

(19)日本国特許庁(JP)

(12)公表特許公報(A)

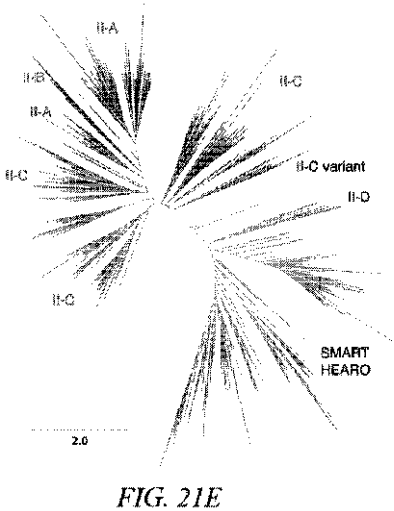
(11)公表番号
特表2024-541402
(P2024-541402A)

(43)公表日 令和6年11月8日(2024.11.8)

(51)国際特許分類		F I		テーマコード (参考)	
C 1 2 N	9/16 (2006.01)	C 1 2 N	9/16	Z Z N A	4 B 0 6 5
C 1 2 N	15/55 (2006.01)	C 1 2 N	15/55		
C 1 2 N	15/09 (2006.01)	C 1 2 N	15/09	1 0 0	
C 1 2 N	15/11 (2006.01)	C 1 2 N	15/11	Z	
C 1 2 N	15/63 (2006.01)	C 1 2 N	15/63	Z	
		審査請求	未請求	予備審査請求	未請求 (全229頁) 最終頁に続く
(21)出願番号	特願2024-529294(P2024-529294)			(71)出願人	522185863
(86)(22)出願日	令和4年11月23日(2022.11.23)				メタゲノミ, インク.
(85)翻訳文提出日	令和6年7月11日(2024.7.11)				アメリカ合衆国 9 4 6 0 8 カリフォル
(86)国際出願番号	PCT/US2022/080437				ニア州 エメリービル パーク・アベニュー
(87)国際公開番号	WO2023/097282				ー 1 5 4 5
(87)国際公開日	令和5年6月1日(2023.6.1)			(74)代理人	110003797
(31)優先権主張番号	63/282,999				弁理士法人清原国際特許事務所
(32)優先日	令和3年11月24日(2021.11.24)			(72)発明者	トーマス, ブライアン シー.
(33)優先権主張国・地域又は機関	米国(US)				アメリカ合衆国 9 4 6 0 8 カリフォル
(31)優先権主張番号	63/289,981				ニア州 エメリービル パーク・アベニュー
(32)優先日	令和3年12月15日(2021.12.15)			(72)発明者	ー 1 5 4 5
(33)優先権主張国・地域又は機関	米国(US)				ブラウン, クリストファー
(31)優先権主張番号	63/356,908				アメリカ合衆国 9 4 6 0 8 カリフォル
	最終頁に続く				ニア州 エメリービル パーク・アベニュー
					ー 1 5 4 5
					最終頁に続く

(54)【発明の名称】 エンドヌクレアーゼ系

(57)【要約】
本開示は、S M A l l A R c h a e a l - a s s o c i a T e d (S M A R T) ヌクレアーゼ系を含む、未培養微生物に由来するエンドヌクレアーゼ酵素系、及びそのような酵素又はそのバリエントを使用する方法を提供する。
【選択図】 図 2 1 の E



【特許請求の範囲】

【請求項 1】

操作されたヌクレアーゼ系であって、

(a) R u v Cドメイン及びH N Hドメインを含むエンドヌクレアーゼであって、前記エンドヌクレアーゼが、未培養微生物に由来する、エンドヌクレアーゼと、

(b) 前記エンドヌクレアーゼと複合体を形成するように構成された、操作されたガイドリボ核酸構造であって、

(i) 標的デオキシリボ核酸配列にハイブリダイズするように構成されたガイドリボ核酸配列、及び

(i i) 前記エンドヌクレアーゼに結合するように構成されたリボ核酸配列を含む、操作されたガイドリボ核酸構造と、を含み、

前記エンドヌクレアーゼが、約96 k D a以下、約80 k D a以下、約70 k D a以下、又は約60 k D a以下の分子量を有し、

(1) 前記エンドヌクレアーゼが、配列番号1～198、221～459、463～612、617～668、674～675、975～1002、1260～1321、若しくはそれらのバリエーションのうちのいずれか1つからのアルギニンリッチ領域若しくはP F 1 4 2 3 9相同性を有するドメインと少なくとも70%、少なくとも75%、少なくとも80%、若しくは少なくとも90%の配列同一性を有するアルギニンリッチ領域若しくはP F 1 4 2 3 9相同性を有するドメインを含むか、

(2) 前記エンドヌクレアーゼが、配列番号1～198、221～459、463～612、617～668、674～675、975～1002、1260～1321、若しくはそれらのバリエーションのうちのいずれか1つからのR E Cドメインと少なくとも70%、少なくとも75%、少なくとも80%、若しくは少なくとも90%の配列同一性を有するR E Cドメインを含むか、又は

(3) 前記エンドヌクレアーゼが、配列番号1～198、221～459、463～612、617～668、674～675、975～1002、1260～1321、若しくはそれらのバリエーションのうちのいずれか1つと少なくとも70%、少なくとも75%、少なくとも80%、若しくは少なくとも90%の配列同一性を有する配列を含む、操作されたヌクレアーゼ系。

【請求項 2】

前記エンドヌクレアーゼが、古細菌エンドヌクレアーゼである、請求項1に記載の操作されたヌクレアーゼ系。

【請求項 3】

前記エンドヌクレアーゼが、配列番号1～198、221～459、463～612、617～668、674～675、975～1002、1260～1321のうちのいずれか1つと少なくとも70%、少なくとも75%、少なくとも80%、又は少なくとも90%の配列同一性を有する配列を含む、請求項1又は2に記載の操作されたヌクレアーゼ系。

【請求項 4】

前記エンドヌクレアーゼが、R R x R Rモチーフを含むアルギニンリッチ領域又はP F 1 4 2 3 9相同性を有するドメインを更に含む、請求項1～3のいずれか一項に記載の操作されたヌクレアーゼ系。

【請求項 5】

前記アルギニンリッチ領域又はP F 1 4 2 3 9相同性を有する前記ドメインが、配列番号1～198、221～459、463～612、617～668、674～675、975～1002、1260～1321のうちのいずれか1つの前記アルギニンリッチ領域又はP F 1 4 2 3 9を有する前記ドメインと少なくとも85%、少なくとも90%、又は少なくとも95%の同一性を有する、請求項4に記載の操作されたヌクレアーゼ系。

【請求項 6】

前記エンドヌクレアーゼが、R E C（認識）ドメインを更に含む、請求項1～5のい

10

20

30

40

50

れか一項に記載の操作されたヌクレアーゼ系。

【請求項 7】

前記 R E C ドメインが、配列番号 1～198、221～459、463～612、617～668、674～675、975～1002、1260～1321 のうちのいずれか 1 つの R E C ドメインと少なくとも 85 %、少なくとも 90 %、又は少なくとも 95 % の同一性を有する、請求項 6 に記載の操作されたヌクレアーゼ系。

【請求項 8】

前記エンドヌクレアーゼが、B H（ブリッジヘリックス）ドメイン、W E D（ウェッジ）ドメイン、又は P I（P A M 相互作用）若しくは T I（T A M 相互作用）ドメインを更に含む、請求項 1～7 のいずれか一項に記載の操作されたヌクレアーゼ系。

10

【請求項 9】

前記 B H ドメイン、前記 W E D ドメイン、又は前記 P I ドメインが、配列番号 1～198、221～459、463～612、617～668、674～675、975～1002、1260～1321 のうちのいずれか 1 つの B H ドメイン、W E D ドメイン、又は P I ドメインと少なくとも 85 %、少なくとも 90 %、又は少なくとも 95 % の同一性を有する、請求項 8 に記載の操作されたヌクレアーゼ系。

【請求項 10】

操作されたヌクレアーゼ系であって、

（a）R u v C－I ドメイン及び H N H ドメインを含むエンドヌクレアーゼと、

（b）前記エンドヌクレアーゼと複合体を形成するように構成された、操作されたガイドリボ核酸構造であって、

20

（i）標的デオキシリボ核酸配列にハイブリダイズするように構成されたガイドリボ核酸配列、及び

（i i）前記エンドヌクレアーゼに結合するように構成されたりボ核酸配列を含む、操作されたガイドリボ核酸構造と、を含み、

前記エンドヌクレアーゼが、配列番号 674～675、975～1002、1260～1321、又はそれらのバリエーションのうちのいずれか 1 つと少なくとも 70 %、少なくとも 75 %、少なくとも 80 %、又は少なくとも 90 % の配列同一性を有する配列を含む、操作されたヌクレアーゼ系。

【請求項 11】

前記エンドヌクレアーゼが、古細菌エンドヌクレアーゼである、請求項 10 に記載の操作されたヌクレアーゼ系。

30

【請求項 12】

前記エンドヌクレアーゼが、R R x R R モチーフを含むアルギニンリッチ領域又は P F 14239 相同性を有するドメインを更に含む、請求項 10 又は 11 に記載の操作されたヌクレアーゼ系。

【請求項 13】

前記アルギニンリッチ領域又は P F 14239 相同性を有する前記ドメインが、配列番号 674～675、975～1002、1260～1321 のうちのいずれか 1 つのアルギニンリッチ領域と少なくとも 85 %、少なくとも 90 %、又は少なくとも 95 % の同一性を有する、請求項 12 に記載の操作されたヌクレアーゼ系。

40

【請求項 14】

前記エンドヌクレアーゼが、R E C（認識）ドメインを更に含む、請求項 10～13 のいずれか一項に記載の操作されたヌクレアーゼ系。

【請求項 15】

前記 R E C ドメインが、配列番号 674～675、975～1002、1260～1321 のうちのいずれか 1 つの R E C ドメインと少なくとも 85 %、少なくとも 90 %、又は少なくとも 95 % の同一性を有する、請求項 14 に記載の操作されたヌクレアーゼ系。

【請求項 16】

前記エンドヌクレアーゼが、B H ドメイン、W E D ドメイン、及び P I ドメインを更に

50

含む、請求項 10～15 のいずれか一項に記載の操作されたヌクレアーゼ系。

【請求項 17】

前記 B H ドメイン、前記 W E D ドメイン、又は前記 P I ドメインが、配列番号 674～675、975～1002、1260～1321 のうちのいずれか 1 つの B H ドメイン、W E D ドメイン、又は P I ドメインと少なくとも 85 %、少なくとも 90 %、又は少なくとも 95 % の同一性を有する、請求項 16 に記載の操作されたヌクレアーゼ系。

【請求項 18】

前記エンドヌクレアーゼが、培養されていない微生物に由来する、請求項 10～17 のいずれか一項に記載の操作されたヌクレアーゼ系。

【請求項 19】

前記エンドヌクレアーゼに結合するように構成された前記リボ核酸配列が、配列番号 199～200、460～461、若しくは 669～673 のうちのいずれか 1 つと少なくとも 80 % の配列同一性を有する配列、又は配列番号 201～203、613～616、677～686、1003～1022、若しくは 1231～1259 のうちのいずれか 1 つの非縮重ヌクレオチドと少なくとも 80 % の配列同一性を有する配列を含む、請求項 1～18 のいずれか一項に記載の操作されたヌクレアーゼ系。

【請求項 20】

前記ガイド核酸構造が、配列番号 201～203、613～616、677～686、1003～1022、又は 1231～1259 のうちのいずれか 1 つの非縮重ヌクレオチドと少なくとも 80 % の同一性を有する配列を含む、請求項 19 に記載の操作されたヌクレアーゼ系。

【請求項 21】

操作されたヌクレアーゼ系であって、

(a) 操作されたガイドリボ核酸構造であって、

(i) 標的デオキシリボ核酸配列にハイブリダイズするように構成されたガイドリボ核酸配列、及び

(i i) エンドヌクレアーゼに結合するように構成されたリボ核酸配列を含み、前記リボ核酸配列が、配列番号 199～200、460～461、若しくは 669～673 のうちのいずれか 1 つと少なくとも 80 % の配列同一性を有する配列、又は配列番号 677～686、1006～1012、若しくは 1231～1259 のうちのいずれか 1 つの非可変ヌクレオチドと少なくとも 80 % の配列同一性を有する配列を含む、操作されたガイドリボ核酸構造と、

(b) 前記操作されたガイドリボ核酸に結合するように構成された R N A ガイドエンドヌクレアーゼと、を含む、操作されたヌクレアーゼ系。

【請求項 22】

前記 R N A ガイドエンドヌクレアーゼが、古細菌エンドヌクレアーゼである、請求項 21 に記載の操作されたヌクレアーゼ系。

【請求項 23】

前記エンドヌクレアーゼが、約 120 k D a 以下、100 k D a 以下、90 k D a 以下、又は 60 k D a 以下の分子量を有する、請求項 21 又は 22 に記載の操作されたヌクレアーゼ系。

【請求項 24】

前記操作されたガイドリボ核酸構造が、少なくとも 2 つのリボ核酸ポリヌクレオチドを含む、請求項 1～23 のいずれか一項に記載の操作されたヌクレアーゼ系。

【請求項 25】

前記操作されたガイドリボ核酸構造が、前記エンドヌクレアーゼに結合するように構成された前記ガイドリボ核酸配列及び前記リボ核酸配列を含む単一のリボ核酸ポリヌクレオチドを含む、請求項 1～24 のいずれか一項に記載の操作されたヌクレアーゼ系。

【請求項 26】

前記ガイドリボ核酸配列が、原核生物、細菌、古細菌、真核生物、真菌、植物、哺乳類

10

20

30

40

50

、又はヒトゲノム配列と相補的である、請求項 1～25 のいずれか一項に記載の操作されたヌクレアーゼ系。

【請求項 27】

前記ガイドリボ核酸配列が、約 14～約 28 ヌクレオチド長、約 18～約 26 ヌクレオチド長、約 22～約 26 ヌクレオチド長、又は約 24 ヌクレオチド長である、請求項 1～26 のいずれか一項に記載の操作されたヌクレアーゼ系。

【請求項 28】

前記ガイドリボ核酸配列が、配列番号 462、676、又は 1229～1230 のうちのいずれか 1つと少なくとも 80% の配列同一性を有する配列を含む、請求項 1～27 のいずれか一項に記載の操作されたヌクレアーゼ系。

10

【請求項 29】

前記エンドヌクレアーゼが、前記エンドヌクレアーゼの N 末端又は C 末端の近位に 1つ以上の核局在化配列 (NLS) を含む、請求項 1～28 のいずれか一項に記載の操作されたヌクレアーゼ系。

【請求項 30】

前記 NLS が、配列番号 205～220 のうちのいずれか 1つから選択される配列を含む、請求項 1～29 のいずれか一項に記載の操作されたヌクレアーゼ系。

【請求項 31】

5' から 3' の方向に、前記標的デオキシリボ核酸配列に対して 5' 側の、少なくとも 20 個のヌクレオチドの配列を含む第 1 の相同アーム、少なくとも 10 個のヌクレオチドの合成 DNA 配列、及び前記標的配列に対して 3' 側の、少なくとも 20 個のヌクレオチドの配列を含む第 2 の相同アームを含む、一本鎖又は二本鎖の DNA 修復鋳型を更に含む、請求項 1～30 のいずれか一項に記載の操作されたヌクレアーゼ系。

20

【請求項 32】

前記第 1 又は第 2 の相同アームが、少なくとも 40、80、120、150、200、300、500、又は 1,000 個のヌクレオチドの配列を含む、請求項 31 に記載の操作されたヌクレアーゼ系。

【請求項 33】

前記系が、 Mg^{2+} の供給源を更に含む、請求項 1～32 のいずれか一項に記載の操作されたヌクレアーゼ系。

30

【請求項 34】

前記エンドヌクレアーゼ及び前記エンドヌクレアーゼに結合するように構成された前記リボ核酸配列が、同じ門内の別個の種に由来する、請求項 1～33 のいずれか一項に記載の操作されたヌクレアーゼ系。

【請求項 35】

前記エンドヌクレアーゼが、配列番号 2～24 のうちのいずれか 1つと少なくとも 70% の配列同一性を有する配列を含み、前記ガイド RNA 構造が、ステム及びループを含むヘアピンを含むと予測される RNA 配列を含み、前記ステムが、少なくとも 10 対のリボヌクレオチド及び介在するマルチループを含む、請求項 1～34 のいずれか一項に記載の操作されたヌクレアーゼ系。

40

【請求項 36】

前記ガイド RNA 構造が、第 2 のステム及び第 2 のループを更に含み、前記第 2 のステムが、少なくとも 5 対のリボヌクレオチドを含む、請求項 35 に記載の操作されたエンドヌクレアーゼ系。

【請求項 37】

前記ガイド RNA 構造が、少なくとも 2 つのヘアピンを含む RNA 構造を更に含む、請求項 35 に記載の操作されたエンドヌクレアーゼ系。

【請求項 38】

a) 前記エンドヌクレアーゼが、配列番号 1～198、221～459、463～612、617～668、674～675、975～1002、1260～1321、又はそ

50

これらのバリエーションのうちのいずれか1つと少なくとも70%、少なくとも80%、又は少なくとも90%同一の配列を含み、

b) 前記ガイドRNA構造が、配列番号199~200、460~461、若しくは669~673のうちのいずれか1つ、又は配列番号201~203、613~616、677~686、1006~1012、若しくは1231~1259のうちのいずれか1つの非可変ヌクレオチドと少なくとも70%、少なくとも80%、又は少なくとも90%同一の配列を含む、請求項1~37のいずれか一項に記載の操作されたヌクレアーゼ系。

【請求項39】

a) 前記エンドヌクレアーゼが、配列番号1~198、221~459、463~612、617~668、674~675、975~1002、1260~1321のうちのいずれか1つと少なくとも70%、少なくとも80%、又は少なくとも90%同一の配列を含み、

10

b) 前記ガイドRNA構造が、クラス2、II型sgRNA又はtracr配列と少なくとも70%、少なくとも80%、又は少なくとも90%同一の配列を含む、請求項1~37のいずれか一項に記載の操作されたヌクレアーゼ系。

【請求項40】

前記配列同一性が、BLASTP、CLUSTALW、MUSCLE、MAFFT、又はSmith-Waterman相同性検索アルゴリズムのパラメータを用いたCLUSTALWによって決定される、請求項1~39のいずれか一項に記載の操作されたヌクレアーゼ系。

20

【請求項41】

前記配列同一性が、3のワード長(W)、10の期待値(E)のパラメータ、及び11の存在、1の延長でギャップコストを設定しているBLOSUM62スコアリングマトリックスを使用し、かつ条件付き組成スコアマトリックス調整を使用している、前記BLASTP相同性検索アルゴリズムによって決定される、請求項40に記載の操作されたヌクレアーゼ系。

【請求項42】

前記エンドヌクレアーゼが、Cas9エンドヌクレアーゼ、Cas14エンドヌクレアーゼ、Cas12aエンドヌクレアーゼ、Cas12bエンドヌクレアーゼ、Cas12cエンドヌクレアーゼ、Cas12dエンドヌクレアーゼ、Cas12eエンドヌクレアーゼ、Cas13aエンドヌクレアーゼ、Cas13bエンドヌクレアーゼ、Cas13cエンドヌクレアーゼ、又はCas13dエンドヌクレアーゼではない、請求項1~41のいずれか一項に記載の操作されたヌクレアーゼ系。

30

【請求項43】

前記エンドヌクレアーゼが、Cas9エンドヌクレアーゼと80%未満の同一性を有する、請求項42に記載の操作されたヌクレアーゼ系。

【請求項44】

操作されたヌクレアーゼ系であって、

(a) ANGG (配列番号1029)、NARAA (配列番号1030)、ATGAA (配列番号1031)、ATGA (配列番号1032)、又はWTGG (配列番号1033)のうちのいずれか1つを含む標的隣接モチーフ(TAM)配列に対して選択的であるように構成されたエンドヌクレアーゼであって、前記エンドヌクレアーゼが、配列番号1~198、221~459、463~612、617~668、674~675、975~1002、1260~1321のうちのいずれか1つのTAM相互作用ドメインと少なくとも75%、80%、又は90%の同一性を有するTAM相互作用ドメインを含む、エンドヌクレアーゼと、

40

(b) 操作されたガイドRNAであって、前記操作されたガイドRNAが、前記エンドヌクレアーゼと複合体を形成するように構成されており、標的核酸配列にハイブリダイズするように構成されたスパーサー配列を含む、操作されたガイドRNAと、を含む、操作されたヌクレアーゼ系。

50

【請求項 45】

前記 T A M 相互作用ドメインが、配列番号 674 の T A M 相互作用ドメイン若しくはそのバリエーションと少なくとも 80% の配列同一性、又は配列番号 675 の T A M 相互作用ドメイン若しくはそのバリエーションと少なくとも 80% の配列同一性を有する配列を含む、請求項 44 に記載の操作されたヌクレアーゼ系。

【請求項 46】

前記エンドヌクレアーゼ系が、真核生物、真菌、植物、哺乳類、又はヒトゲノムポリヌクレオチド配列と相補的な配列を含む、請求項 44 又は 45 に記載の操作されたヌクレアーゼ系。

【請求項 47】

前記ガイド RNA が、30～280 ヌクレオチド長である、請求項 44～46 のいずれか一項に記載の操作されたヌクレアーゼ系。

【請求項 48】

5' から 3' の方向に、前記標的デオキシリボ核酸配列に対して 5' 側の、少なくとも 20 個のヌクレオチドの配列を含む第 1 の相同アーム、少なくとも 10 個のヌクレオチドの合成 DNA 配列、及び前記標的配列に対して 3' 側の、少なくとも 20 個のヌクレオチドの配列を含む第 2 の相同アームを含む、一本鎖又は二本鎖の DNA 修復鋳型を更に含む、請求項 44～47 のいずれか一項に記載の操作されたヌクレアーゼ系。

【請求項 49】

前記第 1 又は第 2 の相同アームが、少なくとも 40 個のヌクレオチドの配列を含む、請求項 48 に記載の操作されたヌクレアーゼ系。

【請求項 50】

前記第 1 及び第 2 の相同アームが、真核生物のゲノム配列に対して相同である、請求項 48 又は 49 に記載の操作されたヌクレアーゼ系。

【請求項 51】

前記一本鎖又は二本鎖 DNA 修復鋳型が、導入遺伝子ドナーを含む、請求項 48～50 のいずれか一項に記載の操作されたヌクレアーゼ系。

【請求項 52】

1 つ又は 2 つの一本鎖 DNA セグメントに隣接した二本鎖 DNA セグメントを含む DNA 修復鋳型を更に含む、請求項 44～51 のいずれか一項に記載の操作されたヌクレアーゼ系。

【請求項 53】

(前記一本鎖 DNA セグメントが、前記二本鎖 DNA セグメントの 5' 端にコンジュゲートされている、請求項 52 に記載の操作されたヌクレアーゼ系。

【請求項 54】

前記一本鎖 DNA セグメントが、前記二本鎖 DNA セグメントの 3' 端にコンジュゲートされている、請求項 52 に記載の操作されたヌクレアーゼ系。

【請求項 55】

前記一本鎖 DNA セグメントが、4～10 個のヌクレオチド塩基の長さを有する、請求項 52～54 のいずれか一項に記載の操作されたヌクレアーゼ系。

【請求項 56】

前記一本鎖 DNA セグメントが、前記スペーサー配列内の配列と相補的なヌクレオチド配列を有する、請求項 52～55 のいずれか一項に記載の操作されたヌクレアーゼ系。

【請求項 57】

前記二本鎖 DNA 配列が、バーコード、オープンリーディングフレーム、エンハンサー、プロモーター、タンパク質コード配列、miRNA コード配列、RNA コード配列、又は導入遺伝子を含む、請求項 52～56 のいずれか一項に記載の操作されたヌクレアーゼ系。

【請求項 58】

前記二本鎖 DNA 配列が、ヌクレアーゼ切断部位に隣接している、請求項 52～57 の

10

20

30

40

50

いずれか一項に記載の操作されたヌクレアーゼ系。

【請求項 59】

操作されたヌクレアーゼ系であって、

(a) NRRを含むプロトスペーサー隣接モチーフ(PAM)配列に対して選択的であるように構成されたエンドヌクレアーゼであって、前記エンドヌクレアーゼが、配列番号1313~1318のうちのいずれか1つのPAM相互作用ドメインと少なくとも75%、80%、又は90%の同一性を有するPAM相互作用ドメインを含む、エンドヌクレアーゼと、

(b) 操作されたガイドRNAであって、前記操作されたガイドRNAが、前記エンドヌクレアーゼと複合体を形成するように構成されており、標的核酸配列にハイブリダイズするように構成されたスペーサー配列を含む、操作されたガイドRNAと、を含む、操作されたヌクレアーゼ系。

10

【請求項 60】

前記TAM相互作用ドメインが、配列番号674のTAM相互作用ドメイン若しくはそのバリエーションと少なくとも80%の配列同一性、又は配列番号675のTAM相互作用ドメイン若しくはそのバリエーションと少なくとも80%の配列同一性を有する配列を含む、請求項59に記載の操作されたヌクレアーゼ系。

【請求項 61】

前記エンドヌクレアーゼ系が、真核生物、真菌、植物、哺乳類、又はヒトゲノムポリヌクレオチド配列と相補的な配列を含む、請求項59又は60に記載の操作されたヌクレアーゼ系。

20

【請求項 62】

前記ガイドRNAが、30~280ヌクレオチド長である、請求項59~61のいずれか一項に記載の操作されたヌクレアーゼ系。

【請求項 63】

5'から3'の方向に、前記標的デオキシリボ核酸配列に対して5'側の、少なくとも20個のヌクレオチドの配列を含む第1の相同アーム、少なくとも10個のヌクレオチドの合成DNA配列、及び前記標的配列に対して3'側の、少なくとも20個のヌクレオチドの配列を含む第2の相同アームを含む、一本鎖又は二本鎖のDNA修復鋳型を更に含む、請求項59~62のいずれか一項に記載の操作されたヌクレアーゼ系。

30

【請求項 64】

前記第1又は第2の相同アームが、少なくとも40個のヌクレオチドの配列を含む、請求項63に記載の操作されたヌクレアーゼ系。

【請求項 65】

前記第1及び第2の相同アームが、真核生物のゲノム配列に対して相同である、請求項63又は64に記載の操作されたヌクレアーゼ系。

【請求項 66】

前記一本鎖又は二本鎖DNA修復鋳型が、導入遺伝子ドナーを含む、請求項59~65のいずれか一項に記載の操作されたヌクレアーゼ系。

【請求項 67】

1つ又は2つの一本鎖DNAセグメントに隣接した二本鎖DNAセグメントを含むDNA修復鋳型を更に含む、請求項59~66のいずれか一項に記載の操作されたヌクレアーゼ系。

40

【請求項 68】

前記一本鎖DNAセグメントが、前記二本鎖DNAセグメントの5'端にコンジュゲートされている、請求項67に記載の操作されたヌクレアーゼ系。

【請求項 69】

前記一本鎖DNAセグメントが、前記二本鎖DNAセグメントの3'端にコンジュゲートされている、請求項67に記載の操作されたヌクレアーゼ系。

【請求項 70】

50

前記一本鎖DNAセグメントが、4～10個のヌクレオチド塩基の長さを有する、請求項67～69のいずれか一項に記載の操作されたヌクレアーゼ系。

【請求項71】

前記一本鎖DNAセグメントが、前記スペーサー配列内の配列と相補的なヌクレオチド配列を有する、請求項67～70のいずれか一項に記載の操作されたヌクレアーゼ系。

【請求項72】

前記二本鎖DNA配列が、バーコード、オープンリーディングフレーム、エンハンサー、プロモーター、タンパク質コード配列、miRNAコード配列、RNAコード配列、又は導入遺伝子を含む、請求項67～71のいずれか一項に記載の操作されたヌクレアーゼ系。

【請求項73】

前記二本鎖DNA配列が、ヌクレアーゼ切断部位に隣接している、請求項67～72のいずれか一項に記載の操作されたヌクレアーゼ系。

【請求項74】

操作された単一ガイドリボ核酸ポリヌクレオチドであって、

a) 標的DNA分子中の標的配列と相補的であるヌクレオチド配列を含むDNA標的化セグメントと、

b) ハイブリダイズして二本鎖RNA(dsRNA)二重鎖を形成するヌクレオチドの2つの相補的なストレッチを含む、タンパク質結合セグメントと、を含み、

前記ヌクレオチドの2つの相補的なストレッチが、介在するヌクレオチドと互いに共有結合的に連結しており、

前記操作されたガイドリボ核酸ポリヌクレオチドが、配列番号674～675、975～1002、1260～1321のうちのいずれか1つ、又はそれらのバリエーションと少なくとも75%の配列同一性を有するバリエーションを含むエンドヌクレアーゼと複合体を形成するように構成されている、操作された単一ガイドリボ核酸ポリヌクレオチド。

【請求項75】

前記DNA標的化セグメントが、前記ヌクレオチドの2つの相補的なストレッチの両方の5'側に位置付けられている、請求項74に記載の操作されたガイドリボ核酸ポリヌクレオチド。

【請求項76】

a) 前記タンパク質結合セグメントが、配列番号199～200、460～461、若しくは669～673のうちのいずれか1つと少なくとも70%、少なくとも80%、若しくは少なくとも90%同一性を有する配列を含むか、又は

b) 前記タンパク質結合セグメントが、配列番号201～203、613～616、677～686、1003～1022、若しくは1231～1259のうちのいずれか1つの非可変ヌクレオチドと少なくとも70%、少なくとも80%、若しくは少なくとも90%同一性を有する配列を含む、請求項74又は75に記載の操作されたガイドリボ核酸ポリヌクレオチド。

【請求項77】

a) 前記エンドヌクレアーゼが、配列番号674～675、975～1002、1260～1321、又はそれらのバリエーションのうちのいずれか1つと少なくとも70%、少なくとも80%、又は少なくとも90%同一性の配列を含む、

b) 前記ガイドRNA構造が、クラス2、II型sgRNAと少なくとも70%、少なくとも80%、又は少なくとも90%同一性の配列を含む、請求項74又は75に記載の操作されたガイドリボ核酸ポリヌクレオチド。

【請求項78】

前記エンドヌクレアーゼが、前記エンドヌクレアーゼに結合された塩基エディタ又はヒストンエディタを更に含む、請求項1～77のいずれか一項に記載の操作されたガイドリボ核酸ポリヌクレオチド。

【請求項79】

10

20

30

40

50

前記塩基エディタが、アデノシンデアミナーゼである、請求項 78 に記載の操作されたガイドリボ核酸ポリヌクレオチド。

【請求項 80】

前記アデノシンデアミナーゼが、ADAR1 又は ADAR2 を含む、請求項 79 に記載の操作されたガイドリボ核酸ポリヌクレオチド。

【請求項 81】

前記塩基エディタが、シトシンデアミナーゼである、請求項 78 に記載の操作されたガイドリボ核酸ポリヌクレオチド。

【請求項 82】

前記シトシンデアミナーゼが、APOBEC1、APOBEC2、APOBEC3A、APOBEC3B、APOBEC3C、APOBEC3D、APOBEC3F、APOBEC3G、APOBEC3H、又は APOBEC4 を含む、請求項 81 に記載の操作されたガイドリボ核酸ポリヌクレオチド。

10

【請求項 83】

請求項 74 ～ 82 のいずれか一項に記載の操作されたガイドリボ核酸ポリヌクレオチドをコードする、デオキシリボ核酸ポリヌクレオチド。

【請求項 84】

生物における発現に最適化された操作された核酸配列を含む核酸であって、前記核酸が、RuvCドメイン及びHNHドメインを含むエンドヌクレアーゼをコードし、前記エンドヌクレアーゼが、未培養微生物に由来し、前記エンドヌクレアーゼが、約 120 kDa

20

以下、100 kDa 以下、90 kDa 以下、60 kDa 以下、又は 30 kDa 以下の分子量を有し、前記エンドヌクレアーゼが、配列番号 674 ～ 675、975 ～ 1002、1260 ～ 1321、又はそれらと少なくとも 70 % の配列同一性を有するそれらのバリエーションを含む、核酸。

【請求項 85】

前記エンドヌクレアーゼが、前記エンドヌクレアーゼの N 末端又は C 末端の近位に 1 つ以上の核局在化配列 (NLS) をコードする配列を更に含む、請求項 84 に記載の核酸。

【請求項 86】

前記 NLS が、配列番号 205 ～ 220 から選択される配列を含む、請求項 85 に記載の核酸。

30

【請求項 87】

前記生物が、原核生物、細菌、真核生物、真菌、植物、哺乳類、齧歯類、又はヒトである、請求項 84 ～ 86 のいずれか一項に記載の核酸。

【請求項 88】

前記生物が、原核生物又は細菌であり、前記生物は、前記エンドヌクレアーゼが由来する生物とは異なる生物である、請求項 87 に記載の核酸。

【請求項 89】

前記生物が、前記未培養微生物ではない、請求項 87 に記載の核酸。

【請求項 90】

RuvC-ドメイン及びHNHドメインを含むRNAガイドエンドヌクレアーゼをコードする核酸配列を含むベクターであって、前記エンドヌクレアーゼが、未培養微生物に由来し、前記エンドヌクレアーゼが、約 120 kDa 以下、100 kDa 以下、90 kDa 以下、又は 60 kDa 以下の分子量を有し、前記RNAガイドエンドヌクレアーゼが、任意選択的に、古細菌であり、

40

前記RNAガイドエンドヌクレアーゼが、配列番号 674 ～ 675、975 ～ 1002、1260 ～ 1321、又はそれらと少なくとも 70 % の配列同一性を有するそれらのバリエーションを含む、ベクター。

【請求項 91】

前記エンドヌクレアーゼが、RRxRRモチーフを含むアルギニンリッチ領域又はPF

50

1 4 2 3 9 相同性を有するドメインを更に含む、請求項 9 0 に記載のベクター。

【請求項 9 2】

前記エンドヌクレアーゼが、R E C（認識）ドメインを更に含む、請求項 9 0 又は 9 1 に記載のベクター。

【請求項 9 3】

前記エンドヌクレアーゼが、B H ドメイン、W E D ドメイン、及び標的隣接モチーフ（T A M）相互作用（T I）ドメインを更に含む、請求項 9 1 又は 9 2 に記載のベクター。

【請求項 9 4】

前記 T I ドメインが、配列番号 1 ～ 1 9 8、2 2 1 ～ 4 5 9、4 6 3 ～ 6 1 2、6 1 7 ～ 6 6 8、6 7 4 ～ 6 7 5、9 7 5 ～ 1 0 0 2、1 2 6 0 ～ 1 3 2 1 のうちのいずれか 1 つの T I ドメインを含む、請求項 9 3 に記載のベクター。

10

【請求項 9 5】

請求項 8 4 ～ 8 9 のうちのいずれか一項に記載の核酸を含む、ベクター。

【請求項 9 6】

前記エンドヌクレアーゼと複合体を形成するように構成された、操作されたガイドリボ核酸構造をコードする核酸を更に含み、前記操作されたガイドリボ核酸構造が、

a）標的デオキシリボ核酸配列にハイブリダイズするように構成されたガイドリボ核酸配列と、

b）前記エンドヌクレアーゼに結合するように構成された t r a c r リボ核酸配列と、を含む、請求項 9 0 ～ 9 5 のいずれか一項に記載のベクター。

20

【請求項 9 7】

前記ベクターが、プラスミド、ミニサークル、C E L i D、アデノ随伴ウイルス（A A V）由来ビリオン、又はレンチウイルスである、請求項 9 0 ～ 9 6 のいずれか一項に記載のベクター。

【請求項 9 8】

請求項 9 0 ～ 9 7 のいずれか一項に記載のベクターを含む、細胞。

【請求項 9 9】

前記細胞が、細菌、古細菌、真菌、真核、哺乳類、又は植物の細胞である、請求項 9 8 に記載の細胞。

【請求項 1 0 0】

30

前記細胞が、細菌細胞である、請求項 9 9 に記載の細胞。

【請求項 1 0 1】

請求項 9 8 ～ 1 0 0 のいずれか一項に記載の細胞を培養することを含む、エンドヌクレアーゼを製造する、方法。

【請求項 1 0 2】

二本鎖デオキシリボ核酸ポリヌクレオチドを結合、切断、マーキング、又は修飾するための方法であって、

（a）前記二本鎖デオキシリボ核酸ポリヌクレオチドを、前記エンドヌクレアーゼ及び前記二本鎖デオキシリボ核酸ポリヌクレオチドに結合するように構成された操作されたガイドリボ核酸構造との複合体におけるエンドヌクレアーゼと接触させることを含み、

40

（b）前記二本鎖デオキシリボ核酸ポリヌクレオチドが、標的隣接モチーフ（T A M）を含み、

前記エンドヌクレアーゼが、約 1 2 0 k D a 以下、1 0 0 k D a 以下、9 0 k D a 以下、又は 6 0 k D a 以下の分子量を有する、方法。

【請求項 1 0 3】

前記エンドヌクレアーゼが、前記二本鎖デオキシリボ核酸ポリヌクレオチドを切断し、前記 T A M が、配列番号 1 0 2 3 ～ 1 0 4 4 のうちのいずれか 1 つを含む、請求項 1 0 2 に記載の方法。

【請求項 1 0 4】

前記エンドヌクレアーゼが、前記 T A M から前記二本鎖デオキシリボ核酸ポリヌクレオ

50

チドを、5～7個のヌクレオチド、5個のヌクレオチド、6個のヌクレオチド、又は7個のヌクレオチド切断する、請求項102又は103に記載の方法。

【請求項105】

前記エンドヌクレアーゼが、配列番号1～198、221～459、463～612、617～668、674～675、975～1002、1260～1321のうちのいずれか1つと少なくとも70%、少なくとも75%、少なくとも80%、又は少なくとも90%の配列同一性を有するバリエーションを含む、請求項102に記載の方法。

【請求項106】

二本鎖デオキシリボ核酸ポリヌクレオチドを結合、切断、マーキング、又は修飾する方法であって、

(a) 前記二本鎖デオキシリボ核酸ポリヌクレオチドを、前記エンドヌクレアーゼ及び前記二本鎖デオキシリボ核酸ポリヌクレオチドに結合するように構成された操作されたガイドリボ核酸構造との複合体におけるRNAガイド古細菌エンドヌクレアーゼと接触させることを含み、

(b) 前記二本鎖デオキシリボ核酸ポリヌクレオチドが、プロトスペーサー隣接モチーフ(PAM)を含み、

前記エンドヌクレアーゼが、配列番号674～675、975～1002、1260～1321のうちのいずれか1つと少なくとも70%、少なくとも75%、少なくとも80%、又は少なくとも90%の配列同一性を有するバリエーションを含む、方法。

【請求項107】

前記エンドヌクレアーゼが、前記二本鎖デオキシリボ核酸ポリヌクレオチドを切断し、前記PAMが、NGGを含む、請求項106に記載の方法。

【請求項108】

前記エンドヌクレアーゼが、前記二本鎖デオキシリボ核酸ポリヌクレオチドを、前記PAMから6～9又は7個のヌクレオチド切断する、請求項106又は107に記載の方法。

【請求項109】

前記エンドヌクレアーゼが、Cas9エンドヌクレアーゼ、Cas14エンドヌクレアーゼ、Cas12aエンドヌクレアーゼ、Cas12bエンドヌクレアーゼ、Cas12cエンドヌクレアーゼ、Cas12dエンドヌクレアーゼ、Cas12eエンドヌクレアーゼ、Cas13aエンドヌクレアーゼ、Cas13bエンドヌクレアーゼ、Cas13cエンドヌクレアーゼ、又はCas13dエンドヌクレアーゼではない、請求項102～108のいずれか一項に記載の方法。

【請求項110】

前記エンドヌクレアーゼが、未培養微生物に由来する、請求項102～109のいずれか一項に記載の方法。

【請求項111】

前記二本鎖デオキシリボ核酸ポリヌクレオチドが、原核細胞、古細菌、細菌、真核生物、植物、真菌、哺乳類、齧歯類、又はヒト二本鎖デオキシリボ核酸ポリヌクレオチドである、請求項102～110のいずれか一項に記載の方法。

【請求項112】

前記二本鎖デオキシリボ核酸ポリヌクレオチドは、前記エンドヌクレアーゼが由来した種以外の種由来の原核生物、古細菌、又は細菌二本鎖デオキシリボ核酸ポリヌクレオチドである、請求項111に記載の方法。

【請求項113】

標的核酸遺伝子座を修飾する方法であって、前記方法が、請求項244～73のいずれか一項に記載の前記操作されたヌクレアーゼ系を前記標的核酸遺伝子座に送達することを含み、前記エンドヌクレアーゼが、前記操作されたガイドリボ核酸構造と複合体を形成するように構成されており、前記複合体が、前記標的核酸遺伝子座への前記複合体の結合時に、前記複合体が前記標的核酸遺伝子座を修飾するように構成されている、方法。

10

20

30

40

50

【請求項 1 1 4】

前記標的核酸遺伝子座を修飾することが、前記標的核酸遺伝子座を結合、ニックング、切断、又はマーキングすることを含む、請求項 1 1 3 に記載の方法。

【請求項 1 1 5】

前記標的核酸遺伝子座が、デオキシリボ核酸（DNA）又はリボ核酸（RNA）を含む、請求項 1 1 3 又は 1 1 4 に記載の方法。

【請求項 1 1 6】

前記標的核酸が、ゲノム真核生物 DNA、古細菌 DNA、ウイルス DNA、又は細菌 DNA を含む、請求項 1 1 5 に記載の方法。

【請求項 1 1 7】

前記標的核酸が、細菌 DNA を含み、前記細菌 DNA が、前記エンドヌクレアーゼが由来した種とは異なる細菌又は古細菌種に由来する、請求項 1 1 6 に記載の方法。

【請求項 1 1 8】

前記標的核酸遺伝子座が、インビトロである、請求項 1 1 3 ～ 1 1 7 のいずれか一項に記載の方法。

【請求項 1 1 9】

前記標的核酸遺伝子座が、細胞内にある、請求項 1 1 3 ～ 1 1 7 のいずれか一項に記載の方法。

【請求項 1 2 0】

前記エンドヌクレアーゼ及び前記操作されたガイド核酸構造が、別個の核酸分子によってコードされる、請求項 1 1 9 に記載の方法。

【請求項 1 2 1】

前記細胞が、原核細胞、細菌細胞、古細菌細胞、真核細胞、真菌細胞、植物細胞、動物細胞、哺乳類細胞、齧歯類細胞、霊長類細胞、又はヒト細胞である、請求項 1 1 9 又は 1 2 0 に記載の方法。

【請求項 1 2 2】

前記細胞が、前記エンドヌクレアーゼが由来した種とは異なる種に由来する、請求項 1 1 9 又は 1 2 0 に記載の方法。

【請求項 1 2 3】

前記操作されたヌクレアーゼ系を前記標的核酸遺伝子座に送達することが、請求項 8 4 ～ 8 9 のいずれか一項に記載の核酸又は請求項 9 0 ～ 9 7 のいずれか一項に記載のベクターを送達することを含む、請求項 1 1 9 ～ 1 2 2 のいずれか一項に記載の方法。

【請求項 1 2 4】

前記操作されたヌクレアーゼ系を前記標的核酸遺伝子座に送達することが、前記エンドヌクレアーゼをコードするオープンリーディングフレームを含む核酸を送達することを含む、請求項 1 1 3 ～ 1 2 3 のいずれか一項に記載の方法。

【請求項 1 2 5】

前記核酸が、前記エンドヌクレアーゼをコードする前記オープンリーディングフレームが作動可能に連結されているプロモーターを含む、請求項 1 2 4 に記載の方法。

【請求項 1 2 6】

前記操作されたヌクレアーゼ系を前記標的核酸遺伝子座に送達することが、前記エンドヌクレアーゼをコードする前記オープンリーディングフレームを含有するキャッピングされた mRNA を送達することを含む、請求項 1 2 2 ～ 1 2 5 のいずれか一項に記載の方法。

【請求項 1 2 7】

前記操作されたヌクレアーゼ系を前記標的核酸遺伝子座に送達することが、翻訳されたポリペプチドを送達することを含む、請求項 1 1 3 ～ 1 2 0 のいずれか一項に記載の方法。

【請求項 1 2 8】

前記操作されたヌクレアーゼ系を前記標的核酸遺伝子座に送達することが、リボ核酸（

10

20

30

40

50

RNA) pol III プロモーターに作動可能に連結された前記操作されたガイドリボ核酸構造をコードするデオキシリボ核酸 (DNA) を送達することを含む、請求項 113 ~ 126 のいずれか一項に記載の方法。

【請求項 129】

前記エンドヌクレアーゼが、前記標的遺伝子座で、又は前記標的遺伝子座の近位で、一本鎖切断又は二本鎖切断を誘導する、請求項 113 ~ 128 のいずれか一項に記載の方法。

【請求項 130】

前記エンドヌクレアーゼが、プロトスペーサー隣接モチーフ (PAM) から 5' 側の前記標的遺伝子座に対して近位の二本鎖切断を誘導する、請求項 129 に記載の方法。

【請求項 131】

前記エンドヌクレアーゼが、前記 PAM から 5' 側の 6 ~ 8 個のヌクレオチド又は 7 個のヌクレオチドの二本鎖切断を誘導する、請求項 130 に記載の方法。

【請求項 132】

前記操作されたヌクレアーゼ系が、前記標的遺伝子座内又はその近位にヌクレオチド塩基の化学修飾を誘導する、請求項 113 ~ 128 のいずれか一項に記載の方法。

【請求項 133】

前記化学修飾が、アデノシン又はシトシンヌクレオチドの脱アミノ化である、請求項 132 に記載の方法。

【請求項 134】

前記エンドヌクレアーゼが、前記エンドヌクレアーゼに結合された塩基エディタを更を含む、請求項 113 ~ 128 のいずれか一項に記載の操作されたヌクレアーゼ系。

【請求項 135】

前記塩基エディタが、アデノシンデアミナーゼである、請求項 134 に記載の操作されたヌクレアーゼ系。

【請求項 136】

前記アデノシンデアミナーゼが、ADAR1 又は ADAR2 を含む、請求項 135 に記載の操作されたヌクレアーゼ系。

【請求項 137】

前記塩基エディタが、シトシンデアミナーゼである、請求項 134 に記載の操作されたヌクレアーゼ系。

【請求項 138】

前記シトシンデアミナーゼが、APOBEC1、APOBEC2、APOBEC3A、APOBEC3B、APOBEC3C、APOBEC3D、APOBEC3F、APOBEC3G、APOBEC3H、又は APOBEC4 を含む、請求項 137 に記載の操作されたヌクレアーゼ系。

【請求項 139】

細胞中の TRAC 遺伝子座を破壊する方法であって、前記細胞に組成物を接触させることを含み、組成物が、

(a) 配列番号 1 ~ 198、221 ~ 459、463 ~ 612、617 ~ 668、674 ~ 675、975 ~ 1002、1260 ~ 1321、又はそれらのバリエーションのうちのいずれか 1 つと少なくとも 75 % の同一性を有するエンドヌクレアーゼと、

(b) 操作されたガイド RNA であって、前記操作されたガイド RNA が、前記エンドヌクレアーゼと複合体を形成するように構成されており、前記操作されたガイド RNA が、前記遺伝子座の領域にハイブリダイズするように構成されたスペーサー配列を含む、操作されたガイド RNA と、を含み、

前記操作されたガイド RNA が、配列番号 1079 ~ 1082、1145 ~ 1166、及び 1169 ~ 1170 のうちのいずれか 1 つにハイブリダイズするように構成されている、方法。

【請求項 140】

前記操作されたガイドRNAが、配列番号1123～1144又は1167～1168のうちのいずれか1つと少なくとも約80%の同一性を有する配列を含む、請求項139に記載の方法。

【請求項141】

前記操作されたガイドRNAが、配列番号1123～1144又は1167～1168のうちのいずれか1つの修飾ヌクレオチドを含む、請求項139又は140に記載の方法。

【請求項142】

前記操作されたガイドRNAが、配列番号1145～1166又は1169～1170のうちのいずれか1つと相補的な配列と少なくとも約80%の同一性を有する配列を含む、請求項139～141のいずれか一項に記載の方法。

【請求項143】

前記エンドヌクレアーゼが、配列番号582、988、990、993、996、999、又は1002のうちのいずれか1つと少なくとも約75%、80%、又は90%の配列同一性を有する、請求項139～142のいずれか一項に記載の方法。

【請求項144】

前記領域が、配列番号 配列番号1023～1044のうちのいずれか1つを含むプロト Spacer 隣接モチーフ (PAM) に対して5'側である、請求項139～143のいずれか一項に記載の方法。

【請求項145】

配列番号1123～1144又は1167～1168のうちのいずれか1つと少なくとも約80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、又は100%同一の配列を含む、単離されたRNA分子。

【請求項146】

配列番号1123～1144又は1167～1168のうちのいずれか1つに列挙される化学修飾のパターンを更に含む、請求項145に記載の単離されたRNA分子。

【請求項147】

細胞のTRAC遺伝子座を修飾するための、請求項145又は146に記載の単離されたRNA分子の使用。

【請求項148】

細胞中のAVS1遺伝子座を破壊する方法であって、前記細胞に組成物を接触させることを含み、組成物が、

(a) 配列番号1～198、221～459、463～612、617～668、674～675、975～1002、1260～1321、又はそれらのバリエーションのうちのいずれか1つと少なくとも約75%の同一性を有するエンドヌクレアーゼと、

(b) 操作されたガイドRNAであって、前記操作されたガイドRNAが、前記エンドヌクレアーゼと複合体を形成するように構成されており、前記操作されたガイドRNAが、前記遺伝子座の領域にハイブリダイズするように構成されたSpacer配列を含む、操作されたガイドRNAと、を含み、

前記操作されたガイドRNAが、配列番号1105～1122のうちのいずれか1つにハイブリダイズするように構成されている、方法。

【請求項149】

前記操作されたガイドRNAが、配列番号1087～1104のうちのいずれか1つと少なくとも約80%の同一性を有する配列を含む、請求項148に記載の方法。

【請求項150】

前記操作されたガイドRNAが、配列番号1087～1104のうちのいずれか1つの修飾ヌクレオチドを含む、請求項148又は149に記載の方法。

【請求項151】

前記操作されたガイドRNAが、配列番号1105～1122のうちのいずれか1つと

10

20

30

40

50

相補的な配列と少なくとも約 80 % の同一性を有する配列を含む、請求項 148 ~ 141 のいずれか一項に記載の方法。

【請求項 152】

前記エンドヌクレアーゼが、配列番号 582、988、990、993、996、999、又は 1002 のうちのいずれか 1 つと少なくとも約 75 %、80 %、又は 90 % の配列同一性を有する、請求項 148 ~ 151 のいずれか一項に記載の方法。

【請求項 153】

前記エンドヌクレアーゼが、配列番号 582 と少なくとも約 75 %、80 %、又は 90 % の配列同一性を有する、請求項 148 ~ 151 のいずれか一項に記載の方法。

【請求項 154】

前記領域が、配列番号 1023 ~ 1044 のうちのいずれか 1 つを含むプロトSpacer 隣接モチーフ (PAM) に対して 5' 側である、請求項 148 ~ 152 のいずれか一項に記載の方法。

【請求項 155】

配列番号 1087 ~ 1104 のうちのいずれか 1 つと少なくとも約 80 %、81 %、82 %、83 %、84 %、85 %、86 %、87 %、88 %、89 %、90 %、91 %、92 %、93 %、94 %、95 %、96 %、97 %、98 %、99 %、又は 100 % 同一の配列を含む、単離された RNA 分子。

【請求項 156】

配列番号 1087 ~ 1104 のうちのいずれか 1 つに列挙される化学修飾のパターンを更に含む、請求項 155 に記載の単離された RNA 分子。

【請求項 157】

細胞の TRAC 遺伝子座を修飾するための、請求項 155 又は 156 に記載の単離された RNA 分子の、使用。

【請求項 158】

操作されたヌクレアーゼ系であって、

(a) RuvC ドメイン及び HNH ドメインを含むエンドヌクレアーゼであって、前記エンドヌクレアーゼが、配列番号 25 ~ 27、30、126、582、594、118、128、396、530、618、620、621、653、656、657、656、又はそれらのバリエーションのうちのいずれか 1 つと少なくとも 80 % の配列同一性を有する配列を含む、エンドヌクレアーゼと、

(b) 前記エンドヌクレアーゼと複合体を形成するように構成された、操作されたガイドリボ核酸構造であって、

(i) 標的デオキシリボ核酸配列にハイブリダイズするように構成されたガイドリボ核酸配列、及び

(ii) 前記エンドヌクレアーゼに結合するように構成されたリボ核酸配列を含む、操作されたガイドリボ核酸構造と、を含み、

前記エンドヌクレアーゼに結合するように構成された前記リボ核酸配列が、配列番号 677 ~ 681、686、1006 ~ 1008、1011 ~ 1014、又は 1231 ~ 1259 のうちのいずれか 1 つの前記非縮重ヌクレオチドと少なくとも 80 % の配列同一性を有する配列を含む、操作されたヌクレアーゼ系。

【請求項 159】

前記操作されたガイドリボ核酸構造が、前記エンドヌクレアーゼに結合するように構成された前記ガイドリボ核酸配列及び前記リボ核酸配列を含む単一のリボ核酸ポリヌクレオチドを含む、請求項 158 に記載の操作されたヌクレアーゼ系。

【請求項 160】

前記ガイドリボ核酸配列が、原核生物、細菌、古細菌、真核生物、真菌、植物、哺乳類、又はヒトゲノム配列と相補的である、請求項 158 又は 159 に記載の操作されたヌクレアーゼ系。

【請求項 161】

10

20

30

40

50

前記エンドヌクレアーゼが、前記エンドヌクレアーゼのN末端又はC末端の近位に1つ以上の核局在化配列（NLS）を含む、請求項158～160のいずれか一項に記載の操作されたヌクレアーゼ系。

【請求項162】

前記NLSが、配列番号205～220のうちのいずれか1つから選択される配列を含む、請求項161に記載の操作されたヌクレアーゼ系。

【請求項163】

5'から3'の方向に、前記標的デオキシリボ核酸配列に対して5'側の、少なくとも20個のヌクレオチドの配列を含む第1の相同アーム、少なくとも10個のヌクレオチドの合成DNA配列、及び前記標的配列に対して3'側の、少なくとも20個のヌクレオチドの配列を含む第2の相同アームを含む、一本鎖又は二本鎖のDNA修復鋳型を更に含む、請求項158～162のいずれか一項に記載の操作されたヌクレアーゼ系。

10

【請求項164】

前記エンドヌクレアーゼ及び前記エンドヌクレアーゼに結合するように構成された前記リボ核酸配列が、同じ門内の別個の種に由来する、請求項158～163のいずれか一項に記載の操作されたヌクレアーゼ系。

【請求項165】

前記エンドヌクレアーゼが、Cas9エンドヌクレアーゼ、Cas14エンドヌクレアーゼ、Cas12aエンドヌクレアーゼ、Cas12bエンドヌクレアーゼ、Cas12cエンドヌクレアーゼ、Cas12dエンドヌクレアーゼ、Cas12eエンドヌクレアーゼ、Cas13aエンドヌクレアーゼ、Cas13bエンドヌクレアーゼ、Cas13cエンドヌクレアーゼ、又はCas13dエンドヌクレアーゼではない、請求項158～164のいずれか一項に記載の操作されたヌクレアーゼ系。

20

【請求項166】

前記エンドヌクレアーゼが、付随的なssDNA切断活性を呈さない、請求項158～165のいずれか一項に記載の操作されたヌクレアーゼ系。

【請求項167】

操作されたヌクレアーゼ系であって、

(a) RuvCDメイン及びHNHドメインを含むエンドヌクレアーゼであって、本明細書に説明されるエンドヌクレアーゼエフェクター配列のうちのいずれか1つ又はそのバリエーションと少なくとも80%の配列同一性を有する配列を含む、エンドヌクレアーゼと、

30

(b) 前記エンドヌクレアーゼと複合体を形成するように構成された、操作されたガイドリボ核酸構造であって、

(i) 標的デオキシリボ核酸配列にハイブリダイズするように構成されたガイドリボ核酸配列、及び

(ii) 前記エンドヌクレアーゼに結合するように構成されたリボ核酸配列を含む、操作されたガイドリボ核酸構造と、を含み、

前記エンドヌクレアーゼが、本明細書に説明されるsgRNA配列のうちのいずれかの非縮重ヌクレオチド又はそのバリエーションと少なくとも80%の配列同一性を有する配列を含む、操作されたヌクレアーゼ系。

40

【請求項168】

本明細書に説明されるsgRNA配列のうちのいずれかの非縮重ヌクレオチドと少なくとも約80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、又は100%同一の配列を含む、単離されたRNA分子。

【請求項169】

本明細書に説明される配列のうちのいずれかを含む、核酸。

【請求項170】

本明細書に説明される核酸配列のうちのいずれかを含む、ベクター。

【発明の詳細な説明】

50

【技術分野】

【0001】

相互参照

本出願は、2021年11月24日に出願された米国仮特許出願第63/282,999号、2021年12月15日に出願された同第63/289,981号、及び2022年6月29日に出願された同第63/356,908号の利益を主張するものであり、その各々は、参照によりその全体が本明細書に組み込まれる。

【0002】

本出願は、PCT出願第PCT/US21/24945号に関し、その全体が参照により本明細書に組み込まれる。

【0003】

配列表

本出願は、XML形式で電子的に提出された配列表を含み、その全体が参照により本明細書に組み込まれる。2022年11月23日に作成された当該XMLコピーは、55921-741__601__SL.xmlと名付けられ、1,897,990バイトのサイズである。

【背景技術】

【0004】

Cas酵素は、関連するクラスター化され規則的に間隔があいた短い回文配列反復（CRISPR）ガイドリボ核酸（RNA）とともに、原核生物の免疫系の広範な（細菌の約45%、古細菌の約84%）成分であると見られ、CRISPR-RNAガイド核酸切断による感染性ウイルス及びプラスミドなどの非自己核酸からこうした微生物を保護する働きをする。CRISPR RNAエレメントをコードするデオキシリボ核酸（DNA）エレメントは、構造及び長さが比較的保存され得るが、それらのCRISPR関連（Cas）タンパク質は、高度に多様であり、非常に様々な核酸相互作用ドメインを含有する。CRISPR DNAエレメントは、早くも1987年に観察されているが、CRISPR/Cas複合体のプログラム可能なエンドヌクレアーゼ切断能力は、比較的最近でしか認識されておらず、多様なDNA操作及び遺伝子編集用途における組換えCRISPR/Cas系の使用につながっている。これらの酵素の有用性により、多種多様なバイオテクノロジー、遺伝子編集、及び治療用途のために再利用されている。その単一エフェクターアーキテクチャにより、ゲノム工学のために現在再利用されている系の大部分は、CRISPRクラス2カテゴリーに属する。

【発明の概要】

【0005】

多くのクラス2 Casエフェクターの大きなサイズ（約1200アミノ酸超）は、治療用途のための送達を困難にする。したがって、SMART（SMALL ARCHAEA L-ASSOCIATED）ヌクレアーゼ系と呼ばれる新規の推定ガイドDNAヌクレアーゼに関する方法、組成物、及び系が本明細書に説明される。これらのエンドヌクレアーゼエフェクターは、それらの小さいサイズ（約400aa～約1050aa）、RuvC及びHNH触媒ドメインの存在、並びに一緒になって新規の生化学的機構を示唆する他の予測されるタンパク質特徴によって定義される。

【0006】

いくつかの態様では、本開示は、操作されたヌクレアーゼ系であって、（a）RuvCドメイン及びHNHドメインを含むエンドヌクレアーゼであって、エンドヌクレアーゼが、未培養微生物に由来する、エンドヌクレアーゼと、（b）エンドヌクレアーゼと複合体を形成するように構成された、操作されたガイドリボ核酸構造であって、（i）標的デオキシリボ核酸配列にハイブリダイズするように構成されたガイドリボ核酸配列、及び（ii）エンドヌクレアーゼに結合するように構成されたリボ核酸配列を含む、操作されたガイドリボ核酸構造と、を含み、エンドヌクレアーゼが、約96kDa以下、約80kDa以下、約70kDa以下、又は約60kDa以下の分子量を有し、（1）エンドヌクレア

10

20

30

40

50

ーゼが、配列番号1～198、221～459、463～612、617～668、674～675、975～1002、1260～1321、若しくはそれらのバリエーションのいずれか1つからのアルギニンリッチ領域若しくはPF14239相同性を有するドメインと少なくとも70%、少なくとも75%、少なくとも80%、少なくとも81%、少なくとも82%、少なくとも83%、少なくとも84%、少なくとも85%、少なくとも86%、少なくとも87%、少なくとも88%、少なくとも89%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、若しくは100%の配列同一性を有するアルギニンリッチ領域若しくはPF14239相同性を有するドメインを含むか、(2) エンドヌクレアーゼが、配列番号1～198、221～459、463～612、617～668、674～675、975～1002、1260～1321、若しくはそれらのバリエーションのいずれか1つからのRECドメインと少なくとも70%、少なくとも75%、少なくとも80%、少なくとも81%、少なくとも82%、少なくとも83%、少なくとも84%、少なくとも85%、少なくとも86%、少なくとも87%、少なくとも88%、少なくとも89%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、若しくは100%の配列同一性を有するRECドメインを含むか、又は(3) エンドヌクレアーゼが、配列番号1～198、221～459、463～612、617～668、674～675、975～1002、1260～1321、若しくはそれらのバリエーションのいずれか1つと少なくとも70%、少なくとも75%、少なくとも80%、少なくとも81%、少なくとも82%、少なくとも83%、少なくとも84%、少なくとも85%、少なくとも86%、少なくとも87%、少なくとも88%、少なくとも89%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、若しくは100%の配列同一性を有する配列を含む、操作されたヌクレアーゼ系を提供する。いくつかの実施形態では、(1) エンドヌクレアーゼが、配列番号674～675、975～1002、若しくは1260～1321、又はそれらのバリエーションのいずれか1つからのアルギニンリッチ領域若しくはPF14239相同性を有するドメインと少なくとも70%、少なくとも75%、少なくとも80%、少なくとも81%、少なくとも82%、少なくとも83%、少なくとも84%、少なくとも85%、少なくとも86%、少なくとも87%、少なくとも88%、少なくとも89%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、又は100%の配列同一性を有するアルギニンリッチ領域又はPF14239相同性を有するドメインを含むか、(2) エンドヌクレアーゼが、配列番号674～675、975～1002、若しくは1260～1321、又はそれらのバリエーションのいずれか1つからのRECドメインと少なくとも70%、少なくとも75%、少なくとも80%、少なくとも81%、少なくとも82%、少なくとも83%、少なくとも84%、少なくとも85%、少なくとも86%、少なくとも87%、少なくとも88%、少なくとも89%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、又は100%の配列同一性を有するRECドメインを含むか、あるいは(3) エンドヌクレアーゼが、配列番号674～675、975～1002、若しくは1260～1321、又はそれらのバリエーションのいずれか1つと少なくとも70%、少なくとも75%、少なくとも80%、少なくとも81%、少なくとも82%、少なくとも83%、少なくとも84%、少なくとも85%、少なくとも86%、少なくとも87%、少なくとも88%、少なくとも89%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99

10

20

30

40

50

%、又は100%の配列同一性を有する配列を含む。いくつかの実施形態では、エンドヌクレアーゼが、古細菌エンドヌクレアーゼである。いくつかの実施形態では、エンドヌクレアーゼが、配列番号1~198、221~459、463~612、617~668、674~675、975~1002、1260~1321のうちのいずれか1つと少なくとも70%、少なくとも75%、少なくとも80%、少なくとも81%、少なくとも82%、少なくとも83%、少なくとも84%、少なくとも85%、少なくとも86%、少なくとも87%、少なくとも88%、少なくとも89%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、又は100%の配列同一性を有する配列を含む。いくつかの実施形態では、エンドヌクレアーゼが、RRxRRモチーフを含むアルギニンリッチ領域又はPF14239相同性を有するドメインを更に含む。いくつかの実施形態では、アルギニンリッチ領域又はPF14239相同性を有するドメインが、配列番号1~198、221~459、463~612、617~668、674~675、975~1002、1260~1321のうちのいずれか1つのアルギニンリッチ領域又はPF14239相同性を有するドメインと少なくとも80%、少なくとも81%、少なくとも82%、少なくとも83%、少なくとも84%、少なくとも85%、少なくとも86%、少なくとも87%、少なくとも88%、少なくとも89%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、又は100%の配列同一性を有する。いくつかの実施形態では、エンドヌクレアーゼが、REC（認識）ドメインを更に含む。いくつかの実施形態では、RECドメインが、配列番号1~198、221~459、463~612、617~668、674~675、975~1002、1260~1321のうちのいずれか1つのRECドメインと少なくとも80%、少なくとも81%、少なくとも82%、少なくとも83%、少なくとも84%、少なくとも85%、少なくとも86%、少なくとも87%、少なくとも88%、少なくとも89%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、又は100%の配列同一性を有する。いくつかの実施形態では、エンドヌクレアーゼが、BH（ブリッジヘリックス）ドメイン、WED（ウェッジ）ドメイン、又はPI（PAM相互作用）ドメイン若しくはTI（TAM相互作用）ドメインを更に含む。いくつかの実施形態では、WEDドメイン、又はPIDドメインが、配列番号1~198、221~459、463~612、617~668、674~675、975~1002、1260~1321のうちのいずれか1つのBHドメイン、WEDドメイン、又はPIDドメインと少なくとも80%、少なくとも81%、少なくとも82%、少なくとも83%、少なくとも84%、少なくとも85%、少なくとも86%、少なくとも87%、少なくとも88%、少なくとも89%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、又は100%の配列同一性を有する。

10

20

30

40

50

【0007】

いくつかの態様では、本開示は、操作されたヌクレアーゼ系であって、(a) Ruvc-Iドメイン及びHNHドメインを含むエンドヌクレアーゼと、(b) エンドヌクレアーゼと複合体を形成するように構成された、操作されたガイドリボ核酸構造であって、(i) 標的デオキシリボ核酸配列にハイブリダイズするように構成されたガイドリボ核酸配列、及び(ii) エンドヌクレアーゼに結合するように構成されたリボ核酸配列を含む、操作されたガイドリボ核酸構造と、を含み、エンドヌクレアーゼが、配列番号674~675、975~1002、1260~1321のうちのいずれか1つ、又はそのバリエーションと少なくとも70%、少なくとも75%、少なくとも80%、少なくとも81%、少なくとも82%、少なくとも83%、少なくとも84%、少なくとも85%、少なくとも86%、少なくとも87%、少なくとも88%、少なくとも89%、少なくとも90%、少な

くとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、又は100%の配列同一性を有する配列を含む、操作されたヌクレアーゼ系を提供する。いくつかの実施形態では、エンドヌクレアーゼが、古細菌エンドヌクレアーゼである。いくつかの実施形態では、エンドヌクレアーゼが、RRxRRモチーフを含むアルギニンリッチ領域又はPF14239相同性を有するドメインを更に含む。いくつかの実施形態では、アルギニンリッチ領域又はPF14239相同性を有するドメインが、配列番号674~675、975~1002、1260~1321のうちのいずれか1つのアルギニンリッチ領域と少なくとも80%、少なくとも81%、少なくとも82%、少なくとも83%、少なくとも84%、少なくとも85%、少なくとも86%、少なくとも87%、少なくとも88%、少なくとも89%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、又は100%の配列同一性を有する。いくつかの実施形態では、エンドヌクレアーゼが、REC（認識）ドメインを更に含む。いくつかの実施形態では、RECドメインが、配列番号674~675、975~1002、1260~1321のうちのいずれか1つのRECドメインと少なくとも80%、少なくとも81%、少なくとも82%、少なくとも83%、少なくとも84%、少なくとも85%、少なくとも86%、少なくとも87%、少なくとも88%、少なくとも89%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、又は100%の配列同一性を有する。いくつかの実施形態では、エンドヌクレアーゼが、BHドメイン、WEDドメイン、及びPIDドメインを更に含む。いくつかの実施形態では、BHドメイン、WEDドメイン、又はPIDドメインが、配列番号674~675、975~1002、1260~1321のうちのいずれか1つのBHドメイン、WEDドメイン、又はPIDドメインと少なくとも80%、少なくとも81%、少なくとも82%、少なくとも83%、少なくとも84%、少なくとも85%、少なくとも86%、少なくとも87%、少なくとも88%、少なくとも89%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、又は100%の配列同一性を有する。いくつかの実施形態では、エンドヌクレアーゼが、未培養微生物に由来する。いくつかの実施形態では、エンドヌクレアーゼに結合するように構成されたりボ核酸配列が、配列番号199~200、460~461、又は669~673のうちのいずれか1つと少なくとも80%、少なくとも81%、少なくとも82%、少なくとも83%、少なくとも84%、少なくとも85%、少なくとも86%、少なくとも87%、少なくとも88%、少なくとも89%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、若しくは100%の配列同一性を有する配列を含むか、又は配列番号201~203、613~616、677~686、1003~1022、若しくは1231~1259のうちのいずれか1つの非縮重ヌクレオチドと少なくとも80%の配列同一性を有する配列を含む。いくつかの実施形態では、ガイド核酸構造が、配列番号201~203、613~616、677~686、1003~1022、又は1231~1259のうちのいずれか1つの非縮重ヌクレオチドと少なくとも80%、少なくとも81%、少なくとも82%、少なくとも83%、少なくとも84%、少なくとも85%、少なくとも86%、少なくとも87%、少なくとも88%、少なくとも89%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、又は100%の配列同一性を有する配列を含む。

【0008】

いくつかの態様では、本開示は、操作されたヌクレアーゼ系であって、（a）操作され

10

20

30

40

50

たガイドリボ核酸構造であって、(i) 標的デオキシリボ核酸配列にハイブリダイズするように構成されたガイドリボ核酸配列、及び(ii) エンドヌクレアーゼに結合するように構成されたリボ核酸配列を含み、リボ核酸配列が、配列番号199~200、460~461、若しくは669~673のうちのいずれか1つと少なくとも80%の配列同一性を有する配列、又は配列番号677~686、1006~1012、若しくは1231~1259のうちのいずれか1つの非可変ヌクレオチドと少なくとも80%、少なくとも81%、少なくとも82%、少なくとも83%、少なくとも84%、少なくとも85%、少なくとも86%、少なくとも87%、少なくとも88%、少なくとも89%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、若しくは100%の配列同一性を有する配列を含む、操作されたガイドリボ核酸構造と、(b) 操作されたガイドリボ核酸に結合するように構成されたRNAガイドエンドヌクレアーゼと、を含む、操作されたヌクレアーゼ系を提供する。いくつかの実施形態では、RNAガイドエンドヌクレアーゼが、古細菌エンドヌクレアーゼである。いくつかの実施形態では、エンドヌクレアーゼが、約120kDa以下、100kDa以下、90kDa以下、又は60kDa以下の分子量を有する、方法。いくつかの実施形態では、操作されたガイドリボ核酸構造が、少なくとも2つのリボ核酸ポリヌクレオチドを含む。いくつかの実施形態では、操作されたガイドリボ核酸構造が、エンドヌクレアーゼに結合するように構成されたガイドリボ核酸配列及びリボ核酸配列を含む単一のリボ核酸ポリヌクレオチドを含む。いくつかの実施形態では、ガイドリボ核酸配列は、原核生物、細菌、古細菌、真核生物、真菌、植物、哺乳類、又はヒトゲノム配列に相補的である。いくつかの実施形態では、ガイドリボ核酸配列が、約14~約28ヌクレオチド長、約18~約26ヌクレオチド長、約22~約26ヌクレオチド長、又は約24ヌクレオチド長である。いくつかの実施形態では、ガイドリボ核酸配列が、配列番号462、676、又は1229~1230のうちのいずれか1つと少なくとも80%、少なくとも81%、少なくとも82%、少なくとも83%、少なくとも84%、少なくとも85%、少なくとも86%、少なくとも87%、少なくとも88%、少なくとも89%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、又は100%の配列同一性を有する配列を含む。いくつかの実施形態では、エンドヌクレアーゼが、エンドヌクレアーゼのN末端又はC末端の近位に1つ以上の核局在化配列(NLS)を含む。いくつかの実施形態では、NLSが、配列番号205~220のうちのいずれか1つから選択される配列を含む。いくつかの実施形態では、系が、5'から3'の方向に、標的デオキシリボ核酸配列に対して5'側の、少なくとも20個のヌクレオチドの配列を含む第1の相同アーム、少なくとも10個のヌクレオチドの合成DNA配列、及び標的配列に対して3'側の、少なくとも20個のヌクレオチドの配列を含む第2の相同アームを含む、一本鎖又は二本鎖のDNA修復鋳型を更に含む。いくつかの実施形態では、第1又は第2の相同アームが、少なくとも40、80、120、150、200、300、500、又は1,000ヌクレオチドの配列を含む。いくつかの実施形態では、系が、 Mg^{2+} の供給源を更に含む。いくつかの実施形態では、エンドヌクレアーゼ及びエンドヌクレアーゼに結合するように構成されたリボ核酸配列が、同じ門内の別個の種に由来する。いくつかの実施形態では、エンドヌクレアーゼが、配列番号2~24のうちのいずれか1つと少なくとも70%の配列同一性を有する配列を含み、ガイドRNA構造が、ステム及びループを含むヘアピンを含むと予測されるRNA配列を含み、ステムが、少なくとも10対のリボヌクレオチド及び介在するマルチループを含む。いくつかの実施形態では、ガイドRNA構造が、第2のステム及び第2のループを更に含み、第2のステムが、少なくとも5対のリボヌクレオチドを含む。いくつかの実施形態では、ガイドRNA構造が、少なくとも2つのヘアピンを含むRNA構造を更に含む。いくつかの実施形態では、a) エンドヌクレアーゼが、配列番号1~198、221~459、463~612、617~668、674~675、975~1002、1260~1321のうちのいずれか1つ、

10

20

30

40

50

又はそれらのバリエーションと少なくとも70%、少なくとも80%、少なくとも81%、少なくとも82%、少なくとも83%、少なくとも84%、少なくとも85%、少なくとも86%、少なくとも87%、少なくとも88%、少なくとも89%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、又は100%の配列同一性を有する配列を含み、b) ガイドRNA構造が、配列番号199~200、460~461、若しくは669~673のうちのいずれか1つ、又は配列番号201~203、613~616、677~686、1006~1012、若しくは1231~1259のうちのいずれか1つの非可変ヌクレオチドと少なくとも70%、少なくとも80%、少なくとも81%、少なくとも82%、少なくとも83%、少なくとも84%、少なくとも85%、少なくとも86%、少なくとも87%、少なくとも88%、少なくとも89%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、若しくは100%の配列同一性を有する配列を含む。いくつかの実施形態では、a) エンドヌクレアーゼが、配列番号1~198、221~459、463~612、617~668、674~675、975~1002、1260~1321のうちのいずれか1つと少なくとも70%、少なくとも80%、少なくとも81%、少なくとも82%、少なくとも83%、少なくとも84%、少なくとも85%、少なくとも86%、少なくとも87%、少なくとも88%、少なくとも89%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、又は100%の配列同一性を有する配列を含み、b) ガイドRNA構造が、クラス2、II型sgRNA又はtracr配列と少なくとも70%、少なくとも80%、少なくとも81%、少なくとも82%、少なくとも83%、少なくとも84%、少なくとも85%、少なくとも86%、少なくとも87%、少なくとも88%、少なくとも89%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、又は100%の配列同一性の配列を含む。いくつかの実施形態では、配列同一性は、BLASTP、CLUSTALW、MUSCLE、MAFFT、又はSmith-Waterman相同性検索アルゴリズムのパラメータを用いたCLUSTALWによって決定される。いくつかの実施形態では、配列同一性は、3のワード長(W)、10の期待値(E)のパラメータ、及び11の存在、1の延長でギャップコストを設定しているBLOSUM62スコアリングマトリックスを使用し、かつ条件付き組成スコアマトリックス調整を使用している、BLASTP相同性検索アルゴリズムによって決定される。いくつかの実施形態では、エンドヌクレアーゼは、Cas9エンドヌクレアーゼ、Cas14エンドヌクレアーゼ、Cas12aエンドヌクレアーゼ、Cas12bエンドヌクレアーゼ、Cas12cエンドヌクレアーゼ、Cas12dエンドヌクレアーゼ、Cas12eエンドヌクレアーゼ、Cas13aエンドヌクレアーゼ、Cas13bエンドヌクレアーゼ、Cas13cエンドヌクレアーゼ、又はCas13dエンドヌクレアーゼではない。いくつかの実施形態では、エンドヌクレアーゼは、Cas9エンドヌクレアーゼと80%未満の同一性を有する。

【0009】

いくつかの態様では、本開示は、操作されたヌクレアーゼ系であって、(a) ANGG (配列番号1029)、NARAA (配列番号1030)、ATGAAA (配列番号1031)、ATGA (配列番号1032)、又はWTGG (配列番号1033)のうちのいずれか1つを含む標的隣接モチーフ(TAM)配列に対して選択的であるように構成されたエンドヌクレアーゼであって、エンドヌクレアーゼが、配列番号1~198、221~459、463~612、617~668、674~675、975~1002、1260~1321のうちのいずれか1つのTAM相互作用ドメインと少なくとも75%、少なくとも80%、少なくとも81%、少なくとも82%、少なくとも83%、少なくとも8

10

20

30

40

50

4 %、少なくとも85 %、少なくとも86 %、少なくとも87 %、少なくとも88 %、少なくとも89 %、少なくとも90 %、少なくとも91 %、少なくとも92 %、少なくとも93 %、少なくとも94 %、少なくとも95 %、少なくとも96 %、少なくとも97 %、少なくとも98 %、少なくとも99 %、又は100 %の配列同一性を有するTAM相互作用ドメインを含む、エンドヌクレアーゼと、(b)操作されたガイドRNAであって、操作されたガイドRNAがエンドヌクレアーゼと複合体を形成するように構成されており、操作されたガイドRNAが標的核酸配列にハイブリダイズするように構成されたスペーサー配列を含む、操作されたガイドRNAと、を含む、操作されたヌクレアーゼ系を提供する。いくつかの実施形態では、TAM相互作用ドメインが、配列番号674のTAM相互作用ドメイン若しくはそのバリエーションと少なくとも80 %、少なくとも81 %、少なくとも82 %、少なくとも83 %、少なくとも84 %、少なくとも85 %、少なくとも86 %、少なくとも87 %、少なくとも88 %、少なくとも89 %、少なくとも90 %、少なくとも91 %、少なくとも92 %、少なくとも93 %、少なくとも94 %、少なくとも95 %、少なくとも96 %、少なくとも97 %、少なくとも98 %、少なくとも99 %、若しくは100 %の配列同一性、又は配列番号675のTAM相互作用ドメイン若しくはそのバリエーションと少なくとも80 %、少なくとも81 %、少なくとも82 %、少なくとも83 %、少なくとも84 %、少なくとも85 %、少なくとも86 %、少なくとも87 %、少なくとも88 %、少なくとも89 %、少なくとも90 %、少なくとも91 %、少なくとも92 %、少なくとも93 %、少なくとも94 %、少なくとも95 %、少なくとも96 %、少なくとも97 %、少なくとも98 %、少なくとも99 %、若しくは100 %の配列同一性を有する配列を含む。いくつかの実施形態では、エンドヌクレアーゼ系が、真核生物、真菌、植物、哺乳類、又はヒトゲノムポリヌクレオチド配列と相補的な配列を含む。いくつかの実施形態では、ガイドRNAが、30 ~ 280ヌクレオチド長である。いくつかの実施形態では、系が、5'から3'の方向に、標的デオキシリボ核酸配列に対して5'側の、少なくとも20個のヌクレオチドの配列を含む第1の相同アーム、少なくとも10個のヌクレオチドの合成DNA配列、及び標的配列に対して3'側の、少なくとも20個のヌクレオチドの配列を含む第2の相同アームを含む、一本鎖又は二本鎖のDNA修復鋳型を更に含む。いくつかの実施形態では、第1又は第2の相同アームが、少なくとも40個のヌクレオチドの配列を含む。いくつかの実施形態では、第1及び第2の相同アームは、真核生物のゲノム配列に対して相同である。いくつかの実施形態では、一本鎖又は二本鎖DNA修復鋳型が、導入遺伝子ドナーを含む。いくつかの実施形態では、系が、1つ又は2つの一本鎖DNAセグメントに隣接した二本鎖DNAセグメントを含むDNA修復鋳型を更に含む。いくつかの実施形態では、一本鎖DNAセグメントが、二本鎖DNAセグメントの5'端にコンジュゲートされている。いくつかの実施形態では、一本鎖DNAセグメントは、二本鎖DNAセグメントの3'端にコンジュゲートされている。いくつかの実施形態では、一本鎖DNAセグメントは、4 ~ 10個のヌクレオチド塩基の長さを有する。いくつかの実施形態では、一本鎖DNAセグメントは、スペーサー配列内の配列と相補的なヌクレオチド配列を有する。いくつかの実施形態では、二本鎖DNA配列は、バーコード、オープンリーディングフレーム、エンハンサー、プロモーター、タンパク質コード配列、miRNAコード配列、RNAコード配列、又は導入遺伝子を含む。いくつかの実施形態では、二本鎖DNA配列は、ヌクレアーゼ切断部位に隣接している。

【0010】

いくつかの態様では、本開示は、操作されたヌクレアーゼ系であって、(a)NR Rを含むプロトスペーサー隣接モチーフ(PAM)配列に対して選択的であるように構成されたエンドヌクレアーゼであって、エンドヌクレアーゼが、配列番号1313 ~ 1318のうちのいずれか1つのPAM相互作用ドメインと少なくとも75 %、少なくとも80 %、少なくとも81 %、少なくとも82 %、少なくとも83 %、少なくとも84 %、少なくとも85 %、少なくとも86 %、少なくとも87 %、少なくとも88 %、少なくとも89 %、少なくとも90 %、少なくとも91 %、少なくとも92 %、少なくとも93 %、少なくとも94 %、少なくとも95 %、少なくとも96 %、少なくとも97 %、少なくとも98

10

20

30

40

50

%、少なくとも99%、又は100%の配列同一性を有するPAM相互作用ドメインを含む、エンドヌクレアーゼと、(b)操作されたガイドRNAであって、操作されたガイドRNAが標的核酸配列にハイブリダイズするように構成されており、操作されたガイドRNAが標的核酸配列にハイブリダイズするように構成されたスペーサー配列を含む、操作されたガイドRNAと、を含む、操作されたヌクレアーゼ系を提供する。いくつかの実施形態では、TAM相互作用ドメインが、配列番号674のTAM相互作用ドメイン若しくはそのバリエーションと少なくとも80%の配列同一性、又は配列番号675のTAM相互作用ドメイン若しくはそのバリエーションと少なくとも80%の配列同一性を有する配列を含む。いくつかの実施形態では、エンドヌクレアーゼ系が、真核生物、真菌、植物、哺乳類、又はヒトゲノムポリヌクレオチド配列と相補的な配列を含む。いくつかの実施形態では、ガイドRNAが、30~280ヌクレオチド長である。いくつかの実施形態では、系が、5'から3'の方向に、標的デオキシリボ核酸配列に対して5'側の、少なくとも20個のヌクレオチドの配列を含む第1の相同アーム、少なくとも10個のヌクレオチドの合成DNA配列、及び標的配列に対して3'側の、少なくとも20個のヌクレオチドの配列を含む第2の相同アームを含む、一本鎖又は二本鎖のDNA修復鋳型を更に含む。いくつかの実施形態では、第1又は第2の相同アームが、少なくとも40個のヌクレオチドの配列を含む。いくつかの実施形態では、第1及び第2の相同アームは、真核生物のゲノム配列に対して相同である。いくつかの実施形態では、一本鎖又は二本鎖DNA修復鋳型が、導入遺伝子ドナーを含む。いくつかの実施形態では、系が、1つ又は2つの一本鎖DNAセグメントに隣接した二本鎖DNAセグメントを含むDNA修復鋳型を更に含む。いくつかの実施形態では、一本鎖DNAセグメントが、二本鎖DNAセグメントの5'端にコンジュゲートされている。いくつかの実施形態では、一本鎖DNAセグメントは、二本鎖DNAセグメントの3'端にコンジュゲートされている。いくつかの実施形態では、一本鎖DNAセグメントは、4~10個のヌクレオチド塩基の長さを有する。いくつかの実施形態では、一本鎖DNAセグメントは、スペーサー配列内の配列と相補的なヌクレオチド配列を有する。いくつかの実施形態では、二本鎖DNA配列は、バーコード、オープンリーディングフレーム、エンハンサー、プロモーター、タンパク質コード配列、miRNAコード配列、RNAコード配列、又は導入遺伝子を含む。いくつかの実施形態では、二本鎖DNA配列は、ヌクレアーゼ切断部位に隣接している。

【0011】

いくつかの態様では、本開示は、操作されたガイドリボ核酸ポリヌクレオチドであって、(a)標的DNA分子中の標的配列に相補的であるヌクレオチド配列を含むDNA標的化セグメントと、(b)ハイブリダイズして二本鎖RNA(dsRNA)二重鎖を形成するヌクレオチドの2つの相補的なストレッチを含むタンパク質結合セグメントと、を含み、ヌクレオチドの2つの相補的なストレッチが、介在するヌクレオチドと互いに共有結合しており、操作されたガイドリボ核酸ポリヌクレオチドが、配列番号674~675、975~1002、1260~1321のうちのいずれか1つ又はそれらのバリエーションと少なくとも75%、少なくとも80%、少なくとも81%、少なくとも82%、少なくとも83%、少なくとも84%、少なくとも85%、少なくとも86%、少なくとも87%、少なくとも88%、少なくとも89%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、又は100%の配列同一性を有するバリエーションを含むエンドヌクレアーゼと複合体を形成するように構成されている、操作されたガイドリボ核酸ポリヌクレオチドを提供する。いくつかの実施形態では、DNA標的化セグメントは、ヌクレオチドの2つの相補的なストレッチの両方の5'側に位置付けられている。いくつかの実施形態では、a)タンパク質結合セグメントが、配列番号199~200、460~461、若しくは669~673のうちのいずれか1つと少なくとも70%、少なくとも80%、少なくとも81%、少なくとも82%、少なくとも83%、少なくとも84%、少なくとも85%、少なくとも86%、少なくとも87%、少なくとも88%、少なくとも89%、少なくとも90%、少なくとも91%、少なくと

も 92%、少なくとも 93%、少なくとも 94%、少なくとも 95%、少なくとも 96%、少なくとも 97%、少なくとも 98%、少なくとも 99%、若しくは 100% の配列同一性を有する配列を含むか、b) タンパク質結合セグメントが、配列番号 201~203、613~616、677~686、1003~1022、若しくは 1231~1259 のうちのいずれか 1 つの非可変ヌクレオチドと少なくとも 70%、少なくとも 80%、少なくとも 81%、少なくとも 82%、少なくとも 83%、少なくとも 84%、少なくとも 85%、少なくとも 86%、少なくとも 87%、少なくとも 88%、少なくとも 89%、少なくとも 90%、少なくとも 91%、少なくとも 92%、少なくとも 93%、少なくとも 94%、少なくとも 95%、少なくとも 96%、少なくとも 97%、少なくとも 98%、少なくとも 99%、若しくは 100% の配列同一性の配列を含む。いくつかの実施形態では、a) エンドヌクレアーゼが、配列番号 674~675、975~1002、1260~1321 のうちのいずれか 1 つ、又はそれらのバリエーションと少なくとも 70%、少なくとも 80%、少なくとも 81%、少なくとも 82%、少なくとも 83%、少なくとも 84%、少なくとも 85%、少なくとも 86%、少なくとも 87%、少なくとも 88%、少なくとも 89%、少なくとも 90%、少なくとも 91%、少なくとも 92%、少なくとも 93%、少なくとも 94%、少なくとも 95%、少なくとも 96%、少なくとも 97%、少なくとも 98%、少なくとも 99%、又は 100% の配列同一性を有する配列を含み、b) ガイド RNA 構造が、クラス 2、II 型 sgRNA と少なくとも 70%、少なくとも 80%、少なくとも 81%、少なくとも 82%、少なくとも 83%、少なくとも 84%、少なくとも 85%、少なくとも 86%、少なくとも 87%、少なくとも 88%、少なくとも 89%、少なくとも 90%、少なくとも 91%、少なくとも 92%、少なくとも 93%、少なくとも 94%、少なくとも 95%、少なくとも 96%、少なくとも 97%、少なくとも 98%、少なくとも 99%、又は 100% の配列同一性を有する配列を含む。いくつかの実施形態では、エンドヌクレアーゼが、エンドヌクレアーゼに結合された塩基エディタ又はヒストンエディタを更に含む。いくつかの実施形態では、塩基エディタが、アデノシンデアミナーゼである。いくつかの実施形態では、アデノシンデアミナーゼが、ADAR1 又は ADAR2 を含む。いくつかの実施形態では、塩基エディタが、シトシンデアミナーゼである。いくつかの実施形態では、シトシンデアミナーゼが、APOBEC1、APOBEC2、APOBEC3A、APOBEC3B、APOBEC3C、APOBEC3D、APOBEC3F、APOBEC3G、APOBEC3H、又は APOBEC4 を含む。

10

20

30

【0012】

いくつかの態様では、本開示は、本明細書に説明される操作されたガイドリボ核酸ポリヌクレオチドのうちのいずれかをコードするデオキシリボ核酸ポリヌクレオチドを提供する。

【0013】

いくつかの態様では、本開示は、生物における発現に最適化された操作された核酸配列を含む核酸を提供し、核酸が、RuvCドメイン及びHNHドメインを含むエンドヌクレアーゼをコードし、エンドヌクレアーゼが、未培養微生物に由来し、エンドヌクレアーゼが、約 120 kDa 以下、100 kDa 以下、90 kDa 以下、60 kDa 以下、又は 30 kDa 以下の分子量を有し、エンドヌクレアーゼが、配列番号 674~675、975~1002、1260~1321、又はそれらと少なくとも 70%、少なくとも 80%、少なくとも 81%、少なくとも 82%、少なくとも 83%、少なくとも 84%、少なくとも 85%、少なくとも 86%、少なくとも 87%、少なくとも 88%、少なくとも 89%、少なくとも 90%、少なくとも 91%、少なくとも 92%、少なくとも 93%、少なくとも 94%、少なくとも 95%、少なくとも 96%、少なくとも 97%、少なくとも 98%、少なくとも 99%、若しくは 100% の配列同一性を有するそれらのバリエーションを含む。いくつかの実施形態では、エンドヌクレアーゼは、エンドヌクレアーゼの N 末端又は C 末端の近位に 1 つ以上の核局在化配列 (NLS) をコードする配列を更に含む。いくつかの実施形態では、NLS は、配列番号 205~220 から選択される配列を含む。いく

40

50

つかの実施形態では、生物は、原核生物、細菌、真核生物、真菌、植物、哺乳類、齧歯類、又はヒトである。いくつかの実施形態では、生物が、原核生物又は細菌であり、生物は、エンドヌクレアーゼが由来する生物とは異なる生物である。いくつかの実施形態では、生物が、未培養微生物ではない。

【0014】

いくつかの態様では、本開示は、RuvC-ドメイン及びHNHドメインを含むRNAガイドエンドヌクレアーゼをコードする核酸配列を含むベクターを提供し、エンドヌクレアーゼが、未培養微生物に由来し、エンドヌクレアーゼが、約120kDa以下、100kDa以下、90kDa以下、又は60kDa以下の分子量を有し、RNAガイドエンドヌクレアーゼが、任意選択的に、古細菌であり、RNAガイドエンドヌクレアーゼが、配列番号674～675、975～1002、1260～1321、又はそれらと少なくとも70%、少なくとも80%、少なくとも81%、少なくとも82%、少なくとも83%、少なくとも84%、少なくとも85%、少なくとも86%、少なくとも87%、少なくとも88%、少なくとも89%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、若しくは100%の配列同一性を有するそれらのバリエーションを含む。いくつかの実施形態では、エンドヌクレアーゼが、RRxRRモチーフを含むアルギニンリッチ領域又はPF14239相同性を有するドメインを更に含む。いくつかの実施形態では、エンドヌクレアーゼが、REC（認識）ドメインを更に含む。いくつかの実施形態では、エンドヌクレアーゼが、BHドメイン、WEDドメイン、及び標的隣接モチーフ（TAM）相互作用（TI）ドメインを更に含む。いくつかの実施形態では、TIドメインが、配列番号1～198、221～459、463～612、617～668、674～675、975～1002、1260～1321のうちのいずれか1つのTIドメインを含む。

【0015】

いくつかの態様では、本開示は、本明細書に説明される核酸のうちのいずれかを含むベクターを提供する。いくつかの実施形態では、ベクターが、エンドヌクレアーゼと複合体を形成するように構成された操作されたガイドリボ核酸構造をコードする核酸を更に含み、操作されたガイドリボ核酸構造が、a) 標的デオキシリボ核酸配列にハイブリダイズするように構成されたガイドリボ核酸配列と、b) エンドヌクレアーゼに結合するように構成されたリボ核酸配列と、を含む。いくつかの実施形態では、ベクターは、プラスミド、ミニサークル、CELLiD、アデノ随伴ウイルス（AAV）由来ビリオン、又はレンチウイルスである。

【0016】

いくつかの態様では、本開示は、本明細書に説明されるベクターのうちのいずれかを含む細胞を提供する。いくつかの実施形態では、細胞は、細菌、古細菌、真菌、真核生物、哺乳類、又は植物細胞である。いくつかの実施形態では、細胞が、細菌細胞である。

【0017】

いくつかの態様では、本開示は、本明細書に説明される細胞のうちのいずれかを培養することを含む、エンドヌクレアーゼを製造する方法を提供する。

【0018】

いくつかの態様では、本開示は、二本鎖デオキシリボ核酸ポリヌクレオチドを結合、切断、マーキング、又は修飾するための方法であって、(a) 二本鎖デオキシリボ核酸ポリヌクレオチドを、エンドヌクレアーゼ及び二本鎖デオキシリボ核酸ポリヌクレオチドに結合するように構成された操作されたガイドリボ核酸構造との複合体におけるエンドヌクレアーゼと接触させることを含み、(b) 二本鎖デオキシリボ核酸ポリヌクレオチドが、標的隣接モチーフ（TAM）を含み、エンドヌクレアーゼが、約120kDa以下、100kDa以下、90kDa以下、又は60kDa以下の分子量を有する、方法を提供する。いくつかの実施形態では、エンドヌクレアーゼが、二本鎖デオキシリボ核酸ポリヌクレオチドを切断し、TAMが、配列番号1023～1044のうちのいずれか1つを含む。い

くつかの実施形態では、エンドヌクレアーゼが、T A Mから二本鎖デオキシリボ核酸ポリヌクレオチドを、5～7個のヌクレオチド、5個のヌクレオチド、6個のヌクレオチド、又は7個のヌクレオチド切断する。いくつかの実施形態では、エンドヌクレアーゼが、配列番号1～198、221～459、463～612、617～668、674～675、975～1002、1260～1321のうちのいずれか1つと少なくとも70%、少なくとも75%、少なくとも80%、少なくとも81%、少なくとも82%、少なくとも83%、少なくとも84%、少なくとも85%、少なくとも86%、少なくとも87%、少なくとも88%、少なくとも89%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、又は100%の配列同一性を有するバリエーションを含む。

10

【0019】

いくつかの態様では、本開示は、二本鎖デオキシリボ核酸ポリヌクレオチドを結合、切断、マーキング、又は修飾するための方法であって、(a)二本鎖デオキシリボ核酸ポリヌクレオチドを、エンドヌクレアーゼ及び二本鎖デオキシリボ核酸ポリヌクレオチドに結合するように構成された操作されたガイドリボ核酸構造との複合体におけるRNAガイド古細菌エンドヌクレアーゼと接触させることを含み、(b)二本鎖デオキシリボ核酸ポリヌクレオチドが、プロトスペーサー隣接モチーフ(PAM)を含み、エンドヌクレアーゼが、配列番号674～675、975～1002、1260～1321のうちのいずれか1つと少なくとも70%、少なくとも75%、少なくとも80%、少なくとも81%、少なくとも82%、少なくとも83%、少なくとも84%、少なくとも85%、少なくとも86%、少なくとも87%、少なくとも88%、少なくとも89%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、又は100%の配列同一性を有するバリエーションを含む、方法を提供する。いくつかの実施形態では、エンドヌクレアーゼが、二本鎖デオキシリボ核酸ポリヌクレオチドを切断し、PAMが、NGGを含む。いくつかの実施形態では、エンドヌクレアーゼが、PAMから二本鎖デオキシリボ核酸ポリヌクレオチドを、6～9個又は7個のヌクレオチド切断する。いくつかの実施形態では、エンドヌクレアーゼは、Cas9エンドヌクレアーゼ、Cas14エンドヌクレアーゼ、Cas12aエンドヌクレアーゼ、Cas12bエンドヌクレアーゼ、Cas12cエンドヌクレアーゼ、Cas12dエンドヌクレアーゼ、Cas12eエンドヌクレアーゼ、Cas13aエンドヌクレアーゼ、Cas13bエンドヌクレアーゼ、Cas13cエンドヌクレアーゼ、又はCas13dエンドヌクレアーゼではない。いくつかの実施形態では、エンドヌクレアーゼは、未培養微生物に由来する。いくつかの実施形態では、二本鎖デオキシリボ核酸ポリヌクレオチドは、原核生物、古細菌、細菌、真核生物、植物、真菌、哺乳類、齧歯類、又はヒト二本鎖デオキシリボ核酸ポリヌクレオチドである。いくつかの実施形態では、二本鎖デオキシリボ核酸ポリヌクレオチドは、エンドヌクレアーゼが由来した種以外の種由来の原核生物、古細菌、又は細菌二本鎖デオキシリボ核酸ポリヌクレオチドである。

20

30

【0020】

いくつかの態様では、本開示は、標的核酸遺伝子座を修飾する方法を提供し、方法は、本明細書に説明される操作されたヌクレアーゼ系のうちのいずれかを標的核酸遺伝子座に送達することを含み、エンドヌクレアーゼは、操作されたガイドリボ核酸構造と複合体を形成するように構成されており、複合体は、複合体が標的核酸遺伝子座に結合すると、複合体が標的核酸遺伝子座を修飾するように構成されている。いくつかの実施形態では、標的核酸遺伝子座が、標的核酸遺伝子座を結合、ニックング、切断、又はマーキングすることを含む。いくつかの実施形態では、標的核酸遺伝子座は、デオキシリボ核酸(DNA)又はリボ核酸(RNA)を含む。いくつかの実施形態では、標的核酸遺伝子座は、ゲノム真核生物DNA、古細菌DNA、ウイルスDNA、又は細菌DNAを含む。いくつかの実施形態では、標的核酸が、細菌DNAを含み、細菌DNAが、エンドヌクレアーゼが由来

40

50

した種とは異なる細菌又は古細菌種に由来する。いくつかの実施形態では、標的核酸遺伝子座は、インビトロである。いくつかの実施形態では、標的核酸遺伝子座は、細胞内にある。いくつかの実施形態では、エンドヌクレアーゼ及び操作されたガイド核酸構造が、別個の核酸分子によってコードされる。いくつかの実施形態では、細胞は、原核細胞、細菌細胞、古細菌細胞、真核細胞、真菌細胞、植物細胞、動物細胞、哺乳類細胞、齧歯類細胞、霊長類細胞、又はヒト細胞である。いくつかの実施形態では、細胞が、エンドヌクレアーゼが由来した種とは異なる種に由来する。いくつかの実施形態では、操作されたヌクレアーゼ系を標的核酸遺伝子座に送達することは、本明細書に説明される核酸のうちのいずれか又は本明細書に説明されるベクターのうちのいずれかを送達することを含む。いくつかの実施形態では、操作されたヌクレアーゼ系を標的核酸遺伝子座に送達することは、エンドヌクレアーゼをコードするオープンリーディングフレームを含む核酸を送達することを含む。いくつかの実施形態では、核酸は、エンドヌクレアーゼをコードするオープンリーディングフレームが作動可能に連結されているプロモーターを含む。いくつかの実施形態では、操作されたヌクレアーゼ系を標的核酸遺伝子座に送達することは、エンドヌクレアーゼをコードするオープンリーディングフレームを含有するキャッピングされたmRNAを送達することを含む。いくつかの実施形態では、操作されたヌクレアーゼ系を標的核酸遺伝子座に送達することは、翻訳されたポリペプチドを送達することを含む。いくつかの実施形態では、標的核酸遺伝子座に操作されたヌクレアーゼ系を送達することは、リボ核酸(RNA)ポリマーに作動可能に連結された操作されたガイドリボ核酸をコードするデオキシリボ核酸(DNA)を送達することを含む。いくつかの実施形態では、エンドヌクレアーゼは、標的遺伝子座で、又は標的遺伝子座の近位で、一本鎖切断又は二本鎖切断を誘導する。いくつかの実施形態では、エンドヌクレアーゼが、プロト Spacer 隣接モチーフ(PAM)から5'側の標的遺伝子座に対して近位の二本鎖切断を誘導する。いくつかの実施形態では、エンドヌクレアーゼが、PAMから5'側の6~8個のヌクレオチド又は7個のヌクレオチドの二本鎖切断を誘導する。いくつかの実施形態では、操作されたヌクレアーゼ系が、標的遺伝子座内又はその近位にヌクレオチド塩基の化学修飾を誘導する。いくつかの実施形態では、化学修飾が、アデノシン又はシトシンヌクレオチドの脱アミノ化である。いくつかの実施形態では、エンドヌクレアーゼが、エンドヌクレアーゼに結合された塩基エディタを更に含む。いくつかの実施形態では、塩基エディタが、アデノシンデアミナーゼである。いくつかの実施形態では、アデノシンデアミナーゼが、ADAR1又はADAR2を含む。いくつかの実施形態では、塩基エディタが、シトシンデアミナーゼである。いくつかの実施形態では、シトシンデアミナーゼが、APOBEC1、APOBEC2、APOBEC3A、APOBEC3B、APOBEC3C、APOBEC3D、APOBEC3F、APOBEC3G、APOBEC3H、又はAPOBEC4を含む。

10

20

30

40

50

【0021】

いくつかの態様では、本開示は、細胞内のTRAC遺伝子座を破壊する方法であって、細胞に組成物を接触させることを含み、組成物が、(a)配列番号1~198、221~459、463~612、617~668、674~675、975-1002、1260-1321のうちのいずれか1つ、又はそれらのバリエーションと少なくとも75%、少なくとも80%、少なくとも81%、少なくとも82%、少なくとも83%、少なくとも84%、少なくとも85%、少なくとも86%、少なくとも87%、少なくとも88%、少なくとも89%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、又は100%を有するエンドヌクレアーゼと、(b)操作されたガイドRNAと、を含み、操作されたガイドRNAが、エンドヌクレアーゼと複合体を形成するように構成され、操作されたガイドRNAが、遺伝子座の領域にハイブリダイズするように構成されたSpacer配列を含み、操作されたガイドRNAが、配列番号1079~1082、1145~1166、及び1169~1170のうちのいずれか1つにハイブリダイズするように構成されている、方法を提供する。いくつかの実

施形態では、操作されたガイドRNAが、配列番号1123～1144又は1167～1168のうちのいずれか1つと少なくともおよそ少なくとも80%、少なくとも81%、少なくとも82%、少なくとも83%、少なくとも84%、少なくとも85%、少なくとも86%、少なくとも87%、少なくとも88%、少なくとも89%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、又は100%の配列同一性を有する配列を含む。いくつかの実施形態では、操作されたガイドRNAが、配列番号1123～1144又は1167～1168のうちのいずれか1つの修飾ヌクレオチドを含む。いくつかの実施形態では、操作されたガイドRNAが、配列番号1145～1166又は1169～1170のうちのいずれか1つと相補的な配列と少なくともおよそ少なくとも80%、少なくとも81%、少なくとも82%、少なくとも83%、少なくとも84%、少なくとも85%、少なくとも86%、少なくとも87%、少なくとも88%、少なくとも89%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、又は100%の配列同一性を有する配列を含む。いくつかの実施形態では、エンドヌクレアーゼが、配列番号582、988、990、993、996、999、又は1002のうちのいずれか1つと少なくとも75%、少なくとも80%、少なくとも81%、少なくとも82%、少なくとも83%、少なくとも84%、少なくとも85%、少なくとも86%、少なくとも87%、少なくとも88%、少なくとも89%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、又は100%の配列同一性を有する。いくつかの実施形態では、領域が、配列番号1023～1044のうちのいずれか1つを含むプロト Spacer 隣接モチーフ (PAM) に対して5'側である。

10

20

30

40

50

【0022】

いくつかの態様では、本開示は、配列番号1123～1144又は1167～1168のうちのいずれか1つと少なくとも約80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、又は100%同一の配列を含む、単離されたRNA分子を提供する。いくつかの実施形態では、配列番号1123～1144又は1167～1168のいずれか1つに列挙される化学修飾のパターンを更に含む、単離されたRNA分子。

【0023】

いくつかの態様では、本開示は、細胞のTRAC遺伝子座を修飾するための、本明細書に説明される単離されたRNA分子のうちのいずれかの使用を提供する。

【0024】

いくつかの態様では、本開示は、細胞内のAAVS1遺伝子座を破壊する方法であって、細胞に組成物を接触させることを含み、組成物が、(a) 配列番号1～198、221～459、463～612、617～668、674～675、975～1002、1260～1321のうちのいずれか1つ、又はそれらのバリエーションと少なくとも75%、少なくとも80%、少なくとも81%、少なくとも82%、少なくとも83%、少なくとも84%、少なくとも85%、少なくとも86%、少なくとも87%、少なくとも88%、少なくとも89%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、又は100%を有するエンドヌクレアーゼと、(b) 操作されたガイドRNAと、を含み、操作されたガイドRNAが、エンドヌクレアーゼと複合体を形成するように構成され、操作されたガイドRNAが、遺伝子座の領域にハイブリダイズするように構成されたSpacer配列を含み、操作されたガイドRNAが、配列番号1105～1122のうちのいずれか1つにハイブリダイズするように構成されている、方法を提供する。いくつかの実施形態では、操作されたガイドRNAが、配列

番号1087～1104のうちのいずれか1つと少なくとも80%、少なくとも81%、少なくとも82%、少なくとも83%、少なくとも84%、少なくとも85%、少なくとも86%、少なくとも87%、少なくとも88%、少なくとも89%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、又は100%の配列同一性を有する配列を含む。いくつかの実施形態では、操作されたガイドRNAが、配列番号1087～1104のうちのいずれか1つの修飾ヌクレオチドを含む。いくつかの実施形態では、操作されたガイドRNAが、配列番号1105～1122のうちのいずれか1つと相補的な配列と少なくとも約80%の同一性を有する配列を含む。いくつかの実施形態では、エンドヌクレアーゼが、配列番号582、988、990、993、996、999、又は1002のうちのいずれか1つと少なくとも75%、少なくとも80%、少なくとも81%、少なくとも82%、少なくとも83%、少なくとも84%、少なくとも85%、少なくとも86%、少なくとも87%、少なくとも88%、少なくとも89%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、又は100%の配列同一性を有する。いくつかの実施形態では、エンドヌクレアーゼが、配列番号582と少なくとも約75%、80%、又は90%の配列同一性を有する。いくつかの実施形態では、領域が、配列番号1023～1044のうちのいずれか1つを含むプロトスペーサー隣接モチーフ（PAM）に対して5'側である。

10

20

【0025】

いくつかの態様では、本開示は、配列番号1087～1104のうちのいずれか1つと少なくとも約80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、又は100%同一の配列を含む、単離されたRNA分子を提供する。いくつかの実施形態では、RNA分子が、配列番号1087～1104のうちのいずれか1つに列挙される化学修飾のパターンを含む。

【0026】

いくつかの態様では、本開示は、操作されたヌクレアーゼ系であって、(a) RuvCドメイン及びHNHドメインを含むエンドヌクレアーゼであって、配列番号25～27、30、126、582、594、118、128、396、530、618、620、621、653、656、657、656のうちのいずれか1つ又はそのバリエーションと少なくとも80%、少なくとも81%、少なくとも82%、少なくとも83%、少なくとも84%、少なくとも85%、少なくとも86%、少なくとも87%、少なくとも88%、少なくとも89%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、又は100%の配列同一性を有する配列を含む、エンドヌクレアーゼと、(b) エンドヌクレアーゼと複合体を形成するように構成された操作されたガイドリボ核酸構造であって、(i) 標的デオキシリボ核酸配列にハイブリダイズするように構成されたガイドリボ核酸配列、及び(ii) エンドヌクレアーゼに結合するように構成されたリボ核酸配列を含む、操作されたガイドリボ核酸構造と、を含み、エンドヌクレアーゼに結合するように構成されたリボ核酸配列が、配列番号677～681、686、1006～1008、1011～1014、又は1231～1259のうちのいずれか1つの非縮重ヌクレオチドと少なくとも80%、少なくとも81%、少なくとも82%、少なくとも83%、少なくとも84%、少なくとも85%、少なくとも86%、少なくとも87%、少なくとも88%、少なくとも89%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、又は100%の配列同一性を有する配列を含む、操作されたヌクレアーゼ系を提供する。いくつかの実施形態では、操作されたガイドリボ核酸構造が、エンドヌクレアーゼに結合する

30

40

50

ように構成されたガイドリボ核酸配列及びリボ核酸配列を含む単一のリボ核酸ポリヌクレオチドを含む。いくつかの実施形態では、ガイドリボ核酸配列は、原核生物、細菌、古細菌、真核生物、真菌、植物、哺乳類、又はヒトゲノム配列に相補的である。いくつかの実施形態では、エンドヌクレアーゼは、エンドヌクレアーゼのN末端又はC末端の近位に1つ以上の核局在化配列(NLS)を含む。いくつかの実施形態では、NLSが、いずれか1つの配列番号205~220から選択される配列を含む。いくつかの実施形態では、系が、5'から3'の方向に、標的デオキシリボ核酸配列に対して5'側の、少なくとも20個のヌクレオチドの配列を含む第1の相同アーム、少なくとも10個のヌクレオチドの合成DNA配列、及び標的配列に対して3'側の、少なくとも20個のヌクレオチドの配列を含む第2の相同アームを含む、一本鎖又は二本鎖のDNA修復鋳型を更に含む。いくつかの実施形態では、エンドヌクレアーゼ及びエンドヌクレアーゼに結合するように構成されたリボ核酸配列が、同じ門内の別個の種に由来する。いくつかの実施形態では、エンドヌクレアーゼは、Cas9エンドヌクレアーゼ、Cas14エンドヌクレアーゼ、Cas12aエンドヌクレアーゼ、Cas12bエンドヌクレアーゼ、Cas12cエンドヌクレアーゼ、Cas12dエンドヌクレアーゼ、Cas12eエンドヌクレアーゼ、Cas13aエンドヌクレアーゼ、Cas13bエンドヌクレアーゼ、Cas13cエンドヌクレアーゼ、又はCas13dエンドヌクレアーゼではない。いくつかの実施形態では、エンドヌクレアーゼは、付随的なssDNA切断活性を呈さない。

10

20

30

40

50

【0027】
一部の態様では、本開示は、操作されたヌクレアーゼ系であって、(a) RuvCドメイン及びHNHドメインを含むエンドヌクレアーゼであって、エンドヌクレアーゼが、本明細書に説明されるエンドヌクレアーゼエフェクター配列のうちのいずれか1つ、又はそれらのバリエーションと少なくとも80%の配列同一性を有する配列を含む、エンドヌクレアーゼと、(b) エンドヌクレアーゼと複合体を形成するように構成された操作されたガイドリボ核酸構造であって、(i) 標的デオキシリボ核酸配列にハイブリダイズするように構成されたガイドリボ核酸配列、及び(ii) エンドヌクレアーゼに結合するように構成されたリボ核酸配列を含む、操作されたガイドリボ核酸構造と、を含み、エンドヌクレアーゼが、本明細書に説明されるsgRNA配列のうちのいずれかの非縮重ヌクレオチド、又はそれらのバリエーションと少なくとも80%の配列同一性を有する配列を含む、操作されたヌクレアーゼ系を提供する。

【0028】
いくつかの態様では、本開示は、本明細書に説明されるsgRNA配列のうちのいずれかの非縮重ヌクレオチドと少なくとも約80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、又は100%同一の配列を含む、単離されたRNA分子を提供する。

【0029】
いくつかの態様では、本開示は、本明細書に説明される配列のうちのいずれかを含む核酸を提供する。

【0030】
いくつかの態様では、本開示は、本明細書に説明される核酸配列のうちのいずれかを含むベクターを提供する。

【0031】
いくつかの態様では、本開示は、操作されたヌクレアーゼ系であって、(a) RuvCドメイン及びHNHドメインを含むエンドヌクレアーゼであって、当該エンドヌクレアーゼが、未培養微生物に由来する、エンドヌクレアーゼと、(b) 当該エンドヌクレアーゼと複合体を形成するように構成された操作されたガイドリボ核酸構造であって、(i) 標的デオキシリボ核酸配列にハイブリダイズするように構成されたガイドリボ核酸配列、及び(ii) 当該エンドヌクレアーゼに結合するように構成されたtracrRNA配列を含む、操作されたガイドリボ核酸構造と、を含み、当該エンドヌクレアーゼが、約96

k D a 以下の分子量を有する、操作されたヌクレアーゼ系を提供する。いくつかの実施形態では、当該エンドヌクレアーゼが、古細菌エンドヌクレアーゼである。いくつかの実施形態では、当該エンドヌクレアーゼが、クラス 2、I I 型 C a s エンドヌクレアーゼである。いくつかの実施形態では、当該エンドヌクレアーゼが、配列番号 1 ~ 1 9 8、2 2 1 ~ 4 5 9、4 6 3 ~ 6 1 2、6 1 7 ~ 6 6 8、6 7 4 ~ 6 7 5、9 7 5 ~ 1 0 0 2、1 2 6 0 ~ 1 3 2 1 のうちのいずれか 1 つ、又はそれらのバリエーションと少なくとも 7 0 %、少なくとも 7 5 %、少なくとも 8 0 %、又は少なくとも 9 0 % の配列同一性を有する配列を含む。いくつかの実施形態では、当該エンドヌクレアーゼが、R R x R R モチーフを含むアルギニンリッチ領域又は P F 1 4 2 3 9 相同性を有するドメインを更に含む。いくつかの実施形態では、当該アルギニンリッチ領域又は P F 1 4 2 3 9 相同性を有する当該ドメインが、配列番号 1 ~ 1 9 8、2 2 1 ~ 4 5 9、4 6 3 ~ 6 1 2、6 1 7 ~ 6 6 8、6 7 4 ~ 6 7 5、9 7 5 ~ 1 0 0 2、1 2 6 0 ~ 1 3 2 1 のうちのいずれか 1 つ、又はそれらのバリエーションのアルギニンリッチ領域又は P F 1 4 2 3 9 相同性を有するドメインと少なくとも 8 5 %、少なくとも 9 0 %、又は少なくとも 9 5 % の同一性を有する。いくつかの実施形態では、当該エンドヌクレアーゼが、R E C (認識) ドメインを更に含む。いくつかの実施形態では、当該 R E C ドメインが、配列番号 1 ~ 1 9 8、2 2 1 ~ 4 5 9、4 6 3 ~ 6 1 2、6 1 7 ~ 6 6 8、6 7 4 ~ 6 7 5、9 7 5 ~ 1 0 0 2、1 2 6 0 ~ 1 3 2 1 のうちのいずれか 1 つ、又はそれらのバリエーションの R E C ドメインと少なくとも 8 5 %、少なくとも 9 0 %、又は少なくとも 9 5 % の同一性を有する。いくつかの実施形態では、当該エンドヌクレアーゼが、B H (ブリッジヘリックス) ドメイン、W E D (ウェッジ) ドメイン、及び P I (P A M 相互作用) ドメインを更に含む。いくつかの実施形態では、当該 B H ドメイン、当該 W E D ドメイン、又は当該 P I ドメインが、配列番号 1 ~ 1 9 8、2 2 1 ~ 4 5 9、4 6 3 ~ 6 1 2、6 1 7 ~ 6 6 8、6 7 4 ~ 6 7 5、9 7 5 ~ 1 0 0 2、1 2 6 0 ~ 1 3 2 1 のうちのいずれか 1 つ、又はそれらのバリエーションの B H ドメイン、W E D ドメイン、又は P I ドメインと少なくとも 8 5 %、少なくとも 9 0 %、又は少なくとも 9 5 % の同一性を有する。

10

20

30

40

50

【0032】

いくつかの態様では、本開示は、操作されたヌクレアーゼ系であって、(a) R u v C -I ドメイン及び H N H ドメインを含むエンドヌクレアーゼと、(b) 当該エンドヌクレアーゼと複合体を形成するように構成された、操作されたガイドリボ核酸構造であって、(i) 標的デオキシリボ核酸配列にハイブリダイズするように構成されたガイドリボ核酸配列、及び (i i) 当該エンドヌクレアーゼに結合するように構成されたリボ核酸配列を含む、操作されたガイドリボ核酸構造と、を含み、当該エンドヌクレアーゼが、配列番号 1 ~ 1 9 8、2 2 1 ~ 4 5 9、4 6 3 ~ 6 1 2、6 1 7 ~ 6 6 8、6 7 4 ~ 6 7 5、9 7 5 ~ 1 0 0 2、1 2 6 0 ~ 1 3 2 1 のうちのいずれか 1 つ、又はそのバリエーションと少なくとも 7 0 %、少なくとも 7 5 %、少なくとも 8 0 %、又は少なくとも 9 0 % の配列同一性を有する配列を含む、操作されたヌクレアーゼ系を提供する。いくつかの実施形態では、当該エンドヌクレアーゼが、古細菌エンドヌクレアーゼである。いくつかの実施形態では、当該エンドヌクレアーゼが、クラス 2、I I 型 C a s エンドヌクレアーゼである。いくつかの実施形態では、当該エンドヌクレアーゼが、R R x R R モチーフを含むアルギニンリッチ領域又は P F 1 4 2 3 9 相同性を有するドメインを更に含む。いくつかの実施形態では、当該アルギニンリッチ領域又は P F 1 4 2 3 9 相同性を有する当該ドメインが、配列番号 1 ~ 1 9 8、2 2 1 ~ 4 5 9、4 6 3 ~ 6 1 2、6 1 7 ~ 6 6 8、6 7 4 ~ 6 7 5、9 7 5 ~ 1 0 0 2、1 2 6 0 ~ 1 3 2 1 のうちのいずれか 1 つ、又はそれらのバリエーションのアルギニンリッチ領域と少なくとも 8 5 %、少なくとも 9 0 %、又は少なくとも 9 5 % の同一性を有する。いくつかの実施形態では、当該エンドヌクレアーゼが、R E C (認識) ドメインを更に含む。いくつかの実施形態では、当該 R E C ドメインが、配列番号 1 ~ 1 9 8、2 2 1 ~ 4 5 9、4 6 3 ~ 6 1 2、6 1 7 ~ 6 6 8、6 7 4 ~ 6 7 5、9 7 5 ~ 1 0 0 2、1 2 6 0 ~ 1 3 2 1 のうちのいずれか 1 つの R E C ドメイン、又はそれらのバリエーションの R E C ドメインと少なくとも 8 5 %、少なくとも 9 0 %、又は少なくとも 9

5 %の同一性を有する。いくつかの実施形態では、当該エンドヌクレアーゼが、B Hドメイン、W E Dドメイン、及びP Iドメインを更に含む。いくつかの実施形態では、当該B Hドメイン、当該W E Dドメイン、又は当該P Iドメインが、配列番号1 ~ 198、221 ~ 459、463 ~ 612、617 ~ 668、674 ~ 675、975 ~ 1002、1260 ~ 1321のうちのいずれか1つ、又はそれらのバリエーションのB Hドメイン、W E Dドメイン、又はP Iドメインと少なくとも85 %、少なくとも90 %、又は少なくとも95 %の同一性を有する。いくつかの実施形態では、当該エンドヌクレアーゼは、未培養微生物に由来する。いくつかの実施形態では、当該エンドヌクレアーゼに結合するように構成された当該リボ核酸配列が、配列番号199 ~ 200、460 ~ 461、若しくは669 ~ 673のうちのいずれか1つと少なくとも80 %の配列同一性を有する配列、又は配列番号201 ~ 203、613 ~ 616、677 ~ 686、1003 ~ 1022、若しくは1231 ~ 1259のうちのいずれか1つの非縮重ヌクレオチドと少なくとも80 %の配列同一性を有する配列を含む。いくつかの実施形態では、当該ガイド核酸構造が、配列番号201 ~ 203、613 ~ 616、677 ~ 686、1003 ~ 1022、又は1231 ~ 1259のうちのいずれか1つの非縮重ヌクレオチドと少なくとも80 %の同一性を有する配列を含む。

10

20

30

40

50

【0033】

いくつかの態様では、本開示は、操作されたヌクレアーゼ系であって、(a) 操作されたガイドリボ核酸構造であって、(i) 標的デオキシリボ核酸配列にハイブリダイズするように構成されたガイドリボ核酸配列、及び(ii) エンドヌクレアーゼに結合するように構成されたリボ核酸配列を含み、当該リボ核酸配列が、配列番号199 ~ 200、460 ~ 461、若しくは669 ~ 673のうちのいずれか1つと少なくとも80 %の配列同一性を有する配列、又は配列番号201 ~ 203、613 ~ 616、677 ~ 686、1003 ~ 1022、若しくは1231 ~ 1259のうちのいずれか1つの非可変ヌクレオチドと少なくとも80 %の配列同一性を有する配列を含む、操作されたガイドリボ核酸構造と、(b) 当該操作されたガイドリボ核酸に結合するように構成されたRNAガイドエンドヌクレアーゼと、を含む、操作されたヌクレアーゼ系を提供する。いくつかの実施形態では、当該RNAガイドエンドヌクレアーゼが、古細菌エンドヌクレアーゼである。いくつかの実施形態では、当該エンドヌクレアーゼが、約120 kDa以下、100 kDa以下、90 kDa以下、又は60 kDa以下の分子量を有する、方法。いくつかの実施形態では、当該操作されたガイドリボ核酸構造は、少なくとも2つのリボ核酸ポリヌクレオチドを含む。いくつかの実施形態では、当該操作されたガイドリボ核酸構造は、当該ガイドリボ核酸配列及び当該tracrRNA配列を含む単一のリボ核酸ポリヌクレオチドを含む。いくつかの実施形態では、当該ガイドリボ核酸配列は、原核生物、細菌、古細菌、真核生物、真菌、植物、哺乳類、又はヒトゲノム配列に相補的である。いくつかの実施形態では、当該ガイドリボ核酸配列は、15 ~ 24ヌクレオチド長である。いくつかの実施形態では、当該エンドヌクレアーゼは、当該エンドヌクレアーゼのN末端又はC末端の近位に1つ以上の核局在化配列(NLS)を含む。いくつかの実施形態では、当該NLSは、配列番号205 ~ 220から選択される配列を含む。いくつかの実施形態では、系が、5'から3'の方向に、当該標的デオキシリボ核酸配列に対して5'側の、少なくとも20個のヌクレオチドの配列を含む第1の相同アーム、少なくとも10個のヌクレオチドの合成DNA配列、及び当該標的配列に対して3'側の、少なくとも20個のヌクレオチドの配列を含む第2の相同アームを含む、一本鎖又は二本鎖のDNA修復鋳型を更に含む。いくつかの実施形態では、当該第1又は第2の相同アームが、少なくとも40、80、120、150、200、300、500、又は1,000ヌクレオチドの配列を含む。いくつかの実施形態では、当該系が、Mg²⁺の供給源を更に含む。いくつかの実施形態では、当該エンドヌクレアーゼ及び当該tracrRNA配列は、同じ門内の別個の細菌種に由来する。いくつかの実施形態では、当該エンドヌクレアーゼが、配列番号2 ~ 24のうちのいずれか1つと少なくとも70 %の配列同一性を有する配列を含み、当該ガイドRNA構造が、ステム及びループを含むヘアピンを含むと予測されるRNA配列を含み、

当該ステムが、少なくとも12対のリボヌクレオチドを含む。いくつかの実施形態では、当該ガイドRNA構造が、第2のステム及び第2のループを更に含み、第2のステムが、少なくとも5対のリボヌクレオチドを含む。いくつかの実施形態では、当該ガイドRNA構造が、少なくとも2つのヘアピンを含むRNA構造を更に含み。いくつかの実施形態では、当該エンドヌクレアーゼが、配列番号1と少なくとも70%の配列同一性を有する配列を含み、当該ガイドRNA構造が、ステム及びループを含む少なくとも4つのヘアピンを含むと予測されるRNA配列を含む。いくつかの実施形態では、a) 当該エンドヌクレアーゼは、配列番号1、2、10、17、若しくは613~616のうちのいずれか1つと少なくとも70%、少なくとも80%、又は少なくとも90%同一の配列を含み、b) 当該ガイドRNA構造は、配列番号199~200若しくは669~673のうちのいずれか1つ又は配列番号201~203、613~616のうちのいずれか1つの非縮重ヌクレオチドと少なくとも70%、少なくとも80%、又は少なくとも90%同一の配列を含む。いくつかの実施形態では、a) 当該エンドヌクレアーゼは、配列番号1~24、462~488、若しくは501~612のうちのいずれか1つと少なくとも70%、少なくとも80%、又は少なくとも90%同一の配列を含み、b) 当該ガイドRNA構造は、配列番号199~200若しくは669~673のうちのいずれか1つ又は配列番号201~203、613~616のうちのいずれか1つの非縮重ヌクレオチドと少なくとも70%、少なくとも80%、又は少なくとも90%同一の配列を含む。いくつかの実施形態では、a) 当該エンドヌクレアーゼは、配列番号2、10、又は17のうちのいずれか1つと少なくとも70%、少なくとも80%、又は少なくとも90%同一の配列を含み、b) 当該ガイドRNA構造は、配列番号202~203又は613~614のうちのいずれか1つの非縮重ヌクレオチドと少なくとも70%、少なくとも80%、又は少なくとも90%同一の配列を含む。いくつかの実施形態では、a) 当該エンドヌクレアーゼは、配列番号25~198、221~459、又は489~580のうちのいずれか1つと少なくとも70%、少なくとも80%、又は少なくとも90%同一の配列を含み、b) 当該ガイドRNA構造は、クラス2、II型sgRNA又はtracr配列と少なくとも70%、少なくとも80%、又は少なくとも90%同一の配列を含む。いくつかの実施形態では、当該配列同一性は、BLASTP、CLUSTALW、MUSCLE、MAFFT、又はSmith-Waterman相同性検索アルゴリズムのパラメータを用いたCLUSTALWによって決定される。いくつかの実施形態では、当該配列同一性は、3のワード長(W)、10の期待値(E)のパラメータ、及び11の存在、1の延長でギャップコストを設定しているBLOSUM62スコアリングマトリックスを使用し、かつ条件付き組成スコアマトリックス調整を使用している、当該BLASTP相同性検索アルゴリズムによって決定される。いくつかの実施形態では、当該エンドヌクレアーゼは、Cas9エンドヌクレアーゼ、Cas14エンドヌクレアーゼ、Cas12aエンドヌクレアーゼ、Cas12bエンドヌクレアーゼ、Cas12cエンドヌクレアーゼ、Cas12dエンドヌクレアーゼ、Cas12eエンドヌクレアーゼ、Cas13aエンドヌクレアーゼ、Cas13bエンドヌクレアーゼ、Cas13cエンドヌクレアーゼ、又はCas13dエンドヌクレアーゼではない。いくつかの実施形態では、当該エンドヌクレアーゼは、Cas9エンドヌクレアーゼと80%未満の同一性を有する。

10

20

30

40

【0034】

いくつかの態様では、本開示は、操作されたガイドリボ核酸ポリヌクレオチドであって、(a) 標的DNA分子中の標的配列に相補的であるヌクレオチド配列を含むDNA標的化セグメントと、(b) ハイブリダイズして二本鎖RNA(dsRNA)二重鎖を形成するヌクレオチドの2つの相補的なストレッチを含むタンパク質結合セグメントと、を含み、当該ヌクレオチドの2つの相補的なストレッチが、介在するヌクレオチドと互いに共有結合しており、当該操作されたガイドリボ核酸ポリヌクレオチドが、配列番号1~198、221~459、463~612、617~668、674~675、975~1002、1260~1321のうちのいずれか1つ又はそれらのバリエーションと少なくとも75%の配列同一性を有するバリエーションを含むエンドヌクレアーゼと複合体を形成するように

50

構成されている、操作されたガイドリボ核酸ポリヌクレオチドを提供する。いくつかの実施形態では、当該DNA標的化セグメントは、当該ヌクレオチドの2つの相補的なストレッチの両方の5'側に位置付けられている。いくつかの実施形態では、a) 当該タンパク質結合セグメントは、配列番号199~200又は669~673のうちのいずれか1つと少なくとも70%、少なくとも80%、又は少なくとも90%同一を有する配列を含み、b) 当該タンパク質結合セグメントは、配列番号201~203又は613~616のうちのいずれか1つの非縮重ヌクレオチドと少なくとも70%、少なくとも80%、又は少なくとも90%同一を有する配列を含む。いくつかの実施形態では、a) 当該エンドヌクレアーゼは、配列番号2、10、又は17のうちのいずれか1つと少なくとも70%、少なくとも80%、又は少なくとも90%同一の配列を含み、b) 当該ガイドRNA構造は、配列番号200又は配列番号202~203若しくは613~614の非縮重ヌクレオチドのうちの少なくとも1つと少なくとも70%、少なくとも80%、又は少なくとも90%同一の配列を含む。いくつかの実施形態では、a) 当該エンドヌクレアーゼは、配列番号25~198、221~459、又は489~580のうちのいずれか1つと少なくとも70%、少なくとも80%、又は少なくとも90%同一の配列を含み、b) 当該ガイドRNA構造は、クラス2、II型sgRNAと少なくとも70%、少なくとも80%、又は少なくとも90%同一の配列を含む。いくつかの実施形態では、当該エンドヌクレアーゼが、当該エンドヌクレアーゼに結合された塩基エディタ又はヒストンエディタを更に含む。いくつかの実施形態では、当該塩基エディタが、アデノシンデアミナーゼである。いくつかの実施形態では、当該アデノシンデアミナーゼが、ADAR1又はADAR2を含む。いくつかの実施形態では、当該塩基エディタが、シトシンデアミナーゼである。いくつかの実施形態では、当該シトシンデアミナーゼが、APOBEC1、APOBEC2、APOBEC3A、APOBEC3B、APOBEC3C、APOBEC3D、APOBEC3F、APOBEC3G、APOBEC3H、又はAPOBEC4を含む。

10

20

【0035】

いくつかの態様では、本開示は、本明細書に説明される操作されたガイドリボ核酸ポリヌクレオチドのうちのいずれかをコードするデオキシリボ核酸ポリヌクレオチドを提供する。

【0036】

いくつかの態様では、本開示は、生物における発現に最適化された操作された核酸配列を含む核酸であって、当該核酸が、RuvCドメイン及びHNHドメインを含むクラス2、II型Casエンドヌクレアーゼをコードし、当該エンドヌクレアーゼが、未培養微生物に由来し、当該エンドヌクレアーゼが、約120kDa以下、100kDa以下、90kDa以下、60kDa以下、又は30kDa以下の分子量を有する、核酸。いくつかの実施形態では、当該エンドヌクレアーゼが、配列番号1~198、221~459、463~612、617~668、674~675、975~1002、1260~1321、又はそれらと少なくとも70%の配列同一性を有するそれらのバリエーションを含む。いくつかの実施形態では、当該エンドヌクレアーゼは、当該エンドヌクレアーゼのN末端又はC末端の近位に1つ以上の核局在化配列(NLS)をコードする配列を更に含む。いくつかの実施形態では、当該NLSは、配列番号205~220から選択される配列を含む。いくつかの実施形態では、当該生物は、原核生物、細菌、真核生物、真菌、植物、哺乳類、齧歯類、又はヒトである。いくつかの実施形態では、当該生物が、原核生物又は細菌であり、当該生物は、当該エンドヌクレアーゼが由来する生物とは異なる生物である。いくつかの実施形態では、当該生物が、当該未培養微生物ではない。

30

40

【0037】

いくつかの態様では、本開示は、RuvC-Iドメイン及びHNHドメインを含むRNAガイドエンドヌクレアーゼをコードする核酸配列を含むベクターであって、当該エンドヌクレアーゼが、未培養微生物に由来し、当該エンドヌクレアーゼが、約120kDa以下、100kDa以下、90kDa以下、又は60kDa以下の分子量を有し、RNAガイドエンドヌクレアーゼが、任意選択的に、古細菌である、ベクターを提供する。いくつ

50

かの実施形態では、当該エンドヌクレアーゼが、RRxRRモチーフを含むアルギニンリッチ領域又はPF14239相同性を有するドメインを更に含む。いくつかの実施形態では、当該エンドヌクレアーゼが、REC（認識）ドメインを更に含む。いくつかの実施形態では、当該エンドヌクレアーゼが、BHドメイン、WEDドメイン、及びPIDドメインを更に含む。

【0038】

いくつかの態様では、本開示は、本明細書に説明される核酸のうちのいずれかを含むベクターを提供する。いくつかの実施形態では、ベクターが、当該エンドヌクレアーゼと複合体を形成するように構成された操作されたガイドリボ核酸構造をコードする核酸を更に含み、当該操作されたガイドリボ核酸構造が、a) 標的デオキシリボ核酸配列にハイブリダイズするように構成されたガイドリボ核酸配列と、b) 当該エンドヌクレアーゼに結合するように構成されたtracrリボ核酸配列と、を更に含む。いくつかの実施形態では、ベクターは、プラスミド、ミニサークル、CELiD、アデノ随伴ウイルス(AAV)由来ビリオン、又はレンチウイルスである。

【0039】

いくつかの態様では、本開示は、本明細書に説明されるベクターのうちのいずれかを含む細胞を提供する。いくつかの実施形態では、当該細胞は、細菌、古細菌、真菌、真核生物、哺乳類、又は植物細胞である。いくつかの実施形態では、当該細胞が、細菌細胞である。

【0040】

いくつかの態様では、本開示は、本明細書に説明される細胞のうちのいずれかを培養することを含む、エンドヌクレアーゼを製造する方法を提供する。

【0041】

いくつかの態様では、本開示は、二本鎖デオキシリボ核酸ポリヌクレオチドを結合、切断、マーキング、又は修飾するための方法であって、(a) 当該二本鎖デオキシリボ核酸ポリヌクレオチドを、当該エンドヌクレアーゼ及び当該二本鎖デオキシリボ核酸ポリヌクレオチドに結合するように構成された操作されたガイドリボ核酸構造との複合体におけるクラス2、II型Casエンドヌクレアーゼと接触させることを含み、(b) 当該二本鎖デオキシリボ核酸ポリヌクレオチドが、プロトスペーサー隣接モチーフ(PAM)を含み、当該エンドヌクレアーゼが、約120kDa以下、100kDa以下、90kDa以下、又は60kDa以下の分子量を有する、方法を提供する。いくつかの実施形態では、当該エンドヌクレアーゼが、当該二本鎖デオキシリボ核酸ポリヌクレオチドを切断し、当該PAMが、NGGを含む。いくつかの実施形態では、当該エンドヌクレアーゼが、当該PAMから当該二本鎖デオキシリボ核酸ポリヌクレオチドを、6~8個のヌクレオチド又は7個のヌクレオチド切断する。いくつかの実施形態では、当該エンドヌクレアーゼが、配列番号1~198、221~459、463~612、617~668、674~675、975~1002、1260~1321のうちのいずれか1つ、又はそれらのバリエーションと少なくとも70%、少なくとも75%、少なくとも80%、又は少なくとも90%の配列同一性を有するバリエーションを含む。

【0042】

いくつかの態様では、本開示は、二本鎖デオキシリボ核酸ポリヌクレオチドを結合、切断、マーキング、又は修飾するための方法であって、(a) 当該二本鎖デオキシリボ核酸ポリヌクレオチドを、当該エンドヌクレアーゼ及び当該二本鎖デオキシリボ核酸ポリヌクレオチドに結合するように構成された操作されたガイドリボ核酸構造との複合体におけるRNAガイド古細菌エンドヌクレアーゼと接触させることを含み、当該二本鎖デオキシリボ核酸ポリヌクレオチドが、プロトスペーサー隣接モチーフ(PAM)を含み、当該エンドヌクレアーゼが、配列番号1~198、221~459、463~612、617~668、674~675、975~1002、1260~1321のうちのいずれか1つ、又はそれらのバリエーションと少なくとも70%、少なくとも75%、少なくとも80%、少なくとも90%の配列同一性を有するバリエーションを含む、方法を提供する。いくつかの実

施形態では、当該エンドヌクレアーゼが、当該二本鎖デオキシリボ核酸ポリヌクレオチドを切断し、当該PAMが、NGGを含む。いくつかの実施形態では、当該エンドヌクレアーゼが、当該PAMから当該二本鎖デオキシリボ核酸ポリヌクレオチドを、6～8個又は7個のヌクレオチド切断する。いくつかの実施形態では、当該クラス2、II型Casエンドヌクレアーゼは、Cas9エンドヌクレアーゼ、Cas14エンドヌクレアーゼ、Cas12aエンドヌクレアーゼ、Cas12bエンドヌクレアーゼ、Cas12cエンドヌクレアーゼ、Cas12dエンドヌクレアーゼ、Cas12eエンドヌクレアーゼ、Cas13aエンドヌクレアーゼ、Cas13bエンドヌクレアーゼ、Cas13cエンドヌクレアーゼ、又はCas13dエンドヌクレアーゼではない。いくつかの実施形態では、当該クラス2、II型Casエンドヌクレアーゼは、未培養微生物に由来する。いくつかの実施形態では、当該二本鎖デオキシリボ核酸ポリヌクレオチドは、原核生物、古細菌、細菌、真核生物、植物、真菌、哺乳類、齧歯類、又はヒト二本鎖デオキシリボ核酸ポリヌクレオチドである。いくつかの実施形態では、当該二本鎖デオキシリボ核酸ポリヌクレオチドは、当該エンドヌクレアーゼが由来した種以外の種由来の原核生物、古細菌、又は細菌二本鎖デオキシリボ核酸ポリヌクレオチドである。

10

20

30

40

50

【0043】

いくつかの態様では、本開示は、標的核酸遺伝子座を修飾する方法を提供し、当該方法は、本明細書に説明される操作されたヌクレアーゼ系のうちのいずれかを当該標的核酸遺伝子座に送達することを含み、当該エンドヌクレアーゼは、当該操作されたガイドリボ核酸構造と複合体を形成するように構成されており、当該複合体は、当該複合体が当該標的核酸遺伝子座に結合すると、当該複合体が当該標的核酸遺伝子座を修飾するように構成されている。いくつかの実施形態では、当該標的核酸遺伝子座を修飾することが、当該標的核酸遺伝子座を結合、ニックング、切断、又はマーキングすることを含む。いくつかの実施形態では、当該標的核酸遺伝子座は、デオキシリボ核酸(DNA)又はリボ核酸(RNA)を含む。いくつかの実施形態では、当該標的核酸遺伝子座は、ゲノム真核生物DNA、古細菌DNA、ウイルスDNA、又は細菌DNAを含む。いくつかの実施形態では、当該標的核酸が、細菌DNAを含み、当該細菌DNAが、当該エンドヌクレアーゼが由来した種とは異なる細菌又は古細菌種に由来する。いくつかの実施形態では、当該標的核酸遺伝子座は、インビトロである。いくつかの実施形態では、当該標的核酸遺伝子座は、細胞内にある。いくつかの実施形態では、当該エンドヌクレアーゼ及び当該操作されたガイド核酸構造が、別個の核酸分子によってコードされる。いくつかの実施形態では、当該細胞は、原核細胞、細菌細胞、古細菌細胞、真核細胞、真菌細胞、植物細胞、動物細胞、哺乳類細胞、齧歯類細胞、霊長類細胞、又はヒト細胞である。いくつかの実施形態では、当該細胞が、当該エンドヌクレアーゼが由来した種とは異なる種に由来する。いくつかの実施形態では、当該操作されたヌクレアーゼ系を当該標的核酸遺伝子座に送達することは、本明細書に説明される核酸のうちのいずれか又は本明細書に説明されるベクターのうちのいずれかを送達することを含む。いくつかの実施形態では、当該操作されたヌクレアーゼ系を当該標的核酸遺伝子座に送達することが、当該エンドヌクレアーゼをコードするオープンリーディングフレームを含む核酸を送達することを含む。いくつかの実施形態では、当該核酸が、当該エンドヌクレアーゼをコードする当該オープンリーディングフレームが作動可能に連結されているプロモーターを含む。いくつかの実施形態では、当該操作されたヌクレアーゼ系を当該標的核酸遺伝子座に送達することは、当該エンドヌクレアーゼをコードする当該オープンリーディングフレームを含有するキャッピングされたmRNAを送達することを含む。いくつかの実施形態では、当該操作されたヌクレアーゼ系を当該標的核酸遺伝子座に送達することは、翻訳されたポリペプチドを送達することを含む。いくつかの実施形態では、当該操作されたヌクレアーゼ系を当該標的核酸遺伝子座に送達することは、リボ核酸(RNA)pol IIIプロモーターに作動可能に連結された当該操作されたガイドリボ核酸をコードするデオキシリボ核酸(DNA)を送達することを含む。いくつかの実施形態では、当該エンドヌクレアーゼは、当該標的遺伝子座で、又は当該標的遺伝子座の近位で、一本鎖切断又は二本鎖切断を誘導する。いくつかの実施形態では、

当該エンドヌクレアーゼが、プロトスペーサー隣接モチーフ（PAM）から5'側の当該標的遺伝子座に対して近位の二本鎖切断を誘導する。いくつかの実施形態では、当該エンドヌクレアーゼが、当該PAMから5'側の6～8個のヌクレオチド又は7個のヌクレオチドの二本鎖切断を誘導する。いくつかの実施形態では、当該操作されたヌクレアーゼ系が、当該標的遺伝子座内若しくはその近位にヌクレオチド塩基の化学修飾、又は標的遺伝子座内若しくはその近位にヒストンの化学修飾を誘導する。いくつかの実施形態では、当該化学修飾が、アデノシン又はシトシンヌクレオチドの脱アミノ化である。いくつかの実施形態では、当該エンドヌクレアーゼが、当該エンドヌクレアーゼに結合された塩基エディタを更に含む。いくつかの実施形態では、当該塩基エディタが、アデノシンデアミナーゼである。いくつかの実施形態では、当該アデノシンデアミナーゼが、ADAR1又はADAR2を含む。いくつかの実施形態では、当該塩基エディタが、シトシンデアミナーゼである。いくつかの実施形態では、当該シトシンデアミナーゼが、APOBEC1、APOBEC2、APOBEC3A、APOBEC3B、APOBEC3C、APOBEC3D、APOBEC3F、APOBEC3G、APOBEC3H、又はAPOBEC4を含む。

10

20

30

40

50

【0044】

本開示の更なる態様及び利点は、本開示の例示的な実施形態のみが示され、説明される、以下の詳細な説明から、当業者に容易に明らかになるであろう。理解されるように、本開示は、他の異なる実施形態が可能であり、そのいくつかの詳細は、全て本開示から逸脱することなく、様々な明白な点において修正が可能である。したがって、図面及び説明は、本質的に例示とみなされるべきであり、限定的なものとしてみなされるべきではない。

【0045】

参照による組み込み

本明細書において言及される全ての刊行物、特許、及び特許出願は、それぞれ個々の刊行物、特許、又は特許出願が、参照により組み込まれるべきことが具体的かつ個別に示されているのと同じ程度に、参照により本明細書に組み込まれる。

【図面の簡単な説明】

【0046】

本発明の新規の特徴は、添付の特許請求の範囲に特記して記載される。本発明の特徴及び利点のより良好な理解が、本発明の原理が利用される例示的な実施形態を記載する以下の詳細な説明、及び添付の図面（本明細書において「図（Figure）」及び「図（FIG.）」）を参照することによって得られるだろう。

【図1A】異なるクラス及び型のCRISPR/Cas遺伝子座の相同関係を示すデンドログラムを図示する。クラス2、II-A型、II-B型、及びII-C型Cas系と比較して本明細書に説明されるSMART I及びII Cas酵素クラスが示され、これらの系は、II-A、II-B、及びII-Cとは別個のクラスにグループ分けされることを実証する。（図1A）は、Cas9参照配列の文脈でSMART系統樹を示し、SMARTエフェクターは、Cas9参照配列（II-A型、II-B型、及びII-C型）から離れて遠くにクラスター化されており、（図1B）は、SMART酵素のサブグループを例示するSMART系統樹を示す。

【図1B】異なるクラス及び型のCRISPR/Cas遺伝子座の相同関係を示すデンドログラムを図示する。クラス2、II-A型、II-B型、及びII-C型Cas系と比較して本明細書に説明されるSMART I及びII Cas酵素クラスが示され、これらの系は、II-A、II-B、及びII-Cとは別個のクラスにグループ分けされることを実証する。（図1A）は、Cas9参照配列の文脈でSMART系統樹を示し、SMARTエフェクターは、Cas9参照配列（II-A型、II-B型、及びII-C型）から離れて遠くにクラスター化されており、（図1B）は、SMART酵素のサブグループを例示するSMART系統樹を示す。

【図2】本明細書に説明されるSMARTエフェクターの長さ分布を示し、SMART I酵素及びII酵素がCas9様酵素よりも低い分子量でクラスター化されることを示す

。SMARTヌクレアーゼは、1つのピークが約400aa (SMART II)、第2のピークが約750aa (SMART I)の二峰性分布を示す。Cas9ヌクレアーゼはまた、約1,100aa (例えば、SaCas9)及び1,300aa (例えば、SpCas9)のピークを有する二峰性分布を示す。

【図3A】「小さい」II型ヌクレアーゼMG33-1、MG35-236のゲノムコンテキストを図示する。SMARTヌクレアーゼ及びCRISPRアクセサリタンパク質は、濃灰色の矢印で示され、他の遺伝子は、淡灰色の矢印で図示される。ゲノム断片中の全ての遺伝子について予測されるドメインが、矢印の下に灰色のボックスとして示される。以下が示される。(図3A) SMART IIの下流のトランスポサゼTnpA及びTnpBを担持する予測される挿入配列を示す、SMART IIヌクレアーゼMG35-236の上流でコードされたSMART I MG33-1ヌクレアーゼ及びCRISPR遺伝子座のゲノムコンテキスト、(図3B)：環境発現配列決定リードがCRISPRアレイ及び予測されるtracrRNAの下に整列して示され、領域のトランスクリプトーム対象範囲がコンティグ配列の上に例示されている、SMART IヌクレアーゼMG34-1のゲノムコンテキスト、(図3C)環境発現配列決定リードがCRISPRアレイ及び予測されるtracrRNAの下に整列して示され、領域のトランスクリプトーム対象範囲がコンティグ配列の上に例示されている、SMART IヌクレアーゼMG34-16のゲノムコンテキスト、(図3D)ゲノム断片がウイルス特異的遺伝子アノテーションターミナーゼ及びポータルに基づいてファージに由来すると識別された、並びに(図3D)のMG34-16 CRISPRアレイからのスパーサー7によって標的とされるゲノム断片、挿入図は、未知の機能のウイルス遺伝子のC末端を標的とするMG34-16スパーサー7の場所を示し、MG34-16の推定NGG PAMは、スパーサーマッチの下流の灰色のボックスによって強調表示される。

【図3B】「小さい」II型ヌクレアーゼMG33-1、MG35-236のゲノムコンテキストを図示する。SMARTヌクレアーゼ及びCRISPRアクセサリタンパク質は、濃灰色の矢印で示され、他の遺伝子は、淡灰色の矢印で図示される。ゲノム断片中の全ての遺伝子について予測されるドメインが、矢印の下に灰色のボックスとして示される。以下が示される。(図3A) SMART IIの下流のトランスポサゼTnpA及びTnpBを担持する予測される挿入配列を示す、SMART IIヌクレアーゼMG35-236の上流でコードされたSMART I MG33-1ヌクレアーゼ及びCRISPR遺伝子座のゲノムコンテキスト、(図3B)：環境発現配列決定リードがCRISPRアレイ及び予測されるtracrRNAの下に整列して示され、領域のトランスクリプトーム対象範囲がコンティグ配列の上に例示されている、SMART IヌクレアーゼMG34-1のゲノムコンテキスト、(図3C)環境発現配列決定リードがCRISPRアレイ及び予測されるtracrRNAの下に整列して示され、領域のトランスクリプトーム対象範囲がコンティグ配列の上に例示されている、SMART IヌクレアーゼMG34-16のゲノムコンテキスト、(図3D)ゲノム断片がウイルス特異的遺伝子アノテーションターミナーゼ及びポータルに基づいてファージに由来すると識別された、並びに(図3D)のMG34-16 CRISPRアレイからのスパーサー7によって標的とされるゲノム断片、挿入図は、未知の機能のウイルス遺伝子のC末端を標的とするMG34-16スパーサー7の場所を示し、MG34-16の推定NGG PAMは、スパーサーマッチの下流の灰色のボックスによって強調表示される。

【図3C】「小さい」II型ヌクレアーゼMG33-1、MG35-236のゲノムコンテキストを図示する。SMARTヌクレアーゼ及びCRISPRアクセサリタンパク質は、濃灰色の矢印で示され、他の遺伝子は、淡灰色の矢印で図示される。ゲノム断片中の全ての遺伝子について予測されるドメインが、矢印の下に灰色のボックスとして示される。以下が示される。(図3A) SMART IIの下流のトランスポサゼTnpA及びTnpBを担持する予測される挿入配列を示す、SMART IIヌクレアーゼMG35-236の上流でコードされたSMART I MG33-1ヌクレアーゼ及びCRISPR遺伝子座のゲノムコンテキスト、(図3B)：環境発現配列決定リードがCRISPR

10

20

30

40

50

Rアレイ及び予測される *tracrRNA* の下に整列して示され、領域のトランスクリプトーム対象範囲がコンティグ配列の上に例示されている、SMART Iヌクレアーゼ MG34-1 のゲノムコンテキスト、(図3C) 環境発現配列決定リードが CRISPRアレイ及び予測される *tracrRNA* の下に整列して示され、領域のトランスクリプトーム対象範囲がコンティグ配列の上に例示されている、SMART Iヌクレアーゼ MG34-16 のゲノムコンテキスト、(図3D) ゲノム断片がウイルス特異的遺伝子アノテーションターミナーゼ及びポータルに基づいてファージに由来すると識別された、並びに(図3D) の MG34-16 CRISPRアレイからのスパーサー7によって標的とされるゲノム断片、挿入図は、未知の機能のウイルス遺伝子のC末端を標的とする MG34-16 スパーサー7の場所を示し、MG34-16 の推定 NGG PAM は、スパーサーマッチの下流の灰色のボックスによって強調表示される。

10

【図3D】「小さい」II型ヌクレアーゼ MG33-1、MG35-236 のゲノムコンテキストを図示する。SMARTヌクレアーゼ及びCRISPRアクセサリタンパク質は、濃灰色の矢印で示され、他の遺伝子は、淡灰色の矢印で図示される。ゲノム断片中の全ての遺伝子について予測されるドメインが、矢印の下に灰色のボックスとして示される。以下が示される。(図3A) SMART IIの下流のトランスポサゼ TnpA 及び TnpB を担持する予測される挿入配列を示す、SMART IIヌクレアーゼ MG35-236 の上流でコードされた SMART I MG33-1ヌクレアーゼ及びCRISPR遺伝子座のゲノムコンテキスト、(図3B)：環境発現配列決定リードが CRISPRアレイ及び予測される *tracrRNA* の下に整列して示され、領域のトランスクリプトーム対象範囲がコンティグ配列の上に例示されている、SMART Iヌクレアーゼ MG34-1 のゲノムコンテキスト、(図3C) 環境発現配列決定リードが CRISPRアレイ及び予測される *tracrRNA* の下に整列して示され、領域のトランスクリプトーム対象範囲がコンティグ配列の上に例示されている、SMART Iヌクレアーゼ MG34-16 のゲノムコンテキスト、(図3D) ゲノム断片がウイルス特異的遺伝子アノテーションターミナーゼ及びポータルに基づいてファージに由来すると識別された、並びに(図3D) の MG34-16 CRISPRアレイからのスパーサー7によって標的とされるゲノム断片、挿入図は、未知の機能のウイルス遺伝子のC末端を標的とする MG34-16 スパーサー7の場所を示し、MG34-16 の推定 NGG PAM は、スパーサーマッチの下流の灰色のボックスによって強調表示される。

20

30

【図4A】例示的なSMARTエンドヌクレアーゼ(MG33-1(配列番号1)、MG33-2(配列番号463)、MG33-3(配列番号464)、MG34-1(配列番号2)、MG34-9(配列番号10)、MG34-16(配列番号17)、MG102-1(配列番号581)、MG102-2(配列番号582)、MG35-1(配列番号25)、MG35-2(配列番号26)、MG35-3(配列番号27)、MG35-102(配列番号126)、MG35-236(配列番号284)、MG35-419(配列番号222)、MG35-420(配列番号223)、及びMG35-421(配列番号224))の多重配列アラインメントを示し、参照ドメインとして使用された SacA s9 の配列が参照配列の下で長方形として示され、触媒残基が各配列の上で正方形として示される。以下が示される。(図4A) RuvC-I 及びブリッジヘリックスドメインを含有するSMARTエンドヌクレアーゼ領域のアラインメント、(図4B) RuvC-IIドメインを含有する領域のアラインメント、並びに(図4C) RuvC-II 及びHNHドメインを含有する領域のアラインメント

40

【図4B】例示的なSMARTエンドヌクレアーゼ(MG33-1(配列番号1)、MG33-2(配列番号463)、MG33-3(配列番号464)、MG34-1(配列番号2)、MG34-9(配列番号10)、MG34-16(配列番号17)、MG102-1(配列番号581)、MG102-2(配列番号582)、MG35-1(配列番号25)、MG35-2(配列番号26)、MG35-3(配列番号27)、MG35-102(配列番号126)、MG35-236(配列番号284)、MG35-419(配列番号222)、MG35-420(配列番号223)、及びMG35-421(配列番号224))の多重配列アラインメントを示し、参照ドメインとして使用された SacA s9 の配列が参照配列の下で長方形として示され、触媒残基が各配列の上で正方形として示される。以下が示される。(図4A) RuvC-I 及びブリッジヘリックスドメインを含有するSMARTエンドヌクレアーゼ領域のアラインメント、(図4B) RuvC-IIドメインを含有する領域のアラインメント、並びに(図4C) RuvC-II 及びHNHドメインを含有する領域のアラインメント

50

号 2 2 4)) の多重配列アラインメントを示し、参照ドメインとして使用された S a C a s 9 の配列が参照配列の下で長方形として示され、触媒残基が各配列の上で正方形として示される。以下が示される。(図 4 A) R u v C - I 及びブリッジヘリックスドメインを含有するエンドヌクレアーゼ領域のアラインメント、(図 4 B) R u v C - I I I ドメインを含有する領域のアラインメント、並びに(図 4 C) R u v C - I I 及び H N H ドメインを含有する領域のアラインメント

【図 4 C】例示的な S M A R T エンドヌクレアーゼ (M G 3 3 - 1 (配列番号 1)、M G 3 3 - 2 (配列番号 4 6 3)、M G 3 3 - 3 (配列番号 4 6 4)、M G 3 4 - 1 (配列番号 2)、M G 3 4 - 9 (配列番号 1 0)、M G 3 4 - 1 6 (配列番号 1 7)、M G 1 0 2 - 1 (配列番号 5 8 1)、M G 1 0 2 - 2 (配列番号 5 8 2)、M G 3 5 - 1 (配列番号 2 5)、M G 3 5 - 2 (配列番号 2 6)、M G 3 5 - 3 (配列番号 2 7)、M G 3 5 - 1 0 2 (配列番号 1 2 6)、M G 3 5 - 2 3 6 (配列番号 2 8 4)、M G 3 5 - 4 1 9 (配列番号 2 2 2)、M G 3 5 - 4 2 0 (配列番号 2 2 3)、及び M G 3 5 - 4 2 1 (配列番号 2 2 4)) の多重配列アラインメントを示し、参照ドメインとして使用された S a C a s 9 の配列が参照配列の下で長方形として示され、触媒残基が各配列の上で正方形として示される。以下が示される。(図 4 A) R u v C - I 及びブリッジヘリックスドメインを含有するエンドヌクレアーゼ領域のアラインメント、(図 4 B) R u v C - I I I ドメインを含有する領域のアラインメント、並びに(図 4 C) R u v C - I I 及び H N H ドメインを含有する領域のアラインメント

【図 5】例として M G 3 4 - 1 を使用している、S M A R T I エンドヌクレアーゼの例示的なドメイン構成を図示する。(図 5 の A) 3 つの R u v C ドメイン、ブリッジヘリックス (「B H」)、認識ドメイン (「R E C」) を中断する P f a m P F 1 4 2 3 9 に対する相同性を有するドメイン、H N H エンドヌクレアーゼドメイン (「H N H」)、ウェッジドメイン (「W E D」)、及び P A M 相互作用ドメイン (P I) を含む、S M A R T I エンドヌクレアーゼの予測されるドメインアーキテクチャを示す図、並びに(図 5 の B) R u v C 及び H N H 触媒残基が各配列の上で黒色バーとして示され、3 D 空間内で S a C a s の結晶構造と整列する領域が丸いボックスで表され、破線が S M A R T 及び S a C a s 9 の 3 D 構造予測の間の 3 D 空間におけるアラインメントが乏しいか又はアラインメントなしの領域を表す、参照 C a s 9 エンドヌクレアーゼ配列に対する 2 つの S M A R T I エンドヌクレアーゼの多重配列アラインメントの概要が示される。

【図 6】例として M G 3 5 ファミリー酵素 (M G 3 5 - 3、M G 3 5 - 4) を使用している、S M A R T I I エンドヌクレアーゼの例示的なドメイン構成を図示する。(図 6 の A) 3 つの R u v C ドメイン、P f a m P F 1 4 2 3 9 に対する相同性を有するドメイン、H N H エンドヌクレアーゼドメイン、未知のドメイン、及び認識ドメイン (R E C) を含む、S M A R T I I エンドヌクレアーゼの予測されるドメインアーキテクチャを示す図、並びに(図 6 の B) R u v C 及び H N H 触媒残基が各配列の上に黒色バーとして示され、3 D 空間内で S a C a s の結晶構造と整列する領域が丸いボックスで表され、ガイド／標的／P A M 配列の認識に参与し得る 3 D 構造予測から識別された残基が M G 3 5 - 4 1 9 配列 (R R X R R 及び R E C ドメイン内) の上の濃灰色のボックスによって表される、参照 C a s 9 エンドヌクレアーゼ配列に対する 2 つの S M A R T I I エンドヌクレアーゼの多重配列アラインメントの概要が示される。

【図 7 A】S M A R T 酵素の様々な特徴を例示する。(図 7 A) 本明細書に図示される様々な酵素の S M A R T I ドメインと s p C a s 9 の S M A R T I ドメインとの同一性を示し、これらが最大約 3 5 % の配列同一性を有することを示すドットプロット、(図 7 B) 本明細書に説明される酵素の個々の S M A R T I ドメインの長さのドットプロットが示される。

【図 7 B】S M A R T 酵素の様々な特徴を例示する。(図 7 A) 本明細書に図示される様々な酵素の S M A R T I ドメインと s p C a s 9 の S M A R T I ドメインとの同一性を示し、これらが最大約 3 5 % の配列同一性を有することを示すドットプロット、(図 7 B) 本明細書に説明される酵素の個々の S M A R T I ドメインの長さのドットプロット

が示される。

【図 8 A】様々な S M A R T 特異的モチーフ対これらのモチーフが S M A R T 酵素でより一般的に発生する C a s 9 ヌクレアーゼ配列で予測されたモチーフの計数分布を例示し、モチーフは、803 個の参照 C a s 9 配列（I I - A 型、I I - B 型、及び I I - C 型）、84 個の S M A R T I 配列、及び 471 個の S M A R T I I 配列で予測された。（図 8 A）様々な型のクラス 2 C a s 酵素における Z n 結合リボンモチーフ（C X [2 - 4] C 及び C X [2 - 4] H）の計数頻度のボックスプロット、並びに（B）様々な型のクラス 2 C a s 酵素における R R X R R モチーフの計数頻度のヒストグラムが示される。（図 8 A）及び（図 8 B）では、線は平均計数値を辿るが、外れ値は点で表される。

【図 8 B】様々な S M A R T 特異的モチーフ対これらのモチーフが S M A R T 酵素でより一般的に発生する C a s 9 ヌクレアーゼ配列で予測されたモチーフの計数分布を例示し、モチーフは、803 個の参照 C a s 9 配列（I I - A 型、I I - B 型、及び I I - C 型）、84 個の S M A R T I 配列、及び 471 個の S M A R T I I 配列で予測された。（図 8 A）様々な型のクラス 2 C a s 酵素における Z n 結合リボンモチーフ（C X [2 - 4] C 及び C X [2 - 4] H）の計数頻度のボックスプロット、並びに（B）様々な型のクラス 2 C a s 酵素における R R X R R モチーフの計数頻度のヒストグラムが示される。（図 8 A）及び（図 8 B）では、線は平均計数値を辿るが、外れ値は点で表される。

【図 9 A】S M A R T I エンドヌクレアーゼを用いた切断活性のために設計された一本鎖 R N A（s g R N A）の予測されるガイド R N A 構造を例示する。（図 9 A）M G 3 4 - 1 s g R N A 1、（図 9 B）M G 3 4 - 1 s g R N A 2、（図 9 C）M G 3 4 - 9 s g R N A 1、及び（図 9 D）M G 3 4 - 16 s g R N A 1 が示される。

【図 9 B】S M A R T I エンドヌクレアーゼを用いた切断活性のために設計された一本鎖 R N A（s g R N A）の予測されるガイド R N A 構造を例示する。（図 9 A）M G 3 4 - 1 s g R N A 1、（図 9 B）M G 3 4 - 1 s g R N A 2、（図 9 C）M G 3 4 - 9 s g R N A 1、及び（図 9 D）M G 3 4 - 16 s g R N A 1 が示される。

【図 9 C】S M A R T I エンドヌクレアーゼを用いた切断活性のために設計された一本鎖 R N A（s g R N A）の予測されるガイド R N A 構造を例示する。（図 9 A）M G 3 4 - 1 s g R N A 1、（図 9 B）M G 3 4 - 1 s g R N A 2、（図 9 C）M G 3 4 - 9 s g R N A 1、及び（図 9 D）M G 3 4 - 16 s g R N A 1 が示される。

【図 9 D】S M A R T I エンドヌクレアーゼを用いた切断活性のために設計された一本鎖 R N A（s g R N A）の予測されるガイド R N A 構造を例示する。（図 9 A）M G 3 4 - 1 s g R N A 1、（図 9 B）M G 3 4 - 1 s g R N A 2、（図 9 C）M G 3 4 - 9 s g R N A 1、及び（図 9 D）M G 3 4 - 16 s g R N A 1 が示される。

【図 10 A】実施例 1 に説明される S M A R T I ヌクレアーゼの切断特性を図示する。（図 10 A）は、2 つの s g R N A 設計対陰性対照を用いた M G 3 4 - 1 の切断アッセイのライゲーション産物の A g i l e n t T a p e S t a t i o n ゲルを示す。レーン L 3：ラダー。レーン A 4：A p o、s g R N A なし。レーン B 4 及び C 4：試験された M G 3 4 - 1 s g R N A（s g 1：配列番号 612、s g 2：613）。切断産物バンドは矢印で標識されている。レーン G 3 及び H 3：灰色で表示され、この実験には関連しない。

【図 10 B】実施例 1 に説明される S M A R T I ヌクレアーゼの切断特性を図示する。（図 10 B）は、M G 3 4 - 1、3 4 - 9 及び 3 4 - 16 の活性を示すライゲーション産物の P C R ゲルを示す。レーン 1：ラダー。レーン 2～7：M G 3 4 - 1 の 6 つのスペーサー長を有する s g R N A 設計。レーン 8 及び 9：それぞれ、3 4～9 及び 3 4～16 の s g R N A 設計。矢印は切断確認バンドを示す。

【図 11 A】M G 3 4 ヌクレアーゼに対する配列切断選好を例示する。（図 11 A）は、s g R N A 1（上、配列番号 612）及び s g R N A 2（下、配列番号 613）を用いた M G 3 4 - 1 に対するコンセンサス P A M 配列（N G G N）の S e q L o g o 表現を示す。

【図 11 B】M G 3 4 ヌクレアーゼに対する配列切断選好を例示する。（図 11 B）は、

M G 3 4 - 1 が P A M から約 7 位で切断することを選択することを実証する、M G 3 4 - 1 の切断部位の場所を示すヒストグラムを示す。

【図 1 1 C】M G 3 4 ヌクレアーゼに対する配列切断選好を例示する。(図 1 1 C) は、M G 3 4 - 9 の好ましい N G G P A M を示すサンガー配列決定クロマトグラムを示す(ボックスで強調表示される)。矢印は、P A M から 7 位の切断部位を示す。

【図 1 2 A】M G 3 4 - 1 に対する E . c o l i におけるプラスミド標的化実験の結果を例示する。(図 1 2 A) は、プラスミド切断を実証する E . c o l i 株のレプリカプレーティングを示し、M G 3 4 - 1 及び s g R N A を発現する E . c o l i を、s g R N A の標的を含有するカナマイシン耐性プラスミド(+ s p) で形質転換した。成長障害(+ s p) 対陰性対照(標的及び P A M なし(- s p)) を示すプレート四分円は、酵素による標的化及び切断の成功を示す。実験を 2 回繰り返し、3 回実施した。

10

【図 1 2 B】M G 3 4 - 1 に対する E . c o l i におけるプラスミド標的化実験の結果を例示する。(図 1 2 B) A におけるレプリカプレーティング実験からのコロニー形成単位(c f u) 測定値のグラフを示し、標的条件(+ s p) 対非標的対照(- s p) における成長抑制を示し、プラスミドが切断されたことを実証する。(図 1 2 C) は、標的条件(白色バー) 対非標的対照(緑色バー) における E . c o l i 成長抑制を示すコロニー形成単位(c f u) 測定値(対数スケール) のバープロットを示す。

【図 1 2 C】M G 3 4 - 1 に対する E . c o l i におけるプラスミド標的化実験の結果を例示する。(図 1 2 C) は、標的条件(白色バー) 対非標的対照(緑色バー) における E . c o l i 成長抑制を示すコロニー形成単位(c f u) 測定値(対数スケール) のバープロットを示す。各ヌクレアーゼに対するプラスミド干渉アッセイを、S p C a s 9 陽性対照とともに 3 回行った。

20

【図 1 3 A】M G 3 5 - 4 1 9 に対する S M A R T 系の例示的なゲノムコンテキストを示す。S M A R T ヌクレアーゼは、濃灰色の矢印で示され、他の遺伝子は、淡灰色の矢印で図示される。ゲノム断片中の全ての遺伝子について予測されるドメインが、矢印の下に灰色のボックスとして示される。環境発現配列決定リードは、C R I S P R アレイの下(図 1 3 A)、及びエフェクターの上流(図 1 3 B) に整列して示されている。発現を示す領域のトランスクリプトーム対象範囲がコンティグ配列の上に例示される。(図 1 3 A) は、近傍でコードされる S M A R T I I M G 3 5 - 4 1 9 エフェクター及び C R I S P R 遺伝子座のゲノムコンテキストを示す。

30

【図 1 3 B】M G 3 5 - 4 1 9 に対する S M A R T 系の例示的なゲノムコンテキストを示す。S M A R T ヌクレアーゼは、濃灰色の矢印で示され、他の遺伝子は、淡灰色の矢印で図示される。ゲノム断片中の全ての遺伝子について予測されるドメインが、矢印の下に灰色のボックスとして示される。環境発現配列決定リードは、C R I S P R アレイの下(図 1 3 A)、及びエフェクターの上流(図 1 3 B) に整列して示されている。発現を示す領域のトランスクリプトーム対象範囲がコンティグ配列の上に例示される。(図 1 3 B) は、転写された 5' U T R を示す S M A R T I I エフェクター M G 3 5 - 3 のゲノムコンテキストを示す。

【図 1 4】S M A R T I I M G 3 5 - 4 1 9 の 3 D 構造予測を示す。この 3 D モデルは、そのサイズの半分未満であるにもかかわらず、S a C a s 9 結晶構造の領域と良好に整列する。S a C a s 9 鋳型と整列した領域は、触媒ローブ(R u v C - I、H N H、及び R u v C - I I ドメイン) 及び認識(R E C) ローブの短い領域を含む。S M A R T I I 特異的ドメインは、R R X R R モチーフ及び P f a m P F 1 4 2 3 9 に対する相同性を含有するドメインと、未知の機能のドメインと、を含む。

40

【図 1 5】S M A R T I I エフェクターに対する予備切断アッセイの結果を図示する。M G 3 5 - 4 2 0 (配列番号 2 2 3) タンパク質調製物を、遺伝子座全体が発現された T X T L 抽出物における切断活性について試験した。実験は、P A M ライブラリー(d s D N A 標的)、順配向及び逆配向の両方(f w 及び r v) の遺伝子座(c r 1) で予測される反復領域、並びに関連する補因子を潜在的にコードする遺伝子間領域を用いて、タンパク質調製物をインキュベートした。レーン 2 ~ 9 (c r アレイなし) : 反復領域なしの対

50

照実験。A p o：標的P A Mライブラリーを有するタンパク質調製物のみ。ラベル1～2・5は、7つの異なる遺伝子間領域を表す。－I G：対照として遺伝子間領域を含まなかった。ライゲーション産物のP C Rゲルは、d s D N A切断を示唆する推定切断バンド（矢印）を示す。

【図16A】S M A R T系のゲノムコンテキストを図示する。S M A R Tヌクレアーゼは、濃灰色の矢印で示され、他の遺伝子は、淡灰色の矢印で図示される。ゲノム断片中の全ての遺伝子について予測されるドメインが、矢印の下に灰色のボックスとして示される。環境発現配列決定リードは、エフェクターの上流に整列して示されている。図16Aは、S M A R T I I M G 3 5－4 1 9エフェクターのゲノムコンテキストを図示する。

【図16B】S M A R T系のゲノムコンテキストを図示する。S M A R Tヌクレアーゼは、濃灰色の矢印で示され、他の遺伝子は、淡灰色の矢印で図示される。ゲノム断片中の全ての遺伝子について予測されるドメインが、矢印の下に灰色のボックスとして示される。環境発現配列決定リードは、エフェクターの上流に整列して示されている。図16Bは、S M A R T I I M G 3 5－1 0 2エフェクターのゲノムコンテキストを図示する。

【図17A】M G 3 5－4 2 0が活性d s D N Aヌクレアーゼであることを実証するデータを図示する。図17Aは、M G 3 4－4 2 0エフェクターのゲノムコンテキストを図示する。エフェクターは、逆配向の暗い矢印で表され、予測されるP F A Mドメインは、矢印の下の長方形で表され、ガイドR N Aをコードする可能性のある遺伝子間領域は、黒色線上に「I G」として注釈付けされる。C R I S P R様反復領域がコンティグに存在する。

【図17B】M G 3 5－4 2 0が活性d s D N Aヌクレアーゼであることを実証するデータを図示する。図17Bは、T X T Lにおける切断活性について試験された精製タンパク質調製物の結果を図示する。実験は、P A Mライブラリー（d s D N A標的）、順配向及び逆配向の両方（f w及びr v）の遺伝子座（c r 1）で予測されるC R I S P R様反復領域、並びに関連する補因子を潜在的にコードする遺伝子間領域を用いて、タンパク質調製物をインキュベートした。レーン2～9（c r アレイなし）：反復領域なしの対照実験。A p o：標的P A Mライブラリーを有するタンパク質調製物のみ。ラベル1～2・5は、7つの異なる遺伝子間領域を表す。－I G：対照として遺伝子間領域を含まなかった。ライゲーション産物のP C Rゲルは、d s D N A切断を示唆する推定切断バンド（矢印）を示す。「4」と標識されたレーン上に回収されたバンドは、酵素をC R I S P R様領域及びS M A R T I I 5' U T Rとインキュベートすることからの切断バンドを表す。

【図18】M G 3 5－4 2 0の予測されるガイドR N Aを図示する。図18のAは、S M A R T I Iエフェクターとその5' U T Rとのインビトロ転写反応から配列決定されたR N A S e qリードを示す、M G 3 4－4 2 0エフェクターのゲノムコンテキストを図示する。エフェクターは、逆配向の暗い矢印で表され、予測されるP F A Mドメインは、矢印の下の長方形で表され、予測されるガイドR N Aは、黒色線上に注釈付けされる。図18のBは、S M A R T I I M G 3 5－4 2 0推定ガイドR N Aの二次構造表現を図示する。

【図19A】S M A R T I Iエフェクターと関連付けられた保存されたU T R領域の多重配列アラインメント（M S A）を図示する。図19Aは、S M A R T I Iエフェクターの開始コドンのすぐ上流の領域の全長M S Aを図示する。アラインメントの上の同一性パーセントヒストグラムは、保存領域を示す（5' U T RガイドR N A、灰色矢印として注釈付けされる）。

【図19B】S M A R T I Iエフェクターと関連付けられた保存されたU T R領域の多重配列アラインメント（M S A）を図示する。図19Bは、推定ガイドR N Aコード配列内の高度に保存された領域を図示する。同一性パーセントヒストグラム及び配列ロゴ表現が、アラインメントの上に示されている。同一塩基は、黒色のボックスで強調表示されている。

【図20A】M G 3 5エフェクターがs g R N Aを使用する活性d s D N Aヌクレアーゼであることを実証するデータを図示する。図20Aは、インビトロ切断アッセイの結果を

図示する。(sg)を含み、(Apo)sgRNAを含まないエフェクターを、PAMライブラリー(dsDNA標的)とインキュベートされたインビトロ転写／翻訳反応でアッセイした。切断産物を、PCR(予想されるサイズのヌクレアーゼ産生バンドによる成功したRNAガイド切断、矢印)を介して増幅した。

【図20B】MG35エフェクターがsgRNAを使用する活性dsDNAヌクレアーゼであることを実証するデータを図示する。図20Bは、標的隣接モチーフ(TAM)を図示する。

【図21-1】SMART酵素が多様な標的化能力を有する新規のヌクレアーゼであることを実証するデータを図示する。図21のAは、SMARTヌクレアーゼ対SpCas9の予測されるドメインアーキテクチャを図示する。

【図21-2】SMART酵素が多様な標的化能力を有する新規のヌクレアーゼであることを実証するデータを図示する。図21のBは、SMART MG102-2系のゲノムコンテキストを図示する。tracrRNA及びCRISPRアレイの配向を、エフェクターとのインビトロ切断活性によって確認した。

【図21-3】SMART酵素が多様な標的化能力を有する新規のヌクレアーゼであることを実証するデータを図示する。図21のCは、SMART MG34-1系のゲノムコンテキストを図示する。適応モジュール遺伝子(Cas1、Cas2、Cas4、及び推定Csn2)を特定した。環境RNASeqリードを、アレイ及びtracrRNAをコードする遺伝子間領域に対して順配向にマッピングした。遺伝子座にコードされる他の遺伝子は、黄色の矢印で表される。tracrRNA及びCRISPRアレイの配向を、エ

フェクターとのインビトロ切断活性によって確認した。

【図21-4】SMART酵素が多様な標的化能力を有する新規のヌクレアーゼであることを実証するデータを図示する。図21のDは、2つの活性SMART HEAROヌクレアーゼのHEARO RNA二次構造を図示する。コンセンサス標的モチーフ配列のSeqLogo表現が示される。

【図21-5】SMART酵素が多様な標的化能力を有する新規のヌクレアーゼであることを実証するデータを図示する。図21のEは、SMARTヌクレアーゼ対Cas9及びIsdB参照配列のタンパク質系統樹を図示する。SMARTエフェクター及び古細菌Cas9配列(青緑色及び紫色の分岐)は、文書化されたCas9参照配列(II-A型、II-B型、及びII-C型、灰色分岐)と遠く関連している。系統樹を、共有されたRuvc-II/HNH/Ruvc-II IDメインの多重配列アラインメントから推定した。ヌクレアーゼのSMART MG33ファミリー(赤紫色分岐)は、CRISPR II-C型バリエーション系とクラスター化するが、一方で他のCRISPR関連SMARTヌクレアーゼ(青緑色分岐)は、最近II-D型として分類された配列とクラスター化する。SMART HEAROヌクレアーゼ(薄紫色分岐)は、HEARO ORF及びIsdB配列とクラスター化する。図21のFは、SMART CRISPR II型ファミリーの系統クレードを図示する。クレードは、図21のEに図示される系統樹の表現の拡大図である。内部ファミリー分割ノードのローカルサポート値が示されており、0~1の範囲である。活性SMARTヌクレアーゼに対する生化学的切断活性アッセイからのコンセンサス標的モチーフ配列及びsgRNA設計のSeqLogo表現が示される。

【図22A】SMART IのものがdsDNAヌクレアーゼであることを実証するデータを図示する。図22Aは、MG34-1がPAMから7位で優先的にdsDNAを切断することを示す、切断位置選好のヒストグラムを図示する。挿入図は、MG34-1が互い違いの切断を生成し、3位の切断が標的ストランド(TS)上で発生し、6~7位の切断が非標的ストランド(NTS)上で発生することを示す。

【図22B】SMART IのものがdsDNAヌクレアーゼであることを実証するデータを図示する。図22Bは、様々なスペーサー長を有するDNA切断パーセントの分布を図示し、MG34-1に対する18bpスペーサーに対する選好を示す。

【図22C】SMART IのものがdsDNAヌクレアーゼであることを実証するデータを図示する。図22Cは、MG34-1の時系列切断アッセイを図示しており、SpC

10

20

30

40

50

a s 9と比較して速度が遅いことを示唆している。

【図22D】SMART IのものがdsDNAヌクレアーゼであることを実証するデータを図示する。図22Dは、プラスミド標的化アッセイを図示する。左：方法の図は、エフェクターヌクレアーゼ（MG34-1又はMG34-9）及びsgRNA補因子を発現する、操作されたE. coli株を示す。標的スパーサー又は非標的スパーサー（陰性対照）を有する抗生物質耐性遺伝子を含むプラスミドで形質転換されたとき、標的プラスミドに対して成長障害が発生する。中央及び右：MG34-1（中央）又はMG34-9（右）酵素及びsgRNAをコードするプラスミドについて、およそ2倍の成長抑制を示す棒グラフ。

【図23】SMART HNHエンドヌクレアーゼ関連RNA及びORF（HEARO）（35-1、35-2、35-3、35-6、35-102、及びIsCB）並びにSMART（34-1、102-2、102-14、102-35、102-45）ヌクレアーゼの群についてのタンパク質全長にわたるアミノ酸含有率を図示する。アルギニン（R）及びリジン（K）の高含有率は、緑色で強調表示されるが、メチオニン（M）の低含有率は、オレンジ色で強調表示される。Uniref50データベース（Carugo, vol. 17, 12 (2008) : 2187-91）中のほとんどのタンパク質のアミノ酸含有率を比較に使用した。

【図24A】Uniref50データベース内のタンパク質の平均アミノ酸含有量（X軸）対SMARTタンパク質内のアミノ酸含有量のパーセンテージ（Y軸）の散布図を図示する。アルギニン（R）及びリジン（K）含有量は、線形傾向から逸脱する。

【図24B】SMARTタンパク質中のアミノ酸パーセンテージとUniref50データベース中のパーセンテージとの比を示すグラフを図示する。全ての比の平均は0.99であり、SDは0.22である。緑色の線は、正規性を仮定して、平均から2つの標準偏差を示す。

【図25A】SMART酵素がdsDNAヌクレアーゼであることを実証するデータを図示する。図25Aは、次世代シーケンシング（NGS）からの非標的鎖（NTS）上の3つのSMARTヌクレアーゼに対する切断位置選好のヒストグラムを図示する。挿入図は、SMARTヌクレアーゼが互い違いの切断を生成することを示し、3位の切断が標的鎖（TS）上で発生するが、一方で、PAMからの5～7位の切断がNTS上で発生する。TS切断部位を、サンガーランオフシーケンシングを介して決定した。

【図25B】SMART酵素がdsDNAヌクレアーゼであることを実証するデータを図示する。図25Bは、標的対非標的対照におけるE. coli成長抑制を示すコロニー形成単位（cfu）測定値（対数スケール）のバープロットを図示する。各ヌクレアーゼに対するプラスミド干渉アッセイを、SpCas9陽性対照とともに3回行った。

【図25C】SMART酵素がdsDNAヌクレアーゼであることを実証するデータを図示する。図25Cは、様々なスパーサー長を有するインビトロDNA切断効率の測定を図示し、SMARTヌクレアーゼについては18～20bpスパーサーに対する選好を示すが、一方で、SMART HEARO 35-1は24bpスパーサーを選好する。（*）スパーサー長14bp（34-1）並びに30bp（35-1及び102-2）は、評価されなかった。

【図25D】SMART酵素がdsDNAヌクレアーゼであることを実証するデータを図示する。図25Dは、PAMから-1～-13位の標的スパーサーに対する高い特異性を示すミスマッチ殺傷アッセイを図示する。左：標的対非標的対照における、E. coli成長抑制を示すコロニー形成単位（cfu）測定値（対数スケール）のバープロット。右上：ミスマッチ殺傷アッセイの図。ヌクレアーゼ発現及びガイド発現のための2つのプラスミドを含むE. coliは、プロトスパーサー内にミスマッチを有する標的プラスミドのライブラリーで形質転換される。右下：標的スパーサーの各位置におけるミスマッチ耐性を示すヒートマップ。許容されるミスマッチを有する標的スパーサーについては、成長が抑制されると予想される（紫色）。必要な塩基対合を有する位置は、効率的に切断されず、出力ライブラリー（黄色）内で比較

的豊富となる。各ヌクレアーゼに対するライブラリーを用いたプラスミド干渉（殺傷）アッセイを二重で行った。

【図26】MG102-2がヒト細胞における高活性ヌクレアーゼであることを実証するデータを図示する。ヌクレアーゼ活性を、MG102-2 mRNA及びTRAC遺伝子座内の2つのsgRNA標的化部位（ガイドA1及びB1）を、増加する濃度のsgRNA（150、300、及び450 pmol／反応）でヌクレオフェクトすることによって試験した。模擬対照は、mRNA及びガイドの非存在下で標的領域におけるバックグラウンド編集レベルを表す。

【図27】MG102-2及びMG35-1について試験されたスパーサーの各位置にミスマッチを有するスパーサーの対数倍率変化切断活性を示すミスマッチ殺傷アッセイを図示する。

10

【図28】SMARTヌクレアーゼがssDNA上で活性を呈さないことを実証するデータを図示する。

【図29】SMARTヌクレアーゼのガイド及び塩濃度滴定を図示する。MG102-2（レーン1～6）及びSMART HEARO 35-1（レーン7～18）のインビトロ切断アッセイは、直線DNA産物（2500 bp未満）への標的プラスミドDNA（約3500 bp）の切断を示す。

【図30A】ヒト細胞におけるSMART I編集効率を実証するデータを図示する。ヌクレアーゼ活性を、遺伝子座内の多重部位を標的化するSMART I mRNA及びsgRNA（450 pmol／反応）をヌクレオフェクトすることによって試験した。各バーは、特定のスパーサー（ガイド）によって標的化された部位における編集効率を表す。図30Aは、AAVS1遺伝子座を標的化するMG102-2のデータを図示する。

20

【図30B】ヒト細胞におけるSMART I編集効率を実証するデータを図示する。ヌクレアーゼ活性を、遺伝子座内の多重部位を標的化するSMART I mRNA及びsgRNA（450 pmol／反応）をヌクレオフェクトすることによって試験した。各バーは、特定のスパーサー（ガイド）によって標的化された部位における編集効率を表す。図30B、図30C、図30D、図30E、図30F、及び図30Gはそれぞれ、TRAC遺伝子座を標的化するMG102-39、MG102-42、MG102-48、MG33-34、MG102-26、及びMG102-45のデータを図示する。

【図30C】ヒト細胞におけるSMART I編集効率を実証するデータを図示する。ヌクレアーゼ活性を、遺伝子座内の多重部位を標的化するSMART I mRNA及びsgRNA（450 pmol／反応）をヌクレオフェクトすることによって試験した。各バーは、特定のスパーサー（ガイド）によって標的化された部位における編集効率を表す。図30B、図30C、図30D、図30E、図30F、及び図30Gはそれぞれ、TRAC遺伝子座を標的化するMG102-39、MG102-42、MG102-48、MG33-34、MG102-26、及びMG102-45のデータを図示する。

30

【図30D】ヒト細胞におけるSMART I編集効率を実証するデータを図示する。ヌクレアーゼ活性を、遺伝子座内の多重部位を標的化するSMART I mRNA及びsgRNA（450 pmol／反応）をヌクレオフェクトすることによって試験した。各バーは、特定のスパーサー（ガイド）によって標的化された部位における編集効率を表す。図30B、図30C、図30D、図30E、図30F、及び図30Gはそれぞれ、TRAC遺伝子座を標的化するMG102-39、MG102-42、MG102-48、MG33-34、MG102-26、及びMG102-45のデータを図示する。

40

【図30E】ヒト細胞におけるSMART I編集効率を実証するデータを図示する。ヌクレアーゼ活性を、遺伝子座内の多重部位を標的化するSMART I mRNA及びsgRNA（450 pmol／反応）をヌクレオフェクトすることによって試験した。各バーは、特定のスパーサー（ガイド）によって標的化された部位における編集効率を表す。図30B、図30C、図30D、図30E、図30F、及び図30Gはそれぞれ、TRAC遺伝子座を標的化するMG102-39、MG102-42、MG102-48、MG33-34、MG102-26、及びMG102-45のデータを図示する。

50

【図30F】ヒト細胞におけるSMART I編集効率を実証するデータを図示する。ヌクレアーゼ活性を、遺伝子座内の多重部位を標的化するSMART I mRNA及びsgRNA (450 pmol/反応)をヌクレオフェクトすることによって試験した。各バーは、特定のスペーサー(ガイド)によって標的化された部位における編集効率を表す。図30B、図30C、図30D、図30E、図30F、及び図30Gはそれぞれ、TRAC遺伝子座を標的化するMG102-39、MG102-42、MG102-48、MG33-34、MG102-26、及びMG102-45のデータを図示する。

【図30G】ヒト細胞におけるSMART I編集効率を実証するデータを図示する。ヌクレアーゼ活性を、遺伝子座内の多重部位を標的化するSMART I mRNA及びsgRNA (450 pmol/反応)をヌクレオフェクトすることによって試験した。各バーは、特定のスペーサー(ガイド)によって標的化された部位における編集効率を表す。図30B、図30C、図30D、図30E、図30F、及び図30Gはそれぞれ、TRAC遺伝子座を標的化するMG102-39、MG102-42、MG102-48、MG33-34、MG102-26、及びMG102-45のデータを図示する。

【図31】4つのSMART HEAROヌクレアーゼの5' UTRヌクレオチド配列の多重配列アラインメントを図示する。HEARO RNA(ボックス)の開始前の領域は、不十分な類似性を示すが、一方で、第1の構造的ヘアピンの周囲の強い保存が観察される(挿入図)。

【図32A】SMART HEAROヌクレアーゼのゲノムコンテキストを図示する。SMART HEAROヌクレアーゼの大部分はCRISPR関連ではないが(例えば、MG35-104、図32A)、一方で、少数のSMART HEAROヌクレアーゼが、CRISPRアレイと関連付けられている(例えば、図32B及び図32CのMG35-463及びMG35-556)。SMART HEAROヌクレアーゼは、遺伝子の下にRRXRR及びHNH Pfamドメインが注釈付けされている濃灰色の矢印で表される。共分散モデル(CM)から予測されるHEARO RNAは、SMART HEAROエフェクター遺伝子(CM HEARO RNA)の上流に示される。RAR:リピーターアンチリピーター。

【図32B】SMART HEAROヌクレアーゼのゲノムコンテキストを図示する。SMART HEAROヌクレアーゼの大部分はCRISPR関連ではないが(例えば、MG35-104、図32A)、一方で、少数のSMART HEAROヌクレアーゼが、CRISPRアレイと関連付けられている(例えば、図32B及び図32CのMG35-463及びMG35-556)。SMART HEAROヌクレアーゼは、遺伝子の下にRRXRR及びHNH Pfamドメインが注釈付けされている濃灰色の矢印で表される。共分散モデル(CM)から予測されるHEARO RNAは、SMART HEAROエフェクター遺伝子(CM HEARO RNA)の上流に示される。RAR:リピーターアンチリピーター。

【図32C】SMART HEAROヌクレアーゼのゲノムコンテキストを図示する。SMART HEAROヌクレアーゼの大部分はCRISPR関連ではないが(例えば、MG35-104、図32A)、一方で、少数のSMART HEAROヌクレアーゼが、CRISPRアレイと関連付けられている(例えば、図32B及び図32CのMG35-463及びMG35-556)。SMART HEAROヌクレアーゼは、遺伝子の下にRRXRR及びHNH Pfamドメインが注釈付けされている濃灰色の矢印で表される。共分散モデル(CM)から予測されるHEARO RNAは、SMART HEAROエフェクター遺伝子(CM HEARO RNA)の上流に示される。RAR:リピーターアンチリピーター。

【図32D】SMART HEAROヌクレアーゼのゲノムコンテキストを図示する。SMART HEAROヌクレアーゼの大部分はCRISPR関連ではないが(例えば、MG35-104、図32A)、一方で、少数のSMART HEAROヌクレアーゼが、CRISPRアレイと関連付けられている(例えば、図32B及び図32CのMG35-463及びMG35-556)。SMART HEAROヌクレアーゼは、遺伝子の下に

R R X R R 及び H N H P f a m ドメインが注釈付けされている濃灰色の矢印で表される。共分散モデル (C M) から予測される H E A R O RNA は、S M A R T H E A R O エフェクター遺伝子 (C M H E A R O RNA) の上流に示される。R A R : リピーターアンチリピーター。図 3 2 D ~ 図 3 2 G は、3 つの活性ヌクレアーゼ : それぞれ、M G 3 5 - 1 0 4 s g 1、M G 3 5 - 4 6 3 s g 2 (C R I S P R 非依存性)、M G 3 5 - 4 6 3 s g 3 (C R I S P R 関連)、及び M G 3 5 - 5 5 6 二重ガイド H E A R O RNA (C R I S P R 関連) の H E A R O RNA 二次構造を図示する。

【図 3 2 E】S M A R T H E A R O ヌクレアーゼのゲノムコンテキストを図示する。S M A R T H E A R O ヌクレアーゼの大部分は C R I S P R 関連ではないが (例えば、M G 3 5 - 1 0 4、図 3 2 A)、一方で、少数の S M A R T H E A R O ヌクレアーゼが、C R I S P R アレイと関連付けられている (例えば、図 3 2 B 及び図 3 2 C の M G 3 5 - 4 6 3 及び M G 3 5 - 5 5 6)。S M A R T H E A R O ヌクレアーゼは、遺伝子の下に R R X R R 及び H N H P f a m ドメインが注釈付けされている濃灰色の矢印で表される。共分散モデル (C M) から予測される H E A R O RNA は、S M A R T H E A R O エフェクター遺伝子 (C M H E A R O RNA) の上流に示される。R A R : リピーターアンチリピーター。図 3 2 D ~ 図 3 2 G は、3 つの活性ヌクレアーゼ : それぞれ、M G 3 5 - 1 0 4 s g 1、M G 3 5 - 4 6 3 s g 2 (C R I S P R 非依存性)、M G 3 5 - 4 6 3 s g 3 (C R I S P R 関連)、及び M G 3 5 - 5 5 6 二重ガイド H E A R O RNA (C R I S P R 関連) の H E A R O RNA 二次構造を図示する。

【図 3 2 F】S M A R T H E A R O ヌクレアーゼのゲノムコンテキストを図示する。S M A R T H E A R O ヌクレアーゼの大部分は C R I S P R 関連ではないが (例えば、M G 3 5 - 1 0 4、図 3 2 A)、一方で、少数の S M A R T H E A R O ヌクレアーゼが、C R I S P R アレイと関連付けられている (例えば、図 3 2 B 及び図 3 2 C の M G 3 5 - 4 6 3 及び M G 3 5 - 5 5 6)。S M A R T H E A R O ヌクレアーゼは、遺伝子の下に R R X R R 及び H N H P f a m ドメインが注釈付けされている濃灰色の矢印で表される。共分散モデル (C M) から予測される H E A R O RNA は、S M A R T H E A R O エフェクター遺伝子 (C M H E A R O RNA) の上流に示される。R A R : リピーターアンチリピーター。図 3 2 D ~ 図 3 2 G は、3 つの活性ヌクレアーゼ : それぞれ、M G 3 5 - 1 0 4 s g 1、M G 3 5 - 4 6 3 s g 2 (C R I S P R 非依存性)、M G 3 5 - 4 6 3 s g 3 (C R I S P R 関連)、及び M G 3 5 - 5 5 6 二重ガイド H E A R O RNA (C R I S P R 関連) の H E A R O RNA 二次構造を図示する。

【図 3 2 G】S M A R T H E A R O ヌクレアーゼのゲノムコンテキストを図示する。S M A R T H E A R O ヌクレアーゼの大部分は C R I S P R 関連ではないが (例えば、M G 3 5 - 1 0 4、図 3 2 A)、一方で、少数の S M A R T H E A R O ヌクレアーゼが、C R I S P R アレイと関連付けられている (例えば、図 3 2 B 及び図 3 2 C の M G 3 5 - 4 6 3 及び M G 3 5 - 5 5 6)。S M A R T H E A R O ヌクレアーゼは、遺伝子の下に R R X R R 及び H N H P f a m ドメインが注釈付けされている濃灰色の矢印で表される。共分散モデル (C M) から予測される H E A R O RNA は、S M A R T H E A R O エフェクター遺伝子 (C M H E A R O RNA) の上流に示される。R A R : リピーターアンチリピーター。図 3 2 D ~ 図 3 2 G は、3 つの活性ヌクレアーゼ : それぞれ、M G 3 5 - 1 0 4 s g 1、M G 3 5 - 4 6 3 s g 2 (C R I S P R 非依存性)、M G 3 5 - 4 6 3 s g 3 (C R I S P R 関連)、及び M G 3 5 - 5 5 6 二重ガイド H E A R O RNA (C R I S P R 関連) の H E A R O RNA 二次構造を図示する。

【図 3 3 - 1】インビトロの S M A R T H E A R O 切断活性を図示する。S M A R T I I エフェクターを、それらの単一ガイド RNA 及び P A M ライブラリー (d s DNA 標的) とインキュベートされたインビトロ転写 / 翻訳反応でアッセイした。切断産物を、切断部位へのライゲーション及び後続の P C R (予想されるサイズのヌクレアーゼ産生バンドによる成功した RNA ガイド切断、矢印) を介して増幅した。図 3 3 の A について、レーンラベルは次のとおりである。L : ラダー、P C : 陽性対照 (P C) としての M G 3 5 - 1 ヌクレアーゼ、1 : M G 3 5 - 9 4、2 : M G 3 5 - 1 0 4、3 : M G 3 5 - 3 4 6

10

20

30

40

50

、4：MG35-350、5：MG35-423、6：MG35-422、7：MG35-461、8：MG35-465、9：MG35-515。図33のBについて、レーンラベルは次のとおりである。L：ラダー、PC：陽性対照（PC）としてのMG35-1ヌクレアーゼ、10：MG35-517、11：sgRNA設計1によるMG35-518、12：sgRNA設計2によるMG35-518、13：MG35-519、14：sgRNA設計1によるMG35-550、15：sgRNA設計2によるMG35-550、16：MG35-553、17：sgRNA設計1によるMG35-554、18：sgRNA設計2によるMG35-554、19：MG35-555、及び20：MG35-556。

【図33-2】インビトロのSMART HEARO切断活性を図示する。SMART IIEフェクターを、それらの単一ガイドRNA及びPAMライブラリー（dsDNA標的）とインキュベートされたインビトロ転写／翻訳反応でアッセイした。切断産物を、切断部位へのライゲーション及び後続のPCR（予想されるサイズのヌクレアーゼ産生バンドによる成功したRNAガイド切断、矢印）を介して増幅した。図33のCについて、SMART IIEフェクターを、TAM／PAM濃縮プロトコルを介して切断活性についてアッセイした。エフェクターを、それらの単一ガイドRNAの存在下でインビトロ転写／翻訳（IVTT）反応で発現させ、次いで、PAMライブラリー（dsDNA標的）に加えた。切断産物を、切断部位へのライゲーション及び後続のPCR（予想されるサイズのヌクレアーゼ産生バンドによる成功したRNAガイド切断、矢印）を介して増幅した。示される反応は、PCRクリーンアップの前であるため、プライマー及びアダプター二量体バンドは、＜100bpのサイズで観察される。

【図34-1】活性SMART HEAROヌクレアーゼのTAM認識モチーフを図示する。図33のA～図33のCで特定されたバンドのNGS配列決定を使用して、各ヌクレアーゼについてTAM及び好ましい切断位置を生成した。Geneious（Andronescu 2007）によって予測されたワーキングガイドの構造が、はめ込みで示されている。切断は、通常、非標的鎖上の5～10位の間で発生する。

【図34-2】活性SMART HEAROヌクレアーゼのTAM認識モチーフを図示する。図33のA～図33のCで特定されたバンドのNGS配列決定を使用して、各ヌクレアーゼについてTAM及び好ましい切断位置を生成した。Geneious（Andronescu 2007）によって予測されたワーキングガイドの構造が、はめ込みで示されている。切断は、通常、非標的鎖上の5～10位の間で発生する。

【図34-3】活性SMART HEAROヌクレアーゼのTAM認識モチーフを図示する。図33のA～図33のCで特定されたバンドのNGS配列決定を使用して、各ヌクレアーゼについてTAM及び好ましい切断位置を生成した。Geneious（Andronescu 2007）によって予測されたワーキングガイドの構造が、はめ込みで示されている。切断は、通常、非標的鎖上の5～10位の間で発生する。

【図35A】活性SMART HEAROヌクレアーゼのインビトロ切断効率を図示する。図35Aについて、反応産物の超らせん状（未切断）から直線（切断）への遷移によって切断を測定し、Agilent TapeStation上で可視化した。矢印は、酵素による成功した標的化された切断（直線化）の後の初期dsDNA産物（超らせん状）及びdsDNA産物を示す。PE：PURExpress、sgRNA、単一ガイドRNA。

【図35B】活性SMART HEAROヌクレアーゼのインビトロ切断効率を図示する。図35Bは、図35Aからの定量化のバースプロット表現を図示する。DNA：RNP反応なしのDNAのみの対照（陰性対照）、Apo：sgRNA追加なしのRNP反応、Holo：sgRNAを伴うRNP反応。

【図36A】SMART HEAROガイドエンジニアリングを図示する。5つの活性SMART HEARO sgRNAは、それらの配列に1つ以上のポリTトラクトを有した。3つのポリT変異体sgRNAを、活性対元のガイドと比較するように候補ごとに設計した。ガイドをインビトロ転写し、同じ濃度に正規化し、次いで、インビトロ切断効率

10

20

30

40

50

反応で使用了。図 3 6 A は、MG 3 5 - 5 1 8 のポリ T 領域及び操作されたガイド配列を有する例示的なガイド RNA を図示する。

【図 3 6 B】SMART HEARO ガイドエンジニアリングを図示する。5 つの活性 SMART HEARO sgRNA は、それらの配列に 1 つ以上のポリ T トラクトを有した。3 つのポリ T 変異体 sgRNA を、活性対元のガイドと比較するように候補ごとに設計した。ガイドをインビトロ転写し、同じ濃度に正規化し、次いで、インビトロ切断効率反応で使用了。図 3 6 B は、操作された SMART HEARO ガイド RNA 対天然ガイドの切断効率を図示する。Apo: ガイド追加なし (陰性対照)、WT: 天然ガイド RNA。

【図 3 7 A】SMART イヌクレアーゼの系統解析を図示する。系統樹を、グローバル (g-ins-i) 又はローカル (l-ins-i) の多重配列アラインメントからの FastTree 又は RAxML で推定した。系統の不確実性を考慮するために、6 つの再構成された配列を複数のツリー (黒丸で強調表示されたノード: MG 3 4 - 2 6、MG 3 4 - 2 7、MG 3 4 - 2 8、MG 3 4 - 2 9、MG 3 4 - 3 0、及び MG 3 4 - 3 1) から得た。

【図 3 7 B】SMART イヌクレアーゼの系統解析を図示する。系統樹を、グローバル (g-ins-i) 又はローカル (l-ins-i) の多重配列アラインメントからの FastTree 又は RAxML で推定した。系統の不確実性を考慮するために、6 つの再構成された配列を複数のツリー (黒丸で強調表示されたノード: MG 3 4 - 2 6、MG 3 4 - 2 7、MG 3 4 - 2 8、MG 3 4 - 2 9、MG 3 4 - 3 0、及び MG 3 4 - 3 1) から得た。

【図 3 7 C】SMART イヌクレアーゼの系統解析を図示する。系統樹を、グローバル (g-ins-i) 又はローカル (l-ins-i) の多重配列アラインメントからの FastTree 又は RAxML で推定した。系統の不確実性を考慮するために、6 つの再構成された配列を複数のツリー (黒丸で強調表示されたノード: MG 3 4 - 2 6、MG 3 4 - 2 7、MG 3 4 - 2 8、MG 3 4 - 2 9、MG 3 4 - 3 0、及び MG 3 4 - 3 1) から得た。

【図 3 7 D】SMART イヌクレアーゼの系統解析を図示する。系統樹を、グローバル (g-ins-i) 又はローカル (l-ins-i) の多重配列アラインメントからの FastTree 又は RAxML で推定した。系統の不確実性を考慮するために、6 つの再構成された配列を複数のツリー (黒丸で強調表示されたノード: MG 3 4 - 2 6、MG 3 4 - 2 7、MG 3 4 - 2 8、MG 3 4 - 2 9、MG 3 4 - 3 0、及び MG 3 4 - 3 1) から得た。

【図 3 8】再構成された SMART イ MG 3 4 - 3 0 の 3 D 構造予測対活性 MG 3 4 - 1 イヌクレアーゼの予測される構造を図示する。タンパク質の全体的な良好な構造的アラインメントを、2 つの構造間の重複、及び低い RMSD 値によって観察した。

【図 3 9】再構成された SMART イエフェクターが活性イヌクレアーゼであることを実証するデータを図示する。新規の SMART イエフェクターを、PAM 濃縮プロトコルを介して切断活性についてアッセイした。イエフェクターを、他の活性 MG 3 4 イヌクレアーゼからの単一ガイド RNA の存在下でインビトロ転写／翻訳 (IVTT) 反応で発現させ、PAM ライブラリー (dsDNA 標的) に加えた。切断産物を、切断部位へのライゲーション及び後続の PCR 増幅 (予想される 180 bp サイズのイヌクレアーゼ産生バンドによる成功した RNA ガイド切断、矢印) を介して増幅した。MG 3 4 - 2 7 及び MG 3 4 - 2 9 は、3 つの試験されたガイド RNA で明確な活性を示した。

【図 4 0】計算再構成からの活性 SMART イヌクレアーゼの PAM 認識モチーフを図示する。図 3 9 で特定されたバンドの NGS 配列決定を使用して、各イヌクレアーゼについて PAM 及び好ましい切断位置を生成した。切断は、非標的鎖上の PAM からの 6 ~ 8 位の間で発生する。

【0047】

配列表の簡単な説明

本明細書とともに提出される配列表は、本開示による方法、組成物、及び系で使用するための例示的なポリヌクレオチド配列及びポリペプチド配列を提供する。以下は、その中の配列の例示的な説明である。

【0048】

MG33ヌクレアーゼ

配列番号1、463～486、981～988、及び1289～1312は、MG33ヌクレアーゼの完全長ペプチド配列を示す。

【0049】

配列番号199及び669～670は、MG33ヌクレアーゼと機能するように予測されたtracrRNAのヌクレオチド配列を示す。

【0050】

配列番号201及び1003～1005は、MG33ヌクレアーゼと機能するように予測された、予測された単一ガイドRNA (sgRNA) 配列のヌクレオチド配列を示す。

「N」は、可変残基を示し、非N残基は、足場配列を表す。

【0051】

配列番号1023～1028は、MG33ヌクレアーゼと適合するPAM配列を示す。

【0052】

配列番号1045～1054は、本明細書に説明されるMG33ヌクレアーゼのCRISPRリピートを示す。

【0053】

MG34ヌクレアーゼ

配列番号2～24、487～488、及び1313～1321は、MG34ヌクレアーゼの完全長ペプチド配列を示す。

【0054】

配列番号200は、MG34ヌクレアーゼと機能するように予測されたtracrRNAのヌクレオチド配列を示す。

【0055】

配列番号202、203、及び613～616は、MG34ヌクレアーゼと機能するように予測された、予測された単一ガイドRNA (sgRNA) 配列のヌクレオチド配列を示す。「N」は、可変残基を示し、非N残基は、足場配列を表す。

【0056】

配列番号1023～1028は、MG34ヌクレアーゼと適合するPAM配列を示す。

【0057】

配列番号1055～1057は、本明細書に説明されるMG34ヌクレアーゼのCRISPRリピートを示す。

【0058】

MG35ヌクレアーゼ

配列番号25～198、221～459、489～580、617～668、及び674～675は、MG35ヌクレアーゼの完全長ペプチド配列を示す。

【0059】

配列番号460～461は、MG35ヌクレアーゼと同じ遺伝子座に由来するMG35tracrRNAのヌクレオチド配列を示す。

【0060】

配列番号462、676、及び1229～1230は、本明細書に説明されるMG35ヌクレアーゼのCRISPRリピートを示す。

【0061】

配列番号677～686、1006～1012、及び1231～1259は、MG35単一ガイドRNAのヌクレオチド配列を示す。

【0062】

配列番号687～974は、MG35単一ガイドRNAコード配列のヌクレオチド配列

10

20

30

40

50

を示す。

【0063】

配列番号1029～1034は、MG35ヌクレアーゼと適合するPAM配列を示す。

【0064】

配列番号1172～1228は、本明細書に説明されるMG35ヌクレアーゼをコードする遺伝子座のヌクレオチド配列を示す。

【0065】

MG102ヌクレアーゼ

配列番号581～612、989～1002、及び1260～1273は、MG102ヌクレアーゼの完全長ペプチド配列を示す。

【0066】

配列番号672～673は、MG102ヌクレアーゼと同じ遺伝子座に由来するMG102 tracrRNAのヌクレオチド配列を示す

【0067】

配列番号205～220は、本開示によるヌクレアーゼに付加され得る例示的な核局在化配列(NLS)の配列を示す。

【0068】

配列番号1013～1022は、MG102単一ガイドRNAのヌクレオチド配列を示す。

【0069】

配列番号1035～1044は、MG102ヌクレアーゼと適合するPAM配列を示す。

【0070】

配列番号1058～1072は、本明細書に説明されるMG102ヌクレアーゼのCRISPRリピートを示す。

【0071】

配列番号1171は、本明細書に説明されるMG102ヌクレアーゼをコードする遺伝子座のヌクレオチド配列を示す。

【0072】

MG143ヌクレアーゼ

配列番号975は、MG143ヌクレアーゼの完全長ペプチド配列を示す。

【0073】

配列番号1073は、本明細書に説明されるMG143ヌクレアーゼのCRISPRリピートを示す。

【0074】

MG144ヌクレアーゼ

配列番号976～979及び1274～1288は、MG144ヌクレアーゼの完全長ペプチド配列を示す。

【0075】

配列番号1074～1077は、本明細書に説明されるMG144ヌクレアーゼのCRISPRリピートを示す。

【0076】

MG145ヌクレアーゼ

配列番号980は、MG145ヌクレアーゼの完全長ペプチド配列を示す。

【0077】

配列番号1078は、本明細書に説明されるMG145ヌクレアーゼのCRISPRリピートを示す。

【0078】

MG102 TRAC標的化

配列番号1079～1082及び1145～1166は、TRAC標的部位のDNA配

10

20

30

40

50

列を示す。

【0079】

配列番号1083～1086及び1123～1144は、TRACを標的化するために、MG102ヌクレアーゼと機能するように操作されたsgRNAのヌクレオチド配列を示す。

【0080】

MG33 TRAC標的化

配列番号1167～1168は、TRACを標的化するために、MG33ヌクレアーゼと機能するように操作されたsgRNAのヌクレオチド配列を示す。

【0081】

配列番号1169～1170は、TRAC標的部位のDNA配列を示す。

【0082】

AAVS1標的化

配列番号1087～1104は、AAVS1を標的化するために、MG102ヌクレアーゼと機能するように操作されたsgRNAのヌクレオチド配列を示す。

【0083】

配列番号1105～1122は、AAVS1標的部位のDNA配列を示す。

【発明を実施するための形態】

【0084】

本発明の様々な実施形態は本明細書に示され、説明されるが、このような実施形態が、例示の目的でのみ提供されることは、当業者には明らかであろう。多数の変形、変更、及び置換は、本発明から逸脱することなく、当業者にとって生じ得る。本明細書に説明される本発明の実施形態に対する様々な代替が用いられ得ることは、理解されるべきである。

【0085】

本明細書に開示されるいくつかの方法の実践は、別段の示唆がない限り、免疫学、生化学、化学、分子生物学、微生物学、細胞生物学、ゲノミクス、及び組換えDNAの技術を用いる。例えば、Sambrook and Green, Molecular Cloning: A Laboratory Manual, 4th Edition (2012); the series Current Protocols in Molecular Biology (F. M. Ausubel, et al. eds.); the series Methods In Enzymology (Academic Press, Inc.), PCR 2: A Practical Approach (M. J. MacPherson, B. D. Hames and G. R. Taylor eds. (1995)), Harlow and Lane, eds. (1988) Antibodies, A Laboratory Manual, and Culture of Animal Cells: A Manual of Basic Technique and Specialized Applications, 6th Edition (R. I. Freshney, ed. (2010)) (参照により本明細書に完全に組み込まれる)を参照のこと。

【0086】

本明細書において使用される場合、単数形「a」、「an」及び「the」は、文脈が別途明確に示さない限り、複数形も含むことが意図される。更に、用語「含むこと」、「含む」、「有すること」、「有する」、「有する」、又はそのバリエーションが、詳細な説明又は特許請求の範囲のいずれかで使用される限りにおいて、かかる用語は、用語「含むこと」と類似した様式で包含的であることが意図される。

【0087】

用語「約」又は「およそ」は、当業者によって決定される特定の値についての許容可能な誤差範囲内を意味し、これは、値がどのように測定又は決定されるか、すなわち、測定系の限界に部分的に依存する。例えば、「約」は、当該技術分野の慣行によると、1又は2つ以上の標準偏差内を意味し得る。あるいは、「約」は、所与の値の最大20%、最大

10

20

30

40

50

15%、最大10%、最大5%、又は最大1%の範囲を意味し得る。

【0088】

本明細書で使用される場合、「細胞」は概して生物学的細胞を指す。細胞は、生きている生物の基本的な構造、機能、又は生物学的単位であり得る。細胞は、1つ以上の細胞を有する任意の生物を起源とし得る。いくつかの非限定的な例としては、原核細胞、真核生物の細胞、細菌細胞、古細菌細胞、単一細胞の真核生物の細胞、原虫細胞、植物由来の細胞（例えば、植物作物、果実、野菜、穀物、大豆、トウモロコシ、トウモロコシ、小麦、種子、トマト、米、キャッサバ、サトウキビ、カボチャ、乾草、ジャガイモ、綿、大麻、タバコ、顕花植物、針葉樹、裸子植物、シダ、ヒカゲノカズラ、ツノゴケ、ゼニゴケ、コケ由来の細胞）、藻類の細胞（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、脊椎動物（例えば、フルーツフライ、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、両生類、爬虫類、鳥類、哺乳類）由来の細胞、哺乳類（例えば、ブタ、ウシ、ヤギ、ヒツジ、齧歯類、ラット、マウス、非ヒト霊長類、ヒトなど）由来の細胞などが挙げられる。いくつかの場合では、細胞は、天然の生物に由来するものではない（例えば、細胞は、合成的に作製されてもよく、時には人工細胞と呼ばれることがある）。

【0089】

本明細書で使用される場合、用語「ヌクレオチド」は、概して、塩基-糖-リン酸の組み合わせを指す。ヌクレオチドは、合成ヌクレオチドを含んでもよい。ヌクレオチドは、合成ヌクレオチド類似体を含んでもよい。ヌクレオチドは、核酸配列（例えば、デオキシリボ核酸（DNA）及びリボ核酸（RNA））の単量体単位であってもよい。ヌクレオチドという用語は、リボヌクレオシド三リン酸アデノシン三リン酸（ATP）、ウリジン三リン酸（UTP）、シトシン三リン酸（CTP）、グアノシン三リン酸（GTP）及びデオキシリボヌクレオシド三リン酸、例えば、dATP、dCTP、dITP、dUTP、dGTP、dTTP、又はその誘導体を含み得る。かかる誘導体としては、例えば、 $[\alpha S]$ dATP、 γ -デアザ-dGTP及び γ -デアザ-dATP、並びにそれらを含む核酸分子にヌクレアーゼ耐性を付与するヌクレオチド誘導体を挙げることができる。本明細書で使用される場合、ヌクレオチドという用語は、ジデオキシリボヌクレオシド三リン酸（ddNTP）及びその誘導体を指す場合がある。ジデオキシリボヌクレオシド三リン酸の例としては、ddATP、ddCTP、ddGTP、ddITP、及びddTTPが挙げられるが、これらに限定されない。ヌクレオチドは、光学的に検出可能な部分（例えば、フルオロフォア）を含む部分を使用するなど、非標識又は検出可能に標識されてもよい。標識はまた、量子ドットを用いて行われてもよい。検出可能な標識としては、例えば、放射性同位元素、蛍光標識、化学発光標識、生物発光標識、及び酵素標識を挙げることができる。ヌクレオチドの蛍光標識は、フルオレセイン、5-カルボキシフルオレセイン（FAM）、2',7'-ジメトキシ-4',5'-ジクロロ-6-カルボキシフルオレセイン（JOE）、ローダミン、6-カルボキシローダミン（R6G）、N, N, N', N'-テトラメチル-6-カルボキシローダミン（TAMRA）、6-カルボキシ-X-ローダミン（ROX）、4-（4'-ジメチルアミノフェニルアゾ）安息香酸（DABCYL）、カスケードブルー、オレゴングリーン、テキサスレッド、シアニン及び5-（2'-アミノエチル）アミノナフタレン-1-スルホン酸（EDANS）を含むが、これらに限定されない。蛍光標識されたヌクレオチドの具体的な例としては、Perkin Elmer、Foster City、Califから入手可能な[R6G] dUTP、[TAMRA] dUTP、[R110] dCTP、[R6G] dCTP、[TAMRA] dCTP、[JOE] ddATP、[R6G] ddATP、[FAM] ddCTP、[R110] ddCTP、[TAMRA] ddGTP、[ROX] ddTTP、[dR6G] ddATP、[dR110] ddCTP、[dTAMRA] ddGTP、及び[dROX] ddTTP

10

20

30

40

50

；A m e r s h a m、A r l i n g t o n H e i g h t s、I l l. から入手可能なフルオロ結合デオキシヌクレオチド、フルオロ結合Cy3-dCTP、フルオロ結合Cy5-dCTP、フルオロ結合フルオロX-dCTP、フルオロ結合Cy3-dUTP、及びフルオロ結合Cy5-dUTP；B o e h r i n g e r M a n n h e i m、I n d i a n a p o l i s、I n d. から入手可能なフルオレsein-15-dATP、フルオレsein-12-dUTP、テトラメチルローダミン-6-dUTP、IR770-9-dATP、フルオレsein-12-ddUTP、フルオレsein-12-UTP、及びフルオレsein-15-2'-dATP；並びにM o l e c u l a r P r o b e s、E u g e n e、O r e gから入手可能な染色体標識ヌクレオチド、B O D I P Y - F L - 1 4 - U T P、B O D I P Y - F L - 4 - U T P、B O D I P Y - T M R - 1 4 - U T P、B O D I P Y - T M R - 1 4 - d U T P、B O D I P Y - T R - 1 4 - U T P、B O D I P Y - T R - 1 4 - d U T P、カスケードブルー-7-UTP、カスケードブルー-7-dUTP、フルオレsein-12-UTP、フルオレsein-12-dUTP、オレゴングリーン488-5-dUTP、ローダミンググリーン-5-UTP、ローダミンググリーン-5-dUTP、テトラメチルローダミン-6-UTP、テトラメチルローダミン-6-dUTP、テキサスレッド-5-UTP、テキサスレッド-5-dUTP、及びテキサスレッド-12-dUTPを挙げることができる。ヌクレオチドはまた、化学修飾によって標識又は印付けられてもよい。化学修飾された単一ヌクレオチドは、ビオチン-dNTPであり得る。ビオチン化dNTPのいくつかの非限定的な例としては、ビオチン-dATP（例えば、ビオチン-N6-ddATP、ビオチン-14-dATP）、ビオチン-dCTP（例えば、ビオチン-11-dCTP、ビオチン-14-dCTP）、及びビオチン-dUTP（例えば、ビオチン-11-dUTP、ビオチン-16-dUTP、ビオチン-20-dUTP）が挙げられ得る。ヌクレオチドは、ヌクレオチド類似体を含んでもよい。一部の実施形態では、ヌクレオチドアナログは、ヌクレオチドのある特定の化学特性を変化させるが、ヌクレオチドアナログがその意図された機能（例えば、RNA又はDNA中の他のヌクレオチドへのハイブリダイゼーション）を行う能力を保持するように、任意の位置で修飾される天然ヌクレオチドの構造を含んでもよい。誘導体化され得るヌクレオチドの位置の例としては、5位、例えば、5-（2-アミノ）プロピルウリジン、5-プロモウリジン、5-プロピンウリジン、5-プロベニルウリジンなど；6位、例えば、6-（2-アミノ）プロピルウリジン；アデノシン、又はグアノシンについては8位、例えば、8-プロモグアノシン、8-クロログアノシン、8-フルオログアノシンなどが挙げられる。ヌクレオチドアナログはまた、デアザヌクレオチド、例えば、7-デアザーアデノシン：O-及びN-修飾（例えば、アルキル化、例えば、N6-メチルアデノシン）ヌクレオチド、並びにH e r d e w i j n, A n t i s e n s e N u c l e i c A c i d D r u g D e v., 2000 Aug. 10 (4) : 297-310に説明されるものなどの他の複素環修飾されたヌクレオチドアナログを含む。ヌクレオチドアナログはまた、ヌクレオチドの糖部分に対する修飾を含んでもよい。例えば、2' OH基は、H、OR、R、F、Cl、Br、I、SH、SR、NH₂、NHR、NR₂、COOR、又はORから選択される基によって置き換えられてもよく、式中、Rは、置換又は非置換のC1-C6アルキル、アルケニル、アルキニル、アリールなどである。他の可能な修飾としては、米国特許第5,858,988号及び同第6,291,438号に説明されるものが挙げられる。誘導体化され得るヌクレオチドの位置の例としては、5位、例えば、5-（2-アミノ）プロピルウリジン、5-プロモウリジン、5-プロピンウリジン、5-プロベニルウリジンなど；6位、例えば、6-（2-アミノ）プロピルウリジン；アデノシン、又はグアノシンについては8位、例えば、8-プロモグアノシン、8-クロログアノシン、8-フルオログアノシンなどが挙げられる。ヌクレオチドアナログはまた、デアザヌクレオチド、例えば、7-デアザーアデノシン：O-及びN-修飾（例えば、アルキル化、例えば、N6-メチルアデノシン）ヌクレオチド、並びにH e r d e w i j n, A n t i s e n s e N u c l e i c A c i d D r u g D e v., 2000 Aug. 10 (4) : 297-310に説明されるものなどの他の複素環修飾されたヌクレオチドアナ

10

20

30

40

50

ログを含む。

ヌクレオチドアナログはまた、ヌクレオチドの糖部分に対する修飾を含んでもよい。例えば、2' OH基は、H、OR、R、F、Cl、Br、I、SH、SR、NH₂、NHR、NR₂、COOR、又はORから選択される基によって置き換えられてもよく、式中、Rは、置換又は非置換のC₁–C₆アルキル、アルケニル、アルキニル、アリールなどである。他の可能な修飾としては、米国特許第5,858,988号及び同第6,291,438号に説明されるものが挙げられる。

【0090】

用語「ポリヌクレオチド」、「オリゴヌクレオチド」、及び「核酸」は、概して、デオキシリボヌクレオチド若しくはリボヌクレオチド、又はその類似体のいずれかを、一本鎖、二本鎖、又は多本鎖の形態で、任意の長さのヌクレオチドのポリマー形態を指すように互換的に使用される。ポリヌクレオチドは、細胞にとって外因性又は内因性であってもよい。ポリヌクレオチドは、無細胞環境に存在してもよい。ポリヌクレオチドは、遺伝子又はそのフラグメントであってもよい。ポリヌクレオチドは、DNAであってもよい。ポリヌクレオチドは、RNAであってもよい。ポリヌクレオチドは、任意の三次元構造を有してもよく、任意の機能を発揮してもよい。ポリヌクレオチドは、1つ以上の類似体（例えば、改変された骨格、糖、又は核酸塩基）を含んでもよい。存在する場合、ヌクレオチド構造に対する修飾は、ポリマーのアセンブリの前又は後に付与されてもよい。アナログのいくつかの非限定的な例としては、5-ブロモウラシル、ペプチド核酸、異種核酸、モルホリノ、ロックド核酸、グリコール核酸、トレオース核酸、ジデオキシヌクレオチド、コーディセピン、7-デアザ-GTP、フルオロフォア（例えば、糖に結合したローダミン又はフルオレセイン）、チオール含有ヌクレオチド、ビオチン結合ヌクレオチド、蛍光塩基アナログ、CpGアイランド、メチル-7-グアノシン、メチル化ヌクレオチド、イノシン、チオウリジン、シュードウリジン、ジヒドロウリジン、クエオシン、及びワイオシンが挙げられる。ポリヌクレオチドの非限定的な例としては、遺伝子又は遺伝子断片のコード又は非コード領域、結合分析から定義した遺伝子座（遺伝子座）、エクソン、イントロン、メッセンジャーRNA（mRNA）、トランスファーRNA（tRNA）、リボソームRNA（rRNA）、短い干渉RNA（siRNA）、短いヘアピンRNA（shRNA）、マイクロRNA（miRNA）、リボザイム、cDNA、組換えポリヌクレオチド、分岐ポリヌクレオチド、プラスミド、ベクター、任意の配列の単離されたDNA、任意の配列の単離されたRNA、細胞を含まないDNA（cfDNA）及び細胞を含まないRNA（cfRNA）を含む細胞を含まないポリヌクレオチド、核酸プローブ、並びにプライマーが挙げられる。ヌクレオチドの配列は、非ヌクレオチド成分によって中断され得る。

【0091】

用語「トランスフェクション」又は「トランスフェクトされた」は、概して、非ウイルス又はウイルスベースの方法による細胞内への核酸の導入を指す。核酸分子は、完全なタンパク質又はその機能的部分をコードする遺伝子配列であってもよい。例えば、Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, 18.1–18.88（参照により本明細書に完全に組み込まれる）を参照のこと。

【0092】

用語「ペプチド」、「ポリペプチド」、及び「タンパク質」は、本明細書において互換的に使用され、概して、ペプチド結合によって結合された少なくとも2つのアミノ酸残基のポリマーを指す。この用語は、ポリマーの特定の長さを意味しておらず、ペプチドが組換え技術、化学若しくは酵素合成を使用して産生されるか、又は天然に存在するかを暗示又は区別することを意図するものではない。この用語は、天然に存在するアミノ酸ポリマー並びに少なくとも1つの修飾アミノ酸を含むアミノ酸ポリマーに適用する。ある場合では、ポリマーは、非アミノ酸によって中断されてもよい。用語は、全長タンパク質、及び二次又は三次構造を有するか又は有さないタンパク質（例えば、ドメイン）を含む、任意

の長さのアミノ酸鎖を含む。用語はまた、例えば、ジスルフィド結合形成、グリコシル化、脂質形成、アセチル化、リン酸化、酸化、及び標識成分とのコンジュゲーションなどの任意の他の操作によって修飾されたアミノ酸ポリマーを包含する。本明細書で使用される場合、「アミノ酸」及び「複数のアミノ酸」という用語は、概して、修飾アミノ酸及びアミノ酸類似体を含むが、これに限定されない天然及び非天然アミノ酸を指す。修飾アミノ酸は、天然アミノ酸及び非天然アミノ酸を含んでもよく、これは、アミノ酸上に天然に存在しない基又は化学的部分を含むように化学的に修飾されている。アミノ酸類似体は、アミノ酸誘導体を指す場合がある。用語「アミノ酸」は、D-アミノ酸及びL-アミノ酸の両方を含む。

【0093】

本明細書で使用される場合、用語「非天然」は、概して、天然の核酸又はタンパク質に存在しない核酸又はポリペプチド配列を指すことができる。非天然は、親和性タグを指してもよい。非天然は、融合物を指してもよい。非天然は、変異、挿入、又は欠失を含む、天然に存在する核酸又はポリペプチド配列を指してもよい。非天然配列は、非天然配列が融合される核酸配列又はポリペプチド配列によっても呈され得る活性（例えば、酵素活性、メチルトランスフェラーゼ活性、アセチルトランスフェラーゼ活性、キナーゼ活性、ユビキチン化活性など）を呈していてもよく、又はコードしてもよい。非天然核酸又はポリペプチド配列を、遺伝子操作によって天然に生じる核酸又はポリペプチド配列（若しくはそのバリエーション）に連結して、キメラ核酸又はポリペプチドをコードするキメラ核酸又はポリペプチド配列を生成してもよい。

【0094】

本明細書で使用される場合、用語「プロモーター」は、概して、遺伝子の転写又は発現を制御し、RNA転写が開始されるヌクレオチド又はヌクレオチドの領域に隣接するか、又は重複して位置し得る調節DNA領域を指す。プロモーターは、しばしば転写因子と呼ばれるタンパク質因子に結合する特定のDNA配列を含有してもよく、これは、RNAポリメラーゼのDNAへの結合を促進し、これにより、遺伝子転写をもたらす。「コアプロモーター」とも呼ばれる「基礎プロモーター」は、概して、作動可能に連結されたポリヌクレオチドの転写発現を促進するための全ての塩基性要素を含有するプロモーターを指してもよい。真核生物の基礎プロモーターは、必ずしもではないが典型的には、TATA-ボックス又はCAATボックスを含む。

【0095】

本明細書で使用される場合、用語「発現」は、概して、核酸配列若しくはポリヌクレオチドがDNA鋳型から（例えば、mRNA若しくは他のRNA転写物に）転写されるプロセス、又は転写mRNAが続いてペプチド、ポリペプチド、若しくはタンパク質に翻訳されるプロセスを指す。転写物及びコードされたポリペプチドは、「遺伝子産物」と総称され得る。ポリヌクレオチドがゲノムDNAに由来する場合、発現は、真核細胞におけるmRNAのスプライシングを含み得る。

【0096】

本明細書で使用される場合、「作動可能に連結された」、「作動可能な連結」、「作動可能に連結された」、又はその文法的な均等物は、概して、遺伝子エレメント、例えば、プロモーター、エンハンサー、ポリアデニル化配列などの並列化を指し、ここで、エレメントは、それらが予想される様式で作動することを可能にする関係にある。例えば、プロモーター配列又はエンハンサー配列を含み得る、調節エレメントは、調節エレメントが、コード配列の転写を開始するのを助ける場合、コード領域に作動可能に連結される。この機能的関係が維持される限り、制御エレメントとコード領域との間に介在する残基があってもよい。

【0097】

本明細書で使用される場合、「ベクター」は、概して、ポリヌクレオチドを含むか、又はポリヌクレオチドと会合し、細胞へのポリヌクレオチドの送達を媒介するために使用され得る高分子又は高分子の会合を指す。ベクターの例としては、プラスミド、ウイルスベ

10

20

30

40

50

クター、リボソーム、及び他の遺伝子送達ビヒクルが挙げられる。ベクターは、概して、標的中の遺伝子の発現を促進するために遺伝子に作動可能に連結された、遺伝子エレメント、例えば、調節エレメントを含む。

【0098】

本明細書で使用される場合、「発現カセット」及び「核酸カセット」は、ともに発現されるか、又は発現のために作動可能に連結される、核酸配列又はエレメントの組み合わせを指すために概して互換的に使用される。いくつかの場合では、発現カセットは、発現のために作動可能に連結されている調節エレメントと遺伝子又は複数の遺伝子との組み合わせを指す。

【0099】

DNA又はタンパク質配列の「機能的断片」は、概して、全長DNA又はタンパク質配列の生物学的活性と実質的に類似した生物学的活性（機能的又は構造的のいずれか）を保持する断片を指す。DNA配列の生物学的活性は、全長配列に起因する様式で発現に影響を与える能力であってもよい。

【0100】

本明細書で使用される場合、「操作された」物体は、概して、物体がヒトの介入によって修飾されたことを示す。非限定的な実施例によれば、核酸は、その配列を、天然では生じない配列に改変することによって修飾されてもよく、核酸は、ライゲーションされた産物が、オリジナルの核酸に存在しない機能を有するように、天然では関連しない核酸にライゲーションすることによって修飾されてもよく、操作された核酸は、天然では存在しない配列を用いてインビトロで合成されてもよく、タンパク質は、天然では存在しない配列にそのアミノ酸配列を変更することによって修飾されてもよく、操作されたタンパク質は、新しい機能又は特性を獲得してもよい。「操作された」系は、少なくとも1つの操作された構成要素を含む。

【0101】

本明細書で使用される場合、用語「最適に整列された」は、概して、最高の同一性パーセントスコアを与えるか、又は合致した残基の数を最大化する2つのアミノ酸配列のアライメントを指す。

【0102】

本明細書において使用される場合、「合成」及び「人工」は、天然に存在するヒトタンパク質と低い配列同一性（例えば、50%未満の配列同一性、25%未満の配列同一性、10%未満の配列同一性、5%未満の配列同一性、1%未満の配列同一性）を有するタンパク質又はそのドメインを指すために互換的に使用される。例えば、VPRドメイン及びVP64ドメインは、合成トランス活性化ドメインである。

【0103】

本明細書で使用される場合、用語「tracrRNA」又は「tracr配列」は、概して、野生型の例示的なtracrRNA配列（例えば、S. pyogenes S. aureusなど由来のtracrRNA）と少なくとも約5%、10%、20%、30%、40%、50%、60%、70%、80%、90%、95%、若しくは100%の配列同一性又は配列類似性を有する核酸を指すことができる。tracrRNAは、野生型の例示的なtracrRNA配列（例えば、S. pyogenes S. aureusなど由来のtracrRNAなど）と最大約5%、10%、20%、30%、40%、50%、60%、70%、80%、90%、若しくは100%の配列同一性又は配列類似性を有する核酸を指すことができる。tracrRNAは、欠失、挿入、若しくは置換、バリエーション、変異、又はキメラなどのヌクレオチドの変更を含み得るtracrRNAの修飾形態を指すことができる。tracrRNAは、少なくとも6つの連続するヌクレオチドのストレッチにわたって、野生型の例示的なtracrRNA（例えば、S. pyogenes S. aureusなど由来のtracrRNA）配列と少なくとも約60%同一であり得る核酸を指してもよい。例えば、tracrRNA配列は、少なくとも6つの連続するヌクレオチドの区画にわたって野生型の例となるtracrRNA（例えば、S. p

10

20

30

40

50

y o g e n e s S . a u r e u s など由来の t r a c r R N A) 配列と少なくとも約 60 % 同一、少なくとも約 65 % 同一、少なくとも約 70 % 同一、少なくとも約 75 % 同一、少なくとも約 80 % 同一、少なくとも約 85 % 同一、少なくとも約 90 % 同一、少なくとも約 95 % 同一、少なくとも約 98 % 同一、少なくとも約 99 % 同一、又は 100 % 同一であり得る。I I 型 t r a c r R N A 配列は、隣接する C R I S P R アレイ中の反復配列の一部に対して相補性を有する領域を特定することによって、ゲノム配列上で予測することができる。

【0104】

本明細書において使用される場合、「ガイド核酸」は、概して、別の核酸にハイブリダイズし得る核酸を指すことができる。ガイド核酸は、RNAであってもよい。ガイド核酸は、DNAであってもよい。ガイド核酸は、部位特異的に核酸の配列に結合するようにプログラムされてもよい。標的化されるべき核酸、又は標的核酸は、ヌクレオチドを含んでもよい。ガイド核酸は、ヌクレオチドを含んでもよい。標的核酸の一部は、ガイド核酸の一部に対して相補的であってもよい。ガイド核酸に相補的であり、ガイド核酸とハイブリダイズする二本鎖標的ポリヌクレオチドの鎖を、相補鎖と呼ぶことができる。相補鎖に相補的であり、それゆえ、ガイド核酸に相補的ではない二本鎖標的ポリヌクレオチドの鎖は、非相補鎖と呼ばれてもよい。ガイド核酸は、ポリヌクレオチド鎖を含んでもよく、また「単一ガイド核酸」と呼ばれ得る。ガイド核酸は、2つのポリヌクレオチド鎖を含んでもよく、また「ダブルガイド核酸」と呼ばれ得る。そうでない場合、用語「ガイド核酸」は、単一ガイド核酸とダブルガイド核酸の両方を指す、包括的であってもよい。ガイド核酸は、「核酸標的化セグメント」又は「核酸標的化配列」と称され得るセグメントを含み得る。核酸標的化セグメントは、「タンパク質結合セグメント」又は「タンパク質結合配列」又は「C a s タンパク質結合セグメント」と称され得るサブセグメントを含んでもよい。

【0105】

2つ以上の核酸又はポリペプチド配列の文脈における用語「配列同一性」又は「同一性パーセント」は、概して、配列比較アルゴリズムを使用して測定された場合、局所比較ウィンドウ又はグローバル比較ウィンドウにわたって最大対応のために比較及び整列されたとき、同一であるか、又は特定のパーセンテージの、同一であるアミノ酸残基又はヌクレオチドを有する、2つ（例えば、ペアワイズアラインメントで）又はそれ以上（例えば、多重配列アラインメントにおける）の配列を指す。ポリペプチド配列に好適な配列比較アルゴリズムとしては、例えば、3のワード長（W）、10の期待値（E）のパラメータ、及び11の存在、1の延長でギャップコストを設定しているB L O S U M 6 2 スコアリングマトリックスを使用し、かつ30残基より長いポリペプチド配列についての条件付き組成スコアマトリックス調整を使用したB L A S T P；2のワード長（W）、100000の期待値（E）のパラメータ、及びオープンギャップに対して9及び30残基より短い配列についての拡張ギャップに対して1でのP A M 3 0 スコアリング設定ギャップコストを使用したB L A S T P（これらは、<https://blast.ncbi.nlm.nih.gov>で入手可能なB L A S T スイートにおいてB L A S T P についてのデフォルトパラメータである）；又は2の一致、-1のミスマッチ、及び-1のギャップのパラメータを有するS m i t h - W a t e r m a n 相同性検索アルゴリズムのパラメータを用いたC L U S T A L W；デフォルトパラメータを用いたM U S C L E；2のパラメータリツリー及び1000の最大反復を用いたM A F F T；デフォルトパラメータを用いたN o v a f o l d；デフォルトパラメータを用いたH M M E R h m m a l i g n が挙げられる。

【0106】

本明細書において使用される場合、用語「R u v C _ I I I ドメイン」は、概して、R u v C エンドヌクレアーゼドメインの第3の不連続セグメントを指す（R u v C エンドヌクレアーゼドメインは、3つの不連続セグメントであるR u v C _ I、R u v C _ I I、及びR u v C _ I I I から構成される）。R u v C ドメイン又はそのセグメント（例えば、R u

v C __ I、RuvC __ II、又はRuvC __ III）は、概して、文書化されているドメイン配列へのアラインメント、注釈付きドメインを有するタンパク質への構造的アラインメント、又は文書化されているドメイン配列（例えば、RuvC __ IIIのPfam HMM PF18541）に基づいて構築された隠れマルコフモデル（HMM）との比較によって特定することができる。

【0107】

本明細書において使用される場合、用語「HNHドメイン」は、概して、特徴的なヒスチジン残基及びアスパラギン残基を有するエンドヌクレアーゼドメインを指す。HNHドメインは、概して、文書化されているドメイン配列へのアラインメント、注釈付きドメインを有するタンパク質への構造的アラインメント、又は文書化されているドメイン配列（例えば、ドメインHNHのPfam HMM PF01844）に基づいて構築された隠れマルコフモデル（HMM）との比較によって特定することができる。

【0108】

本明細書で使用される場合、用語「ブリッジヘリックスドメイン」又は「BHドメイン」は、概して、標的DNAの結合時に切断活性を開始する上で重要な役割を果たすCas酵素中に存在するアルギニンリッチヘリックスドメインを指す。

【0109】

本明細書で使用される場合、用語「認識ドメイン」又は「RECドメイン」は、概して、gRNAのリピート：アンチリピート二重鎖と相互作用し、Casエンドヌクレアーゼ／gRNA複合体の形成を媒介すると考えられるドメインを指す。

【0110】

本明細書で使用される場合、用語「ウェッジドメイン」又は「WEDドメイン」は、概して、4つのアルファヘリックスに隣接したねじれた5本鎖ベータシートを含むフォールドを指し、これは、概して、Cas酵素に対する歪んだリピート：アンチリピート二重鎖の認識に関与する。WEDドメインは、単一ガイドRNA足場の認識に関与し得る。

【0111】

本明細書で使用される場合、用語「PAM相互作用ドメイン」又は「PIDドメイン」は、概して、ガイドRNAの非相補的DNA鎖上のPAM配列を認識するようにエンドヌクレアーゼDNA複合体に位置付けられたCas酵素に見出されるドメインを指す。

【0112】

概要

固有の機能性及び構造を有する新しいCas酵素の発見は、デオキシリボ核酸（DNA）編集技術を更に破壊し、速度、特異性、機能性、及び使い易さを改善する可能性を付与し得る。微生物におけるクラスター化され規則的に間隔があいた短い回文配列反復（CRISPR）系の予測される保有率及び微生物種の完全な多様性と比較して、文献には機能的に特徴付けられたCRISPR／Cas酵素が比較的少ない。これは、一部、実験室条件下では、膨大な数の微生物種が容易には培養されないためである。多数の微生物種を表す天然の環境ニッチからのメタゲノムシーケンシングは、文書化された新しいCRISPR／Cas系の数を劇的に増加させ、新しいオリゴヌクレオチド編集機能の発見を早める可能性を付与する可能性がある。このようなアプローチの実りのある最近の例は、天然微生物群集のメタゲノム分析からのCasX／CasY CRISPR系の2016年の発見によって示される。

【0113】

CRISPR／Cas系は、微生物における適応免疫系として機能すると説明されているRNA指向ヌクレアーゼ複合体である。それらの自然な状況では、CRISPR／Cas系は、CRISPR（クラスター化され規則的に間隔があいた短い回文配列反復）オペロン又は遺伝子座において生じ、これは概して、2つの部分：（i）RNAベースの標的化エレメントをコードする、同等に短いスペーサー配列によって分けられた短い反復配列（30～40bp）のアレイ；及び（ii）アクセサリタンパク質／酵素と並んだRNAベースの標的化エレメントによって指示されるヌクレアーゼポリペプチドをコードする

CasをコードするORFを含む。特定の標的核酸配列の効率的なヌクレアーゼ標的化は、概して、(i) 標的(標的シード)の最初の6~8個の核酸とcrRNAガイドとの間の相補的なハイブリダイゼーション; 及び(ii) 標的シードの画定された近傍内のプロトスペーサー隣接モチーフ(PAM)配列の存在の両方を必要とする(PAMは、通常、宿主ゲノム内に一般的に表されない配列である)。系の正確な機能及び構成に応じて、CRISPR-Cas系は、共有された機能的特徴及び進化的類似性に基づき、2つのクラス、5つのタイプ、及び16のサブタイプに一般的に構成されている。

【0114】

クラスI CRISPR-Cas系は、大型の複数サブユニットエフェクター複合体を有し、タイプI、III、及びIVを含む。

10

【0115】

タイプI CRISPR-Cas系は、構成要素の面で中程度の複雑さとみなされる。タイプI CRISPR-Cas系では、RNA標的化エレメントのアレイは、リピートエレメントで処理される長い前駆crRNA(プレcrRNA)として転写されて、ヌクレアーゼ複合体を核酸標的に配向させる短い成熟crRNAが、その後、プロトスペーサー隣接モチーフ(PAM)と呼ばれる適切な短いコンセンサス配列に引き続いて放出される。このプロセッシングは、カスケードと呼ばれる大きなエンドヌクレアーゼ複合体のエンドリボヌクレアーゼサブユニット(Cas6)を介して生じ、これは、crRNA指向性ヌクレアーゼ複合体のヌクレアーゼ(Cas3)タンパク質成分も含む。Cas Iヌクレアーゼは、主にDNAヌクレアーゼとして機能する。

20

【0116】

タイプIII CRISPR系は、Csm又はCmrタンパク質サブユニットを含む反復関連ミステリアスタンパク質(RAMP)とともに、Cas10として知られる中心ヌクレアーゼの存在によって特徴付けられ得る。タイプIの系と同様、成熟crRNAは、Cas6様酵素を使用してプレcrRNAから処理される。I及びIII型系とは異なり、III型系は、DNA-RNA二本鎖(RNAポリメラーゼのテンプレートとして使用されるDNA鎖など)を標的化し、切断すると思われる。

【0117】

IV型CRISPR-Cas系は、高度に還元された大きなサブユニットヌクレアーゼ(csf1)、Cas5(csf3)及びCas7(csf2)群のRAMPタンパク質の2つの遺伝子、並びにいくつかの場合では、予測される小さいサブユニットの遺伝子を含む、エフェクター複合体を保有し、そのような系は、一般的に内因性プラスミド上で見出される。

30

【0118】

II型CRISPR-Cas系は、概して、単一ポリペプチド複数ドメインヌクレアーゼエフェクターを有し、II型、V型、及びVI型を含む。

【0119】

II型CRISPR-Cas系は、構成要素の点で最も単純なものとみなされる。II型CRISPR-Cas系では、CRISPRアレイを成熟crRNAにプロセッシングすることは、特別なエンドヌクレアーゼサブユニットの存在を必要としないが、むしろ、アレイ反復配列に相補的な領域を有する小さなトランスコードcrRNA(tracrRNA)を必要とし、tracrRNAは、その対応するエフェクターヌクレアーゼ(例えば、Cas9)及び反復配列の両方と相互作用して、前駆dsRNA構造を形成し、これは、内因性RNase IIIで切断されて、tracrRNA及びcrRNAの両方を装填した成熟エフェクター酵素を生成する。Cas IIヌクレアーゼは、DNAヌクレアーゼである。II型エフェクターは、概して、RuvC様ヌクレアーゼドメインのフォールドに挿入された無関係なHNHヌクレアーゼドメインを有するRNase Hフォールドに適合するRuvC様エンドヌクレアーゼドメインを含む構造を呈する。RuvC様ドメインは、標的(例えば、crRNA相補的)DNA鎖の切断に関与し、一方、HNHドメインは、置換DNA鎖の切断に関与する。

40

50

【0120】

V型CRISPR-Cas系は、RuvC様ドメインを含む、II型エフェクターと類似したヌクレアーゼエフェクター（例えば、Cas12）構造によって特徴付けられる。II型と同様に、大部分の（ただし全てではない）V型CRISPR系は、tracrRNAを使用して、precrRNAを成熟crRNAにプロセスするが、precrRNAを複数のcrRNAに切断するためのRNase IIIを必要とするII型系とは異なり、V型系は、エフェクターヌクレアーゼ自体を使用してprecrRNAを切断することができる。II型CRISPR-Cas系と同様に、V型CRISPR-Cas系が、DNAヌクレアーゼである。II型CRISPR-Cas系とは異なり、いくつかのV型酵素（例えば、Cas12a）は、二本鎖標的配列の第1のcrRNA指向切断によって活性化される、強固な一本鎖非特異的デオキシリボヌクレアーゼ活性を有すると見られる。

10

【0121】

VI型CRISPR-Cas系は、RNAガイドRNAエンドヌクレアーゼを有する。RuvC様ドメインの代わりに、VI型系（例えば、Cas13）の単一ポリペプチドエフェクターは、2つのHEPNリボヌクレアーゼドメインを含む。II及びV型系の両方とは異なり、VI型系もまた、precrRNAをcrRNAに処理するためtracrRNAを必要としないと思われる。しかしながら、V型系と同様に、一部のVI型系（例えば、C2C2）は、標的RNAの第1のcrRNA指向切断によって活性化される、頑強な一本鎖非特異的ヌクレアーゼ（リボヌクレアーゼ）活性を有するようである。

20

【0122】

クラスII CRISPR-Casは、より単純な構築物であるため、設計ヌクレアーゼ／ゲノム編集適用としての操作及び開発に最も広く適用されてきた。

【0123】

インビトロ使用のためのこのような系の早期適用のうちの1つは、Jinekら（Science. 2012 Aug 17; 337(6096): 816-21、参照により本明細書に完全に組み込まれる）において見ることができる。Jinek試験では、最初に(i) S. pyogenes SF370から単離した、組換え発現した、精製された全長Cas9（例えば、クラスII、タイプII Cas酵素）、(ii) 切断されるべき標的DNA配列に相補的な約20ntの5'配列、続いて3' tracr結合配列を生じる、精製された成熟約42ntのcrRNA（T7プロモーター配列を担持する合成DNA鋳型からインビトロで転写された全crRNA）、(iii) T7プロモーター配列を担持する合成DNA鋳型からインビトロで転写された精製tracrRNA、及び(iv) Mg^{2+} を含む系を説明した。Jinekは、後に、(ii)のcrRNAが、(iii)の5'末端にリンカー（例えば、GAAA）によって結合されて、それ自体で標的にCas9を誘導することができる単一の融合合成ガイドRNA（sgRNA）を形成する、改善された操作された系を説明した（図2の上部パネルと下部パネルを比較する）。

30

【0124】

参照により本明細書に完全に組み込まれる、Maliら（Science. 2013 Feb 15; 339(6121): 823-826.）は、後に、(i) C末端核局在化配列（例えば、SV40 NLS）及び適当なポリアデニル化シグナル（例えば、TKpAシグナル）を有する適当な哺乳類プロモーター下でコドン最適化Cas9（例えば、クラスII、タイプII Cas酵素）をコードするORF、並びに(ii) 適当なポリメラーゼIIIプロモーター（例えば、U6プロモーター）下にsgRNA（Gで始まる5'配列、続いて3' tracr結合配列に結合した20ntの相補的標的化核酸配列、リンカー、及びtracrRNA配列を有する）をコードするORFをコードするDNAベクターをもたらすことによって、哺乳類細胞での使用にこの系を適合した。

40

【0125】

MG酵素

一態様では、本開示は、操作されたヌクレアーゼ系を提供する。操作されたヌクレアーゼ系は、(a) エンドヌクレアーゼを含み得る。いくつかの場合、エンドヌクレアーゼは

50

、RuvCドメイン及びHNHドメインを含む。エンドヌクレアーゼは、未培養微生物由来であり得る。エンドヌクレアーゼは、Casエンドヌクレアーゼであり得る。エンドヌクレアーゼは、クラス2エンドヌクレアーゼであり得る。エンドヌクレアーゼは、クラス2、II型Casエンドヌクレアーゼであり得る。操作されたヌクレアーゼ系は、(b)操作されたガイドリボ核酸構造を含み得る。操作されたガイドリボ核酸構造は、エンドヌクレアーゼと複合体を形成するように構成され得る。いくつかの場合、エンドヌクレアーゼと複合体を形成するように構成された操作されたガイドリボ核酸構造は、ガイドリボ核酸配列を含む。ガイドリボ核酸配列は、標的デオキシリボ核酸配列にハイブリダイズするように構成され得る。いくつかの場合、エンドヌクレアーゼと複合体を形成するように構成された操作されたガイドリボ核酸構造は、tracrリボ核酸配列を含む。tracrリボ核酸配列は、エンドヌクレアーゼに結合するように構成され得る。いくつかの場合、エンドヌクレアーゼは、約120kDa以下、約110kDa以下、約100kDa以下、約90kDa以下、約80kDa以下、約70kDa以下、約60kDa以下、約50kDa以下、約40kDa以下、約30kDa以下、約20kDa以下、又は約10kDa以下の分子量を有する。

10

【0126】

いくつかの場合、エンドヌクレアーゼは、配列番号1～198、221～459、463～612、617～668、674～675、975～1002、1260～1321のうちのいずれか1つ、又はそれらのバリエーションと少なくとも50%、少なくとも55%、少なくとも50%、少なくとも55%、少なくとも60%、少なくとも65%、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、又は少なくとも99%の配列同一性を有する配列を含む。

20

【0127】

一態様では、本開示は、操作されたヌクレアーゼ系を提供する。操作されたヌクレアーゼ系は、(a)エンドヌクレアーゼを含み得る。エンドヌクレアーゼは、RuvC-1ドメイン又はRucVドメインを含み得る。エンドヌクレアーゼは、HNHドメインを含み得る。エンドヌクレアーゼは、RuvC-1ドメイン及びHNHドメインを含み得る。エンドヌクレアーゼは、Casエンドヌクレアーゼであり得る。エンドヌクレアーゼは、クラス2エンドヌクレアーゼであり得る。エンドヌクレアーゼは、クラス2、II型Casエンドヌクレアーゼであり得る。操作されたヌクレアーゼ系は、(b)操作されたガイドリボ核酸を含み得る。操作されたガイドリボ核酸構造は、エンドヌクレアーゼと複合体を形成するように構成され得る。エンドヌクレアーゼと複合体を形成するように構成された操作されたガイドリボ核酸構造は、ガイドリボ核酸配列を含み得る。ガイドリボ核酸配列は、標的デオキシリボ核酸配列にハイブリダイズするように構成され得る。エンドヌクレアーゼと複合体を形成するように構成された操作されたガイドリボ核酸構造は、tracrリボ核酸配列を含み得る。tracrリボ核酸配列は、エンドヌクレアーゼに結合するように構成され得る。一部の実施形態では、エンドヌクレアーゼは、配列番号1～198、221～459、463～612、617～668、674～675、975～1002、1260～1321のうちのいずれか1つと少なくとも50%、少なくとも55%、少なくとも50%、少なくとも55%、少なくとも60%、少なくとも65%、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも81%、少なくとも82%、少なくとも83%、少なくとも84%、少なくとも85%、少なくとも86%、少なくとも87%、少なくとも88%、少なくとも89%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、又は少なくとも99%の配列同一性を有する配列を含み得る。エンドヌクレアーゼは、古細菌エンドヌクレアーゼであり得る。エンドヌクレアーゼは、クラス2、II型Casエンドヌクレアーゼであり得る。エンドヌクレアーゼは、RRxRRモチーフを含むアルギニンリッチ領域又はPF1423

30

40

50

9 相同性を有するドメインを含み得る。アルギニンリッチ領域又は P F 1 4 2 3 9 相同性を有するドメインは、配列番号 1 ~ 1 9 8、2 2 1 ~ 4 5 9、4 6 3 ~ 6 1 2、6 1 7 ~ 6 6 8、6 7 4 ~ 6 7 5、9 7 5 ~ 1 0 0 2、1 2 6 0 ~ 1 3 2 1 のうちのいずれか 1 つ、又はそれらのバリエーションのアルギニンリッチ領域又は P F 1 4 2 3 9 相同性を有するドメインと少なくとも 5 0 %、少なくとも 5 5 %、少なくとも 5 0 %、少なくとも 5 5 %、少なくとも 6 0 %、少なくとも 6 5 %、少なくとも 7 0 %、少なくとも 7 5 %、少なくとも 8 0 %、少なくとも 8 1 %、少なくとも 8 2 %、少なくとも 8 3 %、少なくとも 8 4 %、少なくとも 8 5 %、少なくとも 8 6 %、少なくとも 8 7 %、少なくとも 8 8 %、少なくとも 8 9 %、少なくとも 9 0 %、少なくとも 9 1 %、少なくとも 9 2 %、少なくとも 9 3 %、少なくとも 9 4 %、少なくとも 9 5 %、少なくとも 9 6 %、少なくとも 9 7 %、少なくとも 9 8 %、又は少なくとも 9 9 % の配列同一性を有する配列を含み得る。アルギニンリッチドメイン又は P F 1 4 2 3 9 相同性を有するドメインのドメイン境界は、M G 3 4 - 1 又は M G 3 4 - 9 に対する最適なアラインメントによって特定され得る。エンドヌクレアーゼは、R E C ドメインを含み得る。R E C ドメインは、配列番号 1 ~ 1 9 8、2 2 1 ~ 4 5 9、4 6 3 ~ 6 1 2、6 1 7 ~ 6 6 8、6 7 4 ~ 6 7 5、9 7 5 ~ 1 0 0 2、1 2 6 0 ~ 1 3 2 1 のうちのいずれか 1 つ、又はそれらのバリエーションの R E C ドメインと少なくとも 5 0 %、少なくとも 5 5 %、少なくとも 5 0 %、少なくとも 5 5 %、少なくとも 6 0 %、少なくとも 6 5 %、少なくとも 7 0 %、少なくとも 7 5 %、少なくとも 8 0 %、少なくとも 8 1 %、少なくとも 8 2 %、少なくとも 8 3 %、少なくとも 8 4 %、少なくとも 8 5 %、少なくとも 8 6 %、少なくとも 8 7 %、少なくとも 8 8 %、少なくとも 8 9 %、少なくとも 9 0 %、少なくとも 9 1 %、少なくとも 9 2 %、少なくとも 9 3 %、少なくとも 9 4 %、少なくとも 9 5 %、少なくとも 9 6 %、少なくとも 9 7 %、少なくとも 9 8 %、又は少なくとも 9 9 % の配列同一性を有する配列を含み得る。R E C ドメインのドメイン境界は、M G 3 4 - 1 又は M G 3 4 - 9 に対する最適なアラインメントによって特定され得る。エンドヌクレアーゼは、B H (ブリッジヘリックス) を含み得る。B H ドメインは、配列番号 1 ~ 1 9 8、2 2 1 ~ 4 5 9、4 6 3 ~ 6 1 2、6 1 7 ~ 6 6 8、6 7 4 ~ 6 7 5、9 7 5 ~ 1 0 0 2、1 2 6 0 ~ 1 3 2 1 のうちのいずれか 1 つ、又はそれらのバリエーションの B H ドメインと少なくとも 5 0 %、少なくとも 5 5 %、少なくとも 5 0 %、少なくとも 5 5 %、少なくとも 6 0 %、少なくとも 6 5 %、少なくとも 7 0 %、少なくとも 7 5 %、少なくとも 8 0 %、少なくとも 8 1 %、少なくとも 8 2 %、少なくとも 8 3 %、少なくとも 8 4 %、少なくとも 8 5 %、少なくとも 8 6 %、少なくとも 8 7 %、少なくとも 8 8 %、少なくとも 8 9 %、少なくとも 9 0 %、少なくとも 9 1 %、少なくとも 9 2 %、少なくとも 9 3 %、少なくとも 9 4 %、少なくとも 9 5 %、少なくとも 9 6 %、少なくとも 9 7 %、少なくとも 9 8 %、又は少なくとも 9 9 % の配列同一性を有する配列を含み得る。B H ドメインのドメイン境界は、M G 3 4 - 1 又は M G 3 4 - 9 に対する最適なアラインメントによって特定され得る。

10

20

30

40

50

【 0 1 2 8 】

エンドヌクレアーゼは、W E D (ウェッジ) ドメインを含み得る。W E D ドメインは、配列番号 1 ~ 1 9 8、2 2 1 ~ 4 5 9、4 6 3 ~ 6 1 2、6 1 7 ~ 6 6 8、6 7 4 ~ 6 7 5、9 7 5 ~ 1 0 0 2、1 2 6 0 ~ 1 3 2 1 のうちのいずれか 1 つ、又はそれらのバリエーションの W E D ドメインと少なくとも 5 0 %、少なくとも 5 5 %、少なくとも 5 0 %、少なくとも 5 5 %、少なくとも 6 0 %、少なくとも 6 5 %、少なくとも 7 0 %、少なくとも 7 5 %、少なくとも 8 0 %、少なくとも 8 1 %、少なくとも 8 2 %、少なくとも 8 3 %、少なくとも 8 4 %、少なくとも 8 5 %、少なくとも 8 6 %、少なくとも 8 7 %、少なくとも 8 8 %、少なくとも 8 9 %、少なくとも 9 0 %、少なくとも 9 1 %、少なくとも 9 2 %、少なくとも 9 3 %、少なくとも 9 4 %、少なくとも 9 5 %、少なくとも 9 6 %、少なくとも 9 7 %、少なくとも 9 8 %、又は少なくとも 9 9 % の配列同一性を有する配列を含み得る。W E D ドメインのドメイン境界は、M G 3 4 - 1 又は M G 3 4 - 9 に対する最適なアラインメントによって特定され得る。エンドヌクレアーゼは、P I (P A M 相互作用) ドメインを含み得る。P I ドメインは、配列番号 1 ~ 1 9 8、2 2 1 ~ 4 5 9、4 6 3 ~ 6

1 2、6 1 7～6 6 8、6 7 4～6 7 5、9 7 5～1 0 0 2、1 2 6 0～1 3 2 1のうちのいずれか1つ、又はそれらのバリエーションのPIDメインと少なくとも50%、少なくとも55%、少なくとも50%、少なくとも55%、少なくとも60%、少なくとも65%、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも81%、少なくとも82%、少なくとも83%、少なくとも84%、少なくとも85%、少なくとも86%、少なくとも87%、少なくとも88%、少なくとも89%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、又は少なくとも99%の配列同一性を有する配列を含み得る。PIDメインのドメイン境界は、MG34-1又はMG34-9に対する最適なアラインメントによって特定され得る。

10

【0129】

いくつかの場合、エンドヌクレアーゼは、未培養微生物に由来する。いくつかの実施形態では、tracrリボ核酸配列は、配列番号199～200、460～461、若しくは669～673のうちのいずれか1つから少なくとも50%、少なくとも60%、少なくとも70%、少なくとも80%個の連続するヌクレオチドと少なくとも50%、少なくとも55%、少なくとも50%、少なくとも55%、少なくとも60%、少なくとも65%、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、若しくは少なくとも99%の配列同一性を有する配列を含むか、又は配列番号201～203、613～616、677～686、1003～1022、若しくは1231～1259のうちのいずれか1つの非可変ヌクレオチドの少なくとも50%、少なくとも60%、少なくとも70%、少なくとも80%個の連続するヌクレオチドと少なくとも50%、少なくとも55%、少なくとも50%、少なくとも55%、少なくとも60%、少なくとも65%、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、若しくは少なくとも99%の配列同一性を有する配列を含む。

20

【0130】

いくつかの場合、ガイド核酸構造は、配列番号201を含む。いくつかの場合、ガイド核酸構造は、配列番号202を含む。いくつかの場合、ガイド核酸構造は、配列番号203を含む。いくつかの場合、ガイド核酸構造は、配列番号201～203を含む。いくつかの場合、ガイド核酸構造は、配列番号613を含む。いくつかの場合、ガイド核酸構造は、配列番号614を含む。いくつかの場合、ガイド核酸構造は、配列番号615を含む。いくつかの場合、ガイド核酸構造は、配列番号616を含む。

30

【0131】

一態様では、本開示は、操作されたヌクレアーゼ系を提供する。操作されたヌクレアーゼ系は、(a)操作されたガイドリボ核酸構造を含み得る。操作されたガイドリボ核酸構造は、ガイドリボ核酸配列を含み得る。ガイドリボ核酸配列は、標的デオキシリボ核酸配列にハイブリダイズするように構成され得る。操作されたガイドリボ核酸構造は、tracrリボ核酸配列を含み得る。tracrリボ核酸配列は、エンドヌクレアーゼに結合するように構成され得る。いくつかの実施形態では、tracrリボ核酸配列は、配列番号199～200、460～461、若しくは669～673のうちのいずれか1つから少なくとも50%、少なくとも60%、少なくとも70%、少なくとも80%個の連続するヌクレオチドと少なくとも50%、少なくとも55%、少なくとも50%、少なくとも55%、少なくとも60%、少なくとも65%、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、若しくは少なくとも99%の配列同一性を有する配列を含むか、又は配列番号201～203、613～616、677～686、1003～10

40

50

22、若しくは1231～1259のうちのいずれか1つの非可変ヌクレオチドの少なくとも15、少なくとも20、少なくとも25、少なくとも30、少なくとも35、少なくとも40、少なくとも45、少なくとも50、少なくとも60、少なくとも70、少なくとも80個の連続するヌクレオチドと少なくとも50%、少なくとも55%、少なくとも50%、少なくとも55%、少なくとも60%、少なくとも65%、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、若しくは少なくとも99%の配列同一性を有する配列を含む。

【0132】

いくつかの場合、操作されたヌクレアーゼ系は、エンドヌクレアーゼを含み得る。エンドヌクレアーゼは、クラス2エンドヌクレアーゼであり得る。エンドヌクレアーゼは、Casエンドヌクレアーゼであり得る。エンドヌクレアーゼは、クラス2、II型Casエンドヌクレアーゼであり得る。

【0133】

いくつかの場合、エンドヌクレアーゼは、特定の分子量範囲を有する。いくつかの実形態では、エンドヌクレアーゼは、約120kDa以下、約110kDa以下、約105kDa以下、約100kDa以下、約95kDa以下、約90kDa以下、約85kDa以下、約80kDa以下、約75kDa以下、約70kDa以下、約65kDa以下、約60kDa以下、約55kDa以下、約50kDa以下、約45kDa以下、約40kDa以下、約35kDa以下、約30kDa以下、約25kDa以下、約20kDa以下、約15kDa以下、又は約10kDa以下の分子量を有する。いくつかの場合、操作されたガイドリボ核酸構造が、少なくとも2つのリボ核酸ポリヌクレオチドを含む。いくつかの場合、エンドヌクレアーゼは、特定の数の残基を含む。エンドヌクレアーゼは、約1,100個以下の残基、約1,000個以下の残基、約950個以下の残基、約900個以下の残基、約850個以下の残基、約800個以下の残基、約750個以下の残基、約700個以下の残基、約650個以下の残基、約600個以下の残基、約550個以下の残基、約500個以下の残基、約450個以下の残基、約400個以下の残基、又は約350個以下の残基を含み得る。エンドヌクレアーゼは、約700～約1,100個の残基を含み得る。エンドヌクレアーゼは、約400～約600個の残基を含み得る。いくつかの場合、操作されたガイドリボ核酸構造が、単一のリボ核酸ポリヌクレオチドを含む。単一のリボ核酸ポリヌクレオチドは、ガイドリボ核酸配列及びtracrリボ核酸配列を含み得る。

【0134】

いくつかの場合、ガイドリボ核酸配列は、原核生物、細菌、古細菌、真核生物、真菌、植物、哺乳類、又はヒトゲノム配列に相補的である。いくつかの場合、ガイドリボ核酸配列は、原核生物ゲノム配列に相補的である。いくつかの場合、ガイドリボ核酸配列は、細菌ゲノム配列に相補的である。いくつかの場合、ガイドリボ核酸配列は、古細菌ゲノム配列に相補的である。いくつかの場合、ガイドリボ核酸配列は、真核生物ゲノム配列に相補的である。いくつかの場合、ガイドリボ核酸配列は、真菌ゲノム配列に相補的である。いくつかの場合、ガイドリボ核酸配列は、植物ゲノム配列に相補的である。いくつかの場合、ガイドリボ核酸配列は、哺乳類ゲノム配列に相補的である。いくつかの場合、ガイドリボ核酸配列は、ヒトゲノム配列に相補的である。

【0135】

いくつかの場合、ガイドリボ核酸標的化配列又はスペーサーは、10～30ヌクレオチド長、又は12～28ヌクレオチド長、又は15～24ヌクレオチド長である。いくつかの場合、エンドヌクレアーゼは、エンドヌクレアーゼのN末端又はC末端の近位に1つ以上の核局在化配列(NLS)を含む。いくつかの場合、NLSは、配列番号205～220から選択される配列を含む。

【0136】

【表 1】

表 1：本開示による *C a s* エフェクターとともに使用され得る例示的な *N L S* 配列。

供給源	NLS アミノ酸配列	配列番号
SV40 NLS	PKKKRKV	205
ヌクレオプラスミン 双節型	KRPAATKKAGQAKKKK	206
c-myc	PAAKRVKLD	207
c-myc	RQRRNELKRSP	208
hnRNPA1 M9	NQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQ GGY	209
インポーチン-アルファ IBB ドメイン	RMRIZFKNKGKDTAELRRRRVEVSVELRKAKKDEQI LKRRNV	210
筋腫 T タンパク質	VSRKRPRP	211
筋腫 T タンパク質	PPKKARED	212
p53	PQPKKKPL	213
マウス c-abl IV	SALIKKKKKMAP	214
インフルエンザウイルス NS1	DRLRR	215
インフルエンザウイルス NS1	PKQKKRK	216
肝炎ウイルスデルタ 抗原	RKLKKKIKKL	217
マウス Mx1 タンパク 質	REKKKFLKRR	218
ヒトポリ(ADP-リボース) ポリメラーゼ	KRKGDEVDGVDEVAKKKSKK	219
ステロイドホルモン 受容体グルココルチ コイド	RKCLQAGMNLEARKTKK	220

【0137】

1つ以上の保存的アミノ酸置換を有する本明細書に説明される酵素のうちのいずれかのバリエーションが、本開示に含まれる。こうした保存的置換は、ポリペプチドの三次元構造又は機能を破壊することなく、ポリペプチドのアミノ酸配列においてなされ得る。保存的置換は、アミノ酸を、互いに同様の疎水性、極性、及び R 鎖長で置換することによって達成することができる。加えて、又は代替的に、異なる種由来の相同なタンパク質の整列された配列を比較することによって、保存的置換は、コードされたタンパク質の基本的な機能を変化させることなく、種間で変異したアミノ酸残基（例えば、非保存残基）を見つけることによって特定され得る。このような保存的置換されたバリエーションは、本明細書に説明されるエンドヌクレアーゼタンパク質配列のうちのいずれか1つに対して少なくとも約20%、少なくとも約25%、少なくとも約30%、少なくとも約35%、少なくとも約40%、少なくとも約45%、少なくとも約50%、少なくとも約55%、少なくとも約60%、少なくとも約65%、少なくとも約70%、少なくとも約75%、少なくとも約80%、少なくとも約85%、少なくとも約86%、少なくとも約87%、少なくとも約88%、少なくとも約89%、少なくとも約90%、少なくとも約91%、少なくとも約92%、少なくとも約93%、少なくとも約94%、少なくとも約95%、少なくとも約96%、少なくとも約97%、少なくとも約98%、少なくとも約99%の同一性を有するバリエーションを含んでもよい。いくつかの実施形態では、このような保存的に置換されたバ

リアントは、機能的バリエーションである。こうした機能的バリエーションは、エンドヌクレアーゼの1つ以上の重要な活性部位残基又はガイドRNA結合残基の活性が破壊されないように、置換を有する配列を包含することができる。いくつかの実施形態では、本明細書に記載されるタンパク質のうちのいずれかの機能的バリエーションは、図4でコールアウトされる保存された残基又は機能的残基のうちの少なくとも1つの置換を欠いている。いくつかの実施形態では、本明細書に説明されるタンパク質のうちのいずれかの機能的バリエーションは、図4でコールアウトされる保存された残基又は機能的残基の全ての置換を欠いている。また、本開示により、本明細書に説明されるヌクレアーゼのうちのいずれかの改変された活性バリエーションも提供される。こうした改変された活性バリエーションは、本明細書（例えば、図4）で特定されるか、又はRuvCドメインについて一般的に説明される1つ以上の触媒残基における不活性化変異を含み得る。そのような改変された活性バリエーションは、RuvCI、RuvCII、又はRuvCIIIドメインの触媒残基における変化スイッチ変異を含み得る。

10

【0138】

機能的に類似したアミノ酸を提供する保存的置換表は、様々な参考文献から入手可能である（例えば、Creighton, Proteins: Structures and Molecular Properties (W H Freeman & Co.; 2nd edition (December 1993) を参照のこと）。以下の8個の群はそれぞれ、互いに保存的置換であるアミノ酸を含有する。

20

- 1) アラニン (A)、グリシン (G)、
- 2) アスパラギン酸 (D)、グルタミン酸 (E)、
- 3) アスパラギン (N)、グルタミン (Q)、
- 4) アルギニン (R)、リシン (K)、
- 5) イソロイシン (I)、ロイシン (L)、メチオニン (M)、バリン (V)、
- 6) フェニルアラニン (F)、チロシン (Y)、トリプトファン (W)、
- 7) セリン (S)、スレオニン (T)、及び
- 8) システイン (C)、メチオニン (M)

【0139】

特定のドメインと配列同一性を有する本明細書に説明されるエンドヌクレアーゼのうちのいずれかのバリエーションが、本開示に含まれる。ドメインは、アルギニンリッチドメイン（例えば、PF14239 相同性を有するドメイン）、REC（認識）ドメイン、BH（ブリッジヘリックス）ドメイン、WED（ウェッジ）ドメイン、PI（PAM相互作用）ドメイン、PF14239 相同性ドメイン、又は本明細書に説明される任意の他のドメインであり得る。いくつかの実施形態では、これらのドメインのうちの1つ以上を包含する残基は、以下のタンパク質のうちの1つに対するアラインメントによってタンパク質中で特定され（例えば、以下のタンパク質及び関心対象のタンパク質のうちの1つが最適に整列されるとき）、残基境界、例えば、ドメインが説明される。

30

【0140】

【表2】

表2：本明細書に説明されるエンドヌクレアーゼの例示的なドメイン境界

40

	RuvC-I	BH	REC	PF14239 相同性を有するドメイン	RuvC-II	HNH	RuvC-III	WED 及び PI
MG34-1 エフェクター	1-41	42-76	77-281	4-65; 123-339	282-323	324-459	460-551	552-747
MG34-9 エフェクター	1-41	42-76	77-280	4-65; 123-338	281-322	323-490	491-582	583-778

【0141】

いくつかの場合、操作されたヌクレアーゼ系は、一本鎖DNA修復鋳型を更に含む。い

50

くつかの場合、操作されたヌクレアーゼ系は、二本鎖DNA修復鑄型を更に含む。いくつかの場合、一本鎖又は二本鎖DNA修復鑄型は、標的デオキシリボ核酸配列の5'側に少なくとも20個のヌクレオチドの配列を含む第1の相同アームを5'から3'の方向に含む。いくつかの場合、一本鎖又は二本鎖DNA修復鑄型は、5'から3'の方向に、少なくとも10個のヌクレオチドの合成DNA配列を含む。いくつかの場合、一本鎖又は二本鎖DNA修復鑄型は、標的配列の3'側に少なくとも20個のヌクレオチドの配列を含む第2の相同アームを5'から3'の方向に含む。いくつかの場合、一本鎖又は二本鎖のDNA修復鑄型は、5'から3'の方向に、標的デオキシリボ核酸配列に対して5'側の、少なくとも20個のヌクレオチドの配列を含む第1の相同アーム、少なくとも10個のヌクレオチドの合成DNA配列、又は標的配列に対して3'側の、少なくとも20個のヌクレオチドの配列を含む第2の相同アームを含む。

10

【0142】

いくつかの場合では、第1の相同アームは、少なくとも10、少なくとも20、少なくとも30、少なくとも40、少なくとも50、少なくとも60、少なくとも70、少なくとも80、少なくとも90、少なくとも100、少なくとも110、少なくとも120、少なくとも130、少なくとも140、少なくとも150、少なくとも175、少なくとも200、少なくとも250、少なくとも300、少なくとも400、少なくとも500、少なくとも750、又は少なくとも1000個のヌクレオチドの配列を含む。いくつかの場合、操作されたヌクレアーゼ系は、 Mg^{2+} の供給源を更に含む。いくつかの場合、エンドヌクレアーゼ及びtracrリボ核酸配列は、別個の細菌種に由来する。いくつかの場合、エンドヌクレアーゼ及びトラクトリボ核酸配列は、同じ門内の別個の細菌種に由来する。

20

【0143】

いくつかの場合、エンドヌクレアーゼは、配列番号1~24又は462~488のうちのいずれか1つと少なくとも50%、少なくとも55%、少なくとも50%、少なくとも55%、少なくとも60%、少なくとも65%、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、又は少なくとも99%の配列同一性を有する配列を含む。いくつかの場合、ガイドRNA構造は、ヘアピンを含むと予測されるRNA配列を含む。いくつかの場合、ヘアピンは、ステム及びループを含む。いくつかの場合、ステムは、少なくとも12対、少なくとも14対、少なくとも16対、若しくは少なくとも18対、又はリボヌクレオチドを含む。

30

【0144】

いくつかの場合、ガイドRNA構造は、第2のステム及び第2のループを更に含む。いくつかの場合、第2のステムは、少なくとも5対、少なくとも6対、少なくとも7対、少なくとも8対、少なくとも9対、又は少なくとも10対のリボヌクレオチドを含む。いくつかの場合、ガイドRNA構造は、RNA構造を更に含み、このRNA構造は、少なくとも2つのヘアピンを含む。いくつかの場合、エンドヌクレアーゼは、配列番号1と少なくとも50%、少なくとも55%、少なくとも50%、少なくとも55%、少なくとも60%、少なくとも65%、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、又は少なくとも99%の配列同一性を有する配列を含み、ガイドRNA構造は、少なくとも4つのヘアピンを含むと予測されるRNA配列を含む。いくつかの場合、これらの4つのヘアピンの各々は、ステム及びループを含む。

40

【0145】

いくつかの場合、操作されたヌクレアーゼ系は、配列番号1と少なくとも50%、少なくとも55%、少なくとも50%、少なくとも55%、少なくとも60%、少なくとも65%、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少

50

なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、又は少なくとも99%同一の配列を含む。いくつかの場合、操作されたヌクレアーゼ系は、配列番号199又は配列番号201の非可変ヌクレオチドのうちの少なくとも1つと少なくとも50%、少なくとも55%、少なくとも50%、少なくとも55%、少なくとも60%、少なくとも65%、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、又は少なくとも99%同一の配列を含むガイドRNA構造を含む。

【0146】

いくつかの場合、操作されたヌクレアーゼ系は、配列番号1~24又は462~488のうちの少なくとも1つと少なくとも50%、少なくとも55%、少なくとも50%、少なくとも55%、少なくとも60%、少なくとも65%、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、又は少なくとも99%同一の配列を含む。いくつかの場合、操作されたヌクレアーゼ系は、配列番号199~200、460~461、若しくは669~673、又は配列番号201~203、613~616、677~686、1003~1022、若しくは1231~1259のうちのいずれか1つの非可変ヌクレオチドのうちのいずれか1つと少なくとも50%、少なくとも55%、少なくとも50%、少なくとも55%、少なくとも60%、少なくとも65%、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、又は少なくとも99%同一の配列を含む。

【0147】

いくつかの場合、配列同一性は、BLASTP、CLUSTALW、MUSCLE、MAFFT、又はSmith-Waterman相同性検索アルゴリズムパラメータを用いたCLUSTALWによって決定される。いくつかの場合、配列同一性は、3のワード長(W)、10の期待値(E)のパラメータ、及び11の存在、1の延長でギャップコストを設定しているBLOSUM62スコアリングマトリックスを使用し、かつ条件付き組成スコアマトリックス調整を使用している、当該BLASTP相同性検索アルゴリズムによって決定される。

【0148】

いくつかの場合、エンドヌクレアーゼは、Cas9エンドヌクレアーゼ、Cas14エンドヌクレアーゼ、Cas12aエンドヌクレアーゼ、Cas12bエンドヌクレアーゼ、Cas12cエンドヌクレアーゼ、Cas12dエンドヌクレアーゼ、Cas12eエンドヌクレアーゼ、Cas13aエンドヌクレアーゼ、Cas13bエンドヌクレアーゼ、Cas13cエンドヌクレアーゼ、又はCas13dエンドヌクレアーゼではない。いくつかの場合、エンドヌクレアーゼは、Cas9エンドヌクレアーゼと80%未満の同一性、75%未満の同一性、70%未満の同一性、65%未満の同一性、60%未満の同一性、55%未満の同一性、又は50%未満の同一性を有する。

【0149】

一態様では、本開示は、(a) DNA標的化セグメントを含む操作されたガイドRNAを提供する。いくつかの場合、DNA標的化セグメントは、標的DNA分子中の標的配列と相補的であるヌクレオチド配列を含む。いくつかの場合、操作された単一ガイドリボ核酸ポリヌクレオチドは、タンパク質結合セグメントを含む。タンパク質結合セグメントは、ハイブリダイズして二本鎖RNA(dsRNA)二重鎖を形成するヌクレオチドの2つの相補的なストレッチを含む。いくつかの場合、ヌクレオチドの2つの相補的なストレッチが、介在するヌクレオチドと互いに共有結合的に連結されている。いくつかの場合、操

10

20

30

40

50

作されたガイドリボ核酸ポリヌクレオチドは、配列番号1～198、221～459、463～612、617～668、674～675、975～1002、1260～1321のうちのいずれか1つ、又はそれらのバリエーションと少なくとも50%、少なくとも55%、少なくとも50%、少なくとも55%、少なくとも60%、少なくとも65%、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、又は少なくとも99%の配列同一性を有するバリエーションを含むエンドヌクレアーゼと複合体を形成するように構成されている。

【0150】

いくつかの場合、DNA標的化セグメントは、ヌクレオチドの2つの相補的なストレッチの両方の5'側に位置付けられている。いくつかの場合、タンパク質結合セグメント操作されたヌクレアーゼ系は、配列番号199～200、460～461、669～673、又は配列番号201～203、613～616、677～686、1003～1022、若しくは1231～1259のうちのいずれか1つの非可変ヌクレオチドのうちのいずれか1つと少なくとも50%、少なくとも55%、少なくとも50%、少なくとも55%、少なくとも60%、少なくとも65%、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、又は少なくとも99%同一の配列を含む。いくつかの場合では、デオキシリボ核酸ポリヌクレオチドは、本明細書に説明される操作されたガイドリボ核酸ポリヌクレオチドをコードする。

【0151】

一態様では、本開示は、操作された核酸配列を含む核酸を提供する。いくつかの場合では、操作された核酸配列は、生物における発現のために最適化される。いくつかの場合では、核酸は、エンドヌクレアーゼをコードする。エンドヌクレアーゼは、Casエンドヌクレアーゼであり得る。エンドヌクレアーゼは、クラス2エンドヌクレアーゼであり得る。エンドヌクレアーゼは、クラス2、II型Casエンドヌクレアーゼであり得る。いくつかの場合、エンドヌクレアーゼは、RuvCドメイン及びHNHドメインを含む。いくつかの場合、エンドヌクレアーゼは、未培養微生物に由来する。いくつかの場合、エンドヌクレアーゼは、特定の分子量範囲を有する。いくつかの実施形態では、エンドヌクレアーゼは、約120kDa以下、約110kDa以下、約105kDa以下、約100kDa以下、約95kDa以下、約90kDa以下、約85kDa以下、約80kDa以下、約75kDa以下、約70kDa以下、約65kDa以下、約60kDa以下、約55kDa以下、約50kDa以下、約45kDa以下、約40kDa以下、約35kDa以下、約30kDa以下、約25kDa以下、約20kDa以下、約15kDa以下、又は約10kDa以下の分子量を有する。いくつかの場合、操作されたガイドリボ核酸構造が、少なくとも2つのリボ核酸ポリヌクレオチドを含む。いくつかの場合、エンドヌクレアーゼは、特定の数の残基を含む。エンドヌクレアーゼは、約1,100個以下の残基、約1,000個以下の残基、約950個以下の残基、約900個以下の残基、約850個以下の残基、約800個以下の残基、約750個以下の残基、約700個以下の残基、約650個以下の残基、約600個以下の残基、約550個以下の残基、約500個以下の残基、約450個以下の残基、約400個以下の残基、又は約350個以下の残基を含み得る。エンドヌクレアーゼは、約700～約1,100個の残基を含み得る。エンドヌクレアーゼは、約400～約600個の残基を含み得る。いくつかの場合、エンドヌクレアーゼは、配列番号1～198、221～459、463～612、617～668、674～675、975～1002、1260～1321、又はそれらのバリエーションと少なくとも50%、少なくとも55%、少なくとも50%、少なくとも55%、少なくとも60%、少なくとも65%、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%

10

20

30

40

50

、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、又は少なくとも99%の配列同一性を有する、配列番号1～198、221～459、463～612、617～668、674～675、975～1002、1260～1321、又はそれらのバリエーションを含む。いくつかの場合、エンドヌクレアーゼは、当該エンドヌクレアーゼのN末端又はC末端の近位に1つ以上の核局在化配列（NLS）をコードする配列を更に含む。いくつかの場合、NLSは、配列番号205～220から選択される配列を含む。

【0152】

いくつかの場合、生物は、原核生物、細菌、真核生物、真菌、植物、哺乳類、齧歯類、又はヒトである。いくつかの場合では、生物は、原核生物である。いくつかの場合では、生物は、細菌である。いくつかの場合では、生物は、真核生物である。いくつかの場合では、生物は、真菌である。いくつかの場合では、生物は、植物である。いくつかの場合では、生物は、哺乳類である。いくつかの場合では、生物は、齧歯類である。いくつかの場合では、生物は、ヒトである。生物が、原核生物又は細菌である場合、生物は、エンドヌクレアーゼが由来する生物とは異なる生物であり得る。いくつかの場合、生物は、未培養微生物ではない。

【0153】

一態様では、本開示は、核酸配列を含むベクターを提供する。いくつかの場合、核酸は、エンドヌクレアーゼをコードする。いくつかの場合、エンドヌクレアーゼは、Casエンドヌクレアーゼである。いくつかの場合、エンドヌクレアーゼは、クラス2エンドヌクレアーゼである。いくつかの場合、エンドヌクレアーゼは、クラス2、II型Caseエンドヌクレアーゼである。エンドヌクレアーゼは、RuvCドメイン及びHNHドメインを含み得る。いくつかの場合、エンドヌクレアーゼは、未培養微生物に由来する。いくつかの場合、エンドヌクレアーゼは、特定の分子量範囲を有する。いくつかの実施形態では、エンドヌクレアーゼは、約120kDa以下、約110kDa以下、約105kDa以下、約100kDa以下、約95kDa以下、約90kDa以下、約85kDa以下、約80kDa以下、約75kDa以下、約70kDa以下、約65kDa以下、約60kDa以下、約55kDa以下、約50kDa以下、約45kDa以下、約40kDa以下、約35kDa以下、約30kDa以下、約25kDa以下、約20kDa以下、約15kDa以下、又は約10kDa以下の分子量を有する。いくつかの場合、操作されたガイドリボ核酸構造が、少なくとも2つのリボ核酸ポリヌクレオチドを含む。いくつかの場合、エンドヌクレアーゼは、特定の数の残基を含む。エンドヌクレアーゼは、約1,100個以下の残基、約1,000個以下の残基、約950個以下の残基、約900個以下の残基、約850個以下の残基、約800個以下の残基、約750個以下の残基、約700個以下の残基、約650個以下の残基、約600個以下の残基、約550個以下の残基、約500個以下の残基、約450個以下の残基、約400個以下の残基、又は約350個以下の残基を含み得る。エンドヌクレアーゼは、約700～約1,100個の残基を含み得る。エンドヌクレアーゼは、約400～約600個の残基を含み得る。

【0154】

いくつかの態様では、本開示は、プロトスペーサー隣接モチーフ（PAM）の5'側の当該標的遺伝子座の近位に二本鎖切断を誘導するように構成された、本明細書に説明されるエンドヌクレアーゼを提供する。エンドヌクレアーゼが、PAMから6～8個のヌクレオチド又はPAMから7個のヌクレオチドの二本鎖切断を誘導し得る。いくつかの態様では、本開示は、プロトスペーサー隣接モチーフ（PAM）の5'側の当該標的遺伝子座の近位に一本鎖切断を誘導するように構成された、本明細書に説明されるエンドヌクレアーゼを提供する。エンドヌクレアーゼが、PAMから6～8個のヌクレオチド又はPAMから7個のヌクレオチドの一本鎖切断を誘導し得る。いくつかの場合、一本鎖切断を誘導するように構成されたエンドヌクレアーゼは、本明細書に説明されるエンドヌクレアーゼの1つ以上の触媒残基における不活性化変異を含む。

【0155】

いくつかの態様では、本開示は、エンドヌクレアーゼ系によって標的化される標的遺伝子座内又はその近位にヌクレオチド塩基の化学修飾を引き起こすように構成された、本明細書に説明されるエンドヌクレアーゼ系を提供する。この場合、ヌクレオチド塩基の化学修飾は、概して、ヌクレオチドの糖又はリン酸部分の修飾ではなく、塩基対合に関与する化学部分の修飾を指す。化学修飾が、アデノシン又はシトシンヌクレオチドの脱アミノ化である。いくつかの場合、化学修飾を引き起こすように構成されたエンドヌクレアーゼ系は、当該エンドヌクレアーゼにフレーム内で結合又は融合された塩基エディタを有するエンドヌクレアーゼを含む。塩基エディタが融合又は結合されるエンドヌクレアーゼは、エンドヌクレアーゼの少なくとも1つの触媒残基（例えば、RuvCドメイン）における不活性化変異を含み得る。塩基エディタは、当該エンドヌクレアーゼのN末端若しくはC末端に融合されるか、又は化学コンジュゲーションを介して連結され得る。塩基エディタは、限定されるものではないが、RNA特異的アデノシンデアミナーゼ1（ADAR1）、RNA特異的アデノシンデアミナーゼ2（ADAR2）、アポリポタンパク質B mRNA編集酵素触媒サブユニット1（APOBEC1）、アポリポタンパク質B mRNA編集酵素触媒サブユニット2（APOBEC2）、アポリポタンパク質B mRNA編集酵素触媒サブユニット3A（APOBEC3A）、アポリポタンパク質B mRNA編集酵素触媒サブユニット3B（APOBEC3B）、アポリポタンパク質B mRNA編集酵素触媒サブユニット3C（APOBEC3C）、アポリポタンパク質B mRNA編集酵素触媒サブユニット3D（APOBEC3D）、アポリポタンパク質B mRNA編集酵素触媒サブユニット3F（APOBEC3F）、アポリポタンパク質B mRNA編集酵素触媒サブユニット3G（APOBEC3G）、アポリポタンパク質B mRNA編集酵素触媒サブユニット3H（APOBEC3H）、若しくはアポリポタンパク質B mRNA編集酵素触媒サブユニット4（APOBEC4）、又はその機能的断片を含む、任意のアデノシン又はシトシンデアミナーゼを含み得る。塩基エディタは、酵母、真核生物、哺乳類、又はヒト塩基エディタを含み得る。

10

20

30

40

50

【0156】

いくつかの態様では、本開示は、エンドヌクレアーゼ系によって標的化される標的遺伝子座内又はその近位にヒストンの化学修飾を引き起こすように構成された、本明細書に説明されるエンドヌクレアーゼ系を提供する。いくつかの場合、ヒストンの化学修飾を引き起こすように構成されたエンドヌクレアーゼ系は、当該エンドヌクレアーゼにフレーム内で結合又は融合されたヒストンエディタを有するエンドヌクレアーゼを含む。ヒストンエディタは、エンドヌクレアーゼにN末端又はC末端で連結又は融合され得る。いくつかの実施形態では、化学修飾は、メチル化、アセチル化、脱メチル化、又は脱アセチル化を含み得る。ヒストンエディタが融合又は結合されるエンドヌクレアーゼは、エンドヌクレアーゼの少なくとも1つの触媒残基（例えば、RuvCドメイン）における不活性化変異を含み得る。ヒストンエディタは、ヒストンメチルトランスフェラーゼ（例えば、ASH1L、DOT1L、EHMT1、EHMT2、EZH1、EZH2、MLL、MLL2、MLL3、MLL4、MLL5、NSD1、PRDM2、SET、SETBP1、SETD1A、SETD1B、SETD2、SETD3、SETD4、SETD5、SETD6、SETD7、SETD8、SETD9、SETDB1、SETDB2、SETMAR、SMYD1、SMYD2、SMYD3、SMYD4、SMYD5、SUV39H1、SUV39H2、SUV420H1、又はSUV420H2）、ヒストンデメチラーゼ（例えば、KDM1、KDM2、KDM3、KDM4、KDM5、又はKDM6ファミリー）、ヒストンアセチルトランスフェラーゼ（例えば、GNAT又はHATファミリーアセチルトランスフェラーゼ）、又はヒストンデアセチラーゼ（例えば、HDAC1、HDAC2、HDAC3、HDAC4、HDAC5、HDAC6、HDAC7、HDAC8、HDAC9、HDAC10、HDAC11、SIRT1、SIRT2、SIRT3、SIRT4、SIRT5、SIRT6、又はSIRT7）を含み得る。ヒストンエディタは、酵母、真核生物、哺乳類、又はヒトヒストンエディタを含み得る。

【0157】

一態様では、本開示は、本明細書に説明される核酸を含むベクターを提供する。いくつかの場合、ベクターは、操作されたガイドリボ核酸構造をコードする核酸を更に含む。操作されたガイドリボ核酸構造は、エンドヌクレアーゼと複合体を形成するように構成され得る。いくつかの場合、操作されたガイドリボ核酸構造は、ガイドリボ核酸配列を含む。いくつかの場合、ガイドリボ核酸配列が、標的デオキシリボ核酸配列にハイブリダイズするように構成されている。いくつかの場合、操作されたガイドリボ核酸構造は、t r a c r リボ核酸配列を含む。いくつかの場合、t r a c r リボ核酸配列は、エンドヌクレアーゼに結合するように構成されている。いくつかの場合、ベクターは、プラスミド、ミニサークル、C E L i D、アデノ随伴ウイルス（A A V）由来ビリオン、又はレンチウイルスである。

10

【0158】

一態様では、本開示は、本明細書に説明されるベクターのうちのいずれかを含む細胞を提供する。

【0159】

一態様では、本開示は、エンドヌクレアーゼを製造する方法を提供する。方法は、本明細書に説明される細胞のうちのいずれかを培養することを含み得る。

【0160】

一態様では、本開示は、二本鎖デオキシリボ核酸ポリヌクレオチドを結合、切断、マーキング、又は修飾するための方法を提供する。方法は、二本鎖デオキシリボ核酸ポリヌクレオチドをエンドヌクレアーゼと接触させることを含み得る。いくつかの場合、エンドヌクレアーゼは、C a s エンドヌクレアーゼである。いくつかの場合、エンドヌクレアーゼは、クラス2 エンドヌクレアーゼである。いくつかの場合、エンドヌクレアーゼは、クラス2、I I 型 C a s エンドヌクレアーゼである。エンドヌクレアーゼは、操作されたガイドリボ核酸構造と複合体化し得る。いくつかの場合、操作されたガイドリボ核酸構造は、エンドヌクレアーゼ及び二本鎖デオキシリボ核酸ポリヌクレオチドに結合するように構成されている。いくつかの場合、二本鎖デオキシリボ核酸ポリヌクレオチドが、プロトスペーサー隣接モチーフ（P A M）を含む。いくつかの場合、エンドヌクレアーゼは、約120 k D a 以下、約110 k D a 以下、約100 k D a 以下、約90 k D a 以下、約80 k D a 以下、約70 k D a 以下、約60 k D a 以下、約50 k D a 以下、約40 k D a 以下、約30 k D a 以下、約20 k D a 以下、又は約10 k D a 以下の分子量を有する。いくつかの場合、エンドヌクレアーゼは、配列番号1～198、221～459、463～612、617～668、674～675、975～1002、1260～1321のうちのいずれか1つ、又はそれらのバリエーションと少なくとも50%、少なくとも55%、少なくとも50%、少なくとも55%、少なくとも60%、少なくとも65%、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、又は少なくとも99%の配列同一性を有するバリエーションを含む。

20

30

【0161】

一態様では、本開示は、二本鎖デオキシリボ核酸ポリヌクレオチドを結合、切断、マーキング、又は修飾するための方法を提供する。方法は、二本鎖デオキシリボ核酸ポリヌクレオチドをエンドヌクレアーゼと接触させることを含み得る。いくつかの場合、エンドヌクレアーゼは、C a s エンドヌクレアーゼである。いくつかの場合、エンドヌクレアーゼは、クラス2 エンドヌクレアーゼである。いくつかの場合、エンドヌクレアーゼは、クラス2、I I 型 C a s エンドヌクレアーゼである。エンドヌクレアーゼは、操作されたガイドリボ核酸構造と複合体化し得る。いくつかの場合、操作されたガイドリボ核酸構造は、エンドヌクレアーゼ及び二本鎖デオキシリボ核酸ポリヌクレオチドに結合するように構成され得る。いくつかの場合、二本鎖デオキシリボ核酸ポリヌクレオチドが、プロトスペーサー隣接モチーフ（P A M）を含む。いくつかの場合、P A M は、N G G である。いくつかの場合、エンドヌクレアーゼは、配列番号1～198、221～459、463～61

40

50

2、617～668、674～675、975～1002、1260～1321のうちのいずれか1つ、又はそれらのバリエーションと少なくとも50%、少なくとも55%、少なくとも50%、少なくとも55%、少なくとも60%、少なくとも65%、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、又は少なくとも99%の配列同一性を有するバリエーションを含む。

【0162】

いくつかの場合、エンドヌクレアーゼは、Cas9エンドヌクレアーゼ、Cas14エンドヌクレアーゼ、Cas12aエンドヌクレアーゼ、Cas12bエンドヌクレアーゼ、Cas12cエンドヌクレアーゼ、Cas12dエンドヌクレアーゼ、Cas12eエンドヌクレアーゼ、Cas13aエンドヌクレアーゼ、Cas13bエンドヌクレアーゼ、Cas13cエンドヌクレアーゼ、又はCas13dエンドヌクレアーゼではない。いくつかの場合、エンドヌクレアーゼは、未培養微生物に由来する。いくつかの場合、二本鎖デオキシリボ核酸ポリヌクレオチドは、原核生物、古細菌、細菌、真核生物、植物、真菌、哺乳類、齧歯類、又はヒト二本鎖デオキシリボ核酸ポリヌクレオチドである。いくつかの場合、二本鎖デオキシリボ核酸ポリヌクレオチドは、エンドヌクレアーゼが由来する種以外の種由来の原核生物、古細菌、又は細菌二本鎖デオキシリボ核酸ポリヌクレオチドである。

10

【0163】

一態様では、本開示は、標的核酸遺伝子座を修飾する方法を提供する。方法は、本明細書に記載される操作されたヌクレアーゼ系を標的核酸遺伝子座に送達することを含み得る。いくつかの場合、エンドヌクレアーゼが、操作されたガイドリボ核酸構造と複合体を形成するように構成されている。いくつかの場合では、複合体は、標的核酸遺伝子座への複合体の結合時に、複合体が標的核酸遺伝子座を修飾するように構成されている。いくつかの場合、標的核酸遺伝子座を修飾することが、標的核酸遺伝子座を結合、ニックング、切断、又はマーキングすることを含む。

20

【0164】

いくつかの場合、標的核酸遺伝子座は、デオキシリボ核酸(DNA)又はリボ核酸(RNA)を含む。いくつかの場合、標的核酸遺伝子座は、ゲノム真核生物DNA、ウイルスDNA、又は細菌DNAを含む。いくつかの場合、標的核酸遺伝子座は、細菌DNAを含む。細菌DNAは、エンドヌクレアーゼが由来した種とは異なる細菌種に由来し得る。いくつかの場合、標的核酸遺伝子座は、インビトロである。いくつかの場合、標的核酸遺伝子座は、細胞内にある。いくつかの場合、別個の核酸分子上でコードされたエンドヌクレアーゼ及び操作されたガイド核酸構造が提供される。いくつかの場合、細胞は、原核細胞、細菌細胞、真核細胞、真菌細胞、植物細胞、動物細胞、哺乳類細胞、齧歯類細胞、霊長類細胞、又はヒト細胞である。いくつかの場合、細胞が、エンドヌクレアーゼが由来する種とは異なる種に由来する。

30

【0165】

いくつかの場合、操作されたヌクレアーゼ系を標的核酸遺伝子座に送達することは、本明細書に説明される核酸又は本明細書に説明されるベクターを送達することを含む。いくつかの場合、操作されたヌクレアーゼ系を標的核酸遺伝子座に送達することが、エンドヌクレアーゼをコードするオープンリーディングフレームを含む核酸を送達することを含む。いくつかの場合、核酸が、エンドヌクレアーゼをコードするオープンリーディングフレームが作動可能に連結されているプロモーターを含む。いくつかの場合、操作されたヌクレアーゼ系を標的核酸遺伝子座に送達することは、当該エンドヌクレアーゼをコードするオープンリーディングフレームを含有するキャッピングされたmRNAを送達することを含む。いくつかの場合、操作されたヌクレアーゼ系を当該標的核酸遺伝子座に送達することは、翻訳されたポリペプチドを送達することを含む。

40

【0166】

50

いくつかの場合、標的核酸遺伝子座に操作されたヌクレアーゼ系を送達することは、リボ核酸（RNA）pol IIIプロモーターに作動可能に連結された操作されたガイドリボ核酸をコードするデオキシリボ核酸（DNA）を送達することを含む。いくつかの場合、エンドヌクレアーゼは、標的遺伝子座で、又は標的遺伝子座の近位で、一本鎖切断又は二本鎖切断を誘導する。

【0167】

本開示の系は、例えば、核酸編集（例えば、遺伝子編集）、核酸分子への結合（例えば、配列特異的結合）などの様々な用途に使用され得る。こうした系は、例えば、対象において疾患を引き起こす可能性のある遺伝的に受け継がれた変異に対処する（例えば、除去又は置換する）ため、遺伝子を細胞におけるその機能を確認するために不活性化するため、疾患を引き起こす遺伝子エレメントを検出する診断ツールとして（例えば、逆転写されたウイルスRNA又は疾患を引き起こす変異をコードする増幅されたDNA配列の切断を介して）、特異的なヌクレオチド配列（例えば、細菌内の抗生物質耐性をコードする配列）を標的とし検出するためのプローブと組み合わせた不活性化酵素として、ウイルスゲノムを標的化することによってウイルスを不活性化する又は宿主細胞に感染できないようにするため、価値ある低分子、高分子、若しくは二次代謝産物を産生するために生物を操作するために遺伝子を加える又は代謝経路を修正するため、進化的選択のための遺伝子駆動エレメントを確立するため、バイオセンサーとして外来低分子及びヌクレオチドによる細胞の摂動を検出するために、使用され得る。

【実施例】

【0168】

実施例1—メタゲノミクスによる新しいCasエフェクターの発見
メタゲノムマイニング

メタゲノム試料を、堆積物、土壌、及び動物から採取した。デオキシリボ核酸（DNA）を、Zymobiomics DNAミニプレップキットを用いて抽出し、Illumina HiSeq（登録商標）2500で配列決定した。試料を、所有者の同意を得て収集した。Qiagen DNeasy PowerSoil Kit又はZymoBIOMICS DNA Miniprep Kitのいずれかを使用して、試料からDNAを抽出した。DNAを、配列決定ライブラリー調製（Illumina TruSeq）及びIllumina HiSeq 4000又はNovaseq上の配列決定のために、UC BerkeleyのVincent J. Coates Genomics Sequencing Laboratory（400～800bpの標的挿入サイズを有する対の150塩基対（bp）リード）に送った。追加的に、公開されている高温、並びに土壌及び海洋のメタゲノム配列決定データをNCBI SRAからダウンロードした。配列決定の読み取りは、BBMap（Bushnell B., sourceforge.net/projects/bbmap/）を使用してトリミングされ、Megahit（<https://paperpile.com/c/QSZG6K/cLMrh>）と組み立てられた。タンパク質配列は、Prodigal（<https://paperpile.com/c/QSZG6K/BJ6ow>）で予測された。文書化されたII型CRISPRヌクレアーゼのHMMプロファイルを構築し、HMMER3（hmmer.org）を使用して全ての予測されるタンパク質に対して検索した。CRISPRアレイを、組み立てたコンティグ上のMinced（<https://github.com/ctSkennerton/minced>）又は<https://paperpile.com/c/QSZG6K/OPC44>）で予測した。分類法は、Kaiju <https://paperpile.com/c/QSZG6K/nMi6k>を有するタンパク質に割り当てられ、コンティグ分類法は、全てのコードされたタンパク質のコンセンサスを見つけてることによって決定された。

【0169】

予測及び参照（例えばSpCas9、SaCas9、及びAsCas9）II型エフェクタータンパク質を、MAFFT（<https://paperpile.com/c/>

Q S Z G 6 K / s V H N H) と整列させ、FastTree2 (<https://paperpile.com/c/QSZG6K/osZNM>) を使用して系統樹を推定した。新規のファミリーを、この研究から回収された配列から構成されるクレードから特定した。ファミリー内から、候補を、それらが実験室分析のための全ての構成要素を含む（すなわち、それらがCRISPRアレイ及び予測されるtracrRNAを用いた十分に組み立てられ、注釈付けされたコンティグで見出された）場合に、選択した。選択された代表的な配列及び参照配列を、MUSCLE (<https://paperpile.com/c/QSZG6K/ITOLA>) を使用して整列させて、触媒残基及びPAM相互作用残基を特定した。

【0170】

このメタゲノムワークフローは、本明細書に説明されるSMART (SMALL ARCHAEA-ASSOCIATED) エンドヌクレアーゼ系の描写を結果的にもたらした。

【0171】

活性残基シグネチャを含有するSMARTエンドヌクレアーゼの発見

メタゲノムデータから組み立てられた数万の高品質のCRISPR Cas系のマイニングは、RuvC及びHNHドメインの両方を含有する新規のエフェクターを発見したが、それらは異常に小さいサイズ (<900 aa) のものであった (図21のA)。これらのエフェクターヌクレアーゼは、参照点として古細菌Cas9エンドヌクレアーゼと低い配列類似性 (<20% アミノ酸同一性) を示した。エフェクタータンパク質配列の系統解析は、SMART系が、サブタイプA、B、又はCからの十分に試験されたII型系に対する分岐群であることを示した (図1A及び図21のB)。

【0172】

これらのコンパクトな「SMART」エフェクター (約400~1000アミノ酸、図2) は、CRISPRアレイに隣接するゲノム内の遺伝子座に現れた。これらの隣接するSMART遺伝子座の一部はまた、tracrRNA及びCRISPR適応遺伝子 (例えば、スパーサー取得に関与する遺伝子) cas1、cas2、又はcas4を同じオペロン内でコードすると予測される配列を含んでいた (図3及び図21のA)。それらのコンパクトなサイズにもかかわらず、SMARTエフェクターは、参照SaCas9配列と整列したときに、6つの推定HNH及びRuvC触媒残基を含有する (図4)。加えて、3D構造予測は、ガイド及び標的結合、並びにPAMの認識に関与する残基を特定し、SMARTエフェクターが活性dsDNAエンドヌクレアーゼであることを示唆する。

【0173】

SMARTエンドヌクレアーゼの複数の群

重要な触媒及び結合残基の場所に基づいて、SMARTヌクレアーゼは、3つのRuvCドメイン、通常はRRxRRモチーフ (例えば、PF14239相同性を有するドメイン) を含有するアルギニンリッチ領域、HNHエンドヌクレアーゼドメイン、及び推定認識ドメインを含む (図5及び図6)。これらのドメインは、参照配列と低い配列類似性を共有する (図7)。加えて、SMARTエフェクター、並びに参照古細菌配列は、RRxRR及び亜鉛結合リボンモチーフ (CX [2-4] C又はCX [2-4] H) を、Cas9ヌクレアーゼよりも有意に頻繁に含有する (図8)。加えて、Cas9エフェクター配列とは異なり、ほとんどのSMARTエフェクターは、PfamドメインPF14239への有意なヒットを含有し、これは、多くの場合、多様なエンドヌクレアーゼと関連付けられている。SMARTエフェクターサイズ、系統学的関係、並びにオペロン及びドメインアーキテクチャの両方における差異に基づいて、これらの系を2つの主要グループ、SMART I及びSMART IIに分類した。これらの群の顕著な特徴が以下の表3に概説され、クラス2、II型A/B/C Cas酵素と比較した差異も例示する。

【0174】

10

20

30

40

50

【表 3】

表 3：本明細書に説明される SMART I 及び II 群酵素の属性

属性	SMART I	SMART II	II 型：A、B、C
Zn 結合残基	はい	はい	いいえ
ブリッジヘリックス	はい	いいえ	はい
PAM 相互作用及び WED ドメイン	はい	いいえ	はい
RRxRR モチーフ	はい	はい	いいえ
REC ドメイン	PF14239 に対する相 同性を有する新規の ドメイン	C 末端における 新規のドメイン	Cas9 REC ドメイン
PF14239 相同性を有する ドメイン	はい	はい	いいえ
単系統クレード	はい	いいえ	はい
TnpB に関連する	はい	はい	はい
オペロンは IS605Tns リ ピートを含有する	いいえ	時々	いいえ
<900aa	時々	はい	いいえ
CRISPR 関連	はい	時々	はい
RuvC 及び HNH を含有 する	はい	はい	はい

【0175】

SMARTヌクレアーゼは、Cas9のようにRuvC及びHNHドメインを含有するが、RuvC-I、ブリッジヘリックス、及び認識ドメインは、あまり整列しない。SMARTヌクレアーゼと参照配列との間の進化的関係を最良に理解するために、文書化及び分類された完全長SMART、参照II型配列の複数の配列アラインメント（例えば、各々、その全体が参照により本明細書に組み込まれる、Burstein, D. et al. New CRISPR-Cas systems from uncultivated microbes. Nature 2017, 542, 237-241、及びGasunas, G. et al. A catalogue of biochemically diverse CRISPR-Cas9 orthologs. Nat Commun 2020, 11, 5512を参照されたい）、並びに>10,300個の最近報告されたCas9相同体及びIsCB配列（例えば、その全体が参照により本明細書に組み込まれる、Altae-Tran, H. et al. The widespread IS200/605 transposon family encodes diverse programmable RNA-guided endonucleases. Science 2021, 374, 57-65を参照されたい）を生成した。RuvC-II/HNH/RuvC-IIIDメインを包含するトリミングされた十分に整列

した領域を保持した。この最終アラインメントから推定された系統解析は、II-A、II-B、及びII-Cに現在分類される文書化されたCas9エフェクターからクラスター化するエフェクターの分岐クレードを示した（図21のE）。分類されたII型エフェクターに系統的に近いことが見出された2つのSMARTクレードは、CRISPRアレイに隣接してコードされる可能性が高かった（図21のB、図21のC、及び図21のE）。II-C2型エフェクターを有するSMARTヌクレアーゼクラスターのMG33ファミリーは、このクレードを大幅に拡張する（図21のE及び図21のF、モープ分岐）。このファミリーは、SMART酵素の最大である900~1050aaの代表を含有し、それらの長さ分布は、最小の分類されたII-C型酵素と重複する。より離れたSMARTクレード（図21のE及び図21のF、青緑色、緑色、及び黄色の分岐）は、最近II-D型として分類された「早期Cas9」配列を含有する（図21のE及び図21のF、淡灰色の分岐）。これらのCRISPR系は、一般的に、集合的にSMARTと称され得る。

10

【0176】

SMART Iエンドヌクレアーゼ

SMART Iエフェクターのサイズは、約600アミノ酸~1,050アミノ酸の範囲である。それらのゲノムコンテキストにおける一般的な特徴は、適応モジュール遺伝子（例えば、スパーサー取得に関与する遺伝子）及びCRISPRアレイ近くの予測されるtracrRNAであり、その構成は、II型及びV型CRISPR系に類似していた（図3A、図3B、及び図3C）。SMART IエフェクターにおけるRRXRモチーフ含有領域は、固有であるが、Cas9ヌクレアーゼにおけるアルギニンリッチブリッジヘリックスと同様の機能的役割を果たし得る。SaCas9結晶構造に対してモデル化されたとき、SMART Iエフェクターの予測される3D構造は、認識ローブ（多くの場合、PfamdメインPF14239を含有する）及びRuvc-IIDメイン内の非整列領域を示した（図5）。結果は、これらのドメインが他のII型エフェクターと比較して異なる起源を有することを示した。II型エフェクター系統樹におけるそれらの分岐配置、及び文書化されたII型エフェクターとのそれらの低い配列類似性と併せて得られた（図1A及び図21のB）、これらの結果は、SMART IエンドヌクレアーゼがII型CRISPR系の新しい群に属することを示す。CRISPR系の承認された分類に続いて、これらのSMART I系を、II-D型として分類した。

20

30

【0177】

推定単一ガイドRNA（sgRNA）を、SMART I MG34-1系の環境RNA発現データを使用して操作した。加えて、SMART Iリピート及びtracrRNA予測から設計された複数のsgRNAを、PAM濃縮アッセイで、インビトロで試験した。SMART I酵素の場合、この段階で末端修復及び平滑末端ライゲーションを使用してPAM配列の最適な特定を実施し、これらの酵素が互い違いの二本鎖DNA切断を生じ得ることを示唆した。アッセイは、複数のsgRNA設計を有するMG34-1（配列番号2）、MG34-9（配列番号9）、及びMG34-16（配列番号17）のdsDNA切断を確認した（図7、配列番号612~615の使用を図示する）。MG34-1は、標的認識及び切断に対するNGGN PAMに対する選好を実証したが（図8A）、MG102ファミリーのメンバーは、標的認識及び切断のための3'NRC PAMを認識する（図21のC）。切断部位の分析は、7位で優先的な切断を示した（図8B及び図22A）。これらの結果は、他のII型酵素からの切断機構と比較して新規の生化学的機構を示唆し、これは、PAMから2~3位で優先的に切断し、SMART I CRISPR系の新しい分類を支持する。

40

【0178】

いくつかのSMART I系の環境発現データは、CRISPRアレイ及び予測されるtracrRNAをコードする遺伝子間領域のin situ転写を確認した（図3B及び図3C）。加えて、活性CRISPR標的化の事例を、同じ又は関連するメタゲノムから組み立てられた他のゲノム配列と一致するスパーサー配列を検索することによって評価

50

した。これらの線に沿って、SMART I CRISPRアレイにコードされるスペーサーのうちの1つによって標的化されるファージゲノム（図3C及び図3D）を特定した。標的配列に隣接する領域の分析は、GGモチーフを含有する3' PAM配列を示唆する（図3D）。これらの結果は、SMART I CRISPR系が、それらの天然環境でファージ防御に関与するRNAガイドエフェクターとして活性であり、標的化されたDNA又はRNAを切断又は分解するヌクレアーゼとして機能する可能性が高いことを示す。

【0179】

SMART Iエフェクターは、活性なRNAガイドdsDNA CRISPRエンドヌクレアーゼ

推定単一ガイドRNA (sgRNA) を、SMART I MG34-1及びMG34-16系の環境RNA発現データを使用して操作した（図3B及び図3C、並びに図9）。加えて、SMART Iリピート及びtracrRNA予測から設計された複数のsgRNAを、PAM濃縮アッセイで、インビトロで試験した（図10）。アッセイは、複数のsgRNA設計でMG34-1、MG34-9、及びMG34-16のプログラム可能なdsDNA切断を確認した（図10）。MG34-1及びMG34-9は、標的認識及び切断のためにNGGN PAMを必要とする（図11A及び図11C）。切断部位の分析は、7位で優先的な切断を示した（図11B及び図11C）。これらの結果は、Cas9酵素と比較して新規の生化学的切断機構を示唆し、これは、PAMから3位で優先的に切断し、SMART I CRISPR系の新しい分類のための更なる支持を提供する。

【0180】

末端修復手順を伴わないPAM濃縮アッセイは、SMART Iヌクレアーゼに対する活性を示さなかった。PAM濃縮プロトコルにおけるライゲーション前に平滑末端断片を作成するための末端修復の要件は、これらの酵素が互い違いの二本鎖DNA切断を作成することを示す。MG34-1ヌクレアーゼの切断産物の配列決定によって、互い違いの二本鎖切断を確認した（図22A）。これらの結果は、ほとんどの文書化されたII型酵素からの機構と比較して新規の生化学的機構を示唆し、これは、PAMから2～3位で優先的に切断する。精製されたタンパク質を用いたインビトロ切断アッセイは、MG34-1が標的ガイド18bpの長さでの標的化されたDNA切断でより効率的であることを示し、時系列切断アッセイは、MG34-1が、同一のガイドで試験されたとき、参照SpCas9と比較してより遅い速度で切断されることを示す（図22B及び図22C）。

【0181】

E. coliで実施された実験は、系が細胞におけるヌクレアーゼとして機能するために必要な活性を有することを示した。MG34-1及びMG34-9 sgRNAを発現するE. coli株を、sgRNAの標的を含有するカナマイシン耐性プラスミドで形質転換した。抗生物質の存在下で、抗生物質耐性プラスミドの標的化及び切断の成功は、成長欠陥を結果的にもたらすことになる。アッセイは、sgRNAの標的を含有しなかったカナマイシン耐性プラスミドを用いて実施された対照実験と比較して、およそ2倍～10倍の成長抑制を示した（図12及び図22D）。

【0182】

SMART IIエンドヌクレアーゼ

SMART IIエフェクターは、SMART Iエフェクターと比較して小さい（約400アミノ酸～600アミノ酸）サイズ分布を有する。それらのゲノムコンテキストは、異常な反復領域又はCRISPRアレイを示唆した。非CRISPR反復領域は、約10～30bpを超えるサイズの範囲の直接反復を含有する。いくつかの場合、これらは、複数の別個の反復ユニットを含む。時には、一般的なCRISPR識別アルゴリズムは、これらの領域をCRISPR系としてフラグ付けすることになるが、より詳細な検査は、スペーサー配列として識別された領域がアレイ内で反復されることを明らかにする。アレイは、エフェクターにすぐ隣接していないが、それらは、同じゲノム領域にある（図3A、MG35-236、及び図13A、例えば、エフェクター遺伝子から>20kb）。SMART II系オペロンは、概して、適応モジュール遺伝子（例えば、スペーサー取得

に關与する遺伝子)を欠いていた。

【0183】

構造予測は、多くの場合、クラス2、II型Casエフェクターに見られる全ての6つのRuvc及びHNHヌクレアーゼ触媒残基(図6)に加えて、ガイドRNA結合、標的切断、並びにPAMの認識及びPAMとの相互作用に關与するCas酵素の特徴的な残基を特定した。加えて、SMART IIEフェクターは、標的核酸モチーフの認識及び結合に關与する可能性がある、複数のRRXR及び亜鉛結合リボンモチーフ(CX_[2-4]C又はCX_[2-4]H)を含有した。重要な残基の場所に基づいて、SMART IIEヌクレアーゼの予測されるドメイン構造は、RRXRモチーフを含有するアルギニンリッチ領域(例えば、PF14239相同性を有するドメイン)、HNHエンドヌクレアーゼドメイン、未知のドメイン、及び認識ドメイン(REC)の3つのRuvcサブドメインを含んでいた(図6)。SMART IIEフェクターのドメインアーキテクチャは、II型Cas9ヌクレアーゼの文書化されたドメイン構造とは異なっていた(図6及び図14)。

10

【0184】

いくつかのSMART IIE系の環境トランスクリプトームデータは、天然環境中のCRISPRアレイ及び他の反復領域の*in situ*発現を確認した(図13A)。いくつかのSMART IIEフェクターの5'非翻訳領域(UTR)の転写もまた、環境発現データから観察され(図13B及び図16)、この領域が、ヌクレアーゼ活性又はSMART系の調節のいずれかに重要であり得ることを示唆している。

20

【0185】

SMART IIEフェクタータンパク質、反復領域、及び関連する遺伝子間領域を用いて実施された予備インビトロ実験は、これらの酵素が、おそらくプログラム可能な様式でdsDNAを切断する能力を有することを示す(図15及び図17)。結果は、SMART IIEヌクレアーゼ活性がRNA又はDNAガイドであり得ることを示唆しており、これは、CRISPRアレイなどの反復領域を使用するか、又はTIR若しくは5'UTRなどの遺伝子座内にコードされる特徴の認識を介して必要とされ得る。SMART IIEフェクターの5'UTRは、インビトロ転写アッセイで能動的に転写され、高い二次構造を呈する(図18)。SMART IIEフェクターの開始コドンのすぐ上流の領域の多重配列アラインメントは、保存ブロックを実証し(図19)、SMART IIEフェクターと関連付けられた5'UTRが、切断活性のためにDNAを標的化するためにエフェクターのRNAガイドをコードすることを示唆している。

30

【0186】

最近、短いCas9ホモログが、エフェクターの5'UTR領域にコードされたガイドRNAを使用してプログラム可能なdsDNAヌクレアーゼであると報告された(Alt *et al.*, *Science* 2021)。これらの系では、標的化「スペーサー」がエフェクターの転写された5'UTRの上流で特定され、SMART IIE酵素が、その5'UTRにコードされる予測されるガイドRNAの5'末端に「標的スペーサー」を加えることによって、特定のDNA部位を標的化及び切断するように再プログラムされ得ることを示唆している。SMART IIEフェクターの5'UTR領域にコードされるガイドRNAの5'末端に標的スペーサーを付加することは、様々な標的隣接モチーフ(TAM)を用いて、標的化されたdsDNA切断のためのエフェクターを活性化した(図20)。

40

【0187】

いくつかのSMART IIEフェクターは、トランスポサゼTnpA及びTnpBをコードする推定挿入配列(IS)の隣で観察された(図3A)。ISの末端は、予測されるヘアピン構造を有する末端逆位リピート(TIR)を含有すると特定され、ISが組み込まれている可能性が最も高い標的部位重複も特定された)。加えて、いくつかのSMART IIE遺伝子座は、SMART IIEフェクターに隣接する推定TIRをコードした(例えば、図3)。

50

【0188】

SMART HEAROクレードは、ウイルス関連RNAガイドdsDNAヌクレアーゼを含有する

系統解析は、長さが600aa未満のSMARTヌクレアーゼ（図21のE、薄紫色の分岐）が、文書化されたIscB配列（「挿入配列Cas9様」（例えば、その全体が参照により本明細書に組み込まれる、Kapitonov, V. V., Makarova, K. S. & Koonin, E. V. ISC, a Novel Group of Bacterial and Archaeal DNA Transposons That Encode Cas9 Homologs. J Bacteriol 2016, 198, 797-807を参照されたい。））（図21のE、濃灰色の分岐）と一緒にクラスター化し、2つの主要なクレードを形成する。Kapitonovらは、RuvC及びHNHドメインの存在に基づいてCas9とのIscB相同性を報告し、その後、この同じ酵素群においてPLMPドメインを説明した（例えば、その全体が参照により本明細書に組み込まれる、Altae-Tran, H. et al. The widespread IS200/605 transposon family encodes diverse programmable RNA-guided endonucleases. Science 2021, 374, 57-65）。3D構造予測を使用して、これらのタンパク質が、通常、RRXRモチーフを含有するアルギニンリッチ領域を含有することを示した。アルギニンリッチ領域は、Cas9のブリッジヘリックスに類似していることが示唆されたが、この領域もRuvC-ドメインも、参照3D構造のブリッジヘリックス及びRuvC-ドメインと3D空間で良好に整列することは見出されなかった。このようなIscB/SMART酵素は、PAM相互作用ドメインを欠く。代わりに、Zn結合リボンモチーフを含有するC末端「WED/REC」ドメインが、標的モチーフ認識に関与し得る。タンパク質ドメイン、触媒残基、及び3Dモデルは、Cas9との進化的関係を示唆するが、ほとんどのIscB/SMARTエフェクターは、CRISPR関連ではない（例えば、それらのゲノムコンテキストではCRISPRリピートの近位には見つからない）。IscB/SMART系を含む群は、概して、サイズがコンパクトであり（およそ400~600aa）、細菌及び古細菌ゲノムに広く分布している。これらのエフェクターをコードするゲノム断片の16%以上が、ウイルス又はプロファージ由来の可能性が高いと分類され、これらの系の進化にウイルスが関係していることが見出された。

【0189】

SMART系と関連付けられた非コードRNA（ncRNA）に対する検索は、IscB/SMART 5'非翻訳領域（UTR）の65%が、RFamデータベース（RF02033）からのHNHエンドヌクレアーゼ関連RNA及びORF（HEARO）RNAに対するヒットを含有することを見出した。これらのncRNAは、バイオインフォマティクス分析から高度に構造化されたRNAとして最初に説明された（例えば、その全体が参照により本明細書に組み込まれる、Weinberg, Z., Perreault, J., Meyer, M. M. & Breaker, R. R. Exceptional structured noncoding RNAs revealed by bacterial metagenome analysis. Nature 2009, 462, 656-659を参照されたい）が、それらの関連するHEARO ORFの機能は、報告されていなかった（例えば、その全体が参照により本明細書に組み込まれる、Harris, K. A. & Breaker, R. R. Large Noncoding RNAs in Bacteria. Microbiol Spectr 2018, 6を参照されたい）。推定HEARO HNHエンドヌクレアーゼORFもまた、RuvC及びHNH触媒ドメインも含有し、IscB/SMARTエフェクターと一緒にクラスター化することが確認された。したがって、IscB、小さいSMART、及びHEARO ORFは、非Casエンドヌクレアーゼの大きな群を表す。最近、IscBの5'UTRが、dsDNAヌクレアーゼ活性に必要な単一ガイドRNAをコードしていることが報告され

ており、これは、著者がオメガRNAと呼称する（例えば、その全体が参照により本明細書に組み込まれる、Altae-Tran, H. et al. The widespread IS200/605 transposon family encodes diverse programmable RNA-guided endonucleases. Science 2021, 374, 57-65を参照されたい）。機能に対するガイドRNAの要件を確認するために、本発明者らは、インビトロ転写アッセイによって再現された、IscB/SMART/HEARO系の5'UTRのin situ天然発現を観察した。オメガRNA構造は、HEARO RNAと高い構造類似性を共有する。IscB/SMART/HEARO系（広範な分類学的起源及びアルギニン残基の濃縮）を統合する特徴、及びこれらの酵素に関連付けられたガイドRNAの年代的発見の認識において、本発明者らは、SMART HEAROとしてIscB/SMART/HEARO系の広範な機能分類を提唱する（図21のE）。本発明者らは、Altae-Tran及びKannanらによって説明されるように、インビトロでSMART HEARO切断活性を評価し、それらのHEARO RNAの5'「スペーサー」領域を再プログラミングすることによって、必要な標的化モチーフを特定した（図21のD）（例えば、Altae-Tran, H. et al. The widespread IS200/605 transposon family encodes diverse programmable RNA-guided endonucleases. Science 2021, 374, 57-65）。更に、E. coliにおけるプラスミド干渉アッセイは、SMART HEAROヌクレアーゼが、SpCas9と比較して高度に活性であることを示し（MG35-1に対する>570倍の抑制対SpCas9によって示される約98倍の抑制、図25B）、特異性実験は、プロトスペーサー中のミスマッチに対する耐性が低いことを示す（図25D）。

10

20

30

40

50

【0190】

実施例2-本明細書に説明されるエンドヌクレアーゼのPAM配列特定／確認

推定SMARTエンドヌクレアーゼを、E. coli溶解物ベースの発現系（PURExpress, New England Biolabs）で発現した。この系では、エンドヌクレアーゼをE. coliに対してコドン最適化し、T7プロモーター及びC末端Hisタグを有するベクターにクローニングした。遺伝子を、それぞれ、T7プロモーター配列及びターミネーター配列の上流及び下流150bpのプライマー結合部位でPCR増幅した。このPCR産物を、5nMの濃度でNEB PURExpressに加えて、37℃で2時間発現させて、PAMアッセイ用のエンドヌクレアーゼを産生した。

【0191】

本明細書に説明される各SMART Cas酵素と適合性のある推定sgRNAを、配列決定データから組み立てられたコンティグCRISPR遺伝子座に組み立てられたRNAseqリードから特定し、二次構造を、Geneiousソフトウェアパッケージ（<https://www.geneious.com>）のCRISPRアレイからの反復配列とともにRNAseqデータからのtracr領域について決定し、結果として得られたヘリックスをトリミングし、GAAAテトラループと連結した。複数の長さのリピーターアンチリピーターヘリックストリミング、並びに異なるスペーサー長及び異なるtracr終止点を試験した（図12、配列番号612～615を実証する）。次いで、各sgRNAをアセンブリPCRを介して組み立て、SPRIビーズで精製し、製造業者が推奨する短いRNA転写物（HiScribe T7キット、NEB）のプロトコルに従ってインビトロ転写（IVT）した。RNA転写反応物をMonarch RNAキットで洗浄し、Tapestation（Agilent）を介して純度をチェックした。

【0192】

推定ヌクレアーゼによって切断され得るランダムに生成された潜在的なPAM配列を含むプラスミドを配列決定することによって、PAM配列を決定した。この系では、推定ヌクレアーゼをコードするE. coliコドン最適化ヌクレオチド配列を、T7プロモーターの制御下でPCR断片からインビトロで転写及び翻訳した。T7プロモーターとそれに

続く反復スペーサー反復配列から構成される最小CRISPRアレイを有する第2のPCR断片を、同じ反応で転写した。CRISPRアレイ処理が後に続くTXTL系におけるエンドヌクレアーゼ及び反復スペーサー反復配列の発現の成功は、活性なインビトロCRISPRヌクレアーゼ複合体を提供する。

【0193】

8Nの混合縮重塩基が先行する最小アレイに合致するスペーサー配列（潜在的なPAM配列）を含有する標的プラスミドのライブラリーを、TXTL反応の出力（翻訳されたCas酵素の5倍希釈、5nMの8N PAMプラスミドライブラリー、及びPAMライブラリーを標的化する50nMのsgRNAを有する、10mMのTris pH7.5、100mMのNaCl、及び10mMのMgCl₂）とともにインキュベートした。1〜3時間後、反応を停止し、DNAクリーンアップキットを介してDNAを回収した。アダプター配列を、エンドヌクレアーゼによって切断された活性なPAM配列を有するDNAに平滑末端ライゲーションしたが、切断されなかったDNAは、ライゲーションにはアクセスできなかった。次いで、活性PAM配列を含むDNAセグメントを、ライブラリー及びアダプター配列に特異的なプライマーを用いてPCRにより増幅した。PCR増幅産物をゲル上で分解して、切断事象に対応するアンプリコンを特定した。切断反応の増幅したセグメントも、NGSライブラリーの調製のための鋳型として、又はサンガーシーケンシングのための基質として使用した。開始8Nライブラリーのサブセットであるこの得られたライブラリーを配列決定することにより、CRISPR複合体と適合するPAM活性を有する配列を明らかにした。処理されたRNA構築物を用いたPAM試験については、インビトロ転写RNAをプラスミドライブラリーとともに加え、最小CRISPRアレイ／tracr 鋳型を省略したことを除いて、同じ手順を繰り返した。これらのアッセイでは、以下のスペーサー配列を標的として使用した（5' - CGUGAGCCACCCACGUCGCAAGCCUCGAC - 3'）。

【0194】

PAMアッセイから生の配列決定リードを得た後、リードをPhred品質スコア>20によってフィルタリングした。PAMに隣接する骨格から文書化されたDNA配列を表す24bpを参照として使用して、PAM近位領域を見つけ、隣接する8bpを推定PAMとして特定した。PAMとライゲーションアダプターとの間の距離も、各読み取りに対して測定された。参照配列又はアダプター配列と正確に一致しなかった読み取りは除外された。PAM配列を、最も頻度の高い切断部位±2bpを有するPAMが分析に選択的に含まれるように、切断部位頻度によってフィルタリングした。PAMのフィルタリングされたリストを使用して、Logomakerを使用して配列ロゴを生成した（参照により本明細書に組み込まれる、Tareen A, Kinney JB. Logomaker: beautiful sequence logos in Python. Bioinformatics. 2020; 36(7): 2272-2274）。

【0195】

実施例3 - 予測されるRNAフォールディングのプロトコル

Andronescu 2007の方法を使用して、活性単一RNA配列の予測されるRNAフォールディングを37°で計算する。塩基の色は、その塩基の塩基対合の確率に対応し、赤は高い確率であり、青は低い確率である。

【0196】

実施例4 - インビトロ切断効率

エンドヌクレアーゼは、タンパク質分解酵素欠損大腸菌B株における誘導性T7プロモーターからのHisタグ付き融合タンパク質として発現される。エンドヌクレアーゼを、2つの核局在化シグナル（N末端NLSヌクレオプラスミン双節型及びC末端シミアンウィルス40 T抗原NLS PPKKKRK）、マルトース結合タンパク質（MBP）タグ、タバコエッチングウイルス（TEV）プロテアーゼ切断部位、及び6XHisタグを、NからC末端に以下の順序で融合した。6XHis-MBP-TEV-NLS-遺伝子-NLS-STOP。このタンパク質を、自己誘導培地（Magic Media The

rm o F i s h e r) によってNEB Iq E. coli中のpT a cプロモーター下で発現させ、30℃で増殖させ、16℃で誘導した。

【0197】

H i sタグ付きタンパク質を発現している細胞を超音波処理によって溶解し、H i sタグ付きタンパク質を、AKTA Avant FPLC (GE Lifescience) 上のH i sTrap FFカラム (GE Lifescience) でのNi-NTA親和性クロマトグラフィーによって精製した。溶出液をアクリルアミドゲル (Bio-Rad) 上のSDS-PAGEによって分離し、InstantBlue Ultrafast Coomassie (Sigma-Aldrich) で染色した。純度を、ImageLabソフトウェア (Bio-Rad) を用いたタンパク質バンドの濃度測定法を使用して決定した。精製したエンドヌクレアーゼを、50mM Tris-HCl、300mM NaCl、1mM TCEP、5%グリセロールからなる保存緩衝液、pH7.5に透析し、-80℃で保存した。

10

【0198】

スペーサー配列及びPAM配列 (例えば、実施例2で決定した) を含有する標的DNAを、DNA合成によって構築した。PAMが縮重塩基を有する場合の試験のため、単一の代表的なPAMを選択する。標的DNAは、PAM及び一端から700bpに位置するスペーサーを用いたPCR増幅を介してプラスミドに由来する2200bpの直鎖DNAからなる。切断の成功は、700及び1500bpの断片をもたらす。標的DNA、インビトロ転写単一RNA、及び精製した組換えタンパク質を、過剰なタンパク質及びRNAを含む切断緩衝液 (10mMのトリス、100mMのNaCl、10mMのMgCl₂) 中で組み合わせ、5分~3時間、通常は1時間インキュベートする。反応を、RNAse Aの添加及び60分のインキュベーションを介して停止させる。次いで、反応物を1.2%TAEアガロースゲル上で分解し、切断した標的DNAの画分をImageLabソフトウェアで定量化する。

20

【0199】

実施例5-E. coliにおける活性

E. coliは、二本鎖DNA切断を効率的に修復する能力を欠く。したがって、ゲノムDNAの切断は、致命的な事象であり得る。この現象を利用して、スペーサー/標的配列及びPAM配列がそのゲノムDNAに組み込まれた標的株においてエンドヌクレアーゼ及びガイドRNAを組換え発現させることによって、E. coliでエンドヌクレアーゼ活性が試験される。

30

【0200】

細菌細胞におけるヌクレアーゼ活性の試験のために、BL21 (DE3) 株 (NEB) を、T7-駆動エフェクター及びsgRNA (各プラスミド10ng) を含有するプラスミドで形質転換し、播種し、一晚増殖させた。結果的に得られたコロニーを3回にわたって一晚培養し、次いで、SOB中で継代培養し、OD0.4~0.6まで増殖させた。細胞培養物の0.5OD当量を、標準キットプロトコル (Zymo Mix及びGoキット) に従って化学能力を有するようにし、骨格中にスペーサー及びPAMを有するか、又は有していないかのいずれかで、130ngのカナマイシンプラスミドで形質転換した。ヒートショック後、形質転換物を、SOC中、37℃で1時間回収し、ヌクレアーゼ効率を、誘導培地 (抗生物質及び0.05mM IPTGを含むLB寒天プレート) 上で成長させた5倍希釈系列によって決定した。コロニーを希釈系列から定量して、ヌクレアーゼ駆動プラスミド切断に起因する全体的な抑制を測定した。

40

【0201】

そのようなアッセイの結果が、図12に示される。図12では、パネル(A)は、プラスミド切断を実証するE. coli株のレプリカプレーティングを示し、MG34-1及びsgRNAを発現するE. coliを、sgRNAの標的を含有するカナマイシン耐性プラスミド (+sp) で形質転換した。成長障害 (+sp) 対陰性対照 (標的及びPAMなし (-sp)) を示すプレート四分円は、酵素による標的化及び切断の成功を示す。実

50

験を2回繰り返し、3回実施した。図12では、パネルBは、Aにおけるレプリカプレーティング実験からのコロニー形成単位(cfu)測定値のグラフを示し、標的条件(+sp)対非標的対照(-sp)における成長抑制を示し、プラスミドが切断されたことを実証する。図12では、パネルCは、様々なSMARTヌクレアーゼの標的条件(白色バー)対非標的対照(緑色バー)におけるE.coli成長抑制を示すコロニー形成単位(cfu)測定値(対数スケール)のバープロットを示す。各ヌクレアーゼに対するプラスミド干渉アッセイを、SpCas9陽性対照とともに3回行った。

【0202】

ゲノムDNAに組み込まれたPAM配列(例えば、実施例2にあるように決定される)を有する操作した株を、エンドヌクレアーゼをコードするDNAで形質転換する。次いで、形質転換体を、化学能力を有するようにし、標的配列に特異的(「オンターゲット」)、又は標的に非特異的(「非標的」)かのいずれかの50ngのガイドRNA(例えば、crRNA)で形質転換する。ヒートショック後、形質転換物を、SOC中、37℃で2時間回収する。次いで、ヌクレアーゼ効率を、誘導培地上で成長させた5倍希釈液によって決定する。コロニーを、3回、希釈系列から定量化する。

【0203】

実施例6—哺乳類細胞におけるMG-CRISPR複合体のゲノム切断活性を試験する

哺乳類細胞における標的化及び切断活性を示すために、MG-Casエフェクタータンパク質配列を、2つの哺乳類発現ベクター、(a)1つは、C末端のSV40-NLS及び2A-GFPタグを有し、(b)1つは、GFPタグを有さず、2つのSV40-NLS配列を有し、1つはN末端上にあり、1つはC末端上にある、で試験する。NLS配列は、本明細書に説明されるNLS配列のうちのいずれかを含む。いくつかの場合では、エンドヌクレアーゼをコードするヌクレオチド配列を、哺乳類細胞における発現のためにコドン最適化する。

【0204】

標的化配列を付加した対応するcrRNA配列を、第2の哺乳類発現ベクターにクローニングする。2つのプラスミドをHEK293T細胞に共トランスフェクトする。HEK293T細胞への発現プラスミド及びgRNA標的化プラスミドの共トランスフェクションから72時間後、DNAを抽出し、NGSライブラリーの調製に使用する。NHEJパーセントを標的部位のシーケンシングにおけるインデルを介して測定して、哺乳類細胞における酵素の標的化効率を示す。各タンパク質の活性を試験するために、少なくとも10個の異なる標的部位を選択する。

【0205】

実施例7—本明細書に説明されるMGファミリーの予測される活性

in situ発現及びタンパク質配列分析は、これらの酵素が活性ヌクレアーゼであることを示す。それらは、予測されるエンドヌクレアーゼ関連ドメイン(RRXRR及びHNHエンドヌクレアーゼPfamドメインに合致する、図2、図3A、及び図3B)を含有し、予測されるHNH及びRuvC触媒残基(図2、図3A、及び図3B、長方形)を含有する。更に、リボヌクレアーゼH様タンパク質ファミリーに見られるRRXRRモチーフの存在は、潜在的なRNA標的化又はヌクレアーゼ活性を示す(図2を参照されたい)。

【0206】

発現データは、候補MG34-1ヌクレアーゼ、tracrRNA、及びCRISPRアレイのin situ天然活性を確認する(図4)。

【0207】

実施例8—mRNA送達を伴う哺乳類細胞における活性

mRNAを用いた細胞トランスフェクション/形質転換を使用したゲノム編集については、コード配列は、Twist Bioscience又はThermo Fisher Scientific(GeneArt)からのアルゴリズムを使用して最適化されるマウス又はヒトコドンである。カセットは、コードエンドヌクレアーゼ配列に付加された2

つの核局在化シグナル：それぞれ、N末端及びC末端に、SV40及びヌクレオプラスミンで構築される。加えて、ヒト補体3（C3）からの非翻訳領域は、カセット内のコード配列の5'及び3'の両方に付加される。

【0208】

次いで、このカセットが、長いポリAストレッチの上流のmRNA産生ベクターにクローニングされる。mRNA構築物の構成は、以下のとおりであり得る。C3からの5'UTR-SV40 NLS-コドン最適化SMART遺伝子-ヌクレオプラスミンNLS-C3からの3'UTR-107ポリAテール。次いで、mRNAの転写の実行は、操作されたT7 RNAポリメラーゼ（Hi-T7：New England Biolabs）を使用してT7プロモーターによって駆動される。mRNAの5'キャッピングは、CleanCap AG（Trilink Biolabs）を使用して共転写的に発生する。次いで、mRNAは、MEGAclear Transcription Clean-Upキット（Thermo Fisher Scientific）を使用して精製される。

10

【0209】

哺乳類細胞が、転写されたmRNAと、Lipofectamine Messenger Max（Thermo Fisher Scientific）を使用して、関心対象のゲノム領域を標的化する少なくとも10個のガイドのセットとで共トランスフェクトされる。細胞がある期間（例えば、48時間）インキュベートされ、続いてPureLink Genomic DNA抽出キット（Fisher Scientific）を使用してゲノムDNA単離される。関心対象の領域は、特定のプライマーを使用して増幅される。次に、編集の結果の完全な分析のために、CRISPR編集及びNGSの推論を使用してサンガーシーケンシングによって編集が評価される。

20

【0210】

実施例9-SMART IIガイドRNA予測

SMART IIエフェクター配列の開始コドンのすぐ上流400bpを含む領域を、活性（UTR）に必要なガイドRNAを潜在的にコードするとして抽出した。UTR配列をMAFFT（mafft-ginsialgorithm）と整列させ、保存ブロックを示す領域を推定ガイドRNAとして注釈付けした。

【0211】

30

実施例10-活性及びPAM決定アッセイ

RNASeqから、又はUTRアラインメントから予測された推定ガイドRNAを、Geneiousでフォールディングした。標的スパーサーをガイドRNAの5'又は3'末端のいずれかに付加して、単一ガイドRNA（sgRNA）を設計した。sgRNAをアセンブリPCRを介して組み立て、SPRIビーズで精製し、製造業者が推奨する短いRNA転写物（HiScribe T7キット、NEB）のプロトコルに従ってインビトロ転写（IVT）した。RNA反応物をMonarch RNAキットで洗浄し、TapeStation（Agilent）を介して純度をチェックした。

【0212】

40

切断及びPAM決定アッセイを、PURExpress（New England Biolabs）を用いて実施した。簡潔に述べると、タンパク質をE.coliに対してコドン最適化し、T7プロモーター及びC末端Hisタグを有するベクターにクローニングした。遺伝子を、それぞれ、T7プロモーター配列及びターミネーター配列の上流及び下流150bpのプライマー結合部位でPCR増幅した。このPCR産物を、5nMの濃度でNEB PURExpressに加え、37℃で2時間発現させた。その後、PURExpressの5倍希釈、5nMの8N PAMプラスミドライブラリー、及びPAMライブラリーを標的化する50nMのsgRNAを用いて、10mMのTris pH7.5、100mMのNaCl、及び10mMのMgCl₂中で切断反応を組み立てた。

【0213】

PURExpress反応からの切断産物を、AMPure SPRIビーズ（Bec

50

kman Coulter)を用いたクリーンアップを介して回収した。DNAは、クレノウ断片及びdNTP(New England Biolabs)の添加を介して平滑化した。平滑末端生成物を、100倍過剰な二本鎖アダプター配列とライゲーションし、NGSライブラリーの調製のための鋳型として使用し、そこからPAM要件を配列分析から決定した。

【0214】

生のNGS読み取りを、Phred品質スコア>20によってフィルタリングした。PAMに隣接する骨格から文書化されたDNA配列を表す24bpを参照として使用して、PAM近位領域を見つけ、隣接する8bpを推定PAMとして特定した。PAMとライゲーションアダプターとの間の距離も、各読み取りに対して測定された。参照配列又はアダプター配列と正確に一致しなかった読み取りは除外された。PAM配列を、最も頻度の高い切断部位±2bpを有するPAMが分析に選択的に含まれるように、切断部位頻度によってフィルタリングした。PAMのフィルタリングされたリストを使用して、Logomakerを使用して配列ロゴを生成した。

10

【0215】

実施例11-SMARTアミノ酸組成物

SMARTタンパク質配列のアミノ酸組成物を説明するために、SMART配列の群についてのアミノ酸含有率を、各残基が観察された回数を、総タンパク質長さで除算し、100倍したものとして計算した。次いで、アミノ酸組成物を、UniProt50データベース(Carugo, Protein Sci. 2008)からのタンパク質配列の大規模セットに報告された含有率と比較した。両方のタンパク質群、SMART HEARO及びSMART(IE-D型)は、Uniref50タンパク質配列で観察される含有量と比較して、異常に高いアルギニン及びリジンアミノ酸含有量を含む(図23)。

20

【0216】

平均して、SMARTのアルギニン及びリジン組成率は、SMART配列の他の残基について観察された線形傾向から、並びにUniref50データベースのタンパク質の残基組成から逸脱する(図24A)。加えて、SMARTのメチオニン含有量は、Uniref50データベースからのタンパク質で観察された含有量よりも統計的に低いことが観察された(図24B)。

【0217】

SMARTの物理化学的特性を説明するために、等電点、分子量、及び電荷を、R中の「protr」及び「ペプチド」パッケージを有する配列から決定した。SMART配列で観察される高アルギニン及びリジン含有量は、中性pHで高等電点及び電荷に寄与し得る(表4)。

30

【0218】

40

50

【表 4】

表4－SMARTファミリーメンバーの理論的特性

ヌクレアーゼ	長さ(a.a.)	MW(Da)	pI	pH7.2における電荷
MG35-1	428	48300.1	11.1	52.6
MG35-2	524	59310.4	10.4	34.3
MG35-3	423	47899.5	10.8	39.8
MG35-6	428	48373.1	10.9	43.6
MG35-102	424	47544.1	10.8	38.7
IscB (Altae-Tran, 2021)	439	49447.5	11.7	47.2
MG34-1	747	86518.4	10.2	46.9
MG102-2	946	107544.0	9.9	38.8
MG102-14	949	108596.1	10.2	50.6
MG102-35	954	108186.4	9.9	38.3
MG102-45	952	107614.8	10.0	40.5

Rパッケージペプチド (The R Journal. 7 (1), 4–14 (2015)) 及び `prot` (*Bioinformatics*, 2015 Jun 1; 31 (11): 1857–9) を使用して特性を計算した。細胞内 pH が 7.0～7.4 の範囲である傾向があるため、pH 7.2 を選択した (*Biochemical Journal*, 1988, 250 (1): 1–8)。

【0219】

SMARTヌクレアーゼの高アルギニン及びZn結合リボンモチーフ含有量は、これらの酵素が天然変性領域を含有し得、これが、大きなガイドRNA及び標的DNAと相互作用する柔軟性をタンパク質に加え得ることを示唆している。天然変性領域は、その天然の非結合状態で安定した三次構造を欠くタンパク質のセグメントであり（例えば、その全体が参照により本明細書に組み込まれる、Bitard-Feildel, T., Lamiable, A., Mornon, J.-P. & Callebaut, I. Order in Disorder as Observed by the “Hydrophobic Cluster Analysis” of Protein Sequences. *Proteomics* 2018, 18, e1800054 を参照されたい）、ポリアニオン（RNA など）と相互作用する正に荷電されたアルギニン中で濃縮され得（例えば、その全体が参照により本明細書に組み込まれる、Murthy, A. C. et al. Molecular interactions underlying liquid-liquid phase separation of the FUS low complexity domain. *Nat Struct Mol Biol* 2019, 26, 637–648 を参照されたい）、Zn結合リボン間のリンカーとして見出されて、「検索機能」を助け得（例えば、その全体が参照により本明細書に組み込まれる、Dyson, H. J. Roles of intrinsic disorder in protein-nucleic acid interactions. *Mol Biosyst* 2011, 8, 97–104 を参照されたい）、それらの全てがSMARTヌクレアーゼにおいて観察される特徴である。

【0220】

実施例12－ミスマッチ殺傷アッセイ

様々なSMART酵素の特異性を決定するために、E. coli BL21 (DE3) 株 (NEB) を、T7 駆動エフェクター (アンピシリン耐性) 及びそれらのT7-駆動sgRNA (クロラムフェニコール耐性) を含有するプラスミドで形質転換し、播種し、一晚増殖させた、ミスマッチ殺傷アッセイを開発した。結果的に得られたコロニーを、能力を有するようにし、骨格中の標的スペーサー及びPAM、各々が24ntスペーサー及び定常PAMに沿った単一のミスマッチを含有する25個のプラスミドのライブラリー、又はスペーサー若しくはPAMを含まない対照プラスミドの3つの条件で、100ngのカナマイシンプラスミドで形質転換した (図25D)。ヒートショック後、形質転換物を、SOC培地中、37℃で2時間回収した。培養物を播種し、誘導培地 (抗生物質及び0.05mM IPTGを含むLB寒天プレート) 上で、37℃で一晩増殖させた。mini prepキット (Qiagen) を介して生存しているミスマッチコロニーからプラスミドを抽出した。標的領域を、PCRを介して増幅させ、NGSを介して分析した。未処理ライブラリーに対して濃縮されたスペーサーは、ヌクレアーゼによって認識及び切断することができず、したがって、エフェクターがミスマッチを許容しない領域であると考えられる。ミスマッチが許容される場合、酵素は、抗生物質耐性プラスミドを切断すると予想され、成長障害が観察されることになる。MG102-2ヌクレアーゼは、PAMからの標的プラスミドの最初の13位に沿ったミスマッチを許容しないことが観察されたが、一方で、14位から可変ミスマッチ耐性が観察された (図25D及び図27)。これらの結果は、SMARTヌクレアーゼが高度に特異的であり得、付随的なssDNA切断を呈さないことを示唆する (図28)。

10

20

【0221】

実施例13-SMARTヌクレアーゼMG102-2を用いたヒト細胞編集

ATCCからのK562細胞を、ATCCプロトコルに従って培養した。TRAC遺伝子座を標的化する2つのsgRNAを、MG102-2 PAMに基づいて設計し、IDTによって化学的に合成した。遺伝子編集実験については、500ngのインビトロで合成されたMG102-2 mRNA、及び150、300、又は450pmolのいずれかの指示されたsgRNAを、Lonza 4D Nucleofector (プログラムFF-120) を使用して 1.5×10^5 個の細胞において共ヌクレオフェクトした。並行して、細胞をmRNAもガイドも用いずにヌクレオフェクトし、TRACガイドによって標的化された部位におけるバックグラウンドを評価した。QuickExtract (Lucigen、#09050) を使用してゲノムDNA抽出のためのエレクトロポレーション後の72時間に細胞を採取し、Illumina Miseq上で次世代シーケンシングのために処理した。結果的に得られたデータを、インデル計算スクリプトを用いて分析した。

30

【0222】

T細胞受容体アルファ定常遺伝子座 (TRAC) を標的化するヒト細胞へのmRNAを介したSMARTヌクレアーゼの送達は、結果的に、MG102-2ヌクレアーゼを有する2つのTRAC標的部位のうちの1つで90%を超える編集活性をもたらした (図26)。インビトロ実験で観察されたように (図29)、sgRNAの量を増加させることは、両方の標的遺伝子座で編集効率が改善した (図26)。ヒト細胞の核へのMG34-1系の局在化 (核局在化シグナル、NLSと融合される) が確認されたが、このヌクレアーゼについてヌクレアーゼ誘導インデル形成は検出されなかった。

40

【0223】

実施例14-SMARTヌクレアーゼの切断選好

MG34-1及びMG102-2ヌクレアーゼの切断産物を配列決定することは、これらの酵素が互い違いの二本鎖DNA切断を作り出すことを示す (図25A)。切断部位の分析は、PAMから5~7位における選択的な切断を示す (図25A)。これらの結果は、ほとんどのCas9酵素と比較して稀に観察される生化学的切断機構を示唆し、これは、平滑末端、及びPAMから3~5位に優先的である互い違いの切断を作り出す。インビトロ転写/翻訳反応及び精製されたタンパク質を用いたインビトロ切断アッセイは、MG

50

34-1及びMG102-2が18及び20個のヌクレオチドスペーサーで最も効率的であることを示す(図25C)。更に、E. coliプラスミド干渉アッセイを使用してインビボで活性を確認し、指定された標的化スペーサーを有する5つのSMARTヌクレアーゼについて2倍(MG34-9)〜>500倍(MG102-2)の成長抑制を示した(図25B)。

【0224】

実施例15-SMART I酵素はヒト細胞における活性ヌクレアーゼである

ATCCから購入したK562細胞を、ATCCプロトコルに従って培養した。TRAC又はAAVS1遺伝子座を標的化するsgRNAを、MG102-2、MG102-36、MG102-39、MG102-42、MG102-45、及びMG33-34によって認識されるPAMに基づいて設計し、IDTによって化学的に合成した。遺伝子編集実験については、500ngのインビトロで合成されたヌクレアーゼmRNA、及び450pmolの指示されたsgRNAを、Lonza 4D Nucleofector (プログラムFF-120)を使用して 1.5×10^5 個の細胞において共ヌクレオフェクトした。QuickExtract (Lucigen、#09050)を使用してゲノムDNA抽出のためのエレクトロポレーション後の72時間に細胞を採取し、Illumina MiSeq上でアンブリコン次世代シーケンシングのために処理した。結果的に得られたデータを、インハウスインデル計算スクリプトを用いて分析した。

【0225】

本明細書の他の箇所に説明されるように、SMART IヌクレアーゼMG102-2は、mRNAを介して送達されたとき、ヒトゲノムのTRAC遺伝子座の2つの標的部位で活性である。MG102-2 (配列番号582)はまた、ヒトゲノム中のAAVS1遺伝子座(セーフハーバー遺伝子座)でも活性であり、酵素の切断効率は8つの異なる標的部位で82.6%程度高く、>50%の編集効率であることが更に確認された(図30A)。加えて、MG102-39 (配列番号993)、MG102-42 (配列番号996)、及びMG102-48 (配列番号1002)は、mRNAによって送達されたとき、ヒトゲノムのTRAC遺伝子座で>40%の切断活性を示したが(図30B〜図30D)、一方で、MG33-34 (配列番号988)、MG102-36 (配列番号990)、及びMG102-45 (配列番号999)は、TRAC遺伝子座でバックグラウンド(10%)を上回る切断効率を示した(図30E〜図30G)。

【0226】

10

20

30

40

50

【表 5 - 1】

表 5 - 実施例 15 で試験されたガイド RNA 及び標的化配列

配列番号	名称	配列
1087	MG102-2 AAVS1 A5	mU*mU*mC*rUrGrGrGrArGrArGrGrGrUrArGrCrGrCr ArGrGrGrUrUrUrUrCrArArUrCrArArArCrUrGrArAr ArArGrUrUrCrCrGrGrUrUrGrArArArArGrArGrCrArU rCrCrGrUrCrCrGrGrArGrGrGrUrGrCrArCrUrCrCrGr GrGrArUrGrGrGrGrCrArGrUrCrCrCrGrGrCrArCrUrU rGrCrGrUrUrUrUrCrCrCrCrGrGrCrUrUrArCrGrCrUrU rCrGrGrArArArArArGrGrCrCrCrUrUrCrGrGrCrArCr GrUrCrGrArArArGrArCrArGrGrArUrGrUrGrArGrCrC rCrArA*mU*mU*mU
1088	MG102-2 AAVS1 H8	mG*mC*mC*rCrUrGrGrGrArArUrArUrArArGrGrUrGr GrUrCrCrCrGrUrUrUrCrArArUrCrArArArCrUrGrArAr ArArGrUrUrCrCrGrGrUrUrGrArArArArGrArGrCrArU rCrCrGrUrCrCrGrGrArGrGrGrUrGrCrArCrUrCrCrGr GrGrArUrGrGrGrGrCrArGrUrCrCrCrGrGrCrArCrUrU rGrCrGrUrUrUrUrCrCrCrCrGrGrCrUrUrArCrGrCrUrU rCrGrGrArArArArArGrGrCrCrCrUrUrCrGrGrCrArCr GrUrCrGrArArArGrArCrArGrGrArUrGrUrGrArGrCrC rCrArA*mU*mU*mU
1089	MG102-2 AAVS1 H9	mA*mU*mG*rCrUrGrUrCrCrUrGrArArGrUrGrGrArCr ArUrArGrGrGrUrUrUrCrArArUrCrArArArCrUrGrArAr ArArGrUrUrCrCrGrGrUrUrGrArArArArGrArGrCrArU rCrCrGrUrCrCrGrGrArGrGrGrUrGrCrArCrUrCrCrGr GrGrArUrGrGrGrGrCrArGrUrCrCrCrGrGrCrArCrUrU rGrCrGrUrUrUrUrCrCrCrCrGrGrCrUrUrArCrGrCrUrU rCrGrGrArArArArArGrGrCrCrCrUrUrCrGrGrCrArCr GrUrCrGrArArArGrArCrArGrGrArUrGrUrGrArGrCrC rCrArA*mU*mU*mU
1090	MG102-2 AAVS1 D11	mC*mU*mA*rGrArGrArGrGrUrArArGrGrGrGrGrUr ArGrGrGrGrGrUrUrUrCrArArUrCrArArArCrUrGrArA rArArGrUrUrCrCrGrGrUrUrGrArArArArGrArGrCrAr UrCrCrGrUrCrCrGrGrArGrGrGrUrGrCrArCrUrCrCrG rGrGrArUrGrGrGrGrCrArGrUrCrCrCrGrGrCrArCrUr UrGrCrGrUrUrUrUrCrCrCrCrGrGrCrUrUrArCrGrCrUr UrCrGrGrArArArArArGrGrCrCrCrUrUrCrGrGrCrArC rGrUrCrGrArArArGrArCrArGrGrArUrGrUrGrArGrCr CrCrArA*mU*mU*mU
1091	MG102-2 AAVS1 E7	mA*mG*mG*rArArGrGrArGrGrArGrGrCrCrUrArArGr GrArUrGrGrGrUrUrUrCrArArUrCrArArArCrUrGrArAr ArArGrUrUrCrCrGrGrUrUrGrArArArArGrArGrCrArU rCrCrGrUrCrCrGrGrArGrGrGrUrGrCrArCrUrCrCrGr GrGrArUrGrGrGrGrCrArGrUrCrCrCrGrGrCrArCrUrU rGrCrGrUrUrUrUrCrCrCrCrGrGrCrUrUrArCrGrCrUrU rCrGrGrArArArArArGrGrCrCrCrUrUrCrGrGrCrArCr GrUrCrGrArArArGrArCrArGrGrArUrGrUrGrArGrCrC rCrArA*mU*mU*mU

【 0 2 2 7 】

10

20

30

40

50

【表 5 - 2】

配列番号	名称	配列
1092	MG102-2 AAVS1 D7	mA*mU*mA*rUrCrArGrGrArGrArCrUrArGrGrArArGr GrArGrGrArGrUrUrUrCrArArUrCrArArArCrUrGrArAr ArArGrUrUrCrCrGrGrUrUrGrArArArArGrArGrCrArU rCrCrGrUrCrCrGrGrArGrGrGrUrGrCrArCrUrCrCrGr GrGrArUrGrGrGrGrCrArGrUrCrCrCrGrGrCrArCrUrU rGrCrGrUrUrUrUrCrCrCrCrGrGrCrUrUrArCrGrCrUrU rCrGrGrArArArArGrGrCrCrCrUrUrCrGrGrCrArCr GrUrCrGrArArArGrArCrArGrGrArUrGrUrGrArGrCrC rCrArA*mU*mU*mU
1093	MG102-2 AAVS1 B7	mC*mU*mG*rCrCrUrArArCrArGrGrArGrGrUrGrGrGr GrGrUrUrArGrUrUrUrCrArArUrCrArArArCrUrGrArAr ArArGrUrUrCrCrGrGrUrUrGrArArArArGrArGrCrArU rCrCrGrUrCrCrGrGrArGrGrGrUrGrCrArCrUrCrCrGr GrGrArUrGrGrGrGrCrArGrUrCrCrCrGrGrCrArCrUrU rGrCrGrUrUrUrUrCrCrCrCrGrGrCrUrUrArCrGrCrUrU rCrGrGrArArArArGrGrCrCrCrUrUrCrGrGrCrArCr GrUrCrGrArArArGrArCrArGrGrArUrGrUrGrArGrCrC rCrArA*mU*mU*mU
1094	MG102-2 AAVS1 D12	mG*mC*mA*rArGrArGrGrArUrGrGrArGrArGrUrGr GrCrUrArArGrUrUrUrCrArArUrCrArArArCrUrGrArAr ArArGrUrUrCrCrGrGrUrUrGrArArArArGrArGrCrArU rCrCrGrUrCrCrGrGrArGrGrGrUrGrCrArCrUrCrCrGr GrGrArUrGrGrGrGrCrArGrUrCrCrCrGrGrCrArCrUrU rGrCrGrUrUrUrUrCrCrCrCrGrGrCrUrUrArCrGrCrUrU rCrGrGrArArArArGrGrCrCrCrUrUrCrGrGrCrArCr GrUrCrGrArArArGrArCrArGrGrArUrGrUrGrArGrCrC rCrArA*mU*mU*mU
1095	MG102-2 AAVS1 C8	mG*mA*mG*rGrGrGrArCrArGrArUrArArArArGrUrAr CrCrCrArGrGrUrUrUrCrArArUrCrArArArCrUrGrArAr ArArGrUrUrCrCrGrGrUrUrGrArArArArGrArGrCrArU rCrCrGrUrCrCrGrGrArGrGrGrUrGrCrArCrUrCrCrGr GrGrArUrGrGrGrGrCrArGrUrCrCrCrGrGrCrArCrUrU rGrCrGrUrUrUrUrCrCrCrCrGrGrCrUrUrArCrGrCrUrU rCrGrGrArArArArGrGrCrCrCrUrUrCrGrGrCrArCr GrUrCrGrArArArGrArCrArGrGrArUrGrUrGrArGrCrC rCrArA*mU*mU*mU
1096	MG102-2 AAVS1 A8	mG*mU*mG*rGrCrCrCrCrArCrUrGrUrGrGrGrGrUrGr GrArGrGrGrGrUrUrUrCrArArUrCrArArArCrUrGrArA rArArGrUrUrCrCrGrGrUrUrGrArArArArGrArGrCrAr UrCrCrGrUrCrCrGrGrArGrGrGrUrGrCrArCrUrCrCrG rGrGrArUrGrGrGrGrCrArGrUrCrCrCrGrGrCrArCrUr UrGrCrGrUrUrUrUrCrCrCrCrGrGrCrUrUrArCrGrCrUr UrCrGrGrArArArArGrGrCrCrCrUrUrCrGrGrCrArC rGrUrCrGrArArArGrArCrArGrGrArUrGrUrGrArGrCr CrCrArA*mU*mU*mU
1097	MG102-2 AAVS1 G6	mU*mG*mG*rCrUrCrCrArGrGrArArUrGrGrGrGrGr UrGrUrGrUrGrUrUrUrCrArArUrCrArArArCrUrGrArAr ArArGrUrUrCrCrGrGrUrUrGrArArArArGrArGrCrArU rCrCrGrUrCrCrGrGrArGrGrGrUrGrCrArCrUrCrCrGr GrGrArUrGrGrGrGrCrArGrUrCrCrCrGrGrCrArCrUrU rGrCrGrUrUrUrUrCrCrCrCrGrGrCrUrUrArCrGrCrUrU rCrGrGrArArArArGrGrCrCrCrUrUrCrGrGrCrArCr GrUrCrGrArArArGrArCrArGrGrArUrGrUrGrArGrCrC rCrArA*mU*mU*mU

10

20

30

40

【 0 2 2 8 】

【表 5 - 3】

配列番号	名称	配列
1098	MG102-2 AAVS1 E5	mG*mU*mG*rGrCrCrArCrUrGrArGrArArCrCrGrGrGr CrArGrGrUrGrUrUrCrArArUrCrArArArCrUrGrArAr ArArGrUrUrCrCrGrGrUrUrGrArArArArGrArGrCrArU rCrCrGrUrCrCrGrGrArGrGrGrUrGrCrArCrUrCrCrGr GrGrArUrGrGrGrGrCrArGrUrCrCrCrGrGrCrArCrUrU rGrCrGrUrUrUrCrCrCrCrGrGrCrUrUrArCrGrCrUrU rCrGrGrArArArArGrGrCrCrCrUrUrCrGrGrCrArCr GrUrCrGrArArArGrArCrArGrGrArUrGrUrGrArGrCrC rCrArA*mU*mU*mU
1099	MG102-2 AAVS1 G7	mU*mC*mU*rGrUrCrArCrArArUrCrCrUrGrUrCrCr CrUrArGrUrGrUrUrUrCrArArUrCrArArArCrUrGrArAr ArArGrUrUrCrCrGrGrUrUrGrArArArArGrArGrCrArU rCrCrGrUrCrCrGrGrArGrGrGrUrGrCrArCrUrCrCrGr GrGrArUrGrGrGrGrCrArGrUrCrCrCrGrGrCrArCrUrU rGrCrGrUrUrUrCrCrCrCrGrGrCrUrUrArCrGrCrUrU rCrGrGrArArArArGrGrCrCrCrUrUrCrGrGrCrArCr GrUrCrGrArArArGrArCrArGrGrArUrGrUrGrArGrCrC rCrArA*mU*mU*mU
1100	MG102-2 AAVS1 C3	mU*mU*mC*rUrCrCrUrCrUrUrGrGrGrArArGrUrGrUr ArArGrGrArGrUrUrUrCrArArUrCrArArArCrUrGrArAr ArArGrUrUrCrCrGrGrUrUrGrArArArArGrArGrCrArU rCrCrGrUrCrCrGrGrArGrGrGrUrGrCrArCrUrCrCrGr GrGrArUrGrGrGrGrCrArGrUrCrCrCrGrGrCrArCrUrU rGrCrGrUrUrUrCrCrCrCrGrGrCrUrUrArCrGrCrUrU rCrGrGrArArArArGrGrCrCrCrUrUrCrGrGrCrArCr GrUrCrGrArArArGrArCrArGrGrArUrGrUrGrArGrCrC rCrArA*mU*mU*mU
1101	MG102-2 AAVS1 E1	mC*mC*mU*rGrCrCrArGrGrArCrGrGrGrGrCrUrGrGr CrUrArCrUrGrUrUrUrCrArArUrCrArArArCrUrGrArAr ArArGrUrUrCrCrGrGrUrUrGrArArArArGrArGrCrArU rCrCrGrUrCrCrGrGrArGrGrGrUrGrCrArCrUrCrCrGr GrGrArUrGrGrGrGrCrArGrUrCrCrCrGrGrCrArCrUrU rGrCrGrUrUrUrCrCrCrCrGrGrCrUrUrArCrGrCrUrU rCrGrGrArArArArGrGrCrCrCrUrUrCrGrGrCrArCr GrUrCrGrArArArGrArCrArGrGrArUrGrUrGrArGrCrC rCrArA*mU*mU*mU
1102	MG102-2 AAVS1 E2	mA*mA*mA*rUrUrGrGrGrGrArCrUrArGrArArArGrGr UrGrArArGrGrUrUrUrCrArArUrCrArArArCrUrGrArAr ArArGrUrUrCrCrGrGrUrUrGrArArArArGrArGrCrArU rCrCrGrUrCrCrGrGrArGrGrGrUrGrCrArCrUrCrCrGr GrGrArUrGrGrGrGrCrArGrUrCrCrCrGrGrCrArCrUrU rGrCrGrUrUrUrCrCrCrCrGrGrCrUrUrArCrGrCrUrU rCrGrGrArArArArGrGrCrCrCrUrUrCrGrGrCrArCr GrUrCrGrArArArGrArCrArGrGrArUrGrUrGrArGrCrC rCrArA*mU*mU*mU
1103	MG102-2 AAVS1 H6	mG*mG*mG*rUrGrUrGrUrCrArCrArGrArUrArArGr GrArArUrCrGrUrUrUrCrArArUrCrArArArCrUrGrArAr ArArGrUrUrCrCrGrGrUrUrGrArArArArGrArGrCrArU rCrCrGrUrCrCrGrGrArGrGrGrUrGrCrArCrUrCrCrGr GrGrArUrGrGrGrGrCrArGrUrCrCrCrGrGrCrArCrUrU rGrCrGrUrUrUrCrCrCrCrGrGrCrUrUrArCrGrCrUrU rCrGrGrArArArArGrGrCrCrCrUrUrCrGrGrCrArCr GrUrCrGrArArArGrArCrArGrGrArUrGrUrGrArGrCrC rCrArA*mU*mU*mU

【 0 2 2 9 】

10

20

30

40

50

【表 5 - 4】

配列番号	名称	配列
1104	MG102-2 AAVS1 H11	mA*mG*mA*rGrGrUrGrArCrCrGrArArUrCrCrArCr ArGrGrArGrGrUrUrUrCrArArUrCrArArArCrUrGrArAr ArArGrUrUrCrCrGrGrUrUrGrArArArGrArGrCrArU rCrCrGrUrCrCrGrGrArGrGrGrUrGrCrArCrUrCrCrGr GrGrArUrGrGrGrGrCrArGrUrCrCrGrGrCrArCrUrU rGrCrGrUrUrUrUrCrCrCrGrGrCrUrUrArCrGrCrUrU rCrGrGrArArArArGrGrCrCrCrUrUrCrGrGrCrArCr GrUrCrGrArArArGrArCrArGrGrArUrGrUrGrArGrCrC rCrArA*mU*mU*mU
1105	MG102-2 AAVS1 A5	TTCTGGGAGAGGGTAGCGCAGGGT
1106	MG102-2 AAVS1 H8	GCCCTGGGAATATAAGGTGGTCCC
1107	MG102-2 AAVS1 H9	ATGCTGTCCTGAAGTGGACATAGG
1108	MG102-2 AAVS1 D11	CTAGAGAGGTAAGGGGGGTAGGGG
1109	MG102-2 AAVS1 E7	AGGAAGGAGGAGGCCTAAGGATGG
1110	MG102-2 AAVS1 D7	ATATCAGGAGACTAGGAAGGAGGA
1111	MG102-2 AAVS1 B7	CTGCCTAACAGGAGGTGGGGGTTA
1112	MG102-2 AAVS1 D12	GCAAGAGGATGGAGAGGTGGCTAA
1113	MG102-2 AAVS1 C8	GAGGGGACAGATAAAAAGTACCCAG
1114	MG102-2 AAVS1 A8	GTGGCCCCACTGTGGGGTGGAGGG
1115	MG102-2 AAVS1 G6	TGGCTCCAGGAAATGGGGGTGTGT
1116	MG102-2 AAVS1 E5	GTGGCCACTGAGAACCGGGCAGGT
1117	MG102-2 AAVS1 G7	TCTGTCACCAATCCTGTCCCTAGT
1118	MG102-2 AAVS1 C3	TTCTCCTCTTGGGAAGTGTAAGGA
1119	MG102-2 AAVS1 E1	CCTGCCAGGACGGGGCTGGCTACT
1120	MG102-2 AAVS1 E2	AAATTGGGGACTAGAAAGGTGAAG
1121	MG102-2 AAVS1 H6	GGGTGTGTCACCAGATAAGGAATC
1122	MG102-2 AAVS1 H11	AGAGGTGACCCGAATCCACAGGAG
1123	MG102-36 TRAC D12	mG*mC*mC*rArCrUrUrUrCrArGrGrArGrGrArGrAr UrUrCrGrGrGrUrUrUrCrArArUrCrArArArCrUrGrArAr ArArGrUrUrCrCrGrGrUrUrGrArArArArGrArGrCrArU rCrCrGrUrCrUrGrUrCrArGrGrCrArCrUrCrCrGrGr ArUrGrGrGrGrCrArGrUrCrCrCrGrGrCrUrCrUrUrGrC rGrGrUrUrArCrCrGrArUrGrCrGrGrCrArArCrGrUrGr UrCrGrArUrGrUrArGrCrCrArArCrUrGrCrCrArGrArCr ArCrGrUrCrUrUrUrGrArCrArGrGrArUrGrUrGrArG rCrCrCrArU*mU*mU*mU

10

20

30

40

【 0 2 3 0 】

50

【表 5 - 5】

配列番号	名称	配列
1124	MG102-36 TRAC F1	mG*mA*mC*rCrCrUrGrCrCrGrUrGrUrArCrCrArGrCr UrGrArGrArGrUrUrUrCrArArUrCrArArArCrUrGrArAr ArArGrUrUrCrCrGrGrUrUrGrArArArArGrArGrCrArU rCrCrGrUrCrUrGrUrCrArGrGrCrArCrUrCrCrCrGrGr ArUrGrGrGrGrCrArGrUrCrCrCrGrGrCrUrCrUrUrGrC rGrGrUrUrArCrCrGrArUrGrCrGrGrCrArArCrGrUrGr UrCrGrArUrGrUrArGrCrCrArArCrUrGrCrCrArGrArCr ArCrGrUrCrUrUrUrUrGrArCrArGrGrArUrGrUrGrArG rCrCrCrArU*mU*mU*mU
1125	MG102-36 TRAC H6	mU*mU*mG*rArArGrUrCrCrArUrArGrArCrCrUrCrAr UrGrUrCrUrGrUrUrUrCrArArUrCrArArArCrUrGrArAr ArArGrUrUrCrCrGrGrUrUrGrArArArArGrArGrCrArU rCrCrGrUrCrUrGrUrCrArGrGrCrArCrUrCrCrCrGrGr ArUrGrGrGrGrCrArGrUrCrCrCrGrGrCrUrCrUrUrGrC rGrGrUrUrArCrCrGrArUrGrCrGrGrCrArArCrGrUrGr UrCrGrArUrGrUrArGrCrCrArArCrUrGrCrCrArGrArCr ArCrGrUrCrUrUrUrUrGrArCrArGrGrArUrGrUrGrArG rCrCrCrArU*mU*mU*mU
1126	MG102-39 TRAC F4	mG*mC*mU*rGrCrCrCrUrUrArCrCrUrGrGrGrCrUrGr GrGrGrArArGrUrUrUrCrArGrUrUrArCrCrCrUrGrArG rArArArUrCrArGrGrCrUrGrArArGrCrUrGrArArArAr GrArGrCrArUrCrCrGrUrCrCrGrGrArArGrGrUrCrCrA rCrUrCrCrGrGrGrUrUrArGrGrGrCrArGrArUrCrCrGr GrCrUrCrUrUrGrGrUrCrCrUrCrUrCrCrUrGrGrCrCrCr UrUrUrUrCrGrGrGrCrUrCrCrGrArGrArGrGrArArGrC rCrUrUrCrCrGrGrCrArUrGrUrCrUrUrCrGrGrArCrAr GrGrArUrGrUrGrArGrCrCrUrUrU*mU*mU*mU
1127	MG102-39 TRAC A9	mU*mC*mU*rUrGrGrUrUrUrUrArCrArGrArUrArCrGr ArArCrCrUrGrUrUrUrCrArGrUrUrArCrCrCrUrGrArGr ArArArUrCrArGrGrCrUrGrArArGrCrUrGrArArArArG rArGrCrArUrCrCrGrUrCrCrGrGrArArGrGrUrCrCrAr CrUrCrCrGrGrGrUrUrArGrGrGrCrArGrArUrCrCrGrG rCrUrCrUrUrGrGrUrCrCrUrCrUrCrCrUrGrGrCrCrCrU rUrUrUrCrGrGrGrCrUrCrCrGrArGrArGrGrArArGrCr CrUrUrCrCrGrGrCrArUrGrUrCrUrUrCrGrGrArCrArG rGrArUrGrUrGrArGrCrCrUrUrU*mU*mU*mU
1128	MG102-39 TRAC G11	mG*mG*mC*rCrArCrUrUrUrCrArGrGrArGrGrArGrGr ArUrUrCrGrGrUrUrUrCrArGrUrUrArCrCrCrUrGrArGr ArArArUrCrArGrGrCrUrGrArArGrCrUrGrArArArArG rArGrCrArUrCrCrGrUrCrCrGrGrArArGrGrUrCrCrAr CrUrCrCrGrGrGrUrUrArGrGrGrCrArGrArUrCrCrGrG rCrUrCrUrUrGrGrUrCrCrUrCrUrCrCrUrGrGrCrCrCrU rUrUrUrCrGrGrGrCrUrCrCrGrArGrArGrGrArArGrCr CrUrUrCrCrGrGrCrArUrGrUrCrUrUrCrGrGrArCrArG rGrArUrGrUrGrArGrCrCrUrUrU*mU*mU*mU

10

20

30

40

【 0 2 3 1 】

【表 5 - 6】

配列番号	名称	配列
1129	MG102-39 TRAC C11	mC*mA*mG*rCrCrGrCrArGrCrGrUrCrArUrGrArGrCr ArGrArUrUrGrUrUrUrCrArGrUrUrArCrCrCrUrGrArGr ArArArUrCrArGrGrCrUrGrArArGrCrUrGrArArArArG rArGrCrArUrCrCrGrUrCrCrGrGrArArGrGrUrCrCrAr CrUrCrCrGrGrGrUrUrArGrGrGrCrArGrArUrCrCrGrG rCrUrCrUrUrGrGrUrCrCrUrCrUrCrUrGrGrCrCrCrU rUrUrUrCrGrGrGrCrUrCrCrGrArGrArGrGrArArGrCr CrUrUrCrCrGrGrCrArUrGrUrCrUrUrCrGrGrArCrArG rGrArUrGrUrGrArGrCrCrUrUrU*mU*mU*mU
1130	MG102-39 TRAC B6	mC*mC*mA*rGrGrCrCrArCrArGrCrArCrUrGrUrUrGr CrUrCrUrUrGrUrUrUrCrArGrUrUrArCrCrCrUrGrArGr ArArArUrCrArGrGrCrUrGrArArGrCrUrGrArArArArG rArGrCrArUrCrCrGrUrCrCrGrGrArArGrGrUrCrCrAr CrUrCrCrGrGrGrUrUrArGrGrGrCrArGrArUrCrCrGrG rCrUrCrUrUrGrGrUrCrCrUrCrUrCrUrGrGrCrCrCrU rUrUrUrCrGrGrGrCrUrCrCrGrArGrArGrGrArArGrCr CrUrUrCrCrGrGrCrArUrGrUrCrUrUrCrGrGrArCrArG rGrArUrGrUrGrArGrCrCrUrUrU*mU*mU*mU
1131	MG102-39 TRAC B5	mG*mU*mC*rUrUrCrUrGrGrArArUrArArUrGrCrUrGr UrUrGrUrUrGrUrUrUrCrArGrUrUrArCrCrCrUrGrArGr ArArArUrCrArGrGrCrUrGrArArGrCrUrGrArArArArG rArGrCrArUrCrCrGrUrCrCrGrGrArArGrGrUrCrCrAr CrUrCrCrGrGrGrUrUrArGrGrGrCrArGrArUrCrCrGrG rCrUrCrUrUrGrGrUrCrCrUrCrUrCrUrGrGrCrCrCrU rUrUrUrCrGrGrGrCrUrCrCrGrArGrArGrGrArArGrCr CrUrUrCrCrGrGrCrArUrGrUrCrUrUrCrGrGrArCrArG rGrArUrGrUrGrArGrCrCrUrUrU*mU*mU*mU
1132	MG102-39 TRAC G9	mG*mA*mU*rUrGrGrGrUrUrCrCrGrArArUrCrUrCr CrUrCrCrUrGrUrUrUrCrArGrUrUrArCrCrCrUrGrArGr ArArArUrCrArGrGrCrUrGrArArGrCrUrGrArArArArG rArGrCrArUrCrCrGrUrCrCrGrGrArArGrGrUrCrCrAr CrUrCrCrGrGrGrUrUrArGrGrGrCrArGrArUrCrCrGrG rCrUrCrUrUrGrGrUrCrCrUrCrUrCrUrGrGrCrCrCrU rUrUrUrCrGrGrGrCrUrCrCrGrArGrArGrGrArArGrCr CrUrUrCrCrGrGrCrArUrGrUrCrUrUrCrGrGrArCrArG rGrArUrGrUrGrArGrCrCrUrUrU*mU*mU*mU
1133	MG102-39 TRAC D1	mA*mU*mU*rCrUrGrArUrGrUrGrUrArUrArUrCrArCr ArGrArCrArGrUrUrUrCrArGrUrUrArCrCrCrUrGrArGr ArArArUrCrArGrGrCrUrGrArArGrCrUrGrArArArArG rArGrCrArUrCrCrGrUrCrCrGrGrArArGrGrUrCrCrAr CrUrCrCrGrGrGrUrUrArGrGrGrCrArGrArUrCrCrGrG rCrUrCrUrUrGrGrUrCrCrUrCrUrCrUrGrGrCrCrCrU rUrUrUrCrGrGrGrCrUrCrCrGrArGrArGrGrArArGrCr CrUrUrCrCrGrGrCrArUrGrUrCrUrUrCrGrGrArCrArG rGrArUrGrUrGrArGrCrCrUrUrU*mU*mU*mU
1134	MG102-39 TRAC B11	mA*mC*mA*rGrCrCrGrCrArGrCrGrUrCrArUrGrArGr CrArGrArUrGrUrUrUrCrArGrUrUrArCrCrCrUrGrArGr ArArArUrCrArGrGrCrUrGrArArGrCrUrGrArArArArG rArGrCrArUrCrCrGrUrCrCrGrGrArArGrGrUrCrCrAr CrUrCrCrGrGrGrUrUrArGrGrGrCrArGrArUrCrCrGrG rCrUrCrUrUrGrGrUrCrCrUrCrUrCrUrGrGrCrCrCrU rUrUrUrCrGrGrGrCrUrCrCrGrArGrArGrGrArArGrCr CrUrUrCrCrGrGrCrArUrGrUrCrUrUrCrGrGrArCrArG rGrArUrGrUrGrArGrCrCrUrUrU*mU*mU*mU

10

20

30

40

【 0 2 3 2 】

【表 5 - 7】

配列番号	名称	配列
1135	MG102-39 TRAC D4	mA*mA*mA*rGrCrUrGrCrCrCrUrUrArCrCrUrGrGrGrCrUrGrGrGrGrUrUrCrArGrUrUrArCrCrCrUrGrArGrArArArUrCrArGrGrCrUrGrArArGrCrUrGrArArArGrArGrCrArUrCrCrGrUrCrCrGrGrArArGrGrUrCrCrArCrUrCrCrGrGrGrUrUrArGrGrGrCrArGrArUrCrCrGrGrCrUrCrUrUrGrGrUrCrCrUrCrUrCrUrGrGrCrCrCrUrUrUrCrGrGrGrCrUrCrCrGrArGrArGrArArGrCrCrUrUrCrGrGrCrArUrGrUrCrUrUrCrGrGrArCrArGrGrArUrGrUrGrArGrCrCrUrUrU*mU*mU*mU
1136	MG102-39 TRAC F2	mC*mA*mA*rCrArGrUrGrCrUrGrUrGrGrCrCrUrGrGrArGrCrArArGrUrUrUrCrArGrUrUrArCrCrCrUrGrArGrArArArUrCrArGrGrCrUrGrArArGrCrUrGrArArArGrArGrCrArUrCrCrGrUrCrCrGrGrArArGrGrUrCrCrArCrUrCrCrGrGrGrUrUrArGrGrGrCrArGrArUrCrCrGrGrCrUrCrUrUrGrGrUrCrCrUrCrUrCrCrUrGrGrCrCrCrUrUrUrCrGrGrGrCrUrCrCrGrArGrArGrArArGrCrCrUrUrCrCrGrGrCrArUrGrUrCrUrUrCrGrGrArCrArGrGrArUrGrUrGrArGrCrCrUrUrU*mU*mU*mU
1137	MG102-39 TRAC G1	mG*mC*mU*rArGrArCrArUrGrArGrGrUrCrUrArUrGrGrArCrUrUrGrUrUrUrCrArGrUrUrArCrCrCrUrGrArGrArArArUrCrArGrGrCrUrGrArArArArGrArGrCrArArGrCrArUrCrCrGrUrCrCrGrGrArArGrGrUrCrCrArCrUrCrCrGrGrGrUrUrArGrGrGrCrArGrArUrCrCrGrGrCrUrCrUrUrGrGrUrCrCrUrCrUrCrUrGrGrCrCrCrUrUrUrCrGrGrGrCrUrCrCrGrArGrArGrArArGrCrCrUrUrCrCrGrGrCrArUrGrUrCrUrUrCrGrGrArCrArGrGrArUrGrUrGrArGrCrCrUrUrU*mU*mU*mU
1138	MG102-42 TRAC D10	mG*mU*mU*rCrCrGrArArUrCrCrUrCrCrUrCrCrUrGrArArArArGrUrGrUrUrCrArGrCrCrArArCrCrUrGrArArArArGrGrUrGrGrUrGrArCrUrGrArArArArGrArGrCrCrArCrArGrCrCrGrGrCrArGrUrUrCrCrCrGrGrCrCrCrUrGrCrArArGrGrCrArGrCrArCrArGrArGrArArGrCrGrUrGrCrCrGrArArArUrGrGrCrGrCrCrGrGrCrUrUrArUrGrUrGrGrUrGrArGrUrCrCrArUrUrArUrU*mU*mU*mU
1139	MG102-42 TRAC D12	mG*mC*mC*rArCrUrUrUrCrArGrGrArGrGrArGrGrArUrUrCrGrGrGrUrUrUrCrArGrCrCrArArCrCrUrGrArArArArGrGrUrGrGrUrGrArCrUrGrArArArArGrArGrCrCrArCrArGrCrCrGrGrCrArGrUrUrCrCrCrGrGrCrCrCrUrGrCrArArGrGrCrArGrCrArCrArGrArGrArArGrCrGrUrGrCrCrGrArArArUrGrGrCrGrCrCrGrGrCrUrUrArUrGrUrGrGrUrGrArGrUrCrCrArUrUrArUrU*mU*mU*mU
1140	MG102-42 TRAC E12	mC*mA*mG*rGrArGrGrArGrGrArUrUrCrGrGrArArCrCrCrArArUrGrUrUrCrArGrCrCrArArCrCrUrGrArArArArGrGrUrGrGrUrGrArCrUrGrArArArArGrArGrCrCrArCrArGrCrCrGrGrCrArGrUrUrCrCrCrGrGrCrCrCrUrGrCrArArGrGrCrArGrCrArCrArGrArGrArArGrCrGrUrGrCrCrGrArArArUrGrGrCrGrCrCrGrGrCrUrUrArUrGrUrGrGrUrGrArGrUrCrCrArUrUrArUrU*mU*mU*mU

10

20

30

40

【 0 2 3 3 】

【表 5 - 8】

配列番号	名称	配列
		rUrGrCrArArGrGrCrArGrCrArCrArGrArGrArArGrCr GrUrGrCrCrGrArArArUrGrGrCrGrCrGrGrCrUrUrA rUrGrUrGrGrUrGrArGrUrCrCrArUrUrUrArUrU*mU*m U*mU
1141	MG102-45 TRAC B1	mU*mG*mU*rCrCrCrArCrArGrArUrArUrCrCrArGrAr ArCrCrCrUrGrUrUrUrCrArArUrCrArArGrCrUrGrArAr ArArGrCrUrCrCrGrGrUrUrGrArArArArGrArGrCrArU rCrCrGrUrCrUrGrArUrArGrCrCrArUrGrCrArCrUrCrC rGrGrArArUrGrGrGrGrCrArGrUrUrCrCrGrGrCrUrCr UrUrGrCrGrArCrUrCrArArUrGrGrGrUrGrUrArUrGrC rUrCrArUrUrGrArGrCrCrArArCrUrGrUrCrArGrArCrA rCrGrUrCrUrCrUrCrUrGrArGrArCrArGrGrArUrGrUr GrArGrCrCrCrUrUrA*mU*mU*mU
1142	MG102-45 TRAC C11	mC*mU*mU*rCrArArGrGrCrCrCrCrUrCrArCrCrUrCr ArGrCrUrGrGrUrUrUrCrArArUrCrArArGrCrUrGrArAr ArArGrCrUrCrCrGrGrUrUrGrArArArArGrArGrCrArU rCrCrGrUrCrUrGrArUrArGrCrCrArUrGrCrArCrUrCrC rGrGrArArUrGrGrGrGrCrArGrUrUrCrCrGrGrCrUrCr UrUrGrCrGrArCrUrCrArArUrGrGrGrUrGrUrArUrGrC rUrCrArUrUrGrArGrCrCrArArCrUrGrUrCrArGrArCrA rCrGrUrCrUrCrUrCrUrGrArGrArCrArGrGrArUrGrUr GrArGrCrCrCrUrUrA*mU*mU*mU
1143	MG102-48 TRAC A1	mU*mC*mC*rUrCrUrUrGrUrCrCrCrArCrArGrArUrAr UrCrCrArGrGrUrUrUrCrArArUrCrArArCrGrGrArAr ArCrGrGrUrCrCrGrGrUrUrGrArArArArGrArGrCrArU rCrGrGrUrCrUrGrArArGrGrArUrGrCrArCrUrCrCrGr GrGrArUrArGrGrGrCrArGrUrCrCrCrGrGrCrUrCrUrU rGrCrUrGrUrUrUrCrCrCrGrGrUrArArGrArCrCrUrC rGrGrArArGrCrArArGrUrCrCrUrUrCrArGrCrArArGr UrCrGrArArArGrArCrArCrGrArUrGrUrGrArGrCrCrU rArU*mU*mU*mU
1144	MG102-48 TRAC D12	mG*mC*mC*rArCrUrUrUrCrArGrGrArGrGrArGrAr UrUrCrGrGrGrUrUrUrCrArArUrCrArArCrGrGrArAr ArCrGrGrUrCrCrGrGrUrUrGrArArArArGrArGrCrArU rCrGrGrUrCrUrGrArArGrGrArUrGrCrArCrUrCrCrGr GrGrArUrArGrGrGrCrArGrUrCrCrCrGrGrCrUrCrUrU rGrCrUrGrUrUrUrCrCrCrGrGrUrArArGrArCrCrUrC rGrGrArArGrCrArArGrUrCrCrUrUrCrArGrCrArArGr UrCrGrArArArGrArCrArCrGrArUrGrUrGrArGrCrCrU rArU*mU*mU*mU
1145	MG102-36 TRAC D12	GCCACTTTCAGGAGGAGGATTCGG
1146	MG102-36 TRAC F1	GACCCTGCCGTGTACCAGCTGAGA
1147	MG102-36 TRAC H6	TTGAAGTCCATAGACCTCATGTCT
1148	MG102-39 TRAC F4	GCTGCCCTTACCTGGGCTGGGGAA
1149	MG102-39 TRAC A9	TCTTGTTTTTACAGATACGAACCT
1150	MG102-39 TRAC G11	GGCCACTTTCAGGAGGAGGATTCG

【 0 2 3 4 】

10

20

30

40

50

【表 5 - 9】

配列番号	名称	配列
1151	MG102-39 TRAC C11	CAGCCGCAGCGTCATGAGCAGATT
1152	MG102-39 TRAC B6	CCAGGCCACAGCACTGTTGCTCTT
1153	MG102-39 TRAC B5	GTCTTCTGGAATAATGCTGTTGTT
1154	MG102-39 TRAC G9	GATTGGGTTCCGAATCCTCCTCCT
1155	MG102-39 TRAC D1	ATTCTGATGTGTATATCACAGACA
1156	MG102-39 TRAC B11	ACAGCCGCAGCGTCATGAGCAGAT
1157	MG102-39 TRAC D4	AAAGCTGCCCTTACCTGGGCTGGG
1158	MG102-39 TRAC F2	CAACAGTGCTGTGGCCTGGAGCAA
1159	MG102-39 TRAC G1	GCTAGACATGAGGTCTATGGACTT
1160	MG102-42 TRAC D10	GTTCCGAATCCTCCTCCTGAAAGT
1161	MG102-42 TRAC D12	GCCACTTTCAGGAGGAGGATTTCGG
1162	MG102-42 TRAC E12	CAGGAGGAGGATTTCGGAACCCAAT
1163	MG102-45 TRAC B1	TGTCCACAGATATCCAGAACCCT
1164	MG102-45 TRAC C11	CTTCAAGGCCCTCACCTCAGCTG
1165	MG102-48 TRAC A1	TCCTCTTGTCCACAGATATCCAG
1166	MG102-48 TRAC D12	GCCACTTTCAGGAGGAGGATTTCGG
1167	MG33-34 TRAC F6	mA*mC*mC*rCrGrGrCrCrArCrUrUrCrArGrGrArGrGrCrUrUrCrArCrUrCrUrArGrCrGrArArArGrCrUrArGrArGrUrGrArArArGrArArGrCrCrCrArGrGrCrGrCrUrGrCrUrCrCrArGrUrCrCrUrCrGrCrCrGrArUrGrUrArArCrCrCrArGrCrArUrCrGrGrCrArCrCrUrArGrGrUrGrUrArGrGrCrArGrCrCrCrGrCrArGrGrCrCrGrGrUrArCrUrCrGrGrArCrCrCrGrGrCrArArArGrGrCrArArGrGrGrUrU*mG*mG*mU
1168	MG33-34 TRAC E6	mU*mA*mA*rArCrCrCrGrGrCrCrArCrUrUrCrArGrGrCrUrUrCrArCrUrCrUrArGrCrGrArArArGrCrUrArGrArGrUrGrArArArGrArArGrCrCrCrArGrGrCrGrCrUrGrCrUrCrCrArGrUrCrCrUrCrGrCrCrGrArUrGrUrArArCrCrCrArGrCrArUrCrGrGrCrArCrCrUrArGrGrUrGrUrArGrGrCrArGrCrCrCrGrCrArGrGrCrCrGrGrUrArCrUrCrGrGrArCrCrCrGrGrCrArArArGrGrGrCrArArGrGrGrUrU*mG*mG*mU
1169	MG33-34 TRAC F6	ACCCGGCCACTTTCAGGAGG
1170	MG33-34 TRAC E6	TAAACCCGGCCACTTTCAGG

r=天然のリボース塩基、m=2'-O メチル修飾塩基、F=2'フルオロ修飾塩基、*=ホスホロチオエート結合

10

20

30

40

【0235】

実施例 16 - SMART HEARO 酵素は活性ヌクレアーゼである

SMART HEARO ガイド RNA のインシリコ予測

【0236】

新規の SMART HEARO ヌクレアーゼと関連付けられたガイド (HEARO) RNA を特定するために、305 個の推定エフェクターの 5' UTR 領域に対応するヌクレオチド配列を抽出した。これらの 5' UTR ヌクレオチド配列を、パラメータ `mafft-xinsi` (<https://mafft.cbrc.jp/alignment/software/>) を用いて MAFFT (その全体が参照により本明細書に組み込まれる

50

、Kato h K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013, 30 (4), 772-780) で整列させ、保存領域を使用して、HEARO RNA境界を描写した(図31)。加えて、活性SMART HEAROヌクレアーゼのHEARO RNA配列を使用して共分散モデルを生成し、新規のSMART HEAROヌクレアーゼをコードするゲノム断片中の追加のHEARO RNAを予測した。共分散モデルは、mafft-xinsi(<https://mafft.cbrc.jp/alignment/software/>)との活性HEARO RNA配列の多重配列アラインメント(MSA)から構築される。MSAの二次構造をRNAalifold(Vienna Package、<https://www.tbi.univie.ac.at/RNA/>)で決定し、共分散モデルをInfernalパッケージ(<http://eddylab.org/infernal/>)で構築した。候補SMART HEAROヌクレアーゼを含有するコンティグ及び305個の5'UTR領域を、Infernalコマンド「cmsearch」を用いて共分散モデルを使用して検索した。5'UTRアラインメントから、及び新規候補の共分散モデルから予測されるHEARO RNAを、インビトロで試験した。

10

【0237】

インビトロTAM決定アッセイ

5'末端に標的化スパーサーを有するsgRNA(HEARO RNA)を、アセンブリPCRを介して構築し、SPRIビーズで精製するか、又は遺伝子断片(IDT)として注文し、次いで、インビトロ転写(IVT、HiScribe T7キット、New England Biolabs)し、短いRNA転写物について製造業者が推奨するプロトコルに従って実施した。RNA反応物をMonarch RNAキットで洗浄し、Tapestation(Agilent)を介して純度をチェックした。切断及びTAM決定アッセイを、PURExpress(New England Biolabs)を用いて実施した。簡潔に述べると、タンパク質をE.coliに対してコドン最適化し、T7プロモーター及びC末端Hisタグを有するベクターにクローニングした。遺伝子を、それぞれ、T7プロモーター配列及びターミネーター配列の上流及び下流150bpのプライマー結合部位でPCR増幅した。このPCR産物を、5nMの最終濃度でPURExpress(New England Biolabs)に加え、37℃で2時間発現させた。PURExpressの5倍希釈、5nMの8N PAMプラスミドライブラリー、及びPAMライブラリーを標的化する50nMのsgRNAを用いて、10mMのTris pH7.5、100mMのNaCl、及び10mMのMgCl₂中で切断反応を組み立てた。PURExpress反応からの切断産物を、SPRIビーズ(AMPure Beckman Coulter又はHighPrep Sigma-Aldrich)を用いたクリーンアップを介して回収した。DNAは、クレノウ断片及びdNTP(New England Biolabs)の添加を介して平滑化した。平滑末端生成物を、100倍過剰な二本鎖アダプター配列とライゲーションし、NGSライブラリーの調製のための鋳型として使用し、そこからPAM要件を配列分析から決定した。生のNGS読み取りを、Phred品質スコア>20によってフィルタリングした。PAMに隣接する骨格から文書化されたDNA配列を表す14~24bpを参照として使用して、PAM近位領域を見つけ、隣接する8bpを推定標的隣接モチーフ(TAM)として特定した。TAMとライゲーションアダプターとの間の距離も、各読み取りに対して測定した。参照配列又はアダプター配列と正確に一致しなかった読み取りは除外された。TAM配列を、最も頻度の高い切断部位±2bpを有するTAMのみが分析に含まれるように、切断部位頻度によってフィルタリングした。TAMのフィルタリングされたリストを使用して、Logomakerを使用して配列ロゴを生成した(Tareen, A. & Kinney, J. B. Logomaker: beautiful sequence logos in Python. Bioinformatics 2020, 36, 2272-22

20

30

40

50

74)。

【0238】

SMART II (HEARO) エフェクターは、標的化された dsDNA 切断のために、その 5' UTR 領域にコードされるガイド (HEARO) RNA と相互作用する短い (約 400~600 aa 長さ) ヌクレアーゼである (図 32A 及び図 32D)。ほとんどの場合、SMART HEARO 系は、CRISPR 関連ではないが、少数の SMART HEARO ヌクレアーゼが、CRISPR と関連付けられ得る。例えば、SMART HEARO MG35-463 (配列番号 530) は、CRISPR アレイの下流でコードされる (図 32B)。共分散モデルから予測される HEARO ガイド RNA の 5' 末端は、アレイの最後の CRISPR リピートと重複し (図 32B 及び図 32F、sg3)、完全標的化単一ガイド RNA が、アレイの最後のスパーサー及び最後の反復、並びに HEARO RNA を含むことを示唆している (図 32F、sg3)。更に、この候補の共分散モデルは、CRISPR アレイの上流にあり、かつそれとは無関係に、第 2 の HEARO RNA を予測した (図 32B 及び図 32E、sg2)。CRISPR 関連 SMART HEARO 系の別の例は、MG35-556 (配列番号 659) であり (図 32C)、HEARO RNA は、エフェクターの 5' UTR 領域にコードされ、これは、CRISPR リピートのうちの 1 つに相補的なアンチリピートを含有する (図 32C)。これは、二重ガイド RNA ガイド HEARO 系の一例を表し、1 つの CRISPR リピート (標的化スパーサーをその 5' 末端で担持する可能性が高い) は、HEARO RNA の 5' 末端にアニールして、他の単一ガイド HEARO RNA に似た構造にフォールディングする (図 32G)。

10

20

【0239】

切断活性について試験したとき、多くの SMART HEARO ヌクレアーゼは、インビトロ TAM 決定アッセイで活性であり、それらのうちのいくつかは、複数の sgRNA 設計で活性であった (図 33 の A~図 33 の C)。MG35-104 (配列番号 128)、HEARO MG35-463 (配列番号 530)、及び MG35-518 (配列番号 621) は、強いバンド強度の読み出しによって示されるように、最も活性なヌクレアーゼの 1 つであった (図 33)。更に、SMART HEARO MG35-463 (配列番号 530) は、ガイド RNA が 65% ペアワイズヌクレオチド同一性のみを共有するにもかかわらず、その CRISPR 関連 (配列番号 1237) 及び CRISPR 非依存性 (配列番号 1236) の HEARO RNA の両方と機能的である (図 32D、図 32E、及び図 33 の C)。活性 MG35 候補は、多様な TAM を認識し、TAM モチーフから 5 又は 7 位に対する切断選択性を示す (図 34)。

30

【0240】

実施例 17-SMART HEARO 酵素は効率的なヌクレアーゼである

インビトロ切断アッセイ

【0241】

MG35 ヌクレアーゼを、インビトロ転写／翻訳 (IVTT) (New England Biolabs) を使用して、37℃で2時間発現させた。転写を、ヌクレアーゼをコードする線形 DNA 鋳型上の T7 プロモーターによって駆動した。ガイド RNA をインビトロで別々に転写し、選択された濃度、通常、4 µM で 0.4 の間で IVTT 混合物に加えた。インビトロ切断反応を、1×エフェクター緩衝液 (10 mM の Tris-HCl pH 7.5、100 mM の NaCl、10 mM の MgCl₂) 又は 1×New England Biolabs 2.1 緩衝液 (10 mM の Tris-HCl pH 7.9、50 mM の NaCl、10 mM の MgCl₂、100 µg/ml の BSA) 中の 10 µL の反応体積で、5 nM の超らせん状 DNA に 3 µL の RNP 試料を加えることによって実施した。反応物を 37℃で1時間インキュベートし、次いで、0.2 µg の RNase A (New England Biolabs) を加えることによってクエンチし、続いて、37℃で20分間インキュベートした。次いで、4 単位のプロテイナーゼ K (New England Biolabs) を追加し、続いて、55℃で30分間インキュベート

40

50

した。反応を、分析及び可視化のために製造業者が推奨する指示に従って、D5000 Tape stationキット (Agilent) を使用してキャピラリー電気泳動によって分析した。成功した切断は、結果的に、線形 dsDNA に切断される超らせん状の 2200 bp DNA をもたらす。

【0242】

活性ガイドRNA及びTAM認識モチーフを特定した後、SMART HEAROヌクレアーゼを、ヌクレアーゼのそのガイドRNAとのインビトロ転写／翻訳の共発現を介してインビトロ切断効率について試験し、その後、ガイドRNAによって標的化されたスパーサーと、TAM／PAM濃縮スクリーニングで特定されたTAMとを含有する標的プラスミドとのインキュベーションを行った。切断は、切断されていない産物（超らせん状）の切断された直鎖状DNAへの移行によって測定される（図35A及び図35B）。結果は、MG35-104（配列番号128）が、他の活性SMART HEAROヌクレアーゼと比較して、dsDNA切断において非常に効率的であることを示す（図35A及び図35B）。

10

【0243】

実施例18-SMART HEAROガイド操作

いくつかの活性SMART IIヌクレアーゼのガイドRNAは、1つ以上のポリ-T領域を含有し（4つ以上のT塩基を連続して）（図36A）、これは、転写効率を制限し得る。候補当たり3つのポリT変異体sgRNAを、インビトロの切断活性について設計及び試験し、それらの活性を、それらの天然ガイドRNAを用いた候補の活性と比較した（図36A及び図36B）。結果は、MG35-94が変異体ガイドM2及びM3と活性であるが、一方で、MG35-104が全3つのガイド変異M1-M3と活性であることを示し、ガイドM3が他のガイドと比較して最も高い活性を保持する。MG35-518は、試験した全3つの突然変異体と活性であるが、M1が最も高い活性を示す（図36B）。

20

【0244】

30

40

50

【表 6 - 1】

表 6 - 実施例 18 で試験されたバリエントガイド RNA

配列 番号	説明	配列
1258	MG35-94_M2 単一 ガイド RNA	(N20)GUA AUCGUCCA UAAAUAACUUAGGCAACUAAGUAGUU UAAGGUUACCCGCUUUGGUUCUUCGGAACUCCGUUAGGGG CGAAAAUAUAGGUACUCUUGGAUGCAUCUCCAGUCCGAGA CUCUACGGGGAACGAUUAACAGGUCUGAUGGAAAGGCCA GUGUCGUUCCA UUUAAAACCGCUUUCUAACA UUAAGCUAG GAAACCAUUAACUCGCGCAAGCGAAGAU AUGUAACAAUUU
1259	MG35-94_M3 単一 ガイド RNA	(N20)GUA AUCGUCCA UAAAUAACUUAGGCAACUAAGUAGAU UAAGGUUACCCGCUUUGGUUCUUCGGAACUCCGUUAGGGG CGAAAAUAUAGGUACUCUUGGAUGCAUCUCCAGUCCGAGA CUCUACGGGGAACGAUUAACAGGUCUGAUGGAAAGGCCA GUGUCGUUCCA UUUAAAACCGCUUUCUAACA UUAAGCUAG GAAACCAUUAACUCGCGCAAGCGAAGAU AUGUAACAAUUU
1249	MG35-104_M1 単一 ガイド RNA	(N20)GUAAGGAACCCCGUAGCUAAAGCUAGGGGCUAUUCAU CCCCGUCCCUUCGGGCGGGCUUAGAUAGCCGAACCUUACC AGCCUAAGACCUUCGAGGUCUACGUAAUUAAGGUCACGAU ACCUAUCAAUGCGUCGCUAGUUGUUAAGCUCUAUCGCUGGU UGUUAACAUCUGUAAUGGGUUAAGGAAGUGCAAUCAGCC CAACAAGCCUUGAAUACA UUGGCGAAGCGAACAUCACCCA GCAAUGGAGUCCUCAAUCA
1250	MG35-104_M2 単一 ガイド RNA	(N20)GUAAGGAACCCCGUAGCUAAAGCUAGGGGCUAUUCAU CCCCGUCCCUUCGGGCGGGCUUAGUAAGCCGAACCUUACC AGCCUAAGACCUUCGAGGUCUACGUAAUUAAGGUCACGAU ACCUAUCAAUGCGUCGCUAGUUGUUAAGCUCUAUCGCUGGU UGUUAACAUCUGUAAUGGGUUAAGGAAGUGCAAUCAGCC CAACAAGCCUUGAAUACA UUGGCGAAGCGAACAUCACCCA GCAAUGGAGUCCUCAAUCA
1251	MG35-104_M3 単一 ガイド RNA	(N20)GUAAGGAACCCCGUAGCUAAAGCUAGGGGCUAUUCAU CCCCGUCCCUUCGGGCGGGCUUAGAUAGCCGAACCUUACC AGCCUAAGACCUUCGAGGUCUACGUUCUCAAGGUCACGAU ACCUAUCAAUGCGUCGCUAGUUGUUAAGCUCUAUCGCUGGU UGUUAACAUCUGUAAUGGGUUAAGGAAGUGCAAUCAGCC CAACAAGCCUUGAGAACA UUGGCGAAGCGAACAUCACCCA GCAAUGGAGUCCUCAAUCA
1252	MG35-518_M1 単一 ガイド RNA	(N20)AUCAAUAACCAACCCACUAAGUGGGCGGAUUGCUUGA CUCUUAUACAAUGAGUUGAGAAACCGUGAUUGAUUAGCCU CAGUUAUAAACUACGUUAUUUGUAAAUAUUAAGGUACCGU CGGAUGUCCGCCUAGUCCUACGCGCUACGCUUUAUUAUUA AACAGUUCUGAUUGGUAGGAACAGUGUAAUAAAGAUUAA AACUACAAGAUAACAUUGGCGAAGGCAAUAAAGGGUUGU UUAUACCCGCUUACCGCAUUAUUAAACAU

【 0 2 4 5 】

10

20

30

40

50

【表 6 - 2】

配列番号	説明	配列
1253	MG35-518_M2 単一ガイド RNA	(N20)AUCAAUAACCAACCCACUAAGUGGGCGGAUUGCUUGA CUCUAUUACAAUGAGUUGAGAAACCGUGAUUGAUUAGCCU CAGUUAUAAACUACGUUAAUUGUAAAUAUUAUAGGUACCGU CGGAUGUCCGCCUAGUCCUACGCGCUACGCUUUUAUUAUUA AACAGUUCUGAUUGGUAGGAACAGUGUAAUAAAGAUUAUA AACUACAAGAUAAAUUGGCGAAGGCAAUAAAGGGUUUGU UUAUACCCGCUUACCGCAUUAUUAAUAAACAU
1254	MG35-518_M3 単一ガイド RNA	(N20)AUCAAUAACCAACCCACUAAGUGGGCGGAUUGCUUGA CUCUGUUACAAUGAGUUGAGAAACCGUGAUUGAUUAGCCU CAGUUAUAAACUACGUUAAUUGUAAAUAUUAUAGGUACCGU CGGAUGUCCGCCUAGUCCUACGCGCUACGCUUUUAUUAUUA AACAGUUCUGAUUGGUAGGAACAGUGUAAUAAAGAUUAUA AACUACAAGAUAAAUUGGCGAAGGCAAUAAAGGGUUUGU UUAUACCCGCUUACCGCAUUAUUAAUAAACAU
1255	MG35-553_M1 単一ガイド RNA	(N20)GUCAACUACCCACGACUAAAGUCGCGGGCUUGUAAUA AGGAUAGUGCUAUGUACUAGCCUUAUUCAGCCCGGUUGAC UAGCCUAAGCACCAAUUGUGCUACGUUAUGCAGGAAAUAG GUACCUCGGGAUGUACAGCCUAGUCCCGGGCUCUACGGUA UGAGGUUAAACAGCUCUGACGGGUAGGAGCAGUGCUUCAU GCGUUAACCCUGCAAUAACAUUGGCGAAGGCUAACUAAC GGAUGCUGCAUCCGGCUUACAGCAAUAAUGCAGCAGAAAA
1256	MG35-553_M2 単一ガイド RNA	(N20)GUCAACUACCCACGACUAAAGUCGCGGGCUUGUAAUA AGGAUAGUGCUAUGUACUAGCCUUAUUCAGCCCGGUUGAC UAGCCUAAGCACCAAUUGUGCUACGUUAUGCAGGAAAUAG GUACCUCGGGAUGUACAGCCUAGUCCCGGGCUCUACGGUA UGAGGUUAAACAGCUCUGACGGGUAGGAGCAGUGCUUCAU GCGUUAACCCUGCAAUAACAUUGGCGAAGGCUAACUAAC GGAUGCUGCAUCCGGCUUACAGCAAUAAUGCAGCAGAAAA
1257	MG35-553_M3 単一ガイド RNA	(N20)GUCAACUACCCACGACUAAAGUCGCGGGCUUGUAAUU AGGAUAGUGCUAUGUACUAGCCUAAUUCAGCCCGGUUGAC UAGCCUAAGCACCAAUUGUGCUACGUUAUGCAGGAAAUAG GUACCUCGGGAUGUACAGCCUAGUCCCGGGCUCUACGGUA UGAGGUUAAACAGCUCUGACGGGUAGGAGCAGUGCUUCAU GCGUUAACCCUGCAAUAACAUUGGCGAAGGCUAACUAAC GGAUGCUGCAUCCGGCUUACAGCAAUAAUGCAGCAGAAAA

【0246】

実施例 19 - 新規の SMART イヌクレアーゼの計算再構成

新規配列のインシリコ再構成

【0247】

SMART イヌクレアーゼの更なる多様性を生成しようとして、祖先配列再構成アルゴリズムを使用して、分岐ヌクレアーゼ配列を再構成した。祖先配列再構成は、既存のタンパク質配列、及びそれらの間の推定される関係を使用して、古代の現在は絶滅したタンパク質の配列を再構成する計算技術である (Harms, M. & Thornton J. W. Analyzing protein structure and function using ancestral gene reconstruction. Current Opinion in Structural Biology 2010, 20, 360-366)。この技術を使用して、SMART イ MG34 ファミリーの新規配列を計算再構成した。この分析では、190 個の SMART イタンパク質配列を、パラメータ L-INS-i 又は G-INS-i を用いた MAFFT を使用して整列させ (Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: impr

ovements in performance and usability. *Mol Biol Evol.* 2013, 30 (4), 772–780)、系統樹を、Fasttree (Price, M. N., Dehal, P. S., and Arkin, A. P. FastTree 2—Approximately Maximum-Likelihood Trees for Large Alignments. *PLoS ONE* 2010, 5 (3), e9490) 又はRAxML (Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 2014, 30 (9), 1312–1313) のいずれかを使用して構築した (図 37)。系統樹を、SpCas9 及び SaCas9 を使用して根付かせた。配列再構成を、PAML 4.8 の codeml パッケージ (Yang, Z. PAML 4: a program package for phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution* 2007, 24, 1586–1591) を使用して行い、全 4 つのアラインメントの組み合わせ及び系統樹構築方法に適用し、系統における不確実性を考慮した。各再構成ノードに対して挿入及び欠失を手動で特定した。

10

【0248】

インビトロ PAM 決定アッセイ

候補タンパク質を *E. coli* に対してコドン最適化し、T7 プロモーター及び C 末端 His タグを有するベクターにクローニングした。遺伝子を、それぞれ、T7 プロモーター配列及びターミネーター配列の上流及び下流 150 bp のプライマー結合部位で PCR 増幅した。この PCR 産物を、5 nM の最終濃度で PURExpress (New England Biolabs) に加え、37 °C で 2 時間発現させた。PURExpress の 5 倍希釈、5 nM の 8 N PAM プラスミドライブラリー、及び PAM ライブラリーを標的化する 50 nM の sgRNA を用いて、10 mM の Tris pH 7.5、100 mM の NaCl、及び 10 mM の MgCl₂ 中で切断反応を組み立てた。PURExpress 反応からの切断産物を、SPRI ビーズ (AMPure Beckman Coulter 又は HighPrep Sigma-Aldrich) を用いたクリーンアップを介して回収した。DNA は、クレノウ断片及び dNTP (New England Biolabs) の添加を介して平滑化した。平滑末端生成物を、100 倍過剰な二本鎖アダプター配列とライゲーションし、NGS ライブラリーの調製のための鋳型として使用し、そこから PAM 要件を配列分析から決定した。生の NGS 読み取りを、Phred 品質スコア > 20 によってフィルタリングした。PAM に隣接する骨格から文書化された DNA 配列を表す 14 ~ 24 bp を参照として使用して、PAM 近位領域を見つけ、隣接する 8 bp を推定 PAM として特定した。PAM とライゲーションアダプターとの間の距離も、各読み取りに対して測定された。参照配列又はアダプター配列と正確に一致しなかった読み取りは除外された。PAM 配列を、最も頻度の高い切断部位 ± 2 bp を有する PAM のみが分析に含まれるように、切断部位頻度によってフィルタリングした。PAM のフィルタリングされたリストを使用して、Logomaker を使用して配列ロゴを生成した (Tareen, A. & Kinney, J. B. Logomaker: beautiful sequence logos in Python. *Bioinformatics* 2020, 36, 2272–2274)。

20

30

40

【0249】

MG34 ファミリーの 6 つの配列を、高い信頼性で再構成し (表 5 及び図 37)、触媒及び結合ドメインを、複数の配列アラインメント及び 3D 構造予測から確認した (図 38)。

【0250】

50

【表 7】

表 7ー計算由来 MG34 候補対 SMART Iヌクレアーゼ MG34-1 の比較

候補	配列番号	平均支持率	MG34-1 %同一性	長さ (アミノ酸)
MG34-26	1313	0.73	66	768
MG34-27	1314	0.93	78	745
MG34-28	1315	0.75	70	765
MG34-29	1316	0.92	79	748
MG34-30	1317	0.74	79	766
MG34-31	1318	0.74	72	768

平均支持率の値は、再構成された配列の平均確率を 0～1 のスケールで示す。支持率の値 > 0.7 は、再構成された配列における高い信頼性を示す。

【0251】

構造間の主な差異は認識ローブにあり、これは、これらの再構成されたエフェクターが MG34-1 と同様のヌクレアーゼ活性を示し得ることを示唆している。新たに再構成された候補に対する強い支持を考慮すると、3つの活性 MG34ヌクレアーゼからのガイド RNAを用いた PAM濃縮アッセイにおいて、インビトロ切断活性について6つの新規のヌクレアーゼを試験した。MG34-1 sgRNA 1 (配列番号613)、MG34-9 sgRNA 1 (配列番号615)、及びMG34-16 sgRNA 1 (配列番号616)。新規のヌクレアーゼ MG34-27 (配列番号1314) 及びMG34-29 (配列番号1316) は、およそ180bpで予想される切断バンドによって示されるように、全3つの試験されたsgRNAと活性であった(図39)。これらの新規のヌクレアーゼによって標的化されるPAMは、3' nRRである可能性が高く、nGGは、最も一般的に認識されるPAMである(図40)。結果は、新たに再構成されたヌクレアーゼが、他の活性MG34ヌクレアーゼと比較して、より緩和されたPAM認識を有し(例えば、MG34-1は、3' nGG PAMを認識する)、PAMから6～9位で柔軟な切断選好を有することを示す(図40)。

【0252】

10

20

30

40

50

表 8 - 1 配列表に含まれない本明細書で言及される追加のタンパク質及び核酸配列のリスト

I.

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG33 活性エフェクター	1	MG33-1 エフェクター	タンパク質	未知	未培養生物	
MG34 活性エフェクター	2	MG34-1 エフェクター	タンパク質	未知	未培養生物	
MG34 エフェクター	3	MG34-2 エフェクター	タンパク質	未知	未培養生物	
MG34 エフェクター	4	MG34-3 エフェクター	タンパク質	未知	未培養生物	
MG34 エフェクター	5	MG34-4 エフェクター	タンパク質	未知	未培養生物	
MG34 エフェクター	6	MG34-5 エフェクター	タンパク質	未知	未培養生物	
MG34 エフェクター	7	MG34-6 エフェクター	タンパク質	未知	未培養生物	
MG34 エフェクター	8	MG34-7 エフェクター	タンパク質	未知	未培養生物	
MG34 エフェクター	9	MG34-8 エフェクター	タンパク質	未知	未培養生物	
MG34 エフェクター	10	MG34-9 エフェクター	タンパク質	未知	未培養生物	
MG34 エフェクター	11	MG34-10 エフェクター	タンパク質	未知	未培養生物	
MG34 エフェクター	12	MG34-11 エフェクター	タンパク質	未知	未培養生物	
MG34 エフェクター	13	MG34-12 エフェクター	タンパク質	未知	未培養生物	
MG34 エフェクター	14	MG34-13 エフェクター	タンパク質	未知	未培養生物	
MG34 エフェクター	15	MG34-14 エフェクター	タンパク質	未知	未培養生物	
MG34 エフェクター	16	MG34-15 エフェクター	タンパク質	未知	未培養生物	
MG34 エフェクター	17	MG34-16 エフェクター	タンパク質	未知	未培養生物	
MG34 エフェクター	18	MG34-17 エフェクター	タンパク質	未知	未培養生物	
MG34 エフェクター	19	MG34-18 エフェクター	タンパク質	未知	未培養生物	

【表 8 - 2】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG34 エフエクター	20	MG34-19 エフエクター	タンパク質	未知	未培養生物	
MG34 エフエクター	21	MG34-20 エフエクター	タンパク質	未知	未培養生物	
MG34 エフエクター	22	MG34-21 エフエクター	タンパク質	未知	未培養生物	
MG34 エフエクター	23	MG34-22 エフエクター	タンパク質	未知	未培養生物	
MG34 エフエクター	24	MG34-23 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	25	MG35-1 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	26	MG35-2 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	27	MG35-3 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	28	MG35-4 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	29	MG35-5 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	30	MG35-6 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	31	MG35-7 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	32	MG35-8 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	33	MG35-9 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	34	MG35-10 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	35	MG35-11 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	36	MG35-12 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	37	MG35-13 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	38	MG35-14 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	39	MG35-15 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	40	MG35-16 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	41	MG35-17 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	42	MG35-18 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	43	MG35-19 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	44	MG35-20 エフエクター	タンパク質	未知	未培養生物	

【 0 2 5 4 】

10

20

30

40

50

【表 8 - 3】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 エフエクター	45	MG35-21 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	46	MG35-22 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	47	MG35-23 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	48	MG35-24 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	49	MG35-25 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	50	MG35-26 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	51	MG35-27 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	52	MG35-28 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	53	MG35-29 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	54	MG35-30 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	55	MG35-31 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	56	MG35-32 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	57	MG35-33 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	58	MG35-34 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	59	MG35-35 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	60	MG35-36 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	61	MG35-37 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	62	MG35-38 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	63	MG35-39 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	64	MG35-40 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	65	MG35-41 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	66	MG35-42 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	67	MG35-43 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	68	MG35-44 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	69	MG35-45 エフエクター	タンパク質	未知	未培養生物	

【 0 2 5 5 】

10

20

30

40

50

【表 8 - 4】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 エフエクトー	70	MG35-46 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	71	MG35-47 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	72	MG35-48 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	73	MG35-49 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	74	MG35-50 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	75	MG35-51 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	76	MG35-52 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	77	MG35-53 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	78	MG35-54 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	79	MG35-55 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	80	MG35-56 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	81	MG35-57 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	82	MG35-58 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	83	MG35-59 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	84	MG35-60 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	85	MG35-61 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	86	MG35-62 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	87	MG35-63 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	88	MG35-64 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	89	MG35-65 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	90	MG35-66 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	91	MG35-67 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	92	MG35-68 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	93	MG35-69 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	94	MG35-70 エフエクトー	タンパク質	未知	未培養生物	

【 0 2 5 6 】

10

20

30

40

50

【表 8 - 5】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 エフエクター	95	MG35-71 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	96	MG35-72 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	97	MG35-73 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	98	MG35-74 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	99	MG35-75 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	100	MG35-76 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	101	MG35-77 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	102	MG35-78 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	103	MG35-79 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	104	MG35-80 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	105	MG35-81 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	106	MG35-82 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	107	MG35-83 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	108	MG35-84 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	109	MG35-85 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	110	MG35-86 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	111	MG35-87 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	112	MG35-88 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	113	MG35-89 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	114	MG35-90 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	115	MG35-91 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	116	MG35-92 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	117	MG35-93 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	118	MG35-94 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	119	MG35-95 エフエクター	タンパク質	未知	未培養生物	

【 0 2 5 7 】

10

20

30

40

50

【表 8 - 6】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 エフエクター	120	MG35-96 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	121	MG35-97 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	122	MG35-98 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	123	MG35-99 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	124	MG35-100 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	125	MG35-101 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	126	MG35-102 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	127	MG35-103 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	128	MG35-104 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	129	MG35-105 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	130	MG35-106 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	131	MG35-107 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	132	MG35-108 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	133	MG35-109 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	134	MG35-110 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	135	MG35-111 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	136	MG35-112 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	137	MG35-113 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	138	MG35-114 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	139	MG35-115 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	140	MG35-116 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	141	MG35-117 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	142	MG35-118 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	143	MG35-119 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	144	MG35-120 エフエクター	タンパク質	未知	未培養生物	

【 0 2 5 8 】

10

20

30

40

50

【表 8 - 7】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 エフエクトター	145	MG35-121 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	146	MG35-122 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	147	MG35-123 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	148	MG35-124 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	149	MG35-125 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	150	MG35-126 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	151	MG35-127 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	152	MG35-128 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	153	MG35-129 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	154	MG35-130 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	155	MG35-131 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	156	MG35-132 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	157	MG35-133 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	158	MG35-134 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	159	MG35-135 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	160	MG35-136 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	161	MG35-137 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	162	MG35-138 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	163	MG35-139 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	164	MG35-140 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	165	MG35-141 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	166	MG35-142 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	167	MG35-143 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	168	MG35-144 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	169	MG35-146 エフエクトター	タンパク質	未知	未培養生物	

【 0 2 5 9 】

10

20

30

40

50

【表 8 - 8】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 エフエクター	170	MG35-147 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	171	MG35-148 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	172	MG35-149 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	173	MG35-150 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	174	MG35-151 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	175	MG35-152 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	176	MG35-153 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	177	MG35-154 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	178	MG35-155 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	179	MG35-156 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	180	MG35-157 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	181	MG35-158 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	182	MG35-159 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	183	MG35-160 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	184	MG35-161 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	185	MG35-162 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	186	MG35-163 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	187	MG35-164 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	188	MG35-165 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	189	MG35-166 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	190	MG35-167 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	191	MG35-168 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	192	MG35-169 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	193	MG35-170 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	194	MG35-171 エフエクター	タンパク質	未知	未培養生物	

【 0 2 6 0 】

10

20

30

40

50

【表 8 - 9】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 エフェクター	195	MG35-172 エフェクター	タンパク質	未知	未培養生物	
MG35 エフェクター	196	MG35-173 エフェクター	タンパク質	未知	未培養生物	
MG35 エフェクター	197	MG35-174 エフェクター	タンパク質	未知	未培養生物	
MG35 エフェクター	198	MG35-175 エフェクター	タンパク質	未知	未培養生物	
MG33 活性エフェクター-tracrRNA 配列	199	MG33-1 tracrRNA 配列	ヌクレオチド	人工配列		
MG34 活性エフェクター-tracrRNA 配列	200	MG34-1 tracrRNA 配列	ヌクレオチド	人工配列		
推定 sgRNA	201	推定 MG33-1 sgRNA	ヌクレオチド	人工配列		
推定 sgRNA	202	推定 MG34-1 sgRNA	ヌクレオチド	人工配列		
推定 sgRNA	203	推定 MG34-1 sgRNA	ヌクレオチド	人工配列		
標的	204	試験標的配列	ヌクレオチド	人工配列		
NLS	205	SV40 NLS	タンパク質	シミアン空胞形成ウイルス 40T		
NLS	206	ヌクレオプラスミン双節型 (bipartite) NLS	タンパク質	ヒト		
NLS	207	c-myc NLS	タンパク質	ヒト		
NLS	208	c-myc NLS	タンパク質	ヒト		
NLS	209	hnRNP A1 M9 NLS	タンパク質	マウス		
NLS	210	インポーチン-アルファ IBB ドメイン NLS	タンパク質	ヒト		
NLS	211	筋腫 T タンパク質 NLS	タンパク質	マウスポリオマウイルス		
NLS	212	筋腫 T タンパク質 NLS	タンパク質	マウスポリオマウイルス		
NLS	213	p53 NLS	タンパク質	ヒト		

【 0 2 6 1 】

10

20

30

40

50

【表 8 - 1 0】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
NLS	214	マウス c-abl IV NLS	タンパク質	マウス		
NLS	215	インフルエンザウイルス NS1 NLS	タンパク質	インフルエンザウイルス NS1		
NLS	216	インフルエンザウイルス NS1 NLS	タンパク質	インフルエンザウイルス NS1		
NLS	217	肝炎ウイルスデルタ抗原 NLS	タンパク質	肝炎ウイルスデルタ		
NLS	218	マウス Mx1 タンパク質 NLS	タンパク質	マウス		
NLS	219	ヒトポリ (ADP-リボース) ポリメラーゼ NLS	タンパク質	ヒト		
NLS	220	ステロイドホルモン受容体グルココルチコイド NLS	タンパク質	ヒト		
MG35 エフェクター	221	MG35-4 エフェクター	タンパク質	未知	未培養生物	
MG35 エフェクター	222	MG35-419 エフェクター	タンパク質	未知	未培養生物	
MG35 エフェクター	223	MG35-420 エフェクター	タンパク質	未知	未培養生物	
MG35 エフェクター	224	MG35-421 エフェクター	タンパク質	未知	未培養生物	
MG35 エフェクター	225	MG35-176 エフェクター	タンパク質	未知	未培養生物	
MG35 エフェクター	226	MG35-177 エフェクター	タンパク質	未知	未培養生物	
MG35 エフェクター	227	MG35-178 エフェクター	タンパク質	未知	未培養生物	
MG35 エフェクター	228	MG35-179 エフェクター	タンパク質	未知	未培養生物	
MG35 エフェクター	229	MG35-180 エフェクター	タンパク質	未知	未培養生物	
MG35 エフェクター	230	MG35-181 エフェクター	タンパク質	未知	未培養生物	
MG35 エフェクター	231	MG35-183 エフェクター	タンパク質	未知	未培養生物	

【 0 2 6 2 】

10

20

30

40

50

【表 8 - 1 1】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 エフエクトター	232	MG35-184 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	233	MG35-185 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	234	MG35-186 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	235	MG35-187 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	236	MG35-188 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	237	MG35-189 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	238	MG35-190 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	239	MG35-191 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	240	MG35-192 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	241	MG35-193 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	242	MG35-194 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	243	MG35-195 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	244	MG35-196 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	245	MG35-197 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	246	MG35-198 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	247	MG35-199 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	248	MG35-200 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	249	MG35-201 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	250	MG35-202 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	251	MG35-203 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	252	MG35-204 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	253	MG35-205 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	254	MG35-206 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	255	MG35-207 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	256	MG35-208 エフエクトター	タンパク質	未知	未培養生物	

【 0 2 6 3 】

10

20

30

40

50

【表 8 - 1 2】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 エフ エク ター	257	MG35-209 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	258	MG35-210 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	259	MG35-211 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	260	MG35-212 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	261	MG35-213 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	262	MG35-214 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	263	MG35-215 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	264	MG35-216 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	265	MG35-217 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	266	MG35-218 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	267	MG35-219 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	268	MG35-220 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	269	MG35-221 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	270	MG35-222 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	271	MG35-223 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	272	MG35-224 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	273	MG35-225 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	274	MG35-226 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	275	MG35-227 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	276	MG35-228 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	277	MG35-229 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	278	MG35-230 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	279	MG35-231 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	280	MG35-232 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	281	MG35-233 エフ エク ター	タンパク質	未知	未培養生物	

【 0 2 6 4 】

10

20

30

40

50

【表 8 - 1 3】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 エフエクトター	282	MG35-234 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	283	MG35-235 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	284	MG35-236 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	285	MG35-237 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	286	MG35-238 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	287	MG35-239 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	288	MG35-240 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	289	MG35-241 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	290	MG35-242 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	291	MG35-243 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	292	MG35-244 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	293	MG35-245 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	294	MG35-246 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	295	MG35-247 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	296	MG35-248 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	297	MG35-249 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	298	MG35-250 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	299	MG35-251 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	300	MG35-252 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	301	MG35-253 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	302	MG35-254 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	303	MG35-255 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	304	MG35-256 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	305	MG35-257 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	306	MG35-258 エフエクトター	タンパク質	未知	未培養生物	

【 0 2 6 5】

10

20

30

40

50

【表 8 - 1 4】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 エフエクター	307	MG35-259 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	308	MG35-260 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	309	MG35-261 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	310	MG35-262 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	311	MG35-263 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	312	MG35-264 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	313	MG35-265 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	314	MG35-266 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	315	MG35-267 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	316	MG35-268 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	317	MG35-269 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	318	MG35-270 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	319	MG35-271 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	320	MG35-272 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	321	MG35-273 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	322	MG35-274 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	323	MG35-275 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	324	MG35-276 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	325	MG35-277 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	326	MG35-278 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	327	MG35-279 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	328	MG35-280 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	329	MG35-281 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	330	MG35-282 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	331	MG35-283 エフエクター	タンパク質	未知	未培養生物	

【 0 2 6 6 】

10

20

30

40

50

【表 8 - 1 5】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 エフエックター	332	MG35-284 エフエックター	タンパク質	未知	未培養生物	
MG35 エフエックター	333	MG35-285 エフエックター	タンパク質	未知	未培養生物	
MG35 エフエックター	334	MG35-286 エフエックター	タンパク質	未知	未培養生物	
MG35 エフエックター	335	MG35-287 エフエックター	タンパク質	未知	未培養生物	
MG35 エフエックター	336	MG35-288 エフエックター	タンパク質	未知	未培養生物	
MG35 エフエックター	337	MG35-289 エフエックター	タンパク質	未知	未培養生物	
MG35 エフエックター	338	MG35-290 エフエックター	タンパク質	未知	未培養生物	
MG35 エフエックター	339	MG35-291 エフエックター	タンパク質	未知	未培養生物	
MG35 エフエックター	340	MG35-292 エフエックター	タンパク質	未知	未培養生物	
MG35 エフエックター	341	MG35-293 エフエックター	タンパク質	未知	未培養生物	
MG35 エフエックター	342	MG35-294 エフエックター	タンパク質	未知	未培養生物	
MG35 エフエックター	343	MG35-295 エフエックター	タンパク質	未知	未培養生物	
MG35 エフエックター	344	MG35-296 エフエックター	タンパク質	未知	未培養生物	
MG35 エフエックター	345	MG35-297 エフエックター	タンパク質	未知	未培養生物	
MG35 エフエックター	346	MG35-298 エフエックター	タンパク質	未知	未培養生物	
MG35 エフエックター	347	MG35-299 エフエックター	タンパク質	未知	未培養生物	
MG35 エフエックター	348	MG35-300 エフエックター	タンパク質	未知	未培養生物	
MG35 エフエックター	349	MG35-301 エフエックター	タンパク質	未知	未培養生物	
MG35 エフエックター	350	MG35-302 エフエックター	タンパク質	未知	未培養生物	
MG35 エフエックター	351	MG35-303 エフエックター	タンパク質	未知	未培養生物	
MG35 エフエックター	352	MG35-304 エフエックター	タンパク質	未知	未培養生物	
MG35 エフエックター	353	MG35-305 エフエックター	タンパク質	未知	未培養生物	
MG35 エフエックター	354	MG35-307 エフエックター	タンパク質	未知	未培養生物	
MG35 エフエックター	355	MG35-308 エフエックター	タンパク質	未知	未培養生物	
MG35 エフエックター	356	MG35-309 エフエックター	タンパク質	未知	未培養生物	

【 0 2 6 7 】

10

20

30

40

50

【表 8 - 1 6】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 エフ エク ター	357	MG35-310 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	358	MG35-311 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	359	MG35-312 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	360	MG35-313 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	361	MG35-314 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	362	MG35-315 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	363	MG35-316 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	364	MG35-317 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	365	MG35-318 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	366	MG35-319 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	367	MG35-320 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	368	MG35-321 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	369	MG35-322 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	370	MG35-323 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	371	MG35-324 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	372	MG35-325 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	373	MG35-326 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	374	MG35-327 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	375	MG35-328 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	376	MG35-329 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	377	MG35-330 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	378	MG35-331 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	379	MG35-333 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	380	MG35-334 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	381	MG35-335 エフ エク ター	タンパク質	未知	未培養生物	

【 0 2 6 8 】

10

20

30

40

50

【表 8 - 1 7】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 エフエクトー	382	MG35-336 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	383	MG35-337 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	384	MG35-338 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	385	MG35-339 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	386	MG35-340 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	387	MG35-341 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	388	MG35-342 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	389	MG35-343 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	390	MG35-344 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	391	MG35-345 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	392	MG35-346 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	393	MG35-347 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	394	MG35-348 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	395	MG35-349 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	396	MG35-350 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	397	MG35-351 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	398	MG35-352 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	399	MG35-353 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	400	MG35-354 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	401	MG35-355 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	402	MG35-356 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	403	MG35-357 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	404	MG35-358 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	405	MG35-359 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	406	MG35-360 エフエクトー	タンパク質	未知	未培養生物	

【 0 2 6 9 】

10

20

30

40

50

【表 8 - 1 8】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 エフエクター	407	MG35-361 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	408	MG35-362 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	409	MG35-363 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	410	MG35-364 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	411	MG35-365 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	412	MG35-366 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	413	MG35-367 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	414	MG35-368 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	415	MG35-369 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	416	MG35-370 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	417	MG35-371 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	418	MG35-372 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	419	MG35-373 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	420	MG35-374 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	421	MG35-375 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	422	MG35-376 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	423	MG35-377 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	424	MG35-378 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	425	MG35-379 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	426	MG35-384 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	427	MG35-385 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	428	MG35-386 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	429	MG35-387 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	430	MG35-388 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	431	MG35-389 エフエクター	タンパク質	未知	未培養生物	

【 0 2 7 0 】

10

20

30

40

50

【表 8 - 1 9】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 エフエクトー	432	MG35-390 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	433	MG35-391 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	434	MG35-392 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	435	MG35-393 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	436	MG35-394 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	437	MG35-395 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	438	MG35-396 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	439	MG35-397 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	440	MG35-398 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	441	MG35-399 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	442	MG35-400 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	443	MG35-401 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	444	MG35-402 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	445	MG35-403 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	446	MG35-404 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	447	MG35-405 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	448	MG35-406 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	449	MG35-408 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	450	MG35-409 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	451	MG35-410 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	452	MG35-411 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	453	MG35-412 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	454	MG35-413 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	455	MG35-414 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	456	MG35-415 エフエクトー	タンパク質	未知	未培養生物	

【 0 2 7 1 】

10

20

30

40

50

【表 8 - 2 0】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 エフェクター	457	MG35-416 エフェクター	タンパク質	未知	未培養生物	
MG35 エフェクター	458	MG35-417 エフェクター	タンパク質	未知	未培養生物	
MG35 エフェクター	459	MG35-418 エフェクター	タンパク質	未知	未培養生物	
MG35 エフェクター tracrRNA 配列	460	MG35-4 tracrRNA 配列	ヌクレオチド	人工配列		
推定 tracrRNA	461	推定 MG35-3 tracrRNA	ヌクレオチド	人工配列		
反復	462	MG35-3 反復	ヌクレオチド	人工配列		
MG33 エフェクター	463	MG33-2 エフェクター	タンパク質	未知	未培養生物	
MG33 エフェクター	464	MG33-3 エフェクター	タンパク質	未知	未培養生物	
MG33 エフェクター	465	MG33-4 エフェクター	タンパク質	未知	未培養生物	
MG33 エフェクター	466	MG33-5 エフェクター	タンパク質	未知	未培養生物	
MG33 エフェクター	467	MG33-6 エフェクター	タンパク質	未知	未培養生物	
MG33 エフェクター	468	MG33-7 エフェクター	タンパク質	未知	未培養生物	
MG33 エフェクター	469	MG33-8 エフェクター	タンパク質	未知	未培養生物	
MG33 エフェクター	470	MG33-9 エフェクター	タンパク質	未知	未培養生物	
MG33 エフェクター	471	MG33-10 エフェクター	タンパク質	未知	未培養生物	
MG33 エフェクター	472	MG33-11 エフェクター	タンパク質	未知	未培養生物	
MG33 エフェクター	473	MG33-12 エフェクター	タンパク質	未知	未培養生物	
MG33 エフェクター	474	MG33-13 エフェクター	タンパク質	未知	未培養生物	
MG33 エフェクター	475	MG33-14 エフェクター	タンパク質	未知	未培養生物	
MG33 エフェクター	476	MG33-15 エフェクター	タンパク質	未知	未培養生物	
MG33 エフェクター	477	MG33-16 エフェクター	タンパク質	未知	未培養生物	
MG33 エフェクター	478	MG33-17 エフェクター	タンパク質	未知	未培養生物	
MG33 エフェクター	479	MG33-18 エフェクター	タンパク質	未知	未培養生物	
MG33 エフェクター	480	MG33-19 エフェクター	タンパク質	未知	未培養生物	

【 0 2 7 2 】

10

20

30

40

50

【表 8 - 2 1】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG33 エフエクター	481	MG33-20 エフエクター	タンパク質	未知	未培養生物	
MG33 エフエクター	482	MG33-21 エフエクター	タンパク質	未知	未培養生物	
MG33 エフエクター	483	MG33-22 エフエクター	タンパク質	未知	未培養生物	
MG33 エフエクター	484	MG33-23 エフエクター	タンパク質	未知	未培養生物	
MG33 エフエクター	485	MG33-24 エフエクター	タンパク質	未知	未培養生物	
MG33 エフエクター	486	MG33-26 エフエクター	タンパク質	未知	未培養生物	
MG34 エフエクター	487	MG34-23 エフエクター	タンパク質	未知	未培養生物	
MG34 エフエクター	488	MG34-24 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	489	MG35-422 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	490	MG35-423 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	491	MG35-424 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	492	MG35-425 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	493	MG35-426 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	494	MG35-427 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	495	MG35-428 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	496	MG35-429 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	497	MG35-430 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	498	MG35-431 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	499	MG35-432 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	500	MG35-433 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	501	MG35-434 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	502	MG35-435 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	503	MG35-436 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	504	MG35-437 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	505	MG35-438 エフエクター	タンパク質	未知	未培養生物	

【 0 2 7 3 】

10

20

30

40

50

【表 8 - 2 2】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 エフエクトター	506	MG35-439 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	507	MG35-440 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	508	MG35-441 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	509	MG35-442 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	510	MG35-443 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	511	MG35-444 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	512	MG35-445 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	513	MG35-446 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	514	MG35-447 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	515	MG35-448 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	516	MG35-449 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	517	MG35-450 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	518	MG35-451 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	519	MG35-452 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	520	MG35-453 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	521	MG35-454 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	522	MG35-455 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	523	MG35-456 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	524	MG35-457 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	525	MG35-458 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	526	MG35-459 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	527	MG35-460 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	528	MG35-461 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	529	MG35-462 エフエクトター	タンパク質	未知	未培養生物	
MG35 エフエクトター	530	MG35-463 エフエクトター	タンパク質	未知	未培養生物	

【 0 2 7 4 】

10

20

30

40

50

【表 8 - 2 3】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 エフエクトー	531	MG35-464 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	532	MG35-465 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	533	MG35-466 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	534	MG35-467 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	535	MG35-468 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	536	MG35-469 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	537	MG35-470 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	538	MG35-471 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	539	MG35-472 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	540	MG35-473 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	541	MG35-474 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	542	MG35-475 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	543	MG35-476 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	544	MG35-477 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	545	MG35-478 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	546	MG35-479 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	547	MG35-480 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	548	MG35-481 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	549	MG35-482 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	550	MG35-483 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	551	MG35-484 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	552	MG35-485 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	553	MG35-486 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	554	MG35-487 エフエクトー	タンパク質	未知	未培養生物	
MG35 エフエクトー	555	MG35-488 エフエクトー	タンパク質	未知	未培養生物	

【 0 2 7 5 】

10

20

30

40

50

【表 8 - 2 4】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 エフエケター	556	MG35-489 エフエケター	タンパク質	未知	未培養生物	
MG35 エフエケター	557	MG35-490 エフエケター	タンパク質	未知	未培養生物	
MG35 エフエケター	558	MG35-491 エフエケター	タンパク質	未知	未培養生物	
MG35 エフエケター	559	MG35-492 エフエケター	タンパク質	未知	未培養生物	
MG35 エフエケター	560	MG35-493 エフエケター	タンパク質	未知	未培養生物	
MG35 エフエケター	561	MG35-494 エフエケター	タンパク質	未知	未培養生物	
MG35 エフエケター	562	MG35-495 エフエケター	タンパク質	未知	未培養生物	
MG35 エフエケター	563	MG35-496 エフエケター	タンパク質	未知	未培養生物	
MG35 エフエケター	564	MG35-497 エフエケター	タンパク質	未知	未培養生物	
MG35 エフエケター	565	MG35-498 エフエケター	タンパク質	未知	未培養生物	
MG35 エフエケター	566	MG35-499 エフエケター	タンパク質	未知	未培養生物	
MG35 エフエケター	567	MG35-500 エフエケター	タンパク質	未知	未培養生物	
MG35 エフエケター	568	MG35-501 エフエケター	タンパク質	未知	未培養生物	
MG35 エフエケター	569	MG35-502 エフエケター	タンパク質	未知	未培養生物	
MG35 エフエケター	570	MG35-503 エフエケター	タンパク質	未知	未培養生物	
MG35 エフエケター	571	MG35-504 エフエケター	タンパク質	未知	未培養生物	
MG35 エフエケター	572	MG35-505 エフエケター	タンパク質	未知	未培養生物	
MG35 エフエケター	573	MG35-506 エフエケター	タンパク質	未知	未培養生物	
MG35 エフエケター	574	MG35-507 エフエケター	タンパク質	未知	未培養生物	
MG35 エフエケター	575	MG35-508 エフエケター	タンパク質	未知	未培養生物	
MG35 エフエケター	576	MG35-509 エフエケター	タンパク質	未知	未培養生物	
MG35 エフエケター	577	MG35-510 エフエケター	タンパク質	未知	未培養生物	
MG35 エフエケター	578	MG35-511 エフエケター	タンパク質	未知	未培養生物	
MG35 エフエケター	579	MG35-512 エフエケター	タンパク質	未知	未培養生物	
MG35 エフエケター	580	MG35-513 エフエケター	タンパク質	未知	未培養生物	

【 0 2 7 6 】

10

20

30

40

50

【表 8 - 2 5】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG102 エフェクタ —	581	MG102-1 エフェクター	タンパク質	未知	未培養生物	
MG102 エフェクタ —	582	MG102-2 エフェクター	タンパク質	未知	未培養生物	
MG102 エフェクタ —	583	MG102-3 エフェクター	タンパク質	未知	未培養生物	
MG102 エフェクタ —	584	MG102-4 エフェクター	タンパク質	未知	未培養生物	
MG102 エフェクタ —	585	MG102-5 エフェクター	タンパク質	未知	未培養生物	
MG102 エフェクタ —	586	MG102-6 エフェクター	タンパク質	未知	未培養生物	
MG102 エフェクタ —	587	MG102-7 エフェクター	タンパク質	未知	未培養生物	
MG102 エフェクタ —	588	MG102-8 エフェクター	タンパク質	未知	未培養生物	
MG102 エフェクタ —	589	MG102-9 エフェクター	タンパク質	未知	未培養生物	
MG102 エフェクタ —	590	MG102-10 エフェクター	タンパク質	未知	未培養生物	
MG102 エフェクタ —	591	MG102-11 エフェクター	タンパク質	未知	未培養生物	
MG102 エフェクタ —	592	MG102-12 エフェクター	タンパク質	未知	未培養生物	
MG102 エフェクタ —	593	MG102-13 エフェクター	タンパク質	未知	未培養生物	
MG102 エフェクタ —	594	MG102-14 エフェクター	タンパク質	未知	未培養生物	

【 0 2 7 7 】

10

20

30

40

50

【表 8 - 2 6】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG102 エフエクタ —	595	MG102-15 エフエクタ	タンパク質	未知	未培養生物	
MG102 エフエクタ —	596	MG102-16 エフエクタ	タンパク質	未知	未培養生物	
MG102 エフエクタ —	597	MG102-17 エフエクタ	タンパク質	未知	未培養生物	
MG102 エフエクタ —	598	MG102-18 エフエクタ	タンパク質	未知	未培養生物	
MG102 エフエクタ —	599	MG102-19 エフエクタ	タンパク質	未知	未培養生物	
MG102 エフエクタ —	600	MG102-20 エフエクタ	タンパク質	未知	未培養生物	
MG102 エフエクタ —	601	MG102-21 エフエクタ	タンパク質	未知	未培養生物	
MG102 エフエクタ —	602	MG102-22 エフエクタ	タンパク質	未知	未培養生物	
MG102 エフエクタ —	603	MG102-23 エフエクタ	タンパク質	未知	未培養生物	
MG102 エフエクタ —	604	MG102-24 エフエクタ	タンパク質	未知	未培養生物	
MG102 エフエクタ —	605	MG102-25 エフエクタ	タンパク質	未知	未培養生物	
MG102 エフエクタ —	606	MG102-27 エフエクタ	タンパク質	未知	未培養生物	
MG102 エフエクタ —	607	MG102-28 エフエクタ	タンパク質	未知	未培養生物	
MG102 エフエクタ —	608	MG102-29 エフエクタ	タンパク質	未知	未培養生物	

【 0 2 7 8 】

10

20

30

40

50

【表 8 - 2 7】

カテゴリ	配列番号	説明	型	生物	他の情報	コメント又は配列
MG102 エフェクタ	609	MG102-30 エフェクタ	タンパク質	未知	未培養生物	
MG102 エフェクタ	610	MG102-31 エフェクタ	タンパク質	未知	未培養生物	
MG102 エフェクタ	611	MG102-32 エフェクタ	タンパク質	未知	未培養生物	
MG102 エフェクタ	612	MG102-33 エフェクタ	タンパク質	未知	未培養生物	
MG34 sgRNA	613	MG34-1 活性エフェクタ —sgRNA1	ヌクレオチド	未知	未培養生物	
MG34 sgRNA	614	MG34-1 活性エフェクタ —sgRNA2	ヌクレオチド	未知	未培養生物	
MG34 sgRNA	615	MG34-9 活性エフェクタ —sgRNA1	ヌクレオチド	未知	未培養生物	
MG34 sgRNA	616	MG34-16 活性エフェク タ—sgRNA1	ヌクレオチド	未知	未培養生物	
MG35 エフェクタ	617	MG35-514 エフェクタ	タンパク質	未知	未培養生物	
MG35 エフェクタ	618	MG35-515 エフェクタ	タンパク質	未知	未培養生物	
MG35 エフェクタ	619	MG35-516 エフェクタ	タンパク質	未知	未培養生物	
MG35 エフェクタ	620	MG35-517 エフェクタ	タンパク質	未知	未培養生物	
MG35 エフェクタ	621	MG35-518 エフェクタ	タンパク質	未知	未培養生物	
MG35 エフェクタ	622	MG35-519 エフェクタ	タンパク質	未知	未培養生物	
MG35 エフェクタ	623	MG35-520 エフェクタ	タンパク質	未知	未培養生物	
MG35 エフェクタ	624	MG35-521 エフェクタ	タンパク質	未知	未培養生物	
MG35 エフェクタ	625	MG35-522 エフェクタ	タンパク質	未知	未培養生物	
MG35 エフェクタ	626	MG35-523 エフェクタ	タンパク質	未知	未培養生物	
MG35 エフェクタ	627	MG35-524 エフェクタ	タンパク質	未知	未培養生物	

【 0 2 7 9 】

10

20

30

40

50

【表 8 - 2 8】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 エフ エク ター	628	MG35-525 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	629	MG35-526 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	630	MG35-527 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	631	MG35-528 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	632	MG35-529 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	633	MG35-530 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	634	MG35-531 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	635	MG35-532 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	636	MG35-533 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	637	MG35-534 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	638	MG35-535 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	639	MG35-536 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	640	MG35-537 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	641	MG35-538 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	642	MG35-539 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	643	MG35-540 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	644	MG35-541 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	645	MG35-542 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	646	MG35-543 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	647	MG35-544 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	648	MG35-545 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	649	MG35-546 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	650	MG35-547 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	651	MG35-548 エフ エク ター	タンパク質	未知	未培養生物	
MG35 エフ エク ター	652	MG35-549 エフ エク ター	タンパク質	未知	未培養生物	

【 0 2 8 0 】

10

20

30

40

50

【表 8 - 2 9】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 エフエクター	653	MG35-550 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	654	MG35-551 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	655	MG35-552 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	656	MG35-553 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	657	MG35-554 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	658	MG35-555 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	659	MG35-556 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	660	MG35-557 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	661	MG35-558 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	662	MG35-559 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	663	MG35-560 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	664	MG35-561 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	665	MG35-562 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	666	MG35-563 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	667	MG35-564 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	668	MG35-565 エフエクター	タンパク質	未知	未培養生物	
MG33 tracrRNA	669	MG33-2 tracrRNA1	ヌクレオチド	人工配列	MG33 tracrRNA	
MG33 tracrRNA	670	MG33-2 tracrRNA2	ヌクレオチド	人工配列	MG33 tracrRNA	
MG33 tracrRNA	671	MG33-3 tracrRNA1	ヌクレオチド	人工配列	MG33 tracrRNA	
MG102 tracrRNA	672	MG102-1 tracrRNA1	ヌクレオチド	人工配列	MG102 tracrRNA	
MG102 tracrRNA	673	MG102-2 tracrRNA1	ヌクレオチド	人工配列	MG102 tracrRNA	
MG35 エフエクター	674	MG35-566 エフエクター	タンパク質	未知	未培養生物	
MG35 エフエクター	675	MG35-567 エフエクター	タンパク質	未知	未培養生物	
MG35 予測される CRISPR 反復	676	MG35-420 予測される CRISPR 反復	ヌクレオチド	未知	未培養生物	

【 0 2 8 1 】

10

20

30

40

50

【表 8 - 3 0】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 推定単一ガイド RNA	677	MG35-1 エフェクター推定単一ガイド RNA1	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA	678	MG35-1 エフェクター推定単一ガイド RNA2	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA	679	MG35-2 エフェクター推定単一ガイド RNA1	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA	680	MG35-3 エフェクター推定単一ガイド RNA1	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA	681	MG35-3 エフェクター推定単一ガイド RNA2	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA	682	MG35-419 エフェクター推定単一ガイド RNA1	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA	683	MG35-419 エフェクター推定単一ガイド RNA2	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA	684	MG35-420 エフェクター推定単一ガイド RNA1	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA	685	MG35-421 エフェクター推定単一ガイド RNA1	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA	686	MG35-102 エフェクター推定単一ガイド RNA1	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	687	MG35-1、MG35-90 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	688	MG35-2、MG35-50、MG35-51 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	

【 0 2 8 2 】

10

20

30

40

50

【表 8 - 3 1】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 推定単一ゲノム RNA コード配列	689	MG35-3、MG35-85 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	690	MG35-32 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	691	MG35-36、MG35-152、MG35-153、MG35-154、MG35-155 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	692	MG35-37 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	693	MG35-38 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	694	MG35-40、MG35-42、MG35-43 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	695	MG35-41 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	696	MG35-44 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	

【 0 2 8 3 】

10

20

30

40

50

【表 8 - 3 2】

カテゴリ	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 推定単一ガイド RNA コード配列	697	MG35-45、MG35-39、MG35-116、MG35-219 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	698	MG35-46 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	699	MG35-48 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	700	MG35-49 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	701	MG35-52 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	702	MG35-53、MG35-54 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	703	MG35-55 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	704	MG35-56、MG35-287 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	705	MG35-57 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	

【 0 2 8 4 】

10

20

30

40

50

【表 8 - 3 3】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 推定単一ガイド RNA コード配列	706	MG35-58、MG35-59 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	707	MG35-60 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	708	MG35-62 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	709	MG35-63 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	710	MG35-65 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	711	MG35-66 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	712	MG35-67、MG35-71 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	713	MG35-68、MG35-64、MG35-69、MG35-70、MG35-75 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	

【 0 2 8 5 】

10

20

30

40

50

【表 8 - 3 4】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 推定単一ガイド RNA コード配列	714	MG35-72 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	715	MG35-73 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	716	MG35-74 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	717	MG35-77 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	718	MG35-78 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	719	MG35-79、MG35-97、MG35-98 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	720	MG35-80、MG35-81 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	721	MG35-82、MG35-95、MG35-96 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	

【 0 2 8 6 】

10

20

30

40

50

【表 8 - 3 5】

カテゴリ	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 推定単一ゲノム RNA コード配列	722	MG35-86 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	723	MG35-87、MG35-88、MG35-89 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	724	MG35-91、MG35-92 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	725	MG35-93、MG35-94 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	726	MG35-99 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	727	MG35-101 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	728	MG35-102 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	729	MG35-103、MG35-104 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	730	MG35-105 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	

【 0 2 8 7 】

10

20

30

40

50

【表 8 - 3 6】

カテゴリ	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 推定単一ガイド RNA コード配列	731	MG35-106 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	732	MG35-107 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	733	MG35-108 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	734	MG35-109 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	735	MG35-110、MG35-112 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	736	MG35-111 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	737	MG35-113 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	738	MG35-114 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	739	MG35-115 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	

【 0 2 8 8 】

10

20

30

40

50

【表 8 - 3 7】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 推定単一ガイド RNA コード配列	740	MG35-116 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	741	MG35-117、MG35-118、MG35-119 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	742	MG35-120 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	743	MG35-121 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	744	MG35-122 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	745	MG35-123、MG35-124 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	746	MG35-125 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	747	MG35-126、MG35-377 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	748	MG35-127 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	

【 0 2 8 9 】

10

20

30

40

50

【表 8 - 3 8】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 推定単一ガイド RNA コード配列	749	MG35-128 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	750	MG35-129 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	751	MG35-130 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	752	MG35-131 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	753	MG35-147 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	754	MG35-148 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	755	MG35-149 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	756	MG35-150 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	757	MG35-151 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	

【 0 2 9 0 】

10

20

30

40

50

【表 8 - 3 9】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 推定単一ゲノム RNA コード配列	758	MG35-152 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	759	MG35-153 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	760	MG35-156、MG35-161 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	761	MG35-157 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	762	MG35-159、MG35-158 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	763	MG35-160 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	764	MG35-165、MG35-166 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	765	MG35-171 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	766	MG35-214 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	

【 0 2 9 1 】

10

20

30

40

50

【表 8 - 4 0】

カテゴリ	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 推定単一ゲイ ド RNA コード配列	767	MG35-217 エフェクター 推定単一ゲイ ド RNA コ ード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲイ ド RNA コード配列	768	MG35-218 エフェクター 推定単一ゲイ ド RNA コ ード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲイ ド RNA コード配列	769	MG35-220 エフェクター 推定単一ゲイ ド RNA コ ード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲイ ド RNA コード配列	770	MG35-221、MG35-222、 MG35-351、MG35-352、 MG35-353 エフェクター 推定単一ゲイ ド RNA コ ード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲイ ド RNA コード配列	771	MG35-223 エフェクター 推定単一ゲイ ド RNA コ ード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲイ ド RNA コード配列	772	MG35-224 エフェクター 推定単一ゲイ ド RNA コ ード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲイ ド RNA コード配列	773	MG35-225 エフェクター 推定単一ゲイ ド RNA コ ード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲイ ド RNA コード配列	774	MG35-226 エフェクター 推定単一ゲイ ド RNA コ ード配列	ヌクレオチド	未知	未培養生物	

【 0 2 9 2 】

10

20

30

40

50

【表 8 - 4 1】

カテゴリ	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 推定単一ガイド RNA コード配列	775	MG35-227 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	776	MG35-228 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	777	MG35-229 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	778	MG35-230 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	779	MG35-231 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	780	MG35-232 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	781	MG35-233 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	782	MG35-235 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	783	MG35-236 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	

【 0 2 9 3 】

10

20

30

40

50

【表 8 - 4 2】

カテゴリ	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 推定単一ゲノム RNA コード配列	784	MG35-238、MG35-237 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	785	MG35-239、MG35-240、MG35-241 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	786	MG35-242 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	787	MG35-243 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	788	MG35-244 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	789	MG35-245 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	790	MG35-246 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	791	MG35-247 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	792	MG35-248 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	

【 0 2 9 4 】

10

20

30

40

50

【表 8 - 4 3】

カテゴリ	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 推定単一ガイド RNA コード配列	793	MG35-249 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	794	MG35-250、MG35-251 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	795	MG35-252 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	796	MG35-253 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	797	MG35-255 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	798	MG35-256 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	799	MG35-257 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	800	MG35-258 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	801	MG35-259 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	

【 0 2 9 5 】

10

20

30

40

50

【表 8 - 4 4】

カテゴリ	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 推定単一ゲノム RNA コード配列	802	MG35-260 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	803	MG35-262、MG35-263 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	804	MG35-266、MG35-270 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	805	MG35-267 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	806	MG35-268 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	807	MG35-269、MG35-264、MG35-265 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	808	MG35-271 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	809	MG35-272 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	810	MG35-273 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	

【 0 2 9 6 】

10

20

30

40

50

【表 8 - 4 5】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 推定単一ガイド RNA コード配列	811	MG35-274 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	812	MG35-275 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	813	MG35-276 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	814	MG35-277 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	815	MG35-278 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	816	MG35-279 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	817	MG35-280 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	818	MG35-281 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	819	MG35-282、MG35-283 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	

【 0 2 9 7 】

10

20

30

40

50

【表 8 - 4 6】

カテゴリ	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 推定単一ガイド RNA コード配列	820	MG35-284 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	821	MG35-285 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	822	MG35-286 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	823	MG35-292、MG35-293 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	824	MG35-296 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	825	MG35-298、MG35-299 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	826	MG35-300 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	827	MG35-302 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	828	MG35-303 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	

【 0 2 9 8 】

10

20

30

40

50

【表 8 - 4 7】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 推定単一ガイド RNA コード配列	829	MG35-305 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	830	MG35-307 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	831	MG35-308 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	832	MG35-309 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	833	MG35-310 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	834	MG35-311 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	835	MG35-312 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	836	MG35-313 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	837	MG35-314、MG35-261 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	

【 0 2 9 9 】

10

20

30

40

50

【表 8 - 4 8】

カテゴリ	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 推定単一ガイド RNA コード配列	838	MG35-315 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	839	MG35-316 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	840	MG35-317 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	841	MG35-318 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	842	MG35-319 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	843	MG35-321 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	844	MG35-322 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	845	MG35-325 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	846	MG35-326 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	

【 0 3 0 0 】

10

20

30

40

50

【表 8 - 4 9】

カテゴリ	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 推定単一ガイド RNA コード配列	847	MG35-327 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	848	MG35-328 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	849	MG35-329 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	850	MG35-330 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	851	MG35-331 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	852	MG35-332、MG35-333、MG35-335 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	853	MG35-334 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	854	MG35-336 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	855	MG35-340 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	

【 0 3 0 1 】

10

20

30

40

50

【表 8 - 5 0】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 推定単一ゲイ ド RNA コード配列	856	MG35-341 エフェクター 推定単一ゲイ ド RNA コ ード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲイ ド RNA コード配列	857	MG35-342 エフェクター 推定単一ゲイ ド RNA コ ード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲイ ド RNA コード配列	858	MG35-343 エフェクター 推定単一ゲイ ド RNA コ ード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲイ ド RNA コード配列	859	MG35-344、MG35-345 エ フェクター推定単一ゲイ ド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲイ ド RNA コード配列	860	MG35-346、MG35-347 エ フェクター推定単一ゲイ ド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲイ ド RNA コード配列	861	MG35-348 エフェクター 推定単一ゲイ ド RNA コ ード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲイ ド RNA コード配列	862	MG35-349 エフェクター 推定単一ゲイ ド RNA コ ード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲイ ド RNA コード配列	863	MG35-350 エフェクター 推定単一ゲイ ド RNA コ ード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲイ ド RNA コード配列	864	MG35-354 エフェクター 推定単一ゲイ ド RNA コ ード配列	ヌクレオチド	未知	未培養生物	

【 0 3 0 2 】

10

20

30

40

50

【表 8 - 5 1】

カテゴリ	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 推定単一ガイド RNA コード配列	865	MG35-355、MG35-356 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	866	MG35-357 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	867	MG35-358 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	868	MG35-359 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	869	MG35-360、MG35-361 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	870	MG35-362 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	871	MG35-363 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	872	MG35-364 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	873	MG35-365 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	

【 0 3 0 3 】

10

20

30

40

50

【表 8 - 5 2】

カテゴリ	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 推定単一ガイド RNA コード配列	874	MG35-366 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	875	MG35-367 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	876	MG35-368、MG35-369 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	877	MG35-370 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	878	MG35-371 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	879	MG35-372、MG35-373 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	880	MG35-374 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	881	MG35-375 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	882	MG35-376 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	

【 0 3 0 4 】

10

20

30

40

50

【表 8 - 5 3】

カテゴリ	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 推定単一ゲノム RNA コード配列	883	MG35-378 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	884	MG35-379 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	885	MG35-384 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	886	MG35-386 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	887	MG35-388 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	888	MG35-419、MG35-339 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	889	MG35-420、MG35-337 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	890	MG35-421、MG35-338 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	891	MG35-422 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	

【 0 3 0 5 】

10

20

30

40

50

【表 8 - 5 4】

カテゴリ	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 推定単一ゲノム RNA コード配列	892	MG35-423 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	893	MG35-424 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	894	MG35-426 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	895	MG35-427 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	896	MG35-428、MG35-436、MG35-437、MG35-457 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	897	MG35-429、MG35-449 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	898	MG35-430 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	899	MG35-431、MG35-442 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	900	MG35-432 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	

【 0 3 0 6 】

10

20

30

40

50

【表 8 - 5 5】

カテゴリ	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 推定単一ガイド RNA コード配列	901	MG35-433、MG35-425 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	902	MG35-434、MG35-455 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	903	MG35-435 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	904	MG35-438 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	905	MG35-439 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	906	MG35-440 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	907	MG35-441、MG35-443 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	908	MG35-444 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	909	MG35-445 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	

【 0 3 0 7 】

10

20

30

40

50

【表 8 - 5 6】

カテゴリ	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 推定単一ガイド RNA コード配列	910	MG35-446、MG35-448、MG35-456 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	911	MG35-447 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	912	MG35-450 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	913	MG35-451 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	914	MG35-452 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	915	MG35-453 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	916	MG35-454 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	917	MG35-458、MG35-523 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	918	MG35-459 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	

【 0 3 0 8 】

10

20

30

40

50

【表 8 - 5 7】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 推定単一ゲノム RNA コード配列	919	MG35-460 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	920	MG35-461 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	921	MG35-462 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	922	MG35-463 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	923	MG35-464 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	924	MG35-465 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	925	MG35-466 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	926	MG35-510 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲノム RNA コード配列	927	MG35-511、MG35-512 エフェクター推定単一ゲノム RNA コード配列	ヌクレオチド	未知	未培養生物	

【 0 3 0 9 】

10

20

30

40

50

【表 8 - 5 8】

カテゴリ	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 推定単一ガイド RNA コード配列	928	MG35-513 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	929	MG35-514 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	930	MG35-515 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	931	MG35-516 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	932	MG35-517 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	933	MG35-518 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	934	MG35-519 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	935	MG35-520 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	936	MG35-521 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	

【 0 3 1 0 】

10

20

30

40

50

【表 8 - 5 9】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 推定単一ガイド RNA コード配列	937	MG35-522、MG35-526、MG35-546、MG35-548 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	938	MG35-524 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	939	MG35-525、MG35-537 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	940	MG35-527 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	941	MG35-528 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	942	MG35-529 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	943	MG35-530 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	944	MG35-531 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	945	MG35-532 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	

【 0 3 1 1 】

10

20

30

40

50

【表 8 - 6 0】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 推定単一ガイド RNA コード配列	946	MG35-533 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	947	MG35-534 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	948	MG35-535 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	949	MG35-536 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	950	MG35-538 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	951	MG35-539 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	952	MG35-540 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	953	MG35-541 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	954	MG35-542 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	

【 0 3 1 2 】

10

20

30

40

50

【表 8 - 6 1】

カテゴリ	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 推定単一ガイド RNA コード配列	955	MG35-543 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	956	MG35-544 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	957	MG35-545 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	958	MG35-547 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	959	MG35-549 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	960	MG35-550 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	961	MG35-552 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	962	MG35-553 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	963	MG35-554 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	

【 0 3 1 3 】

10

20

30

40

50

【表 8 - 6 2】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 推定単一ガイド RNA コード配列	964	MG35-555 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	965	MG35-556 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	966	MG35-557 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	967	MG35-558 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	968	MG35-559 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	969	MG35-560 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	970	MG35-561 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	971	MG35-562 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ガイド RNA コード配列	972	MG35-563 エフェクター推定単一ガイド RNA コード配列	ヌクレオチド	未知	未培養生物	

【 0 3 1 4 】

10

20

30

40

50

【表 8 - 6 3】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 推定単一ゲイ ド RNA コード配列	973	MG35-564 エフェクター 推定単一ゲイ ド RNA コ ード配列	ヌクレオチド	未知	未培養生物	
MG35 推定単一ゲイ ド RNA コード配列	974	MG35-565 エフェクター 推定単一ゲイ ド RNA コ ード配列	ヌクレオチド	未知	未培養生物	
MG143 エフェクタ ー	975	MG143-1 エフェクター	タンパク質	未知	未培養生物	
MG144 エフェクタ ー	976	MG144-1 エフェクター	タンパク質	未知	未培養生物	
MG144 エフェクタ ー	977	MG144-2 エフェクター	タンパク質	未知	未培養生物	
MG144 エフェクタ ー	978	MG144-3 エフェクター	タンパク質	未知	未培養生物	
MG144 エフェクタ ー	979	MG144-4 エフェクター	タンパク質	未知	未培養生物	
MG145 エフェクタ ー	980	MG145-1 エフェクター	タンパク質	未知	未培養生物	
MG33 エフェクタ ー	981	MG33-27 エフェクター	タンパク質	未知	未培養生物	
MG33 エフェクタ ー	982	MG33-28 エフェクター	タンパク質	未知	未培養生物	
MG33 エフェクタ ー	983	MG33-29 エフェクター	タンパク質	未知	未培養生物	
MG33 エフェクタ ー	984	MG33-30 エフェクター	タンパク質	未知	未培養生物	
MG33 エフェクタ ー	985	MG33-31 エフェクター	タンパク質	未知	未培養生物	
MG33 エフェクタ ー	986	MG33-32 エフェクター	タンパク質	未知	未培養生物	
MG33 エフェクタ ー	987	MG33-33 エフェクター	タンパク質	未知	未培養生物	
MG33 エフェクタ ー	988	MG33-34 エフェクター	タンパク質	未知	未培養生物	

【 0 3 1 5 】

10

20

30

40

50

【表 8 - 6 4】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG102 エフェクタ —	989	MG102-35 エフェクター	タンパク質	未知	未培養生物	
MG102 エフェクタ —	990	MG102-36 エフェクター	タンパク質	未知	未培養生物	
MG102 エフェクタ —	991	MG102-37 エフェクター	タンパク質	未知	未培養生物	
MG102 エフェクタ —	992	MG102-38 エフェクター	タンパク質	未知	未培養生物	
MG102 エフェクタ —	993	MG102-39 エフェクター	タンパク質	未知	未培養生物	
MG102 エフェクタ —	994	MG102-40 エフェクター	タンパク質	未知	未培養生物	
MG102 エフェクタ —	995	MG102-41 エフェクター	タンパク質	未知	未培養生物	
MG102 エフェクタ —	996	MG102-42 エフェクター	タンパク質	未知	未培養生物	
MG102 エフェクタ —	997	MG102-43 エフェクター	タンパク質	未知	未培養生物	
MG102 エフェクタ —	998	MG102-44 エフェクター	タンパク質	未知	未培養生物	
MG102 エフェクタ —	999	MG102-45 エフェクター	タンパク質	未知	未培養生物	
MG102 エフェクタ —	1000	MG102-46 エフェクター	タンパク質	未知	未培養生物	
MG102 エフェクタ —	1001	MG102-47 エフェクター	タンパク質	未知	未培養生物	
MG102 エフェクタ —	1002	MG102-48 エフェクター	タンパク質	未知	未培養生物	

【 0 3 1 6 】

10

20

30

40

50

【表 8 - 6 5】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG33 活性エフェクタ ターsgRNA	1003	MG33-3 活性エフェクタ ターsgRNA	ヌクレオチド	人工配列	該当なし	
MG33 活性エフェクタ ターsgRNA	1004	MG33-31 活性エフェクタ ターsgRNA	ヌクレオチド	人工配列	該当なし	
MG33 活性エフェクタ ターsgRNA	1005	MG33-34 活性エフェクタ ターsgRNA	ヌクレオチド	人工配列	該当なし	
MG35 活性エフェクタ ターsgRNA	1006	MG35-1 活性エフェクタ ターsgRNA	ヌクレオチド	人工配列	該当なし	
MG35 活性エフェクタ ターsgRNA	1007	MG35-2 活性エフェクタ ターsgRNA	ヌクレオチド	人工配列	該当なし	
MG35 活性エフェクタ ターsgRNA	1008	MG35-3 活性エフェクタ ターsgRNA	ヌクレオチド	人工配列	該当なし	
MG35 活性エフェクタ ターsgRNA	1009	MG35-4 活性エフェクタ ターsgRNA	ヌクレオチド	人工配列	該当なし	
MG35 活性エフェクタ ターsgRNA	1010	MG35-5 活性エフェクタ ターsgRNA	ヌクレオチド	人工配列	該当なし	
MG35 活性エフェクタ ターsgRNA	1011	MG35-6 活性エフェクタ ターsgRNA	ヌクレオチド	人工配列	該当なし	
MG35 活性エフェクタ ターsgRNA	1012	MG35-102 活性エフェクタ ターsgRNA	ヌクレオチド	人工配列	該当なし	
MG102 活性エフェクタ ターsgRNA	1013	MG102-2 活性エフェクタ ターsgRNA	ヌクレオチド	人工配列	該当なし	
MG102 活性エフェクタ ターsgRNA	1014	MG102-14 活性エフェクタ ターsgRNA	ヌクレオチド	人工配列	該当なし	
MG102 活性エフェクタ ターsgRNA	1015	MG102-35 活性エフェクタ ターsgRNA	ヌクレオチド	人工配列	該当なし	
MG102 活性エフェクタ ターsgRNA	1016	MG102-36 活性エフェクタ ターsgRNA	ヌクレオチド	人工配列	該当なし	

【 0 3 1 7 】

10

20

30

40

50

【表 8 - 6 6】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG102 活性エフェクター-sgRNA	1017	MG102-39 活性エフェクター-sgRNA	ヌクレオチド	人工配列	該当なし	
MG102 活性エフェクター-sgRNA	1018	MG102-42 活性エフェクター-sgRNA	ヌクレオチド	人工配列	該当なし	
MG102 活性エフェクター-sgRNA	1019	MG102-43 活性エフェクター-sgRNA	ヌクレオチド	人工配列	該当なし	
MG102 活性エフェクター-sgRNA	1020	MG102-45 活性エフェクター-sgRNA	ヌクレオチド	人工配列	該当なし	
MG102 活性エフェクター-sgRNA	1021	MG102-47 活性エフェクター-sgRNA	ヌクレオチド	人工配列	該当なし	
MG102 活性エフェクター-sgRNA	1022	MG102-48 活性エフェクター-sgRNA	ヌクレオチド	人工配列	該当なし	
MG33 活性エフェクター-PAM	1023	MG34-1、MG34-9、MG34-16、MG33-3、MG33-31、MG33-34 活性エフェクター-PAM	ヌクレオチド	人工配列	該当なし	
MG33 活性エフェクター-PAM	1024	MG33-31、MG33-34 活性エフェクター-PAM	ヌクレオチド	人工配列	該当なし	
MG33 活性エフェクター-PAM	1025	MG33-31、MG33-34 活性エフェクター-PAM	ヌクレオチド	人工配列	該当なし	
MG34 活性エフェクター-PAM	1026	MG34-1、MG34-9、MG34-16、MG33-3 活性エフェクター-PAM	ヌクレオチド	人工配列	該当なし	
MG34 活性エフェクター-PAM	1027	MG34-1、MG34-9、MG34-16、MG33-3 活性エフェクター-PAM	ヌクレオチド	人工配列	該当なし	

【 0 3 1 8 】

10

20

30

40

50

【表 8 - 6 7】

カテゴリ	配列番号	説明	型	生物	他の情報	コメント又は配列
MG34 活性エフェクターPAM	1028	MG34-1、MG34-9、MG34-16、MG33-3 活性エフェクターPAM	ヌクレオチド	人工配列	該当なし	
MG35 活性エフェクターPAM	1029	MG35-1 活性エフェクターPAM	ヌクレオチド	人工配列	該当なし	
MG35 活性エフェクターPAM	1030	MG35-2 活性エフェクターPAM	ヌクレオチド	人工配列	該当なし	
MG35 活性エフェクターPAM	1031	MG35-3 活性エフェクターPAM	ヌクレオチド	人工配列	該当なし	
MG35 活性エフェクターPAM	1032	MG35-4 活性エフェクターPAM	ヌクレオチド	人工配列	該当なし	
MG35 活性エフェクターPAM	1033	MG35-5 活性エフェクターPAM	ヌクレオチド	人工配列	該当なし	
MG35 活性エフェクターPAM	1034	MG35-102 活性エフェクターPAM	ヌクレオチド	人工配列	該当なし	
MG102 活性エフェクターPAM	1035	MG102-2 活性エフェクターPAM	ヌクレオチド	人工配列	該当なし	
MG102 活性エフェクターPAM	1036	MG102-14、MG102-35、MG102-36、MG102-42、MG102-43、MG102-45、MG102-47、MG102-48 活性エフェクターPAM	ヌクレオチド	人工配列	該当なし	
MG102 活性エフェクターPAM	1037	MG102-14、MG102-35、MG102-36、MG102-42、MG102-43、MG102-45、MG102-47、MG102-48 活性エフェクターPAM	ヌクレオチド	人工配列	該当なし	

【 0 3 1 9 】

10

20

30

40

50

【表 8 - 6 8】

カテゴリ	配列番号	説明	型	生物	他の情報	コメント又は配列
MG102 活性エフェクター-PAM	1038	MG102-14、MG102-35、MG102-36、MG102-42、MG102-43、MG102-45、MG102-47、MG102-48 活性エフェクター-PAM	ヌクレオチド	人工配列	該当なし	
MG102 活性エフェクター-PAM	1039	MG102-39 活性エフェクター-PAM	ヌクレオチド	人工配列	該当なし	
MG102 活性エフェクター-PAM	1040	MG102-14、MG102-35、MG102-36、MG102-42、MG102-43、MG102-45、MG102-47、MG102-48 活性エフェクター-PAM	ヌクレオチド	人工配列	該当なし	
MG102 活性エフェクター-PAM	1041	MG102-14、MG102-35、MG102-36、MG102-42、MG102-43、MG102-45、MG102-47、MG102-48 活性エフェクター-PAM	ヌクレオチド	人工配列	該当なし	
MG102 活性エフェクター-PAM	1042	MG102-14、MG102-35、MG102-36、MG102-42、MG102-43、MG102-45、MG102-47、MG102-48 活性エフェクター-PAM	ヌクレオチド	人工配列	該当なし	
MG102 活性エフェクター-PAM	1043	MG102-14、MG102-35、MG102-36、MG102-42、MG102-43、MG102-45、MG102-47、MG102-48 活性エフェクター-PAM	ヌクレオチド	人工配列	該当なし	

【 0 3 2 0 】

10

20

30

40

50

【表 8 - 6 9】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG102 活性エフェクターPAM	1044	MG102-14、MG102-35、MG102-36、MG102-42、MG102-43、MG102-45、MG102-47、MG102-48 活性エフェクターPAM	ヌクレオチド	人工配列	該当なし	
	MG33 CRISPR 反復	MG33-1 CRISPR 反復	ヌクレオチド	未知	未培養生物	
	MG33 CRISPR 反復	MG33-2 CRISPR 反復	ヌクレオチド	未知	未培養生物	
	MG33 CRISPR 反復	MG33-3 CRISPR 反復	ヌクレオチド	未知	未培養生物	
	MG33 CRISPR 反復	MG33-27 CRISPR 反復	ヌクレオチド	未知	未培養生物	
	MG33 CRISPR 反復	MG33-28 CRISPR 反復	ヌクレオチド	未知	未培養生物	
	MG33 CRISPR 反復	MG33-29 CRISPR 反復	ヌクレオチド	未知	未培養生物	
	MG33 CRISPR 反復	MG33-30 CRISPR 反復	ヌクレオチド	未知	未培養生物	
	MG33 CRISPR 反復	MG33-31 CRISPR 反復	ヌクレオチド	未知	未培養生物	
	MG33 CRISPR 反復	MG33-33 CRISPR 反復	ヌクレオチド	未知	未培養生物	
	MG33 CRISPR 反復	MG33-34 CRISPR 反復	ヌクレオチド	未知	未培養生物	
	MG34 CRISPR 反復	MG34-1、MG34-9 CRISPR 反復	ヌクレオチド	未知	未培養生物	
	MG34 CRISPR 反復	MG34-1、MG34-9 CRISPR 反復	ヌクレオチド	未知	未培養生物	
	MG34 CRISPR 反復	MG34-25 CRISPR 反復	ヌクレオチド	未知	未培養生物	
	MG102 CRISPR 反復	MG102-1 CRISPR 反復	ヌクレオチド	未知	未培養生物	
	MG102 CRISPR 反復	MG102-2 CRISPR 反復	ヌクレオチド	未知	未培養生物	
	MG102 CRISPR 反復	MG102-3 CRISPR 反復	ヌクレオチド	未知	未培養生物	
	MG102 CRISPR 反復	MG102-10 CRISPR 反復	ヌクレオチド	未知	未培養生物	
	MG102 CRISPR 反復	MG102-14 CRISPR 反復	ヌクレオチド	未知	未培養生物	
	MG102 CRISPR 反復	MG102-35 CRISPR 反復	ヌクレオチド	未知	未培養生物	

【 0 3 2 1 】

10

20

30

40

50

【表 8 - 7 0】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG102 CRISPR 反復	1064	MG102-36 CRISPR 反復	ヌクレオチド	未知	未培養生物	
MG102 CRISPR 反復	1065	MG102-38 CRISPR 反復	ヌクレオチド	未知	未培養生物	
MG102 CRISPR 反復	1066	MG102-39 CRISPR 反復	ヌクレオチド	未知	未培養生物	
MG102 CRISPR 反復	1067	MG102-42 CRISPR 反復	ヌクレオチド	未知	未培養生物	
MG102 CRISPR 反復	1068	MG102-43 CRISPR 反復	ヌクレオチド	未知	未培養生物	
MG102 CRISPR 反復	1069	MG102-44 CRISPR 反復	ヌクレオチド	未知	未培養生物	
MG102 CRISPR 反復	1070	MG102-45 CRISPR 反復	ヌクレオチド	未知	未培養生物	
MG102 CRISPR 反復	1071	MG102-47 CRISPR 反復	ヌクレオチド	未知	未培養生物	
MG102 CRISPR 反復	1072	MG102-48 CRISPR 反復	ヌクレオチド	未知	未培養生物	
MG143 CRISPR 反復	1073	MG143-1 CRISPR 反復	ヌクレオチド	未知	未培養生物	
MG144 CRISPR 反復	1074	MG144-1 CRISPR 反復	ヌクレオチド	未知	未培養生物	
MG144 CRISPR 反復	1075	MG144-2 CRISPR 反復	ヌクレオチド	未知	未培養生物	
MG144 CRISPR 反復	1076	MG144-3、MG144-4 CRISPR 反復	ヌクレオチド	未知	未培養生物	
MG144 CRISPR 反復	1077	MG144-3、MG144-4 CRISPR 反復	ヌクレオチド	未知	未培養生物	
MG145 CRISPR 反復	1078	MG145-1 CRISPR 反復	ヌクレオチド	未知	未培養生物	
MG102-2 ヒト TRAC 標的部位	1079	MG102-2 TRAC A1 24nt	ヌクレオチド	人工配列	該当なし	
MG102-2 ヒト TRAC 標的部位	1080	MG102-2 TRAC B1 24nt	ヌクレオチド	人工配列	該当なし	
MG102-2 ヒト TRAC 標的部位	1081	MG102-2 TRAC A1 20nt	ヌクレオチド	人工配列	該当なし	
MG102-2 ヒト TRAC 標的部位	1082	MG102-2 TRAC B1 20nt	ヌクレオチド	人工配列	該当なし	
MG102-2 ヒト TRAC sgRNA	1083	MG102-2 TRAC A1 24nt sgRNA	ヌクレオチド	人工配列	該当なし	

【 0 3 2 2 】

10

20

30

40

50

【表 8 - 7 1】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG102-2 ヒト TRAC sgRNA	1084	MG102-2 TRAC B1 24nt sgRNA	ヌクレオチド	人工配列	該当なし	
MG102-2 ヒト TRAC sgRNA	1085	MG102-2 TRAC A1 20nt sgRNA	ヌクレオチド	人工配列	該当なし	
MG102-2 ヒト TRAC sgRNA	1086	MG102-2 TRAC B1 20nt sgRNA	ヌクレオチド	人工配列	該当なし	
AAVS1 を標的化する MG102-2 ガイド	1087	MG102-2 AAVS1 A5	ヌクレオチド	人工配列		
AAVS1 を標的化する MG102-2 ガイド	1088	MG102-2 AAVS1 H8	ヌクレオチド	人工配列		
AAVS1 を標的化する MG102-2 ガイド	1089	MG102-2 AAVS1 H9	ヌクレオチド	人工配列		
AAVS1 を標的化する MG102-2 ガイド	1090	MG102-2 AAVS1 D11	ヌクレオチド	人工配列		
AAVS1 を標的化する MG102-2 ガイド	1091	MG102-2 AAVS1 E7	ヌクレオチド	人工配列		
AAVS1 を標的化する MG102-2 ガイド	1092	MG102-2 AAVS1 D7	ヌクレオチド	人工配列		
AAVS1 を標的化する MG102-2 ガイド	1093	MG102-2 AAVS1 B7	ヌクレオチド	人工配列		
AAVS1 を標的化する MG102-2 ガイド	1094	MG102-2 AAVS1 D12	ヌクレオチド	人工配列		
AAVS1 を標的化する MG102-2 ガイド	1095	MG102-2 AAVS1 C8	ヌクレオチド	人工配列		
AAVS1 を標的化する MG102-2 ガイド	1096	MG102-2 AAVS1 A8	ヌクレオチド	人工配列		
AAVS1 を標的化する MG102-2 ガイド	1097	MG102-2 AAVS1 G6	ヌクレオチド	人工配列		

【 0 3 2 3 】

10

20

30

40

50

【表 8 - 7 2】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
AAVS1 を標的化する MG102-2 ガイド	1098	MG102-2 AAVS1 E5	ヌクレオチド	人工配列		
AAVS1 を標的化する MG102-2 ガイド	1099	MG102-2 AAVS1 G7	ヌクレオチド	人工配列		
AAVS1 を標的化する MG102-2 ガイド	1100	MG102-2 AAVS1 C3	ヌクレオチド	人工配列		
AAVS1 を標的化する MG102-2 ガイド	1101	MG102-2 AAVS1 E1	ヌクレオチド	人工配列		
AAVS1 を標的化する MG102-2 ガイド	1102	MG102-2 AAVS1 E2	ヌクレオチド	人工配列		
AAVS1 を標的化する MG102-2 ガイド	1103	MG102-2 AAVS1 H6	ヌクレオチド	人工配列		
AAVS1 を標的化する MG102-2 ガイド	1104	MG102-2 AAVS1 H11	ヌクレオチド	人工配列		
AAVS1 標的部位の DNA 配列	1105	MG102-2 AAVS1 A5	ヌクレオチド	人工配列		
AAVS1 標的部位の DNA 配列	1106	MG102-2 AAVS1 H8	ヌクレオチド	人工配列		
AAVS1 標的部位の DNA 配列	1107	MG102-2 AAVS1 H9	ヌクレオチド	人工配列		
AAVS1 標的部位の DNA 配列	1108	MG102-2 AAVS1 D11	ヌクレオチド	人工配列		
AAVS1 標的部位の DNA 配列	1109	MG102-2 AAVS1 E7	ヌクレオチド	人工配列		
AAVS1 標的部位の DNA 配列	1110	MG102-2 AAVS1 D7	ヌクレオチド	人工配列		
AAVS1 標的部位の DNA 配列	1111	MG102-2 AAVS1 B7	ヌクレオチド	人工配列		

【 0 3 2 4】

10

20

30

40

50

【表 8 - 7 3】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
AAVS1 標的部位の DNA 配列	1112	MG102-2 AAVS1 D12	ヌクレオチド	人工配列		
AAVS1 標的部位の DNA 配列	1113	MG102-2 AAVS1 C8	ヌクレオチド	人工配列		
AAVS1 標的部位の DNA 配列	1114	MG102-2 AAVS1 A8	ヌクレオチド	人工配列		
AAVS1 標的部位の DNA 配列	1115	MG102-2 AAVS1 G6	ヌクレオチド	人工配列		
AAVS1 標的部位の DNA 配列	1116	MG102-2 AAVS1 E5	ヌクレオチド	人工配列		
AAVS1 標的部位の DNA 配列	1117	MG102-2 AAVS1 G7	ヌクレオチド	人工配列		
AAVS1 標的部位の DNA 配列	1118	MG102-2 AAVS1 C3	ヌクレオチド	人工配列		
AAVS1 標的部位の DNA 配列	1119	MG102-2 AAVS1 E1	ヌクレオチド	人工配列		
AAVS1 標的部位の DNA 配列	1120	MG102-2 AAVS1 E2	ヌクレオチド	人工配列		
AAVS1 標的部位の DNA 配列	1121	MG102-2 AAVS1 H6	ヌクレオチド	人工配列		
AAVS1 標的部位の DNA 配列	1122	MG102-2 AAVS1 H11	ヌクレオチド	人工配列		
TRAC を標的化する MG102-36 ガイド	1123	MG102-36 TRAC D12	ヌクレオチド	人工配列		
TRAC を標的化する MG102-36 ガイド	1124	MG102-36 TRAC F1	ヌクレオチド	人工配列		
TRAC を標的化する MG102-36 ガイド	1125	MG102-36 TRAC H6	ヌクレオチド	人工配列		

【 0 3 2 5 】

10

20

30

40

50

【表 8 - 7 4】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
TRAC を標的化する MG102-39 ガイド	1126	MG102-39 TRAC F4	ヌクレオチド	人工配列		
TRAC を標的化する MG102-39 ガイド	1127	MG102-39 TRAC A9	ヌクレオチド	人工配列		
TRAC を標的化する MG102-39 ガイド	1128	MG102-39 TRAC G11	ヌクレオチド	人工配列		
TRAC を標的化する MG102-39 ガイド	1129	MG102-39 TRAC C11	ヌクレオチド	人工配列		
TRAC を標的化する MG102-39 ガイド	1130	MG102-39 TRAC B6	ヌクレオチド	人工配列		
TRAC を標的化する MG102-39 ガイド	1131	MG102-39 TRAC B5	ヌクレオチド	人工配列		
TRAC を標的化する MG102-39 ガイド	1132	MG102-39 TRAC G9	ヌクレオチド	人工配列		
TRAC を標的化する MG102-39 ガイド	1133	MG102-39 TRAC D1	ヌクレオチド	人工配列		
TRAC を標的化する MG102-39 ガイド	1134	MG102-39 TRAC B11	ヌクレオチド	人工配列		
TRAC を標的化する MG102-39 ガイド	1135	MG102-39 TRAC D4	ヌクレオチド	人工配列		
TRAC を標的化する MG102-39 ガイド	1136	MG102-39 TRAC F2	ヌクレオチド	人工配列		
TRAC を標的化する MG102-39 ガイド	1137	MG102-39 TRAC G1	ヌクレオチド	人工配列		
TRAC を標的化する MG102-42 ガイド	1138	MG102-42 TRAC D10	ヌクレオチド	人工配列		
TRAC を標的化する MG102-42 ガイド	1139	MG102-42 TRAC D12	ヌクレオチド	人工配列		

【 0 3 2 6 】

10

20

30

40

50

【表 8 - 7 5】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
TRAC を標的化する MG102-42 ガイド	1140	MG102-42 TRAC E12	ヌクレオチド	人工配列		
TRAC を標的化する MG102-45 ガイド	1141	MG102-45 TRAC B1	ヌクレオチド	人工配列		
TRAC を標的化する MG102-45 ガイド	1142	MG102-45 TRAC C11	ヌクレオチド	人工配列		
TRAC を標的化する MG102-48 ガイド	1143	MG102-48 TRAC A1	ヌクレオチド	人工配列		
TRAC を標的化する MG102-48 ガイド	1144	MG102-48 TRAC D12	ヌクレオチド	人工配列		
TRAC 標的部位の DNA 配列	1145	MG102-36 TRAC D12	ヌクレオチド	人工配列		
TRAC 標的部位の DNA 配列	1146	MG102-36 TRAC F1	ヌクレオチド	人工配列		
TRAC 標的部位の DNA 配列	1147	MG102-36 TRAC H6	ヌクレオチド	人工配列		
TRAC 標的部位の DNA 配列	1148	MG102-39 TRAC F4	ヌクレオチド	人工配列		
TRAC 標的部位の DNA 配列	1149	MG102-39 TRAC A9	ヌクレオチド	人工配列		
TRAC 標的部位の DNA 配列	1150	MG102-39 TRAC G11	ヌクレオチド	人工配列		
TRAC 標的部位の DNA 配列	1151	MG102-39 TRAC C11	ヌクレオチド	人工配列		
TRAC 標的部位の DNA 配列	1152	MG102-39 TRAC B6	ヌクレオチド	人工配列		
TRAC 標的部位の DNA 配列	1153	MG102-39 TRAC B5	ヌクレオチド	人工配列		

【 0 3 2 7 】

10

20

30

40

50

【表 8 - 7 6】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
TRAC 標的部位の DNA 配列	1154	MG102-39 TRAC G9	ヌクレオチド	人工配列		
TRAC 標的部位の DNA 配列	1155	MG102-39 TRAC D1	ヌクレオチド	人工配列		
TRAC 標的部位の DNA 配列	1156	MG102-39 TRAC B11	ヌクレオチド	人工配列		
TRAC 標的部位の DNA 配列	1157	MG102-39 TRAC D4	ヌクレオチド	人工配列		
TRAC 標的部位の DNA 配列	1158	MG102-39 TRAC F2	ヌクレオチド	人工配列		
TRAC 標的部位の DNA 配列	1159	MG102-39 TRAC G1	ヌクレオチド	人工配列		
TRAC 標的部位の DNA 配列	1160	MG102-42 TRAC D10	ヌクレオチド	人工配列		
TRAC 標的部位の DNA 配列	1161	MG102-42 TRAC D12	ヌクレオチド	人工配列		
TRAC 標的部位の DNA 配列	1162	MG102-42 TRAC E12	ヌクレオチド	人工配列		
TRAC 標的部位の DNA 配列	1163	MG102-45 TRAC B1	ヌクレオチド	人工配列		
TRAC 標的部位の DNA 配列	1164	MG102-45 TRAC C11	ヌクレオチド	人工配列		
TRAC 標的部位の DNA 配列	1165	MG102-48 TRAC A1	ヌクレオチド	人工配列		
TRAC 標的部位の DNA 配列	1166	MG102-48 TRAC D12	ヌクレオチド	人工配列		
TRAC を標準化する MG33-34 ガイド	1167	MG33-34 TRAC F6	ヌクレオチド	人工配列		

【 0 3 2 8 】

10

20

30

40

50

【表 8 - 7 7】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
TRACを標的化する MG33-34 ガイド	1168	MG33-34 TRAC E6	ヌクレオチド	人工配列		
TRAC標的部位の DNA 配列	1169	MG33-34 TRAC F6	ヌクレオチド	人工配列		
TRAC標的部位の DNA 配列	1170	MG33-34 TRAC E6	ヌクレオチド	人工配列		
エフェクターをコー ドする MG102 遺伝 子座	1171	エフェクターをコードす る MG102-33 遺伝子座	ヌクレオチド	未知	未培養生物	
エフェクターをコー ドする MG35 遺伝子 座	1172	エフェクターをコードす る MG35-3 遺伝子座	ヌクレオチド	未知	未培養生物	
エフェクターをコー ドする MG35 遺伝子 座	1173	エフェクターをコードす る MG35-7 遺伝子座	ヌクレオチド	未知	未培養生物	
エフェクターをコー ドする MG35 遺伝子 座	1174	エフェクターをコードす る MG35-15 遺伝子座	ヌクレオチド	未知	未培養生物	
エフェクターをコー ドする MG35 遺伝子 座	1175	エフェクターをコードす る MG35-20 遺伝子座	ヌクレオチド	未知	未培養生物	
エフェクターをコー ドする MG35 遺伝子 座	1176	エフェクターをコードす る MG35-46 遺伝子座	ヌクレオチド	未知	未培養生物	
エフェクターをコー ドする MG35 遺伝子 座	1177	エフェクターをコードす る MG35-58 遺伝子座	ヌクレオチド	未知	未培養生物	

【 0 3 2 9 】

10

20

30

40

50

【表 8 - 7 8】

カテゴリ	配列番号	説明	型	生物	他の情報	コメント又は配列
エフェクターをコードする MG35 遺伝子座	1178	エフェクターをコードする MG35-59 遺伝子座	ヌクレオチド	未知	未培養生物	
エフェクターをコードする MG35 遺伝子座	1179	エフェクターをコードする MG35-76 遺伝子座	ヌクレオチド	未知	未培養生物	
エフェクターをコードする MG35 遺伝子座	1180	エフェクターをコードする MG35-99 遺伝子座	ヌクレオチド	未知	未培養生物	
エフェクターをコードする MG35 遺伝子座	1181	エフェクターをコードする MG35-100 遺伝子座	ヌクレオチド	未知	未培養生物	
エフェクターをコードする MG35 遺伝子座	1182	エフェクターをコードする MG35-102 遺伝子座	ヌクレオチド	未知	未培養生物	
エフェクターをコードする MG35 遺伝子座	1183	エフェクターをコードする MG35-103 遺伝子座	ヌクレオチド	未知	未培養生物	
エフェクターをコードする MG35 遺伝子座	1184	エフェクターをコードする MG35-104 遺伝子座	ヌクレオチド	未知	未培養生物	
エフェクターをコードする MG35 遺伝子座	1185	エフェクターをコードする MG35-114 遺伝子座	ヌクレオチド	未知	未培養生物	
エフェクターをコードする MG35 遺伝子座	1186	エフェクターをコードする MG35-132 遺伝子座	ヌクレオチド	未知	未培養生物	

【 0 3 3 0 】

10

20

30

40

50

【表 8 - 7 9】

カテゴリ	配列番号	説明	型	生物	他の情報	コメント又は配列
エフェクターをコードする MG35 遺伝子座	1187	エフェクターをコードする MG35-168 遺伝子座	ヌクレオチド	未知	未培養生物	
エフェクターをコードする MG35 遺伝子座	1188	エフェクターをコードする MG35-176 遺伝子座	ヌクレオチド	未知	未培養生物	
エフェクターをコードする MG35 遺伝子座	1189	エフェクターをコードする MG35-177 遺伝子座	ヌクレオチド	未知	未培養生物	
エフェクターをコードする MG35 遺伝子座	1190	エフェクターをコードする MG35-179 遺伝子座	ヌクレオチド	未知	未培養生物	
エフェクターをコードする MG35 遺伝子座	1191	エフェクターをコードする MG35-201 遺伝子座	ヌクレオチド	未知	未培養生物	
エフェクターをコードする MG35 遺伝子座	1192	エフェクターをコードする MG35-231 遺伝子座	ヌクレオチド	未知	未培養生物	
エフェクターをコードする MG35 遺伝子座	1193	エフェクターをコードする MG35-232 遺伝子座	ヌクレオチド	未知	未培養生物	
エフェクターをコードする MG35 遺伝子座	1194	エフェクターをコードする MG35-233 遺伝子座	ヌクレオチド	未知	未培養生物	
エフェクターをコードする MG35 遺伝子座	1195	エフェクターをコードする MG35-237 遺伝子座	ヌクレオチド	未知	未培養生物	

【 0 3 3 1】

10

20

30

40

50

【表 8 - 8 0】

カテゴリ	配列番号	説明	型	生物	他の情報	コメント又は配列
エフェクターをコードする MG35 遺伝子座	1196	エフェクターをコードする MG35-238 遺伝子座	ヌクレオチド	未知	未培養生物	
エフェクターをコードする MG35 遺伝子座	1197	エフェクターをコードする MG35-240 遺伝子座	ヌクレオチド	未知	未培養生物	
エフェクターをコードする MG35 遺伝子座	1198	エフェクターをコードする MG35-291 遺伝子座	ヌクレオチド	未知	未培養生物	
エフェクターをコードする MG35 遺伝子座	1199	エフェクターをコードする MG35-296 遺伝子座	ヌクレオチド	未知	未培養生物	
エフェクターをコードする MG35 遺伝子座	1200	エフェクターをコードする MG35-298 遺伝子座	ヌクレオチド	未知	未培養生物	
エフェクターをコードする MG35 遺伝子座	1201	エフェクターをコードする MG35-299 遺伝子座	ヌクレオチド	未知	未培養生物	
エフェクターをコードする MG35 遺伝子座	1202	エフェクターをコードする MG35-302 遺伝子座	ヌクレオチド	未知	未培養生物	
エフェクターをコードする MG35 遺伝子座	1203	エフェクターをコードする MG35-309 遺伝子座	ヌクレオチド	未知	未培養生物	
エフェクターをコードする MG35 遺伝子座	1204	エフェクターをコードする MG35-323 遺伝子座	ヌクレオチド	未知	未培養生物	

【 0 3 3 2 】

10

20

30

40

50

【表 8 - 8 1】

カテゴリ	配列番号	説明	型	生物	他の情報	コメント又は配列
エフエクターをコードする MG35 遺伝子座	1205	エフエクターをコードする MG35-326 遺伝子座	ヌクレオチド	未知	未培養生物	
エフエクターをコードする MG35 遺伝子座	1206	エフエクターをコードする MG35-337 遺伝子座	ヌクレオチド	未知	未培養生物	
エフエクターをコードする MG35 遺伝子座	1207	エフエクターをコードする MG35-339 遺伝子座	ヌクレオチド	未知	未培養生物	
エフエクターをコードする MG35 遺伝子座	1208	エフエクターをコードする MG35-344 遺伝子座	ヌクレオチド	未知	未培養生物	
エフエクターをコードする MG35 遺伝子座	1209	エフエクターをコードする MG35-345 遺伝子座	ヌクレオチド	未知	未培養生物	
エフエクターをコードする MG35 遺伝子座	1210	エフエクターをコードする MG35-346 遺伝子座	ヌクレオチド	未知	未培養生物	
エフエクターをコードする MG35 遺伝子座	1211	エフエクターをコードする MG35-347 遺伝子座	ヌクレオチド	未知	未培養生物	
エフエクターをコードする MG35 遺伝子座	1212	エフエクターをコードする MG35-348 遺伝子座	ヌクレオチド	未知	未培養生物	
エフエクターをコードする MG35 遺伝子座	1213	エフエクターをコードする MG35-349 遺伝子座	ヌクレオチド	未知	未培養生物	

【 0 3 3 3 】

10

20

30

40

50

【表 8 - 8 2】

カテゴリ	配列番号	説明	型	生物	他の情報	コメント又は配列
エフェクターをコードする MG35 遺伝子座	1214	エフェクターをコードする MG35-350 遺伝子座	ヌクレオチド	未知	未培養生物	
エフェクターをコードする MG35 遺伝子座	1215	エフェクターをコードする MG35-354 遺伝子座	ヌクレオチド	未知	未培養生物	
エフェクターをコードする MG35 遺伝子座	1216	エフェクターをコードする MG35-357 遺伝子座	ヌクレオチド	未知	未培養生物	
エフェクターをコードする MG35 遺伝子座	1217	エフェクターをコードする MG35-358 遺伝子座	ヌクレオチド	未知	未培養生物	
エフェクターをコードする MG35 遺伝子座	1218	エフェクターをコードする MG35-359 遺伝子座	ヌクレオチド	未知	未培養生物	
エフェクターをコードする MG35 遺伝子座	1219	エフェクターをコードする MG35-364 遺伝子座	ヌクレオチド	未知	未培養生物	
エフェクターをコードする MG35 遺伝子座	1220	エフェクターをコードする MG35-366 遺伝子座	ヌクレオチド	未知	未培養生物	
エフェクターをコードする MG35 遺伝子座	1221	エフェクターをコードする MG35-393 遺伝子座	ヌクレオチド	未知	未培養生物	
エフェクターをコードする MG35 遺伝子座	1222	エフェクターをコードする MG35-404 遺伝子座	ヌクレオチド	未知	未培養生物	

【 0 3 3 4 】

10

20

30

40

50

【表 8 - 8 3】

カテゴリ	配列番号	説明	型	生物	他の情報	コメント又は配列
エフェクターをコードする MG35 遺伝子座	1223	エフェクターをコードする MG35-411 遺伝子座	ヌクレオチド	未知	未培養生物	
エフェクターをコードする MG35 遺伝子座	1224	エフェクターをコードする MG35-418 遺伝子座	ヌクレオチド	未知	未培養生物	
エフェクターをコードする MG35 遺伝子座	1225	エフェクターをコードする MG35-419 遺伝子座	ヌクレオチド	未知	未培養生物	
エフェクターをコードする MG35 遺伝子座	1226	エフェクターをコードする MG35-420 遺伝子座	ヌクレオチド	未知	未培養生物	
エフェクターをコードする MG35 遺伝子座	1227	エフェクターをコードする MG35-421 遺伝子座	ヌクレオチド	未知	未培養生物	
エフェクターをコードする MG35 遺伝子座	1228	エフェクターをコードする MG35-423 遺伝子座	ヌクレオチド	未知	未培養生物	
MG35 予測される CRISPR 反復	1229	MG35-463 CRISPR 反復	ヌクレオチド	人工配列		
MG35 予測される CRISPR 反復	1230	MG35-556 CRISPR 反復	ヌクレオチド	人工配列		
MG35 活性エフェクター-sgRNA	1231	MG35-94 sg1 単一ガイド RNA	ヌクレオチド	人工配列		
MG35 活性エフェクター-sgRNA	1232	MG35-94 sg2 単一ガイド RNA	ヌクレオチド	人工配列		
MG35 活性エフェクター-sgRNA	1233	MG35-94 sg3 単一ガイド RNA	ヌクレオチド	人工配列		

【 0 3 3 5】

10

20

30

40

50

【表 8 - 8 4】

カテゴリ	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 活性エフェクター-sgRNA	1234	MG35-104 sg1 単一ガイド RNA	ヌクレオチド	人工配列		
MG35 活性エフェクター-sgRNA	1235	MG35-350 sg3 単一ガイド RNA	ヌクレオチド	人工配列		
MG35 活性エフェクター-sgRNA	1236	MG35-463 sg2 単一ガイド RNA	ヌクレオチド	人工配列		
MG35 活性エフェクター-sgRNA	1237	MG35-463 sg3 単一ガイド RNA	ヌクレオチド	人工配列		
MG35 活性エフェクター-sgRNA	1238	MG35-515 sg2 単一ガイド RNA	ヌクレオチド	人工配列		
MG35 活性エフェクター-sgRNA	1239	MG35-515 sg3 単一ガイド RNA	ヌクレオチド	人工配列		
MG35 活性エフェクター-sgRNA	1240	MG35-517 sg2 単一ガイド RNA	ヌクレオチド	人工配列		
MG35 活性エフェクター-sgRNA	1241	MG35-518 sg1 単一ガイド RNA	ヌクレオチド	人工配列		
MG35 活性エフェクター-sgRNA	1242	MG35-519 sg1 単一ガイド RNA	ヌクレオチド	人工配列		
MG35 活性エフェクター-sgRNA	1243	MG35-519 sg2 単一ガイド RNA	ヌクレオチド	人工配列		
MG35 活性エフェクター-sgRNA	1244	MG35-519 sg3 単一ガイド RNA	ヌクレオチド	人工配列		
MG35 活性エフェクター-sgRNA	1245	MG35-550 sg1 単一ガイド RNA	ヌクレオチド	人工配列		
MG35 活性エフェクター-sgRNA	1246	MG35-553 sg1 単一ガイド RNA	ヌクレオチド	人工配列		
MG35 活性エフェクター-sgRNA	1247	MG35-554 sg3 単一ガイド RNA	ヌクレオチド	人工配列		

【 0 3 3 6 】

10

20

30

40

50

【表 8 - 8 5】

カテゴリ	配列番号	説明	型	生物	他の情報	コメント又は配列
MG35 活性エフェクター-sgRNA	1248	MG35-554 sg4 単一ガイド RNA	ヌクレオチド	人工配列		
MG35 活性エフェクター-sgRNA	1249	MG35-104_M1 単一ガイド RNA	ヌクレオチド	人工配列		
MG35 活性エフェクター-sgRNA	1250	MG35-104_M2 単一ガイド RNA	ヌクレオチド	人工配列		
MG35 活性エフェクター-sgRNA	1251	MG35-104_M3 単一ガイド RNA	ヌクレオチド	人工配列		
MG35 活性エフェクター-sgRNA	1252	MG35-518_M1 単一ガイド RNA	ヌクレオチド	人工配列		
MG35 活性エフェクター-sgRNA	1253	MG35-518_M2 単一ガイド RNA	ヌクレオチド	人工配列		
MG35 活性エフェクター-sgRNA	1254	MG35-518_M3 単一ガイド RNA	ヌクレオチド	人工配列		
MG35 活性エフェクター-sgRNA	1255	MG35-553_M1 単一ガイド RNA	ヌクレオチド	人工配列		
MG35 活性エフェクター-sgRNA	1256	MG35-553_M2 単一ガイド RNA	ヌクレオチド	人工配列		
MG35 活性エフェクター-sgRNA	1257	MG35-553_M3 単一ガイド RNA	ヌクレオチド	人工配列		
MG35 活性エフェクター-sgRNA	1258	MG35-94_M2 単一ガイド RNA	ヌクレオチド	人工配列		
MG35 活性エフェクター-sgRNA	1259	MG35-94_M3 単一ガイド RNA	ヌクレオチド	人工配列		
MG102 エフェクター	1260	MG102-49 エフェクター	タンパク質	未知	未培養生物	
MG102 エフェクター	1261	MG102-50 エフェクター	タンパク質	未知	未培養生物	

【 0 3 3 7 】

10

20

30

40

50

【表 8 - 8 6】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG102 エフェクタ —	1262	MG102-51 エフェクター	タンパク質	未知	未培養生物	
MG102 エフェクタ —	1263	MG102-52 エフェクター	タンパク質	未知	未培養生物	
MG102 エフェクタ —	1264	MG102-53 エフェクター	タンパク質	未知	未培養生物	
MG102 エフェクタ —	1265	MG102-54 エフェクター	タンパク質	未知	未培養生物	
MG102 エフェクタ —	1266	MG102-55 エフェクター	タンパク質	未知	未培養生物	
MG102 エフェクタ —	1267	MG102-56 エフェクター	タンパク質	未知	未培養生物	
MG102 エフェクタ —	1268	MG102-57 エフェクター	タンパク質	未知	未培養生物	
MG102 エフェクタ —	1269	MG102-58 エフェクター	タンパク質	未知	未培養生物	
MG102 エフェクタ —	1270	MG102-59 エフェクター	タンパク質	未知	未培養生物	
MG102 エフェクタ —	1271	MG102-60 エフェクター	タンパク質	未知	未培養生物	
MG102 エフェクタ —	1272	MG102-61 エフェクター	タンパク質	未知	未培養生物	
MG102 エフェクタ —	1273	MG102-62 エフェクター	タンパク質	未知	未培養生物	
MG144 エフェクタ —	1274	MG144-5 エフェクター	タンパク質	未知	未培養生物	
MG144 エフェクタ —	1275	MG144-6 エフェクター	タンパク質	未知	未培養生物	

【 0 3 3 8 】

10

20

30

40

50

【表 8 - 8 7】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG144 エフエクタ —	1276	MG144-7 エフエクター	タンパク質	未知	未培養生物	
MG144 エフエクタ —	1277	MG144-8 エフエクター	タンパク質	未知	未培養生物	
MG144 エフエクタ —	1278	MG144-9 エフエクター	タンパク質	未知	未培養生物	
MG144 エフエクタ —	1279	MG144-10 エフエクター	タンパク質	未知	未培養生物	
MG144 エフエクタ —	1280	MG144-11 エフエクター	タンパク質	未知	未培養生物	
MG144 エフエクタ —	1281	MG144-12 エフエクター	タンパク質	未知	未培養生物	
MG144 エフエクタ —	1282	MG144-13 エフエクター	タンパク質	未知	未培養生物	
MG144 エフエクタ —	1283	MG144-14 エフエクター	タンパク質	未知	未培養生物	
MG144 エフエクタ —	1284	MG144-15 エフエクター	タンパク質	未知	未培養生物	
MG144 エフエクタ —	1285	MG144-16 エフエクター	タンパク質	未知	未培養生物	
MG144 エフエクタ —	1286	MG144-17 エフエクター	タンパク質	未知	未培養生物	
MG144 エフエクタ —	1287	MG144-18 エフエクター	タンパク質	未知	未培養生物	
MG144 エフエクタ —	1288	MG144-19 エフエクター	タンパク質	未知	未培養生物	
MG33 エフエクタ —	1289	MG33-36 エフエクター	タンパク質	未知	未培養生物	
MG33 エフエクタ —	1290	MG33-37 エフエクター	タンパク質	未知	未培養生物	

【 0 3 3 9 】

10

20

30

40

50

【表 8 - 8 8】

カテゴリー	配列番号	説明	型	生物	他の情報	コメント又は配列
MG33 エフエクタター	I291	MG33-38 エフエクタター	タンパク質	未知	未培養生物	
MG33 エフエクタター	I292	MG33-39 エフエクタター	タンパク質	未知	未培養生物	
MG33 エフエクタター	I293	MG33-40 エフエクタター	タンパク質	未知	未培養生物	
MG33 エフエクタター	I294	MG33-41 エフエクタター	タンパク質	未知	未培養生物	
MG33 エフエクタター	I295	MG33-42 エフエクタター	タンパク質	未知	未培養生物	
MG33 エフエクタター	I296	MG33-43 エフエクタター	タンパク質	未知	未培養生物	
MG33 エフエクタター	I297	MG33-44 エフエクタター	タンパク質	未知	未培養生物	
MG33 エフエクタター	I298	MG33-45 エフエクタター	タンパク質	未知	未培養生物	
MG33 エフエクタター	I299	MG33-46 エフエクタター	タンパク質	未知	未培養生物	
MG33 エフエクタター	I300	MG33-47 エフエクタター	タンパク質	未知	未培養生物	
MG33 エフエクタター	I301	MG33-48 エフエクタター	タンパク質	未知	未培養生物	
MG33 エフエクタター	I302	MG33-49 エフエクタター	タンパク質	未知	未培養生物	
MG33 エフエクタター	I303	MG33-50 エフエクタター	タンパク質	未知	未培養生物	
MG33 エフエクタター	I304	MG33-51 エフエクタター	タンパク質	未知	未培養生物	
MG33 エフエクタター	I305	MG33-52 エフエクタター	タンパク質	未知	未培養生物	
MG33 エフエクタター	I306	MG33-53 エフエクタター	タンパク質	未知	未培養生物	
MG33 エフエクタター	I307	MG33-54 エフエクタター	タンパク質	未知	未培養生物	
MG33 エフエクタター	I308	MG33-55 エフエクタター	タンパク質	未知	未培養生物	
MG33 エフエクタター	I309	MG33-56 エフエクタター	タンパク質	未知	未培養生物	
MG33 エフエクタター	I310	MG33-57 エフエクタター	タンパク質	未知	未培養生物	
MG33 エフエクタター	I311	MG33-58 エフエクタター	タンパク質	未知	未培養生物	
MG33 エフエクタター	I312	MG33-59 エフエクタター	タンパク質	未知	未培養生物	
MG34 エフエクタター	I313	MG34-26 エフエクタター	タンパク質	未知	未培養生物	
MG34 エフエクタター	I314	MG34-27 エフエクタター	タンパク質	未知	未培養生物	
MG34 エフエクタター	I315	MG34-28 エフエクタター	タンパク質	未知	未培養生物	
MG34 エフエクタター	I316	MG34-29 エフエクタター	タンパク質	未知	未培養生物	
MG34 エフエクタター	I317	MG34-30 エフエクタター	タンパク質	未知	未培養生物	
MG34 エフエクタター	I318	MG34-31 エフエクタター	タンパク質	未知	未培養生物	
MG34 エフエクタター	I319	MG34-32 エフエクタター	タンパク質	未知	未培養生物	
MG34 エフエクタター	I320	MG34-33 エフエクタター	タンパク質	未知	未培養生物	
MG34 エフエクタター	I321	MG34-34 エフエクタター	タンパク質	未知	未培養生物	

10

20

30

40

【0340】

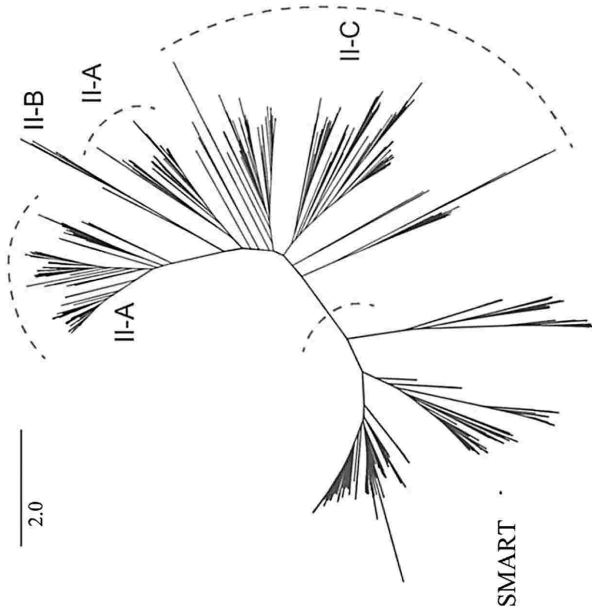
本発明の好ましい実施形態は本明細書に示され、記載されるが、このような実施形態が、例示の目的でのみ提供されることは、当業者には明らかであろう。本発明は、本明細書内で提供される特定の実施例によって限定されることは意図されていない。本発明は前述の説明を参照して記載されているが、本明細書の実施形態の記載及び説明は、限定された意味で解釈されることを意図していない。多数の変形、変更、及び置換は、ここで、本発明から逸脱することなく、当業者にとって生じるであろう。更に、本発明の全ての態様は、様々な条件及び変数に依存する、本明細書に記載される特定の描写、構成又は相対的割合に限定されないことが理解されよう。本明細書に記載される本発明の実施形態に対する

50

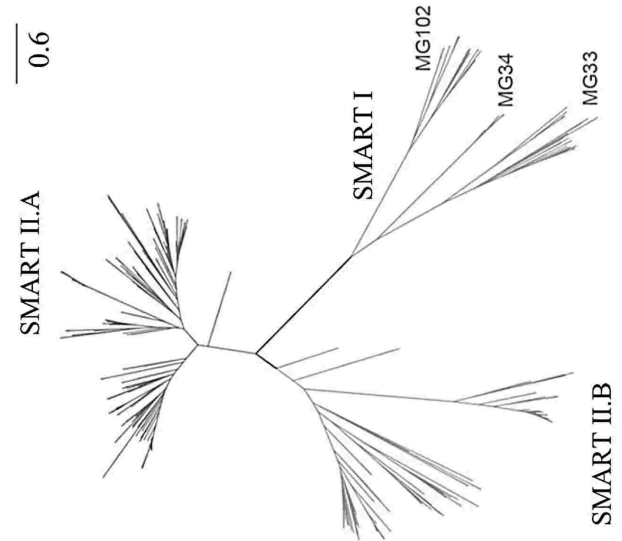
様々な代替が、本発明の実施に用いられ得ることは、理解されるべきである。したがって、本発明は、こうした任意の代替、修正、変形、又は均等物も包含することが企図される。以下の特許請求の範囲は、本発明の範囲を定義し、これらの特許請求の範囲の範囲内の方法及び構造並びにそれらの均等物がそれによって包含されることが意図される。

【図面】

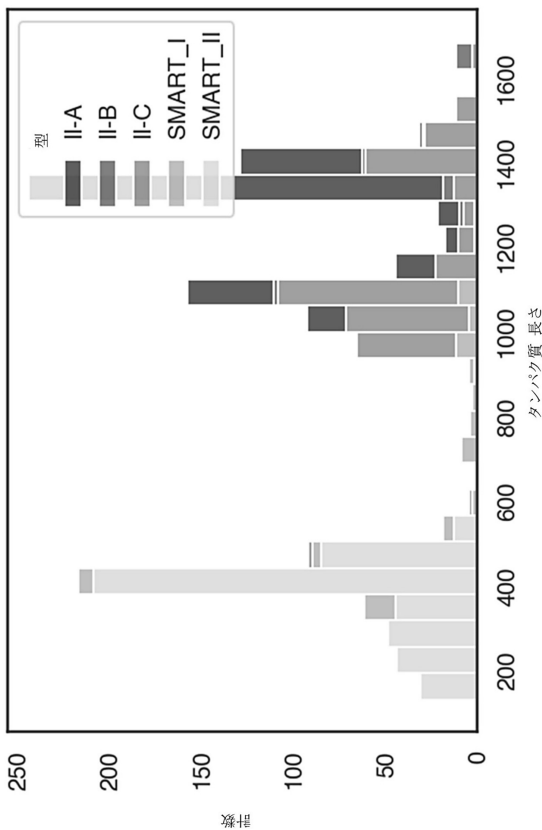
【図 1 A】



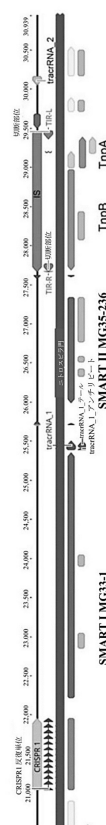
【図 1 B】



【図 2】



【図 3 A】



10

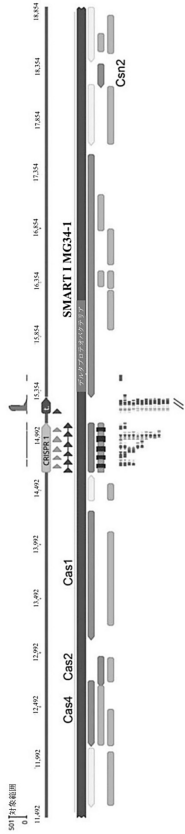
20

30

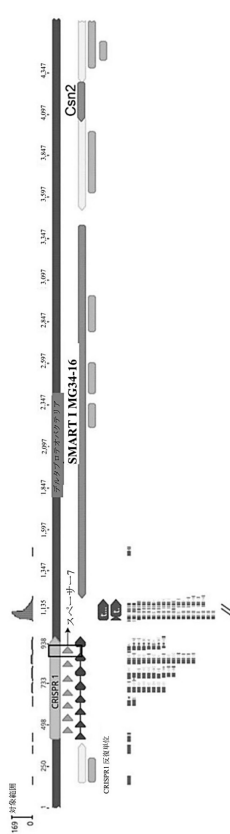
40

50

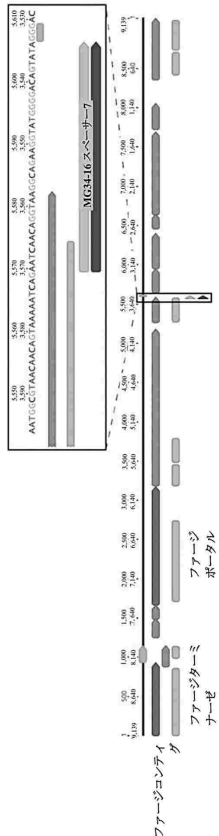
【図 3 B】



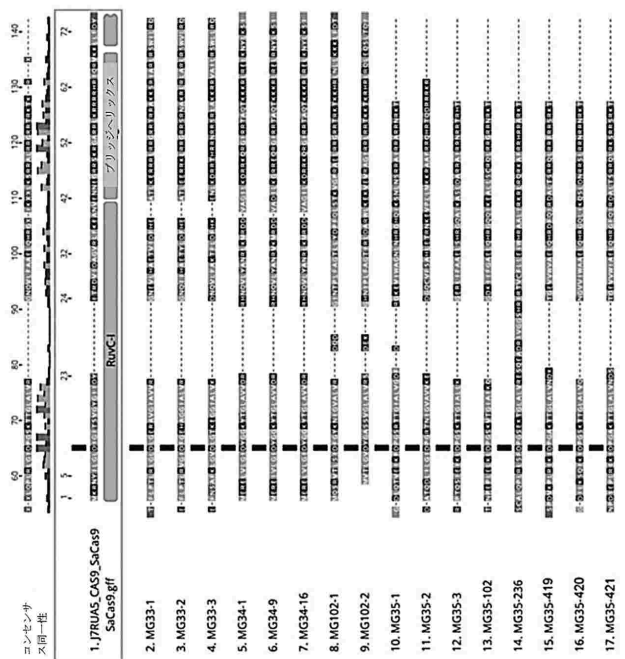
【図 3 C】



【図 3 D】



【図 4 A】



10

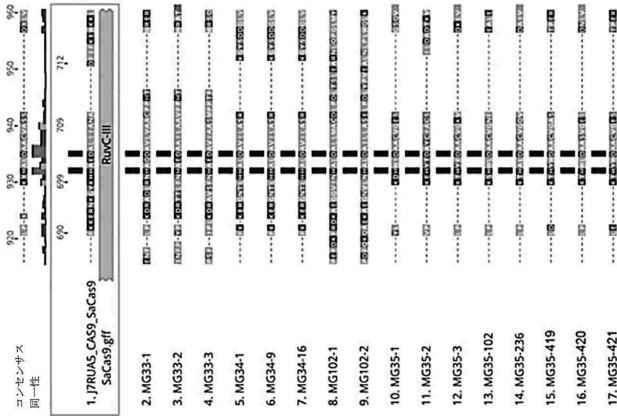
20

30

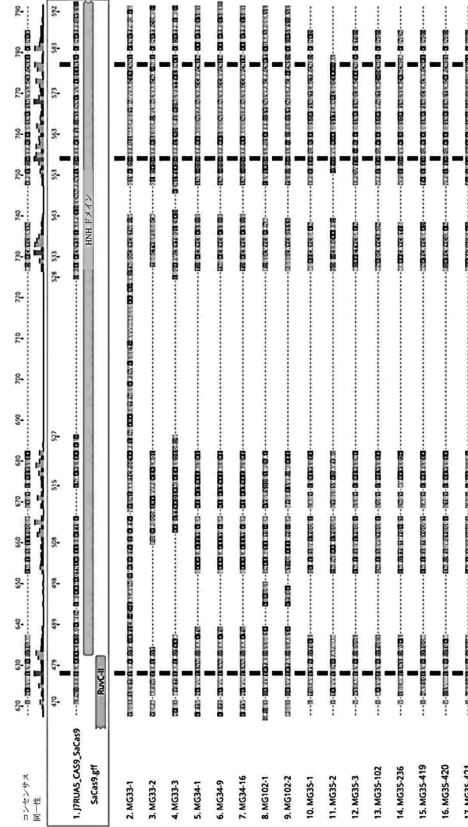
40

50

【図 4 B】



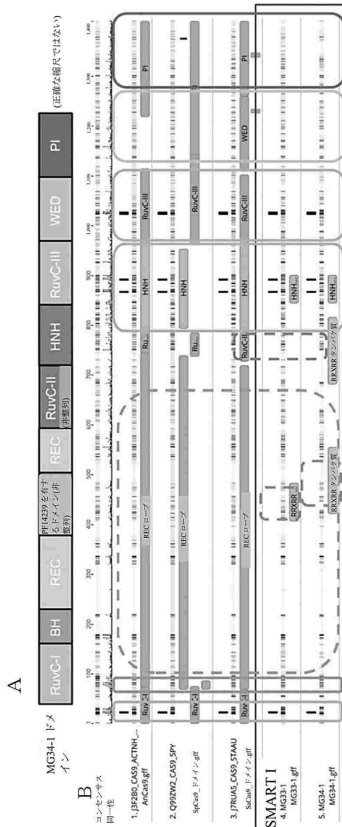
【図 4 C】



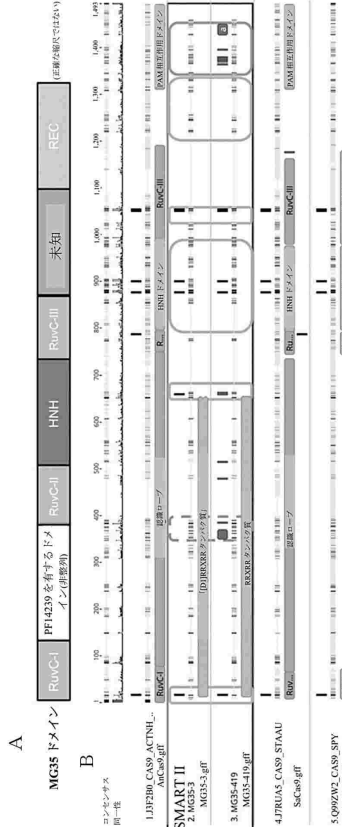
10

20

【図 5】



【図 6】

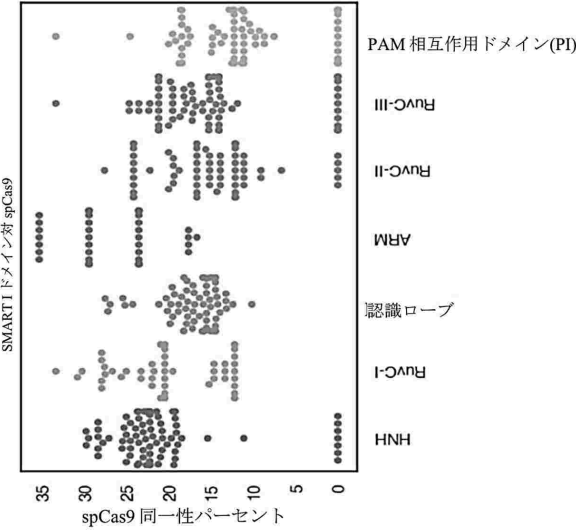


30

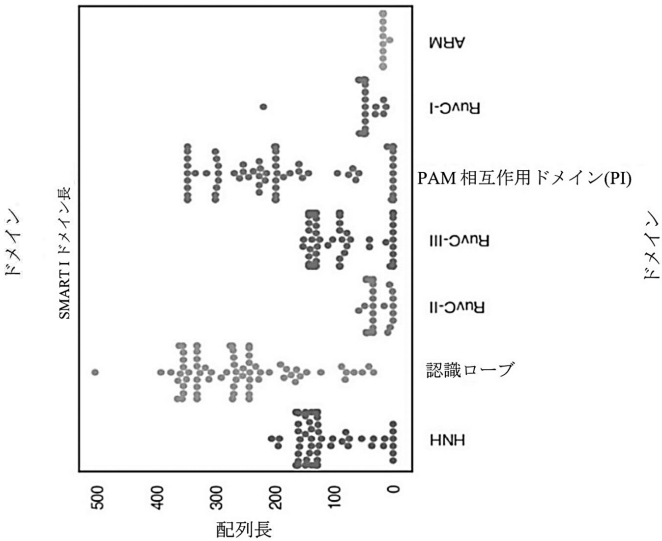
40

50

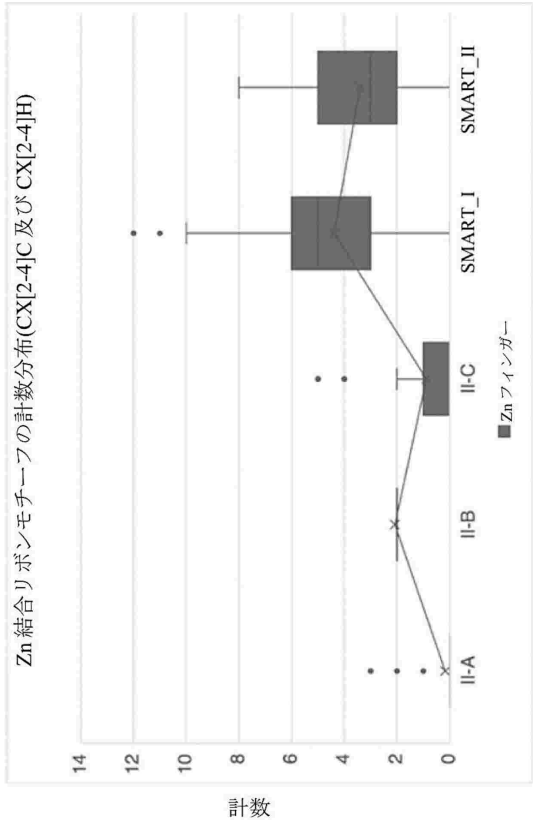
【図 7 A】



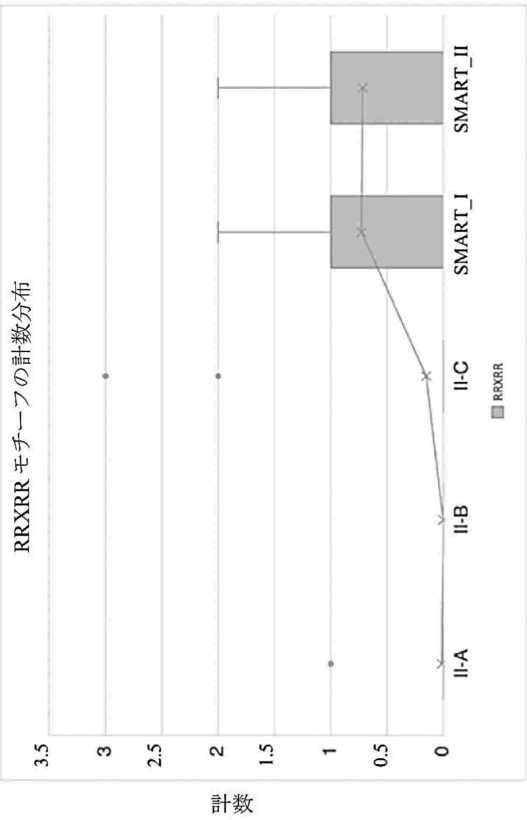
【図 7 B】



【図 8 A】



【図 8 B】



10

20

30

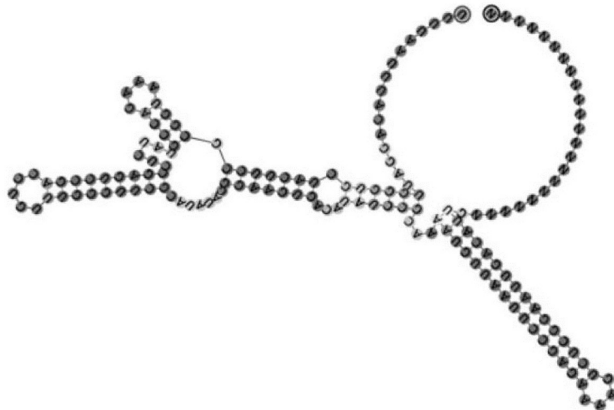
40

50

【図 9 A】



【図 9 B】



10

【図 9 C】



【図 9 D】



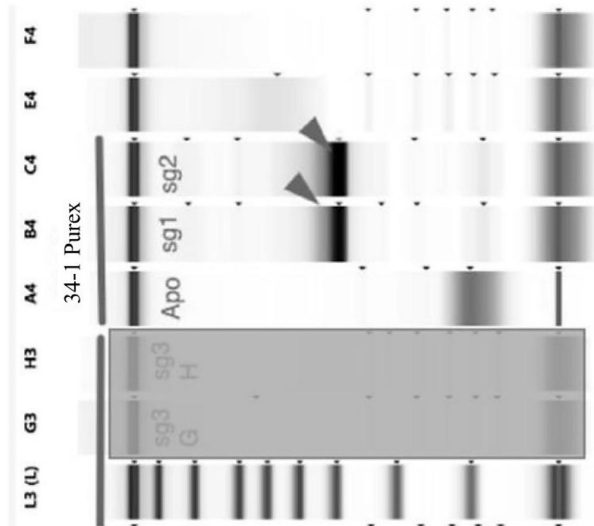
20

30

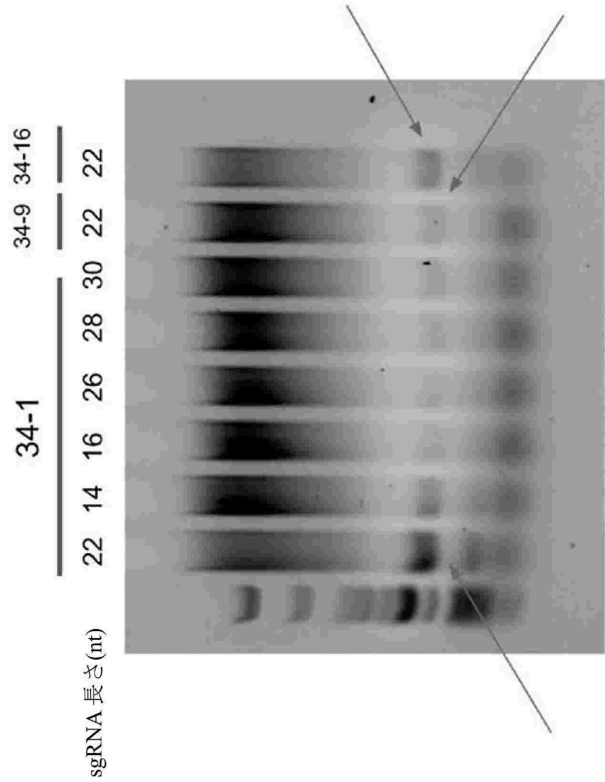
40

50

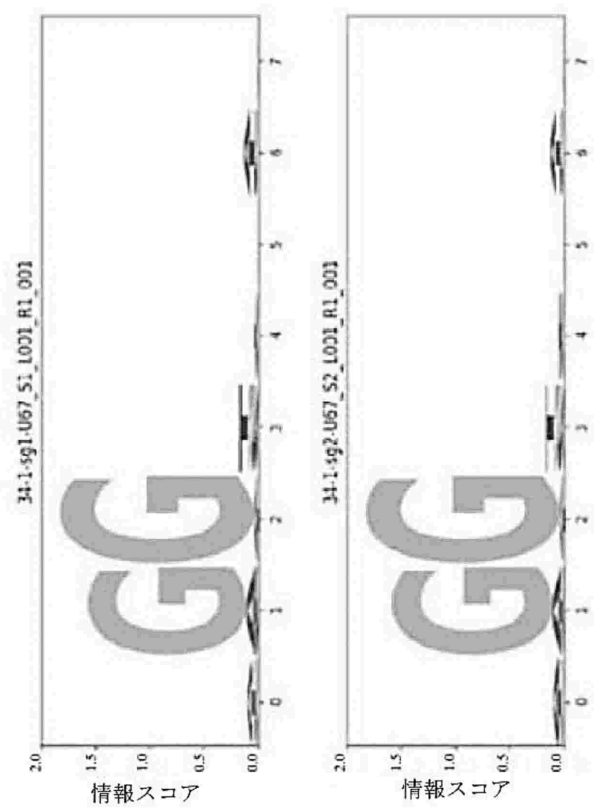
【図 10 A】



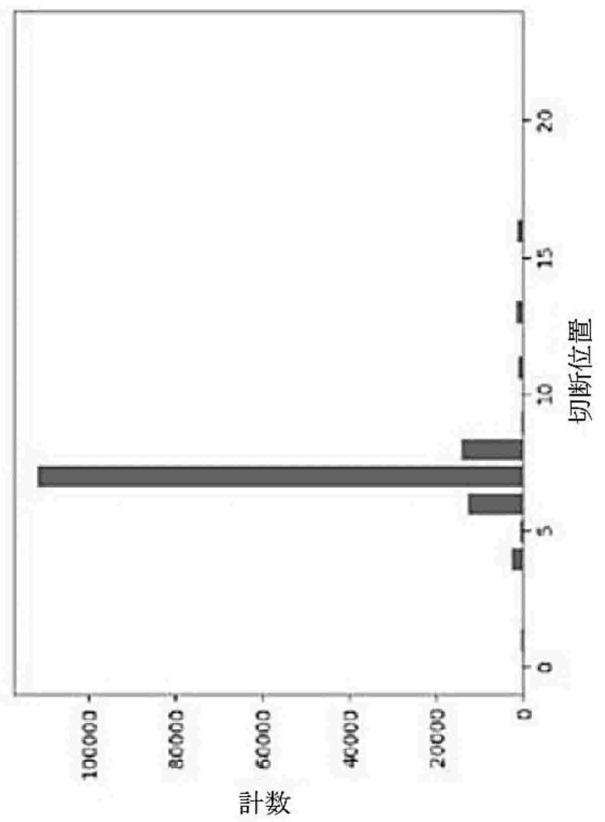
【図 10 B】



【図 11 A】



【図 11 B】



10

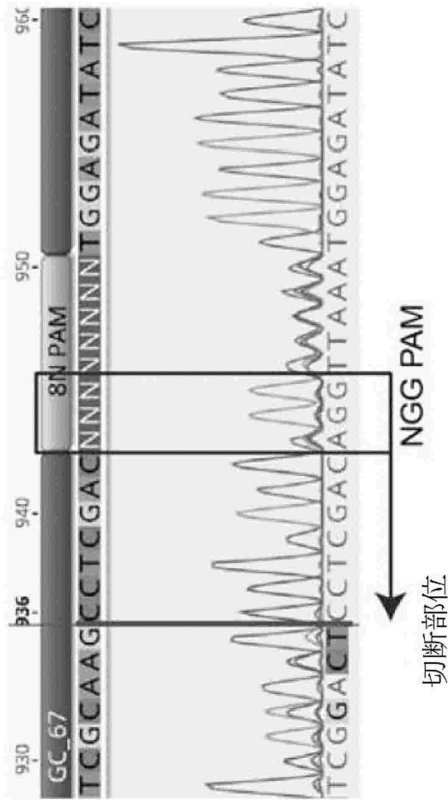
20

30

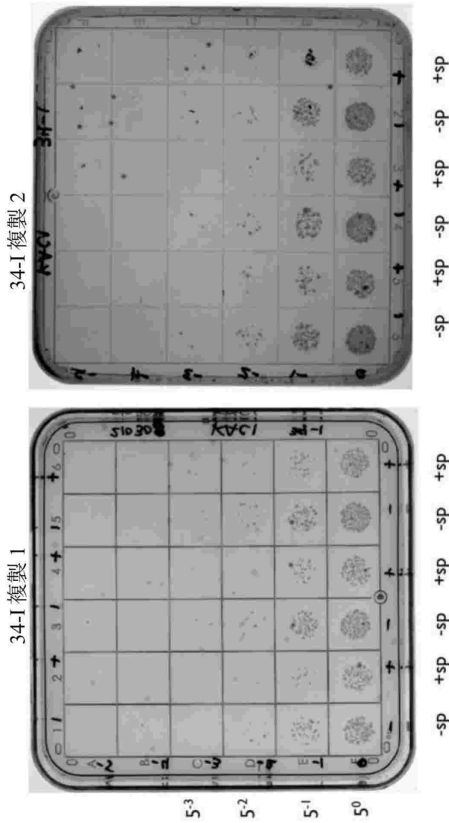
40

50

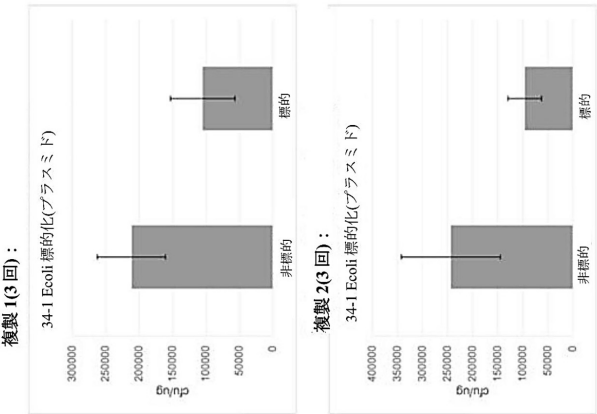
【図 1 1 C）



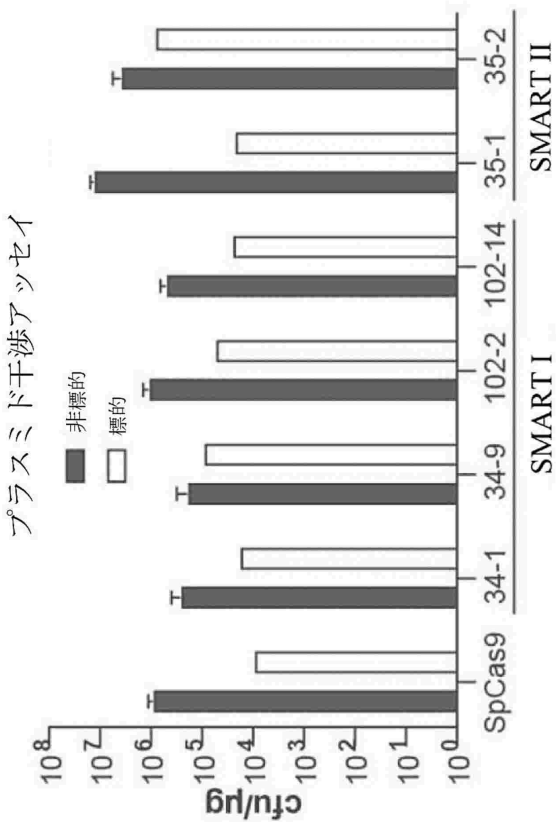
【図 1 2 A）



【図 1 2 B）



【図 1 2 C）



10

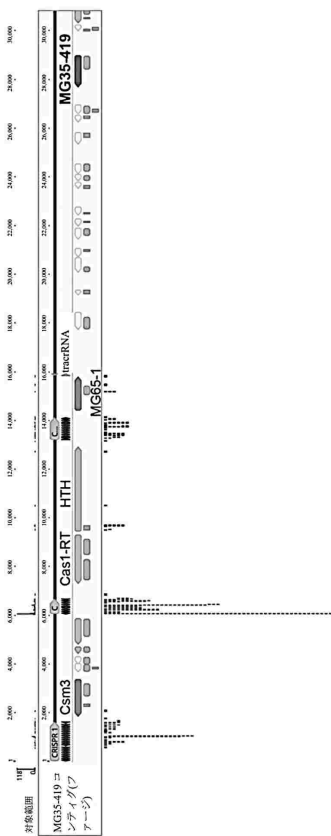
20

30

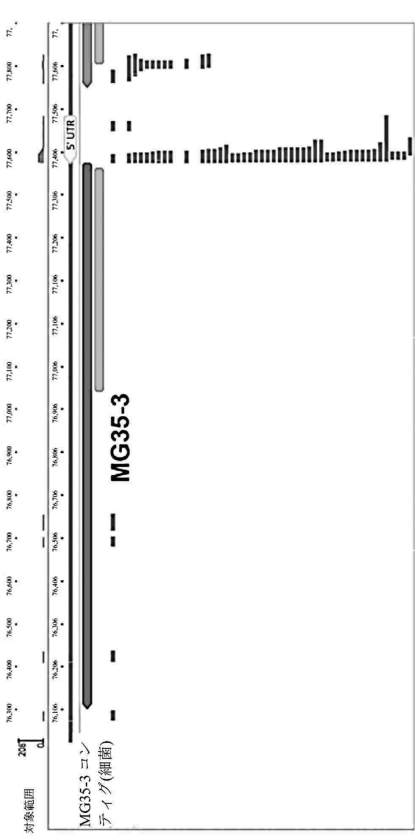
40

50

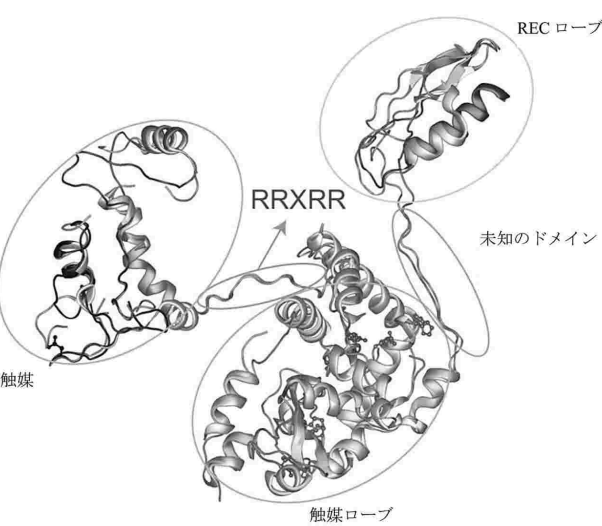
【図 1 3 A】



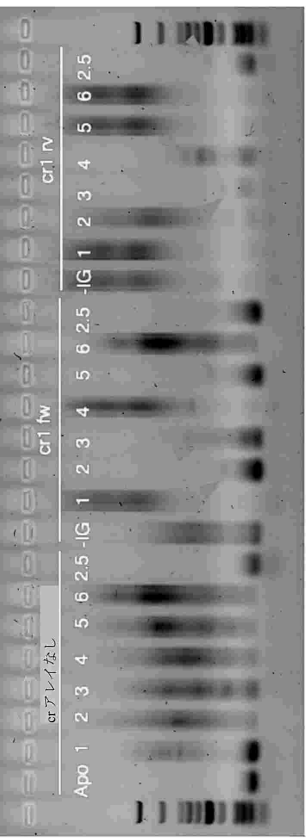
【図 1 3 B】



【図 1 4】



【図 1 5】



10

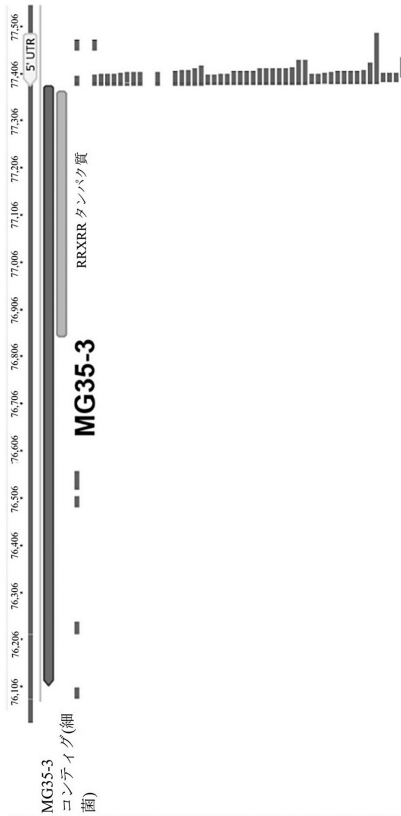
20

30

40

50

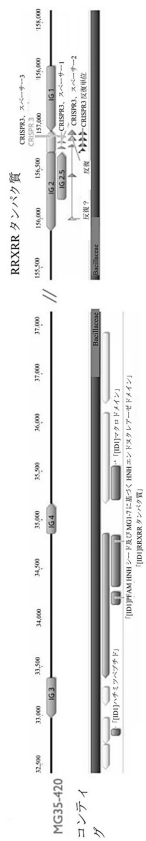
【図 16 A】



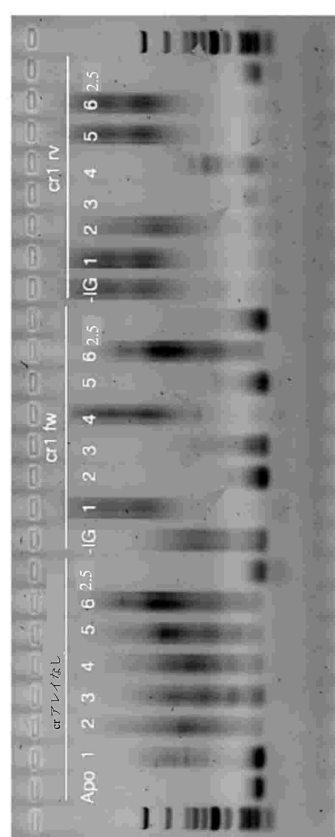
【図 16 B】



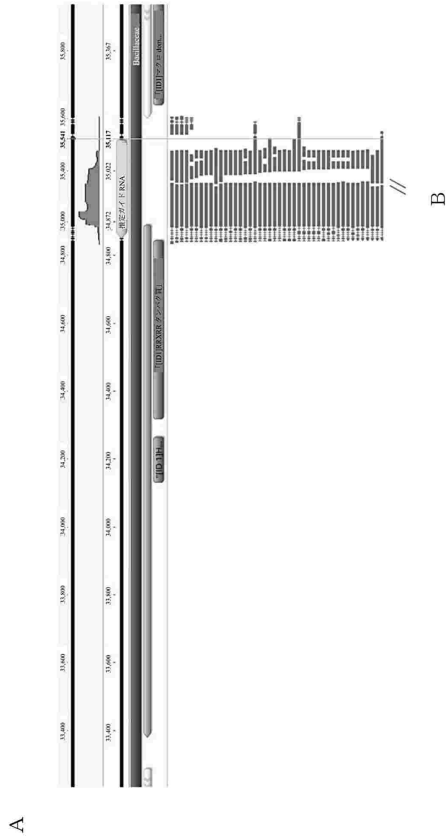
【図 17 A】



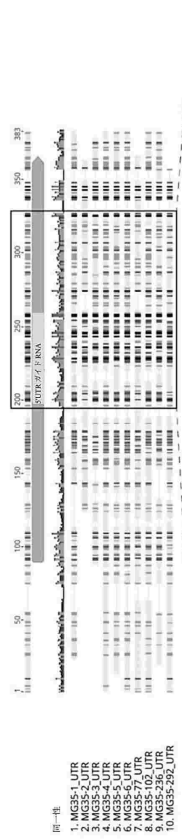
【図 17 B】



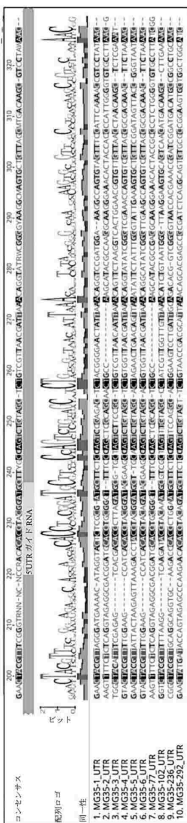
【図 18】



【図 19 A】



【図 19 B】



【図 20 A】



10

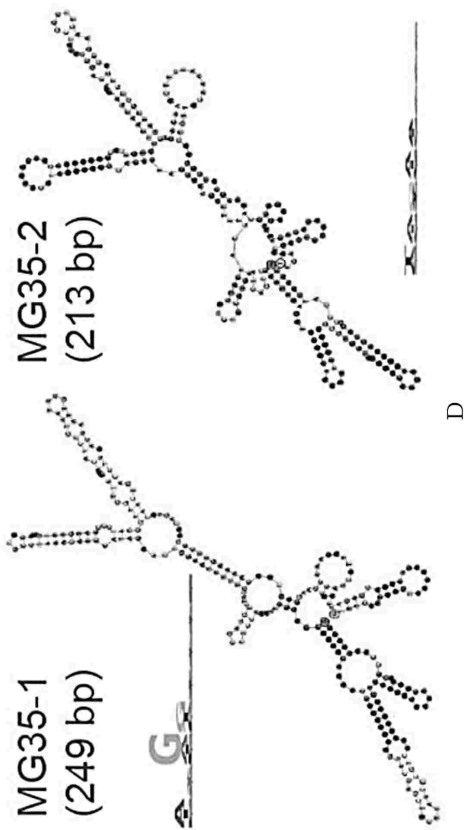
20

30

40

50

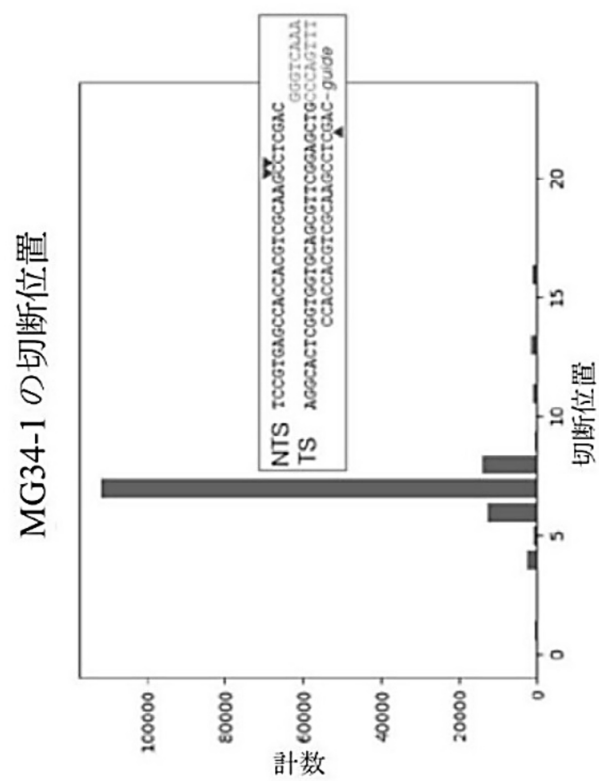
【図 2 1 - 4】



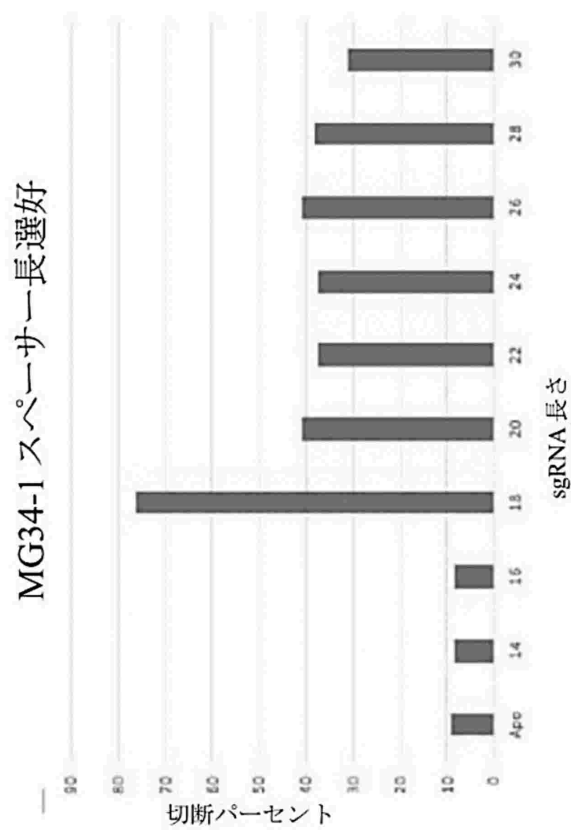
【図 2 1 - 5】



【図 2 2 A】



【図 2 2 B】



10

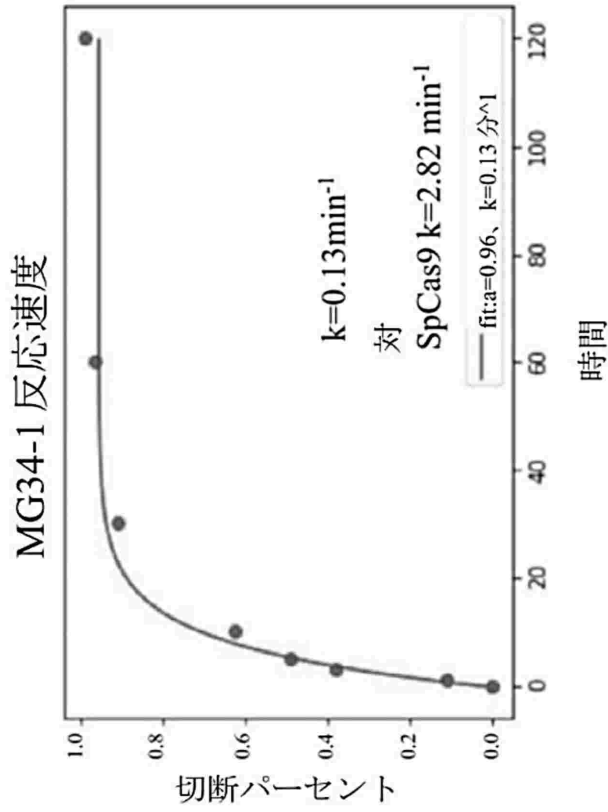
20

30

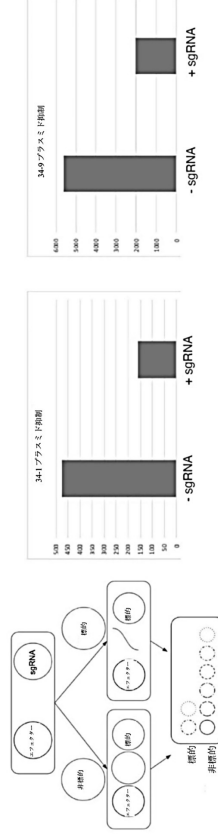
40

50

【図 2 2 C】



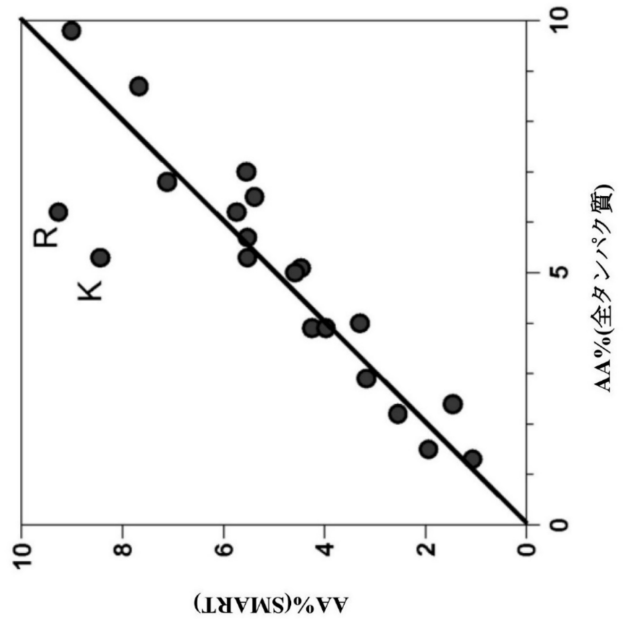
【図 2 2 D】



【図 2 3】

	A	R	N	D	C	E	Q	G	H	I	L	K	M	F	P	S	T	W	Y	V
MG35-1	7.2	9.8	4.2	3.7	2.1	3.7	4.7	8.9	1.9	6.3	7.5	10.0	1.2	3.0	4.4	5.1	5.6	1.2	3.5	5.8
MG35-2	8.4	8.0	5.0	4.6	2.1	4.8	3.6	6.1	4.4	4.8	8.4	7.8	2.7	3.1	4.8	4.6	6.5	0.8	3.2	6.5
MG35-3	7.6	8.5	4.5	3.5	2.1	4.3	4.5	7.1	2.8	5.7	9.7	8.7	1.2	4.3	4.3	5.2	6.9	0.0	3.1	5.2
MG35-6	6.8	9.3	4.7	3.7	2.3	4.0	4.7	7.7	2.3	5.4	9.6	8.6	0.9	3.0	4.7	4.9	7.0	1.4	2.8	6.1
MG35-102	8.0	8.0	4.5	3.3	2.4	4.0	5.4	8.0	2.6	5.7	9.7	8.5	0.7	3.8	5.0	4.7	6.4	1.2	2.8	5.4
IscB (Altae-Tran, 2021)	10.0	11.4	3.2	2.1	1.6	4.6	6.2	7.7	3.4	5.0	10.9	5.9	1.4	2.5	4.3	4.6	5.5	1.4	2.3	6.2
MG34-1	5.5	8.6	5.4	5.8	1.6	9.4	1.7	6.2	2.4	6.2	6.6	12.9	2.0	3.5	3.2	4.4	5.2	0.7	3.1	5.9
MG102-2	6.8	8.7	3.0	5.9	1.7	6.0	4.3	6.7	2.4	5.8	9.3	7.4	1.7	3.4	5.1	7.3	5.0	1.1	3.7	4.9
MG102-14	8.1	11.1	3.7	5.0	1.9	7.7	2.8	6.0	2.3	5.6	9.0	7.0	1.2	3.2	5.0	6.7	4.6	1.1	3.6	4.6
MG102-35	8.0	9.0	4.4	6.1	1.8	7.3	2.8	6.4	1.7	4.9	9.4	8.5	1.6	3.2	4.7	6.9	4.5	0.9	3.4	4.4
MG102-45	8.1	9.6	4.2	5.5	1.8	7.4	3.0	7.6	1.9	5.5	9.1	7.6	1.5	3.3	5.0	6.7	3.7	1.1	3.4	4.3
典型的	8.7	6.2	3.9	5.1	1.5	6.2	3.9	6.8	2.2	5.7	9.8	5.3	2.4	4.0	5.0	7.0	5.3	1.3	2.9	6.5

【図 2 4 A】



10

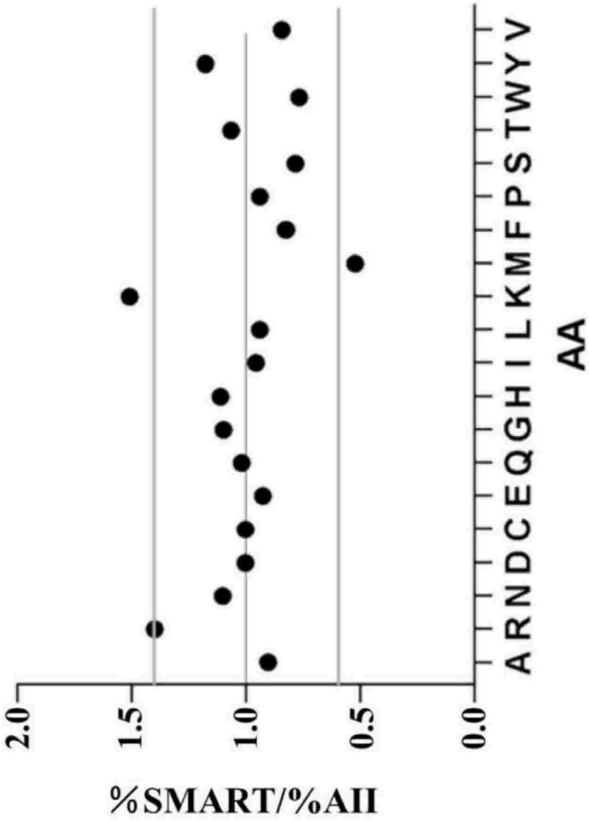
20

30

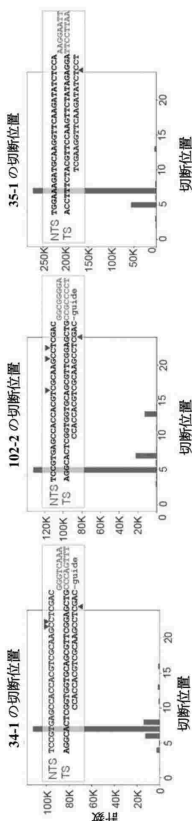
40

50

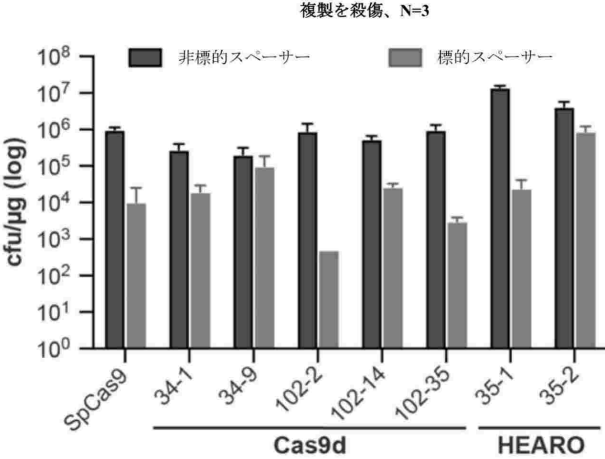
【図 2 4 B】



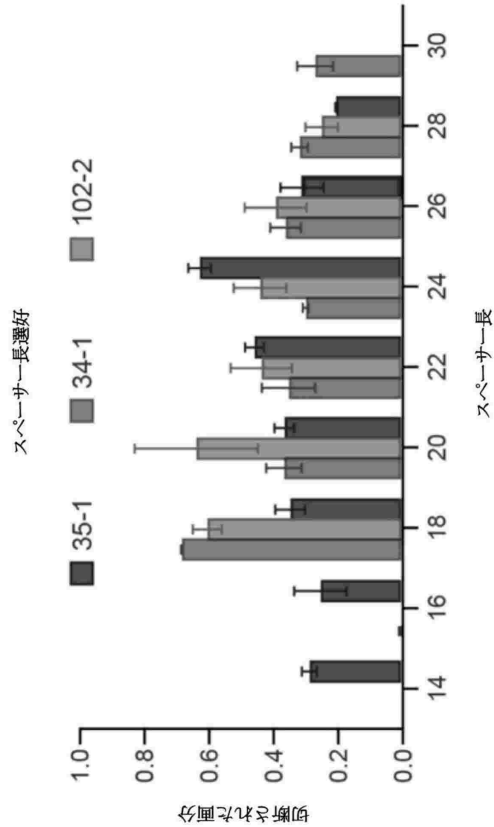
【図 2 5 A】



【図 2 5 B】



【図 2 5 C】



10

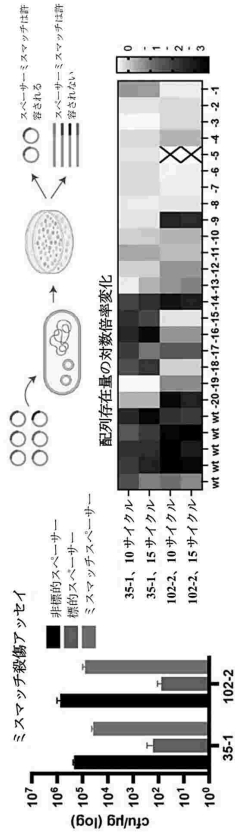
20

30

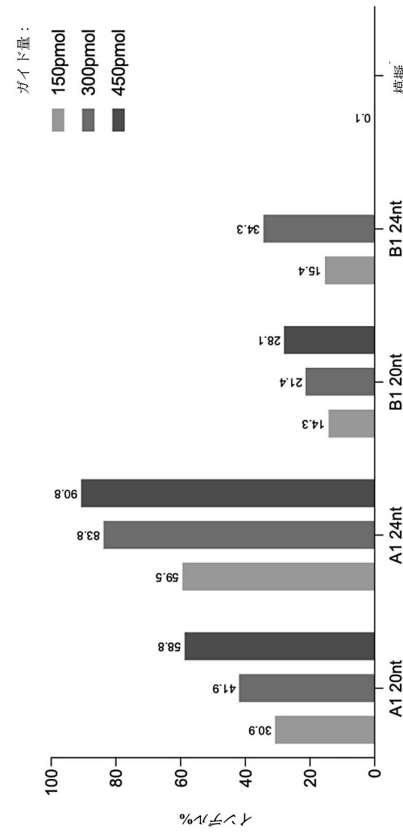
40

50

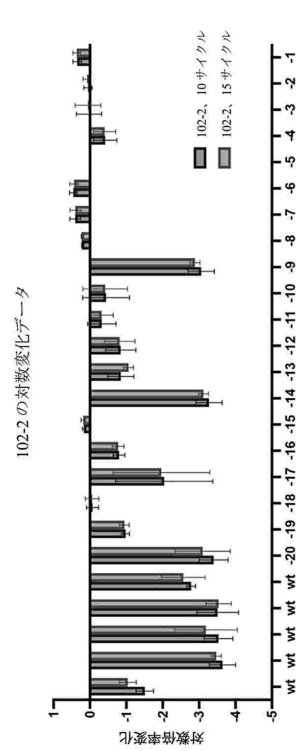
【図 25 D】



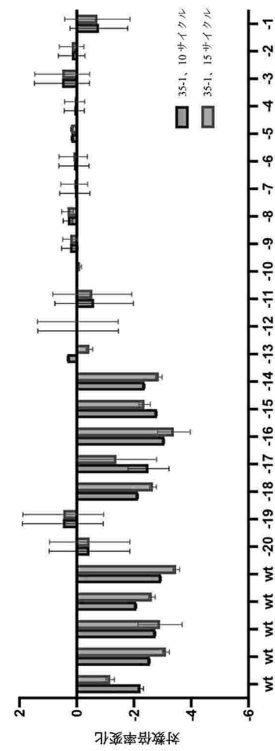
【図 26】



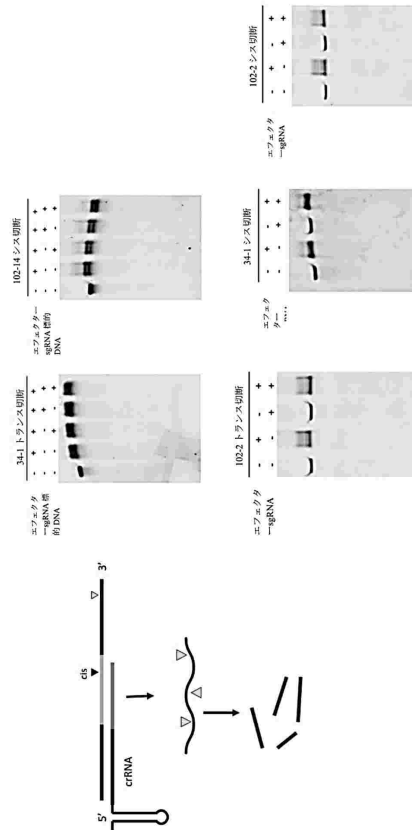
【図 27】



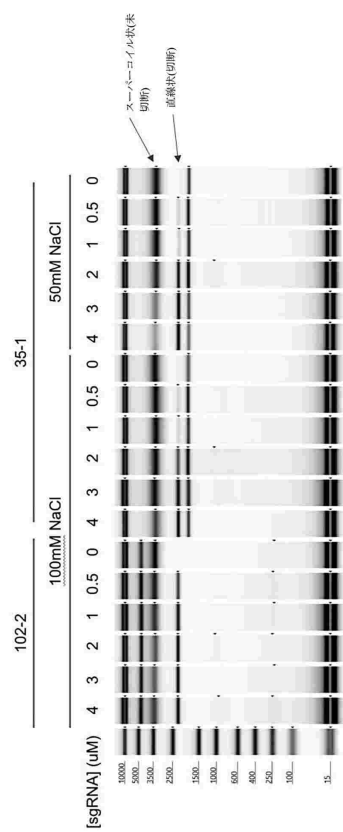
35-1 の対数変化データ



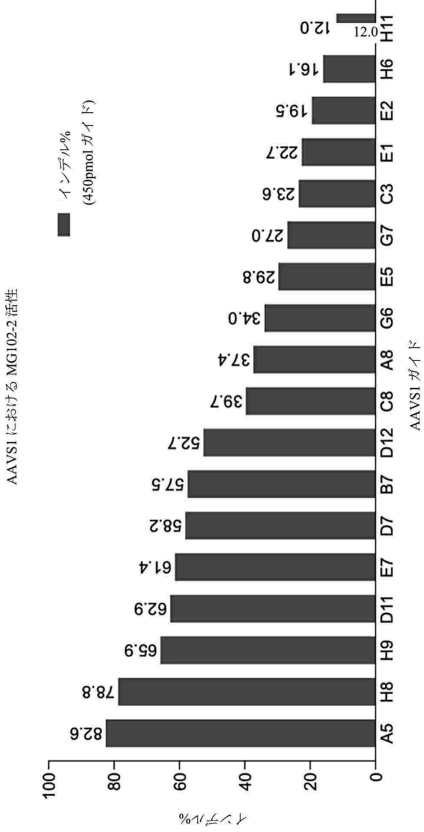
【図 28】



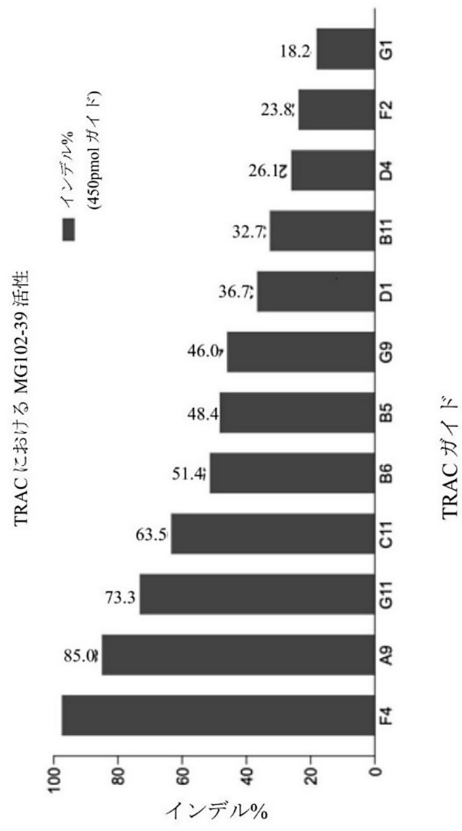
【図 2 9】



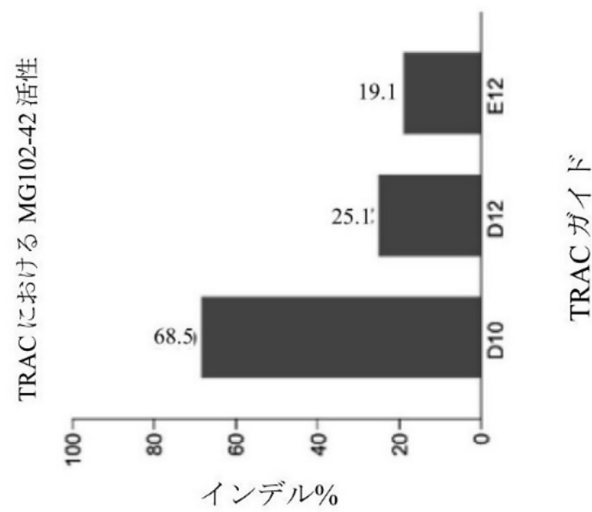
【図 3 0 A】



【図 3 0 B】



【図 3 0 C】



10

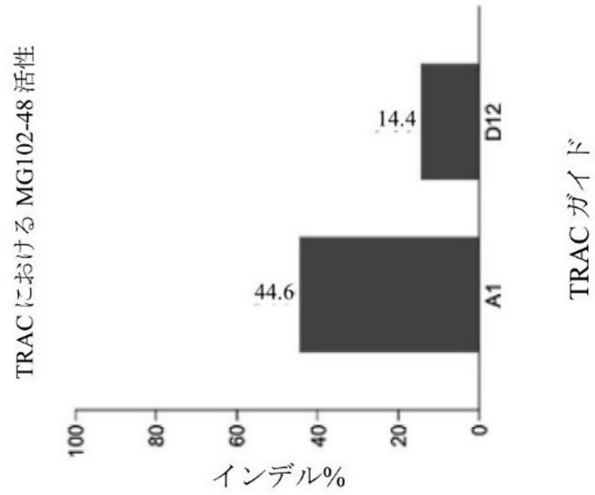
20

30

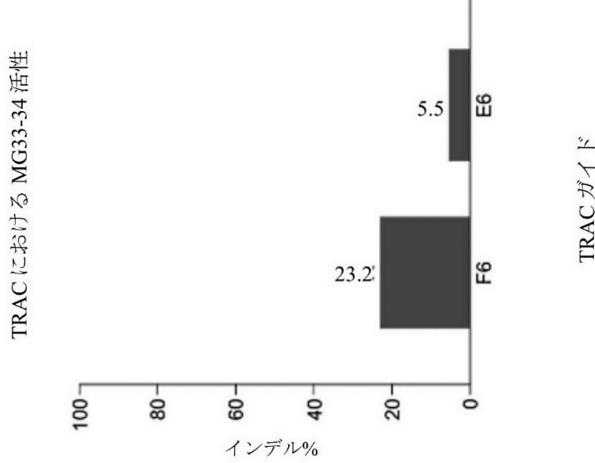
40

50

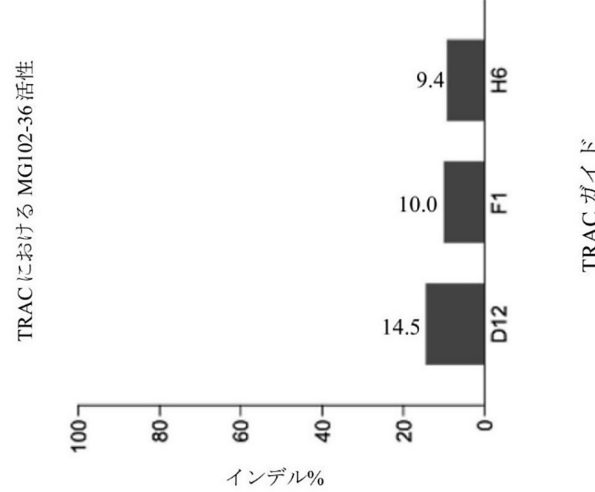
【図 30 D】



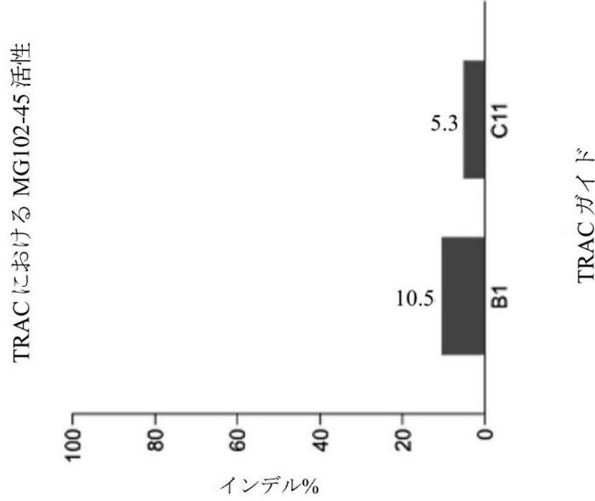
【図 30 E】



【図 30 F】



【図 30 G】



10

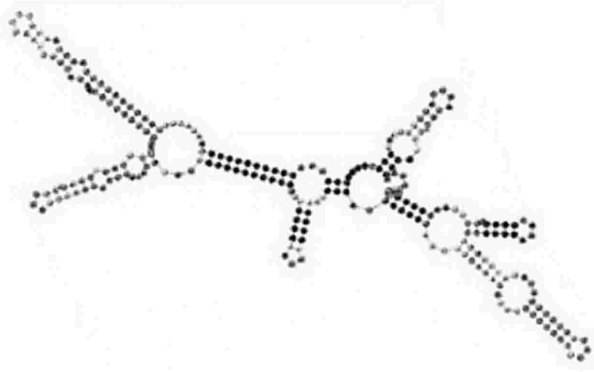
20

30

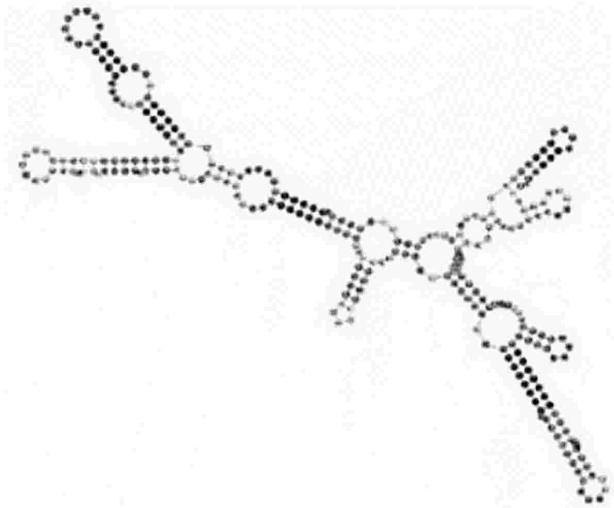
40

50

【図 3 2 D】

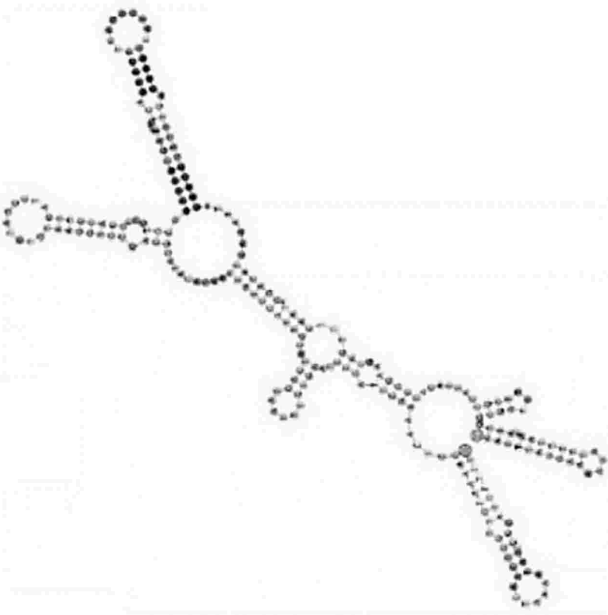


【図 3 2 E】



10

【図 3 2 F】



【図 3 2 G】



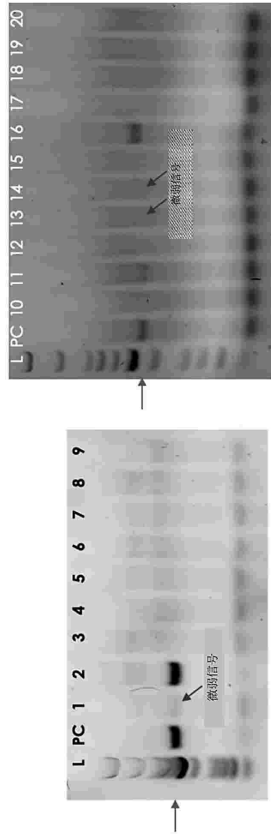
20

30

40

50

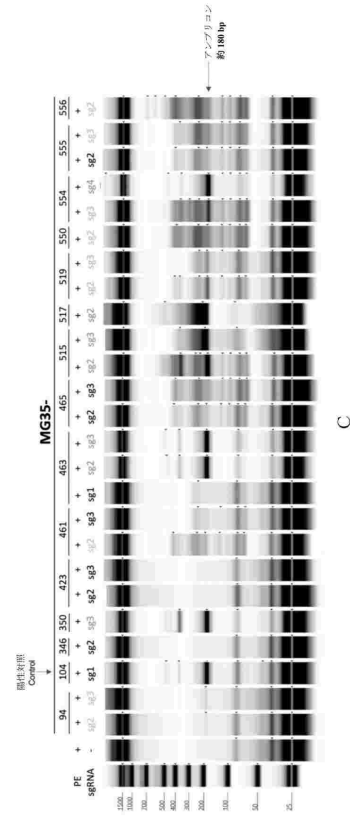
【図 3 3 - 1】



B

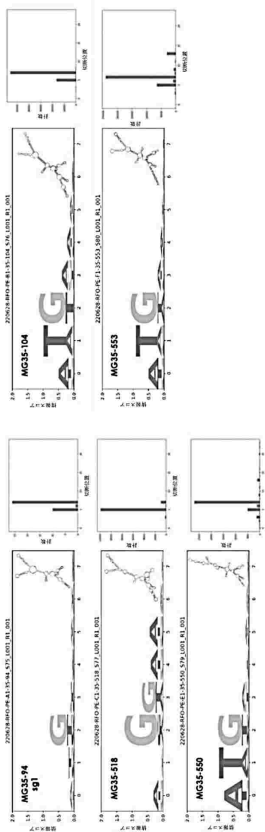
A

【図 3 3 - 2】

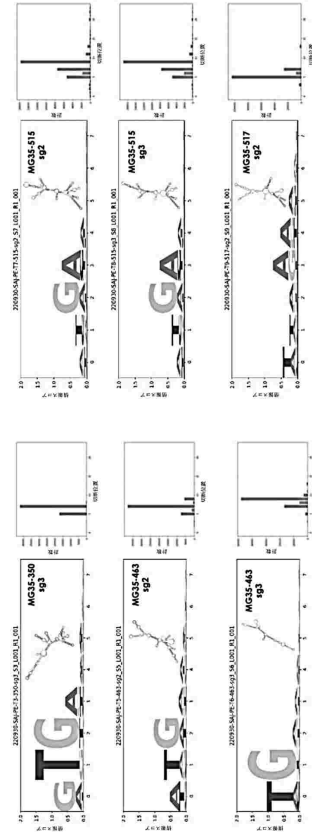


C

【図 3 4 - 1】



【図 3 4 - 2】



10

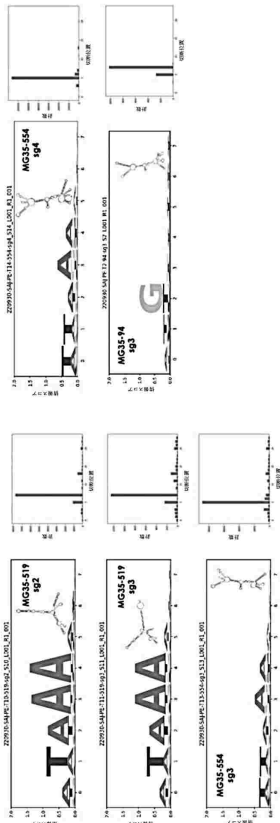
20

30

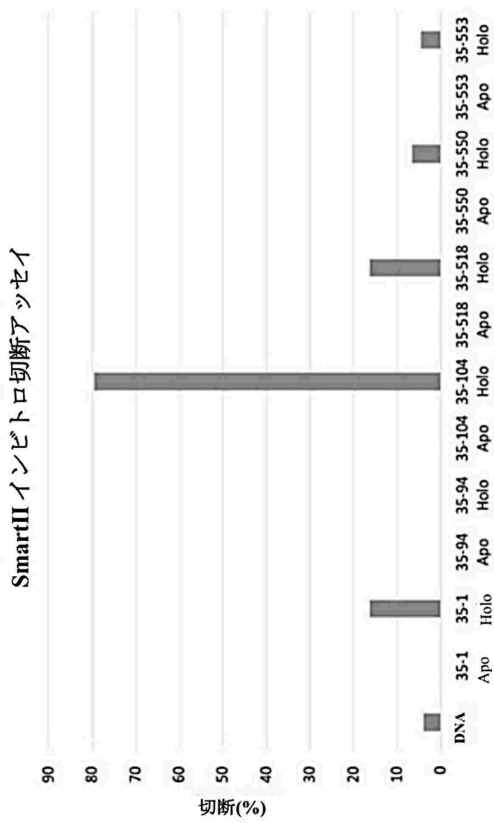
40

50

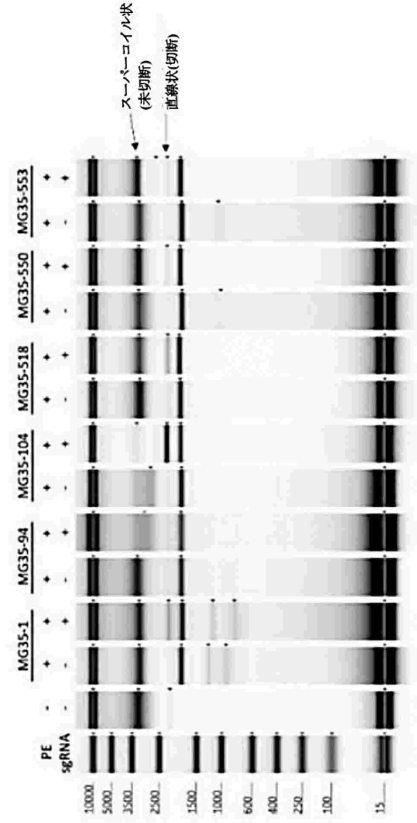
【図 3 4 - 3】



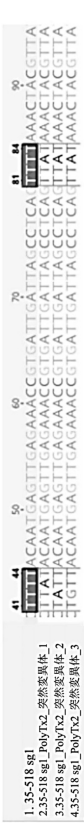
【図 3 5 B】



【図 3 5 A】



【図 3 6 A】



10

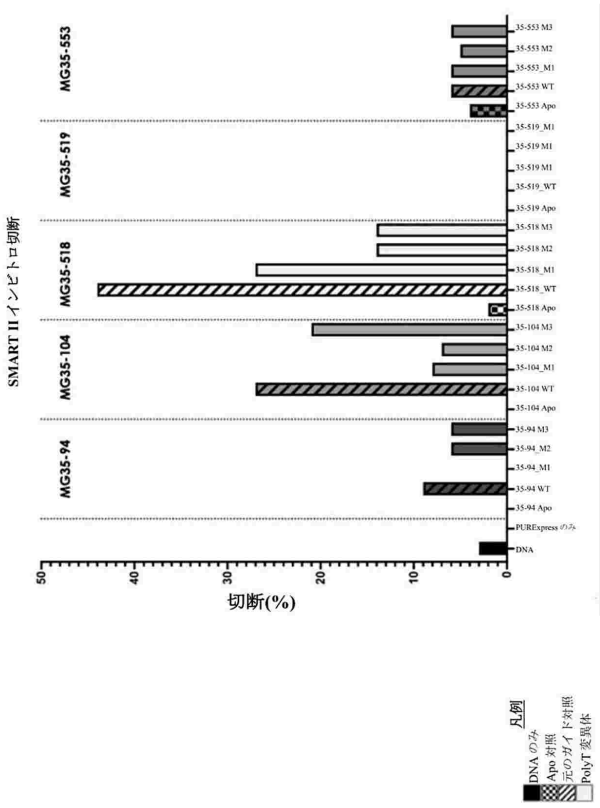
20

30

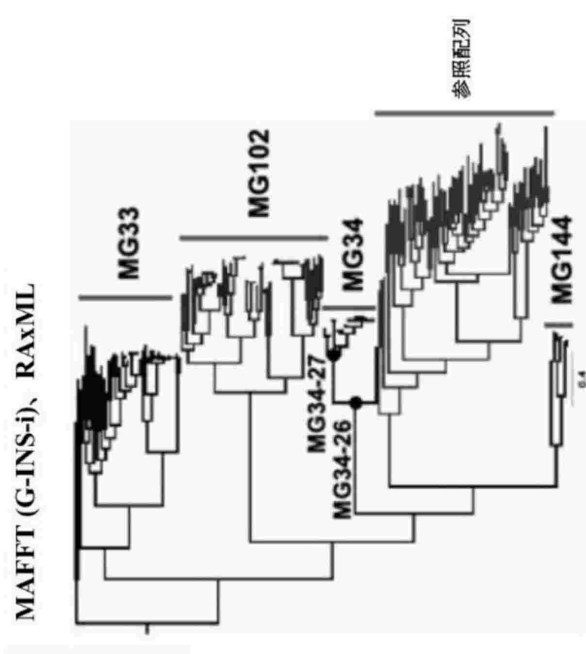
40

50

【 図 3 6 B 】



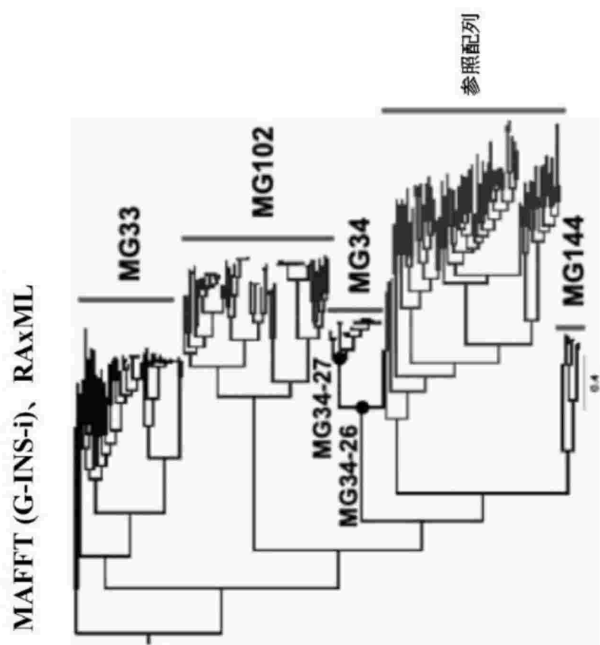
【 図 3 7 A 】



10

20

【 図 3 7 B 】



【 図 】

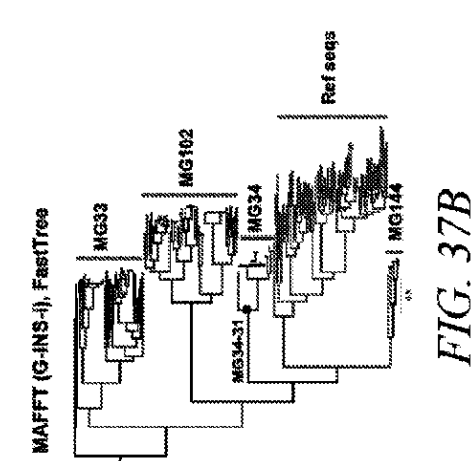


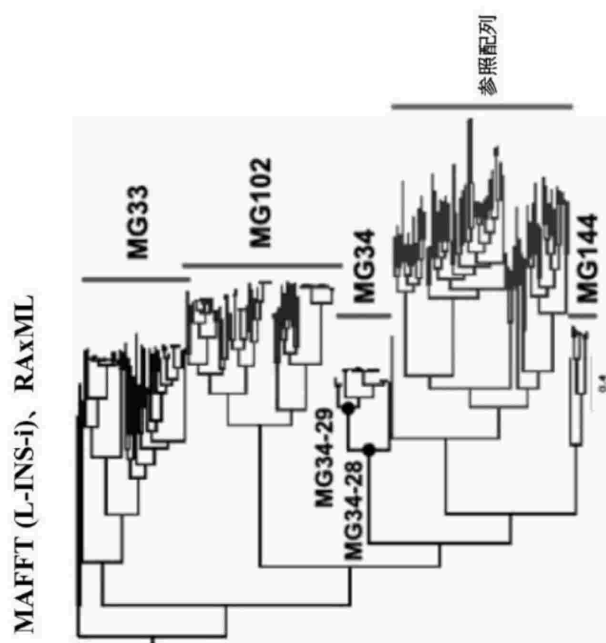
FIG. 37B

30

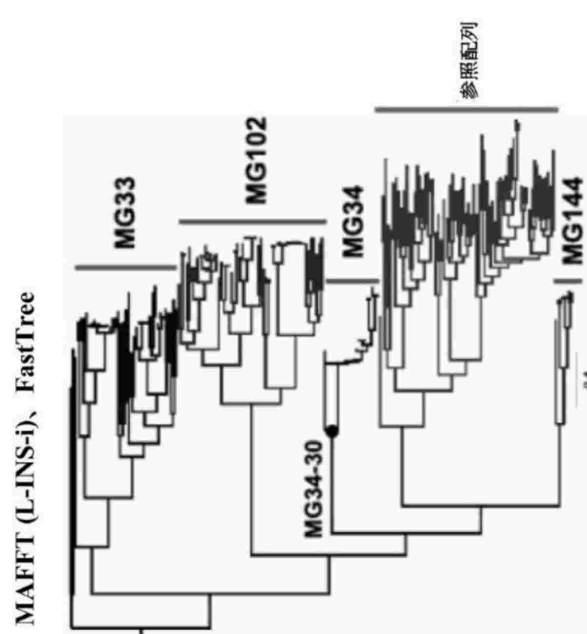
40

50

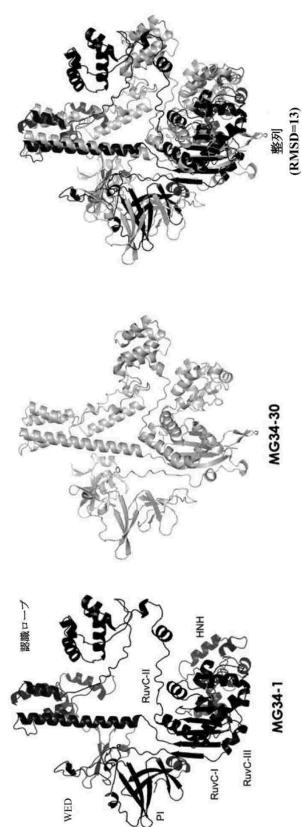
【図 37 C】



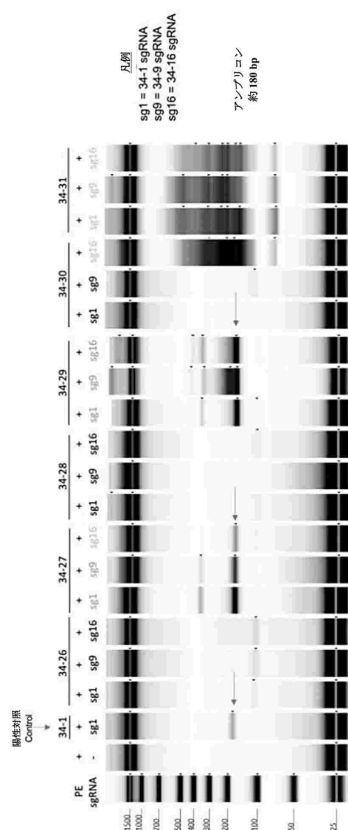
【図 3 7 D】



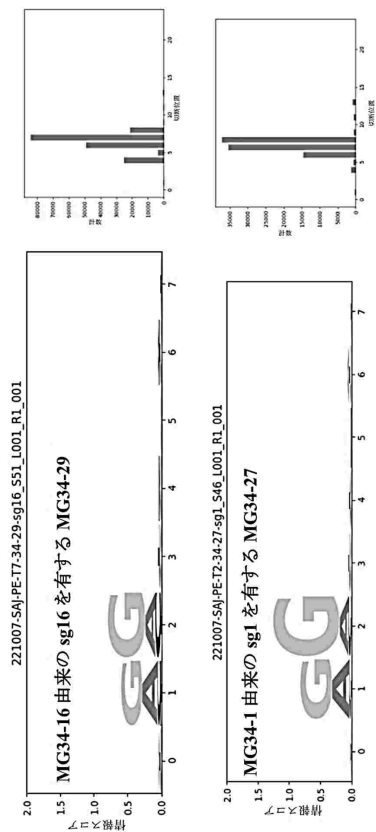
【図 38】



【図 3 9】



【図 40】



【配列表】

2024541402000001.xml

10

20

30

40

50

【国際調査報告】

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2022/080437

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:
- a. ☒ forming part of the international application as filed.
- b. ☐ furnished subsequent to the international filing date for the purposes of international search (Rule 13~~ter~~.1(a)).
☐ accompanied by a statement to the effect that the sequence listing does not go beyond the disclosure in the international application as filed.
2. ☐ With regard to any nucleotide and/or amino acid sequence disclosed in the international application, this report has been established to the extent that a meaningful search could be carried out without a WIPO Standard ST.26 compliant sequence listing.
3. Additional comments:
Sequences searched were SEQ ID NOs: 1, 463-486, 981-988 and 1289-1312.

10

20

30

40

50

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2022/080437

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
the subject matter listed in Rule 39 on which, under Article 17(2)(a)(i), an international search is not required to be carried out, including
2. ☒ Claims Nos.: 60, 113-138, 167-170
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
See Supplemental Box
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a)

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

See Supplemental Box for Details

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-20, 24-27, 29-34, 38-58, 84-101, 105-112, 139-144, 148-152, and 154 (in part)

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2022/080437

A. CLASSIFICATION OF SUBJECT MATTER

C12N 9/22 (2006.01) C12N 15/10 (2006.01) C12N 15/11 (2006.01)

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Genomequest: SEQ ID NOs: 1, 463-486, 981-988 & 1289-1312 at 70%.
PATENW, Medline, CAPUS, Biosis, Embase: keywords MG33, type IIC, SMART, nuclease, endonuclease, archaea, transposon encoded, CRISPR, Cas, and like terms.
PATENW: CPC code C12N2310/20.
DOCDB, DWPI & Internal databases: Applicant & Inventor names.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
	Documents are listed in the continuation of Box C	

☒ Further documents are listed in the continuation of Box C ☒ See patent family annex

* Special categories of cited documents:	
"A" document defining the general state of the art which is not considered to be of particular relevance	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"D" document cited by the applicant in the international application	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier application or patent but published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search
28 February 2023

Date of mailing of the international search report
28 February 2023

Name and mailing address of the ISA/AU

AUSTRALIAN PATENT OFFICE
PO BOX 200, WODEN ACT 2606, AUSTRALIA
Email address: pct@ipaustalia.gov.au

Authorised officer

Jive Bellhouse
AUSTRALIAN PATENT OFFICE
(ISO 9001 Quality Certified Service)
Telephone No. +61 2 6283 2959

INTERNATIONAL SEARCH REPORT		International application No.
C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		PCT/US2022/080437
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2021/202568 A1 (METAGENOMI IP TECHNOLOGIES, LLC) 07 October 2021 see abstract, figures 1, 3 & 4, paragraphs 7, 23-24, 37-40, 95-99, 109, 142, table 3, SEQ ID NOs: 1, 201-220, 463-486, claims	1-20, 24-27, 29-34, 38-58, 84-101, 105-112, 139-144, 148-152, and 154
X	WO 2020/055941 A1 (EXCISION BIOTHERAPEUTICS, INC.) 19 March 2020 see paragraphs 48, 55-60, 132, 134; figure 2, SEQ ID NO: 3, claims 1-2	1-20, 24-27, 29-34, 38-58, 84-101, 105-112, 139-144, 148-152, and 154
X	US 2020/0332273 A1 (METAGENOMI IP TECHNOLOGIES, LLC) 22 October 2020 see Examples 1, 2A, 11 & pars 7, 10-13, 77-85, 213-226, table 1, SEQ ID NO: 280	1-20, 24-27, 29-34, 38-58, 84-101, 105-112, 139-144, 148-152, and 154
A	ALTAE-TRAN, H. et al., "The widespread IS200/IS605 transposon family encodes diverse programmable RNA-guided endonucleases", Science, 9 September 2021, vol. 374, pp. 57-65, DOI: 10.1126/science.abj6856	
P,X	ALIAGA GOLTSMAN, D.S., et al. "Compact Cas9d and HEARO enzymes for genome editing discovered from uncultivated microbes", Nature Communications, 15 December 2022, vol. 13, no. 7602. https://doi.org/10.1038/s41467-022-35257-7 see whole document, & suppl data files 1 & 3	1-20, 24-27, 29-34, 38-58, 84-101, 105-112, 139-144, 148-152, and 154 (in part)

Form PCT/ISA/210 (fifth sheet) (July 2019)

INTERNATIONAL SEARCH REPORT	International application No. PCT/US2022/080437
Supplemental Box	
Continuation of Box II Claims 60, 113-138, 167-170 were found unsearchable because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically: Claim 60 does not comply with Article 6 (clarity) because the feature "said TAM interacting domain" lacks antecedence, and also refers to different nuclease family to its parent claim 59 (See Box VIII); Claims 113-138 do not comply with Article 6 (clarity) because the claim dependency is not clear as a result of a typographical error "claims - 244 -- 73" (See Box VIII); Claims 167-170 do not comply with Rule 6.2(a) because they rely on references to the description and/or drawings (See Box VII Certain Defects). Continuation of: Box III This International Application does not comply with the requirements of unity of invention because it does not relate to one invention or to a group of inventions so linked as to form a single general inventive concept. This Authority has found that there are different inventions based on the following features that separate the claims into distinct groups: • Invention 1 is defined by claims 1-20, 24-27, 29-34, 38-58, 84-101, 105-112, 139-144, 148-152, and 154 (in part). The feature of an endonuclease belonging to the MG33 family (SEQ ID NOs: 1, 463-486, 981-988 and 1289-1312) is specific to this group of claims. • Invention 2 is defined by claims 1-20, 24-27, 29-59, 61-73, 84-101, 105-112, 139-142, 144, 148-151, and 154 (in part). The feature of an endonuclease belonging to the MG34 family (SEQ ID NOs: 2-24, 487-488 and 1313-1321) is specific to this group of claims. • Invention 3 is defined by claims 1-20, 24-34, 38-58, 84-101, 105-112, 139-142, 144, 148-151, 154, 158-166 (in part). The feature of an endonuclease belonging to the MG35 family (SEQ ID NOs: 25-198, 221-459, 489-580, 617-668, 674 and 675) is specific to this group of claims. • Invention 4 is defined by claims 1-20, 24-27, 29-34, 38-58, 84-101, 105-112, 139-142, 144, 148-154, 158-166 (in part). The feature of an endonuclease belonging to the MG102 family (SEQ ID NOs: 581-612, 989-1002 and 1260-1273) is specific to this group of claims. • Invention 5 is defined by claims 1-20, 24-27, 29-34, 38-58, 84-101, 105-112, 139-142, 144, 148-151, and 154 (in part). The feature of an endonuclease belonging to the MG143 family (SEQ ID NO: 975) is specific to this group of claims. • Invention 6 is defined by claims 1-20, 24-27, 29-34, 38-58, 84-101, 105-112, 139-142, 144, 148-151, and 154 (in part). The feature of an endonuclease belonging to the MG144 family (SEQ ID NO: 976-979 and 1274-1288) is specific to this group of claims. • Invention 7 is defined by claims 1-20, 24-27, 29-34, 38-58, 84-101, 105-112, 139-142, 144, 148-151, and 154 (in part). The feature of an endonuclease belonging to the MG145 (SEQ ID NO: 980) is specific to this group of claims. • Invention 8 is defined by claims 21-23 (in full), 24-43 (in part) and are directed to endonuclease systems defined by RNA sequences (SEQ ID NOs: 199, 200, 460, 461, 669-673, 677-686, 1006-1012 or 1231-1259) configured to bind to an endonuclease. • Invention 9 is defined by claims 59, 61-73 (in full) and 106-112 (in part). The feature of an endonuclease selective for a protospacer adjacent motif (PAM) sequence is specific to this group of claims. • Invention 10 is defined by claims 44-58 (in part). The feature of an engineered nuclease system comprising an endonuclease configured to be selective for a target adjacent motif (TAM) selected from SEQ ID NOs: 1029-1033, is specific to these claims. • Invention 11 is defined by claims 74-83, 145-147 and 155-157 (in full) are directed to engineered single guide RNA polynucleotides configured to form a complex with an endonuclease, and uses thereof.	
Form PCT/ISA/210 (Supplemental Box) (July 2019)	

10

20

30

40

50

INTERNATIONAL SEARCH REPORT	International application No. PCT/US2022/080437
Supplemental Box	
• Invention 12 is defined by claims 102-104 (in full) and 105 (in part). These claims are directed to methods for binding, cleaving, marking or modifying a double-stranded DNA by contacting a DNA polynucleotide with an endonuclease system comprising an engineered guide RNA, wherein the DNA polynucleotide comprises a TAM. PCT Rule 13.2, first sentence, states that unity of invention is only fulfilled when there is a technical relationship among the claimed inventions involving one or more of the same or corresponding special technical features. PCT Rule 13.2, second sentence, defines a special technical feature as a feature which makes a contribution over the prior art. When there is no special technical feature common to all the claimed inventions there is no unity of invention. In the above groups of claims, the identified features may have the potential to make a contribution over the prior art but are not common to all the claimed inventions and therefore cannot provide the required technical relationship. The only feature common to all of the claimed inventions and which provides a technical relationship among them is a small RNA-guided DNA nuclease. However this feature does not make a contribution over the prior art because it is disclosed in: ALTAE-TRAN, H. et al., "The widespread IS200/IS605 transposon family encodes diverse programmable RNA-guided endonucleases", Science, 9 September 2021, Vol. 374, pp 57-65 Therefore in the light of this document this common feature cannot be a special technical feature. Therefore there is no special technical feature common to all the claimed inventions and the requirements for unity of invention are consequently not satisfied <i>a posteriori</i>. Form PCT/ISA/210 (Supplemental Box) (July 2019)	

INTERNATIONAL SEARCH REPORT		International application No. PCT/US2022/080437	
Information on patent family members			
This Annex lists known patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.			
Patent Document/s Cited in Search Report		Patent Family Member/s	
Publication Number	Publication Date	Publication Number	Publication Date
WO 2021/202568 A1	07 October 2021	WO 2021202568 A1	07 Oct 2021
		AU 2021248800 A1	13 Oct 2022
		CA 3177051 A1	07 Oct 2021
		EP 4127156 A1	08 Feb 2023
		GB 2608292 A	28 Dec 2022
		KR 20220161383 A	06 Dec 2022
		US 2023051396 A1	16 Feb 2023
WO 2020/055941 A1	19 March 2020	WO 2020055941 A1	19 Mar 2020
		US 2022290177 A1	15 Sep 2022
US 2020/0332273 A1	22 October 2020	US 2020332273 A1	22 Oct 2020
		US 10982200 B2	20 Apr 2021
		AU 2020223370 A1	07 Oct 2021
		AU 2021267379 A1	08 Dec 2022
		CA 3130135 A1	20 Aug 2020
		CA 3177828 A1	11 Nov 2021
		CN 113728098 A	30 Nov 2021
		EP 3924482 A1	22 Dec 2021
		GB 2594339 A	27 Oct 2021
		GB 2594339 B	08 Feb 2023
		JP 2022520428 A	30 Mar 2022
		KR 20210139254 A	22 Nov 2021
		KR 20230021657 A	14 Feb 2023
		US 2020332274 A1	22 Oct 2020
		US 10913941 B2	09 Feb 2021
		US 2022033791 A1	03 Feb 2022
		US 2022220460 A1	14 Jul 2022
WO 2020168291 A1	20 Aug 2020		
WO 2021226363 A1	11 Nov 2021		
End of Annex			
Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001. Form PCT/ISA/210 (Family Annex)(July 2019)			

10

20

30

40

50

フロントページの続き

(51)国際特許分類	F I	テーマコード (参考)
C 1 2 N 1/15 (2006.01)	C 1 2 N 1/15	
C 1 2 N 1/19 (2006.01)	C 1 2 N 1/19	
C 1 2 N 1/21 (2006.01)	C 1 2 N 1/21	
C 1 2 N 5/10 (2006.01)	C 1 2 N 5/10	
(32)優先日	令和4年6月29日(2022.6.29)	
(33)優先権主張国・地域又は機関	米国(US)	
(81)指定国・地域	AP(BW,GH,GM,KE,LR,LS,MW,MZ,NA,RW,SD,SL,ST,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,RU,TJ,TM),EP(AL,AT,BE,BG,CH,CY,CZ,DE,DK,EE,ES,FI,FR,GB,GR,HR,HU,IE,IS,IT,LT,LU,LV,MC,MK,MT,NL,NO,PL,PT,RO,RS,SE,SI,SK,SM,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,KM,ML,MR,NE,SN,TD,TG),AE,AG,AL,AM,AO,AT,AU,AZ,BA,BB,BG,BH,BN,BR,BW,BY,BZ,CA,CH,CL,CN,CO,CR,CU,CV,CV,CZ,DE,DJ,DK,DM,DO,DZ,EC,EE,EG,ES,FI,GB,GD,GE,GH,GM,GT,HN,HR,HU,ID,IL,IN,IQ,IR,IS,IT,JM,JO,JP,KE,KG,KH,KN,KP,KR,KW,KZ,LA,LC,LK,LR,LS,LU,LY,MA,MD,ME,MG,MK,MN,MW,MX,MY,MZ,NA,NG,NI,NO,NZ,OM,PA,PE,PG,PH,PL,PT,QA,RO,RS,RU,RW,SA,SC,SD,SE,SG,SK,SL,ST,SV,SY,TH,TJ,TM,TN,TR,TT,TZ,UA,UG,US,UZ,VC,VN,WS,ZA,ZM,ZW	
(72)発明者	ディヴォート, オードラ	
	アメリカ合衆国 9 4 6 0 8 カリフォルニア州 エメリービル パーク・アベニュー	1 5 4 5
(72)発明者	バターフィールド, クリスティーナ	
	アメリカ合衆国 9 4 6 0 8 カリフォルニア州 エメリービル パーク・アベニュー	1 5 4 5
(72)発明者	アレキサンダー, リサ	
	アメリカ合衆国 9 4 6 0 8 カリフォルニア州 エメリービル パーク・アベニュー	1 5 4 5
(72)発明者	ゴルツマン, ダニエラ エス. エー.	
	アメリカ合衆国 9 4 6 0 8 カリフォルニア州 エメリービル パーク・アベニュー	1 5 4 5
(72)発明者	ラモット, レベッカ	
	アメリカ合衆国 9 4 6 0 8 カリフォルニア州 エメリービル パーク・アベニュー	1 5 4 5
F ターム (参考)	4B065 AA01X AA57X AA72X AA90X AB01 BA02 CA31 CA44 CA46	