

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号  
特許第7572360号  
(P7572360)

(45)発行日 令和6年10月23日(2024.10.23)

(24)登録日 令和6年10月15日(2024.10.15)

(51)国際特許分類		F I	
C 1 2 N	15/55 (2006.01)	C 1 2 N	15/55 Z N A
C 1 2 N	15/09 (2006.01)	C 1 2 N	15/09 1 1 0
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
請求項の数 26 (全129頁) 最終頁に続く			
(21)出願番号	特願2021-538143(P2021-538143)	(73)特許権者	521282169
(86)(22)出願日	令和1年12月20日(2019.12.20)		ライフエディット セラピューティクス
(65)公表番号	特表2022-516130(P2022-516130 A)		, インコーポレイティド
(43)公表日	令和4年2月24日(2022.2.24)		アメリカ合衆国, ノース カロライナ 2
(86)国際出願番号	PCT/US2019/068079		7 7 0 1, ダーラム, モリス ストリート
(87)国際公開番号	WO2020/139783	(74)代理人	3 0 0, スイート 3 0 0
(87)国際公開日	令和2年7月2日(2020.7.2)		100099759
審査請求日	令和4年12月19日(2022.12.19)	(74)代理人	弁理士 青木 篤
(31)優先権主張番号	62/785,391	(74)代理人	100123582
(32)優先日	平成30年12月27日(2018.12.27)		弁理士 三橋 真二
(33)優先権主張国・地域又は機関	米国(US)	(74)代理人	100092624
(31)優先権主張番号	62/790,256		弁理士 鶴田 準一
(32)優先日	平成31年1月9日(2019.1.9)	(74)代理人	100114018
最終頁に続く		(74)代理人	弁理士 南山 知広
		(74)代理人	100117019
		最終頁に続く	

(54)【発明の名称】 遺伝子編集に有用なポリペプチドと利用方法

(57)【特許請求の範囲】

【請求項 1】

RNA誘導型ヌクレアーゼ(RGN)ポリペプチドをコードするポリヌクレオチドを含む核酸分子であって、前記ポリヌクレオチドが、配列番号1と少なくとも95%の配列同一性を有するアミノ酸配列を含むRGNポリペプチドをコードするヌクレオチド配列を含み、RGNポリペプチドをコードする前記ポリヌクレオチドが、このポリヌクレオチドによって異種のプロモータに機能可能に連結されている核酸分子。

【請求項 2】

前記RGNポリペプチドが、標的DNA配列にハイブリダイズ可能なガイドRNA(gRNA)に結合するとき、RNA誘導型配列に特異的なやり方で前記標的DNA配列に結合する、請求項1に記載の核酸分子。

【請求項 3】

前記RGNポリペプチドが配列番号1に記載のアミノ酸配列を含む、請求項1に記載の核酸分子。

【請求項 4】

前記RGNポリペプチドがヌクレアーゼドドであるか、ニッカーゼとして機能し、配列番号568の第10位及び/又は第840位のアミノ酸に対応する位置にアミノ酸置換を含む、並びに/あるいは、配列番号569又は553の第10位のアミノ酸に対応する位置にアミノ酸置換を含む、請求項1に記載の核酸分子。

【請求項 5】

置換されたアミノ酸がアラニンである、請求項4に記載の核酸分子。

【請求項6】

前記RGNポリペプチドが塩基編集ポリペプチドに機能可能に融合されている、請求項4に記載の核酸分子。

【請求項7】

請求項1～4及び6のいずれか1項に記載の核酸分子を含むベクター。

【請求項8】

ガイドRNAをコードする少なくとも1つのヌクレオチド配列をさらに含み、このガイドRNAが、配列番号2と少なくとも95%の配列同一性を有するヌクレオチド配列を含むCRISPR反復を含むCRISPR RNAを含み、前記CRISPR反復はトランス活性化CRISPR RNA (tracrRNA) のアンチ反復とハイブリダイズ可能である、請求項7に記載のベクター。

10

【請求項9】

前記tracrRNAが配列番号3と少なくとも95%の配列同一性を有するヌクレオチド配列を含み、前記tracrRNAの前記アンチ反復は前記CRISPR RNAの前記CRISPR反復とハイブリダイズ可能である、請求項8に記載のベクター。

【請求項10】

請求項1～4及び6のいずれか1項に記載の核酸分子または請求項7～9のいずれか1項に記載のベクターを含む細胞。

【請求項11】

標的DNA配列に結合させるためのシステムであって、

20

該システムは、

a) 1つ以上のガイドRNA (gRNA)、またはこれら1つ以上のgRNAをコードする1つ以上のヌクレオチド配列；及び

b) 配列番号1と少なくとも95%の配列同一性を有するアミノ酸配列を含むRNA誘導型ヌクレアーゼ (RGN) ポリペプチド、またはこのRGNポリペプチドをコードするヌクレオチド配列；

を含み、

前記1つ以上のgRNAをコードする前記ヌクレオチド配列と、前記RGNポリペプチドをコードする前記ヌクレオチド配列が、それぞれ、前記ヌクレオチド配列のそれぞれにとって異種のプロモータに機能可能に連結される、システム。

30

【請求項12】

前記1つ以上のgRNAが前記標的DNA配列にハイブリダイズ可能であり、かつ、

前記1つ以上のgRNAが前記RGNポリペプチドと複合体を形成可能であり、それにより前記RGNポリペプチドを前記標的DNA配列に結合させる、請求項11に記載のシステム。

【請求項13】

前記標的DNA配列が真核細胞の中にある、請求項11に記載のシステム。

【請求項14】

前記RGNポリペプチドがヌクレアーゼドドであるか、ニッカーゼとして機能し、配列番号568の第10位及び／又は第840位のアミノ酸に対応する位置にアミノ酸置換を含む、並びに／あるいは、配列番号569又は553の第10位のアミノ酸に対応する位置にアミノ酸置換を含む、かつ前記RGNポリペプチドが塩基編集ポリペプチドに機能可能に連結されている請求項11に記載のシステム。

40

【請求項15】

置換されたアミノ酸がアラニンである、請求項14に記載の核酸分子。

【請求項16】

1つ以上のドナーポリヌクレオチド、またはこれら1つ以上のドナーポリヌクレオチドをコードする1つ以上のヌクレオチド配列をさらに含み、前記1つ以上のドナーポリヌクレオチドをコードする前記ヌクレオチド配列が、それぞれ、前記ヌクレオチド配列のそれぞれにとって異種のプロモータに機能可能に連結されている、請求項11～13のいずれか1項に記載のシステム。

50

## 【請求項 17】

請求項11～14及び16のいずれか1項に記載のシステムを含む、標的DNA配列に結合させるための組成物。

## 【請求項 18】

標的DNA配列を切断および/または改変するための組成物であって、

a) RNA誘導型ヌクレアーゼ(RGN)ポリペプチド、ここで前記RGNは配列番号1と少なくとも95%の配列同一性を有するアミノ酸配列を含む；及び

b) a)のRGNポリペプチドが前記標的DNA配列を標的とさせることが可能な1つ以上のガイドRNA；  
を含む組成物。

10

## 【請求項 19】

前記1つ以上のガイドRNAが前記標的DNA配列にハイブリダイズし、それにより前記RGNポリペプチドを前記標的DNA配列に結合させ前記標的DNA配列を切断および/または改変させる、請求項18に記載の組成物。

## 【請求項 20】

改変された前記標的DNA配列が、前記標的DNA配列に対し少なくとも1個のヌクレオチドの欠失または変異を含む、請求項19に記載の組成物。

## 【請求項 21】

前記RGNポリペプチドがヌクレアーゼドドであるか、ニッカーゼとして機能し、配列番号568の第10位及び/又は第840位のアミノ酸に対応する位置にアミノ酸置換を含む、並びに/あるいは、配列番号569又は553の第10位のアミノ酸に対応する位置にアミノ酸置換を含み、かつ前記RGNポリペプチドが塩基編集ポリペプチドに機能可能に連結されている、請求項18に記載の組成物。

20

## 【請求項 22】

置換されたアミノ酸がアラニンである、請求項21に記載の核酸分子。

## 【請求項 23】

前記改変された標的DNA配列が、前記標的DNA配列への異種DNAの挿入を含む、請求項19に記載の組成物。

## 【請求項 24】

前記標的DNA配列が細胞の中にある、請求項18に記載の組成物。

30

## 【請求項 25】

前記細胞が真核細胞である、請求項24に記載の組成物。

## 【請求項 26】

前記細胞は、前記RGNポリペプチドが発現して前記標的DNA配列を切断し、改変されたDNA配列を生成させる条件下で更に培養され；そして前記改変されたDNA配列を含む細胞が選択される、請求項24または25に記載の組成物。

## 【発明の詳細な説明】

## 【技術分野】

## 【0001】

本発明は、分子生物学と遺伝子編集の分野に関する。

40

## 【背景技術】

## 【0002】

標的ゲノム編集または標的ゲノム改変が、急速に基礎研究と応用研究のための重要なツールになりつつある。初期の方法は、ヌクレアーゼ(メガヌクレアーゼ、ジンクフィンガー融合タンパク質、またはTALENなど)を操作することを含んでいたため、それぞれの特定の標的配列に対して特異的な操作されたプログラム可能な配列特異的DNA結合ドメインを持つキメラヌクレアーゼを作製する必要があった。RNA誘導型ヌクレアーゼ(クラスター化されて規則的な間隔で配置された短い回文反復(CRISPR)-cas細菌システムのCRISPR-関連(cas)タンパク質など)は、特定の標的配列に特異的にハイブリダイズするガイドRNAとの複合体にすることにより、特定の配列を標的とすることが可能になる。標的

50

異的ガイドRNAの作製は、各標的配列のためのキメラヌクレアーゼの作製よりもコストがかからず、より効率的である。このようなRNA誘導型ヌクレアーゼを用いて配列特異的二本鎖切断を導入することによりゲノムを編集することが可能であり、この二本鎖切断は、エラーを起こしやすい非相同末端結合（NHEJ）を通じて修復されるため、変異が特定のゲノム位置に導入される。あるいは相同組み換え修復を通じて異種DNAをそのゲノム部位に導入することができる。

#### 【0003】

それに加え、RGNはDNA編集アプローチにとって有用である。核酸配列の標的編集（例えば特定の改変をゲノムDNAの中に導入することを可能にする標的切断）によって遺伝子機能と遺伝子発現を研究するための高度にニュアンスに富んだアプローチが可能になる。このような標的編集を活用してヒトにおける遺伝子疾患を標的にすること、または農業で有益な変異を作物のゲノムに導入することもできる。ゲノム編集ツールの開発により、遺伝子編集に基づく哺乳動物用治療薬と農業バイオテクノロジーへの新たなアプローチが提供される。

#### 【発明の概要】

#### 【0004】

興味ある標的配列に結合させるための組成物と方法が提供される。組成物は、興味ある標的配列の切断または改変と、興味ある標的配列の可視化と、興味ある配列の発現変更に利用される。組成物は、RNA誘導型ヌクレアーゼ（RGN）ポリペプチドとそのバリエーション、CRISPR RNA（crRNA）、トランス活性化CRISPR RNA（tracrRNA）、ガイドRNA（gRNA）、デアミナーゼポリペプチド、これらをコードする核酸分子、およびこれら核酸分子を含むベクターと宿主細胞を含んでいる。興味ある標的配列に結合させるためのCRISPRシステムも提供され、このCRISPRシステムは、RNA誘導型ヌクレアーゼポリペプチドと1つ以上のガイドRNAを含んでいる。RNA誘導型DNA結合ポリペプチドとデアミナーゼポリペプチドを含む融合ポリペプチドも提供される。本明細書に開示されている方法は、興味ある標的核酸配列への結合に関するものであり、いくつかの実施態様では、興味あるその標的核酸配列の切断または改変に関する。興味ある標的配列を改変することが、例えば導入されるドナー配列を用いた非相同末端結合または相同組み換え修復の結果として、または塩基編集の結果として可能になる。

#### 【発明を実施するための形態】

#### 【0005】

本発明が関係する分野の当業者は、本明細書に記載されている本発明の改変およびそれ以外の実施態様として、上記の説明とそれに関連する図面に示されている教示の利点を持つものを多数思いつくであろう。したがって本発明が開示されている具体的な実施態様に限定されることはなく、改変やそれ以外の実施態様が添付の実施態様の範囲に含まれることを理解すべきである。特殊な用語を本明細書で使用するが、それらは一般的かつ説明的な意味で用いられているのであり、制限する目的ではない。

#### 【0006】

##### 1. 概要

#### 【0007】

RNA誘導型ヌクレアーゼ（RGN）はゲノム内の単一の部位を標的として操作することを可能にするため、治療と研究の用途のための遺伝子ターゲティングの文脈で有用である。哺乳動物を含む多彩な生物において、RNA誘導型ヌクレアーゼは、例えば非相同末端結合と相同組み換えを刺激することによるゲノム操作のために使用されてきた。本明細書に記載されている組成物と方法は、ポリヌクレオチドに一本鎖切断または二本鎖切断を生じさせること、ポリヌクレオチドを改変すること、ポリヌクレオチド内の特定の部位を検出すること、または特定の遺伝子の発現を変化させることに役立つ。

#### 【0008】

本明細書に開示されているRNA誘導型ヌクレアーゼは、標的配列を改変することによって遺伝子の発現を変化させることができる。特別な実施態様では、RNA誘導型ヌクレアー

ぜは、クラスター化されて規則的な間隔で配置された短い回文反復（CRISPR）RNA誘導型ヌクレアーゼシステムの一部としてのガイドRNA（gRNAまたはsgRNAとも呼ばれる）によって標的配列に向けられる。ガイドRNAはRNA誘導型ヌクレアーゼと複合体を形成してこのRNA誘導型ヌクレアーゼを標的配列と結合させ、いくつかの実施態様では、標的配列の位置に一本鎖切断または二本鎖切断を導入する。標的配列が切断された後、この切断を修復することができるため、修復プロセスの間に標的配列のDNA配列が改変される。したがって本明細書では、RNA誘導型ヌクレアーゼを用いて宿主細胞のDNAの中の標的配列を改変する方法が提供される。例えばRNA誘導型ヌクレアーゼを用いて真核細胞または原核細胞の1つのゲノム遺伝子座の位置にある標的配列を改変することができる。

#### 【0009】

本開示によりさらに、デアミナーゼポリペプチドとそれをコードする核酸分子のほか、DNA結合ポリペプチドとデアミナーゼポリペプチドを含む融合タンパク質が提供される。いくつかの実施態様では、DNA結合ポリペプチドは、メガヌクレアーゼ、ジンクフィンガー融合タンパク質、またはTALENであるか、これらに由来する。いくつかの実施態様では、融合タンパク質は、RNA誘導型DNA結合ポリペプチドとデアミナーゼポリペプチドを含んでいる。いくつかの実施態様では、RNA誘導型DNA結合ポリペプチドはRGNである。いくつかの実施態様では、RGNはII型CRISPR-Casポリペプチドである。別の実施態様では、RGNはV型CRISPR-Casポリペプチドである。さらに別の実施態様では、RGNは、gRNAに結合するCas9ポリペプチドドメインであり、このgRNAが今度は鎖ハイブリダイゼーションを通じて標的核酸配列に結合する。

#### 【0010】

デアミナーゼポリペプチドは、核酸塩基（例えばシチジンなど）を脱アミノ化することのできるデアミナーゼドメインを含んでいる。デアミナーゼによる核酸塩基の脱アミノ化で各残基の位置に点変異を生じさせることができる。これを本明細書では「核酸編集」または「塩基編集」と呼ぶ。したがってRGNポリペプチドバリエーションまたはRGNポリペプチドドメインとデアミナーゼドメインを含む融合タンパク質を用いて核酸配列を狙い通りに編集することができる。

#### 【0011】

このような融合タンパク質は、インビトロでDNAを狙い通りに編集して例えば変異細胞を生成させるのに有用である。これら変異細胞は植物または動物の中に存在することができる。このような融合タンパク質は、狙った変異を導入し、例えば生体外で哺乳動物の細胞（例えば対象から得られてその後同じ対象または別の対象に再導入される細胞）の中にある遺伝子の欠陥を修正することと；狙った変異を導入すること（例えば対象となる哺乳動物の細胞内の疾患関連遺伝子の中にある遺伝子の欠陥を修正すること、またはその遺伝子の中に脱活性化変異を導入すること）にも役立つ可能性がある。このような融合タンパク質は、植物細胞の中に狙った変異を導入し、例えば有益であるか農業で価値のある形質またはアレルを導入するのにも有用である可能性がある。

#### 【0012】

「タンパク質」、「ペプチド」、および「ポリペプチド」という用語は、本明細書では交換可能に用いられ、ペプチド（アミド）結合によって互いに連結された複数のアミノ酸残基のポリマーを意味する。これらの用語は、任意のサイズ、構造、または機能のタンパク質、ペプチド、またはポリペプチドを意味する。典型的には、タンパク質、ペプチド、またはポリペプチドは、長さが少なくとも3個のアミノ酸になる。タンパク質、ペプチド、またはポリペプチドは、個別のタンパク質、またはタンパク質の集団を意味することができる。1つのタンパク質、ペプチド、またはポリペプチドの中の1個以上のアミノ酸を改変することが、例えば化合物（炭化水素基、ヒドロキシル基、リン酸塩基、ファメシル基、イソファメシル基、脂肪酸基、複合体、機能化、またはそれ以外の改変などのためのリンカーなど）を付加することによって可能である。タンパク質、ペプチド、またはポリペプチドは、単一の分子も可能である、または多分子複合体が可能である。タンパク質、ペプチド、またはポリペプチドは、天然のタンパク質またはペプチドの単なる断片が可能

10

20

30

40

50

である。タンパク質、ペプチド、またはポリペプチドは、天然であること、組み換えであること、または合成であること、またはこれらの任意の組み合わせであることが可能である。

#### 【0013】

本明細書で用いられている「融合タンパク質」という用語は、少なくとも2つの異なるタンパク質からのタンパク質ドメインを含むハイブリッドポリペプチドを意味する。1つのタンパク質を融合タンパク質のアミノ末端（N末端）部分、またはカルボキシ末端（C末端）タンパク質に位置させることができるため、それぞれ「アミノ末端融合タンパク質」または「カルボキシ末端融合タンパク質」を形成する。タンパク質は、異なる複数のドメインを含むことができる（例えば1つのリコンビナーゼの核酸結合ドメイン（例えばCas9を標的部位に結合させるCas9のgRNA結合ドメイン）と、核酸切断ドメインまたは触媒ドメイン）。いくつかの実施態様では、タンパク質は、タンパク質性部分（例えば核酸結合ドメインを構成するアミノ酸配列）と有機化合物（例えば核酸切断剤として作用することのできる化合物）を含んでいる。いくつかの実施態様では、タンパク質は、核酸（例えばRNA）との複合体であるか、核酸と会合している。本明細書に提示されているどのタンパク質も、本分野で知られている任意の方法で作製することができる。例えば本明細書に提示されているタンパク質は、組み換えタンパク質の発現と精製を通じて作製することができる。これは、ペプチドリinkerを含む融合タンパク質に特に適している。組み換えタンパク質を発現させて精製する方法は周知であり、その中に含まれるのは、GreenとSambrook、『Molecular Cloning: A Laboratory Manual』（第4版、Cold Spring Harbor Laboratory Press、コールド・スプリング・ハーバー、ニューヨーク州（2012年）；その内容全体が参照によって本明細書に組み込まれている）に記載されている方法である。

#### 【0014】

### II. RNA誘導型ヌクレアーゼ

#### 【0015】

本明細書にはRNA誘導型ヌクレアーゼが提示されている。RNA誘導型ヌクレアーゼ（RGN）という用語は、配列特異的なやり方で特定の標的ヌクレオチド配列に結合するポリペプチドを意味し、このポリペプチドとの複合体を形成したガイドRNA分子によって標的ヌクレオチド配列に向けられて、この標的配列とハイブリダイズする。RNA誘導型ヌクレアーゼは、標的配列に結合したときにこの標的配列を切断できるが、RNA誘導型ヌクレアーゼという用語には、標的配列に結合できるが標的配列の切断はできないヌクレアーゼ-デッドRNA誘導型ヌクレアーゼも包含される。RNA誘導型ヌクレアーゼによって標的配列を切断する結果として、一本鎖を切断すること、または二本鎖を切断することができる。二本鎖核酸分子の単一の鎖を切断することのできるRNA誘導型ヌクレアーゼを本明細書ではニックナーゼと呼ぶ。

#### 【0016】

本明細書に開示されているRNA誘導型ヌクレアーゼには、APG00969、APG03128、APG09748、APG00771、およびAPG02789というRNA誘導型ヌクレアーゼ（そのアミノ酸配列は、それぞれ配列番号1、16、24、35、43、または50で表わされる）と、その活性な断片またはバリエーションでRNA誘導型配列に特異的なやり方で標的ヌクレオチド配列に結合する能力を保持しているものが含まれる。これら実施態様のいくつかでは、RGN APG00969、APG03128、APG09748、APG00771、およびAPG02789の活性な断片またはバリエーションは、一本鎖標的配列または二本鎖標的配列を切断することができる。いくつかの実施態様では、RGN APG00969、APG03128、APG09748、APG00771、またはAPG02789の活性なバリエーションは、配列番号1、16、24、35、43、または50として表わされるアミノ酸配列と配列が少なくとも40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、またはそれよりも多く一致するアミノ酸配列を含んでいる。いくつかの実施態様では、RGN APG00969、APG03128、APG09748、APG00771、またはAPG02789の活性な断片は、配列番号1、16、24、35、43、または50として表わされるアミノ

10

20

30

40

50

酸配列の少なくとも50個、100個、150個、200個、250個、300個、350個、400個、450個、500個、550個、600個、650個、700個、750個、800個、850個、900個、950個、1000個、1050個、またはそれよりも多い個数の連続したアミノ酸残基を含んでいる。本明細書に提示されているRNA誘導型ヌクレアーゼは、少なくとも1つのヌクレアーゼドメイン（例えばDNアーゼドメイン、RNアーゼドメイン）と、ガイドRNAと相互作用する少なくとも1つのRNA認識ドメインおよび/またはRNA結合ドメインを含むことができる。本明細書に提示されているRNA誘導型ヌクレアーゼの中に見いだすことができるさらに別のドメインの非限定的な例に含まれるのは、DNA結合ドメイン、ヘリカーゼドメイン、タンパク質-タンパク質相互作用ドメイン、および二量体化ドメインである。特別な実施態様では、本明細書に提示されているRNA誘導型ヌクレアーゼは、DNA結合ドメイン、ヘリカーゼドメイン、タンパク質-タンパク質相互作用ドメイン、および二量体化ドメインのうちの1つ以上の少なくとも70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%を含むことができる。

#### 【0017】

標的ヌクレオチド配列には、本明細書に提示されているRNA誘導型ヌクレアーゼが結合し、このRNA誘導型ヌクレアーゼが会合したガイドRNAがハイブリダイズする。その後、RNA誘導型ヌクレアーゼがヌクレアーゼ活性を有する場合には、標的配列をこのRNA誘導型ヌクレアーゼによって切断することができる。「切断する」または「切断」という用語は、標的配列の骨格内の少なくとも1つのホスホジエステル結合が加水分解され、その結果として標的核酸内で一本鎖または二本鎖を切断できることを意味する。本開示のRGNは、エンドヌクレアーゼとして機能してポリヌクレオチド内のヌクレオチドを切断すること、またはエキソヌクレアーゼとなってポリヌクレオチドの末端（5'末端および/または3'末端）から次々とヌクレオチドを除去することができる。別の実施態様では、開示されているRGNはポリヌクレオチドの任意の位置にある標的配列のヌクレオチドを切断できるため、エンドヌクレアーゼとエキソヌクレアーゼの両方として機能することができる。本開示のRGNによって標的ポリヌクレオチドが切断される結果として粘着末端または平滑末端になる可能性がある。

#### 【0018】

本明細書に開示されているRNA誘導型ヌクレアーゼとして、細菌または古細菌の種に由来する野生型配列が可能である。あるいはRNA誘導型ヌクレアーゼとして、野生型ポリペプチドのバリエーションまたは断片が可能である。野生型RGNを改変して例えばヌクレアーゼ活性またはPAM特異性を変化させることができる。いくつかの実施態様では、RNA誘導型ヌクレアーゼは天然ではない。

#### 【0019】

いくつかの実施態様では、RNA誘導型ヌクレアーゼはニッカーゼとして機能し、標的ヌクレオチド配列の単一の鎖だけを切断する。このようなRNA誘導型ヌクレアーゼは単一機能のヌクレアーゼドメインを持つ。これら実施態様のいくつかでは、追加のヌクレアーゼドメインが変異していたため、ヌクレアーゼ活性が低下するか消失している。ヌクレアーゼ不活性なRGN、またはニッカーゼRGNは、RNA誘導DNA結合ポリペプチド、RNA誘導型DNA結合タンパク質、または融合タンパク質のRNA誘導型DNA結合ドメインと呼ぶことができる。

#### 【0020】

別の実施態様では、RNA誘導型ヌクレアーゼはヌクレアーゼ活性が完全に消失しているか、低下したヌクレアーゼ活性を示すため、本明細書ではヌクレアーゼ-デッド（不活化）と呼ぶ。アミノ酸配列に変異を導入するための本分野で知られている任意の方法（例えばPCRを媒介とした突然変異誘発と、部位指定突然変異誘発）を利用してニッカーゼRGN、またはヌクレアーゼ-デッドRGNを生成させることができる。例えばアメリカ合衆国特許出願公開第2014/0068797号とアメリカ合衆国特許第9,790,490号を参照されたい（そのそれぞれの全体が、参照によって本明細書に組み込まれている）。

#### 【0021】

10

20

30

40

50

ヌクレアーゼ活性を欠くRNA誘導型ヌクレアーゼを用い、融合したポリペプチド、ポリヌクレオチド、または小分子ペイロードを特定のゲノム位置に送達することができる。これら実施態様のいくつかでは、RGNポリペプチドまたはガイドRNAを検出可能な標識に融合させ、特定の配列を検出できるようにする。非限定的な一例として、ヌクレアーゼ-デッドRGNを検出可能な標識（例えば蛍光タンパク質）に融合させ、疾患に関する特定の配列を標的とすることで、その疾患関連配列の検出が可能になる。

#### 【0022】

あるいはヌクレアーゼ-デッドRGNを特定のゲノム位置に向かわせて望む配列の発現を変化させることができる。いくつかの実施態様では、ヌクレアーゼ-デッドRGN誘導型ヌクレアーゼが標的配列に結合すると、標的とされるゲノム領域内でのRNAポリメラーゼまたは転写因子の結合に干渉することにより、標的配列の発現、または標的配列による転写制御下にある遺伝子の発現が抑制される。別の実施態様では、RGN（例えばヌクレアーゼ-デッドRGN）、またはRGNとの複合体になったガイドRNAは、発現モジュレータをさらに含んでいて、この発現モジュレータは、標的配列に結合すると、標的配列の発現、または標的配列による転写制御下にある遺伝子の発現を抑制するか活性化する。これら実施態様のいくつかでは、発現モジュレータは、標的配列の発現、またはエピジェネティックな機構を通じて調節された遺伝子の発現を変化させる。

#### 【0023】

別の実施態様では、ヌクレアーゼ-デッドRGN、またはニックアーゼ活性だけを有するRGNを特定のゲノム位置に向かわせ、塩基編集ポリペプチド（例えばデアミナーゼポリペプチド、またはその活性なバリエーションまたは断片でヌクレオチド塩基を脱アミノ化するもの）に融合させることを通じて標的ポリヌクレオチドの配列を改変することができるため、その結果として1個のヌクレオチド塩基が別のヌクレオチド塩基に変換される。塩基編集ポリペプチドは、RGNにそのN末端またはC末端の位置で融合させることができる。それに加え、塩基編集ポリペプチドは、ペプチドリンカーを介してRGNに融合させることができる。このような組成物と方法で有用なデアミナーゼポリペプチドの非限定的な一例に含まれるのは、Gaudelli他（2017年）Nature 第551巻：464～471ページ、アメリカ合衆国特許出願公開第2017/0121693号と第2018/0073012号、および国際出願公開第WO/2018/027078号（そのそれぞれの全体が参照によって本明細書に組み込まれている）に記載されているシチジンデアミナーゼまたはアデノシンデアミナーゼ塩基エディタである。

#### 【0024】

ポリペプチドまたはドメインに融合されるRNA誘導型ヌクレアーゼは、リンカーによって隔てること、または接合することができる。本明細書で用いられている「リンカー」という用語は、2つの分子または部分（例えばヌクレアーゼの結合ドメインと切断ドメイン）を連結する化学基または分子を意味する。いくつかの実施態様では、リンカーは、RNA誘導型ヌクレアーゼのgRNA結合ドメインと塩基編集ポリペプチド（デアミナーゼなど）を接合する。いくつかの実施態様では、リンカーは、ヌクレアーゼ-デッドRGNとデアミナーゼを連結する。典型的には、リンカーは、2つの基、分子、またはそれ以外の部分の間に、またはこれらに隣接して位置しており、共有結合を通じてそれぞれに接続されているため、その2つを接続している。いくつかの実施態様では、リンカーは、1個のアミノ酸、または複数個のアミノ酸（例えばペプチドまたはタンパク質）である。いくつかの実施態様では、リンカーは、有機分子、基、ポリマー、または化学的部分である。いくつかの実施態様では、リンカーは、長さが5～100個のアミノ酸、例えば長さが5個、6個、7個、8個、9個、10個、11個、12個、13個、14個、15個、16個、17個、18個、19個、20個、21個、22個、23個、24個、25個、26個、27個、28個、29個、30個、30～35個、35～40個、40～45個、45～50個、50～60個、60～70個、70～80個、80～90個、90～100個、100～150個、または150～200個のアミノ酸である。より長い、またはより短いリンカーも考えられる。

#### 【0025】



本開示のRNA誘導型ヌクレアーゼは、細胞の核へのRGNの輸送を増強するため少なくとも1つの核局在化シグナル(NLS)を含むことができる。核局在化シグナルは本分野で知られており、一般に一連の塩基性アミノ酸を含んでいる(例えばLange他、J. Biol. Chem. (2007年)第282巻:5101~5105ページを参照されたい)。特別な実施態様では、RGNは、2つ、3つ、4つ、5つ、6つ、またはそれよりも多数の核局在化シグナルを含んでいる。核局在化シグナルとして、異種NLSが可能である。本開示のRGNにとって有用な核局在化シグナルの非限定的な例は、SV40大T抗原、ヌクレオパスミン、およびc-Mycの核局在化シグナルである(例えばRay他(2015年)Bioconjug Chem 第26巻(6):1004~1007ページを参照されたい)。特別な実施態様では、RGNは、配列番号10として示されているNLS配列を含んでいる。RGNは、そのN末端、C末端、またはN末端とC末端の両方に1つ以上のNLS配列を含むことができる。例えばRGNは、N末端領域に位置する2つのNLS配列と、C末端領域に位置する4つのNLS配列を含むことができる。

10

#### 【0026】

RGNを標的とするのに、ポリペプチドを細胞内の特定の位置に局在させる本分野で知られた他の局在化シグナルも用いることができ、その非限定的な例に含まれるのは、プラスチド(色素体)局在化配列、ミトコンドリア局在化配列、およびプラスチドとミトコンドリアの両方を標的とする二重標的シグナル配列である(例えばNassouryとMorse(2005年)Biochim Biophys Acta第1743巻:5~19ページ;KunzeとBerger(2015年)Front Physiol dx.doi.org/10.3389/fphys.2015.00259;HerrmannとNeupert(2003年)IUBMB Life第55巻:219~225ページ;Soll(2002年)Curr Opin Plant Biol 第5巻:529~535ページ;CarrieとSmall(2013年)Biochim Biophys Acta第1833巻:253~259ページ;Carrie他(2009年)FEBS J第276巻:1187~1195ページ;Silva-Filho(2003年)Curr Opin Plant Biol第6巻:589~595ページ;PeetersとSmall(2001年)Biochim Biophys Acta第1541巻:54~63ページ;Murtha他(2014年)J Exp Bot第65巻:6301~6335ページ;Mackenzie(2005年)Trends Cell Biol 第15巻:548~554ページ;Glaser他(1998年)Plant Mol Biol第38巻:311~338ページを参照されたい)。

20

#### 【0027】

いくつかの実施態様では、本開示のRNA誘導型ヌクレアーゼは、細胞へのRGNの取り込みを容易にする少なくとも1つの細胞侵入ドメインを含んでいる。細胞侵入ドメインは本分野で知られており、一般に、一連の正に帯電したアミノ酸残基(すなわちポリカチオン性細胞侵入ドメイン)、交互に存在する極性アミノ酸残基と非極性アミノ酸残基(すなわち両親媒性細胞侵入ドメイン)、または疎水性アミノ酸残基(すなわち疎水性細胞侵入ドメイン)を含んでいる(例えばMilletti F.(2012年)Drug Discov Today 第17巻:850~860ページを参照されたい)。細胞侵入ドメインの非限定的な一例は、ヒト免疫不全ウイルス1からのトランス活性化転写アクチベータ(TAT)である。

30

#### 【0028】

核局在化シグナル、プラスチド局在化配列、ミトコンドリア局在化配列、二重標的シグナル配列、および/または細胞侵入ドメインは、RNA誘導型ヌクレアーゼのアミノ末端(N末端)、カルボキシル末端(C末端)、または内部位置に位置させることができる。

40

#### 【0029】

本開示のRGNは、エフェクタドメイン(切断ドメイン、デアミナーゼドメイン、または発現モジュレータドメインなど)に直接的に融合させるか、リンカーペプチドを介して間接的に融合させることができる。このようなドメインは、RNA誘導型ヌクレアーゼのN末端、C末端、または内部位置に位置させることができる。これら実施態様のいくつかでは、融合タンパク質のRGN要素は、ヌクレアーゼ-デッドRGNである。

#### 【0030】

いくつかの実施態様では、RGN融合タンパク質は切断ドメインを含んでいる。切断ドメインは、ポリヌクレオチド(すなわちRNA、DNA、またはRNA/DNAハイブリッド)を切断することのできる任意のドメインであり、その非限定的な例に含まれるのは、制限エン

50

ドヌクレアーゼとホーミングエンドヌクレアーゼ（IIS型エンドヌクレアーゼなど（例えば FokI））である（例えば Belfort 他（1997 年）Nucleic Acids Res. 第 25 巻：3379～3388 ページ；Linn 他（編）『Nucleases』、Cold Spring Harbor Laboratory Press、1993 年を参照されたい）。

【0031】

別の実施態様では、RGN 融合タンパク質は、ヌクレオチド塩基を脱アミノ化して 1 個のヌクレオチド塩基を別のヌクレオチド塩基に変換するデアミナーゼドメインを含んでおり、その非限定的な例に含まれるのは、シチジンデアミナーゼ塩基エディタまたはアデノシンデアミナーゼ塩基エディタである（例えば Gaudelli 他（2017 年）Nature 第 551 巻：464～471 ページ、アメリカ合衆国特許出願公開第 2017/0121693 号と第 2018/0073012 号、アメリカ合衆国特許第 9,840,699 号、および国際出願公開第 WO/2018/027078 号を参照されたい）。さらに別の実施態様では、RGN 融合タンパク質は本発明のデアミナーゼを含むことができ、このデアミナーゼは、配列番号 374～545 または 572～584 のいずれか 1 つのアミノ酸配列、またはその活性なバリエーションを含んでいる。

【0032】

別の実施態様では、本発明のデアミナーゼは、配列番号 374～545 または 572～584 のいずれか 1 つのアミノ酸配列、またはその活性なバリエーションを含んでおり、任意の DNA 結合タンパク質に融合させることができる。いくつかの実施態様では、デアミナーゼは本発明の RGN に融合される。別の実施態様では、デアミナーゼは本分野で知られている RGN に融合される。別の実施態様では、デアミナーゼは、RGN ではない DNA 結合タンパク質（例えば メガヌクレアーゼ、TALEN、または ジンクフィンガーヌクレアーゼなど）に融合される。いくつかの実施態様では、デアミナーゼは、配列番号 374～545 と 572～584 のいずれかのアミノ酸配列と少なくとも 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 95%、少なくとも 96%、少なくとも 97%、少なくとも 98%、または少なくとも 99% 一致するアミノ酸配列を持つ。いくつかの実施態様では、デアミナーゼは、配列番号 374、383、397、399、407、408、411、414、416、420、514、および 572～584 のいずれかのアミノ酸配列と少なくとも 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 95%、少なくとも 96%、少なくとも 97%、少なくとも 98%、または少なくとも 99% 一致するアミノ酸配列を持つ。これら実施態様のいくつかでは、バリエーションデアミナーゼポリペプチドは、配列番号 572～584 のいずれか 1 つとの配列一致があるレベルであり、特定のアミノ酸残基が親配列から変化していない。例えばいくつかの実施態様では、配列番号 572 の 1 つのバリエーションは、配列番号 572 の 102 位に対応する位置にリシンを含み、配列番号 572 の 104 位に対応する位置にチロシンを含み、配列番号 572 の 106 位に対応する位置にトレオニンを含んでいる。特別な実施態様では、配列番号 574 の 1 つのバリエーションは、配列番号 574 の 101 位に対応する位置にグルタミン酸を含み、配列番号 574 の 103 位に対応する位置にセリンを含み、配列番号 574 の 105 位に対応する位置にリシンを含んでいる。いくつかの実施態様では、配列番号 575 の 1 つのバリエーションは、配列番号 575 の 101 位に対応する位置にリシンを含み、配列番号 575 の 103 位に対応する位置にロイシンを含み、配列番号 575 の 105 位に対応する位置にグルタミン酸を含んでいる。いくつかの実施態様では、配列番号 576 の 1 つのバリエーションは、配列番号 576 の 105 位に対応する位置にアラニンを含み、配列番号 576 の 107 位に対応する位置にアルギニンを含んでいる。特別な実施態様では、配列番号 577 の 1 つのバリエーションは、配列番号 577 の 102 位に対応する位置にグリシンを含み、配列番号 577 の 104 位に対応する位置にセリンを含み、配列番号 577 の 106 位に対応する位置にアルギニンを含んでいる。いくつかの実施態様では、配列番号 578 の 1 つのバリエーションは、配列番号 578 の 105 位に対応する位置にセリンを含み、配列番号 578 の 107 位に対応する位置にトレオニンを含んでいる。いくつかの実施態様では、配列番号 579 の 1 つのバリエーションは、配列番号 579 の 102 位に対応する位置にセリンを含み、配列番号 579 の 104 位に対応す

10

20

30

40

50

る位置にグルタミンを含み、配列番号579の106位に対応する位置にグリシンを含んでいる。特別な実施態様では、配列番号580の1つのバリエーションは、配列番号580の111位に対応する位置にグリシンを含んでいる。いくつかの実施態様では、配列番号581の1つのバリエーションは、配列番号581の104位に対応する位置にグルタミンを含み、配列番号581の106位に対応する位置にグリシンを含み、配列番号581の108位に対応する位置にグルタミン酸を含んでいる。いくつかの実施態様では、配列番号582の1つのバリエーションは、配列番号582の102位に対応する位置にアルギニンを含み、配列番号582の104位に対応する位置にトリプトファンを含み、配列番号582の106位に対応する位置にグルタミン酸を含んでいる。いくつかの実施態様では、配列番号583の1つのバリエーションは、配列番号583の104位に対応する位置にアルギニンを含み、配列番号583の106位に対応する位置にセリンを含んでいる。特別な実施態様では、配列番号584の1つのバリエーションは、配列番号584の110位に対応する位置にフェニルアラニンを含み、配列番号584の112位に対応する位置にセリンを含み、配列番号584の114位に対応する位置にトレオニンを含んでいる。

10

#### 【0033】

「デアミナーゼ」という用語は、脱アミノ化反応（すなわちアミノ酸または他の化合物からのアミノ基の除去）の触媒となる酵素を意味する。いくつかの実施態様では、デアミナーゼはシチジンデアミナーゼであり、シチジンまたはデオキシシチジンが加水分解によって脱アミノ化されてウラシルまたはデオキシウラシルになるときの触媒となる。別の実施態様では、デアミナーゼはアデニンデアミナーゼである。アデニンの脱アミノ化によってイノシンが生成し、これがポリメラーゼによってグアニンとして取り扱われる。シチジンデアミナーゼとアデニンデアミナーゼはDNAまたはRNAに対して作用することができるが、これまでのところDNAの中でアデニンを脱アミノ化する天然のアデニンデアミナーゼは知られていない。しかし本明細書には、DNAに対する大きなアデノシンデアミナーゼ活性を持つ天然のタンパク質であるAPG07458（配列番号514）が開示されている。本発明のデアミナーゼは、DNA分子またはRNA分子の編集に用いることができる。本発明のデアミナーゼは、1つのグループとして、DNA分子の中での4つの転位（CからTへ、AからGへ、TからCへ、およびGからAへ）すべてと、RNA分子の中でのCからUへ、AからGへ、およびGからAへという転位をプログラム可能に設定することが可能である。

20

#### 【0034】

本発明のデアミナーゼは一本鎖核酸分子に対して作用する。標的鎖に対するニッカーゼ活性を持つRGNが標的鎖にニック（切れ目）を入れる一方で、それと相補的な非標的鎖はデアミナーゼによって改変される。細胞DNA修復機構は、ニックが入れられた標的鎖を改変されていない非標的鎖を鋳型として用いて修復することにより、変異をDNAの中に導入することができる。

30

#### 【0035】

いくつかの実施態様では、デアミナーゼに融合したヌクレアーゼ不活性なRGNまたはニッカーゼRGNを特定のゲノム位置に向かわせて望む配列の発現を変化させることができる。いくつかの実施態様では、このヌクレアーゼ不活性なRGNまたはニッカーゼRGNを融合タンパク質のRNA誘導型DNA結合ポリペプチド、またはRNA誘導型DNA結合タンパク質、またはタンパク質ドメインと呼ぶことができる。いくつかの実施態様では、この融合タンパク質が標的配列に結合するとヌクレオチド塩基が脱アミノ化され、その結果として1個のヌクレオチド塩基が別のヌクレオチド塩基に変換される。いくつかの実施態様では、RGN融合タンパク質のエフェクタドメインとして発現モジュレータドメインが可能である。発現モジュレータドメインは、転写を上方調節または下方調節する機能を持つドメインである。発現モジュレータドメインとして、エピジェネティックな修飾ドメイン、転写抑制ドメイン、または転写活性化ドメインが可能である。

40

#### 【0036】

これら実施態様のいくつかでは、RGN融合タンパク質の発現モジュレータは、DNAまたはヒストンタンパク質の共有結合を改変してDNA配列を変化させることなくヒストンの構

50

造および/または染色体の構造を変化させ、遺伝子発現の変化(すなわち上方調節または下方調節)に至らせるエピジェネティックな修飾ドメインを含んでいる。エピジェネティックな改変の非限定的な例に含まれるのは、DNA内のリシン残基のアセチル化またはメチル化、アルギニンのメチル化、セリンとトレオニンのリン酸化、およびリシンのユビキチン化とヒストンタンパク質のSUMO化、およびシトシン残基のメチル化とヒドロキシメチル化である。エピジェネティックな改変ドメインの非限定的な例に含まれるのは、ヒストンアセチルトランスフェラーゼドメイン、ヒストンデアセチラーゼドメイン、ヒストンメチルトランスフェラーゼドメイン、ヒストンデメチラーゼドメイン、DNAメチルトランスフェラーゼドメイン、およびDNAデメチラーゼドメインである。

【0037】

10

別の実施態様では、融合タンパク質の発現モジュレータは転写抑制ドメインを含んでいる。転写抑制ドメインは転写制御エレメントおよび/または転写調節タンパク質(RNAポリメラーゼ、転写因子など)と相互作用して少なくとも1つの遺伝子の転写を減少または停止させる。転写抑制ドメインは本分野で知られており、その非限定的な例に含まれるのは、Sp1様リプレッサドメイン、I Bドメイン、およびクリュッペル関連ボックス(KRAB)ドメインである。

【0038】

さらに別の実施態様では、融合タンパク質の発現モジュレータは転写活性化ドメインを含んでいる。転写活性化ドメインは転写制御エレメントおよび/または転写調節タンパク質(例えばRNAポリメラーゼと転写因子)と相互作用して少なくとも1つの遺伝子の転写を増加または活性化させる。転写活性化ドメインは本分野で知られており、その非限定的な例に含まれるのは、単純ヘルペスウイルスVP16活性化ドメインとNFAT活性化ドメインである。

20

【0039】

本開示のRGNポリペプチドとデアミナーゼポリペプチド、またはその融合ポリペプチドは、検出可能な標識または精製タグを含むことができる。検出可能な標識または精製タグは、RNA誘導型ヌクレアーゼのN末端、C末端、または内部位置に、直接位置させること、またはリンカーペプチドを介して間接的に位置させることができる。これら実施態様のいくつかでは、融合タンパク質のRGN要素は、ヌクレアーゼ-デッドRGNである。別の実施態様では、融合タンパク質のRGN要素は、ニックアーゼ活性を有するRGNである。

30

【0040】

検出可能な標識は、可視化すること、または別のやり方で観測することができる。検出可能な標識は、融合タンパク質(例えば蛍光タンパク質)としてRGNに融合させることができる。あるいは検出可能な標識としてRGNポリペプチドとの複合体になった小分子が可能であり、視覚的に、または他の手段によって検出することができる。融合タンパク質として本開示のRGNまたはデアミナーゼに融合させることのできる検出可能な標識には、検出可能な任意のタンパク質ドメインが含まれ、その非限定的な例に含まれるのは、蛍光タンパク質、または特異的抗体を用いて検出できるタンパク質ドメインである。蛍光タンパク質の非限定的な例に含まれるのは、緑色蛍光タンパク質(例えばGFP、EGFP、ZsGreen1)と黄色蛍光タンパク質(例えばYFP、EYFP、ZsYellow1)である。検出可能な小分子標識の非限定的な例に含まれるのは、放射性標識である<sup>3</sup>Hと<sup>35</sup>Sなどである。

40

【0041】

本発明のRGNとデアミナーゼポリペプチド、またはその融合ポリペプチドは精製タグも含むことができる。精製タグは、混合物(例えば生物サンプル、培地)からタンパク質または融合タンパク質を単離するのに使用できる任意の分子である。精製タグの非限定的な例に含まれるのは、myc、マルトース結合タンパク質(MBP)、およびグルタチオン-S-トランスフェラーゼ(GST)である。

【0042】

II. ガイドRNA

【0043】

50

本開示により、ガイドRNAと、それをコードするポリヌクレオチドが提供される。「ガイドRNA」という用語は、標的ヌクレオチド配列と十分な相補性を有するためにその標的配列とハイブリダイズして関係するRNA誘導型ヌクレアーゼを標的ヌクレオチド配列に配列特異的に結合させるヌクレオチド配列を意味する。したがってRGNの各ガイドRNAは、RGNに結合してこのRGNをガイドし、特定の標的ヌクレオチド配列に結合させることのできる1個以上のRNA分子（一般に1個または2個）であり、RGNがニッカーゼ活性またはヌクレアーゼ活性を持つ場合には、標的ヌクレオチド配列も切断する。一般に、ガイドRNAは、CRISPR RNA (crRNA) とトランス活性化CRISPR RNA (tracrRNA) を含んでいる。crRNAとtracrRNAの両方を含む天然のガイドRNAは、一般に、2つの別々のRNA分子を含んでいて、これらRNA分子は、crRNAの反復配列とtracrRNAのアンチ反復配列を通じて互いにハイブリダイズする。

10

#### 【0044】

CRISPRアレイに含まれる天然の直列反復配列は、一般に長さが28～37塩基対の範囲だが、この長さは約23 bp～約55 bpの間で変化する可能性がある。CRISPRアレイに含まれるスペーサ配列は、一般に長さが32～38 bpの範囲だが、この長さは約21 bp～約72 bpの間で変化する可能性がある。各CRISPRアレイは一般に50単位未満のCRISPR反復スペーサ配列を含んでいる。CRISPRは、一次CRISPR転写産物と呼ばれる長い転写産物の一部として転写される。一次CRISPR転写産物はCRISPRアレイの多くを含んでいる。一次CRISPR転写産物はCasタンパク質によって切断されてcrRNAを生成させるが、いくつかの場合にはプレcrRNAを生成させる。プレcrRNAは追加のCasタンパク質によってさらに処理されて成熟crRNAになる。成熟crRNAは、スペーサ配列とCRISPR反復配列を含んでいる。プレcrRNAが処理されて成熟（または処理された）crRNAになるいくつかの実施態様では、成熟に、約1個～約6個以上の5'ヌクレオチド、3'ヌクレオチド、または5'ヌクレオチドと3'ヌクレオチドを除去することが含まれる。興味ある特定の標的ヌクレオチド配列のゲノム編集またはゲノムターゲティングの目的では、プレcrRNA分子が成熟する間に除去されるヌクレオチドは、ガイドRNAの生成または設計に必要とされない。

20

#### 【0045】

CRISPR RNA (crRNA) は、スペーサ配列とCRISPR反復配列を含んでいる。「スペーサ配列」は、興味ある標的ヌクレオチド配列に直接ハイブリダイズするヌクレオチド配列である。スペーサ配列を操作して、興味ある標的ヌクレオチド配列と完全に、または部分的に相補的であるようにする。さまざまな実施態様では、スペーサ配列は、約8個～約30個、またはそれよりも多い個数のヌクレオチドを含むことができる。例えばスペーサ配列は、長さが約8個、約9個、約10個、約11個、約12個、約13個、約14個、約15個、約16個、約17個、約18個、約19個、約20個、約21個、約22個、約23個、約24個、約25個、約26個、約27個、約28個、約29個、約30個、またはそれよりも多い個数のヌクレオチドであることが可能である。いくつかの実施態様では、スペーサ配列は、長さが約10個～約26個のヌクレオチド、または長さが約12個～約30個のヌクレオチドである。特別な実施態様では、スペーサ配列は長さが約30個のヌクレオチドである。いくつかの実施態様では、スペーサ配列とそれに対応する標的配列の間の相補性の程度は、適切なアラインメントプログラムを用いて最適なアラインメントにしたとき、約50%以上、約60%以上、約70%以上、約75%以上、約80%以上、約81%以上、約82%以上、約83%以上、約84%以上、約85%以上、約86%以上、約87%以上、約88%以上、約89%以上、約90%以上、約91%以上、約92%以上、約93%以上、約94%以上、約95%以上、約96%以上、約97%以上、約98%以上、約99%以上、またはそれよりも大きい。特別な実施態様では、スペーサ配列は二次構造が自由であり、それは、本分野で知られている適切な任意のポリヌクレオチド折り畳みアルゴリズムを用いて予測することができる。アルゴリズムの非限定的な例に含まれるのは、mFold（例えばZukerとStiegler（1981年）Nucleic Acids Res. 第9巻：133～148ページを参照されたい）と、RNAfold（例えばGruber他（2008年）Cell第106巻（1）：23～24ページを参照されたい）である。

30

40

#### 【0046】

50

RGNTANパク質は、gRNA内のスペーサ配列とgRNAの標的配列の間のミスマッチに対する感度がさまざまである可能性があり、この感度が切断の効率に影響を与える。

【0047】

CRISPR RNA反復配列は、tracrRNAにハイブリダイズするのに十分な相補性を持つ領域を含むヌクレオチド配列を含んでいる。さまざまな実施態様では、CRISPR RNA反復配列は、約8個～約30個、またはそれよりも多い個数のヌクレオチドを含むことができる。例えばCRISPR反復配列は、長さが約8個、約9個、約10個、約11個、約12個、約13個、約14個、約15個、約16個、約17個、約18個、約19個、約20個、約21個、約22個、約23個、約24個、約25個、約26個、約27個、約28個、約29個、約30個、またはそれよりも多い個数のヌクレオチドであることが可能である。いくつかの実施態様では、CRISPR反復配列は、長さが約21個のヌクレオチドである。いくつかの実施態様では、CRISPR反復配列とそれに対応するtracrRNA配列の間の相補性の程度は、適切なアラインメントプログラムを用いて最適なアラインメントにしたとき、約50%以上、約60%以上、約70%以上、約75%以上、約80%以上、約81%以上、約82%以上、約83%以上、約84%以上、約85%以上、約86%以上、約87%以上、約88%以上、約89%以上、約90%以上、約91%以上、約92%以上、約93%以上、約94%以上、約95%以上、約96%以上、約97%以上、約98%以上、約99%以上、またはそれよりも大きい。特別な実施態様では、CRISPR反復配列は、配列番号2、17、25、36、44、51または63のヌクレオチド配列、またはその活性なバリエーションまたは断片を含んでおり、それがガイドRNAの中に含まれているときには、本明細書に提示されている関連するRNA誘導型ヌクレアーゼを興味ある標的配列に配列特異的に結合させることができる。いくつかの実施態様では、野生型配列の活性なCRISPR反復配列バリエーションは、配列番号2、17、25、36、44、51または63として表わされるヌクレオチド配列と配列が少なくとも40%、45%、約50%、約60%、約65%、約70%、約75%、約80%、約85%、約90%、約91%、約92%、約93%、約94%、約95%、約96%、約97%、約98%、約99%、またはそれよりも多く一致するヌクレオチド配列を含んでいる。いくつかの実施態様では、野生型配列の活性なCRISPR反復配列断片は、配列番号2、17、25、36、44、または51として表わされるヌクレオチド配列の連続した少なくとも5個、6個、7個、8個、9個、10個、11個、12個、13個、14個、15個、16個、17個、18個、19個、または20個のヌクレオチドを含んでいる。

【0048】

いくつかの実施態様では、crRNAは天然には存在しない。これら実施態様のいくつかでは、特異的なCRISPR反復配列は、操作されたスペーサ配列に自然に連結されることなく、このCRISPR反復配列はスペーサ配列にとって異種であると見なされる。いくつかの実施態様では、スペーサ配列は、天然に存在しない操作された配列である。

【0049】

トランス活性化CRISPR RNA分子、すなわちtracrRNA分子は、crRNAのCRISPR反復配列にハイブリダイズするのに十分な相補性を持つ領域（それを本明細書ではアンチ反復領域と呼ぶ）を含むヌクレオチド配列を含んでいる。いくつかの実施態様では、tracrRNA分子はさらに、二次構造（例えばステム-ループ）を持つ領域を含むか、対応するcrRNAにハイブリダイズしたときに二次構造を形成する。特別な実施態様では、tracrRNAのうちでCRISPR反復配列と完全に、または部分的に相補的な領域はこの分子の5'末端にあり、tracrRNAの3'末端は二次構造を含んでいる。II型RGNに関しては、二次構造のこの領域は一般にいくつかのヘアピン構造を含んでいる（その中には、アンチ反復配列の隣に見いだされる連結ヘアピンが含まれる）。連結ヘアピンは、ヘアピンステムの塩基の中に保存されたヌクレオチド配列を有することがしばしばあり、tracrRNA内の連結ヘアピンの中にはモチーフUNANNC（配列番号13；APG00969について）、ANGNNU（配列番号23；APG03128について）、またはUNANNA（配列番号42；APG00771について）が見いだされる。tracrRNAの3'末端には、複数の末端ヘアピンが存在することがしばしばあり、その構造と数はさまざまである可能性があるが、CGリッチでRhoとは独立な転写終了ヘアピンと、その後続く3'末端に位置するUのストリングを含んでいることがしばしばあ

る。例えばBriner他(2014年)Molecular Cell 第56巻:333~339ページ、BrinerとBarrangou(2016年)Cold Spring Harb Protoc; doi: 10.1101/pdb.top090902、およびアメリカ合衆国特許出願公開第2017/0275648号を参照されたい(そのそれぞれの全体が参照によって本明細書に組み込まれている)。

#### 【0050】

さまざまな実施態様では、tracrRNAのうちでCRISPR反復配列と完全に、または部分的に相補的なアンチ反復領域は、約8個~約30個、またはそれよりも多い個数のヌクレオチドを含んでいる。例えばtracrRNAアンチ反復領域とCRISPR反復配列の間で塩基対を形成する領域は、長さが約8個、約9個、約10個、約11個、約12個、約13個、約14個、約15個、約16個、約17個、約18個、約19個、約20個、約21個、約22個、約23個、約24個、約25個、約26個、約27個、約28個、約29個、約30個、またはそれよりも多い個数のヌクレオチドであることが可能である。特別な実施態様では、tracrRNAのうちでCRISPR反復配列と完全に、または部分的に相補的なアンチ反復領域は、長さが約20個のヌクレオチドである。いくつかの実施態様では、CRISPR反復配列とそれに対応するtracrRNAアンチ反復配列の間の相補性の程度は、適切なアラインメントプログラムを用いて最適なアラインメントにしたとき、約50%以上、約60%以上、約70%以上、約75%以上、約80%以上、約81%以上、約82%以上、約83%以上、約84%以上、約85%以上、約86%以上、約87%以上、約88%以上、約89%以上、約90%以上、約91%以上、約92%以上、約93%以上、約94%以上、約95%以上、約96%以上、約97%以上、約98%以上、約99%以上、またはそれよりも大きい。

#### 【0051】

さまざまな実施態様では、tracrRNA全体は、約60個のヌクレオチド~約140個超のヌクレオチドを含むことができる。例えばtracrRNAは、長さが約60個、約65個、約70個、約75個、約80個、約85個、約90個、約95個、約100個、約105個、約110個、約115個、約120個、約125個、約130個、約135個、約140個、またはそれよりも多い個数のヌクレオチドであることが可能である。特別な実施態様では、tracrRNAは、長さが約80個~約90個のヌクレオチドであり、その中には長さが約80個、約81個、約82個、約83個、約84個、約85個、約86個、約87個、約88個、約89個、および約90個のヌクレオチドが含まれる。いくつかの実施態様では、tracrRNAは、長さが約85個のヌクレオチドである。

#### 【0052】

特別な実施態様では、tracrRNAは、配列番号3、18、26、37、45、52、または62のヌクレオチド配列、またはその活性なバリエーションまたは断片を含んでおり、それがガイドRNAの中に含まれているときには、本明細書に提示されている関連するRNA誘導型ヌクレアーゼを興味ある標的配列に配列特異的に結合させることができる。いくつかの実施態様では、野生型配列の活性なtracrRNA配列バリエーションは、配列番号3、18、26、37、45、52、または62として表わされるヌクレオチド配列と配列が少なくとも40%、45%、約50%、約55%、約60%、約65%、約70%、約75%、約80%、約85%、約90%、約91%、約92%、約93%、約94%、約95%、約96%、約97%、約98%、約99%、またはそれよりも多く一致するヌクレオチド配列を含んでいる。いくつかの実施態様では、野生型配列の活性なtracrRNA配列断片は、配列番号3、18、26、37、45、または52として表わされるヌクレオチド配列の連続した少なくとも5個、10個、15個、20個、25個、30個、35個、40個、45個、50個、55個、60個、65個、70個、75個、80個、またはそれよりも多い個数のヌクレオチドを含んでいる。

#### 【0053】

2つのポリヌクレオチド配列は、その2つの配列が厳しい条件下で互いにハイブリダイズするとき、実質的に相補的であると見なすことができる。同様に、RGNは、そのRGNに結合したガイドRNAが厳しい条件下で特定の標的配列に結合する場合に、配列特異的なやり方でその標的配列に結合すると見なされる。「厳しい条件」または「厳しいハイブリダイゼーション条件」とは、2つのポリヌクレオチド配列が、他の配列よりも大きな程度(例

例えばバックグラウンドの少なくとも2倍)の検出可能性で互いにハイブリダイズする条件を意味する。厳しい条件は配列に依存しており、環境が異なれば異なるであろう。典型的には、厳しい条件は、塩の濃度が、pH 7.0~8.3で約1.5 M未満のNaイオン、典型的には約0.01~1.0 MのNaイオン(または他の塩)、かつ温度が短い配列(例えば10~50個のヌクレオチド)では少なくとも約30、長い配列(例えば50個超のヌクレオチド)では少なくとも約60である条件である。厳しい条件は、脱安定剤(ホルムアミドなど)の添加によって実現することもできる。厳しさの低い代表的な条件には、30~35%のホルムアミド、1 MのNaCl、1% SDS(ドデシル硫酸ナトリウム)の緩衝溶液を用いた37 でのハイブリダイゼーションと、50~55 の1x~2x SSC(20x SSC=3.0 MのNaCl/0.3 Mのクエン酸三ナトリウム)の中での洗浄が含まれる。中程度の厳しさの代表的な条件には、40~45%のホルムアミド、1.0 MのNaCl、1% SDSの中での37 でのハイブリダイゼーションと、55~60 の0.5x~1x SSCの中での洗浄が含まれる。厳しさの程度が高い代表的な条件には、50%のホルムアミド、1 MのNaCl、1% SDSの中での37 でのハイブリダイゼーションと、60~65 の0.1x SSCの中での洗浄が含まれる。場合によっては、洗浄用バッファは、約0.1%~約1%のSDSを含むことができる。ハイブリダイゼーションの時間は一般に約24時間未満であり、通常は約4~約12時間である。洗浄時間は、少なくとも平衡に到達するのに十分な長さの時間になろう。

#### 【0054】

T<sub>m</sub>は、相補的な標的配列の50%が、完全に一致した配列に(規定されたイオン強度とpHのもとで)ハイブリダイズする温度である。DNA-DNAハイブリッドに関しては、T<sub>m</sub>は、MeinkothとWahl(1984年) Anal. Biochem. 第138巻: 267~284ページの式から大まかに知ることができ、 $T_m = 81.5 + 16.6(\log M) + 0.41(\%GC) - 0.61(\%form) - 500/L$ となる(ただしMは1価カチオンのモル数であり、%GCは、DNAの中のグアノシンヌクレオチドとシトシンヌクレオチドの割合であり、%formは、ハイブリダイゼーション溶液に含まれるホルムアミドの割合であり、Lは、塩基対の中のハイブリッドの長さである)。一般に、厳しい条件は、規定されたイオン強度とpHにおける具体的な配列とその相補的配列の融点(T<sub>m</sub>)よりも約5 低くなるように選択される。しかし極めて厳しい条件は、融点(T<sub>m</sub>)よりも1、2、3、または4 低い温度でのハイブリダイゼーションおよび/または洗浄を利用することができ; 中程度に厳しい条件は、融点(T<sub>m</sub>)よりも6、7、8、9、または10 低い温度でのハイブリダイゼーションおよび/または洗浄を利用することができ; 厳しさが低い条件は、融点(T<sub>m</sub>)よりも11、12、13、14、15、または20 低い温度でのハイブリダイゼーションおよび/または洗浄を利用することができる。当業者は、ハイブリダイゼーションおよび/または洗浄溶液の厳しさの変動が、上記の式、ハイブリダイゼーション用と洗浄用の組成物、および望むT<sub>m</sub>を利用して本質的に記述されることを理解しているであろう。核酸のハイブリダイゼーションに関する包括的なガイドは、Tijssen(1993年)『Laboratory Techniques in Biochemistry and Molecular Biology - Hybridization with Nucleic Acid Probes』、第I部、第2章(Elsevier社、ニューヨーク); およびAusubel他編(1995年)『Current Protocols in Molecular Biology』、第2章(Greene Publishing and Wiley-Interscience社、ニューヨーク)に見いだされる。Sambrook他(1989年)『Molecular Cloning: A Laboratory Manual』(第2版)、Cold Spring Harbor Laboratory Press、ブレインビュー、ニューヨーク州を参照されたい。

#### 【0055】

ガイドRNAとして、単一ガイドRNA、または二重ガイドRNAシステムが可能である。単一ガイドRNAは、単一のRNA分子上にcrRNAとtracrRNAを含んでいるのに対し、二重ガイドRNAシステムは、2つの異なるRNA分子上に、crRNAとtracrRNAを、crRNAのCRISPR反復配列の少なくとも一部とtracrRNAの少なくとも一部(crRNAのCRISPR反復配列と完全に、または部分的に相補的なものが可能である)を通じて互いにハイブリダイズした状態で含んでいる。ガイドRNAが単一ガイドRNAである実施態様のいくつかでは、crRNAとtracrRNAはリンカーヌクレオチド配列によって隔てられている。一般に、リンカー

10

20

30

40

50



ヌクレオチド配列は、リンカーヌクレオチド配列のヌクレオチドの中に二次構造が形成されることを避けるため、またはリンカーヌクレオチド配列のヌクレオチドを含む二次構造が形成されることを避けるため、相補的な塩基を含んでいない配列である。いくつかの実施態様では、crRNAとtracrRNAの間のリンカーヌクレオチド配列は、長さが少なくとも3個、少なくとも4個、少なくとも5個、少なくとも6個、少なくとも7個、少なくとも8個、少なくとも9個、少なくとも10個、少なくとも11個、少なくとも12個、またはそれよりも多い個数のヌクレオチドである。特別な実施態様では、単一ガイドRNAのリンカーヌクレオチド配列は、長さが少なくとも4個のヌクレオチドである。いくつかの実施態様では、リンカーヌクレオチド配列は、配列番号8または31として表わされるヌクレオチド配列である。別の実施態様では、リンカーヌクレオチド配列は、長さが少なくとも6個のヌクレオチドである。

10

#### 【0056】

単一ガイドRNA、または二重ガイドRNAは、化学的に合成すること、またはインビトロでの転写を通じて合成することができる。RGNとガイドRNAの間の配列特異的結合を明らかにするためのアッセイは本分野で知られており、その非限定的な例に含まれるのは、発現しているRGNとガイドRNAの間のインビトロ結合アッセイである。ガイドRNAは、検出可能な標識（例えばビオチン）をタグとして付けてプル-ダウン検出アッセイで 사용할 ことができ、このアッセイでは、ガイドRNA：RGN複合体が、（例えばストレプトアビジンビーズを用いて）検出可能な標識を通じて捕獲される。ガイドRNAとは無関係な配列または構造を持つ対照ガイドRNAを、RNAへのRGNの非特異的結合の陰性対照として用いることができる。いくつかの実施態様では、ガイドRNAは配列番号4、19、27、38、46、53、64、65、または66であり、その中のスパーサ配列は任意の配列が可能であり、ポリN配列として示される。

20

#### 【0057】

いくつかの実施態様では、ガイドRNAは、標的細胞、オルガネラ（細胞小器官）、または胚の中にRNA分子として導入することができる。ガイドRNAは、インビトロで転写すること、または化学的に合成することができる。別の実施態様では、ガイドRNAをコードするヌクレオチド配列は、標的細胞、オルガネラ、または胚の中に導入される。これら実施態様のいくつかでは、ガイドRNAをコードするヌクレオチド配列は、プロモータ（例えばRNAポリメラーゼIIIプロモータ）に機能可能に連結されている。プロモータとして、天然のプロモータ、またはガイドRNAをコードするヌクレオチド配列にとって異種であるものが可能である。

30

#### 【0058】

さまざまな実施態様では、ガイドRNAは、本明細書に記載されているように、標的細胞、オルガネラ、または胚の中にリボ核タンパク質複合体として導入ことができ、ガイドRNAは、RNA誘導型ヌクレアーゼポリペプチドに結合する。

#### 【0059】

ガイドRNAは、興味ある特定の標的ヌクレオチド配列にハイブリダイズすることを通じ、関連するRNA誘導型ヌクレアーゼをその標的ヌクレオチド配列に向かわせる。標的ヌクレオチド配列は、DNA、RNA、またはその両方の組み合わせを含むことができ、一本鎖または二本鎖が可能である。標的ヌクレオチド配列として、ゲノムDNA（すなわち染色体DNA）、プラスミドDNA、またはRNA分子（例えばメッセンジャーRNA、リボソームRNA、トランスファーRNA、マイクロRNA、小さな干渉性RNA）が可能である。標的ヌクレオチド配列は、インビトロで、または細胞内でRNA誘導型ヌクレアーゼに結合すること（いくつかの場合にはRNA誘導型ヌクレアーゼによって切断すること）ができる。RGNが標的とする染色体配列として、核、プラスチド、またはミトコンドリア染色体配列が可能である。いくつかの実施態様では、標的ヌクレオチド配列は、標的ゲノムの中に1つだけである。

40

#### 【0060】

標的ヌクレオチド配列は、プロトスパーサモチーフ（PAM）に隣接している。プロトス

50

ペーサモチーフは一般に標的ヌクレオチド配列からヌクレオチド約1個～約10個以内の位置にあり、その中には、標的ヌクレオチド配列からヌクレオチド約1個、約2個、約3個、約4個、約5個、約6個、約7個、約8個、約9個、または約10個が含まれる。PAMは標的配列の5'または3'が可能である。いくつかの実施態様では、PAMは、本開示のRGNのための標的配列の3'である。一般に、PAMは約3～4個のヌクレオチドからなるコンセンサス配列だが、特別な実施態様では、長さが2個、3個、4個、5個、6個、7個、8個、9個、またはそれよりも多い個数のヌクレオチドであることが可能である。さまざまな実施態様では、本開示のRGNによって認識されるPAM配列は、配列番号7、22、30、41、または49として表わされるコンセンサス配列を含んでいる。

#### 【0061】

特別な実施態様では、配列番号1、16、24、35、43、または50を持つRNA誘導型ヌクレアーゼ、またはその活性なバリエーションまたは断片は、それぞれ、配列番号7、22、30、41、または49として表わされるPAM配列に隣接した標的ヌクレオチド配列に結合する。これら実施態様のいくつかでは、RGNは、それぞれ配列番号2、17、25、36、44、51、または63として表わされるCRISPR反復配列、またはその活性なバリエーションまたは断片と、それぞれ配列番号3、18、26、37、45、52、または62として表わされるtracrRNA配列、またはその活性なバリエーションまたは断片を含むガイド配列に結合する。RGNシステムは、本明細書の実施例1と表1にさらに詳しく記載されている。

#### 【0062】

所与のヌクレアーゼ酵素に関するPAM配列特異性が酵素の濃度の影響を受けることは周知であり（例えばKarvelis他（2015年）Genome Biol第16巻：253ページを参照されたい）、それは、RGNの発現に用いるプロモータを変えること、または細胞、オルガネラ、または胚に送達されるリボ核タンパク質複合体の量を変えることによって変更することができる。

#### 【0063】

RGNは、対応するPAM配列を認識すると、特定の切断部位にある標的ヌクレオチド配列を切断することができる。本明細書では、切断部位は、標的ヌクレオチド配列内の2個の特定のヌクレオチドで構成されており、その間でヌクレオチド配列がRGNによって切断される。切断部位は、PAMから5'方向または3'方向に向かって1番目と2番目、2番目と3番目、3番目と4番目、4番目と5番目、5番目と6番目、7番目と8番目、または8番目と9番目のヌクレオチドを含むことができる。いくつかの実施態様では、切断部位は、PAMから5'方向または3'方向に向かって10個、11個、12個、13個、14個、15個、16個、17個、18個、19個、または20個のヌクレオチドを超えた位置に存在することが可能である。いくつかの実施態様では、切断部位は、PAMからヌクレオチド4個分離されている。別の実施態様では、切断部位はPAMから少なくとも15個のヌクレオチドだけ離れている。RGNは標的ヌクレオチド配列を切断して粘着末端にすることができるため、いくつかの実施態様では、切断部位は、ポリヌクレオチドのプラス（+）鎖上のPAMからヌクレオチド2個の距離と、ポリヌクレオチドのマイナス（-）鎖上のPAMからヌクレオチド2個の距離に基づいて規定される。

#### 【0064】

### III. 融合タンパク質

#### 【0065】

「リンカー」という用語は、本明細書では、2個の分子または部分（例えばヌクレアーゼの結合ドメインと切断ドメイン）を連結する化学基または分子を意味する。いくつかの実施態様では、リンカーは、RNA誘導型ヌクレアーゼとデアミナーゼを連結する。いくつかの実施態様では、リンカーは、dCas9とデアミナーゼを連結する。典型的には、リンカーは、2つの基、分子、またはそれ以外の部分の間に、またはこれらに隣接して位置しており、共有結合を通じてそれぞれに接続されているため、その2つを接続している。いくつかの実施態様では、リンカーは、1個のアミノ酸、または複数個のアミノ酸（例えばペプチドまたはタンパク質）である。いくつかの実施態様では、リンカーは、有機分子、基

10

20

30

40

50

、ポリマー、または化学的部分である。いくつかの実施態様では、リンカーは、長さが5～100個のアミノ酸、例えば長さが5個、6個、7個、8個、9個、10個、11個、12個、13個、14個、15個、16個、17個、18個、19個、20個、21個、22個、23個、24個、25個、26個、27個、28個、29個、30個、30～35個、35～40個、40～45個、45～50個、50～60個、60～70個、70～80個、80～90個、90～100個、100～150個、または150～200個のアミノ酸である。より長い、またはより短いリンカーも考えられる。

#### 【0066】

本開示のいくつかの側面では、DNA結合ポリペプチドとデアミナーゼポリペプチドを含む融合タンパク質が提供される。DNA結合ポリペプチドとして、DNAに結合する任意のタンパク質またはタンパク質ドメインが可能である。いくつかの実施態様では、融合タンパク質のDNA結合ポリペプチドは、メガヌクレアーゼ、ジンクフィンガー融合タンパク質、またはTALENである。本開示のいくつかの側面により、RNA誘導型DNA結合ポリペプチドとデアミナーゼポリペプチドを含む融合タンパク質が提供される。いくつかの実施態様では、RNA誘導型DNA結合ポリペプチドはRNA誘導型ヌクレアーゼである。いくつかの実施態様では、RNA誘導型ヌクレアーゼは本発明のRGNである。いくつかの実施態様では、RGNは本発明のRGNではない。さらに別の実施態様では、RNA誘導型ヌクレアーゼはCRISPR-Casタンパク質である。さらに別の実施態様では、CRISPR-Casタンパク質はII型CRISPR-Casタンパク質である。別の実施態様では、CRISPR-Casタンパク質はV型CRISPR-Casタンパク質である。別の実施態様では、CRISPR-Casタンパク質はVI型CRISPR-Casタンパク質である。いくつかの実施態様では、RNA誘導型ヌクレアーゼは、ガイドRNAに結合するCas9ドメインであり、このCas9ドメインが今度は鎖ハイブリダイゼーションを通じて標的核酸配列に結合する。いくつかの実施態様では、デアミナーゼポリペプチドとして、核酸塩基（例えばシチジンまたはアデニンなど）を脱アミノ化することのできるデアミナーゼドメインが可能である。いくつかの実施態様では、デアミナーゼポリペプチドは、配列番号374～545または572～584のいずれかから選択されたアミノ酸配列またはそのバリエーションを含んでいる。これら実施態様のいくつかでは、デアミナーゼポリペプチドは、配列番号374、383、397、399、407、408、411、414、416、420、514、および572～584のいずれか1つから選択されたアミノ酸配列またはそのバリエーションを含んでいる。デアミナーゼによって核酸塩基を脱アミノ化すると、各残基の位置に点変異を起こすことができ、そのことによってDNA分子を改変する。この改変操作を本明細書では核酸編集または塩基編集とも呼ぶ。したがってCas9バリエーションまたはCas9ドメインとデアミナーゼを含む融合タンパク質を用いて核酸配列を狙い通りに編集することができる。

#### 【0067】

このような融合タンパク質は、インビトロでDNAを狙い通りに編集して例えば変異細胞を生成させるのに有用である。これら変異細胞は植物または動物の中に存在することができる。このような融合タンパク質は、狙った変異を導入し、例えば生体外で哺乳動物の細胞（例えば対象から得られてその後同じ対象または別の対象に再導入される細胞）の中にある遺伝子の欠陥を修正することと；狙った変異を導入すること（例えば対象となる哺乳動物の細胞内の疾患関連遺伝子の中にある遺伝子の欠陥を修正すること、またはその遺伝子の中に脱活性化変異を導入すること）にも役立つ可能性がある。このような融合タンパク質は、植物細胞の中に狙った変異を導入し、例えば有益であるか農業で価値のある形質またはアレルを導入するのに有用である可能性がある。

#### 【0068】

本明細書で用いられている「ウラシルグリコシラーゼ阻害剤」または「UGI」という用語は、ウラシル-DNAグリコシラーゼ塩基-除去修復酵素を抑制することのできるタンパク質を意味する。いくつかの実施態様では、融合タンパク質は、デアミナーゼに融合されたヌクレアーゼ不活性なRGN（Cas9（dCas9）など）を含んでいる。いくつかの実施態様では、融合タンパク質は、デアミナーゼに融合されたニッカーゼRGN（nCAs9など）を含んでいる。いくつかの実施態様では、融合タンパク質は、デアミナーゼに融合され、さらにUGIドメインに融合されたヌクレアーゼ不活性なRGNまたはニッカーゼRGNを含んでい

る。

【0069】

いくつかの実施態様では、融合タンパク質のニッカーゼRGNは、D10A変異、またはそれとホモロジーが同等な変異（配列番号569；または同様に配列番号553）（この変異はRGNが核酸二重鎖の標的鎖（PAMを含む鎖）だけを切断できるようにする）を含んでいる。いくつかの実施態様では、融合タンパク質のヌクレアーゼ不活性な（「デッド」）RGNは、D10A変異と、H840A変異（この変異はRGNがDNA標的を切断できなくする）またはそれとホモロジーが同等な変異（配列番号568；または同様に配列番号547）を含んでいる。いくつかの実施態様では、融合タンパク質のニッカーゼRGNは、RGNが核酸二重鎖の非標的鎖（PAMを含まない鎖）だけを切断できるようにするH840A変異を含んでいる。H840A変異またはそれと同等な変異を含むニッカーゼRGNは、不活性化されたHNHDメインを有する。D10A変異またはそれと同等な変異を含むニッカーゼRGNは、不活性化されたRuvCドメインを有する。デアミナーゼは非標的鎖に作用する。D10A変異またはそれと同等な変異を含むニッカーゼは不活性なRuvCヌクレアーゼドメインを持ち、DNAの非標的鎖、すなわち塩基編集を望む鎖を切断することはできない。

10

【0070】

いくつかの実施態様では、本明細書に記載されている融合タンパク質のRGNはニッカーゼ活性を持つ。ニッカーゼとして、RGNの断片、またはRGNのニッカーゼバリエーションが可能である。いくつかの実施態様では、本明細書に記載されている融合タンパク質のRGNドメインは少なくとも部分的に脱活性化されたヌクレアーゼ活性を持つため、RNA誘導型DNA結合ポリペプチドと呼ぶことができる。本明細書に記載されている融合タンパク質を利用する方法も提供される。いくつかの実施態様では、RGNはCas9タンパク質である。ヌクレアーゼ不活性なCas9ドメインとニッカーゼCas9ドメインの非限定的な代表例が本明細書に提示されている。適切な1つの代表的なヌクレアーゼ不活性なRGNドメインはD10A/H840A Cas9ドメイン変異体である（例えばQi他、Cell、2013年；第152巻(5)：1173～1183ページを参照されたい；その内容全体が参照によって本明細書に組み込まれている）。追加の適切なヌクレアーゼ不活性なCas9ドメインは、本開示に基づく当業者には明らかであろう。そのような適切な追加の代表的なヌクレアーゼ不活性なCas9ドメインの非限定的な例に含まれるのは、D10A変異ドメイン、D10A/D839A/H840A変異ドメイン、およびD10A/D839A/H840A/N863A変異ドメインである（例えばMali他、Nature Biotechnology、2013年；第31巻(9)：833～838ページを参照されたい；その内容全体が参照によって本明細書に組み込まれている）。それに加え、知られている他のRGNの適切なヌクレアーゼ不活性なRGNドメインを確定することができる（例えば配列番号547、RGN APG08290.1のヌクレアーゼ不活性なバリエーション；アメリカ合衆国特許出願第16/432,321号を参照されたい；その内容全体が参照によって本明細書に組み込まれている）

20

30

【0071】

本開示のいくつかの側面では、(i)ヌクレアーゼ不活性なRGN、またはヌクレアーゼ不活性なドメイン、またはニッカーゼRGN、またはニッカーゼドメインと、(ii)デアミナーゼ酵素またはデアミナーゼドメインを含む融合タンパク質が提供される。いくつかの実施態様では、デアミナーゼ酵素またはデアミナーゼドメインは、DNA編集酵素またはDNA編集ドメインである。いくつかの実施態様では、デアミナーゼ酵素はデアミナーゼ活性を有する。いくつかの実施態様では、デアミナーゼ酵素またはデアミナーゼドメインは、デアミナーゼドメインを含むか、デアミナーゼドメインである。いくつかの実施態様では、デアミナーゼはシチジンデアミナーゼである。いくつかの実施態様では、デアミナーゼはアポリタンパク質B mRNA編集複合体（APOBEC）ファミリーのデアミナーゼである。いくつかの実施態様では、デアミナーゼはAPOBEC1ファミリーのデアミナーゼである。別の実施態様では、デアミナーゼはAPOBEC3ファミリーのデアミナーゼである。いくつかの実施態様では、デアミナーゼは活性化誘導シチジンデアミナーゼ（AID）である。いくつかの実施態様では、デアミナーゼはACF1/ASEデアミナーゼである。いくつかの実施態様では、デアミナーゼはアデノシンデアミナーゼである。いくつかの実施態様では、デ

40

50

アミナーゼはADATファミリーのデアミナーゼである。いくつかの核酸デアミナーゼ酵素と核酸デアミナーゼドメインが本明細書に詳細に記載されている（表17参照）。追加の適切なデアミナーゼ酵素またはデアミナーゼドメインは、本開示に基づく当業者には明らかであろう。これら実施態様のいくつかでは、デアミナーゼポリペプチドは、配列番号374、383、397、399、407、408、411、414、416、420、514、および572～584のいずれか1つから選択されたアミノ酸配列、またはそのバリエーションを含んでいる。

#### 【0072】

本開示により、さまざまな配置の融合タンパク質が提供される。いくつかの実施態様では、デアミナーゼ酵素またはデアミナーゼドメインはRGNドメインのN末端に融合される。いくつかの実施態様では、デアミナーゼ酵素またはデアミナーゼドメインはRGNドメインのC末端に融合される。いくつかの実施態様では、リンカーは、(GGGS)<sub>n</sub>モチーフ（配列番号585）、(G)<sub>n</sub>モチーフ（配列番号586）、(EAAAK)<sub>n</sub>モチーフ（配列番号587）、または(XP)<sub>n</sub>モチーフ（配列番号588）、またはこれらの任意の組み合わせを含んでいる（nは独立に、1～30の間の整数である）。いくつかの実施態様では、nは独立に、1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、または30であり、2つ以上のリンカー、または2つ以上のリンカーモチーフが存在している場合には、これらの任意の組み合わせである。追加の適切なリンカーモチーフとリンカーの配置は、当業者には明らかであろう。いくつかの実施態様では、適切なリンカーモチーフとリンカーの配置に含まれるのは、Chenらの「融合タンパク質リンカー：特性、設計、機能」（Adv Drug Deliv Rev. 2013年；第65巻(10)：1357～1369ページ、その内容全体が参照によって本明細書に組み込まれている）の中に記載されているものである。追加の適切なリンカー配列は、本開示に基づく当業者には明らかであろう。

#### 【0073】

いくつかの実施態様では、本明細書に開示されている代表的な融合タンパク質の一般的なアーキテクチャは、構造：[NH<sub>2</sub>]-[デアミナーゼ酵素またはデアミナーゼドメイン]-[RGNタンパク質またはRGNドメイン][COOH]、または [NH<sub>2</sub>]-[RGNタンパク質またはRGNドメイン]-[デアミナーゼ酵素またはデアミナーゼドメイン][COOH]を含んでいる（これらの構造の中のNH<sub>2</sub>は融合タンパク質のN末端であり、COOHは融合タンパク質のC末端である）。追加の特徴が存在していてもよく、それは例えば、NLSと融合タンパク質の残部の間、および/またはデアミナーゼ酵素またはデアミナーゼドメインとRGNタンパク質またはRGNドメインの間の1つ以上のリンカー配列である。存在していてもよい他の代表的な特徴は、局在化配列（核局在化配列、細胞質局在化配列など）、輸送配列（核外輸送配列など）、または他の局在化配列のほか、融合タンパク質の可溶化、精製、または検出に有用な配列タグである。本明細書に提示されている適切な局在化シグナル配列とタンパク質タグ配列の非限定的な例に含まれるのは、ピオチンカルボキシラーゼ担体タンパク質（BCCP）タグ、myc-タグ、カルモジュリン-タグ、FLAGタグ、ヘマグルチニン（HA）タグ、ポリヒスチジンタグ（ヒスチジンタグまたはHisタグとも呼ばれる）、マルトース結合タンパク質（MBP）タグ、nus-タグ、グルタチオン-S-トランスフェラーゼ（GST）タグ、緑色蛍光タンパク質（GFP）タグ、チオレドキシニン-タグ、Sタグ、Softag（例えばSoftag 1、Softag 3）、Strepタグ、ピオチンリガーゼタグ、FIAsHタグ、V5タグ、およびSBPタグである。追加の適切な配列は当業者には明らかであろう。

#### 【0074】

いくつかの実施態様では、デアミナーゼ酵素、またはデアミナーゼ酵素またはデアミナーゼドメインを有する代表的な融合タンパク質の一般的なアーキテクチャは、構造：[NH<sub>2</sub>]-[NLS]-[RGNタンパク質またはRGNドメイン]-[デアミナーゼ]-[COOH]、[NH<sub>2</sub>]-[NLS]-[デアミナーゼ]-[RGNタンパク質またはRGNドメイン]-[COOH]、[NH<sub>2</sub>]-[RGNタンパク質またはRGNドメイン]-[デアミナーゼ]-[COOH]、または[NH<sub>2</sub>]-[デアミナーゼ]-[RGNタンパク質またはRGNドメイン]-[COOH]を含んでいる（これらの構造の中のNLSは核局在化シグナルであり、NH<sub>2</sub>は融合タンパク質のN末端であり、COOHは融合タンパク質のC末

端である)。いくつかの実施態様では、リンカーが、RGNタンパク質またはRGNドメインとデアミナーゼの間に挿入される。いくつかの実施態様では、NLSは、デアミナーゼ、および/またはRGNタンパク質またはRGNドメインのC末端に位置する。いくつかの実施態様では、NLSは、デアミナーゼとRGNタンパク質またはRGNドメインの間に位置する。追加の特徴(配列タグなど)も存在することができる。「RGNタンパク質またはRGNドメイン」は、ここでは任意のRNA誘導型ヌクレアーゼを表わし、その中にはCRISPR-Casタンパク質とそのバリエーションおよび変異体が含まれ、それを用いて本発明の融合タンパク質を作ることができる。RGNタンパク質として、ヌクレアーゼ不活性なRGNまたはCRISPR-Cas(例えばdCas9(配列番号568)または配列番号547など)、またはRGNニッカーゼまたはCas9ニッカーゼ(例えば配列番号569(または配列番号553)など)が可能である。いくつかの実施態様では、本発明の融合タンパク質はRNA誘導型DNA結合ポリペプチドとデアミナーゼを含んでおり、このデアミナーゼは、配列番号374~545または572~584のいずれかと少なくとも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%、または100%一致するアミノ酸配列、またはその活性なバリエーションを有する。これら実施態様のいくつかでは、融合タンパク質は、配列番号374、383、397、399、407、408、411、414、416、420、514、および572~584のいずれか1つから選択されたアミノ酸配列を含むデアミナーゼポリペプチド、またはそのバリエーションを含んでいる。このような融合タンパク質の例は本明細書の実施例の項に記載されている。

#### 【0075】

デアミナーゼ酵素とデアミナーゼドメインの適切な1つの代表的なタイプは、例えばAPOBECファミリーのシトシンデアミナーゼである。シトシンデアミナーゼ酵素のアポリボタンパク質B mRNA編集複合体(APOBEC)ファミリーは、制御されていて有益なやり方で突然変異誘発を開始させるのに役立つ11種類のタンパク質を包含する(Conticello他、2008年、Genome Biology、第9巻(6):229ページ)。1つのファミリーメンバーである活性化誘導シチジンデアミナーゼ(AID)は、転写に依存していてストランドバイアスがあるやり方でssDNAの中のシトシンをウラシルに変換することによって抗体を成熟させる(Reynaud他、2003年、Nature Immunology、第4巻(7):631~638ページ)。アポリボタンパク質B編集複合体3(APOBEC3)酵素は、逆転写されたウイルスssDNAの中のシトシンを脱アミノ化することを通じてヒト細胞にある種のHIV-1株から保護する(Bhagwat他、2004年、第3巻(1):85~89ページ)。これらタンパク質はすべて、触媒活性のため、 $Zn^{2+}$ 配位モチーフ(HisX-Glu-X<sub>23-26</sub>-Pro-Cys-X<sub>2-4</sub>-Cys;配列番号589)と、結合した水分子を必要とする。Glu残基は、脱アミノ化反応における求核攻撃のため、水分子を活性化して水酸化亜鉛にする。ファミリーの各メンバーは、それ自身の特定の「ホットスポット」(hAIDのWRC(WはAまたはTであり、RはAまたはGである)からhAPOBEC3FのTTCの範囲にわたる)において選択的に脱アミノ化する(Navaratnam他、2006年、Intl J Hematol第83巻(3):195~200ページ)。APOBEC3Gの触媒ドメインの最近の結晶構造から、6つのヘリックスが隣接した5つの鎖状シートコアからなる二次構造が明らかにされ、これはファミリー全体で保存されていると考えられている(Holden他、2008年、Nature第456巻(7218):121~124ページ)。活性な中心ループが、ssDNAの結合と、「ホットスポット」の属性明確化の両方に責任があることが示された(Celico他、2009年、J Biol Chem第284巻(41):27761~27765ページ)。これら酵素の過剰発現はゲノム不安定性およびがんと結びついているため、配列特異的ターゲティングの重要性が注目されている(Pham他、2005年、Biochem第44巻(8):2703~2715ページ)。

#### 【0076】

デアミナーゼ酵素とデアミナーゼドメインの適切な別の1つの代表的なタイプは、アデノシンデアミナーゼである。ADATファミリーのアデノシンデアミナーゼは、RGNまたは断片、またはRGNのドメイン、またはそのバリエーション(例えばヌクレアーゼ不活性なCas9

ドメイン)に融合できるため、Cas9-ADAT融合タンパク質が生成する。本開示には、RGNまたは断片、またはRGNのドメイン、またはそのバリエーションと、デアミナーゼ酵素(例えばシトシンデアミナーゼ(APOBEC酵素など)またはアデノシンデアミナーゼ酵素(ADAT酵素など))の間の融合体の系統的なシリーズが含まれるため、RGN-デアミナーゼ融合体は、デアミナーゼの酵素活性をゲノムDNA内の特定の部位に向ける。認識媒体としてRGNを用いることの利点は2つある。すなわち(1)sgRNA配列を変化させるだけで融合タンパク質の配列特異性を容易に変えられることと;(2)RGN(Cas9など)がdsDNAを変性させることによって対応する標的配列に結合し、その結果として一本鎖のDNAストレッチになり、したがってデアミナーゼが利用できる基質になることである。ヒトとマウスのデアミナーゼドメイン(例えばAIDドメイン)を持つ成功した融合タンパク質が生成している(WO 2010/132092; 参照によって本明細書に組み込まれている)。本明細書に記載されているデアミナーゼとRGNの間の多彩な他の融合タンパク質も考慮される。

10

#### 【0077】

RGN-DNA複合体(RGN-DNAバブルのサイズ)の中で一本鎖であるDNAの部分は明確になっていない。しかし複合体が転写を妨げるように特別に設計されたsgRNAを持つdCas9システムでは、転写妨害は、sgRNAが非鋳型鎖に結合するときだけに起こることが示されている。この結果は、DNA-Cas9複合体内のDNAのある部分がCas9によって守られておらず、融合タンパク質の中のデアミナーゼの標的にされる可能性があることを示唆している(Qi他、2013年、Cell; 第152巻(15): 1173~1183ページ)。したがってCas9(または一般にRGN)とデアミナーゼドメインのN末端融合体とC末端融合体の両方が、本開示の側面によれば有用である。

20

#### 【0078】

いくつかの実施態様では、RGNのデアミナーゼドメインとRNA誘導型DNA結合ドメインはリンカーを介して互いに融合される。特定の用途のためのデアミナーゼ活性にとって最適な長さを実現するため、デアミナーゼドメイン(例えばAID)とRGNドメインの間のリンカーはさまざまな長さで可撓性にする事ができる(例えば(GGGGS)<sub>n</sub>(配列番号590)および(G)<sub>n</sub>の形態の非常に可撓性のあるリンカーから、(EAAAK)<sub>n</sub>(配列番号591)や(XP)<sub>n</sub>の形態のより堅固なリンカーまで)。

#### 【0079】

本開示の複数の側面によるRNA誘導型DNA結合ドメインに融合させることのできる適切ないくつかの代表的な核酸編集酵素とドメイン(例えばデアミナーゼとデアミナーゼドメイン)が提供される(配列番号374~545と572~584)。いくつかの実施態様では、それぞれの配列の活性なドメイン(例えば局在シグナル(核局在化シグナル、核外輸送シグナルなし、細胞質局在化シグナル)のないドメイン)を使用することが理解されよう。

30

#### 【0080】

いくつかの実施態様では、本明細書に提示されている融合タンパク質は、デアミナーゼ酵素の完全長アミノ酸(配列番号374~545または572~584のいずれか1つ)を含んでいる。しかし別の実施態様では、本明細書に提示されている融合タンパク質は核酸編集酵素の完全長配列を含んでおらず、その断片だけを含んでいる。例えばいくつかの実施態様では、本明細書に提示されている融合タンパク質は、RNA誘導型DNA結合ドメインと、デアミナーゼ酵素の断片を含んでいる(例えばこの断片はデアミナーゼドメインを含んでいる)。デアミナーゼドメインの代表的なアミノ酸配列は表17に記載されており、このようなドメインの追加の適切な配列は当業者には明らかであろう。

40

#### 【0081】

本発明の複数の側面に従って利用できる(例えばヌクレアーゼ不活性なRGNドメインまたはニッカーゼRGNドメインに融合させることのできる)追加の適切な核酸編集酵素配列(例えばデアミナーゼ酵素とデアミナーゼドメインの配列)は、本開示に基づく当業者には明らかであろう。いくつかの実施態様では、このような追加の酵素配列に含まれるのは、本明細書に提示されている配列と少なくとも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95%、少なくとも96%、少なくとも

50

97%、少なくとも98%、または少なくとも99%似たデアミナーゼ酵素またはデアミナーゼドメインの配列である。追加の適切なRGNドメイン、バリエーション、および配列も、当業者には明らかであろう。そのような追加の適切なRGNドメインの非限定的な例に含まれるのは、D10A変異ドメイン、D10A/D839A/H840A変異ドメイン、およびD10A/D839A/H840A/N863A変異ドメインである（例えばMali他、「標的特異的スクリーニングのためのCas9転写活性化因子と協同的ゲノム操作のためのペアードニックアーゼ」、Nature Biotechnology、2013年；第31巻(9)：833～838ページを参照されたい；その内容全体が参照によって本明細書に組み込まれている）。

#### 【0082】

RNA誘導型DNA結合ドメインとデアミナーゼドメインを含む融合タンパク質を生成させるための追加の適切な戦略は、本開示と本分野の一般的な知識の組み合わせに基づく当業者には明らかであろう。リンカーを用いて、またはリンカーを用いずに本開示の複数の側面に従う融合タンパク質を作製するのに適した戦略も、本開示と本分野の知識に照らすと当業者には明らかであろう。

#### 【0083】

いくつかの実施態様では、RNA誘導型DNA結合ドメインは、ニックアーゼ活性を有するRGNタンパク質バリエーションである。いくつかの実施態様では、RNA誘導型DNA結合ドメインはRGNニックアーゼである。いくつかの実施態様では、RGNは本発明のRGNである。別の実施態様では、RGNは本発明のRGNではない。RGNニックアーゼとして、二重鎖核酸分子（例えば二重鎖DNA分子であり、二本鎖DNA分子とも呼ばれる）の1本の鎖だけを切断することのできるCas9タンパク質が可能である。いくつかの実施態様では、RGNニックアーゼは二重鎖核酸分子の標的鎖を切断する。これは、RGNニックアーゼが、RGNに結合したgRNAと塩基対になった（相補的な）鎖を切断することを意味する。いくつかの実施態様では、RGNニックアーゼは、D10A変異、またはそれと同等な変異を含んでいる。別の実施態様では、RGNニックアーゼは、H840A変異、またはそれと同等な変異を含んでいる。例えばRGNニックアーゼは、配列番号569として示されているアミノ酸配列を含むことができる。いくつかの実施態様では、RGNニックアーゼはD10A Cas9ニックアーゼである。これは、Cas9のRuvCドメインを不活性化し、その結果として二重鎖核酸分子の標的非塩基編集鎖を切断する。これは、D10A Cas9ニックアーゼが、Cas9に結合したgRNA（例えばsgRNA）と塩基対になった鎖を切断することを意味する。いくつかの実施態様では、Cas9ニックアーゼはH840A変異を含んでいて、この変異がCas9ポリペプチドのHNHドメインを不活性化する。H840A Cas9ニックアーゼは、非標的塩基編集鎖を切断することになる。いくつかの実施態様では、RGNニックアーゼは、配列番号568、569、547、または553と少なくとも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%、または少なくとも99.5%一致するアミノ酸配列を含んでいる。変異してニックアーゼになる追加の適切なRGNタンパク質は、本開示と本分野の知識（例えばアメリカ合衆国特許出願第16/432,321号に開示されているRGN）に基づく当業者には明らかであろうゆえ、本開示の範囲に含まれる。

#### 【0084】

本開示のいくつかの側面は、ウラシルグリコシラーゼ阻害剤（UGI）ドメインを含む融合タンパク質に関する。いくつかの実施態様では、RNA誘導型DNA結合ドメイン（例えばヌクレアーゼ活性なRGNドメイン、またはヌクレアーゼ不活性であるかニックアーゼとして機能するRGNバリエーション）を含む本明細書に提示の融合タンパク質のいずれかをさらに、少なくとも1つのUGIドメインに直接、またはリンカーを介して融合させることができる。いくつかの実施態様では、融合タンパク質はさらに、少なくとも2つのUGIドメインに直接、またはリンカーを介して融合させることができる。本開示のいくつかの側面により、デアミナーゼ-RGN融合タンパク質、デアミナーゼ-ヌクレアーゼ不活性なRGN融合タンパク質、およびデアミナーゼ-ニックアーゼRGN融合タンパク質で、少なくとも1つのUGIドメインにさらに融合されていて、UGIドメインを含まない同様の融合タンパク質と比べてC T

10

20

30

40

50



核酸塩基編集効率が増大したものが提供される。いかなる特定の理論に囚われることも望まないが、U：Gヘテロ二重鎖DNAの存在に対する細胞DNA修復反応が、細胞における核酸塩基編集効率に責任がある可能性がある。例えばウラシルDNAグリコシラーゼ（UDG）は細胞内でDNAからUを除去する触媒となり、それが塩基除去修復を開始させて、最も一般的な結果としてU：G対をC：G対に戻すことができる。

【0085】

本開示では、本発明のデアミナーゼとRNA誘導型DNA結合ドメインを含んでいてUGIドメインにさらに融合された融合タンパク質を考慮する。本開示では、デアミナーゼ、RGNニッカーゼ、またはヌクレアーゼ不活性なRGNポリペプチドを含んでいてUGIドメインにさらに融合された融合タンパク質も考慮する。UGIドメインを使用すると、CからUへの変化の触媒となることができる核酸編集ドメインの編集効率を大きくできることを理解すべきである。例えばUGIドメインを含む融合タンパク質は、C残基を脱アミノ化する効率がより大きい可能性がある。いくつかの実施態様では、融合タンパク質は、構造：[デアミナーゼ]-[オプションのリンカー配列]-[ヌクレアーゼ不活性なRGN]-[オプションのリンカー配列]-[UGI]；[デアミナーゼ]-[オプションのリンカー配列]-[UGI]-[オプションのリンカー配列]-[ヌクレアーゼ不活性なRGN]；[UGI]-[オプションのリンカー配列]-[デアミナーゼ]-[オプションのリンカー配列]-[ヌクレアーゼ不活性なRGN]；[UGI]-[オプションのリンカー配列]-[オプションのリンカー配列]-[ヌクレアーゼ不活性なRGN]-[オプションのリンカー配列]-[デアミナーゼ]；[ヌクレアーゼ不活性なRGN]-[オプションのリンカー配列]-[デアミナーゼ]-[オプションのリンカー配列]-[UGI]；または [ヌクレアーゼ不活性なRGN]-[オプションのリンカー配列]-[UGI]-[オプションのリンカー配列]-[デアミナーゼ]を含んでいる。「ヌクレアーゼ不活性なRGN」は、ヌクレアーゼ不活性へと変異した任意のRGN（任意のCRISPR-Casタンパク質が含まれる）を表わすことを理解すべきである。「UGI」は1つ以上のUGIドメインを表わすことも理解すべきである。

【0086】

別の実施態様では、融合タンパク質は、構造：[デアミナーゼ]-[オプションのリンカー配列]-[RGNニッカーゼ]-[オプションのリンカー配列]-[UGI]；[デアミナーゼ]-[オプションのリンカー配列]-[UGI]-[オプションのリンカー配列]-[RGNニッカーゼ]；[UGI]-[オプションのリンカー配列]-[デアミナーゼ]-[オプションのリンカー配列]-[RGNニッカーゼ]；[UGI]-[オプションのリンカー配列]-[RGNニッカーゼ]-[オプションのリンカー配列]-[デアミナーゼ]；または[RGNニッカーゼ]-[オプションのリンカー配列]-[デアミナーゼ]-[オプションのリンカー配列]-[UGI]；または[RGNニッカーゼ]-[オプションのリンカー配列]-[UGI]-[オプションのリンカー配列]-[デアミナーゼ]を含んでいる。「RGNニッカーゼ」は、ニッカーゼとして活性であるように変異した任意のRGN（任意のCRISPR-Casタンパク質が含まれる）を表わすことを理解すべきである。「UGI」は1つ以上のUGIドメインを表わすことも理解すべきである。

【0087】

いくつかの実施態様では、本明細書に提示されている融合タンパク質はリンカー配列を含んでいない。いくつかの実施態様では、オプションであるリンカー配列の一方または両方が存在している。

【0088】

いくつかの実施態様では、上記の一般構造の中で用いられている「-」は、オプションであるリンカー配列の存在を示す。いくつかの実施態様では、UGIを含む融合タンパク質は、核を標的とする配列（例えば核局在化配列）をさらに含んでいる。いくつかの実施態様では、本明細書に提示されている融合タンパク質は核局在化配列（NLS）をさらに含んでいる。いくつかの実施態様では、NLSは融合タンパク質のN末端に融合される。いくつかの実施態様では、NLSは融合タンパク質のC末端に融合される。いくつかの実施態様では、NLSはUGIタンパク質のN末端に融合される。いくつかの実施態様では、NLSはUGIタンパク質のC末端に融合される。いくつかの実施態様では、NLSはRGNタンパク質のN末端に融合される。いくつかの実施態様では、NLSはRGNタンパク質のC末端に融合される。い

10

20

30

40

50

いくつかの実施態様では、NLSはデアミナーゼのN末端に融合される。いくつかの実施態様では、NLSはデアミナーゼのC末端に融合される。いくつかの実施態様では、NLSは第2のRGNタンパク質のN末端に融合される。いくつかの実施態様では、NLSは第2のRGNタンパク質のC末端に融合される。いくつかの実施態様では、NLSは1つ以上のリンカーを介して融合タンパク質に融合される。いくつかの実施態様では、NLSはリンカーなしで融合タンパク質に融合される。いくつかの実施態様では、NLSは、本明細書に提示されているか、本明細書で参照されているNLS配列のいずれか1つのアミノ酸配列を含む。いくつかの実施態様では、NLSは、配列番号10として示されているアミノ酸配列を含む。

#### 【0089】

いくつかの実施態様では、UGIドメインは、野生型UGI、または配列番号570として示されているUGIを含んでいる。いくつかの実施態様では、本明細書に提示されているUGIは、UGIの断片と、UGIまたはUGI断片と相同なタンパク質を含んでいる。例えばいくつかの実施態様では、UGIドメインは、配列番号570として示されているアミノ酸配列の断片を含んでいる。いくつかの実施態様では、UGI断片は、配列番号570として示されているアミノ酸配列の少なくとも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%、または少なくとも99.5%を含むアミノ酸配列を含んでいる。いくつかの実施態様では、UGIは、配列番号570として示されているアミノ酸配列と相同なアミノ酸配列、または配列番号570として示されているアミノ酸配列の断片と相同なアミノ酸配列を含んでいる。いくつかの実施態様では、UGI、またはUGIの断片、またはUGIまたはUGI断片のホモログを含むタンパク質を「UGIバリエーション」と呼ぶ。UGIバリエーションは、UGIまたはその断片との相同性を共有する。例えばUGIバリエーションは、野生型UGIまたは配列番号570として示されているUGIと少なくとも70%一致する、少なくとも75%一致する、少なくとも80%一致する、少なくとも85%一致する、少なくとも90%一致する、少なくとも95%一致する、少なくとも96%一致する、少なくとも97%一致する、少なくとも98%一致する、少なくとも99%一致する、少なくとも99.5%一致する、または少なくとも99.9%一致する。いくつかの実施態様では、UGIバリエーションはUGIの断片を含んでいて、この断片は、野生型UGIまたは配列番号570として示されているUGIの対応する断片と少なくとも70%一致する、少なくとも80%一致する、少なくとも90%一致する、少なくとも95%一致する、少なくとも96%一致する、少なくとも97%一致する、少なくとも98%一致する、少なくとも99%一致する、少なくとも99.5%一致する、または少なくとも99.9%一致する。

#### 【0090】

適切なUGIタンパク質とヌクレオチド配列が本明細書に提示されるとともに、追加の適切なUGI配列が本分野で知られていて、その中には、例えばWang他、1989年、J. Biol. Chem. 第264巻：1163～1171ページ；Lundquist他、1997年、J. Biol. Chem. 第272巻：21408～21419ページ；Ravishankar他、1998年、Nucleic Acids Res. 第26巻：4880～4887ページ；およびPutnam他、1999年、J. Mol. Biol. 第287巻：331～346ページ（1999年）に記載されているものが含まれる（それぞれの全内容が参照によって本明細書に組み込まれている）。

#### 【0091】

追加のタンパク質としてウラシルグリコシラーゼ阻害剤が可能であることに注意されたい。例えばウラシル-DNAグリコシラーゼ塩基除去修復酵素を抑制する（例えば立体的に阻止する）ことのできる他のタンパク質が本開示の範囲に含まれる。それに加え、塩基除去修復を阻止または抑制するあらゆるタンパク質も本開示の範囲内として。いくつかの実施態様では、DNAに結合するタンパク質が使用される。別の一実施態様では、UGIの代替物が使用される。いくつかの実施態様では、ウラシルグリコシラーゼ阻害剤は、一本鎖DNAに結合するタンパク質である。例えばウラシルグリコシラーゼ阻害剤として、エルウィニア・タスマニエンシス的一本鎖結合タンパク質が可能である。いくつかの実施態様では、ウラシルグリコシラーゼ阻害剤は、ウラシルに結合するタンパク質である。いくつかの実

10

20

30

40

50

施態様では、ウラシルグリコシラーゼ阻害剤は、DNAの中のウラシルに結合するタンパク質である。いくつかの実施態様では、ウラシルグリコシラーゼ阻害剤は、触媒として不活性なウラシルDNAグリコシラーゼタンパク質である。いくつかの実施態様では、ウラシルグリコシラーゼ阻害剤は、触媒として不活性なウラシルDNAグリコシラーゼタンパク質であり、DNAからウラシルを切除しない。他のウラシルグリコシラーゼ阻害剤は当業者には明らかであろうゆえ、本開示の範囲に含まれることに注意すべきである。1つの適切なUGIタンパク質配列が本明細書に提示されており（配列番号570）、追加の適切なUGIタンパク質配列は当業者に知られている。その中に含まれるのは、例えばWang他、1989年、J. Biol. Chem. 第264巻：1163～1171ページ；Lundquist他、1997年、J. Biol. Chem. 第272巻：21408～21419ページ；Ravishankar他、1998年、Nucleic Acids Res. 第26巻：4880～4887ページ；およびPutnam他、1999年、J. Mol. Biol. 第287巻：331～346ページ（1999年）に記載されているものである（それぞれの全内容が参照によって本明細書に組み込まれている）。いくつかの実施態様では、オプシオンのリンカーは、(SGGS)<sub>n</sub>（配列番号592）モチーフを含んでいる（ただしnは、1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、または20である）。いくつかの実施態様では、オプシオンのリンカーは、配列番号546として示されているアミノ酸配列を含んでいる。

#### 【0092】

いくつかの実施態様では、本明細書に提示されている融合タンパク質は、細胞へのRGNの取り込みを容易にする少なくとも1つの細胞侵入ドメインを含んでいる。細胞侵入ドメインは本分野で知られており、一般に、一連の正に帯電したアミノ酸残基（すなわちポリカチオン性細胞侵入ドメイン）、交互に存在する極性アミノ酸残基と非極性アミノ酸残基（すなわち両親媒性細胞侵入ドメイン）、または疎水性アミノ酸残基（すなわち疎水性細胞侵入ドメイン）を含んでいる（例えばMilletti F.（2012年）Drug Discov Today 第17巻：850～860ページを参照されたい）。細胞侵入ドメインの非限定的な一例は、ヒト免疫不全ウイルス1からのトランス活性化転写アクチベータ（TAT）である。

#### 【0093】

核局在化シグナル、プラスチド局在化配列、ミトコンドリア局在化配列、二重標的シグナル配列、および/または細胞侵入ドメインは、RNA誘導型ヌクレアーゼのアミノ末端（N末端）、カルボキシル末端（C末端）、または内部位置に位置させることができる。

#### 【0094】

本発明の別の実施態様は、融合タンパク質と、単一ガイドRNAまたは二重ガイドRNAとしてのガイドRNA（まとめてgRNAと呼ぶ）を含むリボ核タンパク質複合体である。

#### 【0095】

IV. RNA誘導型ヌクレアーゼ、RNA誘導型DNA結合ポリペプチド、デアミナーゼ、CRISPR RNA、tracrRNA、およびgRNAをコードするヌクレオチド

#### 【0096】

本開示により、本開示のRGN、RNA誘導型DNA結合ポリペプチド-デアミナーゼ融合体、デアミナーゼ、CRISPR RNA、tracrRNA、および/またはsgRNAを含むポリヌクレオチドが提供される。本開示のポリヌクレオチドは、配列番号2、17、25、36、44、51、または63のヌクレオチド配列を含むCRISPR反復配列を含むかコードするポリヌクレオチド、またはその活性なバリエーションまたは断片を含んでおり、それは、ガイドRNAの中に含まれているとき、関連するRNA誘導型ヌクレアーゼを興味ある標的配列に配列特異的に結合させることができる。配列番号3、18、26、37、45、52、または62のヌクレオチド配列を含むtracrRNAを含むかコードするポリヌクレオチド、またはその活性なバリエーションまたは断片も開示されており、それは、ガイドRNAの中に含まれているとき、関連するRNA誘導型ヌクレアーゼを興味ある標的配列に配列特異的に結合させることができる。配列番号1、16、24、35、43、または50として表わされるアミノ酸配列を含むRGNをコードするポリヌクレオチドとその活性な断片またはバリエーションであって標的ヌクレアーゼ配列にRNA誘導型配列特異的なやり方で結合する能力を保持しているものも提供される。

## 【 0 0 9 7 】

本開示によりさらに、本明細書で言及されているデアミナーゼ（配列番号374～545と572～584、またはその活性なバリエーション）とDNA結合ポリペプチド（例えばメガヌクレアーゼ、ジンクフィンガー融合タンパク質、またはTALEN）を含む融合タンパク質をコードするポリヌクレオチドが提供される。本開示によりさらに、本明細書で言及されているデアミナーゼと、RNA誘導型DNA結合ポリペプチドを含む融合タンパク質をコードするポリヌクレオチドが提供される。このようなRNA誘導型DNA結合ポリペプチドとして、本発明のRGN、本分野で知られているRGN、CRISPR-Casタンパク質、またはこれらの任意のタンパク質バリエーションが可能である。タンパク質バリエーションは、ニックアーゼ不活性であること、またはニックアーゼであることが可能である。このようなRGNバリエーションの例に含まれるのは、ニックアーゼ不活性なRGN（配列番号568または配列番号547）、またはRGNニックアーゼ変異体（配列番号569または配列番号553）である。RGNヌクレアーゼの別の例は本分野において周知であり、同様の対応する変異が、やはりヌクレアーゼ不活性であるかニックアーゼである変異体バリエーションを作り出すことができる。

10

## 【 0 0 9 8 】

「ポリヌクレオチド」という用語の使用は、本開示ではDNAを含むポリヌクレオチドに限定されない。当業者は、ポリヌクレオチドに、リボヌクレオチド（RNA）と、リボヌクレオチドとデオキシリボヌクレオチドの組み合わせを含めることが可能であることを認識しているであろう。このようなデオキシリボヌクレオチドとリボヌクレオチドに、天然の分子と合成類似体の両方が含まれる。これらの中には、ペプチド核酸（PNA）、PNA-DNAキメラ、架橋型核酸（LNA）、およびホスホチオレート架橋配列が含まれる。本明細書に開示されているポリヌクレオチドにはあらゆる形態の配列も包含され、その非限定的な例に含まれるのは、一本鎖の形態、二本鎖の形態、DNA-RNAハイブリッド、三本鎖構造、ステム-ループ構造などである。

20

## 【 0 0 9 9 】

RGN、デアミナーゼ、または融合タンパク質をコードする核酸分子は、興味ある生物内での発現のためコドンが最適化されているようにすることができる。「コドンが最適化された」コード配列は、特定の宿主細胞の好ましいコドンの使用頻度または転写条件を模倣するように設計されたコドン使用頻度を持つ配列をコードするポリヌクレオチドである。核酸レベルで1個以上のコドンが変化しているが翻訳されるアミノ酸配列は変化していない結果として、その特定の宿主細胞または生物における発現が増加する。核酸分子の全体または一部についてコドン最適化をすることができる。広い範囲の生物について好ましい情報を提供するコドン表とそれ以外の参考文献を本分野で入手することができる（植物が好むコドンの利用に関しては、例えばCampbellとGowri（1990年）Plant Physiol. 第92巻：1～11ページを参照されたい）。本分野では、植物が好む遺伝子を合成するための方法を利用することができる。例えばアメリカ合衆国特許第5,380,831号および第5,436,391号と、Murray他（1989年）Nucleic Acids Res. 第17巻：477～498ページを参照されたい（参照によって本明細書に組み込まれている）。

30

## 【 0 1 0 0 】

本明細書に提示されているRGN、RNA誘導型DNA結合ポリペプチド-デアミナーゼ融合体、デアミナーゼ、crRNA、tracrRNA、および/またはsgRNAをコードするポリヌクレオチドを発現カセットに入れて提供し、インビトロで発現させること、または興味ある細胞、オルガネラ、胚、または生物の中で発現させることができる。カセットは、本明細書に提示されているRGN、RNA誘導型DNA結合ポリペプチド-デアミナーゼ融合体、デアミナーゼ、crRNA、tracrRNA、および/またはsgRNAをコードするポリヌクレオチドに機能可能に連結された5'調節配列と3'調節配列を含むことで、このポリヌクレオチドの発現が可能になる。カセットはさらに、その生物に入れて同時形質転換するための少なくとも1つの追加の遺伝子または遺伝子エレメントを含むことができる。追加の遺伝子またはエレメントが含まれている場合には、これらの要素は機能可能に連結されている。「機能可能に連結された」という表現は、2つ以上のエレメントの間の機能的な連結を意味する

40

50

。例えばプロモータと興味あるコード領域（例えばRGN、RNA誘導型DNA結合ポリペプチド-デアミナーゼ融合体、デアミナーゼ、crRNA、tracrRNA、および/またはsgRNAをコードする領域）の間の機能可能な連結は、興味あるこのコード領域の発現を可能にする機能的な連結である。機能可能に連結されたエレメントは、連続していても不連続でもよい。機能可能に連結されたが、2つのタンパク質コード領域の接合に言及するのに用いられているときには、それらコード領域が同じリーディングフレームの中にあることを意味する。あるいは追加の遺伝子またはエレメントは、複数の発現カセット上に提供することができる。例えば本開示のRGNをコードするヌクレオチド配列は1つの発現カセット上にあるのに対し、crRNA、tracrRNA、または完全なガイドRNAをコードするヌクレオチド配列は別の発現カセット上にあることが可能である。このような発現カセットは、調節領域の転写調節下にあるポリヌクレオチドを挿入するための複数の制限部位および/または組み換え部位とともに提供される。発現カセットは、選択マーカー遺伝子を追加して含むことができる。

#### 【0101】

発現カセットは、転写の5' 3'方向に向かって、興味ある生物の中で機能する転写（と、いくつかの場合には翻訳）開始領域（すなわちプロモータ）、本発明のRGN-、RNA誘導型DNA結合ポリペプチド-デアミナーゼ融合体-、デアミナーゼ-、crRNA-、tracrRNA-、および/またはsgRNAをコードするポリヌクレオチド、および転写（と、いくつかの場合には翻訳）終止領域（すなわち終止領域）を含むことになる。本発明のプロモータは、宿主細胞の中でコード配列の発現を指示または命令することができる。これらの調節領域（例えばプロモータ、転写調節領域、および翻訳終止領域）は、宿主細胞にとって、または互いに内在性または異種であることが可能である。配列に関して本明細書で用いられている「異種の」は、外来種に由来する配列であるか、同じ種に由来する場合には、人為的な介入によって組成および/またはゲノム遺伝子座の中の天然の形態から実質的に改変されている配列である。本明細書で用いられているように、キメラ遺伝子は、コード配列として、このコード配列にとって異種の転写開始領域に機能可能に連結されたコード配列を含んでいる。

#### 【0102】

便利な終止領域（例えばオクトピンシンターゼ終止領域とノバリンシンターゼ終止領域）をA.ツメファシエンスのTiプラスミドから入手することができる。Guerineau他（1991年）Mol. Gen. Genet. 第262巻：141～144ページ；Proudfoot（1991年）Cell第64巻：671～674ページ；Sanfacon他（1991年）Genes Dev. 第5巻：141～149ページ；Mogen他（1990年）Plant Cell第2巻：1261～1272ページ；Munroe他（1990年）Gene第91巻：151～158ページ；Ballas他（1989年）Nucleic Acids Res. 第17巻：7891～7903ページ；およびJoshi他（1987年）Nucleic Acids Res. 第15巻：9627～9639ページも参照されたい。

#### 【0103】

追加の調節シグナルの非限定的な例に含まれるのは、転写開始部位、オペレータ、アクチベータ、エンハンサ、他の調節エレメント、リボソーム結合部位、開始コドン、終止シグナルなどである。例えばアメリカ合衆国特許第5,039,523号と第4,853,331号；EPO 0480762A2；Sambrook他（1992年）『Molecular Cloning: A Laboratory Manual』、Maniatis他編（Cold Spring Harbor Laboratory Press、コールド・スプリング・ハーバー、ニューヨーク州）（今後は「Sambrook 11」と呼ぶ）；Davis他編（1980年）『Advanced Bacterial Genetics』（Cold Spring Harbor Laboratory Press）、コールド・スプリング・ハーバー、ニューヨーク州と、これらの中で引用されている参考文献を参照されたい。

#### 【0104】

発現カセットを用意する際には、さまざまなDNA断片を操作し、適切な向きと、必要に応じて適切なリーディングフレームにされたDNA配列が提供されるようにする。この目的に向け、アダプタまたはリンカーを用いてDNA断片を接合すること、または便利な制限部

10

20

30

40

50

位の提供、余分なDNAの除去、制限部位の除去などのための他の操作を含めることができる。この目的で、インビトロの突然変異誘発、プライマーの修復、制限、アニーリング、再置換（例えば転位（トランジション）と転換（トランスバージョン））を含めることができる。

#### 【0105】

本発明を実施するのに多数のプロモータを使用することができる。プロモータは、望む結果に基づいて選択することができる。核酸を、構成的プロモータ、誘導性プロモータ、増殖段階特異的プロモータ、細胞タイプ特異的プロモータ、組織選択的プロモータ、組織特異的プロモータ、または他のプロモータと組み合わせ、興味ある生物の中で発現させることができる。例えばWO 99/43838と、アメリカ合衆国特許第8,575,425号；第7,790,846号；第8,147,856号；第8,586,832号；第7,772,369号；第7,534,939号；第6,072,050号；第5,659,026号；第5,608,149号；第5,608,144号；第5,604,121号；第5,569,597号；第5,466,785号；第5,399,680号；第5,268,463号；第5,608,142号；および第6,177,611号に記載されているプロモータを参照されたい（これらは参照によって本明細書に組み込まれている）。

10

#### 【0106】

植物の中で発現させるため、構成的プロモータには、CaMV 35Sプロモータ（Odell他（1985年）Nature第313巻：810～812ページ）；イネのアクチン（McElroy他（1990年）Plant Cell第2巻：163～171ページ）；ユビキチン（Christensen他（1989年）Plant Mol. Biol. 第12巻：619～632ページ、Christensen他（1992年）Plant Mol. Biol. 第18巻：675～689ページ）；pEMU（Last他（1991年）Theor. Appl. Genet. 第81巻：581～588ページ）；およびMAS（Velten他（1984年）EMBO J. 第3巻：2723～2730ページ）も含まれる。

20

#### 【0107】

誘導性プロモータの例は、低酸素ストレスまたは低温ストレスによって誘導可能なAdh1プロモータ、熱ストレスによって誘導可能なHsp70プロモータ、光によって誘導可能なPDKプロモータとペブカルボキシラーゼプロモータである。化学的に誘導可能なプロモータも有用であり、それは例えば、薬害軽減剤によって誘導されるIn2-2プロモータ（アメリカ合衆国特許第5,364,780号）、オーキシンによって誘導され、タベート組織特異的だがカルスの中でも活性なAxig1プロモータ（PCT US01/22169）、ステロイド反応性プロモータ（例えばエストロゲンによって誘導されるEREプロモータと、Schena他（1991年）Proc. Natl. Acad. Sci. USA第88巻：10421～10425ページとMcNellis他（1998年）Plant J. 第14巻(2)：247～257ページの中のグルココルチコイド誘導性プロモータを参照されたい）、およびテトラサイクリン誘導性プロモータとテトラサイクリン抑制性プロモータ（例えばGatz他（1991年）Mol. Gen. Genet. 第227巻：229～237ページ、およびアメリカ合衆国特許第5,814,618号と第5,789,156号を参照されたい）である（これらの文献は参照によって本明細書に組み込まれている）。

30

#### 【0108】

組織特異的または組織選択的なプロモータは、特定の組織内で発現コンストラクトを発現させることを目的として用いることができる。いくつかの実施態様では、組織特異的または組織選択的なプロモータは、植物組織の中で活性である。植物の中で発現制御下にあるプロモータの例に含まれるのは、所定の組織（例えば葉、根、果実、種子、または花）の中で選択的に転写を開始するプロモータである。「組織特異的」プロモータは、所定の組織の中でだけ転写を開始するプロモータである。組織特異的発現は、遺伝子の構成的発現とは異なり、遺伝子調節の相互作用のレベルがいくつかあることの結果である。そのため、相同な植物種、または近縁の植物種からのプロモータを選択的に用いて特定の組織内で導入遺伝子の効果的かつ信頼性の高い発現を実現することができる。いくつかの実施態様では、発現は、組織選択的プロモータを含む。「組織選択的」プロモータは、選択的に転写を開始するが、必ずしも全面的に所定の組織の中で、または所定の組織の中だけで転写を開始するのではないプロモータである。

40

50

## 【0109】

いくつかの実施態様では、RGN、RNA誘導型DNA結合ポリペプチド-デアミナーゼ融合体、デアミナーゼ、crRNA、および/またはtracrRNAをコードする核酸分子は、細胞タイプ特異的プロモータを含んでいる。「細胞タイプ特異的」プロモータは、1つ以上の器官内のいくつかのタイプの細胞の中での発現を主に駆動するプロモータである。植物の中で機能する細胞タイプ特異的プロモータが主に活性であることのできる植物細胞のいくつかの例に含まれるのは、例えばBETL細胞、根、葉の中の維管束細胞、茎細胞、および幹細胞である。核酸分子は、細胞タイプ選択的プロモータも含むことができる。「細胞タイプ選択的」プロモータは、たいていは1つ以上の器官内のいくつかのタイプの細胞の中で発現を主に駆動するが、必ずしも全面的に所定の組織の中で、または所定の組織の中だけで発現を駆動するのではないプロモータである。植物の中で機能する細胞タイプ選択的プロモータが選択的に活性であることのできる植物細胞のいくつかの例に含まれるのは、例えばBELT細胞、根、葉の中の維管束細胞、茎細胞、および幹細胞である。

10

## 【0110】

RGN、RNA誘導型DNA結合ポリペプチド-デアミナーゼ融合体、デアミナーゼ、crRNA、tracrRNA、および/またはsgRNAをコードする核酸分子は、例えばインビトロでのmRNA合成のため、ファージRNAポリメラーゼによって認識されるプロモータ配列に機能可能に連結させることができる。このような実施態様では、インビトロで転写されるRNAを精製し、本明細書に記載されている方法でを使用することができる。例えばプロモータ配列として、T7プロモータ配列、T3プロモータ配列、またはSP6プロモータ配列、またはT7プロモータ配列、T3プロモータ配列、またはSP6プロモータ配列のバリエーションが可能である。このような実施態様では、発現したタンパク質および/またはRNAを精製し、本明細書に記載されているゲノム改変の方法でを使用することができる。

20

## 【0111】

いくつかの実施態様では、RGN、RNA誘導型DNA結合ポリペプチド-デアミナーゼ融合体、デアミナーゼ、crRNA、tracrRNA、および/またはsgRNAをコードするポリヌクレオチドは、ポリアデニル化シグナル（例えばSV40ポリAシグナルと、植物の中で機能する他のシグナル）および/または少なくとも1つの転写終止配列に連結することもできる。それに加え、本明細書の別の箇所に記載されているように、RGN、RNA誘導型DNA結合ポリペプチド-デアミナーゼ融合体、またはデアミナーゼをコードする配列は、タンパク質を特定の細胞内位置に移動させることのできる少なくとも1つの核局在化シグナル、少なくとも1つの細胞侵入ドメイン、および/または少なくとも1つのシグナルペプチドをコードする配列に連結することもできる。

30

## 【0112】

RGN、RNA誘導型DNA結合ポリペプチド-デアミナーゼ融合体、デアミナーゼ、crRNA、tracrRNA、および/またはsgRNAをコードするポリヌクレオチドは、1つのベクター、または複数のベクターの中に存在することができる。「ベクター」は、核酸を宿主細胞の中に移す、送達する、または導入するためのポリヌクレオチド組成物を意味する。適切なベクターに含まれるのは、プラスミドベクター、ファージミド、コスミド、人工/ミニ染色体、トランスポゾン、およびウイルスベクター（例えばレンチウイルスベクター、アデノ随伴ウイルスベクター、パキウイルスベクター）である。ベクターは、追加発現制御配列（例えばエンハンサ配列、コザック配列、ポリアデニル化配列、転写終止配列）、選択マーカー配列（例えば抗生剤耐性遺伝子）、複製起点などを含むことができる。追加情報は、『Current Protocols in Molecular Biology』Ausubel他、John Wiley & Sons社、ニューヨーク、2003年、または『Molecular Cloning: A Laboratory Manual』、SambrookとRussell、Cold Spring Harbor Press、コールド・スプリング・ハーバー、ニューヨーク州、第3版、2001年に見いだすことができる。

40

## 【0113】

ベクターは、形質転換された細胞を選択するための選択マーカー遺伝子も含むことができる。選択マーカー遺伝子は、形質転換された細胞または組織を選択するために使用され

50

る。マーカー遺伝子に含まれるのは、抗生剤耐性をコードする遺伝子（ネオマイシンホスホトランスフェラーゼII（NEO）とハイグロマイシンホスホトランスフェラーゼ（HPT）をコードする遺伝子など）のほか、殺虫化合物（例えばグルホシネートアンモニウム、プロモキシニル、イミダゾリノン、および2,4-ジクロロフェノキシ酢酸塩（2,4-D））に対する耐性を与える遺伝子である。

#### 【0114】

いくつかの実施態様では、RGNポリペプチド、RNA誘導型DNA結合ポリペプチド-デアミナーゼ融合体、またはデアミナーゼをコードする配列を含む発現カセットまたは発現ベクターは、ガイドRNAを生成させるため、crRNAおよび/またはtracrRNA、またはそのcrRNAとtracrRNAの組み合わせをコードする配列をさらに含むことができる。crRNAおよび/またはtracrRNAをコードする配列は、興味ある生物または宿主細胞の中でcrRNAおよび/またはtracrRNAを発現させるため、少なくとも1つの転写制御配列に機能可能に連結させることができる。例えばcrRNAおよび/またはtracrRNAをコードするポリヌクレオチドは、RNAポリメラーゼIII（Pol III）によって認識されるプロモータ配列に機能可能に連結させることができる。適切なPol IIIプロモータの非限定的な例に含まれるのは、哺乳動物のU6プロモータ、U3プロモータ、H1プロモータ、および7SL RNAプロモータと、イネのU6プロモータおよびU3プロモータである。

#### 【0115】

示されているように、RGN、RNA誘導型DNA結合ポリペプチド-デアミナーゼ融合体、デアミナーゼ、crRNA、tracrRNA、および/またはsgRNAをコードする発現コンストラクトを用いて興味ある生物を形質転換することができる。形質転換の方法は、ヌクレオチドコンストラクトを興味ある生物の中に導入することを含んでいる。「導入する」とは、ヌクレオチドコンストラクトを宿主細胞の中に導入し、そのコンストラクトが宿主細胞の内部にアクセスできることを意味する。本発明の方法は、ヌクレオチドコンストラクトを宿主生物の中に導入する特別な方法を必要とせず、ヌクレオチドコンストラクトが宿主生物の少なくとも1個の細胞の内部にアクセスできるだけでよい。宿主細胞として真核細胞または原核細胞が可能である。特別な実施態様では、真核細胞は、植物細胞、哺乳動物の細胞、または昆虫細胞である。ヌクレオチドコンストラクトを植物と他の宿主細胞に導入する方法は本分野で知られており、その非限定的な例に含まれるのは、安定に形質転換する方法、一過性に形質転換する方法、およびウイルスを媒介とする方法である。

#### 【0116】

これらの方法により、形質転換された生物（例えば植物であり、その中には植物体全体のほか、植物の器官（例えば葉、幹、根など）、種子、植物細胞、むかご、およびこれらの胚と子孫が含まれる）が得られる。植物細胞は、分化していること、または未分化であることが可能である（例えばカルス、懸濁培養細胞、プロトプラスト、葉細胞、根細胞、師管細胞、花粉）。

#### 【0117】

「トランスジェニック生物」または「形質転換された生物」または「安定に形質転換された」生物または細胞または組織は、本発明のRGN、RNA誘導型DNA結合ポリペプチド-デアミナーゼ融合体、デアミナーゼ、crRNA、および/またはtracrRNAをコードするポリヌクレオチドが組み込まれているか一体化している生物を意味する。他の外来性または内在性の核酸配列またはDNA断片も宿主細胞の中に組み込めることが知られている。アグロバクテリウムを媒介とした形質転換とバイオリスティックを媒介とした形質転換は、植物細胞を形質転換するのに相変わらず利用されている2つの主要なアプローチである。しかし宿主細胞の形質転換は、感染、トランスフェクション、微量注入、電気穿孔、マイクロプロジェクション、バイオリスティックまたは粒子衝突、電気穿孔、シリカ/炭素繊維、超音波を媒介とした方法、PEGを媒介とすること、リン酸カルシウム共沈降、ポリカチオンDMSO技術、DEAEデキストラン手続き、およびウイルスを媒介とすること、リポソームを媒介とすることなどによって実現することができる。RGN、RNA誘導型DNA結合ポリペプチド-デアミナーゼ融合体、デアミナーゼ、crRNA、および/またはtracrRNAをコ



ードするポリヌクレオチドのウイルスを媒介とした導入に含まれるのは、レトロウイルス、レンチウイルス、アデノウイルス、およびアデノ随伴ウイルスを媒介とした導入および発現のほか、カリモウイルス、ジェミニウイルス、およびRNA植物ウイルスの利用である。

#### 【0118】

形質転換のプロトコルのほか、ポリペプチド配列またはポリヌクレオチド配列を植物の中に導入するプロトコルは、形質転換の標的にされる宿主細胞のタイプ（例えば単子葉植物の細胞または双子葉植物の細胞）によって異なる可能性がある。形質転換の方法は本分野で知られており、その中に含まれるのは、アメリカ合衆国特許第8,575,425号；第7,692,068号；第8,802,934号；第7,541,517号に記載されている方法である（これらのそれぞれが参照によって本明細書に組み込まれている）。Rakoczy-Trojanowska, M. (2002年) Cell Mol Biol Lett. 第7巻：849～858ページ；Jones他 (2005年) Plant Methods 第1巻：5ページ；Rivera他 (2012年) Physics of Life Reviews 第9巻：308～345ページ；Bartlett他 (2008年) Plant Methods 第4巻：1～12ページ；Bates, G.W. (1999年) Methods in Molecular Biology 第111巻：359～366ページ；BinnsとThomashow (1988年) Annual Reviews in Microbiology 第42巻：575～606ページ；Christou, P. (1992年) The Plant Journal 第2巻：275～281ページ；Christou, P. (1995年) Euphytica 第85巻：13～27ページ；Tzfira他 (2004年) TRENDS in Genetics 第20巻：375～383ページ；Yao他 (2006年) Journal of Experimental Botany 第57巻：3737～3746ページ；ZupanとZambryski (1995年) Plant Physiology 第107巻：1041～1047ページ；Jones他 (2005年) Plant Methods 第1巻：5ページも参照されたい。

10

20

#### 【0119】

形質転換により、細胞の中に核酸を安定に、または一過性に組み込むことができる。「安定な形質転換」は、宿主細胞の中に導入されたヌクレオチドコンストラクトがその宿主細胞のゲノムに組み込まれ、その子孫によって遺伝させることができることを意味する。「一過性の形質転換」は、ポリヌクレオチドが宿主細胞の中に導入されるがその宿主細胞のゲノムに組み込まれないことを意味する。

#### 【0120】

葉緑体を形質転換する方法は本分野で知られている。例えばSvab他 (1990年) Proc. Natl. Acad. Sci. USA 第87巻：8526～8530ページ；SvabとMaliga (1993年) Proc. Natl. Acad. Sci. USA 第90巻：913～917ページ；SvabとMaliga (1993年) EMBO J. 第12巻：601～606ページを参照されたい。この方法は、選択マーカーを含有するDNAを粒子銃で送達することに頼っており、相同組み換えを通じてそのDNAをプラスチドのゲノムに向かわせる。それに加え、プラスチドの形質転換は、プラスチドが運ぶサイレントな導入遺伝子を、核がコードするプラスチド誘導型RNAポリメラーゼの組織選択的発現によって交差活性化することによって実現できる。このようなシステムは、McBride他 (1994年) Proc. Natl. Acad. Sci. USA 第91巻：7301～7305ページに報告されている。

30

#### 【0121】

形質転換された細胞を従来のやり方に従ってトランスジェニック生物（植物など）の中に入れて増殖させることができる。例えばMcCormick他 (1986年) Plant Cell Reports 第5巻：81～84ページを参照されたい。次いでこれらの植物を増殖させ、同じ形質転換された株、または異なる株を用いて受粉させ、得られたハイブリッドの中から望む表現型の特徴が構成的に発現しているものを特定する。2世代以上増殖させて望む表現型の特徴の発現が安定に維持されて遺伝することを確認した後、種子を回収して望む表現型の特徴が実現したことを確認する。このようにして、本発明により、本発明のヌクレオチドコンストラクト（例えば本発明の発現カセット）がゲノムの中に安定に組み込まれた形質転換種子（「トランスジェニック種子」とも呼ばれる）が提供される。

40

#### 【0122】

あるいは形質転換された細胞を生物の中に導入することができる。これらの細胞はその

50

生物に由来してもよく、その場合には、細胞は生体外アプローチで形質転換される。

【0123】

本明細書に提示されている配列は、任意の植物種（その非限定的な例に、単子葉植物と双子葉植物が含まれる）の形質転換に用いることができる。興味ある植物の非限定的な例に含まれるのは、コーン（トウモロコシ）、モロコシ、コムギ、ヒマワリ、トマト、アブラナ科植物、コショウ、ジャガイモ、木綿、イネ、ダイズ、テンサイ、サトウキビ、タバコ、オオムギ、ナタネ、アブラナ属の種、アルファルファ、ライムギ、雑穀、ベニバナ、ピーナツ、サツマイモ、キャッサバ、コーヒー、ココナツ、パイナップル、柑橘類、ココア、茶、バナナ、アボカド、イチジク、グアバ、マンゴ、オリーブ、パパイア、カシューナッツ、マカダミアナッツ、アーモンド、エンバク、野菜、装飾植物、および針葉樹である。

10

【0124】

野菜の非限定的な例に含まれるのは、トマト、レタス、グリーンピース、ライマメ、エンドウマメ、およびキュウリ属のメンバー（キュウリ、カンタループ、およびマスクメロンなど）である。装飾植物の非限定的な例に含まれるのは、アザレア、アジサイ、ハイビスカス、バラ、チューリップ、ラッパズイセン、ペチュニア、カーネーション、ポインセチア、および菊である。本発明の植物は、栽培作物（例えばトウモロコシ、モロコシ、コムギ、ヒマワリ、トマト、アブラナ科植物、コショウ、ジャガイモ、木綿、イネ、ダイズ、テンサイ、サトウキビ、タバコ、オオムギ、ナタネなど）であることが好ましい。

【0125】

20

本明細書では、植物という用語に含まれるのは、植物細胞、植物プロトプラスト、植物を再生させることのできる植物細胞組織培養物、植物カルス、植物の集団、および植物内で完全であるか植物の一部（胚、花粉、胚珠、種子、葉、花、枝、果実、仁、穂、穂軸、殻、茎、根、根冠、薬など）である植物細胞である。穀粒は、農作物栽培者が増殖または生殖以外の目的で生産した成熟した種子を意味する。再生した植物の子孫、バリエーション、および変異体も本発明の範囲に含まれるが、これらの部分が、導入されたポリヌクレオチドを含んでいることが条件である。さらに、本明細書に開示されている配列を保持している加工された植物製品または副生成物（例えば大豆粕）も提供される。

【0126】

RGN、RNA誘導型DNA結合ポリペプチド-デアミナーゼ融合体、デアミナーゼ、crRNA、および/またはtracrRNAをコードするポリヌクレオチドは、任意の原核生物種を形質転換するのにも使用できる。原核生物種の非限定的な例に含まれるのは、古細菌と細菌（例えばバシラス属の種、クレブシエラ属の種、ストレプトマイセス属の種、リゾビウム属の種、大腸菌属の種、シュドモナス属の種、サルモネラ属の種、赤痢菌属の種、ビブリオ属の種、エルシニア属の種、マイコプラズマ属の種、アグロバクテリウム、ラクトバシラス属の種）である。

30

【0127】

RGN、RNA誘導型DNA結合ポリペプチド-デアミナーゼ融合体、デアミナーゼ、crRNA、および/またはtracrRNAをコードするポリヌクレオチドは、任意の真核生物種を形質転換するのにも使用できる。真核生物種の非限定的な例に含まれるのは、動物（例えば哺乳動物と、昆虫、魚類、鳥類、および爬虫類）、菌類、アメーバ、藻類、および酵母である。

40

【0128】

ウイルスと非ウイルスに基づく従来からの遺伝子導入法を利用して核酸を哺乳動物の細胞または標的組織に導入することができる。このような方法を利用してCRISPRシステムの構成要素をコードする核酸を培養中の細胞または宿主生物に投与することができる。非ウイルスベクター送達システムに含まれるのは、DNAプラスミド、RNA（例えば本明細書に記載されているベクターの転写産物）、裸の核酸、および送達ビヒクル（リボソームなど）との複合体にされた核酸である。ウイルスベクター送達システムに含まれるのはDNAウイルスとRNAウイルスであり、これらのウイルスは細胞に送達された後にエピソード

50

ゲノムまたは組み込まれたゲノムを持つ。遺伝子療法の総説に関しては、Anderson、Science 第256巻：808～813ページ（1992年）；NabelとFeigner、TIBTECH第11巻：211～217ページ（1993年）；MitaniとCaskey、TIBTECH第11巻：162～166ページ（1993年）；Dillon、TIBTECH第11巻：167～175ページ（1993年）；Miller、Nature第357巻：455～460ページ（1992年）；Van Brunt、Biotechnology第6巻(10)：1149～1154ページ（1988年）；Vigne、Restorative Neurology and Neuroscience第8巻：35～36ページ（1995年）；KremerとPerricaudet、British Medical Bulletin第51巻(1)：31～44ページ（1995年）；『Current Topics in Microbiology and Immunology』、DoerflerとBohm（編）の中のHaddada他（1995年）；およびYu他、Gene Therapy第1巻：13～26ページ（1994年）を参照されたい。

10

#### 【0129】

核酸の非ウイルス送達法に含まれるのは、リポフェクション、ヌクレオフェクション、微量注入、バイオリスティック、ウイルスソーム、リポソーム、イムノリポソーム、ポリカチオンまたは脂質：核酸複合体、裸のDNA、人工ビリオン、およびDNAの薬剤増強式取り込みである。リポフェクションは、例えばアメリカ合衆国特許第5,049,386号、第4,946,787号、および第4,897,355号に記載されており、リポフェクション試薬が市販されている（例えばTransfectam（商標）とLipofectin（商標））。ポリヌクレオチドの効率的な受容体認識リポフェクションに適したカチオン性脂質と中性脂質に含まれるのは、FelgnerによるWO 91/17424；WO 91/16024の脂質である。送達は、細胞（例えばインビトロ投与または生体外投与）または標的組織（例えば生体内投与）に対してなすことができる。脂質：核酸複合体の調製は、標的とされるリポソーム（イムノ脂質複合体など）の調製を含め、当業者に周知である（例えばCrystal、Science第270巻：404～410ページ（1995年）；Blaese他、Cancer Gene Ther. 第2巻：291～297ページ（1995年）；Behr他、Bioconjugate Chem. 第5巻：382～389ページ（1994年）；Remy他、Bioconjugate Chem. 第5巻：647～654ページ（1994年）；Gao他、Gene Therapy第2巻：710～722ページ（1995年）；Ahmad他、Cancer Res. 第52巻：4817～4820ページ（1992年）；アメリカ合衆国特許第4,186,183号、第4,217,344号、第4,235,871号、第4,261,975号、第4,485,054号、第4,501,728号、第4,774,085号、第4,837,028号、および第4,946,787号を参照されたい）。

20

#### 【0130】

核酸の送達にRNAウイルスまたはDNAウイルスに基づくシステムを用いるというのは、ウイルスを体内の特定の細胞に向かわせてウイルスのペイロードを核に輸送するための高度に進化した方法を利用することである。ウイルスベクターは患者に直接投与することができる（生体内）。あるいはウイルスベクターをインビトロで細胞を処置するのに用い、この改変された細胞を場合によっては患者に投与することができる（生体外）。ウイルスに基づく従来のシステムには、遺伝子を輸送するためのレトロウイルスベクター、レンチウイルスベクター、アデノウイルスベクター、アデノ随伴ウイルスベクター、および単純ヘルペスウイルスベクターを含めることができると考えられる。宿主のゲノムへの組み込みは、レトロウイルス、レンチウイルス、およびアデノ随伴ウイルスによる遺伝子輸送法を用いることによって可能になり、その結果として挿入された導入遺伝子が長期にわたって発現することがしばしばある。それに加え、大きな遺伝子導入効率が、異なる多くのタイプの細胞と標的組織で観察されてきた。

30

40

#### 【0131】

レトロウイルスの指向性は、外来のエンベロープタンパク質を組み込み、標的細胞の潜在的な標的集団を増殖させることによって変えることができる。レンチウイルスベクターは、非分裂細胞に遺伝子導入または感染することのできるレトロウイルスベクターであり、典型的には高いウイルス力価を生み出す。したがってレトロウイルス遺伝子輸送システムの選択は、標的組織に依存すると考えられる。レトロウイルスベクターは、シス作用性の長い末端反復からなり、6～10 kbまでの外来配列をパッケージする能力を持つ。ベクターの複製とパッケージングには最小のシス作用性LTRで十分であり、このベクターを用

50

いて治療遺伝子を標的細胞の中に組み込み、導入遺伝子を恒久的に発現させる。広く使用されているレトロウイルスベクターに含まれるのは、マウス白血病ウイルス (MuLV)、ギボン白血病ウイルス (GaLV)、サル免疫不全ウイルス (SIV)、ヒト免疫不全ウイルス (HIV) に基づくベクターと、その組み合わせである (例えばBuchscher他、J. Viral. 第66巻: 2731~2739ページ (1992年); Johann他、J. Viral. 第66巻: 1635~1640ページ (1992年); Sommnerfelt他、J. Viral. 第176巻: 58~59ページ (1990年); Wilson他、J. Viral. 第63巻: 2374~2378ページ (1989年); Miller他、J. Viral. 第65巻: 2220~2224ページ (1991年); PCT/US94/05700を参照されたい)。

#### 【0132】

一過性発現が好ましい用途では、アデノウイルスに基づくシステムを用いることができる。アデノウイルスに基づくベクターは、多くのタイプの細胞で非常に高い遺伝子導入効率が可能であり、細胞分裂を必要としない。このようなベクターを用いて高力価と高い発現レベルが得られている。このベクターは、比較的簡単なシステムの中で大量に作製することができる。例えば核酸とペプチドをインピトロで作製するため、そして生体内と生体外の遺伝子療法のため、アデノ随伴ウイルス (「AAV」) ベクターを用いて細胞に標的核酸を導入することもできる (例えばWest他、Virology 第160巻: 38~47ページ (1987年); アメリカ合衆国特許第4,797,368号; WO 93/24641; Katin、Human Gene Therapy第5巻: 793~801ページ (1994年); Muzyczka、J. Clin. Invest. 第94巻: 1351ページ (1994年) を参照されたい)。組み換えAAVベクターの構築は多数の刊行物に記載されており、そうした刊行物に含まれるのは、アメリカ合衆国特許第5,173,414号; Tratschin他、Mol. Cell. Biol. 第5巻: 3251~3260ページ (1985年); Tratschin他、Mol. Cell. Biol. 第4巻: 2072~2081ページ (1984年); HermonatとMuzyczka、PNAS 第81巻: 6466~6470ページ (1984年); およびSamulski他、J. Viral. 第63巻: 03822~3828ページ (1989年) である。典型的にはパッケージング細胞を用い、宿主細胞に感染することのできるウイルス粒子を形成する。このような細胞に含まれるのは、293細胞 (アデノウイルスをパッケージする) と、J2細胞またはPA317細胞 (レトロウイルスをパッケージする) である。

#### 【0133】

遺伝子療法で用いられるウイルスベクターは、通常は、核酸ベクターをウイルス粒子の中にパッケージする細胞系を生成させることによって作製する。ベクターは、典型的には、パッケージングとその後の宿主への組み込みに必要な最小ウイルス配列を含有していて、他のウイルス配列は、発現することになるポリヌクレオチドのための発現カセットによって置き換えられる。失われるウイルス機能は、典型的には、パッケージング細胞系によってトランスで供給される。例えば遺伝子療法で用いられるAAVベクターは、典型的には、パッケージングと宿主ゲノムへの組み込みに必要なAAVゲノムからのITR配列だけを有する。ウイルスDNAは、他のAAV遺伝子 (すなわちrepとcap) をコードするがITR配列を欠くヘルパープラスミドを含有している細胞系の中にパッケージされる。

#### 【0134】

細胞系には、ヘルパーとしてアデノウイルスを感染させることもできる。ヘルパーウイルスは、AAVベクターの複製と、ヘルパープラスミドからのAAV遺伝子の発現を促進する。ヘルパープラスミドは、ITR配列が欠けていることが原因で大量にはパッケージされない。アデノウイルスを用いた汚染は、例えばAAVよりもアデノウイルスのほうがより感受性である熱処理によって減らすことができる。核酸を細胞に送達する追加の方法が当業者には知られている。例えばアメリカ合衆国特許出願公開第2003/0087817号を参照されたい (参照によって本明細書に組み込まれている)。

#### 【0135】

いくつかの実施態様では、本明細書に記載されている1つ以上のベクターを宿主細胞に一過性または非一過性にトランスフェクトする。いくつかの実施態様では、対象で自然に起こるのと同様にして細胞にトランスフェクトする。いくつかの実施態様では、トランスフェクトされた細胞を対象から採取する。いくつかの実施態様では、細胞は、対象 (細胞

10

20

30

40

50

系など)から採取された細胞に由来する。組織培養のための多彩な細胞系が本分野で知られている。細胞系の非限定的な例に含まれるのは、C8161、CCRF-CEM、MOLT、mIMC D-3、NHDF、HeLaS3、Huh1、Huh4、Huh7、HUVEC、HASMC、HEKn、HEKa、Mi aPaCell、Panel、PC-3、TFI、CTLL-2、CIR、Rat6、CVI、RPTE、AIO、T24、182、A375、ARH-77、Calu1、SW480、SW620、SKOV3、SK-UT、CaCo2、P388DI、S EM-K2、WEHI-231、HB56、TIB55、Iurkat、145.01、LRMB、Bcl-1、BC-3、IC21、DLD2、Raw264.7、NRK、NRK-52E、MRC5、MEF、Hep G2、HeLa B、HeLa T4、COS、COS-1、COS-6、COS-M6A、BS-C-1サル腎臓上皮、BALB/3T3マウス胚線維芽細胞、3T3 Swiss、3T3-LI、132-d5ヒト胎児線維芽細胞；10.1マウス線維芽細胞、29 3-T、3T3、721、9L、A2780、A2780ADR、A2780cis、A172、A20、A253、A43 1、A-549、ALC、B16、B35、BCP-I細胞、BEAS-2B、bEnd.3、BHK-21、BR 293、BxPC3、C3H-10TI/2、C6/36、Cal-27、CHO、CHO-7、CHO-IR、CHO-KI、CHO-K 2、CHO-T、CHO Dhfr-/-、COR-L23、COR-L23/CPR、COR-L235010、CORL23/ R 23、COS-7、COV-434、CML TI、CMT、CT26、D17、DH82、DU145、DuCaP、EL 4、EM2、EM3、EMT6/AR1、EMT6/AR10.0、FM3、H1299、H69、HB54、HB55、HCA2、HEK-293、HeLa、Hepalclc7、HL-60、HMEC、HT-29、Iurkat、IY細胞、K562細胞、Ku812、KCL22、KGI、KYOI、LNCap、Ma-Me1 1-48、MC-38、MCF-7、MCF-10A、MDA-MB-231、MDA-MB-468、MDA-MB-435、MDCKII、MDCKII、M OR/0.2R、MONO-MAC 6、MTD-1A、MyEnd、NCI-H69/CPR、NCI-H69/LX10、NC I-H69/LX20、NCI-H69/LX4、NIH-3T3、NALM-1、NW-145、OPCN/OPCT細胞系、Peer、PNT-1A/ PNT 2、RenCa、RIN-5F、RMA/RMAS、Saos-2細胞、Sf-9、SkBr 3、T2、T-47D、T84、THPI細胞系、U373、U87、U937、VCaP、Vero細胞、WM39、WT-49、X63、YAC-1、YARと、これらのトランスジェニックな変種である。当業者 に知られている多彩な供給源からの細胞系を利用することができる(例えばAmerican Ty pe Culture Collection(ATCC)(マナサス、ヴァージニア州)を参照されたい)。

#### 【0136】

いくつかの実施態様では、本明細書に記載されている1つ以上のベクターをトランスフ ェクトされた細胞を用い、1つ以上のベクター由来配列を含む新たな細胞系を確立する。 いくつかの実施態様では、本明細書に記載されているCRISPRシステムの構成要素、また はデアミナーゼ、またはその融合体を(例えば1つ以上のベクターの一過性トランスフェ クション、またはRNAのトランスフェクションによって)一過性にトランスフェクトされ てCRISPR複合体またはデアミナーゼの活性を通じて改変された細胞を用い、改変を含有 するが他の外来性配列は完全に欠けた新たな細胞系を確立する。いくつかの実施態様では、 本明細書に記載されている1つ以上のベクターを一過性または非一過性にトランスフェ クトされた細胞、またはそのような細胞に由来する細胞系を用いて1つ以上の試験化合物 を評価する。

#### 【0137】

いくつかの実施態様では、本明細書に記載されている1つ以上のベクターを用いて非ヒ トトランスジェニック動物またはトランスジェニック植物を作製する。いくつかの実施態 様では、トランスジェニック動物は、哺乳動物(例えばマウス、ラット、またはウサギ) である。いくつかの実施態様では、トランスジェニックヒト細胞が作製される。

#### 【0138】

V. ポリペプチドとポリヌクレオチドのバリエーションと断片

#### 【0139】

本開示により、天然(すなわち野生型)のRNA誘導型ヌクレアーゼとデアミナーゼ(そ のアミノ酸配列は、配列番号1、16、24、35、43、50、374~545、572~590として 表わされる)の活性なバリエーションと断片、およびその活性なバリエーションのほか、天然のCR ISPR反復(例えば配列番号2、17、25、36、44、51、または63として表わされる配列 )の活性なバリエーションと断片、および天然のtracrRNA(例えば配列番号3、18、26、37、 45、52、または62として表わされる配列)の活性なバリエーションと断片、およびこれら

をコードするポリヌクレオチドが提供される。デアミナーゼ（例えば配列番号374～545および572～584として示されている配列）の活性なバリエーションと断片も提供される。

【0140】

バリエーションまたは断片の活性は、興味あるポリヌクレオチドまたはポリペプチドと比べて変化していてもよいが、そのバリエーションと断片は、興味あるそのポリヌクレオチドまたはポリペプチドの機能性を保持していなければならない。例えばバリエーションまたは断片は、興味あるポリヌクレオチドまたはポリペプチドと比べたとき、活性が増大している可能性、活性が低下している可能性、活性のスペクトルが異なっている可能性、または活性が他の何らかの変化をしている可能性がある。

【0141】

天然のRGNポリペプチドの断片とバリエーション（本明細書に開示されているものなど）は、配列特異的なRNA誘導型DNA結合活性を保持することになる。特別な実施態様では、天然のRGNポリペプチドの断片とバリエーション（本明細書に開示されているものなど）は、ヌクレアーゼ活性を保持することになる（一本鎖または二本鎖）。別の実施態様では、天然のデアミナーゼの断片とバリエーション（本明細書に開示されているものなど）は、デアミナーゼ活性を保持することになる。いくつかの実施態様では、デアミナーゼバリエーションは、活性（例えばDNA鋳型に対する活性）が変化しているか、ヌクレオチドに対する活性（例えばアデノシンに対する活性）が天然のデアミナーゼとは異なっている

【0142】

天然のCRISPR反復の断片とバリエーション（本明細書に開示されているものなど）は、（*t* racrRNAを含む）ガイドRNAの一部であるとき、RNA誘導型ヌクレアーゼに結合して（ガイドRNAとの複合体になり）配列特異的なやり方で標的ヌクレオチド配列を標的とするよう導く能力を保持することになる。

【0143】

天然の*tracrRNA*の断片とバリエーション（本明細書に開示されているものなど）は、（CRISPR RNAを含む）ガイドRNAの一部であるとき、RNA誘導型ヌクレアーゼに結合して（ガイドRNAとの複合体になり）配列特異的なやり方で標的ヌクレオチド配列を標的とするよう導く能力を保持することになる。

【0144】

「断片」という用語は、本発明のポリヌクレオチド配列またはポリペプチド配列の一部を意味する。「断片」または「生物活性な部分」には、生物活性を保持する（すなわち、ガイドRNAの中に含まれているとき、RGNに結合してこのRGNを配列特異的に標的ヌクレオチド配列へと向かわせる）のに十分な数の連続したヌクレオチドを含むポリヌクレオチドが含まれる。「断片」または「生物活性な部分」には、生物活性を保持する（すなわち、ガイドRNAとの複合体になったとき、標的ヌクレオチド配列に配列特異的に結合する）のに十分な数の連続したアミノ酸残基を含むポリペプチドが含まれる。RGNタンパク質の断片には、代替の下流開始部位を使用することが理由で完全長配列よりも短いものが含まれる。RGNタンパク質の生物活性な部分として、配列番号1、16、24、35、43、または50の例えば10個、25個、50個、100個、150個、200個、250個、300個、350個、400個、450個、500個、550個、600個、650個、700個、750個、800個、850個、900個、950個、1000個、1050個、またはそれよりも多い個数の連続したアミノ酸残基を含むポリペプチドが可能である。このような生物活性な部分は、組み換え技術によって調製した後、配列特異的なRNA誘導型DNA結合活性について評価することができる。CRISPR反復配列の生物活性な断片は、配列番号2、17、25、36、44、51、または63の少なくとも8個の連続した核酸を含むことができる。CRISPR反復配列の生物活性な部分として、配列番号2、17、25、36、44、51、または63の例えば8個、9個、10個、11個、12個、13個、14個、15個、16個、17個、18個、19個、または20個の連続したヌクレオチドを含むポリヌクレオチドが可能である。*tracrRNA*の生物活性な部分として、配列番号3、18、26、37、45、52、または62の例えば8個、9個、10個、11個、12個、13個、14個、15個、16個、17個、18個、19個、20個、21個、22個、23個、24個、25個

10

20

30

40

50

、26個、27個、28個、29個、30個、35個、40個、45個、50個、55個、60個、65個、70個、75個、80個、またはそれよりも多い個数の連続したヌクレオチドを含むポリヌクレオチドが可能である。デアミナーゼの生物活性な部分として、例えば、配列番号374～545および572～584のいずれか1つの10個、25個、50個、100個、150個、200個、またはそれよりも多い個数の連続したアミノ酸残基を含むポリペプチドが可能である。

#### 【0145】

一般に「バリエーション」は、実質的に似ている配列を意味する。ポリヌクレオチドに関しては、バリエーションは、天然のポリヌクレオチド内の1つ以上の内部部位に1個以上のヌクレオチドの欠失および/または付加、および/または天然のポリヌクレオチド内の1つ以上の内部部位に1個以上のヌクレオチドの置換を含む。本明細書で用いられている「天然」または「野生型」のポリヌクレオチドまたはポリペプチドは、天然のヌクレオチド配列またはアミノ酸配列をそれぞれ含んでいる。ポリヌクレオチドに関しては、保存されたバリエーションに、遺伝暗号の縮重が理由で興味ある遺伝子の天然のアミノ酸配列をコードしている配列が含まれる。このような天然のアレルバリエーションは、分子生物学の周知の技術（例えば下に概略を記載するポリメラーゼ連鎖反応（PCR）とハイブリダイゼーション技術）を利用して同定することができる。バリエーションポリヌクレオチドには、合成に由来するが、興味あるポリペプチドまたはポリヌクレオチドをやはりコードしているポリヌクレオチド（例えば部位指定突然変異誘発によって生成したポリヌクレオチド）も含まれる。一般に、本明細書に開示されている特別なポリヌクレオチドのバリエーションは、本明細書の別の箇所に記載されている配列アラインメントプログラムとパラメータによって調べると、その特別なポリヌクレオチドと配列が少なくとも約40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、またはそれよりも多く一致するであろう。

#### 【0146】

本明細書に開示されている1つの特別なポリヌクレオチド（すなわち参照ポリヌクレオチド）のバリエーションは、バリエーションポリヌクレオチドによってコードされるポリペプチドと、参照ポリヌクレオチドによってコードされるポリペプチドの間の配列一致の割合の比較によって評価することもできる。任意の2つのポリペプチドの間の配列一致の割合は、本明細書の別の箇所に記載されている配列アラインメントプログラムとパラメータを用いて計算することができる。本明細書に開示されているポリヌクレオチドの任意のペアを、これらポリヌクレオチドがコードする2つのポリペプチドが共有する配列の一致率の比較によって評価する場合、コードされる2つのポリペプチドの間の配列一致の割合は、少なくとも約40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、またはそれよりも大きい。

#### 【0147】

特別な実施態様では、本開示のポリヌクレオチドは、配列番号374～545または572～584のいずれかのアミノ酸配列と少なくとも40%、45%、50%、55%、60%、65%、70%、75%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、またはそれよりも多く一致するアミノ酸配列を含むデアミナーゼポリペプチドをコードしている。いくつかの実施態様では、デアミナーゼは、配列番号374、383、397、399、407、408、411、414、416、420、514、および572～584のいずれかのアミノ酸配列と少なくとも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95%、少なくとも96%、少なくとも97%、少なくとも98%、または少なくとも99%配列が一致するアミノ酸配列を有する。これら実施態様のいくつかでは、バリエーションデアミナーゼポリペプチドは配列番号572～584のいずれか1つとの配列一致があるレベルであり、特定のアミノ酸残基が親配列から変化していない。例えばいくつかの実施態様では、配列番号572の1つのバリエーションは、配列番号572の102位に対応する位置にリシンを含み、配列番号5

72の104位に対応する位置にチロシンを含み、配列番号572の106位に対応する位置にトレオニンを含んでいる。特別な実施態様では、配列番号574の1つのバリエーションは、配列番号574の101位に対応する位置にグルタミン酸を含み、配列番号574の103位に対応する位置にセリンを含み、配列番号574の105位に対応する位置にリシンを含んでいる。いくつかの実施態様では、配列番号575の1つのバリエーションは、配列番号575の101位に対応する位置にリシンを含み、配列番号575の103位に対応する位置にロイシンを含み、配列番号575の105位に対応する位置にグルタミン酸を含んでいる。いくつかの実施態様では、配列番号576の1つのバリエーションは、配列番号576の105位に対応する位置にアラニンを含み、配列番号576の107位に対応する位置にアルギニンを含んでいる。特別な実施態様では、配列番号577の1つのバリエーションは、配列番号577の102位に対応する位置にグリシンを含み、配列番号577の104位に対応する位置にセリンを含み、配列番号577の106位に対応する位置にアルギニンを含んでいる。いくつかの実施態様では、配列番号578の1つのバリエーションは、配列番号578の105位に対応する位置にセリンを含み、配列番号578の107位に対応する位置にトレオニンを含んでいる。いくつかの実施態様では、配列番号579の1つのバリエーションは、配列番号579の102位に対応する位置にセリンを含み、配列番号579の104位に対応する位置にグルタミンを含み、配列番号579の106位に対応する位置にグリシンを含んでいる。特別な実施態様では、配列番号580の1つのバリエーションは、配列番号580の111位に対応する位置にグリシンを含んでいる。いくつかの実施態様では、配列番号581の1つのバリエーションは、配列番号581の104位に対応する位置にグルタミンを含み、配列番号581の106位に対応する位置にグリシンを含み、配列番号581の108位に対応する位置にグルタミン酸を含んでいる。いくつかの実施態様では、配列番号582の1つのバリエーションは、配列番号582の102位に対応する位置にアルギニンを含み、配列番号582の104位に対応する位置にトリプトファンを含み、配列番号582の106位に対応する位置にグルタミン酸を含んでいる。いくつかの実施態様では、配列番号583の1つのバリエーションは、配列番号583の104位に対応する位置にアルギニンを含み、配列番号583の106位に対応する位置にセリンを含んでいる。特別な実施態様では、配列番号584の1つのバリエーションは、配列番号584の110位に対応する位置にフェニルアラニンを含み、配列番号584の112位に対応する位置にセリンを含み、配列番号584の114位に対応する位置にトレオニンを含んでいる。

#### 【0148】

本発明のデアミナーゼポリペプチドの生物活性なバリエーションは、約1～15個という少数のアミノ酸残基、約1～10個という少数、例えば約6～10個、5個という少数、4個という少数、3個という少数、2個という少数、または1個という少数のアミノ酸残基が異なっている可能性がある。特別な実施態様では、これらポリペプチドはN末端またはC末端のトランケーションを含むことができ、そこにはポリペプチドのN末端またはC末端から10個、15個、20個、25個、30個、35個、40個、45個、50個、55個、60個、65個、70個、75個、80個、85個、90個、95個、100個、150個、200個、250個、300個、350個、400個、450個、500個、550個、600個、650個、700個、750個、800個、850個、900個、950個、1000個、1050個、またはそれよりも多い個数のアミノ酸の欠失が少なくとも含まれる可能性がある。

#### 【0149】

別の特別な実施態様では、本開示のポリヌクレオチドは、配列番号1、16、24、35、43、または50のアミノ酸配列と少なくとも40%、45%、50%、55%、60%、65%、70%、75%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、またはそれよりも多く一致するアミノ酸配列を含むRNA誘導型ヌクレアーゼポリペプチドをコードしている。

#### 【0150】

本発明のRGNポリペプチドまたはデアミナーゼポリペプチドの生物活性なバリエーションは、約1～15個という少数のアミノ酸残基、約1～10個という少数、例えば約6～10個、5

10

20

30

40

50



個という少数、4個という少数、3個という少数、2個という少数、または1個という少数のアミノ酸残基が異なっている可能性がある。特別な実施態様では、これらポリペプチドはN末端またはC末端のトランケーションを含むことができ、そこにはポリペプチドのN末端またはC末端から10個、15個、20個、25個、30個、35個、40個、45個、50個、55個、60個、65個、70個、75個、80個、85個、90個、95個、100個、150個、200個、250個、300個、350個、400個、450個、500個、550個、600個、650個、700個、750個、800個、850個、900個、950個、1000個、1050個、またはそれよりも多い個数のアミノ酸の欠失が少なくとも含まれる可能性がある。

#### 【0151】

いくつかの実施態様では、本開示のポリヌクレオチドは、配列番号2、17、25、36、44、51、または63として表わされるヌクレオチド配列と少なくとも40%、45%、50%、55%、60%、65%、70%、75%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、またはそれよりも多く一致するヌクレオチド配列を含むCRISPR反復を含むかコードしている。

10

#### 【0152】

本開示のポリヌクレオチドは、配列番号3、18、26、37、45、52、または62として表わされるヌクレオチド配列と少なくとも40%、45%、50%、55%、60%、65%、70%、75%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、またはそれよりも多く一致するヌクレオチド配列を含むtracrRNAを含むかコードすることができる。

20

#### 【0153】

本発明のCRISPR反復またはtracrRNAの生物活性なバリエーションは、約1~25個という少数のヌクレオチド、約1~20個という少数、約1~10個という少数、約6~10個という少数、5個という少数、4個という少数、3個という少数、2個という少数、または1個という少数のヌクレオチドが異なっている可能性がある。いくつかの実施態様では、これらポリヌクレオチドは5'または3'のトランケーションを含むことができ、そこにはポリヌクレオチドの5'末端または3'末端から10個、15個、20個、25個、30個、35個、40個、45個、50個、55個、60個、65個、70個、75個、80個、またはそれよりも多い個数のヌクレオチドの欠失が少なくとも含まれる可能性がある。いくつかの実施態様では、CRISPR反復またはtracrRNAは、欠失および/または挿入の両方によって変化している可能性がある。ヌクレオチドの変異または置換によって変化している可能性もある。

30

#### 【0154】

本明細書に提示されているRGNポリペプチド、DNA結合ポリペプチド-デアミナーゼ融合ポリペプチド、デアミナーゼポリペプチド、CRISPR反復、およびtracrRNAを改変し、バリエーションタンパク質とバリエーションポリヌクレオチドを作製できることがわかる。ヒトが設計した変化を、部位指定突然変異誘発技術を適用して導入することができる。あるいは本明細書に開示されている配列と構造および/または機能が関連した天然の、だがまだ知られていないかまだ同定されていないポリヌクレオチドおよび/またはポリペプチドで、本発明の範囲に入るものを同定することもできる。RGNタンパク質またはデアミナーゼタンパク質の機能を変化させない保存されたアミノ酸置換を、保存されていない領域で実現することができる。あるいはRGNまたはデアミナーゼの活性を向上させるか変化させる改変をなすことができる。

40

#### 【0155】

バリエーションポリヌクレオチドとバリエーションタンパク質には、突然変異と組み換えを起こさせる手続き(DNAシャッフリングなど)から得られる配列とタンパク質も包含される。このような手続きを利用して本明細書に開示されている1つ以上の異なるRGNタンパク質またはデアミナーゼタンパク質(例えば配列番号1、16、24、35、43、50、374~545、および572~584)を操作し、望む特性を有する新たなRGNタンパク質またはデアミナーゼタンパク質を作り出す。このようにして、関連した配列を持つポリヌクレオチドの集

50

団、すなわち実質的に一致した配列を持ち、インピトロまたは生体内で相同組み換えが可能な配列領域を含むポリヌクレオチドの集団から、組み換えポリヌクレオチドのライブラリを生成させる。例えばこのアプローチを用いると、興味あるドメインをコードする配列モチーフを、本明細書に提示されているRGN配列と既知の他のRGN遺伝子の間でシャッフルし、興味ある特性が改善された（例えば酵素の場合には $K_m$ が増大した）タンパク質をコードする新たな遺伝子を得ることができる。本明細書に提示されているデアミナーゼも同様の戦略でシャッフルすることができる。このようなDNAシャッフリングのための戦略は本分野で知られている。例えばStemmer（1994年）Proc. Natl. Acad. Sci. USA 第91巻：10747～10751ページ；Stemmer（1994年）Nature 第370巻：389～391ページ；Cramer他（1997年）Nature Biotech. 第15巻：436～438ページ；Moore他（1997年）J. Mol. Biol. 第272巻：336～347ページ；Zhang他（1997年）Proc. Natl. Acad. Sci. USA 第94巻：4504～4509ページ；Cramer他（1998年）Nature 第391巻：288～291ページ；およびアメリカ合衆国特許第5,605,793号と第5,837,458号を参照されたい。「シャッフルされた」核酸は、シャッフル手続き（例えば本明細書に記載されている任意のシャッフル手続き）によって生成した核酸である。シャッフルされた核酸は、2つ以上の核酸（または文字列）を例えば人工的な、そして場合によっては回帰的なやり方で（物理的または仮想的に）組み換えることによって生成する。一般に、シャッフリングプロセスでは1つ以上のスクリーニング工程を利用して興味ある核酸を同定する。このスクリーニング工程は、任意の組み換え工程の前または後に実施することができる。シャッフリングのいくつかの（だがすべてではない）実施態様では、選択の前に組み換えを多数回実施してスクリーニングするプールの多様性を大きくすることが望ましい。組み換えと選択のプロセス全体を場合によっては回帰的に繰り返す。文脈に応じ、シャッフリングは、組み換えと選択のプロセス全体を意味するか、プロセス全体の組み換え部分だけを意味することができる。

#### 【0156】

本明細書では、2つのポリヌクレオチドまたはポリペプチド配列の文脈における「配列一致」または「一致」は、指定された比較ウィンドウ上で対応が最大になるようにアラインメントしたとき、その2つの配列の中で同じである残基を意味する。配列一致の割合がタンパク質に関して用いられているときには、同じではない残基位置は、保存されたアミノ酸置換の違いであることがしばしばあり、その位置ではアミノ酸残基が似た化学的特性（例えば電荷または疎水性）の別のアミノ酸残基で置換されているため、分子の機能的特性は変化しないと認識されている。配列が保存的置換に関して異なっているときには、その置換の保存的性質に関して修正するため、配列一致の割合を上方に調節することができる。そのような保存的置換が異なる配列は、「配列類似性」または「類似性」を持つと言われる。この調節を行なう手段は当業者に周知である。典型的には、この手段は、保存的置換を完全なミスマッチではなく部分的なミスマッチとして点数化することで、配列一致の割合を大きくすることを含んでいる。したがって例えば一致するアミノ酸には1点を与え、保存されていない置換には0点を与えた場合、保存された置換には0と1の間の点数が与えられる。保存された置換の点数は、例えばプログラムPC/GENE（Intelligenetics社、マウンテン・ビュー、カリフォルニア州）の中に実現されているようにして計算される。

#### 【0157】

本明細書で用いられている「配列一致の割合」は、比較ウィンドウ上で最適なアラインメントにされた2つの配列を比較することによって求まる値を意味する。その場合、ポリヌクレオチド配列のうちで比較ウィンドウの中にある部分は、2つの配列の最適なアラインメントのための（付加または欠失を含まない）参照配列と比較して付加または欠失（すなわちギャップ）を含んでいる可能性がある。この割合は、両方の配列で核酸塩基またはアミノ酸残基が同じである位置の数を明らかにすることによって一致した位置の数を取得し、この一致した位置の数を、比較ウィンドウ内の位置の総数で割り、100を掛けることによって計算され、配列一致の割合が得られる。

#### 【0158】

10

20

30

40

50

特に断わらない限り、本明細書に提示されている配列一致／類似性の値は、GAP Version 10を用いて得られる値（以下のパラメータを用いる：ヌクレオチド配列に関しては、GAP Weightを50、Length Weightを3にし、nwsgapdna.cmpスコアリングマトリックスを用いて得られる％一致と％類似性；アミノ酸配列に関しては、GAP Weightを8、Length Weightを2にし、BLOSUM62スコアリングマトリックスを用いて得られる％一致と％類似性）、またはこれと同等な任意のプログラムを用いて得られる値を意味する。「同等なプログラム」とは、問題の任意の2つの配列について、GAP Version 10によって生成させた対応するアラインメントと比較したとき、ヌクレオチドまたはアミノ酸残基の一致が同じで、配列一致の割合が同じアラインメントを生じさせる任意の配列比較プログラムを意味する。

10

**【0159】**

2つの配列が「最適なアラインメントである」のは、類似性スコアを求めるため、所定のアミノ酸置換マトリックス（例えばBLOSUM62）、ギャップ存在ペナルティ、およびギャップ延長ペナルティを用いてアラインメントさせ、そのペアの配列に関して可能な最高スコアに到達したときである。アミノ酸置換マトリックスと、それを利用して2つの配列の間の類似性を定量化することは、本分野で周知であり、例えば『Atlas of Protein Sequence and Structure』、第5巻、補3（M. O. Dayhoff編）、Natl. Biomed. Res. Found.、ワシントンD.C.の中のDayhoff他（1978年）「タンパク質における進化的変化のモデル」、345～352ページと、Henikoff他（1992年）Proc. Natl. Acad. Sci. USA 第89巻：10915～10919ページに記載されている。BLOSUM62マトリックスは、配列アラインメントプロトコルにおいてデフォルトのスコアリング置換マトリックスとしてしばしば用いられる。ギャップ存在ペナルティは、アラインメントされた配列の1つに単一のアミノ酸ギャップを導入するために課され、ギャップ延長ペナルティは、すでに設けられているギャップの中にそれぞれ追加の空のアミノ酸位置を導入するために課される。アラインメントは、各配列の中で、そのアラインメントが始まるアミノ酸位置と終わるアミノ酸位置によって規定され、場合によっては一方または両方の配列に1つまたは複数のギャップを挿入して、可能な最高スコアが得られるようにする。最適なアラインメントと点数化は手作業で実現することができるが、このプロセスは、コンピュータで実現されるアラインメントアルゴリズム（例えばAltschul 他（1997年）Nucleic Acids Res. 第25巻：3389～3402ページに記載されているギャップ付きBLAST 2.0があり、National Center for Biotechnology Information Website（[www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)）で公開されているため入手できる）を用いると容易になる。最適なアラインメントは、多数のアラインメントを含め、例えばPSI-BLAST（Altschul 他（1997年）Nucleic Acids Res. 第25巻：3389～3402ページに記載されており、[www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)から入手することができる）を用いて用意することができる。

20

30

**【0160】**

参照配列と最適なアラインメントにされたアミノ酸配列に関しては、1個のアミノ酸残基は、アラインメントの中でこの残基とペアになる参照配列中の位置に「対応する」。「位置」は、参照配列の中でN末端に対する相対的な位置に基づいて逐次的に同定される各アミノ酸の番号によって表わされる。最適なアラインメントを求めるときに考慮すべき欠失、挿入、トランケーション、融合などがあるため、一般に、N末端から単純に数えることによって求めたときの試験配列中のアミノ酸残基の番号は、参照配列の中の対応する位置の番号と必ずしも同じにならない。例えばアラインメントされた試験配列の中に欠失が1つ存在している場合には、参照配列の中でその欠失部位の位置に対応するアミノ酸は存在しない。アラインメントされた試験配列の中に挿入が1つ存在している場合には、その挿入は、参照配列の中のどのアミノ酸位置にも対応しない。トランケーションまたは融合がある場合には、参照配列またはアラインメントされた配列の中に、対応する配列中のどのアミノ酸にも対応しない一連のアミノ酸が存在する可能性がある。

40

**【0161】**

VI. 抗体

50

## 【0162】

RGNポリペプチド、本発明のRGNポリペプチドを含むリボ核タンパク質、デアミナーゼ、またはDNA結合デアミナーゼ融合タンパク質（これらの中には、配列番号1、16、24、35、43、50、374～545、および572～584として表わされているアミノ酸配列を含むもの、またはその活性なバリエーションまたは断片が含まれる）に対する抗体も包含される。抗体を作製する方法は本分野で周知である（例えばHarlowとLane（1988年）『Antibodies: A Laboratory Manual』、Cold Spring Harbor Laboratory、コールド・スプリング・ハーバー、ニューヨーク州と；アメリカ合衆国特許第4,196,265号を参照されたい）。これらの抗体は、RGNポリペプチドまたはリボ核タンパク質を検出して単離するためのキットで用いることができる。したがって本開示により、本明細書に記載されているポリペプチドまたはリボ核タンパク質（その中には、例えば配列番号1、16、24、35、43、50、374～545、および572～584のいずれか1つの配列を持つポリペプチドが含まれる）に特異的に結合する抗体を含むキットが提供される。

10

## 【0163】

VII. 興味ある標的配列に結合させるためのシステムとリボ核タンパク質複合体、およびそれを作製する方法

## 【0164】

本開示により、興味ある標的配列に結合させるためのシステムが提供される。このシステムは、少なくとも1つのガイドRNA、またはそれをコードするヌクレオチド配列と、少なくとも1つのRNA誘導型ヌクレアーゼ、またはそれをコードするヌクレオチド配列を含んでいる。ガイドRNAは興味ある標的配列にハイブリダイズし、RGNポリペプチドと複合体も形成することにより、このRGNポリペプチドを標的配列に結合させる。これら実施態様のいくつかでは、RGNは、配列番号1、16、24、35、43、または50のアミノ酸配列、またはその活性なバリエーションまたは断片を含んでいる。さまざまな実施態様では、ガイドRNAは、配列番号2、17、25、36、44、51、または63のヌクレオチド配列を含むCRISPR反復配列、またはその活性なバリエーションまたは断片を含んでいる。特別な実施態様では、ガイドRNAは、配列番号3、18、26、37、45、52、または62のヌクレオチド配列を含むtracrRNA、またはその活性なバリエーションまたは断片を含んでいる。このシステムのガイドRNAとして、単一ガイドRNAまたは二重ガイドRNAが可能である。特別な実施態様では、このシステムは、ガイドRNAにとって異種のRNA誘導型ヌクレアーゼを含んでおり、RGNとガイドRNAが自然に複合体を形成することはない。

20

30

## 【0165】

本開示により、核酸配列を標的としていてその標的核酸配列を改変するシステムも提供される。RNA誘導型DNA結合ポリペプチド（RGNなど）とgRNAは、リボ核ポリペプチド複合体を興味ある核酸配列に向かわせる役割を担っており；デアミナーゼポリペプチドは、標的とされるその核酸配列を改変する役割を持つ。ガイドRNAは興味ある標的配列にハイブリダイズするとともに、RNA誘導型DNA結合ポリペプチドと複合体も形成し、そのことによってこのRNA誘導型DNA結合ポリペプチドを標的配列に結合させる。RNA誘導型DNA結合ポリペプチドは融合タンパク質の1つのドメインであり；他方のドメインは本明細書に記載されているデアミナーゼである。いくつかの実施態様では、RNA誘導型DNA結合ポリペプチドはRGN（Cas9など）である。さらに別の実施態様では、RNA誘導型DNA結合ポリペプチドは、配列番号568、569、547、553のアミノ酸配列、またはその活性なバリエーションまたは断片を含んでいる。RNA誘導型DNA結合ポリペプチドの別の例に含まれるのは、アメリカ合衆国特許出願公開第16/432,321号（その全体が参照によって本明細書に組み込まれている）に記載されているようなRGNである。いくつかの実施態様では、RNA誘導型DNA結合ポリペプチドは、II型 CRISPR-Casポリペプチド、またはその活性なバリエーションまたは断片である。いくつかの実施態様では、RNA誘導型DNA結合ポリペプチドは、V型 CRISPR-Casポリペプチド、またはその活性なバリエーションまたは断片である。別の実施態様では、RNA誘導型DNA結合ポリペプチドは、VI型 CRISPR-Casポリペプチド、またはその活性なバリエーションまたは断片である。別の実施態様では、融合タンパク質

40

50

のDNA結合ドメインは、RNAガイド（例えばZnフィンガーヌクレアーゼ、TALEN、またはメガヌクレアーゼポリペプチド）を必要とせず、それぞれのヌクレアーゼ活性は不活化されている。

【0166】

興味ある標的配列に結合させるための本明細書に提示されているシステムとしてリボ核タンパク質複合体が可能であり、このリボ核タンパク質複合体は、少なくとも1つのタンパク質に結合したRNAの少なくとも1個の分子である。いくつかの実施態様では、本明細書に提示されているリボ核タンパク質複合体は、RNA構成要素としての少なくとも1つのガイドRNAと、タンパク質構成要素としての1つのRNA誘導型ヌクレアーゼを含んでいる。このような核タンパク質複合体は、RGNポリペプチドを自然に発現するとともに興味ある標的配列に対して特異的な特定のガイドRNAを発現するように操作されている細胞または生物から精製することができる。別の実施態様では、本明細書に提示されている核タンパク質複合体は、少なくとも1つのガイドRNAをRNA構成要素として含むとともに、本発明のデアミナーゼとRNA誘導型DNA結合ポリペプチドを含む融合タンパク質をタンパク質構成要素として含んでいる。融合タンパク質の場合、または本発明のRGNの場合には、リボ核タンパク質複合体は、融合タンパク質（または本発明のRGNだけ）とガイドRNAをコードするポリヌクレオチドを用いて形質転換した後にそのRGNポリペプチドとガイドRNAの発現を可能にする条件下で培養した細胞または生物から精製することができる。このようにして、本発明のRGN、本発明のデアミナーゼ、本発明のRGNを含むリボ核タンパク質複合体、本発明の融合タンパク質、または融合タンパク質リボ核タンパク質複合体を作製する方法が提供される。このような方法は、本発明のポリペプチドをコードするヌクレオチド配列と、いくつかの実施態様ではガイドRNAをコードするヌクレオチド配列を含む細胞を、そのポリペプチド（と、いくつかの実施態様ではそのガイドRNA）が発現する条件下で培養することを含んでいる。その後、本発明のRGN、本発明のRGNを含むリボ核タンパク質複合体、本発明の融合タンパク質、または融合タンパク質リボ核タンパク質複合体を、培養された細胞のライセートから精製することができる。

【0167】

本発明のRGN、本発明のデアミナーゼ、本発明のRGNを含むリボ核タンパク質複合体、本発明の融合タンパク質、または融合タンパク質リボ核タンパク質複合体を生物サンプルのライセートから精製する方法は本分野で知られている（例えばサイズ排除クロマトグラフィおよび/またはアフィニティクロマトグラフィ、2D-PAGE、HPLC、逆相クロマトグラフィ、免疫沈降）。特別な方法では、RGNポリペプチドは組み換え産生されたものであり、精製を助けるための精製タグを含んでいる。精製タグの非限定的な例に含まれるのは、グルタチオン-S-トランスフェラーゼ（GST）、キチン結合タンパク質（CBP）、マルトース結合タンパク質、チオレドキシン（TRX）、ポリ(NANP)、タンデムアフィニティ精製（TAP）タグ、myc、AcV5、AU1、AU5、E、ECS、E2、FLAG、HA、nus、Softag 1、Softag 3、Strep、SBP、Glu-Glu、HSV、KT3、S、S1、T7、V5、VSV-G、6×His（配列番号593）、10×His（配列番号594）、ピオチンカルボキシル担体タンパク質（BCCP）、およびカルモジュリンである。一般に、本発明のタグ付きポリペプチドまたはリボ核タンパク質複合体は、固定化された金属アフィニティクロマトグラフィを利用して精製される。本分野で知られている他の同様の方法（他の形態のクロマトグラフィ、または例えば免疫沈降が含まれる）を、単独で、または組み合わせて使用できることがわかるであろう。

【0168】

「単離された」または「精製された」ポリペプチド、またはその生物活性な部分は、天然の環境で見いだされたポリペプチドに通常は付随する構成成分、またはそのポリペプチドと相互作用する構成成分を実質的または本質的に含んでいない。したがって単離または精製されたポリペプチドは、組み換え技術によって作製されるときには他の細胞材料または細胞培地を実質的に含んでおらず、化学合成されるときには化学的前駆体または他の化合物を実質的に含んでいない。細胞材料を実質的に含んでいないタンパク質には、汚染タ

10

20

30

40

50

ンパク質が（乾燥重量で）約30%、20%、10%、5%、または1%よりも少ないタンパク質調製物が含まれる。本発明のタンパク質、またはその生物活性のある部分が組み換えによって作製されるとき、最適には、培地は、化学的前駆体を、または興味あるタンパク質なしの化合物を、（乾燥重量で）約30%、20%、10%、5%、または1%よりも少なく含んでいる。

#### 【0169】

興味ある標的配列に結合させるための、および/または興味ある標的配列を切断するための本明細書に提示されている特別な方法は、インビトロで組み立てられたRGNリボ核タンパク質複合体の使用を含んでいる。RGNリボ核タンパク質複合体のインビトロでの組み立ては、本分野で知られている任意の方法を用いて実施することができ、その方法では、RGNポリペプチドがガイドRNAに結合できる条件下でRGNポリペプチドをガイドRNAに接触させる。本明細書で用いられている「接触」、「接触している」、「接触した」は、望む反応の構成要素をまとめ、望むその反応を実施するのに適した条件下に置くことを意味する。RGNポリペプチドは、インビトロでの翻訳、または化学合成を通じて生成したものを、生物サンプル、細胞ライセート、または培地から精製することができる。ガイドRNAは、インビトロでの転写、または化学合成を通じて生成したものを、生物サンプル、細胞ライセート、または培地から精製することができる。RGNポリペプチドとガイドRNAを溶液（例えば緩衝化生理食塩溶液）中で接触させてRGNリボ核タンパク質複合体をインビトロで組み立てることができる。

#### 【0170】

VIII. 標的配列に結合させる方法、標的配列を切断する方法、または標的配列を改変する方法

#### 【0171】

本開示により、興味ある標的ヌクレオチドに結合させる方法、興味ある標的ヌクレオチドを切断する方法、および/または興味ある標的ヌクレオチド改変する方法が提供される。いくつかの実施態様では、これらの方法は、少なくとも1つのガイドRNAまたはそれをコードするポリヌクレオチドと、少なくとも1つのRGNポリペプチドまたはそれをコードするポリヌクレオチドを含むシステムを、標的配列に、またはこの標的配列を含む細胞、オルガネラ、または胚に送達することを含んでいる。これらの実施態様のいくつかでは、RGNは、配列番号1、16、24、35、43、または50のアミノ酸配列、またはその活性なバリエーションまたは断片を含んでいる。さまざまな実施態様では、ガイドRNAは、配列番号2、17、25、36、44、51、または63のヌクレオチド配列を含むCRISPR反復配列、またはその活性なバリエーションまたは断片を含んでいる。特別な実施態様では、ガイドRNAは、配列番号3、18、26、37、45、52、または62のヌクレオチド配列を含むtracrRNA、またはその活性なバリエーションまたは断片を含んでいる。システムのガイドRNAとして、単一ガイドRNAまたは二重ガイドRNAが可能である。システムのRGNは、ヌクレアーゼ-デッドRGNであること、ニッカーゼ活性を持つこと、または融合ポリペプチドであることが可能である。いくつかの実施態様では、融合ポリペプチドは、塩基編集ポリペプチド（例えばシチジンデアミナーゼまたはアデノシンデアミナーゼ）を含んでいる。特別な実施態様では、RGNおよび/またはガイドRNAは、このRGNおよび/またはガイドRNA（またはこのRGNとガイドRNAの少なくとも一方をコードするポリヌクレオチド）が導入される細胞、オルガネラ、または胚にとって異種である。

#### 【0172】

別の実施態様では、これらの方法は、少なくとも1つのガイドRNAまたはそれをコードするポリヌクレオチドと、本発明のデアミナーゼとRNA誘導型DNA結合ポリペプチドを含む少なくとも1つの融合タンパク質またはそれをコードするポリヌクレオチドを含むシステムを、標的配列に、またはこの標的配列を含む細胞、オルガネラ、または胚に送達することを含んでいる。これら実施態様のいくつかでは、融合タンパク質は、配列番号374～545または572～584のアミノ酸配列のいずれか1つ、またはその活性なバリエーションまたは断片を含んでいる。

10

20

30

40

50

## 【 0 1 7 3 】

いくつかの実施態様では、これらの方法は、標的ポリヌクレオチドを本明細書に開示されているデアミナーゼに接触させることを含んでいる。いくつかの実施態様では、これらの方法は、標的ポリヌクレオチドを、デアミナーゼドメインとDNA結合ドメインを含む融合タンパク質に接触させることを含んでいる。これら実施態様のいくつかでは、これらの方法は、DNA分子を、(a)デアミナーゼドメインとRNA誘導型DNA結合ポリペプチドを含む融合タンパク質(例えばヌクレアーゼ不活性なRGNドメイン)と、(b)(a)の融合タンパク質をDNA鎖の標的ヌクレオチド配列に向かわせるgRNAに接触させることを含んでおり;その間にDNA分子は、ヌクレオチド塩基の脱アミノ化に有効な量かつ適した条件下で融合タンパク質とgRNAに接触する。デアミナーゼがこの方法で用いられるこれら実施態様のいくつかでは、デアミナーゼは、配列番号374、383、397、399、407、408、411、414、416、420、514、および572~584のいずれか1つのアミノ酸配列を持つか、そのバリエーションであり、このバリエーションは、配列番号374、383、397、399、407、408、411、414、416、420、514、および572~584のいずれか1つのアミノ酸配列と少なくとも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95%、少なくとも96%、少なくとも97%、少なくとも98%、または少なくとも99%一致するアミノ酸配列を持つ。これら実施態様のいくつかでは、バリエーションデアミナーゼポリペプチドは、配列番号572~584のいずれか1つとの配列一致があるレベルであり、特定のアミノ酸残基が親配列から変化していない。例えばいくつかの実施態様では、配列番号572の1つのバリエーションは、配列番号572の102位に対応する位置にリシンを含み、配列番号572の104位に対応する位置にチロシンを含み、配列番号572の106位に対応する位置にトレオニンを含んでいる。特別な実施態様では、配列番号574の1つのバリエーションは、配列番号574の101位に対応する位置にグルタミン酸を含み、配列番号574の103位に対応する位置にセリンを含み、配列番号574の105位に対応する位置にリシンを含んでいる。いくつかの実施態様では、配列番号575の1つのバリエーションは、配列番号575の101位に対応する位置にリシンを含み、配列番号575の103位に対応する位置にロイシンを含み、配列番号575の105位に対応する位置にグルタミン酸を含んでいる。いくつかの実施態様では、配列番号576の1つのバリエーションは、配列番号576の105位に対応する位置にアラニンを含み、配列番号576の107位に対応する位置にアルギニンを含んでいる。特別な実施態様では、配列番号577の1つのバリエーションは、配列番号577の102位に対応する位置にグリシンを含み、配列番号577の104位に対応する位置にセリンを含み、配列番号577の106位に対応する位置にアルギニンを含んでいる。いくつかの実施態様では、配列番号578の1つのバリエーションは、配列番号578の105位に対応する位置にセリンを含み、配列番号578の107位に対応する位置にトレオニンを含んでいる。いくつかの実施態様では、配列番号579の1つのバリエーションは、配列番号579の102位に対応する位置にセリンを含み、配列番号579の104位に対応する位置にグルタミンを含み、配列番号579の106位に対応する位置にグリシンを含んでいる。特別な実施態様では、配列番号580の1つのバリエーションは、配列番号580の111位に対応する位置にグリシンを含んでいる。いくつかの実施態様では、配列番号581の1つのバリエーションは、配列番号581の104位に対応する位置にグルタミンを含み、配列番号581の106位に対応する位置にグリシンを含み、配列番号581の108位に対応する位置にグルタミン酸を含んでいる。いくつかの実施態様では、配列番号582の1つのバリエーションは、配列番号582の102位に対応する位置にアルギニンを含み、配列番号582の104位に対応する位置にトリプトファンを含み、配列番号582の106位に対応する位置にグルタミン酸を含んでいる。いくつかの実施態様では、配列番号583の1つのバリエーションは、配列番号583の104位に対応する位置にアルギニンを含み、配列番号583の106位に対応する位置にセリンを含んでいる。特別な実施態様では、配列番号584の1つのバリエーションは、配列番号584の110位に対応する位置にフェニルアラニンを含み、配列番号584の112位に対応する位置にセリンを含み、配列番号584の114位に対応する位置にトレオニンを含んでいる。

10

20

30

40

50

## 【0174】

いくつかの実施態様では、これらの方法は、DNA分子を、(a)デアミナーゼドメインとRNA誘導型DNA結合ポリペプチドを含む融合タンパク質(例えばヌクレアーゼ不活性なRGNドメイン)と、(b)(a)の融合タンパク質をDNA鎖の標的ヌクレオチド配列に向かわせるgRNAに接触させることを含んでおり;その間にDNA分子は、ヌクレオチド塩基の脱アミノ化に有効な量かつ適した条件下で融合タンパク質とgRNAに接触する。いくつかの実施態様では、標的DNA配列は、疾患または障害に関係する配列を含んでおり、その中の問題のヌクレオチド塩基が脱アミノ化する結果として、疾患または障害に関係しない配列になる。いくつかの実施態様では、標的DNA配列は、作物植物の1つのアレルの中に存在し、興味ある形質に関するこのアレルが特別である結果として農業的な価値がより低い植物になる。問題のヌクレオチド塩基が脱アミノ化する結果として、その形質を改善するアレルになり、その植物の農業的価値が上昇する。

10

## 【0175】

いくつかの実施態様では、DNA配列は、疾患または障害に関係するT C点変異またはAG点変異を含んでおり、変異したC塩基またはG塩基の脱アミノ化により、疾患または障害に関係しない配列になる。いくつかの実施態様では、脱アミノ化により、配列内にあって疾患または障害に関係する変異が修正される。

## 【0176】

いくつかの実施態様では、疾患または障害に関係する配列はタンパク質をコードしており、このタンパク質では脱アミノ化によって疾患または障害に関係する配列に終止コドンが導入され、その結果として、コードされるタンパク質のトランケーションが起こる。いくつかの実施態様では、接触操作は、疾患または障害を持つ可能性が大きい対象、疾患または障害を持つ対象、または疾患または障害を持つと診断された対象の生体内で実施される。いくつかの実施態様では、疾患または障害は、ゲノムの中の点変異または単一塩基変異と関係した疾患である。いくつかの実施態様では、疾患は、遺伝子疾患、がん、代謝疾患、またはリソソーム蓄積症である。

20

## 【0177】

本開示の方法が、ガイドRNAおよび/またはRGNポリペプチドをコードするポリヌクレオチド、または本発明のデアミナーゼとRNA誘導型DNA結合ポリペプチドを含む融合タンパク質を送達することを含む実施態様では、細胞または胚をその後、ガイドRNAおよび/またはRGNポリペプチドが発現する条件下で培養することができる。さまざまな実施態様では、この方法は、標的配列をRGNリボ核タンパク質複合体に接触させることを含んでいる。RGNリボ核タンパク質複合体は、ヌクレアーゼ-デッドであるかニックアーゼ活性を持つRGNを含むことができる。いくつかの実施態様では、リボ核タンパク質複合体のRGNは、塩基編集ポリペプチド(例えば本明細書に開示されているデアミナーゼ)を含む融合ポリペプチドである。別の実施態様では、リボ核タンパク質複合体は、本発明のデアミナーゼとRNA誘導型DNA結合ポリペプチドを含む融合ポリペプチドを含んでいる。いくつかの実施態様では、この方法は、標的配列を含む細胞、オルガネラ、または胚の中にRGNリボ核タンパク質複合体を導入することを含んでいる。RGNリボ核タンパク質複合体として、生物サンプルから精製されたもの、組み換え産生の後に精製されたもの、または本明細書に記載されているようにしてインビトロで組み立てられたものが可能である。標的配列、または細胞オルガネラ、または胚に接触させるRGNリボ核タンパク質複合体がインビトロで組み立てられた実施態様では、この方法はさらに、この複合体をインビトロで組み立てた後に、標的配列、細胞、オルガネラ、または胚に接触させることを含むことができる。

30

40

## 【0178】

精製されるかインビトロで組み立てられたRGNリボ核タンパク質複合体は、本分野で知られている任意の方法(その非限定的な例に電気穿孔が含まれる)を用いて細胞、オルガネラ、または胚に導入することができる。あるいはRGNポリペプチド、本発明のデアミナーゼとRNA誘導型DNA結合ポリペプチドを含む融合ポリペプチド、および/またはガイドRNAをコードするか含んでいるポリヌクレオチドは、本分野で知られている任意の方法(

50



例えば電気穿孔)を用いて細胞、オルガネラ、または胚に導入することができる。

【0179】

ガイドRNAは、標的配列に、またはこの標的配列を含む細胞、オルガネラ、または胚に送達されるか接触すると、RGNポリペプチドまたは融合タンパク質を配列特異的なやり方で標的配列に結合させる。RGNがヌクレアーゼ活性を持つ実施態様では、RGNポリペプチドは、興味ある標的配列に結合するとこの標的配列を切断する。標的配列はその後、内在性修復機構(例えば非相同末端結合修復、または提供されるドナーポリヌクレオチドを用いた相同組み換え修復)を通じて改変される。

【0180】

標的配列に対するRNA誘導型DNA結合ポリペプチドの結合を測定する方法は本分野で知られており、方法には、クロマチン免疫沈降アッセイ、ゲル移動度シフトアッセイ、DNAプル-ダウンアッセイ、レポーターアッセイ、およびマイクロプレート捕獲・検出アッセイが含まれる。同様に、標的配列の切断または改変を測定する方法は本分野で知られており、方法には、インビトロまたは生体内での切断アッセイが含まれる。この切断アッセイでは、分解産物の検出を容易にするため標的配列に適切な標識(例えば放射性同位体、蛍光物質)を付着させて、または付着させずに、PCRシーケンシングまたはゲル電気泳動を利用して切断を確認する。あるいは切れ目トリガー式指数関数的増幅反応(NTENPAR)アッセイを利用することができる(例えばZhang他(2016年)Chem. Sci. 第7巻:4951~4957ページを参照されたい)。生体内切断は、Surveyorアッセイ(Guschin他(2010年)Methods Mol Biol第649巻:247~256ページ)を利用して評価することができる。

【0181】

いくつかの実施態様では、本発明の方法は、2つ以上のガイドRNAとの複合体にされた単一のタイプのRGNの使用を含んでいる。これら2つ以上のガイドRNAは、単一の遺伝子の異なる領域、または多数の遺伝子を標的とすることができる。別の実施態様では、これらの方法は、融合タンパク質の一部として、2つ以上のガイドRNAとの複合体になった単一のタイプのRNA結合DNA誘導型ドメインの使用を含んでいる。この多重ターゲティングによって融合タンパク質のデアミナーゼドメインが核酸を改変することが可能になり、そのことによって興味あるゲノムに多数の変異が導入される。

【0182】

ドナーポリヌクレオチドが提供されない実施態様では、RGNポリペプチドによって導入される二本鎖切断を非相同末端結合(NHEJ)修復プロセスによって修復することができる。NHEJはエラーしやすい性質があるため、二本鎖切断の修復によって標的配列が改変される可能性がある。本明細書で核酸分子に関して用いられている「改変」は、核酸分子のヌクレオチド配列の変化を意味し、1個以上のヌクレオチドの欠失、挿入、または置換、またはこれらの組み合わせが可能である。標的配列の改変により、変化したタンパク質産物が発現する可能性、またはコード配列が不活性化する可能性がある。

【0183】

ドナーポリヌクレオチドが存在する実施態様では、導入された二本鎖切断の修復中に、ドナーポリペプチドの中のドナー配列を標的ヌクレオチド配列に組み込むこと、または標的ヌクレオチド配列と交換することができるため、その結果として外来ドナー配列が導入される。そのためドナーポリヌクレオチドは、興味ある標的配列に導入されることが望ましいドナー配列を含んでいる。いくつかの実施態様では、ドナー配列は元の標的ヌクレオチド配列を変化させるため、この新たに組み込まれたドナー配列はRGNによって認識されず、RGNによって切断されることはない。ドナー配列の組み込みは、ドナーポリヌクレオチドの中に、標的ヌクレオチド配列に隣接した配列と実質的に配列が一致するフランキング配列を含めることによって増強することができ、相同組み換え修復プロセスが可能になる。RGNポリペプチドが二本鎖切断で粘着末端を導入する実施態様では、ドナーポリヌクレオチドは、適合したオーバーハングが隣接しているドナー配列を含むことができ、二本鎖切断が修復されている間に非相同組み換え修復プロセスにより、オーバーハングを含む切断された標的ヌクレオチド配列にドナー配列を直接連結することが可能になる。

10

20

30

40

50

## 【0184】

本発明の方法が、ニッカーゼである（すなわち二本鎖ポリヌクレオチドの単一の鎖を切断することだけができる）本発明のRGNの使用を含む実施態様では、この方法は、複数の同じ標的配列または重複している標的配列を標的としていてポリヌクレオチドの異なる鎖を切断する2つのRGNニッカーゼを導入することを含むことができる。例えば二本鎖ポリヌクレオチドのプラス（+）鎖だけを切断するRGNニッカーゼとともに、二本鎖ポリヌクレオチドのマイナス（-）鎖だけを切断する第2のRGNニッカーゼを導入することができる。同様に、いくつかの実施態様では、この方法は、本発明のデアミナーゼとRNA誘導型DNA結合ポリペプチド（例えばRGNなど）を含む融合ポリペプチドの使用を含んでおり、そのRGNはニッカーゼ（例えば配列番号569）である。

10

## 【0185】

さまざまな実施態様では、標的ヌクレオチド配列に結合させてその標的配列を検出する方法が提供され、この方法は、細胞、オルガネラ、または胚の中に少なくとも1つのガイドRNAまたはそれをコードするポリヌクレオチドと、少なくとも1つのRGNポリペプチドまたはそれをコードするポリヌクレオチドを導入し、そのガイドRNAおよび/またはRGNポリペプチド（コード配列を導入する場合；このRGNポリペプチドはヌクレアーゼ-デッドRGNであり、検出可能な標識をさらに含んでいる）を発現させることを含み、この方法はさらに、検出可能な標識を検出することを含んでいる。検出可能な標識は、融合タンパク質（例えば蛍光タンパク質）としてRGNに融合させることができる。あるいは検出可能な標識として、RGNポリペプチドとの複合体にされるか、RGNポリペプチドの中に組み込まれていて、視覚的に、または他の手段によって検出できる小分子が可能である。

20

## 【0186】

本明細書では、標的配列の発現、または標的配列の調節下にある興味ある遺伝子の発現を変化させる方法も提供される。これらの方法は、細胞、オルガネラ、または胚の中に、少なくとも1つのガイドRNAまたはそれをコードするポリヌクレオチドと、少なくとも1つのRGNポリペプチドまたはそれをコードするポリヌクレオチドを導入し、このガイドRNAおよび/またはRGNポリペプチド（コード配列を導入する場合；このRGNポリペプチドはヌクレアーゼ-デッドRGNである）を発現させることを含んでいる。これら実施態様のいくつかでは、ヌクレアーゼ-デッドRGNは、本明細書に記載されているような発現モジュレータドメイン（すなわちエピジェネティックな改変ドメイン、転写活性化ドメイン、または転写抑制ドメイン）を含む融合タンパク質である。いくつかの実施態様では、ヌクレアーゼ-デッドRGNは、本明細書に記載されているデアミナーゼを含む融合タンパク質である。

30

## 【0187】

本開示により、興味ある標的ヌクレオチド配列に結合させる方法、および/または興味ある標的ヌクレオチド配列を改変する方法も提供される。これらの方法は、少なくとも1つのガイドRNAまたはそれをコードするポリヌクレオチドを含むシステムを送達することを含み、少なくとも1つの融合ポリペプチドは、標的配列、またはこの標的配列を含む細胞、オルガネラ、または胚に向かう、本発明のRGNと塩基編集ポリペプチド（例えば本明細書に記載されているデアミナーゼ）またはこの融合ポリペプチドをコードするポリヌクレオチドを含む。

40

## 【0188】

当業者は、本開示の方法のいずれかを利用して単一の標的配列または複数の標的配列を標的にすることができることがわかるであろう。例えば方法は、単一のRGNポリペプチドを多数の異なるガイドRNAと組み合わせて使用することを含んでいて、これらのガイドRNAは、単一の遺伝子および/または複数の遺伝子の中の多数の異なる配列を標的にすることができる。本明細書には、多数の異なるガイドRNAを多数の異なるRGNポリペプチドと組み合わせて導入する方法も包含される。これらのガイドRNAとガイドRNA/RGNポリペプチドシステムは、単一の遺伝子および/または複数の遺伝子の中の多数の異なる配列を標的にすることができる。

## 【0189】

50

1つの側面では、本発明により、上記の方法と組成物に開示されているエレメントの任意の1つ以上を含むキットが提供される。いくつかの実施態様では、キットは、ベクター系と、キットを使用するための指示を含んでいる。いくつかの実施態様では、ベクター系は、(a) tracrメイト配列に機能可能に連結された第1の調節エレメントと、tracrメイト配列の上流にガイド配列を挿入するための1つ以上の挿入部位(ガイド配列は、発現すると、(1) 標的配列にハイブリダイズするガイド配列、および(2) tracr配列にハイブリダイズするtracrメイト配列との複合体になったCRISPR酵素を含むCRISPR複合体を真核細胞の中の標的配列に配列特異的に結合させる) ; および/または(b) 核局在化配列を含む前記CRISPR酵素をコードする酵素コード配列に機能可能に連結された第2の調節エレメントを含んでいる。

10

#### 【0190】

いくつかの実施態様では、キットは、RNA誘導型DNA結合ポリペプチド(RGNポリペプチド)などであり、その例は、ヌクレアーゼ不活性なCas9ドメインである)と、本発明のデアミナーゼと、場合によっては、そのCas9ドメインとそのデアミナーゼの間に位置するリンカーも含む融合タンパク質を含んでいる。それに加え、いくつかの実施態様では、キットは、適切な試薬、バッファ、および/または例えばインビトロまたは生体内でのDNA編集またはRNA編集で融合タンパク質を使用するための指示を含んでいる。エレメントは、個別に、または組み合わせで提供することができ、適切な任意の容器(バイアル、瓶、または試験管など)に入れて提供することができる。いくつかの実施態様では、キットは、核酸配列を狙い通りに編集するのに適したgRNAの設計と使用に関する指示を含んでいる。

20

#### 【0191】

いくつかの実施態様では、キットは、1つ以上の言語による指示を含んでいる。いくつかの実施態様では、キットは、本明細書に記載されている1つ以上のエレメントを利用する方法で使用するための1つ以上の試薬を含んでいる。試薬は、適切な任意の容器に入れて提供することができる。例えばキットは、1つ以上の反应用バッファまたは保管用バッファを提供することができる。試薬は、特定のアッセイで使用できる形態で提供すること、または使用前に1つ以上の他の構成要素を添加する必要がある形態(例えば濃縮液または凍結乾燥形態)で提供することができる。バッファとして任意のバッファが可能であり、その非限定的な例に含まれるのは、炭酸ナトリウムバッファ、炭酸水素ナトリウムバッファ、ホウ酸塩バッファ、Trisバッファ、MOPSバッファ、HEPESバッファ、およびこれらの組み合わせである。いくつかの実施態様では、バッファはアルカリ性である。いくつかの実施態様では、バッファは約7~約10のpHを持つ。

30

#### 【0192】

いくつかの実施態様では、キットは、ベクターの中に挿入するためのガイド配列に対応する1つ以上のオリゴヌクレオチドを含んでいるが、それは、そのガイド配列と調節エレメントを機能可能に連結するためである。いくつかの実施態様では、キットは、相同組み換え鋳型ポリヌクレオチドを含んでいる。1つの側面では、本発明により、CRISPRシステムの1つ以上のエレメントを用いるための方法が提供される。本発明のCRISPR複合体は、標的ポリヌクレオチドを改変するための有効な手段を提供する。本発明のCRISPR複合体は多彩な有用性を持ち、その中には、多くのタイプの細胞の中の標的ポリヌクレオチドの改変(例えば欠失、挿入、転位、不活化、活性化)が含まれる。そのため本発明のCRISPR複合体は、広範な用途を、例えば遺伝子療法、薬のスクリーニング、疾患の診断、および予後予測において持つ。代表的な1つのCRISPR複合体は、標的ポリヌクレオチド内の標的配列にハイブリダイズしたガイド配列との複合体になったCRISPR酵素を含んでいる。

40

#### 【0193】

##### VIII. 標的ポリヌクレオチド

#### 【0194】

1つの側面では、本発明により、真核細胞の中の標的ポリヌクレオチドを改変する方法が提供される。真核細胞は、生体内、生体外、またはインビトロに存在できる。いくつかの実施態様では、この方法は、ヒト、または非ヒト動物、または植物(微細藻類を含む)

50

からの1個の細胞、または細胞の集団をサンプリングし、その1個または複数個の細胞を改変することを含んでいる。培養は、任意の段階で生体外にて実施することができる。その1個または複数個の細胞は、非ヒト動物または植物（微細藻類を含む）に再導入することさえできる。

【0195】

植物栽培者は、望ましい品質（例えば収量、品質、一様性、丈夫さ、および病原体に対する抵抗性）を探すため、天然の変動性を利用して非常に有用な遺伝子を組み合わせる。これらの望ましい品質には、成長、日照時間の好み、温度条件、開花または生殖発生の開始日、脂肪酸含量、昆虫抵抗性、疾患抵抗性、線虫抵抗性、真菌抵抗性、除草剤抵抗性、さまざまな環境因子（旱魃、熱、湿度、寒さ、風、および不利な土壌状態（高塩分を含む））に対する寛容性も含まれる。有用なこれら遺伝子の供給源には、在来種または外来種、家宝種、野生植物の親戚、および誘導変異（例えば突然変異誘発剤で植物材料を処理する）が含まれる。植物栽培者は、本発明を利用することで、変異を誘導するための新たなツールを提供される。したがって当業者は、ゲノムを分析して有用な遺伝子を探し、望む特徴または形質を持つ品種の中で本発明を利用して有用な遺伝子の増加を以前の突然変異誘発剤よりも正確に誘導し、したがって植物栽培プログラムの加速と改良を実現することができる。

【0196】

RGNシステムの標的ポリヌクレオチドとして、真核細胞にとって内在性または外来性の任意のポリヌクレオチドが可能である。例えば標的ポリヌクレオチドとして、真核細胞の核の中に存在するポリヌクレオチドが可能である。標的ポリヌクレオチドとして、遺伝子産物（例えばタンパク質）をコードする配列、または非コード配列（例えば調節ポリヌクレオチドまたはジャンクDNA）が可能である。理論に囚われることは望まないが、標的配列は、PAM（プロトスペーサ隣接モチーフ）、すなわちCRISPR複合体によって認識される短い配列と関係づけられているべきであると考えられる。PAMの正確な配列と長さに関する条件は用いるCRISPR酵素によって異なるが、PAMは、典型的には、プロトスペーサ（すなわち標的配列）に隣接する2～5つの塩基対配列である。

【0197】

CRISPR複合体の標的ポリヌクレオチドは、疾患に関連する遺伝子とポリヌクレオチドのほか、生化学的シグナル伝達経路に関連する遺伝子とポリヌクレオチドを数多く含むことができる。標的ポリヌクレオチドの例に含まれるのは、生化学的シグナル伝達経路に関連する配列、例えば生化学的シグナル伝達経路に関連する遺伝子またはポリヌクレオチドである。標的ポリヌクレオチドの例に含まれるのは、疾患に関連する遺伝子またはポリヌクレオチドである。「疾患に関連する」遺伝子またはポリヌクレオチドは、疾患でない対照の組織または細胞と比べたとき、疾患に冒された組織に由来する細胞の中では異常なレベル、または異常な形態になった転写産物または翻訳産物を生じさせる任意の遺伝子またはポリヌクレオチドを意味する。それは、異常な高レベルで発現する遺伝子である可能性、異常な低レベルで発現する遺伝子である可能性があり、変化した発現は、疾患の発生および/または進行と相関している。疾患に関連した遺伝子は、疾患の起源（例えば原因変異）に直接の責任がある変異または遺伝的バリエーションを有する遺伝子、または疾患の起源に責任がある遺伝子と連鎖平衡にある変異または遺伝的バリエーションを有する遺伝子も意味する。転写または翻訳された産物は、既知であるか未知である可能性があり、さらに、正常なレベルまたは異常なレベルである可能性がある。疾患に関連する遺伝子とポリヌクレオチドの例は、McKusick-Nathans Institute of Genetic Medicine、Johns Hopkins University（バルチモア、メリーランド州）と、National Center for Biotechnology Information、National Library of Medicine（ベセスダ、メリーランド州）から、World Wide Web上で入手することができる。

【0198】

CRISPRシステムは、興味あるゲノム配列を標的とすることが比較的容易であるため特に有用だが、RGNが原因変異に対処するために何ができるかに関する問題が相変わらず残

10

20

30

40

50

されている。1つのアプローチは、RGN（好ましくはRGNの不活性なバリエーションまたはニック・ゼバリエーション）と、塩基編集酵素（例えばシチジンデアミナーゼ塩基エディタまたはアデノシンデアミナーゼ塩基エディタ）、または塩基編集酵素の活性ドメインの間の融合タンパク質を作製するというものである（アメリカ合衆国特許第9,840,699号；参照によって本明細書に組み込まれている）。いくつかの実施態様では、これらの方法は、DNA分子を（a）本発明のRGNと塩基編集ポリペプチド（デアミナーゼなど）を含む融合タンパク質と；（b）（a）の融合タンパク質をDNA鎖の標的ヌクレオチド配列に向かわせるgRNAに接触させることを含んでおり；その間にDNA分子は、ヌクレオチド塩基の脱アミノ化に有効な量かつ適した条件下で融合タンパク質とgRNAに接触する。いくつかの実施態様では、標的DNA配列は、疾患または障害に関係する配列を含んでおり、その中の問題のヌクレオチド塩基が脱アミノ化する結果として、疾患または障害に関係しない配列になる。いくつかの実施態様では、標的DNA配列は、作物植物の1つのアレルの中に存在し、興味ある形質に関するこのアレルが特別である結果として農業的な価値がより低い植物になる。問題のヌクレオチド塩基が脱アミノ化する結果として、その形質を改善するアレルになり、その植物の農業的な価値が上昇する。

#### 【0199】

いくつかの実施態様では、DNA配列は、疾患または障害に関係するT/C点変異またはA/G点変異を含んでおり、変異したC塩基またはG塩基の脱アミノ化により、疾患または障害に関係しない配列になる。いくつかの実施態様では、脱アミノ化により、配列内にあって疾患または障害に関係する変異が修正される。

#### 【0200】

いくつかの実施態様では、疾患または障害に関係する配列はタンパク質をコードしており、このタンパク質では脱アミノ化によって疾患または障害に関係する配列に終止コドンが導入され、その結果として、コードされるタンパク質のトランケーションが起こる。いくつかの実施態様では、接触操作は、疾患または障害を持つ可能性が大きい対象、疾患または障害を持つ対象、または疾患または障害を持つと診断された対象の生体内で実施される。いくつかの実施態様では、疾患または障害は、ゲノムの中の点変異または単一塩基変異と関係した疾患である。いくつかの実施態様では、疾患は、遺伝子疾患、がん、代謝疾患、またはリソソーム蓄積症である。

#### 【0201】

ある種の遺伝疾患を引き起こす遺伝子座、特に本発明のRGNまたはRGN-塩基エディタ融合タンパク質が容易に標的とすることのできる遺伝子座のさらに別の例は、実施例7と、それに対応する表8に見いだすことができる。

#### 【0202】

ハーラー症候群

#### 【0203】

本発明のRGN-塩基エディタ融合タンパク質に頼るアプローチを用いて修正できると考えられる遺伝子疾患の一例は、ハーラー症候群である。ハーラー症候群（MPS-1としても知られる）は、 $\alpha$ -L-イズロニダーゼ（IDUA）が欠損していることの結果であり、リソソームの中に硫酸デルマトンと硫酸ヘパランが蓄積することを分子レベルでの特徴とするリソソーム蓄積症になる。この疾患は一般に、 $\alpha$ -L-イズロニダーゼをコードするIDUA遺伝子の中の変異によって起こる遺伝性の遺伝子障害である。一般的なIDUA変異はW402XとQ70Xであり、両方ともナンセンス変異であるため翻訳が途中で停止する。このような変異は、精密ゲノム編集（PGE）アプローチによってうまく対処される。なぜなら例えば塩基編集アプローチによって1個のヌクレオチドを元に戻すと野生型コード配列が復元され、遺伝子座の内在性調節機構によってタンパク質の発現が制御されるからである。それに加え、ヘテロ接合体は無症状であることが知られているため、これら変異の1つを標的とするPGE療法は、この疾患を持つ患者の多くにとって、修正する必要があるのは変異したアレルのうちの1つだけであるという理由で有用であると考えられる（Bunge他（1994年）Hum. Mol. Genet. 第3巻(6)：861～866ページ；参照によって本明細書に組み込まれて

10

20

30

40

50

いる)。

#### 【0204】

ハーラー症候群に関する現在の治療法には酵素置換療法と骨髄移植が含まれる(Vellodi他(1997年) Arch. Dis. Child. 第76巻(2): 92~99ページ; Peters他(1998年) Blood第91巻(7): 2601~2608ページ; 参照によって本明細書に組み込まれている)。酵素置換療法は、ハーラー症候群の患者の生存と生活の質に対して劇的な効果があったが、このアプローチは、費用と時間がかかる毎週の輸液を必要とする。追加のアプローチに含まれるのは、発現ベクターに載せたIDUA遺伝子の送達、または高発現している遺伝子座(血清アルブミンの遺伝子座など)へのこの遺伝子の挿入である(アメリカ合衆国特許第9,956,247号; 参照によって本明細書に組み込まれている)。しかしこれらのアプローチでは、元のIDUA遺伝子座が復元されて正しいコード配列になることはない。ゲノム編集戦略には多数の利点があると考えられ、その中で非常に注目すべきなのは、遺伝子の発現調節が健康な個人に存在する自然の機構によって制御されると考えられることである。それに加え、塩基編集を利用すると、大規模な染色体再配列、細胞死、または腫瘍抑制機構の破壊によるがん発生につながる可能性がある二本鎖DNAの切断を引き起こす必要がない。この疾患の原因変異を修正する方法を可能にする記述は実施例8に提示されている。記載されている方法は、本発明のRGN-塩基エディタ融合タンパク質を用いる一般的な戦略の一例であり、この融合タンパク質をヒトゲノムの中であって疾患を引き起こすいくつかの変異に向かわせてそれらの変異を修正する。表8に記載されているような疾患を標的とする同様のアプローチも追究できることがわかるであろう。さらに、他の種(特に一般的なペットまたは家畜)で疾患を引き起こす変異を標的とする同様のアプローチも本発明のRGNを用いて展開できることがさらにわかるであろう。一般的なペットと家畜に含まれるのは、イヌ、ネコ、ウマ、ブタ、ウシ、ヒツジ、ニワトリ、ロバ、ヘビ、フェレット、魚類(サケを含む)、およびエビである。

#### 【0205】

フリードライヒ運動失調症

#### 【0206】

本発明のRGNは、原因変異がより複雑なヒトの治療法においても有用である可能性がある。例えばフリードライヒ運動失調症やハンチントン病などのいくつかの疾患は、遺伝子の特定の領域にある3ヌクレオチドモチーフの反復回数が顕著に増加している結果であり、そのことが、発現するタンパク質が機能する能力、または発現する能力に影響を与える。フリードライヒ運動失調症(FRDA)は常染色体劣性疾患であり、脊椎内の神経組織が徐々に変性していく。ミトコンドリア内のフラタキシン(FXN)タンパク質のレベルが低下が、細胞レベルでの酸化的損傷と鉄欠損を引き起こす。低下したFXN発現は、体細胞と生殖系のFXN遺伝子のイントロン1の中にあるGAAトリプレット伸長と結び付けられてきた。FRDA患者では、GAA反復は、70を超えるトリプレットからなることがしばしばあり、ときに1000を超えるトリプレットからなることさえある(非常に一般的なのは600~900)のに対し、この病気でない個人は約40反復以下である(Pandolfo他(2012年) Handbook of Clinical Neurology第103巻: 275~294ページ; Campuzano他(1996年) Science第271巻: 1423~1427ページ; Pandolfo(2002年) Adv. Exp. Med. Biol. 第516巻: 99~118ページ; これらはすべて、参照によって本明細書に組み込まれている)。

#### 【0207】

フリードライヒ運動失調症(FRDA)を引き起こす3ヌクレオチド反復配列の増殖は、FRDA不安定性領域と呼ばれるFXN遺伝子内の特定の遺伝子座の中で起こる。RNA誘導型ヌクレアーゼ(RGN)を用いてFRDA患者の細胞内の不安定性領域を切除することができる。このアプローチは、1) ヒトゲノムの中のアレルを標的とするようにプログラムすることのできるRGNとガイドRNAの配列と; 2) このRGNとガイドの配列を送達する方法を必要とする。ゲノム編集に用いられる多くのヌクレアーゼ(化膿レンサ球菌からの一般に用いられているCas9ヌクレアーゼ(SpCas9)など)は、アデノ随伴ウイルス(AAV)ベク

10

20

30

40

50

ターの中にパッケージするには大きすぎる。特にSpCas9遺伝子とガイドRNAの長さに加え、機能的な発現カセットに必要とされる他の遺伝エレメントの長さを考慮すると、そのことがわかる。そのためSpCas9を利用するアプローチはより困難になる。

#### 【0208】

本発明のコンパクトなRNA誘導型ヌクレアーゼは、FRDA不安定性領域の切除に極めてよく適している。それぞれのRGNは、FRDA不安定性領域の近傍にPAMを必要とする。それに加え、これらRGNのそれぞれは、ガイドRNAとともにAAVベクターの中にパッケージすることができる。2つのガイドRNAをパッケージするには第2のベクターが必要となる可能性はあるが、それでもこのアプローチは、より大きなヌクレアーゼ（例えばSpCas9であり、このタンパク質の配列を2つのベクターに分割せねばならない可能性がある）で必要とされると考えられるベクターよりも有利である。この疾患の原因変異を修正する方法を可能にする記述は実施例9に提示されている。記載されている方法には、本発明のRGNを用いてゲノム不安定性の領域を除去する戦略が含まれる。このような戦略は、同様の遺伝的基礎を持つ他の疾患と障害（ハンチントン病など）に適用することができる。本発明のRGNを用いる同様の戦略は、農業的に、または経済的に重要な非ヒト動物の疾患と障害にも適用することができる。そのような非ヒト動物に含まれるのは、イヌ、ネコ、ウマ、ブタ、ウシ、ヒツジ、ニワトリ、ロバ、ヘビ、フェレット、魚類（サケを含む）、およびエビである。ヘモグロビン異常症

10

#### 【0209】

本発明のRGNは、有利な効果をもたらす可能性がある破壊的変異を導入することもできると考えられる。ヘモグロビン、特にベータグロビン鎖（HBB遺伝子）をコードする遺伝子の中の遺伝子欠陥が、ヘモグロビン異常症として知られる多数の疾患（鎌形赤血球貧血とサラセミアが含まれる）の原因である可能性がある。

20

#### 【0210】

成人では、ヘモグロビンは、2つのアルファ（ $\alpha$ ）様グロビン鎖と、2つのベータ（ $\beta$ ）様グロビン鎖と、4つのヘム基を含むヘテロ四量体である。成人では、この $\alpha_2\beta_2$ 四量体は、ヘモグロビンA（HbA）、または成人ヘモグロビンと呼ばれている。典型的には、アルファグロビン鎖とベータグロビン鎖は約1:1の比で合成され、この比が、ヘモグロビンと赤血球細胞（RBC）の安定性にとって極めて重要であるように見える。発達中の胎児では、ヘモグロビンAよりも酸素への結合親和性が大きい異なる形態のヘモグロビン（胎児ヘモグロビン（HbF））が産生されるため、酸素を母体の血流を通じて胎児のシステムに送達することができる。胎児のヘモグロビンは2つの $\alpha$ グロビン鎖も含有しているが、成人の $\alpha$ グロビン鎖の代わりに2つの胎児ガンマ（ $\gamma$ ）グロビン鎖を有する（すなわち胎児ヘモグロビンは $\alpha_2\gamma_2$ である）。ガンマ-グロビンの産生からベータ-グロビンの産生への切り換えの調節は極めて複雑であり、主に、ガンマグロビン転写の下方調節と、それと同時に起こるベータグロビン転写の上方調節を含んでいる。妊娠約30週の時点では、胎児の体内におけるガンマグロビンの合成は低下し始めるのに対し、ベータグロビンの産生は増加する。新生児のヘモグロビンは、年齢が約10ヶ月になるまではほぼすべてが $\alpha_2\beta_2$ だが、いくつかのHbFが成人期まで残る（全ヘモグロビンの約1~3%）。上に説明したように、ヘモグロビン異常症を持つ患者の大半は、ガンマグロビンをコードする遺伝子が存在したままだが、上に説明したような分娩の頃に起こる正常遺伝子の抑制が原因で発現は比較的少ない。

30

40

#### 【0211】

鎌状赤血球症は、 $\beta$ グロビン遺伝子（HBB）内のV6E変異によって起こり（DNAレベルでGAGからGTGへ）、得られるヘモグロビンは「ヘモグロビンS」または「HbS」と呼ばれる。酸素がより少ない条件では、HbS分子が凝集して繊維状の沈殿物を形成する。これら凝集体がRBCの異常すなわち「鎌状化」を引き起こし、その結果として細胞の可撓性が失われる。この鎌状RBCはもはや毛細血管床に入り込むことができないため、鎌形赤血球患者では血管閉塞危機になる可能性がある。それに加え、鎌状RBCは正常なRBCよりも脆弱で溶血しやすく、最終的に患者は貧血になる。

50

## 【0212】

鎌状赤血球症患者の治療と管理は、抗生剤治療、疼痛管理、および急性期の輸血を含む一生にわたる課題である。1つのアプローチは、ヒドロキシウレアを使用するというものである。ヒドロキシウレアは、ガンマグロビンの産生を増加させることによって効果の一部を及ぼす。慢性ヒドロキシウレア療法の長期にわたる副作用はまだ知られていないが、治療によって望まない副作用が生じることで、患者ごとに効果が異なる可能性がある。鎌状赤血球治療の効果は増大しているにもかかわらず、患者の平均寿命はまだ50代の半ばから後半にすぎず、この疾患に関係する病的状態が患者の生活の質に深刻な影響を及ぼす。

## 【0213】

サラセミア（アルファサラセミアとベータサラセミア）もヘモグロビンに關係する疾患であり、典型的にはグロビン鎖の発現低下が關係する。これは、遺伝子の調節領域における変異を通じて発生するか、グロビンコード配列の中の変異によって機能するグロビントタンパク質の発現またはレベルが低下することから発生する。サラセミアの治療には、通常、輸血と鉄キレート化療法が含まれる。適切なドナーを特定できた場合には、重症患者の治療に骨髓移植も利用されているが、この方法はリスクが大きい可能性がある。

10

## 【0214】

SCDとベータサラセミアの両方を治療するために提案されてきた1つのアプローチは、ガンマグロビンの発現を増加させ、異常な成人ヘモグロビンAをHbFで機能的に置き換えるというものである。上述のように、ヒドロキシウレアを用いたSCD患者の治療は成功すると考えられるが、その理由の一部は、ヒドロキシウレアがガンマグロビンの発現増加に効果があることにある（DeSimone（1982年）Proc Nat'l Acad Sci USA第79巻（14）：4428～4431ページ；Ley他（1982年）N. Engl. J. Medicine、第307巻：1469～1475ページ；Ley他（1983年）Blood第62巻：370～380ページ；Constantoulakis他（1988年）Blood第72巻（6）：1961～1967ページ；これらのすべてが参照によって本明細書に組み込まれている）。HbFの発現増加は、ガンマグロビンの発現調節において役割を果たす産物を産生する遺伝子の同定と關係する。そのような1つの遺伝子はBCL11Aである。BCL11Aは、成人の赤血球前駆細胞で発現しているジンクフィンガータンパク質をコードしており、その発現が下方調節されるとガンマグロビンの発現が増加する（Sankaran他（2008年）Science 第322巻：1839ページ；参照によって本明細書に組み込まれている）。BCL11A遺伝子を標的とする抑制性RNAの使用が提案されている（例えばアメリカ合衆国特許出願公開第2011/0182867号；参照によって本明細書に組み込まれている）が、この技術には潜在的な欠点がある。そのような欠点に含まれるのは、完全なノックダウンを実現できない可能性があること、このようなRNAの送達に問題がある可能性のあること、およびこのRNAが継続的に存在せねばならないため、生涯にわたって多数回の治療が必要とされることである。

20

30

## 【0215】

本発明のRGNを用いてBCL11Aエンハンサ領域を標的としてBCL11Aの発現を妨害することにより、ガンマグロビンの発現を増加させることができる。この狙いを定めた妨害は、非相同末端結合（NHEJ）によって実現することができる。NHEJによって本発明のRGNはBCL11Aエンハンサ領域内の特定の配列を標的として二本鎖を切断し、細胞の機構がその切断を修復し、典型的にはそれと同時に有害な変異を導入する。別の疾患標的について記載したことと同様、本発明のRGNは、相対的にサイズが小さいために生体内送達を目的としてRGNとそれに対応するガイドRNAのための発現カセットを単一のAAVベクターの中にパッケージできることが理由で、他の既知のRGNと比べて利点を有する。この方法を可能にする記述は実施例10に提示されている。本発明のRGNを用いる同様の戦略も、ヒトと、農業的または経済的に重要な非ヒト動物の両方における同様の疾患と障害に適用することができる。

40

## 【0216】

IX. ポリヌクレオチド遺伝子改変を含む細胞

## 【0217】

50



本明細書では、本明細書に記載されているRGN、crRNA、tracrRNA、および/またはデアミナーゼを媒介とする方法を用いて改変された興味ある標的配列を含む細胞と生物が提供される。これら実施態様のいくつかでは、RGNは、配列番号1、16、24、35、43、または50のアミノ酸配列、またはその活性なバリエーションまたは断片を含んでいる。さまざまな実施態様では、ガイドRNAは、配列番号2、17、25、36、44、51、または63のヌクレオチド配列を含むCRISPR反復配列、またはその活性なバリエーションまたは断片を含んでいる。特別な実施態様では、ガイドRNAは、配列番号3、18、26、37、45、52、または62のヌクレオチド配列を含むtracrRNA、またはその活性なバリエーションまたは断片を含んでいる。システムのガイドRNAとして、単一ガイドRNAまたは二重ガイドRNAが可能である。いくつかの実施態様では、デアミナーゼは、配列番号374～545および572～584のいずれか1つのアミノ酸配列、またはその活性なバリエーションまたは断片を含んでいる。

10

**【0218】**

改変された細胞として、真核細胞（例えば哺乳動物の細胞、植物細胞、昆虫細胞）または原核細胞が可能である。本明細書に記載されているRGN、crRNA、および/またはtracrRNAを用いる方法によって改変された少なくとも1つのヌクレオチド配列を含むオルガネラと胚も提供される。遺伝子が改変された細胞、生物、オルガネラ、および胚は、改変されたヌクレオチド配列にとってヘテロ接合性またはホモ接合性であることが可能である。

**【0219】**

細胞、生物、オルガネラ、または胚の染色体が改変される結果として、変化したタンパク質産物または組み込まれた配列の発現の変化（上方調節または下方調節）、不活化、または発現が起こる可能性がある。染色体の改変によって遺伝子の不活化または非機能性タンパク質産物の発現が起こる場合には、遺伝子が改変された細胞、生物、オルガネラ、または胚は「ノックアウト」と呼ばれる。ノックアウト表現型の結果として、欠失変異（すなわち少なくとも1個のヌクレオチドの欠失）、挿入変異（すなわち少なくとも1個のヌクレオチドの挿入）、またはナンセンス変異（すなわち少なくとも1個のヌクレオチドが置換されて終止コドンが導入される）になる可能性がある。

20

**【0220】**

あるいは細胞、生物、オルガネラ、または胚の染色体が改変される結果として、「ノックイン」が生成する可能性がある。これは、あるタンパク質をコードするヌクレオチド配列が染色体に組み込まれることの結果である。これら実施態様のいくつかでは、コード配列が染色体に組み込まれることで野生型タンパク質をコードする染色体配列は不活化するが、外部から導入されたタンパク質は発現する。

30

**【0221】**

別の実施態様では、染色体が改変される結果として、バリエーションタンパク質産物が生成する。発現するバリエーションタンパク質産物は、少なくとも1個のアミノ酸の置換、および/または少なくとも1個のアミノ酸の付加または欠失を有する可能性がある。変化した染色体配列によってコードされるバリエーションタンパク質産物は、野生型タンパク質と比べて改変された特徴または活性（その非限定的な例に、変化した酵素活性または基質特異性が含まれる）を示す可能性がある。

**【0222】**

さらに別の実施態様では、染色体が改変される結果として、タンパク質の発現パターンが変化する可能性がある。非限定的な一例として、タンパク質産物の発現を制御する調節領域内の染色体が変化する結果として、そのタンパク質産物の過剰発現または下方調節、または変化した組織発現パターン、または変化した時間的発現パターンが生じる可能性がある。

40

**【0223】**

冠詞「1つの」は、本明細書では、文法的対象となる1つまたは2つ以上（すなわち少なくとも1つ）の物体を意味するのに用いられる。例として「1つのポリペプチド」は、1つ以上のポリペプチドを意味する。

**【0224】**

50

本明細書で言及されているあらゆる刊行物と特許出願は、本開示が関係する分野の当業者のレベルを示している。すべての刊行物と特許出願が、個々の刊行物または特許出願が具体的かつ個別に参照によって組み込まれていることが示されているかのようにして、参照によって本明細書に組み込まれている。

【0225】

明確な理解を目的として説明と例によって本発明をいくらか詳細に記載してきたが、添付の実施態様の範囲内でいくつかの変更と改変を実施できることは明らかであろう。

【0226】

非限定的な実施態様には、以下のものが含まれる。

【0227】

1. RNA誘導型ヌクレアーゼ (RGN) ポリペプチドをコードするポリヌクレオチドを含む核酸分子であって、前記ポリヌクレオチドが、配列番号1、16、24、35、43、または50と配列が少なくとも95%配列が一致するアミノ酸配列を含むRGNポリペプチドをコードするヌクレオチド配列を含み；

前記RGNポリペプチドが、標的DNA配列にハイブリダイズすることのできるガイドRNA (gRNA) に結合するとき、RNA誘導型配列に特異的なやり方で前記標的DNA配列に結合し；

RGNポリペプチドをコードする前記ポリヌクレオチドが、前記ポリヌクレオチドにとって異種のプロモータに機能可能に連結されている核酸分子。

【0228】

2. 前記RGNポリペプチドが前記標的DNA配列に結合したときに前記標的DNA配列を切断することができる、実施態様1に記載の核酸分子。

【0229】

3. 前記RGNポリペプチドによる切断が二本鎖の切断を生じさせる、実施態様2に記載の核酸分子。

【0230】

4. 前記RGNポリペプチドによる切断が一本鎖の切断を生じさせる、実施態様2に記載の核酸分子。

【0231】

5. 前記RGNポリペプチドがヌクレアーゼドッドであるか、ニッカーゼとして機能する、実施態様1に記載の核酸分子。

【0232】

6. 前記RGNポリペプチドが塩基編集ポリペプチドに機能可能に融合されている、実施態様5に記載の核酸分子。

【0233】

7. 前記塩基編集ポリペプチドがデアミナーゼである、実施態様6に記載の核酸分子。

【0234】

8. 前記デアミナーゼが、配列番号374、383、397、399、407、408、411、414、416、420、514、および572~584のいずれか1つと少なくとも90%配列が一致する、実施態様7に記載の核酸分子。

【0235】

9. 前記RGNポリペプチドが1つ以上の核局在化シグナルを含む、実施態様1~8のいずれか1つに記載の核酸分子。

【0236】

10. 前記RGNポリペプチドが、真核細胞での発現のためコドン最適化されている、実施態様1~9のいずれか1つに記載の核酸分子。

【0237】

11. 前記標的DNA配列が、プロトスペーサ隣接モチーフ (PAM) に隣接した位置にある、実施態様1~10のいずれか1つに記載の核酸分子。

【0238】

10

20

30

40

50

12. 実施態様1～11のいずれか1つに記載の核酸分子を含むベクター。

【0239】

13. 前記標的DNA配列にハイブリダイズすることのできる前記gRNAをコードする少なくとも1つのヌクレオチド配列をさらに含む、実施態様12に記載のベクター。

【0240】

14. 前記gRNAが単一ガイドRNAである、実施態様13に記載のベクター。

【0241】

15. 前記gRNAが二重ガイドRNAである、実施態様13に記載のベクター。

【0242】

16. 前記ガイドRNAが、配列番号2、17、25、36、44、51、または63と少なくとも95%配列が一致するCRISPR反復配列を含むCRISPR RNAを含む、実施態様13～15のいずれか1つに記載のベクター。

10

【0243】

17. 前記ガイドRNAが、配列番号3、18、26、37、45、52、または62と少なくとも95%配列が一致するtracrRNAを含む、実施態様13～16のいずれか1つに記載のベクター。

【0244】

18. 実施態様1～11のいずれか1つに記載の核酸分子または実施態様12～17のいずれか1つに記載のベクターを含む細胞。

【0245】

19. RGNポリペプチドを作製する方法であって、前記RGNポリペプチドが発現する条件下で実施態様18に記載の細胞を培養することを含む方法。

20

【0246】

20. RGNポリペプチドを作製する方法であって、配列番号1、16、24、35、43、または50と少なくとも95%配列が一致するアミノ酸配列を含むRNA誘導型ヌクレアーゼ(RGN)ポリペプチドをコードするヌクレオチド配列を含む異種核酸分子を細胞の中に導入して、

前記RGNポリペプチドを、標的DNA配列にハイブリダイズすることのできるガイドRNA(gRNA)に結合するときにRNA誘導型配列に特異的なやり方で前記標的DNA配列に結合させ、

30

前記RGNポリペプチドが発現する条件下で前記細胞を培養することを含む方法。

【0247】

21. 前記RGNポリペプチドを精製することをさらに含む、実施態様19または20に記載の方法。

【0248】

22. 前記細胞が、RGNリボ核タンパク質複合体を形成するため前記RGNポリペプチドに結合する1つ以上のガイドRNAをさらに発現する、実施態様19または20に記載の方法。

【0249】

23. 前記RGNリボ核タンパク質複合体を精製することをさらに含む、実施態様22に記載の方法。

40

【0250】

24. CRISPR RNA(crRNA)をコードするポリヌクレオチドを含む核酸分子であって、前記crRNAがスペーサ配列とCRISPR反復配列を含み、前記CRISPR反復配列が、配列番号2、17、25、36、44、51、または63少なくとも95%一致する配列を持つ核酸配列を含み；

a) 前記crRNAと；

b) 前記crRNAの前記CRISPR反復配列にハイブリダイズしたトランス活性化CRISPR RNA(tracrRNA)

を含むガイドRNAが、RNA誘導型ヌクレアーゼ(RGN)ポリペプチドに結合するとき、前記crRNAの前記スペーサ配列を通じて配列特異的なやり方で標的DNA配列にハイブリダイ

50

ズすることができ；

crRNAをコードする前記ポリペプチドが、前記ポリヌクレオチドにとって異種のプロモータに機能可能に連結されている核酸分子。

【0251】

25．実施態様24に記載の核酸分子を含むベクター。

【0252】

26．前記tracrRNAをコードするポリヌクレオチドをさらに含む、実施態様25に記載のベクター。

【0253】

27．前記tracrRNAが、配列番号3、18、26、37、45、52、または62と少なくとも95%配列が一致するヌクレオチド配列を含む、実施態様26に記載のベクター。

10

【0254】

28．前記crRNAをコードする前記ポリヌクレオチドと前記tracrRNAをコードする前記ポリヌクレオチドが同じプロモータに機能可能に連結されていて、単一ガイドRNAとしてコードされている、実施態様26または27に記載のベクター。

【0255】

29．前記crRNAをコードする前記ポリヌクレオチドと前記tracrRNAをコードする前記ポリヌクレオチドが、別々のプロモータに機能可能に連結されている、実施態様26または27に記載のベクター。

【0256】

20

30．実施態様25～29のいずれか1つに記載のベクターにおいて、前記ベクターが、前記RGNポリペプチドをコードするポリヌクレオチドをさらに含み、前記RGNポリペプチドが、配列番号1、16、24、35、43、または50と少なくとも95%配列が一致するアミノ酸配列を含む、ベクター。

【0257】

31．配列番号3、18、26、37、45、52、または62と少なくとも95%配列が一致するヌクレオチド配列を含むトランス活性化CRISPR RNA (tracrRNA) をコードするポリヌクレオチドを含む核酸分子であって、

a) 前記tracrRNAと；

b) スペース配列とCRISPR反復配列を含むcrRNA (ただし前記tracrRNAが前記crRNAの前記CRISPR反復配列にハイブリダイズする) を含むガイドRNAが、RNA誘導型ヌクレアーゼ (RGN) ポリペプチドに結合するとき、前記crRNAの前記スペーサ配列を通じて配列特異的なやり方で標的DNA配列にハイブリダイズすることができ；

30

tracrRNAをコードする前記ポリヌクレオチドが、前記ポリヌクレオチドにとって異種のプロモータに機能可能に連結されている核酸分子。

【0258】

32．実施態様31に記載の核酸分子を含むベクター。

【0259】

33．前記crRNAをコードするポリヌクレオチドをさらに含む、実施態様32に記載のベクター。

40

【0260】

34．前記crRNAの前記CRISPR反復配列が、配列番号2、17、25、36、44、51、または63と少なくとも95%配列が一致するヌクレオチド配列を含む、実施態様33に記載のベクター。

【0261】

35．前記crRNAをコードする前記ポリヌクレオチドと前記tracrRNAをコードする前記ポリヌクレオチドが同じプロモータに機能可能に連結されていて、単一ガイドRNAとしてコードされている、実施態様33または34に記載のベクター。

【0262】

50

36．前記crRNAをコードする前記ポリヌクレオチドと前記tracrRNAをコードする前記ポリヌクレオチドが、別々のプロモータに機能可能に連結されている、実施態様33または34に記載のベクター。

【0263】

37．実施態様32～36のいずれか1つに記載のベクターにおいて、前記ベクターが、前記RGNポリペプチドをコードするポリヌクレオチドをさらに含み、前記RGNポリペプチドが、配列番号1、16、24、35、43、または50と少なくとも95%配列が一致するアミノ酸配列を含む、ベクター。

【0264】

38．標的DNA配列に結合させるためのシステムであって、

10

a) 標的DNA配列にハイブリダイズすることのできる1つ以上のガイドRNA、または前記1つ以上のガイドRNA (gRNA) をコードする1つ以上のヌクレオチド配列と；

b) 配列番号1、16、24、35、43、または50と少なくとも95%配列が一致するアミノ酸配列を含むRNA誘導型ヌクレアーゼ (RGN) ポリペプチド、または前記RGNポリペプチドをコードするヌクレオチド配列を含み；

前記1つ以上のガイドRNAをコードする前記ヌクレオチド配列と、前記RGNポリペプチドをコードする前記ヌクレオチド配列が、それぞれ、前記ヌクレオチド配列のそれぞれにとって異種のプロモータに機能可能に連結され；

前記1つ以上のガイドRNAが前記標的DNA配列にハイブリダイズし、

前記1つ以上のガイドRNAが前記RGNポリペプチドと複合体を形成し、そのことによって前記RGNポリペプチドを前記標的DNA配列に結合させるシステム。

20

【0265】

39．前記gRNAが単一ガイドRNA (sgRNA) である、実施態様38に記載のシステム。

【0266】

40．前記gRNAが二重ガイドRNAである、実施態様38に記載のシステム。

【0267】

41．前記gRNAが、配列番号2、17、25、36、44、51、または63と少なくとも95%配列が一致するヌクレオチド配列を含むCRISPR反復配列を含む、実施態様38～40のいずれか1つに記載のシステム。

【0268】

30

42．前記gRNAが、配列番号3、18、26、37、45、52、または62と少なくとも95%配列が一致するヌクレオチド配列を含むtracrRNAを含む、実施態様38～41のいずれか1つに記載のシステム。

【0269】

43．前記標的DNA配列が、プロトスペーサ隣接モチーフ (PAM) に隣接した位置にある、実施態様38～42のいずれか1つに記載のシステム。

【0270】

44．前記標的DNA配列が細胞の中にある、実施態様38～43のいずれか1つに記載のシステム。

【0271】

40

45．前記細胞が真核細胞である、実施態様44に記載のシステム。

【0272】

46．前記真核細胞が植物細胞である、実施態様45に記載のシステム。

【0273】

47．前記真核細胞が哺乳動物の細胞である、実施態様45に記載のシステム。

【0274】

48．前記真核細胞が昆虫細胞である、実施態様45に記載のシステム。

【0275】

49．前記細胞が原核細胞である、実施態様44に記載のシステム。

【0276】

50

50．転写されるときに前記1つ以上のガイドRNAが前記標的DNA配列にハイブリダイズし、前記ガイドRNAが、前記標的DNA配列の切断を引き起こす前記RGNポリペプチドと複合体を形成する、実施態様38～49のいずれか1つに記載のシステム。

【0277】

51．前記切断が二本鎖の切断を生じさせる、実施態様50に記載のシステム。

【0278】

52．前記RGNポリペプチドによる切断が一本鎖の切断を生じさせる、実施態様50に記載のシステム。

【0279】

53．前記RGNポリペプチドがヌクレアーゼドドであるか、ニッカーゼとして機能する、実施態様38～49のいずれか1つに記載のシステム。

10

【0280】

54．前記RGNポリペプチドが塩基編集ポリペプチドに機能可能に融合されている、実施態様53に記載のシステム。

【0281】

55．前記塩基編集ポリペプチドがデアミナーゼである、実施態様54に記載のシステム。

【0282】

56．前記デアミナーゼが、配列番号374、383、397、399、407、408、411、414、416、420、514、および572～584のいずれか1つと少なくとも90%配列が一致する、実施態様55に記載のシステム。

20

【0283】

57．前記RGNポリペプチドが1つ以上の核局在化シグナルを含む、実施態様38～56のいずれか1つに記載のシステム。

【0284】

58．前記RGNポリペプチドが、真核細胞での発現のためコドン最適化されている、実施態様38～57のいずれか1つに記載のシステム。

【0285】

59．前記1つ以上のガイドRNAをコードするヌクレオチド配列とRGNポリペプチドをコードする前記ヌクレオチド配列が、1つのベクター上に位置する、実施態様38～58のいずれか1つに記載のシステム。

30

【0286】

60．1つ以上のドナーポリヌクレオチド、または前記1つ以上のドナーポリヌクレオチドをコードする1つ以上のヌクレオチド配列をさらに含む、実施態様38～59のいずれか1つに記載のシステム。

【0287】

61．標的DNA配列に結合させる方法であって、実施態様38～60のいずれか1つに記載のシステムを、前記標的DNA配列、または前記標的DNA配列を含む細胞に送達することを含む方法。

【0288】

62．前記RGNポリペプチドまたは前記ガイドRNAが検出可能な標識を含み、そのことによって前記標的DNA配列の検出が可能になる、実施態様61に記載の方法。

40

【0289】

63．前記ガイドRNAまたは前記RGNポリペプチドが発現モジュレータをさらに含み、そのことによって前記標的DNA配列、または前記標的DNA配列による転写制御下の遺伝子の発現を変化させる、実施態様61に記載の方法。

【0290】

64．標的DNA配列を切断するか変化させる方法であって、実施態様38～60のいずれか1つに記載のシステムを、前記標的DNA配列、または前記標的DNA配列を含む細胞に送達することを含む方法。

【0291】

50

65．前記改変された標的DNA配列が、前記標的DNA配列への異種DNAの挿入を含む、実施態様64に記載の方法。

【0292】

66．前記改変された標的DNA配列が、前記標的DNA配列からの少なくとも1個のヌクレオチドの欠失を含む、実施態様64に記載の方法。

【0293】

67．前記改変された標的DNA配列が、前記標的DNA配列の中に少なくとも1個のヌクレオチドの変異を含む、実施態様64に記載の方法。

【0294】

68．標的DNA配列に結合させる方法であって、

a) i) 前記標的DNA配列にハイブリダイズすることのできる1つ以上のガイドRNAと

ii) 配列番号1、16、24、35、43、または50と少なくとも95%配列が一致するアミノ酸配列を含むRGNポリペプチド

をインビトロで組み合わせることにより、RNA誘導型ヌクレアーゼ(RGN)リボヌクレオチド複合体を、前記RGNリボヌクレオチド複合体の形成に適した条件下で組み立て、

b) 前記標的DNA配列、または前記標的DNA配列を含む細胞を、インビトロで組み立てた前記RGNリボヌクレオチド複合体に接触させることを含み、

前記1つ以上のガイドRNAが前記標的DNA配列に結合し、そのことによって前記RGNポリペプチドを前記標的DNA配列に結合させる方法。

【0295】

69．前記RGNポリペプチドまたは前記ガイドRNAが検出可能な標識をさらに含み、そのことによって前記標的DNA配列の検出が可能になる、実施態様68に記載の方法。

【0296】

70．前記ガイドRNAまたは前記RGNポリペプチドが発現モジュレータをさらに含み、そのことによって前記標的DNA配列の発現の変更が可能になる、実施態様68に記載の方法。

【0297】

71．標的DNA配列を切断および/または改変する方法であって、前記標的DNA配列を、

a) 配列番号1、16、24、35、43、または50と少なくとも95%配列が一致するアミノ酸配列を含むRNA誘導型ヌクレアーゼ(RGN)ポリペプチド; および

b) 前記RGNを前記標的DNA配列に向かわせることのできる1つ以上のガイドRNAに接触させることを含み、

前記1つ以上のガイドRNAが前記標的DNA配列にハイブリダイズし、そのことによって前記RGNポリペプチドを前記標的DNA配列に結合させ、前記標的DNA配列の切断および/または改変が起こる方法。

【0298】

72．前記改変された標的DNA配列が、前記標的DNA配列への異種DNAの挿入を含む、実施態様71に記載の方法。

【0299】

73．前記改変された標的DNA配列が、前記標的DNA配列からの少なくとも1個のヌクレオチドの欠失を含む、実施態様71に記載の方法。

【0300】

74．前記改変された標的DNA配列が、前記標的DNA配列の中に少なくとも1個のヌクレオチドの変異を含む、実施態様71に記載の方法。

【0301】

75．前記RGNポリペプチドがニッカーゼである、実施態様71~74のいずれか1つに記載の方法。

【0302】

76．前記RGNポリペプチドがヌクレアーゼドドであり、塩基編集ポリペプチドに機能可能に連結されている、実施態様73または74に記載の方法。

【0303】

10

20

30

40

50

77. 前記gRNAが単一ガイドRNAである、実施態様68～76のいずれか1つに記載の方法。
- 【0304】
78. 前記gRNAが二重ガイドRNAである、実施態様68～76のいずれか1つに記載の方法。
- 【0305】
79. 前記gRNAが、配列番号2、17、25、36、44、51、または63と少なくとも95%配列が一致するヌクレオチド配列を含むCRISPR反復配列を含む、実施態様68～78のいずれか1つに記載の方法。
- 【0306】
80. 前記gRNAが、配列番号3、18、26、37、45、52、または62と少なくとも95%配列が一致するヌクレオチド配列を含むtracrRNAを含む、実施態様68～79のいずれか1つに記載の方法。
- 【0307】
81. 前記標的DNA配列が、プロトスペーサ隣接モチーフ(PAM)に隣接した位置にある、実施態様68～80のいずれか1つに記載の方法。
- 【0308】
82. 前記標的DNA配列が細胞の中にある、実施態様61～81のいずれか1つに記載の方法。
- 【0309】
83. 前記細胞が真核細胞である、実施態様82に記載の方法。
- 【0310】
84. 前記真核細胞が植物細胞である、実施態様83に記載の方法。
- 【0311】
85. 前記真核細胞が哺乳動物の細胞である、実施態様83に記載の方法。
- 【0312】
86. 前記真核細胞が昆虫細胞である、実施態様83に記載の方法。
- 【0313】
87. 前記細胞が原核細胞である、実施態様82に記載の方法。
- 【0314】
88. 前記細胞の培養を、前記RGNポリペプチドが発現して前記標的DNA配列を切断し、改変されたDNA配列を生成させる条件下で実施し；前記改変されたDNA配列を含む細胞を選択することをさらに含む、実施態様82～87のいずれか1つに記載の方法。
- 【0315】
89. 実施態様88に記載の方法による改変された標的DNA配列を含む細胞。
- 【0316】
90. 前記細胞が真核細胞である、実施態様89に記載の細胞。
- 【0317】
91. 前記真核細胞が植物細胞である、実施態様90に記載の細胞。
- 【0318】
92. 実施態様91に記載の細胞を含む植物。
- 【0319】
93. 実施態様91に記載の細胞を含む種子。
- 【0320】
94. 前記真核細胞が哺乳動物の細胞である、実施態様90に記載の細胞。
- 【0321】
95. 前記真核細胞が昆虫細胞である、実施態様90に記載の細胞。
- 【0322】
96. 前記細胞が原核細胞である、実施態様89に記載の細胞。
- 【0323】

10

20

30

40

50



97. 遺伝性疾患の原因変異の中に修正を有する遺伝子改変細胞を作製する方法であって、前記細胞の中に、

a) RNA誘導型ヌクレアーゼ (RGN) ポリペプチド (ただし前記RGNポリペプチドは、配列番号1、16、24、35、43、または50と少なくとも95%配列が一致するアミノ酸配列を含む)、または前記RGNポリペプチドをコードするポリヌクレオチド (ただし前記RGNポリペプチドをコードする前記ポリヌクレオチドは、前記細胞内で前記RGNポリペプチドの発現を可能にするためプロモータに機能可能に連結されている) と;

b) ガイドRNA (gRNA) (ただし前記gRNAは、配列番号2、17、25、36、44、51、または63と少なくとも95%配列が一致するヌクレオチド配列を含むCRISPR反復配列を含む)、または前記gRNAをコードするポリヌクレオチド (ただし前記gRNAをコードする前記ポリヌクレオチドは、前記細胞内で前記gRNAの発現を可能にするためプロモータに機能可能に連結されている)

を導入することを含み、

そのことによって前記のRGNとgRNAを前記原因変異のゲノム位置へと向かわせてゲノム配列を改変し、前記原因変異を除去する方法。

【0324】

98. 前記RGNポリペプチドがヌクレアーゼドッドであるか、ニッカーゼとして機能する、実施態様97に記載の方法。

【0325】

99. 前記RGNポリペプチドが塩基編集ポリペプチドに機能可能に連結されている、実施態様98に記載の方法。

【0326】

100. 前記塩基編集ポリペプチドがデアミナーゼである、実施態様99に記載の方法。

【0327】

101. 前記デアミナーゼが、配列番号374、383、397、399、407、408、411、414、416、420、514、および572~584のいずれか1つと少なくとも90%配列が一致するアミノ酸配列を含む、実施態様100に記載の方法。

【0328】

102. 前記細胞が動物細胞である、実施態様97~101のいずれか1つに記載の方法。

【0329】

103. 前記動物細胞が哺乳動物の細胞である、実施態様102に記載の方法。

【0330】

104. 前記細胞が、イヌ、ネコ、マウス、ラット、ウサギ、ウマ、ウシ、ブタ、またはヒトに由来する、実施態様103に記載の方法。

【0331】

105. 前記遺伝性疾患が表8に掲載されている、実施態様102に記載の方法。

【0332】

106. 前記遺伝性疾患がハーラー症候群である、実施態様102に記載の方法。

【0333】

107. 前記gRNAが、配列番号337を標的とするスペーサ配列を含む、実施態様106に記載の方法。

【0334】

108. 疾患の原因となるゲノム不安定性領域に欠失がある遺伝子改変細胞を作製する方法であって、前記細胞の中に、

a) RNA誘導型ヌクレアーゼ (RGN) ポリペプチド (ただし前記RGNポリペプチドは、配列番号1、16、24、35、43、または50と少なくとも95%配列が一致するアミノ酸配列を含む)、または前記RGNポリペプチドをコードするポリヌクレオチド (ただし前記RGNポリペプチドをコードする前記ポリヌクレオチドは、細胞内で前記RGNポリペプチドの発現を可能にするためプロモータに機能可能に連結されている) と;

b) 第1のガイドRNA (gRNA) (ただし前記gRNAは、配列番号2、17、25、36、44

10

20

30

40

50

、51、または63と少なくとも95%配列が一致するヌクレオチド配列を含むCRISPR反復配列を含む)、または前記gRNAをコードするポリヌクレオチド(ただし前記gRNAをコードする前記ポリヌクレオチドは、前記細胞内で前記gRNAの発現を可能にするためプロモータに機能可能に連結され、さらに前記gRNAは、前記ゲノム不安定性領域の5'フラнкを標的とするスペーサ配列を含む)と;

c) 第2のガイドRNA (gRNA) (ただし前記gRNAは、配列番号2、17、25、36、44、51、または63と少なくとも95%配列が一致するヌクレオチド配列を含むCRISPR反復配列を含む)、または前記gRNAをコードするポリヌクレオチド(ただし前記gRNAをコードする前記ポリヌクレオチドは、前記細胞内で前記gRNAの発現を可能にするためプロモータに機能可能に連結され、さらに前記第2のgRNAは、前記ゲノム不安定性領域の3'フラнкを標的とするスペーサ配列を含む)

10

を導入することを含み、

そのことによって前記RGNと前記2つのgRNAが前記ゲノム不安定領域を標的とし、前記ゲノム不安定領域の少なくとも一部が除去される方法。

【0335】

109. 前記細胞が動物細胞である、実施態様108に記載の方法。

【0336】

110. 前記動物細胞が哺乳動物の細胞である、実施態様108に記載の方法。

【0337】

111. 前記細胞が、イヌ、ネコ、マウス、ラット、ウサギ、ウマ、ウシ、ブタ、またはヒトに由来する、実施態様110に記載の方法。

20

【0338】

112. 前記遺伝性疾患がフリードライヒ運動失調症またはハンチントン病である、実施態様109に記載の方法。

【0339】

113. 前記第1のgRNAが、配列番号340、341、342、または343を標的とするスペーサ配列を含む、実施態様112に記載の方法。

【0340】

114. 前記第2のgRNAが、配列番号340、341、342、または343を標的とするスペーサ配列を含む、実施態様113に記載の方法。

30

【0341】

115. BCL11A mRNAとタンパク質の発現が減少した遺伝子改変哺乳動物造血前駆細胞を作製する方法であって、単離されたヒト造血前駆細胞の中に、

a) RNA誘導型ヌクレアーゼ(RGN)ポリペプチド(ただしRGNポリペプチドは、配列番号1、16、24、35、43、または50と少なくとも95%配列が一致するアミノ酸配列を含む)、または前記RGNポリペプチドをコードするポリヌクレオチド(ただし前記RGNポリペプチドをコードする前記ポリヌクレオチドは、前記細胞内で前記RGNポリペプチドの発現を可能にするためプロモータに機能可能に連結されている)と;

b) ガイドRNA (gRNA) (ただし前記gRNAは、配列番号2、17、25、36、44、51、または63と少なくとも95%配列が一致するヌクレオチド配列を含むCRISPR反復配列を含む)、または前記gRNAをコードするポリヌクレオチド(ただし前記gRNAをコードする前記ポリヌクレオチドは、前記細胞内で前記gRNAの発現を可能にするためプロモータに機能可能に連結されている)

40

を導入することを含み、

そのことによって前記のRGNとgRNAを前記細胞の中で発現させてBCL11Aエンハンサ領域の位置で切断する結果として、前記ヒト造血前駆細胞の遺伝子改変が起こり、BCL11AのmRNAおよび/またはタンパク質の発現が減少する方法。

【0342】

116. 前記gRNAが、配列番号350、351、または352を標的とするスペーサ配列を含む、実施態様115に記載の方法。

50

## 【0343】

117. 標的DNA配列に結合させる方法であって、

a) 前記標的DNA配列にハイブリダイズすることのできる1つ以上のガイドRNA、または前記1つ以上のガイドRNA (gRNA) をコードする1つ以上のヌクレオチド配列と；

b) 配列番号1、16、24、35、43、または50と少なくとも95%配列が一致するアミノ酸配列を含むRNA誘導型ヌクレアーゼ (RGN) ポリペプチドを含み、

前記1つ以上のガイドRNAが前記標的DNAにハイブリダイズし、

前記1つ以上のガイドRNAが前記RGNポリペプチドと複合体を形成し、そのことによって前記RGNポリペプチドを前記標的DNAに結合させるシステム。

## 【0344】

118. 前記RGNポリペプチドがヌクレアーゼドッドであるか、ニッカーゼとして機能する、実施態様117に記載のシステム。

## 【0345】

119. 前記RGNポリペプチドが塩基編集ポリペプチドに機能可能に融合されている、実施態様117または118に記載のシステム。

## 【0346】

120. 前記塩基編集ポリペプチドがデアミナーゼである、実施態様119に記載のシステム。

## 【0347】

121. 前記デアミナーゼポリペプチドが、配列番号374、383、397、399、407、408、411、414、416、420、514、および572~584のいずれか1つと少なくとも95%配列が一致するアミノ酸配列を含む、実施態様120に記載のシステム。

## 【0348】

122. デアミナーゼポリペプチドをコードするポリヌクレオチドを含む核酸分子であって、前記ポリヌクレオチドが、配列番号374、383、397、399、407、408、411、414、416、または420と少なくとも90%配列が一致するアミノ酸配列を含むデアミナーゼポリペプチドをコードするヌクレオチド配列を含み；

前記デアミナーゼポリペプチドが、標的ポリヌクレオチドの中の少なくとも1個のヌクレオチドを脱アミノ化し；

デアミナーゼポリペプチドをコードする前記ポリヌクレオチドが前記ポリヌクレオチドにとって異種のプロモータに機能可能に連結されている核酸分子。

## 【0349】

123. 前記デアミナーゼポリペプチドが、真核細胞での発現のためコドン最適化されている、実施態様122に記載の核酸分子。

## 【0350】

124. 前記デアミナーゼポリペプチドが、前記デアミナーゼポリペプチドを前記標的ポリヌクレオチドに局在させるDNA結合ポリペプチドに機能可能に連結されている、実施態様122または123に記載の核酸分子。

## 【0351】

125. 前記DNA結合ポリペプチドが、メガヌクレアーゼ、ジンクフィンガー融合タンパク質、またはTALENである、実施態様124に記載の核酸分子。

## 【0352】

126. 前記DNA結合ポリペプチドがガイドRNAとの複合体になって作用し、したがってRNA誘導型である、実施態様124に記載の核酸分子。

## 【0353】

127. 前記RNA誘導型DNA結合ポリペプチドがRNA誘導型ヌクレアーゼポリペプチドであるか、RNA誘導型ヌクレアーゼポリペプチドに由来する、実施態様126に記載の核酸分子。

## 【0354】

128. 前記RNA誘導型ヌクレアーゼポリペプチドがII型CRISPR-Casポリペプチドであ

10

20

30

40

50

る、実施態様127に記載の核酸分子。

【0355】

129．前記RNA誘導型ヌクレアーゼポリペプチドがV型CRISPR-Casポリペプチドである、実施態様127に記載の核酸分子。

【0356】

130．前記RNA誘導型ヌクレアーゼポリペプチドがニッカーゼである、実施態様126～129のいずれか1つに記載の核酸分子。

【0357】

131．前記RNA誘導型ヌクレアーゼポリペプチドが、配列番号1、16、24、35、43、または50と少なくとも95%配列が一致する、実施態様127に記載の核酸分子。

【0358】

132．前記デアミナーゼポリペプチドがウラシルグリコシラーゼ阻害剤(UGI)ポリペプチドに機能可能に連結されている、実施態様124～131のいずれか1つに記載の核酸分子。

【0359】

133．前記UGIポリペプチドが配列番号570のアミノ酸配列と少なくとも85%一致する、実施態様132に記載の核酸分子。

【0360】

134．前記デアミナーゼポリペプチドが核局在化シグナル(NLS)をさらに含む、実施態様122～133のいずれか1つに記載の核酸分子。

【0361】

135．実施態様122～134のいずれか1つに記載の核酸分子を含むベクター。

【0362】

136．前記標的ポリヌクレオチドにハイブリダイズすることができるガイドRNA(gRNA)をコードして前記RNA誘導型DNA結合ポリペプチドとの複合体になって作用する少なくとも1つのヌクレオチド配列をさらに含む、実施態様126～134のいずれか1つに記載の核酸分子を含むベクター。

【0363】

137．前記gRNAが単一ガイドRNAである、実施態様136に記載のベクター。

【0364】

138．前記gRNAが二重ガイドRNAである、実施態様136に記載のベクター。

【0365】

139．実施態様122～134のいずれか1つに記載の核酸分子または実施態様135～138のいずれか1つに記載のベクターを含む細胞。

【0366】

140．デアミナーゼポリペプチドを作製する方法であって、実施態様139に記載の細胞を前記デアミナーゼポリペプチドが発現する条件下で培養することを含む方法。

【0367】

141．アデノシンデアミナーゼポリペプチドをコードするポリヌクレオチドを含む核酸分子であって、前記ポリヌクレオチドが、

a) 配列番号514と少なくとも90%配列が一致するアミノ酸配列；

b) 配列番号572と少なくとも90%配列が一致して、配列番号572の102位に対応する位置にリシンを含み、配列番号572の104位に対応する位置にチロシンを含み、配列番号572の106位に対応する位置にトレオニンを含むアミノ酸配列；

c) 配列番号573と少なくとも90%配列が一致するアミノ酸配列；

d) 配列番号574と少なくとも90%配列が一致し、配列番号574の101位に対応する位置にグルタミン酸を含み、配列番号574の103位に対応する位置にセリンを含み、配列番号574の105位に対応する位置にリシンを含むアミノ酸配列；

e) 配列番号575と少なくとも90%配列が一致して、配列番号575の101位に対応する位置にリシンを含み、配列番号575の103位に対応する位置にロイシンを含み、配列

10

20

30

40

50

番号575の105位に対応する位置にグルタミン酸を含むアミノ酸配列；

f) 配列番号576と少なくとも90%配列が一致していて、配列番号576の105位に対応する位置にアラニンを含み、配列番号576の107位に対応する位置にアルギニンを含むアミノ酸配列；

g) 配列番号577と少なくとも90%配列が一致していて、配列番号577の102位に対応する位置にグリシンを含み、配列番号577の104位に対応する位置にセリンを含み、配列番号577の106位に対応する位置にアルギニンを含むアミノ酸配列；

h) 配列番号578と少なくとも90%配列が一致していて、配列番号578の105位に対応する位置にセリンを含み、配列番号578の107位に対応する位置にトレオニンを含むアミノ酸配列；

i) 配列番号579と少なくとも90%配列が一致していて、配列番号579の102位に対応する位置にセリンを含み、配列番号579の104位に対応する位置にグルタミンを含み、配列番号579の106位に対応する位置にグリシンを含むアミノ酸配列；

j) 配列番号580と少なくとも90%配列が一致していて、配列番号580の111位に対応する位置にグリシンを含むアミノ酸配列；

k) 配列番号581と少なくとも90%配列が一致していて、配列番号581の104位に対応する位置にグルタミンを含み、配列番号581の106位に対応する位置にグリシンを含み、配列番号581の108位に対応する位置にグルタミン酸を含むアミノ酸配列；

l) 配列番号582と少なくとも90%配列が一致していて、配列番号582の102位に対応する位置にアルギニンを含み、配列番号582の104位に対応する位置にトリプトファンを含み、配列番号582の106位に対応する位置にグルタミン酸を含むアミノ酸配列；

m) 配列番号583と少なくとも90%配列が一致していて、配列番号583の104位に対応する位置にアルギニンを含み、配列番号583の106位に対応する位置にセリンを含むアミノ酸配列；

n) 配列番号584と少なくとも90%配列が一致していて、配列番号584の110位に対応する位置にフェニルアラニンを含み、配列番号584の112位に対応する位置にセリンを含み、配列番号584の114位に対応する位置にトレオニンを含むアミノ酸配列からなるグループから選択されたアミノ酸配列を含むアデノシンデアミナーゼポリペプチドをコードするヌクレオチド配列を含み；

前記アデノシンデアミナーゼポリペプチドが、標的ポリヌクレオチドの中の少なくとも1個のアデノシンを脱アミノ化し；

アデノシンデアミナーゼポリペプチドをコードする前記ポリヌクレオチドが、前記ポリヌクレオチドにとって異種のプロモータに機能可能に連結されている核酸分子。

【0368】

142. 前記アデノシンデアミナーゼポリペプチドが、真核細胞での発現のためコドン最適化されている、実施態様141に記載の核酸分子。

【0369】

143. 前記アデノシンデアミナーゼポリペプチドが、前記デアミナーゼポリペプチドを前記標的ポリヌクレオチドに局在させるDNA結合ポリペプチドに機能可能に連結されている、実施態様141または142に記載の核酸分子。

【0370】

144. 前記DNA結合ポリペプチドが、メガヌクレアーゼ、ジンクフィンガー融合タンパク質、またはTALENである、実施態様143に記載の核酸分子。

【0371】

145. 前記DNA結合ポリペプチドがガイドRNAとの複合体になって作用し、したがってRNA誘導型である、実施態様143に記載の核酸分子。

【0372】

146. 前記RNA誘導型DNA結合ポリペプチドがRNA誘導型ヌクレアーゼポリペプチドであるか、RNA誘導型ヌクレアーゼポリペプチドに由来する、実施態様145に記載の核酸分子。

10

20

30

40

50

## 【0373】

147．前記RNA誘導型ヌクレアーゼポリペプチドがII型CRISPR-Casポリペプチドである、実施態様146に記載の核酸分子。

## 【0374】

148．前記RNA誘導型ヌクレアーゼポリペプチドがV型CRISPR-Casポリペプチドである、実施態様146に記載の核酸分子。

## 【0375】

149．前記RNA誘導型ヌクレアーゼポリペプチドがニッカーゼである、実施態様145～148のいずれか1つに記載の核酸分子。

## 【0376】

150．前記RNA誘導型ヌクレアーゼポリペプチドが、配列番号1、16、24、35、43、50のいずれかと少なくとも95%配列が一致する、実施態様141に記載の核酸分子。

## 【0377】

151．前記デアミナーゼポリペプチドがウラシルグリコシラーゼ阻害剤(UGI)ポリペプチドに機能可能に連結されている、実施態様141～150のいずれか1つに記載の核酸分子。

## 【0378】

152．前記UGIポリペプチドが配列番号570のアミノ酸配列と少なくとも85%一致する、実施態様151に記載の核酸分子。

## 【0379】

153．前記デアミナーゼポリペプチドが核局在化シグナル(NLS)をさらに含む、実施態様141～152のいずれか1つに記載の核酸分子。

## 【0380】

154．実施態様141～153のいずれか1つに記載の核酸分子を含むベクター。

## 【0381】

155．前記標的ポリヌクレオチドにハイブリダイズし、前記RNA誘導型DNA結合ポリペプチドとの複合体となって作用することのできるガイドRNA(gRNA)をコードする少なくとも1つのヌクレオチド配列をさらに含む、実施態様145～153のいずれか1つに記載の核酸分子を含むベクター。

## 【0382】

156．前記gRNAが単一ガイドRNAである、実施態様155に記載のベクター。

## 【0383】

157．前記gRNAが二重ガイドRNAである、実施態様155に記載のベクター。

## 【0384】

158．実施態様141～153のいずれか1つに記載の核酸分子または実施態様154～157のいずれか1つに記載のベクターを含む細胞。

## 【0385】

159．アデノシンデアミナーゼポリペプチドを作製する方法であって、実施態様158に記載の細胞を前記アデノシンデアミナーゼポリペプチドが発現する条件下で培養することを含む方法。

## 【0386】

160．a) 標的ポリヌクレオチドに結合するDNA結合ポリペプチドと；

b) デアミナーゼポリペプチド(ただし前記デアミナーゼポリペプチドは、配列番号374、383、397、399、407、408、411、414、416、または420と少なくとも90%配列が一致するアミノ酸配列を含む)を含み、前記デアミナーゼポリペプチドが、前記標的ポリヌクレオチドの中の少なくとも1個のヌクレオチドを脱アミノ化する融合タンパク質。

## 【0387】

161．前記DNA結合ポリペプチドが、メガヌクレアーゼ、ジンクフィンガー融合タンパク質、またはTALENである、実施態様160に記載の融合タンパク質。

## 【0388】

10

20

30

40

50

162. 前記DNA結合ポリペプチドがガイドRNAとの複合体になって作用し、したがってRNA誘導型である、実施態様160に記載の融合タンパク質。

【0389】

163. 前記RNA誘導型DNA結合ポリペプチドがRNA誘導型ヌクレアーゼポリペプチドであるか、RNA誘導型ヌクレアーゼポリペプチドに由来する、実施態様162に記載の融合タンパク質。

【0390】

164. 前記RNA誘導型ヌクレアーゼポリペプチドがII型CRISPR-Casポリペプチドである、実施態様163に記載の融合タンパク質。

【0391】

165. 前記RNA誘導型ヌクレアーゼポリペプチドがV型CRISPR-Casポリペプチドである、実施態様163に記載の融合タンパク質。

【0392】

166. 前記RNA誘導型ヌクレアーゼポリペプチドがニッカーゼである、実施態様162～165のいずれか1つに記載の融合タンパク質。

【0393】

167. 前記RNA誘導型ヌクレアーゼポリペプチドが、配列番号1、16、24、35、43、50のいずれかと少なくとも95%配列が一致する、実施態様163に記載の融合タンパク質。

【0394】

168. ウラシルグリコシラーゼ阻害剤(UGI)ポリペプチドをさらに含む、実施態様160～167のいずれか1つに記載の融合タンパク質。

【0395】

169. 前記UGIポリペプチドが配列番号570のアミノ酸配列と少なくとも85%一致する、実施態様168に記載の融合タンパク質。

【0396】

170. 核局在化シグナル(NLS)をさらに含む、実施態様160～169のいずれか1つに記載の融合タンパク質。

【0397】

171. a) 標的ポリヌクレオチドに結合するDNA結合ポリペプチドと；

b) i) 配列番号514と少なくとも90%配列が一致するアミノ酸配列；

ii) 配列番号572と少なくとも90%配列が一致していて、配列番号572の102位に対応する位置にリシンを含み、配列番号572の104位に対応する位置にチロシンを含み、配列番号572の106位に対応する位置にトレオニンを含むアミノ酸配列；

iii) 配列番号573と少なくとも90%配列が一致するアミノ酸配列；

iv) 配列番号574と少なくとも90%配列が一致していて、配列番号574の101位に対応する位置にグルタミン酸を含み、配列番号574の103位に対応する位置にセリンを含み、配列番号574の105位に対応する位置にリシンを含むアミノ酸配列；

v) 配列番号575と少なくとも90%配列が一致していて、配列番号575の101位に対応する位置にリシンを含み、配列番号575の103位に対応する位置にロイシンを含み、配列番号575の105位に対応する位置にグルタミン酸を含むアミノ酸配列；

vi) 配列番号576と少なくとも90%配列が一致していて、配列番号576の105位に対応する位置にアラニンを含み、配列番号576の107位に対応する位置にアルギニンを含むアミノ酸配列；

vii) 配列番号577と少なくとも90%配列が一致していて、配列番号577の102位に対応する位置にグリシンを含み、配列番号577の104位に対応する位置にセリンを含み、配列番号577の106位に対応する位置にアルギニンを含むアミノ酸配列；

viii) 配列番号578と少なくとも90%配列が一致していて、配列番号578の105位に対応する位置にセリンを含み、配列番号578の107位に対応する位置にトレオニンを含むアミノ酸配列；

ix) 配列番号579と少なくとも90%配列が一致していて、配列番号579の102位に対応

10

20

30

40

50

する位置にセリンを含み、配列番号579の104位に対応する位置にグルタミンを含み、配列番号579の106位に対応する位置にグリシンを含むアミノ酸配列；

x) 配列番号580と少なくとも90%配列が一致していて、配列番号580の111位に対応する位置にグリシンを含むアミノ酸配列；

xi) 配列番号581と少なくとも90%配列が一致していて、配列番号581の104位に対応する位置にグルタミンを含み、配列番号581の106位に対応する位置にグリシンを含み、配列番号581の108位に対応する位置にグルタミン酸を含むアミノ酸配列；

xii) 配列番号582と少なくとも90%配列が一致していて、配列番号582の102位に対応する位置にアルギニンを含み、配列番号582の104位に対応する位置にトリプトファンを含み、配列番号582の106位に対応する位置にグルタミン酸を含むアミノ酸配列；

xiii) 配列番号583と少なくとも90%配列が一致していて、配列番号583の104位に対応する位置にアルギニンを含み、配列番号583の106位に対応する位置にセリンを含むアミノ酸配列；

xiv) 配列番号584と少なくとも90%配列が一致していて、配列番号584の110位に対応する位置にフェニルアラニンを含み、配列番号584の112位に対応する位置にセリンを含み、配列番号584の114位に対応する位置にトレオニンを含むアミノ酸配列からなるグループから選択されたアミノ酸配列を含むアデノシンデアミナーゼポリペプチドを含み；

前記アデノシンデアミナーゼポリペプチドが、標的ポリヌクレオチドの中の少なくとも1個のアデノシンを脱アミノ化する融合タンパク質。

#### 【0398】

172. 前記DNA結合ポリペプチドが、メガヌクレアーゼ、ジンクフィンガー融合タンパク質、またはTALENである、実施態様171に記載の融合タンパク質。

#### 【0399】

173. 前記DNA結合ポリペプチドがガイドRNAとの複合体になって作用し、したがってRNA誘導型である、実施態様171に記載の融合タンパク質。

#### 【0400】

174. 前記RNA誘導型DNA結合ポリペプチドがRNA誘導型ヌクレアーゼポリペプチドであるか、RNA誘導型ヌクレアーゼポリペプチドに由来する、実施態様173に記載の融合タンパク質。

#### 【0401】

175. 前記RNA誘導型ヌクレアーゼポリペプチドがII型CRISPR-Casポリペプチドである、実施態様174に記載の融合タンパク質。

#### 【0402】

176. 前記RNA誘導型ヌクレアーゼポリペプチドがV型CRISPR-Casポリペプチドである、実施態様174に記載の融合タンパク質。

#### 【0403】

177. 前記RNA誘導型ヌクレアーゼポリペプチドがニッカーゼである、実施態様173～176のいずれか1つに記載の融合タンパク質。

#### 【0404】

178. 前記RNA誘導型ヌクレアーゼポリペプチドが、配列番号1、16、24、35、43、または50と少なくとも95%配列が一致する、実施態様174に記載の融合タンパク質。

#### 【0405】

179. ウラシルグリコシラーゼ阻害剤(UGI)ポリペプチドをさらに含む、実施態様171～178のいずれか1つに記載の融合タンパク質。

#### 【0406】

180. 前記UGIポリペプチドが配列番号570のアミノ酸配列と少なくとも85%一致する、実施態様179に記載の融合タンパク質。

#### 【0407】

181. 核局在化シグナル(NLS)をさらに含む、実施態様171～180のいずれか1つに

10

20

30

40

50



記載の融合タンパク質。

【0408】

182. 標的ポリヌクレオチド配列を改変するためのシステムであって、

a) 前記標的ポリヌクレオチド配列にハイブリダイズすることのできる1つ以上のガイドRNA、または前記1つ以上のガイドRNA (gRNA) をコードする1つ以上のヌクレオチド配列と；

b) 実施態様162～170と173～181のいずれか1つに記載の融合タンパク質、または前記融合タンパク質をコードするヌクレオチド配列を含み、

前記1つ以上のガイドRNAをコードする前記ヌクレオチド配列と、前記融合タンパク質をコードする前記ヌクレオチド配列が、それぞれ、前記ヌクレオチド配列にとって異種のプロモータに機能可能に連結され、

前記1つ以上のガイドRNAが前記標的ポリヌクレオチド配列にハイブリダイズし、

前記1つ以上のガイドRNAが前記融合タンパク質のRNA誘導型DNA結合ポリペプチドと複合体を形成し、そのことによって前記融合タンパク質を前記標的ポリヌクレオチド配列に結合させ、この標的ポリヌクレオチド配列を改変するシステム。

【0409】

183. 前記gRNAが単一ガイドRNA (sgRNA) である、実施態様182に記載のシステム。

【0410】

184. 前記gRNAが二重ガイドRNAである、実施態様182に記載のシステム。

【0411】

185. 前記標的DNA配列が、プロトスペーサ隣接モチーフ (PAM) に隣接した位置にある、実施態様182～184のいずれか1つに記載のシステム。

【0412】

186. 前記標的ポリヌクレオチド配列が細胞の中にある、実施態様182～185のいずれか1つに記載のシステム。

【0413】

187. 前記細胞が真核細胞である、実施態様186に記載のシステム。

【0414】

188. 前記真核細胞が植物細胞である、実施態様187に記載のシステム。

【0415】

189. 前記真核細胞が哺乳動物の細胞である、実施態様187に記載のシステム。

【0416】

190. 前記真核細胞が昆虫細胞である、実施態様187に記載のシステム。

【0417】

191. 前記細胞が原核細胞である、実施態様186に記載のシステム。

【0418】

192. 標的ポリヌクレオチドを脱アミノ化する方法であって、前記標的ポリヌクレオチドを、配列番号374、383、397、399、407、408、411、414、416、または420と少なくとも90%配列が一致するアミノ酸配列を含むデアミナーゼに接触させることを含み、前記デアミナーゼポリペプチドが前記標的ポリヌクレオチドの中の少なくとも1個のヌクレオチドを脱アミノ化する方法。

【0419】

193. 標的ポリヌクレオチドの中の少なくとも1個のアデノシンを脱アミノ化する方法であって、前記標的ポリヌクレオチドを、

a) 配列番号514と少なくとも90%配列が一致するアミノ酸配列；

b) 配列番号572と少なくとも90%配列が一致していて、配列番号572の102位に対応する位置にリシンを含み、配列番号572の104位に対応する位置にチロシンを含み、配列番号572の106位に対応する位置にトレオニンを含むアミノ酸配列；

c) 配列番号573と少なくとも90%配列が一致するアミノ酸配列；

10

20

30

40

50

d) 配列番号574と少なくとも90%配列が一致していて、配列番号574の101位に対応する位置にグルタミン酸を含み、配列番号574の103位に対応する位置にセリンを含み、配列番号574の105位に対応する位置にリシンを含むアミノ酸配列；

e) 配列番号575と少なくとも90%配列が一致していて、配列番号575の101位に対応する位置にリシンを含み、配列番号575の103位に対応する位置にロイシンを含み、配列番号575の105位に対応する位置にグルタミン酸を含むアミノ酸配列；

f) 配列番号576と少なくとも90%配列が一致していて、配列番号576の105位に対応する位置にアラニンを含み、配列番号576の107位に対応する位置にアルギニンを含むアミノ酸配列；

g) 配列番号577と少なくとも90%配列が一致していて、配列番号577の102位に対応する位置にグリシンを含み、配列番号577の104位に対応する位置にセリンを含み、配列番号577の106位に対応する位置にアルギニンを含むアミノ酸配列；

h) 配列番号578と少なくとも90%配列が一致していて、配列番号578の105位に対応する位置にセリンを含み、配列番号578の107位に対応する位置にトレオニンを含むアミノ酸配列；

i) 配列番号579と少なくとも90%配列が一致していて、配列番号579の102位に対応する位置にセリンを含み、配列番号579の104位に対応する位置にグルタミンを含み、配列番号579の106位に対応する位置にグリシンを含むアミノ酸配列；

j) 配列番号580と少なくとも90%配列が一致していて、配列番号580の111位に対応する位置にグリシンを含むアミノ酸配列；

k) 配列番号581と少なくとも90%配列が一致していて、配列番号581の104位に対応する位置にグルタミンを含み、配列番号581の106位に対応する位置にグリシンを含み、配列番号581の108位に対応する位置にグルタミン酸を含むアミノ酸配列；

l) 配列番号582と少なくとも90%配列が一致していて、配列番号582の102位に対応する位置にアルギニンを含み、配列番号582の104位に対応する位置にトリプトファンを含み、配列番号582の106位に対応する位置にグルタミン酸を含むアミノ酸配列；

m) 配列番号583と少なくとも90%配列が一致していて、配列番号583の104位に対応する位置にアルギニンを含み、配列番号583の106位に対応する位置にセリンを含むアミノ酸配列；

n) 配列番号584と少なくとも90%配列が一致していて、配列番号584の110位に対応する位置にフェニルアラニンを含み、配列番号584の112位に対応する位置にセリンを含み、配列番号584の114位に対応する位置にトレオニンを含むアミノ酸配列からなるグループから選択されたアミノ酸配列を含むアデノシンデアミナーゼポリペプチドに接触させることを含み；

前記アデノシンデアミナーゼポリペプチドが、標的ポリヌクレオチドの中の少なくとも1個のアデノシンを脱アミノ化する方法。

#### 【0420】

194. 標的ポリヌクレオチドを改変する方法であって、前記標的ポリヌクレオチドを実施態様160、161、171、および172のいずれか1つに記載の融合タンパク質に接触させることを含み、前記DNA結合ポリペプチドが前記標的ポリヌクレオチドに結合して前記デアミナーゼが前記標的ポリヌクレオチドの中の少なくとも1個のクレオチドを脱アミノ化する方法。

#### 【0421】

195. 標的ポリヌクレオチドを改変する方法であって、前記標的ポリヌクレオチドを実施態様162~170と173~181のいずれか1つに記載の融合タンパク質に接触させて、1つ以上のガイドRNA (gRNA)、または前記1つ以上のgRNAをコードする1つ以上のヌクレオチド配列を導入することを含み；

前記1つ以上のgRNAが前記標的ポリヌクレオチドにハイブリダイズして前記融合タンパク質の前記RNA誘導型DNA結合ポリペプチドと複合体を形成し、そのことによって前記融合タンパク質を前記標的ポリヌクレオチド配列に結合させ、前記融合タンパク質の前記デ

10

20

30

40

50

アミナーゼポリペプチドが前記標的ポリヌクレオチドの中の少なくとも1個のヌクレオチドを脱アミノ化し、そのことによって前記標的ポリヌクレオチド配列を改変する方法。

【0422】

196．前記標的ポリヌクレオチドが細胞の中にある、実施態様195に記載の方法。

【0423】

197．前記細胞が真核細胞である、実施態様196に記載の方法。

【0424】

198．前記真核細胞が哺乳動物の細胞である、実施態様197に記載の方法。

【0425】

199．前記真核細胞が植物細胞である、実施態様197に記載の方法。

10

【0426】

200．前記標的ポリヌクレオチドの改変がCからTへの点変異を含む、実施態様195～199のいずれか1つに記載の方法。

【0427】

201．C塩基が脱アミノ化する結果として疾患または障害に関係する配列が修正される、実施態様200に記載の方法。

【0428】

202．前記標的ポリヌクレオチドの改変が作物植物のゲノム内のCからTへの変化を含み、C塩基が脱アミノ化する結果として作物植物の農業上の品質を改善する配列になる、実施態様200に記載の方法。

20

【0429】

203．前記標的ポリヌクレオチドの改変がAからGへの点変異を含む、実施態様195～199のいずれか1つに記載の方法。

【0430】

204．A塩基が脱アミノ化する結果として疾患または障害に関係する配列が修正される、実施態様203に記載の方法。

【0431】

205．前記標的ポリヌクレオチドの改変が作物植物のゲノム内のAからGへの変化を含み、A塩基が脱アミノ化する結果として作物植物の農業上の品質を改善する配列になる、実施態様203に記載の方法。

30

【0432】

以下の実施例は説明のために提供されているのであり、制限するためではない。

【0433】

実験

【0434】

実施例1：RNA誘導型ヌクレアーゼの同定

【0435】

6つの異なるCRISPR関連RNA誘導型ヌクレアーゼ（RGN）が同定されたため、それらを下記の表1に記載してある。APG00969、APG03128、およびAPG00771はII-C型RGNである。APG09748、APG02789、およびAPG09106はV-B型RGNである。表1には、各RGNの名称、そのアミノ酸配列、RGCの出所、および処理されたcrRNA配列とtracrRNA配列が示されている。表1には一般的な単一ガイドRNA（sgRNA）配列がさらに示されていて、それが、このsgRNAの核酸標的配列を決めている。スペーサ配列の位置はポリNによって示されている。このポリN配列は、sgRNAの中のスペーサ配列の位置を示しているだけであり、機能するスペーサ配列に必要な長さを示しているわけではない。II-C型RGNシステムは、それぞれ、tracrRNAのヘアピンステムの塩基中に保存された配列を持つ。すなわちAPG00969はUNANNC（配列番号13）を持ち；APG03128はANGNNU（配列番号23）を持ち；APG00771はUNANNA（配列番号42）を持つ。

40

【0436】

50

【表 1】

表1：配列番号とCRISPR関連システムのまとめ

RGN の名称	配列番号	出所	crRNA 反復配列 (配列番号)	tracrRNA (配列番号)	sgRNA (配列番号)
APG00969	1	バチルス属の種	2	3	4
APG03128	16	リゾビウム属の種	17	18	19
APG09748	24	ブレビバチルス属の種	25	26	27
APG00771	35	クリセオバクテリウム属の種	36	37	38
APG02789	43	バチルス属の種	44	45	46
APG09106	50	ブレビバチルス属の種	51	52	53

10

## 【0437】

## 実施例2：ガイドRNAの同定とsgRNAの構築

## 【0438】

調べているRNA誘導型ヌクレアーゼシステムを元々発現する細菌の培養物を対数中期 (OD600が約0.600) まで増殖させ、ペレット化した後、瞬間凍結させた。mirVANA miRNA Isolation Kit (Life Technologies社、カールスバッド、カリフォルニア州) を用いてRNAをペレットから単離し、単離されたRNAからNEBNext Small RNA Library Prep Kit (NEB社、Beverly, MA) を用いてシーケンシングライブラリを調製した。このライブラリ調製物を6%ポリアクリルアミドゲル上で18~65 nt RNA種と90~200 nt RNA種に対応する2つのサイズ分画に分画し、crRNAとtracrRNAをそれぞれ検出した。深いシーケンシング (小さいほうの分画に関する40 bpのペアードエンドと、大きいほうの分画に関する80 bpのペアードエンド) をサービス提供者 (MoGene社、セントルイス、ミズーリ州) がNext Seq 500 (High Outputキット) 上で実施した。Cutadaptを用いてリードを品質でトリミングし、Bowtie2を用いて参照ゲノムにマッピングした。カスタムRNAseqパイプラインをPhytonで書いてcrRNA転写産物とtracrRNA転写産物を検出した。天然反復スペーサアレイの配列カバレッジにより、処理されたcrRNAの境界を決定した。許容可能なBLASTnパラメータを用いてtracrRNAのアンチ反復部分を同定した。RNAシーケンシングの深度から、処理されたtracrRNAの境界を、アンチ反復を含有する転写産物を同定することによって確認した。RNA折り畳みソフトウェアであるNUPACKによる二次構造予測を利用してRNAのマニュアルキュレーションを実施した。あるいはAPG02789に関するtracrRNAは、バイオインフォマティクスを利用し、小さなRNAのシーケンシングを通じてではなく抗反復配列によって決定した。一般にsgRNAカセットはDNA合成によって調製し、APG00771、APG03128、およびAPG00969については一般に以下のように設計した (5' 3') : 20~30 bpのスペーサ配列 -- crRNAの処理された反復部分 -- 4 bpの非相補的リンカー (AAAG ; 配列番号8) -- 処理されたtracrRNA。APG09748、APG09106、およびAPG02789については、sgRNAカセットを以下のように設計した (5' 3') : 処理されたtracrRNA -- 4 bpの非相補的リンカー (AAAG ; 配列番号8) -- crRNAの処理された反復部分 -- 20~30 bpのスペーサ配列。本分野で知られている他の4 bpまたは6 bpの非相補的リンカーもsgRNAの設計に使用できる。インビトロアッセイのため、GeneArt (商標) Precision gRNA Synthesis Kit (ThermoFisher社) を用いてインビトロでsgRNAカセットを転写することによってsgRNAを合成した。それぞれのRGNポリペプチドに関する処理されたcrRNA配列とtracrRNA配列を同定し、表1に示す。PAMライブラリ1と2のために構成したsgRNAに関しては下記の記述を参照されたい。

20

30

40

## 【0439】

## 実施例3：各RGNのためのPAMの条件の決定

## 【0440】

本質的にKleinstiver他 (2015年) Nature 第523巻 : 481~485ページとZetsche他 (2015年) Cell第163巻 : 759~771ページから改変したPAM減少アッセイを利用して

50

各RGNのためのPAMの条件を決定した。簡単に述べると、2つのプラスミドライブラリ（L1とL2）をpUC18骨格（ampR）の中で生成させた。それぞれのライブラリは、8個のランダムなヌクレオチド（すなわちPAM領域）が隣接した明確な30 bpのプロトスペース（標的）配列を含有している。RGN APG00969、APG03128、およびAPG00771に関するライブラリ1とライブラリ2の標的配列とフランキングPAM領域は、それぞれ配列番号14と15である。RGN APG09748、APG02789、およびAPG09106に関するライブラリ1とライブラリ2の標的配列とフランキングPAM領域は、それぞれ配列番号32と33である。

【0441】

これらライブラリを別々に、電気穿孔で大腸菌BL21(DE3)細胞に入れた。この大腸菌細胞の中には、本発明のRGN（大腸菌用にコドンが最適化されている）を含有するとともに、L1またはL2の中のプロトスペースに対応するスペース配列を含むコグネイトsgRNAを含有するpRSF-1b発現ベクターが存在している。十分なライブラリプラスミドを形質転換反応物の中で用いて $10^6$ 超のCFUを得た。pRSF-1b骨格の中のRGNとsgRNAの両方とも、T7プロモータの制御下にあった。この形質転換反応物を1時間かけて回収した後、カルベニシリンとカナマイシンを含有するLB培地の中に入れて希釈し、一晚増殖させた。翌日、この混合物を自己誘導性Overnight Express（商標）Instant TB Medium（Millipore Sigma社）の中に入れて希釈し、RGNとsgRNAを発現させ、さらに4時間または20時間にわたって増殖させた後、細胞を沈降させ、Mini-prepキット（Qiagen社、ジャーマンタウン、メリーランド州）を用いてプラスミドDNAを単離した。適切なsgRNAの存在下では、RGNが認識できるPAMを含有するプラスミドが切断されて、集団からそれらプラスミドが除去される。RGNが認識できないPAMを含有するプラスミド、または適切なsgRNAを含有していない細菌の中に形質転換のため入れられるPAMを含有するプラスミドは生き延びて複製される。切断されないプラスミドのPAM領域とプロトスペース領域をPCRで増幅し、公開されているプロトコル（16sメタゲノムライブラリ調製ガイド15044223B、Illumina社、サン・ディエゴ、カリフォルニア州）に従ってシーケンシングの準備をした。サービス提供者（MoGene社、セントルイス、ミズーリ州）によって深いシーケンシング（80 bp単一エンドリード）がMiSeq（Illumina社）上で実施された。典型的には、アンプリコン1つ当たり1～4M個のリードを取得した。各サンプルでPAM領域を取り出し、カウントした後、全リードに規格化した。プラスミドの切断につながるPAMは、対照と比べたとき（すなわちRGNを含有するが適切なsgRNAを欠く大腸菌をライブラリで形質転換したとき）、より少ないことによって同定された。新規なRGNのためのPAMの条件を表わすため、問題の領域内の全配列についての減少比（サンプル中の頻度／対照中の頻度）を、 $-\log_2$ 変換を用いて富化値に変換した。十分なPAMは、富化値が2.3超（減少比が約0.2未満に対応）であるPAMと定義した。両方のライブラリでこの閾値よりも大きいPAMを回収し、ウェブ・ロゴを生成させるのに使用した。これは、例えばインターネット上でウェブに基づくサービスである「weblogo」として知られるサービスを利用して生成させることができる。PAM配列を同定し、富化された上位のPAMの中に一定のパターンがあるときに報告した。それぞれのRGNについて（富化値（EF）が2.3よりも大きい）PAMを表2に提示する。いくつかのRGNについては、（EFが3.3超である）PAMの非限定的な例も同定された。PAMの方向は、APG00969、APG03128、APG00771については5'-標的-PAM-3'であるのに対し、APG09748、APG09106、およびAPG02789については5'-PAM-標的-3'である。

【0442】

10

20

30

40

【表 2】

表2：PAM決定

RGN の名称	sgRNA L1 (配列番号)	sgRNA L2 (配列番号)	PAM (配列番号)
APG00969	5	6	7
APG03128	20	21	22
APG09748	28	29	30
APG00771	39	40	41
APG02789	47	48	49
APG09106	54	55	30

## 【0443】

10

実施例4：ガイドRNAを操作してヌクレアーゼ活性を増大させる

## 【0444】

配列が互いに非常によく一致していて同じPAMを有するRGN APG09748とAPG09106について、RNA折りたたみ予測を利用して、ガイドRNAの中であってヌクレアーゼ活性を最適化するために変化させることのできる領域を明らかにする。反復：アンチ反復領域を短縮し、G-C塩基対を付加し、G-Uゆらぎ対を除去することにより、反復：アンチ反復領域の中のcrRNA：tracrRNA塩基対合の安定性を増大させた。インビトロ切断アッセイにおいてRGN APG09748を用いて「最適化された」ガイドバリエーションを試験し、野生型gRNAと比較した。

## 【0445】

20

RPNを形成するためのRGNを作製するため、C末端His6タグに融合されたRGN（配列番号593）またはC末端His10タグに融合されたRGN（配列番号594）を含有する発現プラスミドを構成して大腸菌のBL21（DE3）株に導入し、この株を形質転換した。50 µg/mlのカナマイシンを補足したMagicMedia（Thermo Fisher社）を用いて発現させた。溶解と透明化の後、タンパク質を固定化金属アフィニティクロマトグラフィーによって精製し、Qubitタンパク質定量キット（Thermo Fisher社）を利用して定量するか、消滅係数の計算値を用いてUV-可視光により定量した。

## 【0446】

精製されたRGNをsgRNAと約2：1の比率で室温にて20分間インキュベートすることによってリボ核タンパク質（RNP）を調製した。インビトロ切断反応のため、RPNを、好ましいPAM配列が隣接した標的プロトスペーサを含有するプラスミドまたは直線状dsDNAと室温で30分間超インキュベートした。TRAC遺伝子座の中の2つの標的核酸配列、すなわちTRAC11（配列番号60）とTRAC14（配列番号61）を調べた。標的とされる活性に関するgRNAのアッセイを、正しい標的核酸配列（例えばgRNAはTRAC11スペーサ配列を持ち、調べる標的はTRAC11である）ありと、正しい標的核酸配列（例えばgRNAはTRAC11スペーサ配列を持ち、調べる標的はTRAC14である）なしの両方について実施した。プラスミドによる切断で求まる活性をアガロースゲル電気泳動によって評価する。結果を表3に示す。ガイドバリエーションは配列番号56～59として掲載されており、スペーサ配列とともに提示されている。これらガイド配列はAAAA（配列番号31）という非相補的なヌクレオチドリンカーを利用している。反復：アンチ反復結合が増大した最適化されたgRNA（配列番号64；ポリ-Nはスペーサ配列の位置を示す）は、最適化されたtracrRNA（配列番号62）要素と最適化されたcrRNA（配列番号63）要素を有する。最適化されたガイドバリエーションは、以前に野生型ガイドRNAを用いたときには切断が検出されなかった2つの遺伝子座を切断することができた。反復：アンチ反復領域におけるハイブリダイゼーションの最適化を通じ、インビトロでのAPG09748の切断が、TRAC遺伝子座内の複数の標的に関して0%切断から100%切断へと増加した。

## 【0447】

50

## 【表 3】

表3：操作されたガイドバリエーションを用いたAPG09748の編集効率

gRNA バリエーション (配列番号)	ガイドの 設計	調べた標識	ゲル 1 - 2 $\mu$ l ロード		ゲル 2 - 1 $\mu$ l ロード	
			無傷%	切断%	無傷%	切断%
56	最適化	TRAC11	68	32	57	43
56	最適化	TRAC14	100	0	100	0
57	最適化	TRAC11	100	0	100	0
57	最適化	TRAC14	70	30	69	31
58	野生型	TRAC11	100	0	100	0
58	野生型	TRAC14	100	0	100	0
59	野生型	TRAC11	100	0	100	0
59	野生型	TRAC14	100	0	100	0
なし		TRAC11	100	0	100	0
なし		TRAC14	100	0	100	0

## 【0448】

追加の最適化されたgRNAバリエーションを設計して調べた。さらに、スペーサの長さがどのように切断効率に影響するかを明らかにするため、異なる長さのスペーサ配列も調べた。スペーサ配列の外部にあるsgRNAをこのアッセイでは「骨格」と呼ぶ。表4では、これらは「WT」（配列番号53、野生型配列）と表記されているのに対し、3つの最適化されたsgRNAは、V1（配列番号65）、V2（配列番号66）、およびV3（配列番号64）として表記されている。これらの配列はすべて、スペーサ配列の位置を示すポリ-Nを有する。ガイドは、インビトロでの転写（IVT）によりsgRNAとして発現した。野生型sgRNA骨格と比較すると、V1は87.8%が一致し、V2は92.4%が一致し、V3は85.5%が一致している。二重ガイドRNAを表わすがそれ以外は上記の野生型sgRNAおよび最適化されたsgRNAと同様である合成tracrRNA：crRNA二重鎖（「合成」）も作製して調べた。

## 【0449】

アッセイのこのセットでは、RGN APG09106を使用した。それ以外は、インビトロ切断反応の方法は、上に記載してあるのと同様であった。標的にされる核酸配列は、標的1（配列番号67）と標的2（配列番号68）であった。結果を表4に示す。

## 【0450】

## 【表 4】

表4：操作されたガイドバリエーションを用いたAPG09106の編集効率

RNA 供給源	標的	スペーサの長さ	骨格	スペーサの配列番号	切断率（%）
合成	2	18	WT	69	12.3
合成	1	20	WT	70	0
合成	2	20	WT	71	55.0
合成	1	25	WT	72	0
合成	2	25	WT	73	61.4
IVT	2	25	V1	74	1.1
IVT	2	25	V2	75	0.9
IVT	2	25	V3	76	0.7
IVT	2	20	V3	77	21.0
IVT	1	25	V3	78	2.0

## 【0451】

実施例5：哺乳動物の細胞における遺伝子編集活性の実証

## 【0452】

実施例5.1：哺乳動物の細胞におけるAPG02789の活性

## 【0453】

ヒトでの発現にコドンが最適化されたRGNヌクレオチド配列をN末端の核局在化タグ付きで合成し、pcDNA3.1 CMV発現プラスミドに入れてクローニングした。RGNポリペプチドの最終コンストラクトは以下の通りである：N末端-SV40 NLS（配列番号10）-3×FLAGタグ（配列番号11）-RGN配列（配列番号1、16、24、35、43、または50）-ヌクレオプラスミンNLS（配列番号12）-C末端。Herculase II（Agilent Technologies社）

を用い、sgRNA配列の発現を駆動するU6プロモータを含むPCRアンプリコンを生成させる。400 ngのRGN発現プラスミドと100 ngのsgRNA PCR産物を、Lipofectamine 2000試薬（Life Technologies社）を用いて24ウエルのプレートの75～90%の集密度になったHEK293T細胞にトランスフェクトする。トランスフェクション後の72時間にわたって細胞を37℃でインキュベートした後、ゲノムDNAを抽出する。ゲノムDNAは、QuickExtract DNA Extraction Solution（Epicentre社）を製造者のプロトコルに従って用いて抽出する。RGN標的部位に隣接するゲノム領域をPCRで増幅し、産物を、QiaQuickスピンカラム（Qiagen社）を製造者のプロトコルに従って用いて精製する。精製された合計200～500 ngのPCR産物を1 µlの10×Taq DNAポリメラーゼPCRバッファ（Enzymatics社）および超純水と混合して最終体積を10 µlにした後、再アニーリングプロセスを適用することでヘテロ二重鎖の形成が可能になるようにする：95℃で10分間、-2℃/秒の勾配で95℃から85℃へ；-0.25℃/秒の勾配で85℃から25℃へ；そして25℃に1分間保持。再アニーリングの後、産物を、SURVEYOR（登録商標）ヌクレアーゼとSURVEYOR（登録商標）エンハンサS（Integrated DNA Technologies社）を製造者が推奨するプロトコルに従って用いて処理し、4～20% Novex TBEポリアクリルアミドゲル（Life Technologies社）上で分析する。ゲルをSYBR Gold DNA染色（Life Technologies社）で10分間染色し、Gel Docゲルイメージングシステム（Bio-rad社）を用いて画像を取得した。定量は、相対バンド強度に基づく。indel率は、式： $100 \times (1 - (1 - (b + c) / (a + b + c))^{1/2})$ によって求めた（ただしaは消化されなかったPCR産物の積分強度であり、bとcは各切断産物の積分強度である）。

#### 【0454】

RGN APG02789について、上に記載されているようにして方法を実施した。ヒトゲノム内の多数の異なる遺伝子をRNA誘導型切断の標的とした。これら遺伝子座は、sgRNAの配列番号とともに下記の表5に含まれている。RGN活性の1つの指標であるindel率も示されている。

#### 【0455】

#### 【表5】

表5：哺乳動物の細胞におけるAPG02789の活性

sgRNA	遺伝子標的	INDEL 率 (%)
79	DNMT1	N. D.
80	VEGFA	N. D.
81	EMX1	45.6
82	AurkB	44.3
83	AurkB	31.6
84	AurkB	25.5
85	HPRT1	N. D.
86	HPRT1	15.7
87	HPRT1	N. D.
88	RelA	N. D.
89	RelA	N. D.
90	RelA	N. D.

#### 【0456】

実施例5.2：哺乳動物の細胞におけるAPG09106の活性

#### 【0457】

RGN発現カセットを作製し、哺乳動物で発現させるためのベクターに導入した。RGN APG00969、APG03128、APG09748、APG09106、およびAPG02789（それぞれ配列番号357～361）は、それぞれ、ヒトにおける発現のためにコドンが最適化されており、発現するタンパク質は、N末端の位置でSV40核局在化配列（配列番号10）と3×FLAGタグ（配列番号11）に機能可能に融合され、C末端の位置で核プラスミンNLS配列（配列番号12）に機能可能に融合されていた。タンデムに機能可能に融合された2コピーのNLS配



列を使用した。それぞれの発現カセットは、サイトメガロウイルス (CMV) プロモータ (配列番号334) の制御下にあった。本分野では、CMV転写エンハンサ (配列番号335) もCMVプロモータを含むコンストラクトの中に含めることができることが知られている。単一のgRNAをコードしてそれぞれがヒトRNAポリメラーゼIII U6プロモータ (配列番号336) の制御下にある複数のガイドRNA発現コンストラクトを作製し、発現ベクターに導入した。ガイドは、AurkB遺伝子の領域を標的とした。1つのRNA誘導型ヌクレアーゼで特定の残基が変異してタンパク質のヌクレアーゼ活性が増大した。特にAPG09106のT849残基がアルギニンに変異した (配列番号362)。この点変異が哺乳動物の細胞における編集率を増大させた。

#### 【0458】

10

上に記載したコンストラクトを哺乳動物の細胞の中に導入した。トランスフェクションの1日前、 $1 \times 10^5$ 個のHEK293T細胞 / ウエル (Sigma社) を24ウエルの皿の中でダルベッコ改変イーグル培地 (DMEM) + 10% (vol/vol) ウシ胎仔血清 (Gibco社) と1% ペニシリン-ストレプトマイシン (Gibco社) の中に播種した。翌日、細胞が50~60%の集密度になったとき、ウエル1つにつき1.5  $\mu$ lのリポフェクタミン3000 (Thermo Scientific社) を製造者の指示に従って用い、500 ngのRGN発現プラスミドと500 ngの単一gRNA発現プラスミドを同時にトランスフェクトした。48時間増殖させた後、ゲノムDNA単離キット (Machery-Nagel社) を製造者の指示に従って用いて全ゲノムDNAを回収した。

#### 【0459】

20

次いで全ゲノムDNAを分析し、AurkB標的における編集率を求めた。PCR増幅とそれに続く増幅されたゲノム標的部位の分析に用いるためのオリゴヌクレオチドを生成させた (配列番号363と364)。すべてのPCR反応を、各プライマーを0.5  $\mu$ M含む20  $\mu$ Lの反応物の中で、10  $\mu$ Lの2 $\times$  Master Mix Phusion High-Fidelity DNAポリメラーゼ (Thermo Scientific社) を用いて実施した。PCR#1プライマー (配列番号363と364) を用いて各標的遺伝子を包含する大きなゲノム領域を最初に増幅した。そのとき利用したプログラムは、98 $^{\circ}$ C、1分間; [98 $^{\circ}$ C、10秒間; 62 $^{\circ}$ C、15秒間; 72 $^{\circ}$ C、5分間] を30サイクル; 72 $^{\circ}$ C、5分間; 12 $^{\circ}$ C、永遠にであった。

#### 【0460】

30

次いで各ガイドに特異的なプライマー (PCR#2プライマー) を用いてこのPCR反応物1マイクロリットルをさらに増幅した。そのとき利用したプログラムは、98 $^{\circ}$ C、1分間; [98 $^{\circ}$ C、10秒間; 67 $^{\circ}$ C、15秒間; 72 $^{\circ}$ C、30秒間] を35サイクル; 72 $^{\circ}$ C、5分間; 12 $^{\circ}$ C、永遠にであった。PCR#2のためのプライマーは、Illumina シークエンシングのためのNextera Read 1 Transposase Adapterオーバーハング配列とNextera Read 2 Transposase Adapterオーバーハング配列を含んでいる。

#### 【0461】

40

第2のPCR増幅の後、PCRクリナップキット (Zymo社) を製造者の指示に従って用いてDNAをクリーンにした後、水の中で溶離させた。200~500 mgの精製されたPCR2#産物を20  $\mu$ Lの反応物の中で2  $\mu$ Lの10 $\times$  NEBバッファ2および水と組み合わせ、アニールしてヘテロ二重鎖DNAを形成した。そのとき利用したプログラムは、95 $^{\circ}$ C、5分間; 95 $^{\circ}$ Cから85 $^{\circ}$ Cへ2 $^{\circ}$ /秒の速度で冷却; 85 $^{\circ}$ Cから25 $^{\circ}$ Cへ0.1 $^{\circ}$ /秒の速度で冷却; 12 $^{\circ}$ C、永遠にであった。アニーリングの後、5  $\mu$ LのDNAを無酵素対照として取り出し、1  $\mu$ LのT7エンドヌクレアーゼI (NEB社) を添加し、この反応物を37 $^{\circ}$ Cで1時間インキュベートした。インキュベーションの後、5 $\times$  FlashGelローディング染料 (Lonza社) を添加し、5  $\mu$ Lの各反応物と対照を、ゲル電気泳動を利用して2.2% アガロースFlashGel (Lonza社) によって分析した。ゲルを可視化した後、非相同末端結合 (NHEJ) の割合を、式: %NHEJイベント =  $100 \times [1 - (1 - \text{切断された割合})^{1/2}]$  を用いて求めた (ただし (切断された割合) は、(消化された産物の密度) / (消化された産物 + 消化されなかった親バンドの密度) と定義される)。

#### 【0462】

50

いくつかのサンプルについて、SURVEYOR (登録商標) を用いて哺乳動物の細胞におけ

る発現後の結果を分析した。トランスフェクション後の72時間にわたって細胞を37℃でインキュベートした後、ゲノムDNAを抽出した。ゲノムDNAは、QuickExtract DNA 抽出溶液（Epicentre社）を製造者のプロトコルに従って用いて抽出した。RGN標的部に隣接するゲノム領域をPCR増幅し、産物を、QiaQuickスピンカラム（Qiagen社）を製造者のプロトコルに従って用いて精製した。合計200～500 ngの精製されたPCR産物を1 μlの10×Taq DNAポリメラーゼPCRバッファ（Enzymatics社）および超純水と混合して最終体積を10 μlにし、再アニーリングプロセスを適用することでヘテロ二重鎖の形成が可能になるようにした：95℃で10分間、-2℃/秒の勾配で95℃から85℃へ；-0.25℃/秒の勾配で85℃から25℃へ；そして25℃に1分間保持。

#### 【0463】

再アニーリングの後、産物を、SURVEYOR（登録商標）ヌクレアーゼとSURVEYOR（登録商標）エンハンサS（Integrated DNA Technologies社）を製造者が推奨するプロトコルに従って用いて処理し、4～20% Novex TBE ポリアクリルアミドゲル（Life Technologies社）上で分析した。ゲルをSYBR Gold DNA染色（Life Technologies社）で10分間染色し、Gel Docゲルイメージングシステム（Bio-rad社）を用いて画像を取得した。定量は、相対バンド強度に基づいた。indel率は、式： $100 \times (1 - (1 - (b + c) / (a + b + c))^{1/2})$ によって求めた（ただしaは消化されなかったPCR産物の積分強度であり、bとcは各切断産物の積分強度である）。

#### 【0464】

それに加え、Illuminaオーバーハング配列を含有するPCR#2からの産物を用い、Illumina 16S Metagenomic Sequencing Libraryプロトコルに従ってライブラリを調製した。深いシーケンシングをサービス提供者（MOGene社）がIllumina Mi-Seqプラットフォーム上で実施した。典型的にはアンプリコン1つにつき250 bpのペアードエンドリードを200,000個（2×100,000個のリード）生成させる。CRISPResso（Pinello他、2016年 Nature Biotech、第34巻：695～697ページ）を利用してこれらのリードを分析し、編集率を計算した。出力のアラインメントをマニュアルキュレーションで実施し、挿入部位と欠失部位を確認するとともに、組み換え部位のマイクロ相同性部位を同定した。編集率を表6に示す。すべての実験をヒト細胞で実施した。「標的配列」は、遺伝子標的内で標的とされる配列である。それぞれの標的配列について、ガイドRNAは、相補的なRNA標的配列と、使用するRGNに依存する適切なsgRNAを含んでいた。実験の選択された内訳をガイドRNAごとに表7.1と表7.2に示す。

#### 【0465】

#### 【表6】

表6：AurkB遺伝子標的に関する全編集率

RGN	ガイド RNA の ID	標的配列（配列番号）	サンプルにおける全編集率	サンプルにおける欠失率	サンプルにおける挿入率
APG09106	830	371	0.55%	100%	
APG09106	831	372	0.60%	54%	46%
APG09106 T849R	830	371	2.97%	98%	2.00%
APG09106 T849R	831	372	2.36%	100%	

#### 【0466】

各ガイドについて、具体的な挿入と欠失を表7.1と表7.2に示す。これらの表では、標的配列は太字の大文字によって特定される。8量体PAM領域には二重下線が引かれており、認識される主要なヌクレオチドは太字にされている。挿入は小文字によって特定される。欠失は点線（---）で示されている。INDEL位置は、標的配列のPAM近位端から計算され、縁部が位置0である。位置は、縁部の標的側にある場合にはプラス（+）であり、縁部のPAM側にある場合にはマイナス（-）である。

#### 【0467】

【表 7 - 1】

表 7.1 : RGN APG09106を用いた場合のガイド831に関する具体的な挿入と欠失

ガイド	リード の数	%リ ード	INDEL 率(%)	タイプ	INDEL の位置	サイ ズ
GTCTGATTGCCTGTCGTTGCCCTCCAGATCA TGGAGGAGTTGGCAGA (配列番号 373)	92294	99.40				
GTCTGATTGCCTGTCGTTGCCCTCCCA----- ---AGGAGTTGGCAGA (配列番号 595)	263	0.28	54.22	欠失	+19	8
GTCTGATTGCCTGTCGTTGCCCTaagtgatt aagcattgtctcagagatttttGGAGGAGTTGGC AGA (配列番号 596)	222	0.24	45.77	欠失	+13	20

10

【 0 4 6 8 】

【表 7 - 2】

表 7.2 : APG09106 T849Rを用いた場合のガイド831に関する具体的な挿入と欠失

ガイド	リード の数	%リ ード	INDEL 率(%)	タイプ	INDEL の位置	サイ ズ
GTCTGATTGCCTGTCGTTGCCCTCCAGATCAT GGAGGAGTTGGCAGA (配列番号 373)	189881	97.64				
GTCTGATTGCCTGTCGTTGCCCT-----T GGAGGAGTTGGCAGA (配列番号 597)	602	0.309	13.129	欠失	+14	10
GTCTGATTGCCTGTCGTTGCCCTCCAGATC- GGAGGAGTTGGCAGA (配列番号 598)	394	0.202	8.593	欠失	+23	2
GTCTGATTGCCTGTCGTTGCCCTCCAGAT--- --AGGAGTTGGCAGA (配列番号 599)	399	0.205	8.702	欠失	+22	5
GTCTGATTGCCTGTCGTTGCCCTaTc-----T G--GGAGTTGGCAGA (配列番号 600)	379	0.194	8.266	欠失と変異	+16	10
GTCTGATTGCCTGTCGTTGCCCTC-----T GGAGGAGTTGGCAGA (配列番号 601)	350	0.179	7.633	欠失	+16	8
GTCTGAT-----T GGAGGAGTTGGCAGA (配列番号 602)	309	0.158	6.739	欠失	-1	26
GTCTGATTGCCTGTCGTTGCCCTC----- GGAGGAGTTGGCAGA (配列番号 603)	280	0.143	6.106	欠失	+16	9
GTCTGATTGCCTGTCGTTGCCCTCC-----a GGAGGAGTTGGCAGA (配列番号 604)	274	0.140	5.976	欠失と変異	+17	7
GTCTGATTGCCTGTCGTTGCC----- ---GGAGTTGGCAGA (配列番号 605)	251	0.129	5.474	欠失	+13	15
GTCTGATTGCCTGTCGTTGCC-----ATCAT GGAGGAGTTGGCAGA (配列番号 606)	250	0.128	5.452	欠失	+13	7
GTCTGATTGCCTGTCGTTGCCCTC-----CAT GGAGGAGTTGGCAGA (配列番号 607)	231	0.118	5.038	欠失	+16	6
GTCTGATTGCCTGTCGTTGCCCTCCCA----- -----GTACT (配 列番号 608)	218	0.112	4.754	欠失	+19	30
GTCTGATTGCCTGTCGTTGCCCT-----aATCtT GGAGGAGTTGGCAGA (配列番号 609)	206	0.105	4.492	欠失と変異	+14	5
GTCTGATTGCCTGTCGTTGCC-----Tggg ATGGAGGAGTTGGCAGA (配列番号 610)	162	0.083	3.533	欠失と変異	+13	8
GTCTGATTGCCTGTCGTTGCCCTC----- -----AGTTGGCAGA (配列番号 611)	158	0.081	3.446	欠失	+16	14
GTCTGATTGCCTGTCGTTGCCCT-----TCAT GGAGGAGTTGGCAGA (配列番号 612)	122	0.062	2.660	欠失	+14	7

20

30

40

【 0 4 6 9 】

実施例6：植物細胞における遺伝子編集活性の実証

【 0 4 7 0 】

本発明のRGNのRNA誘導型ヌクレアーゼ活性を、Li他(2013年)Nat. Biotech. 第31巻: 688~691ページから改変したプロトコルを用いて植物細胞で実証する。簡単に述べると、N末端SV40核局在化シグナルをコードする核酸配列に機能可能に連結された本発明のRGN(配列番号1、16、24、35、43、または50)の植物コドン最適化バージョンを一

50

過性形質転換ベクターの中の強力な構成的35Sプロモータの後ろにクローニングする。適切なPAM配列に隣接して植物PDS遺伝子内の1つ以上の部位を標的とするsgRNAを第2の一過性形質転換ベクターの中の植物U6プロモータの後ろにクローニングする。PEGを媒介とした形質転換を利用してこれら発現ベクターをニコチアナ・ベンサミアナの葉肉プロトプラストの中に導入する。形質転換されたプロトプラストを暗所で36時間までインキュベートする。DNeasy植物ミニキット（Qiagen社）を用いてゲノムDNAをプロトプラストから単離する。RGN標的の部位に隣接するゲノム領域をPCRで増幅し、産物を、QiaQuickスピンカラム（Qiagen社）を製造者のプロトコルに従って用いて精製する。合計200～500 ngの精製されたPCR産物を1 μlの10×Taq DNAポリメラーゼPCRバッファ（Enzymatics社）および超純水と混合して最終体積を10 μlにし、再アニーリングプロセスを適用することでヘテロ二重鎖の形成が可能になるようにする：95℃で10分間、-2℃/秒の勾配で95℃から85℃へ；-0.25℃/秒の勾配で85℃から25℃へ；そして25℃に1分間保持。

#### 【0471】

再アニーリングの後、産物を、SURVEYOR（登録商標）ヌクレアーゼとSURVEYOR（登録商標）エンハンサS（Integrated DNA Technologies社）を製造者の推奨されるプロトコルに従って用いて処理し、4～20%Novex TBEポリアクリルアミドゲル（Life Technologies社）上で分析する。ゲルをSYBR Gold DNA染色（Life Technologies社）で10分間染色し、Gel Docゲルイメージングシステム（Bio-rad社）を用いて画像を取得する。定量は、相対バンド強度に基づく。indel率は、式： $100 \times (1 - (1 - (b + c) / (a + b + c))^{1/2})$ によって求める（ただしaは消化されなかったPCR産物の積分強度であり、bとcは各切断産物の積分強度である）。

#### 【0472】

#### 実施例7：疾患標的の同定

#### 【0473】

臨床バリエーションのデータベースをNCBI ClinVarデータベースから取得した。これは、world wide webのNCBI ClinVarのウェブサイトを通じて入手することができる。病原性一塩基多型（SNP）をこのリストから同定した。ゲノム遺伝子座情報を利用して、各SNPと重複する領域にあるCRISPR標的と各SNPの周囲の領域にあるCRISPR標的を同定した。原因変異（「Cast Mut.」）を標的とする本発明のRGNと組み合わせた塩基編集を利用して修正することのできる一群のSNPが、表8に掲載されている。表8には、各疾患の1つの通称だけを掲載してある。「RS#」は、NCBIウェブサイトのSNPデータベースにおけるRSアクセッション番号に対応する。アレルIDは原因アレルのアクセッション番号に対応し、染色体アクセッション番号は、NCBIウェブサイトを通じて見られるアクセッション参照情報も提供している。表8には、各疾患について、掲載されているRGNに適したゲノム標的配列の情報も提示されている。標的配列の情報は、本発明の対応するRGNにとって必要なsgRNAを作製するためのプロトスペーサ配列も提供する。

#### 【0474】

10

20

30

40

50

【表 8 - 1】

表8：本発明のRGNにのっての疾患標的

疾患	RG#	RGN	Cast Mut	アレル ID	染色体アクセッション番号	遺伝子の記号	標的（配列番号）
ABCA4 関連疾患	1800553	APG00969	C>T	22927	NC_000001.10, NC_000001.11	ABCA4	91
ABCA4 関連疾患	1800553	APG03128	C>T	22927	NC_000001.10, NC_000001.11	ABCA4	92
シュタルガルト病 1	1800728	APG00969	A>G	98777	NC_000001.10, NC_000001.11	ABCA4	93
シュタルガルト病 1	1800728	APG09748, APG09106, APG02789	A>G	98777	NC_000001.10, NC_000001.11	ABCA4	94
糖原病 1A 型	1801175	APG09748, APG09106, APG02789	C>T	27037	NC_000017.10, NC_000017.11	G6PC	95
重症複合免疫不全症	3218716	APG00969	C>T	52071	NC_000014.8, NC_000014.9	MYH7	96
重症複合免疫不全症	3218716	APG09748, APG09106, APG02789	C>T	52071	NC_000014.8, NC_000014.9	MYH7	97
遺伝性がん易罹患性症候群	5030818	APG03128	C>T	17256	NC_000003.11, NC_000003.12	VHL	98
フェニルケトン尿症	5030851	APG00969	G>A	15628	NC_000012.11, NC_000012.12	PAH	99
フェニルケトン尿症	5030858	APG00969	G>A	15616	NC_000012.11, NC_000012.12	PAH	100
フェニルケトン尿症	5030858	APG09748, APG09106, APG02789	G>A	15616	NC_000012.11, NC_000012.12	PAH	101
高フェニルアラニン血症	5030860	APG00969	T>C	15632	NC_000012.11, NC_000012.12	PAH	102
高フェニルアラニン血症	5030860	APG03128	T>C	15632	NC_000012.11, NC_000012.12	PAH	103

10

20

30

40

50

【表 8 - 2】

疾患	RS#	RGN	Cast Mut	アレリ D	染色体アクセッション番号	遺伝子の記号	標的 (配列番号)
高フェニアラニン血症	5030860	AP609748, AP609106, AP602789	T>G	15632	NC_000012.11, NC_000012.12	PAH	104
CBS 欠損症	5742905	AP600969	A>G	15159	NC_000021.8, NC_000021.9	CBS	105
CBS 欠損症	5742905	AP603128	A>G	15159	NC_000021.8, NC_000021.9	CBS	106
先天性小頭症	11556217	AP600969	C>T	34125	NC_000011.9, NC_000011.10	DHCR7	107
$\alpha$ 1-アンチトリプシン欠損症	28929474	AP600969	C>T	33006	NC_000014.8, NC_000014.9	SERPINA1	108
$\alpha$ 1-アンチトリプシン欠損症	28929474	AP603128	C>T	33006	NC_000014.8, NC_000014.9	SERPINA1	109
$\alpha$ 1-アンチトリプシン欠損症	28929474	AP609748, AP609106, AP602789	C>T	33006	NC_000014.8, NC_000014.9	SERPINA1	110
肢体型筋ジストロフィー、2D 型	28933693	AP600969	C>T	24476	NC_000017.10, NC_000017.11	SGCA	111
肢体型筋ジストロフィー、2D 型	28933693	AP603128	C>T	24476	NC_000017.10, NC_000017.11	SGCA	112
限局性皮質異形成 I1 型	28934872	AP600969	G>A	27436	NC_000016.9, NC_000016.10	TSC2	113
高グロブリン血症 D-周期性発熱	28934897	AP600969	G>A	26968	NC_000012.11, NC_000012.12	MVK	114
高グロブリン血症 D-周期性発熱	28934897	AP603128	G>A	26968	NC_000012.11, NC_000012.12	MVK	115
MECP2 関連疾患	28934906	AP600969	G>A	26850	NC_000023.10, NC_000023.11	MECP2	116
MECP2 関連疾患	28934906	AP609748, AP609106, AP602789	G>A	26850	NC_000023.10, NC_000023.11	MECP2	117
MECP2 関連疾患	28935468	AP600969	G>A	26863	NC_000023.10, NC_000023.11	MECP2	118
MECP2 関連疾患	28935468	AP600771	G>A	26863	NC_000023.10, NC_000023.11	MECP2	119
封入体筋炎 2	28937594	AP600969	A>G	21064	NC_000009.11, NC_000009.12	GNE	120

10

20

30

40

50

【表 8 - 3】

疾患	RS#	RGN	Cast Mut	アレル ID	染色体アクセッション番号	遺伝子の記号	標的（配列番号）
封入体筋炎 2	28937594	APG03128	A>G	21064	NC_000009.11, NC_000009.12	GNE	121
封入体筋炎 2	28937594	APG09748, APG09106, APG02789	A>G	21064	NC_000009.11, NC_000009.12	GNE	122
封入体筋炎 2	28937594	APG00771	A>G	21064	NC_000009.11, NC_000009.12	GNE	123
先天性グリコシル化異常症	28939378	APG00969	C>T	19763	NC_000016.9, NC_000016.10	ALG1	124
家族性地中海熱	28940579	APG00969	A>G	17579	NC_000016.9, NC_000016.10	MEFV	125
家族性高コレステロール血症	28942080	APG00969	G>A	18735	NC_000019.9, NC_000019.10	LDLR	126
家族性高コレステロール血症	28942080	APG03128	G>A	18735	NC_000019.9, NC_000019.10	LDLR	127
MUTYH 関連ポリポーシス	34612342	APG00969	T>C	20332	NC_000001.10, NC_000001.11	MUTYH	128
MUTYH 関連ポリポーシス	36053393	APG00969	C>T	20333	NC_000001.10, NC_000001.11	MUTYH	129
MUTYH 関連ポリポーシス	36053393	APG03128	C>T	20333	NC_000001.10, NC_000001.11	MUTYH	130
心筋症	36211715	APG00969	C>T	29159	NC_000014.8, NC_000014.9	MYH7	131
心筋症	36211715	APG03128	C>T	29159	NC_000014.8, NC_000014.9	MYH7	132
フォン・ヴィレブランド病	41276738	APG00969	C>T	15335	NC_000012.11, NC_000012.12	VWF	133
フォン・ヴィレブランド病	41276738	APG03128	C>T	15335	NC_000012.11, NC_000012.12	VWF	134
フォン・ヴィレブランド病	41276738	APG09748, APG09106, APG02789	C>T	15335	NC_000012.11, NC_000012.12	VWF	135
乳がんおよび／または卵巣がん	41293455	APG00969	G>A	32714	NC_000017.10, NC_000017.11	BRCA1	136
乳がんおよび／または卵巣がん	41293455	APG09748, APG09106, APG02789	G>A	32714	NC_000017.10, NC_000017.11	BRCA1	137
乳がんおよび／または卵巣がん	41293465	APG00969	G>A	70268	NC_000017.10, NC_000017.11	BRCA1	138
乳がんおよび／または卵巣がん	41293465	APG03128	G>A	70268	NC_000017.10, NC_000017.11	BRCA1	139

10

20

30

40

50

【表 8 - 4】

疾患	RS#	RGN	Cast Mut	アレル ID	染色体アクセッション番号	遺伝子の記号	標的 (配列番号)
乳がんおよび/または卵巣がん	45580035	APG00969	G>T	67431	NC_000013.10, NC_000013.11	BRCA2	140
乳がんおよび/または卵巣がん	55770810	APG00969	G>A	70063	NC_000017.10, NC_000017.11	BRCA1	141
MECP2 関連疾患	61749721	APG00969	G>A	26868	NC_000023.10, NC_000023.11	MECP2	142
MECP2 関連疾患	61749721	APG03128	G>A	26868	NC_000023.10, NC_000023.11	MECP2	143
MECP2 関連疾患	61750240	APG00969	G>A	26854	NC_000023.10, NC_000023.11	MECP2	144
MECP2 関連疾患	61750240	APG03128	G>A	26854	NC_000023.10, NC_000023.11	MECP2	145
MECP2-関連疾患	61750240	APG09748, APG09106, APG02789	G>A	26854	NC_000023.10, NC_000023.11	MECP2	146
シュタルガルト病 1	61751374	APG00969	G>A	22933	NC_000001.10, NC_000001.11	ABCA4	147
シュタルガルト病 1	61751374	APG03128	G>A	22933	NC_000001.10, NC_000001.11	ABCA4	148
家族性地中海熱	61752717	APG00969	T>C	17577	NC_000016.9, NC_000016.10	MEFV	149
MEFV 関連疾患	61752717	APG09748, APG09106, APG02789	T>C	17577	NC_000016.9, NC_000016.10	MEFV	150
MEFV 関連疾患	61752717	APG00771	T>C	17577	NC_000016.9, NC_000016.10	MEFV	151
フェニルケトン尿症	62508698	APG03128	G>T	15619	NC_000012.11, NC_000012.12	PAH	152
乳がんおよび/または卵巣がん	62625307	APG00969	G>A	69596	NC_000017.10, NC_000017.11	BRCA1	153
乳がんおよび/または卵巣がん	62625307	APG03128	G>A	69596	NC_000017.10, NC_000017.11	BRCA1	154
乳がんおよび/または卵巣がん	62625307	APG09748, APG09106, APG02789	G>A	69596	NC_000017.10, NC_000017.11	BRCA1	155
乳がんおよび/または卵巣がん	62625308	APG00969	G>A	32710	NC_000017.10, NC_000017.11	BRCA1	156
乳がんおよび/または卵巣がん	62625308	APG03128	G>A	32710	NC_000017.10, NC_000017.11	BRCA1	157

10

20

30

40

50



【表 8 - 5】

疾患	RS#	RGN	Cast Mut	アレル ID	染色体アクセッション番号	遺伝子の記号	標的 (配列番号)
乳がんおよび／または卵巣がん	62625308	APG09748, APG09106, APG02789	G>A	32710	NC_000017.10, NC_000017.11	BRCA1	158
		APG09748, APG09106, APG02789	C>T	95218	NC_000003.11, NC_000003.12	MLH1	159
		APG09748, APG09106, APG02789	C>T	94826	NC_000002.11, NC_000002.12	MSH6	160
遺伝性がん易罹患性症候群	63749843	APG09748, APG09106, APG02789	C>T	94826	NC_000002.11, NC_000002.12	MSH6	161
		APG09748, APG09106, APG02789	C>T	96029	NC_000002.11, NC_000002.12	MSH2	162
		APG09748, APG09106, APG02789	C>T	96378	NC_000002.11, NC_000002.12	MSH2	163
遺伝性がん易罹患性症候群	63750636	APG09748, APG09106, APG02789	C>T	96378	NC_000002.11, NC_000002.12	MSH2	164
		APG09748, APG09106, APG02789	C>T	96378	NC_000002.11, NC_000002.12	MSH2	165
		APG09748, APG09106, APG02789	C>T	23992	NC_000001.10, NC_000001.11	CPT2	166
カルニチンパルミトイルトランスフェラーゼ II 欠損症	74315294	APG09748, APG09106, APG02789	C>T	23992	NC_000001.10, NC_000001.11	CPT2	167
		APG09748, APG09106, APG02789	C>T	23992	NC_000001.10, NC_000001.11	CPT2	168
		APG09748, APG09106, APG02789	C>T	22161	NC_000007.13, NC_000007.14	CFTR	169

10

20

30

40

50

【表 8 - 6】

疾患	RS#	RGN	Cast Mut	アレル ID	染色体アクセッション番号	遺伝子の記号	標的（配列番号）
RET 関連疾患	74799832	APG00969	T>C	28958	NC_000010.10, NC_000010.11	RET	170
UDP-グルコース-ヘキソース-1-リン酸ウリジリルトランスフェラーゼ欠損症	75391579	APG00969	A>G	18653	NC_000009.11, NC_000009.12	GALT	171
UDP-グルコース-ヘキソース-1-リン酸ウリジリルトランスフェラーゼ欠損症	75391579	APG03128	A>G	18653	NC_000009.11, NC_000009.12	GALT	172
UDP-グルコース-ヘキソース-1-リン酸ウリジリルトランスフェラーゼ欠損症	75391579	APG09748, APG09106, APG02789	A>G	18653	NC_000009.11, NC_000009.12	GALT	173
嚢胞性線維症	75527207	APG09748, APG09106, APG02789	G>A	22159	NC_000007.13, NC_000007.14	CFTR	174
難聴、X 染色体連鎖	76434661	APG00969	C>T	53916	NC_000013.10, NC_000013.11	GJB2	175
難聴、X 染色体連鎖	76434661	APG03128	C>T	53916	NC_000013.10, NC_000013.11	GJB2	176
アミロイド生成性トランスサイレチンアミロイドーシス	76992529	APG00969	G>A	28465	NC_000018.9, NC_000018.10	TTR	177
アミロイド生成性トランスサイレチンアミロイドーシス	76992529	APG03128	G>A	28465	NC_000018.9, NC_000018.10	TTR	178
アミロイド生成性トランスサイレチンアミロイドーシス	76992529	APG09748, APG09106, APG02789	G>A	28465	NC_000018.9, NC_000018.10	TTR	179
アミロイド生成性トランスサイレチンアミロイドーシス	76992529	APG0771	G>A	28465	NC_000018.9, NC_000018.10	TTR	180
嚢胞性線維症	77010898	APG00969	G>A	22168	NC_000007.13, NC_000007.14	CFTR	181
嚢胞性線維症	77010898	APG09748, APG09106, APG02789	G>A	22168	NC_000007.13, NC_000007.14	CFTR	182

10

20

30

40

50

【表 8 - 7】

疾患	RS#	RGN	Cast Mut	アレリ D	染色体アクセシ ン番号	遺伝子の 記号	標的 (配 列番号)
異染性白質ジストロフィー	80338815	AP600969	C>T	18090	NC_000022.10, NC_000022.11	ARSA	183
異染性白質ジストロフィー	80338815	AP609748, AP609106, AP602789	C>T	18090	NC_000022.10, NC_000022.11	ARSA	184
カウデン症候群 3	80338844	AP600969	C>T	21935	NC_000011.9, NC_000011.10	SDHD	185
カウデン症候群 3	80338844	AP603128	C>T	21935	NC_000011.9, NC_000011.10	SDHD	186
カウデン症候群 3	80338844	AP609748, AP609106, AP602789	C>T	21935	NC_000011.9, NC_000011.10	SDHD	187
カウデン症候群 3	80338844	AP600771	C>T	21935	NC_000011.9, NC_000011.10	SDHD	188
スミス-レムリー-オピッツ症候群	80338853	AP600969	G>A	21822	NC_000011.9, NC_000011.10	DHCR7	189
スミス-レムリー-オピッツ症候群	80338853	AP603128	G>A	21822	NC_000011.9, NC_000011.10	DHCR7	190
スミス-レムリー-オピッツ症候群	80338853	AP609748, AP609106, AP602789	G>A	21822	NC_000011.9, NC_000011.10	DHCR7	191
高チロシン血症	80338901	AP600969	G>A	26909	NC_000015.9, NC_000015.10	FAH	192
高チロシン血症	80338901	AP603128	G>A	26909	NC_000015.9, NC_000015.10	FAH	193
高チロシン血症	80338901	AP600771	G>A	26909	NC_000015.9, NC_000015.10	FAH	194
難聴、X 染色体連鎖	80338940	AP603128	C>T	32068	NC_000013.10, NC_000013.11	GJB2	195
難聴、X 染色体連鎖	80338945	AP600969	A>G	32055	NC_000013.10, NC_000013.11	GJB2	196
乳がんおよび／または卵巣がん	80356962	AP600969	C>T	70247	NC_000017.10, NC_000017.11	BRCA1	197
乳がんおよび／または卵巣がん	80356962	AP603128	C>T	70247	NC_000017.10, NC_000017.11	BRCA1	198
乳がんおよび／または卵巣がん	80356962	AP609748, AP609106, AP602789	C>T	70247	NC_000017.10, NC_000017.11	BRCA1	199

10

20

30

40

50

【表 8 - 8】

疾患	RS#	RGN	Cast Mut	アレル ID	染色体アクセッション番号	遺伝子の記号	標的 (配列番号)
乳がんおよび/または卵巣がん	80356969	AP603128	G>A	70213	NC_000017.10, NC_000017.11	BRCA1	200
	80357123	AP609748, AP609106, AP602789	G>A	70147	NC_000017.10, NC_000017.11	BRCA1	201
先天性遺伝子疾患	80358259	AP600969	A>G	18006	NC_000018.9, NC_000018.10	NPC1	202
先天性遺伝子疾患	80358259	AP609748, AP609106, AP602789	A>G	18006	NC_000018.9, NC_000018.10	NPC1	203
	80359212	AP600969	C>T	67494	NC_000013.10, NC_000013.11	BRCA2	204
乳がんおよび/または卵巣がん	80359212	AP609748, AP609106, AP602789	C>T	67494	NC_000013.10, NC_000013.11	BRCA2	205
	104886457	AP600969	G>A	27086	NC_000009.11, NC_000009.12	FANCC	206
ファンコニー貧血	104886457	AP609748, AP609106, AP602789	G>A	27086	NC_000009.11, NC_000009.12	FANCC	207
	104893915	AP600969	C>T	19128	NC_000005.9, NC_000005.10	SLC26A2	208
SLC26A2 関連疾患	104893915	AP603128	C>T	19128	NC_000005.9, NC_000005.10	SLC26A2	209
SLC26A2 関連疾患	104893915	AP609748, AP609106, AP602789	C>T	19128	NC_000005.9, NC_000005.10	SLC26A2	210
	104894313	AP600969	C>T	18816	NC_000011.9, NC_000011.10	TYR	211
眼皮膚白皮症	104894368	AP609748, AP609106, AP602789	C>T	29104	NC_000012.11, NC_000012.12	MYL2	212
心筋症							

10

20

30

40

50

【表 8 - 9】

疾患	RS#	RGN	Cast Mut	アレル ID	染色体アクセッション番号	遺伝子の記号	標的（配列番号）
難聴、X染色体連鎖	104894396	AP609748, AP609106, AP602789	C>T	32041	NC_000013.10, NC_000013.11	GJB2	213
		AP609748, AP609106, AP602789	C>T	20146	NC_000017.10, NC_000017.11	SGSH	214
		AP609748, AP609106, AP602789	C>T	20146	NC_000017.10, NC_000017.11	SGSH	215
先天性遺伝子疾患	104894635	AP609748, AP609106, AP602789	C>T	20146	NC_000017.10, NC_000017.11	SGSH	216
		AP609748, AP609106, AP602789	C>T	17588	NC_000016.9, NC_000016.10	MEFV	217
		AP609748, AP609106, AP602789	C>T	17588	NC_000016.9, NC_000016.10	MEFV	218
家族性地中海熱	104895097	AP609748, AP609106, AP602789	C>T	17588	NC_000016.9, NC_000016.10	MEFV	219
		AP609748, AP609106, AP602789	A>G	21124	NC_000009.11, NC_000009.12	ELP1	220
		AP609748, AP609106, AP602789	A>G	21124	NC_000009.11, NC_000009.12	ELP1	221
家族性自律神経失調症	111033171	AP609748, AP609106, AP602789	A>G	18235	NC_000007.13, NC_000007.14	SBDS	222
		AP609748, AP609106, AP602789	A>G	18235	NC_000007.13, NC_000007.14	SBDS	223
		AP609748, AP609106, AP602789	A>G	18235	NC_000007.13, NC_000007.14	SBDS	223
シュバツハマン症候群	113993993	AP609748, AP609106, AP602789	C>T	28535	NC_000015.9, NC_000015.10	POLG	224
		AP609748, AP609106, AP602789	C>T	28541	NC_000015.9, NC_000015.10	POLG	225
		AP609748, AP609106, AP602789	C>T	28541	NC_000015.9, NC_000015.10	POLG	226

10

20

30

40

50

【表 8 - 10】

疾患	RS#	RGN	Cast Mut	アレリ ID	染色体アクセション番号	遺伝子の記号	標的 (配列番号)
極長鎖アシル-CoA デヒドロゲナーゼ欠損症	113994167	APG00969	T>C	33877	NC_000017.10, NC_000017.11	ACADVL	227
極長鎖アシル-CoA デヒドロゲナーゼ欠損症	113994167	APG03128	T>C	33877	NC_000017.10, NC_000017.11	ACADVL	228
極長鎖アシル-CoA デヒドロゲナーゼ欠損症	113994167	APG00771	T>C	33877	NC_000017.10, NC_000017.11	ACADVL	229
糖原病	116987552	APG00969	G>A	17337	NC_000011.9, NC_000011.10	PYGM	230
糖原病	116987552	APG03128	G>A	17337	NC_000011.9, NC_000011.10	PYGM	231
糖原病	116987552	APG00771	G>A	17337	NC_000011.9, NC_000011.10	PYGM	232
RYR1 関連疾患	118192172	APG00969	C>T	28003	NC_000019.9, NC_000019.10	RYR1	233
RYR1 関連疾患	118192172	APG09748, APG09106, APG02789	C>T	28003	NC_000019.9, NC_000019.10	RYR1	234
神経セロイドリポフスチン症 2	119455955	APG09748, APG09106, APG02789	G>A	17682	NC_000011.9, NC_000011.10	TPP1	235
中鎖アシル-CoA デヒドロゲナーゼ欠損症	121434274	APG09748, APG09106, APG02789	G>A	18627	NC_000001.10, NC_000001.11	ACADM	236
家族性高コレステロール血症	121908026	APG00969	C>T	18725	NC_000019.9, NC_000019.10	LDLR	237
家族性高コレステロール血症	121908026	APG03128	C>T	18725	NC_000019.9, NC_000019.10	LDLR	238
原発性高シュウ酸尿症	121908529	APG00969	G>A	38436	NC_000002.11, NC_000002.12	AGXT	239
原発性高シュウ酸尿症	121908529	APG03128	G>A	38436	NC_000002.11, NC_000002.12	AGXT	240

10

20

30

40

50

【表 8 - 1 1】

疾患	RS#	RGN	Cast Mut	アレル ID	染色体アクセッション番号	遺伝子の記号	標的 (配列番号)
原発性高シュウ酸尿症	121908529	AP609748, AP609106, AP602789	G>A	38436	NC_000002.11, NC_000002.12	AGXT	241
心臓-顔-皮膚症候群	121908595	AP600969	A>G	28390	NC_000015.9, NC_000015.10	MAP2K1	242
心筋症	121908987	AP600969	C>T	21885	NC_000007.13, NC_000007.14	PRKAG2	243
心筋症	121908987	AP609748, AP609106, AP602789	C>T	21885	NC_000007.13, NC_000007.14	PRKAG2	244
カウデン症候群	121909219	AP600969	C>T	22852	NC_000010.10, NC_000010.11	PTEN	245
カウデン症候群	121909219	AP609748, AP609106, AP602789	C>T	22852	NC_000010.10, NC_000010.11	PTEN	246
FGFR3 関連疾患	121913482	AP600969	C>T	31371	NC_000004.11, NC_000004.12	FGFR3	247
FGFR3 関連疾患	121913482	AP603128	C>T	31371	NC_000004.11, NC_000004.12	FGFR3	248
心筋症	121913625	AP600969	G>A	29128	NC_000014.8, NC_000014.9	MYH7	249
心筋症	121913628	AP609748, AP609106, AP602789	C>T	29131	NC_000014.8, NC_000014.9	MYH7	250
低ホスファターゼ症	121918007	AP600969	G>A	28709	NC_000001.10, NC_000001.11	ALPL	251
メチルマロン酸血症	121918241	AP600969	C>T	16462	NC_000001.10, NC_000001.11	MMACHC	252
メチルマロン酸血症	121918241	AP603128	C>T	16462	NC_000001.10, NC_000001.11	MMACHC	253
メチルマロン酸血症	121918241	AP600771	C>T	16462	NC_000001.10, NC_000001.11	MMACHC	254
先天性遺伝子疾患	121918243	AP609748, AP609106, AP602789	G>A	16464	NC_000001.10, NC_000001.11	MMACHC	255

10

20

30

40

50

【表 8 - 1 2】

疾患	RS#	RGN	Cast Mut	アレル ID	染色体アクセッション番号	遺伝子の記号	標的 (配列番号)
PTPN11 関連疾患	121918457	AP600969	C>T	28370	NC_000012.11, NC_000012.12	PTPN11	256
PTPN11 関連疾患	121918457	AP609748, AP609106, AP602789	C>T	28370	NC_000012.11, NC_000012.12	PTPN11	257
PTPN11 関連疾患	121918457	AP600771	C>T	28370	NC_000012.11, NC_000012.12	PTPN11	258
B リンパ芽球性白血病／リンパ腫、ICD-0 サブタイプなし	121918459	AP600969	A>G	28372	NC_000012.11, NC_000012.12	PTPN11	259
若年性骨髄単球性白血病	121918462	AP609748, AP609106, AP602789	C>T	28373	NC_000012.11, NC_000012.12	PTPN11	260
若年性骨髄単球性白血病	121918466	AP603128	A>G	28379	NC_000012.11, NC_000012.12	PTPN11	261
若年性骨髄単球性白血病	121918466	AP609748, AP609106, AP602789	A>G	28379	NC_000012.11, NC_000012.12	PTPN11	262
若年性骨髄単球性白血病	121918466	AP600771	A>G	28379	NC_000012.11, NC_000012.12	PTPN11	263
ムコ多糖症 I 型	121965019	AP600969	G>A	26947	NC_000004.11, NC_000004.12	IDUA	264
ムコ多糖症 I 型	121965019	AP603128	G>A	26947	NC_000004.11, NC_000004.12	IDUA	265
ムコ多糖症 I 型	121965020	AP600969	C>T	26948	NC_000004.11, NC_000004.12	IDUA	266
ムコ多糖症 I 型	121965020	AP603128	C>T	26948	NC_000004.11, NC_000004.12	IDUA	267
ムコ多糖症 I 型	121965020	AP609748, AP609106, AP602789	C>T	26948	NC_000004.11, NC_000004.12	IDUA	268
神経セロイドリポフスチン症 1	137852700	AP600969	G>A	23943	NC_000001.10, NC_000001.11	PPT1	269

10

20

30

40

50



【表 8 - 1 3】

疾患	RS#	RGN	Cast Mut	アレル ID	染色体アクセッション番号	遺伝子の記号	標的 (配列番号)
神経セロイドリポフスチン症 1	137852700	APG09748, APG09106, APG02789	G>A	23943	NC_000001.10, NC_000001.11	PPT1	270
		137852944	G>A	19147	NC_000006.11, NC_000006.12	PKHD1	271
多発性嚢胞腎	137853007	APG09748, APG09106, APG02789	G>A	20631	NC_000022.10, NC_000022.11	CHEK2	272
		137854568	C>T	15837	NC_000005.9, NC_000005.10	APC	273
大腸がん	137854568	APG09748, APG09106, APG02789	C>T	15837	NC_000005.9, NC_000005.10	APC	274
		137854568	C>T	15837	NC_000005.9, NC_000005.10	APC	275
ブルガダ症候群	137854601	APG09748, APG09106, APG02789	C>T	24416	NC_000003.11, NC_000003.12	SCN5A	276
		137929307	G>A	171217	NC_000019.9, NC_000019.10	LDLR	277
家族性高コレステロール血症	137929307	APG09748, APG09106, APG02789	G>A	171217	NC_000019.9, NC_000019.10	LDLR	278
		141659620	G>A	21858	NC_000016.9, NC_000016.10	SPG7	279
症性対麻痺	180177035	APG09748, APG09106, APG02789	T>C	29012	NC_000007.13, NC_000007.14	BRAF	280
		180177035	T>C	29012	NC_000007.13, NC_000007.14	BRAF	281
心臓-顔-皮膚症候群	180177035	APG09748, APG09106, APG02789	T>C	29012	NC_000007.13, NC_000007.14	BRAF	282
		180177035	G>A	132139	NC_000016.10, NC_000016.9	PALB2	283

10

20

30

40

50

【表 8 - 1 4】

疾患	RS#	RGN	Cast Mut	アレリ D	染色体アクセッション番号	遺伝子の記号	標的 (配列番号)
家族性乳がん	180177083	AP609748, AP609106, AP602789	G>A	132139	NC_000016.10, NC_000016.9	PALB2	284
MYBPC3 関連疾患	200411226	AP600969	C>T	174776	NC_000011.9, NC_000011.10	MYBPC3	285
MYBPC3 関連疾患	200411226	AP603128	C>T	174776	NC_000011.9, NC_000011.10	MYBPC3	286
MYBPC3 関連疾患	200411226	AP609748, AP609106, AP602789	C>T	174776	NC_000011.9, NC_000011.10	MYBPC3	287
MYBPC3 関連疾患	200411226	AP600771	C>T	174776	NC_000011.9, NC_000011.10	MYBPC3	288
RYR1 関連疾患	200563280	AP600969	C>T	169564	NC_000019.9, NC_000019.10	RYR1	289
RYR1 関連疾患	200563280	AP609748, AP609106, AP602789	C>T	169564	NC_000019.9, NC_000019.10	RYR1	290
心筋症	371898076	AP600969	C>T	52045	NC_000014.8, NC_000014.9	MYH7	291
乳がん-卵巣がん、家族性 4	387906843	AP600969	G>A	39241	NC_000017.10, NC_000017.11	RAD51D	292
乳がん-卵巣がん、家族性 4	387906843	AP603128	G>A	39241	NC_000017.10, NC_000017.11	RAD51D	293
MYBPC3 関連疾患	387907267	AP600969	G>A	45725	NC_000011.9, NC_000011.10	MYBPC3	294
MYBPC3 関連疾患	387907267	AP603128	G>A	45725	NC_000011.9, NC_000011.10	MYBPC3	295
MYBPC3 関連疾患	387907267	AP609748, AP609106, AP602789	G>A	45725	NC_000011.9, NC_000011.10	MYBPC3	296
PTPN11 関連疾患	397507547	AP600969	A>G	49032	NC_000012.11, NC_000012.12	PTPN11	297
デスモイド疾患、遺伝性	397515734	AP609748, AP609106, AP602789	C>T	51418	NC_000005.9, NC_000005.10	APC	298

10

20

30

40

50

【表 8 - 15】

疾患	RS#	RGN	Cast Mut	アレル ID	染色体アクセッション番号	遺伝子の記号	標的 (配列番号)
マルファン症候群／ロイス-デイツ症候群／家族性胸部解離性大動脈瘤	397515757	APG00969	C>T	51454	NC_000015.9, NC_000015.10	FBN1	299
マルファン症候群／ロイス-デイツ症候群／家族性胸部解離性大動脈瘤	397515757	APG09748, APG09106, APG02789	C>T	51454	NC_000015.9, NC_000015.10	FBN1	300
MYBPC3 関連疾患	397516074	APG00969	C>T	51962	NC_000011.9, NC_000011.10	MYBPC3	301
MYBPC3 関連疾患	397516074	APG03128	C>T	51962	NC_000011.9, NC_000011.10	MYBPC3	302
免疫不全症 14	397518423	APG00969	G>A	94255	NC_000001.10, NC_000001.11	PIK3CD	303
免疫不全症 14	397518423	APG09748, APG09106, APG02789	G>A	94255	NC_000001.10, NC_000001.11	PIK3CD	304
先天性遺伝子疾患	398123009	APG09748, APG09106, APG02789	C>T	48180	NC_000011.9, NC_000011.10	PACS1	305
リンパ芽球性白血病／リンパ腫、ICD-0 サブタイプなし	529008617	APG03128	G>A	152318	NC_000001.10, NC_000001.11	MUTYH	306
リンパ芽球性白血病／リンパ腫、ICD-0 サブタイプなし	529008617	APG09748, APG09106, APG02789	G>A	152318	NC_000001.10, NC_000001.11	MUTYH	307
家族性乳がん	587780021	APG03128	G>A	133177	NC_000002.11, NC_000002.12	BARD1	308
家族性乳がん	587780021	APG09748, APG09106, APG02789	G>A	133177	NC_000002.11, NC_000002.12	BARD1	309

10

20

30

40

50

【表 8 - 1 6】

疾患	RS#	RGN	Cast Mut	アレル ID	染色体アクセッション番号	遺伝子の記号	標的 (配列番号)
マルファン症候群／ロイス-ディーツ症候群／家族性胸部解離性大動脈瘤	727503054	APG00969	A>G	175979	NC_000015.9, NC_000015.10	FBN1	310
マルファン症候群／ロイス-ディーツ症候群／家族性胸部解離性大動脈瘤	727503054	APG03128	A>G	175979	NC_000015.9, NC_000015.10	FBN1	311
家族性高コレステロール血症	746118995	APG00969	C>T	228192	NC_000019.9, NC_000019.10	LDLR	312
家族性高コレステロール血症	746118995	APG03128	C>T	228192	NC_000019.9, NC_000019.10	LDLR	313
家族性高コレステロール血症	746118995	APG09748, APG09106, APG02789	C>T	228192	NC_000019.9, NC_000019.10	LDLR	314
家族性高コレステロール血症	746118995	APG00771	C>T	228192	NC_000019.9, NC_000019.10	LDLR	315
家族性高コレステロール血症	765696008	APG00969	G>A	228162	NC_000019.10, NC_000019.9	LDLR	316
家族性高コレステロール血症	765696008	APG03128	G>A	228162	NC_000019.10, NC_000019.9	LDLR	317
家族性高コレステロール血症	769370816	APG00969	G>A	228176	NC_000019.10, NC_000019.9	LDLR	318
家族性高コレステロール血症	769370816	APG03128	G>A	228176	NC_000019.10, NC_000019.9	LDLR	319
家族性高コレステロール血症	769370816	APG09748, APG09106, APG02789	G>A	228176	NC_000019.10, NC_000019.9	LDLR	320
家族性高コレステロール血症	769370816	APG00771	G>A	228176	NC_000019.10, NC_000019.9	LDLR	321
遺伝性がん易罹患性症候群	786201042	APG00969	C>T	181998	NC_000002.12, NC_000002.11	MSH6	322

【 0 4 7 5】

実施例 8：ハーラー症候群の原因となる変異を標的とする

【 0 4 7 6】

以下に、ハーラー症候群（MPS-1とも呼ばれる）のための可能性のある治療法として、この疾患を持つ患者の大きな集団内でハーラー症候群の原因となる変異を修正するRNA誘導型塩基編集システムを用いた治療法を記述する。このアプローチは、単一のAAVベクターにパッケージして広い範囲の組織のタイプに送達することのできるRNA誘導型の塩基編集融合タンパク質を利用している。疾患のある遺伝子座を標的とするため、使用する正確な調節エレメントと塩基編集ドメインに応じ、塩基編集融合タンパク質と単一ガイドRNA

の両方をコードする単一のベクターを作製することも可能であろう。

【0477】

実施例8.1：理想的なPAMを有するRGNの同定

【0478】

遺伝子疾患MPS-1はリソソーム蓄積症であり、リソソームの中に硫酸ダルマタンと硫酸ヘパランが蓄積することを分子レベルでの特徴とする。この疾患は、一般に、 $\alpha$ -L-イズロニダーゼをコードするIDUA遺伝子（NCBI参照配列NG\_008103.1）の中の変異によって起こる遺伝性の遺伝子疾患である。この疾患は、 $\alpha$ -L-イズロニダーゼが欠損していることの結果である。北欧系の個人の研究で見いだされた最も一般的なIDUA変異はW402XとQ70Xであり、両方ともナンセンス変異である結果として翻訳が途中で終了する（Bunge他（1994年）、Hum. Mol. Genet、第3巻(6)：861～866ページ；参照によって本明細書に組み込まれている）。単一のヌクレオチドを元に戻すと野生型コード配列が回復し、その結果として遺伝子座の内在性調節機構によりタンパク質発現が制御されると考えられる。

10

【0479】

ヒトIdua遺伝子のW402X変異がMPS-1Hのケースの大きな割合を占めている。塩基エディタは、ガイドRNAのプロトスペーサ成分の結合部位よりも狭い配列ウインドウを標的とすることができるため、標的遺伝子座から特定の距離にPAM配列が存在することが、この戦略の成功にとって不可欠である。標的変異は、塩基編集タンパク質との相互作用中は露出した非標的鎖（NTS）上に存在していなければならないという制約と、RGNドメインのフットプリントは、PAMに近い領域へのアクセスを阻止することになるという制約があるため、アクセス可能な遺伝子座はPAMから10～30 bpであると考えられる。このウインドウ内にある他の近傍のアデノシン塩基の編集と突然変異誘発を回避するため、さまざまなリンカーをスクリーニングする。理想的なウインドウは、PAMから12～16 bpである。

20

【0480】

RGN APG00969は適合する1つのPAM配列を有する。APG00969は5'-nnARV-3'（配列番号7）というPAM配列を持ち、サイズがコンパクトであるため、単一のAAVベクターを通じて送達できる可能性がある。この送達アプローチは、他のアプローチと比べて多くの利点を与える（広い範囲の組織（肝臓、筋肉、CNS）へのアクセス、およびよく確立された安全性プロファイルと製造技術など）。

【0481】

30

化膿レンサ球菌からのCas9（SpyCas9）は、W402X遺伝子座の近くに存在するNGG（配列番号323）というPAM配列を必要とするが、SpyCas9のサイズが理由で単一のAAVベクターにパッケージすることはできないため、このアプローチの上記の利点が失われる。二重送達戦略を利用できるが（例えばRyu他（2018年）、Nat. Biotechnol.、第36巻(6)：536～539ページ；参照によって本明細書に組み込まれている）、作製が極めて複雑で大きなコストがかかることになる。それに加え、二重ウイルスベクター送達は、遺伝子修正の効率を顕著に低下させる。なぜなら所与の細胞内で成功する編集は、両方のベクターを感染させることと、この細胞の中で融合タンパク質を組み立てることを必要とするからである。

【0482】

40

黄色ブドウ球菌からの一般に用いられているCas9オルソログ（SauCas9）はSpyCas9と比べてサイズがかなり小さいが、より複雑なPAMを必要とする（NGRRT（配列番号324））。しかしこの配列は、原因遺伝子座の塩基編集に役立つと予想される範囲内にはない。

【0483】

実施例8.2：RGN融合コンストラクトとsgRNA配列

【0484】

以下のドメイン、すなわち1）DNA切断活性を不活化する変異（「デッド」または「ニックアーゼ」）があるRGNドメインと；2）塩基編集に役立つアデノシンデアミナーゼを有する融合タンパク質をコードするDNA配列を、分子生物学の標準的な技術を利用して作製

50

する。下記の表（表9）に記載されているコンストラクトは、塩基編集活性ドメインを有する融合タンパク質（この実施例では、デッドRGN APG00969（配列番号327）のN末端に機能可能に融合したAPG02312（配列番号325））の変異したバリエーション）を含んでいる。塩基編集活性ドメインとして、本発明の任意のアデノシンデアミナーゼが可能であろう（例えば配列番号514または572～584など）。本分野では、RGNのC末端に塩基編集酵素を有する融合タンパク質も作製できることが知られている。それに加え、融合タンパク質のRGNと塩基エディタは、典型的には、リンカーアミノ配列によって隔てられている。本分野では、標準的なリンカーの長さは15～30個の範囲のアミノ酸であることが知られている。さらに本分野では、RGNと塩基編集酵素の間のいくつかの融合タンパク質は、塩基編集効率を向上させることのできる少なくとも1つのウラシルグリコシラーゼ阻害剤（UGI）ドメイン（配列番号570）も含むことができることが知られている（アメリカ合衆国特許第10,167,457号；参照によって本明細書に組み込まれている）。したがって融合タンパク質は、RGN APG00969またはそのバリエーション、アデノシンデアミナーゼ、および場合によっては少なくとも1つのUGIを含むことができる。

【0485】

【表9】

表9：RNAを標的とする塩基編集のためのコンストラクト

配列番号	コンストラクト	RGN	デッド（D）またはニックナーゼ（N）	塩基エディタ	リンカー（配列番号）
326	Nuc-ADAT-リンカー-dAPG00969-リンカー-SV40	APG00969	D	ADAT	546

【0486】

RGNがアクセスできる編集部位はPAM配列によって決まる。RGNを塩基編集ドメインと組み合わせるとき、編集の標的にされる残基は非標的鎖（NTS）の上に存在していなければならない。なぜならRGNが遺伝子座に会合している間は、NTSは一本鎖だからである。多数のヌクレアーゼとそれに対応するガイドRNAを評価することで、この特定の遺伝子座のための最適な遺伝子編集ツールを選択することが可能になる。ヒトIdua遺伝子の中で上記のコンストラクトが標的にできる可能性のあるいくつかのPAM配列は、W402X変異の原因となる変異ヌクレオチドの近傍にある。1）疾患遺伝子座の位置で非コードDNA鎖と相補的な「スペーサ」と、2）ガイドRNAをRGNと会合させるのに必要なRNA配列を含有するガイドRNA転写産物をコードする配列も作製する。このようなsgRNAは例えば配列番号356によってコードすることができる。当業者が考案することのできるこのsgRNAまたは同様のsgRNAについて、上記の塩基エディタ、または異なるRGNデアミナーゼ融合体を持つ塩基エディタを興味ある遺伝子座に向かわせる効率を評価することができる。

【0487】

実施例8.3：ハーラー疾患の患者からの細胞における活性を調べるアッセイ

【0488】

遺伝子型戦略を検証し、上記のコンストラクトを評価するため、ハーラー疾患の患者からの線維芽細胞を使用する。実施例5に記載されているベクターと同様にして、融合タンパク質コード配列とsgRNAコード配列の上流の適切なプロモータを含有するベクターを、ヒト細胞でこれらを発現させるために設計する。ヒト細胞の中で高レベルに発現することが知られているか、線維芽細胞の中で特異的によく発現することのできるプロモータとそれ以外のDNAエレメント（例えばエンハンサまたはターミネータ）も使用できることが認識されている。標準的な技術（例えば実施例5に記載されているのと同様のトランスフェクション）を利用してベクターを線維芽細胞にトランスフェクトする。あるいは電気穿孔を利用することができる。細胞を1～3日間培養する。標準的な技術を利用してゲノムDNA（gDNA）を単離する。編集効率は、精製されたgDNAに関するqPCR遺伝子型決定アッセイおよび/または次世代シーケンシングによって求める。これについては下にさらに詳しく記載する。

## 【0489】

Taqman (商標) qPCR分析では、野生型アレルと変異アレルに対して特異的なプローブを用いる。これらプローブは蛍光体を担持しており、これら蛍光体は、そのスペクトル励起特性および/または発光特性により、qPCR装置を用いて分離される。PCR用のプライマーとプローブを含有する遺伝子型決定キットは、市販のものを入手する(すなわちSNP ID rs121965019のためのThermo Fisher Taqman (商標) SNP遺伝子型決定アッセイ ID C\_\_27862753\_10)か、設計することができる。設計したプライマーとプローブのセットの一例を表10に示す。

## 【0490】

## 【表10】

10

表10: RT-PCR用のプライマーとプローブ

名称	配列	配列番号
順増幅プライマー	5' -GACTCCTTCACCAAG-3'	328
逆増幅プライマー	5' -GTAGATCAGCACCG-3'	329
野生型プローブ	5' -CTCTGGGCCGAAGT-3'	330
W402X プローブ	5' -CTCTAGGCCGAAGT-3'	331

## 【0491】

編集実験の後、標準的な方法と上記のプライマーおよびプローブを用いてgDNAをqPCRで分析する。予想される結果を表11に示す。このインビトロシステムは、コンストラクトを適切に評価して、さらに研究するための編集効率が大きいコンストラクトを選択するのに利用することができる。このシステムは、W402X変異を持つ細胞および持たない細胞と比較して評価され、好ましくはこの変異に関してヘテロ接合であるいくつかの細胞と比較して評価されることになる。Ct値は、染料(Sybrグリーンなど)を用い、参照遺伝子、またはその遺伝子座の全増幅と比較されることになる。

20

## 【0492】

## 【表11】

表11: 予想されるqPCRの結果

遺伝子型	塩基エディタのトランスフェクション	予想されるPCRの結果
Idua <sup>WT/WT</sup>	なし	ホモ接合野生型
Idua <sup>WT/W402X</sup>	なし	ヘテロ接合: 50%野生型、50%W402X
Idua <sup>W402X/W402X</sup>	なし	ホモ接合 W402X
Idua <sup>W402X/W402X</sup>	あり	さまざま

30

## 【0493】

組織は次世代シーケンシングによって分析することもできる。下に示したようなプライマー結合部位(表12)、または当業者が同定可能な他の適切なプライマー結合部位を使用することができる。PCR増幅の後、Illumina 16S Metagenomic Sequencing Libraryのプロトコルに従い、Illumina Nextera XTオーバーハング配列を含有する産物からライブラリを調製する。深いシーケンシングをIllumina Mi-Seqプラットフォーム上で実施する。典型的には、アンプリコン1つ当たり250 bpのペアードエンドリードを200,000個(2×100,000個のリード)生成させる。CRISPResso (Pinello他、2016年)を利用してこれらのリードを分析し、編集率を計算する。出力アラインメントのマニュアルキュレーションを実施して挿入部位と欠失部位を確認するとともに、組み換え部位の位置のマイクロ相同性部位を同定する。

40

## 【0494】

50

## 【表 1 2】

表12：NGSプライマー結合部位

方向	配列	配列番号
順	5' -ACTTCCTCCAGCC-3'	332
逆	5' -GAACCCGGGCTTA-3'	333

## 【0 4 9 5】

トランスフェクトされた細胞と対照細胞の細胞ライセートのウエスタンブロッティングを抗IDUA抗体を用いて実施することによって完全長タンパク質の発現を確認し、基質である4-メチルウンベリフェリル $\alpha$ -L-イズロニドを用いた細胞ライセートの酵素活性アッセイによってこの酵素が触媒活性であることを確認する（Hopwood他、Clin.Chim. Acta（1979年）、第92巻(2)：257～265ページ；参照によって本明細書に組み込まれている）。これらの実験を実施し、元のIdua<sup>W402X/W402X</sup>細胞系（トランスフェクションなし）と、塩基編集コンストラクトとランダムなガイド配列をトランスフェクトされたIdua<sup>W402X/W402X</sup>細胞系と、野生型IDUAを発現する細胞系を比較する。

## 【0 4 9 6】

実施例8.4：マウスモデルにおける疾患治療の検証

## 【0 4 9 7】

この治療アプローチの効果を確認するため、類似アミノ酸にナンセンス変異を有するマウスモデルを使用する。このマウス系統は、Idua遺伝子（遺伝子ID：15932）の中に、ハーラー症候群の患者における変異と相同な変異に対応するW392X変異を有する（Bunge他（1994年）、Hum. Mol. Genet. 第3巻(6)：861～866ページ；参照によって本明細書に組み込まれている）。この遺伝子座は、ヒトにおける遺伝子座とは明確に異なるヌクレオチド配列を含んでいて、前の実施例に記載した塩基エディタを用いた修正に必要なPAM配列を欠いているため、異なる融合タンパク質を設計してヌクレオチドを修正する必要がある。この動物での疾患改善により、遺伝子送達ベクターがアクセスできる組織内にある変異を修正するこの治療アプローチの有効性を検証することができる。

## 【0 4 9 8】

この変異に関してホモ接合性であるマウスは、ハーラー症候群の患者と似た表現型の特徴を多数示す。上記の塩基編集RGN融合タンパク質（表9）とRNAガイド配列を、マウスでタンパク質の発現とRNAの転写を可能にする発現ベクターに組み込む。研究の設計を下記の表13に示す。この研究は、塩基編集融合タンパク質とRNAガイド配列を含む高用量の発現ベクターで治療した群、低用量の同じ発現ベクターで治療した群、塩基編集融合タンパク質またはRNAガイド配列を含まない発現ベクターで治療したモデルマウスである対照、および空にした同じベクターで治療した野生型マウスである第2の対照を含んでいる。

## 【0 4 9 9】

## 【表 1 3】

表13：マウスモデルにおけるゲノム編集実験

群	マウス系統	数	治療
1	Idua-W392X <sup>1</sup>	5 以上	低用量のベクター
2	Idua-W392X	5 以上	高用量のベクター
3	Idua-W392X	5 以上	ビヒクル
4	129/Sv（野生型）	5	ビヒクル

## 【0 5 0 0】

評価するエンドポイントに含まれるのは、体重、尿GAG排泄、血清IDUA酵素活性、興味ある組織内のIDUA活性、組織の病態、SNPの修正を確認するための興味ある組織の遺伝子型、および挙動評価と神経学的評価である。いくつかのエンドポイントは消失であるため、研究が終了する前に追加の群を追加して例えば組織の病態と組織IDUA活性を評価することができる。エンドポイントの追加例は、ハーラー症候群動物モデルを確立した公開さ



れている論文に見いだすことができる (Shull他 (1994年)、Proc. Natl. Acad. Sci. U.S.A.、第91巻(26): 12937~12941ページ; Wang他 (2010年)、Mol. Genet. Metab.、第99巻(1): 62~71ページ; Hartung他 (2004年)、Mol. Ther.、第9巻(6): 866~875ページ; Liu他 (2005年)、Mol. Ther.、第11巻(1): 35~47ページ; Clark eta他 (1997年)、Hum. Mol. Genet. 第6巻(4): 503~511ページ; これらのすべてが参照によって本明細書に組み込まれている)。

#### 【0501】

可能な1つの送達ベクターは、アデノ随伴ウイルス (AAV) を利用している。ベクターを作製するときに含めるのは、CMVエンハンサ (配列番号335) とCMVプロモータ (配列番号334)、または他の適切なエンハンサとプロモータの組み合わせ、場合によってはコザック配列が前に置かれていて3'末端の位置でターミネータ配列とポリアデニル化配列 (例えばLevitt, N.; Briggs, D.; Gil, A.; Proudfoot, N. J. 「効率的な合成ポリ(A)部位の定義」 Genes Dev. 1989年、第3巻(7)、1019~1025ページに記載されている最小配列) に機能可能に融合された塩基エディタ-dRGN融合タンパク質コード配列 (例えば上記のようなNuc-ADAT-リンカー-dAPG19748-リンカー-SV40) である。このベクターはさらに、単一ガイドRNAをコードしていてその5'末端の位置にはヒトU6プロモータ (配列番号336)、または小さな非コードRNAの作製に適した別のプロモータが機能可能に連結された発現カセットを含むとともに、AAVカプシドにパッケージするために必要な本分野で周知の逆転末端反復 (ITR) 配列をさらに含んでいる。作製とウイルスのパッケージングは、標準的な方法 (例えばアメリカ合衆国特許第9,587,250号に記載されている方法; 参照によって本明細書に組み込まれている) によって実施する。

#### 【0502】

他の可能なウイルスベクターに含まれるのはアデノウイルスベクターとレンチウイルスベクターであり、これらは一般に用いられていて、パッケージングの能力と条件が異なる同様のエレメントを含有すると考えられる。非ウイルス送達法も利用できる。それは例えば脂質ナノ粒子によって封止されたmRNAとsgRNA (Cullis, P. R.とAllen, T. M. (2013年)、Adv. Drug Deliv. Rev. 第65巻(1): 36~48ページ; Finn他 (2018年)、Cell Rep. 第22巻(9): 2227~2235ページ、両方とも参照によって本明細書に組み込まれている)、プラスミドDNAのハイドロダイナミック注入 (Suda TとLiu D、2007年、Mol. Ther. 第15巻(12): 2063~2069ページ、参照によって本明細書に組み込まれている)、またはsgRNAのリボ核タンパク質複合体で金ナノ粒子と会合しているもの (Lee, K.; Conboy, M.; Park, H. M.; Jiang, F.; Kim, H. J.; Dewitt, M. A.; Mackley, V. A.; Chang, K.; Rao, A.; Skinner, C.他、Nat. Biomed. Eng. 2017年、第1巻(11): 889~90ページ) である。

#### 【0503】

実施例8.5: ヒト化遺伝子座を有するマウスモデルにおける疾患の修正

#### 【0504】

ヒトの治療に使用できると考えられるのと同じ塩基エディタコンストラクトの効果を評価するには、W392の近傍のヌクレオチドを変化させてヒトのW402の周辺の配列に一致させたマウスモデルが必要とされる。これは多彩な技術によって実現することができ、その中には、RGNとHDR鋳型を使用してマウスの胚の中にある遺伝子座を切断して置き換えることが含まれる。

#### 【0505】

アミノ酸保存の程度が大きいので、表14に示されているように、マウス遺伝子座の中の大半のヌクレオチドを、サイレント変異を有するヒト配列のヌクレオチドに変えることができる。得られた操作されたマウスゲノムでコード配列が変化する塩基変化だけが、導入された終止コドンの後に生じる。

#### 【0506】

10

20

30

40

【表 14】

表14：ヒト化マウス遺伝子座を生じさせるヌクレオチド変異

	ヒト (W402X)		マウス (W392X)		ヒト化マウス	
特徴	ヌクレオチド (配列番号 337)	コードされるアミノ酸 (配列番号 614 と 615)	ヌクレオチド (配列 ID 番 338)	コードされるアミノ酸 (配列番号 616 と 617)	ヌクレオチド (配列番号 339)	コードされるアミノ酸 (配列番号 616 と 615)
プロトスパーサ	G	E	A	G	G	G
	G	E	G	E	G	E
	A		A		A	
	G		A		G	
	C	Q	C	Q	C	Q
	A		A		A	
	G		A		G	
	C	L	C	L	C	L
	T		T		T	
	C		C		C	
	T	終止	T	終止	T	終止
	A		A		A	
	G		G		G	
	G	A	G	A	G	A
	C		C		C	
	C		A		C	
	G	E	G	E	G	E
	A		A		A	
	A		G		A	
	G	V	G	V	G	V
	T		T		T	
	G		C		G	
	T	S	T	S	T	S
	C		C		C	
	G		A		G	
PAM、重要でない	C	Q	A	K	C	Q
	A		A		A	
	G		G		G	
	G	A	G	A	G	A
PAM、重要	C		C		C	
	C		T		C	

## 【0507】

このマウス系統の操作が終わると、同様の実験を実施例8.4に記載されているようにして実施する。

## 【0508】

実施例9：フリードライヒ運動失調症の原因となる変異を標的とする

## 【0509】

フリードライヒ運動失調症 (FRDA) を引き起こす3ヌクレオチド反復配列の増殖は、FXN遺伝子内の所定の遺伝子座 (FRDA不安定性領域と呼ばれる) に起こる。RNA誘導型ヌクレアーゼ (RGN) を用いてFRDA患者の細胞内の不安定性領域を切除することができる。このアプローチは、1) ヒトゲノム内のアレルを標的とするようにプログラムすることのできるRGNとガイドRNAの配列と；2) このRGNとガイドRNAの配列を送達する方法を必要とする。ゲノム編集に使用される多くのヌクレアーゼ (一般に用いられている化膿レンサ球菌からのCas9ヌクレアーゼ (SpCas9) など) は、アデノ随伴ウイルス (AAV) ベクターの中にパッケージするには大きすぎる。特に、SpCas9遺伝子とガイドRNAの長さに加え、機能的な発現カセットに必要とされる他の遺伝エレメントの長さを考慮すると、そのことがわかる。そのためSpCas9を用いるアプローチの実現可能性は小さい。

## 【0510】

本発明のコンパクトなRNA誘導型ヌクレアーゼ (特にAPG09748とAPG09106) は、F

RDA不安定性領域の切除に極めてよく適している。それぞれのRGNは、FRDA不安定性領域の近傍にPAMを必要とする。それに加え、これらRGNのそれぞれは、ガイドRNAとともにAAVベクターの中にパッケージすることができる。2つのガイドRNAをパッケージするには第2のベクターが必要となる可能性が大きい、それでもこのアプローチは、より大きなヌクレアーゼ（例えばSpCas9であり、そのタンパク質配列を2つのベクターに分割せねばならない可能性がある）で必要とされると考えられるベクターよりも有利である。

#### 【0511】

表15は、APG09748またはAPG09106をFRDA不安定性領域の5'フランクと3'フランクに向かわせるのに適したゲノム標的配列の位置のほか、ゲノム標的のためのsgRNAの配列を示している。RGNは、この遺伝子座の位置に来ると、FA不安定性領域を切除すると考

10

#### 【0512】

#### 【表15】

表15：RGNシステムのためのゲノム標的配列

ガイドの番号	FRDA不安定性領域に対する位置	ゲノム標的配列（配列番号）	sgRNA（配列番号）
1	5'	340	344
2	5'	341	345
3	3'	342	346
4	3'	343	347

20

#### 【0513】

実施例10：鎌形赤血球疾患の原因となる変異を標的とする

#### 【0514】

BCL11エンハンサ領域内の標的配列（配列番号348）は、胎児ヘモグロビン（HbF）を増加させて鎌形赤血球疾患の症状を治癒または緩和する機構を提供することができる。例えばゲノムワイドの関連研究により、BCL11AにおいてHbFのレベル上昇と関係する一群の遺伝子バリエーションが同定されている。これらのバリエーションは、BCL11A内の段階特異的で系列限定エンハンサ領域として機能する非コード領域内に見いだされるSNPの集合である。さらなる探索から、赤血球細胞でBCL11Aを発現させるのにこのBCL11Aエンハンサが必要とされることが明らかになった（Bauer他（2013年）Science 第343巻：253～257ページ；参照によって本明細書に組み込まれている）。このエンハンサ領域はBCL11A遺伝子のイントロン2の中に見いだされ、イントロン2の中では、（調節能力と関係するクロマチンの状態を示していることがしばしばある）3つのDNアーゼI高感受性領域が同定された。これら3つの領域は、BCL11Aの転写開始位置からの距離（単位はキロ塩基）に従って「+62」、「+58」、「+55」として同定された。これらエンハンサ領域は、長さが大まかに350個（+55）；550個（+58）；350個（+62）のヌクレオチドである（Bauer他、2013年）。

30

#### 【0515】

実施例10.1：好ましいRGNシステムを同定する

40

#### 【0516】

ここでは、BCL11AがHBB遺伝子座（成人のヘモグロビンの中で $\beta$ -グロビンを作るのに責任がある遺伝子）内の対応する結合部位に結合するのを妨害するRGNシステムを用いた $\beta$ -ヘモグロビン症の治療の可能性を記述する。このアプローチは、哺乳動物の細胞の中でより効率的なNHEJを利用する。それに加え、このアプローチは、生体内送達のため単一のAAVベクターの中にパッケージできる十分に小さなサイズのヌクレアーゼを利用する。

#### 【0517】

ヒトBCL11エンハンサ領域内のGATA1エンハンサモチーフ（配列番号348）は、成人のヒト赤血球の中でBCL11Aの発現を低下させると同時にHbFを再発現させるためにRNA誘導型ヌクレアーゼ（RGN）を用いて妨害するのに理想的な標的である（Wu他（2019年

50

) Nat Med 第387巻：2554ページ)。このGATA1部位を取り囲む遺伝子座においてAPG09748またはAPG09106と適合性のあるいくつかのPAM配列は容易にわかる。これらヌクレアーゼは5'-DTTN-3' (配列番号30) というPAM配列を持ち、サイズがコンパクトであるため、適切なガイドRNAとともに単一のAAVベクターまたはアデノウイルスベクターの中に入れて送達できる可能性がある。この送達アプローチは、他のアプローチと比べて多数の利点 (例えば造血幹細胞へのアクセス、およびよく確立された安全性プロファイルと製造技術) を与える。

#### 【0518】

化膿レンサ球菌からの一般に用いられているCas9ヌクレアーゼ (SpyCas9) は、5'-NGG-3' (配列番号323) というPAM配列を必要とし、そのいくつかはGATA1モチーフの近くに存在する。しかしSpyCas9のサイズが原因で単一のAAVベクターまたはアデノウイルスベクターにパッケージすることはできないため、このアプローチの上記の利点が消える。二重送達戦略を採用できるが、作製が極めて複雑で大きなコストがかかると考えられる。それに加え、二重ウイルスベクター送達は、遺伝子修正の効率を顕著に低下させる。なぜなら所与の細胞内で編集が成功するには両方のベクターを感染させる必要があるからである。

#### 【0519】

実施例5に記載されている発現カセットと同様にして、ヒトコドンが最適化されたAPG09748 (配列番号349) またはAPG09106 (配列番号360) をコードする発現カセットを作製する。RGN APG09748またはAPG09106のためのガイドRNAを発現する発現カセットも作製する。これらのガイドRNAは、1) BCL11Aエンハンサ遺伝子座 (標的配列) 内の非コードDNA鎖またはコードDNA鎖と相補的なプロトスペーサ配列と、2) ガイドRNAをRGNと会合させるのに必要なRNA配列を含んでいる。APG09748またはAPG09106が標的とする可能性があるいくつかのPAM配列がBCL11A GATA1エンハンサモチーフを取り囲んでいるため、可能性のあるいくつかのガイドRNAコンストラクトを作製し、BCL11A GATA1エンハンサ配列のロバスタな切断とNHEJを媒介とした妨害を生じさせる最良のプロトスペーサ配列を決定する。表16に提示されているsgRNAを用いてRGNをこの遺伝子座に向かわせるため、表16の標的ゲノム配列を評価する。

#### 【0520】

#### 【表16】

表16：APG09748を用いたBCL11A GATA1エンハンサ遺伝子座のための標的配列

ガイド	標的ゲノム配列 (配列番号)	sgRNA (配列番号)
1	350	353
2	351	354
3	352	355

#### 【0521】

APG09748またはAPG09106がBCL11Aエンハンサ領域を中断する挿入または欠失を生じさせる効率を評価するため、ヒト細胞系 (ヒト胚性腎細胞 (HEK細胞) など) を使用する。RGN発現カセット (例えば実施例5に記載されているもの) を含むDNAベクターを作製する。表16のガイドRNA配列をコードする配列を含む発現カセットを含む別のベクターも作製する。このような発現カセットはさらに、実施例5に記載されているようなヒトRNAポリメラーゼIII U6プロモータ (配列番号336) を含むことができる。あるいはRGNとガイドRNAの両方の発現カセットを含む単一のベクターを使用することができる。標準的な技術 (例えば実施例5に記載されている技術) を利用してこのベクターをHEK細胞に導入し、細胞を1~3日間にわたって培養する。この培養期間の後、ゲノムDNAを単離し、実施例5に記載されているように、T7エンドヌクレアーゼIによる消化および/または直接的なシーケンシングを利用して挿入または欠失の頻度を求める。

#### 【0522】

DNA内で標的BCL11A領域を包含する領域を、Illumina Nextera XTオーバーハング配

列を含有するプライマーを用いてPCRによって増幅する。これらのPCRアンプリコンで、T7エンドヌクレアーゼIによる消化を利用してNHEJの形成を調べるか、Illumina 16Sメタゲノムシーケンシングライブラリのプロトコルに従うライブラリの調製、または同様の次世代シーケンシング(NGS)ライブラリの調製を実施する。深いシーケンシングの後、生成したリードをCRISPRessoによって分析し、編集率を計算する。出力アラインメントのマニピュレーションを実施して挿入部位と欠失部位を確認する。この分析により、好ましいRGNとそれに対応する好ましいガイドRNA(sgRNA)が同定される。この分析により、APG09748とAPG09106の両方が同様に好ましい結果になる可能性がある。それに加え、この分析により、2つ以上の好ましいガイドRNAが存在すると判断される可能性、または表16のすべての標的ゲノム配列が同様に好ましいと判断される可能性がある。

10

#### 【0523】

実施例10.2：胎児ヘモグロビンの発現を調べるアッセイ

#### 【0524】

この実施例では、APG09748またはAPG09106が生成させてBCL11Aエンハンサ領域を中断する挿入または欠失を、胎児ヘモグロビンの発現に関して調べる。健康なヒトのドナーCD34<sup>+</sup>造血幹細胞(HSC)を使用する。実施例8.3に記載されているのと同様の方法を利用してこれらHSCを培養し、好ましいRGNと好ましいsgRNAをコードする領域を含む発現カセットを含むベクターを導入する。電気穿孔の後、確立されたプロトコル(例えばGiarratana他(2004年)Nat Biotechnology 第23巻:69~74ページ;参照によって本明細書に組み込まれている)を利用してこれらの細胞をインビトロで赤血球に分化させる。次いでHbFの発現を、抗ヒトHbF抗体を用いたウエスタンブロッティングを利用して測定するか、高性能液体クロマトグラフィ(HPLC)を通じて定量する。BCL11Aエンハンサ遺伝子座の中断が成功すると、ガイドなしでRGNだけが電気穿孔されたHSCと比べてHbFの産生が増加することが予想される。

20

#### 【0525】

実施例10.3：鎌形赤血球形成の減少を調べるアッセイ

#### 【0526】

この実施例では、APG09748またはAPG09106が生成させてBCL11Aエンハンサ領域を中断する挿入または欠失を、鎌形赤血球形成の減少に関して調べる。鎌形赤血球症を患っている患者からのドナーCD34<sup>+</sup>造血幹細胞(HSC)を用いる。実施例8.3に記載されているのと同様の方法を利用してこれらHSCを培養し、好ましいRGNと好ましいsgRNAをコードする領域を含む発現カセットを含むベクターを導入する。電気穿孔の後、確立されたプロトコル(例えばGiarratana他(2004年)Nat Biotechnology 第23巻:69~74ページ)を利用してこれらの細胞をインビトロで赤血球に分化させる。次いでHbFの発現を、抗ヒトHbF抗体を用いたウエスタンブロッティングを利用して測定するか、高性能液体クロマトグラフィ(HPLC)を通じて定量する。BCL11Aエンハンサ遺伝子座の中断が成功すると、ガイドなしでRGNだけが電気穿孔されたHSCと比べてHbFの産生が増加することが予想される。

30

#### 【0527】

分化したこれら赤血球にメタ重亜硫酸塩を添加することによって鎌形赤血球の形成が誘導される。顕微鏡を用いて鎌形赤血球と正常な赤血球の数をカウントする。鎌形赤血球の数は、APG09748またはAPG09106とsgRNAで処理した細胞では、処理しない細胞、またはRGNだけで処理した細胞と比べて少ないことが予想される。

40

#### 【0528】

実施例10.4：マウスモデルにおける疾患治療の検証

#### 【0529】

APG09748またはAPG09106を用いたBCL11A遺伝子座の中断の効果を評価するため、鎌形赤血球貧血の適切なヒト化マウスモデルを用いる。好ましいRGNと好ましいsgRNAをコードする発現カセットをAAVベクターまたはアデノウイルスベクターの中にパッケージ

50

ジする。特にアデノウイルスAd5/35型が、HSCを標的とするのに有効である。鎌形赤血球アレルを持つヒト化HBB遺伝子座を含有する適切なマウスモデル（例えばB6;FVB-Tg(LCR-HBA2,LCR-HBB\*E26K)53Hhb/JまたはB6.Cg-Hbatm1Paz Hbbtm1Tow Tg(HBA-HBBs)41Paz/HhbJ）を選択する。これらのマウスを顆粒球コロニー刺激因子だけで治療するか、プレリキサホルとの組み合わせで治療し、HBCを循環の中に移動させる。次に、RGNとガイドを有するプラスミドを担持するAAVまたはアデノウイルスを静脈内注射し、マウスを一週間回復させる。これらのマウスから得られた血液を、メタ重亜硫酸塩を用いてインビトロ鎌形赤血球アッセイで調べ、マウスを長期にわたって追跡して死亡率と造血機能をモニタする。RGNとガイドRNAを担持するAAVまたはアデノウイルスを用いた治療は、両方の発現カセットが欠けたウイルス、またはRGN発現カセットだけを担持するウイルスで治療したマウスと比べて鎌形赤血球を減少させ、死亡率を低下させ、造血機能を改善することが予想される。

10

【0530】

#### 実施例11：デアミナーゼの同定

【0531】

細菌培養物を標準的な実験室環境において液体培地の中で増殖させた。培養物を飽和するまで増殖させた（16～24時間）後、DNAを調製した。DNAは、洗浄剤を用いた溶解によって細菌細胞から抽出した後、シリカマトリックスに結合させ、エタノールパッファで洗浄した。精製されたDNAは、わずかにアルカリ性の水性パッファを用いてシリカマトリックスから溶離させた。

20

【0532】

シーケンシングのためのDNAの純度と濃度を分光測光法によって調べた。Nextera XTライブラリ調製キットを製造者のプロトコルに従って用いてシーケンシングライブラリを調製した。配列データは、HiSeq 2000上でIllumina HiSeq 2000システムのユーザーガイドプロトコルに従って生成させた。

【0533】

CLC Bio Assembly Cellソフトウェアパッケージを用いてシーケンシングのリードを組み立ててドラフトゲノムにした。組み立て後、遺伝子コールをいくつかの方法によって作り、得られた遺伝子配列を調べてデアミナーゼ遺伝子の新規なホモログを同定した。新規な遺伝子は、BLASTとドメイン組成によって同定した。触媒ドメインD/H/C-[X]-E-[X15-45]-P-C-[X2]-C（配列番号613）がすべての酵素で予測された。それに加え、NCBIデータベースで同定された配列（シーケンシングされた真核生物のゲノムからの仮想デアミナーゼ）も調べた。同定された268個のAPOBECドメインタンパク質は65%の相同性でクラスターを形成し、47個の候補がデアミナーゼ活性アッセイのために選択された。同定された392個のADATドメインタンパク質は65%の相同性でクラスターを形成しており、配列の長さに基づいてさらに選択された。125個のADATドメインタンパクがデアミナーゼ活性アッセイのために選択された。これらADATドメインタンパクはすべて、長さがアミノ酸220個未満である。

30

【0534】

表17は、47個の選択されたAPOBECドメインタンパク質と125個のADATドメインタンパク質を示している。配列番号とAPG ID（各ポリペプチドのための独自の同定コードを与える）が示されている。入手できる場合には、NCBIとUniprotのアクセッション番号が示されている。「ドメインの説明」は、デアミナーゼがADATドメインを持つかAPOBECドメインを持つかを示している。同定されたポリペプチド配列内のADATドメインまたはAPOBECドメインの始点と終点も示されている。

40

【0535】

【表 17 - 1】

表17： 同定されたデアミナーゼ

配列 番号	APG ID	NCBI アクセ ッション番号	Uniprot アクセ ッション番号	ドメインの説明	ドメイ ン始点	ドメイ ン終点
374	APG008 68	N/A	B2XR68_FELCA	APOBEC ドメイン	69	103
375	APG010 21	WP_003372040	N/A	APOBEC ドメイン	334	370
376	APG011 79	N/A	N/A	APOBEC ドメイン	137	170
377	APG011 80	N/A	AOA2K6U5H6_SA IBB	APOBEC ドメイン	56	90
378	APG015 27	N/A	AOA2I0LXZ8_CO LLI	APOBEC ドメイン	63	95
379	APG016 50	N/A	AOA3B3ZFB9_9G OBI	APOBEC ドメイン	212	243
380	APG016 89	N/A	AOA0A1X9Q9_ZE UCU	APOBEC ドメイン	71	102
381	APG022 07	N/A	N/A	APOBEC ドメイン	321	357
382	APG022 82	N/A	N/A	APOBEC ドメイン	131	165
383	APG023 16	WP_002598246	N/A	APOBEC ドメイン	311	347
384	APG024 72	N/A	AOA1V4JAP2_PA TFA	APOBEC ドメイン	54	89
385	APG028 10	N/A	N/A	APOBEC ドメイン	131	164
386	APG030 38	N/A	N/A	APOBEC ドメイン	136	170
387	APG032 37	N/A	AOA2I3GCB3_NO MLE	APOBEC ドメイン	262	275
388	APG032 60	N/A	N/A	APOBEC ドメイン	138	172
389	APG033 31	N/A	M3W3R0_FELCA	APOBEC ドメイン	54	88
390	APG035 26	N/A	G1RYY7_NOMLE	APOBEC ドメイン	263	299
391	APG036 83	N/A	AOA0KOMJ25_HU MAN	APOBEC ドメイン	58	81
392	APG038 57	N/A	U3JMS2_FICAL	APOBEC ドメイン	94	117
393	APG040 50	N/A	AOA340X469_LI PVE	APOBEC ドメイン	267	301
394	APG041 17	N/A	G1TLT9_RABIT	APOBEC ドメイン	69	103
395	APG046 13	N/A	AOA2R2X2H4_PT EAL	APOBEC ドメイン	74	108

10

20

30

40

50

【表 17 - 2】

配列 番号	APG ID	NCBI アクセ ッション番号	Uniprot アクセ ッション番号	ドメインの説明	ドメイ ン始点	ドメイ ン終点
396	APG052 00	N/A	AOA2R2X2I7_PT EVA	APOBEC ドメイン	74	108
397	APG052 41	N/A	ABEC1_MONDO	APOBEC ドメイン	60	95
398	APG057 31	WP_015392428	N/A	APOBEC ドメイン	316	352
399	APG058 40	N/A	G8GPV9_SAG0E	APOBEC ドメイン	74	108
400	APG058 74	N/A	AOA2R2X2I2_PT EVA	APOBEC ドメイン	84	124
401	APG061 19	N/A	ABC3G_LAGLA	APOBEC ドメイン	254	288
402	APG065 44	N/A	AOA218ULD2_9P ASE	APOBEC ドメイン	57	89
403	APG066 62	N/A	AOA2R2X2I3_PT EVA	APOBEC ドメイン	84	124
404	APG067 19	N/A	AOA2U3Y3M5_LE PWE	APOBEC ドメイン	61	96
405	APG070 92	N/A	N/A	APOBEC ドメイン	143	176
406	APG072 77	N/A	N/A	APOBEC ドメイン	130	163
407	APG072 80	N/A	AOA1S3FTE2_DI POR	APOBEC ドメイン	88	123
408	APG073 86	N/A	F6M3K7_MACMU	APOBEC ドメイン	271	305
409	APG076 74	N/A	E2RL86_CANLF	APOBEC ドメイン	53	87
410	APG077 74	N/A	AOA151P6M4_AL LMI	APOBEC ドメイン	62	97
411	APG083 60	N/A	AOA287AD63_PI G	APOBEC ドメイン	81	115
412	APG085 01	N/A	N/A	APOBEC ドメイン	315	351
413	APG086 16	N/A	AOA340WXU3_LI PVE	APOBEC ドメイン	156	190
414	APG092 60	N/A	AOA1S3T3Q2_SA LSA	APOBEC ドメイン	88	106
415	APG096 64	N/A	AOA2K6NP10_RH IRO	APOBEC ドメイン	73	107
416	APG096 88	N/A	G1TVM9_RABIT	APOBEC ドメイン	92	127
417	APG096 93	N/A	AOA2K5XQK6_MA NLE	APOBEC ドメイン	257	290
418	APG097 10	N/A	F7EWS7_RAT	APOBEC ドメイン	85	122
419	APG097	N/A	S4RNJ9_PETMA	APOBEC ドメイン	72	117

10

20

30

40

50



【表 17 - 3】

配列 番号	APG ID	NCBI アクセ ッション番号	Uniprot アクセ ッション番号	ドメインの説明	ドメイ ン始点	ドメイ ン終点
	39					
420	APG099 80	N/A	A0A2R2Z4D8_PT EAL	AP0BEC ドメイン	68	102
421	APG008 35	WP_006418834	N/A	ADAT ドメイン	64	97
422	APG009 21	WP_005583886	N/A	ADAT ドメイン	58	91
423	APG009 70	WP_013486656	N/A	ADAT ドメイン	58	91
424	APG009 90	WP_008706930	N/A	ADAT ドメイン	62	95
425	APG013 40	WP_009534123	N/A	ADAT ドメイン	89	122
426	APG014 99	N/A	A0A0V0J9J5_SC HSO	ADAT ドメイン	46	85
427	APG015 93	N/A	N/A	ADAT ドメイン	65	98
428	APG016 03	WP_005610988	N/A	ADAT ドメイン	71	104
429	APG016 12	N/A	N/A	ADAT ドメイン	60	92
430	APG017 55	N/A	H0V5A6_CAVPO	ADAT ドメイン	71	110
431	APG019 74	WP_011736131	N/A	ADAT ドメイン	64	97
432	APG021 73	WP_003322111	N/A	ADAT ドメイン	55	88
433	APG022 81	N/A	N/A	ADAT ドメイン	29	68
434	APG023 12	WP_016147568	N/A	ADAT ドメイン	53	86
435	APG023 34	WP_011244898	N/A	ADAT ドメイン	55	88
436	APG023 39	WP_005584387	N/A	ADAT ドメイン	51	84
437	APG024 10	WP_009215532	N/A	ADAT ドメイン	76	109
438	APG024 12	N/A	N/A	ADAT ドメイン	64	101
439	APG024 20	WP_015561774	N/A	ADAT ドメイン	59	92
440	APG024 42	WP_017549959	N/A	ADAT ドメイン	63	96
441	APG025 91	N/A	A0A2P8YEP2_BL AGE	ADAT ドメイン	50	89
442	APG026 00	WP_005358896	N/A	ADAT ドメイン	76	109

10

20

30

40

50

【表 17 - 4】

配列 番号	APG ID	NCBI アクセ ッション番号	Uniprot アクセ ッション番号	ドメインの説明	ドメイ ン始点	ドメイ ン終点
443	APG027 51	N/A	N/A	ADAT ドメイン	57	90
444	APG027 86	WP_005997489	N/A	ADAT ドメイン	55	88
445	APG028 13	WP_013282182	N/A	ADAT ドメイン	76	109
446	APG030 10	WP_007203795	N/A	ADAT ドメイン	56	89
447	APG030 46	N/A	N/A	ADAT ドメイン	59	91
448	APG030 93	WP_013270915	N/A	ADAT ドメイン	53	86
449	APG031 10	WP_009677004	N/A	ADAT ドメイン	67	100
450	APG031 20	WP_015923794	N/A	ADAT ドメイン	53	86
451	APG031 40	WP_006701669	N/A	ADAT ドメイン	61	94
452	APG032 24	WP_005353279	N/A	ADAT ドメイン	54	87
453	APG033 36	N/A	N/A	ADAT ドメイン	57	90
454	APG033 90	N/A	ROMCR5_NOSB1	ADAT ドメイン	51	82
455	APG034 67	N/A	N/A	ADAT ドメイン	10	43
456	APG034 68	WP_010285589	N/A	ADAT ドメイン	53	86
457	APG034 74	WP_009643257	N/A	ADAT ドメイン	53	86
458	APG035 42	WP_005838301	N/A	ADAT ドメイン	51	84
459	APG035 57	WP_013497948	N/A	ADAT ドメイン	55	88
460	APG036 05	N/A	AOA1D1XUG0_9A RAE	ADAT ドメイン	61	103
461	APG036 91	WP_012446504	N/A	ADAT ドメイン	57	90
462	APG038 56	WP_019678962	N/A	ADAT ドメイン	52	85
463	APG038 76	N/A	AOA179V3P5_BL AGS	ADAT ドメイン	66	99
464	APG039 80	WP_015537265	N/A	ADAT ドメイン	52	85
465	APG040 36	WP_008982263	N/A	ADAT ドメイン	52	85
466	APG042	N/A	MORA73_RAT	ADAT ドメイン	71	107

10

20

30

40

50

【表 17 - 5】

配列 番号	APG ID	NCBI アクセ ッション番号	Uniprot アクセ ッション番号	ドメインの説明	ドメイ ン始点	ドメイ ン終点
	73					
467	APG042 83	WP_005999210	N/A	ADAT ドメイン	46	79
468	APG043 38	WP_007506011	N/A	ADAT ドメイン	56	89
469	APG044 30	WP_009220054	N/A	ADAT ドメイン	53	86
470	APG044 36	WP_006908309	N/A	ADAT ドメイン	67	100
471	APG045 14	N/A	N/A	ADAT ドメイン	57	91
472	APG045 71	WP_004636848	N/A	ADAT ドメイン	55	88
473	APG047 58	N/A	N/A	ADAT ドメイン	56	98
474	APG047 88	N/A	A0A1C7M4C9_GR IFR	ADAT ドメイン	67	106
475	APG047 95	WP_016424318	N/A	ADAT ドメイン	53	86
476	APG048 15	N/A	N/A	ADAT ドメイン	43	75
477	APG048 52	N/A	N/A	ADAT ドメイン	69	102
478	APG049 68	WP_012157903	N/A	ADAT ドメイン	53	86
479	APG050 22	WP_006785792	N/A	ADAT ドメイン	52	85
480	APG052 11	WP_002441139	N/A	ADAT ドメイン	53	86
481	APG053 14	WP_008301499	N/A	ADAT ドメイン	66	99
482	APG053 20	WP_004832879	N/A	ADAT ドメイン	53	86
483	APG054 15	WP_013171002	N/A	ADAT ドメイン	58	91
484	APG054 61	WP_008909404	N/A	ADAT ドメイン	52	85
485	APG054 86	WP_018131372	N/A	ADAT ドメイン	100	133
486	APG055 07	WP_009061460	N/A	ADAT ドメイン	56	89
487	APG055 73	N/A	N/A	ADAT ドメイン	55	94
488	APG055 82	N/A	A0A2U4CP64_TU RTR	ADAT ドメイン	71	110
489	APG056 94	WP_008590218	N/A	ADAT ドメイン	53	86

10

20

30

40

50

【表 17 - 6】

配列 番号	APG ID	NCBI アクセ ッション番号	Uniprot アクセ ッション番号	ドメインの説明	ドメイ ン始点	ドメイ ン終点
490	APG057 03	N/A	N/A	ADAT ドメイン	43	75
491	APG057 59	N/A	A0A1V4JLA3_PA TFA	ADAT ドメイン	57	96
492	APG058 44	WP_013248301	N/A	ADAT ドメイン	57	90
493	APG059 22	N/A	N/A	ADAT ドメイン	44	98
494	APG061 20	WP_005841945	N/A	ADAT ドメイン	54	87
495	APG062 19	N/A	N/A	ADAT ドメイン	4	53
496	APG062 25	N/A	N/A	ADAT ドメイン	46	79
497	APG063 34	WP_014116671	N/A	ADAT ドメイン	52	85
498	APG063 93	WP_005489257	N/A	ADAT ドメイン	53	86
499	APG065 72	WP_011339999	N/A	ADAT ドメイン	46	79
500	APG066 38	WP_006875450	N/A	ADAT ドメイン	52	85
501	APG066 45	WP_013656057	N/A	ADAT ドメイン	53	86
502	APG067 03	N/A	V8NRA1_OPHHA	ADAT ドメイン	24	63
503	APG067 76	WP_003147612	N/A	ADAT ドメイン	53	86
504	APG068 61	WP_010624847	N/A	ADAT ドメイン	59	92
505	APG069 51	WP_012809557	N/A	ADAT ドメイン	56	89
506	APG069 53	N/A	A0A0A9YIX6_LY GHE	ADAT ドメイン	54	88
507	APG069 73	WP_010630866	N/A	ADAT ドメイン	86	119
508	APG070 45	N/A	N/A	ADAT ドメイン	67	100
509	APG071 28	WP_005345192	N/A	ADAT ドメイン	61	94
510	APG071 64	WP_006525269	N/A	ADAT ドメイン	53	86
511	APG072 64	WP_013276874	N/A	ADAT ドメイン	55	88
512	APG073 31	WP_012939070	N/A	ADAT ドメイン	63	96
513	APG074	N/A	ADAT2_DANRE	ADAT ドメイン	70	109

10

20

30

40

50

【表 17 - 7】

配列 番号	APG ID	NCBI アクセ ッション番号	Uniprot アクセ ッション番号	ドメインの説明	ドメイ ン始点	ドメイ ン終点
	49					
514	APG074 58	WP_004035644	N/A	ADAT ドメイン	53	86
515	APG076 14	WP_009015861	N/A	ADAT ドメイン	57	90
516	APG076 67	N/A	N/A	ADAT ドメイン	61	94
517	APG077 06	N/A	Q16JL7_AEDAE	ADAT ドメイン	60	99
518	APG077 33	WP_011345539	N/A	ADAT ドメイン	52	85
519	APG078 61	N/A	AOA1X0QE95_9M ICR	ADAT ドメイン	54	86
520	APG079 00	WP_019468226	N/A	ADAT ドメイン	53	86
521	APG079 52	WP_017471628	N/A	ADAT ドメイン	52	85
522	APG079 64	N/A	N/A	ADAT ドメイン	57	87
523	APG079 75	N/A	N/A	ADAT ドメイン	71	137
524	APG080 10	WP_016475938	N/A	ADAT ドメイン	69	102
525	APG080 54	WP_015357810	N/A	ADAT ドメイン	51	84
526	APG083 11	WP_013485444	N/A	ADAT ドメイン	58	91
527	APG084 47	WP_015912349	N/A	ADAT ドメイン	52	85
528	APG084 94	WP_013779369	N/A	ADAT ドメイン	59	92
529	APG086 13	N/A	AOA1S9RJZ5_9E URO	ADAT ドメイン	69	102
530	APG087 66	N/A	AOA034WM03_BA CDO	ADAT ドメイン	71	110
531	APG087 99	WP_002835051	N/A	ADAT ドメイン	54	87
532	APG088 10	N/A	N/A	ADAT ドメイン	58	97
533	APG088 93	WP_014394804	N/A	ADAT ドメイン	53	86
534	APG089 39	WP_004098361	N/A	ADAT ドメイン	56	89
535	APG089 52	WP_006789980	N/A	ADAT ドメイン	53	86
536	APG089 55	WP_014969075	N/A	ADAT ドメイン	53	86

10

20

30

40

50

【表 17 - 8】

配列番号	APG ID	NCBI アクセッション番号	Uniprot アクセッション番号	ドメインの説明	ドメイン始点	ドメイン終点
537	APG09011	WP_011937002	N/A	ADAT ドメイン	71	104
538	APG09180	WP_007222730	N/A	ADAT ドメイン	57	90
539	APG09204	WP_006588115	N/A	ADAT ドメイン	57	90
540	APG09216	WP_019878409	N/A	ADAT ドメイン	56	89
541	APG09352	WP_015558044	N/A	ADAT ドメイン	54	87
542	APG09356	WP_018923319	N/A	ADAT ドメイン	54	87
543	APG09546	WP_009796283	N/A	ADAT ドメイン	61	94
544	APG09753	N/A	N/A	ADAT ドメイン	29	70
545	APG09981	N/A	N/A	ADAT ドメイン	83	116

10

## 【0536】

## 実施例12：デアミナーゼ活性を求めるアッセイ

## 【0537】

20

## 実施例12.1：推定デアミナーゼの選択

## 【0538】

表17に示されている選択されたタンパク質のコード配列を、細菌の中で発現させるのに最適なコドンにし、合成した後、本分野において周知で5'末端の位置でT7プロモータに機能可能に連結された標準的な細菌発現ベクターに導入した。

## 【0539】

## 実施例12.2：デアミナーゼ活性アッセイ

## 【0540】

デアミナーゼ活性アッセイは、Garibyan他（DNA Repair 第2巻：593～608ページ、2003年）に基づく。大腸菌のrpoB遺伝子に変異がある結果としてRNAポリメラーゼのサブユニットが変化することによって抗生剤リファンピシン（Rif）に対して耐性になる。

30

## 【0541】

実施例12.1に記載されているようにして、推定デアミナーゼをコードする細菌発現ベクターをT7 Express大腸菌細胞（NEBioLabs社）に導入した。これらの細胞は、ウラシルDNAグリコシラーゼ阻害剤（UGI；配列番号570）とカルベニシリン選択マーカーをコードする別の発現プラスミドも含有することができる。これらの細胞を飽和するまで増殖させた後、自己誘導培地（MagicMedia（商標）、Thermo Fisher Scientific社）の接種源として使用し、さらに5時間にわたって増殖させた。細胞を希釈し、UGI発現プラスミドも細胞に含まれているかどうかに応じ、カナマイシンを有するLBの上に、またはカナマイシンとカルベニシリンを有するLBの上に播種した。これらの希釈プレートを用いて全細胞をカウントした。同じ細胞を、リファンピシンを有するLBの上、またはリファンピシンとカルベニシリンを有するLBの上にも播種し、rpoB遺伝子に変異を導入することに成功した推定デアミナーゼ発現ベクターを同定した。

40

## 【0542】

リファンピシンを含有するLBプレート上で増殖させた細菌コロニーからデアミナーゼ発現ベクターを単離し、このアッセイを少なくとも2回繰り返した。確認後、細菌細胞をシーケンシングした。予想に反し、評価のために選択されたタンパク質を含有する47個のAPOBECドメインのうちで9個だけがデアミナーゼ活性を示した。評価のために選択されたタンパク質を含有する125個のADATドメインのどれもデアミナーゼ活性を示さなかった。これは、デアミナーゼ活性をポリペプチドのアミノ酸配列に基づいて予測することは

50

できず、その代わりに経験的に判断せねばならないことを示唆している。このアッセイによって同定された9個の活性なデアミナーゼに関する結果を表18に示す。リファンピシン耐性コロニーからのrpoB遺伝子をシーケンシングして誘導された変異を同定した。変異率は、それぞれの活性なデアミナーゼを含む耐性コロニーの数をコロニーの総数と比較することによって計算した。

【0543】

【表18】

表18：活性なデアミナーゼの変異率のNGS分析

APG の ID	配列番号	変異率
APG05241	397	42%
APG07280	407	36.90%
APG09260	414	35.30%
APG08360	411	36.90%
APG09980	420	30.50%
APG07386	408	36.80%
APG09688	416	19.20%
APG05840	399	25.80%
APG02316	383	48.40%

【0544】

実施例13：細菌細胞における塩基編集活性

【0545】

その後の活性スクリーニングにおいて、APG00868（配列番号374）も活性なデアミナーゼであると同定された。同定された10個の活性なデアミナーゼのコード配列を、N末端にNLS（配列番号10）を含んでいてC末端の位置で表18の活性なデアミナーゼに機能可能に連結された、C末端の位置でリンカー配列（配列番号546）に機能可能に連結された、C末端の位置でRNA誘導型DNA結合タンパク質、すなわちヌクレアーゼ不活性なRNA誘導型ヌクレアーゼ（RGN）dAPG08290.1バリエーション（配列番号547）に機能可能に連結された、C末端の位置で第2のNLSに機能可能に連結された、C末端の位置でTEV部位（配列番号548）に機能可能に連結された、C末端の位置で10×His（配列番号594）タグに機能可能に連結された融合タンパク質を産生する発現カセットに導入した。実施例12で同定された選択されたデアミナーゼとAPG00868について、細菌細胞の中で標的とする塩基編集活性を調べた。

【0546】

この活性アッセイは実施例12と非常によく似ていた。しかしこれらの実験では、これらのデアミナーゼを不活性なRGNに連結し、rpoB遺伝子の特定の領域を標的にして狙ったCからTへの変異を導入することが可能になるようにした。それに加え、RGN-デアミナーゼ融合体を標的とするガイドRNAの発現が可能な発現カセットを含むベクターを作製した。4つの異なるガイドRNAをこれらの実験で使用した。第1のガイド（表19では「非標的」と記載）（配列番号549）は、RGN-デアミナーゼ融合体を、rpoB遺伝子ではないゲノム細菌DNA内領域へと導いた。標的1（配列番号550）は、rpoBタンパク質にR529C変異を導入すると考えられるrpoB遺伝子内領域であった。標的2（配列番号551）は、rpoBタンパク質にA532V変異を導入すると考えられるrpoB遺伝子内領域であり、標的3（配列番号552）は、rpoBタンパク質にQ513R変異を導入すると考えられるrpoB遺伝子内領域であった。標的1と2の望む変異は、GCペアからATペアへと塩基が編集された結果であると

考えられる。標的3の望む変異は、ATペアからGCペアへと塩基が編集された結果であると考えられる。他の可能な変異もこれらの標的の中に見いだすことができる。

#### 【0547】

融合タンパク質発現ベクターを、rpoB遺伝子上の興味ある位置を標的とするガイドRNAの発現が可能な発現カセットを含むベクターとともにT7 Express大腸菌細胞（NEBioLabs社）に導入した。これらの細胞を飽和するまで増殖させた後、自己誘導培地（MagicMedia（商標）、Thermo Fisher Scientific社）の接種源として使用し、さらに5時間にわたって増殖させた。細胞を希釈し、カナマイシンを有するLBの上に播種した。これらの希釈プレートを用いて全細胞をカウントした。同じ細胞を、リファンピシンを有するLBの上にも播種し、rpoB遺伝子の中に変異を有するコロニーを同定した。「非標的」と、標的デアミナーゼ-RGN融合タンパク質（「標的」）に関する変異率を計算し、表19に示してある。標的デアミナーゼ-RGN融合体の変異率の増加率（「%増加」）も、rpoB遺伝子を標的としないデアミナーゼ-RGN融合体と比較して表19に示されている。デアミナーゼとして機能することが知られている陽性対照である哺乳動物のAPOBECも含まれていた。

#### 【0548】

#### 【表19】

表19：標的デアミナーゼと非標的デアミナーゼの変異率

デアミナーゼ APG の ID	ガイド RNA のタイプ	標的（配列番号）	変異率	%増加
APG05241	非標的	549	$1.71 \times 10^{-8}$	—
APG05241	標的	550	$1.42 \times 10^{-7}$	8.31
APG05241	標的	551	$5.95 \times 10^{-9}$	0.35
APG05241	標的	552	$3.62 \times 10^{-8}$	2.12
APG08360	非標的	549	$4.00 \times 10^{-11}$	—
APG08360	標的	550	$4.74 \times 10^{-10}$	11.84
APG08360	標的	551	$2.38 \times 10^{-11}$	0.60
APG08360	標的	552	$4.70 \times 10^{-8}$	2175.00
APG00868	非標的	549	$1.82 \times 10^{-9}$	—
APG00868	標的	550	$1.43 \times 10^{-7}$	78.68
APG00868	標的	551	$3.79 \times 10^{-9}$	2.09
APG00868	標的	552	$1.18 \times 10^{-6}$	650.00
対照	非標的	549	$1.53 \times 10^{-8}$	—
対照	標的	550	$6.37 \times 10^{-6}$	417.36
対照	標的	551	$2.59 \times 10^{-7}$	16.98
対照	標的	552	$3.79 \times 10^{-7}$	24.86

#### 【0549】

実施例14：哺乳動物の細胞における塩基編集活性

#### 【0550】

同定された活性なデアミナーゼのコード配列を、哺乳動物の細胞の中で発現させるのに最適なコドンにした後、N末端にNLS（配列番号10）を含んでいてC末端の位置で3×FLAGタグ（配列番号11）に機能可能に連結された、C末端の位置で本発明のデアミナーゼに機能可能に連結された、C末端の位置でアミノ酸リンカー（配列番号546）に機能可能に連結された、C末端の位置でRNA誘導型DNA結合ポリペプチド、すなわちニッカーゼとして機能するように変異したRGN（nAPG07433.1；配列番号553）に機能可能に連結された、C末端の位置で第2のNLSに機能可能に連結された融合タンパク質を産生する発現カセットに導入した。それに加え、APG07386のN末端とC末端の断片（それぞれ、配列番号554としてのAPG07386-NTDと、配列番号555としてのAPG07386-CTD）を個別に発現カセットに導入し、各断片のデアミナーゼ-RGN融合体を産生させた。これら発現カセットを、それぞれ、哺乳動物の細胞の中でこの融合タンパク質を発現させることのできるベクターに導入した。ガイドRNAを発現してデアミナーゼ-RGN融合タンパク質を所定のゲノム位置へと向かわせることのできるベクターも作製した。これらガイドRNAは、デアミナーゼ-RGN融合タンパク質を、塩基編集の標的とされるゲノム配列へと導くことができる。



配列番号556～561は、調べたガイドRNAをコードしている。

【0551】

実施例14.1：デアミナーゼの効率と配列特異性

【0552】

上記のデアミナーゼ-RGN融合タンパク質とガイドRNAを発現することのできるベクターを、リポフェクションまたは電気穿孔を利用してHEK293T細胞にトランスフェクトした。リポフェクションでは、増殖培地（DMEM + 10%ウシ胎仔血清 + 1%ペニシリン/ストレプトマイシン）の中でのトランスフェクションの前日に細胞を $1 \times 10^5$ 細胞/ウエルで24ウエルのプレートに播種した。Lipofectamine（登録商標）3000試薬（Thermo Fisher Scientific社）を製造者の指示に従って用いて500 ngのデアミナーゼ-RGN融合体発現ベクターと1  $\mu$ gのガイドRNA発現ベクターをトランスフェクトした。電気穿孔では、Neon（登録商標）Transfection System（Thermo Fisher Scientific社）を製造者の指示に従って用いて細胞に電気穿孔した。

10

【0553】

リポフェクションまたは電気穿孔の24～48時間後、トランスフェクトされた細胞または電気穿孔された細胞からゲノムDNAを回収し、DNAをシーケンシングして分析し、標的とされる塩基編集変異の存在を探した。

【0554】

下記の表20は、各デアミナーゼについて、APG07386のC末端断片とN末端断片を含め、シチジン塩基の編集率を示している。番号の行は、標的とされるゲノム配列内のシチジン塩基の位置を、RGNのPAMを基準にして示している。各位置におけるCヌクレオチドの編集率は、多数の標的の平均値として示されている。標的の数（n）は、列の下それぞれの位置に掲載されている。標準偏差を括弧の中に入れて示してある。このアッセイでは、APG09980、APG07386-CTD、APG05840、APG05241、APG07280、APG09688、およびAPG00868が、少なくとも1個のシチジンの少なくともあるレベルのシチジン塩基編集活性を示す。

20

【0555】

30

40

50

【表 2 0 - 1】

表20：哺乳動物の細胞におけるCヌクレオチドの編集率

APG の ID	標的内の位置																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
APG05241	0% (0.01% (0.25% (0.17% (1.367% (0.4% (6.55% (6.4% (4.7% (5.9% (1.1% (0% (0.22.3% (0% (0) 12.5% (0% (0) 15% (0.4% (0.1% (1.225% (1.8) (1.8)																			
APG07280	0% (0.08% (0.18% (0.067% (0.12% (0% (0) 1.675% (6.45% (0.5% (1.15% (0% (0) 0% (0.13.9% (0% (0) 12.95% (0% (0) 7.5% (0.25% (0.1% (1.575% (2.36) (2.36)																			
APG09260	0% (0% (0) 0% (0) 0% (0) 0% (0) 0% (0) 0.075% (0.05% (0% (0) 0% (0) 0% (0) 0% (0) 0% (0) 0% (0) 0% (0) 0% (0) 0% (0) 0.1% (0% (0) (0.14)																			
APG08360	0% (0.02% (0.04% (0.167% (0.29% (0% (0) 0.025% (0% (0) 1.4% (0% (0) 0% (0) 0% (0) 0% (0) 0% (0) 0% (0) 4.65% (0% (0) 6.3% (0.7% (0.05% (1.75% (3.18) (3.18)																			
APG09980	0% (0.1% (0.22% (1.267% (1.36% (16% (0.12.875% (8.55% (6.9% (5.7% (0% (0) 0% (0) 0% (0.21.1% (3.4% (12.4% (0% (0) 11.1% (1.5% (0% (0) 1.375% (2.55) (2.55)																			
APG07386	0% (0% (0) 0% (0) 1.033% (1.79% (0% (0) 0.025% (0.05% (0% (0) 6.6% (0.25% (0.65% (0.92% (0% (0) 11.25% (3.2% (4.1% (1.7% (0% (0) 2.925% (4.22) (4.22)																			
APG07386-CTD	0% (0.2% (0.45% (2.233% (3.44% (1.3% (2.15% (1.95% (2.1% (2.35% (5.4% (0% (0.21.8% (0% (0) 23.3% (2.1% (8.8% (3.25% (0% (0) 1.125% (1.37) (1.37)																			
APG07386-NTD	0% (0% (0) 0% (0) 0% (0) 0% (0) 0.1% (0.175% (0.45% (0% (0) 0% (0) 0% (0) 0% (0) 0% (0) 0% (0) 0% (0) 0% (0) 0% (0) 0% (0) 0% (0) (0)																			
APG09688	0% (0.02% (0.04% (0.267% (0.46% (0% (0) 1.1% (9.5% (0.6% (0.2% (0% (0) 0% (0) 0% (0.21.9% (0% (0) 19.95% (0% (0) 10.6% (0.4% (0% (0) 1.325% (2.32) (2.32)																			
APG05840	0% (0.1% (0.22% (0.15% (0.29% (15.3% (10.475% (13.25% (7.1% (16.7% (1.9% (4% (0.19.1% (12.5% (0% (0) 18.9% (4.2% (29.2% (3.5% (0% (0) 2.825% (2.84) (2.84)																			
APG02316	0.1% (0.02% (0.04% (0% (0) 0.1% (0.075% (0.1% (0% (0) 0% (0) 0% (0) 0% (0) 0% (0.3.6% (0% (0) 0% (0) 0% (0) 0% (0) 0% (0) 0% (0) 0% (0) (0.05) (0.05)																			

10

20

30

40

50

【表 2 0 - 2】

APG の ID	標的内の位置																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
APG00868	0% (0.02%) (0.39)	0% (0.15%) (0.21)	0.967% (1.34)	---	29.4% (0)	22.85% (7.35)	8.4% (2.12)	7.4% (0)	4% (2.69)	7.8% (14.6%) (0.99)	0% (0)	4.6% (6.2%) (0)	0% (0)	0% (0)	9.3% (52.7%) (.8)	0% (0)	12.8% (12.8%) (0)	1.85% (1.63%) (1.7)	5.3% (5.3%) (1.7)	1.2% (1.45%) (45)
対照 1	0.1% (0) (0)	0.12% (0.22)	0% (0)	0.567% (0.9)	---	0.3% (0)	13.65% (5.84)	0% (19.02)	0% (0)	2.5% (3.11)	3.15% (4.17)	0% (0)	0.35% (0.9)	0% (0)	18.45% (24.25%) (0)	0% (0)	28.6% (28.6%) (0)	1.55% (1.63%) (1.98)	1.4% (1.4%) (3.39)	2.375% (2.375%) (3.39)
対照 2	0% (0.052%) (0.52)	0% (0)	0.7% (0.89)	---	37.2% (0)	12.75% (5.71)	10.1% (9.9)	1.8% (7.45%) (0)	0.95% (8.84%) (0.07)	0.95% (0.07)	40.6% (40.6%) (0)	2.8% (2.8%) (0)	0% (0)	14.5% (14.5%) (0)	3.6% (3.6%) (.71)	0% (0)	4.6% (4.6%) (0)	1% (0.57%) (0.92)	0.65% (0.65%) (0.92)	1.425% (1.425%) (0.98)
標的の数	n=1	n=5	n=2	n=3	n=0	n=1	n=4	n=2	n=1	n=2	n=2	n=1	n=1	n=1	n=2	n=1	n=1	n=2	n=2	n=4

【0556】

実施例 14.2：標的とされる塩基編集の蛍光アッセイ

【0557】

シトシンデアミナーゼを用いてH66のコドンの第1位をCからTに変化させてヒスチジン (CAT) から野生型チロシン (TAT) 残基へと戻せるようにするため、青色蛍光タンパク質 (BFP、配列番号562) へと蛍光をシフトさせるY66H変異を含有する増強型緑色蛍光タンパク質 (EGFP) を有するベクターを構成した。CからTへの変換が成功するとEGFPが発現してそれを定量することができる。デアミナーゼ-RGN融合タンパク質をY66H変異の周辺領域に向かわせるガイドRNA (配列番号563) を発現することのできる第2のベクタ

10

20

30

40

50

ーも作製した。

【0558】

このBFP EGFPレポータベクターを、デアミナーゼ-RGN融合タンパク質とガイドRNAを発現することのできるベクターとともに、リポフェクションまたは電気穿孔を利用してHEK293T細胞にトランスフェクトした。リポフェクションでは、増殖培地（DMEM + 10 % ウシ胎仔血清 + 1 % ペニシリン / ストレプトマイシン）の中でのトランスフェクションの前日に細胞を  $1 \times 10^5$  細胞 / ウエルで24ウエルのプレートに播種した。Lipofectamine（登録商標）3000試薬（Thermo Fisher Scientific社）を製造者の指示に従って用い、BFPレポータベクター、デアミナーゼ-RGN融合発現ベクター、およびガイドRNA発現ベクターをそれぞれ500 ngトランスフェクトした。電気穿孔では、Neon（登録商標）トランスフェクションシステム（Thermo Fisher Scientific社）を製造者の指示に従って用いて細胞に電気穿孔した。

10

【0559】

リポフェクションまたは電気穿孔の24～48時間後、顕微鏡で細胞を調べてGFP<sup>+</sup>細胞の存在を探すことによりGFPの発現を求めた。目視検査の後、GFP<sup>+</sup>細胞とGFP<sup>-</sup>細胞の割合を求めることができる。表21に報告されているデアミナーゼ-RGN融合タンパク質を発現している哺乳動物の細胞で蛍光を観察した。当業者は、RIPAバッファを用いて細胞を溶解させ、得られたライセートを蛍光プレートリーダーで分析してBFPとGFPの蛍光強度を求められることもわかるであろう。それに加え、細胞を分類することによって分析し、BFP<sup>+</sup>細胞、GFP<sup>+</sup>細胞、およびGFP<sup>-</sup>細胞の正確な比率を求めることができる。

20

【0560】

【表21】

表21：哺乳動物シトシンデアミナーゼ蛍光レポータの結果

デアミナーゼの ID	相対 GFP <sup>+</sup> 細胞
APG09260	N. D.
APG09980	+
APG07386-CTD	++
APG05840	+
APG00868	++

N. D. = 検出されず；+ = わずかなGFP<sup>+</sup>細胞が検出された；++ = 多くのGFP<sup>+</sup>細胞が検出された

30

【0561】

実施例15：アデノシン塩基エディタとシトシン塩基エディタを創出するためのデアミナーゼの多様化

【0562】

DNA配列の中のA残基またはT残基をGまたはCに変えることができるアデノシン塩基エディタ（ABE）は、自然界における存在は知られていない。DNA鋳型として作用するABEを同定することを目的として、以下の多様化戦略を利用してデアミナーゼバリエーションを生成させた。

【0563】

第1の戦略は、エラーしやすいPCR酵素によるランダムな突然変異誘発であり、これはGaudelliら（Nature、2017年、doi:10.1038/nature24644；参照によって本明細書に組み込まれている）に似ている。プライマーを設計し、表17の125個のADAT酵素に対するヌクレオチド配列を生成させた。GenMorph II Random Mutagenesis Kit（Agilent Technologies社）を製造者の指示に従って用い、エラーしやすいPCRを実施した。ZR-96 DNA Clean-up Kit（商標）（Zymo Research社）を製造者の指示に従って用い、変異したADAT PCR産物を精製した。このアプローチから、約1000万個のADATバリエーションが生成した。

40

【0564】

追求した第2の戦略は遺伝子シャッフリングであり、これはStemmer, W. P. C.（Proc. Natl. Acad. Sci. USA、1994年）に似ている。Phusion（登録商標）High-Fidelity D

50

NA Polymerase (NEBiolabs社) を主に製造者の指示に従って用い、表17の125個のADAT酵素に関するコード配列をPCRで増幅した後、DNA Clean & Concentrator (商標) -5 (Zymo Research社) を製造者の指示に従って用いて精製した。異なるADATコード配列の精製されたDNAをプールし、Cutsmart (登録商標) バッファの中で酵素RsaI、AfeI、BsaAI、およびBsaHIによる制限消化を37℃で60分間実施した。消化後、反応物をゲル電気泳動によって分離させ、消化された断片をZymoclean (商標) Gel DNA Recovery Kitを用いて精製した。次に、精製された断片をPCR増幅を利用して組み立て、組み立てられた増幅産物を精製した。このアプローチから、約200万個のADATバリエーションが生成した。

#### 【0565】

最後の戦略は、核酸と相互作用する可能性がある構造残基と相同な残基を標的とするものであった。ADAT様酵素については、RNAとの複合体になった黄色ブドウ球菌のtRNAアデノシンデアミナーゼであるTadA (RCSB Protein Data Bank ID No: 2B3J; Losey 他、2006年、Nat. Struct. Mol. Biol. 第13巻: 153~159ページ) の結晶構造を注意深く調べたところ、大腸菌のTadA (UniProt P68398と、GenBankアクセッション番号NP\_417054) と相同な残基P48、L84、A106、D108、およびK110の位置で相互作用する可能性があることが明らかになった。次いで、相同なこれら残基の飽和突然変異誘発を、表17からの125個のADATドメインデアミナーゼで実施した。このアプローチから、約100万個のADATバリエーションが生成した。

#### 【0566】

当業者は、これらのアプローチが排他的ではなく、組み合わせで次々と続く各改良ラウンドの出力に適用できることがわかるであろう。飽和突然変異誘発または回帰的飽和突然変異誘発 (ReetzとCarballeira、2007年、Nature Protocols、第2巻(4): 891~903ページ) は回帰的であることが知られている。これは、1つのラウンドからの最高性能の複数の候補を選択し、1つの最適な候補が同定されるまで変異とスクリーニングのさらなるラウンドを実施することを意味する。

#### 【0567】

実施例16: アデノシン塩基エディタとシトシン塩基エディタを創出するための活性な多様化されたデアミナーゼの決定

#### 【0568】

実施例16.1: 多様化されたデアミナーゼを探す細菌活性アッセイ

#### 【0569】

上記の実施例15の多様化されたラウンドからのADAT変異バリエーション産物を、N末端にNLS (配列番号10) を含んでいてC末端の位置で変異デアミナーゼに機能可能に連結された、C末端の位置でリンカー配列 (配列番号546) に機能可能に連結された、C末端の位置でRNA誘導型DNA結合タンパク質、すなわちヌクレアーゼ不活性なRNA誘導型ヌクレアーゼ (RGN) dAPG08290.1バリエーション (配列番号547) に機能可能に連結された、C末端の位置で第2のNLSに機能可能に連結された、C末端の位置でTEV部位 (配列番号548) に機能可能に連結された、C末端の位置で10x Hisタグ (配列番号594) に機能可能に連結された融合タンパク質を産生する発現カセットに導入した。

#### 【0570】

この活性アッセイは実施例13と非常によく似ていた。しかしこれらの実験では、H193がシトシン塩基編集選択のためにH193R (配列番号566) に変異するか、アデノシン塩基編集選択のためにH193Y (配列番号567) に変異した脱活性化されたクロラムフェニコール遺伝子を持つコンストラクトを作製した。それぞれのプラスミドは、RGN-デアミナーゼ融合体をプラスミドの適切な領域に向かわせるためのガイドRNAも含有していた。(配列番号566について) CGからTAへの変換、または(配列番号567について) ATからGCへの変換が成功すると、細菌細胞は、クロラムフェニコールを含有する培地の中で生き延びることができると考えられる。

#### 【0571】

融合タンパク質発現ベクターを、脱活性化されたクロラムフェニコール遺伝子を含んでいてsgRNAを標的とするベクターとともにT7 Express大腸菌細胞（NEBioLabs社）に導入した。これらの細胞を飽和するまで増殖させた後、自己誘導培地（Media（商標）、Thermo Fisher Scientific社）の接種源として使用し、さらに5時間にわたって増殖させた。細胞を希釈し、カナマイシンとカルベニシリンを有するLBの上に播種した。これらの希釈プレートを用いて全細胞をカウントした。同じ細胞を、カナマイシン、カルベニシリン、およびクロラムフェニコールを有するLBの上にも播種し、活性な変異デアミナーゼを有するコロニーを同定した。代表的なクローンを採取し、シーケンシングした後、個別に調べた。それに加え、同じ細胞を、カナマイシンとリファンピシンを有するLBの上にも播種し、「非標的」変異率を測定した。標的にされるデアミナーゼ-RGN融合タンパク質に関するH193Rプラスミド（「CBE」）への変異率（「Mutn Rate」）と、標的にされるデアミナーゼ-RGN融合タンパク質に関するH193Yプラスミド（「ABE」）への変異率を計算して表22に示してある。デッドRGN（dAPG08290.1；配列番号547）だけを含む細胞で観察されたバックグラウンドと比較したアデノシン塩基編集の相対率（「ABE Rel Rate」）、シトシン塩基編集の相対率（「CBE Rel Rate」）、およびオフターゲットの相対率（「Off-T Rel Rate」）も示してある。シトシンデアミナーゼとして機能することが知られている陽性対照である哺乳動物のデアミナーゼも含まれていた。

【 0 5 7 2 】

【表 2 2】

表22：選択された変異デアミナーゼの変異率

デアミナーゼの ID	配列番号	ABE Mutn Rate	CBE Mutn Rate	オフターゲット	ABE Rel Rate	CBE Rel Rate	Off-T Rel Rate
APG05241	397	4.44E-09	5.25E-05	3.64E-06	14	1110	170
APG08360	411	0.00E+00	8.72E-07	1.40E-08	0	18	1
APG09980	420	4.57E-09	6.35E-06	1.10E-05	14	134	514
APG00868	374	6.57E-10	2.27E-04	5.02E-07	2	4796	23
APOBEC3A（陽性対照）	571	0.00E+00	6.06E-04	8.45E-06	0	12809	395
APG07164_T102K_D104Y_K106T	572	6.46E-08	5.76E-05	1.16E-08	204	1218	1
NTerm_APG03542(1-88)&Cterm_APG02410(114-179)	573	9.93E-08	3.56E-08	1.69E-08	313	1	1
APG07458	514	5.13E-07	2.31E-08	2.60E-08	1621	0	1
APG06334_A101E_D103S_A105K	574	3.56E-08	8.13E-08	ND	112	2	ND
APG03856_A101K_D103L_K105E	575	1.62E-08	7.30E-09	ND	51	0	ND
APG08799_D105A_K107R	576	5.74E-07	3.16E-06	2.60E-08	1812	67	1
APG02312_A102G_D104S_K106R	577	1.10E-07	4.46E-07	2.56E-08	347	9	1
APG09352_D105S_K107T	578	1.36E-06	9.34E-08	4.34E-08	4293	2	2
APG02312_A102S_D104Q_K106G	579	1.36E-06	9.34E-08	4.34E-08	4293	2	2
APG03140_D111G	580	1.24E-06	4.48E-08	4.63E-08	3921	1	2
APG03557_A104Q_D106G_K108R	581	1.12E-06	6.90E-07	3.80E-08	3522	15	2
APG07164_T102R_D104W_K106E	582	3.83E-08	1.81E-07	2.24E-08	121	4	1
APG02312_D104R_K106S	583	2.11E-06	6.53E-08	1.64E-08	6663	1	1
APG03140_A110F_D112S_K114T	584	4.16E-07	3.50E-08	4.61E-08	1314	1	2
dAPG08290.1	547	3.17E-10	4.73E-08	2.14E-08	1	1	1

【 0 5 7 3 】

## 実施例16.2：標的にされる塩基編集の蛍光アッセイ

## 【0574】

アデノシンデアミナーゼを用いてW58コドンの第3位をAからGに変化させて終止（TGA）から野生型トリプトファン（TGG）残基へと戻せるようにするため、未熟な終止コドン（配列番号564）を生じさせるW58\*変異を含有する増強型緑色蛍光タンパク質（EGFP）を有するベクターを構成した。AからGへの変換が成功するとEGFPが発現してそれを定量することができる。デアミナーゼ-RGN融合タンパク質をW58\*変異の周辺領域に向かわせるガイドRNA（配列番号565）を発現することのできる第2のベクターも作製した。

## 【0575】

リポフェクションまたは電気穿孔を利用して、このデッドEGFP EGFP受容体ベクターを、デアミナーゼ-RGN融合タンパク質とガイドRNAを発現することのできるベクターとともにHEK293T細胞にトランスフェクトした。リポフェクションでは、増殖培地（DMEM + 10% ウシ胎仔血清 + 1% ペニシリン / ストレプトマイシン）の中でのトランスフェクションの前日に細胞を  $1 \times 10^5$  細胞 / ウエルで24ウエルのプレートに播種した。Lipofectamine（登録商標）3000試薬（Thermo Fisher Scientific社）を製造者の指示に従って用い、デッドEGFPレポーターベクター、デアミナーゼ-RGN融合発現ベクター、およびガイドRNA発現ベクターをそれぞれ500 ngトランスフェクトした。電気穿孔では、Neon（登録商標）トランスフェクションシステム（Thermo Fisher Scientific社）を製造者の指示に従って用いて細胞に電気穿孔した。

## 【0576】

リポフェクションまたは電気穿孔の24～48時間後、顕微鏡で細胞を調べ、GFP<sup>+</sup>細胞の存在を探すことによってGFPの発現を求めた。目視検査の後、GFP<sup>+</sup>細胞とGFP<sup>-</sup>細胞の割合を求めることができる。表23に報告されているデアミナーゼ-RGN融合タンパク質を発現している哺乳動物の細胞で蛍光を観察した。当業者は、RIPAバッファを用いて細胞を溶解させ、得られたライセートを蛍光プレートリーダーで分析してGFPの蛍光強度を求められることもわかるであろう。それに加え、細胞を分類することによって分析し、GFP<sup>+</sup>細胞とGFP<sup>-</sup>細胞の正確な比率を求めることができる。

## 【0577】

## 【表23】

表23：哺乳動物のアデノシンデアミナーゼ蛍光受容体の結果

デアミナーゼの ID	配列番号	相対 GFP <sup>+</sup> 細胞
APG07164 T102K D104Y K106T	572	+
NTerm_APG03542 (1-88) & Cterm_APG02410 (114-179)	573	+
APG06334_A101E_D103S_A105K	574	++
APG03856_A101K_D103L_K105E	575	+
APG08799 D105A_K107R	576	++
APG09352_D105S_K107T	578	+
APG02312_A102S_D104Q_K106G	579	+
APG03140_D111G	580	+
APG03557_A104Q_D106G_K108R	581	+++
APG02312_D104R_K106S	583	+
APG03140_A110F_D112S_K114T	584	+

+ = わずかなGFP<sup>+</sup>細胞が検出された；++ = 多くのGFP<sup>+</sup>細胞が検出された；+++ = 最も多くのGFP<sup>+</sup>細胞が観察された

## 【配列表】

0007572360000001.app

## フロントページの続き

(51)国際特許分類 F I  
C 1 2 N 5/10 (2006.01) C 1 2 N 5/10  
C 1 2 N 9/14 (2006.01) C 1 2 N 9/14

(33)優先権主張国・地域又は機関

米国(US)

(31)優先権主張番号 62/790,258

(32)優先日 平成31年1月9日(2019.1.9)

(33)優先権主張国・地域又は機関

米国(US)

(31)優先権主張番号 62/790,261

(32)優先日 平成31年1月9日(2019.1.9)

(33)優先権主張国・地域又は機関

米国(US)

(31)優先権主張番号 62/790,262

(32)優先日 平成31年1月9日(2019.1.9)

(33)優先権主張国・地域又は機関

米国(US)

(31)優先権主張番号 62/790,266

(32)優先日 平成31年1月9日(2019.1.9)

(33)優先権主張国・地域又は機関

米国(US)

(31)優先権主張番号 62/932,169

(32)優先日 令和1年11月7日(2019.11.7)

(33)優先権主張国・地域又は機関

米国(US)

## 前置審査

弁理士 渡辺 陽一

(74)代理人 100108903

弁理士 中村 和広

(72)発明者 タイソン ディー・ボーエン

アメリカ合衆国, ノース カロライナ 27560, モリスビル, ウォルトンズ クリーク ロード  
137

(72)発明者 アレクサンドラ ブライナー クローリー

アメリカ合衆国, ノース カロライナ 27511, ケーリー, スコット プレイス 1220

(72)発明者 テッド ディー・エリック

アメリカ合衆国, ノース カロライナ 27705, ダーラム, アルムウッド プレイス 5

(72)発明者 マーク ムーア

アメリカ合衆国, ノース カロライナ 27713, ダーラム, クリムゾン オークス ドライブ 18

(72)発明者 ロドルフ バーランゴ

アメリカ合衆国, ロース カロライナ 27606-4562, ローリー, テリントン レーン 5  
944

(72)発明者 マイケル ラスナー

アメリカ合衆国, オレゴン 94960, ポートランド, サウスイースト ベルモント ストリート  
4162, アpartment 301

審査官 鳥居 敬司

(56)参考文献 特表2015-523856(JP, A)

Nucleic Acids Research, 2014年, Vol.42, No.10, pp.6091-6105

(58)調査した分野 (Int.Cl., D B 名)

C 1 2 N 15/00-15/90



---

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