

(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 I
C 1 2 N 15/09 (2006. 01)	C 1 2 N 15/09 Z N A Z
C 1 2 N 1/15 (2006. 01)	C 1 2 N 15/09 1 1 O
C 1 2 N 1/19 (2006. 01)	C 1 2 N 1/15
C 1 2 N 5/10 (2006. 01)	C 1 2 N 1/19
	C 1 2 N 5/10

請求項の数 21 (全 62 頁)

(21) 出願番号	特願2016-517954 (P2016-517954)	(73) 特許権者	507044516
(86) (22) 出願日	平成26年6月4日 (2014. 6. 4)		プレジデント アンド フェローズ オブ ハーバード カレッジ
(65) 公表番号	特表2016-521554 (P2016-521554A)		アメリカ合衆国 マサチューセッツ O 2 1 3 8, ケンブリッジ, クインシー ストリート 1 7
(43) 公表日	平成28年7月25日 (2016. 7. 25)	(74) 代理人	100079049
(86) 国際出願番号	PCT/US2014/040868		弁理士 中島 淳
(87) 国際公開番号	W02014/197568	(74) 代理人	100084995
(87) 国際公開日	平成26年12月11日 (2014. 12. 11)		弁理士 加藤 和詳
審査請求日	平成29年6月5日 (2017. 6. 5)	(72) 発明者	チャーチ、 ジョージ エム. アメリカ合衆国 O 2 4 4 6 マサチュー セッツ州 ブルックリン ケント ストリ ート 2 1 8
(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)		

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-null) 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。

【請求項 7】

前記標的核酸が疾患または有害な健康状態に関連する、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 RNA c r RNA 融合体である、請求項 1 に記載の方法。

【請求項 9】

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 1 に記載の方法。

【請求項 10】

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 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 MS2 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 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、

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、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 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、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。

【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に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 RNA c r RNA 融合体である、請求項 13 に記載の細胞。

【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】

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

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 RNase III. 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~E2586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); Sapranaukas, 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

10

20

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 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に関連する。

【0010】

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スメイト配列(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はトレーサースメイト配列とtracr配列を連結(link)してtracrRNA crRNA融合体を作製するリンカー核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

【0011】

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。

20

【0012】

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。

【0013】

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。

30

【0014】

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

【0015】

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

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 t r a c r 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 t r a c r 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 t r a c r 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 t r a c r 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 t r a c r 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 t r a c r 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 5 B 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 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、II 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、II 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、II 型 CRISPR システムの RNA 誘導型ヌクレアーゼ欠損 DNA 結合タンパク質をコードする前記外来核酸は、II 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に関連する。

【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 RNA c r RNA 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に関連する。

【 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 RNA c r RNA 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 】

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に関連する。

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及び前記少なくとも1種類のCas9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Cas9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

30

【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。

40

【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R

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 10 ケースをコードする第 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、それぞれが 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 20 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 30 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 40 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る。

【 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 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。

【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。

【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。

【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ドスケー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。

【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。

【図3-3】図3C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

10

20

30

40

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。

【0060】

【図 5 - 1】図 5 A は Cas 9 - gRNA 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 RNA 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 RNA と相同組換えのパーセンテージ率を示す。

10

【0061】

【図 6 - 1】図 6 A は OCT 4 遺伝子のガイド RNA 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 qPCR によって内在性遺伝子の転写活性化を示す。

【0062】

【図 7 - 1】図 7 A は REX 1 遺伝子のガイド RNA 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。

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

【0063】

【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 TALE 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 TALE がガイドされた試料から観察された分布状態。

【0064】

【図 9 - 1】図 9 A は Cas 9 - gRNA 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ープの分析データを示す。図 9 B は PAM 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、Cas 9 - gRNA 複合体のターゲティングランドスケー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、Cas 9 - gRNA 複合体のターゲティングランドスケー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. pyogenes) の Cas 9 の予想される PAM が NGG であり、また NAG であることを確認する、ヌクレアーゼ介在性 HR アッセイのデータを示す。

【0065】

【図 10 - 1】図 10 A は 18 - mer の TALE 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 HR アッセイのデータを示す。

40

【図 10 - 2】図 10 A は 18 - mer の TALE 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 HR 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 (18 - mer、14 - mer および 10 - mer) の TALE のターゲティングランドスケー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 10 - mer の TALE 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 10 - mer の TALE のヒートプロットデータを示す。

【0066】

【図 11 - 1】図 11 A はデザインされたガイド RNA を示す。

50

【図 1 1 - 2】図 1 1 B は様々なガイド R N A 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】

【 0 0 6 7 】

本開示の実施形態は、標的核酸の制御法において D N A 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 D N A 結合タンパク質を使用することに基づく。そのような D N A 結合タンパク質が様々な目的のために D N A に結合することは当業者に既知である。そのような D N A 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 D N A 結合タンパク質にはここでガイド R N A と称される R N A によってガイドされ得るものが含まれる。この態様によると、ガイド R N A と R N A 誘導型 D N A 結合タンパク質は D N A 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、D N A 結合タンパク質はヌクレアーゼ欠損 D N A 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 D N A 結合タンパク質ヌクレアーゼ活性を有する D N A 結合タンパク質の改変(alteration)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 D N A 結合タンパク質は当業者に知られており、例えば I I 型 C R I S P R システムに存在する、C a s 9 タンパク質のような、ヌクレアーゼ活性を有する天然の D N A 結合タンパク質を含む。そのような C a s 9 タンパク質と I I 型 C R I S P 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 M a k a r o v a ら著、Nature Reviews, Microbiology 誌、第 9 巻、2 0 1 1 年 6 月、4 6 7 ~ 4 7 7 頁を参照されたい。

【 0 0 6 8 】

ヌクレアーゼ活性を有する例示的な D N A 結合タンパク質は二本鎖 D N A にニックを入れるかまたは二本鎖 D N A 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 D N A 結合タンパク質から生じ得る。そのような例示的な D N A 結合タンパク質はそれぞれのドメインが二本鎖 D N A 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 M c r A - H N H ヌクレアーゼ関連ドメインおよび R u v C 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 D N A 結合タンパク質は自然界において M c r A - H N H ヌクレアーゼ関連ドメインと R u v C 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 D N A 結合タンパク質はヌクレアーゼ活性を不活化するために改変されるか、あるいは改造される。そのような改変または改造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。そのような改造には、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 D N A 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化するための他の改造は、本開示に基づいて当業者にすぐに明らかになる。したがって、ヌクレアーゼ欠損 D N A 結合タンパク質にはヌクレアーゼ活性を不活化するように改造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 D N A 結合タンパク質はヌクレアーゼ活性が不活化されていても D N A 結合能を保持している。したがって、前記 D N A 結合タンパク質は D N A 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よい。したがって、前記 D N A 結合タンパク質は D N A 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、不活化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よい。

【 0 0 6 9 】

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 D N A 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化されているように改造または改変さ

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Jinke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・エフィシエンシ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 Bielefeld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・クロップステッティイ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・オパカスB4 uid36573株；アシドサーマス・セルロリチカス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；サーモノスポラ・カーバタDSM43183株；ピフィドバクテリウム・デンティウムBd1株；ピ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・キャステンホルツィイ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・ミヌトゥム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・エクイ亜種ズーエビデミカスMGCS10565株；ストレプトコッカス・ガロリチカス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 (S t r e p t o c o c c u s t h e r m o p h i l e
 s) 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 (S t r e p t o c o c
 c u s t h e r m o p h i l e s) L M D - 9 株；ストレプトコッカス・サーモフィル
 ス (S t r e p t o c o c c u s t h e r m o p h i l e s) 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
 イカム 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；ユウバクテリアム・
 レクター 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・ベネトランス；マイコプラズマ・シノピアエ 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；ロドシュードモナス・パルス
 トリス B i s B 1 8 株；ロドシュードモナス・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・ルブラム 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；フェルミネフォロバクター・エイセニア E F 0 1 - 2
 株；ナイセリア・メニンギティデス (N e i s s e r i a m e n i n g i t i d e s)
 0 5 3 4 4 2 株；ナイセリア・メニンギティデス (N e i s s e r i a m e n i n g i
 t i d e s) a l p h a 1 4 株；ナイセリア・メニンギティデス (N e i s s e r i a
 m e n i n g i t i d e s) 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 Doyley 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；ウォリネラ・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 . p y o g e n e s) 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 巻、6 0 2 ~ 6 0 7 頁 (2 0 1 1 年) を参照されたい。

【 0 0 7 3 】

MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAE
 ATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRL EESFLVEEDKKHERHPIFG
 NIVDEVAYHEKYPTIYHLRKKLVDSTDKADRLIYLALAHMIKFRGHFLIEGDLNPDNSD
 VDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRL ENLIAQLPGEKKNGLFGN
 LIALSLGLTPNFKSNFDLAEDAKLQLSKD TYDDDLNLLAQIGDQYADLFLAAKNLSDAI
 LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYA
 GYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELH
 AILRRQEDFYFPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEE
 VVDKGASQAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFL
 SGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
 IKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWG
 RLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSL
 HEHIANLAGSPAIKKGILQTVKVVDDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRER
 MKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDH
 IVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVKKMKNYWRQLLNAKLITQRKFDNL
 TKAERGGSELDDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS
 KLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVVDVR
 K
 MIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDF
 ATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVA
 YSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLI IKLPK
 YSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEQNEQKQLFVE
 QHKHYLDEIIEQISEFSKRVILADANLDKVL SAYNKHRRDKPIREQAENIHLFTLTNLGA
 PAAFKYFDTTIDRKRYTSTKEVL DATLIHQSI TGLYETRIDLSQLGGD-

10

20

30

【 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。

40

【 0 0 7 5 】

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 巻、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 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 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_{m 1} ; D 1 0 と H 8 4 0 についてアラニン

で置換した Cas9_{m2}; D10、H840、および D839 についてアラニンで置換した Cas9_{m3}; 並びに、D10、H840、D839、および N863 についてアラニンで置換した Cas9_{m4}。

【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ープが未改変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ら著、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と組み合わせさせた時に、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

NA 部分の 3' 末端のどちらかに無作為配列挿入物を有する gRNA は機能性を保持し、一方でキメラ gRNA の tracrRNA スキャフォー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 gRNA の柔軟性についてまとめた図 5A ~ B を参照されたい。図 5A は Cas9 gRNA 活性を測定するための相同組換え (HR) アッセイの模式図である。図 5B に示されるように、キメラ gRNA の crRNA 部分の 5' 末端または tracrRNA 部分の 3' 末端のどちらかに無作為配列挿入物を有する gRNA は機能性を保持し、一方でキメラ gRNA の tracrRNA スキャフォールド部分に挿入がなされると機能が失われる。gRNA 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 gRNA が長くなって半減期が増したことに起因するものであり得る。

10

【0084】

VP64 を gRNA に結合させるため、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 ステムループをその gRNA の 3' 末端に取り付けた。ここにその全体が参照により援用される Fuscó ら著、Current Biology 誌：第 CB13 巻、161 ~ 167 頁 (2003 年) を参照されたい。これらのキメラ gRNA は Cas9N と MS2-VP64 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。

【0085】

図 1A は RNA 誘導性転写活性化の模式図である。図 1A に示されるように、転写活性化可能な Cas9N 融合タンパク質を作製するため、VP64 活性化ドメインを Cas9N の C 末端に直接連結した。図 1B に示されるように、転写活性化可能な gRNA 連結物を作製するため、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 ステムループをその gRNA の 3' 末端に取り付けた。これらのキメラ gRNA は Cas9N と MS2-VP64 融合タンパク質と共に発現した。図 1C 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 gRNA 標的部位を有し、且つ、対照 TALE-TF 標的部位を共有する。図 1D に示されるように、蛍光活性化細胞選別 (FACS) および免疫蛍光分析 (IF) の両方によってアッセイすると、Cas9N-VP64 融合体は RNA 誘導性転写活性化を示す。具体的には、対照 TALE-TF が両方のレポーターを活性化した一方で、Cas9N-VP64 融合体は gRNA 配列特異的にレポーターを活性化する。図 1E に示されるように、3 種類全ての要素、すなわち、Cas9N、MS2-VP64、および適切な MS2 アプタマー結合部位を有する gRNA が存在するときのみ、レポーターコンストラクトの gRNA 配列特異的転写活性化が FACS および IF の両方によって観察された。

20

30

【0086】

ある特定の態様によれば、Cas9N、1 つ以上の gRNA 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、多能性の維持に両方とも厳密に制御されている遺伝子である ZFP42 (REX1) および POU5F1 (OCT4) が含まれた。図 1F 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 (DNase 高感受性部位が緑色で強調されている) を標的とする 10 種類の gRNA を REX1 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照により援用される Takahashi ら著、Cell 誌、第 131 巻、861 ~ 872 頁 (2007 年) を参照のこと) 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 qPCR により転写活性化を直接的にアッセイした。

40

【0087】

図 6A ~ C は Cas9N-VP64 を使用する RNA 誘導性 OCT4 制御に関する。図

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gRNA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gRNA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gRNA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-VP64およびgRNA+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gRNA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gRNA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gRNA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、gRNA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巻、243～245頁(2013年)およびPerez-Pineraら著、Nature Methods誌、第10巻、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gRNA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、gRNA連結アプローチは、各gRNA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gRNA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賛助Biology of the Cell誌、第100巻、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-gRNA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巻、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。ライブラリーの各要素がd T 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ブラリーを設計する。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。一つは全てのライブ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」結合部位ライブ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P a t w a r d h a nら著、N a t u r e B i o t e c h n o l o g y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ブラリーを配列解析して24bpのd T o m a t o転写物タグとライブ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、d T 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ブラリー要素にわたって共有されているので対照TFは全てのライブ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、C a s 9 g R N A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ープより明らかである。図2 Cに示されるように、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。特に、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ばかりでなくN A G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g R N A標的配列の3'末端に近くに局在するときのみ、2塩基のミスマッチの導入によってC a s 9 g R N A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 T A L E は、C a s 9 - g R N A 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 T A L E 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。T A L E 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 T A L E によるターゲティングのランドスケー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 T A L E 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 H R 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 T A L E のターゲティングランドスケープを分析した。T A L E が短くなるほど（14-mer および 10-mer）次第にそれらの T A L E 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 T A L E 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 T A L E 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 T A L E 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 F o k 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 K i m ら著、P r o c e e d i n g s o f t h e N a t i o n a l A c a d e m y o f S c i e n c e s o f t h e U n i t e d S t a t e s o f A m e r i c a 誌、第 93 巻、1156 ~ 1160 頁（1996 年）および P a t t a n a y a k ら著、N a t u r e M e t h o d s 誌、第 8 巻、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 b p 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 C a s 9 結合配列または T A L E 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 C a s 9 / g R N A または T A L E 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 T A L E 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 m R N A 分子から作られた c D N A 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と一緒に配列解析し、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 2 B と 2 E、および図 9 A と図 10 B 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 8 B 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 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 8 B 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。

【 0 0 9 7 】

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 9 A 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 9 B 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 9 C 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 ~ 2 3 と標識されている）。図 9 D に示されるように、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 であることがヌクレアーゼ介在性 H R アッセイを用いて確認された。

【 0 0 9 8 】

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で（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、6 7 1 ~ 6 7 6 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。

【 0 0 9 9 】

図 3 A ~ 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 3 A 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 を読み枠から外し、且つ、下流の m C h e r r y 配列を読み枠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

10

20

30

40

50

た。相補ストランドにニックを入れる Cas9 D10A 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】

図 11A ~ B は Cas9 D10A 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断導入時の 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。200bp 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（Cas9 D10A 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 200bp 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。

【0101】

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同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同在複合体と呼ばれる。

【0102】

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】

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っ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）核酸および外来性の（または外来の）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と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または外来性DNAが含まれる。

10

【0104】

外来核酸（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、遺伝子転換、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。

20

【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。

30

【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

40

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・ウェブサイト・ツールキット、[tuebingen.mpg.de/hhpred](http://toolkit.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

50

各構造中の2つのMg配位性側鎖を特定した。それらは3M7KのD92とN113、および4H9DのD53とN77であった。これらの残基はCas9のD839とN863に対応した。Pad残基D92とN113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Cas9突然変異D839AとN863Aを作製した。さらに、HHPredはCas9とサーマス・サーモフィルスのRuvC(PDB ID: 4EP4)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Cas9中のRuvCドメインの機能を除去する以前に報告された突然変異D10A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。4EP4ではD7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Cas9 D10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Cas9のRuvC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

【0109】

[実施例2]

プラスミド構築

Quickchangeキット(アジレントテクノロジーズ社)を使用してCas9突然変異体を作製した。標的gRNA発現コンストラクトは(1)IDTから個々のgBlockとして直接注文され、pCR-BluntII-TOPO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)Genewiz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Gibson構築を用いてgRNAクローニングベクターに構築された(プラスミド番号41824)。Addgene社のEGFPレンチベクターに構築される、終止コドンを含むGFPP配列と適切な断片の融合PCR構築によって切断型GFPPを伴うHR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26777)。次にこれらのレンチベクターを使用してGFPPレポーター安定株を構築した。この研究で使用されたTALEN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照により援用されるSanjanaら著、Nature Protocols誌、第7巻、171~192頁(2012年)を参照されたい。標準的PCR融合プロトコルの手法を用いてCas9NとMS2 VP64の融合を実施した。OCT4およびREX1用のプロモーターシフターゼコンストラクトはAddgene社から入手した(プラスミド番号17221およびプラスミド番号17222)。

20

30

【0110】

[実施例3]

細胞培養および形質移入

10%ウシ胎児血清(FBS、インビトロジェン)、ペニシリン/ストレプトマイシン(pen/strep、インビトロジェン)、および非必須アミノ酸(NEAA、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(DMEM、インビトロジェン)中でHEK293T細胞を培養した。加湿恒温器中に37 °Cおよび5%のCO₂で細胞を維持した。

【0111】

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2000を使用して 0.4×10^6 個の細胞に2 µgのCas9プラスミド、2 µgのgRNAおよび/または2 µgのDNA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、FACS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DNAイーザー・キット(Qiagen)を使用して約 1×10^6 個の細胞のゲノムDNAを抽出した。これらのためにPCRを実施して前記細胞に由来するゲノムDNAを用いてターゲティング領域を増幅し、MiSeqパーソナルシーケンサー(Illumina社)により200,000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シングした。シーケンシングデータを分析してNHJ効率を推定した。

40

【0112】

50

転写活性化アッセイを伴う形質移入については、 0.4×10^6 個の細胞に (1) $2 \mu\text{g}$ の Cas9 N V P 64 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 Cas9 N プラスミド、 $2 \mu\text{g}$ の MS2 - V P 64、 $2 \mu\text{g}$ の gRNA - 2 x 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、その後これらの RNA を RT - PCR により分析した。ここで、OCT4 と RE X 1 についてインビトロジェンの標準的 taqman 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】

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、 0.4×10^6 個の細胞に (1) $2 \mu\text{g}$ の Cas9 N V P 64 プラスミド、 $2 \mu\text{g}$ の gRNA および $0.25 \mu\text{g}$ のレポーター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 $0.25 \mu\text{g}$ のレポーター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 $0.25 \mu\text{g}$ のレポーター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 (インビトロジェン) 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。

【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 8 A 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 8 A (第 1 段) および 8 B を参照されたい。

【0115】

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 8 A、第 3 段、左) とレポーター遺伝子の cDNA 配列 (図 8 A、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、novoalign バージョン 2.07.17 (ワールド・ワイド・ウェブサイト 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 8 A、第 3 段、左を参照のこと)。novoalign 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 8 A、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 Novoalign 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。

(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

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 *nov o a l i g n* によってA、C、GまたはTと呼ばなくてはならない。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 *nov o a l i g n* コールの可能性が生じなかったことに留意されたい。(i i i) 同様に、これらの領域には *nov 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

【 0 1 1 7 】

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・ウェットサイト *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79bpの共通セグメントにまで合併し、その後で *nov o a l i g n* (上記のバージョン) を使用してこれらの79bp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8A、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)24bp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8bp試料バーコード配列対によって挟まれる同一の63bp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

【 0 1 1 8 】

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8A、第4段、左)。A、C、およびGの塩基から構成される24bp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 $2 \cdot 8 e - 11$)、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く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

40

【 0 1 1 9 】

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（

【 0 1 2 0 】

【数 1】

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

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。

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、タグとの関連はこれらの 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。

【 0 1 2 1 】

[実施例 5]

C a s 9_N V P 6 4 を用いた RNA 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（アプタマー改変短鎖ガイド RNA）連結アプローチにより、各 s g R N A 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 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 DNA 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 ± 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 DNA 鎖内を標的とする s g R N A によって制御さ

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十分である塩基ミスマッチ対は「×」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、および N . S . (有意性無し) は $P \geq 0.05 / n$ であり、n は比較数である (表 2 を参照のこと) 。

【 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

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の評価した。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、および N . S . (有意性無し) は $P \geq 0.05 / n$ 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、スパーサーの 3 ' 末端に局在する。

【 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 中のスパーサーの 3 ' 末端の 1 2 b p 内に 1 塩基のミスマッチがあっ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。ヌ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 であるスパーサー配列内に標的に対する 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 中のスパーサーの 3 ' 末端の 1 2 b p 内に 1 塩基のミスマッチがあっても検出可能なターゲティングが生じるものの、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。

【 0 1 2 5 】

10

20

30

40

50

[実施例 9]

有効性確認、5' gRNA 短縮

図 16 A ~ 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。ヌ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 gRNA 1 (図 16 A ~ B) と gRNA 3 (図 16 C - D) を試験した。1 ~ 3 bp 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 \pm SEM (N = 3) である。

【 0 1 2 6 】

[実施例 10]

有効性確認、化膿レンサ球菌 (*S. pyogenes*) の PAM

図 17 A ~ 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 であることがヌクレアーゼ介在性 HR アッセイを用いて確認された。データは平均値 \pm 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 190 K の Cas9 標的であって、ターゲティング配列の最後の 13 nt を共有する代替性 NGG 標的を有しない Cas9 標的を、前の 13 nt 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。

【 0 1 2 7 】

[実施例 11]

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ヌクレアーゼ介在性 HR アッセイ (図 18 A ~ B) を用いて確認された。図 18 A ~ B に示されるように、ヌ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。

【 0 1 2 8 】

[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 19 A ~ 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 19 A : 図 2 に記載されるアプローチを改変したものを用い、一連の 6 NI リピートまたは 6 NH 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ープを分析した。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 19 B ~ 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もある。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 を使用するとヌクレアーゼ用途において FokI の二量体化が標的外効果のさらなる減少を可能にする。

【 0 1 2 9 】

[実施例 13]

オフセット・ニック形成、天然遺伝子座

図 20 A ~ 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

10

20

30

40

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DSB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AAVS1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20A：200bp長のDNA鎖を範囲とする8種類のgRNA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(s1~4)とアンチセンスストランドを標的とする4種類(as1~4)でAAVS1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。図20B：単一のgRNAでは検出可能なNHEJ事象がガイドされなかったが、オフセット・ニックを誘導してDSB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NHEJ事象を生じさせる。サンガーシーケンシングクロ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

【0130】

[実施例14]

オフセット・ニック形成、NHEJプロファイル

図21A~Cはオフセット・ニック形成およびNHEJ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gRNAの位置と共に示されている。さらに、相同組換え(HR)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NHEJ事象がもたらされた(図3B)。5'オーバーハングを形成するとNHEJの促進に加えてHRの堅固な誘導が観察された。3'オーバーハングの形成はHR率の改善を引き起こさなかった(図3C)。

20

【0131】

[実施例15]

表1

内在性遺伝子制御のためのgRNA標的

Cas9 gRNA介在性活性化実験において使用されたREX1、OCT4、SOX2およびNANOGのプロモーター中の標的を列記する。

30

【0132】

【表 1】

gRNA Name	gRNA Target
REX1 1	ctggcgggatcaactcgcgggtt agg
REX1 2	cctcgyccctccaaaagtgcct agg
REX1 3	acgctgattcctgcagatca ggg
REX1 4	ccagggaatacgtatccacca ggg
REX1 5	gcccaccccaagcgatcaaa tgg
REX1 6	aaataatacatctctaaggta agg
REX1 7	gctactggggaggctgagggc agg
REX1 8	tagcaatacagtcacattaa tgg
REX1 9	ctcatgtgatccccccgtct cgg
REX1 10	ccgggcagagagtgaaaggcg cgg
OCT4 1	ttccttccctctcccggtgct tgg
OCT4 2	tctctgcaaagccctggag agg
OCT4 3	aatgcagttgcggagtgcag tgg
OCT4 4	cctcagcctcctaaggctggc ggg
OCT4 5	gagtcacaaatccctcttact agg
OCT4 6	gagtgctctggatttgggata agg
OCT4 7	cagcacctcatctccagtg agg
OCT4 8	tctaaaaccaggggaatcat ggg
OCT4 9	cacaaggcagccagggatcc agg
OCT4 10	gatggcaagctgagaaacac tgg
OCT4 11	tgaaatgcacgcatacaatt agg
OCT4 12	ccagtcacagacctggccttc tgg
OCT4 13	cccagaaaaacagaccctga agg
OCT4 14	aagggttgagcacttgttta ggg
OCT4 15	atgtctgagttttggctgag agg
OCT4 16	ggtcccttgaaaggggagta ggg
OCT4 17	tggcagctctactcttgaaga tgg
OCT4 18	ggcacagtgccagaggtctg tgg
OCT4 19	taaaaataaaaaactaaca ggg
OCT4 20	tctgtggggacctgcaactg agg
OCT4 21	ggccagaggtcaaggctagt ggg
SOX2 1	cacgacccgaacccttctta cgg
SOX2 2	gttgaatgaagacagctctag tgg
SOX2 3	taagaacagagcaagttacg tgg
SOX2 4	tgtaaggttaagagagagag cgg
SOX2 5	tgacacacccaaactcctgcac tgg
SOX2 6	tttaaccacttccttcgaaa agg
SOX2 7	gtggctggcaggctggctct ggg
SOX2 8	ctcccccgccctcccccgcg cgg
SOX2 9	caaaacccggcagcgaggct ggg
SOX2 10	aggagccgcgcgcgctgat tgg
NANOG 1	cacacacaccacacagagat ggg
NANOG 2	gaagaagctaaagagccaga ggg
NANOG 3	atgagaatttcaataacctc agg
NANOG 4	tcccgctctgttgccaggc tgg
NANOG 5	cagacacccacccatgoy tgg
NANOG 6	tcccaatttactgggattac agg
NANOG 7	tgatttaaaagtgggaaacg tgg
NANOG 8	tctagttccccacctagtct ggg
NANOG 9	gattaactgagaattcacaa ggg
NANOG 10	cgcaggagggtgggtcta agg

10

20

30

40

【 0 1 3 3 】

[実施例 1 6]

表 2

C a s 9 g R N A および T A L E 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カ

50

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 $P \geq 0.05/n$ であり、ここでnは比較数である。表2（b）図2D中のシード領域の統計学的特徴解析：2つの突然変異が20bp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Cas9NVP64+gRNAが結合したときの発現値の間の分離度を表す \log_{10} （P値）。最大の $-\log_{10}$ （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8～9bp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」の節を参照されたい。

【0134】

【表 2】

Figure	Expression level comparison:		t-test	P-value	Symbol
	mutations vs.	mutations			
2b	0	1	1-samp	7.8E-05	**
	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-237	***
	6	7	2-samp	1.7E-43	***
	7	8	2-samp	3.7E-02	N.S.
2c	0	1	1-samp	8.9E-01	N.S.
	1	2	2-samp	1.9E-08	***
	2	3	2-samp	5.0E-147	***
	3	4	2-samp	0	***
	4	5	2-samp	0	***
	5	6	2-samp	4.2E-67	***
	6	7	2-samp	1.6E-03	*
	7	8	2-samp	4.7E-01	N.S.
57a	0	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-58	***
	4	5	2-samp	8.5E-101	***
	5	6	2-samp	6.8E-94	***
	6	7	2-samp	1.8E-61	***
	7	8	2-samp	5.1E-24	***
57d and 58d	0	1	1-samp	2.3E-18	***
	1	2	2-samp	2.4E-08	***
	2	3	2-samp	6.3E-04	***
	3	4	2-samp	4.0E-141	***
	4	5	2-samp	1.9E-20	***
	5	6	2-samp	1.2E-03	*
	6	7	2-samp	3.8E-05	***
	7	8	2-samp	9.4E-01	N.S.
58c	0	1	1-samp	7.2E-03	N.S.
	1	2	2-samp	5.0E-01	N.S.
	2	3	2-samp	3.9E-04	***
	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.
S13a (left)	0	1	1-samp	7.3E-01	N.S.
	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	0	***
S13a (middle)	0	1	1-samp	9.4E-02	N.S.
	1	2	2-samp	5.2E-09	***
	2	3	2-samp	7.9E-86	***
	3	4	2-samp	2.9E-53	***
	4	5	2-samp	3.5E-10	***
	5	6	2-samp	0	***
	6	7	2-samp	0	***
	7	8	2-samp	0	***
S13a (right)	0	1	1-samp	1.3E-13	***
	1	2	2-samp	1.1E-04	***
	2	3	2-samp	3.7E-08	***
	3	4	2-samp	0	***
	4	5	2-samp	0	***
	5	6	2-samp	0	***
	6	7	2-samp	0	***
	7	8	2-samp	0	***

【 0 1 3 5 】

【 実施例 1 7 】

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_N 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_{m4} VP64 融合タンパク質形式と C

10

20

30

40

50

a s 9_{m4} 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_{m4}^{VP64}

```
gccaccATGGACAAGAAGTACTCCATTGGGCTCGCTATCGGCACAAACAGCGTCGGCTGG
GCCGTCATTACGGACGAGTACAAGGTGCCGAGCAAAAAATTCAAAGTTCTGGGCAAT
ACCGATCGCCACAGCATAAAGAAGAACCTCATTGGCGCCCTCCTGTTCTGACTCCGGGG
AGACGGCCGAAGCCACGCGGCTCAAAAGAACAGCACGGCGCAGATATACCCGCAGAA
AGAATCGGATCTGCTACCTGCAGGAGATCTTTAGTAATGAGATGGCTAAGGTGGATGA
CTCTTTCTTCCATAGGCTGGAGGAGTCCTTTTTTGGTGGAGGAGGATAAAAAGCACGAG
CGCCACCCAATCTTTGGCAATATCGTGACGAGGTGGCGTACCATGAAAAGTACCCAA
CCATATATCATCTGAGGAAGAAGCTTGTAGACAGTACTGATAAGGCTGACTTGCGGTT
GATCTATCTCGCGCTGGCGCATATGATCAAATTTTCGGGGACACTTCCTCATCGAGGGG
GACCTGAACCCAGACAACAGCGATGTGCACAACTCTTTATCCAACGGTTTCAGACTT
ACAATCAGCTTTTTCGAAGAGAACCCGATCAACGCATCCGGAGTTGACGCCAAAGCAA
TCCTGAGCGCTAGGCTGTCCAAATCCCGGCGGCTCGAAAACCTCATCGCACAGCTCCC
TGGGGAGAAGAAGAACGGCCTGTTTGGTAATCTTATCGCCCTGTCACTCGGGCTGACC
CCCAACTTTAAATCTAACTTCGACCTGGCCGAAGATGCCAAGCTTCAACTGAGCAAAG
ACACCTACGATGATGATCTCGACAATCTGCTGGCCCAGATCGGCGACCAGTACGCAGA
CCTTTTTTTTGGCGGCAAAGAACCTGTCAGACGCCATTCTGCTGAGTGATATTCTGCGAG
TGAACACGGAGATCACCAAAGCTCCGCTGAGCGCTAGTATGATCAAGCGCTATGATG
AGCACCAACAAGACTTGACTTTGCTGAAGGCCCTTGTGACAGCAACTGCCTGAGAA
GTACAAGGAAATTTTCTTCGATCAGTCTAAAAATGGCTACGCCGGATACATTGACGGC
GGAGCAAGCCAGGAGGAATTTTACAAATTTATTAAGCCCATCTTGGAATAAATGGAC
GGCACCGAGGAGCTGCTGGTAAAGCTTAACAGAGAAGATCTGTTGCGCAAACAGCGC
ACTTTTCGACAATGGAAGCATCCCCACCAGATTACCTGGGCGAACTGCACGCTATCC
TCAGGCGGCAAGAGGATTTCTACCCCTTTTTTGAAAGATAACAGGGAAAAGATTGAGA
AAATCCTCACATTTTCGGATACCCTACTATGTAGGCCCCCTCGCCCGGGGAAATTCCAG
ATTCGCGTGGATGACTCGCAAATCAGAAGAGACCATCACTCCCTGGAACCTTCGAGGAA
GTCGTGGATAAGGGGGCCTCTGCCCAGTCCTTCATCGAAAGGATGACTAACTTTGATA
AAAATCTGCCTAACGAAAAGGTGCTTCCTAAACACTCTCTGCTGTACGAGTACTTCAC
AGTTTATAACGAGCTCACCAAGGTCAAATACGTCACAGAAGGGATGAGAAAGCCAGC
ATTCCTGTCTGGAGAGCAGAAGAAAGCTATCGTGGACCTCCTCTTCAAGACGAACCGG
AAAGTTACCGTGAAACAGCTCAAAGAAGACTATTTCAAAAAGATTGAATGTTTCGACT
CTGTTGAAATCAGCGGAGTGGAGGATCGCTTCAACGCATCCCTGGGAACGTATCACGA
TCTCCTGAAAATCATTAAGACAAGGACTTCCTGGACAATGAGGAGAACGAGGACAT
```

10

20

30

40

TCTTGAGGACATTGTCCTCACCTTACGTTGTTTGAAGATAGGGAGATGATTGAAGAA
CGCTTGAAAACCTTACGCTCATCTCTTCGACGACAAAGTCATGAAACAGCTCAAGAGGC
GCCGATATACAGGATGGGGGCGGCTGTCAAGAAAACCTGATCAATGGGATCCGAGACA
AGCAGAGTGGAAGACAATCCTGGATTTTCTTAAGTCCGATGGATTTGCCAACCGGAA
CTTCATGCAGTTGATCCATGATGACTCTCTCACCTTTAAGGAGGACATCCAGAAAGCA
CAAGTTTCTGGCCAGGGGGACAGTCTTCACGAGCACATCGCTAATCTTGCAGGTAGCC
CAGCTATCAAAAAGGGAATACTGCAGACCGTTAAGGTCGTGGATGAACTCGTCAAAG
TAATGGGAAGGCATAAGCCCCGAGAATATCGTTATCGAGATGGCCCCGAGAGAACCAAA
CTACCCAGAAGGGACAGAAGAACAGTAGGGAAAGGATGAAGAGGATTGAAGAGGGT
ATAAAAGAACTGGGGTCCCAAATCCTTAAGGAACACCCAGTTGAAAACACCCAGCTT
CAGAATGAGAAGCTCTACCTGTACTACCTGCAGAACGGCAGGGACATGTACGTGGAT
CAGGAACCTGGACATCAATCGGCTCTCCGACTACGACGTGGCTGCTATCGTGCCCCAGT
CTTTTCTCAAAGATGATTCTATTGATAATAAAGTGTTGACAAGATCCGATAAAgcTAGA
GGGAAGAGTGATAACGTCCCCTCAGAAGAAGTTGTCAAGAAAATGAAAAATTATTGG
CGGCAGCTGCTGAACGCCAAACTGATCACACAACGGAAGTTCGATAATCTGACTAAG
GCTGAACGAGGTGGCCTGTCTGAGTTGGATAAAGCCGGCTTCATCAAAAGGCAGCTTG
TTGAGACACGCCAGATCACCAAGCACGTGGCCCCAAATTCTCGATTACGCATGAACAC
CAAGTACGATGAAAATGACAAACTGATTGAGAGGTGAAAGTTATTACTCTGAAGTCT
AAGCTGGTCTCAGATTTTCAGAAAGGACTTTTCAGTTTTATAAGGTGAGAGAGATCAACA
ATTACCACCATGCGCATGATGCCTACCTGAATGCAGTGGTAGGCACTGCACTTATCAA
AAAATATCCCAAGCTTGAATCTGAATTTGTTTTACGGAGACTATAAAGTGTACGATGTT
AGGAAAATGATCGCAAAGTCTGAGCAGGAAATAGGCAAGGCCACCGCTAAGTACTTC
TTTTACAGCAATATTATGAATTTTTTCAAGACCGAGATTACACTGGCCAATGGAGAGA
TTCGGAAGCGACCACTTATCGAAACAAACGGAGAAACAGGAGAAATCGTGTGGGACA
AGGGTAGGGATTTTCGCGACAGTCCGGAAGGTCTGTCCATGCCGCAGGTGAACATCGT
TAAAAAGACCGAAGTACAGACCGGAGGCTTCTCCAAGGAAAGTATCCTCCCGAAAAG
GAACAGCGACAAGCTGATCGCACGCAAAAAGATTGGGACCCCAAGAAATACGGCGG
ATTCGATTCTCCTACAGTCGCTTACAGTGTACTGGTTGTGGCCAAAGTGGAGAAAGGG
AAGTCTAAAAAACTCAAAAGCGTCAAGGAACTGCTGGGCATCACAATCATGGAGCGA
TCAAGCTTCGAAAAAAACCCCATCGACTTTCTCGAGGCGAAAGGATATAAAGAGGTC
AAAAAAGACCTCATCATTAAGCTTCCCAAGTACTCTCTCTTTGAGCTTGAAAACGGCC
GGAAACGAATGCTCGCTAGTGCGGGCGAGCTGCAGAAAGGTAACGAGCTGGCACTGC
CCTCTAAATACGTTAATTTCTTGTATCTGGCCAGCCACTATGAAAAGCTCAAAGGGTCT
CCCGAAGATAATGAGCAGAAGCAGCTGTTTCGTGGAACAACACAAACTACCTTGAT
GAGATCATCGAGCAAATAAGCGAATTCTCCAAAAGAGTGATCCTCGCCGACGCTAAC
CTCGATAAGGTGCTTTCTGCTTACAATAAGCACAGGGATAAGCCCATCAGGGAGCAGG

10

20

30

40

CAGAAAACATTATCCACTTGTCTTACTCTGACCAACTTGGGCGCGCCTGCAGCCTTCAA
 GTACTTCGACACCACCATAGACAGAAAGCGGTACACCTCTACAAAGGAGGTCCTGGA
 CGCCACACTGATTTCATCAGTCAATTACGGGGCTCTATGAAACAAGAATCGACCTCTCT
 CAGCTCGGTGGAGACAGCAGGGGCTGACCCCAAGAAGAAGAGGAAGGTGGAGGCCAG
 CGGTTCCGGACGGGCTGACGCATTGGACGATTTTGATCTGGATATGCTGGGAAGTGAC
 GCCCTCGATGATTTTGACCTTGACATGCTTGGTTCCGGATGCCCTTGATGACTTTGACCT
 CGACATGCTCGGCAGTGACGCCCTTGATGATTTTCGACCTGGACATGCTGATTAACCTCT
 AGATGA

10

【 0 1 3 7 】

>Cas9_{m4}^{VP64}N Sequences

gccaccATGCCCAAGAAGAAGAGGAAGGTGGGAAGGGGGATGGACAAGAAGTACTCCA
 TTGGGCTCGCTATCGGCACAAACAGCGTCGGCTGGGCCGTCATTACGGACGAGTACAA
 GGTGCCGAGCAAAAAATTCAAAGTTCTGGGCAATACCGATCGCCACAGCATAAAGAA
 GAACCTCATTGGCGCCCTCCTGTTCTGACTCCGGGGAGACGGCCGAAGCCACGCGGCTC
 AAAAGAACAGCACGGCGCAGATATACCCGCAGAAAGAATCGGATCTGCTACCTGCAG
 GAGATCTTTAGTAATGAGATGGCTAAGGTGGATGACTCTTTCTTCCATAGGCTGGAGG
 AGTCCTTTTTTGGTGGAGGAGGATAAAAAGCACGAGCGCCACCCAATCTTTGGCAATAT
 CGTGGACGAGGTGGCGTACCATGAAAAGTACCCAACCATATATCATCTGAGGAAGAA
 GCTTGTAGACAGTACTGATAAGGCTGACTTGCGGTTGATCTATCTCGCGCTGGCGCAT
 ATGATCAAATTTTCGGGGACACTTCCTCATCGAGGGGGACCTGAACCCAGACAACAGC
 GATGTGCACAAACTCTTTATCCAAGTGGTTTCACTTACAATCAGCTTTTTCGAAGAGA
 ACCCGATCAACGCATCCGGAGTTGACGCCAAAGCAATCCTGAGCGCTAGGCTGTCCAA
 ATCCCGGCGGCTCGAAAACCTCATCGCACAGCTCCCTGGGGAGAAGAAGAACGGCCT
 GTTTGGTAATCTTATCGCCCTGTCACTCGGGCTGACCCCAACTTTAAATCTAACTTCG
 ACCTGGCCGAAGATGCCAAGCTTCAACTGAGCAAAGACACCTACGATGATGATCTCG
 ACAATCTGCTGGCCCAGATCGGCGACCAAGTACGCAGACCTTTTTTTGGCGGCAAAGAA
 CCTGTCAGACGCCATTCTGCTGAGTGATATTCTGCGAGTGAACACGGAGATCACCAAA
 GCTCCGCTGAGCGCTAGTATGATCAAGCGCTATGATGAGCACCACCAAGACTTGACTT
 TGCTGAAGGCCCTTGTCAGACAGCAACTGCCTGAGAAGTACAAGGAAATTTTCTTCGA
 TCAGTCTAAAAATGGCTACGCCGGATACATTGACGGCGGAGCAAGCCAGGAGGAATT
 TTACAAATTTATTAAGCCCATCTTGGAATAAATGGACGGCACCGAGGAGCTGCTGGTA
 AAGCTTAACAGAGAAGATCTGTTGCGCAAACAGCGCACTTTCGACAATGGAAGCATC
 CCCCACCAGATTCACCTGGGCGAACTGCACGCTATCCTCAGGCGGCAAGAGGATTTCT
 ACCCCTTTTTTGAAAGATAACAGGGGAAAAGATTGAGAAAATCCTCACATTTTCGGATACC

20

30

40

CTACTATGTAGGCCCCCTCGCCCGGGGAAATTCCAGATTTCGCGTGGATGACTCGCAAA
TCAGAAGAGACCATCACTCCCTGGAACCTTCGAGGAAGTCGTGGATAAGGGGGCCTCT
GCCCAGTCCTTCATCGAAAGGATGACTAACTTTGATAAAAATCTGCCTAACGAAAAGG
TGCTTCCTAAACACTCTCTGCTGTACGAGTACTTCACAGTTTATAACGAGCTCACCAAG
GTCAAATACGTCACAGAAGGGATGAGAAAGCCAGCATTCTGTCTGGAGAGCAGAAG
AAAGCTATCGTGGACCTCCTCTTCAAGACGAACCGGAAAGTTACCGTGAAACAGCTCA
AAGAAGACTATTTCAAAAAGATTGAATGTTTCGACTCTGTTGAAATCAGCGGAGTGGA
GGATCGCTTCAACGCATCCCTGGGAACGTATCACGATCTCCTGAAAATCATTAAAGAC
AAGGACTTCCTGGACAATGAGGAGAACGAGGACATTCTTGAGGACATTGTCCTCACCC
TTACGTTGTTTGAAGATAGGGAGATGATTGAAGAACGCTTGAAAACCTTACGCTCATCT
CTTCGACGACAAAGTCATGAAACAGCTCAAGAGGCGCCGATATACAGGATGGGGGGCG
GCTGTCAAGAAAACCTGATCAATGGGATCCGAGACAAGCAGAGTGGAAGACAATCCT
GGATTTTCTTAAGTCCGATGGATTTGCCAACCGGAACTTCATGCAGTTGATCCATGATG
ACTCTCTCACCTTTAAGGAGGACATCCAGAAAGCACAAGTTTCTGGCCAGGGGGACAG
TCTTCACGAGCACATCGCTAATCTTGCAGGTAGCCCAGCTATCAAAAAGGGAATACTG
CAGACCGTTAAGGTCGTGGATGAACTCGTCAAAGTAATGGGAAGGCATAAGCCCGAG
AATATCGTTATCGAGATGGCCCGAGAGAACCAAACCTACCCAGAAGGGACAGAAGAAC
AGTAGGGAAAGGATGAAGAGGATTGAAGAGGGTATAAAAGAACTGGGGTCCCAAAT
CCTTAAGGAACACCCAGTTGAAAACACCCAGCTTCAGAATGAGAAGCTCTACCTGTAC
TACCTGCAGAACGGCAGGGACATGTACGTGGATCAGGAACTGGACATCAATCGGCTC
TCCGACTACGACGTGGCTGCTATCGTGCCCCAGTCTTTTCTCAAAGATGATTCTATTGA
TAATAAAGTGTTGACAAGATCCGATAAAgcTAGAGGGAAGAGTGATAACGTCCCCTCA
GAAGAAGTTGTCAAGAAAATGAAAAATTATTGGCGGCAGCTGCTGAACGCCAAACTG
ATCACACAACGGAAGTTCGATAATCTGACTAAGGCTGAACGAGGTGGCCTGTCTGAGT
TGGATAAAGCCGGCTTCATCAAAAGGCAGCTTGTTGAGACACGCCAGATCACCAAGC
ACGTGGCCCAAATTCTCGATTACGCATGAACACCAAGTACGATGAAAATGACAAACT
GATTCGAGAGGTGAAAGTTATTACTCTGAAGTCTAAGCTGGTCTCAGATTTTCAGAAAG
GACTTTTCAGTTTTTATAAGGTGAGAGAGATCAACAATTACCACCATGCGCATGATGCCT
ACCTGAATGCAGTGGTAGGCACTGCACCTTATCAAAAAATATCCCAAGCTTGAATCTGA
ATTTGTTTACGGAGACTATAAAGTGTACGATGTTAGGAAAATGATCGCAAAGTCTGAG
CAGGAAATAGGCAAGGCCACCGCTAAGTACTTCTTTTACAGCAATATTATGAATTTTT
TCAAGACCGAGATTACACTGGCCAATGGAGAGATTTCGGAAGCGACCACTTATCGAAA
CAAACGGAGAAAACAGGAGAAATCGTGTGGGACAAGGGTAGGGATTTTCGCGACAGTCC
GGAAGGTCCTGTCCATGCCGCAGGTGAACATCGTTAAAAAGACCGAAGTACAGACCG
GAGGCTTCTCCAAGGAAAGTATCCTCCCGAAAAGGAACAGCGACAAGCTGATCGCAC
GCAAAAAAGATTGGGACCCCAAGAAATACGGCGGATTTCGATTCTCCTACAGTCGCTTA

10

20

30

40

CAGTGTACTGGTTGTGGCCAAAGTGGAGAAAGGGAAGTCTAAAAAACTCAAAAGCGT
CAAGGAACTGCTGGGCATCACAATCATGGAGCGATCAAGCTTCGAAAAAAACCCCAT
CGACTTTCTCGAGGCGAAAGGATATAAAGAGGTCAAAAAAGACCTCATCATTAAGCTT
CCCAAGTACTCTCTCTTTGAGCTTGAAAACGGCCGGAACGAATGCTCGCTAGTGCGG
GCGAGCTGCAGAAAGGTAACGAGCTGGCACTGCCCTCTAAATACGTTAATTTCTTGTA
TCTGGCCAGCCACTATGAAAAGCTCAAAGGGTCTCCCGAAGATAATGAGCAGAAGCA
GCTGTTTCGTGGAACAACACAAACACTACCTTGATGAGATCATCGAGCAAATAAGCGA
ATTCTCCAAAAGAGTGATCCTCGCCGACGCTAACCTCGATAAGGTGCTTTCTGCTTAC
AATAAGCACAGGGATAAGCCCATCAGGGAGCAGGCAGAAAACATTATCCACTTGTTT
ACTCTGACCAACTTGGGCGCGCCTGCAGCCTTCAAGTACTTCGACACCACCATAGACA
GAAAGCGGTACACCTCTACAAAGGAGGTCCTGGACGCCACACTGATTCATCAGTCAAT
TACGGGGCTCTATGAAACAAGAATCGACCTCTCTCAGCTCGGTGGAGACAGCAGGGCT
GACCCCAAGAAGAAGAGGGAAGGTGGAGGCCAGCGGTTCCGGACGGGCTGACGCATTG
GACGATTTTGATCTGGATATGCTGGGAAGTGACGCCCTCGATGATTTTGACCTTGACA
TGCTTGGTTTCGGATGCCCTTGATGACTTTGACCTCGACATGCTCGGCAGTGACGCCCTT
GATGATTTGACCTGGACATGCTGATTAACCTCTAGATGA

10

20

【 0 1 3 8 】

>Cas9_{m4}^{VP64}C

gccaccATGGACAAGAAGTACTCCATTGGGCTCGCTATCGGCACAAACAGCGTCGGCTGG
GCCGTCATTACGGACGAGTACAAGGTGCCGAGCAAAAAATTCAAAGTTCTGGGCAAT
ACCGATCGCCACAGCATAAAGAAGAACCTCATTGGCGCCCTCCTGTTGACTCCGGGG
AGACGGCCGAAGCCACGCGGCTCAAAAGAACAGCACGGCGCAGATATACCCGCAGAA
AGAATCGGATCTGCTACCTGCAGGAGATCTTTAGTAATGAGATGGCTAAGGTGGATGA
CTCTTTCTTCCATAGGCTGGAGGAGTCCTTTTTGGTGGAGGAGGATAAAAAGCACGAG
CGCCACCCAATCTTTGGCAATATCGTGGACGAGGTGGCGTACCATGAAAAGTACCCAA
CCATATATCATCTGAGGAAGAAGCTTGTAGACAGTACTGATAAGGCTGACTTGCGGTT
GATCTATCTCGCGCTGGCGCATATGATCAAATTTGCGGGACACTTCCTCATCGAGGGG
GACCTGAACCCAGACAACAGCGATGTCGACAAACTCTTTATCCAACCTGGTTCAGACTT
ACAATCAGCTTTTCGAAGAGAACCCGATCAACGCATCCGGAGTTGACGCCAAAGCAA
TCCTGAGCGCTAGGCTGTCAAATCCCGGCGGCTCGAAAACCTCATCGCACAGCTCCC
TGGGGAGAAGAAGAACGGCCTGTTTGGTAATCTTATCGCCCTGTCACTCGGGCTGACC
CCCAACTTTAAATCTAACTTCGACCTGGCCGAAGATGCCAAGCTTCAACTGAGCAAAG
ACACCTACGATGATGATCTCGACAATCTGCTGGCCCAGATCGGCGACCAGTACGCAGA
CCTTTTTTTGGCGGCAAAGAACCTGTCAGACGCCATTCTGCTGAGTGATATTCTGCGAG

30

40

TGAACACGGAGATCACCAAAGCTCCGCTGAGCGCTAGTATGATCAAGCGCTATGATG
AGCACCACCAAGACTTGACTTTGCTGAAGGCCCTTGTCAGACAGCAACTGCCTGAGAA
GTACAAGGAAATTTTCTTCGATCAGTCTAAAAATGGCTACGCCGGATACATTGACGGC
GGAGCAAGCCAGGAGGAATTTTACAAATTTATTAAGCCCATCTTGGA AAAAATGGAC
GGCACCAGGAGCTGCTGGTAAAGCTTAAACAGAGAAGATCTGTTGCGCAAACAGCGC
ACTTTGACAATGGAAGCATCCCCACCAGATTCACCTGGGCGAACTGCACGCTATCC
TCAGGCGGCAAGAGGATTTCTACCCCTTTTTGAAAGATAACAGGGAAAAGATTGAGA
AAATCCTCACATTTTCGGATACCCTACTATGTAGGCCCCCTCGCCCGGGGAAATTCCAG
ATTCGCGTGGATGACTCGCAAATCAGAAGAGACCATCACTCCCTGGAACCTTCGAGGAA
GTCGTGGATAAGGGGGCCTCTGCCCAGTCCTTCATCGAAAGGATGACTAACTTTGATA
AAAATCTGCCTAACGAAAAGGTGCTTCTAAACACTCTCTGCTGTACGAGTACTTCAC
AGTTTATAACGAGCTCACCAAGGTCAAATACGTCACAGAAGGGATGAGAAAGCCAGC
ATTCCTGTCTGGAGAGCAGAAGAAAGCTATCGTGGACCTCCTCTTCAAGACGAACCGG
AAAGTTACCGTGAAACAGCTCAAAGAAGACTATTTCAAAAAGATTGAATGTTTCGACT
CTGTTGAAATCAGCGGAGTGGAGGATCGCTTCAACGCATCCCTGGGAACGTATCACGA
TCTCCTGAAAATCATTAAAGACAAGGACTTCCTGGACAATGAGGAGAACGAGGACAT
TCTTGAGGACATTGTCCTCACCTTACGTTGTTTGAAGATAGGGAGATGATTGAAGAA
CGCTTGAAAACCTACGCTCATCTCTTCGACGACAAAGTCATGAAACAGCTCAAGAGGC
GCCGATATACAGGATGGGGGCGGCTGTCAAGAAAACCTGATCAATGGGATCCGAGACA
AGCAGAGTGGAAGACAATCCTGGATTTTCTTAAGTCCGATGGATTTGCCAACCGGAA
CTTCATGCAGTTGATCCATGATGACTCTCTCACCTTTAAGGAGGACATCCAGAAAGCA
CAAGTTTCTGGCCAGGGGGACAGTCTTCACGAGCACATCGCTAATCTTGCAGGTAGCC
CAGCTATCAAAAAGGGAATACTGCAGACCGTTAAGGTCGTGGATGAACTCGTCAAAG
TAATGGGAAGGCATAAGCCCGAGAATATCGTTATCGAGATGGCCCGAGAGAACCAAA
CTACCCAGAAGGGACAGAAGAACAGTAGGGAAAGGATGAAGAGGATTGAAGAGGGT
ATAAAAGAACTGGGGTCCCAAATCCTTAAGGAACACCCAGTTGAAAACACCCAGCTT
CAGAATGAGAAGCTCTACCTGTACTACCTGCAGAACGGCAGGGACATGTACGTGGAT
CAGGAACCTGGACATCAATCGGCTCTCCGACTACGACGTGGCTGCTATCGTGCCCCAGT
CTTTTCTCAAAGATGATTCTATTGATAATAAAGTGTTGACAAGATCCGATAAAgcTAGA
GGGAAGAGTGATAACGTCCCCCTCAGAAGAAGTTGTCAAGAAAATGAAAAATTATTGG
CGGCAGCTGCTGAACGCCAACTGATCACACAACGGAAGTTCGATAATCTGACTAAG
GCTGAACGAGGTGGCCTGTCTGAGTTGGATAAAGCCGGCTTCATCAAAAGGCAGCTTG
TTGAGACACGCCAGATCACCAAGCACGTGGCCCAAATTCTCGATTACGCATGAACAC
CAAGTACGATGAAAATGACAACTGATTCGAGAGGTGAAAGTTATTACTCTGAAGTCT
AAGCTGGTCTCAGATTTTCAGAAAGGACTTTCAGTTTTATAAGGTGAGAGAGATCAACA
ATTACCACCATGCGCATGATGCCTACCTGAATGCAGTGGTAGGCACTGCACTTATCAA

10

20

30

40

AAAATATCCCAAGCTTGAATCTGAATTTGTTTACGGAGACTATAAAGTGTACGATGTT
 AGGAAAATGATCGCAAAGTCTGAGCAGGAAATAGGCAAGGCCACCGCTAAGTACTTC
 TTTTACAGCAATATTATGAATTTTTTTCAAGACCGAGATTACACTGGCCAATGGAGAGA
 TTCGGAAGCGACCACTTATCGAAACAAACGGAGAAACAGGAGAAATCGTGTGGGACA
 AGGGTAGGGATTTTCGCGACAGTCCGGAAGGTCTGTCCATGCCGCAGGTGAACATCGT
 TAAAAAGACCGAAGTACAGACCGGAGGCTTCTCCAAGGAAAGTATCCTCCCGAAAAG
 GAACAGCGACAAGCTGATCGCACGCAAAAAAGATTGGGACCCCAAGAAATACGGCGG
 ATTCGATTCTCCTACAGTCGCTTACAGTGTACTGGTTGTGGCCAAAGTGGAGAAAGGG
 AAGTCTAAAAAACTCAAAAGCGTCAAGGAAGTGTGGGCATCACAATCATGGAGCGA
 TCAAGCTTCGAAAAAAACCCCATCGACTTTCTCGAGGCGAAAGGATATAAAGAGGTC
 AAAAAAGACCTCATCATTAAGCTTCCCAAGTACTCTCTCTTTGAGCTTGAAAACGGCC
 GGAAACGAATGCTCGCTAGTGCGGGCGAGCTGCAGAAAGGTAACGAGCTGGCACTGC
 CCTCTAAATACGTTAATTTCTTGTATCTGGCCAGCCACTATGAAAAGCTCAAAGGGTCT
 CCCGAAGATAATGAGCAGAAGCAGCTGTTCGTGGAACAACACAAACACTACCTTGAT
 GAGATCATCGAGCAAATAAGCGAATTCTCCAAAAGAGTGATCCTCGCCGACGCTAAC
 CTCGATAAGGTGCTTTCTGCTTACAATAAGCACAGGGATAAGCCCATCAGGGAGCAGG
 CAGAAAACATTATCCACTTGTTTACTCTGACCAACTTGGGCGCGCCTGCAGCCTTCAA
 GTACTTCGACACCACCATAGACAGAAAGCGGTACACCTCTACAAAGGAGGTCCTGGA
 CGCCACACTGATTCATCAGTCAATTACGGGGCTCTATGAAACAAGAATCGACCTCTCT
 CAGCTCGGTGGAGACAGCAGGGCTGACCCCAAGAAGAAGAGGAAGGTGGAGGCCAG
 CGGTTCCGGACGGGCTGACGCGATTGGACGATTTTGATCTGGATATGCTGGGAAGTGAC
 GCCCTCGATGATTTTGACCTTGACATGCTTGGTTCGGATGCCCTTGATGACTTTGACCT
 CGACATGCTCGGCAGTGACGCCCTTGATGATTTTCGACCTGGACATGCTGATTAAGTCT
 AGAGCGGCCGCGAGATCCAAAAAAGAAGAGAAAGGTAGATCCAAAAAAGAAGAGAAA
 GGTAGATCCAAAAAAGAAGAGAAAGGTAGATACGGCCGCATAG

10

20

30

【 0 1 3 9 】

B . MS2 活性化因子コンストラクトと 2 × MS2 アプ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_{VP64}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。

【 0 1 4 0 】

>MS2_{VP64}N

gccaccATGGGACCTAAGAAAAAGAGGAAGGTGCGGCCGCTTCTAGAATGGCTTCTAA
 CTTTACTCAGTTCGTTCTCGTCGACAATGGCGGAAGTGGCGACGTGACTGTCGCCCCA
 AGCAACTTCGCTAACGGGATCGCTGAATGGATCAGCTCTAACTCGCGTTCACAGGCTT

40

ACAAAGTAACCTGTAGCGTTCGTCAGAGCTCTGCGCAGAATCGCAAATACACCATCAA
 AGTCGAGGTGCCTAAAGGCGCCTGGCGTTCGTAATAATATGGAACCTAACCATTTCCA
 ATTTTCGCCACGAATTCCGACTGCGAGCTTATTGTTAAGGCAATGCAAGGTCTCCTAA
 AAGATGGAAACCCGATTCCCTCAGCAATCGCAGCAAACCTCCGGCATCTACGAGGCCA
 GCGGTTCCGGACGGGCTGACGCATTGGACGATTTTGATCTGGATATGCTGGGAAGTGA
 CGCCCTCGATGATTTTGACCTTGACATGCTTGGTTCCGGATGCCCTTGATGACTTTGACC
 TCGACATGCTCGGCACTGACGCCCTTGATGATTTTCGACCTGGACATGCTGATTAACCTCT
 AGATGA

10

【 0 1 4 1 】

>MS2_{VP64C}

gccaccATGGGACCTAAGAAAAAGAGGAAGGTGGCGGCCGCTTCTAGAATGGCTTCTAA
 CTTTACTCAGTTCGTTCTCGTCGACAATGGCGGAACCTGGCGACGTGACTGTGCCCCA
 AGCAACTTCGCTAACGGGATCGCTGAATGGATCAGCTCTAACTCGCGTTCACAGGCTT
 ACAAAGTAACCTGTAGCGTTCGTCAGAGCTCTGCGCAGAATCGCAAATACACCATCAA
 AGTCGAGGTGCCTAAAGGCGCCTGGCGTTCGTAATAATATGGAACCTAACCATTTCCA
 ATTTTCGCCACGAATTCCGACTGCGAGCTTATTGTTAAGGCAATGCAAGGTCTCCTAA
 AAGATGGAAACCCGATTCCCTCAGCAATCGCAGCAAACCTCCGGCATCTACGAGGCCA
 GCGGTTCCGGACGGGCTGACGCATTGGACGATTTTGATCTGGATATGCTGGGAAGTGA
 CGCCCTCGATGATTTTGACCTTGACATGCTTGGTTCCGGATGCCCTTGATGACTTTGACC
 TCGACATGCTCGGCACTGACGCCCTTGATGATTTTCGACCTGGACATGCTGATTAACCTCT
 AGACCGCGCCCGCAGATCCAAAAAAGAGAGAAAGGTAGATCCAAAAAAGAGAGAAA
 GGTAGATCCAAAAAAGAGAGAAAGGTAGATACCGCCCGCATAG

20

【 0 1 4 2 】

>gRNA_{2XMS2}

TGTACAAAAAAGCAGGCTTTAAAGGAACCAATTCAGTCGACTGGATCCGGTACCAAG
 GTCGGGCAGGAAGAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACA
 AGGCTGTTAGAGAGATAATTAGAATTAATTTGACTGTAAACACAAAGATATTAGTACA
 AAATACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTAAAATTATGTT
 TAAAATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTGATTTCTTGGCTTTA
 TATATCTTGTGGAAAGGACGAAACACCGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
 CTAGAAATAGCAAGTTAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACC
 GAGTCGGTGCTCTGCAGGTCGACTCTAGAAAACATGAGGATCACCCATGCTCTGCAGTA
 TTCCCGGGTTCATTAGATCCTAAGGTACCTAATTGCCTAGAAAACATGAGGATCACCC
 ATGCTCTGCAGGTCGACTCTAGAAATTTTTTCTAGAC

40

【 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 標の配列、 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 】

>TF Reporter 1

TAGGGATAACAGGGTAATAGTGTCCCCTCCACCCCACAGTGGGGCGAGGTAAGGCGTG
 TACGGTGGGAGGCCTATATAAGCAGAGCTCGTTTAAGTGAACCGTCAGATCGCCTGGAG
 AATTCgccaccaigGACTACAAGGATGACGACGATAAACTTCCGGTGGCGGACTGGGTTT
 CACCGTGAGCAAGGGCGAGGAGGTCATCAAAGAGTTTCATGCGCTTCAAGGTGCGCAT
 GGAGGGCTCCATGAACGGGCCACGAGTTTCGAGATCGAAGGGCGAGGGCGAAGGGCCGCCC
 CTACGAGGGGCACCCAGACCGCCAAGCTGAAGGTGACCAAGGGCGGGCCCCCTGCCCTT
 CGCCTGGGACATCCTGTCCCCCAGTTTCATGTACGGCTCCAAAGGCGTACGTGAAGCAC
 CCGCGCGACATCCCCGATTACAAGAAGCTGTCTTCCCCGAGGGCTTCAAGTGGGAGC
 GCGTGATGAACTTCGAGGACGGCGGTCTGGTGACCGTGACCCAGGACTCCTCCCTGCA
 GGACCGCACGCTGATCTACAAGGTGAAGATGCGCGGCACCAACTTCCCCCCCCGACGG
 CCCCCTAATGCAGAAAGAACCATGGGCTGGGAGGGCTCCACCGAGCGCCTGTACCC
 CCGCGACGGCGTGCTGAAGGGCGAGATCCACCAGGGCCCTGAAGCTGAAGGACGGCGG
 CCACTACCTGGTGGAGTTCAAGACCATCTACATGGCCAAGAAAGCCCGTGCAACTGCCC
 GGCTACTACTACGTGGACACCAAGCTGGACATCACCTCCCACAACGAGGACTACACCA
 TCGTGGAACAAGTACGAGCGCTCCGAGGGGCCGCCACCACTGTTCCTGTACGGCATGGA
 CGAGCTGTACAAGTAA

10

20

【 0 1 4 5 】

>TF Reporter 2

TAGGGATAACAGGGTAATAGTGGGGGCCACTAGGGACAGGATTGGCGAGGTAAGGCGTG
 TACGGTGGGAGGCCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAG
 AATTCgccaccatgGACTACAAGGATGACGACGATAAACTTCCGGTGGCGGACTGGGTTT
 CACCGTGAGCAAGGGCGAGGAGGTCATCAAAGAGTTTCATGCGCTTCAAGGTGCGCAT
 GGAGGGCTCCATGAACGGGCCACGAGTTTCGAGATCGAAGGGCGAGGGCGAGGGCCGCCC
 CTACGAGGGGCACCCAGACCGCCAAGCTGAAGGTGACCAAGGGCGGGCCCCCTGCCCTT
 CGCCTGGGACATCCTGTCCCCCAGTTTCATGTACGGCTCCAAAGGCGTACGTGAAGCAC
 CCGCGCGACATCCCCGATTACAAGAAGCTGTCTTCCCCGAGGGCTTCAAGTGGGAGC
 GCGTGATGAACTTCGAGGACGGCGGTCTGGTGACCGTGACCCAGGACTCCTCCCTGCA
 GGACGGCACGCTGATCTACAAGGTGAAGATGCGCGGCACCAACTTCCCCCCCCGACGG
 CCCCCTAATGCAGAAAGAACCATGGGCTGGGAGGGCTCCACCGAGCGCCTGTACCC
 CCGCGACGGCGTGCTGAAGGGCGAGATCCACCAGGGCCCTGAAGCTGAAGGACGGCGG
 CCACTACCTGGTGGAGTTCAAGACCATCTACATGGCCAAGAAAGCCCGTGCAACTGCCC
 GGCTACTACTACGTGGACACCAAGCTGGACATCACCTCCCACAACGAGGACTACACCA
 TCGTGGAACAAGTACGAGCGCTCCGAGGGGCCGCCACCACTGTTCCTGTACGGCATGGA
 CGAGCTGTACAAGTAA

30

40

【 0 1 4 6 】

D. TALEおよびCas9-gRNAの特異性アッセイに使用されるレポーターライブラリーの全般的形式が下に提供されている(Iscel对照TF標的配列、gRNA/TALE標的部位配列(gRNAについては23bp、およびTALE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dTomato配列が強調されている)。

【0147】

> Specificity Reporter Libraries

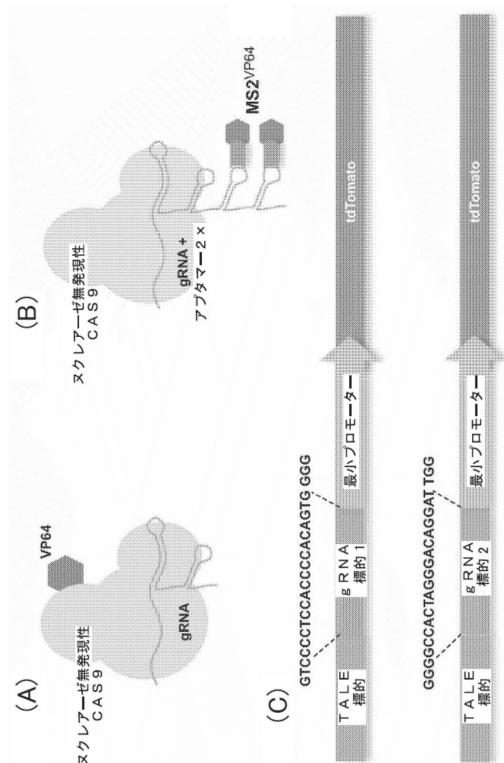
```

TAGGGATAACAGGCTAATAGTNNNNNNNNNNNNNNNNNNNNNNNNNNNNCGAGGTAGGCGT
GTACGGTGGGAGGCCCTATATAAGCAGAGCTCGTTTGTGTAACCGTCAGATCGCCTGGA
GAATTCgccaccatgGACTACAAGGATGACGACGATAAANNNNNNNNNNNNNNNNNNNNNNNN
NNNNACTTCCGGTGGCGGACTGGGTTCCACCGTGGAGCAAGGGCGAGGAGGTCATCAA
AGAGTTCATGCGCTTCAAGGTGCGCATGGAAGGGCTCCATGAACGGCCACGAGTTCGA
GATCGAGGGGCGAGGGCCGAGGGGCCGCCCTACGAGGGCACCCAGACCGCCAAGCTGAA
GGTGACCAAGGGGCGGCCCTTCCCTTCCGCTGGGACATCCTGTCCCCCAATTTCATG
TACGGCTCCAAGGCGTACGTGAAGCACCCCGCCGACATCCCCGATTACAAGAAGCTGT
CCTTCCCCGAGGGCTTCAAGTGGGAGCGCGTGTATGAACCTTCGAGGACGGCGGCTCTGGT
GACCGTGACCCAGGACTCCTCCCTGCAAGGACGGCACGCTGATCTACAAGGTGAAGAT
GCGCGGCACCAACTTCCCCCCCCGACGGCCCCCGTAATGCAGAAGAAGACCATGGGCTG
GGAGGCCCTCCACCGAGCGCTGTACCCCCCGGACGGCGTGTCTGAAGGGCGAGATCCA
CCAGGCCCTGAAGCTGAAGGACGCGCGGCCACTACCTGTTGGAGTTCAAGACCATCTA
CATGCCCCAAGAAGCCCCGTGCAACTGCCCCGCTACTACTACGTGGACACCAAGCTGGA
CATCACCTCCCACAACGAGGACTACACCATCGTGGAAACAGTACGAGCGCTCCGAAGG
CCGCCACCACCTGTTTCTGTACGGCATGGACGAGCTGTACAAGTAAGAATTC
  
```

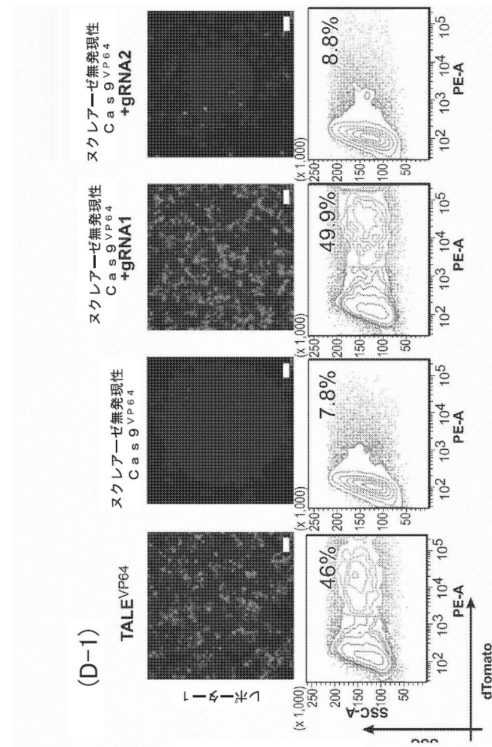
10

20

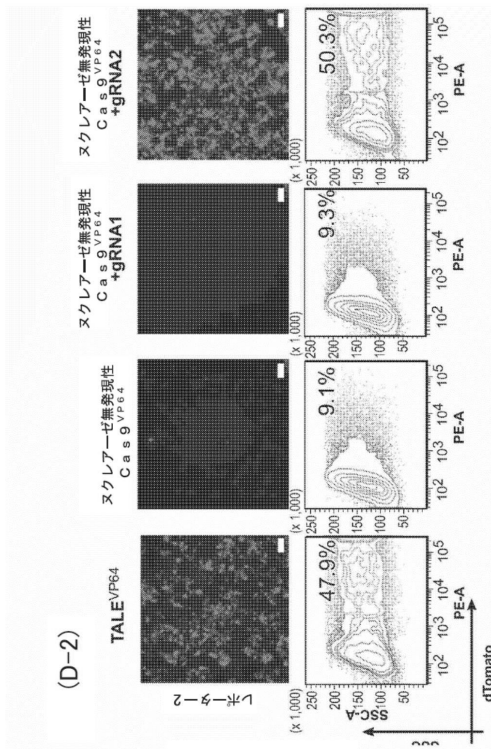
【図1-1】



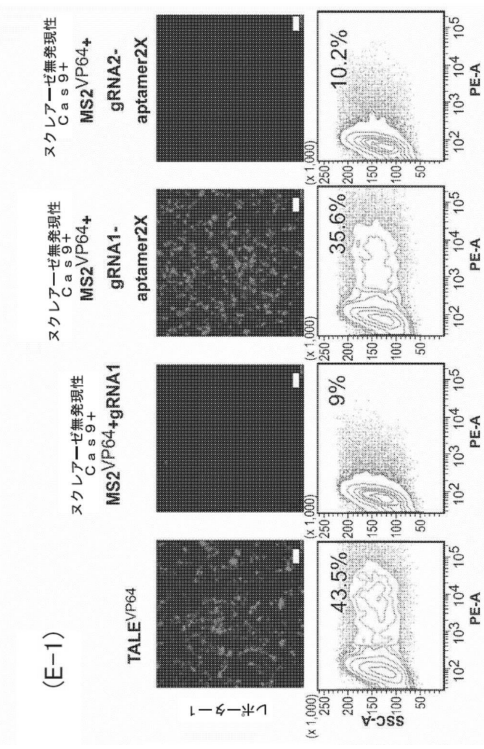
【図1-2】



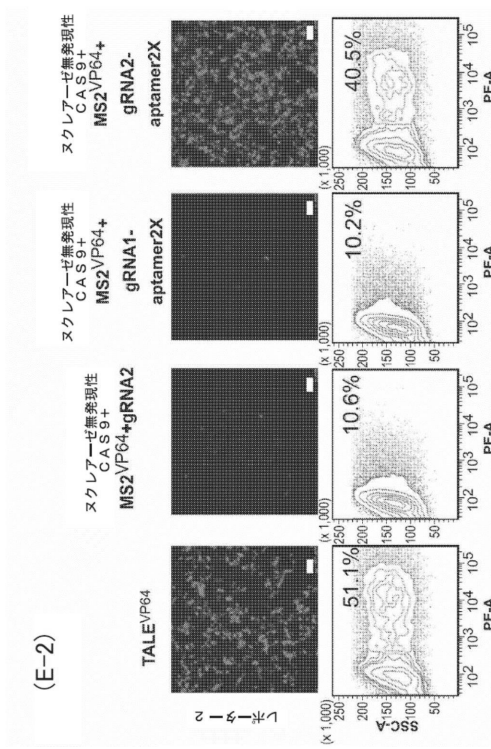
【図 1 - 3】



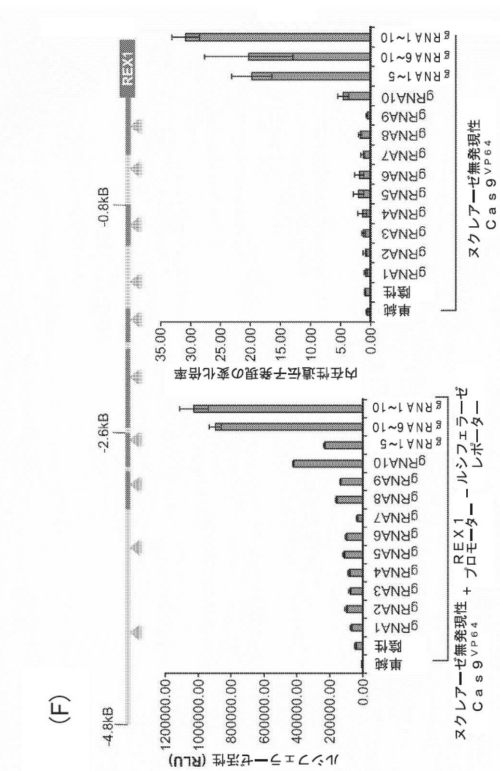
【図 1 - 4】



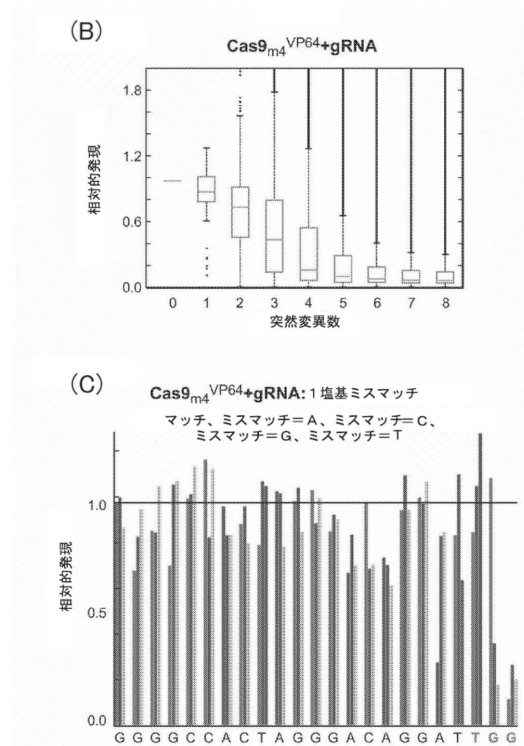
【図 1 - 5】



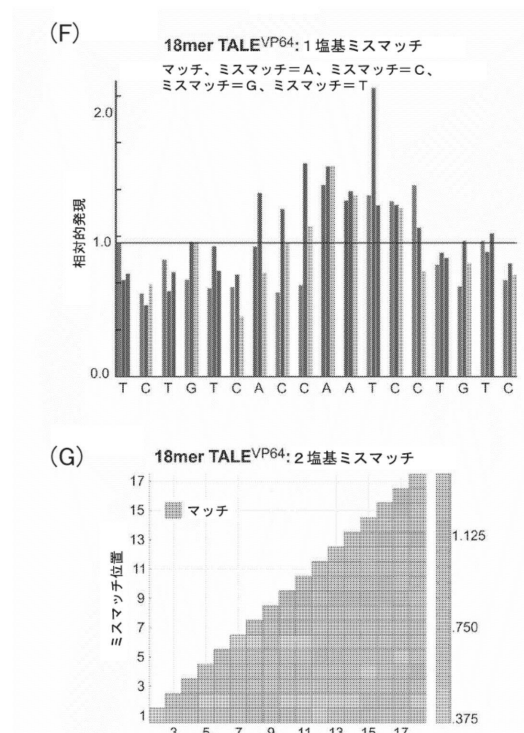
【図 1 - 6】



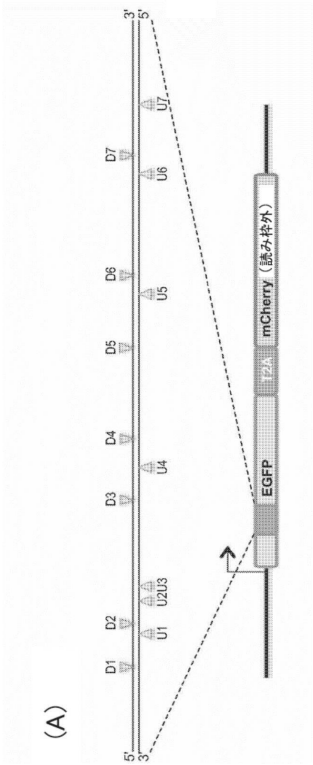
【 図 2 - 2 】



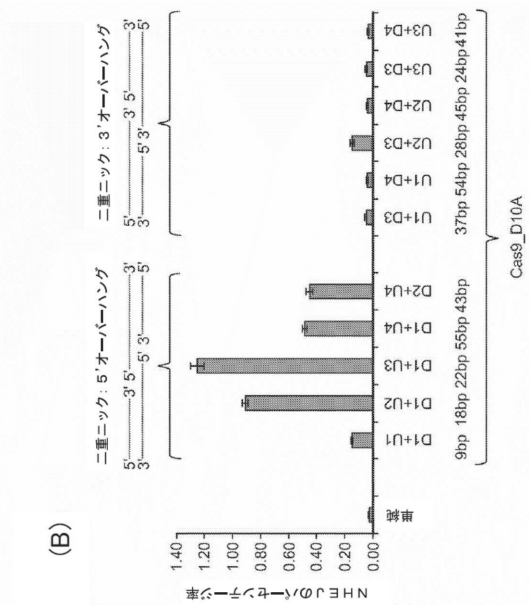
【 図 2 - 4 】



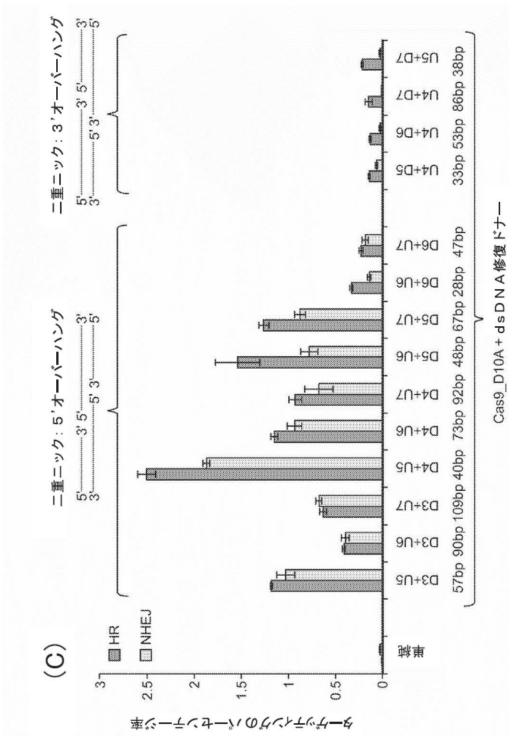
【図 3 - 1】



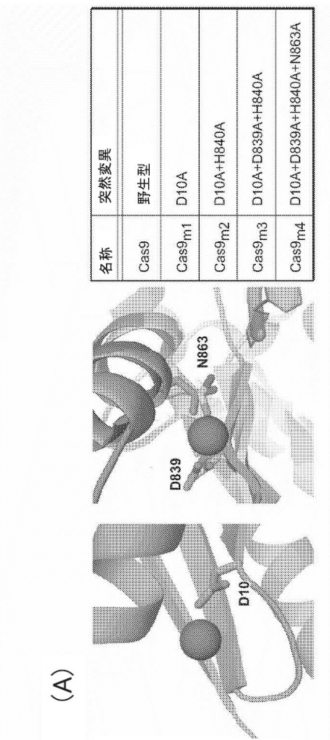
【図 3 - 2】



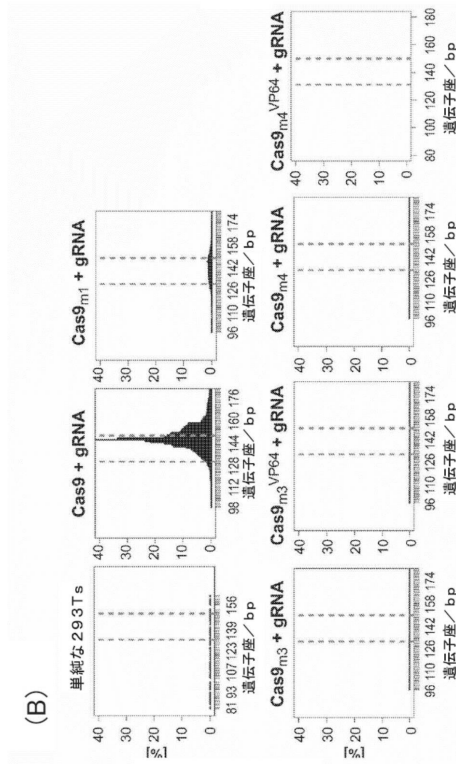
【図 3 - 3】



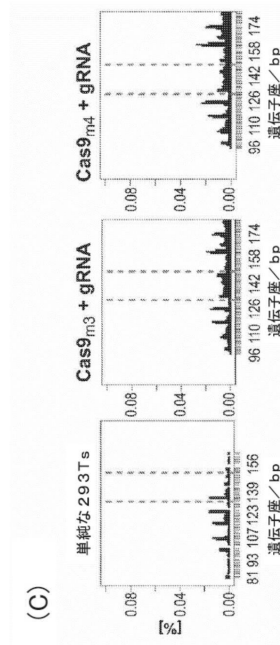
【図 4 - 1】



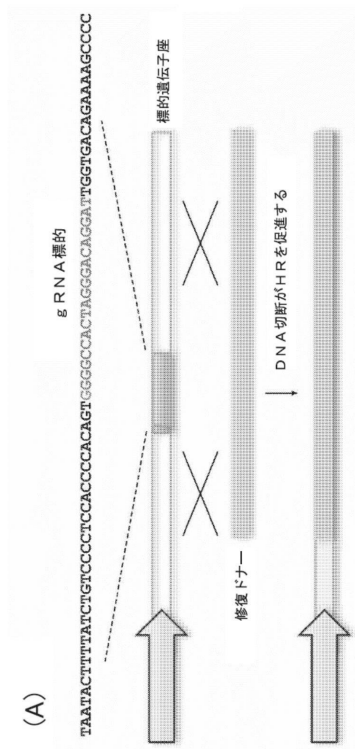
【図 4 - 2】



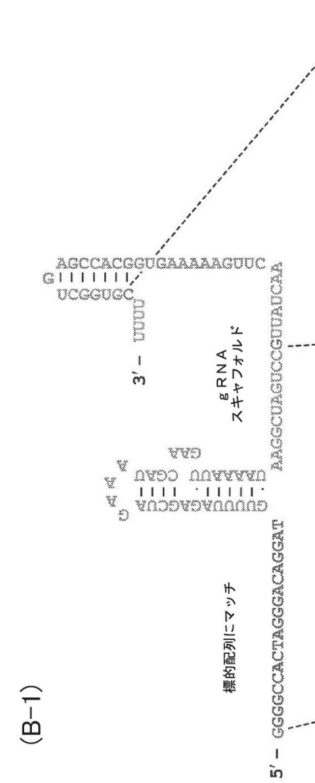
【図 4 - 3】



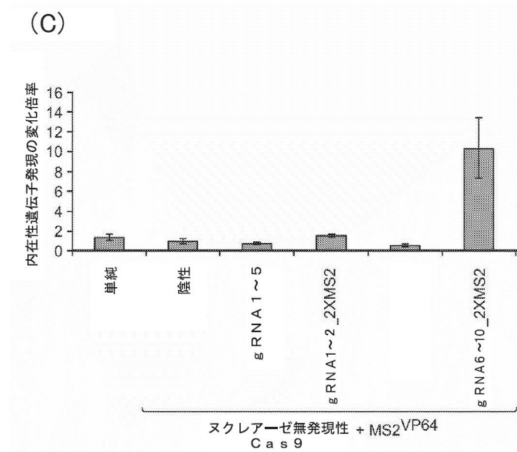
【図 5 - 1】



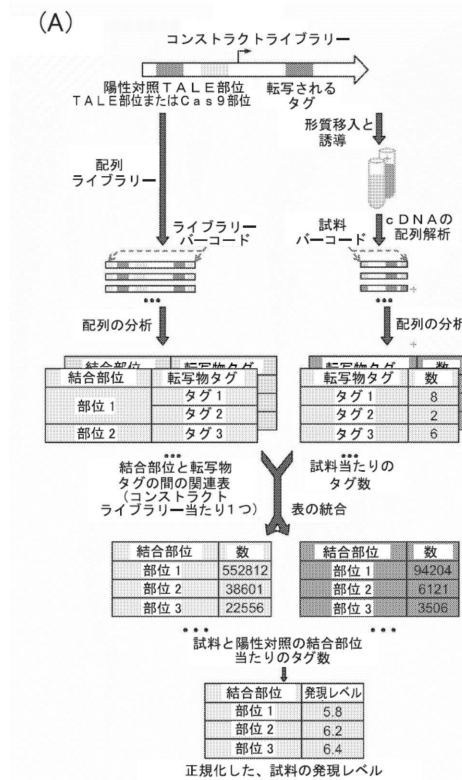
【図 5 - 2】



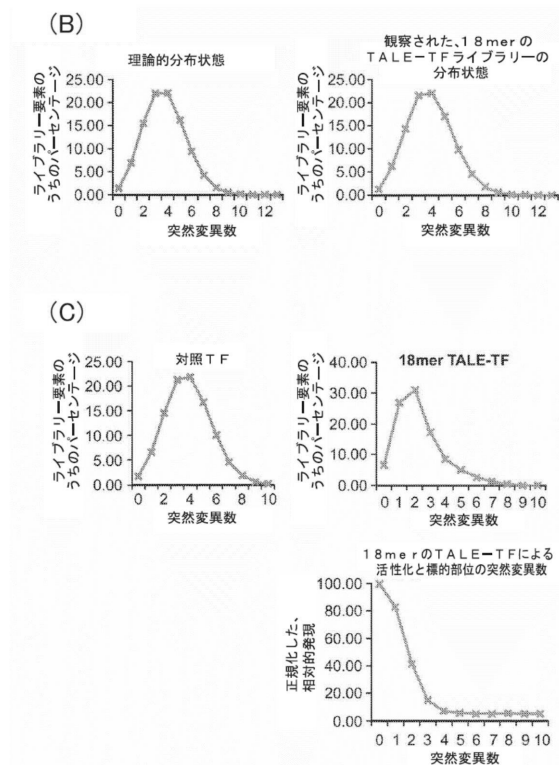
【図 7 - 2】



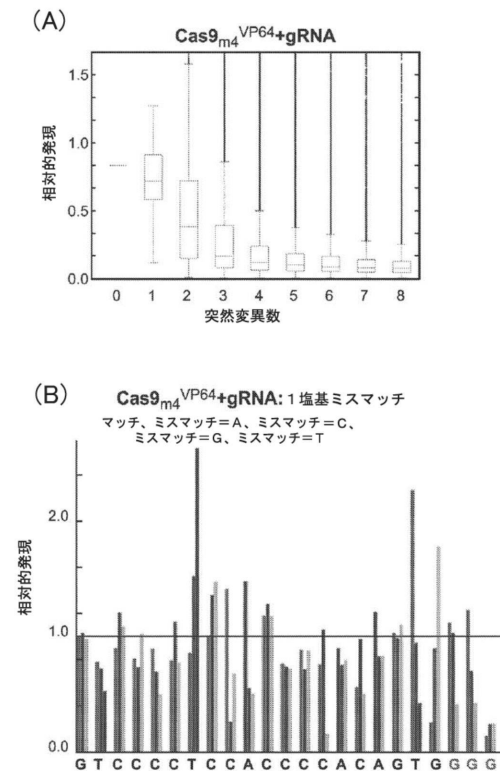
【図 8 - 1】



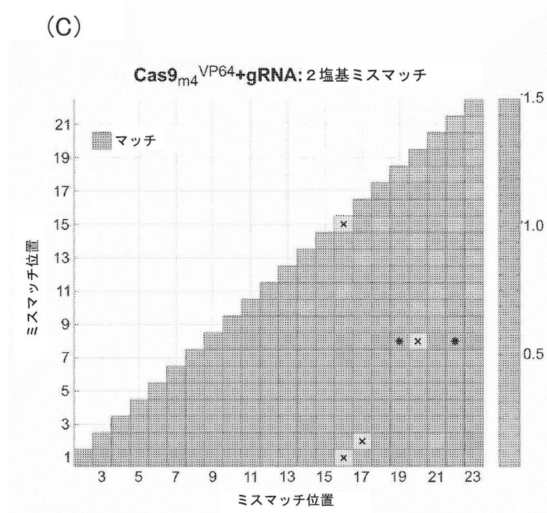
【図 8 - 2】



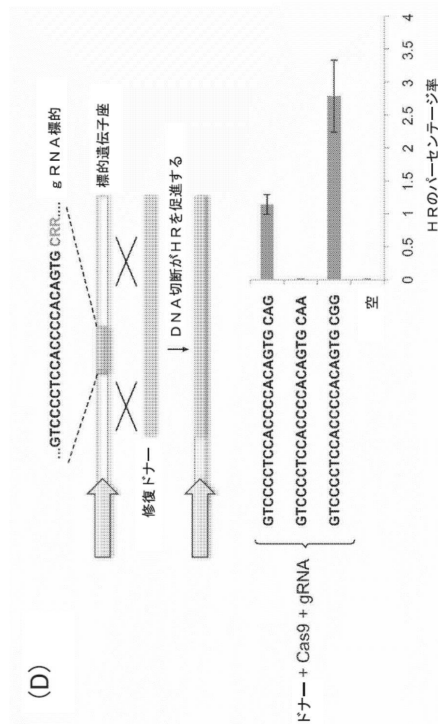
【図 9 - 1】



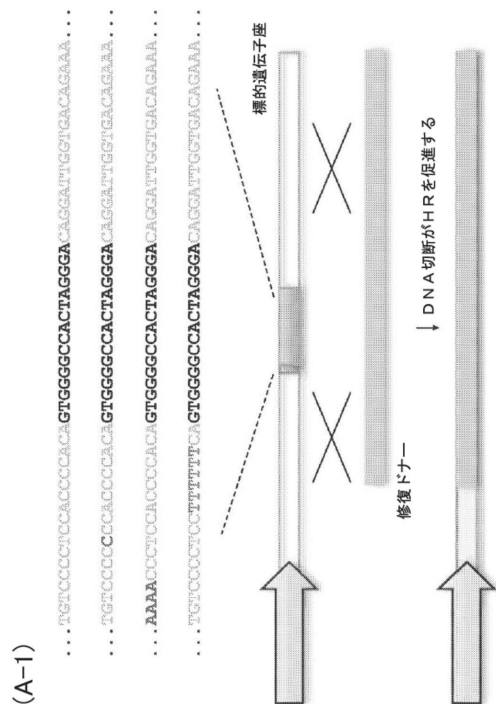
【図 9 - 2】



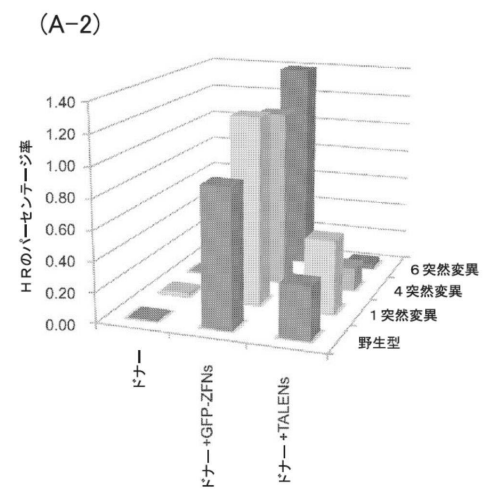
【図 9 - 3】



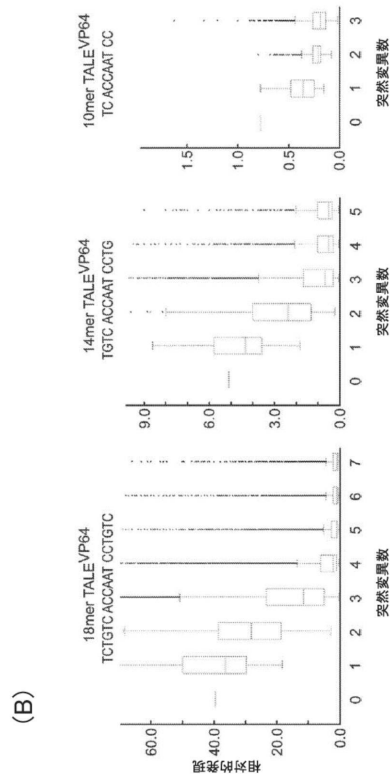
【図 10 - 1】



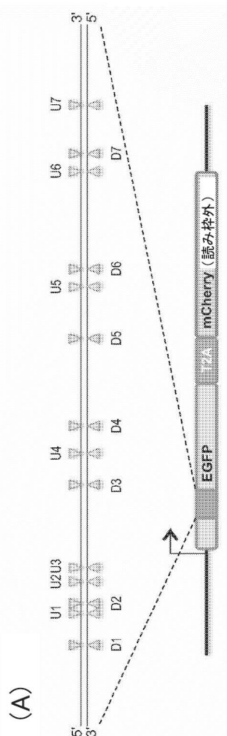
【図 10 - 2】



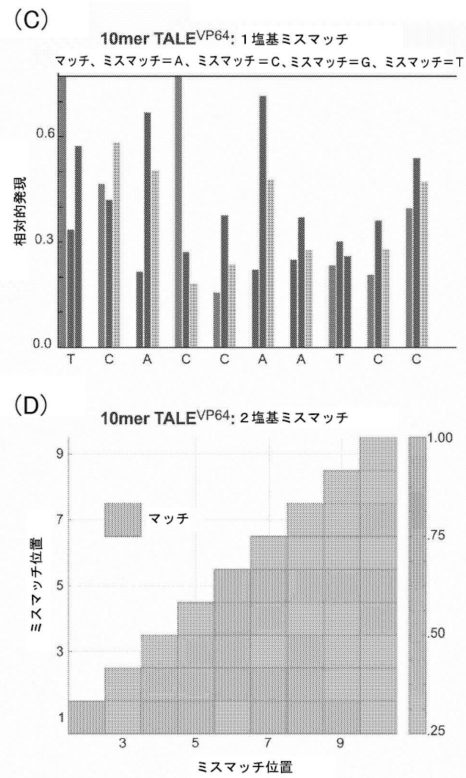
【図 10 - 3】



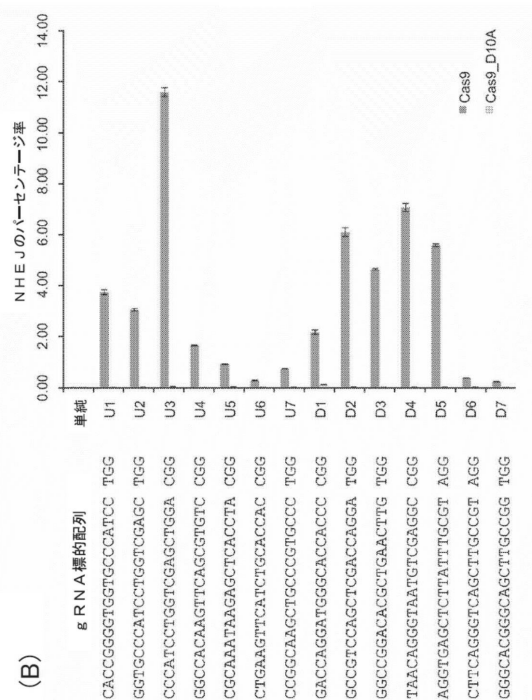
【図 11 - 1】



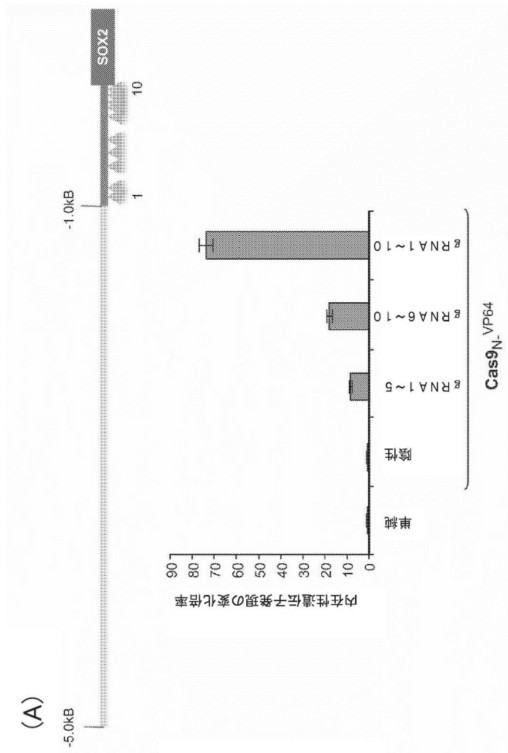
【図 10 - 4】



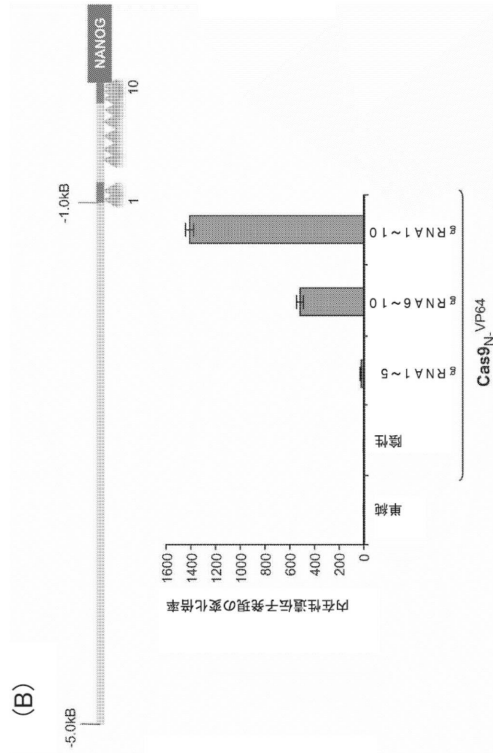
【図 11 - 2】



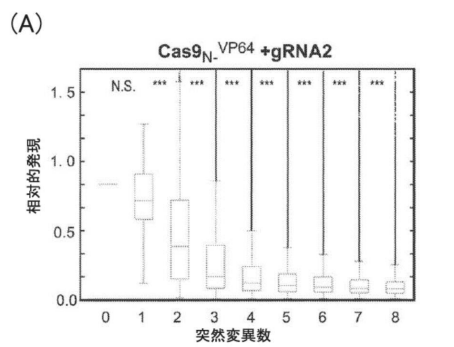
【図 12 - 1】



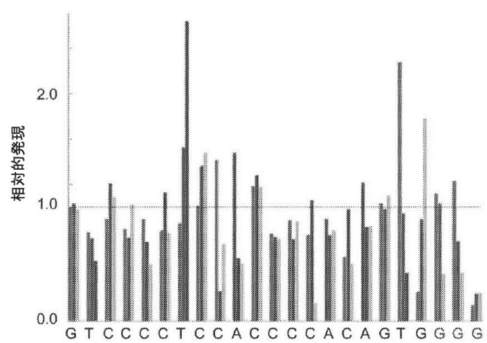
【図 12 - 2】



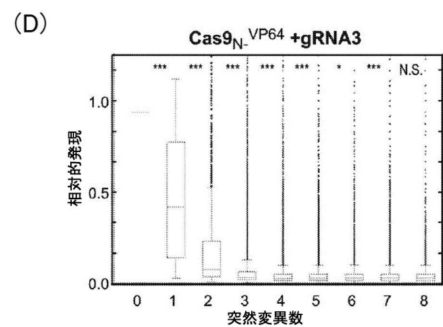
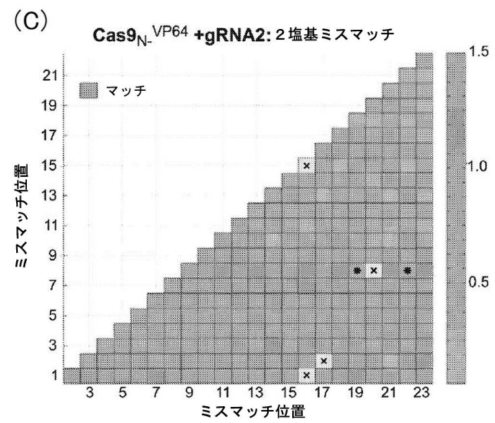
【図 13 - 1】



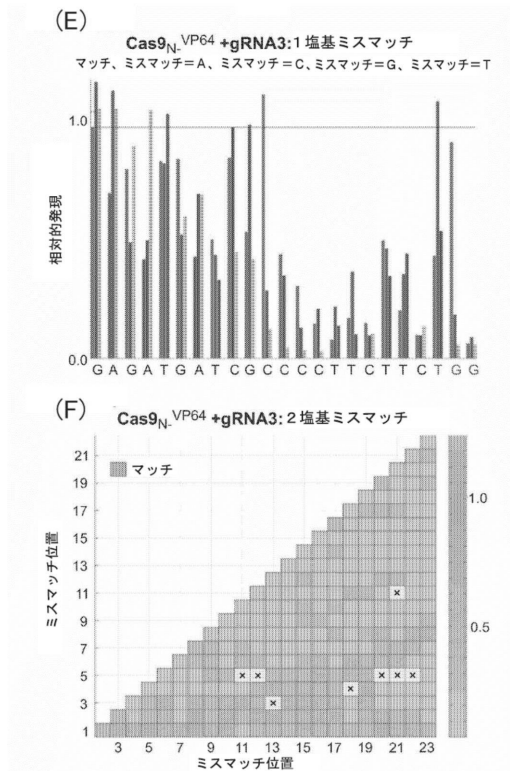
(B) Cas9_N^{VP64} + gRNA2: 1塩基ミスマッチ
マッチ、ミスマッチ=A、ミスマッチ=C、ミスマッチ=G、ミスマッチ=T



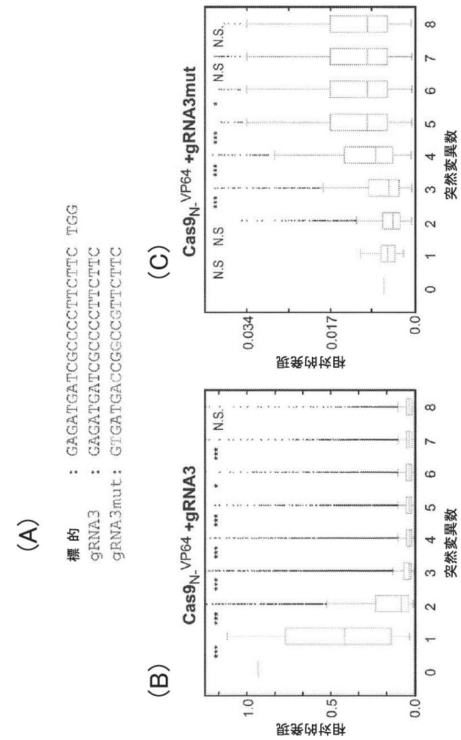
【図 13 - 2】



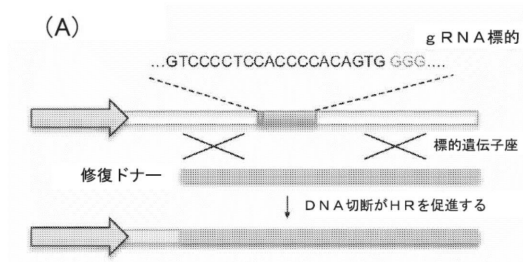
【図 13 - 3】



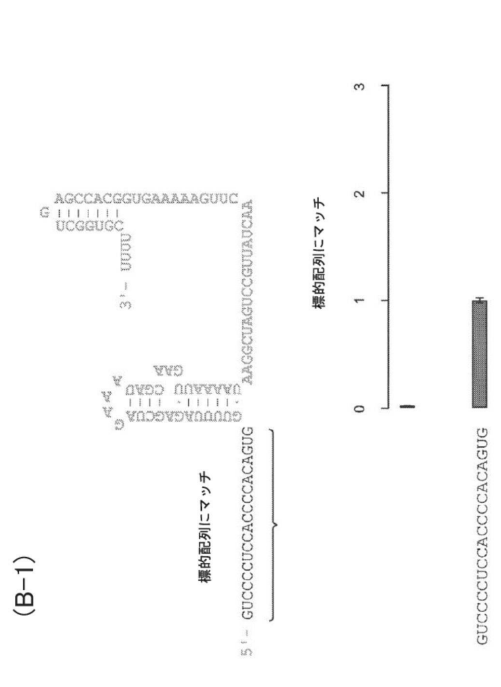
【図 14】



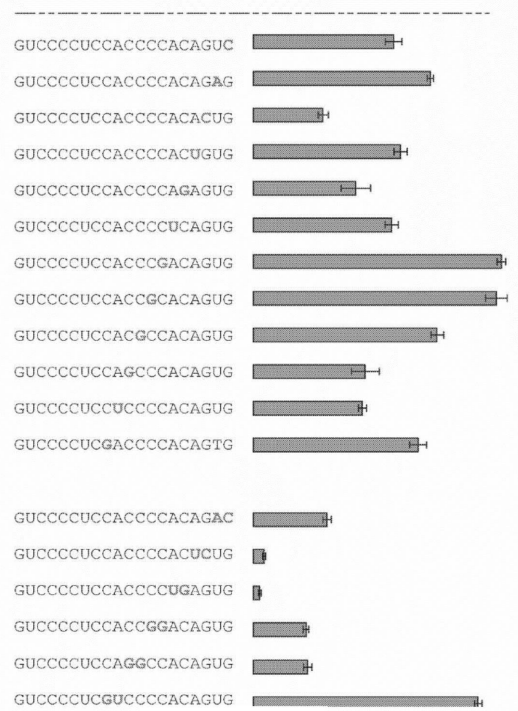
【図 15 - 1】



【図 15 - 2】

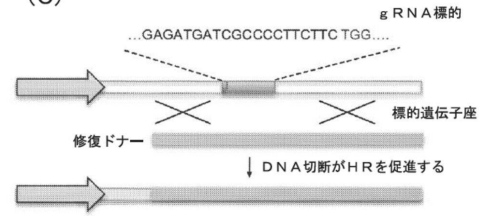


(B-2)

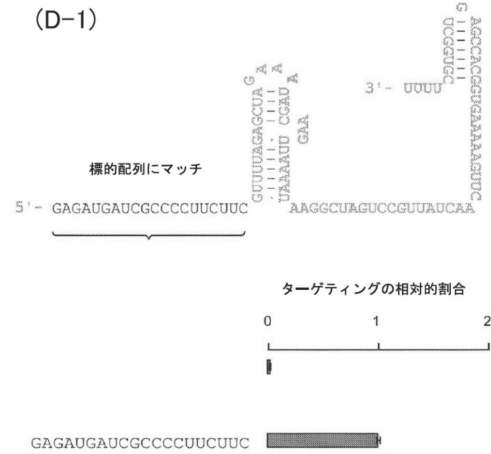


【 図 1 5 - 4 】

(C)

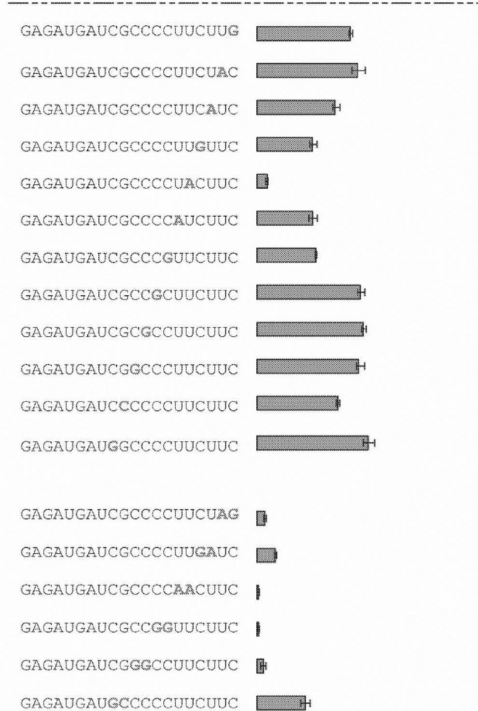


(D-1)



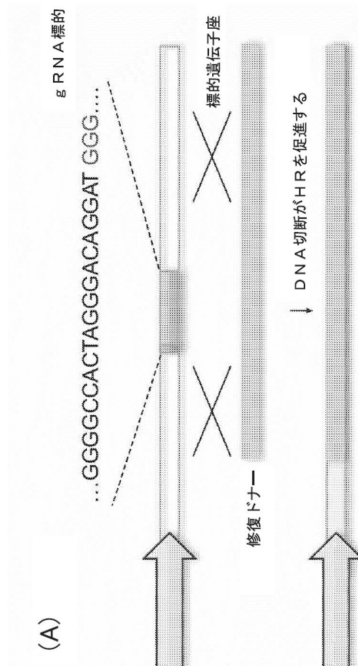
【 図 1 5 - 5 】

(D-2)



【 図 1 6 - 1 】

標的



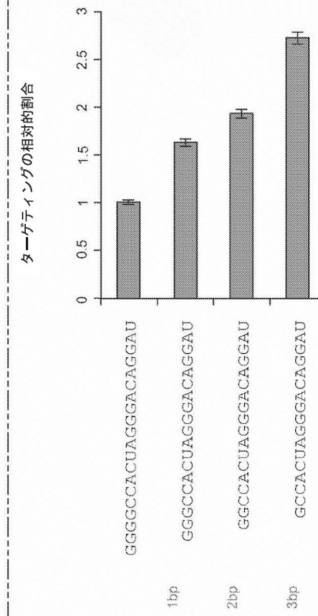
【図 16 - 2】

(B-1)



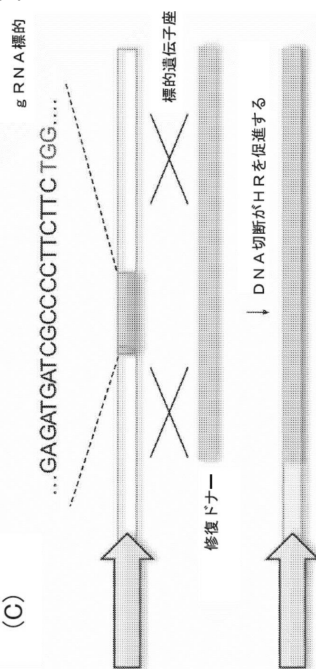
【図 16 - 3】

(B-2)



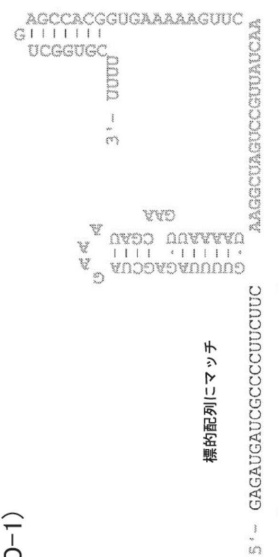
【図 16 - 4】

(C)

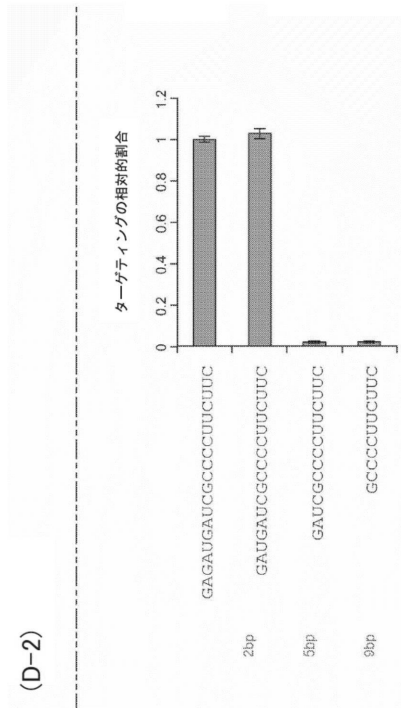


【図 16 - 5】

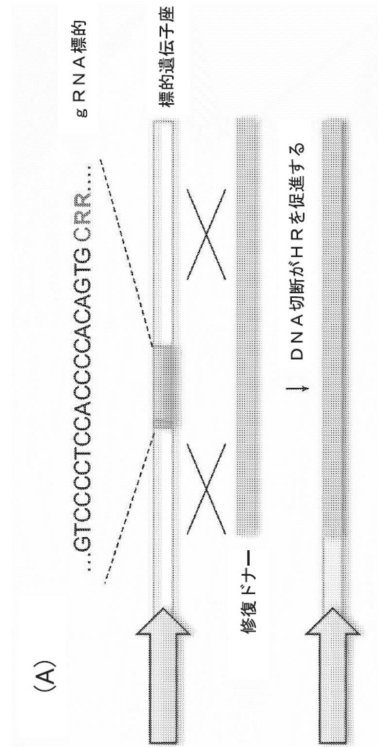
(D-1)



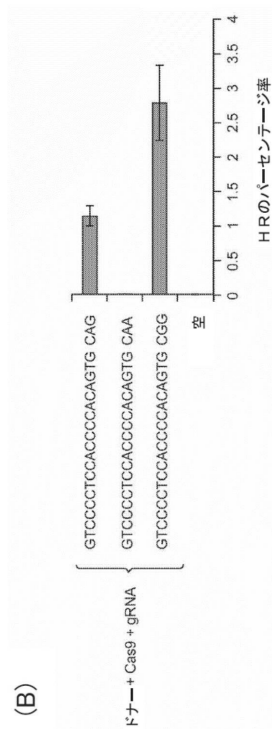
【図 16 - 6】



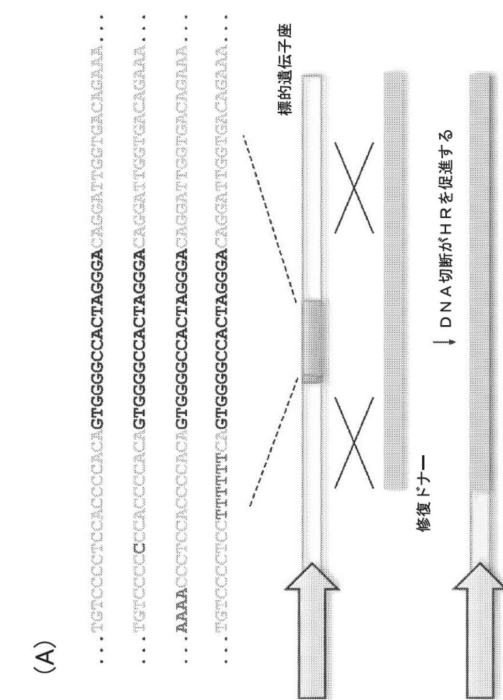
【図 17 - 1】



【図 17 - 2】



【図 18 - 1】



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 ビーコン ストリート 88
- (72)発明者 エスベルト、 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)
CPlus/MEDLINE/EMBASE/BIOSIS(STN)