(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7705382号
(P7705382)

(45)発行日　令和7年7月9日(2025.7.9)

(24)登録日　令和7年7月1日(2025.7.1)

(51)国際特許分類		FI		
C12N 15/09 (2006.01)		C12N	15/09	110
C12N 1/21 (2006.01)		C12N	1/21	
C12N 5/10 (2006.01)		C12N	5/10	
C12Q 1/6816(2018.01)		C12Q	1/6816	Z
C12N 15/85 (2006.01)		C12N	15/85	Z ZNA

請求項の数　29　（全134頁）　最終頁に続く

(21)出願番号　特願2022-515511(P2022-515511)
(86)(22)出願日　令和2年9月9日(2020.9.9)
(65)公表番号　特表2022-547524(P2022-547524A)
(43)公表日　令和4年11月14日(2022.11.14)
(86)国際出願番号　PCT/US2020/049923
(87)国際公開番号　WO2021/050534
(87)国際公開日　令和3年3月18日(2021.3.18)
審査請求日　令和5年9月7日(2023.9.7)
(31)優先権主張番号　62/897,859
(32)優先日　令和1年9月9日(2019.9.9)
(33)優先権主張国・地域又は機関
米国(US)

(73)特許権者　518015608
アーバー バイオテクノロジーズ， インコーポレイテッド
アメリカ合衆国 マサチューセッツ ０２１４０， ケンブリッジ， エイコーン パーク ドライブ ２０， タワー ５００
(74)代理人　100078282
弁理士　山本 秀策
(74)代理人　100113413
弁理士　森下 夏樹
(74)代理人　100181674
弁理士　飯田 貴敏
(74)代理人　100181641
弁理士　石川 大輔
(74)代理人　230113332

最終頁に続く

(54)【発明の名称】　新規ＣＲＩＳＰＲ ＤＮＡターゲティング酵素及びシステム

(57)【特許請求の範囲】

【請求項１】

エンジニアリングされた天然に存在しないクラスター化して規則的な配置の短い回文配列リピート（ＣＲＩＳＰＲ）－Ｃａｓシステムであって、

（ａ）ＣＲＩＳＰＲ関連タンパク質が配列番号２４１のアミノ酸配列を含む、ＣＲＩＳＰＲ関連タンパク質又は前記ＣＲＩＳＰＲ関連タンパク質をコードする核酸；及び

（ｂ）ダイレクトリピート配列と標的核酸へのハイブリダイゼーション能を有するスペーサー配列とを含むＲＮＡガイド

を含み、

前記ＣＲＩＳＰＲ関連タンパク質が、前記ＲＮＡガイドに結合し、前記スペーサー配列に相補的な前記標的核酸配列を修飾することができ、

前記ＣＲＩＳＰＲ関連タンパク質が、配列番号１０、配列番号４、配列番号１２、又は配列番号１４に記載のアミノ酸配列と少なくとも９５％同一のアミノ酸配列を含む、ＣＲＩＳＰＲ－Ｃａｓシステム。

【請求項２】

前記ＣＲＩＳＰＲ関連タンパク質が、配列番号１０に記載のアミノ酸配列と少なくとも９５％同一のアミノ酸配列を含む、請求項１に記載のシステム。

【請求項３】

前記ＣＲＩＳＰＲ関連タンパク質が、配列番号４に記載のアミノ酸配列と少なくとも９５％同一のアミノ酸配列を含む、請求項１に記載のシステム。

【請求項４】
前記ＣＲＩＳＰＲ関連タンパク質が、少なくとも１つのＲｕｖＣドメイン又は少なくとも１つの分割されたＲｕｖＣドメインを含む、請求項１～３のいずれか一項に記載のシステム。
【請求項５】
前記ＣＲＩＳＰＲ関連タンパク質が、
ａ）触媒残基を含む；
ｂ）前記標的核酸を切断する；
ｃ）ペプチドタグ、蛍光タンパク質、塩基編集ドメイン、ＤＮＡメチル化ドメイン、ヒストン残基修飾ドメイン、局在化因子、転写修飾因子、光ゲート制御因子、化学誘導性因子、若しくはクロマチン可視化因子を更に含む；又は
ｄ）自己プロセシングである、
請求項１～４のいずれか一項に記載のシステム。
【請求項６】
前記ＣＲＩＳＰＲ関連タンパク質をコードする前記核酸が、
ａ）細胞での発現にコドン最適化される；
ｂ）プロモーターに作動可能に連結されている；又は
ｃ）ベクター内にある、
請求項１～５のいずれか一項に記載のシステム。
【請求項７】
前記標的核酸がＤＮＡ分子である、請求項１～６のいずれか一項に記載のシステム。
【請求項８】
前記ＣＲＩＳＰＲ関連タンパク質及びＲＮＡガイドによる前記標的核酸の認識により、前記標的核酸の修飾が生じる、請求項１～７のいずれか一項に記載のシステム。
【請求項９】
ドナー鋳型核酸を更に含む、請求項１～８のいずれか一項に記載のシステム。
【請求項１０】
前記ドナー鋳型核酸がＤＮＡ分子である、請求項９に記載のシステム。
【請求項１１】
前記システムが、
ａ）ナノ粒子、リポソーム、エキソソーム、微小胞、若しくは遺伝子銃を含む送達組成物中に存在する；又は
ｂ）細胞内にある、
請求項１～１０のいずれか一項に記載のシステム。
【請求項１２】
（ａ）ＣＲＩＳＰＲ関連タンパク質又は前記ＣＲＩＳＰＲ関連タンパク質をコードする核酸であって、前記ＣＲＩＳＰＲ関連タンパク質が、配列番号２４１のアミノ酸配列を含み、前記ＣＲＩＳＰＲ関連タンパク質が、配列番号１０、配列番号４、配列番号１２、又は配列番号１４に記載のアミノ酸配列と少なくとも９５％同一のアミノ酸配列を含む、ＣＲＩＳＰＲ関連タンパク質又は前記ＣＲＩＳＰＲ関連タンパク質をコードする核酸；及び
（ｂ）ダイレクトリピート配列と標的核酸へのハイブリダイゼーション能を有するスペーサー配列とを含むＲＮＡガイド
を含む、細胞。
【請求項１３】
請求項１～１１のいずれか一項に記載のシステムを、エキソビボで細胞内の標的核酸に結合させる方法であって、
（ａ）前記システムを提供すること；及び
（ｂ）前記システムを前記細胞に送達すること
を含み、前記細胞が前記標的核酸を含み、前記ＣＲＩＳＰＲ関連タンパク質が前記ＲＮＡガイドに結合し、前記スペーサー配列が前記標的核酸に結合する、方法。

【請求項１４】
前記細胞が原核細胞、真核細胞、哺乳動物細胞、又はヒト細胞である、請求項１２に記載の細胞。
【請求項１５】
標的核酸を修飾する方法であって、
（ａ）ＣＲＩＳＰＲ関連タンパク質又は前記ＣＲＩＳＰＲ関連タンパク質をコードする核酸であって、前記ＣＲＩＳＰＲ関連タンパク質が、配列番号２４１のアミノ酸配列を含み、前記ＣＲＩＳＰＲ関連タンパク質が、配列番号１０、配列番号４、配列番号１２、又は配列番号１４に記載のアミノ酸配列と少なくとも９５％同一のアミノ酸配列を含む、ＣＲＩＳＰＲ関連タンパク質又は前記ＣＲＩＳＰＲ関連タンパク質をコードする核酸；及び
（ｂ）ダイレクトリピート配列と前記標的核酸へのハイブリダイゼーション能を有するスペーサー配列とを含むＲＮＡガイド；
を含み、
前記ＣＲＩＳＰＲ関連タンパク質がＲＮＡガイドへの結合能を有し、
前記ＣＲＩＳＰＲ関連タンパク質及びＲＮＡガイドによる前記標的核酸の認識により、前記標的核酸の修飾が生じる、
エンジニアリングされた天然に存在しないＣＲＩＳＰＲ－Ｃａｓシステムを、標的核酸に送達することを含む方法。
【請求項１６】
前記ＣＲＩＳＰＲ関連タンパク質が、以下の配列の１つ以上：
（ａ）ＰＸ$_1$Ｘ$_2$Ｘ$_3$Ｘ$_4$Ｆ（配列番号２１６）（ここで、Ｘ$_1$はＬ又はＭ又はＩ又はＣ又はＦであり、Ｘ$_2$はＹ又はＷ又はＦであり、Ｘ$_3$はＫ又はＴ又はＣ又はＲ又はＷ又はＹ又はＨ又はＶであり、Ｘ$_4$はＩ又はＬ又はＭである）；
（ｂ）ＲＸ$_1$Ｘ$_2$Ｘ$_3$Ｌ（配列番号２１７）（ここで、Ｘ$_1$はＩ又はＬ又はＭ又はＹ又はＴ又はＦであり、Ｘ$_2$はＲ又はＱ又はＫ又はＥ又はＳ又はＴであり、Ｘ$_3$はＬ又はＩ又はＴ又はＣ又はＭ又はＫである）；
（ｃ）ＮＸ$_1$ＹＸ$_2$（配列番号２１８）（ここで、Ｘ$_1$はＩ又はＬ又はＦであり、Ｘ$_2$はＫ又はＲ又はＶ又はＥである）；
（ｄ）ＫＸ$_1$Ｘ$_2$Ｘ$_3$ＦＡＸ$_4$Ｘ$_5$ＫＤ（配列番号２１９）（ここで、Ｘ$_1$はＴ又はＩ又はＮ又はＡ又はＳ又はＦ又はＶであり、Ｘ$_2$はＩ又はＶ又はＬ又はＳであり、Ｘ$_3$はＨ又はＳ又はＧ又はＲであり、Ｘ$_4$はＤ又はＳ又はＥであり、Ｘ$_5$はＩ又はＶ又はＭ又はＴ又はＮである）；
（ｅ）ＬＸ$_1$ＮＸ$_2$（配列番号２２０）（ここで、Ｘ$_1$はＧ又はＳ又はＣ又はＴであり、Ｘ$_2$はＮ又はＹ又はＫ又はＳである）；
（ｆ）ＰＸ$_1$Ｘ$_2$Ｘ$_3$Ｘ$_4$ＳＱＸ$_5$ＤＳ（配列番号２２１）（ここで、Ｘ$_1$はＳ又はＰ又はＡであり、Ｘ$_2$はＹ又はＳ又はＡ又はＰ又はＥ又はＹ又はＱ又はＮであり、Ｘ$_3$はＦ又はＹ又はＨであり、Ｘ$_4$はＴ又はＳであり、Ｘ$_5$はＭ又はＴ又はＩである）；
（ｇ）ＫＸ$_1$Ｘ$_2$ＶＲＸ$_3$Ｘ$_4$ＱＥＸ$_5$Ｈ（配列番号２２２）（ここで、Ｘ$_1$はＮ又はＫ又はＷ又はＲ又はＥ又はＴ又はＹであり、Ｘ$_2$はＭ又はＲ又はＬ又はＳ又はＫ又はＶ又はＥ又はＴ又はＩ又はＤであり、Ｘ$_3$はＬ又はＲ又はＨ又はＰ又はＴ又はＫ又はＰ又はＱ又はＳ又はＡであり、Ｘ$_4$はＧ又はＱ又はＮ又はＲ又はＫ又はＥ又はＩ又はＴ又はＳ又はＣであり、Ｘ$_5$はＲ又はＷ又はＹ又はＫ又はＴ又はＦ又はＳ又はＱである）；及び
（ｈ）Ｘ$_1$ＮＧＸ$_2$Ｘ$_3$Ｘ$_4$ＤＸ$_5$ＮＸ$_6$Ｘ$_7$Ｘ$_8$Ｎ（配列番号２２３）（ここで、Ｘ$_1$はＩ又はＫ又はＶ又はＬであり、Ｘ$_2$はＬ又はＭであり、Ｘ$_3$はＮ又はＨ又はＰであり、Ｘ$_4$はＡ又はＳ又はＣであり、Ｘ$_5$はＶ又はＹ又はＩ又はＦ又はＴ又はＮであり、Ｘ$_6$はＡ又はＳであり、Ｘ$_7$はＳ又はＡ又はＰであり、Ｘ$_8$はＭ又はＣ又はＬ又はＲ又はＮ又はＳ又はＫ又はＬである）
を含む、請求項１～１１のいずれか一項に記載のシステム。
【請求項１７】
前記ダイレクトリピート配列が、

ｉ）配列番号５７～９０、配列番号１１８～１５１、又は配列番号２１３のいずれか１つに記載のヌクレオチド配列と少なくとも８０％同一のヌクレオチド配列を含む；又は
ｉｉ）以下の配列の１つ以上：
（ａ）$X_1X_2TX_3X_4X_5X_6X_7X_8$（配列番号２２４）（ここで、$X_1$はＡ又はＣ又はＧであり、$X_2$はＴ又はＣ又はＡであり、$X_3$はＴ又はＧ又はＡであり、$X_4$はＴ又はＧであり、$X_5$はＴ又はＧ又はＡであり、$X_6$はＧ又はＴ又はＡであり、$X_7$はＴ又はＧ又はＡであり、$X_8$はＡ又はＧ又はＴである）；
（ｂ）$X_1X_2X_3X_4X_5X_6X_7X_8X_9$（配列番号２２６）（ここで、$X_1$はＴ又はＣ又はＡであり、$X_2$はＴ又はＡ又はＧであり、$X_3$はＴ又はＣ又はＡであり、$X_4$はＴ又はＡであり、$X_5$はＴ又はＡ又はＧであり、$X_6$はＴ又はＡであり、$X_7$はＡ又はＴであり、$X_8$はＡ又はＧ又はＣ又はＴであり、$X_9$はＧ又はＡ又はＣである）；及び
（ｃ）$X_1X_2X_3$ＡＣ（配列番号２２８）（ここで、X_1はＡ又はＣ又はＧであり、X_2はＣ又はＡであり、X_3はＡ又はＣである）
を含む、
請求項１～１１又は１６のいずれか一項に記載のシステム。
【請求項１８】
前記ＣＲＩＳＰＲ関連タンパク質が、
（ａ）配列番号４に記載のアミノ酸配列と少なくとも９５％の同一性を有するタンパク質であり、前記ダイレクトリピート配列が、配列番号６０に記載のヌクレオチド配列と少なくとも８０％同一のヌクレオチド配列を含む；
（ｂ）配列番号４に記載のアミノ酸配列と少なくとも９５％の同一性を有するタンパク質であり、前記ＣＲＩＳＰＲ関連タンパク質が、ＰＡＭ配列の認識能を有し、前記ＰＡＭ配列が、５’－ＮＴＴＮ－３’、５’－ＮＴＴＲ－３’又は５’－ＮＮＲ－３’として記載される核酸配列を含み、「Ｎ」が任意のヌクレオチドであり、「Ｒ」がＡ又はＧである；
（ｃ）配列番号１０に記載のアミノ酸配列と少なくとも９５％の同一性を有するタンパク質であり、前記ダイレクトリピート配列が、配列番号６２又は配列番号２１３に記載のヌクレオチド配列と少なくとも９５％同一のヌクレオチド配列を含む；又は
（ｄ）配列番号１０に記載のアミノ酸配列と少なくとも９５％の同一性を有するタンパク質であり、前記ＣＲＩＳＰＲ関連タンパク質が、ＰＡＭ配列の認識能を有し、前記ＰＡＭ配列が、５’－ＮＴＴＮ－３’又は５’－ＲＴＴＲ－３’として記載される核酸配列を含み、「Ｎ」が任意のヌクレオチドであり、「Ｒ」がＡ又はＧである、
請求項１～１１、１６又は１７のいずれか一項に記載のシステム。
【請求項１９】
前記ダイレクトリピート配列が、チミンとして示される１つ以上の場所に、ウラシルを含む、請求項１～１１又は１６～１８のいずれか一項に記載のシステム。
【請求項２０】
前記スペーサー配列が、１５ヌクレオチド～５５ヌクレオチドを含む、請求項１～１１又は１６～１９のいずれか一項に記載のシステム。
【請求項２１】
前記システムがｔｒａｃｒＲＮＡを含まない、請求項１～１１又は１６～２０のいずれか一項に記載のシステム。
【請求項２２】
前記標的核酸がＤＮＡ分子である、請求項１～１１又は１６～２１のいずれか一項に記載のシステム。
【請求項２３】
（ａ）前記標的核酸の前記修飾が、二本鎖切断イベントである；
（ｂ）前記標的核酸の前記修飾が、一本鎖切断イベントである；
（ｃ）前記標的核酸の前記修飾により、挿入イベントが生じる；又は
（ｄ）前記標的核酸の前記修飾により、欠失イベントが生じる、
請求項１～１１又は１６～２２のいずれか一項に記載のシステム。

【請求項２４】
（ａ）標的核酸のターゲティング及び編集方法；
（ｂ）核酸の認識に応じた一本鎖核酸の非特異的分解方法；
（ｃ）二本鎖標的のスペーサー相補鎖の認識に応じた二本鎖標的の非スペーサー相補鎖のターゲティング及びニッキング方法；
（ｄ）二本鎖標的核酸のターゲティング及び切断方法；
（ｅ）試料中の標的核酸の検出方法；
（ｆ）二本鎖核酸の特異的編集方法；
（ｇ）二本鎖核酸の塩基編集方法；
（ｈ）細胞における遺伝子型特異的又は転写状態特異的細胞死又は休眠の誘導方法；
（ｉ）二本鎖核酸標的におけるインデルの作成方法；
（ｊ）二本鎖核酸標的への配列の挿入方法；又は
（ｋ）二本鎖核酸標的における配列の欠失又は逆位形成方法
である、インビトロ又はエキソビボでの方法における、請求項１～１１又は１６～２３のいずれか一項に記載のシステムの使用。
【請求項２５】
哺乳動物細胞における標的核酸中への挿入又は欠失を導入する方法における使用のための、請求項１～１１又は１６～２３のいずれか一項に記載のエンジニアリングされた天然に存在しないＣＲＩＳＰＲ－Ｃａｓシステムを含む組成物。
【請求項２６】
前記ＣＲＩＳＰＲ関連タンパク質が、配列番号４又は１０に記載のアミノ酸配列と少なくとも９５％同一のアミノ酸配列を含む、請求項２５に記載の組成物。
【請求項２７】
前記ダイレクトリピートが、配列番号６０に記載のヌクレオチド配列と少なくとも８０％同一のヌクレオチド配列を含む、請求項２５又は２６に記載の組成物。
【請求項２８】
前記標的核酸がＰＡＭ配列に隣接しており、前記ＰＡＭ配列が、５’－ＮＴＴＮ－３’、５’－ＮＴＴＲ－3’又は５’－ＮＮＲ－３’として記載される核酸配列を含み、「Ｎ」が任意のヌクレオチドであり、「Ｒ」がＡ又はＧである、請求項２５～２７のいずれか一項に記載の組成物。
【請求項２９】
（ｉ）前記トランスフェクションが一過性トランスフェクションである；又は
（ｉｉ）前記細胞がヒト細胞である、
請求項２５～２８のいずれか一項に記載の組成物。
【発明の詳細な説明】
【技術分野】
【０００１】
関連出願
本願は、２０１９年９月９日に出願された米国仮特許出願第６２／８９７，８５９号の優先権の利益を主張し、その全内容は、本明細書によって参照により援用される。
【０００２】
配列表
本出願は、ＡＳＣＩＩ形式で電子的に提出され、全体として参照により本明細書で援用される配列表を含む。２０２０年９月９日に作成された前記ＡＳＣＩＩコピーは、Ａ２１８６－７０２８ＷＯ＿ＳＬ．ｔｘｔというファイル名で、サイズは４７５，５１１バイトである。
【０００３】
本開示は、新規の、クラスター化して規則的な配置の短い回文配列リピート（Ｃｌｕｓｔｅｒｅｄ Ｒｅｇｕｌａｒｌｙ Ｉｎｔｅｒｓｐａｃｅｄ Ｓｈｏｒｔ Ｐａｌｉｎｄｒｏｍｉｃ Ｒｅｐｅａｔｓ：ＣＲＩＳＰＲ）及びＣＲＩＳＰＲ関連（Ｃａｓ）遺伝子を使

用する遺伝子発現のゲノム編集及び調節のためのシステム及び方法に関する。
【背景技術】
【０００４】
近年、ゲノムシーケンシング技術及び解析の進歩により、原核生物の生合成経路からヒト病理に及ぶまでの多種多様な自然領域における生物学的活性の遺伝的基礎に関して重要な洞察が得られている。得られた大量の情報を完全に理解し、評価するためには、対応したゲノム及びエピゲノム操作のシーケンス技術の規模、有効性、及び容易さの向上が必要となる。このような新規技術は、バイオテクノロジー、農業、及びヒト治療薬を含めた数多くの領域における新規適用の開発を加速させることになる。
【０００５】
クラスター化して規則的な配置の短い回文配列リピート（ＣＲＩＳＰＲ）及びＣＲＩＳＰＲ関連（Ｃａｓ）遺伝子は、まとめてＣＲＩＳＰＲ－Ｃａｓ又はＣＲＩＳＰＲ／Ｃａｓシステムとして知られ、外来の遺伝的エレメントから特定の種を防御する古細菌及び細菌の適応免疫系である。ＣＲＩＳＰＲ－Ｃａｓシステムは極めて多様な一群のタンパク質エフェクター、非コードエレメント、並びに遺伝子座構成を含み、その幾つかの例がエンジニアリングされ、適合されることにより、重要なバイオテクノロジーの進歩が生み出されている。
【０００６】
宿主防御に関与するこのシステムの成分には、核酸を修飾する能力を有する１つ以上のエフェクタータンパク質と、エフェクタータンパク質をファージ核酸上の特異的配列に標的化することを担うＲＮＡガイドエレメントとが含まれる。ＲＮＡガイドはＣＲＩＳＰＲＲＮＡ（ｃｒＲＮＡ）で構成され、１つ又は複数のエフェクタータンパク質による標的核酸の操作を実現するために追加的なトランス活性化型ＲＮＡ（ｔｒａｃｒＲＮＡ）を必要とすることもある。ｃｒＲＮＡは、ｃｒＲＮＡへのタンパク質結合を担うダイレクトリピートと、所望の核酸標的配列に相補的なスペーサー配列とからなる。ＣＲＩＳＰＲシステムは、ｃｒＲＮＡのスペーサー配列を修飾することにより、別のＤＮＡ又はＲＮＡ標的を標的化するよう再プログラム化し得る。
【０００７】
ＣＲＩＳＰＲ－Ｃａｓシステムは、大きく２つのクラスに分けることができる：クラス１システムは、一緒になってｃｒＲＮＡの周りに複合体を形成する複数のエフェクタータンパク質で構成され、クラス２システムは、ＲＮＡガイドと複合体化して核酸基質を標的化する単一のエフェクタータンパク質からなる。クラス２システムのシングルサブユニットのエフェクター組成は、エンジニアリング及び適用移行に一層簡便な成分セットを提供し、従ってこれまでプログラム可能なエフェクターの重要な供給源となっている。それにも関わらず、核酸及びポリヌクレオチド（即ち、ＤＮＡ、ＲＮＡ、又は任意のハイブリッド、誘導体、又は修飾体）を修飾するための、その独自の特性によって新規適用を実現する、より小さなエフェクター及び／又はユニークなＰＡＭ配列要件を有するエフェクターなどの現在のＣＲＩＳＰＲ－Ｃａｓシステムを越える更なるプログラム可能なエフェクター及びシステムが依然として必要とされている。
【発明の概要】
【課題を解決するための手段】
【０００８】
この開示は、最初にゲノムデータベースから計算により同定され、その後、エンジニアリングされ、実験的に検証された、新規の単一エフェクタークラス２　ＣＲＩＳＰＲ－Ｃａｓシステムのための非天然のエンジニアリングされたシステム及び組成物を提供する。特に、これらのＣＲＩＳＰＲ－Ｃａｓシステムの成分の同定により、非天然環境、例えば、システムが最初に発見されたもの以外の細菌、又は哺乳動物細胞などの真核細胞における使用が可能になる。これらの新規エフェクターは、既存のクラス２　ＣＲＩＳＰＲエフェクターのオルソログ及びホモログと比較して配列及び機能が異なる。
【０００９】

一態様において、本開示は、配列番号１～５６のいずれか１つに記載のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含むＣＲＩＳＰＲ関連タンパク質；及びダイレクトリピート配列と標的核酸へのハイブリダイゼーション能を有するスペーサー配列とを含むＲＮＡガイドを含む、ＣＬＵＳＴ．０９１９７９のエンジニアリングされた天然に存在しないクラスター化して規則的な配置の短い回文配列リピート（ＣＲＩＳＰＲ）－Ｃａｓシステムを提供し、ここで、ＣＲＩＳＰＲ関連タンパク質は、ＲＮＡガイドに結合し、スペーサー配列に相補的な標的核酸配列を修飾することができる。一態様において、本開示は、ＣＲＩＳＰＲ関連タンパク質が配列番号１～５６のいずれか１つに記載のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含む、ＣＲＩＳＰＲ関連タンパク質又はＣＲＩＳＰＲ関連タンパク質をコードする核酸；及びダイレクトリピート配列と標的核酸へのハイブリダイゼーション能を有するスペーサー配列とを含むＲＮＡガイド、又はＲＮＡガイドをコードする核酸を含む、ＣＬＵＳＴ．０９１９７９のエンジニアリングされた天然に存在しないクラスター化して規則的な配置の短い回文配列リピート（ＣＲＩＳＰＲ）－Ｃａｓシステムを提供し、ここで、ＣＲＩＳＰＲ関連タンパク質は、ＲＮＡガイドに結合し、スペーサー配列に相補的な標的核酸配列を修飾することができる。

【００１０】

一部の態様において、本開示は、ＣＲＩＳＰＲ関連タンパク質が配列番号２４１のアミノ酸配列を含む、ＣＲＩＳＰＲ関連タンパク質又はＣＲＩＳＰＲ関連タンパク質をコードする核酸；及びダイレクトリピート配列と標的核酸へのハイブリダイゼーション能を有するスペーサー配列とをＲＮＡガイドを含む、ＣＬＵＳＴ．０９１９７９の、エンジニアリングされた天然に存在しない、クラスター化して規則的な配置の短い回文配列リピート（ＣＲＩＳＰＲ）－Ｃａｓシステムを提供し、ここで、ＣＲＩＳＰＲ関連タンパク質は、ＲＮＡガイドに結合し、スペーサー配列に相補的な標的核酸配列を修飾することができる。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号４、配列番号１０、配列番号１２、又は配列番号１４に記載のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含む。

【００１１】

本明細書に記載されるシステムのいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、少なくとも１つ（例えば、１つ、２つ、又は３つ）のＲｕｖＣドメイン又は少なくとも１つの分割されたＲｕｖＣドメインを含む。

【００１２】

本明細書に記載されるシステムのいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、以下の配列の１つ以上：（ａ）ＰＸ$_1$Ｘ$_2$Ｘ$_3$Ｘ$_4$Ｆ（配列番号２１６）（ここで、Ｘ$_1$はＬ又はＭ又はＩ又はＣ又はＦであり、Ｘ$_2$はＹ又はＷ又はＦであり、Ｘ$_3$はＫ又はＴ又はＣ又はＲ又はＷ又はＹ又はＨ又はＶであり、Ｘ$_4$はＩ又はＬ又はＭである）；（ｂ）ＲＸ$_1$Ｘ$_2$Ｘ$_3$Ｌ（配列番号２１７）（ここで、Ｘ$_1$はＩ又はＬ又はＭ又はＹ又はＴ又はＦであり、Ｘ$_2$はＲ又はＱ又はＫ又はＥ又はＳ又はＴであり、Ｘ$_3$はＬ又はＩ又はＴ又はＣ又はＭ又はＫである）；（ｃ）ＮＸ$_1$ＹＸ$_2$（配列番号２１８）（ここで、Ｘ$_1$はＩ又はＬ又はＦであり、Ｘ$_2$はＫ又はＲ又はＶ又はＥである）；（ｄ）ＫＸ$_1$Ｘ$_2$Ｘ$_3$ＦＡＸ$_4$Ｘ$_5$ＫＤ（配列番号２１９）（ここで、Ｘ$_1$はＴ又はＩ又はＮ又はＡ又はＳ又はＦ又はＶであり、Ｘ$_2$はＩ又はＶ又はＬ又はＳであり、Ｘ$_3$はＨ又はＳ又はＧ又はＲであり、Ｘ$_4$はＤ又はＳ又はＥであり、Ｘ$_5$はＩ又はＶ又はＭ又はＴ又はＮである）；（ｅ）ＬＸ$_1$ＮＸ$_2$（配列番号２２０）（ここで、Ｘ$_1$はＧ又はＳ又はＣ又はＴであり、Ｘ$_2$はＮ又はＹ又

はＫ又はＳである）；（ｆ）ＰＸ$_1$Ｘ$_2$Ｘ$_3$Ｘ$_4$ＳＱＸ$_5$ＤＳ（配列番号２２１）（ここで、Ｘ$_1$はＳ又はＰ又はＡであり、Ｘ$_2$はＹ又はＳ又はＡ又はＰ又はＥ又はＹ又はＱ又はＮであり、Ｘ$_3$はＦ又はＹ又はＨであり、Ｘ$_4$はＴ又はＳであり、Ｘ$_5$はＭ又はＴ又はＩである）；（ｇ）ＫＸ$_1$Ｘ$_2$ＶＲＸ$_3$Ｘ$_4$ＱＥＸ$_5$Ｈ（配列番号２２２）（ここで、Ｘ$_1$はＮ又はＫ又はＷ又はＲ又はＥ又はＴ又はＹであり、Ｘ$_2$はＭ又はＲ又はＬ又はＳ又はＫ又はＶ又はＥ又はＴ又はＩ又はＤであり、Ｘ$_3$はＬ又はＲ又はＨ又はＰ又はＴ又はＫ又はＰのＱ又はＳ又はＡであり、Ｘ$_4$はＧ又はＱ又はＮ又はＲ又はＫ又はＥ又はＩ又はＴ又はＳ又はＣであり、Ｘ$_5$はＲ又はＷ又はＹ又はＫ又はＴ又はＦ又はＳ又はＱである）；及び（ｈ）Ｘ$_1$ＮＧＸ$_2$Ｘ$_3$Ｘ$_4$ＤＸ$_5$ＮＸ$_6$Ｘ$_7$Ｘ$_8$Ｎ（配列番号２２３）（ここで、Ｘ$_1$はＩ又はＫ又はＶ又はＬであり、Ｘ$_2$はＬ又はＭであり、Ｘ$_3$はＮ又はＨ又はＰであり、Ｘ$_4$はＡ又はＳ又はＣであり、Ｘ$_5$はＶ又はＹ又はＩ又はＦ又はＴ又はＮであり、Ｘ$_6$はＡ又はＳであり、Ｘ$_7$はＳ又はＡ又はＰであり、Ｘ$_8$はＭ又はＣ又はＬ又はＲ又はＮ又はＳ又はＫ又はＬである）を含む。本明細書に記載されるシステムのいずれかの一部の実施形態において、配列番号２１６の配列はＮ末端配列である。本明細書に記載されるシステムのいずれかの一部の実施形態において、配列番号２１９の配列はＣ末端配列である。本明細書に記載されるシステムのいずれかの一部の実施形態において、配列番号２２０の配列はＣ末端配列である。本明細書に記載されるシステムのいずれかの一部の実施形態において、配列番号２２１の配列はＣ末端配列である。本明細書に記載されるシステムのいずれかの一部の実施形態において、配列番号２２２の配列はＣ末端配列である。本明細書に記載されるシステムのいずれかの一部の実施形態において、配列番号２２３の配列はＣ末端配列である。

【００１３】

本明細書に記載されるシステムのいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、以下の配列の１つ以上：（ａ）ＥＣＰＩＴＫＤＶＩＮＥＹＫ（配列番号２９０）；（ｂ）ＮＬＴＳＩＴＩＧ（配列番号２３１）；（ｃ）ＮＹＲＴＫＩＲＴＬＮ（配列番号２３２）；（ｄ）ＩＳＹＩＥＮＶＥＮ（配列番号２３３）；（ｅ）ＥＬＬＳＶＥＱＬＫ（配列番号２３４）；（ｆ）ＨＩＮＳＭＴＩＮＩＱＤＦＫＩＥ（配列番号２３５）；（ｇ）ＫＥＮＳＬＧＦＩＬ（配列番号２３６）；（ｈ）ＧＮＲＱＩＫＫＧ（配列番号２３７）；（ｉ）ＤＶＮＦＫＨＡ（配列番号２３８）；（ｊ）ＧＹＩＮＬＹＫＹＬＬＥＨ（配列番号２３９）；（ｋ）ＫＥＱＶＬＳＫＬＬＹ（配列番号２４０）；（ｌ）ＥＹＩＹＶＳＣＶＮＫＬＲＡＫＹＶＳＹＦＩＬＫＥＫＹＹＥＫＱＫＥＹＤＩＥＭＧＦ（配列番号２４１）；（ｍ）ＤＤＳＴＥＳＫＥＳＭＤＫＲＲ（配列番号２４２）；（ｎ）ＮＶＱＱＤＩＮＧＣＬＫＮＩＩＮＹ（配列番号２４３）；（ｏ）ＡＬＥＮＬＥＮＳＮＦＥＫ（配列番号２４４）；（ｐ）ＱＶＬＰＴＩＫＳＬＬ（配列番号２４５）；（ｑ）ＹＨＫＬＥＮＱＮ（配列番号２４６）；（ｒ）ＡＳＤＫＶＫＥＹＩＥ（配列番号２４７）；（ｓ）ＴＮＥＮＮＥＩＶＤＡＫＹＴ（配列番号２４８）；（ｔ）ＡＮＦＦＮＬＭＭＫＳＬＨＦＡＳ（配列番号２４９）；（ｕ）ＬＬＳＮＮＧＫＴＱＩＡＬＶＰＳＥ（配列番号２５０）；（ｖ）ＨＩＮＧＬＮＡＤＦＮＡＡＮＮＩＫＹＩ（配列番号２５１）、又は前述のいずれかに対して１つ、２つ、若しくは３つ以下の配列差異（例えば、置換）を有する配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号４と少なくとも７０％同一の配列を有する。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号１０と少なくとも７０％同一の配列を有する。

【００１４】

本明細書に記載されるシステムのいずれかの一部の実施形態において、ダイレクトリピート配列は、配列番号５７～９０、配列番号１１８～１５１、又は配列番号２１３のいずれか１つに記載のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。本明細書に記載されるシステムのいずれかの一部の実施形態において、ダイレクトリピート配列は、配列番号５７～９０、配列番号１１８～１５１、又は配列番号

２１３のいずれか１つに記載のヌクレオチド配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。

【００１５】

本明細書に記載されるシステムのいずれかの一部の実施形態において、ダイレクトリピート配列は、以下の配列の１つ以上：（ａ）$X_1X_2TX_3X_4X_5X_6X_7X_8$（配列番号２２４）（ここで、$X_1$はＡ又はＣ又はＧであり、$X_2$はＴ又はＣ又はＡであり、$X_3$はＴ又はＧ又はＡであり、$X_4$はＴ又はＧであり、$X_5$はＴ又はＧ又はＡであり、$X_6$はＧ又はＴ又はＡであり、$X_7$はＴ又はＧ又はＡであり、$X_8$はＡ又はＧ又はＴである）（例えば、ＡＴＴＧＴＴＧＤＡ（配列番号２２５））；（ｂ）$X_1X_2X_3X_4X_5X_6X_7X_8X_9$（配列番号２２６）（ここで、$X_1$はＴ又はＣ又はＡであり、$X_2$はＴ又はＡ又はＧであり、$X_3$はＴ又はＣ又はＡであり、$X_4$はＴ又はＡであり、$X_5$はＴ又はＡ又はＧであり、$X_6$はＴ又はＡであり、$X_7$はＡ又はＴであり、$X_8$はＡ又はＧ又はＣ又はＴであり、$X_9$はＧ又はＡ又はＣである）（例えば、ＴＴＴＴＷＴＡＲＧ（配列番号２２７））；及び（ｃ）$X_1X_2X_3$ＡＣ（配列番号２２８）（ここで、X_1はＡ又はＣ又はＧであり、X_2はＣ又はＡであり、X_3はＡ又はＣである）（例えば、ＡＣＡＡＣ（配列番号２２９））を含む。本明細書に記載されるシステムのいずれかの一部の実施形態において、配列番号２２４はダイレクトリピートの５’末端に近接している。本明細書に記載されるシステムのいずれかの一部の実施形態において、配列番号２２８はダイレクトリピートの３’末端に近接している。

【００１６】

本明細書に記載されるシステムのいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質はプロトスペーサー隣接モチーフ（ＰＡＭ）の認識能を有し、ここで、ＰＡＭ配列は核酸配列を含み、これは５’－ＮＴＴＮ－３’、５’－ＮＴＴＲ－３’、５’－ＲＴＴＲ－３’、５’－ＴＮＮＴ－３’、５’－ＴＮＲＴ－３’、５’－ＴＳＲＴ－３’、５’－ＴＧＲＴ－３’、５’－ＴＮＲＹ－３’、５’－ＴＴＮＲ－３’、５’－ＴＴＹＲ－３’、５’－ＴＴＴＲ－３’、５’－ＴＴＣＶ－３’、５’－ＤＴＹＲ－３’、５’－ＷＴＴＲ－３’、５’－ＮＮＲ－３’、５’－ＮＹＲ－３’、５’－ＹＹＲ－３’、５’－ＴＹＲ－３’、５’－ＴＴＮ－３’、５’－ＴＴＲ－３’、５’－ＣＮＴ－３’、５’－ＮＧＧ－３’、５’－ＢＧＧ－３’、又は５’－Ｒ－３’として記載される核酸配列を含み、「Ｎ」は任意のヌクレオチドであり、「Ｂ」はＣ又はＧ又はＴであり、「Ｄ」はＡ又はＧ又はＴであり、「Ｒ」はＡ又はＧであり、「Ｓ」はＧ又はＣであり、「Ｖ」はＡ又はＣ又はＧであり、「Ｗ」はＡ又はＴであり、「Ｙ」はＣ又はＴである。

【００１７】

本明細書に記載されるシステムのいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号１に記載のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、ダイレクトリピート配列は、配列番号５７に記載のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。本明細書に記載されるシステムのいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号１に記載のアミノ酸配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、ダイレクトリピート配列は、配列番号５７に記載のヌクレオチド配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。本明細書に記載されるシステムのいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質はプロトスペーサー隣接モチーフ（ＰＡＭ）配列の認識能を有し、ここで、ＰＡＭ配列は、５’－ＴＮＮＴ－３’又は５’－ＴＮＲＴ－３’として記載される核酸配列を含み、「Ｎ」は任意のヌクレオチドであり、「Ｒ」はＡ又はＧである。

【００１８】
本明細書に記載されるシステムのいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号４に記載のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、ダイレクトリピート配列は、配列番号６０に記載のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。本明細書に記載されるシステムのいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号４に記載のアミノ酸配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、ダイレクトリピート配列は、配列番号６０に記載のヌクレオチド配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。本明細書に記載されるシステムのいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質はプロトスペーサー隣接モチーフ（ＰＡＭ）配列の認識能を有し、ここで、ＰＡＭ配列は、５’－ＮＴＴＮ－３’、５’－ＮＴＴＲ－３’（例えば、５’－ＴＴＴＧ－３’）、又は５’－ＮＮＲ－３’として記載される核酸配列を含み、「Ｎ」は任意のヌクレオチドであり、「Ｒ」はＡ又はＧである。

【００１９】
本明細書に記載されるシステムのいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号１０に記載のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、ダイレクトリピート配列は、配列番号６２又は配列番号２１３に記載のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。本明細書に記載されるシステムのいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号１０に記載のアミノ酸配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、ダイレクトリピート配列は、配列番号６２又は配列番号２１３に記載のヌクレオチド配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。本明細書に記載されるシステムのいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質はプロトスペーサー隣接モチーフ（ＰＡＭ）配列の認識能を有し、ここで、ＰＡＭ配列は、５’－ＮＴＴＮ－３’又は５’－ＲＴＴＲ－３’（例えば、５’－ＡＴＴＧ－３’又は５’－ＧＴＴＡ－３’）として記載される核酸配列を含み、「Ｎ」は任意のヌクレオチドであり、「Ｒ」はＡ又はＧである。

【００２０】
本明細書に記載されるシステムのいずれかの一部の実施形態において、ＲＮＡガイドのスペーサー配列は約１５ヌクレオチド～約５５ヌクレオチドを含む。本明細書に記載されるシステムのいずれかの一部の実施形態において、ＲＮＡガイドのスペーサー配列は２０～４５ヌクレオチドを含む。

【００２１】
本明細書に記載されるシステムのいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は触媒残基（例えば、アスパラギン酸又はグルタミン酸）を含む。本明細書に記載されるシステムのいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、標的核酸を切断する。本明細書に記載されるシステムのいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、ペプチドタグ、蛍光タンパク質、塩基編集ドメイン、ＤＮＡメチル化ドメイン、ヒストン残基修飾ドメイン、局在化因子、転写修飾因

子、光ゲート制御因子、化学誘導性因子、又はクロマチン可視化因子を更に含む。
【００２２】
本明細書に記載されるシステムのいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質をコードする核酸は、細胞（例えば、真核細胞、例えば、哺乳動物細胞、例えば、ヒト細胞）での発現にコドン最適化される。本明細書に記載されるシステムのいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質をコードする核酸は、プロモーターに作動可能に連結されている。本明細書に記載されるシステムのいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質をコードする核酸は、ベクター内にある。一部の実施形態において、ベクターは、レトロウイルスベクター、レンチウイルスベクター、ファージベクター、アデノウイルスベクター、アデノ随伴ベクター、又は単純ヘルペスベクターを含む。
【００２３】
本明細書に記載されるシステムのいずれかの一部の実施形態において、標的核酸はＤＮＡ分子である。本明細書に記載されるシステムのいずれかの一部の実施形態において、標的核酸はＰＡＭ配列を含む。
【００２４】
本明細書に記載されるシステムのいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、非特異的ヌクレアーゼ活性を有する。
【００２５】
本明細書に記載されるシステムのいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質及びＲＮＡガイドによる標的核酸の認識により、標的核酸の修飾が生じる。本明細書に記載されるシステムのいずれかの一部の実施形態において、標的核酸の修飾は、二本鎖切断イベントである。本明細書に記載されるシステムのいずれかの一部の実施形態において、標的核酸の修飾は、一本鎖切断イベントである。本明細書に記載されるシステムのいずれかの一部の実施形態において、標的核酸の修飾により、挿入イベントが生じる。本明細書に記載されるシステムのいずれかの一部の実施形態において、標的核酸の修飾により、欠失イベントが生じる。本明細書に記載されるシステムのいずれかの一部の実施形態において、標的核酸の修飾により、細胞毒性又は細胞死が生じる。
【００２６】
本明細書に記載されるシステムのいずれかの一部の実施形態において、システムはドナー鋳型核酸を更に含む。本明細書に記載されるシステムのいずれかの一部の実施形態において、ドナー鋳型核酸はＤＮＡ分子である。本明細書に記載されるシステムのいずれかの一部の実施形態において、ドナー鋳型核酸はＲＮＡ分子である。
【００２７】
本明細書に記載されるシステムのいずれかの一部の実施形態において、ＲＮＡガイドは、任意選択でｔｒａｃｒＲＮＡ及び／又はモジュレーターＲＮＡを含む。本明細書に記載されるシステムのいずれかの一部の実施形態において、システムはｔｒａｃｒＲＮＡを更に含む。本明細書に記載されるシステムのいずれかの一部の実施形態において、システムはｔｒａｃｒＲＮＡを含まない。本明細書に記載されるシステムのいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は自己プロセシングである。本明細書に記載されるシステムのいずれかの一部の実施形態において、システムはモジュレーターＲＮＡを更に含む。
【００２８】
本明細書に記載されるシステムのいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号１のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ｔｒａｃｒＲＮＡ配列は、配列番号１５２、配列番号１５３、又は配列番号１５４のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％

、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。

【００２９】

本明細書に記載されるシステムのいずれかの一部の実施形態において、システムは、ナノ粒子、リポソーム、エキソソーム、微小胞、又は遺伝子銃を含む送達組成物中に存在する。

【００３０】

本明細書に記載されるシステムのいずれかの一部の実施形態において、システムは細胞内にある。一部の実施形態において、細胞は真核細胞である。一部の実施形態において、細胞は哺乳動物細胞である。一部の実施形態において、細胞はヒト細胞である。一部の実施形態において、細胞は原核細胞である。

【００３１】

別の態様において、本開示は細胞を提供し、ここで、細胞は、配列番号１～５６のいずれか１つに記載のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含むＣＲＩＳＰＲ関連タンパク質；及びダイレクトリピート配列と標的核酸へのハイブリダイゼーション能を有するスペーサー配列とを含むＲＮＡガイドを含む。別の態様において、本開示は細胞を提供し、ここで、細胞は、ＣＲＩＳＰＲ関連タンパク質が配列番号１～５６のいずれか１つに記載のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含む、ＣＲＩＳＰＲ関連タンパク質又はＣＲＩＳＰＲ関連タンパク質をコードする核酸；及びダイレクトリピート配列と標的核酸へのハイブリダイゼーション能を有するスペーサー配列とを含むＲＮＡガイド、又はＲＮＡガイドをコードする核酸を含む。

【００３２】

本明細書に記載される細胞のいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、少なくとも１つ（例えば、１つ、２つ、又は３つ）のＲｕｖＣドメイン又は少なくとも１つの分割されたＲｕｖＣドメインを含む。

【００３３】

本明細書に記載される細胞のいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、以下の配列の１つ以上：（ａ）ＰＸ$_1$Ｘ$_2$Ｘ$_3$Ｘ$_4$Ｆ（配列番号２１６）（ここで、Ｘ$_1$はＬ又はＭ又はＩ又はＣ又はＦであり、Ｘ$_2$はＹ又はＷ又はＦであり、Ｘ$_3$はＫ又はＴ又はＣ又はＲ又はＷ又はＹ又はＨ又はＶであり、Ｘ$_4$はＩ又はＬ又はＭである）；（ｂ）ＲＸ$_1$Ｘ$_2$Ｘ$_3$Ｌ（配列番号２１７）（ここで、Ｘ$_1$はＩ又はＬ又はＭ又はＹ又はＴ又はＦであり、Ｘ$_2$はＲ又はＱ又はＫ又はＥ又はＳ又はＴであり、Ｘ$_3$はＬ又はＩ又はＴ又はＣ又はＭ又はＫである）；（ｃ）ＮＸ$_1$ＹＸ$_2$（配列番号２１８）（ここで、Ｘ$_1$はＩ又はＬ又はＦであり、Ｘ$_2$はＫ又はＲ又はＶ又はＥである）；（ｄ）ＫＸ$_1$Ｘ$_2$Ｘ$_3$ＦＡＸ$_4$Ｘ$_5$ＫＤ（配列番号２１９）（ここで、Ｘ$_1$はＴ又はＩ又はＮ又はＡ又はＳ又はＦ又はＶであり、Ｘ$_2$はＩ又はＶ又はＬ又はＳであり、Ｘ$_3$はＨ又はＳ又はＧ又はＲであり、Ｘ$_4$はＤ又はＳ又はＥであり、Ｘ$_5$はＩ又はＶ又はＭ又はＴ又はＮである）；（ｅ）ＬＸ$_1$ＮＸ$_2$（配列番号２２０）（ここで、Ｘ$_1$はＧ又はＳ又はＣ又はＴであり、Ｘ$_2$はＮ又はＹ又はＫ又はＳである）；（ｆ）ＰＸ$_1$Ｘ$_2$Ｘ$_3$Ｘ$_4$ＳＱＸ$_5$ＤＳ（配列番号２２１）（ここで、Ｘ$_1$はＳ又はＰ又はＡであり、Ｘ$_2$はＹ又はＳ又はＡ又はＰ又はＥ又はＹ又はＱ又はＮであり、Ｘ$_3$はＦ又はＹ又はＨであり、Ｘ$_4$はＴ又はＳであり、Ｘ$_5$はＭ又はＴ又はＩである）；（ｇ）ＫＸ$_1$Ｘ$_2$ＶＲＸ$_3$Ｘ$_4$ＱＥＸ$_5$Ｈ（配列番号２２２）（ここで、Ｘ$_1$はＮ又はＫ又はＷ又はＲ又はＥ又はＴ又はＹであり、Ｘ$_2$はＭ又はＲ又はＬ又はＳ又はＫ又はＶ又はＥ又はＴ又はＩ又はＤであり、Ｘ$_3$はＬ又はＲ又はＨ又はＰ又はＴ又はＫ又はＰのＱ又はＳ又はＡであり、Ｘ$_4$はＧ又はＱ又はＮ又はＲ又はＫ又はＥ又はＩ又はＴ又はＳ又はＣであり、Ｘ$_5$はＲ又はＷ又はＹ又はＫ又はＴ又はＦ又はＳ又はＱである）；及び（ｈ）Ｘ$_1$

ＮＧＸ$_2$Ｘ$_3$Ｘ$_4$ＤＸ$_5$ＮＸ$_6$Ｘ$_7$Ｘ$_8$Ｎ（配列番号２２３）（ここで、Ｘ$_1$はＩ又はＫ又はＶ又はＬであり、Ｘ$_2$はＬ又はＭであり、Ｘ$_3$はＮ又はＨ又はＰであり、Ｘ$_4$はＡ又はＳ又はＣであり、Ｘ$_5$はＶ又はＹ又はＩ又はＦ又はＴ又はＮであり、Ｘ$_6$はＡ又はＳであり、Ｘ$_7$はＳ又はＡ又はＰであり、Ｘ$_8$はＭ又はＣ又はＬ又はＲ又はＮ又はＳ又はＫ又はＬである）を含む。本明細書に記載される細胞のいずれかの一部の実施形態において、配列番号２１６の配列はＮ末端配列である。本明細書に記載される細胞のいずれかの一部の実施形態において、配列番号２１９の配列はＣ末端配列である。本明細書に記載される細胞のいずれかの一部の実施形態において、配列番号２２０の配列はＣ末端配列である。本明細書に記載される細胞のいずれかの一部の実施形態において、配列番号２２１の配列はＣ末端配列である。本明細書に記載される細胞のいずれかの一部の実施形態において、配列番号２２２の配列はＣ末端配列である。本明細書に記載される細胞のいずれかの一部の実施形態において、配列番号２２３の配列はＣ末端配列である。

【００３４】

本明細書に記載される細胞のいずれかの一部の実施形態において、ダイレクトリピート配列は、配列番号５７～９０、配列番号１１８～１５１、又は配列番号２１３のいずれか１つに記載のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。本明細書に記載される細胞のいずれかの一部の実施形態において、ダイレクトリピート配列は、配列番号５７～９０、配列番号１１８～１５１、又は配列番号２１３のいずれか１つに記載のヌクレオチド配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。

【００３５】

本明細書に記載される細胞のいずれかの一部の実施形態において、ダイレクトリピート配列は、以下の配列の１つ以上：（ａ）Ｘ$_1$Ｘ$_2$ＴＸ$_3$Ｘ$_4$Ｘ$_5$Ｘ$_6$Ｘ$_7$Ｘ$_8$（配列番号２２４）（ここで、Ｘ$_1$はＡ又はＣ又はＧであり、Ｘ$_2$はＴ又はＣ又はＡであり、Ｘ$_3$はＴ又はＧ又はＡであり、Ｘ$_4$はＴ又はＧであり、Ｘ$_5$はＴ又はＧ又はＡであり、Ｘ$_6$はＧ又はＴ又はＡであり、Ｘ$_7$はＴ又はＧ又はＡであり、Ｘ$_8$はＡ又はＧ又はＴである）（例えば、ＡＴＴＧＴＴＧＤＡ（配列番号２２５））；（ｂ）Ｘ$_1$Ｘ$_2$Ｘ$_3$Ｘ$_4$Ｘ$_5$Ｘ$_6$Ｘ$_7$Ｘ$_8$Ｘ$_9$（配列番号２２６）（ここで、Ｘ$_1$はＴ又はＣ又はＡであり、Ｘ$_2$はＴ又はＡ又はＧであり、Ｘ$_3$はＴ又はＣ又はＡであり、Ｘ$_4$はＴ又はＡであり、Ｘ$_5$はＴ又はＡ又はＧであり、Ｘ$_6$はＴ又はＡであり、Ｘ$_7$はＡ又はＴであり、Ｘ$_8$はＡ又はＧ又はＣ又はＴであり、Ｘ$_9$はＧ又はＡ又はＣである）（例えば、ＴＴＴＴＷＴＡＲＧ（配列番号２２７））；及び（ｃ）Ｘ$_1$Ｘ$_2$Ｘ$_3$ＡＣ（配列番号２２８）（ここで、Ｘ$_1$はＡ又はＣ又はＧであり、Ｘ$_2$はＣ又はＡであり、Ｘ$_3$はＡ又はＣである）（例えば、ＡＣＡＡＣ（配列番号２２９））を含む。本明細書に記載される細胞のいずれかの一部の実施形態において、配列番号２２４はダイレクトリピートの５’末端に近接している。本明細書に記載される細胞のいずれかの一部の実施形態において、配列番号２２８はダイレクトリピートの３’末端に近接している。

【００３６】

本明細書に記載される細胞のいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号１に記載のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、ダイレクトリピート配列は、配列番号５７に記載のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。本明細書に記載される細胞のいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号１に記載のアミノ酸配列と少なくとも９５％（例えば、９５％、９６％、９７

％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、ダイレクトリピート配列は、配列番号５７に記載のヌクレオチド配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。本明細書に記載される細胞のいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質はプロトスペーサー隣接モチーフ（ＰＡＭ）配列の認識能を有し、ここで、ＰＡＭ配列は、５’－ＴＮＮＴ－３’又は５’－ＴＮＲＴ－３’として記載される核酸配列を含み、「Ｎ」は任意のヌクレオチドであり、「Ｒ」はＡ又はＧである。

【００３７】

本明細書に記載される細胞のいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号４に記載のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、ダイレクトリピート配列は、配列番号６０に記載のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。本明細書に記載される細胞のいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号４に記載のアミノ酸配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、ダイレクトリピート配列は、配列番号６０に記載のヌクレオチド配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。本明細書に記載される細胞のいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質はプロトスペーサー隣接モチーフ（ＰＡＭ）配列の認識能を有し、ここで、ＰＡＭ配列は、５’－ＮＴＴＮ－３’、５’－ＮＴＴＲ－３’（例えば、５’－ＴＴＴＧ－３’）、又は５’－ＮＮＲ－３’として記載される核酸配列を含み、「Ｎ」は任意のヌクレオチドであり、「Ｒ」はＡ又はＧである。

【００３８】

本明細書に記載される細胞のいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号１０に記載のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、ダイレクトリピート配列は、配列番号６２又は配列番号２１３に記載のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。本明細書に記載される細胞のいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号１０に記載のアミノ酸配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、ダイレクトリピート配列は、配列番号６２又は配列番号２１３に記載のヌクレオチド配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。本明細書に記載される細胞のいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質はプロトスペーサー隣接モチーフ（ＰＡＭ）配列の認識能を有し、ここで、ＰＡＭ配列は、５’－ＮＴＴＮ－３’又は５’－ＲＴＴＲ－３’（例えば、５’－ＡＴＴＧ－３’又は５’－ＧＴＴＡ－３’）として記載される核酸配列を含み、「Ｎ」は任意のヌクレオチドであり、「Ｒ」はＡ又はＧである。

【００３９】

本明細書に記載される細胞のいずれかの一部の実施形態において、スペーサー配列は約１５ヌクレオチド～約５５ヌクレオチドを含む。本明細書に記載される細胞のいずれかの一部の実施形態において、スペーサー配列は２０～４５ヌクレオチドを含む。

【００４０】

本明細書に記載される細胞のいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は触媒残基（例えば、アスパラギン酸又はグルタミン酸）を含む。本明細書に記載される細胞のいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、標的核酸を切断する。本明細書に記載される細胞のいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、ペプチドタグ、蛍光タンパク質、塩基編集ドメイン、ＤＮＡメチル化ドメイン、ヒストン残基修飾ドメイン、局在化因子、転写修飾因子、光ゲート制御因子、化学誘導性因子、又はクロマチン可視化因子を更に含む。

【００４１】

本明細書に記載される細胞のいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質をコードする核酸は、細胞（例えば、真核細胞、例えば、哺乳動物細胞、例えば、ヒト細胞）での発現にコドン最適化される。本明細書に記載される細胞のいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質をコードする核酸は、プロモーターに作動可能に連結されている。本明細書に記載される細胞のいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質をコードする核酸は、ベクター内にある。一部の実施形態において、ベクターは、レトロウイルスベクター、レンチウイルスベクター、ファージベクター、アデノウイルスベクター、アデノ随伴ベクター、又は単純ヘルペスベクターを含む。

【００４２】

本明細書に記載される細胞のいずれかの一部の実施形態において、ＲＮＡガイドは、任意選択でｔｒａｃｒＲＮＡ及び／又はモジュレーターＲＮＡを含む。本明細書に記載される細胞のいずれかの一部の実施形態において、細胞はｔｒａｃｒＲＮＡを更に含む。本明細書に記載される細胞のいずれかの一部の実施形態において、細胞はｔｒａｃｒＲＮＡを含まない。本明細書に記載される細胞のいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は自己プロセシングである。本明細書に記載される細胞のいずれかの一部の実施形態において、細胞はモジュレーターＲＮＡを更に含む。

【００４３】

本明細書に記載される細胞のいずれかの一部の実施形態において、細胞は真核細胞である。本明細書に記載される細胞のいずれかの一部の実施形態において、細胞は哺乳動物細胞である。本明細書に記載される細胞のいずれかの一部の実施形態において、細胞はヒト細胞である。本明細書に記載される細胞のいずれかの一部の実施形態において、細胞は原核細胞である。

【００４４】

本明細書に記載される細胞のいずれかの一部の実施形態において、標的核酸はＤＮＡ分子である。本明細書に記載される細胞のいずれかの一部の実施形態において、標的核酸はＰＡＭ配列を含む。

【００４５】

本明細書に記載される細胞のいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、非特異的ヌクレアーゼ活性を有する。

【００４６】

本明細書に記載される細胞のいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質及びＲＮＡガイドによる標的核酸の認識により、標的核酸の修飾が生じる。本明細書に記載される細胞のいずれかの一部の実施形態において、標的核酸の修飾は、二本鎖切断イベントである。本明細書に記載される細胞のいずれかの一部の実施形態において、標的核酸の修飾は、一本鎖切断イベントである。本明細書に記載される細胞のいずれかの一部の実施形態において、標的核酸の修飾により、挿入イベントが生じる。本明細書に記載される細胞のいずれかの一部の実施形態において、標的核酸の修飾により、欠失イベントが生じる。本明細書に記載される細胞のいずれかの一部の実施形態において、標的核酸の修飾により、細胞毒性又は細胞死が生じる。

【００４７】

別の態様において、本開示は、（ａ）システムを提供すること；及び（ｂ）システムを

細胞に送達することを含む、本明細書に記載されるシステムを細胞内の標的核酸に結合する方法を提供し、ここで、細胞は標的核酸を含み、ＣＲＩＳＰＲ関連タンパク質はＲＮＡガイドに結合し、スペーサー配列は標的核酸に結合する。一部の実施形態において、細胞は真核細胞、例えば、哺乳動物細胞、例えば、ヒト細胞である。

【００４８】

別の態様において、本開示は、標的核酸を修飾する方法であって、配列番号１～５６のいずれか１つに記載のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含むＣＲＩＳＰＲ関連タンパク質；及びダイレクトリピート配列と標的核酸へのハイブリダイゼーション能を有するスペーサー配列とを含むＲＮＡガイドを含む、エンジニアリングされた天然に存在しないＣＲＩＳＰＲ－Ｃａｓシステムを、標的核酸に送達することを含む方法を提供し、ここで、ＣＲＩＳＰＲ関連タンパク質はＲＮＡガイドへの結合能を有し、ＣＲＩＳＰＲ関連タンパク質及びＲＮＡガイドによる標的核酸の認識により、標的核酸の修飾が生じる。別の態様において、本開示は、標的核酸を修飾する方法であって、ＣＲＩＳＰＲ関連タンパク質が配列番号１～５６のいずれか１つに記載のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含む、ＣＲＩＳＰＲ関連タンパク質又はＣＲＩＳＰＲ関連タンパク質をコードする核酸；及びダイレクトリピート配列と標的核酸へのハイブリダイゼーション能を有するスペーサー配列とを含むＲＮＡガイドを含む、エンジニアリングされた天然に存在しないＣＲＩＳＰＲ－Ｃａｓシステムを、標的核酸に送達することを含む方法を提供し、ここで、ＣＲＩＳＰＲ関連タンパク質はＲＮＡガイドへの結合能を有し、ＣＲＩＳＰＲ関連タンパク質及びＲＮＡガイドによる標的核酸の認識により、標的核酸の修飾が生じる。

【００４９】

本明細書に記載される方法のいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、以下の配列の１つ以上：（ａ）ＰＸ$_1$Ｘ$_2$Ｘ$_3$Ｘ$_4$Ｆ（配列番号２１６）（ここで、Ｘ$_1$はＬ又はＭ又はＩ又はＣ又はＦであり、Ｘ$_2$はＹ又はＷ又はＦであり、Ｘ$_3$はＫ又はＴ又はＣ又はＲ又はＷ又はＹ又はＨ又はＶであり、Ｘ$_4$はＩ又はＬ又はＭである）；（ｂ）ＲＸ$_1$Ｘ$_2$Ｘ$_3$Ｌ（配列番号２１７）（ここで、Ｘ$_1$はＩ又はＬ又はＭ又はＹ又はＴ又はＦであり、Ｘ$_2$はＲ又はＱ又はＫ又はＥ又はＳ又はＴであり、Ｘ$_3$はＬ又はＩ又はＴ又はＣ又はＭ又はＫである）；（ｃ）ＮＸ$_1$ＹＸ$_2$（配列番号２１８）（ここで、Ｘ$_1$はＩ又はＬ又はＦであり、Ｘ$_2$はＫ又はＲ又はＶ又はＥである）；（ｄ）ＫＸ$_1$Ｘ$_2$Ｘ$_3$ＦＡＸ$_4$Ｘ$_5$ＫＤ（配列番号２１９）（ここで、Ｘ$_1$はＴ又はＩ又はＮ又はＡ又はＳ又はＦ又はＶであり、Ｘ$_2$はＩ又はＶ又はＬ又はＳであり、Ｘ$_3$はＨ又はＳ又はＧ又はＲであり、Ｘ$_4$はＤ又はＳ又はＥであり、Ｘ$_5$はＩ又はＶ又はＭ又はＴ又はＮである）；（ｅ）ＬＸ$_1$ＮＸ$_2$（配列番号２２０）（ここで、Ｘ$_1$はＧ又はＳ又はＣ又はＴであり、Ｘ$_2$はＮ又はＹ又はＫ又はＳである）；（ｆ）ＰＸ$_1$Ｘ$_2$Ｘ$_3$Ｘ$_4$ＳＱＸ$_5$ＤＳ（配列番号２２１）（ここで、Ｘ$_1$はＳ又はＰ又はＡであり、Ｘ$_2$はＹ又はＳ又はＡ又はＰ又はＥ又はＹ又はＱ又はＮであり、Ｘ$_3$はＦ又はＹ又はＨであり、Ｘ$_4$はＴ又はＳであり、Ｘ$_5$はＭ又はＴ又はＩである）；（ｇ）ＫＸ$_1$Ｘ$_2$ＶＲＸ$_3$Ｘ$_4$ＱＥＸ$_5$Ｈ（配列番号２２２）（ここで、Ｘ$_1$はＮ又はＫ又はＷ又はＲ又はＥ又はＴ又はＹであり、Ｘ$_2$はＭ又はＲ又はＬ又はＳ又はＫ又はＶ又はＥ又はＴ又はＩ又はＤであり、Ｘ$_3$はＬ又はＲ又はＨ又はＰ又はＴ又はＫ又はＰのＱ又はＳ又はＡであり、Ｘ$_4$はＧ又はＱ又はＮ又はＲ又はＫ又はＥ又はＩ又はＴ又はＳ又はＣであり、Ｘ$_5$はＲ又はＷ又はＹ又はＫ又はＴ又はＦ又はＳ又はＱである）；及び（ｈ）Ｘ$_1$ＮＧＸ$_2$Ｘ$_3$Ｘ$_4$ＤＸ$_5$ＮＸ$_6$Ｘ$_7$Ｘ$_8$Ｎ（配列番号２２３）（ここで、Ｘ$_1$はＩ又はＫ又はＶ又はＬであり、Ｘ$_2$はＬ又はＭであり、Ｘ$_3$はＮ又はＨ又はＰであり、Ｘ$_4$はＡ又はＳ又はＣであり、Ｘ$_5$はＶ又はＹ又はＩ又はＦ又はＴ又はＮであり、Ｘ$_6$はＡ又はＳであり、Ｘ$_7$はＳ又はＡ又はＰであり、Ｘ$_8$はＭ又はＣ又はＬ又はＲ又はＮ又はＳ又はＫ又はＬ

である）を含む。本明細書に記載される方法のいずれかの一部の実施形態において、配列番号２１６の配列はＮ末端配列である。本明細書に記載される方法のいずれかの一部の実施形態において、配列番号２１９の配列はＣ末端配列である。本明細書に記載される方法のいずれかの一部の実施形態において、配列番号２２０の配列はＣ末端配列である。本明細書に記載される方法のいずれかの一部の実施形態において、配列番号２２１の配列はＣ末端配列である。本明細書に記載される方法のいずれかの一部の実施形態において、配列番号２２２の配列はＣ末端配列である。本明細書に記載される方法のいずれかの一部の実施形態において、配列番号２２３の配列はＣ末端配列である。

【００５０】

本明細書に記載される方法のいずれかの一部の実施形態において、ダイレクトリピート配列は、配列番号５７～９０、配列番号１１８～１５１、又は配列番号２１３のいずれか１つに記載のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。本明細書に記載される方法のいずれかの一部の実施形態において、ダイレクトリピート配列は、配列番号５７～９０、配列番号１１８～１５１、又は配列番号２１３のいずれか１つに記載のヌクレオチド配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。

【００５１】

本明細書に記載される方法のいずれかの一部の実施形態において、ダイレクトリピート配列は、以下の配列の１つ以上：（ａ）$X_1X_2TX_3X_4X_5X_6X_7X_8$（配列番号２２４）（ここで、$X_1$はＡ又はＣ又はＧであり、$X_2$はＴ又はＣ又はＡであり、$X_3$はＴ又はＧ又はＡであり、$X_4$はＴ又はＧであり、$X_5$はＴ又はＧ又はＡであり、$X_6$はＧ又はＴ又はＡであり、$X_7$はＴ又はＧ又はＡであり、$X_8$はＡ又はＧ又はＴである）（例えば、ＡＴＴＧＴＴＧＤＡ（配列番号２２５））；（ｂ）$X_1X_2X_3X_4X_5X_6X_7X_8X_9$（配列番号２２６）（ここで、$X_1$はＴ又はＣ又はＡであり、$X_2$はＴ又はＡ又はＧであり、$X_3$はＴ又はＣ又はＡであり、$X_4$はＴ又はＡであり、$X_5$はＴ又はＡ又はＧであり、$X_6$はＴ又はＡであり、$X_7$はＡ又はＴであり、$X_8$はＡ又はＧ又はＣ又はＴであり、$X_9$はＧ又はＡ又はＣである）（例えば、ＴＴＴＴＷＴＡＲＧ（配列番号２２７））；及び（ｃ）$X_1X_2X_3$ＡＣ（配列番号２２８）（ここで、X_1はＡ又はＣ又はＧであり、X_2はＣ又はＡであり、X_3はＡ又はＣである）（例えば、ＡＣＡＡＣ（配列番号２２９））を含む。本明細書に記載される方法のいずれかの一部の実施形態において、配列番号２２４はダイレクトリピートの５’末端に近接している。本明細書に記載される方法のいずれかの一部の実施形態において、配列番号２２８はダイレクトリピートの３’末端に近接している。

【００５２】

本明細書に記載される方法のいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号１に記載のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、ダイレクトリピート配列は、配列番号５７に記載のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。本明細書に記載される方法のいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号１に記載のアミノ酸配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、ダイレクトリピート配列は、配列番号５７に記載のヌクレオチド配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。本明細書に記載される方法のいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパ

ク質はプロトスペーサー隣接モチーフ（ＰＡＭ）配列の認識能を有し、ここで、ＰＡＭ配列は、５’－ＴＮＮＴ－３’又は５’－ＴＮＲＴ－３’として記載される核酸配列を含み、「Ｎ」は任意のヌクレオチドであり、「Ｒ」はＡ又はＧである。

【００５３】

本明細書に記載される方法のいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号４に記載のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、ダイレクトリピート配列は、配列番号６０に記載のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。本明細書に記載される方法のいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号４に記載のアミノ酸配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、ダイレクトリピート配列は、配列番号６０に記載のヌクレオチド配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。本明細書に記載される方法のいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質はプロトスペーサー隣接モチーフ（ＰＡＭ）配列の認識能を有し、ここで、ＰＡＭ配列は、５’－ＮＴＴＮ－３’、５’－ＮＴＴＲ－３’（例えば、５’－ＴＴＴＧ－３’）、又は５’－ＮＮＲ－３’として記載される核酸配列を含み、「Ｎ」は任意のヌクレオチドであり、「Ｒ」はＡ又はＧである。

【００５４】

本明細書に記載される方法のいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号１０に記載のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、ダイレクトリピート配列は、配列番号６２又は配列番号２１３に記載のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。本明細書に記載される方法のいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号１０に記載のアミノ酸配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、ダイレクトリピート配列は、配列番号６２又は配列番号２１３に記載のヌクレオチド配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。本明細書に記載される方法のいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質はプロトスペーサー隣接モチーフ（ＰＡＭ）配列の認識能を有し、ここで、ＰＡＭ配列は、５’－ＮＴＴＮ－３’又は５’－ＲＴＴＲ－３’（例えば、５’－ＡＴＴＧ－３’又は５’－ＧＴＴＡ－３’）として記載される核酸配列を含み、「Ｎ」は任意のヌクレオチドであり、「Ｒ」はＡ又はＧである。

【００５５】

本明細書に記載される方法のいずれかの一部の実施形態において、スペーサー配列は、約１５ヌクレオチド～約５５ヌクレオチドを含む。本明細書に記載される方法のいずれかの一部の実施形態において、スペーサー配列は、２０～４５ヌクレオチドを含む。

【００５６】

本明細書に記載される方法のいずれかの一部の実施形態において、ＲＮＡガイドは、任意選択でｔｒａｃｒＲＮＡ及び／又はモジュレーターＲＮＡを含む。本明細書に記載される方法のいずれかの一部の実施形態において、システムはｔｒａｃｒＲＮＡを更に含む。本明細書に記載される方法のいずれかの一部の実施形態において、システムはｔｒａｃｒ

ＲＮＡを含まない。本明細書に記載される方法のいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は自己プロセシングである。本明細書に記載される方法のいずれかの一部の実施形態において、システムはモジュレーターＲＮＡを更に含む。

【００５７】

本明細書に記載される方法のいずれかの一部の実施形態において、標的核酸はＤＮＡ分子である。本明細書に記載される方法のいずれかの一部の実施形態において、標的核酸はＰＡＭ配列を含む。

【００５８】

本明細書に記載される方法のいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、非特異的ヌクレアーゼ活性を有する。

【００５９】

本明細書に記載される方法のいずれかの一部の実施形態において、標的核酸の修飾は、二本鎖切断イベントである。本明細書に記載される方法のいずれかの一部の実施形態において、標的核酸の修飾は、一本鎖切断イベントである。本明細書に記載される方法のいずれかの一部の実施形態において、標的核酸の修飾により、挿入イベントが生じる。本明細書に記載される方法のいずれかの一部の実施形態において、標的核酸の修飾により、欠失イベントが生じる。本明細書に記載される方法のいずれかの一部の実施形態において、標的核酸の修飾により、細胞毒性又は細胞死が生じる。

【００６０】

別の態様において、本開示は、標的核酸の編集方法を提供し、この方法は、本明細書に記載されるシステムを標的核酸に接触させることを含む。別の態様において、本開示は、標的核酸の発現を改変する方法を提供し、この方法は、本明細書に記載されるシステムを標的核酸に接触させることを含む。別の態様において、本開示は、標的核酸のある部位におけるペイロード核酸の挿入を標的化する方法であって、本明細書に記載されるシステムを標的核酸に接触させることを含む、方法を提供する。別の態様において、本開示は、標的核酸の部位からのペイロード核酸の切出しを標的化する方法であって、本明細書に記載されるシステムを標的核酸に接触させることを含む、方法を提供する。別の態様において、本開示は、ＤＮＡ標的核酸の認識時に一本鎖ＤＮＡを非特異的に分解する方法を提供し、この方法は、本明細書に記載されるシステムを標的核酸に接触させることを含む。

【００６１】

本明細書に提供されるシステム又は方法のいずれかの一部の実施形態において、接触は、直接接触又は間接接触を含む。本明細書に提供されるシステム又は方法のいずれかの一部の実施形態において、間接的に接触することは、ＲＮＡガイド及び／又はＣＲＩＳＰＲ関連タンパク質の産生を可能にする条件下で、本明細書に記載されるＲＮＡガイド又はＣＲＩＳＰＲ関連タンパク質をコードする１つ以上の核酸を投与することを含む。本明細書に提供されるシステム又は方法のいずれかの一部の実施形態において、接触は、インビボでの接触又はインビトロでの接触を含む。本明細書に提供されるシステム又は方法のいずれかの一部の実施形態において、標的核酸をシステムと接触させることは、ＣＲＩＳＰＲ関連タンパク質及びガイドＲＮＡが標的核酸に到達することを可能にする条件下で、核酸を含む細胞をシステムと接触させることを含む。本明細書に提供されるシステム又は方法のいずれかの一部の実施形態において、インビボで細胞をシステムと接触させることは、ＣＲＩＳＰＲ関連タンパク質及びガイドＲＮＡが細胞に到達するか又は細胞内で産生されることを可能にする条件下で、細胞を含む対象にシステムを投与することを含む。

【００６２】

別の態様において、本開示は、（ａ）標的核酸のターゲティング及び編集方法；（ｂ）核酸の認識に応じた一本鎖核酸の非特異的分解方法；（ｃ）二本鎖標的のスペーサー相補鎖の認識に応じた二本鎖標的の非スペーサー相補鎖のターゲティング及びニッキング方法；（ｄ）二本鎖標的核酸のターゲティング及び切断方法；（ｅ）試料中の標的核酸の検出方法；（ｆ）二本鎖核酸の特異的編集方法；（ｇ）二本鎖核酸の塩基編集方法；（ｈ）細胞における遺伝子型特異的又は転写状態特異的細胞死又は休眠の誘導方法；（ｉ）二本鎖

核酸標的におけるインデルの作成方法；（ｊ）二本鎖核酸標的への配列の挿入方法；又は（ｋ）二本鎖核酸標的における配列の欠失又は逆位形成方法である、インビトロ又はエキソビボでの方法における使用のための、本明細書に提供されるシステムを提供する。

【００６３】

別の態様において、本開示は（ａ）ＣＲＩＳＰＲ関連タンパク質が、配列番号１～５６のいずれか１つに記載のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含む、ＣＲＩＳＰＲ関連タンパク質をコードする核酸配列；及び（ｂ）ダイレクトリピート配列と標的核酸へのハイブリダイゼーション能を有するスペーサー配列とを含むＲＮＡガイド（又はＲＮＡガイドをコードする核酸）のトランスフェクションを含む、哺乳動物細胞における標的核酸中への挿入又は欠失を導入する方法を提供し、ここで、ＣＲＩＳＰＲ関連タンパク質は、ＲＮＡガイドへの結合能を有し、ＣＲＩＳＰＲ関連タンパク質及びＲＮＡガイドによる標的核酸の認識により、標的核酸の修飾が生じる。

【００６４】

本明細書に提供される方法のいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号４に記載のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含む。本明細書に提供される方法のいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号４に記載のアミノ酸配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含む。本明細書に提供される方法のいずれかの一部の実施形態において、ダイレクトリピートは、配列番号６０に記載のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。本明細書に提供される方法のいずれかの一部の実施形態において、ダイレクトリピートは、配列番号６０に記載のヌクレオチド配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。本明細書に提供される方法のいずれかの一部の実施形態において、標的核酸はＰＡＭ配列に隣接し、ＰＡＭ配列は、５’－ＮＴＴＮ－３’、５’－ＮＴＴＲ－３’（例えば、５’－ＴＴＴＧ－３’）、又は５’－ＮＮＲ－３’として記載される核酸配列を含み、ここで、「Ｎ」は任意のヌクレオチドであり、「Ｒ」はＡ又はＧである。

【００６５】

本明細書に提供される方法のいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号１０に記載のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含む。本明細書に提供される方法のいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号１０に記載のアミノ酸配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含む。本明細書に提供される方法のいずれかの一部の実施形態において、ダイレクトリピートは、配列番号６２又は配列番号２１３に記載のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。本明細書に提供される方法のいずれかの一部の実施形態において、ダイレクトリピートは、配列番号６２又は配列番号２１３に記載のヌクレオチド配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。本明細書に提供される方法のいずれかの一部の実施形態において、標的核酸はＰＡＭ配列に隣接し、ＰＡＭ配列は、５’

－ＮＴＴＮ－３’又は５’－ＲＴＴＲ－３’（例えば、５’－ＡＴＴＧ－３’又は５’－ＧＴＴＡ－３’）として記載される核酸配列を含み、ここで、「Ｎ」は任意のヌクレオチドであり、「Ｒ」はＡ又はＧである。

【００６６】

本明細書に提供される方法のいずれかの一部の実施形態において、トランスフェクションは一過性トランスフェクションである。本明細書に提供される方法のいずれかの一部の実施形態において、細胞はヒト細胞である。

【００６７】

別の態様において、本開示は、（ａ）ＣＲＩＳＰＲ関連タンパク質又はＣＲＩＳＰＲ関連タンパク質をコードする核酸及び（ｂ）ダイレクトリピート配列とスペーサー配列とを含むＲＮＡガイドを含む組成物を提供し；ここで、ＣＲＩＳＰＲ関連タンパク質は、以下のアミノ酸配列の１つ以上：（ｉ）ＰＸ$_1$Ｘ$_2$Ｘ$_3$Ｘ$_4$Ｆ（配列番号２１６）（ここで、Ｘ$_1$はＬ又はＭ又はＩ又はＣ又はＦであり、Ｘ$_2$はＹ又はＷ又はＦであり、Ｘ$_3$はＫ又はＴ又はＣ又はＲ又はＷ又はＹ又はＨ又はＶであり、Ｘ$_4$はＩ又はＬ又はＭである）；（ｉｉ）ＲＸ$_1$Ｘ$_2$Ｘ$_3$Ｌ（配列番号２１７）（ここで、Ｘ$_1$はＩ又はＬ又はＭ又はＹ又はＴ又はＦであり、Ｘ$_2$はＲ又はＱ又はＫ又はＥ又はＳ又はＴであり、Ｘ$_3$はＬ又はＩ又はＴ又はＣ又はＭ又はＫである）；（ｉｉｉ）ＮＸ$_1$ＹＸ$_2$（配列番号２１８）（ここで、Ｘ$_1$はＩ又はＬ又はＦであり、Ｘ$_2$はＫ又はＲ又はＶ又はＥである）；（ｉｖ）ＫＸ$_1$Ｘ$_2$Ｘ$_3$ＦＡＸ$_4$Ｘ$_5$ＫＤ（配列番号２１９）（ここで、Ｘ$_1$はＴ又はＩ又はＮ又はＡ又はＳ又はＦ又はＶであり、Ｘ$_2$はＩ又はＶ又はＬ又はＳであり、Ｘ$_3$はＨ又はＳ又はＧ又はＲであり、Ｘ$_4$はＤ又はＳ又はＥであり、Ｘ$_5$はＩ又はＶ又はＭ又はＴ又はＮである）；（ｖ）ＬＸ$_1$ＮＸ$_2$（配列番号２２０）（ここで、Ｘ$_1$はＧ又はＳ又はＣ又はＴであり、Ｘ$_2$はＮ又はＹ又はＫ又はＳである）；（ｖｉ）ＰＸ$_1$Ｘ$_2$Ｘ$_3$Ｘ$_4$ＳＱＸ$_5$ＤＳ（配列番号２２１）（ここで、Ｘ$_1$はＳ又はＰ又はＡであり、Ｘ$_2$はＹ又はＳ又はＡ又はＰ又はＥ又はＹ又はＱ又はＮであり、Ｘ$_3$はＦ又はＹ又はＨであり、Ｘ$_4$はＴ又はＳであり、Ｘ$_5$はＭ又はＴ又はＩである）；（ｖｉｉ）ＫＸ$_1$Ｘ$_2$ＶＲＸ$_3$Ｘ$_4$ＱＥＸ$_5$Ｈ（配列番号２２２）（ここで、Ｘ$_1$はＮ又はＫ又はＷ又はＲ又はＥ又はＴ又はＹであり、Ｘ$_2$はＭ又はＲ又はＬ又はＳ又はＫ又はＶ又はＥ又はＴ又はＩ又はＤであり、Ｘ$_3$はＬ又はＲ又はＨ又はＰ又はＴ又はＫ又はＰのＱ又はＳ又はＡであり、Ｘ$_4$はＧ又はＱ又はＮ又はＲ又はＫ又はＥ又はＩ又はＴ又はＳ又はＣであり、Ｘ$_5$はＲ又はＷ又はＹ又はＫ又はＴ又はＦ又はＳ又はＱである）；及び（ｖｉｉｉ）Ｘ$_1$ＮＧＸ$_2$Ｘ$_3$Ｘ$_4$ＤＸ$_5$ＮＸ$_6$Ｘ$_7$Ｘ$_8$Ｎ（配列番号２２３）（ここで、Ｘ$_1$はＩ又はＫ又はＶ又はＬであり、Ｘ$_2$はＬ又はＭであり、Ｘ$_3$はＮ又はＨ又はＰであり、Ｘ$_4$はＡ又はＳ又はＣであり、Ｘ$_5$はＶ又はＹ又はＩ又はＦ又はＴ又はＮであり、Ｘ$_6$はＡ又はＳであり、Ｘ$_7$はＳ又はＡ又はＰであり、Ｘ$_8$はＭ又はＣ又はＬ又はＲ又はＮ又はＳ又はＫ又はＬである）を含み、ＣＲＩＳＰＲ関連タンパク質は、ＲＮＡガイドに結合し、スペーサー配列に相補的な標的核酸配列を修飾することができる。

【００６８】

本明細書に記載される組成物のいずれかの一部の実施形態において、ダイレクトリピート配列は、以下の配列の１つ以上：（ａ）Ｘ$_1$Ｘ$_2$ＴＸ$_3$Ｘ$_4$Ｘ$_5$Ｘ$_6$Ｘ$_7$Ｘ$_8$（配列番号２２４）（ここで、Ｘ$_1$はＡ又はＣ又はＧであり、Ｘ$_2$はＴ又はＣ又はＡであり、Ｘ$_3$はＴ又はＧ又はＡであり、Ｘ$_4$はＴ又はＧであり、Ｘ$_5$はＴ又はＧ又はＡであり、Ｘ$_6$はＧ又はＴ又はＡであり、Ｘ$_7$はＴ又はＧ又はＡであり、Ｘ$_8$はＡ又はＧ又はＴである）（例えば、ＡＴＴＧＴＴＧＤＡ（配列番号２２５））；（ｂ）Ｘ$_1$Ｘ$_2$Ｘ$_3$Ｘ$_4$Ｘ$_5$Ｘ$_6$Ｘ$_7$Ｘ$_8$Ｘ$_9$（配列番号２２６）（ここで、Ｘ$_1$はＴ又はＣ又はＡであり、Ｘ$_2$はＴ又はＡ又はＧであり、Ｘ$_3$はＴ又はＣ又はＡであり、Ｘ$_4$はＴ又はＡであり、Ｘ$_5$はＴ又はＡ又はＧであり、Ｘ$_6$はＴ又はＡであり、Ｘ$_7$はＡ又はＴであり、Ｘ$_8$はＡ又はＧ又はＣ又はＴであり、Ｘ$_9$はＧ又はＡ又はＣである）（例えば、ＴＴＴＴＷＴＡＲＧ（配列番号２２７））；及び（ｃ）Ｘ$_1$Ｘ$_2$Ｘ$_3$ＡＣ（配列番号２２８）（ここで、Ｘ$_1$はＡ又はＣ又はＧであり、Ｘ$_2$はＣ又はＡであり、Ｘ$_3$はＡ又はＣである）（例えば、ＡＣＡＡＣ（配列番号２２９））を含む。本明細書に記載される組成物のいずれかの一部の実施形態において、配列番号２２

４はダイレクトリピートの５’末端に近接している。本明細書に記載される組成物のいずれかの一部の実施形態において、配列番号２２８はダイレクトリピートの３’末端に近接している。

【００６９】

本明細書に記載される組成物のいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、少なくとも１つ（例えば、１つ、２つ、又は３つ）のＲｕｖＣドメイン又は少なくとも１つの分割されたＲｕｖＣドメインを含む。

【００７０】

本明細書に記載される組成物のいずれかの一部の実施形態において、ＲＮＡガイドのスペーサー配列は、約１５ヌクレオチド～約５５ヌクレオチドを含む。本明細書に記載される組成物のいずれかの一部の実施形態において、ＲＮＡガイドのスペーサー配列は、２０～４５ヌクレオチドを含む。

【００７１】

本明細書に記載される組成物のいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は触媒残基（例えば、アスパラギン酸又はグルタミン酸）を含む。本明細書に記載される組成物のいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、標的核酸を切断する。本明細書に記載される組成物のいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、ペプチドタグ、蛍光タンパク質、塩基編集ドメイン、ＤＮＡメチル化ドメイン、ヒストン残基修飾ドメイン、局在化因子、転写修飾因子、光ゲート制御因子、化学誘導性因子、又はクロマチン可視化因子を更に含む。

【００７２】

本明細書に記載される組成物のいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質をコードする核酸は、細胞（例えば、真核細胞、例えば、哺乳動物細胞、例えば、ヒト細胞）での発現にコドン最適化される。本明細書に記載される組成物のいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質をコードする核酸は、プロモーターに作動可能に連結されている。本明細書に記載される組成物のいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質をコードする核酸は、ベクター内にある。一部の実施形態において、ベクターは、レトロウイルスベクター、レンチウイルスベクター、ファージベクター、アデノウイルスベクター、アデノ随伴ベクター、又は単純ヘルペスベクターを含む。

【００７３】

本明細書に記載される組成物のいずれかの一部の実施形態において、標的核酸はＤＮＡ分子である。本明細書に記載される組成物のいずれかの一部の実施形態において、標的核酸はＰＡＭ配列を含む。

【００７４】

本明細書に記載される組成物のいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、非特異的ヌクレアーゼ活性を有する。

【００７５】

本明細書に記載される組成物のいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質及びＲＮＡガイドによる標的核酸の認識により、標的核酸の修飾が生じる。本明細書に記載される組成物のいずれかの一部の実施形態において、標的核酸の修飾は、二本鎖切断イベントである。本明細書に記載される組成物のいずれかの一部の実施形態において、標的核酸の修飾は、一本鎖切断イベントである。本明細書に記載される組成物のいずれかの一部の実施形態において、標的核酸の修飾により、挿入イベントが生じる。本明細書に記載される組成物のいずれかの一部の実施形態において、標的核酸の修飾により、欠失イベントが生じる。本明細書に記載される組成物のいずれかの一部の実施形態において、標的核酸の修飾により、細胞毒性又は細胞死が生じる。

【００７６】

本明細書に記載される組成物のいずれかの一部の実施形態において、システムはドナー鋳型核酸を更に含む。本明細書に記載される組成物のいずれかの一部の実施形態において

、ドナー鋳型核酸はＤＮＡ分子である。本明細書に記載される組成物のいずれかの一部の実施形態において、ドナー鋳型核酸はＲＮＡ分子である。

【００７７】

本明細書に記載される組成物のいずれかの一部の実施形態において、ＲＮＡガイドは任意選択でｔｒａｃｒＲＮＡを含む。本明細書に記載される組成物のいずれかの一部の実施形態において、システムはｔｒａｃｒＲＮＡを更に含む。本明細書に記載される組成物のいずれかの一部の実施形態において、システムはｔｒａｃｒＲＮＡを含まない。本明細書に記載される組成物のいずれかの一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は自己プロセシングである。

【００７８】

本明細書に記載される組成物のいずれかの一部の実施形態において、システムは、ナノ粒子、リポソーム、エキソソーム、微小胞、又は遺伝子銃を含む送達組成物中に存在する。

【００７９】

本明細書に記載される組成物のいずれかの一部の実施形態において、組成物は細胞内にある。一部の実施形態において、細胞は真核細胞である。一部の実施形態において、細胞は哺乳動物細胞である。一部の実施形態において、細胞はヒト細胞である。一部の実施形態において、細胞は原核細胞である。

【００８０】

本明細書に記載されるエフェクターは、限定はされないが、１）新規核酸編集特性及び制御機構、２）送達戦略におけるより高い汎用性に比したより小さいサイズ、３）遺伝子型によって惹起される細胞死などの細胞過程、及び４）プログラム可能なＲＮＡ誘導型ＤＮＡ挿入、切出し、及び動員、及び５）非ヒト共生源を介した既存の免疫の差別化されたプロファイルを含めた更なる特徴を提供する。例えば、実施例１、４、及び５並びに図１～３及び５～１１Ｄを参照されたい。本明細書に記載される新規ＤＮＡターゲティングシステムがゲノム及びエピゲノム操作技法のツールボックスに加わることにより、特異的でプログラムされた摂動への幅広い適用が実現する。

【００８１】

本発明の他の特徴及び利点は、以下の詳細な説明から、及び特許請求の範囲から明らかであろう。

【００８２】

図面は、ＣＬＵＳＴ．０９１９７９と呼ばれるタンパク質クラスターの分析結果を表す一連の概略表現である。

【図面の簡単な説明】

【００８３】

【図１Ａ】図１Ａ、図１Ｂ、図１Ｃ、図１Ｄ、図１Ｅ、図１Ｆ、図１Ｇ、図１Ｈ、図１Ｉ、図１Ｊ、図１Ｋ、及び図１Ｌは、配列番号１～４、１４、１５、１７～１９、２１～２５、２７～３３、３５～４９、５１～５６のエフェクターのアラインメントをまとめて示す。

【図１Ｂ】同上。

【図１Ｃ】同上。

【図１Ｄ】同上。

【図１Ｅ】同上。

【図１Ｆ】同上。

【図１Ｇ】同上。

【図１Ｈ】同上。

【図１Ｉ】同上。

【図１Ｊ】同上。

【図１Ｋ】同上。

【図１Ｌ】同上。

【図２】ＣＬＵＳＴ．０９１９７９エフェクターのＲｕｖＣドメインを示す概略図であり

、これは、表６に示される配列のコンセンサス配列に基づく。
【図３】配列番号５７、５８、６０、６２、６３、７０、７２～７４、７６、７７、８０、８３、８４、８６～８８、９０、１２８、１３０、１３９、及び２１３のダイレクトリピート配列のアラインメントを示す。コンセンサス配列（配列番号２３０）はアラインメントの上段に示す。
【図４Ａ】実施例４に記載されているインビボネガティブ選択スクリーニングアッセイの構成要素の概略表現である。２個のＤＲが隣接する且つＪ２３１１９によって発現されるｐＡＣＹＣ１８４又は大腸菌（Ｅ．ｃｏｌｉ）必須遺伝子の両鎖から均一にサンプリングした非代表的スペーサーを含むＣＲＩＳＰＲアレイライブラリを設計した。
【図４Ｂ】実施例４に記載されているインビボネガティブ選択スクリーニングワークフローの概略表現である。ＣＲＩＳＰＲアレイライブラリは、エフェクタープラスミドにクローニングした。エフェクタープラスミド及び非コードプラスミドを大腸菌（Ｅ．ｃｏｌｉ）に形質転換し、続いてｐＡＣＹＣ１８４又は大腸菌（Ｅ．ｃｏｌｉ）必須遺伝子からの転写物に対して干渉を付与するＣＲＩＳＰＲアレイのネガティブ選択用に成長させた。エフェクタープラスミドの標的化シーケンシングを使用して、枯渇したＣＲＩＳＰＲアレイを同定した。スモールＲＮＡｓｅｑを更に実施して、成熟ｃｒＲＮＡ及び潜在的ｔｒａｃｒＲＮＡの要件を特定した。
【図５】非コード配列を有する、ｐＡＣＹＣ１８４及びダイレクトリピート転写方向を標的とするスペーサーについてのエンジニアリングされた組成物の枯渇活性の程度を示す、ＣＬＵＳＴ．０９１９７９　ＡＵＸＯ０１３９８８８８２（配列番号１に記載のエフェクター）のグラフである。「順」方向のダイレクトリピート（５’－ＡＣＴＡ…ＡＡＣＴ－［スペーサー］－３’）及び「逆」方向のダイレクトリピート（５’－ＡＧＴＴ…ＴＡＧＴ－［スペーサー］－３’）の枯渇の程度が示される。
【図６Ａ】図６Ａは、ｐＡＣＹＣ１８４プラスミド上の位置による、非コード配列を有する、ＣＬＵＳＴ．０９１９７９　ＡＵＸＯ０１３９８８８８２の枯渇及び非枯渇標的の密度を示すグラフ表現である。図６Ｂは、大腸菌（Ｅ．ｃｏｌｉ）株、Ｅ．Ｃｌｏｎｉ上の位置による、非コード配列を有する、ＣＬＵＳＴ．０９１９７９　ＡＵＸＯ０１３９８８８８２の枯渇及び非枯渇標的の密度を示すグラフ表現である。トップ鎖上及びボトム鎖上の標的を別々に、アノテートされた遺伝子の向きに関して示す。バンドの大きさは枯渇の程度を示し、明るいバンドはヒット閾値の３に近い。グラジエントは、相対的な転写物の存在度を示すＲＮＡシーケンシングのヒートマップである。
【図６Ｂ】同上。
【図７】ＣＬＵＳＴ．０９１９７９　ＡＵＸＯ０１３９８８８８２（非コード配列を含む）のＰＡＭ配列の予測としての、Ｅ．Ｃｌｏｎｉの枯渇した標的に隣接する配列のＷｅｂＬｏｇｏである。
【図８】非コード配列を有する、ｐＡＣＹＣ１８４及びダイレクトリピート転写方向を標的とするスペーサーについてのエンジニアリングされた組成物の枯渇活性の程度を示す、ＣＬＵＳＴ．０９１９７９　ＳＲＲ３１８１１５１（配列番号４に記載のエフェクター）のグラフである。「順」方向のダイレクトリピート（５’－ＧＴＴＧ…ＣＡＧＧ－［スペーサー］－３’）及び「逆」方向のダイレクトリピート（５’－ＣＣＴＧ…ＣＡＡＣ－［スペーサー］－３’）の枯渇の程度がそれぞれ実線及び破線で示される。
【図９Ａ】図９Ａは、ｐＡＣＹＣ１８４プラスミド上の位置による、非コード配列を有する、ＣＬＵＳＴ．０９１９７９　ＳＲＲ３１８１１５１の枯渇及び非枯渇標的の密度を示すグラフ表現である。図９Ｂは、大腸菌（Ｅ．ｃｏｌｉ）株、Ｅ．Ｃｌｏｎｉ上の位置による、非コード配列を有する、ＣＬＵＳＴ．０９１９７９　ＳＲＲ３１８１１５１の枯渇及び非枯渇標的の密度を示すグラフ表現である。トップ鎖上及びボトム鎖上の標的を別々に、アノテートされた遺伝子の向きに関して示す。バンドの大きさは枯渇の程度を示し、明るいバンドはヒット閾値の３に近い。グラジエントは、相対的な転写物の存在度を示すＲＮＡシーケンシングのヒートマップである。
【図９Ｂ】同上。

【図１０】ＣＬＵＳＴ．０９１９７９　ＳＲＲ３１８１１５１（非コード配列を有する）のＰＡＭ配列の予測としての、Ｅ．Ｃｌｏｎｉの枯渇した標的に隣接する配列のＷｅｂＬｏｇｏである。
【図１１Ａ】図１１Ａは、ＨＥＫ２９３細胞の配列番号２０６のＡＡＶＳ１標的遺伝子座及び配列番号２０８のＶＥＧＦＡ標的遺伝子座において、配列番号４のエフェクターによって誘導されるインデルを示す。図１１Ｂは、ＨＥＫ２９３細胞の配列番号２５３、２５５、２５７、２５９、及び２７５のＡＡＶＳ１標的遺伝子座、配列番号２６３、２６５、２６７、２６９、２７１、２７３、及び２７７のＶＥＧＦＡの標的遺伝子座、並びに配列番号２６１のＥＭＸ１標的遺伝子座において、配列番号４のエフェクターによって誘導されるインデルを示す。図１１Ｃは、ＨＥＫ２９３細胞の配列番号２１０のＡＡＶＳ１標的遺伝子座、配列番号２１２のＡＡＶＳ１標的遺伝子座、及び配列番号２１５のＶＥＧＦＡ標的遺伝子座において、配列番号１０のエフェクターによって誘導されるインデルを示す。図１１Ｄは、ＨＥＫ２９３細胞の配列番号２７９、２８１、２８５、及び２８７のＡＡＶＳ１標的遺伝子座、配列番号２８３のＶＥＧＦＡ標的遺伝子座、並びに配列番号２８９のＥＭＸ１標的遺伝子座において、配列番号１０のエフェクターによって誘導されるインデルを示す。
【図１１Ｂ】同上。
【図１１Ｃ】同上。
【図１１Ｄ】同上。
【発明を実施するための形態】
【００８４】
ＣＲＩＳＰＲ－Ｃａｓシステムは、天然に多様であり、プログラム可能なバイオテクノロジーに生かすことのできる様々な活性機構及び機能要素が含まれている。天然では、これらのシステムが外来ＤＮＡ及びウイルスに対する効率的な防御を実現する一方で、自己と非自己の判別を提供して自己標的化を回避している。エンジニアリングされた設定では、これらのシステムは、分子技術の多様なツールボックスを提供し、ターゲティング空間の境界を画定する。本明細書に記載される方法を用いて、シングルサブユニットのクラス２エフェクターシステム内に、ＲＮＡによるプログラムが可能な核酸操作の能力を拡張する、更なる機構及びパラメータが発見された。
【００８５】
特に定義しない限り、本明細書で使用される全ての科学技術用語は、本発明が属する技術分野の当業者が一般的に理解するのと同じ意味を有する。本発明の実施又は試験においては、本明細書に記載されるものと同様の又は等価な方法及び材料を使用し得るが、好適な方法及び材料を以下に記載する。本明細書において言及される刊行物、特許出願、特許、及び他の参考文献は全て、全体として参照により援用される。矛盾が生じる場合、定義を含め、本明細書が優先するものとする。加えて、材料、方法、及び例は例示に過ぎず、限定することを意図するものではない。出願人は、特許法の標準的な慣行に従って、「を含む」、「から本質的になる」、又は「からなる」という移行句を使用して、任意の開示された発明を代替的に請求する権利を留保する。
【００８６】
本明細書で使用される場合、単数形「ａ」、「ａｎ」、及び「ｔｈｅ」は、文脈が明らかに他のものを示さない限り、複数の指示対象を含む。例えば、「核酸」への言及は、１つ以上の核酸を意味する。
【００８７】
本明細書において、「好ましくは」、「適切に」、「一般的に」、及び「典型的に」などの用語は、請求される発明の範囲を制限するため、又は特定の特徴が請求される発明の構造又は機能にとって重大、必須、又は更に重要であることを示唆するために使用されるものではないことに留意されたい。むしろ、これらの用語は、本発明の特定の実施形態において利用できる、又は利用できない、代替又は追加の特徴を強調することを単に意図するものである。

【００８８】
本発明を説明及び定義する目的で、用語「実質的に」は、任意の定量的比較、値、測定、又は他の表現に帰することができる固有の不確実性の程度を表すために本明細書で使用されることに留意されたい。用語「実質的に」はまた、問題となっている主題の基本的な機能に変化をもたらすことなく、記載される参照物から定量的表現が変化し得る程度を表すために本明細書で使用される。
【００８９】
用語「ＣＲＩＳＰＲ－Ｃａｓシステム」は、本明細書で使用されるとき、ＣＲＩＳＰＲエフェクターをコードする配列、ＲＮＡガイド、並びに他の配列及びＣＲＩＳＰＲ遺伝子座からの転写物を含む、ＣＲＩＳＰＲエフェクターの発現に関与する、又はその活性を導く核酸及び／又はタンパク質を指す。
【００９０】
用語「ＣＲＩＳＰＲ関連タンパク質」、「ＣＲＩＳＰＲ－Ｃａｓエフェクター」、「ＣＲＩＳＰＲエフェクター」、「エフェクター」、「エフェクタータンパク質」、「ＣＲＩＳＰＲ酵素」などは、本明細書で同義的に使用されるとき、酵素活性を実行するタンパク質又はＲＮＡガイドによって指定される核酸上の標的部位に結合するタンパク質を指す。一部の実施形態において、ＣＲＩＳＰＲエフェクターは、エンドヌクレアーゼ活性、ニッカーゼ活性、及び／又はエキソヌクレアーゼ活性を有する。
【００９１】
用語「ＲＮＡガイド」、「ガイドＲＮＡ」、「ｇＲＮＡ」、及び「ガイド配列」は、本明細書で使用されるとき、ＤＮＡ及び／又はＲＮＡなどの標的核酸への本明細書に記載されるエフェクターのターゲティングを促進する任意のＲＮＡ分子を指す。例示的な「ＲＮＡガイド」としては、限定はされないが、ｃｒＲＮＡ、並びにｔｒａｃｒＲＮＡ及び／又はモジュレーターＲＮＡのいずれかとハイブリダイズ又は融合したｃｒＲＮＡが挙げられる。一部の実施形態において、ＲＮＡガイドは、単一のＲＮＡ分子に融合された、又は別個のＲＮＡ分子としての、ｃｒＲＮＡ及びｔｒａｃｒＲＮＡの両方を含む。一部の実施形態において、ＲＮＡガイドは、単一のＲＮＡ分子に融合された、又は別個のＲＮＡ分子としての、ｃｒＲＮＡ及びモジュレーターＲＮＡを含む。一部の実施形態において、ＲＮＡガイドは、単一のＲＮＡ分子に融合された、又は別個のＲＮＡ分子としての、ｃｒＲＮＡ、ｔｒａｃｒＲＮＡ、及びモジュレーターＲＮＡを含む。
【００９２】
用語「ＣＲＩＳＰＲエフェクター複合体」、「エフェクター複合体」、又は「監視複合体」は、本明細書で使用されるとき、ＣＲＩＳＰＲエフェクター及びＲＮＡガイドを含む複合体を指す。ＣＲＩＳＰＲエフェクター複合体は、１つ以上のアクセサリータンパク質を更に含み得る。１つ以上のアクセサリータンパク質は、非触媒的及び／又は非標的結合であり得る。ｃｒＲＮＡは、ｔｒａｃｒＲＮＡにハイブリダイズする配列を含み得る。次にはｃｒＲＮＡ：ｔｒａｃｒＲＮＡ二重鎖は、ＣＲＩＳＰＲエフェクターに結合し得る。本明細書で使用されるとき、用語「プレｃｒＲＮＡ」は、ＤＲ－スペーサー－ＤＲ配列を含むプロセシングされていないＲＮＡ分子を指す。本明細書で使用される場合、「成熟ｃｒＲＮＡ」という用語は、プレｃｒＲＮＡの処理された形態を指す。成熟ｃｒＲＮＡは、ＤＲスペーサー配列を含んでもよく、ここで、ＤＲはプレｃｒＲＮＡのＤＲの短縮型であり、且つ／又はスペーサーはプレｃｒＲＮＡのスペーサーの短縮型である。
【００９３】
用語「ＣＲＩＳＰＲ　ＲＮＡ」及び「ｃｒＲＮＡ」は、本明細書で使用されるとき、核酸配列を特異的に認識するためにＣＲＩＳＰＲエフェクターによって使用されるガイド配列を含むＲＮＡ分子を指す。ｃｒＲＮＡ「スペーサー」配列は核酸標的配列に相補的であり、核酸標的配列への部分的又は全体的結合能を有する。
【００９４】
用語「トランス活性化ｃｒＲＮＡ」又は「ｔｒａｃｒＲＮＡ」は、本明細書で使用されるとき、ＣＲＩＳＰＲエフェクターが特定の標的核酸に結合するために必要な構造及び／

又は配列モチーフを形成する配列を含むＲＮＡ分子を指す。
【００９５】
用語「ＣＲＩＳＰＲアレイ」は、本明細書で使用されるとき、最初のＣＲＩＳＰＲリピートの最初のヌクレオチドから始まって最後の（末端）ＣＲＩＳＰＲリピートの最後のヌクレオチドで終わる、ＣＲＩＳＰＲリピートとスペーサーとを含む核酸（例えば、ＤＮＡ）セグメントを指す。典型的には、ＣＲＩＳＰＲアレイ中の各スペーサーは２つのリピート間に位置する。用語「ＣＲＩＳＰＲリピート」、「ＣＲＩＳＰＲダイレクトリピート」、及び「ダイレクトリピート」は、本明細書で使用されるとき、複数の短い定方向に反復する配列を指し、これはＣＲＩＳＰＲアレイ内で配列変異をごく僅かしか又は全く示さない。
【００９６】
用語「モジュレーターＲＮＡ」は、本明細書に記載されるとき、ＣＲＩＳＰＲエフェクター又はＣＲＩＳＰＲエフェクターを含む核タンパク質複合体の活性を調節する（例えば、増加又は減少させる）任意のＲＮＡ分子を指す。一部の実施形態において、モジュレーターＲＮＡは、ＣＲＩＳＰＲエフェクター又はＣＲＩＳＰＲエフェクターを含む核タンパク質複合体のヌクレアーゼ活性を調節する。
【００９７】
本明細書で使用されるとき、用語「標的核酸」は、ＲＮＡガイド中のスペーサーの全体又は一部に相補的なヌクレオチド配列を含む核酸を指す。一部の実施形態において、標的核酸は遺伝子を含む。一部の実施形態において、標的核酸は非コード領域（例えば、プロモーター）を含む。一部の実施形態において、標的核酸は一本鎖である。一部の実施形態において、標的核酸は二本鎖である。「転写活性部位」は、本明細書で使用されるとき、活発に転写されている核酸配列中の部位を指す。
【００９８】
本明細書で使用されるとき、用語「プロトスペーサー隣接モチーフ」又は「ＰＡＭ」は、エフェクター及びＲＮＡガイドを含む複合体が結合する標的配列に隣接するＤＮＡ配列を指す。一部の実施形態において、酵素活性のためにＰＡＭが必要である。本明細書で使用されるとき、用語「隣接」は、複合体のＲＮＡガイドが、ＰＡＭに直接隣接する標的配列に特異的に結合、相互作用、又は会合する場合を含む。このような場合、標的配列とＰＡＭの間にヌクレオチドは存在しない。用語「隣接」はまた、標的化部分が結合する標的配列とＰＡＭとの間に少数（例えば、１つ、２つ、３つ、４つ、又は５つ）のヌクレオチドが存在する場合を含む。本明細書で使用されるとき、用語「ＰＡＭ配列を認識すること」は、標的核酸への、ＣＲＩＳＰＲ関連タンパク質及びｃｒＲＮＡを含む複合体の結合を指し、ここで、標的核酸はＰＡＭ配列に隣接する。
【００９９】
用語「活性化したＣＲＩＳＰＲエフェクター複合体」、「活性化したＣＲＩＳＰＲ複合体」、及び「活性化した複合体」は、本明細書で使用されるとき、標的核酸を修飾することができるＣＲＩＳＰＲエフェクター複合体を指す。一部の実施形態において、活性化したＣＲＩＳＰＲ複合体は、活性化したＣＲＩＳＰＲ複合体が標的核酸に結合した後、標的核酸を修飾することができる。一部の実施形態において、活性化したＣＲＩＳＰＲ複合体の標的核酸への結合により、コラテラル切断などの追加の切断イベントが生じる。
【０１００】
用語「切断イベント」は、本明細書で使用されるとき、ＤＮＡ及び／又はＲＮＡなどの核酸の切断を指す。一部の実施形態において、切断イベントは、本明細書で使用されるとき、本明細書に記載されるＣＲＩＳＰＲシステムのヌクレアーゼによって作り出される標的核酸における切断を指す。一部の実施形態において、切断イベントは二本鎖ＤＮＡ切断である。一部の実施形態において、切断イベントは一本鎖ＤＮＡ切断である。一部の実施形態において、切断イベントは、コラテラル核酸の切断を指す。
【０１０１】
用語「コラテラル核酸」は、本明細書で使用されるとき、活性化したＣＲＩＳＰＲ複合

体によって非特異的に切断される核酸基質を指す。用語「コラテラルＤＮアーゼ活性」は、本明細書でＣＲＩＳＰＲエフェクターに言及して使用されるとき、活性化したＣＲＩＳＰＲ複合体の非特異的ＤＮアーゼ活性を指す。用語「コラテラルＲＮアーゼ活性」は、本明細書でＣＲＩＳＰＲエフェクターに言及して使用されるとき、活性化したＣＲＩＳＰＲ複合体の非特異的ＲＮアーゼ活性を指す。

【０１０２】

用語「ドナー鋳型核酸」は、本明細書で使用されるとき、本明細書に記載されるＣＲＩＳＰＲエフェクターが標的核酸を修飾した後、標的配列又は標的近位配列に、鋳型化された変更を行うために使用できる核酸分子を指す。一部の実施形態において、ドナー鋳型核酸は二本鎖核酸である。一部の実施形態において、ドナー鋳型核酸は一本鎖核酸である。一部の実施形態において、ドナー鋳型核酸は線状である。一部の実施形態において、ドナー鋳型核酸は環状（例えば、プラスミド）である。一部の実施形態において、ドナー鋳型核酸は外因性核酸分子である。一部の実施形態において、ドナー鋳型核酸は内因性核酸分子（例えば、染色体）である。

【０１０３】

本明細書で使用される場合、「ポリヌクレオチド」、「ヌクレオチド」、「オリゴヌクレオチド」、及び「核酸」という用語は、ＤＮＡ、ＲＮＡ、それらの誘導体、又はそれらの組み合わせを含む核酸を指すために同義的に使用され得る。当業者に周知の方法を使用して、本発明による遺伝子発現構築物及び組換え細胞を構築することができる。これらの方法としては、インビトロ組換えＤＮＡ技術、合成技術、インビボ組換え技術、及びポリメラーゼ連鎖反応（ＰＣＲ）技術が挙げられる。例えば、Ｍａｎｉａｔｉｓ ｅｔ ａｌ．，１９８９，ＭＯＬＥＣＵＬＡＲ ＣＬＯＮＩＮＧ：Ａ ＬＡＢＯＲＡＴＯＲＹ ＭＡＮＵＡＬ，Ｃｏｌｄ Ｓｐｒｉｎｇ Ｈａｒｂｏｒ Ｌａｂｏｒａｔｏｒｙ，Ｎｅｗ Ｙｏｒｋ；Ａｕｓｕｂｅｌ ｅｔ ａｌ．，１９８９，ＣＵＲＲＥＮＴ ＰＲＯＴＯＣＯＬＳ ＩＮ ＭＯＬＥＣＵＬＡＲ ＢＩＯＬＯＧＹ，Ｇｒｅｅｎｅ Ｐｕｂｌｉｓｈｉｎｇ Ａｓｓｏｃｉａｔｅｓ ａｎｄ Ｗｉｌｅｙ Ｉｎｔｅｒｓｃｉｅｎｃｅ，Ｎｅｗ Ｙｏｒｋ、及びＰＣＲ Ｐｒｏｔｏｃｏｌｓ：Ａ Ｇｕｉｄｅ ｔｏ Ｍｅｔｈｏｄｓ ａｎｄ Ａｐｐｌｉｃａｔｉｏｎｓ （Ｉｎｎｉｓ ｅｔ ａｌ．，１９９０，Ａｃａｄｅｍｉｃ Ｐｒｅｓｓ，Ｓａｎ Ｄｉｅｇｏ，Ｃａｌｉｆ．）に記載される技術を参照されたい。

【０１０４】

用語「遺伝子修飾」又は「遺伝子エンジニアリング」は、広義には、細胞のゲノム又は核酸の操作を指す。同様に、用語「遺伝子エンジニアリングされた」及び「エンジニアリングされた」は、操作されたゲノム又は核酸を含む細胞を指す。遺伝子修飾の方法としては、例えば、異種遺伝子発現、遺伝子又はプロモーターの挿入又は欠失、核酸変異、遺伝子発現又は不活性化の変化、酵素エンジニアリング、定向進化、知識ベースの設計、ランダム突然変異誘発法、遺伝子シャッフリング、及びコドン最適化が挙げられる。

【０１０５】

用語「組換え」は、核酸、タンパク質、又は細胞が遺伝子修飾、エンジニアリング、又は組換えの産物であることを示す。一般に、用語「組換え」は、複数の供給源に由来する遺伝物質を含むか、又はそれらによってコードされる、核酸、タンパク質、又は細胞を指す。本明細書で使用されるとき、用語「組換え」という用語はまた、内因性核酸又はタンパク質の変異型を含む、変異核酸又はタンパク質を含む細胞を説明するために使用され得る。用語「組換え細胞」及び「組換え宿主」は、同義的に使用することができる。一部の実施形態において、組換え細胞は、本明細書に開示されるＣＲＩＳＰＲエフェクターを含む。ＣＲＩＳＰＲエフェクターは組換え細胞における発現のためにコドン最適化することができる。一部の実施形態において、本明細書に開示される組換え細胞は、ＲＮＡガイドを更に含む。一部の実施形態において、本明細書に開示される組換え細胞のＲＮＡガイドは、ｔｒａｃｒＲＮＡを含む。一部の実施形態において、本明細書に開示される組換え細胞は、モジュレーターＲＮＡを含む。一部の実施形態において、組換え細胞は、大腸菌（Ｅ．ｃｏｌｉ）細胞などの原核細胞である。一部の実施形態において、組換え細胞は、ヒ

ト細胞を含む哺乳動物細胞などの真核細胞である。

【０１０６】

ＣＬＵＳＴ．０９１９７９の同定

この出願は、本明細書において「ＣＬＵＳＴ．０９１９７９」と呼ばれる新規タンパク質ファミリーの同定、エンジニアリング、及び使用に関する。図２に示すように、ＣＬＵＳＴ．０９１９７９のタンパク質は、ＲｕｖＣドメイン（ＲｕｖＣ　Ｉ、ＲｕｖＣ　ＩＩ、及びＲｕｖＣ　ＩＩＩと表示される）を含む。表５に示すように、ＣＬＵＳＴ．０９１９７９のエフェクターのサイズは、約７００アミノ酸～約８００アミノ酸の範囲である。したがって、以下に示すように、ＣＬＵＳＴ．０９１９７９のエフェクターは、当技術分野で知られているエフェクターよりも小さい。例えば、表１を参照されたい。

【０１０７】

【表１】

表1.公知のCRISPR-Cas システムエフェクターのサイズ。

エフェクター	サイズ(aa)
StCas9	1128
SpCas9	1368
SaCas9	1053
FnCpf1	1300
AsCpf1	1307
LbCpf1	1246
C2c1	1127 (平均)
CasX	982 (平均)
CasY	1189 (平均)
C2c2	1232 (平均)

【０１０８】

ＣＬＵＳＴ．０９１９７９のエフェクターは、他の特定の機能との強力な共出現パターンを呈するタンパク質を検索及び同定するために、計算方法及びアルゴリズムを使用して同定された。特定の実施形態において、これらの計算的方法は、ＣＲＩＳＰＲアレイにごく近接して共出現するタンパク質を同定することに関するものであった。本明細書に開示される方法は、非コード及びタンパク質コードの両方の（例えば、細菌遺伝子座の非コード範囲にあるファージ配列の断片；又はＣＲＩＳＰＲ　Ｃａｓ１タンパク質）、他の特徴にごく近接した範囲内に天然に出現するタンパク質の同定においても有用である。本明細書に記載される方法及び計算は１つ以上の計算装置で実施されてもよいことが理解される。

【０１０９】

ゲノム又はメタゲノムデータベースから一組のゲノム配列が入手された。データベースは、ショートリード、又はコンティグレベルデータ、又はアセンブルされたスキャフォールド、又は生物の完全ゲノム配列を含んだ。同様に、データベースは、原核生物、若しくは真核生物からのゲノム配列データを含んでもよく、又はメタゲノム環境試料からのデータを含んでもよい。データベースリポジトリの例としては、国立バイオテクノロジー情報

センター（Ｎａｔｉｏｎａｌ Ｃｅｎｔｅｒ ｆｏｒ Ｂｉｏｔｅｃｈｎｏｌｏｇｙ Ｉｎｆｏｒｍａｔｉｏｎ：ＮＣＢＩ）のＲｅｆＳｅｑ、ＮＣＢＩのＧｅｎＢａｎｋ、ＮＣＢＩの全ゲノムショットガン（Ｗｈｏｌｅ Ｇｅｎｏｍｅ Ｓｈｏｔｇｕｎ：ＷＧＳ）、及びジョイントゲノム研究所（Ｊｏｉｎｔ Ｇｅｎｏｍｅ Ｉｎｓｔｉｔｕｔｅ：ＪＧＩ）の統合微生物ゲノム（Ｉｎｔｅｇｒａｔｅｄ Ｍｉｃｒｏｂｉａｌ Ｇｅｎｏｍｅｓ：ＩＭＧ）が挙げられる。

【０１１０】

一部の実施形態において、指定される最小長さのゲノム配列データの選択には、最小サイズ要件が課される。特定の例示的実施形態において、最小コンティグ長さは、１００ヌクレオチド、５００ｎｔ、１ｋｂ、１．５ｋｂ、２ｋｂ、３ｋｂ、４ｋｂ、５ｋｂ、１０ｋｂ、２０ｋｂ、４０ｋｂ、又は５０ｋｂであってもよい。

【０１１１】

一部の実施形態において、公知の又は予測されるタンパク質は、完全な又は選択された一組のゲノム配列データから抽出される。一部の実施形態において、公知の又は予測されるタンパク質は、ソースデータベースによって提供されるコード配列（ＣＤＳ）アノテーションを抽出することから取られる。一部の実施形態において、予測タンパク質は、計算的方法を適用してヌクレオチド配列からタンパク質を同定することにより決定される。一部の実施形態では、ＧｅｎｅＭａｒｋスイートを使用してゲノム配列からタンパク質が予測される。一部の実施形態では、Ｐｒｏｄｉｇａｌを使用してゲノム配列からタンパク質が予測される。一部の実施形態では、同じ一組の配列データに対して複数のタンパク質予測アルゴリズムが用いられ、得られる一組のタンパク質から重複が排除されてもよい。

【０１１２】

一部の実施形態において、ＣＲＩＳＰＲアレイはゲノム配列データから同定される。一部の実施形態では、ＰＩＬＥＲ－ＣＲを使用してＣＲＩＳＰＲアレイが同定される。一部の実施形態では、ＣＲＩＳＰＲ認識ツール（ＣＲＩＳＰＲ Ｒｅｃｏｇｎｉｔｉｏｎ Ｔｏｏｌ：ＣＲＴ）を使用してＣＲＩＳＰＲアレイが同定される。一部の実施形態において、ＣＲＩＳＰＲアレイは、最小限の回数（例えば２、３、又は４回）繰り返されるヌクレオチドモチーフを同定する発見的手法によって同定され、ここで繰り返されるモチーフの連続する出現間の間隔は、指定される長さ（例えば、５０、１００、又は１５０ヌクレオチド）を超えない。一部の実施形態では、同じ一組の配列データに対して複数のＣＲＩＳＰＲアレイ同定ツールが用いられ、得られる一組のＣＲＩＳＰＲアレイから重複が排除されてもよい。

【０１１３】

一部の実施形態において、ＣＲＩＳＰＲアレイ（本明細書では「ＣＲＩＳＰＲ近位タンパク質クラスター」と呼ばれる）にごく近接しているタンパク質が同定される。一部の実施形態において、近接性はヌクレオチド距離として定義され、２０ｋｂ、１５ｋｂ、又は５ｋｂ以内であってもよい。一部の実施形態において、近接性は、タンパク質とＣＲＩＳＰＲアレイとの間にあるオープンリーディングフレーム（ＯＲＦ）の数として定義され、特定の例示的距離は、１０、５、４、３、２、１、又は０個のＯＲＦであり得る。ＣＲＩＳＰＲアレイとごく近接した範囲内にあると同定されたタンパク質は、次に相同タンパク質クラスターにまとめられる。一部の実施形態において、ｂｌａｓｔｃｌｕｓｔを使用してＣＲＩＳＰＲ近位タンパク質クラスターが形成される。特定の他の実施形態において、ｍｍｓｅｑｓ２を使用してＣＲＩＳＰＲ近位タンパク質クラスターが形成される。

【０１１４】

ＣＲＩＳＰＲ近接タンパク質クラスターのメンバー間に強力な共出現パターンを確立するため、予め編成された完全な一組の公知の及び予測されるタンパク質に対してタンパク質クラスターの各メンバーのＢＬＡＳＴ検索が実施されてもよい。一部の実施形態では、ＵＢＬＡＳＴ又はｍｍｓｅｑｓ２を使用して類似のタンパク質が検索されてもよい。一部の実施形態では、ファミリー内のタンパク質の代表的なサブセットについてのみ検索が実施されてもよい。

【０１１５】
一部の実施形態では、ＣＲＩＳＰＲ近接タンパク質クラスターがメトリックによって順位付けされるか又はフィルタリングされることにより共出現が決定される。１つの例示的メトリックは、特定のＥ値閾値に至るまでのＢＬＡＳＴマッチの数に対するタンパク質クラスター内の要素の数の比である。一部の実施形態では、一定のＥ値閾値が使用されてもよい。他の実施形態では、Ｅ値閾値は、タンパク質クラスターの最も離れたメンバーによって決定されてもよい。一部の実施形態において、大域的な一組のタンパク質がクラスター化され、共出現メトリックは、含まれる１つ又は複数の大域的クラスターの要素の数に対するＣＲＩＳＰＲ近接タンパク質の要素の数の比である。
【０１１６】
一部の実施形態において、手動でのレビュープロセスを用いることにより、クラスター中のタンパク質の天然に存在する遺伝子座構造に基づいてエンジニアリングされるシステムの潜在的機能性及び最小限の一組の成分が評価される。一部の実施形態において、手動でのレビューにはタンパク質クラスターの図解表現が役立ち得るとともに、これは、ペアワイズでの配列類似性、系統樹、供給源生物／環境、予測される機能性ドメイン、及び遺伝子座構造の図解描写を含む情報を含み得る。一部の実施形態において、遺伝子座構造の図解描写は、高い代表性を有する近隣タンパク質ファミリーをフィルタリングし得る。一部の実施形態において、代表性は、含んでいる１つ又は複数の大域的クラスターの１つ又は複数のサイズに対する関連する近隣タンパク質の数の比によって計算されてもよい。特定の例示的実施形態において、タンパク質クラスターの図解表現は、天然に存在する遺伝子座のＣＲＩＳＰＲアレイ構造の描写を含み得る。一部の実施形態において、タンパク質クラスターの図解表現は、推定ＣＲＩＳＰＲアレイの長さに対する保存されたダイレクトリピートの数、又は推定ＣＲＩＳＰＲアレイの長さに対するユニークなスペーサー配列の数の描写を含み得る。一部の実施形態において、タンパク質クラスターの図解表現は、ＣＲＩＳＰＲアレイとの推定エフェクターの様々な共出現メトリックの描写を含み、新規ＣＲＩＳＰＲ－Ｃａｓシステムを予測し、及びその成分を同定し得る。
【０１１７】
ＣＬＵＳＴ．０９１９７９のプール型スクリーニング
本明細書で同定された、エンジニアリングされたＣＬＵＳＴ．０９１９７９ ＣＲＩＳＰＲ－Ｃａｓシステムの活性、メカニズム、及び機能パラメータを効率的に検証するために、実施例４で説明されるように、大腸菌（Ｅ．ｃｏｌｉ）におけるプール型スクリーニングアプローチを使用した。第一に、ＣＲＩＳＴ．０９１９７９ＣＲＩＳＰＲ－Ｃａｓシステムの保存タンパク質及び非コードエレメントの計算的同定から、ＤＮＡ合成及び分子クローニングを用いて個別の成分を単一の人工発現ベクター（一実施形態ではｐＥＴ－２８ａ＋骨格をベースとする）にアセンブルする。第２の実施形態では、エフェクター及び非コードエレメントをｍＲＮＡ転写物に転写し、異なるリボソーム結合部位を用いて個々のエフェクターを翻訳する。
【０１１８】
第二に、天然のｃｒＲＮＡ及びターゲティングスペーサーを、第２のプラスミドｐＡＣＹＣ１８４を標的化する非天然スペーサーを含むプロセシングされていないｃｒＲＮＡのライブラリに置き換える。このｃｒＲＮＡライブラリをエフェクター及び非コードエレメントを含むベクター骨格（例えばｐＥＴ－２８ａ＋）にクローニングし、その後、続いてこのライブラリをｐＡＣＹＣ１８４プラスミド標的と共に大腸菌（Ｅ．ｃｏｌｉ）に形質転換する。結果的に、得られる各大腸菌（Ｅ．ｃｏｌｉ）細胞は、ただ１つのターゲティングアレイを含む。代替的実施形態では、非天然スペーサーを含むプロセシングされていないｃｒＲＮＡのライブラリが、Ｂａｂａ ｅｔ ａｌ．（２００６）Ｍｏｌ．Ｓｙｓｔ．Ｂｉｏｌ．２：２００６．０００８；及びＧｅｒｄｅｓ ｅｔ ａｌ．（２００３）Ｊ．Ｂａｃｔｅｒｉｏｌ．１８５（１９）：５６７３－８４（これらの各々の内容全体は参照により本明細書に援用される）に記載されるものなどの資料から引用される大腸菌（Ｅ．ｃｏｌｉ）必須遺伝子を更に標的化する。この実施形態において、必須遺伝子機能を破壊す

る新規ＣＲＩＳＰＲ－Ｃａｓシステムの正の標的化された活性は、細胞死又は成長停止を生じさせる。一部の実施形態において、必須遺伝子ターゲティングスペーサーをｐＡＣＹＣ１８４標的と組み合わせることができる。

【０１１９】

第三に、抗生物質選択下で大腸菌（Ｅ．ｃｏｌｉ）を成長させる。一実施形態において、三重抗生物質選択：エンジニアリングされたＣＲＩＳＰＲエフェクターシステムを含むｐＥＴ－２８ａ＋ベクターの形質転換の成功を確認するためのカナマイシン、並びにｐＡＣＹＣ１８４標的ベクターの同時形質転換の成功を確認するためのクロラムフェニコール及びテトラサイクリンが用いられる。ｐＡＣＹＣ１８４は通常、クロラムフェニコール及びテトラサイクリンに対する耐性を付与するため、抗生物質選択下では、このプラスミドを標的化する新規ＣＲＩＳＰＲ－Ｃａｓシステムの正の活性により、エフェクター、非コードエレメント、及びｃｒＲＮＡライブラリの特異的活性エレメントを活性に発現する細胞が排除されることになる。典型的には、生存細胞の集団は、形質転換の１２～１４時間後に分析される。一部の実施形態において、生存細胞の分析は、形質転換後６～８時間、形質転換後８～１２時間、形質転換後最大２４時間、又は形質転換後２４時間を超えて、行われる。早い時点と比較した後の時点における生存細胞集団を調べると、不活性ｃｒＲＮＡと比較してシグナルの枯渇が生じる。

【０１２０】

一部の実施形態において、二重抗生物質選択が用いられる。クロラムフェニコール又はテトラサイクリンのいずれかを抜き取って選択圧を除去すると、ターゲティング基質、配列特異性、及び効力に関する新規情報を得ることができる。例えば、選択された遺伝子又は選択されていない遺伝子におけるｄｓＤＮＡの切断により、選択された遺伝子及び選択されていない遺伝子の両方の枯渇が観察される大腸菌（Ｅ．ｃｏｌｉ）におけるネガティブ選択が生じ得る。ＣＲＩＳＰＲ－Ｃａｓシステムが転写又は翻訳に干渉する場合（例えば、結合又は転写物の切断によって）、選択は、選択されていない耐性遺伝子というよりむしろ、選択された耐性遺伝子の標的に対してのみ観察される。

【０１２１】

一部の実施形態では、カナマイシンのみを使用して、エンジニアリングされたＣＲＩＳＰＲ－Ｃａｓシステムを含むｐＥＴ－２８ａ＋ベクターの形質転換の成功が確認される。この実施形態は、成長の変化を観察するためにカナマイシン以外の更なる選択が必要ないため、大腸菌（Ｅ．ｃｏｌｉ）必須遺伝子を標的化するスペーサーを含むライブラリに好適である。この実施形態では、クロラムフェニコール及びテトラサイクリン依存性が取り除かれ、ライブラリ中のそれらの標的（存在する場合）が、ターゲティング基質、配列特異性、及び効力に関するネガティブ又はポジティブの更なる情報源を提供する。

【０１２２】

ｐＡＣＹＣ１８４プラスミドは、ＣＲＩＳＰＲ－Ｃａｓシステムの活性に影響を及ぼし得る多様な一組の特徴及び配列を含むため、プール型スクリーンからの活性ｃｒＲＮＡをｐＡＣＹＣ１８４にマッピングすることにより、種々の活性機構及び機能パラメータを示唆するものであり得る活性パターンが提供される。このようにして、異種原核生物種における新規ＣＲＩＳＰＲ－Ｃａｓシステムの再構成に必要な特徴をより包括的に試験し、研究することができる。

【０１２３】

本明細書に記載されるインビボプール型スクリーンの重要な利点としては、以下が挙げられる：

（１）汎用性－プラスミド設計により、複数のエフェクター及び／又は非コードエレメントを発現させることが可能になる；ライブラリクローニング戦略により、計算的に予測されたｃｒＲＮＡの両方の転写方向の発現が実現する；

（２）活性機構及び機能パラメータの包括的試験により、核酸切断を含めた多様な干渉機構を評価し；転写、プラスミドＤＮＡ複製などの特徴の共出現；及びｃｒＲＮＡライブラリについてのフランキング配列を調べて、４Ｎの複雑さ等価のＰＡＭを確実に決定する

ことができる；
（３）感度－ｐＡＣＹＣ１８４は低コピープラスミドであり、僅かな干渉率であっても、プラスミドによってコードされる抗生物質耐性を除去することができるため、ＣＲＩＳＰＲ－Ｃａｓ活性について高感度を実現する；及び
（４）効率－ＲＮＡシーケンシングについてより高い速度及びスループットを実現する最適化された分子生物学ステップは、タンパク質発現試料をスクリーンにおける生存細胞から直接採取することができる。
【０１２４】
このインビボプール型スクリーンを用いてその作動可能なエレメント、機構及びパラメータ、並びにその内因性細胞環境の外部でエンジニアリングされたシステムにおいて活性であり再プログラム化されるその能力を評価することにより、本明細書に記載される新規ＣＲＩＳＴ．０９１９７９ＣＲＩＳＰＲ－Ｃａｓファミリーを評価した。
【０１２５】
ＣＲＩＳＰＲエフェクターの活性及び修飾
一部の実施形態において、ＣＲＩＳＴ．０９１９７９のＣＲＩＳＰＲエフェクター及びＲＮＡガイドは、他の成分を含み得る二元複合体を形成する。二元複合体は、ＲＮＡガイド中のスペーサー配列に相補的な核酸基質（即ち、配列特異的基質又は標的核酸）への結合時に活性化される。一部の実施形態において、配列特異的基質は二本鎖ＤＮＡである。一部の実施形態において、配列特異的基質は一本鎖ＤＮＡである。一部の実施形態において、配列特異的基質は一本鎖ＲＮＡである。一部の実施形態において、配列特異的基質は二本鎖ＲＮＡである。一部の実施形態において、配列特異性は、ＲＮＡガイド（例えば、ｃｒＲＮＡ）中のスペーサー配列と標的基質の完全な一致を必要とする。他の実施形態において、配列特異性は、ＲＮＡガイド（例えば、ｃｒＲＮＡ）中のスペーサー配列と標的基質の部分的な（連続的又は非連続的な）一致を必要とする。
【０１２６】
一部の実施形態において、本発明のＣＲＩＳＰＲエフェクターは酵素活性、例えば、ヌクレアーゼ活性を広範なｐＨ条件にわたり有する。一部の実施形態において、ヌクレアーゼは酵素活性、例えば、ヌクレアーゼ活性を約３．０～約１２．０のｐＨにおいて有する。一部の実施形態において、ＣＲＩＳＰＲエフェクターは酵素活性を約４．０～約１０．５のｐＨにおいて有する。一部の実施形態において、ＣＲＩＳＰＲエフェクターは酵素活性を約５．５～約８．５のｐＨにおいて有する。一部の実施形態において、ＣＲＩＳＰＲエフェクターは酵素活性を約６．０～約８．０のｐＨにおいて有する。一部の実施形態において、ＣＲＩＳＰＲエフェクターは酵素活性を約７．０のｐＨにおいて有する。
【０１２７】
一部の実施形態において、本発明のＣＲＩＳＰＲエフェクターは酵素活性、例えば、ヌクレアーゼ活性を約１０℃～約１００℃の温度範囲において有する。一部の実施形態において、本発明のＣＲＩＳＰＲエフェクターは酵素活性を約２０℃～約９０℃の温度範囲において有する。一部の実施形態において、本発明のＣＲＩＳＰＲエフェクターは酵素活性を約２０℃～約２５℃の温度において又は約３７℃の温度において有する。
【０１２８】
一部の実施形態において、二元複合体は標的基質への結合時に活性化した状態になる。一部の実施形態において、活性化した複合体は、「複数回の代謝回転」活性を呈し、従って標的基質への作用時（例えば、それの切断時）、活性化した複合体は活性化した状態のままである。一部の実施形態において、活性化した二元複合体は「単回の代謝回転」活性を呈し、従って標的基質への作用時、二元複合体は不活性状態に戻る。一部の実施形態において、活性化した二元複合体は非特異的な（即ち、「コラテラル」）切断活性を呈し、従って複合体は非標的核酸を切断する。一部の実施形態において、非標的核酸は、ＤＮＡ分子（例えば、一本鎖又は二本鎖ＤＮＡ）である。一部の実施形態において、非標的核酸は、ＲＮＡ分子（例えば、一本鎖又は二本鎖ＲＮＡ）である。
【０１２９】

本発明のＣＲＩＳＰＲエフェクターが標的核酸（例えば、ゲノムＤＮＡ）中で二本鎖分解又は一本鎖分解を誘導する一部の実施形態において、二本鎖分解は、相同組換え（ＨＤＲ）、非相同末端結合（ＮＨＥＪ）、又は代替非相同末端結合（Ａ－ＮＨＥＪ）を含む細胞内因性ＤＮＡ修復経路を刺激し得る。ＮＨＥＪは相同性鋳型を必要とせずに、切断された標的核酸を修復し得る。これにより、標的遺伝子座における１つ以上のヌクレオチドの欠失又は挿入が生じ得る。ＨＤＲは、ドナーＤＮＡなどの相同性鋳型を用いて起こり得る。相同性鋳型は、標的核酸切断部位をフランキングする配列と相同の配列を含み得る。一部の例において、ＨＤＲは外因性ポリヌクレオチド配列を切断標的遺伝子座中に挿入し得る。ＮＨＥＪ及び／又はＨＤＲに起因する標的ＤＮＡの修飾は、例えば、突然変異、欠失、変更、組込み、遺伝子修正、遺伝子置換、遺伝子タグ付け、トランス遺伝子ノックイン、遺伝子破壊、及び／又は遺伝子ノックアウトをもたらし得る。

【０１３０】

一部の実施形態において、本明細書に記載されるＣＲＩＳＰＲエフェクターは、Ｈｉｓタグ、ＧＳＴタグ、ＦＬＡＧタグ、又はｍｙｃタグを含めた１つ以上のペプチドタグに融合することができる。一部の実施形態において、本明細書に記載されるＣＲＩＳＰＲエフェクターは、蛍光タンパク質（例えば、緑色蛍光タンパク質又は黄色蛍光タンパク質）など、検出可能部分に融合することができる。一部の実施形態において、本開示のＣＲＩＳＰＲエフェクター及び／又はアクセサリータンパク質は、タンパク質を組織、細胞、又は細胞の領域に侵入又は局在化させるペプチド又は非ペプチド部分に融合される。例えば、本開示のＣＲＩＳＰＲエフェクターは、ＳＶ４０（シミアンウイルス４０）ＮＬＳ、ｃ－Ｍｙｃ　ＮＬＳ、又は他の好適な単節型ＮＬＳなどの核局在化配列（ＮＬＳ）を含んでもよい。ＮＬＳは、ＣＲＩＳＰＲエフェクターのＮ末端及び／又はＣ末端に融合されてもよく、且つ単独で融合されても（即ち、単一のＮＬＳ）、又はコンカテマー化されてもよい（例えば、２、３、４個等のＮＬＳの鎖）。

【０１３１】

一部の実施形態において、少なくとも１つの核外輸送シグナル（ＮＥＳ）が、ＣＲＩＳＰＲエフェクターをコードする核酸配列に取り付けられている。一部の実施形態において、Ｃ末端及び／又はＮ末端ＮＬＳ又はＮＥＳは、真核細胞、例えばヒト細胞における最適な発現及び核ターゲティングのために取り付けられる。

【０１３２】

ＣＲＩＳＰＲエフェクターにタグが融合される実施形態において、かかるタグは、例えば、液体クロマトグラフィー又は固定化したアフィニティー若しくはイオン交換試薬を利用するビーズ分離による、ＣＲＩＳＰＲエフェクターの親和性ベース又は電荷ベースの精製を促進し得る。非限定的な例として、本開示の組換えＣＲＩＳＰＲエフェクターはポリヒスチジン（Ｈｉｓ）タグを含み、精製のため、固定化された金属イオンを含むクロマトグラフィーカラムにロードされる（例えば、樹脂上に固定化されたキレートリガンドによってキレートされたZn^{2+}、Ni^{2+}、Cu^{2+}イオン、この樹脂は、個々に調製された樹脂又は市販の樹脂若しくはＧＥ　Ｈｅａｌｔｈｃａｒｅ　Ｌｉｆｅ　Ｓｃｉｅｎｃｅｓ、Ｍａｒｌｂｏｒｏｕｇｈ，Ｍａｓｓａｃｈｕｓｅｔｔｓによって商品化されているＨｉｓＴｒａｐ　ＦＦカラムなどの既製のカラムであってもよい。ローディングステップの後、カラムは任意選択で、例えば１つ以上の好適な緩衝溶液を使用してリンスされ、次にＨｉｓタグが付加されたタンパク質が好適な溶出緩衝液を使用して溶出される。それに代えて又は加えて、本開示の組換えＣＲＩＳＰＲエフェクターがＦＬＡＧタグを利用する場合、かかるタンパク質は、本業界で公知の免疫沈降法を用いて精製されてもよい。タグが付加された本開示のＣＲＩＳＰＲエフェクター又はアクセサリータンパク質について他の好適な精製方法が当業者には明らかであろう。

【０１３３】

本明細書に記載されるタンパク質（例えば、ＣＲＩＳＰＲエフェクター又はアクセサリータンパク質）は、核酸分子又はポリペプチドのいずれとしても送達又は使用することができる。核酸分子を使用する場合、ＣＲＩＳＰＲエフェクターをコードする核酸分子をコ

ドン最適化することができる。核酸は、任意の目的の生物、詳細にはヒト細胞又は細菌での使用にコドン最適化することができる。例えば、核酸は、マウス、ラット、ウサギ、イヌ、家畜、又は非ヒト霊長類を含めた任意の非ヒト真核生物向けにコドン最適化することができる。コドン使用表が、例えば、ｗｗｗ．ｋａｚｕｓａ．ｏｒｊｐ／ｃｏｄｏｎ／で利用可能な「コドン使用データベース（Ｃｏｄｏｎ　Ｕｓａｇｅ　Ｄａｔａｂａｓｅ）」において容易に利用可能であり、これらの表を幾つもの方法で適合させることができる。Ｎａｋａｍｕｒａ　ｅｔ　ａｌ．Ｎｕｃｌ．Ａｃｉｄｓ　Ｒｅｓ．２８：２９２（２０００）（全体として参照により本明細書に援用される）を参照のこと。特定の配列を特定の宿主細胞での発現にコドン最適化するためのコンピュータアルゴリズムもまた、Ｇｅｎｅ　Ｆｏｒｇｅ（Ａｐｔａｇｅｎ；Ｊａｃｏｂｕｓ，ＰＡ）など、利用可能である。

【０１３４】

一部の例では、真核生物（例えば、ヒト、又は他の哺乳類細胞）細胞で発現させるためのＣＲＩＳＰＲエフェクターをコードする本開示の核酸は、１つ以上のイントロン、即ち、第１の端部（例えば、５’末端）にスプライスドナー配列を含み、且つ第２の端部（例えば、３’末端）にスプライスアクセプター配列を含む１つ以上の非コード配列を含む。本開示の様々な実施形態において、限定なしに、シミアンウイルス４０（ＳＶ４０）イントロン、β－グロビンイントロン、及び合成イントロンを含め、任意の好適なスプライスドナー／スプライスアクセプターを使用することができる。それに代えて又は加えて、ＣＲＩＳＰＲエフェクター又はアクセサリータンパク質をコードする本開示の核酸は、ＤＮＡコード配列の３’末端に、ポリアデニル化（ポリＡ）シグナルなどの転写終結シグナルを含み得る。一部の例では、ポリＡシグナルは、ＳＶ４０イントロンなどのイントロンにごく近接して、又はそれに隣接して位置する。

【０１３５】

非活性化／不活性化ＣＲＩＳＰＲエフェクター

本明細書に記載されるＣＲＩＳＰＲエフェクターが、減少したヌクレアーゼ活性、例えば、野生型ＣＲＩＳＰＲエフェクターと比較したとき少なくとも５０％、少なくとも６０％、少なくとも７０％、少なくとも８０％、少なくとも９０％、少なくとも９５％、少なくとも９７％、又は１００％のヌクレアーゼ不活性化となるようにＣＲＩＳＰＲ酵素を修飾することができる。ヌクレアーゼ活性は、当技術分野で周知の幾つかの方法、例えば、タンパク質のヌクレアーゼドメインへの突然変異の導入によって減少させることができる。一部の実施形態において、ヌクレアーゼ活性の触媒残基が同定され、それらのアミノ酸残基を異なるアミノ酸残基（例えば、グリシン又はアラニン）に置換することによりヌクレアーゼ活性を減少させてもよい。

【０１３６】

不活性化されたＣＲＩＳＰＲエフェクターは、１つ以上の機能的ドメインを含むか、又はそれと関連づけられ得る（例えば、融合タンパク質、リンカーペプチド、「ＧＳ」リンカーなどを介して）。こうした機能性ドメインは様々な活性、例えば、メチラーゼ活性、デメチラーゼ活性、転写活性化活性、転写抑制活性、転写放出因子活性、ヒストン修飾活性、ＲＮＡ切断活性、ＤＮＡ切断活性、核酸結合活性、及びスイッチ活性（例えば、光誘導性）を有することができる。一部の実施形態において、機能性ドメインは、クルッペル関連ボックス（ＫＲＡＢ）、ＶＰ６４、ＶＰ１６、Ｆｏｋ１、Ｐ６５、ＨＳＦ１、ＭｙｏＤ１、及びビオチン－ＡＰＥＸである。

【０１３７】

不活性化されたＣＲＩＳＰＲエフェクター上に１つ以上の機能性ドメインを位置させることは、その機能性ドメインが帰属する機能的効果による影響を標的に及ぼすのに正しい空間上の向きとなることを可能にする。例えば、機能性ドメインが転写活性化因子（例えば、ＶＰ１６、ＶＰ６４、又はｐ６５）である場合、その転写活性化因子は、それが標的の転写に影響を及ぼすことが可能になる空間上の向きに置かれる。同様に、転写リプレッサーが標的の転写に影響を及ぼすように位置し、及びヌクレアーゼ（例えば、Ｆｏｋ１）が標的を切断又は部分的に切断するように位置する。一部の実施形態において、機能性ド

メインはＣＲＩＳＰＲエフェクターのＮ末端に位置する。一部の実施形態において、機能性ドメインはＣＲＩＳＰＲエフェクターのＣ末端に位置する。一部の実施形態において、不活性化されたＣＲＩＳＰＲエフェクターは、Ｎ末端に第１の機能性ドメイン及びＣ末端に第２の機能性ドメインを含むように修飾される。

【０１３８】

スプリット酵素

本開示はまた、本明細書に記載されるＣＲＩＳＰＲエフェクターのスプリットバージョンも提供する。スプリットバージョンのＣＲＩＳＰＲエフェクターは送達に有利であり得る。一部の実施形態において、ＣＲＩＳＰＲエフェクターは酵素の２つの部分に分割され、それらが一緒になって実質的に機能性のＣＲＩＳＰＲエフェクターを含む。

【０１３９】

分割は、１つ又は複数の触媒ドメインが影響を受けないような方法で行われ得る。ＣＲＩＳＰＲエフェクターはヌクレアーゼとして機能してもよく、又は本質的に（例えば、その触媒ドメインにある１つ又は複数の突然変異に起因して）触媒活性がごく僅かしかない又は全くないＲＮＡ結合タンパク質である不活性化された酵素であってもよい。

【０１４０】

一部の実施形態では、ヌクレアーゼローブ及びα－ヘリックスローブが別個のポリペプチドとして発現する。これらのローブはそれ自体には相互作用を及ぼさないが、ＲＮＡガイドがそれらを複合体へと動員し、その複合体が完全長ＣＲＩＳＰＲエフェクターの活性を再現し、部位特異的ＤＮＡ切断を触媒する。修飾されたＲＮＡガイドを使用すると、二量化が妨げられることによりスプリット酵素活性が無効になり、誘導性の二量化システムの開発が可能となる。スプリット酵素については、例えば、Ｗｒｉｇｈｔ ｅｔ ａｌ．“Ｒａｔｉｏｎａｌ ｄｅｓｉｇｎ ｏｆ ａ ｓｐｌｉｔ－Ｃａｓ９ ｅｎｚｙｍｅ ｃｏｍｐｌｅｘ，”Ｐｒｏｃ．Ｎａｔ’ｌ．Ａｃａｄ．Ｓｃｉ．，１１２．１０（２０１５）：２９８４－２９８９（全体として参照により本明細書に援用される）に記載されている。

【０１４１】

一部の実施形態において、スプリット酵素は、例えばラパマイシン感受性二量化ドメインを利用することにより、二量化パートナーに融合されてもよい。これにより、ＣＲＩＳＰＲエフェクター活性を時間的に制御するための化学誘導性ＣＲＩＳＰＲエフェクターの作成が可能になる。このようにして２つの断片に分割されていることによりＣＲＩＳＰＲエフェクターを化学誘導性にすることができ、ＣＲＩＳＰＲエフェクターの制御された再アセンブルにはラパマイシン感受性二量化ドメインを使用することができる。

【０１４２】

分割点は、典型的にはインシリコで設計され、コンストラクトにクローニングされる。この過程でスプリット酵素に突然変異が導入されてもよく、非機能性ドメインが除去されてもよい。一部の実施形態において、スプリットＣＲＩＳＰＲエフェクターの２つの部分又は断片（即ち、Ｎ末端及びＣ末端断片）は、例えば野生型ＣＲＩＳＰＲエフェクターの配列の少なくとも７０％、少なくとも８０％、少なくとも９０％、少なくとも９５％、又は少なくとも９９％を含む完全なＣＲＩＳＰＲエフェクターを形成することができる。

【０１４３】

自己活性化型又は不活性化型酵素

本明細書に記載されるＣＲＩＳＰＲエフェクターは、自己活性化型又は自己不活性化型であるように設計されてもよい。一部の実施形態において、ＣＲＩＳＰＲエフェクターは自己不活性化型である。例えば、ＣＲＩＳＰＲエフェクターをコードするコンストラクトに標的配列を導入することができる。従ってＣＲＩＳＰＲエフェクターが標的配列を切断するとともに、それによって酵素をコードするコンストラクトがその発現を自己不活性化し得る。自己不活性化ＣＲＩＳＰＲシステムの構築方法については、例えば、Ｅｐｓｔｅｉｎ ｅｔ ａｌ．，“Ｅｎｇｉｎｅｅｒｉｎｇ ａ Ｓｅｌｆ－Ｉｎａｃｔｉｖａｔｉｎｇ ＣＲＩＳＰＲ Ｓｙｓｔｅｍ ｆｏｒ ＡＡＶ Ｖｅｃｔｏｒｓ，”Ｍｏｌ．Ｔｈｅｒ．，２４（２０１６）：Ｓ５０（全体として参照により本明細書に援用される）に記載されてい

る。
【０１４４】
一部の他の実施形態では、弱いプロモーター（例えば、７ＳＫプロモーター）の制御下で発現する更なるＲＮＡガイドが、ＣＲＩＳＰＲエフェクターをコードする核酸配列を標的化して、その発現を（例えば、核酸の転写及び／又は翻訳を妨げることにより）妨げ及び／又は阻止することができる。ＣＲＩＳＰＲエフェクターと、ＲＮＡガイドと、ＣＲＩＳＰＲエフェクターをコードする核酸を標的化するＲＮＡガイドとを発現するベクターを細胞にトランスフェクトすると、ＣＲＩＳＰＲエフェクターをコードする核酸の効率的な破壊につながり、ＣＲＩＳＰＲエフェクターレベルを低下させることができ、従ってゲノム編集活性を制限することができる。
【０１４５】
一部の実施形態において、ＣＲＩＳＰＲエフェクターのゲノム編集活性は、哺乳類細胞における内因性ＲＮＡシグネチャ（例えば、ｍｉＲＮＡ）を通じて調節することができる。ＣＲＩＳＰＲエフェクターをコードするｍＲＮＡの５’－ＵＴＲにｍｉＲＮＡ相補配列を用いることにより、ＣＲＩＳＰＲエフェクタースイッチを作ることができる。このスイッチは、標的細胞中のｍｉＲＮＡに選択的且つ効率的に応答する。従って、このスイッチは、異種細胞集団内で内因性ｍｉＲＮＡ活性を感知することによりゲノム編集を差次的に制御し得る。従って、このスイッチシステムは、細胞内ｍｉＲＮＡ情報に基づく細胞型選択的なゲノム編集及び細胞エンジニアリングのフレームワークを提供し得る（Ｈｉｒｏｓａｗａ ｅｔ ａｌ.“Ｃｅｌｌ－ｔｙｐｅ－ｓｐｅｃｉｆｉｃ ｇｅｎｏｍｅ ｅｄｉｔｉｎｇ ｗｉｔｈ ａ ｍｉｃｒｏＲＮＡ－ｒｅｓｐｏｎｓｉｖｅ ＣＲＩＳＰＲ－Ｃａｓ９ ｓｗｉｔｃｈ，”Ｎｕｃｌ．Ａｃｉｄｓ Ｒｅｓ．，２０１７ Ｊｕｌ ２７；４５（１３）：ｅ１１８）。
【０１４６】
誘導性ＣＲＩＳＰＲエフェクター
ＣＲＩＳＰＲエフェクターは、誘導性、例えば、光誘導性又は化学誘導性であってもよい。この機構により、ＣＲＩＳＰＲ酵素中の機能性ドメインを活性化させることが可能になる。光誘導能は、当該技術分野において公知の様々な方法により、例えば、スプリットＣＲＩＳＰＲエフェクターにおいてＣＲＹ２ＰＨＲ／ＣＩＢＮ対が用いられる融合複合体を設計することにより実現し得る（例えば、Ｋｏｎｅｒｍａｎｎ ｅｔ ａｌ.“Ｏｐｔｉｃａｌ ｃｏｎｔｒｏｌ ｏｆ ｍａｍｍａｌｉａｎ ｅｎｄｏｇｅｎｏｕｓ ｔｒａｎｓｃｒｉｐｔｉｏｎ ａｎｄ ｅｐｉｇｅｎｅｔｉｃ ｓｔａｔｅｓ，”Ｎａｔｕｒｅ，５００．７４６３（２０１３）：４７２を参照のこと）。化学誘導能は、例えば、スプリットＣＲＩＳＰＲエフェクターにおいてＦＫＢＰ／ＦＲＢ（ＦＫ５０６結合タンパク質／ＦＫＢＰラパマイシン結合ドメイン）対が用いられる融合複合体を設計することにより実現し得る。ラパマイシンは融合複合体の形成に必要であり、従ってＣＲＩＳＰＲエフェクターを活性化する（例えば、Ｚｅｔｓｃｈｅ ｅｔ ａｌ.“Ａ ｓｐｌｉｔ－Ｃａｓ９ ａｒｃｈｉｔｅｃｔｕｒｅ ｆｏｒ ｉｎｄｕｃｉｂｌｅ ｇｅｎｏｍｅ ｅｄｉｔｉｎｇ ａｎｄ ｔｒａｎｓｃｒｉｐｔｉｏｎ ｍｏｄｕｌａｔｉｏｎ，”Ｎａｔｕｒｅ Ｂｉｏｔｅｃｈ．，３３．２（２０１５）：１３９－１４２を参照のこと）。
【０１４７】
更に、ＣＲＩＳＰＲエフェクターの発現は、誘導性プロモーター、例えば、テトラサイクリン又はドキシサイクリン制御下での転写活性化（Ｔｅｔ－Ｏｎ及びＴｅｔ－Ｏｆｆ発現システム）、ホルモン誘導性遺伝子発現システム（例えば、エクジソン誘導性遺伝子発現システム）、及びアラビノース誘導性遺伝子発現システムによって調節することができる。ＲＮＡとして送達される場合、ＲＮＡターゲティングエフェクタータンパク質の発現は、小分子様テトラサイクリンを感知することのできるリボスイッチによって調節されてもよい（例えば、Ｇｏｌｄｆｌｅｓｓ ｅｔ ａｌ.“Ｄｉｒｅｃｔ ａｎｄ ｓｐｅｃｉｆｉｃ ｃｈｅｍｉｃａｌ ｃｏｎｔｒｏｌ ｏｆ ｅｕｋａｒｙｏｔｉｃ ｔｒａｎｓｌａｔｉｏｎ ｗｉｔｈ ａ ｓｙｎｔｈｅｔｉｃ ＲＮＡ－ｐｒｏｔｅｉｎ ｉｎｔｅｒａｃ

ｔｉｏｎ，”Ｎｕｃｌ．Ａｃｉｄｓ　Ｒｅｓ．，４０．９（２０１２）：ｅ６４－ｅ６４を参照のこと）。

【０１４８】

誘導性ＣＲＩＳＰＲエフェクター及び誘導性ＣＲＩＳＰＲシステムの様々な実施形態が、例えば、米国特許第８８７１４４５号明細書、米国特許出願公開第２０１６０２０８２４３号明細書、及び国際公開第２０１６２０５７６４号パンフレット（これらの各々は、本明細書において全体として参照により援用される）に記載されている。

【０１４９】

機能性突然変異

本明細書に記載されるとおりのＣＲＩＳＰＲエフェクターに様々な突然変異又は修飾を導入して特異性及び／又はロバスト性を改善することができる。一部の実施形態において、プロトスペーサー隣接モチーフ（ＰＡＭ）を認識するアミノ酸残基が同定される。本明細書に記載されるＣＲＩＳＰＲエフェクターは、例えば、ＰＡＭを認識するアミノ酸残基を他のアミノ酸残基と置換することによって異なるＰＡＭを認識するように更に修飾することができる。一部の実施形態において、ＣＲＩＳＰＲエフェクターは、例えば、５’－ＮＴＴＮ－３’、５’－ＮＴＴＲ－３’、５’－ＲＴＴＲ－３’、５’－ＴＮＮＴ－３’、５’－ＴＮＲＴ－３’、５’－ＴＳＲＴ－３’、５’－ＴＧＲＴ－３’、５’－ＴＮＲＹ－３’、５’－ＴＴＮＲ－３’、５’－ＴＴＹＲ－３’、５’－ＴＴＴＲ－３’、５’－ＴＴＣＶ－３’、５’－ＤＴＹＲ－３’、５’－ＷＴＴＲ－３’、５’－ＮＮＲ－３’、５’－ＮＹＲ－３’、５’－ＹＹＲ－３’、５’－ＴＹＲ－３’、５’－ＴＴＮ－３’、５’－ＴＴＲ－３’、５’－ＣＮＴ－３’、５’－ＮＧＧ－３’、５’－ＢＧＧ－３’、又は５’－Ｒ－３’を認識し得、ここで、「Ｎ」は任意のヌクレオチドであり、「Ｂ」はＣ又はＧ又はＴであり、「Ｄ」はＡ又はＧ又はＴであり、「Ｒ」はＡ又はＧであり、「Ｓ」はＧ又はＣであり、「Ｖ」はＡ又はＣ又はＧであり、「Ｗ」はＡ又はＴであり、「Ｙ」はＣ又はＴである。

【０１５０】

一部の実施形態において、本明細書に記載されるＣＲＩＳＰＲエフェクターは、１つ以上のアミノ酸残基を突然変異させることにより、１つ以上の機能活性が改変され得る。例えば、一部の実施形態において、ＣＲＩＳＰＲエフェクターは、１つ以上のアミノ酸残基を突然変異させることにより、そのヘリカーゼ活性が改変される。一部の実施形態において、ＣＲＩＳＰＲエフェクターは、１つ以上のアミノ酸残基を突然変異させることにより、そのヌクレアーゼ活性（例えば、エンドヌクレアーゼ活性又はエキソヌクレアーゼ活性）が改変される。一部の実施形態において、ＣＲＩＳＰＲエフェクターは、１つ以上のアミノ酸残基を突然変異させることにより、ＲＮＡガイドと機能的に関連するその能力が改変される。一部の実施形態において、ＣＲＩＳＰＲエフェクターは、１つ以上のアミノ酸残基を突然変異させることにより、標的核酸と機能的に関連するその能力が改変される。

【０１５１】

一部の実施形態において、本明細書に記載されるＣＲＩＳＰＲエフェクターは標的核酸分子の切断能を有する。一部の実施形態において、ＣＲＩＳＰＲエフェクターは標的核酸分子の両方の鎖を切断する。しかしながら、一部の実施形態において、ＣＲＩＳＰＲエフェクターは、１つ以上のアミノ酸残基を突然変異させることにより、その切断活性が改変される。例えば、一部の実施形態において、ＣＲＩＳＰＲエフェクターは、ＣＲＩＳＰＲエフェクターが標的核酸を切断する能力を増加させる１つ以上の突然変異を含んでもよい。別の例において、一部の実施形態において、ＣＲＩＳＰＲエフェクターは、酵素が標的核酸の切断能を有しないものとなる１つ以上の突然変異を含んでもよい。他の実施形態において、ＣＲＩＳＰＲエフェクターは、この酵素が標的核酸の鎖の切断能（即ち、ニッカーゼ活性）を有するものとなる１つ以上の突然変異を含んでもよい。一部の実施形態において、ＣＲＩＳＰＲエフェクターは、ＲＮＡガイドがハイブリダイズする鎖に相補的な標的核酸の鎖を切断する能力を有する。一部の実施形態において、ＣＲＩＳＰＲエフェクターは、ＲＮＡガイドがハイブリダイズする標的核酸の鎖を切断する能力を有する。

【０１５２】

一部の実施形態において、本明細書に開示されるＣＲＩＳＰＲエフェクターの１つ以上の残基は、アルギニン部分に変異している。一部の実施形態において、本明細書に開示されるＣＲＩＳＰＲエフェクターの１つ以上の残基は、グリシン部分に変異している。一部の実施形態において、本明細書に開示されるＣＲＩＳＰＲエフェクターの１つ以上の残基は、本明細書に開示されるＣＲＩＳＰＲエフェクターの系統的アラインメントのコンセンサス残基に基づいて変異する。

【０１５３】

一部の実施形態において、本明細書に記載されるＣＲＩＳＰＲエフェクターは、１つ以上の所望の機能活性（例えば、ヌクレアーゼ活性及び機能的にＲＮＡガイドと相互作用する能力）を保持しつつ酵素のサイズを縮小させるため、１つ以上のアミノ酸残基に欠失を含むようにエンジニアリングされてもよい。このトランケート型ＣＲＩＳＰＲエフェクターは有利には、負荷に制限のある送達システムとの組み合わせで用いられてもよい。

【０１５４】

一態様において、本開示は、図２に示されるドメイン構成を維持しつつ、本明細書に記載される核酸配列（ｎｕｃｌｅｉｃ　ｓｅｑｕｅｎｃｅｓ）と少なくとも１０％、１５％、２０％、２５％、３０％、３５％、４０％、４５％、５０％、５５％、６０％、６５％、７０％、７５％、８０％、８５％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、又は９９％同一の核酸配列を提供する。別の態様において、本開示はまた、図２に示されるドメイン構成を維持しつつ、本明細書に記載されるアミノ酸配列と少なくとも１０％、１５％、２０％、２５％、３０％、３５％、４０％、４５％、５０％、５５％、６０％、６５％、７０％、７５％、８０％、８５％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、又は９９％同一のアミノ酸配列も提供する。

【０１５５】

一部の実施形態において、核酸配列は、本明細書に記載される配列と同じである一部分（例えば、少なくとも１、２、３、４、５、６、７、８、９、１０、１１、１２、１３、１４、１５、２０、３０、４０、５０、６０、７０、８０、９０、又は１００ヌクレオチド、例えば、連続又は非連続ヌクレオチド）を少なくとも有する。一部の実施形態において、核酸配列は、本明細書に記載される配列と異なる一部分（例えば、少なくとも１、２、３、４、５、６、７、８、９、１０、１１、１２、１３、１４、１５、２０、３０、４０、５０、６０、７０、８０、９０、又は１００ヌクレオチド、例えば、連続又は非連続ヌクレオチド）を少なくとも有する。

【０１５６】

一部の実施形態において、アミノ酸配列は、本明細書に記載される配列と同じである一部分（例えば、少なくとも１、２、３、４、５、６、７、８、９、１０、１１、１２、１３、１４、１５、２０、３０、４０、５０、６０、７０、８０、９０、又は１００アミノ酸残基、例えば、連続又は非連続アミノ酸残基）を少なくとも有する。一部の実施形態において、アミノ酸配列は、本明細書に記載される配列と異なる一部分（例えば、少なくとも１、２、３、４、５、６、７、８、９、１０、１１、１２、１３、１４、１５、２０、３０、４０、５０、６０、７０、８０、９０、又は１００アミノ酸残基、例えば、連続又は非連続アミノ酸残基）を少なくとも有する。

【０１５７】

２つのアミノ酸配列、又は２つの核酸配列のパーセント同一性を決定するには、それらの配列が最適な比較を目的としてアラインメントされる（例えば、最適なアラインメントとなるように第１及び第２のアミノ酸又は核酸配列の一方又は両方にギャップが導入されてもよく、及び比較を目的として非相同配列が無視されてもよい）。一般に、比較を目的としてアラインメントされる参照配列の長さは、参照配列の長さの少なくとも８０％でなければならず、及び一部の実施形態では、参照配列の長さの少なくとも９０％、９５％、又は１００％である。次に、対応するアミノ酸位置又はヌクレオチド位置にあるアミノ酸残基又はヌクレオチドが比較される。第１の配列におけるある位置が第２の配列における

対応する位置と同じアミノ酸残基又はヌクレオチドによって占有されているとき、次にはそれらの分子は当該の位置において同一である。２つの配列間のパーセント同一性は、２つの配列を最適にアラインメントするために導入する必要があるギャップの数、及び各ギャップの長さを考慮に入れた、それらの配列によって共有される同一の位置の数の関数である。本開示の目的上、配列の比較及び２つの配列間におけるパーセント同一性の決定は、ギャップペナルティーを１２、ギャップ伸長ペナルティーを４、及びフレームシフトギャップペナルティーを５としたＢｌｏｓｓｕｍ ６２スコアリング行列を用いて達成することができる。

【０１５８】

一部の実施形態において、ヌクレアーゼは$PX_1X_2X_3X_4F$（配列番号２１６）として記載される配列を含み、ここで、X_1はＬ又はＭ又はＩ又はＣ又はＦであり、X_2はＹ又はＷ又はＦであり、X_3はＫ又はＴ又はＣ又はＲ又はＷ又はＹ又はＨ又はＶであり、X_4はＩ又はＬ又はＭである。一部の実施形態において、配列番号２１６に記載の配列はＮ末端配列である。一部の実施形態において、ヌクレアーゼは$RX_1X_2X_3L$（配列番号２１７）として記載される配列を含み、ここで、X_1はＩ又はＬ又はＭ又はＹ又はＴ又はＦであり、X_2はＲ又はＱ又はＫ又はＥ又はＳ又はＴであり、X_3はＬ又はＩ又はＴ又はＣ又はＭ又はＫである。一部の実施形態において、ヌクレアーゼはNX_1YX_2（配列番号２１８）として記載される配列を含み、ここで、X_1はＩ又はＬ又はＦであり、X_2はＫ又はＲ又はＶ又はＥである。一部の実施形態において、ヌクレアーゼは$KX_1X_2X_3FAX_4X_5KD$（配列番号２１９）として記載される配列を含み、ここで、X_1はＴ又はＩ又はＮ又はＡ又はＳ又はＦ又はＶであり、X_2はＩ又はＶ又はＬ又はＳであり、X_3はＨ又はＳ又はＧ又はＲであり、X_4はＤ又はＳ又はＥであり、X_5はＩ又はＶ又はＭ又はＴ又はＮである。本明細書に記載されるシステムのいずれかの一部の実施形態において、配列番号２１９の配列はＣ末端配列である。一部の実施形態において、ヌクレアーゼはLX_1NX_2（配列番号２２０）として記載される配列を含み、ここで、X_1はＧ又はＳ又はＣ又はＴであり、X_2はＮ又はＹ又はＫ又はＳである。本明細書に記載されるシステムのいずれかの一部の実施形態において、配列番号２２０の配列はＣ末端配列である。一部の実施形態において、ヌクレアーゼは$PX_1X_2X_3X_4SQX_5DS$（配列番号２２１）として記載される配列を含み、ここで、X_1はＳ又はＰ又はＡであり、X_2はＹ又はＳ又はＡ又はＰ又はＥ又はＹ又はＱ又はＮであり、X_3はＦ又はＹ又はＨであり、X_4はＴ又はＳであり、X_5はＭ又はＴ又はＩである。本明細書に記載されるシステムのいずれかの一部の実施形態において、配列番号２２１の配列はＣ末端配列である。一部の実施形態において、ヌクレアーゼは$KX_1X_2VRX_3X_4QEX_5H$（配列番号２２２）として記載される配列を含み、ここで、X_1はＮ又はＫ又はＷ又はＲ又はＥ又はＴ又はＹであり、X_2はＭ又はＲ又はＬ又はＳ又はＫ又はＶ又はＥ又はＴ又はＩ又はＤであり、X_3はＬ又はＲ又はＨ又はＰ又はＴ又はＫ又はＰのＱ又はＳ又はＡであり、X_4はＧ又はＱ又はＮ又はＲ又はＫ又はＥ又はＩ又はＴ又はＳ又はＣであり、X_5はＲ又はＷ又はＹ又はＫ又はＴ又はＦ又はＳ又はＱである。本明細書に記載されるシステムのいずれかの一部の実施形態において、配列番号２２２の配列はＣ末端配列である。一部の実施形態において、ヌクレアーゼは$X_1NGX_2X_3X_4DX_5NX_6X_7X_8N$（配列番号２２３）として記載される配列を含み、ここで、X_1はＩ又はＫ又はＶ又はＬであり、X_2はＬ又はＭであり、X_3はＮ又はＨ又はＰであり、X_4はＡ又はＳ又はＣであり、X_5はＶ又はＹ又はＩ又はＦ又はＴ又はＮであり、X_6はＡ又はＳであり、X_7はＳ又はＡ又はＰであり、X_8はＭ又はＣ又はＬ又はＲ又はＮ又はＳ又はＫ又はＬである。本明細書に記載されるシステムのいずれかの一部の実施形態において、配列番号２２３の配列はＣ末端配列である。

【０１５９】

ＲＮＡガイド及びＲＮＡガイド修飾

一部の実施形態において、本明細書に記載されるＲＮＡガイドは、ウラシル（Ｕ）を含む。一部の実施形態において、本明細書に記載されるＲＮＡガイドは、チミン（Ｔ）を含む。一部の実施形態において、本明細書に記載されるＲＮＡガイドのダイレクトリピート

配列は、ウラシル（Ｕ）を含む。一部の実施形態において、本明細書に記載されるＲＮＡガイドのダイレクトリピート配列は、チミン（Ｔ）を含む。一部の実施形態において、表２又は表８によるダイレクトリピート配列は、表２又は表８の対応する配列においてチミンとして示される１つ以上の場所に、ウラシルを含む配列を含む。

【０１６０】

一部の実施形態において、ダイレクトリピートは、内因性ＣＲＩＳＰＲアレイにおいて繰り返される配列の１つのコピーのみを含む。一部の実施形態において、ダイレクトリピートは、内因性ＣＲＩＳＰＲアレイに見られる１つ以上のスペーサー配列に隣接する（例えば、フランキング）完全長配列である。一部の実施形態において、ダイレクトリピートは、内因性ＣＲＩＳＰＲアレイに見られる１つ以上のスペーサー配列に隣接する（例えば、フランキング）完全長配列の一部（例えば、プロセシングされた部分）である。

【０１６１】

スペーサー及びダイレクトリピート

ＲＮＡガイドのスペーサー長さは約１５～５５ヌクレオチドの範囲であってもよい。ＲＮＡガイドのスペーサー長さは約２０～４５ヌクレオチドの範囲であってもよい。一部の実施形態において、ＲＮＡガイドのスペーサー長さは、少なくとも１５ヌクレオチド、少なくとも１６ヌクレオチド、少なくとも１７ヌクレオチド、少なくとも１８ヌクレオチド、少なくとも１９ヌクレオチド、少なくとも２０ヌクレオチド、少なくとも２１ヌクレオチド、又は少なくとも２２ヌクレオチドである。一部の実施形態において、スペーサー長さは、１５～１７ヌクレオチド、１５～２３ヌクレオチド、１６～２２ヌクレオチド、１７～２０ヌクレオチド、２０～２４ヌクレオチド（例えば、２０、２１、２２、２３、又は２４ヌクレオチド）、２３～２５ヌクレオチド（例えば、２３、２４、又は２５ヌクレオチド）、２４～２７ヌクレオチド、２７～３０ヌクレオチド、３０～４５ヌクレオチド（例えば、３０、３１、３２、３３、３４、３５、４０、又は４５ヌクレオチド）、３０又は３５～４０ヌクレオチド、４１～４５ヌクレオチド、４５～５０ヌクレオチド、又はそれ以上である。

【０１６２】

一部の実施形態において、ＲＮＡガイドのダイレクトリピート長さは少なくとも１６ヌクレオチドであり、又は１６～２０ヌクレオチド（例えば、１６、１７、１８、１９、又は２０ヌクレオチド）である。一部の実施形態において、ＲＮＡガイドのダイレクトリピート長さは約１９～約４０ヌクレオチドである。

【０１６３】

例示的なダイレクトリピート配列（例えば、プレｃｒＲＮＡ（例えば、プロセシングされていないｃｒＲＮＡ）のダイレクトリピート配列又は成熟ｃｒＲＮＡ（例えば、プロセシングされたｃｒＲＮＡのダイレクトリピート配列））を表２に示す。表８も参照されたい。

【０１６４】

【表２】

表2. 例示的なcrRNA配列のダイレクトリピート配列。

エフェクター	ダイレクトリピート配列
配列番号1	ACTATGTTGGAATACATTTTTATAGGTATTTACAACT（配列番号57）
配列番号2	ATTGTTGGAATATCACTTTTGTAGGGTATTCACAAC（配列番号58）
配列番号3	AATGTTGTTCACCCTTTTT（配列番号59）
配列番号4	CCTGTTGTGAATACTCTTTTATAGGTATCAAACAAC（配列番号60）
配列番号10	ATTGTTGTAGACACCTTTTTATAAGGATTGAACAAC（配列番号62） CTTGTTGTATATGTCCTTTTATAGGTATTAAACAAC（配列番号213）
配列番号14	GTTGTTTAATACCTATAAAAGAATATATACAACAAG （配列番号128）
配列番号15	CTTGTTGTATATACTCTTTTATAGGTATTAAACAAC（配列番号63）
配列番号17	GTTGTATCCACCGTATAAAACATAGTGTCCAACATC （配列番号130）
配列番号18	GATGTTGTTATGCTGTTTTTGTAAGTAATAAACAAC（配列番号70）
配列番号21	ATTGTTGTACGAACCATTTTATATGGTAATAACAAC（配列番号72）
配列番号22	ACTGTAAAACCCCTGCAGATGAAAGGAAAGTACAACAGT （配列番号73）
配列番号23	ATCATGTTGTACATACTATTTTTTAAGTATTAAACAACTA （配列番号74）
配列番号24	CTTGTTGTATATACTCTTTTATAGGTATTAAACAAC（配列番号63）
配列番号27	ATTGTTGGGGTACTTCTTTTATAGGGTACTCACAAC（配列番号76）
配列番号28	ATTGTTGTAGACCTTGTGTTTTAGGGGTCTAACAACG （配列番号77）
配列番号29	GTTGTAAATACATCTCATATTGTATTCCAACACAGT（配列番号139）
配列番号31	ATTGTTGGAATATCACTTTTGTAGGGTATTCACAAC（配列番号58）
配列番号32	AATTGTTGAGATACCGTTTTTTATGGTATTGGCAAC（配列番号80）
配列番号35	ATTGTTGTAGACCTTGTGTTTTAGGGGTCTAACAACG （配列番号77）
配列番号36	GTTGTAAATACATCTCATATTGTATTCCAACACAGT（配列番号139）
配列番号38	AATTGTTGAGATACCGTTTTTTATGGTATTGGCAAC（配列番号80）
配列番号39	ATTGTTGGAATATCACTTTTGTAGGGTATTCACAAC（配列番号58）
配列番号41	ATTGTGTTGGGATACACTTTTATAGGTATTTACAAC（配列番号83）
配列番号42	TATTGTTGAATACCTTTCTTATAAAGGTAATTACAAC（配列番号84）
配列番号44	ATTGTTGAATGTATTCTTTTTTAGGACAGATACAAC（配列番号86）
配列番号45	GTTGTATCCACCGTATAAAACATAGTGTCCAACATC （配列番号130）
配列番号46	TATTGTTGAATACCTTTCTTATAAAGGTAATTACAAC（配列番号84）
配列番号47	ATTGTTGAATGGTATCTTTTATAGACTGATTACAACT（配列番号87）
配列番号48	ATTGTTGGATAATAGGTTTTTTATCTTAATTACAAC（配列番号88）
配列番号51	TATTGTTGAATACCTTTCTTATAAAGGTAATTACAAC（配列番号84）
配列番号53	TATTGTTGAATACCTTTCTTATAAAGGTAATTACAAC（配列番号84）
配列番号55	ATTGTTGGATAATAGGTTTTTTATCTTAATTACAAC（配列番号88）
配列番号56	ATTGTTGTAGATACCTTTTTGTAAGGATTGAACAAC（配列番号90）

【０１６５】

一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号１のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ダイレクトリピート配列は、配列番号５７のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号２

のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ダイレクトリピート配列は、配列番号５８のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号３のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ダイレクトリピート配列は、配列番号５９のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号４のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ダイレクトリピート配列は、配列番号６０のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号１０のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ダイレクトリピート配列は、配列番号６２又は配列番号２１３のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号１４のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ダイレクトリピート配列は、配列番号１２８のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号１５のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ダイレクトリピート配列は、配列番号６３のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号１７のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ダイレクトリピート配列は、配列番号１３０のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態に

おいて、ＣＲＩＳＰＲ関連タンパク質は、配列番号１８のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ダイレクトリピート配列は、配列番号７０のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号２１のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ダイレクトリピート配列は、配列番号７２のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号２２のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ダイレクトリピート配列は、配列番号７３のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号２３のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ダイレクトリピート配列は、配列番号７４のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号２４のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ダイレクトリピート配列は、配列番号６３のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号２７のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ダイレクトリピート配列は、配列番号７６のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号２８のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ダイレクトリピート配列は、配列番号７７のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオ

チド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号２９のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ダイレクトリピート配列は、配列番号１３９のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号３１のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ダイレクトリピート配列は、配列番号５８のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号３２のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ダイレクトリピート配列は、配列番号８０のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号３５のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ダイレクトリピート配列は、配列番号７７のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号３６のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ダイレクトリピート配列は、配列番号１３９のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号３８のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ダイレクトリピート配列は、配列番号８０のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号３９のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ダイレクトリピート配列は、配列番号５８のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、

９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号４１のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ダイレクトリピート配列は、配列番号８３のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号４２のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ダイレクトリピート配列は、配列番号８４のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号４４のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ダイレクトリピート配列は、配列番号８６のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号４５のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ダイレクトリピート配列は、配列番号１３０のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号４６のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ダイレクトリピート配列は、配列番号８４のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号４７のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ダイレクトリピート配列は、配列番号８７のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号４８のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ダイレクトリピート配列は、配列番号８８のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％

、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号５１のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ダイレクトリピート配列は、配列番号８４のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号５３のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ダイレクトリピート配列は、配列番号８４のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号５５のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ダイレクトリピート配列は、配列番号８８のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号５６のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ダイレクトリピート配列は、配列番号９０のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。

【０１６６】

一部の実施形態において、ＲＮＡガイドは図３に記載のダイレクトリピート配列を含む。例えば、一部の実施形態において、ＲＮＡガイドは、図３に示されるコンセンサス配列のダイレクトリピート又は図３に示されるコンセンサス配列の一部を含む。一部の実施形態において、ＲＮＡガイドは、$X_1X_2TX_3X_4X_5X_6X_7X_8$（配列番号２２４）として記載される配列を有するダイレクトリピートを含み、ここで、X_1はＡ又はＣ又はＧであり、X_2はＴ又はＣ又はＡであり、X_3はＴ又はＧ又はＡであり、X_4はＴ又はＧであり、X_5はＴ又はＧ又はＡであり、X_6はＧ又はＴ又はＡであり、X_7はＴ又はＧ又はＡであり、X_8はＡ又はＧ又はＴである。例えば、一部の実施形態において、ＲＮＡガイドは、ＡＴＴＧＴＴＧＤＡ（配列番号２２５）として記載される配列を有するダイレクトリピートを含む。一部の実施形態において、配列番号２２４はダイレクトリピートの５’末端に近接している。一部の実施形態において、配列番号２２５はダイレクトリピートの５’末端に近接している。一部の実施形態において、ＲＮＡガイドは、$X_1X_2X_3X_4X_5X_6X_7X_8X_9$（配列番号２２６）として記載される配列を有するダイレクトリピートを含み、ここで、X_1はＴ又はＣ又はＡであり、X_2はＴ又はＡ又はＧであり、X_3はＴ又はＣ又はＡであり、X_4はＴ又はＡであり、X_5はＴ又はＡ又はＧであり、X_6はＴ又はＡであり、X_7はＡ又はＴであり、X_8はＡ又はＧ又はＣ又はＴであり、X_9はＧ又はＡ又はＣである。例えば、一部の実施形態において、ＲＮＡガイドは、ＴＴＴＴＷＴＡＲＧ（配列番号２２７）として記載される配列を有するダイレクトリピートを含む。一部の実施形態において、ＲＮＡガイドは、$X_1X_2X_3AC$（配列番号２２８）として記載される配列を有する

ダイレクトリピートを含み、ここで、X_1はＡ又はＣ又はＧであり、X_2はＣ又はＡであり、X_3はＡ又はＣである。例えば、一部の実施形態において、ＲＮＡガイドは、ＡＣＡＡＣ（配列番号２２９）として記載される配列を有するダイレクトリピートを含む。一部の実施形態において、配列番号２２８はダイレクトリピートの３’末端に近接している。一部の実施形態において、配列番号２２９はダイレクトリピートの３’末端に近接している。

【０１６７】

一部の実施形態において、ＲＮＡガイドのスペーサーは、表３のＰＡＭ配列に隣接する標的核酸に結合する。例えば、一部の実施形態において、本明細書に開示されるエフェクターとＲＮＡガイドとの複合体は、表３に示されるＰＡＭ配列に隣接する標的核酸に結合する。

【０１６８】

【表３】

表3.CLUST.091979 エフェクターに対応するPAM 配列。

エフェクター	PAM 配列
配列番号 1	5’-TTNT-3’ 5’-TNRT-3’
配列番号 2	5’-TTR-3’ 5’-WTTR-3’
配列番号 4	5’-NNR-3’ 5’-NTTN-3’ 5’-NTTR-3’ 5’-TTTN-3’ 5’-TTTG-3’
配列番号 10	5’-NTTN-3’ 5’-RTTR-3’ 5’-ATTR-3’ 5’-RTTG-3’ 5’-ATTG-3’ 5’-GTTA-3’
配列番号 14	5’-TTN-3’ 5’-TTY-3’ 5’-YYR-3’
配列番号 15	5’-CNT-3’
配列番号 21	5’-TTCV-3’ 5’-TTYR-3’
配列番号 23	5’-GTA-3’
配列番号 24	5’-CNT-3’
配列番号 27	5’-TTR-3’ 5’-YYR-3’ 5’-TYR-3’
配列番号 28	5’-NGG-3’ 5’-BGG-3’ 5’-CGG-3’ 5’-GG-3’
配列番号 31	5’-TTR-3’
配列番号 32	5’-TYR-3’
配列番号 35	5’-NGG-3’

【０１６９】

【表４】

	5’-BGG-3’ 5’-CGG-3’ 5’-GG-3’
配列番号 38	5’-TYR-3’
配列番号 39	5’-TTR-3’
配列番号 41	5’-TYR-3’
配列番号 42	5’-TYR-3’ 5’-TTYR-3’ 5’-DTYR-3’
配列番号 44	5’-TTNR-3’ 5’-TTTR-3’
配列番号 46	5’-TYR-3’ 5’-TTYR-3’ 5’-DTYR-3’
配列番号 48	5’-YYR-3’ 5’-TTR-3’ 5’-TTG-3’
配列番号 51	5’-TYR-3’ 5’-TTYR-3’ 5’-DTYR-3’
配列番号 53	5’-TYR-3’ 5’-TTYR-3’ 5’-DTYR-3’
配列番号 55	5’-YYR-3’ 5’-TTR-3’ 5’-TTG-3’
配列番号 56	5’-TTG-3’ 5’-NYR-3’ 5’-TYR-3’

【０１７０】
一部の実施形態において、ＲＮＡガイドはｔｒａｃｒＲＮＡを更に含む。一部の実施形態において、ｔｒａｃｒＲＮＡは必要でない（例えば、ｔｒａｃｒＲＮＡは任意選択である）。一部の実施形態において、ｔｒａｃｒＲＮＡは、表９に示される非コード配列の一部である。例えば、一部の実施形態において、ｔｒａｃｒＲＮＡは表４の配列である。
【０１７１】

【表５】

表4.例示的なtracrRNA配列。

エフェクター	tracrRNA配列
配列番号1	ATTGGGACTTCCGGAAGTAAAATATCCACCTGAGGATTTTAGGACAT ATAATTTCTAATAAAAATGAACGGAAAAATTTCCGTTCATTTTTTTTT TGTTTATT (配列番号152) TATTGGGACTTCCGGAAGTAAAATATCCACCTGAGGATTTTAGGACA TATAATTTCTAATAAAAATGAACGGAAAAATTTCCGTTCATTTTTTTT TTGTTTATTG (配列番号153) GACGAGAACGGAGTGTGGCTCCTGAGGAAAAACGACAAACATCCAA CATATTTTATCTACCAGAACGGAACACTCTATCAATATGAGGAAGAT TGATTAGTTGATGTTTTCATAATAATTTTATCTGGAATTTGAAAAGAT TCCAGATTTTTTTTTTATTTCG (配列番号154)
配列番号2	GCAATCAACAAGACTTTCATTTTCAAGGCAAAATGCGATAAGAACG ATGTCATATCGTTATGGGAA (配列番号155) GATGCTCCGAAAACGTGGTTGTTCGGACAACAAAAAAATGAATGTT TCTAATGTATTAA (配列番号156) GACGGAAAAATAAATGAGGATGGTATGTTTGTTGAAAACTTGGAAT AATTCTGTATATACCAATTAGAAT (配列番号157) TGTTGATTGCTGATTCTTCGTTGTTTGATTTGTGTTGTGCCATAATCT TAAAATT (配列番号158)

【０１７２】

【表６】

配列番号3	CGCAAGATATAAGGCAATCGGAAACGGATGGACAGTTGATGTAATT TCACATATTTTTAAGAATTTGAAAAATTAATTTGGTA (配列番号159) GGACATTTCGTAAATCATATGGAGATACGGAGTTCAAGTCAATTGA AGAGCTTCCTGAATTTAGAGATAACATACTTATACAACTAGATTGAT TG (配列番号160) ATCAATACATAGATGATGAGAAATGGAGAAAAAAATTTGTTCGCCC AACAAACACTAAT (配列番号161)
配列番号14	CTGGTAATACTGTAAAATCTCCGTGTATAGGGCAAGTAATTGTAACT GGGGTAATTCTATCTACTATTATAGTTTTAGAA (配列番号162)
配列番号17	CAGAAGTCGTTCAAGTTCAAGGTCAAAACGGACAAGGAGACGGTCG AATTATTCAG (配列番号163) GGGAGGGTGACATTCAGAAGTCGTTCAAGTTCAAGGTCAAAACGGA CAAGGAGACGGTCGAATTAT (配列番号164) AAGTGTCTTCAACACATTGAAGAAAACTCTCGGTGCAATATATGGA AAGCTCGATGAAAACGGAAATTTTATTGAGAATGAATGTAATAAGT AACTGGAATA (配列番号165) CCGTGGGAGGATTTGGATTTGGTTGAAGACATCAGAAAAATTTTCGA AATGGAATAGAGGGAACCGGAATTTTTTCCGGTTTTTCTTTGTCCTTT CGA (配列番号166)
配列番号18	CAGAGTAACCTTTCCTGATATGTTGTTACACATTTTTGTAAGTGTTAA ACAACTGACGCATTGATATTGCCTTGTCTATTAA (配列番号167) CAATCGCGAGTTTATACTGAAATGTTGTTACACTGTTTTTGTAAGTGT TAAACAACCTTGCACAAATGTCATCTACCAGTAC (配列番号168)
配列番号21	CCGAGCGACCCACAAACCTATTGTCGTACGCATCATTTCACATGATA ATAACAACGAATATTCCTGCAAGCATGATTT (配列番号169) TATGACATTATGATATTGTTGTATGCATCATTTCACATGGTAATAAC AACGAAGAGAAACACCGAGCGACCCACAAA (配列番号170) ACATCTTTTATGACATTATGATATTGTTGTATGCATCATTTCACATGG TAATAACAACGAAGAGAAACACCGAGCGACCCACAAA (配列番号171)
配列番号22	GCTAAAATATAGTCCTGTGGATGTTGAATACATTTCTTTTAAGTGTA CTTACAACCAACGCTGTACACATTGCTAATGGATG (配列番号172) TGCTAAAATATAGTCCTGTGGATGTTGAATACATTTCTTTTAAGTGT ACTTACAACCAACGCTGTACACATTGCTAATGGATG (配列番号173)

【０１７３】

【表７】

	CAACACCAAGGCTGAGGCAAAGAAGAGGGCTGATGATATGAACAA ACAGAATAGGGTCATACACCAGCTGTCTGTTTATTTGTGTCC (配列番号174) AATTAGACTGATAAACAAAGAATAATGAGAACTATAATAGGGAGGT GTACCCCCGAATTTAAGCCAGTGGAGAACCATACAAACCTATCATAT AG (配列番号175)
配列番号23	TGGGTATGCGTTGTTTAATACTTAAAAAAATGTATGTACAACATGTC TGTGGAAAGTCTTTCTATTGTATAT (配列番号176) CGTTGTTTAATACTTAAAAAAATGTATGTACAACATGTCTGTGGAAA GTCTTTCTATTGTATATAGGA (配列番号177) TGGGTATGCGTTGTTTAATACTTAAAAAAATGTATGTACAACATGTC TGTGGAAAGTCTTTCTATTGTATATAGGAATTTTATATAATTATTTAA TTATCAATGAATTATATTAGTAT (配列番号178) GGTGGGTATGCGTTGTTTAATACTTAAAAAAATGTATGTACAACATG TCTGTGGAAAG (配列番号179)
配列番号27	AATGAACGAGATTGTTGGGATATACCTTTTATAGGATTTTCACAACA TCTGAGTTGTTTGATGTTAAAAACTT (配列番号180) GATAAAAATGAACGAGATTGTTGGGATATACCTTTTATAGGATTTTC ACAACATCTGAGTTGTTTGATGTTAAAAACTTT (配列番号181)
配列番号29	GCTAATATAAAGATTGTACTGTGTTGAGATACACTTTTAGAGGTATT TACAACAAAATGCGTGATATGGAAATGA (配列番号182) ATACCAACATAAATACAGGTCTTGCTGTTTCTGGTCGGTCGTAAACA CCTCTAAAAGGATTGTTTCGACATAGGTTACTGACGCTTCAAG (配列番号183) AATGAAGAAATAACTGTGTTGAGATACACTTTTAGAGGTATTTACAA CACCATATAAACCTGACCATCTCCT (配列番号184)
配列番号31	AGGAAGATGTCAGACGTTTTTATTGTTGGAATACTCGTTTTTTACGG TATTTACAACTGCCCCGTAGCGGAATCAAAATACCAC (配列番号185) ATGTCAGACGTTTTTATTGTTGGAATACTCGTTTTTTACGGTATTTAC AACTGCCCCGTAGCGGAATCAAAATACC (配列番号186) AAATAACAAAAATTCTGGACGGGAAAGGAAGATGTCAGACGTTTTT ATTGTTGGAATACTCGTTTTTTACGGTATTTACAACTGCCCCGTAGCG GAATC (配列番号187)

【０１７４】

【表８】

	ATAACAAAAATTCTGGACGGGAAAGGAAGATGTCAGACGTTTTTAT TGTTGGAATACTCGTTTTTTACGGTATTTACAACTGCCCCGTAGCGG AAT (配列番号188)
配列番号32	TATTGCAACTATTACAACAAACTTAGCGAATGGATTGGCAAAGATAT GTATAACACGCCG (配列番号189) ATTGCAACTATTACAACAAACTTAGCGAATGGATTGGCAAAGATAT GTATAACACGCCG (配列番号190)
配列番号36	GCTAATATAAAGATTGTACTGTGTTGAGATACACTTTTAGAGGTATT TACAACAAAATGCGTGATATGGAAATGA (配列番号182) ATACCAACATAAATACAGGTCTTGCTGTTTCTGGTCGGTCGTAAACA CCTCTAAAAGGATTGTTTCGACATAGGTTACTGACGCTTCAAG (配列番号183) AATGAAGAAATAACTGTGTTGAGATACACTTTTAGAGGTATTTACAA CACCATATAAACCTGACCATCTCCT (配列番号184)
配列番号38	TATTGCAACTATTACAACAAACTTAGCGAATGGATTGGCAAAGATAT GTATAACACGCCG (配列番号189) ATTGCAACTATTACAACAAACTTAGCGAATGGATTGGCAAAGATAT GTATAACACGCCG (配列番号190)
配列番号39	AGGAAGATGTCAGACGTTTTTATTGTTGGAATACTCGTTTTTTACGG TATTTACAACTGCCCCGTAGCGGAATCAAAATACCAC (配列番号185) ATGTCAGACGTTTTTATTGTTGGAATACTCGTTTTTTACGGTATTTAC AACTGCCCCGTAGCGGAATCAAAATACC (配列番号186) AAATAACAAAAATTCTGGACGGGAAAGGAAGATGTCAGACGTTTTT ATTGTTGGAATACTCGTTTTTTACGGTATTTACAACTGCCCCGTAGCG GAATC (配列番号187) ATAACAAAAATTCTGGACGGGAAAGGAAGATGTCAGACGTTTTTAT TGTTGGAATACTCGTTTTTTACGGTATTTACAACTGCCCCGTAGCGG AAT (配列番号188)
配列番号41	GTATGATGACAGAAGAAACACGGAAGACAATAGAGAGCGTCATAGT GGTTCTCGGCATAGCAATCATGCTG (配列番号191) ATGATGACAGAAGAAACACGGAAGACAATAGAGAGCGTCATAGTG GTTCTCGGCATAGCAATCATGCTGGCAGCCGCCGTCCGAATAATGAC GCAGAACAAAGCAATTGTGAAATATG (配列番号192) AGAAGGTACTGCCGCCTTATGACCGACGAGAACGGAGTGTGGCTCC TGAGGAAAAAC (配列番号193)

【０１７５】

【表９】

	GACGAGAACGGAGTGTGGCTCCTGAGGAAAAACGACAAACATCCAACATATTTTATCTACCAGAACGGAACACTCTATCAATATGAGGAAGATTGATTAGTTGATGTTTTCATAATAATTTTATCTGGAATTTGAAAAGATTCCAGATTTTTTTTTTATTTCG (配列番号194)
配列番号43	TCGTTGAATACGATATCGCCGAAACAATTGATTGGAGAAGTACGCTTTGTTTCAAGACATGGAATACGTATGGTTCTCCTCAATGGGACTCGAAGATCAAGAA (配列番号197) ATCGTTGAATACGATATCGCCGAAACAATTGATTGGAGAAGTACGCTTTGTTTCAAGACATGGAATACGTATGGTTCTCCTCAATGGGACTCGAAGATCAAGAACCAG (配列番号198) GAGCTTTTCTGGCAATGTAGACATTAAAGCTGGTATCGTTGAATACGATATCGCCGAAACAATTGATTGGAGA (配列番号199)
配列番号44	TTTTTGTTATATATTTGTCCTGTTAGGTTAAATCACCGCGCCTGATGACGAAGTCGGTGGTAGAATTAGACTAATATTAAATATGTCTCATG (配列番号195) CCTATTAGATATTCCGTATTTCTTTAAGACTGTTATAATACAAATATACTACAAATCATGCAATTTTTGATTTTTAACAAAA (配列番号196)
配列番号45	CAGAAGTCGTTCAAGTTCAAGGTCAAAACGGACAAGGAGACGGTCGAATTATTCAG (配列番号163) GGGAGGGTGACATTCAGAAGTCGTTCAAGTTCAAGGTCAAAACGGACAAGGAGACGGTCGAATTAT (配列番号164) AAGTGTCTTCAACACATTGAAGAAAACTCTCGGTGCAATATATGGAAAGCTCGATGAAAACGGAAATTTTATTGAGAATGAATGTAATAAGTAACTGGAATA (配列番号165) CCGTGGGAGGATTTGGATTTGGTTGAAGACATCAGAAAAATTTTCGAAATGGAATAGAGGGAACCGGAATTTTTTCCGGTTTTTCTTTGTCCTTTCGA (配列番号166)
配列番号48	TTTTTCATTGTTCTCAAATTGTTGGATAATGTTTTGTGTGTTTCATTTTTGTCATTGTGTCACCTTAACTGACAAGGTGGCACATTTTTTATGTCAAT (配列番号200) TTTTCATTGTTCTCAAATTGTTGGATAATGTTTTGTGTGTTTCATTTTTGTCATTGTGTCACCTTAACTGACAAGGTGGCACATTTTTTATGTCAATA (配列番号201) AATATATCTGCTAAGGTCATATTTTTCATTGTTCTCAAATTGTTGGATAATGTTTTGTGTGTTTCATTTTTGTCATTGTGTCACCTTAACTGACAAGGTGGCACATTTTTTATGTCAATATG (配列番号202)

【０１７６】

【表１０】

配列番号52	TCGTTGAATACGATATCGCCGAAACAATTGATTGGAGAAGTACGCTT TGTTTCAAGACATGGAATACGTATGGTTCTCCTCAATGGGACTCGAA GATCAAGAA (配列番号197) ATCGTTGAATACGATATCGCCGAAACAATTGATTGGAGAAGTACGC TTTGTTTCAAGACATGGAATACGTATGGTTCTCCTCAATGGGACTCG AAGATCAAGAACCAG (配列番号198) GAGCTTTTCTGGCAATGTAGACATTAAAGCTGGTATCGTTGAATACG ATATCGCCGAAACAATTGATTGGAGA (配列番号199)
配列番号55	TTTTTCATTGTTCTCAAATTGTTGGATAATGTTTTGTGTGTTTCATTTT TGTCATTGTGTCACCTTAACTGACAAGGTGGCACATTTTTTATGTCA AT (配列番号200) TTTTCATTGTTCTCAAATTGTTGGATAATGTTTTGTGTGTTTCATTTTT GTCATTGTGTCACCTTAACTGACAAGGTGGCACATTTTTTATGTCAA TA (配列番号201) AATATATCTGCTAAGGTCATATTTTTCATTGTTCTCAAATTGTTGGAT AATGTTTTGTGTGTTTCATTTTTGTCATTGTGTCACCTTAACTGACAA GGTGGCACATTTTTTATGTCAATATG (配列番号202)
配列番号56	ACAAATTTTTGATTATGGCACACAAAAAGAACATAGGAGCAGAGAT AGTAAAAACTTACTCTTTTAAGGTGAAGA (配列番号203) TTATTTTATAGGATAATAGAGCTAACAAGCATTAACAATTATTAAAA CGATTTATATTGAAAATAAATTTTGTGGGAATATTTATTTTTACTACC TTTGCATCGTAATACAATTAAACAAATTTTTGATTATGGCA (配列番号204)

【０１７７】

一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号１のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ｔｒａｃｒＲＮＡ配列は、配列番号１５２、配列番号１５３、又は配列番号１５４のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号２のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ｔｒａｃｒＲＮＡ配列は、配列番号１５５、配列番号１５６、配列番号１５７、又は配列番号１５８のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号３のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ｔｒａｃｒＲＮＡ配列は、配列番号１５９、配列番号１６０、又は配列番

号１６１のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号１４のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ｔｒａｃｒＲＮＡ配列は、配列番号１６２のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号１７のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ｔｒａｃｒＲＮＡ配列は、配列番号１６３、配列番号１６４、配列番号１６５、又は配列番号１６６のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号１８のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ｔｒａｃｒＲＮＡ配列は、配列番号１６７又は配列番号１６８のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号２１のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ｔｒａｃｒＲＮＡ配列は、配列番号１６９、配列番号１７０、又は配列番号１７１のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号２２のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ｔｒａｃｒＲＮＡ配列は、配列番号１７２、配列番号１７３、配列番号１７４、又は配列番号１７５のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号２３のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ｔｒａｃｒＲＮＡ配列は、配列番号１７６、配列番号１７７、配列番号１７８、又は配列番号１７９のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号２７のアミノ酸配列と少な

くとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ｔｒａｃｒＲＮＡ配列は、配列番号１８０又は配列番号１８１のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号２９のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ｔｒａｃｒＲＮＡ配列は、配列番号１８２、配列番号１８３、又は配列番号１８４のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号３１のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ｔｒａｃｒＲＮＡ配列は、配列番号１８５、配列番号１８６、配列番号１８７、又は配列番号１８８のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号３２のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ｔｒａｃｒＲＮＡ配列は、配列番号１８９又は配列番号１９０のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号３６のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ｔｒａｃｒＲＮＡ配列は、配列番号１８２、配列番号１８３、又は配列番号１８４のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号３８のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ｔｒａｃｒＲＮＡ配列は、配列番号１８９又は配列番号１９０のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号３９のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ｔｒａｃｒＲＮＡ配列は、配列番号１８５、配列番号１８６、配列番号１８７、又は配列番号１８８のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８

３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号４１のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ｔｒａｃｒＲＮＡ配列は、配列番号１９１、配列番号１９２、配列番号１９３、又は配列番号１９４のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号４３のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ｔｒａｃｒＲＮＡ配列は、配列番号１９７、配列番号１９８、又は配列番号１９９のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号４４のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ｔｒａｃｒＲＮＡ配列は、配列番号１９５又は配列番号１９６のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号４５のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ｔｒａｃｒＲＮＡ配列は、配列番号１６３、配列番号１６４、配列番号１６５、又は配列番号１６６のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号４８のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ｔｒａｃｒＲＮＡ配列は、配列番号２００、配列番号２０１、又は配列番号２０２のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号５２のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ｔｒａｃｒＲＮＡ配列は、配列番号１９７、配列番号１９８、又は配列番号１９９のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号５５のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、

８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ｔｒａｃｒＲＮＡ配列は、配列番号２００、配列番号２０１、又は配列番号２０２のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。一部の実施形態において、ＣＲＩＳＰＲ関連タンパク質は、配列番号５６のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含み、ｔｒａｃｒＲＮＡ配列は、配列番号２０３又は配列番号２０４のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む。

【０１７８】

ＲＮＡガイド配列は、ＣＲＩＳＰＲ複合体の形成及び標的への結合の成功は許容するが、同時にヌクレアーゼ活性の成功は許容しない（即ち、ヌクレアーゼ活性のない／インデルを生じさせない）ような方法で修飾されてもよい。こうした修飾ガイド配列は、「デッドガイド」又は「デッドガイド配列」と称される。こうしたデッドガイド又はデッドガイド配列はヌクレアーゼ活性の点で触媒的に不活性又はコンホメーション的に不活性であってもよい。デッドガイド配列は、典型的には、活性なＲＮＡ切断を生じるそれぞれのガイド配列よりも短い。一部の実施形態において、デッドガイドは、ヌクレアーゼ活性を有するそれぞれのガイドＲＮＡと比べて５％、１０％、２０％、３０％、４０％、又は５０％短い。ＲＮＡガイドのデッドガイド配列は、１３～１５ヌクレオチド長（例えば、１３、１４、又は１５ヌクレオチド長）、１５～１９ヌクレオチド長、又は１７～１８ヌクレオチド長（例えば、１７ヌクレオチド長）であってもよい。

【０１７９】

従って、一態様において、本開示は、本明細書に記載されるとおりの機能性ＣＬＵＳＴ．０９１９７９ＣＲＩＳＰＲエフェクターと、ＲＮＡガイドとを含む天然に存在しない又はエンジニアリングされたＣＲＩＳＰＲシステムを提供し、ここでＲＮＡガイドはデッドガイド配列を含み、従ってＲＮＡガイドは、検出可能な切断活性なしにＣＲＩＳＰＲシステムが細胞の目的のゲノム遺伝子座に向けられるような標的配列へのハイブリダイズ能を有する。デッドガイドの詳細な説明は、例えば、国際公開第２０１６０９４８７２号パンフレット（全体として参照により本明細書に援用される）に記載される。

【０１８０】

誘導性ＲＮＡガイド

ＲＮＡガイドは、誘導性システムの成分として作成することができる。このシステムの誘導可能な性質により、遺伝子編集又は遺伝子発現の時空間的制御が可能となる。一部の実施形態において、誘導性システムの刺激としては、例えば、電磁放射線、音響エネルギー、化学エネルギー、及び／又は熱エネルギーが挙げられる。

【０１８１】

一部の実施形態において、ＲＮＡガイドの転写は、誘導性プロモーター、例えば、テトラサイクリン又はドキシサイクリン制御下での転写活性化（Ｔｅｔ－Ｏｎ及びＴｅｔ－Ｏｆｆ発現システム）、ホルモン誘導性遺伝子発現システム（例えば、エクジソン誘導性遺伝子発現システム）、及びアラビノース誘導性遺伝子発現システムによって調節することができる。誘導性システムの他の例としては、例えば、小分子２ハイブリッド転写活性化システム（ＦＫＢＰ、ＡＢＡ等）、光誘導性システム（フィトクロム、ＬＯＶドメイン、又はクリプトクロム）、又は光誘導性転写エフェクター（ＬＩＴＥ）が挙げられる。これらの誘導性システムは、例えば、国際公開第２０１６２０５７６４号パンフレット及び米国特許第８７９５９６５号明細書（これらはそれぞれ全体として参照により本明細書に援

用される）に記載されている。
【０１８２】
化学修飾
ガイドＲＮＡのリン酸骨格、糖、及び／又は塩基に化学修飾を適用することができる。ホスホロチオエートなどの骨格修飾はリン酸骨格上の電荷を修飾し、オリゴヌクレオチドの送達及びヌクレアーゼ耐性に役立つ（例えば、Ｅｃｋｓｔｅｉｎ，“Ｐｈｏｓｐｈｏｒｏｔｈｉｏａｔｅｓ，ｅｓｓｅｎｔｉａｌ ｃｏｍｐｏｎｅｎｔｓ ｏｆ ｔｈｅｒａｐｅｕｔｉｃ ｏｌｉｇｏｎｕｃｌｅｏｔｉｄｅｓ，”Ｎｕｃｌ．Ａｃｉｄ Ｔｈｅｒ．，２４（２０１４），ｐｐ．３７４－３８７を参照のこと）；２’－Ｏ－メチル（２’－ＯＭｅ）、２’－Ｆ、及びロックド核酸（ＬＮＡ）などの糖修飾は、塩基対合及びヌクレアーゼ耐性の両方を亢進させる（例えば、Ａｌｌｅｒｓｏｎ ｅｔ ａｌ．“Ｆｕｌｌｙ ２‘－ｍｏｄｉｆｉｅｄ ｏｌｉｇｏｎｕｃｌｅｏｔｉｄｅ ｄｕｐｌｅｘｅｓ ｗｉｔｈ ｉｍｐｒｏｖｅｄ ｉｎ ｖｉｔｒｏ ｐｏｔｅｎｃｙ ａｎｄ ｓｔａｂｉｌｉｔｙ ｃｏｍｐａｒｅｄ ｔｏ ｕｎｍｏｄｉｆｉｅｄ ｓｍａｌｌ ｉｎｔｅｒｆｅｒｉｎｇ ＲＮＡ，”Ｊ．Ｍｅｄ．Ｃｈｅｍ．，４８．４（２００５）：９０１－９０４を参照のこと）。化学修飾塩基、とりわけ２－チオウリジン又はＮ６－メチルアデノシンなどは、より強い塩基対合又はより弱い塩基対合のいずれも可能にすることができる（例えば、Ｂｒａｍｓｅｎ ｅｔ ａｌ．，“Ｄｅｖｅｌｏｐｍｅｎｔ ｏｆ ｔｈｅｒａｐｅｕｔｉｃ－ｇｒａｄｅ ｓｍａｌｌ ｉｎｔｅｒｆｅｒｉｎｇ ＲＮＡｓ ｂｙ ｃｈｅｍｉｃａｌ ｅｎｇｉｎｅｅｒｉｎｇ”Ｆｒｏｎｔ．Ｇｅｎｅｔ．，２０１２ Ａｕｇ ２０；３：１５４を参照のこと）。加えて、ＲＮＡは、蛍光色素、ポリエチレングリコール、又はタンパク質を含めた種々の機能性部分との５’末端及び３’末端の両方のコンジュゲーションに適している。
【０１８３】
化学的に合成されるＲＮＡガイド分子には、幅広い種類の修飾を適用することができる。例えば、オリゴヌクレオチドを２’－ＯＭｅで修飾してヌクレアーゼ耐性を改善すると、ワトソン・クリック塩基対合の結合エネルギーを変化させることができる。更には、２’－ＯＭｅ修飾は、オリゴヌクレオチドがトランスフェクション試薬、タンパク質又は細胞中の任意の他の分子とどのように相互作用するかに影響を及ぼし得る。これらの修飾の効果は経験的試験によって決定することができる。
【０１８４】
一部の実施形態において、ＲＮＡガイドは１つ以上のホスホロチオエート修飾を含む。一部の実施形態において、ＲＮＡガイドは、塩基対合を亢進させること及び／又はヌクレアーゼ耐性を増加させることを目的として１つ以上のロックド核酸を含む。
【０１８５】
これらの化学修飾の概要については、例えば、Ｋｅｌｌｅｙ ｅｔ ａｌ．，“Ｖｅｒｓａｔｉｌｉｔｙ ｏｆ ｃｈｅｍｉｃａｌｌｙ ｓｙｎｔｈｅｓｉｚｅｄ ｇｕｉｄｅ ＲＮＡｓ ｆｏｒ ＣＲＩＳＰＲ－Ｃａｓ９ ｇｅｎｏｍｅ ｅｄｉｔｉｎｇ，”Ｊ．Ｂｉｏｔｅｃｈｎｏｌ．２０１６ Ｓｅｐ １０；２３３：７４－８３；国際公開第２０１６２０５７６４号パンフレット；及び米国特許第８７９５９６５号明細書（この各々が全体として参照により援用される）を参照することができる。
【０１８６】
配列修飾
本明細書に記載されるＲＮＡガイド、ｔｒａｃｒＲＮＡ及びｃｒＲＮＡの配列及び長さは最適化することができる。一部の実施形態において、ＲＮＡガイドの最適化された長さは、プロセシングされた形態のｔｒａｃｒＲＮＡ及び／若しくはｃｒＲＮＡを同定することによるか、又はｃｒＲＮＡのＲＮＡガイドについての経験的な長さ研究によって決定されてもよい。
【０１８７】
ＲＮＡガイドはまた、１つ以上のアプタマー配列も含むことができる。アプタマーは、特異的な標的分子に結合することのできるオリゴヌクレオチド又はペプチド分子である。

アプタマーは、遺伝子エフェクター、遺伝子アクチベーター、又は遺伝子リプレッサーに特異的であってもよい。一部の実施形態において、アプタマーはタンパク質に特異的であり、次にはそのタンパク質が特異的遺伝子エフェクター、遺伝子アクチベーター、又は遺伝子リプレッサーに特異的であって、それを動員し／それに結合するものであってもよい。エフェクター、アクチベーター、又はリプレッサーは融合タンパク質の形態で存在することができる。一部の実施形態において、ＲＮＡガイドは、同じアダプタータンパク質に特異的な２つ以上のアプタマー配列を有する。一部の実施形態において、２つ以上のアプタマー配列は異なるアダプタータンパク質に特異的である。アダプタータンパク質としては、例えば、ＭＳ２、ＰＰ７、Ｑβ、Ｆ２、ＧＡ、ｆｒ、ＪＰ５０１、Ｍ１２、Ｒ１７、ＢＺ１３、ＪＰ３４、ＪＰ５００、ＫＵ１、Ｍ１１、ＭＸ１、ＴＷ１８、ＶＫ、ＳＰ、ＦＩ、ＩＤ２、ＮＬ９５、ＴＷ１９、ＡＰ２０５、φＣｂ５、φＣｂ８ｒ、φＣｂ１２ｒ、φＣｂ２３ｒ、７ｓ、及びＰＲＲ１を挙げることができる。従って、一部の実施形態において、アプタマーは、本明細書に記載されるとおりのアダプタータンパク質のうちのいずれか１つに特異的に結合する結合タンパク質から選択される。一部の実施形態において、アプタマー配列はＭＳ２ループである。アプタマーの詳細な説明については、例えば、Ｎｏｗａｋ ｅｔ ａｌ．，“Ｇｕｉｄｅ ＲＮＡ ｅｎｇｉｎｅｅｒｉｎｇ ｆｏｒ ｖｅｒｓａｔｉｌｅ Ｃａｓ９ ｆｕｎｃｔｉｏｎａｌｉｔｙ，”Ｎｕｃｌ．Ａｃｉｄ．Ｒｅｓ．，２０１６ Ｎｏｖ １６；４４（２０）：９５５５－９５６４；及び国際公開第２０１６２０５７６４号パンフレット（これらはそれぞれ全体として参照により本明細書に援用される）を参照することができる。

【０１８８】

ガイド：標的配列一致要件

ＣＲＩＳＰＲシステムでは、ガイド配列とその対応する標的配列との間の相補性の程度は、約５０％、６０％、７５％、８０％、８５％、９０％、９５％、９７．５％、９９％、又は１００％であってもよい。オフターゲット相互作用を減少させるため、例えば、相補性が低い標的配列と相互作用するガイドを減少させるため、ＣＲＩＳＰＲシステムに突然変異を導入して、ＣＲＩＳＰＲシステムが標的配列と、８０％、８５％、９０％、又は９５％より高い相補性を有するオフターゲット配列との間を区別できるようにしてもよい。一部の実施形態において、相補性の程度は、８０％～９５％、例えば、約８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、又は９５％である（例えば、１８ヌクレオチドを有する標的と、１、２、又は３個のミスマッチを有する１８ヌクレオチドのオフターゲットとの間を区別する）。従って、一部の実施形態において、ガイド配列とその対応する標的配列との間の相補性の程度は、９４．５％、９５％、９５．５％、９６％、９６．５％、９７％、９７．５％、９８％、９８．５％、９９％、９９．５％、又は９９．９％より高い。一部の実施形態において、相補性の程度は１００％である。

【０１８９】

当該分野では、機能性となるのに十分な相補性があるならば、完全な相補性は要件とならないことが公知である。スペーサー／標的に沿ったミスマッチの位置を含め、スペーサー配列と標的配列との間へのミスマッチ、例えば１個以上のミスマッチ、例えば１又は２個のミスマッチの導入により、切断効率の調節を生かすことができる。ミスマッチ、例えば二重ミスマッチが中心寄りに位置するほど（即ち、３’末端又は５’末端にあるのでない）；切断効率がより大きい影響を受ける。従って、スペーサー配列に沿ったミスマッチ位置の選択により、切断効率を調節することができる。例えば、標的の１００％未満の切断が（例えば、細胞集団中で）所望される場合、スペーサー配列にスペーサーと標的配列との間の１又は２個のミスマッチを導入してもよい。

【０１９０】

ＣＲＩＳＰＲシステムの使用方法

本明細書に記載されるＣＲＩＳＰＲシステムには、非常に多数の細胞型における標的ポリヌクレオチドの修飾（例えば、欠失、挿入、転位、不活性化、又は活性化）を含め、多

種多様な有用性がある。このＣＲＩＳＰＲシステムは、例えば、ＤＮＡ／ＲＮＡ検出（例えば、特異的高感度酵素レポーターアンロッキング（ｓｐｅｃｉｆｉｃ ｈｉｇｈ ｓｅｎｓｉｔｉｖｉｔｙ ｅｎｚｙｍａｔｉｃ ｒｅｐｏｒｔｅｒ ｕｎｌｏｃｋｉｎｇ：ＳＨＥＲＬＯＣＫ））、核酸の追跡及び標識、エンリッチメントアッセイ（バックグラウンドからの所望の配列の抽出）、循環腫瘍ＤＮＡの検出、次世代ライブラリの調製、薬物スクリーニング、疾患診断及び予後判定、及び様々な遺伝的障害の治療において、幅広い範囲にわたる適用を有する。

【０１９１】

ＤＮＡ／ＲＮＡ検出

一態様において、本明細書に記載されるＣＲＩＳＰＲシステムは、ＤＮＡ／ＲＮＡ検出において使用することができる。シングルエフェクターＲＮＡ誘導型ＤＮアーゼをＣＲＩＳＰＲ ＲＮＡ（ｃｒＲＮＡ）で再プログラム化することにより、特異的一本鎖ＤＮＡ（ｓｓＤＮＡ）センシング用のプラットフォームがもたらされ得る。そのＤＮＡ標的の認識時、活性化したＶ型単一エフェクターＤＮＡガイドＤＮａｓｅは、近隣非標的ｓｓＤＮＡの「コラテラル」切断に関与する。このｃｒＲＮＡによってプログラム化されるコラテラル切断活性により、ＣＲＩＳＰＲシステムが特異的ＤＮＡの存在を標識ｓｓＤＮＡの非特異的分解によって検出することが可能となる。

【０１９２】

ＤＮＡ検出適用においては、コラテラルｓｓＤＮＡ活性をレポーターと組み合わせることができ、例えば、ＤＮＡエンドヌクレアーゼ標的化ＣＲＩＳＰＲトランスレポーター（ＤＮＡ Ｅｎｄｏｎｕｃｌｅａｓｅ－Ｔａｒｇｅｔｅｄ ＣＲＩＳＰＲ ｔｒａｎｓ ｒｅｐｏｒｔｅｒ：ＤＥＴＥＣＴＲ）法と呼ばれる方法などであり、これは、アトモル濃度のＤＮＡ検出感度を実現する（例えば、Ｃｈｅｎ ｅｔ ａｌ．，Ｓｃｉｅｎｃｅ，３６０（６３８７）：４３６－４３９，２０１８を参照のこと）（これは全体として参照により本明細書に援用される）。本明細書に記載される酵素を使用する一つの適用は、インビトロ環境における非特異的ｓｓＤＮＡの分解である。フルオロフォア及び消光剤に連結した「レポーター」ｓｓＤＮＡ分子もまた、未知のＤＮＡ試料（一本鎖又は二本鎖のいずれか）と共にこのインビトロシステムに加えることができる。未知のＤＮＡ片中に標的配列を認識すると、このエフェクター複合体がレポーターｓｓＤＮＡを切断し、蛍光リードアウトが生じる。

【０１９３】

他の実施形態において、ＳＨＥＲＬＯＣＫ法（特異的高感度酵素レポーターアンロッキング（Ｓｐｅｃｉｆｉｃ Ｈｉｇｈ Ｓｅｎｓｉｔｉｖｉｔｙ Ｅｎｚｙｍａｔｉｃ Ｒｅｐｏｒｔｅｒ ＵｎＬＯＣＫｉｎｇ））もまた、核酸増幅及びレポーターｓｓＤＮＡのコラテラル切断に基づいたアトモル濃度（又は単一分子）感度のインビトロ核酸検出プラットフォームを提供し、標的のリアルタイム検出を可能にする。ＳＨＥＲＬＯＣＫにおけるＣＲＩＳＰＲの使用方法については、例えば、Ｇｏｏｔｅｎｂｅｒｇ，ｅｔ ａｌ．“Ｎｕｃｌｅｉｃ ａｃｉｄ ｄｅｔｅｃｔｉｏｎ ｗｉｔｈ ＣＲＩＳＰＲ－Ｃａｓ１３ａ／Ｃ２ｃ２，”Ｓｃｉｅｎｃｅ，３５６（６３３６）：４３８－４４２（２０１７）（これは全体として参照により本明細書に援用される）に詳細に記載される。

【０１９４】

一部の実施形態において、本明細書に記載されるＣＲＩＳＰＲシステムは、マルチプレックス化したエラーロバストな蛍光インサイチュハイブリダイゼーション（ｍｕｌｔｉｐｌｅｘｅｄ ｅｒｒｏｒ－ｒｏｂｕｓｔ ｆｌｕｏｒｅｓｃｅｎｃｅ ｉｎ ｓｉｔｕ ｈｙｂｒｉｄｉｚａｔｉｏｎ：ＭＥＲＦＩＳＨ）において使用することができる。こうした方法については、例えば、Ｃｈｅｎ ｅｔ ａｌ．，“Ｓｐａｔｉａｌｌｙ ｒｅｓｏｌｖｅｄ，ｈｉｇｈｌｙ ｍｕｌｔｉｐｌｅｘｅｄ ＲＮＡ ｐｒｏｆｉｌｉｎｇ ｉｎ ｓｉｎｇｌｅ ｃｅｌｌｓ，”Ｓｃｉｅｎｃｅ，２０１５ Ａｐｒ ２４；３４８（６２３３）：ａａａ６０９０（これは全体として参照により本明細書に援用される）に記載されている。

【０１９５】
核酸の追跡及び標識
細胞過程は、タンパク質、ＲＮＡ、及びＤＮＡの間での分子相互作用網に依存する。タンパク質－ＤＮＡ及びタンパク質－ＲＮＡ相互作用の正確な検出は、かかる過程を理解する鍵である。インビトロ近接性標識技法は、レポーター基、例えば光活性化可能な基と組み合わせたアフィニティータグを用いることにより、インビトロで目的のタンパク質又はＲＮＡの近くにあるポリペプチド及びＲＮＡを標識する。紫外線照射後、光活性化可能な基がタグ付加分子にごく近接したタンパク質及び他の分子と反応し、それによってそれらを標識する。標識された相互作用分子は、続いて回収し、同定することができる。このＲＮＡターゲティングエフェクタータンパク質を使用して、例えば、プローブを選択のＲＮＡ配列に標的化することができる。こうした適用はまた、動物モデルにおいても疾患又は培養が困難な細胞型のインビボイメージングに適用することができる。核酸の追跡及び標識方法については、例えば、米国特許第８７９５９６５号明細書；国際公開第２０１６２０５７６４号パンフレット；及び国際公開第２０１７０７０６０５号パンフレット（これらの各々は、本明細書において全体として参照により援用される）に記載されている。
【０１９６】
ハイスループットスクリーニング
本明細書に記載されるＣＲＩＳＰＲシステムは、次世代シーケンシング（ＮＧＳ）ライブラリの調製に使用することができる。例えば、費用対効果の高いＮＧＳライブラリを作成するため、ＣＲＩＳＰＲシステムを使用して標的遺伝子のコード配列を破壊することができ、同時に次世代シーケンシングによって（例えば、Ｉｏｎ　Ｔｏｒｒｅｎｔ　ＰＧＭシステムで）、ＣＲＩＳＰＲエフェクターがトランスフェクトされたクローンをスクリーニングすることができる。ＮＧＳライブラリの調製方法に関する詳細な説明については、例えば、Ｂｅｌｌ　ｅｔ　ａｌ．，“Ａ　ｈｉｇｈ－ｔｈｒｏｕｇｈｐｕｔ　ｓｃｒｅｅｎｉｎｇ　ｓｔｒａｔｅｇｙ　ｆｏｒ　ｄｅｔｅｃｔｉｎｇ　ＣＲＩＳＰＲ－Ｃａｓ９　ｉｎｄｕｃｅｄ　ｍｕｔａｔｉｏｎｓ　ｕｓｉｎｇ　ｎｅｘｔ－ｇｅｎｅｒａｔｉｏｎ　ｓｅｑｕｅｎｃｉｎｇ，”ＢＭＣ　Ｇｅｎｏｍｉｃｓ，１５．１（２０１４）：１００２（これは全体として参照により本明細書に援用される）を参照することができる。
【０１９７】
エンジニアリングされた細胞
微生物（例えば、大腸菌（Ｅ．ｃｏｌｉ）、酵母、及び微細藻類）は、合成生物学に広く用いられている。合成生物学の発展には、様々な臨床応用を含め、幅広い有用性がある。例えば、プログラム可能なＣＲＩＳＰＲシステムを使用して、例えば癌関連ＲＮＡを標的転写物として用いる標的化した細胞死のため、毒性ドメインのタンパク質を分割することができる。更に、例えばキナーゼ又は酵素などの適切なエフェクターとの融合複合体により、合成生物系においてタンパク質間相互作用が関わる経路に影響を及ぼすことができる。
【０１９８】
一部の実施形態において、ファージ配列を標的化するＲＮＡガイド配列を微生物に導入することができる。従って、本開示はまた、微生物（例えば産生菌株）にファージ感染に対する「ワクチンを接種する」方法も提供する。
【０１９９】
一部の実施形態において、本明細書に提供されるＣＲＩＳＰＲシステムを使用して微生物をエンジニアリングすることにより、例えば、収率を改善し又は発酵効率を改善することができる。例えば、本明細書に記載されるＣＲＩＳＰＲシステムを使用して酵母などの微生物をエンジニアリングすることにより、発酵性糖からバイオ燃料若しくはバイオポリマーを生成し、又は発酵性糖源としての農業廃棄物に由来する植物由来のリグノセルロースを分解することができる。より詳細には、本明細書に記載される方法を使用して、バイオ燃料生産に必要な内因性遺伝子の発現を修飾し、及び／又はバイオ燃料合成を妨げ得る内因性遺伝子を修飾することができる。これらの微生物エンジニアリング方法については

、例えば、Ｖｅｒｗａａｌ ｅｔ ａｌ．，“ＣＲＩＳＰＲ／Ｃｐｆ１ ｅｎａｂｌｅｓ ｆａｓｔ ａｎｄ ｓｉｍｐｌｅ ｇｅｎｏｍｅ ｅｄｉｔｉｎｇ ｏｆ Ｓａｃｃｈａｒｏｍｙｃｅｓ ｃｅｒｅｖｉｓｉａｅ，”Ｙｅａｓｔ，２０１７ Ｓｅｐ ８．ｄｏｉ：１０．１００２／ｙｅａ．３２７８；及びＨｌａｖｏｖａ ｅｔ ａｌ．，“Ｉｍｐｒｏｖｉｎｇ ｍｉｃｒｏａｌｇａｅ ｆｏｒ ｂｉｏｔｅｃｈｎｏｌｏｇｙ－ｆｒｏｍ ｇｅｎｅｔｉｃｓ ｔｏ ｓｙｎｔｈｅｔｉｃ ｂｉｏｌｏｇｙ，”Ｂｉｏｔｅｃｈｎｏｌ．Ａｄｖ．，２０１５ Ｎｏｖ １；３３：１１９４－２０３（これらはそれぞれ全体として参照により本明細書に援用される）に記載されている。

【０２００】

一部の実施形態において、本明細書で提供されるＣＲＩＳＰＲシステムは、真核細胞又は真核生物をエンジニアリングするために使用することができる。例えば、本明細書に記載のＣＲＩＳＰＲシステムは、植物細胞、真菌細胞、哺乳動物細胞、爬虫類細胞、昆虫細胞、鳥類細胞、魚類細胞、寄生虫細胞、節足動物細胞、無脊椎動物細胞、脊椎動物細胞、げっ歯類細胞、マウス細胞、ラット細胞、霊長類細胞、非ヒト霊長類細胞、又はヒト細胞に限定されない真核細胞をエンジニアリングするために使用することができる。一部の実施形態において、真核細胞はインビトロ培養物中にある。一部の実施形態において、真核細胞はインビボである。一部の実施形態において、真核細胞はエキソビボである。

【０２０１】

一部の実施形態において、細胞は細胞系に由来する。組織培養のための広範な細胞系は当技術分野で知られている。細胞系の例としては、限定されるものではないが、２９３Ｔ、ＭＦ７、Ｋ５６２、ＨｅＬａ、及びそれらのトランスジェニック変種が挙げられる。細胞系は、当業者に知られた様々な供給源から入手可能である（例えば、Ａｍｅｒｉｃａｎ Ｔｙｐｅ Ｃｕｌｔｕｒｅ Ｃｏｌｌｅｃｔｉｏｎ（ＡＴＣＣ）（Ｍａｎａｓｓａｓ，Ｖａ．）を参照されたい）。一部の実施形態において、１つ以上の核酸（例えば、ヌクレアーゼポリペプチドコードベクター及びＲＮＡガイド）によってトランスフェクトされた細胞を使用して、１つ以上のベクター由来配列を含む新たな細胞系を樹立して、標的核酸又は標的遺伝子座の修飾を含む新たな細胞系を樹立する。一部の実施形態において、細胞は不死又は不死化細胞である。

【０２０２】

一部の実施形態において、細胞は初代細胞である。一部の実施形態において、細胞は、分化全能性幹細胞（例えば、万能）、多能性幹細胞、複能性幹細胞、少能性幹細胞、又は単能性幹細胞などの幹細胞である。一部の実施形態において、細胞は誘導多能性幹細胞（ｉＰＳＣ）であり、又はｉＰＳＣに由来する。一部の実施形態において、細胞は分化細胞である。例えば、一部の実施形態において、分化細胞は筋細胞、（例えば、ミオサイト）、脂肪細胞（例えば、アジポサイト）、骨細胞（例えば、骨芽細胞、骨細胞、破骨細胞）、血球（例えば、単球、リンパ球、好中球、好酸球、好塩基球、マクロファージ、赤血球、又は血小板）、神経細胞（例えば、ニューロン）、上皮細胞、免疫細胞（例えば、リンパ球、好中球、単球、又はマクロファージ）、肝細胞（例えば、ヘパトサイト）、線維芽細胞、又は生殖細胞である。一部の実施形態において、細胞は最終分化細胞である。例えば、一部の実施形態において、最終分化細胞は神経細胞、アジポサイト、心筋細胞、骨格筋細胞、表皮細胞、又は腸細胞である。一部の実施形態において、細胞は哺乳動物細胞、例えば、ヒト細胞又はネズミ細胞である。一部の実施形態において、ネズミ細胞は野生型マウス、免疫抑制マウス、又は疾患特異的マウスモデルに由来する。

【０２０３】

遺伝子ドライブ

遺伝子ドライブは、特定の遺伝子又は遺伝子群の遺伝形質に有利な偏りが出る現象である。本明細書に記載されるＣＲＩＳＰＲシステムを使用して遺伝子ドライブを構築することができる。例えば、遺伝子の特定のアレルを標的化して破壊することにより、細胞に第２のアレルをコピーさせて配列を固定するようにＣＲＩＳＰＲシステムを設計することができる。このコピーのため、第１のアレルが第２のアレルに変換されることになり、子孫

に第２のアレルが遺伝する可能性が高くなる。どのように本明細書に記載されるＣＲＩＳＰＲシステムを使用して遺伝子ドライブを構築するかに関する詳細な方法については、例えば、Ｈａｍｍｏｎｄ ｅｔ ａｌ．，“Ａ ＣＲＩＳＰＲ－Ｃａｓ９ ｇｅｎｅ ｄｒｉｖｅ ｓｙｓｔｅｍ ｔａｒｇｅｔｉｎｇ ｆｅｍａｌｅ ｒｅｐｒｏｄｕｃｔｉｏｎ ｉｎ ｔｈｅ ｍａｌａｒｉａ ｍｏｓｑｕｉｔｏ ｖｅｃｔｏｒ Ａｎｏｐｈｅｌｅｓ ｇａｍｂｉａｅ，”Ｎａｔ．Ｂｉｏｔｅｃｈｎｏｌ．，２０１６ Ｊａｎ；３４（１）：７８－８３（これは全体として参照により本明細書に援用される）に記載されている。

【０２０４】

プール型スクリーニング

本明細書に記載されるとおり、プール型ＣＲＩＳＰＲスクリーニングは、細胞増殖、薬剤耐性、及びウイルス感染などの生物学的機構に関与する遺伝子を同定するための強力なツールである。細胞がバルクで本明細書に記載されるＲＮＡガイドコードベクターのライブラリによって形質導入され、選択的チャレンジの適用前及び適用後にｇＲＮＡの分布が測定される。プール型ＣＲＩＳＰＲスクリーンは、細胞生存及び増殖に影響を及ぼす機構に対して良好に機能し、個々の遺伝子の活性の測定にまで（例えば、エンジニアリングされたレポーター細胞株を使用することにより）拡張することができる。一度に１つの遺伝子のみが標的化されるアレイ化されたＣＲＩＳＰＲスクリーンでは、ＲＮＡ－ｓｅｑをリードアウトとして使用することが可能になる。一部の実施形態において、本明細書に記載されるとおりのＣＲＩＳＰＲシステムは単一細胞ＣＲＩＳＰＲスクリーンに使用することができる。プール型ＣＲＩＳＰＲスクリーニングに関する詳細な説明については、例えば、Ｄａｔｌｉｎｇｅｒ ｅｔ ａｌ．，“Ｐｏｏｌｅｄ ＣＲＩＳＰＲ ｓｃｒｅｅｎｉｎｇ ｗｉｔｈ ｓｉｎｇｌｅ－ｃｅｌｌ ｔｒａｎｓｃｒｉｐｔｏｍｅ ｒｅａｄ－ｏｕｔ，”Ｎａｔ．Ｍｅｔｈｏｄｓ．，２０１７ Ｍａｒ；１４（３）：２９７－３０１（これは全体として参照により本明細書に援用される）を参照することができる。

【０２０５】

飽和突然変異誘発（「バッシング（ｂａｓｈｉｎｇ）」）

本明細書に記載されるＣＲＩＳＰＲシステムはインサイチュー飽和突然変異誘発に使用することができる。一部の実施形態では、プール型ＲＮＡガイドライブラリを使用して、特定の遺伝子又は調節エレメントに関するインサイチュー飽和突然変異誘発を実施することができる。かかる方法では、決定的な最小の特徴及びそれらの遺伝子又は調節エレメント（例えば、エンハンサー）の個別的な脆弱性を明らかにすることができる。これらの方法については、例えば、Ｃａｎｖｅｒ ｅｔ ａｌ．，“ＢＣＬ１１Ａ ｅｎｈａｎｃｅｒ ｄｉｓｓｅｃｔｉｏｎ ｂｙ Ｃａｓ９－ｍｅｄｉａｔｅｄ ｉｎ ｓｉｔｕ ｓａｔｕｒａｔｉｎｇ ｍｕｔａｇｅｎｅｓｉｓ，”Ｎａｔｕｒｅ，２０１５ Ｎｏｖ １２；５２７（７５７７）：１９２－７（これは全体として参照により本明細書に援用される）に記載されている。

【０２０６】

治療上の適用

一部の実施形態において、本明細書に記載されるＣＲＩＳＰＲシステムを使用して、標的核酸を編集して、標的核酸を修飾することができる（例えば、１つ以上のアミノ酸残基を挿入、欠失、又は変異させることによって）。例えば、一部の実施形態において、本明細書に記載されるＣＲＩＳＰＲシステムは、望ましい核酸配列を含む外因性ドナー鋳型核酸（例えば、ＤＮＡ分子又はＲＮＡ分子）を含む。本明細書に記載されるＣＲＩＳＰＲシステムで誘導される切断イベントの分解時に、細胞の分子機構は、切断イベントを修復及び／又は分解する際に、外因性ドナー鋳型核酸を利用することができる。或いは、細胞の分子機構は、切断イベントを修復及び／又は分解する際に、内因性鋳型を利用することができる。一部の実施形態において、本明細書に記載されるＣＲＩＳＰＲシステムは、挿入、欠失、及び／又は点突然変異を生じる標的核酸を修飾するために使用され得る。一部の実施形態において、挿入は、傷のない挿入である（すなわち、標的核酸への意図された核酸配列の挿入は、切断イベントの分解時に追加の意図されない核酸配列を生じない）。ド

ナー鋳型核酸は、二本鎖又は一本鎖核酸分子（例えば、ＤＮＡ又はＲＮＡ）であってもよい。外因性ドナー鋳型核酸の設計方法については、例えば、国際公開第２０１６０９４８７４号パンフレット（この内容全体が参照により本明細書に明示的に援用される）に記載されている。

【０２０７】

別の態様において、本開示は、ＲＮＡ配列特異的干渉；ＲＮＡ配列特異的遺伝子調節；ＲＮＡ、ＲＮＡ産物、ｌｎｃＲＮＡ、非コードＲＮＡ、核ＲＮＡ、又はｍＲＮＡのスクリーニング；突然変異誘発；ＲＮＡスプライシングの阻害；蛍光インサイチュハイブリダイゼーション；育種；細胞休眠の誘導；細胞周期停止の誘導；細胞成長及び／又は細胞増殖の減少；細胞アネルギーの誘導；細胞アポトーシスの誘導；細胞壊死の誘導；細胞死の誘導；又はプログラムされた細胞死の誘導からなる群から選択される方法における本明細書に記載されるシステムの使用を提供する。

【０２０８】

本明細書に記載されるＣＲＩＳＰＲシステムは、様々な治療上の適用を有し得る。一部の実施形態において、新規ＣＲＩＳＰＲシステムは、様々な疾患及び障害、例えば、遺伝性障害（例えば、単一遺伝子疾患）又はヌクレアーゼ活性によって治療することができる疾患（例えば、Ｐｃｓｋ９ターゲティング又はＢＣＬ１１ａターゲティング）を治療するために使用することができる。一部の実施形態において、本明細書に記載される方法は、対象、例えば、ヒト患者などの哺乳類の治療に用いられる。哺乳類対象はまた、イヌ、ネコ、ウマ、サル、ウサギ、ラット、マウス、雌ウシ、ヤギ、又はヒツジなど、家畜化された哺乳類であってもよい。

【０２０９】

この方法は、病態又は疾患が感染性であることを含んでもよく、ここで感染性病原体は、ヒト免疫不全ウイルス（ＨＩＶ）、単純ヘルペスウイルス１型（ＨＳＶ１）、及び単純ヘルペスウイルス２型（ＨＳＶ２）からなる群から選択される。

【０２１０】

一態様において、本明細書に記載されるＣＲＩＳＰＲシステムは、ＲＮＡ、毒性ＲＮＡ、及び／又は変異ＲＮＡ（例えば、スプライシング欠陥又はトランケーション）の過剰発現によって引き起こされる疾患を治療するために使用することができる。例えば、有毒なＲＮＡの発現は、核封入体の形成及び脳、心臓、又は骨格筋の遅発性変性変化に関連し得る。一部の実施形態において、障害は筋強直性ジストロフィーである。筋緊張性ジストロフィーにおいて、有毒なＲＮＡの主な病原性効果は、結合タンパク質を隔離し、選択的スプライシングの調節を損なうことである（例えば、Ｏｓｂｏｒｎｅ ｅｔ ａｌ．，“ＲＮＡ－ｄｏｍｉｎａｎｔ ｄｉｓｅａｓｅｓ，” Ｈｕｍ．Ｍｏｌ．Ｇｅｎｅｔ．，２００９ Ａｐｒ １５；１８（８）：１４７１－８１を参照）。筋強直性ジストロフィー（ｄｙｓｔｒｏｐｈｉａ ｍｙｏｔｏｎｉｃａ（ＤＭ））は、極めて広範囲の臨床的特徴を生じるため、遺伝学者にとって特に興味深いものである。現在ＤＭ１型（ＤＭ１）と呼ばれている古典的な形態のＤＭは、細胞質ゾルプロテインキナーゼをコードする遺伝子であるＤＭＰＫの３’非翻訳領域（ＵＴＲ）におけるＣＴＧリピートの拡大によって引き起こされる。本明細書に記載されるＣＲＩＳＰＲシステムは、過剰発現されたＲＮＡ又は毒性ＲＮＡ、例えば、ＤＭＰＫ遺伝子、又はＤＭ１骨格筋、心臓、又は脳において誤調節された選択的スプライシングのいずれかを標的とすることができる。

【０２１１】

本明細書に記載されるＣＲＩＳＰＲシステムは、プラダー・ウィリ症候群、脊髄性筋萎縮症（ＳＭＡ）、先天性角化異常症などの様々な疾患を引き起こすＲＮＡ依存性機能に影響を与えるトランス作用性変異を標的とすることもできる。本明細書に記載されるＣＲＩＳＰＲシステムを使用して治療できる疾患のリストは、Ｃｏｏｐｅｒ ｅｔ ａｌ．，“ＲＮＡ ａｎｄ ｄｉｓｅａｓｅ，” Ｃｅｌｌ，１３６．４（２００９）：７７７－７９３及び国際公開第２０１６２０５７６４号パンフレットに要約される（これらはそれぞれ全体として参照により本明細書に援用される）。

【０２１２】
本明細書に記載されるＣＲＩＳＰＲシステムは、例えば、原発性加齢性タウオパチー（ＰＡＲＴ）／神経原線維変化（ＮＦＴ）優勢老人性認知症（アルツハイマー病（ＡＤ）で見られるものと同様のＮＦＴを伴うが、プラークを伴わない）、ボクシング認知症（慢性外傷性脳症）、及び進行性核上性麻痺などの原発性及び続発性タウオパチーを含む、様々なタウオパチーの治療にも使用できる。タウオパチー及びこれらの疾患を治療する方法の有用なリストは、例えば、国際公開第２０１６２０５７６４号パンフレット（本明細書において全体として参照により援用される）に記載されている。
【０２１３】
本明細書に記載されるＣＲＩＳＰＲシステムは、スプライシングの欠陥及び疾患を引き起こし得るシス作用性スプライシングコードを破壊する変異を標的化するためにも使用できる。これらの疾患としては、例えば、ＳＭＮ１遺伝子の欠失に起因する運動ニューロン変性疾患（例えば、脊髄性筋萎縮症）、デュシェンヌ型筋ジストロフィー（ＤＭＤ）、前頭側頭型認知症、及び第１７染色体に関連するパーキンソニズム（ＦＴＤＰ－１７）、及び嚢胞性線維症が挙げられる。
【０２１４】
本明細書に記載されるＣＲＩＳＰＲシステムは、特にＲＮＡウイルスに対する抗ウイルス活性のために更に使用することができる。エフェクタータンパク質は、ウイルスＲＮＡ配列を標的化するために選択された適切なＲＮＡガイドを使用してウイルスＲＮＡを標的化することができる。
【０２１５】
更に、インビトロＲＮＡセンシングアッセイを使用して特定のＲＮＡ基質を検出することができる。ＲＮＡターゲティングエフェクタータンパク質は、生細胞でのＲＮＡベースのセンシングに使用できる。適用例は、例えば、疾患特異的ＲＮＡのセンシングによる診断である。
【０２１６】
本明細書に記載されるＣＲＩＳＰＲシステムの治療用途の詳細な説明は、例えば、米国特許第８７９５９６５号明細書、欧州特許第３００９５１１号明細書、国際公開第２０１６２０５７６４号パンフレット、及び国際公開第２０１７０７０６０５号パンフレット（これらの各々は、本明細書において全体として参照により援用される）に見出すことができる。
【０２１７】
植物における適用
本明細書に記載されるＣＲＩＳＰＲシステムは、植物において幅広い種類の有用性がある。一部の実施形態において、ＣＲＩＳＰＲシステムを使用して植物のゲノムをエンジニアリングすることができる（例えば、生産を向上させる、所望の翻訳後修飾を有する生産品にする、又は工業製品を生産するための遺伝子を導入する）。一部の実施形態において、ＣＲＩＳＰＲシステムを使用して、植物に所望の形質を（例えば、ゲノムに対する遺伝性修飾を伴い又は伴わず）導入し、又は植物細胞若しくは全植物における内因性遺伝子の発現を調節することができる。
【０２１８】
一部の実施形態において、本ＣＲＩＳＰＲシステムを使用して、特異的タンパク質、例えば、アレルゲンタンパク質（例えば、ピーナッツ、ダイズ、レンズマメ、エンドウマメ、サヤマメ、及びヤエナリ中のアレルゲンタンパク質）をコードする遺伝子を同定、編集、及び／又はサイレンシングすることができる。タンパク質をコードする遺伝子を同定、編集、及び／又はサイレンシングする方法に関する詳細な説明については、例えば、Ｎｉｃｏｌａｏｕ ｅｔ ａｌ．，“Ｍｏｌｅｃｕｌａｒ ｄｉａｇｎｏｓｉｓ ｏｆ ｐｅａｎｕｔ ａｎｄ ｌｅｇｕｍｅ ａｌｌｅｒｇｙ，”Ｃｕｒｒ．Ｏｐｉｎ．Ａｌｌｅｒｇｙ Ｃｌｉｎ．Ｉｍｍｕｎｏｌ．，１１（３）：２２２－８（２０１１）、及び国際公開第２０１６２０５７６４号パンフレット（これらはそれぞれ全体として参照により本明細書に援

用される）に記載されている。
【０２１９】
ＣＲＩＳＰＲシステムの送達
本開示及び当技術分野の知識を通じて、本明細書に記載されるＣＲＩＳＰＲシステム、又はその成分、その核酸分子、又はその成分をコードする若しくは提供する核酸分子は、ベクター、例えば、プラスミド、又はウイルス送達ベクターなどの様々な送達システムによって送達することができる。本明細書に開示されるＣＲＩＳＰＲエフェクター及び／又はいずれかのＲＮＡ（例えば、ＲＮＡガイド）は、適切なベクター、例えば、プラスミド、又はアデノ随伴ウイルス（ＡＡＶ）、レンチウイルス、アデノウイルス、及び他のウイルスベクターなどのウイルスベクター、又はそれらの組み合わせを使用して送達することができる。エフェクター及び１つ以上のＲＮＡガイドは、１つ以上のベクター、例えばプラスミド又はウイルスベクターにパッケージングすることができる。
【０２２０】
一部の実施形態において、ベクター、例えばプラスミド又はウイルスベクターは、例えば、筋肉内注射、静脈内投与、経皮投与、鼻腔内投与、経口投与、又は粘膜投与によって目的の組織に送達される。かかる送達は、１回用量又は複数回用量のいずれによるものであってもよい。当業者は、本明細書において実際に送達される投薬量が、ベクターの選択、標的細胞、生物、組織、治療対象の全般的な状態、求められる形質転換／修飾の程度、投与経路、投与様式、及び求められる形質転換／修飾の種類を含むがこれらに限定されず、種々の要因に応じて大きく異なり得ることを理解する。
【０２２１】
特定の実施形態において、送達は、アデノウイルスによるものであり、これは少なくとも1×10^{5}粒子（粒子単位、ｐｕとも称される）のアデノウイルスを含有する１回用量におけるものであることができる。一部の実施形態において、好ましくは用量は少なくとも約1×10^{6}粒子、少なくとも約1×10^{7}粒子、少なくとも約1×10^{8}粒子、及び少なくとも約1×10^{9}粒子のアデノウイルスである。送達方法及び用量については、例えば、国際公開第２０１６２０５７６４号パンフレット及び米国特許第８４５４９７２号明細書（これらはそれぞれ全体として参照により本明細書に援用される）に記載されている。
【０２２２】
一部の実施形態において、送達はプラスミドによるものである。投薬量は、応答を引き出すのに十分な数のプラスミドであり得る。ある場合には、プラスミド組成物中のプラスミドＤＮＡの好適な分量は、約０．１～約２ｍｇであってもよい。プラスミドは概して、（ｉ）プロモーター；（ｉｉ）プロモーターに作動可能に連結された、核酸ターゲティングＣＲＩＳＰＲエフェクターをコードする配列；（ｉｉｉ）選択可能マーカー；（ｉｖ）複製起点；及び（ｖ）（ｉｉ）の下流にある且つそれに作動可能に連結された転写ターミネーターを含むことになる。プラスミドはまた、ＣＲＩＳＰＲ複合体のＲＮＡ成分もコードすることができるが、代わりに、これらのうちの１つ以上が異なるベクターにコードされてもよい。投与頻度は、医学又は獣医学の実践者（例えば、医師、獣医師）、又は当業者の範囲内にある。
【０２２３】
別の実施形態において、送達はリポソーム又はリポフェクチン製剤などによるものであり、当業者に公知の方法によって調製することができる。かかる方法については、例えば、国際公開第２０１６２０５７６４号パンフレット及び米国特許第５５９３９７２号明細書；同第５５８９４６６号明細書；及び同第５５８０８５９号明細書（これらの各々は、本明細書において全体として参照により援用される）に記載されている。
【０２２４】
一部の実施形態において、送達はナノ粒子又はエキソソームによるものである。例えば、エキソソームはＲＮＡ送達に特に有用であることが示されている。
【０２２５】
本明細書に記載されるＣＲＩＳＰＲシステムの１つ以上の成分を細胞に導入する更なる

手段は、細胞透過性ペプチド（ＣＰＰ）の使用によるものである。一部の実施形態では、細胞透過性ペプチドがＣＲＩＳＰＲエフェクターに連結される。一部の実施形態では、ＣＲＩＳＰＲエフェクター及び／又はＲＮＡガイドが１つ以上のＣＰＰとカップリングされ、細胞内部へと輸送する（例えば、植物プロトプラスト）。一部の実施形態では、ＣＲＩＳＰＲエフェクター及び／又は１つ又は複数のＲＮＡガイドが、細胞送達のため１つ以上のＣＰＰにカップリングされている１つ以上の環状又は非環状ＤＮＡ分子によってコードされる。

【０２２６】

ＣＰＰは、生体分子を受容体非依存的に細胞膜を越えて輸送する能力を有するタンパク質又はキメラ配列のいずれかに由来する３５アミノ酸未満の短鎖ペプチドである。ＣＰＰは、カチオン性ペプチド、疎水性配列を有するペプチド、両親媒性ペプチド、プロリンリッチな抗微生物配列を有するペプチド、及びキメラ又は双節型ペプチドであってもよい。ＣＰＰの例としては、例えば、Ｔａｔ（これはＨＩＶ　１型によるウイルス複製に必要な核転写活性化因子タンパク質である）、ペネトラチン、カポジ線維芽細胞成長因子（ＦＧＦ）シグナルペプチド配列、インテグリンβ３シグナルペプチド配列、ポリアルギニンペプチドＡｒｇ配列、グアニンリッチ分子輸送体、及びスイートアローペプチドが挙げられる。ＣＰＰ及びその使用方法については、例えば、Ｈａｅｌｌｂｒｉｎｋ ｅｔ ａｌ．，“Ｐｒｅｄｉｃｔｉｏｎ ｏｆ ｃｅｌｌ－ｐｅｎｅｔｒａｔｉｎｇ ｐｅｐｔｉｄｅｓ，”Ｍｅｔｈｏｄｓ Ｍｏｌ．Ｂｉｏｌ．，２０１５；１３２４：３９－５８；Ｒａｍａｋｒｉｓｈｎａ ｅｔ ａｌ．，“Ｇｅｎｅ ｄｉｓｒｕｐｔｉｏｎ ｂｙ ｃｅｌｌ－ｐｅｎｅｔｒａｔｉｎｇ ｐｅｐｔｉｄｅ－ｍｅｄｉａｔｅｄ ｄｅｌｉｖｅｒｙ ｏｆ Ｃａｓ９ ｐｒｏｔｅｉｎ ａｎｄ ｇｕｉｄｅ ＲＮＡ，”Ｇｅｎｏｍｅ Ｒｅｓ．，２０１４ Ｊｕｎ；２４（６）：１０２０－７；及び国際公開第２０１６２０５７６４号パンフレット（これらの各々は、本明細書において全体として参照により援用される）に記載されている。

【０２２７】

本明細書に記載されるＣＲＩＳＰＲシステムのための様々な送達方法はまた、例えば、米国特許第８７９５９６５号明細書、欧州特許第３００９５１１号明細書、国際公開第２０１６２０５７６４号パンフレット、及び国際公開第２０１７０７０６０５号パンフレット（これらの各々は、本明細書において全体として参照により援用される）にも記載されている。

【実施例】

【０２２８】

以下の例に本発明を更に記載するが、これらの例は、特許請求の範囲に記載される本発明の範囲を限定するものではない。

【０２２９】

実施例１－ＣＬＵＳＴ．０９１９７９　ＣＲＩＳＰＲ－Ｃａｓシステムの成分の同定

このタンパク質ファミリーは、上記の計算方法を使用して同定された。ＣＬＵＳＴ．０９１９７９システムは、腸、ウシ腸、ヒト腸、ヒツジ腸、陸生、糞便、及び哺乳動物消化系環境に限定されない環境から採取された無培養のメタゲノム配列に見られるＣＲＩＳＰＲシステムに関連するシングルエフェクターを含む（表５）。例示的なＣＬＵＳＴ．０９１９７９エフェクターには、以下の表５及び表６に示されるものが含まれる。図１Ａ～図１Ｌに示すように、配列番号１～４、１４、１５、１７～１９、２１～２５、２７～３３、３５～４９、５１～５６に記載のエフェクター配列をアラインメントして配列類似性の領域を同定した。バーグラフは配列類似性を示し、最も高いバーは最大配列類似性を有する残基を示す。非限定的な配列類似性の領域を表７に示す。配列類似性の領域は、本明細書に開示されるエフェクターがヌクレアーゼの代表的な保存Ｃ末端ＲｕｖＣドメインを有するファミリーであることを示す。

【０２３０】

【表１１】

表51.代表的なCLUST.091979 エフェクタータンパク質

供給源	エフェクターアクセッション	スペーサー数	エフェクターサイズ	配列番号
腸メタゲノム	AUXO013988882_8|P	4	775	1
ウシ腸メタゲノム	SRR094437_845781_4|M	11	786	2
腸メタゲノム	SRR1221442_316828_61|P	2	774	3
ウシ腸メタゲノム	SRR3181151_741875_3|M	8	756	4
ウシ腸メタゲノム	SRR5371369_1764679_7|P	7	746	5
ウシ腸メタゲノム	SRR5371371_1138852_2|M	3	733	6
ウシ腸メタゲノム	SRR5371379_2478682_1|M	9	744	7
ウシ腸メタゲノム	SRR5371385_201181_1|P	4	754	8
ウシ腸メタゲノム	SRR5371385_201181_1|M	4	746	9
ウシ腸メタゲノム	SRR5371401_1055766_58|M	15	745	10
ウシ腸メタゲノム	SRR5371439_988701_11|M	5	744	11
ウシ腸メタゲノム	SRR5371497_203858_6|M	5	745	12
ウシ腸メタゲノム	SRR5371501_2762794_1|M	2	712	13
陸生メタゲノム	SRR5678926_1309611_3|P	6	741	14
糞便メタゲノム	SRR6059713_382107_4|P	4	752	15
糞便メタゲノム	SRR6060192_2608084_13|P	16	766	16
ヒツジ腸メタゲノム	SRR7634052_1662339_24|M	8	784	17
腸メタゲノム	AUXO017332817_2|M	5	782	18
ヒト腸メタゲノム	OQVL01000914_15|P	6	735	19
哺乳動物消化系-アジアゾウ糞便-アジアゾウ(elephas maximus)	3300001598|EMG_10017415_6|P	2	774	20
哺乳動物消化系-ウシ及びヒツジルーメン	3300021254|Ga0223824_10022219_2|P	3	755	21
哺乳動物消化系-ウシ及びヒツジルーメン	3300021431|Ga0224423_10015012_2|P	11	789	22
哺乳動物消化系-糞便	3300012973|Ga0123351_1009859_3|P	6	766	23
哺乳動物消化系-	3300012979|Ga0123348_10005323_4|M	4	752	24

【０２３１】

【表１２】

糞便				
哺乳動物消化系-ルーメン-ウシ (bos taurus)	3300028797|Ga0265301_10000251_12|M	26	814	25
哺乳動物消化系-ルーメン-ウシ (bos taurus)	3300028797|Ga0265301_10000251_10|P	26	776	26
哺乳動物消化系-ルーメン-ウシ (bos taurus)	3300028797|Ga0265301_10009039_3|M	2	778	27
哺乳動物消化系-ルーメン-ウシ (bos taurus)	3300028887|Ga0265299_10000013_320|P	8	772	28
哺乳動物消化系-ルーメン-ウシ (bos taurus)	3300028887|Ga0265299_10000026_77|P	2	781	29
哺乳動物消化系-ルーメン-ウシ (bos taurus)	3300028887|Ga0265299_10000133_30|M	11	798	30
哺乳動物消化系-ルーメン-ウシ (bos taurus)	3300028887|Ga0265299_10011526_3|M	15	786	31
哺乳動物消化系-ルーメン-ウシ (bos taurus)	3300028887|Ga0265299_10012919_3|P	10	781	32
哺乳動物消化系-ルーメン-ウシ (bos taurus)	3300028914|Ga0265300_10009460_3|M	2	798	33
哺乳動物消化系-ルーメン-ウシ (bos taurus)	3300031853|Ga0326514_10013355_6|M	4	724	34
哺乳動物消化系-ルーメン-ウシ (bos taurus)	3300031993|Ga0310696_10000014_323|P	8	772	35
哺乳動物消化系-ルーメン-ウシ	3300031993|Ga0310696_10000226_76|P	2	781	36

【０２３２】

【表１３】

(bos taurus)				
哺乳動物消化系-ルーメン-ウシ (bos taurus)	3300031993|Ga0310696_10000447_27|M	11	798	37
哺乳動物消化系-ルーメン-ウシ (bos taurus)	3300031993|Ga0310696_10026614_2|M	2	781	38
哺乳動物消化系-ルーメン-ウシ (bos taurus)	3300031993|Ga0310696_10030100_3|M	14	786	39
哺乳動物消化系-ルーメン-ウシ (bos taurus)	3300031998|Ga0310786_10000003_467|M	9	798	40
哺乳動物消化系-ルーメン-ヒツジ (ovis aries)	AUXO013988882|Ga0247611_10000101_23|P	6	771	41
哺乳動物消化系-ルーメン-ヒツジ (ovis aries)	3300028805|Ga0247608_10000186_37|P	7	764	42
哺乳動物消化系-ルーメン-ヒツジ (ovis aries)	3300028805|Ga0247608_10000895_42|M	8	768	43
哺乳動物消化系-ルーメン-ヒツジ (ovis aries)	3300028805|Ga0247608_10006074_1|M	10	789	44
哺乳動物消化系-ルーメン-ヒツジ (ovis aries)	3300028833|Ga0247610_10000007_379|M	8	784	45
哺乳動物消化系-ルーメン-ヒツジ (ovis aries)	3300028833|Ga0247610_10004486_2|M	7	764	46
哺乳動物消化系-ルーメン-ヒツジ (ovis aries)	3300028888|Ga0247609_10000668_74|M	11	758	47
哺乳動物消化系-ルーメン-ヒツジ	3300028888|Ga0247609_10003329_9|M	8	785	48

【０２３３】

【表１４】

(ovis aries)				
哺乳動物消化系-ルーメン-ヒツジ (ovis aries)	3300028888|Ga0247609_10016480_8|M	2	805	49
哺乳動物消化系-ルーメン-ヒツジ (ovis aries)	3300031992|Ga0310694_10000010_351|M	8	784	50
哺乳動物消化系-ルーメン-ヒツジ (ovis aries)	3300031992|Ga0310694_10022272_2|M	7	764	51
哺乳動物消化系-ルーメン-ヒツジ (ovis aries)	3300031994|Ga0310691_10000084_157|M	8	768	52
哺乳動物消化系-ルーメン-ヒツジ (ovis aries)	3300031994|Ga0310691_10000270_20|M	7	764	53
哺乳動物消化系-ルーメン-ヒツジ (ovis aries)	3300032030|Ga0310697_10001273_44|P	2	805	54
哺乳動物消化系-ルーメン-ヒツジ (ovis aries)	3300032030|Ga0310697_10005481_13|P	8	785	55
ブタ腸メタゲノム	OBLI01003123_14|M	4	735	56

【０２３４】
【表１５】

表62.代表的なCLUST.091979エフェクタータンパク質のアミノ酸配列

>AUXO013988882_8|P
[腸メタゲノム]
MGNTTKKGNLTKTYLFKANLSEQDFKLWRSIVEEYQRYKEVLSKWVCDHLTTMKIGDILPYIDRYSKKIDNKTGEYPENTYYSL
CEEHKDEPLYKIFQFDSNCRNNALYEVIRKINCDLYTGNILNLGETYYRRNGFVKRVLANYATKISGMKPSVRKRKVTSDSTEE
EIRNQVVYEIFNNNIKNEKDFKGVLEYAESKCKTNEAYVERIRLLYDFYIKHTDEIKEYVEYICVEQLKEFCGVKVNRSKSSMN
INIQNFSITRVDGKCTYILHLPIGKKVYDIKLWGNRQVVLNVDGTPVDIIDIINRHGESIDIIFKNGDIYFSFVVSEDFKKDDF
EIGNVVGVDVNTKHMLIQTNIVDNGNVDGFFNIYKELVNDKEFSECVSKEDLELFKELSKYVSFCPIECQFLFTRYAEQKGILV
YEKLRLAEKILTSVLDRSFEKYNGIDCNIANYISNVRMLRSKCKSYFTLKMKYKELQHKYDNEMGYVDTFSDSCVEMDSRRKEN

【０２３５】

【表１６】

PFVQTNEAMELIGKMESVAQDIIGCRDNIITYAYNVFRRNGYDTVGLENLESSQFERFSSVRSPKSLLNYHHLKGKHIDFIDSD
ECSVKVNKDLYNFTLEDDGTISDITLSDKGKYRNDLSMFYNQIIKTIHFADIKDKFIQLGNNGNVQTVLVPSYFTSQMNSKTHK
IYVVNVKNERTGKTEQKLANKNMVRLGQERHINGLNADVNASMNIAYIVENKEMRNAMCTNPKSETGYSVPFLTSRIKKQNIMV
VELKKMGMVEVLNEKSTEI (配列番号1)

>SRR094437_845781_4|M
[ウシ腸メタゲノム]
MAQHKSNNEESAINKTFIFKAKCDKNDVISLWEPAAKEYCDYYNKVSKWIADNLITMKIGDLAQYITNQNSKYYTAVTNKKKKD
LPLYRIFQKGFSSQCADNALYCAIKSINPENYKGNSLGIGESDYRRFGYIQSVVSNFRTKMSSLKATVKWKKFDVNNVDDETLK
IQTIYDVDKYGIETAKEFKELIETLKTRVETPQLNDTIARLECLCDYYSKNEKAINNEIETMAIADLQKFGGCQRKSLNAFTIH
KQDSLMEKVGNTSFRLQLPFRKKTYVINLLGNRQVVNFVNGKRVDLIDIAENHGDLVTFNIKNGVLFVHLTSPIVFDKDVRDIR
NVVGIDVNIKHSMLATSIKDVGNVKGYINLYKELLNDDEFVSTCNESELALYRQMSENVNFGILETDSLFERIVNQSKGGCLKN
KLIRRELAMQKVFERITKTNKDQNIVDYVNYVKMMRAKCKASYILKEKYDEKQKEYYVKMGFTDESTESKETMDKRREEFPFVN
TDTAKELLVKQNNIRQDIIGCRDNIVTYAFNVFKNNEYDTLSVEYLDSSQFDKRRIATPKSLLKYHKFEGKTKDEVENMMKSEK
LSNAYYTFKYENDVVSDIDYSDEGNLRRSKLNFGNWIIKSIHFADIKDKFVQLSNNNKMNIVFCPSAFSSQMDSITHTLYYVEK
ITKNKKGKEKKKYVLANKKMVRTQQEKHINGLNADYNSACNLKYIALNDELRDKMTDRFKASKKIKTMYNIPAYNIKSNFKKNL
SAKTIQTFRELGHYRDGKINEDGMFVENLE (配列番号2)

>SRR1221442_316828_61|P
[腸メタゲノム]
MLNIKNNGESVDMNTIELAMKEYNRYYNICSDWICNNLMTPIGSLYQYIDDKCKNNAYAQNLIAEEWKDKPLYYMFYKGYNANN
CANAICCAIRSQVPEVNKAENILNLSYTYYFRNGVIKSVISNYASKMRILSDKQIKYCIVSENTPDKILIEQCILELKRRHEDL
KDWEENLKYLILKGNESAITRFTILKDFYSKNIERVKEEREIMAIAELKDFGGCRRKDDKLSMCIQSAGNSKDIKVSRVKTTHN
YTELVDDYTENFNIKFSALDFNVMGRRDVVKTKLNKTEDDSNTWGGTELLVDIINNHGCSLTFKLVDDKLYVDIPIDTEHINKT
TDFKKSVGIDVNLKHSLLNTDILDNGGINGYINIYKKLLADDAFMSACTKADLVNYIDIAKTVTFCPIEADFIISNVVEKYLHM
KDNTNKMEIAFSSVLMNIRKELEIKLLHSSKEESPLIRKQIIYINCIICLRNELKQYAIAKHRYYKKQQEYDTLCDTLHGVDYK
QIHPYAQSKEGAEQMKKMKTIENNLIANRNNIIEYAYTVFELNNFDLIALENITKDIMEDKKKRKSFPSINSLLKYHKVINCTE
DNINDNETYQKFAKYYNVSYENGKVTGATLSQEGNKVKLKDDFYDKLLKVLHFTSIKDYFTTLSNKRKIAVAHVPAYYTSQIDS
IDNKICMIKSTDKNGKSTYKIADKTIVRPTQEKHINGLNADYNAARNINFIVADEKWRKKFVRPTNTNKPLYNSPVFSPAVKSE
GGTIKNLQILSATKTIIL (配列番号3)

>SRR3181151_741875_3|M
[ウシ腸メタゲノム]
MTTKQVKSIVLKVKNTNECPITKDVINEYKKYYNICSEWIKDNLTSITIGDIASFLKEATNKDTIPTYINMGLSEEWKYKPIYH
LFTDDYHEKSANNLLYAYFKEKNLDCYNGNILNLSETYYRRNGYFKSVVGNYRTKIRTLNYKIKRKNVDENSTNEDIELQVMYE
IAKRKLNIKKDWENYISYIENVENINIKNIDRYNLLYKHFCENESTINCKMELLSVEQLKEFGGCVMKQHINSMTINIQDFKIE
NKENSLGFILNLPLNKKKYQIELWGNRQIKKGNKDNYKTLVDFINTYGQNIIFTIKNNKIYVVFSYECELKEKEINFDKIVGID
VNFKHALFVASERDKNPLQDNNQLKGYINLYKYLLEHNEFTSLLTKEELDIYKEIAKGVTFCPLEYNLLFTRIENKGGKSNDKE
QVLSKLLYSLQIKLKNENKIQEYIYVSCVNKLRAKYVSYFILKEKYYEKQKEYDIEMGFTDDSTESKESMDKRRLEFPFRNTQI
ANGFLEKLSNVQQDINGCLKNIINYAYKVFEQNGFGVIALENLENSNFEKTQVLPTIKSLLEYHKLENQNINNINASDKVKEYI
EKEYYELTTNENNEIVDAKYTKKGIIKVKKANFFNLMMKSLHFASNKDEFILLSNNGKTQIALVPSEYTSQMDSIEHCLYVDKN
GKKVDKKKVRQKQETHINGLNADFNAANNIKYIIENENLRKLFCGKLKVSGYNTPILDATKKGQFNILAELKKQNKIKIFEIEK
(配列番号4)

>SRR5371369_1764679_7|P
[ウシ腸メタゲノム]
MASHKKTESNQIIKTFPFKLKNANGLSLDVLNDAITEYQNYYNICSDWIKDHLTMKISELYKYIPDEKKNSGYALTLISDEWKD
KPMYMMFKKGYPANNRDNAIYETLNTCNTEHYTGNILNFPDTYYRRFGYVASTISNYVTKISKMSTGSRSKNISNDSDVDTIME
QVIYEMEHNGWTSVKDWENQMEYLESKTDSNPNFVYRMTTLYEFYKSHIDEVNSKMETMSIDLLIKFGGCRRKDSKKSMYIMGG
SNTPFDITQIGDNSLNIKFSKNLNVDVFGRYDVIKDNTLLVDIINGHGASFVLKIINDEIYIDINVSVPFDKKIATTNKVVGID
VNIKHMLLATNILDDGNVKGYVNIYKEVINDSDFKKVCNSTVMKYFTDFSKFVTFCPLEFDFLFSRVCNQKGIYNDNSVMEKSF
SDVLNKLKWNFIETGDNTKRIYIENVMKLRTQMKAYAIVKNAYYKQQSEYDFGKSEEFIQEHPFSNTDKGIEILHKLDNISKKI
LGCRNNIIQYSYNLFEINGYDMISLEKLTSSQFKKKSFPTVNSLLKYHKILGCTQEEMEKKDIYSVIKKGYYDIIFDNDVVTDA
KLSTKGELSKFKDDFFNLMIKSIHFADIKDYFITLSNNGTAGVSLVPSFFTSQMDSIDHKIYFVQDNKSGKLKLANKHKVRSSQ
EKHINGLNADYNAARNIAYIMENTECRNMFMKQSRTDKSLYNKPSYETFIKTQGSAVAKLKKEGFMKILDEASV
(配列番号5)

>SRR5371371_1138852_2|M
[ウシ腸メタゲノム]
MAHKKNIGAEIVKTYSFKVKNTNGITMEKLMNAIDEYQSYYNLCSDWICKNLTTMTIGDLDRYIPEKAKDNIYATVLLDEVWKN
QPLYKIFGKKYSSNNRNNALYCALSSVIDMTKENVLGFSKTHYIRNGYILNVISNYASKLSKLNTGVKSRAIKETSDEATIIEQ
VIYEMEHNKWESIEDWKNQIEYLNSKTDYNPTYMERMKTLSAYYSTHKSEVDAKMQEMAVENLVKFGGCRRNNSKKSMFIMGSN

【０２３６】

【表１７】

TTNYTISYIGDNCFNINFANILNFDVYGRRDVVKNGEVLVDIMANHGDSIVLKIVNGELYADVPCSVTLNKVESNFDKVVGIDV
NMKHMLLSTSVTDNGSSDFVNIYKEMSNNAEFMALCPEKDRKYYKDISQYVTFAPLELDLLFSRISKQGEVKMEKAYSEILESL
KWKFFANGDNKNRIYVESIQKIRQQIKALCVIKNAYYEQQSAYDIDKTQEYIETHPFSLTEKGMSIKSKMDKICQTIIGCRNNI
IDLAYSFFERNGYSIIGLEKLTSSQFKNTKSMPTCKSLLNLHKVLGHTLSELETLPINDIVKYYTFTTDNEGRITDASLSEKGK
IRKMKDRFLNQAIKAIHFADVKDYFATLSNNGQTGIFFVPSQFTSQMDSNTHNLYFEVDKNGGLKMASKDKTRPKQEYHRNGLP
ADYNAARNIAYIGLDETMRNTFLKKVNSNKSLYNQPIYDTGIKKTAGVFSRMKKLKRYEII (配列番号6)

>SRR5371379_2478682_1|M
[ウシ腸メタゲノム]
MIKSIKLKVKGDCPITKDVINEYKEYYNRCSDWIKNNLTSITIGEIGKFLQDVTGKTTGYIEVALSDKWKDKPMYYLFTDQYDT
NHANNLLYSFIQENNLDGYDGNSLNISGTYYRKQGYFKLVSSNYRTKIRTLNCKIKRKKVDVDSTSEDIESQVMYEIINRSLNK
KSDWDSFISYIENVENPNIDSINRYTLLRDYFCDNEDVIKNKIELLSIEQLKDFGGCIMKQHINTMSLNIQHFKIEEKENSLGF
ILYLPLNKKQYQIELWGHRQIKKGSKESCETLVDFINTYGENIVFTINNDELYVVFSYESEFGKEETNFEKSVGLDINFKHALF
VTSELDNDQFDGYINLYKYILSHSEFTNLLTEDERKDYEELSKVVTFCPFENQLLFARYDKMSKFCKKEQVLSKLLYSLQKKLK
NENRTKEYIYVSCVNKLRAKYISYFILREKYDEKNKEYDIEMGFVDDSTESKESMDKRRFENPFRNTLVANELLAKMSKVQQDI
NGCMSNIINYVYKVFEQNGYNIIALENLENSNFEKRQVLPTIKSLLKYHKLENQNINDIKASDKIKEYIENGYYSFTTNENNEI
VDAKYTAKGDIKVKNAKFFNLMMKILHFASIKDEFVLLSNNGKSQIALVPPEYTSQMDSIDHCIYMTENDKGKIVKVDKRKVRT
KQERHINGLNADFNAANNIKYIVSNEKWRNVFCTPKKAKYNTPALDATKKGQFRILDDMKKLNATKLLEIEK (配列番号7)

>SRR5371385_201181_1|P
[ウシ腸メタゲノム]
MYQLNQYIMASHKKTESNQIIKTFSFKIKNANGLSLDVLNDAITEYQNYYNICSDWIKDHLTMKISELYKYIPDEKKNSGYALT
LISDEWKDKPMYMMFKKGYPANNRDNAIYETLNTCNTEHYTGNILNFSDTYYRRFGYVASAISNYVTKISKMSTGSRYKNISND
SDVDTIMEQVIYEMEHNGWTSVKDWENQMEYLESKTDSNPNFVYRMTTLYEFYKSHIDEVNSKMETMSIDSLIKFGGCRRKDSK
KSMYIMGGSNTPFDITQIGGNSLNIKFSKNLNVDVFGRYDVIKDNTLLVDIINGHGASFVLKIINDEIYIDINVSVPFDKKIAT
TNKVVGIDVNIKHMLLATNILDDGNVKGYVNIYKEVINDSDFKKVCNSTVMKYFTDFSKFVTFCPLEFDFLFSRVCNQKGIYND
NSAMEKSFSDVLNKLKWNFIETGDNTKRIYIENVMKLRSQMKAYAIVKNAYYKQQSEYDFGKSEEFIQEHPFSNTDKGIEILHK
LDNISKKILGCRNNIIQYSYNLFEINGYDMISLEKLTSSQFKKKPFPTVNSLLKYHKILGCTQEEMEKKDIYSVIKKGYYDIIF
DNGVVIDAKLSAKGELSKFKDDFFNLMIKSIHFADIKDYFITLSNNGTAGVSLVPSYFTSQMDSIDHKIYFVQDNKSGKLKLAN
KHKVRSSQEKHINGLNADYNAARNIAYIMENTECRNMFMKQSRTDKSLYNKPSYETFIKTQGSAVSKLKKDGFVKILDEASV
(配列番号8)

>SRR5371385_201181_1|M
[ウシ腸メタゲノム]
MASHKKTESNQIIKTFSFKIKNANGLSLDVLNDAITEYQNYYNICSDWIKDHLTMKISELYKYIPDEKKNSGYALTLISDEWKD
KPMYMMFKKGYPANNRDNAIYETLNTCNTEHYTGNILNFSDTYYRRFGYVASAISNYVTKISKMSTGSRYKNISNDSDVDTIME
QVIYEMEHNGWTSVKDWENQMEYLESKTDSNPNFVYRMTTLYEFYKSHIDEVNSKMETMSIDSLIKFGGCRRKDSKKSMYIMGG
SNTPFDITQIGGNSLNIKFSKNLNVDVFGRYDVIKDNTLLVDIINGHGASFVLKIINDEIYIDINVSVPFDKKIATTNKVVGID
VNIKHMLLATNILDDGNVKGYVNIYKEVINDSDFKKVCNSTVMKYFTDFSKFVTFCPLEFDFLFSRVCNQKGIYNDNSAMEKSF
SDVLNKLKWNFIETGDNTKRIYIENVMKLRSQMKAYAIVKNAYYKQQSEYDFGKSEEFIQEHPFSNTDKGIEILHKLDNISKKI
LGCRNNIIQYSYNLFEINGYDMISLEKLTSSQFKKKPFPTVNSLLKYHKILGCTQEEMEKKDIYSVIKKGYYDIIFDNGVVIDA
KLSAKGELSKFKDDFFNLMIKSIHFADIKDYFITLSNNGTAGVSLVPSYFTSQMDSIDHKIYFVQDNKSGKLKLANKHKVRSSQ
EKHINGLNADYNAARNIAYIMENTECRNMFMKQSRTDKSLYNKPSYETFIKTQGSAVSKLKKDGFVKILDEASV (配列番号9)

>SRR5371401_1055766_58|M
[ウシ腸メタゲノム]
MIKSIQLKVKGECPITKDVINEYKEYYNNCSDWIKNNLTSITIGEMAKFLQSLSDKEVAYISMGLSDEWKDKPLYHLFTKKYHT
KNADNLLYYYIKEKNLDGYKGNTLNISNTSFRQFGYFKLVVSNYRTKIRTLNCKIKRKKIDADSTSEDIEMQVMYEIIKYSLNK
KSDWDNFISYIENVENPNIDNINRYKLLRECFCENENMIKNKLELLSVEQLKKFGGCIMKPHINSMTINIQDFKIEEKENSLGF
ILHLPLNKKQYQIELLGNRQIKKGTKEIHETLVDITNTHGENIVFTIKNDNLYIVFSYESEFEKEEVNFAKTVGLDVNFKHAFF
VTSEKDNCHLDGYINLYKYLLEHDEFTNLLTEDERKDYEELSKVVTFCPFENQLLFARYNKMSKFCKKEQVLSKLLYALQKKLK
DENRTKEYIYVSCVNKLRAKYVSYFILKEKYYEKQKEYDIEMGFVDDSTESKESMDKRRTEYPFRNTPVANELLSKLNNVQQDI
NGCLKNIINYIYKIFEQNGYKVVALENLENSNFEKKQVLPTIKSLLKYHKLENQNVNDIKASDKVKEYIENGYYELMTNENNEI
VDAKYTEKGAMKVKNANFFNLMMKSLHFASVKDEFVLLSNNGKTQIALVPSEFTSQMDSTDHCLYMKKNDKGKLVKADKKEVRT
KQERHINGLNADFNAANNIKYIVENEVWRGIFCTRPKKTEYNVPSLDTTKKGPSAILNMLKKIEAIKVLETEK (配列番号10)

>SRR5371439_988701_11|M
[ウシ腸メタゲノム]
MIKSIVFKVKGDCPITKDVIKEYKEYYNRCSEWIKNNLTSITIGEIGKFLQDTMGKTHGYIKVALSDEWKDKPMYYLFTEKYDT

【０２３７】

【表１８】

KHANNLLYYFIQENNLDRYEGNSLNIPSYYYKREGYFKLVTSNYRTKIRTLNCKIKRKKIDVDSTCVDIENQVIYEIIKKGLNK
KSDWDNYISYIENIEMPNIDSINRYKLLRDYFCENENVIKNKIELLSIEQLKNFGGCIMKQHINTMILNIKRLKIEEKENSLGF
ILHLPLNKKQYQIELWGNRQIKKGTKESNETLVDFINTYGEDVVFTIKKNELYAKFSYECEFEKEETNFEKSVGLDINFKHALF
VTSELDDDQFYGYINLYKYILSHSEFTNLLTEDEKKDYEDLSNAITFCPFENQLLFTRYDKKSKLYKKEQVLSKILYSLQKKLK
DENRKQEYIYVSCVNKLRAKYVSYFILKEKYNEKQKEYDIEMGFVDDSTESKESMDKRRYEYPFRNTPVANELLEKMNNVQQDI
SGCLKNIINYAYKVFEQNGYNIVALENLENSNFEKRNVLPTIKSLLKYHKLENQNITDIKASDKIKEYIENGYYELITNENNEI
IDAKYTENGDIKVKNARFFNLMMKSLHFASIKDEFVLLSNNGKSQIALVPSEYTSQMDSTDHCIYMTENDKGKLVKVDKRKVRT
KQERHINGLNADFNAANNIKYIVENEKWRKVFCAPQKAKYNTPTLDATKKGQFRILEDLKKLKATKLLEIGK (配列番号11)

>SRR5371497_203858_6|M
[ウシ腸メタゲノム]
MIKSIQLKVKGECPITKDVINEYKEYYNNCSDWIKNNLTSITIGEMAKFLQSLSDKEVAYISMGLSDEWKDKPLYHLFTKKYHT
KNADNLLYYYIKEKNLDGYKGNTLNISNTSFRQFGYFKLVVSNYRTKIRTLNCKIKRKKIDADSTSEDIEMQVMYEIIKYSLNK
KSDWDNFISYIENVENPNIDNINRYKLLRECFCENENMIKNKLELLSVEQLKKFGGCIMKPHINSMTINIQDFKIEEKENSLGF
ILHLPLNKKQYQIELLGNRQIKKGTKESHETLVDITNTHGENIVFTIKNDNLYIVFSYESEFEKEEVNFAKTVGLDVNFKHAFF
VTSEKDNCHLDGYINLYKYLLEHDEFTNLLTEDERKDYEELSKVVTFCPFENQLLFARYNKMSKFCKKEQVLSKLLYALQKKLK
DENRTKEYIYVSCVNKLRAKYVSYFILKEKYYEKQKEYDIEMGFVDDSTESKESMDKRRTEYPFRNTPVANELLSKLNNVQQDI
NGCLKNIINYIYKIFEQNGYKVVALENLENSNFEKKQVLPTIKSLLKYHKLENQNVNDIKASDKVKEYIENGYYELMTNENNEI
VDAKYTEKGAMKVKNANFFNLMMKSLHFASVKDEFVLLSNNGKTQIALVPSEFTSQMDSTDHCLYMKKNDKGKLVKADKKEVRT
KQERHINGLNADFNAANNIKYIVENEVWRGIFCTRPKKTEYNVPSLDTTKKGPSAILNMLKKIEAVKILETEK (配列番号12)

>SRR5371501_2762794_1|M
[ウシ腸メタゲノム]
MKNNLTTVTIGEMAKFLQETTGKNVTYITMGLSEEWKDKPLYHLFYGKYHTKNADNLLYYFIKAKKLDEYDGNMLNLGDTYYRQ
FGYFKLVVSNYRTKIRTLNLNVKRKRVDVDSTSEDIESQVMYEIVKRNLNTISDWENYISYIEDVETPNIDNINRYKFLQNYFC
ENEEDIKNKIEFLSIEQLKDFGGCIMKPHINSMTINIQDFKIEEIENSLGFVLQLPLNKKYHQIELYGNRQVKKGTKENYKTLV
DIINTHGENIVFTIENNELYVVFSYEYELKKKDINFEKMAGIDVNFKHALFVTSETDNNQLNHYINLYKHILEHNEFTTLLTDS
ERKDYEEIAKTVTFCPFEYQLLFTRFDKNSNANVKEQALSKILYDLQKKLKSQNKIKEYIYVSCVNKLRAKYVSYFILKEKYYE
KQKEYDIQMGFVDDSTESKSSMVKRRVEYPFRNTPVANALLAIVNNVQQDINGCLKNIINYAYKVFELNDYNVVALENLENANF
EKKQVIPTIKSLLKYHKLEMQNINDIKANDTIKKYIENEYYQLITNENNEIVNAIYTPKGITKLKYANFFNLLMKSLHFASIKD
EFILLSNNGNTNIALVPHEYTSQMDSIDHCIYMVQNDKGNLVKAHKTKVRTKQEKHINGLNADFNAANNIKYIVENEKWRNIFC
KIPKKIEYNTPVLDVTKKGQSNIIKTLKNLNATKILEIKK (配列番号13)

>SRR5678926_1309611_3|P
[陸生メタゲノム]
MKKSIKFKVKGNCPITKDVINEYKEYYNKCSDWIKNNLTSITIGEMAKFLQETLGKDVAYISMGLSDEWKDKPLYHLFTKKYHT
NNADNLLYYYIKEKNLDGYKGNTLNIGNTFFRQFGYFKLVVSNYRTKIRTLNCEIKRKKIDADSTSEDIEMQTMYEIIKHNLNK
KTDWDEFISYIENVENPNIDNINRYKLLRKCFCENENMIKNKLELLSIEQLKNFGGCIMKQHINSMTLIIQHFKIEEKENSLGF
ILNLPLNKKQYQIELWGNRQVNKGTKERDAFLNTYGENIVFIINNDELYVVFSYEYELEKEEANFVKTVGLDVNFKHAFFVTSE
KDNCHLDGYINLYKYLLEHDEFTNLLTNDEKKDYEELSKVVTFCPFENQLLFARYNKMSKFCKKEQVLSKLLYALQKQLKDENR
TKEYIYVSCVNKLRAKYVSYFILKEKYYEKQKEYDIEMGFVDDSTESKESMDKRRTEFPFRNTPVANELLSKLNNVQQDINGCL
KNIINYIYKIFEQNGYKIVALENLENSNFEKKQVLPTIKSLLKYHKLENQNVNDIKASDKVKEYIENGYYELITNENNEIVDAK
YTEKGAMKVKNANFFNLMMKSLHFASVKDEFVLLSNNGKTQIALVPSEFTSQMDSTDHCLYMKKNDKGKLVKADKKEVRTKQEK
HINGLNADFNAANNIKYIVENEVWREIFCTRPKKAEYNVPSLDTTKKGPSAILHMLKKIEAIKILETEK (配列番号14)

>SRR6059713_382107_4|P
[糞便メタゲノム]
MAKSIMKKSIKFKVKGNSPINEDIINEYKGYYNTCSNWINNNLTSITIGEMGKFLKDVMRKTTGYIDVALSDEWKDKPMYYLFT
KKYNPKHANNLLYYFIKEKKLDKFNGNILNVPEYYYRKEGYFKLVAGNYRTKINTLNFKIKSKKVDANSLSEDIEMQTIYEIVK
RGLNKKSDWDSYISYIECVQNPNIDNINRYKLLRDYFCENEDVIKNKIEILSIEQIKEFGGCIMKPHINSMTFGIQKFKIEEIE
NSLGFTFNLPLNKNNYKIELWGHRQLKKGNKESNVNVSLDDFINTYGQNVVFTIKRKKLYIVFSYDYEFERGECNFEKSVGLDV
NFKHSLFVTSEIDNNQFDGYINLYKYILSNNEFTSLLTDSERKDYEDLANIVTFCPFEYQLLFSRYDKLSKISEKEKVLSKILY
SLQKKLKNEKRTKEYIYVSCVNKLRAKYVSYFKLKQKYNEKQKEYDIEMGFVDDSTESKESMDKRRFENPFINTPVAKELLEKM
NNVKQDINGCKKNIVVYAYKVLEQNGYNIIALENLENSNFEKIRVLPKIKSLLEYHKFENKNINDIKNSDKYKEFIEPGYFELI
TNENNEIIDAKYTQKGDIKIKNADFINIMIKALNFASIKDEFILLSHNGKSQIALVPAEYTSQMDSIDHCIYMTKNDKGKLVKV
DKRKVRTKQERHINGLNADFNAACNIKYIVTNEDWRKVFCIKPKKEDYNTPLLDATKNGQFRILDKLKKLNATKLLEMEK
(配列番号15)

>SRR6060192_2608084_13|P

【０２３８】

【表１９】

[糞便メタゲノム]
MANKKFKLTKNEVVKSFVLKVANQKKCAITNETLQEYKNYYNKVSQWINNNLTKMTIGDLIQYAPTVSKKGKKQPDGTMVYDTP
LYVTYAMSDEWKNKPLYYIFKKEYNTNNANNLLYEAIRNLNVDEYDGNQLNFNSTYYRTQGYVNRVFSNYRTKINTLDIKIKKS
KVDENSDVETLELQTMYEINKLNLKTNKDWEERLQYLTMQENPNQNTIDRTKILFNYFINNNDTIFQKMEELSIKQLTEFGGCK
MKDNTTSMTINIQDFKIKRKENSIGYIMTIPFNKKNVDVELYGHKQTIKGHKNSYTEIVDIVNKHGNTITFKIKNNQLFAIITS
DTEVTKPEPQYEKIVGVDVNIKHTLMVTSEKDNGKLKGYINLYKEVLKNDEFKKLLNKTELDNFKSLSQIVTFCPIEYDFLFSR
IFDDENTKKELAFSNVLYDIQKQLKNTNNILQYNYIACVNKLRAKYKAYFVLKMSYMKQQKIYDTNMGFFDISTESKETMDQRR
SLYPFINTEIAQNIITKMNNVQQDINGCLKNIFKYTYTVFENNNYDTIVLENLENANFEKHNPLPNITSLLKYHKVQGLTIQEA
EQHEKVGNLIQNDNYIFQLNEDNKIINADYSQKAYYKVCKALFFNQAIKTLHFASVKDEMIKLSNNNKVCVAIIPPEYTSQIDS
NTHKLYFINKDGKLLKADKKTVRKTQEKHINGLNADFNAASNIKYIVQNETWRNLFTNKTNNTYGLPILTPSKKGQSNIITQLM
KINATQELVV (配列番号16)

>SRR7634052_1662339_24|M
[ヒツジ腸メタゲノム]
MYNSKKKGEGDIQKSFKFKVKTDKETVELFRKAAVEYSEYYKRLTTFLCERLTDMTWGEVASFIPEKYRKNEYYKYLIKEENKD
LPLYKMFTKAASSMFIDHSIERYVEALNPEGNTGNILGFCKSSYVRGGYLKNVVSNIRTKFATLKTGIKYKKFNPAEDDEETIL
GQTVFEMEKRGLEFKCDFEKTIKYLNEKGKTQEAERLQCLMEYFSTNTDKINEYRESLVLDDIRKFGGCNRSKSNSFSVTLEKA
DIKEDGLTGYTMKVSKKLKEIHLLGHRRVVEVVNGRRVNLVDICGDKSGDSKVFVVDGDNLYVCISAPVKFSKNGMEAKKYIGV
DMNMKHSIISVSDNASDMKGFLNIYKELLKDEGFRKTLNATELEKYEKLAEGVNIGIIEYDGLYERIVKQKKENSVDGLKVQAE
KKLIEREAAIERVLDKLRKGTSDTDTENYINYNKILRAKIKSAYILKDKYYEMLGKYDSERAGSGDLSEENKIKYKDEFNETEK
GKEILGKLNNVYKDIIGCRDNIVTYAVNLFIRNGYDTVALEYLESSQMKARRIPSTGGLLKGHKLEGKPEGEVTAYLKANKIPK
SYYSFEYDGNGMLTDVKYSDMGEKARGRNRFKNLVPKFLRWASIKDKFVQLSNYKDIQMVYVPSPYTSQTDSRTHSLYYIETVK
VDEKTGKEKKEHIVAPKESVRTEQESFVNGMNADTNSANNIKYIFENETLRDKFLKRTKDGTEMYNRPAFDLKECYKKNSNVSV
FNTLKKTLGAIYGKLDENGNFIENECNK (配列番号17)

>AUXO017332817_2|M
[腸メタゲノム]
MAGHSKIKENHIMKAFLMKVKETRKKQWQSNFIRSEIAKFTNYYNGLSKFIADRLLDDMVTTLAPLIEEKKRNSEYYKYLTNGD
WDGKPLYFIFKEGFNSTNADNILANSLVRVYCEQNYTGNGFGLSYSYYVVIGFAKEVIANYRSSFQKPKVKIKKKKLSENPTED
ELIEQCIYTIYYEFNEKKDIQKWKDEIKFLKERGESKETRLKRIQTLFEFYKDKSHKELVDERVANLVVDNIKEFGGCKRDIDC
PSMGIQIQHNFDISINEKRNGYTICFGPNKKNLTKLEVFGNRMVLLNGEEIVDLPNTHGEKLTLIDRGNAIYAAITAQVPFEKH
MPDGNKTVGIDLNLKHSVFATSIVDNGKLAGYISIYKELLKDDEFVKYCPKDLLRFMKDASKYVFFAPIEIELLRSRVIYNKGY
ACVENYENVYKAEVAFVNVIKRLQSQCEANGDAQGALYMSYLSKMRAQLKNYINLKLAYYDHQSAYDLKMGFTDISTESKETMD
ERRKLFPFNKEKEAQEILAKMKNISNVIIACRNNIAVYMYKMFERNGYDFIGLEKLESSQMKKRQSRSFPTVKSLLNYHKLAGM
TMDEIKKQEVSSNIKKGFYDLEFDADGKLYGAKYSNKGNVHFIEDEFYISGLKAIHFADMKDYFVRLSNNGKVSVALVPPSFTS
QMDSVEHKFFMKKNANGKLIVADKKDVRSCQEKHKINGLNADYNAACNIGFIVEDDYMRESLLGSPTGGTYDTAYFDTKIQGSK
GVYDKIKENGETYIAVLSDDVITAEV (配列番号18)

>OQVL01000914_15|P
[ヒト腸メタゲノム]
MAHKKNVGAEIVKTYSFKVKNTNGITMEKLMNAIDEFQSYYNLCSDWICKNLTTMTIGDLDQYIPEKAKGNTYATVLLDEAWKN
QPLYKIFGKKYSSNNRNNALYCALSSVIDMTKENVLGFSKTHYIRNDYILNVISNYASKLSKLNTGVKSRAIKETSDEATIIEQ
VIYEMEHNKWESIEDWKNQIEYLNSKTDYNPTYMERMKTLSAYYSTHKSEVDAKMQEMAVENLVKFGGCRRNNSKKSMFIMGSN
TTNYTISYIGGNSFNINFANILNFDVYGRRDVVKNGEVLVDIMANHGDSIVLKIVNGELYADVPCSVTLNKVESNFDKVVGIDV
NMKHMLLSTSITDNGSSDFLNIYKEMSNNAEFMALCPEEDRKYYKDISKYVTFAPLELDLLFSRISKQGKVKMEKVYSEILEAL
KWKFFANGDNKNRIYVESIQKIRQQIKALCVIKNAYYEQQSAYDIDKTQEYIETHPFSLTEKGMSIKSKMDKICQTIIGCRNNI
IDYAYSFFERNGYSIIGLEKLTSSQFEKTKSMPTCKSLLNFHKVLGHTLSELETLPINDVVKKGYYTFTTDNEGKITDASLSEK
GKVRKMKDDFFNQAIKAIHFADVKDYFATLSNNGQTGIFFVPSQFTSQMDSNTHNLYFENAKNGGLKLAPKYKVRQTQEYHLNG
LPADYNAARNIAYIGLDETMRNTFLKKANSNKSLYNQPIYDTGIKKTAGVFSRMKKLKRYEII (配列番号19)

>3300001598|EMG_10017415_6|P
[哺乳動物消化系-アジアゾウ糞便-アジアゾウ(elephas maximus)]
MLNIKNNGESVDMNTIELAMKEYNRYYNICSDWICNNLMTPIGSLYQYIDDKCKNNAYAQNLIAEEWKDKPLYYMFYKGYNANN
CANAICCAIRSQVPEVNKAENILNLSYTYYFRNGVIKSVISNYASKMRILSDKQIKYCIVSENTPDKILIEQCILELKRRHEDL
KDWEENLKYLILKGNESAITRFTILKDFYSKNIERVKEEREIMAIAELKDFGGCRRKDDKLSMCIQSAGNSKDIKVSRVKTTHN
YTELVDDYTENFNIKFSALDFNVMGRRDVVKTKLNKTEDDSNTWGGTELLVDIINNHGCSLTFKLVDDKLYVDIPIDTEHINKT
TDFKKSVGIDVNLKHSLLNTDILDNGGINGYINIYKKLLADDAFMSACTKADLVNYIDIAKTVTFCPIEADFIISNVVEKYLHM
KDNTNKMEIAFSSVLMNIRKELEIKLLHSSKEESPLIRKQIIYINCIICLRNELKQYAIAKHRYYKKQQEYDTLCDTLHGVDYK
QIHPYAQSKEGAEQMKKMKTIENNLIANRNNIIEYAYTVFELNNFDLIALENITKDIMEDKKKRKSFPSINSLLKYHKVINCTE
DNINDNETYQKFAKYYNVSYENGKVTGATLSQEGNKVKLKDDFYDKLLKVLHFTSIKDYFTTLSNKRKIAVAHVPAYYTSQIDS
IDNKICMIKSTDKNGKSTYKIADKTIVRPTQEKHINGLNADYNAARNINFIVADEKWRKKFVRPTNTNKPLYNSPVFSPAVKSE
GGTIKNLQILSATKTIIL (配列番号20)

【０２３９】

【表２０】

>3300021254|Ga0223824_10022219_2|P
[哺乳動物消化系-ウシ及びヒツジルーメン]
MAHVRTKNEGNMAKTYSFKVRETNLKKDVMIEYNEYYNRLSDWICGNLTKMTIGELAELVPEKKRNTSYYLAATDEKWINEPMY
KLFTDEYTKKSSFTDPLVANSNNCDNLILTATDVLNPEGYEGNLLSLCKSTYRTFGYAKQIISNMKTKIGALKPNVKRRVLGEN
PTYDEKMIQVLYEMYNNGIADVTGFNDRIKYLKKQETPNEKLISRMKMLRDFFKENRNDIMDKCRIMAVEQLVSFGGCKRNING
ASMTLRNQCISVKRKDGCQGYVVAIPVGTKNSIVFDLYGRRDVIKDGVELVDVCGKHTDTITIKSVNGELFLDMPVAINFEKKS
GKCTKTVGIDVNTKHMLIQTSVKDNGKFDYYVNLYKIFAEDEELNKILGDDEVMVNIKKNAENLSFLPLEMDLLYSRILDGPQK
YKLAEDRITELLKQWGINFDAGCMSQERIYVQCVRKLRGNLKRLLYLQNKYYEAQQEYDKKMGFDDKSTDSKETMDKRRWESPF
RNTEEGTKLYDEINTYQNRIIGIRNSIIDYAYLVLEYNGYDNLSLEYLTSSQFKVNKTFPTTNSLLKYHKLQGKTKTEAEKCDA
YISHKSKYKLSLKDGVIDSIDYSAEGLKQIKKDRSRNIIIKAIHFADVKDRFVLSSNNGNASVTFVPSYHTSQIDSTDHKMFVT
NKGKIVDKRKVRQIQETHVNGLNSDFNAARNIQYISENEEWRNALCKPTENMYNEPIYVPLVKSQNGMFKAIKKLGATKIWQE
(配列番号21)

>3300021431|Ga0224423_10015012_2|P
[哺乳動物消化系-ウシ及びヒツジルーメン]
MAHRNKNLAENCINKTFSFKVKAEKEEINSKWIPAIKEYTAYYNRISDWICDRLTNTTVGELIGIIGYKTDKKGNALAYIKDGS
SEKYRNLPLYCMFKKNFPATTADNIMYQVIEKLGVDKYNGNSLGLSGTYYRRIGYIANVIGNYRTKVRGMKASVKYRNFDPNDV
TEDVLENQTIFEINKNGFECKGDFEKHIEYLKNRELTDRLNKLILRMECLYNYYVEHEDAVKAKMENYAIESFKTFGGCHRNSN
RSMSIQFTNNSPLEIKKVGKTSFDLYMPINGEVACLQLMGNKQAVCVGENGERCDLVDIVNSHSKTITIKIINGEMYVDIPCVV
NFEKKDEDTIKSVGVDVNIKHEILATSVIDNGQLNGYFNIYKELINNKEFVDTFNGDIKAFEAFKDNAAYVTFGLLEPDLLFTR
FYERSGFEKDDRHIKLRERERILTGILKRIGQEHSDVDVRNYVRFVNMLRSKYESYFVLKNKYYEKMQEFDSTQNYVDVSTASK
ETMDKRRFDNPFRNTEVANELLGKIDNVLGDIKGCMANIITYAFKVLQKNGYNTIGLEYLDSSQFENMRTLTPTSILKYHKMEG
KSVDAVESWIKENKIPSNRYDFIYEDNHLTDVLLNSNGIAYQKKNLFMNLVIKAISFADIKNKFVQLSNNTNVSILFAPAAFTS
QMDSNRHVIYTVKNNKGKLALVDKKRVRPNQEKHINGLHSGYNAACNVKFICDNEFFRNTMTISNKGKNLYSQPTYDIKEAYKK
NAGCKVINDFIKNGNAVICCIENNKLIETNGRQ (配列番号22)

>3300012973|Ga0123351_1009859_3|P
[哺乳動物消化系-糞便]
MANKKFKLTKNEVVKSFVLKVANQKKCAITNETLQEYKNYYNKVSQWINNNLTKMTIGDLIQYAPTVSKKGKKQPDGTMVYDTP
LYVTYAMSDEWKNKPLYYIFKKEYNTNNANNLLYEAIRNLNVDEYDGNQLNFNSTYYRTQGYVNRVFSNYRTKINTLDIKIKKS
KVDENSDVETLEPQTMYEINKLNLKTNKDWEERLQYLTMQENPNQNTIDRTKILFNYFINNNDTIFQKMEELSIKQLTEFGGCK
MKDNTTSMTINIQDFKIKRKENSIGYIMTIPFNKKNVDVELYGHKQTIKGHKNSYTEIVDIVNKHGNTITFKIKNNQLFAIITS
DTEVTKPEPQYEKIVGVDVNIKHTLMVTSEKDNGKLKGYINLYKEVLKNDEFKKLLNKTELDNFKSLSQIVTFCPIEYDFLFSR
IFDDENTKKELAFSNVLYDIQKQLKNTNNILQYNYIACVNKLRAKYKAYFVLKMSYMKQQKIYDTNMGFFDISTESKETMDQRR
SLYPFINTEIAQNIITKMNNVQQDINGCLKNIFKYTYTVFENNNYDTIVLENLENANFEKHNPLPNITSLLKYHKVQGLTIQEA
EQHEKVGNLIQNDNYIFQLNEDNKIINADYSQKAYYKVCKALFFNQAIKTLHFASVKDEMIKLSNNNKVCVAIIPPEYTSQIDS
NTHKLYFINKDGKLLKADKKTVRKTQEKHINGLNADFNAASNIKYIVQNETWRNLFTNKTNNTYGLPILTPSKKGQSNIITQLM
KINATQELVV (配列番号23)

>3300012979|Ga0123348_10005323_4|M
[哺乳動物消化系-糞便]
MAKSIMKKSIKFKVKGNSPINEDIINEYKGYYNTCSNWINNNLTSITIGEMGKFLKDVMRKTTGYIDVALSDEWKDKPMYYLFT
KKYNPKHANNLLYYFIKEKKLDKFNGNILNVPEYYYRKEGYFKLVAGNYRTKINTLNFKIKSKKVDANSLSEDIEMQTIYEIVK
RGLNKKSDWDSYISYIECVQNPNIDNINRYKLLRDYFCENEDVIKNKIEILSIEQIKEFGGCIMKPHINSMTFGIQKFKIEEIE
NSLGFTFNLPLNKNNYKIELWGHRQLKKGNKESNVNVSLDDFINTYGQNVVFTIKRKKLYIVFSYDYEFERGECNFEKSVGLDV
NFKHSLFVTSEIDNNQFDGYINLYKYILSNNEFTSLLTDSERKDYEDLANIVTFCPFEYQLLFSRYDKLSKISEKEKVLSKILY
SLQKKLKNEKRTKEYIYVSCVNKLRAKYVSYFKLKQKYNEKQKEYDIEMGFVDDSTESKESMDKRRFENPFINTPVAKELLEKM
NNVKQDINGCKKNIVVYAYKVLEQNGYNIIALENLENSNFEKIRVLPKIKSLLEYHKFENKNINDIKNSDKYKEFIEPGYFELI
TNENNEIIDAKYTQKGDIKIKNADFINIMIKALNFASIKDEFILLSHNGKSQIALVPAEYTSQMDSIDHCIYMTKNDKGKLVKV
DKRKVRTKQERHINGLNADFNAACNIKYIVTNEDWRKVFCIKPKKEDYNTPLLDATKNGQFRILDKLKKLNATKLLEMEK
(配列番号24)

>3300028797|Ga0265301_10000251_12|M
[哺乳動物消化系-ルーメン-ウシ(bos taurus)]
MVKVFINVFLSEKNQITTNIFDTEKISNSYINHINHQFMATHKKTDNQTIVKAYVMKAKMSKHDIERVWKPTIDEYINYYNKLS
DWICKNLTSVTIGDLLKYVGEKQINKGVGYYTYFIDEQKTDLPLYTLFTDCPKTHADNLLFEAVRKINPENYNGNLLSLFETGY
RRNGYFDNVISNYRTKMTTLKINPKYKRFSSENMPTDEVLLEQTVYEVTKNDFKNDDDWKKSIDYMKQKSEPNTALIFRMETLF
DYWKDHKQDVEQYINQKRVECLKDFGGCKRRADGLSMVILLNKKLTKIEADGLTSYKLTTNLFGGKYMINIFGHRALVSVCNGE
RAENENIDICNKHGERFTFKIENGNLFVALTADYNYEKQPNLPKNIVGVDINIKHSMLNSSIEDKGKVKGYVNLYKEFLSDKNF
RKTITSDEELNQYIELSKYATFGITELDSLFARATDTEKSILCKRELAMQDVFEKLEKRYKDDHKIKFYLGSTQKLRAQYISYF
KIKEAYNRKQQEYDLAHGKTDNPDEVYKSDFINEPSAKEMLVKLNRIERKIIGCRNNIVTYAFNVIKNNGYDTIGVEYLTSSQF

【０２４０】

【表２１】

EKKRRLPSIKSLLNYRKLLGKPKDEWNLKEWNDVYMCYRPELDDAGNIMNFTITNEGIKRNKESTFYNSFIKAIHFADVKDKFA
QLTNNNTMNTVFIPSSFTSQIDSKTRKLYLLEYTEKCDNGKTKKVVKFINKRVLRKIQEQHLNGMNADNNAARNIRDITKNLRD
VFTKKQTDKNCYNSAEFMIQTKFKKRLPQATVFGELNRNGYVKVLTQEEYDELTKSAK (配列番号25)

>3300028797|Ga0265301_10000251_10|P
[哺乳動物消化系-ルーメン-ウシ(bos taurus)]
MATHKKTDNQTIVKAYVMKAKMSKHDIERVWKPTIDEYINYYNKLSDWICKNLTSVTIGDLLKYVGEKQINKGVGYYTYFIDEQ
KTDLPLYTLFTDCPKTHADNLLFEAVRKINPENYNGNLLSLFETGYRRNGYFDNVISNYRTKMTTLKINPKYKRFSSENMPTDE
VLLEQTVYEVTKNDFKNDDDWKKSIDYMKQKSEPNTALIFRMETLFDYWKDHKQDVEQYINQKRVECLKDFGGCKRRADGLSMV
ILLNKKLTKIEADGLTSYKLTTNLFGGKYMINIFGHRALVSVCNGERAENENIDICNKHGERFTFKIENGNLFVALTADYNYEK
QPNLPKNIVGVDINIKHSMLNSSIEDKGKVKGYVNLYKEFLSDKNFRKTITSDEELNQYIELSKYATFGITELDSLFARATDTE
KSILCKRELAMQDVFEKLEKRYKDDHKIKFYLGSTQKLRAQYISYFKIKEAYNRKQQEYDLAHGKTDNPDEVYKSDFINEPSAK
EMLVKLNRIERKIIGCRNNIVTYAFNVIKNNGYDTIGVEYLTSSQFEKKRRLPSIKSLLNYRKLLGKPKDEWNLKEWNDVYMCY
RPELDDAGNIMNFTITNEGIKRNKESTFYNSFIKAIHFADVKDKFAQLTNNNTMNTVFIPSSFTSQIDSKTRKLYLLEYTEKCD
NGKTKKVVKFINKRVLRKIQEQHLNGMNADNNAARNIRDITKNLRDVFTKKQTDKNCYNSAEFMIQTKFKKRLPQATVFGELNR
NGYVKVLTQEEYDELTKSAK (配列番号26)

>3300028797|Ga0265301_10009039_3|M
[哺乳動物消化系-ルーメン-ウシ(bos taurus)]
MAHKGEKEGYQIKTLKFKVRSHDIGKSLYDIVNEYTNYYNKVSKWICDNLDTPIGELSKNISEKRHNSKYYRATNDPNWKNEPM
WKIFTKKFSNGETFSEQGKNDKLANLSNCDNILSYSIIDYNIDGYTGNILGLTDTSYRLNGYISNCISNYKTKIRTAKPKVRST
AITEHSTVEEKTNNTIYEMVRKGFMSPNDFKNQIKYLTEKENPNDKLIDRLSILHSFYTENEEDVNNAFSRMSVEMLKNNNGCT
RNGDKKTLNISSIDYKVTRKEGCDGYILSFGSRNQKYNIDLWGRRDTISNGKELIDLSEHGEPLTITSENGDYYVCMTVDVPFE
KKSTGSTEKVASVDVNTKHTMLSTDVIDDGTLKGYLNIYKKLLLDTELTSLLHKQDFDDMKELSHNVCFGPIEYNFLLSRILDL
DAYEKKVEDRITHSMKEMLKTETDERNKMYLGSVIKMRALLKVYISTKNRYHKEQQSYDESMGFTDTSTASKDTMDKRRFENPF
SETETGKKLNNDLSALSKKIIGCRDNIVRYAYTTLQDNGYTMIGVEDLNSSTFANTRNPFPTIKSLLNYHHLSGKTPEEARNID
TYSKFSDHYTLTTDEEGKITDAKYTKKAETKIKKKRARDTIIKAIHFAEVKDVMCVMSNNGTASVAFEPSYFSSQMDSATHKVY
TTRNKKGKDVIASKETVRPRQEKHINGMNCDINSPKNLSYLITNEEFREMFLTPTKNGYNEPFYKSRVKSAASMMSGLKKLGAT
MPLTDENAIFSTPKPKKNIGKQ (配列番号27)

>3300028887|Ga0265299_10000013_320|P
[哺乳動物消化系-ルーメン-ウシ(bos taurus)]
MGNKVQSNETIVKTYTFKVREFISGATHEIMKSAIKQYIEDSNNLSDWINNQLTNKTICEVGALIPIEKRETSYYKSTVDELWA
NKPCFKMFTNDFTKEENFATRNIGNGKNCKNIITSAYKSTVNPSFRNVLDLTEKVYFSDGYGANVCSNYKTKLRTLKPAKIKLV
SSLSDCDDNTLTEQVIREKQKYGYSTPKDFEKRIEYLNEKEKSEQNSKIIERLQKLYEFYDNNTKLVEEKELELSVKSLVEFGG
CRRGEKTMTLNLPDIGYEIQRKDDKYGYIFTLKCSKKRKIIIDVWGSKATIDSNGNDKVDIINTHGKSINFKIINNEMYIDITV
DVPFAKRKLGIKKVVGIDVNTKHMLMATNIKVTDSIKGYVNLYKEFLNSKEIMDVASPETKKNFEDMSMFVNFCPIEYNTMFAL
IFKLNNGDIRTEQAIRRTLHQLSKKFSDGNHETERIYVQNVFSIREQLKHFILLSNRYYSEQSDYDTKMGFIDENTTSNATMDK
RRFDKSLMFRYTQRGRQLYEERIECGRKITEIRDNIITYARNVFVLNGYDTIALEYLTNATIQKPTRPTSPKSLLDYFKLKGKP
VVEAEKNERITKNRKYYNLIPDENDNVINIEYTEEGKVAIKKSIARDHIMKAVHFAEVKDKFIQLSNNGKTQVALVPSNYTSQM
NSETHTVYLMKNPKTKKLVIMDKDKVRPIQEKYKLNGLNADFNSARNIAYIVENEILRNSFLKEETKKYTYNTPLFTPRLKSSE
KIITELKKLGMTTVIE (配列番号28)

>3300028887|Ga0265299_10000026_77|P
[哺乳動物消化系-ルーメン-ウシ(bos taurus)]
MANKSTKGNLPKTIIMKANLSPDGFTQWERVVKEYQAYKDTLSKWVAQNLTAMKIGDLLPYLDKYSKKTNKETGERPVNVYYQL
CEQHKDEPLYKLFTYDSNSRNNAMYEIIRKTNCDGYKGNILGISETHYRRNGFVKNILANYTTKISTLELSERKRKIDSDSPED
LIRSQVVYEMQKNNIKDAKGFKSIIEYLKSKKEVNIQYLERLQILYEYFKNHENEIKEYITLAAVEQLKSFGGVRVNNEKSSMN
LEIQGFSITRVDGACTYILHLPINGKIHGIKLWGNRQVVVNKDGTPVDILDLTNQHGSTINITIKNGEIYFAFTVTSDFVKPEH
QIKNVVGVDVNTKHMLMQSNITDNGNVKGYFNIYKVLVEDRRFTSLLSEEQLKYFCELANIVSFCPIETEFLFARYAEYKKMSN
NAEMRQIEKVFSDILDEQYKKYKDIDTSIANYISYVRKLRSQCCAYFKLKMKYKELQRQFDKEQDYKDLSTESKETMDKRRWEN
PFRNTPEASKLIKKMDNVSRQLIGCRDNIITYAYRVFEKNGYDTISLENLESSQFENNDHVIAPKSLLEYHHLKGKTMNYLLSD
ECKVRITTKDGKVKEWYHVELNDKDEIDNIFLTPEGETEKEKNLFNNMVIKIVHFADIKDKFIQLGNYNKLQTVLVPSYFTSQM
DSKTHSVYVVETANTKTSKKELKLVSKKRVRRQQEWHINGLNADYNAACNIAHIAKNIELRQIMCKTPQTKNGYSSPVLTSKVK
SQVEMVRELKKMGKTILYSNDSLPF (配列番号29)

>3300028887|Ga0265299_10000133_30|M
[哺乳動物消化系-ルーメン-ウシ(bos taurus)]
MAHRKKKDDEATLSYKFKVKVIEGDLTADDITKCIAENAEQGNHFSEFIHKNLTSKTIGEFASQLPVEKRQFGYYQYAIGGTMP
AKKNASDEDKPKGELIDWSKKPFYVLFSKGYSATHAVNLIFNVYLNSEEGKAFSAKNSMNLSKSQFAYSGFVQIVCANYASMLA
NARPDKIKFEEITEATDDGTKKMQVVREMAERYLMKPKNFASRIEYLEANNTKGKFDKTIQRLRLLQPFFEKNEEGITELYYDL
SVKALEHSGQCTYKGGRTISILEIGDIRISRKENAKGYLLTIPINRKSVVFDLYGRKDTIGGDGRDLIDIMNTHGSSLQFTADG

【０２４１】

【表２２】

NDIYLTITATKNFIKEKPTFNEDTVLGGDVNIKHSYTVFSTSPKDIPDFVNFYEYFAKDGEIMKLAPKPMWDYIVAAATKFLTI
LPIETPAISATVYGKRTEEGISRATFRETQKLIALEKAIERVMKQVFDKYNDGKHPLEAIYIGNAIKYRRLIKGYLAQKKKYYS
AHSEYDKAMGYTDDDTDRKENMDERRFDDSKKFRYTPEAQALLDTMHTIEKKIVGCVSNAISYAYHKFDENGFNVIALENLTSA
TFAKKYKSDKPESIKKLLNFDKLLGKTLDEAKASKSISKHPNWYELVADENGCVSDIRITDEGQSATYRSLVTETIMKVSHFAE
TKDRFIGLANSGRLQVGLVPSQYTSYIDSTTHTLYAVIEDGKTVLAPKEVVRASQERHINGLNADYNSALNLKYMITDENFRKT
FTSETSADKFGWGKPMFSPTTRSQDEVFSAIKKIGAITVLED (配列番号30)

>3300028887|Ga0265299_10011526_3|M
[哺乳動物消化系-ルーメン-ウシ(bos taurus)]
MAQHKSNNEESAINKTFIFKAKCEKNDVISLWEPAAKEYGDYYNKVSKWIADNLITMKIGDLAQYITNQNSKYYTAVTNKKKKD
LPLYRIFQKGFSSQCADNALYCAIKSINPENYKGNSLGIGESDYRRFGYIQSVVSNFRTKMSSLKVSVKYKKFDVSNVDDETLK
IQTIYDVDKYGIETAKEFKELIETLKTRVETPQLNDTIARLKCLCDYYSKNEKAINNEIETMAIADLQKFGGCQRKSLNAFTIH
KQDSLMEKVGNTSFRLQLSFRKKTYVINLLGNRQVVNFVNGKRVDLIDIAENHGDLITFNIKNGELFLHITSPIVFDKDVRDIR
NVVGIDVNIKHSMLATSIKDDGNVKGYINLYKELLNDDVFVSTCNESELALYRQMSENVNFGILETDSLFERIVNQSKGGCLKN
KLIRRELAMQKVFERITKTNKDQNIVDYVNYVKMMRAKCKASYILKEKYDEKQKEYYVKMGFTDESTESKETMDKRREEFPFVN
TDTAKELLVKQNNIRQDIIGCRDNIVTYAFNVFKNNEYDTLSVEYLDSSQFDKRRIPTPKSLLKYHKFEGKTKDEVENMMKSEK
LSNAYYTFKYENDVVSDIDYSDEGNLRRSKLNFGNWIIKAIHFADIKDKFVQLSNNNKMNIVFCPSAFSSQMDSITHTLYYVEK
ITKNKKGKEKKKYVLANKKMVRTQQETHINGLNADYNSACNLKYIALNYELRDKMTDRFKASKKIKTMYNIPAYNIKSNFKKNL
SAKTIQTFRELGHYRDGKINEDGMFVEILE (配列番号31)

>3300028887|Ga0265299_10012919_3|P
[哺乳動物消化系-ルーメン-ウシ(bos taurus)]
MAHKNSDGENTINKTFIFKVKCEKNDIISFWKPAAEEYCNYYNKLSEWIGKNLISMKIGDLAKYIDNPKSKYYLSVTDENKKDL
PLYKIFQKGFSSIDADNALYCAIDKLNPEGYNGNILGVGKSDYRRNGYVSSVIGNFRTKMVSLKANVRWKKIDIGNVDEETLRR
QTICDVEKYRIESEKDFRDLIDILKAREETPRLKEKISRLELLYDYYSKNTKTIKSEMENMAISDLQKFGGCVRKSLNTITIHK
QDSKIEKEGNTSFRLHMVFNKKPYTITLLGNRQVVKYIDGKRVDIVNIVEKHGDWITFNIKNGELFVHLTKCVEFSKGQKEIKK
AAGVDVNIKHAMLAASIVDDGQLKGYVNLYRELIEDDDFVSTFGDSDSGKTELGMYQKMAKTVFFGVLEVESLFERVVNQQSGW
KLDNQLIRRERAMEKVFDRIVKTTSNKHIIDYVNYVKMLRAKYKAYFILDEKYHEKQREYDLSMGFTDESDERRELYPFINTET
AKEILGKKRNVEQDLIGCRDNIVTYAFNVLRNNGYDTISVEYLDSSQFDKRRMPTPKSLLEYHKFKGKTQDEVERLMSEKKFAK
TNYDIHYDGENKVDGIVYSKEGELRQKKLNFMNLVIKAIHFADIKDKFAQLCNNNDVNVVFGPSAFTSQMDSETHSLYYVEKET
NGKNGKTGKKFVLADKKSVRRRQETHINGLNADFNAARNLEYIASNPELLERMTKRTKSGKDMYNTPSWNIRQEFKKNLSVRTI
NTFRELGNVKYGKINNEGLFVEDDV (配列番号32)

>3300028914|Ga0265300_10009460_3|M
[哺乳動物消化系-ルーメン-ウシ(bos taurus)]
MAHRKKKDDEATLSYKFKVKVIEGDLTADDITKCIAENAEQGNHFSEFIHKNLTSKTIGEFASQLPAEKRQFGYYQYAIGGTMP
AKKNASDEDKPKGELIDWSKKPFYVLFSKGYSATHAVNLIFNVYLNSEEGKAFSAKNSMNLSKSQFAYSGFVQIVCANYASMLA
NARPDKIKFEEITEATDDGTKKMQVVREMAERYLMKPKNFASRIEYLEANNTKGKFDKTIQRLRLLQPFFEKNEESITELYYDL
SVKALEHSGQCTYKGGRTISILEIGDIRISRKENAKGYLLTIPINRKSVVFDLYGRKDTIGGDGRDLIDIMNTHGSSLQFTADE
NDIYLTITATKNFIKEKPTFNEDTVLGGDVNIKHSYTVFSASPKDIPDFVNFYEYFAKDGEIMKLAPKPMWDYIVAAATKFLTI
LPIETPAISATVYGKRTEEGISRATFRETQKLIALEKAIERVMKQVFDKYNDGKHPLEAIYIGNAIKYRRLIKGYLAQKKKYYS
AHSEYDKAMGYTDDDTDRKENMDERRFDDSKKFRYTPEAQALLDTMHTIEKKIVGCVSNAISYAYHKFDENGFNVIALENLTSA
TFAKKYKSDKPESIKKLLNFDKLLGKTLDEAKASKSISKHPNWYELVADENGCVSDIRITDEGQSATYRSLVTETIMKVSHFAE
TKDRFIGLANSGRLQVGLVPSQYTSYIDSTTHTLYAVIEDGKTVLAPKEVVRASQERHINGLNADYNSALNLKYMITDENFRKT
FTSETSADKFGWGKPMFSPTTRSQDEVFSAIKKIGAITVLED (配列番号33)

>3300031853|Ga0326514_10013355_6|M
[哺乳動物消化系-ルーメン-ウシ(bos taurus)]
MVTTLAPLIEEKKRDSEYYKYLTNGDWDGKPLYFIFKEGFNSTNADNILANSLVRVYCEQNYTGNGFGLSYSYYVVIGFAKEVI
ANYRSSFQKPKVKIKKKKLSENPTEDELIEQCIYTIYYEFNEKKDIKKWKDEIKFLKERGESKETRLKRIQTLFEFYKDKNHKE
LVDERVANLVVDNIKEFGGCKRDIGCPSMGIQIQHNFDISINEKRNGYTICFGPNKKNLTKLEVFGNRMVLLNGEEIVDLPNTH
GEKLTLIDRGNAIYAALTAQVPFEKHMPDGNKTVGIDLNLKHSVFATSIVDNGKLAGYISIYKELLKDDEFVKYCPKDLLRFMK
DASKYVFFAPIEIELLRSRVIYNKGYACVENYENVYKAEVAFVNVIKRLQSQCEANGDAQGALYMSYLSKMRAQLKNYINLKLA
YYDHQSAYDLKMGFNDISAESKETIDERRKLFPFSKEKEAQEILAKMKNISNVIIACRNNIAVYMYKMFERNGYDFIGLEKLES
SQMKKRQSRSFPTVKSLLNYHKLAGMTMDEIKKQEVSSNIKKGFYDLEFDADGKLYGAKYSNKGNVHFIEDEFYISGLKAIHFA
DMKDYFVRLSNNGKVSVALVPPSFTSQMDSVEHKFFMKKNANGKLIVADKKDVRSCQEKHKINGLNADYNAACNIGFIVEDDYM
RESLLGSPTGGTYDTAYFDTKIQGSKGVYDKIKENGETYIAVLSDDVITAEE (配列番号34)

>3300031993|Ga0310696_10000014_323|P
[哺乳動物消化系-ルーメン-ウシ(bos taurus)]
MGNKVQSNETIVKTYTFKVREFISGATHEIMKSAIKQYIEDSNNLSDWINNQLTNKTICEVGALIPIEKRETSYYKSTVDELWA
NKPCFKMFTNDFTKEENFATRNIGNGKNCKNIITSAYKSTVNPSFRNVLDLTEKVYFSDGYGANVCSNYKTKLRTLKPAKIKLV

【０２４２】

【表２３】

SSLSDCDDNTLTEQVIREKQKYGYSTPKDFEKRIEYLNEKEKSEQNSKIIERLQKLYEFYDNNTKLVEEKELELSVKSLVEFGG
CRRGEKTMTLNLPDIGYEIQRKDDKYGYIFTLKCSKKRKIIIDVWGSKATIDSNGNDKVDIINTHGKSINFKIINNEMYIDITV
DVPFAKRKLGIKKVVGIDVNTKHMLMATNIKVTDSIKGYVNLYKEFLNSKEIMDVASPETKKNFEDMSMFVNFCPIEYNTMFAL
IFKLNNGDIRTEQAIRRTLHQLSKKFSDGNHETERIYVQNVFSIREQLKHFILLSNRYYSEQSDYDTKMGFIDENTTSNATMDK
RRFDKSLMFRYTQRGRQLYEERIECGRKITEIRDNIITYARNVFVLNGYDTIALEYLTNATIQKPTRPTSPKSLLDYFKLKGKP
VVEAEKNERITKNRKYYNLIPDENDNVINIEYTEEGKVAIKKSIARDHIMKAVHFAEVKDKFIQLSNNGKTQVALVPSNYTSQM
NSETHTVYLMKNPKTKKLVIMDKDKVRPIQEKYKLNGLNADFNSARNIAYIVENEILRNSFLKEETKKYTYNTPLFTPRLKSSE
KIITELKKLGMTTVIE (配列番号35)

>3300031993|Ga0310696_10000226_76|P
[哺乳動物消化系-ルーメン-ウシ(bos taurus)]
MANKSTKGNLPKTIIMKANLSPDGFTQWERVVKEYQAYKDTLSKWVAQNLTAMKIGDLLPYLDKYSKKTNKETGERPVNVYYQL
CEQHKDEPLYKLFTYDSNSRNNAMYEIIRKTNCDGYKGNILGISETHYRRNGFVKNILANYTTKISTLELSERKRKIDSDSPED
LIRSQVVYEMQKNNIKDAKGFKSIIEYLKSKKEVNIQYLERLQILYEYFKNHENEIKEYITLAAVEQLKSFGGVRVNNEKSSMN
LEIQGFSITRVDGACTYILHLPINGKIHGIKLWGNRQVVVNKDGTPVDILDLTNQHGSTINITIKNGEIYFAFTVTSDFVKPEH
QIKNVVGVDVNTKHMLMQSNITDNGNVKGYFNIYKVLVEDRRFTSLLSEEQLKYFCELANIVSFCPIETEFLFARYAEYKKMSN
NAEMRQIEKVFSDILDEQYKKYKDIDTSIANYISYVRKLRSQCCAYFKLKMKYKELQRQFDKEQDYKDLSTESKETMDKRRWEN
PFRNTPEASKLIKKMDNVSRQLIGCRDNIITYAYRVFEKNGYDTISLENLESSQFENNDHVIAPKSLLEYHHLKGKTMNYLLSD
ECKVRITTKDGKVKEWYHVELNDKDEIDNIFLTPEGETEKEKNLFNNMVIKIVHFADIKDKFIQLGNYNKLQTVLVPSYFTSQM
DSKTHSVYVVETANTKTSKKELKLVSKKRVRRQQEWHINGLNADYNAACNIAHIAKNIELRQIMCKTPQTKNGYSSPVLTSKVK
SQVEMVRELKKMGKTILYSNDSLPF (配列番号36)

>3300031993|Ga0310696_10000447_27|M
[哺乳動物消化系-ルーメン-ウシ(bos taurus)]
MAHRKKKDDEATLSYKFKVKVIEGDLTADDITKCIAENAEQGNHFSEFIHKNLTSKTIGEFASQLPVEKRQFGYYQYAIGGTMP
AKKNASDEDKPKGELIDWSKKPFYVLFSKGYSATHAVNLIFNVYLNSEEGKAFSAKNSMNLSKSQFAYSGFVQIVCANYASMLA
NARPDKIKFEEITEATDDGTKKMQVVREMAERYLMKPKNFASRIEYLEANNTKGKFDKTIQRLRLLQPFFEKNEEGITELYYDL
SVKALEHSGQCTYKGGRTISILEIGDIRISRKENAKGYLLTIPINRKSVVFDLYGRKDTIGGDGRDLIDIMNTHGSSLQFTADG
NDIYLTITATKNFIKEKPTFNEDTVLGGDVNIKHSYTVFSTSPKDIPDFVNFYEYFAKDGEIMKLAPKPMWDYIVAAATKFLTI
LPIETPAISATVYGKRTEEGISRATFRETQKLIALEKAIERVMKQVFDKYNDGKHPLEAIYIGNAIKYRRLIKGYLAQKKKYYS
AHSEYDKAMGYTDDDTDRKENMDERRFDDSKKFRYTPEAQALLDTMHTIEKKIVGCVSNAISYAYHKFDENGFNVIALENLTSA
TFAKKYKSDKPESIKKLLNFDKLLGKTLDEAKASKSISKHPNWYELVADENGCVSDIRITDEGQSATYRSLVTETIMKVSHFAE
TKDRFIGLANSGRLQVGLVPSQYTSYIDSTTHTLYAVIEDGKTVLAPKEVVRASQERHINGLNADYNSALNLKYMITDENFRKT
FTSETSADKFGWGKPMFSPTTRSQDEVFSAIKKIGAITVLED (配列番号37)

>3300031993|Ga0310696_10026614_2|M
[哺乳動物消化系-ルーメン-ウシ(bos taurus)]
MAHKNSDGENTINKTFIFKVKCEKNDIISFWKPAAEEYCNYYNKLSEWIGKNLISMKIGDLAKYIDNPKSKYYLSVTDENKKDL
PLYKIFQKGFSSIDADNALYCAIDKLNPEGYNGNILGVGKSDYRRNGYVSSVIGNFRTKMVSLKANVRWKKIDIGNVDEETLRR
QTICDVEKYRIESEKDFRDLIDILKAREETPRLKEKISRLELLYDYYSKNTKTIKSEMENMAISDLQKFGGCVRKSLNTITIHK
QDSKIEKEGNTSFRLHMVFNKKPYTITLLGNRQVVKYIDGKRVDIVNIVEKHGDWITFNIKNGELFVHLTKCVEFSKGQKEIKK
AAGVDVNIKHAMLAASIVDDGQLKGYVNLYRELIEDDDFVSTFGDSDSGKTELGMYQKMAKTVFFGVLEVESLFERVVNQQSGW
KLDNQLIRRERAMEKVFDRIVKTTSNKHIIDYVNYVKMLRAKYKAYFILDEKYHEKQREYDLSMGFTDESDERRELYPFINTET
AKEILGKKRNVEQDLIGCRDNIVTYAFNVLRNNGYDTISVEYLDSSQFDKRRMPTPKSLLEYHKFKGKTQDEVERLMSEKKFAK
TNYDIHYDGENKVDGIVYSKEGELRQKKLNFMNLVIKAIHFADIKDKFAQLCNNNDVNVVFGPSAFTSQMDSETHSLYYVEKET
NGKNGKTGKKFVLADKKSVRRRQETHINGLNADFNAARNLEYIASNPELLERMTKRTKSGKDMYNTPSWNIRQEFKKNLSVRTI
NTFRELGNVKYGKINNEGLFVEDDV (配列番号38)

>3300031993|Ga0310696_10030100_3|M
[哺乳動物消化系-ルーメン-ウシ(bos taurus)]
MAQHKSNNEESAINKTFIFKAKCEKNDVISLWEPAAKEYGDYYNKVSKWIADNLITMKIGDLAQYITNQNSKYYTAVTNKKKKD
LPLYRIFQKGFSSQCADNALYCAIKSINPENYKGNSLGIGESDYRRFGYIQSVVSNFRTKMSSLKVSVKYKKFDVSNVDDETLK
IQTIYDVDKYGIETAKEFKELIETLKTRVETPQLNDTIARLKCLCDYYSKNEKAINNEIETMAIADLQKFGGCQRKSLNAFTIH
KQDSLMEKVGNTSFRLQLSFRKKTYVINLLGNRQVVNFVNGKRVDLIDIAENHGDLITFNIKNGELFLHITSPIVFDKDVRDIR
NVVGIDVNIKHSMLATSIKDDGNVKGYINLYKELLNDDVFVSTCNESELALYRQMSENVNFGILETDSLFERIVNQSKGGCLKN
KLIRRELAMQKVFERITKTNKDQNIVDYVNYVKMMRAKCKASYILKEKYDEKQKEYYVKMGFTDESTESKETMDKRREEFPFVN
TDTAKELLVKQNNIRQDIIGCRDNIVTYAFNVFKNNEYDTLSVEYLDSSQFDKRRIPTPKSLLKYHKFEGKTKDEVENMMKSEK
LSNAYYTFKYENDVVSDIDYSDEGNLRRSKLNFGNWIIKAIHFADIKDKFVQLSNNNKMNIVFCPSAFSSQMDSITHTLYYVEK
ITKNKKGKEKKKYVLANKKMVRTQQETHINGLNADYNSACNLKYIALNYELRDKMTDRFKASKKIKTMYNIPAYNIKSNFKKNL
SAKTIQTFRELGHYRDGKINEDGMFVEILE (配列番号39)

>3300031998|Ga0310786_10000003_467|M

【０２４３】

【表２４】

[哺乳動物消化系-ルーメン-ウシ(bos taurus)]
MAHRKKKDDEATLSYKFKVKVIEGDLTADDITKCIAENAEQGNHFSEFIHKNLTSKTIGEFASQLPAEKRQFGYYQYAIGGTMP
AKKNASDEDKPKGELIDWSKKPFYVLFSKGYSATHAVNLIFNVYLNSEEGKAFSAKNSMNLSKSQFAYSGFVQIVCANYASMLA
NARPDKIKFEEITEATDDGTKKMQVVREMAERYLMKPKNFASRIEYLEANNTKGKFDKTIQRLRLLQPFFEKNEESITELYYDL
SVKALEHSGQCTYKGGRTISILEIGDIRISRKENAKGYLLTIPINRKSVVFDLYGRKDTIGGDGRDLIDIMNTHGSSLQFTADE
NDIYLTITATKNFIKEKPTFNEDTVLGGDVNIKHSYTVFSASPKDIPDFVNFYEYFAKDGEIMKLAPKPMWDYIVAAATKFLTI
LPIETPAISATVYGKRTEEGISRATFRETQKLIALEKAIERVMKQVFDKYNDGKHPLEAIYIGNAIKYRRLIKGYLAQKKKYYS
AHSEYDKAMGYTDDDTDRKENMDERRFDDSKKFRYTPEAQALLDTMHTIEKKIVGCVSNAISYAYHKFDENGFNVIALENLTSA
TFAKKYKSDKPESIKKLLNFDKLLGKTLDEAKASKSISKHPNWYELVADENGCVSDIRITDEGQSATYRSLVTETIMKVSHFAE
TKDRFIGLANSGRLQVGLVPSQYTSYIDSTTHTLYAVIEDGKTVLAPKEVVRASQERHINGLNADYNSALNLKYMITDENFRKT
FTSETSADKFGWGKPMFSPTTRSQDEVFSAIKKIGAITVLED (配列番号40)

>AUXO013988882|Ga0247611_10000101_23|P
[哺乳動物消化系-ルーメン-ヒツジ(ovis aries)]
MANKRTDTTINLNKTVIMLTNMLPEVRAMFQAGIRQAQAYADLVNKWICSNLTNKIGEVLLPYIDNKNCVYYELCYKYKEAPLY
TIFMKGKFDLNSRNNALYCAVVAQNIDNYSGNIFGFSQSDYRRNGYCKVVFSNYATKMSSLKPSIKKVTINEESTEETIQSQVI
YEMFTNGRQWGKPEYFAEHLKYLEMKDNVSDKLMFRMKTLCEYYQTHTDLIDTMAMNAGVEALKQFEGLKLNRDKFSMTITTNS
TSPYTLTRVAGTCAYNLHIPCRKRSYDIRLWGNRQTVRWVNGELVDIADIINQHGQTIIFTIKNGNVYVHIPYGLNFEKTEHEI
KNVVGVDVNTKHMLMQTSIKDNGWVKGYVNIYKALVEDEEFVKYISKSDLKLYKDLSKYVSFCPLELNLLYTRYLSKKGLPFNE
ADNNAEKCVEKVLNNLVKQYEGDDVHVVNYIHNVKKLRALCKASFVLYKKYAELQKAFDDAQGYNDQSTETKETMDKRRWENPF
IQTREAQELIAKMDNAVAGIIGCRDNIITYAYKVFGDNNYDTVGLENLTTSQFDNYSTVKSPKSLLSYYGLLGQQVDSDKYNAV
MTESNKDWYDFKTDGDGNITDITLTAAGEAQKAKSLFNNKVLKNIHFADVKDKFIQLGNNGSIQTVLVPPSYTSQMDSKTHTIY
VKETVDPKNKNKKKLKLVDKKLVRHGQEYHKNGLNADINAALNIAYIVENQEMREVMCLHPSKKDGVYDQPFLKATTKYPATVA
GILLKMGKTTNWGEK (配列番号41)

>3300028805|Ga0247608_10000186_37|P
[哺乳動物消化系-ルーメン-ヒツジ(ovis aries)]
MNKSYVFKSNVAIDDIMSLFEPAIEEYINYYNRTSDFICDNLTSMKIGDLANYIKNKENVYCKFVLNDDIKDLPLYKIFSLNLN
SSQKKNADNALYEAIKVLNADGYKGKNILGLGDTYFRRNGYVKNVISNYRTKFVTLKPNVKYSKIDINSVTEQLIKTQTIFEVV
NKKIESETDFENLITYFKNRETPNDEKIKRLELLFDYYTKHKNEINEEIEKHAVESLKSFNGCRRNGNRKTMTVQMQKMLLKKH
GLTSYILHLVLDKKPYDINLMGNRQTVKVDNNGNRVDLVDISSKHGYDLTFEVKGKTLFFTFSSEKDFSKKEQEIKNILGIDIN
TKHSMLATSITDNGKVKGYINIYVELLKNKDFVSTLNKEELAYYTEMAKFVSFGLLEIPSLFERVSNQYDKKNNVSITDETLLK
REIAISQTLDNLAKKYRDKNCKIASYIDYTKMLRSKYKSYFILKQKYYEKNHEYDDKMGFSDISTNSKETMDPRRFENPFINTD
IAKGLIVKLENVKCDIVGCRDNIIKYAYDVIVLNGFDTIGLEYLDSSNFERDRLPFPTAKSLMTYYGFEGKKYSEIDKSVFNTK
YYNFIFNENETIKDISYSVYGLKEIQKKRFKNLVIKAIGFADIKDKFVQLSNNTNMNVIFVPAAFTSQMDSNTHKIYVKEIMDK
NNKKQLQLIDKRKVRTKQEFHINGLNADFNAANNIKYIAENNDLLLTMCTKTKENNRYGNPLYNIKDTFKKKIPSSILNIFKKK
DMYQIICD (配列番号42)

>3300028805|Ga0247608_10000895_42|M
[哺乳動物消化系-ルーメン-ヒツジ(ovis aries)]
MFRIFAALKLTNMGHVRLQKREGEVYKTYKLKVKSFSGNVDIKAGIVEYDQKFNNVSQWIADHLTSMTIGEAASRISPHKMDSQ
YAMTSLSDEWKDQPLYKIFTRGFGGMNADNLIIECTKTEENCKYDKEKSLGFSESVFRTFGFAANASSDMKSRMTQAKVKIGRK
NIDEDSADDEKCLQAIYEIQKNELLTDDNWKDRIGYLEMKGDQERELERTTILYDYYRANRTTVLDKLDNLKVETLSKFRGSKR
KSDRKILTLNGISYDIKRKEGCQGFELKFSVDKNHMEFDLLGHRALIKNGEMLVDIENCHGSQLSLEIDGDDMYAIISMRTFCE
KNESKLEKIIGADVNIKHMFLMTSEKDDGNTKCYVNLYRELLSDSDFTDVLNKEEYEIFSELSKYVMFGLIETPYLGSRVIGTT
QHEKIVEDKITSGMKKIAIRLFQEGKVRERIYVQNVLKIRALLKALFSTKLAYSNEQKIYDNLMRFGEKDDRRKDEGFHTTCRG
TSLRSEMDMLSKKILACRDNIVEYGYYVIGLNGFDGISLENLESSTFMDVKISYPSCNSMLDHFKLKGKTIEEAENHETVGKFI
KKGYYVMTLVNGKINDINYSEKAVMLHKKNLLYDTVIKSTHFADVKDKFVELSNNGKVSVVIVPPYFSSQMDSVTHKVFTEEIV
VQKKSSNGKVRKTKKTVLVDKRKVRKTQESHINGLNADYNAALNLKYIAETTIDWRSTLCFKTWNTYGSPQWDSKIKNQKTMIDR
LDSLGAIELKNW (配列番号43)

>3300028805|Ga0247608_10006074_1|M
[哺乳動物消化系-ルーメン-ヒツジ(ovis aries)]
MSHEFNKNKGENEISKTFIFKTKCGKNDITSLWVPAMEEYCTYYNRVSKWICDNLTEMRIGDLAQYIDNHGSAYYSAVTDITKK
DLPLYKIFKKGFSGLCADNALYCAIAKLNPEGYDGNMFGLSETYYRRQGYIANVFGNYRTKMNAGLKVGCAKWKKFDTNDVDDE
ILMEQVIVDVVKYDIDSKNEFKEYIEVLKCREENPKLLETIERLECLYGYYSQHEEDIKKKIEELVVEELKTFGGCVRKSMTSC
TITVQDFVMERIGNTGYRINLTFNKKPYVLGLLGNRQVVRYVDGDRVELVDIVNNHGNQITFNLKNGELFVHLTSGVDFSKEES
SMENIVGVDVNIKHSMLASSIVDDGNVNGYINIYKELVNDDEFVSTFGDSESGLNELELYRQMAESVNFGLMETDSLFERYVEQ
WKGSDSDSRLARRERVVGKVFDRIVKTNGDVHVVNYIHAVKMLRAKCKAYFVLKQKYYEKQKEYDDAHGYTDESTASKETMDKR
RFENPFVETDVAKELLGKLACVEQDIIGCRDNIVTYAFNVFRRNGYDTISLEYLDSSQFKKIGMGAPTPKSLLKYHKLEGKTVE
EVESIISEKGLKKNLYVFKFGDNGLLSDIEYSDEGLIRKKKADFGNIITKAIHFADIKDKFVQLTNNSDMGVVFCPSAFTSQMD
SKTHRLYFVEGLDGNGKNKYVLANKWSVRRQQERHINGLNADFNSACNCQHIAYDPILRDAMTIKVEAGKGMYNKPSYDIRKKF

【０２４４】

【表２５】

KKNLSAATLKTFIKLGNTVKGMIVNGQFVEMES (配列番号44)

>3300028833|Ga0247610_10000007_379|M
[哺乳動物消化系-ルーメン-ヒツジ(ovis aries)]
MYNSKKKGEGDIQKSFKFKVKTDKETVELFRKAAVEYSEYYKRLTTFLCERLTDMTWGEVASFIPEKYRKNEYYKYLIKEENKD
LPLYKMFTKAASSMFIDHSIERYVEALNPEGNTGNILGFCKSSYVRGGYLKNVVSNIRTKFATLKTGIKYKKFNPAEDDEETIL
GQTVFEMEKRGLEFKCDFEKTIKYLNEKGKTQEAERLQCLMEYFSTNTDKINEYRESLVLDDIRKFGGCNRSKSNSFSVTLEKA
DIKEDGLTGYTMKVSKKLKEIHLLGHRRVVEVVNGRRVNLVDICGDKSGDSKVFVVDGDNLYVCISAPVKFSKNGMEAKKYIGV
DMNMKHSIISVSDNASDMKGFLNIYKELLKDEGFRKTLNATELEKYEKLAEGVNIGIIEYDGLYERIVKQKKENSVDGLKVQAE
KKLIEREAAIERVLDKLRKGTSDTDTENYINYNKILRAKIKSAYILKDKYYEMLGKYDSERAGSGDLSEENKIKYKDEFNETEK
GKEILGKLNNVYKDIIGCRDNIVTYAVNLFIRNGYDTVALEYLESSQMKARRIPSTGGLLKGHKLEGKPEGEVTAYLKANKIPK
SYYSFEYDGNGMLTDVKYSDMGEKARGRNRFKNLVPKFLRWASIKDKFVQLSNYKDIQMVYVPSPYTSQTDSRTHSLYYIETVK
VDEKTGKEKKEHIVAPKESVRTEQESFVNGMNADTNSANNIKYIFENETLRDKFLKRTKDGTEMYNRPAFDLKECYKKNSNVSV
FNTLKKTLGAIYGKLDENGNFIENECNK (配列番号45)

>3300028833|Ga0247610_10004486_2|M
[哺乳動物消化系-ルーメン-ヒツジ(ovis aries)]
MNKSYVFKSNVAIDDIMSLFEPAIEEYINYYNRTSDFICDNLTSMKIGDLANYIKNKENVYCKFVLNDDIKDLPLYKIFSLNLN
SSQKKNADNALYEAIKVLNADGYKGKNILGLGDTYFRRNGYVKNVISNYRTKFVTLKPNVKYSKIDINSVTEQLIKTQTIFEVV
NKKIESETDFENLITYFKNRETPNDEKIKRLELLFDYYTKHKNEINEEIEKHAVESLKSFNGCRRNGNRKTMTVQMQKMLLKKH
GLTSYILHLVLDKKPYDINLMGNRQTVKVDNNGNRVDLVDISSKHGYDLTFEVKGKTLFFTFSSEKDFSKKEQEIKNILGIDIN
TKHSMLATSITDNGKVKGYINIYVELLKNKDFVSTLNKEELAYYTEMAKFVSFGLLEIPSLFERVSNQYDKKNNVSITDETLLK
REIAISQTLDNLAKKYRDKNCKIASYIDYTKMLRSKYKSYFILKQKYYEKNHEYDDKMGFSDISTNSKETMDPRRFENPFINTD
IAKGLIVKLENVKCDIVGCRDNIIKYAYDVIVLNGFDTIGLEYLDSSNFERDRLPFPTAKSLMTYYGFEGKKYSEIDKSVFNTK
YYNFIFNENETIKDISYSVYGLKEIQKKRFKNLVIKAIGFADIKDKFVQLSNNTNMNVIFVPAAFTSQMDSNTHKIYVKEIMDK
NNKKQLQLIDKRKVRTKQEFHINGLNADFNAANNIKYIAENNDLLLTMCTKTKENNRYGNPLYNIKDTFKKKIPSSILNIFKKK
DMYQIICD (配列番号46)

>3300028888|Ga0247609_10000668_74|M
[哺乳動物消化系-ルーメン-ヒツジ(ovis aries)]
MAHKTKESEKLVKSFKLKVDISNCEIEKKWIPSFEEYTNYYNGVSNWICENLISMKIGDLGQYIKNTESVYYKFITDESISNLP
LYKIFTLKQTQNVDNALFCAIKEINPEKYNGNSIGLGETDYRRFGYVQCVISNYRTKIGTMKASIKYKTLPENQSYDVIFEQTM
YEMIDKSLEKKEDWENIISNYKAKQTENTSKINRMETLYSFFIEHSEEIIEKSNLVAIEQLALFNGCKRKSLSTMTIHSQHSKL
QKNGLTSFVFCINQKIGSINLFGNRQLVSVDENGNRNDIIDICNNYGDFITFQIKNGKMFIILTAKVDFDKENIEIKNVVGADV
NIKHNMIASSIIDNGNVFGYINIYKELLNDEDFCSSCTNEELDIYKEISKSVNFGLLECESLFSRVSAQIYKENESISKLDDRF
LRREKSIENVLNRLSKQYRYKDCKIATYIDYTKIMRDSYKSYFIIKEKYYEKQKEYDISMGYVDESTNSKKTMDKRRFENPFIE
TETAKNILSKLNRIESRLIGCRNNITNYAFDVFKNNGFDTIALEYLDSSQFDKTKVLTPISMLKYHKFEGKSIEEVKTLNVKFS
MDNYEFEFDNNGKITNISFSQLGKREVMKTNFFNLIIKAIHFAEIKDKFIQLSNNKPINIVLVPSAFSSQMDSKDHKLYVDENG
KLINKRKVRKQQERHINGLNADFNAACNLSYLAKNNELLEKVCLKRKKFGKASYSVPYWNVKDAFKKNVSSNMIATIKKMNMVK
VF (配列番号47)

>3300028888|Ga0247609_10003329_9|M
[哺乳動物消化系-ルーメン-ヒツジ(ovis aries)]
MAHKTNNGENTINKTFIFKAKCEKNDIISLWKPAAEEYCNYYNKLSKWIGDSLTTMKIGDLAQYITNQNSAYYLAVTNDSKKDL
PLYKIFQKGFSSQCADNALYSAIKAINPENYNGNSLEIGETDYRRFGYVQSVIGNFRTKMSSLKVSVKYKKFDVNDVDEETLKT
QTIYDVDKYGIESIKDFNEFIEVLKLREETPQLNEKITRLECLCGYYSKNEENIKNEIETMAISDLQKFGGCQRKSLNTLTIHK
QNSLMEKVGNTSFTLQLSFNKKPYTINLLGNRQVVKFVDGKRVDLIDITEKHGDWVTFNIKNDELFVHLTSPIDFEKEVCEIKN
AVGVDVNIKHNMLATSIKDDGNVKGYINLYKELVNDCDFISTCNEDEFDLYRQMSESVNFGILETDSLFERVVNQSKGGCLNNK
FIRRELAMQKVFDNITKTNKDQNIVDYVNYVKMLRAKYKAYFILKEKYYEKQKEYDIKMGFTDVSTESKETMDKRRMEFPFVNT
DTAKELLAKLNNIEQDLIGCRDNIVTYAFNIFKNNGYDTLAVEYLDSAQFDKRRMPTPTSLLKYHKFEGKTKDEVEDMMKSKKF
SNAYYTFKFENDVVSNIEYSNDGIWKQKQLNFGNLIIKAIHFADIKDKFVQLCNNNKMNIVFCPSAFTSQMDSITHTLYYVEKI
TKKKNGKEEKKYVLANKKMVRTQQETHINGLNADYNSACNLKYIALNDELRNEMTDTFKVTNRQKTMYGIPAYNIKRGFKKNLS
AKTINTFRKLGHYRDGKINEDGMFVETLA (配列番号48)

>3300028888|Ga0247609_10016480_8|M
[哺乳動物消化系-ルーメン-ヒツジ(ovis aries)]
MAHKTNNGENTINKTFIFKAKCDNNDIISLWKPAMEEYCTYYNKLSQWICNNLTSMKVKDLFAYLDDKQKTKPCVDKKTGETKI
GVGYYRYFIENNKEDMPLYWLFTKNCSSSHADNLLFEFVRKVNHEEYNGNSLGMGETDYRRFGYFQNVISNFRTKMSSLKATTK
WKKFDVNDVDEDTLKNQTIYDVDKYGIESVNDFNERIDILKIREETEQTKDKIARLECLCKYYKEHEEDIKNEIATMAIADLQK
FGGCQRKSMNTLTIHKQDSPMEKVGNTSFNLRLTFNKKPYTLNLLGNRQVVKFVGGKRIDLINITENHGDWITFNIKNNELFVH
MTSPVDFEKEVCEIKNAVGVDVNIKHMMLATSIVDDGNVKGYINLYRELVNNNDFIATFGNSKNGHQGLEIYEQMAENVNFGIL
ETESLFERVVNQSNGGELNNQLIRREIAMQKVFDNITKTNNDKNIVNYVNYVKMLRAKYKAYFILKEKYYEKQKEYDDMMGFND

【０２４５】

【表２６】

ESTENKEMMDKRRFEFSFINTDTAQELLIKLNKVEQDLIGCRDNIVTYAFNVFKTNGYDTLAVEYLDSAQFDKAKMPTPKSLLK
YHKFEGKTIDEVKEMMNNKNFTNAYYNFKFENEIVKDIEYSTDGIWRQKKLNFMNLIIKAIHFADIKDKFVQLCNNNSMNVVFC
PSAFTSQMDSITHSLYYIEKTSKTKNGKEKKQYVLANKKMVRTQQEKHINGLNADFNSACNLKYIALDEELRNAMTDEFNPKKQ
KTMYGVPAYNIKNGFKKNLSTKTINTFRTLGHYRDGKINEDGVFVENLA (配列番号49)

>3300031992|Ga0310694_10000010_351|M
[哺乳動物消化系-ルーメン-ヒツジ(ovis aries)]
MYNSKKKGEGDIQKSFKFKVKTDKETVELFRKAAVEYSEYYKRLTTFLCERLTDMTWGEVASFIPEKYRKNEYYKYLIKEENKD
LPLYKMFTKAASSMFIDHSIERYVEALNPEGNTGNILGFCKSSYVRGGYLKNVVSNIRTKFATLKTGIKYKKFNPAEDDEETIL
GQTVFEMEKRGLEFKCDFEKTIKYLNEKGKTQEAERLQCLMEYFSTNTDKINEYRESLVLDDIRKFGGCNRSKSNSFSVTLEKA
DIKEDGLTGYTMKVSKKLKEIHLLGHRRVVEVVNGRRVNLVDICGDKSGDSKVFVVDGDNLYVCISAPVKFSKNGMEAKKYIGV
DMNMKHSIISVSDNASDMKGFLNIYKELLKDEGFRKTLNATELEKYEKLAEGVNIGIIEYDGLYERIVKQKKENSVDGLKVQAE
KKLIEREAAIERVLDKLRKGTSDTDTENYINYNKILRAKIKSAYILKDKYYEMLGKYDSERAGSGDLSEENKIKYKDEFNETEK
GKEILGKLNNVYKDIIGCRDNIVTYAVNLFIRNGYDTVALEYLESSQMKARRIPSTGGLLKGHKLEGKPEGEVTAYLKANKIPK
SYYSFEYDGNGMLTDVKYSDMGEKARGRNRFKNLVPKFLRWASIKDKFVQLSNYKDIQMVYVPSPYTSQTDSRTHSLYYIETVK
VDEKTGKEKKEHIVAPKESVRTEQESFVNGMNADTNSANNIKYIFENETLRDKFLKRTKDGTEMYNRPAFDLKECYKKNSNVSV
FNTLKKTLGAIYGKLDENGNFIENECNK (配列番号50)

>3300031992|Ga0310694_10022272_2|M
[哺乳動物消化系-ルーメン-ヒツジ(ovis aries)]
MNKSYVFKSNVAIDDIMSLFEPAIEEYINYYNRTSDFICDNLTSMKIGDLANYIKNKENVYCKFVLNDDIKDLPLYKIFSLNLN
SSQKKNADNALYEAIKVLNADGYKGKNILGLGDTYFRRNGYVKNVISNYRTKFVTLKPNVKYSKIDINSVTEQLIKTQTIFEVV
NKKIESETDFENLITYFKNRETPNDEKIKRLELLFDYYTKHKNEINEEIEKHAVESLKSFNGCRRNGNRKTMTVQMQKMLLKKH
GLTSYILHLVLDKKPYDINLMGNRQTVKVDNNGNRVDLVDISSKHGYDLTFEVKGKTLFFTFSSEKDFSKKEQEIKNILGIDIN
TKHSMLATSITDNGKVKGYINIYVELLKNKDFVSTLNKEELAYYTEMAKFVSFGLLEIPSLFERVSNQYDKKNNVSITDETLLK
REIAISQTLDNLAKKYRDKNCKIASYIDYTKMLRSKYKSYFILKQKYYEKNHEYDDKMGFSDISTNSKETMDPRRFENPFINTD
IAKGLIVKLENVKCDIVGCRDNIIKYAYDVIVLNGFDTIGLEYLDSSNFERDRLPFPTAKSLMTYYGFEGKKYSEIDKSVFNTK
YYNFIFNENETIKDISYSVYGLKEIQKKRFKNLVIKAIGFADIKDKFVQLSNNTNMNVIFVPAAFTSQMDSNTHKIYVKEIMDK
NNKKQLQLIDKRKVRTKQEFHINGLNADFNAANNIKYIAENNDLLLTMCTKTKENNRYGNPLYNIKDTFKKKIPSSILNIFKKK
DMYQIICD (配列番号51)

>3300031994|Ga0310691_10000084_157|M
[哺乳動物消化系-ルーメン-ヒツジ(ovis aries)]
MFRIFAALKLTNMGHVRLQKREGEVYKTYKLKVKSFSGNVDIKAGIVEYDQKFNNVSQWIADHLTSMTIGEAASRISPHKMDSQ
YAMTSLSDEWKDQPLYKIFTRGFGGMNADNLIIECTKTEENCKYDKEKSLGFSESVFRTFGFAANASSDMKSRMTQAKVKIGRK
NIDEDSADDEKCLQAIYEIQKNELLTDDNWKDRIGYLEMKGDQERELERTTILYDYYRANRTTVLDKLDNLKVETLSKFRGSKR
KSDRKILTLNGISYDIKRKEGCQGFELKFSVDKNHMEFDLLGHRALIKNGEMLVDIENCHGSQLSLEIDGDDMYAIISMRTFCE
KNESKLEKIIGADVNIKHMFLMTSEKDDGNTKCYVNLYRELLSDSDFTDVLNKEEYEIFSELSKYVMFGLIETPYLGSRVIGTT
QHEKIVEDKITSGMKKIAIRLFQEGKVRERIYVQNVLKIRALLKALFSTKLAYSNEQKIYDNLMRFGEKDDRRKDEGFHTTCRG
TSLRSEMDMLSKKILACRDNIVEYGYYVIGLNGFDGISLENLESSTFMDVKISYPSCNSMLDHFKLKGKTIEEAENHETVGKFI
KKGYYVMTLVNGKINDINYSEKAVMLHKKNLLYDTVIKSTHFADVKDKFVELSNNGKVSVVIVPPYFSSQMDSVTHKVFTEEIV
VQKKSSNGKVRKTKKTVLVDKRKVRKTQESHINGLNADYNAALNLKYIAETIDWRSTLCFKTWNTYGSPQWDSKIKNQKTMIDR
LDSLGAIELKNW (配列番号52)

>3300031994|Ga0310691_10000270_20|M
[哺乳動物消化系-ルーメン-ヒツジ(ovis aries)]
MNKSYVFKSNVAIDDIMSLFEPAIEEYINYYNRTSDFICDNLTSMKIGDLANYIKNKENVYCKFVLNDDIKDLPLYKIFSLNLN
SSQKKNADNALYEAIKVLNADGYKGKNILGLGDTYFRRNGYVKNVISNYRTKFVTLKPNVKYSKIDINSVTEQLIKTQTIFEVV
NKKIESETDFENLITYFKNRETPNDEKIKRLELLFDYYTKHKNEINEEIEKHAVESLKSFNGCRRNGNRKTMTVQMQKMLLKKH
GLTSYILHLVLDKKPYDINLMGNRQTVKVDNNGNRVDLVDISSKHGYDLTFEVKGKTLFFTFSSEKDFSKKEQEIKNILGIDIN
TKHSMLATSITDNGKVKGYINIYVELLKNKDFVSTLNKEELAYYTEMAKFVSFGLLEIPSLFERVSNQYDKKNNVSITDETLLK
REIAISQTLDNLAKKYRDKNCKIASYIDYTKMLRSKYKSYFILKQKYYEKNHEYDDKMGFSDISTNSKETMDPRRFENPFINTD
IAKGLIVKLENVKCDIVGCRDNIIKYAYDVIVLNGFDTIGLEYLDSSNFERDRLPFPTAKSLMTYYGFEGKKYSEIDKSVFNTK
YYNFIFNENETIKDISYSVYGLKEIQKKRFKNLVIKAIGFADIKDKFVQLSNNTNMNVIFVPAAFTSQMDSNTHKIYVKEIMDK
NNKKQLQLIDKRKVRTKQEFHINGLNADFNAANNIKYIAENNDLLLTMCTKTKENNRYGNPLYNIKDTFKKKIPSSILNIFKKK
DMYQIICD (配列番号53)

>3300032030|Ga0310697_10001273_44|P
[哺乳動物消化系-ルーメン-ヒツジ(ovis aries)]
MAHKTNNGENTINKTFIFKAKCDNNDIISLWKPAMEEYCTYYNKLSQWICNNLTSMKVKDLFAYLDDKQKTKPCVDKKTGETKI
GVGYYRYFIENNKEDMPLYWLFTKNCSSSHADNLLFEFVRKVNHEEYNGNSLGMGETDYRRFGYFQNVISNFRTKMSSLKATTK
WKKFDVNDVDEDTLKNQTIYDVDKYGIESVNDFNERIDILKIREETEQTKDKIARLECLCKYYKEHEEDIKNEIATMAIADLQK

【０２４６】

【表２７】

FGGCQRKSMNTLTIHKQDSPMEKVGNTSFNLRLTFNKKPYTLNLLGNRQVVKFVGGKRIDLINITENHGDWITFNIKNNELFVH
MTSPVDFEKEVCEIKNAVGVDVNIKHMMLATSIVDDGNVKGYINLYRELVNNNDFIATFGNSKNGHQGLEIYEQMAENVNFGIL
ETESLFERVVNQSNGGELNNQLIRREIAMQKVFDNITKTNNDKNIVNYVNYVKMLRAKYKAYFILKEKYYEKQKEYDDMMGFND
ESTENKEMMDKRRFEFSFINTDTAQELLIKLNKVEQDLIGCRDNIVTYAFNVFKTNGYDTLAVEYLDSAQFDKAKMPTPKSLLK
YHKFEGKTIDEVKEMMNNKNFTNAYYNFKFENEIVKDIEYSTDGIWRQKKLNFMNLIIKAIHFADIKDKFVQLCNNNSMNVVFC
PSAFTSQMDSITHSLYYIEKTSKTKNGKEKKQYVLANKKMVRTQQEKHINGLNADFNSACNLKYIALDEELRNAMTDEFNPKKQ
KTMYGVPAYNIKNGFKKNLSTKTINTFRTLGHYRDGKINEDGVFVENLA (配列番号54)

>3300032030|Ga0310697_10005481_13|P
[哺乳動物消化系-ルーメン-ヒツジ(ovis aries)]
MAHKTNNGENTINKTFIFKAKCEKNDIISLWKPAAEEYCNYYNKLSKWIGDSLTTMKIGDLAQYITNQNSAYYLAVTNDSKKDL
PLYKIFQKGFSSQCADNALYSAIKAINPENYNGNSLEIGETDYRRFGYVQSVIGNFRTKMSSLKVSVKYKKFDVNDVDEETLKT
QTIYDVDKYGIESIKDFNEFIEVLKLREETPQLNEKITRLECLCGYYSKNEENIKNEIETMAISDLQKFGGCQRKSLNTLTIHK
QNSLMEKVGNTSFTLQLSFNKKPYTINLLGNRQVVKFVDGKRVDLIDITEKHGDWVTFNIKNDELFVHLTSPIDFEKEVCEIKN
AVGVDVNIKHNMLATSIKDDGNVKGYINLYKELVNDCDFISTCNEDEFDLYRQMSESVNFGILETDSLFERVVNQSKGGCLNNK
FIRRELAMQKVFDNITKTNKDQNIVDYVNYVKMLRAKYKAYFILKEKYYEKQKEYDIKMGFTDVSTESKETMDKRRMEFPFVNT
DTAKELLAKLNNIEQDLIGCRDNIVTYAFNIFKNNGYDTLAVEYLDSAQFDKRRMPTPTSLLKYHKFEGKTKDEVEDMMKSKKF
SNAYYTFKFENDVVSNIEYSNDGIWKQKQLNFGNLIIKAIHFADIKDKFVQLCNNNKMNIVFCPSAFTSQMDSITHTLYYVEKI
TKKKNGKEEKKYVLANKKMVRTQQETHINGLNADYNSACNLKYIALNDELRNEMTDTFKVTNRQKTMYGIPAYNIKRGFKKNLS
AKTINTFRKLGHYRDGKINEDGMFVETLA (配列番号55)

>OBLI01003123_14|M
[ブタ腸メタゲノム]
MAHKKNIGAEIVKTYSFKVKNTNGITMEKLMAAIDEYQSYYNLCSDWICKNLTTMTIGDLDRYIPEKSKDNIYATVLLDEVWKN
QPLYKIFGKKYSANNRNNALYCALSSVIDMNKENVLGFSKTHYVRNGYILNVISNYASKLSKLNTGVKSRAIKETSDEATIIEQ
VIYEMEHNKWESIEDWKNQIEYLNSKTDYNPTYMERMKTLSAYYSEHKSEIDAKMQEMAVENLVKFGGCRRNNSKKSMFIMGSN
HTNYTISYIGENCFNINFANILNFDVYGRRDVVKNGEVLVDIMANHGDSIVLKIVNGELYADVPCSVTLNKVESNFDKVVGIDV
NMKHMLLSTSVTDNGSLDFLNIYKEMSNNAEFMALCPEKDRKYYKDISQYVTFAPLELDLLFSRISKQDKVKMEKAYSEILEAL
KWKFFANGDNKNRIYVESIQKIRQQIKALCVIKNAYYEQQSAYDIDKTQEYIETHPFSLTEKGMSIKSKMDKICQTIIGCRNNI
IDYAYSFFERNGYTIIGLEKLTSSQFEKTKSMPTCKSLLNFHKVLGHTLSELETLPINDVVKKGYYAFTTDNEGRITDASLSEK
GKVRKMKDDFFNQAIKAIHFADVKDYFATLSNNGQTGIFFVPSQFTSQMDSNTHNLYFENAKNGGLKLASKSKVRKSQEYHLNG
LPADYNAARNIAYIGLDEIMRNTFLKKANSNKSLYNQPIYDTGIKKTAGVFSRMKKLKKYKVI (配列番号56)

【０２４７】

【表２８】

表73. CLUST.091979 エフェクターの保存配列。

配列	残基	位置
$PX_1X_2X_3X_4F$ (配列番号 216)	X_1 はL 又はM 又はI 又はC 又はF である X_2 はY 又はW 又はF である X_3 はK 又はT 又はC 又はR 又はW 又はY 又はH 又はV である X_4 はI 又はL 又はM である	N 末端
$RX_1X_2X_3L$ (配列番号 217)	X_1 はI 又はL 又はM 又はY 又はT 又はF である X_2 はR 又はQ 又はK 又はE 又はS 又はT である X_3 はL 又はI 又はT 又はC 又はM 又はK である	中間配列
NX_1YX_2 (配列番号 218)	X_1 はI 又はL 又はF である X_2 はK 又はR 又はV 又はE である	中間配列
$KX_1X_2X_3FAX_4X_5KD$ (配列番号 219)	X_1 はT 又はI 又はN 又はA 又はS 又はF 又はV である X_2 はI 又はV 又はL 又はS である X_3 はH 又はS 又はG 又はR である X_4 はD 又はS 又はE である X_5 はI 又はV 又はM 又はT 又はN である	C 末端
LX_1NX_2 (配列番号 220)	X_1 はG 又はS 又はC 又はT である X_2 はN 又はY 又はK 又はS である	C 末端
$PX_1X_2X_3X_4SQX_5DS$ (配列番号 221)	X_1 はS 又はP 又はA である X_2 はY 又はS 又はA 又はP 又はE 又はY 又はQ 又はN である X_3 はF 又はY 又はH である	C 末端

【０２４８】

【表２９】

	X_4 はＴ又はＳである X_5 はＭ又はＴ又はＩである	
$KX_1X_2VRX_3X_4QEX_5H$ (配列番号 222)	X_1 はＮ又はＫ又はＷ又はＲ又はＥ又はＴ又はＹである X_2 はＭ又はＲ又はＬ又はＳ又はＫ又はＶ又はＥ又はＴ又はＩ又はＤである X_3 はＬ又はＲ又はＨ又はＰ又はＴ又はＫ又はＰのＱ又はＳ又はＡである X_4 はＧ又はＱ又はＮ又はＲ又はＫ又はＥ又はＩ又はＴ又はＳ又はＣである X_5 はＲ又はＷ又はＹ又はＫ又はＴ又はＦ又はＳ又はＱである	Ｃ末端
$X_1NGX_2X_3X_4DX_5NX_6X_7X_8N$ (配列番号 223)	X_1 はＩ又はＫ又はＶ又はＬである X_2 はＬ又はＭである X_3 はＮ又はＨ又はＰである X_4 はＡ又はＳ又はＣである X_5 はＶ又はＹ又はＩ又はＦ又はＴ又はＮである X_6 はＡ又はＳである X_7 はＳ又はＡ又はＰである X_8 はＭ又はＣ又はＬ又はＲ又はＮ又はＳ又はＫ又はＬである	Ｃ末端

【０２４９】
これらのシステムのためのダイレクトリピート配列及びスペーサー長さの例を表８に示す。
【０２５０】

【表３０】

表84.代表的なCLUST.091979ダイレクトリピートのヌクレオチド配列及びスペーサー長さ

CLUST.091979エフェクタータンパク質アクセッション	ダイレクトリピートヌクレオチド配列	スペーサー長さ
AUXO013988882_8|P (配列番号1)	ACTATGTTGGAATACATTTTTATAGGTATTTACAACT (配列番号57) AGTTGTAAATACCTATAAAAATGTATTCCAACATAGT (配列番号118)	28 - 29
SRR094437_845781_4|M (配列番号2)	ATTGTTGGAATATCACTTTTGTAGGGTATTCACAAC (配列番号58) GTTGTGAATACCCTACAAAAGTGATATTCCAACAAT (配列番号119)	30 - 31
SRR1221442_316828_61|P (配列番号3)	AATGTTGTTCACCCTTTTT (配列番号59) AAAAAGGGTGAACAACATT (配列番号120)	47
SRR3181151_741875_3|M (配列番号4)	CCTGTTGTGAATACTCTTTTATAGGTATCAAACAAC (配列番号60)	26 - 30

【０２５１】

【表３１】

	GTTGTTTGATACCTATAAAAGAGTATTCACAACAGG (配列番号121)	
SRR5371369_1764679_7|P (配列番号5)	ATTGTTGTAACTCTTATTTTGTATGGAGTAAACAAC (配列番号61) GTTGTTTACTCCATACAAAATAAGAGTTACAACAAT (配列番号122)	30
SRR5371371_1138852_2|M (配列番号6)	ATTGTTGTAGACACCTTTTTATAAGGATTGAACAAC (配列番号62) GTTGTTCAATCCTTATAAAAAGGTGTCTACAACAAT (配列番号123)	29 - 43
SRR5371379_2478682_1|M (配列番号7)	CTTGTTGTATATACTCTTTTATAGGTATTAAACAAC (配列番号63) GTTGTTTAATACCTATAAAAGAGTATATACAACAAG (配列番号124)	29 - 38
SRR5371385_201181_1|P (配列番号8)	ATTGTTGTAACTCTTATTTTGTATGGAGTAAACAAC (配列番号61) GTTGTTTACTCCATACAAAATAAGAGTTACAACAAT (配列番号122)	25 - 30
SRR5371385_201181_1|M (配列番号9)	ATTGTTGTAACTCTTATTTTGTATGGAGTAAACAAC (配列番号61) GTTGTTTACTCCATACAAAATAAGAGTTACAACAAT (配列番号122)	25 - 30
SRR5371401_1055766_58|M (配列番号10)	CTTGTTGTATATGTCCTTTTATAGGTATT (配列番号64) AATACCTATAAAAGGACATATACAACAAG (配列番号125)	30 - 51
SRR5371439_988701_11|M (配列番号11)	CTTGTTGTATATACTCTTTTATAGGTATTAAACAAC (配列番号63) GTTGTTTAATACCTATAAAAGAGTATATACAACAAG (配列番号124)	29 - 30
SRR5371497_203858_6|M (配列番号12)	CTTGTTGTATATGTCTTTTTATAGGTATTGAACAAC (配列番号65) GTTGTTCAATACCTATAAAAAGACATATACAACAAG (配列番号126)	30
SRR5371501_2762794_1|M (配列番号13)	TACTCTTTTTTAGGTAATGAACAAC (配列番号66) GTTGTTCATTACCTAAAAAAGAGTA (配列番号127)	41
SRR5678926_1309611_3|P (配列番号14)	CTTGTTGTATATATTCTTTTATAGGTATTAAACAAC (配列番号67) GTTGTTTAATACCTATAAAAGAATATATACAACAAG (配列番号128)	24 - 37
SRR6059713_382107_4|P (配列番号15)	CTTGTTGTATATACTCTTTTATAGGTATTAAACAAC (配列番号63) GTTGTTTAATACCTATAAAAGAGTATATACAACAAG (配列番号124)	28 - 31
SRR6060192_2608084_13|P (配列番号16)	CATGTTGTACATACTATTTTTTAAGTATTAAACAAC (配列番号68) GTTGTTTAATACTTAAAAAATAGTATGTACAACATG (配列番号129)	27 - 42

【０２５２】

【表３２】

SRR7634052_1662339_24|M (配列番号17)	GATGTTGGACACTATGTTTTATACGGTGGATACAAC (配列番号69) GTTGTATCCACCGTATAAAACATAGTGTCCAACATC (配列番号130)	30
AUXO017332817_2|M (配列番号18)	GATGTTGTTATGCTGTTTTTGTAAGTAATAAACAAC (配列番号70) GTTGTTTATTACTTACAAAAACAGCATAACAACATC (配列番号131)	29 - 30
OQVL01000914_15|P (配列番号19)	ATTGTTGTAGACCTCTTTTTATAAGGATTGAACAAC (配列番号71) GTTGTTCAATCCTTATAAAAAGAGGTCTACAACAAT (配列番号132)	30
3300001598|EMG_10017415_6|P (配列番号20)	AATGTTGTTCACCCTTTTT (配列番号59) AAAAAGGGTGAACAACATT (配列番号120)	47

【０２５３】

【表３３】

3300021254|Ga0223824_10022219_2|P (配列番号21)	ATTGTTGTACGAACCATTTTATATGGTAATAACAAC (配列番号72) GTTGTTATTACCATATAAAATGGTTCGTACAACAAT (配列番号133)	29 - 30
3300021431|Ga0224423_10015012_2|P (配列番号22)	ACTGTAAAACCCCTGCAGATGAAAGGAAAGTACAACAGT (配列番号73) ACTGTTGTACTTTCCTTTCATCTGCAGGGGTTTTACAGT (配列番号134)	27 - 42
3300012973|Ga0123351_1009859_3|P (配列番号23)	ATCATGTTGTACATACTATTTTTTAAGTATTAAACAACTA (配列番号74) TAGTTGTTTAATACTTAAAAAATAGTATGTACAACATGAT (配列番号135)	26 - 29
3300012979|Ga0123348_10005323_4|M (配列番号24)	CTTGTTGTATATACTCTTTTATAGGTATTAAACAAC (配列番号63) GTTGTTTAATACCTATAAAAGAGTATATACAACAAG (配列番号124)	28 - 31
3300028797|Ga0265301_10000251_12|M (配列番号25)	ATTGTTGAATGGCTATGTTTGTATGCTATTTACAAC (配列番号75) GTTGTAAATAGCATACAAACATAGCCATTCAACAAT (配列番号136)	28 - 30
3300028797|Ga0265301_10000251_10|P (配列番号26)	ATTGTTGAATGGCTATGTTTGTATGCTATTTACAAC (配列番号75) GTTGTAAATAGCATACAAACATAGCCATTCAACAAT (配列番号136)	28 - 30
3300028797|Ga0265301_10009039_3|M (配列番号27)	ATTGTTGGGGTACTTCTTTTATAGGGTACTCACAAC (配列番号76) GTTGTGAGTACCCTATAAAAGAAGTACCCCAACAAT (配列番号137)	29 - 30
3300028887|Ga0265299_10000013_320|P (配列番号28)	ATTGTTGTAGACCTTGTGTTTTAGGGGTCTAACAACG (配列番号77) CGTTGTTAGACCCCTAAAACACAAGGTCTACAACAAT (配列番号138)	29 - 30
3300028887|Ga0265299_10000026_77|P (配列番号29)	ACTGTGTTGGAATACAATATGAGATGTATTTACAAC (配列番号78) GTTGTAAATACATCTCATATTGTATTCCAACACAGT (配列番号139)	30
3300028887|Ga0265299_10000133_30|M (配列番号30)	ATTGTTGTGGCATACCGCAAGGCGGATGCTGACAAC (配列番号79) GTTGTCAGCATCCGCCTTGCGGTATGCCACAACAAT (配列番号140)	26 - 34
3300028887|Ga0265299_10011526_3|M (配列番号31)	ATTGTTGGAATATCACTTTTGTAGGGTATTCACAAC (配列番号58) GTTGTGAATACCCTACAAAAGTGATATTCCAACAAT (配列番号119)	30 - 31
3300028887|Ga0265299_10012919_3|P (配列番号32)	AATTGTTGAGATACCGTTTTTTATGGTATTGGCAAC (配列番号80) GTTGCCAATACCATAAAAAACGGTATCTCAACAATT (配列番号141)	28 - 43
3300028914|Ga0265300_10009460_3|M (配列番号33)	ATTGTTGTGGCATACCGTATTACGGGTGCTGACAA (配列番号81)	31

【０２５４】

【表３４】

	TTGTCAGCACCCGTAATACGGTATGCCACAACAAT (配列番号142)	
3300031853|Ga0326514_10013355_6|M (配列番号34)	GATGTTGTTATGCTGTTTTTGTAAGTAATAAACAAC (配列番号70) GTTGTTTATTACTTACAAAAACAGCATAACAACATC (配列番号131)	28 - 30
3300031993|Ga0310696_10000014_323|P (配列番号35)	ATTGTTGTAGACCTTGTGTTTTAGGGGTCTAACAACG (配列番号77) CGTTGTTAGACCCCTAAAACACAAGGTCTACAACAAT (配列番号138)	29 - 30
3300031993|Ga0310696_10000226_76|P (配列番号36)	ACTGTGTTGGAATACAATATGAGATGTATTTACAAC (配列番号78) GTTGTAAATACATCTCATATTGTATTCCAACACAGT (配列番号139)	30
3300031993|Ga0310696_10000447_27|M (配列番号37)	ATTGTTGTGGCATACCGCAAGGCGGATGCTGACAAC (配列番号79) GTTGTCAGCATCCGCCTTGCGGTATGCCACAACAAT (配列番号140)	26 - 34
3300031993|Ga0310696_10026614_2|M (配列番号38)	AATTGTTGAGATACCGTTTTTTATGGTATTGGCAAC (配列番号80) GTTGCCAATACCATAAAAAACGGTATCTCAACAATT (配列番号141)	30
3300031993|Ga0310696_10030100_3|M (配列番号39)	ATTGTTGGAATATCACTTTTGTAGGGTATTCACAAC (配列番号58) GTTGTGAATACCCTACAAAAGTGATATTCCAACAAT (配列番号119)	30 - 31
3300031998|Ga0310786_10000003_467|M (配列番号40)	ATTGTTGTGGCATACCGTATTACGGGTGCTGACAAC (配列番号82) GTTGTCAGCACCCGTAATACGGTATGCCACAACAAT (配列番号143)	25 - 31
AUXO013988882|Ga0247611_10000101_23|P (配列番号41)	ATTGTGTTGGGATACACTTTTATAGGTATTTACAAC (配列番号83) GTTGTAAATACCTATAAAAGTGTATCCCAACACAAT (配列番号144)	29 - 31
3300028805|Ga0247608_10000186_37|P (配列番号42)	TATTGTTGAATACCTTTCTTATAAAGGTAATTACAAC (配列番号84) GTTGTAATTACCTTTATAAGAAAGGTATTCAACAATA (配列番号145)	29 - 46
3300028805|Ga0247608_10000895_42|M (配列番号43)	TGTTGTAAATGGCTTTTTATGGGCAACGAACAACTC (配列番号85) GAGTTGTTCGTTGCCCATAAAAAGCCATTTACAACA (配列番号146)	28 - 45
3300028805|Ga0247608_10006074_1|M (配列番号44)	ATTGTTGAATGTATTCTTTTTTAGGACAGATACAAC (配列番号86) GTTGTATCTGTCCTAAAAAAGAATACATTCAACAAT (配列番号147)	28 - 30
3300028833|Ga0247610_10000007_379|M (配列番号45)	GATGTTGGACACTATGTTTTATACGGTGGATACAAC (配列番号69)	30

【０２５５】

【表３５】

	GTTGTATCCACCGTATAAAACATAGTGTCCAACATC (配列番号130)	
3300028833|Ga0247610_10004486_2|M (配列番号46)	TATTGTTGAATACCTTTCTTATAAAGGTAATTACAAC (配列番号84) GTTGTAATTACCTTTATAAGAAAGGTATTCAACAATA (配列番号145)	29 - 46
3300028888|Ga0247609_10000668_74|M (配列番号47)	ATTGTTGAATGGTATCTTTTATAGACTGATTACAACT (配列番号87) AGTTGTAATCAGTCTATAAAAGATACCATTCAACAAT (配列番号148)	29 - 41
3300028888|Ga0247609_10003329_9|M (配列番号48)	ATTGTTGGATAATAGGTTTTTTATCTTAATTACAAC (配列番号88) GTTGTAATTAAGATAAAAAACCTATTATCCAACAAT (配列番号149)	29 - 30
3300028888|Ga0247609_10016480_8|M (配列番号49)	ACTGTTGAATAGTTGATTTTATATCCTATTTACAAC (配列番号89) GTTGTAAATAGGATATAAAATCAACTATTCAACAGT (配列番号150)	29 - 30
3300031992|Ga0310694_10000010_351|M (配列番号50)	GATGTTGGACACTATGTTTTATACGGTGGATACAAC (配列番号69) GTTGTATCCACCGTATAAAACATAGTGTCCAACATC (配列番号130)	30
3300031992|Ga0310694_10022272_2|M (配列番号51)	TATTGTTGAATACCTTTCTTATAAAGGTAATTACAAC (配列番号84) GTTGTAATTACCTTTATAAGAAAGGTATTCAACAATA (配列番号145)	29 - 46
3300031994|Ga0310691_10000084_157|M (配列番号52)	TGTTGTAAATGGCTTTTTATGGGCAACGAACAACTC (配列番号85) GAGTTGTTCGTTGCCCATAAAAAGCCATTTACAACA (配列番号146)	28 - 45
3300031994|Ga0310691_10000270_20|M (配列番号53)	TATTGTTGAATACCTTTCTTATAAAGGTAATTACAAC (配列番号84) GTTGTAATTACCTTTATAAGAAAGGTATTCAACAATA (配列番号145)	29 - 46
3300032030|Ga0310697_10001273_44|P (配列番号54)	ACTGTTGAATAGTTGATTTTATATCCTATTTACAAC (配列番号89) GTTGTAAATAGGATATAAAATCAACTATTCAACAGT (配列番号150)	29 - 30
3300032030|Ga0310697_10005481_13|P (配列番号55)	ATTGTTGGATAATAGGTTTTTTATCTTAATTACAAC (配列番号88) GTTGTAATTAAGATAAAAAACCTATTATCCAACAAT (配列番号149)	29 - 30
OBLI01003123_14|M (配列番号56)	ATTGTTGTAGATACCTTTTTGTAAGGATTGAACAAC (配列番号90) GTTGTTCAATCCTTACAAAAAGGTATCTACAACAAT (配列番号151)	30

【０２５６】

実施例２－トランス活性化型ＲＮＡエレメントの同定

エフェクタータンパク質及びｃｒＲＮＡに加え、本明細書に記載される一部のＣＲＩＳＰＲシステムは、トランス活性化型ＲＮＡ（ｔｒａｃｒＲＮＡ）と呼ばれるロバストな酵素活性を活性化させる追加のスモールＲＮＡも含み得る。このようなｔｒａｃｒＲＮＡは、典型的には、ｃｒＲＮＡにハイブリダイズする相補的領域を含む。ｃｒＲＮＡ－ｔｒａｃｒＲＮＡハイブリッドはエフェクターとの複合体を形成し、プログラム可能な酵素活性の活性化をもたらす。

・ｔｒａｃｒＲＮＡ配列は、ｃｒＲＮＡのダイレクトリピート部分と相同のショート配列

モチーフのためのＣＲＩＳＰＲアレイをフランキングするゲノム配列を検索することによって同定することができる。検索方法には、完全ダイレクトリピート（ＤＲ）又はＤＲ部分配列についての完全又は縮重配列マッチングが含まれる。例えば、長さｎ個のヌクレオチドのＤＲは、重複する６～１０ｎｔ　ｋｍｅｒのセットに分解することができる。これらのｋｍｅｒを、ＣＲＩＳＰＲ遺伝子座をフランキングする配列とアラインメントすることができ、１以上のｋｍｅｒアラインメントと相同の領域を、ｔｒａｃｒＲＮＡとしての実験的検証のためのＤＲ相同性領域として同定することができる。或いは、ＲＮＡ同時フォールド自由エネルギーを完全ＤＲ又はＤＲ部分配列及びＣＲＩＳＰＲシステムのエレメントをフランキングするゲノム配列からのショートｋｍｅｒ配列について計算することができる。低い最小自由エネルギー構造を有するフランキング配列エレメントを、ｔｒａｃｒＲＮＡとしての実験的検証のためのＤＲ相同性領域として同定することができる。

・ｔｒａｃｒＲＮＡエレメントは、ＣＲＩＳＰＲ関連遺伝子又はＣＲＩＳＰＲアレイにごく近接して高頻度で生じる。ｔｒａｃｒＲＮＡエレメントを同定するためにＤＲ相同性領域の検索の代替策として、ＣＲＩＳＰＲエフェクター又はＣＲＩＳＰＲアレイをフランキングする非コード配列を、ｔｒａｃｒＲＮＡの直接実験的検証のためのクローニング又は遺伝子合成によって単離することができる。

・ｔｒａｃｒＲＮＡエレメントの実験的検証は、非天然種において異種発現されるＣＲＩＳＰＲシステム又は合成配列のための宿主生物のスモールＲＮＡシーケンシングを使用して実施することができる。由来ゲノム遺伝子座からのスモールＲＮＡ配列のアラインメントを使用してＤＲ相同性領域を含む発現ＲＮＡ産物及び完全ｔｒａｃｒＲＮＡエレメントに典型的な固定化プロセシングを同定することができる。

・ＲＮＡシーケンシングによって同定された完全ｔｒａｃｒＲＮＡ候補は、ｃｒＲＮＡ及びエフェクターをｔｒａｃｒＲＮＡ候補と組み合わせて又は組み合わせずに発現させ、エフェクター酵素活性の活性化をモニタリングすることによってインビトロ又はインビボで検証することができる。

・エンジニアリングされた構築物において、ｔｒａｃｒＲＮＡの発現は、哺乳動物細胞における発現のためのＵ６、Ｕ１、及びＨ１プロモーター又は細菌における発現のためのＪ２３１１９プロモーターを含むがこれらに限定されないプロモーターによってドライブすることができる。

・一部の例において、ｔｒａｃｒＲＮＡはｃｒＲＮＡと融合させ、シングルＲＮＡガイドとして発現させることができる。

・システムは、表９に列記される非コード配列内に含まれるｔｒａｃｒＲＮＡを含み得る。例えば、一部の実施形態において、システムは、配列番号１５２～２０４のいずれか１つに記載されるｔｒａｃｒＲＮＡを含む。

【０２５７】

【表３６】

表95.代表的なCLUST.091979システムの非コード配列

>3300028887|Ga0265299_10012919_3|P
TATATCGTGGCCGAATATGTTAACGCGGACGACGTCCGTCTTGTGAAGTTTCAGGACGAGGATTTCGACAGGCTTCTTGACAAG
GTTAGAGAATGGAACAAGAAACATCTTGTTGTTGGAAATCGGAACTTCGAAGAAAAATTTGCGTAATCCAAAAATTTTCCGTAT
ATTTGCGGCGTGAAATTAAAAATATGTTTAACTAAAAACAAAGATTATGGCACACAAGAATCCTGATGGGGAGAACACCATCAA
CAAAACTTTTATTTTCAAAGTGAAATGCGAGAAGAATGATATTATATCGTTCTGGAAACCCGCAGCTGAAGAGTATTGCAACTA
TTACAACAAACTTAGCGAATGGATTGGCAAAGATATGTATAACACGCCGTCATGGAACATCCGGCAAGAGTTCAAGAAGAATTT
AAGTGTTAGAACCATAAACACGTTTCGTGAGCTTGGCAATGTGAAATACGGCAAAATCAACAATGAAGGGCTTTTTGTCGAAGA
CGATGTGTAAACATTAAGATTTCCATACGACAGGATTCAAAAAAACGTTCTTTGAAATATTGGATTGGTGGCAAGAGGCTGTTT
TTTTTAGGCTAAAAAGTTGTGTAAATAGCAGAAACACAGAACATAACATAAAATCT (配列番号91)

>3300028797|Ga0265301_10000251_12|M
AACTGCTACAATTCTGCCGAGTTTATGATTCAGACAAAATTCAAAAAAAGACTTCCGCAAGCAACCGTTTTTGGTGAATTGAAC
AGAAACGGGTATGTTAAAGTATTGACCCAAGAAGAATATGACGAACTCACAAAATCAGCAAAATAATTTATTACTGATTGAAAA
ATAAAGCGTTCTTTGACATATTGTATAACAAACAAGCATTTTTGTAAGAGATAACCCATTTCATTTTATTGATATACAATGAAA
TGAAAAGAATAT (配列番号92)

>SRR094437_845781_4|M
GATAAATTTGCCCGTAATGTTATCGGGTTCAAGTCATATCACGAACTGCTTGATAATGCTATCATAAAAGAAAAATTACAACGG
GAATTTGGTTATGAAGATGCTCCGAAAACGTGGTTGTTCGGACAACAAAAAAATGAATGTTTCTAATGTATTAAAACAATAATT
CAATTACAATTTTAAGATTATGGCACAACACAAATCAAACAACGAAGAATCAGCAATCAACAAGACTTTCATTTTCAAGGCAAA
ATGCGATAAGAACGATGTCATATCGTTATGGGAACCAGCGGCAAAGGAATACTGCGACTATTATAACAAAGTGAGCAAGTGGAT
TAAAACTATGTATAACATACCCGCATATAACATTAAGTCCAATTTCAAGAAAAATTTGAGCGCCAAAACAATTCAAACTTTTAG
AGAACTTGGACACTACCGTGACGGAAAAATAAATGAGGATGGTATGTTTGTTGAAAACTTGGAATAATTCTGTATATACCAATT
AGAATTGAAAAAAAAAACGCTCTTTGACATATTGTTTTCTACATAAAAACAAGATTTTACACAACGCAATACATCATAAAGTGTT
GCGTTATAACAAATAACAAAAATTCT (配列番号93)

>3300021254|Ga0223824_10022219_2|P
TTTATTCAATGCGAACCAGAGGTCTTGACGCATGAATCTGGCTATACATATCGTTATGCGACCGACGAAGAGAAAATATTGATT
AAAAGATGCAAATATTGAATAGGCAATTTTAAATTGTGAAAAAAAAAATGATTGAATATAAGTTTACGTTTGAACTGGATGGAC
ATCTATCGGCGTACGATTTTGTTACGTTGCAAGAACGGTTTGAAAGGGAATTGAATCCTTATTTTGATGATGGGAGCATATCTG
GTACTCTTTCTTATGCAAATGATGATTAATATGCAAATAATATGGCACATGTAAGAACAAAAAATGAAGGAAACATGGCAAAAA
CATATTCTTTTAAGGTCAGAGAAACAAACCTTAAAAAGGATGTGATGATTGAATATAACGAATATTATAACAGGTTATCCGATT
GGATATGTGGCAATTTAACCAAAATCTCGGAAAATGAAGAATGGAGGAATGCCTTATGCAAACCAACAGAAAACATGTACAACG
AACCGATTTACGTTCCCTTGGTTAAATCACAGAACGGAATGTTCAAGGCAATTAAAAAATTGGGCGCAACGAAGATATGGCAAG
AATAGAAAGACCGATTTTTAAATCTGAAATCACTTCTAACGAATTGTATACTAAAGAAATATAAAGAATATACATCTTTTATGA
CATTATGATATTGTTGTATGCATCATTTCACATGGTAATAACAACGAAGAGAAACACCGAGCGACCCACAAACCTATTGTCGTA
CGCATCATTTCACATGATAATAACAACGAATATTCCTGCAAGCATGATTTAACAATTTTTAAGAACCTGGTGGTTTCTCCGTTG
GGTTCTTTTTAGTATCTTTGCCTTGTTGAAACAAATAAAACAAATTGAATTATGATTTATAAAGGCAAAGAAATAGACGAAAGT
TACCACATCAATAAATGGGAAGATGAAGAGATTTACTCTGGTCCAACCCATTATGAATCATTCGAAGCCGATGAAATAAAAGAG
TTCTACCTCAAGGCACTTGCAAAGGAAAAGGAA (配列番号94)

>AUXO017332817_2|M
GTGCGCATATACACTCAATTCGCCGATGACCGTGTGTACGCGAAGGATTGTATCGACGGATTCTTTAGTATAAGACAAGATACC
GAAATGCGCCTCGTGTATAAAAATGAGATAGCACGCGGGCTTGAGTGTATCAATATTGTAAGATAGTAGTTTTCTGTTATTTTA
CATATTGATGTGTTTTGGCATGGTTTTTGTTAAAATATAATCTAGCAGTATTGAGACTGCGGAGTAACGTGTCTAACTGTTTCA
TTATAAGCAGTAAAGACTAATATTTTTATATCTTAAACTTATTTTTATTATGGCTGGTCACAGCAAAATCAAAGAAAATCACAT
TATGAAGGCGTTTCTTATGAAAGTAAAAGAAACGCGAAAAAAAACAGTGGCAATCAAATTTTATTAGAAGTGAGATTGCTAAGTT
TACAAATTATTACAATGGGCTGTCAAAGTTCCTTCTTGGAAGCCCGACTGGAGGGACATATGACACTGCATATTTTGATACAAA
GATTCAAGGCTCCAAGGGGGTATATGATAAGATTAAAGAAAACGGAGAAACTTATATTGCAGTATTAAGTGATGACGTTATTAC
GGCAGAGGTGTAAAATCCTCTGCCAACATCGCAAGTAACTCATTGAAAATTAGTTAAATGCGAATGCCAACAAAAGTGAACGAA
CTGACTTGTAAAGCAGGATGTTGTTATATCTTTTTGTAGATAATAAGCAACAAGATACAATCAATCGCGAGTTTATACTGAAAT
GTTGTTACACTGTTTTTGTAAGTGTTAAACAACCTTGCACAAATGTCATCTACCAGTACAATAGATGTTGTTATACTGTTTTGT
AGGTATTAAACAACCATTGCGCAGACTGACAGAGTAACCTTTCCTGATATGTTGTTACACATTTTTGTAAGTGTTAAACAACTG
ACGCATTGATATTGCCTTGTCTATTAAGAATGTTGTTATGCTCTTTTTATTGGTATAAACAACCGAGCAACTGGTACTCAAATT

【０２５８】

【表３７】

TTAAATACTGTCGCGCTATGTTATGTACATCGAACAGCTACCACTCAATGGCTTTGTTTGCAACCGTGATTAATTCAATCGCGG
TTGCATTTGTTTTATGATGTGTTTTTGTATATATTATGTATATATGGAAAAGGAAAACAGGGTATCGGAGTTATGGAGCAAGTT
CTCTGATATTGACTTGCGCCGAAGCCAAATGACATATATGCCAATAAGAGGTAGTAAAAGATACGGCAGAAGAATAAAACGTAG
TGACATCGAGTACGAGTACAGATATCTGTATAGAGCAAACAAACATTGGTAATATGACCGTAGCTAAATTATCAAGTAATCATA
AGCCAGCGTGCCTTGGACGAATCTCAGCTTTAAACACCCCGATTAGATTTGAGTGTCGGGCTGGTAATAGTATAAGGCCTGGCA
ACATAGAGTATAGCTATAAAAGATGGAAAACGTCGTAATTTCAACTATGCACAACCCGCATACGCTGGCTTATTACCAAGGTAA
GCTGGCTCCTATGCATTTCAGACAAGATACAGG (配列番号95)

>3300021431|Ga0224423_10015012_2|P
AGCCTGTATACAGGGACAAGGTTAAGTACAACACCAAGGCTGAGGCAAAGAAGAGGGCTGATGATATGAACAAACAGAATAGGG
TCATACACCAGCTGTCTGTTTATTTGTGTCCTAAATGTCATAAGTGGCATATAGGTAGGAGCAGTGTGGAGAGTGTGCGCAGGG
AAGGGTACTTTAGTCAGATTTGAAATTAATTGTTATATGGCGCATAGAAATAAAAACCTAGCAGAAAACTGCATTAACAAAACA
TTCAGTTTTAAAGTCAAAGCCGAAAAAGAGGAGATAAATTCAAAATGGATTCCAGCCATTAAAGAATATACTGCTTATTATAAC
AGGATAAGTGACTGGATAAACCTGTATTCACAGCCTACTTATGATATTAAGGAAGTTTATAAGAAAAACGCTGGTTGCAAAGTG
ATAAACGACTTCATTAAAAACGGTAACGCCGTTATATGTTGTATCGAAAATAACAAACTAATTGAGACAAATGGAAGACAATAG
TTCAAATTTTAAATGTAAAACAGTCATTAATGTATTAATATATAATACATAGCAAAAATCCAGATGTTGAATACATTTCTTTTA
AGTGTACTTACAACGCGGTGGCATTGCTAAAATATAGTCCTGTGGATGTTGAATACATTTCTTTTAAGTGTACTTACAACCAAC
GCTGTACACATTGCTAATGGATGATGACGATATAGAGGTGTTGAACTACCTTAATGAAAACTACACCAATGAAAACATTGAGTA
TATACGCGGTTGGTGGATGGATGACGACGATAAACTCCAGACACTTGACAGGTTTTTGAAAAATTTTTCAATATAGACCTGTCA
CTGTTGCGGCTATAAGAAGACCGATTTGACACTGAAAGACCGATACTGGGTTTGCCCCGAATGCGGTGCAAAACTAGACCGCGA
TACCAATGCAGGAATAAACATTAAGAATGAGACAATTAGACTGATAAACAAAGAATAATGAGAACTATAATAGGGAGGTGTACC
CCCGAATTTAAGCCAGTGGAGAACCATACAAACCTATCATATAGGGGTTCAATGAATCTGGAATTTCTGACAAAAACAGGGTTT
AACAGCCAGTGTACCAATGACTAACACAGGACATATAAAGACAAATCTAACAATAAAAAAAAATATTGACCAATTCTGCAGAAA
AAACAGGTTGGTTTCGGTTATGTTGGTGAATAAAGACAGTTAGATTAATTTTATATGGAAATGAAAATAGAGACAAAAGACGAG
AACATCTACGTATTCATCTATGCCAAGTCCGCCTACTTCGGCAATACATTTGAATATGGCGGCACATTTTCCGTCGGCAAGGAC
GACAACTGGAACGATGTGAGAGGCCACGTTACCGAA (配列番号96)

>AUXO013988882|Ga0247611_10000101_23|P
GACAACATCCTGGTCAAGACCGAGGTTAACAGAAGGTACTGCCGCCTTATGACCGACGAGAACGGAGTGTGGCTCCTGAGGAAA
AACGACAAACATCCAACATATTTTATCTACCAGAACGGAACACTCTATCAATATGAGGAAGATTGATTAGTTGATGTTTTCATA
ATAATTTTATCTGGAATTTGAAAAGATTCCAGATTTTTTTTTTATTTCGACTGTACAAAAAACAGGTTCCGTTGCGTTATATAG
GTGTAAATTAAAAATTCAGTCAAACAAAAATTGGAATAAAATATGGCTAACAAGAGAACAGACACAACAATCAACCTTAACAAA
ACCGTTATAATGTTAACGAACATGCTGCCAGAAGTACGGGCAATGTTTCAGGCGGGAATACGCCAGGCTCAAGTTTATGCAGAC
TTGGTGAACAAGTGGATATGTTCACAGGAAATGAGAGAGGTTATGTGTCTCCATCCGTCAAAAAAGGACGGGGTGTACGACCAA
CCGTTCCTGAAAGCTACAACCAAATACCCAGCCACGGTAGCTGGTATCCTGCTTAAGATGGGAAAAACAACCAATTGGGGTGAG
AAATAATACCCACCCGCCCCATTTTTTTACACTGATTAGTTCTTTGACTTATTGATTTATATTGGTTTACACAAATTATCGACA
CAATAAATAAAAAAAAATTGTATATTAGTAGTATGATGACAGAAGAAACACGGAAGACAATAGAGAGCGTCATAGTGGTTCTCGG
CATAGCAATCATGCTGGCAGCCGCCGTCCGAATAATGACGCAGAACAAAGCAATTGTGAAATATGATGAACAGGTTGAAACCAT
GCAAACTTGCATA (配列番号97)

>AUXO013988882_8|P
ATGGAAGTTGTACGTGGTGGAAATCAATGGGAGGTTTATGACAATTACGATGAGACTATGAAAGCATCAAAAAATGTAAGGTCT
GTATTGGGACTTCCGGAAGTAAAATATCCACCTGAGGATTTTAGGACATATAATTTCTAATAAAAATGAACGGAAAAATTTCCG
TTCATTTTTTTTTTGTTTATTGGTGAAAAAATAGTATCTTTGTAAAAAATAAATGTTAAAATATTTTTTATGGGAAATACTACA
AAAAAAGGAAATTTGACGAAGACTTATTTATTCAAAGCCAATCTTTCAGAACAAGACTTTAAATTATGGAGGTCTATTGTTGAA
GAGTATCAAAGATATAAGGAAGTGTTGAGTAAATGGGTATGTGACCATCTTAGAAATGCAATGTGTACGAACCCGAAAAGTGAG
ACTGGATATTCTGTACCGTTCTTGACTTCAAGAATCAAGAAACAGAACATTATGGTTGTAGAATTGAAAAAAATGGGCATGGTT
GAAGTCTTGAATGAAAAATCAACAGAAATTTAAGAAAAAAATATTTATATAATGTACTGAAAATAAGTAAATAATAAATATTGT
GTAAAAAACTTGATATTTTTTTTTTGTTATCTTTATAATATAAAATAAAATGTAAATATGAAAAATCTGTTAAAACTCAAAGAA
CAAATCAAGGATTACAAACATCTTCAGTTTGTGTTGGAGAAAGAAGATGAATCTGAACTCCATTATAGATGTATGACTGAAGAT
TTTTCGTTCAAGGTATCTGAAGAAAAAGACGGAACACTT (配列番号98)

>SRR3181151_741875_3|M
TTATAAACATCTAAAAAGAAAGACTTATGACAACAAAACAAGTTAAATCAATCGTTTTAAAAGTAAAAAACACTAATGAATGCC
CTATTACAAAAGATGTAATAAATGAATATAAAAAATATTATAATATATGTAGTGAATGGATTAAAGATAATCTAACAAGTATTA
CTATTGGAAACGAAAATTTACGAAAATTATTTTGTGGTAAACTTAAAGTAAGTGGATATAATACACCAATATTAGACGCAACAA
AAAAAGGTCAATTTAATATATTGGCAGAATTAAAAAAACAGAATAAAATTAAAATATTTGAAATAGAAAAATAAGTCTTATGAT
TACAAAAATAATAGATTTCAAACATTTTTTTTAATTCTATTTTATTGACTAATTCATTGAAATATAAATAATTACAAATAACCC
(配列番号99)

>3300028805|Ga0247608_10000186_37|P
GATAGATATAGTATTGCAGCATTTCTGGCTTGCGAATCATCAGCAATGCAAAAATGTGACTATTGGAACAATGATGATGCCCAA

【０２５９】

【表３８】

GATTACATAAGAAACTACAAAGAGGCTTATAGTAATGCAGTAAGACTTGCGTTTTTTAATGATTAAGCAACACGCTTAACATTG
TCAAATGTAACGACATTAAGTGCGTGTTTCATAAGGGCAGCGAACCTTTCGCCGCCCTTCTTTTTTTGTTGCTGTAACGGAATT
ATGTTTACTTTTGTGCCATCAAGTATATAGTTCCCTTAATAAATTGTATATTAATTAAAAGTTTGGCACAATATTTGATGCGTA
CAAATTAAAATAAAAACATTTTGAATTTTAAAATTTAATTTGTAATTTTAAATAAGAAAGTTTTATTTAACTAAAATAAAAAAA
ATGAATAAATCTTATGTTTTTAAGTCGAATGTGGCTATTGATGACATTATGTCTTTATTTGAACCGGCAATTGAAGAGTACATA
AACTATTACAATAGAACCAGCGATTTCATTTGTGATAATCTTACATCAATGAAAATCGGAGATTTGTTGCTTCTAACAATGTGT
ACTAAGACAAAAGAAAATAATAGATACGGTAACCCCCTCTATAATATCAAAGATACTTTTAAAAAGAAAATACCATCTTCAATA
CTTAATATATTCAAAAAAAAGGATATGTATCAAATAATATGTGATTAATTATGCCTTTTTTTAATAAAAAATTGTTAAATAATA
CTTTGTTTATTAATAAATTATAAATATCACAGTAAACTATTAGGGATTTGTAAAATTTATGGAAATTATATACATGATGGCACT
AAGATTTGGTTATTAAGAAATTTTTCTGTATAAGTATAATAACCTATTTATAATTATAATTGAATAAAATGTATAATATGGAAA
ACACAGGCTTTTATACAGTTTCAAATATTGAAACTTCTCATAAGCCAACCGAAAATTCTAATGACGAAATTCTTAGGATTTTCA
ATAAAAGAAGGCCTTATTGCCCTTCAGACTTTAAGAAGCAACATTTTATT (配列番号100)

>3300028833|Ga0247610_10000007_379|M
AGGCTCAACCTCCTCAACCCGATTTATCTTGAGATCGCCAAGTACGGACACTTCGGGAGGAAGAGCTATGTGAAGGACGGCATC
AAGTACTTCCCGTGGGAGGATTTGGATTTGGTTGAAGACATCAGAAAAATTTTCGAAATGGAATAGAGGGAACCGGAATTTTTT
CCGGTTTTTCTTTGTCCTTTCGAAAATAAATAGTATCTTTGTAAAAAAACAACAGATTATGTACAATAGTAAGAAGAAGGGGGA
GGGTGACATTCAGAAGTCGTTCAAGTTCAAGGTCAAAACGGACAAGGAGACGGTCGAATTATTCAGAAAGGCCGCAGTCGAATA
CTCGGAATACTACAAGAGGCTGACAACATTCCTCTGTGAGATGTATAACAGACCAGCGTTTGACTTGAAGGAGTGCTACAAGAA
AAATTCCAATGTAAGTGTCTTCAACACATTGAAGAAAACTCTCGGTGCAATATATGGAAAGCTCGATGAAAACGGAAATTTTAT
TGAGAATGAATGTAATAAGTAACTGGAATAAAAGAAATTAGACAGAGTAA (配列番号101)

>3300028887|Ga0265299_10011526_3|M
TTGTATTGGTTGCTGTATGGCGACGGAAGTGACATATATGATGACGGGTGGTTTGACTGTGTTCATAATTTTGCCCGTAATGTT
ATCGGGTTTCAGTCATATCACGAACTGCTTGATAATGCTATTATAAAAGAAAAATTACAACGGTAATTTGGTTATGAAGATGCT
CCGAAAACGTGGTTGTTCGGACAACAAAAAAATGAATGTTTCTAATGTATTAAAACAATAATTCAATTACAATTTTAAGATTAT
GGCACAACACAAATCAAACAACGAAGAATCAGCAATCAACAAGACTTTCATTTTCAAGGCAAAATGCGAGAAGAACGATGTCAT
ATCGTTATGGGAACCAGCAGCAAAGGAATACGGCGACTATTATAACAAAGTGAGCAAGTGGATTAAAACTATGTATAACATACC
CGCATATAACATTAAGTCCAATTTCAAGAAAAATTTGAGCGCCAAAACAATTCAAACTTTTAGAGAACTTGGACACTACCGTGA
CGGAAAAATAAATGAGGATGGTATGTTTGTTGAAATTTTGGAATAATTCTGTATATACCAATTAGAATTGAAAAAAAAACGCTC
TTTGACATATTGTTTTCTACATAAAAACAAGATTTTACACAACGCAATACATCATAAAGTGTTGCGTTATAACAAATAACAAAA
ATTCTGGACGGGAAAGGAAGATGTCAGACGTTTTTATTGTTGGAATACTCGTTTTTTACGGTATTTACAACTGCCCCGTAGCGG
AATCAAAATACCACCGCATTGTTGGAGTACAAGTTTTACACGGTATTCACAGTACGAACACCGAATGAACTGAAAAAAATAAAC
CCGACCTTGCAACCGTAGATATAAATAAAGCAATACAAAATTTGAAACTATGGCACACATTAAAAAAATTGACGAAATGGCAAG
TCAAACTGTTTCACTCCGTTCTGACGCATTGTTCAAAAAAGCGTTTGAGGAATTTGAAAAGGAGTTGAAAGAAGTTCTCAAATC
GCACAACAATATCATTTATTGTGGAGGTGAT (配列番号102)

>3300028797|Ga0265301_10009039_3|M
CTCATCAAATTGTACAAGTCGTTGACGGACACTGAATTTGACAAGAAGAAAAATCATCAATGATGTCTACGACGGCACTTTTGAG
ATAATCCTCAAATACCCAAAGAAGAAGAACGGGACATTCGTGTTCTGGAAACATTACAAGAAGTAACACAATGATACACAGTAT
GTTGTAAGAAATAAGATTTAGGCTTTAATTTTAATATATGAAAATATGGCACACAAAGGAGAAAAGGAAGGCTACCAAATCAAG
ACACTGAAGTTCAAGGTACGCTCGCATGACATCGGGAAATCACTTTATGATATTGTCAACGAATACACCAACTACTATAACAAA
GTAAGCAAATGGATATGTGACAACCTTGGTTACAACGAGCCATTCTACAAGTCAAGGGTGAAAAGCGCCGCCTCCATGATGTCA
GGATTGAAAAAACTGGGCGCCACCATGCCATTGACGGATGAAAATGCCATTTTTTCAACACCAAAACCGAAGAAAAACATTGGA
AAACAATAATTTACACAAAGTCTACGGCGGGAATCGTGATAAAAATGAACGAGATTGTTGGGATATACCTTTTATAGGATTTTC
ACAACATCTGAGTTGTTTGATGTTAAAAACTTTAACTAATAAGGCAAGAAGTCCCATTCCTTCAGGTGGGGGTAGTTCATTTGT
TGGGATACTCGTTTCACACGGTATTCACAACTTCCAACCAACCATTAAAAAACCTTCAAATATTGTTGGAGTACCCGTTTTATA
CGGTGCAAAGCCTCCCCGACGATTTCAAGTTCCTGTACGAAGATGTCAATTTTGGATAGCAACTGTTACCAATAAACATATTCA
AAAGTAATCAAATATATTCAAAAACAACTCGTATAAATATATAAAGTTCGTGATATTTATTATAAAGAAGCCGAAGGAGAGAGC
GGTTTCCGAACAATAAAGATATACAGAGGTTTTATTCTTGACGGCACTCTCTCCTTTAGCCGCAAGTTTAATTCCTCTTTTTTA
TTGCACTATGGTCATCGACAGCAAATATACCAAGACATTCAAGTCAAACGGACTGACCCATCAGAAATATGACGAGTTGCTCTC
GTTTGCTTCTATGCTGCGTGACCATAAGAACACCATCTCCGAATATGTCAATGCCAACCTTGAACACTACCTCGAATACTCAAA
ACTCGACTTCCTTAAGGAAATGCGTGCGAGGTACAAGGATGTCGTTCCGAGTTCGTTTGACGCTCAACTCTACACG
(配列番号103)

>OBLI01003123_14|M
AGAATCTGTCCTATATGTGGGAAACATTGCGAATATGAGGAAATGGAGGGCGACCACATTGTTCCATGGTCAAAGGGCGGTAAA
ACCGATATAGGCAACCTCCAAATGCTATGCAAGAAGTGCAATCACGAAAAGTCCAATAGATATTAGTGGCGTAATCAAAAATTT
GTTTGTGTTGAGGAAAAGCAGTGAAAAAAAACATTGTTTTTCCTCAATTTTTATTTGCATAATTCAAATAATTTTTTATTTTAT
AGGATAATAGAGCTAACAAGCATTAACAATTATTAAAACGATTTATATTGAAAATAAATTTTGTGGGAATATTTATTTTTACTA
CCTTTGCATCGTAATACAATTAAACAAATTTTTGATTATGGCACACAAAAAGAACATAGGAGCAGAGATAGTAAAAACTTACTC
TTTTAAGGTGAAGAATACCAATGGTATCACAATGGAAAAATTAATGGCCGCCATTGATGAGTATCAGTCGTACTATAACCTTTG

【０２６０】

【表３９】

CAGTGATTGGATATGCAAGGGTCTTGACGAAATAATGAGGAATACTTTTCTGAAAAAAGCAAATAGCAATAAATCATTGTATAA
TCAGCCAATCTACGATACGGGTATCAAGAAAACCGCAGGTGTGTTTCCTAGAATGAAAAAATTAAAGAAATATAAAGTTATCTG
AAATAAAATATGTATTTTTCTTTGTGGAAATACCTATTAATAGACTGATTTCTAATAAGTTATAAGAAATACTGTATGTAGTAA
ATAAGATATCATATTTTTGCGGAGAGGCACATGGAGTATGCTATAGGGTTTTTGCTACCGAGCAGAAAGCAAAAGAAAAAATGC
AGGGATGATATCATTTCATTCTTGCATTTTGCTTATACATATTCAATCAAGTATCATTTTCTGTTTTTACTATTATCCTATAAA
ATAAAATTTTCCTCAACATTTCCAAATTTAATTTGCAATAATTTTTTTTGATAAAAAGTGCAAATAAATTTTATAGATTCAAAA
CTTTTGATTAACTTTGTAACAAGAAAAACATTAAGGATTATGGGTTACACATATTTTAGGGTTACTGATGAAAGGGCAAGGGAT
GTTATGCCAAAGGCGGCTGAAATCATAAAGGATATTTTC (配列番号104)

>3300028887|Ga0265299_10000133_30|M
CTTCACCTCGTACAGCCGACAATAAGTTTCGCTTGGACTGAACTTATGTGCGCCTGCGCATTCATAGCGGGTGGCGTATCAGGC
TATCTCATCAAGGGCAAGATGCCAAACGACGGGAACAAGTACCAGTCGGTAGAGGGAAAGGAATAGGACAAAAAAAAACACATC
ACCCCCAGCGCATCGGGCGCGGAGGTCGGGTGTGCATATAACGGTGTCTGTGGCGCAACTGGTAGCGCAGTGGATTGTGGTTCC
AAAGGTTGCGAGTTCGAGCCTCGCCAGACACCCATTATCACACGGAAGCATTGGATGGAAGTGCAAGTACCTACTGGGAACTTC
CTGAAAGCGCAAGCAAAGTCGAGGTCTAACGGTACTTATGACCGAGGTAATGGCGGGGCGTTGGTTCGAGTCCAACACAATGTT
TCCATTTACACGGAGAGTTGCAGGAGTGGTAACTGGTCAGATTGCTAATCTGAAGCCCACCTCGTTGTGGCAGGGGTCCGAATC
CCTTACTCTCCGCCAAGCAACATACCCGCAGAGTAGTCGCGTATATTCTGTCGGTGTGGTCAGAAAGAAGTGAATGTGATGCGA
ACGCGCGAAACCATCGCATTTAGAGTCCGAATCTCCTCTGCGGTAGCCAGTCCGCATAGTTTAATCAGGTTAAAACATTCTGAC
GCTTTTTTAAATCGCGGGAGTAGTTCAGTGGTAGAACATCGGCTTCCCAAGCCGAGGGTCGCGGGTTCGAGTCCCGTTTCCGC
TCAACACATAGGCTGTGGACAAGGTGGGCGAAAGTATTTTTTCCATAGTTTTACACCAACGCCCGCCTTTTCCTAAACGCATTG
GAGAGATAGAGGACTTGCCTTCTAAACAAGCAGTACGGGGGAACTTGCATCCGACCTCCGTTTCAATGCGGTAGAACTCCGCTC
CCGTGACAGCGACGAATGATGCAATAGCGGTTCACGAGATACCTCAAGAAACTTCATTTTTCAAAAGCCACAATAGTTCAACTG
GTAGAACGGCGGTATCGTAAACCGCAGGTTGCTGGTTCAATTCCTGCTTGTGGCTCAACAATTTCGGGGGCTTGCAACGCTGCC
ACTGCGGGTGGAAGCCAGCGACAAGAACTTGTGTGAAGCCGAAACGCAGTCCTTCGGGAGAGGGGCGAAGGGGCAAGCGAGATG
TGTCCCACTTTTTTAAAGTAACAGGCTTTAATAAATATTTATCATTCCCGAAAGGCTGTGCGGAACAGCCTCTCGGCTTTTACG
GGGATTTAGTTCAGTTGGTAGAACATCTGGTTCGCAATCAGAAGGTCGCGGGTTCGACTCCCGCAATCTCCACAAATATAAATA
TAGTATTGCCCTGTGGTGCAATCGGTAACACACCAGATTCTGAATCTGGAATTTCGAGTTCGAGCCTCGGTGGGGCAACACAAT
AGGCAGCCGTACTGCCGAATACAAGCCTGTGGAGAACCCAACCGTGGATGACCGTTGCCTATGCAACCTAAAAAGCGGTGGTTC
TGTGAAGCAGGAAGCGGAAATACAATATTCCGCATACGGTGGTGGTGTAATCGGTAACATAACAATATCCGAAAAGTTTAAACC
ATACACCCGACGATTATTTTTATTCATTGTTAGCGACCGCCGTGAGGCGGACGCAGGCTGGCGGTCGGATAATGACGCATAATG
GCGGTTGTGAAAGCCGACGGAAAGCACTACATCGTTAAGTGCCAGCCACCATAATAGGCAGCCGTACTGCCGAATTTAAGCCTG
TGGAGAACCCAACCGTGGATGACCGTTGCGTAAGCAACCTAAAAAGCGATGGTTCTGCGAAGCAGGAAGGAAATGCCCAATTTA
TTAGGTTTTTCCATACGGTATGACAGCCTCTAACTGTAGCGCATTACAAAACAAACGCTACCATTACATAAATGGTCAGAGGCA
TAACGCCGAGCGCAGGTATGGTATGCGTTCAAGTCGCAGTCACGGAAGCCCCAGATAAAAATGGGAGGTGCTTGCGGTCAAGCG
AGTGGTCAGCGGGCTTGCACTCGGTGTGGCAACAATGGTCGTTTCCGAACTTACGACCATTCAAAAAGATAAGGTAGTGGCTTG
TGAGTGAAAAGAAACTCTCGATACGCTCCTTTCGTCTAACGGTCAGGACGCGAGATTCTCAATCTCGTAATGCGGGTTCGATTC
CCGCAGGGAGTACAATGGCGAACACACGACAATCCAAACTGAAGGGGAACTGGAAAACCCTCGCTCCGAGATAACATCAGCGCA
GAGAGGTTGGTGAGGCAACCGTAAAAGTAATCCTGTGTGCAAGCAAGAAGGAAGTTCGGGTTCAAGTCCCGATGAGGATTATTG
TTGAAGAGGGATATGATTCAACCATAGCACTTATGGTGCTGTGCAAGGGTTATAGGCAGCCGTACTGCCGAATACAAGCCTGTG
GAGAACCCAACAGTGGATGACCGTTGCCTATGCAACCTAAAAAGCGGTGGTTCTGCGAAGCAGGAAGGAAATGCCCAATTTATT
AGGTTTTTCCATACGGTATCACTACTCGCGGTGGATGTGGAAATAACCGCGATTTGGTCAGTTGGTGAAGTTGGTTATCATACC
TGCCTGTCACGCAGGTGTTCACGAGTTCGAGCCTCGTACTGACCGCAGACAAAGACAAAGAACGAGAGGACTTGTATGACTTGC
AAATGTCACGGACTCAAACAAGAAAAGTTTATAGGCTATTAGAGGATGACTGTTTCTTTAATTTGTTTTCTTGTACTGAAGGTC
ATCACTGCCGTGCCACCAAGCCGTGCAAGTCCAAATGGTGCGTTAGTTCAGTTGGTTAGAATGCCAGCCTGTCACGCTGGAGGT
CGCGGGTTCGATTCCCGCACGCACCGCAATAATCTGGATATAGGCAAATTACACATATCATATGTCGCCCCGCGTAATCATAGA
CGACACTGCGGACGACAGCGGCGAGAATGTCGAAAGGCTCGACAGCATAATGACATTCGACATCACCGACACCCCGATATACGA
AGGCGGGGAGGAACTTGAGATAAACGCAAAATTCAACAGATAGAAATAATTAAAACAAACGGCAATGGCACACAGAAAAAAGAA
AGATGACGAAGCAACGCTATCGTACAAGTTCAAGGTAAAGGTCATAGAGGGCGACCTGACGGCAGACGACATAACGAAGTGTAT
CGCGGAAAACGCGGAGCAGGGCAACCATTTCTCCGAGTTCATACACGATGAGAATTTCAGGAAGACCTTCACATCCGAGATCAG
CGCGGACAAGTTCGGATGGGGCAAGCCGATGTTCAGCCCGACCACCAGAAGTCAGGACGAAGTGTTCTCCGCGATAAAGAAAAT
CGGGGCGATAACCGTGCTGGAAGATTAGCGCATATTATTCTCATATCTAAAATTGGAAGGACACCTGCGGACGCGGGTGTCCTT
TTTTCTTAAAATGCCAATTTATAAATAATATATAACTTATATTTATTGTACTTTTTTTGTTTAACTAAAACACATAGACAAATA
TGGAAATTCAACAGATTAGGTTTATAAACCCAGTTGATTTTGAAGAAACAATCGTTAATGTACCCACGGAGAAGGGCGAAAGAT
TCCTGAGAACAAAAATCTATACGGACGAGTATTCACCCGAAACATTCATAAAACTCTGCGAGAAG (配列番号105)

>3300028888|Ga0247609_10000668_74|M
TGGCGATTATTCTTACGGCAAAGGCCTTATCCATGCATACATAAATCGAGACATCAAAAGTTTTTGCTTGCCAAACACTTTAAT
ATGTGAATGCCATATACCAAAACATACCAGATATATTACTGATTACTCAGGTACAAATATAGCCGCAAAGAAAATCATCATCGA
CAAAGTTGTCTGGGAGAAGGTATGTATAAAAACATAATGGTATTAGGGGAGAAATTTTCTTGGACGGAATGAATATAATTTCAT
ACCAACACCGTGCATTGATTAAACTAAATTAAATTATCAAGCATAAAAAGTTTGGCACGGTTTTTGATATAGTAAATTTGTATT
TAAAATTTTTAATATGGCACACAAAACTAAAGAATCAGAAAAATTAGTAAAGTCTTTCAAATTAAAAGTAGACATTAGCAATTG
CGAAATTGAAAAGAAATGGATTCCTTCTTTTGAAGAATACACAAATTATTATAATGGAGTAAGTAATTGGATTTGTGAACTATT

【０２６１】

【表４０】

AGAAAAAGTTTGCCTGAAAAGAAAAAAATTTGGAAAGGCTTCTTATTCAGTACCATATTGGAACGTTAAAGACGCATTTAAGAA
AAACGTTAGCTCAAACATGATTGCTACAATTAAAAAAATGAATATGGTAAAGGTTTTTTAATGCGTGATTATGGCGTTTTTTAA
ACATAAAATCATTTATAATATATTGAAAAACATTTTATTATATAAAATATGCATCTTAGTGAAACCGTGTTTTCGTATAGATTG
CTGGATTATACTTTTTTATAGGATAATTACAGCTCGAACTTCTTTGATGGCATTAATAAGATATTGTTGGATTAT
(配列番号106)

>3300028805|Ga0247608_10000895_42|M
ATCATGGCTGAAAGCGTCCGCCTGATTGCAGAGCAAACCGCAAGCCCGAAGGTTGTCATCAAGAGCCGTTACGCTCTGGTCGAC
GCAGGTTTCTATCCTGAGTTGAACTATGTGACCTTCTTCGTGAACACTCCAGATCAACTGGTTTAATCACTGCGGGTAGCAAGC
GATTGACTACGGAAGGCCGATTCGATAGAGTCGGTCTTCTTTTTTTTTTGTATATTTTCTTTTTTTGGTTTGGAAATGTTCCGT
ATATTTGCAGCACTAAAACTAACCAATATGGGACATGTACGTTTGCAAAAAAGAGAGGGAGAGGTTTATAAGACCTACAAACTT
AAAGTAAAGAGCTTTTCTGGCAATGTAGACATTAAAGCTGGTATCGTTGAATACGATATCGCCGAAACAATTGATTGGAGAAGT
ACGCTTTGTTTCAAGACATGGAATACGTATGGTTCTCCTCAATGGGACTCGAAGATCAAGAACCAGAAAACGATGATCGATCGA
CTGGATTCGTTGGGTGCAATAGAATTGAAAAACTGGTGATTTTGATCATGGTTTTGAAACAAAATATTGATTTTTCGTTCTTTG
ACATGCTTGTTAAAAATTGAGTATCAGTTTAATATAAAGAATATAT (配列番号107)

>OQVL01000914_15|P
GGAAACAATTATAACGATGCCTACAAAACGTTAATTCAAATGAGAGACAAAGGAATTTTAACGCAGGAAGTTGTAAATGTATTT
ACCCTATTGAAAGGGCGGTATATTAAAGAAAAAGAATACGGAACACAATATAATACTATCAATTAAATTTTTTGGTAGTTTCAT
TTGGAATTGCCAATTATTTTTTTATTTTATAGAATAATAGAGCCAACAAGCATTAGCAATTATTAAATCGATTTATATTGAAAA
TAAATTTTGTGGGAATATTTATTTTTACTATCTTTGCATCGTAAGATAATTACAAAACATTAACAACATTTATTAAACAATTAA
ACAAATTTTAATTATGGCGCACAAAAAGAACGTAGGAGCAGAGATAGTAAAAACTTACTCTTTTAAGGTAAAGAATACCAATGG
TATCACAATGGAAAAATTGATGAACGCCATTGACGAGTTTCAGTCATACTATAACCTTTGTAGCGATTGGATATGCAAGGGTCT
TGACGAAACAATGAGGAACACTTTTCTGAAAAAAGCAAATAGCAATAAATCATTGTATAATCAGCCAATCTACGATACGGGTAT
CAAGAAGACCGCAGGTGTGTTTTCCAGAATGAAAAAATTAAAGAGATATGAAATTATCTAAAATAAAATATGAATTTTTCTTTG
CGGAAATACCTTTTAATAGATTGATTTCTAATAAGTTATAAGAAATACAATAGATACTGAAGGAAAATCAAAGTGTAATCAAAA
ATTTGTTTGTGTTGAGGAAGCAGTGAAGAAATTTCATTGTTTCCTCAATTTTTATTTGCATAATCCAAAAAGTTTTTTATTTTA
TAGGATAATAAGACTAACAAATCTCAACGACTATTAAAACGATTTATATAAAAAAAGTTTTGCAGTTCCAATCTTTTTTGCTAT
CTTTGCAGTGTTGAAAGACAACAAAGATTTAAGTTTAACAAACAAATACTTTTTATTACATATTTTAATTTTTTTGTATTATGA
CAATAGAAGAAAAAGCAAGGGAAGAATACCCTTATATAACCCCATCTGATGGGTATGAATGCCATGATTATAATGAAGCCGCTA
AAGACGGTTTTATTGAGGGGGCAAAATGGATGCTTGAAAAAGCCGCTGAATGGTTTAAGAAT (配列番号108)

>3300028888|Ga0247609_10003329_9|M
ATATGGGCAAAGCGTGATAAAATTGAAAACAAATATGTCAAAGAACCATTAAAACGAGTCAATGAAGATATGTGGTGGATGTAC
TATGTTTATGAATGGAATGTGTTTTATGTGCTTGAAGAAAATGTCCATCCATATATGAAAAAATAAATTTTACCACACATATTA
TTATTCGTGTCATGCCGATGAGGTTTGGCACGATTTTTGTTTATATGGAGAGACATAATGTCAGTCAATACATGACAACTTGTC
ACAATAACTGACATTAAAAGTTTGGCACAATATTTGCTTATAAGAAAAACGAACAAGTAAAATTAAAATTTTATAGATTATGGC
ACACAAAACAAACAACGGAGAAAACACCATCAACAAAACTTTCATCTTCAAAGCAAAATGCGAGAAGAACGATATTATATCGTT
ATGGAAACCCGCAGCAGAAGAGTATTGCAACTATTATAACAAATTGAGCAAATGGATTGGTAAAACAATGTACGGCATTCCTGC
ATATAACATCAAAAGAGGTTTTAAGAAGAATTTAAGTGCCAAAACTATAAACACATTTAGAAAACTTGGACACTATCGTGATGG
AAAAATAAATGAGGATGGCATGTTTGTTGAAACTTTGGCATAGAATTTGCATATACCAATTAGAATTGAAAAAATCGCTCTTTG
ACACACTGAAACATACAAAAACACCACAATTTTTTAATCCTTTTCTATTTGTATTTTATTGAAATAAAATGTATTATAGTAATA
TATCTGCTAAGGTCATATTTTTCATTGTTCTCAAATTGTTGGATAATGTTTTGTGTGTTTCATTTTTGTCATTGTGTCACCTTA
ACTGACAAGGTGGCACATTTTTTATGTCAATATGTCAGTTGAGGTTTTGGCATAATTTTTGTATAATGGTAAATGGATAAGAAT
TGAAATTACAATGACAACAAAACAAAGGTTAATAAAGAGAATAAACAAGGCATTCGGATTTGAATTAACGGATGCAACACCTTG
TTTCCACCATCAAGGTAGAAGATGGGGAAGCGGTGGTTTC (配列番号109)

>3300028805|Ga0247608_10006074_1|M
GAAGGCGGCGCGTTTGAAATCGCTAACGTAATTGAAAATGCCAAGAAGCAGAATCTCGGGGAGGGTGGATACAAGGAATTGTGC
AATGATTTCCTGAAACATGCGAGGGAAACGTTTTTCAGTGGGAAATACGAACACCATTCTTGGTAGTGGATTTGTTATTTTGGT
AAATATAATTAACGCGGCATTGTCGTCAGTGAATATAATATTGCATTTCGACAGTATTTTATAAGTATTTTGACTTATAAACAG
TATTTATAAGTTATTCGGCTTATAGGTTAATTAGCCTATAGATGTTGTTTATAGGTTGGATGACCTATAGTGCCAAGTTTTGAA
GAAATCGTTATAGTCATCGTTCTGCCCTATTAGATATTCCGTATTTCTTTAAGACTGTTATAATACAAATATACTACAAATCAT
GCAATTTTTGATTTTTAACAAAAATTAAGAAATAGGGTATTATTGTGTATTGTTTTTTGTTATATATTTGTCCTGTTAGGTTAA
ATCACCGCGCCTGATGACGAAGTCGGTGGTAGAATTAGACTAATATTAAATATGTCTCATGAATTTAACAAGAATAAAGGTGAG
AATGAGATTAGCAAGACCTTTATTTTCAAAACAAAATGCGGGAAGAATGATATTACATCATTATGGGTTCCCGCGATGGAGGAG
TATTGCACGTATTACAACAGGGTAAGCAAATGGGGGAAAGGTATGTACAACAAGCCGTCATATGACATACGGAAGAAATTCAAG
AAGAACTTGAGTGCGGCTACTTTGAAAACTTTCATTAAGTTGGGAAACACGGTGAAAGGGATGATTGTCAACGGACAGTTTGTT
GAAATGGAATCATAGGTTGACAGAAACGGAAAATCGGTTTGTTTGTTAGAAGAATATTTGTTGAAATTCATTTTTCTTTTGCTA
ACGTATATACAAATAACTGTAATAGAATATCTTATATAAGATAT (配列番号110)

>3300012973|Ga0123351_1009859_3|P

【０２６２】

【表４１】

ACAAATGAAATTATGGGACAAGTAAAACTTAATAAACCTCTTCTGTATATCAAAATATTGACTATCTTTAGACATAACCTTGTC
AAATAATAAATCTAAATTACTCTTTTCCTTTTCTTTTTTAAATAATTTCATATTAAATATTCCCATAATTTATTAATATATTTT
TTTTTCATTACTTATTTCTCTGTTATATAAATAGTTACATAAAAAAATTAAAACTATTTTTTAAAAAGTCTTGTGTATATAAAA
AAAATATAGTACCTTTGCACCCGAAATCAAGATTTAATCCTGTTTTCATATTATATTTATCAATTTTATACTAATTAATAAACT
TATGGCAAATAAAAAATTTAAACTTACAAAAAATGAAGTCGTGAAATCATTCGTACTCAAAGTTGCTAACCAAAAAAAATGTGC
TATCACTAACGAAACACTTCAAGAATATAAAAACTATTATAATAAGGTAAGTCAGTGGATTAATAACATCGTACAAAATGAAAC
GTGGAGAAATCTATTTACTAACAAAACCAATAATACATATGGATTACCTATACTAACACCTTCAAAAAAAGGACAATCTAATAT
CATTACACAATTAATGAAAATTAATGCAACACAAGAACTTGTTGTATAATATAATCTATTTTTAAATTTATAATACTAATATAA
TTCATTGATAATTAAATAATTATATAAAATTCCTATATACAATAGAAAGACTTTCCACAGACATGTTGTACATACATTTTTTTA
AGTATTAAACAACGCATACCCACCAATGGTACACGAAAATTTTCATGTTGTACATACTATTTTTAGGTATTAAACAACTCACTG
TTTTGACGATTAATATAGGCATGTTGTACATACTCTTTTTAGATATTAACAACCTGTAAACAATAACAATATTTACAACAATAA
TCCATTTTTGAAATAATGAAAAATTTTCTGGAAAAATTTTTTAACAAGTCTGTTTTTGAAATAATGAAAAAATTTCTGGAAAAA
TTTTTTTAACAAACCCATTTTTGATTGGTTCATTTTTTATTGGAAAATTAGTGTGTGGAACTACCCACCCGTATATGAGCAAGT
GTTATGGGGTGTAACGTGGGGAGGGTTACATAGGGGGGTCTTTGGTAGGGGGTACATAGGTAGGGTAATAATGGGGTCTTTGGT
AGGGGGTACATAGGTAGTCCCCATATATTATTATAAAAAGTAAAATAAATGATATATGCAAGAGTTTTTGAAAATTTATTTTTA
TTTTGCTACTTAGACTTTACAAAAAGTAGATATATAGTATTTTCTTTTCAAAATATTTTGTAGTTTGGAAAAAAAGCAGTACCT
TTGCACACGGAAACGAAAAACAAGTTTAACCTATTAAATTTTTAGTTTATGGCAATAAACATTTTGACTTATTCTGCTATGGCA
GAAAAATCTTGGGAAAATTTTATGCGTGAAAATTGCGGTTACGAGCGCATTAGTACATTTTATAGTGATTTCACTATTGCAGAC
CATTGTGGTGGTGTAAACGCAATAAAAGAC (配列番号111)

>3300012979|Ga0123348_10005323_4|M
GATGTGAATGAAGAATTTCTTGGTGGCTTGCGAAGCACTATGACATATCTTGGAGCAAAGAGATTGAAAGATATTCCGAAATGT
TGCGTTTTCTATCGTGTAAATCATCAGTTGAATACAATTTATGAGAATACAACGATAGGAAAATAATATAAATTTTATATTATT
TTGAGAAAAAGAGTCTAAATTTGGGCTCTTTTTTCGTTTTTTATGAAAAAATATGAAAAAAGTTTGTAAAAAATTTGTAATATT
GAAAAAATAGTATTATATTTGTATCAAATTTAAAAATAAAATATAAATATGGCAAAATCAATAATGAAAAAATCAATTAAATTC
AAAGTAAAAGGAAATAGTCCAATAAACGAAGATATTATAAATGAGTATAAAGGTTATTATAATACCTGTAGTAATTGGATTAAT
AATAATTTAACAAGCATAACTATTGGTGAAAATGAAGACTGGAGAAAAGTGTTTTGTATCAAACCAAAAAAAGAAGATTACAAT
ACACCTTTATTGGATGCTACGAAAAATGGTCAATTTAGAATACTTGACAAGTTGAAAAAATTAAATGCTACTAAATTATTAGAA
ATGGAAAAATAATAAATATATACAATAAATTTATATAATTTTGTCTATTTTTAATTTTAGTTCATTAGATAATATGTTCATAAA
TTCATTGACATATAATTATAAATAAATATATATGCAATAAAATTCGAGAGACATTTCATCAGAGATGTCTCTTTTTTATTTTTT
GTTATATTTATATTATGAATATTAGATTGGAACTCATAAAGACAAAGGATAAACAGAACATTGCAAAGCGTATAGTGGAAAGCA
ATCACTCATATGTTCCAACCTGGCGTAGTGTAGGACGAAGGATAGATTATCTTATTTATTTGGATAATGATGTTGTCGGA
(配列番号112)

>3300028888|Ga0247609_10016480_8|M
GTGAACTATATCTACGAATCAATCGAAGGAATATTGACAAAAACAATGAATCCAACCACTTTACAGGATATCATCCTTAACGGA
ATCACATATACACCAGTGGAAGACAACACAACAACATGCGACGGATGTGAATTTAAAGACACATAAGGCCAATGTATGCTAACA
CACCTATTCGATAACGACATGGTCCAAAACTGCCTCAAGGAAAAAAACGGCGTTGCAGATATCATATATGTCAAAAAAGAAAAT
TAATCGGAATCTTGATTTGGATTTTAATATTATTTGTTGTATAATTACAATAGAAAGAAAATTTTGTATATTTTAAAATTTGTA
AATTAAAATTTAGAAAAATGGCACACAAAACAAACAACGGAGAAAATACAATCAATAAAACTTTTATTTTCAAAGCAAAGTGCG
ATAATAACGATATTATATCGTTATGGAAACCCGCAATGGAAGAGTATTGTACTTATTACAATAAATTAAGCCAATGGATTTGCA
AGACAATGTATGGAGTACCAGCTTACAACATTAAAAACGGTTTCAAAAAAAATCTGAGCACAAAGACAATCAATACGTTTAGAA
CGCTTGGCCACTATCGTGACGGAAAAATAAACGAAGACGGCGTATTCGTTGAAAACCTGGCATAATAAGGAGTAAAAAAATGTT
CTTTGATATTCTGACACAAATGAAAAAACAATCAAAAATTTATTTCTGTTTTGCTTGTAATTTATTGAAATAAAATGTATTATA
TAGAAATATGTCGGTGGATAATAGTCAAATAGTCTGTTGACTGTTGAATAGTAAGTTTTTTACTCTATTGACAACAGGTGATGT
GGATGGAACATACAAAGTTTATTGTTGAGTAATAGGTTTTACACTTTTACCACAACTTTAGTGATTTTATGTATAAAATAATTA
AAATCATATATAAAAATTTTTCCAGAAAGTAGTACTTATTGAATTAAAATTATATTGTGAAAAATGGTTTTTGATTTTAATTTT
ATTTGTTGTATAATTGAAATGTAATTTAATTTAGAATTGTATAAATAAAAAACGTAAAAATGAGACTGCCAACAGAAATTTATG
AGTCAGGCACAATGGTTAGTAAGATATCGGAAAAACCATTTAAATCAGGTTTAAGGGTTAATACTGTAAAGTCTGTAGTTGAAC
ATCCACATAAGATTGACCCGAATACTAATAAGGGTGTTCCA (配列番号113)

>3300028887|Ga0265299_10000013_320|P
GACTACGACTGGTTCTCAAATGTGTACGGCGCCATCAGGGAGGAACGTGAGAAAATGAGAAGGGAAGAGGAGGAACGCAGGAAG
AACGAACCCAAGACGGTGAAAACCAAAGAGGTTGACTTGTTCGGGGATGATGACCTGCCGTTCTAATAAAAAAAAAAACAAACC
TCTCCGAAATTGAACGTATCAACTTCGGAGAGGTTATATAGGGTGATGGAAATGTTAAATAAAAAGTTTAAAAATAACTATGGG
AAACAAAGTACAAAGTAATGAAACAATAGTTAAGACTTATACATTTAAAGTGCGTGGATTCATAAGTGGTGCTACCCACGAAAT
AATGAAATCAGCCATAAAACAATATATAGAAGATTCTAACAATCTATCAGATTGGATTAATGTAGAGAATGAAATACTTAGGAA
CTCTTTCCTTAAAGAAGAGACTAAAAAATACACTTATAATACACCATTATTCACTCCCAGACTTAAGTCATCGGAAAAAATAAT
AACAGAATTGAAAAAATTGGGTATGACTACGGTTATAGAATAACCATTACACATTTTTTTCATAACAAACGTTCTTTAACATAT
TGGAAAATAAGAAAATACGATATTCATATAAAAATCCGTCCCACACAAAATTAATGTAATATCTTAGTTTTGTTACATCAACAC
TATATAATTAAAAAAATAAAAAAATATTTTGTGGATTCAAAAAATCATTATATATTTGCGTCCGAAAATTAACACTTATGTCAA
ACAAATTTAAAATGTAAAAGAACTATGCAAACAGAAACACAGAATTTCACAGGCGAGTTGAGAGCAATCAACACAACAATGGGT

【表４２】

TCAAGCAAGAGCTACAAGACAATCTGCCGTTGCGCACTTGACATCCTCAAGGGATATATCGTTACGCACGACATTAGGGACAAC
TTCTCA (配列番号114)

>3300028887|Ga0265299_10000026_77|P
ACAGAGGGTGTATGGATAGGCATGAACCACCAAGGCAAAATACTGATGGCTTGCAGGGAGGCTTTGTGTAACAACTGTGAACCC
CCGATTGATTACAAGGCACTGAACGATGCCGAGATATATTTTTATGGAAAAGAAGTTAAATTTTAAAAATTAAAAGATATGGCG
AACAAAAGCACAAAAGGAAACCTGCCCAAGACAATCATAATGAAGGCAAACCTTAGCCCCGATGGTTTCACTCAATGGGAAAGG
GTTGTAAAAGAATACCAAGCCTACAAAGACACGTTGAGTAAATGGGTAGCCCAAAATCTCAGACAAATAATGTGCAAGACACCG
CAGACAAAGAACGGCTACTCATCACCTGTGCTCACCTCAAAGGTTAAAAGCCAAGTGGAAATGGTAAGAGAATTGAAAAAAATG
GGAAAAACCATTCTTTATTCCAATGATTCACTTCCTTTTTGAAACTAAAATGTCTTATGTGTATTTGAATTATAGGCTAATATA
AAGATTGTACTGTGTTGAGATACACTTTTAGAGGTATTTACAACAAAATGCGTGATATGGAAATGAAGAAATAACTGTGTTGAG
ATACACTTTTAGAGGTATTTACAACACCATATAAACCTGACCATCTCCTGAATCTCGCCCGACACGGATAATGTTAGATATGTT
CACAATACAACTGCATGTGCTATTCAAGAAAAAATAGTATATTTACAATATGTTGGTGCATAATATTAGATGTGCTTACACAAC
GCAGACCTGAAAAGCCAGGATAAAAGTATGCGGGATTGTGTTTTTAGAACACTGTTCAATCCGCTGTATGTCGCTTGAAGCGTC
AGTAACCTATGTCGAAACAATCCTTTTAGAGGTGTTTACGACCGACCAGAAACAGCAAGACCTGTATTTATGTTGGTATACGGT
TCTTTTTAGGGGATTAGTAGTTGAATCCCTTTTCACCCTTGGTGTTCACGGGTTGTGAGACATTCTTCATACCCATGCGTGTCT
TCTCAGCCATCTTACCGAAAGTTATAGGCACAATATGTTCAATGCCTGCCTGCTGAGCATTGTAGCATATATCAGACAG
(配列番号115)

>SRR1221442_316828_61|P
AGAATGCTTTCCCCAATTGAATGTGAAAGACTACAGACACTGCCAGATAACTATACCGAAGGTGTTAGCAAATGCGCAAGATAT
AAGGCAATCGGAAACGGATGGACAGTTGATGTAATTTCACATATTTTTAAGAATTTGAAAAATTAATTTGGTATTTTGAAATAT
TTGACTTATTTTTGCAACATAAAATTTAAAACAAATTTATATGGCACACGCGAAAAAAAAATTTTGACAAAGGAAAGCAAATAA
CAAAAACGTTCTCTTTCAAGGTGTTAAATATTAAGAACAATGGCGAATCAGTTGATATGAATACTATAGAATTAGCCATGAAAG
AGTACAATAGGTATTATAACATTTGTAGTGATTGGATTTGCAACAATCTAATGACGCCAATTGGTTCCCTATATCAATACATAG
ATGATGAGAAATGGAGAAAAAAATTTGTTCGCCCAACAAACACTAATAAACCGTTGTATAACTCTCCAGTTTTCTCCCCTGCTG
TAAAATCTGAAGGTGGTACTATTAAAAATCTCCAAATTTTAAGCGCAACAAAGACCATAATTCTTTGATTTAATTATTAATACA
TATATCGTTCGTAAATTTAATACAACCACAACCAAATATGATAATTTGCATAATTAAAAAAATTCACATATCTTTGTAGCATAA
AAACAAATAGAGAAAAAATGACACTTTACAGATTTACACTTTTAGGCAATACACAAATTTATGTATATGCTGGCACGTTTGAAG
ATGCTCTCAGGACATTTCGTAAATCATATGGAGATACGGGATTCAAGTCAATTGAAGAGCTTCCTGAATTTAGAGATAACATAC
TTATACAACTAGATTGATTGAAACAAACGTCAATTACCCACCACTGAAGTAGTGGGTTTCTTTGCAGTGATTTTATGAAAACGA
TAGAAGACAGAGCAGACATAGCAAGCGATATTGCTAAAAGAGAATTTGAAGAAGATAGTTATTGGAGTCATTACGCAGACGATA
TGGTAACATCTGCTTTTGTTGAAGGATGCTATAAAGGCTATATTTCAGGTGCGACA (配列番号116)

>SRR5678926_1309611_3|P
AAGGAGATAGATTATGACAGGGAAGGTAATATCACAAATATATATCTTTACTATGAGTCAGATAGTTTATGGAATGAAAAATTT
GAATTTATATTAACATTAGATGGTTATGAATTAAAGATACCTATTTTTATAGTAAGTGTAAGATAGTTTTGGCACGGAAATTGC
AGTAATGTTTTCCTGTCAAGAACAAATAAAATAAAAAATATGAAAAAATCAATTAAATTCAAAGTAAAAGGAAATTGTCCAATA
ACCAAAGATGTTATAAATGAATATAAAGAATATTATAATAAATGCAGTGATTGGATTAAGAATAATTTAACAAGCATAACTATT
GGGGAAATGGCAAAATTTCTCAATGAAGTGTGGAGAGAAATATTTTGTACAAGGCCTAAAAAGGCAGAATATAACGTTCCATCG
TTGGATACAACAAAAAAAGGACCATCTGCAATATTGCATATGTTGAAAAAAATCGAGGCAATTAAAATATTAGAAACAGAAAAG
TAGTGACTATAGATATAAACTTCTATGATAGATATCTGTTTTTTAATTCTATTATGCAATATAATATATTGAAATATAAACAAT
TATAAATAAAACGGGTGTATACAACAAGTTTTTTGTTTTTCTTATTCATTATCTGTATATTTGTATTATAAACAAATACAAATA
TGTATAATGAATCAGGAATATATTGCTATAAAAACAAAATAAACGGAAAATTATATATTGGACAGGCGCTAAATCTTAAAAGAA
GATATTTAAACTTTTAAATATCAACCACAGATATGCGGGTCAAGTAATAGAAAACGCACGTAAAAAATATGGTGTAGATAACT
TTGAATATTCAATCCTTACTCACTGTCCAGTAGACGAATTAAATTATTGGGAAGCATTTTATGTAGAAAGATTAAATTGTGTCA
CACCCCACGGTTATAATATGACTAATGGGGGCGATTCAGTATATACTTCTACACAAGCATTTAAAGATGCACAAACTGAAAAGT
TGAAGCAAACTATTCTATCTAAGAATCCTAATCTTAATGTCAGCAAAGTAAAATATGAAGGTAATAGAATTTCAGTTATAATTA
CTTGCCCAATACATGGCACATTTAAAAAAACGCCTGATTACTTTAGAAATCCAGAAATAAATGATTTGTGTTGTCCTAAATGTG
TGAGGGAAGATATAAGACAAAAGACTGAAGATAGTTTCTTTAAACAAGCAACAAAGAAATGGGGAGATAAGTATGATTATTCTA
AAACTATAATAGTAGATAGAATTACCCCAGTTACAATTACTTGCCCTATACACGGAGATTTTACAGTATTACCAGGGAACCATG
TGTGTAAAGATAAAAATACTGGAGGATGCCAACAATGTAGTGAAGAAAGACAACATATTGAATCATTAGAAAAAGGTAGCGTGA
AGGTCATTAAGATGATAAAGAAAAAGTTTGGAAACAAATATTCATTAGATAAATTCGAATATAGGGGAGATAAAGAAAAAGTAA
TTCTTATTTGCCCTATTCATGGAGAATTTTCAATGACGCCAGGTAATTTAAGATATAGCAACGGTTGTCCACAATGCACTTTAG
AAAATGCTTATCGTATAAAAT (配列番号117)

【０２６４】

実施例３－酵素活性の新規ＲＮＡモジュレーターの同定

エフェクタータンパク質及びｃｒＲＮＡに加え、本明細書に記載される一部のＣＲＩＳＰＲシステムは、本明細書においてＲＮＡモジュレーターと称される、エフェクター活性を活性化又は調節するための追加のスモールＲＮＡも含み得る。

・ＲＮＡモジュレーターは、ＣＲＩＳＰＲ関連遺伝子又はＣＲＩＳＰＲアレイにごく近接して生じると予測される。ＲＮＡモジュレーターを同定及び検証するため、ＣＲＩＳＰＲエフェクター又はＣＲＩＳＰＲアレイをフランキングする非コード配列を、直接実験的検証のためのクローニング又は遺伝子合成によって単離することができる。

・ＲＮＡモジュレーターの実験的検証は、非天然種において異種発現されるＣＲＩＳＰＲ

システム又は合成配列のための宿主生物のスモールＲＮＡシーケンシングを使用して実施することができる。由来ゲノム遺伝子座に対するスモールＲＮＡ配列のアラインメントを使用してＤＲ相同性領域を含む発現ＲＮＡ産物及び固定化プロセシングを同定することができる。
・ＲＮＡシーケンシングによって同定された候補ＲＮＡモジュレーターは、ｃｒＲＮＡ及びエフェクターを候補ＲＮＡモジュレーターと組み合わせて又は組み合わせずに発現させ、エフェクター酵素活性の変化をモニタリングすることによってインビトロ又はインビボで検証することができる。
・エンジニアリングされた構築物において、ＲＮＡモジュレーターは、哺乳動物細胞における発現のためのＵ６、Ｕ１、及びＨ１プロモーター、又は細菌における発現のためのＪ２３１１９プロモーターを含むがこれらに限定されないプロモーターによってドライブすることができる。
・一部の例において、ＲＮＡモジュレーターはｃｒＲＮＡ、ｔｒａｃｒＲＮＡのいずれか、又はその両方と人工的に融合させ、シングルＲＮＡエレメントとして発現させることができる。

【０２６５】

実施例４－エンジニアリングされたＣＬＵＳＴ．０９１９７９　ＣＲＩＳＰＲ－Ｃａｓシステムの機能検証

　ＣＬＵＳＴ．０９１９７９　ＣＲＩＳＰＲ－Ｃａｓシステムの成分を特定した後、ＡＵＸＯ０１３９８８８８２と称されるメタゲノムソース（配列番号１）及びＳＲＲ３１８１１５１と称されるメタゲノムソース（配列番号４）からの遺伝子座を機能検証のために選択した。

【０２６６】

ＤＮＡ合成及びエフェクターライブラリクローニング

　例示的なＣＬＵＳＴ．０９１９７９　ＣＲＩＳＰＲ－Ｃａｓシステムの活性を試験するために、ｐＥＴ２８ａ（＋）ベクターを使用してシステムを設計及び合成した。簡潔に言えば、ＣＬＵＳＴ．０９１９７９　ＡＵＸＯ０１３９８８８８２エフェクター（表６に示される配列番号１）をコードする大腸菌（Ｅ．ｃｏｌｉ）コドン最適化核酸配列及びＣＬＵＳＴ．０９１９７９　ＳＲＲ３１８１１５１エフェクター（表６に示される配列番号４）をコードする大腸菌（Ｅ．ｃｏｌｉ）コドン最適化核酸配列を合成し（Ｇｅｎｓｃｒｉｐｔ）、ｐＥＴ－２８ａ（＋）（ＥＭＤ－Ｍｉｌｌｉｐｏｒｅ）に由来するカスタム発現システムに個別にクローニングした。ベクターは、ｌａｃプロモーター及び大腸菌（Ｅ．ｃｏｌｉ）リボソーム結合配列の制御下にあるＣＬＵＳＴ．０９１９７９エフェクターをコードする核酸を含んでいた。ベクターはまた、ＣＬＵＳＴ．０９１９７９エフェクターのオープンリーディングフレームに続くＪ２３１１９プロモーターによってドライブされるＣＲＩＳＰＲアレイライブラリのアクセプター部位も含んでいた。表９に示すように、ＣＬＵＳＴ．０９１９７９　ＡＵＸＯ０１３９８８８８２エフェクター（配列番号１）に使用される非コード配列は配列番号９８に記載され、ＣＬＵＳＴ．０９１９７９　ＳＲＲ３１８１１５１エフェクター（配列番号４）に使用される非コード配列は配列番号９９に記載される。ＣＬＵＳＴ．０９１９７９エフェクターが非コード配列なしでｐＥＴ２８ａ（＋）に個別にクローニングされた、追加の条件が試験された。図４Ａを参照されたい。

【０２６７】

　「リピート－スペーサー－リピート」配列を含むオリゴヌクレオチドライブラリ合成（ＯＬＳ）プールが計算的に設計され、ここで、「リピート」は、エフェクターに関連するＣＲＩＳＰＲアレイに見られるコンセンサスダイレクトリピート配列に相当し、「スペーサー」は、ｐＡＣＹＣ１８４プラスミド又は大腸菌（Ｅ．ｃｏｌｉ）必須遺伝子をタイリング（ｔｉｌｉｎｇ）する配列に相当する。特に、表８に示すように、ＣＬＵＳＴ．０９１９７９　ＡＵＸＯ０１３９８８８８２エフェクター（配列番号１）に使用されるリピート配列は配列番号５７に記載され、０９１９７９　ＳＲＲ３１８１１５１エフェクター（配列番号４）に使用されるリピート配列は配列番号６０に記載される。スペーサー長さは

、内因性ＣＲＩＳＰＲアレイに見られるスペーサー長さの最頻値によって決定した。リピート－スペーサー－リピート配列には、前述のＣＲＩＳＰＲアレイライブラリアクセプター部位、及びより大規模なプールからの特異的リピート－スペーサー－リピートライブラリの特定の増幅を実現するユニークなＰＣＲプライミング部位への断片の双方向クローニングを実現する、制限部位が付加された。

【０２６８】

次に、Ｇｏｌｄｅｎ Ｇａｔｅアセンブリ法を使用して、リピート－スペーサー－リピートライブラリをプラスミドにクローニングした。簡潔に言えば、本発明者らは初めに、ユニークなＰＣＲプライマーを使用してＯＬＳプール（Ａｇｉｌｅｎｔ Ｇｅｎｏｍｉｃｓ）から各リピート－スペーサー－リピートを増幅し、ＢｓａＩを使用してプラスミド骨格を事前に線形化して、潜在的バックグラウンドを低減した両方のＤＮＡ断片は、Ｇｏｌｄｅｎ Ｇａｔｅアセンブリマスターミックス（Ｎｅｗ Ｅｎｇｌａｎｄ Ｂｉｏｌａｂｓ）に添加する前に、Ａｍｐｕｒｅ ＸＰ（Ｂｅｃｋｍａｎ Ｃｏｕｌｔｅｒ）で精製し、製造者の指示に従ってインキュベートした。Ｇｏｌｄｅｎ Ｇａｔｅ反応物を更に精製及び濃縮して、細菌スクリーニングの後続のステップで最大の形質転換効率を実現した。

【０２６９】

異なるリピート－スペーサー－リピートエレメントとＣＲＩＳＰＲエフェクターとを含むプラスミドライブラリを、Ｌｕｃｉｇｅｎが推奨するプロトコルに従ってＧｅｎｅ Ｐｕｌｓｅｒ Ｘｃｅｌｌ（登録商標）（Ｂｉｏ－ｒａｄ）を使用してＥ．Ｃｌｏｎｉエレクトロコンピテントな Ｅ．ｃｏｌｉ（Ｌｕｃｉｇｅｎ）に電気穿孔した。ライブラリを、精製ｐＡＣＹＣ１８４プラスミドで共形質転換するか、又はｐＡＣＹＣ１８４を含むＥ．Ｃｌｏｎｉエレクトロコンピテントな大腸菌（Ｅ．ｃｏｌｉ）（Ｌｕｃｉｇｅｎ）に直接形質転換し、ＢｉｏＡｓｓａｙ（登録商標）ディッシュ（Ｔｈｅｒｍｏ Ｆｉｓｈｅｒ）のクロラムフェニコール（Ｆｉｓｈｅｒ）、テトラサイクリン（Ａｌｆａ Ａｅｓａｒ）、カナマイシン（Ａｌｆａ Ａｅｓａｒ）を含む寒天培地に播種し、３７℃で１０～１２時間インキュベートした。近似コロニー数を推定して細菌プレート上に十分なライブラリ提示を確保した後、細菌を回収し、ＱＩＡｐｒｅｐ Ｓｐｉｎ Ｍｉｎｉｐｒｅｐ（登録商標）キット（Ｑｉａｇｅｎ）を使用してプラスミドＤＮＡを抽出し、「出力ライブラリ」を作成した。Ｉｌｌｕｍｉｎａシーケンシングケミストリーに適合性のあるバーコード及び部位を含むカスタムプライマーを使用してＰＣＲを実行することにより、形質変換前の「入力ライブラリ」及び回収後の「出力ライブラリ」の両方からバーコード付きの次世代シーケンシングライブラリを生成し、これをプールし、Ｎｅｘｔｓｅｑ５５０（Ｉｌｌｕｍｉｎａ）にロードして、エフェクターを評価した。一貫性を確保するために、各スクリーンに対して少なくとも２つの独立したバイオロジカルレプリケートが実施された。図４Ｂを参照されたい。

【０２７０】

細菌スクリーンシーケンシング解析

Ｉｌｌｕｍｉｎａ ｂｃｌ２ｆａｓｔｑを使用してスクリーン入力及び出力ライブラリの次世代シーケンシングデータをデマルチプレックス化した。各試料について得られたｆａｓｔｑファイル中のリードが、スクリーニングプラスミドライブラリ用のＣＲＩＳＰＲアレイエレメントを含んだ。ＣＲＩＳＰＲアレイのダイレクトリピート配列を用いてアレイの向きを決定し、スペーサー配列をソース（ｐＡＣＹＣ１８４又はＥ．Ｃｌｏｎｉ）又は陰性対照配列（ＧＦＰ）にマッピングすることにより対応する標的を決定した。各試料について、所与のプラスミドライブラリ中の各ユニークなアレイエレメントのリード総数（r_a）をカウントし、以下のとおり規格化した：（r_a+1）／全てのライブラリアレイエレメントの総リード数。所与のアレイエレメントに関する規格化出力リード数を規格化入力リード数で除すことにより、枯渇スコアを計算した。

【０２７１】

酵素活性及び細菌細胞死を生じさせる特異的パラメータを同定するため、本発明者らは次世代シーケンシング（ＮＧＳ）を用いて入力及び出力プラスミドライブラリのＰＣＲ産

物中における個々のＣＲＩＳＰＲアレイ（即ち、リピート－スペーサー－リピート）の表現を定量化し、比較した。アレイの枯渇率は、規格化された出力リード数を規格化された入力リード数で割ったものとして定義された。枯渇率が０．３未満（３倍を超える枯渇）の場合、アレイは「強力に枯渇した」と見なし、図５及び図８に破線で示した。バイオロジカルレプリケートにわたるアレイ枯渇率を計算する際には、全実験にわたる所与のＣＲＩＳＰＲアレイについての最大枯渇率の値をとった（即ち、強力に枯渇したアレイは、全てのバイオロジカルレプリケートで強力に枯渇していなければならない）。各スペーサー標的について、アレイ枯渇率及び以下の特徴：標的鎖、転写物ターゲティング、ＯＲＩターゲティング、標的配列モチーフ、フランキング配列モチーフ、及び標的二次構造を含む行列を作成した。この行列中の異なる特徴がＣＬＵＳＴ．０９１９７９システムについての標的枯渇を説明する程度を調査した。

【０２７２】

図５及び図８は、所与の標的について、スクリーン出力対スクリーン入力におけるシーケンシングリードの規格化された比率をプロットすることによる、非コード配列と共にエンジニアリングされたＣＬＵＳＴ．０９１９７９組成物の干渉活性の程度を示す。結果は各ＤＲ転写方向につきプロットされる。組成物の機能的スクリーニングにおいて、活性ＲＮＡガイドと複合体を形成した活性エフェクターは、クロラムフェニコール及びテトラサイクリンに対する大腸菌（Ｅ．ｃｏｌｉ）耐性を付与するｐＡＣＹＣ１８４の能力に干渉し、細胞死及びプール内のスペーサーエレメントの枯渇をもたらす。初期ＤＮＡライブラリ（画面入力）と生存形質転換大腸菌（画面出力）のディープシーケンスの結果を比較すると、活性でプログラム可能なＣＲＩＳＰＲシステムを可能にする特定の標的配列及びＤＲ転写方向が示唆される。スクリーンはまた、エフェクター複合体がＤＲの１つの方向でのみ活性であることも示す。このように、スクリーンは、ＣＬＵＳＴ．０９１９７９　ＡＵＸＯ０１３９８８８２エフェクターがＤＲの「順」方向（５’－ＡＣＴＡ…ＡＡＣＴ－［スペーサー］－３’）で活性であったこと（図５）、及びＣＬＵＳＴ．０９１９７９　ＳＲＲ３１８１１５１エフェクターがＤＲの「逆」方向（５’－ＣＣＴＧ…ＣＡＡＣ－［スペーサー］－３’）で活性であったこと（図８）を示した。

【０２７３】

図６Ａ及び図６Ｂは、それぞれ、ｐＡＣＹＣ１８４及び大腸菌（Ｅ．ｃｏｌｉ）Ｅ．Ｃｌｏｎｉ必須遺伝子を標的化するＣＬＵＳＴ．０９１９７９　ＡＵＸＯ０１３９８８８２エフェクター（＋非コード配列）についての強力に枯渇した標的の位置を示す。同様に、図９Ａ及び図９Ｂは、それぞれ、ｐＡＣＹＣ１８４及び大腸菌（Ｅ．ｃｏｌｉ）Ｅ．Ｃｌｏｎｉ必須遺伝子を標的化するＣＬＵＳＴ．０９１９７９　ＳＲＲ３１８１１５１エフェクターについての強力に枯渇した標的の位置を示す。枯渇した標的の隣接配列を分析して、ＣＬＵＳＴ．０９１９７９　ＡＵＸＯ０１３９８８８２及びＣＬＵＳＴ．０９１９７９　ＳＲＲ３１８１１５１のＰＡＭ配列を決定した。ＣＬＵＳＴ．０９１９７９　ＡＵＸＯ０１３９８８８２及びＣＬＵＳＴ．０９１９７９　ＳＲＲ３１８１１５１のＰＡＭ配列のＷｅｂＬｏｇｏ表現（Ｃｒｏｏｋｓ　ｅｔ　ａｌ．，Ｇｅｎｏｍｅ　Ｒｅｓｅａｒｃｈ　１４：１１８８－９０，２００４）は、それぞれ図７及び図１０に示され、位置「２０」は、標的の５’末端に隣接するヌクレオチドに対応する。

【０２７４】

したがって、ＣＬＵＳＴ．０９１９７９　ＣＲＩＳＰＲ－Ｃａｓの複数のエフェクターがインビボで活性を示す。

【０２７５】

実施例５－ＣＬＵＳＴ．０９１９７９による哺乳動物遺伝子のターゲティング

この実施例は、一過性トランスフェクションによって哺乳動物細胞に導入されたＣＬＵＳＴ．０９１９７９からのヌクレアーゼを使用する複数の標的に対するインデル評価を説明する。

【０２７６】

配列番号４、配列番号８、及び配列番号１０のエフェクターをｐｃｄａ３．１骨格（Ｉ

ｎｖｉｔｒｏｇｅｎ）にクローニングした。次に、プラスミドをマキシプレップし、１μｇ／μＬに希釈した。ＲＮＡガイドの調製では、ｃｒＲＮＡをコードするｄｓＤＮＡ断片は、標的配列の足場及びＵ６プロモーターを含むウルトラマー（ｕｌｔｒａｍｅｒ）によって誘導された。ウルトラマーを１０ｍＭのＴｒｉｓ・ＨＣｌにｐＨ７．５で再懸濁し、最終ストック濃度を１００μＭにした。続いてワーキングストックを１０μＭに希釈し、再度１０ｍＭ　Ｔｒｉｓ・ＨＣｌを使用して、ＰＣＲ反応の鋳型として使用した。ｃｒＲＮＡの増幅は、次の成分を用いた５０μＬの反応物で行われた：前述の鋳型０．０２μｌ、フォワードプライマー２．５μｌ、リバースプライマー２．５μｌ、ＮＥＢ　ＨｉＦｉポリメラーゼ２５μＬ、及び水２０μｌ。サイクリング条件は、１×（９８℃で３０秒）、３０×（９８℃で１０秒、６７℃で１５秒）、１×（７２℃で２分）であった。ＰＣＲ産物は１．８Ｘ　ＳＰＲＩ処理でクリーンアップされ、２５ｎｇ／μＬに規格化された。調製されたｃｒＲＮＡ配列及びそれらの対応する標的配列は表１０に示される。配列番号２０５、配列番号２０７、配列番号２５２、配列番号２５４、配列番号２５６、配列番号２５８、配列番号２６０、配列番号２６２、配列番号２６４、配列番号２６６、配列番号２６８、配列番号２７０、配列番号２７２、配列番号２７４、及び配列番号２７６の成熟ｃｒＲＮＡのダイレクトリピート配列は配列番号６０に記載される。配列番号２０９及び配列番号２１４の成熟ｃｒＲＮＡのダイレクトリピートは配列番号６２に記載される。配列番号２１１、配列番号２７８、配列番号２８０、配列番号２８２、配列番号２８４、配列番号２８６、及び配列番号２８８の成熟ｃｒＲＮＡのダイレクトリピートは配列番号２１３に記載される。

【０２７７】

【表４３】

表10.一過性トランスフェクションアッセイのRNAガイド及び標的配列。

エフェクター配列	成熟crRNA配列	標的配列	PAM配列
配列番号4	CCTGTTGTGAATACTCTTTTATAGGTATCAAACAACGGAAGTGGTTGGTCAGCATGGATTA (配列番号配列番号205)	AAVS1: GGAAGTGGTTGGTCAGCATGGATTA (配列番号206)	5'-TTTG-3'
配列番号4	CCTGTTGTGAATACTCTTTTATAGGTATCAAACAACTGTGAAGTGACCTGGGAGCTAACTG (配列番号207)	VEGFA: TGTGAAGTGACCTGGGAGCTAACTG (配列番号208)	5-TTTG-3'
配列番号4	CCTGTTGTGAATACTCTTTTATAGGTATCAAACAACGAGAGGTGAGGGACTTGGGGGGTAA (配列番号252)	AAVS1: GAGAGGTGAGGGACTTGGGGGGTAA (配列番号253)	5-TTTG-3'
配列番号4	CCTGTTGTGAATACTCTTTTATAGGTATCAAACAACTGAGAATGGTGCGTCCTAGGTGTTC (配列番号254)	AAVS1: TGAGAATGGTGCGTCCTAGGTGTTC (配列番号255)	5-TTTG-3'
配列番号4	CCTGTTGTGAATACTCTTTTATAGGTATCAAACAACGCAGCCTGTGCTGACCCATGCAGTC (配列番号256)	AAVS1: GCAGCCTGTGCTGACCCATGCAGTC (配列番号257)	5-TTTG-3'
配列番号4	CCTGTTGTGAATACTCTTTTATAGGTATCAAACAACGGAAGTGGTTGGTCAGCATGGATTA (配列番号258)	AAVS1: GGAAGTGGTTGGTCAGCATGGATTA (配列番号259)	5-TTTG-3'
配列番号4	CCTGTTGTGAATACTCTTTTATAGGTATCAAACAACAGCCAGTGTTGCTAGTCAAGGGCAG (配列番号260)	EMX1: AGCCAGTGTTGCTAGTCAAGGGCAG (配列番号261)	5-TTTG-3'
配列番号4	CCTGTTGTGAATACTCTTTTATAGGTATCAAACAACTTGACATTGTCCACACCTGGAATCG (配列番号262)	VEGFA: TTGACATTGTCCACACCTGGAATCG (配列番号263)	5-TTTG-3'
配列番号4	CCTGTTGTGAATACTCTTTTATAGGTATCAAACAACGAAATCTATTGAGGCTCTGGAGAGA (配列番号264)	VEGFA: GAAATCTATTGAGGCTCTGGAGAGA (配列番号265)	5-TTTG-3'
配列番号4	CCTGTTGTGAATACTCTTTTATAGGTATCAAACAACGGAAGCTGGATGAGCCTGGTCCATG (配列番号266)	VEGFA: GGAAGCTGGATGAGCCTGGTCCATG (配列番号267)	5-TTTG-3'
配列番号4	CCTGTTGTGAATACTCTTTTATAGGTATCAAACAACCCCATACTGGGGACCAAGGAAGTGT (配列番号268)	VEGFA: CCCATACTGGGGACCAAGGAAGTGT (配列番号269)	5-TTTG-3'

【０２７８】

【表４４】

配列番号4	CCTGTTGTGAATACTCTTTTATA GGTATCAAACAACATGATGCTT TGCCGTAACCCTTCGT (配列番号270)	VEGFA: ATGATGCTTTGCCGTAACCC TTCGT (配列番号271)	5-TTTG-3’
配列番号4	CCTGTTGTGAATACTCTTTTATA GGTATCAAACAACAAGAGTCAT TGCCCCACTTTACCCT (配列番号272)	VEGFA: AAGAGTCATTGCCCCACTTT ACCCT (配列番号273)	5-TTTG-3’
配列番号4	CCTGTTGTGAATACTCTTTTATA GGTATCAAACAACGAGAGGTG AGGGACTTGGGGGGTAA (配列番号274)	AAVS1: GAGAGGTGAGGGACTTGGG GGGTAA (配列番号275)	5-TTTG-3’
配列番号4	CCTGTTGTGAATACTCTTTTATA GGTATCAAACAACGTGAAGTTC TAAACTTCATATTACC (配列番号276)	VEGFA: GTGAAGTTCTAAACTTCATA TTACC (配列番号277)	5-TTTG-3’
配列番号10	ATTGTTGTAGACACCTTTTTATA AGGATTGAACAACAACCCCCGT CTACCTGCCCACAGGG (配列番号209)	AAVS1: AACCCCCGTCTACCTGCCCA CAGGG (配列番号210)	5-ATTG-3’
配列番号10	CTTGTTGTATATGTCCTTTTATA GGTATTAAACAACGTAGAGGGA GAAATGGAATCCATAT (配列番号211)	AAVS1: GTAGAGGGAGAAATGGAAT CCATAT (配列番号212)	5’-GTTA-3’
配列番号10	ATTGTTGTAGACACCTTTTTATA AGGATTGAACAACGCACCAACG GGTAGATTTGGTGGTG (配列番号214)	VEGFA: GCACCAACGGGTAGATTTG GTGGTG (配列番号215)	5-ATTG-3’
配列番号10	CTTGTTGTATATGTCCTTTTATA GGTATTAAACAACGTAGAGGGA GAAATGGAATCCATAT (配列番号278)	AAVS1: GTAGAGGGAGAAATGGAAT CCATAT (配列番号279)	5’-GTTA-3’
配列番号10	CTTGTTGTATATGTCCTTTTATA GGTATTAAACAACGAGTCGCTT TAACTGGCCCTGGCTT (配列番号280)	AAVS1: GAGTCGCTTTAACTGGCCCT GGCTT (配列番号281)	5-ATTG-3’
配列番号10	CTTGTTGTATATGTCCTTTTATA GGTATTAAACAACTCCACACCT GGAATCGGCTTTCAGC (配列番号282)	VEGFA: TCCACACCTGGAATCGGCTT TCAGC (配列番号283)	5-ATTG-3’
配列番号10	CTTGTTGTATATGTCCTTTTATA GGTATTAAACAACAACCCCCGT CTACCTGCCCACAGGG (配列番号284)	AAVS1: AACCCCCGTCTACCTGCCCA CAGGG (配列番号285)	5-ATTG-3’
配列番号10	CTTGTTGTATATGTCCTTTTATA GGTATTAAACAACGTAGAGGGA GAAATGGAATCCATAT (配列番号286)	AAVS1: GTAGAGGGAGAAATGGAAT CCATAT (配列番号287)	5’-GTTA-3’
配列番号10	CTTGTTGTATATGTCCTTTTATA GGTATTAAACAACGACCCATGG GAGCAGCTGGTCAGAG (配列番号288)	EMX1: GACCCATGGGAGCAGCTGG TCAGAG (配列番号289)	5’-GTTA-3’

【０２７９】

トランスフェクションの約１６時間前に、ＤＭＥＭ／１０％ＦＢＳ＋Ｐｅｎ／Ｓｔｒｅｐ中の２５，０００個のＨＥＫ２９３Ｔ細胞１００μｌを９６ウェルプレートの各ウェルにプレーティングした。トランスフェクションの日、細胞は７０～９０％コンフルエントであった。トランスフェクトするウェルごとに、０．５μｌのＬｉｐｏｆｅｃｔａｍｉｎｅ２０００と９．５μｌのＯｐｔｉ－ＭＥＭの混合物を調製し、次に室温で５～２０分間インキュベートした（溶液１）。インキュベーション後、ｌｉｐｏｆｅｃｔａｍｉｎｅ：ＯｐｔｉＭＥＭ混合物を、１８２ｎｇのエフェクタープラスミド及び１４ｎｇのｃｒＲＮＡ及び最大１０μＬの水を含む別の混合物に添加した（溶液２）。陰性対照の場合、ｃｒ

ＲＮＡは溶液２に含まれていなかった。溶液１と溶液２の混合物をピペッティングにより上下に混合し、次に室温で２５分間インキュベートした。インキュベーション後、２０μＬの溶液１と溶液２の混合物を、細胞を含む９６ウェルプレートの各ウェルに滴下した。トランスフェクションの７２時間後、各ウェルの中央に１０μＬのＴｒｙｐＬＥを添加して細胞をトリプシン処理し、約５分間インキュベートする。次に、１００μＬのＤ１０培地を各ウェルに加え、混合して細胞を再懸濁した。次に細胞を５００ｇで１０分間スピンダウンし、上清を廃棄した。ＱｕｉｃｋＥｘｔｒａｃｔ緩衝液を、元の細胞懸濁液量の１／５に加えた。細胞を６５℃で１５分間、６８℃で１５分間、９８℃で１０分間インキュベートした。

【０２８０】

次世代シーケンシング用の試料は、２ラウンドのＰＣＲによって調製された。第１のラウンド（ＰＣＲ１）は、標的に応じて特定のゲノム領域を増幅するために使用された。ＰＣＲ１産物はカラム精製により精製した。ＰＣＲラウンド２（ＰＣＲ２）は、Ｉｌｌｕｍｉｎａのアダプター及びインデックスを付加するために行われた。次に、反応物をプールし、カラム精製によって精製した。シーケンシングの実行は、１５０サイクルのＮｅｘｔＳｅｑ　ｖ２．５の中出力又は高出力キットで行われた。

【０２８１】

図１１Ａ、図１１Ｂ、図１１Ｃ、及び図１１Ｄは、それぞれ、配列番号４又は配列番号１０のエフェクターでトランスフェクションした後のＨＥＫ２９３Ｔ細胞におけるＡＡＶＳ１、ＶＥＧＦＡ、及びＥＭＸ１標的遺伝子座のパーセントインデルを示す。バーは２つのバイオ複製物で測定されたインデルの平均パーセントを反映する。配列番号４及び配列番号１０のエフェクターについて、パーセントインデルは、標的のそれぞれにおいて陰性対照のパーセントインデルよりも高かった。

【０２８２】

図１１Ａに示すように、配列番号４のエフェクター及び配列番号２０５のｃｒＲＮＡによって形成される複合体は配列番号２０６のＡＡＶＳ１標的において活性であり、配列番号４のエフェクター及び配列番号２０７のｃｒＲＮＡによって形成される複合体は配列番号２０８のＶＥＧＦＡ標的において活性であった。図１１Ｂに示すように、配列番号４のエフェクター及び配列番号２５２のｃｒＲＮＡによって形成される複合体は配列番号２５３のＡＡＶＳ１標的において活性であり、配列番号４のエフェクター及び配列番号２５４のｃｒＲＮＡによって形成される複合体は配列番号２５５のＡＡＶＳ１標的において活性であり、配列番号４のエフェクター及び配列番号２５６のｃｒＲＮＡによって形成される複合体は配列番号２５７のＡＡＶＳ１標的において活性であり、配列番号４のエフェクター及び配列番号２５８のｃｒＲＮＡによって形成される複合体は配列番号２５９のＡＡＶＳ１標的において活性であり、配列番号４のエフェクター及び配列番号２７４のｃｒＲＮＡによって形成される複合体は配列番号２７５のＡＡＶＳ１標的において活性であった。図１１Ｂに示すように、配列番号４のエフェクター及び配列番号２６０のｃｒＲＮＡによって形成される複合体は配列番号２６１のＥＭＸ１標的において活性であった。同様に図１１Ｂに示すように、配列番号４のエフェクター及び配列番号２６２のｃｒＲＮＡによって形成される複合体は配列番号２６３のＶＥＧＦＡ１標的において活性であり、配列番号４のエフェクター及び配列番号２６４のｃｒＲＮＡによって形成される複合体は配列番号２６５のＶＥＧＦＡ１標的において活性であり、配列番号４のエフェクター及び配列番号２６６のｃｒＲＮＡによって形成される複合体は配列番号２６７のＶＥＧＦＡ１標的において活性であり、配列番号４のエフェクター及び配列番号２６８のｃｒＲＮＡによって形成される複合体は配列番号２６９のＶＥＧＦＡ１標的において活性であり、配列番号４のエフェクター及び配列番号２７０のｃｒＲＮＡによって形成される複合体は配列番号２７１のＶＥＧＦＡ１標的において活性であり、配列番号４のエフェクター及び配列番号２７２のｃｒＲＮＡによって形成される複合体は配列番号２７３のＶＥＧＦＡ１標的において活性であり、配列番号４のエフェクター及び配列番号２７４のｃｒＲＮＡによって形成される複合体は配列番号２７５のＶＥＧＦＡ１標的において活性であった。配列番号４のエ

フェクターは、図１１Ａ及び図１１Ｂにおける標的のそれぞれについて５’－ＴＴＴＧ－３’ＰＡＭを利用した。

【０２８３】

図１１Ｃに示すように、配列番号１０のエフェクター及び配列番号２０９のｃｒＲＮＡによって形成される複合体は配列番号２１０のＡＡＶＳ１標的において活性であり、配列番号１０のエフェクター及び配列番号２１１のｃｒＲＮＡによって形成される複合体は配列番号２１２のＡＡＶＳ１標的において活性であり、配列番号１０のエフェクター及び配列番号２１４のｃｒＲＮＡによって形成される複合体は配列番号２１５のＶＥＧＦＡ標的において活性であった。図１１Ｄに示すように、配列番号１０のエフェクター及び配列番号２７８のｃｒＲＮＡによって形成される複合体は配列番号２７９のＡＡＶＳ１標的において活性であり、配列番号１０のエフェクター及び配列番号２８０のｃｒＲＮＡによって形成される複合体は配列番号２８１のＡＡＶＳ１標的において活性であり、配列番号１０のエフェクター及び配列番号２８４のｃｒＲＮＡによって形成される複合体は配列番号２８５のＡＡＶＳ１標的において活性であり、配列番号１０のエフェクター及び配列番号２８６のｃｒＲＮＡによって形成される複合体は配列番号２８７のＡＡＶＳ１標的において活性であった。同様に図１１Ｄに示すように、配列番号１０のエフェクター及び配列番号２８８のｃｒＲＮＡによって形成される複合体は配列番号２８９のＥＭＸ１標的において活性であり、配列番号１０のエフェクター及び配列番号２８２のｃｒＲＮＡによって形成される複合体は配列番号２８３のＶＥＧＦＡ標的において活性であった。配列番号１０のエフェクターは、図１１Ｃ及び図１１Ｄにおける標的について５’－ＡＴＴＧ－３’ＰＡＭ及び５’－ＧＴＴＡ－３’ＰＡＭを利用した。

【０２８４】

この実施例は、ＣＬＵＳＴ．０９１９７９ファミリーのヌクレアーゼが哺乳動物細胞において活性を有することを示唆している。

【０２８５】

他の実施形態

本発明はその詳細な説明を伴い説明されているが、前述の説明は例示であり、添付の特許請求の範囲によって定義される本発明の範囲を限定する意図はないことが理解されるべきである。他の態様、利点、及び変形例が、以下の特許請求の範囲内にある。

特定の実施形態では、例えば以下の項目が提供される。

（項目１）

ＣＬＵＳＴ．０９１９７９のエンジニアリングされた天然に存在しないクラスター化して規則的な配置の短い回文配列リピート（ＣＲＩＳＰＲ）－Ｃａｓシステムであって、

（ａ）ＣＲＩＳＰＲ関連タンパク質が配列番号２４１のアミノ酸配列を含む、ＣＲＩＳＰＲ関連タンパク質又は前記ＣＲＩＳＰＲ関連タンパク質をコードする核酸；及び

（ｂ）ダイレクトリピート配列と標的核酸へのハイブリダイゼーション能を有するスペーサー配列とを含むＲＮＡガイド

を含み、

前記ＣＲＩＳＰＲ関連タンパク質が、前記ＲＮＡガイドに結合し、前記スペーサー配列に相補的な前記標的核酸配列を修飾することができる、ＣＲＩＳＰＲ－Ｃａｓシステム。

（項目２）

前記ＣＲＩＳＰＲ関連タンパク質が、配列番号４、配列番号１０、配列番号１２、又は配列番号１４に記載のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含む、項目１に記載のシステム。

（項目３）

ＣＬＵＳＴ．０９１９７９のエンジニアリングされた天然に存在しないクラスター化して規則的な配置の短い回文配列リピート（ＣＲＩＳＰＲ）－Ｃａｓシステムであって、

（ａ）ＣＲＩＳＰＲ関連タンパク質が配列番号１～５６のいずれか１つに記載のアミノ酸

配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含む、ＣＲＩＳＰＲ関連タンパク質又は前記ＣＲＩＳＰＲ関連タンパク質をコードする核酸；及び
（ｂ）ダイレクトリピート配列と標的核酸へのハイブリダイゼーション能を有するスペーサー配列とを含むＲＮＡガイド
を含み、
前記ＣＲＩＳＰＲ関連タンパク質が、前記ＲＮＡガイドに結合し、前記スペーサー配列に相補的な前記標的核酸配列を修飾することができる、ＣＲＩＳＰＲ－Ｃａｓシステム。
（項目４）
前記ＣＲＩＳＰＲ関連タンパク質が、少なくとも１つのＲｕｖＣドメイン又は少なくとも１つの分割されたＲｕｖＣドメインを含む、項目３に記載のシステム。
（項目５）
前記ＣＲＩＳＰＲ関連タンパク質が、以下の配列の１つ以上：
（ａ）ＰＸ$_1$Ｘ$_2$Ｘ$_3$Ｘ$_4$Ｆ（配列番号２１６）（ここで、Ｘ$_1$はＬ又はＭ又はＩ又はＣ又はＦであり、Ｘ$_2$はＹ又はＷ又はＦであり、Ｘ$_3$はＫ又はＴ又はＣ又はＲ又はＷ又はＹ又はＨ又はＶであり、Ｘ$_4$はＩ又はＬ又はＭである）；
（ｂ）ＲＸ$_1$Ｘ$_2$Ｘ$_3$Ｌ（配列番号２１７）（ここで、Ｘ$_1$はＩ又はＬ又はＭ又はＹ又はＴ又はＦであり、Ｘ$_2$はＲ又はＱ又はＫ又はＥ又はＳ又はＴであり、Ｘ$_3$はＬ又はＩ又はＴ又はＣ又はＭ又はＫである）；
（ｃ）ＮＸ$_1$ＹＸ$_2$（配列番号２１８）（ここで、Ｘ$_1$はＩ又はＬ又はＦであり、Ｘ$_2$はＫ又はＲ又はＶ又はＥである）；
（ｄ）ＫＸ$_1$Ｘ$_2$Ｘ$_3$ＦＡＸ$_4$Ｘ$_5$ＫＤ（配列番号２１９）（ここで、Ｘ$_1$はＴ又はＩ又はＮ又はＡ又はＳ又はＦ又はＶであり、Ｘ$_2$はＩ又はＶ又はＬ又はＳであり、Ｘ$_3$はＨ又はＳ又はＧ又はＲであり、Ｘ$_4$はＤ又はＳ又はＥであり、Ｘ$_5$はＩ又はＶ又はＭ又はＴ又はＮである）；
（ｅ）ＬＸ$_1$ＮＸ$_2$（配列番号２２０）（ここで、Ｘ$_1$はＧ又はＳ又はＣ又はＴであり、Ｘ$_2$はＮ又はＹ又はＫ又はＳである）；
（ｆ）ＰＸ$_1$Ｘ$_2$Ｘ$_3$Ｘ$_4$ＳＱＸ$_5$ＤＳ（配列番号２２１）（ここで、Ｘ$_1$はＳ又はＰ又はＡであり、Ｘ$_2$はＹ又はＳ又はＡ又はＰ又はＥ又はＹ又はＱ又はＮであり、Ｘ$_3$はＦ又はＹ又はＨであり、Ｘ$_4$はＴ又はＳであり、Ｘ$_5$はＭ又はＴ又はＩである）；
（ｇ）ＫＸ$_1$Ｘ$_2$ＶＲＸ$_3$Ｘ$_4$ＱＥＸ$_5$Ｈ（配列番号２２２）（ここで、Ｘ$_1$はＮ又はＫ又はＷ又はＲ又はＥ又はＴ又はＹであり、Ｘ$_2$はＭ又はＲ又はＬ又はＳ又はＫ又はＶ又はＥ又はＴ又はＩ又はＤであり、Ｘ$_3$はＬ又はＲ又はＨ又はＰ又はＴ又はＫ又はＰのＱ又はＳ又はＡであり、Ｘ$_4$はＧ又はＱ又はＮ又はＲ又はＫ又はＥ又はＩ又はＴ又はＳ又はＣであり、Ｘ$_5$はＲ又はＷ又はＹ又はＫ又はＴ又はＦ又はＳ又はＱである）；及び
（ｈ）Ｘ$_1$ＮＧＸ$_2$Ｘ$_3$Ｘ$_4$ＤＸ$_5$ＮＸ$_6$Ｘ$_7$Ｘ$_8$Ｎ（配列番号２２３）（ここで、Ｘ$_1$はＩ又はＫ又はＶ又はＬであり、Ｘ$_2$はＬ又はＭであり、Ｘ$_3$はＮ又はＨ又はＰであり、Ｘ$_4$はＡ又はＳ又はＣであり、Ｘ$_5$はＶ又はＹ又はＩ又はＦ又はＴ又はＮであり、Ｘ$_6$はＡ又はＳであり、Ｘ$_7$はＳ又はＡ又はＰであり、Ｘ$_8$はＭ又はＣ又はＬ又はＲ又はＮ又はＳ又はＫ又はＬである）
を含む、項目３又は４に記載のシステム。
（項目６）
前記ダイレクトリピート配列が、配列番号５７～９０、配列番号１１８～１５１、又は配列番号２１３のいずれか１つに記載のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む、項目３～５のいずれか一項に記載のシステム。
（項目７）
前記ダイレクトリピート配列が、配列番号５７～９０、配列番号１１８～１５１、又は

配列番号２１３のいずれか１つに記載のヌクレオチド配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む、項目６に記載のシステム。

（項目８）

前記ダイレクトリピート配列が、以下の配列の１つ以上：

（ａ）$X_1X_2TX_3X_4X_5X_6X_7X_8$（配列番号２２４）（ここで、$X_1$はＡ又はＣ又はＧであり、$X_2$はＴ又はＣ又はＡであり、$X_3$はＴ又はＧ又はＡであり、$X_4$はＴ又はＧであり、$X_5$はＴ又はＧ又はＡであり、$X_6$はＧ又はＴ又はＡであり、$X_7$はＴ又はＧ又はＡであり、$X_8$はＡ又はＧ又はＴである）；

（ｂ）$X_1X_2X_3X_4X_5X_6X_7X_8X_9$（配列番号２２６）（ここで、$X_1$はＴ又はＣ又はＡであり、$X_2$はＴ又はＡ又はＧであり、$X_3$はＴ又はＣ又はＡであり、$X_4$はＴ又はＡであり、$X_5$はＴ又はＡ又はＧであり、$X_6$はＴ又はＡであり、$X_7$はＡ又はＴであり、$X_8$はＡ又はＧ又はＣ又はＴであり、$X_9$はＧ又はＡ又はＣである）；及び

（ｃ）$X_1X_2X_3AC$（配列番号２２８）（ここで、X_1はＡ又はＣ又はＧであり、X_2はＣ又はＡであり、X_3はＡ又はＣである）

を含む、項目３～７のいずれか一項に記載のシステム。

（項目９）

前記ＣＲＩＳＰＲ関連タンパク質が、配列番号１に記載のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、前記ダイレクトリピート配列が、配列番号５７に記載のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む、項目３～８のいずれか一項に記載のシステム。

（項目１０）

前記ＣＲＩＳＰＲ関連タンパク質が、配列番号１に記載のアミノ酸配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、前記ダイレクトリピート配列が、配列番号５７に記載のヌクレオチド配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む、項目９に記載のシステム。

（項目１１）

前記ＣＲＩＳＰＲ関連タンパク質が、配列番号１に記載のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、前記ＣＲＩＳＰＲ関連タンパク質が、プロトスペーサー隣接モチーフ（ＰＡＭ）配列の認識能を有し、前記ＰＡＭ配列が、５’－ＴＮＮＴ－３’又は５’－ＴＮＲＴ－３’として記載される核酸配列を含み、「Ｎ」は任意のヌクレオチドであり、「Ｒ」はＡ又はＧである、項目３～８のいずれか一項に記載のシステム。

（項目１２）

前記ＣＲＩＳＰＲ関連タンパク質が、配列番号１に記載のアミノ酸配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）同一性を有するタンパク質であり、前記ＣＲＩＳＰＲ関連タンパク質が、ＰＡＭ配列の認識能を有し、前記ＰＡＭ配列が、５’－ＴＮＮＴ－３’又は５’－ＴＮＲＴ－３’として記載される核酸配列を含み、「Ｎ」は任意のヌクレオチドであり、「Ｒ」はＡ又はＧである、項目１１に記載のシステム。

（項目１３）

前記ＣＲＩＳＰＲ関連タンパク質が、配列番号４に記載のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８

９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一性を有するタンパク質であり、前記ダイレクトリピート配列が、配列番号６０に記載のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む、項目３～８のいずれか一項に記載のシステム。

（項目１４）

前記ＣＲＩＳＰＲ関連タンパク質が、配列番号４に記載のアミノ酸配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、前記ダイレクトリピート配列が、配列番号６０に記載のヌクレオチド配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む、項目１３に記載のシステム。

（項目１５）

前記ＣＲＩＳＰＲ関連タンパク質が、配列番号４に記載のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、前記ＣＲＩＳＰＲ関連タンパク質が、ＰＡＭ配列の認識能を有し、前記ＰＡＭ配列が、５’－ＮＴＴＮ－３’、５’－ＮＴＴＲ－３’（例えば、５’－ＴＴＴＧ－３’）、又は５’－ＮＮＲ－３’として記載される核酸配列を含み、「Ｎ」は任意のヌクレオチドであり、「Ｒ」はＡ又はＧである、項目３～８のいずれか一項に記載のシステム。

（項目１６）

前記ＣＲＩＳＰＲ関連タンパク質が、配列番号４に記載のアミノ酸配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、前記ＣＲＩＳＰＲ関連タンパク質が、ＰＡＭ配列を認識することができ、前記ＰＡＭ配列が、５’－ＮＴＴＮ－３’、５’－ＮＴＴＲ－３’（例えば、５’－ＴＴＴＧ－３’）、又は５’－ＮＮＲ－３’として記載される核酸配列を含み、ここで、「Ｎ」は任意のヌクレオチドであり、「Ｒ」はＡ又はＧである、項目１５に記載のシステム。

（項目１７）

前記ＣＲＩＳＰＲ関連タンパク質が、配列番号１０に記載のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、前記ダイレクトリピート配列が、配列番号６２又は配列番号２１３に記載のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む、項目３～８のいずれか一項に記載のシステム。

（項目１８）

前記ＣＲＩＳＰＲ関連タンパク質が、配列番号１０に記載のアミノ酸配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、前記ダイレクトリピート配列が、配列番号６２又は配列番号２１３に記載のヌクレオチド配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む、項目１７に記載のシステム。

（項目１９）

前記ＣＲＩＳＰＲ関連タンパク質が、配列番号１０に記載のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、前記ＣＲＩＳＰＲ関連タンパク質が、ＰＡＭ配列の認識能を有し、前記ＰＡＭ配列が、５’－ＮＴＴＮ－３’又は５’－ＲＴＴＲ－３’（例えば、５’－ＡＴＴＧ－３’又は５’－ＧＴＴＡ－３’）として記載される核

酸配列を含み、「Ｎ」は任意のヌクレオチドであり、「Ｒ」はＡ又はＧである、項目３～８のいずれか一項に記載のシステム。
（項目２０）
前記ＣＲＩＳＰＲ関連タンパク質が、配列番号１０に記載のアミノ酸配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、前記ＣＲＩＳＰＲ関連タンパク質が、ＰＡＭ配列の認識能を有し、前記ＰＡＭ配列が、５’－ＮＴＴＮ－３’又は５’－ＲＴＴＲ－３’（例えば、５’－ＡＴＴＧ－３’又は５’－ＧＴＴＡ－３’）として記載される核酸配列を含み、「Ｎ」は任意のヌクレオチドであり、「Ｒ」はＡ又はＧである、項目１９に記載のシステム。
（項目２１）
前記ＲＮＡガイドの前記スペーサー配列が、約１５ヌクレオチド～約５５ヌクレオチドを含む、項目１～２０のいずれか一項に記載のシステム。
（項目２２）
前記ＲＮＡガイドの前記スペーサー配列が、２０～４５ヌクレオチドを含む、項目２１に記載のシステム。
（項目２３）
前記ＣＲＩＳＰＲ関連タンパク質が、触媒残基（例えば、アスパラギン酸又はグルタミン酸）を含む、項目１～２２のいずれか一項に記載のシステム。
（項目２４）
前記ＣＲＩＳＰＲ関連タンパク質が、前記標的核酸を切断する、項目１～２３のいずれか一項に記載のシステム。
（項目２５）
前記ＣＲＩＳＰＲ関連タンパク質が、ペプチドタグ、蛍光タンパク質、塩基編集ドメイン、ＤＮＡメチル化ドメイン、ヒストン残基修飾ドメイン、局在化因子、転写修飾因子、光ゲート制御因子、化学誘導性因子、又はクロマチン可視化因子を更に含む、項目１～２４のいずれか一項に記載のシステム。
（項目２６）
前記ＣＲＩＳＰＲ関連タンパク質をコードする前記核酸が、細胞での発現にコドン最適化される、項目１～２５のいずれか一項に記載のシステム。
（項目２７）
前記ＣＲＩＳＰＲ関連タンパク質をコードする前記核酸が、プロモーターに作動可能に連結されている、項目１～２６のいずれか一項に記載のシステム。
（項目２８）
前記ＣＲＩＳＰＲ関連タンパク質をコードする前記核酸が、ベクター内にある、項目１～２７のいずれか一項に記載のシステム。
（項目２９）
前記ベクターが、レトロウイルスベクター、レンチウイルスベクター、ファージベクター、アデノウイルスベクター、アデノ随伴ベクター、又は単純ヘルペスベクターを含む、項目２８に記載のシステム。
（項目３０）
前記標的核酸がＤＮＡ分子である、項目１～２９のいずれか一項に記載のシステム。
（項目３１）
前記ＣＲＩＳＰＲ関連タンパク質が、非特異的ヌクレアーゼ活性を含む、項目１～３０のいずれか一項に記載のシステム。
（項目３２）
前記ＣＲＩＳＰＲ関連タンパク質及びＲＮＡガイドによる前記標的核酸の認識により、前記標的核酸の修飾が生じる、項目１～３１のいずれか一項に記載のシステム。
（項目３３）
前記標的核酸の前記修飾が、二本鎖切断イベントである、項目３２に記載のシステム。
（項目３４）

前記標的核酸の前記修飾が、一本鎖切断イベントである、項目３２に記載のシステム。
（項目３５）
前記標的核酸の前記修飾により、挿入イベントが生じる、項目３２に記載のシステム。
（項目３６）
前記標的核酸の前記修飾により、欠失イベントが生じる、項目３２に記載のシステム。
（項目３７）
前記標的核酸の前記修飾により細胞毒性又は細胞死が生じる、項目３２～３６のいずれか一項に記載のシステム。
（項目３８）
ドナー鋳型核酸を更に含む、項目１～３０のいずれか一項に記載のシステム。
（項目３９）
前記ドナー鋳型核酸がＤＮＡ分子である、項目３８に記載のシステム。
（項目４０）
前記ドナー鋳型核酸がＲＮＡ分子である、項目３８に記載のシステム。
（項目４１）
前記ＲＮＡガイドが任意選択でｔｒａｃｒＲＮＡを含む、項目１～４０のいずれか一項に記載のシステム。
（項目４２）
前記システムがｔｒａｃｒＲＮＡを含まない、項目１～４０のいずれか一項に記載のシステム。
（項目４３）
前記ＣＲＩＳＰＲ関連タンパク質が自己プロセシングである、項目１～４２のいずれか一項に記載のシステム。
（項目４４）
前記システムが、ナノ粒子、リポソーム、エキソソーム、微小胞、又は遺伝子銃を含む送達組成物中に存在する、項目１～４３のいずれか一項に記載のシステム。
（項目４５）
細胞内にある、項目１～４３のいずれか一項に記載のシステム。
（項目４６）
前記細胞が真核細胞である、項目４５に記載のシステム。
（項目４７）
前記細胞が原核細胞である、項目４５に記載のシステム。
（項目４８）
（ａ）ＣＲＩＳＰＲ関連タンパク質が配列番号１～５６のいずれか１つに記載のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含む、ＣＲＩＳＰＲ関連タンパク質又は前記ＣＲＩＳＰＲ関連タンパク質をコードする核酸；及び
（ｂ）ダイレクトリピート配列と標的核酸へのハイブリダイゼーション能を有するスペーサー配列とを含むＲＮＡガイド
を含む、細胞。
（項目４９）
前記ＣＲＩＳＰＲ関連タンパク質が、以下の配列の１つ以上：
（ａ）ＰＸ$_1$Ｘ$_2$Ｘ$_3$Ｘ$_4$Ｆ（配列番号２１６）（ここで、Ｘ$_1$はＬ又はＭ又はＩ又はＣ又はＦであり、Ｘ$_2$はＹ又はＷ又はＦであり、Ｘ$_3$はＫ又はＴ又はＣ又はＲ又はＷ又はＹ又はＨ又はＶであり、Ｘ$_4$はＩ又はＬ又はＭである）；
（ｂ）ＲＸ$_1$Ｘ$_2$Ｘ$_3$Ｌ（配列番号２１７）（ここで、Ｘ$_1$はＩ又はＬ又はＭ又はＹ又はＴ又はＦであり、Ｘ$_2$はＲ又はＱ又はＫ又はＥ又はＳ又はＴであり、Ｘ$_3$はＬ又はＩ又はＴ又はＣ又はＭ又はＫである）；
（ｃ）ＮＸ$_1$ＹＸ$_2$（配列番号２１８）（ここで、Ｘ$_1$はＩ又はＬ又はＦであり、Ｘ$_2$はＫ

又はＲ又はＶ又はＥである）；
（ｄ）ＫＸ$_1$Ｘ$_2$Ｘ$_3$ＦＡＸ$_4$Ｘ$_5$ＫＤ（配列番号２１９）（ここで、Ｘ$_1$はＴ又はＩ又はＮ又はＡ又はＳ又はＦ又はＶであり、Ｘ$_2$はＩ又はＶ又はＬ又はＳであり、Ｘ$_3$はＨ又はＳ又はＧ又はＲであり、Ｘ$_4$はＤ又はＳ又はＥであり、Ｘ$_5$はＩ又はＶ又はＭ又はＴ又はＮである）；
（ｅ）ＬＸ$_1$ＮＸ$_2$（配列番号２２０）（ここで、Ｘ$_1$はＧ又はＳ又はＣ又はＴであり、Ｘ$_2$はＮ又はＹ又はＫ又はＳである）；
（ｆ）ＰＸ$_1$Ｘ$_2$Ｘ$_3$Ｘ$_4$ＳＱＸ$_5$ＤＳ（配列番号２２１）（ここで、Ｘ$_1$はＳ又はＰ又はＡであり、Ｘ$_2$はＹ又はＳ又はＡ又はＰ又はＥ又はＹ又はＱ又はＮであり、Ｘ$_3$はＦ又はＹ又はＨであり、Ｘ$_4$はＴ又はＳであり、Ｘ$_5$はＭ又はＴ又はＩである）；
（ｇ）ＫＸ$_1$Ｘ$_2$ＶＲＸ$_3$Ｘ$_4$ＱＥＸ$_5$Ｈ（配列番号２２２）（ここで、Ｘ$_1$はＮ又はＫ又はＷ又はＲ又はＥ又はＴ又はＹであり、Ｘ$_2$はＭ又はＲ又はＬ又はＳ又はＫ又はＶ又はＥ又はＴ又はＩ又はＤであり、Ｘ$_3$はＬ又はＲ又はＨ又はＰ又はＴ又はＫ又はＰのＱ又はＳ又はＡであり、Ｘ$_4$はＧ又はＱ又はＮ又はＲ又はＫ又はＥ又はＩ又はＴ又はＳ又はＣであり、Ｘ$_5$はＲ又はＷ又はＹ又はＫ又はＴ又はＦ又はＳ又はＱである）；及び
（ｈ）Ｘ$_1$ＮＧＸ$_2$Ｘ$_3$Ｘ$_4$ＤＸ$_5$ＮＸ$_6$Ｘ$_7$Ｘ$_8$Ｎ（配列番号２２３）（ここで、Ｘ$_1$はＩ又はＫ又はＶ又はＬであり、Ｘ$_2$はＬ又はＭであり、Ｘ$_3$はＮ又はＨ又はＰであり、Ｘ$_4$はＡ又はＳ又はＣであり、Ｘ$_5$はＶ又はＹ又はＩ又はＦ又はＴ又はＮであり、Ｘ$_6$はＡ又はＳであり、Ｘ$_7$はＳ又はＡ又はＰであり、Ｘ$_8$はＭ又はＣ又はＬ又はＲ又はＮ又はＳ又はＫ又はＬである）
を含む、項目４８に記載の細胞。
（項目５０）
前記ダイレクトリピート配列が、配列番号５７～９０、配列番号１１８～１５１、又は配列番号２１３のいずれか１つに記載のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む、項目４８又は４９に記載の細胞。
（項目５１）
前記ダイレクトリピート配列が、配列番号５７～９０、配列番号１１８～１５１、又は配列番号２１３のいずれか１つに記載のヌクレオチド配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む、項目５０に記載の細胞。
（項目５２）
前記ダイレクトリピート配列が、以下の配列の１つ以上：
（ａ）Ｘ$_1$Ｘ$_2$ＴＸ$_3$Ｘ$_4$Ｘ$_5$Ｘ$_6$Ｘ$_7$Ｘ$_8$（配列番号２２４）（ここで、Ｘ$_1$はＡ又はＣ又はＧであり、Ｘ$_2$はＴ又はＣ又はＡであり、Ｘ$_3$はＴ又はＧ又はＡであり、Ｘ$_4$はＴ又はＧであり、Ｘ$_5$はＴ又はＧ又はＡであり、Ｘ$_6$はＧ又はＴ又はＡであり、Ｘ$_7$はＴ又はＧ又はＡであり、Ｘ$_8$はＡ又はＧ又はＴである）；
（ｂ）Ｘ$_1$Ｘ$_2$Ｘ$_3$Ｘ$_4$Ｘ$_5$Ｘ$_6$Ｘ$_7$Ｘ$_8$Ｘ$_9$（配列番号２２６）（ここで、Ｘ$_1$はＴ又はＣ又はＡであり、Ｘ$_2$はＴ又はＡ又はＧであり、Ｘ$_3$はＴ又はＣ又はＡであり、Ｘ$_4$はＴ又はＡであり、Ｘ$_5$はＴ又はＡ又はＧであり、Ｘ$_6$はＴ又はＡであり、Ｘ$_7$はＡ又はＴであり、Ｘ$_8$はＡ又はＧ又はＣ又はＴであり、Ｘ$_9$はＧ又はＡ又はＣである）；及び
（ｃ）Ｘ$_1$Ｘ$_2$Ｘ$_3$ＡＣ（配列番号２２８）（ここで、Ｘ$_1$はＡ又はＣ又はＧであり、Ｘ$_2$はＣ又はＡであり、Ｘ$_3$はＡ又はＣである）
を含む、項目４８～５１のいずれか一項に記載の細胞。
（項目５３）
前記ＣＲＩＳＰＲ関連タンパク質が、配列番号１に記載のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、前記ダイレクトリピート配列が

、配列番号５７に記載のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む、項目４８～５２のいずれか一項に記載の細胞。

（項目５４）

前記ＣＲＩＳＰＲ関連タンパク質が、配列番号１に記載のアミノ酸配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、前記ダイレクトリピート配列が、配列番号５７に記載のヌクレオチド配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む、項目５３に記載の細胞。

（項目５５）

前記ＣＲＩＳＰＲ関連タンパク質が、配列番号１に記載のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、前記ＣＲＩＳＰＲ関連タンパク質が、ＰＡＭ配列の認識能を有し、前記ＰＡＭ配列が、５’－ＴＮＮＴ－３’又は５’－ＴＮＲＴ－３’として記載される核酸配列を含み、「Ｎ」が任意のヌクレオチドであり、「Ｒ」がＡ又はＧである、項目４８～５２のいずれか一項に記載の細胞。

（項目５６）

前記ＣＲＩＳＰＲ関連タンパク質が、配列番号１に記載のアミノ酸配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、前記ＣＲＩＳＰＲ関連タンパク質がＰＡＭ配列の認識能を有し、前記ＰＡＭ配列が、５’－ＴＮＮＴ－３’又は５’－ＴＮＲＴ－３’として記載される核酸配列を含み、「Ｎ」が任意のヌクレオチドであり、「Ｒ」がＡ又はＧである、項目５５に記載の細胞。

（項目５７）

前記ＣＲＩＳＰＲ関連タンパク質が、配列番号４に記載のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、前記ダイレクトリピート配列が、配列番号６０に記載のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む、項目４８～５２のいずれか一項に記載の細胞。

（項目５８）

前記ＣＲＩＳＰＲ関連タンパク質が、配列番号４に記載のアミノ酸配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、前記ダイレクトリピート配列が、配列番号６０に記載のヌクレオチド配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む、項目５７に記載の細胞。

（項目５９）

前記ＣＲＩＳＰＲ関連タンパク質が、配列番号４に記載のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、前記ＣＲＩＳＰＲ関連タンパク質がＰＡＭ配列の認識能を有し、前記ＰＡＭ配列が、５’－ＮＴＴＮ－３’、５’－ＮＴＴＲ－３’（例えば、５’－ＴＴＴＧ－３’）、又は５’－ＮＮＲ－３’として記載される核酸配列を含み、「Ｎ」が任意のヌクレオチドであり、「Ｒ」がＡ又はＧである、項目４８～５２のいずれか一項に記載の細胞。

（項目６０）

前記ＣＲＩＳＰＲ関連タンパク質が、配列番号４に記載のアミノ酸配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、前記ＣＲＩＳＰＲ関連タンパク質がＰＡＭ配列の認識能を有し、前記ＰＡＭ配列が、５’－ＮＴＴＮ－３’、５’－ＮＴＴＲ－３’（例えば、５’－ＴＴＴＧ－３’）、又は５’－ＮＮＲ－３’として記載される核酸配列を含み、「Ｎ」が任意のヌクレオチドであり、「Ｒ」がＡ又はＧである、項目５９に記載の細胞。

（項目６１）

前記ＣＲＩＳＰＲ関連タンパク質が、配列番号１０に記載のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、前記ダイレクトリピート配列が、配列番号６２又は配列番号２１３に記載のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む、項目４８～５２のいずれか一項に記載の細胞。

（項目６２）

前記ＣＲＩＳＰＲ関連タンパク質が、配列番号１０に記載のアミノ酸配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、前記ダイレクトリピート配列が、配列番号６２又は配列番号２１３に記載のヌクレオチド配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む、項目６１に記載の細胞。

（項目６３）

前記ＣＲＩＳＰＲ関連タンパク質が、配列番号１０に記載のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、前記ＣＲＩＳＰＲ関連タンパク質が、ＰＡＭ配列の認識能を有し、前記ＰＡＭ配列が、５’－ＮＴＴＮ－３’又は５’－ＲＴＴＲ－３’（例えば、５’－ＡＴＴＧ－３’又は５’－ＧＴＴＡ－３’）として記載される核酸配列を含み、「Ｎ」が任意のヌクレオチドであり、「Ｒ」がＡ又はＧである、項目４８～５２のいずれか一項に記載の細胞。

（項目６４）

前記ＣＲＩＳＰＲ関連タンパク質が、配列番号１０に記載のアミノ酸配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、前記ＣＲＩＳＰＲ関連タンパク質がＰＡＭ配列の認識能を有し、前記ＰＡＭ配列が、５’－ＮＴＴＮ－３’又は５’－ＲＴＴＲ－３’（例えば、５’－ＡＴＴＧ－３’又は５’－ＧＴＴＡ－３’）として記載される核酸配列を含み、「Ｎ」が任意のヌクレオチドであり、「Ｒ」がＡ又はＧである、項目６３に記載の細胞。

（項目６５）

前記スペーサー配列が、約１５ヌクレオチド～約５５ヌクレオチドを含む、項目４８～６４のいずれか一項に記載の細胞。

（項目６６）

前記スペーサー配列が、２０～４５ヌクレオチドを含む、項目６５に記載の細胞。

（項目６７）

前記細胞がｔｒａｃｒＲＮＡを更に含む、項目４８～６６のいずれか一項に記載の細胞。

（項目６８）

前記システムがｔｒａｃｒＲＮＡを含まない、項目４８～６６のいずれか一項に記載の細胞。

（項目６９）

前記細胞が真核細胞、例えば、哺乳動物細胞、例えば、ヒト細胞である、項目４８～６８のいずれか一項に記載の細胞。

（項目７０）
前記細胞が原核細胞である、項目４８～６９のいずれか一項に記載の細胞。
（項目７１）
項目１～４７のいずれか一項に記載のシステムを、細胞内の標的核酸に結合させる方法であって、
（ａ）前記システムを提供すること；及び
（ｂ）前記システムを前記細胞に送達すること
を含み、前記細胞が前記標的核酸を含み、前記ＣＲＩＳＰＲ関連タンパク質が前記ＲＮＡガイドに結合し、前記スペーサー配列が前記標的核酸に結合する、方法。
（項目７２）
前記細胞が真核細胞、例えば、哺乳動物細胞、例えば、ヒト細胞である、項目７１に記載の方法。
（項目７３）
標的核酸を修飾する方法であって、
（ａ）ＣＲＩＳＰＲ関連タンパク質が配列番号１～５６のいずれか１つに記載のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含む、ＣＲＩＳＰＲ関連タンパク質又は前記ＣＲＩＳＰＲ関連タンパク質をコードする核酸；及び
（ｂ）ダイレクトリピート配列と前記標的核酸へのハイブリダイゼーション能を有するスペーサー配列とを含むＲＮＡガイド；
を含み、
前記ＣＲＩＳＰＲ関連タンパク質がＲＮＡガイドへの結合能を有し、
前記ＣＲＩＳＰＲ関連タンパク質及びＲＮＡガイドによる前記標的核酸の認識により、前記標的核酸の修飾が生じる、
エンジニアリングされた天然に存在しないＣＲＩＳＰＲ－Ｃａｓシステムを、標的核酸に送達することを含む方法。
（項目７４）
前記ＣＲＩＳＰＲ関連タンパク質が、以下の配列の１つ以上：
（ａ）$PX_1X_2X_3X_4F$（配列番号２１６）（ここで、X_1はＬ又はＭ又はＩ又はＣ又はＦであり、X_2はＹ又はＷ又はＦであり、X_3はＫ又はＴ又はＣ又はＲ又はＷ又はＹ又はＨ又はＶであり、X_4はＩ又はＬ又はＭである）；
（ｂ）$RX_1X_2X_3L$（配列番号２１７）（ここで、X_1はＩ又はＬ又はＭ又はＹ又はＴ又はＦであり、X_2はＲ又はＱ又はＫ又はＥ又はＳ又はＴであり、X_3はＬ又はＩ又はＴ又はＣ又はＭ又はＫである）；
（ｃ）NX_1YX_2（配列番号２１８）（ここで、X_1はＩ又はＬ又はＦであり、X_2はＫ又はＲ又はＶ又はＥである）；
（ｄ）$KX_1X_2X_3FAX_4X_5KD$（配列番号２１９）（ここで、X_1はＴ又はＩ又はＮ又はＡ又はＳ又はＦ又はＶであり、X_2はＩ又はＶ又はＬ又はＳであり、X_3はＨ又はＳ又はＧ又はＲであり、X_4はＤ又はＳ又はＥであり、X_5はＩ又はＶ又はＭ又はＴ又はＮである）；
（ｅ）LX_1NX_2（配列番号２２０）（ここで、X_1はＧ又はＳ又はＣ又はＴであり、X_2はＮ又はＹ又はＫ又はＳである）；
（ｆ）$PX_1X_2X_3X_4SQX_5DS$（配列番号２２１）（ここで、X_1はＳ又はＰ又はＡであり、X_2はＹ又はＳ又はＡ又はＰ又はＥ又はＹ又はＱ又はＮであり、X_3はＦ又はＹ又はＨであり、X_4はＴ又はＳであり、X_5はＭ又はＴ又はＩである）；
（ｇ）$KX_1X_2VRX_3X_4QEX_5H$（配列番号２２２）（ここで、X_1はＮ又はＫ又はＷ又はＲ又はＥ又はＴ又はＹであり、X_2はＭ又はＲ又はＬ又はＳ又はＫ又はＶ又はＥ又はＴ又はＩ又はＤであり、X_3はＬ又はＲ又はＨ又はＰ又はＴ又はＫ又はＰのＱ又はＳ又はＡであり、X_4はＧ又はＱ又はＮ又はＲ又はＫ又はＥ又はＩ又はＴ又はＳ又はＣであ

り、X_5はＲ又はW又はＹ又はＫ又はＴ又はＦ又はＳ又はＱである）；及び
（ｈ）X_1ＮＧ$X_2X_3X_4$ＤX_5Ｎ$X_6X_7X_8$Ｎ（配列番号２２３）（ここで、X_1はＩ又はＫ又はＶ又はＬであり、X_2はＬ又はＭであり、X_3はＮ又はＨ又はＰであり、X_4はＡ又はＳ又はＣであり、X_5はＶ又はＹ又はＩ又はＦ又はＴ又はＮであり、X_6はＡ又はＳであり、X_7はＳ又はＡ又はＰであり、X_8はＭ又はＣ又はＬ又はＲ又はＮ又はＳ又はＫ又はＬである）
を含む、項目７３に記載の方法。
（項目７５）
前記ダイレクトリピート配列が、配列番号５７～９０、配列番号１１８～１５１、又は配列番号２１３のいずれか１つに記載のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む、項目７３又は７４に記載の方法。
（項目７６）
前記ダイレクトリピート配列が、配列番号５７～９０、配列番号１１８～１５１、又は配列番号２１３のいずれか１つに記載のヌクレオチド配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む、項目７５に記載の方法。
（項目７７）
前記ダイレクトリピート配列が、以下の配列の１つ以上：
（ａ）X_1X_2Ｔ$X_3X_4X_5X_6X_7X_8$（配列番号２２４）（ここで、X_1はＡ又はＣ又はＧであり、X_2はＴ又はＣ又はＡであり、X_3はＴ又はＧ又はＡであり、X_4はＴ又はＧであり、X_5はＴ又はＧ又はＡであり、X_6はＧ又はＴ又はＡであり、X_7はＴ又はＧ又はＡであり、X_8はＡ又はＧ又はＴである）；
（ｂ）$X_1X_2X_3X_4X_5X_6X_7X_8X_9$（配列番号２２６）（ここで、$X_1$はＴ又はＣ又はＡであり、$X_2$はＴ又はＡ又はＧであり、$X_3$はＴ又はＣ又はＡであり、$X_4$はＴ又はＡであり、$X_5$はＴ又はＡ又はＧであり、$X_6$はＴ又はＡであり、$X_7$はＡ又はＴであり、$X_8$はＡ又はＧ又はＣ又はＴであり、$X_9$はＧ又はＡ又はＣである）；及び
（ｃ）$X_1X_2X_3$ＡＣ（配列番号２２８）（ここで、X_1はＡ又はＣ又はＧであり、X_2はＣ又はＡであり、X_3はＡ又はＣである）
を含む、項目７３～７６のいずれか一項に記載の方法。
（項目７８）
前記ＣＲＩＳＰＲ関連タンパク質が、配列番号１に記載のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、前記ダイレクトリピート配列が、配列番号５７に記載のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む、項目７３～７７のいずれか一項に記載の方法。
（項目７９）
前記ＣＲＩＳＰＲ関連タンパク質が、配列番号１に記載のアミノ酸配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、前記ダイレクトリピート配列が、配列番号５７に記載のヌクレオチド配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む、項目７８に記載の方法。
（項目８０）
前記ＣＲＩＳＰＲ関連タンパク質が、配列番号１に記載のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９

９％又は１００％）の同一性を有するタンパク質であり、前記ＣＲＩＳＰＲ関連タンパク質が、ＰＡＭ配列の認識能を有し、前記ＰＡＭ配列が、５’－ＴＮＮＴ－３’又は５’－ＴＮＲＴ－３’として記載される核酸配列を含み、「Ｎ」が任意のヌクレオチドであり、「Ｒ」がＡ又はＧである、項目７３～７７のいずれか一項に記載の方法。

（項目８１）

前記ＣＲＩＳＰＲ関連タンパク質が、配列番号１に記載のアミノ酸配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、前記ＣＲＩＳＰＲ関連タンパク質が、ＰＡＭ配列の認識能を有し、前記ＰＡＭ配列が、５’－ＴＮＮＴ－３’又は５’－ＴＮＲＴ－３’として記載される核酸配列を含み、「Ｎ」が任意のヌクレオチドであり、「Ｒ」がＡ又はＧである、項目８０に記載の方法。

（項目８２）

前記ＣＲＩＳＰＲ関連タンパク質が、配列番号４に記載のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、前記ダイレクトリピート配列が、配列番号６０に記載のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む、項目７３～７７のいずれか一項に記載の方法。

（項目８３）

前記ＣＲＩＳＰＲ関連タンパク質が、配列番号４に記載のアミノ酸配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、前記ダイレクトリピート配列が、配列番号６０に記載のヌクレオチド配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む、項目８２に記載の方法。

（項目８４）

前記ＣＲＩＳＰＲ関連タンパク質が、配列番号４に記載のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、前記ＣＲＩＳＰＲ関連タンパク質が、ＰＡＭ配列の認識能を有し、前記ＰＡＭ配列が、５’－ＮＴＴＮ－３’、５’－ＮＴＴＲ－３’（例えば、５’－ＴＴＴＧ－３’）、又は５’－ＮＮＲ－３’として記載される核酸配列を含み、「Ｎ」が任意のヌクレオチドであり、「Ｒ」がＡ又はＧである、項目７３～７７のいずれか一項に記載の方法。

（項目８５）

前記ＣＲＩＳＰＲ関連タンパク質が、配列番号４に記載のアミノ酸配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、前記ＣＲＩＳＰＲ関連タンパク質が、ＰＡＭ配列の認識能を有し、前記ＰＡＭ配列が、５’－ＮＴＴＮ－３’、５’－ＮＴＴＲ－３’（例えば、５’－ＴＴＴＧ－３’）、又は５’－ＮＮＲ－３’として記載される核酸配列を含み、「Ｎ」が任意のヌクレオチドであり、「Ｒ」がＡ又はＧである、項目８４に記載の方法。

（項目８６）

前記ＣＲＩＳＰＲ関連タンパク質が、配列番号１０に記載のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、前記ダイレクトリピート配列が、配列番号６２又は配列番号２１３に記載のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は

１００％）同一のヌクレオチド配列を含む、項目７３～７７のいずれか一項に記載の方法。
（項目８７）
前記ＣＲＩＳＰＲ関連タンパク質が、配列番号１０に記載のアミノ酸配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、前記ダイレクトリピート配列が、配列番号６２又は配列番号２１３に記載のヌクレオチド配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む、項目８６に記載の方法。
（項目８８）
前記ＣＲＩＳＰＲ関連タンパク質が、配列番号１０に記載のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、前記ＣＲＩＳＰＲ関連タンパク質が、ＰＡＭ配列の認識能を有し、前記ＰＡＭ配列が、５’－ＮＴＴＮ－３’又は５’－ＲＴＴＲ－３’（例えば、５’－ＡＴＴＧ－３’又は５’－ＧＴＴＡ－３’）として記載される核酸配列を含み、「Ｎ」が任意のヌクレオチドであり、「Ｒ」がＡ又はＧである、項目７３～７７のいずれか一項に記載の方法。
（項目８９）
前記ＣＲＩＳＰＲ関連タンパク質が、配列番号１０に記載のアミノ酸配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）の同一性を有するタンパク質であり、前記ＣＲＩＳＰＲ関連タンパク質がＰＡＭ配列の認識能を有し、前記ＰＡＭ配列が、５’－ＮＴＴＮ－３’又は５’－ＲＴＴＲ－３’（例えば、５’－ＡＴＴＧ－３’又は５’－ＧＴＴＡ－３’）として記載される核酸配列を含み、「Ｎ」が任意のヌクレオチドであり、「Ｒ」がＡ又はＧである、項目８８に記載の方法。
（項目９０）
前記スペーサー配列が、約１５ヌクレオチド～約５５ヌクレオチドを含む、項目７３～８９のいずれか一項に記載の方法。
（項目９１）
前記スペーサー配列が、２０～４５ヌクレオチドを含む、項目９０に記載の方法。
（項目９２）
前記システムがｔｒａｃｒＲＮＡを更に含む、項目７３～９１のいずれか一項に記載の方法。
（項目９３）
前記システムがｔｒａｃｒＲＮＡを含まない、項目７３～９１のいずれか一項に記載の方法。
（項目９４）
前記標的核酸がＤＮＡ分子である、項目７３～９３のいずれか一項に記載の方法。
（項目９５）
前記ＣＲＩＳＰＲ関連タンパク質が、非特異的ヌクレアーゼ活性を含む、項目７３～９４のいずれか一項に記載の方法。
（項目９６）
前記標的核酸の前記修飾が、二本鎖切断イベントである、項目７３～９５のいずれか一項に記載の方法。
（項目９７）
前記標的核酸の前記修飾が、一本鎖切断イベントである、項目７３～９６のいずれか一項に記載の方法。
（項目９８）
前記標的核酸の前記修飾により、挿入イベントが生じる、項目７３～９７のいずれか一項に記載の方法。
（項目９９）
前記標的核酸の前記修飾により、欠失イベントが生じる、項目７３～９８のいずれか一

項に記載の方法。
（項目１００）
前記標的核酸の修飾により、細胞毒性又は細胞死が生じる、項目７３～９９のいずれか一項に記載の方法。
（項目１０１）
標的核酸の編集方法であって、項目１～４７のいずれか一項に記載のシステムを前記標的核酸に接触させることを含む、方法。
（項目１０２）
標的核酸の発現を改変する方法であって、項目１～４７のいずれか一項に記載のシステムを前記標的核酸に接触させることを含む、方法。
（項目１０３）
標的核酸のある部位におけるペイロード核酸の挿入を標的化する方法であって、前記標的核酸を項目１～４７のいずれか一項に記載のシステムと接触させることを含む、方法。
（項目１０４）
標的核酸のある部位からのペイロード核酸の切出しを標的化する方法であって、前記標的核酸を項目１～４７のいずれか一項に記載のシステムと接触させることを含む、方法。
（項目１０５）
ＤＮＡ標的核酸の認識時に一本鎖ＤＮＡを非特異的に分解する方法であって、前記標的核酸を項目１～４７のいずれか一項に記載のシステムと接触させることを含む、方法。
（項目１０６）
試料中の標的核酸の検出方法であって、
（ａ）項目１～４７のいずれか一項に記載のシステム及び標識されたレポーター核酸を前記試料に接触させることであって、スペーサー配列が前記標的核酸にハイブリダイズすると、前記標識されたレポーター核酸の切断が起こること；及び
（ｂ）前記標識されたレポーター核酸の切断によって生成される検出可能シグナルを測定することであって、それにより前記試料中における前記標的核酸の存在を検出することを含む、方法。
（項目１０７）
（ａ）標的核酸のターゲティング及び編集方法；
（ｂ）核酸の認識に応じた一本鎖核酸の非特異的分解方法；
（ｃ）二本鎖標的のスペーサー相補鎖の認識に応じた二本鎖標的の非スペーサー相補鎖のターゲティング及びニッキング方法；
（ｄ）二本鎖標的核酸のターゲティング及び切断方法；
（ｅ）試料中の標的核酸の検出方法；
（ｆ）二本鎖核酸の特異的編集方法；
（ｇ）二本鎖核酸の塩基編集方法；
（ｈ）細胞における遺伝子型特異的又は転写状態特異的細胞死又は休眠の誘導方法；
（ｉ）二本鎖核酸標的におけるインデルの作成方法；
（ｊ）二本鎖核酸標的への配列の挿入方法；又は
（ｋ）二本鎖核酸標的における配列の欠失又は逆位形成方法
である、インビトロ又はエキソビボでの方法における、項目１～４７のいずれか一項に記載のシステムの使用。
（項目１０８）
哺乳動物細胞における標的核酸中への挿入又は欠失を導入する方法であって、
（ａ）ＣＲＩＳＰＲ関連タンパク質が、配列番号１～５６のいずれか１つに記載のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含む、ＣＲＩＳＰＲ関連タンパク質をコードする核酸配列；及び
（ｂ）ダイレクトリピート配列と前記標的核酸へのハイブリダイゼーション能を有するス

ペーサー配列とを含むＲＮＡガイド（又はＲＮＡガイドをコードする核酸）
のトランスフェクションを含み、
前記ＣＲＩＳＰＲ関連タンパク質が前記ＲＮＡガイドへの結合能を有し；
前記ＣＲＩＳＰＲ関連タンパク質及びＲＮＡガイドによる前記標的核酸の認識により、前記標的核酸の修飾が生じる方法。
（項目１０９）
前記ＣＲＩＳＰＲ関連タンパク質が、配列番号４に記載のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含む、項目１０８に記載の方法。
（項目１１０）
前記ＣＲＩＳＰＲ関連タンパク質が、配列番号４に記載のアミノ酸配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含む、項目１０９に記載の方法。
（項目１１１）
前記ダイレクトリピートが、配列番号６０に記載のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む、項目１０８に記載の方法。
（項目１１２）
前記ダイレクトリピートが、配列番号６０に記載のヌクレオチド配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む、項目１１１に記載の方法。
（項目１１３）
前記標的核酸がＰＡＭ配列に隣接しており、前記ＰＡＭ配列が、５’－ＮＴＴＮ－３’、５’－ＮＴＴＲ－３’（例えば、５’－ＴＴＴＧ－３’）、又は５’－ＮＮＲ－３’として記載される核酸配列を含み、「Ｎ」が任意のヌクレオチドであり、「Ｒ」がＡ又はＧである、項目１０８に記載の方法。
（項目１１４）
前記ＣＲＩＳＰＲ関連タンパク質が、配列番号１０に記載のアミノ酸配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含む、項目１０８に記載の方法。
（項目１１５）
前記ＣＲＩＳＰＲ関連タンパク質が、配列番号１０に記載のアミノ酸配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）同一のアミノ酸配列を含む、項目１１４に記載の方法。
（項目１１６）
前記ダイレクトリピートが、配列番号６２又は配列番号２１３に記載のヌクレオチド配列と少なくとも８０％（例えば、８１％、８２％、８３％、８４％、８５％、８６％、８７％、８８％、８９％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む、項目１０８に記載の方法。
（項目１１７）
前記ダイレクトリピートが、配列番号６２又は配列番号２１３に記載のヌクレオチド配列と少なくとも９５％（例えば、９５％、９６％、９７％、９８％、９９％又は１００％）同一のヌクレオチド配列を含む、項目１１６に記載の方法。
（項目１１８）
前記標的核酸がＰＡＭ配列に隣接しており、前記ＰＡＭ配列が、５’－ＮＴＴＮ－３’又は５’－ＲＴＴＲ－３’（例えば、５’－ＡＴＴＧ－３’又は５’－ＧＴＴＡ－３’）として記載さ

れる核酸配列を含み、「Ｎ」が任意のヌクレオチドであり、「Ｒ」がＡ又はＧである、項目１０８に記載の方法。
（項目１１９）
前記トランスフェクションが一過性トランスフェクションである、項目１０８～１１８のいずれか一項に記載の方法。
（項目１２０）
前記細胞がヒト細胞である、項目１０８～１１９のいずれか一項に記載の方法。
（項目１２１）
（ａ）ＣＲＩＳＰＲ関連タンパク質又は前記ＣＲＩＳＰＲ関連タンパク質をコードする核酸、及び
（ｂ）ダイレクトリピート配列とスペーサー配列とを含むＲＮＡガイド
を含む組成物であって、
前記ＣＲＩＳＰＲ関連タンパク質が、以下のアミノ酸配列の１つ以上：
（ｉ）ＰＸ$_1$Ｘ$_2$Ｘ$_3$Ｘ$_4$Ｆ（配列番号２１６）（ここで、Ｘ$_1$はＬ又はＭ又はＩ又はＣ又はＦであり、Ｘ$_2$はＹ又はＷ又はＦであり、Ｘ$_3$はＫ又はＴ又はＣ又はＲ又はＷ又はＹ又はＨ又はＶであり、Ｘ$_4$はＩ又はＬ又はＭである）；
（ｉｉ）ＲＸ$_1$Ｘ$_2$Ｘ$_3$Ｌ（配列番号２１７）（ここで、Ｘ$_1$はＩ又はＬ又はＭ又はＹ又はＴ又はＦであり、Ｘ$_2$はＲ又はＱ又はＫ又はＥ又はＳ又はＴであり、Ｘ$_3$はＬ又はＩ又はＴ又はＣ又はＭ又はＫである）；
（ｉｉｉ）ＮＸ$_1$ＹＸ$_2$（配列番号２１８）（ここで、Ｘ$_1$はＩ又はＬ又はＦであり、Ｘ$_2$はＫ又はＲ又はＶ又はＥである）；
（ｉｖ）ＫＸ$_1$Ｘ$_2$Ｘ$_3$ＦＡＸ$_4$Ｘ$_5$ＫＤ（配列番号２１９）（ここで、Ｘ$_1$はＴ又はＩ又はＮ又はＡ又はＳ又はＦ又はＶであり、Ｘ$_2$はＩ又はＶ又はＬ又はＳであり、Ｘ$_3$はＨ又はＳ又はＧ又はＲであり、Ｘ$_4$はＤ又はＳ又はＥであり、Ｘ$_5$はＩ又はＶ又はＭ又はＴ又はＮである）；
（ｖ）ＬＸ$_1$ＮＸ$_2$（配列番号２２０）（ここで、Ｘ$_1$はＧ又はＳ又はＣ又はＴであり、Ｘ$_2$はＮ又はＹ又はＫ又はＳである）；
（ｖｉ）ＰＸ$_1$Ｘ$_2$Ｘ$_3$Ｘ$_4$ＳＱＸ$_5$ＤＳ（配列番号２２１）（ここで、Ｘ$_1$はＳ又はＰ又はＡであり、Ｘ$_2$はＹ又はＳ又はＡ又はＰ又はＥ又はＹ又はＱ又はＮであり、Ｘ$_3$はＦ又はＹ又はＨであり、Ｘ$_4$はＴ又はＳであり、Ｘ$_5$はＭ又はＴ又はＩである）；
（ｖｉｉ）ＫＸ$_1$Ｘ$_2$ＶＲＸ$_3$Ｘ$_4$ＱＥＸ$_5$Ｈ（配列番号２２２）（ここで、Ｘ$_1$はＮ又はＫ又はＷ又はＲ又はＥ又はＴ又はＹであり、Ｘ$_2$はＭ又はＲ又はＬ又はＳ又はＫ又はＶ又はＥ又はＴ又はＩ又はＤであり、Ｘ$_3$はＬ又はＲ又はＨ又はＰ又はＴ又はＫ又はＰのＱ又はＳ又はＡであり、Ｘ$_4$はＧ又はＱ又はＮ又はＲ又はＫ又はＥ又はＩ又はＴ又はＳ又はＣであり、Ｘ$_5$はＲ又はＷ又はＹ又はＫ又はＴ又はＦ又はＳ又はＱである）；及び
（ｖｉｉｉ）Ｘ$_1$ＮＧＸ$_2$Ｘ$_3$Ｘ$_4$ＤＸ$_5$ＮＸ$_6$Ｘ$_7$Ｘ$_8$Ｎ（配列番号２２３）（ここで、Ｘ$_1$はＩ又はＫ又はＶ又はＬであり、Ｘ$_2$はＬ又はＭであり、Ｘ$_3$はＮ又はＨ又はＰであり、Ｘ$_4$はＡ又はＳ又はＣであり、Ｘ$_5$はＶ又はＹ又はＩ又はＦ又はＴ又はＮであり、Ｘ$_6$はＡ又はＳであり、Ｘ$_7$はＳ又はＡ又はＰであり、Ｘ$_8$はＭ又はＣ又はＬ又はＲ又はＮ又はＳ又はＫ又はＬである）
を含み、
前記ＣＲＩＳＰＲ関連タンパク質が前記ＲＮＡガイドに結合し、前記スペーサーが標的核酸に結合する、組成物。

【図面】

【図１Ａ】

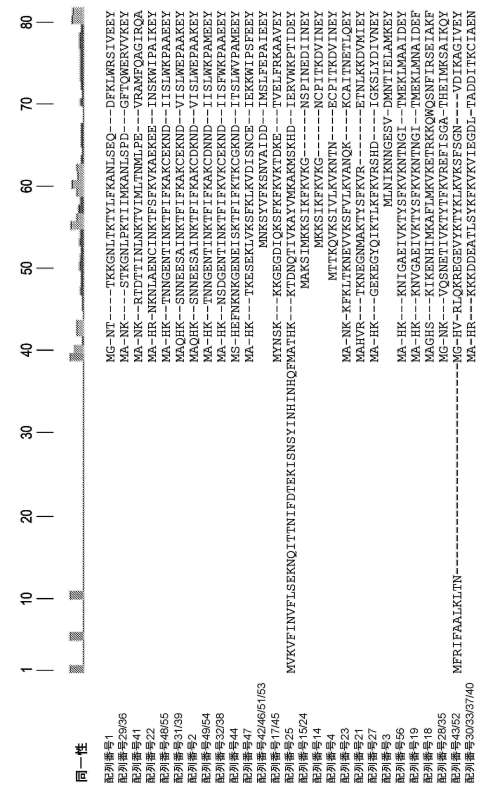


図 1A

【図１Ｂ】

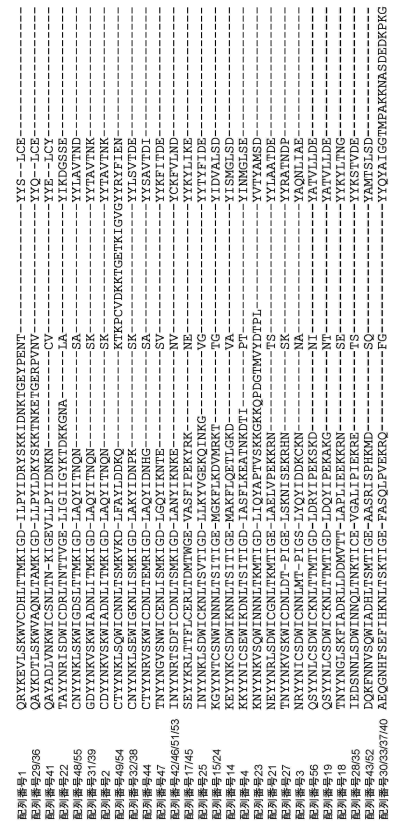


図 1B

【図１Ｃ】

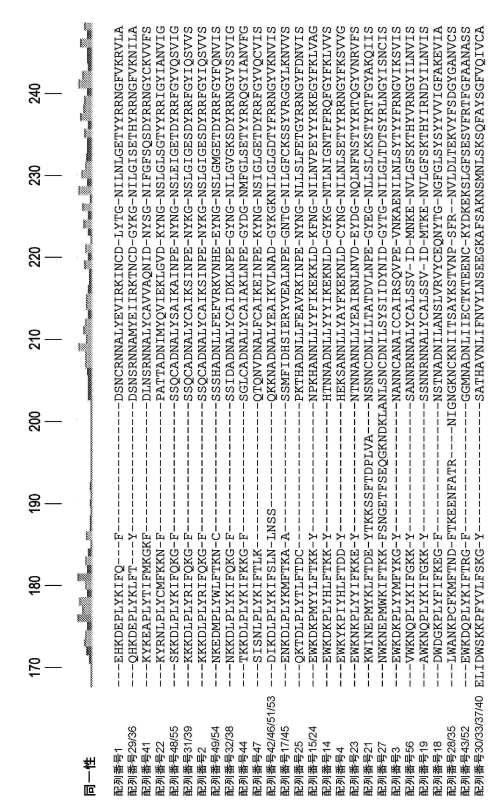


図 1C

【図１Ｄ】

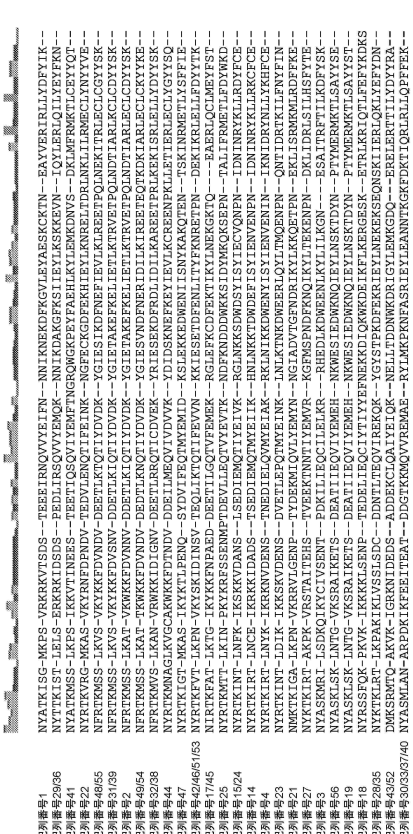


図 1D

【図１Ｅ】

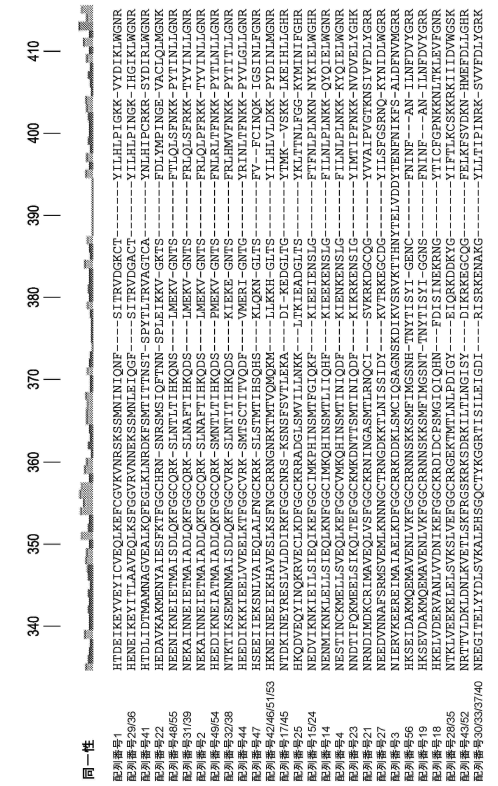


図1E

【図１Ｆ】

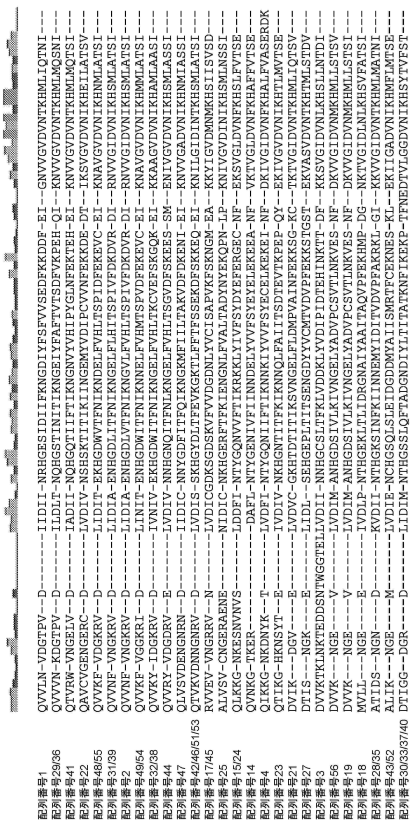


図1F

【図１Ｇ】

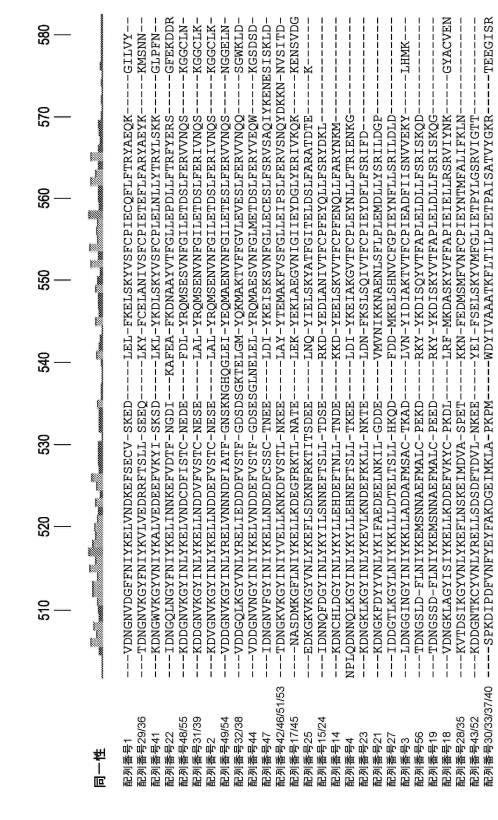


図1G

【図１Ｈ】

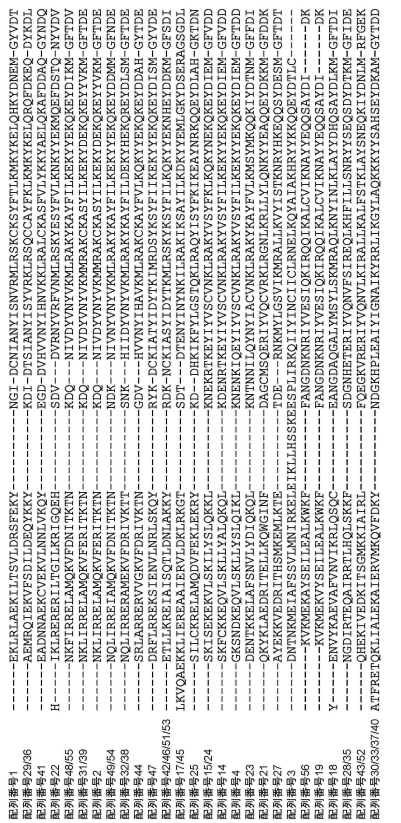


図1H

【図１Ｉ】

図１I

【図１Ｊ】

図１J

【図１Ｋ】

図１K

【図１Ｌ】

図１L

【図２】

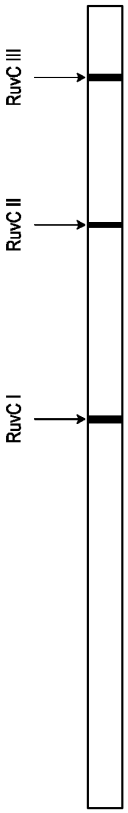


Fig. 2

【図３】

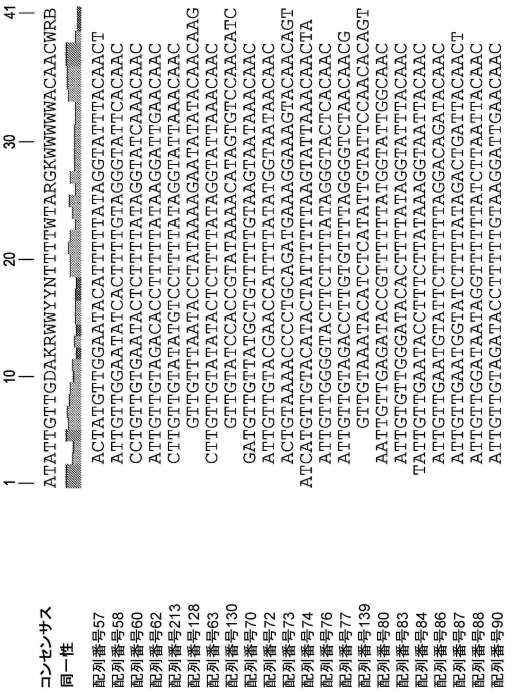


図３

【図４Ａ】

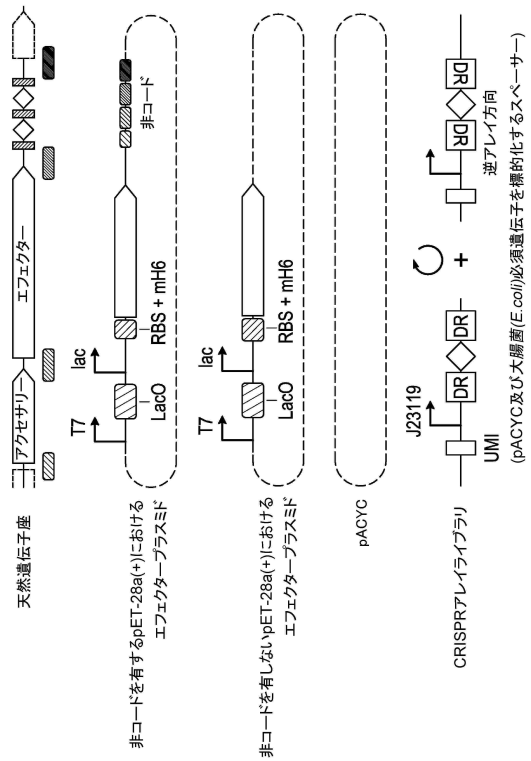


図4A

【図４Ｂ】

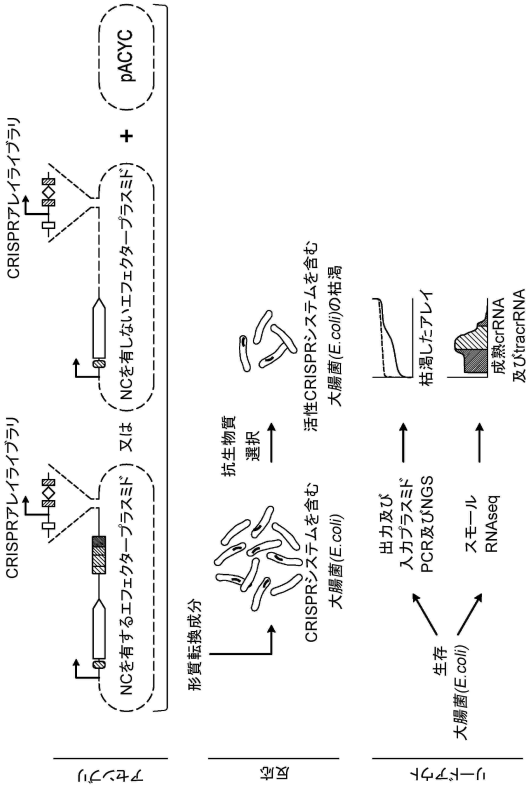


図4B

【図５】

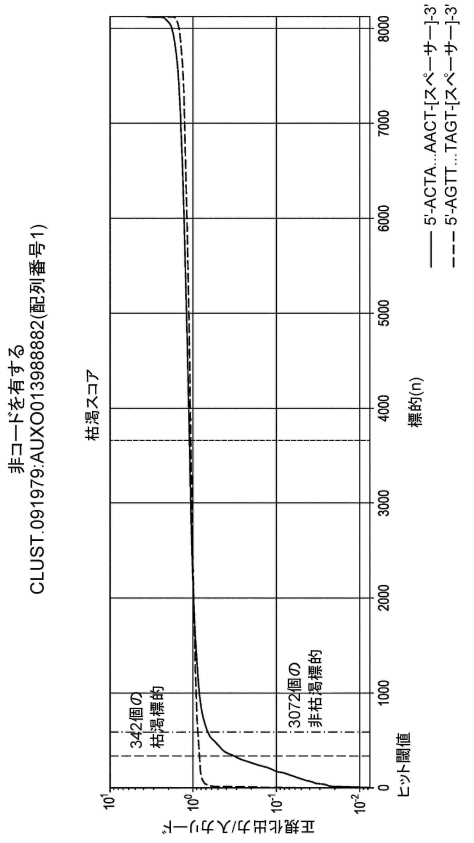


図５

【図６Ａ】

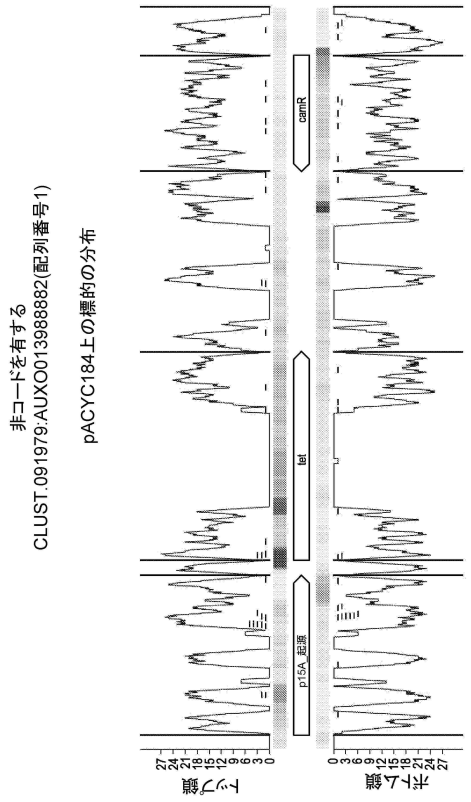


図６Ａ

【図６Ｂ】

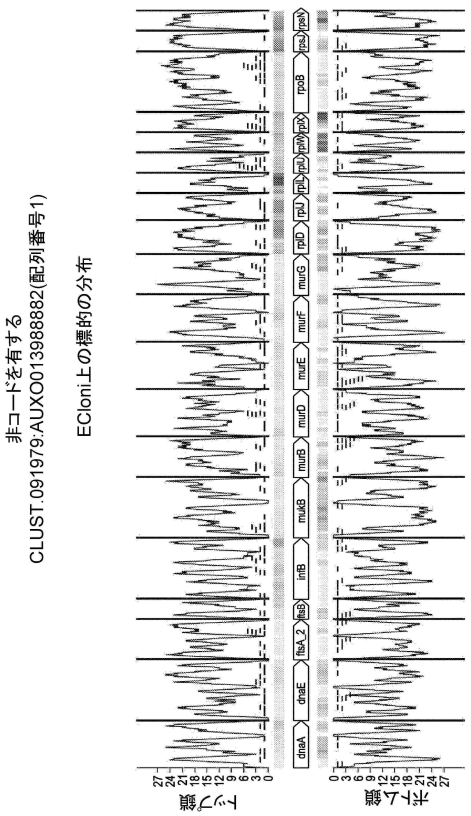
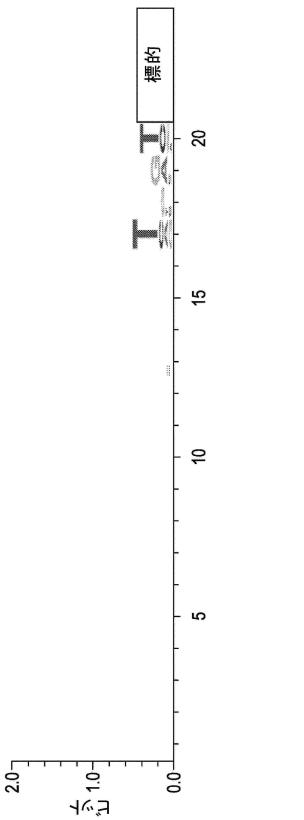


図６Ｂ

【図７】


非コードを有する
CLUST.091979:AUXO013988882(配列番号1)
標的
T
C
G
A
T
C
G
A
T
C
G
A

【図８】

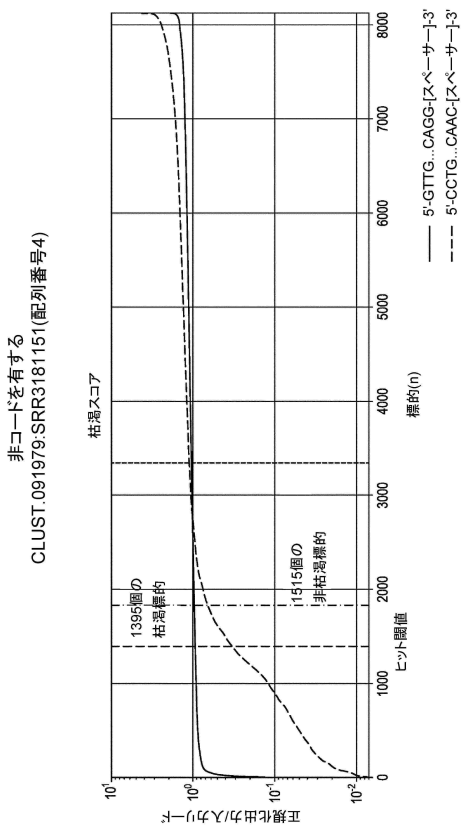


図8

【図９Ａ】

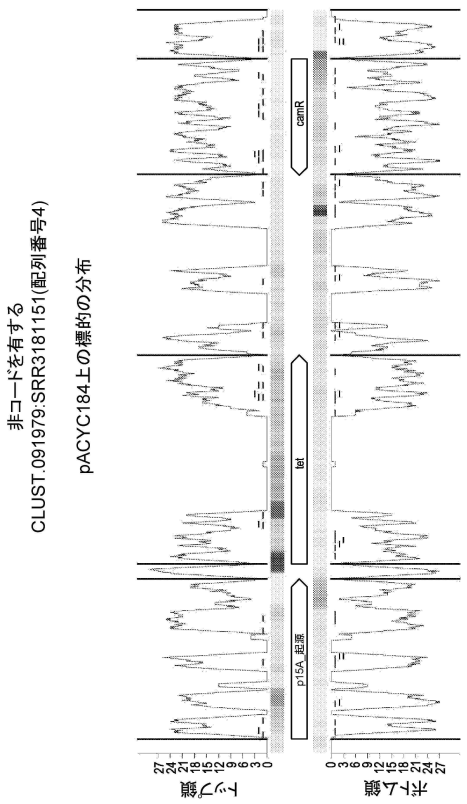


図9A

【図９Ｂ】

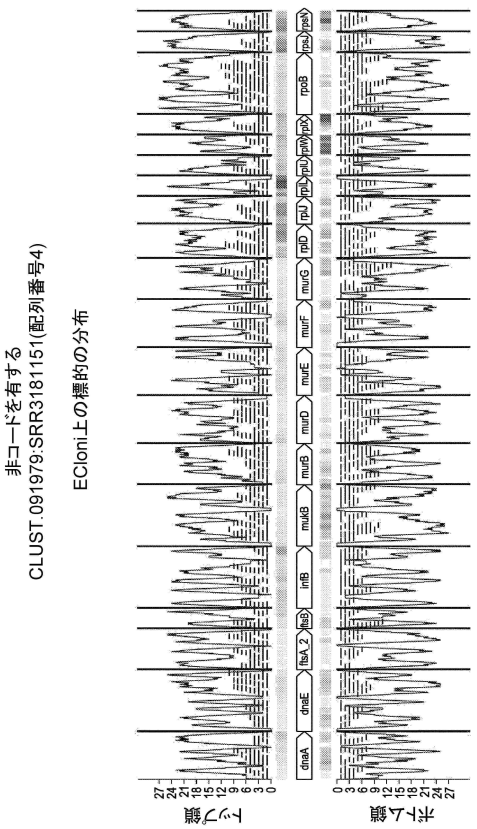


図9B

【図１０】

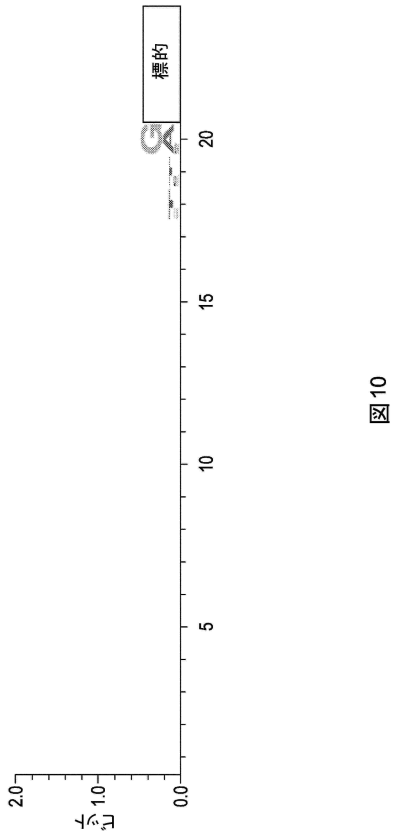


図10

【図１１Ａ】

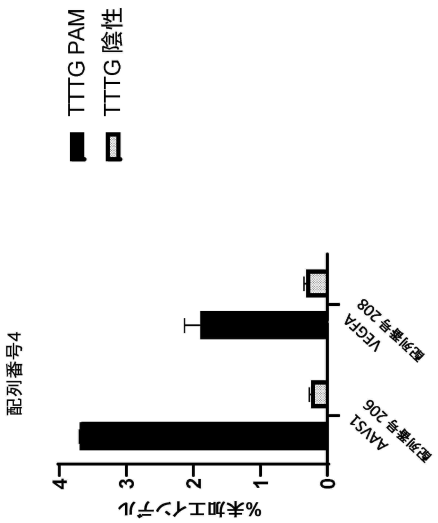


図11A

【図１１Ｂ】

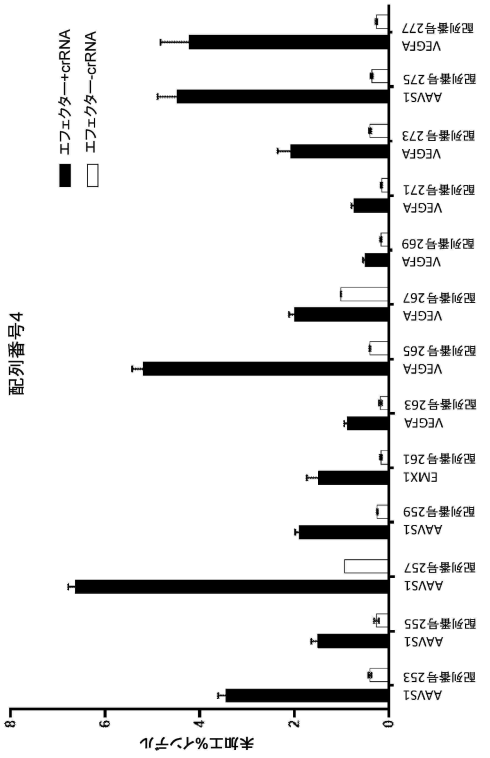


図11B

【図１１Ｃ】

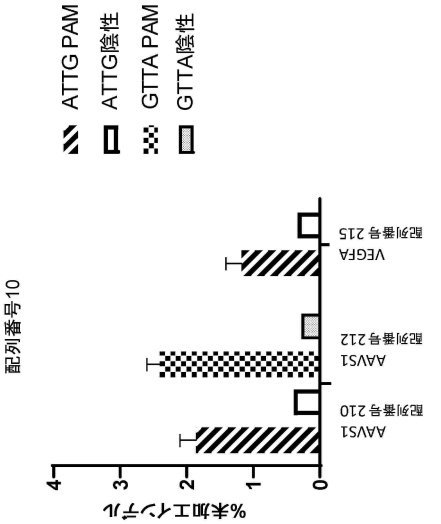


図11C

【図１１Ｄ】

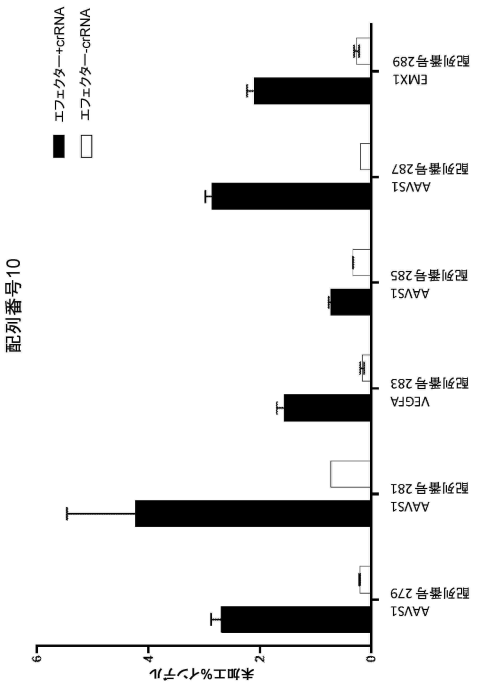


図11D

【配列表】
0007705382000001.app

フロントページの続き

(51)国際特許分類		FI		
C12N	*15/861 (2006.01)*	C12N	15/861	Z
C12N	*15/864 (2006.01)*	C12N	15/864	100Z
C12N	*15/867 (2006.01)*	C12N	15/867	Z
C12N	*15/869 (2006.01)*	C12N	15/869	Z

弁護士　山本 健策
(72)発明者　スコット， デイビッド エー．
アメリカ合衆国 マサチューセッツ ０２１４０， ケンブリッジ， エイコーン パーク ドライブ ２０， タワー ５００
(72)発明者　チェン， デイビッド アール．
アメリカ合衆国 マサチューセッツ ０２１４０， ケンブリッジ， エイコーン パーク ドライブ ２０， タワー ５００
(72)発明者　ヤン， ウィンストン エックス．
アメリカ合衆国 マサチューセッツ ０２１４０， ケンブリッジ， エイコーン パーク ドライブ ２０， タワー ５００
(72)発明者　ディトマソ， ティア エム．
アメリカ合衆国 マサチューセッツ ０２１４０， ケンブリッジ， エイコーン パーク ドライブ ２０， タワー ５００
審査官　上條 のぶよ

(56)参考文献
The CRISPR Journal，2018年，Vol.1, No.2，p.171-181
Frontiers in Microbiology，2017年，Vol.8，Article 2036
Nature，2017年，Vol.542, No.7640，p.237-241
Genome Biol. Evol.，2018年，Vol.10 No.3，p.707-715

(58)調査した分野　(Int.Cl.，ＤＢ名)
Ｃ１２Ｎ　１５／００－９０
Ｃ１２Ｎ　５／００－２８
Ｃ１２Ｎ　１／００－３８
Ｃ１２Ｑ　１／６８－６８９７
ＪＳＴＰｌｕｓ／ＪＭＥＤＰｌｕｓ／ＪＳＴ７５８０（ＪＤｒｅａｍＩＩＩ）
ＣＡｐｌｕｓ／ＲＥＧＩＳＴＲＹ／ＭＥＤＬＩＮＥ／ＥＭＢＡＳＥ／ＢＩＯＳＩＳ（ＳＴＮ）
ＧｅｎＢａｎｋ／ＥＭＢＬ／ＤＤＢＪ／ＧｅｎｅＳｅｑ