(19)日本国特許庁(JP)　(12)公表特許公報(A)　(11)公表番号
特表2023-541990
(P2023-541990A)
(43)公表日　令和5年10月4日(2023.10.4)

(51)国際特許分類		FI		テーマコード（参考）
Ｃ１２Ｎ　15/62　(2006.01)		Ｃ１２Ｎ	15/62	ＺＺＮＡ　４Ｈ０４５
Ｃ１２Ｎ　15/55　(2006.01)		Ｃ１２Ｎ	15/55	
Ｃ１２Ｎ　15/09　(2006.01)		Ｃ１２Ｎ	15/09	１００
Ｃ０７Ｋ　19/00　(2006.01)		Ｃ０７Ｋ	19/00	
Ｃ１２Ｎ　9/78　(2006.01)		Ｃ１２Ｎ	9/78	

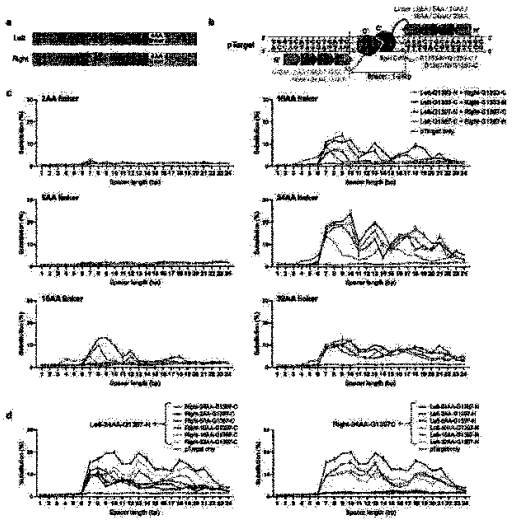
審査請求　有　予備審査請求　未請求　（全154頁）　最終頁に続く

(21)出願番号　特願2023-517882(P2023-517882)
(86)(22)出願日　令和3年9月17日(2021.9.17)
(85)翻訳文提出日　令和5年5月11日(2023.5.11)
(86)国際出願番号　PCT/KR2021/012872
(87)国際公開番号　WO2022/060185
(87)国際公開日　令和4年3月24日(2022.3.24)
(31)優先権主張番号　10-2020-0120399
(32)優先日　令和2年9月18日(2020.9.18)
(33)優先権主張国・地域又は機関
韓国(KR)
(31)優先権主張番号　10-2020-0159920
(32)優先日　令和2年11月25日(2020.11.25)
(33)優先権主張国・地域又は機関
韓国(KR)
(31)優先権主張番号　10-2021-0013263
最終頁に続く

(71)出願人　519444052
インスティチュート　フォー　ベーシック　サイエンス
ＩＮＳＴＩＴＵＴＥ　ＦＯＲ　ＢＡＳＩＣ　ＳＣＩＥＮＣＥ
大韓民国，３４１２６ テジョン，ユソン－グ，エクスポ－ロ ５５
(74)代理人　100139594
弁理士　山口 健次郎
(72)発明者　キム，チンス
大韓民国，０８８３１ ソウル，クァナク－グ，チョンリョン－ギル，７８，＃１０３－７０２
(72)発明者　イム，カヨン
大韓民国，０５５０２ ソウル，ソンパ
最終頁に続く

(54)【発明の名称】　標的化された脱アミノ酵素およびそれを用いた塩基編集

(57)【要約】
本発明は分離された形態のシトシンまたはアデニン脱アミノ酵素またはその変異体、非毒性全長シトシン脱アミノ酵素またはその変異体、それを含む融合タンパク質、塩基編集用組成物およびそれを用いて塩基を編集する方法に関する。



【特許請求の範囲】
【請求項１】
(i) DNA結合タンパク質；及び
(ii) シトシン脱アミノ酵素(deaminase)またはその変異体由来の第1の分割体及び第2の分割体を含む融合タンパク質であって、
前記第1の分割体および第2の分割体は、それぞれ前記DNA結合タンパク質に結合する形態である、融合タンパク質。
【請求項２】
(i) DNA結合タンパク質；及び
(ii) シトシン脱アミノ酵素またはその変異体由来の非毒性全長シトシン脱アミノ酵素を含む、融合タンパク質。
【請求項３】
前記第1の分割体と第2の分割体のそれぞれは、シトシン脱アミノ酵素活性がないことを特徴とする、請求項1に記載の融合タンパク質。
【請求項４】
前記第1の分割体は、配列番号1の配列のうちN-末端からG33、G44、A54、N68、G82、N98、G108からなる群より選ばれる1つ以上までの配列を含む、請求項1に記載の融合タンパク質。
【請求項５】
前記シトシン脱アミノ酵素は、二本鎖DNAに作用する脱アミノ酵素(DddA)またはそのオソログ(orthologue)に由来したことを特徴とする、請求項1又は2に記載の融合タンパク質。
【請求項６】
前記第2の分割体は、配列番号1の配列のうちG34、P45、G55、N69、T83、A99、A109からなる群より選ばれる1つ以上からC-末端までの配列を含む、請求項1に記載の融合タンパク質。
【請求項７】
前記シトシン脱アミノ酵素の変異体は、配列番号1の配列のうちN-末端からG44までの配列を含む第1の分割体のうち、3、5、10、11、13、14、15、16、17、18、19、28、30および31位からなる群より選ばれる1つ以上のアミノ酸が他のアミノ酸に置換された、請求項1に記載の融合タンパク質。
【請求項８】
前記シトシン脱アミノ酵素の変異体は、配列番号1の配列のうちP45からC-末端までの配列を含む第2の分割体のうち13、16、17、20、21、28、29、30、31、32、33、56、57、58および60位からなる群より選ばれる1つ以上のアミノ酸が他のアミノ酸に置換された、請求項1に記載の融合タンパク質。
【請求項９】
前記シトシン脱アミノ酵素の変異体は、配列番号1の配列のうちN-末端からG108までの配列を含む第1の分割体のうち、87、88、91、92、95、100、101、102および103位からなる群より選ばれる1つ以上のアミノ酸が他のアミノ酸に置換された、請求項1に記載の融合タンパク質。
【請求項１０】
前記シトシン脱アミノ酵素の変異体は、配列番号1の配列のうち、A109からC-末端までの配列を含む第2の分割体のうち、13、14、15、及び16位からなる群より選ばれる1つ以上のアミノ酸が他のアミノ酸に置換された、請求項1に記載の融合タンパク質。
【請求項１１】
前記非毒性全長シトシン脱アミノ酵素は、配列番号1の野生型シトシン脱アミノ酵素のうち、37、59、109、129からなる群より選ばれる1つ以上の位置のアミノ酸が他のアミノ酸に置換された、請求項2に記載の融合タンパク質。
【請求項１２】

前記他のアミノ酸はアラニンである、請求項7～11のいずれか1項に記載の融合タンパク質。
【請求項１３】
前記非毒性全長シトシン脱アミノ酵素は、配列番号12～22からなる群より選ばれる少なくとも1つである、請求項2に記載の融合タンパク質。
【請求項１４】
DNA結合タンパク質は、ジンクフィンガータンパク質(zinc finger protein)、TALEタンパク質またはCRISPR関連ヌクレアーゼである、請求項1又は2に記載の融合タンパク質。
【請求項１５】
前記DNA結合タンパク質は、2～40個のアミノ酸残基を含むペプチドリンカーを介してシトシン脱アミノ酵素またはその変異体と連結することを特徴とする、請求項1又は2に記載の融合タンパク質。
【請求項１６】
前記リンカーは、以下を含む、請求項15に記載の融合タンパク質：
2a.aリンカー：GSと、
5a.aリンカー：TGEKQ(配列番号8)と、
10a.aリンカー：SGAQGSTLDF(配列番号9)と、
16a.aリンカー：SGSETPGTSESATPES(配列番号10)と、
24a.aリンカー：SGTPHEVGVYTLSGTPHEVGVYTL(配列番号115)または
32a.aリンカー：GSGGSSGGSSGSETPGTSESATPESSGGSSGGS (配列番号11)
【請求項１７】
前記第1の分割体及び第2の分割体のそれぞれは、ジンクフィンガータンパク質のN-末端またはC-末端に結合した、請求項1に記載の融合タンパク質。
【請求項１８】
前記シトシン脱アミノ酵素に単一のTALEアレイ(array)または第1のTALEアレイ及び第2のTALEアレイのそれぞれが結合する、請求項1又は2に記載の融合タンパク質。
【請求項１９】
(iii) アデニン脱アミノ酵素をさらに含む、請求項1又は2に記載の融合タンパク質。
【請求項２０】
前記アデニン脱アミノ酵素は、大腸菌TadAの変異体であって、デオキシ－アデニン脱アミノ酵素(deoxy-adenine deaminase)である、請求項19に記載の融合タンパク質。
【請求項２１】
前記アデニン脱アミノ酵素は、DNA結合タンパク質またはシトシン脱アミノ酵素またはその変異体のN-末端またはC-末端に結合した、請求項19に記載の融合タンパク質。
【請求項２２】
請求項1～21のいずれか1項に記載の融合タンパク質をコードする、核酸。
【請求項２３】
リボ核酸またはDNAである、請求項22に記載の核酸。
【請求項２４】
請求項1～21のいずれか1項に記載の融合タンパク質または請求項22に記載の核酸を含む、塩基編集用組成物。
【請求項２５】
UGI(uracil glycosylase inhibitor)をさらに含む、請求項24に記載の組成物。
【請求項２６】
請求項1～21のいずれか1項に記載の融合タンパク質または請求項22に記載の核酸を含む、真核細胞の塩基編集用組成物。
【請求項２７】
請求項1～21のいずれか1項に記載の融合タンパク質または請求項22に記載の核酸；及び

NLS(nuclear localization signal)ペプチドまたはそれをコードする核酸を含む、植物細胞の塩基編集用組成物。
【請求項２８】
請求項1～21のいずれか1項に記載の融合タンパク質または請求項22に記載の核酸；及び
葉緑体伝達ペプチド(chloroplast transit peptide)またはそれをコードする核酸を含む、植物細胞の塩基編集用組成物。
【請求項２９】
請求項1～21のいずれか1項に記載の融合タンパク質または請求項22に記載の核酸；及び
ミトコンドリア伝達タンパク質(MitochondrialTargeting Signal, MTS)またはそれをコードする核酸を含む、植物細胞の塩基編集用組成物。
【請求項３０】
核外輸送信号(Nuclear Export Signal, NES)またはそれをコードする核酸をさらに含む、請求項29に記載の組成物。
【請求項３１】
前記融合タンパク質は、以下を通じて植物細胞に伝達されることを特徴とする、請求項27～30のいずれか1項に記載の組成物：
Gene gunによる注入(Bombardment)；
PEG媒介プロトプラストトランスフェクション(Protoplast transfection, PEG-mediated)；
電気穿孔によるプロトプラストトランスフェクション(Protoplast transfection, electroporation)；または
微小注射によるプロトプラスト注射(Protoplast injection, microinjection)。
【請求項３２】
前記核酸は、以下を通じて植物細胞に伝達されることを特徴とする、請求項31に記載の組成物：
Agrobacterium、例えばAgrobacterium tumefaciens、 Agrobacterium rhizogeneなどを用いた形質転換；
ウイルス形質感染；
Gene gunによる注入(Bombardment)；
PEG－媒介プロトプラストトランスフェクション(Protoplast transfection, PEG-mediated)；
電気穿孔によるプロトプラストトランスフェクション(Protoplast transfection, electroporation)；または
微小注射によるプロトプラスト注射(Protoplast injection、microinjection)。
【請求項３３】
前記植物のミトコンドリア、葉緑体または色素体(plastid：白色体、有色体)の、請求項27～30のいずれか1項に記載の塩基編集用組成物。
【請求項３４】
野生型DNA塩基配列を切断するが、編集された塩基配列は切断できないTAL(Transcription Activator-Like)エフェクター(TALE)；及びFokI核酸加水分解酵素(FokI Nuclease)またはそれをコードする核酸、またはZFN(zinc finger Nuclease)またはそれをコードする核酸をさらに含む、請求項27～30のいずれか1項に記載の組成物。
【請求項３５】
請求項27～30のいずれか1項に記載の組成物を処理する段階を含む、真核細胞の核、ミトコンドリア、または色素体DNAの塩基を編集する方法。
【請求項３６】
野生型DNA塩基配列を切断するが、編集された塩基配列は切断できないTALENまたはZFNまたはこれをコードする核酸をさらに含むことにより塩基編集の効率を高める、請

求項35に記載の方法。
【請求項３７】
請求項27～30のいずれか1項に記載の組成物を植物細胞に処理する段階を含む、植物細胞の塩基編集方法。
【請求項３８】
請求項1～21のいずれか1項に記載の融合タンパク質または請求項22に記載の核酸；及び
NLS(nuclear localization signal)ペプチドまたはそれをコードする核酸を植物細胞に処理する段階を含む、植物細胞の塩基編集方法。
【請求項３９】
請求項1～21のいずれか1項に記載の融合タンパク質または請求項22に記載の核酸；及び
葉緑体伝達ペプチド(chloroplast transit peptide)またはそれをコードする核酸を植物細胞に処理する段階を含む、植物細胞の塩基編集方法。
【請求項４０】
請求項1～21のいずれか1項に記載の融合タンパク質または請求項22に記載の核酸；及び
ミトコンドリア伝達タンパク質(Mitochondrial Targeting Signal, MTS)またはそれをコードする核酸を植物細胞に処理する段階を含む、植物細胞の塩基編集方法。
【請求項４１】
請求項1～21のいずれか1項に記載の融合タンパク質または請求項22に記載の核酸；及び
NLS(nuclear localization signal)ペプチドまたはそれをコードする核酸を含む、動物細胞の塩基編集用組成物。
【請求項４２】
請求項1～21のいずれか1項に記載の融合タンパク質または請求項22に記載の核酸；及び
ミトコンドリア伝達タンパク質(MitochondrialTargeting Signal, MTS)またはそれをコードする核酸を含む、動物細胞の塩基編集用組成物。
【請求項４３】
核外輸送信号(Nuclelar Export Signal)またはそれをコードする核酸をさらに含む、請求項42に記載の組成物。
【請求項４４】
野生型DNA塩基配列を切断するが編集された塩基配列は切断できないTAL(Transcription Activator-Like)エフェクター(TALE)；及びFokI核酸加水分解酵素(FokI Nuclease)またはそれをコードする核酸、またはZFNまたはそれをコードする核酸をさらに含む、請求項42に記載の組成物。
【請求項４５】
請求項41又は42に記載の組成物を動物細胞に処理する段階を含む、動物細胞の塩基編集方法。
【請求項４６】
請求項1～21のいずれか1項に記載の融合タンパク質または請求項22に記載の核酸；及びNLS(nuclear localization signal)ペプチドまたはそれをコードする核酸を動物細胞に処理する段階を含む、動物細胞の塩基編集方法。
【請求項４７】
請求項1～21のいずれか1項に記載の融合タンパク質または請求項22に記載の核酸；及びミトコンドリア伝達タンパク質(Mitochondrial Targeting Signal, MTS)またはそれをコードする核酸を動物細胞に処理する段階を含む、動物細胞の塩基編集方法。
【請求項４８】
野生型DNA塩基配列を切断するが、編集された塩基配列は切断できないTALENまたは

ZFNまたはそれをコードする核酸をさらに含むことにより塩基編集の効率を高める、請求項46又は47に記載の方法。

【請求項４９】

請求項19に記載の融合タンパク質またはそれをコードする核酸を含む原核細胞または真核細胞におけるA-to-G塩基編集用組成物であって、前記DNA結合タンパク質は、ジンクフィンガータンパク質(zinc finger protein)、TALEタンパク質またはCRISPR関連ヌクレアーゼであり、前記融合タンパク質のシトシン脱アミノ酵素またはその変異体はバクテリア由来であり、二本鎖DNA特異的であることを特徴とする、組成物。

【請求項５０】

請求項19に記載の融合タンパク質またはそれをコードする核酸を含む原核細胞または真核細胞におけるA-to-G塩基編集用組成物であって、前記DNA結合タンパク質はジンクフィンガータンパク質(zinc finger protein)、TALEタンパク質またはCRISPR関連ヌクレアーゼであり、

前記融合タンパク質のシトシン脱アミノ酵素またはその変異体はバクテリア由来であり、二本鎖DNA特異的なものであり、

前記シトシン脱アミノ酵素またはその変異体のN-末端にDNA結合タンパク質が結合しており、前記融合タンパク質のアデニン脱アミノ酵素のC-末端にDNA結合タンパク質が結合している、組成物。

【請求項５１】

請求項19に記載の融合タンパク質またはそれをコードする核酸およびUGI(uracil glycosylase inhibitor)を含む原核細胞または真核細胞におけるC-to-T塩基編集用組成物であって、前記DNA結合タンパク質は、ジンクフィンガータンパク質(zinc finger protein)、TALEタンパク質またはCRISPR関連ヌクレアーゼであり、前記融合タンパク質のシトシン脱アミノ酵素またはその変異体はバクテリア由来であり、二本鎖DNA特異的であることを特徴とする、組成物。

【請求項５２】

請求項19に記載の融合タンパク質またはそれをコードする核酸を原核細胞または真核細胞に処理する段階を含む原核細胞または真核細胞におけるA-to-G塩基編集方法であって、前記DNA結合タンパク質は、ジンクフィンガータンパク質(zinc finger protein)、TALEタンパク質またはCRISPR関連ヌクレアーゼであり、前記融合タンパク質のシトシン脱アミノ酵素またはその変異体はバクテリア由来であり、二本鎖DNA特異的であることを特徴とする、方法。

【請求項５３】

請求項19に記載の融合タンパク質またはそれをコードする核酸を原核細胞または真核細胞に処理する段階を含む原核細胞または真核細胞におけるA-to-G塩基編集方法であって、前記DNA結合タンパク質はジンクフィンガータンパク質(zinc finger protein)、TALEタンパク質またはCRISPR関連ヌクレアーゼであり、

前記融合タンパク質のシトシン脱アミノ酵素またはその変異体はバクテリア由来であり、二本鎖DNA特異的なものであり、

前記シトシン脱アミノ酵素またはその変異体のN-末端にDNA結合タンパク質が結合しており、前記融合タンパク質のアデニン脱アミノ酵素のC-末端にDNA結合タンパク質が結合している、方法。

【請求項５４】

請求項19に記載の融合タンパク質またはそれをコードする核酸およびUGI(uracil glycosylase inhibitor)を原核細胞または真核細胞に処理する段階を含む原核細胞または真核細胞におけるC-to-T塩基編集方法であって、前記DNA結合タンパク質は、ジンクフィンガータンパク質(zinc finger protein)、TALEタンパク質またはCRISPR関連ヌクレアーゼであり、前記融合タンパク質のシトシン脱アミノ酵素またはその変異体はバクテリア由来であり、二本鎖DNA特異的であることを特徴とする、方法。

【発明の詳細な説明】

【技術分野】
【０００１】
本発明は、分離された形態のシトシンまたはアデニン脱アミノ酵素またはその変異体、非毒性全長シトシン脱アミノ酵素またはその変異体、それを含む融合タンパク質、塩基編集用組成物およびそれを用いて塩基を編集する方法に関する。
【背景技術】
【０００２】
DNA結合タンパク質と脱アミノ酵素が連結された融合タンパク質は、DNA二本鎖切断(DSBs)を生成せずに遺伝体で標的化されたヌクレオチド置換または塩基編集(base editing)、遺伝的障害を誘発する点突然変異を編集するか、原核細胞、ヒトおよび他の真核細胞に目的とする単一のヌクレオチド変異を導入するように標的化された方式(targeted manner)で単一のヌクレオチド転換を可能にする。
【０００３】
標的部位への小さな挿入または欠失(indels)を誘導するCRISPR-Cas9のようなヌクレアーゼとは異なり、標的部位における数個のヌクレオチド(window of several nucleotides)内で一塩基を変換させる。したがって、培養細胞、動物および植物において遺伝疾患を誘発する点突然変異を編集するか、一塩基多型(SNP)を生成することができる。
【０００４】
DNA結合タンパク質と脱アミノ酵素が連結された融合タンパク質として、1) S.pyogenesに由来する触媒的に欠損したCas9(catalytically-deficient Cas9; dCas9)またはD10A Cas9ニッカーゼ(nCas9)と、ラットのシトシン脱アミノ酵素であるrAPOBEC1を含むベースエディタ(Base editors; BEs); 2) dCas9またはnCas9と、ウミヤツメ(sea lamprey)のAID(activation-induced cytidine deaminase)orthologであるPmCDA1またはヒトAIDを含むTarget-AID; 3) MS2-結合タンパク質に融合した過活性化されたAID変異体を募集するためにMS2 RNAヘアピンに連結されたsgRNAsとdCas9を含むCRISPR-Xなどが挙げられる。
【０００５】
このように、ZFN(zinc finger Nuclease)、TALEN(transcription activator-like effector nuclease)、CRISPR(clustered regularly interspaced short palindromic repeat)システム及び核酸分解効率のないクリスパー連関タンパク質9(Cas9)変異体と塩基脱アミノ酵素タンパク質の塩基編集技術などのプログラム(設定)遺伝体編集道具は、塩基配列の変化を通じて植物の遺伝的研究と作物形質改善において発展してきた。しかし、かかる道具は、ミトコンドリアと葉緑体を含む植物細胞器官のDNA配列を編集するには適してないが、その主な理由は、ガイドRNAを小器官に伝達するか、2つの化合物を小器官で同時に発現することが難しいからである。植物小器官は、光合成および呼吸に必要な多数の必須的な遺伝子を暗号化している。かかる小器官の遺伝子を編集する方法や道具は、このような遺伝子の機能研究や作物の生産性と形質改善に非常に必要である。例えば、ミトコンドリアatp6遺伝子の標的突然変異は繁殖に有用な特性である雄性不妊を引き起こす可能性があり、葉緑体ゲノムの16S rRNA遺伝子の特定の点突然変異は抗生剤耐性を引き起こす可能性がある。
【０００６】
バクテリア毒素$DddA_{tox}$は、Burkholderia cenocepaciaに由来するバクテリア毒素の酵素機能を担う部分であって、二重らせんDNAのシトシンを脱アミノ化することができる。脱アミノ酵素の例として、$DddA_{tox}$は細胞に毒性があるため、宿主細胞での毒性を避けるために、$DddA_{tox}$は2つの不活性化された分割体(split)に分けて使用するが、各々の半体は、DNAに結合できるように考案されたDNA結合タンパク質に連結して機能を備えたDdCBE対として使用できる。
【０００７】
原理的には、かかる脱アミノ化の酵素反応は、2つの不活性化された半体がDNA結合タンパク質によって目標DNAの上に近く存在してからこそ活性化されるようになる。した

がって、シトシンからチミンへの塩基編集は、2つのDNA結合タンパク質の結合部位の中間で行われるようになる。DNA結合タンパク質を用いてTALE(transcription activator-like effector)アレイ(array)に結合した2つの不活性形態は、TALEによって標的DNAの近所で結合するときに機能を有する。シトシンからチミンへの塩基変換は、2つのTALE結合部位間の14～18個の塩基間の領域で誘導される。かかる2つに分けた分割形態の$DddA_{tox}$は、実験に多くの制約が伴われる。

【０００８】

全長$DddA_{tox}$は、毒性のために一般的な大腸菌を用いたクローニングは不可であり、毒性を妨げる免疫(immunity)遺伝子を共に発現する大腸菌を用いてクローニングする。

【０００９】

一方、ミトコンドリアDNAは、細胞呼吸に非常に重要な役割を果たすが、ミトコンドリアの酸化的リン酸化(OXPHOS)メカニズムを通じて行われる。OXPHOSメカニズムは生存に必須なので、ミトコンドリアDNAの突然変異は、多くのエネルギーを求める組織のような様々な臓器や筋肉に深刻な機能不全を引き起こす可能性がある。主にヒトのミトコンドリア疾患において、正常のミトコンドリアDNAと一塩基突然変異を有するミトコンドリアDNAは共存するが、これは、ミトコンドリアDNAの数的なヘテロ状態(ヘテロプラスミ)を引き起こす。突然変異と正常のミトコンドリアDNAのバランスが臨床的な症状のあるミトコンドリア疾患の発達を決定する。インビトロ及びインビボにおいて、プログラミング可能な核酸加水分解酵素を、突然変異ミトコンドリアDNAは切断し、正常のミトコンドリアDNAを除去しない方式で使用していた。しかし、かかる核酸加水分解酵素は、特定の突然変異をミトコンドリアに導入するか引き戻すことはできないが、ミトコンドリアにおいて、DNAの二本鎖切断が核のように非相同末端結合や相同組換えを通じて効率的に復旧されないからである。

【００１０】

ミトコンドリア塩基編集を通じて、これまで試みることができなかった多様な疾患に対するモデルを作るか、それを治療する治療剤を製作することに役立つことができる。かかる側面から、高効率のミトコンドリア塩基編集酵素の開発の必要性が増大しつつある。

【００１１】

かかる技術的背景の下で、脱アミノ化酵素の残基を置換して非選択的な塩基編集を減らすか、新規の全長脱アミノ酵素を通じて毒性を持たないようにし、目的とするCBE(cytosine base editor)またはABE(adenine base editor)として使用でき、DNAを編集できることを確認して、本発明を完成した。

【発明の概要】

【発明が解決しようとする課題】

【００１２】

本発明の目的は、DNA結合タンパク質、分離された形態のシトシンまたはアデニン脱アミノ酵素またはその変異体、非毒性全長シトシン脱アミノ酵素、またはその変異体を含む融合タンパク質を提供することにある。

【００１３】

本発明の目的は、前記融合タンパク質をコードする核酸を提供することにある。

【００１４】

本発明の目的は、前記融合タンパク質または核酸を含む塩基編集用組成物を提供することにある。

【００１５】

本発明の目的は、前記組成物を細胞に処理する段階を含む塩基編集方法を提供することにある。

【００１６】

前記目的を達成するために、本発明は、(i) DNA結合タンパク質；及び(ii) シトシン脱アミノ酵素(deaminase)またはその変異体由来の第1の分割体および第2の分割体を含む融合タンパク質であって、前記第1の分割体と第2の分割体は、それぞれ前記DNA結合

タンパク質に結合する形態の融合タンパク質を提供する。
【００１７】
本発明は、(i) DNA結合タンパク質；及び(ii) シトシン脱アミノ酵素またはその変異体由来の非毒性全長シトシン脱アミノ酵素と、を含む融合タンパク質を提供する。
【００１８】
本発明は、(i) DNA結合タンパク質；及び(ii)シトシン脱アミノ酵素(deaminase)またはその変異体と、(iii)アデニン脱アミノ酵素を含む融合タンパク質であって、前記シトシン脱アミノ酵素又はその変異体は、(a) 非毒性全長シトシン脱アミノ酵素または(b) シトシン脱アミノ酵素またはその変異体由来の第1の分割体と第2の分割体を含み、前記第1の分割体と第2の分割体は、それぞれ前記DNA結合タンパク質に結合する形態である、融合タンパク質を提供する。
【００１９】
本発明は、前記融合タンパク質をコードする核酸を提供する。
本発明は、前記融合タンパク質または核酸を含む塩基編集用組成物を提供する。
本発明は、前記融合タンパク質または核酸を含む真核細胞の塩基編集用組成物を提供する。
【００２０】
本発明は、また、前記融合タンパク質または核酸；及びNLS(nuclear localization signal)ペプチドまたはそれをコードする核酸と、を含む植物細胞の塩基編集用組成物を提供する。
【００２１】
本発明は、また、前記融合タンパク質または核酸；及び葉緑体伝達ペプチド(chloroplast transit peptide)またはそれをコードする核酸と、を含む植物細胞の塩基編集用組成物を提供する。
【００２２】
本発明は、また、前記融合タンパク質または核酸；及びミトコンドリア伝達タンパク質(MitochondrialTargeting Signal, MTS)またはそれをコードする核酸と、を含む植物細胞の塩基編集用組成物を提供する。
【００２３】
場合によっては、本発明は、また、核外輸送信号タンパク質またはそれをコードする核酸をさらに含む、植物細胞の塩基編集用組成物を提供する。
【００２４】
本発明は、また、前記組成物を植物細胞に処理する段階を含む植物細胞の塩基編集方法を提供する。
【００２５】
本発明は、また、前記融合タンパク質または核酸；及びNLS(nuclear localization signal)ペプチドまたはそれをコードする核酸を植物細胞に処理する段階と、を含む植物細胞の塩基編集方法を提供する。
【００２６】
本発明は、また、前記融合タンパク質または核酸；及び葉緑体伝達ペプチド(chloroplast transit peptide)またはそれをコードする核酸を植物細胞に処理する段階と、を含む植物細胞の塩基編集方法を提供する。
【００２７】
本発明は、また、前記融合タンパク質または核酸；及びミトコンドリア伝達タンパク質(Mitochondrial Targeting Signal, MTS)またはそれをコードする核酸を植物細胞に処理する段階と、を含む植物細胞の塩基編集方法を提供する。
【００２８】
本発明は、また、前記融合タンパク質または核酸；及びNLS(nuclear localization signal)ペプチドまたはそれをコードする核酸と、を含む動物細胞の塩基編集用組成物を提供する。

【００２９】
本発明は、また、前記融合タンパク質または核酸；及びミトコンドリア伝達タンパク質(Mitochondrial Targeting Signal, MTS)またはそれをコードする核酸と、を含む動物細胞の塩基編集用組成物を提供する。
【００３０】
本発明は、また、場合によっては核外輸送信号タンパク質またはそれをコードする核酸をさらに含む、動物細胞の塩基編集用組成物を提供する。
【００３１】
本発明は、また、前記組成物を動物細胞に処理する段階を含む動物細胞の塩基編集方法を提供する。
【００３２】
本発明は、また、前記融合タンパク質または核酸と、NLS(nuclear localization signal)ペプチドまたはそれをコードする核酸を動物細胞に処理する段階と、を含む動物細胞の塩基編集方法を提供する。
【００３３】
本発明は、また、前記融合タンパク質または核酸と、ミトコンドリア伝達タンパク質(Mitochondrial Targeting Signal, MTS)またはそれをコードする核酸を動物細胞に処理する段階と、を含む動物細胞の塩基編集方法を提供する。
【００３４】
本発明は、また、前記融合タンパク質またはそれをコードする核酸を含む原核細胞または真核細胞におけるA-to-G塩基編集用組成物であって、前記DNA結合タンパク質は、ジンクフィンガータンパク質(zinc finger protein)、TALEタンパク質またはCRISPR連関ヌクレアーゼであり、前記融合タンパク質のシトシン脱アミノ酵素またはその変異体はバクテリア由来であり、二本鎖DNA特異的であることを特徴とする組成物を提供する。
【００３５】
本発明は、また、前記融合タンパク質またはそれをコードする核酸を含む原核細胞または真核細胞におけるA-to-G塩基編集用組成物であって、前記DNA結合タンパク質はジンクフィンガータンパク質、TALEタンパク質またはCRISPR連関ヌクレアーゼであり、前記融合タンパク質のシトシン脱アミノ酵素またはその変異体はバクテリア由来であり、二本鎖DNA特異的なものであり、前記シトシン脱アミノ酵素またはその変異体のN-末端にDNA結合タンパク質が結合しており、前記融合タンパク質のアデニン脱アミノ酵素のC-末端にDNA結合タンパク質が結合している組成物を提供する。
【００３６】
本発明は、また、前記融合タンパク質またはそれをコードする核酸およびUGI(uracil glycosylase inhibitor)を含む原核細胞または真核細胞におけるC-to-T塩基編集用組成物であって、前記DNA結合タンパク質は、ジンクフィンガータンパク質、TALEタンパク質またはCRISPR連関ヌクレアーゼであり、前記融合タンパク質のシトシン脱アミノ酵素またはその変異体は、全長非毒性シトシン脱アミノ酵素であり、前記融合タンパク質のシトシン脱アミノ酵素またはその変異体はバクテリア由来であり、二本鎖DNA特異的であることを特徴とする組成物を提供する。
【００３７】
本発明は、また、前記融合タンパク質またはそれをコードする核酸およびUGI(uracil glycosylase inhibitor)を含む原核細胞または真核細胞におけるC-to-T塩基編集用組成物であって、前記DNA結合タンパク質はジンクフィンガータンパク質、TALEタンパク質またはCRISPR連関ヌクレアーゼであり、前記融合タンパク質のシトシン脱アミノ酵素またはその変異体は、第1の分割体および第2の分割体を含む分割シトシン脱アミノ酵素であり、前記融合タンパク質のシトシン脱アミノ酵素またはその変異体はバクテリア由来であり、二本鎖DNA特異的であることを特徴とする組成物を提供する。
【００３８】
本発明は、また、前記融合タンパク質またはそれをコードする核酸を原核細胞または真

核細胞に処理する段階を含む原核細胞または真核細胞におけるA-to-G塩基編集方法であって、前記DNA結合タンパク質は、ジンクフィンガータンパク質 、TALEタンパク質またはCRISPR連関ヌクレアーゼであり、前記融合タンパク質のシトシン脱アミノ酵素またはその変異体はバクテリア由来であり、二本鎖DNA特異的であることを特徴とする方法を提供する。

【００３９】

本発明は、また、前記融合タンパク質またはそれをコードする核酸を原核細胞または真核細胞に処理する段階を含む原核細胞または真核細胞におけるA-to-G塩基編集方法であって、前記DNA結合タンパク質はジンクフィンガータンパク質、TALEタンパク質またはCRISPR連関ヌクレアーゼであり、

【００４０】

前記融合タンパク質のシトシン脱アミノ酵素またはその変異体はバクテリア由来であり、二本鎖DNA特異的なものであり、

【００４１】

前記シトシン脱アミノ酵素またはその変異体のN-末端にDNA結合タンパク質が結合しており、前記融合タンパク質のアデニン脱アミノ酵素のC-末端にDNA結合タンパク質が結合している、方法を提供する。

【００４２】

本発明は、また、前記融合タンパク質またはそれをコードする核酸およびUGI(uracil glycosylase inhibitor)を原核細胞または真核細胞に処理する段階を含む原核細胞または真核細胞におけるC-to-T塩基編集方法であって、前記DNA結合タンパク質は、ジンクフィンガータンパク質、TALEタンパク質またはCRISPR連関ヌクレアーゼであり、前記融合タンパク質のシトシン脱アミノ酵素またはその変異体はバクテリア由来であり、二本鎖DNA特異的であることを特徴とする方法を提供する。

【図面の簡単な説明】

【００４３】

【図１】pTargetプラスミドを用いたZFD最適化の結果を示した図である。 a、ZFD構造。Split-DddAtoxの半分はZFP(C型)のC-末端に融合する。b、pTargetライブラリーを使用したZFDプラットフォームの最適化。pTargetプラスミドは、サイズが1～24bp(赤色で示す)の範囲のスペーサーおよびZFP DNA結合部位(緑色で示す)を含む。 ZFD構造には、多様な長さのAAリンカー(黄色とオレンジ色で示す)と他のDddAtox分割サイトおよび方向(青色で示す)が含まれている。c-d、ZFD活性は、(b)に説明された変数の影響を調べるために、pTargetライブラリーのターゲット部位で測定した。左側および右側のZFDで、同じ(c)または異なる(d)長さのリンカーのあるZFD対を試験した。塩基編集頻度はpTargetプラスミドの該当領域にターゲットディープシーケンシングにより測定した。データは、n=2の生物学的に独立したサンプルから得られた平均(s.e.m.)の平均±標準誤差で示す。

【図２】pTargetプラスミドにおける多様なリンカーを有するZFD効率を確認した結果を示した図である。 a、ヒートマップに描かれた1-24bpsの長さを有する多様なpTargetスペーサーにおいて、C/Gから非C/Gに編集される頻度。ZFPと分割されたDddAtoxとの間にある多様な種類のリンカーとDddAtoxが分割された部位など多様なZFD構成を試験した。b、各ZFD対の全体的な活性。x軸に使用された命名法で、下には左側のZFDが、上には右側のZFDを示す。c、スペーサー長による塩基編集効率。「AA」は、リンカーにあるアミノ酸の数を示す。データは、n=2の生物学的に独立したサンプルから得られた平均(s.e.m.)の平均±標準誤差で示す。

【図３】24AAリンカーと他のリンカーを有するZFDの効率を確認した結果を示した図である。 ヒートマップに示されたC /Gから非C /Gへの編集効率に対するZFDリンカー長の影響。ZFD対の左側のZFDは24個のAAリンカーで固定し、右側のZFDは可変長のリンカーを含み、その反対の場合も同様である。エラーバーは、n=2個の生物学的に独立したサンプルの平均(s.e.m.)の標準誤差。

【図４】インビボ核を標的するZFDの活性を確認した結果を示した図である。 a、核DNAターゲットZFDの構造。Split-DddAtoxの半分は、ZFPのC-末端(C型)またはN-末端(N型)に融合する。ZFD対はCCまたはNC構成で設計されており、それぞれC型の左側のZFDとC型の右側のZFDまたはN型の左側のZFDとC型の右側のZFDで構成される。b、HEK 293T細胞の核DNAターゲット部位でZFDによって誘導された塩基編集頻度。データは、n=3の生物学的に独立したサンプルで平均±s.e.mで示す。c-f、HEK 293T細胞のNUMBL(c)、INPP5D-2(d)、TRAC-CC(e)およびTRAC-NC(f)ターゲット部位のスペーサー内の各塩基位置におけるZFD誘導塩基編集効率。データは、n=3の生物学的に独立したサンプルで平均±s.e.mで示す。g、電気穿孔または直接送達方式でZFDタンパク質またはZFDエンコーディングプラスミを伝達した後、K562細胞におけるZFD誘導塩基編集頻度。1つまたは4つのNLSを有するZFDタンパク質を試験し、左側および右側のZFDを等モラルとして使用した。電気穿孔法の場合、Amaxa 4D-Nucleofectorを用いて行った。直接伝達法の場合、K562細胞を左側および右側のZFDタンパク質を含む細胞培地と共にインキュベーションした。細胞は同じ方式で1回(1x)または2回(2x)処理された。生物学的に独立したn＝2 のサンプルでは、データは平均±s.e.mで示す。

【図５】核DNAを標的するZFDの構成を示した模式図である。a-d、4つの可能なZFD構成。NCとCNの構成は構造的に同一であるが、左側及び右側のZFDの構成の類型が異なる。

【図６】インビボ核をターゲットするZFDの欠失率を確認した結果を示した図である。試験された全てのZFDは、0.4%未満の頻度で欠失を生成した。データは、n=3の生物学的に独立したサンプルで平均±s.e.mで示す。

【図７】インビトロZFD組換えタンパク質に対する活性試験の結果を示した図である。a、TRACサイトを対象とするZFD対の精製段階。GSTタグ付きタンパク質は、グルタチオンセファロースビーズを用いてE. coli細胞溶解物中で精製された。ポリアクリルアミドゲル電気泳動を使用して精製段階をモニターした。ゲルは、クーマシーブルーで染色された。レーン1、分子量マーカー。レーン2、タンパク質発現がIPTGに誘導されなかった細胞のサンプル。レーン3、タンパク質発現がIPTGに誘導された細胞のサンプル。レーン4、超音波処理後の可溶性分画。レーン5、超音波処理後の不溶性分画。レーン6、カラムの通過分画。レーン6、洗浄分画。レーン7、溶出分画。代表マーカーのサイズは左側に示す。赤いボックスはZFDタンパク質を示す。 b、左側および右側のZFD結合サイト。赤い矢印はZFD誘導脱アミノ化に対して可能なサイトを示す。c、TRAC部位を含むPCRアンプリコンに対するZFD活性の概要。TRAC-NC ZFD対は、まず、シトシンを脱アミノ化してウラシル(赤色で示す)を生成する。その後、USER酵素はウラシルを切断して隙間を生成する(赤い三角形で示す)。d、アガロースゲル電気泳動で分析した未処理PCRアンプリコン(左側)とZFD対で処理されたPCRアンプリコン(右側)。

【図８】ミトコンドリアDNAターゲットZFDの様々な配列(configuration)の模式図である。a-d、4つの可能なmitoZFD構成。従来のZFDのNLSはMTSとNESに置き換えられた。NCおよびCN構成は構造的に同一であるものの、左側および右側のZFDの構成の類型が異なる。

【図９】mitoZFDのミトコンドリア遺伝子塩基編集効率を確認した結果を示した図である。 a、HEK 293T細胞のmitoZFDおよびTALE?DdCBEによって誘導されたmtDNAの塩基編集頻度。データは、n=2の生物学的に独立したサンプルから得られた平均(s.e.m.)の平均±標準誤差で示す。b-g、HEK 293T細胞において、ND2(b)、ND4L(c)、COX2(d)、ND6(e)およびND1(f)ターゲット部位のスペーサー内の各塩基位置におけるmitoZFD誘導塩基編集効率とND1(g)標的部位におけるTALE-DdCBEの誘導塩基編集効率。データは、n=2の生物学的に独立したサンプルから得られた平均(s.e.m.)の平均±標準誤差で示す。mitoZFDとTALE-DdCBEにより導入されたND1遺伝子のDNA変化とアミノ酸変化の比較。各突然変異対立遺伝子に対するシークエンシングリードの頻度(%)は、ターゲットディープシークエンシングによって測定した。ZFD対とTALE-DdCBE対のスペーサー領域は青い点線で示した。

【図１０】MT-ZFDが処理されたHEK293T細胞から分離した単細胞由来クローン集団の塩基編集効率を確認した結果を示した図である。　対立遺伝子分析のために、単一細胞由来クローンを得た。各単一細胞由来クローンにおいて、C/G-to-non-C/G編集頻度は、ターゲットディープシークエンシングによって決定した。a、ND1ターゲットmitoZFDで処理したHEK 293T細胞集団の単一細胞由来クローン。b、ND2ターゲットmitoZFDで処理したHEK 293T細胞集団の単一細胞由来クローン。c、未処理HEK 293T細胞集団の単一細胞由来クローン。ZFP結合サイトは緑色で示す。mitoZFD誘導編集を経たクローンの高い編集頻度は赤色で示す。

【図１１】MT?ZFDが処理されたHEK293T細胞から分離した単細胞由来クローン集団の塩基編集効率を確認した結果を示した図である。　高頻度の塩基編集を示す単一細胞由来クローンの対立遺伝子の分析。表は、ND1における塩基編集の結果として変化したアミノ酸を示す。上部の参照シーケンスにおいて、赤い文字はスペーサーを示す。対立遺伝子において、赤い文字はアミノ酸配列の変化を示す。(*は停止コドンを示す。)

【図１２ａ】及び

【図１２ｂ】MT-ZFDが処理されたHEK293T細胞から分離した単細胞由来クローン集団の塩基編集効率を確認した結果を示した図である。高頻度の塩基編集を示す単一細胞由来クローンの対立遺伝子の分析。表は、ND2における塩基編集の結果として変化したアミノ酸を示す。上部の参照シーケンスにおいて、赤い文字はスペーサーを示す。対立遺伝子において、赤い文字はアミノ酸配列の変化を示す。

【図１３】ZFDとTALE-DdCBEの組合による塩基編集効率を確認した結果を示した図である。　a、mitoZFDおよびTALE-DdCBE対の結合領域のDNA配列。TALE-DdCBEによって認識されるサイトは緑色で強調表示され、mitoZFDについては青色で強調表示される。上部配列はmtDNA heavy strandを示し、下部配列はmtDNA light strandを示す。b、ZFD、TALE-DdCBEおよびZFD/DdCBE混合対によって編集されたシトシンの頻度。ターゲットディープシークエンシングを使用してデータを得たし、データは、n=2の生物学的に独立したサンプルから得られた平均(s.e.m.)の平均±標準誤差で示す。c、各塩基位置において、塩基編集活性のヒートマップ。赤いボックスは、各構成のスペーサーの領域を示す。青い矢印はmtDNAの位置を示す。

【図１４】ZFD濃度別、mRNA vs プラスミドによる編集効率を確認した結果を示した図である。ND1ターゲットmitoZFDのミトコンドリアゲノム全体のターゲット特異性は、ZFDをエンコードするmRNAまたはプラスミドの濃度に応じて異なる特異性を有する。mtDNA全体のシークエンスにより定められたオンおよびオフターゲットの塩基編集頻度。示された濃度のND1-ターゲットmitoZFD-エンコーディングプラスミドまたはmRNAでトランスフェクションしたHEK 293T細胞の結果をグラフ点で示す。赤い矢印はターゲットサイトを示し、赤い点はターゲットサイトの塩基編集頻度を示し、灰色の点はコントロールにも存在するSNPを示す。データは、n=2の生物学的に独立したサンプルから得られた平均(s.e.m.)の平均±標準誤差で示す。

【図１５】ZFD濃度別、mRNA vs プラスミドによる編集効率を確認した結果を示す。a、上部の図は、ND1サイトにおけるZFD結合を示す。ZFD結合サイトは緑色で示す。スペーサーの間にあるターゲットシトシンは赤色で示す。図14の全体のmtDNAシークエンスデータで定めらえたオンターゲット活性。トランスフェクションされたmitoZFDをエンコードするプラスミドまたはmRNAの量が減少するにつれて活性が減少する。b、各プラスミドまたはmRNA量に対して〉1%の頻度で編集されるC/Gサイトの数。c、各プラスミドまたはmRNA濃度によるミトコンドリアゲノムのすべてのC/Gの平均C/G-to-T/A編集頻度。データは、n=2の生物学的に独立したサンプルから得られた平均(s.e.m.)の平均±標準誤差で示す。

【図１６】ZFD濃度別、mRNA vs プラスミドによる編集効率を確認した結果を示す。mtDNA全体のシークエンスにより定められたオンおよびオフターゲット塩基編集頻度。示された濃度のND2-ターゲットmitoZFD-エンコーディングプラスミドまたはmRNAでトランスフェクションしたHEK 293T細胞の結果をグラフ点で示した。赤い矢印はター

ゲットサイトを示し、赤い点はターゲットサイトの塩基編集頻度を示し、灰色の点はコントロールにも存在するSNPを示す。データは、n=2の生物学的に独立したサンプルから得られた平均(s.e.m.)の平均±標準誤差で示す。

【図１７】ZFD濃度別、mRNA vs プラスミドによる編集効率を確認した結果を示した図である。 a、上部の図はND2サイトにおけるZFD結合を示す。ZFD結合サイトは緑色で示す。スペーサー間にあるターゲットシトシンは赤色で示す。図16の全体のmtDNAシークエンスデータで定められたオンターゲット活性。トランスフェクションされたmitoZFDをエンコードするプラスミドまたはmRNAの量が減少するにつれて、活性が減少する。b、各プラスミドまたはmRNA量に対して〉1%の頻度で塩基編集されるC/Gサイトの数。c、各プラスミドまたはmRNA濃度によるミトコンドリアゲノムのすべてのC/Gに対する平均C/G-to-T/A編集頻度。データは、n=2の生物学的に独立したサンプルから得られた平均(s.e.m.)の平均±標準誤差で示す。

【図１８】ミトコンドリア全体のシーケンシング/QQ変異体を作製し、編集効率を確認した結果を示した図である。 a、QQ mitoZFD変異体は、ZFDの各ジンクフィンガーにR(-5)Q突然変異を含めて非特異的DNA接触を除去する。(ジンクフィンガーフレームワークの位置-5にRがない場合は、近くのKまたはRがQに変換される。) b、mitoZFD処理細胞のmtDNA全体のシークエンス。オンおよびオフサイトの編集頻度は、それぞれ赤と黒の点で示す。データは、n=2の生物学的に独立したサンプルから得られた平均(s.e.m.)の平均±標準誤差で示す。〉1%の効率でC/GからT/Aへ塩基編集されたものはすべて示す。c-d、編集効率および特異性は、伝達されたZFDエンコーディングmRNAの用量によって異なる。c、ミトコンドリアゲノムのすべてのC/Gに対する平均C/G-to-T/A編集頻度。d、塩基編集頻度が1%を超える編集されたC/Gの数。

【図１９】植物ベースの塩基編集器のゴールデンゲート組立システム。cp-DdCBEおよびmt-DdCBE構造に対するゴールデンゲートアセンブリの概略図である。対象配列の各位置に対して総424個のセット(= 6 x64 3部 +2 x16 2部 +2 x 4 1部)からTALEサブアレイプラスミドを選択し、所望のベクターと混合して特定配列を対象とするDdCBEをコードするプラスミドを生成した。

【図２０】植物の葉緑体とミトコンドリア塩基編集。a、b、c、d、16s rDNA(a、b)およびpsbA(c、d)におけるcp-DdCBEによって誘導される葉緑体塩基編集頻度とパターン。分割DdCBE G1333とG1397の対をレタスと菜種プロトプラストにトランスフェクションさせた。e、fATP6遺伝子においてmt-DdCBEによって誘導されたミトコンドリア塩基編集効率とパターン。分割DdCBE G1333とG1397の対をレタスと菜種プロトプラストにトランスフェクションさせた。a、c、e、TALE結合領域は青色で示し、スペーサーのシトシンはオレンジ色で示した。全てのグラフにおいて、エラーバーは平均±標準偏差を示す。3つの独立した生物学的複製物。b、d、f、変換されたヌクレオチドは赤色で表示。編集された対立遺伝子%(平均±標準偏差)は、3つの独立した実験から得られた。

【図２１】DdCBEによる植物器官のDNA編集。a、植物器官変異誘発の概略図 である。b、代表的なサンガーシーケンシングクロマトグラムを含めて、スペクチノマイシンの非存在下で培養されたcp-DdCBEトランスフェクションカルス中のC・GからT・Aに変換されることに対する効率。変換されたヌクレオチドの左側に赤色で示す。矢印はクロマトグラムの置換ヌクレオチドを示す。c、DdCBE駆動植物器官変異誘発の要約。変異カルスは、シミュレートされたカルスの頻度よりもはるかに高い編集頻度を有するものと示される。d、16srDNAを対象にしてcp-DdCBEをコードするmRNAをレタスプロトプラストにトランスフェクションした後に誘導されるCからTへの変換頻度。エラーバーは平均±s.dである。n=3の独立した生物学的複製物。e、2.5ヶ月目のスペクトマイシン耐性カルスの編集頻度およびパターン。f、代表的なサンガーシーケンシングクロマトグラムを用いてDdCBE mRNAをトランスフェクションしたストレプトマイシン耐性植物のうち、C・GからT・Aへの変換効率。矢印はクロマトグラムの置換ヌクレオチドを示す。スケールバー：1mm。

【図２２】葉緑体とミトコンドリア塩基編集戦略。cp-DdCBEおよびmt-DdCBE前タン

パク質(preprotein、前駆体タンパク質)のそれぞれにCTP(chloroplast transit peptide)またはMTS (Mitochondrial targeting signal)が含まれているため、植物細胞で翻訳された後、葉緑体およびミトコンドリアに輸送される。前タンパク質は細胞小器官の外膜と内膜を通過し、CTPとMTSはそれぞれ間質処理ペプチダーゼ(peptidase)とミトコンドリア処理ペプチダーゼによって切断された上、cp-DdCBEとmt-DdCBE(成熟タンパク質)が最終形態(conformation)を形成する。

【図２３】レタスプロトプラストにおけるDdCBEプラスミドを通じて編集することに対する時間経過。トランスフェクションされたプロトプラストを各時点で収集し、ターゲットディープシークエンシングによって編集効率性を分析した。頻度(平均±標準偏差)は、3回の独立した実験で得られた。

【図２４】psbB遺伝子の塩基編集頻度。葉緑体psbB遺伝子を標的とするcp-DdCBE対Left-G1333-N+Right-G1333-Cをコードするプラスミドを菜種プロトプラストにトランスフェクションした後、スペーサー領域の塩基編集効率をターゲットディープシークエンシングにより分析した。TALE結合領域、対象シトシン、および変換されたヌクレオチドはそれぞれ青色、オレンジ色、および赤色で示した。頻度(平均±標準偏差)は、n=3回の独立した実験で計算した。

【図２５】ミトコンドリアRPS14遺伝子の塩基編集効率。RPS14遺伝子を標的とするmt-DdCBE対Left-G1333-N+Right-G1333-Cをコードするプラスミドを菜種プロトプラストにトランスフェクションした後、CからTへの変換効率をターゲットディープシークエンシングにより分析した。TALE結合領域、対象シトシン、および変換されたヌクレオチドはそれぞれ青色、オレンジ色、および赤色で示した。頻度(平均±標準偏差)は、n=3回の独立した実験で計算した。

【図２６】カルスにおける葉緑体ゲノムを対象とする塩基編集効率。4週間培養後、レタスと菜種カルスにおいて、16srDNAとpsbAのうち標的部位に対してDdCBEを介して塩基編集した頻度とパターン。スペーサー領域の変換されたヌクレオチドは赤色で示す。

【図２７】カルスにおけるミトコンドリアゲノムを対象とする塩基編集効率性。菜種カルス中のATP6とRPS14遺伝子の標的部位においてDdCBEを介した塩基編集の頻度とパターンが、ターゲットディープシークエンシングによって確認された。対象スペーサー領域の変換されたヌクレオチドは赤色で示す。

【図２８】DNAを含まない塩基編集。レタスプロトプラストにDdCBE mRNAをトランスフェクションした後の16srDNA中のターゲットサイトにおける葉緑体塩基編集の頻度とパターン。プロトプラストを7日間培養した後、ターゲットディップシークエンシングを行った。ターゲットスペーサー領域の変換されたヌクレオチドは赤色で示す。

【図２９】プロトプラストおよびカルス中にDdCBE mRNAまたはDNA配列が存在しないことを示すゲル電気泳動の結果を示した図である(Mはマーカー)。

【図３０】16srDNA変異の選択。赤い矢印はストレプトマイシン耐性の緑色のカルスを示す。

【図３１】抗生剤耐性カルスまたは苗木において、DdCBE標的部位の付近にオフターゲット変異は見付からなかった。(a)、(b) オフターゲット活性は、ターゲットディップシークエンシングにより分析した。TALE結合部位とスペーサー領域は、それぞれ緑色と赤色の下線で示した。(a) DdCBEプラスミドをトランスフェクションしたレタスプロトプラストで培養したスペクチノマイシン耐性カルス。(b) ストレプトマイシン耐性の苗木から得たシュート(shoot)。

【図３２】オンターゲットサイトと最も相同性の高い5つのサイトにおけるオフターゲットの活性分析。TALE結合部位に最大9つの不一致を含めて、レタス葉緑体ゲノムの16SrRNA遺伝子特異的DdCBEの5つの潜在的なオフターゲット部位が選択された。TALE結合配列及び不一致ヌクレオチドは、それぞれ青色と赤色で示した。オフターゲット変異頻度は、ターゲットディップシーケンシングを使用してDdCBEプラスミドまたはDdCBE mRNAにトランスフェクションされたプロトプラストおよび薬剤耐性カルスまたはスーツで測定した。頻度(平均±標準偏差)は、3回の独立した実験で得られた。

【図３３】DdCBE組立およびミトコンドリアDNA編集の模式図である。a、効率的なDdCBEの構築のためのone-pot Golden-gateの組立の説明。最終プラスミド作製のための左側および右側のモジュールを作製するために、総424の配列(64個の3つの配列認識配列×6+16個の2つの配列認識配列×2+4個の1つの配列認識配列×2)および発現ベクターを混合した。b、マウスミトコンドリアDNAの目標遺伝子ND5とDdCBEが相互作用する模式図である。TALE結合部位は灰色で、塩基編集範囲は黒色で示した。それぞれのrepeat variable diresidue moduleはオレンジ色、青色、緑色および黄色で示したが、これは順にアデニン「NI」、チミン「NG」、グアニン「NN」、およびシトシン認識のための「HD」である。

【図３４】DdCBEによる塩基編集から生じたマウスミトコンドリアND5点突然変異。a、DdCBE脱アミノ化酵素をもって媒介されたシトシンからチミンへの塩基編集目標配列およびNIH3T3細胞における効率。目標配列において、翻訳コドンは下線で示し、編集可能な部位は赤色で示した。DdCBEトランスフェクションのための組合は、leftまたはright、-G1333または-G1397、および-Nまたは-Cと表記されている。Left-G1333-N+Right-G1333-C、left-G1333-C+right-G1333-N、left-G1397-N+right-G1397-C、およびleft-G1397-C+right-G1397-NのC10突然変異に対するP値は、それぞれ順に0.0012、0.0003、0.0014、および0.0009であり、C13突然変異は、順に0.0116、0.0076、0.0030、および0.0003である(*p〈0.05および**p〈0.01、Student’s two-tailed t testを使用)。b、これによるマウス胚盤胞における塩基編集効率。配列分析データは、left-G1397-Nとright-G1397-C-DdCBE mRNAを微小注入したコンジュゲートから発達した胚盤胞から得た。c、新生児の突然変異配列の整列図。ターゲットディップシーケンシングは、生まれた直後に得られた尾部位および生後7日、14日後につま先で得られた組織から遺伝体DNAを抽出して行った。 編集された塩基は赤色で示した。突然変異ミトコンドリアゲノムの編集頻度を示した。d、成体F0マウス(sipup-1)の多様な組織における編集効率。シーケンシングデータは、生後50日に各組織から得た。全てのグラフにおいて、濃い灰色のバーと明るい灰色のバーは、それぞれm.C12539T(C10)およびm.G12542A(C13)突然変異の編集頻度を示す。エラーバーは、n=3個の生物学的に独立したサンプルに対する平均(s.e.m.)の標準誤差である。

【図３５】突然変異ミトコンドリアDNAの生殖細胞への伝達。a、mtDNA突然変異の生殖腺伝達を観察するために雌F0(sipup-3)マウスと野生型C57BL6/J雄を交配してF1子孫(101、102)を得た後、ターゲットディープシークエンシングを行った。編集された塩基は赤色で示した。突然変異ミトコンドリアゲノムの編集頻度を示す。b、ゲノムDNAのターゲットディープシークエンシングを使用して得られたF1子(101)の様々な組織における塩基編集効率。濃い灰色と明るい灰色のバーは、それぞれm.C12539T(C10)およびm.G12542A(C13)突然変異の頻度を示す。エラーバーは、n=3個の生物学的に独立したサンプルの平均(s.e.m.)の標準誤差である。

【図３６】DdCBEによって生成されたマウスミトコンドリアND5 G12918A突然変異。a、ND5タンパク質でD393N変化を引き起こすm.G12918A点突然変異を作るためのDdCBE標的。目標コドンに下線で示し、可能な編集部位を赤色で示す。b、NIH3T3細胞におけるDdCBEを用いたシトシン－チミン塩基編集の効率。トランスフェクションされたDdCBE対の組合が示されている。エラーバーはs.e.m.である。n=3個の生物学的に独立したサンプルの場合(n.s.有意味ではない、*p〈0.05、**p〈0.01、Student’s two-tailed t testを使用)。 left -G1333-N + right -G1333-C、left -G1333-C + right -G1333-N、left -G1397-N + right -G1397-C、left -G1397-C + right -G1397- NのC6突然変異に対するP値は、それぞれ0.0052、0.0099、0.0027、および0.0040である。n.s.に対するP値は0.4971である。c、m.G12918A マウス胚盤胞における点突然変異の塩基編集効率。シーケンシングデータは、1細胞期胚にleft -G1397-Cおよびright-G1397-N-DdCBEをコードするmRNAを微小注入した後、培養して発達した胚盤胞で得た。d、ND5点突然変異を有するマウス(F0)。DdCBE mRNAの微小注入後に発達したND5点突然変異を保有しているF0子。新生児で確認された突然変異

配列の整列。編集された塩基は赤色で示し、右側は突然変異ミトコンドリア遺伝子の編集頻度を示す。

【図３７】シトシン脱アミノ化酵素をもって媒介された塩基編集によって産生されたマウスミトコンドリアND5ナンセンス突然変異。a、m.C12336Tナンセンス突然変異とm.G12341Aサイレント突然変異を生成するためのDdCBE目標配列。m.C12336T(C9)突然変異はND5タンパク質でQ199stop突然変異を生成することに対し、m.G12341A(C14)は沈黙するQ200Q突然変異を引き起こす。転写トリプレットは下線で示し、可能な編集位置は赤色で示す。b、NIH3T3細胞でナンセンス突然変異を生成するシトシン－チミン塩基編集の効率性。トランスフェクションされたDdCBE対の組合を示す。濃い灰色と明るい灰色のバーは、それぞれm.C12336T(C9)とm.G12341A(C14)突然変異の頻度を示す。エラーバーはs.e.m.を示す。n=3個の生物学的に独立したサンプルの場合(n.s.有意味ではない、*p〈0.05、**p〈0.01、 Student’s two-tailed t testを使用)。Left-G1333-N+Right-G1333-C、left-G1333-C+right-G1333-N、left-G1397-N+right-G1397-C、left-G1397-C+right-G1397-N の C9突然変異に対するP値は0.0065、0.1143、0.0266および0.0037であり、C14突然変異に対するP値はそれぞれ0.0077、0.0144、0.0406および0.0214である。c、マウス胚盤胞における編集効率。配列分析データは、コンジュゲートにleft-G1333-Nおよびright-G1333-C-DdCBEをエンコードするmRNAを微小注入した後に発達した胚盤胞から得た。濃い灰色と明るい灰色のバーは、それぞれC9およびC14突然変異の頻度を示す。d、新生児の突然変異配列整列。編集された塩基は赤色で示し、右側は突然変異ミトコンドリアゲノムの編集頻度を示す。e、野生型と編集マウスのサンガー配列分析クロマトグラム。赤い矢印は置換されたヌクレオチドを示す。

【図３８】DdCBE構造を生成するゴールデンゲートクローニングの概略図である。全ての反応は、1つのチューブで同時に起こる。矢印は連続的な反応過程を意味しない。BsaI酵素を使用して、空の発現ベクターとモジュールベクターを切断して、互換凝集末端を含み、線形化されたバックボーンおよびTALEモジュール挿入物が欠落するようにした。次に、T4 DNAリガーゼをバックボーンと6つのモジュールインサートを組み合わせて、最終DdCBE構造を作製した。8つのDdCBE複製バックボーンプラスミドを使用した。SOD2MTSの場合、Left-G1333-N、Left-G1333-C、Left-G1397-N、Left-G1397-Cである。 COX8A MTSの場合、Right-G1333-N、Right-G1333-C、Right-G1397-N、Right-G1397-Cである。

【図３９】ND5変異マウス(F0)。(a) ND5サイレント突然変異マウス、(b) ND5 G12918A突然変異マウス、(c) DdCBEmRNA微小注入後、発生したND5ナンセンス突然変異マウス。

【図４０】(a) DdCBE-NESを含むベクター模式図とNES配列。(b) マウスm.G12918 ND5遺伝子および核内の4番染色体のND5-like配列。ミトコンドリアTrnAおよび核内の5番染色体配列。ミトコンドリアRnr2および核内の6番染色体配列。(c) NIH3T3細胞株を用いたDdCBEおよびDdCBE-NESのND5遺伝子における編集効率。(d) NIH3T3細胞株を用いたDdCBEおよびDdCBE-NESのTrnA遺伝子における編集効率。(e) NIH3T3細胞株を用いたDdCBEおよびDdCBE-NESのRnr2遺伝子における編集効率。オレンジグ色のグラフはDdCBE、灰色のグラフはDdCBE-NESの編集効率である。(f) mitoTALEN TALE配列のDNA認識配列。(g) mitoTALENを処理するか処理してない実験群におけるDdCBE塩基編集効率。すべてのグラフはn＝2であり、エラーバーは平均標準誤差である。

【図４１】DdCBE-NESおよびmitoTALENを用いたマウス胚および個体における向上された編集効率。(a) DdCBE-NESを用いた胚盤胞における複数のミトコンドリアDNA標的(mtND5、mtTrnA、mtRNR2)塩基編集効率。 (b) DdCBE、DdCBE-NESおよびmitoTALENを用いたm.G12918A塩基編集効率の比較。(c) マウス個体におけるm.G12918A塩基編集効率の比較。すべてのグラフはn〉=3であり、エラーバーは平均標準誤差である。(n.s.=統計的に無意味である、*p〈0.05、**p〈0.01、***p〈0.001をSt

udent’s two-tailed t testを用いて得た。)
【図４２】(a) DdCBEタンパク質改良に対する模式図である。(b、c) $DddA_{tox}$脱アミノ化酵素の結晶構造。結合表面部位の残基はバーで示した。(b) G1397-N分割及びG1397-C分割は、それぞれ順に紫色と水色で示した。(b) G1333-N分割とG1333-C分割は、それぞれ順にオレンジ色と緑色で示した。(d、e) 結合表面突然変異中のG1397-NおよびG1397-C (d)とG1333-NとG1333-C (e)は赤色で示した。
【図４３】(a) G1397結合表面突然変異の塩基編集効率のグラフ。編集範囲および目標シトシンを上部に示す。突然変異と野生型/TALE-free $DddA_{tox}$タンパク質は、横に示すように、コトランスフェクションされた。Left-DdCBEの場合、TALE-free $DddA_{tox}$タンパク質はG1397-Nであり、Right-DdCBEの場合はTALE-free $DddA_{tox}$タンパク質はG1397-Cである。(b) DdCBEと突然変異の目標シトシン－チミン(グアニン－アデニン)塩基編集効率を示したヒートマップ。
【図４４】(a) G1333結合表面突然変異の塩基編集効率のグラフ。編集範囲および目標シトシンを上部に示す。突然変異と野生型/TALE-free $DddA_{tox}$タンパク質は、横に示すように、コトランスフェクションされた。Left-DdCBEの場合、TALE-free $DddA_{tox}$タンパク質はG1333-Nであり、Right-DdCBEの場合はTALE-free $DddA_{tox}$タンパク質はG1333-Cである。(b) DdCBEと突然変異の目標シトシン－チミン(グアニン－アデニン)塩基編集効率を示したヒートマップ。
【図４５】Wild typeと新規完全長DddAのアミノ酸塩基配列比較結果を示した図である。
【図４６】全長DddAを動物または植物細胞に伝達した形態を示した図である。
【図４７】ヒト細胞ゲノムコンテキストROR1サイト(a)、HEK3サイト(b)、TYRO3サイト(c)においてTCモチフ(motif)のシトシンがチミンに置換される活性を確認した結果を示した図である。
【図４８】全長DddAの利点を示した図である。
【図４９】ヒト細胞ゲノムコンテキストTRACサイト1(a)、TRACサイト2(b)、FANCF(c)、HBB(d)における全長DddAの活性を測定した結果を示した図である。
【図５０】DddA-dCas9(D10A、H840A)-UGIを用いたヒト細胞ゲノムコンテキスTYRO3(a)、ROR1(b)、HEK3(c)、EMX1サイト2(d)、TRACサイト1(e)、HBB(f)におけるDddAの活性を測定した結果を示した図である。
【図５１】HEK293T細胞における全長DddAtoxの塩基編集効率。(a) 構造ベースの方式で全長DddAtoxをスクリーニングする概略図である。赤色アラニンは、正電荷を有するアミノ酸残基がアラニンに置き換えられたことを示す。(b) アラニンに置換されたDddA変異体のE.coli形質転換体を示し、E1347Aを対照群として活性部位の突然変異体として使用した。(c) TYRO3サイトにおけるDddA AAAAAおよびCBEの編集及び挿入欠失の頻度。(d) TYRO3サイトの対立遺伝子頻度がCからTへの置換を赤色で示す。プロトスペーサーは青色で示し、プロトスペーサー隣接モチーフ(PAM)はオレンジ色で示す。
【図５２】無毒性DddA GSVG。(a) エラーが発生しやすいPCRに基づく無毒性の全長DddAtox変異体スクリーニングの概略図である。Cas9、nCas9(D10A)、nCas9(H840A)およびdCas9(D10A、H840A)のN-末端およびC-末端に融合した(b)および対立遺伝子(c)の編集頻度。プロトスペーサーは青色で示し、プロトスペーサー隣接モチーフ(PAM)はオレンジ色で示す。
【図５３】TYRO3サイト(a)、ROR1サイト1(b)、HEK3サイト(c)で正電荷を有するアミノ酸残基がアラニンに置換されたDddAtox変異体がnCas9(D10A)のN-末端に編集された頻度。プロトスペーサーは青色で示し、プロトスペーサー隣接モチーフ(PAM)はオレンジ色で示す。
【図５４】ROR1サイト1(a)、ROR1サイト2(b)、ROR1サイト3(c)、FANCFサイト(d)、HBBサイト(e)、HEK3サイト(f)、TRAC5サイト1(g)、およびEMX1サイト(h、I、j)はそれぞれ。プロトスペーサーは青色で示し、プロトスペーサー隣接モチーフ(PAM)はオレンジ色で示す。CからTへの置換は赤色で示す。DddAのターゲットウィンドウは

、プロトスペーサー5'アップストリームでカウントし、負の数で示す。
【図５５】HeLa細胞におけるAAAAAおよびE1347AのTYRO3サイト(a)、ROR1サイト1(b)、ROR1サイト2(c)、ROR1サイト3(d)、FANCFサイト(e)、HBサイト(f)、HEK3サイト(g)、TRAC5サイト1(h)、TRAC5サイト2(i)、およびEMX1サイト2(j)における編集頻度。プロトスペーサーは青色で示し、プロトスペーサー隣接モチーフ(PAM)はオレンジ色で示す。DddAのターゲットウィンドウはプロトスペーサー5'アップストリームでカウントされ、負の数で示す。標的シトシンは赤色で示す。
【図５６】TYRO3サイト(a)およびROR1サイト1(b)におけるAAAAAおよびE1347Aの時間従属塩基編集および挿入欠失比率を示した図である。
【図５７】EMX1サイト2(a)、FANCFサイト(b)、TRAC5サイト1(c)、TRAC5サイト2(d)、ROR1サイト1(e)、ROR1サイト2(f)、ROR1サイト3(g)及びHBBサイト(h)におけるnCas9(D10A)、nCas9(H840A)およびdCas9のN-末端に融合したGSVGの編集、挿入欠失および対立遺伝子の頻度。プロトスペーサーは青色で示し、プロトスペーサー隣接モチーフ(PAM)はオレンジ色で示す。CからTへの置換は赤色で示す。GSVGのターゲットウィンドウはプロトスペーサー5'アップストリームでカウントし、負の数で示す。
【図５８】EMX1サイト2(a)、EMX1サイト4(b)、ROR1サイト2(c)、およびHBBサイト(d)におけるnCas9(D10A)、nCas9(H840A)およびdCas9のC-末端に融合したGSVGの編集、挿入欠失および対立遺伝子の頻度。プロトスペーサーは青色で示し、プロトスペーサー隣接モチーフ(PAM)はオレンジ色で示す。GからAへの置換は赤色で示す。GSVGのターゲットウィンドウは、プロトスペーサーの最初の位置で3'ダウンストリームでカウントして示す。
【図５９】TYRO3(a)及びEMX1サイト2(b)におけるnCas9(H840A)のC-末端に融合したE1347A、GSVG、SSVG、GSAG、およびGSVSの時間従属編集および挿入欠失の頻度。
【図６０】HEK293T細胞におけるmDdCBEのミトコンドリア塩基編集。ND4(a)及びND6(b)の編集効率性。標的シトシンとTALE結合部位はそれぞれ赤色と灰色で示した。DddAtoxの半分のみがTALEアレイに融合し、残りの半分はTALEのない場合のND4(c、d)およびND6(e、f)における編集効率性。左側のTALE配列と右側のTALE配列は、それぞれLとRで示す。参照ゲノムとのND6 TALEアレイミスマッチは紫色の下線で示す。
【図６１】(a) 従来のZFP DNA結合タンパク質を用いたジンクフィンガーシトシン(ZFD)の模式図である。 (b) アデニンデアミナーゼをZFDに挿入する位置(赤い矢印が挿入位置である) (c) 作製したZFDdABEの核DNA Tracサイトにおける塩基編集効率(C-to-T) (d) 作製したZFDdABEの核DNA Tracサイトにおける塩基編集効率(A-to-G) (WT-ZFDは、アデニンデアミナーゼを有しない分割DddAtoxのみを有するC-to-Tデアミナーゼである) (e) ミトコンドリアDNAを標的とするZF-DdABEのND1サイトにおける効率(C-to-T) (f) ミトコンドリアDNAを標的とする ZF-DdABE のND1サイトにおける効率(A-to-G)
【図６２】(a) TALEと分割DddAtoxを用いたDdABEの模式図である。(構成要素としては、分割DddAtox、アデニンデアミナーゼ及びTALEアレイがある) (b) ミトコンドリアND4サイトを標的とするTALEにアデニンデアミナーゼのみを付着したときの塩基編集効率 (c) ミトコンドリアND1サイトを標的とするTALE-分割DddAtoxにアデニンデアミナーゼを付着したときの塩基編集効率 (d) 左側はDdCBEの対を右側にはTALE-分割DddAtoxにアデニンデアミナーゼを付着したとき、塩基編集効率を1つのヌクレオチド単位で見たもの(緑色のボックスはTALEが付着する部位である) (e) 左側のTALE-分割DddAtoxにアデニンデアミナーゼを付着し、右側はDdCBEの対を用いたとき、塩基編集効率を1つのヌクレオチド単位で見たもの(緑色のボックスはTALEが付着する部位である)
【図６３】(a) UGIがないときとUGIがあるときのミトコンドリアND1サイトターゲットとDdABEのC-to-TとA-to-Gの塩基編集効率(赤色のボックスはアデニンデアミナーゼ

を示す）（b）UGIがないときとUGIがあるときのミトコンドリアND4サイトターゲットと DdABEのC-to-TとA-to-Gの塩基編集効率(赤色のボックスはアデニンデアミナーゼを示す）（c）ミトコンドリアND1サイトターゲットとするDdABEの構成中、最も高い効率を有するものの1つのヌクレオチド単位で見たもの(緑色のボックスはTALEが付着する部位である）（d）ミトコンドリアND4サイトターゲットとするDdABEの構成中、最も高い効率を有するものの1つのヌクレオチド単位で見たもの(緑色のボックスはTALEが付着する部位である)

【図６４】(a）上：1つのTALEモジュールに全てのコンストラクトがあるシングルTALEモジュールの模式図である(構成要素としては、完全長DddAtox、アデニンデアミナーゼおよびTALEアレイがある)。 下：2つのTALEモジュールを使用するデュアルTALEモジュール(構成要素としては、完全長DddAtoxとTALEアレイが一方に、他方にはアデニンデアミナーゼとTALEアレイがある）（b）ミトコンドリアND1サイトをターゲットとするシングルモジュールとデュアルモジュールDdABEの塩基編集効率 （c）ミトコンドリアND4サイトをターゲットとするシングルモジュールとデュアルモジュールDdABEの塩基編集効率

【図６５】ND1サイトをターゲットするシングルモジュールの塩基編集効率(構成要素としては、TALEアレイ、アデニンデアミナーゼ(AD)、および完全長DddAtox(GSVG、AAAA、E1347Aはそれぞれバリアントである))を確認した結果を示した図である。

【図６６】ND1サイトをターゲットするデュアルモジュールの塩基編集効率(構成要素としては、TALEアレイ、アデニンデアミナーゼ(AD)および完全長 DddAtox(GSVG、AAAA、E1347Aはそれぞれバリアントである))を確認した結果を示した図である。

【図６７】(a)ND1サイトをターゲットするTALE結合タンパク質にTadA(AD)アデニン脱アミン酵素のみを付着したときの塩基編集効率を示した図である。 (b)ND4サイトをターゲットするTALE結合タンパク質にTadA(AD)アデニン脱アミン酵素のみを付着したときの塩基編集効率を示した図である。

【図６８】ND1サイトをターゲットするTALEにおけるデュアルモジュール、シングルモジュール及びsplit-DddA-ADに対するアデニン及びシトシン塩基編集効率を示した図である(下から)。両側にUGIがある場合、シトシン塩基編集のみが起こる。一方にADが付着する場合には、シトシンとアデニンの塩基編集がいずれも起こり、UGIがない場合には、選択的にアデニン塩基編集のみが起こることが見られる。同様に、デュアルモジュールとシングルモジュールにおいても、選択的にアデニン塩基編集のみが起こることが分かる。

【発明を実施するための最良の形態】

【００４４】

他の方式で定義されない限り、本明細書で使用されたすべての技術的および科学的用語は、本発明が属する技術分野において熟練した専門家により通常理解されるものと同じ意味を有する。一般に、本明細書で使用される命名法は、当該技術分野において周知であり、通常使用されるものである。

【００４５】

本明細書で使用される「編集」は、「編集」と互いに互換的に使用でき、特異的ゲノム標的の選択的欠失によって核酸配列を変更させる方法を意味する。このような特異的ゲノム標的は、染色体領域、遺伝子、プロモーター、オープンリーディングフレームまたは任意の核酸配列を含むが、これらに限らない。

【００４６】

本明細書で使用される「一塩基」とは、核酸配列内の1つと、1つだけのヌクレオチドを意味する。一塩基編集の文脈で使用される場合、これは、核酸配列内の特異的位置の塩基が異なる塩基に置換されることを意味する。このような置換は、制限なく置換または変形を含む、多くの機序によって起こり得る。

【００４７】

本明細書で使用される「標的」または「標的部位」は、任意の組成および/または長さ

の事前に確認された核線配列を意味する。かかる標的部位は、染色体領域、遺伝子、プロモーター、オープンリーディングフレームまたは任意の核酸配列を含むが、これらに限らない。

【００４８】

本明細書で使用される「オンターゲット」は、プログラム可能なDNA結合領域および/または単一ガイドRNA配列と完全に相補的であり得る特異的ゲノム標的の下位配列を意味する。

【００４９】

本明細書で使用される「オフターゲット」は、プログラム可能なDNA結合領域および/または単一ガイドRNA配列と部分的に相補的であり得る特異的ゲノム標的の下位配列を意味する。

【００５０】

１．分割シトシン脱アミノ酵素(Split Deaminase)

本発明による融合タンパク質は、シトシン脱アミノ酵素(deaminase)またはその変異体を含む融合タンパク質であって、前記シトシン脱アミノ酵素またはその変異体は、シトシン脱アミノ酵素またはその変異体由来の第1の分割体と第2の分割体を含み、前記第1の分割体と第2の分割体はそれぞれ前記DNA結合タンパク質に結合している形態を有する。

【００５１】

前記シトシン脱アミノ酵素は、アミノ基離脱酵素であって、シトシン(C)をウリジン(U)に転換させることができる。

【００５２】

前記シトシン脱アミノ酵素であり得る。前記シトシン脱アミノ酵素として、例えば、APOBEC1(apolipoprotein B editing complex 1)、AID(activation-induced deaminase)の形態が存在するが、ほとんどのDNA脱アミノ酵素は、一本鎖DNAにのみ作用するため、DNA結合タンパク質に連結して塩基編集をすることには適しない場合がある。具体的には、前記シトシン脱アミノ酵素は、二本鎖DNAに作用する脱アミノ酵素(DddA)またはそのオソログ(orthologue)に由来したものであってもよい。より具体的には、前記シトシン脱アミノ酵素は、二本鎖DNA特異的バクテリアシトシン脱アミノ酵素であってもよい。

【００５３】

前記シトシン脱アミノ酵素は、分割された形態であって、前記シトシン脱アミノ酵素は単離された第1の分割体と第2の分割体を含み、前記第1の分割体と第2の分割体のそれぞれは脱アミノ酵素の活性を有さない。

【００５４】

全長シトシン脱アミノ酵素のうち、tox分割体に該当する配列番号1の配列を含んでもよい。前記シトシン脱アミノ酵素は第1の分割体と第2の分割体を含み、前記第1の分割体と第2の分割体のそれぞれは脱アミノ酵素の活性を有さない。

【００５５】

[配列番号1]

GSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGGPTPYPNYANAGHVEGQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKMTVVPPEGAIPVKRGATGETKVFTGNSNSPKSPTKGGC

【００５６】

一実施形態において、前記シトシン脱アミノ酵素の第1の分割体または第2の分割体は、配列番号1の配列中、N-末端でG33、G44、A54、N68、G82、N98、G108からなる群より選ばれる1つ以上までの配列を含んでもよい。前記シトシン脱アミノ酵素の第1の分割体または第2の分割体は、配列番号1の配列中、G34、P45、G55、N69、T83、A99、A109からなる群より選ばれる1つ以上からC-末端までの配列を含んでもよい。

【００５７】

具体的には、前記シトシン脱アミノ酵素は、配列番号23の第1の分割体(G1333-N)と配列番号24の第2の分割体(G1333-C)を含むか、配列番号25の第1の分割体(G1397-N)と配列番号26の第2の分割体(G1397-C)を含むか、配列番号23の第1の分割体(G1333-N)と配列番号26の第2の分割体(G1397-C)を含むか、配列番号25の第1の分割体(G1397-N)と配列番号 24の第2の分割体(G1333-C)を含んでもよい。

【００５８】

(配列番号23) Wild type DddAtox G1333-N

GSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGG (配列番号27)

GGCTCTGGTTCCTACGCCCTGGGTCCATATCAGATTAGTGCTCCCCAACTCCCCGCC
TACAACGGTCAGACAGTGGGGACCTTTTACTATGTCAACGACGCCGGGGGATTGGA
ATCCAAGGTTTTCTCTAGCGGTGGG

【００５９】

(配列番号24) Wild type G1333-C

PTPYPNYANAGHVEGQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAK
MTVVPPEGAIPVKRGATGETKVFTGNSNSPKSPTKGGC (配列番号28)

CCAACACCTTATCCTAACTACGCTAACGCCGGGCACGTCGAGGGGCAGTCAGCTCT
TTTTATGAGAGATAACGGCATTAGCGAAGGGCTTGTGTTCCATAATAATCCTGAGG
GCACCTGTGGCTTCTGTGTAAATATGACCGAAACACTTCTGCCTGAGAACGCTAAA
ATGACTGTCGTACCACCCGAAGGCGCAATCCCAGTTAAACGGGGCGCAACCGGCGA
AACCAAAGTATTCACCGGAAACAGCAATAGTCCAAAGTCCCCCACCAAGGGAGGTT
GC

【００６０】

(配列番号25) Wild type DddAtox G1397-N

GSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGGPTPYPNYANAGHV
EGQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKMTVVPPEG

【００６１】

(配列番号29)

GGTAGCTACGCACTTGGTCCTTACCAGATTAGCGCACCCCAACTCCCCGCCTATAA
TGGTCAAACCGTCGGGACCTTTTACTACGTAAACGATGCTGGTGGGCTGGAATCCA
AAGTATTCTCCTCAGGGGGCCCTACACCCTACCCCAACTACGCCAATGCTGGTCAT
GTAGAAGGGCAGTCAGCACTGTTTATGCGCGATAATGGTATAAGCGAGGGGTTGGT
CTTCCATAACAACCCAGAGGGTACTTGTGGCTTCTGTGTGAATATGACTGAAACCC
TTCTGCCCGAAAATGCCAAGATGACTGTCGTCCCACCTGAAGGC

【００６２】

(配列番号26) Wild type DddAtox G1397-C

AIPVKRGATGETKVFTGNSNSPKSPTKGGC

【００６３】

(配列番号30)

GCCATACCTGTGAAGCGGGGAGCAACAGGGGAGACAAAGGTGTTCACAGGCAACTC
TAACAGTCCAAAGAGCCCCACCAAAGGCGGGTGT

G1333N、G1333C、G1397NおよびG1397Cが組み合わされて分割形態の脱アミノ酵素として使用することができる。具体的には、Left-G1333-N+Right-G133-C、Left-G1397-N+Right-G1397-C、Left-G1397-N+Right-G1333-C、またはLeft-G1333-N+Right-G1397-Cの形態で使用することができる。

【００６４】

２．変異体

本出願の発明者らは、アミノ酸残基が置換されたDdCBE突然変異を通じて、前記DdCBE突然変異が所望しない塩基編集を抑制しようとした。DdCBEの非標的効果を減らすことができる高精度のDddA由来シトシン塩基編集器を提示する。かかる非標的塩基編集の効果は、TALEとDNAとの間の相互作用とは関係なく、$DddA_{tox}$脱アミノ化酵素分割体

の自発的な結合によって引き起こされる現象である。

【００６５】

このとき、突然変異を導入するアミノ酸残基は、DddAtoxの二量体(dimer)が互いに結合する表面に位置する接触部位である。$DddA_{tox}$分割体間の表面に位置したアミノ酸残基をアラニンに置換してHF-DdCBEを作製し、HF-DdCBEは、TALEに連結された2つの脱アミノ化酵素対がDNAに結合できない場合、適切な機能を果たさないようにした。全体のミトコンドリアゲノム分析により、ヒトのミトコンドリアDNAで多数の所望しない非標的C-to-T転換を引き起こす従来のDdCBEとは異なり、HF-DdCBEは非常に効率的かつ精密であることを確認した。

【００６６】

DddAtoxの場合、理論的に分割形態の二量体(split dimer)がすべてDNAの標的部位に適用(recruit)されなければ塩基編集を起こすことができないが、実際の実験の結果、一方のDdCBEはDNAに結合しているが、他方は結合しないDdCBEを使用する場合にも標的塩基編集が起こる。これを解決するために、分割形態の二量体が互いに結合するタンパク質表面の残基を置換して、所望しない位置におけるDdCBE対の結合を防止しようとする。

【００６７】

これに基づき、本発明は、シトシン脱アミノ化酵素DddAtoxのアミノ酸残基を置換して非選択的な塩基編集を減らす新しい変異体に関する。

【００６８】

前記シトシン脱アミノ酵素は、配列番号23の第1の分割体(G1333-N)と配列番号24の第2の分割体(G1333-C)を含むか、配列番号25の第1の分割体(G1397-N)と配列番号26の第2の分割体(G1397-C)を含む。

【００６９】

前記シトシン脱アミノ酵素の変異体は、配列番号23の第1の分割体のうち、3、5、10、11、13、14、15、16、17、18、19、28、30及び31位からなる群より選ばれる1つ以上のアミノ酸が他のアミノ酸に置換されていてもよい。前記配列番号24の第2の分割体のうち、13、16、17、20、21、28、29、30、31、32、33、56、57、58、および60位からなる群より選ばれる1つ以上のアミノ酸が他のアミノ酸に置換されたものである。

【００７０】

本発明による変異体は、前記配列番号25の第1の分割体のうち、87、88、91、92、95、10.0、101、102および103位からなる群より選ばれる1つ以上または前記配列番号26の第2の分割体のうち、13、14、15および16位からなる群より選ばれる1つ以上のアミノ酸が他のアミノ酸に置換されたものである。

【００７１】

前記「他のアミノ酸」は、アラニン、イソロイシン、ロイシン、メチオニン、フェニルアラニン、プロリン、トリプトファン、バリン、アスパラギン酸、システイン、グルタミン、グリシン、セリン、スレオニン、チロシン、グルタミン酸、アルギン、ヒスチジン、ライシン、前記アミノ酸の公知の全ての変異体の中で、野生型タンパク質が元の変異位置に有するアミノ酸を除いたアミノ酸の中から選ばれるアミノ酸を意味する。一例では、前記「他のアミノ酸」はアラニンであってもよい。

【００７２】

具体的には、配列番号23の第1の分割体のうち、Y3A、L5A、I10A、S11A、V13A、G14A、T15A、F16A、Y17A、Y18A、V19A、K28A、F30A、S31Aからなる群より選ばれる少なくとも1つ(それぞれY1292A 、L1294A、I1299A、S1300A、V1312A、G1313A、T1314A、F1315A、Y1316A、Y1317A、V1318A、K1327A、F1329A、S1330Aに該当)のアミノ酸置換を含んでもよい。

【００７３】

配列番号24の第2の分割体のうち、V13A、Q16A、S17A、F20A、M21A、E28A

、G29A、L30A、V31A、F32A、H33A、K56A、M57A、T58A、V60A(それぞれV1346A、Q1349A、S1350A、F1353A、M1354A、E1361A、G1362A、L1363A、V1364A、F1365A、H1366A、K1389A、M1390A、T1391A、V1393A に該当)からなる群より選ばれる1つ以上のアミノ酸置換を含んでもよい。

【００７４】

具体的には、配列番号25の第1の分割体のうち、 C87A、V88A、T91A、E92A、L95A、K100A、M101A、T102A、V103A(それぞれC1376A、V1377A、T1380A、E1381A、L1384A、K1389A、M1390A、T1391A、V1392Aに該当)からなる群より選ばれる1つ以上のアミノ酸置換を含んでもよい。

【００７５】

前記配列番号26の第2の分割体のうち、K13A、V14A、F15A、T16A(それぞれK1410A、V1411A、F1412A、T1413Aに該当)からなる群より選ばれる1つ以上のアミノ酸置換を含んでもよい。

【００７６】

[DddAtox G1333-N変異体]

Type	SEQ	No.
Y1292A	GSAALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGG	31
	GGCTCTGGTTCCgcaGCCCTGGGTCCATATCAGATTAGTGCTCCCCAACTCCCCGCC TACAACGGTCAGACAGTGGGGACCTTTTACTATGTCAACGACGCCGGGGGATTGGAA TCCAAGGTTTTCTCTAGCGGTGGG	32
L1294A	GSYAAGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGG	33
	GGCTCTGGTTCCTACGCCgcaGGTCCATATCAGATTAGTGCTCCCCAACTCCCCGCC TACAACGGTCAGACAGTGGGGACCTTTTACTATGTCAACGACGCCGGGGGATTGGAA TCCAAGGTTTTCTCTAGCGGTGGG	34
I1299A	GSYALGPYQASAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGG	35
	GGCTCTGGTTCCTACGCCCTGGGTCCATATCAGgcaAGTGCTCCCCAACTCCCCGCC TACAACGGTCAGACAGTGGGGACCTTTTACTATGTCAACGACGCCGGGGGATTGGAA TCCAAGGTTTTCTCTAGCGGTGGG	36
S1300A	GSYALGPYQIAAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGG	37
	GGCTCTGGTTCCTACGCCCTGGGTCCATATCAGATTgcaGCTCCCCAACTCCCCGCC TACAACGGTCAGACAGTGGGGACCTTTTACTATGTCAACGACGCCGGGGGATTGGAA TCCAAGGTTTTCTCTAGCGGTGGG	38
V1312A	GSYALGPYQISAPQLPAYNGQTAGTFYYVNDAGGLESKVFSSGG	39
	GGCTCTGGTTCCTACGCCCTGGGTCCATATCAGATTAGTGCTCCCCAACTCCCCGCC TACAACGGTCAGACAgcaGGGACCTTTTACTATGTCAACGACGCCGGGGGATTGGAA TCCAAGGTTTTCTCTAGCGGTGGG	40
G1313A	GSYALGPYQISAPQLPAYNGQTVATFYYVNDAGGLESKVFSSGG	41
	GGCTCTGGTTCCTACGCCCTGGGTCCATATCAGATTAGTGCTCCCCAACTCCCCGCC TACAACGGTCAGACAGTGgcaACCTTTTACTATGTCAACGACGCCGGGGGATTGGAA TCCAAGGTTTTCTCTAGCGGTGGG	42
T1314A	GSYALGPYQISAPQLPAYNGQTVGAFYYVNDAGGLESKVFSSGG	43
	GGCTCTGGTTCCTACGCCCTGGGTCCATATCAGATTAGTGCTCCCCAACTCCCCGCC TACAACGGTCAGACAGTGGGGgcaTTTTACTATGTCAACGACGCCGGGGGATTGGAA TCCAAGGTTTTCTCTAGCGGTGGG	44
F1315A	GSYALGPYQISAPQLPAYNGQTVGTAYYVNDAGGLESKVFSSGG	45
	GGCTCTGGTTCCTACGCCCTGGGTCCATATCAGATTAGTGCTCCCCAACTCCCCGCC TACAACGGTCAGACAGTGGGGACCgcaTACTATGTCAACGACGCCGGGGGATTGGAA TCCAAGGTTTTCTCTAGCGGTGGG	46
Y1316A	GSYALGPYQISAPQLPAYNGQTVGTFAYVNDAGGLESKVFSSGG	47
	GGCTCTGGTTCCTACGCCCTGGGTCCATATCAGATTAGTGCTCCCCAACTCCCCGCC TACAACGGTCAGACAGTGGGGACCTTTgcaTATGTCAACGACGCCGGGGGATTGGAA TCCAAGGTTTTCTCTAGCGGTGGG	48
Y1317A	GSYALGPYQISAPQLPAYNGQTVGTFYAVNDAGGLESKVFSSGG	49
	GGCTCTGGTTCCTACGCCCTGGGTCCATATCAGATTAGTGCTCCCCAACTCCCCGCC TACAACGGTCAGACAGTGGGGACCTTTTACgcaGTCAACGACGCCGGGGGATTGGAA TCCAAGGTTTTCTCTAGCGGTGGG	50
V1318A	GSYALGPYQISAPQLPAYNGQTVGTFYYANDAGGLESKVFSSGG	51

	GGCTCTGGTTCCTACGCCCTGGGTCCATATCAGATTAGTGCTCCCCAACTCCCCGCC TACAACGGTCAGACAGTGGGGACCTTTTACTATgcaAACGACGCCGGGGGATTGGAA TCCAAGGTTTTCTCTAGCGGTGGG	52
K1327A	GSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESAVFSSGG	53
	GGCTCTGGTTCCTACGCCCTGGGTCCATATCAGATTAGTGCTCCCCAACTCCCCGCC TACAACGGTCAGACAGTGGGGACCTTTTACTATGTCAACGACGCCGGGGGATTGGAA TCCgcaGTTTTCTCTAGCGGTGGG	54
F1329A	GSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVASSGG	55
	GGCTCTGGTTCCTACGCCCTGGGTCCATATCAGATTAGTGCTCCCCAACTCCCCGCC TACAACGGTCAGACAGTGGGGACCTTTTACTATGTCAACGACGCCGGGGGATTGGAA TCCAAGGTTgcaTCTAGCGGTGGG	56
S1330A	GSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFASGG	57
	GGCTCTGGTTCCTACGCCCTGGGTCCATATCAGATTAGTGCTCCCCAACTCCCCGCC TACAACGGTCAGACAGTGGGGACCTTTTACTATGTCAACGACGCCGGGGGATTGGAA TCCAAGGTTTTCgcaAGCGGTGGG	58

【００７７】

[DddAtox G1333-C変異体]

Type	SEQ	No.
V1346A	PTPYPNYANAGHAEGQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKM TVVPPEGAIPVKRGATGETKVFTGNSNSPKSPTKGGC	59
	CCAACACCTTATCCTAACTACGCTAACGCCGGGCACgcaGAGGGGCAGTCAGCTCTT TTTATGAGAGATAACGGCATTAGCGAAGGGCTTGTGTTCCATAATAATCCTGAGGGC ACCTGTGGCTTCTGTGTAAATATGACCGAAACACTTCTGCCTGAGAACGCTAAAATG ACTGTCGTACCACCCGAAGGCGCAATCCCAGTTAAACGGGGCGCAACCGGCGAAACC AAAGTATTCACCGGAAACAGCAATAGTCCAAAGTCCCCCACCAAGGGAGGTTGC	60
Q1349A	PTPYPNYANAGHVEGASALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKM TVVPPEGAIPVKRGATGETKVFTGNSNSPKSPTKGGC	61
	CCAACACCTTATCCTAACTACGCTAACGCCGGGCACGTCGAGGGGgcaTCAGCTCTT TTTATGAGAGATAACGGCATTAGCGAAGGGCTTGTGTTCCATAATAATCCTGAGGGC ACCTGTGGCTTCTGTGTAAATATGACCGAAACACTTCTGCCTGAGAACGCTAAAATG ACTGTCGTACCACCCGAAGGCGCAATCCCAGTTAAACGGGGCGCAACCGGCGAAACC AAAGTATTCACCGGAAACAGCAATAGTCCAAAGTCCCCCACCAAGGGAGGTTGC	62
S1350A	PTPYPNYANAGHVEGQAALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKM TVVPPEGAIPVKRGATGETKVFTGNSNSPKSPTKGGC	63
	CCAACACCTTATCCTAACTACGCTAACGCCGGGCACGTCGAGGGGCAGgcaGCTCTT TTTATGAGAGATAACGGCATTAGCGAAGGGCTTGTGTTCCATAATAATCCTGAGGGC ACCTGTGGCTTCTGTGTAAATATGACCGAAACACTTCTGCCTGAGAACGCTAAAATG ACTGTCGTACCACCCGAAGGCGCAATCCCAGTTAAACGGGGCGCAACCGGCGAAACC AAAGTATTCACCGGAAACAGCAATAGTCCAAAGTCCCCCACCAAGGGAGGTTGC	64
F1353A	PTPYPNYANAGHVEGQSALAMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKM TVVPPEGAIPVKRGATGETKVFTGNSNSPKSPTKGGC	65
	CCAACACCTTATCCTAACTACGCTAACGCCGGGCACGTCGAGGGGCAGTCAGCTCTT gcaATGAGAGATAACGGCATTAGCGAAGGGCTTGTGTTCCATAATAATCCTGAGGGC ACCTGTGGCTTCTGTGTAAATATGACCGAAACACTTCTGCCTGAGAACGCTAAAATG ACTGTCGTACCACCCGAAGGCGCAATCCCAGTTAAACGGGGCGCAACCGGCGAAACC AAAGTATTCACCGGAAACAGCAATAGTCCAAAGTCCCCCACCAAGGGAGGTTGC	66
M1354A	PTPYPNYANAGHVEGQSALFARDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKM TVVPPEGAIPVKRGATGETKVFTGNSNSPKSPTKGGC	67
	CCAACACCTTATCCTAACTACGCTAACGCCGGGCACGTCGAGGGGCAGTCAGCTCTT TTTgcaAGAGATAACGGCATTAGCGAAGGGCTTGTGTTCCATAATAATCCTGAGGGC ACCTGTGGCTTCTGTGTAAATATGACCGAAACACTTCTGCCTGAGAACGCTAAAATG ACTGTCGTACCACCCGAAGGCGCAATCCCAGTTAAACGGGGCGCAACCGGCGAAACC AAAGTATTCACCGGAAACAGCAATAGTCCAAAGTCCCCCACCAAGGGAGGTTGC	68
E1361A	PTPYPNYANAGHVEGQSALFMRDNGISAGLVFHNNPEGTCGFCVNMTETLLPENAKM TVVPPEGAIPVKRGATGETKVFTGNSNSPKSPTKGGC	69
	CCAACACCTTATCCTAACTACGCTAACGCCGGGCACGTCGAGGGGCAGTCAGCTCTT TTTATGAGAGATAACGGCATTAGCgcaGGGCTTGTGTTCCATAATAATCCTGAGGGC ACCTGTGGCTTCTGTGTAAATATGACCGAAACACTTCTGCCTGAGAACGCTAAAATG ACTGTCGTACCACCCGAAGGCGCAATCCCAGTTAAACGGGGCGCAACCGGCGAAACC AAAGTATTCACCGGAAACAGCAATAGTCCAAAGTCCCCCACCAAGGGAGGTTGC	70

G1362A	PTPYPNYANAGHVEGQSALFMRDNGISEALVFHNNPEGTCGFCVNMTETLLPENAKM TVVPPEGAIPVKRGATGETKVFTGNSNSPKSPTKGGC	71
	CCAACACCTTATCCTAACTACGCTAACGCCGGGCACGTCGAGGGGCAGTCAGCTCTT TTTATGAGAGATAACGGCATTAGCGAAgcaCTTGTGTTCCATAATAATCCTGAGGGC ACCTGTGGCTTCTGTGTAAATATGACCGAAACACTTCTGCCTGAGAACGCTAAAATG ACTGTCGTACCACCCGAAGGCGCAATCCCAGTTAAACGGGGCGCAACCGGCGAAACC AAAGTATTCACCGGAAACAGCAATAGTCCAAAGTCCCCCACCAAGGGAGGTTGC	72
L1363A	PTPYPNYANAGHVEGQSALFMRDNGISEGAVFHNNPEGTCGFCVNMTETLLPENAKM TVVPPEGAIPVKRGATGETKVFTGNSNSPKSPTKGGC	73
	CCAACACCTTATCCTAACTACGCTAACGCCGGGCACGTCGAGGGGCAGTCAGCTCTT TTTATGAGAGATAACGGCATTAGCGAAGGGgcaGTGTTCCATAATAATCCTGAGGGC ACCTGTGGCTTCTGTGTAAATATGACCGAAACACTTCTGCCTGAGAACGCTAAAATG ACTGTCGTACCACCCGAAGGCGCAATCCCAGTTAAACGGGGCGCAACCGGCGAAACC AAAGTATTCACCGGAAACAGCAATAGTCCAAAGTCCCCCACCAAGGGAGGTTGC	74
V1364A	PTPYPNYANAGHVEGQSALFMRDNGISEGLAFHNNPEGTCGFCVNMTETLLPENAKM TVVPPEGAIPVKRGATGETKVFTGNSNSPKSPTKGGC	75
	CCAACACCTTATCCTAACTACGCTAACGCCGGGCACGTCGAGGGGCAGTCAGCTCTT TTTATGAGAGATAACGGCATTAGCGAAGGGCTTgcaTTCCATAATAATCCTGAGGGC ACCTGTGGCTTCTGTGTAAATATGACCGAAACACTTCTGCCTGAGAACGCTAAAATG ACTGTCGTACCACCCGAAGGCGCAATCCCAGTTAAACGGGGCGCAACCGGCGAAACC AAAGTATTCACCGGAAACAGCAATAGTCCAAAGTCCCCCACCAAGGGAGGTTGC	76
F1365A	PTPYPNYANAGHVEGQSALFMRDNGISEGLVAHNNPEGTCGFCVNMTETLLPENAKM TVVPPEGAIPVKRGATGETKVFTGNSNSPKSPTKGGC	77
	CCAACACCTTATCCTAACTACGCTAACGCCGGGCACGTCGAGGGGCAGTCAGCTCTT TTTATGAGAGATAACGGCATTAGCGAAGGGCTTGTGgcaCATAATAATCCTGAGGGC ACCTGTGGCTTCTGTGTAAATATGACCGAAACACTTCTGCCTGAGAACGCTAAAATG ACTGTCGTACCACCCGAAGGCGCAATCCCAGTTAAACGGGGCGCAACCGGCGAAACC AAAGTATTCACCGGAAACAGCAATAGTCCAAAGTCCCCCACCAAGGGAGGTTGC	78
H1366A	PTPYPNYANAGHVEGQSALFMRDNGISEGLVFANNPEGTCGFCVNMTETLLPENAKM TVVPPEGAIPVKRGATGETKVFTGNSNSPKSPTKGGC	79
	CCAACACCTTATCCTAACTACGCTAACGCCGGGCACGTCGAGGGGCAGTCAGCTCTT TTTATGAGAGATAACGGCATTAGCGAAGGGCTTGTGTTCgcaAATAATCCTGAGGGC ACCTGTGGCTTCTGTGTAAATATGACCGAAACACTTCTGCCTGAGAACGCTAAAATG ACTGTCGTACCACCCGAAGGCGCAATCCCAGTTAAACGGGGCGCAACCGGCGAAACC AAAGTATTCACCGGAAACAGCAATAGTCCAAAGTCCCCCACCAAGGGAGGTTGC	80
K1389A	PTPYPNYANAGHVEGQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAAM TVVPPEGAIPVKRGATGETKVFTGNSNSPKSPTKGGC	81
	CCAACACCTTATCCTAACTACGCTAACGCCGGGCACGTCGAGGGGCAGTCAGCTCTT TTTATGAGAGATAACGGCATTAGCGAAGGGCTTGTGTTCCATAATAATCCTGAGGGC ACCTGTGGCTTCTGTGTAAATATGACCGAAACACTTCTGCCTGAGAACGCTgcaATG ACTGTCGTACCACCCGAAGGCGCAATCCCAGTTAAACGGGGCGCAACCGGCGAAACC AAAGTATTCACCGGAAACAGCAATAGTCCAAAGTCCCCCACCAAGGGAGGTTGC	82
M1390A	PTPYPNYANAGHVEGQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKA TVVPPEGAIPVKRGATGETKVFTGNSNSPKSPTKGGC	83
	CCAACACCTTATCCTAACTACGCTAACGCCGGGCACGTCGAGGGGCAGTCAGCTCTT TTTATGAGAGATAACGGCATTAGCGAAGGGCTTGTGTTCCATAATAATCCTGAGGGC	84

	ACCTGTGGCTTCTGTGTAAATATGACCGAAACACTTCTGCCTGAGAACGCTAAAgca ACTGTCGTACCACCCGAAGGCGCAATCCCAGTTAAACGGGGCGCAACCGGCGAAACC AAAGTATTCACCGGAAACAGCAATAGTCCAAAGTCCCCCACCAAGGGAGGTTGC	
T1391A	PTPYPNYANAGHVEGQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKM AVVPPEGAIPVKRGATGETKVFTGNSNSPKSPTKGGC	85
	CCAACACCTTATCCTAACTACGCTAACGCCGGGCACGTCGAGGGGCAGTCAGCTCTT TTTATGAGAGATAACGGCATTAGCGAAGGGCTTGTGTTCCATAATAATCCTGAGGGC ACCTGTGGCTTCTGTGTAAATATGACCGAAACACTTCTGCCTGAGAACGCTAAAATG gcaGTCGTACCACCCGAAGGCGCAATCCCAGTTAAACGGGGCGCAACCGGCGAAACC AAAGTATTCACCGGAAACAGCAATAGTCCAAAGTCCCCCACCAAGGGAGGTTGC	86
V1393A	PTPYPNYANAGHVEGQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKM TVAPPEGAIPVKRGATGETKVFTGNSNSPKSPTKGGC	87
	CCAACACCTTATCCTAACTACGCTAACGCCGGGCACGTCGAGGGGCAGTCAGCTCTT TTTATGAGAGATAACGGCATTAGCGAAGGGCTTGTGTTCCATAATAATCCTGAGGGC ACCTGTGGCTTCTGTGTAAATATGACCGAAACACTTCTGCCTGAGAACGCTAAAATG ACTGTCgcaCCACCCGAAGGCGCAATCCCAGTTAAACGGGGCGCAACCGGCGAAACC AAAGTATTCACCGGAAACAGCAATAGTCCAAAGTCCCCCACCAAGGGAGGTTGC	88

【００７８】

[DddAtox G1397-N変異体]

Type	SEQ	No.
C1376A	GSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGGPTPYPNYANAGHV EGQSALFMRDNGISEGLVFHNNPEGTCGFAVNMTETLLPENAKMTVVPPEG	89
	GGTAGCTACGCACTTGGTCCTTACCAGATTAGCGCACCCCAACTCCCCGCCTATAAT GGTCAAACCGTCGGGACCTTTTACTACGTAAACGATGCTGGTGGGCTGGAATCCAAA GTATTCTCCTCAGGGGGCCCTACACCCTACCCCAACTACGCCAATGCTGGTCATGTA GAAGGGCAGTCAGCACTGTTTATGCGCGATAATGGTATAAGCGAGGGGTTGGTCTTC CATAACAACCCAGAGGGTACTTGTGGCTTCgcaGTGAATATGACTGAAACCCTTCTG CCCGAAAATGCCAAGATGACTGTCGTCCCACCTGAAGGC	90
V1377A	GSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGGPTPYPNYANAGHV EGQSALFMRDNGISEGLVFHNNPEGTCGFCANMTETLLPENAKMTVVPPEG	91
	GGTAGCTACGCACTTGGTCCTTACCAGATTAGCGCACCCCAACTCCCCGCCTATAAT GGTCAAACCGTCGGGACCTTTTACTACGTAAACGATGCTGGTGGGCTGGAATCCAAA GTATTCTCCTCAGGGGGCCCTACACCCTACCCCAACTACGCCAATGCTGGTCATGTA GAAGGGCAGTCAGCACTGTTTATGCGCGATAATGGTATAAGCGAGGGGTTGGTCTTC CATAACAACCCAGAGGGTACTTGTGGCTTCTGTgcaAATATGACTGAAACCCTTCTG CCCGAAAATGCCAAGATGACTGTCGTCCCACCTGAAGGC	92
T1380A	GSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGGPTPYPNYANAGHV EGQSALFMRDNGISEGLVFHNNPEGTCGFCVNMAETLLPENAKMTVVPPEG	93
	GGTAGCTACGCACTTGGTCCTTACCAGATTAGCGCACCCCAACTCCCCGCCTATAAT GGTCAAACCGTCGGGACCTTTTACTACGTAAACGATGCTGGTGGGCTGGAATCCAAA GTATTCTCCTCAGGGGGCCCTACACCCTACCCCAACTACGCCAATGCTGGTCATGTA GAAGGGCAGTCAGCACTGTTTATGCGCGATAATGGTATAAGCGAGGGGTTGGTCTTC CATAACAACCCAGAGGGTACTTGTGGCTTCTGTGTGAATATGgcaGAAACCCTTCTG CCCGAAAATGCCAAGATGACTGTCGTCCCACCTGAAGGC	94
E1381A	GSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGGPTPYPNYANAGHV EGQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTATLLPENAKMTVVPPEG	95
	GGTAGCTACGCACTTGGTCCTTACCAGATTAGCGCACCCCAACTCCCCGCCTATAAT GGTCAAACCGTCGGGACCTTTTACTACGTAAACGATGCTGGTGGGCTGGAATCCAAA GTATTCTCCTCAGGGGGCCCTACACCCTACCCCAACTACGCCAATGCTGGTCATGTA GAAGGGCAGTCAGCACTGTTTATGCGCGATAATGGTATAAGCGAGGGGTTGGTCTTC CATAACAACCCAGAGGGTACTTGTGGCTTCTGTGTGAATATGACTgcaACCCTTCTG CCCGAAAATGCCAAGATGACTGTCGTCCCACCTGAAGGC	96
L1384A	GSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGGPTPYPNYANAGHV EGQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLAPENAKMTVVPPEG	97
	GGTAGCTACGCACTTGGTCCTTACCAGATTAGCGCACCCCAACTCCCCGCCTATAAT GGTCAAACCGTCGGGACCTTTTACTACGTAAACGATGCTGGTGGGCTGGAATCCAAA GTATTCTCCTCAGGGGGCCCTACACCCTACCCCAACTACGCCAATGCTGGTCATGTA GAAGGGCAGTCAGCACTGTTTATGCGCGATAATGGTATAAGCGAGGGGTTGGTCTTC CATAACAACCCAGAGGGTACTTGTGGCTTCTGTGTGAATATGACTGAAACCCTTgca CCCGAAAATGCCAAGATGACTGTCGTCCCACCTGAAGGC	98
K1389A	GSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGGPTPYPNYANAGHV EGQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAAMTVVPPEG	99

	GGTAGCTACGCACTTGGTCCTTACCAGATTAGCGCACCCCAACTCCCCGCCTATAAT GGTCAAACCGTCGGGACCTTTTACTACGTAAACGATGCTGGTGGGCTGGAATCCAAA GTATTCTCCTCAGGGGGCCCTACACCCTACCCCAACTACGCCAATGCTGGTCATGTA GAAGGGCAGTCAGCACTGTTTATGCGCGATAATGGTATAAGCGAGGGGTTGGTCTTC CATAACAACCCAGAGGGTACTTGTGGCTTCTGTGTGAATATGACTGAAACCCTTCTG CCCGAAAATGCCgcaATGACTGTCGTCCCACCTGAAGGC	100
M1390A	GSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGGPTPYPNYANAGHV EGQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKATVVPPEG	101
	GGTAGCTACGCACTTGGTCCTTACCAGATTAGCGCACCCCAACTCCCCGCCTATAAT GGTCAAACCGTCGGGACCTTTTACTACGTAAACGATGCTGGTGGGCTGGAATCCAAA GTATTCTCCTCAGGGGGCCCTACACCCTACCCCAACTACGCCAATGCTGGTCATGTA GAAGGGCAGTCAGCACTGTTTATGCGCGATAATGGTATAAGCGAGGGGTTGGTCTTC CATAACAACCCAGAGGGTACTTGTGGCTTCTGTGTGAATATGACTGAAACCCTTCTG CCCGAAAATGCCAAGgcaACTGTCGTCCCACCTGAAGGC	102
T1391A	GSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGGPTPYPNYANAGHV EGQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKMAVVPPEG	103
	GGTAGCTACGCACTTGGTCCTTACCAGATTAGCGCACCCCAACTCCCCGCCTATAAT GGTCAAACCGTCGGGACCTTTTACTACGTAAACGATGCTGGTGGGCTGGAATCCAAA GTATTCTCCTCAGGGGGCCCTACACCCTACCCCAACTACGCCAATGCTGGTCATGTA GAAGGGCAGTCAGCACTGTTTATGCGCGATAATGGTATAAGCGAGGGGTTGGTCTTC CATAACAACCCAGAGGGTACTTGTGGCTTCTGTGTGAATATGACTGAAACCCTTCTG CCCGAAAATGCCAAGATGgcaGTCGTCCCACCTGAAGGC	104
V1392A	GSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGGPTPYPNYANAGHV EGQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKMTAVPPEG	105
	GGTAGCTACGCACTTGGTCCTTACCAGATTAGCGCACCCCAACTCCCCGCCTATAAT GGTCAAACCGTCGGGACCTTTTACTACGTAAACGATGCTGGTGGGCTGGAATCCAAA GTATTCTCCTCAGGGGGCCCTACACCCTACCCCAACTACGCCAATGCTGGTCATGTA GAAGGGCAGTCAGCACTGTTTATGCGCGATAATGGTATAAGCGAGGGGTTGGTCTTC CATAACAACCCAGAGGGTACTTGTGGCTTCTGTGTGAATATGACTGAAACCCTTCTG CCCGAAAATGCCAAGATGACTgcaGTCCCACCTGAAGGC	106

【００７９】

[DddAtox G1397-C変異体]

Type	SEQ	No.
K1410A	AIPVKRGATGETAVFTGNSNSPKSPTKGGC	107
	GCCATACCTGTGAAGCGGGGAGCAACAGGGGAGACAgcaGTGTTCACAGGCAACTCT AACAGTCCAAAGAGCCCCACCAAAGGCGGGTGT	108
V1411A	AIPVKRGATGETKAFTGNSNSPKSPTKGGC	109
	GCCATACCTGTGAAGCGGGGAGCAACAGGGGAGACAAAGgcaTTCACAGGCAACTCT AACAGTCCAAAGAGCCCCACCAAAGGCGGGTGT	110
F1412A	AIPVKRGATGETKVATGNSNSPKSPTKGGC	111
	GCCATACCTGTGAAGCGGGGAGCAACAGGGGAGACAAAGGTGgcaACAGGCAACTCT AACAGTCCAAAGAGCCCCACCAAAGGCGGGTGT	112
T1413A	AIPVKRGATGETKVFAGNSNSPKSPTKGGC	113
	GCCATACCTGTGAAGCGGGGAGCAACAGGGGAGACAAAGGTGTTCgcaGGCAACTCT AACAGTCCAAAGAGCCCCACCAAAGGCGGGTGT	114

【００８０】

本発明によるシトシン脱アミノ酵素変異体は、前記表のように記載されたアミノ酸配列からなる群より選ばれる1つ以上の配列を含んでもよい。本発明によるシトシン脱アミノ酵素変異体は、非特異的標的部位内の多数の塩基に所望しない編集を引き起こす効果を低減する可能性を示唆する。

【００８１】

3．全長シトシン脱アミノ酵素

本出願の発明者らは、細胞内に毒性を引き起こし、分離された形態(split)で使用する野生型シトシン脱アミノ酵素(Cytosine deaminase)$DddA_{tox}$の正電荷を有するアミノ酸を変形して作製した全長DddAを用いた新たなプログラム可能なシトシン脱アミノ酵素を開発した。

【００８２】

本発明は、(i) DNA結合タンパク質；及び(ii) シトシン脱アミノ酵素またはその変異体を含む融合タンパク質であって、前記シトシン脱アミノ酵素またはその変異体は、非毒性全長シトシン脱アミノ酵素である、融合タンパク質に関する。

【００８３】

$DddA_{tox}$のC-末端には正電荷を有するアミノ酸が(KRKKK)特異的に集まっている。DNAは負電荷を有するため、タンパク質の正電荷性質を有するアミノ酸と結合する。該正電荷を有するアミノ酸を電極のないアミノ酸に置換し、$DddA_{tox}$がDNAと結合する力を弱め、細胞内毒性を低下または除去しようとする。すなわち、正電荷を有するアミノ酸を極性のないアミノ酸に置換して毒性を有さない組合は、大腸菌を用いたクローニングが可能であり、全長DddAを確保することができる。

【００８４】

野生型$DddA_{tox}$は、細胞内毒性のために2つに分けた分離された形態を用いるが、これは実験に多くの制約が伴われる。特に、Cas9を用いる場合、他のPAMを認識するオルソゴナル(orthogonal)Cas9変異体を用いる。この場合、PAMが制限的に存在するため、所望の位置に正確にシトシンをチミンに脱アミノ化することが困難な場合が多い。また、最も高い活性が表れるターゲットウィンドウ(target window)は、2つのCas9変異体がバインディングする間の40bpで、この部位の所望しないシトシンまで脱アミノ化される。しかし、全長DddAは分離されないためPAMの制約を受けない。さらに、ターゲット位置で10bp内のTCモチフで最も活性が高いため、正確性も高い。

【００８５】

Cas9がターゲットサイト(target site)に結合してなすR-loopにあるACA、GC、CC

モチーフにあるシトシンをチミンに脱アミノ化することを確認した。これは分離された形態では確認されない活性である。

【００８６】

全長DddAは、Cas9だけでなく、TALEモジュール、ジンクフィンガータンパク質を用いて所望の位置のシトシンをチミンに置換することができる。従来のDddAtoxは分離された形態であるので、対(pair)として伝達すべきであったが、全長DddAはTALEモジュール、ジンクフィンガータンパク質の1つのモジュールのみを用いることができるので、ターゲット位置を制約なく選択できる。また、ゲノムサイト(genomic site)のみならず、ミトコンドリア(mitochondria)や植物葉緑体(chloroplast)または色素体などのDNAをターゲットして、特定DNAのシトシンをチミンに変換することができる。

【００８７】

また、サイズが小さく、遺伝子治療に用いられるウイルスベクターであるAAVに全ての構造体を入れることができる。従来使用していたCBE(cytosine base editor)は、Cas9がターゲットサイトとバインディングしてなすR-loopにあるシトシンをチミンに置換するが、今回発明された全長DddAは、R-loopの外にあるシトシンを脱アミノ化する。したがって、従来のCBEをもっては編集に制約がある位置の場合、シトシンをチミンに変換することができる。

【００８８】

これに基づき、前記非毒性全長シトシン脱アミノ酵素は、配列番号1の野生型脱アミノ酵素のうち、1つ以上、2つ以上、3つ以上、4つ以上又は5つ以上のアミノ酸が他のアミノ酸に置換されていてもよい。

【００８９】

前記「他のアミノ酸」は、アラニン、イソロイシン、ロイシン、メチオニン、フェニルアラニン、プロリン、トリプトファン、バリン、アスパラギン酸、システイン、グルタミン、グリシン、セリン、スレオニン、チロシン、グルタミン酸、アルギン、ヒスチジン、ライシン、前記アミノ酸の公知の全ての変異体の中で、野生型タンパク質が元の変異位置に有するアミノ酸を除くアミノ酸の中から選ばれるアミノ酸を意味する。一例では、前記「他のアミノ酸」はアラニンであり得る。

【００９０】

前記非毒性全長DddAは、種類に応じて次の配列番号12ないし配列番号18からなる群より選ばれる配列を含んでもよい。

【００９１】

Wild type (SEQ ID No: 1)

GSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGGPTPYPNYANAGHV
EGQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKMTVVPPEGAIPVK
RGATGETKVFTGNSNSPKSPTKGGC

【００９２】

A1341D KRKKA (SEQ ID No: 12)

GSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGGPTPYPNYDNAGH
VEGQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKMTVVPPEGAIPV
KRGATGETKVFTGNSNSPKSPTAGGC

【００９３】

AAAAA (SEQ ID No: 13)

GSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGGPTPYPNYANAGHV
EGQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKMTVVPPEGAIPVA
AGATGETAVFTGNSNSPASPTAGGC

【００９４】

AAAAK (SEQ ID No: 14)

GSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGGPTPYPNYANAGHV
EGQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKMTVVPPEGAIPVA

AGATGETAVFTGNSNSPASPTKGGC

【００９５】

AAKAA (SEQ ID No: 15)

GSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGGPTPYPNYANAGHV
EGQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKMTVVPPEGAIPVA
AGATGETKVFTGNSNSPASPTAGGC

【００９６】

AAKAK (SEQ ID No: 16)

GSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGGPTPYPNYANAGHV
EGQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKMTVVPPEGAIPVA
AGATGETKVFTGNSNSPASPTKGGC

【００９７】

KAAAA (SEQ ID No: 17)

GSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGGPTPYPNYANAGHV
EGQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKMTVVPPEGAIPVK
AGATGETAVFTGNSNSPASPTAGGC

【００９８】

E1347A (SEQ ID No: 18)

GSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGGPTPYPNYANAGHV
AGQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKMTVVPPEGAIPVK
RGATGETKVFTGNSNSPKSPTKGGC

【００９９】

前記全長脱アミノ酵素変異体は、配列番号1のアミノ酸配列のうち、以下からなる群より選ばれる1つ以上の置換を含んでもよい：

37位のSがGに置換、

59位のGがSに置換、

109位のAがVに置換、及び

129位のSがGに置換。

【０１００】

一実施形態では、前記全長脱アミノ酵素変異体は、配列番号1のアミノ酸配列のうち、37位のSがGに置換、59位のGがSに置換、109位のAがVに置換、及び129位のSがGに置換を含む、配列番号19の配列を含んでもよい。

【０１０１】

GSVG (SEQ ID No: 19)

GSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLEGKVFSSGGPTPYPNYANAGH
VESQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKMTVVPPEGVIPV
KRGATGETKVFTGNSNGPKSPTKGGC

【０１０２】

前記全長DddA GSVGは、一般的な大腸菌を用いたクローニングが可能である。全長DddA GSVGがターゲットサイトにあるTCモチフのシトシンをチミンに脱アミノ化することをヒト細胞ゲノムレベル(human cell genomic context)で確認した。全長DddA GSVGをCas9のN-末端とC-末端にそれぞれクローニングすることができる。同じターゲットサイトでCas9のN-末端に連結されたDddA GSVGは、シトシンをチミンに置換することができる。Cas9のC-末端に連結されたDddA GSVGは、置換を所望するグアニンの相補配列にあるTCモチフのシトシンをチミンに置換し、置換を所望するグアニンの相補配列にあるTCモチフのシトシンをチミンに置換し、置換を所望するグアニンをアデニンに置換することをヒト細胞ゲノムレベで確認した。

【０１０３】

一実施形態では、前記全長脱アミノ酵素変異体は、配列番号20ないし22からなる群より選ばれる配列を含んでもよい。

【０１０４】
SSVG (SEQ ID No: 20)
GSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGGPTPYPNYANAGHV
ESQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKMTVVPPEGVIPVK
RGATGETKVFTGNSNGPKSPTKGGC
【０１０５】
GSAG (SEQ ID No: 21)
GSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLEGKVFSSGGPTPYPNYANAGH
VESQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKMTVVPPEGAIPV
KRGATGETKVFTGNSNGPKSPTKGGC
【０１０６】
GSVS (SEQ ID No: 22)
GSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLEGKVFSSGGPTPYPNYANAGH
VESQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKMTVVPPEGVIPV
KRGATGETKVFTGNSNSPKSPTKGGC
【０１０７】
４．DNA結合タンパク質
前記DNA結合タンパク質は、例えば、ジンクフィンガータンパク質(zinc finger protein)、TALE(Transcription activator-like effector)タンパク質、CRISPR関連ヌクレアーゼ、またはこれらの2つ以上の組合であってもよい。
【０１０８】
前記ジンクフィンガータンパク質のジンクフィンガーモチーフはDNA結合ドメインを有し、フィンガーのC-末端部分はDNA配列を特異的に認識する。3～6個のジンクフィンガーモチーフを含むDNA結合タンパク質がDNA配列を認識する。
【０１０９】
一実施形態では、前記シトシン脱アミノ酵素の第1の分割体及びシトシン脱アミノ酵素の第2の分割体のそれぞれは、ジンクフィンガータンパク質のN-末端またはC-末端に結合することができる。
【０１１０】
シトシン脱アミノ酵素の第1の分割体N-末端にジンクフィンガータンパク質(ZF-Left)のC-末端が結合し、シトシン脱アミノ酵素の第2の分割体N-末端にジンクフィンガータンパク質(ZF-Right)C-末端が結合(CC configuration)、
シトシン脱アミノ酵素の第1の分割体C-末端にジンクフィンガータンパク質(ZF-Left)のN-末端が結合し、シトシン脱アミノ酵素の第2の分割体N-末端にジンクフィンガータンパク質(ZF-Right)C-末端が結合(NC configuration)、
シトシン脱アミノ酵素の第1の分割体N-末端にジンクフィンガータンパク質(ZF-Left)のC-末端が結合し、シトシン脱アミノ酵素の第2の分割体C-末端にジンクフィンガータンパク質(ZF-Right)N-末端が結合(CN configuration)、または、
シトシン脱アミノ酵素の第1の分割体C-末端にジンクフィンガータンパク質(ZF-Left)のN-末端が結合し、シトシン脱アミノ酵素の第2の分割体C-末端にジンクフィンガータンパク質(ZF-Right)N-末端が結合(NN configuration)することができる。
【０１１１】
前記ZF-Leftは、以下の配列番号2の配列を含んでもよい。
【０１１２】
[配列番号2]
GIHGVPAAMAERPFQCRICMRNFSDRSNLSRHIRTHTGEKPFACDICGRKFAISSNLN
SHTKIHTGSQKPFQCRICMRNFSRSDNLARHIRTHTGEKPFACDICGRKFATSGNLTR
HTKIHLR.
【０１１３】
前記ZF-Rightは、以下の配列番号3の配列を含んでもよい。

【０１１４】
[配列番号3]
GIHGVPAAMAERPFQCRICMRNFSRSDNLSVHIRTHTGEKPFACDICGRKFAQKINL
QVHTKIHTGEKPFQCRICMRNFSRSDVLSEHIRTHTGEKPFACDICGRKFAQRNHRT
THTKIHLR.
【０１１５】
ZFは、DNAターゲットによって配列が異になってもよい。ZFは、DNAターゲット配列に応じてカスタマイズできる。ZFは、DNA 3bpを認識するため、通常3～6個のZFを繋ぎ合わせて、9～18bp DNAを認識するZF組合を作製することができる。例えば、GNN、TNN、CNNまたはANNなどのモジュールが含まれたライブラリーを用いて作製することができる。
【０１１６】
場合によっては、前記ジンクフィンガータンパク質は、リンカーを介して脱アミノ酵素と連結され得る。前記リンカーは、2～40個のアミノ酸残基を含むペプチドリンカーであり得る。前記リンカーは、例えば、2aa、5aa、10aa、16aa、24aaまたは32aaの長さのリンカーなどであってもよいが、これに限定されるものではない。
【０１１７】
一実施形態では、リンカーは以下を含んでもよい。
2a.aリンカー：GS、
5a.aリンカー：TGEKQ(配列番号8)、
10a.aリンカー：SGAQGSTLDF(配列番号9)、
16a.aリンカー：SGSETPGTSESATPES(配列番号10)、
24a.aリンカー：SGTPHEVGVYTLSGTPHEVGVYTL(配列番号115)、または、
32a.aリンカー：GSGGSSGGSSGSETPGTSESATPESSGGSSGGS (配列番号11)
【０１１８】
本発明による具体的な実施形態では、分割脱アミノ酵素とジンクフィンガータンパク質がリンカーを介して連結され得て、第1の分割体を含む半分の脱アミノ酵素N-末端にジンクフィンガータンパク質が結合し、第2の分割体を含む半分の脱アミノ酵素N-末端にジンクフィンガータンパク質が結合する。このとき、左側と右側のZFPが付着する位置の間であるスペーサーでCからTへの塩基転換が起こり得る。左側および右側のZFPの全てが24a.aリンカーを介して第1の分割体を含む半分の脱アミノ酵素および第2の分割体を含む半分の脱アミノ酵素のそれぞれと連結される場合、高い編集効率を示すことを確認した。
【０１１９】
前記のTALエフェクター(TALE)は、33個から34個の繰り返されたアミノ酸配列が繰り返された形態からなり、ドメイン(RVD、Repeated variant domain)が約9個程度繰り返される。ドメイン当たり、1つのヌクレオチドを認識することができ、12～13番目のアミノ酸配列に従って特異的なDNA配列と結合することができる(HD→Cytosine、NI→Adenine、NG→Thymine、NN→Guanine)。TALエフェクター(TALE)は、標的部位内の一本のDNA鎖を認識する。標的部位間の距離は12～14ヌクレオチドであり得る。
【０１２０】
前記TALEドメインは、1つ以上のTALE-Repeatの組合によって配列-特異的方式によりヌクレオチドに結合するタンパク質ドメインを意味する。少なくとも1つ以上のTALE-Repeat、具体的には、1～30個のTALE-Repeatを含むが、これに限定されるものではない。TALE-Repeatは、TALEドメイン内の特定のヌクレオチド配列を認識する部位である。
【０１２１】
前記TALEドメインは、骨格構造としてのTALEのN-末端包含領域とTALEのC-末端包含領域が含む。TALEのN-末端包含の第1のTALEは、配列番号4または5によってコード化することができる。前記TALEのC-末端包含むの第2のTALEは配列番号6又は7により

コード化することができる。

【０１２２】

【表１】

Name	Sequence	No.
第1 TALE N-term	GATATAGCTGATTTGAGGACCCTCGGGTACTCACAGCAGCAACAAGAAAAAATCAAACCAAA GGTACGTAGCACGGTTGCCCAACATCATGAAGCCCTAGTTGGTCACGGATTTACTCATGCCC ATATTGTTGCACTGAGTCAACACCCTGCTGCATTAGGCACAGTGGCTGTGAAATATCAAGAT ATGATTGCTGCTCTTCCTGAGGCCACACACGAGGCAATTGTGGGTGTTGGAAAACAATGGTC TGGTGCGCGAGCGCTTGAAGCTCTCCTAACCGTAGCTGGAGAGTTGAGAGGCCCACCGCTTC AGCTTGACACAGGGCAGCTGTTGAAGATAGCAAAGAGGGGGGGTGTAACTGCTGTTGAAGCA GTCCATGCATGGAGAAATGCATTAACTGGAGCTCCA	4
	DIADLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVKYQD MIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEA VHAWRNALTGAP	269
第2 TALE N-term	GATATTGCTGATCTTCGAACATTGGGTTATTCTCAGCAGCAACAAGAGAAAATCAAACCGAA AGTCCGTTCAACCGTAGCTCAACATCATGAAGCCCTTGTTGGTCATGGATTTACTCATGCTC ACATTGTTGCTCTAAGTCAGCACCCTGCCGCGCTAGGCACAGTTGCCGTCAAGTATCAAGAT ATGATAGCTGCTTTACCTGAAGCAACCCACGAGGCAATAGTGGGGGTTGGGAAGAGGGGTGC AGGCGCCAGGGCACTGGAGGCACTCTTAACAGTAGCTGGAGAACTTCGGGGACCACCATTGC AGCTTGATACTGGACAACTCTTGAAGATCGCAAAAAGAGGGGGTGTGACTGCTGTGGAAGCT GTTCATGCATGGAGAAATGCTTTGACGGGAGCACCCCTG	5
	DIADLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVKYQD MIAALPEATHEAIVGVGKRGAGARALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEA VHAWRNALTGAPL	270
第1 TALE C-term	AGCATAGTTGCTCAGCTAAGTAGACCTGATCCAGCACTTGCTGCATTGACGAACGATCATCT CGTAGCGCTGGCTTGTTTGGGCGGCAGGCCAGCCTTGGACGCCGTGAAGAAAGGATTAGGT	6
	SIVAQLSRPDPALAALTNDHLVALACLGGRPALDAVKKGLG	271
第2 TALE C-term	TCGATAGTAGCACAATTAAGCAGGCCTGACCCTGCACTTGCTGCTCTTACAAATGATCATCT TGTTGCTCTGGCCTGTTTGGGTGGGCGTCCCGCGTTGGATGCTGTTAAGAAAGGACTTGGT	7
	SIVAQLSRPDPALAALTNDHLVALACLGGRPALDAVKKGLG	271

【０１２３】

切断部位に基づき、前記TALEドメインが結合する位置に応じて、単一のTALEアレイ又は第1のTALEアレイおよび第2のTALEアレイがそれぞれ結合することができる。

【０１２４】

前記シトシン脱アミノ酵素の第1の分割体に第1TALEが結合(left TALE)し、シトシン脱アミノ酵素の第2の分割体に第2のTALE(right TALE)が結合することができる。それぞれN’-TALE-第1の分割体-C’およびN’-TALE-第2の分割体?C’構造を有する。

【０１２５】

前記シトシン脱アミノ酵素が全長である場合、シトシン脱アミノ酵素のN-末端に単一のモジュールTALEが結合することができる。N-C方向に単一のTALEモジュールおよびシトシン脱アミノ酵素が含まれる。デュアルモジュールTALEが含まれてもよく、このとき、全長シトシン脱アミノ酵素のN-末端に第1のTALEが結合し、第2のTALEが別個に含まれてもよい。N-C方向に第1のTALEモジュールおよびシトシン脱アミノ酵素が含まれ、N’-TALE-シトシン脱アミノ酵素-C’とN’-TALE-C’の構造を有する。

【０１２６】
TALEアレイは、ターゲットDNA配列に合わせてカストマイズすることができる。TALEアレイは、33～35個のアミノ酸残基からなるモジュールが反復的に配列して構成されるが、これらは、植物病原菌Xanthmonasに由来するものであって、モジュール1つがA、C、G、T塩基1つずつをそれぞれ認識してDNAに結合する。各々のモジュールの塩基特異性は、12、13番目のアミノ酸残基によって定められるが、それらをrepeat variable di-residue(RVD)という。例えば、RVDがNNのモジュールはGを認識し、NIはA、HDはC、NGはTを認識する。TALEアレイは、少なくとも14個から最大18個以上のモジュールで構成され、ターゲットDNA配列15bp～20bpを認識できるように設計することができる。
【０１２７】
前記CRISPR関連ヌクレアーゼに関して、CRISPRアレイには2つのRNAが暗号化されているが、一方はcrRNA(CRISPR RNA)であり、他方はtracRNA(trans activating CRPSPR RNA)である。crRNAは、プロトスペース部分で転写され、tracRNAと結合して三次構造を形成する。この2種のRNAは、外部DNAを認識して切断するのに役立つ。
【０１２８】
前記Casタンパク質は、Cas1、Cas1B、Cas2、Cas3、Cas4、Cas5、Cas6、Cas7、Cas8、Cas9、Cas10、Cas12a、Cas12b、Cas12c、Cas12d、Cas12e、Cas12g、Cas12h、Cas12i、Cas12j、Cas13a、Cas13b、Cas13c、cas13d、Cas14、Csy1、Csy2、Csy3、Cse1、Cse2、Csc1、Csc2、Csa5、Csn2、CsMT2、Csm3、Csm4、Csm5、Csm6、Cmr1、Cmr3、Cmr4、Cmr5、Cmr6、Csb1、Csb2、Csb3、Csx17、Csx14、Csx10、Csx16、CsaX、Csx3、Csx1、Csx15、Csf1、Csf2、Csf3またはCsf4のエンドヌクレアーゼであってもよいが、これに限定されるものではない。
【０１２９】
前記Casタンパク質は、コリネバクター(Corynebacter)、ステレラ(Sutterella)、レジオネラ(Legionella)、トレポネーマ(Treponema)、フィリファクター(Filifactor)、ユーバクテリウム(Eubacterium)、ストレプトコッカス(Streptococcus：Streptococcus pyogenes)、ラクトバチルス(Lactobacillus)、ミコプラズマ(Mycoplasma)、バクテロイド(Bacteroides)、フラビイボラ属 (Flaviivola)、フラボバクテリウム(Flavobacterium)、アゾスピリルム (Azospirillum)、グルコナセトバクター(Gluconacetobacter)、ナイセリア(Neisseria) ロゼブリア(Roseburia)、パービバキュラム(Parvibaculum)、 スタフィロコッカス(Staphylococcus：Staphylococcus aureus)、 ニトラティフラクター(Nitratifractor)、コリネバクテリウム(Corynebacterium)及びカンピロバクター(Campylobacter)からなる群より選ばれるCasタンパク質のオソログを含む微生物属に由来し、これらから単純分離されたものまたは組換えされたものであってもよい。
【０１３０】
変異した形態のCasタンパク質を含んでもよい。DNA二本鎖を切断するエンドヌクレアーゼ活性を失うように変異したことを意味することができ、例えば、エンドヌクレアーゼ活性を失い、ニッカーゼ活性を有するように変異した変異標的特異的ヌクレアーゼおよびエンドヌクレア活性とニッカーゼ活性をいずれも失うように変異した形態の1つ以上であってもよい。
【０１３１】
ニッカーゼ活性を有するものである場合、前記シトシン脱アミノ酵素による塩基変換(例えば、シトシンがウラジンに変換)と同時または順序に関係なく順次に、前記塩基変換が起こった鎖またはその逆鎖(例えば、塩基変換が起こった鎖の逆鎖)でニック(nick)が導入されることがある(例えば、PAMが位置する鎖の逆鎖において、PAM配列の5’末端方向に3番目のヌクレオチドと4番目のヌクレオチドとの間に該当する位置にニックが導入される)。このような変異(例えば、アミノ酸置換など)は、触媒活性ドメイン(例えば、

Cas9の場合はRuvC触媒ドメイン)で起こり得る。ストレプトコッカス・ピオゲネス(Streptococcus pyogenes)由来Cas9の場合、変異は触媒活性を有するアスパラギン酸残基(catalytic aspartate residue；10位のアスパラギン酸(D10)など)、762位のグルタミン酸(E762)、840位のヒスチジン(H840)、854位のアスパラギン(N854)、863位のアスパラギン(N863)、986位のアスパラギン酸(D986)などからなる群より選ばれる1つ以上の任意の他のアミノ酸に置換された突然変異を含んでもよい。このとき、置換される任意の他のアミノ酸はアラニンであってもよいが、これに限定されない。

【０１３２】

場合によっては、ストレプトコッカス・ピオゲネス由来Cas9タンパク質の1135位のアスパラギン酸(D1135)、1335位のアルギニン(R1335)、および1337位のトレオニン(T1337)のうちの1つ以上、例えば、3つがいずれも他のアミノ酸に置換されて、野生型Cas9のPAM配列(NGG)と異なるNGA(Nは、A、T、G、およびCから選ばれる任意の塩基である)を認識するように変異したものであってもよい。

【０１３３】

例えば、ストレプトコッカス・ピオゲネス由来Cas9タンパク質のアミノ酸配列のうち、

(1) D10、H840、またはD10+H840、

(2) D1135、R1335、T1337、またはD1135+R1335+T1337、または、

(3) (1)と(2)残基の両方でアミノ酸置換が起こり得る。

【０１３４】

前記「他のアミノ酸」は、アラニン、イソロイシン、ロイシン、メチオニン、フェニルアラニン、プロリン、トリプトファン、バリン、アスパラギン酸、システイン、グルタミン、グリシン、セリン、スレオニン、チロシン、グルタミン酸、アルギニン、ヒスチジン、ライシン、前記アミノ酸の公知の全ての変異体の中で、野生型タンパク質が元の変異位置に有するアミノ酸を除くアミノ酸の中から選択されるアミノ酸を意味する。一例では、前記「他のアミノ酸」は、アラニン、バリン、グルタミン、またはアルギニンでってもよい。

【０１３５】

場合によっては、ガイドRNAをさらに含むことができる。前記ガイドRNAは、例えば、CRISPR RNA(crRNA)、トランス活性化crRNA(tracrRNA)、および単一のガイドRNA(シングルガイドRNA；sgRNA)からなる群より選ばれる1種以上であってもよく、具体的に、crRNAとtracrRNAが互いに結合した二本鎖crRNA：tracrRNA複合体、またはcrRNAまたはその一部とtracrRNAまたはその一部がオリゴヌクレオチドリンカーで連結された一本鎖ガイドRNA(sgRNA)であってもよい。

【０１３６】

5. アデニン脱アミノ酵素の付加

本発明は、(i) DNA結合タンパク質、(ii) シトシン脱アミノ酵素またはその変異体と、(iii) アデニン脱アミノ酵素を含む融合タンパク質であって、前記シトシン脱アミノ酵素又はその変異体は、(a) 非毒性全長シトシン脱アミノ酵素又は(b) シトシン脱アミノ酵素又はその変異体由来である第1の分割体と第2の分割体を含み、前記第1の分割体と第2の分割体は、それぞれ前記DNA結合タンパク質に結合する形態である融合タンパク質に関する。

【０１３７】

本出願の発明者らは、DNAに結合できるTALEまたはZFPタンパク質に、DddAtoxシトシン脱アミノ酵素とA-to-G転換を起こし得るアデニン脱アミノ酵素を連結してA塩基を編集できる塩基編集器(base editor)を作製した。

【０１３８】

従来のDddAtoxを活用したデアミナーゼ(DdCBE)の場合、DNA バインディングモジュールとしてTALEリピートを使用するシトシン脱アミノ酵素である。C-to-Tのみを引き起こすDdCBEとは異なり、DdABEはA-to-Gを引き起こすことができるので、異なる

変異パターンを作り出すことができる。

【０１３９】

DdABEは、それ自体で二本鎖DNAを認識して脱アミノ化を引き起こすため、RNAなどの付加的なコンポーネントがない。ミトコンドリアやクロロプラストなどの場合、RNAをデリバリーするメカニズムは知られていないので、Crispr システムを適用できなかった。しかし、このようなコンポーネントのないDdABEは、セル内のゲノムDNAをターゲットできるだけでなく、ミトコンドリアやクロロプラストのような小器官のDNAをターゲットして、特定のDNAのA-to-G転換を引き起こすができる。

【０１４０】

現在ミトコンドリアや細胞小器官をターゲットする遺伝子編集技術は、DdCBEが唯一である。したがって、従来のすべての技術を合わせて導入できる変異はC-to-Tしかできなかったが、DdABEはA-to-Gを引き起こすことができるため、変異を導入できるスペクトルがはるかに多様になる。これにより、これまで不可能であったミトコンドリア疾患モデルの作製または治療を可能にする。

【０１４１】

従来のDdCBEのような場合、2つのTALEモジュール(それぞれ左側と右側に付着する)が必要なので、遺伝子治療において、遺伝子キャパシティーの低いウイルスベクターであるAAVに搭載することはできない。しかし、DdABEは1つのTALEモジュールのみを使用できる「シングルモジュール」として使用できるため、AAVウイルスに搭載でき、遺伝子治療に強みを持つ。

【０１４２】

DdABEは、必要に応じて分割されたDddAtoxをもって使用することもでき、完全長DddAtoxバリアントをもって使用することもできるため、互換性が高い。

【０１４３】

前記アデニン脱アミノ酵素は、例えば、APOBEC1(apolipoprotein B editing complex 1)、AID (activation-induced deaminase) 及びtadA(tRNA-specific adenosine deaminase)からなる群より選択することができ、具体的には、tadA (tRNA-specific adenosine deaminase)であってもよい。前記アデニン脱アミノ酵素は、例えば、大腸菌TadA の変異体としてデオキシーアデニン脱アミノ酵素(deoxy-adenine deaminase)であってもよい。

【０１４４】

前記シトシン脱アミノ酵素が分割形態で含まれ、前記DNA結合タンパク質がジンクフィンガータンパク質であって、シトシン脱アミノ酵素の第1の分割体C-末端にジンクフィンガータンパク質(ZF-Left)のN-末端が結合し、シトシン脱アミノ酵素の第2の分割体N-末端にジンクフィンガータンパク質(ZF-Right)C-末端が結合(NC configuration)した構造で、アデニン脱アミノ酵素は、ジンクフィンガータンパク質(ZF-Left)のC-末端、シトシン脱アミノ酵素の第1の分割体のN-末端またはC-末端、ジンクフィンガータンパク質(ZF-Right)のN-末端、シトシン脱アミノ酵素の第2の分割体のN-末端またはC-末端に結合することができる。

【０１４５】

アデニン脱アミノ酵素は、シトシン脱アミノ酵素の第1の分割体N-末端にジンクフィンガータンパク質(ZF-Left)のC-末端が結合し、シトシン脱アミノ酵素の第2の分割体N-末端にジンクフィンガータンパク質(ZF-Right)C-末端が結合する(CC configuration)；シトシン脱アミノ酵素の第1の分割体N-末端にジンクフィンガータンパク質(ZF-Left)のC-末端が結合し、シトシン脱アミノ酵素の第2の分割体C-末端にジンクフィンガータンパク質(ZF-Right)N-末端が結合(CN configuration)；または、シトシン脱アミノ酵素の第1の分割体C-末端にジンクフィンガータンパク質(ZF-Left)のN-末端が結合し、シトシン脱アミノ酵素の第2の分割体C-末端にジンクフィンガータンパク質(ZF-Right)N-末端が結合(NN configuration)においても、ジンクフィンガータンパク質(ZF-Left)のC-末端、シトシン脱アミノ酵素の第1の分割体のN-末端またはC-末端、ジンクフィンガー

タンパク質(ZF-Right)のN-末端、シトシン脱アミノ酵素の第2の分割体のN-末端またはC-末端に結合することができる。

【０１４６】

前記シトシン脱アミノ酵素が分割形態で含まれ、前記DNA結合タンパク質が TALEの場合、前記シトシン脱アミノ酵素の第1の分割体に第1のTALEが結合し、前記シトシン脱アミノ酵素の第2の分割体に第2のTALEが結合し、それぞれN’-TALE-第1の分割体DDDA-C’およびN’-TALE-第2の分割体DDDA-C’構造を有する。アデニン脱アミノ酵素は、シトシン脱アミノ酵素の第1の分割のN-末端またはC-末端またはシトシン脱アミノ酵素の第2の分割体N-末端またはC-末端に結合することができる。

【０１４７】

前記シトシン脱アミノ酵素が全長形態で含まれ、前記DNA結合タンパク質がTALEの場合、単一のTALEモジュールを含むことができ、N-C方向に単一のTALEモジュールとシトシン脱アミノ酵素が含まれ、このとき、アデニン脱アミノ酵素は、単一のTALEモジュールのC-末端方向、シトシン脱アミノ酵素のN-末端またはC-末端に結合することができる。

【０１４８】

前記シトシン脱アミノ酵素が全長形態で含まれ、前記DNA結合タンパク質がTALEの場合、デュアルTALEモジュールを含むことができ、N-C方向に第1のTALEモジュールおよびシトシン脱アミノ酵素が含まれ、アデニン脱アミノ酵素および第2のTALEを含む第2の分割体がさらに含まれてもよく、N’-TALE-シトシン脱アミノ酵素-C’およびN’-TALE-アデニン脱アミノ酵素-C’の構造でアデニン脱アミノ酵素はTALEのN末端またはC末端に結合することができる。

【０１４９】

場合によっては、塩基編集効率を高めることができるUGI(uracil DNA glycosylase inhibitor)をさらに含むことができる。UGIは、DNAからUを除去する作用を触媒する突然変異DNAを修復する酵素UDG(Uracil DNA glycosylase)活性を阻害し、塩基編集効率を高めることができる。

【０１５０】

本発明は、前記融合タンパク質またはそれをコードする核酸を含む原核細胞または真核細胞におけるA-to-G塩基編集用組成物であって、前記DNA結合タンパク質は、ジンクフィンガータンパク質、TALEタンパク質またはCRISPR関連ヌクレアーゼであり、前記融合タンパク質のシトシン脱アミノ酵素またはその変異体は、バクテリア由来であり、二本鎖DNA特異的であることを特徴とする組成物に関する。

【０１５１】

本発明は、前記融合タンパク質またはそれをコードする核酸を含む原核細胞または真核細胞におけるA-to-G塩基編集用組成物であって、前記DNA結合タンパク質は、ジンクフィンガータンパク質、TALEタンパク質またはCRISPR関連ヌクレアーゼであり、

前記融合タンパク質のシトシン脱アミノ酵素またはその変異体はバクテリア由来であり、二本鎖DNA特異的なものであり、前記シトシン脱アミノ酵素またはその変異体のN-末端にDNA結合タンパク質が結合しており、前記融合タンパク質のアデニン脱アミノ酵素のC-末端にDNA結合タンパク質が結合している、組成物に関する。

【０１５２】

本発明は、前記融合タンパク質またはそれをコードする核酸およびUGI(uracil glycosylase inhibitor)を含む原核細胞または真核細胞におけるC-to-T塩基編集用組成物であって、前記DNA結合タンパク質は、ジンクフィンガータンパク質、TALEタンパク質またはCRISPR関連ヌクレアーゼであり、前記融合タンパク質のシトシン脱アミノ酵素またはその変異体はバクテリア由来であり、二本鎖DNA特異的であることを特徴とする組成物に関する。

【０１５３】

具体的には、本発明は、1）DNA結合タンパク質、2）全長二本鎖DNA特異的バクテリ

アシトシン脱アミノ酵素またはその変異体、および3) E. coli TadA由来デオキシアデニン脱アミノ酵素を含む原核細胞および真核細胞におけるA-to-G塩基編集用組成物(UGIなし)であって、前記DNA結合タンパク質は、ジンクフィンガータンパク質または転写活性化因子類似エフェクター(TALE)アレイまたは触媒欠損CRISPR-Cas9(nCas9またはdCas9)またはCas12aであり、全長二本鎖DNA特異的バクテリアシトシン脱アミノ酵素Burkholderia cenocepacia由来DddAtoxである組成物である。

【０１５４】

本発明は、1) 全長二本鎖DNA特異的バクテリアシトシン脱アミノ酵素またはその変異体に機能可能に連結された左側のDNA結合タンパク質2) E. coli TadA 由来デオキシアデニン脱アミノ酵素に機能可能に連結された右側のDNA結合タンパク質を含む原核細胞および真核細胞におけるA-to-G塩基編集用組成物(UGIなし)であって、左側または右側のDNA結合タンパク質は、ジンクフィンガータンパク質または転写活性化因子類似エフェクター(TALE)アレイまたは触媒欠損CRISPR-Cas9(nCas9またはdCas9)またはCas12aであり、全長二本鎖DNA特異的バクテリアシトシン脱アミノ酵素Burkholderia cenocepacia由来DddAtoxである組成物である。

【０１５５】

本発明は、また、1) DNA結合タンパク質2) 全長二本鎖DNA特異的バクテリアシトシン脱アミノ酵素またはその変異体、3) E. coli TadA 由来デオキシアデニン脱アミノ酵素および4) UGI(uracil glycosylase inhibitor)を含む原核細胞および真核細胞におけるA-to-GおよびC-to-T塩基編集用組成物であって、前記DNA結合タンパク質は、ジンクフィンガータンパク質または転写活性化因子類似エフェクター(TALE)アレイまたは触媒欠損CRISPR-Cas9(nCas9またはdCas9)またはCas12aであり、全長二本鎖DNA特異的バクテリアシトシン脱アミノ酵素Burkholderia cenocepacia由来DddAtoxである組成物である。

【０１５６】

本発明は、さらに、1) DNA結合タンパク質、2) 分割された二本鎖DNA特異的バクテリアシトシン脱アミノ酵素またはその変異体、および3) E. coli TadA 由来デオキシアデニン脱アミノ酵素を含む原核細胞および真核細胞におけるA-to-G塩基編集用組成物(UGIなし)であって、前記DNA結合タンパク質は、ジンクフィンガータンパク質または転写活性化因子類似エフェクター(TALE)アレイまたは触媒欠損CRISPR-Cas9(nCas9またはdCas9)またはCas12aであり、分割された二本鎖DNA特異的バクテリアシトシン脱アミノ酵素Burkholderia cenocepacia由来DddAtoxである組成物である。

【０１５７】

本発明は、さらに、1) DNA結合タンパク質2) 分割された二本鎖DNA特異的バクテリアシトシン脱アミノ酵素またはその変異体3) E. coli TadA 由来デオキシアデニン脱アミノ酵素および4) UGI(uracil glycosylase inhibitor)を含む原核細胞および真核細胞におけるA-to-GおよびC-to-T塩基編集用組成物であって、前記DNA結合タンパク質は、ジンクフィンガータンパク質または転写活性化因子類似エフェクター(TALE)アレイまたは触媒欠損CRISPR-Cas9(nCas9またはdCas9)またはCas12aであり、分割された二本鎖DNA特異的バクテリアシトシン脱アミノ酵素Burkholderia cenocepacia由来DddAtoxである組成物である。

【０１５８】

本発明は、前記融合タンパク質またはそれをコードする核酸を原核細胞または真核細胞に処理する段階を含む原核細胞または真核細胞におけるA-to-G塩基編集方法であって、前記DNA結合タンパク質は、ジンクフィンガータンパク質、TALEタンパク質またはCRISPR関連ヌクレアーゼであり、前記融合タンパク質のシトシン脱アミノ酵素またはその変異体はバクテリア由来であり、二本鎖DNA特異的であることを特徴とする方法である。

【０１５９】

本発明は、前記融合タンパク質またはそれをコードする核酸を原核細胞または真核細胞に処理する段階を含む原核細胞または真核細胞におけるA-to-G塩基編集方法であって、

前記DNA結合タンパク質は、ジンクピンカータンパク質、TALEタンパク質またはCRISPR関連ヌクレアーゼであり、前記融合タンパク質のシトシン脱アミノ酵素またはその変異体はバクテリア由来であり、二本鎖DNA特異的であり、前記シトシン脱アミノ酵素またはその変異体のN-末端にDNA結合タンパク質が結合しており、前記融合タンパク質のアデニン脱アミノ酵素のC-末端にDNA結合タンパク質が結合している、方法である。

【０１６０】

本発明は、前記融合タンパク質またはそれをコードする核酸およびUGI(uracil glycosylase inhibitor)を原核細胞または真核細胞に処理する段階を含む原核細胞または真核細胞におけるC-to-T塩基編集方法であって、前記DNA結合タンパク質は、ジンクフィンガータンパク質、TALEタンパク質またはCRISPR関連ヌクレアーゼであり、前記融合タンパク質のシトシン脱アミノ酵素またはその変異体はバクテリア由来であり、二本鎖DNA特異的であることを特徴とする方法である。

【０１６１】

本発明による組成物または方法に含まれた構成の具体的な配列は以下の通りである。

ND1-ZFP-Right-1397C-AD (図61f-g: SEQ ID No: 410)

MLGFVGRVAAAPASGALRRLTPSASLPPAQLLLRAAPTAVHPVRDYAAQDYKDDDD
KVDEMTKKFGTLTIHDTEKAAEFGIRIPGEKPFQCRICMRNFSDSGNLRVHIRTHTGE
KPYKCPDCGKSFSQSSSLIRHQRTHTGEKPYECDHCGKSFSQSSHLNVHKRTHTGEK
PYRCKYCDRSFSISSNLQRHVRNIHLRSGTPHEVGVYTLSGTPHEVGVYTLAIPVKRG
ATGETKVFTGNSNSPKSPTKGGCSGSETPGTSESATPESSEVEFSHEYWMRHALTLA
KRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRL
IDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNSKRGAAGSLMNVLNYPGMNHRVE
ITEGILADECAALLCDFYRMPRQVFNAQKKAQSSIN

【０１６２】

ND1-ZFP-Left-1397C-UGI (図61f-g: SEQ ID No: 411)

MLGFVGRVAAAPASGALRRLTPSASLPPAQLLLRAAPTAVHPVRDYAAQDYKDDDD
KVDEMTKKFGTLTIHDTEKAAEFGIRIPGEKPFQCRICMRNFSDSGNLRVHIRTHTGE
KPYKCPDCGKSFSQSSSLIRHQRTHTGEKPYECDHCGKSFSQSSHLNVHKRTHTGEK
PYRCKYCDRSFSISSNLQRHVRNIHLRSGTPHEVGVYTLSGTPHEVGVYTLAIPVKRG
ATGETKVFTGNSNSPKSPTKGGCSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVI
GNKPESDILVHTAYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKML

【０１６３】

ND1-ZFP-Right-1397N-UGI (図61f-g: SEQ ID No: 412)

MLGFVGRVAAAPASGALRRLTPSASLPPAQLLLRAAPTAVHPVRDYAAQYPYDVPD
YAVDEMTKKFGTLTIHDTEKAAEFGIHGVPAAMGGSYALGPYQISAPQLPAYNGQTV
GTFYYVNDAGGLESKVFSSGGPTPYPNYANAGHVEGQSALFMRDNGISEGLVFHNN
PEGTCGFCVNMTETLLPENAKMTVVPPEGSGTPHEVGVYTLSGTPHEVGVYTLYKC
PECGKSFSTKNSLTEHQRTHTGEKPYKCPECGKSFSSKKALTEHQRTHTGEKPYECN
YCGKTFSVSSTLIRHQRIHTGEKPYRCKYCDRSFSISSNLQRHVRNIHLRSGGSTNLS
DIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLTSDAP
EYKPWALVIQDSNGENKIKML

【０１６４】

Trac-ZFP-LEFT-1397N-UGI (図61d: SEQ ID No: 413)

MAPKKKRKVGIHGVPAAMGGSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLES
KVFSSGGPTPYPNYANAGHVEGQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTET
LLPENAKMTVVPPEGSGTPHEVGVYTLSGTPHEVGVYTLFQCRICMRKFATSGSLTR
HTKIHTGEKPFQCRICMRNFSRSDHLSTHIRTHTGEKPFACDICGRKFATSSNRTKHT
KIHTHPRAPIPKPFQCRICMRNFSRSDNLSEHIRTHTGEKPFACDICGRKFAWHSSLR
VHTKIHLRSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYD
ESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKML

【０１６５】

Trac-ZFP-right-1397C-AD (図61d: SEQ ID No: 414)

MAPKKKRKVGIHGVPAAMAERPFQCRICMRNFSRSDHLSTHIRTHTGEKPFACDICG
RKFADRSHLARHTKIHTGSQKPFQCRICMRKFALKQHLNEHTKIHTGEKPFQCRICM
RNFSQSGNLARHIRTHTGEKPFACDICGRKFAHNSSLKDHTKIHLRSGTPHEVGVYT
LSGTPHEVGVYTLAIPVKRGATGETKVFTGNSNSPKSPTKGGCSGSETPGTSESATPE
SSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNSKRG
AAGSLMNVLNYPGMNHRVEITEGILADECAALLCDFYRMPRQVFNAQKKAQSSIN

【０１６６】

ND1-Left-TALE-1397N-UGI (図62: SEQ ID No: 415)

MALSRAVCGTSRQLAPVLGYLGSRQKHSLPDYPYDVPDYAGYPYDVPDYAGYPYDV
PDYAGIRIQDLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPA
ALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDT
GQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPDQVVAIASHDGGKQALETVQRL
LPVLCQAHGLTPDQVVAIASNGGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNIG
GKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQAHGLT
PDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQR
LLPVLCQAHGLTPAQVVAIASNGGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNI
GGKQALETVQRLLPVLCQAHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQAHGL
TPAQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQ
RLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPAQVVAIAS
NGGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNGGGKQALETVQRLLPVLCQAH
GLTPEQVVAIASNGGGKQALESIVAQLSRPDPALAALLVGSGSYALGPYQISAPQLPA
YNGQTVGTFYYVNDAGGLESKVFSSGGPTPYPNYANAGHVEGQSALFMRDNGISEG
LVFHNNPEGTCGFCVNMTETLLPENAKMTVVPPEGSGGSTNLSDIIEKETGKQLVIQ
ESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLTSDAPEYKPWALVIQDSN
GENKIKML

【０１６７】

ND1-Left-TALE-1397C-UGI (図62: SEQ ID No: 416)

MALSRAVCGTSRQLAPVLGYLGSRQKHSLPDYPYDVPDYAGYPYDVPDYAGYPYDV
PDYAGIRIQDLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPA
ALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDT
GQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPDQVVAIASHDGGKQALETVQRL
LPVLCQAHGLTPDQVVAIASNGGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNIG
GKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQAHGLT
PDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQR
LLPVLCQAHGLTPAQVVAIASNGGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNI
GGKQALETVQRLLPVLCQAHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQAHGL
TPAQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQ
RLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPAQVVAIAS
NGGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNGGGKQALETVQRLLPVLCQAH
GLTPEQVVAIASNGGGKQALESIVAQLSRPDPALAALLVGSAIPVKRGATGETKVFTG
NSNSPKSPTKGGCSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILV
HTAYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKML

【０１６８】

ND1-Right-TALE-1397N-UGI (図62: SEQ ID No: 417)

MASVLTPLLLRGLTGSARRLPVPRAKIHSLDYKDHDGDYKDHDIDYKDDDDKGIRIQ
DLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVK
YQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIAKR

GGVTAVEAVHAWRNALTGAPLNLTPDQVVAIASNNGGKQALETVQRLLPVLCQDH
GLTPAQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETV
QRLLPVLCQAHGLTPDQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPAQVVAIA
SNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQA
HGLTPAQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNIGGKQALET
VQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPDQVVAI
ASNNGGKQALETVQRLLPVLCQAHGLTPAQVVAIASHDGGKQALETVQRLLPVLCQ
DHGLTPDQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPDQVVAIASHDGGKQAL
ETVQRLLPVLCQAHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPAQVV
AIASHDGGKQALETVQRLLPVLCQAHGLTPAQVVAIASHDGGKQALETVQRLLPVLC
QDHGLTPAQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQA
LESIVAQLSRPDPALAALLVGSGSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGL
ESKVFSSGGPTPYPNYANAGHVEGQSALFMRDNGISEGLVFHNNPEGTCGFCVNMT
ETLLPENAKMTVVPPEGSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPE
SDILVHTAYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKML

【０１６９】

ND1-Right-TALE-1397C-UGI（図62：SEQ ID No：418）

MASVLTPLLLRGLTGSARRLPVPRAKIHSLDYKDHDGDYKDHDIDYKDDDDKGIRIQ
DLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVK
YQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIAKR
GGVTAVEAVHAWRNALTGAPLNLTPDQVVAIASNNGGKQALETVQRLLPVLCQDH
GLTPAQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETV
QRLLPVLCQAHGLTPDQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPAQVVAIA
SNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQA
HGLTPAQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNIGGKQALET
VQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPDQVVAI
ASNNGGKQALETVQRLLPVLCQAHGLTPAQVVAIASHDGGKQALETVQRLLPVLCQ
DHGLTPDQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPDQVVAIASHDGGKQAL
ETVQRLLPVLCQAHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPAQVV
AIASHDGGKQALETVQRLLPVLCQAHGLTPAQVVAIASHDGGKQALETVQRLLPVLC
QDHGLTPAQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQA
LESIVAQLSRPDPALAALLVGSAIPVKRGATGETKVFTGNSNSPKSPTKGGCSGGSTN
LSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLTSDA
PEYKPWALVIQDSNGENKIKML

【０１７０】

ND4-Left-TALE-AD（図62b：SEQ ID No：419）

MALSRAVCGTSRQLAPVLGYLGSRQKHSLPDYPYDVPDYAGYPYDVPDYAGYPYDV
PDYAMDIADLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPA
ALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDT
GQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNNGGKQALETVQAL
LPVLCQAHGLTPQQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPQQVVAIASNG
GGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQALLPVLCQAHGLT
PEQVVAIASNNGGKQALETVQALLPVLCQAHGLTPEQVVAIASNGGGKQALETVQA
LLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNI
GGKQALETVQALLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGL
TPEQVVAIASHDGGKQALETVQALLPVLCQAHGLTPEQVVAIASNIGGKQALETVQA
LLPVLCQAHGLTPEQVVAIASHDGGKQALETVQALLPVLCQAHGLTPEQVVAIASN
NGGKQALETVQRLLPVLCQAHGLTPQQVVAIASNGGGKQALETVQRLLPVLCQAHG
LTPEQVVAIASNGGGKQALETVQALLPVLCQAHGLTPEQVVAIASHDGGKQALETV
QALLPVLCQAHGLTPQQVVAIASNGGGRPALESIVAQLSRPDPALAALTNDHLVALA

CLGGRPALDAVKKGLGGSSGSETPGTSESATPESSEVEFSHEYWMRHALTLAKRARD
EREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATL
YVTFEPCVMCAGAMIHSRIGRVVFGVRNSKRGAAGSLMNVLNYPGMNHRVEITEGI
LADECAALLCDFYRMPRQVFNAQKKAQSSIN

【０１７１】

ND4-Right-TALE-AD（図62b：SEQ ID No：420）

MASVLTPLLLRGLTGSARRLPVPRAKIHSLDYKDHDGDYKDHDIDYKDDDDKMDIA
DLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVK
YQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIAKR
GGVTAVEAVHAWRNALTGAPLNLTPEQVVAIASHDGGKQALETVQALLPVLCQAHG
LTPQQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPQQVVAIASNGGGKQALETV
QRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQALLPVLCQAHGLTPEQVVAIAS
NGGGKQALETVQALLPVLCQAHGLTPEQVVAIASNIGGKQALETVQALLPVLCQAH
GLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETV
QALLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIAS
NIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQALLPVLCQAH
GLTPEQVVAIASNNGGKQALETVQALLPVLCQAHGLTPEQVVAIASNIGGKQALETV
QALLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPQQVVAIA
SNIGGKQALETVQRLLPVLCQAHGLTPQQVVAIASNNGGKQALETVQRLLPVLCQA
HGLTPQQVVAIASNGGGRPALESIVAQLSRPDPALAALTNDHLVALACLGGRPALDA
VKKGLGGSSGSETPGTSESATPESSEVEFSHEYWMRHALTLAKRARDEREVPVGAVL
VLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMC
AGAMIHSRIGRVVFGVRNSKRGAAGSLMNVLNYPGMNHRVEITEGILADECAALLCD
FYRMPRQVFNAQKKAQSSIN

【０１７２】

ND1-Left-TALE-1333N-AD（図63：SEQ ID No：421）

MALSRAVCGTSRQLAPVLGYLGSRQKHSLPDYPYDVPDYAGYPYDVPDYAGYPYDV
PDYAGIRIQDLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPA
ALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDT
GQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPDQVVAIASHDGGKQALETVQRL
LPVLCQAHGLTPDQVVAIASNGGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNIG
GKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQAHGLT
PDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQR
LLPVLCQAHGLTPAQVVAIASNGGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNI
GGKQALETVQRLLPVLCQAHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQAHGL
TPAQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQ
RLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPAQVVAIAS
NGGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNGGGKQALETVQRLLPVLCQAH
GLTPEQVVAIASNGGGKQALESIVAQLSRPDPALAALLVGSGSYALGPYQISAPQLPA
YNGQTVGTFYYVNDAGGLESKVFSSGGSGSETPGTSESATPESSEVEFSHEYWMRHA
LTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQ
NYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNSKRGAAGSLMNVLNYPGMN
HRVEITEGILADECAALLCDFYRMPRQVFNAQKKAQSSIN

【０１７３】

ND1-Left-TALE-1333C-AD（図63：SEQ ID No：422）

MALSRAVCGTSRQLAPVLGYLGSRQKHSLPDYPYDVPDYAGYPYDVPDYAGYPYDV
PDYAGIRIQDLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPA
ALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDT
GQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPDQVVAIASHDGGKQALETVQRL
LPVLCQAHGLTPDQVVAIASNGGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNIG

GKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQAHGLT
PDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQR
LLPVLCQAHGLTPAQVVAIASNGGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNI
GGKQALETVQRLLPVLCQAHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQAHGL
TPAQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQ
RLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPAQVVAIAS
NGGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNGGGKQALETVQRLLPVLCQAH
GLTPEQVVAIASNGGGKQALESIVAQLSRPDPALAALLVGSPTPYPNYANAGHVEGQ
SALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKMTVVPPEGAIPVKRGA
TGETKVFTGNSNSPKSPTKGGCSGSETPGTSESATPESSEVEFSHEYWMRHALTLAK
RARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLI
DATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNSKRGAAGSLMNVLNYPGMNHRVEI
TEGILADECAALLCDFYRMPRQVFNAQKKAQSSIN

【０１７４】

ND1-Right-TALE-1333N-AD（図63：SEQ ID No：423）

MASVLTPLLLRGLTGSARRLPVPRAKIHSLDYKDHDGDYKDHDIDYKDDDDKGIRIQ
DLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVK
YQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIAKR
GGVTAVEAVHAWRNALTGAPLNLTPDQVVAIASNNGGKQALETVQRLLPVLCQDH
GLTPAQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETV
QRLLPVLCQAHGLTPDQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPAQVVAIA
SNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQA
HGLTPAQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNIGGKQALET
VQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPDQVVAI
ASNNGGKQALETVQRLLPVLCQAHGLTPAQVVAIASHDGGKQALETVQRLLPVLCQ
DHGLTPDQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPDQVVAIASHDGGKQAL
ETVQRLLPVLCQAHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPAQVV
AIASHDGGKQALETVQRLLPVLCQAHGLTPAQVVAIASHDGGKQALETVQRLLPVLC
QDHGLTPAQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQA
LESIVAQLSRPDPALAALLVGSGSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGL
ESKVFSSGGSGSETPGTSESATPESSEVEFSHEYWMRHALTLAKRARDEREVPVGAV
LVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVM
CAGAMIHSRIGRVVFGVRNSKRGAAGSLMNVLNYPGMNHRVEITEGILADECAALLC
DFYRMPRQVFNAQKKAQSSIN

【０１７５】

ND1-Right-TALE-1333C-AD（図63：SEQ ID No：424）

MASVLTPLLLRGLTGSARRLPVPRAKIHSLDYKDHDGDYKDHDIDYKDDDDKGIRIQ
DLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVK
YQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIAKR
GGVTAVEAVHAWRNALTGAPLNLTPDQVVAIASNNGGKQALETVQRLLPVLCQDH
GLTPAQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETV
QRLLPVLCQAHGLTPDQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPAQVVAIA
SNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQA
HGLTPAQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNIGGKQALET
VQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPDQVVAI
ASNNGGKQALETVQRLLPVLCQAHGLTPAQVVAIASHDGGKQALETVQRLLPVLCQ
DHGLTPDQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPDQVVAIASHDGGKQAL
ETVQRLLPVLCQAHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPAQVV
AIASHDGGKQALETVQRLLPVLCQAHGLTPAQVVAIASHDGGKQALETVQRLLPVLC
QDHGLTPAQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQA

LESIVAQLSRPDPALAALLVGSPTPYPNYANAGHVEGQSALFMRDNGISEGLVFHNN
PEGTCGFCVNMTETLLPENAKMTVVPPEGAIPVKRGATGETKVFTGNSNSPKSPTKG
GCSGSETPGTSESATPESSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRV
IGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIH
SRIGRVVFGVRNSKRGAAGSLMNVLNYPGMNHRVEITEGILADECAALLCDFYRMPR
QVFNAQKKAQSSIN

【０１７６】

ND1-Left-TALE-1397C-AD (図62-63: SEQ ID No: 425)

MALSRAVCGTSRQLAPVLGYLGSRQKHSLPDYPYDVPDYAGYPYDVPDYAGYPYDV
PDYAGIRIQDLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPA
ALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDT
GQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPDQVVAIASHDGGKQALETVQRL
LPVLCQAHGLTPDQVVAIASNGGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNIG
GKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQAHGLT
PDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQR
LLPVLCQAHGLTPAQVVAIASNGGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNI
GGKQALETVQRLLPVLCQAHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQAHGL
TPAQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQ
RLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPAQVVAIAS
NGGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNGGGKQALETVQRLLPVLCQAH
GLTPEQVVAIASNGGGKQALESIVAQLSRPDPALAALTNDHLVALACLGGRPALDAV
KKGLLVGSAIPVKRGATGETKVFTGNSNSPKSPTKGGCSGSETPGTSESATPESSEVE
FSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIM
ALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNSKRGAAGSL
MNVLNYPGMNHRVEITEGILADECAALLCDFYRMPRQVFNAQKKAQSSIN

【０１７７】

ND1-Right-TALE-1397C-AD (図62-63: SEQ ID No: 426)

MASVLTPLLLRGLTGSARRLPVPRAKIHSLDYKDHDGDYKDHDIDYKDDDDKGIRIQ
DLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVK
YQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIAKR
GGVTAVEAVHAWRNALTGAPLNLTPDQVVAIASNNGGKQALETVQRLLPVLCQDH
GLTPAQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETV
QRLLPVLCQAHGLTPDQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPAQVVAIA
SNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQA
HGLTPAQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNIGGKQALET
VQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPDQVVAI
ASNNGGKQALETVQRLLPVLCQAHGLTPAQVVAIASHDGGKQALETVQRLLPVLCQ
DHGLTPDQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPDQVVAIASHDGGKQAL
ETVQRLLPVLCQAHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPAQVV
AIASHDGGKQALETVQRLLPVLCQAHGLTPAQVVAIASHDGGKQALETVQRLLPVLC
QDHGLTPAQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQA
LESIVAQLSRPDPALAALTNDHLVALACLGGRPALDAVKKGLGGSGSAIPVKRGATG
ETKVFTGNSNSPKSPTKGGCSGSETPGTSESATPESSEVEFSHEYWMRHALTLAKRA
RDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDA
TLYVTFEPCVMCAGAMIHSRIGRVVFGVRNSKRGAAGSLMNVLNYPGMNHRVEITE
GILADECAALLCDFYRMPRQVFNAQKKAQSSIN

【０１７８】

ND1-Left-TALE-1333N (図63: SEQ ID No: 427)

MALSRAVCGTSRQLAPVLGYLGSRQKHSLPDYPYDVPDYAGYPYDVPDYAGYPYDV
PDYAGIRIQDLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPA

ALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDT
GQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPDQVVAIASHDGGKQALETVQRL
LPVLCQAHGLTPDQVVAIASNGGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNIG
GKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQAHGLT
PDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQR
LLPVLCQAHGLTPAQVVAIASNGGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNI
GGKQALETVQRLLPVLCQAHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQAHGL
TPAQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQ
RLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPAQVVAIAS
NGGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNGGGKQALETVQRLLPVLCQAH
GLTPEQVVAIASNGGGKQALESIVAQLSRPDPALAALTNDHLVALACLGGRPALDAV
KKGLGGSGSGSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGG

【0179】

ND1-Left-TALE-1333C (図63: SEQ ID No: 428)

MALSRAVCGTSRQLAPVLGYLGSRQKHSLPDYPYDVPDYAGYPYDVPDYAGYPYDV
PDYAGIRIQDLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPA
ALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDT
GQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPDQVVAIASHDGGKQALETVQRL
LPVLCQAHGLTPDQVVAIASNGGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNIG
GKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQAHGLT
PDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQR
LLPVLCQAHGLTPAQVVAIASNGGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNI
GGKQALETVQRLLPVLCQAHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQAHGL
TPAQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQ
RLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPAQVVAIAS
NGGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNGGGKQALETVQRLLPVLCQAH
GLTPEQVVAIASNGGGKQALESIVAQLSRPDPALAALTNDHLVALACLGGRPALDAV
KKGLGGSGSPTPYPNYANAGHVEGQSALFMRDNGISEGLVFHNNPEGTCGFCVNMT
ETLLPENAKMTVVPPEGAIPVKRGATGETKVFTGNSNSPKSPTKGGC

【0180】

ND1-Left-TALE-1397N (図62-63: SEQ ID No: 429)

MALSRAVCGTSRQLAPVLGYLGSRQKHSLPDYPYDVPDYAGYPYDVPDYAGYPYDV
PDYAGIRIQDLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPA
ALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDT
GQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPDQVVAIASHDGGKQALETVQRL
LPVLCQAHGLTPDQVVAIASNGGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNIG
GKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQAHGLT
PDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQR
LLPVLCQAHGLTPAQVVAIASNGGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNI
GGKQALETVQRLLPVLCQAHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQAHGL
TPAQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQ
RLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPAQVVAIAS
NGGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNGGGKQALETVQRLLPVLCQAH
GLTPEQVVAIASNGGGKQALESIVAQLSRPDPALAALTNDHLVALACLGGRPALDAV
KKGLGGSGSGSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGGPTPY
PNYANAGHVEGQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKMTVV
PPEG

【0181】

ND1-Right-TALE-1333N (図63: SEQ ID No: 430)

MASVLTPLLLRGLTGSARRLPVPRAKIHSLDYKDHDGDYKDHDIDYKDDDDKGIRIQ

DLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVK
YQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIAKR
GGVTAVEAVHAWRNALTGAPLNLTPDQVVAIASNNGGKQALETVQRLLPVLCQDH
GLTPAQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETV
QRLLPVLCQAHGLTPDQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPAQVVAIA
SNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQA
HGLTPAQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNIGGKQALET
VQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPDQVVAI
ASNNGGKQALETVQRLLPVLCQAHGLTPAQVVAIASHDGGKQALETVQRLLPVLCQ
DHGLTPDQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPDQVVAIASHDGGKQAL
ETVQRLLPVLCQAHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPAQVV
AIASHDGGKQALETVQRLLPVLCQAHGLTPAQVVAIASHDGGKQALETVQRLLPVLC
QDHGLTPAQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQA
LESIVAQLSRPDPALAALLVGSGSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGL
ESKVFSSGG

【０１８２】

ND1-Right-TALE-1333C（図63：SEQ ID No：431）

MASVLTPLLLRGLTGSARRLPVPRAKIHSLDYKDHDGDYKDHDIDYKDDDDKGIRIQ
DLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVK
YQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIAKR
GGVTAVEAVHAWRNALTGAPLNLTPDQVVAIASNNGGKQALETVQRLLPVLCQDH
GLTPAQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETV
QRLLPVLCQAHGLTPDQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPAQVVAIA
SNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQA
HGLTPAQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNIGGKQALET
VQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPDQVVAI
ASNNGGKQALETVQRLLPVLCQAHGLTPAQVVAIASHDGGKQALETVQRLLPVLCQ
DHGLTPDQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPDQVVAIASHDGGKQAL
ETVQRLLPVLCQAHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPAQVV
AIASHDGGKQALETVQRLLPVLCQAHGLTPAQVVAIASHDGGKQALETVQRLLPVLC
QDHGLTPAQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQA
LESIVAQLSRPDPALAALLVGSPTPYPNYANAGHVEGQSALFMRDNGISEGLVFHNN
PEGTCGFCVNMTETLLPENAKMTVVPPEGAIPVKRGATGETKVFTGNSNSPKSPTKG
GC

【０１８３】

ND1-Right-TALE-1397N（図62-63：SEQ ID No：432）

MASVLTPLLLRGLTGSARRLPVPRAKIHSLDYKDHDGDYKDHDIDYKDDDDKGIRIQ
DLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVK
YQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIAKR
GGVTAVEAVHAWRNALTGAPLNLTPDQVVAIASNNGGKQALETVQRLLPVLCQDH
GLTPAQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETV
QRLLPVLCQAHGLTPDQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPAQVVAIA
SNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQA
HGLTPAQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNIGGKQALET
VQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPDQVVAI
ASNNGGKQALETVQRLLPVLCQAHGLTPAQVVAIASHDGGKQALETVQRLLPVLCQ
DHGLTPDQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPDQVVAIASHDGGKQAL
ETVQRLLPVLCQAHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPAQVV
AIASHDGGKQALETVQRLLPVLCQAHGLTPAQVVAIASHDGGKQALETVQRLLPVLC
QDHGLTPAQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQA

LESIVAQLSRPDPALAALLVGSGSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGL
ESKVFSSGGPTPYPNYANAGHVEGQSALFMRDNGISEGLVFHNNPEGTCGFCVNMT
ETLLPENAKMTVVPPEG

【０１８４】

ND4-LEFT-TALE-1333N-AD (図63: SEQ ID No: 433)

MALSRAVCGTSRQLAPVLGYLGSRQKHSLPDYPYDVPDYAGYPYDVPDYAGYPYDV
PDYAMDIADLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPA
ALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDT
GQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNNGGKQALETVQAL
LPVLCQAHGLTPQQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPQQVVAIASNG
GGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQALLPVLCQAHGLT
PEQVVAIASNNGGKQALETVQALLPVLCQAHGLTPEQVVAIASNGGGKQALETVQA
LLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNI
GGKQALETVQALLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGL
TPEQVVAIASHDGGKQALETVQALLPVLCQAHGLTPEQVVAIASNIGGKQALETVQA
LLPVLCQAHGLTPEQVVAIASHDGGKQALETVQALLPVLCQAHGLTPEQVVAIASN
NGGKQALETVQRLLPVLCQAHGLTPQQVVAIASNGGGKQALETVQRLLPVLCQAHG
LTPEQVVAIASNGGGKQALETVQALLPVLCQAHGLTPEQVVAIASHDGGKQALETV
QALLPVLCQAHGLTPQQVVAIASNGGGRPALESIVAQLSRPDPALAALTNDHLVALA
CLGGRPALDAVKKGLGGSGSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKV
FSSGGSGSETPGTSESATPESSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLN
NRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGA
MIHSRIGRVVFGVRNSKRGAAGSLMNVLNYPGMNHRVEITEGILADECAALLCDFYR
MPRQVFNAQKKAQSSIN

【０１８５】

ND4-LEFT-TALE-1333C-AD (図63: SEQ ID No: 434)

MALSRAVCGTSRQLAPVLGYLGSRQKHSLPDYPYDVPDYAGYPYDVPDYAGYPYDV
PDYAMDIADLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPA
ALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDT
GQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNNGGKQALETVQAL
LPVLCQAHGLTPQQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPQQVVAIASNG
GGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQALLPVLCQAHGLT
PEQVVAIASNNGGKQALETVQALLPVLCQAHGLTPEQVVAIASNGGGKQALETVQA
LLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNI
GGKQALETVQALLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGL
TPEQVVAIASHDGGKQALETVQALLPVLCQAHGLTPEQVVAIASNIGGKQALETVQA
LLPVLCQAHGLTPEQVVAIASHDGGKQALETVQALLPVLCQAHGLTPEQVVAIASN
NGGKQALETVQRLLPVLCQAHGLTPQQVVAIASNGGGKQALETVQRLLPVLCQAHG
LTPEQVVAIASNGGGKQALETVQALLPVLCQAHGLTPEQVVAIASHDGGKQALETV
QALLPVLCQAHGLTPQQVVAIASNGGGRPALESIVAQLSRPDPALAALTNDHLVALA
CLGGRPALDAVKKGLGGSPTPYPNYANAGHVEGQSALFMRDNGISEGLVFHNNPEG
TCGFCVNMTETLLPENAKMTVVPPEGAIPVKRGATGETKVFTGNSNSPKSPTKGGCS
GSETPGTSESATPESSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGE
GWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRI
GRVVFGVRNSKRGAAGSLMNVLNYPGMNHRVEITEGILADECAALLCDFYRMPRQV
FNAQKKAQSSIN

【０１８６】

ND4-LEFT-TALE-1397C-AD (図63: SEQ ID No: 435)

MALSRAVCGTSRQLAPVLGYLGSRQKHSLPDYPYDVPDYAGYPYDVPDYAGYPYDV
PDYAMDIADLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPA

ALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDT
GQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNNGGKQALETVQAL
LPVLCQAHGLTPQQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPQQVVAIASNG
GGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQALLPVLCQAHGLT
PEQVVAIASNNGGKQALETVQALLPVLCQAHGLTPEQVVAIASNGGGKQALETVQA
LLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNI
GGKQALETVQALLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGL
TPEQVVAIASHDGGKQALETVQALLPVLCQAHGLTPEQVVAIASNIGGKQALETVQA
LLPVLCQAHGLTPEQVVAIASHDGGKQALETVQALLPVLCQAHGLTPEQVVAIASN
NGGKQALETVQRLLPVLCQAHGLTPQQVVAIASNGGGKQALETVQRLLPVLCQAHG
LTPEQVVAIASNGGGKQALETVQALLPVLCQAHGLTPEQVVAIASHDGGKQALETV
QALLPVLCQAHGLTPQQVVAIASNGGGRPALESIVAQLSRPDPALAALTNDHLVALA
CLGGRPALDAVKKGLGGSGSAIPVKRGATGETKVFTGNSNSPKSPTKGGCSGSETPG
TSESATPESSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRA
IGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVF
GVRNSKRGAAGSLMNVLNYPGMNHRVEITEGILADECAALLCDFYRMPRQVFNAQK
KAQSSIN

【０１８７】

ND4-Right-TALE-1333N-AD（図63：SEQ ID No：436）

MASVLTPLLLRGLTGSARRLPVPRAKIHSLDYKDHDGDYKDHDIDYKDDDDKMDIA
DLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVK
YQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIAKR
GGVTAVEAVHAWRNALTGAPLNLTPEQVVAIASHDGGKQALETVQALLPVLCQAHG
LTPQQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPQQVVAIASNGGGKQALETV
QRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQALLPVLCQAHGLTPEQVVAIAS
NGGGKQALETVQALLPVLCQAHGLTPEQVVAIASNIGGKQALETVQALLPVLCQAH
GLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETV
QALLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIAS
NIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQALLPVLCQAH
GLTPEQVVAIASNNGGKQALETVQALLPVLCQAHGLTPEQVVAIASNIGGKQALETV
QALLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPQQVVAIA
SNIGGKQALETVQRLLPVLCQAHGLTPQQVVAIASNNGGKQALETVQRLLPVLCQA
HGLTPQQVVAIASNGGGRPALESIVAQLSRPDPALAALTNDHLVALACLGGRPALDA
VKKGLGGSGSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGGSGSET
PGTSESATPESSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWN
RAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRV
VFGVRNSKRGAAGSLMNVLNYPGMNHRVEITEGILADECAALLCDFYRMPRQVFNA
QKKAQSSIN

【０１８８】

ND4-Right-TALE-1333C-AD（図63：SEQ ID No：437）

MASVLTPLLLRGLTGSARRLPVPRAKIHSLDYKDHDGDYKDHDIDYKDDDDKMDIA
DLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVK
YQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIAKR
GGVTAVEAVHAWRNALTGAPLNLTPEQVVAIASHDGGKQALETVQALLPVLCQAHG
LTPQQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPQQVVAIASNGGGKQALETV
QRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQALLPVLCQAHGLTPEQVVAIAS
NGGGKQALETVQALLPVLCQAHGLTPEQVVAIASNIGGKQALETVQALLPVLCQAH
GLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETV
QALLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIAS
NIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQALLPVLCQAH
GLTPEQVVAIASNNGGKQALETVQALLPVLCQAHGLTPEQVVAIASNIGGKQALETV
QALLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPQQVVAIA
SNIGGKQALETVQRLLPVLCQAHGLTPQQVVAIASNNGGKQALETVQRLLPVLCQA
HGLTPQQVVAIASNGGGRPALESIVAQLSRPDPALAALTNDHLVALACLGGRPALDA
VKKGLGGSGSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGGSGSET
PGTSESATPESSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWN
RAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRV
VFGVRNSKRGAAGSLMNVLNYPGMNHRVEITEGILADECAALLCDFYRMPRQVFNA
QKKAQSSIN

GLTPEQVVAIASNNGGKQALETVQALLPVLCQAHGLTPEQVVAIASNIGGKQALETV
QALLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPQQVVAIA
SNIGGKQALETVQRLLPVLCQAHGLTPQQVVAIASNNGGKQALETVQRLLPVLCQA
HGLTPQQVVAIASNGGGRPALESIVAQLSRPDPALAALTNDHLVALACLGGRPALDA
VKKGLGGSPTPYPNYANAGHVEGQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTE
TLLPENAKMTVVPPEGAIPVKRGATGETKVFTGNSNSPKSPTKGGCSGSETPGTSES
ATPESSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLH
DPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRN
SKRGAAGSLMNVLNYPGMNHRVEITEGILADECAALLCDFYRMPRQVFNAQKKAQS
SIN

【０１８９】

ND4-Right-TALE-1397C-AD (図63: SEQ ID No: 438)

MASVLTPLLLRGLTGSARRLPVPRAKIHSLDYKDHDGDYKDHDIDYKDDDDKMDIA
DLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVK
YQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIAKR
GGVTAVEAVHAWRNALTGAPLNLTPEQVVAIASHDGGKQALETVQALLPVLCQAHG
LTPQQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPQQVVAIASNGGGKQALETV
QRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQALLPVLCQAHGLTPEQVVAIAS
NGGGKQALETVQALLPVLCQAHGLTPEQVVAIASNIGGKQALETVQALLPVLCQAH
GLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETV
QALLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIAS
NIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQALLPVLCQAH
GLTPEQVVAIASNNGGKQALETVQALLPVLCQAHGLTPEQVVAIASNIGGKQALETV
QALLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPQQVVAIA
SNIGGKQALETVQRLLPVLCQAHGLTPQQVVAIASNNGGKQALETVQRLLPVLCQA
HGLTPQQVVAIASNGGGRPALESIVAQLSRPDPALAALTNDHLVALACLGGRPALDA
VKKGLGGSGSAIPVKRGATGETKVFTGNSNSPKSPTKGGCSGSETPGTSESATPESSE
VEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAE
IMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNSKRGAAG
SLMNVLNYPGMNHRVEITEGILADECAALLCDFYRMPRQVFNAQKKAQSSIN

【０１９０】

ND1-Left-TALE-AD-GSVG (図64: SEQ ID No: 439)

MALSRAVCGTSRQLAPVLGYLGSRQKHSLPDYPYDVPDYAGYPYDVPDYAGYPYDV
PDYAGIRIQDLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPA
ALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDT
GQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPDQVVAIASHDGGKQALETVQRL
LPVLCQAHGLTPDQVVAIASNGGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNIG
GKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQAHGLT
PDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQR
LLPVLCQAHGLTPAQVVAIASNGGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNI
GGKQALETVQRLLPVLCQAHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQAHGL
TPAQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQ
RLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPAQVVAIAS
NGGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNGGGKQALETVQRLLPVLCQAH
GLTPEQVVAIASNGGGKQALESIVAQLSRPDPALAALLVSGSETPGTSESATPESSEV
EFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEI
MALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNSKRGAAGS
LMNVLNYPGMNHRVEITEGILADECAALLCDFYRMPRQVFNAQKKAQSSINLVGSG
SYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLEGKVFSSGGPTPYPNYANAGHV
ESQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKMTVVPPEGVIPVK

RGATGETKVFTGNSNGPKSPTKGGC

【０１９１】

ND1-Left-TALE-AD-E1347A（図64：SEQ ID No：440）

MALSRAVCGTSRQLAPVLGYLGSRQKHSLPDYPYDVPDYAGYPYDVPDYAGYPYDV
PDYAGIRIQDLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPA
ALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDT
GQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPDQVVAIASHDGGKQALETVQRL
LPVLCQAHGLTPDQVVAIASNGGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNIG
GKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQAHGLT
PDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQR
LLPVLCQAHGLTPAQVVAIASNGGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNI
GGKQALETVQRLLPVLCQAHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQAHGL
TPAQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQ
RLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPAQVVAIAS
NGGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNGGGKQALETVQRLLPVLCQAH
GLTPEQVVAIASNGGGKQALESIVAQLSRPDPALAALLVSGSETPGTSESATPESSEV
EFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEI
MALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNSKRGAAGS
LMNVLNYPGMNHRVEITEGILADECAALLCDFYRMPRQVFNAQKKAQSSINLVGSG
SYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGGPTPYPNYANAGHVA
GQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKMTVVPPEGAIPVKR
GATGETKVFTGNSNSPKSPTKGGC

【０１９２】

ND1-Left-TALE-AD-AAAAA（図64：SEQ ID No：441）

MALSRAVCGTSRQLAPVLGYLGSRQKHSLPDYPYDVPDYAGYPYDVPDYAGYPYDV
PDYAGIRIQDLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPA
ALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDT
GQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPDQVVAIASHDGGKQALETVQRL
LPVLCQAHGLTPDQVVAIASNGGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNIG
GKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQAHGLT
PDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQR
LLPVLCQAHGLTPAQVVAIASNGGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNI
GGKQALETVQRLLPVLCQAHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQAHGL
TPAQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQ
RLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPAQVVAIAS
NGGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNGGGKQALETVQRLLPVLCQAH
GLTPEQVVAIASNGGGKQALESIVAQLSRPDPALAALLVSGSETPGTSESATPESSEV
EFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEI
MALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNSKRGAAGS
LMNVLNYPGMNHRVEITEGILADECAALLCDFYRMPRQVFNAQKKAQSSINLVGSG
SYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGGPTPYPNYANAGHVE
GQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKMTVVPPEGAIPVAA
GATGETAVFTGNSNSPASPTAGGC

【０１９３】

ND1-Right-TALE-AD（図64：SEQ ID No：442）

MASVLTPLLLRGLTGSARRLPVPRAKIHSLDYKDHDGDYKDHDIDYKDDDDKGIRIQ
DLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVK
YQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIAKR
GGVTAVEAVHAWRNALTGAPLNLTPDQVVAIASNNGGKQALETVQRLLPVLCQDH
GLTPAQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETV

QRLLPVLCQAHGLTPDQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPAQVVAIA
SNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQA
HGLTPAQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNIGGKQALET
VQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPDQVVAI
ASNNGGKQALETVQRLLPVLCQAHGLTPAQVVAIASHDGGKQALETVQRLLPVLCQ
DHGLTPDQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPDQVVAIASHDGGKQAL
ETVQRLLPVLCQAHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPAQVV
AIASHDGGKQALETVQRLLPVLCQAHGLTPAQVVAIASHDGGKQALETVQRLLPVLC
QDHGLTPAQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQA
LESIVAQLSRPDPALAALLVSGSETPGTSESATPESSEVEFSHEYWMRHALTLAKRAR
DEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDAT
LYVTFEPCVMCAGAMIHSRIGRVVFGVRNSKRGAAGSLMNVLNYPGMNHRVEITEG
ILADECAALLCDFYRMPRQVFNAQKKAQSSIN

【0194】

ND1-Left-TALE-GSVG (図64: SEQ ID No: 443)

MALSRAVCGTSRQLAPVLGYLGSRQKHSLPDYPYDVPDYAGYPYDVPDYAGYPYDV
PDYAGIRIQDLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPA
ALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDT
GQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPDQVVAIASHDGGKQALETVQRL
LPVLCQAHGLTPDQVVAIASNGGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNIG
GKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQAHGLT
PDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQR
LLPVLCQAHGLTPAQVVAIASNGGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNI
GGKQALETVQRLLPVLCQAHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQAHGL
TPAQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQ
RLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPAQVVAIAS
NGGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNGGGKQALETVQRLLPVLCQAH
GLTPEQVVAIASNGGGKQALESIVAQLSRPDPALAALLVGSGSYALGPYQISAPQLPA
YNGQTVGTFYYVNDAGGLEGKVFSSGGPTPYPNYANAGHVESQSALFMRDNGISEG
LVFHNNPEGTCGFCVNMTETLLPENAKMTVVPPEGVIPVKRGATGETKVFTGNSNG
PKSPTKGGC

【0195】

ND1-Left-TALE-E1347A (図64: SEQ ID No: 444)

MALSRAVCGTSRQLAPVLGYLGSRQKHSLPDYPYDVPDYAGYPYDVPDYAGYPYDV
PDYAGIRIQDLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPA
ALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDT
GQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPDQVVAIASHDGGKQALETVQRL
LPVLCQAHGLTPDQVVAIASNGGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNIG
GKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQAHGLT
PDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQR
LLPVLCQAHGLTPAQVVAIASNGGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNI
GGKQALETVQRLLPVLCQAHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQAHGL
TPAQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQ
RLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPAQVVAIAS
NGGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNGGGKQALETVQRLLPVLCQAH
GLTPEQVVAIASNGGGKQALESIVAQLSRPDPALAALLVGSGSYALGPYQISAPQLPA
YNGQTVGTFYYVNDAGGLESKVFSSGGPTPYPNYANAGHVAGQSALFMRDNGISEG
LVFHNNPEGTCGFCVNMTETLLPENAKMTVVPPEGAIPVKRGATGETKVFTGNSNS
PKSPTKGGC

【0196】

ND1-Left-TALE-AAAAA (図64 SEQ ID No: 445)

MALSRAVCGTSRQLAPVLGYLGSRQKHSLPDYPYDVPDYAGYPYDVPDYAGYPYDV
PDYAGIRIQDLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPA
ALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDT
GQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPDQVVAIASHDGGKQALETVQRL
LPVLCQAHGLTPDQVVAIASNGGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNIG
GKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQAHGLT
PDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQR
LLPVLCQAHGLTPAQVVAIASNGGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNI
GGKQALETVQRLLPVLCQAHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQAHGL
TPAQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQ
RLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPAQVVAIAS
NGGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNGGGKQALETVQRLLPVLCQAH
GLTPEQVVAIASNGGGKQALESIVAQLSRPDPALAALLVGSGSYALGPYQISAPQLPA
YNGQTVGTFYYVNDAGGLESKVFSSGGPTPYPNYANAGHVEGQSALFMRDNGISEG
LVFHNNPEGTCGFCVNMTETLLPENAKMTVVPPEGAIPVAAGATGETAVFTGNSNS
PASPTAGGC

【0197】

ND1-TALE-left (SEQ ID No: 446)

DLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVK
YQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIAKR
GGVTAVEAVHAWRNALTGAPLNLTPDQVVAIASHDGGKQALETVQRLLPVLCQAH
GLTPDQVVAIASNGGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNIGGKQALETV
QRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPDQVVAIA
SHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQA
HGLTPAQVVAIASNGGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNIGGKQALET
VQRLLPVLCQAHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPAQVVAI
ASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQ
AHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNGGGKQAL
ETVQRLLPVLCQAHGLTPAQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVV
AIASNGGGKQALESIVAQLSRPDPALAAL

【0198】

ND1-TALE-Right (SEQ ID No: 447)

DLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVK
YQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIAKR
GGVTAVEAVHAWRNALTGAPLNLTPDQVVAIASNNGGKQALETVQRLLPVLCQDH
GLTPAQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETV
QRLLPVLCQAHGLTPDQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPAQVVAIA
SNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQA
HGLTPAQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNIGGKQALET
VQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPDQVVAI
ASNNGGKQALETVQRLLPVLCQAHGLTPAQVVAIASHDGGKQALETVQRLLPVLCQ
DHGLTPDQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPDQVVAIASHDGGKQAL
ETVQRLLPVLCQAHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPAQVV
AIASHDGGKQALETVQRLLPVLCQAHGLTPAQVVAIASHDGGKQALETVQRLLPVLC
QDHGLTPAQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQA
LESIVAQLSRPDPALAAL

【0199】

ND4-TALE-left (SEQ ID No: 448)

MDIADLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGT

VAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLL
KIAKRGGVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNNGGKQALETVQALLPVL
CQAHGLTPQQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPQQVVAIASNGGGKQ
ALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQALLPVLCQAHGLTPEQV
VAIASNNGGKQALETVQALLPVLCQAHGLTPEQVVAIASNGGGKQALETVQALLPVL
CQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQA
LETVQALLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQV
VAIASHDGGKQALETVQALLPVLCQAHGLTPEQVVAIASNIGGKQALETVQALLPVL
CQAHGLTPEQVVAIASHDGGKQALETVQALLPVLCQAHGLTPEQVVAIASNNGGKQ
ALETVQRLLPVLCQAHGLTPQQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQ
VVAIASNGGGKQALETVQALLPVLCQAHGLTPEQVVAIASHDGGKQALETVQALLP
VLCQAHGLTPQQVVAIASNGGGRPALESIVAQLSRPDPALAALTNDHLVALACLGGR
PALDAVKKGLG

【0200】

ND4-TALE-Right (SEQ ID No: 449)

MDIADLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGT
VAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLL
KIAKRGGVTAVEAVHAWRNALTGAPLNLTPEQVVAIASHDGGKQALETVQALLPVL
CQAHGLTPQQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPQQVVAIASNGGGKQ
ALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQALLPVLCQAHGLTPEQ
VVAIASNGGGKQALETVQALLPVLCQAHGLTPEQVVAIASNIGGKQALETVQALLPV
LCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQ
ALETVQALLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQ
VVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQALLPV
LCQAHGLTPEQVVAIASNNGGKQALETVQALLPVLCQAHGLTPEQVVAIASNIGGKQ
ALETVQALLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPQQ
VVAIASNIGGKQALETVQRLLPVLCQAHGLTPQQVVAIASNNGGKQALETVQRLLPV
LCQAHGLTPQQVVAIASNGGGRPALESIVAQLSRPDPALAALTNDHLVALACLGGRP
ALDAVKKGLG

【0201】

TRAC-ZFP-LEFT (SEQ ID No: 450)

FQCRICMRKFATSGSLTRHTKIHTGEKPFQCRICMRNFSRSDHLSTHIRTHTGEKPFA
CDICGRKFATSSNRTKHTKIHTHPRAPIPKPFQCRICMRNFSRSDNLSEHIRTHTGEK
PFACDICGRKFAWHSSLRVHTKIHLR

【0202】

TRAC-ZFP-Right (SEQ ID No: 451)

FQCRICMRNFSRSDHLSTHIRTHTGEKPFACDICGRKFADRSHLARHTKIHTGSQKPF
QCRICMRKFALKQHLNEHTKIHTGEKPFQCRICMRNFSQSGNLARHIRTHTGEKPFA
CDICGRKFAHNSSLKDHTKIHLR

【0203】

ND1-ZFP-Left (SEQ ID No: 452)

FQCRICMRNFSDSGNLRVHIRTHTGEKPYKCPDCGKSFSQSSSLIRHQRTHTGEKPY
ECDHCGKSFSQSSHLNVHKRTHTGEKPYRCKYCDRSFSISSNLQRHVRNIHLR

【0204】

ND1-ZFP-Right (SEQ ID No: 453)

YKCPECGKSFSTKNSLTEHQRTHTGEKPYKCPECGKSFSSKKALTEHQRTHTGEKPY
ECNYCGKTFSVSSTLIRHQRIHTGEKPYRCKYCDRSFSISSNLQRHVRNIHLR

【0205】

G1333N-DddAtox (SEQ ID No: 454)

GSGSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGG

【０２０６】
G1333C-DddAtox (SEQ ID No: 455)
GSPTPYPNYANAGHVEGQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPEN
AKMTVVPPEGAIPVKRGATGETKVFTGNSNSPKSPTKGGC
【０２０７】
G1397N-DddAtox (SEQ ID No: 456)
GSGSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGGPTPYPNYANAG
HVEGQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKMTVVPPEG
【０２０８】
G1397C-DddAtox (SEQ ID No: 457)
GSAIPVKRGATGETKVFTGNSNSPKSPTKGGC
【０２０９】
Adenine deaminase (AD: ABE 8e) (SEQ ID No: 458)
SEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAH
AEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNSKRGA
AGSLMNVLNYPGMNHRVEITEGILADECAALLCDFYRMPRQVFNAQKKAQSSIN
【０２１０】
Full length DddAtox variant GSVG (SEQ ID No: 459)
GSGSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLEGKVFSSGGPTPYPNYANAG
HVESQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKMTVVPPEGVIP
VKRGATGETKVFTGNSNGPKSPTKGGC
【０２１１】
Full length DddAtox variant E1347A (SEQ ID No: 460)
GSGSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGGPTPYPNYANAG
HVAGQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKMTVVPPEGAIP
VKRGATGETKVFTGNSNSPKSPTKGGC
【０２１２】
Full length DddAtox variant AAAAA (SEQ ID No: 461)
GSGSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGGPTPYPNYANAG
HVEGQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKMTVVPPEGAIP
VAAGATGETAVFTGNSNSPASPTAGGC
【０２１３】
UGI (SEQ ID No: 462)
TNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLTS
DAPEYKPWALVIQDSNGENKIKML
【０２１４】
Single module
(TALE)-linker-AD-GSVG (SEQ ID No: 463)
SGSETPGTSESATPESSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIG
EGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSR
IGRVVFGVRNSKRGAAGSLMNVLNYPGMNHRVEITEGILADECAALLCDFYRMPRQ
VFNAQKKAQSSINLVGSGSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLEGKVF
SSGGPTPYPNYANAGHVESQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLP
ENAKMTVVPPEGVIPVKRGATGETKVFTGNSNGPKSPTKGGC
【０２１５】
(TALE)-linker-AD-E1347A (SEQ ID No: 464)
SGSETPGTSESATPESSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIG
EGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSR
IGRVVFGVRNSKRGAAGSLMNVLNYPGMNHRVEITEGILADECAALLCDFYRMPRQ
VFNAQKKAQSSINLVGSGSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVF

SSGGPTPYPNYANAGHVAGQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLP
ENAKMTVVPPEGAIPVKRGATGETKVFTGNSNSPKSPTKGGC

【０２１６】

(TALE)-linker-AD-AAAAA (SEQ ID No: 465)

SGSETPGTSESATPESSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIG
EGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSR
IGRVVFGVRNSKRGAAGSLMNVLNYPGMNHRVEITEGILADECAALLCDFYRMPRQ
VFNAQKKAQSSINLVGSGSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVF
SSGGPTPYPNYANAGHVEGQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLP
ENAKMTVVPPEGAIPVAAGATGETAVFTGNSNSPASPTAGGC

【０２１７】

Dual module

(TALE) -linker-AD (SEQ ID No: 466)

SGSETPGTSESATPESSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIG
EGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSR
IGRVVFGVRNSKRGAAGSLMNVLNYPGMNHRVEITEGILADECAALLCDFYRMPRQ
VFNAQKKAQSSIN

【０２１８】

(TALE)-GSVG (SEQ ID No: 467)

GSGSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLEGKVFSSGGPTPYPNYANAG
HVESQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKMTVVPPEGVIP
VKRGATGETKVFTGNSNGPKSPTKGGC

【０２１９】

(TALE)-E1347A (SEQ ID No: 468)

GSGSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGGPTPYPNYANAG
HVAGQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKMTVVPPEGAIP
VKRGATGETKVFTGNSNSPKSPTKGGC

【０２２０】

(TALE)-AAAAA (SEQ ID No: 469)

GSGSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGGPTPYPNYANAG
HVEGQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKMTVVPPEGAIP
VAAGATGETAVFTGNSNSPASPTAGGC

【０２２１】

(TALE)-1397C-AD (SEQ ID No: 470)

GSAIPVKRGATGETKVFTGNSNSPKSPTKGGCSGSETPGTSESATPESSEVEFSHEYW
MRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGG
LVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNSKRGAAGSLMNVLNY
PGMNHRVEITEGILADECAALLCDFYRMPRQVFNAQKKAQSSIN

【０２２２】

(TALE)-1333N-AD (SEQ ID No: 471)

GSGSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGGSGSETPGTSESA
TPESSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHD
PTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNS
KRGAAGSLMNVLNYPGMNHRVEITEGILADECAALLCDFYRMPRQVFNAQKKAQSS
IN

【０２２３】

(TALE)-1333C-AD (SEQ ID No: 472)

GSPTPYPNYANAGHVEGQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPEN
AKMTVVPPEGAIPVKRGATGETKVFTGNSNSPKSPTKGGCSGSETPGTSESATPESSE
VEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAE

IMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNSKRGAAG
SLMNVLNYPGMNHRVEITEGILADECAALLCDFYRMPRQVFNAQKKAQSSIN
【０２２４】
Linker(AAはアミノ酸を意味する)
8AA: SGGGLGST (SEQ ID No: 473)
16AA: SGSETPGTSESATPES (SEQ ID No: 474)
32AA: SGGSSGGSSGSETPGTSESATPESSGGSSGGS (SEQ ID No: 475)
【０２２５】
SOD2 MTS-3xHA: MALSRAVCGTSRQLAPVLGYLGSRQKHSLPDYPYDVPDYAG
YPYDVPDYAGYPYDVPDYA (SEQ ID No: 476)
COX8A MTS-3xFLAG:
MASVLTPLLLRGLTGSARRLPVPRAKIHSLDYKDHDGDYKDHDIDYKDDDDK (SE
Q ID No: 477)
【０２２６】
５．伝達
本発明による融合タンパク質は、例えば、微小注入法(microinjection)、電気穿孔法(electroporation)、DEAE－デキストラン処理、リポフェクション(lipofection)、ナノ粒子－媒介形質注入、タンパク質伝達ドメイン媒介導入、およびPEG－媒介形質注入などのような当業界における様々な方法によって細胞に伝達されることができるが、これに限定されない。
【０２２７】
本発明は、他の観点から、前記融合タンパク質をコードする核酸に関する。
前記核酸に関して、「ポリヌクレオチド」、「ヌクレオチド」、「ヌクレオチド配列」、および「オリゴヌクレオチド」は相互交換的に使用される。任意の長さのヌクレオチドのポリマー形態、デオキシリボヌクレオチドまたはリボヌクレオチド、またはそれらの類似体を含んでもよい。ポリヌクレオチドは、任意の三次元構造を有することができ、公知または非公知の任意の機能を行うことができる。ポリヌクレオチドは、1つ以上のヌクレオチド、例えば、メチル化されたヌクレオチド及びヌクレオチド類似体を含むことができる。ヌクレオチド構造に対する変形は、ポリマーのアセンブリの前または後に可能である。
【０２２８】
前記ポリヌクレオチドは、RNA配列、DNA配列、またはそれらの組合(RNA-DNA組合配列)であってもよい。
【０２２９】
前記融合タンパク質を発現させるための手段として、プラスミドベクター、コスミドベクター、バクテリオファージベクターなど公知の発現ベクターを用いることができ、ベクターは、DNA組換え技術を用いた任意の公知の方法に従って当業者が容易に製造することができる。
【０２３０】
前記ベクターは、プラスミドベクターまたはウイルスベクターであってもよく、前記ウイルスベクターは、具体的にアデノウイルス(Adenovirus)、アデノ付属ウイルス(Adeno-associated virus)、レンチウイルス(Lentivirus)、レトロウイルス(Retrovirus)ベクターであってもよいが、これに限定されるものではない。
【０２３１】
組換え発現ベクターは、宿主細胞における核酸の発現に適した形態で核酸を含むことができ、これは、組換え発現ベクターが発現のために使用されるように、つまり、発現される核酸配列に機能可能に連結されるように、宿主細胞を基盤にして選択できる1つ以上の調節要素を含むことを意味する。
【０２３２】
組換え発現ベクター内で、「機能可能に連結された」は、関心のヌクレオチド配列がヌ

クレオチド配列の発現を許容する方式で(例えば、インビトロ転写/翻訳システムで、またはベクターが宿主細胞に導入される場合は宿主細胞で)調節整要素に接続されることを意味する。

【０２３３】

組換え発現ベクターは、T7プロモーターを含めてメッセンジャーRNA合成に適した形態を含むことができ、これは、インビトでmRNA合成をできるように、すなわち、T7ポリメラーゼによってメッセンジャーRNAが合成できる1つ以上の調節要素を含むことを意味する。

【０２３４】

「調節要素」は、プロモーター、エンハンサー、内部リボソーム進入部位(IRES)、およびその他の発現調整要素(例えば、転写終結信号、例えば、ポリアデニル化信号およびポリ－U配列)を含んでもよい。調節要素は、多くの類型の宿主細胞においてヌクレオチド配列の誘導または常に発現を指示する要素、および特定の宿主細胞においてのみヌクレオチド配列の発現を指示する要素(例えば、組織－特異的調節配列)を含む。組織－特異的プロモーターは、所望の関心組織、例えば、筋肉、ニューロン、骨、皮膚、血液、特定の器官(例えば、肝臓、膵臓)、または特定の細胞型(例えば、リンパ球)で主に発現を指示することができる。調節要素は、また、組織または細胞－類型特異的であり得るかあり得ない細胞－周期依存性または発達段階－依存的方式のような一時的－依存的方式で発現を指示することができる。

【０２３５】

場合によっては、ベクターは、1つ以上のpol IIIプロモーター、1つ以上のpol IIプロモーター、1つ以上のpol Iプロモーター、またはそれらの組合を含む。pol IIIプロモーターの例には、非限定的にU6およびH1プロモーターが含まれる。pol IIプロモーターの例は、非限定的に、レトロウイルスラウス肉腫ウイルス(RSV)LTRプロモーター(任意にRSVエンハンサーを有する)、サイトメガロウイルス(CMV)プロモーター(選択的にCMVエンハンサーを有する)(例えば、Boshart et. al(1985)Cell 41:521-530)、SV40プロモーター、ジヒドロ葉酸還元酵素プロモーター、β－アクチンプロモーター、ホスホグリセリン酸キナーゼ(PGK)プロモーター、およびEF1αプロモーターを含む。

【０２３６】

「調節要素」には、エンハンサー、例えば、WPRE；CMVエンハンサー；HTLV-IのLTRのR-U5’セグメント；SV40エンハンサー；及びウサギβ?グロビンのエクソン2と3との間のイントロン配列などが含まれ得る。発現ベクターのデザインは、形質転換される宿主細胞の選択、所望の発現水準などのような要因に左右され得ることが当業者によって認識される。ベクターは、宿主細胞内に導入され、本明細書に記載の核酸によってエンコーディングされる融合タンパク質またはペプチドを含む転写物、タンパク質またはペプチドを生成することができる(例えば、クラスタリングされた規則的な間隔の短い回文繰り返し部(CRISPR)転写物、タンパク質、酵素、その突然変異体、その融合タンパク質など)。有益なベクターは、レンチウイルスおよびアデノ－関連ウイルスを含み、そのようなベクターの類型は、また、特定の種類の細胞を標的化するために選択することができる。

【０２３７】

前記ベクターは、それぞれ局所注入法(例えば、病変または標的部位直接注入)、電気穿孔法(electroporation)、リポフェクション、ウイルスベクター、ナノ粒子(nanoparticles)だけでなく、PTD(Protein translocation domain)融合タンパク質法などによりインビボまたは細胞内に伝達され得る。

【０２３８】

前記核酸は、リボ核酸、例えば、メッセンジャーリボ核酸mRNAの形態で注入して、細胞の遺伝子塩基編集、例えば、動物細胞または植物細胞に制限なく遺伝子塩基編集が可能である。

【０２３９】

本発明による核酸は、mRNA形態であってもよく、mRNA形態で伝達する場合、DNA

を用いたベクター形態で伝達することに比べて、mRNAへの転写過程が不要であるため、遺伝子編集の開始が速い。一時的なタンパク質発現(transient protein expression)可能性が高い。

【０２４０】

本出願の発明者らは、植物小器官遺伝子編集のために、シトシン塩基編集器をリボ核酸、例えば、メッセンジャーリボ核酸の形態で植物細胞に注入したとき、プラスミドに伝達することに比べて非標的効果が減少することを確認した。植物小器官遺伝子編集において、シトシン塩基編集器をmRNA形態で植物細胞に形質転換したとき、プラスミドに比べて非標的効果において利点があることを最初に立証した。

【０２４１】

前記mRNAは直接送達されるか、キャリアを介して伝達することができる。場合によっては、核酸切断酵素および/または切断因子のmRNAを化学変形(chemical modification)するか、または合成自己－複製RNA(synthetic self-replicative RNA)の形態で直接伝達することができる。

【０２４２】

細胞にmRNAを伝達する方法、またはインビボでmRNAをヒト、動物のような有機体の細胞に伝達する方法を含み、インビトロまたはインビボでmRNA分子を細胞に伝達することができる方法が考えられる。例えば、脂質(例えば、リポソーム、マイセルなど)を含むか、ナノ粒子またはナノチューブを含むか、カチオン性化合物(例えば、ポリエチレンイミンまたはPEI)を含むか、またはカチオン性化合物(例えば、ポリエチレンイミンまたはPEI)を含んで、mRNA分子を細胞に伝達することができる。場合によっては、mRNAを細胞に伝達するために、遺伝子銃またはバイオリスティック(biolistic)粒子伝達システムのようなボリスティック法(bolistics method)を使用することができる。

【０２４３】

前記キャリアは、例えば、細胞透過性ペプチド(CPP)、ナノ粒子、またはポリマーを含むことができるが、これに限定されるものではない。

【０２４４】

前記CPPは、様々な分子カーゴ(ナノサイズ粒子から小さな化学分子およびDNAの大きな断片まで)の細胞吸収を容易にする短いペプチドである。

【０２４５】

前記ナノ粒子に関して、本発明による組成物は、高分子ナノ粒子、金属ナノ粒子、金属/無機ナノ粒子、または脂質ナノ粒子を介して伝達され得る。前記高分子ナノ粒子は、例えば、ローリングサークル増幅により合成されたDNAナノクルー、糸状DNAナノ粒子であってもよい。DNAナノクルー、糸状DNAナノ粒子にmRNAをローディングし、エンドソーム脱出能を向上させるためにPEIをコーティングすることができる。これらの複合体が細胞膜に結合し、内在化してからエンドソーム脱出を通じて核に移動し、伝達されるようにすることができる。

【０２４６】

前記金属ナノ粒子に関して、金粒子を連結し、カチオン性エンドソーム脱出ポリマー(endosomal disruptive polymer)と複合体を形成して細胞に送達することができる。前記カチオン性エンドソーム脱出ポリマーは、例えば、ポリエチレンイミン、ポリ(アラジン)、ポリ(リジン)、ポリ(ヒスチジン)、ポリ－[２－{(２－アミノエチル)アミノ}－エチル－アスパルトアミド](pAsp(DET))、ポリ(エチレングリコール)(PEG)とポリ(アルギニン)のブロックコポリマー、PEGとポリ(リジン)のブロックコポリマー、またはPEGとポリ{N-[N-(2-アミノエチル)-2-アミノエチル]アスパルトアミド}(PEG-pAsp DET))のブロックコポリマーであってもよい。

【０２４７】

前記金属/無機ナノ粒子に関して、例えば、ZIF-8(zeolitic imidazolate framework-8)を介してmRNAをカプセル化することができる。

【０２４８】

場合によっては、前記mRNAは、アニオンを有し、カチオンを有する物質と結合してナノ粒子を形成することができ、それは細胞へ受容体媒介エンドサイトーシス(receptor-mediated endocytosis)またはパゴサイトーシス(receptor-mediated endocytosis)などを通じて浸透することができる。

【０２４９】

カチオン性ポリマーとして、ポリアリルアミン(PAH)、ポリエチレンイミン(PEI)ポリ(Ｌ－リジン)(PLL)、ポリ(Ｌ－アルギニン)(PLA)、ポリビニルアミン単一またはコポリマー、ポリ(ビニルベンジル－トリ－Ｃ１－Ｃ４－アルキルアンモニウム塩)、脂肪族または芳香脂肪族ジハロゲン化物および脂肪族Ｎ、Ｎ、Ｎ’、Ｎ’－テトラ－Ｃ１－Ｃ４?アルキル?アルキレンジアミンのポリマー、ポリ(ビニルピリジン)またはポリ(ビニルピリジニウム塩)、ポリ(Ｎ,Ｎ－ジアリル－Ｎ，Ｎ－ジ－Ｃ１－Ｃ４－アルキル－アンモニウムハロゲン化物)、四級化ジ－Ｃ１－Ｃ４－アルキル?アミノエチルアクリレートまたはメタクリレートの単一またはコポリマー、ポリクアド(POLYQUAD)^{TM}、ポリアミノアミドなどが含まれ得る。

【０２５０】

カチオン性脂質として、カチオン性リポソーム製剤を含んでもよい。リポソームの脂質二重層は、カプセル化された核酸を分解から保護し、核酸に結合できる抗体による特異的中和を防ぐことができる。エンドソーム成熟中にエンドソームメムブレインとリポソームが融合し、カチオン性脂質－核酸切断酵素の効率的なエンドソーム脱出が可能である。カチオン性脂質として、ポリエチレンイミン、ポリアミドアミン(PAMAM)性化樹枝状体、リポフェクチン(DOTMAとDOPEの組合)、リポフェクターゼ、リポフェクタミン(LIPOFECTAMINE)(R)(例えば、リポフェクタミン(R)2000、リポフェクタミン(R)3000、リポフェクタミン(R) RNAiMAX、リポフェクタミン(R)LTX)、セント－レッド(SAINT-RED)(Synvolux Therapeutics社、オランダ・グロニンゲン所在)、DOPE、シトペクチン(Gilead Sciences社、カリフォルニアポスターシティ所在)、およびユーペクチン(JBL社、カリフォルニアサンルイス・オビスポ所在)を含んでもよい。代表的なカチオン性リポソームは、Ｎ－[１－(２,３－ジオレオオキシ)－プロピル]－Ｎ,Ｎ,Ｎ－トリメチルアンモニウムクロライド(DOTMA)、Ｎ-[１－(２,３－ジオレオオキシ)－プロピル]－Ｎ,Ｎ,Ｎ－トリメチルアンモニウムメチルスルフェート(DOTAP)、３β－[Ｎ－(Ｎ’,Ｎ’－ジメチルアミノエタン)カルバモイル]コレステロール(DC-hol)、２,３－ジオレイルオキシ－Ｎ－[２(スペルミンカルボキサミド)エチル]－Ｎ,Ｎ－ジメチル－１－プロパンアミドトリフルオロアセテート(DOSPA)、１,２－ジミリスチルオキシプロピル－３－ジメチル－ヒドロキシエチルアンモニウムブロミド、又はジメチルジオクタデシルアンモニウムブロミド(DDAB)から製造することができる。

【０２５１】

脂質ナノ粒子に関して、キャリアヘリポソームを用いて伝達することができる。リポソームは、内部水性区画を取り囲む単一または多重のラメラ脂質二重層と比較的不透過性の外側親脂性リン脂質二重層からなる球状小胞構造である。リポソーム剤形は、主に天然のリン脂質および脂質、例えば１,２－ジステアロリル－ｓｎ－グリセロ－３－ホスファチジルコリン(DSPC)、スフィンゴミエリン、ホスファチジルコリンまたはモノシアロガングリオシドなどを含んでもよい。場合によっては、血漿中の不安定性を解消するために、脂質膜にコレステロールまたは１,２－ジオレオイル－ｓｎ－グリセロ－３－ホスホエタノールアミン(DOPE)を追加することができる。コレステロールの添加は、カプセル化された生活性化合物の血漿中への迅速な放出を減少させるか、または１,２－ジオレオイル－ｓｎ－グリセロ－３－ホスホエタノールアミン(DOPE)は安定性を増加させる。

【０２５２】

７．塩基編集

本発明は、他の観点から、前記融合タンパク質または前記核酸を含む塩基編集用組成物に関する。

本発明は、他の観点から、前記組成物を細胞に処理する段階を含む塩基編集方法に関す

る。
DNA結合タンパク質、例えば、TALEまたはZFPがターゲットDNAに結合した後、融合タンパク質のシトシン脱アミノ酵素は、シトシンのアミノ基を加水分解してウラシルの形態に変形させる。ウラシルは、アデニンと塩基対をなすことができるので、細胞内のDNA複製過程を経て、シトシン－グアニン塩基対はウラシル－アデニン塩基対を経て最終的にチミン－アデニン塩基対に編集され得る。また、融合タンパク質のアデニン脱アミノ酵素は、アデニンのアミノ基を加水分解してヒポキサンチンの形態に変換させる。ヒポキサンチンはシトシンと塩基対をなすことができるので、同様に、細胞内のDNA複製過程で、アデニン－チミン塩基対は、ヒポキサンチン－シトシン塩基対を経てグアニン－シトシン塩基対に編集され得る。

【０２５３】
前記細胞は、真核細胞(例えば、酵母などの真菌、真核動物および/または真核植物由来細胞(例えば、胚細胞、幹細胞、体細胞、生殖細胞など)、真核動物(例えば、ヒト、サルなどの霊長類、犬、豚、牛、羊、ヤギ、マウス、ラットなど)、または真核植物(例えば、緑藻類などの藻類、コーン、大豆、小麦、稲など)であってもよいが、これに限定されない。

【０２５４】
(1) 植物細胞DNAの塩基編集
本発明は、植物細胞DNAの塩基編集用組成物または塩基編集方法に関する。前記融合タンパク質またはそれをコードする核酸；およびNLSペプチド、葉緑体伝達ペプチド(chloroplast transit peptide)、ミトコンドリア伝達タンパク質(MitochondrialTargeting Signal, MTS)、核外輸送信号タンパク質またはそれをコードする核酸を含む、植物細胞の塩基編集用組成物に関する。

【０２５５】
本発明は、また、前記融合タンパク質または核酸；及びNLSペプチドまたはそれをコードする核酸を含む植物細胞の塩基編集用組成物を提供する。

【０２５６】
本発明は、また、前記融合タンパク質または核酸；及び葉緑体伝達ペプチドまたはそれをコードする核酸を含む植物細胞の塩基編集用組成物を提供する。

【０２５７】
本発明は、また、前記融合タンパク質または核酸；及びミトコンドリア伝達タンパク質またはそれコードする核酸を含む植物細胞の塩基編集用組成物を提供する。

【０２５８】
場合によっては、本発明は、核外輸送信号タンパク質またはそれをコードする核酸をさらに含む、植物細胞の塩基編集用組成物を提供する。

【０２５９】
具体的には、植物細胞核DNA、ミトコンドリアDNAまたは葉緑体DNAの編集用組成物または塩基編集方法に関する。

【０２６０】
具体的には、前記融合タンパク質は以下を介して植物細胞に伝達され得る。
Gene gunを用いた注入(Bombardment)；
PEG-媒介プロトプラストトランスフェクション(Protoplast transfection, PEG-mediated)；
電気穿孔によるプロトプラストトランスフェクション(Protoplast transfection, electroporation)；または
微小注射によるプロトプラスト注射(Protoplast injection, microinjection)。

【０２６１】
本発明による融合タンパク質をコードするポリヌクレオチド配列は、RNA配列、DNA配列、またはそれらの組合(RNA-DNA組合配列)であってもよい。

【０２６２】

前記融合タンパク質をコードするポリヌクレオチドは、以下を介して植物細胞に伝達され得る。

【０２６３】

アグロバクテリウム、例えば、アグロバクテリウム・ツメファシエンスなどを用いた形質転換

-バイナリベクトル、

-ウイルスベクター：ジェミニウイルス(Geminivirus)、
タバコモザイクウイルス(tobacco rattle virus, TRV)、トマトモザイクウイルス (tomato mosaic virus, ToMV)、フォックステールモザイクウイル(foxtail mosaic virus, FoMV)、barley yellow striate mosaic virus(BYSMV)、Sonchus yellow net rhabdovirus(SYNV)など；

ウイルスを用いたトランスフェクション；

注入(Bombardment, Gene gun)；

PEG-媒介プロトプラストトランスフェクション(Protoplast transfection, PEG-mediated)；

電気穿孔によるプロトプラストトランスフェクション(Protoplast transfection, electroporation)；または

微小注射によるプロトプラスト注射(Protoplast injection, microinjection)。

【０２６４】

前記ウイルスは、例えば、ウイルスベクター：ジェミニウイルス(Geminivirus)、 タバコモザイクウイルス(tobacco rattle virus, TRV)、トマトモザイクウイルス (tomato mosaic virus, ToMV)、フォックステールモザイクウイル(foxtail mosaic virus, FoMV)、オオムギ縞萎縮ウイルス(barley yellow striate mosaic virus, BYSMV)、Sonchus yellow net rhabdovirus(SYNV)等が挙げられる。

【０２６５】

前記ベクターは、それぞれ局所注入法(例えば、病変または標的部位直接注入)、電気穿孔法(electroporation)、リポフェクション、ウイルスベクター、ナノ粒子だけでなく、PTD(Protein translocation domain)融合タンパク質法などを通じて細胞内に伝達され得る。

【０２６６】

前記植物器官に輸送されるタンパク質またはそれをコードする核酸に関して、前記植物器官はミトコンドリア、葉緑体、あるいは色素体(プラスチド：白色体、有色体)であってもよい。

【０２６７】

前記植物器官に輸送されるタンパク質は、例えば、葉緑体伝達ペプチドまたはミトコンドリア伝達タンパク質であってもよい。

【０２６８】

例えば、葉緑体伝達タンパク質(CTP)またはミトコンドリア伝達信号タンパク質(MTS)が結合して植物細胞内の葉緑体とミトコンドリアに伝達される。葉緑体とミトコンドリアに伝達されるとき、N-末端のCTPまたはMTSタンパク質を除く残りの部分は、前タンパク質形態で葉緑体とミトコンドリアの内部に伝達される。葉緑体とミトコンドリアの内部への進入過程で伝達タンパク質部分が離れていくようになり、葉緑体とミトコンドリアをターゲティングして特定部分を塩基編集することができる。

【０２６９】

前記融合タンパク質またはそれをコードする核酸の外に、葉緑体伝達タンパク質(CTP)またはそれをコードする核酸またはミトコンドリア伝達信号タンパク質(MTS)またはそれをコードする核酸を植物細胞に処理または伝達して、植物ミトコンドリア、葉緑体、有色体または白色体DNAの塩基を編集することができる。

【０２７０】

ミトコンドリア遺伝子の編集時に核外輸送配列を塩基編集タンパク質に付着した場合、

より高効率で塩基を編集することができる。前記核外輸送信号タンパク質は、例えばMVM(Mirute virus of mice)由来であり得るが、これに限定されるものではない。前記核外輸送信号タンパク質は、例えば、配列番号31のアミノ酸配列を含んでもよいが、これに限定されるものではない。

【０２７１】

VDEMTKKFGTLTIHDTEK (配列番号31)

本発明は、また、野生型DNA塩基配列を切断するが、編集された塩基配列は切断できないTAL(Transcription Activator-Like)エフェクター(TALE)ドメインおよびFokI核酸加水分解酵素(FokI Nuclease)またはそれをコードする核酸またはZFNまたはそれをコードする核酸、すなわち、ミトコンドリア核酸加水分解酵素であるmitoTALEN(Mitochondrial TALE Nuclease)またはそれをコードする核酸またはZFNまたはそれをコードする核酸をさらに含むことにより、ミトコンドリア配列切断タンパク質を同時に用いる場合でも、より高効率のミトコンドリア塩基編集が期待できる。

【０２７２】

(2)動物細胞DNAの塩基編集

本発明は、動物細胞DNAの塩基編集用組成物または塩基編集方法に関する。前記融合タンパク質またはそれをコードする核酸；及びNLSペプチド、 ミトコンドリア伝達タンパク質(MTS)、核外輸送信号タンパク質またはそれをコードする核酸を含む、動物細胞の塩基編集用組成物に関する。

【０２７３】

本発明は、また、前記融合タンパク質または核酸；及びNLSペプチドまたはそれをコードする核酸を含む、動物細胞の塩基編集用組成物を提供する。

【０２７４】

本発明は、また、前記融合タンパク質または核酸；及びミトコンドリア伝達タンパク質(MTS)またはそれをコードする核酸を含む、動物細胞の塩基編集用組成物を提供する。本発明は、さらに、場合によっては核外輸送信号をタンパク質またはそれをコードする核酸をさらに含む、動物細胞の塩基編集用組成物を提供する。

【０２７５】

前記動物細胞は非ヒト動物細胞であり、核外輸送信号タンパク質またはそれをコードする核酸；及び/またはミトコンドリア伝達信号タンパク質(MTS)またはそれをコードする核酸を処理して、非ヒト動物細胞ミトコンドリアDNAの塩基を編集することができる。

【０２７６】

例えば、前記融合タンパク質またはそれをコードする核酸の外に、ミトコンドリア伝達タンパク質(MTS)が結合してミトコンドリアに伝達される。ミトコンドリアに伝達されるとき、N-末端のMTSタンパク質を除く残りの部分は、全タンパク質の形でミトコンドリアの内部に伝達される。ミトコンドリアの内部への進入過程で伝達タンパク質の部分が離れていくことになり、ミトコンドリアをターゲティングして特定部分を塩基編集することができる。

【０２７７】

ミトコンドリア伝達信号、TALエフェクターおよびシトシン脱アミノ化酵素(cytosine deaminas, $DddA_{tox}$)を含むTALE-DdCBE(TALE DddA-derived cytosine base editor)またはそれをコードする核酸への核外輸送信号(Nuclear Export Signal, NES)またはそれをコードする核酸が結合した非ヒト動物細胞ミトコンドリアDNAの塩基編集用組成物または編集方法に関する。核外輸送信号により、ミトコンドリア－核類似配列に対する核DNA塩基編集を減少させることができる。

【０２７８】

本発明によれば、核外輸送信号タンパク質またはそれをコードする核酸を含み、より効率的な動物ミトコンドリアDNA塩基編集を達成することができる。また、核外輸送信号によりミトコンドリア－核有糸配列に対する核DNA塩基編集を減少させ、ミトコンドリアDNAのみが編集されるようにできる。

【０２７９】
前記核外輸送信号タンパク質は、例えばMVM(Mirute virus of mice)由来であり得るが、これに限定されるものではない。前記核外輸送信号タンパク質は、例えば、VDEMTKKFGTLTIHDTEK(配列番号31)のアミノ酸配列を含んでもよいが、これに限定されるものではない。
【０２８０】
前記(1) 核外輸送信号タンパク質またはそれをコードする核酸および(2) DNA結合タンパク質、脱アミノ酵素またはその変異体またはそれらをコードする核酸と同時に、あるいは順次に編集される前、野生型DNA塩基配列を切断するが、編集された塩基配列は切断できないTALエフェクター(TALE)ドメインおよびFokI核酸加水分解酵素またはそれをコードする核酸またはZFNまたはそれをコードする核酸を真核細胞に注入することができる。
【０２８１】
特に、真核細胞ミトコンドリア遺伝子の塩基編集に関して、核外輸送信号タンパク質またはそれをコードする核酸および/またはミトコンドリア伝達信号タンパク質(MTS)またはそれをコードする核酸を含んでもよい。
【０２８２】
本発明によると、動物ミトコンドリア遺伝子編集時に核外輸送配列を塩基編集タンパク質に付着した場合、より高効率で塩基が編集され、動物胚では、特に、核内類似配列の非特異的塩基編集も抑制されることが表れた。
【０２８３】
本発明は、また、ミトコンドリア核酸加水分解酵素であるmitoTALEN(Mitochondrial TALE Nuclease)またはそれをコードする核酸をさらに含むことにより、ミトコンドリア配列切断タンパク質を同時に用いる場合でも、より高効率のミトコンドリア塩基編集が期待できる。ミトコンドリアDNA核酸加水分解酵素であるmitoTALENを用いて、ミトコンドリアDNAを切断することができ、野生型(Wild-type)ミトコンドリアゲノムを切断して動物において塩基編集されたゲノムを高効率で得られる。
【０２８４】
本発明は、さらに、ミトコンドリア核酸加水分解酵素であるmitoTALENまたはそれをコードする核酸をさらに含むことにより、ミトコンドリア配列切断タンパク質を同時に用いる場合でも、より高効率のミトコンドリア塩基編集が期待できる。具体的には、ミトコンドリア伝達信号が結合したTALエフェクタードメインおよびFokI核酸加水分解酵素(FokI Nuclease)またはZFNまたはそれをコードする核酸を有する融合タンパク質(mitoTALEN)またはそれをコードする核酸を含んでもよい。
【０２８５】
ミトコンドリアDNA核酸加水分解酵素であるmitoTALENを用いて、ミトコンドリアDNAを切断することができ、野生型ミトコンドリアゲノムを切断して動物から塩基編集されたゲノムを高効率で得られる。
【０２８６】
場合によっては、塩基編集効率を高められるUGI(uracil DNA glycosylase inhibitor)をさらに含んでもよい。UGIは、DNAからUを除去する作用を触媒する突然変異DNAを修復する酵素UDG (Uracil DNA glycosylase)活性を阻害し、塩基編集効率を高めることができる。
【０２８７】
具体的に、分割されたバクテリア毒素DddAtox、DNAへ結合できるように考案されたTALEおよびウラシルグリコシル化阻害剤(UGI)からなるDddA－由来シトシン塩基編集器(DdCBEs)は、目標したシトシン－チミン塩基変形をミトコンドリアDNAで可能にした。実施例によると、マウス胚において高効率のミトコンドリアDNA編集が可能であることが示した。ミトコンドリア遺伝子中のNADH脱水およびユビキノンへの電子伝達を触媒するNADH脱水素化酵素のサブユニットをコードするMT-ND5(ND5)を目標とした

が、これは、m.G12918Aのように、ヒトのミトコンドリア疾患に関連した突然変異とm.C12336Tのような早期終止コドンを作る突然変異を含む。これにより、ミトコンドリア疾患モデルをマウスで作製することができ、ミトコンドリア疾患を治療できる可能性を提示した。

【０２８８】

前記(1)核外輸送信号タンパク質またはそれをコードする核酸に(2)DNA結合タンパク質、脱アミノ酵素またはその変異体またはそれをコードする核酸が連結され、前記(2)に核酸加水分解酵素である(3)mitoTALENまたはそれをコードする核酸を連結することができる。(1)～(3)を伝達するために、単一または複数の送達手段を同一または異なる構成で組み合わせて用いることができる。

【０２８９】

前記(1)は、第1の伝達手段に含まれ、(2)は第2の伝達手段に含まれ、(3)は第3の伝達手段に含まれてもよい。各々の送達システムは同時にウイルス伝達手段であるか、一方はウイルス伝達手段であり、他方は非ウイルス伝達手段であるか、同時に非ウイルス送達手段であってもよい。

【０２９０】

前記(1)～(3)の核外輸送信号タンパク質、DdCBE、mitoTALENを混合して伝達することができる。

【０２９１】

前記(1)～(3)の1つ以上は、核外輸送信号タンパク質、DdCBE、 mitoTALENに伝達し、一部は、(1)～(3)をコードするDNA配列がベクター上に位置するようにして伝達することができる。

【０２９２】

前記(1)～(3)をコードするDNA配列が同じベクター上に位置するようにし、1つのベクターを介して同時に伝達することができるか、別々の異なるベクター上に位置するようにして伝達することができる。

【０２９３】

本発明による動物は、ヒトまたは非ヒト動物を含んでもよい。前記非ヒト形質転換動物は、昆虫類、環状動物、軟体動物、腕足類、線形動物、腔腸動物、海綿類、脊索類、脊椎動物であってもよく、前記脊椎動物は、魚類、両生類、爬虫類、鳥類又は哺乳類であってもよく、前記昆虫類は、ショウジョウバエ (Drosophila)であってもよく、前記線形動物は、カエノラブディティス・エレガンス(C.elegans)であってもよく、前記魚類は、ゼブラフィッシュ(zebrafish)であってもよく、前記哺乳類は、霊長目、食肉目、食虫目、齧歯目、偶蹄目、 奇蹄目または長鼻目であってもよく、前記齧歯目は、ラットまたはマウスであってもよい。

【０２９４】

本発明による組成物をヒトまたは非ヒト動物の胚に導入し、前記胚を代理母体に移植して妊娠させて塩基編集された動物を製造することができる。動物の受精卵に本発明による組成物を導入して培養することができる。

【０２９５】

前記得られた受精卵を代理母に移植して分娩させることができる。前記非ヒト形質転換動物は、分娩後に形質転換可否を確認する段階をさらに含んでもよい。前記非ヒト形質転換動物を交配して子孫形質転換動物を産生することができる。

【０２９６】

前記「子孫」とは、前記非ヒト形質転換動物と交配して生存可能な全ての子形質転換動物を意味し、より具体的には、前記形質転換動物を親として形質転換動物同士または前記形質転換動物と正常動物を交配して生産したF1世代、F1世代の動物と正常動物を交配して生産したF2世代およびそれ以降世代を意味するが、これに限定されるものではない。

【０２９７】

前記交配は、前記形質転換動物または正常動物と交配することを特徴とすることができ

る。本発明は、前記形質転換動物または子孫形質転換動物から分離した細胞、組織および副産物を含んでもよい。前記副産物は、前記形質転換ウサギに由来する物質全てを意味するが、好ましくは、血液、血清、尿、糞便、唾液、臓器及び皮膚からなる群より選ばれることを特徴とすることができるが、これに限定されるものではない。

【発明の実施のための形態】

【０２９８】

以下、実施例により本発明をさらに詳細に説明する。これらの実施例は単に本発明を例示するためのものであり、本発明の範囲がこれらの実施例によって制限されるものと解釈されないことは当業界において通常の知識を有する者にとって自明である。

【０２９９】

実施例1．ジンクフィンガー脱アミノ酵素(Zinc Finger Deaminase：ZFD)

核DNAまたはミトコンドリアDNAの塩基編集は、医生命科学研究、医学および生命工学に広く有用である。ZFDプラットフォームは、DNA結合タンパク質、分割バクテリア毒性脱アミノ酵素DddAtox(interbacterial toxin deaminase DddAtox)およびウラシルグリコシラーゼ阻害剤を含む。UGIは、ヒト細胞において所望しない小さな挿入および欠失を誘導せずに標的化されたCからTへの塩基転換を触媒する。公開的に使用可能なジンクフィンガーリソースを使用してZFDをエンコードするプラスミドを作製し、核DNAで最大60%およびミトコンドリアDNAで30%の頻度で塩基編集を達成した。ZFDはCRISPRベースの塩基編集とは異なり、DNAを切断して一本鎖または二本鎖切断を作製しないため、標的位置での非相同末端付着(error-prone non-homologous end joinig)による所望しない挿入および欠失が発生しない。さらに、大腸菌(E. coli)で精製された組換えられたZFDタンパク質は、自発的にヒト細胞を通過し、ターゲット塩基を他の塩基に転換させる(base conversion)。　これは、遺伝子のない遺伝子治療の原理検証(proof-of-princile)を立証する。

真核細胞および生物のゲノム編集のための技術には、ジンクフィンガーヌクレアーゼ(ZFNs)、transcription activator-like effector(TALE)ヌクレアーゼ　(TALENs)、TALE-linked split interbacterial deaminase toxin DddA-derived cytosine base editors (a.k.a. DdCBEs)、CRISPR-Cas9、および切断する活性がなくなったCas9-linked deaminases (a.k.a base editors)が含まれているが、これに限らない。これらの道具は、原則として2つの機能単位で構成される。DNA結合部分と触媒部分である。したがって、ジンクフィンガーアレイまたはTALEアレイは、DNA結合部分として機能し、一方、ヌクレアーゼ(ZFNおよびTALENの場合はFokI)または脱アミノ酵素(DdCBEにおける分割$DddA_{tox}$およびCBEにおけるAPOBEC1)が触媒単位として機能する。Crispr-Cas9は、ヌクレアーゼ機能とRNAに誘導されるDNA結合タンパク質(RNA-guided DNA binding protein)機能の両方を有する。カスタマイズされた、プログラム可能なZFN、TALEN、およびCas9のようなヌクレアーゼは、DNAを切断して二本鎖切断(double strand break)を引き起こす。これを回復しながら、ターゲット化された方式で遺伝子のノックアウトおよびノックインを引き起こす。しかし、プログラム可能なヌクレアーゼが誘導した二本鎖切断は、標的部位で所望しない遺伝子の大きな損失(large deletion)、p53活性化および標的部位と非標的部位の同時DSB修復過程で染色体の再配列を引き起こす可能性がある。対照的に、シトシンおよびアデニン塩基編集器(CBEおよびABE)を含むプログラム可能な塩基編集器は、DSBを引き起こさないので、細胞内でこのような所望しない現象を回避でき、修復テンプレート(repair template)またはドナーDNA(donor DNA)なしで効率的に一塩基置換(single nucleotide conversition)を触媒することができる。しかし、CBEやABEはCas9ニッカーゼ変異体を含むため、DNA　片側の鎖を切断してニックまたは一本鎖切断(single-strand break)を引き起こして、依然として遺伝子標的部位で所望しない挿入－欠失が起こる。

【０３００】

これは、細胞内で核DNAとミトコンドリアDNAにおけるCをTへの塩基置換を触媒することができる。カスタマイズ(custom-designed) されたDdCBEを使用して、マウスの

ミトコンドリアDNA編集と植物における葉緑体DNA編集を示した。ヒト細胞と異なる真核細胞に欠失がなく、正確な塩基編集を行うジンクフィンガー脱アミノ酵素(ZFD)を、分割されたDddAtoxをカスタマイズされたジンクフィンガータンパク質に連結して作製した。ジンクフィンガーアレイ(Zinc finger array：2x0.3～0.6k base pair)は、TALEアレイ(2x1.7～2k base pair)やS.pyogenes Cas9(4.1 k base pair)と比較してサイズが小さいので、インビボ研究よよび遺伝子治療の応用のために、ZFDがエンコーディングされた遺伝子は、容易にウイルス制限された搭載スペース(cargo space)を有するウイルスベクター、例えば、AAVに搭載することができる。Taleアレイとは異なり、ジンクフィンガーアレイは、C-末端またはN-末端部分に体積が大きいドメインを有さないので、エンジニアリングに優しい。分割されたDddAtoxの半分は、ジンクフィンガータンパク質のC-末端部分またはN-末端部分に融合することができる。また、ジンクフィンガータンパク質には独自の細胞透過能力がある。これは、ヒト細胞に核酸なしに(nucleic acid-free)、遺伝子編集を行えるようにする。これらの特性は、ジンクフィンガータンパク質がDNA結合モジュールとして核または他の細胞小器官の塩基を編集するための理想的なプラットフォームになるようにする。

【０３０１】

1-1．材料と方法

プラスミドの製作

哺乳類における発現のためp3s-ZFDプラスミドは、p3s-ABE7.10プラスミド(addgene, #113128)をHindIIIおよびXhoI(NEB)酵素で切断した後に変形して作製した。酵素によって切断されたp3sプラスミドと合成された挿入DNAは、HiFi DNA組立キット(NEB)を用いて組み立てられた。MTS-、ZFP(Toolgen、SangamoおよびBarbasモジュール-、分割DddA-、またはUGIをエンコードするすべての挿入DNAはIDTにより合成された。pTargetプラスミドは、ZFD活性のためのスペーサー配列の最適長を決定するために設計された。様々の長さのスペーサーと2つのZFP結合部位を含む各pTargetプラスミドは、2つの酵素(EcoRIおよびBamHI、NEB)で分解されたpRGS-CCR5-NHEJレポータープラスミドにZFP結合配列とスペーサー配列を挿入して構成された。大腸菌におけるタンパク質産生のためのpET-ZFDプラスミドは、pET-Hisx6-rAPOBEC1-XTEN-nCas9-UGI-NLSプラスミド(addgene, #89508)をNcoIおよびXhoI(NEB)で切断した後、変形して生成された。ZFD配列は、PCRを用いてp3s-ZFDプラスミドから増幅し、Hisx6タグおよびGSTタグ配列はオリゴヌクレオチド(Macrogen)で合成した。タンパク質精製のためのすべてのプラスミドは、ZFDおよびタグをエンコードする配列を酵素で切断されたpETプラスミドにHiFi DNA組立キット(NEB)を使用して挿入することにより作製した。化学的に収容可能な DH5a大腸菌細胞へのプラスミド形質転換をし、プラスミドを製造社のプロトコルに従ってAccuPrep Plasmid Mini Extraction Kit (Bioneer)で精製した。サンガーシーケンシング(Sanger sequencing)で配列全体を確認した後、所望のプラスミドを選択した。

【０３０２】

HEK293T 細胞培養およびトランスフェクション

HEK293T細胞(ATCC CRL-11268)を、10％ウシ胎児血清(Welgene)と1％抗生剤－抗真菌溶液(Welgene)が添加されたDulbeccoのModified Eagle Medium(Welgene)で培養した。HEK293T細胞(7.5×10^4)を48ウェルプレートに播種した。18～24時間後、細胞が70～80％程度に成長すると、左側および右側のZFDをエンコーディングしているプラスミド(各500ng)をLipofectamine 2000(1.5μL、Invitrogen)を用いてトランスフェクションするか、pTargetプラスミド(10ng)と共にトランスフェクションした。細胞をトランスフェクション処理の96時間harvestした後、細胞溶解緩衝液(50mM Tris-HCl、pH8.0(Sigma-Aldrich)、1mM EDTA(Sigma-Aldrich)、0.005％ナトリウムドデシルスルフェート(Sigma-Aldrich))、5μLのプロテイナーゼK(Qiagen)を10.0?μLを入れた後、55℃で1時間、95℃で10分間インキュベーションした。全体mtDNAシークエンスを行うために、HEK293T細胞にND1またはND2ターゲットm

itoZFD対をエンコードするプラスミドまたはmRNAを連続的に希釈した濃度でトランスフェクションした。構造体の量(ng)は、7.5×10^{4}当たり伝達されたものである。mtDNAは、トランスフェクション後、96時間経て、細胞から抽出した。

【０３０３】

K562細胞培養およびトランスフェクション

K562細胞を、10%ウシ胎児血清(Welgene)および1%抗生剤－質抗真菌溶液(Welgene)が添加されたRPMI 1640培養液で培養した。電気穿孔によるK562細胞へのZFD伝達のために、プログラムFF-120(Lonza)を備えたAmaxa 4D-Nucleofector™ X Unitシステムを使用した。16ウェルNucleocuvette™Stripを使用するとき、各サンプルに添加された基質溶液の最大体積は2μLであった。タンパク質のような場合、左側及び右側のZFDそれぞれ220pmol(最大用量の場合)又は110pmol(最大用量の半分)を、プラスミドの場合、左側及び右側のZFDをエンコードするプラスミド500ngをK562細胞(1x105)にトランスフェクションした。処理後96時間経て、細胞を100gで5分間遠心分離して収集し、100μLの細胞溶解緩衝液(50mM Tris-HCl、pH8.0(Sigma-Aldrich)、1M EDTA(Sigma-Aldrich)、0.005%ナトリウムドデシルスルフェート(Sigma-Aldrich))、5μLのプロテイナーゼK(Qiagen)を添加して55℃で1時間、95℃で10分間インキュベーションした。ZFDまたはZFDエンコーディングプラスミドをK562細胞に直接伝達するために、以前にZFNの直接伝達に使用された方法を参照した。左側および右側のZFDタンパク質の混合物(最終濃度50μM)または左側および右側のZFDをエンコードするプラスミドの混合物(それぞれ500ng)を、100mM L－アルギニンおよび90μM ZnCl2を含むpH7.4の無血清培地で希釈し、最終体積を20μLとした。K562細胞(1×105)を100gで5分間遠心分離し、上清液を除去した。次いで、希釈されたZFD溶液に細胞を解放し、37℃で1時間インキュベーションした。インキュベーション後、細胞を100gで5分間遠心分離した上、新しい培養培地を入れた。細胞を30℃(一時的な低体温条件)または37℃で18時間維持し、次いで37℃で2日間増殖させた。一部の細胞は、前記過程に従って2回処理した。細胞を処理後96時間経て、分析した。

【０３０４】

ZFDタンパク質の発現と精製

それぞれC-末端GSTタグを有するZFD各対をエンコードするプラスミドをRosetta(DE3)コンピテント細胞に形質転換させた後、それをカナマイシンを含むLB寒天プレートで培養した。一晩培養した後、単一コロニーを選択し、37℃で50μg/mlカナマイシンおよび10.0μM ZnCl2を含む液体培地で一晩(前培養)培養した。翌日、前培養の一部を大量の液体培地に移し、吸光度A600nm、=～0.5～0.70になるまで37℃で培養した。培養物を約1時間氷上に置き、次いで0.5mMのイソプロピルβ－D－1－チオガラクトピラノシド(IPTG; GoldBio)を添加してZFDタンパク質発現を誘導し、培養物を18℃で14時間インキュベーションした。

【０３０５】

タンパク質精製過程としては、溶解浴液(50 mM Tris-HCl(Sigma-Aldrich), 500 mM NaCl(Sigma-Aldrich)、1 mM MgCl2(Sigma-Aldrich)、10 mM 1,4-dithiothreitol(DTT; GoldBio)、1% Triton X-10 Sigma-Aldrich)、10% グリセロール、1 mM フェニルメチルスルホニルフルオリド(Sigma-Aldrich)、1mg/ml 鶏卵白の上から抽出したリゾチーム(Sigma-Aldrich)、100μM ZnCl2(Sigma-Aldrich)、100mM Sigma-Aldrich)、pH 8.0)で細胞を溶解させ、細胞をさらに溶解させるために超音波処理(3 min total, 5 s on, 10 s off)を行った。これらを行った後、溶液は遠心分離(13,000 rpm)して上清液のみを抽出した。上清液は、Glutathione Sepharose 4B (GE healthcare)を添加して1時間インキュベーションした。インキュベーション後、resin-lysate混合物をカラムに注ぎ、洗浄緩衝液(50?mM Tris-HCl(Sigma-Aldrich)、500?mM NaCl(Sigma-Aldrich)、10?mM DTT(GoldBio)、 1 mM MgCl2(Sigma-Aldrich)、100?μM ZnCl2(Sigma-Aldrich)、10%グリセロール100 mM arginine(Sigma-Aldrich)、pH 8.0)で3回洗浄した。レジンに付着しているタ

ンパク質は、溶出バッファ(50?mM Tris-HCl(Sigma-Aldrich)、500?mM NaCl(Sigma-Aldrich)、1?mM MgCl2(Sigma-Aldrich)、40 mM glutathione(Sigma-Aldrich)、10% グリセロール、1?mM DTT(GoldBio)、100?μM ZnCl2(Sigma-Aldrich)、100 mM arginine(Sigma-Aldrich)、pH 8.0) でレジンから分離した。最後に、抽出されたタンパク質は、～15ng/μL(200～240pmol/μL、タンパク質のサイズに応じて)の濃度に濃縮された。

【０３０６】

ZFDによるPCRアンプリコンのインビトロ脱アミノ化

TRAC部位を含むアンプリコンは、PCRを用いて製造された。8μgのアンプリコンを37℃で1～2時間10.0 μM ZnCl2を含むNEB3.1緩衝液で2μgの各ZFDタンパク質(Left-G1397NおよびRight-G1397C)と一緒にインキュベーションした。反応後、4μLのプロテイナーゼK溶液(Qiagen)を入れて、55℃で30分　間インキュベーションしてZFDを除去し、アンプリコンは、PCR精製キット(MGmed)を用いて精製された。1μgの精製されたアンプリコンを37℃で1時間2つのUSER酵素(NEB)と一緒にインキュベーションした。次いで、アンプリコンを4μLのプロテイナーゼK溶液と一緒にインキュベーションし、PCR精製キットを用いて再び精製した。精製されたPCRアガロースゲルで電気泳動し、イメージ化した。

【０３０７】

ターゲットディープシーケンシング

標的およびオフターゲット部位の塩基編集比率の分析のために、標的部位は、重畳された1次PCR、2次PCR増幅、およびTruSeq HT Dualインデックス含有プライマーを用いた3次PCRをPrimeSTAR(R) GXL DNAポリメラーゼ(TAKARA)を用いてディープシーケンシングライブラリーを生成した。ライブラリーは、Illumina MiniSeqを使用してペアエンド(paired-end)シークエンシングをした。

【０３０８】

mRNAの調製

ZFD配列のアップストリームにT7 RNAポリメラーゼプロモーターを含むDNA鋳型は、順方向および逆方向プライマー(順方向：5'-CATCAATGGGCGTGGATAG-3' SEQ ID No: 116、逆方向：5'-CATCAATGGGCGTGGATAG-3' SEQ ID No: 117、逆方向：5'-GACACCTACTCAGACAATGC-3' SEQ ID No: 118)によってPCR作製された。mRNAは、mMESSAGE mMACHINE™ T7 ULTRA転写キット(Thermo Fisher)を用いてインビトロで合成された。インビトロで転写されたmRNAは、製造社のプロトコルに従って、MEGAclear™ Transcription Clean-Up Kit(Thermo Fisher)を用いて精製した。

【０３０９】

全体のミトコンドリアゲノムシークエンシング

全体のミトコンドリアゲノムシークエンシングの場合、3つの段階が必要である。1．分離されたミトコンドリアからmtDNA抽出：最初に、3×105　HEK293T細胞をトリプシン処理し、ND1またはND2標的mitoZFD対でトランスフェクション後、96時間後に遠心分離(500g、4分、4℃)によって収集した。次いで細胞をリン酸緩衝生理食塩水(Welgene)で洗浄し、遠心分離して再び収集した。上清液を除去し、製造社のプロトコルに従ってMitochondria Isolation Kit for Cultured Cells(Thermo Fisher)の試薬ベースの方法を用いて培養された細胞からミトコンドリアを分離した。mtDNAをDNeasy Blood&Tissue Kit(Qiagen)を用いて分離したミトコンドリアから抽出した。2．NGSライブラリーの生成：抽出したmtDNAからNGSライブラリーを生成するために、Nextera™ DNA CD Indexes(Illumina)が含まれたIllumina DNA Prepキットを使用した。3．NGS：ライブラリをプーリングし、MiniSeqシーケンサー(Illumina)にロードした。平均シークエンシングdepthは50より大きかった。

【０３１０】

ミトコンドリアゲノムの全体DNA編集の分析

全体のミトコンドリアゲノムシーケンシングのNGSデータを分析するために、BWAを用いてFastqファイルをGRCh38.p13(リリースv102)参照ゲノムに整列し、読み取りペアリング情報とフラグを修正してSAMtools(v.1.9)でBAMファイルを生成した。次いで、REDItools(v.1.2.1)のREDItoolDenovo.pyスクリプトを用いて、ミトコンドリアゲノムのすべてのシトシンおよびグアニンのうち、塩基編集率が1%以上の位置を識別した。細胞株の塩基編集率が50%以上の位置を単一のヌクレオチド変異とみなし、全てのサンプルから除外した。オフターゲット分析のために各ZFDの標的部位を除いた。編集頻度が≧1%の残りの位置をオフターゲットと見なし、編集されたC/Gヌクレオチドの数を計算した。ミトコンドリアゲノムのオフターゲットの平均C/G からT/Aの塩基編集頻度を計算するために、すべてのC/G について平均化しました。特異性の比率は、オンターゲット平均値をオフターゲット平均値で割って計算した。ミトコンドリアゲノム全体のグラフは、標的および標的から離れた位置で塩基編集率を表示して生成された。

【０３１１】

1-2．ZFD構造の最適化

ヒトおよび他の真核細胞で塩基編集を行うためのZFDを開発するために、まず半分のDddAtoxに付着するジングフィンガータンパク質(ZFP)のアミノ酸リンカーとスペーサー長を最適化しようとした。左側と右側のZFPが付着する位置の間のスペーサーでCからTへの塩基転換が起こる。適切に特性化されたヒトCCR5遺伝子をターゲットするZFN対を選んだ。これを用いて、2、5、10、16、24、および32の様々なアミノ酸リンカーを有するZFDを作製し、ZFDの左側と右側のZFP結合部位を有し、繰り返しTC配列が存在する1から24ベース対の長さを有する多様なスペーサーのある一連のターゲットプラスミドを作製した(図1a、b、および表1)。

【０３１２】

【表１】

【表１】 pTargetライブラリー配列

Spacer Length (bp)	Left-ZFD DNA binding region	Spacer sequence	Right-ZFD DNA binding region
1	GTCATCCTCATC (SEQ ID NO.: 119)	C	AAACTGCAAAAG(SEQ ID NO.: 120)
2	GTCATCCTCATC	TC	AAACTGCAAAAG
3	GTCATCCTCATC	TCG (SEQ ID NO.: 121)	AAACTGCAAAAG
4	GTCATCCTCATC	CTGA (SEQ ID NO.: 122)	AAACTGCAAAAG
5	GTCATCCTCATC	CTGAT (SEQ ID NO.: 123)	AAACTGCAAAAG
6	GTCATCCTCATC	CTGATC (SEQ ID NO.: 124)	AAACTGCAAAAG
7	GTCATCCTCATC	CTGATCG (SEQ ID NO.: 125)	AAACTGCAAAAG
8	GTCATCCTCATC	CTGATCGA (SEQ ID NO.: 126)	AAACTGCAAAAG
9	GTCATCCTCATC	CTGATCGAT (SEQ ID NO.: 127)	AAACTGCAAAAG
10	GTCATCCTCATC	CTGATCGATC (SEQ ID NO.: 128)	AAACTGCAAAAG
11	GTCATCCTCATC	CTGATCGATCG (SEQ ID NO.: 129)	AAACTGCAAAAG
12	GTCATCCTCATC	CTGATCGATCGA (SEQ ID NO.: 130)	AAACTGCAAAAG
13	GTCATCCTCATC	CTGATCGATCGAT (SEQ ID NO.: 131)	AAACTGCAAAAG
14	GTCATCCTCATC	CTGATCGATCGATC (SEQ ID NO.: 132)	AAACTGCAAAAG
15	GTCATCCTCATC	CTGATCGATCGATCG (SEQ ID NO.: 133)	AAACTGCAAAAG
16	GTCATCCTCATC	CTGATCGATCGATCGA (SEQ ID NO.: 134)	AAACTGCAAAAG
17	GTCATCCTCATC	CTGATCGATCGATCGAT (SEQ ID NO.: 135)	AAACTGCAAAAG
18	GTCATCCTCATC	CTGATCGATCGATCGATC (SEQ ID NO.: 136)	AAACTGCAAAAG
19	GTCATCCTCATC	CTGATCGATCGATCGATCG (SEQ ID NO.: 137)	AAACTGCAAAAG
20	GTCATCCTCATC	CTGATCGATCGATCGATCGA (SEQ ID NO.: 138)	AAACTGCAAAAG
21	GTCATCCTCATC	CTGATCGATCGATCGATCGAT (SEQ ID NO.: 139)	AAACTGCAAAAG
22	GTCATCCTCATC	CTGATCGATCGATCGATCGATC (SEQ ID NO.: 140)	AAACTGCAAAAG
23	GTCATCCTCATC	CTGATCGATCGATCGATCGATCG (SEQ ID NO.: 141)	AAACTGCAAAAG
24	GTCATCCTCATC	CTGATCGATCGATCGATCGATCGA (SEQ ID NO.: 142)	AAACTGCAAAAG

【０３１３】

DddAtoxは2つの位置(G1333とG1397)で分割することができ、各半分は左側または右側のZFPに融合することができる。生成された24個(=6個のlinkerx2個の分割位置xZFP融合位置(左側または右側))ZFD構成の塩基編集効率を各24個のスペーサーを有するターゲットプラスミドについて測定した。Hek293T細胞において、トランスフェクション(transfection)の第4 日次にディープシークエンシングを通じて測定した。

【０３１４】

[構造]

・Left-ZFD: SV40 NLS-ZFP(S162-left)-linker-DddAtox half-4aa linker-UGI

・Right-ZFD: SV40 NLS-ZFP(S162-right)-linker-DddAtox half-4aa linker-UGI
- SV40 NLS:
PKKKRKV (SEQ ID No: 478)
- ZFP(S162-left)
GIHGVPAAMAERPFQCRICMRNFSDRSNLSRHIRTHTGEKPFACDICGRKFAISSNLNSHTKIHTGSQKPFQCRICMRNFSRSDNLARHIRTHTGEKPFACDICGRKFATSGNLTRHTKIHLR (SEQ ID NO: 2)
- ZFP(S162-right)
GIHGVPAAMAERPFQCRICMRNFSRSDNLSVHIRTHTGEKPFACDICGRKFAQKINLQVHTKIHTGEKPFQCRICMRNFSRSDVLSEHIRTHTGEKPFACDICGRKFAQRNHRTTHTKIHLR (SEQ ID NO: 3)

【０３１５】

- ジンクフィンガータンパク質とDddAtox halfとの間のリンカー：
- 2aa: GS
- 5aa: TGEKP (SEQ ID No: 479)
- 10aa: SGAQGSTLDF (SEQ ID No: 9)
- 16aa: SGSETPGTSESATPES (SEQ ID No: 10)
- 24aa: SGTPHEVGVYTLSGTPHEVGVYTL (SEQ ID No: 115)
- 32aa: GSGGSSGGSSGSETPGTSESATPESSGGSSGGS (SEQ ID No: 11)
- Split-DddAtox G1333-N
GSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGG (SEQ ID No: 27)
- Split-DddAtox G1333-C
PTPYPNYANAGHVEGQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKMTVVPPEGAIPV KRGATGETKVFTGNSNSPKSPTKGGC (SEQ ID No: 272)

【０３１６】

- Split-DddAtox G1397-N
GSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGGPTPYPNYANAGHVEGQSALFMRDN GISEGLVFHNNPEGTCGFCVNMTETLLPENAKMTVVPPEG (SEQ ID No: 273)

【０３１７】

- Split-DddAtox G1397-C
AIPVKRGATGETKVFTGNSNSPKSPTKGGC (SEQ ID No: 26)
- 4aa linker
SGGS (SEQ ID NO: 480)
- UGI
TNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLTSDAPEYKPWA LVIQDSNGENKIKML (SEQ ID NO: 481)

【０３１８】

短いリンカーを有するZFD(2および5アミノ酸(AA)リンカー)は、効率が低いかない。一方、10AA以上のリンカーを有するZFDでは、4ベース対以上のスペーサーにおいて、CからTへの塩基編集効率が1%～24%の範囲と確認された(図1cおよび図2a、c)。ZFD対のうち24AAリンカーを有することが最も高い編集効率を示した。最適のリンカーの組合を見付けるために、ZFD対の左側ZFDに24AAリンカーを固定し、右側を多様なリンカー長を有するZFDで組み合わせて塩基編集効率を確認し、同様に、逆にしてみた(図1dおよび図3)。両方とも24AAリンカーを有することが最も効率が良いことが分かった。さらに、DddAtoxがG1397部分で分割されたときがG1333部分で分割されたときよりもさら良い効率を有することを確認した(図1cおよび図2a、b)。これらの最も高い効率を有するZFD対は、7～21bpのスペーサにおいて、＞6.7%の高効率でサイトシンを編集する(図1cおよび図2a、c)。

【０３１９】

1-3．インビボ核DNAターゲットにおける塩基編集

次に、ヒト細胞に24AAリンカーを有するZFDがインビボ染色体ターゲット位置でも、CからTへの塩基編集を触媒できるかどうかを調べた。総8個の遺伝子内にある11箇所(2箇所のZFD対が1箇所)をターゲットする22対のZFDが作製された(図4)。その中で、14対のZFDは、公開的に使用可能なジンクフィンガーリソースから入手して使用した。以前に特性化された(previously-characterized)ZFN(CCR5およびTRACに特定した)を改造して他の8対のZFDを作った。その理由は、ZFNのような場合は、スペーサ長が5～7bpでターゲットDNAを切断することに対し、ZFDは最小7以上のスペーサで機能するため、このようなZFN対で1つまたは2つのジンクフィンガーを着脱可能に機能できるZFDを作った。FokIヌクレアーゼは、部分はZFPにN-末端またはC-末端部分に融合することができ、4つの異なる構成を有するZFNを作ることができるため、また、他の構成を有する2対のZFD(図4bのTrac-NC)を作って、分割されたDddAtoxの半分も従来のZFP C-末端だけでなくZFPのN-末端部分にも融合できるかどうかを試験した(NC構成は図4a及び図5に示す)

【０３２０】

[構造]

・C type: SV40 NLS-Zinc finger protein-24aa linker-$DddA_{tox}$half-4aa linker-UGI

・N type: SV40 NLS-$DddA_{tox}$ half-24aa linker-Zinc finger protein-4aa linker-UGI

- SV40 NLS

PKKKRKV (SEQ ID No: 478)

- 24aa linker

SGTPHEVGVYTLSGTPHEVGVYTL

- Split-$DddA_{tox}$ G1397-N

GSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGGPTPYPNYANAGHVEGQSALFMRDN GISEGLVFHNNPEGTCGFCVNMTETLLPENAKMTVVPPEG

- Split-$DddA_{tox}$ G1397-C

AIPVKRGATGETKVFTGNSNSPKSPTKGGC

- 4aa linker

SGGS

- UGI

TNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLTSDAPEYKPWA LVIQDSNGENKIKML

- ZFP

CCR5-1 Left (C type) [S162 ZFN-Left]

GIHGVPAAMAERPFQCRICMRNFSDRSNLSRHIRTHTGEKPFACDICGRKFAISSNLNSHTKIHTGSQKPFQCRICMRNFSRSDNLARHIRTHTGEKPFACDICGRKFATSGNLTRHTKIHLR

【０３２１】

CCR5-1 Right (C type) [S162 ZFN-Right]

GIHGVPAAMAERPFQCRICMRNFSRSDNLSVHIRTHTGEKPFACDICGRKFAQKINLQVHTKIHTGEKPFQCRICMRNFSRSDVLSEHIRTHTGEKPFACDICGRKFAQRNHRTTHTKIHLR

【０３２２】

CCR5-2 Left (C type) [S162 ZFN-Left]

GIHGVPAAMAERPFQCRICMRNFSDRSNLSRHIRTHTGEKPFACDICGRKFAISSNLNSHTKIHTGSQKPFQCRICMRNFSRSDNLARHIRTHTGEKPFACDICGRKFATSGNLTRHTKIHLR (SEQ ID NO: 2)

【０３２３】
CCR5-2 Right (C type) [Modifying S162 ZFN-Right with additional ZF using the Barbas set of zinc finger modules]
GIHGVPAAMAERPFQCRICMRNFSQSGDLRRHIRTHTGEKPFACDICGRKFARSDNLSVHTKIHTGSQKPFQCRICMRNFSQKINLQVHIRTHTGEKPFACDICGRKFARSDVLSEHTKIHLR (SEQ ID NO: 482)
【０３２４】
TRAC-Left (C type) [Adapted from Paschon, D.E. et al., 2019]
GIHGVPAAMAERPFQCRICMRNFSDQSNLRAHIRTHTGEKPFACDICGRKFATSSNRKTHTKIHTGSQKPFQCRICMRNFSLQQTLADHIRTHTGEKPFACDICGRKFAQSGNLARHTKIHLR (SEQ ID NO: 483)
【０３２５】
TRAC-Left (N type) [Adapted from Paschon, D.E. et al., 2019]
FQCRICMRKFATSGSLTRHTKIHTGEKPFQCRICMRNFSRSDHLSTHIRTHTGEKPFACDICGRKFATSSNRTKHTKIHTHPRAPIPKPFQCRICMRNFSRSDNLSEHIRTHTGEKPFACDICGRKFAWHSSLRVHTKIHLR (SEQ ID NO: 484)
【０３２６】
TRAC-Right (C type) [From Paschon, D.E. et al., 2019]
GIHGVPAAMAERPFQCRICMRNFSRSDHLSTHIRTHTGEKPFACDICGRKFADRSHLARHTKIHTGSQKPFQCRICMRKFALKQHLNEHTKIHTGEKPFQCRICMRNFSQSGNLARHIRTHTGEKPFACDICGRKFAHNSSLKDHTKIHLR (SEQ ID NO: 485)
【０３２７】
MFAP1 Left (C type) [De novo designed using the Toolgen set of zinc finger modules]
GIRIPGEKPYSCGICGKSFSDSSAKRRHCILHTGEKPYTCSDCGKAFRDKSCLNRHRRTHTGEKPYKCMECGKAFNRRSHLTRHQRIHTGEKPYECNYCGKTFSVSSTLIRHQRIHLR (SEQ ID NO: 486)
【０３２８】
MFAP1 Right (C type) [Modifying S162 ZFN-Right with additional ZF using the Barbas set of zinc finger modules]
GIRERPYACPVESCDRRFSTSGSLVRHIRIHTGQKPFQCRICMRNFSRSDELTRHIRTHTGEKPFACDICGRKFARSDHLTTHTKIHTGEKPFQCRICMRKFAQSSNLVRHTKIHLR (SEQ ID NO: 487)
【０３２９】
CCDC28B Left (C type) [Modifying S162 ZFN-Right with additional ZF using the Barbas set of zinc finger modules]
GIRERPYACPVESCDRRFSDPGHLVRHIRIHTGQKPFQCRICMRNFSRSDELTRHIRTHTGEKPFACDICGRKFARSDHLTTHTKIHTGEKPFQCRICMRKFARSDKLVRHTKIHLR (SEQ ID NO: 488)
【０３３０】
CCDC28B Right (C type) [De novo designed using the Toolgen set of zinc finger modules]
GIRIPGEKPYECNYCGKTFSVSSTLIRHQRIHTGEKPYHCDWDGCGWKFARSDELTRHYRKHTGEKPFECKDCGKAFIQKSNLIRHQRTHTGEKPYHCDWDGCGWKFARSDELTRHYRKHLR (SEQ ID NO: 489)
【０３３１】
KDM4B Left (C type) [Modifying S162 ZFN-Right with additional ZF using the Barbas set of zinc finger modules]
GIRERPYACPVESCDRRFSDCRDLARHIRIHTGQKPFQCRICMRNFSRSDELTRHIRTHTGEKPFACDICGRKFARSDHLTTHTKIHTGEKPFQCRICMRKFARSDKLVRHTKIH

LR (SEQ ID NO: 490)

【0332】

KDM4B Right (C type) [De novo designed using the Toolgen set of zinc finger modules]

GIRIPGEKPFECKDCGKAFIQKSNLIRHQRTHTGEKPYRCEECGKAFRWPSNLTRHKRIHTGEKPYRCEECGKAFRWPSNLTRHKRIHTGEKPYSCGICGKSFSDSSAKRRHCILHLR (SEQ ID NO: 491)

【0333】

NUMBL Left (C type) [Modifying S162 ZFN-Right with additional ZF using the Barbas set of zinc finger modules]

GIRERPYACPVESCDRRFSDCRDLARHIRIHTGQKPFQCRICMRNFSRSDELTRHIRTHTGEKPFACDICGRKFARSDHLTTHTKIHTGEKPFQCRICMRKFARSDKLVRHTKIHLR (SEQ ID NO: 492)

【0334】

NUMBL Right (C type) [De novo designed using the Toolgen set of zinc finger modules]

GIRIPGEKPYTCSDCGKAFRDKSCLNRHRRTHTGEKPYKCGQCGKFYSQVSHLTRHQKIHTGEKPFECKDCGKAFIQKSNLIRHQRTHTGEKPYHCDWDGCGWKFARSDELTRHYRKHLR (SEQ ID NO: 493)

【0335】

INPP5D-1 Left (C type) [Modifying S162 ZFN-Right with additional ZF using the Barbas set of zinc finger modules]

GIRERPYACPVESCDRRFSRSDKLVRHIRIHTGQKPFQCRICMRNFSRSDELTRHIRTHTGEKPFACDICGRKFARSDHLTTHTKIHTGEKPFQCRICMRKFARSDKLVRHTKIHLR (SEQ ID NO: 494)

【0336】

INPP5D-1 Right (C type) [De novo designed using the Toolgen set of zinc finger modules]

GIRIPGEKPYTCSDCGKAFRDKSCLNRHRRTHTGEKPYECNYCGKTFSVSSTLIRHQRIHTGEKPYTCSDCGKAFRDKSCLNRHRRTHTGEKPYTCSDCGKAFRDKSCLNRHRRTHLR (SEQ ID NO: 495)

【0337】

INPP5D-2 Left (C type) [Modifying S162 ZFN-Right with additional ZF using the Barbas set of zinc finger modules]

GIRERPYACPVESCDRRFSRSDKLVRHIRIHTGQKPFQCRICMRNFSRSDELTRHIRTHTGEKPFACDICGRKFARSDHLTTHTKIHTGEKPFQCRICMRKFARSDKLVRHTKIHLR (SEQ ID NO: 496)

【0338】

INPP5D-2 Right (C type) [De novo designed using the Toolgen set of zinc finger modules]

GIRIPGEKPYECNYCGKTFSVSSTLIRHQRIHTGEKPYTCSDCGKAFRDKSCLNRHRRTHTGEKPYTCSDCGKAFRDKSCLNRHRRTHTGEKPYHCDWDGCGWKFARSDELTRHYRKH (SEQ ID NO: 497)

【0339】

DVL3 Left (C type) [De novo designed using Barbas zinc finger modules]

GIHGVPAAMAERPFQCRICMRNFSTSGHLVRHIRTHTGEKPFACDICGRKFATSGHLVRHTKIHTGEKPFQCRICMRNFSTSGELVRHIRTHTGEKPFACDICGRKFAQSSNLVRHTKIHLR (SEQ ID NO: 498)

【0340】

DVL3 Right (C type) [S162 ZFN-left]

GIHGVPAAMAERPFQCRICMRNFSDRSNLSRHIRTHTGEKPFACDICGRKFAISSNLNSHTKIHTGSQKPFQCRICMRNFSRSDNLARHIRTHTGEKPFACDICGRKFATSGNLTRHTKIHLR (SEQ ID NO: 499)

【０３４１】

HEK293T細胞において、NC構成を有するZFDを含んでZFDのCからTへの塩基編集効率は1.0%～60%を示す。一方、挿入－欠失は＜0．4%程度であって、ほとんど起こらなかった(図4ｂおよび図6)。前記プラスミドをベースとした実験で読み取れるように、スペーサー5bpを有するCCR5ターケットするZFDは効率が非常に劣る。スペーサー長が少なくとも7以上のターゲットで他の20個のZFD対は平均編集効率が12.0±3.4%を示し、これは、平均8.3±2.2%を示すCas9ベースの塩基編集遺伝子(Cas9-derived Base editor2)と比べ合せられるほどである。さらに、CからTへの塩基編集は、TC conxtextのみならず、ACやGCCでも起こった(図4c～f)。ACはNUMBLのC_6で、GCCはINPP5D-2のC_7でそれぞれC塩基編集効率が4.58%と1.85%まで表れた。

【０３４２】

		Left-ZFD binding region	Spacer region	Right-ZFD binding region		ZFD configuration
MFAP1	5’-	GACGGCCCCAGC	CGTAGACT	GAATGGGCGGTT	-3’	C-C
	3’-	CTGCCGGGGTCG	GCATCTGA	CTTACCCGCCAA	-5’	
CCDC28B	5’-	TGCCGCCCAGTC	CTGGCTGGAG	GTGGAAGCGGCT	-3’	C-C
	3’-	ACGGCGGGTCAG	GACCGACCTC	CACCTTCGCCGA	-5’	
KDM4B	5’-	GGCCGCCCACCC	GGTCCCCACT	GTCGGTGGTGAA	-3’	C-C
	3’-	CCGGCGGGTGGG	CCAGGGGTGA	CAGCCACCACTT	-5’	
NUMBL	5’-	GGCCGCCCACCC	CCAGACAG	GCGGAAGGAGCC	-3’	C-C
	3’-	CCGGCGGGTGGG	GGTCTGTC	CGCCTTCCTCGG	-5’	
INPP5D-1	5’-	CCCCGCCCACCC	CGACGCCGCG	GCCGCCGCTGCC	-3’	C-C
	3’-	GGGGCGGGTGGG	GCTGCGGCGC	CGGCGGCGACGG	-5’	
INPP5D-2	5’-	CCCCGCCCACCC	CGACGCC	GCGGCCGCCGCT	-3’	C-C
	3’-	GGGGCGGGTGGG	GCTGCGG	CGCCGGCGGCGA	-5’	
DVL3	5’-	ACCACCAGCTTC	TTTGACTCA	GATGAGGATGAC	-3’	C-C
	3’-	TGGTGGTCGAAG	AAACTGAGT	CTACTCCTACTG	-5’	
CCR5-1	5’-	GTCATCCTCATC	CTGAT	AAACTGCAAAAG	-3’	C-C
	3’-	CAGTAGGAGTAG	GACTA	TTTGACGTTTTC	-5’	
CCR5-2	5’-	GTCATCCTCATC	CTGATAAA	CTGCAAAAGGCT	-3’	C-C
	3’-	CAGTAGGAGTAG	GACTATTT	GACGTTTTCCGA	-5’	
TRAC-CC	5’-	GTGATTGGGTTC	CGAATCCTCCT	CCTGAAAGTGGCCGG	-3’	C-C
	3’-	CACTAACCCAAG	GCTTAGGAGGA	GGACTTTCACCGGCC	-5’	
TRAC-NC	5’-	TGTCAGTGATTGGGTT	CCGAATCCTCCT	CCTGAAAGTGGCCGG	-3’	N-C
	3’-	ACAGTCACTAACCCAA	GGCTTAGGAGGA	GGACTTTCACCGGCC	-5’	

【０３４３】

1-4．精製されたZFDタンパク質をヒト細胞に直接伝達

遺伝子編集タンパク質をエンコーディングしているプラスミドDNAを伝達する代わりに、精製された遺伝子編集タンパク質自体を細胞に伝達することがオフターゲット影響を低減し、外部DNAによる先天的免疫応答(innate immune response)を避け、外部プラスミドDNAのインビボゲノムに挿入して入ることを防ぐことができる。他のグループは、インビトロおよびインビボにおいて、ZFPが哺乳動物細胞を自発的に通過して入られることを示した。ZFDのタンパク質媒介塩基編集を証明するために、効率が高かったTRAC遺伝子をターゲットするZFD対を選び、1個または4個のNLSを有するZFD組換えタン

パク質を大腸菌から精製した。まず、ZFDタンパク質の塩基編集効率をTRACサイトを有するPCRアンプリコン(amplicon)を用いてインビトロで実験を行い、効率が非常に高いことを確認した。効率はUracil DNA glycosylaseとDNA glycosylase-lyase EndoNuclease VIIIの混合物であるウラシル特異的切断試薬(USER)を通して遺伝子が切断されることで確認できた(図7)。トランスフェクションが困難な細胞であるヒト白血病細胞K562細胞へTRAC-NC ZFDタンパク質を、2つの方式、電気穿孔または電気穿孔なしで直接伝達(direct delivery)により伝達した。ZFDタンパク質は、非常に効率的だった。CからTへの塩基編集は26.5%(electroporation)および17%(direct delivery)を示した(図4g)。まとめると、これらの結果は、ZFDをエンコードするプラスミドまたは精製された組換えZFDタンパク質がヒト細胞の核DNAを塩基編集することに用いられることを示す。

【０３４４】

1-5．ZFDのミトコンドリアDNA塩基編集

クリスパーベースシステムとは異なり、カスタマイズされた(custom-designed)DNA結合タンパク質に分割DddAtoxを融合したシステムの大きな利点は、かかるプログラム可能な塩基編集はさみが、ミトコンドリアDNAのような細胞小器官DNAを編集することに用いることができるとのことである。ZFDをミトコンドリアに伝達するために、MTSとNESをミトコンドリア遺伝子をターゲットするように設計された9つのZFDのN-末端部分に連結してmitoZFDを作製した(図8)。ZFDのZFP部分は、公開的に使用可能なジングフィンガーリソースから得られた。ZFDは、スペーサー長が7～15bpであるように作製し、DNA結合部位は左側と右側のいずれも12bpとなるようにした。

【０３４５】

[構造]

・C type: MTS-FLAG tag-NES-Zinc finger protein-24aa linker-DddA$_{tox}$half-4aa linker-UGI

・N type: MTS-HA tag-NES-DddA$_{tox}$ half-24aa linker-Zinc finger protein-4aa linker-UGI

- MTS (Mitochondrial Targeting Sequence of human mitochondrial ATP synthase F1β subunit)

MLGFVGRVAAAPASGALRRLTPSASLPPAQLLLRAAPTAVHPVRDYAAQ (SEQ ID No:274)

- FLAG tag (C type)

DYKDDDDK (SEQ ID No:275)

- HA tag (N type)

YPYDVPDYA (SEQ ID No:276)

【０３４６】

- NES (Nuclear export signal)

VDEMTKKF (Minute virus if mice; MVM NES)

- Split-DddA$_{tox}$ G1397-N

GSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGGPTPYPNYANAGHVEGQSALFMRDN GISEGLVFHNNPEGTCGFCVNMTETLLPENAKMTVVPPEG

- Split-DddA$_{tox}$ G1397-C

AIPVKRGATGETKVFTGNSNSPKSPTKGGC

- 4aa linker

SGGS

- UGI

TNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLTSDAPEYKPWA LVIQDSNGENKIKML

【０３４７】

- ZFP

＊ジンクフィンガーは、TGEKPリンカーで連結されている。(ZF1-linker-ZF2-linker-ZF3-linker-ZF4).

＊ND2－ターゲットmitoZFDは、QQ変異体を有している。この変異体は、RまたはKアミノ酸をQに置換したものである。赤文字で表示されている。

	Type	ZF1	No.	ZF2	No.	ZF3	No.	ZF4	No.
ND1_ Left	C	FQCRICM RNFSDSG NLRVHIRT H	500	YKCPDCG KSFSQSS SLIRHQRT H	518	YECDHCG KSFSQSS HLNVHKR TH	536	YRCKYCD RSFSISSN LQRHVRNI H	554
ND1_ Right	N	YKCPECG KSFSTKN SLTEHQRT H	501	YKCPECG KSFSSKK ALTEHQRT H	519	YECNYCG KTFSVSST LIRHQRIH	537	YRCKYCD RSFSISSN LQRHVRNI H	555
ND2_ Left	C	YKCPECG KSFSRED NLHTHQRT H	502	YHCDWD GCGWKF ARSDELT RHYRKH	520	YKCPECG KSFSRED NLHTHQRT H	538	YRCKYCD RSFSISSN LQRHVRNI H	556
ND2_ Right	N	YECDHCG KSFSQSS HLNVHKR TH	503	YHCDWD GCGWKF ARSDELT RHYRKH	521	YKCPDCG KSFSQSS SLIRHQRT H	539	YKCGQCG KFYSQVS HLTRHQKI H	557
Cox1_Left	C	YECHDCG KSFRQST HLTQHRRI H	504	YSCGICG KSFSDSS AKRRHCIL H	522	YECHDCG KSFRQST HLTQHRRI H	540	YKCPDCG KSFSQSS SLIRHQRT H	558
Cox1_Right	N	YKCPDCG KSFSQSS SLIRHQRT H	505	YKCGQCG KFYSQVS HLTRHQKI H	523	YKCPDCG KSFSQSS SLIRHQRT H	541	YECDHCG KSFSQSS HLNVHKR TH	559
Cox2_Left	C	YKCPECG KSFSDPG HLVRHQR TH	506	YRCKYCD RSFSISSN LQRHVRNI H	524	YRCKYCD RSFSISSN LQRHVRNI H	542	YKCPECG KSFSRED NLHTHQRT H	560
Cox2_Right	N	YGCHLCG KAFSKSS NLRRHEMI H	507	FQCKTCQ RKFSRSD HLKTHTRT H	525	YKCGQCG KFYSQVS HLTRHQKI H	543	YKCPECG KSFSRED NLHTHQRT H	561
ND4L_Left	C	YKCPECG KSFSDPG HLVRHQR TH	508	YKCPECG KSFSRED NLHTHQRT H	526	YKCPECG KSFSRED NLHTHQRT H	544	YRCKYCD RSFSDSS NLQRHVR NIH	562
ND4L_Right	C	YKCPECG KSFSDPG HLVRHQR TH	509	FHCGYCE KSFSVKD YLTKHIRT H	527	YKCPECG KSFSSPA DLTRHQRT H	545	FQCRICM RNFSDSG NLRVHIRT H	563
ND4_ Left	C	FHCGYCE KSFSVKD YLTKHIRT H	510	YECVQCG KGFTQSS NLITHQRV H	528	YTCSYCG KSFTQSN TLKQHTRI H	546	YRCKYCD RSFSISSN LQRHVRNI H	564
ND4_ Right	C	FHCGYCE KSFSVKD YLTKHIRT H	511	YKCPECG KSFSTHLD LIRHQRTH	529	YKCPECG KSFSSKK ALTEHQRT H	547	FQCRICM RNFSDSG NLRVHIRT H	565
ND5_ Left	C	FHCGYCE KSFSVKD YLTKHIRT H	512	YECDHCG KSFSQSS HLNVHKR TH	530	FQCRICM RNFSDSG NLRVHIRT H	548	YECDHCG KSFSQSS HLNVHKR TH	566
ND5_ Right	N	FACPECP KRFMRSD NLTQHIKT H	513	YECNYCG KTFSVSST LIRHQRIH	531	YECDHCG KAFSVSS NLNVHRRI H	549	YECDHCG KAFSVSS NLNVHRRI H	567
ND6_ Left	C	YECHDCG KSFRQST HLTQHRRI H	514	YECHDCG KSFRQST HLTQHRRI H	532	YKCGQCG KFYSQVS HLTRHQKI H	550	YHCDWD GCGWKF ARSDELT RHYRKH	568
ND6_ Right	N	YRCEECG KAFRWPS NLTRHKRI H	515	YKCPECG KSFSRED NLHTHQRT H	533	YGCHLCG KAFSKSS NLRRHEMI H	551	YKCPECG KSFSRED NLHTHQRT H	569
CYB_ Left	C	FECKDCG KAFIQKSN LIRHQRTH	516	FHCGYCE KSFSVKD YLTKHIRT H	534	YTCSDCG KAFRDKS CLNRHRR TH	552	YTCSYCG KSFTQSN TLKQHTRI H	570
CYB_Right	C	YTCSYCG KSFTQSN TLKQHTRI H	517	YKCDECG KNFTQSS NLIVHKRIH	535	YKCPECG KSFSDPG HLVRHQR TH	553	YECDHCG KSFSQSS HLNVHKR TH	571

【０３４８】

		Left-ZFD binding region	Spacer region	Right-ZFD binding region		ZFD configuration
ND1	5’-	GTTTACTCAATC	CTCTGATC	AGGGTGAGCATC	-3’	C-N
	3’-	CAAATGAGTTAG	GAGACTAG	TCCCACTCGTAG	-5’	
ND2	5’-	CTACGCCTAATC	TACTCCACCTCAA	TCACACTACTCC	-3’	C-N
	3’-	GATGCGGATTAG	ATGAGGTGGAGTT	AGTGTGATGAGG	-5’	
Cox1	5’-	TCTGACTCTTAC	CTCCCTCTCTCC	TACTCCTGCTCG	-3’	C-N
	3’-	AGACTGAGAATG	GAGGGAGAGAGG	ATGAGGACGAGC	-5’	
Cox2	5’-	GCCATCATCCTA	GTCCTCATCGCC	CTCCCATCCCTA	-3’	C-N
	3’-	CGGTAGTAGGAT	CAGGAGTAGCGG	GAGGGTAGGGAT	-5’	
ND4L	5’-	GCCCTACTAGTC	TCAATCTCC	AACACATATGGC	-3’	C-C
	3’-	CGGGATGATCAG	AGTTAGAGG	TTGTGTATACCG	-5’	
ND4	5’-	ATATTTTATATC	TTCTTCGA	AACCACACTTAT	-3’	C-C
	3’-	TATAAAATATAG	AAGAAGCT	TTGGTGTGAATA	-5’	
ND5	5’-	ATATCGGTTTCA	TCCTCGC	CTTAGCATGATT	-3’	C-N
	3’-	TATAGCCAAAGT	AGGAGCG	GAATCGTACTAA	-5’	
ND6	5’-	TCTTCTTCCCAC	TCATCCTA	ACCCTACTCCTA	-3’	C-N
	3’-	AGAAGAAGGGTG	AGTAGGAT	TGGGATGAGGAT	-5’	
CYB	5’-	TTCATAGGCTAT	GTCCTCCCG	TGAGGCCAAATA	-3’	C-C
	3’-	AAGTATCCGATA	CAGGAGGGC	ACTCCGGTTTAT	-5’	

【０３４９】

ＨＥＫ２９３Ｔ細胞において、ｍｉｔｏＺＦＤによるミトコンドリアＤＮＡ塩基編集効率は、２．６％～３０％(平均１４±３％)である(図９ａ)。ＮＣ構成を有するｍｉｔｏＺＦＤ(１６±４．５％、ｎ＝６)は、ＣＣ構成(８．３±２．７％、ｎ＝３)を有するほど効率が出た。スペーサー部分でＴＣコンテキストを有するほとんどのサイトシンは様々な効率で塩基編集された(図９ｂ～ｇ)。さらに、ＮＤ２部にあるＡＣＣ(C_8及びC_9)コンテキストを有するサイトシンもそれぞれ７．４％と１９．９％程度の効率を示した(図９ｂ)。これは、ＺＦＤを媒介とするＣからＴへの塩基編集がＴＣモチフに限定されないことを示す。

【０３５０】

次に、ミトコンドリアDNA(mtDNA)変異細胞から単細胞由来クローン集団(single-cell derived clonal populations)を分離し、mitoZFDが毒性がなく、変異体mtDNAがクローン集団でも維持されることを証明しようとした。ND1-specific mitoZFDを処理したHEK293T細胞から分離した単細胞由来30個のクローン集団のうち、5個ND1遺伝子の塩基編集効率が35%から98%に出た(図10a)。同様に、ND2-specific mitoZFD を処理したHEK293T細胞から分離した単細胞由来36個のクローン集団のうち、7個はND2遺伝子の塩基編集効率が26%から76%に出た(図10b)。その他の全てのクローン集団については0.4%～1.0%と低い効率を示し、ほとんどはシーケンスエラーのように見える。ZFDを処理していない細胞でも類似な効率が出た(図10c)。このような結果は、mitoZFDがヘテロプラスミックな変異を細胞個体で均一に引き起こさないことを示す。ZFDを処理した細胞において、ほとんどは野生型であり、一方、ヘテロプラスミック変異を有するものは、98%までの変異率を有する。この変化は、クローン増殖(clonal expansion)を行っても維持し続ける(図11および図12)。

【0351】

1-6. mitoZFDsおよびTALEベースのDdCBEs

作製したND1-specific mitoZFDが同じ遺伝子をターゲットするTALEベースDdCBEと変異パターンが異なることが分かった(図9f-h)。2つのmitoZFDのような場合、C_5またはC_8位のサイトシンにおいて、CからTへの塩基編集を行うことに対し(図9f)、DdCBEは、C_8 、C_9 、およびC_{11}の部位を塩基編集する(図9g)。その結果、mitoZFDによるアミノ酸の変化は、完全にDdCBEによるアミノ酸変化と相違する(図9h)。mitoZFDが付着する左側と右側の部分がスペーサー8-bpの長さに分けられたことに対し、DdCBEのような場合、スペーサー長が16-bpの長さに分けられているため、かかる相違により、異なる変異パターンを有することがある。これらの結果は、ミトコンドリアDNAにおいて、mizoZFDとDdCBEが互いに相補的に様々な変異を生み出すことができることを示す。

【0352】

変異パターンをより多くするために、ZFD単一体とDdCBE単一体が混合し、混合対(hybrid pair)として使用できるかどうかを試験した。ND1遺伝子をターゲットする10個の混合対は、HEK293T細胞において平均17±3.4%の塩基編集効率で良好な活性を示す(図13)。実際に、混合対の1つ(TALE-L/ZFD-R1)は 2つのDdCBE対および10個の同じ場所をターゲットするZFD対よりも優れた効率を示し、最高41%の塩基編集効率を示した(図13ｂ)。さらに、混合対は、DdCBEおよびmitoZFDとは異なる変異パターンを示した(図13c)。いくつか(例えば、ZFD-L1/TALE-RおよびZFD-L2/TALE-R)は、他の付随的な変異なしに、1箇所のみでCからTへの変異が起こした。対照的に、ほとんどのmitoZFD対とDdCBE対は、スペーサー内の様々な場所でCからTへの変化を引き起こす。これらの結果は、ZFD/DdCBE混合対が特異的な変異パターンを作ることができ、ZFD対やDdCBE対では作れない特定の変異を生み出すことができることを示す。

【0353】

1-7. mitoZFDのミトコンドリアゲノムワイドターゲットの特異性

mitoZFDがミトコンドリアDNAでオフターゲットを引き起こすことを確認するためにND1またはND2遺伝子をターゲットする2対のmitoZFDを処理した細胞から遺伝子を抽出し、ミトコンドリアの全体の遺伝子シーケンシング(whole mitochondrialgenome sequencing)をした。多様な量(5～500ng)のmitoZFD対をエンコードするmRNAまたはプラスミドをHEK293T細胞にトランスフェクションした。予想の通り、オンターゲット効率は、量に依存的だった(dose dependent)。 高い濃度(100、200、および500ng)のmRNAまたはプラスミドは、オンターゲットで＞30%の効率を示したが、数百個の1%を超えるオフターゲットも発生させた(図14-17)。低濃度(5および10ng)では、これらのオフターゲットがほとんどなくなったが、オンターゲットにおける効率も明らかに低くなった。中間濃度である50ng量のmRNAが最も適切であった。数百個のオフターゲットを引き起こすことなく、高いオンターゲット効率を維持した。残りのオフターゲット

をなくすために、R(-5)Q変異を各ジンクフィンガー部分に導入して非特異的DNA接触を除外しようとした。作製されたZFD変異体(図18でQQと示した)は、高いオンターゲット活性は維持しながら、オフターゲットがZFDを処理してない細胞のmtDNAに比べてほとんどない高精度な特異性を有した(図18aおよびb)。

【０３５４】

塩基編集は、DNAの二本鎖切断を引き起こすことなく、DNA修復テンプレートなしでターゲット塩基を編集できる比較的新しい技術である。塩基編集は、細胞、動物および植物において、CからTまたはAからGへの塩基編集を引き起こし、一塩基変異(SNP)の機能的影響を研究することができ、治療アプリケーションで疾患を引き起こす点変異(point mutation)を修復することができる。2種の塩基編集技術が開発された。CRISPRベースのアデニンとサイトシン塩基編集はさみとDddAベースの塩基編集技術である。CRISPRベースの塩基編集はさみは、DNA結合ユニットとしてcatalytically-impaired Cas9またはCas12aを構成しており、デアミナーゼ はラットまたは E.coli 由来の単一らせんDNA特異的な(single-strand DNA specific)ものを有している。一方、DdCBEのような場合、TALEDNA結合アレイと二重らせんDNA特異的(double-strand DNA specific)DddAtoxを構成している。

【０３５５】

DdCBEと比較して、ZFDはサイズがより小さい。なぜなら、ZFDに含まれるジンクフィンガータンパク質のサイズが小さいのに対し、DdCBEに含まれたTALEアレイはサイズが大きいからである。その結果、小さな搭載空間を有しているAAVベクターにZFD対をエンコーディングしている遺伝子は容易に搭載できることに対し、DdCBE対をエンコードする遺伝子はそうではない。さらに、簡素なZFP特性により、エンジニアリングが容易である。半分のDddAtoxをZFPのN-末端またはC-末端部分に融合できるので、ZFP結合部位の上流あるいは下流に機能できるように作製できる。さらに、ZFD組換えタンパク質は、電気穿孔またはリポフェクションなしで自発的にヒトの細胞に貫通して入れる。これは、gene-free遺伝子治療を可能にする。ZFD対またはZFD/DdCBE混合対は独自の変異パターンを作ることができる。これは、DdCBEのみを使用しては得られないターンであった。これらの特性は、ZFDがミトコンドリア疾患のモデリングおよび治療に新しい強力なプラットフォームになるようにする。

【０３５６】

実施例2．TALE-DdCBE植物葉緑体、ミトコンドリア遺伝子編集

ミトコンドリアと葉緑体を含む植物小器官には、それぞれ呼吸と光合成に必須的な多くの遺伝子を暗号化する自体ゲノムがある。植物遺伝学と生命工学に対する満たされていない条件である植物小器官の遺伝子の編集は、このような細胞器官のDNAを標的化するための適切な道具の不足により制限されてきた。DddA－由来シトシン塩基編集プラスミド(DdCBE)を組み立てるために、16個の発現プラスミド(生成タンパク質をミトコンドリアに伝達するための8個および葉緑体に伝達のための他の8個)と424個のTALEサブアレイプラスミドからなるゴールデンゲートクローニングシステムを開発し、ミトコンドリアと葉緑体に効率的に増進された点突然変異を作るために完成されたDdCBEプラスミドを使用した。DdCBE塩基編集は、レタスまたは菜種のカリにおいて最大25%(ミトコンドリア)と38%(葉緑体)の効率を誘導した。DdCBEを暗号化するプラスミドにおいて原因となる非標的変異を避けるためにDdCBEメッセンジャーRNAをレタスプロトプラストに伝達し、DNAなしで塩基編集を葉緑体で示した。さらに、葉緑体16S rRNA遺伝子の点突然変異を導入して、ストレプトマイシンまたはスペクチノマイシンに耐性のある最大99%の編集効率を有するレタスカリと幼植物体を作った。

【０３５７】

DdCBEsは、バクテリアシトシンジアミナーゼ毒性物質DddAtoxに由来する分離された非毒性ドメインと特定の位置に合わせてデザインされたTALEアレイおよびUGIから構成され、ヘテロダイマーでターゲットDNAのTALEタンパク質結合位置間のスペーサー空間でシトシンをチミンに置換するシトシンデアミネーション反応を起こす機能をする。ミ

トコンドリアと葉緑体で発現するDdCBEプラスミドの結合を迅速かつ便利に示し、高効率の植物における小器官の塩基編集を立証することをDdCBEsの結果として使用した。

【０３５８】

2-1．方法

植物プロトプラストにおける発現プラスミドの作製

DdCBEゴールデンゲートデスティネーションベクターは、ギブソンアセンブリ法を用いて作製された。TAL N-ターミナルドメイン、HAタグ、FLAGタグ、TAL C-ターミナルドメイン、分割DddAtox、およびUGIをエンコードする配列は、双子葉植物(シロイヌナズナ)発現にコドン最適化し、Integrated DNA Technologyによって合成された。AtinfAとAtRbcSからのCTPとATPaseデルタサブユニットとATPaseガンマサブユニットからのMTSをエンコードする配列は、シロイヌナズナcDNAから増幅された。植物発現のために、バックボーンプラスミドにおいて、哺乳類CMVプロモーターをPcUbiプロモーターとpea3Aターミネーターに置き換えて使用した。インビトDdCBE mRNA転写のためのベクターの作製には、T7プロモーターカセットがDdCBEゴールデンゲートデスティネーションベクターでPcUbiプロモーターとDdCBEエンコーディング領域との間に導入された。

【０３５９】

TALEアレイ遺伝子は、単一のゴールデンゲートアセンブリを用いて作製された。DdCBE発現プラスミドは、424個のTALEアレイプラスミドとデスティネーションベクターを用いて、BsaI分解とT4ライゲーションのゴールデンゲートアセンブリとして作製された。単一のゴールデンゲートクローニングを以下のように行った。20繰り返しの37℃と50℃各5分、最終50℃15分反応後、80℃5分。全ての植物プロトプラスト形質導入ベクターは、プラスミドプラスミディフラップキット(キアゼン)を用いて精製された。ベクターのDNAとアミノ酸配列は以下の通りである。

【０３６０】

【表２】

Name	Sequence	No.
パセリのユビキチンプロモーター	CTAGCAACGATTGTACAATTGCTTCTTTAAAAAAGGAAGAAAGAAAGAAAGAAAAGAA TCAACATCAGCGTTAACAAACGGCCCCGTTACGGCCCAAACGGTCATATAGAGTAACG GCGTTAAGCGTTGAAAGACTCCTATCGAAATACGTAACCGCAAACGTGTCATAGTCAG ATCCCCTCTTCCTTCACCGCCTCAAACACAAAAATAATCTTCTACAGCCTATATATAC AACCCCCCCTTCTATCTCTCCTTTCTCACAATTCATCATCTTTCTTTCTCTACCCCCA ATTTTAAGAAATCCTCTCTTCTCCTCTTCATTTTCAAGGTAAATCTCTCTCTCTCTCT CTCTCTCTGTTATTCCTTGTTTTAATTAGGTATGTATTATTGCTAGTTTGTTAATCTG CTTATCTTATGTATGCCTTATGTGAATATCTTTATCTTGTTCATCTCATCCGTTTAGA AGCTATAAATTTGTTGATTTGACTGTGTATCTACACGTGGTTATGTTTATATCTAATC AGATATGAATTTCTTCATATTGTTGCGTTTGTGTGTACCAATCCGAAATCGTTGATTT TTTTCATTTAATCGTGTAGCTAATTGTACGTATACATATGGATCTACGTATCAATTGT TCATCTGTTTGTGTTTGTATGTATACAGATCTGAAAACATCACTTCTCTCATCTGATT GTGTTGTTACATACATAGATATAGATCTGTTATATCATTTTTTTTATTAATTGTGTAT ATATATATGTGCATAGATCTGGATTACATGATTGTGATTATTTACATGATTTTGTTAT TTACGTATGTATATATGTAGATCTGGACTTTTTGGAGTTGTTGACTTGATTGTATTTG TGTGTGTATATGTGTGTTCTGATCTTGATATGTTATGTATGTGCAGC	162
エンドウ3Aターミネーター	CAGGCCTCCCAGCTTTCGTCCGTATCATCGGTTTCGACAACGTTCGTCAAGTTCAATG CATCAGTTTCATTGCCCACACACCAGAATCCTACTAAGTTTGAGTATTATGGCATTGG AAAAGCTGTTTTCTTCTATCATTTGTTCTGCTTGTAATTTACTGTGTTCTTTCAGTTT TTGTTTTCGGACATCAAAATGCAAATGGATGGATAAGAGTTAATAAATGATATGGTCC TTTTGTTCATTCTCAAATTATTATTATCTGTTGTTTTTACTTTAATGGGTTGAATTTA AGTAAGAAAGGAACTAACAGTGTGATATTAAGGTGCAATGTTAGACATATAAAACAGT CTTTCACCTCTCTTTGGTTATGTCTTGAATTGGTTTGTTTCTTCACTTATCTGTGTAA TCAAGTTTACTATGAGTCTATGATCAAGTAATTATGCAATCAAGTTAAGTACAGTATA GGCTT	163
T7 promoter	TAATACGACTCACTATAGG	164

【０３６１】

ＤｄＣＢＥの構成に対する具体的なアミノ酸配列およびＴＡＬＥリピートに対する具体的なアミノ酸配列は、以下の通りである。

【表３】

Name	Sequence	No.
AtinfA CTP-3xHA-N terminal domain	MLQLCSTFRPQLLLPCQFRFTNGVLIPQINYVASNSVVNIRPMIRCQRASGGRGGANR SKPAKPQVKEGSNKTVIEGLVTESLPNGMFRVDLENGDNILGYICGKIRKNFIRILPG DKVKVEMSVYDSTKGRIIFRMSSRDYPYDVPDYAGYPYDVPDYAGYPYDVPDYAMDIA DLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVKYQ DMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIAKRGGV TAVEAVHAWRNALTGAPL	165
AtRbcS CTP-3xFLAG-N terminal domain	MASSMLSSATMVASPAQATMVAPFNGLKSSAAFPATRKANNDITSITSNGGRVNCMQV WPPIGKKKFETLSYLPDLTDSEDYKDHDGDYKDHDIDYKDDDDKMDIADLRTLGYSQQ QQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVKYQDMIAALPEAT HEAIVGVGKRGAGARALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWR NALTGAPL	166
AtATPase gamma subunit MTS-3xHA-N terminal domain	MAMAVFRREGRRLLPSIAARPIAAIRSPLSSDQEEGLLGVRSISTQVVRNRMKSVKNI QKITKAMKMVAASKLRAVQYPYDVPDYAGYPYDVPDYAGYPYDVPDYAMDIADLRTLG YSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVKYQDMIAAL PEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAV HAWRNALTGAPL	167
AtATPase delta subunit MTS-3xFLAG-N terminal domain	MFKQASRLLSRSVAAASSKSVTTRAFSTELPSTLDSDYKDHDGDYKDHDIDYKDDDDK MDIADLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVA VKYQDMIAALPEATHEAIVGVGKRGAGARALEALLTVAGELRGPPLQLDTGQLLKIAK RGGVTAVEAVHAWRNALTGAPL	168
Half (NG) domain-C-terminal domain-G1333-N-UGI	GLTPEQVVAIASNGGKQALESIVAQLSRPDPALAALTNDHLVALACLGGRPALDAVK KGLGGSGSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGGSGGSTNLS DIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLTSDAP EYKPWALVIQDSNGENKIKML*	169
Half (NG) domain-C-terminal domain-G1333-C-UGI	GLTPEQVVAIASNGGKQALESIVAQLSRPDPALAALTNDHLVALACLGGRPALDAVK KGLGGSPTPYPNYANAGHVEGQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLP ENAKMTVVPPEGAIPVKRGATGETKVFTGNSNSPKSPTKGGCSGGSTNLSDIIEKETG KQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLTSDAPEYKPWALV IQDSNGENKIKML*	170

Half (NG) domain-C-terminal domain-G1397-N-UGI	GLTPEQVVAIASNGGGKQALESIVAQLSRPDPALAALTNDHLVALACLGGRPALDAVK KGLGGSGSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGGPTPYPNYA NAGHVEGQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKMTVVPPEGSG GSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVML LTSDAPEYKPWALVIQDSNGENKIKML*	171
Half (NG) domain-C-terminal domain-G1397-C-UGI	GLTPEQVVAIASNGGGKQALESIVAQLSRPDPALAALTNDHLVALACLGGRPALDAVK KGLGGSAIPVKRGATGETKVFTGNSNSPKSPTKGGCSGGSTNLSDIIEKETGKQLVIQ ESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLTSDAPEYKPWALVIQDSNG ENKIKML*	172
16s rDNA Left TALE repeat - GGGGAGTACG TTCGCAAG	NLTPAQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNNGGKQALETVQ RLLPVLCQDH	173
	GLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNNGGKQALETVQ RLLPVLCQDH	174
	GLTPAQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQ RLLPVLCQAH	175
	GLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNIGGKQALETVQ RLLPVLCQAH	176
	GLTPAQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNNGGKQALETVQ RLLPVLCQDH	177
	GLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNGGGKQALETVQ RLLPVLCQDH	178
	GLTPAQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNNGGKQALETVQ RLLPVLCQDH	179
	GLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNIGGKQALETVQ RLLPVLCQDH	180
	GLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQ RLLPVLCQAH	181
16s rDNA Right TALE repeat - TGTGCGGGCC CCCGTCAA	NLTPAQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNNGGKQALETVQ RLLPVLCQAH	182
	GLTPAQVVAIASNGGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNNGGKQALETVQ RLLPVLCQAH	183
	GLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNNGGKQALETVQ RLLPVLCQAH	184
	GLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNNGGKQALETVQ RLLPVLCQDH	185
	GLTPDQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPAQVVAIASHDGGKQALETVQ RLLPVLCQAH	186
	GLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQ RLLPVLCQDH	187

	GLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQ RLLPVLCQDH	188
	GLTPAQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPAQVVAIASHDGGKQALETVQ RLLPVLCQDH	189
	GLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNIGGKQALETVQ RLLPVLCQDH	190
psbA Left TALE repeat - GCAACAGGAG CTGAATAT	NLTPAQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPDQVVAIASHDGGKQALETVQ RLLPVLCQDH	191
	GLTPAQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNIGGKQALETVQ RLLPVLCQDH	192
	GLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNIGGKQALETVQ RLLPVLCQAH	193
	GLTPAQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNNGGKQALETVQ RLLPVLCQDH	194
	GLTPAQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNNGGKQALETVQ RLLPVLCQAH	195
	GLTPAQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNGGGKQALETVQ RLLPVLCQAH	196
	GLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNIGGKQALETVQ RLLPVLCQDH	197
	GLTPDQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNGGGKQALETVQ RLLPVLCQAH	198
	GLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNGGGKQALETVQ RLLPVLCQAH	199
psbA Right TALE repeat - AGTTTCCGTC TGGGTATG	NLTPAQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQ RLLPVLCQAH	200
	GLTPDQVVAIASNGGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNGGGKQALETVQ RLLPVLCQDH	201
	GLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPDQVVAIASHDGGKQALETVQ RLLPVLCQAH	202
	GLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQ RLLPVLCQDH	203
	GLTPAQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPAQVVAIASHDGGKQALETVQ RLLPVLCQAH	204
	GLTPAQVVAIASNGGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQ RLLPVLCQDH	205
	GLTPAQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNNGGKQALETVQ RLLPVLCQDH	206
	GLTPDQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQ RLLPVLCQAH	207
	GLTPAQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNNGGKQALETVQ RLLPVLCQDH	208

psbB Left TALE repeat - AAATGAATTC CAAAAATC	NLTPDQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNIGGKQALETVQ RLLPVLCQDH	209
	GLTPAQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNGGGKQALETVQ RLLPVLCQAH	210
	GLTPAQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNIGGKQALETVQ RLLPVLCQDH	211
	GLTPAQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNGGGKQALETVQ RLLPVLCQAH	212
	GLTPDQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPAQVVAIASHDGGKQALETVQ RLLPVLCQDH	213
	GLTPDQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNIGGKQALETVQ RLLPVLCQAH	214
	GLTPDQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNIGGKQALETVQ RLLPVLCQDH	215
	GLTPAQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNIGGKQALETVQ RLLPVLCQDH	216
	GLTPAQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPAQVVAIASHDGGKQALETVQ RLLPVLCQDH	217
psbB Right TALE repeat - GAACGTACAG GAAAACC	NLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNIGGKQALETVQ RLLPVLCQDH	218
	GLTPDQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPDQVVAIASHDGGKQALETVQ RLLPVLCQDH	219
	GLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNGGGKQALETVQ RLLPVLCQAH	220
	GLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQ RLLPVLCQDH	221
	GLTPAQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNNGGKQALETVQ RLLPVLCQAH	222
	GLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNIGGKQALETVQ RLLPVLCQAH	223
	GLTPDQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNIGGKQALETVQ RLLPVLCQDH	224
	GLTPAQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPAQVVAIASHDGGKQALETVQ RLLPVLCQDH	225
	GLTPAQVVAIASHDGGKQALETVQRLLPVLCQAH	226
ATP6 (lettuce, rapeseed) Left TALE repeat -	NLTPDQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNIGGKQALETVQ RLLPVLCQDH	227
	GLTPDQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQ RLLPVLCQAH	228
	GLTPAQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQ RLLPVLCQDH	229

TATTGGCATTACTATAG	GLTPAQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAH	230
	GLTPDQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPAQVVAIASHDGGKQALETVQRLLPVLCQDH	231
	GLTPDQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDH	232
	GLTPDQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDH	233
	GLTPAQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNNGGKQALETVQRLLPVLCQDH	234
	GLTPDQVVAIASNNGGKQALETVQRLLPVLCQDH	235
ATP6 (lettuce) Right TALE repeat - GAAAAAATGAAGCCCA	NLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNIGGKQALETVQRLLPVLCQDH	236
	GLTPAQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQAH	237
	GLTPAQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNIGGKQALETVQRLLPVLCQAH	238
	GLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNGGGKQALETVQRLLPVLCQAH	239
	GLTPAQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNIGGKQALETVQRLLPVLCQDH	240
	GLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQAH	241
	GLTPAQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDH	242
	GLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQAH	243
ATP6 (rapeseed) Right TALE repeat - TAAAAAATGAAGCCCA	NLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDH	244
	GLTPDQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQAH	245
	GLTPAQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNIGGKQALETVQRLLPVLCQAH	246
	GLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNGGGKQALETVQRLLPVLCQAH	247
	GLTPAQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNIGGKQALETVQRLLPVLCQDH	248
	GLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQAH	249
	GLTPAQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPAQVVAIASHDGGKQALETVQRLLPVLCQAH	250

	GLTPAQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQAH	251
RPS14 Left TALE repeat - CGTTCTGTATCTGAG	NLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDH	252
	GLTPAQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAH	253
	GLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNGGGKQALETVQRLLPVLCQAH	254
	GLTPDQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNGGGKQALETVQRLLPVLCQDHG	255
	LTPAQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAH	256
	GLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNGGGKQALETVQRLLPVLCQAH	257
	GLTPAQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNIGGKQALETVQRLLPVLCQAH	258
	GLTPDQVVAIASNNGGKQALETVQRLLPVLCQAH	259
RPS14 Right TALE repeat - AATCCACGAAAAACGA	NLTPDQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDH	260
	GLTPDQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPAQVVAIASHDGGKQALETVQRLLPVLCQDH	261
	GLTPDQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQAH	262
	GLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNNGGKQALETVQRLLPVLCQAH	263
	GLTPAQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQAH	264
	GLTPAQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNIGGKQALETVQRLLPVLCQAH	265
	GLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQAH	266
	GLTPAQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQAH	267

【０３６２】

mRNAインビトロ転写

DdCBE DNAテンプレートは、フュージョンDNAポリメラーゼを用いたPCRによって調製された(Thermo Fisher)。DdCBE mRNAは、インビトmRNA合成キットを用いて合成精製した(Enzynomics)。

【０３６３】

プロトプラストの抽出と形質導入

レタスの種子を70％エタノールに30秒および0.4％ラックス溶液に15分間表面滅菌した後、滅菌蒸留水で3回すすぐ。レタス種子は、0.5倍のMS培地と2％スクロースが含ま

れた培地に16時間の明条件、8時間の暗条件下に25℃で播種する。菜種種子は、70%エタノールに3分および1.0%ラックス溶液に30分表面殺菌した後、滅菌蒸留水で3回すすぐ。菜種種子は、1倍MS培地と3%スクロースが含まれた培地に16時間の明条件8時間の暗条件下に25℃で播種する。

【0364】

プロトプラスト抽出と形質導入は、従来の研究に従って行う。7日となったレタスと14日となった菜種のもち葉を、酵素溶液に揺れる(40rpm)暗条件で3時間処理する。プロトプラスト－酵素混合物は、同じ量のW5溶液を用いて洗浄した後、無傷のプロトプラストを80gで7分間遠心分離してスクロース溶液から得る。プロトプラストはW5溶液で4℃に1時間処理後、ポリエチレングリコールを用いた形質導入を行う。

【0365】

レタスプロトプラストと菜種プロトプラストは、MMG溶液に再懸濁した後、プラスミドまたはmRNAをPEGを用いた形質導入を進行し、20分間常温で培養する。PEG－原形質体混合物は、同じ量のW5溶液を用いて3回すすぐが、毎回ゆっくりと反転しながらその後10分間培養する。プロトプラストはその後100gで5分間遠心分離してペレット状にする。

【0366】

プロトプラストの培養

レタスプロトプラストは、DdCBEをエンコードするプラスミドが導入された後、レタスプロトプラスト培養培地(LPCM)によって再浮遊した。培地を含むプロトプラストは、培地とローメルティングアガロース2.4%が含まれた培地と1：1で混合した後、直ちに6ウェル皿に載せる。その後、混合物が個体化したら、埋め込まれたプロトプラスト上に1mlの液体培地を載せた後、25 ℃の暗条件で1週間培養する。初期培養後、上に載せた液体培地を新鮮な培地に毎週交換し、埋め込まれたプロトプラストは16時間の弱い明条件、8時間の暗条件で1週間培養した後、16時間の明条件、8時間の暗条件で2週間培養する。プロトプラストから誘導された小さなカルスは、再分化培地に4週間、25℃の16時間の明条件、8時間の暗条件で培養する。塩基編集効率の分析を準備する際には、プロトプラストは埋め込みなしで液体培地中で25℃の暗条件で1週間培養した。抗生剤耐性を実験するために、1ヵ月埋め込まれた小さなカルスは50mg/L濃度のストレプトマイシンまたは50mg/L濃度のスペクチノマイシンが含まれた再分化培地で25℃、16時間の明条件、8時間の暗条件で4週間培養した。4週間後、抗生剤耐性緑色カルスまたは不定芽は、200mg/L濃度のストレプトマイシンまたは50mg/L濃度のスペクチノマイシンが含まれた新しい再分化培地に移した。

【0367】

DdCBEをエンコードするプラスミドを導入された菜種プロトプラストは、菜種培養培地に再浮遊した。プロトプラスト－培地混合物は、6ウェル皿に移した後、25℃の暗条件で2週間培養する。2週間後、プロトプラストは16時間の弱い明条件、8時間の暗条件で3週間培養した。培地は非常に新鮮な培地に交換した。

【0368】

DNAとRNAの抽出

液体培地で培養された細胞や形質転換されたカルスの総DNAまたはRNAは、ジエンイジ植物ミニキットまたはアルエンイジ植物ミニキットを用いて抽出した。培養した細胞やカルスを10,000rpmで1分間遠心分離して収穫した。総RNAからのcDNAは、RNA to cDNA EcoDry Premix(TAKARA)を用いて逆転写された。

【0369】

ディープシーケンシング

ターゲット領域は、フュージョン酵素と適切なプライマーを用いて増幅した(サプテーブル1) DNA配列分析ライブラリーを作製するために、3ラウンドのPCR(第一、 nested PCR、第二、PCR、第三、インデクシング PCR)を行った。同量のDNAを集めた後、MiniSeq(イルミナ)装置を用いて配列分析した。ペアインドシーケンスファイルは、Cas-a

nalyzerとコンピュータプログラムのソースコードによって分析された。

【０３７０】

2-2．結果

葉緑体をターゲットとするDdCBE(cp-DdCBE)とミトコンドリアをターゲットとするDdCBE(mt-DdCBE)を作製するために、ゴールデンゲートアセンブリシステムを開発した(図19)。発現プラスミドは、結合されたタンパク質で、葉緑体伝達ペプチドまたはミトコンドリアターゲティング配列、TALEのN-またはC-末端ドメイン、分割DddAtox(G1333N、G1333C、G1397N、およびG1397C)およびUGIを暗号化し、これは、双子葉植物の発現に最適化されており、それは、パセリユビキチン(PcUbi)プロモーターとpea3Aターミネーターで調節される。カスタマイズされたTALEDNA結合配列とDdCBEプラスミドは、発現用ベクターと6つのTALEサブアレイプラスミドをE-tubeに混合して単一のサブクローニング段階にすることができる。総424個(6X64 3 配列を認識+2X16 2配列を認識+2X41配列を認識) モジュール型TALEサブアレイプラスミドは、5'の末に保存されたTを含んで16-20塩基配列の長さを認識するcp-DdCBEとmt-DdCBEを作製できる。結果的に、機能的に DdCBEヘテロダイマは、32から40 bp DNA塩基配列を認識する。

【０３７１】

DdCBEsが葉緑体塩基編集に促進できるか否かを知るために、30Sリボソームサブユニットの16S rRNA化合物の暗号化する葉緑体16S rRNA遺伝子に適したcp-DdCBEsプラスミド4対を作製し、レタスと菜種プロトプラストにそれぞれの対を一緒に形質注入した後、7日後、ディープシーケンシングにより塩基編集効率を測定した(図20a、b)。最も高効率を示したcp-DdCBE対は、(Left-G1397-N+Right-G1397-C)2つのTALE結合部位間の15bpのスペーサー領域にレタスプロトプラストでは30%、菜種プロトプラストでは15%の効率でC・GからT・Aへの置換を誘導した(図20b)。従来哺乳類細胞およびマウスの結果と同様に、cp-DdCBEによって5'-TCモチフであるシトシン(C9、C13)はチミンに優先的に置換された。興味深いことに、5'-ACモチーフであるシトシン(C7)は、レタスプロトプラストにおいて、他のcp-DdCBE(Left-G1333-N+Right-G1333-C)によって4.2%の効率でチミンに交替された。さらに、レタスプロトプラストにおいて、cp-DdCBEによる塩基編集の14日間の培養過程で持続することを調べた(図24)。編集効率は最大10日まで持続的に増加し、培養期間中維持された。

【０３７２】

光系IIにおいて、各々D1とCP-47光合成タンパク質暗号化する葉緑体遺伝子であるpsbAとpsbBについて、2つの追加的な実験を行った。(図20c、dおよび図25)。Ps遺伝子をターゲットとするcp-DdCBEのうち、最も効率の高いものは(Left-G1397-C+Right-G1397-N)レタスプロトプラストにおいて、C・GからT・Aへの置換を最大25%まで誘導できた(図20d)。この塩基編集はさみによって、もっぱら5'-TCCの2つのシトシン(C11、C12)のみが効率的にチミンに置換された。これは、5'-TCCが最初の5'-TTCに置換された後、次いで5'-TTTになる可能性がある。菜種プロトプラストにおいて、他の組合(Left-G1333-N+Right-G1333-C)が4つのシトシン位置(C3、C4、C11、C12)で最も高い効率を示し、最大3.5%(C3)まで効率を示した。菜種の遺伝子にはC3とC4は5'-TCC配列であるが、レタスでは比較的一塩基多型のために5'-ACC配列を有しており、DdCBEにより菜種遺伝子では2つのシトシン(C3、C4)で効率的な編集効率の原因となるが、レタス遺伝子ではそうではない。さらに、菜種プロトプラストのpsbB遺伝子をターゲットとするcp-DdCBEの組合は、TCC配列を有する2つのシトシンを0.36%から4.1%の効率で置換した(図25)。まとめると、かかる結果は、DddAtoxの分割された位置(G1333 vs. G1397)と方向(Left-G1333-N vs. Left-G1333-C)を含むシトシン位置と配列によって編集効率が定められ、cp-DdCBEは、植物の葉緑体遺伝子に対して効率的な塩基編集が可能であることを立証した。

【０３７３】

次に、カスタマイズされたmt-DdCBEを用いて植物のミトコンドリアDNAの塩基編集

を達成しようとした。そのために、レタスと菜種のatp6遺伝子をターゲットとし、菜種のrps14遺伝子をターゲットとするmt-DdCBEをエンコードするプラスミド(ゴールデンゲートクローニングシステムを利用)を作製し、プラスミドをレタスと菜種のプロトプラストに導入した後、導入7日後、ディープシークエンシングにより塩基編集効率を測定した(図20e、f、図26)。最も効率の高いmt-DdCBEの組合は(レタスではLeft-G1397-N+Right-G1397-C、菜種ではLeft-G1397-C+Right-G1397-N)atp6遺伝子ターゲット位置でレタスプロトプラストは23%、プロトプラストは23%のC・GをT・Aに置換を変化させた(図20)。さらに、mt-DdCBEの組合は、rps14ターゲット位置で菜種プロトプラストから11%のC？GをT？Aへ置換を誘導した。これらの結果は、mt-DdCBEが植物ミトコンドリアDNAに対して塩基編集が良好に行われることを示す。

【０３７４】

DdCBEによるcpDNAとmtDNAの編集が再分化の過程中に維持されるかどうかを調べるために、注入後4週が経過してDdCBE処理されたプロトプラストから再分化されたレタスと菜種カルスを収集し(図21ａ)、ディープシークエンシングとサンガーシークエンシングを用いて各々のカルスの塩基編集効率を測定した(図21b、図27)。DdCBEによって塩基編集が誘導された葉緑体またはミトコンドリアの遺伝子は、26個のうち22個のレタスカルスと14個のうち7個の菜種カルスがそれぞれ最大38%と25%の効率で測定された(図21c)。さらに、葉緑体psbA遺伝子の塩基編集は、レタスカルスにおいて最大3.9%の効率を示した(図27)。さらに、菜種カルスにおけるミトコンドリア塩基編集は、atp6とrps14でそれぞれ最大25%と1.9%と測定された(図27)。これらの結果は、植物プロトプラストにおけるDdCBEの発現に耐えることができ、プロトプラストへの再分化の過程でDdCBEによって小器官の塩基編集を誘導した。

【０３７５】

次に、プラスミドではなくインビトロ転写されたcp-DdCBE mRNAを用いて小器官でDNAフリー塩基編集を立証しようとした。レタスで16S rRNA遺伝子をターゲットとするcp-DdCBEをエンコードするインビトロ転写物をプロトプラストに注入した後、ターゲット位置の塩基編集効率を分析した(図21a)。プロトプラスト中のC-to-T変異は最大25%まで測定された(図21d 、図28)。予想の通り、DdCBE mRNAとDNA配列をプロトプラストに導入した7日後に消失した(図29)。この方法は、ホストゲノム中にプラスミドDNAの一部が潜在的に組み込まれるのを避けることができると言える。

【０３７６】

プロトプラストから再分化したカルスで小器官の編集が安定的に維持されることに助けを受け、葉緑体DNAで16S rRNA遺伝子の編集を通じて16S rRNAに不可逆的に結合してタンパク質の合成を阻害するストレプトマイシンとスペクチノマイシン抗生剤耐性を調査した。16S rRNA遺伝子の数カ所の一塩基多型は、原核生物および真核生物においてストレプトマイシン耐性が一般的に観察され、特に、16S rRNA C860T(大腸菌はC912位)の突然変異はタバコにおいてストレプトマイシン耐性を有するようにする。タバコのC860T点突然変異はレタスのC9位と同一である(図20a、21b、22b、22d)。DdCBE処理されたプロトプラストから再分化して得られたレタスカルスをストレプトマイシンとスペクチノマイシンが添加された培地に移した。モック(Mock)処理群は、抗生剤にさらされたとき白く変化し、これはカルスのプロトプラスト機能障害を示す。一方、DdCBE処理されたカルスは緑色を維持し、これらの抗生剤に耐性を示す。耐性を有するレタスカリと幼植物体におけるDdCBEによる編集効率を分析した。C860T変異のような、C9位のCにおけるT置換は、薬物処理から得られたカルスとシュートで最大98.6%観察された(図21e、f)。興味深いことに、近くのC13位のCでT編集がスペクニオマイシンのない場合は最大20%の効率を示したが、抗生剤がある場合は見られなかったが、薬物処理によってこの変異が選抜されたことを立証する。まとめると、これらの結果は、プロトプラスト中のDdCBEによって植物小器官の変異が誘導されたことは、細胞分裂と植物発達の後にも維持することができ、葉緑体の編集が薬物選抜でホモプラスミを達成することができると言える。

【０３７７】
さらに、16S rRNA位置をターゲットとするTALEジアミナーゼのオフターゲットをプロトプラスト、カルス、スーツで分析した。単一細胞に由来し、抗生剤耐性を有するカルスとスーツにおいて、ターゲットサイトのターゲット周辺(両側50塩基対)(図31)または葉緑体遺伝子で配列が類似することに基づいて選択した5つの潜在的オフターゲットの位置(図32)。オフターゲットが見えないことが測定された。一方、DdCBEをエンコードするプラスミドをプロトプラストに導入したとき、TCがTTに変化するオフターゲットが、5つの潜在的オフターゲットのうち3箇所で1.2%から4.1%の低効率で誘導された。TALEジアミナーゼをエンコードするプラスミドの代わりにインビトロ転写物(mRNA)を使用した場合、プロトプラストにおけるオフターゲットの効率が大幅に減少した(図22)。これらの結果からみると、過発現または長期的なプラスミド由来のDdCBEの発現は、オフターゲットの変異を増加させる可能性があり、一時的な、mRNAを使用するmRNA由来の発現は、オフターゲットの塩基編集を回避するのに好ましい。
【０３７８】
要約すると、424TALEサブアレイプラスミドと16個の発現プラスミドを使用するゴールデンゲートクローニングシステムを開発して、植物中の小器官の塩基編集を行うためにDdCBEをエンコードするプラスミドを組み立てることができる。葉緑体DNAの3遺伝子、ミトコンドリアDNA2遺伝子をターゲットするカスタマイズされたDdCBEは、レタスおよび菜種プロトプラストにおいて高効率のCからのT置換を達成した。特に、植物の小器官編集を細胞分裂と植物発達での編集が維持された。さらに、葉緑体16S rRNA遺伝子の変異により、ほぼホモプラスミ(99%)の抗生剤耐性レタスカルスと植物体を確保した。抗生剤選抜なしでは、ミトコンドリアで25%、葉緑体で38%の編集効率を示した。ゴールデンゲートクローニングシステムが植物の小器官DNA編集に価値のある資源になると期待する。
【０３７９】
実施例3．TALE?DdCBE動物DNA編集
分割されたバクテリア毒素DddAtox、DNAへの結合を可能にするように考案されたTALEおよびUGIを含むDddA由来シトシン塩基編集器(DdCBEs)は、目標としたシトシン－チミン塩基変形をミトコンドリアDNAで可能にした。マウス胚において高効率のミトコンドリアDNA編集が可能であることを示した。ミトコンドリア遺伝子のうち、NADH脱水およびユビキノンへの電子伝達を触媒するNADH脱水素化酵素のサブユニットをコードするMT-ND5(ND5)を目指したが、これは、m.G12918Aのようにヒトのミトコンドリア疾患に関連した突然変異と、m.C12336Tのような早期終止コドンを生成する突然変異を含めた。これを通して、ミトコンドリア疾患モデルをマウスにおいて作製でき、ミトコンドリア疾患を治療できる可能性を提示した。
【０３８０】
3-1．方法
プラスミドの組み立て。TALEN(transcription activator-like effector Nucleases)システムを、DddA半体が含まれた発現プラスミドおよび最終TALE-DddAtox構造を作製するために借用した。TALENシステムの発現プラスミドにおいて、核内輸送信号とFokI二量体中、モノマーをミトコンドリア伝達信号(MTS)、DddA脱アミノ化酵素半体、およびウラシルグリコシル化阻害剤(UGI)に置換した。MTS、DddA、およびUGIをエンコードする配列はIDTで合成した。発現ベクターを作製するために、Gibson assemblyに必要なDNA断片をQ5 DNAポリメラーゼ(NEB)を用いて増幅し、その後に精製した。精製された遺伝子断片は、HiFi DNA assembly kit (NEB)を用いて組み立て、E coli DH5a(Enzynomics)に化学的形質転換させた後、サンガーシーケンシングにより確認した。したがって、8つの異なる発現プラスミドを得たが、これには、ゴールデンゲートクローニングのためのBsaI制限酵素部位が、N-末端ドメインとC-末端ドメインをエンコードする配列の間にある。DdCBEプラスミドの組み立てのために、表現プラスミドにモジュールベクター(それぞれTALE配列をコードする)、BsaI-HFv2(10U)、T4 D

NAリガーゼ(200 U)、および反応バッファを1つのチューブに収めた。以後、制限酵素及びライゲイズ反応はサーモサイクラーにて37℃5分、50℃5分の20回繰り返した後、50℃を15分、80 ℃を5分間処理した。接合されたプラスミドは、化学的形質転換を通じてE coli DH5aに導入され、最終構造体は、サンガーシーケンシングにより確認された。細胞株の導入のために、プラスミドはミディプレップした。

【０３８１】

哺乳類細胞株の培養及びトランスフェクション。NIH3T3(CRL-1658、American Type Culture Collection(ATCC))細胞株は37℃の5％CO_2環境で培養された。細胞株は10％(v/v)エレガントな血清強化したDMEM(Gibco)の培地で抗生剤なしで生長し、マイコプラズマ試験は行われなかった。リポフェクションのために、細胞は、12ウェル細胞培養プレート(SPL, Seoul, Korea)に1．5×10^4細胞密度でトランスフェクションの18～24時間前に生長を開始した。Lipofectamine 3000(Invitrogen)を用いて、各々のDdCBE分割体を500ng使用して総1,000ngのプラスミドDNAを導入した。細胞をトランスフェクションの4日後に回収した。

【０３８２】

mRNAの作製。mRNAの鋳型は、Q5 DNAポリメラーゼ(NEB)を用いてPCR増幅し、以下のプライマーを使用した(F: 5′-CATCAA TGGGCGTGGATAG-3′SEQ ID No. 268, R: 5′-GACACCTACTCAGACAATGC-3 SEQ ID No. 269)。DdCBE mRNAは、in vitro RNA transcription kit (mMESSAGE mMACHINE T7 Ultra kit, Ambion)を用いて合成した後、MEGAclear kit(Ambion)を用いて精製した。

【０３８３】

動物。ラットを含む実験は、すべて基礎科学研究院の動物管理および使用委員会の承認を得て行われた。過排卵したC57BL/6JメスをC57BL/6Jオスと交配させ、ICRストレインメスを代理母として使用した。マウスは、特定の病原体のない施設で12時間の昼夜サイクルを維持する条件、恒温および恒湿条件下で飼育した(20～26℃、40～60％)。

【０３８４】

マウスコンジュゲートに微小注入。微小注入直前の過程である過排卵と胚採集、微小注入は前回の論文に記載されているとおり進行した。微小注入のために、left DdCBE mRNA(300 ng /μl)およびright DdCBE mRNA(300 ng/μl)を含む混合物はDEPC処理された注入バッファ(0.25 mM EDTA、10 mM Tris、pH 7.4)に希釈し、コンジュゲート細胞質にNikon ECLIPSE Tiマイクロマニピュレーターとa FemtoJet 4iマイクロインジェクター(Eppendorf)を用いて注入した。微小注入後、胚をKSOM + AA(Millipore)マイクロドロップに入れ、37℃、5％ CO_2条件で4日間培養した。2細胞期の胚は、0.5-d.p.c pseudo-pregnant代理母の卵管に移植された。

【０３８５】

遺伝子型の分析。胚盤胞段階の胚および組織は分解バッファ(25mM NaOH、0.2mM EDTA、pH10)に入れ、95℃で20分間培養した後、HEPES(free acids、pH調節なし)を用いて最終濃度が50mMとなるようにpHを7.4に調節した。マウス子の遺伝体DNA分離のためには、DNeasy Blood & Tissue Kits(Qiagen)を用い、サンガーシーケンシング及びターゲットディープシークエンシングで分析した。

【０３８６】

高処理配列分析のためのミトコンドリアDNA分離。12ウェルプレートに培養されたNIH3T3細胞のミトコンドリアを分離するために、細胞培養培地を除去した後、200μlのMitochondrial isolation buffer A(ScienCell)を培養プレートに入れた。細胞をセルリフターを用いて掻き取った後、マイクロチューブに入れて使い捨てすりこぎを用いて細胞を砕いた。15回の粉砕後、よく砕かれたホモジネートを1,000×g、4℃で5分間遠心分離した。上清液を新しいマイクロチューブに入れ、10,000× g、4℃で20分間遠心分離した。沈殿物を20μlの溶解液(25mM NaOH, 0.2mM EDTA, pH10)に解放した後、95℃で20分間煮込んだ。pHを下げるために、2μlの1M HEPES(free acids、without pH adjustment)をミトコンドリア溶解液に加えた。このように調製された

溶液の1μlを、高処理配列分析のためのPCR鋳型鎖として使用した。

【０３８７】

高処理配列分析。ディープシーケンシングライブラリーを作製するために、Q5 DNAポリメラーゼを用いてネステッド1次PCRおよび2次PCRを行い、最終インデックス配列を追加した。ライブラリーは、MiniSeq(Illumina)を用いてフェアドエンドリード配列分析に使用した。全体のミトコンドリアゲノム分析のために、分離されたミトコンドリアDNAをtagmentation DNA prep kit(Illumina)を用いて製造社のプロトコルに従って調製した。全ての分析において、フェアエンド配列配列の結果は、1つのfastqjoinファイルを用いて合わせ、CRISPR RGEN Tools(http://www.rgenome.net/)を用いて分析した。

【０３８８】

データの分析と表示。Microsoft Excel(2019)及びPowerPoint(2019)は、絵、グラフ、及び表の作成に使用された。ゲノム配列整列、プライマーの作製およびクローニングの考案には、Geneious(version 2021.0.1)およびSnapgene 5.2.3を使用し、基準配列としてNC_005089を使用した。

【０３８９】

3-2．結果

DdCBEプラスミドの組み立て。DdCBEでカストマイズされたTALE配列の組み立てをより容易にするために、分割されたDddAtoxの半分をコードする発現プラスミドを構成し、総424個(6×64個の3個の配列認識＋2×16個の2個の配列認識＋2×4個の1個の配列認識)を用いるゴールデンゲートクローニングシステムを使用した(図33a)。以下の表4に示すように、6つのTALEプラスミドと発現プラスミドを同じチューブに混ぜてすぐに使用できるDdCBEプラスミドを作製し、15.5～18.5個のrepeat variable diresidue配列を有するようにした(図38)。

【０３９０】

【表４】

	Left TALE Target Sequence	Right TALE Target Sequence
ND5 Silent Muation	5'-T TTTCCTACTGGTCCGAT T-3' SEQ ID NO: 277	5'-T TAAAATAAAGTTATTAT T-3' SEQ ID NO: 280
ND5 G12918A Mutation	5'-T TGCAGGTATTAATTGCT T-3' SEQ ID NO: 278	5'-T TCCTAACAGGGTTCTAC T-3' SEQ ID NO: 281
ND5 Nonsense Mutation	5'-T TCCCTAAACATAAACTCA T-3' SEQ ID NO: 279	5'-T TGTTGTTGGAGAATA T-3' SEQ ID NO: 282

【０３９１】

DdCBE構成の配列は、以下の表5の通りである。その結果、DdCBEは、5'末端の保存されたチミン配列を含み17～20個のDNA配列を認識する。したがって、機能を備えたDdCBE対は、総32個から40個のDNA塩基配列を認識する。

【０３９２】

【表５】

Left-SOD2 MTS-3×HA-N terminal domain
MALSRAVCGTSRQLAPVLGYLGSRQKHSLPDYPYDVPDYAGYPYDVPDYAGYPY
DVPDYAGIRIQDLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQ
HPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQ
LDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPL SEQ ID No: 283

Right-COX8A MTS-3×FLAG-N terminal domain
MASVLTPLLLRGLTGSARRLPVPRAKIHSLDYKDHDGDYKDHDIDYKDDDDKGIRIQ
DLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVK
YQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIAKR
GGVTAVEAVHAWRNALTGAPL SEQ ID No: 284

C terminal half domain (NG)-G1333-N-UGI
GLTPEQVVAIASNGGGKQALESIVAQLSRPDPALAALTNDHLVALACLGGRPALDA
VKKGLGGSGSGSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGGS
GGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVML
LTSDAPEYKPWALVIQDSNGENKIKML SEQ ID No: 285

C terminal half domain (NG)-G1333-C-UGI
GLTPEQVVAIASNGGGKQALESIVAQLSRPDPALAALTNDHLVALACLGGRPALDA
VKKGLGGSGSPTPYPNYANAGHVEGQSALFMRDNGISEGLVFHNNPEGTCGFCV
NMTETLLPENAKMTVVPPEGAIPVKRGATGETKVFTGNSNSPKSPTKGGCSGGST
NLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLTSD
APEYKPWALVIQDSNGENKIKML SEQ ID No: 286

C terminal half domain (NG)-G1397-N-UGI
GLTPEQVVAIASNGGGKQALESIVAQLSRPDPALAALTNDHLVALACLGGRPALDA
VKKGLGGSGSGSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGGP
TPYPNYANAGHVEGQSALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAK
MTVVPPEGSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAY
DESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKML SEQ ID No: 287

C terminal half domain (NG)-G1397-C-UGI
GLTPEQVVAIASNGGGKQALESIVAQLSRPDPALAALTNDHLVALACLGGRPALDA
VKKGLGGSGSAIPVKRGATGETKVFTGNSNSPKSPTKGGCSGGSTNLSDIIEKETG
KQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLTSDAPEYKPWALV
IQDSNGENKIKML SEQ ID No: 288

ND5 silent mutation Left TALE repeat
NLTPDQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNGGGKQALET
VQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPDQVVAI
ASHDGGKQALETVQRLLPVLCQAHGLTPAQVVAIASHDGGKQALETVQRLLPVLC
QDHGLTPAQVVAIASNGGGKQALETVQRLLPVLCQDHGLTPEQVVAIASNIGGKQA
LETVQRLLPVLCQAHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPDQV
VAIASNGGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNNGGKQALETVQRLLPV
LCQDHGLTPAQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNGGG
KQALETVQRLLPVLCQAHGLTPAQVVAIASHDGGKQALETVQRLLPVLCQDHGLTP

DQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRL
LPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNGG
GKQALETVQRLLPVLCQD SEQ ID No: 289

ND5 silent mutation Right TALE repeat
NLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNIGGKQALETV
QRLLPVLCQDHGLTPAQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPDQVVAIAS
NIGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDH
GLTPDQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNIGGKQALET
VQRLLPVLCQAHGLTPAQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPAQVVAIA
SNIGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNNGGKQALETVQRLLPVLCQD
HGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNGGGKQALE
TVQRLLPVLCQDHGLTPAQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPDQVVAI
ASNGGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNGGGKQALETVQRLLPVLC
QAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNGGGKQA
LETVQRLLPVLCQD SEQ ID No: 290

ND5 point mutation Left TALE repeat
NLTPAQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNNGGKQALET
VQRLLPVLCQAHGLTPAQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPAQVVAI
ASNIGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNNGGKQALETVQRLLPVLCQ
DHGLTPAQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNGGGKQAL
ETVQRLLPVLCQAHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVV
AIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVL
CQAHGLTPAQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNIGGKQA
LETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPAQV
VAIASNGGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNNGGKQALETVQRLLPV
LCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNGGG
KQALETVQRLLPVLCQD SEQ ID No: 291

ND5 point mutation Right TALE repeat
NLTPAQVVAIASNGGGKQALETVQRLLPVLCQDHGLTPAQVVAIASHDGGKQALET
VQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAI
ASNGGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNIGGKQALETVQRLLPVLCQ
DHGLTPAQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQAL
ETVQRLLPVLCQAHGLTPAQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPAQVV
AIASNNGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNNGGKQALETVQRLLPVL
CQDHGLTPAQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNGGGK
QALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPD
QVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNGGGKQALETVQRLL
PVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPAQVVAIASHDG
GKQALETVQRLLPVLCQD SEQ ID No: 292

ND5 nonsense mutation Left TALE repeat
NLTPAQVVAIASNGGGKQALETVQRLLPVLCQDHGLTPAQVVAIASHDGGKQALET
VQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPDQVVAI
ASHDGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNGGGKQALETVQRLLPVLC
QDHGLTPAQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQAL
ETVQRLLPVLCQAHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPAQVV

AIASHDGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNIGGKQALETVQRLLPVLC
QDHGLTPAQVVAIASNGGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQA
LETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPDQV
VAIASNIGGKQALETVQRLLPVLCQAHGLTPAQVVAIASHDGGKQALETVQRLLPVL
CQAHGLTPAQVVAIASNGGGKQALETVQRLLPVLCQDHGLTPEQVVAIASHDGGK
QALETVQRLLPVLCQAHGLTPAQVVAIASNIGGKQALETVQRLLPVLCQD SEQ ID No: 293

ND5 nonsense mutation Right TALE repeat
NLTPAQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNNGGKQALET
VQRLLPVLCQAHGLTPAQVVAIASNGGGKQALETVQRLLPVLCQDHGLTPAQVVAI
ASNGGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNNGGKQALETVQRLLPVLC
QAHGLTPAQVVAIASNGGGKQALETVQRLLPVLCQDHGLTPEQVVAIASNGGGKQ
ALETVQRLLPVLCQAHGLTPAQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQ
VVAIASNNGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNIGGKQALETVQRLLPV
LCQDHGLTPAQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNIGGK
QALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPD
QVVAIASNGGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLL
PVLCQA SEQ ID No: 294

【０３９３】

インビトロでのミトコンドリア塩基編集。ゴールデンゲートクローニングシステムを用いてインビボミトコンドリアDNA編集を試みるために、Mus musculusミトコンドリアNADH－ユビキノン酸化還元酵素鎖5タンパク質をエンコードするND5遺伝子を選定した。ND5タンパク質は、NADH脱水素酵素(ユビキノン)の核心サブユニットであり、NADHから呼吸鎖に電子を伝達することを触媒する。ヒトにおいて、ND5遺伝子の突然変異は、 MELAS(mitochondrial encephalomyopathy, lactic acidosis, and stroke-like episodes)と関連があると知られており、Leigh症候群またはLHON(Leber's hereditary optic neuropathy)の一部の症状とも関連があると知られている。ヒトの機能不全を模倣するために、ミトコンドリア遺伝子に遺伝的変形のあるマウスモデルを作製しようとした。

【０３９４】

まず、いくつかの種類のDdCBEプラスミドを組み立てたが、これは、2つのサイレント突然変異であるm.C12539Tとm.G12542Aを引き起こすように考案された。このようなプラスミドをNIH3T3マウス細胞株にトランスフェクションし、3日後に塩基編集頻度を測定した。予想の通り、目標範囲のシトシン塩基は最大19％の効率でチミンに編集された(図34a)。DddAtoxは「TC」配列のシトシンのみを排他的に脱アミノ化すると従来に報告されたが、実験結果でもTCコンテキストの2つのシトシンのみが編集された。欠失または他の種類の点突然変異は、編集目標範囲で有意味に生成されなかった。

【０３９５】

インビボでのミトコンドリアの塩基編集。最も有効なDdCBE対(left-G1397-Nおよびright-G1397-C)をインビトロ実験に使用した。このDdCBE対をエンコードするインビトロ転写体をC57BL6/J胚の1細胞期に微小注入した4日後、総32個の胚のうち、9個の編集成功した胚が得られた(28％、表6)。

【０３９６】

【表６】

Type of mutagenesis	Number of examined embryos	Number of blastocysts (%)	Number of transferred embryos	Number of offspring (%)	Number of edited/total blastocysts (%)	Number of edited/total offspring (%)
N/A (buffer injection)	20	11(55)	NA	NA	0/11(0)	NA
ND5 silent	56	32(57)	30	4(13)	9(28)	3 (75)
ND5 G12918A	79	44(55)	50	11(22)	11(25)	4(36)
ND5 STOP	68	37(54)	120	27 (23)	19(51)	9 (33)

【０３９７】

ＴＡＬＥ-ＤｄｄＡｔｏｘ脱アミノ化酵素は、Ｃ・ＧからＴ・Ａへの塩基転位を効率的に生成したが、ｍ.Ｃ１２５３９では2.2から２５％の効率で、ｍ.Ｇ１２５４２では0.63から5.8％でその効率が示された。その後、ＤｄＣＢＥが注入された胚を代理母に移植し、ｍ.Ｃ１２５３９Ｔとｍ.Ｇ１２５４２Ｔを有する子孫を得た(図３９)。４匹の子(Ｆ０)のうち３匹がＣ・ＧからＴ・Ａへ編集され、その範囲は１～２７ ％までだった(図３４ｃ)。２匹の子は、つま先と尾で類似な突然変異数値を示し、生まれてから１４日後もその効率を維持していた。さらに、これらのミトコンドリアＤＮＡ突然変異は、生後５０日が経過した成体Ｆ０マウスの様々な組織において検出された(図３４ｄ)。これらの結果は、１細胞基複合体中でＤｄＣＢＥによって生じたミトコンドリアＤＮＡのヘテロプラスミが発達と分化の過程で維持されることを示唆する。

【０３９８】

ＤｄＣＢＥによって生じた突然変異が次世代に伝達されることを確認するために、雌Ｆ０マウスと野生型Ｃ５７ＢＬ６/Ｊ雄を交配させてＦ１子孫を作った。ｍ.Ｃ１２５３９Ｔおよびｍ.Ｇ１２５４２Ｔ突然変異は、２匹の子において６～２６ ％の効率で観察された。また、これらのミトコンドリア編集は、１１個の異なる組織で類似な数値で観察された(図３５ｂ)。

【０３９９】

ＤｄＣＢＥで媒介されたＭＴ-ＮＤ５ Ｇ１２９１８Ａ突然変異。次に、ｍ.Ｇ１２９１８Ａ突然変異を作製しようとしたが、この突然変異は人にミトコンドリア疾患を起こすこともある。注目すべき点は、この突然変異はいくつかのミトコンドリア疾患、例えば、Ｌｅｉｇｈ症候群、ＭＥＬＡＳ症候群、およびＬＨＯＮ症候群などを引き起こす。この位置のシトシン塩基は隣にチミンがあるため、ＤｄＣＢＥを用いた塩基編集が可能である(図３６ａ)。４対のＤｄＣＢＥを組み立て、ＮＩＨ３Ｔ３で最大6.4％まで編集が可能であることを確認した(図３６ｂ)。その後、最も効率の良いＤｄＣＢＥの組合をマウスコンジュゲートに微小注入し、その効率を胚盤胞で観察した。４４個のうち１１個(２５％)の胚がｍ.Ｇ１２９１８Ａ突然変異を有し、その効率は0.25から２３％までだった(図３６ｃ)。次に、ＤｄＣＢＥが微小注入された胚を代理母に移植してＧ１２９１８Ａ突然変異を有する子孫を得た(図３９ｂ)。１１匹の新しく生まれたマウスのうち４匹がこの突然変異を3.9～３１.６％程度有することを確認した(図３６ｄ)。生まれた直後に見られる表現型はなかったが、おそらく子が非常に幼く、また正常ミトコンドリアＤＮＡと突然変異ＤＮＡが異型状態で存在するため、これらの結果は、ＤｄＣＢＥがミトコンドリア疾患を有する動物モデルを作ることができることを示唆する。

【０４００】

ＭＴ-ＮＤ５ナンセンス突然変異。最後に、ＮＤ５の機能喪失突然変異がマウスで維持可能であるかどうかを遺伝子にナンセンス突然変異を作って確認した。ｍ.Ｃ１２３３６を目標シトシンとして、早期終止コドンをＮＤ５タンパク質の第１９９位に導入した(Ｑ１９９*；図３７ａ)。まず、４つのＤｄＣＢＥ組合をＮＩＨ３Ｔ３細胞株にトランスフェクトして塩基編集効率を確認し、最も効果的なＤｄＣＢＥ対がナンセンス突然変異を約5.7％効率で起こすことを確認した(図３７ｂ)。このＤｄＣＢＥはシトシン－チミン編集を引き起こし、わずかに効率が低いもののｍ.Ｇ１２３４１Ａのサイレント突然変異を生成する突然変異(Ｑ２００Ｑ)も目標範囲で

編集されることを確認した。37個のうち19個(＝51%)のマウス胚においても、m.C12336Tとm.G12341A突然変異を順に全て32%および23%の効率で確認した(図37c)。

【0401】

この結果に基づき、マウス胚を代理母に移植し、m.C12336Tとm.G12341A突然変異を有する子孫を得ることができた(図39c)。総27個のうち、9匹のF0マウス(23%)がC・G to T・A 編集を示し、その効率範囲は0.22から57 %であるが(図37d、e)、これを通じて、ND5のナンセンス突然変異が胚淘汰を起こさないことが分かる。

【0402】

実施例4．動物ミトコンドリアDNA編集

4-1．核外輸送信号が結合した動物ミトコンドリア塩基編集用発現ベクターの作製

TALE?DdCBEに核外輸送信号が結合したタンパク質を動物細胞で発現できるようにするベクター(図 40)を作製した。ベクターには、サイトメガロウイルスプロモーター(CMV promoter)を使用した。N-末端からミトコンドリア伝達信号、タンパク質精製/検出タグ、TALアレイN-末端ドメイン、繰り返し部位、C-末端ドメイン、DddAシトシン脱アミノ化酵素半体(split half)、ウラシルグリコシラーゼ阻害剤および核外輸送信号を含む(図40a)。核外輸送信号は、例えば、MVM(Mirute virus of mice)のNS2タンパク質由来を使用することができるが、他の配列も可能である。これにより発現されたタンパク質は核外に放出され、ミトコンドリアに伝達されて塩基編集が行われる。このとき使用されたターゲットDNAは、ミトコンドリアND5遺伝子-4番染色体ND5-like遺伝子、ミトコンドリア TrnA-5番染色体、ミトコンドリアRnr2-6番染色体である(図40b)。

【0403】

4-2．動物細胞株におけるDdCBE-NES

NIH3T3細胞株(ATCC CRL-1658)を、トランスフェクションの1日前の午後1.5×10^4/ウェルで1mlの細胞増殖培地(DMEM+10%Bovine Calf Serum)を含む12ウェルプレートに分注した。翌朝、DNAを処理してないMock、DdCBE、DdCBE-MVM NESプラスミド投入実験群をLipofectamine 3000の製造社のプロトコルに従って細胞にトランスフェクションした。その後、3日間二酸化炭素培養器(37℃、5% CO_2)で培養後、細胞を取り込み、 Qiagen Blood&Tissue Kitを用いて全体DNAを精製した。その後、ミトコンドリア遺伝子特異的なPCRプライマーを用いて増幅し、イルミナMiniseq装備を用いて次世代配列分析(Next-generation Sequencing)を行ってからその結果をCas-analyzer(www.rgenome.net)を用いて塩基編集効率を確認した。

【0404】

図40c、d、eは、DdCBE-NESをNIH3T3マウス細胞株にトランスフェクションし、マウスミトコンドリア遺伝子であるND5、TrnA、Rnr2に突然変異を作製したものである。DdCBEの組合別に効率が異なることが分かる。

【0405】

4-3．動物細胞株におけるmitoTALEN

図40fに示された配列を認識するTALENを作製し、この構造体にMTSを連結してミトコンドリアにDdCBEと共に導入した。実験方法は実施例4?2と同様である。TALE認識部位に意図的にミスマッチを与え、野生型mtDNAとは1つのミスマッチで、突然変異mtDNAとは2つのミスマッチが起こるようにした。TALEは、1つの配列ミスマッチは認識できない場合があるからである。その結果、DdCBEのみを処理したものに比べてTALENと共に処理した場合、+2ミスマッチ実験群で効率がわずかに上昇することを確認した。

【0406】

4-4．動物胚におけるDdCBE-NES

DdCBE発現ベクターとDdCBE-NES発現ベクターを鋳型として、PCR方式によりT7プロモーターとDdCBEまたはDdCBE-NES発現部分を増幅させたDNAを確保する。このDNAを鋳型としてT7ポリメラーゼを用いてmRNAを合成する。

【0407】

一対のDdCBE mRNAまたは一対のDdCBE-NES mRNAを微小注入溶液に混合し、マ

ウス受精卵に微小注入する。

4日間培養した受精卵は胚盤胞状態になり、この胚盤胞を溶解する。これを鋳型として使用し、ミトコンドリアDNAで核DNAとは区別される標的位置部分をPCRを用いて増幅した後、さらにインデックスおよびシークエンシングアダプターを追加のPCRにより増幅した。イルミナMiniseq装備を用いて高処理量シーケンシング(high-throughput sequencing)を行い、その結果をCas-analyzer(www.rgenome.net)を用いて塩基編集効率を確認した。また、ミトコンドリア標的部分が類似した配列を有する核内DNAをPCRを用いて増幅、シークエンシングを行った。

【０４０８】

その結果、DdCBEはミトコンドリアDNAだけでなく核内の類似DNA配列でも変異を誘発した。(ミトコンドリア：13.1%、核：3.2%)DdCBE-NESの場合、ミトコンドリア標的の変異効率を18.2%に高めるとともに、核DNA変異効率を0.2%に下げた。(図41a)。他のターゲットであるTrnAまたはRnr2の場合、核DNA突然変異は起こらなかったが、ミトコンドリアDNAの場合、統計的に有意味に塩基編集効率が増加することが分かった。(*p〈0.05、**p〈0.01、n.s.統計的に有意味ではない)

【０４０９】

4-5：動物胚におけるDdCBEおよびmitoTALEN

DdCBEを用いたND5遺伝子位置に、C-to-Tの転換後に編集されたミトコンドリアDNAが細胞内に占める比重を高めるために、編集されていないミトコンドリアDNA配列を切断するTALENを一緒に注入した。微小注入法とシーケンシング配列確認法は実施例4-2及び4-4と同様である。DdCBEを単独で微小注入したグループは11%の編集効率を示し、mitoTALENと同時に処理した場合33.3%に上昇して統計的に有意味な編集効率の上昇を得た。また、DdCBE-NESを単独で微小注入したグループは20.5%の編集効率を示し、mitoTALENと同時に処理した場合36.8%の効率を観察した。これも統計的に有意味であった(図41b)。

【０４１０】

微小注入された受精卵を代理母に移植し、新しく生まれたマウスの子でも同様にDdCBEは10.9%の編集効率を示したが、DdCBE-NESおよびmitoTALENを一緒に処理した場合、23.4%の効率が得られた(図 41c)。

【０４１１】

動物ミトコンドリア遺伝子の編集時に核外輸送配列を塩基編集タンパク質に付着した場合、より高い効率で塩基が編集され、動物胚では、特に、核内類似配列の非特異的塩基編集も抑制されることが示された。また、ミトコンドリア配列切断タンパク質を同時に用いる場合でも、より高い効率のミトコンドリア塩基編集が期待できる。

【０４１２】

実施例5．分割$DddA_{tox}$脱アミノ化酵素変異体

DdCBEの非標的効果を減らすことができる高精度のDddA由来シトシン塩基エディタを提示する。この非標的塩基編集効果は、TALEとDNAとの間の相互作用とは関係なく、DddAtox脱アミノ化酵素分割体の自発的な結合によって引き起こされる現象である。したがって、DddAtox分割体間の表面に位置するアミノ酸残基をアラニンに置換してHF-DdCBEを作製し、HF-DdCBEは、TALEに連結された2つの脱アミノ化酵素対がDNAに結合できない場合、適切な機能を果たさないようにした。全体のミトコンドリアゲノム分析により、ヒトのミトコンドリアDNAにおいて多数の所望しない非標的C-to-T転換を引き起こす従来のDdCBEとは異なり、HF-DdCBEは非常に効率的かつ精密であることを確認した。

【０４１３】

5-1．方法

プラスミドの作製。DdCBE発現プラスミドに点突然変異を導入した。プラスミドをQ5 Site-Directed Mutagenesis(NEB)用突然変異作製プライマー(表7)を用いて増幅し、その結果はサンガーシーケンシングで確認した。

【０４１４】
【表７】

Name	Sequence	No.
Left-G1333-N_F1315A_F	AGTGGGGACCGCATACTATGTCAACGAC	295
Left-G1333-N_F1315A_R	GTCTGACCGTTGTAGGCG	296
Left-G1333-N_F1329A_F	ATCCAAGGTTGCATCTAGCGGTGGGTCCGG	297
Left-G1333-N_F1329A_R	TCCAATCCCCCGGCGTCG	298
Left-G1333-N_G1313A_F	CAGACAGTGGCAACCTTTTACTATGTCAACGAC	299
Left-G1333-N_G1313A_R	ACCGTTGTAGGCGGGGAG	300
Left-G1333-N_I1299A_F	TCCATATCAGGCAAGTGCTCCCCAACTCCC	301
Left-G1333-N_I1299A_R	CCCAGGGCGTAGGAACCA	302
Left-G1333-N_K1327A_F	ATTGGAATCCGCAGTTTTCTCTAGCGGTGGGTCCGG	303
Left-G1333-N_K1327A_R	CCCCCGGCGTCGTTGACA	304
Left-G1333-N_L1294A_F	CTCTGGTTCCTACGCCGCAGGTCCATATCAGATTAG	305
Left-G1333-N_L1294A_R	CTAATCTGATATGGACCTGCGGCGTAGGAACCAGAG	306
Left-G1333-N_S1300A_F	ATATCAGATTGCAGCTCCCCAACTCCCCGC	307
Left-G1333-N_S1300A_R	GGACCCAGGGCGTAGGAA	308
Left-G1333-N_S1330A_F	CAAGGTTTTCGCAAGCGGTGGGTC	309
Left-G1333-N_S1330A_R	GATTCCAATCCCCCGGCG	310
Left-G1333-N_T1314A_F	GGTCAGACAGTGGGGGCATTTTACTATGTC	311
Left-G1333-N_T1314A_R	GACATAGTAAAATGCCCCCACTGTCTGACC	312
Left-G1333-N_V1312A_F	GGTCAGACAGCAGGGACCTTTTACTATG	313
Left-G1333-N_V1312A_R	GTTGTAGGCGGGGAGTTG	314
Left-G1333-N_V1318A_F	TTTTACTATGCAAACGACGCCGGGG	315

Left-G1333-N_V1318A_R	GGTCCCCACTGTCTGACC	316
Left-G1333-N_Y1292A_F	CTCTGGTTCCGCAGCCCTGGGTCCATATCAGATTAGTGCTCC	317
Left-G1333-N_Y1292A_R	CCAGAGCCCCCCAACCCC	318
Left-G1333-N_Y1316A_F	CAGTGGGGACCTTTGCATATGTCAACGACGC	319
Left-G1333-N_Y1316A_R	GCGTCGTTGACATATGCAAAGGTCCCCACTG	320
Left-G1333-N_Y1317A_F	GACCTTTTACGCAGTCAACGACGCCGGG	321
Left-G1333-N_Y1317A_R	CCCACTGTCTGACCGTTG	322
Left-G1397-N_C1376A_F	TGTGGCTTCGCAGTGAATATGACTGAAACCCTTCTG	323
Left-G1397-N_C1376A_R	CATATTCACTGCGAAGCCACAAGTACCCTCTG	324
Left-G1397-N_E1381A_F	GTGTGAATATGACTGCAACCCTTCTGCCCGAAAATG	325
Left-G1397-N_E1381A_R	CATTTTCGGGCAGAAGGGTTGCAGTCATATTCACAC	326
Left-G1397-N_K1389A_F	GAAAATGCCGCAATGACTGTCGTCCCACCTGAAGG	327
Left-G1397-N_K1389A_R	GACAGTCATTGCGGCATTTTCGGGCAGAAGG	328
Left-G1397-N_L1384A_F	GAAACCCTTGCACCCGAAAATGCCAAGATGAC	329
Left-G1397-N_L1384A_R	ATTTTCGGGTGCAAGGGTTTCAGTCATATTCACACAGAAG	330
Left-G1397-N_M1390A_F	AATGCCAAGGCAACTGTCGTCCCACCTGAAGG	331
Left-G1397-N_M1390A_R	GACGACAGTTGCCTTGGCATTTTCGGGCAGAAGG	332
Left-G1397-N_T1380A_F	GTGAATATGGCAGAAACCCTTCTGCCCGAAAATG	333
Left-G1397-N_T1380A_R	AAGGGTTTCTGCCATATTCACACAGAAGCCACAAG	334
Left-G1397-N_T1391A_F	CGAAAATGCCAAGATGGCAGTCGTCCCACCTGAAG	335
Left-G1397-N_T1391A_R	CTTCAGGTGGGACGACTGCCATCTTGGCATTTTCG	336
Left-G1397-N_V1377A_F	GGTACTTGTGGCTTCTGTGCAAATATGACTGAAACC	337
Left-G1397-N_V1377A_R	GGTTTCAGTCATATTTGCACAGAAGCCACAAGTACC	338

Left-G1397-N_V1392A_F	GAAAATGCCAAGATGACTGCAGTCCCACCTGAAGG	339
Left-G1397-N_V1392A_R	CCTTCAGGTGGGACTGCAGTCATCTTGGCATTTTC	340
Right-G1333-C_E1361A_F	GGCATTAGCGCAGGGCTTGTGTTCCATAATAATCC	341
Right-G1333-C_E1361A_R	CACAAGCCCTGCGCTAATGCCGTTATCTCTCATAAAAAGAGC	342
Right-G1333-C_F1353A_F	GTCAGCTCTTGCAATGAGAGATAACGGCATTAGCGAAGGG	343
Right-G1333-C_F1353A_R	TGCCCCTCGACGTGCCCG	344
Right-G1333-C_F1365A_F	GGGCTTGTGGCACATAATAATCCTGAGGGCACCTG	345
Right-G1333-C_F1365A_R	GATTATTATGTGCCACAAGCCCTTCGCTAATG	346
Right-G1333-C_G1362A_F	CATTAGCGAAGCACTTGTGTTCC	347
Right-G1333-C_G1362A_R	CCGTTATCTCTCATAAAAAG	348
Right-G1333-C_H1366A_F	GCTTGTGTTCGCAAATAATCCTGAGGGCAC	349
Right-G1333-C_H1366A_R	CCTTCGCTAATGCCGTTA	350
Right-G1333-C_K1389A_F	TGAGAACGCTGCAATGACTGTCG	351
Right-G1333-C_K1389A_R	GGCAGAAGTGTTTCGGTC	352
Right-G1333-C_L1363A_F	TAGCGAAGGGGCAGTGTTCCATAATAATCCTGAGGGC	353
Right-G1333-C_L1363A_R	ATGGAACACTGCCCCTTCGCTAATGCCGTTATCTC	354
Right-G1333-C_M1354A_F	AGCTCTTTTTGCAAGAGATAACGGCATTAGCGAAGG	355
Right-G1333-C_M1354A_R	GACTGCCCCTCGACGTGC	356
Right-G1333-C_M1390A_F	GAACGCTAAAGCAACTGTCGTACCAC	357
Right-G1333-C_M1390A_R	TCAGGCAGAAGTGTTTCG	358
Right-G1333-C_Q1349A_F	CGTCGAGGGGGCATCAGCTCTTTTTATGAGAGATAACGGC	359
Right-G1333-C_Q1349A_R	TGCCCGGCGTTAGCGTAG	360
Right-G1333-C_S1350A_F	CGAGGGGCAGGCAGCTCTTTTTATG	361

Right-G1333-C_S1350A_R	ACGTGCCCGGCGTTAGCG	362
Right-G1333-C_T1391A_F	CGCTAAAATGGCAGTCGTACCAC	363
Right-G1333-C_T1391A_R	TTCTCAGGCAGAAGTGTTTC	364
Right-G1333-C_V1346A_F	GCCGGGCACGCAGAGGGGCAGT	365
Right-G1333-C_V1346A_R	GTTAGCGTAGTTAGGATAAGGTGTTG	366
Right-G1333-C_V1364A_F	CGAAGGGCTTGCATTCCATAATAATCC	367
Right-G1333-C_V1364A_R	CTAATGCCGTTATCTCTC	368
Right-G1333-C_V1393A_F	AATGACTGTCGCACCACCCGAAG	369
Right-G1333-C_V1393A_R	TTAGCGTTCTCAGGCAGAAG	370
Right-G1397-C_F1412A_F	ACAAAGGTGGCAACAGGCAACTCTAACAGTCCAAAG	371
Right-G1397-C_F1412A_R	GTTGCCTGTTGCCACCTTTGTCTCCCCTGTTGCTCC	372
Right-G1397-C_K1410A_F	CAACAGGGGAGACAGCAGTGTTCACAGGCAA	373
Right-G1397-C_K1410A_R	TTGCCTGTGAACACTGCTGTCTCCCCTGTTG	374
Right-G1397-C_T1413A_F	GGAGACAAAGGTGTTCGCAGGCAACTCTAACAGTCC	375
Right-G1397-C_T1413A_R	GGACTGTTAGAGTTGCCTGCGAACACCTTTGTCTCC	376
Right-G1397-C_V1411A_F	GAGACAAAGGCATTCACAGGCAACTCTAACAGTC	377
Right-G1397-C_V1411A_R	GCCTGTGAATGCCTTTGTCTCCCCTGTTGCTCC	378

【０４１５】

結合表面突然変異の組み立てのために、ミニプレップされた突然変異発現プラスミドにモジュールベクター(それぞれTALE配列をコードする)、BsaI-HFv2(10U)、T4 DNAリガーゼ(200U)および反応バッファを1本のチューブに入れた。その後、制限酵素およびライゲイズ反応はサーモサイクラーにて37℃5分、50℃5分の20回繰り返した後、50℃を15分、80℃を5分間処理した。結合したプラスミドは、化学的形質転換によりE coli DH5aに導入され、最終構造体はサンガーシーケンシングにより確認された。細胞株の導入のために、プラスミドはミディプレップした。

【０４１６】

哺乳類細胞株の培養およびトランスフェクション。HEK 293T/17(CRL-11268, American Type Culture Collection (ATCC))細胞株は、37℃の5% CO_2環境で培養

した。細胞株は、DMEM supplemented with 10 % (v/v) fetal bovine serum (Gibco)の培地で抗生剤なしで増殖し、マイコプラズマ試験はしなかった。リポフェクションのために、細胞は24ウェルの細胞培養プレート(SPL, Seoul, Korea)にて1×10^5の細胞密度でトランスフェクションの18～24時間前に生長を開始した。Lipofectamine 2000(Invitrogen)を用いて、各々のDdCBE分割体を500ng使用して、総1,000ngのプラスミドDNAを導入した。細胞はトランスフェクションの4日後に回収された。

【０４１７】

高処理配列分析のためのゲノムおよびミトコンドリアDNA分離。ゲノムDNAを分離するために細胞培養培地を除去した後、細胞を底から分離するためにDNeasy Blood & Tissue Kit(Qiagen)のプロテイナーゼKが添加されたリシスバッファを細胞培養プレートに入れた。その後、製造社のプロトコルに従ってゲノムDNAを分離した。全体のミトコンドリアゲノム配列分析のためには、200μlのMitochondrial isolation buffer A (ScienCell)を細胞培養培地を除去した培養プレートに入れた。細胞をセルリフターを用いて掻き取った後、マイクロチューブに入れて使い捨てすりこぎを用いて細胞を砕いた。20回の粉砕後、よく砕かれたホモジネートを1,000×g、4℃で5分間遠心分離した。上清液を新しいマイクロチューブに入れ、10,000 ×g、4℃で20分間遠心分離した。沈殿物を10μlの溶解液(25mM NaOH, 0.2mM EDTA, pH10)に解放した後、95℃で20分間煮込んだ。pHを下げるために、1μlの1M HEPES(free acids, without pH adjustment)をミトコンドリア溶解液に加えた。このように調製した溶液の1μlを、高処理配列分析のためのPCR鋳型鎖として使用した。

【０４１８】

高処理配列分析。ディープシーケンシングライブラリーを作製するために、Q5 DNAポリメラーゼを用いてネステッド1次PCRおよび2次PCRを行い、最終インデックス配列を追加した。ライブラリーは、MiniSeq(Illumina)を用いてフェアドエンドリード配列分析に使用した。全体のミトコンドリアゲノム分析のためには、分離されたミトコンドリアDNAをtagmentation DNA prep kit (Illumina)を用いて製造社のプロトコルに従って調製した。全ての分析において、フェアエンド配列分析の結果は、1つのfastqjoinファイルを用いて合わせ、CRISPR RGEN Tools(http://www.rgenome.net/)を用いて分析した。

【０４１９】

5-2．結果

植物体内で葉緑体編集を試みる場合、非標的塩基突然変異が葉緑体遺伝体上に現れ、したがって、DdCBEの正確性について疑問を提起した。DdCBEの非標的塩基編集の2つの理由を考えてみた。まず、第一の理由は、TALEタンパク質とDNAとの間の非特異的な結合であり、第二の理由は、自然で意図せず起こるDddAtox半体同士の結合である(図42a)。本研究ではDddAtox分割半体に焦点を当て、分割された2つのタンパク質の結合表面を改善して、所望しないDddAtox半体の結合を防止するようにした。

【０４２０】

まず、ミトコンドリアのND1(mtND1)遺伝子を対象として、各々のサブユニット(Left-TALEもしくはRight-TALE)がそれぞれ単独でもDNAに結合、TALEのないTALE-free DddAtoxまたは結合配列のないTALE-DddAtoxの他の半体と相互作用してシトシンーチミン塩基編集を引き起こすかどうかを確認してみた。ヒトの腎胚細胞株(HEK293T)において、ヒトのミトコンドリアND1(mtND1)遺伝子を目標配列とするDdCBE対(Left-TALE:G1397N (N-末端側のG1397 DddAtox半体でTALE配列のC-末端に結合して、全体の認識配列のうち左側部位を引き受ける構造)+Right-TALE：G1397C(C-末端側G1397 DddAtox半体でTALE配列のC-末端に結合して全認識配列のうち右側部位を引き受ける構造))は、60.7%の効率で標的配列のC11を効果的にシトシンからチミンに編集した(図43a)。また、各々のサブユニットと対を成すことができるTALE-free DddAtox半体と共に処理したとき、元のDdCBE対よりは劣るが、塩基編集が起こることを観察した。したがって、Left-TALEとRight-TALEのそれぞれは、目標ND1配列に結合し、

31%または8.1%の効率で塩基を編集することが分かる(図43)。言い換えると、両方のTALEタンパク質とそれぞれ融合したDdCBE対は、一方のTALEとのみ結合した場合よりも2.0倍(=60.7%/31%)または7.5倍(=60.7%/8.1%)効率が高いことが分かる。明らかに、TALEと結合したDddAtoxのN-末端半体は、TALE配列のないDddAtoxのC-末端半体と結合して脱アミノ化反応を可能にし、その反対の場合も同様であった。

【0421】

DddAtoxは、2つの位置(G1333とG1397)に分かれて半体を形成するので、G1333位置でmtND1遺伝子を目標とするDdCBE対を作製(Left-TALE：G1333-NおよびRight-TALE：G1333-C)し、Left-TALE：G1333-NおよびRight-TALE：G133-C構造体を、対をなすTALE-free DddAtox半体と結合するように、それぞれ独立して用いてシトシン－チミン編集を起こすのかを確認した。予想の通り、各々のTALE融合DdCBEは、C8位で32.7%(左側TALE結合体)または18.1%(右側TALE結合体)の塩基編集効率を示し、これは、元のDdCBE対が示した56.1%と比較できる。したがって、TALEが結合したDdCBE対は、それぞれ別々に作用する場合に比べて1.7倍(=56.1%/32.7%)あるいは3.1倍(=56.1%/18.1%)の効率で作用することが分る。まとめると、これらの結果は、DdCBEが一方のTALE配列のみを結合する位置でも所望しない非標的塩基編集を引き起こすことを提示する。TALEタンパク質は、少数の不一致配列があってもDNAと結合するため、DdCBEは細胞小器官または核内遺伝子に非標的塩基編集を引き起こす可能性が存在する。

【0422】

分割されたDddAtox半体が互いに結合することによって生じる非標的塩基編集を減らすために、高精度のDdCBEを開発しようとした。分割二量体の結合表面部位を改良すれば、自己結合を防止または阻害できると考えた。したがって、PyMOLソフトウェアでは、Pythonコード(InterfaceResidues.py)を用いて1平方オングストロームの範囲内の2つの分割体DddAtox(分割位置G1333およびG1397)で結合表面のアミノ酸残基を見付けた。その結果、G1397-N(G1397位置で分割したN-末端DddAtox半体)における9個のアミノ酸残基を、G1397-C(G1397位置で分割したC-末端DddAtox半体)で4個の残基を、G1333-N(G1333位置で分割したN-末端DddAtox半体)で14個のアミノ酸を、G1333-C(G1333位置で分割したC-末端DddAtox半体) で15個のアミノ酸残基を見付けることができた(図42bおよび42c)。

以後、見付けたアミノ酸残基をアラニンに置換した複数の突然変異DddAtox半体を作製し、これらの結合表面突然変異DdCBEをそれぞれ野生型DdCBE対またはTALEのないDddAtox対と共にHEK293T細胞株に処理して塩基編集効率を観察した。C1376A、M1390AおよびF1412Aを含むG1397分割のいくつかの結合表面突然変異では、TALE-free DddAtoxのみならず、野生型とカップリングしたときにもシトシン－チミン編集効率を目標位置で観察できなかったが、これは、このような突然変異は他のDddAtox半体と近く位置しても相互作用できないことを意味する。一部の突然変異、例えば、E1381AおよびV1377Aの場合、 TALE-free 半体と結合して高効率の塩基編集を示したが、これは、このような突然変異が二量体の相互作用とは関係ない変化であることを示す。

【0423】

重要なことに、一部の突然変異、例えば、K1389A、K1410A、およびT1413Aの場合、野生型のDdCBE対と結合ときには高い活性を示したものの、TALE-free 対と使用するとぉには活性が低かった。一例として、K1410A突然変異は53.2%の効率を示し、野生型DdCBE対(60.7%)に比べて類似な結果を得たが、TALE-free と共に使用するときには、0.9%の効率を示し、59.1倍の差(＝53.2%/0.9%)を示した。前記のように、野生型の場合は、7.5倍(= 60.7%/8.1%)の差を見せる。また、これらの突然変異は、野生型のDdCBEよりも選択的に塩基を編集した。変異体は、編集範囲内でC_8 、C_9 、またはC_{13}よりもC_{11}に対して強い選択性を示したが、これは、野生型において4つのシトシンを均等に6.7%以上に全て編集したものと比較することができる(図43b)。

【0424】

さらに、G1333において、29個の突然変異(14個のG1333Nおよび15個のG1333C)をスクリーニングして、いくつかの適切な結合表面突然変異を得ることができた(図44)。多数の突然変異は、野生型とカップリングしたときに効率が低下するか(I1299A、Y1316A、Y1317A、およびF1329A)、TALE-free とカップリングしたときに効率が増加したもの(S1300AおよびT1314A)であった。しかし、注目すべき点は、 K1389A、T1391A、およびV1393Aのような突然変異は、野生型との結合では塩基編集がよく起こり、TALE-free とは低い効率を示した。一例として、K1389Aの場合は38倍の差(＝45.4%/1.2%)を示すことに対し、野生型DdCBE対は約3.1倍の差(＝56.1%/18.1%)に過ぎない。 また、K1389Aの場合、野生型より効率が良いように観察されることもあった。さらに、これらの突然変異は、C_{11}またはC_{13}ではないC_8位に編集効率が集中すると観察されたが、これは、野生型DdCBEの19%以上のすべてのシトシン位置に編集効率が上昇することと比較できる。また、注目すべき点は、G1397-CのK1410AはC8を優先的に編集するのに対し、G1333-CのK1389Aの場合、C11を選択的により高効率に編集することが分かる。一方、野生型のDdCBE対(G1333または G1397 $DddA_{tox}$分割)は、いずれもこの選択性が低いことが分かる。このような結果は、前記の結合表面突然変異はDdCBEにおける標的部位内の多数の塩基に所望しない編集を引き起こす効果を減らす可能性を示す。

【０４２５】

実施例6．全長脱アミノ化酵素

分割されたバクテリア毒素DddAtox、TALEアレイ、ウラシルグリコシラーゼ阻害剤(UGI)を含む DddA由来シトシン塩基編集酵素(DdCBE)は、真核細胞の核、ミトコンドリアDNA(mtDNA)、植物の葉緑体DNAの標的シトシンをチミンに変換可能にする。バクテリアに毒性を誘発するDddAtoxは、Burkholderia cenocepaciaに由来し、酵素で二本鎖DNA(double-strand DNA)内でシトシンを脱アミノ化する。宿主細胞の毒性を避けるために、DddAtoxは不活性を有する半分に分け、それぞれ TALE DNA 結合タンパク質に結合してDdCBEを作製する。TALEアレイに結合された2つの不活性形態は、TALEによって標的DNAに近接して結合されるときに機能を有する。シトシンからチミンへの塩基変換は、2つのTALE結合部位間の14～18個の塩基間の領域で誘導される。
DdCBEは、細胞小器官DNAを編集できないCRISPR由来塩基編集器とは異なり、核および細胞小器官DNAの両方で標的化された塩基編集を可能にするが、それを誘導するために1つではなく2つのTALE構成が必要であるという欠点がある。第一の欠点は、TALEが5'末端と3'末端の両方においてチミンに標的DNA部位に結合する必要があるため、2つのTALEアレイを使用すると標的可能部位が制限される。第二の欠点は、1つではなく2つのTALE構成を伝達することはしばしば非効率的かつ挑戦的である。遺伝子治療に広く用いられるアデノ関連ウイルス(AAV)ベクター(用量、～4.7kbps)のように用量が制限されたウイルスベクターは、二量体DdCBEの組合が大きすぎる(2×4.1kbps、プロモーターおよびpolyA信号を含む)ので、2つのDdCBEエンコーディング配列を収容できない。より大きな用量の単一のベクターに2つのTALEアレイDNAをクローニングすることは、2つのTALEアレイ配列の高い類似性のために難しい場合がある。最後に、1つではなく2つのTALEアレイを使用すると、オフターゲット効果が悪化する可能性がある。分割DddAtoxのある二量体DdCBEのかかる限界を克服するために、無毒性全長DddAtox形態のmDdCBE(モノマーDdCBE)を核および小器官の標的DNAにおいてシトシンからチミンへの転換を誘導するDdCBEを提示する。

【０４２６】

6-1．方法

プラスミドの作製。DddA変異体は、合成された全長DddAtox(gBlock, IDT)を鋳型として使用し、表8のプライマーおよびQ5 DNAポリメラーゼ(NEB)を用いてPCR増幅した。これらのPCR産物は、Apobec1がBamHIとSma I(NEB)で切断されたp3s-BE3部位でギブソンアセンブリ(NEB)を用いてクローニングした。TALE-DddAtox(Addgene #158093、#158095、#157842、#157841)は、プラスミドをBamHIとSma Iで

切断し、DddA変異体を表8のプライマーを用いてPCR増幅し、ギブソンアセンブリを用いてクローニングした。確保したプラスミドを化学的方法で製造されたE. coli DH5aに熱衝撃法で形質転換し、生存したコロニーのプラスミド塩基配列をサンガーシーケンシングにより分析した。最終プラスミドは細胞トランスフェクションのためにミディプレップ(Macherey-Nagel)した。

【０４２７】

【表８】

Name	Sequence	No.
DddA F	cccaagcttgccaccatgggatccggcagctacgccctgggtccgtatcag	380
DddA R	cccgggagtctcgctgccgctgcaaccgcctttggtcgggctcttcggg	381
K1402A R	cccgggagtctcgctgccgctgcaaccgcctttggtcgggctcttcgggctgttg ctgttaccggtaaacactttggtttcaccggtagcgccgcgCGCcactggaatgg	382
R1403A R	cccgggagtctcgctgccgctgcaaccgcctttggtcgggctcttcgggctgttg ctgttaccggtaaacactttggtttcaccggtagcgccGGCcttcttcactgg	383
K1402A R1403A R	cccgggagtctcgctgccgctgcaaccgcctttggtcgggctcttcgggctgttg ctgttaccggtaaacactttggtttcaccggtagcgccGGCCGCcactggaatgg	384
K1410A R	cccgggagtctcgctgccgctgcaaccgcctttggtcgggctcttcgggctgttg ctgttaccggtaaacacGGCggtttcacc	385
K1420A R	cccgggagtctcgctgccgctgcaaccgcctttggtcgggctGGCcgggctgttg ctg	386
K1424A R	cccgggagtctcgctgccgctgcaaccgccGGCggtcgggctcttcgggg	387
E1347 F	ctatgccaatgccggtcatgtgGCCggtcagagcgccctgttcatg	388
E1347 R	catgaacagggcgctctgaccGGCcacatgaccggcattggcatag	389
error_prone PCR F	tgagccagctgggcggcgacagcggcagcgagactcccgggacctcagagtccgc cacacccgaaagtggcagctacgccctgggtcc	390
error_prone PCR R	aatagggccctctagatgcatgctcgagttagcaaccgcctttggtcggg	391
TAIL Right DddA F	tgaagaaaggtctgggcggatccggcagctacgccctgggtccgtatcag	392
TAIL_Left_DddA_F	tgaagaaaggcctgggtggatccggcagctacgccctgggtccgtatcag	393
TAIL DddA R	tcagattagttgagccgccagagcaaccgcctttggtcgggctcttcgggg	394
TAIL UGI F	tctggcggctcaactaatctgagcgacatcattg	395
TAIL UGI R	cagatccgaaaatggatatacaagctccc	396

【０４２８】

ランダム突然変異誘発。GeneMorph ii Random mutagenesis kit(Agilent)を用いて製造社のプロトコルに従って合成された全長DddAtox(gBlock, IDT)を鋳型としてエラーが発生しやすいPCRを行った。要約すると、DddAtox DNA 1ng、100ng、および700ngを鋳型として使用して、それぞれ0～16 mutations/kbのランダム変異を導入した。全長DddAtox gBlockは、表8のプライマーを用いて予めPCR増幅した。全体のPCR結果物を合わせ、Sma1およびXho1で切断したp3s-UGI-Cas9(H840A)にギブソンアセンブリ(NEB)を用いてクローニングした。化学的方法で製造されたE. coli DH5aに熱衝撃法で細胞にプラスミドを形質転換し、生存したコロニーが有するプラスミド塩基配列をサンガーシーケンシングにより分析した。分析したプラスミドのうち、コーディングフレームを有するp3s-UGI-nCas9(H840A)-DddAtoxプラスミドを、sgRNAと共にHEK293T細胞にトランスフェクトした後、ターゲットディープシークエンシング

による編集活性を確認した。

【０４２９】

哺乳動物細胞培養およびトランスフェクション。HEK293T(ATCC, CRL-11268)細胞およびHeLa(ATCC, CCL-2)細胞を5%CO_2で37℃に維持して培養した。細胞を、10%(v/v)ウシ胎児血清(Welgene)および1%ペニシリン／ストレプトマイシン(Welgene)を補充したDMEMで培養した。細胞をトランスフェクションの24時間前に3X105細胞(HEK293T)および4x104細胞(HeLa)の密度で48ウェルプレート(Corning)に播種し、Lipofectamine 2000(Invitrogen)およびCas9融合DddAプラスミド(750ng)とsgRNA(250ng)でトランスフェクションした。TALE-DddAは、200ngプラスミドとLipofectamine 2000を用いてHEK293T細胞にトランスフェクションした。sgRNA配列を表9に示す。

【０４３０】

【表９】

Name	sgRNA sequence 5` to 3`	No.
HEK3	GGCCCAGACTGAGCACGTGA	397
TYRO3	GGCCACACTAGCGTTGCTGC	398
ROR1 site1	GCCATAGATGGTGGACCGAA	399
ROR1 site2	CCATCTATGGCTCTCGGCTG	400
ROR1 site3	CCGCAGCCGAGAGCCATAGA	401
FANCF	GGAATCCCTTCTGCAGCACC	402
HBB	CTTGCCCCACAGGGCAGTAA	403
EMX1-1	TGCCCCTCCCTCCCTGGCCC	404
EMX1-2	CCCTGGCCCAGGTGAAGGTG	405
EMX1-3	GTGAAGGTGTGGTTCCAGAAC	406
EMX1-4	AAAGTACAAACGGCAGAAGC	407
TRAC5 site1	GTGGTAGCGGAACTCACTAAG	408
TRAC5 site2	CACCCAGCCTGCTCTGCCTT	409

【０４３１】

ゲノムおよびミトコンドリアDNAの調製。Cas9融合DddA変異体にトランスフェクションされた細胞をトランスフェクションの2日後に収穫し、TALE-DddAにトランスフェクションされた細胞を3日後に収穫した。ゲノムおよびミトコンドリアDNAは、DNeasy血液および組織キット(Qiagen)を用いて分離した。大規模分析のために、5μLのプロテイナーゼK(Qiagen)を含む100μLの細胞溶解緩衝液(50mM Tris-HCl, pH8.0(Sigma-Aldrich)、1mM EDTA(Sigma-Aldrich)、0.005% sodium dodecyl sulfate (Sigma-Aldrich))を用いてDNAを抽出した。溶解物を55℃で1時間反応後、95℃で10分間反応させた。

【０４３２】

6-2．結果

野生型と新規全長DddAのアミノ酸塩基配列を比較し、図45に変更されたアミノ酸を灰色のボックスで示した。

【０４３３】

図46に示すように、DddAをCas9のN-末端の前に16個のアミノ酸からなるリンカーを用いて連結し、C-末端には、UGI(uracil glycosylase inhibitor)とNLS(nuclear localization signal)を4個のアミノ酸からなるリンカーを用いて連結した。逆に、DddAをCas9のC-末端の後に16個のアミノ酸からなるリンカーを用いて連結し、N-末端には UGIとNLSを4個のアミノ酸からなるリンカーを用いて連結した。

【０４３４】

本発明では、DddA-Cas9(D10A、D10AとH840A)-UGIを用いて実験した。DddA

をDNA結合タンパク質であるジンクフィンガータンパク質、TALEモジュールに結合し、1つのモジュールのみでシトシンをチミンに置換することができる。従来のスプリットには2つのモジュールが必要であるが、完全長DddAは1つのモジュールのみでも可能である。これら2つのDNA結合タンパク質は、NLS(nuclear localization signal)、MTS(Mitochondrialtargeting sequence)、CTP(chloroplast transit peptide)を連結してゲノム領域だけでなく、Cas9では不可能なミトコンドリアと植物の葉緑体のシトシンをチミンに置換することができる。図47に示すように、ヒト細胞ジェノミックコンテキストROR1サイト(a)、HEK3サイト(b)、TYRO3サイト(c)において、TCモチフのシトシンがチミンに置換される活性を確認した。ターゲット位置で25 bp 離れたTCモチフのシトシンがチミンに置換された活性を確認した(a)。A1341D KRKKAは、CCモチフの2番目のシトシンがチミンに置換する活性を確認した(a、b)。触媒突然変異であるE1347Aも、TCモチフのシトシンがチミンに置換される活性を確認した(a、b、c)。 赤い下線は、Cas9がバインディングする部位を示した。 効率は、全体のシークエンシングリード中の欠失のないリード中でシトシンがチミンに置換された割合をパーセントで示した。また、全体のシークエンシングリード中欠失の割合をパーセントで示した。

【０４３５】

図48に示すように、赤い四角ボックスは分割してDddAtoxの活性を確認した部分を示したもので、同じ部分を全長DddAを用いて3つのターゲットサイトに分けて活性を測定した。分割体は、PAMが異なるorthogonal Cas9を使用して2つのCas9の間にあるシトシンをチミンに置換するが、この場合、所望のシトシンをチミンに正確に置換することは困難である。しかし、完全長DddAは、同じターゲットサイトにCas9がバインディングする部分を3つに分けてターゲットでき、所望のシトシンをチミンに精巧に置換することができる。効率は、全体のシークエンシングリード中の欠失のないリード中でシトシンがチミンに置換された割合をパーセントで示した。また、全体のシークエンシングリードのうち欠失の割合をパーセントで示した。

【０４３６】

図49に示すように、ヒト細胞ゲノムコンテキストTRACサイト1 (a) 、TRACサイト2(b)、FANCF(c)、HBB(d)における全長DddAの活性を測定した。 赤い下線は、Cas9がバインディングする部位を示した。効率は、全体のシークエンシングリード中の欠失のないリード中でシトシンがチミンに置換された割合をパーセントで示した。また、全体のシークエンシングリードのうち欠失の割合をパーセントで示した。

【０４３７】

図50に示すように、DddA-dCas9(D10A, H840A)-UGIを用いたヒト細胞ゲノムコンテキストTYRO3 (a)、ROR1(b)、HEK3(c)、EMX1 site 2(d)、TRACサイト1(e)、HBB(f)におけるDddAの活性を測定した。効率は、全体のシークエンシングリード中のシトシンがチミンに置換された割合をパーセントで示した。欠失は観察されなかった。

【０４３８】

塩基編集に有用な無毒性の全長DddAtox変異体を得るために、構造に基づく特定部位の変異、ランダム突然変異の2つの方式をとった。最初の接近方式では、DNA結合が減少するか触媒活性が減少したDddAtox変異体を不活性CRISPR-Cas9(dCas9)またはニッカーゼ(nCas9)変異体に融合し、ヒト由来の細胞内の標的シトシンがチミンに置換される新しい塩基編集器を開発した。このために、DddAtoxの正電荷を有するアミノ酸をアラニンに置換して発現ベクターにサブクローニングしようとした(図51a)。このような変異体が負電荷を有するdsDNAに結合することを弱め、潜在的に毒性を避けることができると推論した。ほとんどのアラニン置換変異体はE. coli形質転換体を形成できなかった(図51b)。得られた形質転換体から分離されたプラスミドDNA塩基分析結果、タンパク質を形成する領域において、様々なフレーム変異突然変異が誘導された。これらの全長DddAtox変異体は哺乳類発現プロモーターの制御下にあるが、E. coliで弱く発現してアポトーシスをもたらした。幸いなことに、フレームシフト突然変異のないいくつかの三重または四重または五重(「AAAAA」という)アラニン置換変異体を得ることができた。さ

らに、活性部位突然変異E1347Aも成功的にクローニングされた。

【0439】

次に、D10A nCas9またはdCas9およびUGIに融合したAAAAA変異体がヒト胚腎臓293T(HEK293T)細胞において塩基編集を誘導できるか否かを調べた(図51c、d)。ラットAPBEC1脱アミノ酵素、ウラシルグリコシラーゼ阻害剤(UGI)とdCas9(またはD10A nCas9)からなる塩基編集器2(または3)は、プロトスペーサー領域内の狭い領域で活性を有するが、AAAAA変異体はプロトスペーサー領域のすぐ5'上方でシトシンからチミンへの最大43%の置換を誘導した。予想外に、E1347A突然変異は、37%(nCas9融合)または16%(dCas9融合)の頻度で同じC_{-3}位で塩基編集を誘導し(図51c)、これは、E1347A突然変異がDddAtoxの完全な脱アミノ不活性化を誘導せず、ヒト由来細胞において塩基編集を達成するのに十分高い脱アミノ活性を有することを確認した。しかし、5重AAAAA突然変異と結合したE1347A変異体は塩基編集を誘導できなかった。また、dCas9またはnCas9およびUGIに融合された移動突然変異のないE1347A、AAAAAおよびアラニンに置換された他の変異体(図53)が、プロトスペーサー領域の上方に最大25ベース離れた位置で編集が起こることを確認し、他の多くのサイトでは、最大26%の編集効率があることを確認した(図54)。さらに、最大60%の編集頻度でHeLa細胞において非常に効率的であった(図55)。そのような融合タンパク質によって誘導された塩基編集は、最大21日間細胞で維持され、これは、そのような塩基編集が細胞毒性のないことを示唆する(図56)。

【0440】

シトシン塩基編集器の編集ウィンドウの変更のために、アラニンに置換された変異体はH840A nCas9のC-末端に融合しようとした。予想外に、フレームシフト突然変異のない無傷の構成物を得ることに失敗した。したがって、エラーが発生しやすいPCRを行い、DddAtoxコーディング配列にランダム突然変異を導入し、4点突然変異S1326G、G1348S、A1398V、S1418G(「GSVG」という)のある無毒性全長DddAtox変異体を得ることができた(S1326G、G1348S、A1398V、S1418Gそれぞれ配列番号269のアミノ酸配列のうち、37位のSがGに置換；59位のGがSに置換；109位のAがVに置換；及び129位のSがGに置換することを含む、配列番号276の配列を含む、図52a)。さらに、この変異体はdCas9、D10A nCas9およびCas9のC-末端とdCas9、nCas9およびCas9のN-末端に融合した。ヒト由来細胞における野生型Cas9を除くこれらの融合タンパク質は、最大38%の効率で様々な部位でシトシンからチミンへの転換を誘導した(図52b、57、58)。興味深いことに、dCas9、D10A nCas9、およびH840A nCas9のC-末端に融合されたGSVG変異体を含む融合タンパク質は、プロトスペーサー隣接モチーフ(PAM)の3'ダウンストリームでシトシン塩基編集を示すことに対し、N-末端に融合された同じ変異体を含む融合タンパク質であるdCas9およびnCas9は、プロトスペーサーの5'アップストリームで塩基編集を誘導した(図52c)。予想の通り、Cas9を含む融合タンパク質は、塩基置換よりは挿入欠失を誘導した。

【0441】

どの突然変異がGSVG変異体にとって重要であるかを見付けるために、SSVG、GGVG、GSAG、およびGSVSの4つの復帰体をサイト指定突然変異誘発によって作製しようとした。SSVG、GSAGおよびGSVS復帰体は得られたが、nCas9のC-末端に融合したGGVG変異体は得られなかった。G1348は、触媒部位の核心部位であるE1347のすぐ隣にある。G1348S突然変異は、触媒活性を低下させ、大腸菌の細胞毒性を回避することができた。最大21日間、トランスフェクションされた細胞の2つの標的部位にて3つの復帰体とGSVG変異体の編集頻度を測定した。GSAGおよびGSVSによって誘導されたシトシンからチミンへの編集の頻度はトランスフェクション後の3日から21日まで徐々に～2倍減少し、これら2つの復帰体はやや細胞毒性である一方、GSVGおよびSSVGは維持された(図59)。これらの結果は、GSVG変異体においてG1348Sが必須であり、S1326Gが中性である一方、A1398VおよびS1418Gは細胞毒性を低下させることを示唆する。

【0442】

まとめると、前記結果は、dsDNAに対する親和度減少(AAAAA)または弱化された脱アミノ活性(E1347AおよびおそらくはGSVG)または減少した細胞毒性(GSVG)を有する無毒性、全長DddAtox変異体がdCas9またはnCas9に融合して新たに変更された編集ウィンドウを有する塩基編集器を発明した。このような塩基編集基は、以下dCas9-mDdBE(dCas9のC-末端に融合した全長のモノマーDddAtox変異体からなるDddA由来の塩基編集器)、nCas9-mDdBE、mDdCE-dCas9およびmDdCE-nCas9と名付けることができ、BE2またはBE3において到達できないプロトスペーサーの上流または下流位置で塩基編集に用いられるBE2またはBE3以後の塩基編集器である。

【０４４３】

さらに、無毒性の全長DddAtox変異体をミトコンドリアDNA編集に使用できるか否かを調べた。多数の変異体の中で、GSVGおよびE1347A変異体の2つのみがミトコンドリア遺伝子、ND4およびND6に結合するように設計されたTALEアレイのC-末端に成功的に融合した。GSVG変異体を含むモノマーDdCBE(mDdCBE)は、もともと分割されたDdCBE対と同等の最大31%(ND4)(図60a)および27%(ND6)(図60b)の標的ヌクレオチド位置で塩基編集を達成した。E1347Aを含むmDdCBEも最大7.2%(ND4)と8.9%(ND6)の編集比で効率性は低下するが、目標シトシンをチミンに転換した。興味深いことに、ND4遺伝子に特化した元のDdCBE対(G1333分割)は、C4位で編集効率が0.8%であるのに対し、GSVGを含む2つのmDdCBEは26%と31%の高い編集効率を示した。この結果は、分割二量体DdCBEとmDdCBEが異なる突然変異パターンを有する可能性があることを示し、mDdCBEが二量体DdCBEと相補的であり得るので、与えられた標的部位で多様な突然変異を誘導できることを示唆する。

【０４４４】

分割二量体DdCBEに比べてmDdCBEの潜在的な利点の1つは、非特異的TALE-DNA相互作用によるオフターゲット効果が二量体DdCBEに比べて半分であることである。分割DddAtoxを有する二量体DdCBEは、1つのサブユニットのみを結合できる半分の部位で機能することができ、これは所望しないオフターゲット突然変異を引き起こす。DdCBE対の不活性DddAtox半分は、他の不活性半分を募集して機能的脱アミノ酵素を形成することができる。この仮説を確認するために、二量体DdCBEの1つのサブユニットをエンコードするプラスミドとTALEのないDddAtox半分をエンコードするプラスミドをHEK293T細胞に共同トランスフェクションし、両方のミトコンドリア標的部位で編集頻度を測定した。予想の通り、シトシンからチミンへの編集は、0.7～3.6%の頻度で対象サイトで観察された(図60c～f)。この結果は、DdCBE対で分割されたDddAtox半分が半分の部位で互いに相互作用して所望しないオフターゲット突然変異を引き起こす可能性があり、mDdCBEが二量体DdCBEによって誘導されたオフターゲット突然変異の半分を回避できることを示唆する。

【０４４５】

実施例7．DdABEを用いたヒト細胞における高効率のA-to-G塩基編集

DddA由来シトシン塩基編集(DdCBE)によるミトコンドリアDNA塩基編集は、様々な細胞株(cell line)と動物の疾患モデルを作製することを可能にし、ミトコンドリア遺伝病を治療するための新しい道を開いた。しかし、DdCBEは、ほぼTC-toTT塩基編集のみ限定的に引き起こすため、すべての場合の数で1/8程度だけカバーすることができる。そこで、TALE(transcription activator-like effector)に2種類の脱アミノ酵素を連結したTALEDを開発した。ここでTALEは、所望のDNA部分に結合できるようにカスタマイズし、触媒能力を失ったDddAtoxシトシン脱アミノ酵素変異体およびE.coli 由来のDNAアデニン脱アミノ酵素であるTadAタンパク質を含んでいる。TALEDは、ヒトミトコンドリアで従来のTCコンテキストに関してのみシトシン塩基編集が可能だった従来の塩基編集技術とは異なり、すべてのA-to-Gに対する塩基編集を可能にする。実際に、カスタマイズされたTALEDは、ヒト細胞の複数のターゲットで高い効率(50%程度まで)でアデニン塩基編集を引き起こすことができた。

【０４４６】

従来はなかった新しい塩基編集技術を開発するために、多数のTadA変異体の中からABE8eのTadA変異体(TadA *)を選択した。これは、高効率でアデニン編集ができるだけでなく、様々な多様なDNA結合タンパク質と互換性を有するように改良されたものであるため、実際のTALEまたはZFPにも互換性が良くなるからである。

【０４４７】

初めてND1またはND4を標的できるようにカスタマイズされたTALEにTadA *とMTSを融合し、実際にミトコンドリアDNAに塩基編集を引き起こすことができるかどうかを実験した。ターゲットディープシークエンスの結果、融合タンパク質のアデノシン塩基編集効率が非常に低いとはいえ、あることを示した。ND1サイトでは最大1.2％効率(図67a)で、ND4サイトでは0.6％効率(図67b)でアデニン塩基編集を引き起こした。たとえ非常に低い効率であるが、TadA *がDNA一本鎖のみ特異的に機能すると知られたことに比べてTALEに融合していると二本鎖ターゲットDNAでも塩基編集を起こすことが分かった。

【０４４８】

ミトコンドリアDNAでアデニン塩基編集が起こり得るという結果とともに、かかる効率をさらに上げるために、既に知られているDddAtoxタンパク質を融合することを考えた。DddAtoxタンパク質は、シトシン脱アミンをするタンパク質として、Burkholderia cenocepaciaに由来するinterbacterial toxinである。このタンパク質は、DNA二本鎖に機能するタンパク質であるため、このタンパク質を使用すればTadA *アデニン脱アミノ酵素がターゲットDNAによりよくアクセスできるようにするからである。従来のDddAtoxを使用するDdCBEのような場合、DddAtoxタンパク質を半分に分けて左側のDNA部分に付着するTALE(Left-TALE(L-TALE)と右側のDNA部分に付着するRight-TALE(R-TALE))にそれぞれ分かれて分割形態で融合し、シトシン塩基効率を上げるウラシルグリコシラーゼ阻害剤(UGI)まで融合(TALE-Split DddAtox-UGI)されている。このようにDddAtoxを分割して使用する理由は、全長のタンパク質を使用すると細胞毒性が発生するためである。まず、ND1サイトをターゲットするDdCBEの片側にUGIの代わりにTadA *を取り付けて、L-TALE-Split-DddAtox-TadA *、R-TALE-1397C-TadA *の形でそれぞれ作って実験してみた。不思議なことに、一方の部分にTadA *があることを、他方の部分には1397NとUGIがあることを対にしてヒト細胞に伝達したとき、A-to-GとC-to-T塩基編集の両方が起こることを確認できた(図62c)。従来のDdCBEのような場合、約20％シトシン塩基編集が起こり、アデニン塩基編集は全く起こらない一方、一方の部分にTadA *がある場合、シトシン塩基編集が半分レベルに減少し、アデニン塩基編集が約10％程度起こることを確認した(図62c)。アデニン塩基編集とシトシン塩基編集効率は類似に起こった(図62c)。

【０４４９】

もちろん、シトシン塩基編集とアデニン塩基編集が全て起こることは、ランダムに突然変異を引き起こすのに有用かもしれないが、疾患を治療することにおいて、特にC-to-T突然変異で発生するLOHNやMEALSのようなミトコンドリア遺伝病を治療するためには、アデニンのみ塩基編集を引き起こすことが望ましい。したがって、このように同時に発生するシトシン塩基編集をなくすためにUGIをなくした。DdCBEでは、UGIの場合、シトシン脱アミノ酵素DddAtoxがCをUに脱アミノ化すると、DNA修復過程中に細胞内にある修復タンパク質であるウラシルグリコシラーゼ(uracil glycosylase)によってUが再修復されるのを防止しようと、ウラシルグリコシラーゼを阻害するための阻害剤用途として融合されたものである。したがって、このようなUGIを除去すれば、アデニン塩基編集効率は維持し、シトシン塩基編集を抑制できると考えた。驚くべきことに、UGIのないND1標的TALE脱アミノ酵素対は、シトシン塩基編集をほとんど起こさず(〈0.5％)、高効率でアデニン塩基編集(約50％)のみを引き起こすことを確認した(図63a、c)。これはUGIがあるものよりもはるかに高い効率です。さらに、ND4を標的するTALE脱アミノ酵素対でも確認し、ND1を標的することと同様に高い効率(約35％)でアデニン塩基のみを編集することを確認した(図63b、d)。このようにDddAtoxシステムとTadA *を融

合したDNA二本鎖で機能する新しいアデニンデアミナーゼであるTALEDを開発し、世界初でヒトミトミトコンドリアでアデニン塩基編集を行うことができた。また、初めてTALEにTadA *だけを付着したものより、最終的に50倍程度高くアデニン塩基を編集することができた。

【０４５０】

次に、さらに触媒活性をなくした全長E1347A DddAtox変異体または開発した触媒活性は維持し、細胞毒性はなくした変異体(AAAAAとGSVG)を活用してアデニン塩基編集を引き起こそうとした。シトシン塩基編集を起こすことが目的ではなく、DNA二本鎖でアデニン塩基編集を引き起こすことが目的であるため、シトシン塩基編集活性がなくなった全長E1347A DddAtox変異体を用いてもよいし、UGIがなければシトシン塩基編集の効率がなくなることを確認したため、同時に細胞毒性だけをなくした変異体も用いることができた。全長変異体を含むTALEDを2つのタイプにした(図64ａ)。一番目は、1つのTALEにTadA *(AD)と全長DddAtox変異体の両方を含むこと(mTALED)、二番目は、それぞれのTALEにTadA *(AD)と全長DddAtox変異体をそれぞれ別々に融合すること(dTALED)の2つのタイプを作って試験した(図64a)。驚くべきことに、ND1を標的する2つのタイプのTALEDは、いずれも高効率でアデニン塩基編集を引き起こすことを確認した(図64b)。mTALEDの場合、最大約45％の効率を示し、dTALEDも約50％程度のアデニン塩基編集効率を示した(図64b)。ND1サイトだけでなくND4サイトでも同様に実験を行い、類似に高い効率でアデニン塩基編集を引き起こした(図64c)。また、シトシン塩基編集活性がなくなった全長E1347A DddAtox変異体を用いた場合も、高効率のアデニン塩基編集を引き起こすことを確認した(図 64b、c)。これは、シトシン脱アミン能力が失われたにもかかわらず、十分にTadA *をDNA二本鎖に近づけることを助ける役割は維持されることを示している。前記の結果を詳細に一塩基レベルで見たとき(図65、66)、TALEがDNAに結合するすぐ近くでアデニン塩基編集が起こることが分かる。また、2つのTALEを用いる場合、その間(Spacer)でのみ塩基編集が起こることになるが、不思議にも1つのTALEが用いるmTALEDでも類似なターゲット長を有することが見られた(図65、66)。

【０４５１】

また、システムがジンクフィンガータンパク質(ZFP)システムでも機能するのか、核DNAでも機能するのかが気になった。したがって、核DNAをターゲットとするNC型のZFPを作製し、ここに分割されたDddAtoxとTadA *を融合した(図61a)。このとき、ZFPに様々な位置にTadA *を融合させた(図61b)。様々な形態のうち、核DNAにおいて最大10％程度のアデニン塩基を編集する個体を作ることができた(図61d)。UGIが片側に存在するため、シトシン塩基編集効率も高効率で存在することが見られる(図61c)。ヒト細胞核DNAにおいてZFP-DddAtox-TadA *システムが機能することがわかり、さらに、ミトコンドリアでも機能するのか実験した。このとき、核DNAで最も高効率で機能した構造を参考にした。核DNAに伝達させるNuclear Localization Signal(NLS)の代わりにミトコンドリアに向けるMitochondrial Targeting Sequence(MTS)を取り付け、ND1サイトをターゲットするZFP(NCタイプ)に、左側のZFPには1397Nが、右側のZFPには1397CとともにTadA *を融合した形で実験をした。その結果、約3％効率でアデニン塩基編集が起こることが分かった(図61.g)。たとえTALEDよりアデニン塩基編集効率が低かったが、タンパク質間を連結するリンカーなど様々な条件を最適化すると、このZFPシステムを用いても良い効率のアデニン塩基編集を起こすことができるだろう。

【０４５２】

現在まで、遺伝子編集技術は眩しい発展をしてきた。CRISPRベースの遺伝子はさみ(CRISPR-Cas9、base editor、prime editorなど)は、オフターゲットを改善し、効率を高め、多様に発展してきた。しかし、このような多くの発展にもミトコンドリア遺伝病の治療方法は限界があった。その理由は、CRISPRベースの技術は、実際に触媒するタンパク質と標的へガイドとして役立つgRNAで構成されている。しかし、タンパク質とは異なり、gRNAをミトコンドリアに送る方法の欠如で、まだミトコンドリア遺伝子を扱う

技術は、DNAにして切断を起こし、ミトコンドリアDNAを取り除くこと以外は存在しなかった。しかし、米国のDavid R. Liuグループで初めてミトコンドリアで塩基編集を引き起こすことができるDdCBEを披露した。DdCBEは、DNA二本鎖で機能するシトシン脱アミノ酵素DddAtoxを含んでいるため、DNA結合タンパク質TALEに融合してタンパク質のみを送り、塩基編集を引き起こすことができた。しかし、このようなDdCBEは、シトシンの中でもTC型のシトシンでのみ制限的に塩基編集を起こすため、実際に疾患モデルの製作または遺伝病の治療において限界点が多かった。それ故、世界初ミトコンドリアでアデニン塩基編集を起こすことができるTALEDを作り、TALEDは最大50%程度の高い効率を有し、周辺ターゲット部分で多様な種類のアデニンで塩基編集を起こした。また、UGIがある場合には、シトシンとアデニン塩基を同時に編集することができるため、ランダム変異を起こすときに有用に使用できることに対し、UGIがない場合には、シトシン塩基編集は起こらず、アデニン塩基編集のみを起こすため特異的なアデニン塩基編集技術として使用できる。また、ZFPシステムでも適用でき、核DNAに関してもアデニン塩基編集が可能であることを示した。このようなTALEDの開発は、多くのミトコンドリア遺伝病に解決策を提示し、それに該当する疾患モデルを作ることを可能にし、まだ開拓されていない多くのミトコンドリア遺伝子関連研究をすることに有用な道具になると言える。

【０４５３】

以上、本発明の内容の特定の部分を詳細に説明したところ、当業界における通常の知識を有する者にとって、このような具体的な記述は単に好ましい実施形態に過ぎず、これにより本発明の範囲が制限されるものではない点は明白である。したがって、本発明の実質的な範囲は、添付の特許請求の範囲とそれらの等価物によって定義されると言える。

【産業上の利用可能性】

【０４５４】

本発明によると、遺伝子編集時にシトシン脱アミノ酵素分割形態(split)結合表面の特定のアミノ酸残基を置換することにより、所望しないシトシン脱アミノ酵素の非選択性を低減することができる。

【０４５５】

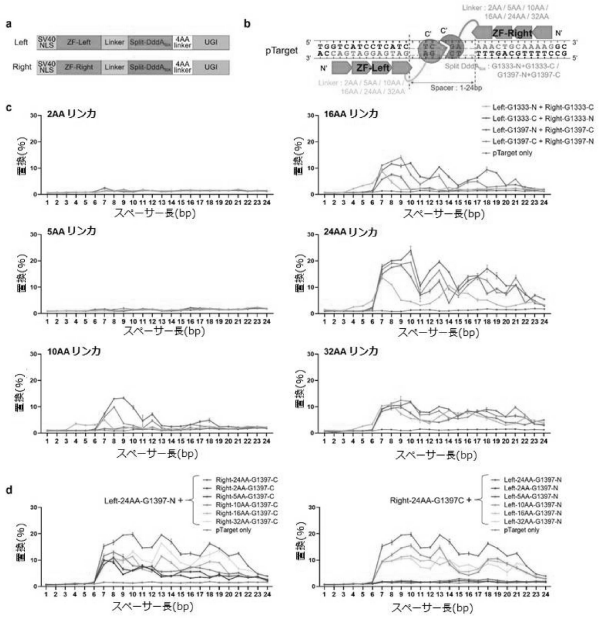
全長シトシン脱アミノ酵素に関連して、従来使用されるシトシン塩基編集(cytosine base editor)をもって編集し難しい部分の編集が可能である。従来のシトシン塩基編集手段のうち、脱アミノ酵素(deaminase)として使用するApobec1は癌誘発遺伝子として知られているので、治療目的としての使用に制約が伴われるが、今回開発された全長脱アミノ酵素はかかる問題点がないと判断される。

【０４５６】

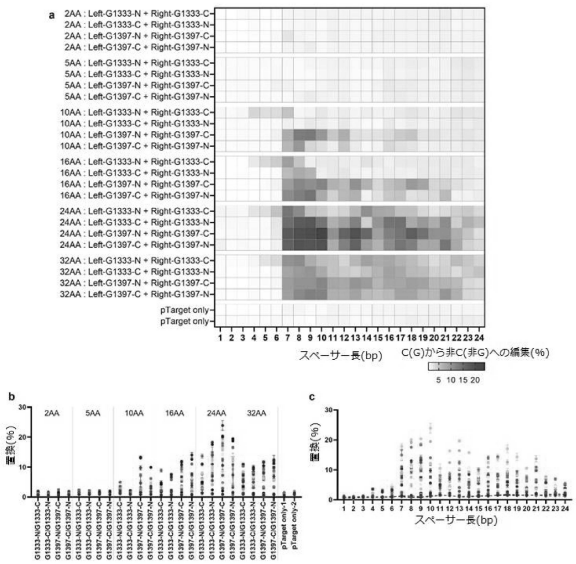
DNA結合タンパク質を含めて2.5Kb程度に小さく、AAVベクターを用いた遺伝子治療に使用可能であり、mRNA、RNP伝達が容易であり、原核生物を用いた有用物質の生産などに使用可能である。

【図面】

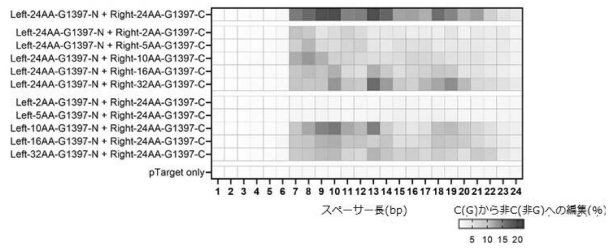
【図１】



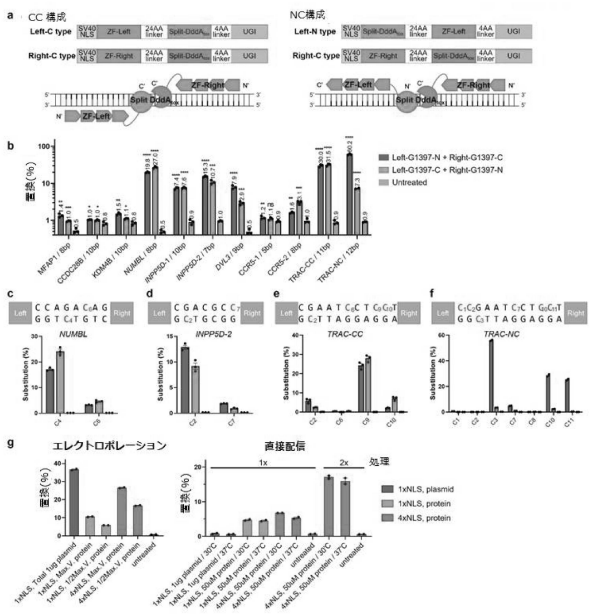
【図２】



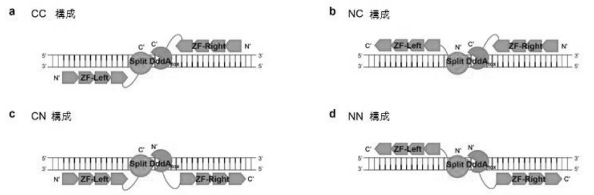
【図３】



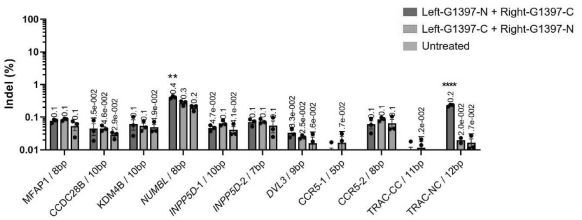
【図４】



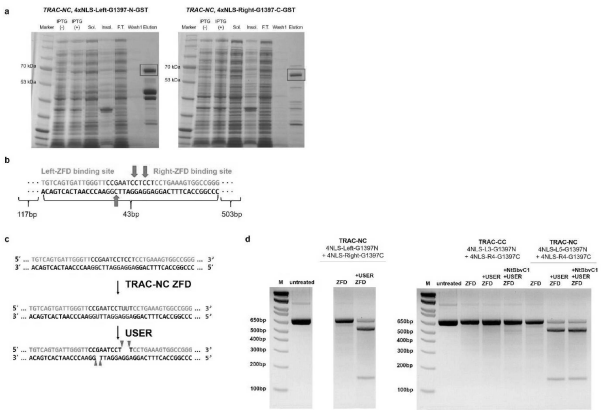
【図５】



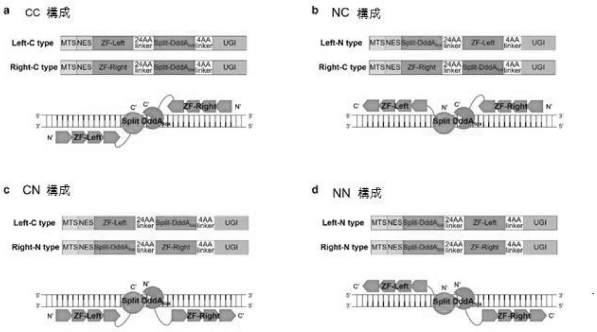
【図６】



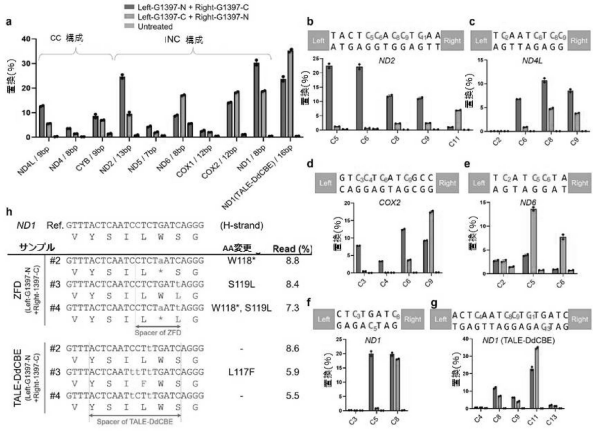
【図７】



【図８】



【図９】



【図１０】

a

MT-ND1 GTTTACTCAATC CTCTGATC$_8$ AGGGTGAGCATC
CAAATGAGTTAG GAGAC$_2$TAG TCCCACTCGTAG

	sub (%)		sub (%)
ND1-1	0.6	ND1-16	83.2
ND1-2	0.4	ND1-17	0.4
ND1-3	34.8	ND1-18	0.5
ND1-4	0.6	ND1-19	0.4
ND1-5	0.5	ND1-20	0.5
ND1-6	0.6	ND1-21	0.5
ND1-7	0.5	ND1-22	0.6
ND1-8	0.5	ND1-23	0.5
ND1-9	0.5	ND1-24	0.7
ND1-10	98.1	ND1-25	0.5
ND1-11	0.5	ND1-26	0.5
ND1-12	0.5	ND1-27	47.7
ND1-13	0.5	ND1-28	0.6
ND1-14	0.6	ND1-29	0.6
ND1-15	0.6	ND1-30	94.7

b

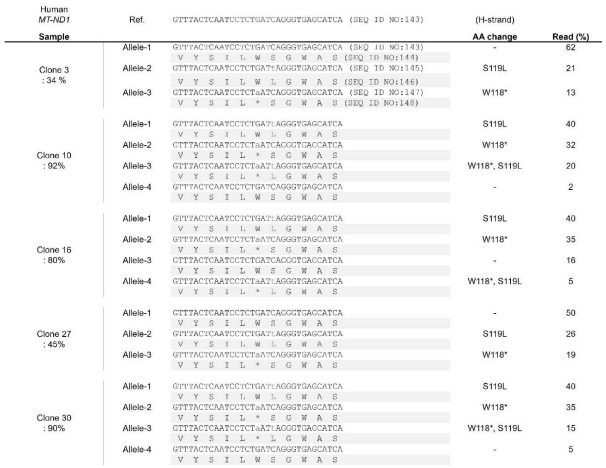
MT-ND2 CTACGCCTAATC TACTCCACCTCAA TCACACTACTCC
GATGCGGATTAG ATGAGGTGGAGTT AGTGTGATGAGG

	sub (%)		sub (%)
ND2-1	0.6	ND2-19	0.9
ND2-2	0.7	ND2-20	0.7
ND2-3	0.7	ND2-21	67.5
ND2-4	0.6	ND2-22	1.0
ND2-5	0.6	ND2-23	71.3
ND2-6	0.7	ND2-24	75.7
ND2-7	0.6	ND2-25	0.6
ND2-8	0.6	ND2-26	0.7
ND2-9	0.7	ND2-27	0.7
ND2-10	0.9	ND2-28	0.8
ND2-11	0.7	ND2-29	0.7
ND2-12	0.8	ND2-30	0.6
ND2-13	0.6	ND2-31	45.1
ND2-14	0.7	ND2-32	57.9
ND2-15	0.6	ND2-33	52.6
ND2-16	0.7	ND2-34	25.5
ND2-17	0.7	ND2-35	0.9
ND2-18	0.6	ND2-36	0.8

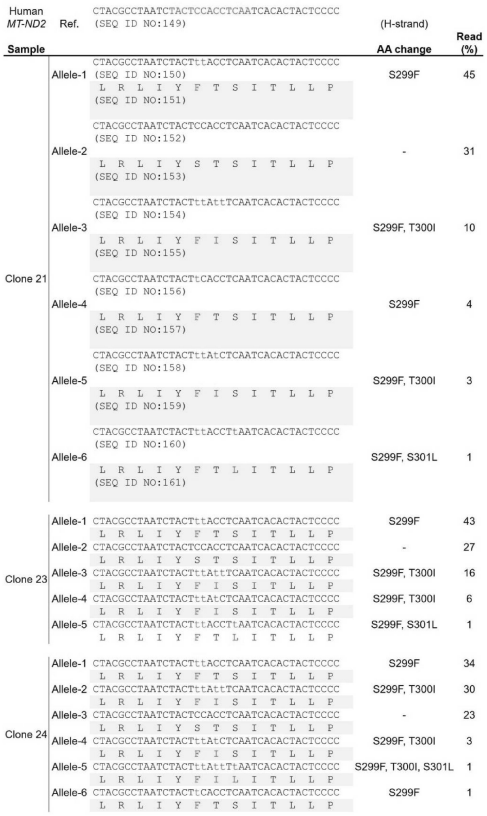
c

	ND1 region sub (%)	ND2 region sub (%)
untreated-1	0.5	0.8
untreated-2	0.6	0.7
untreated-3	0.6	0.7
untreated-4	1.1	0.8
untreated-5	0.5	0.7
untreated-6	0.5	0.7
untreated-7	0.6	0.9
untreated-8	0.5	0.8
untreated-9	0.5	0.9
untreated-10	0.6	0.8
untreated-11	0.5	0.9
untreated-12	0.6	0.7
untreated-13	0.5	0.7
untreated-14	0.6	0.7
untreated-15	0.5	0.7
untreated-16	0.5	0.8
untreated-17	0.5	0.9
untreated-18	0.4	0.6

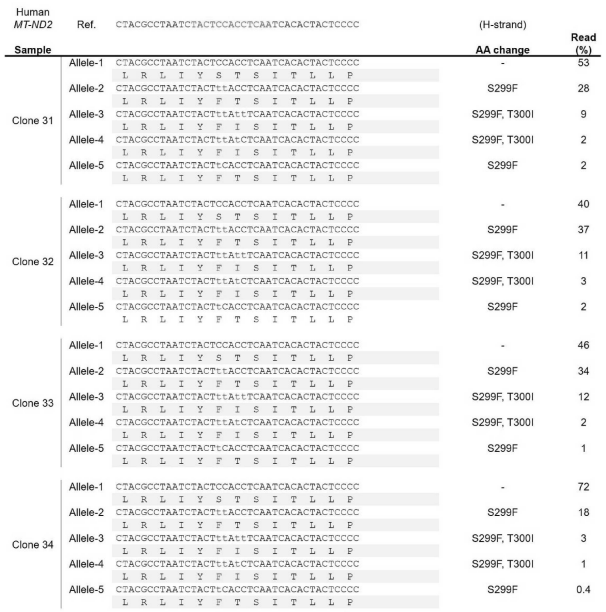
【図１１】



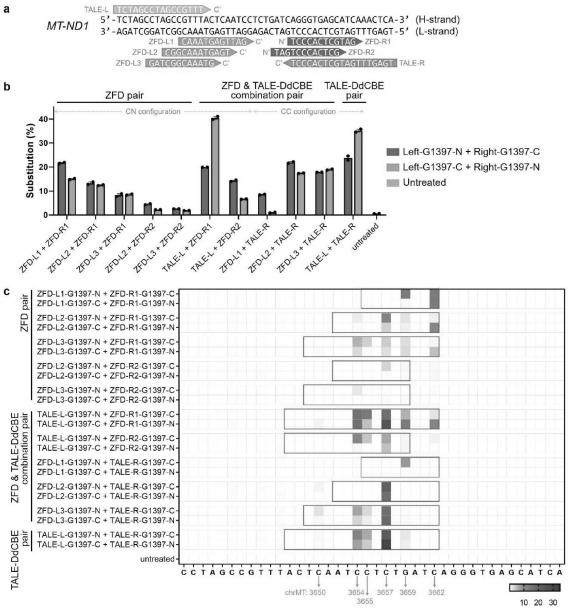
【図１２ａ】



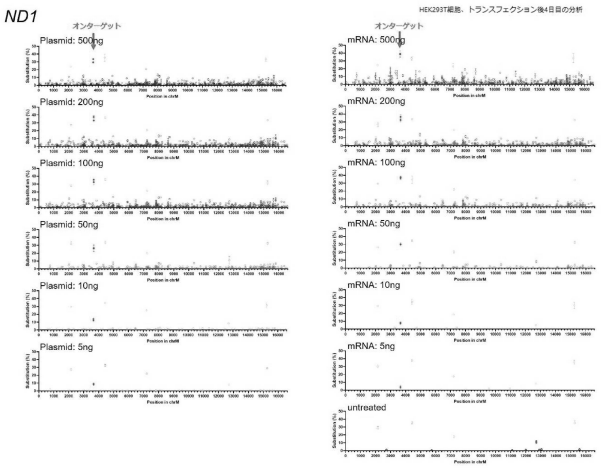
【図１２ｂ】



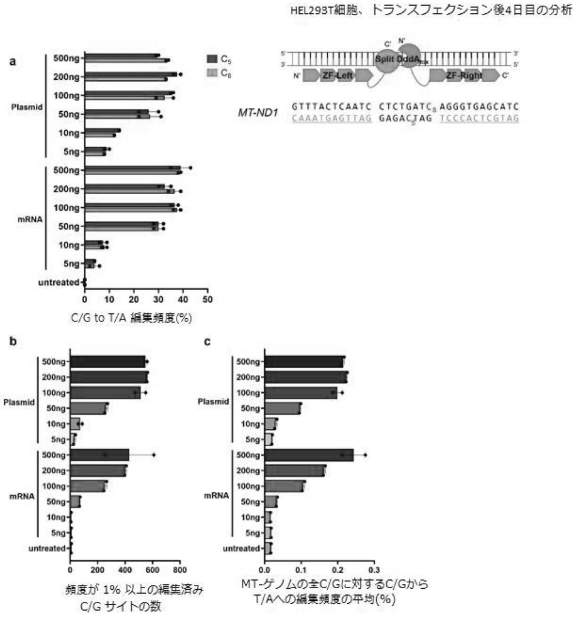
【図１３】



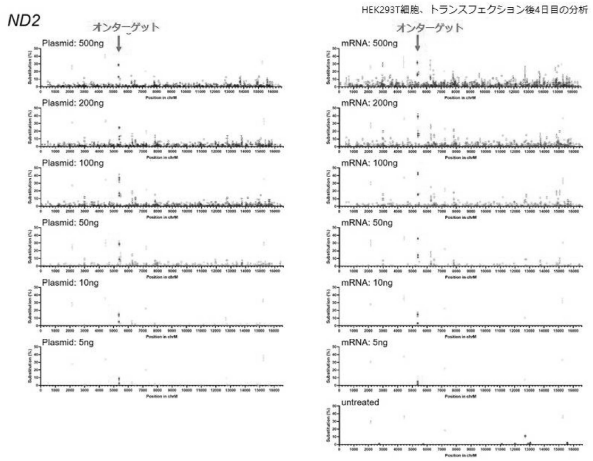
【図１４】



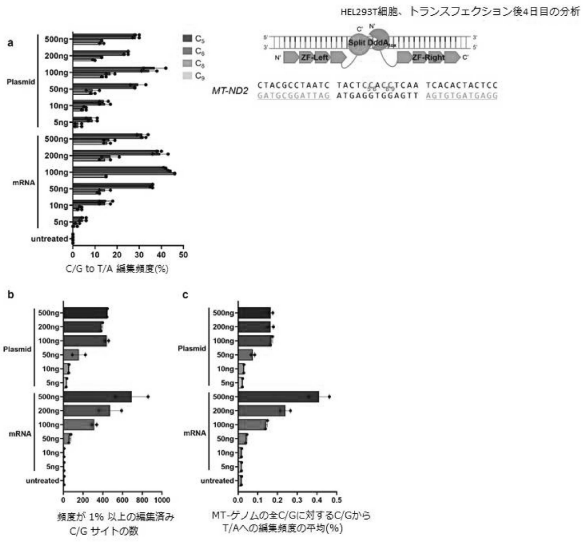
【図１５】



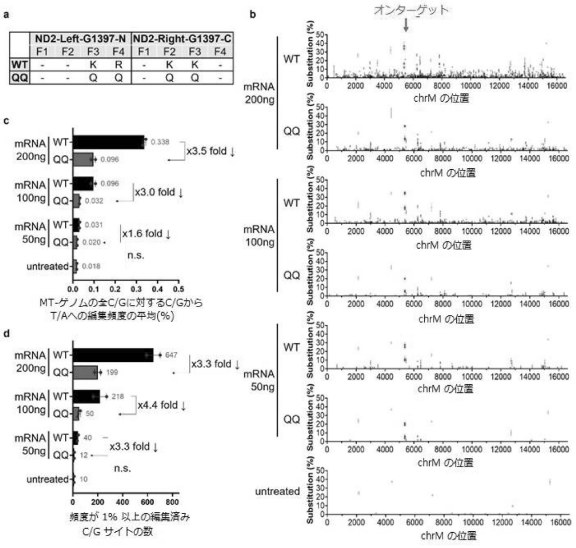
【図１６】



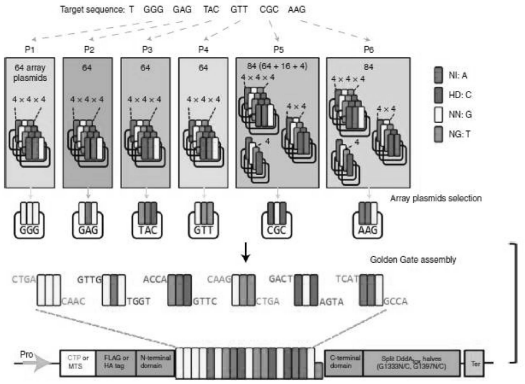
【図１７】



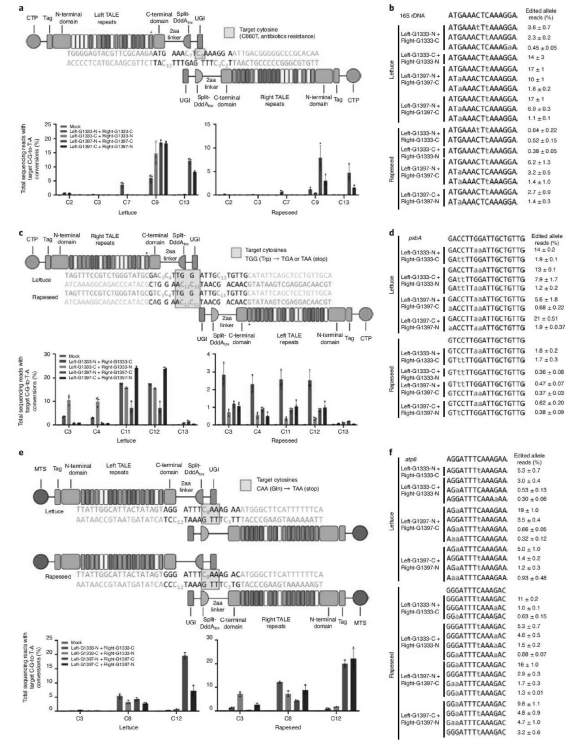
【図１８】



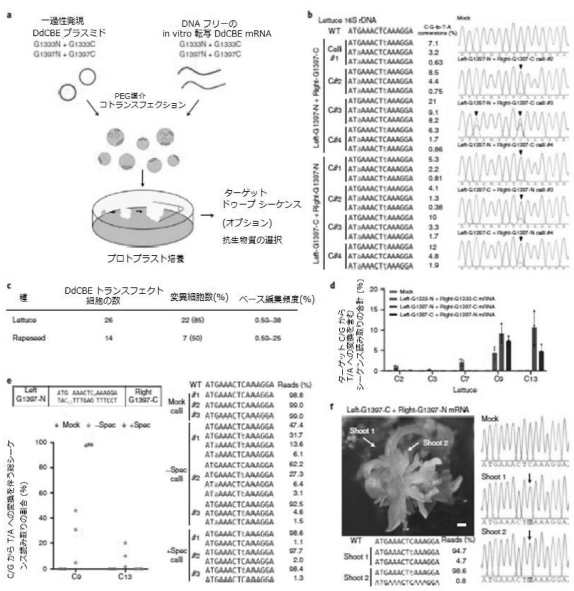
【図１９】



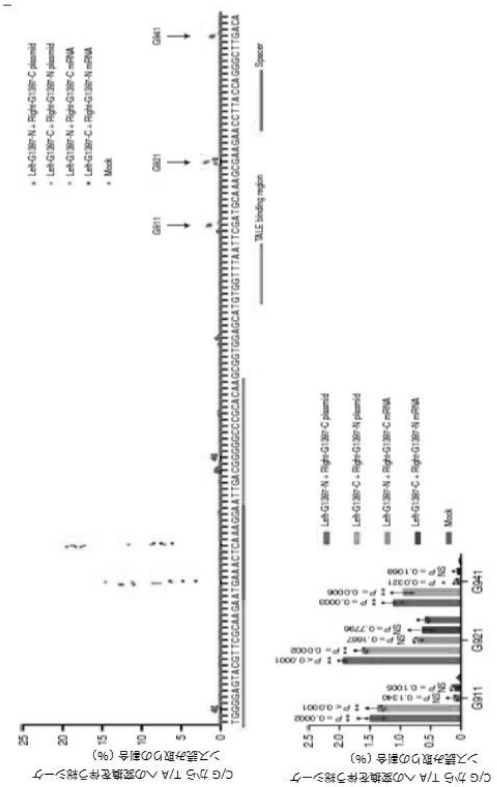
【図２０】



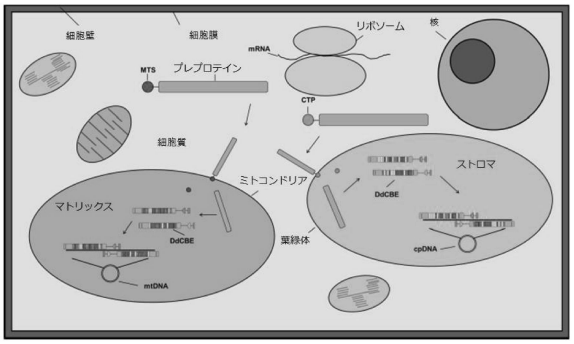
【図２１】



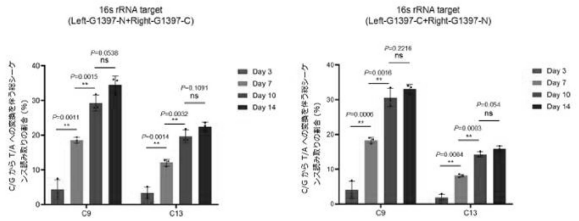
【図２２】



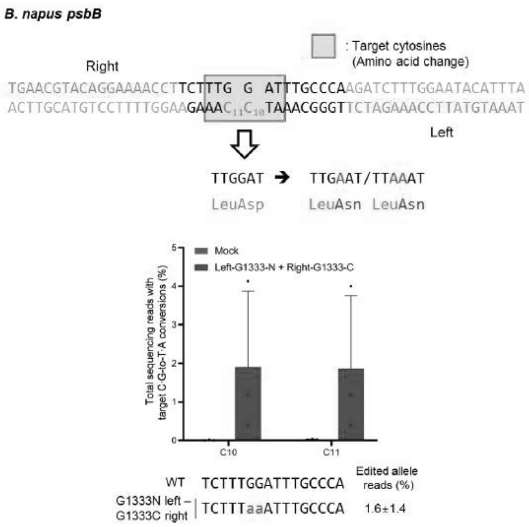
【図２３】



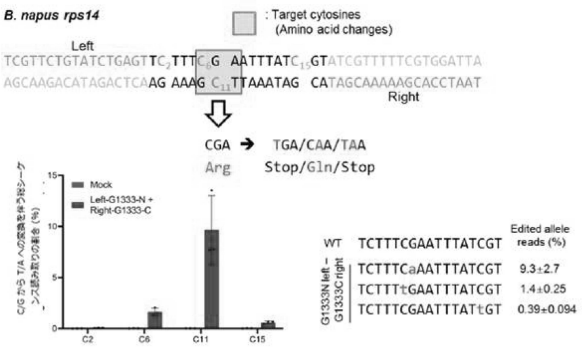
【図２４】



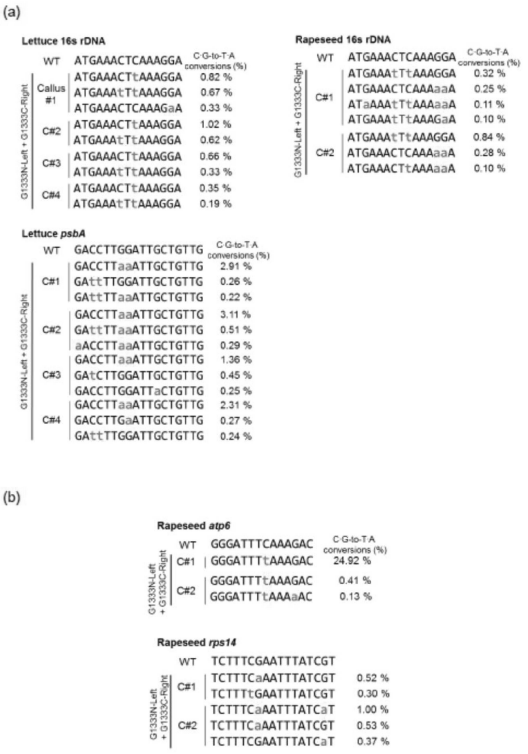
【図２５】



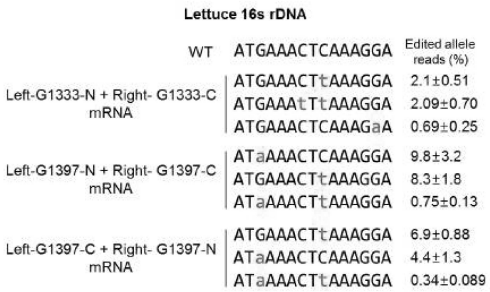
【図２６】



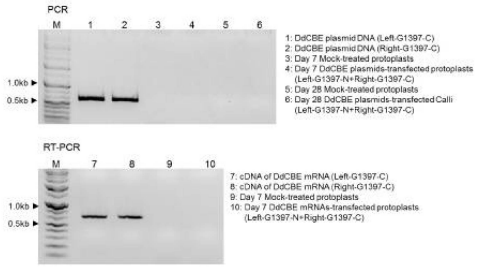
【図２７】



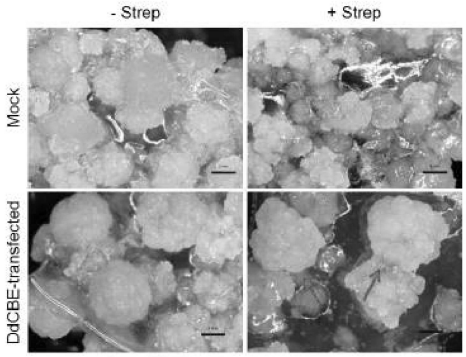
【図２８】



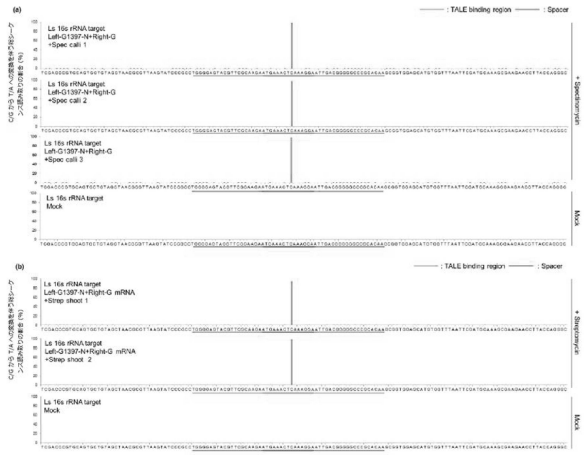
【図２９】



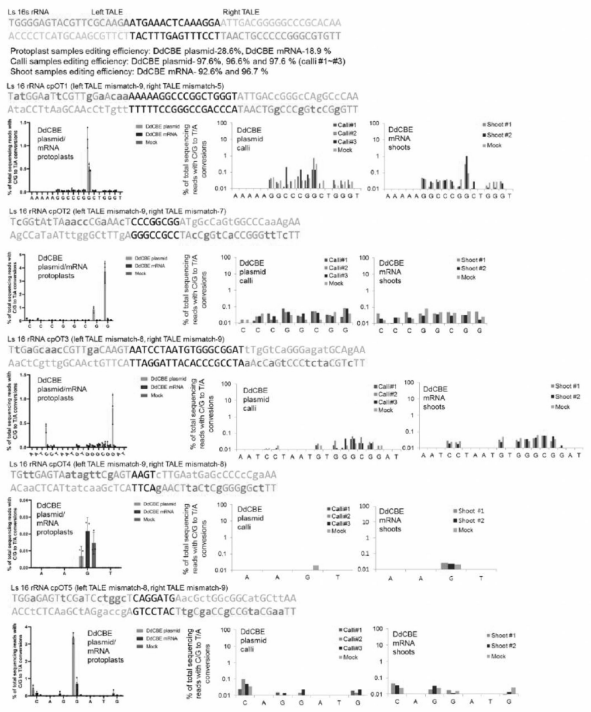
【図３０】



【図３１】



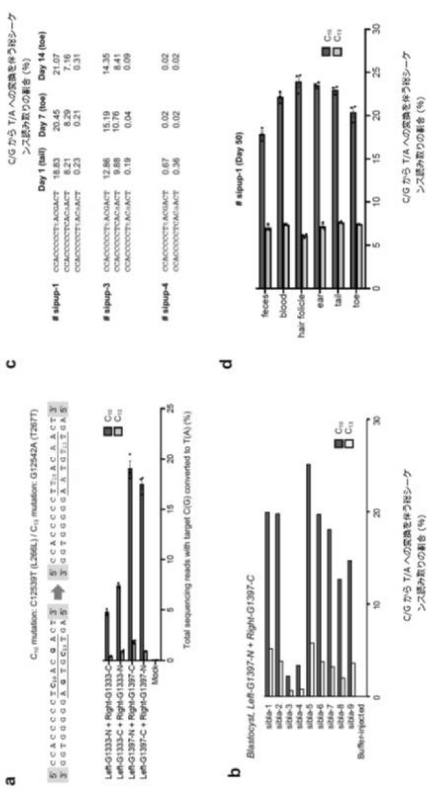
【図３２】



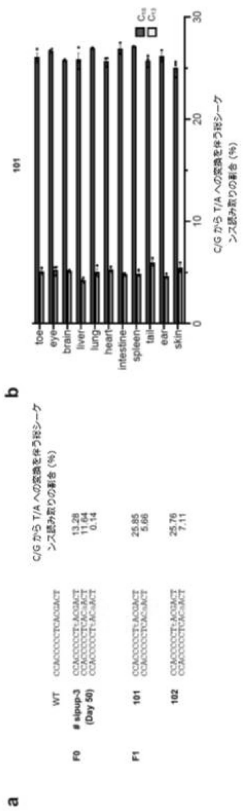
【図３３】



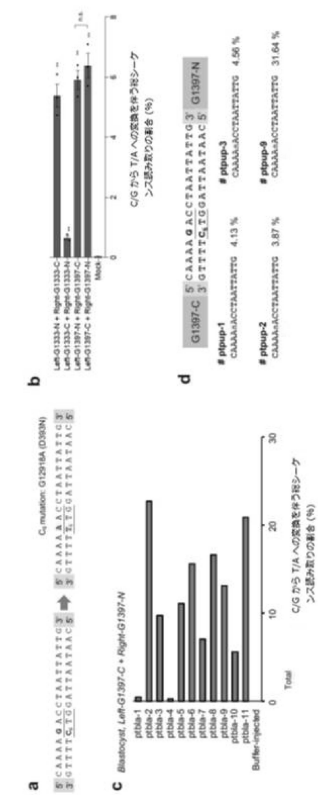
【図３４】



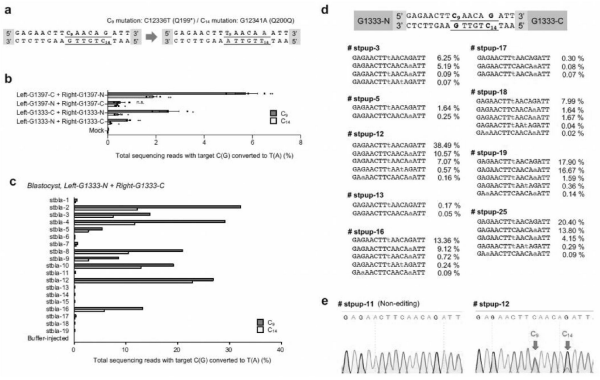
【図３５】



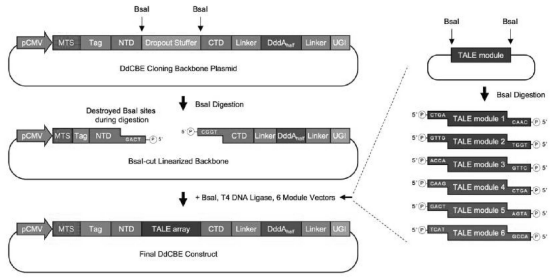
【図３６】



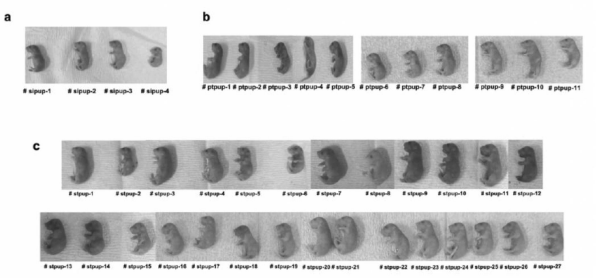
【図３７】



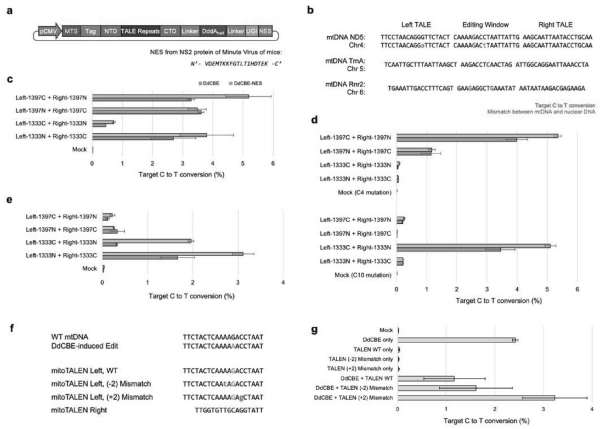
【図３８】



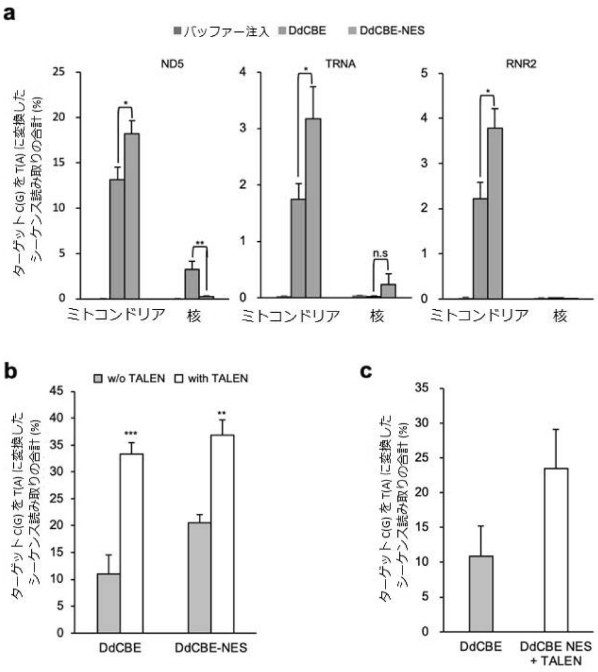
【図３９】



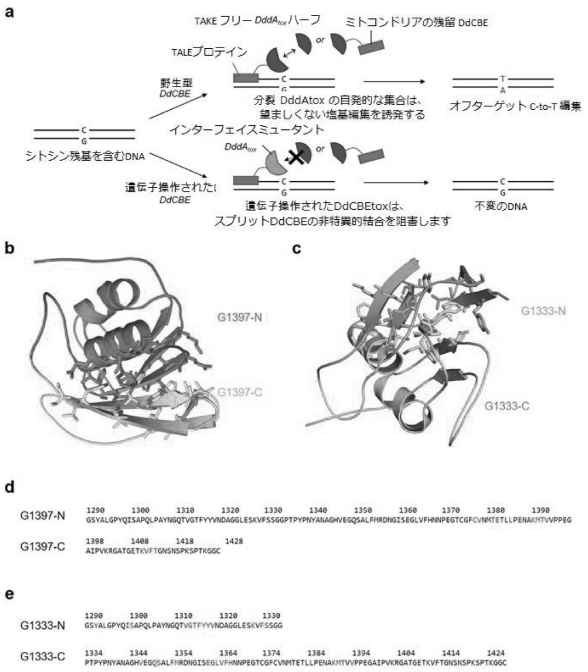
【図４０】



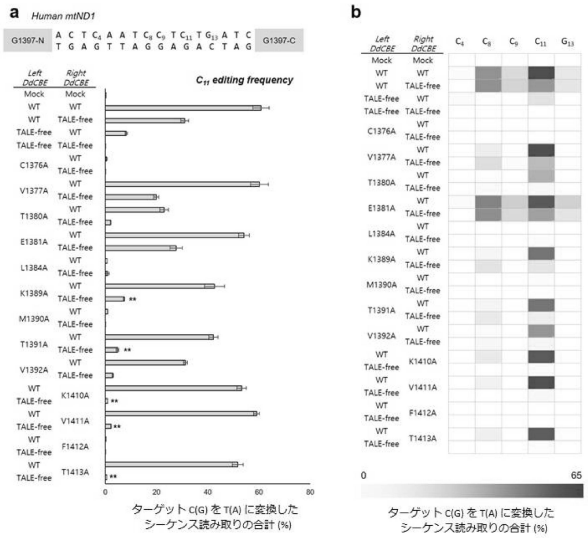
【図４１】



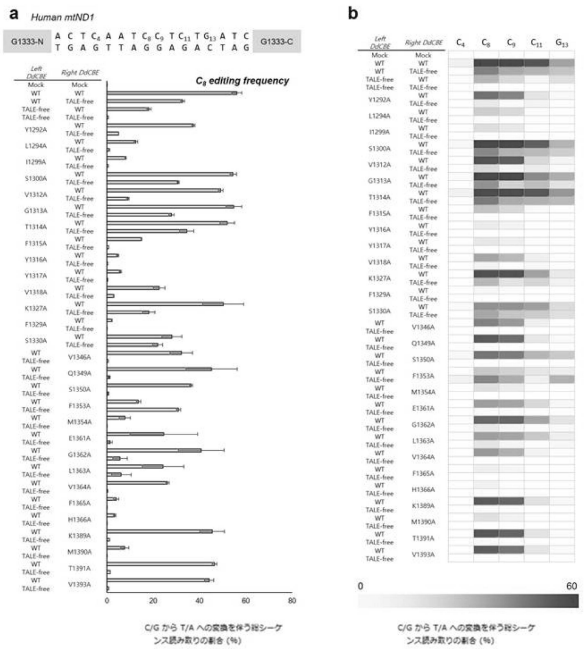
【図４２】



【図４３】



【図４４】



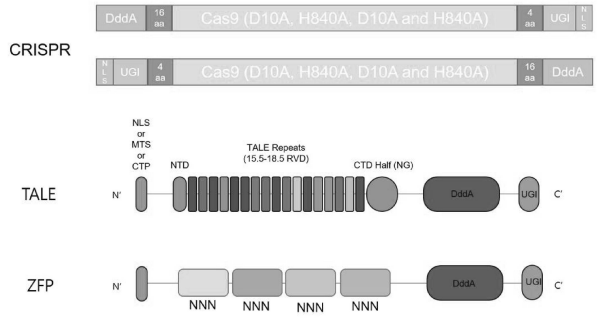
【図４５】

```
Wild Type       GSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGGPTPYPNYANAGHVEGQ  60
AAKAK           GSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGGPTPYPNYANAGHVEGQ  60
AAKAA           GSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGGPTPYPNYANAGHVEGQ  60
AAAAA           GSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGGPTPYPNYANAGHVEGQ  60
AAAAK           GSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGGPTPYPNYANAGHVEGQ  60
KAAAA           GSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGGPTPYPNYANAGHVEGQ  60
A1341D_KRKKA    GSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGGPTPYPNYDNAGHVEGQ  60

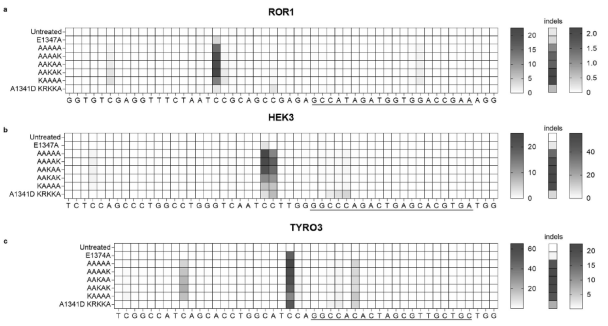
Wild Type       SALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKMTVVPPEGAIPVKRGATGET  120
AAKAK           SALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKMTVVPPEGAIPVAAGATGET  120
AAKAA           SALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKMTVVPPEGAIPVAAGATGET  120
AAAAA           SALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKMTVVPPEGAIPVAAGATGET  120
AAAAK           SALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKMTVVPPEGAIPVAAGATGET  120
KAAAA           SALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKMTVVPPEGAIPVKAGATGET  120
A1341D_KRKKA    SALFMRDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKMTVVPPEGAIPVKRGATGET  120


Wild Type       KVFTGNSNSPKSPTKGGC    138
AAKAK           KVFTGNSNSPASPTKGGC    138
AAKAA           KVFTGNSNSPASPTAGGC    138
AAAAA           AVFTGNSNSPASPTAGGC    138
AAAAK           AVFTGNSNSPASPTKGGC    138
KAAAA           AVFTGNSNSPASPTAGGC    138
A1341D_KRKKA    KVFTGNSNSPKSPTAGGC    138
```

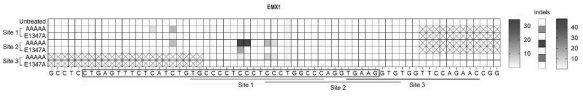
【図４６】



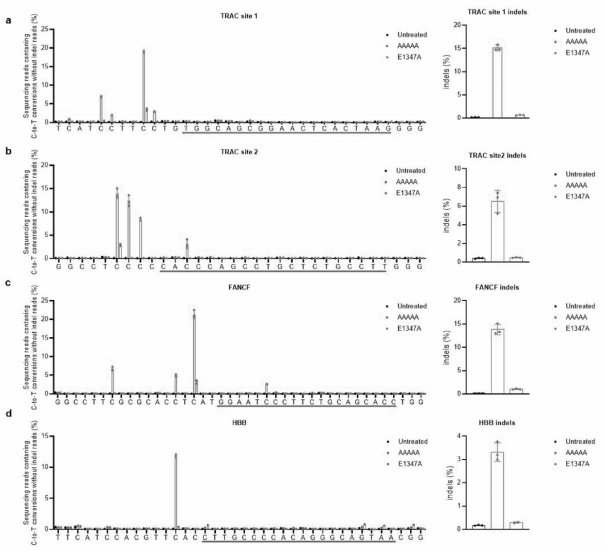
【図４７】



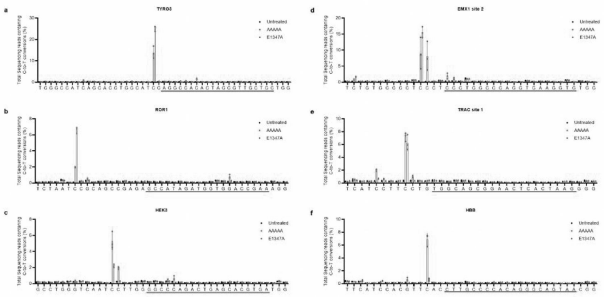
【図４８】



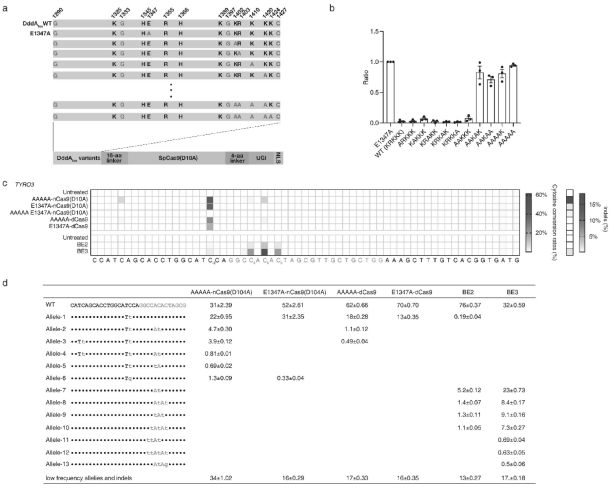
【図４９】



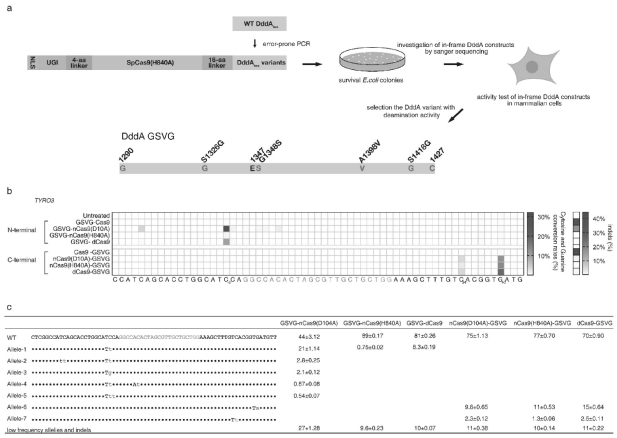
【図５０】



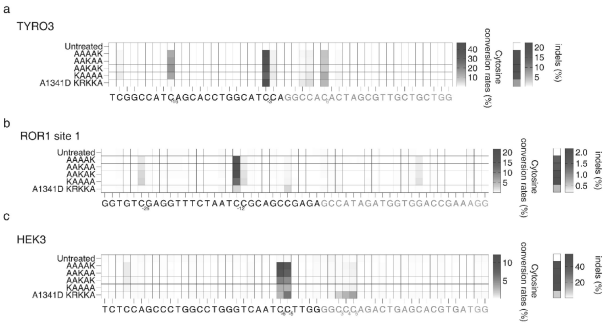
【図５１】



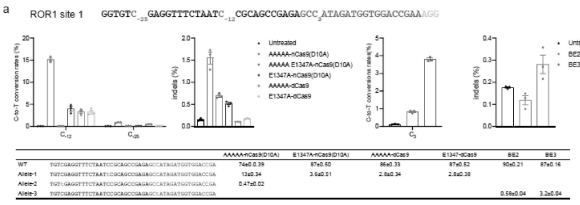
【図５２】



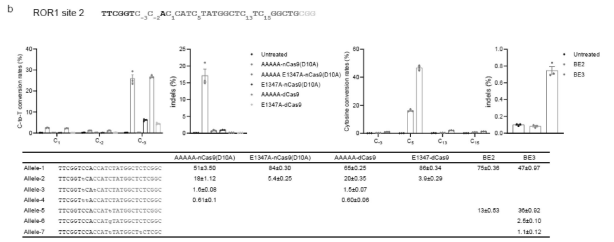
【図５３】



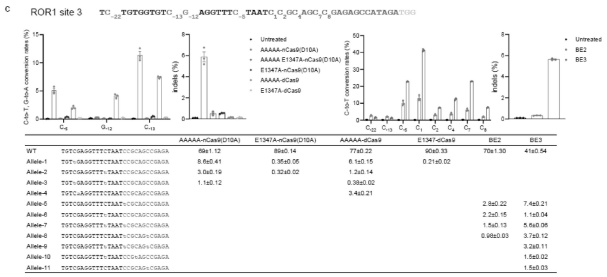
【図５４ａ】



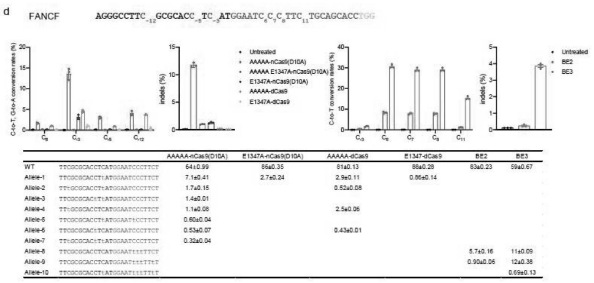
【図５４ｂ】



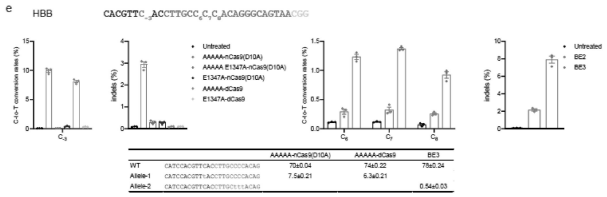
【図５４ｃ】



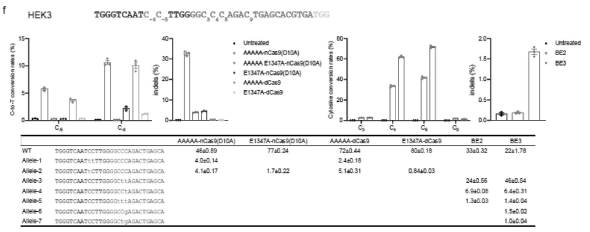
【図５４ｄ】



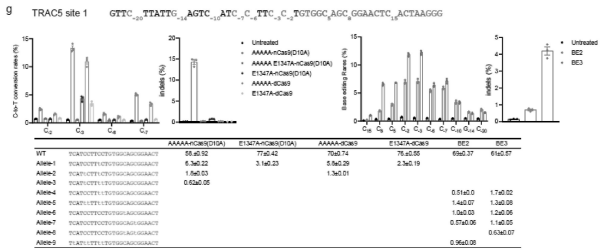
【図５４ｅ】



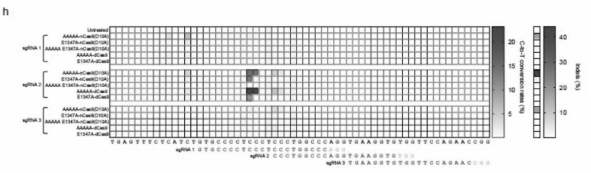
【図５４ｆ】



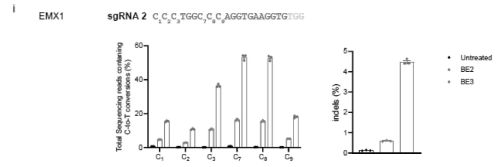
【図５４ｇ】



【図５４ｈ】



【図５４ｉ】

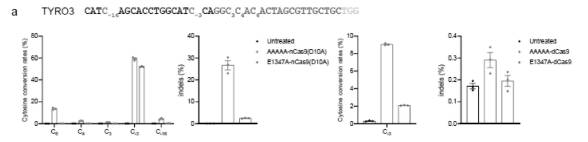


【図５４ｊ】

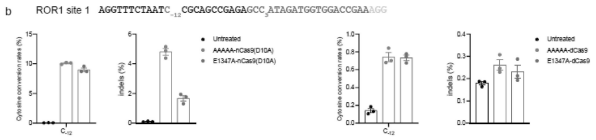
j

		AAAAA-nCas9(D10A)	E1347A-nCas9(D10A)	AAAAA-dCas9	E1347A-dCas9	BE2	BE3
WT	TCTCATCTGTGCCCCTCCCTGGCCCAGGTG	34±1.32	65±0.44	62±0.36	73±1.02	55±0.28	22±0.66
Allele-1	TCTCATCTGTGCCCCTC-CCTGGCCCAGGTG	6.7±0.16		12±0.55			
Allele-2	TCTCATCTGTGCCCCTC-CCTGGCCCAGGTG	1.7±0.11		0.78±0.02			
Allele-3	TCTCATCTGTGCCCCTC-CCTGGCCCAGGTG	1.2±0.05		1.6±0.02			
Allele-4	TCTCATCTGTGCCCCTC-CCTGGCCCAGGTG	0.87±0.06		1.4±0.03			
Allele-5	TCTCATCTGTGCCCCTC-CCTCCCTGGCCCAGGTG		10±0.56	1.4±0.05	7.8±0.71		
Allele-6	TCTCATCTGTGCCCCTCCCTCCCTGGCCCAGGTG		0.95±0.02		0.60±0.01		
Allele-7	TCTCATCTGTGCCCCTCCCTCCCTGGCCCAGGTG					3.2±0.14	10±0.09
Allele-8	TCTCATCTGTGCCCCTCCCTCCCTGGCCCAGGTG					2.6±0.14	6.2±0.17
Allele-9	TCTCATCTGTGCCCCTCCCTCCCTGGCCCAGGTG					1.3±0.06	2.6±0.11
Allele-10	TCTCATCTGTGCCCCTCCCTCCCTGGCCCAGGTG					1.1±0.06	3.7±0.10
Allele-11	TCTCATCTGTGCCCCTCCCTCCCTGGCCCAGGTG						3.5±0.14
Allele-12	TCTCATCTGTGCCCCTCCCTCCCTGGCCCAGGTG						2.4±0.15
Allele-13	TCTCATCTGTGCCCCTCCCTCCCTGGCCCAGGTG						2.3±0.18

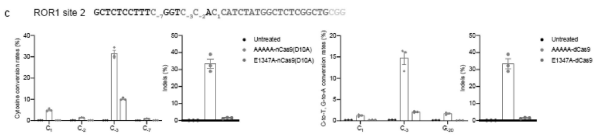
【図５５ａ】



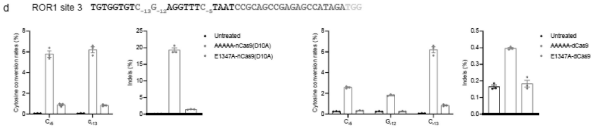
【図５５ｂ】



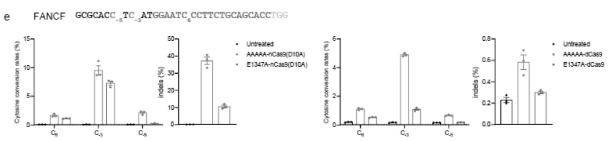
【図５５ｃ】



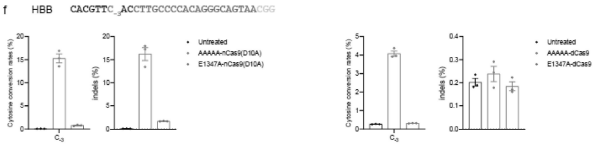
【図５５ｄ】



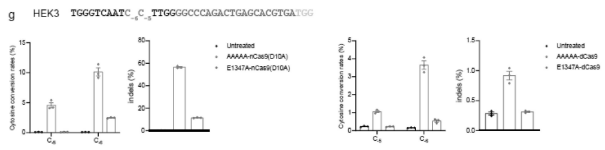
【図５５ｅ】



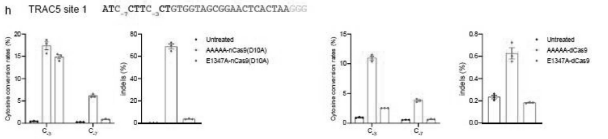
【図５５ｆ】



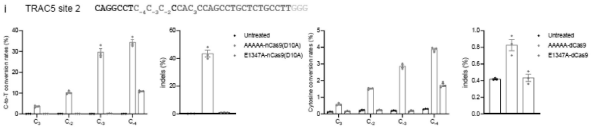
【図５５ｇ】



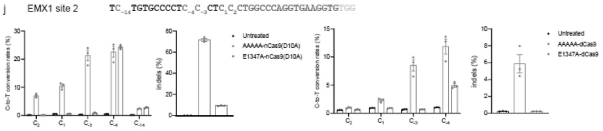
【図５５ｈ】



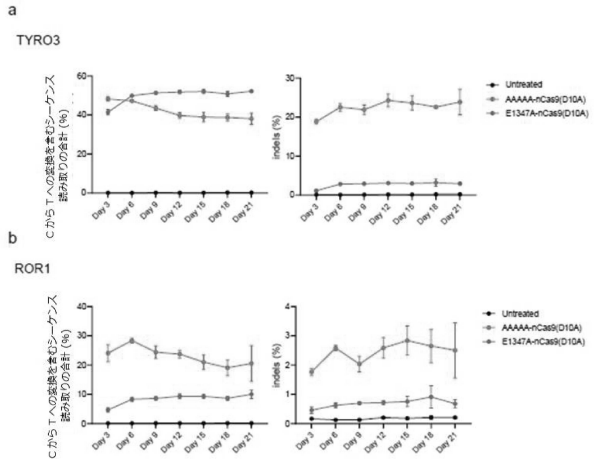
【図５５ｉ】



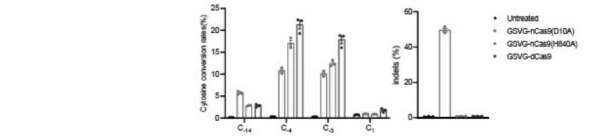
【図５５ｊ】



【図５６】

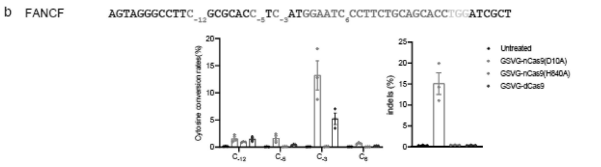


【図５７ａ】



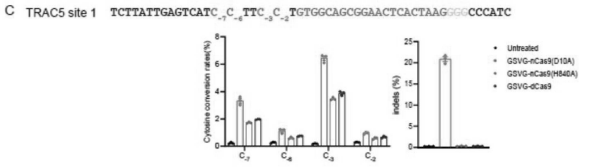
		GSVG-nCas9(D10A)	GSVG-nCas9(H840A)	GSVG-dCas9
WT	TCTCATCTGTGCCCCTCCCTCCCTGGCCCAGGTG	33±1.3	67±1.2	63±1.5
Allele-1	TCTCATcTGTGCCCCTssCTCCCTGGCCCAGGTG	3.9±0.13	1.1±0.09	1.5±0.26
Allele-2	TCTCATCTGTGCCCCTssCTCCCTGGCCCAGGTG	3.5±0.31	8.5±0.59	12±0.77
Allele-3	TCTCATCTGTGCCCCTcCCTCCCTGGCCCAGGTG		3.6±0.36	2.6±0.16
Allele-4	TCTCATcTGTGCCCCTCCCTCCCTGGCCCAGGTG		0.66±0.02	
Allele-5	TCTCATCTGTGCCCCTcCCTCCCTGGCCCAGGTG			0.68±0.25

【図５７ｂ】



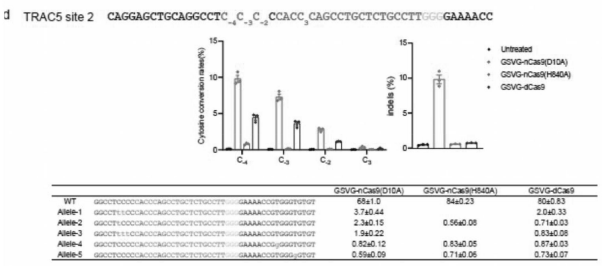
		GSVG-nCas9(D10A)	GSVG-nCas9(H840A)	GSVG-dCas9
WT	TTCGCGCACCTCATGGAATCCCTTCT	84±4.9	90±0.08	84±1.9
Allele-1	TTCGCGCACCTcATGGAATCCCTTCT	8.9±1.7		4.0±1.4
Allele-2	TTCGCGCACcTcATGGAATCCCTTCT	0.95±0.23		
Allele-3	TTcGCGCACCTcATGGAATCCCTTCT	0.81±0.18	0.78±0.11	0.32±0.04
Allele-4	TTTGCGCACCTgATGGAATCCCTTCT	0.47±0.17		
Allele-5	TTcGCGCACCTCATGGAATCCCTTCT	0.38±0.11		0.90±0.33
Allele-6	TTCGCGCACCTcATGGAATcCCTTCT	0.38±0.04		

【図５７ｃ】



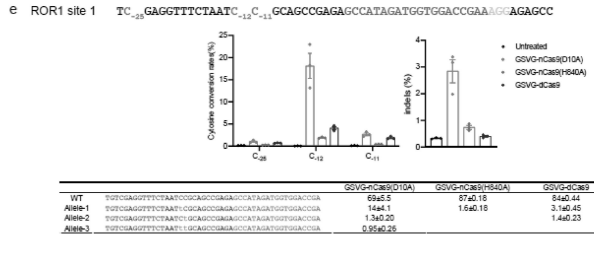
		GSVG-nCas9(D10A)	GSVG-nCas9(H840A)	GSVG-dCas9
WT	GAGTCATCCTTCCTGTGGCAGCGGAA	81±1.4	82±0.38	81±0.26
Allele-1	GAGTCATCCTTcCTGTGGCAGCGGAA	2.9±0.18	2.1±0.08	2.0±0.21
Allele-2	GAGTCATcCTTcCTGTGGCAGCGGAA	1.2±0.19		0.53±0.02
Allele-3	GAGTCATcCTTCCTGTGGCAGCGGAA		0.53±0.06	0.43±0.11

【図５７ｄ】



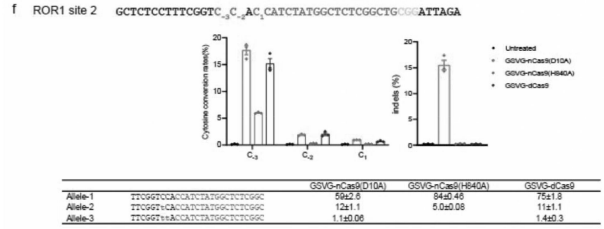
		GSVG-nCas9(D10A)	GSVG-nCas9(H840A)	GSVG-dCas9
WT	GGCCTCCCCCACCCAGCCTGCTCTGCCTTGGGGAAAACCGTGGGTGTGT	68±1.0	84±0.23	80±0.83
Allele-1	GGCCTt1CCCACCCAGCCTGCTCTGCCTTGGGGAAAACCGTGGGTGTGT	3.7±0.44		2.0±0.33
Allele-2	GGCCTcCCCCACCCAGCCTGCTCTGCCTTGGGGAAAACCGTGGGTGTGT	2.3±0.15	0.56±0.08	0.71±0.03
Allele-3	GGCCTt1CCACCCAGCCTGCTCTGCCTTGGGGAAAACCGTGGGTGTGT	1.9±0.22		0.83±0.08
Allele-4	GGCCTCCCCCACCCAGCCTGCTCTGCCTTGGGGAAAACCGgGGGTGTGT	0.82±0.12	0.83±0.05	0.87±0.03
Allele-5	GGCCTCCCCCACCCAGCCTGCTCTGCCTTGGGGAAAACCGTGGGgGTGT	0.59±0.09	0.71±0.06	0.73±0.07

【図５７ｅ】



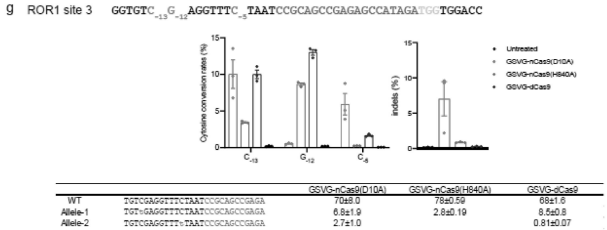
		GSVG-nCas9(D10A)	GSVG-nCas9(H840A)	GSVG-dCas9
WT	TGTCGAGGTTTCTAATCCGCAGCCGAGAGCCATAGATGGTGGACCGA	69±5.5	87±0.18	84±0.44
Allele-1	TGTCGAGGTTTCTAATcCGCAGCCGAGAGCCATAGATGGTGGACCGA	14±4.1	1.6±0.18	3.1±0.45
Allele-2	TGTCGAGGTTTCTAATcTGCAGCCGAGAGCCATAGATGGTGGACCGA	1.3±0.20		1.4±0.23
Allele-3	TGTCGAGGTTTCTAATt1GCAGCCGAGAGCCATAGATGGTGGACCGA	0.95±0.26		

【図５７ｆ】



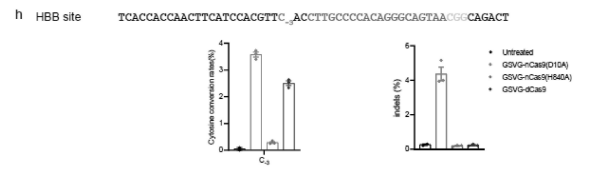
		GSVG-nCas9(D10A)	GSVG-nCas9(H840A)	GSVG-dCas9
Allele-1	TTCGGTCCACCATCTATGGCTCTCGGC	59±2.6	84±0.46	75±1.8
Allele-2	TTCGGTcCACCATCTATGGCTCTCGGC	12±1.1	5.0±0.08	11±1.1
Allele-3	TTCGGTttACCATCTATGGCTCTCGGC	1.1±0.06		1.4±0.3

【図５７ｇ】



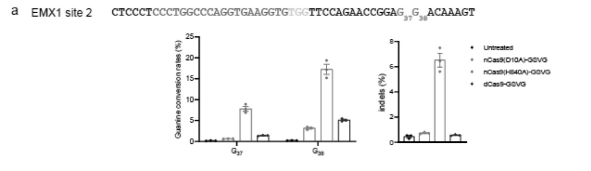
		GSVG-nCas9(D10A)	GSVG-nCas9(H840A)	GSVG-dCas9
WT	TGTCGAGGTTTCTAATCCGCAGCCGAGA	70±8.0	78±0.59	68±1.6
Allele-1	TGTcGAGGTTTCTAATCCGCAGCCGAGA	6.8±1.9	2.8±0.19	8.5±0.8
Allele-2	TGTCGAGGTTTcTAATCCGCAGCCGAGA	2.7±1.0		0.81±0.07
Allele-3	TGTcGAGGTTTcTAATCCGCAGCCGAGA	1.6±0.95		
Allele-4	TGTCGAGGTTTgTAATCCGCAGCCGAGA	0.4±0.13		
Allele-5	TGTCaAGGTTTCTAATCCGCAGCCGAGA		7.5±0.27	11±0.6

【図５７ｈ】



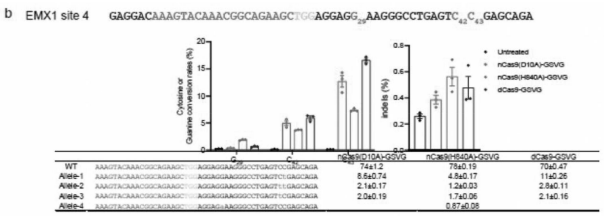
		GSVG-nCas9(D10A)	GSVG-nCas9(H840A)	GSVG-dCas9
WT	ACCACCCCCAACTTCATCCACGTTCACCTTGCCCCACAG	80±0.34	88±0.35	85±0.46
Allele-1	ACCACCCCCAACTTCATCCACGTTcACCTTGCCCCACAG	2.9±0.17		2.2±0.13
Allele-2	ACCCCCCCCAACTTCATCCACGTTCACCTTGCCCCACAG		0.50±0.06	

【図５８ａ】



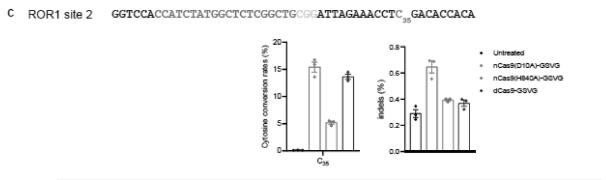
		nCas9(D10A)-GSVG	nCas9(H840A)-GSVG	dCas9-GSVG
WT	CCCTGGCCCAGGTGAAGGTGTGGGTTCCAGAACCGGAGcACAAAGT	84±0.35	57±1.6	83±0.21
Allele-1	CCCTGGCCCAGGTGAAGGTGTGGGTTCCAGAACCGGAGcACAAAGT	2.1±0.15	8.0±0.45	3.1±0.07
Allele-2	CCCTGGCCCAGGTGAAGGTGTGGGTTCCAGAACCGGAaacACAAAGT		6.5±0.49	1.1±0.05

【図５８ｂ】



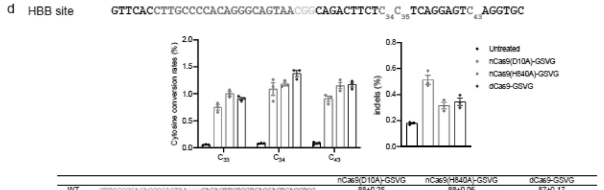
		nCas9(D10A)-GSVG	nCas9(H840A)-GSVG	dCas9-GSVG
WT	AAAGTACAAACGGCAGAAGC---AGGAGGAAGGGCCTGAGTCCGAGCAGA	74±1.2	78±0.19	70±0.47
Allele-1	AAAGTACAAACGGCAGAAGC---AGGAGGAAGGGCCTGAGTCCGAGCAGA	8.5±0.74	4.8±0.17	11±0.26
Allele-2	AAAGTACAAACGGCAGAAGC---AGGAGGAAGGGCCTGAGTTCGAGCAGA	2.1±0.17	1.2±0.03	2.8±0.11
Allele-3	AAAGTACAAACGGCAGAAGC---AGGAGGAAGGGCCTGAGTCCGAGCAGA	2.0±0.19	1.7±0.06	2.1±0.16
Allele-4	AAAGTACAAACGGCAGAAGC---AGGAGGAAGGGCCTGAGTCCGAGCAGA		0.67±0.08	

【図５８ｃ】



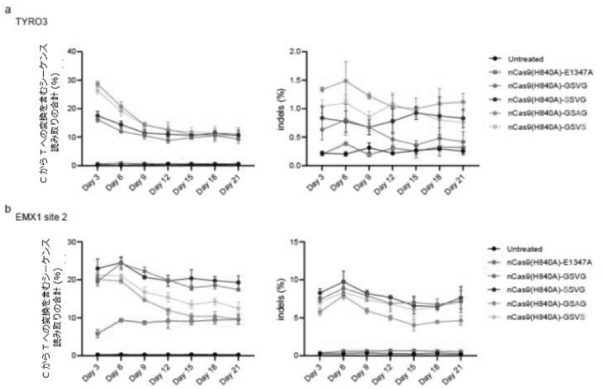
		nCas9(D10A)-GSVG	nCas9(H840A)-GSVG	dCas9-GSVG
WT	CCATCTATGGCTCTCGGCTG---ATTAGAAACCTCGACACCACA	76±0.84	86±0.32	78±0.39
Allele-1	CCATCTATGGCTCTCGGCTG---ATTAGAAACCTCGACACCACA	14±0.84	4.5±0.24	12±0.41

【図５８ｄ】

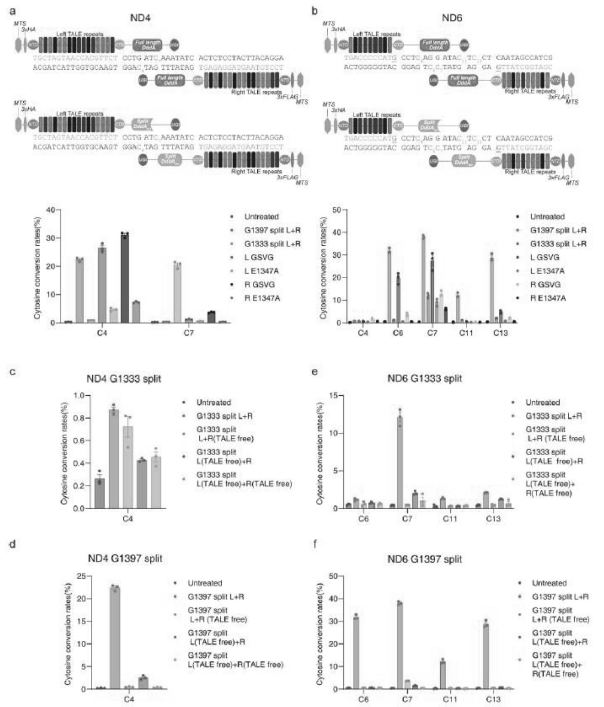


		nCas9(D10A)-GSVG	nCas9(H840A)-GSVG	dCas9-GSVG
WT	CTTGCCCCACAGGGCAGTAA---CAGACTTCTCCTCAGGAGTCAGGTGC	88±0.25	88±0.06	87±0.17
Allele-1	CTTGCCCCACAGGGCAGTAA---CAGACTTCTCCTCAGGAGTCAGGTGC	0.59±0.04	0.60±0.07	0.80±0.04
Allele-2	CTTGCCCCACAGGGCAGTAA---CAGACTTCTCTCTCAGGAGTCAGGTGC	0.42±0.05	0.53±0.03	0.52±0.03
Allele-3	CTTGCCCCACAGGGCAGTAA---CAGACTTCTCTCTCAGGAGTCAGGTGC			0.46±0.03

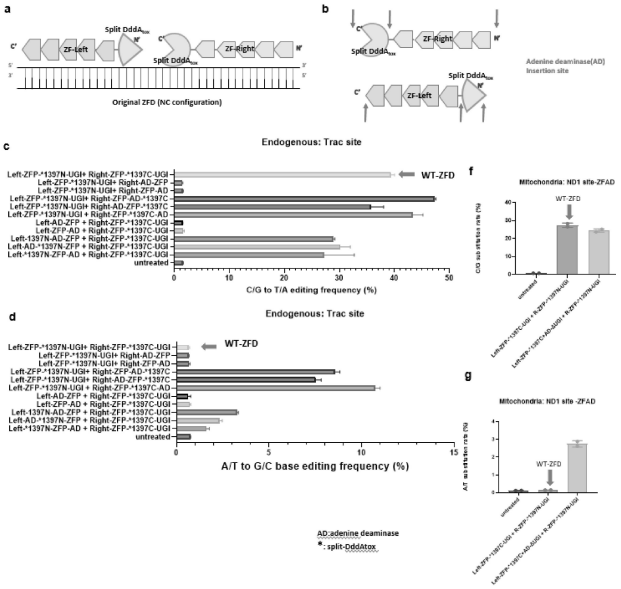
【図５９】



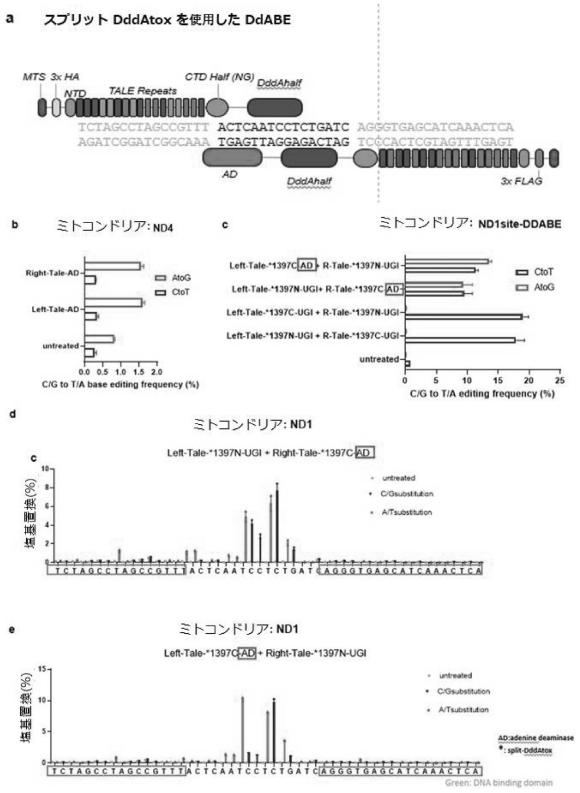
【図６０】



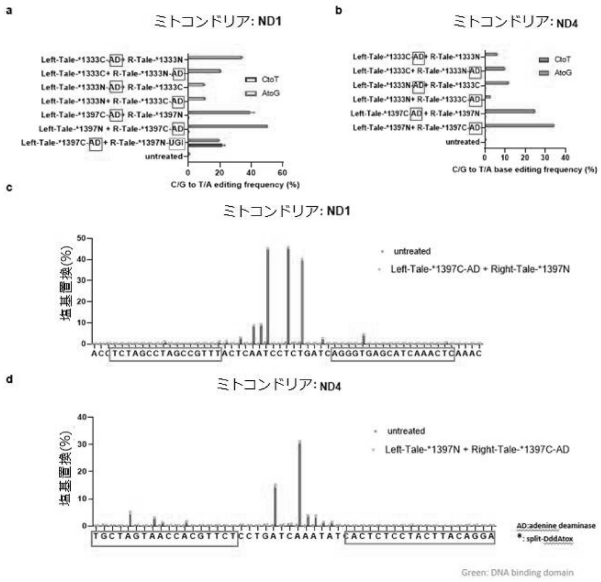
【図６１】



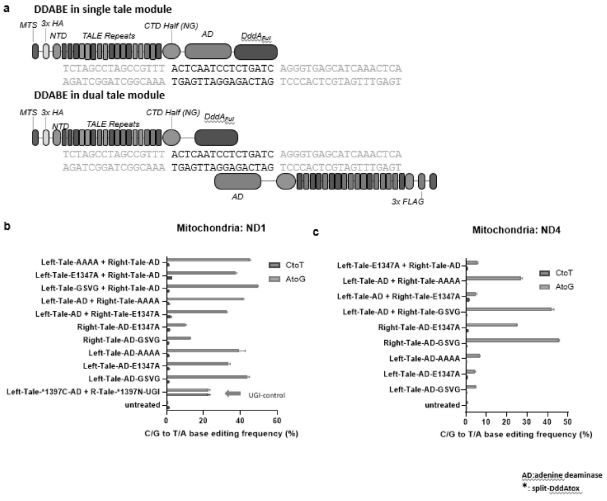
【図６２】



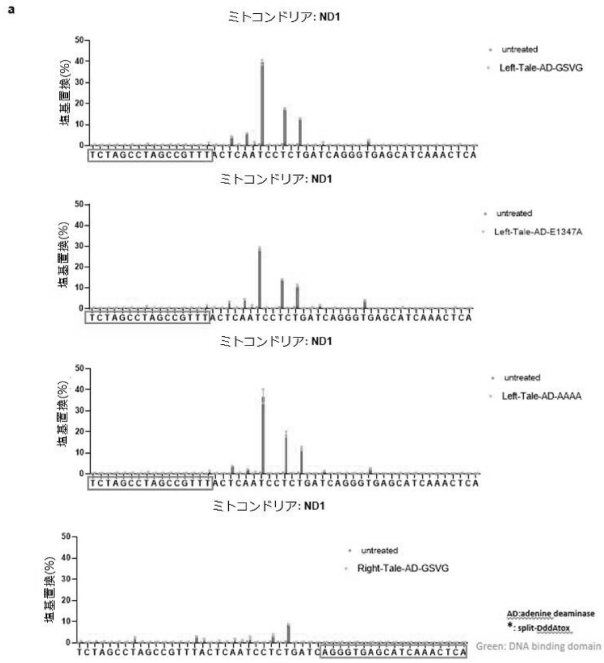
【図６３】



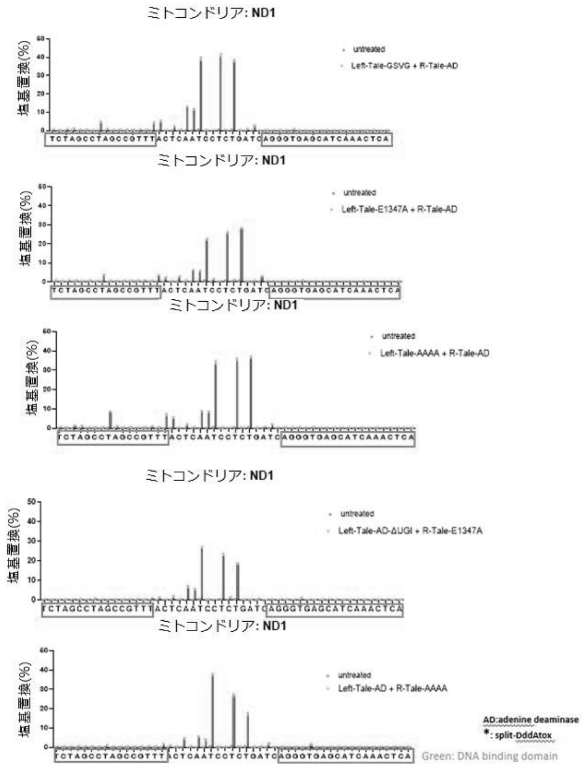
【図６４】



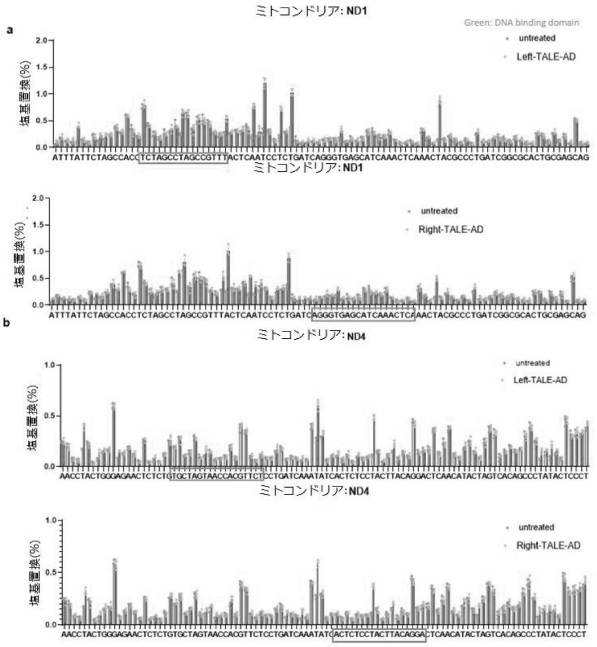
【図６５】



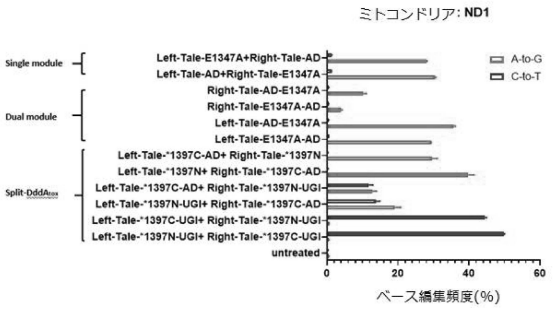
【図６６】



【図６７】



【図６８】



【配列表】

<u>20235419900000001.app</u>

【国際調査報告】

INTERNATIONAL SEARCH REPORT

International application No.
PCT/KR2021/012872

A. CLASSIFICATION OF SUBJECT MATTER

C12N 9/78(2006.01)i; **C12N 15/52**(2006.01)i; **C12N 9/22**(2006.01)i; **C12N 15/10**(2006.01)i; **C12N 15/90**(2006.01)i; **C12N 15/82**(2006.01)i; **C12N 15/62**(2006.01)i

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

C12N 9/78(2006.01); C12N 15/10(2006.01); G16B 30/00(2019.01)

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

Korean utility models and applications for utility models: IPC as above
Japanese utility models and applications for utility models: IPC as above

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

eKOMPASS (KIPO internal) & keywords: 시토신 탈아미노효소(cytosine deaminase), 아데닌 탈아미노효소(adenine cytosine deaminase), 융합단백질(fusion protein), DNA 결합 단백질(DNA binding protein), 분할체(split), 전장(full-length)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
	MOK, B. Y. et al. A bacterial cytidine deaminase toxin enables CRISPR-free mitochondrial base editing. Nature. July 2020, vol. 583, no. 7817, pp. 631-637. HHS Public Access Author manuscript version. inner pp. 1-43.	
X	See abstract; page 4, paragraphs [0003]-[0004]; page 5, paragraph [0001]; page 6, paragraphs [0001]-[0002] and [0005]; page 26; and figures 1-3.	1-6,14-18
Y		19-21,49-51
A		7-13
	KR 10-2019-0034628 A (PRESIDENT AND FELLOWS OF HARVARD COLLEGE) 02 April 2019 (2019-04-02)	
Y	See claims 1-5 and 114-116; and paragraphs [0003], [0009], [0013] and [0816].	19-21,49-51
	KR 10-2019-0044157 A (INDUSTRY-ACADEMIC COOPERATION FOUNDATION GYEONGSANG NATIONAL UNIVERSITY) 30 April 2019 (2019-04-30)	
Y	See abstract; claims 1-3; and paragraphs [0004]-[0005] and [0016].	19-21,49-51

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

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

Date of the actual completion of the international search	Date of mailing of the international search report
23 December 2021	**23 December 2021**
Name and mailing address of the ISA/KR **Korean Intellectual Property Office** **Government Complex-Daejeon Building 4, 189 Cheongsa-ro, Seo-gu, Daejeon 35208** Facsimile No. **+82-42-481-8578**	Authorized officer Telephone No.

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

INTERNATIONAL SEARCH REPORT

International application No.
PCT/KR2021/012872

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	KR 10-2018-0029937 A (TOOLGEN INCORPORATED et al.) 21 March 2018 (2018-03-21) See claims 1-8.	1-21,49-51
A	KR 10-2019-0025565 A (SIGMA-ALDRICH CO., LLC) 11 March 2019 (2019-03-11) See entire document.	1-21,49-51

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

INTERNATIONAL SEARCH REPORT

International application No.
PCT/KR2021/012872

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

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:
   a. ☑ forming part of the international application as filed:
      ☑ in the form of an Annex C/ST.25 text file.
      ☐ on paper or in the form of an image file.
   b. ☐ furnished together with the international application under PCT Rule 13*ter*.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.
   c. ☐ furnished subsequent to the international filing date for the purposes of international search only:
      ☐ in the form of an Annex C/ST.25 text file (Rule 13*ter*.1(a)).
      ☐ on paper or in the form of an image file (Rule 13*ter*.1(b) and Administrative Instructions, Section 713).

2. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No.
PCT/KR2021/012872

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

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

1. ☑ Claims Nos.: **45-48,52-54**
because they relate to subject matter not required to be searched by this Authority, namely:

Claims 45-48 and 52-54 pertain to a method for treatment of the human body by surgery or therapy, and thus pertain to a subject matter on which the International Searching Authority is not required to carry out an international search under the provisions of PCT Article 17(2)(a)(i) and PCT Rule 39.1(iv).

2. ☑ Claims Nos.: **23,25,30,32,36,43-44**
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

Claims 23, 25, 30, 32, 36 and 43-44 refer to multiple dependent claims not meeting the requirement of PCT Rule 6.4(a).

3. ☑ Claims Nos.: **22,24,26-29,31,33-35,37-42,45-48**
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/KR2021/012872

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
KR 10-2019-0034628 A	02 April 2019	AU 2017-306676 A1	28 February 2019
		CA 3032699 A1	08 February 2018
		CN 110214183 A	06 September 2019
		EP 3494215 A1	12 June 2019
		GB 2568182 A	08 May 2019
		IL 264565 A	28 February 2019
		JP 2019-536425 A	19 December 2019
		SG 11201900907 A	27 February 2019
		US 10113163 B2	30 October 2018
		US 10947530 B2	16 March 2021
		US 2018-0073012 A1	15 March 2018
		US 2019-0093099 A1	28 March 2019
		US 2021-0317440 A1	14 October 2021
		WO 2018-027078 A1	08 February 2018
KR 10-2019-0044157 A	30 April 2019	None	
KR 10-2018-0029937 A	21 March 2018	EP 3530737 A1	28 August 2019
		JP 2019-526271 A	19 September 2019
		KR 10-2026421 B1	27 September 2019
		US 2020-0131536 A1	30 April 2020
		WO 2018-052247 A1	22 March 2018
KR 10-2019-0025565 A	11 March 2019	AU 2017-274145 A1	22 November 2018
		AU 2017-274145 B2	23 July 2020
		AU 2020-244497 A1	29 October 2020
		AU 2020-244497 B2	12 November 2020
		AU 2021-200636 A1	04 March 2021
		BR 112018074531 A2	19 March 2019
		BR 112018074531 B1	19 January 2021
		CA 3026321 A1	07 December 2017
		CN 109983124 A	05 July 2019
		DK 3272867 T3	02 December 2019
		DK 3604527 T3	31 May 2021
		EP 3272867 A1	24 January 2018
		EP 3272867 B1	25 September 2019
		EP 3604527 A1	05 February 2020
		EP 3604527 B1	12 May 2021
		EP 3907286 A1	10 November 2021
		ES 2760477 T3	14 May 2020
		GB 2552861 A	14 February 2018
		GB 2552861 B	15 May 2019
		GB 2578802 A	27 May 2020
		GB 2578802 B	16 September 2020
		GB 2582731 A	30 September 2020
		GB 2582731 B	30 December 2020
		IL 262854 A	31 December 2018
		IL 262854 B	30 June 2020
		IL 275244 A	30 July 2020
		JP 2019-517795 A	27 June 2019
		JP 2021-121193 A	26 August 2021
		JP 6878468 B2	26 May 2021
		LT 3604527 T	25 June 2021

Form PCT/ISA/210 (patent family annex) (July 2019)

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/KR2021/012872

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
		PL	3272867	T3	31 January 2020
		PL	3604527	T3	06 September 2021
		PT	3272867	T	04 December 2019
		PT	3604527	T	02 June 2021
		SG	11201810003	A	28 December 2018
		US	10266851	B2	23 April 2019
		US	2017-0349913	A1	07 December 2017
		US	2019-0169650	A1	06 June 2019
		US	2019-0169651	A1	06 June 2019
		WO	2017-209809	A1	07 December 2017

Form PCT/ISA/210 (patent family annex) (July 2019)

국 제 조 사 보 고 서

국제출원번호
PCT/KR2021/012872

A. 발명이 속하는 기술분류(국제특허분류(IPC))

C12N 9/78(2006.01)i; **C12N 15/52**(2006.01)i; **C12N 9/22**(2006.01)i; **C12N 15/10**(2006.01)i; **C12N 15/90**(2006.01)i; **C12N 15/82**(2006.01)i; **C12N 15/62**(2006.01)i

B. 조사된 분야

조사된 최소문헌(국제특허분류를 기재)

C12N 9/78(2006.01); C12N 15/10(2006.01); G16B 30/00(2019.01)

조사된 기술분야에 속하는 최소문헌 이외의 문헌

한국등록실용신안공보 및 한국공개실용신안공보: 조사된 최소문헌란에 기재된 IPC
일본등록실용신안공보 및 일본공개실용신안공보: 조사된 최소문헌란에 기재된 IPC

국제조사에 이용된 전산 데이터베이스(데이터베이스의 명칭 및 검색어(해당하는 경우))

eKOMPASS(특허청 내부 검색시스템) & 키워드: 시토신 탈아미노효소(cytosine deaminase), 아데닌 탈아미노효소(adenine cytosine deaminase), 융합단백질(fusion protein), DNA 결합 단백질(DNA binding protein), 분할체(split), 전장(full-length)

C. 관련 문헌

카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항
X	MOK, B. Y. 등, `A bacterial cytidine deaminase toxin enables CRISPR-free mitochondrial base editing`, Nature, 2020.07, 583권, 7817호, 페이지 631-637, HHS Public Access Author manuscript version 내부페이지 1-43 초록; 페이지 4, 단락 3-4; 페이지 5, 단락 1; 페이지 6, 단락 1-2, 5; 페이지 26; 도면 1-3	1-6,14-18
Y		19-21,49-51
A		7-13
Y	KR 10-2019-0034628 A (프레지던트 엔드 펠로우즈 오브 하바드 칼리지) 2019.04.02 청구항 1-5, 114-116; 단락 [0003], [0009], [0013], [0816]	19-21,49-51
Y	KR 10-2019-0044157 A (경상대학교산학협력단) 2019.04.30 요약; 청구항 1-3; 단락 [0004]-[0005], [0016]	19-21,49-51
A	KR 10-2018-0029937 A (주식회사 툴젠 등) 2018.03.21 청구항 1-8	1-21,49-51

☑ 추가 문헌이 C(계속)에 기재되어 있습니다. ☑ 대응특허에 관한 별지를 참조하십시오.

* 인용된 문헌의 특별 카테고리:

"A" 특별히 관련이 없는 것으로 보이는 일반적인 기술수준을 정의한 문헌
"D" 본 국제출원에서 출원인이 인용한 문헌
"E" 국제출원일보다 빠른 출원일 또는 우선일을 가지나 국제출원일 이후에 공개된 선출원 또는 특허 문헌
"L" 우선권 주장에 의문을 제기하는 문헌 또는 다른 인용문헌의 공개일 또는 다른 특별한 이유(이유를 명시)를 밝히기 위하여 인용된 문헌
"O" 구두 개시, 사용, 전시 또는 기타 수단을 언급하고 있는 문헌
"P" 우선일 이후에 공개되었으나 국제출원일 이전에 공개된 문헌
"T" 국제출원일 또는 우선일 후에 공개된 문헌으로, 출원과 상충하지 않으며 발명의 기초가 되는 원리나 이론을 이해하기 위해 인용된 문헌
"X" 특별한 관련이 있는 문헌. 해당 문헌 하나만으로 청구된 발명의 신규성 또는 진보성이 없는 것으로 본다.
"Y" 특별한 관련이 있는 문헌. 해당 문헌이 하나 이상의 다른 문헌과 조합하는 경우로 그 조합이 당업자에게 자명한 경우 청구된 발명은 진보성이 없는 것으로 본다.
"&" 동일한 대응특허문헌에 속하는 문헌

국제조사의 실제 완료일	국제조사보고서 발송일
2021년12월23일(23.12.2021)	**2021년12월23일(23.12.2021)**

ISA/KR의 명칭 및 우편주소	심사관
대한민국 특허청 **(35208) 대전광역시 서구 청사로 189, 4동 (둔산동, 정부대전청사)**	**한인호**
팩스 번호 **+82-42-481-8578**	전화번호 **+82-42-481-5373**

서식 PCT/ISA/210 (두 번째 용지) (2019년 7월)

국제조사보고서

국제출원번호
PCT/KR2021/012872

C. 관련 문헌

카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항
A	KR 10-2019-0025565 A (시그마-알드리치 컴퍼니., 엘엘씨) 2019.03.11 전체 문헌	1-21,49-51

서식 PCT/ISA/210 (두 번째 용지) (2019년 7월)

국 제 조 사 보 고 서

국제출원번호
PCT/KR2021/012872

제1기재란 핵산염기 및/또는 아미노산 서열(첫 번째 용지의 1.c의 계속)

1. 국제출원에 개시된 핵산염기 및/또는 아미노산 서열과 관련하여, 국제조사는 다음에 기초하여 수행되었습니다.

   a. ☑ 아래의 형태로 출원시 국제출원의 일부를 구성하는 서열목록
      - ☑ 부록 C/ST.25 텍스트 파일
      - ☐ 서면 혹은 이미지 파일

   b. ☐ PCT 규칙 13의3.1(a)에 따라 국제출원과 함께 국제조사만을 목적으로 부록 C/ST.25 텍스트 파일의 형태로 제출된 서열목록

   c. ☐ 국제조사만을 목적으로 국제출원일 이후에 아래 형태로 제출된 서열목록
      - ☐ 부록 C/ST.25 텍스트 파일 (규칙 13의3.1(a))
      - ☐ 서면 혹은 이미지 파일 (규칙 제13의3.1(b) 및 시행세칙 713).

2. ☐ 추가로 서열목록에 대하여 하나 이상의 버전이나 사본이 제출된 경우, 후속 버전 또는 추가된 사본에 기재되어 있는 정보가 출원시 출원의 일부를 구성하는 정보와 동일하거나 또는 출원시의 개시범위를 벗어나지 않는다는 진술서가 제출되었습니다.

3. 추가 의견:

국 제 조 사 보 고 서

국제출원번호
PCT/KR2021/012872

제2기재란 일부 청구항을 조사할 수 없는 경우의 의견(첫 번째 용지의 2의 계속)

PCT 제17조(2)(a)의 규정에 따라 다음과 같은 이유로 일부 청구항에 대하여 본 국제조사보고서가 작성되지 아니하였습니다.

1. ☑ 청구항: **45-48,52-54**
이 청구항은 본 기관이 조사할 필요가 없는 대상에 관련됩니다. 즉,

청구항 45-48, 52-54는 수술 또는 치료에 의한 사람의 치료방법에 관한 것이므로 PCT 조약 제 17조(2)(a)(i) 및 조약규칙 39.1(iv)의 규정에 의하여 국제조사기관이 국제 조사할 의무가 없는 대상에 해당됩니다.

2. ☑ 청구항: **23,25,30,32,36,43-44**
이 청구항은 유효한 국제조사를 수행할 수 없을 정도로 소정의 요건을 충족하지 아니하는 국제출원의 부분과 관련됩니다. 구체적으로는,

청구항 23, 25, 30, 32, 36, 43-44는 PCT 규칙 제 6.4조(a)의 규정을 충족시키지 않는 다수 종속청구항을 인용하고 있습니다.

3. ☑ 청구항: **22,24,26-29,31,33-35,37-42,45-48**
이 청구항은 종속청구항이나 PCT규칙 6.4(a)의 두 번째 및 세 번째 문장의 규정에 따라 작성되어 있지 않습니다.

국 제 조 사 보 고 서
대응특허에 관한 정보

국제출원번호
PCT/KR2021/012872

국제조사보고서에서 인용된 특허문헌	공개일	대응특허문헌	공개일
KR 10-2019-0034628 A	2019/04/02	AU 2017-306676 A1	2019/02/28
		CA 3032699 A1	2018/02/08
		CN 110214183 A	2019/09/06
		EP 3494215 A1	2019/06/12
		GB 2568182 A	2019/05/08
		IL 264565 A	2019/02/28
		JP 2019-536425 A	2019/12/19
		SG 11201900907 A	2019/02/27
		US 10113163 B2	2018/10/30
		US 10947530 B2	2021/03/16
		US 2018-0073012 A1	2018/03/15
		US 2019-0093099 A1	2019/03/28
		US 2021-0317440 A1	2021/10/14
		WO 2018-027078 A1	2018/02/08
KR 10-2019-0044157 A	2019/04/30	없음	
KR 10-2018-0029937 A	2018/03/21	EP 3530737 A1	2019/08/28
		JP 2019-526271 A	2019/09/19
		KR 10-2026421 B1	2019/09/27
		US 2020-0131536 A1	2020/04/30
		WO 2018-052247 A1	2018/03/22
KR 10-2019-0025565 A	2019/03/11	AU 2017-274145 A1	2018/11/22
		AU 2017-274145 B2	2020/07/23
		AU 2020-244497 A1	2020/10/29
		AU 2020-244497 B2	2020/11/12
		AU 2021-200636 A1	2021/03/04
		BR 112018074531 A2	2019/03/19
		BR 112018074531 B1	2021/01/19
		CA 3026321 A1	2017/12/07
		CN 109983124 A	2019/07/05
		DK 3272867 T3	2019/12/02
		DK 3604527 T3	2021/05/31
		EP 3272867 A1	2018/01/24
		EP 3272867 B1	2019/09/25
		EP 3604527 A1	2020/02/05
		EP 3604527 B1	2021/05/12
		EP 3907286 A1	2021/11/10
		ES 2760477 T3	2020/05/14
		GB 2552861 A	2018/02/14
		GB 2552861 B	2019/05/15
		GB 2578802 A	2020/05/27
		GB 2578802 B	2020/09/16
		GB 2582731 A	2020/09/30
		GB 2582731 B	2020/12/30
		IL 262854 A	2018/12/31
		IL 262854 B	2020/06/30
		IL 275244 A	2020/07/30
		JP 2019-517795 A	2019/06/27
		JP 2021-121193 A	2021/08/26
		JP 6878468 B2	2021/05/26
		LT 3604527 T	2021/06/25

서식 PCT/ISA/210 (대응특허 추가용지) (2019년 7월)

국제조사보고서
대응특허에 관한 정보

국제출원번호
PCT/KR2021/012872

국제조사보고서에서 인용된 특허문헌	공개일	대응특허문헌	공개일
		PL 3272867 T3	2020/01/31
		PL 3604527 T3	2021/09/06
		PT 3272867 T	2019/12/04
		PT 3604527 T	2021/06/02
		SG 11201810003 A	2018/12/28
		US 10266851 B2	2019/04/23
		US 2017-0349913 A1	2017/12/07
		US 2019-0169650 A1	2019/06/06
		US 2019-0169651 A1	2019/06/06
		WO 2017-209809 A1	2017/12/07

서식 PCT/ISA/210 (대응특허 추가용지) (2019년 7월)

フロントページの続き

(51)国際特許分類	FI	テーマコード（参考）
C12N 9/16 (2006.01)	C12N 9/16 Z	

(32)優先日 令和3年1月29日(2021.1.29)
(33)優先権主張国・地域又は機関
韓国(KR)
(31)優先権主張番号 10-2021-0016788
(32)優先日 令和3年2月5日(2021.2.5)
(33)優先権主張国・地域又は機関
韓国(KR)
(31)優先権主張番号 10-2021-0049348
(32)優先日 令和3年4月15日(2021.4.15)
(33)優先権主張国・地域又は機関
韓国(KR)
(31)優先権主張番号 10-2021-0050497
(32)優先日 令和3年4月19日(2021.4.19)
(33)優先権主張国・地域又は機関
韓国(KR)
(31)優先権主張番号 10-2021-0085473
(32)優先日 令和3年6月30日(2021.6.30)
(33)優先権主張国・地域又は機関
韓国(KR)
(31)優先権主張番号 10-2021-0085474
(32)優先日 令和3年6月30日(2021.6.30)
(33)優先権主張国・地域又は機関
韓国(KR)
(31)優先権主張番号 10-2021-0092056
(32)優先日 令和3年7月14日(2021.7.14)
(33)優先権主張国・地域又は機関
韓国(KR)
(31)優先権主張番号 10-2021-0114750
(32)優先日 令和3年8月30日(2021.8.30)
(33)優先権主張国・地域又は機関
韓国(KR)
(81)指定国・地域 AP(BW,GH,GM,KE,LR,LS,MW,MZ,NA,RW,SD,SL,ST,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,RU,TJ,TM),EP(AL,AT,BE,BG,CH,CY,CZ,DE,DK,EE,ES,FI,FR,GB,GR,HR,HU,IE,IS,IT,LT,LU,LV,MC,MK,MT,NL,NO,PL,PT,RO,RS,SE,SI,SK,SM,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,KM,ML,MR,NE,SN,TD,TG),AE,AG,AL,AM,AO,AT,AU,AZ,BA,BB,BG,BH,BN,BR,BW,BY,BZ,CA,CH,CL,CN,CO,CR,CU,CZ,DE,DJ,DK,DM,DO,DZ,EC,EE,EG,ES,FI,GB,GD,GE,GH,GM,GT,HN,HR,HU,ID,IL,IN,IR,IS,IT,JO,JP,KE,KG,KH,KN,KP,KR,KW,KZ,LA,LC,LK,LR,LS,LU,LY,MA,MD,ME,MG,MK,MN,MW,MX,MY,MZ,NA,NG,NI,NO,NZ,OM,PA,PE,PG,PH,PL,PT,QA,RO,RS,RU,RW,SA,SC,SD,SE,SG,SK,SL,ST,SV,SY,TH,TJ,TM,TN,TR,TT,TZ,UA,UG,US,UZ,VC,VN,WS,ZA,ZM,ZW

ーグ，オリンピックーロ，１３５，＃２４０－９０２
(72)発明者 チョ，ソンイク
大韓民国，０５３３４ ソウル，カンドンーグ，クチョンミョンーロ ２８ーギル，１５－４，＃３０２
(72)発明者 カン，ボムチャン
大韓民国，３４１２５ テジョン，ユソンーグ，エクスポーロ １２３ボンーギル，３４，＃１０２－６０４
(72)発明者 イ，ソンヒョン

大韓民国，２２６９５ インチョン，ソーグ，スンハクーロ ４９５ボンーギル，７，＃１０５－５０６
(72)発明者 イ，ヒョンジ
大韓民国，３４１４０ テジョン，ユソンーグ，ユソンーグ，オウンーロ，５７，＃１１８－３０５
(72)発明者 モク，ヨングン
大韓民国，０８８３３ ソウル，クァナクーグ，クァナクーロ，１３３，＃１０１－１０３
(72)発明者 イ，ジミン
大韓民国，０８６０７ ソウル，クムチョンーグ，ポッコッロ ６ーギル，３，＃１０３－１９０１
(72)発明者 チュン，ユージン
大韓民国，０５６４２ ソウル，ソンパーグ，カラクーロ ３３ーギル，１３，＃４０３

Ｆターム（参考） 4H045 AA10 AA20 AA30 BA10 BA41 CA11 DA89 EA65 FA74