、及びマウスメタロチオネインIが挙げられる。適切なベクター及びプロモーターの選択は、十分に当業者のレベルの範囲内である。発現ベクターはまた、翻訳開始及び転写ターミネーターに対するリボソーム結合部位を含んでいてもよい。発現ベクターはまた、発現を増幅するための適切な配列を含んでいてもよい。発現ベクターはまた、CasYタンパク質に融合することができ、それによってキメラCasYポリペプチドを生じる、タンパク質タグ（例えば、6xHisタグ、ヘマグルチニンタグ、蛍光タンパク質など）をコードするヌクレオチド配列を含んでいてもよい。

10

【0163】

いくつかの実施形態では、CasYガイドRNA及び/またはCasY融合ポリペプチドをコードするヌクレオチド配列は、誘導性プロモーターに機能的に連結される。いくつかの実施形態では、CasYガイドRNA及び/またはCasY融合タンパク質をコードするヌクレオチド配列は、構成的プロモーターに機能的に連結される。

【0164】

プロモーターは構成的に活性なプロモーター（すなわち、構成的に活性/「オン」状態であるプロモーター）であっても、誘導性プロモーター（すなわち、活性/「オン」または不活性/「オフ」の状態が、外部刺激、例えば特定の温度、化合物、またはタンパク質の存在によって制御されるプロモーター）であっても、空間的に制限されたプロモーター（すなわち、転写制御要素、エンハンサーなど）（例えば、組織特異的プロモーター、細胞型特異的プロモーターなど）であっても、一時的に制限されるプロモーター（すなわち、胚発生の特定の段階または生物学的過程の特定の段階（例えば、マウスにおける毛包サイクル）の間、プロモーターが「オン」状態または「オフ」状態にある）であってもよい。

20

【0165】

好適なプロモーターは、ウイルスに由来するものであるため、ウイルスプロモーターと呼ばれる場合がある。あるいは好適なプロモーターは、原核生物または真核生物を含む任意の生物に由来するものであってもよい。好適なプロモーターを使用して、任意のRNAポリメラーゼ（例えば、pol I、pol II、pol III）によって発現を駆動することができる。例示的なプロモーターとして、SV40初期プロモーター、マウス乳癌ウイルスの長末端反復配列（LTR）プロモーター；アデノウイルス主要後期プロモーター（Ad MLP）；単純ヘルペスウイルス（HSV）プロモーター、CMV最初期プロモーター領域（CMV IE）などのサイトメガロウイルス（CMV）プロモーター、ラウス肉腫ウイルス（RSV）プロモーター、ヒトU6核内低分子プロモーター（U6）（Miyagishi et al., Nature Biotechnology 20, 497-500 (2002)）、改良型U6プロモーター（例えば、Xia et al., Nucleic Acids Res. 2003 Sep 1; 31(17)）、ヒトH1プロモーター（H1）などが挙げられるが、これらに限定されない。

30

40

【0166】

場合によって、CasYガイドRNAをコードするヌクレオチド配列は、真核生物細胞において機能できるプロモーター（例えば、U6プロモーター、改良型U6プロモーター、H1プロモーターなど）に機能的に連結される（その制御下に置かれる）。当業者であれば理解しているであろうが、U6プロモーター（例えば真核細胞において）または別のPol IIIプロモーターを使用して核酸（例えば発現ベクター）からRNA（例えばガイドRNA）を発現する場合、いくつかのT（RNAではUをコードする）が連続して存在する場合に、RNAの変異を要する場合がある。これは、DNAのTの列（例えば5つのT）が、ポリメラーゼIII（Pol III）に対するターミネーターとして作用する可能性があるためである。したがって、真核細胞のガイドRNAの転写を確実にするには

50

、場合によりガイドRNAをコードする配列を変更して、Tの連続を排除することが必要となる場合がある。場合によって、CasYタンパク質（例えば、野生型CasYタンパク質、ニックアーゼCasYタンパク質、dCasYタンパク質、キメラCasYタンパク質など）をコードするヌクレオチド配列は、真核生物細胞において機能できるプロモーター（例えば、CMVプロモーター、EF1プロモーター、エストロゲン受容体制御性プロモーターなど）に機能的に連結される。

【0167】

誘導性プロモーターの例としては、T7 RNAポリメラーゼプロモーター、T3 RNAポリメラーゼプロモーター、イソプロピル-ベータ-D-チオガラクトピラノシド（IPTG）制御性プロモーター、ラクトース誘導性プロモーター、熱ショックプロモーター、テトラサイクリン制御性プロモーター、ステロイド制御性プロモーター、金属制御性プロモーター、エストロゲン受容体制御性プロモーターなどが挙げられるが、これらに限定されない。したがって誘導性プロモーターは、ドキシサイクリン；エストロゲン及び/またはエストロゲン類似体；IPTGなどを含むがこれに限定されない分子によって制御することができる。

10

【0168】

使用に適した誘導性プロモーターには、本明細書に記載される、または当業者に公知である任意の誘導性プロモーターを含む。誘導性プロモーターの例として、化学的/生化学的制御及び物理的制御プロモーター、例えば、アルコール制御性プロモーター、テトラサイクリン制御性プロモーター（例えば、アンヒドロテトラサイクリン（aTc）応答性プロモーター及び他のテトラサイクリン応答性プロモーター系。これにはテトラサイクリン抑制因子タンパク質（tetR）、テトラサイクリンオペレーター配列（tetO）、及びテトラサイクリントランス活性化因子融合タンパク質（tTA）を含む）、ステロイド制御性プロモーター（例えば、ラットのグルココルチコイド受容体、ヒトエストロゲン受容体、蛾のエクジソン受容体に基づいたプロモーター、及びステロイド/レチノイド/甲状腺受容体スーパーファミリーからのプロモーター）、金属制御性プロモーター（例えば、酵母、マウス、及びヒトからのメタロチオネイン（金属イオンに結合して封鎖するタンパク質）遺伝子に由来するプロモーター）、病原性制御プロモーター（例えば、サリチル酸、エチレン、またはベンゾチアジアゾール（BTH）によって誘導される）、温度/熱誘導性プロモーター（例えば、熱ショックプロモーター）、及び光制御性プロモーター（例えば、植物細胞由来の光応答性プロモーター）が挙げられるが、これらに限定されない。

20

30

【0169】

場合によって、プロモーターは、多細胞生物において、特異的細胞のサブセットでプロモーターが活性（すなわち、「オン」）であるような空間的に制限されたプロモーター（すなわち、細胞型特異的プロモーター、組織特異的プロモーターなど）である。空間的に制限されたプロモーターはまた、エンハンサー、転写制御要素、制御配列などと呼ばれる場合もある。プロモーターが標的宿主細胞（例えば、真核細胞；原核細胞）において機能的である限り、利便な空間的に制限されたプロモーターであればいずれでも使用することができる。

40

【0170】

場合によって、プロモーターは可逆的プロモーターである。可逆的な誘導性プロモーターを含む、好適な可逆的プロモーターは当技術分野で公知である。そのような可逆的プロモーターは、多くの生物、例えば真核生物及び原核生物から単離して得ることができる。第1の生物に由来する可逆的プロモーターを第2の生物で使用する、例えば、第1の原核生物と第2の真核生物、第1の真核生物と第2の原核生物などで使用するための改変は、当技術分野で周知されている。そのような可逆的プロモーター、及びそのような可逆的プロモーターに基づくが、追加の制御タンパク質も含むシステムとして、アルコール制御性プロモーター（例えば、アルコールデヒドロゲナーゼI（alcA）遺伝子プロモーター、アルコールトランス活性化因子タンパク質（AlcR）に応答性のプロモーターなど）

50

、テトラサイクリン制御性プロモーター（例えば、Tet Activator、Tet ON、Tet OFFなどを含むプロモーター系）、ステロイド制御性プロモーター（例えば、ラットのグルココルチコイド受容体プロモーター系、ヒトのエストロゲン受容体プロモーター系、レチノイドプロモーター系、甲状腺プロモーター系、エクジソンプロモーター系、ミフェプリストンプロモーター系など）、金属制御性プロモーター（例えば、メタロチオネインプロモーター系など）、病原関連制御性プロモーター（例えば、サリチル酸制御性プロモーター、エチレン制御性プロモーター、ベンゾチアジアゾール制御性プロモーターなど）、温度制御性プロモーター（例えば、熱ショック誘導性プロモーター（例えば、HSP-70、HSP-90、ダイズ熱ショックプロモーターなど））、光制御性プロモーター、合成誘導性プロモーターなどが挙げられるが、これらに限定されない。

10

【0171】

核酸（例えば、ドナーポリヌクレオチド配列を含む核酸、CasYタンパク質及び/またはCasYガイドRNAをコードする1つ以上の核酸など）を宿主細胞に導入する方法は当技術分野で公知であり、任意の利便な方法を使用して、核酸（例えば発現構築物）を細胞へと導入することができる。好適な方法として、例えば、ウイルス感染、トランスフェクション、リポフェクション、エレクトロポレーション、リン酸カルシウム沈降、ポリエチレンイミン（PEI）介在型トランスフェクション、DEAE-デキストラン介在型トランスフェクション、リポソーム介在型トランスフェクション、パーティクルガン技術、リン酸カルシウム沈降、直接マイクロインジェクション、ナノ粒子介在型核酸送達などが挙げられる。

20

【0172】

細胞への組換え発現ベクターの導入は、細胞の生存を促進する任意の培地及び任意の培養条件下で行なうことができる。標的細胞への組換え発現ベクターの導入は、*in vivo*または*ex vivo*で実施することができる。標的細胞への組換え発現ベクターの導入は、*in vitro*で実施することができる。

【0173】

いくつかの実施形態では、CasYタンパク質をRNAとして提供することができる。RNAは、直接的な化学合成によって提供することも、または（例えば、CasYタンパク質をコードする）DNAから*in vitro*で転写することもできる。合成した後は、核酸を細胞に導入するための周知の技術（例えば、マイクロインジェクション、エレクトロポレーション、トランスフェクションなど）のいずれかによってRNAを細胞に導入することができる。

30

【0174】

核酸は、十分に開発されたトランスフェクション技術（例えばAngel and Yanik (2010) PLoS ONE 5 (7) : e11756を参照）、ならびに市販されているQiagen製のTransMessenger（登録商標）試薬、Stemgent製のStemfect（商標）RNA Transfection Kit、及びMirus Bio LLC製のTransIT（登録商標）-mRNA Transfection Kitを使用して細胞に供給することができる。Beumer et al. (2008) PNAS 105 (50) : 19821 - 19826も参照のこと。

40

【0175】

標的宿主細胞にベクターを直接供給してもよい。換言すれば、ベクターが細胞によって取り込まれるように、本発明の核酸を含むベクター（例えば、ドナー鑄型配列を有し、CasYガイドRNAをコードする組換え発現ベクター；CasYタンパク質をコードする組換え発現ベクターなど）と細胞を接触させる。プラスミドである核酸ベクターと細胞を接触させるための方法として、エレクトロポレーション、塩化カルシウムトランスフェクション、マイクロインジェクション、及びリポフェクションが挙げられ、当技術分野で周知である。ウイルスベクター送達では、本発明のウイルス発現ベクターを含むウイルス粒子と細胞を接触させることができる。

【0176】

50

レトロウイルス、例えば、レンチウイルスが本開示の方法での使用に適している。一般的に使用されるレトロウイルスベクターは「機能欠損型」、すなわち増殖性感染に必要なウイルスタンパク質を産生することができない。代わりに、ベクターの複製にパッケージング細胞株の増殖を必要とする。対象となる核酸を含むウイルス粒子を生成するには、その核酸を含むレトロウイルス核酸を、パッケージング細胞株によってウイルスカプシドにパッケージングする。パッケージング細胞株が異なると、カプシドに組み込まれるエンベロープタンパク質（エコトロピック、アンホトロピック、またはゼノトロピック）が異なる。このエンベロープタンパク質は、細胞に対するウイルス粒子の特異性を決定する（マウス及びラットではエコトロピック；ヒト、イヌ、及びマウスを含むほとんどの哺乳動物細胞型ではアンホトロピック；ならびにマウス細胞を除くほとんどの哺乳動物細胞型ではゼノトロピック）。適切なパッケージング細胞株を使用すると、パッケージングされるウイルス粒子によって細胞を確実に標的化することができる。本発明のベクター発現ベクターをパッケージング細胞株に導入する方法、及びパッケージング株によって生成されるウイルス粒子を回収する方法は、当技術分野で周知されている。核酸はまた、直接マイクロインジェクション（例えば、RNAの注入）によって導入することができる。

10

【0177】

CasYガイドRNA及び/またはCasYポリペプチドをコードする核酸を標的宿主細胞へ供給するために使用されるベクターは、対象となる核酸の発現の駆動、すなわち転写活性化に適したプロモーターを含み得る。換言すれば、場合によって、対象となる核酸はプロモーターに機能的に連結される。これは遍在的に作用するプロモーター、例えばCMV - -アクチンプロモーター、または特定の細胞集団で活性であるプロモーター、もしくはテトラサイクリンなどの薬物の存在に応答するプロモーターのような誘導性プロモーターを含み得る。転写活性化とは、標的細胞における基礎レベルよりも、転写を10倍、100倍、より一般的には1000倍増加させることを意図する。加えて、CasYガイドRNA及び/またはCasYタンパク質をコードする核酸を細胞に供給するために使用されるベクターには、CasYガイドRNA及び/またはCasYタンパク質が取り込まれた細胞を同定するために、標的細胞の選択マーカーをコードする核酸配列を含んでもよい。

20

【0178】

CasYポリペプチドまたはCasY融合ポリペプチドをコードするヌクレオチド配列を含む核酸は、場合によってRNAである。したがって、CasY融合タンパク質はRNAとして細胞に導入することができる。細胞にRNAを導入する方法は、当技術分野で公知であり、例えば、直接注入、トランスフェクション、またはDNAの導入に使用される任意の他の方法を含み得る。CasYタンパク質は、代わりにポリペプチドとして細胞に供給されてもよい。そのようなポリペプチドは、必要に応じて生成物の可溶性を増加させるポリペプチドドメインに融合されてもよい。ドメインは、規定されたプロテアーゼ切断部位、例えば、TEVプロテアーゼによって切断されるTEV配列を介してポリペプチドに連結することができる。リンカーはまた、1つ以上の柔軟な配列、例えば1~10個のグリシン残基を含んでもよい。いくつかの実施形態では、融合タンパク質の切断は、生成物の可溶性を維持する緩衝液中、例えば0.5~2Mの尿素の存在下、可溶性を高めるポリペプチド及び/またはポリヌクレオチドの存在下などで行われる。対象となるドメインとして、エンドソーム分解ドメイン、例えばインフルエンザHAドメイン；及び産生を助けるポリペプチド、例えばIF2ドメイン、GSTドメイン、GRPEドメインなどが挙げられる。ポリペプチドは、安定性を改善するように製剤化することができる。例えば、ペプチドをPEG化することで、ポリエチレンオキシ基によって血流中での持続時間の延長をもたらすことができる。

30

40

【0179】

加えてまたはその代わりに、本開示のCasYポリペプチドを、細胞による取り込みを促進するポリペプチド透過性ドメインに融合することができる。いくつかの透過性ドメインが当技術分野で公知であり、これをペプチド、ペプチド模倣物、及び非ペプチド担体を

50

含む本開示の非組み込み型ポリペプチドに使用することができる。例えば、透過性ペプチドは、ペネトラチンと呼ばれる、*Drosophila melanogaster*の転写因子*Antennapedia*の3番目のアルファヘリックスに由来するものであってよく、アミノ酸配列RQIKIWFQNRRMKWKK(配列番号133)を含む。別の例として、透過性ペプチドは、HIV-1 tatの塩基性領域のアミノ酸配列を含み、これには、例えば天然のtatタンパク質の49~57アミノ酸を含み得る。他の透過性ドメインとして、ポリアルギニンモチーフ、例えば、HIV-1 revタンパク質の34~56アミノ酸の領域、ノナアルギニン、オクタアルギニンなどが挙げられる。(例えば、Futaki et al. (2003) *Curr Protein Pept Sci.* 2003 Apr; 4(2): 87-9及び446; ならびにWender et al. (2000) *Proc. Natl. Acad. Sci. U.S.A.* 2000 Nov. 21; 97(24): 13003-8; 公開された米国特許出願第20030220334号; 第20030083256号; 第20030032593号; 及び第20030022831号を参照のこと。転位ペプチド及びペプトイドの教示のために参照により本明細書に明確に組み込まれる)。ノナアルギニン(R9)配列は、特性決定されているより効率的なPTDのいずれかである(Wender et al. 2000; Uemura et al. 2002)。融合がなされる部位を選択して、ポリペプチドの生物学的活性、分泌、または結合特性を最適化することができる。最適な部位は日常的な実験によって決定されるであろう。

10

20

30

40

50

【0180】

本開示のCasYポリペプチドは、*in vitro*で、または真核細胞もしくは原核細胞によって産生することができる、さらにアンフォールディング、例えば熱変性、ジチオスレート還元などにより処理すること、さらに当技術分野で公知の方法を使用してリフォールディングすることもできる。

【0181】

一次配列を変更しないことを目的とする修飾には、ポリペプチドの化学誘導体化、例えば、アシル化、アセチル化、カルボキシル化、アミド化などを含む。これにはグリコシル化修飾、例えば、その合成及び処理工程またはさらなる処理工程、例えば、哺乳動物のグリコシル化または脱グリコシル化酵素など、グリコシル化に影響を及ぼす酵素にポリペプチドを曝露することによってポリペプチドのグリコシル化パターンを変更することによりなされるものを含む。また、アミノ酸残基をリン酸化した配列、例えばホスホチロシン、ホスホセリン、またはホスホスレオニンも包含される。

【0182】

本開示の実施形態に包含するのに適するものは、通常分子生物学的技術及び合成化学を使用して修飾し、それによってタンパク質分解耐性を改善した、標的配列特異性を変更した、可溶性を最適化した、タンパク質の活性(例えば、転写調節活性、酵素活性など)を変更した、または適合性を高めた核酸(例えば、CasYガイドRNAをコードする核酸、CasY融合タンパク質をコードする核酸など)及びタンパク質(例えば、野生型タンパク質または変異体タンパク質由来CasY融合タンパク質)である。そのようなポリペプチドの類似体として、天然L-アミノ酸以外、例えばD-アミノ酸または非天然合成アミノ酸の残基を含むものが挙げられる。アミノ酸残基の一部または全部がD-アミノ酸に置換されていてもよい。

【0183】

本開示のCasYポリペプチドは、当技術分野で公知の従来法を用いて、*in vitro*合成により調製することができる。種々の市販の合成装置、例えば、Applied Biosystems, Inc., Beckmanなどによる自動合成装置が提供されている。合成装置を使用して、天然アミノ酸が非天然アミノ酸に置換されていてもよい。具体的な配列及び調製方法は、利便性、経済性、要求される純度などによって決定されることになる。

【0184】

必要に応じて、合成過程または発現時に種々の基をペプチドに導入して、他の分子との、または表面との連結を可能にすることができる。したがって、チオエーテルの生成にシステインを、金属イオン錯体との結合にヒスチジンを、アミドまたはエステル形成にカルボキシル基を、アミドの形成にアミノ基を使用することなどが可能である。

【0185】

本開示の Cas Y ポリペプチドはまた、組換え合成の従来法に従って単離及び精製されていてもよい。発現宿主から溶解物を調製し、その溶解物を高速液体クロマトグラフィー (HPLC)、排除クロマトグラフィー、ゲル電気泳動、アフィニティクロマトグラフィー、またはその他の精製技術を使用して精製することができる。ほとんどの場合、使用される組成物は、生成物の調製方法及びその精製と関連する夾雑物に対して、目的生成物が 20 重量%以上、より一般的には 75 重量%以上、好ましくは 95 重量%以上を占め、また治療目的には、通常 99.5 重量%以上を占める。通常、パーセンテージは総タンパク質を基準とする。したがって、場合によって、本開示の Cas Y ポリペプチド、または Cas Y 融合ポリペプチドは、少なくとも純度 80%、少なくとも純度 85%、少なくとも純度 90%、少なくとも純度 95%、少なくとも純度 98%、または少なくとも純度 99% である (例えば、夾雑物、非 Cas Y タンパク質、または他の巨大分子などを含まない)。

【0186】

標的核酸 (例えば、ゲノム DNA) に対する切断または任意の所望する修飾、または標的核酸と会合したポリペプチドに対する任意の所望する修飾を誘導するには、本開示の Cas Y ガイド RNA 及び/または Cas Y ポリペプチド、及び/またはドナー鋳型配列は、それらが核酸またはポリペプチドとして導入されるかどうかにかかわらず、約 30 分 ~ 約 24 時間、例えば、1 時間、1.5 時間、2 時間、2.5 時間、3 時間、3.5 時間、4 時間、5 時間、6 時間、7 時間、8 時間、12 時間、16 時間、18 時間、20 時間、または約 30 分 ~ 約 24 時間の間の他の任意の期間、細胞に供給され、約 1 日 ~ 約 4 日ごと、例えば、1.5 日ごと、2 日ごと、3 日ごと、または約 1 日 ~ 約 4 日ごとの間の他の任意の頻度でこれを繰り返すことができる。作用物質 (複数可) は対象細胞に 1 回以上、例えば 1 回、2 回、3 回、または 3 回超供給されてもよく、各接触事象の後、ある程度の時間、例えば 16 ~ 24 時間、細胞を作用物質 (複数可) とインキュベートし、その後、培地を新鮮な培地と交換し、細胞をさらに培養することができる。

【0187】

2 つ以上の異なる標的化複合体 (例えば、同一または異なる標的核酸内の異なる配列に相補的な 2 つの異なる Cas Y ガイド RNA) を細胞に供給する場合、各複合体を (例えば、2 つのポリペプチド及び/または核酸などとして) 同時に供給するか、または同時に送達することができる。あるいは、例えば最初に供給される標的化複合体に続いて、第 2 の標的化複合体を供給する、またはその逆の順番で、それらを連続して供給することができる。

【0188】

標的細胞への DNA ベクターの送達を改善するには、例えば、リポプレックス及びポリプレックスを使用することによって、DNA を損傷から保護し、細胞内への進入を促進することができる。したがって、場合によって、本開示の核酸 (例えば、本開示の組換え発現ベクター) を、ミセルまたはリポソームのような組織構造の脂質で覆うことができる。組織構造が DNA と複合体を形成している場合、それをリポプレックスと呼ぶ。脂質には、アニオン性 (負電荷)、中性、またはカチオン性 (正電荷) の 3 種類がある。カチオン性脂質を利用するリポプレックスは、遺伝子導入での有用性が実証されている。カチオン性脂質は、その正電荷により、負電荷の DNA と自然に複合体を形成する。また、電荷の結果として細胞膜と相互作用する。その後、リポプレックスのエンドサイトーシスが発生し、DNA は細胞質に放出される。カチオン性脂質は、細胞による DNA の分解を阻止する。

【0189】

ポリマーとDNAとの複合体をポリプレックスと呼ぶ。ほとんどのポリプレックスはカチオン性ポリマーからなり、その産生はイオン相互作用によって制御される。ポリプレックスとリボプレックスの作用方法の大きな違いの一つは、ポリプレックスはDNA積荷を細胞質に放出できないため、この目的のために、不活性化アデノウイルスのようなエンドソーム溶解剤（エンドサイトーシス中に作られるエンドソームを溶解する）との同時トランスフェクションを行う必要があることである。ただし、必ずしもそうでない場合もあり、ポリエチレンイミンのようなポリマーは、キトサン及びトリメチルキトサンと同様、独自のエンドソーム分解方法を有している。

【0190】

デンドリマー、すなわち球状の形状を有する高度に分岐した巨大分子もまた、幹細胞の遺伝子改変に使用することができる。デンドリマー粒子の表面を官能化して、その特性を変化させることができる。具体的には、カチオン性デンドリマー（すなわち、正の表面電荷を有するもの）を構築することができる。DNAプラスミドなどの遺伝子材料が存在する場合、電荷の相補性により、核酸とカチオン性デンドリマーとが一時的に会合される。デンドリマーと核酸の複合体が目的地に到達すると、エンドサイトーシスにより、複合体を細胞内に取り込むことができる。

【0191】

場合によって、本開示の核酸（例えば、発現ベクター）には、対象となるガイド配列に対する挿入部位を含む。例えば、核酸は、対象となるガイド配列に対する挿入部位を含み得るが、この挿入部位は、ガイド配列が所望の標的配列にハイブリダイズするように変化するとき、変化しないCasYガイドRNAの部分をコードするヌクレオチド配列（例えば、ガイドRNAのCasY結合局面に寄与する配列、例えば、CasYガイドRNAのdsRNA二重鎖（複数可）に寄与する配列。ガイドRNAのこの部分は、ガイドRNAの「骨格」または「定常領域」と呼ばれることもある）に直接隣接する。したがって、場合によって、本発明の核酸（例えば、発現ベクター）には、ガイドRNAのガイド配列部分をコードする部分が挿入配列（挿入部位）であることを除いて、CasYガイドRNAをコードするヌクレオチド配列を含む。挿入部位は、目的配列の挿入に使用される任意のヌクレオチド配列である。種々の技術で使用される「挿入部位」は当業者に公知であり、任意の利便な挿入部位を使用することができる。挿入部位は、核酸配列を操作するための任意の方法に応じたものであり得る。例えば、場合によって、挿入部位は、多重クローニング部位（MCS）（例えば、1つ以上の制限酵素認識配列を含む部位）、ライゲーション非依存性クローニングのための部位、組換えによるクローニングのための部位（例えば、att部位に基づく組換え）、CRISPR/Cas（例えばCas9）を用いる技術によって認識されるヌクレオチド配列などである。

【0192】

挿入部位は、任意の望ましい長さであってよく、挿入部位の種類に依存し得る（例えば、その部位が（及びいくつの部位が）1つ以上の制限酵素認識配列を含むかどうか、その部位がCRISPR/Casタンパク質に対する標的部位を含むかどうかなどに依存し得る）。場合によって、本発明の核酸の挿入部位は、長さが3ヌクレオチド（nt）以上（例えば、長さが5nt以上、8nt以上、10nt以上、15nt以上、17nt以上、18nt以上、19nt以上、20nt以上、または25nt以上、または30nt以上）である。場合によって、本発明の核酸の挿入部位の長さは、2～50ヌクレオチド（nt）（例えば、2～40nt、2～30nt、2～25nt、2～20nt、5～50nt、5～40nt、5～30nt、5～25nt、5～20nt、10～50nt、10～40nt、10～30nt、10～25nt、10～20nt、17～50nt、17～40nt、17～30nt、17～25nt）の範囲の長さを有する。場合によって、本発明の核酸の挿入部位の長さは、5～40ntの範囲の長さを有している。

【0193】

核酸修飾

いくつかの実施形態では、本発明の核酸（例えば、CasYガイドRNA）は、新機能

10

20

30

40

50

または改良機能（例えば、改善された安定性）をもつ核酸を提供するように、1つ以上の修飾、例えば塩基修飾、骨格修飾などを有する。ヌクレオシドは塩基-糖の組み合わせである。ヌクレオシドの塩基部分は通常、複素環塩基である。そのような複素環塩基のうち最も一般的な2つのクラスがプリン及びピリミジンである。ヌクレオチドは、ヌクレオシドの糖部分に共有結合したリン酸基をさらに含むヌクレオシドである。ペントフラノシル糖を含むヌクレオシドの場合、リン酸基は、糖の2'、3'、または5'位のヒドロキシル部分に結合する可能性がある。オリゴヌクレオチドを形成する場合、そのリン酸基が、隣接するヌクレオシドを互いに共有結合させて直鎖の高分子化合物を形成する。それに続いて、この直鎖の高分子化合物のそれぞれの末端がさらに結合され、環状化合物を形成することも可能であるが、直鎖状化合物が好適である。また、直鎖状化合物は、内部ヌクレオチド塩基の相補性を有する場合があります、その結果、完全にまたは部分的に二本鎖の化合物を生成するような方法で折り畳まれる場合がある。オリゴヌクレオチド内では、リン酸基は一般に、オリゴヌクレオチドのヌクレオシド間骨格を形成すると言われる。RNA及びDNAの通常の結合または骨格は、3'と5'とのホスホジエステル結合である。

10

20

30

40

50

【0194】

好適な核酸修飾として、2'-Oメチル修飾ヌクレオチド、2'-フルオロ修飾ヌクレオチド、ロックド核酸(LNA)修飾ヌクレオチド、ペプチド核酸(PNA)修飾ヌクレオチド、ホスホロチオエート架橋を有するヌクレオチド、及び5'キャップ(例えば、7-メチルグアニル酸キャップ(m7G))が挙げられるが、これらに限定されない。追加詳細及びその他の修飾は、以下に記載されている。

【0195】

2'-O-メチル修飾ヌクレオチド(2'-O-メチルRNAとも称する)は、転写後修飾として起こる、tRNA及び他の低分子RNAに見られる天然のRNA修飾である。2'-O-メチルRNAを含むオリゴヌクレオチドを直接合成することができる。この修飾により、RNA:RNA二重鎖のT_mは増加するが、RNA:DNAの安定性にはわずかにしか変化が生じない。これは、一本鎖リボヌクレアーゼによる攻撃に対して安定であり、通常、DNAよりも5~10倍DNaseに反応しにくい。この修飾は、標的メッセージに対する安定性及び結合親和性を増加させる手段としてアンチセンスオリゴによく使用される。

【0196】

2'-フルオロ修飾ヌクレオチド(例えば、2'-フルオロ塩基)は、結合親和性(T_m)を増加させ、また天然のRNAと比較した場合に、ある程度の相対的なヌクレアーゼ耐性を付与するフッ素修飾リボースを有する。この修飾は、血清または他の体液中での安定性を改善するためにリボザイム及びsiRNAによく使用される。

【0197】

LNA塩基は、C3'エンド位置に塩基をロックするリボース骨格への修飾を有し、RNAがA型らせん二重鎖形状をとりやすくする。この修飾はT_mを著しく増加させ、また極めてヌクレアーゼに耐性である。3'末端以外の任意の位置で、複数のLNA挿入をオリゴに配置することができる。アンチセンスオリゴからハイブリダイゼーションプローブ、さらにはSNP検出及び対立遺伝子特異的PCRに至る用途が記載されている。また、LNAによってT_mの大きな増加が得られることから、プライマー二量体の形成、ならびに自己ヘアピン形成の増加をもたらすこともできる。場合によって、単一のオリゴに組み込まれるLNAの数は10塩基以下である。

【0198】

ホスホロチオエート(PS)結合(すなわち、ホスホロチオエート架橋)では、核酸(例えば、オリゴ)のリン酸骨格の非架橋酸素が硫黄原子に置換されている。この修飾により、ヌクレオチド間結合はヌクレアーゼ分解に対して耐性になる。ホスホロチオエート結合を、オリゴの5'末端または3'末端の最後の3~5ヌクレオチドの間に導入することで、エキソヌクレアーゼ分解を阻害することができる。オリゴ内に(例えば、オリゴ全体にわたって)ホスホロチオエート結合を含めると、エンドヌクレアーゼによる攻撃もまた軽

減できる効果がある。

【0199】

いくつかの実施形態では、本発明の核酸は、2'-O-メチル修飾ヌクレオチドである1つ以上のヌクレオチドを有する。いくつかの実施形態では、本発明の核酸（例えば、dsRNA、siNAなど）は、1つ以上の2'フルオロ修飾ヌクレオチドを有する。いくつかの実施形態では、本発明の核酸（例えば、dsRNA、siNAなど）は、1つ以上のLNA塩基を有する。いくつかの実施形態では、本発明の核酸（例えば、dsRNA、siNAなど）は、ホスホロチオエート結合により連結された1つ以上のヌクレオチドを有する（すなわち、本発明の核酸は、1つ以上のホスホロチオエート架橋を有する）。いくつかの実施形態では、本発明の核酸（例えば、dsRNA、siNAなど）は、5'キ
10
ャップ（例えば、7-メチルグアニル酸キップ（m7G））を有する。いくつかの実施形態では、本発明の核酸（例えば、dsRNA、siNAなど）は、修飾ヌクレオチドの組み合わせを有する。例えば、本発明の核酸（例えば、dsRNA、siNAなど）は、5'キップ（例えば、7-メチルグアニル酸キップ（m7G））を有し、さらに他の修飾（例えば、2'-O-メチルヌクレオチド及び/または2'フルオロ修飾ヌクレオチド及び/またはLNA塩基及び/またはホスホロチオエート架橋）を有する1つ以上のヌクレオチドを有する可能性がある。

【0200】

修飾骨格及び修飾ヌクレオシド間結合

修飾を含む好適な核酸（例えば、CasYガイドRNA）の例として、修飾骨格または非天然ヌクレオシド間結合を含む核酸がある。修飾骨格を有する核酸には、骨格にリン原子を保持するもの及び骨格にリン原子をもたないものを含む。
20

【0201】

リン原子をそこに含む好適な修飾オリゴヌクレオチド骨格としては、例えば、ホスホロチオエート、キラルホスホロチオエート、ホスホロジチオエート、ホスホトリエステル、アミノアルキルホスホトリエステル、メチル及び他のアルキルホスホネート（3'-アルキレンホスホネート、5'-アルキレンホスホネート、及びキラルホスホネートを含む）、ホスフィネート、ホスホロアミダート（3'-アミノホスホロアミダート及びアミノアルキルホスホロアミダートを含む）、ホスホロジアミダート、チオノホスホロアミダート、チオノアルキルホスホネート、チオノアルキルホスホトリエステル、セレノホスフェート、及びボラノホスフェートが挙げられ、これらは通常の3'-5'結合を有するもの、2'-5'結合されたこれらの類似体、ならびに1つ以上の反転極性を有する（ヌクレオチド間結合が、3'と3'、5'と5'、または2'と2'の結合である）ものを含む。反転極性を有する好適なオリゴヌクレオチドは、最も3'側のヌクレオチド間結合に単一の3'と3'の結合、すなわち塩基性であり得る（核酸塩基がないか、またはその代わりにヒドロキシル基をもつ）単一の反転ヌクレオシド残基を含む。様々な塩（例えば、カリウムまたはナトリウムなど）、混合塩、及び遊離酸形態も含まれる。
30

【0202】

いくつかの実施形態では、本発明の核酸は、1つ以上のホスホロチオエート結合及び/またはヘテロ原子ヌクレオシド間結合、特に-CH₂-NH-O-CH₂-、-CH₂-N(CH₃)-O-CH₂-（メチレン（メチルイミノ）またはMMI骨格と呼ばれる）、-CH₂-O-N(CH₃)-CH₂-、-CH₂-N(CH₃)-N(CH₃)-CH₂-、及び-O-N(CH₃)-CH₂-CH₂-を含む（ここで、天然のホスホジエステルヌクレオチド間連結は、-O-P(=O)(OH)-O-CH₂-として表される）。MMI型ヌクレオシド間結合は、上記で引用した米国特許第5,489,677号に記載されており、その開示内容全体が参照により本明細書に組み込まれる。好適なアミドヌクレオシド間結合は、米国特許第5,602,240号に記載されており、その開示内容全体が参照により本明細書に組み込まれる。
40

【0203】

また、例えば米国特許第5,034,506号に記載されるようなモルホリノ骨格構造
50

を有する核酸も適している。例えば、いくつかの実施形態では、本発明の核酸は、リボース環の代わりに、6員のモルホリノ環を含む。これらの実施形態のいくつかは、ホスホジエステル結合の代わりに、ホスホロジアミダートまたは他の非ホスホジエステルヌクレオシド間結合である。

【0204】

そこにリン原子を含まない好適な修飾ポリヌクレオチド骨格は、短鎖アルキルもしくはシクロアルキルヌクレオシド間結合、混合ヘテロ原子及びアルキルもしくはシクロアルキルヌクレオシド間結合、または1つ以上の短鎖ヘテロ原子もしくは複素環ヌクレオシド間結合によって形成される骨格を有する。これには、モルホリノ結合（ヌクレオシドの糖部分から部分的に形成される）；シロキサノ骨格；スルフィド骨格、スルホキシド骨格、及びスルホン骨格；ホルムアセチル骨格及びチオホルムアセチル骨格；メチレンホルムアセチル骨格及びチオホルムアセチル骨格；リボアセチル骨格；アルケンを含む骨格；スルファメート骨格；メチレンイミノ骨格及びメチレンヒドラジド骨格；スルホネート骨格及びスルホンアミド骨格；アミド骨格；ならびにN、O、S、及びC H₂ 構成要素部分の混合を有するその他の骨格をもつものを含む。

【0205】

模倣物

本発明の核酸は、核酸模倣物であり得る。ポリヌクレオチドに適用される場合、用語「模倣物」は、フラノース環のみまたはフラノース環とヌクレオチド間結合の両方が、非フラノース基で置換されているポリヌクレオチドを含むことを意図し、またフラノース環のみの置換は当技術分野で糖代用物であると称される。適切な標的核酸とのハイブリダイゼーションにおいて、複素環塩基部分または修飾複素環塩基部分は維持される。そのような核酸の1つが、優れたハイブリダイゼーション特性を有することが示されているポリヌクレオチド模倣物であり、ペプチド核酸（PNA）と呼ばれる。PNAでは、ポリヌクレオチドの糖骨格が、アミドを含む骨格、特にアミノエチルグリシン骨格で置換されている。ヌクレオチドは保持され、骨格のアミド部分のアザ窒素原子に直接的または間接的に結合される。

【0206】

優れたハイブリダイゼーション特性を有することが報告されているポリヌクレオチド模倣物の1つが、ペプチド核酸（PNA）である。PNA化合物の骨格は、アミドを含む骨格をPNAに与える2つ以上の連結されたアミノエチルグリシン単位である。複素環塩基部分は、骨格のアミド部分のアザ窒素原子に直接的または間接的に結合される。PNA化合物の調製について記載している代表的な米国特許として、限定はされないが、米国特許第5,539,082号；第5,714,331号；及び第5,719,262号が挙げられ、各開示内容全体が参照により本明細書に組み込まれる。

【0207】

研究されているポリヌクレオチド模倣物の別のクラスは、モルホリノ環に結合された複素環塩基を有する連結されたモルホリノ単位（モルホリノ核酸）を基にしている。モルホリノ核酸において、モルホリノモノマー単位を結合する連結基がいくつか報告されている。連結基のクラスの1つは、非イオン性オリゴマー化合物を得られるように選択されている。非イオン性モルホリノ系オリゴマー化合物は、細胞タンパク質との望ましくない相互作用を有する可能性が低い。モルホリノ系ポリヌクレオチドは、細胞タンパク質との望ましくない相互作用を生じさせる可能性が低い非イオン性のオリゴヌクレオチド模倣物である（Dwaine A. Braasch and David R. Corey, *Biochemistry*, 2002, 41(14), 4503-4510）。モルホリノ系ポリヌクレオチドは、米国特許第5,034,506号に記載されており、その開示内容全体が参照により本明細書に組み込まれる。モノマーサブユニットを結合する多種多様な連結基を有する、モルホリノクラスのポリヌクレオチドに属する様々な化合物が調製されている。

【0208】

50

ポリヌクレオチド模倣物のさらなるクラスは、シクロヘキセニル核酸 (CeNA) と呼ばれる。DNA/RNA分子に通常存在するフラノース環が、シクロヘキセニル環で置換されている。古典的なホスホロアミダイト化学に従って、CeNA DMT保護ホスホロアミダイトモノマーが調製され、オリゴマー化合物の合成に使用されてきた。CeNAで修飾された特異的部分を有する完全修飾CeNAオリゴマー化合物及びオリゴヌクレオチドが調製及び研究されている (Wang et al., J. Am. Chem. Soc., 2000, 122, 8595-8602を参照のこと。その開示内容全体が参照により本明細書に組み込まれる)。一般に、DNA鎖へのCeNAモノマーの組み込みは、DNA/RNAハイブリッドの安定性を増加させる。RNA及びDNAと複合体を形成するCeNAオリゴアデニル酸は、天然複合体と同様の安定性を補う。研究によれば、天然の核酸構造にCeNA構造を組み込むと、立体構造の容易な適合が進行されることがNMR及び円偏光二色性によって示された。

10

【0209】

さらなる修飾には、糖環の4'炭素原子に2'-ヒドロキシル基が結合し、それによって2'-C, 4'-C-オキシメチレン結合を形成し、それによって二環式糖部分を形成するロックド核酸 (LNA) を含む。この連結基は、メチレン (-CH₂-)、すなわち、nが1または2である、2'酸素原子と4'炭素原子を架橋する基であり得る (Singh et al., Chem. Commun., 1998, 4, 455-456。その開示内容全体が参照により本明細書に組み込まれる)。LNA及びLNA類似体は、相補的DNA及びRNAとの非常に高い二重鎖熱安定性 (T_m = +3~+10)、3'-エキソヌクレアーゼ分解に対する安定性、及び良好な溶解特性を示す。LNAを含む強力かつ非毒性のアンチセンスオリゴヌクレオチドが記載されている (例えば、Wahlestedt et al., Proc. Natl. Acad. Sci. U.S.A., 2000, 97, 5633-5638。その開示内容全体が参照により本明細書に組み込まれる)。

20

【0210】

LNAモノマーのアデニン、シトシン、グアニン、5-メチル-シトシン、チミン、及びウラシルの合成及び調製に加え、それらのオリゴマー合成、及び核酸認識特性が記載されている (例えば、Koshkin et al., Tetrahedron, 1998, 54, 3607-3630。その開示内容全体が参照により本明細書に組み込まれる)。LNA及びその調製については、WO98/39352及びWO99/14226、ならびに米国出願第20120165514号、第20100216983号、第20090041809号、第20060117410号、第20040014959号、第20020094555号、及び第20020086998号にも記載されており、その開示内容全体が参照により本明細書に組み込まれる。

30

【0211】

修飾糖部分

本発明の核酸はまた、1つ以上の置換された糖部分を含み得る。好適なポリヌクレオチドは、OH; F; O-, S-, もしくはN-アルキル; O-, S-, もしくはN-アルケニル; O-, S-, もしくはN-アルキニル; またはO-アルキル-O-アルキルから選択される糖置換基を含み、ここでのアルキル、アルケニル、及びアルキニルは、置換または非置換のC₁~C₁₀アルキルまたはC₂~C₁₀アルケニル及びアルキニルであり得る。特に、O((CH₂)_nO)_mCH₃、O(CH₂)_nOCH₃、O(CH₂)_nNH₂、O(CH₂)_nCH₃、O(CH₂)_nONH₂、及びO(CH₂)_nON((CH₂)_nCH₃)₂ (ここで、n及びmは1~約10である) が適している。他の好適なポリヌクレオチドは、C₁~C₁₀低級アルキル、置換低級アルキル、アルケニル、アルキニル、アルカリル、アラルキル、O-アルカリルまたはO-アラルキル、SH、SCH₃、OCN、Cl、Br、CN、CF₃、OCF₃、SOCH₃、SO₂CH₃、ONO₂、NO₂、N₃、NH₂、ヘテロシクロアルキル、ヘテロシクロアルカリル、アミノアルキルアミノ、ポリアルキルアミノ、置換シリル、RNA切断基、レポーター基、インターカレーター、オリゴヌクレオチドの薬物動態特性を改善するための基、または

40

50

オリゴヌクレオチドの薬力学特性を改善するための基、及び同様の特性を有する他の置換基から選択される糖置換基を含む。好適な修飾には、2'-メトキシエトキシ(2'-O-CH₂CH₂OCH₃、2'-O-(2'-メトキシエチル)または2'-MOEとしても知られる)(Martin et al., Helv. Chim. Acta, 1995, 78, 486-504。その開示内容全体が参照により本明細書に組み込まれる)、すなわちアルコキシアルコキシ基を含む。さらに好適な修飾には、本明細書で以下の実施例に記載するような、2'-DMAOEとしても知られる2'-ジメチルアミノオキシエトキシ、すなわちO(CH₂)₂ON(CH₃)₂基、及び2'-ジメチルアミノエトキシエトキシ(当技術分野で2'-O-ジメチル-アミノ-エトキシ-エチルまたは2'-DMAEOEとしても知られる)、すなわち2'-O-CH₂-O-CH₂-N(CH₃)₂を含む。

【0212】

他の好適な糖置換基として、メトキシ(-O-CH₃)、アミノプロポキシ(-O-CH₂CH₂CH₂NH₂)、アリル(-CH₂-CH=CH₂)、-O-アリル(-O-CH₂-CH=CH₂)、及びフルオロ(F)が挙げられる。2'-糖置換基は、アラビノ位置(上)であっても、またはリボ位置(下)であってもよい。好適な2'-アラビノ修飾は、2'-Fである。同様の修飾はまた、オリゴマー化合物の他の位置、特に3'末端ヌクレオチドまたは2'-5'結合オリゴヌクレオチドの糖の3'位置、及び5'末端ヌクレオチドの5'位置でなされ得る。オリゴマー化合物はまた、ペントフラノシル糖の代わりにシクロブチル部分などの糖模倣物を有していてもよい。

【0213】

塩基の修飾及び置換

本発明の核酸はまた、(当技術分野で単に「塩基」と呼ばれることが多い)核酸塩基の修飾または置換を含み得る。本明細書で使用される場合、「非修飾」または「天然」核酸塩基には、プリン塩基のアデニン(A)及びグアニン(G)、ならびにピリミジン塩基のチミン(T)、シトシン(C)、及びウラシル(U)を含む。修飾核酸塩基には、他の合成核酸塩基及び天然核酸塩基を含み、例えば、5-メチルシトシン(5-me-C)、5-ヒドロキシメチルシトシン、キサントシン、ヒポキサントシン、2-アミノアデニン、アデニン及びグアニンの6-メチル及び他のアルキル誘導体、アデニン及びグアニンの2-プロピル及び他のアルキル誘導体、2-チオウラシル、2-チオチミン及び2-チオシトシン、5-ハロウラシル及びシトシン、5-プロピニル(-C=C-CH₃)ウラシル及びシトシン、ならびにピリミジン塩基の他のアルキニル誘導体、6-アゾウラシル、シトシン及びチミン、5-ウラシル(シュードウラシル)、4-チオウラシル、8-ハロ、8-アミノ、8-チオール、8-チオアルキル、8-ヒドロキシル及び他の8-置換アデニン及びグアニン、5-ハロ、特に5-プロモ、5-トリフルオロメチル及び他の5-置換ウラシル及びシトシン、7-メチルグアニン及び7-メチルアデニン、2-F-アデニン、2-アミノアデニン、8-アザグアニン及び8-アザアデニン、7-デアザグアニン及び7-デアザアデニン、ならびに3-デアザグアニン及び3-デアザアデニンが挙げられる。さらなる修飾核酸塩基として、フェノキサジンシチジン(1H-ピリミド(5,4-b)(1,4)ベンゾキサジン-2(3H)-オン)、フェノチアジンシチジン(1H-ピリミド(5,4-b)(1,4)ベンゾチアジン-2(3H)-オン)などの三環ピリミジン、置換フェノキサジンシチジン(例えば、9-(2-アミノエトキシ)-H-ピリミド(5,4-(b)(1,4)ベンゾキサジン-2(3H)-オン)、カルバゾールシチジン(2H-ピリミド(4,5-b)インドール-2-オン)、ピリドインドールシチジン(H-ピリド(3',2':4,5)ピロロ(2,3-d)ピリミジン-2-オン)などのGクランプが挙げられる。

【0214】

複素環塩基部分にはまた、プリン塩基またはピリミジン塩基が他の複素環で置換されているもの、例えば7-デアザ-アデニン、7-デアザグアニン、2-アミノピリジン、及び2-ピリドンを含み得る。さらなる核酸塩基として、米国特許第3,687,808

号に記載のもの、The Concise Encyclopedia Of Polymer Science And Engineering, pages 858 - 859, Kroschwitz, J. I., ed. John Wiley & Sons, 1990に記載のもの、Englisch et al., Angewandte Chemie, International Edition, 1991, 30, 613によって記載されるもの、及びSanghvi, Y. S., Chapter 15, Antisense Research and Applications, pages 289 - 302, Crooke, S. T. and Lebleu, B., ed., CRC Press, 1993によって記載されるものが挙げられ、その開示内容全体が参照により本明細書に組み込まれる。これらのうち、ある種の核酸塩基は、オリゴマー化合物の結合親和性を増加させるのに有用である。このような核酸塩基として、5 - 置換ピリミジン、6 - アザピリミジン、及びN - 2、N - 6及びO - 6置換プリン、例えば、2 - アミノプロピルアデニン、5 - プロピニルウラシル、及び5 - プロピニルシトシンが挙げられる。5 - メチルシトシン置換は、核酸の二重鎖安定性を0.6 ~ 1.2 上昇させることが示されており (Sanghvi et al., eds., Antisense Research and Applications, CRC Press, Boca Raton, 1993, pp. 276 - 278。その開示内容全体が参照により本明細書に組み込まれる)、例えば、2' - O - メトキシエチル糖修飾と組み合わせた場合、好適な塩基置換である。

10

20

30

40

50

【0215】

コンジュゲート

本発明の核酸について可能な別の修飾は、オリゴヌクレオチドの活性、細胞分布、または細胞取り込みを増強する1つ以上の部分またはコンジュゲートを、ポリヌクレオチドに化学的に連結することを伴う。これらの部分またはコンジュゲートには、第1級または第2級ヒドロキシル基などの官能基に共有結合するコンジュゲート基を含み得る。コンジュゲート基としては、インターカレーター、レポーター分子、ポリアミン、ポリアミド、ポリエチレングリコール、ポリエーテル、オリゴマーの薬力学的特性を向上させる基、及びオリゴマーの薬物動態特性を向上させる基が挙げられるが、これらに限定されない。好適なコンジュゲート基として、コレステロール、脂質、リン脂質、ビオチン、フェナジン、葉酸、フェナントリジン、アントラキノン、アクリジン、フルオレセイン、ローダミン、クマリン、及び色素が挙げられるが、これらに限定されない。薬力学的特性を向上させる基には、取り込みを改善する基、分解に対する耐性を高める基、及び/または標的核酸との配列特異的ハイブリダイゼーションを強化する基が含まれる。薬物動態特性を向上させる基には、本発明の核酸の取り込み、分布、代謝、または排出を改善する基が含まれる。

【0216】

コンジュゲート部分として、限定はされないが、脂質部分が挙げられ、これには例えば、コレステロール部分 (Letsinger et al., Proc. Natl. Acad. Sci. USA, 1989, 86, 6553 - 6556)、コール酸 (Manoharan et al., Bioorg. Med. Chem. Lett., 1994, 4, 1053 - 1060)、チオエーテル、例えばヘキシル - S - トリチルチオール (Manoharan et al., Ann. N. Y. Acad. Sci., 1992, 660, 306 - 309; Manoharan et al., Bioorg. Med. Chem. Lett., 1993, 3, 2765 - 2770)、チオコレステロール (Oberhauser et al., Nucl. Acids Res., 1992, 20, 533 - 538)、脂肪族鎖、例えばドデカンジオール残基もしくはウンデシル残基 (Saison - Behmoaras et al., EMBO J., 1991, 10, 1111 - 1118; Kabanov et al., FEBS Lett., 1990, 259, 327 - 330; Svinarchuk et al., Biochimie, 1993, 75, 49 - 54)、リン脂質、例えばジ - ヘキサデシル - rac - グリセロールもしくはトリエチルアンモニウム 1, 2 - ジ - O - ヘキサデシル - rac - グリセロール -

3-H-ホスホネート (Manoharan et al., Tetrahedron Lett., 1995, 36, 3651-3654; Shea et al., Nucl. Acids Res., 1990, 18, 3777-3783)、ポリアミンもしくはポリエチレングリコール鎖 (Manoharan et al., Nucleosides & Nucleotides, 1995, 14, 969-973)、またはアダマンタン酢酸 (Manoharan et al., Tetrahedron Lett., 1995, 36, 3651-3654)、パルミチル部分 (Mishra et al., Biochim. Biophys. Acta, 1995, 1264, 229-237)、またはオクタデシルアミンもしくはヘキシルアミノ-カルボニル-オキシコレステロール部分 (Crooke et al., J. Pharmacol. Exp. Ther., 1996, 277, 923-937) などがある。

【0217】

コンジュゲートには、「タンパク質導入ドメイン」すなわちPTD (細胞膜透過性ペプチド、CPPとも呼ばれる) を含み得るが、これは脂質二重層、ミセル、細胞膜、オルガネラ膜、または小胞膜の透過を促進するポリペプチド、ポリヌクレオチド、炭水化物、または有機もしくは無機化合物を指す場合がある。極性小分子から大きな巨大分子及び/またはナノ粒子までを範囲とし得る別の分子に結合したPTDは、例えば、細胞外空間から細胞内空間への、またはサイトゾルからオルガネラ (例えば、核) 内への移行といった、分子の膜透過を促進する。いくつかの実施形態では、PTDは、外因性ポリヌクレオチドの3'末端に共有結合されている。いくつかの実施形態では、PTDは、外因性ポリヌクレオチドの5'末端に共有結合されている。例示的なPTDとして、最小のウンデカペプチドタンパク質導入ドメイン (YGRKKRRQRRR: 配列番号112を含むHIV-1 TATの47~57残基に対応); 細胞への直接導入に十分な複数のアルギニン (例えば、3、4、5、6、7、8、9、10、または10~50のアルギニン) を含むポリアルギニン配列; VP22ドメイン (Zender et al. (2002) Cancer Gene Ther. 9(6): 489-96); Drosophila Antennapediaタンパク質導入ドメイン (Noguchi et al. (2003) Diabetes 52(7): 1732-1737); 短縮ヒトカルシトニンペプチド (Trehin et al. (2004) Pharm. Research 21: 1248-1256); ポリリジン (Wender et al. (2000) Proc. Natl. Acad. Sci. USA 97: 13003-13008); RRQRRTSKLMKR (配列番号113); トランスポータン GWTLNSAGYLLGKINLKALAAALAKKIL (配列番号114); KALAWEAKLAKALAKALAKHLAKALAKALKCEA (配列番号115); 及び RQIKIWFQNRRMKWKK (配列番号116) が挙げられるが、これらに限定されない。例示的なPTDとして、限定はされないが、YGRKKRRQRRR (配列番号117)、RKKRRQRRR (配列番号118); 3アルギニン残基~50アルギニン残基までからなるアルギニンホモポリマーが挙げられ、例示的なPTDドメインアミノ酸配列には、限定はされないが、以下のいずれかが含まれる: YGRKKRRQRRR (配列番号119); RKKRRQR (配列番号120); YARAAARQARA (配列番号121); THRLPRRRRRR (配列番号122); 及び GGRRARRRRRRR (配列番号123)。いくつかの実施形態では、PTDは、活性化可能なCPP (ACPP) である (Aguilera et al. (2009) Integr Biol (Camb) June; 1(5-6): 371-381)。ACPPには、対となるポリアニオン (例えば、Glu9すなわち「E9」) に切断可能なリンカーを介して接続されたポリカチオン性CPP (例えば、ARG9すなわち「R9」) が含まれ、これが実効電荷をほぼゼロに低下させ、それによって細胞への接着及び取り込みを阻害する。リンカーの切断時に、ポリアニオンが遊離して、ポリアルギニン及びそれに備わる接着性が局所的に露出され、それによってACPPの膜透過が「活性化」する。

【0218】

標的細胞への構成要素の導入

本開示の CasY ガイド RNA (またはそれをコードするヌクレオチド配列を含む核酸) 及び/または CasY ポリペプチド (またはそれをコードするヌクレオチド配列を含む核酸) 及び/または本開示の CasY 融合ポリペプチド (または本開示の CasY 融合ポリペプチドをコードするヌクレオチド配列を含む核酸) 及び/またはドナーポリヌクレオチド (ドナー鑄型) は、周知の様々な方法のいずれかによって宿主細胞に導入することができる。

【0219】

種々の化合物及び方法のいずれかを使用して、本開示の CasY システムを標的細胞に送達することができる。例えば、CasY システムは以下を含む： a) 本開示の CasY 10
 ポリペプチド及び CasY ガイド RNA； b) 本開示の CasY ポリペプチド、CasY
 ガイド RNA、及びドナー鑄型核酸； c) 本開示の CasY 融合ポリペプチド及び Cas
 Y ガイド RNA； d) 本開示の CasY 融合ポリペプチド、CasY ガイド RNA、及び
 ドナー鑄型核酸； e) 本開示の CasY ポリペプチドをコードする mRNA 及び Cas Y
 ガイド RNA； f) 本開示の CasY ポリペプチドをコードする mRNA、CasY ガイ
 ド RNA、及びドナー鑄型核酸； g) 本開示の CasY 融合ポリペプチドをコードする m
 RNA 及び Cas Y ガイド RNA； h) 本開示の CasY 融合ポリペプチドをコードする
 mRNA、CasY ガイド RNA、及びドナー鑄型核酸； i) 本開示の CasY ポリペプ
 チドをコードするヌクレオチド配列及び CasY ガイド RNA をコードするヌクレオチド
 20
 配列を含む組換え発現ベクター； j) 本開示の CasY ポリペプチドをコードするヌクレ
 オチド配列、CasY ガイド RNA をコードするヌクレオチド配列、及びドナー鑄型核酸
 をコードするヌクレオチド配列を含む組換え発現ベクター； k) 本開示の CasY 融合ポリ
 ペプチドをコードするヌクレオチド配列及び CasY ガイド RNA をコードするヌクレ
 オチド配列を含む組換え発現ベクター； l) 本開示の CasY 融合ポリペプチドをコード
 するヌクレオチド配列、CasY ガイド RNA をコードするヌクレオチド配列、及びドナ
 ー鑄型核酸をコードするヌクレオチド配列を含む組換え発現ベクター； m) 本開示の Cas
 Y ポリペプチドをコードするヌクレオチド配列を含む第 1 の組換え発現ベクター、及び
 Cas Y ガイド RNA をコードするヌクレオチド配列を含む第 2 の組換え発現ベクター；
 n) 本開示の CasY ポリペプチドをコードするヌクレオチド配列を含む第 1 の組換え発
 30
 現ベクター、ならびに CasY ガイド RNA をコードするヌクレオチド配列及びドナー鑄
 型核酸を含む第 2 の組換え発現ベクター； o) 本開示の CasY 融合ポリペプチドをコー
 ドするヌクレオチド配列を含む第 1 の組換え発現ベクター、及び CasY ガイド RNA を
 コードするヌクレオチド配列を含む第 2 の組換え発現ベクター； p) 本開示の CasY 融
 合ポリペプチドをコードするヌクレオチド配列を含む第 1 の組換え発現ベクター、なら
 びに CasY ガイド RNA をコードするヌクレオチド配列及びドナー鑄型核酸を含む第 2 の
 組換え発現ベクター； q) 本開示の CasY ポリペプチドをコードするヌクレオチド配列
 、第 1 の CasY ガイド RNA をコードするヌクレオチド配列、及び第 2 の CasY ガイ
 ド RNA をコードするヌクレオチド配列を含む組換え発現ベクター；もしくは r) 本開示
 の CasY 融合ポリペプチドをコードするヌクレオチド配列、第 1 の CasY ガイド RN
 A をコードするヌクレオチド配列、及び第 2 の CasY ガイド RNA をコードするヌクレ
 40
 オチド配列を含む組換え発現ベクター；または (a) ~ (r) のいずれか 1 つの何らかの
 変形例。非限定的な例として、本開示の CasY システムは、脂質と組み合わせることが
 できる。別の非限定的な例として、本開示の CasY システムは、粒子と組み合わせる、
 または粒子に製剤化することができる。

【0220】

宿主細胞に核酸を導入する方法は、当技術分野で公知であり、任意の利便な方法を使用
 して、本発明の核酸 (例えば、発現構築物/ベクター) を標的細胞 (例えば、原核細胞、
 真核細胞、植物細胞、動物細胞、哺乳動物細胞、ヒト細胞など) に導入することができる
 。好適な方法として、例えば、ウイルス感染、トランスフェクション、コンジュゲーション、
 プロトプラスト融合、リポフェクション、エレクトロポレーション、リン酸カルシウ
 50

ム沈降、ポリエチレンイミン（PEI）介在型トランスフェクション、DEAE-デキストラン介在型トランスフェクション、リポソーム介在型トランスフェクション、パーティクルガン技術、リン酸カルシウム沈降、直接マイクロインジェクション、ナノ粒子介在型核酸送達（例えば、Panyam et al Adv Drug Deliv Rev. 2012 Sep 13. pii: S0169-409X(12)00283-9. doi: 10.1016/j.addr.2012.09.023を参照）などが挙げられる。

【0221】

場合によって、本開示のCasYポリペプチドは、CasYポリペプチドをコードする核酸（例えば、mRNA、DNA、プラスミド、発現ベクター、ウイルスベクターなど）として提供される。場合によって、本開示のCasYポリペプチドは、タンパク質として（例えば、関連するガイドRNAなしで、または関連するガイドRNAあり、すなわちリボ核タンパク質複合体として）直接提供される。本開示のCasYポリペプチドは、任意の利便な方法によって細胞に導入する（細胞に供給する）ことができる。そのような方法は、当業者に公知である。例示的な一例として、本開示のCasYポリペプチドは、（例えば、CasYガイドRNAまたはCasYガイドRNAをコードする核酸ありまたはなしで、及びドナーポリヌクレオチドありまたはなしで）細胞内に直接注入することができる。別の例として、本開示のCasYポリペプチドとCasYガイドRNA（RNP）との複合体を予め形成して、それを細胞（例えば、真核細胞）に導入することができる（例えば、注入によって、ヌクレオフェクションによって；1つ以上の構成要素にコンジュゲートされたタンパク質導入ドメイン（PTD）、例えばCasYタンパク質にコンジュゲートされたPTD、ガイドRNAにコンジュゲートされたPTD、本開示のCasYポリペプチド及びガイドRNAにコンジュゲートされたPTDを介してなど）。

10

20

【0222】

場合によって、本開示のCasY融合ポリペプチド（例えば、融合パートナーに融合されたdCasY、融合パートナーに融合されたニックーゼCasYなど）は、CasY融合ポリペプチドをコードする核酸（例えば、mRNA、DNA、プラスミド、発現ベクター、ウイルスベクターなど）として提供される。場合によって、本開示のCasY融合ポリペプチドは、タンパク質として（例えば、関連するガイドRNAなしで、または関連するガイドRNAあり、すなわちリボ核タンパク質複合体として）直接提供される。本開示のCasY融合ポリペプチドは、任意の利便な方法によって細胞に導入する（細胞に供給する）ことができる。そのような方法は、当業者に公知である。例示的な一例として、本開示のCasY融合ポリペプチドは、（例えば、CasYガイドRNAをコードする核酸ありまたはなしで、及びドナーポリヌクレオチドありまたはなしで）細胞内に直接注入することができる。別の例として、本開示のCasY融合ポリペプチドとCasYガイドRNA（RNP）との複合体を予め形成して、それを細胞に導入することができる（例えば、注入によって、ヌクレオフェクションによって；1つ以上の構成要素にコンジュゲートされたタンパク質導入ドメイン（PTD）、例えばCasY融合タンパク質にコンジュゲートされたPTD、ガイドRNAにコンジュゲートされたPTD、本開示のCasY融合ポリペプチド及びガイドRNAにコンジュゲートされたPTDを介してなど）。

30

40

【0223】

場合によって、核酸（例えば、CasYガイドRNA；本開示のCasYポリペプチドをコードするヌクレオチド配列を含む核酸など）は、粒子内の、または粒子と会合する細胞（例えば、標的宿主細胞）及び/またはポリペプチド（例えば、CasYポリペプチド；CasY融合ポリペプチド）に送達される。場合によって、本開示のCasYシステムは、粒子内の、または粒子と会合する細胞に送達される。用語「粒子」及び「ナノ粒子」は、必要に応じて同義に使用することができる。本開示のCasYポリペプチド及び/またはCasYガイドRNAをコードするヌクレオチド配列、本開示のCasYポリペプチドをコードするヌクレオチド配列を含むmRNA、及びガイドRNAを含む組換え発現ベクターは、粒子または脂質エンベロープを使用して、例えば、CasYポリペプチド及び

50

C a s Y ガイド R N A を、例えば複合体（例えば、リボ核タンパク質（R N P）複合体）として同時に送達してもよく、粒子、例えば脂質またはリポイドと親水性ポリマー、例えばカチオン性脂質と親水性ポリマーを含む送達粒子を介して送達することができる。例えば、その場合のカチオン性脂質は、1, 2 - ジオレオイル - 3 - トリメチルアンモニウム - プロパン（D O T A P）または1, 2 - ジテトラデカノイル - s n - グリセロ - 3 - ホスホコリン（D M P C）を含み、及び/または親水性ポリマーは、エチレングリコールまたはポリエチレングリコール（P E G）を含み、及び/または粒子はさらにコレステロールを含む（例えば、製剤1 = D O T A P 100、D M P C 0、P E G 0、コレステロール0；製剤番号2 = D O T A P 90、D M P C 0、P E G 10、コレステロール0；製剤番号3 = D O T A P 90、D M P C 0、P E G 5、コレステロール5から得られる粒子）。例えば、複数工程の処理を用いて粒子を形成することができ、その工程では、C a s Y ポリペプチド及びC a s Y ガイド R N A を、例えばモル比1 : 1で、例えば室温で、例えば30分間、例えばヌクレアーゼを含まない滅菌1倍リン酸緩衝生理食塩水（P B S）中で一緒に混合し、かつ、製剤に応じて、D O T A P、D M P C、P E G、及びコレステロールを別々にアルコール（例えば、100%エタノール）に溶解し、この2つの溶液を一緒に混合して、複合体を含有する粒子を形成する。

【0224】

本開示のC a s Y ポリペプチド（または本開示のC a s Y ポリペプチドをコードするヌクレオチド配列を含むm R N A；または本開示のC a s Y ポリペプチドをコードするヌクレオチド配列を含む組換え発現ベクター）及び/またはC a s Y ガイド R N A（またはC a s Y ガイド R N A をコードする1つの以上の発現ベクターなどの核酸）は、粒子または脂質エンベロープを使用して同時に送達することができる。例えば、リン脂質二重層シェルに封入されたポリ（ α -アミノエステル）（P B A E）コアをもつ、生分解性のコア-シェル構造のナノ粒子を使用することができる。場合によって、自己組織化生体接着性ポリマーを用いた粒子/ナノ粒子が使用される。そのような粒子/ナノ粒子を、例えば脳への、ペプチドの経口送達、ペプチドの静脈内送達、及びペプチドの経鼻送達に利用してもよい。疎水性薬物の経口吸収及び眼内送達のような他の実施形態もまた企図される。疾患の部位を保護するように、及び疾患の部位に送達するように設計されたポリマーエンベロープに関連する分子エンベロープ技術を使用することができる。様々な因子、例えば標的組織に応じて、約5mg/kg用量を単回投与または複数回投与で用いることができる。

【0225】

リポイド化合物（例えば、米国特許出願第20110293703号に記載されているようなもの）もポリヌクレオチドの投与に有用であり、本開示のC a s Y ポリペプチド、本開示のC a s Y 融合ポリペプチド、本開示のR N P、本開示の核酸、または本開示のC a s Y システムの送達に使用することができる（例えば、C a s Y システムは以下を含む：a）本開示のC a s Y ポリペプチド及びC a s Y ガイド R N A；b）本開示のC a s Y ポリペプチド、C a s Y ガイド R N A、及びドナー鑄型核酸；c）本開示のC a s Y 融合ポリペプチド及びC a s Y ガイド R N A；d）本開示のC a s Y 融合ポリペプチド、C a s Y ガイド R N A、及びドナー鑄型核酸；e）本開示のC a s Y ポリペプチドをコードするm R N A 及びC a s Y ガイド R N A；f）本開示のC a s Y ポリペプチドをコードするm R N A、C a s Y ガイド R N A、及びドナー鑄型核酸；g）本開示のC a s Y 融合ポリペプチドをコードするm R N A 及びC a s Y ガイド R N A；h）本開示のC a s Y 融合ポリペプチドをコードするm R N A、C a s Y ガイド R N A、及びドナー鑄型核酸；i）本開示のC a s Y ポリペプチドをコードするヌクレオチド配列及びC a s Y ガイド R N A をコードするヌクレオチド配列を含む組換え発現ベクター；j）本開示のC a s Y ポリペプチドをコードするヌクレオチド配列、C a s Y ガイド R N A をコードするヌクレオチド配列、及びドナー鑄型核酸をコードするヌクレオチド配列を含む組換え発現ベクター；k）本開示のC a s Y 融合ポリペプチドをコードするヌクレオチド配列及びC a s Y ガイド R N A をコードするヌクレオチド配列を含む組換え発現ベクター；l）本開示のC a s Y 融合ポリペプチドをコードするヌクレオチド配列、C a s Y ガイド R N A をコードするヌ

クレオチド配列、及びドナー鑄型核酸をコードするヌクレオチド配列を含む組換え発現ベクター； m) 本開示の CasY ポリペプチドをコードするヌクレオチド配列を含む第 1 の組換え発現ベクター、及び CasY ガイド RNA をコードするヌクレオチド配列を含む第 2 の組換え発現ベクター； n) 本開示の CasY ポリペプチドをコードするヌクレオチド配列を含む第 1 の組換え発現ベクター、ならびに CasY ガイド RNA をコードするヌクレオチド配列及びドナー鑄型核酸を含む第 2 の組換え発現ベクター； o) 本開示の CasY 融合ポリペプチドをコードするヌクレオチド配列を含む第 1 の組換え発現ベクター、及び CasY ガイド RNA をコードするヌクレオチド配列を含む第 2 の組換え発現ベクター； p) 本開示の CasY 融合ポリペプチドをコードするヌクレオチド配列を含む第 1 の組換え発現ベクター、ならびに CasY ガイド RNA をコードするヌクレオチド配列及びドナー鑄型核酸を含む第 2 の組換え発現ベクター； q) 本開示の CasY ポリペプチドをコードするヌクレオチド配列、第 1 の CasY ガイド RNA をコードするヌクレオチド配列、及び第 2 の CasY ガイド RNA をコードするヌクレオチド配列を含む組換え発現ベクター； もしくは r) 本開示の CasY 融合ポリペプチドをコードするヌクレオチド配列、第 1 の CasY ガイド RNA をコードするヌクレオチド配列、及び第 2 の CasY ガイド RNA をコードするヌクレオチド配列を含む組換え発現ベクター； または (a) ~ (r) のいずれか 1 つの何らかの変形例)。一態様では、アミノアルコールリピド化合物は、細胞または対象に送達される薬剤と組み合わせられ、微粒子、ナノ粒子、リポソーム、またはミセルを形成する。アミノアルコールリピド化合物を、他のアミノアルコールリピド化合物、ポリマー（合成または天然）、界面活性剤、コレステロール、炭水化物、タンパク質、脂質などと組み合わせ、粒子を形成することができる。その後、このような粒子を必要に応じて医薬賦形剤と組み合わせて、医薬組成物を形成することができる。

【0226】

ポリ（ベータ - アミノアルコール）（PBA A）を、本開示の CasY ポリペプチド、本開示の CasY 融合ポリペプチド、本開示の RNP、本開示の核酸、または本開示の CasY システムの標的細胞への送達に使用することができる。米国特許公開番号第 20130302401 号は、コンビナトリアル重合を用いて調製されたポリ（ベータ - アミノアルコール）（PBA A）の 1 クラスに関する。

【0227】

糖ベースの粒子、例えば、WO2014118272（参照により本明細書に組み込まれる）及び Nair, J K et al., 2014, Journal of the American Chemical Society 136(49), 16958 - 16961 に関連して記載される、GalNAc を使用してもよく、本開示の CasY ポリペプチド、本開示の CasY 融合ポリペプチド、本開示の RNP、本開示の核酸、または本開示の CasY システムの標的細胞への送達に、これを使用することができる。

【0228】

場合によって、脂質ナノ粒子（LNP）を、本開示の CasY ポリペプチド、本開示の CasY 融合ポリペプチド、本開示の RNP、本開示の核酸、または本開示の CasY システムの標的細胞への送達に使用することができる。RNA などの負電荷ポリマーは、イオン性脂質が正電荷を示すような低 pH 値（例えば、pH 4）で LNP に充填することができる。しかしながら、生理的 pH 値では、LNP は、循環時間が長いと共溶性となる低表面電荷を示す。4 種のイオン性カチオン性脂質、すなわち、1, 2 - ジリノレイル - 3 - ジメチルアンモニウム - プロパン（DLinDAP）、1, 2 - ジリノレイルオキシ - 3 - N, N - ジメチルアミノプロパン（DLinDMA）、1, 2 - ジリノレイルオキシ - ケト - N, N - ジメチル - 3 - アミノプロパン（DLinKDMA）、及び 1, 2 - ジリノレイル - 4 - (2 - ジメチルアミノエチル) - [1, 3] - ジオキソラン（DLinKC2 - DMA）が注目されている。LNP の調製については、Rosin et al. (2011) Molecular Therapy 19: 1286 - 2200 に記載されている。カチオン性脂質、1, 2 - ジリノレイル - 3 - ジメチルアンモニウム - プロパン（DLinDAP）、1, 2 - ジリノレイルオキシ - 3 - N, N - ジメチルアミノ

プロパン (DLinK - DMA)、1, 2 - ジリノレイル - 4 - (2 - ジメチルアミノエチル) - [1, 3] - ジオキソラン (DLinKC2 - DMA)、(3 - o - [2" - (メトキシポリエチレングリコール2000)スクシノイル] - 1, 2 - ジミリストイル - sn - グリコール (PEG - S - DMG)、及び R - 3 - [(オメガ - メトキシ - ポリ(エチレングリコール)2000)カルバモイル] - 1, 2 - ジミリストイルオキシプロピル - 3 - アミン (PEG - C - DOMG) を使用してもよい。核酸 (例えば、CasY ガイド RNA; 本開示の核酸など) を DLinDAP、DLinDMA、DLinK - DMA、及び DLinKC2 - DMA を含有する LNP (モル比 40 : 10 : 40 : 10 のカチオン性脂質 : DSPC : CHOL : PEG - DMG または PEG - C - DOMG) に封入してもよい。場合によって、0.2% の SP - DiOC18 が組み込まれる。

10

【0229】

Spherical Nucleic Acid (SNA (商標)) 構築物及び他のナノ粒子 (特に金ナノ粒子) を、本開示の CasY ポリペプチド、本開示の CasY 融合ポリペプチド、本開示の RNP、本開示の核酸、または本開示の CasY システムの標的細胞への送達に使用することができる。例えば、J. Am. Chem. Soc. 2011 133 : 9254 - 9257、Hao et al., Small. 2011 7 : 3158 - 3162、Zhang et al., ACS Nano. 2011 5 : 6962 - 6970、Cutler et al., J. Am. Chem. Soc. 2012 134 : 1376 - 1391、Young et al., Nano Lett. 2012 12 : 3867 - 71、Zheng et al., Proc. Natl. Acad. Sci. USA. 2012 109 : 11975 - 80、Mirkin, Nanomedicine 2012 7 : 635 - 638、Zhang et al., J. Am. Chem. Soc. 2012 134 : 16488 - 1691、Weintraub, Nature 2013 495 : S14 - S16、Choi et al., Proc. Natl. Acad. Sci. USA. 2013 110 (19) : 7625 - 7630、Jensen et al., Sci. Transl. Med. 5, 209ra152 (2013)、及び Mirkin, et al., Small, 10 : 186 - 192 を参照のこと。

20

【0230】

RNA を含む自己組織化ナノ粒子は、ポリエチレングリコール (PEG) の遠位端に結合された Arg - Gly - Asp (RGD) ペプチドリガンドで PEG 化されたポリエチレンイミン (PEI) を用いて構築することができる。

30

【0231】

一般に、「ナノ粒子」とは、1000 nm 未満の直径を有する任意の粒子を指す。場合によって、本開示の CasY ポリペプチド、本開示の CasY 融合ポリペプチド、本開示の RNP、本開示の核酸、または本開示の CasY システムの標的細胞への送達に使用するのに適したナノ粒子は、500 nm 以下、例えば、25 nm ~ 35 nm、35 nm ~ 50 nm、50 nm ~ 75 nm、75 nm ~ 100 nm、100 nm ~ 150 nm、150 nm ~ 200 nm、200 nm ~ 300 nm、300 nm ~ 400 nm、または 400 nm ~ 500 nm の直径を有する。場合によって、本開示の CasY ポリペプチド、本開示の CasY 融合ポリペプチド、本開示の RNP、本開示の核酸、または本開示の CasY システムの標的細胞への送達に使用するのに適したナノ粒子は、25 nm ~ 200 nm の直径を有する。場合によって、本開示の CasY ポリペプチド、本開示の CasY 融合ポリペプチド、本開示の RNP、本開示の核酸、または本開示の CasY システムの標的細胞への送達に使用するのに適したナノ粒子は、100 nm 以下の直径を有する。場合によって、本開示の CasY ポリペプチド、本開示の CasY 融合ポリペプチド、本開示の RNP、本開示の核酸、または本開示の CasY システムの標的細胞への送達に使用するのに適したナノ粒子は、35 nm ~ 60 nm の直径を有する。

40

【0232】

本開示の CasY ポリペプチド、本開示の CasY 融合ポリペプチド、本開示の RNP

50

、本開示の核酸、または本開示の CasY システムの標的細胞への送達に使用するのに適したナノ粒子は、例えば、固体ナノ粒子（例えば、銀、金、鉄、チタンなどの金属、非金属、脂質ベースの固体、ポリマー）、ナノ粒子の懸濁液、またはそれらの組み合わせのような異なる形態で提供することができる。ハイブリッド構造（例えば、コア-シェルナノ粒子）とならび、金属、誘電体、及び半導体ナノ粒子を作製してもよい。また、電子エネルギー準位の量子化が起きるほどナノ粒子が小さい（通常は 10 nm より小さい）場合、半導体材料で作製されたナノ粒子を標識化量子ドットにすることができる。そのようなナノスケール粒子は、薬物担体または造影剤といった生物医学用途に使用され、本開示における同様の目的に応用することができる。

【0233】

半固体及びソフトナノ粒子もまた、本開示の CasY ポリペプチド、本開示の CasY 融合ポリペプチド、本開示の RNP、本開示の核酸、または本開示の CasY システムの標的細胞への送達に使用するのに適している。典型的な半固体性質のナノ粒子がリポソームである。

【0234】

場合によって、本開示の CasY ポリペプチド、本開示の CasY 融合ポリペプチド、本開示の RNP、本開示の核酸、または本開示の CasY システムの標的細胞への送達に、エクソソームが使用される。エクソソームは、RNA 及びタンパク質を運搬して、脳及び他の標的器官に RNA を送達することができる内因性ナノ小胞である。

【0235】

場合によって、本開示の CasY ポリペプチド、本開示の CasY 融合ポリペプチド、本開示の RNP、本開示の核酸、または本開示の CasY システムの標的細胞への送達に、リポソームが使用される。リポソームは、内部の水性区画を取り巻く単層または多層脂質二重層と、比較的透過性である外側の親油性リン脂質二重層で構成される球状小胞構造体である。リポソームは、いくつかの異なる種類の脂質から作製できるが、リン脂質がリポソームの作製に最もよく使用される。脂質膜を水溶液と混合すると、リポソーム形成が自発的に起こるが、ホモジナイザー、ソニケーター、または押出成形装置を使用して振盪という形で力を加えることによりリポソーム形成を促進することもできる。いくつかの他の添加剤をリポソームに添加して、その構造及び特性を改質することができる。例えば、コレステロールまたはスフィンゴミエリンのいずれかをリポソーム混合物に添加すると、リポソーム構造の安定化、及びリポソーム内の積荷の漏出防止に役立てることができる。リポソーム製剤は主に、1, 2 - ジステアロイル - sn - グリセロ - 3 - ホスファチジルコリン (DSPC)、スフィンゴミエリン、卵ホスファチジルコリン、及びモノシアロガングリオシドのような天然のリン脂質及び脂質で構成することができる。

【0236】

安定な核酸-脂質粒子 (SNALP) を、本開示の CasY ポリペプチド、本開示の CasY 融合ポリペプチド、本開示の RNP、本開示の核酸、または本開示の CasY システムの標的細胞への送達に使用することができる。SNALP 製剤は、脂質 3 - N - [(メトキシポリ(エチレングリコール)2000)カルバモイル] - 1, 2 - ジミリスチルオキシ - プロピルアミン (PEG - C - DMA)、1, 2 - ジリノレイルオキシ - N, N - ジメチル - 3 - アミノプロパン (DLiNDMA)、1, 2 - ジステアロイル - sn - グリセロ - 3 - ホスホコリン (DSPC)、及びコレステロールを、モル比 2 : 40 : 10 : 48 で含有し得る。SNALP リポソームは、25 : 1 の脂質 / siRNA 比、及び 48 / 40 / 10 / 2 のコレステロール / D - Lin - DMA / DSPC / PEG - C - DMA モル比を使用して、D - Lin - DMA 及び PEG - C - DMA を、ジステアロイルホスファチジルコリン (DSPC)、コレステロール、及び siRNA とともに製剤化することによって調製することができる。得られる SNALP リポソームは、大きさが約 80 ~ 100 nm であり得る。SNALP は、合成コレステロール (Sigma - Aldrich, St Louis, Mo., USA)、ジパルミトイルホスファチジルコリン (Avanti Polar Lipids, Alabaster, Ala., USA)

10

20

30

40

50

、 3 - N - [(w - メトキシポリ (エチレングリコール) 2 0 0 0) カルバモイル] - 1 , 2 - ジミリスチルオキシプロピルアミン、及びカチオン性 1 , 2 - ジリノレイルオキシ - 3 - N , Nジメチルアミノプロパンを含んでもよい。S N A L P は、合成コレステロール (S i g m a - A l d r i c h) 、 1 , 2 - ジステアロイル - s n - グリセロ - 3 - ホスホコリン (D S P C ; A v a n t i P o l a r L i p i d s I n c .) 、 P E G - c D M A 、 及び 1 , 2 - ジリノレイルオキシ - 3 - (N ; N - ジメチル) アミノプロパン (D L i n D M A) を含んでもよい。

【 0 2 3 7 】

アミノ脂質 2 , 2 - ジリノレイル - 4 - ジメチルアミノエチル - [1 , 3] - ジオキソラン (D L i n - K C 2 - D M A) のような他のカチオン性脂質を、本開示の C a s Y ポリペプチド、本開示の C a s Y 融合ポリペプチド、本開示の R N P 、本開示の核酸、または本開示の C a s Y システムの標的細胞への送達に使用することができる。以下の脂質組成物：それぞれ 4 0 / 1 0 / 4 0 / 1 0 モル比のアミノ脂質、ジステアロイルホスファチジルコリン (D S P C) 、コレステロール、及び (R) - 2 , 3 - ビス (オクタデシルオキシ) プロピル - 1 - (メトキシポリ (エチレングリコール) 2 0 0 0) プロピルカルバメート (P E G 脂質) 、 ならびに総脂質比が約 0 . 0 5 (w / w) の F V I I s i R N A を用いた予備成形小胞を企図することができる。7 0 ~ 9 0 n m の範囲の狭い粒度分布及び 0 . 1 1 ± 0 . 0 4 (n = 5 6) の低い多分散指数を確保するために、ガイド RNA を添加する前に、最大 3 回、8 0 n m 膜を通して押し出してもよい。極めて強力なアミノ脂質 1 6 を含有する粒子を使用してもよく、その場合、1 6 、 D S P C 、コレステロール、及び P E G 脂質という 4 つの脂質成分のモル比 (5 0 / 1 0 / 3 8 . 5 / 1 . 5) をさらに最適化して、i n v i v o 活性を増強することができる。

【 0 2 3 8 】

本開示の C a s Y システムまたはその構成要素 (複数可) またはそれをコードする核酸と脂質を製剤化して、脂質ナノ粒子 (L N P) を形成してもよい。好適な脂質として、限定はされないが、D L i n - K C 2 - D M A 4 、 C 1 2 - 2 0 0 及び補脂質ジステロイルホスファチジルコリン、コレステロール、ならびに P E G - D M G が挙げられ、自発的な小胞形成手順を用いて本開示の C a s Y システムまたはその構成要素と製剤化することができる。成分のモル比は、約 5 0 / 1 0 / 3 8 . 5 / 1 . 5 (D L i n - K C 2 - D M A または C 1 2 - 2 0 0 / ジステロイルホスファチジルコリン / コレステロール / P E G - D M G) であり得る。

【 0 2 3 9 】

本開示の C a s Y システム、またはその構成要素は、米国公開出願第 2 0 1 3 0 2 5 2 2 8 1 号及び第 2 0 1 3 0 2 4 5 1 0 7 号及び第 2 0 1 3 0 2 4 4 2 7 9 号にさらに記載されるような P L G A マイクロスフィアに封入して送達することができる。

【 0 2 4 0 】

超荷電タンパク質を、本開示の C a s Y ポリペプチド、本開示の C a s Y 融合ポリペプチド、本開示の R N P 、本開示の核酸、または本開示の C a s Y システムの標的細胞への送達に使用することができる。超荷電タンパク質とは、異常に高い正または負の正味理論電荷を有する、遺伝子操作されたタンパク質または天然タンパク質のクラスである。超負荷電のタンパク質と超正荷電のタンパク質はいずれも、熱的または化学的に誘発される凝集に耐える能力を示す。超正荷電のタンパク質はまた、哺乳動物細胞を透過することができる。このようなタンパク質、例えばプラスミド D N A 、 R N A 、または他のタンパク質と積荷を会合することで、i n v i t r o と i n v i v o との両方で、このような巨大分子の哺乳動物細胞への機能的送達を可能にすることができる。

【 0 2 4 1 】

細胞透過性ペプチド (C P P) を、本開示の C a s Y ポリペプチド、本開示の C a s Y 融合ポリペプチド、本開示の R N P 、本開示の核酸、または本開示の C a s Y システムの標的細胞への送達に使用することができる。C P P は通常、リジンまたはアルギニンのような正荷電アミノ酸が、高い相対存在量で含まれているか、または極性 / 荷電アミノ酸と

非極性の疎水性アミノ酸の交互パターンを含む配列を有しているアミノ酸組成物を有する。

【0242】

本開示のCasYポリペプチド、本開示のCasY融合ポリペプチド、本開示のRNP、本開示の核酸（例えば、CasYガイドRNA、CasYガイドRNAをコードする核酸、CasYポリペプチドをコードする核酸、ドナー鋳型など）、または本開示のCasYシステムを、標的細胞（例えば、in vivoの標的細胞。その場合の標的細胞とは、血中の標的細胞、組織内の標的細胞、器官内の標的細胞などである）に送達するために埋込型装置を使用することができる。本開示のCasYポリペプチド、本開示のCasY融合ポリペプチド、本開示のRNP、本開示の核酸、または本開示のCasYシステムの標的細胞（例えば、in vivoの標的細胞。その場合の標的細胞とは、血中の標的細胞、組織内の標的細胞、器官内の標的細胞などである）への送達に使用するのに適した埋込型装置には、CasYポリペプチド、CasY融合ポリペプチド、RNP、またはCasYシステム（またはその構成要素、例えば本開示の核酸）を含む容器（例えば、リザーバー、マトリックスなど）を含み得る。

10

【0243】

好適な埋込型装置には、例えば装置本体として使用される、マトリックスなどのポリマー基材、及び場合によって金属または追加のポリマーなどの追加の足場材料、及び可視化能及びイメージングを向上させる材料を含み得る。埋込型の送達装置は、局所的かつ長期間にわたる放出をもたらす際に有利となる可能性があり、送達されるポリペプチド及び/または核酸が、標的部位、例えば細胞外マトリックス（ECM）、腫瘍周囲の血管系、罹患組織などに直接放出される。好適な埋込型の送達装置として、薬物送達系が固定されないまたは取り付けられない、腹腔などの腔への送達及び/または他の任意の種類の種類に使用する際に適しており、生体安定性及び/または分解性及び/または生体吸収性のポリマー基材（例えば必要に応じてマトリックスであってよい）を含む装置が挙げられる。場合によって、好適な埋込型の薬物送達装置は、分解性ポリマーを含み、主要な放出メカニズムがバルク崩壊である。場合によって、好適な埋込型の薬物送達装置は、非分解性または遅分解性ポリマーを含み、主要な放出メカニズムがバルク崩壊ではなく拡散であり、それによって、外側部分は膜として機能し、かつ内側部分は薬物リザーバーとして機能し、実質的に長期間（例えば、約1週間～約数か月間）周囲の影響を受けない。放出メカニズムが異なる別のポリマーの組み合わせも必要に応じて使用することができる。全放出期間のかなりの期間にわたって濃度勾配を事実上、一定に維持することができるため、拡散速度が事実上、一定である（「ゼロモード」拡散と称する）。「一定」という用語は、治療有効性の下限を超えて維持されるが、さらに必要に応じて初期バーストを特徴とする場合がある拡散速度、及び/または変動する、例えばある程度まで増減する場合がある拡散速度を意味する。拡散速度は長期間にわたってそのように維持することができ、治療有効期間、例えば有効サイレンシング期間を最適化するような、ある水準までを一定とみなすことができる。

20

30

【0244】

場合によって、埋込型の送達系は、ヌクレオチドを用いた治療薬を分解から保護するように設計される。分解は、本質的に化学的であるか、または対象の体内の酵素及び他の因子からの攻撃に起因するかを問わない。

40

【0245】

装置の埋込部位、すなわち標的部位は、最大の治療効果を得るように選択することができる。例えば、送達装置は、腫瘍環境内もしくは腫瘍環境付近、または腫瘍に関連する血液供給源に埋め込むことができる。標的位置は、例えば、1)パーキンソン病またはアルツハイマー病での大脳基底核、白質及び灰白質のような変性部位の脳；2)（筋萎縮性側索硬化症（ALS）の場合のような）脊椎；3)子宮頸部；4)活動性及び慢性炎症性関節；5)（乾癬の場合のような）真皮；7)交感神経及び感覚神経部位（鎮痛効果のため）；7)骨；8)急性または慢性感染部位；9)腔内；10)内耳-聴覚系、内耳迷路、

50

前庭系；11) 気管内；12) 心臓内；冠動脈、心外膜；13) 尿路または膀胱；14) 胆管系；15) 腎臓、肝臓、脾臓を含むが、これに限定されない実質組織；16) リンパ節；17) 唾液腺；18) 歯肉；19) 関節内（関節内部）；20) 眼内；21) 脳組織；22) 脳室；23) 腹腔を含む腔（例えば限定されないが、卵巣癌の場合）；24) 食道内；及び25) 直腸内；ならびに26) 血管系内部であり得る。

【0246】

他の種類の組織埋込及び/または挿入及び/または組織のサンプリングのために、埋込などの挿入方法を必要に応じて予め使用することができる。場合により、そのような方法に変更を加えずに、または必要に応じて重要でない変更のみを加えて代替的に使用してもよい。そのような方法には場合により、小線源治療法、生検、超音波を用いる及び/または

10

【0247】

改変宿主細胞

本開示は、本開示の CasY ポリペプチド、及び/または本開示の CasY ポリペプチドをコードするヌクレオチド配列を含む核酸を含む改変細胞を提供する。本開示は、本開示の CasY ポリペプチドを含む改変細胞であり、通常は本開示の CasY ポリペプチドを含んでいない細胞である改変細胞を提供する。本開示は、本開示の CasY ポリペプチドをコードするヌクレオチド配列を含む核酸を含む改変細胞（例えば遺伝子改変細胞）を提供する。本開示は、本開示の CasY ポリペプチドをコードするヌクレオチド配列を含む mRNA で遺伝子改変された遺伝子改変細胞を提供する。本開示は、本開示の CasY ポリペプチドをコードするヌクレオチド配列を含む組換え発現ベクターで遺伝子改変された遺伝子改変細胞を提供する。本開示は、a) 本開示の CasY ポリペプチドをコードするヌクレオチド配列；及び b) 本開示の CasY ガイド RNA をコードするヌクレオチド配列を含む組換え発現ベクターで遺伝子改変された遺伝子改変細胞を提供する。本開示は、a) 本開示の CasY ポリペプチドをコードするヌクレオチド配列；b) 本開示の CasY ガイド RNA をコードするヌクレオチド配列；及び c) ドナー鋳型をコードするヌクレオチド配列を含む組換え発現ベクターで遺伝子改変された遺伝子改変細胞を提供する。

20

【0248】

本開示の CasY ポリペプチド、及び/または本開示の CasY ポリペプチドをコードするヌクレオチド配列を含む核酸、及び/または本開示の CasY ガイド RNA のレシピエントの役割を果たす細胞は、例えば、in vitro の細胞；in vivo の細胞；ex vivo の細胞；初代細胞；がん細胞；動物細胞；植物細胞；藻類細胞；真菌細胞などを含む、種々の細胞のいずれかであり得る。本開示の CasY ポリペプチド、及び/または本開示の CasY ポリペプチドをコードするヌクレオチド配列を含む核酸、及び/または本開示の CasY ガイド RNA のレシピエントの役割を果たす細胞を、「宿主細胞」または「標的細胞」と称する。宿主細胞または標的細胞は、本開示の CasY システムのレシピエントであってよい。宿主細胞または標的細胞は、本開示の CasY RNP のレシピエントであってよい。宿主細胞または標的細胞は、本開示の CasY システムの単一構成要素のレシピエントであってよい。

30

40

【0249】

細胞（標的細胞）の非限定的な例として、原核細胞、真核細胞、細菌細胞、古細菌細胞、単細胞真核生物の細胞、原生動物細胞、植物由来細胞（例えば、作物（植物）、果実、野菜、穀物、大豆、トウモロコシ（corn、maize）、小麦、種子、トマト、イネ、キャッサバ、サトウキビ、カボチャ、干し草、ジャガイモ、綿、大麻、タバコ、顕花植物、針葉樹、裸子植物、被子植物、シダ類、ヒカゲノカズラ類、ツノゴケ類、ゼニゴケ類、セン類、双子葉植物、単子葉植物など由来の細胞）、藻類細胞（例えば、Botryococcus braunii、Chlamydomonas reinhardtii、Nannochloropsis gaditana、Chlorella pyrenoidosa、Sargassum patens、C. agardh など）、海藻（

50

例えばケルブ)、真菌細胞(例えば、酵母細胞、キノコ由来の細胞)、動物細胞、無脊椎動物(例えば、ショウジョウバエ、刺胞動物、棘皮動物、線虫など)由来の細胞、脊椎動物(例えば、魚類、両生類、爬虫類、鳥類、哺乳類)由来の細胞、哺乳動物(例えば、有蹄類(例えば、ブタ、ウシ、ヤギ、ヒツジ);齧歯類(例えば、ラット、マウス);非ヒト霊長類;ヒト;ネコ科動物(例えば、ネコ);イヌ科動物(例えば、イヌ)など)由来の細胞などが挙げられる。場合によって、細胞は、天然の生物に由来しない細胞である(例えば、細胞は、合成的に作製された細胞であり得る:これは人工細胞とも呼ばれる)。

【0250】

細胞は*in vitro*の細胞(例えば、樹立培養細胞株)であってよい。細胞は*ex vivo*の細胞(個体由来の培養細胞)であってよい。細胞は*in vivo*の細胞(例えば、個体内の細胞)であってよい。細胞は単離細胞であってよい。細胞は生物の体内細胞であってよい。細胞は生物であってよい。細胞は細胞培養物(例えば、*in vitro*の細胞培養物)中の細胞であってよい。細胞は細胞の集合のうちの一つであってよい。細胞は、原核細胞であってよい、または原核細胞に由来するものであってもよい。細胞は、細菌細胞であってよい、または細菌細胞に由来するものであってもよい。細胞は、古細菌細胞であってよい、または古細菌細胞に由来するものであってもよい。細胞は、真核細胞であってよい、または真核細胞に由来するものであってもよい。細胞は、植物細胞であってよい、または植物細胞に由来するものであってもよい。細胞は、動物細胞であってよい、または動物細胞に由来するものであってもよい。細胞は、無脊椎動物細胞であってよい、または無脊椎動物細胞に由来するものであってもよい。細胞は、脊椎動物細胞であってよい、または脊椎動物細胞に由来するものであってもよい。細胞は、哺乳動物細胞であってよい、または哺乳動物細胞に由来するものであってもよい。細胞は、齧歯類細胞であってよい、または齧歯類細胞に由来するものであってもよい。細胞は、ヒト細胞であってよい、またはヒト細胞に由来するものであってもよい。細胞は、微生物細胞であってよい、または微生物細胞に由来するものであってもよい。細胞は、真菌細胞であってよい、または真菌細胞に由来するものであってもよい。細胞は昆虫細胞であってよい。細胞は節足動物細胞であってよい。細胞は原生動物細胞であってよい。細胞は蠕虫細胞であってよい。

【0251】

好適な細胞として、幹細胞(例えば、胚性幹(ES)細胞、人工多能性幹(iPS)細胞);生殖細胞(例えば、卵母細胞、精子、卵原細胞、精原細胞など);体細胞、例えば線維芽細胞、オリゴデンドロサイト、グリア細胞、造血細胞、ニューロン、筋肉細胞、骨細胞、肝細胞、膵臓細胞などが挙げられる。

【0252】

好適な細胞として、ヒト胚性幹細胞、胎児心筋細胞、筋線維芽細胞、間葉系幹細胞、自家移植拡張心筋細胞、脂肪細胞、全能性細胞、多能性細胞、血液幹細胞、筋芽細胞、成体幹細胞、骨髄細胞、間葉系細胞、胚性幹細胞、実質細胞、上皮細胞、内皮細胞、中皮細胞、線維芽細胞、骨芽細胞、軟骨細胞、外因性細胞、内因性細胞、幹細胞、造血幹細胞、骨髄由来前駆細胞、心筋細胞、骨格細胞、胎児細胞、未分化細胞、多能性前駆細胞、単能性前駆細胞、単球、心筋芽細胞、骨格筋芽細胞、マクロファージ、毛細血管内皮細胞、異種細胞、同種細胞、及び出生後幹細胞が挙げられる。

【0253】

場合によって細胞は、免疫細胞、ニューロン、上皮細胞、及び内皮細胞、または幹細胞である。場合によって、免疫細胞は、T細胞、B細胞、単球、ナチュラルキラー細胞、樹状細胞、またはマクロファージである。場合によって、免疫細胞は細胞傷害性T細胞である。場合によって、免疫細胞はヘルパーT細胞である。場合によって、免疫細胞は制御性T細胞(Treg)である。

【0254】

場合によって、細胞は幹細胞である。幹細胞には成体幹細胞を含む。成体幹細胞は、体性幹細胞とも呼ばれる。

【0255】

10

20

30

40

50

成体幹細胞は、分化組織に常在しているが、自己再生及び複数の細胞型（通常は、幹細胞が見られる組織に特有の細胞型）の発生能といった特性を保有する。体性幹細胞の多数の例が、当業者に知られており、これには、筋幹細胞；造血幹細胞；上皮幹細胞；神経幹細胞；間葉系幹細胞；乳腺幹細胞；腸管幹細胞；中胚葉幹細胞；内皮幹細胞；嗅覚幹細胞；神経堤幹細胞などが含まれる。

【0256】

対象となる幹細胞には哺乳動物の幹細胞を含むが、その場合の用語「哺乳動物」とは、哺乳動物として分類されるいずれかの動物を指し、ヒト；非ヒト霊長類；家畜及び農場動物；ならびに動物園の動物、実験動物、競技用動物、または愛玩動物、例えば、イヌ、ウマ、ネコ、ウシ、マウス、ラット、ウサギなどを含む。場合によって、幹細胞はヒト幹細胞である。場合によって、幹細胞は齧歯類（例えばマウス；ラット）幹細胞である。場合によって、幹細胞は非ヒト霊長類の幹細胞である。

10

【0257】

幹細胞は、1つ以上の幹細胞マーカー、例えば、SOX9、KRT19、KRT7、LGR5、CA9、FXRD2、CDH6、CLDN18、TSPAN8、BPIFB1、OLFM4、CDH17、及びPPARGC1Aを発現することができる。

【0258】

いくつかの実施形態では、幹細胞は造血幹細胞（HSC）である。HSCは、骨髄、血液、臍帯血、胎児肝、及び卵黄嚢から単離することができる中胚葉由来の細胞である。HSCは、CD34⁺及びCD3⁻を特徴とする。HSCは、*in vivo*の赤血球、好中球マクロファージ、巨核球、及びリンパ系の造血細胞株を増殖することができる。*in vitro*で、少なくともいくつかの自己再生細胞分裂を受けるとしてHSCを誘導することができる、*in vivo*で見られるものと同じ系統へと分化するようにHSCを誘導することができる。したがって、HSCは赤血球系細胞、巨核球、好中球、マクロファージ、及びリンパ系細胞のうち1つ以上に分化するように誘導することができる。

20

【0259】

他の実施形態では、幹細胞は神経幹細胞（NSC）である。神経幹細胞（NSC）は、ニューロン及びグリア（オリゴデンドロサイト及びアストロサイトを含む）に分化することができる。神経幹細胞は、多様な分化が可能な多能性幹細胞であり、特定の条件下で、神経幹細胞である娘細胞、または神経芽細胞またはグリア芽細胞となり得る神経前駆細胞、例えば、それぞれ1つ以上の種類のニューロン及びグリア細胞となることが約束された細胞を産生することができる。NSCの入手方法は、当技術分野において公知である。

30

【0260】

他の実施形態では、幹細胞は間葉系幹細胞（MSC）である。MSCは本来、胚中胚葉に由来し、成体骨髄から単離され、分化して筋肉、骨、軟骨、脂肪、骨髄間質、及び腱を形成することができる。MSCの単離方法は当技術分野で公知であり、任意の公知の方法を使用してMSCを得ることができる。例えば、ヒトMSCの単離について記載している米国特許第5,736,396号を参照のこと。

【0261】

細胞は場合によって植物細胞である。植物細胞は、単子葉植物の細胞であってよい。細胞は、双子葉植物の細胞であってよい。

40

【0262】

場合によって、細胞は植物細胞である。例えば、細胞は、主要な農業植物、例えば、大麦、豆（食用乾燥）、キャノーラ、トウモロコシ、綿（ピマ）、綿（高地）、亜麻仁、干し草（アルファルファ）、干し草（アルファルファ以外）、カラスムギ、ピーナッツ、イネ、モロコシ、大豆、テンサイ、サトウキビ、ヒマワリ（油）、ヒマワリ（油以外）、サツマイモ、タバコ（パーレー種）、タバコ（熱風乾燥）、トマト、小麦（デュラム）、小麦（春）、小麦（冬）などの細胞であり得る。別の例として、細胞は、野菜作物の細胞であり、限定されないが、例えば、アルファルファモやし、アロエの葉、クズウコン、クワイ、アーティチョーク、アスパラガス、タケノコ、バナナの花、モヤシ、豆、ビート茎葉

50

部、ビート、ニガウリ、チンゲンサイ、ブロッコリー、ブロッコリーレイブ(ラピニ)、芽キャベツ、キャベツ、キャベツスプラウト、サボテンの葉(ノパレス)、ヒョウタン、カルドン、ニンジン、カリフラワー、セロリ、ハヤトウリ、チョロギ(クローヌ)、白菜、キンサイ、ニラ、サイシン、キクナ(春菊)、コラード若葉、トウモロコシの茎、スイートコーン、キュウリ、ダイコン、タンポポの葉、タロイモ、豆苗(エンドウ若葉)、トウガン(冬瓜)、ナス、エンダイブ、キクヂシャ、ゼンマイシダ、グンバイナズナ、チコリ、ガイチョイ(芥菜)、ガイラン、ガランガ(シャム、タイショウガ)、ニンニク、ショウガの根、ゴボウ、緑色野菜、ハノーバーサラダの葉、ウアウソントレ、キクイモ、ヒカマ、ケールの葉、コールラビ、シロザ(アマランサス)、レタス(ビブ)、レタス(ボストン)、レタス(ボストンレッド)、レタス(グリーンリーフ)、レタス(アイスバーグ)、レタス(ロコロッサ)、レタス(オークリーフグリーン)、レタス(オークリーフレッド)、レタス(加工)、レタス(レッドリーフ)、レタス(ロメイン)、レタス(ルビーロメイン)、レタス(ロシアンレッドマスタード)、ヒシの実、ロボク、ササゲ、レンコン、ノヂシャ、リュウゼツラン(アガーベ)の葉、タロイモ、メスキュランミックス、ミズナ、モープ(moap)(ヘチマ)、ムー、マクア(ファジースカッシュ)、マッシュルーム、マスタード、ナガイモ、オクラ、空心菜、九条ネギ、オポ(ロングスカッシュ)、観賞用トウモロコシ、観賞用ヒョウタン、パセリ、パースニップ、エンドウマメ、トウガラシ(ベルタイプ)、トウガラシ、カボチャ、エンダイブ、カイワレナ、ラディッシュ、アブラナ、アブラナ、ルパーブ、ロメイン(ベイビーレッド)、ルタバガ、アッケシソウ(シービーンズ)、ヘチオク(トカド/十角ヘチマ)、ハウレンソウ、スクオッシュ、ストローベイル、サトウキビ、サツマイモ、スイスチャード、タマリンド、タロイモ、タロイモの葉、タロイモの茎、ターサイ、テベグアヘ(グアヘ)、ティンドラ、トマテイヨ、トマト、トマト(チェリー)、トマト(グレープ種)、トマト(プラム種)、ターメリック、カブ上菜、カブ、オオクログワイ、ヤンピ、ヤムイモ(ナメス)、ユーチョイ、ユカ(キャッサバ)などを含む。

10

20

30

40

【0263】

細胞は、場合によって節足動物細胞である。例えば、細胞は、例えば、Chelicerata、Myriapodia、Hexipodia、Arachnida、Insecta、Archaeognatha、Thysanura、Palaeoptera、Ephemeroptera、Odonata、Anisoptera、Zygoptera、Neoptera、Exopterygota、Plecoptera、Embiodoptera、Orthoptera、Zoraptera、Dermaptera、Dicyptera、Notoptera、Grylloblattidae、Mantophasmatidae、Phasmatodea、Blattaria、Isoptera、Mantodea、Parapneuroptera、Psocoptera、Thysanoptera、Phthiraptera、Hemiptera、EndopterygotaもしくはHolometabola、Hymenoptera、Coleoptera、Strepsiptera、Raphidioptera、Megaloptera、Neuroptera、Mecoptera、Siphonaptera、Diptera、Trichoptera、またはLepidopteraの亜目、科、亜科、属、亜属、または種の細胞であり得る。

【0264】

細胞は、場合によって昆虫細胞である。例えば、場合によって、細胞は、カ、バッタ、半翅目の昆虫、ハエ、ノミ、ハチ、スズメバチ、アリ、シラミ、ガ、または甲虫の細胞である。

【0265】

キット

本開示は、本開示のCasYシステムまたは本開示のCasYシステムの構成要素を含むキットを提供する。

【0266】

50

本開示のキットは、a) 本開示の CasY ポリペプチド及び CasY ガイド RNA ; b) 本開示の CasY ポリペプチド、CasY ガイド RNA、及びドナー 鋳型核酸 ; c) 本開示の CasY 融合ポリペプチド及び CasY ガイド RNA ; d) 本開示の CasY 融合ポリペプチド、CasY ガイド RNA、及びドナー 鋳型核酸 ; e) 本開示の CasY ポリペプチドをコードする mRNA 及び CasY ガイド RNA ; f) 本開示の CasY ポリペプチドをコードする mRNA、CasY ガイド RNA、及びドナー 鋳型核酸 ; g) 本開示の CasY 融合ポリペプチドをコードする mRNA 及び CasY ガイド RNA ; h) 本開示の CasY 融合ポリペプチドをコードする mRNA、CasY ガイド RNA、及びドナー 鋳型核酸 ; i) 本開示の CasY ポリペプチドをコードするヌクレオチド配列及び CasY ガイド RNA をコードするヌクレオチド配列を含む組換え発現ベクター ; j) 本開示の CasY ポリペプチドをコードするヌクレオチド配列、CasY ガイド RNA をコードするヌクレオチド配列、及びドナー 鋳型核酸をコードするヌクレオチド配列を含む組換え発現ベクター ; k) 本開示の CasY 融合ポリペプチドをコードするヌクレオチド配列及び CasY ガイド RNA をコードするヌクレオチド配列を含む組換え発現ベクター ; l) 本開示の CasY 融合ポリペプチドをコードするヌクレオチド配列、CasY ガイド RNA をコードするヌクレオチド配列、及びドナー 鋳型核酸をコードするヌクレオチド配列を含む組換え発現ベクター ; m) 本開示の CasY ポリペプチドをコードするヌクレオチド配列を含む第 1 の組換え発現ベクター、及び CasY ガイド RNA をコードするヌクレオチド配列を含む第 2 の組換え発現ベクター ; n) 本開示の CasY ポリペプチドをコードするヌクレオチド配列を含む第 1 の組換え発現ベクター、ならびに CasY ガイド RNA をコードするヌクレオチド配列及びドナー 鋳型核酸を含む第 2 の組換え発現ベクター ; o) 本開示の CasY 融合ポリペプチドをコードするヌクレオチド配列を含む第 1 の組換え発現ベクター、及び CasY ガイド RNA をコードするヌクレオチド配列を含む第 2 の組換え発現ベクター ; p) 本開示の CasY 融合ポリペプチドをコードするヌクレオチド配列を含む第 1 の組換え発現ベクター、ならびに CasY ガイド RNA をコードするヌクレオチド配列及びドナー 鋳型核酸を含む第 2 の組換え発現ベクター ; q) 本開示の CasY ポリペプチドをコードするヌクレオチド配列、第 1 の CasY ガイド RNA をコードするヌクレオチド配列、及び第 2 の CasY ガイド RNA をコードするヌクレオチド配列を含む組換え発現ベクター ; もしくは r) 本開示の CasY 融合ポリペプチドをコードするヌクレオチド配列、第 1 の CasY ガイド RNA をコードするヌクレオチド配列、及び第 2 の CasY ガイド RNA をコードするヌクレオチド配列を含む組換え発現ベクター ; または (a) ~ (r) のいずれか 1 つの何らかの変形例を含み得る。

【0267】

本開示のキットは、a) 本開示の CasY システムの上記のような構成要素を含んでもよく、または本開示の CasY システム ; ならびに b) 例えば、i) 緩衝液 ; ii) プロテアーゼ阻害剤 ; iii) ヌクレアーゼ阻害剤 ; iv) 検出可能な標識を顕色または可視化するために必要な試薬 ; v) 陽性対照及び / または陰性対照の標的 DNA ; vi) 陽性対照及び / または陰性対照の CasY ガイド RNA などの 1 つ以上の追加の試薬を含んでもよい。本開示のキットは、a) 本開示の CasY システムの上記のような構成要素を含んでもよく、または本開示の CasY システム ; 及び b) 治療薬を含んでもよい。

【0268】

本開示のキットは、a) 標的核酸内の標的ヌクレオチド配列にハイブリダイズする CasY ガイド RNA の部分をコードするヌクレオチド配列を含む核酸を挿入するための挿入部位 ; 及び b) CasY ガイド RNA の CasY 結合部分をコードするヌクレオチド配列を含む、組換え発現ベクターを含んでもよい。本開示のキットは、a) 標的核酸内の標的ヌクレオチド配列にハイブリダイズする CasY ガイド RNA の部分をコードするヌクレオチド配列を含む核酸を挿入するための挿入部位 ; b) CasY ガイド RNA の CasY 結合部分をコードするヌクレオチド配列 ; 及び c) 本開示の CasY ポリペプチドをコードするヌクレオチド配列を含む、組換え発現ベクターを含んでもよい。

【0269】

有用性

本開示の CasY ポリペプチドまたは本開示の CasY 融合ポリペプチドは、(例えば、CasY ガイド RNA と組み合わせ、場合によってはさらにドナー鑄型と組み合わせ) 種々の方法に使用することができる。例えば、本開示の CasY ポリペプチドを使用して、(i) 標的核酸 (DNA または RNA; 一本鎖または二本鎖) を修飾する (例えば、切断、例えばニック; メチル化など); (ii) 標的核酸の転写を調節する; (iii) 標的核酸を標識する; (iv) (例えば、単離、標識、イメージング、追跡などを目的として) 標的核酸に結合する; (v) 標的核酸と会合するポリペプチド (例えば、ヒストン) を修飾することなどが可能である。このように、本開示は標的核酸を修飾する方法を提供する。場合によって、標的核酸を修飾するための本開示の方法は、a) 本開示の CasY ポリペプチド; 及び b) 1 つ以上 (例えば、2 つ) の CasY ガイド RNA と、標的核酸を接触させることを含む。場合によって、標的核酸を修飾するための本開示の方法は、a) 本開示の CasY ポリペプチド; b) CasY ガイド RNA; 及び c) ドナー核酸 (例えばドナー鑄型) と、標的核酸を接触させることを含む。場合によって、接触工程は *in vitro* の細胞に行われる。場合によって、接触工程は *in vivo* の細胞に行われる。場合によって、接触工程は *ex vivo* の細胞に行われる。

10

【0270】

CasY ポリペプチドを用いる方法には、(会合する CasY ガイド RNA によって、そこを標的とすることにより) 標的核酸内の特定の領域に CasY ポリペプチドを結合することを含むため、この方法は一般に、本明細書で結合する方法 (例えば、標的核酸を結合する方法) とも称する。しかしながら、結合する方法の結果が標的核酸の結合にすぎない場合もあれば、この方法が異なる最終結果をもたらし得る場合もあると理解されるべきである (例えば、この方法は、標的核酸の修飾、例えば切断/メチル化など、標的核酸からの転写の調節; 標的核酸の翻訳の調節; ゲノム編集; 標的核酸と会合するタンパク質の調節; 標的核酸の単離などをもたらし得る)。

20

【0271】

好適な方法の例としては、例えば、Jinek et al., Science. 2012 Aug 17; 337(6096): 816-21; Chylinski et al., RNA Biol. 2013 May; 10(5): 726-37; Ma et al., Biomed Res Int. 2013; 2013: 270805; Hou et al., Proc Natl Acad Sci USA. 2013 Sep 24; 110(39): 15644-9; Jinek et al., Elife. 2013; 2: e00471; Pattanayak et al., Nat Biotechnol. 2013 Sep; 31(9): 839-43; Qi et al., Cell. 2013 Feb 28; 152(5): 1173-83; Wang et al., Cell. 2013 May 9; 153(4): 910-8; Auer et al., Genome Res. 2013 Oct 31; Chen et al., Nucleic Acids Res. 2013 Nov 1; 41(20): e19; Cheng et al., Cell Res. 2013 Oct; 23(10): 1163-71; Cho et al., Genetics. 2013 Nov; 195(3): 1177-80; DiCarlo et al., Nucleic Acids Res. 2013 Apr; 41(7): 4336-43; Dickinson et al., Nat Methods. 2013 Oct; 10(10): 1028-34; Ebin et al., Sci Rep. 2013; 3: 2510; Fujii et al., Nucleic Acids Res. 2013 Nov 1; 41(20): e187; Hu et al., Cell Res. 2013 Nov; 23(11): 1322-5; Jiang et al., Nucleic Acids Res. 2013 Nov 1; 41(20): e188; Larson et al., Nat Protoc. 2013 Nov; 8(11): 2180-96; Mali et al., Nat Methods. 2013 Oct; 10(10): 957-63; Nakayama et

30

40

50

al., *Genesis*. 2013 Dec; 51(12): 835-43; Ran et al., *Nat Protoc*. 2013 Nov; 8(11): 2281-308; Ran et al., *Cell*. 2013 Sep 12; 154(6): 1380-9; Upadhyay et al., *G3 (Bethesda)*. 2013 Dec 9; 3(12): 2233-8; Walsh et al., *Proc Natl Acad Sci USA*. 2013 Sep 24; 110(39): 15514-5; Xie et al., *Mol Plant*. 2013 Oct 9; Yang et al., *Cell*. 2013 Sep 12; 154(6): 1370-9; ならびに米国特許及び出願：第8,906,616号；第8,895,308号；第8,889,418号；第8,889,356号；第8,871,445号；第8,865,406号；第8,795,965号；第8,771,945号；第8,697,359号；第20140068797号；第20140170753号；第20140179006号；第20140179770号；第20140186843号；第20140186919号；第20140186958号；第20140189896号；第20140227787号；第20140234972号；第20140242664号；第20140242699号；第20140242700号；第20140242702号；第20140248702号；第20140256046号；第20140273037号；第20140273226号；第20140273230号；第20140273231号；第20140273232号；第20140273233号；第20140273234号；第20140273235号；第20140287938号；第20140295556号；第20140295557号；第20140298547号；第20140304853号；第20140309487号；第20140310828号；第20140310830号；第20140315985号；第20140335063号；第20140335620号；第20140342456号；第20140342457号；第20140342458号；第20140349400号；第20140349405号；第20140356867号；第20140356956号；第20140356958号；第20140356959号；第20140357523号；第20140357530号；第20140364333号；及び第20140377868号を参照のこと（各文献はその全体が参照により本明細書に組み込まれる）。

【0272】

例えば、本開示は、（限定はされないが）標的核酸を切断する方法；標的核酸を編集する方法；標的核酸からの転写を調節する方法；標的核酸を単離する方法；標的核酸に結合する方法；標的核酸をイメージングする方法、標的核酸を修飾する方法などを提供する。

【0273】

本明細書で使用される場合、例えば、CasYポリペプチドまたはCasY融合ポリペプチドと、「標的核酸を接触させる」及び「標的核酸を接触させること」という用語／語句は、標的核酸を接触させるための方法すべてを包含する。例えば、CasYポリペプチドを、タンパク質、（CasYポリペプチドをコードする）RNA、または（CasYポリペプチドをコードする）DNAとして細胞に供給することもでき、一方、CasYガイドRNAをガイドRNAとしてまたはガイドRNAをコードする核酸として供給することもできる。したがって、例えば、細胞内（例えば、*in vitro*の細胞内部、*in vivo*の細胞内部、*ex vivo*の細胞内部）で方法を実行する場合、標的核酸を接触させることを含む方法は、活性状態／最終状態（例えば、CasYポリペプチドのタンパク質（複数可）形態；CasY融合ポリペプチドのタンパク質形態；場合によってガイドRNAのRNA形態）である構成要素の一部または全部を細胞に導入することを包含し、また、1つ以上の構成要素をコードする1つ以上の核酸（例えば、CasYポリペプチドまたはCasY融合ポリペプチドをコードするヌクレオチド配列（複数可）を含む核酸（複数可）、ガイドRNA（複数可）をコードするヌクレオチド配列（複数可）を含む核酸（複数可）、ドナー鋳型をコードするヌクレオチド配列を含む核酸など）を細胞に導入することを包含する。方法はまた、細胞外の*in vitro*で行うこともできるため、標

的核酸を接触させることを含む方法には、（特に明記しない限り）*in vitro*の細胞外、*in vitro*の細胞内部、*in vivo*の細胞内部、*ex vivo*の細胞内部での接触を包含する。

【0274】

場合によって、標的核酸を修飾するための本開示の方法は、標的細胞にCasY遺伝子座、例えば、CasYポリペプチドをコードするヌクレオチド配列、ならびにCasY遺伝子座を含む細胞（例えば、場合によって、天然状態（自然界で起こる状態）でCasY遺伝子座を含む細胞）に由来するCasYをコードするヌクレオチド配列を含む核酸を導入することを含み、その場合、標的細胞は通常では（天然状態では）CasY遺伝子座を含まない。しかしながら、コードされたcrRNA（複数可）に対するガイド配列をコードする1つ以上のスペーサー配列を、対象となる1つ以上の標的配列を標的とするように修飾することができる。したがって、例えば、場合によって、標的核酸を修飾するための本開示の方法は、標的細胞にCasY遺伝子座、例えば、起源細胞（例えば、場合によって、天然状態（自然界で起こる状態）でCasY遺伝子座を含む細胞）から得られる核酸を導入することを含み、その場合の核酸は、100ヌクレオチド（nt）～5kbの長さ（例えば、100nt～500nt、500nt～1kb、1kb～1.5kb、1.5kb～2kb、2kb～2.5kb、2.5kb～3kb、3kb～3.5kb、3.5kb～4kb、または4kb～5kbの長さ）を有し、CasYポリペプチドをコードするヌクレオチド配列を含む。上記のように、いくつかのそのような場合に、コードされたcrRNA（複数可）に対するガイド配列をコードする1つ以上のスペーサー配列を、対象となる1つ以上の標的配列を標的とするように修飾することができる。場合によって、この方法は、標的細胞に：i）CasY遺伝子座；及びii）ドナーDNA鋳型を導入することを含む。場合によって、標的核酸は、*in vitro*の無細胞組成物中にある。場合によって、標的核酸は、標的細胞中に存在する。場合によって、標的核酸は、原核細胞である標的細胞中に存在する。場合によって、標的核酸は、真核細胞である標的細胞中に存在する。場合によって、標的核酸は、哺乳動物細胞である標的細胞中に存在する。場合によって、標的核酸は、植物細胞である標的細胞中に存在する。

【0275】

場合によって、標的核酸を修飾するための本開示の方法は、本開示のCasYポリペプチドまたは本開示のCasY融合ポリペプチドと、標的核酸を接触させることを含む。場合によって、標的核酸を修飾するための本開示の方法は、CasYポリペプチド及びCasYガイドRNAと、標的核酸を接触させることを含む。場合によって、標的核酸を修飾するための本開示の方法は、CasYポリペプチド、第1のCasYガイドRNA、及び第2のCasYガイドRNAと、標的核酸を接触させることを含む。場合によって、標的核酸を修飾するための本開示の方法は、本開示のCasYポリペプチド、ならびにCasYガイドRNA及びドナーDNA鋳型と、標的核酸を接触させることを含む。

【0276】

対象となる標的核酸及び標的細胞

本開示のCasYポリペプチドまたは本開示のCasY融合ポリペプチドは、CasYガイドRNAに結合されている場合、標的核酸と結合することができ、また場合によっては標的核酸と結合してそれを修飾することができる。標的核酸は、任意の核酸（例えば、DNA、RNA）であってよく、二本鎖または一本鎖であってよく、どの種類（例えば、染色体（ゲノムDNA）、染色体由来、染色体DNA、プラスミド、ウイルス、細胞外、細胞内、ミトコンドリア、葉緑体、直鎖、環状など）の核酸であってよく、（例えば、CasYガイドRNAが、標的核酸の標的配列にハイブリダイズするヌクレオチド配列を含み、それによって標的核酸を標的化できる限り）どの生物由来のものであってもよい。

【0277】

標的核酸はDNAであってもRNAであってもよい。標的核酸は、二本鎖（例えば、d

s DNA、ds RNA)であっても、一本鎖(例えば、ss RNA、ss DNA)であってもよい。場合によって、標的核酸は一本鎖である。場合によって、標的核酸は一本鎖RNA(ss RNA)である。場合によって、標的ss RNA(例えば、標的細胞のss RNA、ウイルスss RNAなど)は、mRNA、rRNA、tRNA、ノンコーディングRNA(ncRNA)、長鎖ノンコーディングRNA(lncRNA)、及びマイクロRNA(miRNA)から選択される。場合によって、標的核酸は一本鎖DNA(ss DNA)(例えば、ウイルスDNA)である。上述するように、場合によって、標的核酸は一本鎖である。

【0278】

標的核酸は、場所を問わず、例えば、in vitroの細胞外、in vitroの細胞内部、in vivoの細胞内部、ex vivoの細胞内部に局在するものであってよい。好適な標的細胞(例えば、ゲノムDNAなどの標的核酸を含み得る標的細胞)として、細菌細胞;古細菌細胞;単細胞真核生物の細胞;植物細胞;藻類細胞、例えば、Botryococcus braunii、Chlamydomonas reinhardtii、Nannochloropsis gaditana、Chlorella pyrenoidosa、Sargassum patens、C. agardhなど);真菌細胞(例えば酵母細胞);動物細胞;無脊椎動物(例えば、ショウジョウバエ、刺胞動物、棘皮動物、線虫など)由来の細胞;昆虫(例えば、カ、ハチ、農業害虫など)の細胞;クモ類(例えば、クモ、ダニなど)の細胞;脊椎動物(例えば、魚類、両生類、爬虫類、鳥類、哺乳類)由来の細胞;哺乳動物由来の細胞(例えば、齧歯類由来の細胞;ヒト由来の細胞;非ヒト哺乳動物由来の細胞);齧歯類(例えば、マウス、ラット)の細胞;ウサギ目(例えば、ウサギ)の細胞;有蹄類(例えば、ウシ、ウマ、ラクダ、ラマ、ビクーナ、ヒツジ、ヤギなど)の細胞;海洋哺乳動物(例えば、クジラ、アザラシ、ゾウアザラシ、イルカ、アシカなど)の細胞が挙げられるが、これらに限定されない。いかなる種類の細胞も対象となり得る(例えば、幹細胞、例えば、胚性幹(ES)細胞、人工多能性幹(iPS)細胞、生殖細胞(例えば、卵母細胞、精子、卵原細胞、精原細胞など)、成体幹細胞、体細胞、例えば線維芽細胞、造血細胞、ニューロン、筋肉細胞、骨細胞、肝細胞、膵臓細胞;任意の段階の胚のin vitroまたはin vivo胚細胞、例えば、1細胞期、2細胞期、4細胞期、8細胞期などのゼブラフィッシュ胚など)。

【0279】

細胞は、樹立細胞株由来のものであってよく、または初代細胞であってもよい。ここで、「初代細胞」、「初代細胞株」、及び「初代培養物」は、本明細書で同義に使用され、対象に由来しており、培養物の限定回数(すなわち分割)により、in vitroで増殖させることができる細胞及び細胞培養物を指す。例えば、初代培養物は、0回、1回、2回、4回、5回、10回、または15回継代されていてもよいが、危機的段階を経験するほどの回数は継代されていない培養物である。通常、初代細胞株のin vitroでの継代は10回未満に保たれる。標的細胞は単細胞生物であってもよく、及び/または培養で増殖させることもできる。細胞が初代細胞である場合、細胞を任意の利便な方法により個体から採取することができる。例えば、アフエレーシス、白血球アフエレーシス、密度勾配分離などによって、白血球を利便に採取できる一方、皮膚、筋肉、骨髄、脾臓、肝臓、膵臓、肺、腸、胃などの組織由来の細胞を生検によって利便に採取することができる。

【0280】

上記の用途のいくつかでは、本発明の方法が、in vivo及び/またはex vivo及び/またはin vitroの有糸分裂細胞または有糸分裂後細胞において、標的核酸の切断、標的核酸の修飾を誘導する、及び/または(例えば、可視化、回収、及び/または分析などのために)標的核酸に結合させる際に用いられる場合がある(例えば、標的mRNAによってコードされるタンパク質の産生を妨害する、標的DNAを切断またはその他の方法により修飾する、標的細胞を遺伝子改変するなど)。ガイドRNAは標的核酸にハイブリダイズすることによって特異性を与えるため、開示される方法で対象となる

10

20

30

40

50

有糸分裂細胞及び/または有糸分裂後細胞は、どの生物に由来する細胞をも含み得る（例えば、細菌細胞、古細菌細胞、単細胞真核生物の細胞、植物細胞、藻類細胞、例えば、*Botryococcus braunii*、*Chlamydomonas reinhardtii*、*Nannochloropsis gaditana*、*Chlorella pyrenoidosa*、*Sargassum patens*、*C. agardh*など、真菌細胞（例えば酵母細胞）、動物細胞、無脊椎動物（例えば、ショウジョウバエ、刺胞動物、棘皮動物、線虫など）由来の細胞、脊椎動物（例えば、魚類、両生類、爬虫類、鳥類、哺乳類）由来の細胞、哺乳動物由来の細胞、齧歯類由来の細胞、ヒト由来の細胞）。場合によって、本発明のCasYタンパク質（及び/またはタンパク質をコードする核酸、例えばDNA及び/またはRNA）、及び/またはCasYガイドRNA（及び/またはガイドRNAをコードするDNA）、及び/またはドナー鋳型、及び/またはRNPは、個体（例えば、哺乳動物、ラット、マウス、ブタ、霊長類、非ヒト霊長類、ヒトなど）に導入することができる（すなわち、標的細胞は*in vivo*であり得る）。場合によって、そのような投与は、例えば、標的細胞のゲノムを編集することによる疾患の治療及び/または予防を目的としたものであり得る。

10

【0281】

植物細胞は単子葉植物の細胞、及び双子葉植物の細胞を含む。この細胞は、根細胞、葉細胞、木部の細胞、師部の細胞、形成層の細胞、頂端分裂組織細胞、柔細胞、厚角組織細胞、厚膜組織細胞などであり得る。植物細胞には、小麦、トウモロコシ、イネ、モロコシ、キビ、大豆などの農作物の細胞を含む。植物細胞には、農産果実及びナッツ植物、例えば、アプリコット、オレンジ、レモン、リンゴ、プラム、ナシ、アーモンドなどを実らせる植物の細胞を含む。

20

【0282】

標的細胞のさらなる例は、上記の「改変細胞」という表題のセクションで列挙されている。細胞（標的細胞）の非限定的な例として、原核細胞、真核細胞、細菌細胞、古細菌細胞、単細胞真核生物の細胞、原生動物細胞、植物由来細胞（例えば、作物（植物）、果実、野菜、穀物、大豆、トウモロコシ（*corn*、*maize*）、小麦、種子、トマト、イネ、キャッサバ、サトウキビ、カボチャ、干し草、ジャガイモ、綿、大麻、タバコ、顕花植物、針葉樹、裸子植物、被子植物、シダ類、ヒカゲノカズラ類、ツノゴケ類、ゼニゴケ類、セン類、双子葉植物、単子葉植物など由来の細胞）、藻類細胞（例えば、*Botryococcus braunii*、*Chlamydomonas reinhardtii*、*Nannochloropsis gaditana*、*Chlorella pyrenoidosa*、*Sargassum patens*、*C. agardh*など）、海藻（例えばケルプ）、真菌細胞（例えば、酵母細胞、キノコ由来の細胞）、動物細胞、無脊椎動物（例えば、ショウジョウバエ、刺胞動物、棘皮動物、線虫など）由来の細胞、脊椎動物（例えば、魚類、両生類、爬虫類、鳥類、哺乳類）由来の細胞、哺乳動物（例えば、有蹄類（例えば、ブタ、ウシ、ヤギ、ヒツジ）；齧歯類（例えば、ラット、マウス）；非ヒト霊長類；ヒト；ネコ科動物（例えば、ネコ）；イヌ科動物（例えば、イヌ）など）由来の細胞などが挙げられる。場合によって、細胞は、天然の生物に由来しない細胞である（例えば、細胞は、合成的に作製された細胞であり得る：これは人工細胞とも呼ばれる）。

30

40

【0283】

細胞は*in vitro*の細胞（例えば、樹立培養細胞株）であってよい。細胞は*ex vivo*の細胞（個体由来の培養細胞）であってよい。細胞は*in vivo*の細胞（例えば、個体内の細胞）であってよい。細胞は単離細胞であってよい。細胞は生物の体内細胞であってよい。細胞は生物であってよい。細胞は細胞培養物（例えば、*in vitro*細胞培養物）中の細胞であってよい。細胞は細胞の集合のうちの一つであってよい。細胞は、原核細胞であってよい、または原核細胞に由来するものであってもよい。細胞は、細菌細胞であってよい、または細菌細胞に由来するものであってもよい。細胞は、古細菌細胞であってよい、または古細菌細胞に由来するものであってもよい。細胞は、真核細胞であってよい。

50

ても、または真核細胞に由来するものであってもよい。細胞は、植物細胞であっても、または植物細胞に由来するものであってもよい。細胞は、動物細胞であっても、または動物細胞に由来するものであってもよい。細胞は、無脊椎動物細胞であっても、または無脊椎動物細胞に由来するものであってもよい。細胞は、脊椎動物細胞であっても、または脊椎動物細胞に由来するものであってもよい。細胞は、哺乳動物細胞であっても、または哺乳動物細胞に由来するものであってもよい。細胞は、齧歯類細胞であっても、または齧歯類細胞に由来するものであってもよい。細胞は、ヒト細胞であっても、またはヒト細胞に由来するものであってもよい。細胞は、微生物細胞であっても、または微生物細胞に由来するものであってもよい。細胞は、真菌細胞であっても、または真菌細胞に由来するものであってもよい。細胞は昆虫細胞であってよい。細胞は節足動物細胞であってよい。細胞は原生動物細胞であってよい。細胞は蠕虫細胞であってよい。

【0284】

好適な細胞として、幹細胞（例えば、胚性幹（ES）細胞、人工多能性幹（iPS）細胞）；生殖細胞（例えば、卵母細胞、精子、卵原細胞、精原細胞など）；体細胞、例えば線維芽細胞、オリゴデンドロサイト、グリア細胞、造血細胞、ニューロン、筋肉細胞、骨細胞、肝細胞、膵臓細胞などが挙げられる。

【0285】

好適な細胞として、ヒト胚性幹細胞、胎児心筋細胞、筋線維芽細胞、間葉系幹細胞、自家移植拡張心筋細胞、脂肪細胞、全能性細胞、多能性細胞、血液幹細胞、筋芽細胞、成体幹細胞、骨髄細胞、間葉系細胞、胚性幹細胞、実質細胞、上皮細胞、内皮細胞、中皮細胞、線維芽細胞、骨芽細胞、軟骨細胞、外因性細胞、内因性細胞、幹細胞、造血幹細胞、骨髄由来前駆細胞、心筋細胞、骨格細胞、胎児細胞、未分化細胞、多能性前駆細胞、単能性前駆細胞、単球、心筋芽細胞、骨格筋芽細胞、マクロファージ、毛細血管内皮細胞、異種細胞、同種細胞、及び出生後幹細胞が挙げられる。

【0286】

場合によって細胞は、免疫細胞、ニューロン、上皮細胞、及び内皮細胞、または幹細胞である。場合によって、免疫細胞は、T細胞、B細胞、単球、ナチュラルキラー細胞、樹状細胞、またはマクロファージである。場合によって、免疫細胞は細胞傷害性T細胞である。場合によって、免疫細胞はヘルパーT細胞である。場合によって、免疫細胞は制御性T細胞（Treg）である。

【0287】

場合によって、細胞は幹細胞である。幹細胞には成体幹細胞を含む。成体幹細胞は、体性幹細胞とも呼ばれる。

【0288】

成体幹細胞は、分化組織に常在しているが、自己再生及び複数の細胞型（通常は、幹細胞が見られる組織に特有の細胞型）の発生能といった特性を保有する。体性幹細胞の多数の例が、当業者に知られており、これには、筋幹細胞；造血幹細胞；上皮幹細胞；神経幹細胞；間葉系幹細胞；乳腺幹細胞；腸管幹細胞；中胚葉幹細胞；内皮幹細胞；嗅覚幹細胞；神経堤幹細胞などが含まれる。

【0289】

対象となる幹細胞には哺乳動物の幹細胞を含むが、その場合の用語「哺乳動物」とは、哺乳動物として分類されるいずれかの動物を指し、ヒト；非ヒト霊長類；家畜及び農場動物；ならびに動物園の動物、実験動物、競技用動物、または愛玩動物、例えば、イヌ、ウマ、ネコ、ウシ、マウス、ラット、ウサギなどを含む。場合によって、幹細胞はヒト幹細胞である。場合によって、幹細胞は齧歯類（例えばマウス；ラット）幹細胞である。場合によって、幹細胞は非ヒト霊長類の幹細胞である。

【0290】

幹細胞は、1つ以上の幹細胞マーカー、例えば、SOX9、KRT19、KRT7、LGR5、CA9、FXFD2、CDH6、CLDN18、TSPAN8、BPIFB1、OLFM4、CDH17、及びPPARGC1Aを発現することができる。

10

20

30

40

50

【0291】

いくつかの実施形態では、幹細胞は造血幹細胞（HSC）である。HSCは、骨髄、血液、臍帯血、胎児肝、及び卵黄嚢から単離することができる中胚葉由来の細胞である。HSCは、CD34⁺及びCD3⁻を特徴とする。HSCは、*in vivo*の赤血球、好中球マクロファージ、巨核球、及びリンパ系の造血細胞株を再増殖することができる。*in vitro*で、少なくともいくつかの自己再生細胞分裂を受けるようにHSCを誘導することができる。したがって、HSCは赤血球系細胞、巨核球、好中球、マクロファージ、及びリンパ系細胞のうち1つ以上に分化するように誘導することができる。

【0292】

他の実施形態では、幹細胞は神経幹細胞（NSC）である。神経幹細胞（NSC）は、ニューロン及びグリア（オリゴデンドロサイト及びアストロサイトを含む）に分化することができる。神経幹細胞は、多様な分化が可能な多能性幹細胞であり、特定の条件下で、神経幹細胞である娘細胞、または神経芽細胞またはグリア芽細胞となり得る神経前駆細胞、例えば、それぞれ1つ以上の種類のニューロン及びグリア細胞となることが約束された細胞を産生することができる。NSCの入手方法は、当技術分野において公知である。

【0293】

他の実施形態では、幹細胞は間葉系幹細胞（MSC）である。MSCは本来、胚中胚葉に由来し、成体骨髄から単離され、分化して筋肉、骨、軟骨、脂肪、骨髄間質、及び腱を形成することができる。MSCの単離方法は当技術分野で公知であり、任意の公知の方法を使用してMSCを得ることができる。例えば、ヒトMSCの単離について記載している米国特許第5,736,396号を参照のこと。

【0294】

細胞は場合によって植物細胞である。植物細胞は、単子葉植物の細胞であってよい。細胞は、双子葉植物の細胞であってよい。

【0295】

場合によって、細胞は植物細胞である。例えば、細胞は、主要な農業植物、例えば、大麦、豆（食用乾燥）、キャノーラ、トウモロコシ、綿（ピマ）、綿（高地）、亜麻仁、干し草（アルファルファ）、干し草（アルファルファ以外）、カラスムギ、ピーナッツ、イネ、モロコシ、大豆、テンサイ、サトウキビ、ヒマワリ（油）、ヒマワリ（油以外）、サツマイモ、タバコ（パーレー種）、タバコ（熱風乾燥）、トマト、小麦（デュラム）、小麦（春）、小麦（冬）などの細胞であり得る。別の例として、細胞は、野菜作物の細胞であり、限定されないが、例えば、アルファルファもやし、アロエの葉、クズウコン、クワイ、アーティチョーク、アスパラガス、タケノコ、バナナの花、モヤシ、豆、ビート茎葉部、ビート、ニガウリ、チンゲンサイ、ブロッコリー、ブロッコリーレイブ（ラピニ）、芽キャベツ、キャベツ、キャベツスプラウト、サボテンの葉（ノパレス）、ヒョウタン、カルドン、ニンジン、カリフラワー、セロリ、ハヤトウリ、チョロギ（クローヌ）、白菜、キンサイ、ニラ、サイシン、キクナ（春菊）、コラード若葉、トウモロコシの茎、スイートコーン、キュウリ、ダイコン、タンポポの葉、タロイモ、豆苗（エンドウ若葉）、トウガン（冬瓜）、ナス、エンダイブ、キクヂシャ、ゼンマイシダ、グンバイナズナ、チコリ、ガイチョイ（芥菜）、ガイラン、ガランガ（シャム、タイショウガ）、ニンニク、ショウガの根、ゴボウ、緑色野菜、ハノーバーサラダの葉、ウアウソントレ、キクイモ、ヒカマ、ケールの葉、コールラビ、シロザ（アマランサス）、レタス（ピブ）、レタス（ボストン）、レタス（ボストンレッド）、レタス（グリーンリーフ）、レタス（アイスバーグ）、レタス（ロロロッサ）、レタス（オークリーフグリーン）、レタス（オークリーフレッド）、レタス（加工）、レタス（レッドリーフ）、レタス（ロメイン）、レタス（ルビーロメイン）、レタス（ロシアンレッドマスタード）、ヒシの実、ロボク、ササゲ、レンコン、ノヂシャ、リュウゼツラン（アガーベ）の葉、タロイモ、メスキュランミックス、ミズナ、モーブ（moap）（ヘチマ）、ムー、マクア（ファジースカッシュ）、マッシュルーム、マスタード、ナガイモ、オクラ、空心菜、九条ネギ、オポ（ロングスカッシュ

10

20

30

40

50

ユ)、観賞用トウモロコシ、観賞用ヒョウタン、パセリ、パースニップ、エンドウマメ、トウガラシ(ベルタイプ)、トウガラシ、カボチャ、エンダイブ、カイワレナ、ラディッシュ、アブラナ、アブラナ、ルバーブ、ロメイン(ベイビーレッド)、ルタバガ、アケシソウ(シービーンズ)、ヘチオク(トカドノ十角ヘチマ)、ホウレンソウ、スクオッシュ、ストローベイル、サトウキビ、サツマイモ、スイスチャード、タマリンド、タロイモ、タロイモの葉、タロイモの茎、ターサイ、テペグアヘ(グアヘ)、ティンドラ、トマテイヨ、トマト、トマト(チェリー)、トマト(グレープ種)、トマト(プラム種)、ターメリック、カブ上菜、カブ、オオクログワイ、ヤンピ、ヤムイモ(ナメス)、ユーチョイ、ユカ(キャッサバ)などを含む。

【0296】

細胞は、場合によって節足動物細胞である。例えば、細胞は、例えば、Chelicerata、Myriapodia、Hexipodia、Arachnida、Insecta、Archaeognatha、Thysanura、Palaeoptera、Ephemeroptera、Odonata、Anisoptera、Zygoptera、Neoptera、Exopterygota、Plecoptera、Embiodoptera、Orthoptera、Zoraptera、Dermaptera、DICTYOPTERA、Notoptera、Grylloblattidae、Mantophasmatidae、Phasmatodea、Blattaria、Isoptera、Mantodea、Parapneuroptera、Psocoptera、Thysanoptera、Phthiraptera、Hemiptera、EndopterygotaもしくはHolometabola、Hymenoptera、Coleoptera、Strepsiptera、Raphidioptera、Megaloptera、Neuroptera、Mecoptera、Siphonaptera、Diptera、Trichoptera、またはLepidopteraの亜目、科、亜科、属、亜属、または種の細胞であり得る。

【0297】

細胞は、場合によって昆虫細胞である。例えば、場合によって、細胞は、カ、バッタ、半翅目の昆虫、ハエ、ノミ、ハチ、スズメバチ、アリ、シラミ、ガ、または甲虫の細胞である。

【0298】

標的細胞への構成要素の導入

Cas9ガイドRNA(またはそれをコードするヌクレオチド配列を含む核酸)及び/またはCas9融合ポリペプチド(またはそれをコードするヌクレオチド配列を含む核酸)及び/またはドナーポリヌクレオチドは、周知の様々な方法のいずれかによって宿主細胞に導入することができる。

【0299】

細胞に核酸を導入する方法は、当技術分野で公知であり、任意の利便な方法を使用して、核酸(例えば、発現構築物)を標的細胞(例えば、真核細胞、ヒト細胞、幹細胞、前駆細胞など)に導入することができる。好適な方法は本明細書の別の箇所で詳細に記載しており、例えば、ウイルスまたはバクテリオファージ感染、トランスフェクション、コンジュゲーション、プロトプラスト融合、リポフェクション、エレクトロポレーション、リン酸カルシウム沈降、ポリエチレンジイミン(PEI)介在型トランスフェクション、DEAE-デキストラン介在型トランスフェクション、リポソーム介在型トランスフェクション、パーティクルガン技術、リン酸カルシウム沈降、直接マイクロインジェクション、ナノ粒子介在型核酸送達(例えば、Panyam et al Adv Drug Deliv Rev. 2012 Sep 13. pii: S0169-409X(12)00283-9. doi: 10.1016/j.addr.2012.09.023を参照)などが挙げられる。構成要素の一部または全部を、例えばヌクレオフェクションなどの公知の方法を使用して、組成物(例えば、Cas9ポリペプチド、Cas9ガイドRNA、ドナーポリヌクレオチドなどの任意の利便な組み合わせを含む)として細胞に導入すること

10

20

30

40

50

ができる。

【0300】

ドナーポリヌクレオチド（ドナー鑄型）

CasYタンパク質は、CasYガイドRNAによって誘導されて、場合によって、二本鎖DNA（dsDNA）標的核酸内の部位特異的二本鎖切断（DSB）または一本鎖切断（SSB）（例えば、CasYタンパク質がニッカーゼ変異体である場合）を生成し、非相同末端結合（NHEJ）または相同指向性組換え（HDR）のいずれかによって切断を修復する。

【0301】

場合によって、（CasYタンパク質及びCasYガイドRNAとの）標的DNAの接触は、非相同末端結合または相同組換え修復が許容される条件下で行う。したがって、場合によって、本発明の方法は、（例えば、細胞へのドナーポリヌクレオチドの導入によって）標的DNAをドナーポリヌクレオチドと接触させ、ドナーポリヌクレオチド、ドナーポリヌクレオチドの一部、ドナーポリヌクレオチドのコピー、またはドナーポリヌクレオチドのコピーの一部を標的DNAに組み込むことを含む。場合によって、この方法は、ドナーポリヌクレオチドと細胞を接触させることを含まず、かつ、標的DNAは、標的DNA内のヌクレオチドが欠失されるように改変される。

10

【0302】

場合によって、CasYガイドRNA（またはそれをコードするDNA）及びCasYタンパク質（またはRNAもしくはDNAなどの、それをコードする核酸、例えば1つ以上の発現ベクター）は、少なくとも標的DNA配列と相同性を有するセグメントを含むドナーポリヌクレオチド配列と同時投与される（例えば、標的核酸と接触させる、細胞に投与されるなど）。本発明の方法は、核酸材料を標的DNA配列に付加する、すなわち挿入または置換する（例えば、核酸、例えばタンパク質をコードする核酸、siRNA、miRNAなどを「ノックイン」する）、タグ（例えば、6xHis、蛍光タンパク質（例えば、緑色蛍光タンパク質；黄色蛍光タンパク質など）、ヘマグルチニン（HA）、FLAGなど）を付加する、遺伝子に制御配列（例えば、プロモーター、ポリアデニル化シグナル、内部リボソーム進入配列（IRES）、2Aペプチド、開始コドン、終止コドン、スプライスシグナル、移行シグナルなど）を付加する、核酸配列を修飾する（例えば、変異を導入する、正しい配列を導入することにより疾患原因変異を除去する）などのために使用することができる。したがって、CasYガイドRNA及びCasYタンパク質を含む複合体は、部位特異的、すなわち「標的指向性」方法でのDNA修飾、例えば遺伝子ノックアウト、遺伝子ノックイン、遺伝子編集、遺伝子タグ付けなどが望ましい、いずれのin vitroまたはin vivo用途、例えば、例えば疾患の治療、または抗ウイルス、抗病原性、もしくは抗がん療法としての遺伝子療法、農業における遺伝子組換え作物の生産、治療、診断または研究を目的とした細胞によるタンパク質の大規模製造、iPS細胞の誘導、生物学的研究、欠失または置換などのための病原体遺伝子の標的化などでの使用に有用である。

20

30

【0303】

標的配列が切断された場所のゲノムにポリヌクレオチド配列を挿入することが望まれる用途でも、ドナーポリヌクレオチド（ドナー配列を含む核酸）を細胞に供給することができる。「ドナー配列」または「ドナーポリヌクレオチド」または「ドナー鑄型」とは、（例えば、dsDNAの切断後、標的DNAのニックング後、標的DNAの二重ニックング後などに）CasYタンパク質によって切断された部位に挿入される核酸配列を意味する。ドナーポリヌクレオチドは、相同性を有するゲノム配列との間の相同組換え修復を維持するため、標的部位のゲノム配列に対して十分な相同性、例えば、標的部位に隣接するヌクレオチド配列、例えば、標的部位の約50塩基以内、例えば、約30塩基以内、約15塩基以内、約10塩基以内、約5塩基以内、または標的部位とすぐ隣接するヌクレオチド配列と、70%、80%、85%、90%、95%、または100%の相同性を備え得る。ドナーとゲノム配列との間の配列相同性が、約25、50、100、または200ヌク

40

50

レオチド、または200ヌクレオチド超（または10～200ヌクレオチド、もしくはそれ以上の任意の整数値）の場合、相同組換え修復を維持することができる。ドナーポリヌクレオチドは、任意の長さ、例えば10ヌクレオチド以上、50ヌクレオチド以上、100ヌクレオチド以上、250ヌクレオチド以上、500ヌクレオチド以上、1000ヌクレオチド以上、5000ヌクレオチド以上などのものであり得る。

【0304】

ドナー配列は通常、置換するゲノム配列と同一ではない。むしろ、（例えば遺伝子修正、例えば、疾患の原因となる塩基対または疾患の原因ではない塩基対の変換のための）相同組換え修復を維持するのに十分な相同性が存在する限り、ドナー配列は、ゲノム配列に対して、少なくとも1つ以上の単一塩基の変更、挿入、欠失、逆位、または再配置を含んでいてもよい。いくつかの実施形態では、ドナー配列は、相同性のある2つの領域が隣接する非相同配列を含み、それによって標的DNA領域と2つの隣接配列との間の相同組換え修復の結果、標的領域に非相同配列が挿入される。ドナー配列にはまた、対象となるDNA領域に相同ではなく、対象となるDNA領域への挿入を意図しない配列を含むベクター骨格を含み得る。一般に、ドナー配列の相同領域（複数可）は、組換えを必要とするゲノム配列に対して、少なくとも50%の配列同一性を有する。ある特定の実施形態では、60%、70%、80%、90%、95%、98%、99%、または99.9%の配列同一性が存在する。ドナーポリヌクレオチドの長さに応じて、1%～100%の任意の値の配列同一性が存在し得る。

【0305】

ドナー配列は、ゲノム配列など、例えば制限部位、ヌクレオチド多型、選択マーカー（例えば、薬剤耐性遺伝子、蛍光タンパク質、酵素など）と比較して、ある一定の配列相違性を含んでいてもよく、切断部位でのドナー配列の挿入の成功を評価するために使用することも、または場合によって他の目的（例えば、標的ゲノム遺伝子座での発現を表すため）に使用することもできる。場合によって、コード領域に位置する場合、このようなヌクレオチド配列の相違により、アミノ酸配列が変更されないか、またはアミノ酸がサイレント変異されることになる（すなわち、タンパク質の構造または機能に影響を与えない）。あるいは、これらの配列の相違には、後で活性化されてマーカー配列を除去することができる、FLP、loxP配列などの隣接する組換え配列を含み得る。

【0306】

場合によって、ドナー配列は、一本鎖DNAとして細胞に供給される。場合によって、ドナー配列は、二本鎖DNAとして細胞に供給される。これは、直鎖形態または環状形態で細胞に導入することができる。直鎖形態で導入された場合、ドナー配列の末端は、任意の利便な方法によって（例えば、エキソヌクレアーゼ分解から）保護されていてもよく、そのような方法は当業者に公知である。例えば、1つ以上のジデオキシヌクレオチド残基を直鎖分子の3'末端に付加すること、及び/または自己相補性オリゴヌクレオチドを一方または両方の末端にライゲートすることができる。例えば、Chang et al. (1987) Proc. Natl. Acad. Sci. USA 84: 4959-4963; Nehls et al. (1996) Science 272: 886-889を参照のこと。外因性ポリヌクレオチドを分解から保護するための追加の方法には、末端アミノ基（複数可）の付加、及び修飾ヌクレオチド間結合、例えばホスホロチオエート、ホスホロアミダート、及びO-メチルリボースまたはデオキシリボース残基などの使用が挙げられるが、これらに限定されない。直鎖ドナー配列の末端を保護する代わりに、相同性領域の外側にさらに配列の長さを加えることで、組換えに影響を与えることなく分解することができる。ドナー配列は、例えば、複製起点、プロモーター、及び抗生物質抵抗性をコードする遺伝子のような付加的な配列を有するベクター分子の一部として細胞に導入することができる。さらに、ドナー配列は、裸の核酸として、リポソームまたはポロキサマーなどの薬剤と複合体化された核酸として導入することができ、あるいはCasYガイドRNA及び/またはCasY融合ポリペプチド及び/またはドナーポリヌクレオチドをコードする核酸について本明細書の他の箇所に記載するように、ウイルス（例えば、アデノ

ウイルス、AAV)によって送達することができる。

【0307】

トランスジェニック非ヒト生物

上述するように、場合によって、本開示の核酸(例えば、組換え発現ベクター)(例えば、本開示のCasYポリペプチドをコードするヌクレオチド配列を含む核酸;本開示のCasY融合ポリペプチドをコードするヌクレオチド配列を含む核酸など)は、本開示のCasYポリペプチドまたはCasY融合ポリペプチドを産生するトランスジェニック非ヒト生物を作製するための導入遺伝子として使用される。本開示は、本開示のCasYポリペプチドまたはCasY融合ポリペプチドをコードするヌクレオチド配列を含むトランスジェニック非ヒト生物を提供する。

10

【0308】

トランスジェニック非ヒト動物

本開示は、トランスジェニック非ヒト動物を提供する。この動物は、CasYポリペプチドまたはCasY融合ポリペプチドをコードするヌクレオチド配列を含む核酸を含んでいる導入遺伝子を含む。いくつかの実施形態では、トランスジェニック非ヒト動物のゲノムは、本開示のCasYポリペプチド、またはCasY融合ポリペプチドをコードするヌクレオチド配列を含む。場合によって、トランスジェニック非ヒト動物は、遺伝子改変されるホモ接合体である。場合によって、トランスジェニック非ヒト動物は、遺伝子改変されるヘテロ接合体である。いくつかの実施形態では、トランスジェニック非ヒト動物は、脊椎動物、例えば、魚類(例えば、サケ、マス、ゼブラフィッシュ、キンギョ、フグ、洞窟魚など)、両生類(カエル、イモリ、サンショウウオなど)、鳥類(例えば、ニワトリ、七面鳥など)、爬虫類(例えば、ヘビ、トカゲなど)、非ヒト哺乳動物など(例えば、有蹄動物、例えば、ブタ、ウシ、ヤギ、ヒツジ、など;ウサギ目(例えばウサギ);齧歯類(例えば、ラット、マウス);非ヒト霊長類など)である。場合によって、トランスジェニック非ヒト動物は無脊椎動物である。場合によって、トランスジェニック非ヒト動物は昆虫(例えば、カ;農業害虫など)である。場合によって、トランスジェニック非ヒト動物はクモ類である。

20

【0309】

本開示のCasYポリペプチドまたはCasY融合ポリペプチドをコードするヌクレオチド配列は、(例えば、核酸が宿主細胞ゲノムに無作為に組み込まれる場合)未知のプロモーターの制御下にあってもよく(すなわち、機能的に連結される)、既知のプロモーターの制御下にあってもよい(すなわち、機能的に連結される)。好適な既知のプロモーターは任意の既知のプロモーターであってよく、構造的に活性なプロモーター(例えば、CMVプロモーター)、誘導性プロモーター(例えば、熱ショックプロモーター、テトラサイクリン制御性プロモーター、ステロイド制御性プロモーター、金属制御性プロモーター、エストロゲン受容体制御性プロモーターなど)、空間制限及び/または時間制限されたプロモーターなど(例えば、組織特異的プロモーター、細胞型特異的プロモーターなど)を含む。

30

【0310】

トランスジェニック植物

上述するように、場合によって、本開示の核酸(例えば、組換え発現ベクター)(例えば、本開示のCasYポリペプチドをコードするヌクレオチド配列を含む核酸;本開示のCasY融合ポリペプチドをコードするヌクレオチド配列を含む核酸など)は、本開示のCasYポリペプチドまたはCasY融合ポリペプチドを産生するトランスジェニック植物を作製するための導入遺伝子として使用される。本開示は、本開示のCasYポリペプチドまたはCasY融合ポリペプチドをコードするヌクレオチド配列を含むトランスジェニック植物を提供する。いくつかの実施形態では、トランスジェニック植物のゲノムは、本発明の核酸を含む。いくつかの実施形態では、トランスジェニック植物は、遺伝子改変されるホモ接合体である。いくつかの実施形態では、トランスジェニック植物は、遺伝子改変されるヘテロ接合体である。

40

50

【0311】

外因性核酸を植物細胞に導入する方法は当技術分野で周知されている。そのような植物細胞は、上記で定義するように「形質転換された」とみなされる。好適な方法として、ウイルス感染（二本鎖DNAウイルスなど）、トランスフェクション、コンジュゲート、プロトプラスト融合、エレクトロポレーション、パーティクルガン技術、リン酸カルシウム沈降、直接マイクロインジェクション、炭化ケイ素ウィスカー技術、*Agrobacterium* 介在性形質転換などが挙げられる。方法の選択は一般に、形質転換される細胞の種類、及び形質転換が行われる状況（すなわち、*in vitro*、*ex vivo*、または *in vivo*）に応じて異なる。

【0312】

土壌細菌 *Agrobacterium tumefaciens* を用いた形質転換法は、維管束植物に外因性核酸分子を導入するために特に有用である。*Agrobacterium* の野生型形態には、宿主植物での腫瘍形成性クラウンゴールの増殖の発生を誘導する Ti（腫瘍誘発）プラスミドが含まれている。Ti プラスミドの腫瘍誘発性 T-DNA 領域を植物ゲノムに転写するには、Ti プラスミドが病原性遺伝子ならびに T-DNA 境界をコードする必要がある。これらは、転送されるべき領域の境界を示す一連の DNA 直列反復配列である。*Agrobacterium* 系のベクターは Ti プラスミドの改変形態であり、対象となる核酸配列によって腫瘍誘導機能が置き換えられて植物宿主に導入される。

【0313】

Agrobacterium 介在性の形質転換は一般に、融合ベクターシステムまたはバイナリーベクターシステムを用いる。このシステムでは、Ti プラスミドの構成要素が、*Agrobacterium* 宿主に恒久的に常在して病原性遺伝子を保有するヘルパーベクターと、T-DNA 配列によって境界が規定された対象となる遺伝子を含むシャトルベクターとに分かれている。種々のバイナリーベクターが当技術分野において周知され、市販されており、例えば、Clontech (Palo Alto, Calif.) 製のものがある。例えば、葉組織、根の外植片、子葉部、茎片または塊茎などの、培養植物細胞または損傷組織と *Agrobacterium* との共培養方法もまた、当技術分野において周知されている。例えば、Glick and Thompson, (eds.), *Methods in Plant Molecular Biology and Biotechnology*, Boca Raton, Fla.: CRC Press (1993) を参照のこと。

【0314】

マイクロプロジェクトイル介在性形質転換もまた、本発明のトランスジェニック植物の産生に使用することができる。Klein et al. (Nature 327:70-73 (1987)) によって最初に記載されたこの方法は、塩化カルシウム、スペルミジン、またはポリエチレングリコールを用いた沈降によって、所望する核酸分子で被覆した金またはタングステンなどのマイクロプロジェクトイルを利用する。マイクロプロジェクトイル粒子は、BIOLISTIC PD-1000 (Biorad; Hercules Calif.) などの装置を使用して、被子植物組織に高速で打ち込まれる。

【0315】

本開示の核酸（例えば、本開示の CasY ポリペプチドまたは CasY 融合ポリペプチドをコードするヌクレオチド配列を含む核酸（例えば、組換え発現ベクター））は、核酸を植物細胞（複数可）に進入させることができるような方法で、例えば、*in vivo* または *ex vivo* のプロトコールによって植物に導入することができる。「*in vivo*」とは、核酸が植物の生体に投与されること、例えば浸入することを意味する。「*ex vivo*」とは、細胞または外植片を植物の外部で改変した後、そのような細胞または器官を植物に再生させることを意味する。植物細胞の安定な形質転換またはトランスジェニック植物の樹立に適したベクターがいくつか記載されており、それには、Weissbach and Weissbach, (1989) *Methods for Pl*

10

20

30

40

50

ant Molecular Biology Academic Press、及び Gelvin et al., (1990) Plant Molecular Biology Manual, Kluwer Academic Publishersに記載されているものが含まれる。具体例として、Agrobacterium tumefaciensのTiプラスミドに由来するもの、ならびに Herrera-Estrella et al. (1983) Nature 303:209, Bevan (1984) Nucleic Acid Res. 12:8711-8721, Klee (1985) Bio/Technology 3:637-642に記載されているものが挙げられる。あるいは、非Tiベクターを使用して、遊離DNA送達技術を用いて植物及び細胞にDNAを転写することができる。これらの方法を使用して、小麦、イネ(Christou (1991) Bio/Technology 9:957-9及び4462)、及びトウモロコシ(Gordon-Kamm (1990) Plant Cell 2:603-618)などのトランスジェニック植物を生産することができる。未熟胚もまた、パーティクルガンを使用した直接DNA送達技術(Weeks et al. (1993) Plant Physiology 102:1077-1084; Vasil (1993) Bio/Technology 10:667-674; Wan and Lemeaux (1994) Plant Physiology 104:37-48)、及びAgrobacterium介在性DNA転写(Ishida et al. (1996) Nature Biotech 14:745-750)において、単子葉植物に適切な標的組織であり得る。葉緑体にDNAを導入するための例示的な方法は、バイオリスティックボンバードメント、プロトプラストのポリエチレングリコール形質転換、及びマイクロインジェクションである(Danieli et al. Nat. Biotechnol 16:345-348, 1998; Staub et al. Nat. Biotechnol 18:333-338, 2000; O'Neill et al. Plant J. 3:729-738, 1993; Knoblauch et al. Nat. Biotechnol 17:906-909; 米国特許第5,451,513号、第5,545,817号、第5,545,818号、及び第5,576,198号; 国際出願第WO95/16783号; ならびに Boynton et al., Methods in Enzymology 217:510-536 (1993)、Svab et al., Proc. Natl. Acad. Sci. USA 90:913-917 (1993)、及び McBride et al., Proc. Natl. Acad. Sci. USA 91:7301-7305 (1994))。バイオリスティックボンバードメント、プロトプラストのポリエチレングリコール形質転換、及びマイクロインジェクションの方法に適した任意のベクターは、葉緑体形質転換のための標的化ベクターとして適している。任意の二本鎖DNAベクターは、特にAgrobacteriumを利用しない導入方法の場合に、形質転換ベクターとして使用することができる。

【0316】

遺伝子改変することができる植物には、穀物、飼料作物、果実、野菜、脂肪種子作物、ヤシ、森林、及び蔓植物を含む。改変できる植物の具体例は以下の通りである：トウモロコシ、バナナ、ピーナッツ、フィールドピー、ヒマワリ、トマト、キャノーラ、タバコ、小麦、大麦、オート麦、ジャガイモ、大豆、綿、カーネーション、モロコシ、ハウチワマメ、及びイネ。

【0317】

本開示は、形質転換された植物細胞、組織、植物、及び形質転換された植物細胞を含む産物を提供する。本発明の形質転換細胞及び組織、ならびにそれを含む産物の特徴は、ゲノムに組み込まれた本発明の核酸の存在、及び植物細胞による本開示のCasYポリペプチドまたはCasY融合ポリペプチドの産生である。本発明の組換え植物細胞は、組換え細胞の集団として、または組織、種子、植物全体、茎、果実、葉、根、花、茎、塊茎、穀粒、家畜飼料、植物用地などとして有用である。

【0318】

本開示の CasY ポリペプチドまたは CasY 融合ポリペプチドをコードするヌクレオチド配列は、（例えば、核酸が宿主細胞ゲノムに無作為に組み込まれる場合）未知のプロモーターの制御下にあってもよく（すなわち、機能的に連結される）、既知のプロモーターの制御下にあってもよい（すなわち、機能的に連結される）。好適な既知のプロモーターは任意の既知のプロモーターであってよく、構成的に活性なプロモーター、誘導性プロモーター、空間制限及び/または時間制限されたプロモーターなどを含む。

【0319】

CRISPR RNA 誘導型エンドヌクレアーゼの同定方法

CRISPR RNA 誘導型エンドヌクレアーゼの同定方法を提供する。例えば、いくつかの実施形態では、そのような方法には、複数のメタゲノムヌクレオチド配列において、Cas1 ポリペプチドをコードするヌクレオチド配列を検出する工程を含む。Cas1 タンパク質は、当技術分野において公知であり、クラス2 CRISPR システムの CRISPR 遺伝子座の近傍に存在している。このような CRISPR システムには、エンドヌクレアーゼとして機能し、タンパク質の複合体との相互作用を必要とせずに適切に機能する単一のエフェクタータンパク質が含まれる。Cas1 タンパク質自体は、CRISPR 遺伝子座への新規標的配列の取り込みに関与しているため、この方法による同定に望ましいエフェクタータンパク質ではないが、CRISPR 遺伝子の近傍に Cas1 タンパク質が存在することは、遺伝子座付近に存在する他の Cas タンパク質の少なくとも1つがエフェクタータンパク質（RNA 誘導型エンドヌクレアーゼ）であり得ることの指標である。

10

20

【0320】

本明細書で使用される場合、用語「メタゲノミクス」とは、試料、例えば未知量の原核生物（細菌/古細菌）を含む試料、及び以前に発見及び/または特性決定されていない原核生物を含み得る試料などの環境試料中の複数の微生物（例えば、細菌、古細菌など）から回収された核酸の並行分析を意味する。任意の利便な方法によって、そのような試料から核酸を回収することができ、一般に、分析前には、任意の所与の核酸分子の起源となる微生物が不明であるような試料全体から核酸を網羅的に回収する。いくつかの実施形態では、試料は、微生物の未知の混合物及び/または未知量の微生物を含有する。核酸をさらに配列決定して、複数のメタゲノム配列を得ることができる。場合によって、CRISPR RNA 誘導型エンドヌクレアーゼを同定する本発明の方法は、試料（例えば、環境試料）を単離する工程を含む。場合によって、CRISPR RNA 誘導型エンドヌクレアーゼを同定する本発明の方法には、試料から核酸を単離する工程、及び/または試料から複数のメタゲノムヌクレオチド配列を得るために試料をアッセイする工程を含む。

30

【0321】

CRISPR RNA 誘導型エンドヌクレアーゼを同定する本発明の方法は、Cas1 タンパク質が同定された後に、Cas1 をコードするヌクレオチド配列の近傍にある CRISPR 配列（リピート-スペーサー-リピート配列）を検出する工程を含み得る。次に、この方法は、検出された CRISPR 配列を含む CRISPR 遺伝子座を（例えば、複数のメタゲノムヌクレオチド配列が由来する核酸試料から）発現ベクターにクローニングして、組換え CRISPR 遺伝子座発現ベクターを生成する工程を含み得る。さらに、組換え CRISPR 遺伝子座発現ベクターが標的核酸を切断する能力についてアッセイすることにより、CRISPR 遺伝子座の機能を試験することができる。任意の利便なアッセイを使用することができる。いくつかの実施形態では、アッセイ工程には、組換え CRISPR 遺伝子座発現ベクター及び標的核酸を、細胞、例えば、E. coli 細胞などの異種宿主細胞に導入することを含む。例えば、下記の実施例（図5）の PAM 欠失アッセイを参照のこと。場合によって、アッセイ工程は、宿主細胞（例えば、E. coli 細胞）の集団にプラスミドライブラリを導入することを含み、その場合、ライブラリの各プラスミドは、標的配列の 5' 及び/または 3' をランダム抽出した 4~10（例えば、5~10、5~8、6~10、6~8、5、6、7、8）個のヌクレオチドを有する。宿主細胞には、試験される組換え CRISPR 遺伝子座発現ベクターが予め含まれていてもよく、ま

40

50

たは組換え C R I S P R 遺伝子座発現ベクターが、ライブラリの後に導入されてもよい。機能的である試験 C R I S P R 遺伝子座、すなわち機能的な C R I S P R R N A 誘導型エンドヌクレアーゼを含む試験 C R I S P R 遺伝子座のみが、標的配列を有するプラスミドを切断する能力をもたらすことになる。標的配列の 5' 及び 3' ランダム配列を含める理由は、実験の開始時に目的とするエンドヌクレアーゼに必要な P A M 配列が未知である可能性があるためである。

【 0 3 2 2 】

発現ベクターが、標的核酸（例えば、適切な標的配列及び P A M を有するもの、例えば、C R I S P R 配列の少なくとも 1 つのスペーサーに適合する標的配列など）を切断することができる場合、その C R I S P R 遺伝子座は、候補の C R I S P R R N A 誘導型エンドヌクレアーゼをコードするヌクレオチド配列を含む。したがって、C R I S P R R N A 誘導型エンドヌクレアーゼをコードする C R I S P R 遺伝子座のオープンリーディングフレームをさらに同定することができる。場合によって、以前に未知であった C R I S P R R N A 誘導型エンドヌクレアーゼを同定することが望ましく、それゆえ、場合によって、同定されるポリペプチドは、既知の C R I S P R R N A 誘導型エンドヌクレアーゼのポリペプチドのアミノ酸配列に対して 20% 未満のアミノ酸配列同一性（例えば、15% 未満、10% 未満、5% 未満のアミノ酸配列同一性）を有する。

10

【 0 3 2 3 】

本開示の非限定的な態様の例

上述する本発明の主題となる、実施形態を含む態様は、単独で、または 1 つ以上の他の態様もしくは実施形態と組み合わせる利益をもたらすことができる。前述の記載、本開示の特定の非限定的な態様に限定することなく、以下の番号 1 ~ 1 2 3 を提供する。本開示を読めば当業者には明らかであるように、個別番号の態様のそれぞれを使用しても、または前後の個別番号の態様のいずれかと組み合わせてもよい。これは、態様のそのような組み合わせすべてに対する支持を示すことを意図しており、以下に明示的に示す態様の組み合わせに限定されない。

20

【 0 3 2 4 】

1 . a) C a s Y ポリペプチド、または前記 C a s Y ポリペプチドをコードする核酸分子、及び

b) C a s Y ガイド R N A、または前記 C a s Y ガイド R N A をコードする 1 つ以上の D N A 分子を含む、組成物。

30

2 . 前記 C a s Y ポリペプチドが、配列番号 1 または配列番号 2 に記載されるアミノ酸配列（または配列番号 1 ~ 8 のいずれか 1 つに記載されるアミノ酸配列）に対して、50% 以上の同一性を有するアミノ酸配列を含む、1 に記載の組成物。

3 . 前記 C a s Y ガイド R N A が、配列番号 1 1 ~ 1 5 のいずれか 1 つに記載される c r R N A 配列と 80% 以上の同一性を有するヌクレオチド配列を含む、1 または 2 に記載の組成物。

4 . 前記 C a s Y ポリペプチドが N L S 配列に融合されている、1 または 2 に記載の組成物。

5 . 前記組成物が脂質を含む、1 ~ 4 のいずれかに 1 つに記載の組成物。

40

【 0 3 2 5 】

6 . a) 及び b) がリポソーム内にある、1 ~ 4 のいずれか 1 つに記載の組成物。

7 . a) 及び b) が粒子内にある、1 ~ 4 のいずれか 1 つに記載の組成物。

8 . 緩衝液、ヌクレアーゼ阻害剤、及びプロテアーゼ阻害剤のうち 1 つ以上を含む、1 ~ 7 のいずれか 1 つに記載の組成物。

9 . 前記 C a s Y ポリペプチドが、配列番号 1 または配列番号 2 に記載されるアミノ酸配列（または配列番号 1 ~ 8 のいずれか 1 つに記載されるアミノ酸配列）に対して、85% 以上の同一性を有するアミノ酸配列を含む、1 ~ 8 のいずれか 1 つに記載の組成物。

10 . 前記 C a s Y ポリペプチドが、二本鎖標的核酸分子の一方の鎖のみを切断することができるニッカーゼである、1 ~ 9 のいずれか 1 つに記載の組成物。

50

【0326】

11. 前記 CasY ポリペプチドが、触媒不活性な CasY ポリペプチド (dCasY) である、1~9 のいずれか 1 つに記載の組成物。

12. 前記 CasY ポリペプチドが、配列番号 1 の D672、E769、及び D935 から選択されるものに対応する位置に 1 つ以上の変異を含む、10 または 11 に記載の組成物。

13. DNA ドナー鑄型をさらに含む、1~12 のいずれかに 1 つに記載の組成物。

14. 異種ポリペプチドに融合された CasY ポリペプチドを含む、CasY 融合ポリペプチド。

15. 前記 CasY ポリペプチドが、配列番号 1 または配列番号 2 に記載されるアミノ酸配列 (または配列番号 1~8 のいずれか 1 つに記載されるアミノ酸配列) に対して、50% 以上の同一性を有するアミノ酸配列を含む、14 に記載の CasY 融合ポリペプチド。

【0327】

16. 前記 CasY ポリペプチドが、配列番号 1 または配列番号 2 に記載されるアミノ酸配列 (または配列番号 1~8 のいずれか 1 つに記載されるアミノ酸配列) に対して、85% 以上の同一性を有するアミノ酸配列を含む、14 に記載の CasY 融合ポリペプチド。

17. 前記 CasY ポリペプチドが、二本鎖標的核酸分子の一方の鎖のみを切断することができるニッカーゼである、14~16 のいずれか 1 つに記載の CasY 融合ポリペプチド。

18. 前記 CasY ポリペプチドが、触媒不活性な CasY ポリペプチド (dCasY) である、14~17 のいずれか 1 つに記載の CasY 融合ポリペプチド。

19. 前記 CasY ポリペプチドが、配列番号 1 の D672、E769、及び D935 から選択されるものに対応する位置に 1 つ以上の変異を含む、17 または 18 に記載の CasY 融合ポリペプチド。

20. 前記異種ポリペプチドが、前記 CasY ポリペプチドの N 末端及び / または C 末端に融合される、14~19 のいずれか 1 つに記載の CasY 融合ポリペプチド。

【0328】

21. NLS を含む、14~20 のいずれか 1 つに記載の CasY 融合ポリペプチド。

22. 前記異種ポリペプチドが、標的細胞または標的細胞型の細胞表面部分への結合性を備える標的化ポリペプチドである、14~21 のいずれか 1 つに記載の CasY 融合ポリペプチド。

23. 前記異種ポリペプチドが、標的 DNA を修飾する酵素活性を示す、14~21 のいずれか 1 つに記載の CasY 融合ポリペプチド。

24. 前記異種ポリペプチドが、ヌクレアーゼ活性、メルトランスフェラーゼ活性、デメチラーゼ活性、DNA 修復活性、DNA 損傷活性、脱アミノ化活性、ジスムターゼ活性、アルキル化活性、脱プリン活性、酸化活性、プリミジン二量体を形成する活性、インテグラーゼ活性、トランスポザーゼ活性、リコンビナーゼ活性、ポリメラーゼ活性、リガーゼ活性、ヘリカーゼ活性、フォトリアーゼ活性、及びグリコシラーゼ活性から選択される 1 つ以上の酵素活性を示す、23 に記載の CasY 融合ポリペプチド。

25. 前記異種ポリペプチドが、ヌクレアーゼ活性、メルトランスフェラーゼ活性、デメチラーゼ活性、脱アミノ化活性、脱プリン活性、インテグラーゼ活性、トランスポザーゼ活性、及びリコンビナーゼ活性から選択される 1 つ以上の酵素活性を示す、24 に記載の CasY 融合ポリペプチド。

【0329】

26. 前記異種ポリペプチドが、標的核酸と会合する標的ポリペプチドを修飾する酵素活性を示す、14~21 のいずれか 1 つに記載の CasY 融合ポリペプチド。

27. 前記異種ポリペプチドが、ヒストン修飾活性を示す、26 に記載の CasY 融合ポリペプチド。

28. 前記異種ポリペプチドが、メチルトランスフェラーゼ活性、デメチラーゼ活性、アセチルトランスフェラーゼ活性、デアセチラーゼ活性、キナーゼ活性、ホスファターゼ活性、ユビキチンリガーゼ活性、脱ユビキチン化活性、アデニル化活性、脱アデニル化活性、SUMO化活性、脱SUMO化活性、リボシル化活性、脱リボシル化活性、ミリストイル化活性、脱ミリストイル化活性、グリコシル化活性（例えば、O-GlcNAcトランスフェラーゼによる）、及び脱グリコシル化活性から選択される1つ以上の酵素活性を示す、26または27に記載のCasY融合ポリペプチド。

29. 前記異種ポリペプチドが、メチルトランスフェラーゼ活性、デメチラーゼ活性、アセチルトランスフェラーゼ活性、及びデアセチラーゼ活性から選択される1つ以上の酵素活性を示す、28に記載のCasY融合ポリペプチド。

30. 前記異種ポリペプチドが、エンドソーム放出ポリペプチドである、14~21のいずれか1つに記載のCasY融合ポリペプチド。

【0330】

31. 前記エンドソーム放出ポリペプチドが、GLFXALLXLLXSLWXLLLXA（配列番号94）及びGLFHALLHLHSLWHLLLHA（配列番号95）から選択されるアミノ酸配列を含み、ここで、各Xは独立して、リジン、ヒスチジン、及びアルギニンから選択される、30に記載のCasY融合ポリペプチド。

32. 前記異種ポリペプチドが、葉緑体輸送ペプチドである、14~21のいずれか1つに記載のCasY融合ポリペプチド。

33. 前記葉緑体輸送ペプチドが、MASMISSSAVTTVSRASRGQSAA
 MAPFGGLKSM TGF PVRKVNTDIT SITSNGGRVKCMQVWPP
 IGKKKFETLSYLPPLTRDSRA（配列番号83）；MASMISSSAV
 TTVSRASRGQSAA MAPFGGLKSM TGF PVRKVNTDIT SITS
 NGGRVKS（配列番号84）；MASSMLSSATMVASPAQATMVAPF
 NGLKSSA AF PATRKANN DIT SITSNGGRVNCMQVWPPIEK
 KKFETLSYLPDLTDSGGRVNC（配列番号85）；MAQVSRICNG
 VQNPSLISNLSKSSQRKSPLSVSLKTQQHPRAYPISSSWG
 LKKS GMTLIGSELRPLKVMSSVSTAC（配列番号86）；MAQVS
 RICNGVWNPSLISNLSKSSQRKSPLSVSLKTQQHPRAYPI
 SSSWGLKKS GMTLIGSELRPLKVMSSVSTAC（配列番号87）；
 MAQINNMAQGIQTLNPN SNFHKPQVPKSSSFLVFGSKKLK
 NSANSM LV LKKDSIFMQLFC SFRISASVATAC（配列番号88）
 ；MAALVTSQLATSGTVLSVTD RFR RP GFQGLRPNPADAA
 LGMRTV GASAA PKQSRKPHRFDRRCLSMVV（配列番号89）；M
 AALTTSQLATSATGFGIADR SAPSSLRHGFQGLKPRSPA
 GGDATSLSVTT SARATPKQQR SVQRGSRFP SVVVC（配列番
 号90）；MASSVLS SA AVATRSNVAQANMVAPFTGLKSAASF
 PVS RKQNL DIT S IASNGGRVQC（配列番号91）；MESLAATSV
 FAPSRVAVPAARALVRAGTVVPTRR TSSSTS GTS GVKCSAA
 VTPQASPVISRSA AAA（配列番号92）；及びMGA AATSMQSLKF
 SNRLVPPSRRLSPVPNNVTCNNLPKSAAPVRTVKCCASSW
 NSTINGAAATTNGASAASS（配列番号93）から選択されるアミノ酸配列
 を含む、32に記載のCasY融合ポリペプチド。

34. 前記異種ポリペプチドが、転写を増加させるかまたは減少させるタンパク質である、14~21のいずれか1つに記載のCasY融合ポリペプチド。

35. 前記異種ポリペプチドが転写抑制因子ドメインである、34に記載のCasY融合ポリペプチド。

【0331】

36. 前記異種ポリペプチドが転写活性化ドメインである、34に記載のCasY融合ポリペプチド。

10

20

30

40

50

37. 前記異種ポリペプチドがタンパク質結合ドメインである、14~21のいずれか1つに記載のCasY融合ポリペプチド。

38. 14~37のいずれか1つに記載のCasY融合ポリペプチドをコードする核酸分子。

39. 前記CasY融合ポリペプチドをコードする前記ヌクレオチド配列が、プロモーターに機能的に連結される、38に記載の核酸分子。

40. 前記プロモーターが、真核細胞において機能的である、39に記載の核酸分子。

【0332】

41. 前記プロモーターが、植物細胞、真菌細胞、動物細胞、脊椎動物の細胞、ハエ細胞、脊椎動物の細胞、哺乳動物細胞、霊長類細胞、非ヒト霊長類細胞、及びヒト細胞のうち1つ以上において機能的である、40に記載の核酸分子。 10

42. 前記プロモーターが、構成的プロモーター、誘導性プロモーター、細胞型特異的プロモーター、及び組織特異的プロモーターのうち1つ以上である、39~41のいずれか1つに記載の核酸分子。

43. 前記DNA分子が組換え発現ベクターである、38~42のいずれか1つに記載の核酸分子。

44. 前記組換え発現ベクターが、組換えアデノ随伴ウイルスベクター、組換えレトロウイルスベクター、または組換えレンチウイルスベクターである、43に記載の核酸分子。

45. 前記プロモーターが、原核細胞において機能的である、39に記載の核酸分子。 20

【0333】

46. 前記核酸分子がmRNAである、38に記載の核酸分子。

47. (a) CasYガイドRNA、及び

(b) CasYポリペプチド

をコードする1つ以上の核酸分子。

48. 前記CasYポリペプチドが、配列番号1または配列番号2に記載されるアミノ酸配列（または配列番号1~8のいずれか1つに記載されるアミノ酸配列）に対して、50%以上の同一性を有するアミノ酸配列を含む、47に記載の1つ以上の核酸分子。

49. 前記CasYポリペプチドが、配列番号1または配列番号2に記載されるアミノ酸配列（または配列番号1~8のいずれか1つに記載されるアミノ酸配列）に対して、85%以上の同一性を有するアミノ酸配列を含む、47に記載の1つ以上の核酸分子。 30

50. 前記CasYガイドRNAが、配列番号11~15のいずれか1つに記載されるcrRNA配列と80%以上の同一性を有するヌクレオチド配列を含む、47~49のいずれか1つに記載の1つ以上の核酸分子。

【0334】

51. 前記CasYポリペプチドがNLS配列に融合されている、47~50のいずれか1つに記載の1つ以上の核酸分子。

52. 前記1つ以上の核酸分子が、プロモーターに機能的に連結される前記CasYガイドRNAをコードするヌクレオチド配列を含む、47~51のいずれか1つに記載の1つ以上の核酸分子。 40

53. 前記1つ以上の核酸分子が、プロモーターに機能的に連結される前記CasYポリペプチドをコードするヌクレオチド配列を含む、47~52のいずれか1つに記載の1つ以上の核酸分子。

54. 前記プロモーターが、前記CasYガイドRNAをコードする前記ヌクレオチド配列に機能的に連結される、及び/または前記プロモーターが、前記CasYポリペプチドをコードする前記ヌクレオチド配列に機能的に連結され、真核生物において機能的である、52または53に記載の1つ以上の核酸分子。

55. 前記プロモーターが、植物細胞、真菌細胞、動物細胞、脊椎動物の細胞、ハエ細胞、脊椎動物の細胞、哺乳動物細胞、霊長類細胞、非ヒト霊長類細胞、及びヒト細胞のうち1つ以上において機能的である、54に記載の1つ以上の核酸分子。 50

【0335】

56. 前記プロモーターが、構成的プロモーター、誘導性プロモーター、細胞型特異的プロモーター、及び組織特異的プロモーターのうち1つ以上である、53～55のいずれか1つに記載の1つ以上の核酸分子。

57. 前記1つ以上の核酸分子が、1つ以上の組換え発現ベクターである、47～56のいずれか1つに記載の1つ以上の核酸分子。

58. 前記1つ以上の組換え発現ベクターが、1つ以上のアデノ随伴ウイルスベクター、1つ以上の組換えレトロウイルスベクター、または1つ以上の組換えレンチウイルスベクターから選択される、57に記載の1つ以上の核酸分子。

59. 前記プロモーターが、原核細胞において機能的である、53に記載の1つ以上の核酸分子。 10

60. a) CasYポリペプチド、または前記CasYポリペプチドをコードする核酸分子、

b) CasY融合ポリペプチド、または前記CasY融合ポリペプチドをコードする核酸分子、及び

c) CasYガイドRNA、または前記CasYガイドRNAをコードする核酸分子のうち1つ以上を含む真核細胞。

【0336】

61. 前記CasYポリペプチドをコードする前記核酸分子を含み、前記核酸分子が前記細胞のゲノムDNAに組み込まれている、60に記載の真核細胞。 20

62. 前記真核細胞が、植物細胞、哺乳動物細胞、昆虫細胞、クモ類細胞、真菌細胞、鳥類細胞、爬虫類細胞、両生類細胞、無脊椎動物細胞、マウス細胞、ラット細胞、霊長類細胞、非ヒト霊長類細胞、またはヒト細胞である、60または61に記載の真核細胞。

63. CasY融合ポリペプチド、または前記CasY融合ポリペプチドをコードする核酸分子を含む、細胞。

64. 前記細胞が原核細胞である、63に記載の細胞。

65. 前記CasY融合ポリペプチドをコードする前記核酸分子を含み、前記核酸分子が前記細胞のゲノムDNAに組み込まれている、63または64に記載の細胞。

【0337】

66. 標的核酸を修飾する方法であって、 30

前記標的核酸を

a) CasYポリペプチド、及び

b) 前記標的核酸の標的配列にハイブリダイズするガイド配列を含むCasYガイドRNAと接触させることを含み、

前記接触が、前記CasYポリペプチドによる前記標的核酸の修飾をもたらす、方法。

67. 前記修飾が、前記標的核酸の切断である、66に記載の方法。

68. 前記標的核酸が、二本鎖DNA、一本鎖DNA、RNA、ゲノムDNA、及び染色体外DNAから選択される、66または67に記載の方法。

69. 前記接触が、細胞外の*in vitro* (インビトロ)で行われる、66～68のいずれかに記載の方法。 40

70. 前記接触が、培養物中の細胞内で行われる、66～68のいずれかに記載の方法。

【0338】

71. 前記接触が、*in vivo* (インビボ)の細胞内で行われる、66～68のいずれかに記載の方法。

72. 前記細胞が真核細胞である、70または71に記載の方法。

73. 前記細胞が、植物細胞、真菌細胞、哺乳動物細胞、爬虫類細胞、昆虫細胞、鳥類細胞、魚類細胞、寄生生物細胞、節足動物細胞、無脊椎動物の細胞、脊椎動物の細胞、齧歯類細胞、マウス細胞、ラット細胞、霊長類細胞、非ヒト霊長類細胞、及びヒト細胞から選択される、72に記載の方法。 50

74. 前記細胞が原核細胞である、70または71に記載の方法。

75. 前記接触の結果、ゲノムが編集される、66~74のいずれか1つに記載の方法。

【0339】

76. 前記接触が、(a)前記CasYポリペプチド、または前記CasYポリペプチドをコードする核酸分子、及び(b)前記CasYガイドRNA、または前記CasYガイドRNAをコードする核酸分子を細胞に導入することを含む、66~75のいずれか1つに記載の方法。

77. 前記接触が、DNAドナー鋳型を前記細胞に導入することをさらに含む、76に記載の方法。

78. 前記CasYガイドRNAが、配列番号11~15のいずれか1つに記載されるcrRNA配列と80%以上の同一性を有するヌクレオチド配列を含む、66~77のいずれか1つに記載の方法。

79. 前記CasYポリペプチドがNLS配列に融合されている、66~78のいずれか1つに記載の方法。

80. 標的DNAからの転写を調節する、標的核酸を修飾する、または標的核酸と会合するタンパク質を修飾する方法であって、

前記標的核酸を

a)異種ポリペプチドに融合されたCasYポリペプチドを含む、CasY融合ポリペプチド、及び

b)前記標的核酸の標的配列にハイブリダイズするガイド配列を含むCasYガイドRNAと接触させることを含む、方法。

【0340】

81. 前記CasYガイドRNAが、配列番号11~15のいずれか1つに記載されるcrRNA配列と80%以上の同一性を有するヌクレオチド配列を含む、80に記載の方法。

82. 前記CasY融合ポリペプチドがNLS配列を含む、80または81に記載の方法。

83. 前記修飾が、前記標的核酸の切断ではない、80~82のいずれかに記載の方法。

84. 前記標的核酸が、二本鎖DNA、一本鎖DNA、RNA、ゲノムDNA、及び染色体外DNAから選択される、80~83のいずれかに記載の方法。

85. 前記接触が、細胞外の*in vitro*(インビトロ)で行われる、80~84のいずれかに記載の方法。

【0341】

86. 前記接触が、培養物中の細胞内で行われる、80~84のいずれかに記載の方法。

87. 前記接触が、*in vivo*(インビボ)の細胞内で行われる、80~84のいずれかに記載の方法。

88. 前記細胞が真核細胞である、86または87に記載の方法。

89. 前記細胞が、植物細胞、真菌細胞、哺乳動物細胞、爬虫類細胞、昆虫細胞、鳥類細胞、魚類細胞、寄生生物細胞、節足動物細胞、無脊椎動物の細胞、脊椎動物の細胞、齧歯類細胞、マウス細胞、ラット細胞、霊長類細胞、非ヒト霊長類細胞、及びヒト細胞から選択される、88に記載の方法。

90. 前記細胞が原核細胞である、86または87に記載の方法。

【0342】

91. 前記接触が、(a)前記CasY融合ポリペプチド、または前記CasY融合ポリペプチドをコードする核酸分子、及び(b)前記CasYガイドRNA、または前記CasYガイドRNAをコードする核酸分子を細胞に導入することを含む、80~90のいずれか1つに記載の方法。

10

20

30

40

50

92. 前記 CasY ポリペプチドが、触媒不活性な CasY ポリペプチド (dCasY) である、80~91のいずれか1つに記載の方法。

93. 前記 CasY ポリペプチドが、配列番号1の D672、E769、及び D935 から選択されるものに対応する位置に1つ以上の変異を含む、80~92のいずれか1つに記載の方法。

94. 前記異種ポリペプチドが、標的 DNA を修飾する酵素活性を示す、80~93のいずれか1つに記載の方法。

95. 前記異種ポリペプチドが、ヌクレアーゼ活性、メチルトランスフェラーゼ活性、デメチラーゼ活性、DNA 修復活性、DNA 損傷活性、脱アミノ化活性、ジスムターゼ活性、アルキル化活性、脱プリン活性、酸化活性、ピリミジン二量体を形成する活性、インテグラーゼ活性、トランスポザーゼ活性、リコンビナーゼ活性、ポリメラーゼ活性、リガーゼ活性、ヘリカーゼ活性、フォトリアーゼ活性、及びグリコシラーゼ活性から選択される1つ以上の酵素活性を示す、94に記載の方法。

【0343】

96. 前記異種ポリペプチドが、ヌクレアーゼ活性、メチルトランスフェラーゼ活性、デメチラーゼ活性、脱アミノ化活性、脱プリン活性、インテグラーゼ活性、トランスポザーゼ活性、及びリコンビナーゼ活性から選択される1つ以上の酵素活性を示す、95に記載の方法。

97. 前記異種ポリペプチドが、標的核酸と会合する標的ポリペプチドを修飾する酵素活性を示す、80~93のいずれか1つに記載の方法。

98. 前記異種ポリペプチドが、ヒストン修飾活性を示す、97に記載の方法。

99. 前記異種ポリペプチドが、メチルトランスフェラーゼ活性、デメチラーゼ活性、アセチルトランスフェラーゼ活性、デアセチラーゼ活性、キナーゼ活性、ホスファターゼ活性、ユビキチンリガーゼ活性、脱ユビキチン化活性、アデニル化活性、脱アデニル化活性、SUMO 化活性、脱SUMO 化活性、リボシル化活性、脱リボシル化活性、ミリスチル化活性、脱ミリスチル化活性、グリコシル化活性 (例えば、O-GlcNAc トランスフェラーゼによる)、及び脱グリコシル化活性から選択される1つ以上の酵素活性を示す、97または98に記載の方法。

100. 前記異種ポリペプチドが、メチルトランスフェラーゼ活性、デメチラーゼ活性、アセチルトランスフェラーゼ活性、及びデアセチラーゼ活性から選択される1つ以上の酵素活性を示す、99に記載の方法。

【0344】

101. 前記異種ポリペプチドが、転写を増加させるかまたは減少させるタンパク質である、80~93のいずれか1つに記載の方法。

102. 前記異種ポリペプチドが転写抑制因子ドメインである、101に記載の方法。

103. 前記異種ポリペプチドが転写活性化ドメインである、101に記載の方法。

104. 前記異種ポリペプチドがタンパク質結合ドメインである、80~93のいずれか1つに記載の方法。

- 105. a) CasY ポリペプチド、
- b) CasY 融合ポリペプチド、及び
- c) CasY ガイド RNA

のうち1つ以上をコードするヌクレオチド配列を含む導入遺伝子をゲノムに含む、トランスジェニック多細胞非ヒト生物。

【0345】

106. 前記 CasY ポリペプチドが、配列番号1または配列番号2に記載されるアミノ酸配列 (または配列番号1~8のいずれか1つに記載されるアミノ酸配列) に対して、50%以上のアミノ酸配列同一性を有するアミノ酸配列を含む、105に記載のトランスジェニック多細胞非ヒト生物。

107. 前記 CasY ポリペプチドが、配列番号1または配列番号2に記載されるアミノ酸配列 (または配列番号1~8のいずれか1つに記載されるアミノ酸配列) に対して、

10

20

30

40

50

85%以上のアミノ酸配列同一性を有するアミノ酸配列を含む、105に記載のトランスジェニック多細胞非ヒト生物。

108. 前記生物が、植物、単子葉植物、双子葉植物、無脊椎動物、昆虫、節足動物、クモ類、寄生生物、蠕虫、刺胞動物、脊椎動物、魚類、爬虫類、両生類、有蹄動物、鳥類、ブタ、ウマ、ヒツジ、齧歯類、マウス、ラット、または非ヒト霊長類である、105~107のいずれか1つに記載のトランスジェニック多細胞非ヒト生物。

109. a) CasYポリペプチド及びCasYガイドRNA、

b) CasYポリペプチド、CasYガイドRNA、及びDNAドナー鑄型、

c) CasY融合ポリペプチド及びCasYガイドRNA、

d) CasY融合ポリペプチド、CasYガイドRNA、及びDNAドナー鑄型、

e) CasYポリペプチドをコードするmRNA、及びCasYガイドRNA、

f) CasYポリペプチドをコードするmRNA、CasYガイドRNA、及びDNAドナー鑄型、

g) CasY融合ポリペプチドをコードするmRNA、及びCasYガイドRNA、

h) CasY融合ポリペプチドをコードするmRNA、CasYガイドRNA、及びDNAドナー鑄型、

i) i) CasYポリペプチドをコードするヌクレオチド配列、及びii) CasYガイドRNAをコードするヌクレオチド配列を含む1つ以上の組換え発現ベクター、

j) i) CasYポリペプチドをコードするヌクレオチド配列、ii) CasYガイドRNAをコードするヌクレオチド配列、及びiii) DNAドナー鑄型を含む1つ以上の組換え発現ベクター、

k) i) CasY融合ポリペプチドをコードするヌクレオチド配列、及びii) CasYガイドRNAをコードするヌクレオチド配列を含む1つ以上の組換え発現ベクター、ならびに

l) i) CasY融合ポリペプチドをコードするヌクレオチド配列、ii) CasYガイドRNAをコードするヌクレオチド配列、及びDNAドナー鑄型を含む1つ以上の組換え発現ベクターを含む、CasYシステム。

110. 前記CasYポリペプチドが、配列番号1または配列番号2に記載されるアミノ酸配列(または配列番号1~8のいずれか1つに記載されるアミノ酸配列)に対して、50%以上のアミノ酸配列同一性を有するアミノ酸配列を含む、109に記載のCasYシステム。

【0346】

111. 前記CasYポリペプチドが、配列番号1または配列番号2に記載されるアミノ酸配列(または配列番号1~8のいずれか1つに記載されるアミノ酸配列)に対して、85%以上のアミノ酸配列同一性を有するアミノ酸配列を含む、109に記載のCasYシステム。

112. 前記ドナー鑄型核酸が、8ヌクレオチド~1000ヌクレオチドの長さを有する、109~111のいずれかに記載のCasYシステム。

113. 前記ドナー鑄型核酸が、25ヌクレオチド~500ヌクレオチドの長さを有する、109~111のいずれかに記載のCasYシステム。

114. 109~113のいずれか1つに記載のCasYシステムを含むキット。

115. 前記キットの構成要素が同じ容器内にある、114に記載のキット。

【0347】

116. 前記キットの構成要素が別の容器内にある、114に記載のキット。

117. 109~116のいずれか1つに記載のCasYシステムを含む滅菌容器。

118. 前記容器が注射器である、117に記載の滅菌容器。

119. 109~116のいずれか1つに記載のCasYシステムを含む埋込型装置。

120. 前記CasYシステムがマトリックス内にある、119に記載の埋込型装置。

【0348】

121. 前記CasYシステムがリザーバー内にある、119に記載の埋込型装置。

122. CRISPR RNA誘導型エンドヌクレアーゼを同定する方法であって、複数のメタゲノムヌクレオチド配列において、Cas1ポリペプチドをコードするヌクレオチド配列を検出すること、

前記Cas1をコードするヌクレオチド配列の近傍にあるCRISPR配列を検出すること、

検出されたCRISPR配列を含むCRISPR遺伝子座を、前記複数のメタゲノムヌクレオチド配列が由来する核酸試料から発現ベクターにクローニングして、組換えCRISPR遺伝子座発現ベクターを生成すること、

前記組換えCRISPR遺伝子座発現ベクターが標的核酸を切断する能力についてアッセイすること（標的核酸を切断する能力を有するCRISPR遺伝子座は、CRISPR RNA誘導型エンドヌクレアーゼをコードするヌクレオチド配列を含む）、及び

前記CRISPR遺伝子座内において、既知のCRISPR RNA誘導型エンドヌクレアーゼポリペプチドのアミノ酸配列に対して20%未満のアミノ酸配列同一性を有するポリペプチドをコードするオープンリーディングフレームを同定することを含む、方法。

123. 前記アッセイすることが、前記組換えCRISPR遺伝子座発現ベクター及び標的核酸を細胞に導入することを含む、122に記載の方法。

【実施例】

【0349】

以下の実施例は、本発明の作製方法及び使用方法の完全な開示及び説明を当業者に提供するために記載するものであり、発明者が発明とみなすものの範囲を限定することを意図せず、また以下の実験が、すべてまたは唯一の実施された実験であることを意味することも意図しない。使用した数字（例えば、量、温度など）に関しては正確性を確保するよう努めたが、一部の実験誤差及び偏差が考慮されるべきである。特に記載のない限り、部は重量部であり、分子量は重量平均分子量であり、温度は摂氏度であり、圧力は大気圧または大気圧付近である。例えば、bp、塩基対（複数可）；kb、キロベース（複数可）；pl、ピコリットル（複数可）；sまたはsec、秒（複数可）；min、分（複数可）；hまたはhr、時間（複数可）；aa、アミノ酸（複数可）；kb、キロベース（複数可）；bp、塩基対（複数可）；nt、ヌクレオチド（複数可）；i.m.、筋肉内（筋肉内に）；i.p.、腹腔内（腹腔内に）；s.c.、皮下（皮下に）などの標準的な略語を使用する場合がある。

【0350】

実施例1

本明細書に記載の研究には、地下水、堆積物、及び酸性鉱山排水からの微生物群集のメタゲノム試料の分析を含む。培養生物には存在しない新たなクラス2のCRISPR-Casシステムを同定した。

【0351】

図3. CasYドメイン及び類似性検索。（パネルA）HHpredを使用した、Accp1との遠隔ホモログアライメントから推定したCasYドメインの模式図。保存される触媒残基を、タンパク質の上部に赤いバーでマークしている。CasYは、C末端領域のRuvC分割ドメイン（RuvC-I、RuvC-II、及びRuvC-III）、及び大きい新規N末端ドメインで構成される。模式図の下に、以下の検索に基づく上位ヒットを示す：（1）NCBI（モデル及び環境タンパク質を含むNRデータベース）の全タンパク質に対するBLAST検索。（2）Makarova et al. Nat Rev Microbiol. 2015 Nov; 13(11): 722-36、及びShmakov et al. Mol Cell. 2015 Nov 5; 60(3): 385-97）に記載されるCasタンパク質すべてを使用して構築されたモデルに基づく、プロファイル隠れマルコフモデル（HMM）の検索。（3）HHpredに基づく遠隔ホモログ検索。ヒットは、その有意性に基づいて色分けし、ヒット範囲及びE値を示している。注目すべき点として、CasYは局所的なヒットのみであった。CasYの812 N末端アミノ酸は、極めてマイナーな部分ヒットを1つだけ有した。総合すると、これら

の知見は、CasYが新規のCasタンパク質であることを示している。(パネルB)異なるCasYを含むCRISPR遺伝子座骨格を配列データから構築した。

【0352】

実施例2

図4. CasY及びC2c3遺伝子座の模式図。干渉タンパク質を緑で、取り込みタンパク質を赤で示している。RNA構造を利用して折り畳まれるリピートを右側に示す。5'末端に強度のヘアピンが表れており、CasYによるCRISPR配列の自己プロセッシングを示唆している。

【0353】

図5(パネルA~D) CasYによるPAM依存性プラスミド干渉。(パネルA) CasYを用いてPAM欠失アッセイを実施した。CasY CRISPR遺伝子座を含むE.coliを、標的配列の5'または3'をランダム抽出した7ヌクレオチドのプラスミドライブラリで形質転換した。標的プラスミドを選別して、形質転換体をプールした。ランダムな領域を増幅して調製し、ディープシーケンシングした。欠失配列を同定し、これを使用してPAMのロゴを生成した。(パネルB)生成したCasY.1のPAMロゴは、標的の5'隣接配列5'-TA-3'を含む配列に対して高い選択性を示した。3'PAMは検出されなかった。(パネルC)4種類のPAMを直接アッセイして、PAM欠失アッセイから決定されるPAMを確認した。(パネルD)生成したCasY.2のPAMロゴは、標的の5'隣接配列5'-YR-3'及び/または5'-TR-3'(例えば、5'-DTR-3')(それぞれ下限閾値及び上限閾値)を含む配列に対して選択性を示した(ここで、YはTまたはCであり、RはAまたはGであり、DはA、G、またはTである)。3'PAMは検出されなかった。

【0354】

図6.(パネルA)天然のCasYガイドRNAからの「リピート」配列(CasY遺伝子座Y1~Y6に対応)。(パネルB)DNA切断を誘導するCasY RNAの図。CasYタンパク質はリピート領域のcrRNA(CasYガイドRNA)に結合する(黒、リピート;赤、スペーサー)。ガイドRNAのガイド配列と、正しいプロトスペーサー隣接モチーフ(PAM)を含む標的配列(青)との塩基対形成により、標的DNAの二本鎖切断が生じる。

【0355】

実施例3:未培養微生物からの新規CRISPR-Casシステム

CRISPR-Cas適応免疫系は、部位特異的DNA切断が可能であるプログラム可能な酵素を提供することにより、ゲノム工学に変革をもたらした。しかしながら、現在のCRISPR-Cas技術は、培養細菌に由来するシステムのみに基づいているため、単離されていない生物に由来する大部分の酵素は未開発のままである。本明細書に示すデータは、古細菌の生命ドメインにおいて初めて報告されたCas9を含む新たなCRISPR-Casシステムが、培養に依存しないゲノム解決型のメタゲノミクスを使用して同定されることを示している。この異なったCas9酵素は、現行のCRISPR-Casシステムの一部としてはほとんど研究されていないナノ古細菌において見出された。細菌には、以前は未知であった2つのシステム、CRISPR-CasX及びCRISPR-CasYが発見された。これは、まだ解明されていない最も合理的なシステムに属する。注目すべきは、必要なすべての機能構成要素が、メタゲノミクスによって同定され、その結果、E.coliにおける頑強なRNA誘導型DNA干渉活性の検証を可能にしたことである。本明細書のデータは、生きた細胞での実験と併せて、環境微生物群集を調べることによって、ゲノムの従来にない多様性を知ることが可能であり、その内容が、微生物を用いたバイオテクノロジーの守備範囲を広げるであろうことを示している。

【0356】

結果

地下水、堆積物、及び酸性鉱山排水の微生物群集から得たテラベース規模のメタゲノムデータセットを分析し、培養生物には存在しない新たなクラス2のCRISPR-Cas

10

20

30

40

50

システムを探索した。最初に古細菌ドメインの Cas9 タンパク質を同定し、非培養細菌において、2つの新たな CRISPR-Cas システム、CRISPR-Cas X 及び CRISPR-Cas Y を発見した (図 7)。注目すべきは、古細菌 Cas9 と Cas Y の両方が、既知の単離標本がない系統からの生物のゲノムにおいて排他的にコードされていることであった。

【0357】

古細菌 Cas9 の最初の同定

CRISPR-Cas9 の特徴の一つは、細菌ドメインにのみ存在が推定されることであった。したがって、酸鉱山排水 (AMD) のメタゲノムデータセットにおいて、ナノ古細菌 ARMAN-1 (*Candidatus Micrarchaeum acidiphilum* ARMAN-1) 及び ARMAN-4 (*Candidatus Parvarchaeum* ARMAN-4) のゲノムにコードされた Cas9 タンパク質が発見されたことは驚くべきことであった。これらの知見から、Cas9 を含む CRISPR システムが別の生命ドメインでも発生することがわかる。

10

【0358】

ARMAN-4 cas9 遺伝子は、ゲノム内容が同じであるが、(25 kbp 超のいくつかの DNA 配列コンティグの中央部に位置しているにもかかわらず) 他の隣接する cas 遺伝子がなく、隣接する CRISPR リピート-スペーサー単位を1つだけもつ16種類の試料に見出された (図 13)。一般的な CRISPR 配列、及び一般的な CRISPR インテグラーゼをコードする cas1 の欠失は、新しいスペーサーを取り込む能力をもたないシステムであることを示す。スペーサー配列に対する標的は同定できなかったが、数年にわたって採取された試料の遺伝子座が保全されているとすれば、「単一標的」の CRISPR-Cas システムでの機能をこの時点で排除することはできない。

20

【0359】

逆に、15種類の試料から回収された ARMAN-1 における CRISPR-Cas 遺伝子座は、cas1、cas2、cas4、及び cas9 遺伝子に隣接する大きな CRISPR 配列を含む。大部分が保存された末端 (おそらく最も古いスペーサーからなる) 及び多くの異なったスペーサーが組み込まれた可変領域をもつ、多数の代替 ARMAN-1 CRISPR 配列を再構成した (図 8 A 及び図 14)。スペーサー含量のこの超可変性に基づくと、これらのデータは、ARMAN-1 CRISPR-Cas9 システムが試料

30

【0360】

注目すべきことに、ARMAN-1 CRISPR-Cas9 システムの56の推定スペーサー標的 (プロトスペーサー) は、単一の10 kbp のゲノム断片上に位置した。高密度の短い仮想タンパク質をコードしていることを考慮すると、これは ARMAN-1 ウイルスである可能性が高い (図 8 B)。実際、クライオ電子断層撮影法による再構成では、ARMAN 細胞に結合したウイルス粒子が特定されることが多かった。ARMAN-1 プロトスペーサーはまた、ARMAN-2 (別のナノ古細菌) のゲノム内の推定トランスポゾン、及び同じ生態系の I-プラズマのものを含む、*Thermoplasma talle* 古細菌のゲノム内の推定流動要素に由来していた (図 15)。ARMAN と *Thermoplasma talle* 細胞との間に直接的な原形質「架橋」が観察され、両者の密接な関係を示唆した。したがって、ARMAN-1 CRISPR-Cas9 は、これらの生物間でのトランスポゾンの増殖を抑制することができる。これは、真核生物の生殖系列での転移に対する piRNA 介在性の防御を連想させる役割である。

40

【0361】

現行の DNA 標的 CRISPR-Cas システムは、自己と非自己を区別するために、標的配列に隣接する 2~4 bp のプロトスペーサー隣接モチーフ (PAM) を使用する。ゲノム標的配列に隣接する配列を調べると、実際に ARMAN-1 において「NGG」PAM の高い選択性が明らかであった (図 8 C)。Cas9 はまた、RNA 誘導型 DNA 切断のために、2つの個別の転写物、CRISPR RNA (crRNA) 及びトランス活

50

性化CRISPR RNA (tracrRNA)を用いる。ARMAN-1とARMAN-4両方のCRISPR-Cas9システムの近傍に推定tracrRNAが同定された(図16)。crRNA-tracrRNA誘導複合体の成熟を担う宿主因子、RNase IIIが欠損しているため、II型CRISPRシステムは古細菌に存在しないことが以前に示唆されていた。注目すべきことに、ARMAN-1ゲノム(95%完全と推定)に同定されたRNase IIIホモログはなく、CRISPR配列に予測される内部プロモーターもなく、今まで未確認のガイドRNAの産生メカニズムであることが示唆された。E. coli及び酵母の両方から精製したARMAN-1及びARMAN-4 Cas9タンパク質の切断活性を試験する生化学的実験、及びin vivo E. coli標的アッセイでは、検出可能な活性はまったく示されなかった(図21及び図17を参照)

10

【0362】

CRISPR-CasXは、新たな二重RNA誘導型CRISPRシステムであるCas9に加えて、発見され、実験的に検証されているクラス2のCasエフェクタータンパク質のファミリーは、Cpf1、C2c1、及びC2c2の3つのみである。小さいDNA断片にのみ同定された別の遺伝子c2c3もまた、このようなタンパク質ファミリーをコードすることが示唆されている。新たな種類のクラス2のCRISPR-Casシステムが、地下水試料及び堆積物試料から繰り返し回収した2種類の細菌ゲノムに見出された。属する種族が異なる2つの生物、Deltaproteobacteria及びPlanctomycetesにおけるこのシステムの高度な保存性は、最近の近縁種間転写を示唆している。この新たに記載されるシステムは、Cas1、Cas2、Cas4、及び本明細書でCasXと称する特性が未知の約980aaのタンパク質を含む。各CasXと関連するCRISPR配列は、酷似した37塩基対のリピート、33~34塩基対のスペーサー、及びCasオペロンとCRISPR配列との間の推定tracrRNAを有していた(図7B)。BLAST検索によると、トランスポザーゼに対して弱い類似性(e値 $> 1 \times 10^{-4}$)のみを示し、類似性はCasXC末端の特定領域に限定された。遠隔相同性の検出及びタンパク質モデリングでは、V型CRISPR-Casシステムに見られるものを連想させる組織をもつCasXC末端近傍のRuvCDドメインが同定された(図18)。CasXTタンパク質の残部(630個のN末端アミノ酸)は、どの既知のタンパク質とも検出可能な類似性を示さず、これが新たなクラス2のエフェクターであることを示唆した。tracrRNAとCas1、Cas2、及びCas4の個々のタンパク質との組み合わせは、V型システムの中でも独特である。さらに、CasXは、任意の既知のV型タンパク質よりもかなり小さく、Cpf1、C2c1、及びC2c3の通常サイズ1200aa超と比較して980aaである。

20

30

【0363】

次なる疑問点は、その小さなサイズと非標準遺伝子座の含量にもかかわらず、CasXがCas9及びCpf1酵素と類似するRNA誘導型DNA標的化が可能であるかどうかということであった。この可能性を試験するために、casX、短いリピート-スペーサー配列、及び間にあるノンコーディング領域を含む最小のCRISPR-CasX遺伝子座をコードするプラスミドを合成した。E. coliで発現させると、この最小遺伝子座は、メタゲノム解析で同定された標的配列をもつプラスミドによる形質転換を遮断した(図9A~C、図19)。さらに、ミニ遺伝子座におけるスペーサー配列がプラスミド標的のプロトスペーサー配列と一致した場合にのみ、形質転換の干渉が起こった。CasXのPAM配列を同定するために、標的部位に隣接する5'または3'いずれかのランダム配列を含むプラスミドを使用して、E. coliにおいて形質転換アッセイを繰り返した。この分析により、プロトスペーサー配列のすぐ5'側に位置する配列「TTCN」に対して、厳密な選択性が示された(図9D)。3'のPAM選択性は観察されなかった(図19)。この知見と一致して、「TTCA」は、環境試料で同定された推定Deltaproteobacteria CRISPR-CasXのプロトスペーサー上流に見られる配列であった。注目すべきは、CasXTタンパク質の高度な相同性と同様、CRISPR-

40

50

Cas Xの両方の遺伝子座が同じPAM配列を共有していることである。

【0364】

V型CRISPR遺伝子座には、一本鎖RNA及び二本鎖RNA誘導型システムいずれの例も存在する。環境メタトランスクリプトームデータを使用して、Cas XがDNAの標的化活性にtracrRNAを必要とするかどうかを決定した。この分析により、Cas 2オープンリーディングフレームとCRISPR配列との間にコードされたCRISPRリピートと相補的な配列をもつノンコーディングRNA転写物が明らかになった(図10A)。さらにトランスクリプトームマッピングは、CRISPR RNA(crRNA)が、CRISPR-Cas 9システムで起こるcrRNAプロセッシングと同様に、22 ntのリピート及び20 ntの隣接スペーサーを含むようにプロセッシングされることを示唆している(図10A)。また、crRNA-tracrRNA二重鎖のRNAs III介在性のプロセッシングと一致する、2 ntの3'オーバーハングが同定された(図10B)。推定tracrRNAに対するCas X活性の依存性を決定するために、上述した最小のCRISPR-Cas X遺伝子座からこの領域を欠失させ、プラスミド干渉アッセイを繰り返した。推定tracrRNAコード配列をCas Xプラスミドから欠失させると、存在時に観察された頑強な形質転換干渉が無効化された(図10C)。これらの結果を総合すると、Cas Xは、新たな機能性DNA標的化二本鎖RNA誘導型CRISPR酵素であると確定される。

10

【0365】

CRISPR-Cas Y、分離株を欠く細菌系統でのみ見出されたシステム

20

特定の候補門放散(candidate phyla radiation: CPR)細菌のゲノムにコードされた別の新たなクラス2のCasタンパク質を同定した。これらの細菌は通常、細胞の大きさが小さく(クライオTEMデータ及び濾過による濃縮に基づく)、ゲノムが極小であり、生合成能力が限られている。このことは、これらの細菌がほぼ間違いなく共生生物であることを示している。本明細書でCas Yと称する、新たな約1200 aaのCasタンパク質は、Cas 1及びCRISPR配列程度しか含まない最小のCRISPR-Casシステムに属すると思われる(図11A)。CRISPR配列のほとんどは、17~19 ntの異常に短いスペーサーを有するが、Cas 1を欠く、あるシステム(Cas Y. 5)はそれよりも長いスペーサー(27~29 nt)を有している。同定されたCas Yタンパク質の6つの例は、公開データベースのどのタンパク質とも有意な配列類似性を有していなかった。公開されたCasタンパク質3、4から構築されたプロファイルモデル(HMM)を使用した高感度検索は、6つのCas Yタンパク質のうち4つが、RuvCドメインと重複するC末端領域及びN末端の小領域(約45 aa)に、C2c3に対する局所類似性(e値 4×10^{-11} ~ 3×10^{-18})を有することを示した(図18参照)。C2c3は、分類学的関係のない短いコンティグに同定された推定V型Casエフェクターであり、実験的には確認されていない。Cas Y同様、C2c3は、短いスペーサーとCas 1をもつ配列の隣に見られたが、他のCasタンパク質をもつ配列には見られない。注目すべきは、現在の研究で同定されたCas Yタンパク質のうち2つが、他のCas Yタンパク質と有意な配列類似性(Blastの最適ヒット: e値 6×10^{-85} 、 7×10^{-75})を共有しているにもかかわらず、C2c3に対して有意な類似性を有さなかったことである。

30

40

【0366】

実験的に検証されたどのCRISPR遺伝子座に対してもCRISPR-Cas Yの相同性が低いことを考慮すると、次なる疑問点は、このシステムがRNA誘導型DNA干渉を付与するかどうかであったが、スペーサー長が短いため、そのような活性に必要とされる可能なPAMモチーフに関して信頼性の高い情報が存在しなかった。これに対処するために、CRISPR-Cas Y. 1遺伝子座全体を、短縮型CRISPR配列を用いて合成し、プラスミドベクターでE. coliに導入した。次に、形質転換アッセイにおいて、配列内のスペーサー配列と一致し、隣接するランダムな5'または3'領域を含む配列をもつ標的プラスミドを使用してこの細胞を攻撃し、可能性のあるPAMを同定した。形

50

質転換体の分析では、標的配列にすぐ隣接する5' T Aを含む配列の欠失が明らかになった(図11B)。この同定されたP A M配列を使用して、C a s Y . 1 遺伝子座を単一のP A Mを含むプラスミドに対して試験した。同定された5' T A P A M配列を含む標的が存在する場合のみプラスミド干渉を示した(図11C)。したがって、これらのデータは、C R I S P R - C a s YがD N A干渉活性を有することを示している。

【0367】

考察

未培養細菌及び古細菌由来のゲノムにおける新たなクラス2のC R I S P R - C a s 適応免疫系を同定し、特性決定した。現行のC R I S P R 遺伝子座に共通する、C a s 1の進化分析(図12A)は、本明細書に記載の古細菌C a s 9システムが明確にはどの既存のI I型サブタイプにも分類されないことを示唆した。C a s 1の系統(ならびにc a s 4の存在)は、I I - B型のシステムと一緒に群化されるが、C a s 9の配列は、I I - C型タンパク質との類似性が高かった(図20)。したがって、古細菌I I型システムはI I - C型及びI I - B型システムの融合体として生じた可能性がある(図12B)。同様に、C a s 1系統発生分析では、C R I S P R - C a s XシステムのC a s 1は、他のどの既知のV型システムからも遠いことが示された。V型システムは、祖先I型システムからの適応モジュールを有するトランスポゾン融合(C a s 1 - C a s 2)の結果であると示唆されている。したがって、C R I S P R - C a s Xシステムは、前述したV型システムに生じたものとは異なる融合事象の後に生じると仮定される。驚くべきことに、C R I S P R - C a s Yと推定C 2 c 3システムの両方が、C R I S P R 遺伝子座へのD N Aの組み込みに不可欠であると考えられるタンパク質、C a s 2を欠いていると思われる。すべてのC R I S P R - C a s システムが、C a s 1とC a s 2の両方を含んだ祖先I型システムの子孫であると考えられていることを考慮すると、C R I S P R - C a s Y及びC 2 c 3システムはいずれも、それ以外のC R I S P R - C a s システムとは異なる祖先を有するか、あるいは両システムの進化の歴史の過程でC a s 2が失われた可能性がある。

【0368】

本明細書に記載する、古細菌でのC a s 9の発見、及び以前は未知であった2つのC R I S P R - C a s システムの細菌での発見には、複雑な天然微生物群集から得た網羅的なD N A及びR N A配列データセットを使用した。C a s X及びC a s Yの場合、未アセンブルの配列情報から明らかにされていない機能の予測には、ゲノム内容が不可欠であった。さらに、メタゲノムデータ支援型の機能試験の分析を通じて、推定t r a c r R N Aの同定、ならびに標的ウイルス配列を明らかにした。興味深いことに、これまでに同定された最もコンパクトなC R I S P R - C a s 遺伝子座のいくつかは、極小ゲノムを有する生物で発見された。ゲノムサイズが小さい結果、このような生物は基本的な代謝条件が群集の他のメンバーによって左右されるため、大部分が従来による培養による方法の範囲から除外されたままであった。干渉に必要とされるタンパク質の数が限定されるため、これらの最小システムは、新たなゲノム編集ツールの開発に特に有益である。重要なことに、C R I S P R - C a s システムに関するメタゲノムの発見は、i n s i l i c oでの観察に限定されるものではなく、その機能を試験することができる実験環境に導入可能であることが本明細書で示されている。今後、生命が存在する事実上すべての環境をゲノム解決型メタゲノム法によって実証できるとすれば、本明細書に記載の複合的な計算 - 実験的アプローチは、既知のC R I S P R - C a s システムの多様性を大いに拡大し、生物学的研究及び臨床応用のための新技術を提供するものと期待される。

【0369】

方法

メタゲノミクスとメタトランスクリプトーム

次の異なる3箇所から得たメタゲノム試料を分析した：(1)2006年~2010年にカリフォルニア州アイアンマウンテン、リッチモンド鉱山から採取した酸性鉱山排水(A M D)試料、(2)2007年~2013年にコロラド州ライフル近郊コロラド川沿い

10

20

30

40

50

の Rifle Integrated Field Research (IFRC) の敷地から採取した地下水及び堆積物試料、(3) 2009年及び2014年にユタ州コロラド高原のCO₂ 噴出冷間欠泉、Crystal Geysersから採取した地下水。

【0370】

AMDのデータに関しては、DNA抽出方法及びショートリード配列決定が、Deneff and Banfield (2012) 及び Miller et al. (2011) によって報告された。Rifleのデータに関しては、DNA及びRNAの抽出、ならびに配列決定、アセンブル、及び再構成されたゲノムが、Anantharaman et al. (2016) 及び Brown et al. (2015) によって記載された。Crystal Geysersからの試料に関しては、Probst et al (2016) 及び Emerson et al. (2015) によって記載された方法に従っている。簡潔には、PowerSoil DNA単離キット (MoBio Laboratories Inc., Carlsbad, CA, USA) を使用して試料からDNAを抽出した。Brown et al. (2015) によって記載されているように、6つの2011年のRifle地下水試料からRNAを回収して、0.2 µmフィルターから抽出した。Illumina HiSeq 2000プラットフォームでDNAを配列決定し、5500XL SOLiDプラットフォームでメタトランスクリプトームcDNAを配列決定した。Crystal Geysersのデータの新規報告、及びAMDのデータの再分析のために、IDBA-UDを用いて配列をアセンブルした。配列決定範囲及び遺伝子発現をそれぞれ決定するために使用されるDNA及びRNA (cDNA) のリードマッピングを、Bowtie 2を使用して実施した。Prodigalを使用してアセンブルした骨格に基づいてオープンリーディングフレーム (ORF) を予測した。ABAWACA、ABAWACA2 (<https://github.com/CK7>) Maxbin 2の組み合わせを使用して、差次的な重複度の存在量パターンに基づいて、Crystal Geysersのデータセットから得た骨格をピニングし、Emergent Self-Organizing Maps (ESOM) を使用してテトラヌクレオチド頻度をピニングした。ゲノムは、GC含量%、分類学的関係、及びゲノムの完全性を用いて手動でキュレーションした。骨格構築エラーを、ra2.py (<https://github.com/christophertbrown>) を使用して補正した。

10

20

30

【0371】

CRISPR - Cas 計算解析

種々の試料からアセンブルしたコンティグを、Makarova et al. 及び Shmakov et al. によるアライメントに基づいて、HMMer スイートを使用して構築された隠れマルコフモデル (HMM) プロファイルを用いてスキャンして既知のCasタンパク質を探索した。CrisprFinder ソフトウェアのローカルバージョンを使用して、CRISPR配列を同定した。cas1 遺伝子に隣接する10のORFのうち1つが、800 aaより大きい、特性が未知のタンパク質をコードしているかどうか、及び同じコンティグに既知のcas干渉遺伝子が同定されていないかどうか、Cas1及びCRISPR配列の両方を含む遺伝子座をさらに分析した。この大きなタンパク質をクラス2のCasエフェクター候補としてさらに分析した。エフェクター候補を、MCCLを使用して配列類似性に基づいたタンパク質ファミリーに群化した。これらのファミリーのそれぞれを代表するHMMを構築し、類似するCasタンパク質のメタゲノムデータセットの検索に使用することにより、これらのタンパク質ファミリーを拡張した。タンパク質ファミリーが事実上、新しいことを確認するために、BLASTを使用して、NCBIの非重複(nr) 及びメタゲノム(env_nr) タンパク質データベースに対して既知のホモログを検索し、ならびにUniProt Knowledge Baseに対してHMMを検索した。全長ヒット (タンパク質の長さの25%超) がないタンパク質のみを新規タンパク質とみなした。推定Casタンパク質の遠隔相同性検索を、HH-suiteのHHpredを使用して実施した。ハイスコアのHHpredヒットを使用して、決定された結晶構造及びJPred 4によって予測した二次構造との比較に基づいてドメ

40

50

インアーキテクチャを推定した。新たに発見された Cas タンパク質を含む、HMM データベースは、補足データ 1 に記載している。

【0372】

スペーサー配列は、CrisprFinder を使用してアセンブルしたデータから決定した。CRASS を使用して、関連する試料の短い DNA リードで追加のスペーサーの配置を特定した。次に、スペーサー標的 (プロトスペーサー) を、スペーサーに対する mismatches が 1 以下のヒットの関連するメタゲノムアセンブルに対する BLAST 検索 (「-task blastn-short」を使用) によって同定した。関連するリピートを含んだコンティグに属するヒットは除外した (CRISPR 配列をプロトスペーサーとして同定することを回避するため)。プロトスペーサーに隣接する領域をアライメントすることにより、プロトスペーサー隣接モチーフ (PAM) を同定し、WebLogo を使用して可視化した。RNA 構造は mFold を使用して予測した。アセンブルしたデータからのスペーサー、リピート、及び隣接配列を手動でアライメントすることにより、CRISPR 配列の多様性を分析した。手動アライメント及びコンティグの可視化は、Genious 9.1 を用いて実施した。

【0373】

新たに同定されたシステムの Cas 1 及び Cas 9 タンパク質の系統発生分析のため、Makarova et al. 及び Shmakov et al. によるタンパク質とともに使用した。CD-HIT を使用して、90% 以上同一性のあるタンパク質をまとめて群化することにより、非重複セットを編集した。MAFFT でアライメントを作成し、RAxML を使用して、代入モデル PROTGAMMALG 及び 100 ブートストラップサンプリングで最尤系統発生を構成した。casposon に至る分岐を用いて Cas 1 のツリーのルートを決定した。FigTree 1.4.1 (<http://tree.bio.ed.ac.uk/software/figtree/>) 及び iTOL v3 を使用して、ツリーを可視化した。

【0374】

異種プラスミドの作製

Cas X の取り込みに関連するタンパク質を除去し、Cas X と Cas Y 両方の CRISPR 配列のサイズを減少させることによって、メタゲノムコンティグから最小 CRISPR 干渉プラスミドを作製した。最小遺伝子座は Gblocks (Integrated DNA Technology) として合成し、Gibson Assembly を使用してアセンブルした。

【0375】

PAM 欠失アッセイ

以前の記載に改変を加えて PAM 欠失アッセイを実施した。7 nt のランダムな PAM 領域をもつ標的を含む DNA オリゴヌクレオチドをプライマーとアニーリングすることによって、ランダムな PAM 配列を含むプラスミドライブラリをアセンブルし、クレノウ断片 (NEB) で伸長させた。二本鎖 DNA を EcoRI 及び NcoI で消化して、pUC19 骨格にライゲーションした。ライゲーションしたライブラリを DH5 に形質転換して、 10^8 超の細胞を回収し、プラスミドを抽出し、精製した。CRISPR 遺伝子座を保有するエレクトロコンピテント E.coli、または遺伝子座をもたない対照プラスミドに、プールされたライブラリ 200 ng を形質転換した。カルベニシリン (100 mg L^{-1}) 及びクロラムフェニコール (30 mg L^{-1}) を含有する選択培地に形質転換細胞をプレーティングし、25 で 30 時間培養した。プラスミド DNA を抽出し、アダプターを付加して PAM 配列を増幅し、Illumina で配列決定した。7 nt の PAM 領域を抽出し、各 7 nt の配列について PAM 頻度を算出した。指定閾値を超える欠失のある PAM 配列を使用して、WebLogo を作成した。

【0376】

プラスミド干渉

メタゲノム配列解析または PAM 欠失アッセイから同定された推定標的を pUC19 プ

10

20

30

40

50

ラスミドにクローニングした。10 ngの標的プラスミドを、CRISPR遺伝子座プラスミドを含むエレクトロコンピテント*E. coli* (NEB安定)に形質転換した。細胞を25℃で2時間回復させ、適切な希釈液を選択培地にプレーティングした。プレートを25℃でインキュベートし、コロニー形成単位を計数した。プラスミド干渉実験はすべて3連で実施し、エレクトロコンピテント細胞は複製物ごとに個別に調製した。

【0377】

ARMAN - Cas9タンパク質の発現及び精製

ARMAN - 1 (AR1) 及びARMAN - 4 (AR4) から得たCas9の発現構築物を、*E. coli*用にコドン最適化されたgBlocks (Integrated DNA Technologies) からアセンブルした。アセンブルした遺伝子を、N末端His₆-MBPまたはHis₆融合タンパク質として、pET系の発現ベクターにクローニングした発現ベクターをBL21 (DE3) *E. coli*細胞に形質転換し、LBブロス中にて37℃で増殖させた。タンパク質を発現させるため、中間対数期の間、0.4 mMのIPTG (イソプロピル-β-D-1-チオガラクトピラノシド) で細胞を誘導し、16℃で一晩インキュベートした。以降の工程はすべて、4℃で実施した。細胞ペレットを溶解緩衝液 (50 mM Tris-HCl pH8、500 mM NaCl、1 mM TCEP、10 mM イミダゾール) 0.5% Triton X-100に再懸濁し、Completeプロテアーゼ阻害混合物 (Roche) を補充した後、超音波処理により溶解した。溶解物を15000 gで40分間遠心分離して清澄化し、Superflow Ni-NTAアガロース (Qiagen) にバッチで適用した。樹脂を洗浄緩衝液A (50 mM Tris-HCl pH8、500 mM NaCl、1 mM TCEP、10 mM イミダゾール)、続いて5カラム体積の洗浄緩衝液B (50 mM Tris-HCl pH8、1 M NaCl、1 mM TCEP、10 mM イミダゾール) で十分に洗浄した。溶出緩衝液 (50 mM Tris-HCl pH8、500 mM NaCl、1 mM TCEP、300 mM イミダゾール) で、Ni-NTA樹脂からタンパク質を溶出した。洗浄緩衝液Aに対する一晩の透析中に、TEVプロテアーゼによってHis₆-MBPタグを除去した。第2のNi-NTAアガロースカラムを通して、切断されたCas9をアフィニティタグから除去した。IEX緩衝液A (50 mM Tris-HCl pH7.5、300 mM NaCl、1 mM TCEP、5% グリセロール) にタンパク質を透析した後、5 mL Heparin HiTrapカラム (GE Life Sciences) にかけた。Cas9をNaCl (0.3~1.5 M) の線形勾配で溶出した。画分をプールし、30 kDaのスピン濃縮器 (Thermo Fisher) で濃縮した。必要に応じて、Superdex 200 pgカラム (GE Life Sciences) でのサイズ排除クロマトグラフィーによってCas9をさらに精製し、その後の切断アッセイに備えてIEX緩衝液A中に保存した。酵母を発現させるため、AR1-Cas9を、Gal1/10 His₆-MBP TEV Ura S. cerevisiae発現ベクター (Addgeneプラスミド#48305) にクローニングした。ベクターをBY4741 URA3株に形質転換し、培養物を培地中にて30℃で増殖させた。OD₆₀₀が約0.6のとき、2% w/vガラクトースでタンパク質発現を誘導し、16℃で一晩インキュベートした。タンパク質精製を上記の通り実施した。

【0378】

RNAの*in vitro*転写とオリゴヌクレオチド精製

T7プロモーター配列を含む合成DNA鋳型を使用して、以前の記載のように65℃、*in vitro*転写反応を実施した。*in vitro*転写したガイドRNA及び標的RNAまたはDNAをすべて変性PAGEによって精製した。20 mM Tris-HCl pH7.5及び100 mM NaCl中で、95℃で1分間インキュベートすることにより二本鎖標的RNA及びDNAをハイブリダイズさせた後、室温まで徐冷した。非変性PAGEによってハイブリッドを精製した。

【0379】

*in vitro*切断アッセイ

精製したDNA及びRNAオリゴヌクレオチドを、1倍PNK緩衝液中にて、37℃で30分間、T4ポリヌクレオチドキナーゼ(NEB)及び[γ -³²P]ATP(Perkin-Elmer)を用いて放射性標識した。PNKを65℃で20分間熱失活させ、illustra Microspin G-25カラム(GE Life Sciences)を使用して遊離ATPを標識反応物から除去した。1倍リフォールディング緩衝液(50mM Tris HCl pH7.5、300mM NaCl、1mM TCEP、5%グリセロール)中で、CrRNAとtracrRNAsを等モル量で混合し、70℃で5分間インキュベートした後、室温まで徐冷した。最終金属濃度1mMまで反応物を補充し、引き続き50℃で5分間加熱した。室温まで徐冷した後、リフォールディングしたガイドを氷上に置いた。緩衝液、塩濃度の指定がない限り、Cas9は、1倍切断緩衝液(50mM Tris HCl pH7.5、300mM NaCl、1mM TCEP、5%グリセロール、5mM二価金属)中にて、37℃で10分間、等モル量のガイドと再構成した。放射性標識標的よりも10倍過剰のCas9ガイド複合体を用いて、1倍切断緩衝液中にて37℃または指定温度で、切断反応を実施した。50mM EDTAを補充した等体積のゲルローディング緩衝液で反応をクエンチした。切断産物を10%変性PAGEで分離し、リン光によって可視化した。

10

【0380】

in vivo E. coli 干渉アッセイ

AR1-Cas9及びAR4-Cas9のE. coli形質転換アッセイは、以前に公開された66通りに実施した。簡潔には、ガイドRNAで形質転換したE. coliをエレクトロコンピテントにした。次に、野生型または触媒不活性のCas9(dCas9)をコードする9fmolのプラスミドで細胞を形質転換した。回収した細胞の希釈系列を、選択抗生物質を入れたLBプレートにプレーティングした。37℃で16時間経過後にコロニーを計数した。

20

【0381】

表1. CRISPR-Casシステムが同定された生物及びゲノム位置に関する詳細、ならびに再構成したスペーサーの数及び平均長ならびにリピート長に関する情報(NA、該当なし)。ARMAN-1のスペーサーは16の試料から再構成した。

【0382】

30

40

50

【表 1】

表 1

分類学的群	Casエフェクター	NCBIアクセッション番号	座位	リピート長	スペーサー数	スペーサー平均長
ARMAN-1	Cas9	MOEG01000017	1827..7130	36	271	34.5
ARMAN-4	Cas9	KY040241	11779..14900	36	1	36
Deltaproteobacteria	CasX	MGPG01000094	4319..9866	37	5	33.6
Planctomycetes	CasX	MHYZ01000150	1..5586	37	7	32.3
Candidatus Katanobacteria	CasY.1	MOEH01000029	459..5716	26	14	17.1
Candidatus Vogelbacteria	CasY.2	MOEJ01000028	7322..13087	26	18	17.3
Candidatus Vogelbacteria	CasY.3	MOEK01000006	1..4657	26	12	17.3
Candidatus Parcubacteria	CasY.4	KY040242	1..5193	25	13	18.4
Candidatus Komelibacteria	CasY.5	MOEI01000022	2802..7242	36	8	26
Candidatus Kerfeldbacteria	CasY.6	MHKD01000036	11503..15366	NA	NA	NA

10

20

30

【0383】

本発明をその具体的な実施形態を参照しながら説明してきたが、当業者は、本発明の真の趣旨及び範囲を逸脱することなく、種々の変更を加えることができる、及び等価物に代替できるものと理解されるべきである。加えて、特定の状況、材料、物質の組成、方法、方法の工程、または工程が、本発明の目的、趣旨及び範囲に適合するように多くの変更を加えることができる。そのような変更はすべて、本明細書に添付の特許請求の範囲の範囲内であることを意図する。

40

【0384】

関連出願の相互参照

本出願は、2016年9月30日出願の米国仮特許出願第62/402,849号の利益を主張するものであり、その全体が参照により本明細書に組み込まれる。

【0385】

テキストファイルとして提供される配列表の参照による組み込み

配列表を2017年9月28日作成のテキストファイル「BERK-343WO__SeqList__ST25.txt」(ファイルサイズ244KB)として本明細書とともに提出する。テキストファイルの内容は、その全体が参照により本明細書に組み込まれる。

50

【 図面 】

【 図 1 - 1 】

>CasY1
 MRKKLFKGYLHNNKRLVYTLGAAVRSIKYPLVAPNKTNLNLSEKI IYDYEHLFGPLNVA
 SYRNSNRYSLVDFWDLGKAGIYQWKSSTSLIDL ISKLEGSKSPESEKIQEIDFELKKNL
 DDKSQKFDI ILLNTGIRSSSNRSLRGRFLKCFKCEFRDTEIEAVCDKWSKDLIIVEGKS
 IIVSKQFLWEEFEGKIFPHFDHNDLPLKLTFFVPESELEFSPHLPANLCLERLKKFDIS
 RESLLGLDNNFSAFSNYFNLFLNLSRGEIKKIVTAVLAVSKSWENEPLEKRLHFLSEK
 AKLGLYPKLTSWADYRMIIGGKIKSWSNYTEQLIKVREDLKKHQIALDKLQEDLKKVV
 DSSLRQIEAQREALLPLDMLKEKDFDDLELYRFLISDFPKSLNLSYQRYIQTEEBER
 KEDRDVTKKYKDLNLSRNPREFEGESKKEQPNKFINKSLPTIDVGLKILEDIRNALETV
 SVRKPSSIIEEYVTKQLEKLSRKYKINAFNSRNFQKITEQVLRKYNNGELPKI SEVFYRY
 PRESHVAIRILPVKISNPKRDISYLLDKYQISPDWKNNSPGEVVDLIEIYKLTGLWLLSC
 NKDFMSDFSSYDLKLFPEAASLKNFGSCLSGYYL SKMIFNCITSEIKGMI TLYTRDFEV
 VRYVTQMGISNQKFLPCLVGEKQTKNFSRNGVLIIEEKGDLGEEKNQEKCLIFKDKTDF
 AKAVEIIFKNNIWRIRTSKYQIQFLNRLFKTKEDWLMNVLSEPSLVEEGBEVSVDK
 DKLLPLKKEKCEERLYSLPLNLPVATDYKEQSAEIEQRNTYLGFDVGEFGVYAVVR
 IVRDRLELLSNGFLKDPALRKIRERVDMMKQVMVAVFSSSTAVARVREMAIHSLRNQI
 HSIALAYKAKI IYGISISNPFNGNRMMAKIRKISVDVYRESGADTLVSEMIWGKNKQ
 MGNHISYSYATSYTCNCARCPTFFELIINDKYEKGGDEPIFNVDGDEKVKRGLQKSLGK
 TIKGKVELKSIKEYARPIREVLLEGEDEVQLLKKRNGSYIYRCPFCGYKTDAGIQAALN
 IACRGYISDNKADAVKEGERKLDYILEVRKLWEKNGAVLSAKFL (配列番号1)

>CasY2
 MQKVRKTLSEVHKNPYGTQVNRNAKTGYSLQIERLSYTGKEGMRSEFKI PLENKNKEVFDEF
 WKIRNDYISQVGLNLSDWYEHYQEKQEHYSLADFWKIDSLRAGVIFAHKETE IKNLSK
 IRGDKSIVDKFNASIKKHKADLVALYDIKALYDFLTSADARRGLKTEEEFFNSKRWLFPFK
 FRKDKNAVDLWVKFPIQLGNDKDLNPTKFKIGDPNPOIKYDHTFFHQIDNFRDLERTI
 TPKELSITYKFKLGNKDLGSDTEEDQLKMLVGFHNNHGAFSKYFNASLEAFRGRDLS
 LVEQIINNSPWNHSHRKELEKRIIFLQVQSKIKETELGKPHYELASFGGKFEWGSKNK
 RQEBEVKRLQFGYENKKGQKFKIVGNKDELKIRGTDEYEIKAI SKETIGLQKCLL
 LEQLKDSVDDYTLSLYRQILVELIRLNLVEFQETPELIGKSEKDEKDAKKNRKRYP
 QIFKDIKILPNFGETKQMYKFKIRASADILYEGINFIQIDKQITQNLPCFKNDKERI
 EFTKQFETLRRKYIYMNSSRHFHVI EGI INNRKLIEMKKRENSLKTFSDFSKVLSKFL
 LKKGKKEYENEVYFYIIPKARDQRRIKLVIIDINGNSVGIILQDLVQKLPKWDI IKN
 DMGELLDAIEIEKVRGLI LALYCEHFKKIKKELSLDLFASAYQYLEEDDPEELSGTN
 LGRFLQSLVCSIEKGAINKISRTVEYIERYVQPMNTEKYNPILLINKEGKATWHIAAKDDL
 SKKGGGTVAMNQIKFQKFGQDYKVTNMLQKDFDLTSLKYHLQFLSKTLDFFVGSQSW
 NKNMINDLNSYSFIEFQKVKWDLNFDLPIKIKPSENSDRRLVSVI PFVJPKPKQ
 RKDLQTRVNYMGIIDIEGYLAWTIINIDLKKNKIKKISQKQIYEPLTHKVRDYYATIKD
 NQVRCFTGMPDTPKARLRENAITSLRNQVHDIAAMRYDAKPVYEFESNFEI TSGNSKVKVI
 DSVKRADIGRQNNTEADNTEVNLVWGKTSKQFSGS QI GAYATS YICSGFGYSPIYEFENS
 KSGDEGARDNLQMKKISDPSLEDFTQGNPVYKTRFDYKYNQRLQKTDGKDEGKWT
 HRGNTAIYACQKCRHISDRPDIQASVYIALKQVYVDFYKDEMDGDLIQGDNKDKRKNVNL
 NRLIGVHKDVPININKRLITSLDINLL (配列番号2)

【 図 1 - 3 】

>CasY5
 MAESKQMQCRKCGASMKYEVIGLKKSCRYMPCDCGNHTSARKIQNKKKRD
 KKYGSASQAQSRIAVAGALYPPKQVQITKYTKYPADLNGEVHDSGVAEKIQAQI QEDEI
 GLLGPPSEYACWIASQKQSEFYSVVDWFWDVACAGGVFAYS GARLLSTVQLSGEESVLR
 AALASSPFVDDINLQAQEFKFLAVSRRTQDKLKGRI GECEAEGRLEALGIKDRMREFFQA
 IDVAGTAGORFAAKLIKIFGISQMPKQWNNDSGLTYVCLPDYVYVPEENRADDQVLLR
 LREIAYCMGIEDEAGFEHLGIDPGALSNSFNNGNPKRGLRLLNNDI IALANMNSAMTYP
 WEGRKGLIERLAWLKHRAEGLYLKEPHFGNSWADHRSRIFSR IAGWLSGCGAKGLKIAKD
 QISGVRTDLFLKRLLDVAPQASPSDFIASI SADRFLLEAAESSQDPAEQVRALYAFHL
 NAPAVRSIANKAVQRSDSQEWL I KELDVAHDHLEFNKAFPFFSDT GKKKKGGANSAPSE
 EYETETESIQQPEDAQBQEVNGQEGNASKNQKFFQRI PRFFGEGRSRYRLLTEAPQVDF
 MFCNNMRAIFMQLSEOPKRAPRDFKFLQNRILQKLYKQFTLNARSNKCRALLLESVILSWG
 EPHYTYGANEKFRLEHAEASERSDDYVVOQALEIARRLFLFGFEWRDCSAGERVDLVEI
 HKKALISFLLAITQAEVSVGSYNLGNSTVSRVLSVAGTDTLTYGTQLEEFNATVLSQMRG
 LAIRLSSQELKDFDVQLESSCQDNLQHLVYRASRDLAACKRATCPAELDPKI LVL PVG
 AF IASVMKMI ERGDEPLAGAYLHRPHSFGWQIRVRGVAEVMGDQGTALAFQKTEPSEFP
 KIKPFSAQYGVPLWLNSSYSQSYLDGFTSOPKNWSMRVLPAQGSVRVEQVRLIWNLQ
 AGKMRLESGARAFMPVPVFSFRPSGSDAEVAPNRYLGLFPHSGGIEYAVVDVLSAG
 FKILERGTIAVNGFSQKRGERQEEAHRKQRRI SDI GRKKPVQAEVDAANE LHRKYTDV
 ATRLGCRIVVQWAPQPKGPTAQTAVYARAVRTEAPRSGNQEDHARMKSSWGYTWGTYW
 EKRRPEDIILGISTQVYWTGGIGESCPAVAVALLGHIRATSTQTEWEKEEVVFGRLKFFP
 S (配列番号6)

>CasY6
 MAESKQMQCRKCGASMKYEVIGLKKSCRYMPCDCGNHTSARKIQNKKKRD
 KKYGSASQAQSRIAVAGALYPPKQVQITKYTKYPADLNGEVHDSGVAEKIQAQI QEDEI
 GLLGPPSEYACWIASQKQSEFYSVVDWFWDVACAGGVFAYS GARLLSTVQLSGEESVLR
 AALASSPFVDDINLQAQEFKFLAVSRRTQDKLKGRI GECEAEGRLEALGIKDRMREFFQA
 IDVAGTAGORFAAKLIKIFGISQMPKQWNNDSGLTYVCLPDYVYVPEENRADDQVLLR
 LREIAYCMGIEDEAGFEHLGIDPGALSNSFNNGNPKRGLRLLNNDI IALANMNSAMTYP
 WEGRKGLIERLAWLKHRAEGLYLKEPHFGNSWADHRSRIFSR IAGWLSGCGAKGLKIAKD
 QISGVRTDLFLKRLLDVAPQASPSDFIASI SADRFLLEAAESSQDPAEQVRALYAFHL
 NAPAVRSIANKAVQRSDSQEWL I KELDVAHDHLEFNKAFPFFSDT GKKKKGGANSAPSE
 EYETETESIQQPEDAQBQEVNGQEGNASKNQKFFQRI PRFFGEGRSRYRLLTEAPQVDF
 MFCNNMRAIFMQLSEOPKRAPRDFKFLQNRILQKLYKQFTLNARSNKCRALLLESVILSWG
 EPHYTYGANEKFRLEHAEASERSDDYVVOQALEIARRLFLFGFEWRDCSAGERVDLVEI
 HKKALISFLLAITQAEVSVGSYNLGNSTVSRVLSVAGTDTLTYGTQLEEFNATVLSQMRG
 LAIRLSSQELKDFDVQLESSCQDNLQHLVYRASRDLAACKRATCPAELDPKI LVL PVG
 AF IASVMKMI ERGDEPLAGAYLHRPHSFGWQIRVRGVAEVMGDQGTALAFQKTEPSEFP
 KIKPFSAQYGVPLWLNSSYSQSYLDGFTSOPKNWSMRVLPAQGSVRVEQVRLIWNLQ
 AGKMRLESGARAFMPVPVFSFRPSGSDAEVAPNRYLGLFPHSGGIEYAVVDVLSAG
 FKILERGTIAVNGFSQKRGERQEEAHRKQRRI SDI GRKKPVQAEVDAANE LHRKYTDV
 ATRLGCRIVVQWAPQPKGPTAQTAVYARAVRTEAPRSGNQEDHARMKSSWGYTWSTY
 EKRRPEDIILGISTQVYWTGGIGESCPAVAVALLGHIRATSTQTEWEKEEVVFGRLKFFP
 S (配列番号7)

【 図 1 - 2 】

>CasY3
 MKAKKSFYNQKRKFKGRYRHLDERIAYS GGIGSMRSIKYELKDSYGIAGLRNR
 IADATISDNKWLGNINLNDLVEWRSSKDTQIEDDRESSLGLFWLBAALRGLFVFSKQS
 HAPDFENFALQDLFEFLDDDLKHLVDRKKWCDFIKIGPKPTNDQGRLLKQIKNLLKGNK
 REEIEKTLNESDDELKEKINR IADVFAKNKSQDYTFI FLKDKPNTKRYPRINDVQVAFCH
 PDFEEITERDRTKTLDLINRPNKRYEITENKDKDKTSNRMALYSLNQGYIPRVLNDLFL
 FVKDNEDDFSQFLSDLENFFSFSNEQIKI I KERLKLKLYAEPI PGKPOLADKWDYASD
 FGGKLESWYSNRIEKLKIPESVSDLRNNLEKIRNVLKKNASKLLELQSKI IEYTRDY
 GVSFEKPEI I KPSWINTKDQGGKVFVYAKMADREPIEKLDMWADLRSQNLNRYQDNKV
 SFKKGKKEELGVLDPALNKAKNKSTKNENGWQOKLSEITQSAPELFFGEGNRVRENEV
 YNLKDLFSEIKNVENILMSSEADLKNKIEIYKEDGAKGNVNLVNLARFYARFNEDGY
 GGWNKVTVLENLAREAGTDFSKYGNNNRNRAGRYFLNGRERQVTLIKFEKSI TEVEKIL
 ELVKLPSLLEADAYRDLVNNENKHLRVDIQLSKT IMAVLVSHSDKEKIGGNYI HSKLSG
 YNALISKRPDI SRYSVQTTNGTQCKLAIKGGKSKGNEIDRYFYAFQPFKNDKSKINLK
 IKNNSHKNDIDPNDNENKINALQVYSNRYIQDFLFFFEKHQKKTSLVEGSGSPTIAEKS
 TLDWGSNPNRVGPKRSDEKRVVPSQPTLIPDDDEKERRKEMI KTKNRF I D IGEY
 GLAWSLIEVDNGDKNNGRIRFVLESGFITDNQGVKLVKVVKSQRQKIRQFTTSPDTKIR
 LRESLIGSYNQKLESIMVAKKANLSFEYVSGVGFVGRKRVAKIYDSIKRGSVRKDDNNSQ
 NDQSWGKKGINWSEFETAAGTSQFCHCEKRSGLAVDIEEVELKDNVDFKVKINDG
 EVRLGKGGKGRSAGEIKGKELFGPVKDMRPNVDGLGMKIKRKYKLLDLRWDVSRYGNM
 AIFIPCYVDCHHISHADKQAAFNIAV (配列番号3)

>CasY4
 MSKRHRPISVGVGRYLRHAQRLEYTGKSGAMRTIKYPLYSPPSGGRTVPREIVSAINDDYV
 GYGLSNFDDLYNAEKREKVEYVSLDFWYDCVQYGAVFSYATPGLLNVAEVRGGSYLE
 TTKLKGSHLYDELQIDKVIKFLNKKELSRANGSLDKLKKDIIDCFKAEYRERHKQCNKL
 ADIKNAKDKAGASLGERQKLFDRDFGISEQSNKPSF'NPNLNTCCLLPDP'VNNNR
 NRGVEINFLKLEYAQKLDKNEGSLEWMEYIIGNSGTSFNFGEFLGRLENKIKELAR
 KAMMIDITAWRQEQEELKRLRLAAILTKLREPKDFNHWGGYRSINGKLSLWLNQY
 INQTVYKIDEDLKGKDKLKKAKEMINFRGSDTKEAVVSSLDEKIEI VPDSDSDEKPY
 DIPALAIYRRFLSDGRITLNRVQREDQVQALIKERLEAEKPKKPKRKKKSDADEKET
 IDPKEFLFKLAKPLKLVNPNFYGDSKRELYKKNYKNAIYTDALWAVEKIKYKSAFSSLSK
 SFFDFFDKDF I KRLQKIFSVYRRFNTPDKWIPVNSGPAFYCDIVSLAENEVLYKPKQS
 RSKLSAAIDKNRVLSPSTENIAKAGIALARELSVAGFDWDLKKEHEEYIDLIELHTK
 ALALLAVTETQLDI SADFVENGTVKDMTRDGNLVEGRFLEMFSQSIVSELRGLA
 GLMSRKEFT'RAAIQTMMGQAEALYIIPHEQSAKI'PTPKMSRAFLLDPAEPAFSALEP
 ELSSEKSLLLKQMYRYPHYFGEYELTRTQGGIDVAGBNALRLEKSPVKRKEIKCKQYI
 LGRGQNKILVLYVRSYQYQFLWFLRHPKVVQDVAVSGSFLIDKPKVTRKNYDALYT
 ALRFPVSGSERVSVSPPTIFPEKSAEEOEQRYLGI D IGEYIATYLAITGDSALLDQN
 FISDQLKTLREEVKGLKLDQRRTFAMPSTKIAR IRESLVHSLNRNHHLLKHKAKIV
 YEELVSRFEGEKQKIKKVIATLKKADVYSEIADKNLQTTVWGLAVASEIASYTSQDF
 GACKLWARAEMQVDEITITQELIGTVRVIKGGTLIDAKDLFMRPITFDENTFFPKYRC
 CDKHIISKMRGNSCLFICPFCRANADAIQASQITALLRVVKEEKKVYDFERFRKLKN
 IKVLQMKKI (配列番号4)

【 図 1 - 4 】

>CasY7
 MKRIINLSLVAALRLLFRGKGSSELVTKVYPLVSPVQGAVEELAEAIRHNDLHLFGQKEI
 VDLMEKDEGTQVYSVDFWLDLTLRLGMFFSPANAKITLKGKNSDQVSPFRVLEQSPFI
 FLAGRLKVEPAERILSVEIRKIGKRENRVNYADVETCFITGLSDEKQS IQLANDIW
 DADKHEEQRLKADFFAIPLIKDPKAVTEEDPENE TAGQKPLELVCVCLVPELYTRGFS
 ISDKFLVQRLTLRLDKMSDTADBCLEVYGI EKEGKNGNSLLG'FLKNLQDGFQEIYFQF
 MLGSSVWGQKEDVLRERLDBLAEKVKRIPKPKFAEGWSHRMLFHGQLKSDSNFFRFL
 NRETRELESIKSDIQHATMLISYVEKGGYHPQLSQQRYKMLQELPALTQWLEIEMT
 HMSAEVRSYIMIHKSVAAGFLPDLLESIDRDKDRFLLSIFPRI PKDKKTEIVAVELPG
 EPEEGYLFETANLFRNLENPKHVPRFMAERIPEDWTRILRSAPVFDGMVKQWQVNHQL
 VESGALYQFNEFLRQLQAMLTVYKRDQTEKFLKLLAVDACPRLVDFFLGGNDIIFK
 SQDPRKQWQTVIPLSVPADVYACEGLAIRLRETLGFEPWNLKHEHREDFLRHLQLLGN
 LLEWIRDAKLVKLEDMNPNVCQEVYEAARKAIDLPLEIIFGFEVPIFLNGYLFSELRLQLE
 LLLRKSVMTSYSVKTGSPNRLFQVLYLPLNPSDEPKKNSNFRQERLDTPTGLSRRLFD
 LTRRDAKGLLTDVPTQELKTMAGFYDHLFGFKLPCCLAAMSNHGGSSKMVLAQPKKG
 VASNGFEPIPPAHFVFRVRSWPELKYLEGGLYLPEDPTLTLAETSVCVSGSVVA
 DLKLNLTILGRVGEFRVADQPKLPIIPEKESIEGKTYLGLDAGERSVGFVAVITV
 DGGYEQRLGVHEDTQMLAQVVASKSKKPEVFPQIRKGTFRQERIRKSLRGCYWNFY
 HALMITYRKYRKHVESVSSGLVQWLRAPQDKLKKADLPKGGKGNVGDKRRKSSAQD
 TLWGFAPSKKEQQIAFEVQAAGSSQFCLKCGWFWLGMREVNVRQESGVLDWNRN IFT
 FLESSEKVEYGFSPQOLEKGFPRDIEPTKMMVRDFMRPMPDRKGRPAAYERVFLGR
 HRRYRFDKVFERRFGRSALFICPRVCGFNHDSSESQSAVALLIGYIADKEGMSGKGLVY
 VRLAELMAEWKLLKLESRVEEQSSAQ (配列番号5)

>CasY8
 MKRIAKFRHDKPVKREASWGYRVHKNRI INKVTRSIKYPLVVKDEWKKRLIDDAHDYRWLWV
 PINSYDWRDPNQYSILEFWIDFLCVGGVQSSHSIMCRALAIQLSGGVSFEQEWKDLSPFVRAN
 LIQGIKPAEFIGFLTAEFRSSSNPNKFI SKFPEGNSDELESITNEFASVDFIKAKDILSRKS
 LPSCKIAPNLWEKAVGSHSTNELKLLTKYTRVMLVAEPPSHSDRVFSQTLQSDNDDEPILTG
 PLPSHKVQKASYLPEIPEIREVNLDKISKLDLAKSLAVEVQKKSLELTSDFKQIEMQSEYAF
 GISTSFNLSNFLGILIRTLRNPAPAILKQDIALCALPKDILKILITLDLWLCRQALPENPFEI
 TNWAEYRSYLGKIKSWFSNYENPFPIQAASSQNNRKKLGNRSAILRNLKKEAPEKARE
 TPKDQKGTLEKIDLAYRLGSSISEPEVLQCDGELKLYQGFNDELVLNLETINQKFDQAKRD IKAK
 KEKESFEKLRNLSPLRPIPEFFGERAKKGYQKARVSKLARHLBCELDNARLFAKVEESA
 SEKQERILDLWRTSDFLPIRKSQKDPWSWRYIARVATGYKYVWSEYSRKRVY I DKP1AQ
 NPELKEIISWFLNLLKDAFSDRNLDFKGLSSKVMIRIARIMAGILDRDREGLWGLQAMTSLKINL
 LHPSFSVPVTDLSLKDAFYRAFFSELEGLNIGRSLI IERI TLQSQQSKNKKTRRPLMPEPFI
 TNEDKVEFLAPKFEFTKKNVGTTRVYNSPDEVNWLSP I RSKSQGLSFMFRCLSEDAKIMTTSQ
 GCSYIEVFKLLBAQEAVLSIHDCDI IPRAPVSIPTLRESEBETKPDWKNRIMFGDIEGAV
 AYCVELKGTDSIEILDCGIVRNGAHRVLEKVDRLKRSQMTFEGMATS IIAARESSLVGNRY
 RLHALKHLGAKLVYVEVSAFESGNNRILKVEYTLKRSQDSTGDEKARNKRIWGTNAVGD
 QIGAGWTSQCAKCGRSFGADLKAGNFGVAVVPEKVEDSKGHYAYHEPPEFDGLKVRGFLKPN
 KISDQKELAKAVHYMRPPLVALGKRLKPNARYRGNSSLFRCPSCDGGTADADTQAANYI
 AVKQLYPKPKGYPERKRWQDFVILKPKPEKSLFDKQYRPN (配列番号8)

10

20

30

40

50

【 2 - 1 】

A

1. Casy1	1	2	3	4	5	6	7	8	9
1. Casy1	M	--R	K	L	F	K	---	---	G
2. Casy2	M	Q	K	V	R	T	L	S	---
3. Casy3	M	K	A	K	S	F	Y	N	---
4. Casy4	M	--S	R	---	---	---	---	---	G
5. Casy5	M	A	S	K	O	V	O	C	R
6. Casy6	M	A	S	K	O	V	O	C	R
7. Casy7	M	--K	E	L	I	N	---	---	S

1. Casy1	11	21	24	32	39	48
1. Casy1	I	H	N	E	P	E
2. Casy2	S	I	O	J	E	S
3. Casy3	E	H	D	E	S	S
4. Casy4	E	H	A	D	E	S
5. Casy5	A	O	S	O	R	J
6. Casy6	A	O	S	O	R	J
7. Casy7	K	V	A	A	L	E

1. Casy1	64	65	67	76	86	96
1. Casy1	P	I	N	V	A	S
2. Casy2	G	L	N	L	S	D
3. Casy3	G	L	N	L	S	D
4. Casy4	G	L	N	L	S	D
5. Casy5	G	L	N	L	S	D
6. Casy6	G	L	N	L	S	D
7. Casy7	G	L	N	L	S	D

【 2 - 2 】

A

1. Casy1	106	113	123	133	137	144
1. Casy1	E	G	S	P	S	E
2. Casy2	R	G	S	P	S	E
3. Casy3	R	G	S	P	S	E
4. Casy4	R	G	S	P	S	E
5. Casy5	R	G	S	P	S	E
6. Casy6	R	G	S	P	S	E
7. Casy7	R	G	S	P	S	E

1. Casy1	154	164	174	181	191
1. Casy1	R	G	F	L	K
2. Casy2	---	---	---	---	---
3. Casy3	---	---	---	---	---
4. Casy4	---	---	---	---	---
5. Casy5	---	---	---	---	---
6. Casy6	---	---	---	---	---
7. Casy7	---	---	---	---	---

1. Casy1	197	201	209	219	230	231
1. Casy1	K	I	F	P	K	---
2. Casy2	D	P	E	S	T	---
3. Casy3	---	---	---	---	---	---
4. Casy4	---	---	---	---	---	---
5. Casy5	---	---	---	---	---	---
6. Casy6	---	---	---	---	---	---
7. Casy7	---	---	---	---	---	---

【 2 - 3 】

A

1. Casy1	237	244	251	265	275
1. Casy1	L	K	F	D	I
2. Casy2	L	G	S	D	T
3. Casy3	L	G	S	D	T
4. Casy4	L	G	S	D	T
5. Casy5	L	G	S	D	T
6. Casy6	L	G	S	D	T
7. Casy7	L	G	S	D	T

1. Casy1	285	292	302	310	320	330
1. Casy1	V	S	R	S	S	---
2. Casy2	N	S	P	K	G	---
3. Casy3	F	E	S	---	---	---
4. Casy4	I	H	A	D	E	S
5. Casy5	M	I	P	---	---	---
6. Casy6	M	I	P	---	---	---
7. Casy7	S	Y	V	G	---	---

1. Casy1	340	343	353	363	373
1. Casy1	T	E	L	I	E
2. Casy2	L	R	E	---	---
3. Casy3	I	E	L	---	---
4. Casy4	I	E	L	---	---
5. Casy5	A	S	K	---	---
6. Casy6	A	S	K	---	---
7. Casy7	F	R	I	---	---

【 2 - 4 】

A

1. Casy1	383	392	395	396	398
1. Casy1	K	E	K	D	F
2. Casy2	K	E	K	D	F
3. Casy3	K	E	K	D	F
4. Casy4	K	E	K	D	F
5. Casy5	K	E	K	D	F
6. Casy6	K	E	K	D	F
7. Casy7	K	E	K	D	F

1. Casy1	412	413	419	424
1. Casy1	N	G	S	---
2. Casy2	N	G	S	---
3. Casy3	N	G	S	---
4. Casy4	N	G	S	---
5. Casy5	N	G	S	---
6. Casy6	N	G	S	---
7. Casy7	N	G	S	---

1. Casy1	430	433	443	453	463
1. Casy1	---	---	---	---	---
2. Casy2	---	---	---	---	---
3. Casy3	---	---	---	---	---
4. Casy4	---	---	---	---	---
5. Casy5	---	---	---	---	---
6. Casy6	---	---	---	---	---
7. Casy7	---	---	---	---	---

【 2 - 5 】

A

704	711	721	729	737	747
1. Casy1	GEER	--NOBKCLIFKDDTDFAR--	ABEVEIFKANN--	MRPETSKEVIOFIRKLEKFT	
2. Casy2	SKKIGG	GTVAMNQIKGNFFGQDYKVFMLQDKR--	FDLITSKYHLQFJKTKTDTG		
3. Casy3	SKIN	--LQVIKNNSHK--	NIDFNENKIKWALQWNSYIOFIDWFFEKH		
4. Casy4	GVAENALRLKSPVKRILKCKO--	YRDLGRGONKTI--	VLYWRSYVTOFIDWFFHRP		
5. Casy5	GMDQ--	GTALAFQPTSEPPK--	IPPSAOPYL--	LMLNSYVSOYVINGFSDP	
6. Casy6	HPSSKMMVVIKPKGVASNIG--	FPIPPDPAHP--	RRRKSWMFELKYHJLFLP		
7. Casy7					

【 2 - 7 】

A

704	711	721	729	737	747
1. Casy1	GEER	--NOBKCLIFKDDTDFAR--	ABEVEIFKANN--	MRPETSKEVIOFIRKLEKFT	
2. Casy2	SKKIGG	GTVAMNQIKGNFFGQDYKVFMLQDKR--	FDLITSKYHLQFJKTKTDTG		
3. Casy3	SKIN	--LQVIKNNSHK--	NIDFNENKIKWALQWNSYIOFIDWFFEKH		
4. Casy4	GVAENALRLKSPVKRILKCKO--	YRDLGRGONKTI--	VLYWRSYVTOFIDWFFHRP		
5. Casy5	GMDQ--	GTALAFQPTSEPPK--	IPPSAOPYL--	LMLNSYVSOYVINGFSDP	
6. Casy6	HPSSKMMVVIKPKGVASNIG--	FPIPPDPAHP--	RRRKSWMFELKYHJLFLP		
7. Casy7					

【 2 - 5 】

A

497	497	497	497	497	497	
1. Casy1	GLKTE	DNRNAPETVS	YK--	PPSTTE	YVTKO--	LEKSRKVKINAFNS--
2. Casy2	GINF	DDQDKQ	TQNLIPC--	PKNDRK	IEIETKQ--	FETRRKRYL--
3. Casy3	FIKNE	MLMSBE	AEEDLN--	IKIEXK	DGAKG	YVNVVAFAR--
4. Casy4	LKWA	VEKYK	SAPSSS	LNNSFD	TDFD	DEFIKR--
5. Casy5	FCNN	RAIF	MOESOP	PKKA--	PRDF	CKLONS--
6. Casy6	FCNN	RAIF	MOESOP	PKKA--	PRDF	CKLONS--
7. Casy7	MVXQ	WQVQ	VQVQ	YESPOL	---	YQFNSELRQK--

【 2 - 6 】

A

550	551	561	571	577	587	
1. Casy1	I--	-----	LPVKISNPR	PTSYLLDKYQ	LSPLN--	SNPNQVVDITETITL
2. Casy2	INPKAR	QORR	KIVL	DINGNS	VGILO	VLQKPKMDDI
3. Casy3	IKF	--EK--	--SITVEK	LEL	KLPSL	DEMRDLV
4. Casy4	AAI	--DKNR	VRLP	STENAK	AGAT	ELSVAGR
5. Casy5	SDP	--P--	-----	YVQK	PLP	PLP
6. Casy6	SDP	--P--	-----	YVQK	PLP	PLP
7. Casy7	P--	-----	LSVPAD	VYFAC	EGTAL--	RLRETLGFEH--

【 2 - 6 】

A

597	607	615	625	630	640	
1. Casy1	TTCG	MLTSCN	KPSS	SMDESSY--	RLKLP	PEAASLKNF--
2. Casy2	RLG	MLTSCN	KPSS	SMDESSY--	RLKLP	PEAASLKNF--
3. Casy3	IMAL	MLTSCN	KPSS	SMDESSY--	RLKLP	PEAASLKNF--
4. Casy4	ADAL	MLTSCN	KPSS	SMDESSY--	RLKLP	PEAASLKNF--
5. Casy5	ADAL	MLTSCN	KPSS	SMDESSY--	RLKLP	PEAASLKNF--
6. Casy6	ADAL	MLTSCN	KPSS	SMDESSY--	RLKLP	PEAASLKNF--
7. Casy7	LGNI	FWHR	DAK	AVWAKLE--	BWMN	PCVQEV

【 2 - 7 】

A

809	818	827	836	842	851	
1. Casy1	LNEI	VEADYK	EOSA	HE--	THEQR	TYLGHV
2. Casy2	FTVTK	PK--	--OIKR	RD--	--DTRV	NYGTH
3. Casy3	FTVTK	PK--	--OIKR	RD--	--DTRV	NYGTH
4. Casy4	FTVTK	PK--	--OIKR	RD--	--DTRV	NYGTH
5. Casy5	FTVTK	PK--	--OIKR	RD--	--DTRV	NYGTH
6. Casy6	FTVTK	PK--	--OIKR	RD--	--DTRV	NYGTH
7. Casy7	FTVTK	PK--	--OIKR	RD--	--DTRV	NYGTH

【 2 - 8 】

A

861	870	876	886	896	906	
1. Casy1	FK	PA	LRK	HE--	RVD	DMK--
2. Casy2	FYV	PK	YK	MD	--YVAT	IK--
3. Casy3	FTVTK	PK--	--OIKR	RD--	--DTRV	NYGTH
4. Casy4	FTVTK	PK--	--OIKR	RD--	--DTRV	NYGTH
5. Casy5	FTVTK	PK--	--OIKR	RD--	--DTRV	NYGTH
6. Casy6	FTVTK	PK--	--OIKR	RD--	--DTRV	NYGTH
7. Casy7	FTVTK	PK--	--OIKR	RD--	--DTRV	NYGTH

【 2 - 8 】

A

906	916	926	936	942	947	
1. Casy1	AY	KK	LY	VEI	SISNS	FETG
2. Casy2	RY	DA	VP	TH	FE	NSFETG
3. Casy3	RY	DA	VP	TH	FE	NSFETG
4. Casy4	RY	DA	VP	TH	FE	NSFETG
5. Casy5	RY	DA	VP	TH	FE	NSFETG
6. Casy6	RY	DA	VP	TH	FE	NSFETG
7. Casy7	RY	DA	VP	TH	FE	NSFETG

【 2 - 9 】

A

	1,002	1,009	1,019	1,028	1,042
1. Casy1	DEFFENVGDIIIEKKYRGFLQKSLGKTKKGGKVLKSKYVAPPIREV				
2. Casy2	RDNVYQMK--KLRPSLE--DFLQGNVYKTFDFDKYKNDQR				
3. Casy3	NDNFFVKIN--DGEVRLGKGGWRSGEKIKGKLFPPVGDAMPNVDGL--GKIVKR				
4. Casy4	MELIIGTVR--VTKGGTLDIKDFMPPPIFDENDTFFPKY				
5. Casy5					
6. Casy6					
7. Casy7	NRSVTFLEIESSGEKVYGFSPQQLKGFREDIEFKKVVDFMPPMPDRKGRPAAAY				

	1,043	1,046	1,055	1,070	1,080
1. Casy1	---LLEGEDVEQLLK---RRNSYLYKPPF--GGYKTDADIQALNIAQRCVYISD				
2. Casy2	---LQKTGDKDGEMKT---HRNTAVYKOK--GRHISDIDIQKSYWIKKQVDFD				
3. Casy3	KYL---KLDLDRWVS---RYENMAVYKPPYVGRHHISDIDIQKSYWIKKQVDFD				
4. Casy4	RDF---CDKHHISKK---MRENSCLHPPF--GRANADADIQKSYWIKKQVDFD				
5. Casy5					
6. Casy6					
7. Casy7	ERFVLRGRRHRRYRFDKVFRRFRSALHPRVGGNFGNFDHSEQSAVVLALGVIAD				

【 2 - 10 】

B

	1,090	1,100	1,110	1,114	1,122	1,125
1. Casy1	NAKDAVKEGERKLDYILEVRKLEWE					
2. Casy2	FYRDKEMEGDLIQGDNKDKKRNELNRLIGVHEDVPHHNEENLITSLDILL					
3. Casy3						
4. Casy4	EKKVE---					
5. Casy5	TSQTEW---					
6. Casy6	TSQTEW---					
7. Casy7	KEGMSGKK---					

	818	827	836	844	850
1. Casy1	HEQRNTHLGLGVVDFP--GVAVAVVQR--VRDR--				
2. Casy2	QTRVNYGGLDIIIGY--GLAMTHINLDLKNK--				
3. Casy3	HKTKNRGLDIIIGY--GLAMTHINLDLKNK--				
4. Casy4	GORNLGDIIGY--GLAVTATE--TGDS--				
5. Casy5	VLAPNRGLDFPHSG--GLAVAVDV--LDSA--				
6. Casy6	VLAPNRGLDFPHSG--GLAVAVDV--LDSA--				
7. Casy7	GKTKRGLDIIIGY--GLAMTHINLDLKNK--				
8. AsCp1_5843	EHPETPILGIDRGR--NLIIVITV--HDSIGKI--				
7. LbcP1_5106	HDDNPNYVGLDGRGR--NLIIVITV--VDGKGNL---				

10

20

【 2 - 11 】

B

	899	909	919	928	937
1. Casy1	NOHSDIAHAKKIKIYIYISISNFIIGNRN--AKLRSIKVSVYVRS				
2. Casy2	NOVHDIAHAKKIKIYIYISISNFIIGNRN--AKLRSIKVSVYVRS				
3. Casy3	NOVHDIAHAKKIKIYIYISISNFIIGNRN--AKLRSIKVSVYVRS				
4. Casy4	NOVHDIAHAKKIKIYIYISISNFIIGNRN--AKLRSIKVSVYVRS				
5. Casy5	NOVHDIAHAKKIKIYIYISISNFIIGNRN--AKLRSIKVSVYVRS				
6. Casy6	NOVHDIAHAKKIKIYIYISISNFIIGNRN--AKLRSIKVSVYVRS				
7. Casy7	NOVHDIAHAKKIKIYIYISISNFIIGNRN--AKLRSIKVSVYVRS				
8. AsCp1_5843	NOVHDIAHAKKIKIYIYISISNFIIGNRN--AKLRSIKVSVYVRS				

	942	955	966	980
1. Casy1	---ESGATLVSEMIW---KKN--KQMGNHISY			
2. Casy2	---QNTEANTEVNLVW---KTS--KQFGSSQICGY			
3. Casy3	---KDNNSQNDQSM---KKN--KQFGSSQICGY			
4. Casy4	---EIDAKNLQTVW---KLA--VASEHIS			
5. Casy5	---SGNQHARMKSM---YTWG---TYWEKKKPEDILGISTQVY			
6. Casy6	---SGNQHARMKSM---YTWG---TYWEKKKPEDILGISTQVY			
7. Casy7	KNGVDKRRSSAODLW---GAFSKEE--QOIAFEVQAA			
8. AsCp1_5843	---LNCVLKDYPAEKVGVLPYQLTDQTFSAKMGTSGLFFVYVAP			
7. LbcP1_5106	---LNYMVDKKNPCATGGALKGYQITNKFESFKMSMTQNGFFIFYVAPW			

【 2 - 12 】

B

	970	980	984	995
1. Casy1	ATSYTCCNCRATPFPELHNDKYEK			
2. Casy2	ATSYTCCNCRATPFPELHNDKYEK			
3. Casy3	ATSYTCCNCRATPFPELHNDKYEK			
4. Casy4	ATSYTCCNCRATPFPELHNDKYEK			
5. Casy5	ATSYTCCNCRATPFPELHNDKYEK			
6. Casy6	ATSYTCCNCRATPFPELHNDKYEK			
7. Casy7	ATSYTCCNCRATPFPELHNDKYEK			
8. AsCp1_5843	ATSYTCCNCRATPFPELHNDKYEK			
7. LbcP1_5106	ATSYTCCNCRATPFPELHNDKYEK			

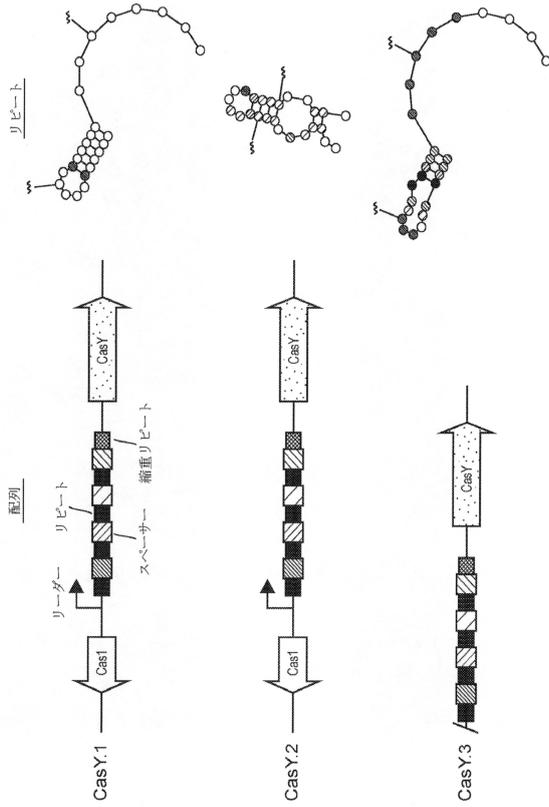
	1,004	1,005	1,012	1,016
1. Casy1	VG---			
2. Casy2	MKK---			
3. Casy3	VKI---			
4. Casy4	VRV---			
5. Casy5	---			
6. Casy6	---			
7. Casy7	FLIE---			
8. AsCp1_5843	FKNRRNLSFQGLPGFMPAWDIVFERNETQDAXGTPFIAGKRIVPVIN			
7. LbcP1_5106	NFSTTDADYIKKWKLYSYGNRIRIFAA			

30

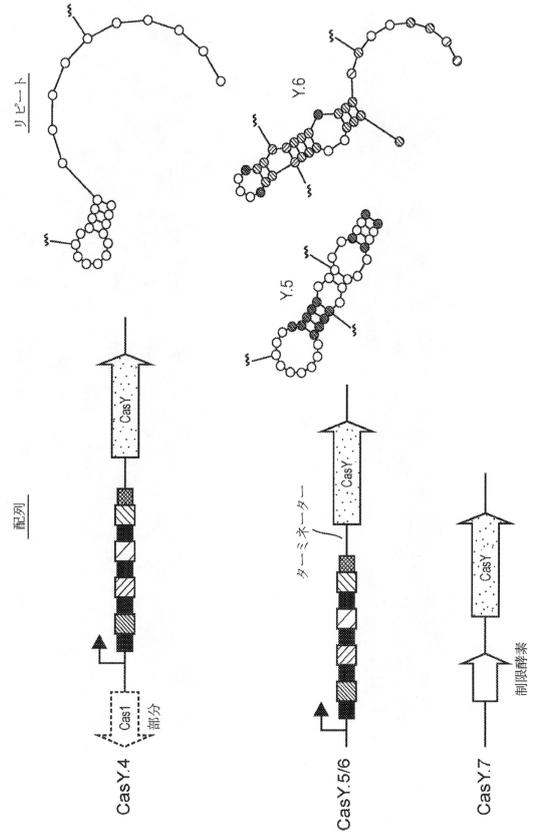
40

50

【 図 4 - 1 】



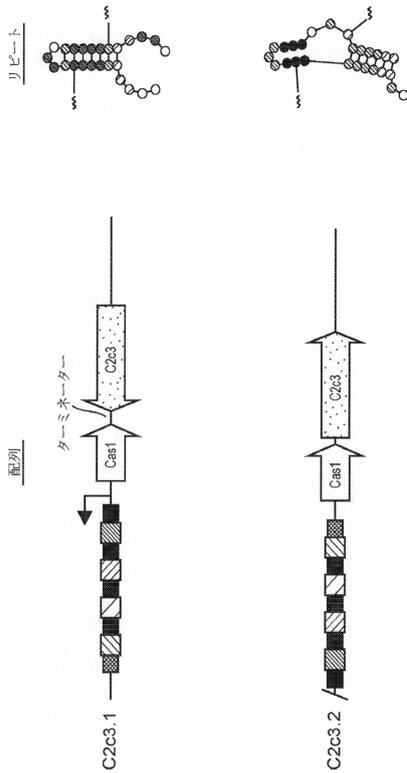
【 図 4 - 2 】



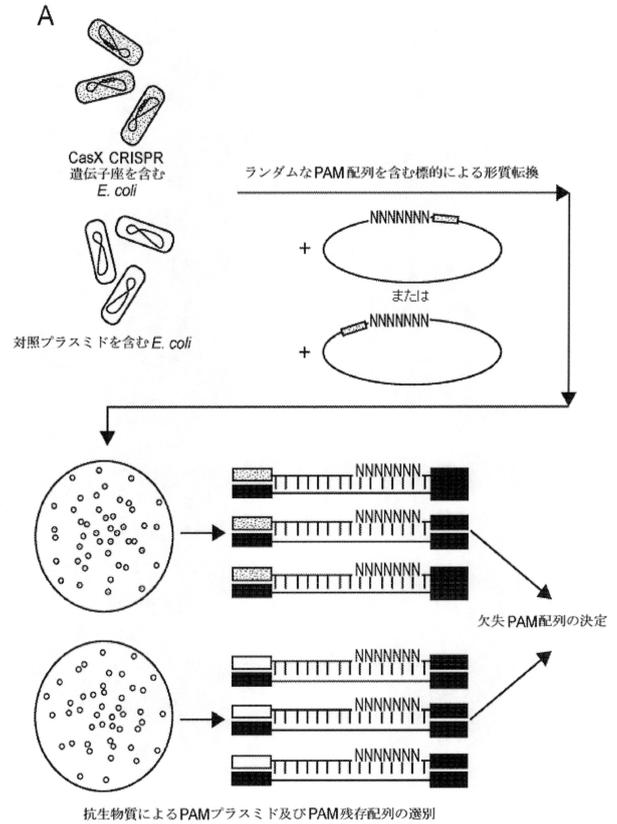
10

20

【 図 4 - 3 】



【 図 5 - 1 】

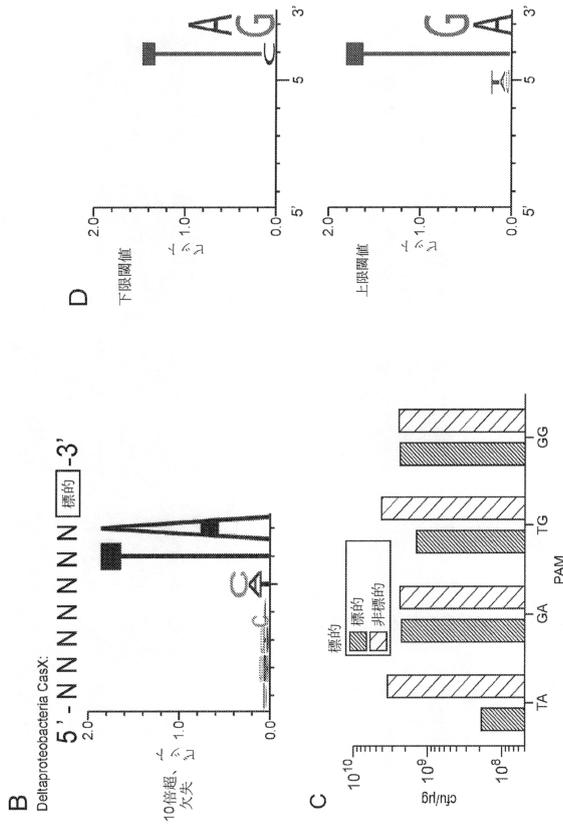


30

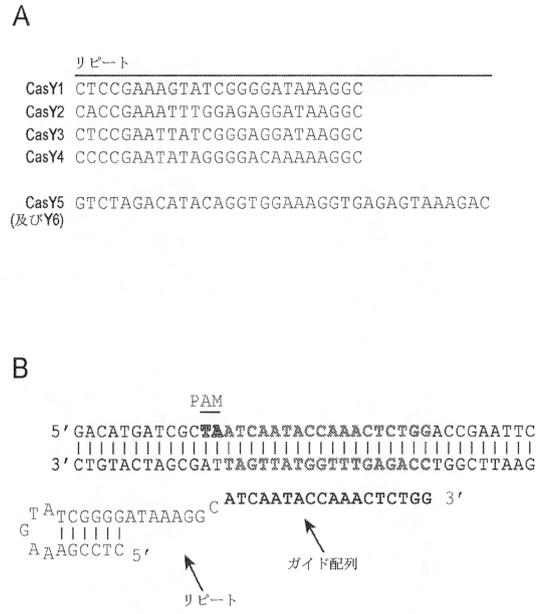
40

50

【 図 5 - 2 】



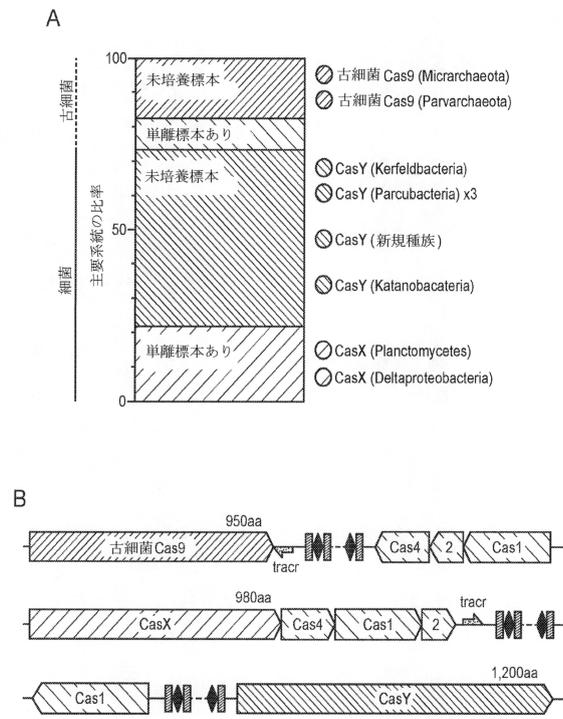
【 図 6 】



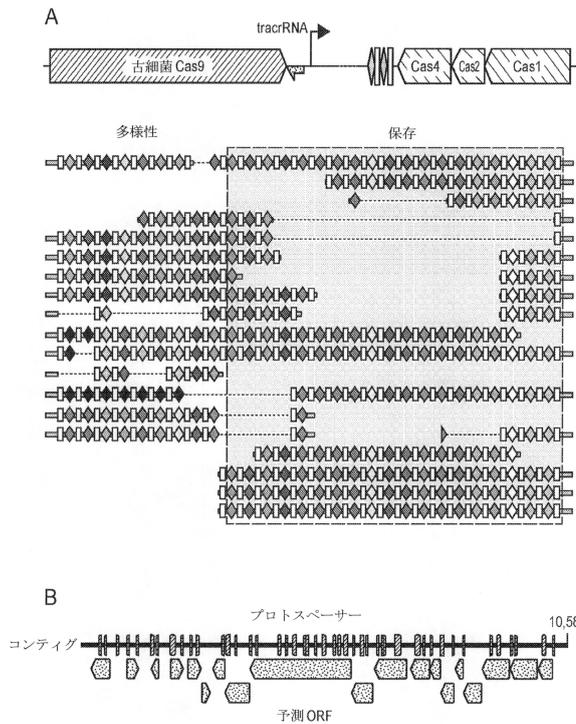
10

20

【 図 7 】



【 図 8 - 1 】

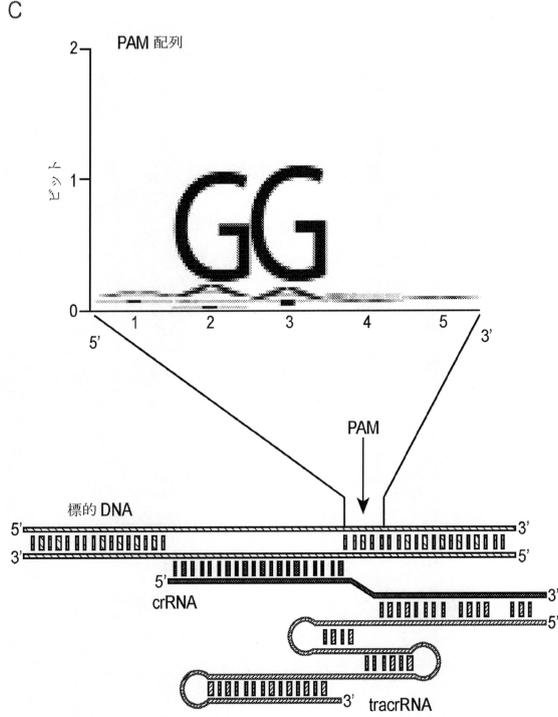


30

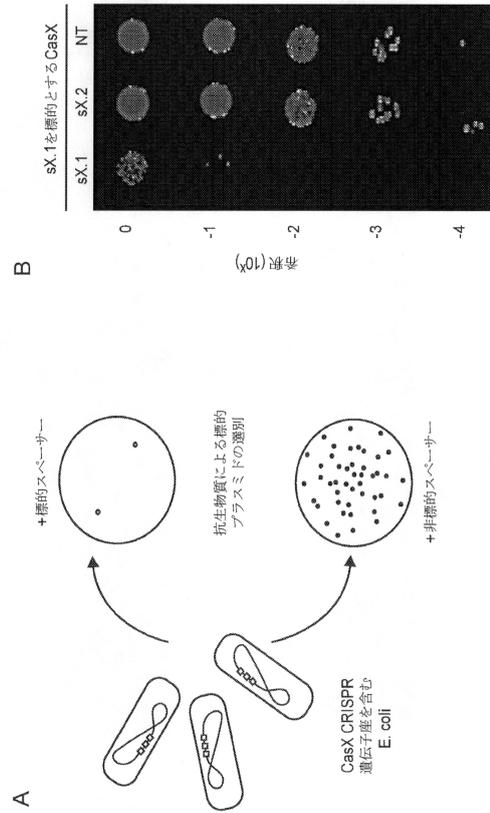
40

50

【 図 8 - 2 】



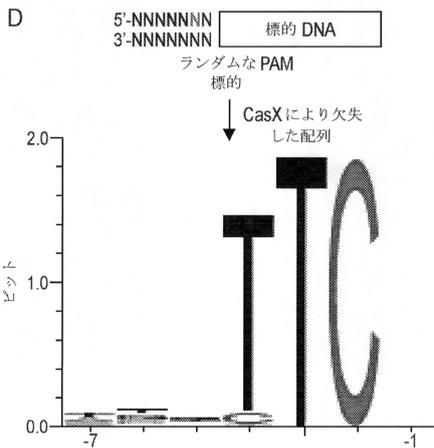
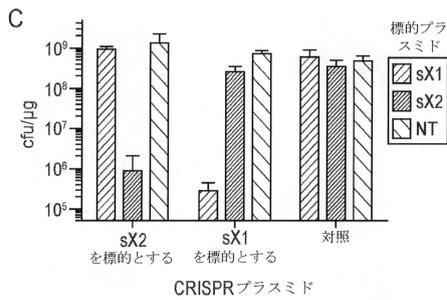
【 図 9 - 1 】



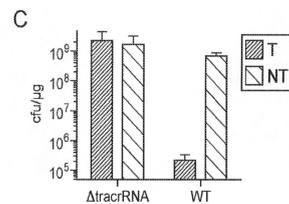
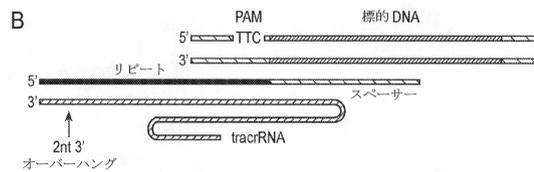
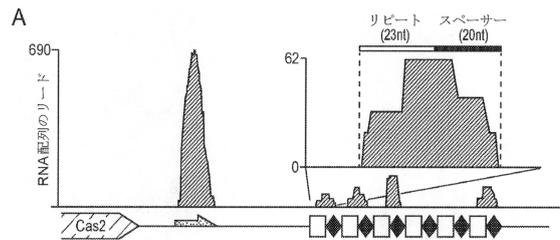
10

20

【 図 9 - 2 】



【 図 10 】

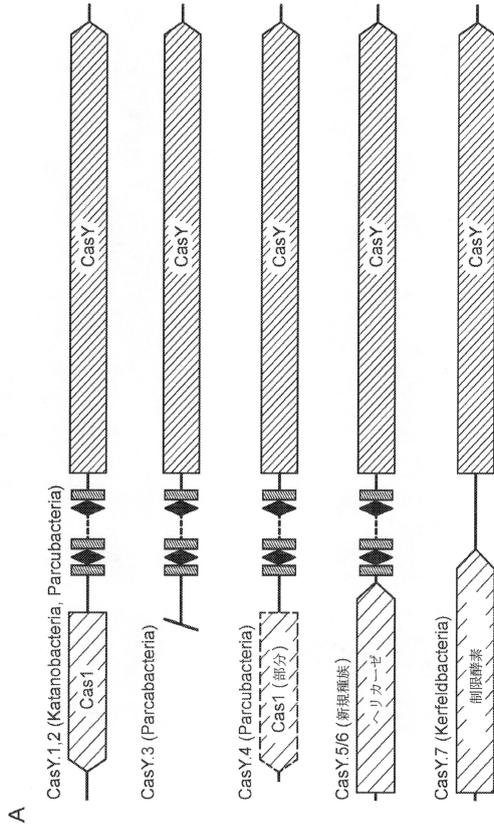


30

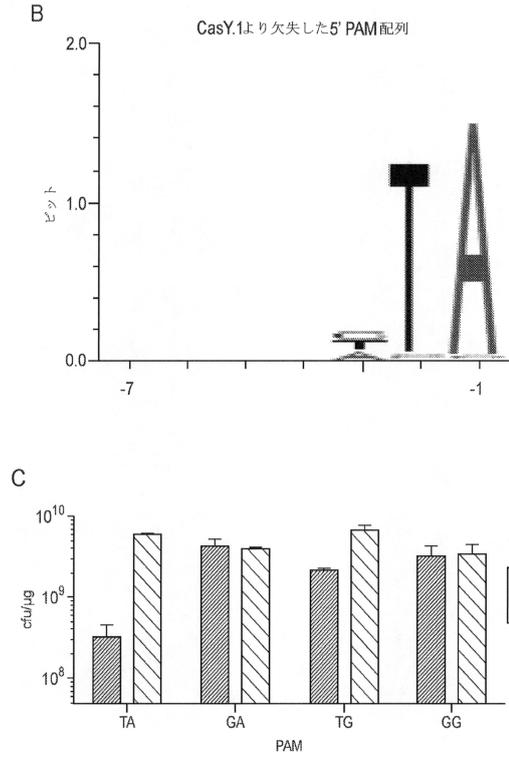
40

50

【 図 1 1 - 1 】



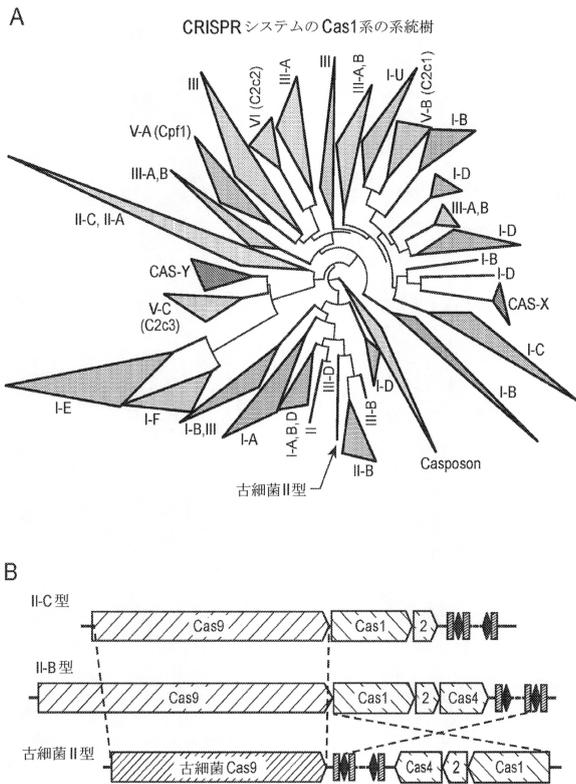
【 図 1 1 - 2 】



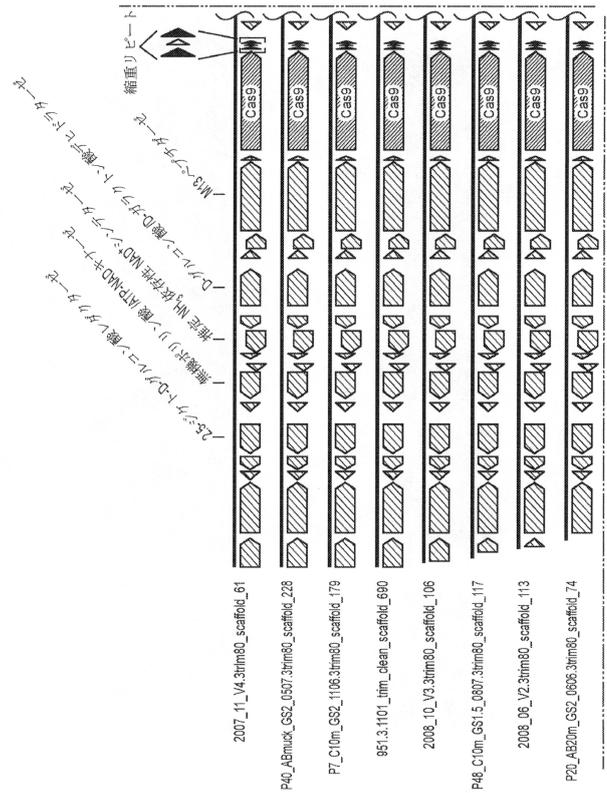
10

20

【 図 1 2 】



【 図 1 3 - 1 】

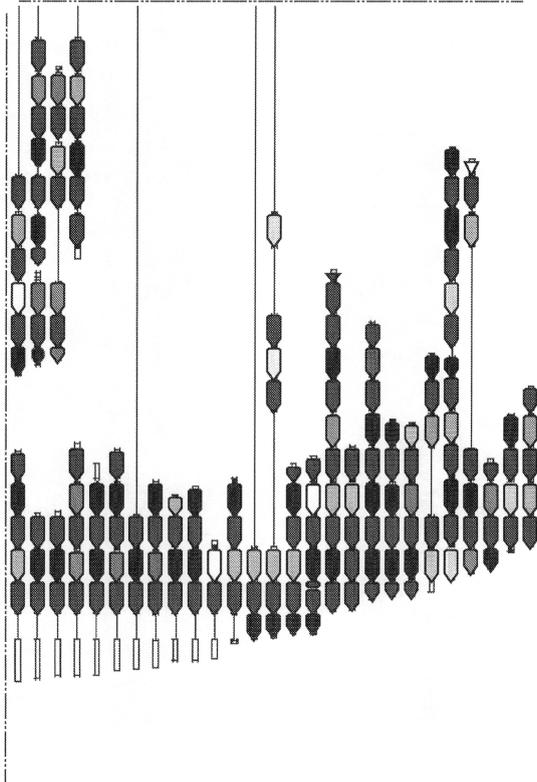


30

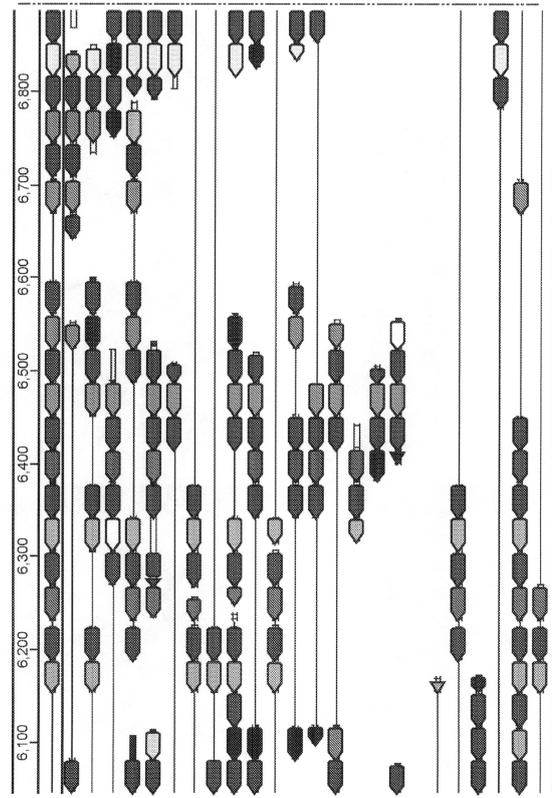
40

50

【 図 1 4 - 2 】



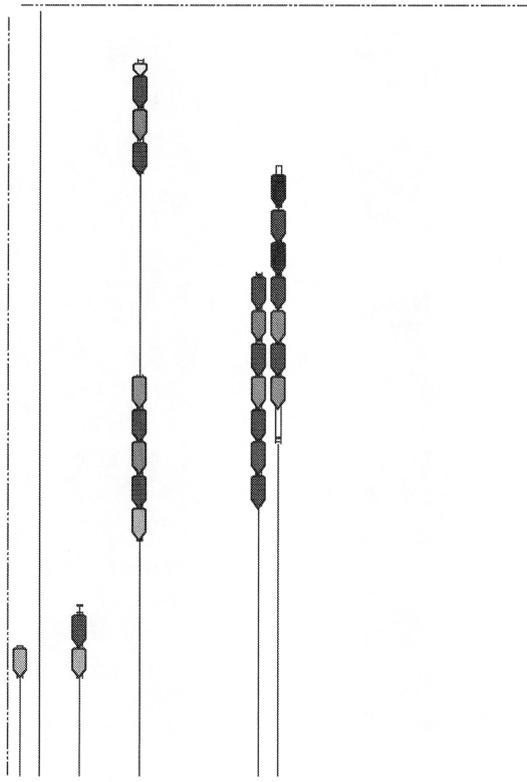
【 図 1 4 - 3 】



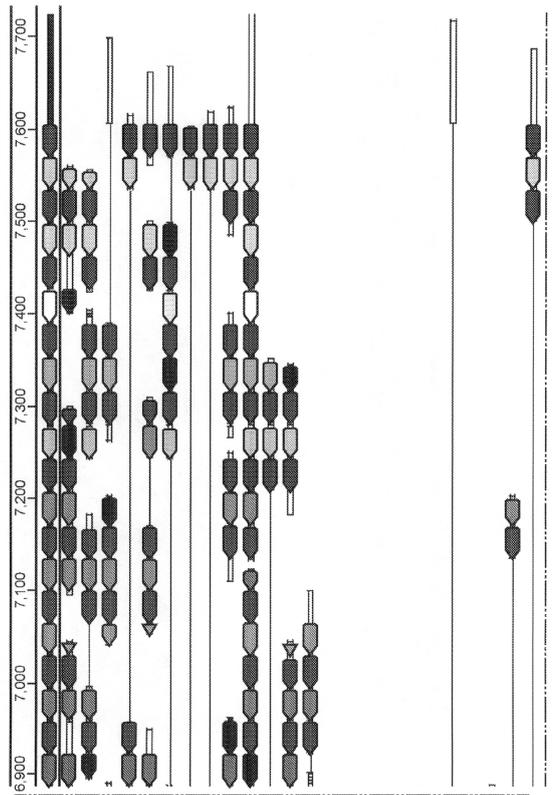
10

20

【 図 1 4 - 4 】



【 図 1 4 - 5 】

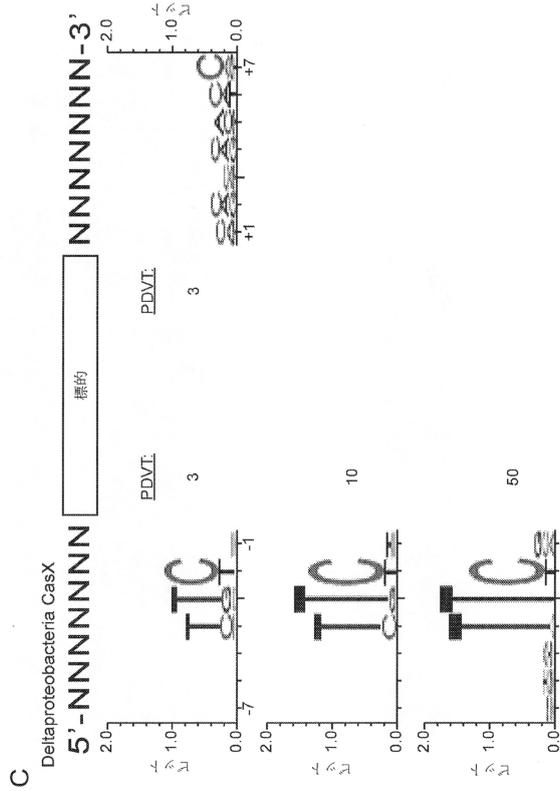


30

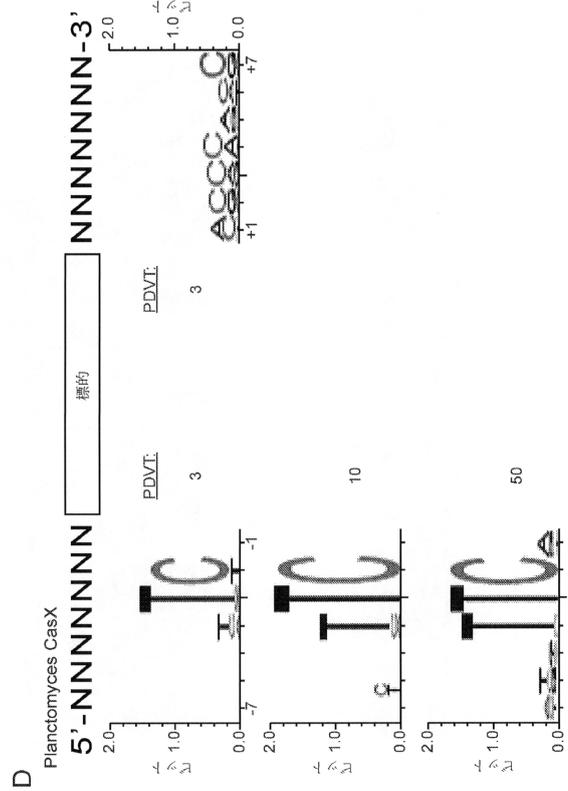
40

50

【 図 19 - 2 】



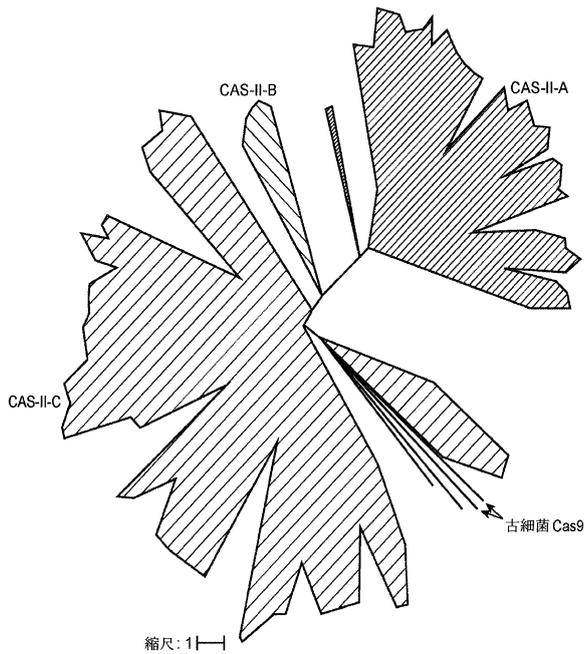
【 図 19 - 3 】



10

20

【 図 20 】



【 図 21 】

ARMAN-1由来 Cas9 をアッセイする際の *In vitro* 切断条件

タンパク質精製	緩衝液	塩 (mM)	金属	ガイド	標的	温度
AR1-Cas9 #1	Tris pH 7.5	300	Mg ²⁺ Mn ²⁺ Zn ²⁺	cr:69 cr:104 cr:179	dsDNA ssDNA DNA-パプレ ssRNA dsDNA	37
AR1-Cas9 #1	Tris pH 7.5	100-500	Mg ²⁺	cr:69 cr:104 cr:179	dsDNA	37
AR1-Cas9 #1	Tris pH 7.5	300	Mg ²⁺ Mn ²⁺ Zn ²⁺	cr:69 cr:104 cr:179	dsDNA	30-48
AR1-Cas9 #1	MOPS pH 6 pH 6.5 pH 7.0 pH 7.5	300	Mg ²⁺	cr:69 cr:104 cr:179	dsDNA	37
AR1-Cas9 #1	ヒン酸 pH 5 pH 5.5 pH 6	300	Mg ²⁺	cr:69 cr:104 cr:179	dsDNA	37
AR1-Cas9 #1	Tris pH 7.5	300	Mg ²⁺ Mn ²⁺ Zn ²⁺	cr:69 cr:104 cr:179	プラスミド	37-50
AR1-Cas9 #2	Tris pH 7.5	300	Mg ²⁺ Mn ²⁺ Zn ²⁺	cr:69 cr:104 cr:179	dsDNA	37
AR1-Cas9 #3	Tris pH 7.5	300	Mg ²⁺ Mn ²⁺ Zn ²⁺	cr:69 cr:104 cr:179	dsDNA	37
AR1-Cas9 #4	Tris pH 7.5	300	Mg ²⁺ Mn ²⁺ Zn ²⁺	cr:69 cr:104 cr:179	dsDNA	37
AR1-Cas9 #5	Tris pH 7.5	300	Mg ²⁺ Mn ²⁺ Zn ²⁺	cr:69 cr:104 cr:179	dsDNA	37
AR1-Cas9 #6	Tris pH 7.5	300	Mg ²⁺ Mn ²⁺ Zn ²⁺	cr:69 cr:104 cr:179	ssDNA dsDNA	37
AR4-Cas9 #1	Tris pH 7.5	300	Mg ²⁺ Mn ²⁺ Zn ²⁺	sgRNA-122	dsDNA	37
AR4-Cas9 #2	Tris pH 7.5	300	Mg ²⁺ Mn ²⁺ Zn ²⁺	sgRNA-122	dsDNA	37
AR4-Cas9 #3	Tris pH 7.5	300	Mg ²⁺ Mn ²⁺ Zn ²⁺	sgRNA-122	dsDNA	37
AR4-Cas9 #4	Tris pH 7.5	300	Mg ²⁺ Mn ²⁺ Zn ²⁺	sgRNA-122	dsDNA	37

30

40

50

【配列表】

2023030067000001.app

10

20

30

40

50

フロントページの続き

(72)発明者 アメリカ合衆国 9 4 7 0 5 カリフォルニア州, バークリー, ヴィチェンテ ロード 1 6 4
バンフィールド, ジリアン エフ.

(72)発明者 アメリカ合衆国 9 4 7 0 8 カリフォルニア州, バークリー, ヒル ロード 1 6 0
バースタイン, デイビッド

(72)発明者 アメリカ合衆国 9 4 7 2 0 カリフォルニア州, バークリー, マッコーン ホール 3 0 7

(72)発明者 ハリントン, ルーカス ベンジャミン
アメリカ合衆国 9 4 7 2 0 カリフォルニア州, バークリー, スタンリー ホール 7 3 1, エム
エス シャープ 3 2 2 0