

(19) 日本国特許庁(JP)

## (12) 特許公報(B2)

(11) 特許番号

特許第6621738号  
(P6621738)

(45) 発行日 令和1年12月18日(2019.12.18)

(24) 登録日 令和1年11月29日(2019.11.29)

(51) Int.Cl.	F 1
C 12 N 15/09	(2006.01) C 12 N 15/09 Z N A Z
C 12 N 1/15	(2006.01) C 12 N 15/09 1 1 O
C 12 N 1/19	(2006.01) C 12 N 1/15
C 12 N 5/10	(2006.01) C 12 N 1/19 C 12 N 5/10

請求項の数 21 (全 62 頁)

(21) 出願番号	特願2016-517954 (P2016-517954)
(86) (22) 出願日	平成26年6月4日(2014.6.4)
(65) 公表番号	特表2016-521554 (P2016-521554A)
(43) 公表日	平成28年7月25日(2016.7.25)
(86) 国際出願番号	PCT/US2014/040868
(87) 国際公開番号	W02014/197568
(87) 国際公開日	平成26年12月11日(2014.12.11)
審査請求日	平成29年6月5日(2017.6.5)
(31) 優先権主張番号	61/830,787
(32) 優先日	平成25年6月4日(2013.6.4)
(33) 優先権主張国・地域又は機関	米国(US)

(73) 特許権者	507044516 プレジデント アンド フェローズ オブ ハーバード カレッジ アメリカ合衆国 マサチューセッツ 02 138, ケンブリッジ, クインシー ストリート 17
(74) 代理人	100079049 弁理士 中島 淳
(74) 代理人	100084995 弁理士 加藤 和詳
(72) 発明者	チャーチ、ジョージ エム. アメリカ合衆国 02446 マサチュー セツ州 ブルックリン ケント ストリ ート 218

最終頁に続く

(54) 【発明の名称】 RNA誘導性転写制御

## (57) 【特許請求の範囲】

## 【請求項 1】

ガイド RNA および RNA 結合ドメインの標的を含み、該ガイド RNA が標的核酸を含む DNA に相補的である 1 種類以上のキメラガイド RNA をコードする第 1 外来核酸を細胞に導入すること。

前記 1 種類以上のキメラガイド RNA によってガイドされるヌクレアーゼ欠損 (nuclease - nuclease) Cas9 タンパク質をコードする第 2 外来核酸を前記細胞に導入すること。

転写制御因子タンパク質またはドメインと前記 RNA 結合ドメインとの融合体をコードする第 3 外来核酸を前記細胞に導入することを含み、

前記 1 種類以上のキメラガイド RNA、前記ヌクレアーゼ欠損 Cas9 タンパク質、および前記転写制御因子タンパク質またはドメインと前記 RNA 結合ドメインとの融合体が発現し、前記 RNA 結合ドメインと前記 RNA 結合ドメインの標的とが互いに結合し、

前記 1 種類以上のキメラガイド RNA、前記ヌクレアーゼ欠損 Cas9 タンパク質、および前記転写制御因子タンパク質またはドメインと RNA 結合ドメインとの融合体が前記 DNA に共局在し、前記転写制御因子タンパク質またはドメインが前記標的核酸の発現を制御する。

細胞において標的核酸の発現を制御する方法(ただし、該方法が人体内で行われる場合を除く)。

## 【請求項 2】

10

20

前記細胞が真核細胞である、請求項 1 に記載の方法。

【請求項 3】

前記細胞が酵母細胞、植物細胞、または動物細胞である、請求項 1 に記載の方法。

【請求項 4】

前記転写制御因子タンパク質またはドメインが転写活性化因子である、請求項 1 に記載の方法。

【請求項 5】

前記転写制御因子タンパク質またはドメインが前記標的核酸の発現を上方制御する、請求項 1 に記載の方法。

【請求項 6】

前記転写制御因子タンパク質またはドメインが、ヒト以外の対象における疾患または有害な健康状態を治療するために前記標的核酸の発現を上方制御する、請求項 1 に記載の方法。

10

【請求項 7】

前記標的核酸が疾患または有害な健康状態と関連する、請求項 1 に記載の方法。

【請求項 8】

前記ガイド RNAが t r a c r R N A c r R N A 融合体である、請求項 1 に記載の方法。

【請求項 9】

前記 D N A がゲノム D N A 、ミトコンドリア D N A 、ウイルス D N A 、または外来性 D N A である、請求項 1 に記載の方法。

20

【請求項 10】

前記RNA 結合ドメインの標的が 2 コピーの M S 2 バクテリオファージ・コートプロテイン結合性 RNA ステムループを含む、請求項 1 に記載の方法。

【請求項 11】

前記RNA 結合ドメインが M S 2 バクテリオファージ・コートプロテインを含む、請求項 1 に記載の方法。

【請求項 12】

前記RNA 結合ドメインの標的がアプタマーを含む、請求項 1 に記載の方法。

【請求項 13】

ガイド RNA および RNA 結合ドメインの標的を含み、該ガイド RNA が標的核酸を含む D N A に相補的である 1 種類以上のキメラガイド RNAをコードする第 1 外来核酸、

ヌクレアーゼ欠損 C a s 9 タンパク質をコードする第 2 外来核酸、および

転写制御因子タンパク質またはドメインと前記 RNA 結合ドメインとの融合体をコードする第 3 外来核酸を含み、

前記 1 種類以上のキメラガイド RNA、前記ヌクレアーゼ欠損 C a s 9 タンパク質、および前記転写制御因子タンパク質またはドメインと前記 RNA 結合ドメインとの融合体が、前記標的核酸に対する共局在複合体の構成要素である、

細胞。

【請求項 14】

前記細胞が真核細胞である、請求項 13 に記載の細胞。

40

【請求項 15】

前記転写制御因子タンパク質またはドメインが転写活性化因子である、請求項 13 に記載の細胞。

【請求項 16】

前記転写制御因子タンパク質またはドメインが、疾患または有害な健康状態を治療するため前記標的核酸の発現を上方制御する、請求項 13 に記載の細胞。

【請求項 17】

前記ガイド RNAが t r a c r R N A c r R N A 融合体である、請求項 13 に記載の細胞。

50

**【請求項 18】**

前記DNAがゲノムDNA、ミトコンドリアDNA、ウイルスDNA、または外来性DNAである、請求項13に記載の細胞。

**【請求項 19】**

前記RNA結合ドメインの標的が2コピーのMS2バクテリオファージ・コートプロテイン結合性RNAシステムループを含む、請求項13に記載の細胞。

**【請求項 20】**

前記RNA結合ドメインがMS2バクテリオファージ・コートプロテインを含む、請求項13に記載の細胞。

**【請求項 21】**

10

前記RNA結合ドメインの標的がアプタマーを含む、請求項13に記載の細胞。

**【発明の詳細な説明】****【技術分野】****【0001】****関連出願データ**

本願は2013年6月4日に出願された、ここに全ての目的のために全体の参照により本明細書に援用される、米国仮特許出願番号第61/830787号の優先権を主張する。

**【0002】****政府権益の説明**

20

本発明は米国国立衛生研究所助成金番号P50 HG005550と米国エネルギー省の助成金番号DE-FG02-02ER63445の国庫補助により行われた。米国政府は本発明について一定の権利を有する。

**【背景技術】****【0003】**

細菌と古細菌のCRISPR CasシステムはCasタンパク質と複合体になっている短鎖ガイドRNAに依存して、侵入してきた外来核酸の内に存在する相補配列の分解を導く。Deltcheva, E.ら著、CRISPR RNA maturation by trans-encoded small RNA and host factor RNaseIII. Nature誌、第471巻、602~607頁(2011年)、Gasiunas, G., Barrangou, R., Horvath, P. およびSiksnys, V.著、Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proceedings of the National Academy of Sciences of the United States of America誌、第109巻、E2579~2586頁(2012年)、Jinek, M.ら著、A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science誌、第337巻、816~821頁(2012年); Sapranauskas, R.ら著、The Streptococcus thermophilus CRISPR/Cas system provides immunity in Escherichia coli. Nucleic acids research誌、第39巻、9275~9282頁(2011年)、およびBhaya, D., Davison, M. およびBarrangou, R.著、CRISPR-Cas system in bacteria and archaea: versatile small RNAs for adaptive defense and regulation. Annual review of genetics誌、第45巻、273~297頁(2011年)を参照されたい。最近の、化膿レンサ球菌(S. pyogenes)のII型CRISPRシステムのインビトロにおける再構成により、通常

30

40

50

トランスにコードされるtracrRNA(「trans-activating CRISPR RNA」)と融合したcrRNA(「CRISPR RNA」)は、そのcrRNAに一致する標的DNA配列をCas9タンパク質に配列特異的に切断させるのに充分であることが示された。標的部位に相同であるgRNAの発現によりCas9の動員と標的DNAの分解が引き起こされる。H. Deveauら著、Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. Journal of Bacteriology誌、第190巻、1390頁(2008年2月)を参照されたい。

【発明の概要】

【発明が解決しようとする課題】

10

【0004】

本開示の態様は、ガイドRNAと、DNA結合タンパク質と、二本鎖DNA標的配列との複合体に関するものである。ある特定の態様によると、本開示の範囲内のDNA結合タンパク質には、ガイドRNAとの複合体を形成するタンパク質、およびその複合体を二本鎖DNA配列にガイドするガイドRNAとの複合体であって、そのDNA配列に結合する前記複合体を形成するタンパク質が含まれる。本開示のこの態様を前記RNAとDNA結合タンパク質の二本鎖DNAへの共局在、または二本鎖DNAとの共局在と呼ぶことができる。このようにDNA結合タンパク質 ガイドRNA複合体を使用して、標的DNAの発現を制御するために標的DNAに転写制御因子タンパク質またはドメインを局在させることができる。

20

【課題を解決するための手段】

【0005】

ある特定の態様によると、細胞において標的核酸の発現を制御する方法であって、前記標的核酸を含むDNA(デオキシリボ核酸)に相補的である1種類以上のRNA(リボ核酸)をコードする第1外来核酸を前記細胞に導入すること、前記DNAに結合し且つ前記1種類以上のRNAによってガイドされるRNA誘導型ヌクレアーゼ欠損(nuclease-null)DNA結合タンパク質をコードする第2外来核酸を前記細胞に導入すること、転写制御因子タンパク質またはドメインをコードする第3外来核酸を前記細胞に導入することを含み、前記1種類以上のRNA、前記RNA誘導型ヌクレアーゼ欠損DNA結合タンパク質、および前記転写制御因子タンパク質またはドメインが発現し、前記1種類以上のRNA、前記RNA誘導型ヌクレアーゼ欠損DNA結合タンパク質、および前記転写制御因子タンパク質またはドメインが前記DNAに共局在し、且つ、前記転写制御因子タンパク質またはドメインが前記標的核酸の発現を制御する前記方法が提供される。

30

【0006】

一つの態様によると、RNA誘導型ヌクレアーゼ欠損DNA結合タンパク質をコードする前記外来核酸は、前記RNA誘導型ヌクレアーゼ欠損DNA結合タンパク質に融合した前記転写制御因子タンパク質またはドメインをさらにコードする。一つの態様によると、1種類以上のRNAをコードする前記外来核酸はRNA結合ドメインの標的をさらにコードし、前記転写制御因子タンパク質またはドメインをコードする前記外来核酸は、前記転写制御因子タンパク質またはドメインに融合したRNA結合ドメインをさらにコードする。

40

【0007】

一つの態様によると、前記細胞は真核細胞である。一つの態様によると、前記細胞は酵母細胞、植物細胞、または動物細胞である。一つの態様によると、前記細胞は哺乳類細胞である。

【0008】

一つの態様によると、前記RNAは約10ヌクレオチドから約500ヌクレオチドの間である。一つの態様によると、前記RNAは約20ヌクレオチドから約100ヌクレオチドの間である。

【0009】

50

一つの態様によると、前記転写制御因子タンパク質またはドメインは転写活性化因子である。一つの態様によると、前記転写制御因子タンパク質またはドメインは前記標的核酸の発現を上方制御する。一つの態様によると、前記転写制御因子タンパク質またはドメインは疾患または有害な健康状態を治療するために前記標的核酸の発現を上方制御する。一つの態様によると、前記標的核酸は疾患または有害な健康状態と関連する。

【 0 0 1 0 】

一つの態様によると、前記1種類以上のRNAはガイドRNAである。一つの態様によると、前記1種類以上のRNAはtracrRNA-crrNA融合体である。一つの態様によると、前記ガイドRNAはスペーサー配列とトレーサーメイト配列(tracer mate sequence)を含む。前記ガイドRNAは、その一部がtracrメイト配列にハイブリダイズするtracr配列を含んでもよい。前記ガイドRNAはトレーサーメイト配列とtracr配列を連結(link)してtracrRNA-crrNA融合体を作製するリンクカーネ酸配列を含んでもよい。前記スペーサー配列は、例えばハイブリダイゼーションにより、標的DNAに結合する。

10

[ 0 0 1 1 ]

一つの態様によると、前記ガイドRNAは短縮型スペーサー配列を含む。一つの態様によると、前記ガイドRNAは1塩基の5'末端短縮を有する短縮型スペーサー配列を含む。一つの態様によると、前記ガイドRNAは2塩基の5'末端短縮を有する短縮型スペーサー配列を含む。一つの態様によると、前記ガイドRNAは3塩基の5'末端短縮を有する短縮型スペーサー配列を含む。一つの態様によると、前記ガイドRNAは4塩基の5'末端短縮を有する短縮型スペーサー配列を含む。したがって、前記スペーサー配列は1塩基～4塩基の5'末端短縮をその配列に有し得る。

30

[ 0 0 1 2 ]

ある特定の実施形態によると、前記スペーサー配列は前記標的核酸配列にハイブリダイズする約16個から約20個の間のヌクレオチドを含んでよい。ある特定の実施形態によると、前記スペーサー配列は前記標的核酸配列にハイブリダイズする約20ヌクレオチドを含んでよい。

[ 0 0 1 3 ]

ある特定の態様によると、前記リンカー核酸配列は約4個から約6個の間の核酸を含んでよい。

88

[ 0 0 1 4 ]

ある特定の態様によると、前記 tracr 配列は約 60 個から約 500 個の間の核酸を含んでよい。ある特定の態様によると、前記 tracr 配列は約 64 個から約 500 個の間の核酸を含んでよい。ある特定の態様によると、前記 tracr 配列は約 65 個から約 500 個の間の核酸を含んでよい。ある特定の態様によると、前記 tracr 配列は約 66 個から約 500 個の間の核酸を含んでよい。ある特定の態様によると、前記 tracr 配列は約 67 個から約 500 個の間の核酸を含んでよい。ある特定の態様によると、前記 tracr 配列は約 68 個から約 500 個の間の核酸を含んでよい。ある特定の態様によると、前記 tracr 配列は約 69 個から約 500 個の間の核酸を含んでよい。ある特定の態様によると、前記 tracr 配列は約 70 個から約 500 個の間の核酸を含んでよい。ある特定の態様によると、前記 tracr 配列は約 80 個から約 500 個の間の核酸を含んでよい。ある特定の態様によると、前記 tracr 配列は約 90 個から約 500 個の間の核酸を含んでよい。ある特定の態様によると、前記 tracr 配列は約 100 個から約 500 個の間の核酸を含んでよい。

40

[ 0 0 1 5 ]

ある特定の態様によると、前記 tracr 配列は約 60 個から約 200 個の間の核酸を含んでよい。ある特定の態様によると、前記 tracr 配列は約 64 個から約 200 個の間の核酸を含んでよい。ある特定の態様によると、前記 tracr 配列は約 65 個から約 200 個の間の核酸を含んでよい。ある特定の態様によると、前記 tracr 配列は約 66 個から約 200 個の間の核酸を含んでよい。ある特定の態様によると、前記 tracr 配列は約 67 個から約 200 個の間の核酸を含んでよい。

50

配列は約 67 個から約 200 個の間の核酸を含んでよい。ある特定の態様によると、前記 tracr 配列は約 68 個から約 200 個の間の核酸を含んでよい。ある特定の態様によると、前記 tracr 配列は約 69 個から約 200 個の間の核酸を含んでよい。ある特定の態様によると、前記 tracr 配列は約 70 個から約 200 個の間の核酸を含んでよい。ある特定の態様によると、前記 tracr 配列は約 80 個から約 200 個の間の核酸を含んでよい。ある特定の態様によると、前記 tracr 配列は約 90 個から約 200 個の間の核酸を含んでよい。ある特定の態様によると、前記 tracr 配列は約 100 個から約 200 個の間の核酸を含んでよい。

【0016】

例示的なガイド RNA が図 5B に示されている。

10

【0017】

一つの態様によると、前記 DNA はゲノム DNA、ミトコンドリア DNA、ウイルス DNA、または外来性 DNA である。

【0018】

ある特定の態様によると、細胞において標的核酸の発現を制御する方法であって、前記標的核酸を含む DNA (デオキシリボ核酸) に相補的である 1 種類以上の RNA (リボ核酸) をコードする第 1 外来核酸を前記細胞に導入すること、前記 DNA に結合し且つ前記 1 種類以上の RNA によってガイドされる I I 型 CRISPR システムの RNA 誘導型ヌクレアーゼ欠損 DNA 結合タンパク質をコードする第 2 外来核酸を前記細胞に導入すること、転写制御因子タンパク質またはドメインをコードする第 3 外来核酸を前記細胞に導入することを含み、前記 1 種類以上の RNA、I I 型 CRISPR システムの前記 RNA 誘導型ヌクレアーゼ欠損 DNA 結合タンパク質、および前記転写制御因子タンパク質またはドメインが発現し、前記 1 種類以上の RNA、I I 型 CRISPR システムの前記 RNA 誘導型ヌクレアーゼ欠損 DNA 結合タンパク質、および前記転写制御因子タンパク質またはドメインが前記 DNA に共局在し、且つ、前記転写制御因子タンパク質またはドメインが前記標的核酸の発現を制御する前記方法が提供される。

20

【0019】

一つの態様によると、I I 型 CRISPR システムの RNA 誘導型ヌクレアーゼ欠損 DNA 結合タンパク質をコードする前記外来核酸は、I I 型 CRISPR システムの前記 RNA 誘導型ヌクレアーゼ欠損 DNA 結合タンパク質に融合した前記転写制御因子タンパク質またはドメインをさらにコードする。一つの態様によると、1 種類以上の RNA をコードする前記外来核酸は RNA 結合ドメインの標的をさらにコードし、前記転写制御因子タンパク質またはドメインをコードする前記外来核酸は、前記転写制御因子タンパク質またはドメインに融合した RNA 結合ドメインをさらにコードする。

30

【0020】

一つの態様によると、前記細胞は真核細胞である。一つの態様によると、前記細胞は酵母細胞、植物細胞、または動物細胞である。一つの態様によると、前記細胞は哺乳類細胞である。

【0021】

一つの態様によると、前記 RNA は約 10 ヌクレオチドから約 500 ヌクレオチドの間である。一つの態様によると、前記 RNA は約 20 ヌクレオチドから約 100 ヌクレオチドの間である。

40

【0022】

一つの態様によると、前記転写制御因子タンパク質またはドメインは転写活性化因子である。一つの態様によると、前記転写制御因子タンパク質またはドメインは前記標的核酸の発現を上方制御する。一つの態様によると、前記転写制御因子タンパク質またはドメインは疾患または有害な健康状態を治療するために前記標的核酸の発現を上方制御する。一つの態様によると、前記標的核酸は疾患または有害な健康状態と関連する。

【0023】

一つの態様によると、前記 1 種類以上の RNA はガイド RNA である。一つの態様によ

50

ると、前記 1 種類以上の RNA は t r a c r R N A c r R N A 融合体である。

【 0 0 2 4 】

一つの態様によると、前記 DNA はゲノム DNA 、ミトコンドリア DNA 、ウイルス DNA 、または外来性 DNA である。

【 0 0 2 5 】

ある特定の態様によると、細胞において標的核酸の発現を制御する方法であって、前記標的核酸を含む DNA ( デオキシリボ核酸 ) に相補的である 1 種類以上の RNA ( リボ核酸 ) をコードする第 1 外来核酸を前記細胞に導入すること、前記 DNA に結合し且つ前記 1 種類以上の RNA によってガイドされるヌクレアーゼ欠損 Cas9 タンパク質をコードする第 2 外来核酸を前記細胞に導入すること、転写制御因子タンパク質またはドメインをコードする第 3 外来核酸を前記細胞に導入することを含み、前記 1 種類以上の RNA 、前記ヌクレアーゼ欠損 Cas9 タンパク質、および前記転写制御因子タンパク質またはドメインが発現し、前記 1 種類以上の RNA 、前記ヌクレアーゼ欠損 Cas9 タンパク質、および前記転写制御因子タンパク質またはドメインが前記 DNA に共局在し、前記転写制御因子タンパク質またはドメインが前記標的核酸の発現を制御する前記方法が提供される。

【 0 0 2 6 】

一つの態様によると、ヌクレアーゼ欠損 Cas9 タンパク質をコードする前記外来核酸は、前記ヌクレアーゼ欠損 Cas9 タンパク質に融合した前記転写制御因子タンパク質またはドメインをさらにコードする。一つの態様によると、1 種類以上の RNA をコードする前記外来核酸は RNA 結合ドメインの標的をさらにコードし、前記転写制御因子タンパク質またはドメインをコードする前記外来核酸は前記転写制御因子タンパク質またはドメインに融合した RNA 結合ドメインをさらにコードする。

【 0 0 2 7 】

一つの態様によると、前記細胞は真核細胞である。一つの態様によると、前記細胞は酵母細胞、植物細胞、または動物細胞である。一つの態様によると、前記細胞は哺乳類細胞である。

【 0 0 2 8 】

一つの態様によると、前記 RNA は約 10 ヌクレオチドから約 500 ヌクレオチドの間である。一つの態様によると、前記 RNA は約 20 ヌクレオチドから約 100 ヌクレオチドの間である。

【 0 0 2 9 】

一つの態様によると、前記転写制御因子タンパク質またはドメインは転写活性化因子である。一つの態様によると、前記転写制御因子タンパク質またはドメインは前記標的核酸の発現を上方制御する。一つの態様によると、前記転写制御因子タンパク質またはドメインは疾患または有害な健康状態を治療するために前記標的核酸の発現を上方制御する。一つの態様によると、前記標的核酸は疾患または有害な健康状態と関連する。

【 0 0 3 0 】

一つの態様によると、前記 1 種類以上の RNA はガイド RNA である。一つの態様によると、前記 1 種類以上の RNA は t r a c r R N A c r R N A 融合体である。

【 0 0 3 1 】

一つの態様によると、前記 DNA はゲノム DNA 、ミトコンドリア DNA 、ウイルス DNA 、または外来性 DNA である。

【 0 0 3 2 】

一つの態様によると、標的核酸を含む DNA に相補的である 1 種類以上の RNA をコードする第 1 外来核酸、 RNA 誘導型ヌクレアーゼ欠損 DNA 結合タンパク質をコードする第 2 外来核酸、および転写制御因子タンパク質またはドメインをコードする第 3 外来核酸を含む細胞であって、前記 1 種類以上の RNA 、前記 RNA 誘導型ヌクレアーゼ欠損 DNA 結合タンパク質、および前記転写制御因子タンパク質またはドメインが前記標的核酸に対する共局在複合体の構成要素である前記細胞が提供される。

【 0 0 3 3 】

10

20

30

40

50

一つの態様によると、RNA誘導型ヌクレアーゼ欠損DNA結合タンパク質をコードする前記外来核酸は、RNA誘導型ヌクレアーゼ欠損DNA結合タンパク質に融合した前記転写制御因子タンパク質またはドメインをさらにコードする。一つの態様によると、1種類以上のRNAをコードする前記外来核酸はRNA結合ドメインの標的をさらにコードし、前記転写制御因子タンパク質またはドメインをコードする前記外来核酸は、前記転写制御因子タンパク質またはドメインに融合したRNA結合ドメインをさらにコードする。

【0034】

一つの態様によると、前記細胞は真核細胞である。一つの態様によると、前記細胞は酵母細胞、植物細胞、または動物細胞である。一つの態様によると、前記細胞は哺乳類細胞である。

10

【0035】

一つの態様によると、前記RNAは約10ヌクレオチドから約500ヌクレオチドの間である。一つの態様によると、前記RNAは約20ヌクレオチドから約100ヌクレオチドの間である。

【0036】

一つの態様によると、前記転写制御因子タンパク質またはドメインは転写活性化因子である。一つの態様によると、前記転写制御因子タンパク質またはドメインは前記標的核酸の発現を上方制御する。一つの態様によると、前記転写制御因子タンパク質またはドメインは疾患または有害な健康状態を治療するために前記標的核酸の発現を上方制御する。一つの態様によると、前記標的核酸は疾患または有害な健康状態と関連する。

20

【0037】

一つの態様によると、前記1種類以上のRNAはガイドRNAである。一つの態様によると、前記1種類以上のRNAはtracrRNA　crRNA融合体である。

【0038】

一つの態様によると、前記DNAはゲノムDNA、ミトコンドリアDNA、ウイルスDNA、または外来性DNAである。

【0039】

ある特定の態様によると、前記RNA誘導型ヌクレアーゼ欠損DNA結合タンパク質はII型CRISPRシステムのRNA誘導型ヌクレアーゼ欠損DNA結合タンパク質である。ある特定の態様によると、前記RNA誘導型ヌクレアーゼ欠損DNA結合タンパク質はヌクレアーゼ欠損Cas9タンパク質である。

30

【0040】

一つの態様によると、細胞内でDNA標的核酸を改変させる方法であって、それぞれ前記DNA標的核酸中の隣り合う部位に相補的である2種類以上のRNAをコードする第1外来核酸を前記細胞に導入すること、前記2種類以上のRNAによってガイドされる少なくとも1種類のRNA誘導型DNA結合タンパク質ニッカーゼ(nickase)をコードする第2外来核酸を前記細胞に導入することを含み、前記2種類以上のRNA及び前記少なくとも1種類のRNA誘導型DNA結合タンパク質ニッカーゼが発現し、前記少なくとも1種類のRNA誘導型DNA結合タンパク質ニッカーゼが前記2種類以上のRNAと共に前記DNA標的核酸に共局在し、前記DNA標的核酸にニック(nick)を入れて2つ以上の隣接するニックを生じさせる前記方法が提供される。

40

【0041】

一つの態様によると、細胞内でDNA標的核酸を改変させる方法であって、それぞれ前記DNA標的核酸中の隣り合う部位に相補的である2種類以上のRNAをコードする第1外来核酸を前記細胞に導入すること、前記2種類以上のRNAによってガイドされるII型CRISPRシステムの少なくとも1種類のRNA誘導型DNA結合タンパク質ニッカーゼをコードする第2外来核酸を前記細胞に導入することを含み、前記2種類以上のRNA及び前記II型CRISPRシステムの少なくとも1種類のRNA誘導型DNA結合タンパク質ニッカーゼが発現し、前記II型CRISPRシステムの少なくとも1種類のRNA誘導型DNA結合タンパク質ニッカーゼが前記2種類以上のRNAと共に前記DNA

50

標的核酸に共局在し、前記DNA標的核酸にニックを入れて2つ以上の隣接するニックを生じさせる前記方法が提供される。

【0042】

一つの態様によると、細胞内でDNA標的核酸を改変させる方法であって、それぞれ前記DNA標的核酸中の隣り合う部位に相補的である2種類以上のRNAをコードする第1外来核酸を前記細胞に導入すること、1つの不活性型ヌクレアーゼドメインを有し、且つ、前記2種類以上のRNAによってガイドされる少なくとも1種類のCas9タンパク質ニックeraseをコードする第2外来核酸を前記細胞に導入することを含み、前記2種類以上のRNA及び前記少なくとも1種類のCas9タンパク質ニックeraseが発現し、前記少なくとも1種類のCas9タンパク質ニックeraseが前記2種類以上のRNAと共に前記DNA標的核酸に共局在し、前記DNA標的核酸にニックを入れて2つ以上の隣接するニックを生じさせる前記方法が提供される。

10

【0043】

DNA標的核酸を改変させる前記方法によると、前記2つ以上の隣接するニックは二本鎖DNAの同じストランド上に存在する。一つの態様によると、前記2つ以上の隣接するニックは二本鎖DNAの同じストランド上に存在して相同組換えを引き起こす。一つの態様によると、前記2つ以上の隣接するニックは二本鎖DNAの異なるストランド上に存在する。一つの態様によると、前記2つ以上の隣接するニックは二本鎖DNAの異なるストランド上に存在して二本鎖切断を生じさせる。一つの態様によると、前記2つ以上の隣接するニックは二本鎖DNAの異なるストランド上に存在し、二本鎖切断を生じさせて非相同末端結合を引き起こす。一つの態様によると、前記2つ以上の隣接するニックは二本鎖DNAの異なるストランド上に存在し、且つ、その位置が互いにずれて(offset with respect to one another)いる。一つの態様によると、前記2つ以上の隣接するニックは二本鎖DNAの異なるストランド上に存在し、且つ、その位置が互いにずれており、且つ、二本鎖切断を生じさせる。一つの態様によると、前記2つ以上の隣接するニックは二本鎖DNAの異なるストランド上に存在し、且つ、その位置が互いにずれており、且つ、二本鎖切断を生じさせて非相同末端結合を引き起こす。一つの態様によると、前記方法はドナー核酸配列をコードする第3外来核酸を前記細胞に導入することをさらに含み、前記2つ以上のニックにより前記標的核酸の前記ドナー核酸配列との相同組換えが引き起こされる。

20

【0044】

一つの態様によると、細胞内でDNA標的核酸を改変させる方法であって、それぞれ前記DNA標的核酸中の隣り合う部位に相補的である2種類以上のRNAをコードする第1外来核酸を前記細胞に導入すること、前記2種類以上のRNAによってガイドされる少なくとも1種類のRNA誘導型DNA結合タンパク質ニックeraseをコードする第2外来核酸を前記細胞に導入することを含み、前記2種類以上のRNA及び前記少なくとも1種類のRNA誘導型DNA結合タンパク質ニックeraseが発現し、前記少なくとも1種類のRNA誘導型DNA結合タンパク質ニックeraseが前記2種類以上のRNAと共に前記DNA標的核酸に共局在し、前記DNA標的核酸にニックを入れて2つ以上の隣接するニックを生じさせ、前記2つ以上の隣接するニックが二本鎖DNAの異なるストランド上に存在し、二本鎖切断を生じさせて前記標的核酸の断片化を引き起こすことによって前記標的核酸の発現を妨げる前記方法が提供される。

30

【0045】

一つの態様によると、細胞内でDNA標的核酸を改変させる方法であって、それぞれ前記DNA標的核酸中の隣り合う部位に相補的である2種類以上のRNAをコードする第1外来核酸を前記細胞に導入すること、前記2種類以上のRNAによってガイドされるII型CRISPRシステムの少なくとも1種類のRNA誘導型DNA結合タンパク質ニックeraseをコードする第2外来核酸を前記細胞に導入することを含み、前記2種類以上のRNA及び前記II型CRISPRシステムの少なくとも1種類のRNA誘導型DNA結合タンパク質ニックeraseが発現し、前記II型CRISPRシステムの少なくとも1種類のR

40

50

N A 誘導型 D N A 結合タンパク質ニッカーゼが前記 2 種類以上の R N A と共に前記 D N A 標的核酸に共局在し、前記 D N A 標的核酸にニックを入れて 2 つ以上の隣接するニックを生じさせ、前記 2 つ以上の隣接するニックが二本鎖 D N A の異なるストランド上に存在し、二本鎖切断を生じさせて前記標的核酸の断片化を引き起こすことによって前記標的核酸の発現を妨げる前記方法が提供される。

【 0 0 4 6 】

一つの態様によると、細胞内で D N A 標的核酸を改変させる方法であって、それぞれ前記 D N A 標的核酸中の隣り合う部位に相補的である 2 種類以上の R N A をコードする第 1 外来核酸を前記細胞に導入すること、1 つの不活性型ヌクレアーゼドメインを有し且つ前記 2 種類以上の R N A によってガイドされる少なくとも 1 種類の C a s 9 タンパク質ニッカーゼをコードする第 2 外来核酸を前記細胞に導入することを含み、前記 2 種類以上の R N A 及び前記少なくとも 1 種類の C a s 9 タンパク質ニッカーゼが発現し、前記少なくとも 1 種類の C a s 9 タンパク質ニッカーゼが前記 2 種類以上の R N A と共に前記 D N A 標的核酸に共局在し、前記 D N A 標的核酸にニックを入れて 2 つ以上の隣接するニックを生じさせ、前記 2 つ以上の隣接するニックが二本鎖 D N A の異なるストランド上に存在し、二本鎖切断を生じさせて前記標的核酸の断片化を引き起こすことによって前記標的核酸の発現を妨げる前記方法が提供される。

【 0 0 4 7 】

一つの態様によると、それが D N A 標的核酸中の隣り合う部位に相補的である 2 種類以上の R N A をコードする第 1 外来核酸、および少なくとも 1 種類の R N A 誘導型 D N A 結合タンパク質ニッカーゼをコードする第 2 外来核酸を含む細胞であって、前記 2 種類以上の R N A と前記少なくとも 1 種類の R N A 誘導型 D N A 結合タンパク質ニッカーゼが前記 D N A 標的核酸に対する共局在複合体の構成要素である前記細胞が提供される。

【 0 0 4 8 】

一つの態様によると、前記 R N A 誘導型 D N A 結合タンパク質ニッカーゼは I I 型 C R I S P R システムの R N A 誘導型 D N A 結合タンパク質ニッカーゼである。一つの態様によると、前記 R N A 誘導型 D N A 結合タンパク質ニッカーゼは 1 つの不活性型ヌクレアーゼドメインを有する C a s 9 タンパク質ニッカーゼである。

【 0 0 4 9 】

一つの態様によると、前記細胞は真核細胞である。一つの態様によると、前記細胞は酵母細胞、植物細胞、または動物細胞である。一つの態様によると、前記細胞は哺乳類細胞である。

【 0 0 5 0 】

一つの態様によると、前記 R N A は約 1 0 個から約 5 0 0 個の間のヌクレオチドを含む。一つの態様によると、前記 R N A は約 2 0 個から約 1 0 0 個の間のヌクレオチドを含む。

【 0 0 5 1 】

一つの態様によると、前記標的核酸は疾患または有害な健康状態と関連する。

【 0 0 5 2 】

一つの態様によると、前記 2 種類以上の R N A はガイド R N A である。一つの態様によると、前記 2 種類以上の R N A は t r a c r R N A c r R N A 融合体である。

【 0 0 5 3 】

一つの態様によると、前記 D N A 標的核酸はゲノム D N A 、ミトコンドリア D N A 、ウイルス D N A 、または外来性 D N A である。

【 0 0 5 4 】

本発明のある特定の実施形態のその他の特徴および利点は、後続の実施形態とそれらの図面の説明において、および特許請求の範囲からさらに十分に明らかとなろう。

【 0 0 5 5 】

本特許または特許出願のファイルは彩色図面を含む。彩色図面を有する本特許または特許出願公開の写しは、請求および所要の手数料を納付すれば、特許庁によって提供される

10

20

30

40

50

。本実施形態の前述その他の特徴および利点は、添付されている図面と併せて、後続の例示的な実施形態の詳細な説明からより十分に理解されよう。

【図面の簡単な説明】

【0056】

【図1-1】図1AはRNA誘導性転写活性化の模式図である。図1BはRNA誘導性転写活性化の模式図である。図1Cはレポーターコンストラクトのデザインを示す。

【図1-2】図1Dは蛍光活性化細胞選別(FACS)と免疫蛍光分析(IF)の両方によってアッセイしたときにCas9N-VP64融合体がRNA誘導性転写活性化を示すことを表すデータを示す。

【図1-3】図1Dは蛍光活性化細胞選別(FACS)と免疫蛍光分析(IF)の両方によってアッセイしたときにCas9N-VP64融合体がRNA誘導性転写活性化を示すことを表すデータを示す。 10

【図1-4】図1EはCas9N、MS2-VP64、および適切なMS2アプタマー結合部位を有するgRNAの存在下でのレポーターコンストラクトのgRNA配列特異的転写活性化を示すFACSとIFによるアッセイデータを示す。

【図1-5】図1EはCas9N、MS2-VP64、および適切なMS2アプタマー結合部位を有するgRNAの存在下でのレポーターコンストラクトのgRNA配列特異的転写活性化を示すFACSとIFによるアッセイデータを示す。 20

【図1-6】図1Fは個々のgRNAと複数のgRNAによる転写誘導を表すデータを示す。 20

【0057】

【図2-1】図2AはCas9 gRNA複合体とTALEによるターゲティングのランダスケイプを評価するための方法を示す。

【図2-2】図2BはCas9 gRNA複合体が平均してその標的配列中の1~3個の突然変異に対して許容的であることを表すデータを示す。図2CはCas9 gRNA複合体が、PAM配列に局在しているものを除いて点突然変異にほとんど感受性が無いことを表すデータを示す。

【図2-3】図2Dは2塩基のミスマッチの導入によってCas9 gRNA複合体活性が顕著に損なわれることを表すヒートプロットデータを示す。図2Eは18-merのTALEが平均してその標的配列中の1~2個の突然変異に対して許容的であることを表すデータを示す。 30

【図2-4】図2Fは18-merのTALEがCas9 gRNA複合体と同様にその標的中の1塩基のミスマッチにほとんど感受性が無いことを表すデータを示す。図2Gは2塩基のミスマッチの導入によって18-merのTALEの活性が顕著に損なわれることを表すヒートプロットデータを示す。

【0058】

【図3-1】図3AはガイドRNAデザインの模式図である。

【図3-2】図3Bは5'オーバーハングを生じさせるオフセット・ニックと3'オーバーハングを生じさせるオフセット・ニックについての非相同末端結合のパーセンテージ率を示すデータを表す。 40

【図3-3】図3Cは5'オーバーハングを生じせるオフセット・ニックと3'オーバーハングを生じせるオフセット・ニックについてのターゲティングのパーセンテージ率を示すデータを表す。

【0059】

【図4-1】図4AはPDB ID: 4EP4(青色)のRuvCの位置D7にある金属配位性残基の模式図(左)、およびPDB ID: 3M7K(オレンジ色)と4H9D(青緑色)のHNHエンドヌクレアーゼドメインの模式図であって3M7Kの配位しているMgイオン(灰色の四角)とDNA(紫)を含む模式図(中央)、および分析した突然変異体のリスト(右)である。

【図4-2】図4BはCas9突然変異体m3およびm4、およびまた、それらの突然変 50

異体のそれぞれの V P 6 4 との融合体の検出不可能であるヌクレアーゼ活性を示すデータを表す。

【図 4 - 3】図 4 C は図 4 B のデータの高感度検査である。

【0 0 6 0】

【図 5 - 1】図 5 A は C a s 9 - g R N A 活性を測定するための相同組換えアッセイの模式図である。

【図 5 - 2】図 5 B は無作為配列挿入物を有するガイド R N A と相同組換えのパーセンテージ率を示す。

【図 5 - 3】図 5 B は無作為配列挿入物を有するガイド R N A と相同組換えのパーセンテージ率を示す。 10

【0 0 6 1】

【図 6 - 1】図 6 A は O C T 4 遺伝子のガイド R N A の模式図である。図 6 B はプロモーター・ルシフェラーゼレポーターコンストラクトの転写活性化を示す。

【図 6 - 2】図 6 C は q P C R によって内在性遺伝子の転写活性化を示す。

【0 0 6 2】

【図 7 - 1】図 7 A は R E X 1 遺伝子のガイド R N A の模式図である。図 7 B はプロモーター・ルシフェラーゼレポーターコンストラクトの転写活性化を示す。 20

【図 7 - 2】図 7 C は q P C R によって内在性遺伝子の転写活性化を示す。

【0 0 6 3】

【図 8 - 1】図 8 A は正規化された発現レベルの算出のための高レベルの特異性分析処理フローを模式的に示す図である。 20

【図 8 - 2】図 8 B はバイアスがかけられたコンストラクトライブラリー内で生じたミスマッチ数毎の結合部位のパーセンテージの分布状態のデータを示す。左：理論的分布状態。右：実際の T A L E コンストラクトライブラリーから観察された分布状態。図 8 C はミスマッチ数毎の結合部位に集まったタグの数のパーセンテージの分布状態のデータを示す。左：陽性対照試料から観察された分布状態。右：対照ではない T A L E がガイドされた試料から観察された分布状態。

【0 0 6 4】

【図 9 - 1】図 9 A は C a s 9 - g R N A 複合体の標的配列中の 1 ~ 3 個の突然変異に対する許容性を示す、当該複合体のターゲティングランドスケイプの分析データを示す。図 9 B は P A M 配列に局在しているものを除く点突然変異に対する非感受性を示す、 C a s 9 - g R N A 複合体のターゲティングランドスケイプの分析データを示す。 30

【図 9 - 2】図 9 C は 2 塩基のミスマッチの導入によって活性が顕著に損なわれることを示す、 C a s 9 - g R N A 複合体のターゲティングランドスケイプの分析のヒートプロットデータを示す。

【図 9 - 3】図 9 D は化膿レンサ球菌 ( S . p y o g e n e s ) の C a s 9 の予想される P A M が N G G であり、また N A G であることを確認する、ヌクレアーゼ介在性 H R アッセイのデータを示す。

【0 0 6 5】

【図 10 - 1】図 10 A は 1 8 - m e r の T A L E がそれらの標的配列中の複数の突然変異を許容することを確認する、ヌクレアーゼ介在性 H R アッセイのデータを示す。 40

【図 10 - 2】図 10 A は 1 8 - m e r の T A L E がそれらの標的配列中の複数の突然変異を許容することを確認する、ヌクレアーゼ介在性 H R アッセイのデータを示す。

【図 10 - 3】図 10 B は 3 つの異なるサイズ ( 1 8 - m e r 、 1 4 - m e r および 1 0 - m e r ) の T A L E のターゲティングランドスケイプの分析データを示す。

【図 10 - 4】図 10 C はほぼ 1 塩基ミスマッチ解像度を示す 1 0 - m e r の T A L E のデータを示す。図 10 D はほぼ 1 塩基ミスマッチ解像度を示す 1 0 - m e r の T A L E のヒートプロットデータを示す。

【0 0 6 6】

【図 11 - 1】図 11 A はデザインされたガイド R N A を示す。 50

【図11-2】図11Bは様々なガイドRNAの非相同末端結合のパーセンテージ率を示す。

【発明を実施するための形態】

【0067】

本開示の実施形態は、標的核酸の制御法においてDNAに転写制御因子タンパク質またはドメインを共局在させるためにDNA結合タンパク質を使用することに基づく。そのようなDNA結合タンパク質が様々な目的のためにDNAに結合することは当業者に既知である。そのようなDNA結合タンパク質は天然物であり得る。本開示の範囲内に含まれるDNA結合タンパク質にはここでガイドRNAと称されるRNAによってガイドされ得るもののが含まれる。この態様によると、ガイドRNAとRNA誘導型DNA結合タンパク質はDNAにおいて共局在複合体を形成する。ある特定の態様によると、DNA結合タンパク質はヌクレアーゼ欠損DNA結合タンパク質であり得る。この態様によると、ヌクレアーゼ欠損DNA結合タンパク質ヌクレアーゼ活性を有するDNA結合タンパク質の改変(a Iteration)または改造(modification)の結果から生じ得る。そのようなヌクレアーゼ活性を有するDNA結合タンパク質は当業者に知られており、例えばI I型CRISP Rシステムに存在する、Cas9タンパク質のような、ヌクレアーゼ活性を有する天然のDNA結合タンパク質を含む。そのようなCas9タンパク質とI I型CRISP Rシステムは当技術分野においてよく記述されている。ここにその全体が参照により援用される全ての追加情報を含むMakarovaら著、Nature Reviews, Microbiology誌、第9巻、2011年6月、467～477頁を参照されたい。

【0068】

ヌクレアーゼ活性を有する例示的なDNA結合タンパク質は二本鎖DNAにニックを入れるかまたは二本鎖DNAを切断するように機能する。そのようなヌクレアーゼ活性はヌクレアーゼ活性を示す1つ以上のポリペプチド配列を有するDNA結合タンパク質から生じ得る。そのような例示的なDNA結合タンパク質はそれぞれのドメインが二本鎖DNAの特定のストランドを切断すること、またはニックを入れることを担っている2つの別個のヌクレアーゼドメインを有していてよい。当業者に知られているヌクレアーゼ活性を有する例示的なポリペプチド配列にはMcRA-HNHヌクレアーゼ関連ドメインおよびRuvC様ヌクレアーゼドメインが含まれる。したがって、例示的なDNA結合タンパク質は自然界においてMcRA-HNHヌクレアーゼ関連ドメインとRuvC様ヌクレアーゼドメインのうちの1つ以上を含むタンパク質である。ある特定の態様によると、前記DNA結合タンパク質はヌクレアーゼ活性を不活性化するために改変されるか、あるいは改変される。そのような改変または改変にはヌクレアーゼ活性またはヌクレアーゼドメインを不活性化するための1つ以上のアミノ酸の変更が含まれる。そのような改変には、ヌクレアーゼ活性を示す単数または複数のポリペプチド配列、すなわち、ヌクレアーゼドメインが、前記DNA結合タンパク質に存在しないように、ヌクレアーゼ活性を示す前記単数または複数のポリペプチド配列、すなわち、ヌクレアーゼドメインを除去することが含まれる。ヌクレアーゼ活性を不活性化するための他の改変は、本開示に基づいて当業者にすぐに明らかになろう。したがって、ヌクレアーゼ欠損DNA結合タンパク質にはヌクレアーゼ活性を不活性化するように改変されたポリペプチド配列、またはヌクレアーゼ活性を不活性化するための単数または複数のポリペプチド配列の除去が含まれる。そのヌクレアーゼ欠損DNA結合タンパク質はヌクレアーゼ活性が不活性化されてもDNA結合能を保持している。したがって、前記DNA結合タンパク質はDNA結合に必要な単数または複数のポリペプチド配列を含んでいるが、ヌクレアーゼ活性を示すヌクレアーゼ配列の1つ以上または全てを欠失していてよい。したがって、前記DNA結合タンパク質はDNA結合に必要な単数または複数のポリペプチド配列を含んでいるが、不活性ヌクレアーゼ活性を示すヌクレアーゼ配列の1つ以上または全てを有していてよい。

【0069】

一つの態様によると、2つ以上のヌクレアーゼドメインを有するDNA結合タンパク質はヌクレアーゼドメインの1つを除いて全てが不活性化されているように改変または改変さ

10

20

30

40

50

れていてよい。そのような改造型または改変型DNA結合タンパク質は、そのDNA結合タンパク質が二本鎖DNAの一方のストランドだけを切断する、またはニックを入れるものである限り、DNA結合タンパク質ニッカーゼと呼ばれる。RNAによってDNAにガイドされるとき、DNA結合タンパク質ニッカーゼはRNA誘導型DNA結合タンパク質ニッカーゼと呼ばれる。

【0070】

例示的なDNA結合タンパク質としては、ヌクレアーゼ活性を欠くII型CRISPRシステムのRNA誘導型DNA結合タンパク質がある。例示的なDNA結合タンパク質としてはヌクレアーゼ欠損Cas9タンパク質がある。例示的なDNA結合タンパク質としてはCas9タンパク質ニッカーゼがある。

10

【0071】

化膿レンサ球菌(*S. pyogenes*)ではCas9は、そのタンパク質中の2つの触媒ドメイン、すなわち、DNAの相補ストランドを切断するHNHドメインと非相補ストランドを切断するRuvC様ドメインが介在する処理を介して、プロトスペーサー隣接モチーフ(PAM)の3bp上流に平滑末端二本鎖切断を形成する。ここにその全体が参照により援用されるJinkenら著、*Science*誌、第337巻、816~821頁(2012年)を参照されたい。Cas9タンパク質は、Makarovaら著、*Nature Reviews, Microbiology*誌、第9巻、2011年6月、467~477頁の追加情報中に見出される次のものを含む、多数のII型CRISPRシステムに存在することが知られている:メタノコッカス・マリパルディスC7株;コリネバクテリウム・ジフテリアエ;コリネバクテリウム・エフィシエンスYS-314株;コリネバクテリウム・グルタミカムATCC13032 Kitasato株;コリネバクテリウム・グルタミカムATCC13032 Bieliefeld株;コリネバクテリウム・グルタミカムR株;コリネバクテリウム・クロッペンステッティイDSM44385株;マイコバクテリウム・アブセサスATCC19977株;ノカルディア・ファルシニカIFM10152株;ロドコッカス・エリスロポリスPR4株;ロドコッカス・ジョスティイRHA1株;ロドコッカス・オバカスB4 uid36573株;アシドサーマス・セルロリティカス11B株;アルスロバクター・クロロフェノリカスA6株;クリベラ・フラビダDSM17836 uid43465株;サーモノスボラ・カーバタDSM43183株;ビフィドバクテリウム・デンティウムBd1株;ビフィドバクテリウム・ロングムDJ010A株;スラッキア・ヘリオトリニレデューセンスDSM20476株;パーセフォネラ・マリナEX-H1株;バクテロイデス・フラギリスNCTC9434株;カプノサイトファガ・オクラセアDSM7271株;フラボバクテリウム・サイクロフィルムJIP02-86株;アッカーマンシア・ムシニフィラATCC BAA835株;ロゼイフレクサス・キャステンホルツイイDSM13941株;ロゼイフレクサスRS1株;シネコシスティスPCC6803株;エルシミクロビウム・ミヌトウムPei191株;非培養性シロアリ1群細菌系統型Rs-D17株;フィプロバクター・サクシノゲネスS85株;バチルス・セレウスATCC10987株;リステリア・イノキュア;ラクトバチルス・カゼイ;ラクトバチルス・ラムノーサスGG株;ラクトバチルス・サリバリウスUCC118株;ストレプトコッカス・アガラクティアエA909株;ストレプトコッカス・アガラクティアエNEM316株;ストレプトコッカス・アガラクティアエ2603株;ストレプトコッカス・ディスガラクティアエ亜種エクイシミリスGGS124株;ストレプトコッカス・エクイ亜種ズーエピデミカスMGCS10565株;ストレプトコッカス・ガロリティカスUCN34 uid46061株;ストレプトコッカス・ゴルドニイChallis subst CH1株;ストレプトコッカス・ミュータンスNN2025 uid46353株;ストレプトコッカス・ミュータンス;ストレプトコッカス・ピオゲネスM1 GAS株;ストレプトコッカス・ピオゲネスMGAS5005株;ストレプトコッカス・ピオゲネスMGAS2096株;ストレプトコッカス・ピオゲネスMGAS9429株;ストレプトコッカス・ピオゲネスMGAS10270株;ストレプトコッカス・ピオゲネスMGAS6180株;ストレプトコッカス・ピオゲネスMG

20

30

40

50

A S 3 1 5 株 ; ストレプトコッカス・ピオゲネス S S I - 1 株 ; ストレプトコッカス・ピオゲネス M G A S 1 0 7 5 0 株 ; ストレプトコッカス・ピオゲネス N Z 1 3 1 株 ; ストレプトコッカス・サーモフィルス ( *Streptococcus thermophiles* ) C N R Z 1 0 6 6 株 ; ストレプトコッカス・サーモフィルス ( *Streptococcus thermophiles* ) L M D - 9 株 ; ストレプトコッカス・サーモフィルス ( *Streptococcus thermophiles* ) L M G 1 8 3 1 1 株 ; クロストリジウム・ボツリヌム A 3 L o c h M a r e e 株 ; クロストリジウム・ボツリヌム B E k l u n d 1 7 B 株 ; クロストリジウム・ボツリヌム B a 4 6 5 7 株 ; クロストリジウム・ボツリヌム F L a n g e l a n d 株 ; クロストリジウム・セルロリティカム H 1 0 株 ; フィネゴルディア・マグナ A T C C 2 9 3 2 8 株 ; ユウバクテリウム・レクターレ A T C C 3 3 6 5 6 株 ; マイコプラズマ・ガリセプティカム ; マイコプラズマ・モービレ 1 6 3 K 株 ; マイコプラズマ・ペネトランス ; マイコプラズマ・シノビアエ 5 3 株 ; ストレプトバチルス・モニリフォルミス D S M 1 2 1 1 2 株 ; ブラジリゾビウム B T A i 1 株 ; ニトロバクター・ハンブルゲンシス X 1 4 株 ; ロドシードモナス・パルストリス B i s B 1 8 株 ; ロドシードモナス・パルストリス B i s B 5 株 ; バルビバクラム・ラバメンティボランス D S - 1 株 ; ディノロセオバクター・シバエ D F L 1 2 株 ; グルコンアセトバクター・ジアゾトロフィクス P a l 5 F A P E R J 株 ; グルコンアセトバクター・ジアゾトロフィクス P a l 5 J G I 株 ; アゾスピリルム B 5 1 0 u i d 4 6 0 8 5 株 ; ロドスピリラム・ルプラム A T C C 1 1 1 7 0 株 ; ジアフォロバクター T P S Y u i d 2 9 9 7 5 株 ; フエルミネフォロバクター・エイセニアエ E F 0 1 - 2 株 ; ナイセリア・メニンギティデス ( *Neisseria meningitidis* ) 0 5 3 4 4 2 株 ; ナイセリア・メニンギティデス ( *Neisseria meningitidis* ) a l p h a 1 4 株 ; ナイセリア・メニンギティデス ( *Neisseria meningitidis* ) Z 2 4 9 1 株 ; デスルホビブリオ・サレキシゲンス D S M 2 6 3 8 株 ; カンピロバクター・ジェジュニ亜種ドレイ 2 6 9 9 7 株 ; カンピロバクター・ジェジュニ 8 1 1 1 6 株 ; カンピロバクター・ジェジュニ ; カンピロバクター・ラリ R M 2 1 0 0 株 ; ヘリコバクター・ヘパティカス ; ウオリネラ・サクシノゲネス ; トルモナス・アウエンシス D S M 9 1 8 7 株 ; シュードアルテロモナス・アトランティカ T 6 c 株 ; シュワネラ・ペアレアナ A T C C 7 0 0 3 4 5 株 ; レジオネラ・ニューモフィラ P a r i s 株 ; アクチノバチルス・サクシノゲネス 1 3 0 Z 株 ; パスツレラ・ムルトシダ株 ; フランシセラ・ツラレンシス亜種ノビシダ U 1 1 2 株 ; フランシセラ・ツラレンシス亜種ホラルクティカ ; フランシセラ・ツラレンシス F S C 1 9 8 株 ; フランシセラ・ツラレンシス亜種ツラレンシス ; フランシセラ・ツラレンシス W Y 9 6 - 3 4 1 8 株 ; およびトレボネーマ・デンティコーラ A T C C 3 5 4 0 5 株。したがって、本開示の態様は、I I 型 C R I S P R システムに存在する C a s 9 タンパク質であって、ヌクレアーゼ欠損になっている、または本明細書に記載されるようなニッカーゼになっている C a s 9 タンパク質に関する。

## 【 0 0 7 2 】

C a s 9 タンパク質は文献中では当業者によって C s n 1 と呼ばれることがあり得る。本明細書に記載される実験対象である化膿レンサ球菌 ( *S. pyogenes* ) C a s 9 タンパク質配列を以下に示す。ここにその全体が参照により援用される D e l t c h e v a ら著、 N a t u r e 誌、第 4 7 1 卷、 6 0 2 ~ 6 0 7 頁 ( 2 0 1 1 年 ) を参照されたい。

## 【 0 0 7 3 】

MDKKYSIGLDIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETA  
 ATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLLEESFLVEEDKKHERHPIFG  
 NIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSD  
 VDKLEIQLVOTYNQLFEENPINASGVDAKAILSARLSKSRRLLENLIAQLPGEKKNGLFGN  
 LIJALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAI  
 LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYA  
 GYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNRDLLRKQRTFDNGSIPHQIHLGELH  
 AILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEE 10  
 VVDKGASAQSFIERMINTFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFL  
 SGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI  
 IKDKDFLDNEENEDILEDIVLTLLFEDREMIEERLKTYAHLFDDKVMKQLKRRRTGWG  
 RLSRKLINGIRDQSGKTILDFLKSDGFANRNMQLIHDDSLTFKEDIQKAQVSGQGDSL  
 HEHIANLAGSPAIIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRER  
 MKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDWDH  
 IVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVKKMKNYWRQLLNAKLITQRKFDSL  
 TKAERGGISELDKAGFIKRQLVETRQITKHVAQILDLSRMNTKYDENDKLIREVKVITLKS 20  
 KLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTLAIKKYPKLESEFVYGDYKVYDVR  
 K  
 MIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDF  
 ATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKYGGFDSPPTVA  
 YSVLVVAKVEKGKSKKLKSYKELLGITIMERSSEKNPIDFLEAKGYKEVKKDLIHKLPK  
 YSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLAHYEKLKGSPEDNEOKQLFVE

30  
 QHKHYLDEIIEQISEFSKRVILADANLDKVL SAYNKHRDKPIREQAENIHLFTLTNLGA  
 PAAFKYFDTTIDRKRYTSTKEVLDATLIHQSIITGLYETRIDLSQLGGD-

#### 【 0 0 7 4 】

本明細書に記載されるRNA誘導性ゲノム制御方法のある特定の態様によると、ヌクレアーゼ活性を低下させるように、実質的に低下させるように、または除去するようにCas9を改変する。一つの態様によると、RuvCヌクレアーゼドメインまたはHNHヌクレアーゼドメインを改変することによってCas9ヌクレアーゼ活性を低下させる、実質的に低下させる、または除去する。一つの態様によると、RuvCヌクレアーゼドメインを不活性化する。一つの態様によると、HNHヌクレアーゼドメインを不活性化する。一つの態様によると、RuvCヌクレアーゼドメインとHNHヌクレアーゼドメインを不活性化する。追加の態様によると、RuvCヌクレアーゼドメインとHNHヌクレアーゼドメインが不活性化されているCas9タンパク質が提供される。追加の態様によると、RuvCヌクレアーゼドメインとHNHヌクレアーゼドメインが不活性化されている限りにおいてヌクレアーゼ欠損Cas9タンパク質が提供される。追加の態様によると、RuvCヌクレアーゼドメインまたはHNHヌクレアーゼドメインのどちらかが不活性化され、それによって残りのヌクレアーゼドメインがヌクレアーゼ活性について活性型のままでいる、Cas9ニッカーゼが提供される。このように二本鎖DNAの一方のストランドだけが切断されるか、またはニックを入れられる。

#### 【 0 0 7 5 】

10

20

30

40

50

追加の態様によると、ヌクレアーゼ欠損 C a s 9 タンパク質を提供するために C a s 9 中の 1 つ以上のアミノ酸が改変されているかあるいは除去されているヌクレアーゼ欠損 C a s 9 タンパク質が提供される。一つの態様によると、それらのアミノ酸には D 1 0 と H 8 4 0 が含まれる。J i n k e ら著、S c i e n c e 誌、第 3 3 7 卷、8 1 6 ~ 8 2 1 頁 (2 0 1 2 年) を参照されたい。追加の態様によると、それらのアミノ酸には D 8 3 9 と N 8 6 3 が含まれる。一つの態様によると、D 1 0 、 H 8 4 0 、 D 8 3 9 および H 8 6 3 のうちの 1 つ以上または全てがヌクレアーゼ活性を低下させるか、実質的に低下させるか、または除去するアミノ酸で置換される。一つの態様によると、D 1 0 、 H 8 4 0 、 D 8 3 9 および H 8 6 3 のうちの 1 つ以上または全てがアラニンで置換される。一つの態様によると、D 1 0 、 H 8 4 0 、 D 8 3 9 のうちの 1 つ以上または全てが、アラニンなどの、ヌクレアーゼ活性を低下させるか、実質的に低下させるか、または除去するアミノ酸で置換された C a s 9 タンパク質は、ヌクレアーゼ欠損 C a s 9 または C a s 9 N と称され、当該タンパク質はヌクレアーゼ活性が低下若しくは削減されているか、またはヌクレアーゼ活性が検出のレベル内に無いか若しくは実質的に無い。この態様によると、C a s 9 N のヌクレアーゼ活性は公知のアッセイを用いて検出できない場合があり得る、すなわち、公知のアッセイの検出レベル未満であり得る。

## 【 0 0 7 6 】

一つの態様によると、前記ヌクレアーゼ欠損 C a s 9 タンパク質は、D N A に結合し且つ R N A によってガイドされる当該タンパク質の能力を保持する、そのホモログおよびオーソログを含む。一つの態様によると、前記ヌクレアーゼ欠損 C a s 9 タンパク質は、化膿レンサ球菌 ( S . p y o g e n e s ) 由来の天然 C a s 9 について示されている配列であって、D 1 0 、 H 8 4 0 、 D 8 3 9 および H 8 6 3 のうちの 1 つ以上または全てがアラニンで置換されている前記配列、および、その配列に対して少なくとも 3 0 % 、 4 0 % 、 5 0 % 、 6 0 % 、 7 0 % 、 8 0 % 、 9 0 % 、 9 5 % 、 9 8 % または 9 9 % の相同性を有するタンパク質配列であって、R N A 誘導型 D N A 結合タンパク質などのD N A 結合タンパク質である、前記配列を含む。

## 【 0 0 7 7 】

一つの態様によると、前記ヌクレアーゼ欠損 C a s 9 タンパク質は、R u v C ヌクレアーゼドメインと H N H ヌクレアーゼドメインのタンパク質配列を除いた化膿レンサ球菌 ( S . p y o g e n e s ) 由来の天然 C a s 9 について示されている配列、およびまた、その配列に対して少なくとも 3 0 % 、 4 0 % 、 5 0 % 、 6 0 % 、 7 0 % 、 8 0 % 、 9 0 % 、 9 5 % 、 9 8 % または 9 9 % の相同性を有するタンパク質配列であって、R N A 誘導型 D N A 結合タンパク質などのD N A 結合タンパク質である前記配列を含む。このように本開示の態様はD N A 結合を担う、例えばガイド R N A との共局在とD N A 結合を担う、タンパク質配列、及び当該タンパク質配列と相同なタンパク質配列を含み、且つ、ヌクレアーゼ欠損 C a s 9 タンパク質を作製するためには、R u v C ヌクレアーゼドメインと H N H ヌクレアーゼドメインは不活性化されるかまたは天然の C a s 9 タンパク質のタンパク質配列から除去されるかのどちらかであり得るので、( D N A 結合に必要とされない限り ) これらのドメインのタンパク質配列を含む必要が無い。

## 【 0 0 7 8 】

本開示の目的のため、C a s 9 に対する相同性を有する公知のタンパク質構造の中の金属配位性残基を図 4 A に示す。C a s 9 配列中の位置に基づいて残基が標識されている。左は P D B I D : 4 E P 4 の R u v C 構造 ( 青色 ) 。C a s 9 配列中の D 1 0 に対応する位置 D 7 は、M g イオン配位位置の中で強調されている。中央は P D B I D : 3 M 7 K ( オレンジ色 ) と 4 H 9 D ( 青緑色 ) の H N H エンドヌクレアーゼドメインの構造であり、3 M 7 K の配位している M g イオン ( 灰色の四角 ) と D N A ( 紫 ) とを含む。3 M 7 K 中の残基 D 9 2 および N 1 1 3 および 4 H 9 D の位置 D 5 3 および N 7 7 は C a s 9 のアミノ酸 D 8 3 9 および N 8 6 3 に対する配列相同性を有し、棒で示されている。右は作製されヌクレアーゼ活性について分析された突然変異体のリストである。野生型 C a s 9 、D 1 0 についてアラニンで置換した C a s 9 <sub>m 1</sub> ; D 1 0 と H 8 4 0 についてアラニン

で置換した Cas9<sub>m2</sub> ; D10、H840、およびD839についてアラニンで置換した Cas9<sub>m3</sub> ; 並びに、D10、H840、D839、およびN863についてアラニンで置換した Cas9<sub>m4</sub> 。

#### 【0079】

図4Bに示されるように、標的座位についてディープシーケンシングを行うと、Cas9 突然変異体 : m3 および m4 、および、それらのそれぞれの VP64 との融合体は、検出不可能なヌクレアーゼ活性を示した。gRNA 標的を区別する赤色の線と共に、ゲノム上の位置に対する突然変異頻度をプロットで示す。図4Cは図4Bのデータの高感度検査を示す図であり、突然変異ランドスケイプが未変更座位に類似のプロファイルを示すことが確認される。

10

#### 【0080】

一つの態様によると、転写活性化ドメインをヌクレアーゼ欠損 Cas9 またはガイドRNA のどちらかに連結することによってヒト細胞においての RNA 誘導性ゲノム制御を可能にする、遺伝子改変された Cas9-gRNA システムが提供される。本開示の一つの態様によると、1つ以上の転写制御タンパク質またはドメイン（そのような用語は互換的に使用される）がヌクレアーゼ欠損型 Cas9 または1つ以上のガイドRNA（gRNA）に結びつけられ、あるいはこれ以外のやり方で接続される。それらの転写制御ドメインは標的座位に対応する。したがって、本開示の態様は、転写制御ドメインを Cas9N または gRNA のどちらかに融合し、接続し、または結びつけることによりそのようなドメインを標的座位に局在させるための方法および材料を含む。

20

#### 【0081】

一つの態様によると、転写活性化可能な Cas9N 融合タンパク質が提供される。一つの態様によると、VP64 活性化ドメイン（ここにその全体が参照により援用される Zhang 著、Nature Biotechnology 誌、第29巻、149～153 頁（2011年）を参照のこと）が Cas9N の C 末端に結びつけられ、融合され、接続され、あるいはこれら以外のやり方で連結（tether）される。一つの方法によると、当該 Cas9N タンパク質により標的ゲノムDNA 部位に転写制御ドメインが提供される。一つの方法によると、転写制御ドメインに融合した Cas9N が1つ以上のガイドRNAと共に細胞内に提供される。融合した転写制御ドメインを有する Cas9N が、標的ゲノムDNA に、または標的ゲノムDNA 近傍に結合する。1つ以上のガイドRNA が、標的ゲノムDNA に、または標的ゲノムDNA 近傍に結合する。転写制御ドメインが標的遺伝子の発現を制御する。特定の態様によると、Cas9N-VP64 融合体は、プロモーター近辺の配列を標的とする gRNA と組み合わさった時に、レポーターコンストラクトの転写を活性化し、それによって RNA 誘導性転写を活性化した。

30

#### 【0082】

一つの態様によると、転写活性化可能な gRNA 融合タンパク質が提供される。一つの態様によると、VP64 活性化ドメインが gRNA に結びつけられ、融合され、接続され、あるいはこれら以外のやり方で連結（tether）される。一つの方法によると、その gRNA によりゲノムDNA 部位を標的とする転写制御ドメインが提供される。一つの方法によると、転写制御ドメインに融合した gRNA が Cas9N タンパク質と共に細胞内に提供される。当該 Cas9N は、標的ゲノムDNA に、または標的ゲノムDNA 近傍に結合する。転写制御タンパク質またはドメインと融合した、1つ以上のガイドRNA が、標的ゲノムDNA に、または標的ゲノムDNA 近傍に結合する。転写制御ドメインが標的遺伝子の発現を制御する。特定の態様によると、Cas9N タンパク質と、転写制御ドメインが融合した gRNA とが、レポーターコンストラクトの転写を活性化し、それによって RNA 誘導性転写を活性化した。

40

#### 【0083】

転写制御可能な gRNA 連結物は、無作為配列をその gRNA に挿入し、Cas9 機能についてアッセイすることによりその gRNA のどの領域が改造を許容するか特定することによって作製された。キメラ gRNA の crRNA 部分の 5' 末端または tracrR

50

N A 部分の 3' 末端のどちらかに無作為配列挿入物を有する g R N A は機能性を保持し、一方でキメラ g R N A の t r a c r R N A スキヤフォルド部分への挿入がなされると機能が喪失する。無作為塩基挿入に対する g R N A の柔軟性についてまとめた図 5 A ~ B を参考されたい。図 5 A は C a s 9 g R N A 活性を測定するための相同組換え ( H R ) アッセイの模式図である。図 5 B に示されるように、キメラ g R N A の c r R N A 部分の 5' 末端または t r a c r R N A 部分の 3' 末端のどちらかに無作為配列挿入物を有する g R N A は機能性を保持し、一方でキメラ g R N A の t r a c r R N A スキヤフォルド部分に挿入がなされると機能が失われる。g R N A 配列中の挿入の位置は赤色のヌクレオチドによって示されている。科学的理論に捉われることを望むことなく述べれば、5' 末端での無作為塩基挿入による活性の上昇は g R N A が長くなつて半減期が増したことに起因するものであり得る。

#### 【 0 0 8 4 】

V P 6 4 を g R N A に結合させるため、2 コピーの M S 2 バクテリオファージ・コートプロテイン結合性 R N A ステムループをその g R N A の 3' 末端に取り付けた。ここにその全体が参考により援用される F u s c o ら著、 C u r r e n t B i o l o g y 誌：第 C B 1 3 卷、 1 6 1 ~ 1 6 7 頁 ( 2 0 0 3 年 ) を参考されたい。これらのキメラ g R N A は C a s 9 N と M S 2 - V P 6 4 融合タンパク質と共に発現した。3 種類全ての要素の存在下でレポーターコンストラクトの配列特異的転写活性化が観察された。

#### 【 0 0 8 5 】

図 1 A は R N A 誘導性転写活性化の模式図である。図 1 A に示されるように、転写活性化可能な C a s 9 N 融合タンパク質を作製するため、 V P 6 4 活性化ドメインを C a s 9 N の C 末端に直接連結した。図 1 B に示されるように、転写活性化可能な g R N A 連結物を作製するため、2 コピーの M S 2 バクテリオファージ・コートプロテイン結合性 R N A ステムループをその g R N A の 3' 末端に取り付けた。これらのキメラ g R N A は C a s 9 N と M S 2 - V P 6 4 融合タンパク質と共に発現した。図 1 C は転写活性化をアッセイするために使用されたレポーターコンストラクトのデザインを示している。それらの 2 種類のレポーターは別個の g R N A 標的部位を有し、且つ、対照 T A L E - T F 標的部位を共有する。図 1 D に示されるように、蛍光活性化細胞選別 ( F A C S ) および免疫蛍光分析 ( I F ) の両方によってアッセイすると、 C a s 9 N - V P 6 4 融合体は R N A 誘導性転写活性化を示す。具体的には、対照 T A L E - T F が両方のレポーターを活性化した一方で、 C a s 9 N - V P 6 4 融合体は g R N A 配列特異的にレポーターを活性化する。図 1 E に示されるように、3 種類全ての要素、すなわち、 C a s 9 N 、 M S 2 - V P 6 4 、および適切な M S 2 アブタマー結合部位を有する g R N A が存在するときにのみ、レポーターコンストラクトの g R N A 配列特異的転写活性化が F A C S および I F の両方によって観察された。

#### 【 0 0 8 6 】

ある特定の態様によれば、 C a s 9 N 、1 つ以上の g R N A および転写制御タンパク質またはドメインを使用して内在性遺伝子を制御する方法が提供される。一つの態様によると、内在性遺伝子はあらゆる所望の遺伝子 ( ここでは標的遺伝子と呼ばれる ) であり得る。一つの例示的な態様によると、制御の標的となる遺伝子には、多能性の維持に関与する両方とも厳密に制御されている遺伝子である Z F P 4 2 ( R E X 1 ) および P O U 5 F 1 ( O C T 4 ) が含まれた。図 1 F に示されるように、転写開始部位の上流の約 5 k b 長の D N A 鎖 ( D N a s e 高感受性部位が緑色で強調されている ) を標的とする 1 0 種類の g R N A を R E X 1 遺伝子用にデザインした。プロモーター・ルシフェラーゼレポーターコンストラクト ( ここにその全体が参考により援用される T a k a h a s h i ら著、 C e l 1 誌、第 1 3 1 卷、 8 6 1 ~ 8 7 2 頁 ( 2 0 0 7 年 ) を参考のこと ) を用いて転写活性化をアッセイするか、または内在性遺伝子の q P C R により転写活性化を直接的にアッセイした。

#### 【 0 0 8 7 】

図 6 A ~ C は C a s 9 N - V P 6 4 を使用する R N A 誘導性 O C T 4 制御に関する。図

10

20

30

40

50

6 A に示されるように、転写開始部位の上流の約 5 kb 長の DNA 鎖を標的とする 21 種類の g RNA を OCT4 遺伝子用にデザインした。 DNase 高感受性部位が緑色で強調されている。図 6 B はプロモーター・ルシフェラーゼレポーターコンストラクトを使用して転写活性化を示している。図 6 C は内在性遺伝子の qPCR により直接的に転写活性化を示している。個々の g RNA の導入によって転写が中程度に促進された一方で、複数の g RNA が相乗的に作用して堅固な複数倍の転写活性化を促進した。

#### 【0088】

図 7 A ~ C は Cas9N、MS2 - VPF4 および g RNA + 2 × MS2 アプタマーを使用する RNA 誘導性 REX1 制御に関する。図 7 A に示されるように、転写開始部位の上流の約 5 kb 長の DNA 鎖を標的とする 10 種類の g RNA を REX1 遺伝子用にデザインした。 DNase 高感受性部位が緑色で強調されている。図 7 B はプロモーター・ルシフェラーゼレポーターコンストラクトを使用して転写活性化を示している。図 7 C は内在性遺伝子の qPCR により直接的に転写活性化を示している。個々の g RNA の導入によって転写が中程度に促進された一方で、複数の g RNA が相乗的に作用して堅固な複数倍の転写活性化を促進した。一つの態様では、 g RNA 上に 2 × MS2 アプタマーが存在しないことで転写活性化が引き起こされない。ここにその全体が参照によりそれぞれ援用される Maeder ら著、 Nature Methods 誌、第 10 卷、 243 ~ 245 頁 (2013 年) および Perez-Pinera ら著、 Nature Methods 誌、第 10 卷、 239 ~ 242 頁 (2013 年) を参照されたい。

#### 【0089】

したがって、方法は、標的遺伝子の発現を制御するために Cas9N タンパク質および転写制御タンパク質またはドメインと複数のガイド RNA を使用することに関する。

#### 【0090】

Cas9 連結アプローチと g RNA 連結アプローチの両方が効果的であり、前者は約 1.5 ~ 2 倍高い能力を示した。この差異は 3 要素複合体集合と対照的な 2 要素複合体集合の必要条件に起因するようである。しかしながら、 g RNA 連結アプローチは、各 g RNA が異なる RNA - タンパク質間相互作用対を利用する限り、別個の g RNA によって様々なエフェクタードメインを動員することを原理的に可能にする。ここにその全体が参照により援用される Karyer-Bibens ら著、ヨーロッパ細胞生物学会贊助 Biology of the Cell 誌、第 100 卷、 125 ~ 138 頁 (2008 年) を参照されたい。本開示の一つの態様によると、特異的なガイド RNA および包括的な (generic) Cas9N タンパク質、すなわち、異なる標的遺伝子に対する同一または類似の Cas9N タンパク質を使用して、異なる標的遺伝子を制御することができる。一つの態様によると、同一または類似の Cas9N を使用する多重遺伝子制御の方法が提供される。

#### 【0091】

本開示の方法はヒト細胞の多重遺伝子操作およびエピジェネティック操作を提供するために本明細書に記載されるような Cas9N タンパク質とガイド RNA を使用して標的遺伝子を編集することにも関する。 Cas9 g RNA ターゲティングが論点であるので (ここにその全体が参照により援用される Jiang ら著、 Nature Biotechnology 誌、第 31 卷、 233 ~ 239 頁 (2013 年) を参照のこと)、非常に多種多様な標的配列のバリエーションに対する Cas9 親和性を徹底的に調査するための方法を提供する。したがって、本開示の態様は、天然のヌクレアーゼ活性型 Cas9 を使用する特異性試験によってもたらされる dsDNA 切断毒性および突然変異修復により生じる厄介な問題を回避しつつヒト細胞における Cas9 ターゲティングの直接的ハイスクループット・リードアウトを提供する。

#### 【0092】

その他の本開示の態様は概ね標的遺伝子の転写制御のために DNA 結合タンパク質または DNA 結合システムを使用することに関する。当業者は本開示に基づいて例示的な DNA 結合システムを容易に特定する。そのような DNA 結合システムは、天然の Cas9 タ

10

20

30

40

50

ンパク質にあるようないかなるヌクレアーゼ活性も有する必要が無い。したがって、そのようなDNA結合システムはヌクレアーゼ活性を不活性化する必要が無い。一つの例示的なDNA結合システムはT A L Eである。ゲノム編集ツールとして通常はT A L E - F o k I ダイマーが使用され、ゲノム制御にはT A E L - V P 6 4 融合体が非常に有効であることが示されている。一つの態様に従い、図2 Aに示されている方法を用いてT A L Eの特異性を評価した。ライプラリーの各要素がdT o m a t o 蛍光タンパク質を発現するミニマルプロモーターを含むコンストラクトライプラリーを設計する。転写開始部位mの下流に24bpの(A/C/G)無作為転写物タグを挿入し、一方で2つのTF結合部位をプロモーターの上流に配置する。一つは全てのライプラリー要素によって共有される一定DNA配列であり、二つ目のものは「バイアスがかかった」結合部位ライプラリーを有する可変的特徴であり、それらの結合部位はプログラム可能DNAターゲティング複合体が結合するように設計された標的配列から離れて多数の突然変異の組合せを提示する配列の大きなコレクションに及ぶように設計されている。これは、標的配列ヌクレオチドが79%の頻度で現れ、他のそれぞれのヌクレオチドが7%の頻度で現れるようなヌクレオチド出現頻度を各位置で有するように設計された縮重オリゴヌクレオチドを使用して達成される。ここにその全体が参照により援用されるPatwardhanら著、Nature Biotechnology誌、第30巻、265~270頁(2012年)を参照されたい。次にレポーターライプラリーを配列解析して24bpのdT o m a t o 転写物タグとライプラリー要素中のそれらの対応する「バイアスがかかった」標的部位との間の関係を明らかにする。異なる標的間でのタグの共有が非常にまれであることがそれらの転写物タグの大きな多様性によって確証され、一方で標的配列のバイアスがかかった構成は、ほとんど突然変異を有しない部位がより多くの突然変異を有する部位よりも多くのタグと結合することを意味する。次に、dT o m a t o レポーター遺伝子の転写が共有DNA部位に結合するように設計された対照TF、または標的部位に結合するように設計された標的TFのどちらかによって促進される。促進を受けた細胞に対してRNA配列解析を実施することにより発現した各転写物タグの量を各試料で測定し、次に前もって作製された関連表を使用してそれらの対応する結合部位にその転写物タグをマップする。対照TFの結合部位は全てのライプラリー要素にわたって共有されているので対照TFは全てのライプラリー構成要素を等しく刺激することが予期され、一方で標的TFは、その標的TFが優先的に標的とする構成要素に対して、発現構成要素の分布状態をゆがめると予期される。対照TFについて得られたタグ数で標的TFについて得られたタグ数を除算することにより各結合部位について正規化発現レベルを計算するために、ステップ5においてこの仮説を利用する。

#### 【0093】

図2 Bに示されるように、Cas9-gRNA複合体は平均してその標的配列中の1~3個の突然変異に対して許容的であることがその複合体のターゲティングランドスケイプより明らかである。図2 Cに示されるように、Cas9-gRNA複合体は、PAM配列に局在しているものを除いて点突然変異にもほとんど感受性が無い。特に、化膿レンサ球菌(*S. pyogenes*)のCas9の予想されるPAMがNGGばかりでなくNAGでもあることがこのデータより明らかである。しかしながら、図2 Dに示されるように、これらのミスマッチがさらに8~10塩基ほどgRNA標的配列の3'末端に近くに局在するときにのみ、2塩基のミスマッチの導入によってCas9-gRNA複合体活性は顕著に損なわれる(ヒートプロットでは標的配列の位置は5'末端より1~23と標識されている)。

#### 【0094】

本明細書に記載される転写特異性アッセイを用いて別の広範に使用されるゲノム編集ツールであるT A L Eドメインの突然変異許容性を決定した。図2 Eに示されるように、18-merのT A L EについてのT A L Eオフ・ターゲティングデータより、18-merのT A L Eは平均してその標的配列中の1~2個の突然変異を許容することができ、且つ、その標的中の3塩基ミスマッチ変異体の大多数を活性化することができないことが明

10

20

30

40

50

らかである。図 2 F に示されるように、18-mer の TALE は、Cas9-gRNA 複合体と同様に、その標的中の 1 塩基のミスマッチにほとんど感受性が無い。図 2 G に示されるように、18-mer の TALE の活性は 2 塩基のミスマッチの導入によって顕著に損なわれる。TALE の活性はミスマッチがその標的配列の 5' 末端に近くなるほどそれらのミスマッチに対して感受性が高くなる（ヒートプロットでは標的配列の位置は 5' 末端より 1 ~ 18 と標識されている）。

#### 【0095】

様々なサイズの TALE によるターゲティングのランドスケイプを評価することに関する図 10 A ~ C の主題であるヌクレアーゼアッセイの標的化実験を用いて結果を確認した。図 10 A に示されるように、18-mer の TALE はそれらの標的配列中の複数の突然変異を許容することがヌクレアーゼ介在性 HR アッセイを用いて確認された。図 10 B に示されるように、図 2 に記載されているアプローチを用いて 3 つの異なるサイズ（18-mer、14-mer および 10-mer）の TALE のターゲティングランドスケイプを分析した。TALE が短くなるほど（14-mer および 10-mer）次第にそれらの TALE のターゲティングが特異的になり、活性もほぼ一桁低くなる。図 10 C および 10 D に示されるように、10-mer の TALE はほぼ 1 塩基ミスマッチの感度を示し、2 つのミスマッチを有する標的に対してほとんど全ての活性を失う（ヒートプロットでは標的配列の位置は 5' 末端より 1 ~ 10 と標識されている）。まとめると、ゲノム工学用途においてより短い TALE を設計するほどより高い特異性を得ることができ、一方で TALE ヌクレアーゼ用途において標的外効果を回避するために Fok I の二量体化の必要性が重要であることをこれらのデータは示している。それぞれここにその全体が参照により援用される Kim ら著、Proceedings of the National Academy of Sciences of the United States of America 誌、第 93 卷、1156 ~ 1160 頁（1996 年）および Pattanayak ら著、Nature Methods 誌、第 8 卷、765 ~ 770 頁（2011 年）を参照されたい。

#### 【0096】

図 8 A ~ C は、実験データを例に示されている正規化発現レベルの算出のための高レベルの特異性分析処理フローに関する。図 8 A に示されるように、レポーター遺伝子転写物に組み込まれる分布状態にバイアスがかかっている結合部位配列と無作為配列 24 bp タグを有するコンストラクトライブラリーを構築する（上段）。転写されたタグは、縮合性が高いので、多くが 1 つずつの Cas9 結合配列または TALE 結合配列に対応することになるはずである。コンストラクトライブラリーを配列解析して（第 3 段、左）どのタグがどの結合部位に対応するか確定し、結合部位と転写されたタグの間の関連表を作製する（第 4 段、左）。様々な結合部位について構築された複数のコンストラクトライブラリーについて、ライブラリーバーコードを使用して配列解析してもよい（ここでは水色および明黄色で表される；第 1 ~ 4 段、左）。次にコンストラクトライブラリーを細胞集団に形質移入し、一組の様々な Cas9 / gRNA または TALE 転写因子をそれらの集団の試料中でガイドする（第 2 段、右）。1 つの試料は、コンストラクト内の一定の結合部位配列を標的とする一定の TALE 活性化因子によって常にガイドされる（上段、緑色のボックス）。この試料は陽性対照として機能する（緑色の試料、+ 記号によっても表される）。次にそれらのガイド試料内のレポーター mRNA 分子から作られた cDNA を配列解析し、分析して試料中の各タグのタグ数を得る（第 3 段および第 4 段、右）。コンストラクトライブラリーの配列解析と同様に、試料バーコードを付加することにより陽性対照を含む複数の試料を一緒に配列解析し、分析する。ここで明赤色は、配列解析され、陽性対照（緑色）と一緒に分析された、1 つの非対照試料を表す。各リードには転写されたタグだけが現れ、コンストラクトの結合部位は現れないので、次にコンストラクトライブラリーの配列解析から得られた結合部位とタグの間の関連表を使用して各試料の中の各結合部位から発現するタグの総数を算出する（第 5 段）。次に各非陽性対照試料についての総数を陽性対照試料において得られた総数で除算することにより各非陽性対照試料についての総

10

20

30

40

50

数を各結合部位についての正規化発現レベルに変換する。ミスマッチ数毎の正規化発現レベルのプロットの例が図2Bと2E、および図9Aと図10Bに提供されている。誤りがちなタグ、コンストラクトライブラリーと結合できないタグ、および明らかに複数の結合部位によって共有されるタグに対する幾つかのレベルのフィルタリングは、この全体的処理フローに含まれていない。図8Bはバイアスがかかったコンストラクトライブラリー内で生じたミスマッチ数毎の結合部位のパーセンテージの例示的な分布状態を示している。左：理論的分布状態。右：実際のT A L E コンストラクトライブラリーから観察された分布状態。図8Cはミスマッチ数毎の結合部位に集まったタグの数のパーセンテージの例示的な分布状態を示している。左：陽性対照試料から観察された分布状態。右：対照ではないT A L E がガイドされた試料から観察された分布状態。陽性対照T A L E がコンストラクト内の一定の部位に結合するので、集まったタグ数の分布状態は図8Bにおける結合部位の分布状態をよく反映し、一方でミスマッチが少ない部位ほど高い発現レベルをガイドするのでその分布状態は非対照T A L E 試料について左方にゆがめられている。下：標的T Fについて得られたタグ数を対照T Fについて得られたタグ数で除算してこれらの間の相対的濃縮度を計算することによって平均発現レベルと標的部位における突然変異数との間の関係が明らかになる。

#### 【0097】

これらの結果は異なるC a s 9 g R N A 複合体を使用して作成された特異性データによってさらに再確認される。図9Aに示されるように、異なるC a s 9 g R N A 複合体がその標的配列中の1～3個の突然変異に対して許容的である。図9Bに示されるように、そのC a s 9 g R N A 複合体は、P A M配列に局在しているものを除いて点突然変異にもほとんど感受性が無い。しかしながら、図9Cに示されるように、2塩基のミスマッチの導入によって活性が顕著に損なわれる（ヒートプロットでは標的配列の位置は5'末端より1～23と標識されている）。図9Dに示されるように、化膿レンサ球菌（*S. pyogenes*）のC a s 9の予想されるP A MがN G Gであり、またN A Gであることがスクレアーゼ介在性H R アッセイを用いて確認された。

#### 【0098】

ある特定の態様によると、本明細書に記載される方法に従って結合特異性が増加する。複数の複合体の間の相乗作用がC a s 9 N V P 6 4による標的遺伝子の活性化の要因であり、個々の標的外結合事象の効果は最小限であるべきなのでC a s 9 Nの転写制御用途は自然と非常に特異的である。一つの態様によると、ゲノム編集の方法の中でオフセット・ニックを使用する。ニックの大多数がN H E J事象を引き起こすことはほとんどないの（ここにその全体が参照により援用されるC e r t o ら著、N a t u r e M e t h o d s誌、第8巻、671～676頁（2011年）を参照のこと）、標的外ニック形成の効果が最小になる。対照的に、オフセット・ニックを誘導して二本鎖切断（D S B）を生じさせることは遺伝子破壊の誘導において非常に有効である。ある特定の態様によると、5'オーバーハングは3'オーバーハングと比較してより有意なN H E J事象を引き起こす。同様に、H R事象の総数は5'オーバーハングが生じたときよりも著しく少ないが、3'オーバーハングはN H E J事象よりもH R事象を好む。以上より、相同組換えのためにニックを用い、二本鎖切断を生じさせるためにオフセット・ニックを用いて標的外C a s 9 g R N A活性の効果を最小限にする方法が提供される。

#### 【0099】

図3A～Cは多重オフセット・ニック形成およびガイドR N Aとの標的外結合を減少させるための方法に関する。図3Aに示されるように、トラフィックライトレポーターを使用して標的化ニック導入時または標的化切断導入時のH R事象とN H E J事象を同時にアッセイした。H D R経路により解消されたD N A切断事象はG F P配列を回復し、一方で突然変異誘発性N H E JはG F Pを読み替から外し、且つ、下流のm C h e r r y配列を読み替に入れるフレームシフトの原因になる。そのアッセイについて、200 b p長のD N A鎖を範囲とする14種類のg R N A、すなわち、センスストランドを標的とする7種類（U 1～7）とアンチセンスストランドを標的とする7種類（D 1～7）をデザインし

た。相補ストランドにニックを入れる Cas9D10A 突然変異体を使用して異なる二通りの gRNA の組合せを用いて一連の計画的 5' オーバーハングまたは 3' オーバーハングを誘導した ( それらの 14 種類の gRNA のニック形成部位が示されている ) 。図 3B に示されるように、オフセット・ニックを誘導して二本鎖切断 ( DSB ) を生じさせることは遺伝子破壊の誘導にとって非常に有効である。特に、5' オーバーハングを生じさせるオフセット・ニックは 3' オーバーハングと比べてより多くの NHEJ 事象を引き起こす。図 3C に示されるように、3' オーバーハングの生成は NHEJ 事象よりも HR 事象の比率にとっても好都合であるが、HR 事象の総数は 5' オーバーハングが生じたときよりも著しく少ない。

## 【 0100 】

10

図 11A ~ B は Cas9D10A ニッカーゼ介在性 NHEJ に関する。図 11A に示されるように、トラフィックライトレポーターを使用して標的化ニック導入時または標的化切斷導入時の NHEJ 事象をアッセイした。簡単に説明すると、DNA 切断事象の導入時にその切断が突然変異誘発性 NHEJ を誘発すると GFP が読み枠から外れ、且つ、下流の mCherry 配列が読み枠に入って赤色の蛍光が生じる。200 bp 長の DNA 鎖を範囲とする 14 種類の gRNA 、すなわち、センスストランドを標的とする 7 種類 ( U1 ~ 7 ) とアンチセンスストランドを標的とする 7 種類 ( D1 ~ 7 ) をデザインした。図 11B に示されるように、全ての標的について DSB と堅固な NHEJ を引き起こす野生型 Cas9 と異なり ( Cas9D10A 突然変異体を用いると ) 大半のニックはほとんど NHEJ 事象を引き起こさないことが観察された。14 か所の部位全てが連続的な 200 bp 長の DNA 鎖の範囲内に位置し、10 倍を超えるターゲティング効率の差が観察された。

20

## 【 0101 】

30

ある特定の態様に従い、細胞において標的核酸の発現を制御する方法であって、1つ以上の、2つ以上のまたは複数の外来核酸を前記細胞に導入することを含む前記方法を本明細書に記載する。細胞に導入されるそれらの外来核酸は単数のガイド RNA または複数のガイド RNA 、単数のヌクレアーゼ欠損 Cas9 タンパク質または複数のヌクレアーゼ欠損 Cas9 タンパク質と転写制御因子タンパク質またはドメインをコードする。また、ガイド RNA 、ヌクレアーゼ欠損 Cas9 タンパク質および転写制御因子タンパク質またはドメインは、当該ガイド RNA 、ヌクレアーゼ欠損 Cas9 タンパク質および転写制御因子タンパク質またはドメインが DNA に結合し、標的核酸の発現を制御する限り、当業者によって理解されるように共局在複合体と呼ばれる。ある特定の追加の態様によると、細胞に導入される前記外来核酸は単数のガイド RNA または複数のガイド RNA と Cas9 タンパク質ニッカーゼとをコードする。また、ガイド RNA と Cas9 タンパク質ニッカーゼは、当該ガイド RNA と Cas9 タンパク質ニッカーゼが DNA に結合し、標的核酸にニックを入れる限り、当業者によって理解されるように共局在複合体と呼ばれる。

## 【 0102 】

40

本開示による細胞には本明細書に記載されるように外来核酸を導入することができ、発現させることができるあらゆる細胞が含まれる。本明細書に記載される本開示の基本的な考えは細胞種によって限定されないことが理解されるべきである。本開示による細胞には真核細胞、原核細胞、動物細胞、植物細胞、真菌細胞、古細菌細胞、真正細菌細胞などが含まれる。細胞には酵母細胞、植物細胞、および動物細胞などの真核細胞が含まれる。特定の細胞には哺乳類細胞が含まれる。さらに、細胞には標的核酸を制御するのに有益である、または望ましいあらゆる細胞が含まれる。そのような細胞には疾患または有害な健康状態を引き起こす特定のタンパク質の発現に欠損を有する細胞が含まれ得る。当業者にはそのような疾患または有害な健康状態は容易にわかる。本開示によれば、本明細書に記載される方法と、標的核酸の上方制御と特定のタンパク質の対応的な発現を引き起こす転写活性化因子とによって、当該特定のタンパク質の発現を担う核酸が標的とされ得る。この場合、本明細書に記載される方法は治療処置を提供する。

## 【 0103 】

50

標的核酸には、本明細書に記載される共局在複合体が制御またはニックの形成のどちらかにとて有用であり得るあらゆる核酸配列が含まれる。標的核酸には遺伝子が含まれる。本開示の目的のため、二本鎖DNAなどのDNAは標的核酸を含むことができ、共局在複合体はその共局在複合体が標的核酸に対する所望の作用を有し得るようにその標的核酸で、またはそれに隣接して、またはその近傍でそのDNAに結合するか、あるいはこれら以外のやり方でそのDNAと共局在することができる。そのような標的核酸には内在性の（または天然の）核酸および外來性の（または外來の）核酸が含まれ得る。当業者は標的核酸を含むDNAに共局在するガイドRNAおよびCas9タンパク質を本開示に基づいて容易に特定またはデザインすることができる。当業者は標的核酸を含むDNAに同様に共局在する転写制御因子タンパク質またはドメインをさらに特定することができる。DNAにはゲノムDNA、ミトコンドリアDNA、ウイルスDNAまたは外來性DNAが含まれる。

#### 【0104】

外來核酸（すなわち、細胞の天然の核酸組成物の一部ではない核酸）を、当業者に知られている導入のためのあらゆる方法を用いて細胞に導入してよい。そのような方法には形質移入、形質導入、ウイルス形質導入、微量注入、リポフェクション、ヌクレオフェクション、ナノ粒子ボンバードメント、遭影室転換、接合などが含まれる。当業者は容易に特定可能な文献を用いてそのような方法を容易に理解し、適用する。

#### 【0105】

転写活性化因子である転写制御因子タンパク質またはドメインにはVP16、VP64、および本開示に基づいて当業者によって容易に特定可能な他のものが含まれる。

#### 【0106】

疾患および有害な健康状態は特定のタンパク質の発現の異常な喪失を特徴とするものである。そのような疾患または有害な健康状態は特定のタンパク質の上方制御によって治療され得る。以上より、疾患または有害な健康状態を治療する方法であって、本明細書に記載される共局在複合体が標的核酸を含むDNAに会合その他のやり方で結合し、共局在複合体の転写活性化因子が標的核酸の発現を上方制御する前記方法を提供する。例えば、褐色脂肪分化および代謝取込みの増加を促進するPRDM16と他の遺伝子の上方制御を用いてメタボリックシンドロームまたは肥満を治療することができる。抗炎症性遺伝子の活性化は自己免疫疾患および心血管疾患において有用である。腫瘍抑制遺伝子の活性化は癌の治療に有用である。当業者は本開示に基づいてそのような疾患および有害な健康状態を容易に特定する。

#### 【0107】

後続の実施例を本開示の代表として示す。これらの実施形態および他の同等の実施形態は本開示、図面および添付されている特許請求の範囲を考慮して明らかであるので、これらの実施例は本開示の範囲を限定するものと解釈されてはならない。

#### 【実施例】

#### 【0108】

#### [実施例1]

##### Cas9突然変異体

Cas9のRuvCドメインとHNHドメインの天然の活性を除去することができるCas9の候補突然変異を特定するために、公知の構造を有するCas9と相同である配列を検索した。HHpred（ワールド・ワイド・ウェップサイト・ツールキット、tuebingen.mpg.de/hhpred）を使用して全長タンパク質データバンク（2013年1月）に対してCas9の全長配列をクエリ配列とした。この検索によってCas9のHNHドメインに対する顕著な配列相同性を有する2つの異なるHNHエンドヌクレアーゼであるPadエンドヌクレアーゼと推定エンドヌクレアーゼ（それぞれPDB ID: 3M7Kと4H9D）が返された。これらのタンパク質を調査してマグネシウムイオンの配位に関わる残基を見つけ出した。次に、それらの対応する残基をCas9に対する配列アラインメントの中で特定した。Cas9の中の同じ種類のアミノ酸と整列する

10

20

30

40

50

各構造中の 2 つの M g 配位性側鎖を特定した。それらは 3 M 7 K の D 9 2 と N 1 1 3 、および 4 H 9 D の D 5 3 と N 7 7 であった。これらの残基は C a s 9 の D 8 3 9 と N 8 6 3 に対応した。 P a d 残基 D 9 2 と N 1 1 3 のアラニンへの突然変異によってそのヌクレアーゼが触媒的に欠損になることも報告された。この分析に基づいて C a s 9 突然変異 D 8 3 9 A と N 8 6 3 A を作製した。さらに、 H H p r e d は C a s 9 とサーマス・サーモフィルスの R u v C ( P D B I D : 4 E P 4 ) の N 末端との間の相同性も予測する。この配列アラインメントは C a s 9 中の R u v C ドメインの機能を除去する以前に報告された突然変異 D 1 0 A を含む。この突然変異を適切な突然変異として確認するため、前述したように金属結合性残基を決定した。 4 E P 4 では D 7 がマグネシウムイオンを配位するのに機能する。この位置は C a s 9 D 1 0 に対応する配列相同性を有し、この突然変異が金属結合の除去に役立ち、結果として C a s 9 の R u v C ドメインからの触媒活性の除去に役立つことを確認する。 10

#### 【 0 1 0 9 】

#### 【 実施例 2 】

##### プラスミド構築

Q u i k c h a n g e キット ( アジレントテクノロジーズ社 ) を使用して C a s 9 突然変異体を作製した。標的 g R N A 発現コンストラクトは ( 1 ) I D T から個々の g B l o c k として直接注文され、 p C R - B l u n t I I - T O P O ベクター ( インビトロジェン ) にクローン化されるか、または ( 2 ) G e n e w i z によってカスタム合成されるか、または ( 3 ) オリゴヌクレオチドの G i b s o n 構築を用いて g R N A クローニングベクターに構築された ( プラスミド番号 4 1 8 2 4 ) 。 A d d g e n e 社の E G I P レンチベクターに構築される、終止コドンを有する G F P 配列と適切な断片の融合 P C R 構築によって切断型 G F P を伴う H R レポーター・アッセイ用のベクターを構築した ( プラスミド番号 2 6 7 7 7 ) 。次にこれらのレンチベクターを使用して G F P レポーター安定株を構築した。この研究で使用された T A L E N は標準的なプロトコルを用いて構築された。ここにその全体が参考により援用される S a n j a n a ら著、 N a t u r e P r o t o c o l s 誌、第 7 卷、 1 7 1 ~ 1 9 2 頁 ( 2 0 1 2 年 ) を参照されたい。標準的 P C R 融合プロトコルの手法を用いて C a s 9 N と M S 2 V P 6 4 の融合を実施した。 O C T 4 および R E X 1 用のプロモータールシフェラーゼコンストラクトは A d d g e n e 社から入手した ( プラスミド番号 1 7 2 2 1 および プラスミド番号 1 7 2 2 2 ) 。 20 30

#### 【 0 1 1 0 】

#### 【 実施例 3 】

##### 細胞培養および形質移入

1 0 % ウシ胎児血清 ( F B S 、インビトロジェン ) 、ペニシリン / ストレプトマイシン ( p e n / s t r e p 、インビトロジェン ) 、および非必須アミノ酸 ( N E A A 、インビトロジェン ) を添加した高グルコース・ダルベッコ改変イーグル培地 ( D M E M 、インビトロジェン ) 中で H E K 2 9 3 T 細胞を培養した。加湿恒温器中に 3 7 および 5 % の C O 2 で細胞を維持した。 40

#### 【 0 1 1 1 】

ヌクレアーゼアッセイを伴う形質移入は次の通りであった。製造業者のプロトコルに従ってリポフェクタミン 2 0 0 0 を使用して 0 . 4 × 1 0 6 個の細胞に 2 μ g の C a s 9 プラスミド、 2 μ g の g R N A および / または 2 μ g の D N A ドナープラスミドを形質移入した。形質移入から 3 日後に細胞を回収し、 F A C S によって分析するか、またはゲノム切断の直接的アッセイのために D N A イージー・キット ( Q i a g e n ) を使用して約 1 × 1 0 6 個の細胞のゲノム D N A を抽出した。これらのために P C R を実施して前記細胞に由来するゲノム D N A を用いてターゲティング領域を増幅し、 M i S e q パーソナルシーケンサー ( I l l u m i n a 社 ) により 2 0 0 , 0 0 0 リードを超える包括度でアンブリコンをディープシーケンスした。シーケンシングデータを分析して N H E J 効率を推定した。

#### 【 0 1 1 2 】

転写活性化アッセイを伴う形質移入については、 $0.4 \times 10^6$  個の細胞に(1)  $2 \mu\text{g}$  の Cas9N VP64 プラスミド、 $2 \mu\text{g}$  の gRNA および / または  $0.25 \mu\text{g}$  のレポーターコンストラクトを形質移入するか、または(2)  $2 \mu\text{g}$  の Cas9N プラスミド、 $2 \mu\text{g}$  の MS2-VP64、 $2 \mu\text{g}$  の gRNA -  $2 \times$  MS2 アプタマー および / または  $0.25 \mu\text{g}$  のレポーターコンストラクトを形質移入した。形質移入から 24 ~ 48 時間後に細胞を回収し、FACS または免疫蛍光方法を用いてアッセイするか、またはそれらの全 RNA を抽出し、その後にこれらの RNA を RT-PCR により分析した。ここで、OCT4 と REX1 についてインビトロジエンの標準的 tagman プローブを使用し、各試料の正規化を GAPDH に対して実施した。

## 【0113】

10

Cas9 gRNA 複合体と TALE の特異性プロファイルについての転写活性化アッセイを伴う形質移入については、 $0.4 \times 10^6$  個の細胞に(1)  $2 \mu\text{g}$  の Cas9N VP64 プラスミド、 $2 \mu\text{g}$  の gRNA および  $0.25 \mu\text{g}$  のレポータライブラリーを形質移入するか、または(2)  $2 \mu\text{g}$  の TALE-TF プラスミドおよび  $0.25 \mu\text{g}$  のレポータライブラリーを形質移入するか、または(3)  $2 \mu\text{g}$  の対照 TF プラスミドおよび  $0.25 \mu\text{g}$  のレポータライブラリーを形質移入した。形質移入から 24 時間後に(レポーターの促進が飽和モード状態であることを回避するため)細胞を回収した。RNA イージー・プラス・キット (Qiagen) を使用して全 RNA 抽出を実施し、Superscript-III (インビトロジエン) を使用して標準的 RT-PCR を実施した。転写物タグの標的化 PCR 増幅により次世代シーケンシング用のライブラリーを作製した。

20

## 【0114】

## [実施例4]

Cas9-TF および TALE-TF レポーターの発現レベルの計算のための数理分析および配列分析

この処理のための高レベルの論理フローが図 8A に示されており、その他の詳細がここに提供される。コンストラクトライブラリーの組成の詳細については図 8A (第 1 段) および 8B を参照されたい。

## 【0115】

30

シーケンシング : Cas9 実験についてはコンストラクトライブラリー (図 8A、第 3 段、左) とレポーター遺伝子の cDNA 配列 (図 8A、第 3 段、右) を Illumina MiSeq 上で 150 bp の重複性ペアエンドリードとして入手し、TALE 実験については対応する配列を Illumina HiSeq 上で 51 bp の非重複性ペアエンドリードとして入手した。

## 【0116】

コンストラクトライブラリー配列の処理 : アラインメント : Cas9 実験について、novoa1ign バージョン 2.07.17 (ワールド・ワイド・ウェップサイト novocraft.com/main/index.php) を使用して 8 bp のライブラリーバーコード対によって挟まれる 234 bp のコンストラクトに対応する一組の 250 bp の基準配列に対してリード対を整列させた (図 8A、第 3 段、左を参照のこと)。 novoa1ign に提供された基準配列では 23 bp の縮合性 Cas9 結合部位領域と 24 bp の縮合性転写物タグ領域 (図 8A、第 1 段を参照のこと) を N と指定し、一方でコンストラクトライブラリーバーコードを明確に規定した。TALE 実験については基準配列の長さが 203 bp であり、縮合性結合部位領域の長さが 23 bp に対して 18 bp であったことを除いて同じ方法を用いた。有効性検査 : 生成されたファイルについての、各リード対の左のリードと右のリードをそれぞれ基準配列に対して整列させた Novoa1ign 出力データ。両方とも基準配列に対してユニークに整列しているリード対のみがその他の有効性条件の対象とされ、これらの条件の全てを通過したリード対のみが保持された。有効性条件には以下のものが含まれる。

40

(i) 2 つのコンストラクトライブラリーバーコードの各々が少なくとも 4 か所の位置で

50

基準配列バーコードと整列しなくてはならず、それら2つのバーコードは同じコンストラクトライブラリーのバーコード対について整列しなくてはならない。

(i i) 基準配列のN領域に整列する全ての塩基はn o v o a l i g nによってA、C、GまたはTと呼ばれなくてはならない。C a s 9 実験についてもT A L E 実験についても左のリードと右のリードは基準N領域で重ならず、これらのN塩基のあいまいなn o v o a l i g nコールの可能性が生じなかつたことに留意されたい。(i i i) 同様に、これらの領域にはn o v o a l i g nによってコールされる挿入または欠失が現れてはならない。(i v) (これらの無作為配列はA、C、およびGのみから生じるので) 転写物タグ領域にTが現れてはならない。これらの条件のうちのいずれか1つに違反するリード対を拒絶リード対ファイルに収集した。自作のp e r lスクリプトを使用してこれらの有効性検査を実施した。 10

#### 【0117】

ガイドされた試料レポーター遺伝子c D N A配列の処理：アラインメント：まず（ワールド・ワイド・ウェップサイトg i t h u b . c o m / j s t j o h n / S e q P r e pからダウンロードされた）S e q P r e pを使用して重複するリード対を7 9 b pの共通セグメントにまで合併し、その後でn o v o a l i g n（上記のバージョン）を使用してこれらの7 9 b p共通セグメントを非対合性単一リードとして一組の基準配列に整列させて（図8 A、第3段、右を参照のこと）、その中で（コンストラクトライブラリーの配列解析については）2 4 b pの縮合性転写物タグをNと指定し、一方で試料バーコードを明確に規定した。T A L E 配列領域とC a s 9 c D N A配列領域の両方が8 b p試料バーコード配列対によって挟まれる同一の6 3 b pのc D N A領域に対応した。有効性検査では（a）ここで、それまでのS e q P r e pによるリード対の合併のため、有効性処理はリード対中の両方のリードのユニークなアラインメントについてフィルターをかける必要が無く、合併したリードのユニークなアラインメントについてのみフィルターをかける必要があること、（b）c D N A配列リードには転写物タグだけが現れるので、有効性処理は基準配列のこれらのタグ領域を適用するだけであり、および別個の結合部位領域に適用されることもないこと、を除き、コンストラクトライブラリーの配列解析に適用したのと同じ条件を適用した（上記参照）。 20

#### 【0118】

結合部位と転写物タグの間の関連表の構築：自作のp e r lを使用して有効性検査されたコンストラクトライブラリー配列からこれらの表を作成した（図8 A、第4段、左）。A、C、およびGの塩基から構成される2 4 b pのタグ配列は基本的にコンストラクトライブラリーにわたってユニークであるべきではあるが（共有する確率=約2 . 8 e - 1 1）、結合部位とタグの関連の早期分析により、おそらくは主に結合配列中の配列エラーの組合せ、またはコンストラクトライブラリーを作製するために使用されたオリゴの中のオリゴ合成エラーの組合せが原因で実際には無視できない割合のタグ配列が複数の結合配列によって共有されることが明らかになった。タグの共有に加えて、有効性検査されたリード対中の結合部位と関連することが分かったタグは、バーコードミスマッチに起因して、どのコンストラクトライブラリーにそれらのタグが由来し得るのか明確ではない場合にコンストラクトライブラリーリード対拒絶ファイルに見出される可能性もあった。最後に、タグ配列自体が配列エラーを含む可能性があった。これらのエラー原因に対処するため、タグを3つの属性に分類した：（i）安全対不安全：不安全はコンストラクトライブラリー拒絶リード対ファイルに見出され得るタグを意味した；共有対非共有：共有は複数の結合部位配列との関連が分かったタグを意味する；および2 + 対1のみ：2 + は有効性検査されたコンストラクトライブラリー配列の中で少なくとも2回現れるタグであって、配列エラーを含む可能性が少ないと考えられるタグを意味する。これらの3つの基準を組み合わせて各結合部位と関連するタグの8つのクラスが生じ、最も保証された（しかし、最も少ない）クラスは安定、非共有、2 + タグのみを含み、最も保証されていない（しかし、最も多い）クラスは安全性、共有制、または発生数に関係なく全てのタグを含む。 30

#### 【0119】

10

20

30

40

50

正規化発現レベルの計算：自作の `perl` コードを使用して図 8 A の第 5 ~ 6 段に示されているステップを実施した。まず、コンストラクトライブラリーについて前もって計算した結合部位と転写物タグの間の表を用いて各ガイド試料について得られたタグ数を結合部位ごとに集計した（図 8 C を参照のこと）。次に各試料について集計された各結合部位のタグ数を陽性対照試料について集計されたタグ数で除算して正規化発現レベルを作成した。これらの計算に関係するその他の考慮事項には以下のものが含まれる。

1. 各試料について、結合部位と転写物タグの間の関連表の中に見出すことができない有効性検査された cDNA 遺伝子配列の中に「新規」タグの組が見出された。その後の計算ではこれらのタグを無視した。

2. 結合部位と転写物タグの間の関連表の中の上記のタグの 8 クラスの各々について上記のタグ数の集計を実施した。コンストラクトライブラリー中の結合部位は、中心となる配列に類似の配列は頻繁に生成されるが、ミスマッチ数が増えた配列はだんだん稀にしか生成されないようにバイアスがかけられるので、ミスマッチがほとんど無い結合部位が一般に大多数のタグになり、ミスマッチが多い結合部位ほど少数のタグにしかならなかった。したがって、最も保証されたタグのクラスを使用することが一般的には望ましいが、2 つ以上のミスマッチを有する結合部位の評価は結合部位当たり少数のタグに基づいたものとなり、それらのタグ自体は信頼できるものであってもその保証された数および比率が統計学的にあまり信頼できないものになる可能性があった。そのような場合、全てのタグを使用した。この考慮事項についてのある補償は、n か所のミスマッチ位置について別個に集計されたタグ数の数はミスマッチ位置の組合せ数（

10

20

【0120】

【数 1】

$$\binom{L}{n}^3$$

に等しい）と共に増加し、n と共に劇的に増加するという事実から通用しており、したがって、様々な数 n のミスマッチについて集計されたタグ数の平均（図 2 b、2 e、および図 9 A および 10 B に示される）は 2 以上の n について統計学的に非常に大きな集計されたタグ数の集合に基づいている。

30

3. 最後に、T A L E コンストラクトライブラリーに構築された結合部位は 18 b p であり、タグとの連関はこれらの 18 b p 配列に基づいて指定されるが、幾つかの実験は 18 b p のコンストラクト結合部位領域内の中央の 14 b p または 10 b p の領域に結合するように計画されている T A L E を用いて実施された。これらの T A L E の発現レベルの計算において、タグを関連表中の 18 b p の結合部位の対応する領域に基づいて結合部位について集計し、それによりこの領域の外側にある結合部位のミスマッチを無視した。

【0121】

【実施例 5】

C a s 9 N V P 6 4 を用いた R N A 誘導性 S O X 2 および N A N O G 制御

本明細書に記載される s g R N A (アプタマー-変換短鎖ガイド R N A) 連結アプローチにより、各 s g R N A が異なる R N A - タンパク質間相互作用対を利用する限り、別個の s g R N A により様々なエフェクタードメインを動員することができ、同一の C a s 9 N タンパク質を用いて多重遺伝子制御を行うことができる。図 12 A の S O X 2 遺伝子と図 12 B の N A N O G 遺伝子について、転写開始部位の上流の約 1 k b 長の D N A 鎖を標的とする 10 種類の g R N A をデザインした。D N a s e 高感受性部位が緑色で強調されている。内在性遺伝子の q P C R により転写活性化をアッセイした。両方の例において、個々の g R N A の導入によって転写が中程度に促進された一方で、複数の g R N A が相乗的に作用して堅固な複数倍の転写活性化を促進した。データは平均値  $\pm$  S E M (N = 3) である。図 12 A ~ B に示されるように、2 種類の追加の遺伝子である S O X 2 と N A N O G はプロモーターの上流約 1 k b 長の D N A 鎖内を標的とする s g R N A によって制御さ

40

50

れた。転写開始部位近傍のそれらの s g R N A によって堅固な遺伝子活性化が引き起こされた。

#### 【 0 1 2 2 】

##### [ 実施例 6 ]

###### C a s 9 g R N A 複合体によるターゲティングのランドスケイプの評価

図 2 に記載されているアプローチを用いて 2 種類の追加的 C a s 9 g R N A 複合体 ( 図 1 3 A ~ C ) と ( 図 1 3 D ~ F ) のターゲティングランドスケイプを分析した。それらの 2 種類の g R N A は非常に異なる特異性プロファイルを有し、 g R N A 2 は最大で 2 ~ 3 個のミスマッチを許容し、 g R N A 3 は最大でも 1 個のミスマッチしか許容しない。これらの態様は 1 塩基ミスマッチプロット ( 図 1 3 B 、 1 3 E ) と 2 塩基ミスマッチプロット ( 図 1 3 C 、 1 3 F ) の両方に反映されている。図 1 3 C および 1 3 F では正規化発現レベルを計算するには入手できるデータが不充分である塩基ミスマッチ対は「 x 」を含む灰色のボックスで表されており、一方でデータ表示を改善するため、正規化発現レベルが彩色スケールの一番上を超す外れ値であるミスマッチ対は星印「 \* 」を含む黄色のボックスで表されている。統計学的有意性の記号は、 \* \* \* が  $P < 0.0005/n$  であり、 \* \* が  $P < 0.005/n$  であり、 \* が  $P < 0.05/n$  であり、 N . S . ( 有意性無し ) は  $P > = 0.05/n$  であり、 n は比較数である ( 表 2 を参照のこと ) 。

10

#### 【 0 1 2 3 】

##### [ 実施例 7 ]

###### 有効性確認、レポーターアッセイの特異性

20

図 1 4 A ~ C に示されるように 2 種類の異なる s g R N A : C a s 9 複合体を使用して特異性データを作成した。対応する突然変異型 s g R N A がレポーターライブラリーを刺激することができなかつたので、そのアッセイは評価される s g R N A について特異的であることを確認した。図 1 4 A : 野生型 g R N A 標的配列に対してデザインされたレポーターライブラリーを使用して 2 種類の g R N A ( 野生型および突然変異体 ; 配列の差異が赤色で強調されている ) の特異性プロファイルを評価した。図 1 4 B : 対応する突然変異型 g R N A がレポーターライブラリーを刺激することができなかつたので、このアッセイは評価される g R N A について特異的であることを確認した ( データを図 1 3 D から再プロット ) 。統計学的有意性の記号は、 \* \* \* が  $P < 0.0005/n$  であり、 \* \* が  $P < 0.005/n$  であり、 \* が  $P < 0.05/n$  であり、 N . S . ( 有意性無し ) は  $P > = 0.05/n$  であり、 n は比較数である ( 表 2 を参照のこと ) 。異なる s g R N A は異なる特異性プロファイルを有することができ ( 図 1 3 A 、 1 3 D ) 、具体的には、 s g R N A 2 は最大で 3 個のミスマッチを許容し、 s g R N A 3 は最大でも 1 個のミスマッチしか許容しない。ミスマッチに対する最も高い感受性は、他の位置にあるミスマッチも活性に影響することが観察されてはいたが、スペーサーの 3 ' 末端に局在する。

30

#### 【 0 1 2 4 】

##### [ 実施例 8 ]

###### 有効性確認、 1 塩基および 2 塩基 g R N A ミスマッチ

図 1 5 A ~ D に示されるように、アッセイした s g R N A 中のスペーサーの 3 ' 末端の 1 2 b p 内に 1 塩基のミスマッチがあつても検出可能なターゲティングが生じることが標的化実験により確認された。しかしながら、この領域内の 2 b p のミスマッチは活性の顕著な喪失を引き起こした。スクレアーゼアッセイを用いて 2 種類の独立した g R N A であるスペーサー配列内に標的にする 1 塩基または 2 塩基のミスマッチ ( 赤色で強調 ) を担持する g R N A 2 ( 図 1 5 A ~ B ) と g R N A 3 ( 図 1 5 C ~ D ) を試験した。アッセイした g R N A 中のスペーサーの 3 ' 末端の 1 2 b p 内に 1 塩基のミスマッチがあつても検出可能なターゲティングが生じるもの、この領域内の 2 b p のミスマッチは活性の急速な喪失を引き起こすことが確認された。これらの結果は異なる g R N A 間の特異性プロファイルの差をさらに強調し、図 1 3 の結果と一致する。データは平均値  $\pm$  S E M ( N = 3 ) である。

40

#### 【 0 1 2 5 】

50

## [実施例9]

有効性確認、5'gRNA短縮

図16A～Dに示されるように、スペーサーの5'部分を短縮してもsgRNA活性が維持された。スクレアーゼアッセイを用いて2種類の独立したgRNAであるスペーサーの5'末端の短縮を有するgRNA1(図16A～B)とgRNA3(図16C～D)を試験した。1～3bpの5'末端短縮はよく許容されるが、それよりも大きい欠失は活性の喪失を引き起こすことが観察された。データは平均値±SEM(N=3)である。

## 【0126】

## [実施例10]

有効性確認、化膿レンサ球菌(S. pyogenes)のPAM

10

図17A～Bに示されるように、化膿レンサ球菌(S. pyogenes)Cas9のPAMはNGGであり、およびまた、NAGであることがスクレアーゼ介在性HRアッセイを用いて確認された。データは平均値±SEM(N=3)である。追加の調査により、ヒトエクソン内の作製された一組の約190KのCas9標的であって、ターゲティング配列の最後の13ntを共有する代替性NGG標的を有しないCas9標的を、前の13nt内にミスマッチを有する代替性NAG部位の存在について、またはNGG部位について走査した。わずかに0.4%のみがそのような代替性標的を有しないことがわかった。

## 【0127】

## [実施例11]

20

有効性確認、TAL E突然変異

18-merのTAL Eは標的配列中の複数の突然変異を許容することがスクレアーゼ介在性HRアッセイ(図18A～B)を用いて確認された。図18A～Bに示されるように、スクレアーゼアッセイにおける標的化実験により測定すると、標的の中央におけるある特定の突然変異によってより高いTAL Eの活性が生じる。

## 【0128】

## [実施例12]

TAL Eモノマーの特異性とTAL Eタンパク質の特異性

個々の反復性可変二残基(RVD)の役割を切り離すため、RVDの選択は塩基特異性に寄与しないが、TAL E特異性は全体としてそのタンパク質の結合エネルギーの関数でもあることを確認した。図19A～CはTAL Eモノマー特異性とTAL Eタンパク質特異性の間の比較を示している。図19A：図2に記載されるアプローチを改変したものを行い、一連の6NIリピートまたは6NHリピートを有する2種類の14-merのTAL E-TFのターゲティングランドスケイプを分析した。このアプローチでは、中央に縮合性6-mer配列を有するレポーターの縮小型ライブラリーを作成および使用して、TAL E-TFの特異性をアッセイした。図19B～C：両方の例において、予期される標的配列(すなわち、NIリピートについては6個のAを有し、NHリピートについては6個のGを有するもの)が濃縮されていることが指摘された。これらのTAL Eの各々はなお中央の6-mer標的配列内に1～2個のミスマッチがあることを許容する。モノマーの選択は塩基特異性に寄与しないが、TAL E特異性は全体としてそのタンパク質の結合エネルギーの関数である。一つの態様によると、操作して短くしたTAL Eまたは高親和性モノマーと低親和性モノマーの組成を有するTAL Eはゲノム工学用途においてより高い特異性を生じ、短いTAL Eを使用するとスクレアーゼ用途においてFok Iの二量体化が標的外効果のさらなる減少を可能にする。

30

## 【0129】

## [実施例13]

40

オフセット・ニック形成、天然遺伝子座

図20A～Bはオフセット・ニック形成に関するデータを示している。ゲノム編集の背景ではオフセット・ニックを形成してDSBを生じさせる。大多数のニックは非相同末端結合(NHEJ)介在性インデルを生じることがなく、したがってオフセット・ニックが

50

ガイドされるときに標的外単一ニック形成事象がインデルを生じる比率は非常に低い可能性がある。オフセット・ニックを誘導して D S B を生成することは、挿入されたレポーター遺伝子座と天然の A A V S 1 ゲノム座位の両方で遺伝子破壊をガイドするのに有効である。図 20 A : 2 0 0 b p 長の D N A 鎖を範囲とする 8 種類の g R N A であるターゲティングセンスストランドを標的とする 4 種類 ( s 1 ~ 4 ) とアンチセンスストランドを標的とする 4 種類 ( a s 1 ~ 4 ) で A A V S 1 遺伝子座のターゲティングを行った。相補ストランドにニックを入れる C a s 9 D 1 0 A 突然変異体を使用して異なる二通りの g R N A の組合せを用いて一連の計画的 5 ' オーバーハングまたは 3 ' オーバーハングを誘導した。図 20 B : 単一の g R N A では検出可能な N H E J 事象がガイドされなかつたが、オフセット・ニックを誘導して D S B を生成することが遺伝子破壊の誘導に非常に有効であることがサンガーシーケンシング・ベースアッセイを用いて観察された。特に、5 ' オーバーハングを生じさせるオフセット・ニックは 3 ' オーバーハングと比べてより多くの N H E J 事象を生じさせる。サンガーシーケンシングクローンの数はバーの上で強調されており、予想されるオーバーハングの長さは対応する x 軸の凡例の下に示されている。  
10

【 0 1 3 0 】

【 実施例 1 4 】

オフセット・ニック形成、N H E J プロファイル

図 21 A ~ C はオフセット・ニック形成および N H E J プロファイルに関する。3 種類の異なるオフセット・ニック形成組合せの代表的なサンガーシーケンシングの結果がボックスにより強調されているターゲティング g R N A の位置と共に示されている。さらに、相同組換え ( H R ) 介在性修復の標準モデルと一致して、オフセット・ニックによる 5 ' オーバーハングの作製によって 3 ' オーバーハングよりも堅固な N H E J 事象がもたらされた ( 図 3 B )。5 ' オーバーハングを形成すると N H E J の促進に加えて H R の堅固な誘導が観察された。3 ' オーバーハングの形成は H R 率の改善を引き起こさなかった ( 図 3 C )。  
20

【 0 1 3 1 】

【 実施例 1 5 】

表 1

内在性遺伝子制御のための g R N A 標的

C a s 9 g R N A 介在性活性化実験において使用された R E X 1 、 O C T 4 、 S O X 2 および N A N O G のプロモーター中の標的を列記する。  
30

【 0 1 3 2 】

【表1】

gRNA Name	gRNA Target
REX1 1	ctggggggatcactcggggtt agg
REX1 2	ccctcgcctccaaaaaaytgtt agg
REX1 3	acgtgtattctgcagatcaa ggg
REX1 4	ccaggaaatacgtatccacca ggg
REX1 5	gcacacccaaagcgcataaa tgg
REX1 6	aataatacattctaaggta agg
REX1 7	gtctactggggagggtgtggc agg
REX1 8	tagcaatacaytacacattaa tgg
REX1 9	ctcatgtgtatccccccgtct agg
REX1 10	ccgggcagagactgaaacgcg agg
OCT4 1	ttccttccctctccctgtgt tgg
OCT4 2	tcctctgcctaaageccctggag agg
OCT4 3	aatggaaatgtccggagggtgcag tgg
OCT4 4	ccctcagccctctaaaytgtt agg
OCT4 5	gagtccaaatcctctttact agg
OCT4 6	gagtgtctggatttggata agg
OCT4 7	cagcacccatctccctgg tgg
OCT4 8	tctaaaaccccaaggaaatcat ggg
OCT4 9	cacaaggcagccaggatcc agg
OCT4 10	gatggcaagctgagaaaaacac tgg
OCT4 11	tgaatgtcacgcatacaatt agg
OCT4 12	ccagtcctcagacacttggcttc tgg
OCT4 13	cccgaaaaaacagacccctga agg
OCT4 14	aaagggttggcacttggtttta ggg
OCT4 15	atgtctgggtttttgggttgg agg
OCT4 16	gttcccttgaaggggaaagta ggg
OCT4 17	tggcagtttactttttaaga tgg
OCT4 18	ggcacagtgtccagggttcgg tgg
OCT4 19	taaaaataaaaaaaactaacs ggg
OCT4 20	tctgtgggggacctgtcactg agg
OCT4 21	ggccagagggtcaaggctgtt ggg
SOX2 1	cacgacccgaaacccttctta agg
SOX2 2	gttggaaatgtaaaggacagtctat tgg
SOX2 3	taagaacccggcggcaatgttacg tgg
SOX2 4	tgttaaggtaagagagggatgg agg
SOX2 5	tgtcacacccaaacttctgtac tgg
SOX2 6	tttacccacttcccttggaaa agg
SOX2 7	gtgggttggcagggtggctttt ggg
SOX2 8	ctcccccggccctcccccggcgg agg
SOX2 9	caaaacccggcaggcgagggtt ggg
SOX2 10	aggaggccggccggcgtgtt tgg
NANOG 1	cacacacacccacacccggat ggg
NANOG 2	gaagaaggttaaaggagccaga ggg
NANOG 3	atggaaatttcaataaccttc agg
NANOG 4	tcccgcttgcgttgcggcaggc tgg
NANOG 5	cagacacccacccacccatgtgg tgg
NANOG 6	tcccaatttactgggattac agg
NANOG 7	tgatttaaaaagggtggaaaagg tgg
NANOG 8	tctagttccccacccatgtttt ggg
NANOG 9	gatataactgtggaaattccacaa ggg
NANOG 10	ccggccggccggcgtgggttca agg

【 0 1 3 3 】

### [ 实施例 1 6 ]

表 2

## Cas9 gRNA および TALE の特異性データの統計学的分析の要約

表2 (a) 特定の数の標的部位突然変異を有する標的配列に結合するT A L E またはC a s 9 - V P 6 4 活性化因子の正規化発現レベルの比較のP値。正規化発現レベルが図力

ラムに示されている図の中の箱ひげ図によって表されており、箱は標的部位のミスマッチ数毎のこれらの発現レベルの分布状態を表す。各箱ひげ図の中のミスマッチ数の各連続対について  $t$  検定を用いて  $P$  値を計算したが、その  $t$  検定は一標本  $t$  検定または二標本  $t$  検定のどちらかであった（方法を参照のこと）。各箱ひげ図内の比較数に基づくボンフェローニ補正  $P$  値閾値を使用して統計学的有意性を評価した。統計学的有意性の記号は、  
\* \* が  $P < 0.0005/n$  であり、\* \* \* が  $P < 0.005/n$  であり、\* が  $P < 0.05/n$  であり、および N.S.（有意性無し）は  $P >= 0.05/n$  であり、ここで  $n$  は比較数である。表 2 (b) 図 2 D 中のシード領域の統計学的特徴解析：2 つの突然変異が 20 bp の標的部位の 3' 末端の候補シード領域内の突然変異形成位置対にある標的配列と他の全ての位置対にある標的配列に C a s 9 N V P 6 4 + g R N A が結合したときの発現値の間の分離度を表す  $10g10$  ( $P$  値)。最大の  $-10g10$  ( $P$  値)（上で強調されている）によって示されている最大の分離度は標的部位の最後の 8 ~ 9 bp の中に見出される。これらの位置はこの標的部位の「シード」領域の始まりを表すものと解釈してよい。 $P$  値の計算方法の情報については方法のなかの「シード領域の統計学的分析」の節を参考されたい。  
10

【0134】

【表2】

a Figure	Expression level comparison: mutations vs. mutations		t-test	P-value	Symbol
	0	1			
2b	1	2	2-samp	1.4E-06	***
	2	3	2-samp	4.0E-01	***
	3	4	2-samp	0	***
	4	5	2-samp	0	***
	5	6	2-samp	1.0E-217	***
	6	7	2-samp	1.7E-43	***
	7	8	2-samp	3.7E-02	N.S.
	8	1	1-samp	8.9E-01	N.S.
	1	2	2-samp	1.9E-06	***
	2	3	2-samp	3.0E-147	***
2e	3	4	2-samp	0	***
	4	5	2-samp	0	***
	5	6	2-samp	4.2E-63	***
	6	7	2-samp	1.6E-03	*
	7	8	2-samp	4.7E-01	N.S.
	8	1	1-samp	5.2E-02	N.S.
	1	2	2-samp	2.8E-05	***
	2	3	2-samp	3.5E-21	***
	3	4	2-samp	1.4E-52	***
	4	5	2-samp	8.5E-101	***
S7e	5	6	2-samp	6.8E-94	***
	6	7	2-samp	1.8E-01	***
	7	8	2-samp	5.1E-24	***
	8	1	1-samp	2.3E-18	***
	1	2	2-samp	2.4E-08	***
	2	3	2-samp	6.2E-54	***
	3	4	2-samp	4.8E-141	***
	4	5	2-samp	1.9E-20	***
	5	6	2-samp	1.2E-03	*
	6	7	2-samp	3.8E-02	***
S7d and S8d	7	8	2-samp	9.4E-01	N.S.
	8	1	1-samp	7.2E-03	N.S.
	1	2	2-samp	5.0E-01	N.S.
	2	3	2-samp	3.3E-34	***
	3	4	2-samp	6.5E-153	***
	4	5	2-samp	8.6E-76	***
	5	6	2-samp	1.6E-03	*
	6	7	2-samp	7.1E-01	N.S.
	7	8	2-samp	7.8E-02	N.S.
	8	1	1-samp	7.3E-01	N.S.
S8c	1	2	2-samp	2.4E-06	***
	2	3	2-samp	7.2E-140	***
	3	4	2-samp	0	***
	4	5	2-samp	0	***
	5	6	2-samp	1.0E-72	***
	6	7	2-samp	4.0E-03	*
	7	8	2-samp	7.3E-01	N.S.
	8	1	1-samp	7.3E-01	N.S.
	1	2	2-samp	2.4E-06	***
	2	3	2-samp	7.2E-140	***
S13a (left)	3	4	2-samp	0	***
	4	5	2-samp	0	***
	5	6	2-samp	1.0E-72	***
	6	7	2-samp	4.0E-03	*
	7	8	2-samp	7.3E-01	N.S.
	8	1	1-samp	7.3E-01	N.S.
S13a (middle)	1	2	2-samp	5.8E-03	***
	2	3	2-samp	7.9E-86	***
	3	4	2-samp	2.9E-53	***
	4	5	2-samp	8.5E-10	***
	5	6	2-samp	1.3E-13	***
	6	7	2-samp	3.7E-08	***
S13a (right)	7	8	2-samp	5.8E-03	***
	8	1	1-samp	1.3E-13	***
	1	2	2-samp	1.1E-04	***

seed position	Number position pairs		-log10 p-value
	both in seed	not both in seed	
2	373	19	3.11
3	153	37	1.36
4	126	54	2.01
5	129	79	2.34
6	105	85	5.83
7	91	99	7.54
8	78	112	6.81
9	66	124	7.10
10	55	135	9.72
11	45	145	9.83
12	36	134	10.44
13	28	162	10.72
14	21	169	8.97
15	15	175	9.61
16	10	180	3.34
17	6	184	2.26
18	3	187	1.16

【0135】

[実施例17]

実施例中のタンパク質とRNAの配列

A. m4突然変異体に基づくCas9<sub>N</sub> VP64活性化因子コンストラクトの配列が下に表示されている。最大の活性を示すCas9<sub>m4</sub> VP64融合タンパク質形式とC

10

20

30

40

50

a s 9 <sub>m 4</sub> V P 6 4 N 融合タンパク質形式を有する 3 種類のバージョンを構築した。m 3 突然変異体およびm 2 突然変異体（図 4 A）の対応するベクターも構築した（N L S ドメインと V P 6 4 ドメインが強調されている）。

【 0 1 3 6 】

>Cas9<sub>m4</sub><sup>VP64</sup>

gccaccATGGACAAGAAGTACTCCATTGGCTCGCTATCGGCACAAACAGCGTCGGCTGG  
 GCCGTCATTACGGACGAGTACAAGGTGCCGAGCAAAAAATTCAAAGTTCTGGCAAT  
 ACCGATGCCACAGCATAAAGAAGAACCTCATTGGCGCCCTCCTGTTGACTCCGGGG  
 AGACGGCCGAAGGCCACGCCGCTCAAAGAACAGCACGGCGCAGATATAACCCGAGAA  
 AGAATCGGATCTGCTACCTGCAGGAGATCTTAGTAATGAGATGGCTAAGGTGGATGA  
 CTCTTCTTCCATAGGCTGGAGGAGTCCTTTGGAGGAGATAAAAAGCACGAG  
 CGCCACCCAATCTTGGCAATATCGTGGACGAGGTGGCGTACCATGAAAAGTACCCAA  
 CCATATATCATCTGAGGAAGAACGTTGAGACAGTACTGATAAGGCTGACTGCGGTT  
 GATCTATCTCGCGCTGGCGCATATGATCAAATTCTGGGGACACTTCCTCATCGAGGG  
 GACCTGAACCCAGACAACAGCGATGTCGACAAACTCTTATCCAACACTGGTCAGACTT  
 ACAATCAGCTTTCGAAGAGAACCCGATCAACGCATCCGGAGTTGACGCCAAGCAA  
 TCCTGAGCGCTAGGCTGTCCAATCCGGCGGCTCGAAAACCTCATCGCACAGCTCCC  
 TGGGGAGAAGAAGAACGGCCTGTTGTAATCTTATGCCCTGTCACTCGGGCTGACC  
 CCCAACTTAAATCTAACTCGACCTGGCGAAGATGCCAAGCTCAACTGAGCAAAG  
 ACACCTACGATGATGATCTGACAATCTGCTGGCCAGATGGCGACCAGTACGCAGA  
 CCTTTTTGGCGGCAAAGAACCTGTCAGACGCCATTGCTGAGTGATATTCTGCGAG  
 TGAACACGGAGATCACAAAGCTCCGCTGAGCGCTAGTATGATCAAGCGCTATGATG  
 AGCACCACCAAGACTTGACTTTGCTGAAGGCCCTGTCAGACAGCAACTGCCTGAGAA  
 GTACAAGGAAATTCTCGATCAGTCTAAAGGCTACGCCATCTGGAAAAATGGAC  
 GGCACCGAGGAGCTGCTGGATAAGGCTTAACAGAGAACGACTGTTGCGCAAACAGCGC  
 ACTTTCGACAATGGAAGCATCCCCCACCAGATTCACCTGGCGAACTGCACGCTATCC  
 TCAGGCGGCAAGAGGATTCTACCCCTTTGAAAGATAACAGGGAAAAGATTGAGA  
 AAATCCTCACATTCTGGATACCCACTATGTTAGGCCCCCTGCCGGGGAAATTCCAG  
 ATTGCGTGGATGACTCGCAAATCAGAAGAGACCATCACTCCCTGGAACCTCGAGGAA  
 GTCGTGGATAAGGGGGCTCTGCCAGTCCTCATGAAAGGATGACTAACTTGATA  
 AAAATCTGCCTAACGAAAAGGTGCTCCTAAACACTCTGCTGTACGAGTACTTCAC  
 AGTTTATAACGAGCTCACCAAGGTCAAATACGTCACAGAACGGATGAGAAAGCCAGC  
 ATTCCCTGTCTGGAGAGCAGAACAGCTATCGTGGACCTCCTTCAAGACGAACCGG  
 AAAGTTACCGTGAAACAGCTCAAAGAACACTATTCAAAAAGATTGAATGTTGCACT  
 CTGTTGAAATCAGCGGAGTGGAGGATCGCTCAACGCATCCCTGGAACGTATCACGA  
 TCTCCTGAAAATCATTAAAGACAAGGACTTCCTGGACAATGAGGAGAACGAGGACAT

10

20

30

40

TCTTGAGGACATTGTCCTCACCTTACGTTGTTGAAGATAGGGAGATGATTGAAGAA  
CGCTGAAAACCTACGCTCATCTCTCGACGACAAAGTCATGAAACAGCTAAGAGGC  
GCCGATATACAGGATGGGGCGGCTGTCAAGAAAAGTCAATGGGATCCGAGACA  
AGCAGAGTGGAAAGACAATCCTGGATTTCTTAAGTCCGATGGATTGCCAACCGGAA  
CTTCATGCAGTTGATCCATGACTCTCACCTTAAGGAGGACATCCAGAAAGCA  
CAAGTTCTGCCAGGGGACAGTCTCACGAGCACATCGCTAATCTGCAGGTAGCC  
CAGCTATCAAAAAGGAATACTGCAGACCGTTAAGGTCGTGGATGAACCTCGTCAAAG  
TAATGGGAAGGCATAAGCCCGAGAATATCGTTATCGAGATGGCCGAGAGAACCAAA  
CTACCCAGAAGGGACAGAAGAACAGTAGGGAAAGGATGAAGAGGATTGAAGAGGGT  
ATAAAAGAACTGGGGTCCCAAATCCTTAAGGAACACCCAGTTGAAAACACCCAGCT  
CAGAATGAGAAGCTCTACCTGTACTACCTGCAGAACGGCAGGGACATGTACGTGGAT  
CAGGAACCTGGACATCAATCGGCTCTCCGACTACGACGTGGCTGCTATCGTCCCCAGT  
CTTTCTCAAAGATGATTCTATTGATAATAAGTGTGACAAGATCCGATAAAAgcTAGA  
GGGAAGAGTGATAACGTCCCCCTCAGAAGAACAGTTGCAAGAAAATGAAAATTATTGG  
CGGCAGCTGCTGAACGCCAAACTGATCACACAACGGAAGTTGATAATCTGACTAAG  
GCTGAACGAGGTGGCCTGCTGAGTTGATAAAGCCGGCTCATCAAAAGGCAGCTG  
TTGAGACACGCCAGATCACCAAGCACGTGGCCCAAATTCTCGATTACCGCATGAACAC  
CAAGTACGATGAAAATGACAAACTGATTGAGAGGTGAAAGTTATTACTCTGAAGTCT  
AAGCTGGTCTCAGATTCAAGAAAGGACTTCAGTTATAAGGTGAGAGAGATCAACA  
ATTACCACCATGCGCATGATGCCTACCTGAATGCAGTGGTAGGCAGTGCACCTATCAA  
AAAATATCCAAGCTGAATCTGAATTGTTACGGAGACTATAAGTGTACGATGTT  
AGGAAAATGATCGCAAAGTCTGAGCAGGAAATAGGCAAGGCCACCGCTAAGTACTTC  
TTTACAGCAATTATGAATTTCAGACCCAGATTACACTGGCAATGGAGAGA  
TTCGGAAGCGACCACTATCGAAACAAACGGAGAAACAGGGAGAAATCGTGTGGACA  
AGGGTAGGGATTCGCGACAGTCCCGAAGGTCTGTCCATGCCAGGTGAACATCGT  
AAAAAGACCGAAGTACAGACCGGAGGCTCTCCAAGGAAAGTATCCTCCGAAAAG  
GAACAGCGACAAGCTGATCGCACGCCAAAAGATTGGGACCCCAAGAAATACGGCG  
ATTGATTCTCCTACAGTCGCTTACAGTGTACTGGTTGTCAGGAAAGTGGAGAAAGGG  
AAGTCTAAAAAAACTCAAAAGCGTCAAGGAACGTCTGGCATCACAATCATGGAGCGA  
TCAAGCTCGAAAAAAACCCATCGACTTCTCGAGGGCAGGAAAGGATATAAGAGGTC  
AAAAAAGACCTCATCATTAAGCTTCCAAGTACTCTCTTGTGAGCTGAAACGGCC  
GGAAACGAATGCTCGTAGTGCAGGAGCTGCAGAAAGGTAACGAGCTGGCACTGC  
CCTCTAAATACGTTAATTCTGTATCTGGCAGCCACTATGAAAAGCTCAAAGGGTCT  
CCCGAAGATAATGAGCAGAAGCAGCTGTTGGAACAACACAAACACTACCTTGAT  
GAGATCATCGAGCAAATAAGCGAATTCTCCAAAAGAGTGTGATCCTCGCCACGCTAAC  
CTCGATAAGGTGCTTCTGCTTACAATAAGCACAGGGATAAGCCCATCAGGGAGCAGG

10

20

30

40

CAGAAAACATTATCCACTTGTACTCTGACCAACTGGCGCGCCTGCAGCCTCAA  
 GTACTTCGACACCACCATAGACAGAAAGCGGTACACCTCTACAAAGGAGGTCTGGA  
 CGCCACACTGATTCAATTACGGGCTCTATGAAACAAGAATCGACCTCT  
 CAGCTCGGTGGAGACAGCAGGGCTGACCCCAAGAAGAAGAGGAAAGTGGAGGCCAG  
 CGGTTCCGGACGGCTGACGCATTGGACGATTTGATCTGGATATGCTGGAAAGTGAC  
 GCCCTCGATGATTTGACCTTGACATGCTTGGTTCGGATGCCCTGATGACTTTGACCT  
 CGACATGCTCGGCAGTGACGCCCTTGATGATTTCGACCTGGACATGCTGATTAACCT

AGATGA

10

【 0 1 3 7 】

>Cas9<sub>m4</sub><sup>VP64</sup>N Sequences

gcaccATGCCAAAGAAGAAGAGGAAGGTGGAAAGGGGATGGACAAGAAGTACTCCA  
 TTGGGCTCGCTATCGGCACAAACAGCGTCGGCTGGCCGTATTACGGACGAGTACAA  
 GGTGCCGAGCAAAAATTCAAAGTTCTGGCAATACCGATGCCACAGCATAAAGAA  
 GAACCTCATTGGGCCCTCCTGTTGACTCCGGGAGACGGCGAAGCCACGCGGCTC  
 AAAAGAACAGCACGGCGCAGATATACCGCAGAAAGAACATGGATCTGCTACCTGCAG  
 GAGATCTTAGTAATGAGATGGCTAAGGTGGATGACTCTTCTTCATAGGCTGGAGG  
 AGTCCTTTGGTGGAGGAGGATAAAAGCACGAGCGCCACCAATCTTGGCAATAT  
 CGTGGACGAGGTGGCGTACCATGAAAAGTACCCAAACCATATATCATCTGAGGAAGAA  
 GCTTGTAGACAGTACTGATAAGGCTGACTTGCCTGTTGATCTATCTCGCGCTGGCGCAT  
 ATGATCAAATTGGGGACACTTCCTCATCGAGGGGACCTGAACCCAGAACACAGC  
 GATGTCGACAAACTCTTATCCAACGGTTCAAGACTACATCAGCTTTCGAAGAGA  
 ACCCGATCAACGCATCCGGAGTTGACGCCAAGCAATCCTGAGCGCTAGGCTGTCAA  
 ATCCCGGGCTCGAAAACCTCATCGCACAGCTCCCTGGGAGAACAGAACGGCCT  
 GTTGGTAATCTTATGCCCTGTCACTGGCTGACCCCAACTTAAATCTAACTTCG  
 ACCTGGCGAAGATGCCAAGCTCAACTGAGCAAAGAACACCTACGATGATGATCTG  
 ACAATCTGCTGGCCAGATCGCGACCACTGAGCAGACCTTTGGCGGAAAGAA  
 CCTGTCAGACGCCATTCTGCTGAGTGTATTCTGCGAGTGAACACGGAGATACCAAA  
 GCTCCGCTGAGCGCTAGTATGATCAAGCGCTATGATGAGCACCACCAAGACTGACTT  
 TGCTGAAGGCCCTGTCAGACAGCAACTGCCTGAGAAGTACAAGGAAATTCTCGA  
 TCAGTCTAAAATGGCTACGCCGATACATTGACGGCGAGCAAGCCAGGAGGAATT  
 TTACAAATTATTAAGCCCATTGGAAAAAATGGACGGCACCGAGGAGCTGCTGGTA  
 AAGCTTAACAGAGAAGATCTGTTGCGAACACAGCGCACTTCGACAATGGAAGCATC  
 CCCCACCAAGATTACCTGGCGAACACTGCACGCTATCCTCAGGGCGAACAGAGGATTCT  
 ACCCCTTTGAAAGATAACAGGGAAAAGATTGAGAAAATCCTCACATTCGGATACC

20

30

40

CTACTATGTAGGCCCCCTGCCCGGGAAATTCCAGATTGCGGTGGATGACTCGCAAA  
TCAGAAGAGACCATCACTCCCTGGAACCTCGAGGAAGTCGTGGATAAGGGGGCCTCT  
GCCAGTCCTCATCGAAAGGATGACTAACTTGATAAAAATCTGCCTAACGAAAAGG  
TGCTCCTAAACACTCTGTGAGTACAGTACAGTTATAACGAGCTCACCAAG  
GTCAAATACGTACAGAAGGGATGAGAAAGCCAGCATTCTGTGGAGAGCAGAAG  
AAAGCTATCGTGGACCTCCTCTCAAGACGAACCGGAAAGTTACCGTAAACAGCTCA  
AAGAAGACTATTCAAAAAGATTGAATGTTCGACTCTGTTGAAATCAGCGGAGTGG  
GGATCGCTTCAACGCATCCCTGGGAACGTATCACGATCTCCTGAAAATCATTAAAGAC 10  
AAGGACTTCCTGGACAATGAGGAGAACGAGGACATTCTTGAGGACATTGCTCACCC  
TTACGTTGTTGAAGATAGGGAGATGATTGAAGAACGCTTGAAAACCTACGCTCATCT  
CTTCGACGACAAAGTCATGAAACAGCTCAAGAGGCCGATATACAGGATGGGGCG  
GCTGTCAAGAAAATGATCAATGGGATCCGAGACAAGCAGAGTGGAAAGACAATCCT  
GGATTTCTTAAGTCCGATGGATTGCCAACCGAACCTCATGCAAGTTGATCCATGATG  
ACTCTCACCTTAAGGAGGACATCCAGAAAGCACAAGTTCTGCCAGGGGACAG  
TCTTCACGAGCACATCGCTAACCTGCAGGTAGCCCAGCTATCAAAAGGGAAACTG 20  
CAGACCGTTAAGGTGAGTGAACCGTCAAAGTAATGGGAAGGCATAAGCCCAG  
AATATCGTTATCGAGATGGCCCGAGAGAACCAAACCTACCCAGAAGGGACAGAAC  
AGTAGGGAAAGGATGAAGAGGATTGAAGAGGGTATAAAAGAACTGGGTCCCAAAT  
CCTTAAGGAACACCCAGTTGAAAACACCCAGCTTCAGAATGAGAAGCTCTACCTGTAC  
TACCTGCAGAACGGCAGGGACATGTACGTGGATCAGGAACCTGGACATCAATCGGCTC  
TCCGACTACGACGTGGCTGCTACGTGCCCCAGTCTTCTCAAAGATGATTCTATTGA  
TAATAAAGTGTGACAAGATCCGATAAA<sup>Ag</sup>TAGAGGGAAAGAGTGATAACGTCCCTCA 30  
GAAGAAGTTGTCAAGAAAATGAAAATTATTGGCGGCCAGCTGCTGAACGCCAAACTG  
ATCACACAACGGAAGTCGATAATCTGACTAAGGCTGAACGAGGTGGCCTGTGAGT  
TGGATAAAGCCGGCTTCATCAAAAGGCAGCTTGTGAGACACGCCAGATACCAAGC  
ACGTGGCCCAAATTCTGATTACGCATGAACACCAAGTACGATGAAAATGACAAACT  
GATTGAGAGGTGAAAGTTATTACTCTGAAGTCTAACGCTGGTCTCAGATTCAGAAAG  
GACTTTCAGTTATAAGGTGAGAGAGATCAACAATTACCAACCATGCGCATGATGCCT  
ACCTGAATGCAGTGGTAGGCAGTGCACCTATCAAAAATATCCAAGCTGAACTG  
ATTGTTACGGAGACTATAAGTGTACGATGTTAGGAAAATGATCGCAAAGTCTGAG  
CAGGAAATAGGCAAGGCCACCGCTAAGTACTTCTTTACAGCAATATTGAATT  
TCAAGACCGAGATTACACTGGCAATGGAGAGATTCGGAAGCGACCACTATCGAAA 40  
CAAACGGAGAAACAGGGAGAAATCGTGTGGACAAGGGTAGGGATTCGCGACAGTCC  
GGAAGGTCTGTCCATGCCGCAGGTGAACATCGTAAAAGACCGAAGTACAGACCG  
GAGGCTCTCCAAGGAAAGTATCCTCCGAAAAGGAACAGCGACAAGCTGATCGCAC  
GCAAAAAAGATTGGGACCCAAGAAATACGGCGGATTGATTCTCCTACAGTCGCTTA

CAGTGTACTGGTTGTGGCAAAGTGGAGAAAGGGAAAGTCTAAAAAACTCAAAAGCGT  
 CAAGGAAC TGCTGGCATCACAAATCATGGAGCGATCAAGCTCGAAAAAAACCCAT  
 CGACTTCTCGAGGCAGAAGGATATAAGAGGTCAAAAAAGACCTCATCATTAAGCTT  
 CCCAAGTACTCTCTTGTAGCTTGAACAGGCTGGCACTGCCCTCTAAATACGTTAATTCTTGTA  
 GCGAGCTGCAGAAAGGTAAACGAGCTGGCACTGCCCTCTAAATACGTTAATTCTTGTA  
 TCTGGCCAGCCACTATGAAAAGCTCAAAGGGTCTCCGAAGATAATGAGCAGAAGCA  
 GCTGTTCGTGGAAACAACACAAACACTACCTTGATGAGATCATCGAGCAAATAAGCGA  
 ATTCTCCAAAAGAGTGATCCTCGCCGACGCTAACCTCGATAAGGTGCTTCTGCTTAC 10  
 AATAAGCACAGGGATAAGCCCACAGGGAGCAGGCAGAAAACATTATCCACTTGT  
 ACTCTGACCAACTTGGCGCGCCTGCAGCCTCAAGTACTTCGACACCACCATAGACA  
 GAAAGCGGTACACCTCTACAAAGGAGGTCTGGACGCCACACTGATTCATCAGTCAT  
 TACGGGGCTCTATGAAACAAGAACATCGACCTCTCAGCTCGTGGAGACAGCAGGGCT  
 GACGATTGATCTGGATATGCTGGAAAGTGACGCCCTCGATGATTTGACCTTGACA  
 TGCTTGGTCGGATGCCCTTGATGACTTTGACCTCGACATGCTCGGAGTGACGCCCTT  
 GATGATTTCGACCTGGACATGCTGATTAACCTAGATGA 20  
**【 0 1 3 8 】**

>Cas9<sub>m4</sub><sup>VP64</sup>C

gecaccATGGACAAGAAGTACTCCATTGGCTCGCTATCGGCACAAACAGCGTCGGCTGG  
 GCCGT CATTACGGACGAGTACAAGGTGCCGAGCAAAAAATTCAAAGTTCTGGCAAT  
 ACCGATCGCCACAGCATAAAAGAACGAAACCTCATTGGCGCCCTCCTGTTGACTCCGGGG  
 AGACGGCCGAAGCCACGCGGCTCAAAAGAACAGCACGGCGCAGATATAACCGCAGAA  
 AGAACATCGGATCTGCTACCTGCAGGAGATCTTAGTAATGAGATGGCTAAGGTGGATGA  
 CTCTTCTTCCATAGGCTGGAGGAGTCCTTTGGTGGAGGAGGATAAAAGCACGAG 30  
 CGCCACCCAACTTTGGCAATATCGGACGAGGTGGCGTACCATGAAAAGTACCCAA  
 CCATATATCATCTGAGGAAGAACGCTGTAGACAGTACTGATAAGGCTGACTGCGGTT  
 GATCTATCTCGCGCTGGCGCATATGATCAAATTCTGGGGACACTCCTCATCGAGGGG  
 GACCTGAACCCAGACAACAGCGATGTCGACAAACTCTTATCCAAC TGCTCAGACT  
 ACAATCAGCTTTCGAAGAGAACCGATCAACGCATCCGGAGTTGACGCCAAAGCAA  
 TCCTGAGCGCTAGGCTGTCAAATCCGGCGGCTCGAAAACCTCATCGCACAGCTCCC  
 TGGGGAGAAGAACGGCTGTTGGTAATCTTATGCCCTGTCACTCGGGCTGACC 40  
 CCCAACTTAAACTAACTCGACCTGGCGAAGATGCCAAGCTCAACTGAGCAAAG  
 ACACCTACGATGATGATCTGACAAATCTGCTGGCCAGATGGCGACCAGTACGCAGA  
 CCTTTTTGGCGGCAAAGAACCTGTCAGACGCCATTCTGCTGAGTGATATTCTGCGAG

TGAACACGGAGATCACCAAAGCTCCGCTGAGCGCTAGTATGATCAAGCGCTATGATG  
AGCACCACCAAGACTTGACTTGTGAAGGCCCTGTCAAGACAGCAACTGCCTGAGAA  
GTACAAGGAAATTTCTTCGATCAGTCTAAAATGGCTACGCCGGATACATTGACGGC  
GGAGCAAGCCAGGAGGAATTACAAATTATAAGCCATCTGGAAAAAATGGAC  
GGCACCGAGGAGCTGGTAAAGCTAACAGAGAAGATCTTGCACAAACAGCGC  
ACTTCGACAATGGAAGCATCCCCCACCAGATTCACCTGGCGAACTGCACGCTATCC  
TCAGGCAGCAAGAGGATTCTACCCCTTGAAAGATAACAGGGAAAAGATTGAGA  
AAATCCTCACATTGGATACCTACTATGTAGGCCCCCTGCCCGGGAAATTCCAG 10  
ATTCGCGTGGATGACTCGCAAATCAGAAGAGACCATCACTCCCTGAACTTCGAGGAA  
GTCGTGGATAAGGGGGCCTCTGCCAGTCCTCATCGAAAGGATGACTAACCTTGATA  
AAAATCTGCCTAACGAAAAGGTGCTCCTAACACTCTGCTGTACGAGTACTTCAC  
AGTTATAACGAGCTACCAAGGTCAAATACGTACAGAAGGGATGAGAAAGCCAGC  
ATTCCTGTCTGGAGAGCAGAAGAAAGCTATCGGACCTCCTCTCAAGACGAACCGG  
AAAGTTACCGTAAACAGCTCAAAGAAGACTATTCAAAAAGATTGAATGTTCGACT  
CTGTTGAAATCAGCGGAGTGGAGGATCGCTAACGCATCCCTGGAACGTATCAGA 20  
TCTCCTGAAAATCATTAAAGACAAGGACTTCCTGGACAATGAGGAGAACGAGGACAT  
TCTTGAGGACATTGTCCTCACCTTACGTTGAAAGATAGGGAGATGATTGAAGAA  
CGCTTGAAAACCTACGCTCATCTCTCGACGACAAAGTCATGAAACAGCTCAAGAGGC  
GCCGATATACAGGATGGGGCGGCTGTCAAGAAAATGATCAATGGATCCGAGACA  
AGCAGAGTGGAAAGACAATCCTGGATTTCTAACGTCCGATGGATTGCCAACCGGAA  
CTTCATGCAGTTGATCCATGATGACTCTCACCTTAAGGAGGACATCCAGAAAGCA  
CAAGTTCTGCCAGGGGACAGTCTCACGAGCACATCGCTAACCTGCAGGTAGCC  
CAGCTATCAAAAGGGAAATACTGCAGACCGTTAACGTCGTGGATGAACTCGTCAAAG  
TAATGGGAAGGCATAAGCCCAGAAATACGTTATCGAGATGGCCGAGAGAACCAAA 30  
CTACCCAGAAGGGACAGAAGAACAGTAGGGAAAGGATGAGAGGATTGAAGAGGGT  
ATAAAAGAACTGGGGTCCAAATCCTAACGAAACACCCAGTTGAAAACACCCAGCTT  
CAGAATGAGAAGCTCTACCTGTACTACCTGCAGAACGGCAGGGACATGTACGTGGAT  
CAGGAACCTGGACATCAATCGGCTCTCCGACTACGACGTGGCTGCTACGTCCCCAGT  
CTTTCTCAAAGATGATTCTATTGATAATAAGTGTGACAAGATCCGATAAAAGTAGA 40  
GGGAAGAGTGATAACGTCCTCAGAAGAAGTTGCAAGAAAATGAAAATTATTGG  
CGGCAGCTGCTGAACGCCAAACTGATCACACACAGGAAGTTGATAATCTGACTAAG  
GCTGAACGAGGTGGCCTGTCTGAGTTGATAAGCCGGCTCATCAAAAGGCAGCTG  
TTGAGACACGCCAGATCACCAAGCACGTGGCCAAATTCTCGATTACGCATGAACAC  
CAAGTACGATGAAAATGACAAACTGATTGAGAGGTGAAAGTTATTACTCTGAAGTCT  
AAGCTGGTCTCAGATTCAGAAAGGACTTCAGTTATAAGGTGAGAGAGATCAACA  
ATTACCACCATGCGCATGATGCCTACCTGAATGCAGTGGTAGGCAGTCACCTATCAA

AAAATATCCAAGCTGAATCTGAATTGTTACGGAGACTATAAGTGTACGATGTT  
 AGGAAAATGATCGAAAGTCTGAGCAGGAATAGGCAAGGCCACCGCTAAGTACTTC  
 TTTACAGCAATATTGAATTTCAGACAGGAGATTACACTGGCAATGGAGAGA  
 TTCGGAAGCGACCACTTATCGAAACAAACGGAGAACAGGAGAAATCGTGTGGAC  
 AGGGTAGGGATTTCGCGACAGTCCGGAAAGTCCTGTCCATGCCAGGTGAACATCGT  
 TAAAAAGACCGAAGTACAGACCGGAGGCTCTCCAAGGAAAGTATCCTCCGAAAAG  
 AACAGCGACAAGCTGATCGCACGCAAAAAAGATTGGACCCAAAGAAATACGGCGG  
 ATTCGATTCTCCTACAGTCGCTTACAGTGTACTGGTTGTGGCAAAGTGGAGAAAGGG 10  
 AAGTCTAAAAAAACTCAAAAGCGTCAAGGAAGTGTCTGGCATCACAATCATGGAGCGA  
 TCAAGCTCGAAAAAAACCCATCGACTTCTCGAGGGAAAGGATATAAGAGGTC  
 AAAAAGACCTCATCATTAAGCTTCCAAGTACTCTCTTGAGCTGAAAACGGCC  
 GGAAACGAATGCTCGCTAGTGCAGGCGAGCTGCAGAAAGGTAACGAGCTGGCACTGC  
 CCTCTAAATACGTTAATTCTGTATCTGCCAGCCACTATGAAAAGCTCAAAGGGTCT  
 CCCGAAGATAATGAGCAGAAGCAGCTGTTGGAACAAACACAAACACTACCTTGAT  
 GAGATCATCGAGCAAATAAGCGAATTCTCAAAGAGTGTACCTCGCCAGCCTAAC 20  
 CTCGATAAGGTGCTTCTGCTTACAATAAGCACAGGGATAAGCCATCAGGGAGCAGG  
 CAGAAAACATTATCCACTTGTACTCTGACCAACTTGGCGCGCCTGCAGCCTCAA  
 GTACTTCGACACCACCATAGACAGAAAGCGGTACACCTCTACAAAGGAGGTCTGGA  
 CGCCACACTGATTCACTCAGTCAATTACGGGCTCTATGAAACAAGAATCGACCTCT  
 CAGCTCGGTGGAGACAGCAGGCTGACCCCAAGAAGAAGAGGAAGGTGGAGGCCAG  
 CGGTTCCGGACGGCTGACCGCATGGACGATTTGATCTGGATATGCTGGAAAGTGAC  
 GCCCTCGATGATTTGACCTTGACATGCTTGGTTGGATGCCCTGATGACTTTGACCT  
 CGACATGCTCGGAGTOACCCCCCTGATGATTTGACCTGGACATGCTGATTAACCT  
 AGAGCGGGCGCAGATCCAAAAAGAAGAGAAAGGTAGATCCAAAAAGAAGAGAAA 30  
 GGTAGATCCAAAAAGAAGAGAAAGGTAGATACGGCCGATAG

## 【0139】

B. MS2活性化因子コンストラクトと2×MS2アブタマードメインを有する対応するgRNAバックボーンベクターの配列が下に提供されている(NLS、VP64、gRNAスペーサー、およびMS2結合RNAステムループドメインが強調されている)。最大の活性を示すMS2<sub>VP64</sub>N融合タンパク質形式を有する前者の2種類のバージョンを構築した。

## 【0140】

>MS2<sub>VP64</sub>N

gccaccATGGGACCTAAGAAAAAGAGGAAGGTGGCGGCCGCTCTAGAATGGCTTCTAA  
 CTTTACTCAGTCGTTCTCGTCGACAATGGCGGAAGTGGCGACGTGACTGTCGCCCA  
 AGCAACTTCGCTAACGGATCGCTGAATGGATCAGCTCTAACTCGCGTTCACAGGGCTT

10

20

30

40

ACAAAGTAACCTGTAGCGTCAGAGCTCTGCGAGAATCGAAATACACCATAA  
 AGTCGAGGTGCCTAAAGGCCTGGCGTACTAAATATGGAACTAACCATTCAA  
 ATTTGCCACGAATTCCGACTGCGAGCTATTGTAAGGCAATGCAAGGTCTCTAA  
 AAGATGGAAACCCGATTCCCTCAGCAATCGCAGCAAACCTCCGGCATCTACGAGGCCA  
 GCGGTTCCGGACGGGCTGACGCATTGGACGATTTGATCTGGATATGCTGGAAAGTGA  
 CGCCCTCGATGATTTGACCTTGACATGCTGGATGCCCTGATGACTTTGACC  
 TCGACATGCTCGGCAGTGACGCCCTGATGATTTGACCTGGACATGCTGATTAACCTCT

AGATGA

10

【 0 1 4 1 】

>MS2<sub>VP64C</sub>

gccaccATGGGACCTAAGAAAAAGAGGAAGGTGGCGGCCGCTTCTAGAATGGCTTCTAA  
 CTTTACTCAGTCGTTCTCGTCGACAATGGCGGAACGGCGACTGGCGACGTGACTGTCGCCCA  
 AGCAACTTCGCTAACGGGATCGCTGAATGGATCAGCTCTAACCGCGTTACAGGCTT  
 ACAAAAGTAACCTGTAGCGTCAGAGCTCTGCGCAGAAATCGAAATACACCATAA  
 AGTCGAGGTGCCTAAAGGCCTGGCGTACTAAATATGGAACTAACCATTCAA  
 ATTTGCCACGAATTCCGACTGCGAGCTATTGTAAGGCAATGCAAGGTCTCTAA  
 AAGATGGAAACCCGATTCCCTCAGCAATCGCAGCAAACCTCCGGCATCTACGAGGCCA  
 GCGGTTCCGGACGGGCTGACGCATTGGACGATTTGATCTGGATATGCTGGAAAGTGA  
 CGCCCTCGATGATTTGACCTTGACATGCTGGATGCCCTGATGACTTTGACC  
 TCGACATGCTCGGCAGTGACGCCCTGATGATTTGACCTGGACATGCTGATTAACCTCT  
 AGAGCCCCCGCAAGATCCAAAAAAGAAGAGAAAGGTAGATCCAAAAAAGAAAGAGAAA  
 GGTAGATCCAAAAAAGAAGAGAAAGTAGATAACGGCCGATAG

【 0 1 4 2 】

>gRNA<sub>2XMS2</sub>

20

TGTACAAAAAAGCAGGCTTAAAGGAACCAATTCACTGACTGGATCCGGTACCAAG  
 GTCGGGCAGGAAGAGGGCCTATTCCATGATTCCCTCATATTGCATATACGATACA  
 AGGCTGTTAGAGAGATAATTAGAATTAAATTGACTGTAAACACAAAGATATTAGTACA  
 AAATACGTGACGTAGAAAGTAATAATTCTGGTAGTTGCAGTTAAAATTATGTT  
 TTAAAATGGACTATCATATGCTTACCGTAACCTGAAAGTATTGATTTCTGGCTTA  
 TATATCTTGTGAAAGGACGAAACACCGNNNNNNNNNNNNNNNNNNNNNNNNNNNN  
 CTAGAAATAGCAAGTAAAATAAGGCTAGTCCGTTACACTGAAAAAGTGGCACC  
 GAGTCGGTCTGCAGGTCACTCTAGAAAACATGAGGATCACCCATGCTGAGTA  
 TTCCCGGGTTCTAGATCCTAAGGTACCTAATTGCCTAGAAAACATGAGGATCACCC  
 ATGCTGAGGTCAGGACTCTAGAAATTCTAGAC

30

【 0 1 4 3 】

C . d T o m a t o 蛍光ベース転写活性化レポーター配列が下に記載されている ( I S c e I 対照 T F 標的配列、 g R N A 標的配列、最小 C M V プロモーター配列および F L A G タグ + d T o m a t o 配列が強調されている ) 。

【 0 1 4 4 】

40

>TF Reporter 1

TAGGGATAACAGGGTAATAGTGTCCCCCTCCACCCCACAGTGGGGCGAGGTAGGGTGT  
 TACGGTOGGAGGCCTATATAAAGCAGAGCTCGTTAGTGAACCGTCAGATCOCCTGGAG  
 AATTCCgcaccaatgGACTACAAGGATGACGACGATAAAACTTCCGGTGGCGGACTGGGTT  
 CACCGTGAGCAAGGGCGAGGGAGGTATCAAAGAGTTCATGCCTCAAGGTGGCAT  
 GGAGGGCTCCATGAACCGGCCACGGAGTTGGAGATCGAOGGCAGGGCGAOGGCCGCC  
 CTACGAGGGCACCCAGACCGCCAAGCTGAAGGTGACCAAGGGCGGGCTGGCTT  
 CGGCTGGGACATCCTGTCCCCCTGGAGTTCATOTACGGCTCCAAGGGCTACGTGAAGCAC  
 CCCGCCGACATCCCCGATTACAAGAAGCTGTCTTCCCCGAGGGCTTCAAGTGGAGC  
 GCGTGAATGAGGACGGACGGCGGTCTGGTACCGTGACCCAGGACTCCTCCCTGCA  
 GGACCGCACGCTGATCTACAAAGCTGAAGATCGGCCGACCAACTTCCCCCGACGG  
 CCCCGTAATGCAGAAGAAGACCATGGCTGGAGGGCTCCACCGAGCGCTGTACCC  
 CGGGGACCCGCTGCTGAAGGGCGAGATCCACCAAGGGCTGAAGCTGAAGGACCGCG  
 CCACTACCTGGTGGAGTTCAAGACCATCTACATGGCCAAGAAGCCGTGCAACTGCC  
 GGCTACTACTACGTGGACACCAAGCTGGACATCACCTCCCCACAACGAGGGACTACACCA  
 TCGTGGAACAGTACGAGCGCTCCGAGGGCCGCCACCTGTTCTGTACGGCATGGA  
 CGAGCTGTACAAGTAA

10

20

【 0 1 4 5 】

>TF Reporter 2

TAGGGATAACAGGGTAATAGTGGGGCCACTAGGGACAGGATTGGCGAGGTAGGGTGT  
 TACGGTGGGAGGCCTATATAAAGCAGAGCTCGTTAGTGAACCGTCAGATCGCTGGAG  
 AATTCCgcaccaatgGACTACAAGGATGACGACGATAAAACTTCCGGTGGCGGACTGGGTT  
 CACCGTGAGCAAGGGCGAGGGAGGTATCAAAGAGTTCATGCCTCAAGGTGGCAT  
 GGAGGGCTCCATGAACGGGCCACGGAGTTGGAGATCGAGGGCGAOGGCCGCC  
 CTACGAGGGCACCCAGACCGCCAAGCTGAAGGTGACCAAGGGCGGGCTGGCTT  
 CGGCTGGGACATCCTGTCCCCCTGGAGTTCATGTACGGCTCCAAGGGCTACGTGAAGCAC  
 CCCGCCGACATCCCCGATTACAAGAAGCTGTCTTCCCCGAGGGCTTCAAGTGGAGC  
 GCGTGAATGAGGACGGACGGCGGTCTGGTACCGTGACCCAGGACTCCTCCCTGCA  
 GGACCGCACGCTGATCTACAAAGGTGAAGATCGCCGCCACCAACTTCCCCCGACGG  
 CCCCGTAATGCAGAAGAAGACCATGGGCTGGAGGGCCACCGAGCGGCCGTACCC  
 CGGGGACCCGCTGCTGAAGGGCGAGATCCACCAAGGGCTGAAGCTGAAGGACGGCGG  
 CCACTACCTGGTGGAGTTCAAGACCATCTACATGGCCAAGAAGCCGTGCAACTGCC  
 GGCTACTACTACGTGGACACCAAGCTGGACATCACCTCCCCACAACGAGGGACTACACCA  
 TCGTGGAACAGTACGAGCGCTCCGAGGGCCGCCACCTGTTCTGTACGGCATGGA  
 CGAGCTGTACAAGTAA

30

40

【 0 1 4 6 】

D . T A L E および C a s 9 - g R N A の特異性アッセイに使用されるレポーター ライブリーアの全般的形式が下に提供されている ( I S c e I 対照 T F 標的配列、 g R N A / T A L E 標的部位配列 ( g R N A については 23 b p 、および T A L E については 18

50

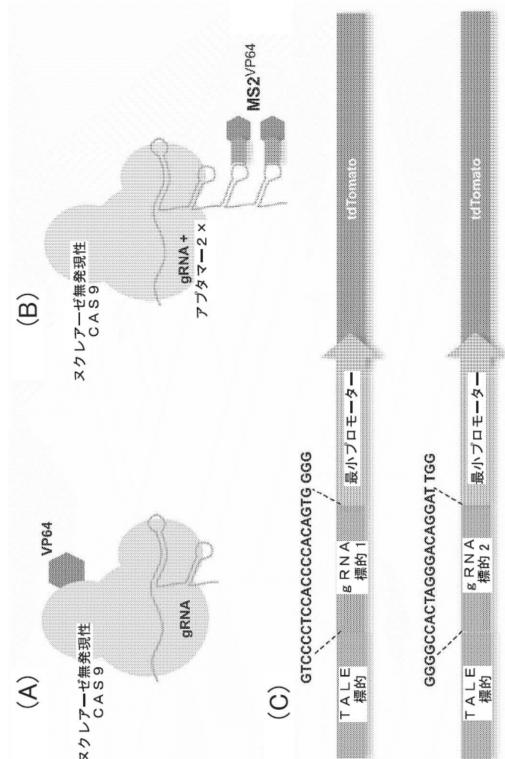
b p )、最小CMVプロモーター配列、RNAバーコード配列、およびdTomo配列が強調されている)。

## 【0147】

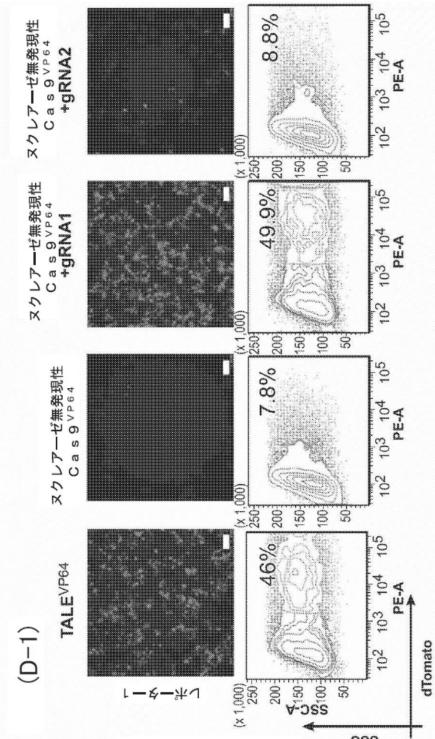
> Specificity Reporter Libraries

TAGGGATAACAGGGTAATAGTNNNNNNNNNNNNNNNNNNNCAGGGTAGGCCT  
 GTACGGTGGAGGCCTATATAACAGAGCTCGTTAGTGAACCGTCAGATCGCCTGGA  
 GAATTGccaccatgGACTACAAGGATGACGACGATAAANNNNNNNNNNNNNNNNN  
 NNNNACTTCCGGTGGCGACTGGGTTCCACCGTGAGCAAGGGCGAGGAGGTCAAA 10  
 AGAGTTCATGCGCTTCAAGGTGCGCATGGAGGGCTCCATGAACGGCCACGAGTCGA  
 GATCGAGGGCGAGGGCGAGGGCCCTACGAGGGCACCCAGACCGCCAAGCTGAA  
 GGTGACCAAGGGCGGCCCCCTGCCCCCTGCGCTOGGACATCTGTCCCCCAAGTTCATG  
 TACGGCTCCAAGGCCTACGTGAAGCACCCGCCGACATCCCCGATTACAAGAAGCTGT  
 CCTTCCCCGAGGGCTTCAAGTGGGAGCGCGTGTGAACCTCGAGGACGGCGGTCTGGT  
 GACCGTGACCCAGGACTCCTCCCTOCAAGGACGGCACCGTGTGATCTACAAGGTOAAGAT  
 GCGCGOCACCAACTTCCCCCGACGGCCCCGTAATGCAGAAGAAGACCATGGCTG 20  
 GGAGGCCTCCACCGAGCGCCTGTACCCCCCGACGGCGGTGTGAAGGGCGAGATCCA  
 CCAGGCGCTGAAAGCTGAAGGACGGCGGACTACCTGGTGGAGTTCAAGACCATCTA  
 CATGGCCAAGAAGGCCGTGCAACTGCGGCGCTACTACTACGTGGACACCCAAGCTGGA  
 CATCACCTCCCACAACGAGGACTACACCATGGAACAGTACGAGGGCTCCGAGGG  
 CGGCCACCACCTGTTCTGTACGGCATGGACGGCTGTACAAGTAAGAATTG

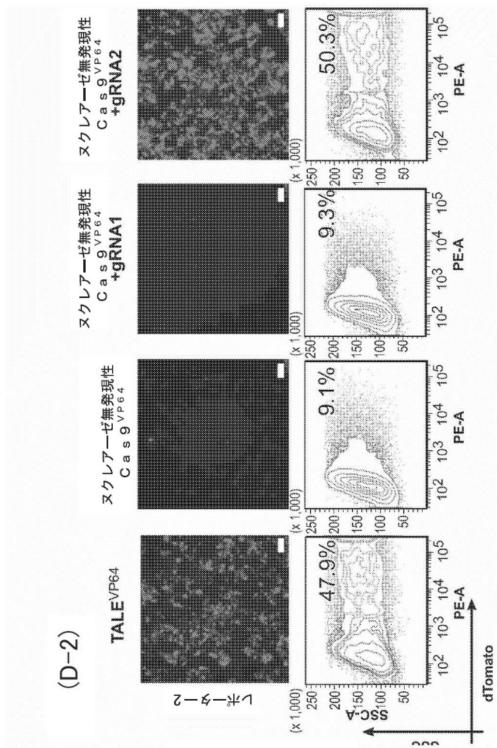
【図1-1】



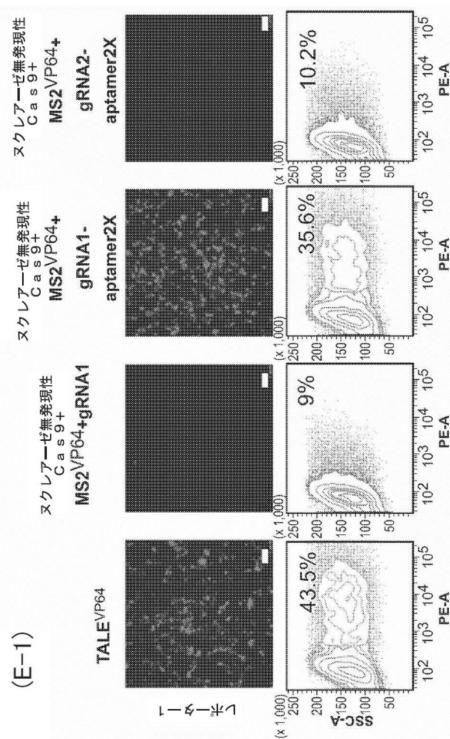
【図1-2】



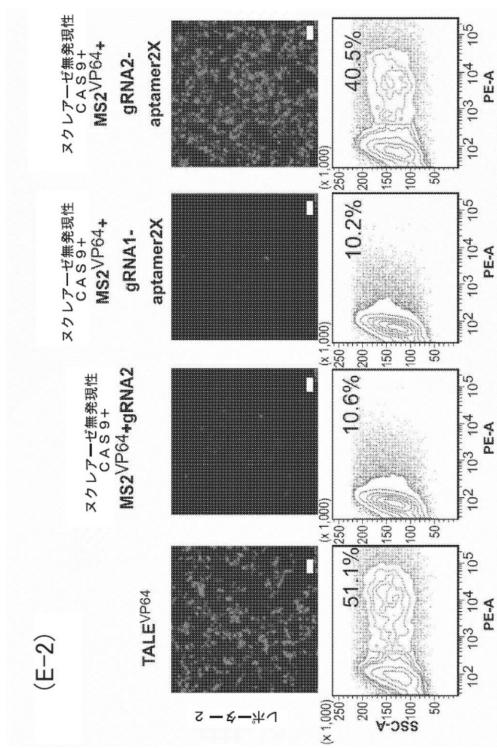
【図 1 - 3】



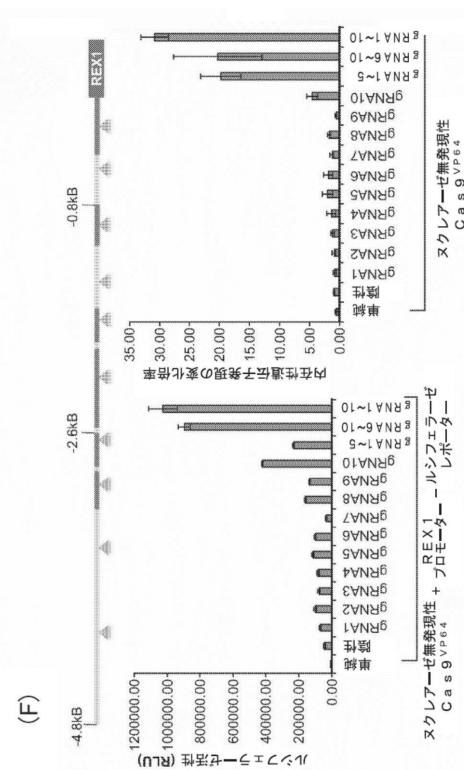
【図 1 - 4】



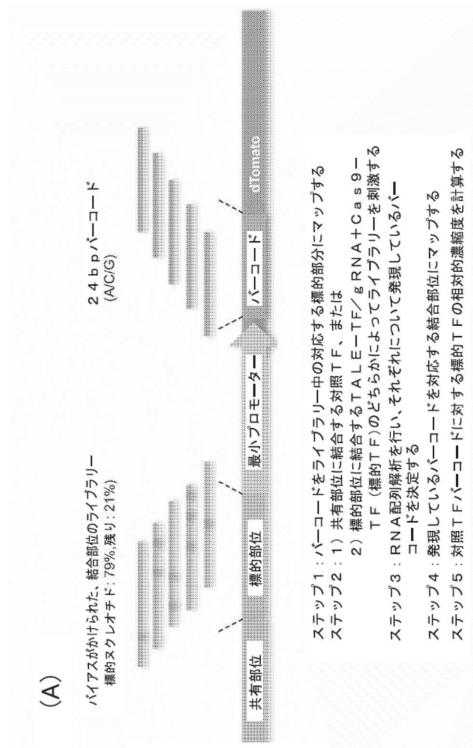
【図 1 - 5】



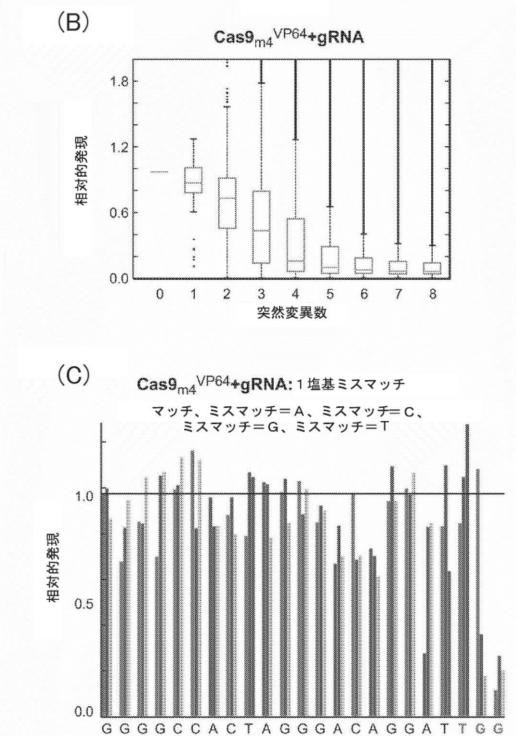
【図 1 - 6】



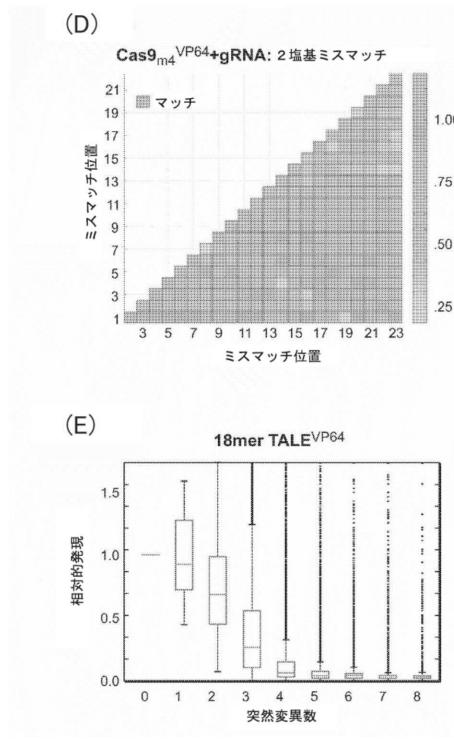
### 【図 2 - 1】



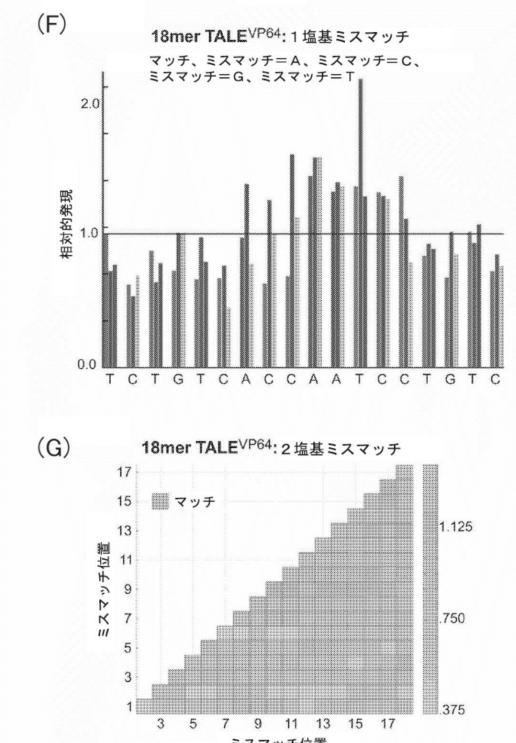
### 【図2-2】



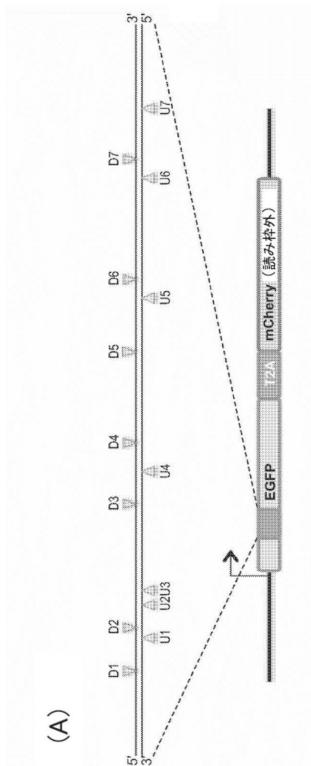
### 【図2-3】



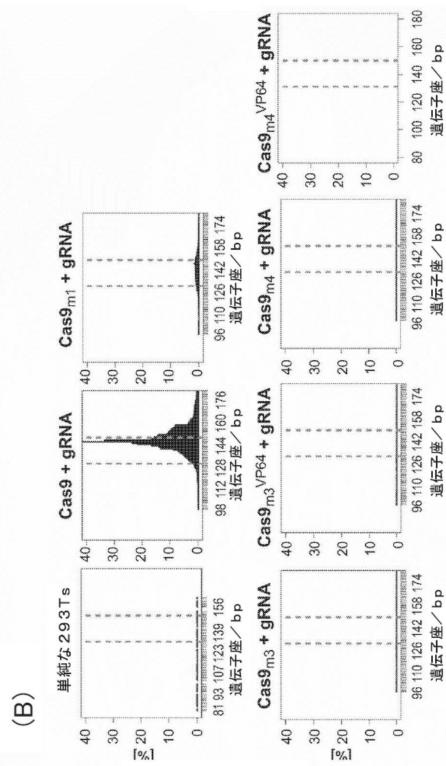
### 【図2-4】



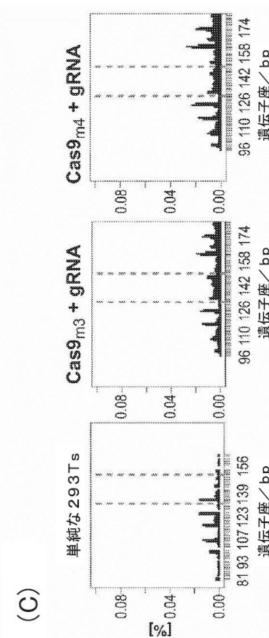
【図3-1】



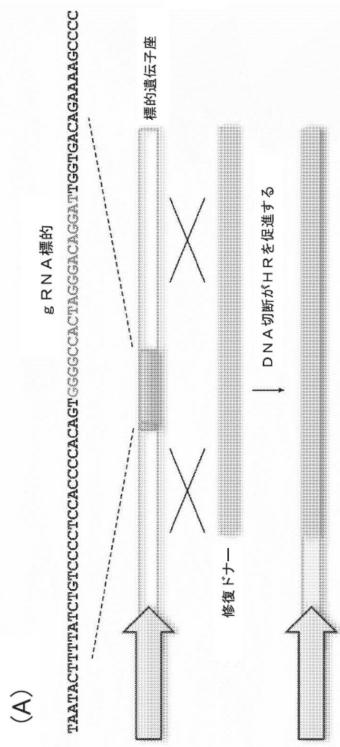
【図4-2】



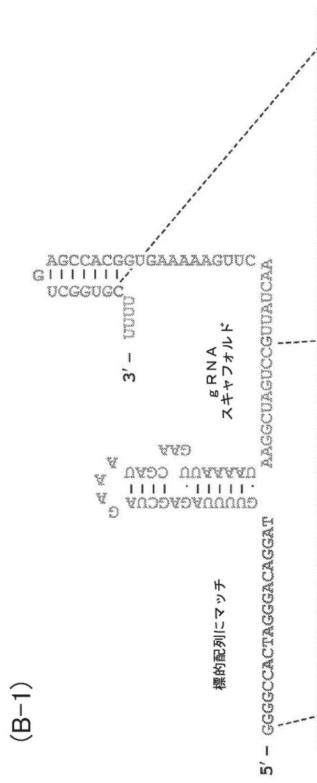
【図4-3】



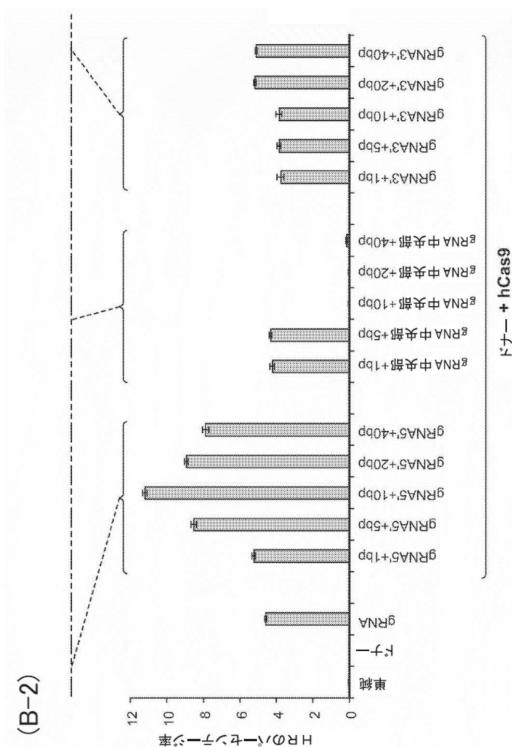
【図5-1】



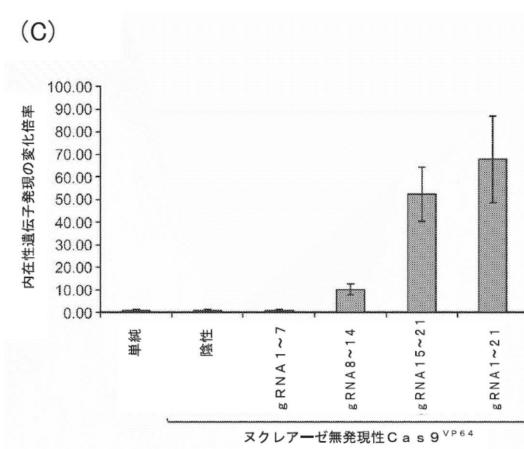
【図5-2】



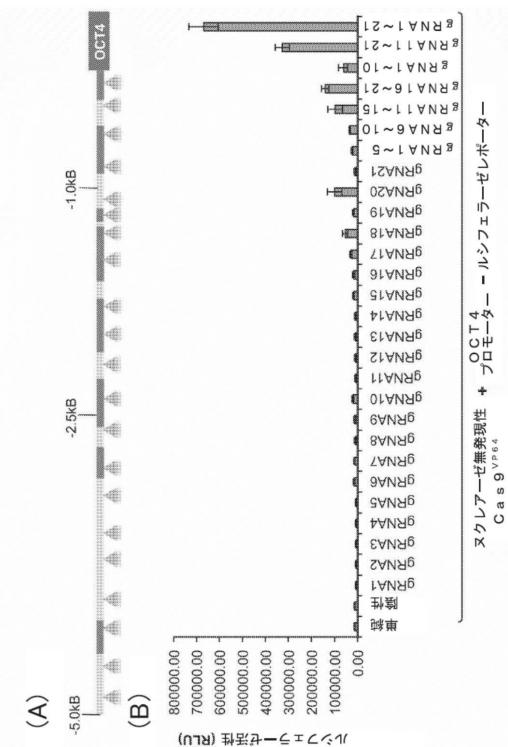
【 図 5 - 3 】



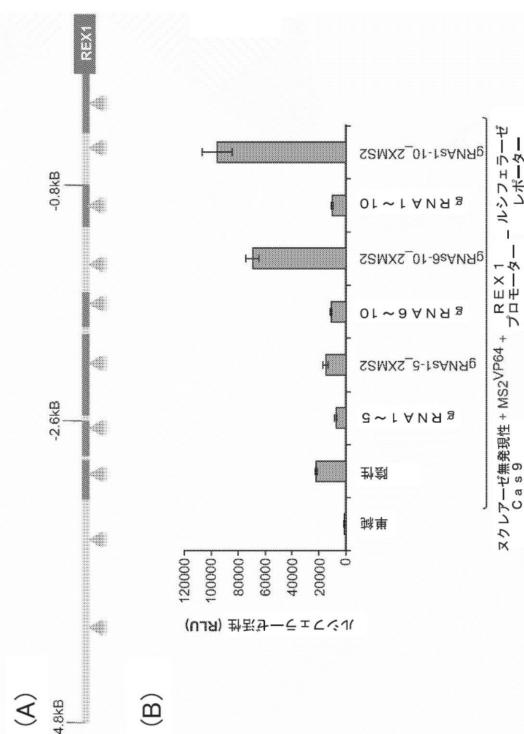
【 図 6 - 2 】



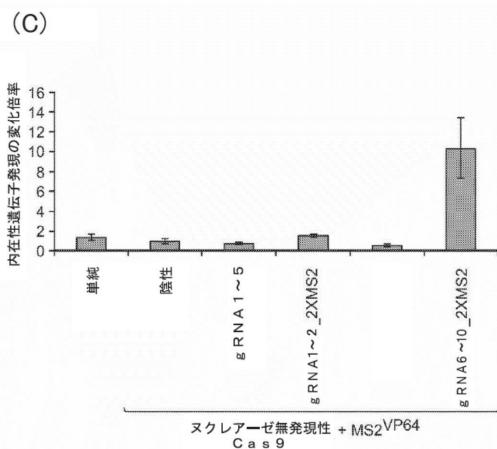
【図 6 - 1】



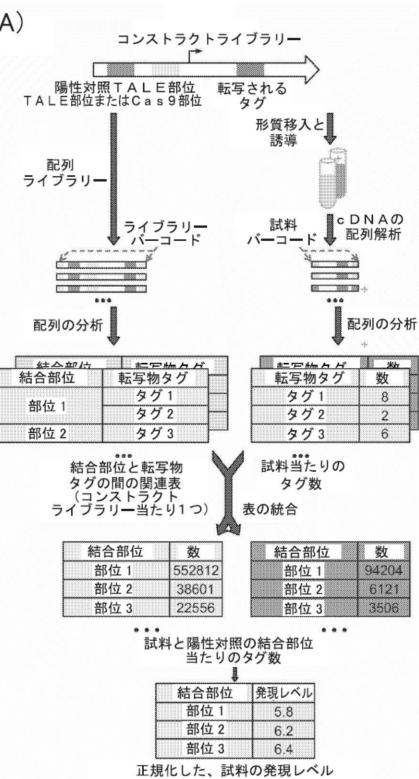
### 【 図 7 - 1 】



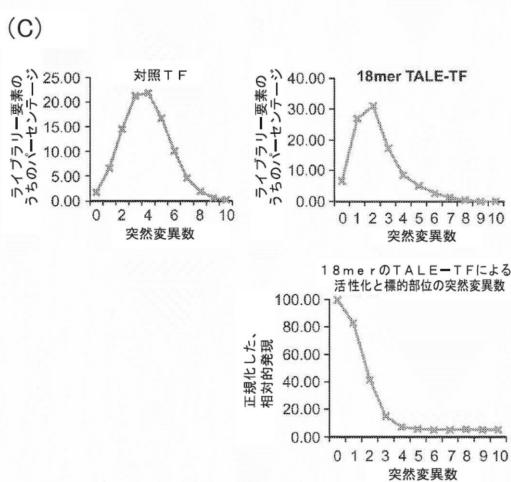
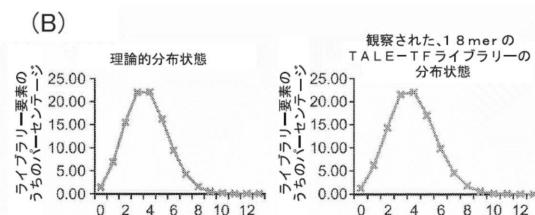
【図7-2】



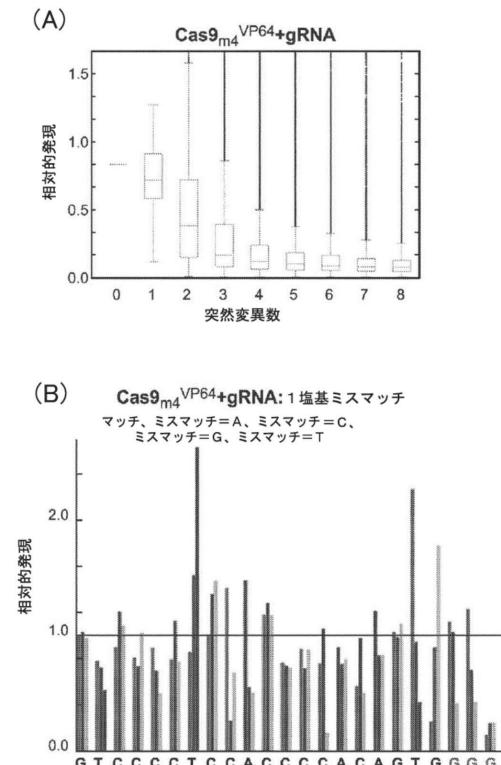
【図8-1】



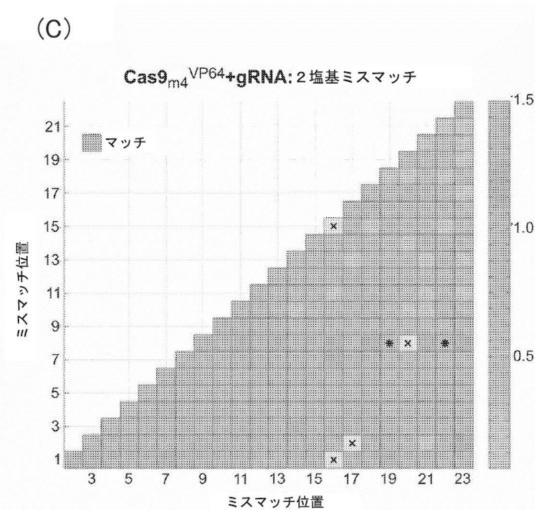
【図8-2】



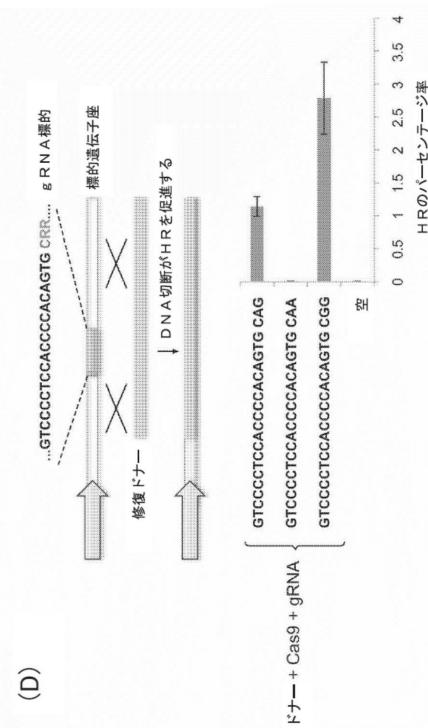
【図9-1】



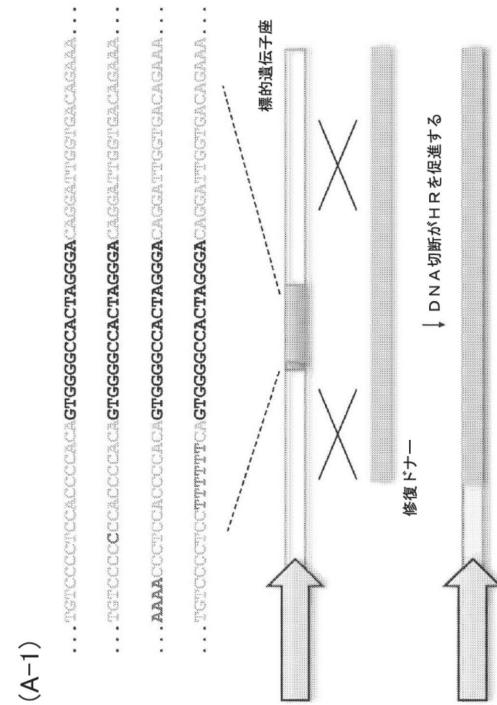
【図 9 - 2】



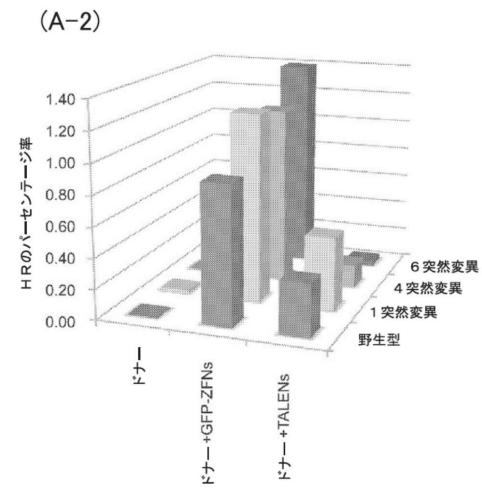
【図 9 - 3】



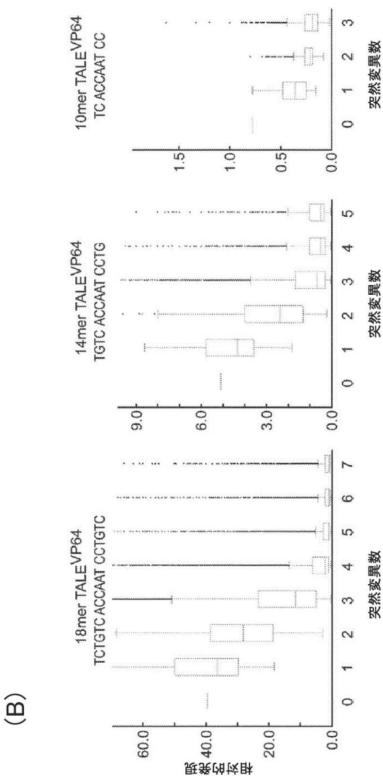
【図 10 - 1】



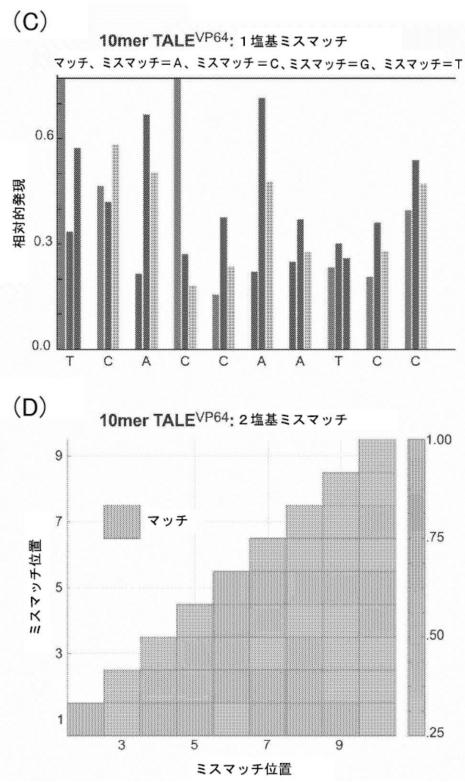
【図 10 - 2】



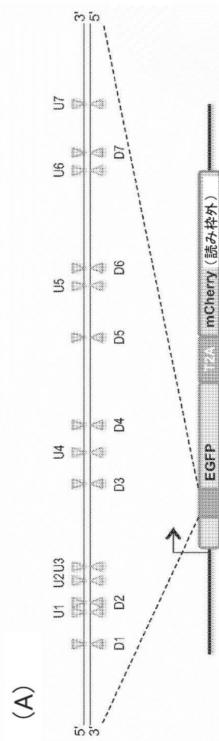
【図10-3】



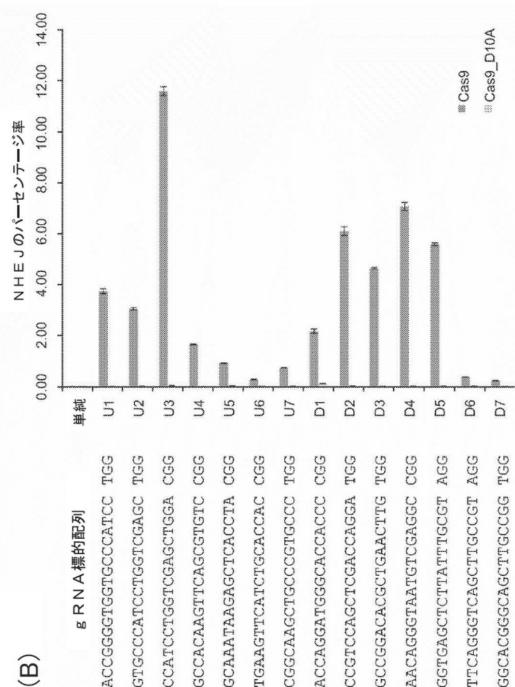
【図10-4】



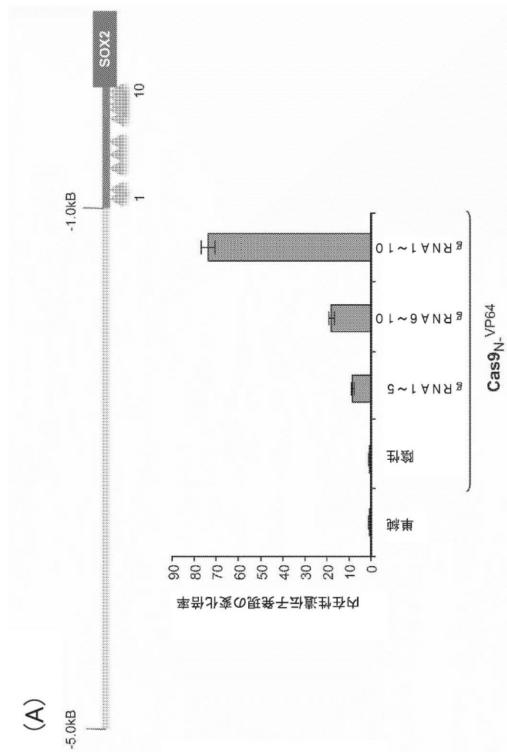
【図11-1】



【図11-2】

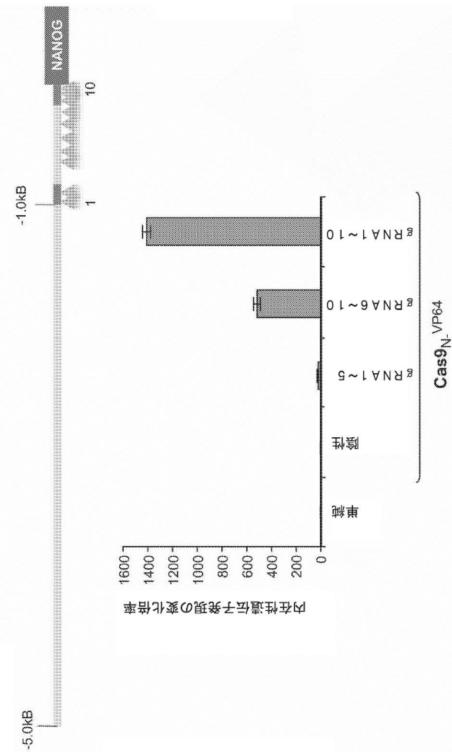


【図12-1】



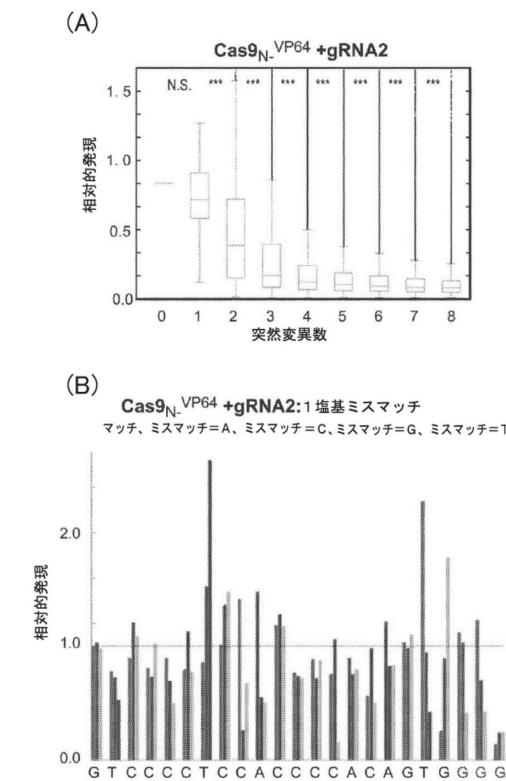
(A)

【図12-2】

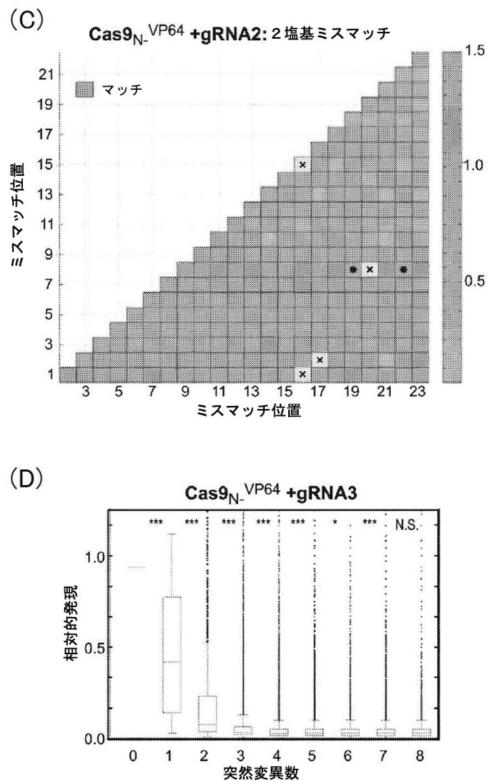


(B)

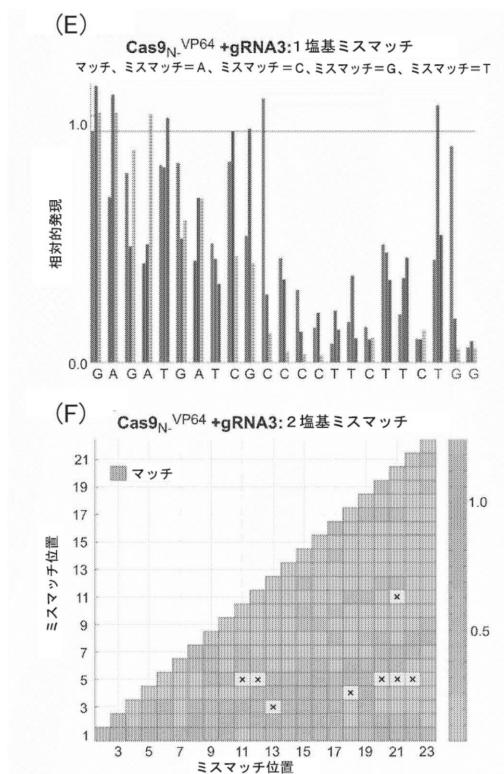
【図13-1】



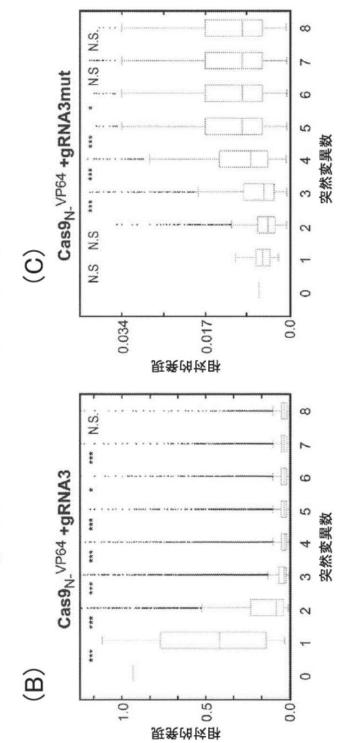
【図13-2】



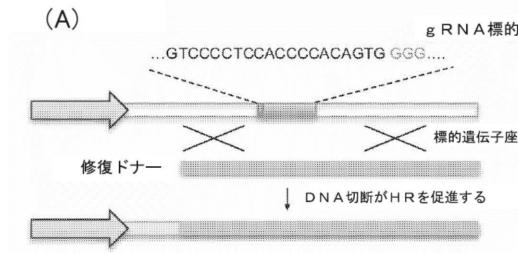
### 【図 1-3】



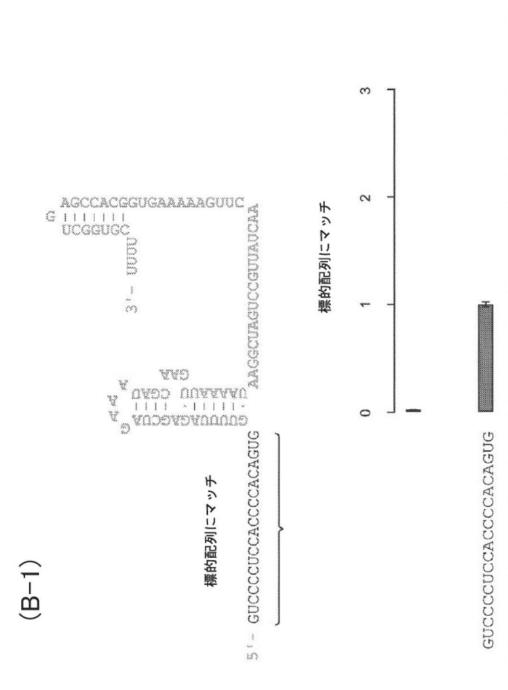
【図14】



【図 15-1】

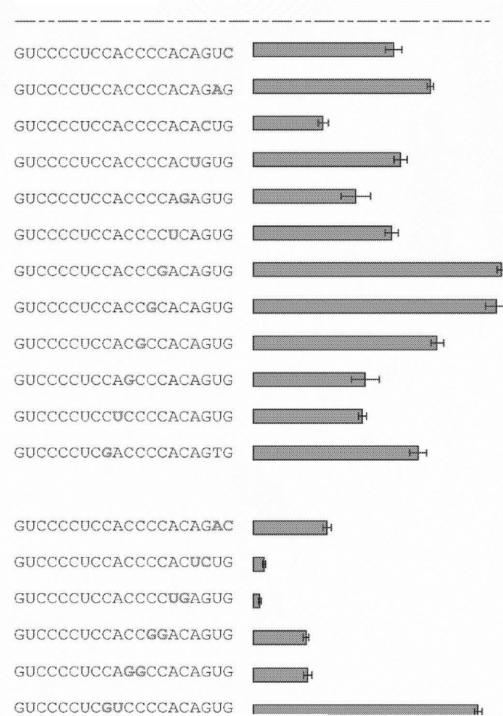


【図 15-2】

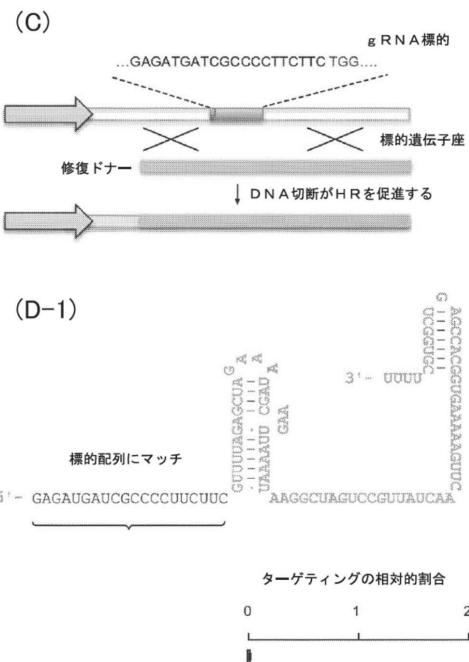


【図 15-3】

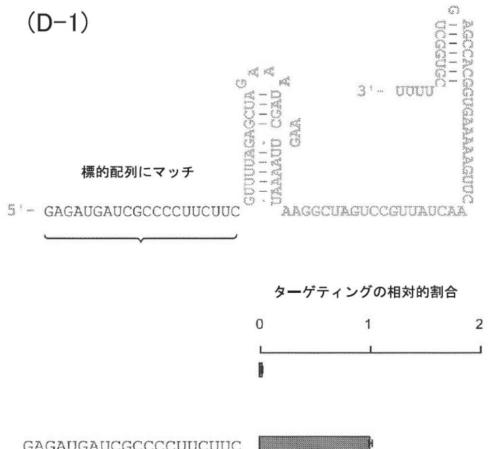
(B-2)



【図 15-4】

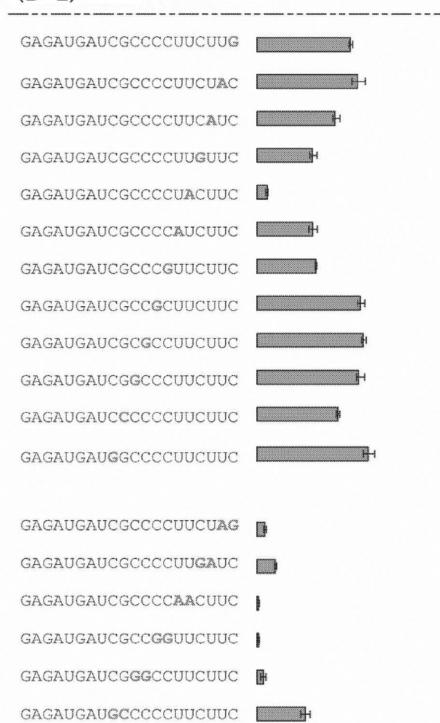


(D-1)

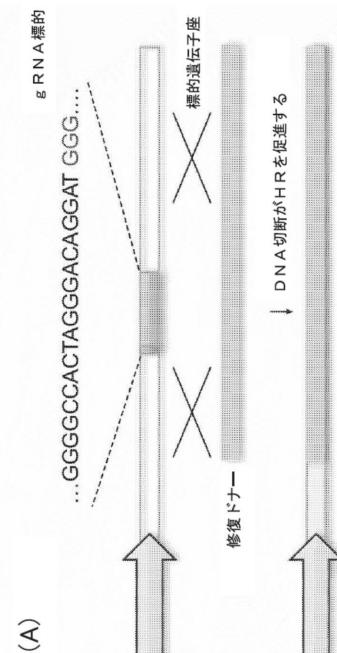


【図 15-5】

(D-2)

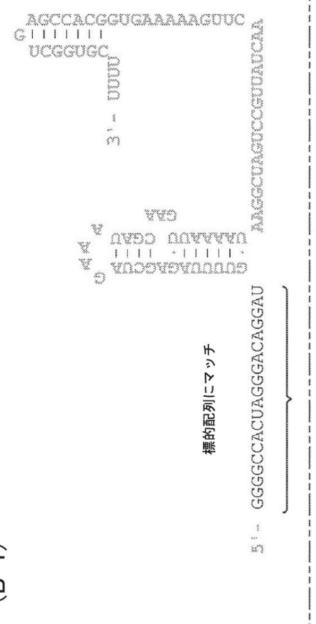


【図 16-1】

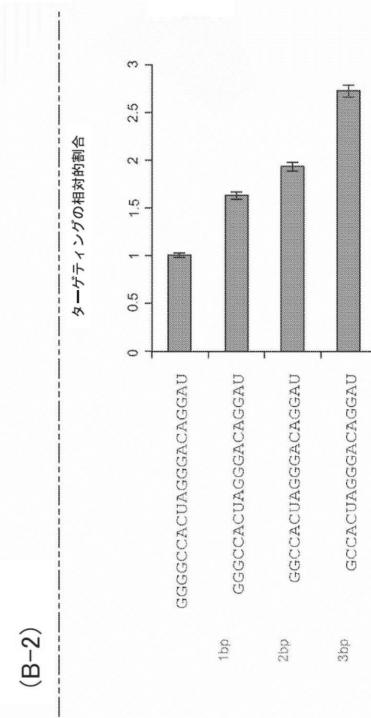


(A)

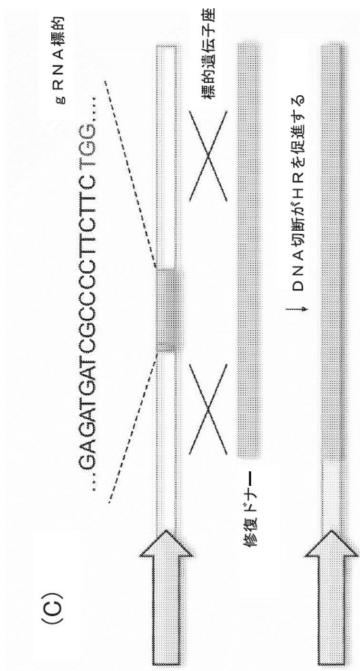
【図 1 6 - 2】



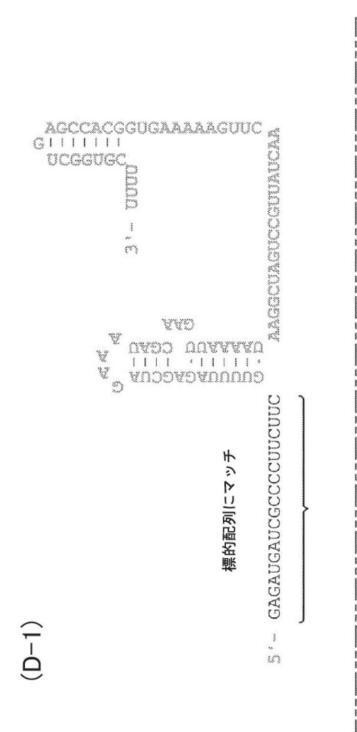
【図 1 6 - 3】



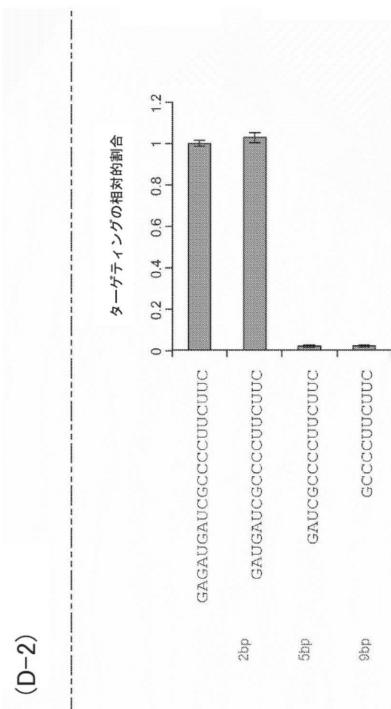
【図 1 6 - 4】



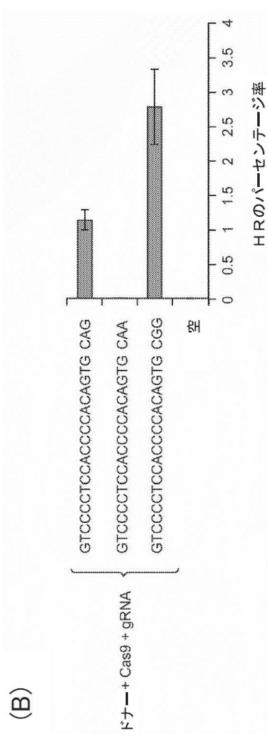
【図 1 6 - 5】



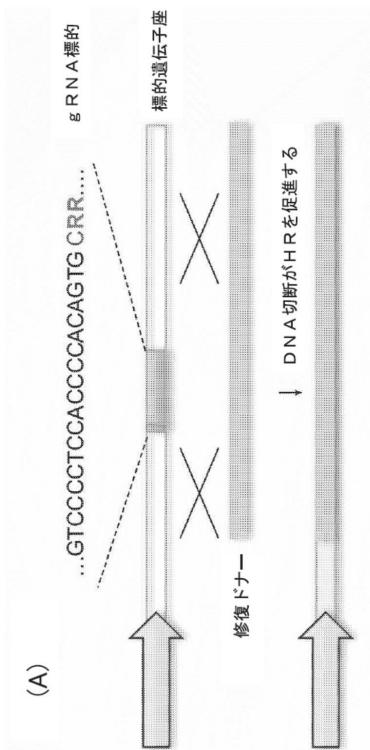
【図 1 6 - 6】



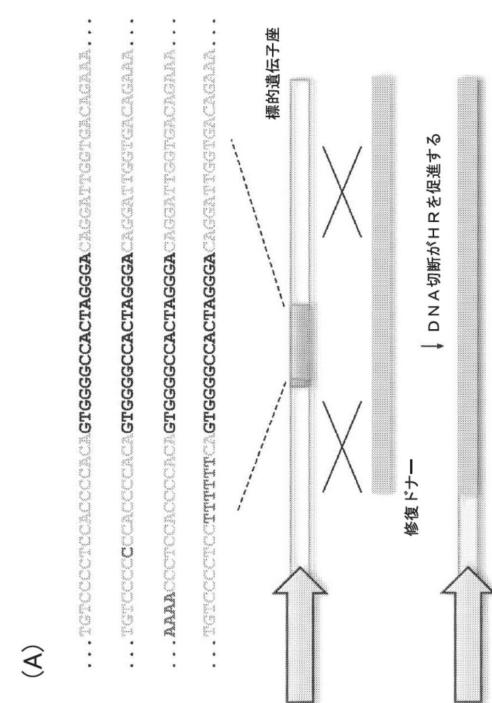
【図 1 7 - 2】



【図 1 7 - 1】

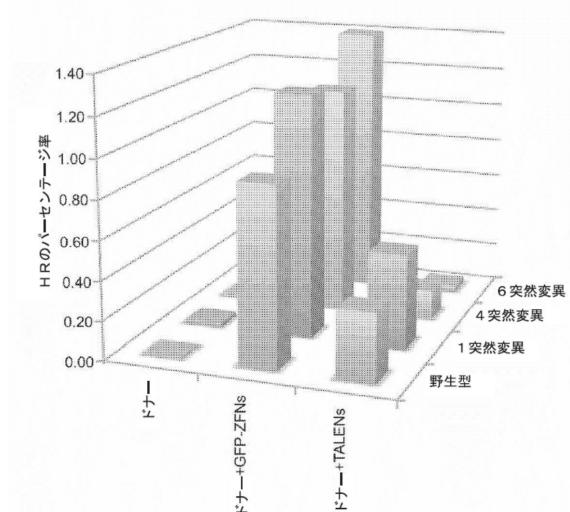


【図 1 8 - 1】

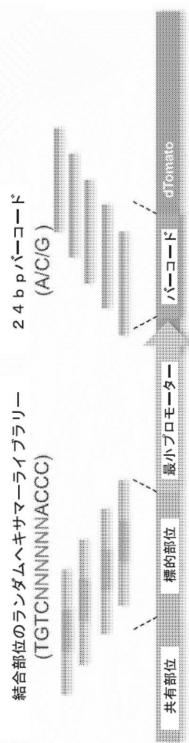


【図18-2】

(B)

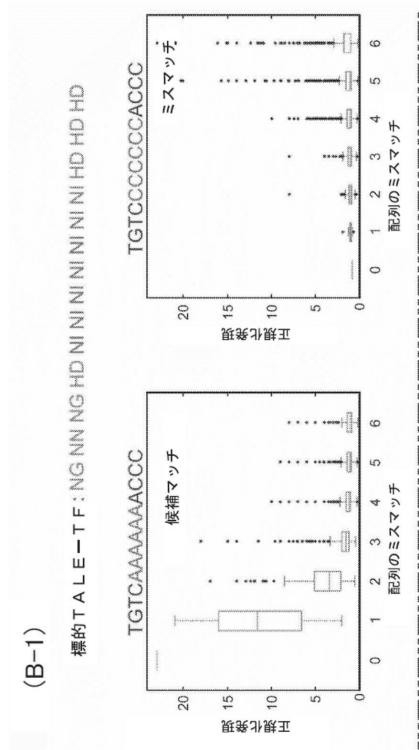


【図19-1】

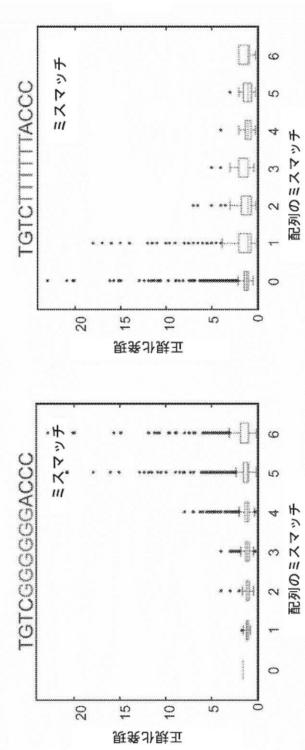


(A)

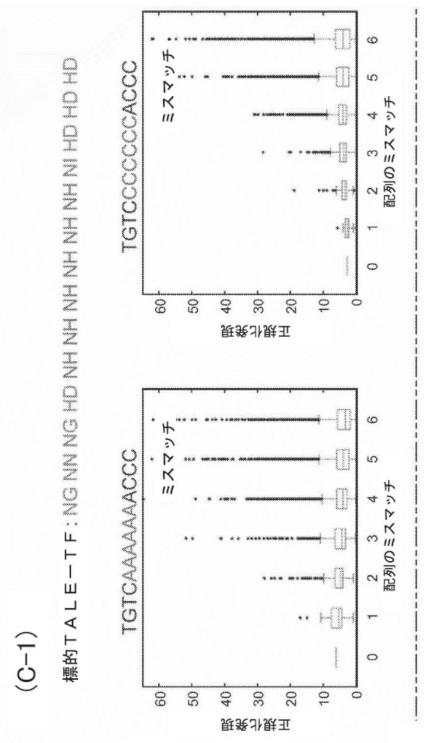
【図19-2】



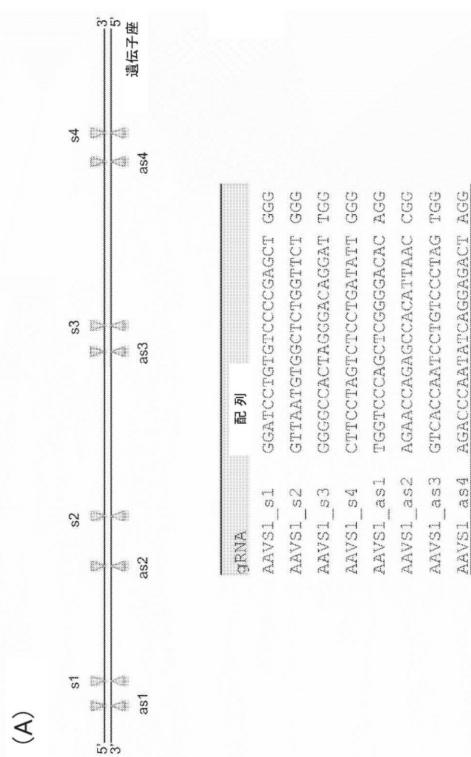
【図19-3】



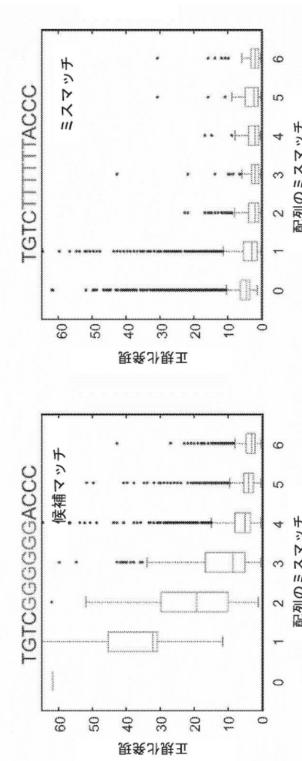
【図19-4】



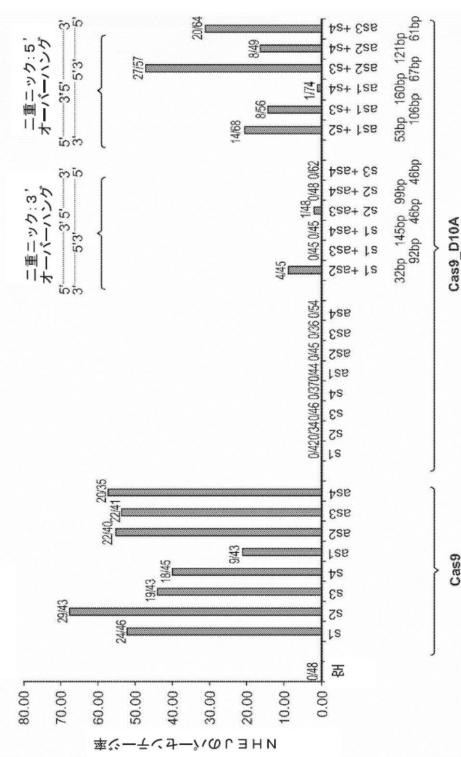
### 【図20-1】



### 【図 19-5】



### 【図20-2】



### 【図2.1-1】

(A)

【図21-2】

【図21-3】

6

---

フロントページの続き

(72)発明者 マリ、 パラシャント ジー。  
アメリカ合衆国 02143 マサチューセッツ州 サマービル ピーコン ストリート 88  
(72)発明者 エスペルト、 ケビン エム。  
アメリカ合衆国 02138 マサチューセッツ州 ケンブリッジ トウブリッジ ストリート  
20 アパート 2

審査官 飯室 里美

(56)参考文献 国際公開第2013/176772 (WO, A1)  
国際公開第2014/089290 (WO, A1)  
国際公開第2014/093622 (WO, A1)  
特表2015-523856 (JP, A)  
特表2016-502840 (JP, A)  
特表2016-501531 (JP, A)  
Cell, 2013.02.28, Vol.152, p.1173-1183  
Nature Methods, 2013-MAR, Vol.10, p.239-242

(58)調査した分野(Int.Cl., DB名)

C12N 15/00  
JSTPlus / JMEDPlus / JST7580 (JDreamIII)  
Caplus / MEDLINE / EMBASE / BIOSIS (STN)