

(19)日本国特許庁(JP)

(12)公表特許公報(A)

(11)公表番号
特表2025-517630
(P2025-517630A)

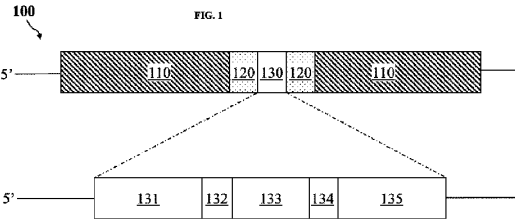
(43)公表日 令和7年6月10日(2025.6.10)

(51)国際特許分類		F I		テーマコード (参考)	
C 1 2 N	15/09 (2006.01)	C 1 2 N	15/09	1 0 0	4 C 0 7 6
C 1 2 N	15/11 (2006.01)	C 1 2 N	15/11	Z Z N A	4 C 0 8 4
A 6 1 P	43/00 (2006.01)	A 6 1 P	43/00	1 1 1	4 C 0 8 6
A 6 1 K	48/00 (2006.01)	A 6 1 K	48/00		
A 6 1 K	31/7105(2006.01)	A 6 1 K	31/7105		
		審査請求	未請求	予備審査請求	未請求 (全281頁) 最終頁に続く
(21)出願番号	特願2024-564803(P2024-564803)		(71)出願人	504256408	
(86)(22)出願日	令和5年5月2日(2023.5.2)			ザ リージェンツ オブ ザ ユニヴァーシ	
(85)翻訳文提出日	令和7年1月6日(2025.1.6)			ティ オブ カリフォルニア	
(86)国際出願番号	PCT/US2023/066470			THE REGENTS OF THE U	
(87)国際公開番号	WO2023/215727			NIVERSITY OF CALIFO	
(87)国際公開日	令和5年11月9日(2023.11.9)			R N I A	
(31)優先権主張番号	63/337,564			アメリカ合衆国 9 4 6 0 7 カリフォル	
(32)優先日	令和4年5月2日(2022.5.2)			ニア州, オークランド, フランクリン	
(33)優先権主張国・地域又は機関	米国(US)			ストリート 1 1 1 1 1 2 階	
(81)指定国・地域	AP(BW,CV,GH,GM,KE,LR,LS,MW,MZ			1 1 1 1 Franklin Stree	
	,NA,RW,SD,SL,ST,SZ,TZ,UG,ZM,ZW),			t, 12th Floor, Oakla	
	EA(AM,AZ,BY,KG,KZ,RU,TJ,TM),EP(			nd, California 9 4 6 0 7	
	AL,AT,BE,BG,CH,CY,CZ,DE,DK,EE,ES,		(74)代理人	100188499	
	FI,FR,GB,GR,HR,HU,IE,IS,IT,LT,LU,LV			弁理士 勝又 政徳	
	最終頁に続く			最終頁に続く	

(54)【発明の名称】 部位特異的ゲノム改変用の多要素システム

(57)【要約】

本発明は、逆転写酵素に関連するプロセスによって、モジュール型の遺伝子編集を行うためのシステム、組成物および方法を含む。より具体的には、本発明は、修飾ヌクレオチドおよび改変ペプチドを使用したシステムおよび方法を提供する。



【特許請求の範囲】

【請求項 1】

ゲノム編集用システムであって、

- (i) 少なくとも 1 つの逆転写酵素コンストラクト (RTC) と、
- (ii) 少なくとも 1 つの遺伝子挿入コンストラクト (GIC)

を含み、

前記 RTC が、少なくとも 1 つの逆転写酵素モジュール (RTC: RT モジュール) を含み、該 RTC: RT モジュールが、逆転写酵素 (RT)、少なくとも 1 つの 5' 側モジュール (RTC: 5' 側モジュール)、および／または少なくとも 1 つの 3' 側モジュール (RTC: 3' 側モジュール) をコードする mRNA を含み、

前記 GIC が、前記少なくとも 1 つの RTC によってコードされるポリペプチドによる逆転写に適した少なくとも 1 つの RNA 鋳型を含み、少なくとも 1 つの GIC: 5' 側モジュールと、少なくとも 1 つの GIC: ペイロードモジュールと、少なくとも 1 つの GIC: 3' 側モジュールを含んでいてもよい、システム。

【請求項 2】

i) 前記 RTC の 5' 側モジュールが、5' 非翻訳領域 (5' UTR)、コザック配列もしくは内部リボソーム侵入部位、非天然翻訳開始コドン、および／または 5' キャップを含み；

ii) 前記 RT モジュールが、ノドジロシトド (*Zonotrichia albicollis*) (ZoAl)、キンカチョウ (*Taeniopygia guttata*) (TaGu)、ノドジロシギダチョウ (*Tinamus guttatus*) (TiGu)、メダカ (*Oryzias latipes*) (OrLa) およびコクヌストモドキ (*Tribolium castaneum*) (B 系統) (TriCasB) からなる群から選択される生物に由来する逆転写酵素をコードする mRNA を含み；

iii) 前記 RTC の 3' 側モジュールが、逆転写酵素の翻訳終止コドン、3' 非翻訳領域 (3' UTR) およびポリ A テールを含み；

iv) 前記 GIC: 5' 側モジュールが、天然レトロエレメントの 5' 領域に由来する配列、rRNA 配列、リボザイム配列、フォールディングモチーフ配列、および／または RNA ポリメラーゼのターミネーター配列を含み；

(v) 前記 GIC: ペイロードモジュールが、少なくとも 1 つの導入遺伝子 ORF または非コード RNA (ncRNA) 配列、前記導入遺伝子のプロモーター配列、前記導入遺伝子の 5' 非翻訳配列、前記導入遺伝子の 3' 非翻訳配列、前記導入遺伝子のポリアデニル化シグナル配列、および／または前記導入遺伝子の ncRNA プロセッシング配列を含み；

(iv) 前記 GIC: 3' 側モジュールが、逆転写酵素認識配列、rRNA 配列、および／または A トラクト配列を含む、

請求項 1 に記載のシステム。

【請求項 3】

前記少なくとも 1 つの逆転写酵素が、非長鎖末端反復 (非 LTR) 型レトロエレメントまたはその改変バリエーションに由来するものである、請求項 1 または 2 に記載のシステム。

【請求項 4】

前記少なくとも 1 つの逆転写酵素が、少なくとも 1 つの DNA 結合ドメイン、少なくとも 1 つの RNA 結合ドメイン、少なくとも 1 つの cDNA 合成ドメイン、少なくとも 1 つのエンドヌクレアーゼドメイン、およびこれらの任意の組み合わせを含む、請求項 1～3 のいずれか 1 項に記載のシステム。

【請求項 5】

前記逆転写酵素が鳥類に由来するものである、請求項 1～4 のいずれか 1 項に記載のシステム。

【請求項 6】

前記逆転写酵素が、ノドジロシトド (ZoAl)、キンカチョウ (TaGu) またはノドジロシギダチョウ (TiGu) に由来するものである、請求項 5 に記載のシステム。

【請求項 7】

前記逆転写酵素が、配列番号18、配列番号20、配列番号27、配列番号29または配列番号25と少なくとも90%の同一性を有するアミノ酸配列を含む、請求項6に記載のシステム。

【請求項8】

前記少なくとも1つの逆転写酵素モジュールが、図2～5に記載の少なくとも1つの構造またはこれらの任意の組み合わせを含むか、これらをコードする、請求項1～7のいずれか1項に記載のシステム。

【請求項9】

前記少なくとも1つの逆転写酵素コンストラクトが、配列番号1～57およびこれらの任意の組み合わせからなる群から選択される少なくとも1つの配列を含むか、この配列をコードするか、この配列によってコードされる、請求項1～8のいずれか1項に記載のシステム。

10

【請求項10】

前記GIC：5'側モジュールの任意の構成要素である少なくとも1つのrRNA配列が、対象の1～30ntのrRNAを含むか、これをコードする、請求項2～9のいずれか1項に記載のシステム。

【請求項11】

前記rRNA配列が、配列番号250～276からなる群から選択される配列を含むか、または配列番号250～276からなる群から選択される配列と比較して1つ、2つもしくは3つのヌクレオチド変化を有する配列を含む、請求項10に記載のシステム。

20

【請求項12】

前記GIC：5'側モジュールがrRNA配列を含んでいない、請求項11に記載のシステム。

【請求項13】

前記GIC：5'側モジュールのリボザイム配列が、少なくとも1つの自己切断型リボザイムを含み、該自己切断型リボザイムが、デルタ肝炎ウイルス（HDV）のリボザイム折り畳み構造を含んでいてもよい、請求項2～12のいずれか1項に記載のシステム。

【請求項14】

前記HDV由来のリボザイムが、配列番号102～127および129～154からなる群から選択される配列と少なくとも90%の同一性を有する配列を含む、請求項13に記載のシステム。

30

【請求項15】

前記GIC：5'側モジュールのリボザイム配列が、少なくとも1つの非長鎖末端反復型レトロエレメントの5'領域に由来するリボザイムを含む、請求項2～12のいずれか1項に記載のシステム。

【請求項16】

前記リボザイムが、配列番号64～65、67、75～76、86、89～101および128からなる群から選択される配列と少なくとも90%の同一性を有する配列を含む、請求項15に記載のシステム。

【請求項17】

前記GIC：5'側モジュールのフォールディングモチーフ配列が、少なくとも1つの自律型フォールディングRNA配列モチーフを含み、該自律型フォールディングRNA配列モチーフが、少なくとも1つのヘアピンモチーフ、少なくとも1つのステムループモチーフ、少なくとも1つの対合したステム4モチーフまたはこれらの任意の組み合わせを含んでいてもよい、請求項2～16のいずれか1項に記載のシステム。

40

【請求項18】

前記フォールディングモチーフ配列が、配列番号278もしくは279を含むか、または配列番号278もしくは279と少なくとも90%の同一性を有する配列を含む、請求項17に記載のシステム。

【請求項19】

50

前記GIC：5'側モジュールが、配列番号60～154からなる群から選択される配列と少なくとも90%の同一性を有する配列を含む、請求項2～18のいずれか1項に記載のシステム。

【請求項20】

前記GIC：3'側モジュールの逆転写酵素認識配列が、少なくとも1つの逆転写酵素と相互作用する少なくとも1つの配列を含む、請求項2～19のいずれか1項に記載のシステム。

【請求項21】

前記GIC：3'側モジュールの逆転写酵素認識配列が、天然レトロエレメントの3'領域に由来するものである、請求項20に記載のシステム。

【請求項22】

前記GIC：3'側モジュールの少なくとも1つの逆転写酵素認識配列が、配列番号200～224からなる群から選択される配列と少なくとも90%の同一性を有する配列を含む、請求項20または21に記載のシステム。

【請求項23】

前記GIC：3'側モジュールのrRNA配列が、1～30ntのrRNAを含む、請求項2～22のいずれか1項に記載のシステム。

【請求項24】

前記rRNA配列が、配列番号280～289、および配列番号280～289に1つまたは2つのヌクレオチド置換を含む配列からなる群から選択される、請求項23に記載のシステム。

【請求項25】

前記GIC：3'側モジュールのAトラクト配列が、1～50個のアデニン塩基を含む、請求項2～24のいずれか1項に記載のシステム。

【請求項26】

前記GIC：3'側モジュールが、配列番号300～329およびこれらの任意の組み合わせからなる群から選択される配列と少なくとも90%の同一性を有する配列を含む、請求項2～25のいずれか1項に記載のシステム。

【請求項27】

前記GIC：3'側モジュールが、ZoA1、TaGu、GeFoまたはTiGuに由来する3'UTR配列を含む、請求項2～25のいずれか1項に記載のシステム。

【請求項28】

前記3'UTR配列が、配列番号202～205および配列番号222～224からなる群から選択される配列と少なくとも90%の同一性を有する配列を含む、請求項27に記載のシステム。

【請求項29】

前記少なくとも1つの導入遺伝子配列が、対象のゲノムへの挿入用の少なくとも1つの目的の配列を含むか、この目的の配列をコードする、請求項2～28のいずれか1項に記載のシステム。

【請求項30】

前記導入遺伝子配列が、少なくとも1つの、mRNA、マイクロRNA、siRNA、rRNA、tRNA、長鎖非コードRNA、細胞質低分子RNA、核内低分子RNA、核小体低分子RNA、カハル体低分子RNA、環状RNA、制御RNA、ペプチド、ポリペプチド、タンパク質、阻害タンパク質、および／または少なくとも1つの導入遺伝子の発現を制御する配列を含むか、これらをコードする、請求項29に記載のシステム。

【請求項31】

前記導入遺伝子が、hTERT、hPAH、ヒト第VIII因子、Bドメインの長さが様々な変異型ヒト第VIII因子、および第IX因子から選択されるタンパク質をコードする、請求項30に記載のシステム。

【請求項32】

10

20

30

40

50

前記導入遺伝子のプロモーター配列が、対象のゲノムにおいて導入遺伝子の発現を促進する少なくとも1つの配列を含む、請求項2～31のいずれか1項に記載のシステム。

【請求項33】

前記導入遺伝子の5'非翻訳配列が、前記導入遺伝子のmRNAの少なくとも1つの5'非翻訳領域を含む、請求項2～32のいずれか1項に記載のシステム。

【請求項34】

前記導入遺伝子の3'非翻訳配列が、前記導入遺伝子のmRNAの少なくとも1つの3'非翻訳領域を含む、請求項2～33のいずれか1項に記載のシステム。

【請求項35】

前記導入遺伝子のポリアデニル化シグナル配列が、少なくとも1つの導入遺伝子ポリアデニル化シグナルを含む、請求項2～34のいずれか1項に記載のシステム。 10

【請求項36】

前記導入遺伝子の非コードRNA(ncRNA)プロセッシング配列が、導入遺伝子から発現された少なくとも1つのncRNAの少なくとも1つの終止シグナル、少なくとも1つの3'プロセッシングシグナル、およびこれらの任意の組み合わせを含む、請求項2～35のいずれか1項に記載のシステム。

【請求項37】

前記少なくとも1つのGIC:ペイロードモジュールが、配列番号411～422、配列番号499～536およびこれらの任意の組み合わせからなる群から選択される配列と少なくとも90%の同一性を有する少なくとも1つの配列を含むか、この配列をコードする、請求項2～36のいずれか1項に記載のシステム。 20

【請求項38】

前記少なくとも1つのGIC:5'側モジュールおよび前記少なくとも1つのGIC:3'側モジュールの少なくとも一方が、他方が由来する非長鎖末端反復型レトロエレメントの由来生物種とは異なる生物種に由来する少なくとも1つの配列を含むか、この配列をコードする、請求項1～37のいずれか1項に記載のシステム。

【請求項39】

前記少なくとも1つの遺伝子挿入コンストラクトが、図6～9に示す少なくとも1つの構造およびこれらの任意の組み合わせを含むか、これらの構造をコードする、請求項1～38のいずれか1項に記載のシステム。 30

【請求項40】

2つの異なる遺伝子挿入コンストラクトを含み、これらの遺伝子挿入コンストラクトに含まれるGIC:ペイロードモジュールの導入遺伝子のORFが互いに異なるものである、請求項1～39のいずれか1項に記載のシステム。

【請求項41】

前記2つの異なる遺伝子挿入コンストラクトが同じRNA鋳型上に存在する、請求項40に記載のシステム。

【請求項42】

前記2つの異なる遺伝子挿入コンストラクトが異なるRNA鋳型上に存在する、請求項40に記載のシステム。 40

【請求項43】

(i) 少なくとも1つの逆転写酵素コンストラクトと、

(ii) 少なくとも1つの遺伝子挿入コンストラクト

を含み、

前記逆転写酵素コンストラクトが、配列番号1～57からなる群から選択される配列と少なくとも90%の同一性を有する少なくとも1つの配列を含むか、この配列によってコードされ；

前記少なくとも1つの遺伝子挿入コンストラクトが、

配列番号60～154からなる群から選択される配列と少なくとも90%の同一性を有する配列を含むGIC:5'側モジュールと、 50

配列番号250～276からなる群から選択される配列、または配列番号250～276からなる群から選択される配列と比較して1つ、2つもしくは3つのヌクレオチド変化を有する配列を含む任意の構成要素であるrRNA配列と、

少なくとも1つの導入遺伝子配列を含むGIC：ペイロードモジュールと、

配列番号300～329からなる群から選択される配列と少なくとも90%の同一性を有する配列を含むGIC：3'側モジュールと、

配列番号200～224からなる群から選択される配列と少なくとも90%の同一性を有する配列を含む、GIC：3'側モジュールの逆転写酵素認識配列と、

配列番号280～289、および配列番号280～289に1つもしくは2つのヌクレオチド置換を含む配列からなる群から選択される、GIC：3'側モジュールのrRNA配列と、

1～100個のアデニン塩基を含む、GIC：3'側モジュールのAトラクト配列を含む、

請求項1～42のいずれか1項に記載のシステム。

【請求項44】

前記GIC：ペイロードモジュールが、配列番号411～422および499～536からなる群から選択される配列と少なくとも90%の同一性を有する少なくとも1つの配列を含む、請求項43に記載のシステム。

【請求項45】

前記少なくとも1つの逆転写酵素コンストラクトおよび前記少なくとも1つの遺伝子挿入コンストラクトの少なくとも一方が、他方が由来するレトロエレメントの由来生物種とは異なる生物種に由来する少なくとも1つの配列を含むか、この配列をコードする、請求項1～44のいずれか1項に記載のシステム。

【請求項46】

前記逆転写酵素コンストラクトのRNAおよび／または前記遺伝子挿入コンストラクトのRNAが、少なくとも1つの修飾ウラシルを含むか、前記逆転写酵素コンストラクトのRNAおよび／または前記遺伝子挿入コンストラクトのRNAを構成するウラシルの100%が修飾ウラシルである、請求項1～45のいずれか1項に記載のシステム。

【請求項47】

前記修飾ウラシルが、5-メチルウリジン、5-メトキシウリジン、シュードウリジン、N¹-メチルシュードウリジンおよび／または2-チオウリジンからなる群から選択される、請求項46に記載のシステム。

【請求項48】

対象のゲノムに少なくとも1つの導入遺伝子を挿入する方法であって、請求項1～47のいずれか1項に記載の遺伝子挿入システム（GIS）の少なくとも1つを有効量で対象に投与する工程を含む、方法。

【請求項49】

前記導入遺伝子が、前記対象のゲノムの1つ以上の標的部位に挿入され、該1つ以上の標的部位が、少なくとも1つのセーフハーバー部位を含んでいてもよい、請求項48に記載の方法。

【請求項50】

任意の構成要素である前記少なくとも1つのセーフハーバー部位が、少なくとも1つのリボソームDNA（rDNA）配列を含み、該少なくとも1つのリボソームDNA配列が、少なくとも1つの28S rDNA配列を含んでいてもよい、請求項49に記載の方法。

【請求項51】

少なくとも1つの送達剤で製剤化された前記遺伝子挿入システムの少なくとも1つを投与する工程を含む、請求項48～50のいずれか1項に記載の方法。

【請求項52】

前記少なくとも1つの送達剤が、少なくとも1つのナノ粒子であり、該少なくとも1つのナノ粒子が、少なくとも1つの脂質ナノ粒子を含んでいてもよい、請求項51に記載の方法。

10

20

30

40

50

【請求項 5 3】

前記導入遺伝子が、90%を超える標的部位特異性で挿入される、請求項 4 8～5 2 のいずれか 1 項に記載の方法。

【請求項 5 4】

前記逆転写酵素コンストラクトのRNAが、ノドジロシトド (ZoA1)、キンカチョウ (TaGu) またはノドジロシギダチョウ (TiGu) に由来する逆転写酵素をコードするか、配列番号18、配列番号20、配列番号27、配列番号29または配列番号25と少なくとも90%の同一性を有するアミノ酸配列を含む、請求項 5 3 に記載の方法。

【請求項 5 5】

前記導入遺伝子が、3ヶ月以上にわたって前記標的部位において発現される、請求項 4 8～5 4 のいずれか 1 項に記載の方法。

【請求項 5 6】

医薬組成物であって、請求項 1～4 7 に記載の遺伝子挿入システムの少なくとも1つを含み、少なくとも1つの添加剤、少なくとも1つの送達剤、少なくとも1つの補助剤、およびこれらの任意の組み合わせのうちの少なくとも1つを含んでいてもよい、医薬組成物。

【請求項 5 7】

治療適応症の治療を必要とする対象において、治療適応症を治療する方法であって、請求項 1～4 7 に記載の遺伝子挿入システムの少なくとも1つ、または請求項 5 6 に記載の医薬組成物を有効量で投与する工程を含み、請求項 4 8～5 2 に記載の方法の少なくとも1つを含んでいてもよい、方法。

【請求項 5 8】

前記治療適応症が、テロメラーゼ活性の欠失によって引き起こされたものである、請求項 5 7 に記載の方法。

【請求項 5 9】

前記少なくとも1つの遺伝子挿入システムが、少なくとも1つのTERT導入遺伝子を含む、請求項 5 7 または 5 8 に記載の方法。

【請求項 6 0】

遺伝子挿入システムの作製用キットであって、請求項 1～4 7 に記載の遺伝子挿入システムを含み、請求項 5 6 に記載の医薬組成物を含んでいてもよく、バッファー、DNAプラスミド、または前記遺伝子挿入システムもしくは前記医薬組成物を作製するためのプロトコルをさらに含んでいてもよい、キット。

【請求項 6 1】

第2鎖にニックを導入して第2鎖を合成する宿主機構を動員する5'側モジュールのデノボ設計を含む方法。

【請求項 6 2】

(a) 所定の長さのrRNAを所定の位置に含み、(b) リボザイム (RZ) のフォールディングが促進され、かつ／または (c) 宿主細胞機構を動員するように構成された前記5'側モジュールのデノボ設計によって、挿入効率が増加する、請求項 6 1 に記載の方法。

【請求項 6 3】

細胞のゲノムに少なくとも1つの導入遺伝子を挿入する方法であって、請求項 1～4 7 のいずれか 1 項に記載の遺伝子挿入システム (GIS) の少なくとも1つを細胞に接触させる工程を含む、方法。

【請求項 6 4】

前記導入遺伝子が、対象のゲノムの1つ以上の標的部位に挿入され、該1つ以上の標的部位が、少なくとも1つのセーフハーバー部位を含んでいてもよい、請求項 6 3 に記載の方法。

【請求項 6 5】

任意の構成要素である前記少なくとも1つのセーフハーバー部位が、少なくとも1つのリボソームDNA (rDNA) 配列を含み、該少なくとも1つのリボソームDNA配列が、少

10

20

30

40

50

なくとも1つの28S rDNA配列を含んでいてもよい、請求項64に記載の方法。

【請求項66】

少なくとも1つの送達剤で製剤化された前記遺伝子挿入システムの少なくとも1つを投与する工程を含む、請求項63～65のいずれか1項に記載の方法。

【請求項67】

前記少なくとも1つの送達剤が、少なくとも1つのナノ粒子であり、前記少なくとも1つのナノ粒子が、少なくとも1つの脂質ナノ粒子を含んでいてもよい、請求項66に記載の方法。

【請求項68】

前記導入遺伝子が、90%を超える標的部位特異性で挿入される、請求項63～67のいずれか1項に記載の方法。

【請求項69】

前記逆転写酵素コンストラクトのRNAが、ノドジロシトド (ZoA1)、キンカチョウ (TaGu) またはノドジロシギダチョウ (TiGU) に由来する逆転写酵素をコードするか、配列番号18、配列番号20、配列番号27、配列番号29または配列番号25と少なくとも90%の同一性を有するアミノ酸配列を含む、請求項68に記載の方法。

【請求項70】

前記導入遺伝子が、3ヶ月以上にわたって前記標的部位において発現される、請求項63～69のいずれか1項に記載の方法。

【請求項71】

前記逆転写酵素コンストラクトと前記遺伝子挿入コンストラクトのモル比が約10:1～1:20である、請求項63～70のいずれか1項に記載の方法。

【請求項72】

インビトロでの方法、エキスピボでの方法またはインビボでの方法である、請求項63～71のいずれか1項に記載の方法。

【請求項73】

前記細胞が、初代細胞、形質転換細胞、上皮細胞、線維芽細胞、ヒト細胞、サル細胞およびマウス細胞からなる群から選択される、請求項63～72のいずれか1項に記載の方法。

【請求項74】

前記細胞が、同種細胞または自家細胞である、請求項63～73のいずれか1項に記載の方法。

【請求項75】

前記自家細胞が、HLAがマッチした細胞である、請求項74に記載の方法。

【発明の詳細な説明】

【技術分野】

【0001】

関連出願の相互参照

本出願は、2022年5月2日に出願された米国仮特許出願第63/337,564号の優先権を主張するものであり、この出願の開示内容は、引用によりその全体があらゆる目的で本明細書に援用される。

【0002】

配列表の参照

本願は電子形式の配列表とともに出願されたものである。この配列表は、[[XXXX.xm l]]というファイル名で[[XX]]に作成されたものであり、[[XX]]バイトのサイズである。この電子形式の配列表に記載の情報は、引用によりその全体が本明細書に援用される。

【0003】

政府支援に関する陳述

本発明は、米国国立衛生研究所によって付与された助成金番号GM139306およびHL156819による政府支援を受けてなされたものである。米国政府は、本発明に関し一定の

10

20

30

40

50

権利を保有する。

【背景技術】

【0004】

真核生物のゲノムへの導入遺伝子の導入は、遺伝子発現の改善、修正および／または改変を可能にすると同時に、疾患の症状の治療や緩和にも役立つ。実際に導入遺伝子を挿入することができれば、機能喪失変異の回復、機能獲得変異の阻害、RNA発現および／またはタンパク質発現の外因性制御、アイソフォーム発現の特異性の導入、組換え遺伝子および組換えタンパク質の発現、ならびにその他の有用な結果を得ることができる。

【0005】

しかし、遺伝物質を細胞に導入してゲノムに挿入する現在の方法には、克服すべき大きなハードルがいまだ存在する。例えば、標的細胞にDNAを送達する方法では、DNAが細胞質を通過する必要がある、これに付随して、破壊的な免疫応答や有害な免疫応答が誘導されることも多い。また、相同組換え（HR）によりDNAがゲノムに導入される部位特異的組み込み方法では、変異原性の二本鎖DNA切断が導入される可能性があり、組み込み部位の対象のゲノムやエピゲノムが破壊される恐れもある。高等真核生物では、特に有糸分裂後の細胞において、DNAの組み込みは非部位特異的であることが多い。これは、細胞周期の大半において相同組換えが抑制され、非相同末端結合が優先されるためである。

【0006】

導入するDNAの長さに柔軟に対応でき、細胞質にDNAを導入することなく、生細胞のゲノムに効果的かつ部位特異的に導入遺伝子を挿入できる方法であれば、ヒト生物学、動物生物学、微生物生物学および植物生物学への大きな貢献となり、研究応用や臨床応用の力強い進展につながると考えられる。

【0007】

このような方法の1つとして、導入遺伝子配列をRNAとして導入し、逆転写酵素（RT）による相補的DNA（cDNA）合成の鋳型とする方法が考えられる。しかし現時点では、哺乳動物細胞に導入されたRNAが、ゲノムに導入遺伝子を挿入するための鋳型として複製されるように誘導できる分子シグナルは同定されていない。

【0008】

このような課題に対処する上で、非長鎖末端反復（非LTR）型レトロエレメント（RE）またはこれと同義の非LTR型レトロトランスポゾンとして知られている一群の遺伝子は有望な解決策となりうる。これらの遺伝子は、宿主ゲノム内で自己増幅が可能であり、非LTR型レトロトランスポゾン逆転写酵素（RT）タンパク質を発現することでその機能を発揮する。このRTタンパク質は、自身のレトロエレメント転写産物RNAに結合して、これを鋳型として使用するとともに、（このRTタンパク質のエンドヌクレアーゼ（EN）ドメインの触媒作用により）ゲノムDNAに導入されたニックをcDNA合成開始のためのプライマーとして使用することで、cDNAを合成する（RTプライマー伸長）。このプロセスは、target-primed reverse transcription（TPRT）として知られており、これによりゲノム内に二本鎖DNAレトロエレメントのコピーが付加される。

【0009】

WO2022/155055では、ヒトゲノムのセーフハーバー部位への部位特異的な導入遺伝子挿入用の2種の構成要素からなるシステムについて述べられている。この2種の構成要素として、非LTR型レトロエレメント逆転写酵素（RT）と鋳型RNAが使用されている。この非LTR型レトロエレメント逆転写酵素（RT）は、5'側に挿入切断する傾向のある天然レトロエレメントの代わりに、導入遺伝子の全長の挿入が可能になるように組み換えられており、この非LTR型レトロエレメント逆転写酵素（RT）に鋳型RNAを適合させている。挿入される第1のDNA鎖の合成機構は、鋳型RNAの3'側モジュールにより誘導されるtarget-primed reverse transcription（TPRT）であり、この逆転写は、3'側モジュールの一部である非天然3'テールによって増強される。鋳型RNAの5'側モジュールは、鋳型RNAに生体安定性を付与し、RTタンパク質に結合する際の鋳型RNAのバイオアベイラビリティを増加させ、第2鎖の合成を誘導する。

【0010】

本開示は、レトロエレメント配列の一部から生体高分子コンストラクトを作製することによって、真核生物細胞のゲノムに、特にヒト細胞のゲノムに導入遺伝子を挿入し、発現させるための組成物および方法を提供する。

【発明の概要】

【課題を説明するための手段】

【0011】

本発明は、非長鎖末端反復（非LTR）型レトロトランスポゾンに由来する構成要素を使用して、target primed reverse transcription (TPRT) により対象のゲノムに導入遺伝子を挿入するための、組成物、方法、ならびに／またはタンパク質およびヌクレオチドの使用、ならびに改変タンパク質および改変ポリヌクレオチドを提供する。

10

【0012】

本発明は、ゲノム編集用システムであって、

(i) ポリヌクレオチド鋳型を逆転写する酵素活性を有するポリペプチドをコードするポリヌクレオチドを含む少なくとも1つの逆転写酵素コンストラクト (RTC) と、

(ii) 前記少なくとも1つのRTCによってコードされるポリペプチドによる逆転写に適した少なくとも1つのポリヌクレオチド鋳型を含む少なくとも1つの遺伝子挿入コンストラクト (GIC)

を含む、システムを提供する。

20

【0013】

いくつかの実施形態において、前記ゲノム編集用システムは、

(i) 少なくとも1つの逆転写酵素コンストラクト (RTC) と、

(ii) 少なくとも1つの遺伝子挿入コンストラクト (GIC)

を含み、

前記RTCは、少なくとも1つの逆転写酵素モジュール (RTC: RTモジュール) を含み、該RTC: RTモジュールは、逆転写酵素 (RT)、少なくとも1つの5'側モジュール (RTC: 5'側モジュール)、および／または少なくとも1つの3'側モジュール (RTC: 3'側モジュール) をコードするmRNAを含み、

前記GICは、前記少なくとも1つのRTCによってコードされるポリペプチドによる逆転写に適した少なくとも1つのRNA鋳型を含み、該GICは、少なくとも1つのGIC: 5'側モジュール、少なくとも1つのGIC: ペイロードモジュール、および／または少なくとも1つのGIC: 3'側モジュールを含む。

30

【0014】

いくつかの実施形態において、前記RTモジュールは、鳥類、節足動物、魚類、尾索動物およびその他の動物（哺乳動物およびヒトを含む）から選択される生物に由来する逆転写酵素をコードするmRNAを含む。

【0015】

いくつかの実施形態において、前記ゲノム編集用システムは、

i) 5'非翻訳領域 (5'UTR)、コザック配列、非天然翻訳開始コドン、および／または5'キャップを含むRTC 5'側モジュール；

40

ii) ノドジロシトド (*Zonotrichia albicollis*) (ZoAl)、キンカチョウ (*Taeniopygia guttata*) (TaGu)、ノドジロシギダチョウ (*Tinamus guttatus*) (TiGu)、メダカ (*Oryzias latipes*) (OrLa) およびコクヌストモドキ (*Tribolium castaneum*) (B系統) (TriCasB) からなる群から選択される生物に由来する逆転写酵素をコードするmRNAを含むRTモジュール；

iii) 逆転写酵素の翻訳終止コドン、3'非翻訳領域 (3'UTR) およびポリAテールを含むRTC 3'側モジュール；

iv) 天然レトロエレメントの5'領域に由来する配列、rRNA配列、リボザイム配列、フォールディングモチーフ配列、および／またはRNAポリメラーゼのターミネーター配列を含むGIC: 5'側モジュール；

50

(v) 少なくとも1つの導入遺伝子ORFまたは非コードRNA (ncRNA) 配列、前記導入遺伝子のプロモーター配列、内部リボソーム侵入部位 (IRES)、前記導入遺伝子の5' 非翻訳配列、前記導入遺伝子の3' 非翻訳配列、前記導入遺伝子のポリアデニル化シグナル配列、および／または前記導入遺伝子のncRNAプロセッシング配列を含むGIC：ペイロードモジュール；ならびに

(iv) 逆転写酵素認識配列、rRNA配列、および／またはAトラクト配列を含むGIC：3' 側モジュールを含む。

【0016】

いくつかの実施形態において、前記少なくとも1つの逆転写酵素コンストラクトは、少なくとも1つの生体高分子を含み、該生体高分子は、少なくとも1つの核酸、少なくとも1つのアミノ酸、およびこれらの任意の組み合わせを含む。いくつかの実施形態において、前記(i)に記載の逆転写酵素コンストラクトのポリヌクレオチドは、逆転写酵素をコードするmRNAを含む。いくつかの実施形態において、前記(ii)に記載の遺伝子挿入コンストラクトのポリヌクレオチドからなる鑄型は、RNAを含む。いくつかの実施形態において、前記(i)に記載の逆転写酵素コンストラクトのポリヌクレオチドは、逆転写酵素をコードするmRNAを含み、前記(ii)に記載の遺伝子挿入コンストラクトのポリヌクレオチドからなる鑄型は、別の(異なる) RNAを含む。いくつかの実施形態において、前記遺伝子挿入コンストラクトは、前記(i)に記載の逆転写酵素をコードするmRNAとは異なるRNA鑄型を含む。

10

20

【0017】

いくつかの実施形態において、前記少なくとも1つの逆転写酵素コンストラクトは、少なくとも1つの逆転写酵素オープンリーディングフレーム (ORF) モジュール (RTC：RTモジュール) を含み、少なくとも1つの5' 非翻訳領域 (UTR) モジュール (RTC：5' 側モジュール)、少なくとも1つの3' UTRモジュール (RTC：3' 側モジュール)、およびこれらの任意の組み合わせを含んでもよい。

【0018】

いくつかの実施形態において、前記少なくとも1つの逆転写酵素モジュールは、少なくとも1つの逆転写酵素を含むか、少なくとも1つの逆転写酵素をコードする。

【0019】

いくつかの実施形態において、前記少なくとも1つの逆転写酵素モジュールは、非長鎖末端反復 (非LTR) 型レトロエレメントに由来する少なくとも1つの逆転写酵素を含むか、これをコードする。

30

【0020】

いくつかの実施形態において、前記少なくとも1つの逆転写酵素モジュールは、非天然翻訳開始コドンを含むか、これをコードする。

【0021】

いくつかの実施形態において、前記少なくとも1つの逆転写酵素は、少なくとも1つのDNA結合ドメイン、少なくとも1つのRNA結合ドメイン、少なくとも1つのcDNA合成ドメイン、少なくとも1つのエンドヌクレアーゼドメイン、およびこれらの任意の組み合わせを含む。

40

【0022】

いくつかの実施形態において、前記少なくとも1つの逆転写酵素ドメイン、前記少なくとも1つの対象DNA結合ドメイン、前記少なくとも1つの鑄型RNA結合ドメイン、および前記少なくとも1つのエンドヌクレアーゼドメイン、ならびにこれらの任意の組み合わせのうちの少なくとも1つは、残りのドメインの少なくとも1つが由来する逆転写酵素の由来生物とは異なる生物種に由来するものである。

【0023】

いくつかの実施形態において、前記逆転写酵素コンストラクトの少なくとも1つの5' 側モジュールは、少なくとも1つのRNAポリメラーゼプロモーター、少なくとも1つの5

50

、非翻訳領域（5'-UTR）、少なくとも1つのコザック配列、少なくとも1つの5' キャップ、およびこれらの任意の組み合わせを含むか、これらをコードする。

【0024】

いくつかの実施形態において、前記逆転写酵素コンストラクトの少なくとも1つの3' 側モジュールは、少なくとも1つの逆転写酵素翻訳終止コドン、少なくとも1つの3' 非翻訳領域（3' UTR）、少なくとも1つのポリAトラクトおよび／またはポリAテール、およびこれらの任意の組み合わせを含むか、これらをコードする。

【0025】

いくつかの実施形態において、前記少なくとも1つの逆転写酵素モジュールは、図2～5に記載の少なくとも1つの構造またはこれらの任意の組み合わせを含むか、これらをコードする。

【0026】

いくつかの実施形態において、前記少なくとも1つの逆転写酵素コンストラクトは、配列番号1～57の少なくとも1つを含むか、配列番号1～57の少なくとも1つをコードするか、配列番号1～57の少なくとも1つによってコードされる。いくつかの実施形態において、前記少なくとも1つの逆転写酵素コンストラクトは、TriCasB、NaViB、OrLa、ZoAl、TiGu、TaGu、GeFo、DroSi、BoMo、DrMerc、DrMe、GaAc、PuPu、AdVa、HyMaA、CiIn、LiPo、TriCan、LeCo、およびこれらの任意の組み合わせからなる群から選択される生物種に由来する逆転写酵素タンパク質をコードするmRNAを含む。

【0027】

いくつかの実施形態において、前記少なくとも1つの遺伝子挿入コンストラクトは、少なくとも1つの核酸生体高分子を含むか、これをコードする。いくつかの実施形態において、前記遺伝子挿入コンストラクトは鋳型RNAを含む。

【0028】

いくつかの実施形態において、前記少なくとも1つの遺伝子挿入コンストラクトは、任意の構成要素である少なくとも1つのGIC：5' 側モジュール、少なくとも1つのGIC：ペイロードモジュール、任意の構成要素である少なくとも1つのGIC：3' 側モジュール、およびこれらの任意の組み合わせを含むか、これらをコードする。

【0029】

いくつかの実施形態において、前記少なくとも1つのGIC：5' 側モジュールは、天然レトロエレメントの5' 領域に由来する少なくとも1つの配列を含むか、これをコードし、少なくとも1つのrRNA配列、少なくとも1つのリボザイム（RZ）配列、少なくとも1つのフォールディングモチーフ配列、またはこれらの任意の組み合わせを含んでもよく、これらをコードしてもよい。

【0030】

いくつかの実施形態において、前記GIC：5' 側モジュールの任意の構成要素である少なくとも1つのrRNA配列は、対象の1～30ntのrRNAを含むか、これをコードする。

【0031】

いくつかの実施形態において、前記GIC：5' 側モジュールの任意の構成要素である少なくとも1つのリボザイム配列は、少なくとも1つの自己切断型リボザイムを含むか、これをコードし、該自己切断型リボザイムは、デルタ肝炎ウイルス（HDV）のリボザイムを含んでもよい。

【0032】

いくつかの実施形態において、前記GIC：5' 側モジュールの任意の構成要素である少なくとも1つのリボザイム配列は、少なくとも1つの非長鎖末端反復型レトロエレメントの5' 領域に由来するリボザイムを含むか、これをコードする。いくつかの実施形態において、前記GIC：5' 側モジュールの任意の構成要素である少なくとも1つのフォールディングモチーフ配列は、少なくとも1つの自律型フォールディングRNA配列モチーフを含むか、これをコードし、該自律型フォールディングRNA配列モチーフは、RZ内において、

10

20

30

40

50

少なくとも1つのヘアピンモチーフ、少なくとも1つのステムループモチーフ、少なくとも1つの対合したステムモチーフ、またはこれらの任意の組み合わせを含んでいてもよい。

【0033】

いくつかの実施形態において、前記GIC：5'側モジュールは、配列番号60～154の少なくとも1つ、または配列番号60～154の少なくとも1つと少なくとも90%の同一性（例えば、少なくとも90%、91%、92%、93%、94%、95%、96%、97%、98%、99%または100%の同一性）を有する配列を含むか、この配列をコードする。いくつかの実施形態において、前記GIC：5'側モジュールは、OrLa、TriCasB、TriCasA、ZoAl、TiGu、DroSi、LeCo、CiIn、FoRa、TriCan、HDV-28、HDV-24、HDV-21、HDV-13、HDV-36、およびこれらの任意の組み合わせからなる群から選択される生物種に由来する配列を含む。

10

【0034】

いくつかの実施形態において、前記少なくとも1つのGIC：3'側モジュールは、少なくとも1つの逆転写酵素認識配列を含むか、これをコードし、少なくとも1つのrRNA配列、少なくとも1つのAトラクト配列、またはこれらの任意の組み合わせを含んでいてもよく、これらをコードしてもよい。

【0035】

いくつかの実施形態において、前記GIC：3'側モジュールの少なくとも1つの逆転写酵素認識配列は、少なくとも1つの逆転写酵素と相互作用する少なくとも1つの配列を含むか、これをコードする。いくつかの実施形態において、前記GIC：3'側モジュールの少なくとも1つの逆転写酵素認識配列は、配列番号200～224からなる群から選択される配列を含む。

20

【0036】

いくつかの実施形態において、前記GIC：3'側モジュールの少なくとも1つの逆転写酵素認識配列は、天然レトロエレメントの3'領域に由来するものである。

【0037】

いくつかの実施形態において、前記GIC：3'側モジュールの任意の構成要素である少なくとも1つのrRNA配列は、1～30ntのrRNAを含むか、これをコードする。

【0038】

いくつかの実施形態において、前記GIC：3'側モジュールの任意の構成要素である少なくとも1つのAトラクト配列は、約1～50個のアデニン塩基からなる配列を含むか、これをコードする。

30

【0039】

いくつかの実施形態において、前記少なくとも1つのGIC：3'側モジュールは、配列番号200～224の少なくとも1つまたは配列番号300～329の少なくとも1つを含むか、これらのいずれか1つをコードする。いくつかの実施形態において、前記GIC：3'側モジュールは、OrLa、TriCasB、TaGu、GeFo、ZoAl、NaViB、DroSi、PuPu、LiPo、BoMo、GaAc、LeCo、CiIn、DrMe、DrNa、DrMer、TriCan、AdVa、HyMaA、およびこれらの任意の組み合わせからなる群から選択される生物種に由来する配列を含む。

40

【0040】

いくつかの実施形態において、前記少なくとも1つのGIC：ペイロードモジュールは、少なくとも1つの導入遺伝子ORF配列を含むか、これをコードし、前記導入遺伝子の少なくとも1つのプロモーター配列、前記導入遺伝子の少なくとも1つの5'非翻訳配列、前記導入遺伝子の少なくとも1つの3'非翻訳配列、前記導入遺伝子の少なくとも1つのポリアデニル化シグナル配列、前記導入遺伝子の少なくとも1つの非コードRNA（ncRNA）プロセッシング配列、少なくとも1つのncRNAプロセッシング配列、および／もしくはその他の3'末端プロセッシング配列もしくは安定化シグナル、またはこれらの任意の組み合わせを含んでいてもよく、これらをコードしてもよい。

50

【0041】

いくつかの実施形態において、前記少なくとも1つの導入遺伝子配列は、対象のゲノムへの挿入用の少なくとも1つの目的の配列を含むか、これをコードする。

【0042】

いくつかの実施形態において、前記導入遺伝子の少なくとも1つのプロモーター配列は、対象のゲノムにおいて導入遺伝子の発現を促進する少なくとも1つの配列を含むか、これをコードする。

【0043】

いくつかの実施形態において、前記少なくとも1つのGIC：ペイロードモジュールは、前記導入遺伝子の少なくとも1つの5'非翻訳配列を含み、この少なくとも1つの5'非翻訳配列は、前記導入遺伝子のmRNAの少なくとも1つの5'非翻訳領域を含むか、これをコードする。

10

【0044】

いくつかの実施形態において、前記導入遺伝子の少なくとも1つの3'非翻訳配列は、前記導入遺伝子のmRNAの少なくとも1つの3'非翻訳領域を含むか、これをコードする。

【0045】

いくつかの実施形態において、前記導入遺伝子の少なくとも1つのポリアデニル化シグナル配列は、前記導入遺伝子の少なくとも1つのポリアデニル化シグナルを含むか、これをコードする。

【0046】

いくつかの実施形態において、前記導入遺伝子の少なくとも1つの非コードRNA(ncRNA)プロセッシング配列および／またはその他の3'末端プロセッシング配列もしくは安定化シグナルは、導入遺伝子から発現された少なくとも1つのncRNAの少なくとも1つの終止シグナル、少なくとも1つの3'プロセッシングシグナル、およびこれらの任意の組み合わせを含むか、これらをコードする。

20

【0047】

いくつかの実施形態において、前記少なくとも1つのGIC：ペイロードモジュールは、配列番号411～422および配列番号499～536ならびにこれらの任意の組み合わせのうちの少なくとも1つと少なくとも90%の同一性(例えば、少なくとも90%、91%、92%、93%、94%、95%、96%、97%、98%、99%または100%の同一性)を有する配列を含むか、この配列をコードする。

30

【0048】

いくつかの実施形態において、前記少なくとも1つのGIC：5'側モジュールおよび前記少なくとも1つのGIC：3'側モジュールの少なくとも一方は、他方が由来する非長鎖末端反復型レトロエレメントの由来生物種とは異なる生物種に由来する少なくとも1つの配列を含むか、この配列をコードする。

【0049】

いくつかの実施形態において、前記少なくとも1つの遺伝子挿入コンストラクトは、本明細書に添付の図面に示された少なくとも1つの構造、例えば、図6～9およびこれらの任意の組み合わせに記載の少なくとも1つの構造を含むか、これらの少なくとも1つの構造をコードする。

40

【0050】

いくつかの実施形態において、前記ゲノム編集用システムは、

(i) 少なくとも1つの逆転写酵素コンストラクトと、

(ii) 少なくとも1つの遺伝子挿入コンストラクト

を含み、

前記少なくとも1つの逆転写酵素コンストラクトは、配列番号1～57からなる群から選択される配列と少なくとも90%の同一性(例えば、少なくとも90%、91%、92%、93%、94%、95%、96%、97%、98%、99%または100%の同一性)を有する少なくとも1つの配列を含むか、この配列をコードするか、この配列によってコードされ、

50

前記少なくとも1つの遺伝子挿入コンストラクトが、配列番号60～154、250～276、278～279、280～289、300～329、400～404、405～407、411～422および499～536からなる群から選択される配列と少なくとも90%の同一性（例えば、少なくとも90%、91%、92%、93%、94%、95%、96%、97%、98%、99%または100%の同一性）を有する少なくとも1つの配列を含む。

いくつかの実施形態において、逆転写酵素タンパク質を生成するためにトランスフェクトされるmRNA配列は、プラスミドから分割して発現され、複数のタンパク質のアミノ酸配列をコードする。

【0051】

いくつかの実施形態において、前記ゲノム編集用システムは、

(i) 少なくとも1つの逆転写酵素コンストラクトと、

(ii) 少なくとも1つの遺伝子挿入コンストラクト

を含み、

前記少なくとも1つの逆転写酵素コンストラクトは、配列番号1～57からなる群から選択される配列と少なくとも90%の同一性（例えば、少なくとも90%、91%、92%、93%、94%、95%、96%、97%、98%、99%または100%の同一性）を有する少なくとも1つの配列を含むか、この配列によってコードされ、

前記少なくとも1つの遺伝子挿入コンストラクトは、

配列番号60～154からなる群から選択される配列と少なくとも90%の同一性（例えば、少なくとも90%、91%、92%、93%、94%、95%、96%、97%、98%、99%または100%の同一性）を有する配列を含むGIC：5'側モジュールと、

配列番号250～276からなる群から選択される配列、または配列番号250～276からなる群から選択される配列と比較して1つ、2つもしくは3つのヌクレオチド変化を有する配列を含む任意の構成要素であるrRNA配列と、

少なくとも1つの導入遺伝子配列を含むGIC：ペイロードモジュールと、

配列番号300～329からなる群から選択される配列と少なくとも90%の同一性（例えば、少なくとも90%、91%、92%、93%、94%、95%、96%、97%、98%、99%または100%の同一性）を有する配列を含むGIC：3'側モジュールと、

配列番号200～224からなる群から選択される配列と少なくとも90%の同一性（例えば、少なくとも90%、91%、92%、93%、94%、95%、96%、97%、98%、99%または100%の同一性）を有する配列を含む、GIC：3'側モジュールの逆転写酵素認識配列と、

配列番号280～289、および配列番号280～289に1つ、2つまたは3つのヌクレオチド置換を含む配列からなる群から選択される、GIC：3'側モジュールのrRNA配列と、

1～100個のアデニン塩基を含む、GIC：3'側モジュールのAトラクト配列を含む。

【0052】

いくつかの実施形態において、前記RTC 5'側モジュールの5'UTRは、配列番号58と少なくとも90%の同一性（例えば、少なくとも90%、91%、92%、93%、94%、95%、96%、97%、98%、99%または100%の同一性）を有する配列を含む。

【0053】

いくつかの実施形態において、前記RTC 3'側モジュールの3'UTRは、配列番号59と少なくとも90%の同一性（例えば、少なくとも90%、91%、92%、93%、94%、95%、96%、97%、98%、99%または100%の同一性）を有する配列を含む。

【0054】

いくつかの実施形態において、前記ゲノム編集用システムは、本明細書に記載の遺伝子挿入コンストラクトの少なくとも1つを含むか、これをコードする、遺伝子挿入コンストラクトの合成用コンストラクト（GIC：合成用コンストラクト）を含む。

【0055】

10

20

30

40

50

いくつかの実施形態において、前記少なくとも1つの逆転写酵素コンストラクトおよび前記少なくとも1つの遺伝子挿入コンストラクトの少なくとも一方は、他方が由来するレトロエレメントの由来生物種とは異なる生物種に由来する少なくとも1つの配列を含むか、この配列をコードする。

【0056】

いくつかの実施形態において、前記ゲノム編集用システムは、(i) 本明細書に記載の少なくとも1つの逆転写酵素コンストラクトと、(ii) 本明細書に記載の少なくとも1つの遺伝子挿入コンストラクトの組み合わせのうちの少なくとも1つを含む。

【0057】

さらに、対象のゲノムに少なくとも1つの導入遺伝子を挿入する方法であって、本開示の遺伝子挿入システム(GIS)の少なくとも1つを有効量で対象に投与する工程を含む、方法を提供する。

【0058】

いくつかの実施形態において、前記導入遺伝子は、前記対象のゲノムの1つ以上の標的部位に挿入され、該1つ以上の標的部位は、少なくとも1つのセーフハーバー部位を含んでいてもよい。

【0059】

いくつかの実施形態において、任意の構成要素である前記少なくとも1つのセーフハーバー部位は、少なくとも1つのリボソームDNA(rDNA)配列を含み、該少なくとも1つのリボソームDNA配列は、少なくとも1つの28S rDNA配列を含んでいてもよい。

【0060】

いくつかの実施形態において、前記少なくとも1つの方法は、少なくとも1つの送達剤で製剤化された前記遺伝子挿入システムの少なくとも1つを投与する工程を含む。

【0061】

いくつかの実施形態において、前記少なくとも1つの送達剤は、少なくとも1つのナノ粒子であり、該少なくとも1つのナノ粒子は、少なくとも1つの脂質ナノ粒子を含んでいてもよい。

【0062】

さらに、医薬組成物であって、請求項に記載の遺伝子挿入システムの少なくとも1つを含み、少なくとも1つの添加剤、少なくとも1つの送達剤、少なくとも1つの補助剤、およびこれらの任意の組み合わせのうちの少なくとも1つを含んでいてもよい、医薬組成物を提供する。

【0063】

さらに、治療適応症の治療を必要とする対象において、治療適応症を治療する方法であって、本開示の遺伝子挿入システムの少なくとも1つ、または本開示の医薬組成物を有効量で前記対象に投与する工程を含む、方法を提供する。

【0064】

いくつかの実施形態において、前記治療適応症は、テロメラーゼ活性の欠失によって引き起こされたものである。

【0065】

いくつかの実施形態において、前記少なくとも1つの遺伝子挿入システムは、少なくとも1つのTERT導入遺伝子を含む。

【0066】

さらに、本開示の遺伝子挿入システムの作製用キットを提供する。いくつかの実施形態において、前記キットは、本開示の医薬組成物を含む。いくつかの実施形態において、前記キットは、バッファー、DNAプラスミド、または前記遺伝子挿入システムもしくは医薬組成物を作製するためのプロトコルをさらに含んでいてもよい。

【0067】

さらに、第2鎖にニックを導入して第2鎖を合成する宿主機構を動員する5'側モジュールのデノボ設計を含む方法を提供する。いくつかの実施形態において、前記方法におい

10

20

30

40

50

て、(a) (本明細書に記載の) 所定の長さのrRNAを所定の位置に含み、(b) リボザイム(RZ)のフォールディングが促進され、かつ／または(c) 宿主細胞機構を動員するように構成された前記5'側モジュールのデノボ設計によって、挿入効率が增加する。

【0068】

別の一態様において、本開示は、細胞のゲノムに少なくとも1つの導入遺伝子を挿入する方法であって、本開示の遺伝子挿入システム(GIS)の少なくとも1つを細胞に接触させる工程を含む、方法を提供する。

【0069】

いくつかの実施形態において、前記導入遺伝子は、前記対象のゲノムの1つ以上の標的部位に挿入され、該1つ以上の標的部位は、少なくとも1つのセーフハーバー部位を含んでいてもよい。いくつかの実施形態において、任意の構成要素である前記少なくとも1つのセーフハーバー部位は、少なくとも1つのリボソームDNA(rDNA)配列を含み、該少なくとも1つのリボソームDNA配列は、少なくとも1つの28S rDNA配列を含んでいてもよい。

【0070】

いくつかの実施形態において、前記方法は、少なくとも1つの送達剤で製剤化された前記遺伝子挿入システムの少なくとも1つを投与する工程を含む。いくつかの実施形態において、前記少なくとも1つの送達剤は、少なくとも1つのナノ粒子であり、該少なくとも1つのナノ粒子は、少なくとも1つの脂質ナノ粒子を含んでいてもよい。

【0071】

いくつかの実施形態において、前記導入遺伝子は、90%を超えるオンターゲットな標的部位特異性(例えば、90%、91%、92%、93%、94%、95%、96%、97%、98%または99%を超える標的部位特異性)で挿入される。

【0072】

いくつかの実施形態において、前記逆転写酵素コンストラクトは、ノドジロシトド(ZoA1)、キンカチョウ(TaGu)もしくはノドジロシギダチョウ(TiGu)に由来する逆転写酵素をコードするRNAを含むか、または配列番号18、配列番号20、配列番号27、配列番号29もしくは配列番号25と少なくとも90%の同一性を有するアミノ酸配列を含む。

【0073】

いくつかの実施形態において、前記導入遺伝子は、3ヶ月以上にわたって前記標的部位において発現される。

【0074】

いくつかの実施形態において、前記細胞と前記遺伝子挿入システムを接触させ、前記逆転写酵素コンストラクトと前記遺伝子挿入コンストラクトのモル比は、約10:1~1:20である。

【0075】

いくつかの実施形態において、前記方法は、インビトロでの方法、エクスピボでの方法またはインピボでの方法である。

【0076】

いくつかの実施形態において、前記細胞は、初代細胞、形質転換細胞、上皮細胞、線維芽細胞、ヒト細胞、サル細胞およびマウス細胞からなる群から選択される。

【0077】

いくつかの実施形態において、前記細胞は、同種細胞または自家細胞である。いくつかの実施形態において、前記自家細胞は、HLAがマッチした細胞である。

【0078】

本発明は、本明細書に記載の特定の実施形態のあらゆる組み合わせが詳細に記載されているものとして、これらの組み合わせを包含する。

【図面の簡単な説明】

【0079】

【図1】 標的挿入部位と天然レトロエレメントを含む対象のゲノムの一例を示すダイアグラムである。拡大図（下図）は、天然のR2レトロエレメントの例示的な構成要素の構造を示す。

【0080】

【図2】 逆転写酵素コンストラクト（RTC）の一例の構造を示すダイアグラムである。

【0081】

【図3】 本発明の逆転写酵素（RT）タンパク質の例示的な各ドメインを示すダイアグラムである。

【0082】

【図4】 DNA結合ドメイン（DB）、RNA結合ドメイン（RB）、逆転写酵素（RT）ドメインおよびエンドヌクレアーゼ（EN）ドメインを含む逆転写酵素（RT）タンパク質ドメインの供給源としての例示的な生物を示す図である。さらに、各RTタンパク質ドメインの組み合わせの例の小さなセットを示すダイアグラムを示す。各ドメインの識別は各生物ごとに定義され、野生型RTは以下の略号で示している。A1はノドジロシトド（*Zonotrichia albicollis*）であり、A2はキンカチョウ（*Taeniopygia guttata*）であり、A3はノドジロシギダチョウ（*Tinamus guttatus*）であり、A4はガラパゴスフィンチ（*Geospiza fortis*）であり、B1はキタノトミヨ（*Pungitis pungitis*）であり、B2はメダカ（*Oryzias latipes*）であり、B3はイトヨ（*Gasterosteus aculeatus*）であり、C1はキョウソヤドリコバチ（*Nasonia vitripennis*）であり、C2はキイロショウジョウバエ（*Drosophila melanogaster*）であり、C3はコクヌストモドキ（*Tribolium castaneum*）（B系統）であり、C4はカイコガ（*Bombyx mori*）であり、C5はオナジショウジョウバエ（*Drosophila simulans*）であり、C6は*Drosophila mercatorum*であり、D1は*Lepidurus couseii*であり、D2はヨーロッパカブトエビ（*Triops cancriformis*）であり、E1はチクビヒドラ（*Hydra magnipapillata*）であり、E2はアメリカカブトガニ（*Limulus polyphemus*）であり、E3は*Adineta vaga*であり、E4はカタユレイボヤ（*Ciona intestinalis*）である。

10

20

【0083】

【図5】 本発明の例示的な一連の逆転写酵素コンストラクト（RTC）を示す一連のダイアグラムである。この逆転写酵素コンストラクト（RTC）は、逆転写酵素の翻訳開始コドン（M）を含む逆転写酵素タンパク質（RT）を含む配列を含むか、これをコードする配列を含む。逆転写酵素コンストラクト（RTC）は、5'非翻訳配列（5'-UTR）、翻訳終止コドン（SC）および／または3'非翻訳配列（3'-UTR）を含んでいてもよい。

30

【0084】

【図6】 遺伝子挿入コンストラクト（中央）の一例の構造を示すダイアグラムである。拡大図は、5'側モジュール（左下）、3'側モジュール（右下）、およびペイロードモジュール（上部）の一例の構造を示す。

【0085】

【図7】 遺伝子挿入コンストラクト（GIC）の5'側モジュール（5'M）の構成要素および3'側モジュール（3'M）の構成要素と、逆転写酵素コンストラクト（RTC）のRTモジュール（RT）の構成要素の供給源となる例示的な生物を示す図である。さらに、ペイロードモジュールを挟むように配置された5'側モジュールと3'側モジュールの可能な組み合わせを有する遺伝子挿入コンストラクトの想定される例と、これらとペアになる逆転写酵素コンストラクト（ペアになるRT）とを含むいくつかのセットを示す図も示している。モジュールの識別は生物名により定義され、野生型のレトロエレメントおよび／または逆転写酵素は以下のように記載している。A1はノドジロシトド（*Zonotrichia albicollis*）であり、A2はキンカチョウ（*Taeniopygia guttata*）であり、A3はノドジロシギダチョウ（*Tinamus guttatus*）であり、A4はガラパゴスフィンチ（*Geospiza fortis*）であり、B1はキタノトミヨ（*Pungitis pungitis*）であり、B2はメダカ（*Oryzias latipes*）であり、B3はイトヨ（*Gasterosteus aculeatus*）であり、C1はキョウソヤドリコバチ（*Nasonia vitripennis*）であり、C2はキイロショウジョウバエ（*Dros*

40

50

ophila melanogaster) であり、C3はコクヌストモドキ (*Tribolium castaneum*) であり、C4はカイコガ (*Bombyx mori*) であり、C5はオナジショウジョウバエ (*Drosophila simulans*) であり、C6は *Drosophila mercatorum* であり、D1は *Lepidurus couseii* であり、D2はヨーロッパカブトエビ (*Triops cancriformis*) であり、E1はチクビヒドラ (*Hydra magnipapillata*) であり、E2はアメリカカブトガニ (*Limulus polyphemus*) であり、E3は *Adineta vaga* であり、E4はカタユウレイボヤ (*Ciona intestinalis*) である。

【0086】

【図8】本発明の遺伝子挿入システム (GIS) により導入遺伝子を挿入した後の対象のゲノムの一例の構造を示す図である。

【0087】

【図9】合成遺伝子挿入コンストラクトの一例の構造を示す図である。

【0088】

【図10】変性PAGEゲルで分離した放射性DNA合成産物の画像である。黒色の実線の囲み枠は、予想された産物長のゲル領域を示す。各レーンの番号は、実施例10の表3に詳述している試験した様々なRTタンパク質に対応している。レーン1の反応は、RTタンパク質を発現しない細胞から精製した陰性コントロールを含む。

【0089】

【図11】図11Aは、同種R2因子3' UTRと同種でないR2因子3' UTRに由来する鋳型RNAに対するRTタンパク質の特異性を検証するための実験計画の一例を示した図である。

図11Bは、同種3' UTRと同種でない3' UTRに対する、カイコガ、オナジショウジョウバエおよびメダカのRTの選択性を調べたスポットプロットの結果である。

【0090】

【図12A-12B】TPRT反応産物の変性PAGEゲルの結果を示す。矢印は、正しいTPRT産物で予想されるサイズを示す。レーンBはカイコガのRTの反応産物を含み、レーンDはオナジショウジョウバエのRTの反応産物を含み、レーンOはメダカのRTの反応産物を含み、レーンNは酵素非存在下の反応産物を含む。図12Aは、TPRT反応の結果を示し、オナジショウジョウバエの3' UTRを含む鋳型と各種RTタンパク質の反応産物（「単独」と表示されたレーン）、およびオナジショウジョウバエの3' UTRと4ntのrRNAを含む鋳型と各種RTタンパク質の反応産物（「R4あり」と表示されたレーン）を示している。図12Bは、TPRT反応の結果を示し、メダカの3' UTRを含む鋳型と各種RTタンパク質の反応産物（「単独」と表示されたレーン）、およびメダカの3' UTRと4ntのrRNAを含む鋳型と各種RTタンパク質の反応産物（「R4あり」と表示されたレーン）を示している。

【0091】

【図13】カイコガ由来RTと各種鋳型のTPRT反応産物の変性PAGEゲルの結果を示す。矢印は、正しいTPRT産物の予想されるサイズを示す。丸印は、内部配列において合成が開始された場合の産物の長さを示す。

【0092】

【図14A-14B】メダカ由来RTと各種鋳型のTPRT反応産物の変性PAGEゲルの結果を示す。

【0093】

【図15】コクヌストモドキ由来RTと各種鋳型のTPRT反応産物の変性PAGEゲルの結果を示す。目的のTPRT産物の長さを矢印で示す。

【0094】

【図16】ノドジロシトド由来RTタンパク質からのTPRT反応産物の変性PAGEゲルの結果を示す。実施例17に示した表8は、変性PAGEゲルの写真に示した各レーンに使用した遺伝子挿入コンストラクトの識別名を示している。実線の囲み枠（上部）は、予想された長さのTPRT産物を示し、破線の囲み枠は、予想された長さの沈殿回収操作コントロールを示し（中間）、二点鎖線の囲み枠は、予想された長さの放射標識標的部位オリゴヌクレオチドを示す（下部）。

10

20

30

40

50

【0095】

【図17】キンカチョウ由来RTタンパク質からのTPRT反応産物の変性PAGEゲルの結果を示す。レーン1は、長さ参照用のラダーを含んでいた。レーン2は、RTタンパク質のみ（鋳型RNAなし）を含んでいた。実施例19に示した表11は、変性PAGEゲルの写真に示したその他のレーンに使用した遺伝子挿入コンストラクトの識別名を示している。実線の囲み枠（上部）は、予想された長さのTPRT産物を示し、破線の囲み枠は、予想された長さの沈殿回収操作コントロールを示し（中間）、二点鎖線の囲み枠は、予想された長さの放射標識標的部オリゴヌクレオチドを示す（下部）。

【0096】

【図18A-18B】コクヌストモドキ由来RTタンパク質と各種鋳型により鋳型化した導入遺伝子を挿入した後のゲノムDNAのPCR増幅産物を示す。図18Aでは、予想された産物長を囲み枠で示している。正確に挿入されたPCR産物はすべて同じサイズであるはずである。図18Bでは、予想された産物長を矢印で示している。5'側モジュールを持たない鋳型（3）と5'側モジュールを持つ鋳型（5_3）とで、正確に挿入されたPCR産物の長さが様々に異なっている。

【0097】

【図19】ゲノムDNAのPCR増幅の結果を示す。上パネルは予想された3'連結部の増幅を示し、下パネルは予想された5'連結部の増幅を示す。「L」で示したレーンは、長さ参照用のラダーを含んでいた。レーン1とレーン9は、TriCasB由来RTを発現するプラスミドや遺伝子挿入コンストラクトをトランスフェクトしていないPCR産物を含んでいた。レーン2～8は、RT発現プラスミドをトランスフェクトせずに、実施例21の表13に示した遺伝子挿入コンストラクトをトランスフェクトした後のPCR産物を含んでいた。レーン10～16は、実施例21の表13に示した遺伝子挿入コンストラクトと、RT発現プラスミドの両方をトランスフェクトした後のPCR産物を含んでいた。予想されたPCR産物の長さのいくつかにアスタリスクで印を付けている。アスタリスクを含むすべての画像については、補足資料の図面を参照されたい。

【0098】

【図20】ゲノムDNAのPCR増幅の結果を示す。レーンA～Jは、実施例24の表16に示した逆転写酵素コンストラクトのmRNAと遺伝子挿入システムのRNAを共トランスフェクトした後に目的の5'連結部を検出した際の予想されたサイズのPCR産物を示す。

【0099】

【図21】導入遺伝子由来GFP陰性クローン細胞集団（上部の2つのパネル）と、導入遺伝子由来GFP陽性クローン細胞集団（下部の2つのパネル）の例示的なFACS分析結果を示す。

【発明を実施するための形態】

【0100】

1. 序論

以下の詳細な説明および本明細書全体を通して、不適切である場合や別段の記載がない限り、「a」および「an」という用語は1つ以上のものを意味し、「または」という用語は「および／または」を意味する。本明細書に記載の実施例および実施形態は、説明のみを目的としたものである。このことから、様々な改良や変更を行えることが当業者に示唆されており、これらの改良や変更は、本願の要旨および範囲ならびに添付の請求項の範囲に含まれる。本明細書に引用された刊行物、特許および特許出願（これらの文献中の引用文献を含む）は、いずれも引用によりその全体があらゆる目的で本明細書に援用される。

【0101】

本発明は、対象のゲノムに導入遺伝子を挿入することを含む、ゲノム編集用および／または遺伝子改変用のシステムおよび方法を提供する。本明細書において、「遺伝子挿入システム（GIS）」と呼ばれる本発明のシステムは、（a）少なくとも1つの逆転写酵素を含むか、これをコードする少なくとも1つの逆転写酵素（RT）コンストラクト（RTC）と、（b）逆転写用の鋳型として使用されるRNAコンストラクトを含むか、これをコード

し、RTCとは別個に発現される少なくとも1つの遺伝子挿入コンストラクト（GIC）という、少なくとも2種の構成要素を含んでもよい（すなわち、2種の構成要素からなるGIS）。本明細書において、「コンストラクト」という用語は、人工的に設計または合成されたあらゆる生体高分子を意味する。この生体高分子は、例えば、核酸（例えば、DNAやRNA）、アミノ酸またはこれらの任意の組み合わせで構成されていてもよい。いくつかの実施形態において、（a）と（b）はいずれもRNAコンストラクトである。いくつかの実施形態において、（a）はアミノ酸コンストラクト（すなわちタンパク質）であり、（b）はRNAコンストラクトである。

【0102】

また、target primed reverse transcription（TPRT）を行うことが可能な組換え逆転写酵素コンストラクト（RTC）が提供される。本明細書において、「target primed reverse transcription（TPRT）」という用語は、逆転写酵素が標的部位の利用可能なDNAの3'末端をプライマーとして利用することでcDNAの合成を開始するプロセスを意味する。

【0103】

さらに、本明細書において提供されるシステムおよび方法は、対象のDNAの配列特異的な位置（本明細書において「標的部位」と呼ぶ）、例えば、セーフハーバー部位に導入遺伝子を挿入可能であってもよい。本明細書において、「セーフハーバー」および「セーフハーバー部位」という用語は、例えば、異種配列の挿入により対象のDNA配列が破壊されても対象の細胞の機能に悪影響が及ばない、対象のゲノムの部位を意味する。本発明において利用される例示的なセーフハーバー部位は、リボソームRNA（rRNA）をコードする対象のゲノムの一部に存在する。セーフハーバー部位として、本明細書において「リボソームDNA（rDNA）座位」と呼ばれる座位によってコードされるRNAポリメラーゼIにより転写されるrRNA前駆体が挙げられ、このrRNA前駆体は、5.8S rRNA、18S rRNAまたは28S rRNAをコードする配列を含む。

【0104】

本開示は、RNAのみを送達することによって、細胞（例えば、ヒト細胞）のゲノムのセーフハーバー部位にDNA導入遺伝子を挿入できるようにプログラムできることを実証するものである。いくつかの実施形態において、挿入される導入遺伝子をコードするRNA鋳型と、RNA鋳型をゲノムDNAに変換するのに必要な逆転写酵素をコードするメッセージRNAとが、細胞に送達される。本発明のRNAのみを送達する方法は、現在も技術革新が進行中の、無毒かつ高効率に特定の種類の細胞を標的とするRNA送達機構を利用することにより、ヒト遺伝子治療への移行がさらに容易となると予想される。

【0105】

いくつかの態様において、プラスミドからの逆転写酵素（RT）の発現と、RNA鋳型のトランスフェクションを組み合わせる。いくつかの実施形態において、天然のR2レトロエレメント配列またはその一部を含む、導入遺伝子の鋳型の5'側モジュールと逆転写酵素とを異種間で組み合わせる使用することによって、切断型レトロエレメント配列を挿入するのではなく、全長配列を部位特異的に挿入できるという利点がある。いくつかの実施形態において、鋳型RNAは、逆転写酵素と同じ生物種に由来するレトロエレメントの3' UTR配列を有する3'側モジュールを含む。いくつかの実施形態において、この3' UTRは、標的部位への特異的な挿入効率を増加させる3'ポリAトラクトをさらに含む。

【0106】

本開示は、先行技術によるシステムおよび方法と比べて、以下の改良点および利点を提供する。

【0107】

本発明者らは以下の事項を実証することができた。

（i）鳥類由来の逆転写酵素（RT）タンパク質は、導入遺伝子の挿入に対して顕著な活性を示し、20%を超えるトランスフェクト細胞において導入遺伝子が機能的に発現された。鳥類由来の逆転写酵素は、鳥類由来の3' UTRと3'ポリAトラクトを含む鋳型RNAの

複製に対して極めて高い選択性を示す。

【0108】

(ii) 鳥類由来のR2レトロエレメントの3' UTRと逆転写酵素タンパク質を異種間で組み合わせることによって、天然の組み合わせよりも効果的になりうる。

【0109】

(iii) デノボ作製され、最適化された非天然の5' 側モジュールは、さらに効果的であることから、部位特異的な挿入効率が1桁以上増加する。

【0110】

(iv) コクヌストモドキ (TriCasA) 由来の天然の5' 側モジュール (TCA、TCA5、TCARZなど) は、鳥類の逆転写酵素タンパク質とは完全に異なるクレードのR2レトロエレメントに由来するものであり、このようなコクヌストモドキ由来の天然の5' 側モジュールはさらに効果的になりうる。

【0111】

(v) 逆転写酵素をプラスミド発現させた後に、鋳型RNAをトランスフェクトするのではなく、2種のRNA系を共トランスフェクトして送達することにより導入遺伝子を挿入する。

【0112】

(vi) 2種のRNA系をトランスフェクトすることによって1個の細胞に複数の導入遺伝子を挿入できることから、1回のRNA投与で遺伝子送達の多重化が可能になる。このような方法によって、複数の治療用導入遺伝子を1個の細胞のゲノムに挿入することが可能となり、この複数の治療用導入遺伝子として、治療用タンパク質をそれぞれコードする複数の導入遺伝子、複数の治療用タンパク質の別々のサブユニットをそれぞれコードする複数の導入遺伝子、治療用タンパク質の組み合わせ、および治療用RNAの組み合わせが挙げられる。

【0113】

(vii) 2種のRNA系を送達することによって、広範囲な種類の細胞において導入遺伝子を発現させることができ、このような細胞として、初代細胞株、非分裂細胞、および分裂速度の遅い細胞が挙げられ、マウス細胞、サル細胞およびヒト細胞が挙げられる。

【0114】

(viii) ゲノムシーケンシングによって、導入遺伝子の挿入が部位特異性であることが示されている。

【0115】

(ix) 挿入された導入遺伝子発現カセットは、数ヶ月間にわたり安定に発現する。

【0116】

レトロエレメント由来の構成要素

【0117】

本発明の逆転写酵素コンストラクト (RTC) および／または遺伝子挿入コンストラクト (GIC) は、少なくとも1つの非長鎖末端反復 (非LTR) 型レトロエレメントの一部に由来してもよく、かつ／または自然界には存在しない構成要素 (同じ意味で「モジュール」とも呼ぶ) を含んでいてもよい。何らかの理論に拘束されることを望むものではないが、図1 (上図) は、天然レトロエレメント100を含む対象のゲノムを示し、この場合、非長鎖末端反復 (非LTR) 型レトロエレメントを示している。この図から見て取れるように、対象のDNA 110は、少なくとも1つの標的挿入部位120を含んでいてもよく、この標的挿入部位に天然レトロエレメント130が挿入されてもよい。拡大図 (下図) において、一例としての天然レトロエレメントの構造をさらに詳しく検討する。この図では、天然レトロエレメントの5' 領域131が、翻訳開始部位132の前に配置されている。天然レトロエレメントの5' 領域は、通常、アミノ酸生体高分子に翻訳されない。天然レトロエレメントの5' 領域は、後の挿入において、この天然レトロエレメント自身の逆転写酵素 (RT) によって認識されて、かつ／または天然レトロエレメントの第2鎖の合成に影響を及ぼす核酸配列を含んでいてもよい。翻訳開始部位132は、アミノ酸に翻訳される第

1のヌクレオチドである。天然レトロエレメントの逆転写酵素のオープンリーディングフレーム133は、逆転写酵素をコードし、この逆転写酵素は、この天然レトロエレメント自身のRNA転写産物を認識し、これに結合し、これを鋳型として使用することにより逆転写を行うことができる。天然レトロエレメントの逆転写酵素のオープンリーディングフレームは、翻訳終止部位134まで伸長しているが、このオープンリーディングフレームと翻訳終止部位は区別される。天然レトロエレメントの3'領域135は、通常、アミノ酸生体高分子に翻訳されず、天然のレトロエレメント自身の逆転写酵素により認識される核酸配列を含んでいてもよい。5'領域131と3'領域135は存在していてもよく、存在していなくてもよく、5'領域131と3'領域135が存在する場合、周囲の標的部位配列を複製し、かつ／または天然レトロエレメントのRNA鋳型によってコードされない配列を含んでいてもよい。

10

【0118】

遺伝子挿入システム（GIS）の各構成要素が由来してもよい好適なレトロエレメントとして、例えば、RLE型、APE型またはPenelope型の非LTR型レトロエレメントが挙げられるが、これらに限定されない。RLE型の非LTR型レトロトランスポゾン、以下に限定されないが、R2、R4、CRE、Genie、HERO、NeSLを含む多数のクレードのうちのいずれか1つに由来するものであってもよい。APE型の非LTR型レトロトランスポゾンは、以下に限定されないが、I、R1、L1、Tx1、CR1、Rex1、Jockey、L2、Tad、RTE、RTEX、Ingi、Vingi、TRAS、SARTまたはこれらの任意の組み合わせを含む多数のクレードのうちのいずれか1つに由来するものであってもよい。いくつかの実施形態において、遺伝子挿入システムの構成要素は、rDNAに挿入されるレトロエレメント、すなわち、いわゆるR因子に由来するものであってもよく、例えば、R1クレードまたはR2クレードのレトロエレメントに由来するものであってもよい。いくつかの実施形態において、R2クレードのレトロエレメントは、標準的なR2レトロエレメントの挿入部位に対して特異性を有していてもよく、あるいは、標準的なR2レトロエレメントの標的配列が変更されることにより上位系統のR2クレードから派生したR8レトロエレメントおよび／またはR9レトロエレメントに由来するものであってもよく、あるいは、標的部位に対する特異性を喪失したことにより派生したR2NSレトロエレメントに由来するものであってもよい。

20

【0119】

遺伝子挿入システム（GIS）の各構成要素は、対象とは遠縁の種を含む何らかの生物種から得られるレトロエレメントの一部またはそのドメインに由来するものであってもよい。例えば、遺伝子挿入システムの各構成要素が由来してもよい好適なレトロエレメントとして、鳥類（例えば、ノドジロシトド（*Zonotrichia albicollis*）、キンカチョウ（*Taeniopygia guttata*）、ノドジロシギダチョウ（*Tinamus guttatus*）およびガラパゴスフィンチ（*Geospiza fortis*））、魚類（例えば、キタノトミヨ（*Pungitius pungitius*）、メダカ（*Oryzias latipes*）、ゼブラフィッシュ（*Danio rerio*）、*Oryzias melastigma*、ウミヤツメ（*Petromyzon marinus*）、ブラウントラウト（*Salmo trutta*）、タイセイヨウサケ（*Salmo salar*）またはイトヨ（*Gasterosteus aculeatus*））、昆虫（例えば、*Drosophila mercatorum*、キイロショウジョウバエ（*Drosophila melanogaster*）、キョウソヤドリコバチ（*Nasonia vitripennis*）、コクヌストモドキ（*Tribolium castaneum*）、オナジショウジョウバエ（*Drosophila simulans*）、トウヨウミツバチ（*Apis cerana*）およびカイコガ（*Bombyx mori*））、甲殻類（例えば、*Lepidurus couesii*およびヨーロッパカブトエビ（*Triops cancriformis*））、その他の無脊椎動物（例えば、アメリカカブトガニ（*Limulus polyphemus*）、チクビヒドラ（*Hydra magnipapillata*）またはAdineta vaga）、脊索動物（例えば、カタユレイボヤ（*Ciona intestinalis*））、哺乳動物、ならびにこれらの任意の組み合わせに見られるレトロエレメントが挙げられる。

30

40

【0120】

いくつかの実施形態において、遺伝子挿入システムの各構成要素は、本明細書に開示さ

50

れたどのような配列の一部またはドメインに由来するものであってもよい。

【0121】

II. 遺伝子挿入システム (GIS) の組成

本開示の記載の全体を通して、対象のゲノムに遺伝物質（例えば導入遺伝子）を挿入するための本発明のシステムを「遺伝子挿入システム (GIS)」と呼ぶ。本開示の遺伝子挿入システムは、複数の生体高分子コンストラクトで構成されていてもよく、これらの生体高分子コンストラクトは、target primed reverse transcription (TPRT) を介して少なくとも1つの導入遺伝子を挿入するために共投与される。これらの生体高分子コンストラクトは、アミノ酸からなる生体高分子、核酸からなる生体高分子、アミノ酸と核酸の両方を含むハイブリッド型生体高分子、またはこれらの任意の組み合わせであってもよい。いくつかの例では、本開示の遺伝子挿入システムは、少なくとも2つの生体高分子からなり、すなわち、少なくとも1つの逆転写酵素コンストラクト (RTC) と、少なくとも1つの遺伝子挿入コンストラクト (GIC) からなる。そのような例において、逆転写酵素コンストラクト (RTC) は、例えば、逆転写酵素を含んでいたり、逆転写酵素をコードすることによって、逆転写を行うための手段を含み、遺伝子挿入コンストラクト (GIC) は、逆転写酵素コンストラクト (RTC) によりcDNAを合成するための鋳型として使用されてもよい。少なくとも1つのRNA配列を含むか、このようなRNA配列をコードする。

10

【0122】

本発明の生体高分子コンストラクトは、それら自体が複数のモジュールで構成されていることから、必要に応じてこれらのモジュールを組み合わせることによって、所望の機能を発揮できるように本開示の遺伝子挿入システムを改変してもよい。本明細書において、「モジュール」という用語は、その機能によって定義されるコンストラクトの一部（例えば、タンパク質の機能性ドメイン）、またはその配列によって定義されるコンストラクトの一部（例えば、アミノ酸配列または核酸配列）を意味する。

20

【0123】

逆転写酵素コンストラクト (RTC)

本発明の遺伝子挿入システムは、非LTR型レトロエレメントに由来する逆転写酵素などの活性な逆転写酵素タンパク質を含むか、これをコードする少なくとも1つの逆転写酵素コンストラクト (RTC) を含む。本明細書において、「RTC」という用語は、少なくとも1つの逆転写酵素 (RT) を含むか、少なくとも1つの逆転写酵素 (RT) をコードする生体高分子コンストラクトを意味する。いくつかの実施形態において、本発明の遺伝子挿入システムで使用される少なくとも1つの逆転写酵素コンストラクト (RTC) は、ポリペプチド、タンパク質、プロタンパク質またはこれらの任意の組み合わせなどのアミノ酸生体高分子を含んでいてもよいが、これらに限定されない。いくつかの実施形態において、本発明の遺伝子挿入システムで使用される少なくとも1つの逆転写酵素コンストラクト (RTC) は、RNA、DNAまたはこれらの任意の組み合わせなどの核酸生体高分子を含んでいてもよいが、これらに限定されない。いくつかの実施形態において、少なくとも1つの逆転写酵素コンストラクト (RTC) は、少なくとも1つのmRNAコンストラクトを含んでいてもよい。

30

【0124】

逆転写酵素コンストラクト (RTC) の構造

本発明の逆転写酵素コンストラクト (RTC) は、少なくとも1つのRTC：逆転写酵素モジュール (RTC：RTモジュール)、任意の構成要素である少なくとも1つの5'側モジュール (RTC：5'側モジュール)、任意の構成要素である少なくとも1つの3'側モジュール (RTC：3'側モジュール)、およびこれらの任意の組み合わせを含んでいてもよい。RTCのいくつかの例において、RTC：5'側モジュールおよびRTC：3'側モジュールは、任意の構成要素であってもよく、一方または両方が存在していなくてもよい。いくつかの実施形態において、少なくとも1つのRTCは、直鎖RNA生体高分子を含んでいてもよく、または直鎖RNA生体高分子として対象に送達してもよい。いくつかの実施形態において、少なくとも1つのRTCは、mRNA生体高分子を含んでいてもよく、またはmRNA

40

50

生体高分子として対象に送達してもよい。

【0125】

図2を参照すると、直鎖RNA生体高分子（例えば、mRNA）である逆転写酵素コンストラクト（RTC）200の例示的な構造が示されている。この図に示すように、mRNA生体高分子であるRTCにおいて、RTC：5'側モジュール210は、RTCの任意の構成要素であり、RTC内に存在する場合、RTCの免疫原性を改変する配列および／またはRTC：RTモジュール220の発現を制御する配列を含んでいてもよい。例えば、RTC：5'側モジュールは、少なくとも1つの5'キャップ（例えば、TriLink社製のClean Cap AGや、m7(3'OMeG)(5')ppp(5')(2'OMeA)pG)、少なくとも1つの5'非翻訳領域（5'-UTR）、少なくとも1つのコザック配列、少なくとも1つのプロモーター、およびこれらの任意の組み合わせを含んでいてもよく、これらをコードしてもよい。開始コドンは、翻訳を開始させることが知られている3塩基の核酸配列であり、RTC：RTモジュールの5'末端を規定する。RTC：RTモジュール（以降で詳述する）は、開始コドンから終止コドンまで伸長した領域を含むが、終止コドンは含まない。任意の構成要素であるRTC：3'側モジュール230は、RTC内に存在する場合、終止コドンからRTCの3'末端まで伸長した領域を含む。RTC：3'側モジュールは、RTC内に存在する場合、RTCの免疫原性を改変する配列および／またはRTC：RTモジュールの発現を制御する配列を含んでいてもよい。例えば、RTC：3'側モジュールは、翻訳終止コドン、3'UTR、ポリアデノシン配列、ポリアデニル化シグナル、またはこれらの任意の組み合わせを含んでいてもよく、これらをコードしてもよい。

10

20

【0126】

いくつかの実施形態において、少なくとも1つのRTCは、プラスミドを含んでいてもよく、またはプラスミドとして対象に送達してもよい。いくつかの実施形態において、少なくとも1つのRTCは、mRNAもしくはpro-mRNAを含んでいてもよく、またはmRNAもしくはpro-mRNAとして対象に送達してもよい。いくつかの実施形態において、少なくとも1つのRTCは、タンパク質を含んでいてもよく、またはタンパク質として対象に送達してもよい。いくつかの実施形態において、少なくとも1つのRTCは、プロタンパク質を含んでいてもよく、またはプロタンパク質として対象に送達してもよい。

【0127】

RTC：RTモジュール

RTCのRTモジュールは、逆転写活性を有する少なくとも1つの化合物または組成物を含むか、これらをコードし、そのような化合物または組成物の具体例として、逆転写酵素（RT）として知られている一群の酵素タンパク質が挙げられるが、これに限定されない。いくつかの実施形態において、RTモジュールは、レトロエレメント遺伝子に見られる少なくとも1つの逆転写酵素（すなわち、レトロエレメント逆転写酵素）に由来する生体高分子を含んでいてもよく、これをコードしてもよい。いくつかの実施形態において、RTC：RTモジュールは、非長鎖末端反復（非LTR）型レトロエレメントに由来する少なくとも1つの逆転写酵素を含むか、これをコードする。

30

【0128】

逆転写酵素

本明細書において、「逆転写酵素（RT）」という用語は、最も広い意味で使用され、逆転写活性を有するあらゆる生体高分子を意味する。いくつかの実施形態において、本明細書で使用されるRTは、ノドジロシトド（*Zonotrichia albicollis*）、キンカチョウ（*Taeniopygia guttata*）、ノドジロシギダチョウ（*Tinamus guttatus*）、ガラパゴスフィンチ（*Geospiza fortis*）、キタノトミヨ（*Pungitis pungitis*）、メダカ（*Oryzias latipes*）、ゼブラフィッシュ（*Danio rerio*）、*Oryzias melastigma*、ウミヤツメ（*Petromyzon marinus*）、ブラウントラウト（*Salmo trutta*）、タイセイヨウサケ（*Salmo salar*）、イトヨ（*Gasterosteus aculeatus*）、*Drosophila mercatorum*、キイロショウジョウバエ（*Drosophila melanogaster*）、キョウソヤドリコバチ（*Nasonia vitripennis*）、コクヌストモドキ（*Tribolium castaneum*）

40

50

、オナジショウジョウバエ (*Drosophila simulans*)、トウヨウミツバチ (*Apis cerana*)、カイコガ (*Bombyx mori*)、*Lepidurus couesii*、ヨーロッパカブトエビ (*Triops cancriformis*)、アメリカカブトガニ (*Limulus polyphemus*)、チクビヒドラ (*Hydra magnipapillata*)、*Adineta vaga*、カタユレイボヤ (*Ciona intestinalis*)、その他の鳥類、その他の節足動物、その他の魚類、その他の尾索動物、その他の動物 (哺乳動物およびヒトを含む) に由来する非LTR型RTもしくはそのゲノムであってもよく、またはこれらの非LTR型RTもしくはそのゲノムに由来するものであってもよい。

【0129】

いくつかの実施形態において、本開示の遺伝子挿入システム (GIS) で使用される少なくとも1つのRTC: RTモジュールは、配列番号1~57の少なくとも1つを含んでいてもよく、配列番号1~57の少なくとも1つをコードしてもよく、配列番号1~57の少なくとも1つによってコードされてもよい。いくつかの実施形態において、少なくとも1つのRTC: RTモジュールは、配列番号1~57の少なくとも1つと少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。いくつかの実施形態において、RTC: RTモジュールは、配列番号1~57と少なくとも90%、91%、92%、93%、94%、95%、96%、97%、98%または99%の同一性を有する非天然配列を含む。

【0130】

いくつかの実施形態において、少なくとも1つのRTC: RTモジュールは、配列番号17~21 (ZoA1 RT配列) の少なくとも1つと少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

【0131】

いくつかの実施形態において、少なくとも1つのRTC: RTモジュールは、配列番号26~29 (TaGu RT配列) の少なくとも1つと少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

【0132】

いくつかの実施形態において、少なくとも1つのRTC: RTモジュールは、配列番号1~5 (TriCasB RT配列) の少なくとも1つと少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

【0133】

いくつかの実施形態において、RTC: RTモジュールは、好適なTPRTアッセイによりTPRT活性があることが示されたタンパク質を含んでいてもよく、このようなタンパク質をコードしてもよい。一例としての好適なTPRTアッセイは、(i) アフィニティー精製の適切なタグ (例えば、FLAGタグ) を有するRTタンパク質をコードする発現プラスミドを細胞集団にトランスフェクトする工程、(ii) 前記細胞集団を溶解し、当技術分野で公知の適切な方法により、発現したタンパク質産物を回収し、精製する工程、(iii) 当技術分野で公知の任意の方法 (例えば、T7 RNAポリメラーゼ) により、組換え鋳型RNAを調製する工程、(iv) 精製RTタンパク質と、組換え鋳型と、末端が放射標識された下の鎖を有する標的部位オリゴヌクレオチド二重鎖DNAを含むヌクレオチド溶液とを、RTによる逆転写が促進される培地中で組み合わせる工程、ならびに(v) 当技術分野で公知の任意の適切な方法 (例えば、変性PAGE) により、産物を回収し、分析する工程を含

むが、これらに限定されない。

【0134】

本発明での使用に好適な逆転写酵素（RT）は、複数の機能性ドメインで構成されていてもよい。いくつかの実施形態において、例えば、図3に示すように、少なくとも1つの逆転写酵素300は、少なくとも1つのDNA結合ドメイン310、少なくとも1つのRNA結合ドメイン320、少なくとも1つのcDNA合成ドメイン330、少なくとも1つのエンドヌクレアーゼドメイン340、およびこれらの任意の組み合わせを含む。この図では、各ドメインに配置することが可能な1つの構成のみを提示していることには留意されたい。いくつかの実施形態において、図示した各ドメインは様々な頻度でこの逆転写酵素（RT）に含まれていてもよく、かつ／またはこれらのドメインはどのような順序で含まれていてもよい。いくつかの実施形態において、DNA結合ドメインまたはRNA結合ドメインは、逆転写酵素（RT）とは異なる種類のポリペプチドに由来するものであってもよく、真核生物のゲノムにおいて知られていない配列（例えば、デノボ作製されたDNA結合ドメインまたはRNA結合ドメイン）であってもよい。

10

【0135】

開始コドン

当技術分野で公知の様々な方法を利用して、逆転写酵素をコードする核酸配列に少なくとも1つの非天然翻訳開始コドンを付加してもよい。この非天然翻訳開始コドンは、非LTR型レトロエレメントに由来する配列において、機能性逆転写酵素の生成を誘導可能な、どのような位置に付加してもよい。例えば、野生型の非LTR型レトロエレメントにおいて、公知の基準点から（例えば、天然レトロエレメント逆転写酵素のORFのアミノ酸配列モチーフから）約1塩基、約10塩基、約50塩基、約100塩基、約150塩基、約200塩基、約250塩基、約300塩基、約350塩基、約400塩基、約450塩基、約500塩基、約550塩基、約600塩基、約650塩基、約700塩基、約750塩基、約800塩基、約850塩基、約900塩基、約950塩基、または約1000塩基以上離れた位置に少なくとも1つの非天然開始コドンを付加してもよい。翻訳開始コドンの位置は、組換え技術を用いた目的の標的細胞におけるタンパク質発現の最適化または調節に従事している当業者には明らかな様々な検討事項のうち、ポリペプチドの長さ、配列組成、活性、生物学的安定性、凝集の回避もしくは局在化を最適化した結果に応じて選択してもよく、かつ／またはタンパク質をコードするmRNAの生物学的安定性が改善されるように選択してもよい。

20

30

【0136】

翻訳開始コドンは、mRNA中の別の配列もしくは別の構造に応じて、またはmRNA中の別の配列もしくは別の構造とは無関係に、リボソームにより翻訳が開始されることが知られている3つのヌクレオチドであればどのようなものであってもよい。いくつかの実施形態において、非天然翻訳開始コドンはAUGである。

【0137】

RTC：5'側モジュール

本発明の逆転写酵素コンストラクト（RTC）は、少なくとも1つのRTC：5'側モジュールを含んでもよい。通常、RTC：5'側モジュールは、翻訳されない生体高分子構成要素を含み、この翻訳されない生体高分子構成要素によって、例えば、遺伝子挿入コンストラクトの免疫原性の改変、標的とする細胞内領域への遺伝子挿入コンストラクトの局在化の補助、遺伝子挿入コンストラクトに含まれるRTC：RTモジュールの発現の制御または改変、同定用の遺伝子挿入コンストラクトの標識、遺伝子挿入コンストラクトの精製の補助、遺伝子挿入コンストラクトの分解の制御、遺伝子挿入コンストラクトの活性および／または機能の外因性または内因性の調節、ならびにこれらの任意の組み合わせを行ってもよいが、これらに限定されない。

40

【0138】

いくつかの実施形態において、少なくとも1つのRTC：5'側モジュールは、少なくとも1つの5' UTRを含んでもよく、これをコードしてもよい。いくつかの実施形態において、少なくとも1つのRTC：5'側モジュールは、少なくとも1つの5' キャップを含

50

んでいてもよく、これをコードしてもよい。いくつかの実施形態において、少なくとも1つのRTC：5'側モジュールは、少なくとも1つのマイクロRNA結合配列を含んでいてもよく、これをコードしてもよい。いくつかの実施形態において、少なくとも1つのRTC：5'側モジュールは、少なくとも1つのRNAポリメラーゼプロモーターを含んでいてもよく、これをコードしてもよい。

【0139】

いくつかの実施形態において、本開示の遺伝子挿入システムにおいて使用される少なくとも1つのRTC：5'側モジュールは、配列番号58の5'UTRを含む。

【0140】

複数の実施形態において、本発明者らは、WHOにより報告されているビオンテック社製のワクチン配列から得られた、トランスフェクトされたmRNA用の5'UTRと3'UTRを1つずつ使用した。さらに、(転写後にポリAポリメラーゼを使用するのではなく)その鋳型によってコードされるポリA領域を使用した。このポリA領域は、30個のアデノシンと10ntのリンカーと70個のアデノシンで構成されていた。さらに、余分な3'末端塩基を付加することなくmRNA転写用の鋳型を切断するため、Type IIS制限部位を導入した。すべてのmRNAは、TriLink社製AG clean cap、すなわち、m7(3'OMeG)(5')ppp(5')(2'OMeA)pGでキャップした。UTRは、例えば、細胞の種類に応じて特異的な翻訳調節を行うため、組織特異的に逆転写酵素が発現されるように選択した。

【0141】

いくつかの実施形態において、RTC：5'側モジュールは、配列番号58と少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

【0142】

RTC：3'側モジュール

本発明の逆転写酵素コンストラクト(RTC)は、少なくとも1つのRTC：3'側モジュールを含んでいてもよい。通常、RTC：3'側モジュールは、翻訳されない生体高分子構成要素を含み、この翻訳されない生体高分子構成要素によって、例えば、遺伝子挿入コンストラクトの免疫原性の改変、標的とする細胞内領域への遺伝子挿入コンストラクトの局在化の補助、遺伝子挿入コンストラクトに含まれるRTC：RTモジュールの発現の制御または改変、同定用の遺伝子挿入コンストラクトの標識、遺伝子挿入コンストラクトの精製の補助、遺伝子挿入コンストラクトの分解の制御、遺伝子挿入コンストラクトの活性および/または機能の外因性または内因性の調節、ならびにこれらの任意の組み合わせを行ってもよいが、これらに限定されない。

【0143】

いくつかの実施形態において、少なくとも1つのRTC：3'側モジュールは、少なくとも1つの3'UTRを含んでいてもよい。いくつかの実施形態において、少なくとも1つのRTC：3'側モジュールは、少なくとも1つのポリAトラクトすなわちポリAテールを含んでいてもよく、これをコードしてもよい。いくつかの実施形態において、少なくとも1つのRTC：3'側モジュールは、少なくとも1つのマイクロRNA結合配列を含んでいてもよく、これをコードしてもよい。

【0144】

いくつかの実施形態において、本開示の遺伝子挿入システムにおいて使用される少なくとも1つのRTC：3'側モジュールは、配列番号59に示される3'UTRとポリAテールを含む。

【0145】

いくつかの実施形態において、RTC：3'側モジュールは、配列番号59と少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有す

る3' UTRを含む。

【0146】

逆転写酵素コンストラクト（RTC）のモジュール性

本発明の逆転写酵素コンストラクト（RTC）は、所望の機能または活性が得られるように、少なくとも1つのRTC：RTモジュール、任意の構成要素としての少なくとも1つのRTC：5'側モジュール、および／または任意の構成要素としての少なくとも1つのRTC：3'側モジュールを組み合わせることによって設計してもよい。いくつかの実施形態において、逆転写酵素コンストラクトは、少なくとも1つのRTC：5'側モジュールを含む。いくつかの実施形態において、逆転写酵素コンストラクトは、少なくとも1つのRTC：3'側モジュールを含む。いくつかの実施形態において、逆転写酵素コンストラクトは、少なくとも1つのRTC：RTモジュールを含む。いくつかの実施形態において、逆転写酵素コンストラクトは、少なくとも1つのRTC：5'側モジュールと、少なくとも1つのRTC：RTモジュールと、少なくとも1つのRTC：3'側モジュールを含む。いくつかの実施形態において、逆転写酵素コンストラクトは、少なくとも1つのRTC：5'側モジュールと少なくとも1つのRTC：RTモジュールを含む。いくつかの実施形態において、逆転写酵素コンストラクトは、少なくとも1つのRTC：RTモジュールと少なくとも1つのRTC：3'側モジュールを含む。

10

【0147】

いくつかの実施形態において、本発明の逆転写酵素コンストラクトは、少なくとも1つのRTC：5'側モジュールと少なくとも1つのRTC：3'側モジュールの両方を含んでいなくてもよい。いくつかの実施形態において、本発明の逆転写酵素コンストラクトは、少なくとも1つのRTC：5'側モジュールと少なくとも1つのRTC：3'側モジュールのいずれかを含んでいなくてもよい。いくつかの実施形態において、本発明の逆転写酵素コンストラクトは、少なくとも1つのRTC：5'側モジュールを含んでいなくてもよい。いくつかの実施形態において、本発明の逆転写酵素コンストラクトは、少なくとも1つのRTC：3'側モジュールを含んでいなくてもよい。

20

【0148】

いくつかの実施形態において、少なくとも1つの逆転写酵素コンストラクトは、（a）配列番号58から選択されるか、配列番号58をコードするか、配列番号58によってコードされる少なくとも1つのRTC：5'側モジュール；（b）配列番号1～57から選択されるか、配列番号1～57のいずれかをコードするか、配列番号1～57のいずれかによってコードされる少なくとも1つのRTC：RTモジュール；および／または（c）配列番号59から選択されるか、配列番号59をコードするか、配列番号59によってコードされる少なくとも1つのRTC：3'側モジュールのどのような組み合わせを含んでいてもよい。

30

【0149】

例示的な逆転写酵素コンストラクト（RTC）

本発明で使用される逆転写酵素コンストラクト（RTC）は、配列番号1～57の少なくとも1つを含んでいてもよく、配列番号1～57の少なくとも1つをコードしてもよく、配列番号1～57の少なくとも1つによってコードされてもよい。いくつかの実施形態において、逆転写酵素コンストラクト（RTC）は、配列番号1～57の少なくとも1つと少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

40

【0150】

いくつかの実施形態において、少なくとも1つの逆転写酵素コンストラクト（RTC）は、配列番号17～21の少なくとも1つと少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

50

【0151】

いくつかの実施形態において、少なくとも1つの逆転写酵素コンストラクト（RTC）は、配列番号26～29の少なくとも1つと少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

【0152】

いくつかの実施形態において、少なくとも1つの逆転写酵素コンストラクト（RTC）は、配列番号24または25と少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

10

【0153】

いくつかの実施形態において、少なくとも1つの逆転写酵素コンストラクト（RTC）は、配列番号1～5の少なくとも1つと少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

【0154】

いくつかの実施形態において、少なくとも1つの逆転写酵素コンストラクト（RTC）は、配列番号35～37の少なくとも1つと少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

20

【0155】

いくつかの実施形態において、少なくとも1つの逆転写酵素コンストラクト（RTC）は、配列番号32～34の少なくとも1つと少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

30

【0156】

いくつかの実施形態において、少なくとも1つの逆転写酵素コンストラクト（RTC）は、図5に示す構造を含む。

【0157】

逆転写酵素コンストラクト（RTC）の調節因子

本発明の逆転写酵素コンストラクト（RTC）は、どのような数の調節因子をさらにも含んでもよく、調節因子は、RTCモジュールのうちのいずれに含まれていてもよい。本明細書において、「調節因子」という用語は、コンストラクトの一部である生体高分子の発現または活性の制御を可能とする配列、領域またはドメインを意味する。

【0158】

例えば、RNAベースのRTCは、どのような数のマイクロRNA（miRNA）結合部位または低分子干渉RNA（siRNA）結合部位を含んでいてもよい。何らかの理論に拘束されることを望むものではないが、RNA干渉（RNAi）結合部位が存在することによって、RNAiトランスクリプトームの存在に基づいて特定の種類の細胞におけるRTタンパク質の発現を阻止してもよい。このようにして、本発明の遺伝子挿入システムは、対象の種類の細胞を標的から外すことができる。本明細書において、「miRNA結合部位またはsiRNA結合部位」という用語は、少なくとも1つのmiRNAまたはsiRNAに相補的なRNA配列を意味する。

40

【0159】

いくつかの実施形態において、RTCは、遺伝子挿入システムによって挿入される導入

50

遺伝子に含まれる少なくとも1つのmiRNAおよび／もしくはsiRNA、またはこの導入遺伝子によってコードされる少なくとも1つのmiRNAおよび／もしくはsiRNAに相補的な少なくとも1つのmiRNA結合部位および／またはsiRNA結合部位を含んでいてもよい。本発明の遺伝子挿入システムは、このようなmiRNA結合部位またはsiRNA結合部位を含むことによって、概して、1回の遺伝子挿入システムの投与により達成される導入遺伝子の挿入数を自身で調節することが可能となってもよく、かつ／または最初の投与後に導入遺伝子の反復挿入を防ぐことが可能になってもよい。このように、遺伝子挿入システムは、対象に対して再投与または共投与を行う能力が向上されていてもよい。

【0160】

遺伝子挿入コンストラクト（GIC）

本発明の遺伝子挿入システムは、少なくとも1つの遺伝子挿入コンストラクト（GIC）を含み、このGICは、通常、対象のゲノムへの挿入を意図した少なくとも1つの目的の配列（すなわち「ペイロード配列」）を含むか、このような配列をコードする。本明細書において、「遺伝子挿入コンストラクト（GIC）」という用語は、少なくとも1つのRNA配列を含むか、少なくとも1つのRNA配列をコードする生体高分子コンストラクトであって、少なくとも1つのRTC：RTモジュールに含まれるか、少なくとも1つのRTC：RTモジュールにコードされる少なくとも1つの逆転写酵素によって、このRNA配列が認識されることにより、このRNA配列が逆転写用の鋳型として機能することができるとを特徴とする生体高分子コンストラクトを意味する。いくつかの実施形態において、本発明の遺伝子挿入システムで使用される少なくとも1つの遺伝子挿入コンストラクトは、RNA、DNAまたはこれらの任意の組み合わせなどの核酸生体高分子を含んでいてもよいが、これらに限定されない。

【0161】

遺伝子挿入コンストラクト（GIC）の構造

本発明の遺伝子挿入コンストラクト（GIC）は、少なくとも1つのGIC：5'側モジュール、少なくとも1つのGIC：ペイロードモジュール、少なくとも1つのGIC：3'側モジュール、およびこれらの任意の組み合わせを含んでいてもよく、これらをコードしてもよい。いくつかの実施形態において、少なくとも1つのGICは、プラスミドを含んでいてもよく、またはプラスミドとして対象に送達してもよい。いくつかの実施形態において、少なくとも1つのGICは、直鎖RNAを含んでいてもよく、または直鎖RNAとして対象に送達してもよい。

【0162】

いくつかの実施形態において、前記少なくとも1つのGIC：5'側モジュールは任意の構成要素である。いくつかの実施形態において、前記少なくとも1つのGIC：3'側モジュールは任意の構成要素である。いくつかの実施形態において、本発明の遺伝子挿入コンストラクト（GIC）は、少なくとも1つのGIC：ペイロードモジュールを含むか、少なくとも1つのGIC：ペイロードモジュールをコードするが、少なくとも1つのGIC：5'側モジュールおよび／または少なくとも1つのGIC：3'側モジュールを含まず、これらをコードしていない。

【0163】

図6を参照すると、例示的な直鎖RNA遺伝子挿入コンストラクト（GIC）400が示されており、任意の構成要素であるGIC：5'側モジュール410は、GIC配列の5'末端からGIC：5'側モジュールの末端420まで伸長している。GIC：ペイロードモジュール430は、GIC：5'側モジュール（存在する場合）の3'側に向かって、GIC：ペイロードモジュールの末端440まで伸長している。さらに、GIC：3'側モジュール450は、GICの3'末端まで伸長している。これらの各特徴の詳細は以下で述べる。

【0164】

GIC：5'側モジュール

本開示の遺伝子挿入コンストラクト（GIC）において使用されるGIC：5'側モジュールは、天然レトロエレメントの5'領域に由来する少なくとも1つの配列を含んでいてもよ

10

20

30

40

50

く、このような配列をコードしてもよい。何らかの理論に拘束されることを望むものではないが、この5'側モジュールは、逆転写酵素の少なくとも1つのRNA結合ドメインと相互作用するRNA配列、導入遺伝子の挿入の際に第2鎖を合成することができるRNA配列、遺伝子挿入コンストラクトの免疫原性を減少させるRNA配列、遺伝子挿入コンストラクトの安定性および／もしくは精製に有用な特徴を提供するRNA配列、ならびにこれらの任意の組み合わせを含んでいてもよく、このようなRNA配列をコードしてもよい。

【0165】

GIC：5'側モジュールの構造

複数の実施形態において、5'側モジュールは、5' rRNA配列およびリボザイム（RZ）配列を含む。いくつかの実施形態において、5' rRNA配列およびRZ配列は、必ずしも完全に分離されている必要はない。いくつかの実施形態において、5'側モジュールは、RZ配列とは分離されていてもよい「フォールディング配列」を含む。いくつかの実施形態において、GIC：5'側モジュールは、少なくとも1つのrRNA配列（またはその他の標的部位配列）、少なくとも1つのリボザイム（RZ）配列、少なくとも1つのフォールディング配列、およびこれらの任意の組み合わせを含んでいてもよく、これらをコードしてもよい。

10

【0166】

図6を再度参照すると、GIC：5'側モジュール410の拡大図（左下）は、例示的な1つのGIC：5'側モジュールの構造を示している。GIC：5' rRNA配列411が、5'側モジュールの5'末端に存在する場合、このGIC：5' rRNA配列411は、標的挿入部位の5'側または標的挿入部位の近傍に位置する対象のDNA配列に相補的なRNA配列を含んでいてもよく、このようなRNA配列をコードしてもよい。GIC：5'側モジュールのリボザイム（RZ）配列412が存在する場合、このGIC：5'側モジュールのRZ配列は、自己切断型リボザイムの折り畳み構造を有する少なくとも1つのRNA配列を含んでいてもよく、このRNA配列は、自己切断することにより、転写された5'リーダー配列から機能性のGICを放出してもよく、このような自己切断が起こらなくてもよい。GIC：5'側モジュールのRZ配列は折り畳まれ、活性である場合、自己切断して、GICの5'末端またはその近傍においてGIC：5' rRNA配列をRZの一部として組み込む。GIC：5'側モジュールの任意の構成要素であるフォールディングモチーフ配列413は、自律的に折り畳まれることが予測または実証されている少なくとも1つのRNA配列を含んでいてもよく、このRNA配列は、GIC：5'側モジュールのRZの折り畳みとペイロード配列の折り畳みとを物理的および／または動態学的に分離するのに有用であると考えられる。さらに、領域414内において、またはGIC 5'側モジュール410とペイロードモジュール430の間の位置420において、細胞内で標的部位に隣接する内因性プロモーター配列から開始される転写を停止させるか、その他の制御を行うためにGIC配列を付加してもよい。いくつかの実施形態において、細胞内において標的部位に隣接する内因性プロモーター配列をペイロードの発現に利用してもよく、この態様は、GIC配列を位置420および／または位置440に付加することによってペイロードの発現を調節する場合の一例である（例えば、宿主プロモーターによるペイロード配列を含むRNA転写産物の翻訳開始または転写終止を行う）。さらに、領域414は、RNAポリメラーゼによる標的挿入部位の遺伝子のリードスルー（読み飛ばし）を阻止するために、RNAポリメラーゼ（RNAP）終止配列を含んでいてもよい。いくつかの実施形態において、RNAPはRNAP I（Pol I）であり、GICペイロードモジュールがリボソームDNAの遺伝子標的部位に組み込まれた際に、この終止配列によってPol Iのリードスルー転写が阻止される。いくつかの実施形態において、RNAP転写終止配列は、5'-AGGTCGACCGAGATGTCCGAGGTCGACCGAGTTGTCCG-3'（配列番号537）で示される配列を含む。

20

30

40

【0167】

GIC：5'側モジュールのrRNA配列

GIC：5'側モジュールの少なくとも1つのrRNA配列は、GIC：5'側モジュールの任意の構成要素である。GIC：5'側モジュールにrRNA配列が存在する場合、このrRNA配列

50

は、標的挿入部位の5'側に位置する少なくとも1つの対象のDNA配列と同種かつ／または相補的なヒトリボソームRNA (rRNA) 配列またはその他の配列を含んでいてもよく、このような配列をコードしてもよい。何らかの理論に拘束されることを望むものではないが、このrRNA配列は、少なくとも1つの内因性DNA修復機構を動員することによって、挿入されたcDNA導入遺伝子の第2鎖の合成を誘導してもよい。いくつかの実施形態において、GIC: 5'側モジュールのrRNA配列は、GIC: 5'側モジュールのRZ配列の5'側に位置する。いくつかの実施形態において、GIC: 5'側モジュールは、rRNAゲノム配列を含む配列を含んでいない。

【0168】

いくつかの実施形態において、GIC: 5'側モジュールの少なくとも1つのrRNA配列は、約1~36ntのrRNAを含んでいてもよく、約1~36ntのrRNAをコードしてもよい。いくつかの実施形態において、GIC: 5'側モジュールの少なくとも1つのrRNA配列は、約1~30ntのrRNAを含んでいてもよく、約1~30ntのrRNAをコードしてもよい。いくつかの実施形態において、GIC: 5'側モジュールの少なくとも1つのrRNA配列は、約1~28ntのrRNAを含んでいてもよく、約1~28ntのrRNAをコードしてもよい。いくつかの実施形態において、GIC: 5'側モジュールの少なくとも1つのrRNA配列は、約1~26ntのrRNAを含んでいてもよく、約1~26ntのrRNAをコードしてもよい。いくつかの実施形態において、GIC: 5'側モジュールの少なくとも1つのrRNA配列は、約1~13ntのrRNAを含んでいてもよく、約1~13ntのrRNAをコードしてもよい。いくつかの実施形態において、GIC: 5'側モジュールの少なくとも1つのrRNA配列は、約1~11ntのrRNAを含んでいてもよく、約1~11ntのrRNAをコードしてもよい。

【0169】

いくつかの実施形態において、GIC: 5'側モジュールの少なくとも1つのrRNA配列は、約1nt、約2nt、約3nt、約4nt、約5nt、約6nt、約7nt、約8nt、約9nt、約10nt、約11nt、約12nt、約13nt、約14nt、約15nt、約16nt、約17nt、約18nt、約19nt、約20nt、約21nt、約22nt、約23nt、約24nt、約25nt、約26nt、約27nt、約28nt、約29nt、約30nt、約31nt、約32nt、約33nt、約34nt、約35nt、または約36ntのrRNAを含んでいてもよく、このような長さのrRNAをコードしてもよい。

【0170】

いくつかの実施形態において、GIC: 5'側モジュールの少なくとも1つのrRNA配列は、約30ntのrRNAを含んでいてもよく、約30ntのrRNAをコードしてもよい。いくつかの実施形態において、GIC: 5'側モジュールの少なくとも1つのrRNA配列は、約36ntのrRNAを含んでいてもよく、約36ntのrRNAをコードしてもよい。いくつかの実施形態において、GIC: 5'側モジュールの少なくとも1つのrRNA配列は、約28ntのrRNAを含んでいてもよく、約28ntのrRNAをコードしてもよい。いくつかの実施形態において、GIC: 5'側モジュールの少なくとも1つのrRNA配列は、約26ntのrRNAを含んでいてもよく、約26ntのrRNAをコードしてもよい。いくつかの実施形態において、GIC: 5'側モジュールの少なくとも1つのrRNA配列は、約13ntのrRNAを含んでいてもよく、約13ntのrRNAをコードしてもよい。いくつかの実施形態において、GIC: 5'側モジュールの少なくとも1つのrRNA配列は、約11ntのrRNAを含んでいてもよく、約11ntのrRNAをコードしてもよい。いくつかの実施形態において、GIC: 5'側モジュールのrRNA配列は、5' Gヌクレオチドを含む。

【0171】

いくつかの実施形態において、GIC: 5'側モジュールの少なくとも1つのrRNA配列は、配列番号250~276の少なくとも1つを含んでいてもよく、これらの配列の少なくとも1つをコードしてもよく、これらの配列の少なくとも1つによってコードされてもよい。いくつかの実施形態において、GIC: 5'側モジュールの少なくとも1つのrRNA配列は、配列番号250~276の少なくとも1つと少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%または70%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。いくつかの実施

形態において、GIC：5'側モジュールの少なくとも1つのrRNA配列は、配列番号250～276からなる群から選択される配列と比較して1つ、2つまたは3つのヌクレオチド変化またはヌクレオチド置換を有する配列を含む。

【0172】

いくつかの実施形態において、GIC：5'側モジュールの少なくとも1つのrRNA配列は、配列番号252と少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%または70%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。いくつかの実施形態において、GIC：5'側モジュールの少なくとも1つのrRNA配列は、配列番号252と比較して1つ、2つまたは3つのヌクレオチド変化を有する配列を含む。

10

【0173】

いくつかの実施形態において、GIC：5'側モジュールの少なくとも1つのrRNA配列は、配列番号254と少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%または70%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。いくつかの実施形態において、GIC：5'側モジュールの少なくとも1つのrRNA配列は、配列番号254と比較して1つ、2つまたは3つのヌクレオチド変化を有する配列を含む。

【0174】

いくつかの実施形態において、GIC：5'側モジュールの少なくとも1つのrRNA配列は、配列番号255と少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%または70%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。いくつかの実施形態において、GIC：5'側モジュールの少なくとも1つのrRNA配列は、配列番号255と比較して1つ、2つまたは3つのヌクレオチド変化を有する配列を含む。

20

【0175】

GIC：5'側モジュールのRZ配列

GIC：5'側モジュールのRZ配列は、GIC：5'側モジュールの任意の構成要素であり、GIC：5'側モジュールに存在する場合、少なくとも1つの自己切断型リボザイムまたは自己切断型リボザイムの折り畳み構造を持つ少なくとも1つの配列を含むか、これらをコードする（これらを合わせて「RZ」と呼ぶ）。何らかの理論に拘束されることを望むものではないが、このようなRZモチーフは、GICに5'OH末端を生成し、この5'OH末端は、例えば、自己切断により生じた5'末端などであってもよく、安定な三次構造では、生成された5'OH末端によって、外因性RNAに対する自然免疫応答が低減されてもよく、5'モノリン酸に依存して切断を開始する5'-3'エキソヌクレアーゼによるGICの分解が低減されてもよく、対象の細胞によって、GICが鋳型RNAとしてではなく、mRNAやその他の望ましくない種類のRNAとして認識されてしまう可能性を低減してもよい。

30

【0176】

いくつかの実施形態において、GIC：5'側モジュールの少なくとも1つのRZ配列は、少なくとも1つの非LTR型レトロエレメントの5'領域に由来するリボザイムを含むか、このようなリボザイムをコードする。いくつかの実施形態において、GIC：5'側モジュールの少なくとも1つのRZ配列は、イトヨ、アメリカカブトガニ、キタノトミヨ、キョウソヤドリコバチ、ガラパゴスフィンチ、メダカ、ノドジロシトド、キンカチョウ、コクヌストモドキ（例えば、A系統またはB系統に由来するR2）、ノドジロシギダチョウ、その他の鳥類、その他の節足動物、その他の魚類、その他の尾索動物、その他の動物または同様のゲノムから得られる非LTR型レトロエレメントの5'領域に由来するリボザイムを含むか、このようなリボザイムをコードする。

40

【0177】

いくつかの実施形態において、GIC：5'側モジュールのRZ配列は、デルタ肝炎ウイルス（HDV）のRZの二次構造および三次構造を形成することが可能なRZを含むか、このようなRZをコードし、このRZは、自然界で見られる配列から改変されたものであっても

50

よく、かつ／または公知のゲノム配列を使用せずにデノボ設計されたものであってもよい。いくつかの実施形態において、HDVにおいて対合したステムP1と対合したステムP2を架橋する折り畳まれたRZ配列は、ジャンクション(J) 1/2とも呼ばれ、その一部または全体が、所望の長さの標的部位配列(例えば、5' rRNA)に含まれるか、ステムループの形成によりさらに保護された所望の標的部位配列に含まれる。いくつかの実施形態では、HDVの折り畳まれたRZ配列の対合したステム4(P4)を設計に組み込むことにより、例えば、PP7ファージまたはMS2ファージのコートタンパク質の天然配列または改変配列に結合させることによって、変性させずに遺伝子挿入コンストラクト(GIC)を精製することを可能としてもよい。いくつかの実施形態において、RZ配列は、非生産的なフォールディングを最小限に抑えたり、排除したりするために設計および最適化される。いくつかの実施形態において、RZ配列は、ウリジンヌクレオチドの数を最小限に抑えるために設計および最適化される。いくつかの実施形態において、RZ配列は、鑄型RNAの合成過程において組み込まれるヌクレオチド類似体で標準的なリボヌクレオチドのすべてまたはその一部が置換できるように設計および最適化される。

10

20

30

40

50

【0178】

いくつかの実施形態において、GIC: 5'側モジュールの少なくとも1つのRZ配列は、配列番号60~154の少なくとも1つを含んでいてもよく、配列番号60~154の少なくとも1つをコードしてもよく、配列番号60~154の少なくとも1つによってコードされてもよい。いくつかの実施形態において、RZ配列は、自律的に折り畳まれて、活性なリボザイムを形成する。いくつかの実施形態において、RZ配列は、その5'末端に内部rRNA配列を含む。いくつかの実施形態において、RZ配列は、5'末端または3'末端が伸長した配列である。いくつかの実施形態において、RZ配列は、触媒不活性型RZ配列である。いくつかの実施形態において、GIC: 5'側モジュールの少なくとも1つのRZ配列は、配列番号60~154の少なくとも1つと少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。いくつかの実施形態において、GIC: 5'側モジュールのRZ配列は、配列番号60~154と少なくとも90%、91%、92%、93%、94%、95%、96%、97%、98%または99%の同一性を有する非天然配列を含む。

【0179】

いくつかの実施形態において、GIC: 5'側モジュールの少なくとも1つのRZ配列は、配列番号60と少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

【0180】

いくつかの実施形態において、GIC: 5'側モジュールの少なくとも1つのRZ配列は、配列番号64と少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

【0181】

いくつかの実施形態において、GIC: 5'側モジュールの少なくとも1つのRZ配列は、配列番号67と少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

【0182】

いくつかの実施形態において、GIC: 5'側モジュールの少なくとも1つのRZ配列は、

配列番号101と少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

【0183】

いくつかの実施形態において、GIC：5'側モジュールの少なくとも1つのRZ配列は、配列番号121と少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

10

【0184】

いくつかの実施形態において、GIC：5'側モジュールの少なくとも1つのRZ配列は、配列番号122と少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

【0185】

いくつかの実施形態において、GIC：5'側モジュールの少なくとも1つのRZ配列は、配列番号137と少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

20

【0186】

GIC：5'側モジュールのフォールディング配列

GIC：5'側モジュールのフォールディング配列は、5'側モジュールの任意の構成要素であり、このフォールディング配列が存在する場合、このフォールディング配列は、特別に設計された構造を有する少なくとも1つのRNA配列モチーフを含む。いくつかの実施形態において、自律型フォールディングRNA配列モチーフは少なくとも1つのヘアピンモチーフを含み、このヘアピンモチーフは、例えば、RZが、後に転写されるペイロード領域と塩基対を形成することによって、RZ配列のミスフォールディングを遮断した後に存在していてもよい。いくつかの実施形態において、生産的な鑄型RNAのフォールディングを改善するように設計された5'側モジュール領域は、ペイロードモジュールまたは3'側モジュールの別の鑄型RNA領域と直接または間接的に塩基対を形成するか、相互作用する。いくつかの実施形態において、鑄型RNAのフォールディングを誘導する少なくとも1つのRNA配列モチーフは、別のステムループモチーフにタンパク質架橋を結合させる少なくとも1つのステムループモチーフを含んでいてもよい。いくつかの実施形態において、5'側モジュールのフォールディング配列は、逆転写酵素をコードするmRNAと鑄型RNAのペ어링を促すことによって、例えば、逆転写酵素をコードするmRNAと鑄型RNAが1：1の化学量論比で個々の送達媒体と一緒にパッケージングされることを促してもよい。いくつかの実施形態において、5'側モジュールのフォールディング配列は、例えば、鑄型RNAの安定化、局在化、および／またはその他の有用な結果を達成することを目的として、標的細胞の内因性RNAと鑄型RNAのペ어링を促してもよい。

30

40

【0187】

いくつかの実施形態において、GIC：5'側モジュールの少なくとも1つのフォールディング配列は、配列番号278～279の少なくとも1つを含んでいてもよく、これらの配列の少なくとも1つをコードしてもよく、これらの配列の少なくとも1つによってコードされてもよい。いくつかの実施形態において、GIC：5'側モジュールの少なくとも1つのフォールディング配列は、配列番号278～279の少なくとも1つと少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を

50

含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。いくつかの実施形態において、GIC：5'側モジュールのフォールディング配列は、配列番号278～279と少なくとも90%、91%、92%、93%、94%、95%、96%、97%、98%または99%の同一性を有する非天然配列を含む。

【0188】

いくつかの実施形態において、GIC：5'側モジュールの少なくとも1つのフォールディング配列は、配列番号278と少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

10

【0189】

いくつかの実施形態において、GIC：5'側モジュールの少なくとも1つのフォールディング配列は、配列番号279と少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

【0190】

GIC：5'側モジュールのモジュール性

本明細書において開示した5'側モジュールの各構成要素は、コンビナトリアルな方法で互いに交換可能に使用して、特定の遺伝子挿入システムに必要な機能性または所望の機能性を有する5'側モジュールを設計してもよい。

20

【0191】

いくつかの実施形態において、少なくとも1つのGIC：5'側モジュールは、GIC：5'側モジュールの少なくとも1つのrRNA配列を含む。いくつかの実施形態において、少なくとも1つのGIC：5'側モジュールは、GIC：5'側モジュールの少なくとも1つのRZ配列を含む。いくつかの実施形態において、少なくとも1つのGIC：5'側モジュールは、GIC：5'側モジュールの少なくとも1つのフォールディング配列を含む。いくつかの実施形態において、少なくとも1つのGIC：5'側モジュールは、GIC：5'側モジュールの少なくとも1つのrRNA配列と、GIC：5'側モジュールの少なくとも1つのRZ配列とを含む。いくつかの実施形態において、少なくとも1つのGIC：5'側モジュールは、GIC：5'側モジュールの少なくとも1つのrRNA配列と、GIC：5'側モジュールの少なくとも1つのRZ配列と、GIC：5'側モジュールの少なくとも1つのフォールディング配列とを含む。

30

【0192】

いくつかの実施形態において、少なくとも1つのGIC：5'側モジュールは、

(a) 配列番号250～276のいずれか1つから選択されるか、これらのいずれか1つをコードするか、これらのいずれか1つによってコードされる少なくとも1つのrRNA配列；

(c) 配列番号60～154のいずれか1つから選択されるか、これらのいずれか1つをコードするか、これらのいずれか1つによってコードされる少なくとも1つのRZ配列；および／または

40

(d) 配列番号278～279のいずれか1つから選択されるか、これらのいずれか1つをコードするか、これらのいずれか1つによってコードされる少なくとも1つのフォールディング配列

をどのような組み合わせで含んでいてもよい。

【0193】

例示的なGIC：5'側モジュール

いくつかの実施形態において、少なくとも1つのGIC：5'側モジュールは、配列番号60～154の少なくとも1つを含んでいてもよく、配列番号60～154の少なくとも1つをコードしてもよく、配列番号60～154の少なくとも1つによってコードされてもよい。

いくつかの実施形態において、少なくとも1つのGIC：5'側モジュールは、配列番号60

50

～154の少なくとも1つと少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。いくつかの実施形態において、GIC：5'側モジュールは、配列番号60～154と少なくとも90%、91%、92%、93%、94%、95%、96%、97%、98%または99%の同一性を有する非天然配列を含む。

【0194】

いくつかの実施形態において、少なくとも1つのGIC：5'側モジュールは、配列番号60、61、77および79～83の少なくとも1つと少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

10

【0195】

いくつかの実施形態において、少なくとも1つのGIC：5'側モジュールは、配列番号62または63と少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

【0196】

いくつかの実施形態において、少なくとも1つのGIC：5'側モジュールは、配列番号121と少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

20

【0197】

いくつかの実施形態において、少なくとも1つのGIC：5'側モジュールは、配列番号117～119の少なくとも1つと少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

30

【0198】

GIC：3'側モジュール

【0199】

本開示の遺伝子挿入コンストラクト（GIC）において使用される3'側モジュールは、天然レトロエレメントの3'UTRに由来する少なくとも1つの配列を含んでいてもよく、このような配列をコードしてもよい。通常、3'側モジュールは、逆転写酵素による遺伝子挿入コンストラクトの認識および結合を促進する構成要素、逆転写用のペイロードモジュールを配置する構成要素、および遺伝子挿入コンストラクトのRNAを安定化させる構成要素を含む。

【0200】

40

GIC：3'側モジュールの構造

いくつかの実施形態において、GIC：3'側モジュールは、少なくとも1つの逆転写酵素（RT）認識配列、任意の構成要素である少なくとも1つのrRNA配列、任意の構成要素である少なくとも1つのAトラクト配列、およびこれらの任意の組み合わせを含んでいてもよく、これらをコードしてもよい。

【0201】

図6を再度参照すると、拡大図（右下）は、GIC：3'側モジュール450の構造の一例を示す。GIC：3'側モジュールの5'末端は、GIC：3'側モジュールの逆転写酵素認識配列451であり、この逆転写酵素認識配列451は、少なくとも1つの逆転写酵素により認識または結合される配列を含んでいてもよく、このような配列をコードしてもよい。GIC：3'

50

側モジュールのrRNA配列452が存在する場合、このrRNA配列452は、GIC：3'側モジュールの逆転写酵素認識配列の3'側に位置していてもよく、標的部領域と同種の配列を含むか、これをコードしてもよく、例えば、TPRTプライマーの3'末端と塩基対を形成できる28S rRNAヌクレオチドであってもよい。さらに、GIC：3'側モジュールのAトラクト配列453が存在する場合、このAトラクト配列453は、長さが制限されていてもよいアデノシンに富む配列またはアデノシンが直列に並んだ配列を含んでいてもよく、この配列の長さは、例えば、10～60ntであってもよく、Aトラクト配列453は、GIC：3'側モジュールの3'末端に配置されていてもよい。

【0202】

GIC：3'側モジュールの逆転写酵素認識配列

GIC：3'側モジュールの逆転写酵素認識配列は、少なくとも1つの逆転写酵素と相互作用する少なくとも1つの配列、または少なくとも1つの逆転写酵素によって認識される少なくとも1つの配列を含んでいてもよく、このような配列をコードしてもよい。何らかの理論に拘束されることを望むものではないが、GIC：3'側モジュールの逆転写酵素認識配列に含まれる少なくとも1つのRNA配列は、レトロエレメント逆転写酵素などの逆転写酵素の少なくとも1つの鋳型RNA結合ドメインと少なくとも一時的に結合してもよい。GIC：3'側モジュールの逆転写酵素認識配列の長さおよび配列のアイデンティティは、遺伝子挿入コンストラクト（GIC）に逆転写酵素を配置させるように構成されていてもよく、その結果、逆転写酵素によって逆転写された第1のヌクレオチドを、挿入される導入遺伝子の意図する3'末端に配置させることができる。本明細書において、GIC：3'側モジュールの逆転写酵素認識配列は、「GIC：3'側モジュールの3'UTR」と呼ぶこともできる。

【0203】

いくつかの実施形態において、少なくとも1つのGIC：3'側モジュールの逆転写酵素認識配列は、天然レトロエレメントの3'領域に由来するものであるか、天然レトロエレメントの3'領域を含む。いくつかの実施形態において、少なくとも1つのGIC：3'側モジュールの逆転写酵素認識配列は、イトヨ、キロショウジョウバエ、アメリカカブトガニ、カタノトミヨ、キョウソヤドリコバチ、ガラパゴスフィンチ、メダカ、ノドジロシトド、キンカチョウ、コクヌストモドキ、ノドジロシギダチョウ、オナジショウジョウバエ、カイコガ、A.vaga、その他の鳥類、その他の節足動物、その他の魚類、その他の尾索動物、その他の動物、またはこれらと同様の生物のゲノムから得られる非LTR型レトロエレメントの3'領域に由来するものである。いくつかの実施形態において、GIC：3'側モジュールの少なくとも1つの逆転写酵素認識配列は、フォールディングの安定性または均一性を増加させることによって、天然レトロエレメントの3'領域から改変されたものである。いくつかの実施形態において、GIC：3'側モジュールの少なくとも1つの逆転写酵素認識配列は、逆転写酵素との相互作用において所望の親和性および／または特異性を有するように設計および／または選択されたものであり、あるいは逆転写用の鋳型として所望の機能を付与する別の機構を有するように設計および／または選択されたものである。いくつかの実施形態において、GIC：3'側モジュールの少なくとも1つの逆転写酵素認識配列は、標的細胞の内因性構成要素と相互作用したり、これに影響を与えたりしないように設計および／または選択されたものであり、かつ／または宿主細胞に有害な影響を及ぼさないように設計および／または選択されたものである。

【0204】

いくつかの実施形態において、GIC：3'側モジュールの少なくとも1つの逆転写酵素認識配列（すなわち、GIC：3'側モジュールの3'UTR配列）は、配列番号200～224の少なくとも1つを含んでいてもよく、配列番号200～224の少なくとも1つをコードしてもよく、配列番号200～224の少なくとも1つによってコードされてもよい。いくつかの実施形態において、GIC：3'側モジュールの少なくとも1つの逆転写酵素認識配列は、配列番号200～221の少なくとも1つと少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35

10

20

30

40

50

%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。いくつかの実施形態において、GIC：3'側モジュールの逆転写酵素認識配列は、配列番号200~224と少なくとも90%、91%、92%、93%、94%、95%、96%、97%、98%または99%の同一性を有する非天然配列である。

【0205】

いくつかの実施形態において、GIC：3'側モジュールの少なくとも1つの逆転写酵素認識配列は、配列番号202と少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコード

10

【0206】

いくつかの実施形態において、GIC：3'側モジュールの少なくとも1つの逆転写酵素認識配列は、配列番号204、222、223または224と少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

【0207】

いくつかの実施形態において、GIC：3'側モジュールの少なくとも1つの逆転写酵素認識配列は、配列番号203と少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコード

20

【0208】

いくつかの実施形態において、GIC：3'側モジュールは、逆転写酵素コンストラクト（RTC）によってコードされる逆転写酵素の由来生物種とは異なる生物種に由来する逆転写酵素認識配列を含む。例えば、いくつかの実施形態において、逆転写酵素認識配列は鳥類の1種に由来するものであってもよく、逆転写酵素は鳥類の別の種に由来するものであってもよい。いくつかの実施形態において、逆転写酵素認識配列は、ノドジロシトド、キンカチョウ、ノドジロシギダチョウおよびガラパゴスフィンチのいずれか1種から選択された鳥類に由来するものであり、逆転写酵素は、選択された鳥類とは異なる鳥類から選択される（例えば、ノドジロシトド、キンカチョウ、ノドジロシギダチョウまたはガラパゴスフィンチ）。いくつかの実施形態において、RTCコンストラクトによってコードされる逆転写酵素は、ノドジロシトド、キンカチョウ、ノドジロシギダチョウおよびガラパゴスフィンチのいずれか1種から選択され、逆転写酵素認識配列は、選択された鳥類とは異なる鳥類から選択される（例えば、ノドジロシトド、キンカチョウ、ノドジロシギダチョウまたはガラパゴスフィンチ）。いくつかの実施形態において、RTCコンストラクトによってコードされる逆転写酵素は、配列番号18または20と少なくとも90%、91%、92%、93%、94%、95%、96%、97%、98%、99%または100%の同一性を有するアミノ酸配列から選択され、逆転写酵素認識配列は、配列番号203、204、205または222~224と少なくとも90%、91%、92%、93%、94%、95%、96%、97%、98%、99%または100%の同一性を有するアミノ酸配列から選択される。いくつかの実施形態において、RTCコンストラクトによってコードされる逆転写酵素は、配列番号27または29と少なくとも90%、91%、92%、93%、94%、95%、96%、97%、98%、99%または100%の同一性を有するアミノ酸配列から選択され、逆転写酵素認識配列は、配列番号202、204、205または222~224と少なくとも90%、91%、92%、93%、94%、95%、96%、97%、98%、99%または100%の同一性を有するアミノ酸配列から選択される。いくつかの実施形態において、RTCコンストラクトによってコードされる逆転写酵素は、配列番号25と少なくとも90%、91%、92%、93%、94%、95%、96%、97%、98%、99%または100%の同一性を有するアミノ酸配列から選択され、逆転写酵素認識配列は、

30

40

50

配列番号202、203、204または222～224と少なくとも90%、91%、92%、93%、94%、95%、96%、97%、98%、99%または100%の同一性を有するアミノ酸配列から選択される。いくつかの実施形態において、RTCコンストラクトによってコードされる逆転写酵素は、配列番号31と少なくとも90%、91%、92%、93%、94%、95%、96%、97%、98%、99%または100%の同一性を有するアミノ酸配列から選択され、逆転写酵素認識配列は、配列番号202、203または205と少なくとも90%、91%、92%、93%、94%、95%、96%、97%、98%、99%または100%の同一性を有するアミノ酸配列から選択される。

【0209】

GIC：3'側モジュールのrRNA配列

GIC：3'側モジュールのrRNA配列、すなわち、非rDNA標的部位に導入されたニックのすぐ下流のTPRTプライマーと塩基対を形成する配列は、3'側モジュールの任意の構成要素であり、3'側モジュールに存在する場合、このrRNA配列は、ヒトリボソームRNA（rRNA）配列を含んでいてもよい。何らかの理論に拘束されることを望むものではないが、GIC：3'側モジュールのrRNA配列の長さや配列のアイデンティティは、本明細書に開示された遺伝子挿入システムが、どの程度の正確性で効率的に対象のゲノムに導入遺伝子を挿入できるのかということに影響を及ぼす。例えば、選択されたGIC：3'側モジュールのrRNA配列の長さによっては、内部配列から逆転写が開始されて、挿入された導入遺伝子が効率よく短縮されたり、オフターゲット部位への挿入が可能になる可能性があり、いずれの場合でも、目的の標的部位における導入遺伝子の挿入効率および挿入の特異性が低下する。RTCおよびGICは、標的部位に導入されたニックのすぐ下流のプライマー配列とGIC：3'側モジュールのrRNA配列との間で形成される塩基対が特定の長さである必要があるように構成される。これによって、標的部位の利用に関して忠実度をさらに向上させ、導入遺伝子が挿入された正確な連結部をさらに効率よく得ることができる。GIC：3' rRNAの最適な長さは20nt未満であり、特に4ntであれば、標的部位のニックに形成された4bpの塩基対全体から強力な刺激を受けることができる。したがって、RTCがランダムにニックを導入する場合、4ntのGIC：3' rRNAを使用すれば、256個のニックのうち1個のみが最適な導入遺伝子の挿入を達成すると考えられる。

【0210】

いくつかの実施形態において、GIC：3'側モジュールの少なくとも1つのrRNA配列は、約1～30ntのrRNAを含んでいてもよく、これをコードしてもよい。いくつかの実施形態において、GIC：3'側モジュールの少なくとも1つのrRNA配列は、約1～20ntのrRNAを含んでいてもよく、これをコードしてもよい。いくつかの実施形態において、GIC：3'側モジュールの少なくとも1つのrRNA配列は、約1～10ntのrRNAを含んでいてもよく、これをコードしてもよい。いくつかの実施形態において、GIC：3'側モジュールの少なくとも1つのrRNA配列は、約1～5ntのrRNAを含んでいてもよく、これをコードしてもよい。

【0211】

いくつかの実施形態において、GIC：3'側モジュールの少なくとも1つのrRNA配列は、約1nt、約2nt、約3nt、約4nt、約5nt、約6nt、約7nt、約8nt、約9nt、約10nt、約11nt、約12nt、約13nt、約14nt、約15nt、約16nt、約17nt、約18nt、約19nt、約20nt、約21nt、約22nt、約23nt、約24nt、約25nt、約26nt、約27nt、約28nt、約29nt、または約30ntのrRNAの一部を含んでいてもよく、このような長さのrRNAの一部をコードしてもよい。

【0212】

いくつかの実施形態において、GIC：3'側モジュールの少なくとも1つのrRNA配列は、約20ntのrRNAを含んでいてもよく、約20ntのrRNAをコードしてもよい。いくつかの実施形態において、GIC：3'側モジュールの少なくとも1つのrRNA配列は、約4ntのrRNAを含んでいてもよく、約4ntのrRNAをコードしてもよい。いくつかの実施形態において、GIC：3'側モジュールの少なくとも1つのrRNA配列は、約10ntのrRNAを含ん

でいてもよく、約10ntのrRNAをコードしてもよい。

【0213】

いくつかの実施形態において、GIC：3'側モジュールの少なくとも1つのrRNA配列は、配列番号280~285の少なくとも1つを含んでいてもよい。いくつかの実施形態において、GIC：3'側モジュールの少なくとも1つのrRNA配列は、配列番号280~289、および配列番号280~289に1つ、2つまたは3つのヌクレオチド置換を含む配列からなる群から選択される。

【0214】

GIC：3'側モジュールのAトラクト配列

GIC：3'側モジュールのAトラクト配列は、3'側モジュールの任意の構成要素であり、3'側モジュールに存在する場合、このAトラクト配列は、直列に並んだ複数のアデノシン（A）を含む末端ポリ配列を含む。何らかの理論に拘束されることを望むものではないが、GIC：3'側モジュールのAトラクト配列は、3'側のさらなるプロセッシングからGICを安定化または保護してもよく、さらに、細胞において、GICがmRNAとして認識されて、リボ核タンパク質の組み立て、輸送および翻訳に関連したGICの分解を起こりにくくする。さらに、GIC：3'側モジュールの少なくとも1つのAトラクト配列は、一般的な一本鎖RNA結合タンパク質による結合からGICを保護してもよく、標的部位プライマーと塩基対を形成できるようにGIC：3' rRNA配列を配置することを補助してもよい。明確化のために述べると、このAトラクト配列は、通常約100~200ntを上回る数のアデノシンが直列に並んだ天然のmRNAポリAテール配列とは異なるものである。

【0215】

いくつかの実施形態において、GIC：3'側モジュールの任意の構成要素である少なくとも1つのAトラクト配列は、約1~50個のアデノシンからなる配列を含むか、約1~50個のアデノシンからなる配列をコードする。例えば、GIC：3'側モジュールの任意構成要素であるAトラクト配列は、約1~50個のアデノシン、約5~50個のアデノシン、約10~50個のアデノシン、約15~50個のアデノシン、約20~50個のアデノシン、約25~50個のアデノシン、約30~50個のアデノシン、約35~50個のアデノシン、約40~50個のアデノシン、約45~50個のアデノシン、約1~45個のアデノシン、約5~45個のアデノシン、約10~45個のアデノシン、約15~45個のアデノシン、約20~45個のアデノシン、約25~45個のアデノシン、約30~45個のアデノシン、約35~45個のアデノシン、約40~45個のアデノシン、約1~40個のアデノシン、約5~40個のアデノシン、約10~40個のアデノシン、約15~40個のアデノシン、約20~40個のアデノシン、約25~40個のアデノシン、約30~40個のアデノシン、約35~40個のアデノシン、約1~35個のアデノシン、約5~35個のアデノシン、約10~35個のアデノシン、約15~35個のアデノシン、約20~35個のアデノシン、約25~35個のアデノシン、約30~35個のアデノシン、約1~30個のアデノシン、約5~30個のアデノシン、約10~30個のアデノシン、約15~30個のアデノシン、約20~30個のアデノシン、約25~30個のアデノシン、約1~25個のアデノシン、約5~25個のアデノシン、約10~25個のアデノシン、約15~25個のアデノシン、約20~25個のアデノシン、約1~20個のアデノシン、約5~20個のアデノシン、約10~20個のアデノシン、約15~20個のアデノシン、約1~15個のアデノシン、約5~15個のアデノシン、約10~15個のアデノシン、約1~10個のアデノシン、約5~10個のアデノシン、または約1~5個のアデノシンからなる配列を含んでいてもよく、このような配列をコードしてもよい。いくつかの実施形態において、GIC：3'側モジュールのAトラクト配列は、約1~100個、約1~90個、約1~80個、約1~70個または約1~60個のアデノシンを含む。

【0216】

いくつかの実施形態において、GIC：3'側モジュールの任意の構成要素である少なくとも1つのAトラクト配列は、約20~25個のアデノシンからなる配列を含むか、約20~25個のアデノシンからなる配列をコードする。

【0217】

10

20

30

40

50

いくつかの実施形態において、GIC：3'側モジュールの任意の構成要素である少なくとも1つのAトラクト配列は、約1個、約2個、約3個、約4個、約5個、約6個、約7個、約8個、約9個、約10個、約11個、約12個、約13個、約14個、約15個、約16個、約17個、約18個、約19個、約20個、約21個、約22個、約23個、約24個、約25個、約26個、約27個、約28個、約29個、約30個、約31個、約32個、約33個、約34個、約35個、約36個、約37個、約38個、約39個、約40個、約41個、約42個、約43個、約44個、約45個、約46個、約47個、約48個、約49個または約50個のアデノシンからなる配列を含むか、このような個数のアデノシンからなる配列をコードする。いくつかの実施形態において、GIC：3'側モジュールのAトラクト配列は、22個のアデノシンを含む。

【0218】

10

GIC：3'側モジュールのモジュール性

本明細書において開示した3'側モジュールの各構成要素は、コンビナトリアルな方法で互いに交換可能に使用して、特定の遺伝子挿入システムに必要な機能性または所望の機能性を有する3'側モジュールを設計してもよい。

【0219】

いくつかの実施形態において、少なくとも1つのGIC：3'側モジュールは、GIC：3'側モジュールの少なくとも1つの逆転写酵素認識配列を含む。いくつかの実施形態において、少なくとも1つのGIC：3'側モジュールは、GIC：3'側モジュールの少なくとも1つのrRNA配列を含む。いくつかの実施形態において、少なくとも1つのGIC：3'側モジュールは、GIC：3'側モジュールの少なくとも1つのAトラクト配列を含む。いくつかの実施形態において、少なくとも1つのGIC：3'側モジュールは、GIC：3'側モジュールの少なくとも1つの逆転写酵素認識配列と、GIC：3'側モジュールの少なくとも1つのrRNA配列とを含む。いくつかの実施形態において、少なくとも1つのGIC：3'側モジュールは、GIC：3'側モジュールの少なくとも1つの逆転写酵素認識配列と、GIC：3'側モジュールの少なくとも1つのAトラクト配列とを含む。いくつかの実施形態において、少なくとも1つのGIC：3'側モジュールは、GIC：3'側モジュールの少なくとも1つの逆転写酵素認識配列と、GIC：3'側モジュールの少なくとも1つのrRNA配列と、GIC：3'側モジュールの少なくとも1つのAトラクト配列とを含む。

20

【0220】

いくつかの実施形態において、少なくとも1つのGIC：3'側モジュールは、

30

(a) 配列番号200~221のいずれか1つから選択されるか、これらのいずれか1つをコードするか、これらのいずれか1つによってコードされる少なくとも1つの逆転写酵素認識配列；

(b) 配列番号280~289のいずれか1つから選択されるか、これらのいずれか1つをコードするか、これらのいずれか1つによってコードされる少なくとも1つのrRNA配列；および／または

(c) 少なくとも1つのAトラクト配列
をどのような組み合わせで含んでもよい。

【0221】

例示的なGIC：3'側モジュール

40

いくつかの実施形態において、少なくとも1つのGIC：3'側モジュールは、配列番号300~329の少なくとも1つを含んでもよく、配列番号300~329の少なくとも1つをコードしてもよく、配列番号300~329の少なくとも1つによってコードされてもよい。いくつかの実施形態において、少なくとも1つのGIC：3'側モジュールは、配列番号300~329からなる群から選択される少なくとも1つの配列と少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでもよく、この配列をコードしてもよく、この配列によってコードされてもよい。いくつかの実施形態において、少なくとも1つのGIC：3'側モジュールは、配列番号300~329からなる群から選択される配列またはこれらの任意の組み合わせと少なくとも

50

90%の同一性（例えば、少なくとも90%、91%、92%、93%、94%、95%、96%、97%、98%または99%の同一性）を有する配列を含む。いくつかの実施形態において、GIC：3'側モジュールは、配列番号300～329と少なくとも90%、91%、92%、93%、94%、95%、96%、97%、98%または99%の同一性を有する非天然配列を含む。

【0222】

いくつかの実施形態において、少なくとも1つのGIC：3'側モジュールは、配列番号313～320の少なくとも1つと少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

10

【0223】

いくつかの実施形態において、少なくとも1つのGIC：3'側モジュールは、「GACGG TAGC TAGGTTTCGCA AGGCAGCCAC AAGCCAAAGA TAGGTAGGGT GCTCATAGT G AGTAGGGACA GTGCCTTTTG ATTCACAACG CGTCAATACC ATCTGACACG GATACCTTA CCGGACTTGT CATGATCTCC CAGACTTGTC CAAGGTGGAC GGGCC ACCTT TACTTAACCC GGAAAAGGAA CATATATTAA TTATATGTGT TCGGAAAA」(配列番号222)、「CCGGACTTGT CATGATCTCC CAGACTTGTC CAAGGTGGAC GGGCCACCTT TACTTAACCC GGAAAAGGAA CATATATTAA TTATATGTGT TCGGAAAA」(配列番号223)、および「CAAGGTGGAC GGGCCACCTT TACTTAACCC GGAAAAGGAA CATATATTAA TTATATGTGT TCGGAAAA」(配列番号224)からなる群から選択される配列と少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよい。いくつかの実施形態において、このような配列として、TAGCaaaaaaaaaaaaaaaaaaaaa(配列番号538)で示される3'側配列が挙げられる。

20

【0224】

いくつかの実施形態において、少なくとも1つのGIC：3'側モジュールは、配列番号315と少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

30

【0225】

いくつかの実施形態において、少なくとも1つのGIC：3'側モジュールは、配列番号307と少なくとも95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

【0226】

いくつかの実施形態において、少なくとも1つのGIC：3'側モジュールは、配列番号316と少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

40

【0227】

GIC：ペイロードモジュール

本発明の遺伝子挿入コンストラクト（GIC）において使用されるGIC：ペイロードモジュールは、逆転写用の鋳型の一部として機能して、本明細書に開示された遺伝子挿入システムにより対象のゲノムに挿入される少なくとも1つのペイロード配列を含むか、このようなペイロード配列をコードする。本明細書において、「ペイロード配列」または単に「ペイロード」という用語は、本発明の少なくとも1つの遺伝子挿入システムにより標的ゲノムに挿入することを目的とした生体高分子配列を意味する。本発明のペイロード配列は

50

、少なくとも1つの導入遺伝子を含んでいてもよい。

【0228】

本明細書において、「導入遺伝子」という用語は、最も広い意味で使用され、本発明の遺伝子挿入システムによって対象のゲノムに挿入されるあらゆる遺伝子配列を意味する。例えば、導入遺伝子として、対象のゲノムにおいて通常は見られない配列、または対象のゲノムにおいて通常見られるが、標的挿入部位には通常は見られない配列が挙げられる。導入遺伝子として、所望の発現産物（例えば、少なくとも1つの、mRNA、マイクロRNA、siRNA、rRNA、tRNA、長鎖非コードRNA、細胞質低分子RNA、核内低分子RNA、核小体低分子RNA、カハール体低分子RNA、環状RNA、ペプチド、ポリペプチド、および／もしくはタンパク質）を含むか、これらをコードする配列、ならびに／または少なくとも1つの導入遺伝子の発現を制御する配列が挙げられるが、これらに限定されない。いくつかの実施形態において、導入遺伝子は、テロメラーゼ逆転写酵素（TERT、例えば、ヒトTERT）、フェニルアラニンヒドロキシラーゼ（PAH、例えば、ヒトPAH）、第VII因子（例えば、ヒト第VIII因子）、Bドメインの長さが様々に異なる変異型第VIII因子（例えば、hFactor VIII N6やhFactor VIII N6変異体）、および第IX因子（例えば、ヒト第IX因子）から選択されるタンパク質をコードする。いくつかの実施形態において、導入遺伝子は、制御RNAをコードする。いくつかの実施形態において、導入遺伝子は、別のタンパク質の阻害剤をコードする。いくつかの実施形態において、前記阻害剤は単鎖抗体である。いくつかの実施形態において、導入遺伝子は、以下の表Xに記載の遺伝子から選択される疾患の治療に使用することができるタンパク質をコードする。

【0229】

10

20

30

40

50

【表 1】

表 X. 代表的な導入遺伝子

疾患	座位	遺伝子名
全色盲(ACHM)	CNGB3	beta 3 subunit of a cyclic nucleotide-gated ion channel
全色盲(ACHM)	CNGA3	alpha 3 subunit of a cyclic nucleotide-gated ion channel
副腎白質ジストロフィー	ABCD1	ALDP タンパク質
眼皮膚白皮症 II 型	OCA2	眼皮膚白皮症 II (OCA2)
β サラセミア	HBB	ヘモグロビンサブユニット β
ブルガダ症候群	SCN5A	電位依存性ナトリウムチャネル α サブユニット 5
カナバン病	ASPA	アスパルトアシラーゼ
シャルコー・マリー・トゥース病	PMP22	末梢ミエリンタンパク質 22
全脈絡膜萎縮(CHM)	REP1	Rab エスコートタンパク質 1
慢性肉芽腫症(CGD)	CYBA	p22-phox (オキシダーゼ) : α サブユニット
CILD1, 内臓逆位あり、または内臓逆位なし(カルタゲナー症候群)	DNAI1	軸系ダイニン中間鎖 1
古典的エーラス・ダンロス症候群(cEDS)	COL5A1/2	V 型コラーゲン
鎖骨頭蓋異形成症(CCD)	RUNX2	RUNX ファミリー転写因子 2
先天性難聴(出生時)	GJB2	ギャップ結合タンパク質 β 2
クリグラール・ナジャー症候群 I 型	UGT1A1	ビリルビン・ウリジンニリン酸グルクロン酸トランスフェラーゼ
嚢胞性線維症	CFTR	CF 膜貫通コンダクタンス調節因子
家族性大腸腺腫症	APC	Wnt シグナル伝達経路の APC 調節因子
ファンコーニ貧血	FANCE	FA 相補性群 E
脆弱 X 症候群	FMR1	脆弱 X メッセンジャーリボ核タンパク質 1
ゴーシェ病 1 型	GBA	グルコシルセラミダーゼ β 1
ヘモクロマトーシス(鉄過剰症)	HFE	鉄恒常性調節因子
血友病 A	F8	第 VIII 凝固因子
ハンチントン病	HTT	ハンチンチン (<i>HTT</i>)
高コレステロール血症 B 型	APOB	アポリポタンパク質 B
低リン血症性くる病	PHEX	X 連鎖性リン酸調節エンドペプチダーゼホモログ
クニースト症候群	COL2A1	II 型コラーゲン α 1 鎖
レーバー先天性黒内障(LCA)	CEP290	中心体タンパク質 290kDa
レーバー先天性黒内障(LCA)	CRB1	crumbs family member 1, photoreceptor morphogenesis associated
レーバー先天性黒内障(LCA)	GUCY2D	グアニル酸シクラーゼ 2D, 膜組織(網膜特異的)
レーベル遺伝性視神経症(LHON)	ND4	NADH デヒドロゲナーゼ 4
レーベル遺伝性視神経症(LHON)	ND1	NADH デヒドロゲナーゼ 1
レッシュ・ナイハン症候群(LNS)	HPRT1	ヒポキサンチングアニンホスホリボシルトランスフェラーゼ

10

20

30

40

50

マルファン症候群	FBN1	フィブリリン 1
中鎖アシル CoA デヒドロゲナーゼ欠損症	ACADM	中鎖アシル CoA デヒドロゲナーゼ
ムコ多糖症(MPS)	IDUA	α -L-イヅロニダーゼ
筋ジストロフィー, ベッカー型	DMD	ジストロフィン
筋ジストロフィー, デュシェンヌ型	DMD	ジストロフィン
筋強直性ジストロフィー1 型	DMPK	筋強直性ジストロフィープロテインキナーゼ
筋強直性ジストロフィー2 型	CNBP	CCHC 型ジンクフィンガー核酸結合タンパク質
神経線維腫症 II 型	NF2	Moesin-Ezrin-Radixin 様(MERLIN)腫瘍抑制因子
神経線維腫症 1 型	NF1	ニューロフィブロミン 1 (NF1)
ニーマン・ピック病 A 型および B 型	SMPD1	スフィンゴミエリナーゼ
パーキンソン病	GBA	グルコシルセラミダーゼ β 1
フェニルケトン尿症 (PKU)	PAH	フェニルアラニンヒドロキシラーゼ(PAH)
多発性嚢胞腎 1 型および 2 型	PKD2	多発性嚢胞腎 2
呼吸窮迫症候群, サーファクタントタンパク質 B(SP-B)欠損症	SFTPC	肺関連サーファクタントタンパク質 C
網膜色素変性症による視野障害	EYS	Eyes Shut Homolog
レット症候群	MECP2	メチル CpG 結合タンパク質 2
ロドプシン関連常染色体優性網膜色素変性症(RHO-adRP)	PRPH2	ペリフェリン 2
ロドプシン関連常染色体優性網膜色素変性症(RHO-adRP)	PRPF31	pre-mRNA プロセシング因子 31
ロドプシン関連常染色体優性網膜色素変性症(RHO-adRP)	RHO	ロドプシン
鎌状赤血球貧血	HBB	ヘモグロビンサブユニット β
非閉塞性精子形成不全	USP9Y	ユビキチン特異的ペプチダーゼ 9Y
脊髄性筋萎縮症	SMN1	Survival Of Motor Neuron 1, Telomeric
シュタルガルト病	ABCA4	ATP 結合カセットサブファミリーA メンバー4
テイ・サックス病	HEXA	ヘキササミニダーゼ A
アッシャー症候群	MYO7A	ミオシン VIIA
卵黄様黄斑ジストロフィー(ベスト病)	BEST1	ベストロフィン 1
フォン・ヒッペルリンドウ病 (VHL)	VHL	フォン・ヒッペルリンドウ病ユビキチン化複合体
X 連鎖性網膜色素変性症(XLRP)	RPGR	網膜色素変性症 GTP アーゼ調節因子
X 連鎖性網膜色素変性症(XLRP)	RP2	網膜色素変性症 2
X 連鎖性網膜分離症(XLRS)	RS1	レチノスキシン
α 1-アンチトリプシン欠損症 (COPD、肺気腫、肝疾患)	SERPINA1	α 1-アンチトリプシン

10

20

30

40

【 0 2 3 0 】

ペイロードモジュールの構造

【 0 2 3 1 】

GIC：ペイロードモジュールは、少なくとも1つの（例えば、1つ、2つ、3つ、またはそれ以上の個数の）導入遺伝子配列を含んでもよく、さらに、前記導入遺伝子の少なくとも1つのプロモーター配列、前記導入遺伝子の少なくとも1つの5' 非翻訳配列、前記導入遺伝子の少なくとも1つの3' 非翻訳配列、前記導入遺伝子の少なくとも1つのポリアデニル化シグナル配列またはポリAテール配列、前記導入遺伝子の少なくとも1つの非コードRNA (ncRNA) プロセシング配列、およびこれらの任意の組み合わせを含ん

50

でいてもよい。

【0232】

図6を再度参照すると、上部の拡大図に例示的なペイロードモジュール430の構造が示されている。導入遺伝子の任意の構成要素であるプロモーター配列431が存在する場合、このプロモーター配列431は、対象の細胞において挿入された導入遺伝子の発現を制御してもよい少なくとも1つのプロモーターを含んでいてもよく、このようなプロモーターをコードしてもよい。導入遺伝子の任意の構成要素である5' UTR配列432は、挿入された導入遺伝子が発現された場合に、導入遺伝子のmRNAの5' UTRをコードする配列を含んでいてもよく、このような配列をコードしてもよい。ペイロードモジュールの導入遺伝子配列433は、本開示の遺伝子挿入システムにより逆転写されて挿入される少なくとも1つの導入遺伝子配列を含んでいてもよく、例えば、この配列は、目的の遺伝子のORFを含んでいてもよく、目的の遺伝子のORFをコードしてもよい。導入遺伝子の任意の構成要素である3' UTR配列434は、発現された導入遺伝子のmRNAの少なくとも1つの3' UTRを含んでいてもよく、この3' UTRをコードしてもよい。同様に、導入遺伝子の任意の構成要素であるポリアデニル化シグナル配列435は、発現された導入遺伝子のmRNAのポリアデニル化シグナルを含んでいてもよく、このポリアデニル化シグナルをコードしてもよい。さらに、導入遺伝子の任意の構成要素である非コードRNA (ncRNA) プロセシング配列436は、導入遺伝子により発現されたnrRNAの終止シグナルおよび／または3' プロセシングシグナルを含んでいてもよく、このようなシグナルをコードしてもよい。

10

【0233】

導入遺伝子のプロモーター配列とRNAP II 5' UTR配列

導入遺伝子のプロモーター配列が存在する場合、このプロモーター配列は、対象のゲノムにおいて導入遺伝子の発現を促進する手段を含む少なくとも1つのプロモーター配列を含んでいてもよく、このようなプロモーター配列をコードしてもよい。遺伝子および／または導入遺伝子の発現を促進するこのような手段は当技術分野において数多く知られており、このような手段として、目的の遺伝子の5' 側への公知のプロモーター配列の挿入が挙げられる。当業者であれば、プロモーター配列の種類は、導入遺伝子や、使用するその他の特定の因子に基づいて選択してもよく、したがって、本開示を実施するにあたり、好適なプロモーターであればどのようなものを使用してもよいことを理解できるであろう。

20

【0234】

本開示で使用される例示的なプロモーターは、構成的なプロモーターであってもよく、誘導型プロモーターであってもよい。いくつかの実施形態において、導入遺伝子のプロモーター配列は、RNAポリメラーゼI~III (RNAP I、RNAP IIまたはIII) の少なくとも1つのプロモーターを含んでいてもよく、このプロモーターをコードしてもよい。いくつかの実施形態において、プロモーターの代わりに、またはプロモーターに加えて、少なくとも1つの導入遺伝子においてプロモーターと同じ領域に、宿主細胞のrDNA RNAP Iにより転写された導入遺伝子のRNA転写産物の遊離を可能にする少なくとも1つのリボザイムまたはその他のモチーフを含んでいてもよく、このようなリボザイムまたはその他のモチーフをコードしてもよい。

30

【0235】

いくつかの実施形態において、導入遺伝子の少なくとも1つのプロモーター配列は、少なくとも1つのヒトU1 snRNAプロモーターを含むか、このプロモーターをコードする。いくつかの実施形態において、導入遺伝子の少なくとも1つのプロモーター配列は、少なくとも1つのヒトU3 snRNAプロモーターを含むか、このプロモーターをコードする。いくつかの実施形態において、導入遺伝子の少なくとも1つのプロモーター配列は、少なくとも1つのヒトU6 snRNAプロモーターを含むか、このプロモーターをコードする。いくつかの実施形態において、導入遺伝子の少なくとも1つのプロモーター配列は、少なくとも1つのtRNAプロモーターを含むか、このプロモーターをコードする。

40

【0236】

導入遺伝子の5' UTR配列が存在する場合、この5' UTR配列は、挿入された導入遺伝

50

子の少なくとも1つのmRNA 5' UTRを含むか、これをコードする。通常、この5' UTR配列は、挿入された導入遺伝子が細胞によって発現された場合に、細胞内のリボソームによってアミノ酸生体高分子に翻訳されない配列を含むか、このような配列をコードする。このような配列として、例えば、天然の状態で導入遺伝子に関連する5' UTR、導入遺伝子に非天然の5' UTR（レトロエレメントの5' 側配列に由来する配列を含む）、公知の野生型遺伝子に関連していなくてもよい「合成」5' UTR、およびこれらの任意の組み合わせが挙げられる。

【0237】

当業者であれば、導入遺伝子の5' UTR配列の選択は、その導入遺伝子のアイデンティティや使用されるその他の特定の因子に左右されることから、あらゆる公知の5' UTR配列や、新たに見出された5' UTR配列でも、導入遺伝子のペイロードモジュールの5' 配列における使用に適している可能性があることを理解できるであろう。

10

【0238】

いくつかの実施形態において、導入遺伝子の少なくとも1つのプロモーター配列は、配列番号400~404および408~409の少なくとも1つを含んでいてもよく、配列番号400~404および408~409の少なくとも1つをコードしてもよく、配列番号400~404および408~409の少なくとも1つによってコードされてもよい。いくつかの実施形態において、導入遺伝子の少なくとも1つのプロモーター配列は、配列番号400~404および408~409の少なくとも1つと少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

20

【0239】

いくつかの実施形態において、導入遺伝子の少なくとも1つのプロモーター配列は、配列番号400と少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

【0240】

いくつかの実施形態において、導入遺伝子の少なくとも1つのプロモーター配列は、配列番号401と少なくとも95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

30

【0241】

いくつかの実施形態において、導入遺伝子の少なくとも1つのプロモーター配列は、配列番号403と少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

40

【0242】

いくつかの実施形態において、導入遺伝子の少なくとも1つのプロモーター配列は、配列番号404と少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含む。

【0243】

いくつかの実施形態において、導入遺伝子の少なくとも1つのプロモーター配列は、配列番号408と少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含む。

50

【0244】

いくつかの実施形態において、導入遺伝子の少なくとも1つのプロモーター配列は、配列番号409と少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含む。

【0245】

いくつかの実施形態において、GIC：ペイロードモジュールは、導入遺伝子のプロモーター配列の5'側に配置されたRNAポリメラーゼ（RNAP）ターミネーター配列を含む。いくつかの実施形態において、RNAPはRNAP I（Pol I）であり、GICペイロードモジュールがリボソームDNA遺伝子の標的部位に組み込まれた際に、この終止配列によってPol Iのリードスルー転写が阻止される。いくつかの実施形態において、RNAPターミネーター配列は、5'-AGGTCGACCGAGATGTCCGAGGTCGACCGAGTTGTCCG-3'（配列番号537）で示される配列を含む。

10

【0246】

導入遺伝子配列

ペイロードモジュールの導入遺伝子配列は、対象のゲノムへの挿入用の少なくとも1つの目的の配列を含むか、この目的の配列をコードする。本明細書において、「目的の配列」は、少なくとも1つの所望の発現産物を含むか、これをコードする生体高分子配列を意味する。いくつかの実施形態において、導入遺伝子は、hTERT、hPAH、hFactor VIII、Bドメインの長さが様々に異なる変異型ヒト第VIII因子（例えば、hFactor VIII N6やhFactor VIII N6変異体）、および第IX因子（例えば、ヒト第IX因子）から選択されるタンパク質をコードする。いくつかの実施形態において、導入遺伝子は、制御RNAをコードする。いくつかの実施形態において、導入遺伝子は、別のタンパク質の阻害剤をコードする。いくつかの実施形態において、前記阻害剤は単鎖抗体である。いくつかの実施形態において、導入遺伝子は、上記の表Xに記載の遺伝子から選択される疾患の治療に使用することができるタンパク質をコードする。

20

【0247】

どのような目的の配列であっても、本開示の実施に好適であってもよく、配列の由来源に限定されず（すなわち、由来源の生物種や、天然配列あるいは人工配列であるのかにかかわらず）、配列の長さにも限定されない。

30

【0248】

いくつかの実施形態において、少なくとも1つの導入遺伝子配列は、配列番号411~422の少なくとも1つを含んでいてもよく、配列番号411~422の少なくとも1つをコードしてもよく、配列番号411~422の少なくとも1つによってコードされてもよい。いくつかの実施形態において、少なくとも1つの導入遺伝子配列は、配列番号411~422の少なくとも1つと少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

【0249】

いくつかの実施形態において、少なくとも1つの導入遺伝子配列は、配列番号419または420と少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

40

【0250】

いくつかの実施形態において、少なくとも1つの導入遺伝子配列は、配列番号421~422の少なくとも1つと少なくとも95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされ

50

てもよい。

【0251】

いくつかの実施形態において、少なくとも1つの導入遺伝子配列は、配列番号518～536の少なくとも1つと少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

【0252】

導入遺伝子の3' UTR配列とポリアデニル化シグナル

導入遺伝子の3' UTR配列が存在する場合、この3' UTR配列は、挿入された導入遺伝子の少なくとも1つのmRNA 3' UTRを含むか、これをコードする。通常、この3' UTR配列は、挿入された導入遺伝子が細胞によって発現された場合に、細胞内のリボソームによってアミノ酸生体高分子に翻訳されない配列を含むか、このような配列をコードする。このような配列として、例えば、天然の状態で導入遺伝子に関連する3' UTR、導入遺伝子に非天然の3' UTR（レトロエレメントの3' 側配列に由来する配列を含む）、公知の野生型遺伝子に関連していない「合成」3' UTR、およびこれらの任意の組み合わせが挙げられる。

【0253】

当業者であれば、導入遺伝子の3' UTR配列の選択は、その導入遺伝子のアイデンティティや使用されるその他の特定の因子に左右されることから、あらゆる公知の3' UTR配列や、新たに見出された3' UTR配列でも、導入遺伝子のペイロードモジュールの3' 配列における使用に適している可能性があることを理解できるであろう。

【0254】

導入遺伝子にポリアデニル化シグナル配列が存在する場合、このポリアデニル化シグナル配列は、導入遺伝子のmRNAの少なくとも1つのポリアデニル化シグナルを含むか、これをコードする。公知または新規なあらゆる好適なポリアデニル化シグナルを、本開示の鑄型モジュールにおいて使用してもよい。明確化のために付言すると、挿入された導入遺伝子に存在するか、挿入された導入遺伝子内にコードされている少なくとも1つのポリアデニル化シグナルは、導入遺伝子のmRNA発現産物またはncRNA発現産物にポリAテールを付加するRNAP IIを提供するものである。

【0255】

いくつかの実施形態において、導入遺伝子の少なくとも1つの3' UTR配列は、配列番号405～407の少なくとも1つから選択される配列を含んでいてもよい。いくつかの実施形態において、導入遺伝子の少なくとも1つの3' UTR配列は、配列番号405～407の少なくとも1つと少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよい。

【0256】

いくつかの実施形態において、導入遺伝子の少なくとも1つの3' UTR配列は、配列番号405と少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

【0257】

いくつかの実施形態において、導入遺伝子の少なくとも1つの3' UTR配列は、配列番号406と少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

【0258】

10

20

30

40

50

いくつかの実施形態において、導入遺伝子の少なくとも1つの3' UTR配列は、配列番号407と少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

【0259】

導入遺伝子の非コードRNA (ncRNA) プロセッシング配列

導入遺伝子のncRNAプロセッシング配列が存在する場合、このncRNAプロセッシング配列は、導入遺伝子から発現されたncRNA（例えば、トランスファーRNA (tRNA)、rRNA、マイクロRNA、siRNA、snRNAなど）の発現またはプロセッシングを制御する配列を含むか、このような配列をコードする。いくつかの実施形態において、少なくとも1つの非コードRNA (ncRNA) プロセッシング配列は、導入遺伝子から発現された少なくとも1つのncRNAの少なくとも1つの終止シグナル、少なくとも1つの3' プロセッシングシグナル、およびこれらの任意の組み合わせを含むか、これらをコードする。

【0260】

いくつかの実施形態において、導入遺伝子の少なくとも1つのncRNAプロセッシング配列は、少なくとも1つのMALAT1 3' 末端プロセッシングシグナルおよび／またはMALAT1 3' 末端保護シグナルを含むか、これらをコードする。いくつかの実施形態において、導入遺伝子の少なくとも1つのncRNAプロセッシング配列は、少なくとも1つのRNA三重鎖形成末端保護構造を含むか、これらをコードする。いくつかの実施形態において、導入遺伝子の少なくとも1つのncRNAプロセッシング配列は、少なくとも1つのエンドヌクレアーゼ動員構造、エンドヌクレアーゼ動員部位またはエンドヌクレアーゼ動員モチーフを含むか、これらをコードする。いくつかの実施形態において、導入遺伝子の少なくとも1つのncRNAプロセッシング配列は、少なくとも1つのポリチミジントラクトを含むか、これをコードする。いくつかの実施形態において、導入遺伝子の少なくとも1つのRNA 3' 末端終結配列および／またはRNA 3' 末端プロセッシング配列は、RNAP IのSalI終端ボックスを含む。

【0261】

ペイロードモジュールのモジュール性

本明細書において開示したGIC：ペイロードモジュールの各構成要素は、コンビナトリアルな方法で互いに交換可能に使用して、特定の遺伝子挿入システムに必要な機能性または所望の機能性を有する3' 側モジュールを設計してもよい。

【0262】

いくつかの実施形態において、少なくとも1つのGIC：ペイロードモジュールは、少なくとも1つの導入遺伝子配列を含んでいてもよく、これをコードしてもよい。いくつかの実施形態において、少なくとも1つのGIC：ペイロードモジュールは、導入遺伝子の少なくとも1つのプロモーター配列を含んでいてもよく、これをコードしてもよい。いくつかの実施形態において、少なくとも1つのGIC：ペイロードモジュールは、導入遺伝子の少なくとも1つの5' UTR配列を含んでいてもよく、これをコードしてもよい。いくつかの実施形態において、少なくとも1つのGIC：ペイロードモジュールは、導入遺伝子の少なくとも1つの3' UTR配列を含んでいてもよく、これをコードしてもよい。いくつかの実施形態において、少なくとも1つのGIC：ペイロードモジュールは、導入遺伝子の少なくとも1つのポリアデニル化シグナル配列を含んでいてもよく、これをコードしてもよい。いくつかの実施形態において、少なくとも1つのGIC：ペイロードモジュールは、導入遺伝子の少なくとも1つのncRNAプロセッシング配列を含んでいてもよく、これをコードしてもよい。

【0263】

いくつかの実施形態において、少なくとも1つのGIC：ペイロードモジュールは、少なくとも1つの導入遺伝子配列、前記導入遺伝子の少なくとも1つのプロモーター配列、前記導入遺伝子の少なくとも1つの5' UTR配列、前記導入遺伝子の少なくとも1つの3' U

10

20

30

40

50

TR配列、前記導入遺伝子の少なくとも1つのポリアデニル化シグナル配列、および／または少なくとも1つのncRNAプロセッシング配列を含んでいてもよく、これらをコードしてもよい。

【0264】

いくつかの実施形態において、少なくとも1つのGIC：ペイロードモジュールは、

(a) 配列番号400～404のいずれか1つ、および配列番号400～404のいずれか1つと少なくとも90%の同一性（例えば、少なくとも90%、91%、92%、93%、94%、95%、96%、97%、98%、99%または100%の同一性）を有する配列から選択される、導入遺伝子の少なくとも1つのプロモーター配列および5' UTR配列；

(b) 配列番号411～422および配列番号499～536のいずれか1つ、ならびに配列番号411～422または499～536と少なくとも90%の同一性（例えば、少なくとも90%、91%、92%、93%、94%、95%、96%、97%、98%、99%もしくは100%の同一性）を有する配列から選択されるか、これらの配列から選択される配列をコードするか、これらの配列から選択される配列によってコードされる、少なくとも1つの導入遺伝子配列；ならびに

(c) 配列番号405～407、および配列番号405～407と少なくとも90%の同一性（例えば、少なくとも90%、91%、92%、93%、94%、95%、96%、97%、98%、99%または100%の同一性）を有する配列から選択される、導入遺伝子の少なくとも1つの3' UTR配列およびポリアデニル化シグナルをどのような組み合わせで含んでいてもよい。

【0265】

例示的なGIC：ペイロードモジュール

いくつかの実施形態において、少なくとも1つのGIC：ペイロードモジュールは、配列番号499～536から選択される少なくとも1つの配列を含んでいてもよく、配列番号499～536から選択される少なくとも1つの配列をコードしてもよく、配列番号499～536から選択される少なくとも1つの配列によってコードされてもよい。いくつかの実施形態において、少なくとも1つのGIC：ペイロードモジュールは、配列番号499～536から選択される少なくとも1つの配列と少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

【0266】

いくつかの実施形態において、少なくとも1つのGIC：ペイロードモジュールは、配列番号419、420、518または519と少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

【0267】

いくつかの実施形態において、少なくとも1つのGIC：ペイロードモジュールは、配列番号421、422、520または521と少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

【0268】

いくつかの実施形態において、少なくとも1つのGIC：ペイロードモジュールは、配列番号522、523、524または525と少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでいてもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

【0269】

遺伝子挿入コンストラクト（GIC）のモジュール性

本明細書において開示した遺伝子挿入コンストラクト（GIC）の各構成要素（すなわち、GIC：5'側モジュール、GIC：3'側モジュールおよびGIC：ペイロードモジュール）は、コンビナトリアルな方法で互いに交換可能に使用して、特定の遺伝子挿入システムに必要な機能性または所望の機能性を有するGICを設計してもよい。

【0270】

いくつかの実施形態において、少なくとも1つの遺伝子挿入コンストラクト（GIC）は、少なくとも1つのGIC：5'側モジュールを含む。いくつかの実施形態において、少なくとも1つの遺伝子挿入コンストラクトは、少なくとも1つのGIC：ペイロードモジュールを含む。いくつかの実施形態において、少なくとも1つの遺伝子挿入コンストラクトは、少なくとも1つのGIC：3'側モジュールを含む。いくつかの実施形態において、少なくとも1つの遺伝子挿入コンストラクトは、少なくとも1つのGIC：5'側モジュールと、少なくとも1つのGIC：ペイロードモジュールを含む。いくつかの実施形態において、少なくとも1つの遺伝子挿入コンストラクトは、少なくとも1つのGIC：5'側モジュールと、少なくとも1つのGIC：3'側モジュールを含む。いくつかの実施形態において、少なくとも1つの遺伝子挿入コンストラクトは、少なくとも1つのGIC：5'側モジュールと、少なくとも1つのGIC：ペイロードモジュールと、少なくとも1つのGIC：3'側モジュールを含む。

10

【0271】

いくつかの実施形態において、少なくとも1つの遺伝子挿入コンストラクトは、GIC：3'側モジュールの逆転写酵素認識配列の由来元と同じ生物種のレトロエレメントに由来するGIC：5'側モジュールレトロエレメント配列を含む少なくとも1つのGIC：5'側モジュールを含む。いくつかの実施形態において、少なくとも1つの遺伝子挿入コンストラクトは、GIC：3'側モジュールの逆転写酵素認識配列の由来元とは異なる生物種のレトロエレメントに由来するGIC：5'側モジュールレトロエレメント配列を含む少なくとも1つのGIC：5'側モジュールを含む。いくつかの実施形態において、少なくとも1つの遺伝子挿入コンストラクトは、GIC：3'側モジュールの逆転写酵素認識配列を含む少なくとも1つの遺伝子挿入コンストラクトに概して有用な、真核生物の生態に天然には見られないGIC：5'側モジュール配列を含む少なくとも1つのGIC：5'側モジュールを含む。

20

【0272】

いくつかの実施形態において、遺伝子挿入コンストラクトは、図7に示した供給源に由来するGIC：5'側モジュール配列と、GIC：3'側モジュール配列の組み合わせを含む。図7において、A1はノドジロシトド（*Zonotrichia albicollis*）であり、A2はキンカチヨウ（*Taeniopygia guttata*）であり、A3はノドジロシギダチヨウ（*Tinamus guttatus*）であり、A4はガラパゴスフィンチ（*Geospiza fortis*）であり、B1はカタノトミヨ（*Pungitis pungitis*）であり、B2はメダカ（*Oryzias latipes*）であり、B3はイトヨ（*Gasterosteus aculeatus*）であり、C1はキョウソヤドリコバチ（*Nasonia vitripennis*）であり、C2はキイロショウジョウバエ（*Drosophila melanogaster*）であり、C3はコクヌストモドキ（*Tribolium castaneum*）であり、C4はカイコガ（*Bombyx mori*）であり、C5はオナジショウジョウバエ（*Drosophila simulans*）であり、C6は*Drosophila mercatorum*であり、D1は*Lepidurus couseii*であり、D2はヨーロッパカブトエビ（*Triops cancriformis*）であり、E1はチクビヒドラ（*Hydra magnipapillata*）であり、E2はアメリカカブトガニ（*Limulus polyphemus*）であり、E3は*Adineta vaga*であり、E4はカタユウレイボヤ（*Ciona intestinalis*）である。

30

40

【0273】

いくつかの実施形態において、少なくとも1つの遺伝子挿入コンストラクトは、
 (a) 少なくとも1つのGIC：5'側モジュールであって、配列番号250~276、配列番号250~276と比較して1つ、2つもしくは3つのヌクレオチド変化もしくはヌクレオチド置換を有する配列、配列番号60~154、配列番号60~154と少なくとも90%の同

50

一性（例えば、少なくとも90%、91%、92%、93%、94%、95%、96%、97%、98%、99%もしくは100%の同一性）を有する配列、配列番号278～279、および配列番号278～279と少なくとも90%の同一性（例えば、少なくとも90%、91%、92%、93%、94%、95%、96%、97%、98%、99%もしくは100%の同一性）を有する配列から選択されるか、これらの配列から選択される配列をコードするか、これらの配列から選択される配列によってコードされる、少なくとも1つのGIC：5'側モジュール；

(b) 少なくとも1つのGIC：ペイロードモジュールであって、配列番号411～422および499～525のいずれか1つ、ならびに配列番号411～422もしくは499～522と少なくとも90%の同一性（例えば、少なくとも90%、91%、92%、93%、94%、95%、96%、97%、98%、99%もしくは100%の同一性）を有する配列から選択されるか、これらの配列から選択される配列をコードするか、これらの配列から選択される配列によってコードされる、少なくとも1つのGIC：ペイロードモジュール；ならびに／または

(c) 少なくとも1つのGIC：3'側モジュールであって、配列番号300～329のいずれか1つ、および配列番号300～329と少なくとも90%の同一性（例えば、少なくとも90%、91%、92%、93%、94%、95%、96%、97%、98%、99%もしくは100%の同一性）を有する配列から選択されるか、これらの配列から選択される配列をコードするか、これらの配列から選択される配列によってコードされる、少なくとも1つのGIC：3'側モジュール

をどのような組み合わせで含んでもよく、このような組み合わせをコードしてもよく、このような組み合わせによってコードされてもよい。

【0274】

例示的な遺伝子挿入コンストラクト（GIC）

いくつかの実施形態において、少なくとも1つの遺伝子挿入コンストラクト（GIC）は、配列番号411～422および499～525の少なくとも1つを含んでもよく、配列番号411～422および499～525の少なくとも1つをコードしてもよく、配列番号411～422および499～525の少なくとも1つによってコードされてもよい。いくつかの実施形態において、少なくとも1つの遺伝子挿入コンストラクトは、配列番号411～422および499～536の少なくとも1つと少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

【0275】

いくつかの実施形態において、少なくとも1つの遺伝子挿入コンストラクトは、配列番号419、420、518または519と少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

【0276】

いくつかの実施形態において、少なくとも1つの遺伝子挿入コンストラクトは、配列番号421、422、520または521と少なくとも100%、99%、98%、97%、96%、95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

【0277】

いくつかの実施形態において、少なくとも1つの遺伝子挿入コンストラクトは、配列番号522、523、524または525と少なくとも95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、10%または5%の相同性を有する配列を含んでもよく、この配列をコードしてもよく、この配列によってコードされてもよい。

【0278】

遺伝子挿入システムの設計とモジュール性

本明細書において開示した遺伝子挿入システム（GIS）の各構成要素（すなわち、逆転写酵素コンストラクト（RTC）と遺伝子挿入コンストラクト（GIC））は、コンビナトリアルな方法で互いに交換可能に使用して、必要な機能性または所望の機能性を有する遺伝子挿入システム（GIS）を設計してもよい。

【0279】

いくつかの実施形態において、少なくとも1つの遺伝子挿入システムは、少なくとも1つの逆転写酵素コンストラクト（RTC）を含んでいてもよい。いくつかの実施形態において、少なくとも1つの遺伝子挿入システムは、少なくとも1つの遺伝子挿入コンストラクト（GIC）を含んでいてもよい。いくつかの実施形態において、少なくとも1つの遺伝子挿入システムは、少なくとも1つの逆転写酵素コンストラクト（RTC）と、少なくとも1つの遺伝子挿入コンストラクト（GIC）を含んでいてもよい。

10

【0280】

遺伝子挿入システムを含む生体高分子の組成

本明細書に開示した遺伝子挿入システム（GIS）の各構成要素を含む生体高分子の組成は、本明細書に開示されたものの中からコンビナトリアルな方法で選択して、必要な機能性または所望の機能性を有する遺伝子挿入システム（GIS）を設計してもよい。

【0281】

いくつかの実施形態において、少なくとも1つの逆転写酵素コンストラクト（RTC）は、RNA生体高分子として少なくとも1つの対象に導入してもよい。いくつかの実施形態において、少なくとも1つの逆転写酵素コンストラクト（RTC）は、mRNA生体高分子として少なくとも1つの対象に導入してもよい。

20

【0282】

いくつかの実施形態において、少なくとも1つの遺伝子挿入コンストラクト（GIC）は、RNA生体高分子として少なくとも1つの対象に導入してもよい。いくつかの実施形態において、少なくとも1つの遺伝子挿入コンストラクト（GIC）は、直鎖RNA生体高分子として少なくとも1つの対象に導入してもよい。

【0283】

いくつかの実施形態において、少なくとも1つの逆転写酵素コンストラクト（RTC）は、RNA生体高分子として少なくとも1つの対象に導入してもよく、少なくとも1つの遺伝子挿入コンストラクト（GIC）は、RNA生体高分子として少なくとも1つの対象に導入してもよい。

30

【0284】

いくつかの実施形態において、少なくとも1つの逆転写酵素コンストラクト（RTC）は、mRNA生体高分子として少なくとも1つの対象に導入してもよく、少なくとも1つの遺伝子挿入コンストラクト（GIC）は、RNA生体高分子として少なくとも1つの対象に導入してもよい。

【0285】

いくつかの実施形態において、少なくとも1つの逆転写酵素コンストラクト（RTC）および／または少なくとも1つの遺伝子挿入コンストラクト（GIC）は、DNA生体高分子として少なくとも1つの対象に導入してもよい。いくつかの実施形態において、少なくとも1つの逆転写酵素コンストラクト（RTC）および／または少なくとも1つの遺伝子挿入コンストラクト（GIC）は、プラスミドとして少なくとも1つの対象に導入してもよい。

40

【0286】

いくつかの実施形態において、少なくとも1つの逆転写酵素コンストラクト（RTC）は、アミノ酸生体高分子として少なくとも1つの対象に導入してもよい。いくつかの実施形態において、少なくとも1つの逆転写酵素コンストラクト（RTC）は、タンパク質として少なくとも1つの対象に導入してもよい。

【0287】

50

いくつかの実施形態において、少なくとも1つの逆転写酵素コンストラクト（RTC）は、アミノ酸生体高分子として少なくとも1つの対象に導入してもよく、少なくとも1つの遺伝子挿入コンストラクト（GIC）は、RNA生体高分子として少なくとも1つの対象に導入してもよい。いくつかの実施形態において、少なくとも1つの逆転写酵素コンストラクト（RTC）は、プラスミドとして少なくとも1つの対象に導入してもよく、少なくとも1つの遺伝子挿入コンストラクト（GIC）は、RNA生体高分子として少なくとも1つの対象に導入してもよい。

【0288】

いくつかの実施形態において、少なくとも1つの逆転写酵素コンストラクト（RTC）は、プラスミドとして少なくとも1つの対象に導入してもよく、少なくとも1つの遺伝子挿入コンストラクト（GIC）は、プラスミドとして少なくとも1つの対象に導入してもよい。いくつかの実施形態において、少なくとも1つの逆転写酵素コンストラクト（RTC）は、RNA（例えば、mRNA）として少なくとも1つの対象に導入してもよく、少なくとも1つの遺伝子挿入コンストラクト（GIC）は、プラスミドとして少なくとも1つの対象に導入してもよい。

【0289】

ペアになる逆転写酵素

本発明の遺伝子挿入システムは、この遺伝子挿入システムに含まれる遺伝子挿入コンストラクト（GIC）および逆転写酵素コンストラクト（RTC）ならびにその両方のうちの少なくとも1つの組成を設計または選択することによって、GICとRTCの間の相互作用を制御することにより、所望の機能に対して最適化してもよい。例えば、GICおよび／またはRTCの組成を改変することによって、PCR、シーケンシング、および／またはペイロードである導入遺伝子の発現を利用して挿入を検出することにより観察されるように、ペイロード全長の挿入の効率、速度および／または忠実度を変化させることができ；シーケンシング、ハイブリダイゼーション、または挿入されたDNAの遺伝子上での位置を可視化するその他の方法により観察されるように、ペイロードが挿入される標的部位の選択における配列特異性および／または染色体位置を変化させることができ；RTCが逆転写用の鋳型として投与されたGICのみを利用する選択性を変化させることなどができる。本明細書において、「ペアになる逆転写酵素」という用語は、特定のGIC配列と組み合わせて投与される特定のRTC：RTモジュール配列を意味する。

【0290】

何らかの理論に拘束されることを望むものではないが、RTCおよびGICの相互作用の改変は、RTC：RTモジュールと、GIC：5'側モジュールおよび／またはGIC：3'側モジュールとを選択することによって達成してもよい。例えば、GICに対するRTCの特異性は、同じ生物種または異なる生物種のレトロエレメントに由来する構成要素を選択することによって改変してもよい。本明細書において、遺伝子挿入システムの2種の構成要素が同じ生物種のレトロエレメントに由来する場合、同種であるという。これに対して、遺伝子挿入システムの2種の構成要素が異なる生物種のレトロエレメントに由来する場合、異種であるという。

【0291】

いくつかの実施形態において、少なくとも1つのRTC：RTモジュールは、GIC：5'側モジュール配列および／またはGIC：3'側モジュール配列が由来する少なくとも1つのレトロエレメントとは異なる生物種のレトロエレメントに由来する少なくとも1つの配列を含むか、このような配列をコードする（本明細書において、「異種ペアの逆転写酵素」と呼ぶ）。

【0292】

いくつかの実施形態において、RTCおよびGICに含まれるレトロエレメント由来の配列はいずれも同じ生物種のレトロエレメントに由来する（本明細書において、「同種ペアの逆転写酵素」と呼ぶ）。

【0293】

10

20

30

40

50

いくつかの実施形態において、同種ペアの逆転写酵素と比較して、異種ペアの逆転写酵素は特異性が増加していてもよい。

【0294】

本明細書において、「特異性」という用語は、ペアになる逆転写酵素が、導入遺伝子の挿入に意図する鋳型RNAを効率的かつ／または選択的に利用する可能性を意味する。

【0295】

いくつかの実施形態において、少なくとも1つの遺伝子挿入システムは、GICと、これとペアになる逆転写酵素の少なくとも1つの組み合わせを含み、このような組み合わせを図7に示している。

【0296】

例示的な遺伝子挿入システム

いくつかの実施形態において、少なくとも1つの遺伝子挿入システムは、

(a) 少なくとも1つの逆転写酵素コンストラクト(RTC)であって、配列番号1~59のいずれか1つ、および配列番号1~59のいずれか1つと少なくとも90%の同一性(例えば、少なくとも90%、91%、92%、93%、94%、95%、96%、97%、98%、99%もしくは100%の同一性)を有する配列から選択されるか、これらの配列から選択される配列をコードするか、これらの配列から選択される配列によってコードされる、少なくとも1つの逆転写酵素コンストラクト(RTC)；ならびに

(b) 少なくとも1つの遺伝子挿入コンストラクト(GIC)であって、配列番号250~276、配列番号250~276と比較して1つ、2つもしくは3つのヌクレオチド変化もしくはヌクレオチド置換を有する配列、配列番号60~154、配列番号60~154と少なくとも90%の同一性(例えば、少なくとも90%、91%、92%、93%、94%、95%、96%、97%、98%、99%もしくは100%の同一性)を有する配列、配列番号278~279、配列番号278~279と少なくとも90%の同一性(例えば、少なくとも90%、91%、92%、93%、94%、95%、96%、97%、98%、99%もしくは100%の同一性)を有する配列、配列番号411~422もしくは499~536、配列番号411~422もしくは499~536と少なくとも90%の同一性(例えば、少なくとも90%、91%、92%、93%、94%、95%、96%、97%、98%、99%もしくは100%の同一性)を有する配列、配列番号300~329、および／または配列番号300~329と少なくとも90%の同一性(例えば、少なくとも90%、91%、92%、93%、94%、95%、96%、97%、98%、99%もしくは100%の同一性)を有する配列のいずれか1つを含む配列から選択されるか、この配列をコードするか、この配列によってコードされる、少なくとも1つの遺伝子挿入コンストラクト(GIC)をどのような組み合わせで含んでもよく、これらをどのような組み合わせでコードしてもよく、これらのどのような組み合わせによりコードされてもよい。

【0297】

III. 製剤および送達機構

核酸

いくつかの実施形態において、逆転写酵素コンストラクト(RTC)または遺伝子挿入コンストラクト(GIC)は、1つ以上の修飾ヌクレオチドを含んでもよく、修飾ヌクレオチドとして、核酸塩基修飾、糖類修飾ヌクレオチド、および／または主鎖修飾が挙げられるが、これらに限定されない。いくつかの実施形態において、RTCコンストラクトまたはGICコンストラクトは、修飾の組み合わせを含んでもよく、例えば、核酸塩基修飾と主鎖修飾の組み合わせを含んでもよい。

【0298】

いくつかの実施形態において、修飾ヌクレオチドは、核酸塩基が修飾されたヌクレオチドであってもよい。「修飾塩基」は、1つ以上の基または原子が置換または付加されることによって修飾された、アデニン、シトシン、チミン、グアニン、ウラシル、キサンチン、イノシン、キューオシンなどのヌクレオチド塩基を意味するが、これらに限定されない。いくつかの実施形態において、修飾ヌクレオチドは、主鎖が修飾されたヌクレオチドであってもよい。

10

20

30

40

50

【0299】

RTCコンストラクトおよび／またはGICコンストラクトは、参照配列、特に、目的の配列と比較して1つ以上の置換、挿入および／または付加、欠失、ならびに共有結合修飾を含んでもよく、このようなRTCコンストラクトおよび／またはGICコンストラクトも、本発明の範囲内に含まれる。

【0300】

いくつかの実施形態において、RTCコンストラクトおよび／またはGICコンストラクトは、1つ以上の転写後修飾（例えば、キャッピング、切断、ポリアデニル化、スプライシング、ポリA配列、メチル化、アシル化、リン酸化、リジン残基およびアルギニン残基のメチル化、アセチル化、チオール基およびチロシン残基のニトロシル化など）を含む。

10

【0301】

RTCコンストラクトおよび／またはGICコンストラクトは、糖類、核酸塩基またはヌクレオシド間結合（例えば、リン酸結合、リン酸ジエステル結合、またはホスホジエステル主鎖）などに有用な修飾を含んでもよい。

【0302】

いくつかの実施形態において、前記修飾は、化学誘導性の修飾または細胞誘導性の修飾を含んでもよい。例えば、細胞内RNA修飾のいくつかの例は、Nat Reviews Mol Cell Biol, 2017, 18:202-210に掲載されている「RNA modifications and structures cooperate to guide RNA-protein interactions」においてLewisおよびPanにより報告されているものがあるが、これらに限定されない。

20

【0303】

いくつかの実施形態において、RNAへの化学的修飾によって免疫回避を増強してもよい。RNAは、当技術分野において確立された方法により合成および／または修飾してもよい。

【0304】

いくつかの実施形態において、少なくとも1つのRNAコンストラクトは、少なくとも1つの修飾ウラシルを含んでもよい。ウラシル修飾の例として、5-メチルウリジン、5-メトキシウリジン、シュードウリジン、N¹-メチルシュードウリジン、および／または2-チオウリジンが挙げられる。いくつかの実施形態において、少なくとも1つのRNAコンストラクトは、少なくとも1つの修飾アデノシンを含んでもよい。アデノシン修飾の例として、2,6-ジアミノプリンデオキシリボヌクレオチドが挙げられる。

30

【0305】

いくつかの実施形態において、1つ以上のRNAにおける糖類修飾（例えば、2'位または4'位）もしくは糖類置換または主鎖修飾は、リン酸ジエステル結合の修飾または置換を含む。

【0306】

送達機構

本発明の遺伝子挿入システム（GIS）は、当技術分野で公知のどのような送達機構を介して対象に導入してもよい。本明細書において、「送達機構」は、遺伝子挿入システム、遺伝子挿入システムの構成要素、または遺伝子挿入システムの産物を対象に導入するために使用される方法または組成物を意味する。送達機構の例として、送達媒体、直接トランスフェクション（トランスフェクション試薬などを使用したもの）、遺伝子挿入システムをあらかじめトランスフェクトした細胞の移植、およびこれらの任意の組み合わせが挙げられるが、これらに限定されない。

40

【0307】

送達媒体

いくつかの実施形態において、本発明の遺伝子挿入システムは、送達媒体で製剤化してもよい。送達媒体は、一般に、細胞外環境での分解からの遺伝子挿入システムの各構成要素の保護、対象の細胞による取り込みの促進、エンドソーム脱出の促進、およびこれらの任意の組み合わせを介して、インビボまたはインビトロにおける対象の細胞のトランスフ

50

エクシオンを容易にしてもよい。送達媒体としては、脂質系ナノ粒子（例えば、脂質ナノ粒子（LNP）、リポソームおよびミセル）ならびに非脂質系ナノ粒子（例えば、ウイルス様粒子（VLP）およびポリマー送達粒子）などのナノ粒子が挙げられるが、これらに限定されない。

【0308】

ナノ粒子

いくつかの実施形態において、前記送達媒体は、少なくとも1種のナノ粒子を含んでいてもよい。本明細書において、通常、「ナノ粒子」という用語は、10~1000nmの大きさの粒子を意味してもよく、例えば、10nm、15nm、20nm、25nm、30nm、35nm、40nm、45nm、50nm、55nm、60nm、65nm、70nm、75nm、80nm、85nm、90nm、95nm、100nm、105nm、110nm、115nm、120nm、125nm、130nm、135nm、140nm、145nm、150nm、155nm、160nm、165nm、170nm、175nm、180nm、185nm、190nm、195nm、200nm、205nm、210nm、215nm、220nm、225nm、230nm、235nm、240nm、245nm、250nm、255nm、260nm、265nm、270nm、275nm、280nm、285nm、290nm、295nm、300nm、305nm、310nm、315nm、320nm、325nm、330nm、335nm、340nm、345nm、350nm、355nm、360nm、365nm、370nm、375nm、380nm、385nm、390nm、395nm、400nm、405nm、410nm、415nm、420nm、425nm、430nm、435nm、440nm、445nm、450nm、455nm、460nm、465nm、470nm、475nm、480nm、485nm、490nm、495nm、500nm、505nm、510nm、515nm、520nm、525nm、530nm、535nm、540nm、545nm、550nm、555nm、560nm、565nm、570nm、575nm、580nm、585nm、590nm、595nm、600nm、605nm、610nm、615nm、620nm、625nm、630nm、635nm、640nm、645nm、650nm、655nm、660nm、665nm、670nm、675nm、680nm、685nm、690nm、695nm、700nm、705nm、710nm、715nm、720nm、725nm、730nm、735nm、740nm、745nm、750nm、755nm、760nm、765nm、770nm、775nm、780nm、785nm、790nm、795nm、800nm、805nm、810nm、815nm、820nm、825nm、830nm、835nm、840nm、845nm、850nm、855nm、860nm、865nm、870nm、875nm、880nm、885nm、890nm、895nm、900nm、905nm、910nm、915nm、920nm、925nm、930nm、935nm、940nm、945nm、950nm、955nm、960nm、965nm、970nm、975nm、980nm、985nm、990nm、995nm、または1000nmの大きさの粒子であってもよい。

10

20

30

【0309】

脂質系粒子

いくつかの実施形態において、前記送達媒体は、少なくとも1つの脂質系ナノ粒子を含んでもよく、脂質系ナノ粒子として、脂質ナノ粒子（LNP）、リポソーム、ミセル、およびこれらの任意の組み合わせが挙げられるが、これらに限定されない。

【0310】

脂質ナノ粒子

いくつかの実施形態において、前記送達媒体は、脂質ナノ粒子（LNP）であってもよい。一般に、LNPは、非LNP環境と接する親水性外表面を含む外部脂質層と、非水性または水性の内部空間（すなわち、それぞれミセル様LNPおよび小胞様LNP）と、少なくとも1つの疎水性膜間空間とを有する。LNP膜は、非ラメラ構造であってもラメラ構造であってもよく、1層、2層、3層、4層、5層または5層より多い数の層で構成されていてもよい。前記LNPは、固体でも半固体でもよい。いくつかの実施形態において、前記LNPの内部空間、膜間空間、外表面またはこれらの任意の組み合わせにおいて、少なくとも1つのカーゴまたはペイロード（遺伝子挿入システム（GIS）など）が含まれていてもよい。

40

【0311】

本発明において有用なLNPは当技術分野で公知であり、概して、イオン化（カチオン

50

性) 脂質、リン脂質、コレステロール、およびポリマー修飾脂質が挙げられる。何らかの理論に拘束されることを望むものではないが、コレステロールは、膜融合を促進してLNPの安定性を補助し；リン脂質は、エンドソーム脱出を促進してLNP二重層に構造化を提供してもよく；ポリマー修飾脂質は、LNPの凝集を抑制し、免疫細胞による非特異的なエンドサイトーシスからLNPを「保護」し；イオン化(カチオン性)脂質は、エンドソーム脱出を促進し、負に荷電したカーゴ(遺伝子挿入システムのポリヌクレオチドなど)と複合体を形成する。

【0312】

いくつかの実施形態において、本発明の遺伝子挿入システムは脂質ナノ粒子(LNP)に組み込んでもよい。いくつかの実施形態において、LNPは、少なくとも1つのカチオン性脂質(例えば、イオン化カチオン性脂質)、少なくとも1つの非カチオン性脂質(例えば、リン脂質)、少なくとも1つのステロール(例えば、コレステロール)、少なくとも1つのポリマー修飾脂質(例えば、PEG脂質)、またはこれらの任意の組み合わせで構成されていてもよい。いくつかの実施形態において、LNPは、少なくとも1つのカチオン性脂質(例えば、イオン化カチオン性脂質)、少なくとも1つの非カチオン性脂質(例えば、リン脂質)、少なくとも1つのステロール(例えば、コレステロール)、および少なくとも1つのポリマー修飾脂質(例えば、PEG脂質)で構成されていてもよい。いくつかの実施形態において、LNPは、少なくとも1つのカチオン性脂質(例えば、イオン化カチオン性脂質)、少なくとも1つの非カチオン性脂質、および少なくとも1つのステロール(例えば、コレステロール)で構成されていてもよい。いくつかの実施形態において、LNPは、少なくとも1つのカチオン性脂質(例えば、イオン化カチオン性脂質)、少なくとも1つの非カチオン性脂質(例えば、リン脂質)、および少なくとも1つのポリマー修飾脂質(例えば、PEG脂質)で構成されていてもよい。いくつかの実施形態において、LNPは、少なくとも1つの非カチオン性脂質(例えば、リン脂質)、少なくとも1つのステロール(例えば、コレステロール)、および少なくとも1つのポリマー修飾脂質(例えば、PEG脂質)で構成されていてもよい。いくつかの実施形態において、LNPは、少なくとも1つのカチオン性脂質(例えば、イオン化カチオン性脂質)および少なくとも1つの非カチオン性脂質(例えば、リン脂質)で構成されていてもよい。いくつかの実施形態において、LNPは、少なくとも1つのカチオン性脂質(例えば、イオン化カチオン性脂質)および少なくとも1つのステロールで構成されていてもよい。いくつかの実施形態において、LNPは、少なくとも1つのカチオン性脂質(例えば、イオン化カチオン性脂質)および少なくとも1つのポリマー修飾脂質(例えば、PEG脂質)で構成されていてもよい。いくつかの実施形態において、LNPは、少なくとも1つの非カチオン性脂質(例えば、リン脂質)および少なくとも1つのステロール(例えば、コレステロール)で構成されていてもよい。いくつかの実施形態において、LNPは、少なくとも1つのカチオン性脂質(例えば、イオン化カチオン性脂質)および少なくとも1つのポリマー修飾脂質(例えば、PEG脂質)で構成されていてもよい。いくつかの実施形態において、LNPは、少なくとも1つのステロール(例えば、コレステロール)および少なくとも1つのポリマー修飾脂質(例えば、PEG脂質)で構成されていてもよい。いくつかの実施形態において、LNPは、少なくとも1つのカチオン性脂質(例えば、イオン化カチオン性脂質)および少なくとも1つの非カチオン性脂質(例えば、リン脂質)で構成されていてもよい。いくつかの実施形態において、LNPは、ステロール(例えば、コレステロール)で構成されていてもよい。いくつかの実施形態において、LNPは、ポリマー修飾脂質(例えば、PEG脂質)で構成されていてもよい。

【0313】

本明細書に記載のLNPは、当技術分野で公知の技術を使用して形成してもよい。一例として、マイクロ流路内において、遺伝子挿入システムを含む酸性水溶液に、脂質を含む有機溶液を混合することによって、遺伝子挿入システムが担持された送達媒体を形成させることができるが、この方法に限定されない。

10

20

30

40

50

【0314】

ミセル

いくつかの実施形態において、前記送達媒体は、少なくとも1種のミセルを含む。いくつかの実施形態において、前記ミセルは、脂質ナノ粒子と同じ成分で構成されていてもよいが、基本的に製造方法が異なる。本明細書において、「ミセル」とは、水性の粒子内空間を有していない小さな粒子を意味する。何らかの理論に拘束されることを望むものではないが、ミセルの粒子内空間には、脂質ヘッド基は含まれず、むしろ、ミセル膜を構成する脂質の疎水性テールと、それに結合する可能性のある遺伝子挿入システム（GIS）が含まれている。

【0315】

リポソーム

いくつかの実施形態において、前記送達媒体は、少なくとも1種のリポソームを含む。いくつかの実施形態において、前記リポソームは、脂質ナノ粒子と同じ成分および同じ分量で構成されていてもよいが、基本的に製造方法が異なる。本明細書において、「リポソーム」とは、水性の内部ナノ粒子空間を包囲する少なくとも1つの脂質二重膜で構成される小胞を意味する。さらに、リポソームは細胞外小胞とは異なり、一般に前駆細胞／宿主細胞に由来するものではない。リポソームとしては、狭い水性空間によって隔てられた複数の同心二重層を含み、直径が数百ナノメートルのもの（すなわち（大型）多層小胞（MLV））、直径が50nmより小さいもの（小型単層小胞（SUV））、直径が50～500nmのもの（大型単層小胞（LUV））などが挙げられる。

【0316】

エクソソーム

いくつかの実施形態において、前記送達媒体は、少なくとも1種のエクソソームを含む。一般に、「エクソソーム」とは、エンドサイトーシスに由来する小型の膜結合体である細胞外小胞を意味する。エクソソーム膜は、一般に脂質二重層で構成されたラメラ構造をとっており、水性の内部ナノ粒子空間を有する。エクソソームには、あらかじめ設定していた成分に加えて、宿主細胞膜／前駆細胞膜の成分が含まれている傾向がある。何らかの理論に拘束されることを望むものではないが、エクソソームは一般に、多胞体が細胞形質膜に融合した後、宿主細胞／前駆細胞から細胞外環境に放出される。

【0317】

ウイルス様粒子

いくつかの実施形態において、前記送達媒体は、少なくとも1種のウイルス様粒子（VLP）を含む。一般に、ウイルス様粒子は、遺伝子挿入システム（GIS）を担持することができるウイルス由来のタンパク質カプシド、コート、シェルまたはシース（本明細書ではこれらはすべて同じものとして同じ意味で使用される）を主成分として構成された非感染性小胞である。いくつかの実施形態において、VLPは、細胞内機構により発現したウイルスカプシドタンパク質配列が自己集合して遺伝子挿入システム（GIS）を取り込むことにより合成されたものであってもよい。いくつかの実施形態において、VLPは、発現に関連する細胞内機構を利用せずに、カプシドおよび遺伝子挿入システム（GIS）の各構成要素を準備して、これらを自己集合させることによって形成されたものであってもよい。

【0318】

VLPの由来となるウイルス科およびウイルス種の例としては、以下に限定されないが、バルボウイルス科、レトロウイルス科、フラビウイルス科、パラミクソウイルス科、アデノ随伴ウイルス、HIV、C型肝炎ウイルス、HPV、バクテリオファージまたはこれらの組み合わせが挙げられる。

【0319】

ポリマー性送達粒子

いくつかの実施形態において、前記送達媒体は、少なくとも1種のポリマー性送達粒子を含んでいてもよい。本明細書において、「ポリマー性送達粒子」は、様々な連結基を介

10

20

30

40

50

して遺伝子挿入システムを構成する部分に結合された可溶性ポリマーで構成された非凝集性送達粒子を意味する。いくつかの実施形態において、ポリマー性送達粒子は、本明細書に記載のポリマーのいずれを含んでいてもよい。

【0320】

いくつかの実施形態において、前記送達媒体は、核酸ナノ粒子（NANP）を含んでいてもよい。一般に、「核酸ナノ粒子」は、相互作用を介して、カーゴ（例えば、遺伝子挿入システムの各構成要素）を担持することができる3次元構造を形成可能な、非コード核酸配列から形成された小型粒子である。

【0321】

カプセル化

いくつかの実施形態において、前記送達媒体は、本明細書に開示された遺伝子挿入システムを完全に封入していてもよい。いくつかの実施形態において、前記送達媒体は、本明細書に開示された遺伝子挿入システムを部分的に封入していてもよい。いくつかの実施形態では、最終的な製剤において、送達媒体に含まれる遺伝子挿入システムの実質的に0%が、送達媒体の外部環境に露出されている（すなわち、遺伝子挿入システムは完全に封入されている）。いくつかの実施形態において、遺伝子挿入システムは送達媒体に結合しているが、少なくともその一部は送達媒体の外部環境に露出されている。

【0322】

いくつかの実施形態において、前記送達媒体は、カプセル化効率により特徴付けられてもよく、すなわち、送達媒体の外部環境に露出されていない遺伝子挿入システムの割合によって特徴付けられてもよい。明確化のために付言すると、約100%のカプセル化効率とは、遺伝子挿入システムの実質的に全体が送達媒体により完全に封入されている送達媒体製剤を意味し、約0%のカプセル化率とは、送達媒体に封入されている遺伝子挿入システムが実質的に存在しない送達媒体を意味し、例えば、遺伝子挿入システムが送達媒体の外部表面に結合された送達媒体などを意味する。いくつかの実施形態において、送達媒体のカプセル化効率は、約100%未満、約95%未満、約85%未満、約80%未満、約75%未満、約70%未満、約65%未満、約60%未満、約55%未満、約50%未満、約45%未満、約40%未満、約35%未満、約30%未満、約25%未満、約20%未満、約15%未満、約10%未満、または約5%未満であってもよい。いくつかの実施形態において、送達媒体のカプセル化効率は、約90~100%、約80~100%、約70~100%、約60~100%、約50~100%、約40~100%、約30~100%、約20~100%、約10~100%、約80~90%、約70~90%、約60~90%、約50~90%、約40~90%、約30~90%、約20~90%、約10~90%、約70~80%、約60~80%、約50~80%、約40~80%、約30~80%、約20~80%、約10~80%、約60~70%、約50~70%、約40~70%、約30~70%、約20~70%、約10~70%、約40~50%、約30~50%、約20~50%、約10~50%、約30~40%、約20~40%、約10~40%、約20~30%、約10~30%、または約10~20%であってもよい。

【0323】

送達媒体であるナノ粒子の物理学的特性

いくつかの実施形態において、前記送達媒体は、その形状によって特徴付けられる。いくつかの実施形態において、前記送達媒体は、実質的に球状であってもよく、実質的に円柱状（すなわち、円筒状）であってもよく、実質的に円盤状であってもよいが、これらに限定されない。

【0324】

いくつかの実施形態において、前記送達媒体は、その大きさによって特徴付けられる。いくつかの実施形態において、前記送達媒体の大きさは、その直径として定義することができる。前記送達媒体の大きさに関連して本明細書において使用される「直径」は、送達媒体の最大円形断面の直径を意味する。いくつかの実施形態において、前記送達媒体の直径は、30nm~約150nmであってもよい。例えば、前記送達媒体の直径は、約40~150nm、約50~150nm、約60~150nm、約70~150nm、約80~150nm、約90~1

10

20

30

40

50

50nm、約100nm以上、約110～150nm、約120～150nm、約130～150nm、約140～150nm、約30～30～140nm、約40～140nm、約50～140nm、約60～140nm、約70～140nm、約80～140nm、約90～140nm、約100～140nm、約110～140nm、約120～140nm、約130～140nm、約140～140nm、約30～140nm、約40～130nm、約50～130nm、約60～130nm、約70～130nm、約80～130nm、約90～130nm、約100～130nm、約110～130nm、約120～130nm、約30～120nm、約40～120nm、約50～120nm、約60～120nm、約70～120nm、約80～120nm、約90～120nm、約100～120nm、約110～120nm、約30～110nm、約40～110nm、約50～110nm、約60～110nm、約70～110nm、約80～110nm、約90～110nm、約100～110nm、約30～100nm、約40～100nm、約50～100nm、約60～100nm、約70～100nm、約80～100nm、約90～100nm、約30～90nm、約40～90nm、約50～90nm、約60～90nm、約70～90nm、約80～90nm、約30～80nm、約40～80nm、約50～80nm、約60～80nm、約70～80nm、約30～70nm、約40～70nm、約50～70nm、約60～70nm、約30～60nm、約40～60nm、約50～60nm、約30～50nm、約40～50nm、または約30～40nmであってもよい。

10

20

30

40

50

【0325】

いくつかの実施形態において、前記送達媒体の集団、例えば、1つの製剤に由来するすべての送達媒体は、その集団に含まれる粒子の物理学的特性（例えば、大きさ、形状または質量）の均一性を測定することによって特性評価してもよい。いくつかの実施形態において、この均一性は、その集団の多分散性指数（PI）として示してもよい。いくつかの実施形態において、この均一性は、その集団の不均衡（D）として示してもよい。本明細書において、「多分散性指数」および「不均衡」という用語は、同じものとして同じ意味で使用してもよい。

【0326】

いくつかの実施形態において、ある製剤に由来する送達媒体集団のPIは約0.1～1である。いくつかの実施形態において、ある製剤に由来する送達媒体集団のPIは、約0.1～1、約0.1～0.8、約0.1～0.6、約0.1～0.4、約0.1～0.2、約0.2～1、約0.2～0.8、約0.2～0.6、約0.2～0.4、約0.4～1、約0.4～0.8、約0.4～0.6、約0.6～1、約0.6～0.8、または約0.8～1である。いくつかの実施形態において、ある製剤に由来する送達媒体集団のPIは、約1未満、約0.5未満、約0.4未満、約0.3未満、約0.2未満、または約0.1未満である。

【0327】

送達の標的化

いくつかの実施形態において、本発明の遺伝子挿入システムを含むように製剤化された送達媒体は、本明細書に記載の標的とする領域、組織、細胞または生理系のいずれかへの本発明の遺伝子挿入システムの局在化を促進してもよい（すなわち、送達媒体は特定の場所を「標的とする」）。いくつかの実施形態において、標的化は、製剤中に含まれる送達媒体の構成成分により達成してもよい。いくつかの実施形態において、送達媒体は、標的化剤を含んでいてもよい。

【0328】

標的化剤

いくつかの実施形態において、前記送達媒体は、少なくとも1種の標的化剤を含んでいてもよい。本明細書において、「標的化剤」という用語は、いくつかの実施形態において、特定の種類または分類の細胞および／またはその他の特定の種類の化合物に特異的に結合する部分、化合物、抗体（例えば、特定の細胞または特定の細胞の種類の細胞を標的とする部分）などを意味してもよい。いくつかの実施形態において、標的化剤は、特定の標的細胞の表面、標的とする細胞表面抗原、標的とする細胞受容体、またはこれらの組み合わせに親和性を有していてもよい（すなわち、これらに対して特異的であってもよい）。

【0329】

いくつかの実施形態において、「標的化剤」は、特定の種類または分類の物質および／

または細胞に暴露された場合に、特定の作用（例えば、切断作用）を発揮することによって、送達媒体を特定の種類または分類の細胞に標的化することができる薬剤を意味してもよい。

【0330】

いくつかの実施形態において、「標的化剤」という用語は、送達媒体の一部であり、それ自身では特定の種類または分類の細胞に対して特異性を有していなくてもよいが、送達媒体の標的に対する特異性に一定の役割を果たす薬剤を意味してもよい。

【0331】

いくつかの実施形態において、送達媒体に少なくとも1種の標的化剤が含まれていることによって、送達媒体によって送達される遺伝子挿入システムが細胞に取り込まれる効率（例えば、その総量または速度）が増加してもよい。いくつかの実施形態において、送達媒体に少なくとも1種の標的化剤が含まれていることによって、送達媒体によって送達される遺伝子挿入システムが細胞に取り込まれる特異性（例えば、その総量または速度）が増加してもよい。本明細書において、「特異性」は、非標的細胞よりも標的細胞による細胞への取り込み効率が高いことを意味する。

【0332】

いくつかの実施形態において、好適な標的化剤として、低分子標的化剤（例えば、糖鎖部分）、抗体、抗体様分子、ペプチド、ビタミン類（例えば、葉酸）、糖類（例えば、ラクトースやガラクトース）、人工親和性分子（例えば、ペプチド模倣体またはアプタマー）、抗体断片、一本鎖可変領域断片（scFv）、細胞表面受容体（例えば、T細胞受容体（TCR）、B細胞受容体（BCR）またはキメラ抗原受容体（CAR））、およびこれらの任意の組み合わせのうちの1種以上が挙げられるが、これらに限定されない。

【0333】

いくつかの実施形態において、標的化剤の標的となりうる細胞表面抗原として、標的細胞の細胞表面分子が挙げられる。好適な細胞表面分子の例として、タンパク質、糖類、脂質、または細胞表面上のその他の抗原が挙げられるが、これらに限定されない。いくつかの実施形態において、細胞表面抗原は細胞内移行する。

【0334】

いくつかの特定の実施形態において、前記送達媒体は、2種以上の標的化剤を含んでもよい。

【0335】

いくつかの実施形態において、少なくとも1種の標的化剤は、ナノ粒子の脂質膜に組み込まれていてもよい。いくつかの実施形態において、少なくとも1種の標的化剤は、ナノ粒子の外表面上に提示されていてもよい。いくつかの実施形態において、少なくとも1種の標的化剤は、ナノ粒子の脂質成分に結合されていてもよい。いくつかの実施形態において、少なくとも1種の標的化剤は、ナノ粒子のポリマー成分に結合されていてもよい。いくつかの実施形態において、標的化剤の一部分を含むモノマー（例えば、標的化剤の重合可能な誘導体、例えば、ペプチドの（アルキル）アクリル酸誘導体）を共重合することにより、ポリマー修飾脂質を形成させて送達媒体を形成させることができる。いくつかの実施形態において、少なくとも1種の標的化剤は、疎水性相互作用または親水性相互作用を介してナノ粒子に繋ぎ止められていることによって、少なくとも1種の標的化剤と、ナノ粒子の膜組織と、該ナノ粒子の内部または外部の水性環境との間で連結されていてもよい。いくつかの実施形態において、少なくとも1種の標的化剤は、ナノ粒子の膜組織のペプチド／タンパク質成分に結合されている。いくつかの実施形態において、少なくとも1種の標的化剤は、ナノ粒子の膜組織の構成要素に結合された好適なリンカー部分に結合されている。いくつかの実施形態において、どのような力や結合の組み合わせでも、ナノ粒子に標的化剤を結合させることができる。

【0336】

いくつかの実施形態において、1種以上の標的化剤は、連結部分を介して送達媒体の少なくとも1つのポリマーに連結されていてもよい。いくつかの実施形態において、この連

10

20

30

40

50

結部分は、切断可能な連結部分（例えば、切断可能な結合を含む連結部分）であってもよい。いくつかの実施形態において、前記連結部分は、特定の酵素（例えば、ホスファターゼやプロテアーゼ）によって切断されてもよい結合を含んでいてもよい。いくつかの実施形態において、前記連結部分は、細胞内pH、酸化還元電位またはその他の細胞内パラメータの変化によって切断可能であってもよい結合を含んでいてもよい。いくつかの実施形態において、前記連結部分は、マトリックスメタロプロテイナーゼ（MMP）への暴露により切断されてもよい結合を含んでいてもよい。

【0337】

直接トランスフェクション

いくつかの実施形態において、本明細書に開示の遺伝子挿入システム（GIS）は、前記送達媒体を使用せずに標的細胞に直接トランスフェクトしてもよい。いくつかの実施形態において、本明細書に開示の遺伝子挿入システム（GIS）は、当技術分野で公知の任意の技術を用いて標的細胞にトランスフェクトしてもよい。そのような技術としては、化学的トランスフェクション法（例えば、リン酸カルシウム接触法）、物理的トランスフェクション法（例えば、エレクトロポレーション法、マイクロインジェクション法、遺伝子銃送達法）などが挙げられるが、これらに限定されない。いくつかの実施形態において、直接トランスフェクションは、以下に限定されないが、Lipofectamine、Lipofectamine 2000、およびこれらの任意の組み合わせなどの脂質系トランスフェクション試薬を用いて実施してもよい。

【0338】

トランスフェクト細胞の移植

いくつかの実施形態において、本発明の遺伝子挿入システムは、（例えば、本明細書に記載の直接トランスフェクションを介して）インビトロにおいて細胞集団に導入し、その後に対象に移植してもよい。いくつかの実施形態において、移植用の細胞集団は、幹細胞であってもよい。いくつかの実施形態において、移植用の細胞集団は、対象に由来する細胞であってもよい。いくつかの実施形態において、移植は、当技術分野で公知の方法で行ってもよい。

【0339】

IV. 医薬組成物および投与経路

本発明は、対象に遺伝子挿入システム（GIS）を投与するための医薬組成物を提供する。いくつかの実施形態において、本発明は、治療適応症の治療において医薬として使用するための医薬組成物を提供する。いくつかの実施形態において、前記医薬組成物は、少なくとも1つの有効成分（例えば、本発明の遺伝子挿入システム（GIS））と、少なくとも1つの薬学的に許容される添加剤、補助剤、担体、希釈剤またはこれらの任意の組み合わせを含む。いくつかの実施形態において、前記医薬組成物は、少なくとも1つの投与経路に適した製剤として調製される。いくつかの実施形態において、前記医薬組成物は、少なくとも1つの有効成分（例えば、遺伝子挿入システム（GIS））の所定の用量が送達されるような製剤として調製される。また、所定のスケジュールで送達されるような製剤として調製することもできる。

【0340】

本明細書において、「医薬組成物」という用語は、少なくとも1つの有効成分を含み、さらに1つ以上の薬学的に許容される添加剤を含んでいてもよい組成物を意味する。本明細書において、「有効成分」という用語は、通常、本明細書に記載の遺伝子挿入システム（GIS）、対象のゲノムへの挿入用の遺伝子挿入システム（GIS）によって送達される遺伝子ペイロード、または本明細書に記載の遺伝子挿入システム（GIS）によって送達される遺伝子ペイロードの発現産物を意味する。

【0341】

医薬製剤および医薬組成物

本発明の遺伝子挿入システムは、1種以上の添加剤を使用して製剤化されることにより、以下を達成してもよい。

- (1) 遺伝子挿入システムまたは該遺伝子挿入システムを含む送達機構の安定性の増加、
- (2) 細胞へのトランスフェクションまたは形質導入の増加、
- (3) 対象の細胞への遺伝子挿入システムの持続導入または遅延導入、
- (4) 生体内分布の改変（例えば、特定の組織または特定の種類の細胞への遺伝子挿入システムの標的化）、
- (5) コードされている遺伝子の発現の増加、
- (6) コードされているタンパク質の放出特性の改変、および／または
- (7) 遺伝子挿入システムおよび／またはそのペイロードの制御可能な発現。

【0342】

医薬製剤は、生理食塩水、リポソーム、脂質ナノ粒子、ポリマー、ペプチド、タンパク質、遺伝子挿入システムをトランスフェクトした細胞（例えば、対象への移入用または移植用）、およびこれらの任意の組み合わせを含んでいてもよいが、これらに限定されない。

【0343】

いくつかの実施形態において、本明細書に記載の医薬組成物の製剤は、薬理学分野で公知のまたは今後開発される任意の方法で調製してもよい。一般に、このような製剤の調製方法は、有効成分を添加剤および／または1種以上の他の補助成分と組み合わせる工程を含む。

【0344】

本明細書に記載の遺伝子挿入システムおよび医薬組成物の製剤は、薬理学分野で公知のまたは今後開発される任意の方法で行ってもよい。一般に、このような製剤の調製方法は、有効成分を添加剤および／または1つ以上のその他の補助成分と混合する工程を含み、必要性および／または所望に応じて、得られた混合物の分割工程、成形工程、および／または所望の単一用量剤形または複数用量剤形への包装工程を含んでいてもよい。

【0345】

本明細書に記載の医薬組成物は、単一の単位用量および／または複数の単位用量として、調製、包装および／または販売を行ってもよい。本明細書において、「単位用量」という用語は、所定量の有効成分を含む、個別の量の医薬組成物を意味する。一般に、有効成分の量は、対象に投与される有効成分の用量と等しく、かつ／またはそのような用量の簡便な分割量であり、例えば、そのような用量の半分または3分の1などである。

【0346】

いくつかの実施形態において、添加剤は、ヒトへの使用および獣医学的使用が承認されたものである。いくつかの実施形態において、添加剤は、米国食品医薬品局により承認されたものであってもよい。いくつかの実施形態において、添加剤は、米国薬局方（USP）、欧州薬局方（EP）、英国薬局方、および／または国際薬局方の基準に適合したものであってもよい。いくつかの実施形態において、薬学的に許容される添加剤の純度は、少なくとも100%、少なくとも99%、少なくとも98%、少なくとも97%、少なくとも96%、または95%であってもよい。いくつかの実施形態において、添加剤は、医薬品グレードのものであってもよい。

【0347】

いくつかの実施形態において、薬学的に許容される添加剤、有効成分、および／または追加の成分の相対量は、本発明の医薬組成物ごとに様々に異なってもよい。いくつかの実施形態において、前記相対量は、治療を受ける対象の体長、状態および／またはアイデンティティに応じて様々に異なってもよい。いくつかの実施形態において、前記相対量は、本発明の医薬組成物の投与経路に応じて様々に異なってもよい。例えば、本発明の医薬組成物は、0.1%~100%（例えば、0.1%~99%、0.5~50%、1~30%、5~80%、または少なくとも80%（w/w））の有効成分を含んでいてもよい。

【0348】

添加剤、希釈剤および不活性成分

10

20

30

40

50

いくつかの実施形態において、本発明の医薬組成物は、当技術分野で公知または新規な添加剤を含んでいてもよい。好適な添加剤の例として、あらゆる種類の、保存剤、等張化剤、増粘剤、乳化剤、溶媒、分散媒、希釈剤またはその他の液体溶媒、分散補助剤または懸濁補助剤、界面活性剤、およびこれらの組み合わせが挙げられるが、これらに限定されない。いくつかの実施形態において、添加剤は、所望の特定の剤形に適するように選択してもよい。

【0349】

いくつかの実施形態において、本明細書に記載の製剤は、少なくとも1種の不活性成分を含んでいてもよい。本明細書において、「不活性成分」という用語は、製剤中に含まれる1種以上の物質であって、本発明の医薬組成物の有効成分の活性に寄与しない物質を意味する。いくつかの実施形態において、本発明の医薬組成物中に含まれるすべての不活性成分もしくはその一部は、米国食品医薬品局（FDA）により承認されたものであってもよく、あるいは本発明の医薬組成物中に含まれるすべての不活性成分は、米国食品医薬品局（FDA）による承認を受けていないものであってもよい。

【0350】

いくつかの実施形態において、本明細書に開示された医薬製剤は、カチオンまたはアニオンを含んでいてもよい。いくつかの実施形態において、本開示の医薬製剤は、 Ca^{2+} 、 Zn^{2+} 、 Mn^{2+} 、 Cu^{2+} 、 Mg^{+} 、これらの任意の組み合わせなどの金属カチオンを含むが、これらに限定されない。いくつかの実施形態において、本開示の医薬製剤は、金属カチオンと複合体化されたポリマーを含んでいてもよい。

【0351】

いくつかの実施形態において、本開示の医薬組成物は、1種以上の薬学的に許容される塩を含んでいてもよい。本明細書において、「薬学的に許容される塩」という用語は、（例えば、遊離塩基を好適な有機酸と反応させることにより）親化合物中に存在する酸部分または塩基部分を塩の形態に変換することによって修飾された本開示の化合物の誘導体を意味する。本発明の薬学的に許容される塩として、例えば、非毒性の無機酸または有機酸を用いて形成された、親化合物の従来の非毒性塩が挙げられる。さらに、薬学的に許容される塩として、カルボン酸などの酸性残基のアルカリ塩または有機塩や、アミンなどの塩基性残基の無機酸塩または有機酸塩が挙げられるが、これらに限定されない。

【0352】

いくつかの実施形態において、本開示の医薬組成物は、少なくとも1種の溶媒を含んでいてもよい。いくつかの実施形態において、溶媒が水である場合、その溶媒和物は一般に「水和物」と呼ばれる。

【0353】

投与経路

本発明の遺伝子挿入システム（GIS）、例えば、本明細書に記載の遺伝子挿入システム（GIS）を含む医薬組成物は、対象の細胞に遺伝子挿入システム（GIS）を組み込むことができる送達経路であれば、いずれの経路で投与してもよい。許容可能な投与経路として、経耳投与（耳内への投与または耳経由の投与）、胆管灌流、頬側投与（頬内側への投与）、心臓灌流、仙骨ブロック、結膜投与、経皮投与、経歯投与（1本の歯または複数本の歯への投与）、歯冠内投与、診断用投与、点耳投与、電気浸透、子宮頸管内膜投与、副鼻腔内投与、気管内投与、浣腸、腸内投与（腸管内への投与）、表皮投与（皮膚の上からの穿刺）、硬膜外投与（硬膜外腔への投与）、羊膜外投与、体外循環装置を介した投与、点眼投与（結膜上への投与）、胃腸管投与、血液透析、浸潤投与、吹送投与（鼻からの吸引）、組織間隙投与、腹腔内投与、羊膜内投与、動脈内投与（動脈内への投与）、関節内投与、胆管内投与、気管支内投与、嚢内投与、心臓内投与（心臓内への投与）、軟骨内投与（軟骨内への投与）、馬尾内投与（馬尾内への投与）、空洞内注射（病的空洞内への投与）、海綿体洞内投与（陰茎基部への投与）、脳内投与（大脳への投与）、脳室内投与（脳室への投与）、大槽内投与（大槽／小脳延髄槽内への投与）、角膜内投与（角膜内への投与）、冠動脈内投与（冠動脈内への投与）、海綿体内投与（陰茎の海綿体の膨張性空間内

への投与)、皮内投与(皮膚自体への投与)、椎間板内投与(椎間板内への投与)、腺管内投与(腺管内への投与)、十二指腸内投与(十二指腸内への投与)、硬膜内投与(硬膜内または硬膜下への投与)、表皮内投与(表皮への投与)、食道内投与(食道への投与)、胃内投与(胃内への投与)、歯肉内投与(歯肉内への投与)、回腸内投与(小腸の遠位部内への投与)、病変内投与(局所病変内への投与または局所病変への直接的な導入)、腔内投与(管の内腔内への投与)、リンパ管内投与(リンパ管内への投与)、髄内投与(骨の髄腔内への投与)、髄膜内投与(髄膜内への投与)、筋肉内投与(筋肉への投与)、心筋内投与(心筋内への投与)、眼内投与(眼の内部への投与)、骨内注入(骨髄内への投与)、卵巣内投与(卵巣内への投与)、実質内投与(脳組織への投与)、心膜内投与(心膜内への投与)、腹腔内投与(腹腔内への注入または注射)、胸膜内投与(胸膜内への投与)、前立腺内投与(前立腺内への投与)、肺内投与(肺内またはその気管支内への投与)、腔内投与(副鼻腔内または眼窩周囲腔内への投与)、脊髄内投与(脊柱内への投与)、滑液包内投与(関節の滑液腔内への投与)、腱内投与(腱内への投与)、精巣内投与(精巣内への投与)、髄腔内投与(脊柱管内への投与)、髄腔内投与(脳脊髄軸の任意の部位の脳脊髄液内への投与)、胸腔内投与(胸腔内への投与)、管腔内投与(臓器の管腔内への投与)、腫瘍内投与(腫瘍内への投与)、中耳内投与(中耳内への投与)、子宮内投与、膣内投与、血管内投与(1本の血管内または複数本の血管内への投与)、静脈内投与(静脈内への投与)、静脈内ボラス、点滴静注、心室内投与(心室内への投与)、膀胱内注入、硝子体内投与(眼球の内部への投与)、イオントフォレーシス(電流を用いて可溶性塩のイオンを生体組織に導入する方法)、洗浄(開放性創傷または体腔の浸漬洗浄またはフラッシュ)、喉頭投与(喉頭への直接投与)、経鼻投与(鼻腔を介した投与)、経鼻胃投与(鼻腔を通して胃内に投与する方法)、神経ブロック、密封療法(外用経路投与を行った後、その部位を閉鎖する被覆材で覆う方法)、経眼投与(外眼部への投与)、経口投与(口腔を介した投与)、口咽頭投与(口腔および咽頭への直接投与)、非経口投与、経皮膚投与、関節周囲投与、硬膜周囲投与、神経周囲投与、歯周投与、フォトフェレーシス、直腸投与、呼吸器投与(局所効果または全身効果を得るための、経口吸入または経鼻吸入による呼吸器内への投与)、球後投与(脳橋の後ろまたは眼球の後ろへの投与)、軟部組織投与、クモ膜下投与、結膜下投与、皮下投与(皮膚の下への投与)、口唇下投与、舌下投与、粘膜下投与、外用投与、経皮投与、経皮投与(全身分布させるための無傷な皮膚を介した拡散)、経粘膜投与(粘膜を介した拡散)、経胎盤投与(胎盤を通した投与または胎盤を介した投与)、経気管投与(気管壁を通した投与)、経鼓室投与(鼓室を介した投与または鼓室を通した投与)、経膣投与、尿管投与(尿管への投与)、尿道投与(尿道への投与)、膣投与ならびに脊髄投与が挙げられるが、これらに限定されない。

10

20

30

40

50

【0354】

いくつかの実施形態において、本開示の医薬組成物は、血管障壁、血液脳関門またはその他の上皮性関門を通過できる方法で投与してもよい。遺伝子挿入システムは、好適な形態であればどのような形態で投与してもよく、具体的には、溶液、懸濁液、固体形態、溶液への溶解に適した固体形態、溶液への懸濁が可能な固体形態、およびこれらの任意の組み合わせが挙げられるが、これらに限定されない。

【0355】

いくつかの実施形態において、遺伝子挿入システムは、複数の部位への投与経路を介して対象に送達してもよい。対象の2つ、3つ、4つ、5つまたは6つ以上の部位に投与してもよい。

【0356】

いくつかの実施形態において、遺伝子挿入システムは、単一の投与経路を介して対象に送達してもよい。

【0357】

いくつかの実施形態において、ボラス注入を使用して遺伝子挿入システムを対象に投与してもよい。

【0358】

いくつかの実施形態において、持続送達方法（すなわち点滴）を使用して数分間、数時間または数日間にわたり遺伝子挿入システムを対象に投与してもよい。注入速度は、以下に限定されないが、対象の特性、所望の分布、使用する製剤などの送達パラメータに応じて変更してもよい。

【0359】

いくつかの実施形態において、遺伝子挿入システムは、以下に限定されないが、皮下注射や静脈内注射などの筋肉内送達経路により送達してもよい。

【0360】

いくつかの実施形態において、遺伝子挿入システムは、以下に限定されないが、消化管投与や頬側投与などの経口投与により送達してもよい。

【0361】

いくつかの実施形態において、遺伝子挿入システムは、以下に限定されないが、点眼剤の硝子体内注射や点眼などの眼内送達経路により送達してもよい。

【0362】

いくつかの実施形態において、遺伝子挿入システムは、以下に限定されないが、点鼻剤や鼻腔用スプレー剤などの鼻腔内送達経路により送達してもよい。

【0363】

いくつかの実施形態において、遺伝子挿入システムは、以下に限定されないが、筋肉内注射、腹腔内注射、静脈内注射、結膜注射、関節注射などの末梢への注射により対象に投与してもよい。

【0364】

いくつかの実施形態において、遺伝子挿入システムは、以下に限定されないが、髄腔内投与や脳室内投与などの脳脊髄液経路への注射により送達してもよい。

【0365】

いくつかの実施形態において、遺伝子挿入システムは、以下に限定されないが、血管内投与などの全身送達経路により送達してもよい。

【0366】

いくつかの実施形態において、遺伝子挿入システムは、実質内投与により対象に投与してもよい。

【0367】

いくつかの実施形態において、遺伝子挿入システムは、外用投与により対象に投与してもよい。

【0368】

いくつかの実施形態において、遺伝子挿入システムは、頭蓋内送達により対象に投与してもよい。

【0369】

いくつかの実施形態において、遺伝子挿入システムは、筋肉内投与により対象に投与してもよい。

【0370】

いくつかの実施形態において、遺伝子挿入システムは、静脈内投与により対象に投与してもよい。

【0371】

いくつかの実施形態において、遺伝子挿入システムは、皮下投与により対象に投与してもよい。

【0372】

いくつかの実施形態において、遺伝子挿入システムは、2種以上の投与経路により送達してもよい。

【0373】

注射剤および非経口投与

いくつかの実施形態において、本明細書に記載の医薬組成物は、非経口投与してもよい

10

20

30

40

50

。非経口投与用および経口投与用の液体剤形として、薬学的に許容される液剤、乳剤、マイクロエマルジョン剤、エリキシル剤、懸濁剤、および／またはシロップ剤が挙げられるが、これらに限定されない。液体剤形は、有効成分に加えて、当技術分野で一般に使用されている不活性希釈剤を含んでいてもよく、このような不活性希釈剤として、例えば、可溶化剤、水またはその他の溶媒、および乳化剤（例えば、ポリエチレングリコール、プロピレングリコール、1,3-ブチレングリコール、テトラヒドロフルフリルアルコール、イソプロピルアルコール、エチルアルコール、炭酸エチル、酢酸エチル、ベンジルアルコール、安息香酸ベンジル、ジメチルホルムアミド、油、グリセロール、およびソルビタン脂肪酸エステル）、およびこれらの任意の組み合わせが挙げられる。油の例として、綿実油、ラッカセイ油、トウモロコシ油、胚芽油、オリーブ油、ヒマシ油、およびゴマ油、ならびにこれらの混合物が挙げられる。いくつかの実施形態において、本開示の医薬組成物は、アルコール、油、グリコール、CREMOPHOR（登録商標）、変性油、ポリソルベート、ポリマー、シクロデキストリンおよび／またはこれらの組み合わせなどの可溶化剤を含む。いくつかの実施形態において、ヒドロキシプロピルセルロースなどの界面活性剤が含まれる。

10

【0374】

いくつかの実施形態において、注射用調製物は、滅菌注射用水溶懸濁液または滅菌注射用油性懸濁液を含んでいてもよい。注射用滅菌溶液は、好適な湿潤剤、分散剤および／または懸濁剤を使用して、公知の技術により製剤化してもよい。滅菌注射用調製物は、非経口的に許容される非毒性の希釈剤および／または溶媒中の滅菌注射用懸濁液、滅菌注射用溶液および／または滅菌注射用エマルジョンであってもよい。いくつかの実施形態において、滅菌注射用調製物は、1,3-ブタンジオール溶液であってもよい。いくつかの実施形態において、許容される媒体および溶媒として、リンゲル溶液（米国薬局方）、水、等張食塩水、および滅菌不揮発性油が挙げられるが、これらに限定されない。いくつかの実施形態において、不揮発性油として、非刺激性の不揮発性油（例えば、合成モノグリセリドまたは合成ジグリセリド）が挙げられる。いくつかの実施形態において、オレイン酸などの脂肪酸を注射剤の調製に使用することもできる。

20

【0375】

いくつかの実施形態において、注射用製剤は、細菌捕捉フィルターを通したろ過、および／または滅菌剤の添加により滅菌処理を行ってもよい。いくつかの実施形態において、滅菌剤は、使用前に滅菌水などの滅菌注射用媒体に溶解または分散可能な滅菌固体組成物の形態であってもよい。

30

【0376】

有効成分の効果を持続させるために、皮下注射または筋肉注射からの有効成分の吸収を遅延させることが望ましいことが多い。いくつかの実施形態において、本開示の医薬組成物を油性溶媒に溶解または懸濁することによって、非経口投与される医薬組成物の吸収を遅延させることができる。いくつかの実施形態において、低水溶性の非晶質成分または結晶性成分の液体懸濁液を使用することによって、有効成分の吸収を遅延させてもよい。有効成分の吸収率は溶解率に依存し、この溶解率は結晶の大きさと結晶形態に依存する。

【0377】

経口投与

いくつかの実施形態において、本明細書に記載の医薬組成物および／または医薬製剤は、経口投与してもよい。経口投与用の固体剤形として、錠剤、カプセル剤、散剤、丸剤および顆粒剤が挙げられる。一般に、固体剤形では、少なくとも1種の薬学的に許容される不活性な添加剤に有効成分を混合し、このような添加剤として、リン酸二カルシウムまたはクエン酸ナトリウム；結合剤（例えば、カルボキシメチルセルロース、アルギン酸塩、ゼラチン、ポリビニルピロリドン、スクロースおよびアラビアゴム）；充填剤または増量剤（例えば、デンプン、ラクトース、スクロース、グルコース、マンニトールおよびケイ酸）；崩壊剤（例えば、カンテン、炭酸カルシウム、バレイショデンプンまたはタピオカデンプン、アルギン酸、特定のケイ酸塩および炭酸ナトリウム）；吸収促進剤（例えば、

40

50

第四級アンモニウム化合物)；保湿剤(例えば、グリセロール)；液状遅延剤(例えば、パラフィン)；吸収剤(例えば、カオリンおよびベントナイトクレイ)；湿潤剤(例えば、セチルアルコールおよびモノステアリン酸グリセリン)；滑沢剤(例えば、タルク、ステアリン酸カルシウム、ステアリン酸マグネシウム、固形ポリエチレングリコール、ラウリル硫酸ナトリウム)；およびこれらの任意の組み合わせが挙げられるが、これらに限定されない。錠剤、カプセル剤および丸剤の場合、これらの剤形は緩衝剤を含んでいてもよい。

【0378】

経口投与用の液体剤形は、前述した非経口投与用の添加剤を含んでいてもよい。経口組成物は、不活性希釈剤の他にも、乳化剤、湿潤剤、懸濁剤、香料、甘味剤および／または芳香剤などの補助剤を含んでいてもよい。

10

【0379】

外用投与または経皮投与

いくつかの実施形態において、本明細書に記載の医薬組成物および／または医薬製剤は、外用投与用に製剤化してもよい。皮膚はアクセスが容易であることから、理想的な送達用標的部位になりうる。いくつかの実施形態において、本明細書に記載の医薬組成物の皮膚への送達経路または皮膚を介した送達経路として、外用塗布(例えば、美容用途および／または局所／領域治療用)、皮内注射(例えば、美容用途および／または局所／領域治療用)、および全身送達(例えば、皮膚領域と真皮外領域の両方に影響を及ぼす皮膚疾患の治療用)が挙げられるが、これらに限定されない。

20

【0380】

いくつかの実施形態において、本明細書に記載の医薬組成物および／または医薬製剤は、本明細書に記載の方法を効果的かつ／または簡便に実施するために、様々な被覆包帯(例えば、絆創膏または創傷被覆材)を用いて送達してもよい。いくつかの実施形態において、被覆材または包帯は、使用者自身が複数回の治療を行えるように、本明細書に記載の医薬組成物を十分な量で含んでいてもよい。

【0381】

外用投与用および／または経皮投与用の剤形として、ローション剤、クリーム剤、軟膏剤、ゲル剤、スプレー剤、ペースト剤、散剤、液剤、吸入剤および／またはパッチ剤が挙げられる。一般に、外用投与用および／または経皮投与用の剤形は、無菌条件下において、薬学的に許容される添加剤、緩衝剤および／または必要に応じて保存剤と有効成分を混合することにより製剤化してもよい。

30

【0382】

いくつかの実施形態において、経皮パッチを使用してもよい。経皮パッチは、本明細書に記載の医薬組成物を身体に制御送達できるというさらなる利点を有していてもよい。一般に、経皮パッチは、本明細書に記載の医薬組成物を適切な媒体に溶解および／または分散することによって調製してもよい。いくつかの実施形態において、送達速度は、ポリママトリックスおよび／もしくはゲルに本発明の医薬組成物を分散することによって、または放出速度を制御する膜を設けることによって、またはこれらを組み合わせによって制御してもよい。

40

【0383】

いくつかの実施形態において、外用投与に好適な製剤として、液体調製物および／または半液体調製物(例えば、リニメント剤およびローション剤)、水中油型エマルジョン剤および／または油中水型エマルジョン剤(例えば、軟膏剤、クリーム剤および／またはペースト剤)、液剤および／または懸濁剤、ならびにこれらの任意の組み合わせが挙げられる。

【0384】

経眼投与または経耳投与

いくつかの実施形態において、本明細書に記載の医薬組成物は、経眼投与、経耳投与またはその両方に好適な製剤の形態であってもよい。一般に、そのような製剤は、点眼剤お

50

よび／または点耳剤の形態であってもよく、具体的には、水性液体添加剤および／または油性液体添加剤に有効成分を加えた液剤および／または懸濁剤が挙げられるが、これらに限定されない。いくつかの実施形態において、そのような点眼剤および／または点耳剤は、塩、緩衝剤、本明細書に記載の1種以上のその他の追加の成分、およびこれらの組み合わせを含んでもよい。いくつかの実施形態において、経眼投与可能な製剤として、リポソーム調製物および／または微結晶形態の有効成分が挙げられる。いくつかの実施形態において、本明細書に記載の医薬組成物は、網膜下投与してもよい。

【0385】

経肺投与

いくつかの実施形態において、本明細書に記載の医薬組成物は、経肺投与に好適な製剤の形態であってもよい。いくつかの実施形態において、経肺投与は、頬腔を介して行われる。いくつかの実施形態において、本明細書に記載の医薬組成物は、有効成分を含む乾燥粒子を含んでもよい。いくつかの実施形態において、経肺投与用の乾燥粒子の直径は、約0.5～7nmの範囲または約1～6nmの範囲であってもよい。

10

【0386】

いくつかの実施形態において、自己噴射型の溶媒／粉体供給容器を使用して本明細書に記載の医薬組成物を投与してもよい。一般に、有効成分は、密閉容器内において低沸点の噴射剤中に溶解および／または懸濁されていてもよい。いくつかの実施形態において、本明細書に記載の医薬組成物は、そのような粉末が噴射剤の流れによって分散される乾燥粉末リザーバーを備えた装置を使用して投与される乾燥粉末の形態であってもよい。乾燥粉末が使用されるいくつかの実施形態において、この粉末は、重量基準で少なくとも98%の粒子が0.5nmより大きい直径を有し、個数基準で少なくとも95%の粒子が7nm未満の直径を有する粒子を含んでもよい。いくつかの実施形態において、重量基準で少なくとも95%の粒子が1nmより大きい直径を有し、個数基準で少なくとも90%の粒子が6nm未満の直径を有する。いくつかの実施形態において、粉末を含む乾燥医薬組成物は、固体微粉末の形態の希釈剤（例えば、糖類）を含んでもよく、簡便に単位用量形態で提供してもよい。

20

【0387】

いくつかの実施形態において、低沸点の噴射剤として、大気圧下で65°Fを下回る沸点を有する液体噴射剤が挙げられる。いくつかの実施形態において、本発明の医薬組成物中の噴射剤の含有量は50%～99.9%（w/w）であってもよく、本発明の医薬組成物中の有効成分の含有量は0.1%～20%（w/w）であってもよい。いくつかの実施形態において、噴射剤は、追加の成分を含んでもよく、具体的には、液体の非イオン性界面活性剤、固体のアニオン性界面活性剤、固体希釈剤（例えば、有効成分を含む粒子と同じ桁数の粒径を有する固体希釈剤が挙げられる）、およびこれらの任意の組み合わせが挙げられるが、これらに限定されない。

30

【0388】

いくつかの実施形態において、経肺送達用に製剤化された医薬組成物は、液剤、懸濁剤およびその組み合わせの液滴の形態であってもよい。そのような製剤は、液剤、懸濁剤またはその組み合わせとして調製、包装および／または販売する場合、噴霧化装置および／またはネブライザー装置を使用して投与してもよい。いくつかの実施形態において、これらの液剤および／または懸濁剤は滅菌されていてもよい。例示的な液剤および／または懸濁剤として、水性アルコール組成物および／または希釈されたアルコール組成物が挙げられる。いくつかの実施形態において、経肺送達用に製剤化された医薬組成物は、香料（例えば、サッカリンナトリウム）、揮発性油、界面活性剤、緩衝剤、保存剤（例えば、ヒドロキシ安息香酸メチル）、およびこれらの任意の組み合わせを含んでもよい。いくつかの実施形態において、経肺投与経路により提供される液滴の平均直径は、約0.1nm～約200nmの範囲であってもよい。

40

【0389】

鼻腔内投与、経鼻投与または頬側投与

50

いくつかの実施形態において、本明細書に記載の医薬組成物は、鼻腔内投与、経鼻投与またはその両方で投与してもよい。いくつかの実施形態において、鼻腔内送達用の医薬組成物は、経肺送達用として本明細書で述べた添加剤を含んでいてもよい。いくつかの実施形態において、鼻腔内投与用の医薬組成物は、有効成分を含む粗い粉末を含み、その平均粒径は、約 $0.2\text{ }\mu\text{m}$ ～ $500\text{ }\mu\text{m}$ である。いくつかの実施形態において、鼻腔内投与用の医薬組成物は、鼻の近くに保持した粉末容器から鼻腔を介して急速に吸入することによって、すなわち、鼻からすするようにして投与してもよい。例示的な医薬製剤は、約 0.1% （ w/w ）～ 100% （ w/w ）の有効成分を含んでいてもよく、本明細書に記載の1種以上の追加の成分をさらに含んでいてもよい。

【0390】

いくつかの実施形態において、本明細書に記載の医薬組成物は、頬側投与に好適な製剤であってもよく、このような製剤として、錠剤、トローチ剤およびその組み合わせが挙げられるが、これらに限定されない。一般に、そのような錠剤またはトローチ剤は、従来の方法を使用して作製してもよく、（非限定的な例として） 0.1% ～ 20% （ w/w ）の有効成分を含んでいてもよく、経口摂取により溶解可能な組成物および経口摂取により崩壊可能な組成物のあらゆる組み合わせを含んでいてもよく、任意の構成要素として本明細書に記載の1種以上の追加の成分を含んでいてもよい。いくつかの実施形態において、頬側投与に好適な医薬組成物は、有効成分をそれぞれ含む粉剤、エアロゾル化された液剤および／もしくは懸濁剤、または噴霧化された液剤および／または懸濁剤のあらゆる組み合わせを含んでいてもよく、これらの剤形は、約 0.1 nm ～ 200 nm の平均粒径の分散粒子および／または約 0.1 nm ～ 200 nm の液滴径の分散液滴の形態であってもよい。いくつかの実施形態において、頬側投与用の医薬組成物は、本明細書に記載の1種以上の追加の成分をさらに含んでいてもよい。

【0391】

デポー投与

いくつかの実施形態において、本明細書に記載の医薬組成物は、徐放用のデポー剤の形態に製剤化してもよい。いくつかの実施形態において、本明細書に記載の医薬組成物は、標的組織内またはその近傍に空間的に保持される。

【0392】

注射用デポー剤形は、一般に、生分解性ポリマー（例えば、ポリ乳酸-ポリグリコリド）中に封入した医薬組成物のマイクロカプセル化マトリックスを形成することによって作製される。一般に、医薬組成物の放出速度は、医薬組成物とポリマーの比率を変更すること、および使用する特定のポリマーの性質を変更することによって制御することができる。好適な生分解性ポリマーとして、ポリ（オルトエステル）およびポリ（無水物）が挙げられるが、これらに限定されない。デポー注射用製剤は、生体組織と適合性を有するリポソームまたはマイクロエマルジョン内に医薬組成物を封入することによって調製される。

【0393】

直腸投与および経膣投与

いくつかの実施形態において、本明細書に記載の医薬組成物は、直腸投与してもよく、経膣投与してもよく、その組み合わせで投与してもよい。一般に、直腸投与または経膣投与用の組成物は坐剤であり、この坐剤は、周囲温度では固体であるが体温により液状となる好適な非刺激性添加剤（例えば、ポリエチレングリコール、カカオ脂、または坐剤用ワックス）と有効成分を混合することにより調製することができる。坐剤が直腔または膣腔で溶解すると有効成分が放出される。

【0394】

用量

本発明の遺伝子挿入システム（GIS）および／または該遺伝子挿入システム（GIS）を含む医薬組成物は、対象において所望の効果（例えば、所望の治療効果や研究結果など）をもたらす量（すなわち、用量）で投与してもよい。いくつかの実施形態において、所望の用量は、対象のパラメータ（例えば、対象の体長、状態もしくは性質）、効果のパラメ

ータ（例えば、要求される応答の程度、治療的な効果の閾値、効果の持続性、もしくは提示される副作用）、またはこれらの任意の組み合わせに基づいて決定してもよい。いくつかの実施形態において、適切な用量は、初回投与前に決定してもよく、少なくとも1つの対象のパラメータを試験する少なくとも1つのアッセイに基づいて決定してもよい。いくつかの実施形態において、適切な用量は、初回投与後に決定してもよく、少なくとも1つの効果パラメータを試験する少なくとも1つのアッセイに基づいて決定してもよい。いくつかの実施形態において、用量は、投与コースを通して変更せずに維持してもよい。いくつかの実施形態において、用量は投与コースを通して、1回、2回、または多数回変更してもよい。

【0395】

いくつかの実施形態において、前記用量は、対象の質量（例えば、mg/kgの単位の質量）に対する有効成分の質量比として表してもよい。例えば、前記用量は、0.1~100mg/kg、1~100mg/kg、2~100mg/kg、3~100mg/kg、4~100mg/kg、5~100mg/kg、6~100mg/kg、7~100mg/kg、8~100mg/kg、9~100mg/kg、10~100mg/kg、15~100mg/kg、20~100mg/kg、25~100mg/kg、30~100mg/kg、35~100mg/kg、40~100mg/kg、45~100mg/kg、50~100mg/kg、55~100mg/kg、60~100mg/kg、65~100mg/kg、70~100mg/kg、75~100mg/kg、80~100mg/kg、85~100mg/kg、90~100mg/kg、95~100mg/kg、0.1~95mg/kg、1~95mg/kg、2~95mg/kg、3~95mg/kg、4~95mg/kg、5~95mg/kg、6~95mg/kg、7~95mg/kg、8~95mg/kg、9~95mg/kg、10~95mg/kg、15~95mg/kg、20~95mg/kg、25~95mg/kg、30~95mg/kg、35~95mg/kg、40~95mg/kg、45~95mg/kg、50~95mg/kg、55~95mg/kg、60~95mg/kg、65~95mg/kg、70~95mg/kg、75~95mg/kg、80~95mg/kg、85~95mg/kg、90~95mg/kg、0.1~90mg/kg、1~90mg/kg、2~90mg/kg、3~90mg/kg、4~90mg/kg、5~90mg/kg、6~90mg/kg、7~90mg/kg、8~90mg/kg、9~90mg/kg、10~90mg/kg、15~90mg/kg、20~90mg/kg、25~90mg/kg、30~90mg/kg、35~90mg/kg、40~90mg/kg、45~90mg/kg、50~90mg/kg、55~90mg/kg、60~90mg/kg、65~90mg/kg、70~90mg/kg、75~90mg/kg、80~90mg/kg、85~90mg/kg、0.1~85mg/kg、1~85mg/kg、2~85mg/kg、3~85mg/kg、4~85mg/kg、5~85mg/kg、6~85mg/kg、7~85mg/kg、8~85mg/kg、9~85mg/kg、10~85mg/kg、15~85mg/kg、20~85mg/kg、25~85mg/kg、30~85mg/kg、35~85mg/kg、40~85mg/kg、45~85mg/kg、50~85mg/kg、55~85mg/kg、60~85mg/kg、65~85mg/kg、70~85mg/kg、75~85mg/kg、80~85mg/kg、0.1~80mg/kg、1~80mg/kg、2~80mg/kg、3~80mg/kg、4~80mg/kg、5~80mg/kg、6~80mg/kg、7~80mg/kg、8~80mg/kg、9~80mg/kg、10~80mg/kg、15~80mg/kg、20~80mg/kg、25~80mg/kg、30~80mg/kg、35~80mg/kg、40~80mg/kg、45~80mg/kg、50~80mg/kg、55~80mg/kg、60~80mg/kg、65~80mg/kg、70~80mg/kg、75~80mg/kg、0.1~75mg/kg、1~75mg/kg、2~75mg/kg、3~75mg/kg、4~75mg/kg、5~75mg/kg、6~75mg/kg、7~75mg/kg、8~75mg/kg、9~75mg/kg、10~75mg/kg、15~75mg/kg、20~75mg/kg、25~75mg/kg、30~75mg/kg、35~75mg/kg、40~75mg/kg、45~75mg/kg、50~75mg/kg、55~75mg/kg、60~75mg/kg、65~75mg/kg、70~75mg/kg、0.1~70mg/kg、1~70mg/kg、2~70mg/kg、3~70mg/kg、4~70mg/kg、5~70mg/kg、6~70mg/kg、7~70mg/kg、8~70mg/kg、9~70mg/kg、10~70mg/kg、15~70mg/kg、20~70mg/kg、25~70mg/kg、30~70mg/kg、35~70mg/kg、40~70mg/kg、45~70mg/kg、50~70mg/kg、55~70mg/kg、60~70mg/kg、65~70mg/kg、0.1~65mg/kg、1~65mg/kg、2~65mg/kg、3~65mg/kg、4~65mg/kg、5~65mg/kg、6~65mg/kg、7~65mg/kg、8~65mg/kg、9~65mg/kg、10~65mg/kg、15~65mg/kg、20~65mg/kg、25~

10

20

30

40

50

65mg/kg、30～65mg/kg、35～65mg/kg、40～65mg/kg、45～65mg/kg、50～65mg/kg、55～65mg/kg、60～65mg/kg、0.1～60mg/kg、1～60mg/kg、2～60mg/kg、3～60mg/kg、4～60mg/kg、5～60mg/kg、6～60mg/kg、7～60mg/kg、8～60mg/kg、9～60mg/kg、10～60mg/kg、15～60mg/kg、20～60mg/kg、25～60mg/kg、30～60mg/kg、35～60mg/kg、40～60mg/kg、45～60mg/kg、50～60mg/kg、55～60mg/kg、0.1～55mg/kg、1～55mg/kg、2～55mg/kg、3～55mg/kg、4～55mg/kg、5～55mg/kg、6～55mg/kg、7～55mg/kg、8～55mg/kg、9～55mg/kg、10～55mg/kg、15～55mg/kg、20～55mg/kg、25～55mg/kg、30～55mg/kg、35～55mg/kg、40～55mg/kg、45～55mg/kg、50～55mg/kg、0.1～50mg/kg、1～50mg/kg、2～50mg/kg、3～50mg/kg、4～50mg/kg、5～50mg/kg、6～50mg/kg、7～50mg/kg、8～50mg/kg、9～50mg/kg、10～50mg/kg、15～50mg/kg、20～50mg/kg、25～50mg/kg、30～50mg/kg、35～50mg/kg、40～50mg/kg、45～50mg/kg、0.1～45mg/kg、1～45mg/kg、2～45mg/kg、3～45mg/kg、4～45mg/kg、5～45mg/kg、6～45mg/kg、7～45mg/kg、8～45mg/kg、9～45mg/kg、10～45mg/kg、15～45mg/kg、20～45mg/kg、25～45mg/kg、30～45mg/kg、35～45mg/kg、40～45mg/kg、0.1～40mg/kg、1～40mg/kg、2～40mg/kg、3～40mg/kg、4～40mg/kg、5～40mg/kg、6～40mg/kg、7～40mg/kg、8～40mg/kg、9～40mg/kg、10～40mg/kg、15～40mg/kg、20～40mg/kg、25～40mg/kg、30～40mg/kg、35～40mg/kg、0.1～35mg/kg、1～35mg/kg、2～35mg/kg、3～35mg/kg、4～35mg/kg、5～35mg/kg、6～35mg/kg、7～35mg/kg、8～35mg/kg、9～35mg/kg、10～35mg/kg、15～35mg/kg、20～35mg/kg、25～35mg/kg、30～35mg/kg、0.1～30mg/kg、1～30mg/kg、2～30mg/kg、3～30mg/kg、4～30mg/kg、5～30mg/kg、6～30mg/kg、7～30mg/kg、8～30mg/kg、9～30mg/kg、10～30mg/kg、15～30mg/kg、20～30mg/kg、25～30mg/kg、0.1～25mg/kg、1～25mg/kg、2～25mg/kg、3～25mg/kg、4～25mg/kg、5～25mg/kg、6～25mg/kg、7～25mg/kg、8～25mg/kg、9～25mg/kg、10～25mg/kg、15～25mg/kg、20～25mg/kg、0.1～20mg/kg、1～20mg/kg、2～20mg/kg、3～20mg/kg、4～20mg/kg、5～20mg/kg、6～20mg/kg、7～20mg/kg、8～20mg/kg、9～20mg/kg、10～20mg/kg、15～20mg/kg、0.1～15mg/kg、1～15mg/kg、2～15mg/kg、3～15mg/kg、4～15mg/kg、5～15mg/kg、6～15mg/kg、7～15mg/kg、8～15mg/kg、9～15mg/kg、10～15mg/kg、0.1～10mg/kg、1～10mg/kg、2～10mg/kg、3～10mg/kg、4～10mg/kg、5～10mg/kg、6～10mg/kg、7～10mg/kg、8～10mg/kg、9～10mg/kg、0.1～9mg/kg、1～9mg/kg、2～9mg/kg、3～9mg/kg、4～9mg/kg、5～9mg/kg、6～9mg/kg、7～9mg/kg、8～9mg/kg、0.1～8mg/kg、1～8mg/kg、2～8mg/kg、3～8mg/kg、4～8mg/kg、5～8mg/kg、6～8mg/kg、7～8mg/kg、0.1～7mg/kg、1～7mg/kg、2～7mg/kg、3～7mg/kg、4～7mg/kg、5～7mg/kg、6～7mg/kg、0.1～6mg/kg、1～6mg/kg、2～6mg/kg、3～6mg/kg、4～6mg/kg、5～6mg/kg、0.1～5mg/kg、1～5mg/kg、2～5mg/kg、3～5mg/kg、4～5mg/kg、0.1～4mg/kg、1～4mg/kg、2～4mg/kg、3～4mg/kg、0.1～3mg/kg、1～3mg/kg、2～3mg/kg、0.1～2mg/kg、1～2mg/kg、または0.1～1mg/kgであってもよい。

10

20

30

40

50

【0396】

投与スケジュール

本発明の遺伝子挿入システムおよび／または該遺伝子挿入システムを含む医薬組成物は、対象において所望の効果（例えば、所望の治療効果や研究結果など）をもたらす頻度（すなわち、投与スケジュール）で投与してもよい。いくつかの実施形態において、投与スケジュールは、本明細書に記載の用量の決定に使用される方法のいずれかにより決定してもよい。いくつかの実施形態において、遺伝子挿入システムは、1回のみ投与してもよい

。

【0397】

いくつかの実施形態において、本発明の遺伝子挿入システムは、2回以上投与してもよい。例えば、本発明の遺伝子挿入システムは、2回、3回、4回、5回、6回、7回、8回、9回、10回、またはそれ以上の回数投与してもよい。いくつかの実施形態において、本発明の遺伝子挿入システムは、対象における治療適応症の治療過程において断続的かつ／または連続的に投与してもよい。いくつかの実施形態において、本発明の遺伝子挿入システムは、一生にわたって対象に繰り返し投与してもよい。

【0398】

V. 使用方法遺伝子挿入システム製剤を送達する際の標的領域、標的組織または標的細胞

本明細書において、対象の少なくとも1つの標的位置に、本明細書に記載の医薬組成物および／または医薬製剤を送達する方法であって、生理系、解剖学的部位、臓器、組織、特定の種類の細胞、細胞集団などの少なくとも1つの標的（1つ以上の標的細胞を含む）に、本明細書に記載の少なくとも1つの医薬組成物および／または医薬製剤を接触させることにより実施される方法が提供される。

【0399】

本明細書に記載の医薬組成物および／または医薬製剤は、標的に位置する少なくとも1つの細胞において、目的とする効果（例えば、少なくとも1つの導入遺伝子を対象のゲノムに挿入するという効果）を達成するのに十分な量の有効成分（例えば、本発明の遺伝子挿入システム）を含む。

【0400】

いくつかの実施形態において、本明細書に記載の医薬組成物および／または医薬製剤は、通常、1種以上の細胞膜透過剤を含むが、薬学的に許容される担体を含んでいてもよい「裸」の製剤（例えば、細胞膜透過剤やその他の薬剤を含まない製剤）も想定される。

【0401】

生理系

いくつかの実施形態において、本明細書に記載の医薬組成物および／または医薬製剤は、生理系を標的とする。

【0402】

いくつかの実施形態において、生理系として、聴覚系、心血管系、中枢神経系、化学受容器系、循環器系、消化器系、内分泌系、排泄器系、外分泌系、生殖器系（genital system）、外皮系、リンパ系、筋肉系、筋骨格系、神経系、末梢神経系、腎臓系、生殖器系（reproductive system）、呼吸器系、泌尿器系、および視覚系が挙げられる。

【0403】

いくつかの実施形態において、本明細書に記載の医薬組成物および／または医薬製剤は、アミン前駆体取り込み脱炭酸系（APUD）（内分泌機能を有し、様々な低分子アミンまたはポリペプチドホルモンを分泌する一連の細胞）を標的とし、APUD系組織として、下垂体組織、副甲状腺組織、甲状腺組織、気管支組織、副腎髄質組織、膵臓組織、胃腸管、頸動脈小体、化学受容器系組織などが挙げられるが、これらに限定されない。

【0404】

臓器

いくつかの実施形態において、本明細書に記載の医薬組成物および／または医薬製剤は、臓器を標的とする。臓器として、肛門管、動脈、上行結腸、膀胱、骨髄、脳、気管支、細気管支、尿道球腺、毛細血管、盲腸、小脳、大脳半球、大脳、子宮頸部、脈絡叢、陰核、脳神経、下行結腸、間脳、十二指腸、耳、腸神経系、精巣上体、食道、外性器、卵管、胆嚢、神経節、味覚器、腸管関連リンパ組織、心臓、回腸、内部生殖器、間質、空腸、閥節、腎臓、大腸、喉頭、靱帯、肝臓、肺、リンパ節、リンパ管、乳腺、延髄、腸間膜、中脳、口腔、呼吸器筋肉、鼻腔、神経、嗅覚器、卵巣、膵臓、耳下腺、陰茎、咽頭、胎盤、脳橋、前立腺、直腸、唾液腺、陰嚢、精嚢、S状結腸、骨格、皮膚、小腸、脊髄神経、脾

臓、胃、皮下組織、舌下腺、顎下腺、歯、腱、精巣、脳幹、脊髄、脳室系、胸腺、舌、扁桃腺、気管、横行結腸、尿管、尿道、子宮、膣、輸精管、静脈および外陰が挙げられる。

【0405】

いくつかの実施形態において、本明細書に記載の医薬組成物および／または医薬製剤は、一方の眼または両方の眼を標的としてもよい。

【0406】

いくつかの実施形態において、本明細書に記載の医薬組成物および／または医薬製剤は、肝臓を標的としてもよい。

【0407】

いくつかの実施形態において、本明細書に記載の医薬組成物および／または医薬製剤は、脳を標的としてもよい。

【0408】

細胞

いくつかの実施形態において、本明細書に記載の医薬組成物および／または医薬製剤は、特定の細胞および／または特定の種類の細胞を標的としてもよい。

【0409】

細胞として、脂肪細胞、アドレナリン作動性神経細胞、 α 細胞、アマクリン細胞、エナメル芽細胞、前房レンズ上皮細胞、下垂体前葉／中葉細胞、アポクリン汗腺細胞、アストロサイト、コルチ器の聴覚内有毛細胞、コルチ器の聴覚外有毛細胞、B細胞、バルトリン腺細胞、角膜、舌、口腔、鼻腔、遠位肛門管、遠位尿道および遠位膣の基底細胞（幹細胞）、嗅上皮基底細胞、バスケット細胞、好塩基球性顆粒球およびその前駆細胞、 β 細胞、betz細胞、骨髓細網組織線維芽細胞、コルチ器の境界細胞、境界細胞、ボウマン腺細胞、褐色脂肪細胞、ブルナー腺細胞、尿道球腺細胞、房状細胞、C細胞、カハール・レチウス細胞、心筋細胞、心筋細胞、カートホイール細胞、グルココルチコイドを産生する束状帯細胞、ミネラルコルチコイドを産生する球状帯細胞、アンドロゲンを産生する網状帯細胞、副腎皮質細胞、セメント芽細胞、房心細胞、耳の耳道腺細胞、シャンデリア細胞、頸動脈小体細胞の化学受容グロムス細胞、主細胞、コリン作動性ニューロン、クロム親和細胞、クラブ細胞、低温感受性の一次知覚性ニューロン、結合組織マクロファージ（すべてのタイプ）、角膜線維芽細胞（角膜の角膜実質細胞）、プロゲステロンを分泌する破裂卵胞の黄体細胞、皮質毛幹細胞、副腎皮質刺激ホルモン分泌細胞、クリスタリンを含むレンズ線維細胞、表皮毛幹細胞、細胞傷害性T細胞、D細胞、 δ 細胞、樹状細胞、ダブルブーケ細胞、導管細胞、エクリン汗腺明細胞、エクリン汗腺暗細胞、精巣輸出管細胞、弾性軟骨の軟骨細胞、内皮細胞、腸管グリア細胞、腸クロム親和細胞、腸クロム親和細胞様細胞、腸内分泌細胞、好酸球性顆粒球およびその前駆細胞、上衣細胞、表皮基底細胞、表皮ランゲルハンス細胞、精巣上体基底細胞、精巣上体主細胞、上皮網状細胞、 ϵ 細胞、赤血球、線維軟骨の軟骨細胞、フォークニューロン、腺窩細胞、G細胞、胆嚢上皮細胞、生殖細胞、litter腺細胞、眼瞼のmoll腺細胞、グリア細胞、ゴルジ細胞、生殖腺間質細胞、性腺刺激ホルモン分泌細胞、顆粒細胞、顆粒膜細胞、顆粒膜黄体細胞、格子細胞および頭方位細胞が挙げられる。

【0410】

いくつかの実施形態において、細胞はがん細胞であってもよい。いくつかの実施形態において、細胞は非がん細胞であってもよい。

【0411】

いくつかの実施形態において、真核細胞は幹細胞であってもよい。様々な種類の幹細胞が当技術分野において知られており、そのすべてを本開示を実施するために使用してもよい。幹細胞の例として、胚性幹細胞、造血幹細胞、神経幹細胞、表皮神経堤幹細胞、人工多能性幹細胞、乳房幹細胞、腸管幹細胞、間葉系幹細胞、嗅覚器成体幹細胞、精巣細胞、および前駆細胞（例えば、神経前駆細胞、血管芽細胞、骨芽細胞、軟骨芽細胞、腭前駆細胞、表皮前駆細胞など）が挙げられるが、これらに限定されない。いくつかの実施形態において、幹細胞は、対象から採取された細胞に由来する胚幹細胞株であってもよい。

10

20

30

40

50

【0412】

いくつかの実施形態において、真核細胞は、ヒト、非ヒト霊長類および／またはその他の哺乳動物（マウスおよび／またはラットを含む）の循環器系に見られる細胞である。例示的な循環器系細胞として、血小板、形質細胞、赤血球、B細胞、T細胞、ナチュラルキラー細胞、マクロファージ、好中球、これらの前駆細胞などが挙げられるが、これらに限定されない。いくつかの実施形態において、少なくとも1つの真核細胞は、これらの循環器系真核細胞のいずれに由来するものであってもよい。

【0413】

いくつかの実施形態において、少なくとも1つの真核細胞は、ナチュラルキラー細胞、またはナチュラルキラー細胞の前駆細胞である。

10

【0414】

いくつかの実施形態において、少なくとも1つの真核細胞は、B細胞または前駆B細胞である。

【0415】

いくつかの実施形態において、真核細胞は植物細胞であってもよい。いくつかの実施形態において、植物細胞は、単子葉植物または双子葉植物の細胞であり、具体的には、ズッキーニ、蒴果植物や落葉樹などの木本植物、コムギ、カブ、トマト、タバコ、ヒマワリ、サトウキビ、テンサイ、イチゴ、ホウレンソウ、大豆、モロコシ、ライ麦、米、ラズベリー、ナタネ、ラディッシュ、カボチャ、ジャガイモ（サツマイモを含む）、プラム、パイナップル、ピーナッツ、エンドウ、パパイア、オートムギ、メロン、マンゴー、トウモロコシ、レタス、レンズマメ、ハーブ、麻、牧草、花、ユーカリ、キュウリ、綿、コーヒー、柑橘類、チコリー、サクランボ、セロリ、カリフラワー、ニンジン、キャノーラ、キャベツ、ブロッコリー、アブラナ、ブラックベリー、豆類、大麦、バナナ、アボカド、アスパラガス、シロイヌナズナ、その他の果実植物、観賞植物、アーモンド、アルファルファ、多年生牧草、飼料作物、その他の野菜、その他の核果（例えば、モモ、ネクタリン、アプリコット、西洋ナシ、プラムなど）、その他の仁果類（例えば、リンゴや西洋ナシなど）、その他の果実、その他の鱗莖類（例えば、ニンニク、タマネギ、ニラなど）、その他の農作物、多年生植物の一部（例えば、球茎；塊茎；根；クラウン；茎；ストロン；ひこばえ；シュート；根なしの挿木苗、挿木苗、およびカルス付きの挿木苗またはカルスを形成する幼植物体などの挿木苗；頂端分裂組織など）、およびこれらの任意の組み合わせ、またはこれらのハイブリッドが挙げられるが、これらに限定されない。本明細書において、「植物」は、種子、実生、若木、根、塊茎、茎、柄、葉および果実などの、植物の物理的な部分を意味する。

20

30

【0416】

腫瘍

いくつかの実施形態において、本明細書に記載の医薬組成物および／または医薬製剤は、腫瘍を標的とする。この腫瘍は、良性腫瘍、前がん性腫瘍または悪性腫瘍であってもよい。

【0417】

導入遺伝子の挿入

本発明は、対象、例えば、ヒト対象に導入遺伝子を導入する方法を提供する。いくつかの実施形態において、この方法は、本明細書に記載の少なくとも1つの遺伝子挿入システムを有効量で前記対象に導入する工程を含む。いくつかの実施形態において、前記方法は、導入遺伝子を含む少なくとも1つの遺伝子挿入システムを有効量で前記対象に導入する工程を含む。

40

【0418】

いくつかの実施形態において、前記方法は、1つ以上の標的挿入部位に導入遺伝子を挿入する工程を含んでもよい。図8を参照すると、挿入された導入遺伝子を有する対象のゲノム領域500が示されている。この例において、対象のゲノムDNAは、標的挿入部位120と、その周囲のゲノムDNA 110を含んでいる。明確化のために付言すると、標的

50

挿入部位は対象のDNAの一部である。5'連結部の510は、対象のDNAと挿入された導入遺伝子520の間の遷移点を示し、これは導入遺伝子の5'末端に相当する。この連結部510は、対象のゲノムと鋳型RNAの5'末端の両方に存在する、標的部位の上流配列の一部または全体の重複部分を有していてもよい。これに対して、3'連結部530は、導入遺伝子の3'末端と対象のDNAの間の遷移点を示している。この連結部530は、対象のゲノムと鋳型RNAの3'側モジュールの両方に存在する、標的部位の下流配列の一部または全体の重複部分を有していてもよい。連結部510および／または連結部530は、別のヌクレオチドをさらに含んでいてもよく、この別のヌクレオチドは、例えば、逆転写酵素により鋳型以外のヌクレオチドが伸長前のプライマーまたはcDNA 3'末端プライマーに付加され、鋳型産物の二本鎖から酵素により解離されて生じたヌクレオチドなどであってもよい。

10

【0419】

標的挿入部位

いくつかの実施形態において、1つ以上の標的挿入部位はセーフハーバー部位を含む。本明細書において、「セーフハーバー部位」は、導入遺伝子を挿入しても意図しない細胞機能の破壊が起こらない対象のゲノムの特定の位置を意味する。一般に、(a) その部位での遺伝物質の挿入が対象の遺伝子の発現を改変しない場合、または(b) その部位での遺伝物質の挿入によって遺伝子の発現が改変されるが、(例えば、対象のゲノムの妨害された遺伝子に多くの反復配列があることから) その変化によって対象の細胞の正常な機能が改変されない場合に、対象のゲノムにおいてその部位がセーフハーバー部位であると特定されてもよい。(b) の場合の例として、ゲノム内において、一部のrRNA遺伝子が妨害を受けても正常な細胞機能が摂動しないような量でリボソームRNA (rRNA) のコード遺伝子が反復している場合が挙げられるが、これに限定されない。

20

【0420】

いくつかの実施形態において、少なくとも1つのセーフハーバー部位および／または標的挿入部位は、少なくとも1つのリボソームDNA (rDNA) 配列を含む。本明細書において、「リボソームDNA」という用語は、rRNAをコードする遺伝子を意味する。いくつかの実施形態において、少なくとも1つのセーフハーバー部位および／または標的挿入部位は、少なくとも1つの28S rDNA配列を含む。

【0421】

導入遺伝子

本発明の方法および組成物は、どのようなペイロード配列（すなわち導入遺伝子）の挿入にでも使用することができ、ペイロード配列の長さやその由来は限定されない。

30

【0422】

いくつかの実施形態において、導入遺伝子は、治療的に活性な遺伝子を含む。本明細書において、「治療的に活性な遺伝子」という用語は、少なくとも1つの治療適応症の治療、緩和または予防に有用な発現産物を発現することができるあらゆる遺伝子を意味する。

【0423】

いくつかの実施形態において、少なくとも1つの導入遺伝子は、少なくとも1つのテロメラーゼ逆転写酵素 (TERT) 遺伝子を含んでいてもよい。いくつかの実施形態において、少なくとも1つの導入遺伝子は、少なくとも1つの短鎖型第VIII因子遺伝子を含んでいてもよい。いくつかの実施形態において、少なくとも1つの導入遺伝子は、少なくとも1つのフェニルアラニンヒドロキシラーゼ (PAH) 遺伝子を含んでいてもよい。

40

【0424】

いくつかの実施形態において、少なくとも1つの導入遺伝子はレポーター遺伝子である。本明細書において、「レポーター遺伝子」は、アッセイにより検出可能な発現産物を発現することができるあらゆる遺伝子を意味する。

【0425】

いくつかの実施形態において、少なくとも1つのレポーター遺伝子は、以下に限定されないが、少なくとも1つの緑色蛍光タンパク質 (GFP)、少なくとも1つの赤色蛍光タンパク質 (RFP)、ルシフェラーゼ酵素 (LUC)、 β -ガラクトシダーゼ (LacZ)、ク

50

ロラムフェニコールアセチルトランスフェラーゼ（cat）などを含んでいてもよく、これらをコードしてもよい。

【0426】

非野生型導入遺伝子

当業者であれば、上記で例示した導入遺伝子の多くは天然型配列または野生型配列であるが、本明細書において開示された遺伝子挿入システムは、野生型遺伝子もしくは天然遺伝子またはその遺伝子配列の一部の挿入に限定されないことを理解できるであろう。本発明の遺伝子挿入システムは、例えば、野生型遺伝子に由来する遺伝子、野生型遺伝子の一部のみを含む遺伝子、様々な種類の野生型遺伝子の一部から組み立てた遺伝子、および／またはその配列が自然界に存在することが知られていない遺伝子を挿入するために使用してもよい。さらに、本発明の遺伝子挿入システムは、対象の細胞においてその発現産物が通常見られない導入遺伝子、および／またはその遺伝子発現が通常見られない導入遺伝子を挿入するために使用してもよい。

10

【0427】

導入遺伝子の調節因子

いくつかの実施形態において、本発明の遺伝子挿入システムは、少なくとも1つの調節因子を含むか、これをコードする少なくとも1つの導入遺伝子を挿入するために使用してもよい。例えば、導入遺伝子の発現産物において、任意の数のmiRNA結合領域および／またはsiRNA結合領域が含まれるように導入遺伝子を設計および／または組換えてもよい。通常、miRNAおよび／またはsiRNAを組み込むことによって、トランスクリプトーム内に相補的なmiRNAまたはsiRNAを含む細胞を導入遺伝子の発現の標的から外すことができる。

20

【0428】

いくつかの実施形態において、導入遺伝子は、少なくとも1つのmiRNAおよび／またはsiRNAを含むか、これらをコードする第1の発現産物と、第1の発現産物と相補的な少なくとも1つのmiRNA結合部位および／またはsiRNA結合部位を含むか、これらをコードする第2の発現産物（またはそれ以上の個数の発現産物）とを含んでいてもよく、これらをコードしてもよい。何らかの理論に拘束されることを望むものではないが、このような態様によって、第2の発現産物の長期的な発現を阻止してもよい。

【0429】

30

抗体

本明細書において、「抗体」という用語は、最も広い意味で使用され、具体的には、様々な実施形態を包含し、所望の生物学的活性を示す限り（例えば、「機能性」である限り）、モノクローナル抗体、ポリクローナル抗体、多重特異性抗体（例えば、少なくとも2つの全長抗体から形成された二重特異性抗体）、および抗体断片（例えば、ダイアボディ）を含むが、これらに限定されない。抗体は、主にアミノ酸からなる分子であり、既知の抗体配列または親抗体配列に由来する少なくとも1つのアミノ酸領域を含む、単量体ポリペプチドまたは多量体ポリペプチドである。抗体は、1つ以上の内因性修飾または非天然修飾（糖類部分、蛍光部分、化学タグなどの付加を含むが、これらに限定されない）を動員するアミノ酸モチーフを含んでいてもよい。本発明の目的を達成するにあたり、「抗体」は、重鎖可変ドメインおよび軽鎖可変ドメイン、ならびにFc領域を含んでいてもよい。

40

【0430】

本発明の遺伝子挿入システムは、少なくとも1つ以上の機能性抗体を含むか、これをコードする導入遺伝子を挿入するために使用してもよい。

【0431】

治療適応症の治療

本発明は、少なくとも1つの治療適応症の治療または予防を必要とする対象において、少なくとも1つの治療適応症を治療または予防する方法を提供する。いくつかの実施形態において、前記方法は、本明細書に記載の少なくとも1つの遺伝子挿入システムを有効量で前記対象に導入する工程を含む。いくつかの実施形態において、前記方法は、少なくとも

50

1つの治療的に活性な導入遺伝子を含む少なくとも1つの遺伝子挿入システムを有効量で前記対象に導入する工程を含む。

【0432】

いくつかの実施形態において、少なくとも1つの治療適応症は、機能喪失を伴う少なくとも1つの遺伝性疾患を含む。いくつかの実施形態において、少なくとも1つの治療適応症の治療方法の少なくとも1つは、機能喪失を伴う遺伝性疾患から対象を回復させる少なくとも1つの導入遺伝子を投与する工程を含む。本明細書において、「回復させる」という用語は、対象が失っていた元の機能を発揮することを可能にする少なくとも1つの組成物を該対象に提供することを意味する。

【0433】

いくつかの実施形態において、前記少なくとも1つの方法は、少なくとも1つのTERT導入遺伝子を含む遺伝子挿入システムを有効量で対象に投与することによって、該対象の不十分なテロメラーゼ活性を回復させる工程を含む。

【0434】

いくつかの実施形態において、本発明の方法および組成物は、対象において不十分なテロメラーゼ機能によって引き起こされた状態を治療または予防するために使用してもよい。いくつかの実施形態において、前記少なくとも1つの方法は、不十分なテロメラーゼ活性を呈している対象に、少なくとも1つのTERT遺伝子を含む少なくとも1つの遺伝子挿入システムを治療有効量で投与する工程を含む。いくつかの実施形態において、少なくとも1つの方法は、不十分なテロメラーゼ活性により疾患を発症する疑いのある対象の少なくとも1つのTERT遺伝子を含む少なくとも1つの遺伝子挿入システムを治療有効量で投与する工程を含む。

【0435】

異種遺伝子の調節

本明細書に記載の医薬製剤および医薬組成物を包含する本発明の遺伝子挿入システムは、異種遺伝子の発現を調節する方法において使用してもよい。明確化のために付言すると、「異種遺伝子」という用語は、本明細書における遺伝子発現の調節に関して使用される場合、対象のゲノムにおいて、本発明の遺伝子挿入システムにより挿入される遺伝子以外の遺伝子を意味する。

【0436】

通常、異種遺伝子の発現を調節する方法は、本発明の遺伝子挿入システムを使用して、その発現産物が別の遺伝子の発現経路に作用する配列を挿入することを含んでいてもよい。例えば、挿入された遺伝子の発現産物は、異種遺伝子からmRNAへの転写、異種遺伝子のmRNAからポリペプチドへの翻訳、細胞質における異種遺伝子のmRNAの分解速度または不活性化速度、またはこれらの任意の組み合わせなどに影響を与えてもよい。

【0437】

いくつかの実施形態において、本発明の少なくとも1つの遺伝子挿入システムは、少なくとも1つのマイクロRNA (miRNA) を含む導入遺伝子、または少なくとも1つのマイクロRNA (miRNA) をコードする導入遺伝子を挿入するために使用してもよい。いくつかの実施形態において、本開示の実施に好適なmiRNAは、当技術分野で公知のmiRNAまたは今後見出されるmiRNAを含んでいてもよい。いくつかの実施形態において、本発明の少なくとも1つの遺伝子挿入システムは、少なくとも1つの人工miRNAを含む導入遺伝子、または少なくとも1つの人工miRNAをコードする導入遺伝子を挿入するために使用してもよく、この人工のmiRNAは、対象内に存在する少なくとも1つの遺伝子発現産物と結合するように設計されている。本明細書において、「人工miRNA」という用語は、所望の標的配列に結合するようにその配列が改変または設計されたmiRNAを意味する。人工miRNAは、当技術分野で公知の様々な方法で設計してもよい。

【0438】

いくつかの実施形態において、本発明の少なくとも1つの遺伝子挿入システムは、少なくとも1つの低分子干渉RNA (siRNA) を含む導入遺伝子、または少なくとも1つの低

10

20

30

40

50

分子干渉RNA (siRNA) をコードする導入遺伝子を挿入するために使用してもよい。本明細書において、「低分子干渉RNA」という用語は、標的遺伝子の少なくとも一部と実質的に同一のヌクレオチド配列を有する二本鎖リボ核酸 (dsRNA) を意味する。一般に、siRNAは通常、21~25ntの長さであるが、これよりも短くても長くてもよく、標的遺伝子のmRNAの分解を促進することによって、標的遺伝子の発現を妨害 (抑制) する。既知のsiRNAまたは今後見出されるsiRNAは、本発明での使用に好適であってもよい。

【0439】

いくつかの実施形態において、本発明の少なくとも1つの遺伝子挿入システムは、少なくとも1つの人工siRNAを含む導入遺伝子、または少なくとも1つの人工siRNAをコードする導入遺伝子を挿入するために使用してもよい。本明細書において、「人工siRNA」という用語は、少なくとも1つの目的の遺伝子と相補性を有するようにその配列が設計されたsiRNAを意味する。

【0440】

いくつかの実施形態において、本発明の少なくとも1つの遺伝子挿入システムは、少なくとも1つの転写因子 (TF) を含む導入遺伝子、または少なくとも1つの転写因子 (TF) をコードする導入遺伝子を挿入するために使用してもよい。本明細書において、「転写因子」という用語は、DNAに結合して、少なくとも1つの遺伝子の転写を改変するか、これに影響を与えるポリペプチドを意味する。既知の転写因子または今後見出される転写因子は、本発明での使用に好適であってもよい。

【0441】

本発明の遺伝子挿入システムは、miRNA、siRNAおよび／または転写因子の任意の組み合わせを含むか、これらをコードする導入遺伝子を挿入するために使用してもよい。例えば、少なくとも1つの遺伝子挿入システムは、少なくとも1つのmiRNAと少なくとも1つのsiRNA；少なくとも1つのmiRNAと少なくとも1つの転写因子；少なくとも1つのsiRNAと少なくとも1つの転写因子；または少なくとも1つmiRNAと、少なくとも1つのsiRNAと、少なくとも1つの転写因子を含むか、これらをコードする導入遺伝子を挿入するために使用してもよい。

【0442】

予防的用途

いくつかの実施形態において、本明細書に記載の医薬組成物および／または医薬製剤は、疾患の予防または治療適応症の進行の安定化に使用してもよい。

【0443】

いくつかの実施形態において、本明細書に記載の医薬組成物および／または医薬製剤は、将来的な治療適応症の発症を予防するための予防策として使用してもよい。

【0444】

いくつかの実施形態において、本明細書に記載の医薬組成物および／または医薬製剤は、治療適応症のさらなる進行を阻止するために使用してもよい。

【0445】

ワクチン

いくつかの実施形態において、本明細書に記載の医薬組成物および／または医薬製剤は、ワクチンとして、かつ／またはワクチンと同様に使用してもよい。本明細書において、「ワクチン」は、特定の治療適応症または感染因子に対する免疫を向上させる生物学的調製物を意味する。

【0446】

いくつかの実施形態において、本明細書に記載の医薬組成物および／または医薬製剤は、以下に限定されないが、皮膚科、中枢神経科、心臓血管科、腫瘍内科、内分泌内科、免疫科、呼吸器科、抗感染症科などの治療領域用のワクチンとして使用してもよく、かつ／またはこのような用途のワクチンと同様に使用してもよい。

【0447】

抗原

本発明の遺伝子挿入システムは、少なくとも1つの抗原を含む導入遺伝子、または少なくとも1つの抗原をコードする導入遺伝子を挿入するために使用してもよく、この抗原は、対象の少なくとも1つの細胞により活性化されてもよく、またはその表面に提示されてもよい。本明細書において、「抗原」という用語は、生物において免疫応答を引き起こす組成物を意味する。例えば、対象の生物において自身に対する抗体を産生させることができる組成物であり、特に、その抗体により対象の生物体内で獲得免疫応答を惹起できる組成物を意味する。抗原は、例えば、ポリペプチド、タンパク質、多糖類、核酸、脂質などの免疫原性物質であってもよい。いくつかの実施形態において、抗原は、細菌、ウイルス、原虫、真菌、プリオンなどの感染因子に由来するものであってもよいが、これらに限定されない。

10

【0448】

いくつかの実施形態において、抗原は、感染因子の一部またはサブユニットを含んでもよく、例えば、コート、コート成分、コートタンパク質、コートポリペプチド、表面成分、表面タンパク質、表面ポリペプチド、莢膜成分、細胞壁成分、鞭毛、線毛、毒素、またはトキソイドを含んでもよい。

【0449】

いくつかの実施形態において、本発明の少なくとも1つの遺伝子挿入システムは、少なくとも1つの治療適応症に対するワクチン接種を対象に行うための少なくとも1つの抗原を含む導入遺伝子、またはこのような抗原をコードする導入遺伝子を挿入するために使用してもよい。

20

【0450】

研究

いくつかの実施形態において、本明細書に記載の医薬組成物および／または医薬製剤は、本明細書に開示された治療適応症に対する診断目的または研究ツールとして使用してもよい。

【0451】

いくつかの実施形態において、本明細書に記載の医薬組成物および／または医薬製剤は、あらゆる研究実験に使用してもよく、例えば、インビボまたはインビトロでの実験に使用してもよい。

【0452】

いくつかの実施形態において、本明細書に記載の医薬組成物および／または医薬製剤は、研究用バイオマーカーの検出に使用してもよい。

30

【0453】

いくつかの実施形態において、本明細書に記載の医薬組成物および／または医薬製剤は、培養細胞に使用してもよい。培養細胞は、当業者に公知のどのような供給源に由来するものであってもよく、安定な細胞株、動物モデルまたはヒト患者もしくはコントロール対象に由来する細胞であってもよいが、これらに限定されない。

【0454】

いくつかの実施形態において、本明細書に記載の医薬組成物および／または医薬製剤は、動物モデル（すなわち、マウス、ラット、ウサギ、ネコ、イヌ、非ヒト霊長類、モルモット、ショウジョウバエ、フェレット、線虫、ゼブラフィッシュ、または当技術分野において公知の研究目的で使用するその他の動物）におけるインビボでの実験に使用してもよい。

40

【0455】

いくつかの実施形態において、本明細書に記載の医薬組成物および／または医薬製剤は、幹細胞および／または細胞分化に使用してもよい。

【0456】

いくつかの実施形態において、本明細書に記載の医薬組成物および／または医薬製剤は、ヒトの研究実験またはヒトの臨床試験に使用してもよい。

【0457】

50

本発明は、対象に関する科学研究および／または医学研究を行う方法を提供する。いくつかの実施形態において、前記方法は、本明細書に記載の少なくとも1つの遺伝子挿入システムを有効量で前記対象に導入する工程を含む。いくつかの実施形態において、前記方法は、少なくとも1つのレポーター導入遺伝子を含む少なくとも1つの遺伝子挿入システムを有効量で前記対象に導入する工程を含む。

【0458】

単独療法および併用療法

いくつかの実施形態において、本明細書に記載の医薬組成物および／または医薬製剤は、疾患の治療用の単独治療薬または併用治療薬として使用してもよい。

【0459】

いくつかの実施形態において、本明細書に記載の医薬組成物および／または医薬製剤は、単独療法として使用してもよい。いくつかの実施形態において、本明細書に記載の医薬組成物および／または医薬製剤は、併用療法に使用してもよい。併用療法は、1種以上の神経保護剤を組み合わせてもよく、神経保護剤は、例えば、神経細胞変性に対する神経保護効果が評価されている低分子化合物、増殖因子、ホルモンなどであってもよい。

【0460】

いくつかの実施形態において、本明細書に記載の医薬組成物および／または医薬製剤は、1種以上のその他の治療剤と組み合わせ使用してもよい。「組み合わせる」とは、複数の治療剤の同時投与および／または共送達用の製剤化を示唆すると解釈すべきではないが、このような送達方法も本発明の範囲内に含まれる。本明細書に記載の医薬組成物および／または医薬製剤、ならびにその他の治療剤は、1種以上のその他の所望の治療薬または医療処置と同時に、その前に、またはその後に投与することができる。通常、各治療剤は、その治療剤に対して規定された用量および／またはタイムスケジュールで投与される。

【0461】

本明細書に記載の医薬組成物および／または医薬製剤と併用してもよい治療剤は、低分子化合物であってもよく、この低分子化合物は、抗酸化剤、抗炎症剤、抗アポトーシス剤、カルシウム調節剤、抗グルタミン酸作動剤、構造タンパク質阻害剤、筋機能に関与する化合物、または金属イオン調節に関与する化合物である。

【0462】

インビボでの遺伝子挿入コンストラクト（GIC）の合成

本発明は、遺伝子挿入システム（GIS）生体高分子、例えば、遺伝子挿入コンストラクト（GIC）生体高分子を合成する方法を提供する。いくつかの実施形態において、この方法は、少なくとも1つのGIC合成用コンストラクトを対象の細胞集団に投与する工程、前記対象の細胞により少なくとも1つのGIS合成用コンストラクトが発現されるのに十分な時間にわたって前記細胞集団を維持する工程、および当技術分野で公知の方法により、前記GIS合成用コンストラクトの発現産物を回収し、精製する工程を含む。

【0463】

いくつかの実施形態において、少なくとも1つのGIC合成用コンストラクトは、本発明のGICを含むか、これをコードする。いくつかの実施形態において、少なくとも1つのGIC合成用コンストラクトは、本発明のGICと、少なくとも1つの組換えRNAをインビボで合成するための手段とを含むか、これらをコードする。そのような手段は、RNAポリメラーゼプロモーター、前記組換えRNAの選択用および精製用の配列、相補的GIC配列、および組換え後RNA産生プロセッシングシグナルを提供またはコードすることを含んでもよい。いくつかの実施形態において、少なくとも1つのGIC合成用コンストラクトは、コードされたRNAを細胞内の内因性機構により産生させることを可能にするDNAプラスミドの形態で投与される。

【0464】

例示的なGIC合成用コンストラクト600を図9に示す。このGIC合成用コンストラクトの5'末端において、RNAPモジュール610は、好適なRNAポリメラーゼプロモーター（

10

20

30

40

50

例えば、T7 RNAPプロモーター)を含んでいてもよい。任意の構成要素である5'リーダーモジュール620が存在する場合、この5'リーダーモジュール620は、RNAPモジュールの3'側に位置する。この5'リーダーモジュール620に含まれる構成要素として、鋳型の5'側モジュールの折り畳みと自己切断を向上させる構成要素であって、かつ／または転写産物を不安定化させ、かつ／もしくは免疫原性の5'末端(例えば、RZによる自己切断の失敗の結果として生じる可能性がある)を有するGIC転写産物を迅速に除去することが可能な構成要素が含まれていてもよい。GICとして使用する前に、発現された5'リーダーモジュールRNAは、RZ自己切断部位630において切断される。5'側モジュールの相補鎖640、鋳型モジュールの相補鎖650、および3'側モジュールの相補鎖660は、それぞれGICの5'側モジュール、鋳型モジュール、および3'側モジュールをコードする。さらに、3'末端は、線状化制限酵素部位670であってもよく、この部位は、制限酵素による切断点であり、この切断点において制限酵素により切断されることによって、GIC RNAが線状化され、すべての余分なベクター構成要素はベクター上に残存する。

10

【0465】

VII. 実施形態の一覧

実施形態1. ゲノム編集用システムであって、

(i) ポリヌクレオチド鋳型を逆転写する酵素活性を有するポリペプチドをコードするポリヌクレオチドを含む少なくとも1つの逆転写酵素コンストラクト(RTC)と、

(ii) 前記少なくとも1つのRTCによってコードされるポリペプチドによる逆転写に適した少なくとも1つのポリヌクレオチド鋳型を含む少なくとも1つの遺伝子挿入コンストラクト(GIC)を含む、システム。

20

【0466】

実施形態2. 前記少なくとも1つの逆転写酵素コンストラクトが、少なくとも1つの生体高分子を含み、該生体高分子が、少なくとも1つの核酸、少なくとも1つのアミノ酸、およびこれらの任意の組み合わせを含む、実施形態1に記載のシステム。

【0467】

実施形態3. 前記少なくとも1つの逆転写酵素コンストラクトが、少なくとも1つの逆転写酵素モジュール(RTC: RTモジュール)を含み、少なくとも1つの5'側モジュール(RTC: 5'側モジュール)、少なくとも1つの3'側モジュール(RTC: 3'側モジュール) 30)、およびこれらの任意の組み合わせを含んでいてもよい、実施形態1または2に記載のシステム。

【0468】

実施形態4. 前記少なくとも1つの逆転写酵素モジュールが、少なくとも1つの逆転写酵素を含むか、これをコードする、実施形態3に記載のシステム。

【0469】

実施形態5. 前記少なくとも1つの逆転写酵素モジュールが、非長鎖末端反復(非LTR)型レトロエレメントに由来する少なくとも1つの逆転写酵素を含むか、これをコードする、実施形態3または4に記載のシステム。

【0470】

実施形態6. 前記少なくとも1つの逆転写酵素モジュールが、非天然翻訳開始コドンを含むか、これをコードする、実施形態4または5に記載のシステム。

40

【0471】

実施形態7. 前記少なくとも1つの逆転写酵素が、少なくとも1つのDNA結合ドメイン、少なくとも1つのRNA結合ドメイン、少なくとも1つのcDNA合成ドメイン、少なくとも1つのエンドヌクレアーゼドメイン、およびこれらの任意の組み合わせを含む、実施形態4~6のいずれか1つに記載のシステム。

【0472】

実施形態8. 少なくとも1つの逆転写酵素ドメイン、少なくとも1つの対象DNA結合ドメイン、少なくとも1つの鋳型RNA結合ドメイン、および少なくとも1つのエンドヌク

50

レアーゼドメイン、ならびにこれらの任意の組み合わせのうちの少なくとも1つが、残りのドメインの少なくとも1つが由来する逆転写酵素の由来生物種とは異なる生物種に由来するものである、実施形態7に記載のシステム。

【0473】

実施形態9. 前記逆転写酵素コンストラクトの任意の構成要素である少なくとも1つの5'側モジュールが、少なくとも1つのRNAポリメラーゼプロモーター、少なくとも1つの5'非翻訳領域(5'-UTR)、少なくとも1つのコザック配列、少なくとも1つの5'キャップ、およびこれらの任意の組み合わせを含むか、これらをコードする、実施形態3に記載のシステム。

【0474】

実施形態10. 前記逆転写酵素コンストラクトの任意の構成要素である少なくとも1つの3'側モジュールが、少なくとも1つの逆転写酵素翻訳終止コドン、少なくとも1つの3'非翻訳領域(3'UTR)、少なくとも1つのポリAテール、およびこれらの任意の組み合わせを含むか、これらをコードする、実施形態3に記載のシステム。

【0475】

実施形態11. 前記少なくとも1つの逆転写酵素モジュールが、図2~5に記載の少なくとも1つの構造またはこれらの任意の組み合わせを含むか、これらをコードする、実施形態1~10のいずれか1つに記載のシステム。

【0476】

実施形態12. 前記少なくとも1つの逆転写酵素コンストラクトが、配列番号1~57およびこれらの任意の組み合わせのうちの少なくとも1つの配列を含むか、この配列をコードするか、この配列によってコードされる、実施形態1~11のいずれか1つに記載のシステム。

【0477】

実施形態13. 前記少なくとも1つの遺伝子挿入コンストラクトが、少なくとも1つの核酸生体高分子を含むか、これをコードする、実施形態1に記載のシステム。

【0478】

実施形態14. 前記少なくとも1つの遺伝子挿入コンストラクトが、任意の構成要素である少なくとも1つのGIC:5'側モジュールと、少なくとも1つのGIC:ペイロードモジュールと、任意の構成要素である少なくとも1つのGIC:3'側モジュール、およびこれらの任意の組み合わせを含むか、これらをコードする、実施形態1または13に記載のシステム。

【0479】

実施形態15. 前記少なくとも1つのGIC:5'側モジュールが、天然レトロエレメントの5'領域に由来する少なくとも1つの配列を含むか、これをコードし、少なくとも1つのrRNA配列、少なくとも1つのリボザイム配列、少なくとも1つのフォールディングモチーフ配列、またはこれらの任意の組み合わせを含んでもよく、これらをコードしてもよい、実施形態14に記載のシステム。

【0480】

実施形態16. 前記GIC:5'側モジュールの任意の構成要素である少なくとも1つのrRNA配列が、対象の1~30ntのrRNAを含むか、これをコードする、実施形態15に記載のシステム。

【0481】

実施形態17. 前記GIC:5'側モジュールの任意の構成要素である少なくとも1つのリボザイム配列が、少なくとも1つの自己切断型リボザイムを含むか、これをコードし、該自己切断型リボザイムが、デルタ肝炎ウイルスのリボザイムを含んでもよく、これをコードしてもよい、実施形態15に記載のシステム。

【0482】

実施形態18. 前記GIC:5'側モジュールの任意の構成要素である少なくとも1つのリボザイム配列が、少なくとも1つの非長鎖末端反復型レトロエレメントの5'領域に由来

10

20

30

40

50

するリボザイムを含むか、これをコードする、実施形態17に記載のシステム。

【0483】

実施形態19. 前記GIC: 5'側モジュールの任意の構成要素である少なくとも1つのフォールディングモチーフ配列が、少なくとも1つの自律型フォールディングRNA配列モチーフを含むか、これをコードし、該自律型フォールディングRNA配列モチーフが、少なくとも1つのヘアピンモチーフ、少なくとも1つのステムループモチーフ、少なくとも1つの対合したステム4モチーフ、またはこれらの任意の組み合わせを含んでいてもよく、これらをコードしてもよい、実施形態15に記載のシステム。

【0484】

実施形態20. 前記GIC: 5'側モジュールが、配列番号60~154、250~276および278~279の少なくとも1つ、またはこれらの任意の組み合わせを含むか、これらのいずれか1つをコードする、実施形態14~19のいずれか1つに記載のシステム。 10

【0485】

実施形態21. 前記少なくとも1つのGIC: 3'側モジュールが、少なくとも1つの逆転写酵素認識配列を含むか、これをコードし、少なくとも1つのrRNA配列、少なくとも1つのAトラクト配列、またはこれらの任意の組み合わせを含んでいてもよく、これらをコードしてもよい、実施形態14に記載のシステム。

【0486】

実施形態22. 前記GIC: 3'側モジュールの少なくとも1つの逆転写酵素認識配列が、少なくとも1つの逆転写酵素と相互作用する少なくとも1つの配列を含むか、これをコードする、実施形態21に記載のシステム。 20

【0487】

実施形態23. 前記GIC: 3'側モジュールの少なくとも1つの逆転写酵素認識配列が、天然レトロエレメントの3'領域に由来するものである、実施形態21または22に記載のシステム。

【0488】

実施形態24. 前記GIC: 3'側モジュールの任意の構成要素である少なくとも1つのrRNA配列が、1~30ntのrRNAを含むか、これをコードする、実施形態21に記載のシステム。

【0489】

実施形態25. 前記GIC: 3'側モジュールの任意の構成要素である少なくとも1つのAトラクト配列が、1~50個のアデニン塩基からなる配列を含むか、これをコードする、実施形態21に記載のシステム。 30

【0490】

実施形態26. 前記少なくとも1つのGIC: 3'側モジュールが、配列番号300~329の少なくとも1つまたはこれらの任意の組み合わせを含むか、これらをコードする、実施形態14および21~25のいずれか1つに記載のシステム。

【0491】

実施形態27. 前記少なくとも1つのGIC: ペイロードモジュールが、少なくとも1つの導入遺伝子配列を含むか、これをコードし、前記導入遺伝子の少なくとも1つのプロモーター配列、前記導入遺伝子の少なくとも1つの5'非翻訳配列、前記導入遺伝子の少なくとも1つの3'非翻訳配列、前記導入遺伝子の少なくとも1つのポリアデニル化シグナル配列、前記導入遺伝子の少なくとも1つの非コードRNA(ncRNA)プロセッシング配列、またはこれらの任意の組み合わせを含んでいてもよく、これらをコードしてもよい、実施形態14に記載のシステム。 40

【0492】

実施形態28. 前記少なくとも1つの導入遺伝子配列が、対象のゲノムへの挿入用の少なくとも1つの目的の配列を含むか、これをコードする、実施形態27に記載のシステム。

【0493】

実施形態29. 前記導入遺伝子の少なくとも1つのプロモーター配列が、対象のゲノム 50

において導入遺伝子の発現を促進する少なくとも1つの配列を含むか、これをコードする、実施形態27に記載のシステム。

【0494】

実施形態30. 前記導入遺伝子の少なくとも1つの5'非翻訳配列を含み、この少なくとも1つの5'非翻訳配列が、前記導入遺伝子のmRNAの少なくとも1つの5'非翻訳領域を含むか、これをコードする、実施形態27に記載のシステム。

【0495】

実施形態31. 前記導入遺伝子の少なくとも1つの3'非翻訳配列が、前記導入遺伝子のmRNAの少なくとも1つの3'非翻訳領域を含むか、これをコードする、実施形態27に記載のシステム。

【0496】

実施形態32. 前記導入遺伝子の少なくとも1つのポリアデニル化シグナル配列が、前記導入遺伝子の少なくとも1つのポリアデニル化シグナルを含むか、これをコードする、実施形態27に記載のシステム。

【0497】

実施形態33. 前記導入遺伝子の少なくとも1つの非コードRNA(ncRNA)プロセッシング配列が、前記導入遺伝子から発現された少なくとも1つのncRNAの少なくとも1つの終止シグナル、少なくとも1つの3'プロセッシングシグナル、およびこれらの任意の組み合わせを含むか、これらをコードする、実施形態27に記載のシステム。

【0498】

実施形態34. 前記少なくとも1つのGIC:ペイロードモジュールが、配列番号499~525の少なくとも1つまたはこれらの任意の組み合わせを含むか、これらをコードする、実施形態14および27~33のいずれか1つに記載のシステム。

【0499】

実施形態35. 前記少なくとも1つのGIC:5'側モジュールおよび前記少なくとも1つのGIC:3'側モジュールの少なくとも一方が、他方が由来する非長鎖末端反復型レトロエレメントの由来生物種とは異なる生物種に由来する少なくとも1つの配列を含むか、この配列をコードする、実施形態13~34のいずれか1つに記載のシステム。

【0500】

実施形態36. 前記少なくとも1つの遺伝子挿入コンストラクトが、図6~9に示す少なくとも1つの構造およびこれらの任意の組み合わせを含むか、これらの構造をコードする、実施形態1および13~35のいずれか1つに記載のシステム。

【0501】

実施形態37.

(i) 配列番号1~57の少なくとも1つに含まれるか、これらのうちの少なくとも1つによってコードされる少なくとも1つの逆転写酵素コンストラクトと、

(ii) 配列番号60~154、250~276、278~279、280~289、300~329、400~404、405~407、411~422および499~536の少なくとも1つの配列に含まれるか、これらのうちの少なくとも1つの配列によってコードされる少なくとも1つの遺伝子挿入コンストラクト

を含む、実施形態1および13~36のいずれか1つに記載のシステム。

【0502】

実施形態38. 実施形態13~37に記載の遺伝子挿入コンストラクトの少なくとも1つを含むか、これをコードする、遺伝子挿入コンストラクトの合成用コンストラクト(GIC:合成用コンストラクト)を含む、実施形態1および13~37のいずれか1つに記載のシステム。

【0503】

実施形態39. 前記少なくとも1つの逆転写酵素コンストラクトおよび前記少なくとも1つの遺伝子挿入コンストラクトの少なくとも一方が、他方が由来するレトロエレメントの由来生物種とは異なる生物種に由来する少なくとも1つの配列を含むか、この配列をコ

10

20

30

40

50

ードする、実施形態1～38のいずれか1つに記載のシステム。

【0504】

実施形態40. (i) 実施形態2～12に記載の少なくとも1つの逆転写酵素コンストラクトと、(ii) 実施形態13～37に記載の少なくとも1つの遺伝子挿入コンストラクトの組み合わせのうちの少なくとも1つを含む、実施形態1～39のいずれか1つに記載のシステム。

【0505】

実施形態41. 対象のゲノムに少なくとも1つの導入遺伝子を挿入する方法であって、実施形態1～40に記載の遺伝子挿入システム(GIS)の少なくとも1つを有効量で投与する工程を含む、方法。

【0506】

実施形態42. 前記導入遺伝子が、前記対象のゲノムの1つ以上の標的部位に挿入され、該1つ以上の標的部位が、少なくとも1つのセーフハーバー部位を含んでいてもよい、実施形態41に記載の方法。

【0507】

実施形態43. 任意の構成要素である前記少なくとも1つのセーフハーバー部位が、少なくとも1つのリボソームDNA(rDNA)配列を含み、該少なくとも1つのリボソームDNA配列が、少なくとも1つの28S rDNA配列を含んでいてもよい、実施形態42に記載の方法。

【0508】

実施形態44. 少なくとも1つの送達剤で製剤化された前記遺伝子挿入システムの少なくとも1つを投与する工程を含む、実施形態40～43のいずれか1つに記載の方法。

【0509】

実施形態45. 前記少なくとも1つの送達剤が、少なくとも1つのナノ粒子であり、該少なくとも1つのナノ粒子が、少なくとも1つの脂質ナノ粒子を含んでいてもよい、実施形態44に記載の方法。

【0510】

実施形態46. 医薬組成物であって、実施形態1～40に記載の遺伝子挿入システムの少なくとも1つを含み、少なくとも1つの添加剤、少なくとも1つの送達剤、少なくとも1つの補助剤、およびこれらの任意の組み合わせのうちの少なくとも1つを含んでいてもよい、医薬組成物。

【0511】

実施形態47. 治療適応症の治療を必要とする対象において、治療適応症を治療する方法であって、実施形態1～40に記載の遺伝子挿入システムの少なくとも1つ、または実施形態46に記載の医薬組成物の少なくとも1つを有効量で投与する工程を含み、実施形態41～45に記載の方法の少なくとも1つを含んでいてもよい、方法。

【0512】

実施形態48. 前記治療適応症が、テロメラーゼ活性の欠失によって引き起こされたものである、実施形態47に記載の方法。

【0513】

実施形態49. 前記少なくとも1つの遺伝子挿入システムが、少なくとも1つのTERT導入遺伝子を含む、実施形態46または47に記載の方法。

【0514】

実施形態50. 遺伝子挿入システムの作製用キットであって、実施形態1～40に記載の遺伝子挿入システムの方法を含み、実施形態46に記載の医薬組成物を含んでいてもよく、バッファー、DNAプラスミド、または前記遺伝子挿入システムもしくは前記医薬組成物を作製するためのプロトコルをさらに含んでいてもよい、キット。

【0515】

VIII. 定義

28S rDNA: 本明細書において、「28S rDNA」という用語は、真核生物の細胞質リ

10

20

30

40

50

ボソームの大サブユニット (LSU) を構成する大型リボソームRNA (rRNA) をコードする対象のゲノムの一部を意味する。

【0516】

3' 連結部：本明細書において、「3' 連結部」という用語は、挿入される配列の3' 末端が、対象のゲノムの5' 末端と連結する位置を意味する。

【0517】

3' 領域：本明細書において、「3' 領域」という用語は、レトロエレメント遺伝子において、オープンリーディングフレームの3' 側に位置する部分を意味する。

【0518】

5' 連結部：本明細書において、「5' 連結部」という用語は、対象のゲノムの3' 末端が 10、挿入される配列の3' 末端と連結する位置を意味する。

【0519】

5' 領域：本明細書において、「5' 領域」という用語は、レトロエレメント遺伝子において、オープンリーディングフレームの5' 側に位置する部分を意味する。

【0520】

活性：本明細書において、「活性」という用語は、何らかの事象が起こっている状態または何らかの事象が行われている状態を意味する。本開示のタンパク質および核酸は活性を有していてもよく、この活性は1つ以上の生物学的事象に関与していてもよい。

【0521】

調整が行われた：本明細書において、「調整が行われた」という用語は、その特性および／または活性を変化させたり、付加したり、除去するための、タンパク質配列またはアミノ酸配列の変化を意味する。 20

【0522】

アッセイ：本明細書において、「アッセイ」という用語を動詞として使用する場合、この用語は最も広い意味で使用され、当技術分野で公知の適切な方法を用いて試験を行う行為を意味する。本明細書において、「アッセイ」という用語を名詞として使用する場合、この用語は、アッセイの対象の特性、状態および／または活性を測定するために使用される試験を意味する。

【0523】

生物学的特性：本明細書において、「生物学的特性」および「特性」という用語は、測定可能または観察可能な、生物、生理系、臓器、組織、細胞または分子の特徴または活性を意味する。 30

【0524】

カーゴ：送達媒体に関連する文脈では、「カーゴ」および「ペイロード」という用語は、通常、対象の細胞、組織、臓器または生理系への送達、これらに対する送達、これらの近傍への送達を意図した化合物または構造（例えば、本発明の遺伝子挿入システム）を意味する。

【0525】

細胞：本明細書において、「細胞」という用語は、可能な限り広い意味を有し、生きた膜結合構造を意味する。 40

【0526】

細胞プロセス：本明細書において、「細胞プロセス」という用語およびこれと文法的に同等な用語は、細胞レベルで行われるプロセスを意味し、ここで細胞レベルとは、1個の細胞に限定されていてもよく、1個の細胞に限定されていなくてもよい。

【0527】

特徴：本明細書において、「特徴」という用語は、通常、ヒト、場所または物に属する特性または品質であり、これらを同定する役割を果たすものを意味する。「特徴」および「特性」という用語は、同じ意味を有し、交換して使用してもよい。

【0528】

付与する：本明細書において、「付与する」という用語およびこれと文法的に同等な用 50

語は、対象に特徴を付加するプロセスを意味する。

【0529】

コンストラクト：本明細書において、「コンストラクト」という名詞は、人工的に設計された生体高分子を意味する。生体高分子の例として、DNA、RNAおよびポリペプチドが挙げられる。通常、本明細書に記載のコンストラクトは、遺伝子挿入システムでの使用を目的として設計されている。

【0530】

分解：本明細書において、「分解」は、時間の経過に伴う組成物の機能喪失を意味する。

【0531】

送達：本明細書において、「送達」という用語は、生きた細胞または生きた生物に、化合物、物質、物体、部分、カーゴまたはペイロードを送達する行為または方法を意味する。別段の指定がない限り、「送達」という用語および「生物学的な送達」という用語は、同じ意味で使用してもよい。

【0532】

送達システム：本明細書において、「送達システム」という用語は、本発明の遺伝子挿入システムと組み合わせて製剤として調製した場合に、標的細胞の細胞質内に遺伝子挿入システムの各構成要素を送達することができる組成物、方法またはこれらの組み合わせを意味する。送達システムの例として、送達媒体で構成されるシステムおよび直接トランスフェクション用システムが挙げられるが、これらに限定されない。

【0533】

由来する：本明細書において、「由来する」という用語は、特定の供給源から単離または入手された、非長鎖末端反復型（非LTR型）レトロトランスポゾンなどの核酸配列またはタンパク質配列を意味する。この用語には、特定の供給源から単離または入手された天然配列が含まれる。さらに、この用語は、同じ機能特性または類似した機能特性を有する供給源から派生した人工バリエーション配列を含み、例えば、このバリエーションは、元の供給源に由来する分子と比較して機能特性が改善するように改変された核酸配列またはアミノ酸配列を含んでいてもよい。

【0534】

設計された：本明細書において、「設計された」という用語は、新規かつ所望の特性および／または活性を有するように、天然の状態または現行の状態に変化を加えた組成物を意味する。

【0535】

DNAおよびRNA：本明細書において、「RNA」、「RNA分子」または「リボ核酸分子」という用語は、リボヌクレオチドのポリマーを意味する。本明細書において、「DNA」、「DNA分子」または「デオキシリボ核酸分子」という用語は、デオキシリボヌクレオチドのポリマーを意味する。DNAおよびRNAは、天然に合成することができる。例えば、DNAは、DNAの複製により天然に合成することができ、RNAは、DNAの転写により天然に合成することができる。また、DNAおよびRNAは、化学的に合成することもできる。DNAおよびRNAは、一本鎖であってもよく（すなわち、ssRNAまたはssDNAであってもよく）、複数鎖であってもよい（例えば、二本鎖、すなわち、dsRNAまたはdsDNAであってもよい）。本明細書において、「mRNA」または「メッセンジャーRNA」という用語は、1本以上のポリペプチド鎖のアミノ酸配列をコードする一本鎖RNAを意味する。デオキシリボヌクレオチドを使用してRNA配列が記載されている場合、チミジン（「T」）をウリジン（「U」）またはウリジン類似体と置換して、DNA配列をRNA配列に変換することができる。

【0536】

DNA修復：本明細書において、「DNA修復」という用語は、細胞内のゲノムに発生した損傷を修正するために、細胞内で実行される内在性プロセスを意味する。

【0537】

10

20

30

40

50

効率的：本明細書において、導入遺伝子の挿入に関連する「効率的」という用語およびこれと文法的に同等な用語は、所望の標的部位にペイロードモジュールの全長を挿入する際の、RTタンパク質とGIC：5'側モジュールとGIC：3'側モジュールの任意の組み合わせの有効性を意味する。

【0538】

エレメント：本明細書において、「エレメント」という用語は、分子、システムまたは方法の一工程の個々の構成要素を意味する。

【0539】

発現産物：本明細書において、「発現産物」という用語は、目的の配列（例えば、mRNA）から転写されたRNA、または目的の配列から転写されたmRNAから翻訳されたポリペプチドを意味する。

10

【0540】

封入する：本明細書において、「封入する」という用語は、内包させること、包囲すること、または格納することを意味する。

【0541】

コードする：本明細書において、「コードする」という用語は、ポリマー高分子（第1の分子）に書き込まれた情報を用いて、この第1の分子とは異なる第2の分子の産生を誘導するプロセスを広く意味する。第2の分子は、第1の分子とは異なる化学的性質を有する化学構造を有していてもよい。

【0542】

20

エンドヌクレアーゼ：本明細書において、「エンドヌクレアーゼ」という用語は、両末端以外のヌクレオチドを分解することによりポリヌクレオチド鎖を切断するタンパク質またはその一部を意味する。

【0543】

エクソソーム：本明細書において、「エクソソーム」は、哺乳動物細胞により分泌される小胞またはRNAの分解に関与する複合体である。

【0544】

エキスピボ：「エキスピボ」という用語は、ドナー対象から細胞を取り出し、本明細書に記載の方法を使用して細胞を改変し、レシピエント対象に細胞を移入することを意味する。この用語には、1つの個体対象から得られた自家細胞（すなわち、この対象は、非改変細胞のドナーであり、かつエキスピボで改変した細胞のレシピエントでもある）、およびレシピエント対象とは異なる個体であるドナー対象から得られた同種細胞が含まれる。同種ドナーおよびレシピエントはHLAがマッチしていてもよい。

30

【0545】

促進する：本明細書において、「促進する」という用語は、最も広い意味で使用され、特定のエレメントを付加することによって、何らかの作用またはプロセスを起しやすくすることを意味する。

【0546】

忠実度：本明細書において、「忠実度」という用語は、目的の遺伝子が対象のゲノムに挿入される際の精度を意味する。「高い忠実度」という用語は、ヌクレオチド同一性、配列長および標的部位の位置に関して比較的少ないエラーで目的の遺伝子が挿入されることに相当する。例えば、約5,000個のヌクレオチドを含む鋳型RNAがRTタンパク質によりコピーされて、ミスマッチ塩基対を生じることなくcDNAを産生することができる場合、この遺伝子挿入は高い忠実度を有する。導入遺伝子の挿入の目的によって異なるが、わずかな数のミスマッチが起こることはあっても、十分に高い忠実度により、機能的な導入遺伝子を産生することができる。

40

【0547】

隣接：本明細書において、「隣接」という用語は、1つのエレメントが、別のエレメントの5'側（5'隣接）または3'側（3'隣接）に位置することを意味する。隣接するエレメント同士は、互いに直接連結されていてもよく、隣接するエレメント同士の間に別のエレ

50

メントが存在していてもよい。

【0548】

製剤：本明細書において、「製剤」は、本明細書に記載の遺伝子挿入システムの少なくとも1つの構成要素と、少なくとも1つの送達剤もしくは薬学的に許容される添加剤またはその両方を含む。

【0549】

機能的／活性型：本明細書において、生体分子に関連して使用される「機能的」という用語は、それを特徴付ける特性および／または活性を示す形態の生体分子を意味する。

【0550】

遺伝子：本明細書において、「遺伝子」という用語は、最も広い意味で使用され、染色体の一部を形成する区別可能なヌクレオチド配列、または染色体の一部を形成してもよい区別可能なヌクレオチド配列を意味し、その順序によって、ポリペプチド内または核酸分子内におけるモノマーの順序が決まる。

10

【0551】

遺伝子挿入コンストラクト：本明細書において、「遺伝子挿入コンストラクト」すなわち「GIC」という用語は、逆転写酵素タンパク質用のRNA鋳型を含むRNAコンストラクトを意味する。

【0552】

遺伝子挿入システム：本明細書において、「遺伝子挿入システム」すなわち「GIS」という用語は、TPRTなどの逆転写を介して、対象のゲノムの特定の位置への遺伝子配列（導入遺伝子）の挿入に使用してもよい各構成要素（モジュール）からなるシステムである。

20

【0553】

GIC:3'側モジュール：本明細書において、「3'側モジュール」という用語は、レトロエレメント遺伝子の3'領域に由来する少なくとも1つのエレメントまたはレトロエレメント遺伝子の3'領域の機能を置換する少なくとも1つのエレメントを含む、遺伝子挿入コンストラクト（GIC）の一部を意味する。

【0554】

GIC:5'側モジュール：本明細書において、「5'側モジュール」という用語は、導入遺伝子の全長の挿入を促進する、遺伝子挿入コンストラクト（GIC）の一部を意味し、レトロエレメント遺伝子の5'領域に由来していてもよく、これに由来していなくてもよい。

30

【0555】

生成する：本明細書において、「生成する」という動詞およびその活用形は、最も広い意味で使用され、特定の産物を得るためのプロセスを意味する。

【0556】

ゲノム：本明細書において、「ゲノム」という用語は、最も広い意味で使用され、細胞内に存在するあらゆる遺伝物質を意味する。

【0557】

デルタ肝炎ウイルス（HDV）リボザイム（RZ）折り畳み構造：本明細書において、「HDV RZ折り畳み構造」という用語は、デルタ肝炎ウイルス（HDV）のリボザイムの折り畳み構造を取ることができ、リボザイムの機能を保持するRNA配列を意味する。

40

【0558】

異種：本明細書において、「異種」という用語は、ある細胞で通常作られない遺伝子またはタンパク質の配列または構造をその細胞に導入する場合に、この導入される遺伝子またはタンパク質の配列または構造を意味する用語である。さらに、この用語は、異なる生物種に由来する核酸配列（DNAもしくはRNA）またはアミノ酸配列を含む、本開示の逆転写酵素コンストラクトまたは遺伝子挿入コンストラクトの個々のエレメント、モジュールまたは一部分を含む。例えば、逆転写酵素コンストラクトまたは遺伝子挿入コンストラクトの5'側モジュールは、鳥類の1種（または第1の種）に由来する配列を含んでいてもよく、この逆転写酵素コンストラクトまたは遺伝子挿入コンストラクトの3'側モジュ

50

ールは、鳥類の別の種（または第2の種）に由来する配列を含んでいてもよい。

【0559】

相同組換え：本明細書において、「相同組換え」という用語は、導入遺伝子と対象のゲノムの間の配列相同性に依存した導入遺伝子挿入プロセスを意味する。

【0560】

インビトロ：本明細書において、「インビトロ」という用語は、生きた細胞または生きた生物の外部で行われる反応またはプロセスを意味する。

【0561】

インビボ：本明細書において、「インビボ」という用語は、生きた細胞または生きた生物の内部またはその表面で行われる反応またはプロセスを意味する。

【0562】

不活性型：本明細書において、生体分子に関連して使用される「不活性型」という用語は、それを特徴付ける特性および／または活性を示さない形態の生体分子を意味する。

【0563】

不活性成分：本明細書において、「不活性成分」という用語は、製剤に含まれる医薬組成物の有効成分の活性に寄与しない1つ以上の薬剤を意味する。いくつかの実施形態において、本発明の製剤に使用してもよい不活性成分のすべてが、米国食品医薬品局（FDA）による承認を受けたものであってもよく、そのうちのいくつかは、FDAによる承認を受けたものであってもよく、本発明の製剤に使用してもよい不活性成分のすべてが、FDAによる承認を受けていないものであってもよい。

【0564】

誘導する：本明細書において、「誘導する」という用語およびこれと文法的に同等な用語は、工程に対して特定の制限を加えることなく、特定の結果がもたらされるプロセスを意味する。

【0565】

導入する：本明細書において、「導入する」という用語は、遺伝物質（DNAであることが多い）を細胞に付加することを意味する。

【0566】

挿入する：本明細書において、「挿入する」という用語は、DNA配列にヌクレオチドを付加することを意味する。

【0567】

連結部：本明細書において、「連結部」という用語は、対象のゲノム内に挿入された導入遺伝子のcDNAと、挿入部位の対象のゲノムのDNAが連結されている位置を意味する。

【0568】

少なくとも1つ：本明細書において、「少なくとも1つ」という用語は、1つ、2つ、3つ、4つまたは5つ以上の改変された物体、例えば、本開示のコンストラクト、モジュールまたは配列を意味する。

【0569】

脂質ナノ粒子：本明細書において、「脂質ナノ粒子」すなわち「LNP」は、1種以上の脂質（例えば、カチオン性脂質、非カチオン性脂質、PEG修飾脂質）を含む送達媒体を意味する。

【0570】

リポソーム：本明細書において、「リポソーム」は、通常、脂質（例えば両親媒性脂質）からなる1層以上の二重層または球状二重層で構成される小胞を意味する。

【0571】

機能喪失：本明細書において、「機能喪失」という用語は、野生型遺伝子の機能が喪失した改変遺伝子産物をもたらす対象の遺伝子の変化を意味する。

【0572】

改変：本明細書において、「改変」は、分子の状態または構造に変化が加えられていることを意味する。分子は、様々な方法で化学的、構造的または機能的に改変されていても

10

20

30

40

50

よい。

【0573】

モジュールシステム：本明細書において、「モジュールシステム」は、互いに対して比較的自律的に機能し、強く相互作用する複数の部分からなる複数のセットに分割可能なシステムを意味する。

【0574】

モチーフ：本明細書において、「モチーフ」という用語は、認識可能な構造を有する生体高分子の配列を意味し、この認識可能な構造は、特有の化学機能または生体機能によって定義されていてもよく、特有の化学機能または生体機能によって定義されていなくてもよい。

10

【0575】

天然：本明細書において、「天然」という用語は、野生型または天然型の化合物、生体分子（例えば、タンパク質もしくは核酸）または組成物を意味する。

【0576】

非LTR型レトロエレメント逆転写酵素：本明細書において、「非LTR型レトロエレメント逆転写酵素（RT）」という用語は、非LTR型レトロエレメントに由来する逆転写活性を有するタンパク質を意味する。

【0577】

非LTR型レトロエレメント：本明細書において、「非LTR型レトロエレメント」という用語は、長鎖末端反復を含んでいない一群のレトロエレメント遺伝子（別名：レトロトランスポゾン）を意味する。

20

【0578】

外側：本明細書において、挿入部位に関連する「外側」という用語は、挿入部位の5'末端または3'末端から約60bpを超えて離れたゲノムの任意の部分の意味する。

【0579】

ペアになる逆転写酵素（RT）：本明細書において、「ペアになるRT」は、挿入ペイロードモジュールを含む少なくとも1つのモジュールと組み合わせて使用される逆転写酵素（RT）を意味する。モジュールは、ペアになるRTと同種であってもよく、これは、このモジュールに含まれるすべてのエレメントとRTが同じレトロエレメント遺伝子に由来することを意味する。モジュールは、ペアになるRTと異種であってもよく、これは、このモジュールに含まれる少なくとも1つのエレメントが、RTと同じレトロエレメント遺伝子に由来していないことを意味する。

30

【0580】

ペイロード：「ペイロード」という用語は、送達媒体に関連する文脈で使用される場合を除き、対象のゲノムに挿入することを意図した遺伝子挿入システム（GIS）に含まれる核酸（例えば、目的の遺伝子）の配列を意味してもよい。

【0581】

相同率：「相同率」または「相同性（%）」という用語は、2つの核酸配列間またはアミノ酸配列間で同一または同じ配列の量を意味する。本明細書で定義するように、「相同率」という用語は、「同一性の割合」または「配列同一性の割合」という用語と同じ意味で使用することができる。

40

【0582】

本明細書において、「同一性の割合」、「配列同一性の割合」または「相同率」は、比較ウィンドウ内の最適化されたアライメントにおいて2つの配列を比較することによって決定され、比較ウィンドウ内の配列の一部は、2つの配列間でのアライメントを最適化するために参照配列（付加や欠失を含まない）と比較して付加または欠失（すなわちギャップ）を有していてもよい。配列同一性の割合は、両方の配列において同一の核酸塩基またはアミノ酸残基が存在する位置の数を測定して、核酸塩基またはアミノ酸残基がマッチした位置の数を算出し、マッチした位置の数を比較ウィンドウ内の位置の総数で割り、得られた結果に100を掛けることにより、配列同一性の割合を算出することができる。

50

【0583】

2つ以上の核酸配列またはポリペプチド配列に関する文脈における「同一」、「同一性」または「相同性」という用語は、2つ以上の同じ配列または部分配列を意味する。後述の配列比較アルゴリズムのうちの1つを使用した測定によって、または手動アライメントと目視試験による測定によって、比較ウィンドウ内または所定の領域での対応が最大になるように比較またはアライメントした場合に、複数の配列において同じヌクレオチドまたはアミノ酸残基が所定の割合である場合、これらの配列は互いに「実質的に同一」である（例えば、所定の領域において、少なくとも20%、少なくとも25%、少なくとも30%、少なくとも35%、少なくとも40%、少なくとも45%、少なくとも50%、少なくとも55%、少なくとも60%、少なくとも65%、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、少なくとも99.5%、または少なくとも99.9%の同一性を有する場合など）。この定義は、試験配列の相補鎖にも適用される。したがって、別段の記載がない限り、本明細書において提供されるすべての核酸配列およびアミノ酸配列は、参照配列と実質的に同一の配列を含む。

10

【0584】

配列比較では、通常、1つの配列が参照配列として使用され、これに対して試験配列を比較する。配列比較アルゴリズムを使用する場合、試験配列と参照配列をコンピュータに入力し、必要に応じて部分配列の座標を指定し、配列アルゴリズムプログラムのパラメータを指定する。一般に、デフォルトのプログラムパラメータが使用され、別のパラメータを指定することもできる。次に、配列比較アルゴリズムにより、プログラムパラメータに基づいて、参照配列に対する試験配列の配列同一性または配列類似性の割合を計算する。

20

【0585】

配列同一性および配列類似性の決定に適したアルゴリズムとして、Altschulら (Nuc. Acids Res. 25:3389-402, 1977) に記載のBLASTアルゴリズム、およびAltschulら (J. Mol. Biol. 215:403-10, 1990) に記載のBLAST 2.0アルゴリズムがある。BLAST分析を行うためのソフトウェアは、米国国立生物工学情報センター (National Center for Biotechnology Information) (<http://www.ncbi.nlm.nih.gov/>) から公的に利用可能である。このアルゴリズムでは、まず、high scoring sequence pairs (HSP) が同定される。データベース配列上の同じ長さのワードとアライメントして、クエリー配列において長さWを有する短いワードを同定し、このワードが、ある程度の陽性値を示す閾値であるTスコアに一致するか、このTスコアを満たした場合にHSPとして報告する。「T」は近隣ワードスコア閾値 (Altschulら、前掲) と呼ばれる。最初に同定された近隣ワードヒットを検索開始用のシードとして使用し、これらのワードを含む長いHSPを同定する。ワードヒットを各配列に沿って両方向に拡張し、累積的なアライメントスコアが増大する限り拡張を続ける。ヌクレオチド配列に対する累積スコアは、M (マッチした1組の残基のリワードスコア; 常に>0) およびN (ミスマッチ残基のペナルティスコア; 常に<0) を使用して計算する。アミノ酸配列では、スコア行列を使用して累積スコアを計算する。ワードヒットの両方向への拡張は、累積アライメントスコアが最大値からXの値分だけ低下し始めたとき、負のスコアを有する1個以上の残基のアライメントが蓄積して累積スコアが0以下になったとき、または一方の配列の末端に到達したときに停止させる。BLASTアルゴリズムのパラメータであるW、TおよびXにより、アライメントの感度と速度が決まる。BLASTNプログラム (ヌクレオチド配列用) では、デフォルトパラメータとして、ワード長 (W) = 11、期待値 (E) = 10、M = 5、N = -4、および両鎖の比較を使用している。アミノ酸配列を比較する場合、BLASTPプログラムでは、デフォルトパラメータとして、ワード長 = 3、期待値 (E) = 10、BLOSUM62スコア行列 (Henikoff and Henikoff, Proc. Natl. Acad. Sci. USA 89:10915, 1989を参照されたい) を使用したアライメント、(B) = 50、期待値 (E) = 10、M = 5、N = -4を使用している。

30

40

50

【0586】

さらに、BLASTアルゴリズムでは、2つの配列間の類似性の統計分析を行う（例えば、Karlin and Altschul, Proc. Natl. Acad. Sci. USA 90:5873-87, 1993を参照されたい）。BLASTアルゴリズムにより提供される類似性の指標の1つとして、最小確率の合計（ $P(N)$ ）があり、これは、2つのヌクレオチド配列間または2つのアミノ酸配列間での一致が偶然に生じる確率の指標を提供する。例えば、試験核酸を参照核酸と比較した際の最小確率の合計が、約0.2未満である場合、一般に、約0.01未満である場合、より一般的には約0.001未満である場合に、核酸が参照配列に類似していると見なされる。

【0587】

ペプチド：本明細書において、「ペプチド」は、50アミノ酸長以下のアミノ酸鎖を意味し、その長さは、例えば、約5アミノ酸長、約10アミノ酸長、約15アミノ酸長、約20アミノ酸長、約25アミノ酸長、約30アミノ酸長、約35アミノ酸長、約40アミノ酸長、約45アミノ酸長または約50アミノ酸長である。

【0588】

医薬組成物：本明細書において、「医薬組成物」という用語は、少なくとも1つの有効成分を含み、さらに1つ以上の薬学的に許容される添加剤を含んでいてもよい組成物を意味する。

【0589】

ポリアデノシン：本明細書において、「ポリアデノシン」という用語は、任意の長さのアデノシンヌクレオチド配列を意味する。

【0590】

ポリアデノシンテール：本明細書において、「ポリアデノシンテール」または「ポリAテール」という用語は、約80ヌクレオチド長以上のアデノシンヌクレオチド配列を意味する。

【0591】

ポリアデノシントラクト：本明細書において、「ポリアデノシントラクト」、「ポリAトラクト」および「Aトラクト」という用語（いずれもPAと略される）は、同じものを示し、同じ意味で使用され、約1~50ヌクレオチド長のアデノシンヌクレオチド配列を意味する。

【0592】

プロモーター：本明細書において、「プロモーター」という用語は、タンパク質の結合を受けて転写を開始するDNA配列を意味する。

【0593】

プロタンパク質：本明細書において、「タンパク質前駆体」、「プロタンパク質」および「プロペプチド」という用語は、翻訳後修飾により活性な形態に変化することができる不活性タンパク質を意味する。

【0594】

保護する：本明細書において、「保護する」という用語およびこれと文法的に同等な用語は、生体高分子の全体またはその一部の分解を防ぐ組成物またはプロセスを意味する。

【0595】

タンパク質：本明細書において、「タンパク質」は、50アミノ酸長を超えるアミノ酸からなる生体高分子を意味する。本明細書において、タンパク質の例として、酵素、逆転写酵素およびエンドヌクレアーゼが挙げられるが、これらに限定されない。

【0596】

領域：本明細書において、「領域」という用語は、ヌクレオチド配列の一部またはアミノ酸配列の一部を意味する。領域は、その長さが不明であっても定義されていなくてもよく、このような場合、その領域が示す機能により規定されるか、または配列内のその他のエレメントに対する相対的な位置により規定される。

【0597】

10

20

30

40

50

レトロエレメント／レトロトランスポゾン：本明細書において、「レトロエレメント」および「レトロトランスポゾン」という用語は、同じ意味で使用され、RNA中間体を介して、自身のゲノム内の新たな場所で複製することができる一群の真核細胞遺伝子を意味する。

【0598】

逆転写酵素：本明細書において、「逆転写酵素」という用語は、RNA鋳型配列からcDNAを合成することができるタンパク質を意味する。

【0599】

逆転写酵素コンストラクト：本明細書において、「逆転写酵素コンストラクト（RTC）」という用語は、前述したように、少なくとも1つの逆転写酵素を含むか、少なくとも1つの逆転写酵素をコードする生体高分子コンストラクトを意味する。

10

【0600】

RTC：RTモジュール：本明細書において、「RTC：RTモジュール」または「逆転写酵素モジュール」という用語は、少なくとも1つの逆転写酵素を含むか、少なくとも1つの逆転写酵素コードする生体高分子コンストラクトを意味する。

【0601】

リボソームDNA：本明細書において、「リボソームDNA（rDNA）」という用語は、RNAP Iによって合成されるリボソームRNA前駆体をコードする対象のゲノムの一部を意味する。

【0602】

リボソームRNA：本明細書において、「リボソームRNA（rRNA）」という用語は、リボソームの構成要素である非コードRNAを意味する。

20

【0603】

セグメント：本明細書において、「セグメント」という用語は、配列の一部を意味する。例えば、ヌクレオチド配列のセグメントは、遺伝子の全長よりも短い、遺伝子の任意の部分を含んでいてもよい。

【0604】

選択的：本明細書において、「選択的」および「選択性」という用語は、酵素、酵素タンパク質、遺伝子などの分子（ただしこれらに限定されない）が、非常に限られた種類、構造、タンパク質配列または遺伝子配列の別の分子に結合する傾向を意味する。

30

【0605】

自己切断型リボザイム：本明細書において、「自己切断型リボザイム」という用語は、配列特異的な分子内切断（または分子間切断）を触媒する一群のRNAを意味する。

【0606】

選択性：本明細書において、「選択性」は、逆転写酵素が、自身とペアになった異種由来のGIC：5'側モジュールまたはGIC：3'側モジュールを効率的に利用できる度合いを意味する。

【0607】

配列：本明細書において、「配列」という用語は、生体高分子のN末端からC末端の方向に記載されたアミノ酸の順序、または生体高分子の5'末端から3'末端の方向に記載されたヌクレオチドの順序を意味する。

40

【0608】

部位特異的：本明細書において、「部位特異的」という用語は、座位、例えば、約60bpの配列からなる座位を意味する。

【0609】

安定性：本明細書において、「安定性」という用語は、組成物がその特性を経時的に保持する能力を意味する。

【0610】

TPRTが行われた：本明細書において、「TPRTが行われた」という用語は、標的部位にニックが導入されることにより生じたプライマーを利用して、cDNAの合成および／ま

50

たは導入遺伝子の挿入が行われたことを意味する。

【0611】

適切な：本明細書において、「適切な」という用語は、特定の目的または使用に対して効果があるか、有効であるか、または適合性があることを意味する。

【0612】

合成の：本明細書において、「合成の」という用語は、人工的に製造、調製および／または生産されたものを意味する。本発明のポリヌクレオチドもしくはポリペプチドまたはその他の分子の合成は、化学合成であってもよく、酵素合成であってもよい。

【0613】

標的細胞：本明細書において、「標的細胞」という用語は、1個以上の目的の細胞を意味する。標的細胞は、インビトロ、インビボもしくはインサイチューで、または生物の組織もしくは臓器に存在してもよい。生物は動物であってもよく、哺乳動物であることが好ましく、ヒトであることがより好ましく、患者であることが最も好ましい。

10

【0614】

Target Primed Reverse Transcription (TPRT)：本明細書において、「Target Primed Reverse Transcription (TPRT)」という用語は、標的部位において、ニックが生じたゲノムDNAの3'末端を逆転写酵素がプライマーとして利用することでcDNAの合成が開始されるプロセスを意味する。

【0615】

鋳型：本明細書において、「鋳型」および「RNA鋳型」という用語は、逆転写酵素によってcDNAに転写されるRNA配列を意味する。

20

【0616】

鋳型末端：本明細書において、「鋳型末端」という用語は、RNA鋳型の5'末端または3'末端を意味する。

【0617】

治療上有効な：本明細書において、「治療上有効な」という用語は、対象における治療適応症を治療または緩和することができる遺伝子または遺伝子産物を意味する。

【0618】

転写：本明細書において、「転写」という用語は、DNA分子を鋳型として用いたRNAポリメラーゼによるRNA分子の形成または合成を意味する。

30

【0619】

トランスフェクション：本明細書において、「トランスフェクション」という用語は、細胞に外因性核酸を導入する方法を意味する。トランスフェクション方法として、化学的方法、物理的处理、およびカチオン性脂質またはカチオン性混合物が挙げられるが、これらに限定されない。

【0620】

導入遺伝子：本明細書において、「導入遺伝子」という用語は、対象のゲノムに挿入される遺伝子を意味する。

【0621】

翻訳：本明細書において、「翻訳」という用語は、RNA鋳型に基づいてリボソームによりポリペプチド分子が形成されることを意味する。

40

【0622】

治療および予防：本明細書において、「治療する」または「予防する」という用語およびこれらから派生した用語は、必ずしも100%の治療もしくは予防または完全な治療もしくは予防を必要としない。実際、当業者が有用性または治療効果が見込まれると認める治療または予防の程度は様々である。さらに、「予防」は、疾患、その症状または状態の発症を遅延させることを含んでいてもよい。

【0623】

非改変：本明細書において、「非改変」という用語は、何らかの方法で変化が加えられる前の物質、化合物または分子を意味する。「非改変」は、生体分子の野生型または天然

50

型を意味する場合もあるが、必ずしもこれらに限定されるものではない。分子は、複数の改変を受ける場合があり、その過程では、改変分子が、次の改変で「非改変」出発分子として使用されることもある。

【0624】

ベクター：本明細書において、「ベクター」という用語は、異種分子を輸送するか、異種分子を形質導入するか、または異種分子の担体として作用する分子または部分を意味する。

【0625】

IX. 同等物および範囲

当業者であれば、本明細書に記載された本開示の具体的な実施形態に対する等価物が多数存在することは理解できるであろう。また、そのことを通常の実験により確認することができるであろう。本発明の範囲は、明細書内の上記の説明に限定されることを意図するものではなく、添付の請求項に規定される通りである。

【0626】

特許請求の範囲において、「a」、「an」および「the」などの冠詞は、別段の指定がない限り、または文脈から明らかな場合を除いて、1つ以上を意味する場合がある。グループ内の1つ以上の構成要素の間に「または」が含まれる請求項または説明は、別段の指定がない限り、または文脈から明らかな場合を除いて、グループ内の1つの構成要素、複数の構成要素またはすべての構成要素が、所定の製品もしくはプロセスに存在するか、使用されるか、または関連するとみなされる。本開示は、グループ内の1つの構成要素が、所定の製品またはプロセスに存在するか、使用されるか、または関連する実施形態を含む。本開示は、グループ内の複数の構成要素またはすべての構成要素が、所定の製品またはプロセスに存在するか、使用されるか、または関連する実施形態を含む。

【0627】

また、「含む (comprising)」という用語は、非限定的であることが意図されており、追加の要素または工程が含まれることを容認するものであるが、追加の要素または工程が含まれることを要求するものではないことに留意されたい。本明細書において、「含む (comprising)」という用語が用いられる場合、「からなる (consisting of)」もまた包含され開示される。

【0628】

範囲が指定されている場合は、端点も含まれる。さらに、別段の指定がない限り、または文脈および当業者の理解に照らして明らかでない限り、範囲として表される値は、本開示の様々な実施形態で記載された範囲内の任意の特定の値または部分範囲が想定され、文脈上別段の明確な指定がない限り、範囲の下限値の10分の1の単位まで想定されることは理解されるであろう。

【0629】

また、従来技術に該当する本発明の特定の実施形態は、請求項のうちのいずれか1つ以上から明示的に除外される場合があることは理解されるであろう。このような実施形態は、当業者にとって公知であるとみなされるため、本明細書において明示的に除外を定めていない場合であっても、除外される場合がある。また、本開示の組成物の任意の特定の実施形態（例えば、任意の抗生物質、治療薬または活性成分；任意の製造方法；任意の使用

方法；など）が、従来技術に存在するものとの関連の有無にかかわらず、任意の理由で、任意の1つ以上の請求項から除外される場合がある。

【0630】

使用されている文言は、限定するためのものではなく説明するためのものであり、広い意味での本開示の真の範囲および趣旨から逸脱しなければ、請求項の範囲内で変更が可能であることは理解されるであろう。

【0631】

本発明について、いくつかの記載された実施形態に関して、ある程度の長さで、ある程度の具体性をもって記載してきたが、本開示は、そのような具体的事項もしくは実施形態

10

20

30

40

50

または任意の特定の実施形態に限定されることを意図するものではなく、添付の請求項を参照し、従来技術に鑑みて、請求項の可能な限り広範な解釈を提供するものとして、つまり、本開示の意図された範囲を効果的に包含するものとして解釈されるものである。

【0632】

本発明を以下の実施例によりさらに説明するが、本発明はこれらの実施例に限定されない。

【実施例】

【0633】

X. 実施例

実施例1. 遺伝子挿入コンストラクト (GIC) のインビトロ転写 (IVT) – ペイロードなしのGIC

例えば、精製された逆転写酵素 (RT) を使用したインビトロでのTPRTアッセイに使用するためのRNAとしての、約1000nt未満の遺伝子挿入コンストラクト (GIC) のRNA生体高分子は、通常、インビトロRNA転写 (IVT) 反応を利用して以下のように調製される。

【0634】

RNA転写用のGIC DNA鋳型は、Q5 DNAポリメラーゼ (NEB社) を用いたPCRにより作製し、カラム精製 (Bio Basic社) により精製する。

【0635】

インビトロ転写 (IVT) 反応は、T7 RNAポリメラーゼ (RNAP) を使用して、同等なRNA精製物が得られる以下の2つのプロトコルのいずれかにより行った。精製された反応成分を使用する第1の方法では、40mM Tris (pH7.9)、2.5mM スペルミジン、26mM MgCl₂、0.01% Triton X-100、約30mM DTT、8mM GTP、4mM その他のrNTP、0.5 μLのRiboLock (サーモサイエンティフィック社)、0.5 μLの無機ピロホスファターゼ (NEB社) および0.5 μLのT7 RNAP (細菌で過剰発現させた後に精製し、20mM KPO₄ (pH7.5)、100mM NaCl、50%グリセロール、10mM DTT、0.1mM EDTAおよび0.2% NaN₃を含む溶液中に50mg/mLで保存したもの) を含む25 μLの反応溶液中において1 μgのDNA鋳型を転写する。反応液を37℃で3~4時間インキュベートした後、1 μLのDNase RQ1 (プロメガ社)、1.5 μLの20mM CaCl₂および2 μLのH₂Oを加える。第2の方法では、反応溶液20 μlあたり1 μgの消化したプラスミドを使用して、NEB社製HiScribe T7キットを製造業者の説明書に従って使用する。反応液を37℃で2時間インキュベートした後、1 μLのDNase RQ1 (プロメガ社)、1.5 μLの20mM CaCl₂および2 μLのH₂Oを加える。

【0636】

次に、当技術分野で公知の一般的な手順に従って、脱塩 (ロシュ社製ミニクイックスピンカラム)、有機溶媒による抽出および沈殿によりRNA産物を精製する。

【0637】

実施例2. 遺伝子挿入コンストラクト (GIC) のインビトロ転写 (IVT) – ペイロードとして導入遺伝子を有するGIC

インビトロRNA転写 (IVT) 反応により、ペイロードとして導入遺伝子発現カセットを含むGIC RNA生体高分子を以下のように調製する。

【0638】

GIC DNA転写用鋳型配列をpUC57-miniバックボーン (配列番号350) にクローニングして、目的のGIC RNA鋳型の上流にT7 RNAPプロモーターを配置し、このGIC RNA鋳型の下流にBbsI部位を配置する。精製したプラスミドDNAを、BbsI-HF (NEB社) を用いて37℃で4時間消化して線状化する。次に、消化されたプラスミドを、キアゲン社製PCR精製カラムで精製し、ヌクレアーゼフリー水で溶出する。

【0639】

インビトロ転写反応は、反応溶液20 μlあたり1 μgの消化したプラスミドを用いて、NEB社製HiScribe T7キットを使用して行う。具体的には、各インビトロ転写反応液には

、2 μ lの各rNTP、2 μ lの10×バッファー、2 μ lのT7ポリメラーゼミックス、1 μ gの消化したプラスミドおよびddH₂Oが含まれ、37℃で2時間インキュベートする。

【0640】

インビトロ転写後、RnaseフリーDnase Iを用いて37℃で30分間処理することによってDNA鋳型を除去する。次に、25：24：1のフェノール：クロロホルム：イソアミルアルコール（pH 6.7）（PCI社）を等量加え、激しくボルテックスし、遠心分離し、水層を回収して、10体積%の3M酢酸ナトリウム（pH 5）と3体積%の100%エタノールで沈殿させることにより、合成したRNAを精製する。70%エタノールで3回洗浄後、RNAペレットを風乾し、1mMのクエン酸ナトリウム（pH6.5）中に溶解する。

【0641】

実施例3. TPRTアッセイ用の逆転写酵素（RT）タンパク質の調製

RTタンパク質は、以下のようにヒト細胞において一過性発現させ、精製することにより作製した。

【0642】

目的の逆転写酵素（RT）をコードするコドン最適化ORF（GenScript社）を、pcDNA3.1 N-DYKプラスミド（GenScript社）（配列番号351）のKpn I部位とXba I部位の間にクローニングして、このベクターにコードされているN末端FLAGタグと融合させる。Kpn I部位は、FLAGタグとRTアミノ酸配列の間にグリシンスレオニンリンカーを付加する。Xba I部位は、3' UTRの開始位置付近の翻訳終止コドンの後ろに位置する。Lipofectamine 3000（インビトロジェン社）を使用して、12 μ gのプラスミドDNAのリバーストランスフェクションを行う。まず、500 μ LのOPTI-MEMと24 μ LのP3000をDNAと穏やかに混合する。次に、500 μ LのOPTI-MEMと24 μ LのLipofectamineを混合し、DNA混合物に加える。Lipofectamine/DNA複合体を室温で10分間インキュベートし、以下のように調製した細胞に加える。簡単に述べると、各トランスフェクションにおいて、10cmのディッシュに入れた80%コンフルエントのHEK 293T細胞（以下「293T」と呼ぶ）を小分けしてLipofectamine/DNA複合体に加え、80%コンフルエントになるように再播種する。18～24時間後に、細胞をトリプシン処理してプレートから回収し、5mLの培地中に再懸濁し、15mLのコニカルチューブ中で約2000×gで3分間スピンドウンする。1mM PMSFを含むPBSでペレットを洗浄し、1.5mLのチューブに移し、4℃、2000×gで1分間スピンドウンして再度ペレットを得る。

【0643】

細胞ペレットの4倍の体積量の1×低張溶解バッファー（HLB）[20mM HEPES（pH8）、2mM MgCl₂、200 μ M EGTA、10%グリセロール、1mM DTT、0.2%セリンプロテアーゼインヒビターカクテル（SPIC、シグマ社）、1mM PMSF] に細胞ペレットを懸濁し、氷上に5分間静置して細胞を膨張させる。次に、液体窒素中で試料をスナップ凍結し、室温のウォーターバスで解凍する操作を3サイクル行って細胞を溶解する。次に、400mM NaClに試料を入れ、穏やかにボルテックスし、氷上でさらに5分間静置する。次に、試料を4℃、17000×gで5分間遠心する。上清を回収し、0.2% NP-40を含む1×HLBバッファーを等量加えることによって、NaClの濃度を200mMに低下させ、NP-40の濃度を0.1%まで上昇させる。試料を穏やかにボルテックスし、4℃、17000×gで10分間遠心する。

【0644】

清澄化した上清を新しいチューブに回収し、ブロッキングおよび平衡化した20 μ LのFLAG抗体樹脂（シグマ社）を加える。試料を4℃で2時間回転振盪してタンパク質を免疫沈降させる。次に、FLAG樹脂を、IPバッファー（1×HLB、200mM NaCl、0.1% NP-40）で合計4回洗浄する（簡単な洗浄2回と、4℃、5分間での振盪洗浄2回）。最終洗浄の後、バッファー全量を30G針で除去し、樹脂を40 μ LのIPバッファー中に再懸濁する。50ng/ μ L 3×FLAGペプチド（シグマ社）を加えて室温で1時間インキュベートすることによってタンパク質を部分的に溶出させる。溶出したタンパク質は液体窒素中でフラッシュ凍結し、次の使用まで-80℃で保存する。

10

20

30

40

50

【0645】

実施例4. 逆転写酵素コンストラクト (RTC) mRNAの作製

RNA (mRNA) RTC生体高分子を以下のように調製する。

【0646】

逆転写酵素 (RT) をコードするコドン最適化ORF (GenScript社) をPCRで増幅して、ORFの前にBamHI部位を付加し、ORFの終結を指定する終止コドンの後ろにXhoI部位を付加する。BamHI部位は、N末端FLAGタグとRT ORFの間にインフレームに挿入され、その連結部にグリシンセリンリンカーが付加される。

【0647】

pUC57-mini (配列番号350) において、5' UTR配列 (配列番号58) と、3' UTRおよび鑄型にコードされたポリアデノシンテールからなる配列 (配列番号59) との間にRT ORFがクローニングされ、その上流にT7 RNAPプロモーター配列が配置され、その下流にBbsI部位が配置される。次に、実施例2と同様にして、BbsIでmRNA転写鑄型プラスミドを線状化し、再精製する。委託会社 (TriLink社) に依頼して、またはTriLink社の試薬とプロトコルを使用して、通常、ウリジンリボヌクレオチド三リン酸 (U) の100%を置換した5-メトキシウリジンリボヌクレオチド三リン酸 (5moU) を用いたAG Clean cap mRNA合成と、シリカ膜を用いた精製を行う。5moUによるウリジンの100%置換と、N¹-メチルシュードウリジンによる置換を比較すると、どちらの修飾ヌクレオチドを使用した場合でも、mRNAの機能は同程度になる。

【0648】

実施例5. インビトロ逆転写酵素活性スクリーニング

当技術分野において標準的なRTアッセイであるRNA鑄型にアニールしたDNAプライマーを使用して、候補タンパク質のインビトロ逆転写酵素活性を以下のように試験する。

【0649】

実施例3と同様にして、逆転写酵素 (RT) タンパク質を調製する。プライマーDNAオリゴ (配列番号352) はIDT社から購入し、鑄型RNA (配列番号353) は実施例1に記載の第1のプロトコルによって作製する。

【0650】

各スクリーニング反応を行うため、試料を65℃で3分間加熱することによって、8 μM DNAオリゴ2 μLと4 μM鑄型RNA 2 μLをアニールさせ、氷上で試料を少なくとも5分間静置する。

【0651】

10×RTバッファー2 μL (50mM MgCl₂、250mM Tris (pH7.5) および750mM KCl)、100mM DTT 2 μL、20%PEG-6K 2 μLならびにヌクレアーゼフリー水5 μLを含む非放射性マスターミックスを作製する。

【0652】

さらに、10mM dA、dCおよびdTTP各1 μL；2mM dGTP 1 μL；前述のアニールしたDNA-RNA 4 μL、ならびに³²P α-dGTP 1 μL (パーキンエルマー社) を含む放射性マスターミックスを作製する。

【0653】

各反応では、非放射性マスターミックス11 μLと、候補RTタンパク質2 μLと、放射性マスターミックス7 μLを混合し、各反応液量を20 μLとする。37℃で30分間反応を進行させた後、70℃で5分間熱失活させる。γ-³²P ATPで5'末端をあらかじめ放射標識した100ntのオリゴヌクレオチド (配列番号290) を含む停止溶液 (50mM Tris (pH 7.5)、20mM EDTAおよび0.2% SDS) 80 μLと、T4ポリヌクレオチドキナーゼ (NEB社) を反応液に加え、次に、PCIで抽出してDNAの精製と濃縮を行った後、エタノール沈澱 (ドライアイス-エタノール浴) を行う。卓上遠心分離機において14,000×gで20分間遠心分離してDNAをペレット化し、75%エタノールで1回洗浄し、風乾し、5 μL H₂O + 5 μL 2×ホルムアミドローディングバッファー中に再懸濁する。

【0654】

10

20

30

40

50

9%尿素-PAGE変性ゲル上で試料を泳動し、乾燥させ、phosphor imagerのスクリーン上で露光させ、翌日にTyphoon Trio Imagerシステムで撮影する。

【0655】

実施例6. インビトロTPRT活性アッセイ

実施例3と同様にして、逆転写酵素（RT）タンパク質を調製する。TPRT用の鋳型RNAは、実施例1で述べたインビトロ転写反応により調製する。下の鎖の5'末端を $\gamma^{32}\text{P}$ ATPで放射標識した64bpまたは84bpの長さの標的部オリゴヌクレオチド二本鎖DNA（それぞれ配列番号291および配列番号292）とT4ポリヌクレオチドキナーゼ（NEB社）を、マグネシウム反応バッファー中でRTタンパク質および鋳型RNAと混合し、37℃で30分間インキュベートする。得られた産物を変性PAGEで分離して、Typhoon Trio Imagerシステムでゲルを撮影する。

10

【0656】

実施例7. 細胞培養およびRNAベースの逆転写酵素コンストラクト（RTC）と遺伝子挿入コンストラクト（GIC）の共トランスフェクション

トランスフェクションを行う直前に、6ウェルプレートにおいて1ウェルあたり125万～250万個の細胞密度で本明細書に記載の哺乳動物細胞株を播種する。

【0657】

5 μl のMessenger Maxを125 μl のOpti-MEMで希釈し、10分間インキュベートする。

【0658】

20

RTC mRNAとGIC RNA（それぞれ実施例4と実施例2で調製したもの）を所定のモル比で混合した後、125 μl のOpti-MEMで希釈する。次に、Messenger MaxのOpti-MEM溶液と、GIC RNAのOpti-MEM溶液をよく混合し、室温で5分間インキュベートする。

【0659】

得られた混合物を、6ウェルプレートの各ウェルに播種した細胞に滴下し、プレートを細胞インキュベーターに戻し、十分な時間が経過してから細胞の分析を行う。

【0660】

実施例8. FACS分析

フローサイトメトリー分析

30

トランスフェクションの翌日に（別段の指定がない限り）、トリプシン処理により細胞を回収して5%FBS含有DMEM培地中に懸濁し、Attune NxTフローサイトメーター（Thermo社）または同等の機器で分析する。生きた単一細胞は、前方散乱および側方散乱でゲートをかける。Attune NxTフローサイトメーターでのmCherryチャンネルはYL2であり、561nmで励起され、発光フィルターは620/15nmである。Attune NxTフローサイトメーターでのeGFPチャンネルはBL1であり、488nmで励起され、発光フィルターは530/30nmである。フローサイトメトリーの結果はFlowJo 10.8.1を使用して分析する。RT mRNAを使用せずにGIC RNAのみでのトランスフェクションをバックグラウンドコントロールとして使用し、定量時に各シグナルからバックグラウンドを差し引く。

40

【0661】

セルソーティング

トランスフェクションの翌日に（別段の指定がない限り）、トリプシン処理により細胞を回収して5%FBS含有DMEM培地中に懸濁し、130 μm のチップを装着したultra-purityモードのソニー社製SH800または同等の機器を使用してソートする。ソートした細胞を遠心分離により回収し、PBSで洗浄する。

【0662】

実施例9. RNAベースの逆転写酵素コンストラクト（RTC）と遺伝子挿入コンストラクト（GIC）の組成

トランスフェクション用の逆転写酵素コンストラクト（RTC）のmRNAは、以下の表1

50

の記載に従って実施例4と同様の方法で作製する。

【0663】

【表2】

表1: 2種のRNA構成要素からなる遺伝子挿入システム(GIS)の逆転写酵素コンストラクト(RTC)

RTC 識別名	RTC: RT モジュールの供給源の生物	配列番号
F-ZoAI RT mRNA	ノドジロシトド (<i>Z. albicollis</i>)	19
F-TaGu RT mRNA	キンカチョウ (<i>T. guttata</i>)	28
F-TriCasB RT mRNA	コクヌストモドキ (<i>T. castaneum</i>)	3
OrLa-3F RT mRNA	メダカ (<i>O. latipes</i>)	10
ZoAI RT mRNA (タグなし)	ノドジロシトド (<i>Z. albicollis</i>)	21
ZoAI_catdead RT mRNA	ノドジロシトド (<i>Z. albicollis</i>)	23
TriCasB RT mRNA (タグなし)	コクヌストモドキ (<i>T. castaneum</i>)	5

10

【0664】

トランスフェクション用の遺伝子挿入コンストラクト (GIC) のRNAは、以下の表2に従って実施例1および実施例2と同様の方法で作製する。

【0665】

20

30

40

50

【表 3】

表 2: 2 種の RNA 構成要素からなる遺伝子挿入システム (GIS) の遺伝子挿入コンストラクト (GIC)

GIC 識別名	5'側 モジュール の供給源の 生物	導入遺伝子 のプロモー ター領域 & 5'UTR	導入遺伝子	導入遺伝 子の 3'UTR & ポリ A シグナル 領域	3'側 モジュール の供給源の 生物	GIC の配列 番号
TriCas_ZoAl	コクヌスト モドキ	CBh	NLSeGFP	SV40LPA	ノドジロシ トド	
TriCas_GeFo	コクヌスト モドキ	CBh	NLSeGFP	SV40LPA	ガラパゴス フィンチ	
TriCas_TaGu	コクヌスト モドキ	CBh	NLSeGFP	SV40LPA	キンカチョ ウ	
TriCasFlipZoAl	コクヌスト モドキ	CBh_Flip	GFP	SV40LPA	ノドジロシ トド	
TriCasBsiZoAl	コクヌスト モドキ	CBh_Bsi	GFP	SV40LPA	ノドジロシ トド	
TCA5_ZoAl	コクヌスト モドキ	CBh	GFP	SV40LPA	ノドジロシ トド	
TCA5_GeFo	コクヌスト モドキ	mPGK	GFP	SV40LPA	ガラパゴス フィンチ	
TCA5_TaGu	コクヌスト モドキ	CBh_Bsi	GFP	SV40LPA	キンカチョ ウ	
TCA5_TiGu	コクヌスト モドキ	CBh_Bsi	GFP	SV40LPA	ノドジロシ ギダチョウ	
TCARZ_GeFo	コクヌスト モドキ	CBh_Bsi	GFP	SV40LPA	ガラパゴス フィンチ	
TCARZ_Cherry_GeFo	コクヌスト モドキ	CBh_Bsi	mCherry	SV40LPA	ガラパゴス フィンチ	
HDVgu5_GeFo	なし	CBh_Bsi	GFP	SV40LPA	ガラパゴス フィンチ	
HDVgu5b_GeFo	なし	CBh_Bsi	GFP	SV40LPA	ガラパゴス フィンチ	
HDVgu5c_GeFo	なし	CBh_Bsi	GFP	SV40LPA	ガラパゴス フィンチ	
HDVgu5d_GeFo	なし	CBh_Bsi	GFP	SV40LPA	ガラパゴス フィンチ	
HDVac11_GeFo	なし	CBh_Bsi	GFP	SV40LPA	ガラパゴス フィンチ	
HDVac11b_GeFo	なし	CBh_Bsi	GFP	SV40LPA	ガラパゴス フィンチ	
HDVac12_GeFo	なし	CBh_Bsi	GFP	SV40LPA	ガラパゴス フィンチ	
HDVac12b_GeFo	なし	CBh_Bsi	GFP	SV40LPA	ガラパゴス フィンチ	
など						

【0666】

実施例 10. 逆転写活性用の候補タンパク質のスクリーニング

逆転写用にスクリーニングした候補 R2 ファミリーレトロエレメントタンパク質 (表 3 参照) を実施例 3 と同様の方法で調製し、実施例 5 と同様の方法で逆転写活性を試験した。いくつかの TPRT タンパク質または RT タンパク質は、アッセイの 1 つのサブセットのみにいて活性が検出された (「低／なし」として示す)。

【0667】

10

20

30

40

50

【表 4】

表 3: 逆転写酵素活性用候補タンパク質

配列番号	供給源の生物種	生物種の 参照コード	図 10 の レーン No.	逆転写酵素 活性
47	<i>Drosophila mercatorum</i>	DrMerc	15	なし
57	<i>Lepidurus couesii</i>	LeCoB	11	なし
55	ヨーロッパカブトエビ (<i>Triops cancriformis</i>)	TriCan	12	なし
43	カタユウレイボヤ (<i>Ciona intestinalis</i>)	Ciln	3	なし
51	イトヨ (<i>Gasterosteus aculeatus</i>)	GaAc	19	なし
49	キイロショウジョウバエ (<i>Drosophila melanogaster</i>)	DrMe	14	低/なし
45	アメリカカブトガニ (<i>Limulus polyphemus</i>)	LiPo	13	低/なし
53	キタノトミヨ (<i>Pungitis pungitis</i>)	PuPu	16	低
7	キョウソヤドリコバチ (<i>Nasonia vitripennis</i>) (B 系統)	NaviB	9	高
9	メダカ (<i>Oryzias latipes</i>)	OrLa	8	低
18	ノドジロシトド (<i>Zonotrichia albicollis</i>)	ZoAl	10	低/中程度
27	キンカチョウ (<i>Taeniopygia guttata</i>)	TaGu	18	高
2	コクヌストモドキ (<i>Tribolium castaneum</i>) (B 系統)	TriCasB	5	高
25	ノドジロシギダチョウ (<i>Tinamus guttatus</i>)	TiGu	17	低
33	オナジショウジョウバエ (<i>Drosophila simulans</i>)	DroSi	4	高
36	カイコガ (<i>Bombyx mori</i>)	BoMo	2	高
39	<i>Adineta vaga</i>	AdVa	7	中程度
41	チクビヒドラ (<i>Hydra magnipapillata</i>)	HyMa	6	なし
31	ガラパゴスフィンチ (<i>Geospiza fortis</i>)	GeFo	NS	低/なし

【0668】

逆転写酵素活性は、生物種間で大きく変動した。図10のPAGEゲルの画像に示された結果から分かるように、予想された長さの最初の逆転写産物が、候補RTタンパク質として評価したTriCasB、DroSi、TaGu、NaViB、BoMo、OrLa、AdVa（タンパク質発現に正規化した場合）、ZoAl、LiPo（様々な程度で検出可能な産物）、PuPuおよびTiGu、ならびにGeFo（様々な程度で検出された産物）の黒色の実線の囲み枠に認められる。Ciln、LeCoB、TriCan、DroMer、DroMe、HyMaおよびGaAcでは、再現可能なRT産物は検出されなかった。DrMeとGeFoでは、非常に低い活性が時折検出された。予想された産物長のバンドの不透明度と、RTタンパク質のFLAGエピトープタグに対する抗体を使用した免疫プロットにより検出された精製タンパク質の量とを組み合わせることによって、相対的な逆転写活性レベルを評価し、各候補タンパク質を、表3に示すように、高い逆転写活性を有するタンパク質、中程度の逆転写活性を有するタンパク質、低い逆転写活性を有するタンパク質、または逆転写活性がないタンパク質（使用したアッセイで検出不能）に分類した。全体的な結果として、候補タンパク質のうち、TriCasB、DroSi、TaGu、NaViBおよびBoMoは、最も高い逆転写酵素能を示したことから、本発明の逆転写酵素コンストラクト（RTC）に組み込むことが可能な候補タンパク質として極め

て有望である。

【0669】

実施例11. 3'側モジュールの特異性を確認するためのインビボRTアッセイ

カイコガ由来の改変RTタンパク質ORF（配列番号35）、オナジショウジョウバエ由来の改変RTタンパク質ORF（配列番号32）またはメダカ由来の改変RTタンパク質ORF（配列番号8）を発現するpcDNA3.1バックボーンプラスミドのうちの1つと、カイコガ（配列番号209）、オナジショウジョウバエ（配列番号210）またはメダカ（配列番号200）のR2因子の3' UTR RNAを発現する別のプラスミドとをそれぞれ組み合わせて、9つのHEK293T細胞集団にトランスフェクトした（図11A参照）。各RTタンパク質と、各3' UTR RNAを共発現させた。

10

【0670】

RTタンパク質プラスミドの転写および翻訳により生成されたRTタンパク質と、転写された3' UTR RNAが会合するのに十分な時間を置いた後、細胞を溶解し、RTタンパク質+RNA鑄型複合体を、FLAGを用いた免疫精製（シグマ社製FLAG抗体固定化樹脂）により精製した。免疫精製前の各細胞溶解液中に存在するRNAと、免疫精製後の各試料中の複合体を形成したRNAをそれぞれ精製した。免疫精製前の各RNA試料と各RT結合RNA試料からそれぞれ同じ分量を採取し、Hybond N+メンブレン（Cytiva社）に格子状にスポットした。各種3' UTR RNAのスポットを含むメンブレンに所望の3' UTR RNAが存在することを確認するために、T4ポリヌクレオチドキナーゼ（NEB社）を用いて、³²Pで5'末端を放射標識した相補的なオリゴヌクレオチドプローブとのハイブリダイゼーションによる検出を行った。すなわち、カイコガのR2の3' UTRを発現する細胞の試料に、カイコガの3' UTR配列が存在するかどうかを確認した（カイコガの3' UTR用プローブは、CATCATGGATTAGGATCGGAAGACCCCG（配列番号539）、GTACGCCGCGCGAAATTGGATCAGTAGATG（配列番号540）およびGAGAAACAGACGGGCCTGATCTACACCC（配列番号541）である）。オナジショウジョウバエのR2の3' UTRを発現する試料に、オナジショウジョウバエの3' UTR配列が存在するかどうかを確認した（オナジショウジョウバエの3' UTR用プローブは、CTATCTGAACCGAAGTTCCGCAACGCCTACGTAC（配列番号542）、CACTGCGTGTGGTCAGTTTTCTAGCATGCACG（配列番号543）およびGATGTTATGCCAAGACAGCAAGCAAATGTTTTGAACCAAACG（配列番号544）である）。メダカのR2の3' UTRを発現する試料に、メダカの3' UTR配列が存在するかどうかを確認した（メダカの3' UTR用プローブは、TTGAGGCGAGTCACCACTCGCTTTCCGG（配列番号545）およびGTGTCCGTCACGGGGACGACATCCGAGTG（配列番号546）である）。

20

30

【0671】

図11Bから分かるように、カイコガの改変RTタンパク質は、同種の3' UTR配列だけでなく、オナジショウジョウバエまたはメダカのR2因子の3' UTR配列とも結合したが、オナジショウジョウバエの改変タンパク質とメダカの改変タンパク質はより高い選択性を示した。カイコガのRTは、ヒト細胞で比較的非特異的なRNA相互作用を示すことが分かった。

【0672】

40

実施例12. カイコガ、オナジショウジョウバエまたはメダカに由来する逆転写酵素のインビトロTPRT特異性

実施例3と同様にして、カイコガ由来のRTタンパク質（配列番号36）、オナジショウジョウバエ由来のRTタンパク質（配列番号33）およびメダカ由来のRTタンパク質（配列番号9）を調製した。さらに、3'末端に4ntのrRNA配列「R4」（配列番号280）が付加された、または「R4」が付加されていない、メダカの3' UTR（配列番号200）に由来するGIC：RT認識配列を含む遺伝子挿入コンストラクト（GIC）と、3'末端に4ntのrRNA配列「R4」（配列番号280）が付加された、または「R4」が付加されていない、オナジショウジョウバエの3' UTR（配列番号210）に由来するGIC：RT認識配列を含む遺伝子挿入コンストラクト（GIC）を、実施例1と同様にして調製した。

50

【0673】

各GICを利用した際の各RTの能力を検証するため、実施例6と同様にしてインビトロTPRTアッセイを行った。

【0674】

オナジショウジョウバエに由来するRTタンパク質は、メダカの3' UTRに由来するGIC：RT認識配列を含む遺伝子挿入コンストラクト（GIC）をTPRTに利用せず、メダカのRTに由来するRTタンパク質は、オナジショウジョウバエの3' UTRに由来するGIC：RT認識配列を含む遺伝子挿入コンストラクト（GIC）をTPRTに利用しなかった。一方、カイコガに由来するRTタンパク質は、両方の遺伝子挿入コンストラクト（GIC）をTPRTに利用することができた（図12）。

10

【0675】

カイコガのRTタンパク質は、TPRTアッセイにおいて、その他の改変R2 RTタンパク質とは対照的に、非特異的に鑄型を複製した（すなわち、同種のGICに対して選択的ではなかった）。例えば、メダカに由来するRTまたはオナジショウジョウバエに由来するRTは、それぞれに同種のGIC：RT認識配列に選択的であったことから、選択性の高い遺伝子挿入システムを設計する場合に好適である可能性がある。

【0676】

実施例13. RT特異性の系統発生スクリーニング

表4に概要を示すように、様々な生物種のレトロエレメントに由来するRTタンパク質を実施例3と同様の方法で調製し、様々な生物種の天然レトロエレメントの3' UTRに由来するGIC：RT認識配列を含む遺伝子挿入コンストラクト（GIC）を実施例1と同様の方法で調製した。このインビトロTPRT比較実験では、すべてのGIC：RT認識配列は、3' 末端に付加された4ntのrRNA配列「R4」（配列番号280）を含んでおり、必要に応じて、T7 RNAPによる転写開始用のグアノシンを5' 末端に付加した。

20

【0677】

各種RTが特定のGIC：RT認識配列を認識する能力を検証するため、実施例6と同様にしてインビトロTPRTアッセイを行った。変性PAGEゲルの予想された産物長のバンドの不透明度から、相対的なTPRT活性レベルを評価し、各候補タンパク質を、高いTPRT活性を有するタンパク質、中程度のTPRT活性を有するタンパク質、低いTPRT活性を有するタンパク質、またはTPRT活性がないタンパク質（アッセイで検出不能）に分類した。

30

【0678】

TPRTアッセイの結果を以下の表4にまとめた。各データ行は、使用した各RTタンパク質を表示しており、各RT配列が由来する供給源としての生物名を含んでいる。各データ列は、使用した各GICを表示しており、各GIC：RT認識配列が由来する供給源としての生物名を含んでいる。マイナス記号（-）で示したセルは、そのRTとGICの組み合わせにおいて、予想された長さの産物が観察されなかったことを示す。プラスマイナス記号（+/-）で示したセルは、少なくとも一部のアッセイにおいて、予想された長さの産物がわずかな量で検出されたことを示す。単一のプラス記号（+）で示したセルは、予想された長さの産物が少量で観察されたことを示し、二重のプラス記号（++）で示したセルは、予想された長さの産物が中程度の量で観察されたことを示し、三重のプラス記号（+++）で示したセルは、予想された長さの産物が多量に観察されたことを示す。

40

【0679】

キンカチョウ、メダカ、ノドジロシトド、ノドジロシギダチョウ、コクヌストモドキ（B系統のR2）およびオナジショウジョウバエのそれぞれに由来するRTタンパク質は、カイコガに由来するRTタンパク質よりも、同種のGIC：RT認識配列を含むGICに対する選択性が高かった。したがって、対象のゲノムへの意図しない鑄型配列の挿入を最小限に抑えたり、これを阻止することを目的とする場合、キンカチョウ、メダカ、ノドジロシトド、ノドジロシギダチョウ、コクヌストモドキおよび／またはオナジショウジョウバエに由来するRTタンパク質は、カイコガ由来のRTタンパク質よりも、本発明の遺伝子挿入システムへの組み込みに好適である可能性がある。

50

【0680】

さらに、ノドジロシトド、キンカチョウおよび／またはノドジロシギダチョウに由来するRTタンパク質は、鳥類由来のGIC：RT認識配列に対する特異性が高かった。したがって、ノドジロシトド、キンカチョウおよび／またはノドジロシギダチョウに由来するRTタンパク質は、対象のゲノムへの意図しない鑄型配列の挿入を阻止できる可能性があるとともに、3'側モジュールを柔軟に組換えることが可能であることから、本発明の遺伝子挿入システムへの組み込みに好ましい可能性がある。

【0681】

【表5】

表4: RTの特異性

		GIC: 3'側モジュールの RT 認識配列 (各列に記載の供給源に由来するもの)																		
配列 番号		217	212	213	214	211	215	208	207	206	204	200	202	203	201	205	210	209	218	219
	供給源の 参照 コード	DrMerc-GIC	LeCoB-GIC	TriCan-GIC	Ciln-GIC	GaAc-GIC	DrMe-GIC	LiPo-GIC	PuPu-GIC	NaviB-GIC	GeFo-GIC	OrLa-GIC	ZoAl-GIC	TaGu-GIC	TriCasB-GIC	TiGu-GIC	DroSi-GIC	BoMo-GIC	AdVa-GIC	HyMa-GIC
7	NaviB-RT	-	-	-	-	-	-	+	+	++	+	+/-	++	+	+++	++	+/-	+/-	+/-	-
9	OrLa-RT	-	+	-	-	-	-	+/-	+	+/-	+	++	+	+	+/-	++	-	+	+/-	-
18	ZoAl-RT	-	-	-	-	-	-	-	-	-	++	-	+	++	-	++	-	-	-	-
27	TaGu-RT	-	-	-	-	-	-	-	-	-	+++	-	+++	+++	-	+++	-	-	-	-
2	TriCasB-RT	+/-	-	-	-	-	+/-	+/-	++	++	++	++	++	++	++	++	+	+	+/-	-
25	TiGu-RT	ND	ND	ND	ND	ND	ND	ND	ND	ND	+	-	+/-	+/-	-	+	ND	ND	ND	ND
33	DroSi-RT	+/-	ND	ND	+/-	ND	+/-	ND	ND	ND	ND	+/-	ND	ND	+	ND	++	+	+	ND
36	BoMo-RT	+	+	+	-	-	++	++	++	++	++	++	++	++	++	++	+++	+++	++	+/-

【0682】

実施例14. カイコガ由来RTを用いたTPRTに対する3'側モジュールの組換えの効果

カイコガに由来するRTタンパク質（配列番号36）を、実施例3と同様の方法で調製した。以下の表5に示す5'隣接配列および／または3'隣接配列を付加したBoMo 3' UTR配列（配列番号209）を含む遺伝子挿入コンストラクト（GIC）を、実施例1と同様の方法で調製した。

【0683】

10

20

30

40

50

【表 6】

表 5: カイコガ由来の遺伝子挿入コンストラクト(GIC)

* T7 RNAP による転写開始用に付加した 5'末端のグアノシンを示す

鋳型の参照コード	GIC: 5' rRNA の長さ	レトロエレメント由来配列の供給源	検証した 3' rRNA の長さ	A トラクトの長さ
GG*-BM3UTR-R3	0 nt	カイコガ	3 nt (配列番号 286)	0 nt
R26_ BM3UTR	26 nt (配列番号 254)	カイコガ	0 nt	0 nt
GG*_ BM3UTR_R4	0 nt	カイコガ	4 nt (配列番号 280)	0 nt
GGG*-R4_ BM3UTR_R4	4 nt (配列番号 275)	カイコガ	4 nt (配列番号 280)	0 nt
R26_ BM3UTR_R4	26 nt (配列番号 254)	カイコガ	4 nt (配列番号 280)	0 nt
R26_ BM3UTR_R4_PA	26 nt (配列番号 254)	カイコガ	4 nt (配列番号 280)	22 nt
R26_ BM3UTR_R20	26 nt (配列番号 254)	カイコガ	20 nt (配列番号 285)	0 nt

10

【0684】

カイコガ由来のRTタンパク質を、各種鋳型および標的部点である64bpまたは84bpのDNA二本鎖（それぞれ配列番号291および配列番号292）のそれぞれと組み合わせて、実施例6と同様にインビトロTPRTアッセイを行った。矢印は、形成されると予想された3'側の連結部の予想されたTPRT産物長が示された領域を示す。

20

【0685】

図13から分かるように、カイコガ由来3' UTR RNAの3'末端から配列を伸長させても、カイコガ由来RTによるtarget primed reverse transcription (TPRT) の効率には大きな影響は認められなかった。具体的には、3'隣接rRNAは、TPRT用の鋳型には不必要であった。4ntのrRNAを3'末端に付加すると、TPRT産物長の均一性が向上したが、鋳型RNAの全長がcDNAに複製されると予想したように、実際のTPRT産物長は増加しなかった。その代わりに、鋳型の長さを4nt伸長することによって、ニックが導入された標的部点をプライマーとして塩基対が形成され、cDNAの合成が開始されると考えられる。

【0686】

3' rRNAの長さを20ntまで伸長することによって、内部配列からのcDNAの合成を開始することが可能になることから（丸印の位置）、4ntの3' rRNAのみを有する鋳型RNAを用いた意図した高精度なTPRT合成と比べて（矢印は、高い忠実度で3'側の連結部が形成された領域を示す）、3'側の連結部の忠実度が低下する。したがって、20ntの3'隣接rRNA配列は、4ntの3'隣接rRNA配列に比べて不利であった。一方、3'隣接rRNAを少なくとも22ntのアデノシン鎖（PA）で伸長しても、正しい産物を合成する効率や正確性が失われないということは注目すべき点である。

30

【0687】

実施例15. メダカ由来RTのTPRT効率に対する3'側モジュールの組換えの効果

メダカに由来するRTタンパク質（配列番号9）を、実施例3と同様の方法で調製した。以下の表6に示す5'隣接配列および／または3'隣接配列を付加したOrLa 3' UTR配列（配列番号200）を含む遺伝子挿入コンストラクト（GIC）を、実施例16と同様の方法で調製した。

40

【0688】

50

【表 7】

表 6: メダカ由来の遺伝子挿入コンストラクト(GIC)

* T7 RNAP による転写開始用に付加した 5'末端のグアノシンを示す

鑄型の参照コード	GIC: 5' rRNA の長さ	レトロエレメント由来領域	GIC: 3'側モジュールの rRNA 配列の長さ	GIC: 3'側モジュールの A トラクト配列の長さ
R26_OL	26 nt (配列番号 254)	メダカ	0 nt	0 nt
R4_OL_R4	4 nt (配列番号 275)	メダカ	4 nt (配列番号 280)	0 nt
R26_OL_R4	26 nt (配列番号 254)	メダカ	4 nt (配列番号 280)	0 nt
R26_OL_R20	26 nt (配列番号 254)	メダカ	20 nt (配列番号 285)	0 nt
R26_OL_R4_PA	26 nt (配列番号 254)	メダカ	4 nt (配列番号 280)	22 nt
GG*-R0-OL3-R0	0 nt	メダカ	0 nt	0 nt
GG*-R0-OL3-R4	0 nt	メダカ	4 nt (配列番号 280)	0 nt
GG*-R0-OL3-R8	0 nt	メダカ	8 nt (配列番号 287)	0 nt
GG*-R0-OL3-R12	0 nt	メダカ	12 nt (配列番号 288)	0 nt
GG*-R0-OL3-R16	0 nt	メダカ	16 nt (配列番号 289)	0 nt
GG*-R0-OL3-R20	0 nt	メダカ	20 nt (配列番号 285)	0 nt

10

20

【0689】

メダカ由来の RTタンパク質と各種鑄型をそれぞれ組み合わせて、実施例6と同様にしてインビトロTPRTアッセイを行った。逆転写産物が形成されたことから、メダカ由来RTがTPRTにおいて生化学的に活性であることが示された。

【0690】

図14 (A) から分かるように、rRNAによる3'伸長部のないメダカの3'UTRは、メダカ由来RTによるTPRTに効率的に使用されておらず、3'隣接rRNAがなくても、カイコガの3'UTR RNAがカイコガ由来RTによりTPRTに効率的に使用されることが示された図13とは異なる結果になった。また、カイコガ由来の構成要素を用いた場合と同様に、3'隣接rRNAを少なくとも22ntのポリアデノシン鎖 (PA) で伸長しても、メダカ由来RTによるTPRTは阻害されず、産物長の均一性は増加した。

30

【0691】

本試験で検証した様々な3' rRNAの長さに対する効果を系統的に調査するため、第2のTPRTアッセイを実施した。

【0692】

図14 (B) から分かるように、この結果からも前述の結果が裏付けられた。3' rRNAの伸長部がない場合、メダカ由来RTによる活性が不十分となり、内部配列からの合成開始が不適切になったが、4ntのrRNAの存在により、TPRTが十分に刺激され、3'側の連結部の正確性が向上した。したがって、本発明の遺伝子挿入コンストラクト (GIC) では、GIC 3'側モジュールのrRNA配列に、本試験で検証した4ntの3' rRNAのみを組み込むことが好ましいと考えられる。GIC 3' rRNA配列の長さを伸長しても、TPRT産物の長さは増加しないことから、GIC 3' rRNA配列は複製されないことが示された。その代わりに、GIC 3' rRNA配列は、ニックが導入された標的部位をDNAプライマーとして塩基対を形成し、cDNAの合成を開始することができる。

40

【0693】

実施例16. コクヌストモドキ由来RTのTPRT効率に対する3'側モジュールの組換えの効果

コクヌストモドキ由来のRTタンパク質 (配列番号2) を実施例3と同様にして調製した

50

。以下の表7に示す5'隣接配列および／または3'隣接配列を付加したTriCasB 3' UTR配列（配列番号201）を含む遺伝子挿入コンストラクト（GIC）を、実施例1と同様の方法で調製した。

【0694】

【表8】

表7: コクヌストモドキ由来の遺伝子挿入コンストラクト(GIC)

鋳型の参照コード	GIC: 5' rRNA 配列の長さ	レトロエレメント 由来領域	GIC: 3'側モジュールの rRNA 配列の長さ	GIC: 3'側モジュールの A トラクト 配列の長さ
R25-TC_UTR-R4	25 nt (配列番号 276)	コクヌストモドキ	4 nt (配列番号 280)	0 nt
R25-TC_UTR-R4_PA	25 nt (配列番号 276)	コクヌストモドキ	4 nt (配列番号 280)	22 nt
R25-TC_UTR-R10	25 nt (配列番号 276)	コクヌストモドキ	10 nt (配列番号 284)	0 nt

10

【0695】

コクヌストモドキ由来のRTタンパク質と各種鋳型をそれぞれ組み合わせて、実施例6と同様にしてインビトロTPRTアッセイを行った。矢印は、目的のTPRT産物の位置を示す。標的部位のDNAは、画像の下部に示されている黒色のバンドとして検出される。逆転写産物が形成されたことから、コクヌストモドキ由来RTがTPRTにおいて生化学的に

20

【0696】

図15から分かるように、4ntを超えるrRNA配列をGIC: 3'側モジュールに付加しても、産物合成の向上は認められなかったが、正確な産物の合成を阻害することなく、少なくとも22ntのポリアデノシン鎖（PA）で3'隣接rRNAを伸長することができた。

【0697】

実施例17. ノドジロシトド由来RTおよびキンカチョウ由来RTのTPRT効率に対する3'側モジュールの組換えの効果

ノドジロシトドに由来するRTタンパク質（配列番号18）を、実施例3と同様の方法で調製した。以下の表8に示す5'隣接配列および／または3'隣接配列を付加した、ノドジロシトド（ZoAl）3' UTR（配列番号202）またはノドジロシギダチョウ（TiGu）3' UTR（配列番号205）またはキンカチョウ（TaGu）3' UTR（配列番号203）の3'側モジュールRT認識配列を含む遺伝子挿入コンストラクト（GIC）を、実施例1と同様の方法で調製した。

30

【0698】

40

50

【表 9】

表 8: 鳥類由来 R2 遺伝子挿入コンストラクト(GIC)

鋳型の参照コード	図 16 のレー ン番号	GIC: 5' rRNA 配列の長さ	GIC: 3'側モジュ ールの RT 認識 配列の供給源	GIC: 3'側モジュ ールの rRNA 配列の長さ	GIC: 3'側モ ジュールの A トラクト の長さ
R26(-28)-ZA3-R0	1	26 nt (配列番号 254)	ノドジロシトド	0 nt	0 nt
R26(-28)-ZA3-R4	2	26 nt (配列番号 254)	ノドジロシトド	4 nt (配列番号 280)	0 nt
R26(-28)-ZA3-R20	3	26 nt (配列番号 254)	ノドジロシトド	20 nt (配列番号 285)	0 nt
R26(-28)-ZA3-R4PA	4	26 nt (配列番号 254)	ノドジロシトド	4 nt (配列番号 280)	22 nt
R26(-28)-TiG3-R0	5	26 nt (配列番号 254)	ノドジロシギダ チョウ	0 nt	0 nt
産物の喪失	6				
R26(-28)-TiG3-R20	7	26 nt (配列番号 254)	ノドジロシギダ チョウ	20 nt (配列番号 285)	0 nt
R26(-28)-TiG3-R4PA	8	26 nt (配列番号 254)	ノドジロシギダ チョウ	4 nt (配列番号 280)	22 nt
R28(-28)-TaG3-R0	9	28 nt (配列番号 252)	キンカチョウ	0 nt	0 nt
R28(-28)-TaG3-R4	10	28 nt (配列番号 252)	キンカチョウ	4 nt (配列番号 280)	0 nt
R28(-28)-TaG3-R20	11	28 nt (配列番号 252)	キンカチョウ	20 nt (配列番号 285)	0 nt
R28(-28)-TaG3-R4PA	12	28 nt (配列番号 252)	キンカチョウ	4 nt (配列番号 280)	22 nt

【0699】

ノドジロシトド由来のRTタンパク質と各種鋳型をそれぞれ組み合わせて、実施例6と同様にインビトロTPRTアッセイを行った。実線の囲み枠はTPRT産物を示し、破線の囲み枠は沈殿回収操作コントロールを示し、二点鎖線の囲み枠は64bpの標的部DNAを示す。これらの結果から、ノドジロシトド由来のRTが、TPRTにおいて生化学的に活性であることが示された。

【0700】

図16から分かるように、ノドジロシトド由来のRTタンパク質は、GIC: 3'側モジュールのrRNA配列を持たない3'側モジュール設計を有する遺伝子挿入コンストラクト(GIC)を効率的に利用できなかったことから、GIC 3' rRNA配列が塩基対合できる標的部位において、cDNAの合成効率が増加することが示された。GIC 3' rRNA配列の長さを伸長しても、TPRT産物の長さは増加しないことから、GIC 3' rRNA配列が複製されないことが示された。GIC 3' rRNA配列は、ニックが導入された標的部位をプライマーとして塩基対を形成し、cDNAの合成を開始しているに違いないと考えられる。4ntの3' rRNA配列と22ntのポリA鎖テールとを含む遺伝子挿入コンストラクト(GIC)、または20ntのrRNA配列を含む遺伝子挿入コンストラクト(GIC)を使用した場合に、合成されたTPRT産物の量が最も多くなった。さらに、ノドジロシトド由来のRTタンパク質は、

検証した数種類の鳥類のそれぞれに由来するGIC：3'側モジュールRT認識配列を含む遺伝子挿入コンストラクト（GIC）を利用することができた。この実験と並行して、キンカチョウに由来するRTタンパク質（配列番号27）を用いた実験を行ったところ、キンカチョウ由来の鳥類RTタンパク質は、数種類の鳥類に由来するGIC：3'側モジュールRT認識配列を含む遺伝子挿入コンストラクト（GIC）を利用することができ、GIC：3' rRNA配列を含む遺伝子挿入コンストラクト（GIC）の利用に対して選択性を示した。

【0701】

以上の結果から、TPRT反応効率を改変するには、遺伝子挿入システムにおいて、ノドジロシドまたはキンカチョウに由来するRTタンパク質に、様々な鳥類に由来するGIC：3'側モジュールRT認識配列と、GIC：3'側モジュールrRNA配列とを組み合わせてもよく、GIC：3'側モジュールAトラクト配列は含まれていても含まれていなくてもよいことがさらに支持された。GIC：3'側モジュールのrRNA配列が、ニックが導入された標的部位をプライマーとして塩基対を形成することができない場合、cDNAの合成は観察されなかった。GIC：3'側モジュールのrRNAと塩基対を形成することが可能な、ニックの下流の標的部位配列を別の配列に改変した場合（変異標的部位；配列番号296）、依然として切断は観察されたものの、GIC：3'側モジュールのrRNAが、プライマー鎖の3'末端との塩基対を形成できなかったことからTPRTは阻害された。したがって、標的部位の正確な配列にニックが導入された場合にのみ、GIC：3'側モジュールのrRNA配列にマッチするプライマー鎖3'末端が形成されて、TPRTによりcDNAが合成される。変異標的部位を使用した場合、GIC：3'側モジュールのrRNA配列を、このニックにより形成されたプライマー鎖の3'末端と塩基対を形成可能な配列に変更したところ、TRPTによるcDNAの合成が回復した。この結果から、GIC：3'側モジュールのRNA配列の機能機構が、プライマー鎖の3'末端領域との塩基対の形成にあることが示された。

【0702】

実施例18. インビボでのヒトゲノムへの導入遺伝子の挿入に対するGIC：3'側モジュールのテールの組換えの効果

パートA：キンカチョウに由来するRTC：RTモジュール

キンカチョウに由来する逆転写酵素コンストラクト（RTC）のmRNA（配列番号28）を、実施例4と同様にして作製した。また、ペイロードとしてのGFP導入遺伝子発現カセットと、同じGIC：5'側モジュールとGIC：3'側モジュールRT認識配列を含む複数の遺伝子挿入コンストラクト（GIC）のRNA（TCA5_CBhBsi_GFP_GeFo3）を実施例2と同様にして作製し、以下の表9に詳細を示した。

【0703】

Lipofectamine Messenger Maxを使用して、RTCと各種GICを（1：1のモル比で）hTERT RPE-1細胞に共トランスフェクトし、24時間後に細胞を回収した。FACS分析により、各処理で得られたGFP陽性細胞の割合を測定し、この結果を以下の表9に報告した。

【0704】

【表10】

表9:インビボにおける3'側モジュールのテールの組換えの効果

GIC: 3'側モジュールのrRNA 配列の長さ (nt)	GIC: 3'側モジュールのA トラクトの長さ (nt)	GIC の配列番号	GFP 陽性細胞の割合 (%)
0	0	500	0.12
0	22	501	0.17
4	0	502	4.05
4	22	503	15.67
20	0	504	6.84
20	22	505	4.23

【0705】

以上の結果から、4ntのGIC：3'側モジュールrRNA配列と22ntのAトラクト配列を含むGIC：3'側モジュールを使用すると、検証したその他の組み合わせよりも導入遺伝子の挿入率が有意に増加したことが示された。4ntのGIC：3'側モジュールrRNA配列を少なくとも含んでいたその他の組み合わせでも、哺乳動物細胞株に導入遺伝子を挿入して発現させることができたことには注目されたい。一方で、20ntのGIC：3'側モジュールrRNA配列を使用すると、22ntの長さのAトラクト配列が阻害性を示した。

【0706】

パートB：キンカチョウ由来RTC：RTモジュールとノドジロシトド由来RTC：RTモジュールの比較

キンカチョウに由来する逆転写酵素コンストラクト（RTC）のmRNA（配列番号28）およびノドジロシトドに由来する逆転写酵素コンストラクト（RTC）のmRNA（配列番号19）を、実施例4と同様にして作製した。また、ペイロードとしてのGFP導入遺伝子発現カセットと、同じGIC：5'側モジュールとGIC：3'側モジュールRT認識配列を含む複数の遺伝子挿入コンストラクト（GIC）のRNA（TCA5_CBhBsi_GFP_GeFo3）を実施例2と同様にして作製し、以下の表10に詳細を示した。

【0707】

Lipofectamine Messenger Maxを使用して、RTCと各種GICを（1：3のモル比で）hTERT RPE-1細胞に共トランスフェクトし、24時間後に細胞を回収した。GFP陽性細胞の割合と、GFP陽性細胞におけるGFP発現強度の中央値をFACS分析により各処理ごとに測定した。この結果を以下の表10に示す。

【0708】

【表11】

表 10:インビボにおける 3'側モジュールのテールの組換えの効果の別の検討

RTC: RT モジュールの 供給源の生物	GIC: 3'側 モジュール の rRNA の 長さ (nt)	GIC: 3'側 モジュール の A トラク トの長さ (nt)	GIC の 配列番号	GFP 陽性細胞 の割合 (%)	GFP の強度 (遺伝子挿入コンス トラクトのみの バックグラウンド に対する 相対蛍光単位)
キンカチョウ	0	0	500	0.093	1705
キンカチョウ	0	22	501	0.17	2098
キンカチョウ	4	0	502	2.84	4570
キンカチョウ	4	22	503	14.708	9011
キンカチョウ	20	0	504	5.342	5003
キンカチョウ	20	22	505	2.235	3835
ノドジロシトド	0	0	500	0	0
ノドジロシトド	0	22	501	0.25	2183
ノドジロシトド	4	0	502	3.83	4260
ノドジロシトド	4	22	503	13.608	7364
ノドジロシトド	20	0	504	4.972	4315
ノドジロシトド	20	22	505	2.075	3145

【0709】

本実験で得られた上記の結果を、キンカチョウに由来するRTC mRNAを使用して実施したパートAの実験の結果と合わせて検討した。その結果、ノドジロシトドに由来するRTC mRNAが、GIC：3'側モジュールrRNA配列およびAトラクトの長さに関して、キン

カチヨウ由来RTC mRNAと同じパターンの効率を示したことが示された。最適なR4A2 2鑄型を使用した場合、ノドジロシトド由来のRTC：RTモジュールによる導入遺伝子の挿入効率は、キンカチヨウ由来のRTC：RTモジュールと比べてわずかに低いだけであった。

【0710】

キンカチヨウ由来のRTC：RTモジュールと、ノドジロシトド由来のRTC：RTモジュールは、いずれも本発明の遺伝子挿入システムの構成要素として有効なものであった。いずれのRTC：RTモジュールも、様々な長さのGIC：3'側モジュールrRNAおよび／またはGIC：3'側モジュールAトラクトを含む遺伝子挿入コンストラクト（GIC）を利用することができ、遺伝子挿入コンストラクトの最適な組成として、約4ntの長さのGIC：3'側モジュールrRNA配列と、約22ntの長さのGIC：3'側モジュールAトラクト配列を採用できる可能性が示された。

10

【0711】

実施例19. キンカチヨウ由来RTのTPRT効率に対する3'側モジュールの組換えの効果

キンカチヨウに由来するRTタンパク質（配列番号27）を、実施例3と同様の方法で調製した。以下の表11の記載に従って、GIC：3'側モジュールrRNA配列としてR4（配列番号280）を付加した様々なGIC：3'側モジュールRT認識配列を含む遺伝子挿入コンストラクト（GIC）を実施例1と同様にして調製し、そのうちの一部には、T7 RNAPによる転写開始用の5'グアノシンを付加した。

20

【0712】

30

40

50

【表 1 2】

表 11: GIC: 3'側モジュールの RT 認識配列に対するキンカチョウ由来 RT の特異性

* T7 RNAP による転写開始用に付加した 5'末端のグアノシンを示す

鋳型の参照コード	図 17 の レーン 番号	GIC: 3'側モジュールの RT 認識配列の供給源と その配列番号(括弧内)	GIC: 3'側 モジュールの rRNA 配列の 長さ	配列 番号
鋳型なしのコントロール	2	NA	NA	NA
GGG*-HM3-R4	3	チクビヒドラ(219)	4 nt	
GGG*-AV3-R4	4	<i>A. vaga</i> (218)	4 nt	
G*-LP3-R4	5	アメリカカブトガニ(208)	4 nt	
G*-ZA3-R4	6	ノドジロシトド (202)	4 nt	
G*-TiG3-R4	7	ノドジロシギダチョウ (205)	4 nt	
G*-TaG3-R4	8	キンカチョウ (203)	4 nt	
G*-GF3-R4	9	ガラパゴスフィンチ (204)	4 nt	
GA3-R4	10	イトヨ(211)	4 nt	
OL3-R4	11	メダカ(200)	4 nt	
G*-PP3-R4	12	キタノトミヨ(207)	4 nt	
GGG*-TCasB3-R4	13	コクヌストモドキ(201)	4 nt	
G*-NVB3-R4	14	キョウソヤドリコバチ (206)	4 nt	
GGG*-CI3-R4	15	カタユウレイボヤ (214)	4 nt	
BM3-R4	16	カイコガ(209)	4 nt	
G*-LCB3-R4	17	<i>L. couesii</i> (212)	4 nt	
G*-TCan3-R4	18	ヨーロッパカブトエビ (213)	4 nt	
G*-DS3-5iA-R4	19	オナジショウジョウバエ (210)	4 nt	
GG*-DMer3-R4	20	<i>D. mercatorum</i> (217)	4 nt	
G*-DMel3-5iA-R4	21	キイロショウジョウバエ (215)	4 nt	
GG*-DN3-R4	22	<i>D. nasuta</i> (216)	4 nt	

【0 7 1 3】

キンカチョウ由来の RT タンパク質と各種鋳型をそれぞれ組み合わせて、実施例 6 と同様にしてインビトロ TPRT アッセイを行った。鋳型配列は、レトロエレメント 3' UTR 配列と、必要に応じて付加した T7 RNAP 転写を補助するための 5' グアノシンと、4 nt の長さの GIC: 3' 側モジュール rRNA 配列とを含んでいたが、GIC: 3' 側モジュールの Aトラクト rRNA 配列は含んでいなかった。実線の囲み枠は予想された TPRT 産物を示し、破線の囲み枠は沈殿回収操作コントロールを示し、二点鎖線の囲み枠は、ニックが導入されていない 64 bp の標的部位 DNA を示す。

【0 7 1 4】

図 17 に示したように、キンカチョウに由来する RT タンパク質は、様々な鳥類に由来する GIC: 3' 側モジュール RT 認識配列を有する遺伝子挿入コンストラクト (GIC) を認識

することができたが、鳥類以外の生物種に由来するGIC：3'側モジュールRT認識配列を含む遺伝子挿入コンストラクト（GIC）の存在下では、TPRT活性はほとんど観察されな
いか、全く観察されなかった。さらに、キンカチョウ由来のRTタンパク質と、検討した
もののうち最も短い鳥類由来のGIC：3'側モジュールRT認識配列を有するガラパゴスフ
ィンチ由来の遺伝子挿入コンストラクト（GIC）とを組み合わせた場合に、高いTPRT活
性が観察された。

【0715】

したがって、本発明の少なくとも1つの遺伝子挿入システムは、特に、鳥類以外の対象
への投与を想定する場合、少なくとも1つのキンカチョウ由来RTタンパク質を含むか、
これをコードする少なくとも1つのRTC：RTモジュールと、少なくとも1つのガラパゴ
スフィンチ由来GIC：3'側モジュールRT認識配列を含むか、これをコードする少なく
とも1つのGICとを含むように設計することが好ましいと考えられる。この組み合わせを遺
伝子挿入システムに組み込むことによって、対象のゲノムへのペイロード配列の挿入効率
が高くなるとともに、遺伝子挿入コンストラクトの特異性も高くなると考えられる。

【0716】

実施例20. インビボでのコクヌストモドキ由来RTによるヒトゲノムへの導入遺伝子の挿
入効率に対する5'側モジュールおよび3'側モジュールの組換えの効果

実施例3と同様にして作製したプラスミドを293T細胞にトランスフェクトして、翻訳
用の単一のAUG開始コドンを提供する合成ORF配列を含む3系統のコクヌストモドキ由来
R2レトロエレメントのうちの1つを改変したタンパク質（配列番号1）を発現させた。
これと並列して、プラスミドをトランスフェクトしていない細胞を陰性コントロールとし
て試験した。48時間後に、Lipofectamine 3000を用いて、以下の表12に記載の組み
合わせを含むように実施例1と同様にして調製した精製遺伝子挿入コンストラクト（GIC
）のRNAを細胞にトランスフェクトした。この2回目のトランスフェクションの翌日に、
トランスフェクトした細胞からゲノムDNAを精製した。

【0717】

10

20

30

40

50

【表 1 3】

表 12: コクヌストモドキ由来の遺伝子挿入コンストラクト(GIC)

鋳型の参照コード	GIC: 5'側 モジュール の rRNA 配列の長さ*	GIC: 5'側 モジュール のレトロエ レメント 配列の 供給源**	GIC: 3'側 モジュール の RT 認識 配列の 供給源	GIC: 3'側 モジュール の rRNA 配列の長さ	GIC: 3'側 モジュールの A トラクト 配列の長さ	配列 番号
R25-TCB3-R4	25 nt	NA	コクヌスト モドキ	4 nt	0 nt	330
R25-TCB3-R10	25 nt	NA	コクヌスト モドキ	10 nt	0 nt	331
R25-TCB3-R4-PA	25 nt	NA	コクヌスト モドキ	4 nt	22 nt	332
R25*-TCB5_TCB3-R4	25 nt*	コクヌスト モドキ	コクヌスト モドキ	4 nt	0 nt	333
R25*-TCB5_TCB3- R10	25 nt*	コクヌスト モドキ	コクヌスト モドキ	10 nt	0 nt	334
R25*-TCB5_TCB3-R4- PA	25 nt*	コクヌスト モドキ	コクヌスト モドキ	4 nt	22 nt	335
R25*-TCB5_TCB3-PA	25 nt*	コクヌスト モドキ	コクヌスト モドキ	0 nt	22 nt	336
R25*-TCB5_TCB3- R10PA	25 nt*	コクヌスト モドキ	コクヌスト モドキ	10 nt	22 nt	337
3'側の 25nt の rRNA のうち 13nt は、GIC:5'側モジュール内に含まれており、自己切断後にそのまま残る。5'側の 12nt は除去される。 **TriCasB の 5'側モジュール配列は、短い天然型 rRNA や進化的に改変された rRNA 配列ではなく、ヒトゲノムにマッチした第 1 のニック標的部位の上流に 13nt の rRNA を含むように天然型から改変されたものである。						

【0 7 1 8】

図 18A に示した 1 つの実験において、各遺伝子挿入コンストラクト (GIC) は、いずれもコクヌストモドキ B 系統の R2 レトロエレメント由来の 5' 側モジュールと、コクヌストモドキ B 系統の R2 レトロエレメント由来の 3' 側モジュールとを有しており (「5_3UTR」)、GIC: 3' 側モジュールの rRNA の長さ (ゼロ、R4 または R10) と、GIC: 3' 側モジュールの 22 nt の A トラクト (PA) の有無の点でそれぞれ異なっていた。様々な種類の細胞集団から得た同量のゲノム DNA を使用して PCR を実施することにより、導入遺伝子が挿入された 3' 側の連結部を検出した (フォワードプライマー: CTCCTGACCAACTAGCTCACTGACTAATTTTAAAC (配列番号 547) およびバースプライマー: CCACTTATTCTACACCTCTCATGTCTCTTCACCG (配列番号 548))。PCR 産物として得られた DNA を非変性アガロースゲルで分離し、臭化エチジウムで検出した。

【0 7 1 9】

22 nt の A トラクト (PA) を有する GIC: 3' 側モジュールをトランスフェクトした細胞において、特に、4 nt の長さの rRNA を有する GIC: 3' 側モジュールをトランスフェクトした細胞において、目的の 3' 側の連結部の予測されたサイズのジャンクション PCR 産物が最も多く認められた。GIC: 3' 側モジュール rRNA を持たない GIC: 3' 側モジュール A トラクトでは、導入遺伝子の挿入を十分に検出できず、これは、導入遺伝子合成用の鋳型

に潜在的に利用可能なものとして、ヒト宿主細胞由来のアデノシンテールを含むmRNAを除外する際に有利である。

【0720】

図18Bに示した別の実験において、各遺伝子挿入コンストラクト（GIC）は、コクヌストモドキB系統のR2レトロエレメント由来の3'側モジュールを有しており、コクヌストモドキB系統のR2レトロエレメント由来の5'側モジュールの有無の点でそれぞれ異なっていた（それぞれ「5_3」または「3」で示す）。さらに、各遺伝子挿入コンストラクト（GIC）は、GIC：3'側モジュールのrRNAの長さ（R4もしくはR10）の点、および／またはGIC：3'側モジュールのAトラクト（PA）の有無の点でもそれぞれ異なっていた。様々な種類の細胞集団から得た同量のゲノムDNAを使用してPCRを実施することにより、導入遺伝子が挿入された3'側および5'側の連結部を検出した。このPCRでは、3'挿入ジャンクションプライマー（フォワードプライマー：CTCCTGACCAACTAGCTCACTGACTAATTTTAAAC（配列番号547）およびリバースプライマー：CCACTTATTCTACACCTCTCATGTCTCTTCACCG（配列番号548））、または5'挿入ジャンクションプライマー（フォワードプライマー：CTAGCAGCCGACTTAGAACTGGTGCGG（配列番号549）およびリバースプライマー：CTTCGTCTTCGGAATCCATGTCCATAGC（配列番号550））を使用した。PCR産物として得られたDNAを、1×TAEで泳動した非変性アガロースゲルで分離し、臭化エチジウムで検出し、BioRad社製molecular imager ChemiDoc XRS+で撮影した。

10

【0721】

左パネルでは、矢印で示すように、GIC：3'側モジュールの4ntの長さのrRNAとGIC：3'側モジュールのAトラクト（PA）をトランスフェクトした細胞において、3'側の正確な連結部の予測されたサイズのPCR産物が最も多く認められた。さらに、コクヌストモドキB系統のR2レトロエレメント由来の5'側モジュールが存在する場合、3'側の連結部を示す産物が増加したことから、挿入された導入遺伝子が多かったことが示された。R4_PA GICでは、不正確なサイズのPCR産物があったとしても、わずかな量しか検出されず、3'連結部の形成の忠実度が高かったことが示された。一方、その他の遺伝子挿入コンストラクトをトランスフェクトした細胞では、さらに別の3'連結部のPCR産物が示された。

20

【0722】

右パネルでは、導入遺伝子の全長の5'連結部の予想されたサイズのPCR産物は、それぞれ矢印で示すように、5'側モジュールの有無に応じて遺伝子挿入コンストラクトごとに様々なサイズで検出された。導入遺伝子の全長が挿入された5'連結部のPCR産物は、GIC：3'側モジュールの4ntの長さのrRNAとGIC：3'側モジュールのAトラクト（PA）をトランスフェクトした細胞において最も多く認められた。さらに、コクヌストモドキB系統のR2レトロエレメント由来の5'側モジュールが存在する場合、5'連結部のPCR産物の長さが長かったにもかかわらず（長さが長いとPCRの効率が低くなりやすい）、5'連結部の産物の量と均一性が増加したことから、挿入された導入遺伝子が多く、挿入の忠実度が高かったことが示された。

30

【0723】

5'連結部と3'連結部の形成は、RTタンパク質の発現とRNA鋳型のトランスフェクションの両方が起こった場合にのみ検出することができた。鋳型RNAの非存在下でRTタンパク質を発現した細胞や、RTタンパク質の発現の非存在下で鋳型RNAをトランスフェクトした細胞では、PCR産物が産生されないか、非特異的なPCR産物がわずかに見られた。

40

【0724】

以上の結果から、GIC：3'側モジュールのrRNA配列の長さが短いほど（例えば、4ntの長さのrRNA配列）、本発明の遺伝子挿入システムがより高いTPRT活性を有することが示され、具体的には、反応収率が高くなり、より特異的に導入遺伝子の連結部（5'連結部と3'連結部の両方）が形成されることが示された。

【0725】

50

実施例 21. インビボでのコクヌストモドキ由来 RT によるヒトゲノムへの導入遺伝子の挿入効率に対する 5' 側モジュールの RZ の組換えの効果

実施例 3 と同様にして、コクヌストモドキ由来の RT タンパク質（配列番号 1）を発現するように 293T 細胞にトランスフェクトした。次に、Lipofectamine 3000 を用いて、以下の表 13 に記載の組み合わせを含むように実施例 1 と同様にして調製した遺伝子挿入コンストラクト（GIC）の RNA を細胞にトランスフェクトした。すべての遺伝子挿入コンストラクト（GIC）には、コクヌストモドキに由来する GIC：3' 側モジュールの RT 認識配列、GIC：3' 側モジュールの 4nt の長さの rRNA 配列、および GIC：3' 側モジュールの 22nt の長さの A トラクト配列が含まれていた（配列番号 338）。これらの遺伝子挿入コンストラクトは、GIC：5' 側モジュールが様々に異なっていた。

10

【0726】

【表 14】

表 13: コクヌストモドキ由来の別の RZ を有する遺伝子挿入コンストラクト(GIC)

鋳型の参照コード	GIC: 5'側モジュールの rRNA 配列の長さ**	図 19 のレーン番号	GIC: 5'側モジュールの RZ 配列の供給源/改変	GIC: 5'側モジュールの RE 配列の供給源	配列番号
TriCasB_5 (配列番号 62)	13 (配列番号 266)	2 & 10	コクヌストモドキ/ なし, 塩基付加*	コクヌストモドキ	
TriCasB_5rzdead (配列番号 63)	13 (配列番号 266)	3 & 11	コクヌストモドキ/ 不活性化	コクヌストモドキ	
TriCasB_5RZ (配列番号 64)	13 (配列番号 266)	4 & 12	コクヌストモドキ/ なし, 塩基付加*	コクヌストモドキ	
TriCasB_5RZmin (配列番号 65)	13 (配列番号 266)	5 & 13	コクヌストモドキ/ 短縮した 5RZ	コクヌストモドキ	
TriCasB_5RZmin+down (配列番号 145)	13 (配列番号 266)	6 & 14	コクヌストモドキ/ TriCasB 5 の天然型 RZ 領域を短縮した 5RZ で置換	コクヌストモドキ	
OrLa_5L (配列番号 60)	26 (配列番号 254)	7 & 15	メダカ/なし	メダカ	
DroSi_5 (配列番号 70)	0	8 & 16	オナジショウジョウバエ/なし	オナジショウジョウバエ	
* TriCasB の 5'側モジュール配列は、短い天然型 rRNA や進化的に改変された rRNA 配列ではなく、ヒトゲノムにマッチした第 1 のニック標的部位の上流に 13nt の rRNA を含むように天然型から改変されたものである。					
** 自己切断後の 5' rRNA の長さ。					

20

30

40

【0727】

実施例 20 と同様にして、トランスフェクトした細胞プールから得たゲノム DNA の 2 つの別の PCR 増幅を行って、3' 挿入連結部（上パネル）と 5' 挿入連結部（下パネル）を検出した。

PCR プライマー：

3' 連結部：

フォワードプライマー：CTCCTGACCAACTAGCTCACTGACTAATTTTAAAC（配列番号 547）、

リバースプライマー：CCACTTATTCTACACCTCTCATGTCTCTTCACCG（配列番号

50

548) ;

5' 連結部:

フォワードプライマー: CCAGGGGAATCCGACTGTTTAATTAAACAAAGC (配列番号551)、

リバースプライマー: GCGACTCGCATCACTGACTTTAATTGGTTG (配列番号552)

【0728】

図19に示したように、挿入により正確な3' 連結部と5' 連結部が形成された導入遺伝子の全長の予想された長さの単一のPCR産物が主に確認されたことから、コクヌストモドキB系統またはメダカのR2レトロエレメントに由来する5' 側モジュール構成要素を含む遺伝子挿入コンストラクト (GIC) が、導入遺伝子の挿入と連結部の忠実度を最も良好に達成したことが示された。コクヌストモドキB系統の5' 側モジュールのRZ活性部位において、単一塩基を変化させたところ、RZ活性が破壊され (TriCasB_5rzdead)、導入遺伝子の挿入効率が著しく低下し、挿入の忠実度が損なわれた。また、コクヌストモドキのGIC: 5' 側モジュールのレトロエレメント配列の全長を含む遺伝子挿入コンストラクト (TriCasB_5) は、コクヌストモドキ由来の5' 側モジュール配列全長のRZ領域のみを含む遺伝子挿入コンストラクト (TriCasB_5RZ) よりも高い導入遺伝子挿入効率を示した。一方、コクヌストモドキ由来の最小限の長さのRZ領域のみを含む遺伝子挿入コンストラクト (TriCasB_5RZmin) は、遺伝子挿入コンストラクト「TriCasB_5」と同程度の性能を示し、「TriCasB_5RZ」よりも良好な性能を示し、「TriCasB_5」から「TriCasB_5RZ」を作製する際に除去したコクヌストモドキのRZ領域の下流の5' UTR配列を戻した「TriCasB_5RZmin+down」よりも良好な性能を示した。

【0729】

さらに、メダカの5' 側モジュール構成要素を含む遺伝子挿入コンストラクト (OrLa_5L) は、コクヌストモドキ由来のRTタンパク質とコクヌストモドキ由来のGIC: 3' 側モジュール構成要素と組み合わせた場合に、「TriCasB_5」と同程度の性能を示したが、これは、オナジショウジョウバエの5' 側モジュール構成要素を含む遺伝子挿入コンストラクト (DroSi_5) の場合には当てはまらなかった。オナジショウジョウバエの5' 側モジュールのRZによる自己切断活性により、5' UTRの5' 側の最初のGIC転写産物において、5' rRNAを含むすべての配列が除去される。自己切断RZ内に保護された5' rRNAがなくとも、最初の第1鎖のcDNA合成は依然として起こったが、5' 連結部の形成に必要な第2鎖の合成と安定な導入遺伝子の挿入効率と正確性が、「TriCasB_5」または「OrLa_5L」を含む遺伝子挿入コンストラクトよりも低下した。これは、5' 連結部のPCR産物の長さの分布がスミア化していることから裏付けられた (図19の下パネルのレーン16)。

【0730】

実施例22. GIC: 5' 側モジュールのrRNAの長さ

F-ZoAl RTの逆転写酵素コンストラクト (RTC) のmRNA (配列番号19) を、実施例4と同様にして作製した。5' 側モジュールの配列のみが異なるGFP導入遺伝子発現カセットを含む遺伝子挿入コンストラクト (GIC) のRNA (配列番号506: CBhBsi_GPF_GeFo_R4A22) を実施例2と同様にして作製した。自己切断型のHDV RZ折り畳み構造を利用できるように最適化されたGIC: 5' 側モジュール配列をデノボ設計により開発し、標的部位のニックの上流のrRNA配列の特定位置でGIC 5' 末端が切断されるように構成した。この切断位置は、例えば、自己切断に必要とされる+1個のグアノシンを得ることができ、かつT7 RNAPの転写産物が自己切断されることが実験により確認された-28位 (HDV-28)、-13位 (HDV-13) または別の位置である。

【0731】

さらに、自己切断型のHDV RZ折り畳み構造を利用できるように最適化してデノボ設計されたGIC: 5' 側モジュール配列を、GIC: 5' 側モジュール中のrRNA配列の量を変えることによって各位置で自己切断するように調節した。例えば、TPRTを介して形成されるニックから-28位において自己切断を誘導するGIC: 5' 側モジュールは、28ntの5' rRNAを含むことができ、あるいは、その3' 末端の境界からrRNA配列をトリミングするこ

とによって、25nt、26ntまたは27ntといった別の長さのrRNAを含むことができた。

【0732】

Lipofectamine Messenger Maxを使用して、RTC mRNAと各種GIC mRNAを1：3のモル比で混合してhTERT RPE-1細胞に共トランスフェクトした。24時間後に、トランスフェクト細胞のプールをフローサイトメトリーで分析して、GFP+細胞の割合（％）を検出した。FACS分析によりGFP陽性細胞の割合を測定した結果を以下の表14に報告した。

【0733】

【表15】

表 14: GIC: 5'側モジュール rRNA 配列の長さの効果

GIC: 5'側モジュールの RZ 配列の ID	GIC: 5'側 モジュール の rRNA の 開始配列 の位置	GIC: 5'側 モジュール の rRNA 配 列の長さ	RZ の 自己切断 効率	GFP 陽性細胞 の割合 (%)	自己切断され た GIC あた りの正規化し た GFP+細胞 の割合(%)
HDV-28(26)gu1 (配列番号 107)	-28	26	76	12.6	17
HDV-28(26)ac2 (配列番号 109)	-28	26	58	10.3	18
HDV-28(28)ac2b (配列番号 113)	-28	28	57	9.5	17
HDV-28(27)ac2c (配列番号 114)	-28	27	59	9.2	16
HDV-28(25)ac2d (配列番号 115)	-28	25	56	10.9	19
HDV-13(13)ac11 (配列番号 116)	-13	13	約 100	2.7	2.7
HDV-13(11)ac11b (配列番号 118)	-13	11	約 100	4.9	4.9

10

20

30

40

【0734】

以上の結果から、導入遺伝子の挿入を成功させるための要点がいくつか判明した。第1に、導入遺伝子の挿入を目的とする場合、本実施例で設計したRZは天然型RZよりも効率が低い。第2に、5'隣接rRNA配列の所定のRZ切断部位（例えば、-28位または-13位）で自己切断させる場合、GIC：5' rRNA配列の長さは、最長のrRNA配列ではなく、それよりも短いrRNA配列を組み込むことによって、導入遺伝子の挿入を向上できるという効果がある（例えば、「ac2」系内での、26ntまたは25ntのrRNAを有するRZバックボーン配列ac2（ac2とac2dの正規化したGFP+細胞の割合はそれぞれ18%と19%であった）と、28ntまたは27ntのrRNA配列を有するRZバックボーン配列ac2（ac2bとac2cの正規化したGFP+細胞の割合はそれぞれ17%と16%であった）との比較）。第3に、上流部位でのRZ切断は、導入遺伝子の挿入効率に影響を及ぼす（例えば、HDV-28（28）もしくはHDV-13（13）において、下の鎖のニックまで伸長したrRNA配列にマッチさせた場合、またはHDV-28（26）もしくはHDV-13（11）において、5'側モジュールのrRNAと下の鎖のニック部位の間のギャップを残すことにより効率を改善した場合に、HDV-13 RZの5'側モジュールは、導入遺伝子の挿入効率の点でHDV-28 RZの5'側モジュールよりも劣っている）。

【0735】

実施例23. GIC：5'側モジュールの組換え

F-ZoAl RTの逆転写酵素コンストラクト（RTC）のmRNA（配列番号19）を、実施例4と同様にして作製した。表20に示すように、5'側モジュールの配列のみが異なるGFP導入遺伝子発現カセットを含む遺伝子挿入コンストラクト（GIC）のRNA（配列番号506：CBhBsi_GPF_GeFo_R4A22）を実施例2と同様にして作製した。

【0736】

50

Lipofectamine Messenger Maxを使用して、RTC mRNAと各種GIC mRNAを1：3のモル比で混合してhTERT RPE-1細胞に共トランスフェクトした。24時間後に、トランスフェクト細胞のプールをFACSで分析して、GFP+細胞の割合（%）を検出した。FACS分析により、各処理で得られたGFP+細胞の割合を測定し、この結果を以下の表15に示した。

【0737】

【表16】

表 15: 組換え GIC: 5'側モジュールの構成要素

5'側モジュール	RZ の前の T7 RNAP 転写産物の 5'リーダー*	GIC: 5'側モ ジュールの rRNA 配列 の長さ	GIC: 5' RZ の自己切断 効率 (%)	GFP 陽性 細胞の割合 (%)
HDV-28(26)gu1 (配列番号 107)	PP7hp	26 nt	60	3.2
HDV-28(28)gu5b (配列番号 121)	PP7hp	28 nt	80	4.4
HDV-28(28)NL (配列番号 121)	なし	28 nt	0	1.8
HDV-28(28)_rzdead (配列番号 126)	PP7hp	28 nt	0	0.44
-28(28) (RZ なし) (配列番号 252)	なし	28 nt	0	0.022
TCARZ-28(28) (配列番号 67)	PP7hp	28 nt	87	3.9
TCA5-28(28) (配列番号 62)	PP7hp	28 nt	89	3.2
TCA5_rzdead (配列番号 63)	PP7hp	28 nt	0	0.29
* PP7hp は、PP7 ファージのコートタンパク質に結合させるための、コンセンサス配列のヘアピンステムループが含まれていることを示す。				

【0738】

以上の結果からいくつかの結論が得られた。第1に、鋳型RNAの上流にrRNAが存在していても、効率的に折り畳まれたRZに組み込まれない場合、導入遺伝子の効率的な挿入は起こらなかった（5'側モジュールの「-28(28)(RZなし)」と、TCA5、TCARZ、デノボ設計したHDV-28バリエーションなどのRZ活性5'側モジュールとの比較）。第2に、T7 RNAP転写産物が、RZによる自己切断を促進する5'リーダー配列を有している場合、5'側モジュールの自己切断型RZ折り畳み構造配列の少なくとも一部によって、導入遺伝子の挿入効率が高くなる（HDV-28(28)NL（リーダーなし）による導入遺伝子の挿入効率と、HDV-28(28)gu5bにおいてPP7ファージのヘアピンのリーダー配列を組み込んで作製したRZ切断型鋳型RNAの同じ配列の比較）。第3に、RZとリーダー配列を有する5'側モジュールにより最適な導入遺伝子挿入効率を達成するには、触媒活性を有するRZが必要とされる（rzdeadとRZ活性型5'側モジュールの比較）。

【0739】

実施例24. 2種のRNA構成要素によるGIS: 5'連結部の忠実度

実施例4と同様にして、逆転写酵素コンストラクト（RTC）のmRNAを調製した。以下の表16の記載に従って、実施例2と同様にして遺伝子挿入コンストラクト（GIC）のRNAを調製した。

【0740】

10

20

30

40

50

【表 17】

表 16: 2 種の RNA 構成要素による 5'連結アッセイに使用した遺伝子挿入コンストラクト(GIC)

RTC mRNA	GIC 識別名	図 20 に示したレーンの記号	5'側モジュールの供給源の生物	3'側モジュールの供給源の生物	GIC の配列番号
なし		A			
メダカ(配列番号 10)	TCA5_OrLa3	B	コクヌストモドキ	メダカ	339
ノドジロシトド (配列番号 19)	TCA5_ZoAl3	C	コクヌストモドキ	ノドジロシトド	340
コクヌストモドキ (配列番号 3)	TCA5_TCB3	D	コクヌストモドキ	コクヌストモドキ	341
コクヌストモドキ、 タグなし (配列番号 5)	TCA5_TCB3	E	コクヌストモドキ	コクヌストモドキ	341
なし		F			
メダカ(配列番号 10)	OrLa5L_OrLa3	G	メダカ	メダカ	343
ノドジロシトド (配列番号 19)	OrLa5L_ZoAl3	H	メダカ	ノドジロシトド	344
コクヌストモドキ (配列番号 3)	OrLa5L_TCB3	I	メダカ	コクヌストモドキ	345
コクヌストモドキ、 タグなし (配列番号 5)	OrLa5L_TCB3	J	メダカ	コクヌストモドキ	345

【0741】

RNAはいずれも、1mMクエン酸ナトリウム (pH6.5) からなる最終バッファー中で調製した。6ウェルプレートの各ウェルあたりのトータルRNAの量を2.5 µgに固定した。トランスフェクション効率コントロールとして、蛍光タンパク質のmRNAを添加する場合 (Uを5moUで100%置換したTrilink社製のmCherry mRNA)、このmRNA 50ngを、RTC mRNAとGIC RNAの混合物に加えた。

【0742】

MessengerMaxの代わりにLipofectamine 3000を使用し、RTCとGICのモル比を1:1としたこと以外は実施例7とほぼ同様にして、RTC mRNAとGIC RNAを293T細胞にトランスフェクトした。各RTC mRNAとともにトランスフェクトした各GIC RNAコンストラクトには、(i) コクヌストモドキA系統またはメダカに由来する5'側モジュール、および(ii) RTタンパク質と同じ生物種に由来し、かつ関連する場合、その生物種の同じレトロエレメント系統に由来する3'側モジュールが含まれていた (例えば、コクヌストモドキB系統のR2由来構成要素に由来するTriCasB RTは、コクヌストモドキA系統のR2由来の5'側モジュール「TCA5」とは異なるTriCasB 3' UTR「TCB」と組み合わせられる)。

【0743】

24時間後、ゲノムDNAを抽出するため、200 µlの変性RIPAバッファー (150mM NaCl、50mM Tris (pH7.5)、1mM EDTA、1%TX-100、0.5%デオキシコール酸Na、0.1%SDSおよび1mM DTT) を用いて細胞ペレットを溶解した。10 µlのRNase Aを添加し、試料を37℃で30分間インキュベートした。次に、5 µlのプロテイナーゼKを加え、試料を50℃で一晩インキュベートした。等量のPCI溶液 (フェノール:クロロホルム:イソアミルアルコール=25:24:1) を加えた。ボルテックスし、5分間遠心分離した後、水層を抽出した。1 µlのグリコゲン (20 µg/µl)、10%体積量の5M塩化ナト

リウムおよび3体積量の100%エタノールを加えた。混合し、 -20°C で30分間インキュベートした後、試料を 4°C で30分間遠心分離した。ゲノムDNAペレットを70%エタノールで3回洗浄した。風乾後、ペレットをTEバッファーに溶解した。500ngのゲノムDNAを用いて、挿入連結部のPCRアッセイを行った。PCR後、6 μl のローディング色素と25 μl のPCR反応液を混合し、混合物の半量を、臭化エチジウムを添加した1×TAEバッファー中の1.2%アガロースゲルの各ウェルにロードした。電気泳動後、BioRad社製molecular imager ChemiDoc XRS+でゲルを撮影した。その画像を図20に示す。

【0744】

上記の分析から、2種のRNA構成要素を含む遺伝子挿入システムを利用して、ヒトゲノムの目的の標的部位に導入遺伝子の全長を挿入できることが示された。

10

【0745】

ノドジロシトドに由来する発現されたRTタンパク質と、同じ生物種に由来するGIC:3'側モジュールRT認識配列とを利用することによって、発現された各種RTタンパク質と、メダカまたはコクヌストモドキB系統に由来するGIC:3'側モジュールRT認識配列とを利用した系よりも、予想されたサイズのPCR産物が多く産生された。また、コクヌストモドキB系統に由来する発現されたRTタンパク質と、同じ生物種に由来するGIC:3'側モジュールRT認識配列とを利用することによって、メダカに由来する発現されたRTタンパク質とGIC:3'側モジュールRT認識配列とを利用した系よりも予想されたサイズのPCR産物が多く産生された。

【0746】

20

2種のGIC:5'側モジュール構成要素の一方を有するGICを使用した各RTCを比較したところ、メダカのR2 5'領域由来の「OrLa5L」と、コクヌストモドキA系統のR2 5'領域由来の「TCA5」は、導入遺伝子の全長を挿入することができたことが示された。これらのGIC:5'側モジュールを、検討した各GIC:3'側モジュールRT認識配列と組み合わせても、結果は変わらなかった。本実施例から、遺伝子挿入システムのRNAのみが送達されたことが示された。

【0747】

実施例25. 複数の細胞株における2種のRNAを含む遺伝子挿入システムの送達

F-ZoAl RTの逆転写酵素コンストラクト(RTC)のmRNA(配列番号19)を、実施例4と同様にして作製した。表17に示すように、GFP導入遺伝子発現カセットを含む遺伝子挿入コンストラクト(GIC)のRNA(TCA5_CBh_NLSGPF_ZoAl3_R4A22(配列番号507)またはTCA5_CBh_NLSGPF_GeFo3_R4A22(配列番号508))を実施例2と同様にして作製した。

30

【0748】

ヒト細胞株であるSK-HEP1、293T、HCT116、hTERT RPE-1、HeLa、Huh7、IMR-90およびHaCaT、サル細胞株であるCos7およびVero、ならびにマウス細胞株であるC2C12を培養し、実施例7と同様にして、RTC mRNAとGIC RNAを1:3のモル比で混合したものを共トランスフェクトした。24時間後に、トランスフェクトした細胞のプールをフローサイトメトリーで分析して、GFP+細胞の割合(%)を検出した。その結果を表17Aおよび表17Bに示した。

40

【0749】

【表18】

表 17: 2 種の RNA を含む遺伝子挿入システムの送達により導入遺伝子を挿入した細胞の種類の一覧

	SK-HEP1	293T	hTERT RPE-1	HeLa	HCT 116	Huh7
GIC: 3' ZoAl	1.15 %	0.19 %	2.12 %	0.26 %	0.36 %	1.02 %

【0750】

50

【表 19】

表 17B:

2 種の RNA を含む遺伝子挿入システムの送達により導入遺伝子を挿入した別の種類の細胞の一覧

	IMR-90	HaCaT	hTERT RPE-1	C2C12	Cos7	Vero
GIC: 3' GeFo	0.52 %	3.26 %	2.59 %	2.77 %	1.08 %	0.52 %

【0751】

いずれの細胞集団でも、少なくともある程度の割合の GFP 発現細胞が示されたことから、RTC と GIC のいずれの組み合わせでも、対象のゲノムへの GFP 発現導入遺伝子の挿入において少なくとも最小限の効果が認められたことが示された。さらに、HeLa や 293T などのヒトがん由来細胞株と比べて、比較的高い割合の GFP+ 細胞が初代 hTERT RPE-1 ヒト細胞株において観察された。

【0752】

複数の細胞株において、2 種の RNA を含む遺伝子挿入システムを送達する追加の実験を行った。F-ZoAl RT をコードする RTC mRNA (N¹-メチルシュードウリジンを用いて作製したもの) を、以下に示す 2 種の異なる遺伝子挿入システム (GIS) RNA 鋳型を別々に用いて共トランスフェクトした。すなわち、RNA 鋳型として、i) 通常のウリジンヌクレオチドで構成された 5' TCA5_RNAPItterm1_sylacO_CBh promoter_eGFP_SV40L PA_sylacO_GeFo3_R4A22、または ii) 導入遺伝子 RNA の発現用の改変 CMV プロモーターを含み、かつシュードUヌクレオチドを含む 5' TCARZ_CMV*promoter_eGFP_minpA_GeFo3_R4A22 のいずれかを使用した。2 種の RNA を送達後の 1 日目 (または 1 日目と 3 日目) に、フローサイトメトリーで導入遺伝子の発現を測定した。GFP+ 細胞の割合 (%) に対する全体的なトランスフェクション効率を比較する手段として、mCherry をコードする mRNA (TriLink 社) を共トランスフェクトした。この結果を以下の表 17C および表 17D に示す。

【0753】

【表 2 0】

表 17C: 通常のウリジンヌクレオチドを含む 5' TCA5_RNAPItem1_sylacO_CBhpromoter_eGFP_SV40LPA_sylacO_GeFo3_R4A22 を RNA 鋳型として使用して、2 種の RNA を含む遺伝子挿入システムの送達により導入遺伝子を挿入した別の種類の細胞の一覧

細胞株	1 日目		3 日目
	GFP%	mCherry%	GFP%
RPEhTERT	20.71	92.657	18.89
ARPE19	19.64	91.32	17.1
293T	0.21	90.866	2.34
HaCat	3.74	84.801	2.47
Hela	0.92	62.78	0.77
Huh7	0	98.12	11.68
IMR90	6.07	75.12	8.51
MRC5	5.42	82.99	5.7
Cos7	3.41	95.18	4.66
Vero	1.91	91.938	2.38
C2C12	9.26	96.98	5.69
G8	1.9	84.338	1.04
C26	1.03	77.744	1.26

10

20

【0 7 5 4】

【表 2 1】

表 17D: 導入遺伝子 RNA の発現用の改変 CMV プロモーターを含み、かつシュード U ヌクレオチドを含む 5' TCARZ_CMV*promoter_eGFP_minpA_GeFo3_R4A22 を RNA 鋳型として使用して、2 種の RNA を含む遺伝子挿入システムの送達により導入遺伝子を挿入した別の種類の細胞の一覧

細胞株	GFP% の平均値	S.D. GFP	GFP 強度 の中央値	mCherry% の平均値	S.D. mCherry
RPE	61.31	0.8386	44537	90.61333	0.1528
ARPE-19	53.57	0.3512	61436	90.57333	0.2517
293T	9.567	0.2003	3511	74.125	0.1732
Hela	11.52	0.1	5827	51.47667	0.6506
IMR90	38.01	0.0577	27271	66.22	0.755
MRC5	40.52	0.1155	28822	71.65333	0.5859
Vero	10.06	0.3	4071	83.20333	0.6429
C2C12	30.53	1.701	5560	78.00667	1.5044

30

40

SD = 標準偏差

【0 7 5 5】

上記のデータから、2 種の RNA を含む遺伝子挿入システムの送達が、ヒト、サルおよびマウスに由来する複数の種類の細胞において機能性を有することが示された。また、上記のデータから、改変 CMV プロモーターとシュード U ヌクレオチドの組み合わせが、導入遺伝子を発現する細胞の割合を増加させることが示された。

【0 7 5 6】

実施例 26. 逆転写酵素コンストラクト (RTC) と遺伝子挿入コンストラクト (GIC) の

50

組み合わせ

hTERT RPE-1細胞株を培養し、ZoAl RT mRNA RTC（配列番号19）、ZoAl RT-dead mRNA RTC（配列番号24）およびTaGu RT mRNA RTC（配列番号28）のいずれか1つと、TCA5_ZoAl3 GIC RNA（配列番号509）、TCA5_GeFo3 GIC RNA（配列番号503）およびTCA5_TaGu3 GIC RNA（配列番号510）のいずれか1つを、RTCとGICの比率を1：3として実施例9と同様にしてトランスフェクトした。

【0757】

5日後に細胞集団を回収し、前述と同様に計数して、GFP陽性細胞の割合とGFP陽性集団におけるGFP強度の中央値を測定し、この結果を以下の表18に報告した。

【0758】

【表22】

表 18: RTC と GIC の組み合わせ

RTC	GIC	GFP 陽性細胞の割合 (%)
F-ZoAl RT mRNA	TCA5_TaGu3	2.38
F-TaGu RT mRNA	TCA5_TaGu3	3.56
F-ZoAl RT mRNA	TCA5_ZoAl3	11.75
F-TaGu RT mRNA	TCA5_ZoAl3	13.28
F-ZoAl RT mRNA	TCA5_GeFo3	11.71
F-TaGu RT mRNA	TCA5_GeFo3	13.87

【0759】

逆転写酵素コンストラクト（RTC）（ZoAl RT mRNAまたはTaGu RT mRNA）と遺伝子挿入コンストラクト（GIC）（TCA5_ZoAl3またはTCA5_GeFo3）の組み合わせを投与することによって、GFP発現細胞が有意に高い割合で得られた。この結果から、導入遺伝子の挿入効率を高くするには、ノドジロシトドまたはガラパゴスフィンチに由来する3'側モジュールRT認識配列を有する遺伝子挿入コンストラクト（GIC）を、ノドジロシトドまたはキンカチョウに由来するRTC：RTモジュールと組み合わせることが好ましいことが示された。さらに、いずれの組み合わせでも、安定な挿入が得られ（挿入部位の5'連結部と3'連結部をPCRにより検出することにより調べた）、導入遺伝子が発現されることが示された。ZoAl RT-dead mRNAをどの遺伝子挿入コンストラクト（GIC）と組み合わせても、バックグラウンドを上回るGFP蛍光は得られなかった。

【0760】

実施例27. 細胞株に応じた逆転写酵素コンストラクト（RTC）と遺伝子挿入コンストラクト（GIC）の比率および3'側モジュール

パートA

上記と同様に、hTERT RPE-1ヒト細胞株、SK-HEP1ヒト細胞株およびHeLaヒト細胞株を培養し、ZoAl RT mRNA RTCと、TCA5_ZoAl3 GIC RNAまたはTCA5_GeFo3 GIC RNAとをトランスフェクトした。

【0761】

5日後に、上記と同様に細胞集団を回収し、計数した。eGFP発現細胞の割合（%）を以下の表19に示す。

【0762】

10

20

30

40

50

【表 2 3】

表 19: RTC と GIC の比率

細胞株	GIC	RTC と GIC の比率					
		RTC なし	1:1	1:3	1:5	1:8	1:10
hTERT RPE-1	TCA5_ZoAl3	0.01%	2.47%	2.8%	2.63%	N.A.	2.3%
hTERT RPE-1	TCA5_GeFo3	0.04%	1.96%	2.48%	2.57%	2.34%	2.28%
SK-HEP1	TCA5_ZoAl3	0.04%	0.38%	0.58%	0.62%	0.64%	0.7%

【0 7 6 3】

10

パート B

上記と同様に、SK-HEP 1細胞株およびHeLa細胞株を培養し、トランスフェクトし、回収し、分析した。その結果を以下の表20に示した。表20に示すように、RTCとGICの比率を様々に変えた。

【0 7 6 4】

【表 2 4】

表 20: RTC と GIC の比率

細胞株	GIC	RTC と GIC の比率					
		RTC なし	3:1	2:1	1:1	1:2	1:3
SK-HEP 1	TCA5_ZoAl3	0.09	1.07	1.58	2.44	3.23	3.60
HeLa	TCA5_ZoAl3	0.04	0.15	0.20	0.26	0.27	0.32

20

【0 7 6 5】

以下の表20Bは、上記と同様に、hTERT RPE-1ヒト細胞を培養し、F-TaGu mRNA RTCまたはF-ZoAl mRNA RTC（いずれも5moUで作製）と、TCA5_ZoAl3 GIC RNAまたはTCA5_GeFo3 GIC RNAとをトランスフェクトした類似の実験の結果を示す。

【0 7 6 6】

30

40

50

【表 2 5】

表 20B: RTC と GIC の比率

RTC mRNA/GIC RNA	モル比:	GFP%	GFP の強度
TaGu/ZoAl3	10:1	6.27	3846
	3:1	11.26	5562
	1:1	14.36	6412
	1:3	14.36	6347
	1:6	14.76	6746
	1:12	13.36	6156
	1:20	11.26	5773
TaGu/GeFo3	10:1	4.5	3405
	3:1	9.1	4330
	1:1	12.21	4841
	1:3	13.91	5146
	1:6	13.31	5413
	1:12	13.51	5600
	1:20	11.61	5323
ZoAl/ZoAl3	10:1	1.89	3014
	3:1	5.35	4540
	1:1	9.96	5676
	1:3	11.46	6347
	1:6	11.06	6521
	1:12	9.96	5911
	1:20	7.82	5487
ZoAl/GeFo3	10:1	0.77	3014
	3:1	4.25	3393
	1:1	8.12	4330
	1:3	10.31	5233
	1:6	10.71	5146
	1:12	9.91	5146
	1:20	7.84	4744

【0767】

最も効果的な導入遺伝子挿入効率を示したRTCとGICの比率はある程度のばらつきがあったが、RTC RNAよりもGIC RNAが多いモル比が最適であった。

【0768】

これらの結果から、2種の構成要素からなる遺伝子挿入システムを介して特定の対象に導入遺伝子を挿入する際の理想的な比率は、構成要素や対象のアイデンティティから予測するのではなく、実験を行うことによって決定する必要があると見られる。hTERT RPE-1細胞に投与することを目的とし、ノドジロシトド由来のRTモジュールを含むRTCと、ノドジロシトド由来のGIC:3'側モジュールRT認識配列を含むGICとを含む遺伝子挿入システムでは、1:3の比率(RTC:GIC)が好ましいと考えられる。hTERT RPE-1細胞に投与することを目的とし、ノドジロシトド由来のRTC:RTモジュールを含むRTCと、ガラパゴスフィンチ由来のGIC:3'側モジュールRT認識配列を含むGICとを含む遺伝子挿入システムでは、1:5の比率(RTC:GIC)が好ましいと考えられる。SK-HEP1細胞

胞またはHeLa細胞に投与することを目的とし、ノドジロシトド由来のRTC：RTモジュールを含むRTCと、ノドジロシトド由来のGIC：RT認識配列を含むGICとを含む遺伝子挿入システムでは、1：3の比率（RTC：GIC）が好ましいと考えられる。

【0769】

実施例29. 導入遺伝子の発現の耐久性

F-ZoAl RTをコードする逆転写酵素コンストラクト（RTC）のmRNA（配列番号19）を、実施例4と同様にして作製した。GFP導入遺伝子発現カセットとしてTCA5_CBh_NLSGFP_ZoAl3（配列番号507）を含む遺伝子挿入コンストラクトのRNAを実施例2と同様にして作製した。

【0770】

上記のRTCとGICを、実施例7で述べた293T細胞培養に共トランスフェクトし、トランスフェクション後の3日目にソートしてGFP+細胞を濃縮した。その翌日に、Fusion Ariaソータープレートホルダーを用いてソートすることにより個々のGFP陽性細胞に分離して、96ウェルプレートの各ウェルに播種した。約3週間増殖させた後、各ウェルの生きたGFP陽性細胞株をスクリーニングし、24ウェルマスタープレートに移し、週に2回分割した。GFP蛍光強度に単一ピークの分布が見られたことから、37個の細胞株がクローン性を有すると見られた（図21）。各細胞株は、GFP陰性クローン細胞株とは明確に区別可能な様々に異なる絶対GFP強度を有していた（図21）。連続培養期間中に、Attune NxT Acoustic Focusingサイトメーターを使用して細胞の分取試料を約週1回スクリーニングした。クローン細胞株として2ヶ月間継代したところ、最初のトランスフェクションからほぼ3ヶ月経過後に、37個の細胞株のうち1つのみが、GFP強度の低下を示したが、その低下は約50%ほどであった。

【0771】

これらの結果から、本発明の遺伝子挿入システムによって哺乳動物細胞のゲノムに挿入された導入遺伝子が、3ヶ月以上にわたって安定に発現できたことが示された。

【0772】

実施例30. 複数の導入遺伝子の挿入および発現

F-ZoAl RTの逆転写酵素コンストラクト（RTC）のmRNA（配列番号19）を、実施例4と同様にして作製した。GFP導入遺伝子発現カセットを含む遺伝子挿入コンストラクト（GIC）RNA（TCA5_CBhBsi_GFP_GeFo3、配列番号503）と、mCherry導入遺伝子発現カセットを含む遺伝子挿入コンストラクト（GIC）RNA（TCA5_CBhBsi_mCherry_GeFo3、配列番号512）を、実施例2と同様に作製した。

【0773】

RTC mRNAと、2種のGICコンストラクトの一方またはこれらの等量混合物とを、RTC mRNAと総GIC鑄型RNAのモル比を1：3として、hTERT RPE-1細胞に共トランスフェクトした。コントロールとして、トランスフェクトしない細胞（陰性コントロール）、RTCのみをトランスフェクトした細胞（RTCコントロール）、またはGFP GICのみまたはmCherry GICのみをトランスフェクトした細胞（GFP鑄型またはmCherry鑄型のみのコントロール）を使用した。3種のGIC（GFP、mCherryまたはこれらの等量混合物）のいずれかとRTCを細胞にトランスフェクトした。24時間後に、フローサイトメトリーにより細胞におけるGFPとmCherryの発現を分析した。目的の導入遺伝子産物を発現する細胞の割合を以下の表21に示した。

【0774】

10

20

30

40

50

【表 2 6】

表 21: 2 種の導入遺伝子の挿入および発現

トランスフェクトした 構成要素	GFP 陽性細胞のみ の割合 (%)	mCherry 陽性細胞 のみの割合 (%)	GFP&mCherry 陽性細胞の割合 (%)
なし	0.0041	0.0055	0.0014
RTC のみ	0.026	0.024	0
GFP GIC のみ	0.043	0.0020	0.0061
mCherry GIC のみ	0.024	0.010	0.017
RTC + GFP GIC	21.7	0.29	0.044
RTC + mCherry GIC	0.3	15.3	0.038
RTC + GFP & mCherry GIC	5.43	3.54	8.94

10

【0 7 7 5】

これらの結果から、本発明の遺伝子挿入システムが、単一の GIC に含まれる 2 種以上の導入遺伝子を対象のゲノムに挿入して、対象の細胞において両方の導入遺伝子を発現させることができることが示された。このことから、単一の GIC を使用することによって、複数の導入遺伝子をゲノムに挿入することができ、その結果、対象の細胞によってペイロードが高発現されると考えられる。複数の導入遺伝子コピーが望ましくない場合、当業者に公知の方法を使用して、最初の導入遺伝子の挿入後にさらなる挿入を停止させることができるネガティブフィードバック機構を導入遺伝子のペイロードに組み込んでもよい。

20

【0 7 7 6】

追加実験として、2 種の GIC 鋳型 RNA を混合して細胞にトランスフェクトする実験と、2 種の導入遺伝子をコードする単一の GIC 鋳型 RNA (「タンデム鋳型」と呼ぶ) を使用した実験を行った。具体的には、F-ZoAl RT (5moU を含む、配列番号 19) をコードする RTC mRNA または N¹-メチルシュードウリジンを含む RT 触媒不活化 ZoAl (「ZoAl RT D」、配列番号 23) をコードする RTC mRNA と、単一の導入遺伝子鋳型としての TCARZ_SV40*_GFP_GeFo3 (配列番号 529) もしくは TCARZ_CMV*_mCherry_GeFo3 (配列番号 31)、またはタンデム鋳型 TCARZ_SV40*_GFP_minPA_CMV*_mCherry_SV40LPA_GeFo3 (配列番号 533) とを細胞に共トランスフェクトした。この結果を以下の表 21B に示す。eGFP+mCherry mRNA は陽性コントロールである。

30

【0 7 7 7】

40

50

【表 2 7】

表 21B. mCherry 陽性細胞、eGFP 陽性細胞または mCherry と eGFP が陽性の細胞の割合

トランスフェクトした 構成要素	mCherry 陽性細胞 のみの割合 (%)	GFP 陽性細胞 のみの割合 (%)	GFP&mCherry 陽性細胞の割合 (%)
ZoAI RTD + CMV-mCherry	0.01 ± 0.004	0.01 ± 0.002	0.0004 ± 0.0004
ZoAI RTD + SV40-eGFP	0.06 ± 0.006	0.005 ± 0.003	0.005 ± 0.002
ZoAI RTD + Tandem SV40- eGFP_CMV-mCherry	0.005 ± 0.002	0.03 ± 0.003	0.008 ± 0.004
ZoAI + CMV-mCherry	69.4 ± 0.88	0.02 ± 0.0009	0.06 ± 0.006
ZoAI + SV40-eGFP	0.002 ± 0.0009	68.0 ± 0.38	0.04 ± 0.005
ZoAI + SV40-eGFP + CMV- mCherry	7.77 ± 0.54	3.3 ± 0.23	52.2 ± 2.8
ZoAI + Tandem SV40- eGFP_CMV-mCherry	12.5 ± 0.20	0.6 ± 0.02	45.6 ± 0.73
eGFP + mCherry mRNA	0.4 ± 0.05	0.3 ± 0.2	97.4 ± 0.2

平均値 ± SEM, n=3

【0 7 7 8】

以上のデータから、2種の導入遺伝子RNAが同じ細胞に挿入されたことから、その2種の導入遺伝子RNAが同じGIC鋳型RNAから送達できることが示された。

【0 7 7 9】

実施例31. 遺伝子挿入システムによる内因性修復機構の動員
パートA－RNA干渉によるMUS81のノックダウン

【0 7 8 0】

F-ZoAIの逆転写酵素コンストラクト(RTC)のmRNA(配列番号19)を、実施例4と同様にして作製した。GFP導入遺伝子発現カセットとしてTCA5_CBhBsi_GFP_GeFo3(配列番号503)を含む遺伝子挿入コンストラクトのRNAを実施例2と同様にして作製した。表22に詳細を示した検証済みの抗MUS81 siRNAと抗MSH2 siRNAをサーモフィッシャーサイエンティフィック社から購入した。Silencer Select Negative Control No. 1 siRNAはインビトロジェン社から購入した。

【0 7 8 1】

10

20

30

40

50

【表 28】

表 22: siRNA 二本鎖の設計

siRNA ID	標的 遺伝子	センス配列	アンチセンス配列
s37038	MUS81	CGCGCUUCGUUUUCAGAAtt (配列番号 553)	UUCUGAAAUACGAAGCGCGtg (配列番号 554)
s37039	MUS81	UGACCUCUCCAAACCCUCUtt (配列番号 555)	AGAGGGUUUGGAGAGGUCAtg (配列番号 556)
s37040	MUS81	GGGAGCACCUGAAUCCUAAtt (配列番号 557)	UUAGGAUUCAGGUGCUCCCgg (配列番号 558)
s8966	MSH2	GGAUAUUACUUUCGUGUAAtt (配列番号 559)	UUACACGAAAGUAAUAUCCaa (配列番号 560)
s8967	MSH2	CGUCGAUUCCCAGAUCUUAAtt (配列番号 561)	UAAGAUCUGGGAAUCGACGaa (配列番号 562)
s8968	MSH2	GAAUCGCAAGGAUAUGAUAtt (配列番号 563)	UAUCAUAUCCUUGCGAUUCtc (配列番号 564)

10

【0782】

20

各siRNA二本鎖は、センス鎖とアンチセンス鎖がアニールされており、オーバーハングを小文字で示している。3種のsiRNA二本鎖を混合して、各siRNA処理に使用した。

【0783】

トランスフェクション用のsiRNAミックスは、37.5 µlのLipofectamine 3000と混合した625 µlのOptiMEM (Gibco社)を含むチューブと、375 pmolのsiRNAと混合した625 µlのOptiMEMを含むチューブを混合することにより調製した。標的用の3種のsiRNAをプールし、陰性コントロールとして375 pmolのSilencer Select Negative Control No. 1 siRNA (インビトロジェン社)を使用した。

【0784】

10分間のインキュベーション後、1.25 mlのsiRNA-脂質複合体混合物をプレートに加え、約450万個のhTERT RPE-1細胞（プレートに接着した場合、約75%コンフルエントに相当）を加えることによって、各ウェルの培地量を10 mlとした（siRNAの最終濃度：37.5 nM）。24時間後に、細胞を1:3の量に分割して約60%コンフルエントまで増殖させ、siRNAの導入の2日後に、2種のRNAの組み合わせを含む遺伝子挿入システムをトランスフェクトした。トランスフェクションの72時間後にqRT-PCRを実施して、標的mRNAのノックダウン効率を測定した。

30

【0785】

まず、hTERT RPE-1細胞に、コントロールとして、抗MUS81 siRNA、抗MSH2 siRNA、またはスクランブルsiRNAをトランスフェクトした。翌日（1）または2日後（2）に、前述の遺伝子挿入システムを細胞にトランスフェクトしないか（陰性コントロール）、遺伝子挿入コンストラクト（GIC）のみをトランスフェクトするか、逆転写酵素コンストラクト（RTC）と遺伝子挿入コンストラクト（GIC）を共トランスフェクトした。

40

【0786】

最後のトランスフェクションの24時間後に、細胞を回収し、実施例8と同様にFACS分析によりGFP発現細胞の割合を測定し、その結果を以下の表23に報告した。

【0787】

50

【表 2 9】

表 23: 遺伝子挿入システムの機能に対する内因性修復機構のノックダウンの効果

トランスフェクトした siRNA	トランスフェクトした 遺伝子挿入システム	1 回目と 2 回目の トランスフェクション の間の日数	GFP 陽性細胞 の割合 (%)
スクランブル	なし	1	0.0016
スクランブル	なし	2	0.073
スクランブル	GIC のみ	1	0.029
スクランブル	GIC のみ	2	0.028
スクランブル	RTC + GIC	1	4.57
スクランブル	RTC + GIC	2	1.48
siMSH2	なし	1	0.024
siMSH2	なし	2	試験せず
siMSH2	GIC のみ	1	0.01
siMSH2	GIC のみ	2	0.017
siMSH2	RTC + GIC	1	4.36
siMSH2	RTC + GIC	2	1.72
siMUS81	なし	1	0.043
siMUS81	なし	2	0.034
siMUS81	GIC のみ	1	0.021
siMUS81	GIC のみ	2	0.06
siMUS81	RTC + GIC	1	3.06
siMUS81	RTC + GIC	2	0.22

【0 7 8 8】

パート B - MUS81 欠損細胞株における遺伝子挿入システムの活性

F-ZoAl RT の逆転写酵素コンストラクト (RTC) の mRNA (配列番号 19) を、実施例 4 と同様にして作製した。GFP 導入遺伝子発現カセットとして TCA5_CBhBsi_GFP_GeFo3 (配列番号 503) を含む遺伝子挿入コンストラクト (GIC) の RNA を実施例 2 と同様にして作製した。

【0 7 8 9】

上記と同様にして、野生型または MUS81 陰性変異型の HTC 116 細胞株に共トランスフェクトした。トランスフェクションの 24 時間後に細胞を回収し、FACS 分析により GFP 発現細胞の割合を測定し、以下の表 24 に報告した。

【0 7 9 0】

【表 3 0】

表 24: MUS81 陰性細胞株における遺伝子挿入システムの活性

細胞株	トランスフェクトした 遺伝子挿入システム	GFP 陽性細胞の割合 (%)
HCT116 MUS81+	なし	0.00677
HCT116 MUS81+	なし	0.028
HCT116 MUS81+	GIC のみ	0.26
HCT116 MUS81-	GIC のみ	0.01
HCT116 MUS81-	RTC + GIC	0.00721
HCT116 MUS81-	RTC + GIC	0.036

【0791】

以上の結果から、本発明の遺伝子挿入システムによる導入遺伝子挿入効率を最大にするには、MUS81の活性が必要であることが示された。したがって、本明細書に記載の遺伝子挿入システムは、内因性のゲノム修復機構（例えば、MUS81）を動員することにより導入遺伝子の挿入を成功させている可能性がある。

【0792】

さらに、ゲノムの修復において機能することが知られている別の酵素であるMSH2の活性を欠失させても、遺伝子挿入システムによる導入遺伝子挿入率は有意に阻害されなかったことから、本発明の遺伝子挿入システムは、MUS81を選択的に動員して導入遺伝子を挿入していると考えられた。MUS81が、天然のレトロエレメントまたは導入遺伝子の挿入機構において機能することは知られていなかったことには注目すべきである。

【0793】

パートC—RNA干渉によるノックダウンとレポーターの共トランスフェクション

F-ZoAl RTをコードする逆転写酵素コンストラクト（RTC）のmRNA（配列番号19）を、実施例4と同様にして作製した。GFP導入遺伝子発現カセット（GIC）としてTCA5_CBhBsi_GFP_GeFo3（配列番号503）を含む遺伝子挿入コンストラクトのRNAを実施例2と同様にして作製した。mCherryをコードするmRNA（TriLink社、#L-7203）、抗MUS81 siRNA（サーモフィッシュャー社のSilencer Select ID番号s37038、s37039およびs37040の等量混合物）、およびSilencer Select Negative Control No. 1 siRNA（インビトロジェン社）を購入した。

【0794】

まず、hTERT RPE-1細胞に、抗MUS81 siRNAまたは陰性コントロールsiRNAをトランスフェクトした。2日後（2）に、遺伝子挿入システムを細胞にトランスフェクトしないか（陰性コントロール）、遺伝子挿入コンストラクト（GIC）のみをトランスフェクトするか、逆転写酵素コンストラクト（RTC）と遺伝子挿入コンストラクト（GIC）とmCherry mRNAを共トランスフェクトした。すべてのトランスフェクションは、Lipofectamine Messenger Maxを使用して行った。このmCherry mRNAは、古典的なキャップ依存性のmRNA翻訳を介して（すなわち遺伝子挿入システムの活性を必要とせずに）mCherryを翻訳するように設計されており、GFP挿入効率が低下した場合のトランスフェクション効率のコントロールとして使用した。

【0795】

最後のトランスフェクションの24時間後に、細胞を回収し、FACS分析によりGFP発現細胞とmCherry発現細胞の割合を測定し、その結果を以下の表25に報告した（該当する場合、バックグラウンドに対するGFP陽性細胞の割合を括弧内に示した）。

【0796】

【表31】

表 25: siRNA による MUS81 のノックダウン

トランスフェクトした iRNA	トランスフェクトした コンストラクト	GFP 陽性細胞 の割合 (%)	GFP 強度 の中央値	mCherry 陽性細胞 の割合 (%)	mCherry 強度の 中央値
スクランブル	なし	0.00279	1968	0.00558	187
スクランブル	GIC のみ	0.025	2299	0.034	321
スクランブル	RTC + GIC + mCherry(mRNA)	3.8	8126	90.666	2307
siMus81	なし	0.069	2163	0.05	291
siMus81	GIC のみ	0.17	2331	0.22	337
siMus81	RTC + GIC + mCherry(mRNA)	0.39	2705	88.78	2460

10

20

30

40

50

【0797】

以上の結果から、パートA、パートBおよびパートCにおいて再現可能に観察された、MUS81の除去または欠失による細胞でのGFP導入遺伝子発現の大きな減少は、細胞へのRNAトランスフェクト能に対するMUS81ノックダウンの効果によるものではなく、遺伝子挿入システムを含むhTERT RPE-1細胞によるトランスフェクトmRNAの翻訳能に対するMUS81ノックダウンの効果によるものでもなかったことが確認された。

【0798】

実施例32. 様々なプロモーターを有する鋳型モジュール

hTERT RPE-1細胞を培養し、N末端に核局在化配列（NLS）を含むか、NLSを含んでおらず、様々な異なる発現コンテキストを有するGFP ORF（配列番号513～517）を含む遺伝子挿入コンストラクト（GIC）と、F-ZoAl RT mRNA逆転写酵素コンストラクト（RTC）（配列番号19）とをトランスフェクトした。転写プロモーターとして、CBh、EFSおよびmPGK（配列番号400～402または408～409）を試験した。ペイロードカセットの転写の方向は、RNAPIと同じ方向またはRNAPIの転写に対して「反転」した逆向き方向であった。「反転」方向では、RNAPIIプロモーターの上流からRNAPI転写終結シグナルカセットが除去されていた。

【0799】

トランスフェクションの翌日と5日後に、FACSによりGFPの合成を観察した（表26A）。比較に際して、いくつかの興味深い点が見出された。まず、RNAPIと同じ方向のCBh_GFPまたはCBh_NLSGFPと、RNAPIと逆向きの方向の[CBh_NLSGFP]flipは、いずれもトランスフェクションの翌日のGFP細胞の割合は同程度であったが、4日後には、RNAPIと逆向きの方向の[CBh_NLSGFP]flipではGFP細胞の割合が減少したのに対し、RNAPIと同じ方向の導入遺伝子では、GFPシグナルが高いまま維持されていた。この結果から、持続的な導入遺伝子の発現では、RNAPIと同じ方向で転写され、かつ／またはRNAPII発現カセットの前にRNAPI転写終結シグナルが配置されていることが好ましいが、一過性発現が所望される場合は、反転した方向の配列が有利であることが示唆された。2点目として、mPGKプロモーターを有するGFP導入遺伝子とEFSプロモーターを有するGFP導入遺伝子の発現が検出できたことから、様々な種類のプロモーターを使用して、生産性の導入遺伝子の発現を誘導できることが示された。

【0800】

【表32】

表 26A: 最適な発現用の導入遺伝子プロモーターとコンテキスト

プロモーター および ORF	GIC の 配列番号	GFP 陽性細胞 の割合 (%) 1 日目	GFP 陽性細胞 の割合 (%) 5 日目
CBh_GFP	513	4.041	4.46
CBh_NLSGFP	514	3.192	2.2
[CBh_NLSGFP]flip	515	3.934	0.6
mPGK_GFP	516	0.614	0.16
EFS_GFP	517	0.963	0.57

【0801】

別の導入遺伝子転写プロモーターを含む遺伝子挿入コンストラクト（GIC）を用いた追加実験を行った。CpG変異とneo3' 5' UTRを有する改変サイトメガロウイルスプロモーター（CMV*、配列番号408）を試験するとともに、改良したTATAボックスを有する改変シミアンウイルス40プロモーター（SV40*、配列番号409）を試験した。これらのプロモーターを遺伝子挿入コンストラクト（GIC）において使用して、GFP発現導入遺伝子を挿入した。ZoAl RTC mRNAと2種のGICコンストラクトの一方を、RTC mRNAと総GIC鋳型RNAのモル比を1：3として、hTERT RPE-1細胞に共トランスフェクトした。24時間後に、フローサイトメトリーにより細胞におけるGFP発現を分析した。目的の導

入遺伝子産物を発現する細胞の割合を以下の表26Bに示した。

【0802】

【表33】

表 26B: 最適な発現用の導入遺伝子プロモーター

プロモーター_レポ ータータンパク質	GIC の 配列番号	通常の U (U))	GFP+細胞 の割合 1 日目	mCherry+細胞 の割合 1 日目
CBh_GFP	513	U	20.7	n.a.
CMV*_GFP	528	U	50.7	n.a.
SV40*_GFP	529	U	44.8	n.a.
CBh_mCherry	512	U	n.a.	19.9
CMV*_mCherry	531	U	n.a.	33.5
SV40*_mCherry	532	U	n.a.	16.6

10

【0803】

実施例33. 挿入部位特異性を決定するための、ゲノムDNAから挿入された導入遺伝子のシーケンシング

F-ZoAl RTの逆転写酵素コンストラクト (RTC) mRNA (配列番号19) またはF-TaGu RTの逆転写酵素コンストラクト (RTC) mRNA (配列番号28) を実施例4と同様にして作製した。5'側モジュールTCA5 (TCA5_CBhBsi_GFP_GeFo3、配列番号503) または5'側モジュールTCARZ (TCARZ_CBhBsi_GFP_GeFo3、配列番号526) を含むGFP導入遺伝子発現カセットを含む遺伝子挿入コンストラクト (GIC) RNAを、実施例2と同様に作製した。

20

【0804】

RTC mRNAとGIC RNAを、RTC mRNAとGIC鋳型RNAのモル比を1:3として、hTERT RPE-1細胞に共トランスフェクトした。24時間後、実施例8と同様にして細胞をソートし、GFP+集団を濃縮した。濃縮したGFP+細胞を回収し、実施例24と同様にしてゲノムDNAを精製した。標準的なライブラリー調製を行うため、カリフォルニア大学バークレー校のFunctional Genomics Laboratoryに1μgのDNAを提出した。また、イルミナ社のホールゲノムショットガン (WGS) シーケンシングを実施するため、Vince nt J. Coates Genomics Sequencing Laboratoryに1μgのDNAを提出した。ヒトWGSは、Kapa Hyper Prep試薬を用いて調製し、ユニークデュアルインデックスY字アダプターを1サイクルのPCRで付加した。シーケンシングは、NovaSeq 6000 S4において、ペアエンドリードを150bpとし、30倍のカバレッジで実施した。

30

【0805】

アダプターをトリミングした後、導入遺伝子配列を含むカスタムコンティグにリードをマッピングした。カスタムコンティグ (配列番号360) の導入遺伝子配列領域にユニークにマッピングされた領域を有し、かつマッピングされていない部分 (「クリッピングされた」部分) も含むリードを、導入遺伝子およびゲノムの連結部配列の候補として評価した。まず、導入遺伝子の3'連結部リード候補を、予想された下流の正確な標的部位に隣接する導入遺伝子配列 (配列番号361) にマッピングして、「標的部位」への挿入 (大部分を占める) をカウントした。次に、正確な標的部位とマッチしなかった3'連結部候補のクリッピングされた領域をヒトrDNAコンセンサス足場全体にマッピングし、不正確に連結されているが、rDNAの標的となる挿入 (「rDNAの不正確な標的部位」) をカウントした。

40

【0806】

rDNAにマッピングされないクリッピング領域を、ヒトゲノムアセンブリGRCh39にマッピングした。ZoAl RTCをトランスフェクトすることにより得られたオフターゲットな

50

挿入連結部リード候補（「不確かなもの」）は、挿入された導入遺伝子の3'末端の特徴を有していなかったことから、広範なシーケンシングライブラリーの増幅中に配列が人工的に再構成されたことが示唆された。オフターゲットな挿入部位は存在しないことが判明した。TaGu RTCのトランスフェクションより得られた7つのオフターゲットな挿入連結部リード候補は、予想された導入遺伝子の3'末端をrDNA以外のヒトゲノム配列に連結したことから、オフターゲットな挿入頻度が最大1%未満であることが示された。

【0807】

【表34】

表 27. ゲノムのシーケンシングに基づく挿入部位の特異性

RTC mRNA (配列番号)	GIC の 配列番号	挿入部位	正確な標的部 位ではない rDNA	不確かなもの (ライブラリー作製時の アーチファクトまたは オフターゲット)
ZoAl (19)	503	531	1	1
ZoAl (19)	526	1033	3	1
TaGu (28)	526	964	5	7

10

【0808】

実施例 34. ウリジン類似体を含む逆転写酵素コンストラクト (RTC) mRNA および遺伝子挿入コンストラクト (GIC) RNA

20

ウリジンヌクレオチドまたは修飾ウリジンヌクレオチドを使用して、F-ZoAl RT の RTC mRNA (配列番号 19) を実施例 4 と同様にして作製した。ウリジンヌクレオチドまたは修飾ウリジンヌクレオチドを使用して、GFP 導入遺伝子発現カセットを含む GIC 鋳型 RNA を実施例 2 と同様にして作製した。各実験の RNA は、ウリジンの 100% が以下に記載のウリジン類似体のいずれかであるか、2 種類のウリジンが記載されている場合は、50% ずつの混合物を含んでいた。以下の表は、ZoAl RT mRNA と、GFP 導入遺伝子発現カセットを含む GIC 鋳型 RNA という 2 種の別々の RNA をトランスフェクションした結果を示す。トランスフェクションの翌日に細胞を回収し、フローサイトメトリーにより GFP 陽性細胞の割合を測定した。

30

【0809】

以下の表 28 は、表に示したウリジン類似体を含む F-ZoAl mRNA と、非修飾ウリジン (ウリジンリボヌクレオチド三リン酸「regU」) を含む GIC 鋳型 RNA である TCA5_CB hBsi_GFP_GeFo3_R4A22 (配列番号 503) のデータを示す。

40

50

【表 3 5】

表 28. ウリジン類似体を含む ZoA1 RT の RTC mRNA

ZoA1 mRNA ウリジン ヌクレオチド	平均 GFP%	S.D.	平均 GFP 強度 の中央値	S.D.
regU	7.17633333	0.79286401	5645.33333	133.881789
regU:N1mpsU 50:50 混合物	13.2296667	0.56862407	9568.66667	520.66528
N1mpsU	13.2296667	0.66583281	9354	733.130957
N1mpsU:psU 50:50 混合物	12.2963333	0.45092498	9160.33333	580.542275
psU	12.463	0.43588989	9086.33333	933.837423
5mU	11.163	0.7	7338.33333	354.425357
5moU	12.9296667	0.37859389	8715.66667	177.902595

略語：ウリジンリボヌクレオチド三リン酸（regU）、5-メトキシウリジンリボヌクレオチド三リン酸（5moU）、5-メチルウリジンリボヌクレオチド三リン酸（5mU）、シュードウリジンリボヌクレオチド三リン酸（psU）、N¹-メチルシュードウリジンリボヌクレオチド三リン酸（N1mpsU）

【0810】

以下の表29は、5moUを含むF-ZoA1 mRNAと、表に示したウリジン類似体を含むGIC鋳型RNAであるTCA5_CBhBsi_GFP_GeFo3_R4A22（配列番号503）のデータを示す。

【表 3 6】

表 29. ウリジン類似体を含む GIC 鋳型 RNA

GIC 鋳型 RNA ウリジン ヌクレオチド	平均 GFP%	SD	平均 GFP 強度 の中央値	SD
regU	12.9296667	0.37859389	8715.66667	177.902595
5moU	1.04333333	0.0321455	1171.66667	41.0528115
5mU	17.81	0.26457513	5845.66667	113.160653
psU	41.44	0.45825757	6311.33333	86.3153134
N1mpsU	30.1433333	0.75055535	3959.66667	307.034743

【0811】

以下の表30は、N¹-メチルシュードウリジンを用いて作製したZoA1 mRNA（配列番号21）と、表30に記載の配列番号で示され、psUを含む6種のGIC鋳型RNA（CBhプロモーター、CMV*プロモーターまたはSV40*プロモーターを有し、GFPまたはmCherryを発現する導入遺伝子）のデータを示す。これらの結果は、表26Bに示した結果と並行して得られた。表30と表26Bの2つの表を比較すると、通常のU鋳型を使用した場合よりも、psU鋳型を使用した場合の方が、導入遺伝子の送達効率が高いことが示された。

【0812】

10

20

30

40

50

【表 3 7】

表 30: 様々なプロモーターをコードする GIC 鋳型 RNA はシュードウリジンの恩恵を受ける

プロモーター_ レポーター タンパク質	GIC の 配列番号	シュード ウリジン (psU)	GFP+細胞 の割合 1 日目	mCherry+細胞 の割合 1 日目
CBh_GFP	513	psU	38	n.a.
CMV*_GFP	528	psU	81.9	n.a.
SV40*_GFP	529	psU	65	n.a.
CBh_mCherry	512	psU	n.a.	42.2
CMV*_mCherry	531	psU	n.a.	70.1
SV40*_mCherry	532	psU	n.a.	51.9

10

【0 8 1 3】

以上の結果から、RTタンパク質をコードするRTC mRNAが修飾ウリジンヌクレオチドを含む場合に、導入遺伝子の発現の増加が観察されることが示された。同様に、GIC鋳型RNAが修飾ウリジンヌクレオチドを含む場合でも、ウリジンがpsUまたはN1mpsUである場合に、導入遺伝子の発現の増加が観察される。

【0 8 1 4】

20

実施例35. ウリジン類似体を含み、切断型GeFo 3' UTRを有するGIC 3' 側モジュールRNAによる導入遺伝子発現細胞の頻度の増加

本実施例は、切断型GeFo 3' UTRを含むGIC 3' 側モジュールとウリジン類似体を含む鋳型RNAによって導入遺伝子発現の頻度が増加することを示す。5moUを用いてF-ZoAl mRNA（配列番号19）を合成し、通常のUまたはシュードUを用いてGIC鋳型RNA（TCARZ_CBh_GFP_GeFo3_R4A22、配列番号526）を合成した。GIC鋳型RNAは、全長3' UTR（GeFo3、配列番号204）を含むか、3種の切断型3' UTR（GeFo217、配列番号222；GeFo98、配列番号223；またはGeFo68、配列番号224）を含んでいた。以下の表31に結果を示す。

【表 3 8】

30

表 31. 切断型 GeFo 3'UTR とシュード U を含む GIC 鋳型 RNA を使用した導入遺伝子の発現

GIC 3' UTR	GFP%	GFP 強度の中央値
GeFo3 regU	8.54	2855
GeFo217 regU	9.7	3129
GeFo98 regU	10.4	3846
GeFo68 regU	9.33	3487
GeFo3 psU	25.86	2642
GeFo217 psU	28.16	2875
GeFo98 psU	29.26	3558
GeFo68 psU	17.96	2001

40

【0 8 1 5】

以上のデータから、切断型3' UTRを含むGIC RNA鋳型により、全長3' UTRと比べて、機能性導入遺伝子タンパク質を発現する細胞の頻度が増加したことが示された。また、以上のデータから、シュードUを含むGIC RNA鋳型により、通常のUを用いて合成した鋳型と比べて、機能性導入遺伝子タンパク質を発現する細胞の頻度が増加したことが示された。

50

配列表

RNA配列が「T」を含む場合、実際のリボヌクレオチドは「U」である。

1. R2A TriCasB* N-DYKプラスミドORF配列: [* = 天然のアミノ酸から改変]

ATGGATTACAAGGATGACGACGATAAGGGTACCCTGCCTTTTCAGAGCAGAAGCTG
CGGCATCTGTCTGAATGCCGGTAAAGGTAACCTCCGCGCGCTGAGCCTGGACGACG
AAGAACGGCACCTGCGTGAACGGCATCCACTGTCCCTGATCCTCTATAAATGCAGC
GATTGCAAGGGCCAGTACAGATCCAAGAGGGCCGCCCTGTGCCACGCCCCCAAGTG
CACCGGACCGACCCCTGATCCTCAAGGCAATGCCCTGCGTTGTCATCTGTGTGGTC
TTGTTTGTAAAAGCCAGAGTGGTGTACCCAGCATTTACGTCATAGACACCCTCTG
GTCAGAAACACCCAGCGGGCAGCTGAAGAAAGCGGTAGAGCCGAACGTGCTGCACT
GCCTCGGCCTCTGCGTCGTAACACCCGTTTCGGTTTTTCAGCGAAGAGGATGAAGCAA
AAATGCTGGAGCTGGAAGTGCGGTTCCAGAACGAGCGTTGTGTTGCAAAATGCATG
CTGCCGTTTTTTCCGAATAGAACTTGCAAGCAGATCCGTGATAAGCGTAATACCGA
TGCATATAAACGGAGAAGAGAACTGTACTTCGAGGGCGTCCGGGTGCAGGACCCTG
CAGGCGCCGAGGACAGCGTTCTCCCTGTTGTGGAAACCGACGAACCCGCCGAGGAA
AATATTCCGCTGGAGTACCCCGAGCTGCCTGGCGATGAAGAGGGTGCTCCTGCCTG
CAGCCAGACTATTCTGAACACAGAAGGTCCGGATGGACTGGGCAGCCCACCGGTGC
CCGTTGAAGAAGAAATGGCAAGTTCGGGTAGCACATCTAATAACGTGGATACCGGT
TGGAGAGAAAGCATTATCACAGCTGCACTCGGCGTTGAAATTCCGAAAGCAATCAG
CCAAGAGCCCGCCGCCGTTATTCAGGAGCTGCAGGATGCTCTGCGCGAGGGCCGTGA
TCGGCGTGTTTTCCGCAGGATCGCCTGGACGAGATGTACGAGCGGGTACTCAAAGTG
GTCAACCCGGATGATACACAGGAACGCCCGAAACGTCAAAGAAAGAAAGGCAAGTC
TCGTAATGCCTTCCGCCGTTATGTGTACAGCCAGACCCAGGACCTGTTCAAGAAAA
ATCCTGGACAGCTGGCACGTTATGTTAGGGAAGATGTGAGATGGCTGGAACAGGGC
CGGGTGCAAGTTGCAGAGAGATGATATTGAAAGAATGTACAACAAGCTGTGGGGCAC
CAAGCCGGATGTGCTGCCTCCCCACTGGGATTATCCACTGCCACTCGACACCGCTG
ATGTTCTGACCCCGATTGAGCTGAAAGAAGTCCGGAAAAGAATATCTCAGACGAAA
CTCAAATCTGCAGCAGGGCCCCGACGGTCTGCAGAAAAGACATCTGGTGCGTCGCGT
TGTGCAAGAAATTCTGCGCCTGCTGTATAACCTGCTGATGTGTTGTGCAATGCAGC
CTACACAGTGGCGGATGAACCGTACCCAACCTGTTACTGAAGCAGGGTAAAGATCCT
CTGGATGTCGCTAGTTATCGTCCTATAACCATCTCCAGCATCCTTTGTCGTCTGTAC
TGGGGTATAATCGACCAGAAGCTGCGTGAGCATGTTTCGTTTTCCACCCACGTCAGAA
AGGCTTCGTGAGCGAGGCAGGTTGTTTTAATAATGTGCAAATTCTGAATGAACTGC
TGCCTCACAGCAAGGGCCAGCACAAAAATCTGGTTGCCGTGTGCCTGGATGTTTCT
AAAGCATTTGATACCGTTCCTCATAGCATCCTCGGCCCCGCCCTGCGCATGAAGGG
CCTGCCGGAACAGGTCGTTTCGTCTGGTTGAAGATAGCTACAAAGATCTGCATACTG
TCGTTAAACAGGGGACCGCAGAAGTGACGCTGAGCCTGCAGCGTGGAGTGAAGCAG
GGCGACCCCTGAGCCCCTTCCTGTTCAACGCCGTGCTGGAGCCGCTGCTGCTGCA
GCTGGAAAGCCATCCTGGTTATAAAGTGGGCGGTGAACTGGCCTCTGTTAGCTGTA
TGGCCTTTGCAGATGATATCTTTCTGATTGCAGCTAATGTTCCGCAGGCCTGTACCC
TGCTGAGGGTCACGGAAGATTATCTGGAAAGACTGGGCATGCGTATCAGCGCCCCT
AAATGTACCAGCTTTGAAATCCGTCCGACCAAAGATAGCTGGTATGTTGCAGATCC
GGGGCTTACACTGACCAAAGGAGAACGTATCCCTGTCGCTGCTGTGGATGCCGTTT
TTAGCTACCTGGGTGTTGAAATTAGCCCTTGGGCAGGTATCACCAGCGAGGGCATC
GAACGGGATTGGCGGGGTACACTGCATCGTGTGCAACGCCTGCCGCTGAAGCCCCA
CCAGAACTGGAAGTATCAGCAGATACCTGGTTCCCTCATTTTCTGTATAAACTGG
TGGTGACCATCCCTAGCATAACCCCTGATTAGACAGCTGGATCAGGAACTGCGGGTT
GTGGTGAAGCAGATCTGTCATCTGCCTCAGAGCACCGCCGACGGCATGATCTATTG
TCGGAGAGTGGACGGCGGTCTGGGTATTCCGAAGCTGGAAATTGTTACCGTGACCA

10

20

30

40

50

GCATACTGAAAGCAGGCCTGAAATTTAGAGATAGCCAGGACAAAATCATGCAGGCA
CTCTGGCTGGCATCAGGTATGAGCAGCCGTCTGAACAGCCTGGCCAAGGCGACCA
AGTACAACCTTGGCCCCGAACAATATTAAAGATCTGGACAGACATAAAGTTGCTC
GTAAGAAAGAAGAACTGGCCCCGATGGGCCAGTTTGACCAGCCAGGGTAAAAGCGTG
AAAAGCTTCGCCGGCAGCCGTACCGCCAATGCATGGCTGATTAACAAGAAGTTACT
GAAGCCCTCTACCTTTATCAGCGCCTTAAGACTGAGAGGCAATGTCGCTGGAGACC
GTGTGGCCCTGAATAGAGCAATCCCGCAGGCCAACCTGATGTGCAGACGTTGCGGT
AGCCAGAGGGAAACTCTGGGCCACATCCTGGGTATCTGTACCAGCACCAAAGCCCT
ACGTATTTACGCCATGATGAGATCAAGAATCTGATCGTGGACGAAGCAGCAAAGA
AGGACGACGAAGTGCTGTTACACTGGAGCCAACCATTCGTCACCCTGTTCTGTTGGT
AACCTGAAACCGGACCTGGTGGTTCAAAACAGAGAAGGCGTGTACGTTGTTGACGT
GACAGTGAGACACGAGGATGGCAACCTGCTTGACACAGGGACGTCAGGATAAACTGG
ACAAGTACGAAGTGCTGCTGCCGATTCTGCAAGAAAGACTGGGTGCTCCTACCGGT
GAGGTTCTGCCGATTGTTGTTGGCACCCGTGGCGCCATGCCTAAAGAGACAGTGGA
AGCCTTGAAGAACTGCGCATTACCGACCGGCAGACCCTGCTCACGATCAGCCTGA
TTGCCCTGAGAATGTCTGTGAAAATTTATCATACCTTCATGGACTATGCAAACGCC
AGACCGCGTCCGGGGCGGCGGTGCAAACCTACCCCCACAGATGATAA

10

2. R2A TriCasB* N-DYKアミノ酸配列: [* = 天然のアミノ酸から改変]

MDYKDDDDKGTLPFQSRSCGICLNAGKGNFRALSLDDEERHLRERHPLSLILYKCS
CKGQYRSKRAALCHAPKCTGPTDPQGNALRCHLCGLVCKSQSGVTQHLRHRHPLV
RNTQRAAEESGRAERAALPRPLRRNTRSVFSEEDEAKMLELEVRFQNERCVAKCM
FFPNRTCKQIRDKRNTDAYKRRRELYFEGVRVQDPAGAEDSVLPVVETDEPAEENIP
LEYPELPGDEEGAPACSQTILNTEGPDGLGSPVPVEEEMASSGSTSNNVD TGWRES
IITAALGVEIPKAISQEPAAVIQELQDALREAVIGVFPQDR LDEMYERVLKVVNPDDT
QERPQRQRKKGKSRNAFRRYVYSQTQDLFKKNPGQLARYVREDVRWLEQGRVQLQ
RDDIERMYNKLWGTKPDVLP PHWDYPLPLDTADVLTPIELKEVRKRISQTKLKSAA
GPDGLQKRHLVRRVVQEILRLLYNLLMCCAMQPTQWRMNRTQLLLKQGDPLDVA
SYRPITISSILCRLYWGIIDQKLREHVRFHPRQKGFVSEAGCFNNVQILNELLRHSKG
QHKNLVAVCLDVSKAFDTPVPHSILGPALRMKGLPEQVVRLVEDSYKDLHTVVKQGT
AEVTL SLQRGVKQGDPLSPFLFNAVLEPLLLQLESHPGYKVGGELASVSCMAFADDI
FLIAANVPQACTLLRVTE DYLERLGM RISAPKCTSF EIRPTKDSWYVADPGLTLTKGE
RIPVA AVDAVFSYLGVEISPWAGITSEGI ERDWRGTLHRVQRLPLKPHQKLELISRYL
VPHFLYKL VVTIPSITLIRQLDQELRVVVKQICHLPQSTADGMIYCRRVDGGLGIPKL
EIVTVTSILKAGLKFRDSQDKIMQALWLASGMSSRLNSLAKATRVQPWPPNNIKDL
RHKVARKKEELARWASLTSQGKSVKSFAGSRTANAWLINKKLLKPSTFISALRLRGN
VAGDRVALNRAIPQANLMCRRCGSQRETLGHILGICTSTKALRISRHDEIKNLIVDEA
AKKDDEVAVTLEPTIRHPVRGNLKPDLVVQNREGVYVVDVTVRHEDGNLLAQGRQD
KLDKYEVL LPIQLERLGAPTGEVLPIVVGTRGAMPKETVEALKKL RITDRQTLTISL
IALRMSVKIYHTFMDYANARPRPGGGANYPHR**

20

30

3. R2A TriCasB* N-DYKトランスフェクトしたmRNA配列: [* = 天然のアミノ酸から改変]

40

AGGAUAAACUAGUAUUCUUCUGGUCCCCACAGACUCAGAGAGAACCCGCCACCA
UGGAUUACAAGGAUGACGACGAUAAGGGAUCCUGCCUUUUCAGAGCAGAAGCUG
CGGCAUCUGUCUGAAUGCCGGUAAAGGUAACUUCGCGCGCUGAGCCUGGACGAC
GAAGAACGGCACCU GCGUGAACGGCAUCCACUGUCCUGAUCCUCUAUAAAUGCA
GCGAUUGCAAGGGCCAGUACAGAUCCAAGAGGGCCGCCUGUGCCACGCCCCCAA
GUGCACCGGACCGACCCUGAUCCUCAAGGCAAUGCCUGCGUUGUCAUCUGUGU
GGUCUUGUUUGUAAAAGCCAGAGUGGUGUUAACCCAGCAUUUACGUCAUAGACACC
CUCUGGUCAGAAACACCCAGCGGGCAGCUGAAGAAAGCGGUAGAGCCGAACGUGC
UGCACUGCCUCGGCCUCUGCGUCGUAACACCCGUUCGGUUUUCAGCGAAGAGGAU

50

GAAGCAAAA AUGCUGGAGCUGGAAGUGCGGUUCCAGAACGAGCGUUGUGUUGCAA
AAUGCAUGCUGCCGUUUUUUCCGAAUAGAACUUGCAAGCAGAUCCGUGAUAAAGCG
UAAUACCGAUGCAUAUAAACGGAGAAGAGAACUGUACUUCGAGGGCGUCCGGGUG
CAGGACCCUGCAGGCGCCGAGGACAGCGUUCUCCUGUUGUGGAAACCGACGAAC
CCGCCGAGGAAAUAUUCGCGUGGAGUACCCCGAGCUGCCUGGCGAUGAAGAGGG
UGCUCUGCCUGCAGCCAGACUAUUCUGAACACAGAAGGUCCGGAUGGACUGGGC
AGCCACCGGUGCCCGUUGAAGAAGAAAUGGCAAGUUCGGGUAGCACAUUAUA
ACGUGGAUACCGGUUGGAGAGAAAGCAUUAUCACAGCUGCACUCGGCGUUGAAAU
UCCGAAAGCAAUCAGCCAAGAGCCCCGCCGCGGUUAUUCAGGAGCUGCAGGAUGCU
CUGCGCGAGGCCGUGAUCGGCGUGUUUCCGCGAGGAUCGCCUGGACGAGAUGUACG
AGCGGGUACUCAAGUGGUCAACCCGGAUGAUACACAGGAACGCCCGAAACGUCA
AAGAAAGAAAGGCAAGUCUCGUAAUGCCUUCGCGCGUUAUGUGUACAGCCAGACC
CAGGACCUGUUCAAGAAAAAUCCUGGACAGCUGGCACGUUAUGUUAGGGGAAGAUG
UGAGAUGGCUGGAACAGGGCCGGGUGCAGUUGCAGAGAGAUGAUUAUUGAAAGAAU
GUACAACAAGCUGUGGGGACCAAGCCGGAUGUGCUGCCUCCCCACUGGGAUUAU
CCACUGCCACUCGACACCGCUGAUGUUCUGACCCCGAUUGAGCUGAAAGAAGUCC
GGAAAAGAAUAUCUCAGACGAAACUCAAUUCUGCAGCAGGGCCCGACGGUCUGCA
GAAAAGACAUCUGGUGCGUCGCGUUGUGCAAGAAAUUCUGCGCCUGCUGUAUAAC
CUGCUGAUGUGUUGUGCAAUGCAGCCUACACAGUGGCGGAUGAACCGUACCCAAC
UGUUACUGAAGCAGGGUAAAGAUAUCCUCUGGAUGUGCGUAGUUUAUCGUCCUAUAAC
CAUCUCCAGCAUCCUUGUGUCGUCUGUACUGGGGUAAUAUCGACCAGAAGCUGCGU
GAGCAUGUUCGUUCCACCCACGUCAGAAAGGCUUCGUGAGCGAGGCAGGUUGUU
UUAUAUAUGUGCAAUUCUGAAUGAACUGCUGCGUCACAGCAAGGGCCAGCACAA
AAUUCUGGUUGCCGUGUGCCUGGAUGUUUCUAAAGCAUUAUGAUACCGUCCUCAU
AGCAUCCUCGGCCCCCGCCUGCGCAUGAAGGGCCUGCCGGAACAGGUUCGUUCGUC
UGGUUGAAGAUAGCUACAAAGAUUCUGCAUACUGUCGUUAAACAGGGGACCGCAGA
AGUGACGCUGAGCCUGCAGCGUGGAGUGAAGCAGGGCGACCCCCUGAGCCCCUUC
CUGUUCAACGCCGUGCUGGAGCCGUGCUGCUGCAGCUGGAAAGCCAUCCUGGUU
AUAAGUGGGCGGUGAACUGGCCUCUGUUAAGCUGUAUGGCCUUGCAGAUGAUUA
CUUUCUGAUUGCAGCUAAUGUUCGCGAGGCCUGUACCCUGCUGAGGGUACGGAA
GAUUAUCUGGAAAGACUGGGCAUGCGUAUCAGCGCCCCUAAAUUGUACCAAGCUUUG
AAUCCGUCCGACCAAAGAUAGCUGGUAUGUUGCAGAUCCGGGGCUUACACUGAC
CAAAGGAGAACGUAUCCUGUCGCGUGCUGUGGAUGCCGUUUUAGCUACCUGGGU
GUUGAAAUUAGCCCUUGGGCAGGUUAUACCCAGCGAGGGCAUCGAACGGGAUUGGC
GGGGUACACUGCAUCGUGUGCAACGCCUGCCGUGAAGCCCCACCAGAAACUGGA
ACUGAUCAGCAGAUACCUGGUUCCUCAUUUUCUGUAUAAACUGGUGGUGACCAUC
CCUAGCAUAACCCUGAUUAGACAGCUGGAUCAGGAACUGCGGGUUGUGGUGAAGC
AGAUCUGUCAUCUGCCUCAGAGCACCGCCGACGGCAUGAUUAUUGUCGGAGAGU
GGACGGCGGUCUGGGUAUUCGGAAGCUGGAAAUUGUUAACCGUGACCAGCAUACUG
AAAGCAGGCCUGAAAUUAGAGAUAGCCAGGACAAAUAUGCAGGCACUCUGGC
UGGCAUCAGGUUAUGAGCAGCCGUCUGAACAGCCUGGCCAAGGCGACCAGAGUACA
ACCUUGGCCCCCGAACAAUAUUAAGAUUCUGGACAGACAUAAAGUUGCUCGUAAAG
AAAGAAGAACUGGCCCGAUGGGCCAGUUUGACCAGCCAGGGUAAAAGCGUGAAAA
GCUUCGCCGGCAGCCGUACCGCCAAUGCAUGGCUGAUUAACAAGAAGUUAUGAA
GCCUCUACCUUUAUCAGCGCCUUAAGACUGAGAGGCAUUGUCGUGGAGACCGU
GUGGCCUGAAUAGAGCAAUCCCGCAGGCCAACCUGAUGUGCAGACGUUGCGGUA
GCCAGAGGGAAACUCUGGGCCACAUCUGGGUUAUCUGUACCAGCACCAAAGCCCU
ACGUAAUUCACGCCAUGAUGAGAUCAAGAAUCUGAUCGUGGACGAAGCAGCAAAG
AAGGACGACGAAGUGGCUGUUAACACUGGAGCCAACCAUUCGUCACCCUGUUCGUG
GUAACCUGAAACCGGACCUGGUGGUUCAAACAGAGAAGGCGUGUACGUUGUUGA

10

20

30

40

50

CGUGACAGUGAGACACGAGGAUGGCAACCUGCUUGCACAGGGGACGUCAGGAUAAA
CUGGACAAGUACGAAGUGCUGCUGCCGAUUCUGCAAGAAAGACUGGGUGCUCCUA
CCGGUGAGGUUCUGCCGAUUGUUGUUGGCACCCGUGGGCGCCAUGCCUAAAGAGAC
AGUGGAAGCCUUGAAGAAACUGCGCAUUACCGACCGGCAGACCCUGCUCACGAUC
AGCCUGAUUGCCCUGAGAAUGUCUGUGAAAAUUUAUCAUACCUUCAUGGACUAUG
CAAACGCCAGACCGCGUCCGGGCGGGCGGUGCAAACUACCCCCACAGAUGAUGACC
UCGAGCUGGUACUGCAUGCACGCAAUGCUAGCUGCCCCUUCUCCGUGCCUGGGUAC
CCCGAGUCUCCCCGACCUCGGGUCCCAGGUAUGCUCCCACCUCACCUGCCCCAC
UCACCACCUCUGCUAGUUC CAGACACCUC CCAAGCACGCAGCAAUGCAGCUCAA
ACGCUUAGCCUAGCCACACCCCCACGGGAAACAGCAGUGAUUAACCUUUAGCAAU
AAACGAAAGUUUAACUAAGCUAUACUAACCC CAGGGUUGGUCAAUUUCGUGCCAG
CCACACCCUGGAGCUAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCAUAUG
ACUAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAA

10

4. R2A TriCasB*アミノ酸配列: [* = 天然のアミノ酸から改変]

MLPFQSRSCGICLNAGKGNFRALSLDDEERHLRERHPLSLILYKCS DCKGQYRSKRA
ALCHAPKCTGPTPDPQGNALRCHLCGLVCKSQSGVTQH LRHRHPLVRNTQR AAEES
GRAERAALPRPLRRNTRS VFSEED EAKMLELEVRFQNERC VAKCMLPFFPNRTCKQI
RDKRNTDAYKRRRELYFEGVRVQDPAGAEDSVLPVVETDEPAEENIPLEYPELPGDE
EGAPACSQTILNTEGPDGLGSPVPVVEEEMASSGSTSNNVD TGWRESIITAALGVEIP
KAISQEPAAVIQELQDALREAVIGVFPQDRLDEMYERVLKVVNPDDTQERP KRQRKK
GKS RNAFR RYVYSQTQDLFKKNPGQLARYVREDVRWLEQGRVQLQRDDIERMYNKL
WGTKPDVLP PHWDYPLPLDTADVLTPIELKEVRKRISQTKLKSAAGPDGLQKRHLV
RRVVQEILRLLYNLLMCCAMQPTQWRMNRTQLLLKQGKDPLDVASYRPITISSILCR
LYWGIIDQKLREHVRFHPRQKGFVSEAGCFNNVQILNELLRHSKGQHKNLVAVCLD
VSKAFDTPHSILGPALRMKGLPEQVVRLVEDSYKDLHTVVKQGTAEVTL SLQRGV
KQGDPLSPFLFNAVLEPLLLQLESHPGYKVG GELASVSCMAFADDIFLIAANVPQAC
TLLRVTEDYLERLGMRISAPKCTSFEIRPTKDSWYVADPGLTLTKGERIPVA AVDAV
FSYLGVEISPWAGITSEGIERDWRGTLHRVQRLPLKPHQKLELISRYLVPHFLYKL VV
TIPSITLIRQLDQELRVVVKQICHLPQSTADGMIYCRRVDGGLGIPKLEIVTVTSILKA
GLKFRDSQDKIMQALWLASGMSSRLNSLAKATRVQPWP PNNIKDLDRHKVARKKEE
LARWASLTSQGKSVKSFAGSRTANAWLINKKLLKPSTFISALRLRGNVAGDRVALNR
AIPQANLMCRRCGSQRET LGHILGICTSTKALRISRHDEIKNLIVDEAAKKDDEVAVT
LEPTIRHPVRGNLKPDLVVQNREGVYVVDVTVRHEDGNLLAQGRQDKLDKYEVLLP
ILQERLGAPTGEVLPIVVGTRGAMPKETVEALKKL RITDRQTLTISLIALRMSVKIY
HTFMDYANARPRPGGGANYPHR**

20

30

5. R2A TriCasB*トランスフェクトしたmRNA配列: [* = 天然のアミノ酸から改変]

AGGAAUAAACUAGUAUUCUUCUGGUCCCCACAGACUCAGAGAGAACCCGCCACCA
UGCUGCCUUUUCAGAGCAGAAGCUGCGGCAUCUGUCUGAAUGCCGGUAAAGGUAA
CUUCCGCGCGCUGAGCCUGGACGACGAAGAACGGCACCUGCGUGAACGGCAUCCA
CUGUCCCUGAUCCUCUAUAAAUGCAGCGAUUGCAAGGGCCAGUACAGAUCCAAGA
GGGCCGCCUGUGCCACGCCCCCAAGUGCACCGGACCGACCCUGAUCCUCAAGGC
AAUGCCCUGCGUUGUCAUCUGUGUGGUCUUGUUUGUAAAAGCCAGAGUGGUGUU
ACCCAGCAUUUACGUCAUAGACACCUCUGGUCAGAAACACCCAGCGGGCAGCUG
AAGAAAGCGGUAGAGCCGAACGUGCUGCACUGCCUCGGCCUCUGCGUCGUAACAC
CCGUUCGGUUUUCAGCGAAGAGGAUGAAGCAAAAAUGCUGGAGCUGGAAGUGCGG
UUC CAGAACGAGCGUUGUGUUGCAAAAUGCAUGCUGCCGUUUUUUCCGAAUAGAA
CUUGCAAGCAGAUCCGUGAUAAAGCGUAAUACCGAUGCAUAUAAACGGAGAAGAGA
ACUGUACUUCGAGGGCGUCCGGGUGCAGGACCCUGCAGGCGCCGAGGACAGCGUU
CUCCCUGUUGUGGAAACCGACGAACCCGCCGAGGAAAAUAUUCGCGUGGAGUACC

40

50

CCGAGCUGCCUGGCGAUGAAGAGGGUGCUCCUGCCUGCAGCCAGACUAUUCUGAA
CACAGAAGGUCCGGAUGGACUGGGCAGCCCACCGGUGCCCGUUGAAGAAGAAUG
GCAAGUUCGGGUAGCACAUCAAUAACGUGGAUACCGGUUGGAGAGAAAGCAUUA
UCACAGCUGCACUCGGCGUUGAAAUUCCGAAAGCAAUCAGCCAAGAGCCCGCCGC
CGUUAUUCAGGAGCUGCAGGAUGCUCUGCGCGAGGCCGUGAUCGGCGUGUUCCG
CAGGAUCGCCUGGACGAGAUGUACGAGCGGGUACUCAAGUGGUCAACCCGGAUG
AUACACAGGAACGCCCGAAACGUCAAAGAAAGAAAGGCAAGUCUCGUAUUGCCUU
CCGCCGUUAUGUGUACAGCCAGACCCAGGACCUGUUCAAGAAAAAUCCUGGACAG
CUGGCACGUUAUGUUAGGGAAGAUGUGAGAUGGCUGGAACAGGGCCGGGUGCAGU
UGCAGAGAGAUGAUUUGAAAGAAUGUACAACAAGCUGUGGGGCACCAAGCCGGA
UGUGCUGCCUCCCCACUGGGAUUAUCCACUGCCACUCGACACCGCUGAUGUUCUG
ACCCCGAUUGAGCUGAAAGAAGUCCGGAAAAGAAUAUCUCAGACGAAACUCAAAU
CUGCAGCAGGGCCCGACGGUCUGCAGAAAAGACAUCUGGUGCGUCGCGUUGUGCA
AGAAAUUCUGCGCCUGCUGUAUAACCGUCUGAUGUGUUGUGCAAUGCAGCCUACA
CAGUGGCGGAUGAACCUGUACCCAACUGUUAUCUGAAGCAGGGUAAAGAUAUCCUCUGG
AUGUCGCUAGUUAUCGUCCUAUAACCAUCUCCAGCAUCCUUUGUCGUCUGUACUG
GGGUAAAUUCGACCAGAAGCUGCGUGAGCAUGUUCGUUUCACCCACGUCAGAAA
GGCUUCGUGAGCGAGGCAGGUUGUUUAAUAAUGUGCAAUUCUGAAUGAACUGC
UGCGUCACAGCAAGGGGCCAGCACAAAAAUUCUGGUUGCCGUGUGCCUGGAUGUUUC
UAAAGCAUUUGAUACCGUUCUCAUAGCAUCCUCGGCCCCGCCUGCGCAUGAAG
GGCCUGCCGGAACAGGUCGUUCGUCUGGUUGAAGAUAGCUACAAAGAUUCUGCAUA
CUGUCGUUAAACAGGGGACCGCAGAAGUGACGCUGAGCCUGCAGCGUGGAGUGAA
GCAGGGCGACCCCCUGAGCCCCUCCUGUUAACGCCGUGCUGGAGCCGCGUCUG
CUGCAGCUGGAAAGCCAUCCUGGUUAUAAAGUGGGCGGUGAACUGGCCUCUGUUA
GCUGUAUGGCCUUUGCAGAUGAUUUCUUCUGAUUUGCAGCUAAUGUUCGCGAGGC
CUGUACCCUGCUGAGGGGUCACGGAAGAUUAUCUGGAAAGACUGGGCAUGCGUAUC
AGCGCCCCUAAUUGUACCAGCUUUGAAAUCCGUCCGACCAAAGAUAGCUGGUUUG
UUGCAGAUCCGGGGCUUACACUGACCAAAGGAGAACGUUACCCUGUCGCGUGCUGU
GGAUGCCGUUUUAGCUACCUGGGUGUUGAAAUAAGCCCUUGGGCAGGUUACACC
AGCGAGGGCAUUCGAACGGGAUUGGCGGGGUACACUGCAUCGUGUGCAACGCCUGC
CGCUGAAGCCCCACCAGAAACUGGAACUGAUCAGCAGAUACCUGGUUCCUCAUUU
UCUGUAUAAACUGGUGGUGACCAUCCCUAGCAUAACCCUGAUUAGACAGCUGGAU
CAGGAACUGCGGGUUGUGGUGAAGCAGAUUCUGUCAUCUGCCUCAGAGCACCGCCG
ACGGCAUGAUCUAUUGUCGGAGAGUGGACGGCGGUCUGGGUAUUCGGAAGCUGGA
AAUUGUUACCGUGACCAGCAUACUGAAAGCAGGCCUGAAAUUUAGAGAUAGCCAG
GACAAAAUCAUGCAGGCACUCUGGCUGGCAUCAGGUUAUGAGCAGCCGUCUGAACA
GCCUGGCCAAGGCGACCAGAGUACAACCUUGGCCCCCGAACAAUUAUAAAGAUUCU
GGACAGACAUAAAGUUGCUCGUAAGAAAGAAAGAACUGGGCCCGAUGGGCCAGUUUG
ACCAGCCAGGGUAAAAGCGUGAAAAGCUUCGCCGGCAGCCGUACCGCCAAUGCAU
GGCUGAUUAACAAGAAGUUACUGAAGCCCUCUACCUUUUAUCAGCGCCUUAAAGACU
GAGAGGCAAUGUCGCGUGGAGACCGUGUGGGCCUGAAUAGAGCAAUCCCGCAGGCC
AACCUGAUGUGCAGACGUUGCGGUAGCCAGAGGGGAAACUCUGGGCCACAUCUGG
GUAUCUGUACCAGCACCAAAGCCCUACGUUUUACGCCAUGAUGAGAUCAAGAA
UCUGAUCGUGGACGAAGCAGCAAAGAAGGACGACGAAGUGGCUGUUACACUGGAG
CCAACCAUUCGUCACCCUGUUCGUGGUAAACUGAAACCGGACCUGGUGGUUCAA
ACAGAGAAGGCGUGUACGUUGUUGACGUGACAGUGAGACACGAGGAUGGCAACCU
GCUUGCACAGGGACGUCAGGAUAAACUGGACAAGUACGAAGUGCUGCUGCCGAUU
CUGCAAGAAAGACUGGGUGCUCCUACCGGUGAGGUUCUGCCGAUUGUUGUUGGCA
CCCGUGGCGCCAUGCCUAAAGAGACAGUGGAAGCCUUGAAGAAACUGCGCAUUAC
CGACCGGCAGACCCUGCUCACGAUCAGCCUGAUUGCCCGAGAAUGUCUGUGAAA

10

20

30

40

50

AUUUAUCAUACCUUCAUGGACUAUGCAAACGCCAGACCGCGUCCGGGCGGCGGUG
CAAACUACCCCCACAGAUGAUGACUCGAGCUGGUACUGCAUGCACGCAAUGCUAG
CUGCCCCUUUCCCGUCCUGGGUACCCCGAGUCUCCCCCGACCUCGGGUCCCAGGU
AUGCUCACCACCUCCACCUGCCCCACUCACCACCUCUGCUAGUUCAGACACCUCCC
AAGCACGCAGCAAUGCAGCUCAAAACGCUUAGCCUAGCCACACCCCCACGGGAAAC
AGCAGUGAUUAACCUUUAGCAAUAAACGAAAGUUUAACUAAGCUAUACUAACCCC
AGGGUUGGUCAAUUUCGUGCCAGCCACACCCUGGAGCUAGCAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAGCAUAUGACUAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

6. R2A NaviB N-DYKプラスミドORF配列:

ATGGATTACAAGGATGACGACGATAAGGGTACCATGGAGATCGTGGCCATCGCCAA
GAATCCACAGGCATGCCCAAAGTGTCTGCAGGGAGGAACCCAGCTGCTGTGCATGG
GCAGCTGGGAGCTGTCCAGACACATCAACAAGGAGCACCCAGCGTGGACGTGACA
TGGGTGTGCGGCGCCTGTTCAGCGGAGATGTACCACACTGCGCAGCTGGTCCTGCCA
CGTGCCTCACTGTAAGGGCCGGCAGGAGCCTAAGGATCTGCCATTCAAGTGCAGAGC
ACTGTTCTCTGAGCTTTGACAGCCAGATCGGCCTGTCCCAGCACGAGAGGCACGTG
CACCCAGAGGTGCGGAATGATAAGAGAGCCGCCGAGGCCAACAAGCCAAAGGGCAA
GAGCGGCAGGAGGCCCTTCCATCTGGTCTGATGAGGACCTGCTGCTGATCAGGGAGC
TGGAGTCCGAGTACCACGGAGCCCGGAATATCAACGAGAAGATCGCAGAGCACTTC
CCTGACAGGACCGGAAGGCAGGTGTCCGATGCACGGAGAAGGAAGGACTATGCCGC
CCTGAGGGGCGCGGAGGACCACAGGGACCTGCAGAGGGAGTGGAGGCAATCGAGG
AGGTGGATGAGGGCGAGATCCCAGAGGGCGAGGAGCTGGTGGCCACAGACGGAGCA
GCACTGGAGTCTGGACCACCTGAGAATGGAGGCAGCGCACCTGCAGAGCAGGTGAA
CGCACCAGCACTGGAGAGCTCCTCTCAGCAGGATAGGGAGTGCTCTCCAGCAGTGG
GCAGCGATGAGCAGATCGAGGACAGCTCCGACGATGACGAGTTTTCTGACGCACTG
GGAGAGATCAGCCTGCCAGAGCCTCTGTCCGTGGAGAGGACCACAATCAGCCCACC
ACCTAGGGATGACTGGAAGGGACCTATGCGGTGGGAGATCTGTAATGCCTCCGAGG
AGGCCGGCTCTTACGCCAACTGGGTGACCGGACTGCAGGAGCTGGTGGCGCAACAAT
GCCCTGTCTGAGATCGGCCTGGATAGCCTGTATGACCAGCTGATCCAGATCATGCG
GCACCCAGCGATGACAATGAGCAGGACAGGCTGCAGCTGAACGCAAGAGGACCAC
CAAGGAGGGGACACCGCAAGAATAGAAGGAGGAGGAGACTGACAGCAGCAGATAGG
AAGAGATTGCGCTTTGCCCGGTGCCAGGACCTGTGGAACAATAACCCTAAGAAGCT
GGCCGAGCTGGTCATCGCCAACGATCTGTCCATCCTGCAGAGGAGGCAGGCACCAG
GAAGAACCGAGACACAGACCCTGTACAATGAGCTGTGGGGAAGGGTGGGACCAAAC
ATCGAGGCACCAAGGAGAACCGAGGACCAATCCCCGTGTCCCGCATCTTCACACC
AATCACCCCCCAGGAGATCATGGGCCGGATCAGGCGCATCAAGAACGATTCTGCAG
CAGGACCTGACGGAGTGACAAAGGATGACCTGAGGGGCGCGGCGTGTCTATCGCA
CTGAGCAAGCTGTTTAATAGCATCCTGCTGGCCGGCTACTATCCAAAGGCCTGGCG
GGAGAACAGAACCACACTGCTGCCTAAGCCAGAGAAGGACCCCGCCGACGTGAAGA
ATTGGAGACCCATCACCATCTCTAGCATGGTGTCCAGGGTGTATTCTGGCCTGCTG
GATCAGCGGGTGAGAGCCGTGATCAAGCAGTGCGACAGGCAGAAAGGGCTTCACAGA
GGAGAATGGCTGTTTTTCCAACATCCAGCTGCTGGATGACGCCGTGTCTAACGCCA
AGAAGGCCGCGGCGGTGATCACCATCCTGGATGTGAGCAAGGCCTTCGACACAGTG
CCACACGCCGTGATCCAGGGCTGCCTGGAGAAGAAGGGCATCCCCGAGACCGTGGC
CGCCTACATCTCCTCTATGTATCGCGATTGTAGCACAGCAATCAGGACCAGGTCCG
GCGACGTGAAGATCGGCATGAAGAGGGGAGTGAAGCAGGGCGACCCTCTGTCCCCA
CTGATCTTTAATCTGGTGTCTGGAGCCACTGCTGGAGAGGCTGCAGGAGACCAGCGG
CGTGGAGATCGAGGGCATGAATCTGTCCTGCGCCGCCTTCGCCGATGACATCGTGT
GTTTTGCAAACACAGCACCTGAGGCAGGAAGGCAGCTGCGCATGGTGGCCGATTAC
CTGGGCAGACTGGACATGTCCCTGTCTGTGAGCAAGTGCATCGCCGTGGAGTACGT

10

20

30

40

50

GCCTCACAGGAAGACATGGTATACCAAGAACCCAGGCCTGGAAGTGAATGGCAACG
CCGTGCCCTCCATCTCTCCTAGCGAGACCTTCAAGTACCTGGGCGCCAAGGTGAGC
CCATGGAAGGGCCTGCTGGAGGGCTTCGAGTCCGATGCCTTTTCGCGAAGTGATCTC
TCGGGTGCAGAGACTGCCCCCTGAAGCCTATGCAGAAGGTGGACCTGCTGCAGATGT
ATATCTTCCCCAGATACACCTACGGCCTGATCACCAGCCCTCCAGCCAAGGCCGTG
CTGAAGACAATCGATAGGATCATCCGGACCAGAATCAAGGAGATCCTGCACCTGCC
AGAGTCCGTGAGCTCCTCTTTTCTGTACACCCCAAGGAAGCAGGGAGGACTGGGAC
TGCTGGAGGTGGAGAAGATGGTGCTGATCGCAGCACTGAGAAATGGACTGAGGGCA
AGGCAGTCTCACGACCCAGTGACCAGGGCCGCCATGAATTCTAACGCCGCCGATGA
CAGACTGAAGAGCTACGCCGATGCACTGAGGCTGCACTGGCCACTGACCACAAAGG
AGCTGGACACATAACAAGTATCAGCTGAGACTGTCCTATGCCCAGAAGTGGGCAGAG
CAGAAGTGGCAGGGACAGGGAGTGGAGGAGTTCGCACAGGACCCCGTGGGCAATTC
TTGGCTGCAGAGATACGATCTGCTGCCTGCCAGCAGGTATATCGACGCCATCAAGC
TGCGGACAAACACCTACCCACCCAGAGCCCTGATGAAGATCATCGATGGCAGGGTG
GACAGCTCCTGCCGCAAGTGTCAGGGCTCTAGCGAGACCCTGGGACACATCCTGGG
AAGGTGTAGATATACAAAGGATAAGAGGATCAGCCGCCACAATGAGATCAAGGACC
TGCTGAAGGCCCGCCTGGCCAAGAACCACCAGGTCATGGATGAGCCTCAGATCACC
GTGCGGGGCCAGAGATTCAAGCCAGACCTGGTGGTGAAGACAAATGAGGGCCGGGT
GCACGTGATCGATGTGACAGTGAGGTACGAGCACCGCACCTATCTGGACGAGGGCC
GCACAGAGAAGATCGGCAAGTACCGGCAGATCCTGAGCACCTGCGGAGAGATCTG
CACTCCAACGCCGAGGAAGTGATCCCCATCGTGATCGGCTCCAGAGGAGCAATCCC
TAGGGAGACCAGGAAGGCACTGTCTAAGCTGGGCATCGGCAAGTCCGACTGGCTGA
CAATCTCTCTGATCGCCCTGAGGTCCTCTCTGGAGATCGTGAACGCCTTTATGGAT
GACTAG

10

20

30

40

7. R2A NaviB N-DYKアミノ酸配列:

MDYKDDDDKGTMEIVAIKPNQACPKCLQGGTQLLCMGSWELSRHINKEHPSVDVT
WVCGACQRRCTTLRSWSCHVPHCKGRQEPKDLFPKCEHCSSLFDSQIGLSQHERHV
HPEVRNDKRAAEANKPKGKSGRRPSIWSDEDLLIRELESEYHGARNINEKIAEHFPD
RTGRQVSDARRRKDYAALRGRGGPQGPAGVEAIEEVDEGEIPEGEELVATDGALE
SGPPENGGSAPAEQVNAPALESSSQQDRECSPAVGSDEQIEDSSDDDEFSDALGEISL
PEPLSVERTTISPPPRDDWKGPMPRWEICNASEEAGSYANWVTGLQELVRNNALSEIG
LDSLYDQLIQIMRHPSDDNEQDRLQLNARGPPRRGHRKNRRRRRLTAADRKRFAFA
RCQDLWNNNPKKLAELVIANDLSILQRRQAPGRTETQTLYNELWGRVGPNIAPRR
TEDPIPVSRIFTPTPQEIIMGRIRRIKND SAAGPDGVTKDDLGRGRVSIALLSKLFNSIL
LAGYYPKAWRENRTTLLPKPEKDPADVKNWRPITISSMVSRVYSGLLDQVRRAVIK
QCDRQKGFTEENGCFNSIQLLDDAVSNAKKAGGVITILDVSKAFDTPHAVIQGCLE
KKGIPETVAAYISSMYRDCSTAIRTRSGDVKIGMKRGVKQGDPLSPLIFNLVLEPLLE
RLQETSGVEIEGMNLSCAAFADDIVCFANTAPEAGRQLRMVADYLGRLDMSLSVSKC
IAVEYVPHRKTWYTKNPGLEVNGNAVPSISPSETFKYLGAKVSPWKGLLEGFESDAF
REVISRVQRLPLKPMQKVDLLQMYIFPRYTYGLITSPPAKAVLKTIDRIIRTRIKEILH
LPESVSSSFLYTFRKQGGGLGLEVEKMLVIAALRNGLRARQSHDPVTRAAMNSNAAD
DRLKSYADALRLHWPLTTKELDTYKYQLRLSYAQKWAEQKWQGGQVEEFAQDPVG
NSWLQRYDLLPASRYIDAIKLRTNTYPTRALMKIIDGRVDSSCRKCKQSSETLGHILG
RCRYTKDKRISRHNKIDLLKARLAKNHQVMDEPQITVRGQRFPDLVVKTNENGRV
HVIDVTVRYEHRTYLDEGRTEKIGKYRQILSTLRRDLHSNAEEVIPIVIGSRGAIPRET
RKALSKLGIGKSDWLTISLIALRSSLEIVNAFMDD*

8. R2A OrLaSyn C-3FプラスミドORF配列:

ATGGGGACCGATACCGTTTATGTGGGTCAAGATTATCCAAGCGGCCTGTCCAAGCG
CGTTCCGGCTCGTCTGGTTGCTGGTCCAATGCTGCGCGAGCGCAGCTGCCATGCCC
ACGTGTTCCGTGCTGGTCACATGTGGAATTGGCGCACCAAGCCTGCCGAGCGGTCGT

50

TGGGACCAGCCGGCCCTGGAGAAGAGCCGCGTGCTGACCCGTAGCGTGGCTACCGC
CACCGATCCGGAGATCACCAGCTACCCGGGCAAGAGCGTGAGCACCAGCACCCAGG
TGCAGGAAGAGGACTGGTGCAGCCGTGAAAGCGGCTGGATCAGCCCGGGTCTGGCT
CCAGAGGAACCGAGCGTGGTGAGCGAGATTACCGCTAGCATGGTGGCTACCATGCG
TGTGGCTACCGAGGAAGTGGTGCTGGAACCGCAGCCGGAGCAGGTGGTGACCATTC
TGCCAGAACATGGTTCGTAACGTGCCGCCGGGTCTGGCTGAACAGGACACCGCGAGC
CCGATTGAGGTGAGCGTGCTGCTGCCGGATCTGGCGGAGAACTGCCCACTGTGCGG
TGTGCCGAGCGGCGGTCTGCGTCTGCTGGGCAAGCACTTCGCGGTGCGTCATGCTG
GTGTGCCAGTGACCTACGAGTGCCGCAAATGCGCGTGCGGTAGCCCGAACAGCCAC
AGCATCAGCTGCCATGTGCCGAAGTGCCGCGGTCTGTGCTCGTATGCCGAGCGGGCGA
CCCGGGTATTGCTTGCGATCTGTGCGAAGCGCGCTTTGCTACCGAGGTGGGTGTGG
CCCAGCACAAAGCGTCACGTGCACCCGGTGGAATGGAACAAAGTGCGCCTGGAGCGT
CGCGGTGCCCCGTGGCGGTGGCATCAAGGCGACCAAACCTGTGGAGCGTGGCTGAAGT
GGAAACCTGATCCGTCTGATTCGTGAGCACGGCGACAGCGGTGCGACCTATCAGC
TGATCGCTGATGAACTGGGTCTGTGGCAAGACCGCCGAGCAGGTGCGTAGCAAGAAA
CGCCTGCTGCGTATTGACACCGCCAGCAACAGCCCGGACGATGCGGAAGTGGAGGA
AGAGCGCCTGGAGAGCCTGGCCGTGCGTAGCAGCAGCCGTAGCCCGCCGAGCCTGG
TGGCCACCCGCGTGCGTGAAGCGGTGGCTCGTGGTGAAAGCGAGGGTGGCGAAGAG
ATCCGCGCCATTGCCGCGCTGATCCGTGACGTGGATCAGAACCCTGCGCTGATTGA
AACCAGCGCCAGCGACATCATTAGCAAGCTGGGCCGTGCGGTGGACGGCCCGAAAC
GCCCCGCGTCCGGTGGTGCGTGAACAGACCCAGGAGAAGGGTTGGGTGCGTCGCCTG
GCCCCGTGCAAACGCGAATACCGTGAGGCTCAGTACCTGTATAGCCGCGACCAGGC
TCGTCTGGCTGCGCAGATCCTGGACGGCGCCGCTAGCCAGGAGTGCGCCCTGCCAG
TGGACCAGGTGTATGGTGCTTTCCGTGAGAAGTGGGAACCGTGCGGCCAGTTCAC
GGTCTGGGCGAATTTGCGACCGGTGCGCGTGCTGATAACTGGGAGTTCTACAGCCC
GATCCTGGCCGCGGAAGTGAAGGAGAACCTGATGCGCATGGCTAATGGTACCGCCC
CAGGTCCAGACCGTATTAGCAAGAAAGCCCTGCTGGATTGGGACCCGCGTGGTGAA
CAGCTGGCCCCGTCTGTATACCACTGGCTGATCGGTGGCGTGATTCCGCGCGTGTT
CAAGGAATGCCGTACCAAGCTGCTGCCGAAGAGCAGCGACCCGGTGGAGCTGCAGG
ATATCGGTGGCTGGCGTCCGGTGACCATTTGGTAGCATGGTGACCCGCGCTGTTTAGC
CGTATCCTGACCATGCGTCTGACCCGCGCGTGCCCGATTAAACCCGCGCCAGCGTGG
CTTCCTGGCCAGCAGCAGCGGTTGCGCCGAGAACCTGCTGATCTTTGACGAGATTG
TGCGTCGCAGCCGTGCGGATGGTGGCCCACTGGCCGTGGTGTTCTGGACTTTGCC
CGTGCCCTTCGACAGCATCAGCCACGAACACATTCTGTGCGTGCTGGAAGAGGGTGG
CCTGGACCGTCACGTGATCGGTCTGATTCTGTAACAGCTACGTGGATTGCGTGACCC
GTGTGGGCTGCGTGGAAGGGTATGACCCCGCCGATCCAGATGAAAGTGGGCGTGAAA
CAGGGTGACCCGATGAGCCCGCTGCTGTTTAACTGGCGATGGACCCGCTGATCCA
CAAGCTGGAAACCGCTGGTACCGGCCTGAAATGGGGCGACCTGAGCATTGCCACCC
TGGCTTTGCGCGACGATCTGGTGCTGGTGAGCGATAGCGAAGAGGGTATGGGCCGT
AGCCTGGGCATTCTGGAGAAGTTCTGCCAGCTGACCGGTCTGCGTGTCAGCCGCG
TAAGTGCCACGGCTTCTTTATGGATAAAGGTGTGGTGAAACGGTTGCGGCACCTGGG
AAATCTGCGGCAGCCCGATCCACATGATTCCGCCGGGTGAGAGCGTGCGCTACCTG
GGTGTGCAAGTGGGTCCGGGTGCTGGTGATGGAGCCAGACCTGATTCCGACCGT
GCACACCTGGATCGAACGCATTAGCGAAGCCCCGCTGAAACCGAGCCAGCGCATGC
GTGTGCTGAACAGCTTCGCTCTGCCGCGTATCATCTACCAGGCCGACCTGGGCAAG
GTGACCGTGACCAAACCTGGCCCAGATCGATGGTATTGTGCGTAAGGCGGTGAAGAA
ATGGCTGCACCTGAGCCCGAGCACCTGCAACGGCCTGCTGTACAGCCGCAACCGTG
ATGGTGCCCTGGGTCTGCTGAAGCTGGAGCGTCTGATCCCGAGCGTGCGCACCAAA
CGTATCTACCGCATGAGCCGTAGCCCGGACATCTGGACCCGTGCGATGACCAGCCA
CAGCGTGAGCAAGAGCGATTGGGAATGCTGTGGGTGCAGGCGGGTGGCGAGCGTG

10

20

30

40

50

GCAGCGCCCCGGTTATGGGTGCGGTGGAAGCGGCTCCAACCGACGTGGAGCGTAGC
CCGGACTACCCAGATTGGCGTCGCGAAGAGAACCTGGCCTGGAGCGCGCTGCGTGT
GCAGGGCGTGGGTGCCGACCAGTTCCGCGGCGACCGTACCAGCAGCAGCTGGATCG
CTGAACCAGCCAGCGTGGGCTTCGCCCAGCGTCACTGGCTGGCGGCTCTGGCGCTG
CGTGCTGGCGTGTATCCAACCCGCGAATTCCTGGCCCGTGGCAAGGAGAAGAGCGG
TGCTGCCTGCCGTCGCTGCCCAGCTCGTCTGGAGAGCTGCAGCCACATTCTGGGCC
AGTGCCCGTTTTGTGCAGGCTAACCGCATTGCCCGTCACAACAAAGTGTGCGTGCTG
CTGGCTACCGAAGCCGAGCGTTTTCGGCTGGACCGTGATCCGCGAATTTCTGTCTGGA
GGACGCTGCCGGCGGCCTGAAGATCCCGGATCTGGTGTGCAAGAAAGCCGACACCG
TGCTGATTGTGGACGTGACCGTGCGCTACGAAATGGACGGTGAAACCCCTGAAGCGT
GCGGCTAGCGAGAAGGTGAAGCACTATCTGCCGGTGGGTGAGCAGATCACCGATAA
AGTGGGTGGCCGCTGCTTCAAAGTGATGGGCTTTCCGGTGGGTGCTCGTGGCAAGT
GGCCGGCCAGCAACAATACCGTGCTGGCCGAACTGGGCGTGCCGGCTGGTCGCATG
CGTACCTTCGCGCGTCTGGTGAGCCGTCGCACCCCTGCTGTATAGCCTGGACATTCT
GCGCGACTTCATGCGTGAACCAGCAGGTCTGTGGCACCCGTGTGGCACTGATTCCAG
CAGCAACCGGCGCAGCAAATATGGGGAGCGACTACAAAGACCATGACGGTGATTAT
AAAGATCATGACATCGATTACAAGGATGACGATGACAAGAAGTGATAA

9. R2A OrLaSyn C-3Fアミノ酸配列: MGTDTVYVGQDYPSGLSKRVPARLVAGPML
RERSCHAHVFRAGHMWNWRTSLPSGRWDQPALEKSRVLT RSVATATDPEITSYPGK
SVSTSTQVQEEDWCSRESGWISPLAPEEPSVSEITASMVATMRVATEEVVLEPQP
EQVVTILPEHGRNVPPGLAEQDTASPIEVSVLLPDLAENCPLCGVPSGGLRLLGKHF
AVRHAGVPVTYECRKCAWRSPNSHSISCHVPKCRGRARMPSGDPGIACDLCEARFA
TEVGVAQHKRHVHPVEWNVRLERRGARGGGIKATKLWSVAEVELIRLIREHGDS
GATYQLIADELGRGKTAEQVRSKRLLRIDTASNPDDAEVEEERLES LAVRSSSRSP
PSLVATR VREAVARGESEGGEEIRAI AALIRDVDQNPCLIETSASDIISKLGRRVDGPK
RPRPVVREQTQEKGWVRR LARRKREYREAQYLYSRDQARLAAQILDGAASQECALP
VDQVYGAFREKWETVGQFHGLGEFRTGARADNWEFYSPILAAEVKENLMRMANGT
APGPDRISK KALLDWDPRGEQLARLYTTWLIGGVIPRVFKECRTKLLPKSSDPVELQ
DIGGWRPV TIGSMVTRLFSRILTMRLTRACPINPRQRGFLASSSGCAENLLIFDEIVR
RSRRDGGPLAVVFDFAFAFDSISHEHILCVLEEGGLDRHVIGLIRNSYVDCVTRVGC
VEG MTPPIQMKVGVKQGDPMSP L LFN LAMDPLIHKLETAGTGLKWGDLSIATLAF A
DDLVLVSDSEEGMGRSLGILEKFCQLTGLRVQPRKCHGFFMDKGVVNGCGTWEICG
SPIHMIPPGESVRYLGVQVGPGRGVM EPDLIPTVHTWIERISEAPLKPSQRMRLNS
FALPRIIYQADLGKVTVT KLAQIDGIVRKAVKKWLHLSPSTCNGLLYSRNRDGGGLL
LKLERLIPSVRTKRIYRMSRSPDIWTRRMTSHSVSKSDWEMLWVQAGGERGSAPVM
GAVEAAPT DVERS PDY PDWRREENLAWSALRVQGVGADQFRGDR TSSSWIAEPASV
GFAQRH WLAALALRAGVYPTREFLARGKEKSGAACRRCPARLESCSHILGQCPFVQA
NRIARHNKVCVLLATEAERFGWTVIREFRLEDAAGGLKIPDLVCKKADTVLIVDVTV
RYEMDGETLKRAASEKV KHYLPVGQQITDKVGGRCFKVMGFPVGARGKWPASNNT
VLAELGV PAGRMR T FARLVSRRTLLYSLDILRDFMREPAGRGTRVALIPAATGAANM
GSDYKDHDGDYKDHDIDYKDDDDKK**

10. R2A OrLaSyn C-3FトランスフェクトしたmRNA配列:

AGGAAUAAACUAGUAUUCUUCUGGUCCCCACAGACUCAGAGAGAACCCGCCACCA
UGGGCACCGAUACCGUUUAUGUGGGUCAAGAUUAUCCAAGCGGCCUGUCCAAGCG
CGUUCGGCUCGUCUGGUUGCUGGUCCA AUGCUGCGGAGCGCAGCUGCCAUGCC
CACGUGUUCGUGCUGGUCACAUGUGGAAUUGGCGCACCAAGCCUGCCGAGCGGUC
GUUGGGACCAAGCCGGCCCCUGGAGAAGAGCCGCGUGCUGACCCGUAGCGUGGCUAC
CGCCACCGAUCCGGAGAUCAACCAGCUACCCGGGCAAGAGCGUGAGCACCAAGCACCC
AGGUGCAGGAAGAGGACUGGUGCAGCCGUGAAAGCGGCUGGAUCAGCCCGGGUCU
GGCUCCAGAGGAACCGAGCGUGGUGAGCGAGAUUACCGCUAGCAUGGUGGCUACC

AUGCGUGUGGCUACCGAGGAAGUGGUGCUGGAACCGCAGCCGGAGCAGGUGGUGA
CCAUUCUGCCAGAACAUUGGUCGUAACGUGCCGCCGGGUCUGGCUGAACAGGACAC
CGCGAGCCCGAUUGAGGUGAGCGUGCUGCUGCCGGAUCUGGCGGAGAACUGCCCA
CUGUGCGGUGUGCCGAGCGGGCGGUCUGCGUCUGCUGGGCAAGCACUUCGCGGUGC
GUCAUGCUGGUGUGGCCAGUGACCUACGAGUGCCGCAAAUGCGCGUGGCGUAGCCC
GAACAGCCACAGCAUCAGCUGCCAUGUGCCGAAGUGCCGCGGUCGUGCUCGUAUG
CCGAGCGGCGACCCGGGUAAUUGCUUGCGAUCUGUGCGAAGCGCGCUUUGCUACCG
AGGUGGGUGUGGCCCAGCACAAAGCGUCACGUGCACCCGGUGGAAUGGAACAAAGU
GCGCCUGGAGCGUCGCGGUGCCCGUGGCGGUGGCAUCAAGGCGACCAAACUGUGG
AGCGUGGCUGAAGUGGAAACCCUGAUCCGUCUGAUUCGUGAGCACGGCGACAGCG
GUGCGACCUAUCAGCUGAUCGUGAUGAACUGGGUCGUGGCAAGACCGCCGAGCA
GGUGCGUAGCAAGAAACGCCUGCUGCGUAUUGACACCGCCAGCAACAGCCCGGAC
GAUGCGGAAGUGGAGGAAGAGCGCCUGGAGAGCCUGGCCGUGCGUAGCAGCAGCC
GUAGCCCGCCGAGCCUGGUGGCCACCCGCGUGCGUGAAGCGGUGGCUCGUGGUGA
AAGCGAGGGUGGCGAAGAGAUCGCGCCAUUGCCGCGCUGAUCCGUGACGUGGAU
CAGAACCCGUGCCUGAUUGAAACAGCGCCAGCGACAUCAUUAGCAAGCUGGGCC
GUCGCGUGGACGGCCCGAAACGCCCGGUGCCGGUGGUGCGUGAACAGACCCAGGA
GAAGGGUUGGGUGCGUCGCCUGGCCCGUCGCAAACGCGAAUACCGUGAGGCUCAG
UACCGUGAUUAGCCGCGACCAGGCUCGUCUGGCUGCGCAGAUCUGGACGGCGCCG
CUAGCCAGGAGUGCGCCUGCCAGUGGACCAGGUGUAUGGUGCUUUCGUGAGAA
GUGGGAAACCGUGGGGCCAGUUCCACGGUCUGGGCGAAUUCGCACCGGUGCGCGU
GCUGAUAAACUGGGAGUUCUACAGCCCGAUCCUGGCCGCGGAAGUGAAGGAGAACC
UGAUGCGCAUGGCUAAUGGUACCGCCCCAGGUCCAGACCGUAUUAAGCAAGAAAGC
CCUGCUGGAUUGGGACCCGCGUGGUGAACAGCUGGCCCGUCUGUAUACCACCUUG
CUGAUCGGUGGCGUGAUUCCGCGCGUGUUAAGGAAUGCCGUACCAAGCUGCUGC
CGAAGAGCAGCGACCCGGUGGAGCUGCAGGAUAUCGGUGGCUGGCGUCCGGUGAC
CAUUGGUAGCAUGGUGACCCGCCUGUUUAGCCGUAUCCUGACCAUGCGUCUGACC
CGCGCGUGCCCGAUUAACCCGCGCCAGCGUGGCUUCCUGGCCAGCAGCAGCGGUU
GCGCCGAGAACCUGCUGAUUCUUGACGAGAUUGUGCGUCGCGAGCCGUCGCGAUGG
UGGCCACUGGCCGUGGUGUUCGUGGACUUGCCCGUGCCUUCGACAGCAUCAGC
CACGAACACAUUCUGUGCGUGCUGGAAGAGGGUGGCCUGGACCGUCACGUGAUCG
GUCUGAUUCGUAACAGCUACGUGGAUUGCGUGACCCGUGUGGGCUGCGUGGAGGG
UAUGACCCCGCCGAUCCAGAUGAAAGUGGGCGUGAAACAGGGUGACCCGAUGAGC
CCGCGUCUGUUUAACCUGGCGAUGGACCCGCGUGAUCCACAAGCUGGAAACCGCUG
GUACCGGCCUGAAAUGGGGCGACCUGAGCAUUGCCACCCUGGCUUUCGCCGACGA
UCUGGUGCUGGUGAGCGAUAGCGAAGAGGGUAUGGGCCGUAGCCUGGGCAUUCUG
GAGAAGUUCUGCCAGCUGACCGGUCUGCGUGUGCAGCCGCGUAAGUGCCACGGCU
UCUUUAUGGAUAAAGGUGUGGUGAACGGUUGCGGCACCUGGGAAAUCUGCGGCAG
CCCGAUCCACAUGAUUCCGCCGGGUGAGAGCGUGCGCUACCUGGGUGUGCAAGUG
GGUCCGGGUCGUGGUGUGAUGGAGCCAGACCUGAUUCCGACCGUGCACACCUGGA
UCGAACGCAUUAAGCGAAGCCCCGCGUGAAACCGAGCCAGCGCAUGCGUGUGCUGAA
CAGCUUCGCUUCUGCCGCGUAUCAUACCAAGGCCGACCUGGGCAAGGUGACCGUG
ACCAAACUGGCCCAGAUUGGAUUGUGCGUAAGGCGGUGAAGAAAUGGCUGC
ACCUGAGCCCGAGCACCUUGCAACGGCCUGCUGUACAGCCGCAACCGUGAUGGUGG
CCUGGGUCUGCUGAAGCUGGAGCGUCUGAUCCCGAGCGUGCGCACCAAACGUAUC
UACCGCAUGAGCCGUAGCCCGGACAUCUGGACCCGUCGCAUGACCAGCCACAGCG
UGAGCAAGAGCGAUUGGGAAAUGCUGUGGGUGCAGGCGGGUGGCGAGCGUGGCAG
CGCCCCGGUUAUGGGUGCGGUGGAAGCGGCUCCAACCGACGUGGAGCGUAGCCCG
GACUACCCAGAUUGGCGUCGCGAAGAGAACCUGGCCUGGAGCGCGCUGCGUGUGC
AGGGCGUGGGUGCCGACCAGUUCGCGGCGACCGUACCAGCAGCAGCUGGAUCGC

10

20

30

40

50

UGAACCAGCCAGCGUGGGCUUCGCCAGCGUCACUGGCUGGCGGCUCUGGCGCUG
CGUGCUGGCGUGUAUCCAACCCGCGAAUUCUGGCCCGUGGCAAGGAGAAGAGCG
GUGCUGCCUGCCGUCGCGUGCCCAGCUCGUCUGGAGAGCUGCAGCCACAUCUGGG
CCAGUGCCCGUUUGUGCAGGCUAACCGCAUUGCCCGUCACAACAAAGUGUGCGUG
CUGCUGGCUACCGAAGCCGAGCGUUUCGGCUGGACCGUGAUCCGCGAAUUUCGUC
UGGAGGACGCGUGCCGGCGGCCUGAAGAUCCTCGAUCUGGUGUGCAAGAAAGCCGA
CACCGUGCUGAUUGUGGACGUGACCGUGCGCUACGAAAUGGACGGUGAAACCCUG
AAGCGUGCGGCUAGCGAGAAGGUGAAGCACUAUCUGCCGGUGGGUCAGCAGAUCA
CCGAUAAAGUGGGUGGCCGCGUGCUUCAAGUGAUGGGCUUUCGGUGGGUGCUCG
UGGCAAGUGGCCGGCCAGCAACAAUACCGUGCUGGCCGAACUGGGCGUGCCGGCU
GGUCGCAUGCGUACCUUCGCGCGUCUGGUGAGCCGUCGCACCCUGCUGUAUAGCC
UGGACAUUCUGCGCGACUUAUGCGUGAACCAGCAGGUCGUGGCACCCGUGUGGC
ACUGAUUCCAGCAGCAACCGGCGCAGCAAAUAUGGGGAGCGACUACAAAGACCAU
GACGGUGAUUAUAAAGAUCAUGACAUCGAUUACAAGGAUGACGAUGACAAGAAGU
GAUGACUCGAGCUGGUACUGCAUGCACGCAAUGCUAGCUGCCCCUUCGCGUCCU
GGGUACCCCGAGUCUCCCCCGACCUCGGGUCCCAGGUUAGCUCCCACCUCCACCU
GCCCCACUCACCACCUCUGCUAGUUCAGACACCUCCCAAGCACGCAGCAAUGCAG
CUCAAAACGCUUAGCCUAGCCACACCCCCACGGGAAACAGCAGUGAUUAACCUUU
AGCAAUAAACGAAAGUUUAACUAAGCUAUACUAACCCCAAGGGUUGGUCAAUUUCG
UGCCAGCCACACCCUGGAGCUAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAG
CAUAUGACUAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAA

10

20

30

40

50

11. R2A OrLa天然C-3FプラスミドORF配列:

ATGGGAACAGATACAGTGTATGTGCGGCCAGGACTACCTTCTGGCTTATCAAAACG
GGTACCAGCACGGTTAGTGGCGGGACCGATGCTGCGAGAGCGAAGCTGTCACGCCC
ATGTGTTTAGGGCTGGACACATGTGGAACCTGGCGAACCAGCCTTCCGAGCGGGCGC
TGGGACCAGCCCGCTTTGGAGAAGTCTCGGGTCCTAACCCGGTCGGTGGCGACGGC
CACCGACCCCGAAATTACCTCTTACCCAGGAAAGTCCGTATCGACAAGTACGCAGG
TTCAGGAGGAGGACTGGTGTAGCCGGGAGAGCGGGTGGATCTCGCCAGGACTTGCT
CCTGAAGAACCCTCGGTGGTGTCCGAAATTACAGCCTCCATGGTAGCGACAATGAG
GGTAGCAACCGAGGAGGTGCTGCTGGAACCACAGCCTGAACAGGTGCTCACAATAC
TGCCGGAGCATGGTCGAAACGTTCTCCGGGGCTGGCAGAACAGGACACCGCCAGC
CCCATAGAAGTCTCGGTGCTCCTCCAGACCTCGCTGAGAACTGCCCATTTGTGTGG
CGTGCCGAGCGGGGGCCTACGCTTGCTCGGGAAGCATTTTGTGTGCTCGACATGCGG
GGGTGCCTGTAACGTATGAGTGCCGTAAGTGTGCGTGGCGGAGCCCCAACAGCCAC
TCAATCTCGTGTACGTCCCCAAATGCCGGGGGCGTGCGCGGATGCCCAGTGGCGA
TCCAGGGATCGCCTGCGATCTCTGTGAAGCCCGGTTTGCCACGGAGGTTGGGGTCG
CCCAACACAAGCGGCACGTTTCATCCGGTGGAGTGGAAACAAGGTGAGGCTGGAAAGG
AGAGGTGCGCGCGGAGGGGGAATTAAGGCGACGAAGCTCTGGAGTGTAGCGGAGGT
AGAGACGCTAATCCGGCTCATCCGTGAGCACGGAGATTAGGTGCCACTTACCAGC
TCATTGCCGATGAGCTGGGAAGGGGGCAAGACGGCCGAACAGGTGAGGAGTAAAAAG
AGGCTCCTGCGCATAGATACGGCAAGCAATAGCCCAGATGATGCAGAGGTTGAGGA
GGAGAGGTTGGAATCTCTGGCGGTTTCGGTCCTCGTCACGGTCACCCCCGAGCCTGG
TGGCGACCAGGGTCAGGGAGGCAGTTGCCAGGGGTGAATCAGAAGGTGGCGAGGAG
ATCAGGGCTATTGCTGCTCTCATTAGGGACGTAGATCAGAATCCTTGTCTGATTGA
AACCTCGGCGTCGGACATCATCTCGAAGCTGGGAAGGAGGGTGGATGGGCCCAAGA
GACCCAGGCCCCGTTGTGAGAGAACAGACCCAAGAGAAGGGATGGGTAAGGCGGCTT
GCCCCGGCGGAAAAGGGAGTACAGAGAAGCGCAGTACCTGTACTCAAGGGATCAAGC
AAGGCTGGCGGCCCAGATCCTCGATGGTGGCGCCAGCCAGGAATGCGCCCTCCCGG
TGGACCAGGTCTACGGAGCGTTCCGTGAGAAATGGGAACCGTAGGGCAGTTCAC

GGACTTGGTGAGTTCCGGACGGGTGCACGCGCAGACAACTGGGAGTTCTACTCTCC
AATTCTGGCGGCTGAGGTGAAAGAAAACCTAATGAGAATGGCTAACGGCACGGCCC
CGGGACCAGACAGGATAAGCAAAAAGGCTCTGCTTGAAGTGGGACCCCCGGGGTGAG
CAACTGGCACGGCTGTACACGACGTGGCTGATCGGTGGGGTCATACCAAGGGTCTT
CAAGGAGTGCAGGACTAAGCTGCTACCGAAATCCAGCGACCCGGTTCGAGTTGCAGG
ACATCGGTGGATGGAGGCCGGTGACGATTGGGTGCGATGGTGAAGTGGCTGTTTCACT
CGGATTCTAACGATGAGGCTAACCCGAGCCTGTCCGATCAATCCGAGGCAGCGCGG
TTTCTTGGCCTCCTCGAGTGGATGCGCGGAAAACCTGTTGATCTTTGACGAGATCG
TCAGGCGCTCGAGGCGGGACGGGGGGCCGCTGGCAGTGGTGTGTTGTGGACTTTGCG
AGGGCCTTTGACTCCATCTCACATGAACATATCCTGTGTGTTCTCGAAGAAGGCGG 10
GCTTGACAGGCACGTTATCGGGTTGATCCGAAACTCGTACGTGGATTGCGTGACCA
GGGTGGGTTGTGTGCGAGGGCATGACACCACCAATAACAAATGAAGGTTGGAGTGAAG
CAGGGAGACCCCATGTCCCCCTTGCTCTTCAACCTGGCTATGGATCCCCCTCATCCAT
AAACTCGAGACGGCCGGAACCTGGACTGAAATGGGGCGATCTTTCAATCGCCACGCT
GGCCTTTGCCGACGATCTGGTGCTGGTGAGTGAAGTCTGAGGAAGGCATGGGGAGGA
GTCTCGGGATTTTGGAGAAGTTTTGCCAACTGACTGGGCTGAGGGTTTCAGCCCAGG
AAGTGTCACGTTTTCTTTATGGACAAGGGCGTGGTGAACGGCTGTGGAACCTGGGA
AATCTGTGGGTACCGATCCACATGATTCCCCCGGGGGAATCAGTTCGTTATTTGG
GAGTCCAGGTAGGCCCGGGGCGCGGCGTGATGGAACCGGATCTTATCCCTACGGTC
CACACGTGGATCGAAAGGATCTCGGAGGCTCCTCTAAAGCCCTCACAAACGCATGAG 20
GGTTTTGAACTCATTGCTCTCCCCCGGATAATTTACCAGGCCGATCTAGGGAAGG
TTACGGTAACCAAATTGGCCCAGATAGATGGGATTGTCCGGAAGGCTGTGAAGAAG
TGGCTCCATTTGTACCATCCACGTGCAATGGACTGCTGTATTACGGAACCGCGA
CGGTGGTTTTGGGCCTCCTAAAGCTGGAAAGACTAATCCCATCCGTGCGCACGAAGC
GTATCTATCGGATGTCCAGGTCTCCGGATATCTGGACACGGCGAATGACCAGCCAT
TCTGTGTCAAATCTGACTGGGAGATGTTGTGGGTCCAAGCGGGAGGTGAGAGGGG
CAGTGACCTGTAATGGGTGCCGTGGAGGCTGCCCCGACCGATGTGGAGAGATCGC
CAGACTACCCAGACTGGCGGCGTGAGGAAAACCTGGCATGGTTCGGCCCTGCGGGTG
CAGGGTGTGGGTGCAGACCAGTTTCGAGGCGACAGGACCAGCAGCTCTTGGATCGC
CGAGCCCGCTTCGGTTGGGTTCGCGCAGCGCCACTGGTTGGCTGCCCTGGCGCTGA 30
GGGCTGGGGTGTATCCGACTCGGGAGTTTCTGGCTCGGGGTAAAGGAAAAGTCAGGA
GCAGCTTGACAGCGCTGCCCGGCCAGGTTGGAATCATGTTACACATACTTGGGCA
ATGTCCGTTTCGTTTCAGGCGAACAGAATTGCGAGGCACAACAAGGTGTGTGTGCTCT
TGGCCACGGAGGCGGAGAGGTTTCGGCTGGACGGTAATAAGGGAGTTCCGTCTTGAG
GACGCCGCTGGCGGTCTCAAGATACCCGACCTGGTTTGCAAGAAGGCCGACACAGT
TCTCATTGTGACGCTGACCGTCCGGTACGAGATGGATGGAGAGACGCTAAAAAGGG
CCGCATCGGAGAAGGTGAAACACTATCTCCAGTAGGGCAACAGATAACGGACAAG
GTCGGAGGGCGTTGCTTTAAAGTCATGGGGTTCCCTGTAGGTGCTAGGGGAAAGTG
GCCGGCGAGCAACAACACAGTTTTTGGCTGAGTTAGGCGTCCCTGCAGGTTCGGATGA
GGACCTTTGCCAGGCTGGTGAGCCGGAGGACTCTTCTTTATTCTTTGGATATATTG 40
AGGGACTTCATGCGTGAGCCGGCCGGCAGGGGAACTCGGGTTGCTCTCATCCCTGC
GGCAACGGGTGCCGCGAATATGGGGAGCGACTACAAAGACCATGACGGTGATTATA
AAGATCATGACATCGATTACAAGGATGACGATGACAAGAAGTGA

12. R2A OrLa天然C-3Fアミノ酸配列:

MGTDTVYVGQDYPSGLSKRVPARLVAGPMLRERSCHAHVFRAGHMWNWRTSLPSG
RWDQPALEKSRVLTRSVATATDPEITSYPGKSVSTSTQVQEEDWCSRESGWISPLA
PEEPSVVSEITASMVATMRVATEEVVLEPQPEQVVITLPEHGRNVPPGLAEQDTASP
IEVSVLLPDLAENCPLCGVPSGGLRLLGKHFAVRHAGVPVTYECRKCAWRSPNSHSI
SCHVPKCRGRARMPSGDPGIACDLCEARFATEVGVAQHKRHVHPVEWNVKVRLEERRG
ARGGGIKATKLWSVAEVETLIRLIREHGD SGATYQLIADELGRGKTAEQVRSKKRLLR 50

IDTASNSPDDAEVEEERLES LAVRSSSRSPPSLVATRVREAVARGESEGGEEIRAI AAL
IRDVDQNPCLIETSASDIISKLGRRVDGPKRPRPVVREQTQEKGWVRRRLARRKREYR
EAQYLYSRDQARLAAQILDGAASQECALPVDQVYGAFREKWETVGQFHGLGEFRTG
ARADNWEFYSPILAAEVKENLMRMANGTAPGPDRIKKALLDWDPRGEQLARLYTT
WLI GGVIPRVFKECRTKLLPKSSDPVELQDIGGWRPV TIGSMVTRLFSRILTMRLTRA
CPINPRQRGFLASSSGCAENLLIFDEIVRRSRRDGGGPLAVVFVDFARAFDSISHEHILC
VLEEGGLDRHVIGLIRNSYVDCVTRVGCVEGMTPIQM KVGVKQGDPMSPLLFNLA
MDPLIHKLETAGTGLKWGDLSIATLAFADDLVLVSDSEEGMGRSLGILEKFCQLTGL
RVQPRKCHGFFMDKGVVNGCGTWEICGSP IHMIPPGESVRYLG VQVGPGRGVM EPD
LIPTVHTWIERISEAPLKPSQRM RVLNSFALPRIIYQADLGKVTVT KLAQIDGIVRKA
VKKWLHLSPSTCNGLLYSRNRDGGGLGLLKLERLIPSVRTKRIYRMSRSPDIWTRRMT
SHSVSKSDWEMLWVQAGGERGSAPVMGAVEAAPT DVERS PDYPDW RREENLAWSA
LRVQGVGADQFRGDR TSSSWIAEPASVGFAQRH WLAALALRAGVYPTREFLARGKE
KSGAACRRCPARLESCSHILGQCPFVQANRIARHNKVCVLLATEAERFGWTVIREFR
LEDAAGGLKIPDLVCKKADTVLIVDVT VRYEMDGETLKRAASEKV KHYLPVGQQITD
KVGGRCFKVMGFPVGARGKWPASNNTVLAELGVPAGRMRTFARLVSRRTLLYSLDI
LRDFMREPAGRGTRVALIPAATGAANMGSDYKDHDGDYKDHDIDYKDDDDKK*

10

13. R2A OrLa合成N-DYKプラスミドORF配列:

ATGGATTACAAGGATGACGACGATAAGGGTACCGATACCGTTTATGTGGGTCAAGA
TTATCCAAGCGGCCTGTCCAAGCGCGTTCCGGCTCGTCTGGTTGCTGGTCCAATGC
TGCGCGAGCGCAGCTGCCATGCCACGTGTTCCGTGCTGGTCACATGTGGAATTGG
CGCACCAGCCTGCCGAGCGGTCTGTTGGGACCAGCCGGCCCTGGAGAAGAGCCGCGT
GCTGACCCGTAGCGTGGCTACCGCCACCGATCCGGAGATCACCAGCTACCCGGGCA
AGAGCGTGAGCACCAGCACCCAGGTGCAGGAAGAGGACTGGTGCAGCCGTGAAAGC
GGCTGGATCAGCCCGGGTCTGGCTCCAGAGGAACCGAGCGTGGTGAGCGAGATTAC
CGCTAGCATGGTGGCTACCATGCGTGTGGCTACCGAGGAAGTGGTGCTGGAACCGC
AGCCGGAGCAGGTGGTGACCATTCTGCCAGAACATGGTCGTAACTGCCGCCGGGT
CTGGCTGAACAGGACACCGCGAGCCCGATTGAGGTGAGCGTGCTGCTGCCGGATCT
GGCGGAGAACTGCCCACTGTGCGGTGTGCCGAGCGGCGGTCTGCGTCTGCTGGGCA
AGCACTTCGCGGTGCGTCATGCTGGTGTGCCAGTGACCTACGAGTGCCGCAAATGC
GCGTGCGTAGCCCGAACAGCCACAGCATCAGCTGCCATGTGCCGAAGTGCCGCGG
TCGTGCTCGTATGCCGAGCGGCGACCCGGGTATTGCTTGCGATCTGTGCGAAGCGC
GCTTTGCTACCGAGGTGGGTGTGGCCAGCACAAAGCGTCACGTGCACCCGGTGGA
TGGAACAAAGTGCGCCTGGAGCGTCGCGGTGCCCGTGCGGTGGCATCAAGGCGAC
CAAAGTGTGGAGCGTGGCTGAAGTGGAACCCCTGATCCGTCTGATTCTGTGAGCACG
GCGACAGCGGTGCGACCTATCAGCTGATCGCTGATGAACTGGGTCTGTGGCAAGACC
GCCGAGCAGGTGCGTAGCAAGAAACGCCTGCTGCGTATTGACACCGCCAGCAACAG
CCCGGACGATGCGGAAGTGGAAGGAAGAGCGCCTGGAGAGCCTGGCCGTGCGTAGCA
GCAGCCGTAGCCCGCCGAGCCTGGTGGCCACCCGCGTGCGTGAAAGCGGTGGCTCGT
GGTGAAAGCGAGGGTGGCGAAGAGATCCGCGCCATTGCCGCGCTGATCCGTGACGT
GGATCAGAAACCCGTGCCTGATTGAAACCAGCGCCAGCGACATCATTAGCAAGCTGG
GCCGTGCGGTGGACGGCCCGAAACGCCCGCGTCCGGTGGTGCGTGAAACAGACCCAG
GAGAAGGGTTGGGTGCGTGCCTGGCCCGTCGCAAACGCGAATACCGTGAGGCTCA
GTACCTGTATAGCCGCGACCAAGGCTCGTCTGGCTGCGCAGATCCTGGACGGCGCCG
CTAGCCAGGAGTGCGCCCTGCCAGTGGACCAGGTGTATGGTGCTTTCCGTGAGAAG
TGGGAAACCGTGGGCCAGTTCCACGGTCTGGGCGAATTTGCGACCGGTGCGCGTGC
TGATAACTGGGAGTTCTACAGCCCGATCCTGGCCGCGGAAGTGAAAGGAGAACCTGA
TGCGCATGGCTAATGGTACCGCCCCAGGTCCAGACCGTATTAGCAAGAAAGCCCTG
CTGGATTGGGACCCGCGTGGTGAACAGCTGGCCCGTCTGTATACCACCTGGCTGAT
CGGTGGCGTGATTCCGCGCGTGTTCAAGGAATGCCGTACCAAGCTGCTGCCGAAGA

20

30

40

50

GCAGCGACCCGGTGGAGCTGCAGGATATCGGTGGCTGGCGTCCGGTGACCATTGGT
AGCATGGTGACCCGCCTGTTTAGCCGTATCCTGACCATGCGTCTGACCCGCGCGTG
CCCGATTAACCCGCGCCAGCGTGGCTTCCTGGCCAGCAGCAGCGGTTGCGCCGAGA
ACCTGCTGATCTTTGACGAGATTGTGCGTCGCAGCCGTGCGGATGGTGGCCCACTG
GCCGTGGTGTTCGTGGACTTTGCCCGTGCCTTCGACAGCATCAGCCACGAACACAT
TCTGTGCGTGCTGGAAGAGGGTGGCCTGGACCGTCACGTGATCGGTCTGATTTCGTA
ACAGCTACGTGGATTGCGTGACCCGTGTGGGCTGCGTGGAGGGTATGACCCCGCCG
ATCCAGATGAAAGTGGGCGTGAAACAGGGTGACCCGATGAGCCCGCTGCTGTTTAA
CCTGGCGATGGACCCGCTGATCCACAAGCTGGAAACCGCTGGTACCGGCCTGAAAT
GGGGCGACCTGAGCATTGCCACCCTGGCTTTCCGCCGACGATCTGGTGCTGGTGAGC
GATAGCGAAGAGGGTATGGGCCGTAGCCTGGGCATTCTGGAGAAGTTCTGCCAGCT
GACCGGTCTGCGTGTGCAGCCGCGTAAGTGCCACGGCTTCTTTATGGATAAAGGTG
TGGTGAACGGTTGCGGCACCTGGGAAATCTGCGGCAGCCCGATCCACATGATTCCG
CCGGGTGAGAGCGTGCGCTACCTGGGTGTGCAAGTGGGTCCGGGTCTGGTGAT
GGAGCCAGACCTGATTCCGACCGTGACACACCTGGATCGAACGCATTAGCGAAGCCC
CGCTGAAACCGAGCCAGCGCATGCGTGTGCTGAACAGCTTCGCTCTGCCGCGTATC
ATCTACCAGGCCGACCTGGGCAAGGTGACCGTGACCAAACCTGGCCCAGATCGATGG
TATTGTGCGTAAGGCGGTGAAGAAATGGCTGCACCTGAGCCCGAGCACCTGCAACG
GCCTGCTGTACAGCCGCAACCGTGATGGTGGCCTGGGTCTGCTGAAGCTGGAGCGT
CTGATCCCGAGCGTGCGCACCAACGTATCTACCGCATGAGCCGTAGCCCGGACAT
CTGGACCCGTCGCATGACCAGCCACAGCGTGAGCAAGAGCGATTGGGAAATGCTGT
GGGTGCAGGCGGGTGGCGAGCGTGGCAGCGCCCCGGTTATGGGTGCGGTGGAAGCG
GCTCCAACCGACGTGGAGCGTAGCCCGGACTACCCAGATTGGCGTCGCGAAGAGAA
CCTGGCCTGGAGCGCGCTGCGTGTGCAGGGCGTGGGTGCCGACCAGTTCGCGGGCG
ACCGTACCAGCAGCAGCTGGATCGCTGAACCAGCCAGCGTGGGCTTCGCCCAGCGT
CACTGGCTGGCGGCTCTGGCGCTGCGTGCTGGCGTGTATCCAACCCGCGAATTCCT
GGCCCGTGGCAAGGAGAAGAGCGGTGCTGCCTGCCGTGCTGCCAGCTCGTCTGG
AGAGCTGCAGCCACATTCTGGGCCAGTGCCCGTTTGTGCAAGGCTAACCGCATTGCC
CGTCACAACAAAGTGTGCGTGCTGCTGGCTACCGAAGCCGAGCGTTTCGGCTGGAC
CGTGATCCGCGAATTTCTGCTGAGGACGCTGCCGGCGGCCTGAAGATCCCGGATC
TGGTGTGCAAGAAAGCCGACACCGTGCTGATTGTGGACGTGACCGTGCGCTACGAA
ATGGACGGTGAAACCCTGAAGCGTGCGGCTAGCGAGAAGGTGAAGCACTATCTGCC
GGTGGGTGAGCAGATCACCGATAAAGTGGGTGGCCGCTGCTTCAAAGTGATGGGCT
TTCCGGTGGGTGCTCGTGGCAAGTGGCCGGCCAGCAACAATACCGTGCTGGCCGAA
CTGGGCGTGCCGGCTGGTTCGCATGCGTACCTTCGCGCGTCTGGTGAGCCGTGCGAC
CCTGCTGTATAGCCTGGACATTCTGCGCGACTTCATGCGTGAACCAGCAGGTCTGTG
GCACCCGTGTGGCACTGATTCCAGCAGCAACCGGCGCAGCAAATTGATAA

10

20

30

14. R2A OrLa 合成 N-DYK アミノ酸配列:

MDYKDDDDKGTDTVYVGQDYP SGLSKRVPARLVAGPMLRERSCHAHVFRAGHMW
NWRTSLPSGRWDQPALEKSRVLTRSVATATDPEITSYPGKSVSTSTQVQEEDWCSRE
SGWISPGLAPEEPSVVSEITASMVATMRVATEEVVLEPQPEQVV TILPEHGRNVPPG
LAEQDTASPIEVSVLLPDLAENCPLCGVPSGGLRLLGKHFAVRHAGVPV TYE CRKCA
WRSPNSHSISCHVPKCRGRARMPSGDPGIACDLCEARFATEVGVAQHKRHVHPVEW
NKVRLERRGARGGGIKATKLWSVAEVETLIRLIREHGD SGATYQLIADELGRGKTAE
QVRSKKRLLRIDTASNPDDAEVEEERLES LAVRSSSRPPSLVATRVREAVARGESE
GGEEIRAI AALIRDVDQNPCLIETSASDIISKLGRRVDGPKRPRPVVREQTQEKGWVR
RLARRKREYREAQYLYSRDQARLAAQILDGAASQECALPVDQVYGAFREK WETV GQ
FHGLGEFRTGARADNWEFYSPILAAEVKENLMRMANGTAPGPDRI SKKALLDWDPR
GEQLARLYTTWLIGGVIPRVFKECRTKLLPKSSDPVELQDIGGWRPV TIGSMVTRLF
SRILTMRLTRACPINPRQRGFLASSSGCAENLLIFDEIVRRSRRDGGPLAVVFVDFAR

40

50

AFDSISHEHILCVLEEGGLDRHVIGLIRNSYVDCVTRVGCVEGMTTPPIQMKVGVKQG
DPMSPLLFLNAMDPLIHKLETAGTGLKWGDLSIATLAFADDLVLVSDSEEGMGRSLG
ILEKFCQLTGLRVQPRKCHGFFMDKGVVNGCGTWEICGSP IHMIPPGESVRYLGVQV
GPGRGVM EPDLIPTVHTWIERISEAPLKPSQRM RVLNSFALPRIIYQADLGKVTVTK
LAQIDGIVRKAVKKWLHLSPSTCNGLLYSRNRDGGGLLLKLERLIPSVRTKRIYRMSR
SPDIWTRRMTSHSVSKSDWEMLWVQAGGERGSAPVMGAVEAAPT DVERS PDYPDW
RREENLAWSALRVQGVGADQFRGDRTSSSWIAEPASVGFAQRH WLAALALRAGVYP
TREFLARGKEKSGAACRRCPARLESCSHILGQCPFVQANRIARHNKVCVLLATEAER
FGWTVIREFRLEDAAGGLKIPDLVCKKADTVLIVDVT VRYEMDGETLKRAASEKVKH
YLPVGQQITDKVGGRCFKVMGFPVGARGKWPASNNTVLAELGVPAGRMRTFARLVS
RRTL LYSLDILRDFMREPAGRGTRVALIPAATGAAN**

10

15. R2A OrLa_s合成N-DYKプラスミドORF配列:

ATGGATTACAAGGATGACGACGATAAGGGTACCGATCTGGCGGAGAACTGCCCACT
GTGCGGTGTGCCGAGCGGCGGTCTGCGTCTGCTGGGCAAGCACTTCGCGGTGCGTC
ATGCTGGTGTGCCAGTGACCTACGAGTGCCGCAAATGCGCGTGCGGTAGCCCGAAC
AGCCACAGCATCAGCTGCCATGTGCCGAAGTGCCGCGGTCTGTGCTCGTATGCCGAG
CGGCGACCCGGGTATTGCTTGCGATCTGTGCGAAGCGCGCTTTGCTACCGAGGTGG
GTGTGGCCCAGCACAAAGCGTCACGTGCACCCGGTGGAATGGAACAAAGTGCGCCTG
GAGCGTCGCGGTGCCCGTGCGGTGGCATCAAGGCGACCAAACCTGTGGAGCGTGGC
TGAAGTGGAACCCCTGATCCGTCTGATTCTGTGAGCACGGCGACAGCGGTGCGACCT
ATCAGCTGATCGCTGATGAACTGGGTCTGTGGCAAGACCGCCGAGCAGGTGCGTAGC
AAGAAACGCCTGCTGCGTATTGACACCGCCAGCAACAGCCCGGACGATGCGGAAGT
GGAGGAAGAGCGCCTGGAGAGCCTGGCCGTGCGTAGCAGCAGCCGTAGCCCGCCGA
GCCTGGTGGCCACCCGCGTGCGTGAAAGCGGTGGCTCGTGGTGAAAGCGAGGGTGGC
GAAGAGATCCGCGCCATTGCCGCGCTGATCCGTGACGTGGATCAGAACCCGTGCCT
GATTGAAACCAGCGCCAGCGACATCATTAGCAAGCTGGGCGCTCGCGTGGACGGCC
CGAAACGCCC GCGTCCGGTGGT GCGTGAAACAGACCCAGGAGAAGGGTTGGGTGCGT
CGCCTGGCCCGTCGCAAACGCGAATACCGTGAGGCTCAGTACCTGTATAGCCGCGA
CCAGGCTCGTCTGGCTGCGCAGATCCTGGACGGCGCCGCTAGCCAGGAGTGCGCCC
TGCCAGTGGACCAGGTGTATGGTGCTTTCCGTGAGAAGTGGGAAACCGTGGGCCAG
TTCCACGGTCTGGGCGAATTTGCGACCGGTGCGCGTGCTGATAACTGGGAGTTCTA
CAGCCCGATCCTGGCCGCGGAAGTGAAGGAGAACCTGATGCGCATGGCTAATGGTA
CCGCCCCAGGTCCAGACCGTATTAGCAAGAAAGCCCTGCTGGATTGGGACCCGCGT
GGTGAACAGCTGGCCCGTCTGTATACCACCTGGCTGATCGGTGGCGTGATTCCGCG
CGTGTTCAAGGAATGCCGTACCAAGCTGCTGCCGAAGAGCAGCGACCCGGTGGAGC
TGCAGGATATCGGTGGCTGGCGTCCGGTGACCATTTGGTAGCATGGTGACCCGCGCTG
TTTAGCCGTATCCTGACCATGCGTCTGACCCGCGCGTGCCCGATTAACCCGCGCCA
GCGTGGCTTCCTGGCCAGCAGCAGCGGTTGCGCCGAGAACCTGCTGATCTTTGACG
AGATTGTGCGTCGCAGCCGTCGCGATGGTGGCCCACTGGCCGTGGTGTTCGTGGAC
TTTGGCCGTGCCTTCGACAGCATCAGCCACGAACACATTCTGTGCGTGCTGGAAGA
GGGTGGCCTGGACCGTCACGTGATCGGTCTGATTCTGTAACAGCTACGTGGATTGCG
TGACCCGTGTGGGCTGCGTGGAGGGTATGACCCCGCCGATCCAGATGAAAGTGGGC
GTGAAACAGGGTGACCCGATGAGCCCGCTGCTGTTTAACCTGGCGATGGACCCGCT
GATCCACAAGCTGGAAACCGCTGGTACCGGCCTGAAATGGGGCGACCTGAGCATTG
CCACCCTGGCTTTCGCCGACGATCTGGTGCTGGTGAGCGATAGCGAAGAGGGTATG
GGCCGTAGCCTGGGCATTCTGGAGAAGTTCTGCCAGCTGACCGGTCTGCGTGTGCA
GCCGCGTAAGTGCCACGGCTTCTTTATGGATAAAGGTGTGGTGAACGGTTGCGGCA
CCTGGGAAATCTGCGGCAGCCCGATCCACATGATTCCGCCGGGTGAGAGCGTGCGC
TACCTGGGTGTGCAAGTGGGTCCGGGTCTGGTGTGATGGAGCCAGACCTGATTCC
GACCGTGACACCTGGATCGAACGCATTAGCGAAGCCCCGCTGAAACCGAGCCAGC

20

30

40

50

GCATGCGTGCTGAACAGCTTCGCTCTGCCGCGTATCATCTACCAGGCCGACCTG
GGCAAGGTGACCGTGACCAAACTGGCCCAGATCGATGGTATTGTGCGTAAGGCGGT
GAAGAAATGGCTGCACCTGAGCCCGAGCACCTGCAACGGCCTGCTGTACAGCCGCA
ACCGTGATGGTGGCCTGGGTCTGCTGAAGCTGGAGCGTCTGATCCCGAGCGTGCGC
ACCAAACGTATCTACCGCATGAGCCGTAGCCCGGACATCTGGACCCGTCGCATGAC
CAGCCACAGCGTGAGCAAGAGCGATTGGGAAATGCTGTGGGTGCAGGCGGGTGGCG
AGCGTGGCAGCGCCCCGGTTATGGGTGCGGTGGAAGCGGCTCCAACCGACGTGGAG
CGTAGCCCGGACTACCCAGATTGGCGTCGCGAAGAGAACCTGGCCTGGAGCGCGCT
GCGTGTGCAGGGCGTGGGTGCCGACCAGTTCGCGGGCGACCGTACCAGCAGCAGCT
GGATCGCTGAACCAGCCAGCGTGGGCTTCGCCCAGCGTCACTGGCTGGCGGCTCTG
GCGCTGCGTGCTGGCGTGTATCCAACCCGCGAATTCTGGCCCGTGGCAAGGAGAA
GAGCGGTGCTGCCTGCCGTGCTGCCAGCTCGTCTGGAGAGCTGCAGCCACATTC
TGGGCCAGTGCCCGTTTGTGTCAGGCTAACCGCATTGCCCGTCACAACAAAGTGTGC
GTGCTGCTGGCTACCGAAGCCGAGCGTTTCGGCTGGACCGTGATCCGCGAATTTCTG
TCTGGAGGACGCTGCCGGCGGCGCTGAAGATCCCGGATCTGGTGTGCAAGAAAGCCG
ACACCGTGCTGATTGTGGACGTGACCGTGCGCTACGAAATGGACGGTGAACCCCTG
AAGCGTGCGGCTAGCGAGAAGGTGAAGCACTATCTGCCGGTGGGTGAGCAGATCAC
CGATAAAGTGGGTGGCCGCTGCTTCAAAGTGATGGGCTTTCCGGTGGGTGCTCGTG
GCAAGTGGCCGGCCAGCAACAATACCGTGCTGGCCGAACTGGGCGTGCCGGCTGGT
CGCATGCGTACCTTCGCGCGTCTGGTGAGCCGTGCGACCCTGCTGTATAGCCTGGA
CATTCTGCGCGACTTCATGCGTGAACCAGCAGGTCGTGGCACCCGTGTGGCACTGA
TTCCAGCAGCAACCGGCGCAGCAAATTGATAA

10

20

30

40

50

16. R2A OrLa_s 合成N-DYKアミノ酸配列:

MDYKDDDDKGTDLAENCPLCGVPSGGLRLLGKHFAVRHAGVPVTYECRKCAWRSP
NSHSISCHVPKCRGRARMPSPGDPGIACDLCEARFATEVGVAQHKRHVHPVEWNKVR
LERRGARGGGIKATKLWSVAEVETLIRLIREHGD SGATYQLIADELGRGKTAEQVRSK
KRLLRIDTASNPDDAEVEEERLESLAVRSSSRSPPSLVATRVREAVARGESEGGEEIR
AIAALIRDVDQNPCLIETSASDIISKLGRRVDGPKRPRPVVREQTQEKGWVRRRLARRK
REYREAQYLYSRDQARLAAQILDGAASQECALPVDQVYGAFREKWETVGQFHLGGE
FRTGARADNWEFYSPILAAEVKENLMRMANGTAPGPDRISKKALLDWDPRGEQLAR
LYTTWLIGGVIPRVFKECRTKLLPKSSDPVELQDIGGWRPV TIGSMVTRLFSRILTM R
LTRACPINPRQRGFLASSSGCAENLLIFDEIVRRSRRDGGPLAVVFVDFARAFDSISHE
HILCVLEEGGLDRHVIGLIRNSYVDCVTRVGCVEGMTPIIQMKVGVKQGDPMSPLLF
NLAMDPLIHKLETAGTGLKWGDL SIATLAFADDLVLVSDSEEGMGRSLGILEKFCQL
TGLRVQPRKCHGFFMDKGVVNGCGTWEICGSP IHMIPPGESVRYLG VQVGPGRGVM
EPDLIPTVHTWIERISEAPLKPSQRM RVLNSFALPRIIYQADLGKVTVT KLAQIDGIV
RKAVKKWLHLSPSTCNGLLYSRNRDGG LGLLKLERLIPSVRTKRIYRMSRSPDIWTR
RMTSHSVSKSDWEMLWVQAGGERGSAPVMGAVEAAPT DVERS PDYPDW RREENLA
WSALRVQGVGADQFRGDRTSSSWIAEPASVGFAQRH WLAALALRAGVYPTREFLAR
GKEKSGAACRRCPARLESCSHILGQCPFVQANRIARHNKVCVLLATEAERFGWTVIR
EFRLEDAAGGLKIPDLVCKKADTVLIVDVT VRYEMDGETLKRAASEKVKHYLPVGQQ
ITDKVGGRCFKVMGFPVGARGKWPASNNTVLAELGVPAGRMRTFARLVSRRTL LYS
LDILRDFMREPAGRGTRVALIPAATGAAN**

17. R2A ZoAl N-DYKプラスミドORF配列:

ATGGATTACAAGGATGACGACGATAAGGGTACCATGAACATCGTGAAGGTGACCGT
GCCCGACAAGAATCCCCCTTGTCTTGCTGTGGCGTGCGCCTGAACAGCGTGCTGG
CACTGATCGAGCACCTGAAGGGCTCCACGGAAGGAGAAGGGTGTGCTTCCGCTGT
GCCAAGTGCGGCCGGGAGAACTTTAATCACACAGCACCGTGTGTCACTACGCCAA
GTGCAAGGGCCACAGATCGAGAGACCACCCGTGGGCGAGTGGATCTGTGAGGTGT
GCGGCAGGGATTTACCACAAAGATCGGCCTGGGCCAGCACAAAGCGCCACATGCAC

GCCATGGTGC GCAACCAGGAGCGGATCGACGCCAGCCAGCCCAAGGAGACCTCCAA
TCGCGGCGCC CACAAGCGGTGCTGGACAAAGGAGGAGGAGGAGCTGCTGATGAAGC
TGGAGGTGCAGTTT GAGAACCACAAGAACATCAATAAGCTGATCGCCGAGCAGCTG
ACCACAAAGACCGCCAAGCAGATCTCCGATAAAGCGCCGGATGCTGCTGAAGAAGGG
CAGAGGCACCACAGGCAACCTGGAGACAGAGCCTGGCATGAGCCACCAGTCCCAGG
CCAAGGTGAAGGACAATGGACTGGGAGGCGATCACCTGCCAGGAGGACCTGTGGTG
GACAAGGGCACCATCGGCAAGCCAGGCCAGCACCTGGACACCGATAATTCTCACC
GATCACAGCAGGCAAGAAGAAGGGAGGAGGACTGCAGGCCAGATACAGAAGGCGCA
TCATGAAGAGGCTGGCCGCCGGCACCATCAACATCTTCCCAAAGGTGTTTAAGGAG
CTGATCAACGACCAGGAGGCCAGGCCCTGATCAATCAGACCACAGAGGACTGTTT
CGGCCTGCTGGATTCCGCCTGCCAGATCAGAACCGCCCTGAGGGAGAAGGGCAAGT
CTCAGGAGGAGCGGCC CAGAAAGCAGTACCAGAAGTGATGAAGAAGCGCGCCATC
AAGCGGGGCGACTATCTGCGGTTCCAGAGACTGTTTACCTGGATAGAGGCAAGCT
GGCCAGGATCATCCTGGACAACACCGAGTCCCTGTCTTGTGATATCAGCCCTTCCG
AGATCTACAGCGTGTTCAAGGCCCGCTGGGAGACACCAGGCCACTTCAATGGCCTG
GGCGATTTTGAGATCAAGGGCAAGGCCAACATAAGGCCTTCCGGGACTTTATCAC
CGCCAAGGAGATCGAGAAGAACGTGAGAGAGATGTCTAAGGGCAGCGCACCCAGGAC
CTGATGGAATCGCCCTGGGCGACATCAAGAAGATGGACCCCGGCTATTCTCGGACC
GCCGAGCTGTTTAATCTGTGGCTGACAGCCGGCGACATCCCAGATATGGTGAGGGG
ATGCAGGACCGTGCTGATCCCAAAGAGCACCACACCCGAGCGCCTGAAGGACATCA
ACAATTGGCGGCCCATCACAATCGGCTCCATCCTGCTGAGACTGTTCTCTAGGATC
ATCACCGCCAGAATGACAAAGGCATGTCCACTGAACCCAAGGCAGAGGGGCTTTAT
CTCCGCACCAGGCTGCTCTGAGAACCTGAAGCTGCTGCAGAGCATCATCAGGACCG
CCAAGAATGAGCACAAGCCCCCTGGGCGTGATCTTCGTGGACATCGCCAAGGCCTTT
GATACAGTGTCACACAGCACATCATCCACGTGCTGCAGCAGCGGAGAGTGGAACCC
CCACATCGTGGGCCTGGTGAACAATATGTACAAGGACATCTCCACCTATGTGACCA
CAAAGAAGAACACCCACACAGACAAGATCCAGATCAGAGTGGGCGTGAAGCAGGGC
GATCCTCTGTCTCCACTGCTGTTCAATCTGGCCATGGACCCACTGCTGTGTAAGCTG
GAGGAGAGCGGCAAGGGCTTCCACAGGGGGCCAGAGCTCCATCACCGCCATGGCCTT
TGCCGACGATCTGGTGCTGCTGTCTGATAGCTGGGAGAACATGAAGGAGAATATCA
AGATCCTGGAGACATTCTGTAATCTGACCGGCCTGAAGACACAGGGCCAGAAGTGC
CACGGCTTTTACATCAAGCCTACCAAGGACTCCTATACAATCAACAATTGCCCAGC
CTGGACCATCAACGGCACACCTCTGAACATGATCAATCCAGGCGAGTCTGAGAAGT
ACCTGGGACTGCAGATCGACCCATGGACCGGCGTGGCCAAGTATGATCTGAGCACA
AAGCTGAAGATCTGGCTGGAGTCCATCGACAGGGCCCCCTGAAGCCTCTGCAGAA
GCTGGATATCCTGAAGACCTACACAATCCCCCGCCTGACCTATCTGGCCGATCACA
GCGAGATGAAGGCAGGAGCACTGGAGGCACTGGACCAGCAGATCAGGACCGCCGTG
AAGGATTGGCTGCACCTGCCTTCTGACATGTGACGCCATCCTGTACGTGTCTAC
CAGGGATGGAGGACTGGGAGTGACAAAGCTGGCAGGACTGATCCCATCTGTGCAAG
CAAGGAGGCTGCACAGGATCGCACAGAGCCCTGACGAGACCATGAAGGATTTCTG
GAGAAGGCC CAGATGGAGAAGATGTACGAGAAGCTGTGGGTGCAGGCCGGCGGCAA
GAAGAAGGGCATGCCATCCATCTGGGAGGGCCCTGCCATGACAGTGCCTCCAACCA
ACACAGGCAATCTGTCTGAGTGGGAGGCCCAAATCCCAAGAGCAAGTATCCTAAG
CCATGCGACTGGCGGAGAAAGGAGCTGAAGAAGTGGACCAAGCTGGAGAGCCAGGG
CAGAGGCGTGAAGAACTTTAGGAATGACACAATCTCCAACGATTGGATTCACTACT
ATAGGCGCATCCACACCGCAAGCTGCTGACCGCCATCCAGCTGCGCGCCAACGTG
TACCCAACAAGGGAGTTCTGGCACGGGGCAGAGGCGATAATTATGTGAAGTTTTG
TCGGCACTGCGAGGCCGACCTGGAGACCTGTGGCCACATCATCGGCTTCTGCCCCG
TGACAAAGGACGCCCGGATCAAGAGACACAACAGGATCTGTGATAGACTGTGCGAG
GAGGCAGCAAAGAGGGGAGTGGGTGGTGTTTAAGGAGCCCCACCTGAGGGACGCCAC

10

20

30

40

50

CACAGAGCTGTTCAAGCCTGATGTGATCTTTGTGAAGGAGGACAGGGCACTGGTGG
TGGATGTGACCGTGCGGTACGAGTCCGCCAAGACCACACTGGAGGCTGCCGCCATG
GAGAAGGTGGATAAGTATAAGCACCTGGAGGCCGAGGTGAAGGAGCTGACAAATGC
CAAGGACGTGGTGTTCATGGGATTTCCACTGGGAGCAAGGGGCAAGTTCTACAAGG
GCAACTTTAATCTGCTGGAGACCCTGGGCCTGCCTAAGACAAGACAGCTGTCTGTG
GCCAAGACCCTGAGCACCTACGCCCTGATGTCTAGCGTGGACATCGTGCACATGTT
CGCCTCCAGGTCTCGCAAGCCTAACGTGTAG

18. R2A ZoAl N-DYKアミノ酸配列:

MDYKDDDDKGTMNIVKVTVPDKNPPCPCCGVRLNSVLALIEHLKGSHGRRRVCFRC
AKCGRENFNHSTVCHYAKCKGPQIERPPVGEWICEVCGRDFTTKIGLGQHKRHMH
AMVRNQERIDASQPKETSNRGAAHKRCWTKEEEELMKLEVQFENHKNINKLIAEQL
TTKTAKQISDKRRMLLKKGRGTTGNLETEPGMSHQSQAKVKDNLGGDHLPGGPVV
DKGTIGKPGQHLDTDNSHQITAGKKKGGGLQARYRRRIMKRLAAGTINIFPKVFKELI
NDQEARPLINQTTEDCFGLLDSACQIRTAALREKGSQEERPRKQYQKWMKKRAIKRG
DYLRQFQRLFHLDRGKLARIILDNTESLSCDISPSEIYSVFKARWETPGHFNGLGDFEIK
GKANNKAFRDFITAKEIEKNVREMSKGSAPGPDGIALGDIKKMDPGYSRTAELFNLW
LTAGDIPDMVRGCRTVLIPKSTTPERLKDINNWRPITIGSILLRLFSRIITARMTKACP
LNPRQRGFISAPGCSENKLLQSIIRTAKEHKKPLGVIFVDIAKAFTVSHQHIIHVLQ
QRRVDPHIVGLVNNMYKDISTYVTTKKNTHTDKIQIRVGKQGDPLSPLLFNLAMD
PLLCKLEESGKGFHRGQSSITAMAFADDLVLLSDSWENMKENIKILETFCNLTGLKT
QGQKCHGFYIKPTKDSYTINNCPAWTINGTPLNMINPGESEKYLGLQIDPWTGVAKY
DLSTKLKIWLESIDRAPLKPLQKLDILKTYTIPRLTYLADHSEMKAAGALEALDQQIRT
AVKDWLHLPSCCTCDAILYVSTRDGGGLGVTKLAGLIPSVQARRLHRIAQSPDETMKDF
LEKAQMEKMYEKLWVQAGGKKKGMPsiWEALPMTVPPTNTGNLSEWEAPNPKSKY
PKPCDWRRKELKKWTKLESQGRGVKNFRNDTISNDWIQYYRRIPHRKLLTAIQLRA
NVYPTREFLARGRDNYVKFCRHCEADLETGHIIGFCPVTKDARIKRHNRCRDLRC
EEAAKREWVVFKEPHLRDATTELFKPDVIFVKEDRALVVDVTVRYESAKTTLAAA
MEKVDKYKHLEAEVKELTNAKDVFVMGFPLGARGKFYKGNFNLLETGLPKTRQLS
VAKTLSTYALMSSVDIVHMFASRSRKP NV*

19. R2A ZoAl N-DYKトランスフェクトしたmRNA配列:

AGGAUAAACUAGUAUUCUUCUGGUCCCCACAGACUCAGAGAGAACCCGCCACCA
UGGAUUACAAGGAUGACGACGAUAAGGGAUCCAACAUCGUGAAGGUGACCGUGCC
CGACAAGAAUCCCCCUUGUCCUUGCUGUGGCGUGCGCCUGAACAGCGUGCUGGCA
CUGAUCGAGCACCUGAAGGGCUGCCACGGAAGGAGAAGGGUGUGCUUCCGCUGUG
CCAAGUGCGGCCGGGAGAACUUUAUACACACAGCACCUGUGUGUCACUACGCCAA
GUGCAAGGGCCACAGAU CGAGAGACCACCCGUGGGCGAGUGGAUCUGUGAGGUG
UGCGGCAGGGAUUUCACCACAAAGAU CGGCCUGGGCCAGCACAAAGCGCCACAUGC
ACGCCAUGGUGCGCAACCAGGAGCGGAUCGACGCCAGCCAGCCCAAGGAGACCUCC
AAUCGCGGCGCCACAAGCGGUGCUGGACAAAGGAGGAGGAGGAGCUGCUGAUGA
AGCUGGAGGUGCAGUUUGAGAACCACAAGAACAUCAAUAAGCUGAUCGCCGAGCA
GCUGACCACAAAGACCGCCAAGCAGAU CUCGUAAGCGCCGGAUGCUGCUGAAG
AAGGGCAGAGGCACCACAGGCAACCUGGAGACAGAGCCUGGCAUGAGCCACCAGU
CCCAGGCCAAGGUGAAGGACAAUGGACUGGGAGGGCGAUCACCUGCCAGGAGGACC
UGUGGUGGACAAGGGCACCAUCGGCAAGCCAGGCCAGCACCUGGACACCGAUAAU
UCUCACCAGAU CACAGCAGGCAAGAAGAAGGGAGGAGGACUGCAGGCCAGAUACA
GAAGGCGCAUCAUGAAGAGGCUGGCCGCGGCCACCAUCAACAUCUCCCAAAGGU
GUUUAAGGAGCUGAUCAACGACCAGGAGGCCAGGCCCCUGAUCAAU CAGACCACA
GAGGACUGUUUCGGCCUGCUGGAUUCGCCUGCCAGAU CAGAACC GCCCUGAGGG
AGAAGGGCAAGUCUCAGGAGGAGCGGCCCAGAAAGCAGUACCAGAAGUGGAUGAA
GAAGCGCGCCAUCAAGCGGGGCGACUAUCUGCGGUUCCAGAGACUGUUU CACCUG

GAUAGAGGCAAGCUGGCCAGGAUCAUCCUGGACAACACCGAGUCCUGUCUUGUG
AUUUCAGCCCUUCCGAGAUUCUACAGCGUGUUAAGGCCCGCUGGGAGACACCAGG
CCACUUCAAUGGCCUGGGCGAUUUUGAGAUCAAGGGCAAGGCCAACAAUAAGGCC
UUCCGGGACUUUAUCACCGCCAAGGAGAUUCGAGAAGAACGUGAGAGAGAUGUCUA
AGGGCAGCGCACCAAGGACCUGAUGGAAUCGCCCUGGGCGACAUCAAGAAGAUGGA
CCCCGGCUAUUCUCGGACCGCCGAGCUGUUUAAUCUGUGGCUGACAGCCGGCGAC
AUCCCAGAUUUGGUGAGGGGAUGCAGGACCGUGCUGAUCCCAAAGAGCACACAC
CCGAGCGCCUGAAGGACAUCAACAAUUGGCGGCCCAUCACAAUCGGCUCCAUCCU
GCUGAGACUGUUCUCUAGGAUCAUACCGCCAGAAUGACAAAGGCAUGUCCACUG
AACCCAAGGCAGAGGGGCUUUAUCUCCGCACCAGGCUGCUCUGAGAACCUGAAGC
UGCUGCAGAGCAUCAUCAGGACCGCCAAGAAUGAGCACAAAGCCCCUGGGCGUGAU
CUUCGUGGACAUCGCCAAGGCCUUUGAUACAGUGUCCCAACAGCACAUCAUCCAC
GUGCUGCAGCAGCGGAGAGUGGACCCCCACAUCGUGGGCCUGGUGAACAAUAUGU
ACAAGGACAUCUCCACCUAUGUGACCACAAAGAAGAACACCCACACAGACAAGAU
CCAGAUCAAGAGUGGGCGUGAAGCAGGGCGAUCCUCUGUCUCCACUGCUGUCAAU
CUGGCCAUGGACCCACUGCUGUGUAAGCUGGAGGAGAGCGGCAAGGGCUUCCACA
GGGGCCAGAGCUCCAUCACCGCCAUGGCCUUUGCCGACGAUCUGGUGCUGCUGUC
UGAUAGCUGGGAGAACAUAGAAGGAGAAUAUCAAGAUCUGGAGACAUUCUGUAAU
CUGACCGGCCUAAAGACACAGGGGCCAGAAGUGCCACGGCUUUUACAUCAAGCCUA
CCAAGGACUCCUAUACAUAACAUAUUGCCCAGCCUGGACCAUCAACGGGCACACC
UCUGAACAUGAUCAAUCCAGGCGAGUCUGAGAAGUACCUGGGACUGCAGAUCGAC
CCAUGGACCGGCGUGGCCAAGUAUGAUCUGAGCACAAAGCUGAAGAUUCUGGCUGG
AGUCCAUCGACAGGGCCCCCUGAAGCCUCUGCAGAAGCUGGAUAUCCUGAAAC
CUACACAAUCCCCCGCCUGACCUAUCUGGCCGAUCACAGCGAGAUGAAGGCAGGA
GCACUGGAGGCACUGGACCAGCAGAUCAAGGACCGCCGUGAAGGAUUGGCUGCACC
UGCCUUCUGCACAUGUGACGCCAUCCUGUACGUGUCUACCAGGGGAUGGAGGACU
GGGAGUGACAAAGCUGGCAGGACUGAUCCCAUCUGUGCAGGCAAGGAGGCUGCAC
AGGAUCGCACAGAGCCCUGACGAGACCAUGAAGGAUUUCCUGGAGAAGGCCCAGA
UGGAGAAGAUGUACGAGAAGCUGUGGGUGCAGGCCCGGCCAAGAAGAAGGGCAU
GCCAUCCAUCUGGGAGGCCUGCCCAUGACAGUGCCUCCAACCAACACAGGCAAU
CUGUCUGAGUGGGAGGCCCAAUCCCAAGAGCAAGUAUCCUAAGCCAUGCGACU
GGCGGAGAAAGGAGCUGAAGAAGUGGACCAAGCUGGAGAGCCAGGGCAGAGGCGU
GAAGAACUUUAGGAUUGACACAAUCUCCAACGAUUGGAUUCAGUACUAUAGGCGC
AUCCCACACCGCAAGCUGCUGACCGCCAUCCAGCUGCGCGCCAACGUGUACCCAAC
AAGGGAGUUCUGGCACGGGGCAGAGGCGAUAAUUAUGUGAAGUUUUGUCGGCAC
UGCGAGGCCGACCUGGAGACCUGUGGCCACAUCAUCGGCUUCUGCCCCGUGACAA
AGGACGCCCAGGAUCAAGAGACACAACAGGAUCUGUGAUAGACUGUGCGAGGAGGC
AGCAAAGAGGGAGUGGGUGGUUUUAAAGGAGCCCCACCUGAGGGACGCCACCACA
GAGCUGUUAAGCCUGAUGUGAUCUUUGUGAAGGAGGACAGGGCACUGGUGGUGG
AUGUGACCGUGCGGUACGAGUCCGCCAAGACCACACUGGAGGCUGCCGCCAUGGA
GAAGGUGGAUAAGUAUAAGCACCUUGGAGGCCGAGGUGAAGGAGCUGACAAAUGCC
AAGGACGUGGUGUUAUGGGGAUUUCCACUGGGAGCAAGGGGCAAGUUCUACAAGG
GCAACUUUAAUCUGCUGGAGACCCUGGGCCUGCCUAAGACAAGACAGCUGUCUGU
GGCCAAGACCCUGAGCACCUACGCCUGAUGUCUAGCGUGGACAUCGUGCACAUG
UUCGCCUCCAGGUCUCGCAAGCCUAACGUGUGAUGACCUCGAGCUGGUACUGCAU
GCACGCAAUGCUAGCUGCCCCUUUCCCGUCCUGGGUACCCCGAGUCUCCCCGAC
CUCGGGUCCCAGGU AUGCUCCCAACCUCACCUGCCCCACUCACCACCUCUGCUAG
UUCAGACACCUCCAAGCACGCAGCAAUGCAGCUAAAACGCUUAGCCUAGCCA
CACCCCCACGGGAACAGCAGUGAUUAACCUUUAGCAAUAAACGAAAGUUUAACU
AAGCUAUACUAACCCCAGGGUUGGUCAAUUUCGUGCCAGCCACACCCUGGAGCUA

10

20

30

40

50

GCAAAAAAAAAAAAAAAAAAAAAAAAAAGCAUAUGACUAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AA

20. R2A ZoAlアミノ酸配列:

MNIVKVTVPDKNPPCPCCGVRLNSVLALIEHLKGS HGRRRVCFRCAKCGRENFNHH
STVCHYAKCKGPQIERPPVG EWICEVCGRDFTTKIGLGQH KRHMHAMVRN QERIDA
SQPKETS NRGAHKRCWTK EEEELLMKLEVQFENHKNINKLIAEQ LTTKTAKQISDKR
RMLLKKGRGTTGNLETEPGMSHQSAKVKDNLGGDHLPGGPVVDKGTIGKPGQH
LDTDNSHQITAGKKKGGGLQARYRRRIMKRLAAGTINIFPKVFKELINDQEARPLINQ
TTEDCFGL LDSACQIR TALREK GKSQEERPRKQYQKWMKKRAIKRGDYLRFQRLFHL
DRGKLARIILDNTESLSCDISPSEIYSVFKARWETPGHFNGLGDFEIKGKANNKAFRD
FITAKEIEKNVREMSKGSAPGPDGIALGD IKKMDPGYSRTAELFNLWLTAGDIPDMV
RGCRTVLIPKSTTPERLKDINNWRPITIGSILLRLFSRIITARMTKACPLNPRQRFIS
APGCSENLLKLLQSIIRTAKNEHKPLGVIFVDIAKAFDTVSHQHIIHVLQQR RVDPHIV
GLVNNMYKDISTYVTTKKNTHTDKIQIRVGVKQGDPLSPLLFN LAMDPLLCKLEESG
KGFHRGQSSITAMAFADDLVLLSDSWENMKENIKILETFCNL TGLKTQGQKCHGYI
KPTKDSYTINNCPAWTINGTPLNMINPGESEKYLGLQIDPWTGVAKYDLSTKLKIWL
ESIDRAPLKPLQLDILKTYTIPRLTYLADHSEM KAGALEALDQQIRTAVKDWLHLP
SCTCDAILYVSTRDGG LGVTKLAGLIPSVQARRLHRIAQSPDET MKDFLEKAQMEKM
YEKLWVQAGGKKKGMP SIWEALPMTVPPTNTGNLSEWEAPNPKSKYPKPCDWRRK
ELKKWTKLESQGRGVKNFRNDTISNDWIQYYRRIPHRKLLTAIQLRANVYPTREFLA
RGRGDNYVKFCRHCEADLET CGHIIGFCPVTKDARIKRHN RICDRLCEEAAKREWVV
FKEPHLRDATTELFKPDVIFVKEDRALVVDVTVRYESA KTTLEAAAMEKV DKKYKHLE
AEVKELTNAKD VVFMGFPLGARGKFYKGNFN LLETGLPKTRQLSVAKTLSTYALMS
SVDIVHMFASRSRKPNV**

10

20

21. R2 ZoAlトランスフェクトしたmRNA配列:

AGGAAUAAACUAGUAUUCUUCUGGUCCCCACAGACUCAGAGAGAACCCGCCACCA
UGAACAUUCGUGAAGGUGACCGUGCCCGACAAGAUCCCCCUUGUCCUUGCUGUGG
CGUGCGCCUGAACAGCGUGCUGGCACUGAUCGAGCACCUGAAGGGC UCCCACGGA
AGGAGAAGGGUGUGCUUCCGCUGUGCCAAGUGCGGGCCGGGAGAACUUUAUACACC
ACAGCACCGUGUGUCACUACGCCAAGUGCAAGGGGCCACAGAU CGAGAGACCACC
CGUGGGCGAGUGGAUCUGUGAGGUGUGCGGCAGGGAUUUCACCACAAAGAU CGGC
CUGGGCCAGCACAAAGCGCCACAUGCACGCCAUGGUGCGCAACCAGGAGCGGAUCG
ACGCCAGCCAGCCCAAGGAGACCUCCAAUUCGCGGGCGCCACAAGCGGUGCUGGACA
AAGGAGGAGGAGGAGCUGCUGAUGAAGCUGGAGGUGCAGUUUGAGAACCACAAGA
ACAUCAAUAAGCUGAUCGCCGAGCAGCUGACCACAAAGACCGCCAAGCAGAUUC
CGAUAAAGCGCCGGAUCUGCUGAAGAAGGGCAGAGGCACCACAGGCAACCUGGAG
ACAGAGCCUGGCAUGAGCCACCAGUCCCAGGCCAAGGUGAAGGACAAUGGACUGG
GAGGCGAUCACCUGCCAGGAGGACCUGUGGUGGACAAGGGCACCAUCGGCAAGCC
AGGCCAGCACCUGGACACCGAUAAUUCUACACAGAUACAGCAGGCAAGAAGAAG
GGAGGAGGACUGCAGGCCAGAUACAGAAGGCGCAUCAUGAAGAGGCUGGCCGCCG
GCACCAUCAACAUCUCCCCAAAGGUGUUUAAGGAGCUGAUAACGACCAGGAGGC
CAGGCCCCUGAUAUAUCAGACCACAGAGGACUGUUUCGGCCUGCUGGAUUCCGCC
UGCCAGAU CAGAACC GCCUGAGGGAGAAGGGCAAGUCUCAGGAGGAGCGGCCCA
GAAAGCAGUACCAGAAGUGGAUGAAGAAGCGCGCCAUCAAGCGGGGCGACUAUCU
GCGGUUCCAGAGACUGUUUCACCUGGAUAGAGGCAAGCUGGCCAGGAUCAUCCUG
GACAACACCGAGUCCUGUCUUGUGAUUAUCAGCCCUCCGAGAU CUACAGCGUGU
UCAAGGCCCGCUGGGAGACACCAGGCCACUUCAAUGGCCUGGGCGAUUUUGAGAU
CAAGGGCAAGGCCAACAAUAAGGCCUUCGGGACUUUAUCACCGCCAAGGAGAU C
GAGAAGAACGUGAGAGAGAUGUCUAAGGGCAGCGCACCAAGGACCUGAUGGAUCG

30

40

50

CCCUGGGCGACAUCAAGAAGAUGGACCCCGGCUAUUCUCGGACCGCCGAGCUGUU
 UAAUCUGUGGCUGACAGCCGGCGACAUCCAGAUUAGGUGAGGGGAUGCAGGACC
 GUGCUGAUCCCAAAGAGCACCACACCCGAGCGCCUGAAGGACAUCAACAAUUGGC
 GGCCCAUCACAAUCGGCUCCAUCUGCUGAGACUGUUCUCUAGGAUCAUACCGC
 CAGAAUGACAAAGGCAUGUCCACUGAACCCTAAGGCAGAGGGGGCUUUAUCUCCGCA
 CCAGGCUGCUCUGAGAACCUGAAGCUGCUGCAGAGCAUCAUCAGGACCGCCAAGA
 AUGAGCACAAAGCCCCUGGGCGUGAUCUUCGUGGACAUCGCCAAGGCCUUGAUAC
 AGUGUCCCACCAGCACAUCAUCCACGUGCUGCAGCAGCGGAGAGUGGACCCCCAC
 AUCGUGGGGCCUGGUGAACAUAUGUACAAGGACAUCUCCACCUAUGUGACCACAA
 AGAAGAACACCCACACAGACAAGAUAUCCAGAUCAAGAGUGGGCGUGAAGCAGGGCGA 10
 UCCUCUGUCUCCACUGCUGUUAUCUGGGCCAUGGACCCACUGCUGUGUAAGCUG
 GAGGAGAGCGGCAAGGGCUUCCACAGGGGCCAGAGCUCCAUCACCGCCAUGGCCU
 UUGCCGACGAUCUGGUGCUGCUGUCUGAUAGCUGGGAGAACAUAGAAGGAGAAUUA
 CAAGAUCUGGAGACAUCUGUAUUCUGACCGGCCUAAAGACACAGGGGCCAGAAG
 UGCCACGGCUUUUACAUCAAGCCUACCAAGGACUCCUAUACAUAACAUAUUGCC
 CAGCCUGGACCAUCAACGGCACACCUCUGAACAUGAUAUCCAGGGCGAGUCUGA
 GAAGUACCUGGGACUGCAGAUCCGACCCAUGGACCGGGCGUGGCCAAGUAUGAUCUG
 AGCACAAAGCUGAAGAUAUCUGGCUGGAGUCCAUCGACAGGGCCCCCUGAAGCCUC
 UGCAGAAGCUGGAUAUCCUGAAAACCUACACAAUCCCCCGCCUGACCUAUCUGGC
 CGAUCACAGCGAGAUGAAGGCAGGAGCACUGGAGGCACUGGACCAGCAGAUCAAG 20
 ACCGCCGUGAAGGAUUGGCUGCACCUGCCUUCUGCACAUGUGACGCCAUCCUGU
 ACGUGUCUACCAGGGAUGGAGGACUGGGAGUGACAAAGCUGGCAGGACUGAUCCC
 AUCUGUGCAGGCAAGGAGGCUGCACAGGAUCGCACAGAGCCUGACGAGACCAUG
 AAGGAUUUCCUGGAGAAGGCCCCAGAUUGGAGAAGAUGUACGAGAAGCUGUGGGUGC
 AGGCCGGCGGCAAGAAGAAGGGCAUGCCAUCUUGGGAGGCCCCUGCCCAUGAC
 AGUGCCUCCAACCAACACAGGCAAUUCUGUCUGAGUGGGAGGCCCCAAAUCCCAAG
 AGCAAGUAUCCUAAGCCAUGCGACUGGCGGAGAAAGGAGCUGAAGAAGUGGACCA
 AGCUGGAGAGCCAGGGCAGAGGCGUGAAGAACUUUAGGAUUGACACAAUCCAA
 CGAUUGGAUUCAGUACUAUAGGCGCAUCCCAACACCGCAAGCUGCUGACCGCCAUC
 CAGCUGCGCGCCAACGUGUACCCAACAAGGGAGUUCUGGGCACGGGGCAGAGGCG 30
 AUAAUUAUGUGAAGUUUUGUCGGCACUGCGAGGCCGACCUGGAGACCUGUGGCCA
 CAUCAUCGGCUUCUGCCCCGUGACAAAGGACGCCCGGAUCAAGAGACACAACAGG
 AUCUGUGAUAGACUGUGCGAGGAGGCAGCAAAGAGGGAGUGGGUGGUUUAAGG
 AGCCCCACCUGAGGGACGCCACCACAGAGCUGUUAAGCCUGAUGUGAUCUUUGU
 GAAGGAGGACAGGGCACUGGUGGUGGAUGUGACCGUGCGGUACGAGUCCGCCAAG
 ACCACACUGGAGGCUGCCGCCAUGGAGAAGGUGGAUAAGUAUAAGCACCUGGAGG
 CCGAGGUGAAGGAGCUGACAAAUGCCAAGGACGUGGUGUUAUGGGAUUUCCACU
 GGGAGCAAGGGGCAAGUUCUACAAGGGCAACUUUAUUCUGCUGGAGACCCUGGGC
 CUGCCUAAGACAAGACAGCUGUCUGUGGCCAAGACCCUGAGCACCUACGCCUGA
 UGUCUAGCGUGGACAUCGUGCACAUGUUCGCCUCCAGGUCUCGCAAGCCUAACGU 40
 GUGAUGACCUCGAGCUGGUACUGCAUGCACGCAUAGCUAGCUGCCCCUUCCCGU
 CCUGGGUACCCCGAGUCUCCCCGACCUCGGGUCCCAGGUUAGCUCCCACCUCCA
 CCUGCCCCACUCACCACCUCUGCUAGUUCAGACACCUCCCAAGCACGCAGCAAU
 GCAGCUCAAAACGCUUAGCCUAGCCACACCCCCACGGGAAACAGCAGUGAUUAAC
 CUUUAGCAAUAAACGAAAGUUUAACUAAGCUAUACUAACCCCAAGGGUUGGUCAAU
 UUCGUGCCAGCCACACCCUGGAGCUAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
 AAAGCAUAUGACUAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

22. R2A ZoAl_catdead アミノ酸配列:

MNIVKVTVPDKNPPCPCCGVRLNSVLALIEHLKGS HGRRRVCFRCAKCGRENFNHH 50

STVCHYAKCKGPQIERPPVGEWICEVCGRDFTTKIGLGQHKRHHAMVRNQERIDA
SQPKETSNRGAHKRCWTKEEEELLMKLEVQFENHKNINKLIAEQLTTKTAKQISDKR
RMLLKKGRGTTGNLETEPGMSHQSAKVKDNLGGDHLPGGPVVDKGTIGKPGQH
LDTDNSHQITAGKKKGGGLQARYRRRIMKRLAAGTINIFPKVFKELINDQEARPLINQ
TTEDCFGLLD SACQIRTA LREKGSQEERPRKQYQKWMKKRAIKRGDYLRFRQLFHL
DRGKLARIILDNTESLSCDISPSEIYSVFKARWETPGHFNGLGDFEIKGKANNKAFRD
FITAKEIEKNVREMSKGSAPGPDGIALGDIKKMDPGYSRTAELFNLWLTAGDIPDMV
RGCRTVLIPKSTTPERLKDINNWRPITIGSILLRLFSRIITARMTKACPLNPRQGRFIS
APGCSENLLKLLQSIIRTAKNEHKPLGVIFVDIAKAFTVSHQHIIHVLQQRVDPHIV
GLVNNMYKDISTYVTTKKNTHTDKIQIRVGVKQGDPLSPLLFNLA MDPLLCKLEESG
KGFHRGQSSITAMAFAAALVLLSDSWENMKENIKILETFCNLTGLKTQGGQKCHGFYI
KPTKDSYTINNCPAWTINGTPLNMINPGESEKYLGLQIDPWTGVAKYDLSTKLKIWL
ESIDRAPLKPLQKLDILKTYTIPRLTYLADHSEM KAGALEALDQQIRTAVKDWLHLP
SCTCDAILYVSTRDGGGLGVTKLAGLIPSVQARRLHRIAQSPDET MKDFLEKAQMEKM
YEKLWVQAGGKKKGMP SIWEALPMTVPPTNTGNLSEWEAPNPKSKYPKPCDWRRK
ELKKWTKLESQGRGVKNFRNDTISNDWIQYYRRIPHRKLLTAIQLRANVYPTREFLA
RGRGDNYVKFCRHCEADLET CGHIIGFCPVT KDARIKRHN RCDRLCEEAAKREWVV
FKEPHLRDATT ELFKPDVIFVKEDRALVVDVTVRYESA KTTLEAAAMEKV D KYKHLE
AEVKELTNAKD VVFMGFPLGARGKFYKGNFNLLET LGLPKTRQLSVAKTLSTYALMS
SVDIVHMFASRSRKPNV**

10

20

23. R2 ZoAl_catdeadトランスフェクトしたmRNA配列:

AGGAAUAAACUAGUAUUCUUCUGGUCCCCACAGACUCAGAGAGAACCCGCCACCA
UGAACAUUCGUGAAGGUGACCGUGCCCGACAAGAAUCCCCCUUGUCCUUGCUGUGG
CGUGCGCCUGAACAGCGUGCUGGCACUGAUCGAGCACCUGAAGGGCUGCCACGGA
AGGAGAAGGGUGUGCUUCCGCUGUGCCAAGUGCGGCCGGGAGAACUUUAUACACC
ACAGCACCGUGUGUCACUACGCCAAGUGCAAGGGCCACAGAU CGAGAGACCACC
CGUGGGCGAGUGGAUCUGUGAGGUGUGCGGCAGGGAUUUCACCACAAAGAU CGGC
CUGGGCCAGCACAAAGCGCCACAUGCACGCCAUGGUGCGCAACCAGGAGCGGAUCG
ACGCCAGCCAGCCCAAGGAGACCUCCAAU CGCGGCCGCCACAAGCGGUGCUGGACA
AAGGAGGAGGAGGAGCUGCUGAUGAAGCUGGAGGUGCAGUUUGAGAACCACAAGA
ACAUCAAUAAGCUGAUCGCCGAGCAGCUGACCACAAAGACCGCCAAGCAGAUUC
CGAU AAGCGCCGGAUCUGCUGAAGAAGGGCAGAGGCACCACAGGCAACCUGGAG
ACAGAGCCUGGCAUGAGCCACCAGUCCAGGCCAAGGUGAAGGACAAUGGACUGG
GAGGCGAUCACCUGCCAGGAGGACCUGUGGUGGACAAGGGCACCAUCGGCAAGCC
AGGCCAGCACCUUGGACACCGAUAAUUCUACACAGAUACAGCAGGCAAGAAGAAG
GGAGGAGGACUGCAGGCCAGAUACAGAAGGCGCAUCAUGAAGAGGCUGGCCGCCG
GCACCAUCAACAUCU UCCCAAAGGUGUUUAAGGAGCUGAUCAACGACCAGGAGGC
CAGGCCCCUGAUCAAUCAGACCACAGAGGACUGUUUCGGCCUGCUGGAUUCCGCC
UGCCAGAU CAGAACC GCCUGAGGGAGAAGGGCAAGUCUCAGGAGGAGCGGCCCA
GAAAGCAGUACCAGAAGUGGAUGAAGAAGCGCGCCAUCAAGCGGGGCGACUAUCU
GCGGUUCCAGAGACUGUUUCACCUGGAUAGAGGCAAGCUGGCCAGGAUCAUCCUG
GACAACACCGAGUCCUGUCUUGUGAUUACAGCCCUCCGAGAUCAACAGCGUGU
UCAAGGCCCGCUGGGAGACACCAGGCCACUUC AAUGGCCUGGGCGAUUUUGAGAU
CAAGGGCAAGGCCAACAAUAAGGCCUUCGGGACUUUAUCACCGCCAAGGAGAU C
GAGAAGAACGUGAGAGAGAUGUCUAAGGGCAGCGCACCAGGACCUGAUGGAUCG
CCCUGGGCGACAUCAAGAAGAUGGACCCCGGCUAUUCUCGGACCGCCGAGCUGUU
UAAUCUGUGGCUGACAGCCGGCGACAUC CAGAU AUGGUGAGGGGAUGCAGGACC
GUGCUGAUCCCAAAGAGCACACACCCGAGCGCCUGAAGGACAUCAACAAUUGGC
GGCCCAUCACAAUCGGCUCCAUCUGCUGAGACUGUUCUCUAGGAUCAUACCGC
CAGAAUGACAAAGGCAUGUCCACUGAACC CAAGGCAGAGGGGCUUUAUCUCCGCA

30

40

50

CCAGGCUGCUCUGAGAACCUGAAGCUGCUGCAGAGCAUCAUCAGGACCGCCAAGA
AUGAGCACAAGCCCCUGGGCGUGAUCUUCGUGGACAUCGCCAAGGCCUUUGAUAC
AGUGUCCCACCAGCACAUCAUCCACGUGCUGCAGCAGCGGAGAGUGGACCCCCAC
AUCGUGGGGCCUGGUGAACAUAUGUACAAGGACAUCUCCACCUAUGUGACCACAA
AGAAGAACACCCACACAGACAAGAUAUCCAGAUUCAGAGUGGGCGUGAAGCAGGGGCGA
UCCUCUGUCUCCACUGCUGUUAUCUGGGCCAUGGACCCACUGCUGUGUAAGCUG
GAGGAGAGCGGCAAGGGCUUCCACAGGGGCCAGAGCUCCAUCACCGCCAUGGCCU
UUGCCGCCGCACUGGUGCUGCUGCUGAUAAGCUGGGAGAACAUGAAGGAGAAUAU
CAAGAUAUCCUGGAGACAUAUCUGUAAUCUGACCGGCCUAAAGACACAGGGGCCAGAAG
UGCCACGGCUUUUACAUAAGCCUACCAAGGACUCCUAUAACAUAACAUAUUGCC
CAGCCUGGACCAUCAACGGCACACCUCUGAACAUGAUAUCCAGGGCGAGUCUGA
GAAGUACCUGGGACUGCAGAUCCGACCCAUGGACCGGGCGUGGGCCAAGUAUGAUCUG
AGCACAAAGCUGAAGAUAUCUGGCUGGAGUCCAUCGACAGGGCCCCCUGAAGCCUC
UGCAGAAGCUGGAUAUCCUGAAAACCUACACAAUCCCCCGCCUGACCUAUCUGGC
CGAUCACAGCGAGAUGAAGGCAGGAGCACUGGAGGCACUGGACCAGCAGAUCAGG
ACCGCCGUGAAGGAUUGGCUGCACCUGCCUUCUGCACAUGUGACGCCAUCCUGU
ACGUGUCUACCAGGGAUGGAGGACUGGGAGUGACAAAGCUGGCAGGACUGAUCCC
AUCUGUGCAGGCAAGGAGGCUGCACAGGAUCGCACAGAGCCUGACGAGACCAUG
AAGGAUUUCCUGGAGAAGGCCCCAGAUGGAGAAGAUGUACGAGAAGCUGUGGGUGC
AGGCCGGCGGCAAGAAGAAGGGCAUGCCAUAUCCUGGGAGGCCCCUGCCCAUGAC
AGUGCCUCCAACCAACACAGGCAAUUCUGUCUGAGUGGGAGGCCCCAAAUCCCAAG
AGCAAGUAUCCUAAGCCAUGCGACUGGCGGAGAAAGGAGCUGAAGAAGUGGACCA
AGCUGGAGAGCCAGGGCAGAGGCGUGAAGAACUUUAGGAUUGACACAAUCUCCAA
CGAUUGGAUUCAGUACUAUAGGCGCAUCCCAACACCGCAAGCUGCUGACCGCCAUC
CAGCUGCGCGCCAACGUGUACCCAACAAGGGAGUUCUGGCACGGGGCAGAGGCG
AUAAUUAUGUGAAGUUUUGUCGGCACUGCGAGGCCGACCUGGAGACCUGUGGCCA
CAUCAUCGGCUUCUGCCCCGUGACAAAGGACGCCCGGAUCAAGAGACACAACAGG
AUCUGUGAUAGACUGUGCGAGGAGGCAGCAAAGAGGGGAGUGGGUGGUUUAAAG
AGCCCCACCUGAGGGACGCCACCACAGAGCUGUUAAGCCUGAUGUGAUCUUUGU
GAAGGAGGACAGGGCACUGGUGGUGGAUGUGACCGUGCGGUACGAGUCCGCCAAG
ACCACACUGGAGGCUGCCGCCAUGGAGAAGGUGGAUAAGUAUAAGCACCUGGAGG
CCGAGGUGAAGGAGCUGACAAAUGCCAAGGACGUGGUGUUAUGGGAUUUCCACU
GGGAGCAAGGGGCAAGUUCUACAAGGGCAACUUUAUUCUGCUGGAGACCCUGGGC
CUGCCUAAGACAAGACAGCUGUCUGUGGCCAAGACCCUGAGCACCUCACGCCUGA
UGUCUAGCGUGGACAUCGUGCACAUGUUCGCCUCCAGGUCUCGCAAGCCUAACGU
GUGAUGACCUCGAGCUGGUACUGCAUGCACGCAAUGCUAGCUGCCCCUUCGCCU
CCUGGGUACCCCGAGUCUCCCCGACCUCGGGUCCCAGGUUAGCUCCCACCUGCA
CCUGCCCCACUCACCACCUCUGCUAGUUCAGACACCUCCCAAGCACGCAGCAAU
GCAGCUCAAAACGCUUAGCCUAGCCACACCCCCACGGGAAACAGCAGUGAUUAAC
CUUUAGCAAUAAACGAAAGUUUAACUAAGCUAUACUAACCCCAGGGUUGGUCAAU
UUCGUGCCAGCCACACCCUGGAGCUAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAGCAUAUGACUAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

24. R2A TiGu N-DYK プラスミドORF配列:

ATGGATTACAAGGATGACGACGATAAGGGTACCATGGTGGGCGGAAGCCCTCCTGG
CGAGGGCGTTAGAAGCCCTGGAATCGCCTCTCCTAGCGGTACAGCCGTGCAGCACG
ACTTCGGCTCTCCAATCCTGGTGCCTGGCGCTGAAGCTGCCGAAGTTTCTACACCT
GTGGTGAAAGTGCCCCAGGACCACCCAGCCTGTCCATGCTGCGGCACACGGGTGGT
GAAGGTGACAGCCCTGTCTGAGCACCTCCGGCGCGCCACGGGAGAAAGCGGGTGC
TGTTCCAGTGCTCTCGGTGCGGCAGAATGAACGAGAAGCACCATTTCTATCGCCTGC

CACTTCCCCAAGTGCCGGGGACCTCCTGTGGAGGAAGGCCCCCTGGGCGCTCCTGA
ATGGTGCTGTGAAGAGTGCGGACAGAAGTTCAACACAAAGTCCGGCCTGAGCCAGC
ATAAGCGCAGCGTGCACCCCCTGACCAGAAATGTGGAAAGAATTGAAGCCGCCCGC
CCCAAGGGTAAGGGCAAAAGAGGAGCCCATAAAGGGCTGTTGGACAGAGGCCGAGGT
GGCACAGCTCATCGAACTGGAGGGCAGATTCAAGAACCAGAGATTTCATCAACAAGC
TGATTGCGGAGCACCTGCCTTCTAAATCCGCCAAGCAGATCAGCGACAAGAGACGG
CAGCTGGCCGCTGCTACCAAGACCTCTAGCCCCGAGAAGAGGGGTGACCAGCAGCAC
AAGCGGAGAGAGCAGTCCTGAGGTGGAGAAAGTGGAGGGAATAAAACGGGAGTACC
GTAGACGGGTGGGCGAGTGGCTGTGCGCCGGTAGTCTGAACGACCAGACCAGTTTT
CAGAAAATCCTGGAAGATGTGGAGAGCGGCAGCGAGATCGTGACCGGGGCCACTGGA
GGAAGTGGCCAGCTTCGCCAGGGGGCAAGCTGGCCGCCGCTAGAGTGCGGCACCACC
GGAAGCATCCCGCCGAGGGCCGTGCCTGCGAGAGAGGAGCAGAGATGGATGAAGCGC
AGAGTCGGACGACGGGGACTGTACCTGAGATTCCAGAGACTGTTTCGCCCTGGACCG
GAGAAAGCTAGCCGGCATCATCCTGGACGATGTGGAATCTATCAAGTGCCCTCTGC
CTATGGAAGAAGTGGCCGACGTGTTCCGGAGGAGGTGGGAAGAAGTGGCCCTTTTC
ACCGGGTCAGGCAGCTTCAGGTCTCTGGGCAAGGCCGACAACGGCGCCTTCAAGCC
TATGATCAGCGCCAAGGAAGTTATGAAGAACGTGCGGGAGATGAGCCGACGGAGCG
CTCCTGGACCTGATGGCCTGTCCCTGAGGGATCTCATGAAGATCGATCCTCAGGGC
TCCAGAATGGCCGAACCTGTTCAATCTGTGGCTGCTGGCTGGAAGAGTGGCAGACCA
GGTGAAGGCCGGCAGAACCGTCCTGATTCTAAGAGCGCCGATCCTGGGAAGATCG
GCAATATCGACAACCTGGCGGCCTATCACCATCGGAAGCGTGCTGCTTAGAATGTTC
AGCAGAATCCTGAGTGCCAGACTGCGGAGAGCCTGTCCTATCAACAGACGGCAGCG
GGGCTTTATCGCCGCCCCCGGCTGCAGCGAGAACCTGAAGTTACTGCAAGCCCTCA
TTAAGAGCGCCAAGCGGGACCACAGAACACTGGGCGTGGTGTTTCGTGGACCTGGCG
AAGGCCTTCGACAGCGTGAAACCACCAGCACATCTTCCAAGTGCTGGTACAGAAGGG
GGTGGATGGCCACATCATCGACATCCTGAGAGATCTTTACACCAACGCCGGCACCT
ACCTGGAATCCGGCTCTCAGAGAAGCGGCTTTATCAAGATCCTGAGAGGGCGTCAAG
CAGGGCGACCCACTGAGCCCAATCCTGTTTAACCTGGCCCTGGACCCCTGCTATG
TAGACTGGAGGACCGGGGCTGGGATATAAGTACGGCGATCAGCAAATCACCAGCC
TGGCATTTCGCCGATGATCTGGCCCTACTGAGCGACAGCTGGGAGGGCATGCAGCAA
AGCATCCGGGTGCTCGAAGAGTTTTGCCAGAGGACAGGCCTGCGAGTGCAAGCTCC
AAAGTGCCACGGCTTCCTGATCAGACCTACCAAAGAATCTTACACCATTAACTGACT
GCGACCCCTTGGACAATCGCCGACATGCAGCTGGATATGATCGACCCAGGCTCCAGC
GAAAAATACCTGGGCCTGGGCATCGACCCTTGGATCGGCCTGAGCCGTCCTGAGCT
GTCCGAAGTGCTGACCAGATGGGTGAAGAACATCGGAGGGCGCCCCCTGAAGCCCC
TGCAAAAGGTGGACATCCTGAGAAGCTACGCCCTGCCTAGACTGCTGTTTATCGCC
GACCACGCCGGCCTGTCCGCCACCTGCCTGCACTCTCTGGATCTTTCCATCAGAAG
CGCTGTGAAGGGCTGGCTGCACCTGCCTCCCAGCACCTGTGATGCCATCATCTACG
TGAGCTACAAGGACGGCGGACTGGGCCTGCCAGACTGGCAAGCCTGATCCCTAAC
GTGCAGGCAAGACGGCTGGTGAGAATCGCCCAAGCGAGGATGACGTGATCCGCAG
CGTGGTGCTGCAGGAGGGAATCCAGGAGGAGATCAGAAAAGTGTGGATCTCTGCCG
GCGGCAGACCTGAGAAGGTCCCTAGCGTGACCGGCGAATTCCCCGTGATGGAAGCC
CAGGCCGCCGACGAGGCACTGTCCGAGTGGGAGAGACGGGGCCCTCGCACCATCTA
CCCCATCCCTTGTAATGGAGAAAGCGCGAGATGGAAAACCTGGACGAATCTCAAAT
CTCAGGGCCACGGCATTAGAAATTTGAGAAATGACAGAATCAGCAACGACTGGCTG
CTGCACTACGGGAGAATCCCTCACAGAAAGCTGATCACAGCTATCCAGCTGAGAGC
CAACGTGTATCCTACAAGAGAGTTCTGGCTAGAGGCCTTGGCGAGGGCGCTCCTA
GAGGCTGCAGACACTGTCCCGCCGAATGGGAATCCTGCAGCCACATCATCGGCTAC
TGCCCTGCCGTGCAGGAGGCCAGAATCAAGAGACACAACGACATCTGCGGGCTGCT
TGCTGAAGAGGCCAGAAAGCTGGGATGGGTGATCTTCATCGAGCCTCACTTGAGAG

10

20

30

40

50

ATAATACCAACGAGCTGTTCAAACCCGACCTGGTGGTGGTGAAGGATCTTGCGCC
AAGGTGGTGGACGTGACCATCAGATACGAGAGCGGGCTAACCACCCTGTCAGATGC
TGCCGCTGAGAAAGCACGAAAATACCAGCACCTGGCGGGCGAGGTGAGAGCCCTGA
CTAGCGCCACAACCGTGGACTTTCTGGGCTTCCCCATCGGTGCCCCGGGGCAAATGG
TACGTTGGAAACAACGGCCTCCTGAGCGACCTGGGCTTCAGCACCCAGTAGAGTTGT
TCGGATCGCCCCGGGCCCTGAGCAAGAAGGCCCTGCTGAGCTCTGTGGACATAATCC
ACATCTTTGCTAGCAGAGCCAGACAGGCCCAGACCAGCGAATAG

25. R2A TiGu N-DYKアミノ酸配列:

MDYKDDDDKGTMVGGSPPGEGVRSPGIASPSGTAVQHDFGSPILVPGAEEAEVSTPV
VKVPQDHPACPCCGTRVVKVTALSEHLRRAHGRKRVLFQCSRCGRMNEKHHSIACH
FPKCRGPPVEEGPLGAPEWCCEECGQKFNTKSGLSQHKRSVHPLTRNVERIEAARPK
GKGKRGAKHGCWTEAEVAQLIELEGRFKNQRFINKLIAEHLPSKSAKQISDKRRQLA
AATKTSSPEKRVTSSTSGESSPEVEKVEGIKREYRRRVGEWLCAGSLNDQTSFQKILE
DVESGSEIVTGPLEELASFARGKLAARVRHHRKHPAEAVPAREEQRW MKRRVGRR
GLYLRFQRLFALDRRKLAGIILDDVESIKCPLPMEEVADVFRRRWEEVAPFTGSGSFR
SLGKADNGAFKPMISAKEVMKNVREMSRRSAPGPDGLSLRDL MKIDPQGSMAELF
NLWLLAGRVPDQVKAGRTVLIPKSADPGKIGNIDNWRPITIGSVLLRMFSRILSARLR
RACPINRRQRGFIAAPGCSENLKLLQALIKSAKRDHRTLGVVFVDLAKAFDSVNHQH
IFQVLVQKGVGDGHIIDILRDLYTNAGTYLESQSQRSGFIKILRGVKQGDPLSPILFNLA
LDPLLCRLEDRLGYKYGDQQITSLAFADDLALLSDSWEGMQQSIRVVEEFCQRTGL
RVQAPKCHGFLIRPTKESYTINDCDPWTIADMQLDMIDPGSSEKYLGLGIDPWIGLS
RPELSEVLTRWVKNIGGAPLKPLQKVDILRSYALPRLFIADHAGLSATCLHSLDLSI
RSAVKGWLHLPPSTCDAIIVSYKDGGLGLPRLASLIPNVQARRLVRIAQSEDDVIRS
VVLQEGIQEEIRKVVWISAGGRPEKVPSTGEFPVMEAQAADSEWERRAPRTIYP
IPCKWRKREMENWTNLKSQGHGIRNFENDRISNDWLLHYGRIPHRKLITAIQLRAN
VYPTREFLARGLGEGAPRGCRHCPAEWESC SHIIGYCPAVQEARIKRHNDICGVLAE
EARKLGWVIFIEPHLRDNTNELFKPDLVLVKGSCAKVVDVTIRYESGLTTLSDAAAE
KARKYQHLAGEVRALTSATTVDLFGFPIGARGKWYVGNNGLLSDLGFTSRVVRIAR
ALS K KALLSSVDIIHIFASRARQAQTSE*

10

20

30

26. R2A TaGu N-DYKプラスミドORF配列:

ATGGATTACAAGGATGACGACGATAAGGGTACCGAGAAGGTGATGGTCACCGTTCC
GGACAAGAACCCTCCATGTCCATGCTGCGGCACCCGGGTGAACAGCGTGCTGAACC
TTATCGAGCACCTGAAGGTCAGCCACGGAAAGCGGGGCGTGTGCTTCCGGTGCGCG
AAGTGTGGGAAGGAAAATAGCAACTACCACAGCGTGGTGTGCCATTTCCCAAATG
TAGAGGACCTGAGACAGAGAAAGCTCCAGCCGGCGAGTGGATTTGTGAAGTGTGCA
ACAGAGACTTCACCACCAAATCGGCCTGGGCCAGCATAAGAGATTAGCTCACCCC
GCTGTTTCGGAATCAGGAGAGAATCGTCGCCAGCCAGCCTAAGGAACTAGCAACAG
AGGCGCTCACAAAAGATGCTGGACCAAAGAGGAGGAGGAAGTCTGATCAGACTGG
AAGCCCAGTTTGAGGGCAACAAGAACATCAACAAGCTGATCGCAGAACACATCACC
ACAAAGACCGCCAAGCAGATCTCCGATAAGAGGAGACTGCTGAGCCGCAAGCCTGC
CGAGGAGCCTAGAGAGGAACCTGGCACATGCCACCACACCAGACGGGCAGCTGCCA
GCCTGAGAACAGAGCCCGAGATGTCCACACGCTCAGGCCGAAGATCGGGATAAT
GGCCCTGGCCGGAGACCCCTGCCTGGTAGAGCCGCCGCCGGCGGCGAGAACAATGGA
CGAGATCAGGAGACACCCTGACAAGGGCAATGGACAGCAGCGGCCTACCAAACAGA
AGTCTGAGGAGCAGCTGCAAGCCTACTACAAAAGACACTGGAGGAGCGGCTGAGC
GCCGGAGCCCTGAATACCTTCCCCGCGCCTTCAAGCAAGTGATGGAAGGCAGAGA
TATCAAGCTGGTGATCAACCAGACAGCCCAGGACTGCTTCGGCTGTCTGGAGAGCA
TCTCTCAGATCAGAACAGCGACCAGAGATAAGAAAGATACCGTGACCAGAGAGAAG
CACCCGAAGAAGCCATTCCAAAAGTGATGAAGGACCGGGCCATAAAGAAGGGCAA
CTACCTGAGGTTCCAGAGATTGTTCTACCTGGACCGAGGCAAACTGGCCAAGATCA

40

50

TCCTGGACGATATCGAATGCCTGAGCTGCGACATCCCCCTGTCTGAGATCTACAGC
GTGTTTAAGACAAGATGGGAAACCACAGGAAGCTTCAAATCTCTGGGCGATTTTAA
AACCTACGGCAAGGCCGACAACACTGCCTTCAGAGAACTGATTACAGCTAAGGAAA
TCGAGAAAAACGTGCAGGAGATGAGCAAGGGCAGCGCCCCCTGGCCCTGACGGTATC
ACCCTGGGAGATGTGGTGAAAATGGACCCCGAGTTCAGCCGCACAATGGAGATCTT
CAACCTGTGGCTGACCACTGGCAAAATCCCTGACATGGTGCGGGGATGCCGCACCG
TGCTGATACCTAAGAGCAGCAAGCCAGATAGACTGAAGGATATCAACAACTGGCGG
CCTATCACAATCGGCAGCATCCTGCTCCGGCTGTTTAGCAGAATCGTAACCGCCCCG
TCTGAGTAAGGCCTGTCTCTGAACCCAGACAGCGGGGCTTCATCCGGGGCCGCCG
GCTGCAGCGAGAACCTCAAACCTGCTGCAGACCATCATCTGGTTCGGCCAAACGAGAG
CACAGACCTCTCGGCGTGGTGTTCGTGGACATCGCCAAAGCTTTTGATACCGTCTC
CCACCAGCACATCATCCACGCCCTGCAGCAGCGGGAAGTGGACCCTCACATCGTGG
GACTGGTGTCCAACATGTACGAGAACATCAGCACTTATATCACCACCAAGAGAAAC
ACCCACACAGACAAGATCCAGATCAGAGTGGGCGTTAAGCAGGGCGATCCTATGTC
TCCTCTGCTGTTCAACCTGGCTATGGACCCTCTGCTCTGTAAACTGGAGGAGTCTG
GCAAGGGATACCACAGAGGCCAAAGCTCCATCACCCTATGGCCTTTGCCGACGAC
CTGGTGCTGCTGTCTGATTCTTGGGAGAACATGAATACCAACATAAGCATCCTGGA
AACATTCTGCAACCTTACAGGCCTGAAAACCCAGGGACAAAAGTGCCACGGCTTTT
ACATCAAGCCTACCAAAGACAGCTACACCATCAACGACTGCGCCGCCTGGACCATC
AACGGCACCCCTTTAAACATGATCGACCCTGGAGAAAGCGAGAAGTATCTGGGCCT
GCAGTTCGACCCCTGGATCGGCATCGCTAGAAGCGGCCTGAGCACAAAGCTGGACT
TTTGGCTGCAGAGAATTGACCAGGCCCTCTGAAGCCTCTGCAGAAAACCGACATC
CTGAAAACCTACACAATCCCTCGGCTGATCTACATTGCCGATCACAGCGAGGTAAA
GACCGCCCTGCTGGAAACCTGGACCAAAAGATCAGAACCGCCGTGAAGGAATGGC
TCCACCTGCCCCCTGCACCTGCGACGCTATCCTGTACAGCAGCACCAGGGACGGC
GGCCTGGGCATACCAAGCTGGCGGGCCTGATCCCCCTCCGTCCAGGCTAGACGGCT
GCATAGAATCGCCCAGAGCAGCGATGACACAATGAAGTGCTTTATGGAAAAGGAAA
AGATGGAACAGCTGCACAAGAAGCTGTGGATTACAGGCCGGTGGCGACAGAGAGAAC
ATTCTAGCATCTGGGAGGCGCCGCCTAGTAGCGAGCCTCCCAACAACGTGTCTAC
CAATTCTGAGTGGGAAGCCCCCTACACAGAAGGACAAGTTCCCTAAGCCTTGTAATT
GGAGAAAGAACGAGTTCAAGAAGTGGACAAAGCTGGCCTCTCAGGGCCGGGGAATT
GTGAATTTGAGCGGGACAAGATCAGCAATCACTGGATTACGTAACAGAGAAGT
CCCACACAGAAAGCTGCTGACGGCCCTTCAGCTGAGAGCCAACGTGTACCCACGC
GGGAGTTCCTGGCCAGAGGTAGACAGGACCAGTACATCAAGGCCTGCAGACATTGT
GATGCTGATATCGAGTCTTGCGCCACATCATCGGCAACTGCCCTGTGACACAGGA
CGCGAGAATCAAAAGACACAACACTACATCTGCGAGCTGCTGCTGGAAGAGGCCAAGA
AGAAGGACTGGGTGGTGTTCAGGAACCCACATCAGAGACAGCAATAAGGAACTC
TATAAACCTGACCTGATCTTCGTGAAGGACGCCCGGGCCCTGGTCGTGGACGTGAC
CGTGAGATACGAGGCCGCCAAGTCTAGCCTGGAGGAGGCTGCCGCCGAGAAAGTGC
GGAAGTACAAGCACCTTGAAACAGAAGTGGGGCACCTGACCAACGCCAAGGACGTC
ACATTCGTGGGCTTCCCCCTGGGCGCCAGAGGCAAATGGCACCAGGATAACTTCAA
GCTGCTGACAGAGCTGGGCCTGAGCAAGTCCCGCCAAGTGAAGATGGCCGAGACCT
TCAGCACCGTGGCCCTGTTCTCCTCTGTGGACATCGTGCACATGTTCTGCTAGCAGA
GCCAGAAAGAGCATGGTGATGTAG

10

20

30

40

50

27. R2A TaGu N-DYKアミノ酸配列:

MDYKDDDDKGTEKVMVTVPDKNPPCPCCGTRVNSVLNLIHLKVSHGKRGVCFRCA
KCGKENSNYHSVVCHFPKCRGPETEKAPAGEWICEVCNRDFTTKIGLGQHKRLAHP
AVRNQERIVASQPKETSNRGAHKRCWTKEEEELLIRLEAQFEGNKNINKLIAEHITTK
TAKQISDKRRLLSRKPAEEPREEPGTCHHTRRAAASLRTEPEMSHHAQAEDRDNGPG
RRPLPGRAAAGGRTMDEIRRHDPKGNGQQRPTKQKSEEQLQAYYKKTLEERLSAGA

LNTFPRAFKQVMEGRDIKLVINQTAQDCFGCLESISQIRTATRDKKDVTREKHPKK
PFQKWMKDRAIKKGNYLRFQRLFYLDGRGLAKIILDDIECLSCDIPLSEIYSVFKTRW
ETTGSFKSLGDFKTYGKADNTAFRELITAKEIEKNVQEMSKGSAPGPDGITLGDVVK
MDPEFSRTMEIFNLWLTGKIPDMVRGCRTVLIPKSSKPDRLKDINNWRPITIGSILL
RLFSRIVTARLSKACPLNPRQRGFIRAAGCSENKLLQTIIWSAKREHRPLGVVFDI
AKAFDTVSHQHIIHALQQREVDPHIVGLVSNMYENISTYITTKRNTHTDKIQIRVGV
KQGDPMSPLLFNLAMDPLLCKLEESGKGYHRGQSSITAMAFADDLVLLSDSWENMN
TNISILETFCNLTGLKTQGQKCHGFYIKPTKDSYTINDCAAWTINGTPLNMIDPGESE
KYLGLQFDPWIGIARSLSTKLDFWLQRIDQAPLKPLQKTDILKTYTIPRLIYIADHS
EVKTALLETLDQKIRTAVKEWLHLPPCTCDAILYSSSTRDGGGLGITKLAGLIPSVQARR
LHRIAQSSDDTMKCFMEKEKMEQLHKKLWIQAGGDRENIPSIWEAPPSSEPPNNVS
TNSEWEAPTQKDKFPKPCNWRKNEFKKWTKLASQGRGIVNFERDKISNHWIQYYRR
IPHRKLLTALQLRANVYPTREFLARGRQDQYIKACRHCDADIESCAHIIGNCPVTQD
ARIKRHNYICELLLEEAKKKDWVVFKEPHIRDSNKELYKPD LIFVKDARALVVDVTV
RYEAAKSSLEEEAAAEKVRKYKHLETEVRHLTNAKDVT FVG FPLGARGKWHQDNFKL
LTELGLSKSRQVKMAETFSTVALFSSVDIVHMFASRARKSMVM*

10

20

30

40

50

28. R2A TaGu N-DYKトランスフェクトしたmRNA配列:

AGGAUAAACUAGUAUUCUUCUGGUCCCCACAGACUCAGAGAGAACCCGCCACCA
UGGAUUACAAGGAUGACGACGAUAAGGGAUCCGAGAAGGUGAUGGUCACCGUUC
GGACAAGAACCUCUUGUCCAUGCUGCGGCACCCGGGUGAACAGCGUGCUGAAC
CUUAUCGAGCACCUGAAGGUCAGCCACGGAAAGCGGGGCGUGUGCUUCCGGUGCG
CGAAGUGUGGGAAGGAAAUAAGCAACUACCACAGCGUGGUGUGCCAUUUCCCAA
AUGUAGAGGACCUGAGACAGAGAAAGCUCCAGCCGGCGAGUGGAUUUGUGAAGUG
UGCAACAGAGACUUCACCACCAAAAUCGGCCUGGGCCAGCAUAAGAGAUUAGCUC
ACCCCGCUGUUCGGAUUCAGGAGAGAAUCGUCGCCAGCCAGCCUAAGGAAACUAG
CAACAGAGGCGCUCACAAAAGAUGCUGGACCAAAGAGGAGGAGGAACUGCUGAUC
AGACUGGAAGCCCAGUUUGAGGGCAACAAGAACAUCAACAAGCUGAUCGCAGAAC
ACAUCACCACAAAGACCGCCAAGCAGAUUCUCCGAUAAGAGGAGACUGCUGAGCCG
CAAGCCUGCCGAGGAGCCUAGAGAGGAACCUGGCACAUGCCACCACACCAGACGGG
CAGCUGCCAGCCUGAGAACAGAGCCCGAGAUGUCCCAACACGCUCAGGCCGAAGA
UCGGGAUAAUGGCCCUUGGCCGGAGACCCUGCCUGGUAGAGCCGCCGCCGGCGGC
AGAACA AUGGACGAGAU CAGGAGACACCCUGACAAGGGCAAUGGACAGCAGCGGC
CUACCAAAACAGAAGUCUGAGGAGCAGCUGCAAGCCUACUACAAAAGACACUGGA
GGAGCGGCUGAGCGCCGGAGCCUGAAUACCUUCCCCCGCGCCUUAAGCAAGUG
AUGGAAGGCAGAGAUUAUCAAAGCUGGUGAUCAACCAGACAGCCAGGACUGCUUCG
GCUGUCUGGAGAGCAUCUCUCAGAU CAGAACAGCGACCAGAGAU AAGAAAGAUAC
CGUGACCAGAGAGAAGCACCCGAAGAAGCCAUUCCAAAAGUGGAUGAAGGACCGG
GCCAUAAAGAAGGGCAACUACCUGAGGUUCCAGAGAUUGUUCUACCUGGACCGAG
GCAAACUGGCCAAGAUAUCCUGGACGAUAUCGAAUGCCUGAGCUGCGACAUCCC
CCUGUCUGAGAU CUACAGCGUGUUUAAGACAAGAUGGGAAACCACAGGAAGCUUC
AAAUUCUCUGGGCGAUUUUAAAACCUACGGCAAGGCCGACAACACUGCCUUCAGAG
AACUGAUUACAGCUAAGGAAAUCGAGAAAAACGUGCAGGAGAUGAGCAAGGGCAG
CGCCCCUGGCCCUGACGGUAUACCCUGGGGAGAUGUGGUGAAAAUGGACCCCGAG
UUCAGCCGCACAAUGGAGAU CUUAACCUGUGGCUGACCACUGGCAAAAUCCUG
ACAUGGUGCGGGGAUGCCGCACCGUGCUGAUACCUAAGAGCAGCAAGCCAGAUAG
ACUGAAGGAUAUCAACAACUGGCGGCCUAUCACAAUCGGCAGCAUCCUGCUCCGG
CUGUUUAGCAGAAUCGUAAACCGCCCGUCUGAGUAAGGCCUGUCCUCUGAACCCCA
GACAGCGGGGCUUCAUCCGGGCGCCGGCUGCAGCGAGAACCUCAAACUGCUGCA
GACCAUCAUCUGGUCGGCCAAACGAGAGCACAGACCUCUCGGCGUGGUGUUCGUG
GACAUCGCCAAAGCUUUUGAUACCGUCUCCCAACAGCACAUCAUCCACGCCCUGC

AGCAGCGGGAAGUGGACCCUCACAUCGUGGGACUGGUGUCCAACAUGUACGAGAA
CAUCAGCACUUAUAUACCAACCAAGAGAAACACCCACACAGACAAGAUCAGAUUC
AGAGUGGGCGUUAAGCAGGGCGAUCCUAUGUCUCCUCUGCUGUUAACCUGGGCUA
UGGACCCUCUGCUCUGUAAACUGGAGGAGUCUGGCAAGGGAUACCACAGAGGCCA
AAGCUCCAUCACCGCUAUGGCCUUUGCCGACGACCUGGUGCUGCUGUCUGAUUCU
UGGGAGAACAUAGAAUACCAACAUAAGCAUCCUGGAAACAUAUCUGCAACCUUACAG
GCCUGAAAACCCAGGGACAAAAGUGCCACGGCUUUUACAUAAGCCUACCAAGA
CAGCUACACCAUCAACGACUGCGCCGCCUGGACCAUCAACGGCACCCCUUUAAC
AUGAUCGACCCUGGAGAAAGCGAGAAGUAUCUGGGCCUGCAGUUCGACCCUGGA
UCGGCAUCGCUAGAAGCGGCCUGAGCACAAAGCUGGACUUUUGGCUGCAGAGAAU
UGACCAGGGCCCCUCUGAAGCCUCUGCAGAAAACCGACAUCUGAAAACCUACACA
AUCCCUCGGCUGAUCUACAUUGCCGAUACAGCGAGGUAAAGACCGCCCUGCUGG
AAACCCUGGACCAAAAGAUCAAGACCGCCGUGAAGGAAUGGCUCCACCUGCCCC
CUGCACCUGCGACGCUAUCUGUACAGCAGCACCCAGGGACGGCGGCCUGGGCAUC
ACCAAGCUGGCGGGCCUGAUCCCCUCCGUCCAGGCUAGACGGCUGCAUAGAAUCG
CCCAGAGCAGCGAUGACACAUAAGUGCUUUUAUGGAAAAGGAAAAGAUGGAACA
GCUGCACAAGAAGCUGUGGAUUCAGGCCGGUGGGCGACAGAGAGAACAUUCCUAGC
AUCUGGGAGGCGCCGCCUAGUAGCGAGCCUCCCAACAACGUGUCUACCAAUUCUG
AGUGGGAAGCCCUACACAGAAGGACAAGUUCCUAAGCCUUGUAAUUGGAGAAA
GAACGAGUUAAGAAGUGGACAAAGCUGGCCUCUCAGGGCCGGGGAAUUGUGAAU
UUCGAGCGGGACAAGAUCAAGCAUACUGGAUUCAGUACUACAGAAGAAUCCAC
ACAGAAAGCUGCUGACGGCCCUUCAGCUGAGAGCCAACGUGUACCCACGCGGGA
GUUCCUGGCCAGAGGUAGACAGGACCAGUACAUAAGGCCUGCAGACAUAUGUGAU
GCUGAUUAUCGAGUCUUGCGCCCAUCAUUCGGCAACUGCCCUGUGACACAGGACG
CGAGAAUCAAAAGACACAACUACAUCUGCGAGCUGCUGCUGGAAGAGGCCAAGAA
GAAGGACUGGGUGGUGUUAAGGAACCCCAUAUCAGAGACAGCAAUAAGGAACUC
UAUAAACCUGACCUGAUCUUCGUGAAGGACGCCCGGGCCUGGUCGUGGACGUGA
CCGUGAGAUACGAGGCCGCCAAGUCUAGCCUGGAGGAGGCUGCCGCCGAGAAAGU
GCGGAAGUACAAGCACCUUGAAACAGAAGUGCGGCACCUGACCAACGCCAAGGAC
GUCACAUCUGUGGGCUUCCCCUGGGCGCCAGAGGCCAAUUGGCACCAGGAUAACU
UCAAGCUGCUGACAGAGCUGGGCCUGAGCAAGUCCCGCCAAGUGAAGAUGGCCGA
GACCUUCAGCACCGUGGCCCUUUCUCCUCUGUGGACAUCGUGCACAUGUUCGCU
AGCAGAGCCAGAAAGAGCAUGGUGAUGUAGUGACCUCGAGCUGGUACUGCAUGCA
CGCAAUGCUAGCUGCCCCUUCUCCGUCCUGGGUACCCCGAGUCUCCCCCGACCUC
GGGUCCCAGGU AUGCUCCACCUCACCUGCCCCACUCACCACCUCUGCUAGUUC
CAGACACCUC CCAAGCACGCAGCAAUGCAGCUAAAACGCUUAGCCUAGCCACACC
CCCACGGGAAACAGCAGUGAUUAACCUUUAAGCAAUAAACGAAAGUUUAACUAAGC
UAUACUAACCCAGGGUUGGUCAAUUUCGUGCCAGCCACACCCUGGAGCUAGCAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCAUAUGACUAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

10

20

30

40

29. R2A TaGuアミノ酸配列:

MVTVPDKNPPCPCCGTRVNSVLNLIHLKVSHGKRGVCFRCAKCGKENSNYHSVVC
HFPKCRGPETEKAPAGEWICEVCNRDFTTKIGLGQHKRLAHPAVRNQERIVASQPKE
TSNRGAHKRCWTKEEEELLIRLEAQFEGNKNINKLIAEHITTKTAKQISDKRRLLSRK
PAEEPREEPGTCHHTRRAAASLRTEPEMSHHAQAEDRDNGPGRRPLPGRAAAGGRT
MDEIRRHDPDKNGGQQRPTKQKSEELQAYYKKTLEERLSAGALNTFPRAFKQVMEG
RDIKLVINQTAQDCFGCLESISQIRTATRDKKDTV TREKHPKKPFQKWMKDRAIKKG
NYLRFQRLFYLD RGKLAKIILDDIECLSCDIPLSEIYSVFKTRWETTGSFKSLGDFKTY
GKADNTAFRELITAKEIEKNVQEMSKGSAPGPDGITLGDVVKMDPEFSRTMEIFNLW
LTTGKIPDMVRGCRTVLIPKSSKPDRLKDINNWRPITIGSILLRFSRIVTARLSKACP

50

LNPRQRGFIRAAGCSENKLLQTIWWSAKREHRPLGVVFDIAKAFDTVSHQHIIHAL
QQREVDPHIVGLVSNMYENISTYITTKRNTHTDKIQIRVGVKQGDPMSPLLFLNAMD
PLLCKLEESGKGYHRGQSSITAMAFADDLVLLSDSWENMNTNISILETFCNLTLGLKT
QGQKCHGFYIKPTKDSYTINDCAAWTINGTPLNMIDPGESEKYLGLQFDPWIGIARS
GLSTKLDFWLQRIDQAPLKPLQKTDILKTYTIPRLIYIADHSEVKTALLETLDQKIRT
AVKEWLHLPPCTCDAILYSSSTRDGGGLGITKLAGLIPSVQARRLHRIAQSSDDTMKCF
MEKEKMEQLHKKLWIQAGGDRENIPSIWEAPPSSEPPNNVSTNSEWEAPTQKDKFP
KPCNWRKNEFKKWTKLASQGRGIVNFERDKISNHWIQYYRRIPHRKLLTALQLRAN
VYPTREFLARGRQDQYIKACRHCADADIESCAHIIGNCPVTQDARIKRHNYICELLLEE
AKKKDWVVFKEPHIRDSNKELYKPDILIFVKDARALVVDVTVRYEAAKSSLEEAAAEK
VRKYKHLETEVRHLTNAKDVT FVG FPLGARGKWHQDNFKLLTELGLSKSRQVKMAE
TFSTVALFSSVDIVHMFASRARKSMVM*

10

30. R2A GeFo N-DYKプラスミドORF配列:

ATGGATTACAAGGATGACGACGATAAGGGTACCATGGGCAACATCGTCATGGTGAC
CTTACCAAACAAGAACCCTCCATGTCCCTGTTGCCGGGTGCGGCTGCACAGCGTG
TGGCTCTGATCGAGCATCTGAAGGGCAGCCACGGAAAAAACGGGCATGCTTCCGG
TGCGTGAAATGCGGACGTGAGAACTTCAACTACCACAGCACCGTG TGCCACATCGC
CAAGTGCAAGGGCCCCAAAGTAGAAAAAGCCCCCTGTGGGCGAGTGGATCTGTGAAG
TTTGCGGGCCGAGATTTCAACCACCAAGATCGGCCTGGGACAACACAAGAGACTGGCT
CATCCTCTTGTGCGGAACCAGGAGCGGATCGACGCTTCTCAGCCTAAGGAAACCAG
CAACAGAGGCGCCCATAGAGATGTTGGACAAAAGAGGAAGAAGAGATGCTGATTA
AGCTGGAGGTGCAGTTCGAAGGCCACAGAAACATCAACAAGCTCATCGCTGAGCAC
CTGACCACAAAGACCAGCAAGCAGATCTCTGACAAAAGGAGACTGCTGCCCAGGAA
GCAGCTGACAGACCTGAGCAAGGGGGTGCGCGGCCAGAAAGGTGCTGGACCCTGGCC
TGAGCCACCAGCCTCAGCTGGGCGTGTTGGACAATGGCCTGGGCGGAGGACACCTA
CCTGGCGGGCCCGGCCGCGGAAGGCAGAACGATCGAACCTCTTGGCCACCACCTTGA
TAAGGACAACGGCCACC GGGAGATCGCCGACCAGCACAAAGGCCGGCAGACTGCAGG
CCCCTACAGAAAGAAGATCCGGAAGCGGCTGAGCGAGGGGCATGATCTCCAACCTC
CCCGAGGTGTTTCGAGCAACTGCTGGACTGCCAGGAGGCCCAGCCTCTGATTAACCA
GGCCGCACAGGACTGCTTCGGCTGCCTGGATTCTGCCAGCCAGATTAGAAAAGCTC
TGCGGAAGCAGAACACCCAGAAGGATCAGGGCGACCAGCCTAAGCGGGCCCGCCAG
AAGTGGATGAAGAAGAGAGCCGTCAAGCGGGGCCACTTCCTGAGATTCCAGAAACT
GTTCCACCTGGATCGGGGCAAGCTGGCTAAGATCATCTTGGACGACGTGGAATGCC
TGTCTTGTGATATCCCTCCAAGTGAGATCTACAGCGTGTTCAAGGCCAGATGGGAG
ACACCCGGACAGTTTCGCTGGCCTGGGCGACTTCGAGATCAACAGAAAGGCCAAACAA
TAAAGCTTTTCAGAGATCTGATTACCGCGAAAGAGATCCTGAAGAACGTGAGAGAAA
TGACAAAGGGCTCTGCCCCCTGGTCCTGATGGCATCGCTCTGGGCGACATCAGAAAG
ATGGACCCCGAGTACACCAGAACTGCCGAGCTGTTTAATCTGTGGCTGACATCTGG
CGAGATCCCCGACATGGTGAGAGGCTGTAGAACAGTGCTGATCCCCAAGAGCAGCA
AACCTGAGCGGCTCAAGGACATCAATAATTGGAGACCCATCACCATCGGGAGCATC
CTGCTGCGGCTGTTTCAGCAGAATCATCACCGCTAGACTGACAAAAGCCTGCCCTCT
GAATCCAAGACAGAGAAGCTTTATCAGCGCCGCGGATGTAGCGAGAACCTGAAGC
TCCTGCAGACCATCATCCGGACCGCCAAGAACGAGCACCGGCCACTGGGCGTTGTG
TTCGTGGACATCGCCAAGGCCTTCGACACCGTGAGCCACCAGCACATCATCCACGT
GCTGCAGAGAAGAAGAGTGGACCCACACATCATCGGCCTCGTGAAGAACATGTACA
AAGATATCAGCACCGTCATCACCAACCAAGAAGAACACCTATACAGACAAGATTTCAG
ATCCAGGTGGGCGTGAAGCAAGGCGATCCCCCTGAGCCCTCTCCTTTTTAACCTGGC
TATGGACCCTCTGCTGTGCAAGCTGGAGGAACACGGCAAGGGATTTACAGAGGGCC
AGAGCAAGATCACAGCTATGGCCTTTGCCGATGACCTCGTGCTGCTGTCCGACAGC
TGGGAGGACATGAACGCCAATATCAAGATACTGGAAACCTTCTGCGACCTGACAGG

20

30

40

50

CCTGAAAACCTCAGGGACAGAAGTGTACCGGTTTCTACATCAAACCTACCAAGGACA
GCTACACAGTGAACAACCTGCGCCGCTGGACGATCAACGGTACACCTCTGAACATG
ATTAACCCTGGAGAGAGCGAGAAGTACCTGGGCCTGCAGTTCGATCCCTGGGTAGG
CATCGCCAAAACAAGCCTGCCTGAGAAGCTGGACTTCTGGCTGGAACGCATTGATC
GGGCCCCCTCTGAAGCCTTTCCAAAAGCTGGACATCCTGAAAACCTACACCATCCCT
AGACTCACCTATGTGGCCGACCACTCTGAGATGAAGGCCGGCGCCCTGGAAGCCCT
GGATAGAACCATCCGGAGCGCCGTGAAGGACTGGCTGCACCTGCCTAGCAGCACCT
GTGACGCGATCCTGTACACCAGTATGAAGGACGGTGGACTGGGCGTGACCAAACCTG
GTGGGCCTGATCCCATCCGTGCAGGCCAGACGACTGCACAGAATCGCCCAAAGCCC
CGAGGAAACCATGAAGGACTTCTCTGGAGAAGGCTCAGATGGAAAAAATGTACGAGA
AGCTGTGGGTGCAGGCCGGAGGCAAGAGAAAGAGAATGCCTTCTATCTGGGAGGCT
CTGCCAGAAGTGGTCCCCTCCATCGATACCGCCACAACCTAGCGAGTGGGAGGCCCC
TAACCCCAAGTCCAAGTACCCTAGACCTTGTAAATTGGCGCCGGAAAGAGTTCAAGA
AGTGGACCAAGCTGATTGCTCAAGGATGGGGCATCCGGTGCTTTAAAGGGGATAAG
ATCTCCAATAACTGGATCAGGCATTATAGATACATTCCTCACAGAAAGCTGCTGAC
AGCCATTCAACTGAGAGCCAGCGTCTACCCACCCGGGAGTTCCTGGCCCGGGGCA
GAGAAGATAACTGCGTGAAAAGCTGCCGGCACTGCGAGGCCGCTGAGGAAAGCTGC
GCCACATCATCGGAATGTGCCCCGTGGTGCGGGACGCCAGAATCAAGAGACACAA
CAGAATCTGCGAGAGGCTCATGGAAGAAGCCGGCAAGCGCGATTGGACCGTGTTTC
AGGAGCCTCATATACGTGACGTGACAAAGGAACTGTACAAGCCCGACCTGATCTTC
GTGAAAGAAGGCCTGGCCCTGGTGGTGGACGTGACCATCCGGTTCGAAAGCACCAA
GACCACCCTGGAAGAGGCAGCCGCTGAGAAGGTGAACAAGTACAAGCACCTGGAAA
CGGAAGTCAGAAATCTGACAAACGCCAAGGACGTGATCTTCATGGGCTTTCCTCTG
GGGGCCAGAGGCCAATGGTACAACAAGAACTTCGAGCTGCTGGACACCCTGGGCCT
GCCTAGAAGCAGGCAGGACATCATCGCCAAAACCCTGTCCACCGATGCTCTGATCT
CTTCCGTGGATATTATCCACATGTTTCGCCAGCAGAGGCAGACGGCAGCACGCCTGA
TGA

31. R2A GeFo N-DYKアミノ酸配列:

MDYKDDDDKGTMGNIVMVTLPNKNPPCPCCRVRLHSVLALIEHLKGS HGKKRACFR
CVKCGRENFNHSTVCHIAKCKGPKVEKAPVGEWICEVCGRDFTTKIGLGQHKRLAH
PLVRNQERIDASQPKETSNRGAHKRCWTKEEEEMLIKLEVQFEGHRNINKLIAEHLT
TKTSKQISDKRRLPRKQLTDL SKGVAGQKVLDPLSHQPQLGVVDNGLGGGHLPG
GPAAEGRTIEPLGHHLDKDNHGREIADQHKAGRLQAHYRKKIRKRLSEG M ISN FPEV
FEQLLD CQEAQPLINQAAQDCFGCLDSASQIRKALRKQNTQKDQGDQPKRPAQKWM
KKRAVKRGHFLRFQKLFHLD RGLAKIILDDVECLSCDIPPSEIYSVF KARWETPGQF
AGLGDFEINRKANNKAFRDLITAKEILKNVREM TKGSAPGPDGIALGDIRKMDPEYT
RTAELFNLWLTSGEIPDMVRGCRTVLIPKSSKPERLKDINNWRPITIGSILLRLFSRII
TARLTACPLNPRQRSFISAAGCSENLLKLLQTIIRTAKNEHRPLGVVFVDIAKAFDTV
SHQHIIHVLQRRRVDPHIIGLVKNMYKDISTVITTKNTYTDKIQIQVGVKQGDPLSP
LLFNLAMDPLLCKLEEHGKGFHRGQSKITAMAFADDLVLLSDSWEDMNANIKILETF
CDLTGLKTQGGQKCHGFYIKPTKDSYTVNNCAAWTINGTPLNMINPGESEKYLGLQF
DPWVGIAKTS LPEKLD FWERIDRAPLPKFQKLDILKTYTIPRLTYVADHSEM KAGA
LEALDRTIRSAVKDWLHLPSS TCDAILYTSMKDGGLGVTKLVGLIPSVQARRLHRIA
QSPEETMKDFLEKAQMEKMYEKLWVQAGGKRKRMP SIWEALPEVVPSIDTATTSE
WEAPNPKSKYPRPCNWR RKEFKKWT KLIAQGWGIRCFKGDKISNNWIRHYRYIPHR
KLLTAIQLRASVYPTREFLARGREDNCVKSCRHCEAAEESCAHIIGMCPVVRDARIKR
HNRICERLMEEAGKRDWTVFQEPHIRDVT KELYKPD LIFVKEGLALVVDVTIRFEST
KTTLEEA AAEKV NKYKHLETEVRNL TNAKD VIFMGFPLGARGQWYNKNFELLDTLG
LPRSRQDIIAKTLSTDALISSVDIIHMFASRGRRQHA**

32. R2D DroSi SynM C_3FプラスミドORF配列:

ATGACCACCCGTCCGAGCGTGGATATCTTCCCGGAAGACCAGTACGAGCCGAATGC
GGCCGCTACCCTGAGCCGTGTGCCATGCACCGTGTGCGGTCGTAGCTTTAACAGCA
AACGCGGTCTGGGCGTGCACATGCGCAGCCGTACCCGGACGAGCTGGACGAGGAA
CGTCGCCGTGTGGATATCAAGGCGCGTTGGAGCGAGGAAGAGAAATGGATGATGGC
TCGCAAGGAAGTGGAGCTGACCGCCAACGGCCACAAGCACATGAACAAACAGCTGG
CTGTGTACTTCGCCAACCCTAGCGTGGAGGCCATTAAGAACTGCGCCAGCGTGGT
GACTATAAGGAAAAAATTGAGCAGATCCGTGGTCAGAGCGCTCTGGTGCCAGAAGT
GGCTAACCTGACCATTCTGTCGTCCGAGCCGTAGCGAGCAGAACCACCAGGTGA
CCACCAGCGAAACCACCCCGATCACCCGTTTGAACAGAGCAACCGTGAGATTCTG
CGTACCCTGCGTGGTTACAGCCAGTGGAGTGCCACAGCAAATGGCGTGCTCAGGA
ACTGCAGACCATCATTGACCGCGCCGAACTGGAGGGCAAGGAGACCACCTGCAGT
GCCTGAGCCTGTACCTGCTGGGTATTTTTCCGGCGCAGGGCGTGCGTCATACCCTG
ACCCGTCCACCACGTCTGTCGCGTAACCGTCTGTGAAAGCCGTCTGTCAGCAGTATGC
TGTGGTGACGCGTAACTGGGATAAGCACAAAGGTCGCTGCATCAAAGCCTGCTGA
ACGGCACCGACGAGAGCGTGATGCCGAGCCAGGAAGTGATGGTGCCGTATTGGCGT
GAGGTGATGACCCAGCCAAGCCCAAGCAGCTGCAGCGGTGAAGTGATTCAGATGGA
TCACAGCCTGGAGCGTGTGTGGAGCGCCATCACCGAACACGACCTGCGTGCTAGCC
GCATTAGCCTGAGCAGCAGCCAGGTCCAGATGGTATCACCCCAAAGAGCGCCCGT
GAGGTGCCGAGCGGTATTATGCTGCGCATCATGAACCTGATTCTGTGGTGCGGCAA
CCTGCCGCACAGCATTCTGTCTGGCCCGCACCGTGTTTATTCCGAAGACCGTGACCG
CGAAACGTCCGCAGGACTTTCGCCCAATCAGCGTGCCGAGCGTGCTGGTGCGTCAG
CTGAACGCTATCCTGGCCACCCGCCTGAACAGCAGCATTAACTGGGACCCGCGTCA
GCGTGTTTTCTTGCCAACCGATGGCTGCGCCGACAACGCTACCATCGTGGACCTGG
TGCTGCGTCACAGCCACAAACACTTCCGCAGCTGCTACATTGCCAACCTGGACGTG
AGCAAGGCCTTCGACAGCCTGAGCCACGCTAGCATCTACGATACCCTGCGTGCTTA
TGGTGCGCCGAAGGGCTTTGTGGACTACGTGCAGAACACCTATGAGGGCGGTGGCA
CCAGCCTGAACGGCGACGGCTGGAGCAGCGAAGAGTTCTGTGCCGGCCCGTGGTG
AAACAGGGCGATCCGCTGAGCCCGATCCTGTTCAACCTGGTTATGGACCGTCTGCT
GCGCAACCTGCCGAGCGAGATCGGTGCTAAGGTGGGTAAACGCCATTACCAACGCCG
CGGCTTTCGCCGACGATCTGGTGCTGTTTGCGGAAACCCGTATGGGTCTGCAGGTG
CTGCTGGACAAGACCCTGGACTTCCTGAGCCTGGTGGGCCTGAAGCTGAACGCCGA
CAAATGCTTTACCGTGGGTATCAAGGGCCAGCCGAAGCAGAAATGCACCGTGCTGG
AGGCGCAGAGCTTCTACGTGGGTAGCCGTGAGATCCCGAGCCTGAAACGCACCGAT
GAATGGAAGTATCTGGGTATTAACCTTTACCGCTACCGGTGCGGTGCGTTGCAACCC
AGCGGAGGACATTGGCCCGAAACTGCAGCGTCTGACCAAGGCGCCGCTGAAACCGC
AGCAGCGTATGTTCTGCTCTGCGCACCGTGCTGATCCCGCAGCTGTACCACAAGCTG
GCGCTGGGTAGCGTGGCTATCGGCGTGCTGCGTAAGACCGATAAACTGATTCTGCTA
CTATGTGCGTCGTTGGCTGAACCTGCCACTGGATGTGCCAATTGCGTTCATTCATG
CTCCGCCGAAAAGCGGTGGTCTGGGTATTCCAAGCCTGCGTTGGGTGGCCCCAATG
CTGCGTCTGCGTCGTCTGAGCAACATCAAATGGCCGCACCTGACCCAGAACGAGGT
GGCTAGCAGCTTTCTGGAAGCGGAGAAACAGCGTGCCCGTGATCGTCTGCTGGCTG
AACAGAACGAGCTGCTGAGCCGTCCGGCGATCGAGAAGTACTGGGCTAACAACTG
TATCTGAGCGTGGATGGTAGCGGTCTGCGTGAAGCCGGTCACTGGGGTCCACAGCA
TGGTTGGGTGAACCAGCCAACCCGTCTGCTGACCGGTAAAGAGTACATTGATGGCA
TCCGTCTGCGCATTAATGCTCTGCCAACCAAGAGCCGTACCACCCGTGGTCGTCTAT
GAACTGGAACGCCAGTGCCGTGCTGGTTGCGATGCCCCGGAAACCACCAACCACAT
CATGCAGAAATGCTATCGTAGCCATGGTCTGTCGTGTGGCTCGTCACAACTGCGTGG
TGAACCGTATCAAGCGCGGTCTGGAAGAGCGTGGCTGCGTGGTGATTGTGGAACCG
AGCCTGCAGTGCGAGAGCGGTCTGAACAAACCGGATCTGGTGGCTCTGCGTCAGGA
CCACATTGATGTGATCGACATTCAGATCGTGACCGACGGCCACAGCATGGACGATG

10

20

30

40

50

CCCACCAGCGTAAGATCAACCGTTACGATCGCCCGGACATTTCGCACCGAACTGCGT
CGTCGTTTTGAGGCGGCCGGTGATATCGAATTTACAGCGCGACCCTGAACTGGCG
TGGTATCTGGAGCGGCCAGAGCGTGAAGCGCCTGATTGCTAAAGGCCTGCTGAGCA
AGTATGACAGCCACATCATTAGCGTGACGGTGATGCGTGGTTCCCTGGGCTGTTTC
AAGCAATTCATGTATCTGAGCGGTTTCTCCCGTGACTGGACCATGGGGAGCGACTA
CAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATG
ACAAGAAGTAGTAA

33. R2D DroSi SynM C_3Fアミノ酸配列:

MTTRPSVDIFPEDQYEPNAAATLSRVPCTVCGRSFNSKRGLGVHMRSRHPDELDEER
RRVDIKARWSEEEKWMMARKEVELTANGHKHMNKQLAVYFANRSVEAIKKLRQRG
DYKEKIEQIRGQSALVPEVANLTIRRRPSRSEQNHQVTTSETTPITPFEQSNREILRTL
RGYSPVECHSKWRAQELQTIIDRAELEGKETTLQCLSLYLLGIFPAQGV RHTLTRPPR
RPRNRRESRRQQYAVVQRNWDKHKGRCIKSLLNGTDESVMPSQEVMPYWREVM T
QPSPSSSCSGEVIQMDHSLERVWSAITEHDLRASRISLSSSPGPDGITPKSAREVPSGI
MLRIMNLILWCGNLPHSIRLARTVFIPKTVTAKRPQDFRPISVPSVLVRQLNAILATR
LNSSINWDPRQRGFLPTDGCADNATIVDLVLRHSHKHFRSCYIANLDVSKAFDSLH
ASIYDTLRAYGAPKGFVDYVQNTYEGGGTSLNGDGSSEEFVPARGVKQGDPLSPIL
FNLVMDRLLRNLPSEIGAKVGNAITNAAAFADDLVLF AETRMGLQVLLDKTLDLFLSL
VGLKLNADKCF TVGIKGQPKQKCTVLEAQS FYVGSREIPSLKRTDEWKYLGINFTAT
GRVRCNPAEDIGPKLQRLTKAPLKPQQRMFALRTVLIPQLYHKLALGSVAIGVLRKT
DKLIRYYVRRWLNLPDVP IAFIHAPPKSGGLGIPSLRWVAPMLRLRLRLSNIKWPHL
TQNEVASSFLEAEKQRARDRL LAEQNELLSRPAIEKYWANKLYLSVDGSGGLREAGH
WGPQHGWVNQPTRLTGKEYIDGIRLRINALPTKSRTTRGRHELERQCRAGCDAPE
TTNHIMQKCYRSHGRRVARHNCVVNR IKRGLEERGCVVIVEPSLQCESGLNKPDLVA
LRQDHIDVIDIQIVTDGHSMDDAHQRKINRYDRPDIRTELRRRFEAAGDIEFHSATLN
WRGIWSGQSVKRLIAKGLLSKYDSHIISVQVMRGS LGCFKQFMYLSGFSRDWTMGS
DYKDHGDYKDHIDYKDDDDKK**

34. R2D DroSi SynM C_3FトランスフェクトしたmRNA配列:

AGGAAUAAACUAGUAUUCUUCUGGUCCCCACAGACUCAGAGAGAACCCGCCACCA
UGACCACCCGUGCCGAGCGUGGAUAUCUUC CCGGAGGACCAGUACGAGCCGAAUGC
GGCCGCUACCCUGAGCCGUGUGCCAUGCACCGUGUGCGGUCGUAGCUUUAACAGC
AAACGCGGUCUGGGCGUGCACAUGC GCAGCCGUCACCCGGACGAGCUGGACGAGG
AACGUCGCCGUGUGGAUAUCAAGGCGCGUUGGAGCGAGGAAGAGAAAUGGAUGAU
GGCUCGCAAGGAAGUGGAGCUGACCGCCAACGGCCACAAGCACAUGAACAAACAG
CUGGCUGUGUACUUCGCCAACCGUAGCGUGGAGGCCAUUAAGAAACUGCGCCAGC
GUGGUGACUAUAAGGAAAAAAUUGAGCAGAUCCGUGGUCAGAGCGCUCUGGUGCC
AGAAGUGGCUAACCUGACCAUUCGUCGUCGUCCGAGCCGUAGCGAGCAGAACCCAC
CAGGUGACCACCAGCGAAACCACCCCGAUACCCCGUUCGAACAGAGCAACCGUGA
GAUUCUGCGUACCCUGCGUGGUUACAGCCCAGUGGAGUGCCACAGCAA AUGGCGU
GCUCAGGAACUGCAGACCAUCAUUGACCGCGCCGAACUGGAGGGGCAAGGAGACCA
CCCUGCAGUGCCUGAGCCUGUACCGUGGUAUUUUUCCGGCGCAGGGGCGUGCG
UCAUACCCUGACCCGUCCACCACGUCGUCCGCGUAACCGUCGUGAAAGCCGUCGU
CAGCAGUAUGCUGUGGUGCAGCGUAACUGGGGAUAAGCACAAAGGUCGCGUCAUCA
AAAGCCUGCUGAACGGCACCGACGAGAGCGUGAUGCCGAGCCAGGAAGUGAUGGU
GCCGUAUUGGCGUGAGGUGAUGACCCAGCCAAGCCCAAGCAGCUGCAGCGGUGAA
GUGAUUCAGAUUGGAUACAGCCUGGAGCGUGUGUGGAGCGCCAUCACCGAACACG
ACCGCGUGCUAGCCGCAUUAAGCCUGAGCAGCAGCCAGGUCCAGAUUGGUUAUCAC
CCCAAAGAGCGCCCGUGAGGUGCCGAGCGGUAUUAUGCUGCGCAUCAUGAACCG
AUUCUGUGGUGCGGCAACCUGCCGCACAGCAUUCGUCUGGCCCGCACCGUGUUA
UCCGAAAACCGUGACCGCGAAACGUCCGCAGGACUUCGCCCCAAUCAGCGUGCC

GAGCGUGCUGGUGCGUCAGCUGAACGCUAUCCUGGCCACCCGCCUGAACAGCAGC
 AUUAACUGGGACCCGCGUCAGCGUGGUUCCUGCCAACCGAUGGCUGCGCCGACA
 ACGCUACCAUCGUGGACCUGGUGCUGCGUCACAGCCACAAACACUUCCGCAGCUG
 CUACAUUGCCAACCUGGACGUGAGCAAGGCCUUCGACAGCCUGAGCCACGCUAGC
 AUCUACGAUACCCUGCGUGCCUAUGGUGCGCCGAAGGGCUUUGUGGACUACGUGC
 AGAACACCUAUGAGGGGCGGUGGCACCAGCCUGAACGGCGACGGCUGGAGCAGCGA
 AGAGUUCGUGCCGGCCCGUGGUGUGAAACAGGGCGAUCCGCUGAGCCCGAUCCUG
 UUCAACCUGGUUAUGGACCGUCUGCUGCGCAACCUGCCGAGCGAGAUCCGGUGCUA
 AGGUGGGUAACGCCAUUACCAACGCCGCGGCUUUCGCCGACGAUCUGGUGCUGUU
 UGCGGAAACCCGUAUGGGUCUGCAGGUGCUGCUGGACAAGACCCUGGACUCCUG 10
 AGCCUGGUGGGCCUGAAGCUGAACGCCGACAAAUUCUUAACCGUGGGUAUCAAGG
 GCCAGCCGAAGCAGAAUAGCACCGUGCUGGAGGCGCAGAGCUUCUACGUGGGUAG
 CCGUGAGAUCCCGAGCCUGAAACGCACCGAUGAAUGGAAGUAUCUGGGUAUUAAC
 UUUACCGCUACCGGUCGCGUGCGUUGCAACCCAGCGGAGGACAUAUGGCCCGAAAC
 UGCAGCGUCUGACCAAGGCGCCGCGUGAAACCGCAGCAGCGUAUGUUCGCUCUGCG
 CACCGUGCUGAUCCCGCAGCUGUACCACAAGCUGGCGCUGGGUAGCGUGGCUAUC
 GCGUGCUGCGUAAGACCGAUAAACUGAUUUCGUACUAUGUGCGUCGUUGGCUGA
 ACCUGCCACUGGAUGUGCCAAUUGCGUUCAUUAUGCUCCGCCGAAAAGCGGUGG
 UCUGGGUAUUAAGCCUGCGUUGGGUGGCCCAAUGCUGCGUCUGCGUCGUCUG
 AGCAACAUCAAAUGGCCGCGACCUGACCCAGAACGAGGUGGCUAGCAGCUUUCUGG 20
 AAGCGGAGAAACAGCGUGCCCGUGAUCGUCUGCUGGCGUAACAGAACGAGCUGCU
 GAGCCGUCCGGCGAUCGAGAAGUACUGGGCUAACAAACUGUAUCUGAGCGUGGAU
 GGUAGCGGUCUGCGUGAAGCCGGUCACUGGGGUCCACAGCAUGGUUGGGUGAACCC
 AGCCAACCCGUCUGCUGACCGGUAAAGAGUACAUAUGAUGGCAUCCGUCUGCGCAU
 UAAUGCUCUGCCAACCAAGAGCCGUACCACCCGUGGUCGUAUGAACUGGAACGC
 CAGUGCCGUGCUGGUUGCGAUGCCCCGGAAACCAACCAACAUAUGCAGAAAU
 GCUAUCGUAGCCAUGGUCGUCGUGUGGCUCGUCACAACUGCGUGGUGAACCGUAU
 CAAGCGCGGUCUGGAAGAGCGUGGCUGCGUGGUGAUUUGUGGAACCGAGCCUGCAG
 UGCGAGAGCGGUCUGAACAACCCGGAUCUGGUGGCUCUGCGUCAGGACCACAUUG
 AUGUGAUCGACAUAUCAGAUUCGUGACCGACGGCCACAGCAUGGACGAUGCCCAACA 30
 GCGUAAGAUAACCGUUAACGAUCGCCCGGACAUAUCGCACCGAACUGCGUCGUCGU
 UUUGAGGCGGCCGGUGAUUAUCGAAUUAACAGCGCGACCCUGAACUGGCGUGGUA
 UCUGGAGCGGCCAGAGCGUGAAGCGCCUGAUUGCUAAAGGCCUGCUGAGCAAGUA
 UGACAGCCACAUAUAUAGCGUGCAGGUGAUGCGUGGUUCCUGGGCUGUUUCAAG
 CAUUAUGUAUCUGAGCGGUUUCUCCCGUGACUGGACCAUGGGGAGCGACUACA
 AAGACCAUGACGGUGAUUAUAAAGAUCAUGACAUCGAUUAACAAGGAUGACGAUGA
 CAAGAAGUGAUGACUCGAGCUGGUACUGCAUGCACGCAAUGCUAGCUGCCCCUUU
 CCCGUCCUGGGUACCCCGAGUCUCCCCGACCUCGGGUCCAGGUUAUGUCCAC
 CUCCACCUGCCCCACUCACCACCUCUGCUAGUUCACAGACACCUCCCAAGCACGCAG
 CAUUGCAGCUCAAAACGCUUAGCCUAGCCACACCCCCACGGGAAACAGCAGUGAU 40
 UAACCUUUAGCAAUAAACGAAAGUUUAACUAAGCUAUACUAACCCAGGGUUGGU
 CAUUUUCGUGCCAGCCACACCCUGGAGCUAGCAAAAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAGCAUAUGACUAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

35. R2D BoMo D69 C_3FプラスミドORF配列:

ATGAAGAAAAGCAACAAGGAGAACCGTCCGGAAGCGAGCGGTCTGCCGCTGGAGAG
 CGAACGTACCGGCGATAACCCGACCGTGCGTGGTAGCGCGGGTGCGGACCCGGTTG
 GTCAGGATGCGCCGGGTTGGACCTGCCAATTCTGCGAGCGTACCTTTAGCACCAAC
 CGTGGTCTGGGCGTGCACAAGCGTCGTGCGCACCCGGTTGAAACCAACACCGACGC
 GCGCGCGATGATGGTGAAACGTCGTTGGCACGGCGAGGAAATCGATCTGCTGGCGC 50

GTACCGAGGCGCGTCTGCTGGCGGAACGTGGCCAGTGCAGCGGTGGCGACCTGTTTC
GGCGCGCTGCCGGGTTTTGGTTCGTACCCTGGAGGCGATTAAAGGTCAACGTCGTCTG
TGAACCGTATCGTGCCTGGTTCAGGCGCATCTGGCGCGTTTTTGGTAGCCAACCGG
GTCCGAGCAGCGGTGGCTGCAGCGCGGAGCCGGATTTTCGTCTGTGCGAGCGGTGCG
GAGGAAGCGGGCGAGGAACGTTGCGCGGAAGATGCGGCGGCGTATGATCCGAGCGC
GGTGGGTCAAATGAGCCCGGATGCGGCGCGTGTGCTGAGCGAACTGCTGGAGGGGTG
CGGGTCGTCTGTGCTGCGTGCCGTGCGATGCGTCCGAAGACCGCGGGTCTGCTGAAC
GACCTGCACGACGATCGTACCGCGAGCGCGCACAAAGACCAGCCGTCAGAAACGTCG
TGCGGAGTACGCGCGTGTGCAAGAACTGTATAAGAAATGCCGTAGCCGTGCGGCGG
CGGAAGTGATCGATGGTGCCTGCGGTGGCGTTGGTTCACAGCCTGGAGGAAATGGAA
ACCTACTGGCGTCCGATTCTGGAACGTGTGAGCGACGCGCCGGGTCCGACCCCGGA
GGCGCTGCACGCGCTGGGTCTGTCGGAATGGCACGGTGGCAACCGTGATTATACCC
AGCTGTGGAAGCCGATCAGCGTTGAGGAAATTAAAGCGAGCCGTTTCGACTGGCGT
ACCAGCCCGGGTCCGGATGGTATCCGTAGCGGCCAGTGGCGTGCGGTGCCGGTTCA
CCTGAAGGCGGAAATGTTCAACGCGTGGATGGCGCGTGGCGAGATCCCGGAAATTC
TGCGTCAATGCCGTACCGTGTGTTGTTCCGAAAGTTGAGCGTCCGGGTGGCCCGGGT
GAATACCGTCCGATCAGCATTGCGAGCATCCCGCTGCGTCACTTCCACAGCATTCT
GGCGCGTCTGCTGCTGGCGTGCTGCCCGCCGGACGCGCGTCAAGCGTGGCTTTATCT
GCGCGGATGGTACCCTGGAGAACAGCGCGGTGCTGGACGCGGTTCTGGGTGATAGC
CGTAAGAAACTGCGTGAATGCCACGTGGCGGTTCTGGACTTCGCGAAGGCGTTTGA
TACCGTGAGCCACGAGGCGCTGGTTGAACTGCTGCGTCTGCGTGGCATGCCGGAGC
AGTTCTGCGGTTACATTGCGCACCTGTATGACACCGCGAGCACCAACCTGGCGGTG
AACAACGAAATGAGCAGCCCGGTGAAAGTTGGCCGTGGTGTTCGTCAAGGCGACCC
GCTGAGCCCGATCCTGTTTAACTGGTTATGGATCTGATTCTGGCGAGCCTGCCGG
AGCGTGTGGGTTACCGTCTGGAGATGGAAGTGGTTAGCGCGCTGGCGTATGCGGAC
GATCTGGTGCTGCTGGCGGGCAGCAAGGTTGGTATGCAGGAAGCATCAGCGCGGT
GGACTGCGTTGGCCGTCAAATGGGTCTGCGTCTGAACTGCCGTAAAAGCGCGGTGC
TGAGCATGATCCCGGATGGTCAACGTAAGAAACACCACTACCTGACCGAGCGTACC
TTCAACATTGGTGGCAAGCCGCTGCGTCAGGTGAGCTGCGTTGAACGTTGGCGTTA
TCTGGGCGTGGACTTTGAGGCGAGCGGTTGCGTTACCCTGGAACACAGCATCAGCA
GCGCGCTGAACAACATTAGCCGTGCGCCGCTGAAACCGCAGCAACGTCTGGAGATC
CTGCGTGCGCACCTGATTCCGCGTTTCCAGCACGGCTTTGTTCTGGGTAAACATCAG
CGACGATCGTCTGCGTATGCTGGATGTGCAGATTTCGTAAGGCGGTTGGTCAATGGC
TGCGTCTGCCGGCGGACGTGCCGAAAGCGTACTATCACGCGGCGGTTCAAGATGGT
GGCCTGGCGATCCCGAGCGTGCGTGCGACCATCCCGGACCTGATTGTTTCGTCTGTTT
TGGTGGCCTGGATAGCAGCCCGTGGAGCGTGGCGCGTGCGGCGGCGAAGAGCGACA
AAATTCGTAAGAAACTGCGTTGGGCGTGGAAGCAGCTGCGTCGTTTCAGCCGTGTG
GATAGCACCAACCAACGTCCGAGCGTTCGTCTGTTTTGGCGTGAGCACCTGCACGC
GAGCGTTGACGGTCGTGAGCTGCGTGAAAGCACCCGTACCCCGACCAAGCAACAAAT
GGATTCGTGAACGTTGCGCGCAGATTACCGGTCTGTGATTTCGTGCAATTTGTTTAC
ACCCACATCAACGCGCTGCCGAGCCGTATTTCGTGGCAGCCGTGGCCGTCTGTGGTGG
CGGTGAGAGCAGCCTGACCTGCCGTGCGGGTTGCAAAGTGCGTGAAACCAACCGCGC
ACATCCTGCAGCAATGCCACCGTACCCACGGCGGTTCGTATCCTGCGTCACAACAAG
ATTGTGAGCTTCGTTGCGAAGGCGATGGAGGAAAACAAATGGACCGTGGAGCTGGA
ACCGCGTCTGCGTACCAGCGTTGGCCTGCGTAAACCGGACATCATTGCGAGCCGTG
ATGGCGTGGGTGTTATCGTGGACGTTTCAGGTGGTTAGCGGTCAACGTAGCCTGGAT
GAGCTGCACCGTGAAAAGCGTAACAAATACGGCAACCACGGTGAGCTGGTTGAGCT
GGTTGCGGGCCGTCTGGGTCTGCCGAAAGCGGAGTGCGTGCGTGCGACCAAGCTGCA
CCATTAGCTGGCGTGGCGTTTGGAGCCTGACCAAGCTATAAAGAGCTGCGTAGCATC
ATTGGTCTGCGTGAACCGACCCCTGCAGATCGTGCCGATTCTGGCGCTGCGTGGCAG

10

20

30

40

50

CCACATGAACTGGACCCGTTTTTAACCAAATGACCAGCGTGATGGGTGGCGGTGTTG
GTGGTGGAGGTAGCGGGGGCAGTGGAGGGATGGGGAGCGACTACAAAGACCATGAC
GGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATGACAAGAAGTAATAA
36. R2D BoMo D69 C_3Fアミノ酸配列:

MKKSNNKENRPEASGLPLESERTGDNPTVRGSAGADPVGQDAPGWTCQFCERTFSTN
RGLGVHKRRRAHPVETNTDAAPMMVKRRWHGEEIDLLARTEARLLAERGQCSGGDLF
GALPGFGRTLLEAIKGQRRREPYRALVQAHLARFGSQPGPSSGGCSAEPDFRRASGAE
EAGEERCAEDAAAYDPSAVGQMSPDAARVLSELLEGAGRRRACRAMRPKTAGRRND
LHDDRTASAHKTSRQKRRAEYARVQELYKKCRSRAAAEVIDGACGGVGHSLSEMET
YWRPILERVSDAPGPTPEALHALGRAEWHGGNRDYTQLWKPISEVEIKASRFDWRT
SPGPDGIRSGQWRAVPVHLKAEMFNAWMARGEIPEILRQCRTVFVPKVERPGGPGE
YRPISIASIPLRHFHSILARRLLACCPDARQRGFICADGTLENSAVLDAVLGDSRKKL
RECHVAVLDFAKAFDTSHEALVELLRLRGMPEQFCGYIAHLYDTASTTLAVNNEM
SSPVKVGRGVRQGDPLSPILFNVVMDLILASLPERVGYRLEMELVSALAYADDLVLL
AGSKVGMQESISAVDCVGRQMGLRLNCRKSAVLSMIPDGHRKKHHYLTERTFNIGG
KPLRQVSCVERWRYLGVDFEASGCVTLEHSISSALNNISRAPLKPQQRLEILRAHLIP
RFQHGFVLGNISDDRLRMLDVQIRKAVGQWLRLPADVPKAYYHAAVQDGGGLAIPSV
RATIPDLIVRRFGGLDSSPWSVARAAAKSDKIRKKLRWAWKQLRRFSRVDSTTQRPS
VRLFWRHLHASVDGRELRESTRTPSTKWIRERCAQITGRDFVQFVHTHINALPSR
IRGSRGRRGGGESSLTCRAGCKVRETTAHILQQCHRTHGGRI LRHNKIVSFVAKAME
ENKWTVELEPRLRTSVGLRKPDIASRDGVGVIVDVQVVSQGQSLDELHREKRKNKYG
NHGELVELVAGRLGLPKAECVRATSCTISWRGVWSLTSYKELRSIIGLREPTLQIVPI
LALRGSHMNWTRFNQMTSVMGGGVGGGGSGGSGGMGSDYKDHDGDYKDHDIDYK
DDDDKK

10

20

30

40

50

37. R2D BoMo D69 N-DYKトランスフェクトしたmRNA配列:

AGGAAUAAACUAGUAUUCUUCUGGUCCCCACAGACUCAGAGAGAACCCGCCACCA
UGGAUUACAAGGAUGACGACGAUAAGGGAUCCAAGAAAAGCAACAAGGAGAACCG
UCCGGAAGCGAGCGGUCUGCCGCUGGAGAGCGAACGUACCGGCGAUAAACCCGACC
GUGCGUGGUAGCGCGGGUGCGGACCCGGUUGGUCAGGAUGCGCCGGGUUGGACCU
GCCAAUUCUGCGAGCGUACCUUUAGCACCAACCGUGGUCUGGGCGUGCACAAGCG
UCGUGCGCACCCGGUUGAAACCAACACCGACGCGGCGCCGAUGAUGGUGAAACGU
CGUUGGCACGGCGAGGAAAUCGAUCUGCUGGCGCGUACCGAGGCGCGUCUGCUGG
CGGAACGUGGCCAGUGCAGCGGUGGCGACCUGUUCGGCGCGCUGCCGGGUUUUGG
UCGUACCCUGGAGGCGAUUAAAGGUCAACGUCGUCGUGAACCGUAUCGUGCGCUG
GUUCAGGCGCAUCUGGCGCGUUUUGGUAGCCAACCGGGUCCGAGCAGCGGUGGCU
GCAGCGCGGAGCCGGAUUUUCGUCGUGCGAGCGGUGCGGAGGAAGCGGGCGAGGA
ACGUUGCGCGGAAGAUGCGGCGGCGUAUGAUCCGAGCGCGGUGGGUCAAUGAGC
CCGGAUGCGGCGCGUGUGCUGAGCGAACUGCUGGAGGGUGCGGGUCGUCGUCGUG
CGUGCCGUGCGAUGCGUCCCAAACCGCGGGUGCUGUAACGACCUGCACGACGA
UCGUACCGCGAGCGCGCACAAGACCAGCCGUCAGAAACGUCGUGCGGAGUACGCG
CGUGUGCAAGAACUGUAUAAGAAAUGCCGUAGCCGUGCGGCGGCGGAAGUGAUCG
AUGGUGCGUGCGGUGGCGUUGGUACACAGCCUGGAGGAAAUGGAAACCUACUGGCG
UCCGAUUCUGGAACGUGUGAGCGACGCGCCGGGUCCGACCCCGGAGGCGCUGCAC
GCGCUGGGUCGUGCGGAUUGGCACGGUGGCAACCGUGAUUAUACCCAGCUGUGGA
AGCCGAUCAGCGUUGAGGAAAUAAGCGAGCCGUUUCGACUGGCGUACCAGCCC
GGGUCCGGAUGGUUAUCCGUAGCGGCCAGUGGCGUGCGGUGCCGGUUCACCUGAAG
GCGGAAAUGUUAACGCGUGGAUGGCGCGUGGCGAGAUCCCGGAAAUUCUGCGUC
AAUGCCGUACCGUGUUGUUCGAAAGUUGAGCGUCCGGGUGGCCCGGGUGAAUA
CCGUCCGAUCAGCAUUGCGAGCAUCCCGCUGCGUCACUCCACAGCAUUCUGGCG
CGUCGUCUGCUGGCGUGCUGCCCGCCGACGCGCGUCAGCGUGGCUUUAUCUGCG

CGGAUGGUACCCUGGAGAACAGCGCGGUGCUGGACGCGGUUCUGGGUGAUAGCCG
UAAGAAACUGCGUGAAUGCCACGUGGCGGUUCUGGACUUCGCGAAGGCGUUUGAU
ACCGUGAGCCACGAGGCGCUGGUUGAACUGCUGCGUCUGCGUGGCAUGCCGGAGC
AGUUCUGCGGUUACAUUGCGCACCUUGUAUGACACCGCGAGCACCAACCCUGGCGGU
GAACAACGAAAUGAGCAGCCCGGUGAAAGUUGGCCGUGGUGUUCGUCAAGGCGAC
CCGCUGAGCCCGAUCCUGUUUAACGUGGUUAUGGAUCUGAUUCUGGCGAGCCUGC
CGGAGCGUGUGGGUUAACGUCUGGAGAUGGAACUGGUUAGCGCGCUGGCGUAUGC
GGACGAUCUGGUGCUGCUGGCGGGCAGCAAGGUUGGUAUGCAGGAAAGCAUCAGC
GCGGUGGACUGCGUUGGCCGUCAAAUGGGUCUGCGUCUGAACUGCCGUAAAAGCG
CGGUGCUGAGCAUGAUCCCGGAUGGUCACCGUAAGAAACACCACUACCUGACCGA
GCGUACCUUCAACAUUGGUGGCAAGCCGCGUCGUCAGGUGAGCUGCGUUGAACGU
UGGCGUUAUCUGGGCGUGGACUUGAGGGCGAGCGGUUGCGUUAACCCUGGAACACA
GCAUCAGCAGCGCGCUGAACAAACAUAUAGCCGUGCGCCGUGAAACCGCAGCAACG
UCUGGAGAUCUGCGUGCGCACCUUGAUUCCGCGUUUCCAGCACGGCUUUGUUCUG
GGUAACAUCAGCGACGAUCGUCUGCGUAUGCUGGAUGUGCAGAUUCGUAAGGCGG
UUGGUCAAUGGCGUGCGUCUGCCGGCGGACGUGCCGAAAGCGUACUAUCACGCGGC
GGUUCAAGAUGGUGGCCUGGCGAUCCCGAGCGUGCGUGCGACCAUCCCGGACCUG
AUUGUUCGUCGUUUUGGUGGCCUGGAUAGCAGCCCGUGGAGCGUGGCGCGUGCGG
CGGCGAAGAGCGACAAAUAUCGUAAGAAACUGCGUUGGGCGUGGAAGCAGCUGCG
UCGUUUCAGCCGUGUGGAUAGCACCAACCAACGUCCGAGCGUUCGUCUGUUUUGG
CGUGAGCACCUGCACGCGAGCGUUGACGGUCGUGAGCUGCGUGAAAGCACCCGUA
CCCCGACCAGCACCAAAUGGAUUCGUGAACGUUGCGCGCAGAUUACCGGUCGUGA
UUUCGUGCAAUUUGUUCACACCCACAUCAACGCGCUGCCGAGCCGUAUUCGUGGC
AGCCGUGGCCGUCGUGGUGGGCGGUGAGAGCAGCCUGACCUGCCGUGCGGGUUGCA
AAGUGCGUGAAACCACCGCGCACAUCCUGCAGCAAUGCCACCGUACCCACGGCGG
UCGUAUCCUGCGUCACAACAAGAUUGUGAGCUUCGUUGCGAAGGCGAUGGAGGAA
AACAAAUGGACCGUGGAGCUGGAACCGCGUCUGCGUACCAGCGUUGGCCUGCGUA
AACCGGACAUCAUUGCGAGCCGUGAUUGGCGUGGGUGUUAUCGUGGACGUUCAGGU
GGUUAAGCGGUCAACGUAGCCUGGAUGAGCUGCACCGUGAAAAGCGUAACAAAUAC
GGCAACCACGGUGAGCUGGUUGAGCUGGUUGCGGGCCGUCUGGGUCUGCCGAAAG
CGGAGUGCGUGCGUGCGACCAAGCUGCACCAUUAAGCUGGCGUGGCGUUUGGAGCCU
GACCAGCUAUAAGAGCUGCGUAGCAUCAUUGGUCUGCGUGAACCGACCCUGCAG
AUCGUGCCGAUUCUGGCGCUGCGUGGCGAGCCACAUGAACUGGACCCGUUUUAACC
AAUAGACCAGCGUGAUUGGUGGGCGGUGUUGGUUAGUGACCUCGAGCUGGUACUGC
AUGCACGCAAUGCUAGCUGCCCCUUCGUGCCGUGGGUACCCCGAGUCUCCCCCG
ACCUCGGGUUCCAGGUUAGCUCCACCUCCACCUCCGCCCCACUCACCACCUUCUGCU
AGUUCCAGACACCUCCCAAGCACGACGCAUAGCAGCUCAAACGCUUAGCCUAGC
CACACCCCCACGGGAAACAGCAGUGAUUAACCUUUAAGCAAUAACGAAAGUUUAA
CUAAGCUAUACUAACCCAGGGUUGGUCAAUUUCGUGCCAGCCACACCCUGGAGC
UAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCAUAUGACUAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAA

38. R9 AdVa プラスミド ORF 配列:

ATGGATTACAAGGATGACGACGATAAGGGTACCGCATTCAACATTCTTACCCTAT
CGGTCTGTGTTGATTGTGAAGTGTGCCATGAGGATATAGGTCAGATGAACCTGCCGA
TTCGTGAACACGCAGTGAGCGTGCACAACATCAATAAATTCAACTACCTGTGTCAG
CTGTGCAGCAAAAGCTACGACACCATCAACTCCGTAAAGGCCATTATGTTGCATG
TCGTCTGCAGAGAATGCTTCTAGTACGACCGCGGTTCCACCAACGTGATTAATA
ACAACCAGCTGGCAATTAACACAAACCAAGTGATTAGCCGTAATCCGCTACAGTGC
GTGGAATGCCTGATGAAGCAGGTGGATTTCTACGCCAAGGACACCAAAGCCCTGGT

TACCCACATGCGGACCAAACATGCCGCCGCTACGAAGAGTCCAAGAAAGTTGCCA
CCAGAAGAGTTGCATGGAGCCCGGATGAGGACCAGATTCTGGCAGAACTGGAAGTG
AAGCTGAAAAAGATCCAGAAAGGACAGCTGCTGAGCCGTCTGGTGGTCGAATATAA
CAAATGCGCCGATAAAAGCAAAGCACCGAGCAGGTCGAAAGATGCCATTTCGCACGC
GGAGACAACAGCACGACTACAAACTGCTGCTTAGAAGCCTGCAGAGCCAACAGCCT
CCTGTGGGTAGCGAGGATAGTGATTCCGACATCAGCTCCAGCAATAATAATCCGCT
TACAACCACCCATAATGTTACCCCGACCCCTGACAGCAGCAATGTTGTGCTGCTGA
TCCAAAAGATTTCGGGAAAGCGTGGACAGCATCGTGAAAATTACAAATCTGAAACTG
AACACCAATATGCTGAATGCAGCAAGCGCATTTATCAACCAGAACAATAACATGGA
CCCGTTGGAGTTAAGCATGAGAGGCATTGAGGAGGACGTTAAAGCCATTAGAGATA
AAGAACTACAGAAACCGACACGTAATGTCCCTAGTAGCACCACCAGCAGAAAGCCG
ACCAGAAATGCCAAGCGGCTGGAAAAAAGCAAGAAGTATGGTTATTATCAGCATCT
TTATTACAACAACAAAAAGAACTGGTGGCTGAGATTCTGGATGGCGAGACATCTG
GCGCCAAGCCTCCGCCTATGAATCTGGTTGAGGATTATTATCGTAATATTTGGAGC
AGATCTACCATTGATGATTCTCCCGTCAACAATATCAAACTGTTAATTCCGATTCT
TATCTTTGCCCAATTTCTCGTGATGAAATCAAACTAGCTCTGAGCAATACCAAGA
AAGATTCCGCCGCAGGACCTGATGCCGTTACCATTAAAGAAGCAAAAGCTATAATC
GATAATCTGTACGTGGCCTATAATATCTGGCTGGGTGTTTCAAGGGCATTCCGGAACA
ACTGAAGCTGAATAAAACCATCCTGATTCCGAAAGGCAACAGCGACCTGAGCCTGC
TGAAAAATTGGCGTCCGATTACCATCTCTTCCATTATTCTGCGTGTTTATAATCGTC
TGCTGGCCTATAGAATGAATAAAATCTTCAAAACAAATGACAAGCAAGTTGGTTTC
AAGCCTGTTAATGGCTGCGGTATTAACATTAGCTGGCTGCATAGCCTGCTGAAGCA
TGCACGTCTGAATAAAAAATAGCATCTACGCTTGTCTGGTTGACGTGTCCAAGGCCT
TTGATAGCGTGAGCCACCAGAGCATCGTGAGAGCTCTGACCATGAATGGAGCACCG
TCTCTGCTGGTTAAACTCATTATGGATCAGTACACTAATGTTAATACCGTAATCAC
ATGTAGCGGTAGCATCAGCAACAAGATTAATATTAGCAGCGGTGTCAAGCAAGGTG
ATCCTCTGAGTTCTCTGCTGTTCAACTTGGTTATTGATGAACTGTTTCGACGTGATCA
AGGATCAGTATGGATATACTATCGACAACATTGGGACCACCAATGCCAGATGCTTT
GCAGATGATCTGACCCTGATCTCTAGCAGCCGTATGGGCATGAACAACTGCTGGA
ACTGACAACAAAGTTCTTCAAGGAACGTGGTCTGAATGTTAACCCAGCAAGTGCA
TGTCATCGGTATGAGCAAGGGTTACAAGGGTAAAAAAGCAAAATTGAAAGCGAG
CCACTGTTTAGCATTACCGACGCACAGATCCCCATGCTGGGCTATATTGACAAAAC
AACACGTTATCTGGGTGTAACTTCACTTCCATTGGCGCAATCGACGCCAAGAGGA
TTAAAAAGGACCTGCAGGATACTCTGGACAAGCTGGAACACCTGAAACTCAAAGCA
CAGTGTAATAATGGACCTGCTACGTACCTACATGATCCCTCGGTTTATGTTTCAGCT
CATACACACCGAGCTGTATCCTAAGCTGCTGATCAAAATGGATATCCTCATCAGAA
AGCTGGCTAAACGTATACTTCATCTGCCTATCAGCACCAGCAGCGAATTTTTTTAT
CTGCCGTTTAAAGAAGGCGGACTGCAGCTGACAAGCCTGAAAGAAGCAGTGGGTCT
GGCCAAAATTAACTGCATAAAAAAATTATGTCCTCAAACGACCCGATGCTGTGCT
ATCTGATCGAGAGCCAGCGTAGTCGTATTGTTGAACACTTCATGAAGGATCTGAAG
CTGGGCGACAGCCTGACCCTGAACGAGATGAATAATATTAAAGAATGTTTTATGAA
AGAGAAGCGAATTAGCTTCGCCCAGAAAATCCACGGCGTGGGCTTTGAAGTTTTCA
GCTCCAGCCCGTTAACCAACCAGTGGATCAACGGTGAGATCAAAACCATGACCACA
AAAACCTATATAAACAGCATCAAGTTGCGTACCAACACCCTGGAAACCCGCGTAAC
GACCTCACGTGGTCTGAACATCATCAAAACCTGTCGGAGATGTCATGTTGCCGATG
AAAGTCTGATGCACGTGCTGCAGTGTTGTAGCTCAACAAAGGGCCTCCGGTACAGC
CGTCATCATAAAATTTGCGCAAAAGTGGCCAACAACTGGTGATGAACGGCTATGG
TGTGTTTTCGTGAAAAGAGCTATCCTGATCCTAACAATAGCGGTAGCTACCTGCGTC
CCGATATCATCGCCGTGAAGAACGGTCATGTGATTGTGCTGGATGTGACCGTAGTT
TACGAGGTGACCGGTGCAACCTTTATCAACGCCTACCAGACCAAAATCAACAAGTA

10

20

30

40

50

CAACGCAATTATGGTGCAGATCGAACAGATGTTTAATTGTGTGAATGGCGAGCTAC
ATGGCTTAGTCATTGGCTCCCGAGGCAGCATACACCACAGCCAGCTGCACATCTGG
CACCAGATGGGTTTTAGCAGTATTGAACTGAAATATGTGGCAATCGGCTGTATGGA
AGATTCTCTGCGCATTATGAGCACCTTCAGCAAGGCCATCACCTGATAA

39. R9 AdVa アミノ酸配列:

MDYKDDDDKGTAFNIPYPIGRVDCEVCHEDIGQMNLPIREHAVSVHNINKFNYLCQ
LCSKSYDTINSVKAHYVACRRQKNASSTTAVPTNVINNNQLAINTNQVISRNPLQCV
ECLMKQVDFYAKDTKALVTHMRTKHAAAYEESKKVATRRVAWSPDEDQILAELEVK
LKKIQKGQLLSRLVVEYNKCADKSKAPSRSKDAIRTRRQQHDYKLLLRSLQSQQPPV
GSEDSDDISSNNNNPLTTTHNVTPTPDSSNVLLIQKIRESVDSIVKITNLKLNTNM
LNAASAFINQNNMMDPLELSMRGIEEDVKAIRDKEQKPTRNVPSSSTTSRKPTRNAK
RLEKSKKYGYQHLYYNNKKKLVAEILDGETSGAKPPPMNLVEDYYRNIWSRSTIDD
SPVNNIKTVNSDSIFAPISRDEIKLALSNTKKDSAAGPDAVTIKEAKAIIDNLYVAYNI
WLGVGQIPEQLKLNKTILIPKGNLSDLSLLKNWRPITISSIILRVYNRLAYRMNKIFK
TNDKQVGFKPVNGCGINISWLHSLK HARLNKNSIYACLVDVSKAFDSVSHQSIVRA
LTMNGAPSLLVKLIMDQYTNVNTVITCSGSISNKINISSGVKQGDPLSSLLFNLVIDE
LFDVIKDQYGYTIDNIGTTNARCFADDLTLISSSRMGMNKLLELTTKFFKERGLNVNP
SKCMSIGMSKGYKGKSKIESEPLFSITDAQIPMLGYIDKTTRYLGVNFTSIGAIDAKR
IKKDLQDTLDKLEHLKLKAQCKMDLLRTYMIPRFMFQLIHTELYPKLLIKMDILIRKL
AKRILHLPISTSSEFFYLPFKEGGLQLTSLKEAVGLAKIKLHKKIMSSNDPMLCYLIES
QRSRIVEHFMKDLKLGDSLTLNEMNNIKECFMKEKRISFAQKIHGVGFVSSSPLTN
QWINGEIKTMTTKTYINSIKLRTNTLETRVTTSRGLNIIKTCRRCHVADES LMHV LQ
CCSSTKGLRYSRHHKICAKVAN KLVMNGYGVFREKSYPDPNNSGSYLRPDIIAVKNG
HVIVLDVTVVVEVTGATFINAYQTKINKYNAIMVQIEQMFNCVNGELHGLVIGSRGS
IHHSQLHIWHQMGSFSSIELKYVAIGCMEDSLRIMSTFSKAIT**

40. R8 HyMaA プラスミドORF配列:

ATGGATTACAAGGATGACGACGATAAGGGTACCGTGATTATACACCCGGGCACCGA
TGGTATTTGGTGTGTGTACCGAATGTGTGGAAATCCATAATAGCGGCAAGGATCTGA
AGCGACACCTGGCCAAACGCCATCCCTCGGTACCATCTCCGGCTACAAGTGCAAC
CTGTGTCCGTTTGTAGCGAGCGGCAACTGAGTGTTGGTACACATTTGCGTTACTG
TCGTGGTGTTAAGGAAGTTGTGAAAAGAGAATTGCGATGTGCATCCTGTTCTTTA
GCAGCGACACCTTCAGCGGCCTTCAGGTGCACATGCAGCGTAAGCACATCGCAGAG
TGGAATGATCAGCTGAAGGAAAAGACAGAGTTGCGCTGGACCGACCGGGAACCTCCG
TGAACCTGGCTGAAAAGGAGCTGACCACACCGTCTCTCCGGTACAACAAGATCTTCT
ACGCCGCTCTGGGCACCTCCAGGACCTACGATGCAGTGCGGAAAATTCGCTATAAT
GATCGATACAAGAGCGCCATCGCCGAAATGAGAAGCCAGATTGCAGATGCAGCAGC
AGCCGCCCCAAGAACGTGATGTGGAGCGTGGTCTCGTGAGCGCCACAGCGACCGTG
GAAAAGAGATGCTGCCTGTTGTTGAAACCAAAAGCGACATCCAGGTTAATAACGAT
ATCAAAAAAGACATTGAGCTGACCCCTAATAGCCGCCAGAAACAGACCAATCTGGC
CCTGGCCAGACCTGCCGTGATCGAGGTGGAAGAGGATCTGGGCAGACAGGACGTGA
AGCAGTACCTGGCGAGCCTGAGACAGGATGATTACACCAGCCCTGCTGAGCGTAGC
ATTTTTGCATACTGCCGCGAAGAAACAACTGGAGCGCCACGAAGCGGCAAGTGCT
GAAAATAAGTCGGACCACCAGAGGTCTGCGCCAGCCGAAAAAAGTGCGTCCATTTG
AGTTTCCGGAGGGCTTCAAACCCAACAGAAATATGAGGAAGTGGCGCAAATATCGT
TTTCTGCAGGAATGCTACCGTGAAAAACGTGCAGAAACAGTGAGCAAAATCCTGGA
CGGCACCTTCATTGACGAACCCGAGGAAGAAATTCGTCCTGAACTGGAGGAGGTTC
AGAGAATGTACATCGACCGTCTGGAAAAGCGTACACAGCTGGATAACCAAAAATC
GTTTCAGACAGATGAAGTGTTCTGCTGCGAGCTACGGCAGAATCACGATCGGTGA
AGTGAGAGATGCTCTGGGCGCAAGCAAAAAGGACAGCGCATCCGGTCCGGATGGGC
TGCTGCTGCAAGACGTGCGGAGGCTGGGACCTCTGCTGCTGTGCAATATCTTCAAC

10

20

30

40

50

ATGTGGTATCTGCACGGCATCCCGGTGGAAGAGAACAGATGTAGGACAATCCTGCT
GTATAAAAGCGGTGATCGACATCTGGCCAGCAACTACCGTCCGGTTACCATCGGCA
ACATGCTGAACCGCCTGTATGCAAAAATATGGGACAAAAGAATTAGAAAAAATGTG
CGTCTGCATGTTAGACAGAAAGCCTTCATCCCGGTTGATGGATGTTTTGAGAATGT
AAAGACCATTTCAGTGCGTTCTACAGAGCTATCGTAAGCGCAAATTAGAACACAACG
TGGTTTTTTATTGATCTGGCCAAAGCATTTGATACCGTGCTGCATGACAGCATCCGG
AAAGCCCTGTGGCGTAAGGGGGTGCCTAGCGGCGTGGTGAAGGTTGTTGATAGCCT
GTACGCCGGAGCCGTAACATCTATCAGCGTTGGCAAGACCAAACGAGAAGCATTT
GCATCAACAGCGGTGTCAAACAGGGCTGCCCTTTGTCACCGCTGCTGTTTAATCTG
ATTCTGGATGAGTTAGCAGAACGGATTGAAGCCACAGGCTGCGGCCTTGATCTTGA
CGGCCACGTTCTGTCTAGCATGGCCTTTGCCGACGACTATGTTCTGCTGGCTAAGG
ATAGCGTGGAATGAACGAACTGATTCGCGTTTGTCTACCTTCTTCAAAGAGAAG
GGCTTGTCTGTGAATCCTGGCAAATGTCAGAGCCTGCGCGTGCTGCCAGTAAAAGA
GAAAAAACGCAGCATGAAGGTACTGGTGCCTCCTCATCGTTGGTGGCGTATTAAGG
ACCAGGATGTGGATATTCCCAGCATGACCTACGATTCTCTCGGCAAGTATCTGGGG
GTTAGCATCGACCCACCGGTAAAATCGCCCTGCCTATTGAGGAATGGAAGAACTG
GATGACCAAACCTGAAAGAGTGCAAGCTGAAACCGGAACAGAAAGTGAAGATTCTGA
AAGAAGTGGTGTGTAGCAGAGTCAATTATGTTCTGCGTATGAGCGAATGTGGTATT
AGCGAACTGCGTAGCTGGACCAGATTTGTTAGAAATTGGGCAAAAACATCATCCA
TCTGCCTACGTGGTGCAGCAGCGATTGGATTCAATTCATTAAAGGTCTGGGTATTCT
CTGATGTGTCTAAGGGCATTGTTATCCAACGGATGAGAGCTAGTGAGAAGATGAGC
ACCAGCGAAGATGGCATCGTTCGTGTTGTGGGAGCACGCCTGGTTCAGAAGAATCG
TGTTCTGTGGGAAAAAGCCGGTTTTGAAGGCATTGAACTGAAAGCTGCACGTAGAC
ACTGCGAAGTTGAGAGACTGAACAACATTGGCAATATCACCAACGGAGTAGCCCTG
AAGACCATCGCAGCCGTGTCCTCTGTGAACCGATACTGGATGATTGAGGACAATCT
GAAAAGCGGTAAATAAAATCCTGGTGTGGAAGGCAATGGCAGGTGCTATCCCGACCA
AAATCAATCTGTCACGGGGCGTTGCCGACCAGACCCTTAAGAAATGTCGACGTTGT
GGTCTAACAGCGGAAACCGATGGTCATATTCTGGCTGGCTGCCACACCAGCTCTGA
TGCATATAGCAAACGACATAATATGCTGTGTGATAAACTGGCCAAAGAGCTGAAAC
TGAATGGCGGTCCGAATCGGCGTGTTTGGCGGGAACGGACCTGCTTTACCAGCACA
GGTCGTCGTTATCGTCCGGATATTATTGTAAAAGATGATAGCAAATCACAGTCAT
CGACATGACCTGCCCGTATGAAAAATCCGAAGGTCATCTGATCCAGTGTGAAAGCG
CGAAAGTGACCAAATATGAACCTCTGAAACTGGATAAGTATTGGACAAGAGAGCTG
GAAGGCGCCAATGGTATTGTGGCAGAAAAAGTTGAGCTGATGGGCCTGGCCATCGG
TGCAATCGGCACCATCATGCGTTCTACCCTGCGCAAACCTGTGTGAACTGAAGAGCG
GTAGAATTGTGAGACGTCTGCAGATGATTGCATGTAACAACAGCGCCCAGATCATC
AAGGGTCATTTGTGCGCGGCCACACGCCGTAATCTGAGATGATAA

10

20

30

41. R8 HyMaAアミノ酸配列:

MDYKDDDDKGTVIIHPGTDGIWCCTECVEIHNSGKDLKRHLAKRHPSVTISGYKCNL
CPFVSERQLSVGTHLRYCRGVKEVVKREFACASCSFSSDTFSGLQVHMQRKHIAEWN
DQLKEKTEFAWTDRELRELAEKELTTPSFRYNKIFYAALGTSRTYDAVRKIRYNDRY
KSAIAEMRSQIADAAAAAQERDVERGLVSAHSDRGKEMLPVVETKSDIQVNNDIKKD
IELTPNSRQKQTNLALARPAVIEVEEDLGRQDVKQYLASLRQDDYTSPAERSIFAYCR
EETNWSATKRQVLKISRTTRGLRQPKKVRPFEPFEGFKPNRNMWKWRKYRFLQECY
REKRAETVSKILDGTFIDEPEEEIRPELEEVRMYIDRLEKRTQLDTTKIVQTDEVFCL
QSYGRITIGEV RDALGASKKDSASGPDGLLLQDVRRLLGPLLLCNIFNMWYLGIPVE
ENRCRTILLYKSGDRHLASNYRPVTIGNMLNRLYAKIWDKRIRKNVRLHVRQKAFIP
VDGCFENVKTIQCVLQSYRKRKLEHNVV FIDLAKAFD TVLHDSIRKALWRKGVPSGV
VKVVD SLYAGAVTSISVGKTKTRSICINSGVKQGCP LSP LFN LILDELAERIEATGCG
LDLDGHVLSSMAFADDYVLLAKDSVEMNELIRVCSTFFKEKGLSVNPGKCQSLRVLP

40

50

VKEKKRSMKVLVRPHRWWRIKDQDVDIPSMTYDSLGLKYLGVSIDPTGKIALPIEEW
KNWMTKLKECKLKPEQKVILKEVVC SRVNYVLRMSECGISELRSWTRFVRNWAKN
IIHLPTWCSSDWIHSIKGLGIPDVSKGIVIQRMRASEKMSTSEDGIVRVVGARLVQKN
RVLWEKAGFEGIELKAARRHCEVERLNNIGNITNGVALKTIAAVSSVNRYWMIEDNL
KSGNKILVWKAMAGAIPTKINLSRGVADQTLKKCRRCGLTAETDGHILAGCHTSSDA
YSKRHNMLCDKLAKELKLNGGPNRRVWRERTCFTSTGRRYRPDIIVKDDSKITVIDM
TCPYEKSEGHLIQCESAKVTKYEPLKLDKYWTRELEGANGIVA EKVELMGLAIGAIGT
IMRSTLRKLCELKSGRIVRRLQMIACNNSAQIIKGHLSRATRRLR**

42. R2D CiInプラスミドORF配列

ATGGATTACAAGGATGACGACGATAAGGGTACCAATACCCTGGGAGAAAACGTGGA 10
CGCCTCAGCCACCACCTCCTCTGCAAACCCTCTGTCTCAGGAACCGGGCTGGTGCG
AGAGCTGTAGCAAACCTGTTCAAGTCCCAGCGTGGTCTGCGGGTACACCAGCGCAGC
AAGCACCCCGAACTGTACCACAGTCAGAATCAGCCGCTGCCACGTTCAAAAGCACG
TTGGAGCGACGAAGAAATGGTTATATTTGCCCGTGAAGAGATTGCTAATCGAAAAA
TCAGATTCAATTAATCAGCATCTGCACAAAGTGTTTCCGCACCGTACCCTGGAGAGC
ATCAAGGGCCTGAGGGGTAAAAACGTTAGATATGCTCGTATCATGGCCGACCTGGA
AGCCGAGATGACCAGCCAGCCAGAAGCAGCAACCTCACTGTGTACCGAGACAAGCG
AGAACCTGGCAAGCAGCAATGTTCTGCCTCAGACCCGTGGCTGGGCAGAGAACCTG
GTTGAAAATATCGATACTGCCCATCTGGCCAACCTGGGTCCGCTGAGCCAGTTTGA
GCCC GGGAACCGTCGTCCAGTACCAAAGAAGCCATTAACACCGAATATAATGATT 20
GGATCTCTAAGTGGCTGCCTAGCGGTGCCGCCACAGAGAAAGACGCGCCAACCCC
CCCAGCACGAAGCTGAATGCACGTGCAACCAGAAGACTGCAATACAGTAGAATCCA
GAATCTGTACAAACTGAACAGATCTGCCTGTGCACAAGAGGTTCTGTCCGGTGCT
GGAAAGTTCAGAGCGGAGAACTGAATTTGAAAGAGGTTTCAGCCGTTTTGGGAAAAG
ATGTTCCGGAAAGAAAGCGCAAAAGACAGACGGAAGCCTAAGCCACAGGTGAAGT
GCTGTGGGGTCTGATGGAACCGCTGACCATCGCAGAAGTTGGCAGCACACTGAAGT
CCACGACCCCTAGCGCACCAAGGTCCGGATAAACTAACACTGGATGGTGTTAAGCGT
ATCCCGATCGCGGAACCTGGTTAGCCACTACAATCTGTGGCTGTATGCAGGTTATCA
GCCTGAAGGCCTGCGGGAAGGCATCACCACTGATTCCGAAAATTAAGGCACAA
GAGATCCCGCGAAACTGCGTCCGATCACCGTGAGTAGTTTCATCTGCCGGATCTTT 30
CATCGTTGTCTGGCTCAGCGTATGGAAACCAGCCTGCCTCTGGGTGAACGTCAAAA
AGCATT CAGAAAGGTGGACGGCATTGTGTCATAATATCTGGTCCCTGCGGAGCCTGA
TCCACAACAGCAAAGACAATCTGAAAGAGCTGAACATTACCTTTCTGGATGTTTCGC
AAAGCCTTCGACAGCATCTCTCATAAGAGCCTGGGAATTGCCGCGGCAAGGCTGGG
CCTGCCTCCTCCGCTAATTACTTACATCAGCAACCTGTATCCGAATTGCAGCACCA
AACTTAAGGTGAATGGCAAAATTTCCAAACCCATAGAGGTCCGTCGTGGCGTGCGA
CAGGGTGATCCTCTGAGTCCTCTGCTGTTTAATGCCGTGATGGATTGGGCCCTGAG
CGAGCTGGACCCTAGAGTAGGTGTT CAGATCGGCGAACAGCGGATTAACCACCTGG
CCTTTGCAGATGATATTATTCTGGTTAGCAGTACCAAGATTGGCATGGTGAGCTCT
ATCAATACCCTGAGCCGTCATCTGGCCAAAAGTGGACTGGAGATCAGTGCCGGTAA 40
AGAAGGAAAATCGGCCAGCATGGCTATTGTTGTGATGGTAAAAAGAAGATGTGGA
CCGTTGATCCGCTGCCCAGATTTAAAGTGAATAGCCAAAAAATTCCGGCACTGAGC
ATCACCCAGCAGTACAAGTACCTGGGGATTAATATCGATGCACAGGGCGCCCGGAA
CGACGCCGCCAGAATTCTGACCGAGGGCCTAGCCGAACCTGAGCCGTGCTCCGTTAA
AACCGCAGCAGAGACTGTATCTGCTCCGCGTGATCTGTTACCAAGCTGCAGCAT
GGTCTGGTGCTCTCTAGCTGCGCCAAGCGTGCTCTGACCTACTTAGACAAGAGCGT
TCGTT CAGCCATCAGGAGATGGCTGACCCTGCCGAAAGATACACCGACAGCTTTTT
ATCATGCAAAAGCATGCGACGGAGGTCTGGGTATAACCCGGCTGGAACATACCAT
CCTATCCTAAAGCGTAACAGAATGATGAAGCTGACCTTAAGCGAAGATCCGGTTAT
CATGGAACCTCGTGAAGCTCACATACTTACCAATTTATTACACAAGTACAGCAACG 50

TGAAACTGCTGAACAGCTGGCCTGTGACCGATAAGGATAGCCTGGCACGGGCTGAG
GCAAGCATGCTGCACACCTCTGTGGATGGCCGGGGCCTGAGCAACTGCAGCGATGT
GCCTAGACAGAGCGATTGGGTTACCAATGGTGCTAGCCTGCTGAGCGGCAGAGATT
TTATTGGTGCAATCAAAGTTCGTGGTAATCTGCTGCCTACCAAAGTGAGCGCAGCT
CGTGGACGTCAACGTGAAATTACCTGCGATTGTTGTCTGCTCCTGAGAGTCTGGG
CCACATCCTACAGACATGTCCGCGGACCTGGGGCCCTCGCATTAGCCGGGCACGATA
GCCTGCTGAAGAGAGTTTCGCAACCAGGCCTGCCTTAAAAACTGGACCCCTATCATC
GAACCGAGCATTCTACAAACATTGGTTTACGGAGACCGGATCTGGTGCTGGCGAA
AGGCAATATCGCTTTCCTGGTGGACGCCACCGTGGTGGCCGATAATGCAAACATGC
AGCTGCAGCATGAAGCCAAAGTGGAATAATATAACAACAGCGACATCAAGGAATGG
ATCAAAGTTCATTGCCCGGGCGTCGACGAGGTTCTGTGTGACCAGCCTGACAGCTAA
TTGGCGGGGCTGTCTGTATGGCGGTAGCGCTAGCTTCTTAACCGAGGACCTGGGTC
TGCCGAAGGCGGAGCTCAGCCTGCTCAGCGCCAAGATTAACGAGAAAGGCTACTAC
CTGTGGTGTGCACATTATAGAGGCACCGCAAGACTGTGGAATCGTCCTCTGAGATC
CTGATAA

10

43. R2D CiInアミノ酸配列

MDYKDDDDKGTNTLGENVDASATTSSANPLSQEPGWCESSKLFKSQRGLRVHQRS
KHPELYHSQNQPLPRSKARWSDEEMVIFAREEIANRKIRFINQHLHKVFPHRTLESI
KGLRGKNVRYARIMADLEAEMTSQPEAATSLCTETSENLASSNVLPQTRGWAENLV
ENIDTAHLANLGPLSQFEPGKPSSTKEAINTFYNDWISKWLPSGAAHRERRANPPS
TKLNARATRRLQYSRIQNLYKLNRSACAQEVLSGAWKVQSGELNLKEVQPFWEKMF
RKESAKDRRKPKPTGEVLWGLMEPLTIAEVGSTLKSTTPSAPGPDKLTLDGVKRIPIA
ELVSHYNLWLYAGYQPEGLREGITTLPKIKGTRDPAKLRPITVSSFICRIFHRCLAQR
METSLPLGERQKAFRKVDGICHNIWSLSLIHNSKDNLKELNITFLDVRKAFFDSISHK
SLGIAAARLGLPPLITYISNLYPNCSTKLKVNGKISKPIEVRRGVRQGDPLSPLLFNA
VMDWALSELDPRVGVQIGEQRINHLAFADDIILVSSTKIGMVSSINTLSRHLAKSGLE
ISAGKEGKSASMAIVVDGKKKMWTVDPLPRFKVNSQKIPALSITQQYKYLGINIDAQ
GARNDAAARILTEGLAELSRAPLKPQQRLYLLRVHLLPKLQHGLVLSSCAKRALTYLD
KSVRSAIRRWLTLPKDTPTAFYHAKACDGGGLGITRLEHTIPIILKRNRMMKLTLSEDP
VIMELVKLTFTNLLHKYSNVKLLNSWPVTDKDSLARAEASMLHTSVDGRGLSNCS
DVPRQSDWVTNGASLLSGRDFIGAIVKVRGNLLPTKVSAARGRQREITCDDCRRPESL
GHILQTCPRTWGPRISRHDSLLKRVNRNQAACLKNWTPIIEPSIPTNIGLRRPDLVLAKG
NIAFLVDATVVADNANMQLQHEAKVEKYNNSDIKEWIKVHCPGVDEVVRTSLTAN
WRGCLYGGASASFLTEDLGLPKAELSLLSAKINEKGYLWCAHYRG TARLWNRPLRS**

20

30

44. R2A LiPoプラスミドORF配列

ATGGGGAGCGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTA
CAAGGATGACGATGACAAGAAGGGTGGAGGTAGCGGGGGCAGTGGAGGGCCCAAGA
AGAAGAGGAAGGTGATGTGCGTGACCGAGGGTGAATGCCTGCCGCCGACCGACAAC
TGCAACCCGAGCGTGGAGAACCAGCTGCCGTGCGTGACCGAGGGTTCGCTTCGAACG
TGTGGGCAGCCTGGTGACCGTGCGTCTGCCATTTTCGCAAGGTGGCCTGCGACCTGT
GCAGCAAGGAGTTCCTGACCTACAGCAAATTTGCTGTGCACCAGGCCAACTTCCAC
AACAGCGAAACCCAGGCTTGCTGCACCTATTGCGGCAAGAGCGATGGCAACCACCA
CAGCATTGCTTGCCATGTGCCAAAATGCCCGTGGCGTTCGCACCGTGACCTTCGCCG
CCAACCTGAGCAACTTTCTGTGCGACCTGTGCAACGATAGCTTTAAGACCAAAAGC
GGTCTGAGCCAGCACAAAGCGCCACAAACACCCGTGCAGCCGTAATGCCGAACGCAT
CCTGAGCCTGGGCGTGCGTACCCCAAGCGCCCGTCCACGTCAGGTGGTGTGGAGCG
AGGAAGAGACCCGCACCTGCGTGAGGTGGAAGTGGTGTACAGCGGTCAGAAGAAC
ATTAACGTGCTGTGCGCCGGTCACCTGCCAGGCAAGACCAAGCAACAAGTGAGCGA
CAAACGTGCGATCTGCACCGCATTCGTAGCAGCAACGTGCACGGTACCCCGACCA
CCCAGAGCCGTGGCGACCCGGTGGAGCAGGTGGAAGAGTACGAAGAGCTGGATTGG

40

50

GAGGGTATGCACCCGTTCCCGGACCCGGATAGCAAGTTTTGCAGCTATCTGGACCA
GCTGCGCGATCAGAAAGGCCTGACCGAGCCGGTGTGGCAGGAGATCGAAATTGTGG
CTCAGGAGTGGGTGGAAAACCTGGCCACGTGCAGAGCAGCTGGAACCACGAGCGT
ACCACCAAGCAGGTGCCAGAAAATAATACCCCGGCTCGTCGCCCGTTCAAGCGTCG
CCTGCACCGTGTGGAGCGTTACAAACGCTTTCAGCGTATGTATGACCTGCAGCGCA
AGCGTCTGGCCGAAGAGATCCTGGATGGCCGTGAAGCGGTGACCTGCAACCTGAAG
AAAGAAGAGATTAAAGACCACTACGATCAGGTGTATGGTGTGAGCAACGACCGTGT
GAGCCTGGATGATTGCCCGCGTCCGCCGGGTGCCAACAACACCGACCTGCTGAAGC
CGTTCACCCCGACCGAGGTGATGGATAGCCTGCAGGGTATGAAGAACGGTGCTCCG
GGCCCGGATAAAAATCACCTGCGCTTCTGCGAGAAGCGTCTGAAAAACGGTATTCA
CGTGAGCCTGGCCAACGTGTTCAACCTGTGGCAGTTTAGCGGCCGTATCCCGGAGT
GCATGAAAAGCAACCGCAGCGTGCTGATTCCGAAGGGCAAAAGCAACCTGCGCGAC
GTGCGTAACTGGCGCCCGATCACCATTAGCAGCATCGTGCTGCGTCTGTACACCCG
CATTCTGGCTCGTCGCCTGGAACGTGCCGTGCAGATCAATCCACGTCAGCGTGTT
TTGTGCCACAGGCGGGCTGCCGCGATAACATTTTTCTGCTGCAGAGCGCGATGCGT
CGCGCTAAGCGTAAAGGTACCCTGGCTCTGGGCCTGCTGGACCTGAGCAAGGCCTT
CGATACCGTGGGCCACAAACACCTGCTGACCAGCCTGGAGCGTTTTGCTGTGCACC
CGCACTTTGTGCGCATCGTGGAAGACATGTACAGCGTTGCAGCACCAGCTTTCGT
GTGGGCAGCCAGAGCACCCGTCCGATTGTGCTGATGCGCGGTGTGAAGCAGGGCGA
CCCGATGAGCCCGATCCTGTTTAACATCGCCCTGGACCCGCTGCTGCGTCAGCTGG
AAGAGGAAAGCCGTGGTTTCATGTTTCGTGAGGGTCAGGCCCCGGTGAGCAGCCTG
GCTTATGCCGACGATATGGCTCTGCTGGCCAAGGACCACGCTAGCCTGCAGAGCAT
GCTGGGTACCGTGGATAAATTCTGCAGCGGCAACGGTCTGGGCCTGAACATCGCGA
AGAGCGCCGGTCTGCTGATTGCGGCGCTAACAAAACCTTTACCGTGAACGACTGC
CCGAGCTGGCTGGTGAACGGCGAGACCCTGCCGATGATCGGCCCGGAACAGACCTA
CCGTTATCTGGGTGCTAGCATCTGCCCGTGGACCGGTATTAACAGCGGCCCGGTGA
AGCCGACCCTGGAGAAATGGATCGCCAACATTACCGAAAGCCCGCTGAAGCCACAC
CAGCGTGTGGACATTCTGTGCAAATACGCGCTGCCGCGTCTGTTCTATCAGCTGGA
GCTGGGTACCCTGAACTTTAAGGAACTGAAAGAACTGGATAGCATGGTGAAGCAGG
CGGTGAAGCGTTGGTGCCATCTGCCAGCTTGCAACCGCCGATGGTCTGCTGTACAGC
CGTCACCGTGATGGCGGTCTGGCCGTGGTGAAGCTGGAAAGCCTGGTGCCGTGCCT
GAAGATCAAAACCAACCTGCGTCTGGTGCACAGCACCGACCCGGTGATTAGCAGCC
TGGCCGAGAGCGATGGTCTGGTGGGTGCTATCGAAGGTATTGCTCAGAAAGCGGGT
CTGCCAATCCCAACCCAGACCAGCGTAGCGGTACCTACCACAGCAACTGGCGCGA
TATGGAGCGTCGCAGCTGGGAACGTCTGGCCCTGCACGGTCAGGGTGTGGAGCTGT
TCAAAGGTAGCCGTAGCGCGAACCACCTGGCTGCCACGTCCAGTGGGCATGAAGCCG
CACCCTGGGTGAAATGCCTGGCGATGCGTGCTAACGTGTATCCGACCAAGCGTGG
TCTGAGCCGCGGCAACCTGAGCAAGAACAAGACAGCGCCAAATGCCGTGGTTGCA
CCAGCATGCGTGAGACCCTGTGCCACCTGAGCGGCCAGTGCCCGAAGCTGAAAAGC
ATGCGTATCCGTCGCCACAACAATCTGCGAGCACCTGATTGCGGAAGCTAGCTT
CAAGGGTTGGAAAGTGCTGCAGGAGCCGACCCTGGTGACCGACAACGGCGAACGTC
GCCGTCCGGATCTGATCTTTCACCGTGACGATAAGGCGGTGGTGGTGGACGTGACC
GTGCGCTACGAGATTAGCAAGGATACCCTGCGTGAAGCCTACGCGAGCAAAGTGCG
CCGTTATGGCTGCCTGACCGAGCAGATCAAAGACCTGACCGGTGCGACCAGCGTGG
TGTTCCACGGCTTTCGATGGGTGCGCGTGGCGCCTGGTTCCAGAAAGCAGCGAC
GTGATGGCTGATCTGAACATCCGTAGCAAGTACTTCGAGGAATTTCTGTGCCGCCG
TACCATTCTGTATACCCTGGATCTGCTGTGGAAAAGCAACAACGAGCAGTATCTGG
AACGTCTGGCGCCGTAATAA

10

20

30

40

45. R2A LiPoアミノ酸配列

MGSDYKDHGDYKDHIDYKDDDDKKGGGSGGSGGPKKKRKVMCVTEGECLPPTD

50

NCNPSVENQLPCVTEGRFERVGS�VTVRLPFRKVACDLCSEFLTYSKFVHVQANFH
NSETQACCTYCGKSDGNHHSIACHVPKCPWRRRTVTFAANLSNFLCDLCNDSFKTKS
GLSQHKRHKHPCSRNAERILSLGVRTPSARPRQVVWSEEETRTLREVEVVYSGQKNI
NVLCAGHLP GKTSKQVSDKRRDLHRIRSSNVHGTPTTQSRGDPVEQVEEYEELDWE
GMHPFPDPDSKFCSYLDQLRDQKGLTEPVWQEIEIVAQEWVENLAHVQSSWNHER
TTKQVPENNTPARRPFKRRLHRVERYKRFQRM YDLQRKRLAEEILDGREAVTCNLK
KEEIKDHYDQVYGVSNDRVSLDDCPRPPGANNTDLLKPFTPTTEVMDSLQGMKNGAP
GPDKITLPFLQKRLKNGIHVSLANVFNLWQFSGRIPECMKSNRSVLIPKGKSNLRDV
RNWRPITISSIVLRLYTRILARRLERAVQINPRQRGFVPQAGCRDNIFLLQSAMRRAK
RKGTLALGLLDLSKAFDTVGHKHLTSLERFAVHPHFVRIVEDMYSGCSTSFVRVGSQ
STRPIVLMRGVKQGDPMSPILFNIALDPLLRLQLEESRGFMFREGQAPVSSLAYADD
MALLAKDHASLQSM LGTVDKFCSGNGLGLNIAKSAGLLIRGANKTFTVNDCPSWL
VNGETLPMIGPEQTYRYLGASICPWTGINS GPVKPTLEKWIANITESPLKPHQRVDILC
KYALPRLFYQLELGT LNFKELKELDSMVKQAVKRWCHLPACTADGLLYSRHRDGG
LAVVKLES LVPCLIKIKTNLRLVHSTDPVISSLAESDGLVG AIEGIAQKAGLPIPTPDQRS
GTYHSNWRDMERRSWERLALHGQGVELFKGSRSANHWLPRPVGMKPHHWVKCLA
MRANVYPTKRGLSRGNLSKNKDSAKCRGCTSMRETLCHLSGQCPKLKSMRIRRH
NKICEHLIAEASFKGWKVLQEPTLVTDNGERRRPDLIFHRDDKAVVVDVTVRYEISK
DTLREAYASKVRRYGC LTEQIKDLTGATSVVFHGFPMGARGAWFPESSDVMADLNIR
SKYFEEFLCRRTILYTLDLLWKSNNEQYLERLAP**

10

20

46. R2D DrMercプラスミドORF配列

ATGGATTACAAGGATGACGACGATAAGGGTACCATGCACCAGGTGCGGAACGCCGA
TGCCAATCTGCAGACCATGTTCCCCTGCAGAGAGTGTGAGCGCAGCTTTCGGACAA
AGATCGGACTGGGAGTGACACATGAGGCACAGGCACAAGGACGAGCTGGATACCGCC
CGGAGAAGGGTGGACGTGAAGGCCAGATGGAATGAGGAGGAGCTGTCCATGATGGC
CAGGAAGGAGCTGGAGCTGACCGCAAACGGAGAGCGCTTCATCAATAAGAAGCTGG
CCGAGATCTTTACAAACAGAAGCGTGGATGCCATCAAGAAGTGCAGACAGAGGGAC
AATTACAAGGCCAAGATCGAGCAGCTGCAGGGACAGGCAGCACTGATCTCCGAGGC
AAACGAGCCACCTACCACACAGAGGAGGCCATCTCTGAGCGAGCTGGAGGTGACCC
CAAGCTCCTCTCACTCCGTGCCAATCGCACCACCACCTATCCACTCTGACGATATCC
TGCTGCAGGAGCTGCAGGGCATGAGCCCTGTGGCCGTGAGAAGGTCCTGGAGAGTG
GAGGTGCTGCAGAGCATCATCGATAGGGCACACATCTCCGGCAAGGAGGCAACCCT
GCAGTGTCTGAGCAACTACCTGCTGGAGATCTTCCCTAACAGAAATGACAGGCCAA
GCTCCGCAACAGTGCCAGCACGCCGGCCTAGGAATAGAAGGATCTCTCGCCGGCAG
CAGTACGCCCCGCTGCATCAAGAGCCTGCTGGACGGAACCGATGAGTCCGCACTGCC
TAACCAGTCTATCATGGAGCCATATTGGAGGCAGGTCATGACCCAGCCATCCCCCT
CTCTGTGTTCCAACACAGTGCCTCGGAAGGGCAATATGCAGGAGGGCGTGTGGTCT
CCAATCACCAGCAGAGATCTGCAGGTGCACAAGGTGCCCCCTGACATCTAGCCCTGG
CCCAGACGGCATCACCTCCCAGACAGCCAGATCTATCCCAATCGGCATCATGCTGA
GGATCGTGAATCTGATCCTGTGGTGCGGCGATCTGCCAGTGCCTTTCGCGATGGCA
AGGACCATCTTTATCCCTAAGACAGTGAGAGCCAACAGGCCACAGGACTTCCGCCC
AATCAGCGTGCCCTCCATCGTGGTGCGCCAGCTGAATGCCATCCTGGCCTCTCGGC
TGACCGCAGCCGTGAGCTGGGACCCAGGCAGAGGGGCTTTCTGCCTACCGACGGA
TGCGCCGATAACGCCACAATCGTGGACCTGGTGCTGCGCGATCACCACAAGCGGTA
CGCCAGCTGTTATATCGCCACCCTGGATGTGAGCAAGGCCTTTGACTCCGTGGCCC
ACGATGCCGTGTTTAAACACCGTGACAGCCTACGGCGCCCCCAAGAGCTTCGTGGAC
TACGTGAGAAGGTGGTATTCCGGCGGGCGGCACATATTTCAATGGCGGGCGATTGGAG
ATCCGAGGAGTTTGTGCCAGCAAGGGGAGTGAAGCAGGGCGACCCACTGTCCCCCG
TGCTGTTTAACTGATCATCGATCGCCTGCTGCGGTCTCTGCCTAAGGACATCGGC
GTGCACGTGGGCAACGCCAAAGTGAATGCCTGTGCCTTCGCCGACGATCTGATGCT

30

40

50

GTTTGCCAGCACCCCTAAGGGCCTGCAGGAGCTGCTGAACACCACAGTGAAGTTCC
TGTCCTCTGTGGGCCTGACCCTGAATGCCGACAAGTGCTTTACAATCTCTATCAAG
GGCCAGCCAAAGCAGAAGGTGACCGTGGTGGAGCAGAGGACATTCTGTATCGGCAG
AGCCAGGGTGCAGCTGAAGCGCTCCGAGGAGTGGAAGTACCTGGGCGATCCACTTTA
CAGCAGATGGAAGGGCAAGGTATAACCCAAGCGAGGACATCGGGCCCTAAGCTGGAG
AGGCTGATGCAGTCCCCTCTGAAGCCACAGCAGAAGCTGTTTCGCACTGAGGACCGT
GCTGGTGCCTCAGCTGTACCACAAGCTGACACTGGGCAGCGTGGCACTGGGAGTGC
TGAGGAAGTGCGATAAGCTGGTGAGATCTTTCGCCAGGAAGCTGCTGGGACTGCCA
CTGGACGTGAGCGTGGCCTTTTATCACGCACCACACTCTTGTGGAGGACTGGGCAT
CCCAAGCGTGCCTGGATCGCACCCATGCTGAGAACCAAGAGGCTGGCCGGCATCA
ATTGGCCACACCTGGAGCAGTCCGAGGTGGCCTCTGCCTTTCTGAGCGAGGAGCTG
AGGAGGGCAAGAGATAGGGCAAAGGCAGGAGTGAACGAGCTGCTGAGCCAGCCCAA
GATCGACACCTACTGGGCGGATCGGCTGTATACATCCGTGGACGGCAATGGCCTGA
GAGAGGCAAGAAGATACGCACCACAGCACGGATGGGTGTCCCAGCCTACCCGGCTG
ATGTCTGGCAAGGCCTATCGCACAGGCATCCAGCTGCGGATCAACGCCCTGCCTAC
CCGCTCTCGGACCACAAGAGGCAGGCACGAGATGAACAGGCAGTGCAGGGCAGGAT
GTGATGCACCAAGCCACAATCACGTGCTGCAGAGATGCCACAGGACACACGGCTCT
CGCGTGAGCCGGCACAATGGCGTGGTGAGCTACCTGAAGAAGGGCCTGGAGACCAG
AGGCTACACAGTGTATTCCGAGCAGTCTCTGCACGGCCAGAACC GCCTGTATAAGC
CCGATATCGTGGCCTTCCGGCACGACTCCACCATCGTGGTGGATGCCAGGTGGTG
ACAGACGGACTGGACCTGGATAGGGCACACCAGTCTAAGGTGGAGATCTACAATCG
CCAGGACCTGCTGACCACACTGAGGTCCGTGTATAGGGCAAGGGAGAACATCGAGG
TGGTGTCTGCCACCCTGAATTGGAGAGGCATCTGGAGCTTTCAGTCCATCACCAGA
CTGAGGACACTGGGCATCCTGACCGCCGGCGACTCCAACGTGATCAGCTCCAGAGT
GGTGTCTGGCAGGGTGTACAGCTTCAAGACATTCATGTTTCACGCCGGCTTTCACC
GGGGCATGGCCTAG

10

20

47. R2D DrMerc アミノ酸配列

MDYKDDDDKGTMHQVRNADANLQTMFPCRECERSFR TKIGLGVHMRHRHKDELDT
ARRRVDVKARWNEEELSMMARKELELTANGERFINKKLAEIFTNRSVDAIKKCRQRD
NYKAKIEQLQGQAALISEANEPPTTQRRPSLSELEVT PSSSHSVPIAPPPHSDDILLQ
ELQGMSPVAVRRSWRVEVLQSIIDRAHISGKEATLQCLSNYLLEIFPNRND RPSSATV
PARRPRNRRISR RQQYARCIKSLLDGTDESALPNQSIMEPYWRQVMTQPSPSLCSNT
VPRKGNMQEGVWSPITSRDLQVHKVPLTSSPGPDGITSQTARSIPIGIMLRIVNLILW
CGDLPVPFRMARTIFIPKTVRANRPQDFRPISVPSIVVRQLNAILASRLTA AVSWDPR
QRGFLPTDGCADNATIVDLVLRDHHKRYASCIYATLDVSKAFDSVAHDAVFNTVTAY
GAPKSFVDYVRRWYSGGGTYFN GGDWRSEEFVPARGVKQGDPLSPVLFNLIIDRLR
SLPKDIGVHVGN AKVNACAFADDLMLFASTPKGLQELLNTTVKFLSSVGLTLNADKC
FTISIKGQPKQKVTVVEQRTFCIGRARVQLKRSEEWKYLGIHFTADGRARYNPSEDIG
PKLERLMQSPLKPQQKLFALRTVLVPQLYHKLT LGSVALGVLRKCDKLVRSFARKLL
GLPLDVSVAFYHAPHSCGGLGIPSVRWIAPMLRTKRLAGINWPHLEQSEVASAFLSE
ELRRARDRAKAGVNELLSQPKIDTYWADRLYTSVDGNGLREARRYAPQHGWVSQPT
RLMSGKAYRTGIQLRINALPTRSRTTRGRHEMN RQCRAGCDAPSHNHVLQRCHRTH
GSRVSRHNGVVSYLKKGLETRGYTVYSEQSLHGQNRVYKPDIVAFRHDSTIVVDAQV
VTDGLDLDRAHQSKVEIYNRQDLLTTLRSVYRARENI EVVSATLNWRGIWSFQSITR
LRTLGLITAGDSNVISSRVVSGRVYSFKTFMFHAGFHRGMA*

30

40

48. R2D DrMe プラスミドORF配列

ATGGGGAGCGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTA
CAAGGATGACGATGACAAGAAGGGTGGAGGTAGCGGGGGCAGTGGAGGGCCCAAGA
AGAAGAGGAAGGTGatgACCACCCGCCCGAGCGTTGACATCTTCCCAGAGGACCAAT
ACGAACCAAATGCCGCAGCGACCCTGAGCCGTGTGCCGTGTACCGTGTGCGGTCGT

50

AGCTTCAACAGCAAACGTGGTCTGGGTGTGCACATGCGCAGCCGTCACCCAGACGA
GCTGGATGAGGAACGTCGCCGTGTGGACATCAAGGCCCGTTGGAGCGATGAGGAAA
AATGGATGATGGCTCGCAAGGAGGTGGAACCTGACCGCCAACGGCTGCAAGCACATT
AACAAACAGCTGGCGGTGTACTTTGCTAACCGTAGCGTGGAAGCCATCAAGAACT
GCGCCAGCGTGGTGACTATAAGGAGAAAATCGAACAGATTCTGTGGTCAGAGCGCCC
TGGCTCCAGAGGTGGCTAACCTGACCATCCGCCGTGCGCCGAGCCGCAGCGAACAG
GATCACCAGGTGACCACCAGCGAAACCACCCCGATTACCCCGTTTGAGCAGAGCAA
CCGTGAAATCCTGCGTACCCTGCGCGGTTACAGCCCGGTGGAATGCCACAGCAAAT
GGCGTGCTCAGGAGCTGCAGACCATCATTGACCGCGCCACCTGGAGGGCAAGGAA
ACCACCCTGCAGTGCCTGAGCCTGTACCTGCTGGGTATCTTTCCGGCGCAGGGCGT
GCGTCACACCCTGACCCGCCCGCCACGTGCGCCGCGTAACCGTCGCGAGAGCCGTC
GCCAGCAGTATGCTGTGGTGCAGCGTAACTGGGACAAGCACAAAGGTCGCTGCATC
AAGAGCCTGCTGAACGGCACCGATGAAAGCGTGATGCCGAGCCAGGAGATCATGGT
GCCGTATTGGCGTGAAGTGATGACCCAGCCAAGCCCAAGCAGCTGCAGCGGTGAAG
TGATCCAGATGGACCACAGCCTGGAACGCGTGTTGGAGCGCGATCACCGAGCAGGAT
CTGCGTGCTAGCCGCGTGAGCCTGAGCAGCAGCCCAGGTCCAGACGGTATCACCCC
AAAGAGCGCCCGTGAGGTGCCAAGCGGTATCATGCTGCGCATTATGAACCTGATCC
TGTGGTGCGGCAACCTGCCGCACAGCATTCTGTCTGGCCCGCACCGTGTTTCATCCCG
AAGACCGTGACCGCGAAACGTCCGCAGGATTTTCGCCCAATCAGCGTGCCAAGCGT
GCTGGTGCGTCAGCTGAACGCGATTCTGGCTACCCGCTGAACAGCAGCATCAACT
GGGACCCGCGTCAGCGTGGTTTCTGCCAACCAGCGTTGCGCTGATAACGCTACC
ATTGTGGATCTGGTGCTGCGTCACAGCCACAAACACTTCCGCAGCTGCTACATCGC
CAACCTGGACGTGAGCAAGGCGTTTGATAGCCTGAGCCACGCTAGCATCTACGACA
CCCTGCGTGCTTATGGTGCCCCGAAGGGCTTTGTGGATTACGTGCAGAACACCTAT
GAGGGTGTTGGCACCAGCCTGAACGGTGACGGCTGGAGCAGCGAGGAATTCTGTGCC
AGCTCGTGGTGTGAAACAGGGCGACCCGCTGAGCCCGATCCTGTTCAACCTGGTTA
TGGATCGTCTGCTGCGCACCCCTGCCGAGCGAAATTGGTGCGAAAGTGGGCAACGCT
ATCACCACGCGCGCGGCTTTTCGCCGACGATCTGGTGCTGTTTGCGGAAACCCGTAT
GGGTCTGCAGGTGCTGCTGGACAAAACCTGGATTTCCTGAGCATTGTGGGCCTGA
AGCTGAACGCCGACAAATGCTTTACCGTGGGTATCAAGGGCCAGCCGAAGCAGAAA
TGCACCGTGCTGGAAGCGCAGAGCTTCTACGTGGGTAGCAGCGAAATTCCGAGCCT
GAAACGTACCGACGAGTGGAAGTATCTGGGTATCAACTTTACCGCTACCGGTCGCG
TGCGTTGCAACCCAGCGGAGGATATCGGCCCGAAACTGCAGCGTCTGACCAAGGCG
CCGCTGAAACCGCAGCAGCGTCTGTTCTGCTCTGCGCACCCGTGCTGATTCCGCAGCT
GTACCACAAGCTGGCCCTGGGTAGCGTGGCGATTGGCGTGCTGCGTAAGACCGACA
AACTGATCCGCTACTATGTGCGTCGCTGGCTGAACCTGCCGCTGGATGTGCCAATT
GCCTTCGTGCACGCTCCACCAAAGAGCGGTGGCCTGGGCATCCCGAGCCTGCGTTG
GGTGGCTCCAATGCTGCGTCTGCGTCGCTGAGCAACATCAAATGGCCGACCTGA
CCCAGAACGAGGTGGCTAGCAGCTTTCTGGAGGCTGAAAAGCAGCGTGCTCGTGAC
CGCCTGCTGGCTGAGCAGAACGAACTGCTGAGCCGTCCGGCGATTGAAAAGTACTG
GGCTAACAACTGTATCTGAGCGTGGATGGTAGCGGCCTGCGTGAGGGTGGTCACT
ACGGTCCACAGCACGGTTGGGTGAGCCAGCCGACCCGCTGCTGACCGGTAAAGAA
TATATGGACGGCATTCTGTCTGCGCATCAACGCTCTGCCAACCAGAGCCGTACCAC
CCGTGGTGCACACGAGCTGGAACGTGAGTCCGTGCTGGTTGCGATGCTCCAGAAA
CCACCAACCACATCATGCAGAAATGCTACCGTAGCCACGGTCGTCGCGTGGCTCGT
CACAACTGCGTGGTGAACCGTATTAAGCGCGGTCTGGAGGAACGTGGCTGCGTTGT
GATCGTGGAGCCGAGCCTGCAGTGCGAAAGCGGTCTGAACAAACCGGACCTGGTGG
CTCTGCGCCAGAACCACATCGACGTGATTGATACCCAGATCGTGACCGATGGCCAC
AGCATGGACGATGCCACACAGCGTAAGATTAACCGTTACGACCGTCCAGATATCCG
TACCGAGCTGCGTCGCCGTTTTCGAAGCTGCTGGCGACATTGAGTTTCACAGCGCGA

10

20

30

40

50

CCCTGAACTGGCGTGGTATTTGGAGCGGCCAGAGCGTGAAGCGCCTGATCGCTAAA
GGTCTGCTGAGCAAGTATGATAGCCACATCATTAGCGTGCAGGTTATGCGTGGCTC
CCTGGGTTGCTTCAAACAGTTCATGTATCTGAGCGGCTTTTCCCGTGATTGGACCT
AATAA

49. R2D DrMeアミノ酸配列

MGSDYKDHGDYKDHHDIDYKDDDDKKGGGSGGSGGPKKKRKVMTTRPSVDIFPEDQ
YEPNAAATLSRVPCTVCGRSFNSKRGLGVHMRSRHPDELDEERRRVDIKARWSDEE
KWMMARKEVELTANGCKHINKQLAVYFANRSVEAIKKLRQRGDYKEKIEQIRGQSA
LAPEVANLTIRRRPSRSEQDHQVTTSETTPITPFEQSNREILRTLRGYSPVECHSKWR
AQELQTIIDRAHLEGKETTLQCLSLYLLGIFPAQGV RHTLTRPPRRPRNRRESRRQQY
AVVQRNWDKHKGRCIKSLLNGTDESVMPSQEIMVPYWREVM TQPSPPSSCSGEVIQ
MDHSLERVWSAITEQDLRASRVSLSSSPGPDGITPKSAREVPSGIMLRIMNLILWCG
NLPHSIRLARTVFIPKTVTAKRPQDFRPISVPSVLVRQLNAILATRLNSSINWDPRQR
GFLPTDGCADNATIVDLVLRHSHKHFRSCYIANLDVSKAFDSLHASIYDTLRAYGAP
KGFVDYVQNTYEGGTS LN GDGWSSEEFVPARGVKQGDPLSPILFNLVMDRLLRTL
PSEIGAKVGNAITNAAAFADDLVLFAETRMGLQVLLDKTLDFLSIVGLKLNADKCF
VGIKGQPKQKCTVLEAQS FYVGSSEIPSLKRTDEWKYLGINFTATGRVRCNPAEDIGP
KLQRLTKAPLKPQQRLFALRTVLIPQLYHKLALGSVAIGVLRKTDKLIRYYVRRWLN
LPLDVPIAFVHAPPKSGGLGIPSLRWVAPMLRLRRLSNIKWPHLTQNEVASSFLEAE
KQRARDRLLAEQNELLSRPAIEKYWANKLYLSVDGSGLREGGHYGPQHGWVSQPTR
LLTGKEYMDGIRLRINALPTKSRTTRGRHELERQCRAGCDAPETTNHIMQKCYRSHG
RRVARHNCVVNRIRKRGLEERGCVVIVEPSLQCESGLNKPDLVALRQNHIDVIDTQIV
TDGHSMDDAHQRKINRYDRPDIRTELRRRFEAAGDIEFHSATLNWRGIWSGQSVKR
LIAKGLLSKYDSHIISVQVMRGS LGCFKQFM YLSGFSRDWT**

10

20

50. R2A GaAcプラスミドORF配列

ATGGATTACAAGGATGACGACGATAAGGGTACCCATATGGCCGGTCCTAGCACCAG
ATCTAGAAGGAAGGGCTGCCTGCCCCCTGTGGACTTCGAGAACCCTAAGAAGAGAA
CAAGACTCATGGCCAAGATGACAAATGGCAACCCTACCAGCCACGTGCCGTGCCCC
GCACCCTGCAGCAATGGCCACGAGGGCGGGCGGAAGAGTGGCTGTGATCGAGGGCCG
GCTTCCAGAACTGAGCGGCAGCAGAATTAGCGGGATT CAGCCTGCCCTACCTGTGG
AAACCAGCTTCGTGGGACAAAGCACCGGCAGAGGGGCGGACGGAGATGCCAACGCC
AACAGCAGCCCTCCCAGCCCCAACCTGGGCGGCAGCGTGGGCATGGTGCCTGCAGT
GCGCGATGGCACACCTCCTCTGGGCGGCCTGGCGAGGACCACAGCCGGGAATGCG
CCGGCGGCAACACCCCTCTGTGGATGCTGGAGGATAGCTTCCGGTGCGACTACTGC
CCTAGAGAGTTTCGGAACAAGAGCCGGAAGAAGCCTGCACATGCGGCGGGCTCACCT
GGCTGAGTACGACGGCGCTGGATTTTGCTGGGGCGAAAGACTGTCCGAGTTTCGCCG
CTACCCGGCTGTGGTCTACCGAGGAAACCAAGAAGCTGGCTGTCTTTTGTGAACGC
GGTGTGCCCAGCCCTTCTGAGTGCAGGGCCATCGCCGCCAGCCTCGGCGCCGGTAA
GACACACCACCAGGTGAGGAGCAAGTGCCGACTGGTGT T TGAGGCTATCAGACGGA
GAGAGCTGCTGGAAGTGGCCGCCGCCACCGAGAGACTGGAAAAGAGCGCACGGCGG
AAGCAGCCTGCCGTGCCCCCAGCACCCGTGCACGGCGTGAGAGGCGTGCTTAGAGG
CCTGCTGGGCAAGAGAGTGCCGCGGGAGGGTGGCACCACAGGAAGCACCTCCGCCC
GTATTGTGAGAAGGGACGACTGCCGGCAGGGCGCTGTGGCCTCCGCCAGCTTGAAC
CTGATCCGGCGGCTGGGCGCAAGGCTACAGGAAGAAGCGGACGTAGGAGAGTCCT
GGGCAGACCTCCAGAAATGGATGTGCGCCGGAGCGTGCGTATGAGAAGAATGCGGC
GGTTCCTGTACAGACTGGCTCGGCTGGGCTGGGCCAAGCTGGCCATGTTCTGTGTTG
GATGGCCAGATGGGCGCTTCTTGTCCTGTTCCACTGGTGGAAAGTGAGCGCCGTGTT
CAGAGAGCGGTGGAGCATCGTCAGAGCCTTCTTGGGCCTGGGACAGTTCGGCGGGCT
TCGGCACCGCCGACAACGCGGGCTTTGGCAAGTTGATAGACCCTGCGGAGGTGAGA
GCTCACCTGCAGAGCATCAAAAACAGAAGCAGCCCTGGCCCTGACGGCATCACCAA

30

40

50

GGTGGCCCTGTCTAAGTGGGACCCTGAGGGAATCAAGCTGGCCACATGTACAGCA
CCTGGCTGGTGTCTGCTGGCATCCCAAAGGTGTTCAAGAAATGCAGAACAACTG
ATCCCTAAGACCGGCGACGTGTCCCTGCATGGCGACGTGGGGCAGTGGCGGCCTAT
CACAATCGCCTCCCTGGTGTGAGACTGTACTCCCGCATCCTGACCGAGAGAATGA
CCGTGGCCTGCCCTAGCCATCCTAGACAAAGAGGATTTCATCGCCAGCCCCGGCTGT
TCTGAAAACCTGATGCTGCTTGAAGGCTGTATGAGCCTGAGCAAGGCCGGCAACGG
CTCCCTGGCTGTTGTGTTTGTGGACTTCGCCAAGGCTTTTCGACACCGTGAGCCACG
AGCACCTGCTGAGCGTTTTAGTTCAGAAGGGCCTGGACCAGCACATGGTGGAAGT
ATCAAGGACAGCTATGAGAACAGCGTGACCAAGGTGCATTGCCAGGAGGGATGCTC
TACCGACATCGCCATGAAAGTGGGCGTGAAACAGGGCGATAGTATGTCTCCACTGC
TGTTCAACCTGGCCCTAGACCCTCTGATCCAGCAGCTGGAAAGAGAAGGCAGGGGC
TTCCCCGTGAACGGAAAGTCTATCACCGCCATGGCCTTTGCCGATGACCTGGCCAT
CGTGAGCGACAGCTGGGAGGGCATGAGAGCCAATCTGGACATCCTGGTGGACTTCT
GTGAGCTGACCGGAATGCGGACCCAGCCAGCAAGTGCCACGGTTTCCTGATTGAA
AAGAGCGGCTCCCGTAGCTACAAGGTGAACAGATGCGAGCCTTGGCTGCTGAACGA
TACCGCCCTGCACATGGTGGGCCCCAAGGAGTCCATCAAGTACCTGGGGGTGCAAG
TGAATCCCTGGACAGGCATCTTCGCCGAAGATACCGTGGCTAAGCTGAGACAGTGG
GTGGTGGCCATCAGCAAGACACCTCTGAGACCTCTGGATAAGGTTAGCCTGCTGTG
TCAGTTCGCCGTGCCTAGGGTGATCTTCGTGGCCGACCACTGTATGCTGAGCGCCA
AAGCCCTGACAGAGATGGACAGAAGCATCCGGCAGGCCGTAAAGAGATGGCTGCAC
CTAGCTCGGTGTACCACCAACGGCCTGCTCTACAGCCGGAAGTCATCCGGCGGCCT
GGGCATCCCTAAACTGAGCATGATCGTCCCTGCGATGCAGGCCAGACGGCTGCTGG
GGCTGAGCCGCTCTAAAGACGAGACTGTAAGATGGATGTTTCTGGAAACCACAGAT
CACGTGGCCTTCGAGAGAGCCTGGCTTCGGGGCCGGCGGAAGCCCCGACGAGGTGCC
AGAGCTGGGCCCTGATCTGGTGGAAAGGATCTCCAGCCGAAGGCAATGCCGATCCCG
TGTCTACAGTGAGACCCAGAAAGAGAATCGTGCCATGTGACTGGAGACAAGTGGAG
TTCGACAGATGGGCCGGCCAACTGGTTCAGGGAAAGGGAATCAGAACATTTCGAGGC
CGATAAGATCTCTAATTGCTGGCTGTATGACTACCCTCCTAACAAAGCTCAAGCCTG
GCGATTTACACGCCGCCGTGCGAGCTGAGAGCCAACGTGTACCCACAAAGAGAGCTC
GCCGGCAGAGGCAGAACCGACACAATCGACGTGTGCTGCCGGCACTGCGGGCAGGC
TCCTGAAACCTGTTGGCACATCCTGGCCCTGTGCCCAAGGTGAAGCGGTGCAGAA
TCCAGAGACACCACAAGGTCTGCCAGGTGCTGGTGGCCGAGGCCGAGAGACACGGC
TGGGAAGTGGAAAGAGAGAAACGGTGGATGCTGCCCAGCGGCGAATGCGTGGCCCC
CGATCTGATCTGCTGGCTGGACGAACTGGCACTGATCGTGGACGTAACCGTTAGAT
ACGAGTTCGACGAGGAATCCCTGGAAAGAGCCAGAATCGAGAAAGAATGCAAGTAC
CGGCCCTGATCCCCGTGATCAGAGCCTCTCGTGTGCAGACCAAGAAAGTCACAGT
TTACGGCTTTCCTCTGGGCGCCCGTGGTAAGTGGCCTGCCAAAAACGAGCTGCTGC
TGGCCGACCTGGGCTGTCAAAGGCCAGGACCAGATCTTTCGCCAAACTTCTGTCC
CGGAGAGTGCTTCTGCACAGCCTGGATGTCATGAGAACCTTCATGCGCTAG

10

20

30

40

50

51. R2A GaAc アミノ酸配列

MDYKDDDDKGTHMAGPSTRSRKGLPPVDFENPKKRTLMAKMTNGNPTSHVPC
PAPCSNGHEGGGRVAVIEGRLPELSGSRSIGIQPALPVETSFVGQSTGRGADGDANAN
SSPPSPNLGGSVGMVPAVRDGTPLLRPGEDHSRECAGGNTPLWMLSDSFRCDYCP
REFGTRAGRSLHMRRAHLAEYDGAGFCWGERLSEFAATRLWSTEETKKLAVFCERG
VPSPSECRAIAASLGAGKTHHQVRSKRLVFIAIRRELLEVAATERLEKSARRKQP
AVPPAPVHGVVRGLRLLGKRVPREGGTTGSTSARIVRRDDCRQGAVASASLNLIRR
LGRKATGRSGRRRLGRPPRMDVRRSVRMRRMRRFLYRLARLGWAKLAMFVLDGQ
MGASCPVPLVEVSAVFRERWSIVRAFLGLGQFGGFGTADNAGFGKLIDPAEVRAHLQ
SIKNRSSPGPDGITKVALSKWDPEGIKLAHMYSTWLVSAGIPKVFKKCRITLIPKTGD
VSLHGDVGVQWRPITIASLVLRLYSRILTERM TVACPSHPRQRGFIASPGCSENLMLE

GCM SLSKAGNGSLAVVFVDFAKAFDTVSHEHLLSVLVQKGLDQHMVELIKDSYENS
VTKVHCQEGCSTDIAMKVGKQGDMSPLLFNLALDPLIQQLEREGRGFPVNGKSIT
AMAFADDLAIVSDSWEGMRANLDILVDFCELTGMRTQPSKCHGFLIEKSGSRSYKVN
RCEPWLLNDTALHMGPKESIKYLVGVQVNPWTGIFAEDTVAKLRQWVVAISKTPLR
PLDKVSLLCQFAVPRVIFVADHCMLSAKALTEMDRSIRQAVKRWLHLARCTTNGLL
YSRKSSGGLGIPKLSMIVPAMQARRLLGLSRKDETVRWMFLETTDHFVAFERAWLR
AGGSPDEVPELGPDLVEGSPAEGNADPVSTVRPRKRIVPCDWRQVEFDRWAGQLVQ
GKGIRTFEADKISNCWLYDYPPNKLKPGDFTA AVQLRANVYPTRELAGRGRTDTIDV
CCRHCGEAPETCWHILALCPKVKRCRIQRHHKVCQVLVAEAERHGWEEVEREKRWM
LPSGECVAPDLICWLDELALIVDVTVRYEFDEESLERARIEKECKYRPLIPVIRASRVQ
TKKVTVYGFPLGARGKWPAKNELLADLGLSKARTRSFAKLLSRRVLLHSLDVMRTF
MR*

10

52. R2A PuPuプラスミドORF配列

ATGGATTACAAGGATGACGACGATAAGGGTACCATGAACATCCCCACCCCTGTGGG
CATCGAGGTGCTGCTGCCTGACCCTAATCCTAACTGCCCTCTGTGCGGCGTCCAGG
TGGGCAAGGTGGGAATGCTGGGCAAACACTTCGCCAGCCGCCACGCCAGAGTGGCC
GTGTCTTATAAGTGTCGGCGCTGTGGCAGGACAAACAGCAACTCCCATTCTATCAG
CTGTACAGTGCCTAAGTGCAAGGGACTCGGCCTGGAAGAGACAGGAAGAAAGGAAG
GCGAATTTGGCTGCGAGCTGTGTTTCAGGACGGTTCAGAACC GGCGTGGGGCTGACC
CAACACCGGCGGCTGGCTCACCTGTGGAATGGAACATGGAAAAGGTTGGCTCTAC
CCAGCCTAGAGGAGGCCTGAAGGGCACAAAGCTGTGGTCCGTGGACGAGCTGAAAA
GCCTGATCGAGCTATCCGTGCGGTTTCGCTGGCGAACCCAATCATAAGATCAACAAG
CTGATCGCTGATGAACTTGGTAGCGGCAAGACCGCCGAGCAGGTCAGAAGCAAGCG
CCGGCTGCTGCAAAAGGAGATGGCCAGCAACGGCCCCCCCCATCGCTGCCGCTGAAG
AGAGCGGCACCTCTCGGGCCAGGCCGATCCGATCGTGCCCCAGAGCCCTCCTAAC
GTGGCCAAGACACGGCTGATGGAGGCCATCAACACCGAGGGCGGCACCGAGGGCGA
GAACGAGCGGAGAGCCATCGCCGACCTGATCAGAGACGTGGACCAGAACCCCGGCC
TGATCGAGTCTAGCGCCCTGGAGCTGATCCAGAGACTGGGAAAGGGCGCCACCAGA
AATAGAGCCGAAAACGGCAGAGAGATCACCAGAGCTAAGAAAGGATGGGAAAGACG
GCTGAGCCAGAAGCGCAGAGAATTCAGAGTGCACCAGAGACTCTACGAGAAGGACC
CAACCAAGCTGGCAGCGCTGGTGCTGGACGGCGATGCCTGTGCTGAGTGCACCATC
CCTGTGACCCAAATCGACGAAGAATACAGACGGCGCTGGGAGGCCGGCGGCGAGTT
CCACGGGCTGGGCAACTTCAGATCTGCCGGAGGCGCCGATAATAGCGAGTTTCA
ACCCCATCCTGGCCACCGAGGTGTCCGAGAACCTGATCAAGATGAAAAACGGCTCT
GCCCCCTGGCCCTGATGGCATCTCCAAGAAGGCCCTGCTTCAATGGGACCCTGCCGG
CGAACAACTGGCTAGACTCTACACTGCCTGGCTGGTGACGGCGTGATCCACGGG
CCTTCAAGGAATGTAGAACCAAGCTGCTGCCAAGAGCGGTGACGCGGATGAGCTG
AAGGACGTCAACGGCTGGAGACCTATCACCATCGGCTCCCTAGTGCTGAGACTTTT
CAGCAGAATCCTGACAATGCGGCTGACCAGAGCCTGCACAATGAACCCTAGACAGA
GAGGTTTTCTGGCCTCTAGCAATGGTTGCGCCGACAATCTGATGGTGCTGAATGGC
GTGATCAAGCGGAGCCGGGAAGAACGGGCTCCTCTGGCCGTCGTGTTCTGTGGACTT
CGCCAAGGCCTTTGATAGCATCTCACACGAGCACTTTCTGCGGGTGCTGGAGCAGC
GGCAGCTGGATAGACATGTGATCGAACTGATTAAGAACAGCTACGTGGACTGCGTG
ACAAGAGTGGGATCTGGCGGCGAACGCACCAGCCCCATCAACATGAAAGTGGGAGT
CAAACAGGGCGACCCAATGAGCCCCCTGCTGTTCAACCTGGCTATGGACCCTTTGA
TCCAAGCCCTGGAGGACGGCCAGAGCGGCCTGTCTTGGGATGGCAGATCCATCAGC
ACCCTGGCCTTCGCCGACGACTTAGTGCTGCTGAGCAGCAGCGTGCGGGGCATGAG
AAGAAACCTGGAGATCCTGGAAGAATTCTGCCAGCTGACCGGCCTGAAGGTGCAGC
CTAAGAAGTGCCACGGCTTCTACTACGACAAGGGCGTGATCAACGACTGTGAGCCT
TGGACCATCGGCGGCACCGAGCATCCACCTGGTGGGCCCTGGCGAGAGCGTGAAGTA

20

30

40

50

CCTGGGAGTTCAGGTGGGCCCCAGAAGAGGCATCACAATGCCTGATCTCTCGCCAC
AGCTGCAGGAGTGGATCGTGCGCATCAATAAGGCACCACTGAAGCCTACGCAGAGG
GTGAAAGTGCTGAATAGCTTCGCCCTGCCTCGACTAATCTACCAGGCTGACCACTG
CGACGTGCAGGGGCACCGCCCTGGCCACACTGGACGGAATGGTGAGAAAGGCCGTGA
AAACATGGCTGCACCTGGCCCCCTAGCACTTGTAAACGGCCTGCTGTACAGCAGAAAT
AGAGACAGCGGCCTGGGCATCGTGAGACTGGAAAGACTGATCCCTAGCCTGCAGGT
CAAGAGAATCCACAGACTGTCCCGGAGCCAGGACCTGATGACCCGGTTAATTACCG
TGAAAACCGTCACCCAGTCTGAGTGGCAGCGCAAGTGGCAAATGGTCGGCGGCGAT
CCTGACAACAGTCCAGTGTTGGAAGTCGAGGAGGGATTCCCCGTGGACGGCAGAAAG
CCTGGTTGTGACCGATTGGAGGCGGGGAGAGAACGAGGCCTGGCTGGCCCTCCAGG
TGCAGGGCATGGGCGCTGGTGAGTTTAAGGGCGATAAAATCAGCAACAGCTGGCTG
TTCGACCCCACTAAGGTGGGCTTCAGACAGCGGCACTACATCGCCGCCCTGGCCCT
TCGGACCAACACCTACCCACAGGCGAGTTCCTGGCACGCGGCAGAGATAAGGCCG
GCGCCGCATGCAGACGGTGCCAGGCCCGGCTGGAGACATGCAGCCACATACTCGGG
CAGTGCCCCTTCGTGAAAGGCAGCAGAATCAGACGGCACCAACAGAATCTGCGAACT
GCTGGCCAACGAAGCCGAGAGAGTGGGCTGGAAGGCTACCAGAGAGTTCAGAGTGA
TTACACCTGAGGGCGGAGTACGGATTCCAGACCTGGTGTGCACCAAGGGCGATAAG
GCCCTGGTGCTGGACGTGACAATTAGATTCGAGATGGCCCCCTGACACCTGTGCG
CGCTGCCGCCGAAAAGGTGAACCACTACCAGCCTGTGCGCTCTCAGATTGCCAGCA
TGGTGGGCGCTAAAGAAGTGAAAGTCATGGGATTCCCTATCGGGCGCCAGAGGAAAG
TGGCCTTCTTGCAACAACAAAGTGCTGACAGAGCTGGGCGTTACCGTGGGCAGACG
GGCCACCTTTGCCAGACTGGTGAGTAGACGCTCCTTGCTGTATAGCTTGGATATCC
TGCGGGGCTTTCGTGTGCGAGCCCGCCCTGTAG

10

20

30

40

53. R2A PuPu アミノ酸配列

MDYKDDDDKGTMTNIPTPVGIEVLLPDPNPNCPLCGVQVGKVGMLGKHFA SRHARV
AVSYKCRRCGR TNSNSHSISCHVPKCKGLGLEETGRKEGEFGCELC SGRFRTGVGLT
QHRRLAHPVEWNMEKVGSTQPRGGLKGTKLWSVDELKSLIELSVRFAGEPNHKINK
LIADELGSGKTAEQVRSKRRL LQKEMASNGPPIAAAEESGTSRAQADPIVPQSPNV
AKTRLMEAINTEGGTEGENERRAIADLIRDVDQNPGLIESSALELIQRLGKGATRNR
ENGREITRAKKGWERRLSQKRREFRVHQRLYEKDP TKLAALVLDGDACAECTIPVTQ
IDEEYRRRWEAGGEFHGLGNFRSAGGADNSEFHNPI LATEVSENLIKMKNGSAPGPD
GISKKALLQWDPAGEQLARLYTAWLVHGVIPRAFKE CRTKLLPKSGDADELKDVNG
WRPITIGSLVLRLFSRILTMRLTRACTMNPQRGFL ASSNGCADNLMVLNGVIKRSR
EERAPLAVVFVDFAKAFDSISHEHFLRVLEQRQLDRH VIELIKNSYVDCVTRVGS GGE
RTSPINMKVG VKQGDPMSPLLFNLAMDPLIQALEDGQ SGLSWDGRSISTLAFADDLV
LLSSSVRGMRRNLEILEEFCQLTGLKVQPKKCHGF YYDKGVINDCEPWTIGGTSIHL
VGPGESVKYLGVQVGPRRGITMPDLSPQLQEWIVR INKAPLKPTQRVKVLNSFALPR
LIYQADHCDVQGTALATLDGMVRKAVKTWLHLAPST CNGLLYSRNRDSGLGIVRLE
RLIPSLQVKRIHRLSRSDLMTRLITVKTVTQSEWQR KWQMVG GDDPNSPVLEVEE
GFPVDGRSLV VTDWRRGENEAWLALQVQGMGAGEF KGDKISNSWLFDP TKVGFRQ
RHYIAALALRTNTYPTGEFLARGRDKAGAACRRCQ ARLETCSHILGQC PFVKGSRI
RHHRICELLANEAE RVGWKATREFRVITPEGGVRI PDLVCTKGDKALVLDVTIRFEMA
PDTLLRAAAEKVNHYQPVASQIASMVGAK EVKVMGFPIGARGKWPSCNNKVLTELG
VTVGRRATFARLVSRRLLYSLDILRAFVCEPAL*

54. TriCan プラスミドORF配列

ATGGATTACAAGGATGACGACGATAAGGGTACCCATATGAGTCAGAAGAGAAGACC
CGAAAAAGCCGTCCCCGATGAAGGAGCCACCGCCCCACGATGTGCCCCAGCCAGACA
AGAGTAAATGTAGCGTGTGCGGCGAGACATTCAAGGGGCCAGCCAGCGTGACCATG
CACATGGTGAAGAAGCACCCCGTGGAGTTTAATGAGCTGAAGATGGCCAAGAAGCC
CGTGCCTAAGAAGGTGAGATGGTCCGAGGAGGAGATCTTCCAGCTGGCCAGGACAG

50

AGGCAGAGCTGACCCTGCAGGGCGTGCGCTTCATCAACGTGGAGCTGCAGAAGATC
TTTCCTGCCAGAGAGATCGAGGGCATCAAGGGCCAGAGGAAGCTGGCCAAGTACAA
GGAGCTGGTGAAGGACCAGCTGGATGAGATCGGCAGGGCCCCAAATCCCCCTGAGC
AGGAGATCGGCGAGGACGTGCCAAGCCCCCTTCAAGGCATGGCTGGAGCTGCTGCTG
GCCCTGCCAAAGACCCCCAATGATTTTCTGGAGCACAAAGCTGGACAACATCATCGT
GCAGGGCCCTGAAGGAGGACGTGAACTCTGATCAGGTGTTCAACGATCTGAATAGCT
ACCTGAAGCTGATCCTGGAGCCATCTGGAAGGGCAAAGAGCGTGCCCGGAGAGATC
ATCCACGGCGACCCTTCTGGCAGCGCCAAGACCTCCGTGACAAAGGCCCTAAGCC
AGCCACAGTGAGCTCCTCTCGGAAGAAGCGGAGAGATGCCGAGTTTGCCAGAATCC
AGCGGCTGTATAGGAAGAACCGCACATCTTGATCAATACCATCCTGGACGGCAAC
ACCAGAGAGCACGAGGGCCCCCAAGAATATGGAGGGCTTCTGGCGCGAGATCTTTGA
GCGGGAGAGCCCTGACGATCCAGACGATCCCGATATCTTCCTGGAGGGAGGAGGCCT
CCGACATCTGGAAGTACATCTCTTTTTATGAGATGTGCAACCTGTATCCACCACCTT
CCACAGCACCAGGACCTGACGGCTTCAGCTCCAAGGATCTGAGGCGCATGACCCCC
AGGGTGCTGAACAAGATCCTGAATCTGCTGCTGCACCTGCGCGATCTGCCTCAGAT
CCTGAAGAGCCACCGGACAGTGCTGATCCCAAAGACAGATCTGCCTACCAAGCCAG
GCGACTTCCGCCCTATCACCATCAGCAACATCCTGGTGCGCCACCTGAACAAGATC
CTGGCCAATCGGGTGTCCACCTGATCCCCATCAATGAGCGGCAGAAAGGCCTTTCT
GCCTATCGATGGCTGCGCCGAGAACATCTTCACACTGGACTTTATCCTGCACCACG
CCAGAACCAAGATCAAGTCCCTGTCTATGGCCATCCTGGACATCAGCAAGGCCTTC
GATAGCGTGTCCCACTCCATCTTTCGGGCCCTGAGAGAGGGCCAGGTGCCCAAT
CGGCTTCATCAAGTTTATCGAGAATTGCTACGGCGGCTGTTTCACCAAGCTGTTTT
GTGGCGGCGTGAAGTATCCCTCTGAGGTGAGCATGAATAGGGGCGTGAAGCAGGGC
GATCCCCTGTCTCCCGTGCTGTTCAACCTGGTCATCGACGGCCTGATCCGCCAGAT
CCCAAGCGCCCTGGGCTTTAACGTGTCTGACCAGGTGAAGGTGAGCTGCATCGCCT
ACGCCGACGATCTGATCCTGATCGCCACCACACGGGCGGCCTGAAGACACTGCTG
GATCTGACCAATAGCTATCTGGCCAAGAGAGGCCTGTCCCTGAACCCCGACAAGTG
TTCCGCCCTGTCTATCGTGGCCAGCGGCAAGCAGAAAGCTGGTGTACATCGCCTCTA
GCGAGCACTTCGATCTGGCCGGCCAGAAAGATGAGAAACCTGAATGTGGGGCGACTCC
TGGAGGTATCTGGGCATCCAGTTTTCTCACCTGGGAAGGGGCAGAGAAGGTGACACC
AGATCTGACCTGTCTGATCAATAGACTGCAGAAGGCCCTCTGAAGCTGCAGCAGA
AGCTGTACGCCCTGAGAATCTATCTGATCCCCAGGCTGATCCACGGCCTGACACTG
TCCAAGACCAACCTGGGCGAGCTGAAGACCCTGGACAAGCTGATCCGCAAGTACAT
CCGGGCCTGGCTGCACCTGCCAGACGATACACCCATGGGCTACTTCTATACCCCTC
TGAAGGCAGGAGGCCTGGGCCTGCCAAGCCTGCGGCTGGTCATCCTGAACAATCGC
CTGGAGAGAATCCTGAGAATGAAGGCCTCCAGGATATCATCGTGAGGACAATCGC
CGAGTCTGAGACACTGGGCGTGGAGATCCGCAAGCTGCACGACCTGCTGTCCATCG
ATGGCACAAATCCTGGACACCTCCGTGAAGATCCAATCTTTTTGGGCGGAGAGACTG
TACTCCTCTTATGACGGCAAGTGCTGTGCAACTCTGCCAATTTCCACCCGGCAA
TAAGTGGATCGGCGAGGATAGCCTGAACCAGCGGTCCACATCTTTGCCGACTGCC
TGAAGCTGAGAATCAACGCCCTGCCTACACGGTCCAGAACCAGGAGGGGCGCCCT
CTGAAGGATAAGCCATGCAGAGCCGGCTGTAGGAATTCTGACGGCGTGAAAGTGAT
CGAGACACTGAACCACATCACCCAGGTGTGCGAGAGGACACACGGAGCAAGGGTGA
AGAGGCACGACAGGCTGGTGGATTTTCGAGTGAAGGGCCTGCAGCGGCCCCACAGG
GTGGTGCTGAAGGAGCCCCACTACAAGACCGTGAATGGCGTGAGGAAGCCCGACAT
CGTGATCAAGATCCCTGACCACACATATATCTGCGATTTCCAGGTGGTGAGCGACA
CCTCCTGTCTGGAGCTGGAGTTTCGCAAGAAGGCCCTGAAGTACGCCGAGGATAAG
GGCCTGTGCGACCAGCTGACCAGGGATCACCCAGGAGAGCTGAGCTTCACCGCCAT
CACCTTCAACACCCGCGGCCTGATGCCAAGAGCTCCGTGACAGCCCTGAGGAAGC
TGGGAATGCCTCCACGGTCCATCATGACCCTGCAGAAGATCTGTATGGAGGGGTCA

10

20

30

40

50

CTGGAAATCTGGAGAATCTTCAATCAGACTACTGCTATGGCACGCAACTAG

55. TriCan アミノ酸配列

MDYKDDDDKGTHMSQKRRPEKAVPDEGATAHDVAQPDKSKCSVCGETFKGPASVT
MHMVKKHPVEFNEFKMAKKPVPPKVRWSEEEIFQLARTEAELTLQGVRFINVELQK
IFPAREIEGIKGQRKLAKYKELVKDQLDEIGRAPNPPEQEIGEDVPSPFKAWLELLLA
LPKTPNDFLEHKLDNIIVQALKEDVNSDQVFNDLNSYLKLILEPSGRAKSVPGEIIHG
DPSGSAKTSVTKAPKPATVSSSRKKRRDAEFARIQRLYRKNRTSCINTILDGNTREHE
APKNMEGFWRERESPDDDDPDIFLEEEASDIWKYISFYEMCNLYPPPSTAPGP
DGFSSKDLRRMTPRVLNKILNLLLHLRDLPQILKSHRTVLIPKTDLPKPGDFRPITI
SNILVRHLNKLILANRVSHLIPINERQKAFLPIDGCAENIFTLDLFIHHARTKIKSLSMA
ILDISKAFDSVSHHSIFRALREARCPIGFIKFIENCYGGCFTKLFCGGVKYPSEVSMNR
GVKQGDPLSPVLFNLVIDGLIRQIPSA LGFNVSDQVKVSCIAYADDLILIIATTRAGLK
TLLDLTNSYLA KRGLSLNPDKCSALSIVASGKQKLVYIASSEHFDLAGQKMRNLNVG
DSWRYLGIQFSLHGRAEKVTPDLTCLINRLQKAPLKLQKLYALRIYLIPRLIHGLTL
SKTNLGEKLTLDKLRKYIRAWLHLPDDTPMGYFYTPLKAGGLGLPSLRLVILNNRL
ERILRMKASQDIIVRTIAESETLGVEIRKLHDLLSIDGTILDTSVKIHSFWAERLYSSY
DGKCLCNSANFPPGNKWIGEDSLNQRSHIFADCLKLRINALPTRSRTARGRPLKDKP
CRAGCRNSDGVKVIETLNHITQVCERTHGARVKRHDRLVDFAVKGLQRPHRVVLKE
PHYKTVNGVRKPDIVIKIPDHTYICDFQVVS DTS CLELEFRKKALKYAEDKGLCDQLT
RDHPGELSFTAITFNT RGLIAKSSVTALRKLGMPPRSIMTLQKICMEGSLEIWRIFNQ
TTAMARN*

10

20

56. LeCoB プラスミド ORF 配列

ATGGATTACAAGGATGACGACGATAAGGGTACCATGAGCGGCAAGAGCTCCAAGCC
TCGCACCGTGTCTAGCGGCTCCTCTAGCCAGGAGACACCCCTTCCGGCTCTAATG
CCTGCGACATCTGTGGCAAGTGCTTCATGAAGCCCGTGGGCCTGTCTCTGCACATG
AGCAAGGTGCACCCTACCCAGTACCACGCCCGCCTGGAGAAGAACCAGCCAAAGGC
CAAGAAGTTCCGGTGGACAGACGAGGATCTGTATTTTCTGGCCAAGAAGGAGGCAG
AGCTGCTGCTGCTGGGAGGCATCAAGTTTATGAATAAGGAGCTGGCCGAGTTCTTT
CCCGAGAAGTCTGTGGATCAGATCAAGGGCCAGCGGAGAAGCGAGACCTACAAGCA
GCAGGTGGTGAGCATCCACTCCGAGCTGCTGAAGCTGCAGGCAGTGGCAGACTCCC
CACCACCTTCTCGCATCCCCGCAAAGGAGGTGTCCGCATGGCTGGATTTCCTGCTG
GCCCTGCCTAAGACCAAGAATAAGTTTTCTGAGGACAAGCTGGATCAGCTGATCAG
AACCGCACAGGAGGGAACACCAGTGCTGAACGACCTGGATCTGTATCTGAGGGAGG
TGCTGGTGCAGCCAACCAGACAGGGAGAGAGGGCAGGCAAAGCCACTGCCACCCCT
AAGTCCTCTCGGGAGAAGAGGGACCGCGAGTACGCCAGAGTGCAAGATTTCTATAG
GAAGAACAAGACCGCCTGTGTGAACGCCATCCTGGATGGCAATAAGAAGTGCGAGA
ACAAGATCCCTGACATCGATGAGTTCTGGAAGGCCATCTTTGAGAGCCAGTCCCCA
CCAGACGCAGAGCCCGTGAGCTACGTGGTGGATGAGGAGCCTAAGAATATCTGGTC
TTGGATCAGCTTCTTTGAGATGAACAGAAATTATCCCGACACCTCTACAAGCCCAG
GACCTGATGGAGTGACCGCAAGGATGCTGAGATCCATCCCTGCCAGGGTGCTGAAC
AAGCTGCTGAATCTGCTGCTGTTTCATCGAGGACCTGCCCGCCGTGTTTAAAGTGTCA
CCGCACCGTGCTGATCCCTAAGGTGGATAACCCAGCACTGCCAGGAGAGTTCCGGC
CAATCACAATCAGCTCCATCATCGTGAGACAGCTGAATAAGATCATCGCAGCAAGG
GTGAGCGAGGGAGTGCTATCAACCCACGCCAGAAGGCCTTTCCGCAGATCGACGG
CTGCGCCGAGAACGTGTTCTGCTGGACTTTATCCTGAGGGATGCCAAGACCAAGA
TCAAGAGCCTGTCTCTGGCCACAGTGGACATCAAGAAGGCCTTCGATAGCGTGTCC
CACCACAGCATCTTTAGGGCAATCAGGGGAGCAAGGTGTCCAGAGAACCTGGTGAA
TTACATCCAGAATTCTTATAGCGGCTGCACCACACAGATCTCCGTGGGCGGCTCCA
TCTCTACCACAAAGATCCTGATGAACAGAGGCGTGAAGCAGGGCGATCCCCTGAGC
CCTGTGCTGTTCAACCTGGTCATCAACGAGATCATCAGGAAGCTGCCTGCCTCCAT

30

40

50

CGGCTACCCAATCAATTCCGAGCTGTCTATCAACTGCATCGCCTATGCCGACGATC
TGATCCTGGTGGCCAATACCAGAGAGGGCCTGAAGCTGCTGCTGAACCTGCTGAAT
GAGGAGCTGCCCCAAGAGGGGCTGGAGCTGAACGCCTCTAAGTGTGTTTGGCCTGAG
CCTGACAGCCCTGGGCAAGCTGAAGAAGACCCACCTGTGCACAAGCGACCAGCTGG
ATCTGCACGGCACCCCTGATCAAGAATCTGACAGCCGAGGAGAGCTGGGTGTACCTG
GGCGTGCCATTCTCCACATCGGCCGCAGCAAGTCCTTTTCTCCCGACCTGGAGGC
CCTGCTGAACAAGCTGCAGAAGTCCCCACTGAAGCTGCAGCAGAAGCTGTTTCGCCC
TGCGCGTGTATCTGATCCCAAGGCTGCTGCACGGACTGGTGCTGTCTAGGGTGGCA
ATCGGCGAGCTGAAGATCATGGACAAGCTGATCCTGAAGCACCTGAGGGTGTGGCT
GAGGCTGCCAAAGGATACCCCCCTGGGCTTCTTTTACTCCCCAGTGAAGCTGGGCG
GCCTGGGCATCAAGAACCTGCGCACAAATGTGCTGAAGTGTAGAAAGCAGAGGATC
GAGCGCATGCTGGTGTCCCCAGACGATGTGGTGCGGCTGGTGGCAGAGTCTGAGAT
CTTTCTGAAGGAGACCGACAAGCTGAAGGATCTGCTGACAATCAATGGCATGTGCC
TGGACACCAGAAACGTGCCTAGGACAGGCAAGAACAATAAGTTCTGGTCCGAGAGA
CTGTACACCTCTTTTGACGGCAAGACACTGGCCTACAGCGAGTATTTACCCAGGG
CGGCGGCTGGATCAGAGAGGATAAGATCCTGCAGCCAGCCACGTGTTTGCCGAGT
GTATCAAGCTGCGGATCAACGCCCTGCCTACCAAGAGCAGAGTGGCCCACGGCAGG
CCAACAAAGGACCGGTCTGTAGAGCCGGCTGCCTGGATGTGCAGAAGGTGCCTGC
CATCGAGACCATCAATCACATCGCCCAGGTGTGCCCAAGAACACACGGAGCAAGGA
TCAAGAGGCACGACAGGCTGGTGCAGTTCCTGTCCCTGAACCTGCGCAAGAATCCT
AAGCGGAACGTGCTGGTGGAGTACAATTTTAGAACCGTGGCCGGCATCAGGAAGCC
AGACATCATCGTGATCGAGGATACAAGGGCAGCAATCCTGGACGTGCAGGTGGTGG
GCGATTCTAGCAACCTGGAGATGGAGTACCTGGAGAAGAGCAGAAAGTATCCAAAT
GATGCCACCCTGTCTATGAGGATCAACGCCCTGCAGAAGCTGTATCCCACCGTGAC
AAGCCTGACCTTCCACGCCGTGACCTTCAACAATAGGGGCCTGATCGCCAAGTCTA
CAGTGGCAGCACTGAGGATGCTGGGAGTGCCTCCACGGTGTATCATGATCCTGTGC
GTGATCAGCCTGGAGAAGACCCTGGAAGTGTGGCGCATGTTCAACCAGAGCACAGC
CTCCGCCCCGGAAGTAG

10

20

30

40

57. LeCoBアミノ酸配列

MDYKDDDDKGTMSGKSSKPRTVSSGSSSQETPPSGSNACDICGKCFMKPVGLSLHM
SKVHPTQYHARLEKNQPKAKKFRWTDLEDLYFLAKKEAELLLLGGIKFMNKELAEFFP
EKSVDQIKGQRRSETYKQQVVSIIHSELLKLQAVADSPPPSRIPAKEVSAWLDFLLALP
KTKNKFSEDKLDQLIRTAQEGTPVLNDLDLYLREVLVQPTRQGERQAKPLPPPKSSR
EKRDREYARVQNFYRKNKTACVNAILDGNKKCENKIPDIDEFWKAIFESQSPDAEP
VSYVVDEEPKNIWSWISFFEMNRNYPDTSTSPGPDGVTARMLRSIPARVLNKLNLNL
LFIEDLPAVFKCHRTVLIPKVDNPALPGEFRPITISSIIVRQLNKIIAARVSEGVPINPR
QKAFRQIDGCAENVFLDFILRDAKTKIKSLSLATVDIKKAFDSVSHHSIFRAIRGARC
PENLVNYIQNSYSGCTTQISVGGSSISTTKILMNRGVKQGDPLSPVLFNLVINEIIRKLP
ASIGYPINSELSINCIAYADDLILVANTREGLKLLNLNLLNEELPKRGLELNASKCFGLS
LTALGKLKKTHLCTSDQLDLHGTLIKNLTAESWVYLGVPFSHIGRSKSFSPDLEALL
NKLQKSPLKLQQKLFALRVYLIPRLLHGLVLSRVAIGELKIMDKLILKHLRVWLRLP
KDTPLGFFYSPVKLGGLGIKNLRTNVLKCRKQRIERMLVSPDDVVRLVAESEIFLKET
DKLKDLLTINGMCLDTRNVPRTGKNNKFWSERLYTSFDGKTLAYSEYFTQGGGWIR
EDKILQPAHVFAECIKLRINALPTKSRVAHGRPTKDRSCRAGCLDVQKVPAIETINHI
AQVCPRTHGARIKRHDRLVQFLSLNLRKNPKRNVLVEYNFRTVAGIRKPDIIVIEDTR
AAILDVQVVGDSSNLEMEYLEKSRKYSNDATLSMRINALQKLYPTVTSLTFHAVTFN
NRGLIAKSTVAALRMLGVPPRCIMILCVISLEKTLEVWRMFNQSTASARK*

58. トランスフェクトしたRT mRNA 5' UTR

AGGAUAUAAACUAGUAUUCUUCUGGUCCCCACAGACUCAGAGAGAACCCGCCACC

59. トランスフェクトしたRT mRNA 3' UTRおよびpolyAテール

50

UGAUGACCUCGAGCUGGUACUGCAUGCACGCAAUGCUAGCUGCCCCUUUCCCGUC
CUGGGUACCCCGAGUCUCCCCGACCUCGGGUCCAGGUAUGCUCCCACCUCCAC
CUGCCCCACUCACCACCUCUGCUAGUUCAGACACCUCCCAAGCACGCAGCAAUG
CAGCUCAAAACGCUUAGCCUAGCCACACCCCCACGGGAAACAGCAGUGAUUAACC
UUUAGCAAUAAACGAAAGUUUAACUAAGCUAUACUAACCCCAGGGUUGGUCAAUU
UCGUGCCAGCCACACCUGGAGCUAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAGCAUAUGACUAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAA

60. OrLa5Lモジュール

10

GGCGGGAGTAACTATGACTCTCTTAACGCACAGGGGACACAGAGCCTGCCCAAGTA
CCGCTCCCGAGGGAGCGGGAAACGGGGGGGTGACTATCCCCTGGGGTCCGGCGAGA
GCGCTGGTCTACGGACCAGGGGTGGCTGTGGGCAGGCTGCTCCTCAGGCCAGTTGA
TTAGTTACGCATGGGCTGTACCTCCACGTGGTCCCGCTGGTAACGACTTGTCGGCT
AAATCAGCCCGCCACCATCTGGGATATGGTTGACCGTCTAACCCCAGTACTCAGG
TCACAAACAAAATGGGAACAGATACAGTGTATGTCGGCCAGGACTACCCTTCTGGC
TTATCAAAACGGGTACCAGCACGGT

61. OrLa5Sモジュール

20

GGCGGGAGTAACTATGACTCTCTTAACGCACAGGGGACACAGAGCCTGCCCAAGTA
CCGCTCCCGAGGGAGCGGGAAACGGGGGGGTGACTATCCCCTGGGGTCCGGCGAGA
GCGCTGGTCTACGGACCAGGGGTGGCTGTGGGCAGGCTGCTCCTCAGGCCAGTTGA
TTAGTTACGCATGGGCTGTACCTCCACGTGGTCCCGCTGGTAACGACTTGTCGGCT
AAATCAGCCCGCCACCATCTGGGATATGGTTGACCGTCTAACCCCAGTACTCAGG
TCACAAACAAA

62. TriCasB5 [すなわちTCA5-28(28)]

30

GACTCTCTTAAGGGAAGACCCCGCCCATGAGGCTTGGAGAGTGTGATCCTGATCAG
ATCACACTTGAAAAGTTATGCTGAGTACGTCCGCGTCGTGAGAGTCGGTAACTGTC
CCAGGATGGTCTGGGATAGGCTAAACCTCAGCAGGGGAAAGTTGTAGGGGCCTGCC
ACCCCTACACTTTATTGGTATGGCATTTCGATACCCCTAACGAAGCCTCGGACTTGG
AGGAGCACGGTTCCCCTCCTCCTCGTATTAGACCAGGAACCAACTGTCCTGACAAC
CCCATTGGACCTATGGGAGCGGACCATGCTATGGACATGGATTCCGAAGACGAAGC
GGGGGCACACGGACCCCCCGCCGATAGTGCTCACTTAACGTCAGGCGAACCCCTTG
AAATCATCTTG

63. TriCasB_5RZdead [すなわちTCA5_rzdead]

40

GACTCTCTTAAGGGAAGACCCCGCCCATGAGGCTTGGAGAGTGTGATCCTGATCAG
ATCACACTTGAAAAGTTATGCTGAGTACGTCCGCGTCGTGAGAGTCGGTAACTGTC
CCAGGATGGTCTGGGATAGGTTAAACCTCAGCAGGGGAAAGTTGTAGGGGCCTGCC
ACCCCTACACTTTATTGGTATGGCATTTCGATACCCCTAACGAAGCCTCGGACTTGG
AGGAGCACGGTTCCCCTCCTCCTCGTATTAGACCAGGAACCAACTGTCCTGACAAC
CCCATTGGACCTATGGGAGCGGACCATGCTATGGACATGGATTCCGAAGACGAAGC
GGGGGCACACGGACCCCCCGCCGATAGTGCTCACTTAACGTCAGGCGAACCCCTTG
AAATCATCTTG

64. TriCasB_5RZ

GACTCTCTTAAGGGAAGACCCCGCCCATGAGGCTTGGAGAGTGTGATCCTGATCAG
ATCACACTTGAAAAGTTATGCTGAGTACGTCCGCGTCGTGAGAGTCGGTAACTGTC
CCAGGATGGTCTGGGATAGGCTAAACCTCAGCA

65. TriCasB_5RZmin

GACTCTCTTAAGGTGCTGAGTACGTCCGCGTCGTGAGAGTCGGTAACTGTCCCAGG
ATGGTCTGGGATAGGCTAAACTCAGCA

66. TriCasA5(28)モジュール

50

GGCGGGAGTAACTATGACTCTCTTAAGGAGTCATAGAGCCAGAACCTCCTCGTGGT
CCCGCTGGGCACAGGGATTAATTTTTCTGTGGCAAATTTGACTGGCTTCAGAGAGC
GTTTTTCGAAGTGGACTGTGTGACTGCGTTCCCCCTTAGTTGCTATATCCGCTTCG
ATTAACATCTCACCTCGACGTATAAGATCATT

67. TriCasA_RZモジュール [すなわちTCARZ-28(28)]

GGCGGG AGTAACTATGACTCTCTTAAGG AGTCATAG AGCCAGA ACC TCCTCGT
GGT CCCGCT GG GCACAGGGATTAATTTTTCTGTGG CAAA TTTGACT GGCTTCAGA

68. ZoA15モジュール

GGCGGGAGTAACTATGACTCTCTTAAGG cgacttgaga aggtctgggtt acaactgggc a
tagct gcagagatcg cgcc tcctcgtggc cccgct ggta agcccttaac agggtgacta agt
cgatctctgc cccagtc caggagccgc tgggtttcac cagcccagcg attccttcca aattcg
gtga aacaaattcc tcggtaaaag ccgctgtggct tattgcctga aacctggccc cccggtttca
gacagggggca aagagttcgg aagtggactg accaccacc ccgaaccga gaggc
gaatct ggtcatgacc caactgtccc aaatcctgggt ccgtctcttg gagcgggggga aggtgc
acag ccactaccct tactcagggt atcggtgggc acccaaacct gtgaagagga cttataa
ca tctagaccaa ccaaattacc cggaattgaa tcagaattag gcccgtgggt gaagttttct t
tagaggttt acaggtcaga tcttaagggg gatgtgcaat ttgaggggat tcattttc
ca gataattggg ggggtactgga ggggtttcct gaggtgtacg aacaactggc accacagcca
aacggggggag acgagttaaa tcatagtctc ccagggggaca gggagggggga tgtacttgag aa
ggatagcagcgaaaagga gaaggaggct gcaccagagg cattgccctc agtgcaaagg gc
ccgcagtg aacagttgcc

69. TiGu5モジュール

GGCGGGAGTAACTATGACTCTCTTAAGG ctggggaccg tggttacaac ccgggcttag c
t gcagagacag tacctccc cgtgggtt ccc gcc ggacccc gtaacatcgg gtgactg
aat ctgtctctgc cccgggagta gttcctcctt gccctattga ccagcggctc cgggctgc
tc aatagtattc taggcgtgaa atatagcgat agtcctagtg gttgtcttac tgggccatag
ccccttgctt caggggtcat tcgcgaagtc tctcaggaga actggggggtg gtgttcttct g
ggtatagct aaacccccta gactgtgtcc gatccatggg gtcctggatc gtgaatttcg
tttcgggtggc gactcagacg ggagaattcc ctgtggatac ggccaggagg gcacct
gtgc cggtaacatc ataccctgag tcggaatgcc acntaccgtt gcccctgaca ttt
tgtaact cggatgtgac tatttgggga ggggttcgcc ctgaaccggt ggactgcttg
ggwgatcttc crgaggygta tgatgcactc ccaggggttg ctgggcctcg ggaay
cgggtg ggtgggagcc cgccgggrga aggggtcagg tcgccaggga ttgcgtcrcc ctctggt
act gcggtccaac atgattttgg gagtcccatc ctcgtaccgg gagccgaagc cgc
cgaggtn tctaccccg tagtgaaggt tcncaagac catccagcat

70. DroSi5モジュール

GGGGATCTGGGGTAATTGCGAGCAGAGGGGGAGTATTTTTCTGTAATTCGTAAGTC
ATATCATATGGTGTGCGGAAGGGGAATTTACTCTGTAATCACAAGTCTCTCCTT
TACTCAAGTCGACTCAAAACCTCCTCGTGGTGGTCCCCGGTAATGCTAAACTTGTT
TAGCAGCTAATTTGAGCGGCAAAACTTTTCCGATGGGCTGGTTACCCAGAGGAAA
TTTACTCATATTGGAACACGAACACAAATAACGAGCCTCGGATATCTTTACACAA
TCTG

71. TriCan5モジュール

Ggtctggcatttgatcgtttccgcctcctcgtggcgccagactgggtaagctgatttatcagtgagcta
agagaaacgatcaccgcaggagtccatctacctgcgctgcacgtgatttcattgtgcacttggcgagtt
atcccttgaggagctcgtcgttagcttttgag

72. LeCoB5モジュール

ggggtagcaattgatcgattcccgcctcctcgtggcgctaccctgggataacctcaaagaaatttgac
ggtaaagctaagaggaattgatcacaccgtgacgaatatcatcagtcacagctgcacgaatccagata
gatataaacaggcgagtaattcttttcgaagctcgtcagtaatcttcccag

10

20

30

40

50

73. FoRa5モジュール

GGCGGGAGTAAGTACTGACTCTCTTAAGGTAGTCCAAAAGCCTTGTCATCTTAGGAT
GACGCGCAGGTTTTTGAGATGTATATTTTTATAGTAGTAGTAAAATACGACTTGTT
AAATTCTTAAGATAAAGAGTCATAGTAGCAAGTGAAGACCTTTTCGTGGTTCCTGT
CATAAGTACCCTACGGGGAAACCAACTTCACCCCTGCGAACATGGAGCTTAGAGGT
TACCAGCCTCCTCGCTCGTCAATTAGCGAAGAGC

74. Ciln5モジュール

GGTAACTATG ACTCTCTTAA cgacggtgaa ccaccttgtc gcggtgtaag agcttttagtg
tctcgaacaa gaaatagctt gtgtgctgtc cttctgggcg gtgcacatac ttcttaacct cccga
ggcca tgccggcggg ggcttttagcc cccggcaggt ttaccatgc cggacgggtt cgagaggt
ag aggccaaact aagagttcac cagcagactt cgcacgcggc tggccactgg ccgaagtta
aacaacaggg ccgcatcttc ccaaactcaa tataatggtgt taagtgaacc gtgccg
atgg gtgaatggcc atgggtaagt tggctcttaa cggtccttgt ggagaagtgg cgaccttca
caatacttca gccttacctt atgcccgggc aacttagggg ggacgtatat ctcccgcgga aaac
atctta tctcatggat aaaaatatat acgaaaacac tacttctcca ggaggggggac cgctttgc
gg tgagaagacg catcggtccg atgtgatcat ccctccaccg gggttcgccc cgtcgacgga
cactgcttcc

10

75. OrLa 5RZop

ggcgggAGTAACTATGACTCTCTTAA cgcacagggg acacagagcc tgcccaagta ccg
ctccga gggagcgga aacggggggg tgactatccc ctgggggtccg gcgagagcgc
tggctctacgg accaggggtg gctgtgggca ggctgctcct caggccagtt gattagttac
gcatgggctgt acc tcc acgt ggt cccgct gg ta acgaccttggtcgg cGaa acagccc

20

76. OrLa 5RZmin

ggcgggAGTAACTATGACTCTCTTAA gggctgt acc tcc acgt ggt cccgct ggtaac
gaccttggtcgg cGaa acagccc

77. OrLa5L_MS2

GGCGGG AGTAACTATGACTCTCTTAA CGCACAGGGGACACAGAGCCTGCCCAAGT
ACCGCTCCCGAGGGAGCGGGAAACGGGGGGGTGACTATCCCCTGGGGTCCGGCGAG
AGCGCTGGTCTACGGACCAGGGGTGGCTGTGGGCAGGCTGCTCCTCAGGCCAGTTG
ATTAGTTACGCATGGGCTGTACCTCCACGTGGTCCCGCTGGTAACGACCAGGAACA
CCGGTCGGCTAAATCAGCCCGCCACCATCTGGGATATGGTTGACCGTCTAACCCC
AGTACTCAGGTCACAAACAAAATGGGAACAGATACAGTGTATGTCGGCCAGGACTA
CCCTTCTGGCTTATCAAAACGGGTACCAGCACGGT

30

78. OrLa5S_PP7

GGCGGG AGTAACTATGACTCTCTTAA CGCACAGGGGACACAGAGCCTGCCCAAGT
ACCGCTCCCGAGGGAGCGGGAAACGGGGGGGTGACTATCCCCTGGGGTCCGGCGAG
AGCGCTGGTCTACGGACCAGGGGTGGCTGTGGGCAGGCTGCTCCTCAGGCCAGTTG
ATTAGTTACGCATGGGCTGTACCTCCACGTGGTCCCGCTGGTA ACGACACGACATA
GGAGTCGGTCGG CTAAATCAGCCC GCCCACCATCTGGGATATGGTTGACCGTCTA
ACCCAGTACTCAGGTCACAAACAAA

40

79. OrLa5L_MS2(CAAA)

GGCGGG AGTAACTATGACTCTCTTAA CGCACAGGGGACACAGAGCCTGCCCAAGT
ACCGCTCCCGAGGGAGCGGGAAACGGGGGGGTGACTATCCCCTGGGGTCCGGCGAG
AGCGCTGGTCTACGGACCAGGGGTGGCTGTGGGCAGGCTGCTCCTCAGGCCAGTTG
ATTAGTTACGCATGGGCTGTACCTCCACGTGGTCCCGCTGGTA ACGACCAGGAACA
CCGGTCGG CAAAATCAGCCC GCCCACCATCTGGGATATGGTTGACCGTCTAACCC
CAGTACTCAGGTCACAAACAAA ATGGGAACAGATACAGTGTATGTCGGCCAGGAC
TACCCTTCTGGCTTATCAAAACGGGTACCAGCACGGT

80. OrLa5L_P2

GGCGGG AGTAACTATGACTCTCTTAA CGCACAGGGGACACAGAGCCTGCCCAAGT

50

ACCGCTCCCGAGGGAGCGGGAAACGGGGGGGTGACTATCCCCTGGGGTCCGGCGAG
AGCGCTGGTCTACGGACCAGGGGTGGCTGTGGGCAGGCTGCTCCTCAGGCCAGTTG
ATTAGTTACGCATGGGCTGTACCTCCACGTGGT CCGCT GGTA ACGACCAGGAA
CACCGGTCGG CTAA ACAGCCCGCCCACCATCTGGGATATGGTTGACCGTCTAACC
CCAGTACTCAGGTCACAAACAAA ATGGGAACAGATACAGTGTATGTCGGGCCAGGA
CTACCCTTCTGGCTTATCAAAACGGGTACCAGCACGGT

81. Orla5L(28)

GGCGGG AGTAACTATGACTCTCTTAAGG CGCACAGGGGACACAGAGCCTGCCCAA
GTACCGCTCCCGAGGGAGCGGGAAACGGGGGGGTGACTATCCCCTGGGGTCCGGCG
AGAGCGCTGGTCTACGGACCAGGGGTGGCTGTGGGCAGGCTGCTCCTCAGGCCAGT
TGATTAGTTACGCATGGGCTGTACCTCCACGTGGTCCCGCTGGTA ACGACCAGGAA
CACCGGTCGG CTAA ATCAGCCCGCCCACCATCTGGGATATGGTTGACCGTCTAAC
CCCAGTACTCAGGTCACAAACAAA ATGGGAACAGATACAGTGTATGTCGGGCCAGG
ACTACCCTTCTGGCTTATCAAAACGGGTACCAGCACGGT

10

82. Orla5L(27)

GGCGGG AGTAACTATGACTCTCTTAAG CGCACAGGGGACACAGAGCCTGCCCAAG
TACCGCTCCCGAGGGAGCGGGAAACGGGGGGGTGACTATCCCCTGGGGTCCGGCGA
GAGCGCTGGTCTACGGACCAGGGGTGGCTGTGGGCAGGCTGCTCCTCAGGCCAGTT
GATTAGTTACGCATGGGCTGTACCTCCACGTGGTCCCGCTGGTA ACGACCAGGAAC
ACCGGTCGG CTAA ATCAGCCCGCCCACCATCTGGGATATGGTTGACCGTCTAACC
CCAGTACTCAGGTCACAAACAAA ATGGGAACAGATACAGTGTATGTCGGGCCAGGA
CTACCCTTCTGGCTTATCAAAACGGGTACCAGCACGGT

20

83. Orla5L(25)

GGCGGG AGTAACTATGACTCTCTTA CGCACAGGGGACACAGAGCCTGCCCAAGTA
CCGCTCCCGAGGGAGCGGGAAACGGGGGGGTGACTATCCCCTGGGGTCCGGCGAGA
GCGCTGGTCTACGGACCAGGGGTGGCTGTGGGCAGGCTGCTCCTCAGGCCAGTTGA
TTAGTTACGCATGGGCTGTACCTCCACGTGGTCCCGCTGGTA ACGACCAGGAACAC
CGGTCGG CTAA ATCAGCCCGCCCACCATCTGGGATATGGTTGACCGTCTAACCC
AGTACTCAGGTCACAAACAAA ATGGGAACAGATACAGTGTATGTCGGGCCAGGACT
ACCCTTCTGGCTTATCAAAACGGGTACCAGCACGGT

30

84. 配列番号65の重複

85. TriCasB 5RZmin+5downstream

gactctc ttaAGG tgctgagt acgtccgcgtcgt gagagtc ggta actgtcccaggatgggtctg
ggatagg cttaa actcagca ggggaaagtgttaggggcctgccaccctacactttattggatggc
attcgatacccctaacgaagcctcggacttggaggagcacggttccccctcctcctcgtattagaccag
gaaccaactgtcctgacaacccattggacatatgggagcggaccatgctatggacatggattccgaa
gacgaagcggggggcacacggaccccccgccgatagtgtcacttaacgtcaggcgaaccccttgaaa
tcactttg

86. TriCasB_RZop

gactctc ttaAGG gctgagt acc tcc gcgt ggt gagagtc ggta acagccauugcacucc
ggctgg cgaa actcagc

40

87. TCA_MS2

GGCGGG AGTAACTATGACTCTCTTAAGG
AGTCATAG AGCCAGA ACC TCCTCGT GGT CCGCT GG GCACAGAGGAACACCC
TGTGG CAAA TTTGACT GGCTTCAGA GAGCGTTTTT CGAAGTGGAC TGTGTGAC
TG CGTTCCCCC TTAGTTGCTA TATCCGCTTC GATTAACATC TCACCTCGAC G
TATAAGATC ATT

88. TriCasA5_catdead

GGCGGG AGTAACTATGACTCTCTTAAGG AGTCATAG AGCCAGA ACC TCCTCGT
GGT CCGCT GG GCACAGGGATTAATTTTTCTGTGG TAAA TTTGACT GGCTTCA

50

GA GAGCGTTTTT CGAAGTGGAC TGTGTGACTG CGTTCCCCC TTAGTTGCTA T
 ATCCGCTTC GATTAACATC TCACCTCGAC GTATAAGATC ATT

89. TCARZ_MS2
 GGC GGG AGTAACTATGACTCTCTTAAGG AGTCATAG AGCCAGA ACC TCCTCGT
 GGT CCCGCT GG ACACAGAGGAACACCCTGTGG CAAA TTTGACT GGC

90. TCARZ_P1
 GGC GGG AGTAACTATGACTCTCTTAAGG AGTCATAG AGCCAGA ACC TCCTCGT
 GGT CCCGCC GG ACACAGAGGAACACCCTGTGG CAAA TTTGACT GGC

91. TCARZ_P2
 GGC GGG AGTAACTATGACTCTCTTAAGG AGTCATAG AGCAGA ACC TCCTCGT G 10
 GT CCCGCT GG ACACAGAGGAACACCCTGTGG CAAA TTTGCT GGC

92. TCARZ(27)
 GGC GGG AGTAACTATGACTCTCTTAAG AGTCATAG AGCCAGA ACC TCCTCGT G
 GT CCCGCT GG ACACAGAGGAACACCCTGTGG CAAA TTTGACT GGC

93. TCARZ(26)
 GGC GGG AGTAACTATGACTCTCTTAA AGTCATAG AGCCAGA ACC TCCTCGT GG
 T CCCGCT GG ACACAGAGGAACACCCTGTGG CAAA TTTGACT GGC

94. TCARZ(25)
 GGC GGG AGTAACTATGACTCTCTT AGTCATAG AGCCAGA ACC TCCTCGT GGT
 CCCGCT GG ACACAGAGGAACACCCTGTGG CAAA TTTGACT GGC 20

95. TCARZ_wob1
 GGC GGG AGTAACTATGACTCTCTTAAGG AGTCATAG AGCAGA ACC CCCCCAC
 GGT CCCGCC GG ACACAGAGGAACACCCTGTGG CAAA TTTGCT GGC

96. TCARZ_wob2
 GGC GGG AGTAACTATGACTCTCTTAAGG AGTCATAG AGCAGA ACC CCCCCAC
 GGT CCCGCC GG GCA ACACAGAGGAACACCCTGTGG CAAA TTTGCT GGC

97. TCARZ_wob3
 GGC GGG AGTAACTATGACTCTCTTAAGG AGTCATAG AGCAGA GCC CCCCCAC
 GGC CCCGCC GG GCA ACACAGAGGAACACCCTGTGG CAAA TTTGCT GGC

98. TCARZ_wob4 30
 GGC GGG AGTAACTATGACTCTCTTAAGG AGTCATAG AGCAGA GCC CCCCCAC
 GGC CCCGCC GG GCA ACACAGAGGAACACCCTGTGG CAAA TTTGCT GGC

99. TCARZ(28)_NL
 GGC GGG AGTAACTATGACTCTCTTAAGG AGTCATAG AGCCAGA ACC TCCTCGT
 GGT CCCGCT GG GCACAGGGATTAATTTTTCTGTGG CAAA TTTGACT GGC TTC
 TAGA

100. TCARZ(28)_L8 [自己切断により最初の8ntが除去される]
 GGGTAAAC GGC GGG AGTAACTATGACTCTCTTAAGG AGTCATAG AGCCAGA A
 CC TCCTCGT GGT CCCGCT GG GCACAGGGATTAATTTTTCTGTGG CAAA TTTG
 ACT GGC ITCTAGA AGGTCGACCAGATGTCCGAG 40

101. TCARZ(26)_L8 [自己切断により最初の8ntが除去される]
 GGGTAAAC GGC GGG AGTAACTATGACTCTCTTAA AGTCATAG AGCCAGA ACC
 TCCTCGT GGT CCCGCT GG GCACAGGGATTAATTTTTCTGTGG CAAA TTTGACT
 GGC ITCTAGA

102. hdvRZ-28
 ggcgggAGTAACTATGACTCTCTTAA ggtccca gcc tcc tcgt ggc tcccgcc ggta a
 cagccauugcacuccggctgg cgaa tgggacc

103. hdvRZ-13_1
 gactctcTTAAGG ggtccca gcc tcctcgt ggc gagagtc ggta acagccauugcacuccg
 gctgg cgaa tgggacc 50

104. hdvRZ-13_2

gactctcTTAAGG ggtccca gcc CCCTCAC ggc gagagtc ggta ACACAGAGGAACA
CCCTGTGG cgaa tgggacc

105. hdvRZ-21

GTA ACTATGACTCTCTTAAGG ggtccca gcc tcc tcgt ggc tagttac ggta acagcc
auugcacuccggctgg cgaa tgggacc

106. hdvRZ+0

ggccggc at ggtccca gcc tcc tcgt ggc gccggcc ggta acagccauugcacuccggct
gg cgaa tgggacc

107. hdvRZ-28(26)_gu1

GGCGGG AGTAACTATGACTCTCTTAA AAAAGAGAATCATAGAA CGTCAGCA GCC
TCCTCGC GGC CCGGCC GGTA AGATTCCGAAAGGAATCG CGAA TGCTGACG AA
AGTCTTC

10

108. hdvRZ-28(26)_ac1

GGCGGG AGTAACTATGACTCTCTTAA AAAAGAGAATCATAGAA CGTCAGCA GCC
CCCTCAC GGC CCGGCC GGTA AGATTCCGAAAGGAATCG CGAA TGCTGACG AA
AGTCTTC

109. hdvRZ-28(26)_ac2

GGCGGG AGTAACTATGACTCTCTTAA AAAAGAGAATCATAGAA CGTCAGCA GCC
CCCTCAC GGC CCGGCC GGTA ACACAGAGGAACACCCTGTGG CGAA TGCTGACG
AAAGTCTTC

20

110. hdvRZ-13(13)_gu1

GACTCTC TTAAGG CGTCAGCAGCCTCC TCGCGGCCCC GCCGGTAAGA TTCCGA
AAGG AATCG CGAA TGCTGACG A AAGTCTTC

111. hdvRZ-13(13)_ac1

GACTCTC TTAAGG CGTC AGCAGCCCC TCACGGCCCC GCCGGTAAGA TTCCG
AAAGG AATCG CGAA TGCTGACG A AAGTCTTC

112. hdvRZ-13(13)_ac2

GACTCTC TTAAGG CGTC AGCAGCCCC TCACGGCCCC GCCGGTAAGA CAGAG
GAACA CCCTGTCG CGAA TGCTGACG A AAGTCTTC

30

113. hdvRZ-28(28)_ac2b

GGCGGG AGTAA CTATGACTCTCTTAAGG AAGAGAATCATAG AA CGTCAGCA G
CCCCCTCAC GGC CCGGCC GGTA ACACAGAGGAACACCCTGTGG CGAA TGCTG
ACG A

114. hdvRZ-28(27)_ac2c

GGCGGG AGTAA CTATGACTCTCTTAAG A AAGAGAATCATAG AA CGTCAGCA G
CCCCCTCAC GGC CCGGCC GGTA ACACAGAGGAACACCCTGTGG CGAA TGCTG
ACG A

115. hdvRZ-28(25)_ac2d

GGCGGG AGTAA CTATGACTCTCTTA CAA AAGAGAATCATAG AA CGTCAGCA G
CCCCCTCAC GGC CCGGCC GGTA ACACAGAGGAACACCCTGTGG CGAA TGCTG
ACG A

40

116. hdvRZ-13(13)_ac11

GACTCTC TTAAGG A CGCAGCA GCCCCCTCAC GGC GAGAGTC GGTA ACACAG
AGGAACACCCTGTGG CGAA TGCTGCG A

117. hdvRZ-13(13)_ac12

GACTCTC TTAAGG A CGCAGCA GCCCCCTCAC GGC GAGAGTT GGTA ACACAG
AGGAACACCCTGTGG CGAA TGCTGCG A

118. hdvRZ-13(11)_ac11b

GACTCTC TTA AAA CGCAGCA GCCCCCTCAC GGC GAGAGTC GGTA ACACAGAG

50

GAACACCCTGTGG CGAA TGCTGCG A
119. hdvRZ-13(11)_ac12b
GACTCTC TTAAAA CGCAGCA GCCCCCTCAC GGC GAGAGTT GGTA ACACAGAG
GAACACCCTGTGG CGAA TGCTGCG A
120. HDVRZ-13(13)_ac5
GACTCTC TTAAGG A CGACGCG GCC CCCACAC GGC GAGAGTC GGCA ACACAG
AGGAACACCCTGTGG CGAA CGCGTCG A
121. HDVRZ-28(28)_gu5b
GGCGGG AGTAA CTATGACTCTCTTAAGG AAAAGAGAATCATAGAA CGTCAGCA
GCCTCCTCGC GGC CCCGCC GGTA ACACAGAGGAACACCCTGTGG CGAA TGCTG
ACG A 10
122. HDV-28(26)_gu5
GGCGGG AGTAA CTATGACTCTCTTAA AAAAGAGAATCATAGAA CGTCAGCA GC
CTCCTCGC GGC CCCGCC GGTA ACACAGAGGAACACCCTGTGG CGAA TGCTGAC
G A TCTAGA
123. HDV-28(22)_gu5
GGCGGG AGTAA CTATGACTCTC AACAAAGAGAATCATAGAA CGTCAGCA GCCT
CCTCGC GGC CCCGCC GGTA ACACAGAGGAACACCCTGTGG CGAA TGCTGACG
A TCTAGA
124. HDV-28(18)_gu5 20
GGCGGG AGTAA CTATGAC AACA GTCATAGAA CGTCAGCA GCCTCCTCGC GGC
CCCGCC GGTA ACACAGAGGAACACCCTGTGG CGAA TGCTGACG A TCTAGA
125. HDV-28(14)_gu5
GGCGGG AG TAACTA AACA TAGTTA GAA CGTCAGCA GCCTCCTCGC GGC CCC
GCC GGTA ACACAGAGGAACACCCTGTGG CGAA TGCTGACG A TCTAGA
126. HDVRZ-28(28)_gu5_NP2 [すなわちHDV-28(28)_rzdead]
GGCGGG AGTAA CTATGACTCTCTTAAGG AAAAGAGAATCATAGAA CGTCAGCA
GCCTCCTCGC GGC CCCGCC GGTA ACACAGAGGAACACCCTGTGG CGAA
127. HDVRZ-28(28)_gu5_CD
GGCGGG AGTAA CTATGACTCTCTTAAGG AAAAGAGAATCATAGAA CGTCAGCA 30
GCCTCCTCGC GGC CCCGCC GGTA ACACAGAGGAACACCCTGTGG AGAA TGCT
GACG A
128. SL-28_26
GG CGGGAGT A ACTAT GA CTCTCTT AACA AAGAGAG A ATAGT A ACTCCC
G TCTAGA
129. HDV-28(28)_gu5b_CDNL
GGCGGG AGTAA CTATGACTCTCTTAAGG AAAAGAGAATCATAGAA CGTCAGCA
GCCTCCTCGC GGC CCCGCC GGTA ACACAGAGGAACACCCTGTGG AGAA TGCT
GACG A TCTAGA
130. HDV-28(28)_gu5b_L8 [自己切断により最初の8ntが除去される] 40
GGGTAAAC GGCGGG AGTAA CTATGACTCTCTTAAGG AAAAGAGAATCATAGAA
CGTCAGCA GCCTCCTCGC GGC CCCGCC GGTA ACACAGAGGAACACCCTGTGG C
GAA TGCTGACG A TCTAGA
131. HDV-28(27)_gu5b_L8 [自己切断により最初の8ntが除去される]
GGGTAAAC GGCGGG AGTAA CTATGACTCTCTTAAG AAAAGAGAATCATAGAA
CGTCAGCA GCCTCCTCGC GGC CCCGCC GGTA ACACAGAGGAACACCCTGTGG C
GAA TGCTGACG A TTCTAGA
132. HDV-28(25)_gu5_L8 [自己切断により最初の8ntが除去される]
GGGTAAAC GGCGGG AGTAA CTATGACTCTCTTAGCA AAGAGAATCATAGAA CG
TCAGCA GCCTCCTCGC GGC CCCGCC GGTA ACACAGAGGAACACCCTGTGG CG 50

AA TGCTGACG A TTCTAGA

133. HDV-28(24)gu5_L8 [自己切断により最初の8ntが除去される]

GGGTAAAC GGC GGG AGTAA CTATGACTCTCTTGAAA AAGAGAATCATAGAA CG
TCAGCA GCCTCCTCGC GGC CCCGCC GGTA ACACAGAGGAACACCCTGTGG CG
AA TGCTGACG A TTCTAGA

134. HDV-28(23)gu5_L8 [自己切断により最初の8ntが除去される]

GGGTAAAC GGC GGG AGTAA CTATGACTCTCTTGAAA AGAGAATCATAGAA CGTC
AGCA GCCTCCTCGC GGC CCCGCC GGTA ACACAGAGGAACACCCTGTGG CGAA
TGCTGACG A TTCTAGA

135. HDV-28(30)gu5_L8 [自己切断により最初の8ntが除去される]

GGGTAAAC GGC GGG AGTAA CTATGACTCTCTTAAGG TA AAGAGAATCATAGAA
CGTCAGCA GCCTCCTCGC GGC CCCGCC GGTA ACACAGAGGAACACCCTGTGG C
GAA TGCTGACG A TTCTAGA

136. HDV-28(32)gu5_L8 [自己切断により最初の8ntが除去される]

GGGTAAAC GGC GGG AGTAA CTATGACTCTCTTAAGG TAGC TAAGAGAATCAT
AGAA CGTCAGCA GCCTCCTCGC GGC CCCGCC GGTA ACACAGAGGAACACCCT
GTGG CGAA TGCTGACG A TTCTAGA

137. HDV-28(28)_gu5b_L8 [自己切断により最初の8ntが除去される]

GGGTAAAC GGC GGG AGTAA CTATGACTCTCTTAAGG AAAAGAGAATCATAGAA
CGTCAGCA GCCTCCTCGC GGC CCCGCC GGTA ACACAGAGGAACACCCTGTGG C
GAA TGCTGACG A TCTAGA

138. HDV-36(36)

GGGTAAA CGGCGGGAGTAACTATGAC TCTCTTAAGG AAAAGAGAATCATAG AA
ACTCCCG AA CGTCAGCA GCCTCCTCGC GGC TTTACCC GGTA ACACAGAGGAA
CACCTGTGG CGAA TGCTGACG A TTCTAGA

139. HDV-24(24)

GGAGTAA CTATGACTCTCTTAAGG AAGAGAATCATCAA CGTCAGCA GCCTCCTC
GC GGC TTA CTCC GGTA ACACAGAGGAACACCCTGTGG CGAA TGCTGACG A T
TCTAGA

140. HDV-21(21)

GTA ACTA TGACTCTCTTAAGG AAGAGAGTAAA CGTCAGCA GCCTCCTCGC GGC
TAGTTAC GGTA ACACAGAGGAACACCCTGTGG CGAA TGCTGACG A TTCTAGA

141. HDV-13(13)

GACTCTC TTAAGG CGTCAGCA GCCTCCTCGC GGC TAGTTAC GGTA ACACAG
AGGAACACCCTGTGG CGAA TGCTGACG A TTCTAGA

142. HDVgrZ

GGCCGGC AT GGTCCCA GCCTCCTCGCT GGC GCCGGCT GG GCA ACATTCCGA
GGGGA CCG TCCCTCGGTAATGG CGAA TGGGACC TCTAGA

143. HDV-28(3co)

GGGTAAAC GGC GGG AGTAA CTATGACTCTCTTAACA CCGAAC TGTTA GAGAA
TCATAGAA CGTCAGCA GCCTCCTCGC GGC CCCGCC GGTA ACACAGAGGAACA
CCCTGTGG CGAA TGCTGACG A TTCTAGA

144. HDV-28(26)_gu6

GGCGGG AGTAA CTATGACTCTCTTAA AAAAGAGAATCATAGAA CGTCAGCG GC
CTCCACGC GGC CCCGCC GGA ACGCAGAGGAACACCCTGCGG CGAA CGCTGACG
A TCTAGA

145. TriCasB_5RZmin+down

gactctc ttaAGG tgctgagt acgtccgcgtcgt gagagtc ggta actgtcccaggatgggtctg
ggatagg ctaa actcagca ggggaaagtgttaggggacctgccaccctacactttattggatggc
attcgatacccctaacgaagcctcggacttggaggagcacggttccccctcctcctcgtattagaccag

10

20

30

40

50

gaaccaactgtcctgacaacccccattggacctatgggagcggaccatgctatggacatggattccgaa
gacgaagcgggggcacacggaccccccgccgatagtgctcacttaacgtcaggcgaaccccttgaa
tcatcttg

146. HDV-28gu6_L8 [自己切断により最初の8ntが除去される]

GGGTAAAC GGCGGG AGTAA CTATGACTCTCTT AA AAAAGAGAATCATAGAA C
GTCAGCG GCCTCCACGC GGC CCCGCC GGA ACGCAGAGGAACACCCTGCGG CG
AA CGCTGACG C TCTAGA

147. HDV-28gu7_NL

GGCGGG AGTAA CTATGACTCTCTT AA AAAAGAGAATCATAGAC CGTCAGCG G
CCTCCACGC GGC CCCGCC GGA ACGCAGAGGAACACCCTGCGG CGAA CGCTGA
CG TCTAGA

10

148. HDV-28ac6_L8 [自己切断により最初の8ntが除去される]

GGGTAAAC GGCGGG AGTAA CTATGACTCTCTT AA AAAAGAGAATCATAGAA C
GTCAGCG GCCCCACAC GGC CCCGCC GGA ACGCAGAGGAACACCCTGCGG CG
AA CGCTGACG A TTCTAGA

149. HDV-28ac7_L8 [自己切断により最初の8ntが除去される]

GGGTAAAC GGCGGG AGTAA CTATGACTCTCTT AA AAAAGAGAATCATAGAC C
GTCAGCG GCCCCACAC GGC CCCGCC GGA ACGCAGAGGAACACCCTGCGG CG
AA CGCTGACG TCTAGA

150. HDV-28ac7_NL

GGCGGG AGTAA CTATGACTCTCTT AA AAAAGAGAATCATAGAC CGTCAGCG G
CCCCACAC GGC CCCGCC GGA ACGCAGAGGAACACCCTGCGG CGAA CGCTGA
CG TCTAGA

20

151. HDV-28gu8

GGCGGG AGTAA CTATGACTCTCTT AA CAAAGAGAATCATAGAA CGTCAGCG G
CCTCCACGC GGC CCCGCC GGA ACGCAGAGGAACACCCTGCGG CGAA CGCTGA
CG C TCTAGA

152. HDV-28ac8

GGCGGG AGTAA CTATGACTCTCTT AA CAAAGAGAATCATAGAA CGTCAGCG G
CCCCACAC GGC CCCGCC GGA ACGCAGAGGAACACCCTGCGG CGAA CGCTGA
CG G TCTAGA

30

153. HDV-28ac8_L8 [自己切断により最初の8ntが除去される]

GGGTAAAC GGCGGG AGTAA CTATGACTCTCTT AA CAAAGAGAATCATAGAA C
GTCAGCG GCCCCACAC GGC CCCGCC GGA ACGCAGAGGAACACCCTGCGG CG
AA CGCTGACG G TCTAGA

154. HDV-28ac9_L8 [自己切断により最初の8ntが除去される]

GGGTAAAC GGCGGG AGTAA CTATGACTCTCTT AA CAAAGAGAATCATAGAA C
GTCTGCG GCCCCACAC GGC CCCGCC GGA ACGCAGAGGAACACCCTGCGG CG
AA CGCAGACG G TCTAGA

40

200. OrLa 3'UTR

TGAGGGGGACAGCTGGGAGTCTCGGCATGATTACAAATCTTGCGCTGCACTCGGAT
GTCGTCCCCGTGACGGACACATTAATCCGGAAAGCGAGTGGTGACTCGCCTCAAG

201. TriCasB 3'UTR

taaaatctcctgaccaactagctcactgactaatttttaaactgtcctgtcttacttgttttacacgtgctc
tgtggcggggccatttacaccccgtcgcaacacaacctgtaaatacttgtgtatgtctgtttatgtccta
atttattatttttaaagacagatcttggccatgggtctcgccaaccaattaaagtcagtgatgagtgca
atgcgaggaagagacctaggcgtgtattttattgctggcatgcgggcgccggagccggtcatctgctat
ggggagcaatggccgggcgatacctccacgtggttcctgtgggtggcccgctcaggagcggttaacc
agcgaaactccgtaaagtccttcttacgagaagggaactccggttaaagatttttccaagcctgtacacg

50

tgattcccttggacaagcaaagtgtggttccctcgagagggcccagggtcaggagttcgcaatagtgg
gctgcaagagttcatgtctgggctacagtgtcaggacgaagagtgggtagtgatcgcaaaatcacgtg
aatagctacccccgcctggcaccactagacaacaacaaggggtacgacagctcttctgtcgaaagt
cgggcgcacacccgtaaaagg

202. ZoAl 3'UTR

taggtagtcacattgcactttctgtaacttgcactgggtgtgggatgtgggcctgggggtgtgggttatg
gggtatatatgtgggatattctgggtgggaatgtccattcactgtatgcctatctttttaataaaaagacg
gtagctagggttcgcgaagcagccacaagccaatagccagttaggtagctcatagtgggtaggtgaca
ggaacctttgactcagaacgcgtccattaacatctagaacggaccaaaccttcggacatgcaccgatta
accggatttgtccaaggtggacggggccacctttacttaacccggaaagggaacatatatagttatatgt
gttcgtaata

10

203. TaGu 3'UTR

taatt caggttat ttt agatgcttag tttttgtacc tttcttgttt tgtttaggat tttgatagt
g ttagtat tttt tata tttttg tacgattgca taatgttctt ttttatacag ttctgtttta ata
aaataga cgatagctag agacgttagg gcagccacaa gccagttagg tagcggatag tag
gtaggaa cagactttta ctatttcata acgcgtcaat taccacctga ttggaccaaa ttca
cgggat ttgtccaagg tggacggggc acctttactt aacccggaaa aggaacatat ata
atttatg tgtgttcgat aaa

204. GeFo 3'UTR

tagg gtagataatc tttgtatagt ggggggggat ctcatgtacc gggtttcttt tatttgattt
tcaataaaac agacggtagc taggttcgca aggagccac aagccaaaga taggtaggg
t gctcatagt agtagggaca gtgccttttg attcacaacg cgtcaatacc atctgacacg
gataccctta ccggacttgt catgatctcc cagacttgct caaggtggac gggccacctt ta
cttaaccc ggaaaaggaa catatatata ttatatgtgt tcggaaaa

20

205. TiGu 3'UTR

tagggg gcttggcatt tctcattgcc tgctcctgaa aggatattggg tcctgctcg cgtgg
taggc agaccattc gtccgagtag ggggcttggc agtntccatt gcctgtgccc gaaag
gacgt gggatcatctg gtctgtctgc ctacacctct ctagacttgt aacatctagt ctgtcaac
aa gatcaaaatt cttcacacag acgaccgagc ttgctcagtc ttcctgtacc cgcagaattt t
gctcttgct ctcttttggc tgtgtccttg acgtgggact attccatctc gtcccaaagt ccg
cgtccaa ttataccgga ttgacaaaag cggacggccc gctttataag ccggaaaagg tgcct
tgtaa aattgcaagg ttcattaaat ag

30

206. NavIB 3'UTR

TGACCTGAACAAAACGTGTTGTCTTGTCTTGTCTAAACTATTTATTCGAAATAAG
GGGAGGCTAACTGCCTGCAAGTTGAACGCGAAAGTTAGACCTTCCCACCTAAAGCC
CAAAAGTGATCGGGGAATGAATCCGCGGGTGACCCAGAGTTGGGTAAACCCTTGA
AACGTTGGAGAAGCGGAAGAGAGTCCCGCCACCGAGCATCGAGTGCTGCGGCGCCC
GAATGAAACCGATCGCGGATGGTGCAAGTCGTAGGACGGGGCACGACCTAAGCCTC
TGTCACGGCGGCGAAGCCAGGAATCACCATGCAAAGGTGTGAACTGGGGCGGATAC
CTCCACGGGGTTTCCCTGGGCATCGCGCGAGCGATGGCCAAAGTCCGCTTTCTCAG
CTACAAAACAAAAATGGTATGAGACTTCGTAAACACTAATTTTTTCCGAGCCTAGCA
GGCTCCCTTGACAACGCTTATGAATCTGGAAAAGGACACAAAGTGGAAAAAGCGCT
GATGGTGGACAAAAGTCAGTTGAGACTTGATATCAGTTGTTTTGACTAAGAATTTT
ATTATCGTTGACTTTTAAATATTTTATTATTGACTGTTAATACTGACTTGGGACC
AAGTCATCTCTGTTACCGGTACCGGTTCTGTCATCAAACCGGAAAGTCCGTCCC
ACGTAATGTGGTAGACGCAGGAG

40

207. PuPu 3'UTR

TAG G GTTCTCCAC CCTCCGGCTG ACGAACAGCT GGTAAACGGG GGGGCGGT
GG GGTGCCTCTC CAGCCGACTG ATACAGGAGT AAGGGACGGT GGGGTCGCAT
CCAGGAAGCG CAGCACCGCG ATGCCGAAAC TGATGTGCAG TATAACACAG AAA

50

GCCTAAA GGGCCAAAAG

208. Lipo 3'UTR

TAAATTTTGTCTCTTTCCCAATGATGTCTACTAGCACGCTGCCGAAGCTAGATAG
ATTGAGGAATCTGCGTAATCTGTAATGATTACGCCTCATGGGCATCTATCGGGTAGC
GTCGACCCTGACGTAAATTGGGT AATAAGAAATAT

209. BoMo 3'UTR

TGAGCCTTGACACAGTAGTCCAGCGGTAAGGGTGTAGATCAGGCCCGTCTGTTTCTC
CCCCGGAGCTCGCTCCCTTGGCTTCCCTTATATATTTTAACATCAGAAACAGACATT
AAACATCTACTGATCCAATTTGCGCGGCGTACGGCCACGATCGGGAGGGTGGGAAT
CTCGGGGGTCTTCCGATCCTAATCCATGATGATTACGACCTGAGTCACTAAAGACG
ATGGCATGATGATCCGGCGATGAAAA

10

210. DroSi 3'UTR

TAG CTAAACGTT TGGTTCAAAA CATTGCTTG CTGTCTTGGC ATAACATCAA
TAAAGGCATA AACATCGCAA AATAATGGTT ATATATAAAT GGCTATGAGG AT
GGTTTTAG TACGTAGGCG TTGCGGAACT TCGGTTTCTAGA TAGAGCAATG AATCG
TGCAT GCTAGGAAAA CTGACCACAC GCAGTGTGG CAGCCCTAGT ATCTTTTCG
AT AGATTTCCAT ACCTCCGCGA TCAAAAAAAA AAAAAAAAAA AAAAA

211. GaAc 3'UTR

GGAGGGGAGTAGGTCTCTACTCTGACCCGAAGGGCCCCCGTTTCAGACCTGATT
CTAGGCTACCTGTGCCTAATTGGGGGGGTCCCAAAGAGATGTTGTCTGTTGTAGAA
GGGTTTGCGCCACTGACTGCACGGAAGGGTGGGCCTCGACAGGTAGGGGTACATG
ACTCCGTGCTGCTCAGCAGACCCGCGCCTCTGAGACCGGGTAGGGCTACTTGAACA
AGCGACGCCCTGGTGTATGTCCGTATCCTAACCTGGTTTGGGAAAGCCGATACCGG
CAATGCCCGCCACAGGTGTGCGGCACCCACGGGATGACGTATGGGCCCCGGGGGA
CCTCATGGATACTCCACTGGACTTGACAAATCCTGGTGTACTGGATGCAGCGACGT
TGGTGACATAAGCAATCGCTAAGTCGGGGTAGGGGAGGTGGGGACCTCGGCACGGC
TG TAGGAACGGGTGTATGGGCTCCGGCAGCCGTCGTCACTCCCATACAACACAGGG
GCTGCATCCTGGTGGCCGGTGCTAGTTGGTTCTGGAAGCCCGCCCGGGCTGGTTTCG
CAGAAGCAGGGTGCGCCCAGGGTAGGTTTGGTATATCTGGGTCCGGTGCGATACCT
ATCGATGGGCAGCGAGGGCCGCCTCGTGACGCGCTGTGTGGAGCTGGAGCCGGCCT
GGGTATGAACAGTTCTTGCGGATGTGGCGTAGCTAGATAGTACCCGTGGTTGTGGG
CGTGGTGTGACCAAATGTTGTCCTGTGTGCACATAGGCCAAGGGTTACGTGGGTG
GCAGTCAGAAGCACCCGCACCTGGAAGTGATTGCCCCGGGATCCCGGCTCTCTGTG
AAGAGCTACCTTGAGGAAAGGTGTTCCGCTGGAACCTCAAGACCCTACAGTAGGGGA
TATCAACTGGCTTTGAGGTGCTGTGATTCCGGAACCAGGGCGAGGGCGAGTACTTA
GAGCATGTCCAAAAGCCCCGGGGAACGTTCCGGGGGCCTGCTTGGGTGCTTGGACCC
ACATCCGTAAAACGATGGATCTCGCGTCGGCGCTCGGGAGAACTTCCCGCATGAAC
GCTGATTGCATGTGAGAACGCCCCACGGCGGGCGGGGCAGGCGCTCCCCCTGGGTG
TAAGGCTCGGGGGGGTACGGCTCCGCTCTAAAAG

20

30

212. LeCoB 3'UTR

Tgataatcgctccatcctgcaactaattatgaatgcaaactgtttaagtgacattagtgatacttacctg
atacttaccttggtatatttgacctatacttaccttggtatctacctgacatatatttatctaaccacct
acctatgatgactcccgcggaaactctcacttaccttattaccacttggtcttttattttctcgttcctta
ttactttgttccttttggtgtagggttctctggtttttggaacggcttccttagccggaattttgtctgatgt
atcttgcttggtccttgaaatatacgaccaggcttgcgctcatttaggctctgggaaa

40

213. TriCan 3'UTR

Tagatgactgccctacccctttgctgccgaagaactactgaagactattgaactaaccaactgtgttaa
gtaagaactaatgcctctttttccctgcatgtatccctgatcagtgacttattttcttttcctgttgccg
ctttgtttagttatttcttttaattactagaattatcttttcgttctccgtctaattgcttttcgggtgtagg
aacggctacctaagctggaagtgggaagtgttttcaatgtactttgtgattatagaaatatatgacc

50

gaggtgcattgtttggcatttcctcgagaaa

214. CiIn 3'UTR

tgacagtaatatgaaaacatcacatctgaccggcacagaatcaccatgccgtaatgcacccaactaag
gattccaatgggtaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa

215. DrMe 3'UTR

tagctaaatcgtttgggttcaaaacatttgcttgctgtcttggcataacatcaataaaggcataaacatcg
caaaataat gggtataatt aaatggctat gaggatgggt ttagtacgtaggcgttgcggaacttcg
gtt catatagagc aatgaatcgt gcatgctagg aaaactgacc acacacagtg ttggcagacc
tagtatcttt cgaagatttc catacctccg cgatcaaaaa aaaaaaaaaaaaaaaaaa

216. DrNa 3'UTR

TAGCACCGAGGTTTACTTTCCTCATTGAAAGTCAACTACACTGGCCAACATCTATT
GTGTACAAAGTAAAGGGGAAATAGTCATATGACGTGCCATCCAAATATGGCATTGT
TGTGTGGGTAAATTCATTTCTCTAC AAAAAAAAAAAAAAAAAAAAAAAAAA

217. DrMerc 3'UTR

TAGCAGCTCTGAAAGCGTATCCTACAAACGCTGCTAACTTGCACAACATCTTAAAA
AGTTATAAAGAAAGCACACGGGTGTCTTAGCTTATCAGATATCTCTGGTATAGCGA
GTCATATAGCGCAAC AAAAAAAAAAAAAAAAAAAAAAAAAA

218. AdVa 3'UTR

tgaac tagtctcctt cttctattag tcagtctaataaatttttct tacattctac atctagttcc at
tattaaat tggatgatac agtgctatct ctgctacact caatgcttaa tcgtatgtta ttgacagt
ct gacacttgat tactcttacg acatatgcac tgtttgcttc agagaaacca ctgttcatat agt
gaagttc ctcagttttc tgttgatata ttcttctttc attctcgctt ctccttttct actgtgttct t
tttatcagt ttttgtgga aaaattgaga ataaataaag t

219. HyMaA 3'UTR

tga taaatgccaa aagttgcttg ggctaaatga tacgtacgct agaaaaagcg acttgctgca
cggatgacgg ttcatcagag cccgatatgt gcatgtcaag gcggcaggga gaatcactag
ttagctgtt ctttccatta cgacttacgc ggtaacgtg gcacgataga ttacaccag
gaaataatac gtgaagggtt ccacatata ctggagtta gatctatgag ggaaacattt
gtaataagtc agtctggtaa cctggcgccg ctgttgagtc aaattaacta tgtcaatact
cattaagtta tcgacttga tatggcatgg ggtgattccg cgttatatca aagtcaaaca
tgatgattgc aatgagaaac taccacgctt ggtcacgttt gtgaggagaa catctcattc
aagcctcccg gatgtcggca cccgctgaca tcttctggct tatgaaaatt ttcattaatt
tttgaagtc atgggaggct tgaaagc

220. DroSi3v2

GCTTGC TGCCTT GGCATAACATCAATA AAGGCA TAAAC TTCCC AAAA G AGG
ATGC TTTTAG TACGTAGGCG TTGCGGAACT TCGGTTTCTAGA TAGAGCAATGAAT
CGTGATGC TAGG AACA CCTAGT ATCTTTGA AGATTG GCATCCT CC GGGAA
CAAAA AAAA AAAA AAAA

221. DroSi3v3

GCTTGC TGCCTT GGCATAACATCAATA AAGGCA TAAAC TTCCC AAAA G AGG
ATGC TTTTAG TACGTAGGCG TTGCGGAACT TCGGTTTCTAGA TAGAGCAATGAAT
CGTGATGC TAGG AACA CCTAGT ATCTTTGA AGATTG GCATCCT CC GGGAA
CAAAG AAAAG AAAAG AAAA

222. GEFO217

GACGGTAGC TAGGTTTCGCA AGGCAGCCAC AAGCCAAAGA TAGGTAGGGT GCTC
ATAGTG AGTAGGGACA GTGCCTTTTG ATTCACAACG CGTCAATACC ATCTGAC
ACG GATACCCTTA CCGGACTTGT CATGATCTCC CAGACTTGTC CAAGGTGGAC
GGGCCACCTT TACTTAACCC GGAAAAGGAA CATATATTAA TTATATGTGT TCG
GAAAA

223. GEFO98

CCGGACTTGT CATGATCTCC CAGACTTGTC CAAGGTGGAC GGGCCACCTT TAC
TTAACCC GGAAAAGGAA CATATATTAA TTATATGTGT TCGGAAAA

224. GEFO68

CAAGGTGGAC GGGCCACCTT TACTTAACCC GGAAAAGGAA CATATATTAA TTA
TATGTGT TCGGAAAA

250. -28(32)

GGCGGGAGTAACTATGACTCTCTTAAGGTAGC

251. -28(30)

GGCGGGAGTAACTATGACTCTCTTAAGGTA

10

252. -28(28)

GGCGGGAGTAACTATGACTCTCTTAAGG

253. -28(27)

GGCGGGAGTAACTATGACTCTCTTAAG

254. -28(26)

GGCGGGAGTAACTATGACTCTCTTAA

255. -28(25)

GGCGGGAGTAACTATGACTCTCTTA

256. -28(24)

GGCGGGAGTAACTATGACTCTCTT

20

257. -28(23)

GGCGGGAGTAACTATGACTCTCT

258. -28(22)

GGCGGGAGTAACTATGACTCTC

259. -28(18)

GGCGGGAGTAACTATGAC

260. -28(14)

GGCGGGAGTAACTA

261. -24(24)

GGAGTAACTATGACTCTCTTAAGG

30

262. -24(22)

GGAGTAACTATGACTCTCTTAA

263. -24(21)

GGAGTAACTATGACTCTCTTA

264. -21(21)

GTAAGTAACTATGACTCTCTTAAGG

265. -21(19)

GTAAGTAACTATGACTCTCTTAA

266. -13(13)

GTAAGTAACTATGACTCTCTTAAGG

40

267. -13(11)

GTAAGTAACTATGACTCTCTTAAGG

268. -36(36)

GGGTAAACGGCGGGAGTAACTATGACTCTCTTAAGG

269. -36(34)

GGGTAAACGGCGGGAGTAACTATGACTCTCTTAA

270. -36(33)

GGGTAAACGGCGGGAGTAACTATGACTCTCTTA

271. -28(26)_L8

GGGTAAAC GGCGGG AGTAA CTATGACTCTCTTAA

50

272. -28(28)_L8

GGGTAAAC GGCGGG AGTAA CTATGACTCTCTTAAGG

273. -28(25)_L8

GGGTAAAC GGCGGG AGTAA CTATGACTCTCTTA

274. -28(24)_L8

GGGTAAAC GGCGGG AGTAA CTATGACTCTCTT

275. BoMo/OrLa R2 upstream R4

UUA

276. R25

GGGAGTAACTATGACTCTCTTAAG

10

278. OrLa_下流

TGGGGTCCGGCGAGAGCGCTGGTCTACGGACCAGGGGTGGCTGTGGGCAGGCTGCT
CCTCAGGCCAGTTGATTAGTTACGCATGGGCTGTACCTCCACGTGGTCCCGCTGGT
AACGACTTGTCGGCTAAATCAGCCCGCCACCATCTGGGATATGGTTGACCGTCTA
ACCCAGTACTCAGGTCACAAACAAAATGGGAACAGATACAGTGTATGTCGGCCAG
GACTACCCTTCTGGCTTATCAAAACGGGTACCAGCACGGT

279. TriCasA 下流

GAGCGTTTTTCGAAGTGGACTGTGTGACTGCGTTCCTCCCTTAGTTGCTATATCCGC
TTCGATTAAACATCTCACCTCGACGTATAAGATCATT

20

280. R4

UAGC

281. R4+1

GUAGC

282. R3

UAG

283. R5

UAGC

284. R10

UAGCCAAAG

285. R20

UAGCCAAAGCCUCGUCAUC

286. R3

UAG

287. R8

UAGCCAAA

288. R12

UAGCCAAAGCC

289. R16

UAGCCAAAGCCUCGU

30

40

290. 沈殿回収操作コントロール

GAAATTAATACGACTCACTATAGGGAGAACTCACTATTTGTTTTTCGCGCCCAGTTG
CAAAAAGTGTCGGCCACCATGGGAACAGATACAGTGTATGTCGG

291. R2 64 bp 標的部位にアニールしたオリゴ

Top: CAAGCGCGGGTAAACGGCGGGAGTAACTATGACTCTCTTAAGGTAGCCAAAT
GCCTCGTCATCT/3Cy5Sp/

Bottom: AGATGACGAGGCATTTGGCTACCTTAAGAGAGTCATAGTTACTCCCGCCG
TTTACCCGCGCTTG/3Cy5Sp/

50

292. R2 84 bp 標的部位にアニールしたオリゴ

Top: CAAGCGCGGGTAAACGGCGGGAGTAACTATGACTCTCTTAAGGTAGCCAAAT
GCCTCGTCATCTAATTAGTGACGCGCATGAAT/3Phos/

Bottom: ATTCATGCGCGTCACTAATTAGATGACGAGGCATTTGGCTACCTTAAGAG
AGTCATAGTTACTCCCGCCGTTTACCCGCGCTTG/3Phos/

293. R9 標的部位にアニールしたオリゴ

Top:

CAAATCGATCGTCAAACATGCGTTAAGGGGCGAAAGACTAATCGAACCATCTAGTA
GCTGGTTCCGTCCGAAGTT

Bottom:

AACTTCGGACGGAACCAGCTACTAGATGGTTCGATTAGTCTTTGCCCCCTTAACGC
ATGTTTGACGATCGATTTG

294. HoSa R8標的部位: [陰性コントロール TPRT]

Top:

TTTGGGTTCCGGGGGGAGTATGGTTGCAAAGCTGAAACTTAAAGGAATTGACGGAA
AGGAGTGG/3Phos/

Bottom:

CCACTCCTTTCCGTCAATTCCTTTAAGTTTCAGCTTTGCAACCATACTCCCCCGGA
ACCCAAA/3Cy5Sp/

295. HyMa R8標的部位:

Top:

TTTGGATTCCGGGGGGTAGTATGGTTGCAAAGCTGAAACTTAAACGGAAGGGCACCA
CCAGGAGT/3Phos/

Bottom:

ACTCCTGGTGGTGCCCTTCCGTTTAAAGTTTCAGCTTTGCAACCATACTACCCCCGGA
ATCCAAA/3Phos/

296. R2-tsrc標的部位

Top:

CAAGCGCGGGTAAACGGCGGGAGTAACTATGACTCTCTGGCTACCTTAAAATGCCT
CGTCATCT

Bottom:

AGATGACGAGGCATTTTAAAGGTAGCCAGAGAGTCATAGTTACTCCCGCCGTTTACC
CGCGCTTG

300. OrLa3_R4A22

TGAGGGGGACAGCTGGGAGTCTCGGCATGATTACAAATCTTGCGCTGCACTCGGAT
GTCGTCCCCGTGACGGACACATTAATCCGGAAAGCGAGTGGTGACTCGCCTCAAG T
AGC AAAAAAAAAAAAAAAAAAAAAA

301. OrLa3_R4A40

TGAGGGGGACAGCTGGGAGTCTCGGCATGATTACAAATCTTGCGCTGCACTCGGAT
GTCGTCCCCGTGACGGACACATTAATCCGGAAAGCGAGTGGTGACTCGCCTCAAG T
AGC AAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAA

302. OrLa3_R4A60

TGAGGGGGACAGCTGGGAGTCTCGGCATGATTACAAATCTTGCGCTGCACTCGGAT
GTCGTCCCCGTGACGGACACATTAATCCGGAAAGCGAGTGGTGACTCGCCTCAAG T
AGC AAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAA
AAAAAAAAA

303. OrLa3_R20A22

TGAGGGGGACAGCTGGGAGTCTCGGCATGATTACAAATCTTGCGCTGCACTCGGAT
GTCGTCCCCGTGACGGACACATTAATCCGGAAAGCGAGTGGTGACTCGCCTCAAG

10

20

30

40

50

UAGCCAAAUGCCUCGUCAUC

AAAAAAAAAAAAAAAAAAAAA

304. OrLa3_R20A40

TGAGGGGGACAGCTGGGAGTCTCGGCATGATTACAAATCTTGCGCTGCACTCGGAT
GTCGTCCCCGTGACGGACACATTAATCCGGAAAGCGAGTGGTGACTCGCCTCAAG

UAGCCAAAUGCCUCGUCAUC

AAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAA

305. OrLa3_R20A60

TGAGGGGGACAGCTGGGAGTCTCGGCATGATTACAAATCTTGCGCTGCACTCGGAT
GTCGTCCCCGTGACGGACACATTAATCCGGAAAGCGAGTGGTGACTCGCCTCAAG

UAGCCAAAUGCCUCGUCAUC

AAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAA
AAAAAA

306. TriCasB3_R4A22

taaaatctcctgaccaactagctcactgactaattttaactgtcctgtcttacttgttttacacgtgctc
tgtggcggggccatttacaccccgtcgcaacacaacctgtaaatacttgtgtatgtctgtttatgtccta
atttattattttaaacagatcttggccatgggtctcggccaaccaattaaagtcagtgatgagtcgca
atgcggagcaagagacctaggcgtgtatttattgttgcatgcggcgccggagccggtcatctgctat
ggggagcaatggccgggaggatacctccacgtgggttcctgtgggtggcccgtcgaggacggtaacc
agcgaaactccgtaaagtccttcttacgagaaggaaactccgggttaaagattttccaagcctgtacacg
tgattcccttgaacaagcaaaagtgtgggttcctcgagagggcccagggtcaggagttcgcaatagtg
gctgcaagagttcatgctgggctacagtgtcaggacgaagagtgggtagtgtatcgcaaaatcacgtg
aatagctacccccgcctggcaccactagacaacaacaaggggtacgacagctcttctgtcgaaagtt
cgggcgcacacccgtaaaagg TAGCAAAAAAAAAAAAAAAAAAAAAA

307. ZoAl3_R4A22

taggtagtcacattgcactttctgtaacttgcactgggtgtgggatgtgggcctgggggtgtgggttatg
gggtatatatgtgggatattctgggtgggaatgtccattcactgtatgcctatctttttaataaaaagacg
gtagctaggttcgcgaagcagccacaagccaatagccagttaggtagctcatagtgggtaggtgaca
ggaacctttgactcagaacgcgtccattaacatctagaacggaccaaaacttcggacatgcaccgatta
accggatttgtccaaggtggacggggccacctttacttaaccggaaagggaacatatatagttatatgt
gttcgtaata TAGC AAAAAAAAAAAAAAAAAAAAAA

308. ZoAl3_R4A40

taggtagtcacattgcactttctgtaacttgcactgggtgtgggatgtgggcctgggggtgtgggttatg
gggtatatatgtgggatattctgggtgggaatgtccattcactgtatgcctatctttttaataaaaagacg
gtagctaggttcgcgaagcagccacaagccaatagccagttaggtagctcatagtgggtaggtgaca
ggaacctttgactcagaacgcgtccattaacatctagaacggaccaaaacttcggacatgcaccgatta
accggatttgtccaaggtggacggggccacctttacttaaccggaaagggaacatatatagttatatgt
gttcgtaata TAGC AAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAA

309. ZoAl3_R4A60

taggtagtcacattgcactttctgtaacttgcactgggtgtgggatgtgggcctgggggtgtgggttatg
gggtatatatgtgggatattctgggtgggaatgtccattcactgtatgcctatctttttaataaaaagacg
gtagctaggttcgcgaagcagccacaagccaatagccagttaggtagctcatagtgggtaggtgaca
ggaacctttgactcagaacgcgtccattaacatctagaacggaccaaaacttcggacatgcaccgatta
accggatttgtccaaggtggacggggccacctttacttaaccggaaagggaacatatatagttatatgt
gttcgtaata TAGC AAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAA A
AAAAAAAAAAAAAAAAAAAAA

310. ZoAl3_R10A22

taggtagtcacattgcactttctgtaacttgcactgggtgtgggatgtgggcctgggggtgtgggttatg
gggtatatatgtgggatattctgggtgggaatgtccattcactgtatgcctatctttttaataaaaagacg
gtagctaggttcgcgaagcagccacaagccaatagccagttaggtagctcatagtgggtaggtgaca

ggaacctttgactcagaacgcgtccattaacatctagaacggaccaaacttcggacatgcaccgatta
accggatttgtccaaggtggacggggccacctttacttaacccggaaagggaacatatatagttatatgt
gttcgtaata UAGCCAAAUG AAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAA
AAAAA

311. ZoAl3_R10A40

taggtagtcacattgcactttctgtaacttgcactgggtgtgggatgtgggcctgggggtgtgggttatg
gggtatatatgtgggatattctgggtgggaatgtccattcactgtatgcctatctttttaataaaaagacg
gtagctaggtttcgcaagcagccacaagccaatagccagttaggtagctcatagtgggtagggtgaca
ggaacctttgactcagaacgcgtccattaacatctagaacggaccaaacttcggacatgcaccgatta
accggatttgtccaaggtggacggggccacctttacttaacccggaaagggaacatatatagttatatgt
gttcgtaata UAGCCAAAUG AAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAA

10

312. ZoAl3_R10A60

taggtagtcacattgcactttctgtaacttgcactgggtgtgggatgtgggcctgggggtgtgggttatg
gggtatatatgtgggatattctgggtgggaatgtccattcactgtatgcctatctttttaataaaaagacg
gtagctaggtttcgcaagcagccacaagccaatagccagttaggtagctcatagtgggtagggtgaca
ggaacctttgactcagaacgcgtccattaacatctagaacggaccaaacttcggacatgcaccgatta
accggatttgtccaaggtggacggggccacctttacttaacccggaaagggaacatatatagttatatgt
gttcgtaata UAGCCAAAUG AAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAA
AAAAA AAAAAAAAAAAAAAAAAAAAAA

20

313. GeFo3_R4A22

tagg gtagataatc tttgtatagt ggggggggat ctcattgtacc gggtttcttt tatttgattt
tcaataaaaac agacggtagc taggttcgca aggcagccac aagccaaaga taggtaggg
t gctcatagtg agtagggaca gtgccttttg attcacaacg cgtcaatacc atctgacacg
gataccctta ccggacttgt catgatctcc cagacttgtc caaggtggac gggccacctt ta
cttaaccc ggaaaaggaa catatatataa ttatatgtgt tcggaaaa TAGC AAAAAA
AAAAAAAAAAAAAAAAAAAA

315. GeFo3_R4

tagg gtagataatc tttgtatagt ggggggggat ctcattgtacc gggtttcttt tatttgattt
tcaataaaaac agacggtagc taggttcgca aggcagccac aagccaaaga taggtaggg
t gctcatagtg agtagggaca gtgccttttg attcacaacg cgtcaatacc atctgacacg
gataccctta ccggacttgt catgatctcc cagacttgtc caaggtggac gggccacctt ta
cttaaccc ggaaaaggaa catatatataa ttatatgtgt tcggaaaa UAGCCAAAUG

30

316. GeFo3_R10A15

tagg gtagataatc tttgtatagt ggggggggat ctcattgtacc gggtttcttt tatttgattt
tcaataaaaac agacggtagc taggttcgca aggcagccac aagccaaaga taggtaggg
t gctcatagtg agtagggaca gtgccttttg attcacaacg cgtcaatacc atctgacacg
gataccctta ccggacttgt catgatctcc cagacttgtc caaggtggac gggccacctt ta
cttaaccc ggaaaaggaa catatatataa ttatatgtgt tcggaaaa UAGCCAAAUG A
AAAAAAAAAAAAAAAAAAAAA

317. GeFo3_R10A22

tagg gtagataatc tttgtatagt ggggggggat ctcattgtacc gggtttcttt tatttgattt
tcaataaaaac agacggtagc taggttcgca aggcagccac aagccaaaga taggtaggg
t gctcatagtg agtagggaca gtgccttttg attcacaacg cgtcaatacc atctgacacg
gataccctta ccggacttgt catgatctcc cagacttgtc caaggtggac gggccacctt ta
cttaaccc ggaaaaggaa catatatataa ttatatgtgt tcggaaaa UAGCCAAAUG A
AAAAAAAAAAAAAAAAAAAAA

40

318. GeFo3_R20

tagg gtagataatc tttgtatagt ggggggggat ctcattgtacc gggtttcttt tatttgattt
tcaataaaaac agacggtagc taggttcgca aggcagccac aagccaaaga taggtaggg
t gctcatagtg agtagggaca gtgccttttg attcacaacg cgtcaatacc atctgacacg

50

gataccctta ccggaacttgt catgatctcc cagacttgtc caaggtggac gggccacctt ta
 cttaaccc ggaaaaggaa catatatataa ttatatgtgt tcggaaaa UAGCCAAAUGC
 CUCGUCAUC

319. GeFo3_R20A15

tagg gtagataatc tttgtatagt ggggggggat ctcatgtacc gggtttcttt tatttgattt
 tcaataaaac agacggtagc taggttcgca aggcagccac aagccaaaga taggtaggg
 t gctcatagt agtagggaca gtgccttttg attcacaacg cgtcaatacc atctgacacg
 gataccctta ccggaacttgt catgatctcc cagacttgtc caaggtggac gggccacctt ta
 cttaaccc ggaaaaggaa catatatataa ttatatgtgt tcggaaaa UAGCCAAAUGC
 CUCGUCAUC AAAAAAAAAAAAAAAAAA

10

320. GeFo3_R20A22

tagg gtagataatc tttgtatagt ggggggggat ctcatgtacc gggtttcttt tatttgattt
 tcaataaaac agacggtagc taggttcgca aggcagccac aagccaaaga taggtaggg
 t gctcatagt agtagggaca gtgccttttg attcacaacg cgtcaatacc atctgacacg
 gataccctta ccggaacttgt catgatctcc cagacttgtc caaggtggac gggccacctt ta
 cttaaccc ggaaaaggaa catatatataa ttatatgtgt tcggaaaa UAGCCAAAUGC
 CUCGUCAUC AAAAAAAAAAAAAAAAAA

321. TaGu3_R4A22

taatt caggttattt agatgcttag tttttgtacc tttcttgttt tgtttaggat tttgatagt
 g ttagtatttt tatatttttg tacgattgca taatgttctt ttttatacag ttctgtttta ata
 aataga cgatagctag agacgttagg gcagccacaa gccagttagg tagcggatag tag
 gtaggaa cagactttta ctatttcata acgctgaat taccacctga tttggaccaa ttca
 cgggat ttgtccaagg tggacgggcc acctttactt aaccgggaaa aggaacatat ata
 atttatg tgtgttcgat aaa TAGCAAAAAAAAAAAAAAAAAA

20

322. TiGu3_R4A22

tagggg gcttggcatt tctcattgcc tgctcctgaa aggatattggg tcctgcgtcg cgtgg
 taggc agaccattc gtccgagtag ggggcttggc agtntccatt gcctgtgccc gaaag
 gacgt gggatcatctg gtctgtctgc ctacacctct ctagacttgt aacatctagt ctgtcaac
 aa gatcaaaatt cttcacacag acgaccgagc ttgctcagtc ttcctgtacc cgcagaattt t
 gctcttgct ctcttttggc tgtgtccttg acgtgggact attccatctc gtcccaaagt ccg
 cgtccaa ttataccgga ttgacaaag cggacggccc gctttataag ccggaaaagg tgcct
 tgtaa aattgcaagg ttcattaaat ag TAGC AAAAAAAAAAAAAAAAAA

30

323. BoMo3_R4A22

TGAGCCTTGACACAGTAGTCCAGCGGTAAGGGTGTAGATCAGGCCCGTCTGTTTCTC
 CCCCGGAGCTCGCTCCCTTGCTTCCCTTATATATTTTAACATCAGAAACAGACATT
 AAACATCTACTGATCCAATTTTCGCCGGCGTACGGCCACGATCGGGAGGGTGGGAAT
 CTCGGGGGTCTTCCGATCCTAATCCATGATGATTACGACCTGAGTCACTAAAGACG
 ATGGCATGATGATCCGGCGATGAAAA TAGC AAAAAAAAAAAAAAAAAA

324. DroSi3_R4A22

TAG CTAACGTT TGGTTCAAAA CATTTGCTTG CTGTCTTGGC ATAACATCAA
 TAAAGGCATA AACATCGCAA AATAATGGTT ATATATAAAT GGCTATGAGG AT
 GGTTTTAG TACGTAGGCG TTGCGGAACT TCGGTTTCTAGA TAGAGCAATG AATCG
 TGCAT GCTAGGAAAA CTGACCACAC GCAGTGTTGG CAGCCCTAGT ATCTTTTCG
 AT AGATTTCCAT ACCTCCGCGA TCAAAAAAAAAA AAAAAAAAAA AAAAA TAGC A
 AAAAAAAAAAAAAAAAAA

40

325. DroSi3_R20A22

TAG CTAACGTT TGGTTCAAAA CATTTGCTTG CTGTCTTGGC ATAACATCAA
 TAAAGGCATA AACATCGCAA AATAATGGTT ATATATAAAT GGCTATGAGG AT
 GGTTTTAG TACGTAGGCG TTGCGGAACT TCGGTTTCTAGA TAGAGCAATG AATCG
 TGCAT GCTAGGAAAA CTGACCACAC GCAGTGTTGG CAGCCCTAGT ATCTTTTCG

50

AT AGATTTCCAT ACCTCCGCGA TCAAAAAAAAAA AAAAAAAAAA AAAAA UAGCC
AAUAGCCUCGUCAUC AAAAAAAAAAAAAAAAAAAAAA

326. DroSi3_R20

TAG CTAACGTT TGGTTCAAAA CATTGCTTG CTGTCTTGGC ATAACATCAA
TAAAGGCATA AACATCGCAA AATAATGGTT ATATATAAAT GGCTATGAGG AT
GGTTTTAG TACGTAGGCG TTGCGGAACT TCGGTTTCAAGA TAGAGCAATG AATCG
TGCAT GCTAGGAAAA CTGACCACAC GCAGTGTTGG CAGCCCTAGT ATCTTTTCG
AT AGATTTCCAT ACCTCCGCGA TCAAAAAAAAAA AAAAAAAAAA AAAAA UAGCC
AAUAGCCUCGUCAUC

327. DroSi3_R4

TAG CTAACGTT TGGTTCAAAA CATTGCTTG CTGTCTTGGC ATAACATCAA
TAAAGGCATA AACATCGCAA AATAATGGTT ATATATAAAT GGCTATGAGG AT
GGTTTTAG TACGTAGGCG TTGCGGAACT TCGGTTTCAAGA TAGAGCAATG AATCG
TGCAT GCTAGGAAAA CTGACCACAC GCAGTGTTGG CAGCCCTAGT ATCTTTTCG
AT AGATTTCCAT ACCTCCGCGA TCAAAAAAAAAA AAAAAAAAAA AAAAA TAGC

328. DroSi3v2_R20

GCTTGC TGCCTT GGCATAACATCAATA AAGGCA TAAAC TTCCC AAAA G AGG
ATGC TTTTAG TACGTAGGCG TTGCGGAACT TCGGTTTCAAGA TAGAGCAATGAAT
CGTGTCATGC TAGG AACA CCTAGT ATCTTTGA AGATTT GCATCCT CC GGGAA
CAAAAAAAAAAAAAAAAAAAAAA UAGCCAAUAGCCUCGUCAUC

329. DroSi3v3_R20

GCTTGC TGCCTT GGCATAACATCAATA AAGGCA TAAAC TTCCC AAAA G AGG
ATGC TTTTAG TACGTAGGCG TTGCGGAACT TCGGTTTCAAGA TAGAGCAATGAAT
CGTGTCATGC TAGG AACA CCTAGT ATCTTTGA AGATTT GCATCCT CC GGGAA
CAAAGAAAAGAAAAGAAAAA UAGCCAAUAGCCUCGUCAUC

330. R25-TCB3-R4

GGG AGT AAC TAT GAC TCT CTT AAG GTAAAATCTCCTGACCAACTAGCTCAC
TGACTAATTTTAAACTGTCTGTCTTACTTGTTTTACACGTGCTCTGTGGCGGGGCC
ATTTACACCCCGTCGCAACACAACCTGTAAATACTTGTGTATGTCTGTTTATGTCTCCT
AATTTATTATTTTAAACAGATCTTGCCATGGTCTCGGCCAACCAATTAAAGTCAG
TGATGCGAGTCGCAATGCGGAGCAAGAGACCTAGGCGTGTATTTATTGCTGGCATG
CGGCGCCGGAGCCGGTCATCTGCTATGGGGAGCAATGGCCGGGCGGATACCTCCAC
GTGGTTCCCTGTGGGTGGCCCGTCGAGGACGGTAACCAGCGAAACTCCGTAAAGTC
CTTCTTACGAGAAGGAACTCCGGTTAAAGATTTTTTCCAAGCCTGTACACGTGATTCT
CCTTGGAACAAGCAAAGTGTGGTTCCCTCGAGAGGGGCCAGGTCAGGAGTTTCGCAA
TAGTGGGCTGCAAGAGTTCATGCTGGGCTACAGTGTCAGGACGAAGAGTGGGTAGT
GATCGCAAAATCACGTGAATAGCTACCCCCCGCCTGGCACCACTAGACAACAACAA
GGGGTACGACAGCTCTTCTGTGAAAGTTCGGGCGCACACCCGTAAAAGGTAGC

331. R25-TCB3-R10

GGGAGTAACTATGACTCTCTTAAGGTAAAATCTCCTGACCAACTAGCTCACTGACT
AATTTTAAACTGTCTGTCTTACTTGTTTTACACGTGCTCTGTGGCGGGGCCATTTA
CACCCCGTCGCAACACAACCTGTAAATACTTGTGTATGTCTGTTTATGTCCTAATTT
ATTATTTTAAACAGATCTTGCCATGGTCTCGGCCAACCAATTAAAGTCAGTGATG
CGAGTCGCAATGCGGAGCAAGAGACCTAGGCGTGTATTTATTGCTGGCATGCGGCG
CCGGAGCCGGTCATCTGCTATGGGGAGCAATGGCCGGGCGGATACCTCCACGTGGT
TCCCTGTGGGTGGCCCGTCGAGGACGGTAACCAGCGAAACTCCGTAAAGTCCTTCT
TACGAGAAGGAACTCCGGTTAAAGATTTTTTCCAAGCCTGTACACGTGATTCCCTTG
GAACAAGCAAAGTGTGGTTCCCTCGAGAGGGGCCAGGTCAGGAGTTTCGCAATAGTG
GGCTGCAAGAGTTCATGCTGGGCTACAGTGTCAGGACGAAGAGTGGGTAGTGATCG

10

20

30

40

50

CAAAATCACGTGAATAGCTACCCCCCGCCTGGCACCAGTACACAACAACAAGGGGT
ACGACAGCTCTTCTGTCGAAAGTTCTGGGCGCACACCCGTAAAAGGTAGc caaatg

332. R25-TCB3-R4-PA

GGGAGTAACTATGACTCTCTTAAGGTAAAATCTCCTGACCAACTAGCTCACTGACT
AATTTTAAACTGTCCTGTCTTACTTGTTTTACACGTGCTCTGTGGCGGGGCCATTTA
CACCCCGTCGCAACACAACCTGTAAATACTTGTGTATGTCTGTTTATGTCCTAATTT
ATTATTTTAAACAGATCTTGCCATGGTCTCGGCCAACCAATTAAAGTCAGTGATG
CGAGTCGCAATGCGGAGCAAGAGACCTAGGCGTGTATTTATTGCTGGCATGCGGCG
CCGGAGCCGGTCATCTGCTATGGGGAGCAATGGCCGGGCGGATACCTCCACGTGGT
TCCCTGTGGGTGGCCCGTCGAGGACGGTAACCAGCGAAACTCCGTAAAGTCCTTCT
TACGAGAAGGAACTCCGGTTAAAGATTTTTTCCAAGCCTGTACACGTGATTCCCTTG
GAACAAGCAAAGTGTGGTTCCCTCGAGAGGGGCCAGGTCAGGAGTTCGCAATAGTG
GGCTGCAAGAGTTCATGCTGGGCTACAGTGTGAGGACGAAGAGTGGGTAGTGATCG
CAAAATCACGTGAATAGCTACCCCCCGCCTGGCACCAGTACACAACAACAAGGGGT
ACGACAGCTCTTCTGTCGAAAGTTCTGGGCGCACACCCGTAAAAGGTAGc aaaaa aa
aaa aaaaa aaaaa aa

10

333. R25*-TCB5_TCB3-R4 [5'側の12ntはRZの自己切断により除去される]

GGGAGTAACTAT GACTCTCTTAAGGGAAGACCCCGCCCATGAGGCTTGGAGAGTG
TGATCCTGATCAGATCACACTTGAAAAGTTATGCTGAGTACGTCCGCGTCGTGAGA
GTCGGTAACTGTCCCAGGATGGTCTGGGATAGGCTAAACCTCAGCAGGGGAAAGTT
GTAGGGGCCTGCCACCCCTACACTTTATTGGTATGGCATTTCGATACCCCTAACGAA
GCCTCGGACTTGGAGGAGCACGGTTCCCTCCTCCTCGTATTAGACCAGGAACCAA
CTGTCCTGACAACCCCATTTGGACCTATGGGAGCGGACCATGCTATGGACATGGATT
CCGAAGACGAAGCGGGGGCACACGGACCCCGCGGATAGTGCTCACTTAACGTCA
GGCGAACCCCTTGAAATCATCTTGTAATAATCTCCTGACCAACTAGCTCACTGACTA
ATTTTAAACTGTCCTGTCTTACTTGTTTTACACGTGCTCTGTGGCGGGGCCATTTAC
ACCCCGTCGCAACACAACCTGTAAATACTTGTGTATGTCTGTTTATGTCCTAATTTA
TTATTTTAAACAGATCTTGCCATGGTCTCGGCCAACCAATTAAAGTCAGTGATGC
GAGTCGCAATGCGGAGCAAGAGACCTAGGCGTGTATTTATTGCTGGCATGCGGCGC
CGGAGCCGGTCATCTGCTATGGGGAGCAATGGCCGGGCGGATACCTCCACGTGGTT
CCCTGTGGGTGGCCCGTCGAGGACGGTAACCAGCGAAACTCCGTAAAGTCCTTCTT
ACGAGAAGGAACTCCGGTTAAAGATTTTTTCCAAGCCTGTACACGTGATTCCCTTGG
AACAAGCAAAGTGTGGTTCCCTCGAGAGGGGCCAGGTCAGGAGTTCGCAATAGTGG
GCTGCAAGAGTTCATGCTGGGCTACAGTGTGAGGACGAAGAGTGGGTAGTGATCGC
AAAATCACGTGAATAGCTACCCCCCGCCTGGCACCAGTACACAACAACAAGGGGTA
CGACAGCTCTTCTGTCGAAAGTTCTGGGCGCACACCCGTAAAAGGTAGc

20

30

334. R25*-TCB5_TCB3-R10 [5'側の12ntはRZの自己切断により除去される]

GGGAGTAACTA TGACTCTCTTAAGGGAAGACCCCGCCCATGAGGCTTGGAGAGTG
TGATCCTGATCAGATCACACTTGAAAAGTTATGCTGAGTACGTCCGCGTCGTGAGA
GTCGGTAACTGTCCCAGGATGGTCTGGGATAGGCTAAACCTCAGCAGGGGAAAGTT
GTAGGGGCCTGCCACCCCTACACTTTATTGGTATGGCATTTCGATACCCCTAACGAA
GCCTCGGACTTGGAGGAGCACGGTTCCCTCCTCCTCGTATTAGACCAGGAACCAA
CTGTCCTGACAACCCCATTTGGACCTATGGGAGCGGACCATGCTATGGACATGGATT
CCGAAGACGAAGCGGGGGCACACGGACCCCGCGGATAGTGCTCACTTAACGTCA
GGCGAACCCCTTGAAATCATCTTGTAATAATCTCCTGACCAACTAGCTCACTGACTA
ATTTTAAACTGTCCTGTCTTACTTGTTTTACACGTGCTCTGTGGCGGGGCCATTTAC
ACCCCGTCGCAACACAACCTGTAAATACTTGTGTATGTCTGTTTATGTCCTAATTTA
TTATTTTAAACAGATCTTGCCATGGTCTCGGCCAACCAATTAAAGTCAGTGATGC
GAGTCGCAATGCGGAGCAAGAGACCTAGGCGTGTATTTATTGCTGGCATGCGGCGC
CGGAGCCGGTCATCTGCTATGGGGAGCAATGGCCGGGCGGATACCTCCACGTGGTT

40

50

CCCTGTGGGTGGCCCGTCGAGGACGGTAACCAGCGAAACTCCGTAAAGTCCTTCTT
ACGAGAAGGAACTCCGGTTAAAGATTTTTTCCAAGCCTGTACACGTGATTCCCTTGG
AACAAGCAAAGTGTGGTTCCTCGAGAGGGCCCAGGTCAGGAGTTCGCAATAGTGG
GCTGCAAGAGTTTCATGCTGGGCTACAGTGTGAGGACGAAGAGTGGGTTAGTGATCGC
AAAATCACGTGAATAGCTACCCCCCGCCTGGCACCCTAGACAACAACAAGGGGTA
CGACAGCTCTTCTGTGCAAAGTTCGGGCGCACACCCCGTAAAAGGTAGc c a a a t g
335. R25*-TCB5_TCB3-R4-PA [5'側の12ntはRZの自己切断により除去される]
GGGAGTAACTAT GACTCTCTTAAGGGAAGACCCCGCCCATGAGGCTTGGAGAGTG
TGATCCTGATCAGATCACACTTGAAAAGTTATGCTGAGTACGTCCGCGTCGTGAGA
GTCGGTAACTGTCCCAGGATGGTCTGGGATAGGCTAAACCTCAGCAGGGGAAAGTT
GTAGGGGCTGCCACCCCTACACTTTATTGGTATGGCATTTCGATACCCCTAACGAA
GCCTCGGACTTGGAGGAGCACGGTTCCCCTCCTCCTCGTATTAGACCAGGAACCAA
CTGTCCTGACAACCCCATTTGGACCTATGGGAGCGGACCATGCTATGGACATGGATT
CCGAAGACGAAGCGGGGGCACACGGACCCCGCCGATAGTGCTCACTTAACGTCA
GGCGAACCCCTTGAAATCATCTTGTAATAATCTCCTGACCAACTAGCTCACTGACTA
ATTTTAACTGTCCTGTCTTACTTGTTTTACACGTGCTCTGTGGCGGGGCCATTTAC
ACCCCGTCGCAACACAACCTGTAAATACTTGTGTATGTCTGTTTATGTCCTAATTTA
TTATTTTAAACAGATCTTGGCCATGGTCTCGGCCAACCAATTAAAGTCAGTGATGC
GAGTCGCAATGCGGAGCAAGAGACCTAGGCGTGTATTTATTGCTGGCATGCGGGCGC
CGGAGCCGGTCATCTGCTATGGGGAGCAATGGCCGGGCGGATACCTCCACGTGGTT
CCCTGTGGGTGGCCCGTCGAGGACGGTAACCAGCGAAACTCCGTAAAGTCCTTCTT
ACGAGAAGGAACTCCGGTTAAAGATTTTTTCCAAGCCTGTACACGTGATTCCCTTGG
AACAAGCAAAGTGTGGTTCCTCGAGAGGGCCCAGGTCAGGAGTTCGCAATAGTGG
GCTGCAAGAGTTTCATGCTGGGCTACAGTGTGAGGACGAAGAGTGGGTTAGTGATCGC
AAAATCACGTGAATAGCTACCCCCCGCCTGGCACCCTAGACAACAACAAGGGGTA
CGACAGCTCTTCTGTGCAAAGTTCGGGCGCACACCCCGTAAAAGGTAGc a a a a a a a a
a a a a a a a a a a a a

10

20

30

40

50

336. R25*-TCB5_TCB3-PA [5'側の12ntはRZの自己切断により除去される]
GGGAGTAACTAT GACTCTCTTAAGGGAAGACCCCGCCCATGAGGCTTGGAGAGTG
TGATCCTGATCAGATCACACTTGAAAAGTTATGCTGAGTACGTCCGCGTCGTGAGA
GTCGGTAACTGTCCCAGGATGGTCTGGGATAGGCTAAACCTCAGCAGGGGAAAGTT
GTAGGGGCTGCCACCCCTACACTTTATTGGTATGGCATTTCGATACCCCTAACGAA
GCCTCGGACTTGGAGGAGCACGGTTCCCCTCCTCCTCGTATTAGACCAGGAACCAA
CTGTCCTGACAACCCCATTTGGACCTATGGGAGCGGACCATGCTATGGACATGGATT
CCGAAGACGAAGCGGGGGCACACGGACCCCGCCGATAGTGCTCACTTAACGTCA
GGCGAACCCCTTGAAATCATCTTGTAATAATCTCCTGACCAACTAGCTCACTGACTA
ATTTTAACTGTCCTGTCTTACTTGTTTTACACGTGCTCTGTGGCGGGGCCATTTAC
ACCCCGTCGCAACACAACCTGTAAATACTTGTGTATGTCTGTTTATGTCCTAATTTA
TTATTTTAAACAGATCTTGGCCATGGTCTCGGCCAACCAATTAAAGTCAGTGATGC
GAGTCGCAATGCGGAGCAAGAGACCTAGGCGTGTATTTATTGCTGGCATGCGGGCGC
CGGAGCCGGTCATCTGCTATGGGGAGCAATGGCCGGGCGGATACCTCCACGTGGTT
CCCTGTGGGTGGCCCGTCGAGGACGGTAACCAGCGAAACTCCGTAAAGTCCTTCTT
ACGAGAAGGAACTCCGGTTAAAGATTTTTTCCAAGCCTGTACACGTGATTCCCTTGG
AACAAGCAAAGTGTGGTTCCTCGAGAGGGCCCAGGTCAGGAGTTCGCAATAGTGG
GCTGCAAGAGTTTCATGCTGGGCTACAGTGTGAGGACGAAGAGTGGGTTAGTGATCGC
AAAATCACGTGAATAGCTACCCCCCGCCTGGCACCCTAGACAACAACAAGGGGTA
CGACAGCTCTTCTGTGCAAAGTTCGGGCGCACACCCCGTAAAAGG a a a a a a a a
a a a a a a a a

337. R25*-TCB5_TCB3-R10PA [5'側の12ntはRZの自己切断により除去される]
GGGAGTAACTAT GACTCTCTTAAGGGAAGACCCCGCCCATGAGGCTTGGAGAGTG

TGATCCTGATCAGATCACACTTGAAAAGTTATGCTGAGTACGTCCGCGTCGTGAGA
GTCGGTAACTGTCCCAGGATGGTCTGGGATAGGCTAAACCTCAGCAGGGGAAAGTT
GTAGGGGCCTGCCACCCCTACACTTTATTGGTATGGCATTTCGATACCCCTAACGAA
GCCTCGGACTTGAGGAGCACGGTTCCCCCTCCTCCTCGTATTAGACCAGGAACCAA
CTGTCCTGACAACCCCATTTGGACCTATGGGAGCGGACCATGCTATGGACATGGATT
CCGAAGACGAAGCGGGGGCACACGGACCCCCCGCGATAGTGCTCACTTAACGTCA
GGCGAACCCCTTGAAATCATCTTGTAATACTCCTGACCAACTAGCTCACTGACTA
ATTTTAAACTGTCCTGTCTTACTTGTTTTACACGTGCTCTGTGGCGGGGCCATTTAC
ACCCCGTCGCAACACAACCTGTAAATACTTGTGTATGTCTGTTTATGTCCTAATTTA
TTATTTTAAACAGATCTTGGCCATGGTCTCGGCCAACCAATTAAAGTCAGTGATGC
GAGTCGCAATGCGGAGCAAGAGACCTAGGCGTGTATTTATTGCTGGCATGCGGGCGC
CGGAGCCGGTCATCTGCTATGGGGAGCAATGGCCGGGCGGATACCTCCACGTGGTT
CCCTGTGGGTGGCCCGTCGAGGACGGTAACCAGCGAAACTCCGTAAAGTCCTTCTT
ACGAGAAGGAACTCCGGTTAAAGATTTTTCCAAGCCTGTACACGTGATTCCCTTGG
AACAAGCAAAGTGTGGTTCCCTCGAGAGGGGCCAGGTCAGGAGTTCGCAATAGTGG
GCTGCAAGAGTTTATGCTGGGCTACAGTGTGAGGACGAAGAGTGGGTAGTGATCGC
AAAATCACGTGAATAGCTACCCCCCGCCTGGCACCACTAGACAACAACAAGGGGTA
CGACAGCTCTTCTGTGCGAAAGTTCGGGCGCACACCCGTAAAAGG TAGccaaatg aa
aaa aaaaa aaaaa aaaaa aa

10

338. TCB3-R4-PA

20

TAAAATCTCCTGACCAACTAGCTCACTGACTAATTTTAAACTGTCCTGTCTTACTTG
TTTTACACGTGCTCTGTGGCGGGGCCATTTACACCCCGTCGCAACACAACCTGTAA
ATACTTGTGTATGTCTGTTTATGTCCTAATTTATTATTTTAAACAGATCTTGGCCAT
GGTCTCGGCCAACCAATTAAAGTCAGTGATGCGAGTCGCAATGCGGAGCAAGAGAC
CTAGGCGTGTATTTATTGCTGGCATGCGGCGCCGGAGCCGGTCATCTGCTATGGGG
AGCAATGGCCGGGCGGATACCTCCACGTGGTTCCCTGTGGGTGGCCCGTCGAGGAC
GGTAACCAGCGAAACTCCGTAAAGTCCTTCTTACGAGAAGGAACTCCGGTTAAAGA
TTTTTCCAAGCCTGTACACGTGATTCCCTTGGAACAAGCAAAGTGTGGTTCCCTCG
AGAGGGGCCAGGTCAGGAGTTCGCAATAGTGGGCTGCAAGAGTTTATGCTGGGCTA
CAGTGTGAGGACGAAGAGTGGGTAGTGATCGCAAAATCACGTGAATAGCTACCCCC
CGCCTGGCACCACTAGACAACAACAAGGGGTACGACAGCTCTTCTGTGCGAAAGTTC
GGGCGCACACCCGTAAAAGGTAGc aaaaa aaaaa aaaaa aaaaa aa

30

339. TCA5_OrLa3 [最初の54ntは自己切断により除去される]

gggagacgggucaaccgcguaggagcggugaccgGAAUUCGGCGGGAGUAACUAU GACUC
UCUUAAGGAGUCAUAGAGCCAGAACCUCUCGUGGUCCCGCUGGGGCACAGGGGAUU
AAUUUUUCUGUGGC AAAUUUGACUGGCUUCAGAGAGCGUUUUUCGAAGUGGACU
GUGUGACUGCGUCCCCCUUAGUUGCUAUAUCCGCUUCGAUUAACAUCUCACCU
CGACGUUAUAGAUAUUUCUAGAGCGGCCGCUAGAGGGGGACAGCUGGGGAGUCUCG
GCAUGAUUAACAAUUCUUGCGCUGCACUCGGAUGUCGUCCCCGUGACGGACACAUA
AAUCCGGAAGCGAGUGGUGACUCGCCUCAAGUAGCAAAAAAAAAAAAAAAAAAAAA
AAA

40

340. TCA5_ZoAl3 [最初の54ntは自己切断により除去される]

gggagacgggucaaccgcguaggagcggugaccgGAAUUCGGCGGGAGUAACUAU GACUC
UCUUAAGGAGUCAUAGAGCCAGAACCUCUCGUGGUCCCGCUGGGGCACAGGGGAUU
AAUUUUUCUGUGGC AAAUUUGACUGGCUUCAGAGAGCGUUUUUCGAAGUGGACU
GUGUGACUGCGUCCCCCUUAGUUGCUAUAUCCGCUUCGAUUAACAUCUCACCU
CGACGUUAUAGAUAUUUCUAGAGCGGCCGCUagguagucacauugcacuuucuguaac
uugcacuggguguggggaugugggccuggggugugggguuaugggguauauaugugggauauuc
uggugggaauguccauucacuguaugccuauuuuuuaauaaaaagacggguagcuagguucgc
gaagcagccacaagccaauagccaguuagguagcucauagugggguaggugacaggaaccuuuga

50

cucagaacgcguccauuaacaucuagaacggaccaaacuucggacaugcaccgauuaaccggau
uuguccaagguggacgggccaccuuuacuuuacccggaaagggaacauauauaguuaauaugug
uucguaauaUAGCAAAAAAAAAAAAAAAAAAAAAA

341. TCA5_TCB3 [最初の54ntは自己切断により除去される]

gggagacgggucaaccgcguaaggagcggugaccgGAAUUCGGCGGGAGUAACUAU GACUC
UCUUAAGGAGUCAUAGAGCCAGAACCUCUCGUGGUCCCGCUGGGGCACAGGGGAUU
AAUUUUUCUGUGGCAAAUUUGACUGGCUUCAGAGAGCGUUUUUCGAAGUGGACU
GUGUGACUGCGUUCUUUUUUAGUUGCUAUAUCCGCUUCGAUUAACAUCUCACCU
CGACGUUAUAGAUAUUUCUAGAGCGGCCGCUAAAAUCUCCUGACCAACUAGCUC
ACUGACUAAUUUUUAAACUGUCCUGUCUUAUUGUUUUACACGUGGCUCUGUGGGCG
GGGCCAUUUACACCCCGUCGCAACACAACCUGUAAAUACUUGUGUAUGUCUGUUU
AUGUCCUAAUUUAUUUAUUUAAACAGAUCUUGGCCAUGGUCUCGGCCAACCAAUU
AAAGUCAGUGAUGCGAGUCGCAAUGCGGAGCAAGAGACCUAGGCGUGUAUUUAUU
GCUGGCAUGCGGGCGCCGGAGCCGGUCAUCUGCUAUGGGGAGCAAUGGCCGGGGCGG
AUACCUCCACGUGGUUCCUGUGGGUGGCCCGUCGAGGACGGUAACCAGCGAAAC
UCCGUAAAGUCCUUCUUAACGAGAAGGAACUCCGGUUAAGAUUUUCCAAGCCUG
UACACGUGAUUCCCUUGGAACAAGCAAAGUGUGGUUCCUCGAGAGGGCCCAGGU
CAGGAGUUCGCAAUAGUGGGCUGCAAGAGUUAUGCUGGGCUACAGUGUCAGGAC
GAAGAGUGGGUAGUGAUCGCAAAAUACGUGAAUAGCUACCCCCCGCCUGGCACC
ACUAGACAACAACAAGGGGUACGACAGCUCUUCUGUCGAAAGUUCGGGCGCACAC
CCGUAAAAGGUAGCAAAAAAAAAAAAAAAAAAAAAA

10

20

343. OrLa5L_OrLa3 [最初の39ntは自己切断により除去される]

gggagacgggucaaccgcguaaggagcggugaccgGAAUUC GGCGGGAGUAACUAUGACUC
UCUUAACGCACAGGGGACACAGAGCCUGCCCAAGUACCGCUCCCGAGGGAGCGGG
AAACGGGGGGGUGACUAUCCCGUGGGGUCCGGCGAGAGCGCUGGUCUACGGACCA
GGGGUGGCUGUGGGCAGGCUGCUCCUCAGGCCAGUUGAUUAGUUAACGCAUGGGCU
GUACCUCCACGUGGUCCCGCUGGUAAACGACUUGUCGGCUAAAUCAGCCCGCCAC
CAUCUGGGGAUAUGGUUGACCGUCUAACCCAGUACUCAGGUCACAAACAAAUGG
GAACAGAUACAGUGUAUGUCGGCCAGGACUACCCUUCUGGCUUAUCAAAACGGGU
ACCAGCACGGUUCUAGAGCGGCCGUGAGGGGGACAGCUGGGAGUCUCGGCAUGA
UUACAAAUUCUUGCGCUGCACUCGGAUGUCGUCCCCGUGACGGACACAUUAUCCG
GAAAGCGAGUGGUGACUCGCCUCAAGUAGCAAAAAAAAAAAAAAAAAAAAAA

30

344. OrLa5L_ZoAl3 [最初の39ntは自己切断により除去される]

gggagacgggucaaccgcguaaggagcggugaccgGAAUUC GGCGGGAGUAACUAUGACUC
UCUUAACGCACAGGGGACACAGAGCCUGCCCAAGUACCGCUCCCGAGGGAGCGGG
AAACGGGGGGGUGACUAUCCCGUGGGGUCCGGCGAGAGCGCUGGUCUACGGACCA
GGGGUGGCUGUGGGCAGGCUGCUCCUCAGGCCAGUUGAUUAGUUAACGCAUGGGCU
GUACCUCCACGUGGUCCCGCUGGUAAACGACUUGUCGGCUAAAUCAGCCCGCCAC
CAUCUGGGGAUAUGGUUGACCGUCUAACCCAGUACUCAGGUCACAAACAAAUGG
GAACAGAUACAGUGUAUGUCGGCCAGGACUACCCUUCUGGCUUAUCAAAACGGGU
ACCAGCACGGUUCUAGAGCGGCCGcuagguagucacauugcacuuucuguaacuugcacu
ggguguggggaugugggccugggguguggguuauugggguauauaugugggauauucugguggg
aanguccaauucacuguaugccuauuuuuuauaaaaagacgguaagcuagguucgcgaagcag
ccacaagccaauagccaguuaagguagcucauaguggguagguagacaggaaccuuugacucagaa
cgcguccauuaacaucuagaacggaccaaacuucggacaugcaccgauuaaccggauuugucca
agguggacgggccaccuuuacuuuacccggaaagggaacauauauaguuaauauguguuucguaa
uaUAGCAAAAAAAAAAAAAAAAAAAAAA

40

345. OrLa5L_TCB3 [最初の39ntは自己切断により除去される]

gggagacgggucaaccgcguaaggagcggugaccgGAAUUC GGCGGGAGUAACUAUGACUC
UCUUAACGCACAGGGGACACAGAGCCUGCCCAAGUACCGCUCCCGAGGGAGCGGG

50

AAACGGGGGGGUGACUAUCCCCUGGGGUCCGGCGAGAGCGCUGGUCUACGGACCA
GGGGUGGCUGUGGGCAGGCUGCUCCUCAGGCCAGUUGAUUAGUUACGCAUGGGCU
GUACCUCCACGUGGUCCCGCUGGUAACGACUUGUCGGCUAAAUCAGCCCGCCAC
CAUCUGGGGAUAUGGUUGACCGUCUAACCCAGUACUCAGGUCACAAACAAAUGG
GAACAGAUACAGUGUAUGUCGGCCAGGACUACCCUUCUGGCUUAUCAAAACGGGU
ACCAGCACGGUUCUAGAGCGGCCGCUAAAUCUCCUGACCAACUAGCUCACUGAC
UAAUUUUAAAACUGUCCUGUCUUAACUUGUUUACACGUGCUCUGUGGGCGGGCCA
UUUACACCCCGUCGCAACACAACCUGUAAAUAACUUGUGUAUGUCUGUUUAUGUCC
UAAUUUAUUUAUUUAAAACAGAUUCUUGGCCAUGGUCUCGGCCAACCAAUAAAGUC
AGUGAUGCGAGUCGCAAUGCGGAGCAAGAGACCUAGGCGUGUAUUUAUUGCUGGC
AUGCGGCGCCGGAGCCGGUCAUCUGCUAUGGGGAGCAAUGGCCGGGCGGAUACCU
CCACGUGGUUCCUGUGGGUGGCCCGUCGAGGACGGUAACCAGCGAAACUCCGUA
AAGUCCUUCUACGAGAAGGAACUCCGGUAAAGAUUUUCCAAGCCUGUACACG
UGAUUCCCUUGGAACAAGCAAAGUGUGGUUCCCUAGAGAGGGCCAGGUCAGGAG
UUCGCAAUAGUGGGCUGCAAGAGUUAUGCUGGGCUACAGUGUCAGGACGAAGAG
UGGGUAGUGAUCGCAAAAUACGUGAAUAGCUACCCCCCGCCUGGCACCACUAGA
CAACAACAAGGGGUACGACAGCUCUUCUGUCGAAAGUUCGGGCGCACACCCGUAA
AAGGUAGCAAAAAAAAAAAAAAAAAAAAAA

10

346-349. 配列なし

350. pUC57-miniプラスミド

20

CCAATGATATCGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGT
AAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCA
CAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACC
AGGCGTTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTA
CCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCAC
GCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCAC
GAACCCCCCGTTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTC
CAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTA
GCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAAGTGGTGGCCTAACTAC
GGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTT
CGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTG
GTTTTTTTGTGTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGAT
CCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACCTCACGTTAAGG
GATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAA
AATGAAGTTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTAC
CAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATA
GTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGG
CCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAG
CAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCC
GCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCCGCCAGT
TAATAGTTTTCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTACGCTCGT
CGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGA
TCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTGAG
AAGTAAGTTGGCCGCGAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTC
TACTGTTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAG
TCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACG
GGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTT
CTTCGGGGCGGAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAA
CCCACTCGTGACCCAACTGATCTTCAGCATCTTTTACTTTACCAGCGTTTCTGGG
TGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAGGGAATAAGGGCGACACGGAA

30

40

50

ATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTATCAGGGTTA
TTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAACAAATAGGGG
TTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTC

351. pcDNA3.1 N-DYKプラスミド

GACGGATCGGGAGATCTCCCGATCCCCTATGGTGCACCTCTCAGTACAATCTGCTCT
GATGCCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGA
GTAGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCA
TGAAGAATCTGCTTAGGGTTAGGCGTTTTTGCCTGCTTCGCGATGTACGGGGCCAGA
TATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTC
ATTAGTTCATAGCCCATATATGGAGTTCGCGGTTACATAACTTACGGTAAATGGCC
CGCCTGGCTGACCGCCCAACGACCCCCGCCCCATTGACGTCAATAATGACGTATGTT
CCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACG
GTAAACTGCCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTA
TTGACGTCAATGACGGTAAATGGCCCCGCTGGCATTATGCCCAGTACATGACCTTA
TGGGACTTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTG
ATGCGGTTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTTGACTCACGGGGATT
TCCAAGTCTCCACCCCATTTGACGTCAATGGGAGTTTTGTTTTGGCACCAAAATCAAC
GGGACTTTTCCAAAATGTCGTAACAACCTCCGCCCCATTGACGCAAAATGGGCGGTAGG
CGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCTCTGGCTAACTAGAGAACCCAC
TGCTTACTGGCTTATCGAAATTAATACGACTCACTATAGGGAGACCCAAGCTGGCT
AGCGTTTTAACTTAAGCTTGGTACCGAGCTCGGATCCGCCACCGAATTCTGCAGAT
ATCCAGCACAGTGGCGGCCGCTCGAGTCTAGAGGGCCCGATTACAAGGATGACGAC
GATAAGTGATAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATC
TGTTGTTTTGCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCCTGT
CCTTTCTTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCAATTCTA
TTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGC
AGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAGAACCAGCTG
GGGCTCTAGGGGGTATCCCCACGCGCCCTGTAGCGGCGCATTAAAGCGCGGCGGGTG
TGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCT
TTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTTCGCCGGCTTTCCCCGTCAAGCTCTA
AATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAA
AAAACCTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTT
TTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAACTG
GAACAACACTCAACCCTATCTCGGTCTATTCTTTTTGATTTATAAGGGATTTTGCCGA
TTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAATAAATTTAACGCGAATTAA
TTCTGTGGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGC
AGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCC
AGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCA
TAGTCCCGCCCCCTAACTCCGCCCATCCCGCCCCCTAACTCCGCCCAGTTCCGCCCAT
CTCCGCCCATGGCTGACTAATTTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCTG
CCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTTGGAGGCCTAGGCTTTTGC
AAAAAGCTCCCGGGAGCTTGTATATCCATTTTCGGATCTGATCAAGAGACAGGATG
AGGATCGTTTTCGCATGATTGAACAAGATGGATTGCACGCAGGTTCTCCGGCCGCTT
GGGTGGAGAGGCTATTCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGAT
GCCGCCGTGTTCCGGCTGTCAGCGCAGGGGCGCCCGGTTCTTTTTGTCAAGACCGA
CCTGTCCGGTGCCCTGAATGAACTGCAGGACGAGGCAGCGCGGCTATCGTGGCTGG
CCACGACGGGCGTTCTTTCGCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGG
GACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCTATCTACCTTGC
TCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTTG
ATCCGGCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAGCACGT

10

20

30

40

50

ACTCGGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGGACGAAGAGCATCAGGG
GCTCGCGCCAGCCGAACGTTCGCCAGGCTCAAGGCGCGCATGCCCGACGGCGAGG
ATCTCGTCGTGACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGC
CGCTTTTCTGGATTTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGA
CATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACC
GCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTTCGCAGCGCATCGCCTTCTATC
GCCTTCTTGACGAGTTCTTCTGAGCGGGACTCTGGGGTTTGGAAATGACCGACCAAG
CGACGCCCAACCTGCCATCACGAGATTTTCGATTCCACCGCCGCCTTCTATGAAAGG
TTGGGCTTCGGAATCGTTTTCCGGGACGCCGGCTGGATGATCCTCCAGCGCGGGGA
TCTCATGCTGGAGTTCTTCGCCCACCCCAACTTGTTTATTGTCAGCTTATAATGGTTA
CAAATAAAGCAATAGCATCACAAATTTACAAATAAAGCATTTTTTTCACTGCATT
CTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGTATACCGTCGA
CCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCTGTGTGAAATTGT
TATCCGCTCACAAATTCACACAACATACGAGCCGGAAGCATAAAGTGTAAGCCTG
GGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTT
TCCAGTCGGGAAACCTGTGCTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGG
AGAGGCGGTTTTCGTATTGGGCGCTCTTCCGCTTCTCGCTCACTGACTCGCTGCG
CTCGGTTCGTTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGT
TATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAA
AAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCC
CCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAG
GACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTT
CCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTCTCCCTTCGGGAAGCGTGGC
GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTTCGGTGTAGGTTCGTTTCGCTCCAA
GCTGGGCTGTGTGCACGAACCCCCCGTTTCAGCCCGACCGCTGCGCCTTATCCGGTA
ACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCC
ACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAA
GTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGC
TGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACC
ACCGCTGGTAGCGGTGGTTTTTTTGTGTTGCAAGCAGCAGATTACGCGCAGAAAAAA
AGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACG
AAAACCTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAG
ATCCTTTTAAATTAATAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAAC
TTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTC
TATTTTCGTTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGG
GAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACC
GGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAAGTG
GTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGA
GTAAGTAGTTCGCCAGTTAATAGTTTTCGCAACGTTGTTGCCATTGCTACAGGCAT
CGTGGTGTACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATC
AAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTC
CTCCGATCGTTGTGAGAAAGTAAGTTGGCCGCGAGTGTTATCACTCATGGTTATGGCA
GCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGT
GAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTG
CCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAGTGCTCA
TCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGA
TCCAGTTCGATGTAACCCACTCGTGACCCCAACTGATCTTCAGCATCTTTTACTTTC
ACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCCAAAAATGCCGCAAAAAAGGGAAT
AAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAA
GCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAA

10

20

30

40

50

AATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTC

352. RTプライマー伸長アッセイ用のDNAプライマー

CAGCACTAGATTTTTTGGGGTTGAATG

353. RTプライマー伸長アッセイ用の鋳型RNA AUACCCGCUUAAUUCAUUCAGAUUC

UGUAAUAGAACUGUCAUUCACCCCAAAAAUCUAGUGCUGAUUAUAAACCUUCACCA

AUUAGGUUCAAAUAAGUGGUAAUGCGGGACAAAAGACUAUCGACAUUUGAUACAC

UAUUUAUCAAUUGGAUGUCUUAUUUUUUUU

360. シーケンスリードのマッピング用の導入遺伝子コンティグ [TriCasA5_CBhBsi_GFP_GeFo3]

AGTCATAGAGCCAGAACCTCCTCGTGGTCCCGCTGGGCACAGGGATTAATTTTTCT
GTGGCAAATTTGACTGGCTTCAGAGAGCGTTTTTTCGAAGTGGACTGTGTGACTGCG
TTCCCCCCTTAGTTGCTATATCCGCTTCGATTAACATCTCACCTCGACGTATAAGAT
CATTTCTAGAAGGTCGACCAGATGTCCGAGGTCGACCAGTTGTCCGTGTGGAATTG
TGAGCGCTCACAATTCCACACGttacataacttacggtaaattggcccgcttgctgaccgccc
acgacccccgcccattgacgtcaatagtaacgccaatagggactttccattgacgtcaatgggtggag
tatttacggtaaactgccacttggcagtacatcaagtgtatcatatgccaaagtacgccccctattgac
gtcaatgacggtaaattggcccgcttgccattgtgcccagtacatgaccttatgggactttcctacttgg
cagtacatctacgtatttagtcatcgtattaccatgggtcgaggtgagccccacgttctgcttcactctcc
catctccccccccctccccaccccccaattttgtatttatttattttttaattttttgtgcagcgatggggg
cggggggggggggggggggggcgcgcgccaggcgggggcgggggcgggggcgagggggcgggggcgga
ggcgggagaggtgcggcggcagccaatcagagcgggcgcgctccgaaagtcttctttatggcgaggcg
gcggcgggcgggcgggccctataaaaagcgaagcgcgcgggcgggcgggagtcgctgcgacgctgccttc
gccccgtgccccgctccgcccgcgctcgcgcccggctctgactgaccgcggttactccca
caggtgagcgggcgggacggcccttctcctccgggctgtaattagctgagcaagaggtaagggttta
agggatgggttggttggtgggtatattaatgtttaattacctggagcacctgcctgaaatcacttttttca
ggttggcggtacggccacccatgggtgagcaaggcgaggagctgttcaccgggggtgggtgcccacctgg
tcgagctggacggcgacgtaaacggccacaagtccagcgtgtccggcgagggcgagggcgatgcc
cctacggcaagctgacctgaagtcatctgcaccacgggcaagctgccccgtgcccctggccccacccctc
gtgaccacccctgacctacggcgtgcagtgcttcagccgctaccccgaccacatgaagcagcacgactt
cttcaagtccgccatgcccgaaggctacgtccaggagcgcaccatcttcttcaaggacgacggcaact
acaagacccgcgcccaggtgaagttcgagggcgacacccctgggtgaaccgcatcgagctgaagggca
tcgacttcaaggaggacggcaacatcctggggcacaaagctggagtacaactacaacagccacaacgt
ctatatcatggccgacaagcagaagaacggcatcaaggtgaacttcaagatccgccacaacatcgag
gacggcgagcgtgcagctcgccgaccactaccagcagaacacccccatcgggcgacggccccgtgctg
ctgccccgacaaccactacctgagcaccacgtccgcccctgagcaaagaccccaacgagaagcgcgatc
acatgggtcctgctggagttcgtgaccgcccggggtcactctcggcagtgagcagctgtacaagtaa
aagcttGatccagacatgataagatacattgatgagtttgacaaaccacaactagaatgcagtga
aaaatgctttatgtgaaatgtgatgctattgctttatgttaaccattataagctgcaataaacaag
tTGTGGAATTGTGAGCGCTCACAATTCCACAGCGGCGCTGAGGTAGATAATCTTTG
TATAGTGGGGGGGGATCTCATGTACCGGGTTTCTTTTATTTGATTTTCAATAAAAC
AGACGGTAGCTAGGTTTCGCAAGGCAGCCACAAGCCAAAGATAGGTAGGGTGCTCAT
AGTGAGTAGGGACAGTGCCTTTTGATTCAACAACGCGTCAATACCATCTGACACGGA
TACCCTTACCGGACTTGTCTATGATCTCCCAGACTTGTCCAAGGTGGACGGGGCCACC
TTTACTTAACCCGGAAGGAACATATATTAATTATATGTGTTCGGAAAA

361. シーケンスリードのマッピング用の、rDNAに隣接する導入遺伝子配列コンティグ

GCTGGGCGCGCGCCGCGGGCTGGACGAGGCGCGCGCCCCCCCCACGCCCGGGGCACC
CCCCCTCGCGGCCCTCCCCCGCCCCACCCGCGCGCGCGCGCTCGCTCCCTCCCCACCCC
GCGCCCTCTCTCTCTCTCTCTCCCCGCTCCCCGCTCCTCCCCCTCCCCGGGGGAGC
GCCGCGTGGGGGCGCGGCGGGGGGAGAAGGGTTCGGGGCGGCAGGGGGCCGCGCGGCG
GCCGCGGGGGCGGCCGCGGGGGGAGGTCCCCGCGAGGGGGGGCCCCGGGGACCCGG

10

20

30

40

50

GGGGCCGGCGGGCGGGCGGGACTCTGGACGCGAGCCGGGGCCCTTCCCGTGGATCGCC
CCAGCTGCGGGCGGGCGTTCGCGGGCCGCCCCGGGGAGCCCGGCGGGCGGGCGCG
CCCCCACCCCCACCCACGTCTCGGTGCGCGCGCGTCCGCTGGGGGCGGGAGCG
GTCGGGCGGGCGGGCGGTTCGGCGGGCGGGCGGGGCGGGGCGGTTTCGTCCCCCGCCCTA
CCCCCCCCGGCCCCGTCCGCCCCCGTTCCCCCCTCCTCCTCGGCGCGGGCGGGCGGGC
GGCGGGCAGGCGGGCGGAGGGGCGCGGGCGGGTCCCCCGCGGGGTCCGCCCCCGG
GGCCGCGGTTCCGCGCGCGCCTCGCCTCGGCCGGCGCCTAGCAGCCGACTTAGAAC
TGGTGCGGACCAGGGGAATCCGACTGTTTAATTAAACAAAGCATCGCGAAGGCC
GCGGCGGGTGTGACGCGATGTGATTTCTGCCCAGTGCTCTGAATGTCAAAGTGAA
GAAATTCAATGAAGCGCGGGTAAACGGCGGGAGTAACATGACTCTCTTAAGGAGT
CATAGAGCCAGAACCTCCTCGTGGTCCCGCTGGGCACAGGGATTAATTTTTCTGTG
GCAAATTTGACTGGCTTCAGAGAGCGTTTTTTCGAAGTGGACTGTGTGACTGCGTTC
CCCCCTTAGTTGCTATATCCGCTTCGATTAACATCTCACCTCGACGTATAAGATCAT
TTCTAGAAGGTCGACCAGATGTCCGAGGTCGACCAGTTGTCCGTGTGGAATTGTGA
GCGCTCACAATTCCACACgttacataacttacggtaaattggcccgctggctgaccgccaacg
accccgcccattgacgtcaatagtaacgccaatagggactttccattgacgtcaatgggtggagtatt
tacggtaaactgcccacttggcagtacatcaagtgtatcatatgccaaagtacgccccctattgacgtca
atgacggtaaattggcccgctggcattgtgcccagtacatgaccttatgggactttcctacttggcagt
acatctacgtattagtcacgtattaccatggtcgaggtgagccccacgttctgcttcactctcccat
ctccccccccctccccaccccccaattttgtattttattttttaattttttgtgcagcgatggggg
cgggggggggggggggcgcgcgccaggcgggggcgggggcgggggcgagggggcgggggcgaggc
ggagaggtgcgggcgccagccaatcagagcgggcgcgctccgaaagtcttctttatggcgaggcgggc
gcgggcgggcgccctataaaaagcgaagcgcgcgggcgggcgggagtcgctgcgacgctgccttcgcc
ccgtgccccgctccgcccgcctcgcgccggccggcctctgactgaccggttactcccacag
gtgagcgggcgggagggcccttctcctcgggctgtaattagctgagcaagaggttaagggtttaagg
gatggttggttggtgggtattaatgtttaattacctggagcacctgcctgaaatcacttttttcaggtt
ggcgtagggccaccatgggtgagcaaggcgaggagctgttcaccgggggtggtgcccacctggtcga
gctggagggcgacgtaaacggccacaagttcagcgtgtccggcgagggcgagggcgatgccaccta
cggcaagctgaccctgaagttcatctgcaccaccggcaagctgcccgtgcccgtggcccacccctcgtga
ccaccctgacctacggcgtgcagtgcttcagccgctaccccgaccacatgaagcagcacgacttcttc
aagtccgcatgcccgaaggctacgtccaggagcgccaccatcttcttcaaggacgacggcaactaca
agaccgcgccgaggtgaagttcgagggcgacaccctggtgaaccgcatcgagctgaagggcacg
acttcaaggaggacggcaacatcctggggcacaagctggagtacaactacaacagccacaacgtcta
tatcatggccgacaagcagaagaacggcatcaaggtgaacttcaagatccgccacaacatcgaggac
ggcagcgtgcagctcgccgaccactaccagcagaacacccccatcggcgacggccccgtgctgctgc
ccgacaaccactacctgagcaccagtcggccctgagcaaaagacccaacgagaagcgcgatcacat
ggtcctgctggagttcgtgaccgcccgggatcactctcggcacgagctgtacaagtaaaag
cttGatccagacatgataagatacattgatgagtttggacaaaccacaactagaatgcagtgaaaaa
atgctttatgttgaaatttgatgctattgctttatgttaaccattataagctgcaataaacaagtTG
TGGAATTGTGAGCGCTCACAATCCACAGCGGCCGCTGAGGTAGATAATCTTTGTA
TAGTGGGGGGGGATCTCATGTACCGGGTTTCTTTTATTTGATTTTCAATAAAACAG
ACGGTAGCTAGGTTTCGAAGGCAGCCACAAGCCAAAGATAGGTAGGGTGCTCATAG
TGAGTAGGGACAGTGCCTTTTGATTCAACGCGTCAATACCATCTGACACGGATA
CCCTTACCGGACTTGTCATGATCTCCAGACTTGTCCAAGGTGGACGGGCCACCTT
TACTTAACCCGGAAAAGGAACATATATTAATTATATGTGTTTCGGAAAATAGCCAAA
TGCCTCGTCATCTAATTAGTGACGCGCATGAATGGATGAACGAGATTCCCACTGTC
CCTACCTACTATCCAGCGAAACCACAGCCAAGGGAACGGGCTTGGCGGAATCAGCG
GGGAAAGAAGACCCTGTTGAGCTTGAAGTCTAGTCTGGCACGGTGAAGAGACATGAG
AGGTGTAGAATAAGTGGGAGGCCCGCGCCCCCGGTGTCCCGCGAGGGGCC
CGGGGCGGGGTCCGCGGCCCTGCGGGCGCCGGTGAAATACCACTACTCTGATCGT

10

20

30

40

50

TTTTTCACTGACCCGGTGAGGCGGGGGGGCGAGCCCGAGGGGCTCTCGCTTCTGGC
GCCAAGCGCCCGCCCGGGCGCGACCCGCTCCGGGGACAGTGCCAGGTGGGGA
GTTTGACTGGGGCGGTACACCTGTCAAACGGTAACGCAGGTGTCCTAAGGCGAGCT
CAGGGAGGACAGAAACCTCCCGTGGAGCAGAAAGGGCAAAAGCTCGCTTGATCTTGA
TTTTCAGTACGAATACAGACCGTGAAAGCGGGGCCTCACGATCCTTCTGACCTTTT
GGGTTTTAAGCAGGAGGTGTCAGAAAAGTTACCACAGGGATAACTGGCTTGTGGCG
GCCAAGCGTTTCATAGCGACGTCGCTTTTTGATCCTTCGATGTCGGCTCTTCCTATCA
TTGTGAAGCAGAATTCGCCAAGCGTTGGATTGTTACCCACTAATAGGGAACGTGA
GCTGGGTTTTAGACCGTCGTGAGACAGGTTAGTTTTACCCTACTGATGATGTGTTGT
TGCCATGGTAATCCTGCTCAGTACGAGAGGAACCGCAGGTTTCAGACA

10

400. CBhプロモーター

Cgt tacataactt acggtaaatg gcccgcctgg ctgaccgccc aacgaccccc gccatttgac
gtcaatagta acgccaatag ggactttcca ttgacgtcaa tgggtggagt atttacggta aact
gcccac ttggcagtac atcaagtgt tcatatgcc agtacgcccc ctattgacgt caatgacg
gt aaatggcccg cctggcattg tgcccagtac atgaccttat gggactttcc tacttggcag ta
catctacg tattagtcat cgctattacc atggtcgagg tgagccccac gttctgcttc actctcc
cca tctccccccc ctccccaccc ccaattttgt atttatttat tttttaatta tttgtgcag cgat
ggggggc gggggggggg gggggggcgcg cgccaggcgg ggcggggcgg ggcgaggggc ggg
gcggggc gaggcggaga ggtgcggcgg cagccaatca gagcggcgcg ctccgaaagt ttcc
ttttat ggcgagggcgg cggcggcggc ggccctataa aaagcgaagc gcgcggcggg cggga
gtcgc tgcgacgtg ctttcgcccc gtgccccgt cgcgcgcgc ctcgcgcgc cgcgccg
gc tctgactgac cgcgttact ccacaggtga gcggggcggga cggcccttct cctccgggct g
taattagct gagcaagagg taagggttta agggatggtt ggttgggtggg gtattaatgt ttaatt
acct ggagcacctg cctgaaatca ctttttttca ggttgg

20

401. EFSプロモーター

Gggcagagcgcacatcgcccacagtccccgagaagttggggggaggggtcggcaattgaaccggtg
cctagagaaggtggcgcggggtaaaactgggaaagtgatgtcgtgtactggctccgcctttttcccgag
ggtggggggagaaccgtatataagtgcagtagtcgccgtgaacgttctttttcgcaacggggttgccgc
cagaacaca

30

Gg

403. EFSBsiWIプロモーター

Gggcagagcgcacatcgcccacagtccccgagaagttggggggaggggtcggcaattgaaccggtg
cctagagaaggtggcgcggggtaaaactgggaaagtgatgtcgtgtactggctccgcctttttcccgag
ggtggggggagaaccgtatataagtgcagtagtcgccgtgaacgttctttttcgcaacggggttgccgc
cagaacacaggcgtacg gccacatg

404. mPGKプロモーター

TTCTACCGGGTAGGGGAGGCGCTTTTCCCAAGGCAGTCTGGAGCATGCGCTTTAGC
AGCCCCGCTGGGCACTTGGCGCTACACAAGTGGCCTCTGGCCTCGCACACATTCCA
CATCCACCGGTAGGCGCCAACCGGCTCCGTTCTTTGGTGGCCCCCTTCGCGCCACCTT
CTACTCCTCCCCTAGTCAGGAAGTTCCCCCCCCGCCCGCAGCTCGCGTCGTGCAGG
ACGTGACAAATGGAAGTAGCACGTCTACTAGTCTCGTGACAGATGGACAGCACCGC
TGAGCAATGGAAGCGGGTAGGCCTTTGGGGCAGCGGCCAATAGCAGCTTTGCTCCT
TCGCTTTCTGGGCTCAGAGGCTGGGAAGGGGTGGGTCCGGGGGCGGGCTCAGGGGC
GGGCTCAGGGGCGGGGCGGGCGCCCGAAGGTCCTCCGGAGGCCCGGCATTCTGCAC
GCTTCAAAAGCGCACGTCTGCCGCGCTGTTCTCCTCTTCCTCATCTCCGGGCCTTTT
GACCT

40

405. SV40 LpA

aacttgtttattgcagcttataatggttacaaataaagcaatagcatcacaatttcacaaataaagcat

50

ttCtttcactgcattctagttgtggtttgtccaaactcatcaatgtatctta

406. min pA

AAATAAAATACGAAATGAATTCGTAATCATGTGTGTTGGTTTCTTGTGT

407. BGH pA

CTAGTTGCCAGCCATCTGTTGTTTGGCCCTCCCCCGTGCCTTCCTTGACCCTGGAAG
GTGCCACTCCCCTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTG
AGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGA
TTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGG

408. CMV*

CGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCC
CATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCAT
TGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGT
GTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCT
GGCATTATGCCCAGTACATGACCTTATGGGACTTTTCTACTTGGCAGTACATCTAC
GTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTTGGCAGTACATCAATGGGCG
TGGATAGGGGTTTGA CTCACGGGGATTTC CAAGTCTCCACCCCATTGACGTCAATG
GGAGTTTGT TTTGGCACCAAAATCAACGGGACTTTTCCAAAATGTCGTAACAACTCC
GCCCCATTGACGCAAATGGGCGGTAGGCGTG TACGGTGGGAGGTCTATATAAGCAG
AGCTcggttagtgaaccgCTTGTCTCGCTCCGGGGAACGCTCGGAAACTCCCGGCCGCC
GCCACCCGCGTCTGTTCTGTTACACAAGGGAAGAAAAGCCGCTGCCGCACTCCGAG
TGTCGTACGGCCACC.

409. SV40*

Ctgtggaatgtgtgtcagttaggggtgtggaaagtc cccaggctcccagcaggcagaagtatgcaa
gcatgcatctcaattagtcagcaaccagggtgtggaaagtc cccaggctcccagcaggcagaagtat
gcaaagcatgcatctcaattagtcagcaaccatagtc cccgcccctaactccgcccattccgcccctaa
ctccgcccagttccgcccattctccgcccattGgctgactaattGCGtCtatAtatgcagaggccgag
gccgcctcggcctctgagctattccagaagtagtgaggaggcCGTACGGCCACC.

408-410. 配列なし

411. NLS-GFP ORF配列

ATG gcc cca aag aag aag cgg aag gtcgtgagcaagggcgaggagctgttcaccgggggt
gggtgcccattcctggctgagctggacggcgacgtaaacggccacaagttcagcgtgtccggcgagggc
gagggcgatgccacctacggcaagctgacctgaagttcatctgcaccaccggcaagctgcccgtgc
cctggcccaccctcgtgaccacctgacctacggcgtgcagtgtttcagccgctaccccgaccacatg
aagcagcacgacttcttcaagtccgccatgccgaaggctacgtccaggagcgcaccatcttcttcaa
ggacgacggcaactacaagaccgcgcggaggtgaagttcagggcgacaccctgggtgaaccgcat
cgagctgaagggcatcgacttcaaggaggacggcaacatcctggggcacaagctggagtacaactac
aacagccacaacgtctatatcatggccgacaagcagaagaacggcatcaaggtgaacttcaagatcc
gccacaacatcgaggacggcagcgtgcagctcgccgaccactaccagcagaacacccccatcggcg
acggccccgtgctgctgcccgacaaccactacctgagcaccagtcgcccctgagcaaagaccccaa
cgagaagcgcgatcacatggctcctgctggagttcgtgaccgccgcccgggatcactctcggcattggac
gagctgtacaagtaa

412. NLS-GFP アミノ酸配列

MAPKKKRKVVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLT LKFI
TTGKL PVPWPTLVTTLT YGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDG
NYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNNSHN VYIMADKQKNGI
KVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNH YLSTQSALS KDPNEKRDH
MVLLEFVTAAGITLGMDELYK-

413. GFP ORF配列

atgggtgagcaagggcgaggagctgttcaccgggggtgggtgcccattcctggctgagctggacggcgacg

10

20

30

40

50

taaacggccacaagttcagcgtgtccggcgagggcgagggcgatgccacctacggcaagctgacct
gaagttcatctgcaccaccggcaagctgcccgtgccctggcccaccctcgtgaccaccctgacctacg
gcgtgcagtgtcttcagccgctaccccgaccacatgaagcagcacgacttcttcaagtccgccatgccc
gaaggctacgtccaggagcgcaccatcttcttcaaggacgacggcaactacaagacccgcgccgagg
tgaagttcgagggcgacaccctgggtgaaccgcatcgagctgaagggcatcgacttcaaggaggacgg
caacatcctggggcacaagctggagtacaactacaacagccacaacgtctatatcatggccgacaag
cagaagaacggcatcaaggtgaacttcaagatccgccacaacatcgaggacggcagcgtgcagctcg
ccgaccactaccagcagaacacccccatcgggcgacggccccgtgctgctgcccgacaaccactacct
gagcaccagtcgcgccctgagcaaagaccccaacgagaagcgcgatcacatgggtcctgctggagttc
gtgaccgccgcgggatcactctcggcacatggacgagctgtacaagtaa

10

414. GFP アミノ酸配列

MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTCLKFICTTGKLPVP
WPTLVTTLTYGVCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEV
KFEGDTLVNRIELKGIDFKEDGNILGHKLEYNNSHNVIYIMADKQKNGIKVNFKIRH
NIEDGSVQLADHYQNTPIGDGPVLLPDNHYLSTQSALS KDPNEKRDHMLLEFVT
AAGITLGMDELYK-

415. Tdtomato ORF 配列

atggtgtccaaaggcgaggaagtgatcaaggaatttatgagattcaaggtcagaatggaaggatctat
gaacggccacgagttcgagatcgaggggagaaggagaaggcagaccctacgagggcacccaaaccg
ctaaattaaaagtgacaaaggcgacacctctgcccgttcgcctgggatatacctgtccccccagttcatgt
acggcagcaaggcctacgtgaagcacctgtgacatccccgattacaaaaagctgtcttttctgag
ggctttaagtgggagagagtgatgaacttcgaggacggaggcctggtgaccgttaccacagacagca
gcctgcaagatggaacactgatctacaaagtgaagatgcggggcactaatttccctccagacggcccc
gtgatgcagaagaaaacatgggctgggaggcctctacagaacggctgtaccccagagacggagtg
ctgaagggcgagattcaccaggcactgaagctgaaagatggcgacactacctgggtggagttcaaga
ccatttacatggccaagaagcctgtgcagctgcctggctactactacgtggacaccaagctggacatc
acctccacaatgaggattataccatcgtggaacagtacgagagaagcgagggcagacaccacctgt
tcctcggccacggcacaggctccaccggctctggcagcagcggcacagccagcagcagggacaaca
acatggccgtgatcaaggagttcatgcggttcaagggtgcggatggaaggcagtatgaacggccatga
gttcgaaatcgagggcgaaaggcgaaaggcagaccttacgagggcacacagacagctaagctgaaggt
gaccaaggaggaggcctctgccattcgcctgggacatcctgagccctcagtttatgtatggcagcaagg
cctacgtgaagcaccgccgacatccccgactacaagaagctgagcttccctgaggggtttcaaatgg
gaaagggtgatgaacttcgaggacggcgagctgggtgacagtgaccaggacagcagcctgcaggat
ggcacctgatctataaggtgaagatgagaggcaccaactttccacctgacggggcccgtgatgcaga
aaaaaacatgggatgggaggccagcaccgagcggctgtaccctagagatggcgttctgaaggcg
agatccaccaggccctgaagctgaaggacggcgccactacctgggtggagttcaagaccatctacat
ggccaagaaacctgtgcagctgcctggatactactacgtggacacaaagctggacatcacctctcaca
acgaggactataccatcgtcgagcagtagaacgcagcgaaggcagacaccatctgttcctgtacgg
tatggacgagctgtacaagtaatga

20

30

416. Tdtomato アミノ酸配列

MVSKGEEVIK EFMRFKVRME GSMNGHEFEI EGEGERPVE GTQTAKLKVT KGG
PLPFAWD ILSPQFMYGS KAYVKHPADI PDYKKLSFPE GFKWERVMNF EDGGLV
TVTQ DSSLQDGT LI YKVKMRGTNF PPDGPVMQKK TMGWEASTER LYPRDGV L
KG EIHQALKLKD GGHYLVFEKT IYMAKKPVQL PGYYYVDTKL DITSHNEDYT I
VEQYERSEG RHHLFLGHGT GSTGSGSSGT ASSEDNNMAV IKEFMRFKVR MEGS
MNGHEF EIEGEGEGRP YEGTQTAKLK VTKGGPLPFA WDILSPQFMY GSKAYVK
HPA DIPDYKKLSF PEGFKWERVM NFEDGGLVTV TQDSSLQDGT LIYKVKMRGT
NFPPDGPVMQ KKTMGWEAST ERLYPRDGV L KGEIHQALKL KDGGHYLVFE KTI
YMAKKPV QLPGYYYVDT KLDITSHNED YTIVEQYERS EGRHHLFLYG MDELYK

40

417. mCherry ORF 配列

50

auggugagcaagggcgaggaggauaacauggccaucaucaaggaguucaugcgcuucaaggug
cacauggagggcuccgugaacggccacgaguucgagaucgagggcgagggcgagggccgccccu
acgagggcacccagaccgccaagcugaaggugaccaagggugggccccugccccuucgccuggga
cauccuguccccucaguucauguacggcucaaggccuacgugaagcaccgcccgcgacaucccc
gacuacuugaagcuguccuucccgaggggcuucaagugggagcgcgugaugaacuucgaggac
ggcggcguggugaccgugacccaggacuccuuccugcaggacggcgaguucaucuacaagguga
agcugcgcgggcaccaacuuccccuccgacggccccguaaugcagaagaagaccaugggcuaggga
ggccuccuccgagcggauguaccccgaggacggcgcccugaaggggcgagaucaagcagaggcug
aagcugaaggacggcgggccacuacgacgcugaggucaagaccaccuacaaggccaagaagcccg
ugcagcugcccggcgccuacaacgucaacaucagguaggacaucaccucccacaacgaggacua
caccaucguggaacaguacgaacgcgcccaggggccgcccacuccaccggcgggcauggacgagcug
uacaaguag

10

418. mCherry アミノ酸配列

MVSKGEEDNMAIIKEFMRFKVHMEGSVNGHEFEIEGEGEGRPYEGTQTAKLKVTKG
GPLPFAWDILSPQFMYGSKAYVKHPADIPDYLKLSFPEGFKWERVMNFDGGVVTV
TQDSSLQDGEFIYKVKLRGTNFPSPDGPVMQKKTMGWEASSERMYPEDGALKGEIKQ
RLKLKDGGHYDAEVKTTYKAKKPVQLPGAYNVNIKLDITSHNEDYTIVEQYERAEGR
HSTGGMDELYK-

419. ヒト TERT ORF 配列

atggtgatgcctcgagcaccggtgcagagccgtgcggagcctgctgcggtctcactacagagagg
tcctccccctggccacgttcgtgcggcgctctgggccctcagggtggtgcgactggtgcagcgcgga
ccccgccgattccgcgccctggtggccagtgccctggtgtgtgtgccttgggatgccagacctcctc
ctgcggctccttcttttagacaggtgtcttctgaaggagctggtggccgggtgctgcagcggtgtg
gcgagcggggagctaagaatgtgttagctttcggtttgcccctgctggacggcgccagaggcggaacc
acctgaagccttcactacctcggtgagaagctacctgcctaacaccgtgacagacgccctgcggggc
agcgggcgccctggggcctgctactgagacgggtgggagatgatgtgctggtgcacctgctggccagat
gcgccctctttgtgctggtggccccctagctgtgcctaccagggtgtgtggccctccccctgtaccagctgg
gcgccgccacgcaggccagacccccctccgcacgccagcgccctaggagacgtctgggctgcgaaa
gagcttggaaccacagtgtgcgagaagccggcgctccccctgggcctccctgccccctggagccagaag
aagaggaggctccgccagcagaagcctcccgctgcccaagcggcctagacggggcgccgccccctga
acccgagcggaacccgggtcgccagggcagttggggccaccccggaaggaccagaggacctagcgga
ccggggcttctgctggtctccccagcccgccagccgaggaagcaaccagcctggaggggcgccctg
agcgggcaccagacactcccaccccttcggttgacagacaccacgcccggcccaccaagcaccagca
gacctccgcgaccttgggacacacccctgtcctcctgtgtacgccgagaccaagcacttcctgtactcct
ctggcgataaggaacagctccggcctagcttctctgctgagctctctgagaccagcttgaccggcgcc
cgcagactcgtggaaacaatcttcctgggcagcagaccttggtatgccaggcacccttagacggctgc
ctagactgccccagagatactggcagatgcgtccactgtttttggaactgctgggcaaccacgctcag
tgtccttacggcgcttctgctgaaaacccactgccccctgagagccgctgtcaccccagccgcccggcgt
gtgcgccagagaaaaagcctcagggaagcgtggctgctcctgaggaagaggatacagaccctagacg
cctggtgcaactgctgcgccaacatagcagcccttggcaggtgtacggcttcgtcagagcttgtctga
gaaggctggtgccccgggactgtggggctccagacacaacgagcggaagattcctgcggaacacaa
agaagttcatcagcctgggcaagcacgccaagctgtctctgagaggattgacctggaagatgagcgt
gagggattgcgctggtgagaaagatccctggagtggtgctgtgtgccagccgcccagacacagactg
cgagaggaaatcctggccaaattcctgcactggctgatgagcgtatatgtggtggaactgctgcggag
cttcttctacgtgaccgagacaacatttcagaagaacagacttttcttaccgggaagtctgtatggag
caagctgcagtcctcgaatcagacagcacctgaagagagttcagctgcgggaactgagcgaggga
gaggtgaggcagcatagagaggctagacctgtctgcttaccagcagactgagattcatcccaaac
ctgatggcctgagacctatcgtgaatatggactacgtggtggcgcccgaaccttcggcgctgagaag
cgggctgaaagactgacatccgggtgaaggccctgttcagcgtgctgaactacgagagagctcgtc
ggcctgggctactgggcgcccagcgtgctgggcctggatgacatccacagagcctggagaacattcgt

20

30

40

50

gctgagagtgagggctcaagatccccaccggagctgtacttcgttaaggtggacgtgacagggcgca
tacgacaccatccctcaggacagactgaccgaggtgattgccagcattatcaagccccagaacaccta
ttgctgagacgggtacgccgtggtgcagaaagccgcacatggccacgtgcggaaagccttcaagtct
cacgtgtccaccctgaccgacctgcagccttacatgcggcaattcgtggcccacctccaggagacaag
ccctctgagagacgccgtggtgatcgagcagtcacagccttaatgaggcctctagcggcctgttcg
acgtgttcctgagatttatgtgccaccacgccgtgagaatcagaggcaaaagctacgtgcagtgccag
ggcatcccccaggatctatcctgagcaccctgctgtgcagcctgtgctacggcgacatggaaaacaa
gctgttcgccggaatccggagagacggcctgctgctgcggctcgttgacgacttccctgctggtgacac
ctcacctgacacacgccaaaaccttcctcagaacactggttcggggcgctccccgagtagcggtgcgtc
gtgaacctgagaaagacagtgggtgaacttccctgtggaagatgaggcactcggcggaacagcgttcg
tgcatgatgccgcccacggcctgttccctggtgcggcctgctgctggacaccagaaccttgagggt
gcagagcgattacagcagctatgccagaaccagcatccgagccagcctaacctttaaccggggccttc
aaagccggcagaaatatgagaagaaagctgttcggagtgctgaggccttaaatgccacagcctgttcct
ggacctgcaggtcaacagcctgcaaacagtgtgcaccaacatctacaagatcctgctgttacaagcct
acagatttcacgcctgtgtgctgcagctgcctttccaccagcaggtgtggaagaacctacgtttttcc
tgagagtgatctctgatacagccagtttgtgctacagcatcctgaaggccaagaacgccggcatgagc
ctgggagccaaggcgctgctgggcctctgccttctgaggccgtgcagtggtgtgccaccaggcctt
tctgctgaagctgaccagacatagagtgcctacgtgcctctcctgggcagcctgagaaccgccccaga
cccagctctctcggaagctgcctggcacaacctgaccgccctggaagctgctgccaatcctgcctg
cctagcgacttcaagaccatcctggactgataa

10

20

420. ヒトTERTアミノ酸配列

MVMPRAPRCRAVRSLLRSHYREVLPLATFVRRRLGPQGWRLVQRGDPAAFRALVAQC
LVCVPWDARPPPAAPSFRQVSLKELVARVLQRLCERGAKNVLAFGFALLDGARGG
PPEAFTTSVRSYLPNTVTDALRGSAGWGLLLRRVGGDVLVHLLARCALFVLVAPSCA
YQVCGPPLYQLGAATQARPPPHASGPRRRLGCERAWNHSVREAGVPLGLPAPGARR
RGGSASRSLPLPKRPRRGAAPEPERTPVGQGSWAHPGRTRGPSDRGFCVVSPARPAE
EATSLEGALSGTRHSHPSVGRQHHAGPPSTSRRPPRPWDTPCPPVYAETKHFLYSSGD
KEQLRPSFLLSSLRPSLTGARRLVETIFLGSRPWMPGTPRRLPRLPQRYWQMRPLFL
ELLGNHAQCPYGVLLKTHCPLRAAVTPAAGVCAREKPGQGSVAAPEEEDTDPRRLVQ
LLRQHSSPWQVYGFVRACLRLRLVPPGLWGSRHNERFLRNTKKFISLGKHAKLSLQ
ELTWKMSVRDCAWLRRSPGVGCVPAAEHRLREEILAKFLHWLMSVYVVELLRSSFFY
VTETTFQKNRLFFYRKSVWSKLQSIGIRQHLKRVQLRELSEA EVRQHREARPALLTS
RLRFIPKPDGLRPIVNMDYVVGARTFRREKRAERLTSRVKALF SVLN YERARRPGLL
GASVLGLDDIHRAWRTFVLRVRAQDPPPELYFVKVDVTGAYDTIPQDRLTEVIASIIK
PQNTYCVRRYAVVQKAAHGHVRKAFKSHVSTLTDLQPYMRQFVAHLQETSPLRDA
VVIEQSSSLNEASSGLFDVFLRFMCHHAVRIRGKSYVQCQGIPQGSILSTLLCSLCYG
DMENKLFAGIRRDGLLLRLVDDFLLVTPHLTHAKTFLRTLVRGVPEYGCVVNLRKTV
VNFPVEDEALGGTAFVQMPAHGLFPWCGLLLDTRTLEVQSDYSSYARTSIRASLT FN
RGFKAGRNMRRKLFGLVRLKCHSLFLDLQVNSLQTVCTNIYKILLQAYRFHACVLQ
LPFHQQVWKNPTFFLRVISDTASLCYSILKAKNAGMSLGAKGAAGPLPSEAVQWLC
HQAFLCLKLTRHRVTYVPLLGLSLRTAQTQLSRKLP GTT LTALEAAANPALPSDFK TIL
D--

30

40

421. ヒトPAH ORF配列

atggtgatgagcaccgccgtgctggagaatcctggcctgggcagaaagctgtccgatttcggccaag
agacaagctacattgaggataattgcaaccagaacggcgctatcagcctgatcttttcttgaaagag
gaagtgggcgctctggccaaggctcctgaggctgttcgaggaaaacgacgtgaacctgaccatctcg
aaagcagacccagccggctgaagaaagatgagtatgagttcttcacccacctggacaagagatctctc
cctgccctgacaaacatcatcaaaatcctaagacacgacattggcgccaccgtgcacgagctgtcaag
agataagaagaaggacaccgtgccttggtttcctagaacctccaggagctggatagattcgccaatc
agatcctgtcttacggcgccgaactggacgcccgaaccacccggccttcaaggacctgtgtaccgggc

50

cagacgcaagcagttcgccgacatcgccctataactaccggcatggccagcctatccctagagtggaaat
acatggaagaagagaaaaaacctggggcacccgttttcaagaccctgaaaagcctgtataagactca
cgcttgctacgagtacaaccacatcttccctctgctggaaaagtactgtggctttcacgaggacaacat
cccacagctggaggacgtgtcccagttcctgcagacctgtaccggcttcgggctgcggccccgtggccg
gactgctgagcagccgggacttcctggggcggactggcttttagagtgttccactgcacacagtacatc
agacacggctccaagcctatgtacacacctgagcccgacatctgccacgagctgtgggacacgtcc
ctctgttttagcgatagatctttcgcccagtttagccaggagatcggcctggccagcttgggcgccccctg
atgaatacatcgagaaaactggctacaatctactggttcacagtggagttcggcctctgcaagcagggc
gacagcatcaaggcctacggcgccggactgctttctagcttcggcgagctgcaatactgtctgtccga
gaagcccaagctgctgcctctggaactggaaaagaccgccatccagaattacaccgtgaccgagttc
cagccccctgtactacgtggctgagagtttcaacgacgctaaggaaaaggtgcggaacttcgccgcca
caatcccaagacctttcagcgtgcggtacgacccctacacccaaagaatcgagggtgctggacaacac
ccagcagctgaagatcctggcagatagcatcaacagcgagatcgggaatcctgtgcagcgccccctgcag
aagattaagtgataag

10

422. ヒトPAHアミノ酸配列

MSTAVLENPGLGRKLSDFGQETSYIEDNCNQNGAISLIFSLKEEVGALAKVLRRLFEEEN
DV
NLTHIESRPSRLKKDEYEFFTHLDKRSLPALTNIIKILRHDIGATVHELSDKKKDTV
PW
FPRTIQELDRFANQILSYGAELDADHPGFKDPVYRARRKQFADIAYNYRHGQPIPRV
EYM
EEEKKTWGTVFKTLKSLYKTHACYEYNHIFLLEKYCGFHEDNIPQLEDVSQFLQTC
TGF
RLRPVAGLLSSRDFLGGLAFRVFHCTQYIRHGSKPMYTPEPDICHELLGHVPLFSDRS
FA
QFSQEIGLASLGAPDEYIEKLATIWFTVEFGLCKQGDSIKAYGAGLLSSFGEQYCLS
E
KPKLLPLELEKTAIQNYTVTEFQPLYYYVAESFNDAKEKVRNFAATIPRPFVRYDPYT
QR
IEVL DNTQQLKILADSINSEIGILCSALQKIK

20

30

423-498. 配列なし

499. XbaI_RNAPI_term_rlacO_CBh_HindIII_Kozak_NLSGFP_HindIII_SV40
LPA_rlacO_NotI
TCTAGAAGGTCGACCAGATGTCCGAGGTCGACCAGTTGTCCGTGTGGAATTGTGAG
CGCTCACAATTCCACA Cgt tacataactt acggtaaattg gcccgccttg ctgaccgccc a
acgaccccc gcccatgtac gtcaatagta acgccaatag ggactttcca ttgacgtcaa tggg
tgaggt atttacggta aactgcccac ttggcagtac atcaagtgt tcatatgcca agtacgcc
cc ctattgacgt caatgacgggt aaatggccc cctggcattg tgcccagtac atgaccttat gg
gactttcc tacttggcag tacatctacg tattagtcac cgctattacc atggtcgagg tgagccc
cac gttctgcttc actctcccca tctccccccc ctccccaccc ccaattttgt atttatttat tttt
taatta ttttgtgcag cgatggggggc ggggggggggg gggggggcgcg cgccaggcggg ggcgg
ggcgg ggcgagggggc gggggcggggc gagggcggaga ggtgcggcgcg cagccaatca gaggc
ggcgcg ctccgaaagt ttccttttat ggcgagggcg cgggcgcgggc ggccctataa aaagcga
agc gcgcggcggg cgggagtcgc tgcgacgctg ctttcgcccc gtgccccgct ccgcccgcgc
ctcgcgccgc ccgccccggc tctgactgac cgcgttactc ccacaggtga gcggggcgggga cg
gcccttct cctccggggt gtaattagct gagcaagagg taagggttta agggatgggt ggttgg
tggg gtattaatgt ttaattacct ggagcacctg cctgaaatca ctttttttca ggttggga agctt
GCCACCATG gcc cca aag aag aag cgg aag gtcgtgagcaaggcgaggagctgttcac
cgggggtgggtgcccattcctggctgagctggacggcgacgtaaacggccacaagttcagcgtgtccggc

40

50

gagggcgagggcgatgccacctacggcaagctgaccctgaagttcatctgcaccaccggcaagctgc
ccgtgccctggcccaccctcgtgaccaccctgacctacggcgtgcagtgtttcagccgctaccccgac
cacatgaagcagcacgactttcttcaagtccgccatgcccgaaggctacgtccaggagcgcaccatctt
cttcaaggacgacggcaactacaagaccgcgccgaggtgaagttcgagggcgacaccctggtgaa
ccgcatcgagctgaagggcatcgacttcaaggaggacggcaacatcctggggcacaagctggagta
aactacaacagccacaacgtctatatcatggccgacaagcagaagaacggcatcaaggtgaacttca
agatccgccacaacatcgaggacggcagcgtgcagctcgccgaccactaccagcagaacacccccat
cggcgacggccccgtgctgctgcccgacaaccactacctgagcaccagtcggccctgagcaaagac
cccaacgagaagcgcgatcacatggtcctgctggagttcgtgaccgccgcccgggatcactctcggcat
ggacgagctgtacaagtaaaagcttGatccagacatgataagatacattgatgagtttggacaaacca
caactagaatgcagtgaaaaaaatgctttatgtgaaatttgtgatgctattgctttatgttaaccatt
ataagctgcaataaacaagtTGTGGAATTGTGAGCGCTCACAAATTCCACAGCGGCCGC
500. TCA5_CBhBsi_GFP_GeFo3

10

gggagacgggucaaccgcguaggagcggugaccgGAAUUCGGCGGGAGUAAUAUGACUC
UCUUAAGGAGUCAUAGAGCCAGAACCUCUGUGGUCCCGCUGGGCACAGGGGAUU
AAUUUUUCUGUGGCAAAUUGACUGGCUUCAGAGAGCGUUUUUCGAAGUGGACU
GUGUGACUGCGUUCUUUUUUAGUUGCUAUAUCCGCUUCGAUUAACAUCUCACCU
CGACGUAAUAGAUAUUUCUAGAAGGUCGACCAGAUGUCCGAGGUCGACCAGUUG
UCCGUGUGGAAUUGUGAGCGCUCACAAUCCACACGUUACAUAAAUACGGUAAA
UGGCCCGCCUGGCUGACCGCCCAACGACCCCGCCCAUUGACGUCAAUAGUAACG
CCAAUAGGGACUUUCCAUUGACGUCAAUGGGUGGAGUAUUUACGGUAAACUGCCC
ACUUGGCAGUACAUAAGUGUAUCAUAUGCCAAGUACGCCCCCUAUUGACGUCAA
UGACGGUAAAUGGCCCGCCUGGCAUUGUGCCCAGUACAUGACCUUAUGGGACUUU
CCUACUUGGCAGUACAUCUACGUAAUAGUCAUCGCUAAUACCAUGGUCGAGGUGA
GCCCCACGUUCUGCUUCACUCUCCCCAUUCUCCCCCCCCUCCCCACCCCCAAUUUU
GUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUU
GGGGCGCGCGCCAGGCGGGGCGGGGCGGGGCGAGGGGCGGGGCGGGGCGAGGGCGG
AGAGGUGCGGCGGCGAGCCAAUCAGAGCGGCGCGCUCCGAAAGUUUCCUUUUUAUGG
CGAGGCGGCGGCGGCGGCGGCCUAAUAAAAGCGAAGCGCGGCGGGCGGGGAGU
CGCUGCGACGCGUCCUUCGCCCCGUGCCCCGCUCCGCGCGCGCCUCGCGCGCCCG
CCCCGGCUCUGACUGACCGCGUACUCCACAGGUGAGCGGGCGGGACGGCCCUU
CUCCUCCGGGCGUAAUAGCUGAGCAAGAGGUAAGGGUUUAAGGGAUGGUUGGU
UGGUGGGGUAAUUAUGUUUAUUAACCUGGAGCACCUGCCUGAAAUACUUUUUU
UCAGGUUGGCGUACGGCCACCAUGGUGAGCAAGGGGCGAGGAGCUGUUCACCGGGG
UGGUGCCCAUCCUGGUCGAGCUGGACGGCGACGUAAACGGCCACAAGUUCAGCGU
GUCCGGCGAGGGCGAGGGCGAUGCCACCUACGGCAAGCUGACCCUGAAGUUAUC
UGCACCACGGCAAGCUGCCCGUGCCCGGCCCCACCCUCGUGACCAACCUGACCUA
CGGCGUGCAGUGCUUCAGCCGCUACCCCGACCAUAUGAAGCAGCACGACUUCUUC
AAGUCCGCCAUGCCCGAAGGCUACGUCCAGGAGCGCACCAUCUUCUUAAGGACG
ACGGCAACUACAAGACCCGCGCCGAGGUGAAGUUCGAGGGCGACACCCUGGUGAA
CCGCAUCGAGCUGAAGGGCAUCGACUUAAGGAGGACGGCAACAUCUGGGGGCAC
AAGCUGGAGUACAACUACAACAGCCACAACGUCUAUAUCAUGGCCGACAAGCAGA
AGAACGGCAUCAAGGUGAACUUAAGAUCGCCACAACAUCGAGGACGGCAGCGU
GCAGCUCGCCGACCACUACCAGCAGAACACCCCAUUCGGCGACGGCCCCGUGCUGC
UGCCCGACAACCACUACCUGAGCACCCAGUCCGCCUGAGCAAAGACCCCAACGAG
AAGCGCGAUCACAUGGUCCUGCUGGAGUUCGUGACCGCCGCCGGGAUCACUCUCG
GCAUGGACGAGCUGUACAAGUAAAAGCUUGAUCCAGACAUGAUAAAGAUACAUGA
UGAGUUUGGACAAACCACAACUAGAAUGCAGUGAAAAAAAUAGCUUUUAUUUGUGAA
AUUUGUGAUGCUAUUGCUUUUAUUUGUAACCAUUAUAAGCUGCAAUAAACAAGUU
GUGGAAUUGUGAGCGCUCACAAUCCACAGCGGCCGCGUGAGGUAGAUAAUCUUUG

20

30

40

50

UAUAGUGGGGGGGGAUCUCAUGUACCGGGUUUCUUUUAUUUGAUUUUCAAUAAA
ACAGACGGUAGCUAGGUUCGCAAGGCAGCCACAAGCCAAAGAUAGGUAGGGUGCU
CAUAGUGAGUAGGGACAGUGCCUUUUGAUUCACAACGCGUCAAUACCAUCUGACA
CGGAUACCCUUACCGGACUUGUCAUGAUCUCCCAGACUUGUCCAAGGUGGACGGG
CCACCUUUACUUAACCCGGAAAAGGAACAUAUAUUAUUAUUAUGUGUUCGGAAAA
501. TCA5_CBhBsi_GFP_GeFo3_A22

gggagacgggucaaccgcguaggagcggugaccgGAAUUCGGCGGGAGUAACUAUGACUC
UCUUAAGGAGUCAUAGAGCCAGAACCUCGUGGUCCTCGUGGGCACAGGGGAUU
AAUUUUUCUGUGGCAAAUUGACUGGCUUCAGAGAGCGUUUUUCGAAGUGGACU
GUGUGACUGCGUUCCTCCCUUAGUUGCUAUAUCCGCUUCGAUUAACAUCUCACCU
CGACGUAAUAGAUAUUUCUAGAAGGUCGACCAGAUGUCCGAGGUCGACCAGUUG
UCCGUGUGGAAUUGUGAGCGCUCACAAUCCACACGUUACAUAACUACGGUAAA
UGGCTCGCCUGGCGACCGCCCAACGACCTCCGCCCAUUGACGUCAAUAGUAAACG
CCAAUAGGGACUUUCCAUUGACGUCAAUGGGUGGAGUAUUUACGGUAAACUGCCC
ACUUGGCAGUACAUAAGUGUAUCAUAUGCCAAGUACGCTCCCUAUUGACGUCAA
UGACGGUAAAUGGCTCGCCUGGCAUUGUGCCAGUACAUGACCUUAUGGGACUUU
CCUACUUGGCAGUACAUCUACGUAAUAGUCAUCGCUAAUACCAUGGUCGAGGUGA
GCTCCACGUUCUGCUUCACUCUCCCAUCUCCCCCCCCUCCCCACCTCCAAUUUU
GUUUUUUAUUUAUUUUUUAAUUAUUUUGUGCAGCGAUUGGGGGCGGGGGGGGGGG
GGGGCGCGCGCCAGGCGGGGGCGGGGGCGGGGGCGAGGGGGCGGGGGCGAGGGCGG
AGAGGUGCGGCGGCGAGCCAAUCAGAGCGGCGCGCUCCGAAAGUUUCCUUUUAUGG
CGAGGCGGCGGCGGCGGCGGCGCCUAUAAAAAGCGAAGCGCGCGGGCGGGGAGU
CGCUGCGACGCGUGCCUUCGCTCCGUGCTCCGCUCCGCGCGCCUCGCGCGCTCCG
CTCCGGCUCUGACUGACCGCGUACUCCACAGGUGAGCGGGCGGGACGGCTCUU
CUCCUCCGGGCGUGAAUUAAGCUGAGCAAGAGGUAAGGGUUUAAGGGAUGGUUGGU
UGGUGGGGUAAUUAUGUUUAUUAUACCUGGAGCACCUGCCUGAAAUACUUUUUU
UCAGGUUGGCGUACGGCCACCAUGGUGAGCAAGGGCGAGGAGCUGUUCACCGGGG
UGGUGCCCAUCCUGGUCGAGCUGGACGGCGACGUAAACGGCCACAAGUUCAGCGU
GUCCGGCGAGGGGCGAGGGCGAUGCCACCUACGGCAAGCUGACCCUGAAGUUAUC
UGCACCACCGGCAAGCUGCTCCGUGCTCCUGGCTCCACCTUCGUGACCACTCCUGACCUA
CGGCGUGCAGUGCUUCAGCTCGUACCTCCGACCACAUGAAGCAGCACGACUUCUUC
AAGUCCGCCAUGCTCCGAAGGCUACGUCCAGGAGCGCACCAUCUUCUUAAGGACG
ACGGCAACUACAAGACCTCGCGCCGAGGUGAAGUUCGAGGGCGACACCTUGGUGAA
CCGCAUCGAGCUGAAGGGCAUCGACUUAAGGAGGACGGCAACAUCUGGGGGCAC
AAGCUGGAGUACAACUACAACAGCCACAACGUCUAUAUCAUGGCCGACAAGCAGA
AGAACGGCAUCAAGGUGAACUUAAGAUCGCGCCACAACAUCGAGGACGGCAGCGU
GCAGCUCGCTGACCACUACCAGCAGAACACCTCCAUUCGGCGACGGCTCCGUGCUGC
UGCTCCGACAACCAUACCUAGACCTCCAGUCCGCTCCUGAGCAAAGACCTCAACGAG
AAGCGCGAUACAUGGUCCUGCUGGAGUUCGUGACCGCTCGCTGGGAUCACUCUCG
GCAUGGACGAGCUGUACAAGUAAAAGCUUGAUCCAGACAUGAUAAAGAUACAUGA
UGAGUUUGGACAAACCAACAACUAGAAUGCAGUGAAAAAAAUGCUUUUAUUUGUGAA
AUUUGUGAUGCUAUUGCUUUUAUUUGUAACCAUUAUAAGCUGCAAUAAACAAGUU
GUGGAAUUGUGAGCGCUCACAAUUCACAGCGGCTCGCUGAGGUAGAUAAUCUUUG
UAUAGUGGGGGGGGAUCUCAUGUACCGGGUUUCUUUUAUUUGAUUUUCAAUAAA
ACAGACGGUAGCUAGGUUCGCAAGGCAGCCACAAGCCAAAGAUAGGUAGGGUGCU
CAUAGUGAGUAGGGACAGUGCCUUUUGAUUCACAACGCGUCAAUACCAUCUGACA
CGGAUACCCUUACCGGACUUGUCAUGAUCUCCCAGACUUGUCCAAGGUGGACGGG
CCACCUUUACUUAACCCGGAAAAGGAACAUAUAUUAUUAUUAUGUGUUCGGAAAA
AAAAAAAAAAAAAAAAAAAAAAAAA

502. TCA5_CBhBsi_GFP_GeFo3_R4

10

20

30

40

50

gggagacgggucaaccgcuaggagcggugaccgGAAUUCGGCGGGAGUAACUAUGACUC
UCUUAAGGAGUCAUAGAGCCAGAACCUCUCGUGGUCCCGCUGGGCACAGGGAAU
AAUUUUUCUGUGGCAAAUUUGACUGGCUUCAGAGAGCGUUUUUCGAAGUGGACU
GUGUGACUGCGUUCUUUUUUAGUUGCUAUAUCCGCUUCGAUUAACAUCUCACCU
CGACGUAUAAGAUAUUAUUAAGAAGGUCGACCAGAUGUCCGAGGUCGACCAGUUG
UCCGUGUGGAAUUGUGAGCGCUCACAAUCCACACGUUACAUAACUUACGGUAAA
UGGCCCGCCUGGCGACCGCCCAACGACCCCGCCCAUUGACGUAUAUAGUAACG
CCAAUAGGGACUUUCCAUAUGACGUCAAUGGGUGGAGUAUUUACGGUAAACUGCCC
ACUUGGCAGUACAUAAGUGUAUAUAUGCCAAGUACGCCCCCUAUUGACGUCAA
UGACGGUAAAUGGCCCGCCUGGCAUUGUGCCCAGUACAUGACCUUAUGGGACUUU
CCUACUUGGCAGUACAUCUACGUAAUAGUCAUCGCUAAUUAACCAUGGUCGAGGUGA
GCCCCACGUUCUGCUUACUCUCCCCAUCUCCCCCCCCUCCCCACCCCCAAUUUU
GUAUUUAUUUAUUUUUUAAUUAUUUUUGUGCAGCGAUUGGGGGCGGGGGGGGGGGG
GGGGCGCGCGCCAGGCGGGGGCGGGGGCGGGGGCGAGGGGGCGGGGGCGAGGGCGG
AGAGGUGCGGGCGGCAGCCAAUCAGAGCGGGCGCGCUCCGAAAGUUUCCUUUUAUGG
CGAGGGCGGGCGGGCGGGCGGGCCUAUAAAAAGCGAAGCGCGGGCGGGCGGGAGU
CGCUGCGACGCGUGCCUUCGCCCCGUGCCCCGCUCCGCGCGCCGCUUGCGCGCCGCCC
CCCCGGCUCUGACUGACCGCGUUAUCUCCACAGGUGAGCGGGCGGGACGGCCCUU
CUCCUCCGGGCGUGAAUUAAGCUGAGCAAGAGGUAAGGGUUUAAGGGAUGGUUGGU
UGGUGGGGUAUUAAUGUUUAUUUACCUGGAGCACCUGCCUGAAAUACAUUUUUU
UCAGGUUGGCGUACGGCCACCAUGGUGAGCAAGGGCGAGGAGCUGUUCACCGGGG
UGGUGCCCAUCCUGGUCGAGCUGGACGGCGACGUAAACGGCCACAAGUUCAGCGU
GUCCGGCGAGGGCGAGGGCGAUGCCACCUACGGCAAGCUGACCCUGAAGUUCAU
UGCACCACCGGCAAGCUGCCCGUGCCCGUGGCCACCCUCGUGACCACCCUGACCUA
CGGGCGUGCAGUGCUUCAGCCGCUACCCCGACCACAUGAAGCAGCACGACUUCUUC
AAGUCCGCCAUGCCCGAAGGCUACGUCCAGGAGCGCACCAUCUUCUUAAGGACG
ACGGCAACUACAAGACCCGCGCCGAGGUGAAGUUCGAGGGCGACACCCUGGUGAA
CCGCAUCGAGCUGAAGGGCAUCGACUUAAGGAGGACGGCAACAUCCUGGGGGCAC
AAGCUGGAGUACAACUACAACAGCCACAACGUCUAUAUCAUGGCCGACAAGCAGA
AGAACGGCAUCAAGGUGAACUUAAGAUCGCGCCACAACAUCGAGGACGGCAGCGU
GCAGCUCGCCGACCACUACCAGCAGAACACCCCAUCGGCGACGGCCCCGUGCUGC
UGCCCGACAACCACUACCUGAGCACCCAGUCCGCCUGAGCAAAGACCCCAACGAG
AAGCGCGAUACAUGGUCCUGCUGGAGUUCGUGACCGCGCGGGGAUACUCUCG
GCAUGGACGAGCUGUACAAGUAAAAGCUUGAUCCAGACAUGAUAAAGAUACAUGA
UGAGUUUGGACAAACCACAACUAGAAUGCAGUGAAAAAAUAGCUUUUAUUUGUGAA
AUUUGUGAUGCUAUUGCUUUUAUUUGUAACCAUUAUAAGCUGCAAUAAACAAGUU
GUGGAAUUGUGAGCGCUCACAAUCCACAGCGGCCGCGUGAGGUAGAUAAUCUUUG
UAUAGUGGGGGGGGAUCUCAUGUACCGGGUUUCUUUUUAUUUGAUUUUCAUAAA
ACAGACGGUAGCUAGGUUCGCAAGGCAGCCACAAGCCAAAGAUAGGUAGGGUGCU
CAUAGUGAGUAGGGACAGUGCCUUUUGAUUACACGCGUCAAUACCAUCUGACA
CGGAUACCCUUAACGGACUUGUCAUGAUCUCCAGACUUGUCCAAGGUGGACGGG
CCACCUUUACUUAACCCGGAAAAGGAACAUAUAUUAAUUAUUGUGUUCGGAAAA
UAGC

10

20

30

40

50

503. TCA5_CBhBsi_GFP_GeFo3_R4A22

gggagacgggucaaccgcuaggagcggugaccgGAAUUCGGCGGGAGUAACUAUGACUC
UCUUAAGGAGUCAUAGAGCCAGAACCUCUCGUGGUCCCGCUGGGCACAGGGAAU
AAUUUUUCUGUGGCAAAUUUGACUGGCUUCAGAGAGCGUUUUUCGAAGUGGACU
GUGUGACUGCGUUCUUUUUUAGUUGCUAUAUCCGCUUCGAUUAACAUCUCACCU
CGACGUAUAAGAUAUUAUUAAGAAGGUCGACCAGAUGUCCGAGGUCGACCAGUUG
UCCGUGUGGAAUUGUGAGCGCUCACAAUCCACACGUUACAUAACUUACGGUAAA

UGGCCCGCCUGGCUGACCGCCCAACGACCCCCGCCCAUUGACGUCAAUAGUAACG
CCAAUAGGGACUUUCCAUGACGUCAAUGGGUGGAGUAUUUACGGUAAACUGCCC
ACUUGGCAGUACAUCAAGUGUAUCAUAUGCCAAGUACGCCCCCUAUUGACGUCAA
UGACGGUAAAUGGCCCGCCUGGCAUUGUGCCCAGUACAUGACCUUAUGGGACUUU
CCUACUUGGCAGUACAUCUACGUAAUAGUCAUCGCUAAUACCAUGGUCGAGGUGA
GCCCCACGUUCUGCUUCACUCUCCCCAUCUCCCCCCCCUCCCCACCCCCAAUUUU
GUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUU
GGGGCGCGCGCCAGGCGGGGCGGGGCGGGGCGAGGGGCGGGGCGGGGCGAGGCGG
AGAGGUGCGGCGGCAGCCAAUCAGAGCGGCGCGCUCCGAAAGUUUCCUUUUUUGG
CGAGGCGGCGGCGGCGGCGGCCUAUAAAAAGCGAAGCGCGGCGGGCGGGAGU
CGCUGCGACGCUGCCUUCGCCCCGUGCCCCGCUCCGCGCGCCUCGCGCCGCCCG
CCCCGGCUCUGACUGACCGCGUACUCCACAGGUGAGCGGGCGGGACGGCCCUU
CUCCUCCGGGCUGUAAUAGCUGAGCAAGAGGUAAGGGUUUAAGGGAUGGUUGGU
UGGUGGGGUAAUAAUGUUUAAUUAACCUGGAGCACCUGCCUGAAAUACUUUUUU
UCAGGUUGGCGUACGGCCACCAUGGUGAGCAAGGGCGAGGAGCUGUUCACCGGGG
UGGUGCCCAUCCUGGUCGAGCUGGACGGCGACGUAAACGGCCACAAGUUCAGCGU
GUCCGGCGAGGGCGAGGGCGAUGCCACCUACGGCAAGCUGACCCUGAAGUUCAUC
UGCACCACCGGCAAGCUGCCCGUGCCCUUGGCCACCCUCGUGACCACCCUGACCUA
CGGCGUGCAGUGCUUCAGCCGCUACCCCGACCACAUGAAGCAGCACGACUUCUUC
AAGUCCGCCAUGCCCGAAGGCUACGUCCAGGAGCGCACCAUCUUCUUAAGGACG
ACGGCAACUACAAGACCCGCGCCGAGGUGAAGUUCGAGGGCGACACCCUGGUGAA
CCGCAUCGAGCUGAAGGGCAUCGACUUAAGGAGGACGGCAACAUCUGGGGCAC
AAGCUGGAGUACAACUACAACAGCCACAACGUCUAUAUCAUGGCCGACAAGCAGA
AGAACGGCAUCAAGGUGAACUUAAGAUCGCCACAACAUCGAGGACGGCAGCGU
GCAGCUCGCCGACCACUACCAGCAGAACACCCCCAUCGGCGACGGCCCCGUGCUGC
UGCCCGACAACCACUACCUGAGCACCCAGUCCGCCCUAGAGCAAAGACCCCAACGAG
AAGCGCGAUCACAUGGUCCUGCUGGAGUUCGUGACCGCCGCCGGGAUCACUCUCG
GCAUGGACGAGCUGUACAAGUAAAAGCUUGAUCCAGACAUGAUAAAGAUACAUGA
UGAGUUUGGACAAACCACAACUAGAAUGCAGUGAAAAAAAUGCUUUUUUUUGUGAA
AUUUUGUGAUGCUAUUGCUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUU
GUGGAAUUGUGAGCGCUCACAAUUCACAGCGGCCGCGUGAGGUAGAUAAUCUUUG
UAUAGUGGGGGGGGAUCUCAUGUACCGGGUUUCUUUUUUUUUUUUUUUUUUUUUU
ACAGACGGUAGCUAGGUUCGCAAGGCAGCCACAAGCCAAAGAUAGGUAGGGUGCU
CAUAGUGAGUAGGGACAGUGCCUUUUGAUUACACAACGCGUCAAUACCAUCUGACA
CGGAUACCCUUACCGGACUUGUCAUGAUCUCCAGACUUGUCCAAGGUGGACGGG
CCACCUUUACUUAACCCGGAAGGAACAUAUAUUAUUAUUAUGUGUUCGGAAAA
UAGCAAAAAAAAAAAAAAAAAAAAAAAAAA

10

20

30

40

50

504. TCA5_CBhBsi_GFP_GeFo3_R20

gggagacgggucaaccgcuaggagcggugaccgGAAUUCGGCGGGAGUAACUAUGACUC
UCUUAAGGAGUCAUAGAGCCAGAACCUCUCGUGGUCCCGCUGGGCACAGGGAUU
AAUUUUUCUGUGGC AAAUUUGACUGGCUUCAGAGAGCGUUUUUCGAAGUGGACU
GUGUGACUGCGUUCUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUU
CGACGUAAUAGAUCAUUUCUAGAAGGUCGACCAGAUGUCCGAGGUCGACCAGUUG
UCCGUGUGGAAUUGUGAGCGCUCACAAUUCACACGUUACAUAACUACGGUAAA
UGGCCCGCCUGGCUGACCGCCCAACGACCCCCGCCCAUUGACGUCAAUAGUAACG
CCAAUAGGGACUUUCCAUGACGUCAAUGGGUGGAGUAUUUACGGUAAACUGCCC
ACUUGGCAGUACAUCAAGUGUAUCAUAUGCCAAGUACGCCCCCUAUUGACGUCAA
UGACGGUAAAUGGCCCGCCUGGCAUUGUGCCCAGUACAUGACCUUAUGGGACUUU
CCUACUUGGCAGUACAUCUACGUAAUAGUCAUCGCUAAUACCAUGGUCGAGGUGA
GCCCCACGUUCUGCUUCACUCUCCCCAUCUCCCCCCCCUCCCCACCCCCAAUUUU

GUAUUUAUUUAUUUUUUAAUUAUUUUGUGCAGCGAUGGGGGCGGGGGGGGGGGG
GGGGCGCGCGCCAGGCGGGGCGGGGCGGGGCGAGGGGCGGGGCGAGGCGG
AGAGGUGCGGCGGCAGCCAAUCAGAGCGGCGCGCUCCGAAAGUUUCCUUUUAUGG
CGAGGCGGCGGCGGCGGCGGCCCCUAUAAAAAGCGAAGCGCGCGGGCGGGGAGU
CGCUGCGACGCUGCCUUCGCCCCGUGCCCCGCUCCGCCGCCGCCUCGCGCCGCCCG
CCCCGGCUCUGACUGACCGCGUACUCCACAGGUGAGCGGGCGGGACGGCCCUU
CUCCUCCGGGCGUGAAUUAAGCUGAGCAAGAGGUAAGGGUUUAAGGGAUGGUUGGU
UGGUGGGGUAAUAAUGUUUAAUUAACCUGGAGCACCUGCCUGAAAUCACUUUUUU
UCAGGUUGGCGUACGGCCACCAUGGUGAGCAAGGGCGAGGAGCUGUUCACCGGGG
UGGUGCCCAUCCUGGUCGAGCUGGACGGCGACGUAAACGGCCACAAGUUCAGCGU
GUCCGGCGAGGGGCGAGGGCGAUGCCACCUACGGCAAGCUGACCCUGAAGUUCAUC
UGCACCACCGGCAAGCUGCCCGUGCCCGUGGCCACCCUCGUGACCACCCUGACCUA
CGGCGUGCAGUGCUUCAGCCGCUACCCCGACCACAUGAAGCAGCACGACUUCUUC
AAGUCCGCCAUGCCCCGAAGGCUACGUCCAGGAGCGCACCAUCUUCUUAAGGACG
ACGGCAACUACAAGACCCGCGCCGAGGUGAAGUUCGAGGGCGACACCCUGGUGAA
CCGCAUCGAGCUGAAGGGCAUCGACUUAAGGAGGACGGCAACAUCUGGGGCGAC
AAGCUGGAGUACAACUACAACAGCCACAACGUCUAUAUCAUGGCCGACAAGCAGA
AGAACGGCAUCAAGGUGAACUUAAGAUCGCCACAACAUCGAGGACGGCAGCGU
GCAGCUCGCCGACCACUACCAGCAGAACACCCCCAUCGGCGACGGCCCCGUGCUGC
UGCCCGACAACCACUACCUGAGCACCCAGUCCGCCUGAGCAAAGACCCCAACGAG
AAGCGCGAUCACAUGGUCCUGCUGGAGUUCGUGACCGCCGCCGGGAUCACUCUCG
GCAUGGACGAGCUGUACAAGUAAAAGCUUGAUCCAGACAUGAUAAAGAUACAUGA
UGAGUUUGGACAAACCACAACUAGAAUGCAGUGAAAAAAUAGCUUUAUUUGUGAA
AUUUGUGAUGCUAUUGCUUUAUUUGUAACCAUUAUAAGCUGCAAUAAACAAGUU
GUGGAAUUGUGAGCGCUCACAAUUCACAGCGGCCGCGUGAGGUAGAUAAUCUUG
UAUAGUGGGGGGGGAUCUCAUGUACCGGGUUUCUUUUAUUUGAUUUUCAUAAA
ACAGACGGUAGCUAGGUUCGCAAGGCAGCCACAAGCCAAAGAUAGGUAGGGUGCU
CAUAGUGAGUAGGGACAGUGCCUUUUGAUUCACAACGCGUCAAUACCAUCUGACA
CGGAUACCCUUAACGGGACUUGUCAUGAUCUCCAGACUUGUCCAAGGUGGACGGG
CCACCUUUAUUAACCCGGAAAAGGAACAUAUAUUAUUAUUGUGUUCGGAAAA
UAGCCAAUUGCCUCGUCAUC

10

20

30

40

50

505. TCA5_CBhBsi_GFP_GeFo3_R20A22

gggagacggucaaccgcuaggagcgguagaccgGAAUUCGGCGGGAGUAACUAUGACUC
UCUUAAGGAGUCAUAGAGCCAGAACCUCUCGUGGUCCCGCUGGGCACAGGGGAUU
AAUUUUUCUGUGGCAAAUUGACUGGCUUCAGAGAGCGUUUUUCGAAGUGGACU
GUGUGACUGCGUUCUUUUUUUAGUUGCUAUAUCCGCUUCGAUUAACAUCUCACCU
CGACGUAAAGAUCAUUUCUAGAAGGUCGACCAGAUGUCCGAGGUCGACCAGUUG
UCCGUGUGGAAUUGUGAGCGCUCACAAUUCACACGUUACAUAACUACGGUAAA
UGGCCCGCCUGGCUGACCGCCCAACGACCCCCGCCCAUUGACGUCAAUAGUAACG
CCAAUAGGGACUUUCCAUUGACGUCAAUGGGUGGAGUAUUUACGGUAAACUGCCC
ACUUGGCAGUACAUAAGUGUAUCAUAUGCCAAGUACGCCCCCUAUUGACGUCAA
UGACGGUAAAUGGCCCGCCUGGCAUUGUGCCCAGUACAUGACCUUAUGGGACUUU
CCUACUUGGCAGUACAUCUACGUAAUAGUCAUCGCUAAUACCAUGGUCGAGGUGA
GCCCCACGUUCUGCUUCACUCUCCCCAUCUCCCCCCCCUCCCCACCCCCAAUUUU
GUAUUUAUUUAUUUUUUAAUUAUUUUGUGCAGCGAUGGGGGCGGGGGGGGGGGG
GGGGCGCGCGCCAGGCGGGGCGGGGCGGGGCGAGGGGCGGGGCGGGGCGAGGCGG
AGAGGUGCGGCGGCAGCCAAUCAGAGCGGCGCGCUCCGAAAGUUUCCUUUUAUGG
CGAGGCGGCGGCGGCGGCGGCCCCUAUAAAAAGCGAAGCGCGCGGGCGGGGAGU
CGCUGCGACGCUGCCUUCGCCCCGUGCCCCGCUCCGCCGCCGCCUCGCGCCGCCCG
CCCCGGCUCUGACUGACCGCGUACUCCACAGGUGAGCGGGCGGGACGGCCCUU

CUCCUCCGGGCUGUAAUAGCUGAGCAAGAGGUAAGGGUUUAAGGGAUGGUUGGU
UGGUGGGGUAAUAAUGUUUAAUUAACUGGAGCACCUGCCUGAAAUACAUUUUUU
UCAGGUUGGCGUACGGCCACCAUGGUGAGCAAGGGCGAGGAGCUGUUCACCGGGG
UGGUGCCCAUCCUGGUCGAGCUGGACGGCGACGUAAACGGCCACAAGUUCAGCGU
GUCCGGCGAGGGGCGAGGGCGAUGCCACCUACGGCAAGCUGACCCUGAAGUUCAUC
UGCACCACCGGCAAGCUGCCCUGGCCACCCUCGUGACCACCCUGACCUA
CGGCGUGCAGUGCUUCAGCCGCUACCCCGACCACAUGAAGCAGCACGACUUCUUC
AAGUCCGCCAUGCCCGAAGGCUACGUCCAGGAGCGCACCAUCUUCUUAAGGACG
ACGGCAACUACAAGACCCGCGCCGAGGUGAAGUUCGAGGGGCGACACCCUGGUGAA
CCGCAUCGAGCUGAAGGGCAUCGACUUAAGGAGGACGGCAACAUCUGGGGGCAC
AAGCUGGAGUACAACUACAACAGCCACAACGUCUAUAUCAUGGCCGACAAGCAGA
AGAACGGCAUCAAGGUGAACUUAAGAUCGCCACAACAUCGAGGACGGCAGCGU
GCAGCUCGCCGACCACUACCAGCAGAACACCCCCAUCGGCGACGGCCCCGUGCUGC
UGCCCGACAACCACUACCUGAGCACCCAGUCCGCCUGAGCAAAGACCCCAACGAG
AAGCGCGAUCACAUGGUCCUGCUGGAGUUCGUGACCGCCGCCGGGAUCACUCUCG
GCAUGGACGAGCUGUACAAGUAAAAGCUUGAUCCAGACAUGAUAAAGAUACAUGA
UGAGUUUGGACAAACCACAACUAGAAUGCAGUGAAAAAAUAGCUUUAUUUGUGAA
AUUUGUGAUGCUAUUGCUUUAUUUGUAACCAUUAUAAGCUGCAAUAAACAAGUU
GUGGAAUUGUGAGCGCUCACAAUUCACAGCGGCCGCGUGAGGUAGAUAAUCUUG
UAUAGUGGGGGGGGAUCUCAUGUACCGGGUUUCUUUUAUUUGAUUUUCAUAAA
ACAGACGGUAGCUAGGUUCGCAAGGCAGCCACAAGCCAAAGAUAGGUAGGGUGCU
CAUAGUGAGUAGGGACAGUGCCUUUGAUUCACAACGCGUCAAUACCAUCUGACA
CGGAUACCCUUAACCGGACUUGUCAUGAUUCUCCAGACUUGUCCAAGGUGGACGGG
CCACCUUUACUUAACCCGGAAAAGGAACAUAUAUUAUUAUUAUGUGUUCGGAAAA
UAGCCAAAUGCCUCGUCAUC AAAAAAAAAAAAAAAAAAAAAA

10

20

30

40

50

506. CBhBsi_GPF_GeFo3_R4A22

AGGUCGACCAGAUGUCCGAGGUCGACCAGUUGUCCGUGUGGAAUUGUGAGCGCUC
ACAAUUCCACACGUUACAUAACUUAACGGUAAAUGGCCCGCCUGGCUGACCGCCCA
ACGACCCCCGCCCAUUGACGUCAAUAGUAACGCCAAUAGGGACUUAUCCAUGACG
UCAUUGGGUGGAGUAUUUACGGUAAACUGCCCACUUGGCAGUACAUAAGUGUAU
CAUAUGCCAAGUACGCCCCCUAUUGACGUCAAUGACGGUAAAUGGCCCGCCUGGC
AUUGUGCCCAGUACAUGACCUUAUGGGACUUAUCCUACUUGGCAGUACAUCUACGU
AUUAGUCAUCGCUAUUAACCAUGGUCGAGGUGAGCCCCACGUUCUGCUUCACUCUC
CCCAUCUCCCCCCCCUCCCCACCCCCAAUUUUGUAUUUAUUUAUUUUUAAUUAU
UUUGUGCAGCGAUGGGGGGCGGGGGGGGGGGGGGGGGGGCGCGCGCCAGGCGGGGCGGG
GCGGGGCGAGGGGCGGGGCGGGGCGAGGCGGAGAGGUGCGGCGGCAGCCAUCAG
AGCGGCGCGCUCCGAAAGUUUCCUUUAUGGCGAGGCGGCGGCGGCGGGCCCU
AUAAAAGCGAAGCGCGCGGGCGGGGAGUCGCUUGCGACGCUUGCCUUCGCCCCG
UGCCCCGCUCCGCGCGCCUCGCGCCGCCCGCCCCGGCUCUGACUGACCGCGUUA
CUCCCACAGGUGAGCGGGCGGGACGGCCCUUCUCCUCCGGGCUGUAAUUAAGCUGA
GCAAGAGGUAAGGGUUUAAGGGAUGGUUGGUUGGUGGGGUAAUUAUGUUUAAUU
ACCUGGAGCACCUGCCUGAAAUACAUUUUUUUCAGGUUGGCGUACGGCCACCAUG
GUGAGCAAGGGCGAGGAGCUGUUCACCGGGGUGGUGCCCAUCCUGGUCGAGCUGG
ACGGCGACGUAAACGGCCACAAGUUCAGCGUGUCCGGCGAGGGCGAGGGCGAUGC
CACCUACGGCAAGCUGACCCUGAAGUUAUCUGCACCACCGGCAAGCUGCCCGUG
CCCUGGCCACCCUCGUGACCACCCUGACCUACGGCGUGCAGUGCUUCAGCCGCU
ACCCCGACCACAUGAAGCAGCACGACUUCUUAAGUCCGCCAUGCCCGAAGGCUA
CGUCCAGGAGCGCACCAUCUUCUUAAGGACGACGGCAACUACAAGACCCGCGCC
GAGGUGAAGUUCGAGGGCGACACCCUGGUGAACCGCAUCGAGCUGAAGGGCAUCG
ACUUCAAGGAGGACGGCAACAUCUGGGGCAACAAGCUGGAGUACAACUACAACAG

CCACAACGUCUAUAUCAUGGCCGACAAGCAGAAGAACGGCAUCAAGGUGAACUUC
AAGAUCCGCCACAACAUCGAGGACGGCAGCGUGCAGCUCGCCGACCACUACCAGCA
GAACACCCCCAUCGGCGACGGCCCCGUGCUGCUGCCCGACAACCACUACCUGAGCA
CCCAGUCCGCCUGAGCAAAGACCCCAACGAGAAGCGCGAUCACAUGGUCCUGCU
GGAGUUCGUGACCGCCGCCGGGAUCACUCUCGGCAUGGACGAGCUGUACAAGUAA
AAGCUUGAUCCAGACAUGAUAAGAUACAUAUGAUGAGUUUGGACAAACCACAACUA
GAAUGCAGUGAAAAAAUAGCUUUAUUUGUGAAAAUUUGUGAUGCUAUUGCUUUAU
UUGUAACCAUUAUAAGCUGCAAUAAACAAGUUGUGGAAUUGUGAGCGCUCACAAU
UCCACAGCGGCCCGCUGAGGUAGAUAAUCUUUGUAUAGUGGGGGGGGAUCUCAUGU
ACCGGGUUUCUUUUAUUUGAUAUUUCAUAUAAAACAGACGGUAGCUAGGUUCGCAA
GGCAGCCACAAGCCAAAGAUAGGUAGGGUGCUCAUAGUGAGUAGGGACAGUGCCU
UUUGAUUCACAACGCGUCAAUACCAUCUGACACGGAUACCCUUACCGGACUUGUC
AUGAUCUCCAGACUUGUCCAAGGUGGACGGGCCACCUUUAACUUAACCCGGAAAA
GGAACAUAUAUUAUUAUUAUGUGUUCGGAAAAUAGCAAAAAAAAAAAAAAAAAAAA
AAA

10

507. TCA5_CBh_NLSGPF_ZoAl3_R4A22

gggagacgggucaaccgcguaggagcggugaccgGAAUUCGGCGGGAGUAACUAUGACUC
UCUUAAGGAGUCAUAGAGCCAGAACCUCUCGUGGUCCCGCUGGGGCACAGGGGAUU
AAUUUUUCUGUGGCAAAUUGACUGGCUUCAGAGAGCGUUUUUCGAAGUGGACU
GUGUGACUGCGUUCUUUUUAGUUGCUAUUAUCCGCUUCGAUUAACAUCUCACCU
CGACGUUAUAGAUAUU TCTAGAAGGTCGACCAGATGTCCGAGGTTCGACCAGTTG
TCCGTGTGGAATTGTGAGCGCTCACAATTCCACA Cgt tacataactt acggtaaattg g
ccgccttg ctgaccgccc aacgaccccc gccattgac gtcaatagta acgccaatag gga
ctttcca ttgacgtcaa tgggtggagt atttacggta aactgcccac ttggcagtac atcaagt
ta tcatatgcca agtacgcccc ctattgacgt caatgacggt aaatggcccg cctggcattg tg
cccagtac atgaccttat gggactttcc tacttggcag tacatctacg tattagtcat cgctatta
cc atggtcgagg tgagccccac gttctgttc acttccccca tctcccccc ctccccacc cc
aattttgt atttatttat tttttaatta ttttgtgcag cgatgggggc gggggggggg ggggggc
gcg cgccaggcgg ggcggggcgg ggcgaggggc ggggcggggc gaggcggaga ggtgcg
gchg cagccaatca gagcggcgcg ctccgaaagt tctttttat ggcgaggcgg cggcggcg
gc ggccctataa aaagcgaagc gcgcggcggg cgggagtcgc tgcgacgctg cttcgcccc
gtgccccgt cgcgcggcg ctcgcggcg cgcggcggc tctgactgac cgcgttact cca
cagggtga gcgggcggga cggcccttct ctcggggct gtaattagct gagcaagagg taaggg
ttta agggatggtt ggttggtggg gtattaatgt ttaattacct ggagcacctg cctgaaatca
cttttttca ggttgga agctt GCCACCATG gcc cca aag aag aag cgg aag gtcgtg
agcaaggggcgaggagctgttcaccggggtggtgcccattcctggtcgagctggacggcgacgtaaac
ggccacaagttcagcgtgtccggcgagggcgagggcgatgccacctacggcaagctgacctgaagt
tcatctgcaccaccggcaagctgcccgtgccctggcccaccctcgtgaccacctgacctacggcgtg
cagtgttccagccgctaccccgaccacatgaagcagcagcacttcttcaagtccgccatgccgaagg
ctacgtccaggagcgcaccatcttcttcaaggacgacggcaactacaagaccgcgcccagggtgaag
ttcgagggcgacaccctggtgaaccgcatcgagctgaaggcgatcgacttcaaggaggacggcaaca
tcctggggcacaagctggagtacaactacaacagccacaacgtctatatcatggccgacaagcagaa
gaacggcatcaaggtgaacttcaagatccgccacaacatcgaggacggcagcgtgcagctcgccgac
cactaccagcagaacacccccatcggcgacggccccgtgctgctgcccgacaaccactacctgagca
cccagtcgcccctgagcaaaagaccccaacgagaagcgcgatcacatggctcctgctggagtctgtgac
cgccgcccggatcactctcgcatggacgagctgtacaagtaaaagcttGatccagacatgataaga
tacattgatgagtttggacaaaccacaactagaatgcagtgaaaaaaatgctttatttgtgaaatttgt
gatgctattgctttatttgttaaccattataagctgcaataaacaagtTGTGGAATTGTGAGCGCTC
ACAATTCCACAGCGGCCGCTaggttagtcacattgcactttctgtaacttgcactgggtgtgggat
gtgggcctgggggtgtgggttatgggggtatataatgtgggatattctggtgggaatgtccattcactgtat

20

30

40

50

gcctatctttttaataaaaaagacggtagctaggttcgcgaagcagccacaagccaatagccagttagg
tagctcatagtgggtaggtgacaggaacctttgactcagaacgcgtccattaacatctagaacggacc
aaacttcggacatgcaccgattaaccggatttgtccaaggtggacggggccacctttacttaacccgga
aagggaacatatatagttatatgtgttcgtaata TAGC AAAAAAAAAAAAAAAAAAAAAA
508. TCA5_CBh_NLSGPF_GeFo3_R4A22

gggagacgggucaaccgcguaggagcggugaccgGAAUUCGGCGGGAGUAACUAUGACUC
UCUUAAGGAGUCAUAGAGCCAGAACCUCUCGUGGUCCCGCUGGGCACAGGGGAUU
AAUUUUUCUGUGGCAAAUUUGACUGGCUUCAGAGAGCGUUUUUCGAAGUGGACU
GUGUGACUGCGUUCUUUUUUAGUUGCUAUUAUCCGCUUCGAUUAACAUCUCACCU
CGACGUUAUAAGAUAUU TCTAGAAGGTCGACCAGATGTCCGAGGTCGACCAGTTG
TCCGTGTGGAATTGTGAGCGCTCACAATTCCACA Cgt tacataactt acggtaaatg g
cccgccctgg ctgaccgccc aacgaccccc gccattgac gtcaatagta acgccaatag gga
ctttcca ttgacgtcaa tgggtggagt atttacggta aactgcccac ttggcagtac atcaagt
ta tcatatgcca agtacgcccc ctattgacgt caatgacggt aaatggcccc cctggcattg tg
cccagtac atgaccttat gggactttcc tacttggcag tacatctacg tattagtcac cgctatta
cc atgggtcaggg tgagccccac gttctgtctt actctccca tctcccccc ctccccacc cc
aattttgt atttatttat tttttaatta ttttgtgcag cgatgggggc gggggggggg ggggggc
gcg cgccaggcgg ggcggggcgg ggcgaggggc ggggcggggc gaggcggaga ggtgcg
gcgg cagccaatca gagcggcgcg ctccgaaagt ttccttttat ggcgagggcgg cggcggcg
gc ggccctataa aaagcgaagc gcgcggcggg cgggagtcgc tgcgacgctg ctttcgcccc
gtgccccgct ccgcccgcgc ctgcgcgcgc ccgccccggc tctgactgac cgcgttact cca
caggtga gcgggcggga cggcccttct cctccgggct gtaattagct gagcaagagg taaggg
ttta agggatgggt ggttgggtgg gtattaatgt ttaattacct ggagcacctg cctgaaatca
ctttttttca ggttggga agctt GCCACCATG gcc cca aag aag aag cgg aag gtcgtg
agcaagggcgaggagctgttcaccggggtgggtgcccatcctggtcgagctggacggcgacgtaaac
ggccacaagttcagcgtgtccggcgagggcgagggcgatgccacctacggcaagctgacctgaagt
tcatctgcaccaccggcaagctgcccgtgccctggcccacctcgtgaccacctgacctacggcgtg
cagtgtctcagccgctaccccgaccacatgaagcagcagcacttcttcaagtccgccatgccgaagg
ctacgtccaggagcgcaccatcttcttcaaggacgacggcaactacaagacccgcgcccagggatgaag
ttcgagggcgacaccttggtgaaccgcatcgagctgaagggcacgacttcaaggaggacggcaaca
tcctggggcacaagctggagtacaactacaacagccacaacgtctatatcatggccgacaagcagaa
gaacggcatcaaggtgaacttcaagatccgccacaacatcgaggacggcagcgtgcagctcgccgac
cactaccagcagaacacccccatcgggcgacggccccgtgctgctgcccgacaaccactacctgagca
cccagtcgcccctgagcaaagaccccaacgagaagcgcgatcacatggctcctgctggagtctgtgac
cgccgcccggatcactctcggcacggagctgtacaagtaaaagcttGatccagacatgataaga
tacattgatgagtttggacaaaccacaactagaatgcagtgaaaaaatgctttatttgtgaaatttgt
gatgctattgctttatttgttaaccattataagctgcaataaacaagtTGTGGAATTGTGAGCGCTC
ACAATTCCACAGCGGCCGCGUGAGGUAGAUAAUCUUUGUAUAGUGGGGGGGGAUCU
CAUGUACCGGGUUCUUUUUAUUUGAUUUUCAUAAAACAGACGGUAGCUAGGUU
CGCAAGGCAGCCACAAGCCAAAGAUAGGUAGGGUGCUCAUAGUGAGUAGGGACAG
UGCCUUUUGAUUCACAACGCGUCAUACCAUCUGACACGGAUACCCUUACCGGAC
UUGUCAUGAUCUCCAGACUUGUCCAAGGUGGACGGGCCACCUUUACUUAACCCG
GAAAAGGAACAUAUAUUAUUAUUAUGUGUUCGGAAAAUAGCAAAAAAAAAAAAAA
AAAAAAAAA

509. TCA5_ZoAl3

gggagacgggucaaccgcguaggagcggugaccgGAAUUCGGCGGGAGUAACUAUGACUC
UCUUAAGGAGUCAUAGAGCCAGAACCUCUCGUGGUCCCGCUGGGCACAGGGGAUU
AAUUUUUCUGUGGCAAAUUUGACUGGCUUCAGAGAGCGUUUUUCGAAGUGGACU
GUGUGACUGCGUUCUUUUUUAGUUGCUAUUAUCCGCUUCGAUUAACAUCUCACCU
CGACGUUAUAAGAUAUUUCUAGAAGGUCGACCAGAUGUCCGAGGUCGACCAGUUG

10

20

30

40

50

UCCGUGUGGAAUUGUGAGCGCUCACAAUUCCACACGUUACAUAACUUACGGUAAA
UGGCCCGCCUGGCUGACCGCCCAACGACCCCCGCCCAUUGACGUCAAUAGUAACG
CCAAUAGGGACUUUCCAUUGACGUCAAUGGGUGGAGUAUUUACGGUAAACUGCCC
ACUUGGCAGUACAUCAAGUGUAUCAUAUGCCAAGUACGCCCCCUAUUGACGUCAA
UGACGGUAAAUGGCCCGCCUGGCAUUGUGCCCAGUACAUGACCUUAUGGGACUUU
CCUACUUGGCAGUACAUCUACGUAAUAGUCAUCGCUAAUUAACCAUGGUCGAGGUGA
GCCCCACGUUCUGCUUCACUCUCCCCAUCUCCCCCCCCUCCCCACCCCCAAUUUU
GUAUUUAUUUAUUUUUUAAUUAAUUUUUGUGCAGCGAUUGGGGGCGGGGGGGGGGG
GGGGCGCGCGCCAGGCGGGGCGGGGCGGGGCGAGGGGCGGGGCGGGGCGAGGCGG
AGAGGUGCGGCGGCAGCCAAUCAGAGCGGCGCGCUCCGAAAGUUUCCUUUUUUGG
CGAGGCGGGCGGCGGCGGCGGCCUAAUAAAAGCGAAGCGCGCGGGCGGGGAGU
CGCUGCGACGCGUGCCUUCGCCCCGUGCCCCGCUCCGCGCGCCGCUUGCGCGCCCG
CCCCGGCUCUGACUGACCGCGUACUCCCACAGGUGAGCGGGCGGGACGGCCCUU
CUCCUCCGGGCGUGAAUAGCUGAGCAAGAGGUAAAGGGUUUAAGGGAUGGUUGGU
UGGUGGGGUAAUAAUGUUUAAUUAACCUGGAGCACCUGCCUGAAAUCACUUUUU
UCAGGUUGGCGUACGGCCACCAUGGUGAGCAAGGGCGAGGAGCUGUUCACCGGGG
UGGUGCCCAUCCUGGUCGAGCUGGACGGCGACGUAAACGGCCACAAGUUCAGCGU
GUCCGGCGAGGGGCGAGGGCGAUGCCACCUACGGCAAGCUGACCCUGAAGUUCAUC
UGCACCACCGGCAAGCUGCCCGUGCCCUUGGCCACCCUCGUGACCACCCUGACCUA
CGGCGUGCAGUGCUUCAGCCGCUACCCCGACCACAUGAAGCAGCACGACUUCUUC
AAGUCCGCCAUGCCCGAAGGCUACGUCCAGGAGCGCACCAUCUUCUUAAGGACG
ACGGCAACUACAAGACCCGCGCCGAGGUGAAGUUCGAGGGCGACACCCUGGUGAA
CCGCAUCGAGCUGAAGGGCAUCGACUUAAGGAGGACGGCAACAUCCUGGGGCAC
AAGCUGGAGUACAACUACAACAGCCACAACGUCUAUAUCAUGGCCGACAAGCAGA
AGAACGGCAUCAAGGUGAACUUAAGAUCGCGCCACAACAUCGAGGACGGCAGCGU
GCAGCUCGCGACCACUACCAGCAGAACACCCCCAUCGGCGACGGCCCCGUGCUGC
UGCCCGACAACCACUACCUGAGCACCCAGUCCGCCUGAGCAAAGACCCCAACGAG
AAGCGCGAUCACAUGGUCCUGCUGGAGUUCGUGACCGCGCGGGGAUCACUCUCG
GCAUGGACGAGCUGUACAAGUAAAAGCUUGAUCCAGACAUGAUAAAGAUACAUGA
UGAGUUUGGACAAACCACAACUAGAAUGCAGUGAAAAAAUAGCUUUUAUUUGUGAA
AUUUGUGAUGCUAUUGCUUUUAUUUGUAACCAUUAUAAGCUGCAAUAAACAAGUU
GUGGAAUUGUGAGCGCUCACAAUUCCACAGCGGCCGtaggtagtcacattgcactttct
gtaacttgactgggtgtgggatgtgggcctggggtgtgggttatgggtatatatgtgggatattctg
gtgggaatgtccattcactgtatgcctatcttttaataaaaagacggtagctaggttcgcgaagcagc
cacaagccaatagccagttaggtagctcatagtgggttaggtgacaggaacctttgactcagaacgcgt
ccattaacatctagaacggaccaaacttcggacatgcaccgattaaccggatttgtccaaggtggacg
ggccacctttacttaaccggaaagggaacatatatagttatatgtgttcgtaata TAGC AAAAA
AAAAAAAAAAAAAAAAAAAA

10

20

30

510. TCA5_TaGu3

gggagacgggucaaccgcguaggagcggugaccgGAAUUCGGCGGGAGUAACUAUGACUC
UCUUAAGGAGUCAUAGAGCCAGAACCUCUCGUGGUCCCGCUGGGCACAGGGAUU
AAUUUUUCUGUGGCAAAUUGACUGGCUUCAGAGAGCGUUUUUCGAAGUGGACU
GUGUGACUGCGUUCCCCCCUUAGUUGCUAUAUCCGCUUCGAUUAACAUCUCACCU
CGACGUAAUAGAUAUUUCUAGAAGGUCGACCAGAUGUCCGAGGUCGACCAGUUG
UCCGUGUGGAAUUGUGAGCGCUCACAAUUCCACACGUUACAUAACUUACGGUAAA
UGGCCCGCCUGGCUGACCGCCCAACGACCCCCGCCCAUUGACGUCAAUAGUAACG
CCAAUAGGGACUUUCCAUUGACGUCAAUGGGUGGAGUAUUUACGGUAAACUGCCC
ACUUGGCAGUACAUCAAGUGUAUCAUAUGCCAAGUACGCCCCCUAUUGACGUCAA
UGACGGUAAAUGGCCCGCCUGGCAUUGUGCCCAGUACAUGACCUUAUGGGACUUU
CCUACUUGGCAGUACAUCUACGUAAUAGUCAUCGCUAAUUAACCAUGGUCGAGGUGA

40

50

GCCCCACGUUCUGCUUCACUCUCCCCAUCUCCCCCCCCUCCCCACCCCCAAUUUU
GUAUUUAUUUAUUUUUUAAUUAUUUUGUGCAGCGAUGGGGGCGGGGGGGGGGGG
GGGGCGCGCGCCAGGCGGGGCGGGGCGGGGCGAGGGGCGGGGCGGGGCGAGGCGG
AGAGGUGCGGCGGCAGCCAAUCAGAGCGGCGCGCUCCGAAAGUUUCCUUUUUAUGG
CGAGGCGGCGGCGGCGGCGGCCUAUAAAAAGCGAAGCGCGCGGGCGGGCGGGAGU
CGCUGCGACGCUGCCUUCGCCCCGUGCCCCGCUCCGCCGCCGCGCCGCGCCGCGG
CCCCGGCUCUGACUGACCGCGUACUCCCCACAGGUGAGCGGGCGGGACGGCCCUU
CUCCUCCGGGCGUGAAUUAAGCUGAGCAAGAGGUAAGGGUUUAAGGGAUGGUUGGU
UGGUGGGGUUAUUAUGUUUAUUAACCUGGAGCACCUGCCUGAAAUCACUUUUUU
UCAGGUUGGCGUACGGCCACCAUGGUGAGCAAGGGGCGAGGAGCUGUUCACCGGGG
UGGUGCCCAUCCUGGUCGAGCUGGACGGCGACGUAAACGGCCACAAGUUCAGCGU
GUCCGGCGAGGGCGAGGGCGAUGCCACCUACGGCAAGCUGACCCUGAAGUUCAUC
UGCACCACCGGCAAGCUGCCCGUGCCCGUGGCCACCCUCGUGACCACCCUGACCUA
CGGCGUGCAGUGCUUCAGCCGCUACCCCGACCACAUGAAGCAGCACGACUUCUUC
AAGUCCGCCAUGCCCGAAGGCUACGUCCAGGAGCGCACCAUCUUCUUAAGGACG
ACGGCAACUACAAGACCCGCGCCGAGGUGAAGUUCGAGGGCGACACCCUGGUGAA
CCGCAUCGAGCUGAAGGGCAUCGACUUAAGGAGGACGGCAACAUCUGGGGGCAC
AAGCUGGAGUACAACUACAACAGCCACAACGUCUAUAUCAUGGCCGACAAGCAGA
AGAACGGCAUCAAGGUGAACUUAAGAUCGCCACAACAUCGAGGACGGCAGCGU
GCAGCUCGCCGACCACUACCAGCAGAACACCCCAUCGGCGACGGCCCCGUGCUGC
UGCCCGACAACCACUACCUGAGCACCCAGUCCGCCCGUGAGCAAAGACCCCAACGAG
AAGCGCGAUCACAUGGUCCUGCUGGAGUUCGUGACCGCCGCCGGGAUCACUCUCG
GCAUGGACGAGCUGUACAAGUAAAAGCUUGAUCCAGACAUGAUAAAGAUACAUGA
UGAGUUUGGACAAACCACAACUAGAAUGCAGUGAAAAAAAUAGCUUUUAUUUGUGAA
AUUUGUGAUGCUAUUGCUUUUAUUUGUAACCAUUAUAAGCUGCAAUAAACAAGUU
GUGGAAUUGUGAGCGCUCACAAUUCACAGCGGCCGC taatt caggttattt agatgc
ttag tttttgtacc tttcttgttt tgtttaggat tttgatagtg ttagtatttt tatattttg
tacgattgca taatgttctt ttttatacag ttctgtttta ataaaataga cgatagctag ag
acgttagg gcagccacaa gccagttagg tagcggatag taggtaggaa cagactttta ct
atttcata acgcgtcaat taccacctga tttggaccaa ttcacgggat ttgtccaagg tg
gacggggc acctttactt aacccggaaa aggaacatat ataatttatg tgtgttcgat aaa
TAGCAAAAAAAAAAAAAAAAAAAAAA

10

20

30

40

50

511. 配列番号503と同じ

512. TCA5_CBhBsi_mCherry_GeFo3_R4A22

gggagacgggucaaccgcgguaggagcggugaccgGAAUUCGGCGGGGAGUAACUAUGACUC
UCUUAAGGAGUCAUAGAGCCAGAACCUCCUCGUGGUCCCGCUGGGCACAGGGAAU
AAUUUUUCUGUGGCAAAUUGACUGGCUUCAGAGAGCGUUUUUCGAAGUGGACU
GUGUGACUGCGUUCCCCCCUAGUUGCUAUAUCCGCUUCGAUUAACAUCUACCU
CGACGUUAUAGAUAUUUCUAGAAGGUCGACCAGAUGUCCGAGGUCGACCAGUUG
UCCGUGUGGAAUUGUGAGCGCUCACAAUCCACACGUUACAUAACUUAACGGUAAA
UGGCCCGCCUGGCUGACCGCCCAACGACCCCCGCCCAUUGACGUCAAUAGUAAAG
CCAAUAGGGACUUUCCAUAUGACGUCAAUGGGUGGAGUAUUUACGGUAAACUGCCC
ACUUGGCAGUACAUAAGUGUAUCAUAUGCCAAGUACGCCCCCUAUUGACGUCAA
UGACGGUAAAUGGCCCGCCUGGCAUUGUGCCCAGUACAUGACCUUAUGGGACUUU
CCUACUUGGCAGUACAUCUACGUAAUAGUCAUCGCUAAUACCAUGGUCGAGGUGA
GCCCCACGUUCUGCUUCACUCUCCCCAUCUCCCCCCCCUCCCCACCCCCAAUUUU
GUAUUUAUUUAUUUUUUAAUUAUUUUGUGCAGCGAUGGGGGCGGGGGGGGGGGG
GGGGCGCGCGCCAGGCGGGGCGGGGCGGGGCGAGGGGCGGGGCGGGGCGAGGCGG
AGAGGUGCGGCGGCAGCCAAUCAGAGCGGCGCGCUCCGAAAGUUUCCUUUUUAUGG
CGAGGCGGCGGCGGCGGCGGCCUAUAAAAAGCGAAGCGCGCGGGCGGGCGGGAGU

CGCUGCGACGCUGCCUUCGCCCCGUGCCCCGCUCGCGCCGCCUCGCGCCGCCCG
CCCCGGCUCUGACUGACCGCGUUAUCCACAGGUGAGCGGGCGGGACGGCCCUU
CUCCUCCGGGCUGUAAUAGCUGAGCAAGAGGUAAGGGUUAAGGGAUGGUUGGU
UGGUGGGGUUAUAAUGUUUAAUUAACCUGGAGCACCUGCCUGAAAUCACUUUUU
UCAGGUUGGCGUACGGCCACCauggugagcaagggcgaggaggauaacauggccauc
aaggaguucaugcgcuucaaggugcacauggagggcuccgugaacggccacgaguucgagau
gagggcgaggcgaggcgccgccccuacgagggcacccagaccgccaagcugaaggugaccaagg
guggccccugcccuucgcuugggacaucuguccccucaguucauguacggcuccaaggccua
cgugaagcaccccgccgacaucccgacuacuugaagcuguccuuccccgaggggcuucaagugg
gagcgcgugaugaacuucgaggacggcgggcguggugaccgugacccaggacuccuuccugcagg
acggcgaguucaucuacaaggugaagcugcgcgccaccaacuuccccuccgacggccccguaau
gcagaagaagaccaugggcuugggagggccuccuccgagcggauguaccccgaggacggcgcccug
aaggggcgagaucaagcagaggcgugaagcugaaggacggcgccacuacgacgcugaggucaaga
ccaccuacaaggccaagaagcccgugcagcugcccggcgccuacaacgucaacaucaaguuggac
aucaccucccacaacgaggacuacaccaucguggaacaguacgaacgcgcccagggcgccacuc
caccggcgggcauggacgagcuguacaaguagAAGCUUGAUCCAGACAUGAUAAAGAUACA
UUGAUGAGUUUGGACAAACCACAACUAGAAUGCAGUGAAAAAAUAGCUUUAUUUG
UGAAAUUUGUGAUGCUAUUGCUUUAUUUGUAACCAUUAUAAGCUGCAAUAAACA
AGUUGUGGAAUUGUGAGCGCUCACAAUCCACAGCGGCCGCUGAGGUAGAUAAUC
UUUGUAUAGUGGGGGGGGAUCUCAUGUACCGGGUUCUUAUUAUUUGAUUUUCAA
UAAAACAGACGGUAGCUAGGUUCGCAAGGCAGCCACAAGCCAAAGAUAGGUAGGG
UGCUCAUAGUGAGUAGGGACAGUGCCUUUGAUUCACAACGCGUCAAUACCAUCU
GACACGGAUACCCUUAACCGGACUUGUCAUGAUCUCCAGACUUGUCCAAGGUGGA
CGGGCCACCUUUAACUUAACCCGGAAAAGGAACAUAUAUUAUUAUUAUGUGUUCGG
AAAAUAGCAAAAAAAAAAAAAAAAAAAAAAA

10

20

30

40

50

513. CBh-GFP

gggagacggucaaccgcuaggagcgugaccgGAAUUCGGCGGGAGUAACUAUGACUC
UCUUAAGGAGUCAUAGAGCCAGAACCUCUCUGUGGUCCCGCUGGGGCACAGGGAUU
AAUUUUUCUGUGGCAAAUUGACUGGCUUCAGAGAGCGUUUUUCGAAGUGGACU
GUGUGACUGCGUUCUUUUUAGUUGCUAUUCCGCUUCGAUUAACAUCUCACCU
CGACGUUAUAGAUCAUUUCUAGAAGGUCGACCAGAUGUCCGAGGUCGACCAGUUG
UCCGUGUGGAAUUGUGAGCGCUCACAAUCCACACGUUACAUAACUUAACGGUAAA
UGGCCCGCCUGGGCUGACCGCCCAACGACCCCGCCCAUUGACGUCAAUAGUAACG
CCAAUAGGGACUUUCCAUUGACGUCAAUGGGUGGAGUAUUUACGGUAAACUGCCC
ACUUGGCAGUACAUAAGUGUAUCAUAUGCCAAGUACGCCCCCUAUUGACGUCAA
UGACGGUAAAUGGCCCGCCUGGCAUUGUGCCCAUUAUGACCUUAUGGGACUUU
CCUACUUGGCAGUACAUCUACGUUUAGUCAUCGCUAUUACCAUGGUCGAGGUGA
GCCCCACGUUCUGCUUCACUCUCCCCAUCUCCCCCCCCUCCCCACCCCCAAUUUU
GUAUUUAUUUAUUUUUUAUUUAUUUUGUGCAGCGAUUGGGGGCGGGGGGGGGGGG
GGGGCGCGCGCCAGGCGGGGCGGGGCGGGGCGAGGGGCGGGGCGGGGCGAGGCGG
AGAGGUGCGGCGGCAGCCAAUCAGAGCGGCGCGCUCCGAAAGUUUCCUUUUUAGG
CGAGGCGGCGGCGGCGGCGGCCCUAUAAAAAGCGAAGCGCGCGGGCGGGGAGU
CGCUGCGACGCUGCCUUCGCCCCGUGCCCCGCUCGCGCGCGCCUCGCGCGCGCCG
CCCCGGCUCUGACUGACCGCGUUAUCCACAGGUGAGCGGGCGGGACGGCCCUU
CUCCUCCGGGCUGUAAUAGCUGAGCAAGAGGUAAGGGUUAAGGGAUGGUUGGU
UGGUGGGGUUAUAAUGUUUAAUUAACCUGGAGCACCUGCCUGAAAUCACUUUUU
UCAGGUUGGCGUACGGCCACCAUGGUGAGCAAGGGCGAGGAGCUGUUCACCGGGG
UGGUGCCCAUCCUGGUCGAGCUGGACGGCGACGUAAACGGCCACAAGUUCAGCGU
GUCCGGCGAGGGCGAGGGCGAUGCCACCUACGGCAAGCUGACCCUGAAGUUCAUC
UGCACCACCGGCAAGCUGCCCGUGCCCGGGCCACCCUCGUGACCACCCUGACCUA

10

20

30

40

50

10

20

30

40

30

40

50

CUCGUCCAUGCCGAGAGUGAUCCCGGCGGGCGGUCACGAACUCCAGCAGGACCAUG
UGAUCGCGCUUCUCGUUGGGGUCUUGGCUCAGGGCGGACUGGGUGUCUCAGGUAGU
GGUUGUCGGGCAGCAGCACGGGGCCGUCGCCGAUGGGGGUGUUCUGCUGGUAGUG
GUCGGCGAGCUGCACGCGUGCCGUCCUCGAUGUUGUGGGCGGAUCUUGAAGUUCACC
UUGAUGCCGUUCUUCUGCUUGUCGGCCAUGAUUAUAGACGUUGUGGGCUGUUGUAG
UUGUACUCCAGCUUGUGCCCCAGGAUGUUGCCGUCCUCCUUGAAGUCGAUGCCCU
UCAGCUCGAUGCGGUUACACAGGGUGUCGCCUCGAACUUCACCUCGGCGCGGGU
CUUGUAGUUGCCGUCGUCCUUGAAGAAGAUGGUGCGCUCCUGGACGUAGCCUUCG
GGCAUGGGCGGACUUGAAGAAGUCGUGCUGCUUCAUGUGGUCGGGGUAGCGGCUGA
AGCACUGCACGCCGUAGGUCAGGGUGGUCACGAGGGUGGGCCAGGGGCACGGGCAG
CUUGCCGGUGGUGCAGAUGAACUUCAGGGUCAGCUUGCCGUAGGUGGCAUCGCCC
UCGCCCUUCGCCGGACACGCUGAACUUGUGGGCCGUUACGUCGCCGUCCAGCUCGA
CCAGGAUGGGCACCAACCCCGGUGAACAGCUCCUCGCCCUUGCUCACCAUGGUGGC
CGUACGCCUGUGUUCUGGGCGGCAAACCCGUUGCGAAAAAGAACGUUCACGGCGAC
UACUGCACUUAUAUACGGUUCUCCCCACCCUCGGGAAAAAGGCGGAGCCAGUAC
ACGACAUCACUUCUCCAGUUAACCCCGCGCCACCUUCUCUAGGCACCGGUUCAAU
UGCCGACCCCUCCCCCAACUUCUCGGGGACUGUGGGCGAUGUGCGCUCUGCCCC
GGACAACUGGUCGACCUCGGACAUCUGGUCGACCUCUAGAUGAGGUAGAUAAUC
UUUGUAUAGUGGGGGGGGAUCUCAUGUACCGGGUUCUUAUUAUUGAUUUUCAA
UAAAACAGACGGUAGCUAGGUUCGCAAGGCAGCCACAAGCCAAAGAUAGGUAGGG
UGCUCAUAGUGAGUAGGGACAGUGCCUUAUGAUUACAAACGCGUCAAUACCAUCU
GACACGGAUACCCUUAACGGACUUGUCAUGAUCUCCAGACUUGUCCAAGGUGGA
CGGGCCACCUUUAACUUAACCCGGAAAAAGGAACAUAUAUUAUUAUUAUGUGUUCGG
AAAAUAGCAAAAAAAAAAAAAAAAAAAAAA

10

20

30

40

50

518. TCA5_CBhBsi_hTERT_GeFo3 RNA 配列

gggagacgggucaaccgcguaaggagcggugaccgGAAUUCGGCGGGAGUAACUAUGACUC
UCUUAAGGAGUCAUAGAGCCAGAACCUCGUGGUCCCGCUGGGCACAGGGGAUU
AAUUUUUCUGUGGCAAAUUGACUGGCUUCAGAGAGCGUUUUUCGAAGUGGACU
GUGUGACUGCGUUCCCCCCUUAGUUGCUAUAUCCGCUUCGAUUAACAUCUCACCU
CGACGUAAUAGAUAUUUCUAGAAGGUCGACCAGAUGUCCGAGGUCGACCAGUUG
UCCGUGUGGAAUUGUGAGCGCUCACAAUCCACACGUUACAUAACUUAACGGUAAA
UGGCCC GCCUGGCUGACCGCCCAACGACCCCCGCCCAUUGACGUCAAUAGUAACG
CCAAUAGGGACUUUCCAUUGACGUCAAUGGGUGGAGUAUUUACGGUAAACUGCCC
ACUUGGCAGUACAUAAGUGUAUCAUAUGCCAAGUACGCCCCCUAUUGACGUCAA
UGACGGUAAAUGGCCCGCCUGGCAUUGUGCCAGUACAUGACCUUAUGGGACUUU
CCUACUUGGCAGUACAUCUACGUAAUAGUCAUCGCUAAUACCAUGGUCGAGGUGA
GCCCCACGUUCUGCUUCACUCUCCCCAUCUCCCCCCCCUCCCCACCCCCAAUUUU
GUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUU
GGGGCGCGCGCCAGGCGGGGCGGGGCGGGGCGAGGGGCGGGGCGGGGCGAGGGCGG
AGAGGUGCGGCGGCAGCCAAUCAGAGCGGCGCGCUCCGAAAGUUUCCUUUUUAUGG
CGAGGCGGCGGCGGCGGCGGCCCUAUA AAAAGCGAAGCGCGGCGGGCGGGAGU
CGCUGCGACGCGUCCUUCGCCCGUGCCCCGCUCCGCCGCCGCGCUGCGCGCCCG
CCCCGGCUCUGACUGACCGCGUACUCCACAGGUGAGCGGGCGGGACGGCCCUU
CUCCUCCGGGCUGUAAUUAAGCUGAGCAAGAGGUAAAGGGUUUAAGGGAUGGUUGGU
UGGUGGGGUAAUUAUGUUUAUUAUACCUGGAGCACCUGCCUGAAAUACUUUUUU
UCAGGUUGGCGUACGgccaccauggugaugccucgagcaccgccgugcagagccgugcggg
gccugcugcgggucacacagagagguccuccccugggccacguucgugcggcgucugggccc
ucagggcugggcgacuggugcagcgcgggcgaccccgccgcauuccgcgcccugguaggccagugc
cuggugugugugccuugggaugccagaccuccuccugcgccuccuucuuuuagacaggugucu
ugucugaaggagcugguaggcccgggugcugcagcgggcugugcagcggggagcuaagaauug

uuagcuuucgguuuugcccugcuggacggcgccagagggcggaaccaccugaagccuucacuaccu
cggugagaagcuaccugccuaacaccgugacagacgcccugcggggcagcggcgccugggggccu
gcuacugagacgggugggagauaugugcuggugcaccugcuggccagaugcggccucuuugu
gcuggugggcccuagcugugccuaccaggugugugggccuccccuguaaccagcuggggcgccgc
acgcaggccagacccccuccgcacgccagcggcccuaggagacgucuggggcugcgaaagagcuu
ggaaccacagugugcgagaagccggcguccccugggccuccugcccccuggagccagaagaag
aggaggcuccgccagcagaagccucccgucgcccaggcgccuagacggggcgccgccccugaac
ccgagcgggaccccgguccggccaggcgaguuuggggccaccccggaaggaccagaggaccuagcga
ccggggcuucugcguggucuccccagccggccagccgaggaagcaaccagccugggagggcgcc
cugagcggcaccagacacucccaccuuccguuggcagacagcaccacgccgggcccaccaagcac
cagcagaccuccgcgaccuugggacacaccuguccuccuguguacgccgagaccaagcacuuc
cuguacuccucugggcgauaaggaaacagcuccggccuagcuuccugcugagcucucugagacca
gcuugaccggcgcccgacagacucguggaaacaauccuuccugggcagcagaccuuggaugccagg
caccuccuagacggcugccuagacugccccagagauacugggcagaugcguccacuguuuuggaa
cugcugggcaaccacgcucaguguccuuacggcgucugcugaaaaccacugcccccugagag
ccgcugucacccagccggcgugugcgccagagaaaagccucagggaagcguggcgucucc
ugaggaagaggauacagaccuagacgccugggugcaacugcugcgccaacauagcagcccuugg
cagguguacggcuucgucagagcuugucugagaaggcuggugccccgggacuguggggcucc
agacacaacgagcggagauuccugcggaacacaaagaaguucaucagccugggcaagcacgcca
agcugucucugcaggaguugaccuggaagaugagcgugagggaauugcgugggcugagaagau
ccccuggagugggcgugugccagccggcgagcacagacugcgagaggaaauccugggccaaa
ccugcacuggcgugaugagcguauaugugguggaacugcugcgaggcuuccuucacgugaccga
gacaacauuucagaagaacagacuuccuucacccggaagucuguauggagcaagcugcagucc
aucggaaucagacagcaccugaagagaguucagcugcggggaacugagcgaggcagaggugaggc
agcauagagaggcuagaccugcucugcuuaccagcagacugagauucaucccaaaaccugaugg
ccugagaccuauugugaauauggacuacguggugggcgcccgaaccuuccggcgugagaagcgg
gcugaaagacugacaucgggugaaggcccuugucagcgugcugaacucagagagcucguc
ggccuggggcuacuggggcgccagcgugcuggggccuggaugacauccacagagccuggagaacau
cgugcugagagugagggcucaagaucaccccgagcuguaucuuuguaagguggacgugac
aggcgcauacgacaccauccucaggacagacugaccgaggugauugccagcauuaucaagccc
cagaacaccuauugcgugagacgguaacgccguggugcagaaagccgcacauggccacgugcggga
aagccuucagucucacguguccacccugaccgaccugcagccuacaugcggcaauucguggc
ccaccuccaggagacaagccucugagagacgccguggugaucgagcaguccagcagccuua
gaggccucuaagcggccugucgacgugucuccugagauuuauuguccaccacgccgugagaau
agaggcaaaagcuacgugcagugccagggcaucccccaaggauucauccugagcaccugcugu
gcagccugugcuacggcgacauggaaaacaagcugucgcccggaaucgggagagacggccugcu
gcugcggcucguugacgacuuccugcuggugacaccucaccugacacacgcaaaaaccuuccuc
agaacacugguucggggcguccccgaguacggcugcgucgugaaccugagaaagacaguggug
aacuucccuguggaagaugaggcacucggcggaacagcgucgugcagaugcccggccacggcc
uguuccccugggugcgggccugcugcuggacaccagaaccucggaggugcagagcgauuacagcag
cuaugccagaaccagcauccgagccagccuaaccuuuaaccggggcuucaaaagccggcagaaau
augagaagaaagcuguucggagugcugaggcuuaaaugccacagccuguuuccuggaccugcag
gucaacagccugcaaacagugugcaccacaucuaacaagaauccugcuguuacaagccuacagau
uucacgccugugugcugcagcugccuuccaccagcagguguggaagaaccuacguuuuucc
ugagagugaucucugauacagccaguuuugugcuacagcauccugaaggccaagaacgccggcau
gagccugggagccaaggcgugcugggccucugccuucugaggccgugcagugggcugugcca
ccaggccuuucugcugaagcugaccagacaauagagugaccuacgugccucuccugggacgcu
agaaccggccagaccagcucucucggaagcugccugggcacaaccucgaccggccuggaagcugc
ugccaauccugcccugccuagcagacuucagaaccuuccuggacugauAAGCUUGAUCCAGAC
AUGAUAAAGAUACAUUGAUGAGUUUGGACAAACCACAACUAGAAUGCAGUGAAAAA

10

20

30

40

50

AAUGCUUUAAUUGUGAGAAUUGUGAUGCUAUUGCUUUAAUUGUAACCAUUAUAA
 GCUGCAAUAAACAAGUUGUGGAAUUGUGAGCGCUCACAAUCCACAGCGGCCGCU
 GAGGUAGAUAAUCUUGUAUAGUGGGGGGGGAUCUCAUGUACCGGGUUCUUUU
 AUUUGAUUUUCAUAAAACAGACGGUAGCUAGGUUCGCAAGGCAGCCACAAGCCA
 AAGAUAGGUAGGGUGUCUAUAGUGAGUAGGGACAGUGCCUUUUGAUUCACAACGC
 GUCAAUACCAUCUGACACGGAUACCCUUACCGGACUUGUCAUGAUCUCCAGACU
 UGUCCAAGGUGGACGGGCCACCUUUAACUUAACCCGGAAAAGGAACAUAUAUAAU
 UAUAUGUGUUCGGAAAAUAGCAAAAAAAAAAAAAAAAAAAAAA

519. TCA5_CBhBsi_hTERT_EFS_TdT_GeFo3 RNA配列

gggagacgggucaaccgcguaggagcggugaccgGAAUUCGGCGGGAGUAACUAUGACUC
 UCUUAAGGAGUCAUAGAGCCAGAACCUCUCGUGGUCCCGCUGGGGCACAGGGAAU
 AAUUUUUCUGUGGC AAAUUGACUGGCUUCAGAGAGCGUUUUUCGAAGUGGACU
 GUGUGACUGCGUUC CCCCCCUAGUUGCUAUAUCCGCUUCGAUUAACAUCUCACCU
 CGACGUAAUAGAUAUUUCUAGAAGGUCGACCAGAUGUCCGAGGUCGACCAGUUG
 UCCGUGUGGAAUUGUGAGCGCUCACAAUUCACACGUUACAUAACUACGGUAAA
 UGGCCCGCCUGGCUGACCGCCCAACGACCCCCGCCCAUUGACGUCAAUAGUAACG
 CCAAUAGGGACUUUCCAUUGACGUCAAUGGGUGGAGUAUUUACGGUAAACUGCCC
 ACUUGGCAGUACAUAAGUGUAUCAUAUGCCAAGUACGCCCCCUAUUGACGUCAA
 UGACGGUAAAUGGCCCGCCUGGCAUUGUGCCCAGUACAUGACCUUAUGGGACUUU
 CCUACUUGGCAGUACAUCUACGUAAUAGUCAUCGCUAAUACCAUGGUCGAGGUGA
 GCCCCACGUUCUGCUUCACUCUCCCCAUCUCCCCCCCCUCCCCACCCCCAAUUUU
 GUAAUUAAUUAAUUUUUAAUUAUUUUGUGCAGCGAUUGGGGGCGGGGGGGGGGGG
 GGGGCGCGCGCCAGGCGGGGCGGGGCGGGGCGAGGGGCGGGGCGGGGCGAGGCGG
 AGAGGUGCGGCGGCAGCCAAUCAGAGCGGCGCGCUCCGAAAGUUUCCUUUUAUGG
 CGAGGCGGCGGCGGCGGCGGCCUAAUAAAAGCGAAGCGCGGCGGGCGGGAGU
 CGCUGCGACGCGUCCUUCGCCCGUGCCCCGCUCCGCGCGCCUCGCGCCGCCCG
 CCCC GGCUUCUGACUGACCGCGUACUCCCACAGGUGAGCGGGCGGGACGGCCCUU
 CUCCUCCGGGCGUGAAUUAAGCUGAGCAAGAGGUAAAGGGUUUAAGGGAUGGUUGGU
 UGGUGGGGUAAUAAUGUUUAAUUAACCUUGGAGCACCUUGCCUGAAAUCACUUUUU
 UCAGGUUGGCGUACGgcccaccauggugaugccucgagcaccgccggugcagagccgugcggg
 gccugcugcgggucacacacagagagguccucccccuggccacguucgugcggcgucuggggcc
 ucagggcugggcgacuggugcagcgcgcgaccccgccgcauuccgcgcccuggugggccagugc
 cuggugugugugccuugggaugccagaccuccuccugcggcuccuucuuuagacaggugucu
 ugucugaaggagcuggugggccgggugcugcagcgugcugcagcggggagcuaagaugug
 uuagcuuucgguuuugcccugcugggacggcgccagaggcggaaccaccugaagccuucacuccu
 cggugagaagcuaccugccuaacaccgugacagacgcccugcggggcagcgggcgccuggggccu
 gcuacugagacggguggggagauaugugcuggugcaccugcuggccagaugcgccucuuugu
 gcuggugggcccuagcugugccuaccaggugugugggccucccccuguaccagcuggggcgccgc
 acgcaggccagaccccccuccgcacgcccagcgggccuaggagacgucugggugcgaaagagcuu
 ggaaccacagugugcgagaagccggcgucccccugggcuuccugccccuggagccagaagaag
 aggaggcuuccgcccagcagaagccucccgugcccaagcgggccuagacggggcgccgcccugaac
 ccgagcgggaccccgugcgccaggcgaguuugggcccaccccggaaggaccagaggaccuagcga
 ccggggcuucugcugggucuccccagcccggccagccgaggaagcaaccagccuggagggcgcc
 cugagcggcaccagacacucccaccuuccguuggcagacagcaccacgcccggcccaccaagcac
 cagcagaccuccgcgaccuugggacacaccuuccuccuguguacgccgagaccaagcacuuc
 cuguacuccucuggcgauaagggaacagcuccggccuagcuuccugcugagcucucugagacca
 gcuugaccggcgcccgacacugggaaacaauccuuccugggacagaccuuggaugccagg
 caccuccuagacggcugccuagacugccccagagauacuggcagaugcguccacuguuuuuggaa
 cugcuggggcaaccacgcucaguguccuuacggcguucugcugaaaaccacugccccugagag
 ccgcugucacccagccgcccggcgugugcgccagagaaaagccucagggaagcguggcugcucc

10

20

30

40

50

ugaggaagaggauacagacccuagacgccugggugcaacugcugcgccaacauagcagcccuugg
cagguguacggcuucgucagagcuugucugagaaggcugggugccccgggacuguggggcuucc
agacacaacgagcggagauuccugcggaaacacaaagaaguucaucagccugggcaagcacgcca
agcugucucugcaggaguugaccuggaagaugagcgugagggaauugcgugggcugagaagaau
ccccuggagugggcugugugccagccgcccagcacagacugcgagaggaaauccugggccaaaau
ccugcacugggcugaugagcguaauaugugguggaacugcugcgaggcuucuuucacugagaccga
gacaacauuucagaagaacagacuucuuucuaaccggaagucuguauggagcaagcugcagucc
aucggaauacagacagcaccugaagagaguucagcugcgggaaacugagcgaggcagaggugaggc
agcauagagaggcuagaccugcucugcuuaccagcagacugagaauucaucccaaaaaccugaugg
ccugagaccuaucgugaauauggacuacguggugggcgcccgaaccuuccggcgugagaagcgg
gcugaaagacugacaucgggugaaggccuguuacgugcgugaacuacgagagagcucguc
ggccuggggcuacugggcgccagcgugcuggggccuggaugacauccacagagccuggagaacaau
cgugcugagagugaggggcucaagaauccccaccggagcuguaucuuuguaaggugggacgugac
aggcgcauacgacaccauuccucaggacagacugaccgaggugauugccagcauuaucaagccc
cagaacaccuaauugcgugagacgguaacggcguggugcagaaagccgcacauggccacgugcgga
aagccuucagucucacguguccaccugaccgaccugcagccuuacaugcgggcaauucguggc
ccaccuccaggagacaagcccucugagagacggcguggugaucgagcaguccagcagccuuau
gaggccucuaagcggccuguuacguguuuccugagauuuauugugccaccacggcgugagaau
agaggcaaaagcuacgugcagugccaggggcauuccccaaggauucuauccugagcaccugcugu
gcagccugugcuacggcgacauggaaaacaagcuguuucggcggaauccggagagacggccugcu
gcugcgggcucguugacgacuuccugcuggugacaccucaccugacacacgccaauuccuuc
agaacacugguucggggcguccccgaguacggcugcgucgugaaccugagaaagacaguggug
aacuucccuguggaagaugaggcacucggcggaacagcguucgugcagaugcccggccacggcc
uguuucccugggugcgggccugcugcuggacaccagaaccucggaggugcagagcgauuacagcag
cuauugccagaaccagcauuccgagccagccuaaccuuuaaccggggcuucaaaagccggcgagaa
augagaagaaagcuguucggagugcugaggcuuaauugccacagccuguuuccugaccugcag
gucaacagccugcaaacagugugcaccacaucuaacaagaauccugcuguuacaagccuacagau
uucacggccugugugcugcagcugccuuuccaccagcagguguggaagaaccuacguuuuucc
ugagagugaucucugauacagccaguuuugugcuacagcauuccugaaggccaaagaacggcgga
gagccugggagccaaggggcgugcugggccucugccuucugaggccgugcagugggcugugcca
ccaggccuuucugcugaagcugaccagacauagagugaccuacgugccucuccugggcagccug
agaaccggccagaccagcucucucggaaagcugccuggcacaaccugaccggccuggaagcugc
ugccaauccugcccugccuagcgacuucagaaccuuccugggacugauAAGCUUGAUCCAGAC
AUGAUAAAGAUACAUAUGAUGAGUUUGGACAAACCACAACUAGAAUGCAGUGAAAAA
AAUGCUIUAUUUGUGAAAAUUUGUGAUGCUAUUGCUUUAUUUGUAACCAUUAUAA
GCUGCAAUAACAAGUUGUGGAAUUGUGAGGCGUCACAAUCCACAGCGGCGCG
ggcagagcgacauccgcccacaguccccgagaaguuggggggagggggucggcaauugaaccggg
gccuagagaaggugggcgcgggguaaaacuggggaaagugaugucguguacugggcuccgcuuuu
cccgaggggugggggagaaccguauauaagugcaguagucggcgugaacguucuuuuucgcaac
ggguuugccgcccagaacacaggcguaacggccaccaugguuguccaaaggcgagggaagugaucag
gaauuuauagagaaucaaggucagaauagggaaggauucaagaacggccacgaguucgagau
ggagaaggagaaggcagaccuacgagggcacccaaacgcuaauuaaaagugacaaaggggcg
gaccucugccguucggcugggauauccuguccccccaguucauguacggcagcaaggccuacgu
gaagcaccucugcugacaucgggaauuacaaaagcugucuuuuuccugaggggcuuuuaguggga
gagagugaugaacuucgaggacggaggccugggugaccguuacccaagacagcagccugcaagau
ggaaacacugaucuacaaagugaagaugcggggcacuaauuuuccuccagacggcccccugaugc
agaagaaaaaccuaggcgugggagggccucuaacagaacggcuguaccccagagacggagugcuga
ggggcgagaauccagggcacugaagcugaagaaggcgggacacuaccuggguggaguucaagacc
auuuuacauggccaagaagccugugcagcugccuggcuacuacuacguggacaccaagcuggaca
uaccuccccacaauagaggauuauaccaucguggaacaaguacgagagaagcgaggggcagacacca

10

20

30

40

50

20

50

ACGCAAGCAGUUCGCCGACAUCGCCUAUAACUACCGGCAUGGCCAGCCUAUCCCU
AGAGUGGAAUACAUGGAAGAAGAGAAAAAACCGGGGCACCGUUUUAAGACCC
UGAAAAGCCUGUAUAAGACUCACGCCUGCUACGAGUACAACCACAUCUUCCCUCU
GCUGGAAAAGUACUGUGGCUUUCACGAGGACAACAUCCACAGCUGGAGGACGUG
UCCCAGUUCUGCAGACCUGUACCGGCUUCCGGCUGCGGGCCCGUGGCCGGACUGC
UGAGCAGCCGGGACUUCUGGGCGGACUGGCUUUUAGAGUGUUCCACUGCACACA
GUACAUCAGACACGGCUCCAAGCCUAUGUACACACCUGAGCCCGACAUCUGCCAC
GAGCUGCUGGGACACGUCCCUCUGUUUAGCGAUAGAUCUUUCGCCCAGUUUAGCC
AGGAGAUCCGGCCUGGCCAGCUUGGGCGCCCCUGAUGAAUACAUCGAGAAACUGGC
UACAAUCUACUGGUUCACAGUGGAGUUCGGCCUCUGCAAGCAGGGCGACAGCAUC
AAGGCCUACGGCGCCGGACUGCUUUCUAGCUUCGGCGAGCUGCAAUACUGUCUGU
CCGAGAAGCCCAAGCUGCUGCCUCUGGAACUGGAAAAGACCGCCAUCCAGAAUUA
CACCGUGACCGAGUUCAGCCCCUGUACUACGUGGCUGAGAGUUUCAACGACGCU
AAGGAAAAGGUGCGGAACUUCGCCGCCACAUAUCCCAAGACCUUUCAGCGUGCGGU
ACGACCCCUACACCCAAAGAAUCGAGGUGCUGGACAACACCCAGCAGCUGAAGAU
CCUGGCAGAUAGCAUACAACAGCGAGAUCCGGAUCCUGUGCAGCGCCCUGCAGAAG
AUUAAGUGAUAAAGCUUGAUCCAGACAUGAUAAAGAUACAUUGAUGAGUUUGGACAA
ACCACAACUAGAAUGCAGUGAAAAAAAUAGCUUUUAUUUGUGAAAUUUUGUGAUGC
AUUGCUUUUAUUUGUAACCAUUAUAAGCUGCAAUAAACAAGUUGUGGAAUUGUGA
GCGCUCACAAUUCACAGCGGCCGCGUGAGGUAGAUAAUCUUUGUAUAGUGGGGGG
GGAUCUCAUGUACCGGGUUCUUUUAUUUGAUUUUCAAUAAAACAGACGGUAGC
UAGGUUCGCAAGGCAGCCACAAGCCAAAGAUAGGUAGGGUGCUCAUAGUGAGUAG
GGACAGUGCCUUUUGAUUCACAACGCGUCAAUACCAUCUGACACGGAUACCCUUA
CCGGACUUGUCAUGAUCUCCAGACUUGUCCAAGGUGGACGGGGCCACCUUUACUU
AACCCGGAAGGAACAUAUAUUAUUAUUAUGUGUUCGGAAAAUAGCAAAAAAAA
AAAAAAAAAAAAAAAAA

10

20

521. TCA5_CBhBsi_PAH_EFS_TdT_GeFo3 RNA 配列

gggagacggucaaccgcgguaggagcggugaccgGAAUUCGGCGGGAGUAACUAUGACUC
UCUUAAGGAGUCAUAGAGCCAGAACCUCUGUGGUCCCGCUGGGGCACAGGGGAUU
AAUUUUUCUGUGGCAAAUUGACUGGCUUCAGAGAGCGUUUUUCGAAGUGGACU
GUGUGACUGCGUUCSCCCCUUAGUUGCUAUAUCCGCUUCGAUUAACAUCUCACCU
CGACGUUAAGAUAUUCUAGAAGGUCGACCAGAUGUCCGAGGUCGACCAGUUG
UCCGUGUGGAAUUGUGAGCGCUCACAAUCCACACGUUACAUAACUACGGUAAA
UGGCCCGCCUGGCUGACCGCCCAACGACCCCCGCCCAUUGACGUCAAUAGUAACG
CCAAUAGGGACUUUCCAUUGACGUCAAUGGGUGGAGUAUUUACGGUAAACUGCCC
ACUUGGCAGUACAUAAGUGUAUCAUAUGCCAAGUACGCCCCCUAUUGACGUCAA
UGACGGUAAAUGGCCCGCCUGGCAUUGUGCCCAGUACAUGACCUUAUGGGACUUU
CCUACUUGGCAGUACAUCUACGUAUUAGUCAUCGCUAUUACCAUGGUCGAGGUGA
GCCCCACGUUCUGCUUCACUCUCCCCAUCUCCCCCCCCUCCCCACCCCCAAUUUU
GUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUGUGCAGCGAUUGGGGGCGGGGGGGGGGG
GGGGCGCGCGCCAGGCGGGGGCGGGGGCGGGGGCGAGGGGGCGGGGGCGAGGGCGG
AGAGGUGCGGCGGCAGCCAAUCAGAGCGGCGCGCUCCGAAAGUUUCCUUUUUAUGG
CGAGGGCGGCGGCGGCGGCGGCCCUAUA AAAAGCGAAGCGCGGCGGGCGGGAGU
CGCUGCGACGCUGCCUUCGCCCGUGCCCCGCUCCGCCGCCGCCUCGCGCCGCCCG
CCCCGGCUCUGACUGACCGCGUUAUCCACAGGUGAGCGGGCGGGACGGCCCUU
CUCCUCCGGGCUGUAAUUAAGCUGAGCAAGAGGUAAGGGUUUAAGGGAUGGUUGGU
UGGUGGGGUUAUUAUGUUUAUUAUACCUGGAGCACCUGCCUGAAAUACUUUUUU
UCAGGUUGGCGUACGgccaaccAUGGUGAUGAGCACCGCCGUGCUGGAGAAUCCUGG
CCUGGGGCAGAAAGCUGUCCGAUUUCGGCCAAGAGACAAGCUACAUAUGAGGAUAAU
UGCAACCAGAACGGCGCUAUCAGCCUGAUCUUUUCUUUGAAAGAGGAAGUGGGCG

30

40

50

CUCUGGCCAAGGUCCUGAGGCUGUUCGAGGAAAACGACGUGAACCUGACCCAUAU
CGAAAGCAGACCCAGCCGGCUGAAGAAAGAUGAGUAUGAGUUCUUCACCCACCUG
GACAAGAGAUCUCUCCCUGCCCUGACAAACAUCAUAAAAUCCUAAGACACGACA
UUGGCGCCACCGUGCAGCAGCUGUCAAGAGAUAGAAGAAGGACACCGUGCCUUG
GUUUCUAGAACCAUCCAGGAGCUGGAUAGAUUCGCCAAUCAGAUCUGUCUUAAC
GGCGCCGAACUGGACGCCGACCACCCCGGCUUCAAGGACCCUGUGUACCGGGGCCAG
ACGCAAGCAGUUCGCCGACAUCGCCUUAUAACUACCGGCAUGGCCAGCCUAUCCCU
AGAGUGGAAUACAUGGAAGAAGAGAAAAAAACCUGGGGCACCGUUUUAAGACCC
UGAAAAGCCUGUAUAAGACUCACGCCUGCUACGAGUACAACCACAUCUUCUCCUCU
GCUGGAAAAGUACUGUGGCUUUCACGAGGACAACAUCCACAGCUGGAGGACGUG
UCCCAGUUCUGCAGACCUGUACCGGCCUUCGGCUGCGGCCCGUGGCCGGACUGC
UGAGCAGCCGGGACUUCUGGGCGGACUGGCUUUUAGAGUGUCCACUGCACACA
GUACAUCAGACACGGCUCCAAGCCUAUGUACACACCUGAGCCCGACAUCUGCCAC
GAGCUGCUGGGACACGUCCUCUGUUUAGCGAUAGAUCUUUCGCCCAGUUUAGCC
AGGAGAUUCGGCCUGGCCAGCUUGGGCGCCCCUGAUGAAUACAUCGAGAAACUGGC
UACAAUCUACUGGUUCACAGUGGAGUUCGGCCUCUGCAAGCAGGGCGACAGCAUC
AAGGCCUACGGCGCCGGACUGCUUUCUAGCUUCGGCGAGCUGCAAUACUGUCUGU
CCGAGAAGCCCAAGCUGCUGCCUCUGGAACUGGAAAAGACCGCCAUCCAGAAUUA
CACCGUGACCGAGUUCAGCCCCUGUACUACGUGGCUGAGAGUUUCAACGACGCU
AAGGAAAAGGUGCGGAACUUCGCCGCCACAUAUCCCAAGACCUUUCAGCGUGCGGU
ACGACCCCUACACCCAAAGAAUCGAGGUGCUGGACAACACCCAGCAGCUGAAGAU
CCUGGCAGAUAGCAUCAACAGCGAGAUCCGGAUCCUGUGCAGCGCCCUGCAGAAG
AUUAAGUGAUAAAGCUUGAUCCAGACAUGAUAAAGAUACAUUGAUGAGUUUGGACAA
ACCACAACUAGAAUGCAGUGAAAAAAAUGCUUUUAUUUGUGAAAAUUUGUGAUGCU
AUUGCUUUUAUUUGUAACCAUUAUAAGCUGCAAUAAACAAGUUGUGGAAUUGUGA
GCGCUCACAAUCCACAGCGGCCGCGgggcagagcgcacaucgcccacaguccccgagaag
uuggggggaggggucggcaauugaaccggugccuagagaagguggcgcgggguaaacugggaa
agugaugucguguacuggcuccgccuuuuucccgagggugggggagaaccguauauaagugca
guagucgccgugaacguucuuuuucgcaacggguuugccgccagaacacaggcguacggccacc
augguguccaaaggcgaggaagugaucaaggaaauuuaugagauucaaggucagaauagggaagga
ucuauagaacggccacgaguucgagauccgagggagaaggagaaggcagaccuacgagggcacc
aaaccgcuaaaauuaaaagugacaaaggcggaaccucugccguucgccugggauauccugucccc
ccaguucauguacggcagcaaggccuacgugaagcaccucgugacaucccccgaauacaaaaag
cugucuuuuuccugagggcuuuuagugggagagagugaugaacuuucgaggacggagggccuggug
accguuacccaagacagcagccugcaagauggaacaacugaucuacaaagugaagaugcggggca
cuauuuuccuccagacggccccgugaugcagaagaaaaccaugggcuugggagggccucuaacaga
acggcguaacccagagacggagugcugaaggcgagauuaccaggcacugaagcugaaagau
ggcggacacuaccuggguggaguucaagaccuuuacauuggccaagaaggccugugcagcugccug
gcuacuacuacguggacaccaagcugggacauaccuccacaaugaggauuauaccaucgugga
acaguacgagagaagcgagggcagacaccaccuguuucccgccacggcacaggcuccaccggc
ucuggcagcagcggcacagccagcagcaggacaacaacauggccgugaaucaaggaguucaugc
gguucaaggugcggauggaaggcagauugaacggccaugaguucgaaauccaggggcgaaggcg
aaggcagaccuuacgagggcacacagacagcuuagcugaaggugaccaaggggagggccucugcc
auucgccuggggacauccugagccucaguuuauuguauggcagcaaggccuacgugaagcacc
gccgacaucccccgaucacaagaagcugagcuucccugagggguuucaaaugggaaaggguuga
acuucgaggacggcggaucuggugacagugacccaggacagcagccugcaggauggcaccucgau
cuauaaggugaagaugagagggaccaacuuccaccugacggggccgugaugcagaaaaaaacc
augggaugggagggccagcaccgagcggcuguaccuagagauggcguucugaaggggcgagauc
caccaggcccugaagcugaaggacggcgccacuaccuggguggaguucaagaccuacuaugg
ccaagaaaccugugcagcugccuggauacuacuacguggacacaaagcugggacauaccucua

10

20

30

40

50

caacgaggacuauaccaucgucgagcaguacgaacgcagcgaaggcagacaccaucuguuuccug
uacgguauggacgagcuguacaaguaauggauuaaauaaaagaucuuuauguucauuag
aucuguguguugguuuuuguguguguggaauugugagcgcucacaaauuccacagcggccgcU
GAGGUAGAUAUAUCUUUGUAUAGUGGGGGGGGAUCUCAUGUACCGGGUUUCUUUU
AUUUGAUUUUCAUA AAAACAGACGGUAGCUAGGUUCGCAAGGCAGCCACAAGCCA
AAGAUAGGUAGGGUGCUCAUAGUGAGUAGGGACAGUGCCUUUUGAUUCACAACGC
GUCAAUACCAUCUGACACGGAUACCCUUAACCGGACUUGUCAUGAUCUCCAGACU
UGUCCAAGGUGGACGGGCCACCUUUA CUUAACCCGGAAAAGGAACAUAUAUUAU
UAUAUGUGUUCGGAAAAUAGCAAAAAAAAAAAAAAAAAAAAAA

522. TCA5_CBhBsi_F8N6_GeFo3 RNA 配列

gggagacggucaaccgcguaggagcggugaccgGAAUUCGGCGGGAGUAACUAUGACUC
UCUUAAGGAGUCAUAGAGCCAGAACCUCUCGUGGUCCCGCUGGGCACAGGGAUU
AAUUUUUCUGUGGCCAAAUUUGACUGGCUUCAGAGAGCGUUUUUCGAAGUGGACU
GUGUGACUGCGUUC CCCCCCUAGUUGCUAUAUCCGCUUCGAUUAACAUCUCACCU
CGACGUUAUAGAUAUUCUAGAAGGUCGACCAGAUGUCCGAGGUUCGACCAGUUG
UCCGUGUGGAAUUGUGAGCGCUCACAAUCCACACGUUACAUAACUUACGGUAAA
UGGCCCGCCUGGCUGACCGCCCAACGACCCCCGCCCAUUGACGUCAAUAGUAACG
CCAAUAGGGACUUUCCA UUGACGUCAAUGGGUGGAGUAUUUACGGUAAACUGCCC
ACUUGGCAGUACAUAAGUGUAUCAUAUGCCAAGUACGCCCCCUAUUGACGUCAA
UGACGGUAAAUGGCCCGCCUGGCAUUGUGCCCAGUACAUGACCUUAUGGGACUUU
CCUACUUGGCAGUACAUCUACGUAUUAGUCAUCGCUAUUACCAUGGUCGAGGUGA
GCCCCACGUUCUGCUUCACUCUCCCCAUCUCCCCCCCCUCCCCACCCCCAAUUUU
GUAUUUAUUUAUUUUUUAAUUAUUUUGUGCAGCGAUUGGGGGCGGGGGGGGGGGG
GGGGCGCGCGCCAGGCGGGGCGGGGCGGGGCGAGGGGCGGGGCGGGGCGAGGGCGG
AGAGGUGCGGCGGCAGCCAAUCAGAGCGGCGCGCUCCGAAAGUUUCCUUUAUGG
CGAGGCGGCGGCGGCGGCGGCCUUAUAAAAGCGAAGCGCGGCGGGCGGGAGU
CGCUGCGACGCUGCCUUCGCCCCGUGCCCCGCUCCGCCGCCGCCUCGCGCCGCCCG
CCCCGGCUCUGACUGACCGCGUUA CUUCCACAGGUGAGCGGGCGGGACGGCCCUU
CUCCUCCGGGCUGUAAUUAAGCUGAGCAAGAGGUAAAGGGUUUAAGGGAUGGUUGGU
UGGUGGGGUUAUAAUGUUUAUUAUACCUGGAGCACCUGCCUGAAAUACAUUUUUU
UCAGGUUGGCGUACGgccaccaugcaaaucgagcuguccacuugcuucuccugugccugc
ugcgcuuuuguuucagugcuaccgcagauacuaccugggagccguggagcuuucuccugggacu
auaugcagagcgaucugggcgagcugccagucgacgccagauuuccaccuagagugccuaagag
uuuccccuucacaccucugugguuacaagaaaaccuguuuguggaguucaccgaccaccug
uucaaauaucgcuagccucggccuccuuggaugggccugcugggccccaccauucaggcugaag
uguacgacaccguggugaucaccucugaagaacauggccuccaccucugucucucugcacgccgu
cggcgugucguacuggaaggccagcagggcgugaaucgacgaccagaccagccaaagagag
aaggaggauaauagguguucccagguggcagccacaccuacguguggcaggugcugaaagag
aacggcccaauggccucagaccucugugccugacauacagcuuuuaucccacguugaccugg
ugaaggaccugaacagcggacugaucggcgcccugcuggugugcagagaaggcucccuggccaa
ggaaaagacacagacacuacacaaguucaucuacuguuugccguguucgacgagggcaagagc
uggcacagcagaccagaauagccugaugcaagacagagaugccgccuccgcuagagcguaggc
cuaaaaugcacaccgugaacggcuacgucaacagauucucugcccggccugauuggcugucaccg
gaaguccguguacuggcacgugaucggaaugggcacaaccccgaggugcacucaauucuccug
gaaggccacacguuccugguagaaaccaccggcaggccucgcucgagauccccuauacau
uucugaccgccagaccuugcucauggaccugggccaguuuucugcuuuucugccacaucagcuc
ccaccagcacgacggcauggaggcuuacgugaagguggaauucuccugaggagccucagcu
gagaauagaagaauaauagaggaagccgaagaauacgacgacgaccugacggacagcgagauggau
guugugaggguugacgaugauaauagcccuuccuuuauccagauccggucaguggcgaagaaa
caccuuaagacaugggugcacuacaucgccgcugaggaggaggauugggacuacgccccccugg

10

20

30

40

50

ugcuggccccugacgaccggagcuacaaaucucaguaccugaacaacggcccucaaagaauccgg
cagaaaguacaagaaggugagauucauggcuuauacagaugaaacauuuaaaaccagagaagcu
auccagcaugagagcggaauccugggaccucugcuguacggcgagguggggcgacaccucugcuga
ucaucuuuaagaaccaggccagcagaccuuacaauaucuaccucacggcauuacagacgugag
accucuguaacagccggcgggcugcccaaggggcgugaagcaccugaaggacuucccaauccugccu
ggagagauucuacaaguacaaguggaccgugacaguggaagauggcccuaccaagucugaccccc
ggugccugacccgguaacuacucgaguuuugugaacauggagcgcgaccuggccucuggacuca
ucggcccauugcugaucugcuacaaagaguuccguggaccagagaggcaaccagaucaugagcgga
caagcguaacgugauccuguuacagcguguuugaugagaauagaagcugguauucugaccgagaa
cauccagcgguuccuccccaaccucgcccggcgugcagcuggaggaccucugaguuccaggccagc
aacaucaugcacucuauuaacggcuacguguucgacucccugcagcuaaagcguuugccugcacg
agguggcuuacugguacaauccugagcaucggcgcccagaccgacuuccuguccguguucuuu
caggcuacacauucaaaacauaagaugguguaugaggacaccucgaccucguuucccuuuuccgg
agagaccguguuauaugagcauggagaaaccucggccuuuggauccuggggcugucacaaauagcgga
cuuucggaaaccgggggaugaccgcccugcugaaagucagcucuugcgacaagaacaccggcgac
uacuacgaggacagcuauagaggacaucuccgccuaccugcugagcaagaacaacgccaucgagc
cucggucguucagccagaacucuaggcaccaccagcaccggcagaagcaauucaacgcaaccacc
aucccguaaaaacgacaucgaaaagaccgauccuugguucgcucacagaaccccaaugcccaaga
uccaaaacguguccucuucugaccugcugaugcugcugagacagagcccaacccucacggccu
gagccugucugaccugcaggaggccaaauacgagacuucucugacgauccuagccccggggcc
aucgacagcaacaacagccugagcgagauagacucauuuccggccucagcuccaccacucugggag
acauggguuuucacaccugaaaagcggccugcagcugagacugaaacgagaagcuaggcaccaccgc
ugccaccgaacugaagaagcuggacuuaagguaagcagcaccagcaacaaccugaucagcacc
aucccuucugacaaccuugccggccggcaccgacaacacaagcagccuggggccuccuucuaugcc
ugugcacuacgacagccagcuggauaccacacuguuuggaaagaaaucuaagcccuucugaccgaa
ucuggcgggccucugagccugagcgaagagaacaacgauagcaaacugcuggagagcgggccuga
ugaacagccaggagucaucuuggggcaagaacgugagcaguaccaggaggauaccccggaacaac
ccugcagagcgaucaggaggagaucauuauaugacacuaucagcguggagauagaagaaga
ggacuucgacaucuaacgacgaggacgaaaaccaaagccucggagcuuccagaagaaaaccaga
cacuauuucaucgccgcuguggagagacugugggacuacggcaugagcuccucuccccacgugc
ugcggaaaccggggccagagcggauucguuccucaguucaagaagucguguuccaagaguucac
cgacggcucuuuuaccagccucuguaauagaggcgaaauaaaacgaacaccuggggccugcugggc
ccuacaucaugagcagaaguugaggauaacaucauggugaccuucagaaaucaggccucucggc
ccuacucuuucuaucuccucgcucaucagcuacgaggaggaccagaggcaaggcgccgaaccag
aaagaacuucgugaagcccaacgaaaccaagacuuaucuuuggaaggugcaaacaccacauggcc
ccuaccaaagacgaguuuugauugcaaggccugggccuacuucagcgauguggaccuggaaaagg
auguccauagcggccugauuggcccccugcugguguccacacaaacacacucaauccugccca
cgggagacaggugaccgugcaggagauucggccuguuucacccaucuuucgacgagacaaagagc
ugguacuucacagaaaacauuggaaagaaauugcagagcccccugcaacauucagauggaggau
ccaccuucagggaauuuacagauuccacgccaucaacggauacaucauggacacacugccagg
acuggguauuggcccgagaucaagaauccggugguaccugcucucuauggggcucaaacgagaa
uauccacuuccauccacuuaagcggccacguguuacaccgugagaaaagaaagaggaguacaagaug
gcccuguaacaauucguaccccgggcguguucgagacaguggaaaugcugccaagcaaggccggga
ucuggcgggguggaauugucuuaauuggcgagcaccugcacggcggaugagcacacuguuucugg
uauacagcaacaagugccagacaccccgggcauggcuucugguacaucaugagauuuucagau
cacagccucugguacagacggacagugggcucccaagcuggccagacugcacuacuccggcucc
aucaacgccuggagcacaaaagagccuuucagcuggaucaaggucgaucugcuggcuccaauga
ucauccacggaaucagaagaccaggcgcccggcagaaaauucugagccucuacaucucgcaguu
cauuaucauguacagccuggacggaaagaaguggcagaccuacagagggaauaguaccggcacc
cugauggguguucucggcaauguggacagcucaggaauuaagcacacaucuucaauccuccaa

10

20

30

40

50

ucauagccagauacauacagacuucuccuacccacuacucuaucagaagcacgcugagaauugga
guggauggggaugugaccugaacagcugcagcaugccucuuggcauggaauucuaaagccaucuc
ugaugcccagauacaccgcuucuauguuacucaccaacauguuuccgcccacauggucccccuaagcaag
gccagacugcaccugcagggcgaggucuaaaugccugggcgccucaggugaacaaccccaaggaau
ggcugcaggucgauuuccagaaaaccaugaaaagugaccggagugaccacccaggggcgugaaaag
ccugcugaccagcauguacgugaaggaguuccugaucucuaagcucccaggaugggccaccagugg
accuguucuuucagaacggcaaagugaaaaguguuccaaggcaaccaggacagcuucaccccag
ucgugaacagccuggauccaccucuccuaccagauaccugagaauuccaucccccaguccugggg
ccaccagauccgcccugcggauggaggugcuggggcugcgaggcccaagaccuguacuAAGCUUG
AUCCAGACAUGAUAAAGAUACAUGAUGAGUUUGGACAAACCACAACUAGAAUGCA
GUGAAAAAAAUAGCUUUUUUUGUGAAAUUUUGUGAUGCUAUUUGCUUUUUUUUGUAAC
CAUUUAUAAGCUGCAAUAACAAGUUGUGGAAUUGUGAGCGCUCACAAUCCACAG
CGGCCGCUGAGGUAGAUAAUCUUUGUAUAGUGGGGGGGGAUCUCAUGUACCGGGU
UUCUUUUUAUUUGAUUUUCAAUAAAACAGACGGUAGCUAGGUUCGCAAGGCAGCCA
CAAGCCAAAGAUAGGUAGGGUGCUCAUAGUGAGUAGGGACAGUGCCUUUUGAUUC
ACAACGCGUCAAUACCAUCUGACACGGAUACCCUUACCGGACUUGUCAUGAUCUC
CCAGACUUGUCCAAGGUGGACGGGCCACCUUUACUUAACCCGGAAAAGGAACAUA
UAUUAAUUUAUUGUGUUCGGAAAAUAGCAAAAAAAAAAAAAAAAAAAAAA

10

523. TCA5_CBhBsi_F8N6_EFS_TdT_GeFo3 RNA配列

gggagacgggucaaccgcguaggagcgugaccgGAAUUCGGCGGGAGUAACUAUGACUC
UCUUAAGGAGUCAUAGAGCCAGAACCUCUCGUGGUCCCGCUGGGCACAGGGAUU
AAUUUUUCUGUGGCAAAUUUGACUGGCUUCAGAGAGCGUUUUUCGAAGUGGACU
GUGUGACUGCGUUCUUUUUUAGUUGCUAUUACCGCUUCGAUUAACAUCUCACCU
CGACGUUAUAGAUCAUUUCUAGAAGGUCGACCAGAUGUCCGAGGUCGACCAGUUG
UCCGUGUGGAAUUGUGAGCGCUCACAAUUCACACGUUACAUAACUUACGGUAAA
UGGCCCGCCUGGCUGACCGCCCAACGACCCCCGCCCAUUGACGUCAAUAGUAACG
CCAAUAGGGACUUUCCAUUGACGUCAAUGGGUGGAGUAUUUACGGUAAACUGCCC
ACUUGGCAGUACAUAAGUGUAUCAUAUGCCAAGUACGCCCCCUAUUGACGUCAA
UGACGGUAAAUGGCCCGCCUGGCAUUGUGCCCAGUACAUGACCUUAUGGGACUUU
CCUACUUGGCAGUACAUCUACGUUUUAGUCAUCGCUAUUACCAUGGUCGAGGUGA
GCCCCACGUUCUGCUUCACUCUCCCCAUCUCCCCCCCCUCCCCACCCCCAAUUUU
GUAUUUAUUUAUUUUUUAAUUUAUUUUGUGCAGCGAUUGGGGGCGGGGGGGGGGGG
GGGGCGCGCGCCAGGCGGGGCGGGGCGGGGCGAGGGGCGGGGCGGGGCGAGGCGG
AGAGGUGCGGCGGCAGCCAAUCAGAGCGGCGCGCUCCGAAAGUUUCCUUUUUUGG
CGAGGCGGCGGCGGCGGCGGCCUUAUAAAAGCGAAGCGCGCGGGCGGGGAGU
CGCUGCGACGCGUGCCUUCGCCCCGUGCCCCGCUCCGCGCGCCGCGCGCCCG
CCCCGGCUCUGACUGACCGCGUUACUCCACAGGUGAGCGGGCGGGACGGCCCUU
CUCCUCCGGGCGUGAAUUAGCUGAGCAAGAGGUAAAGGGUUUAAGGGAUGGUUGGU
UGGUGGGGUUAUUAAUGUUUAUUUACCUGGAGCACCUGCCUGAAAUCACUUUUU
UCAGGUUGGCGUACGgcccaccaugcaauucgagcuguccacuugcuucuccugugccugc
ugcgcuuuuguuucagugcuacccgcagauacuaccugggagccguggagcuuucuccugggacu
auaugcagagcgauccggcgagcugccagucgacgccagauuuccaccuagagugccuaagag
uuuccccuucacaccucugugguguaacaagaaaaccuguuuguggaguucaccgaccaccug
uucaauaucgcuuagccucggccuccuuggaugggccugcuggggccccaccauucaggcugaag
uguacgacaccguggugauccaccugaagaacauggccuccaccugucucucugcacgccgu
cggcgugucguacuggaaggccagcgagggcgcuuacgacgaccagaccagccaaagagag
aaggaggauauaagguguucccagguggcagccacaccuacguguggcaggugcugaaagag
aacggcccaaugggccucagaccucugugccugacauacagcuauuuauccacguugaccugg
ugaaggaccugaacagcggaugaucggcgcccugcuggugugcagagaaggcuuccugggcaa
ggaaaagacacagacacuacacaaguucaucuuacuguuugccguguucgacgagggcaagagc

20

30

40

50

uggcacagcgagaccaagaauagccugaugcaagacagagaugccgccuccgcuagagcguggc
cuaaaauugcacaccgugaacggcuacgucaacagaucucugcccggccugauuggcugucaccg
gaaguccguguacuggcacgugaucggaauggggcacaaccccggaggugcacucaauucucg
gaaggccacacguuccugguagaaaaccaccggcaggccucgcucgagaucuccccuauacau
uucugaccgcccagaccuugcucauggaccuggggccaguuuucugcuuuucugccacaucagcuc
ccaccagcacgacggcauggaggccuuacgugaagguggauucuuuguccugaggagccucagcu
gagaauagaagaauaauagaggaaagccgaagauuacgacgacgaccugacggacagcgagauggau
guugugaggguuugacgaugauaauagcccuuccuuuauccagauccgggucaguggcgaaagaa
caccuuaagacauugggugcacuacauccgcccugagggaggaggauugggacuacgccccccug
ugcuggccccugacgaccggagcuacaaaucucaguaccugaacaacggcccucaaagaaucgg
cagaaaguacaagaaggugagauucauggccuuauacagaugaaacauuuuaaaaccagagaagcu
auccagcaugagagcggaauccugggaccucugcuguaacggcgagguggggcgacaccucugcuga
ucaucuuuaagaaccaggccagcagaccuuacaauaucuaccucacggcauuacagacgugag
accucuguaacagccggcgccugcccgaaggcgugaagcaccugaaggacuucccaauccugccu
ggagagaucuucaaguacaaguggaccgugacaguggaagauggcccuaccaagucugaccccc
ggugccugacccgguaucucugaguuuugugaacauggagcgcgaccugggccucugggacuca
ucggcccauugcugaucugcuacaaagaguccguggaccagagaggcaaccagaucaugagcg
caagcguaacgugauccuguuacguguuugaugagaauagaagcugguauucugaccgagaa
cauccagcgguuccucccaacccugccggcgugcagcugggaggaccucugaguuccaggccagc
aacaucaugcacucuauuaacggcuacguguucgacuccugcagcuaaagcguuugccugcacg
agguggccuuacugguacauccugagcaucggcgcccagaccgacuuccuguccguguucuuu
caggcuacacauucaaaacauaagaugguguaugaggacaccucugaccucguuccccuuuuccgg
agagaccguguucaugagcauggagaaccucggccuuuggauccuggggcugucacaaauagcg
cuuucggaaaccgggggaugaccgcccugcugaaagucagcucuuugcgacaagaacaccggcgac
uacuacgaggacagcuauagaggacaucuccgcccuaaccugcugagcaagaacaacgccaucgagc
cucggugcguucagccagaacucuaggcaccaccagcaccggcgagaagcaauucaacgcaaccacc
aucccuagaaaacgacauccgaaaagaccgauccuugguucgucacagaaccccaaugcccaaga
uccaaaacguguccucuucugaccugcugaugcugcugagacagagcccaacccucacggccu
gagccugucugaccugcaggaggccaaaauacgagacuucucugacgauccuagccccggggcc
aucgacagcaacaacagccugagcgagaugacuauuuccggccucagcuccaccacucugggag
acauggguuuucacaccugaaagcgggccugcagcugagacugaacgagaagcuaggcaccaccgc
ugccaccgaacugaagaagcuggacuuaagguaagcagcaccagcaacaaccugauacgacacc
aucccuucugacaaccuugccggccggcaccgacaacacaagcagccuggggccuccuuaaugcc
ugugcacuacgacagccagcuggauaccacacuguuuggaaagaaaucuaagcccuugaccgaa
ucuggcgggccucugagccugagcgaagagaacaacgauagcaaacugcuggagagcgggccuga
ugaacagccaggagucacuugggggcaagaacgugagcaguaccaggaggauacccgggacaac
ccugcagagcgaucaggaggagaucgauuaugaugacacuacagcugggagauagaagaaaga
ggacuucgacaucuaacgacgaggacgaaaaccaaagccucggagcuuccagaagaaaaccaga
cacuauuucaucgcccugugggagagacugugggacuacggcaugagcuccucuccccacgugc
ugcggaaccggggccagagcggaucuguuuccuaguucaagaaagucguguuccaagaguucac
cgacggcucuuuuaaccagccucuguaauagaggcgaaauaaacgaacaccuggggccugcugggc
ccuuacauacagagcagaaguugaggauaacaucaugguagccuucagaaaucaggccucucggc
ccuacucuuuucuaucuccucgcucaucagcuacgaggaggaccagaggcaaggcgccgaaccag
aaagaacuucgugaagcccaacgaaaccaagacuauuuuggaaggugcaaacaccacauggcc
ccuaccaaagacgaguuuugauugcaaggccugggccuacuucagcgauguggaccuggaaaagg
auguccauagcggccugauuggcccccugcuggugugccacacaaacacacucaauccugccca
cgggagacaggugaccgugcaggaguucgcccuuguuucaccaucuuucgacgagacaaagagc
ugguacuucacagaaaacauggaaagaaauugcagagcccuugcaacaauucagauggaggau
ccaccuucaaggaaaaauacagauuccacgccaucaacggauacaucauggacacacugccagg
acugguuauggcccagggaucagagaauccggugguaccugcucucuauugggcucaaaacgagaa

10

20

30

40

50

uauccacuccauccacuuaagcggccacguguuacaccgugagaaaagaaagaggaguacaagaug
gcccuguaacaauucguaccccggcguguucgagacaguggaaaugcugccaagcaaggccggga
ucuggcggguggaauugucuuaauuggcgagcaccugcacgccggaauagagcacacuguuucugg
uauacagcaacaagugccagacaccccguggcauggcuucuggucacauacagagauuuucagau
cacagccucugguacaguacggacagugggcuuccaagcugggccagacugcacuacuccggcucc
aucaacgccugggagcacaaaagagccuucagcuggaucaaggucgaucugcuggcuccaauga
ucauccacggaaucaagacccaggcgccccggcagaaaauucucgagccucuacauucgcaguu
cauuaucauguacagccuggacggaaaagaaguggcagaccuacagaggaaaauaguaccggcacc
cugauggguguucuuucggcaauuguggacagcucaggaauuaagcacacaucuucaauccuccaa
ucauagccagauacauacagacuucacuccuacccacuacucuaucagaagcacgcugagaaugga
guggauggggaugugaccugaacagcugcagcaugccucuuggcauggaaucuaaagccaucuc
ugaugcccagauacaccgcuucuauguuacuuacccaacauguucgccacaugguccccuagcaag
gccagacugcaccugcagggcgaggucuaauugccuggcgggccucaggugaacaaccccaaggaa
ggcugcaggucgauuuuccagaaaaccaugaaagugaccggagugaccacccaggggcgugaaaag
ccugcugaccagcauguacgugaaggaguuccugaucucuagcucccagggaugggccaccagugg
accuguuucuuucagaacggcaaaagugaaaguguuuccaaggcaaccaggacagcuuaccccg
ucgugaacagccuggauccaccucuccuuaccagauaccugagaauccaaucccccaguccugggu
ccaccagaucgcccugcggauggaggugcugggcuugcgaggcccaagaccuguaucuaAGCUUG
AUCCAGACAUGAUAAAGAUACAUAUGAUGAGUUUGGACAAACCAACUAAGAAUGCA
GUGAAAAAAAUAGCUUUUAUUUGUGAAAAUUUGUGAUGCUAUUUGCUUUUAUUUGUAAC
CAUUUAUAAGCUGCAAUAACAAGUUGUGGAAUUGUGAGCGCUCACAAUCCACAG
CGGCCGCgggcagagcgcaaucgcccacaguccccgagaaguuggggggaggggucggcaau
ugaaccggugccuagagaaggugcgcgggguaaaacugggaaagugaugucguguacuggcuc
cgccuuuuucccgaggguggggggagaaccguauauaagugcagauagucgccgugaacguucuu
uuucgcaacggguuugcccgccagaacacaggcguaacggccaccaugguguccaaaggcgaggaa
gugaucuaaggaaauuaugagauucaaggucagaauggaaggauucaugaacggccacgaguuc
gagaucgaggggagaaggagaaggcagaccuacgagggcacccaaacgcuaauuaaaaguga
caaaggcggaaccucugccguucgcuugggauauccuguccccccaguucauguacggcagcaa
ggccuacgugaagcacccugcugacaucuccgaaucacaaaaagcugucuuuuccugaggggcuuu
aagugggagagagugaugaacuucgaggacggaggccugggugaccguuacccaagacagcagcc
ugcaagauggaacaucugaucuacaaagugaagaugcggggcacuaauuucccuccagacggccc
cgugaugcagaagaaaaccaugggcuugggagggccucuaacagaacggcuguaccccagagacgga
gugcugaaggggcgagauucaccagggcacugaagcugaaagauggcgggacacuaccugguggag
uucaagaccauuuacauuggccaagaagccugugcagcugccugggcuacuacuacugggacacca
agcugggacauacuccuccacaaugaggauuauaccaucugugaacaguacgagagaagcgaggg
cagacaccaccuguuuccucggccacggcacaggcuccaccggcucugggcagcagcgggcacagcca
gcagcgaggacaacaauuggccgugaucuaaggaguucaugcgguucaaggugcggauggaag
gcaguaugaacggccaugaguucgaaaucgagggcgaaaggcgaaaggcagaccuucgaggggcac
acagacagcuaagcugaaggugaccaaggggaggccucugccaauucgccugggacaucugagc
ccucaguuuauuguauuggcagcaaggccuacgugaagcaccgccgacauccccgacuacaaga
agcugagcuucccugaggguuucaaaugggaaaggguugaugaacuucgaggacggcggauggg
ugacagugaccacaggacagcagccugcaggauggcaccucgauucuaaaggugaagaugagagg
caccacuuccaccugacggggcccugugaugcagaaaaaaaccauggggaugggaggccagcacc
gagcgggcuguacccuagagauggcguuucugaaggggcgagauccaccaggcccugaagcugaagg
acggcgggccacuaccugguggaguucaagaccaucuaucauggccaagaaaccugugcagcugcc
uggauacuacuacguggacacaaagcuggacauacccucacaaacgaggacuauaccaucguc
gagcaguacgaacgcagcgaaggcagacaccaucuguuuccuguaacgguauggacgagcuguaca
aguuaugauuaauuaaaauaaaagaucuuuauguucauuagaucuguguguugguuuuugug
uguguggaaauugugagcgucacaaauuccacagcgggccgcUGAGGUAGAUAAUCUUUGUA
UAGUGGGGGGGGAUCUCAUGUACCGGGUUUCUUUAUUUGAUUUUCAAUAAAAC

10

20

30

40

50

AGACGGUAGCUAGGUUCGCAAGGCAGCCACAAGCCAAAGAUAGGUAGGGUGCUCA
UAGUGAGUAGGGACAGUGCCUUUUGAUUCACAACGCGUCAAUACCAUCUGACACG
GAUACCCUUAACGGACUUGUCAUGAUCUCCCAGACUUGUCCAAGGUGGACGGGCC
ACCUUUACUUAACCCGGAAAAGGAACAUAUAUUAUUAUUAUGUGUUCGGAAAAUA
GCAAAAAAAAAAAAAAAAAAAAAA

524. TCA5_CBhBsi_F8N6mutant_GeFo3 RNA配列

gggagacggucaaccgcguaggagcggugaccgGAAUUCGGCGGGAGUAACUAUGACUC
UCUUAAGGAGUCAUAGAGCCAGAACCUCGUGGUGCCCGCUGGGGCACAGGGGAUU
AAUUUUUCUGUGGGCAAUUUGACUGGCUUCAGAGAGCGUUUUUCGAAGUGGACU
GUGUGACUGCGUUCSCCCCUUAGUUGCUAUAUCCGCUUCGAUUAACAUCUCACCU
CGACGUAAUAGAUAUUUCUAGAAGGUCGACCAGAUGUCCGAGGUCGACCAGUUG
UCCGUGUGGAAUUGUGAGCGCUCACAAUCCACACGUUACAUAACUUAACGGUAAA
UGGCCCGCCUGGCGACCGCCCAACGACCCCGCCCAUUGACGUCAAUAGUAAACG
CCAAUAGGGACUUUCCAUUGACGUCAAUGGGUGGAGUAUUUACGGUAAACUGCCC
ACUUGGCAGUACAUAAGUGUAUCAUAUGCCAAGUACGCCCCCUAUUGACGUCAA
UGACGGUAAAUGGCCCGCCUGGCAUUGUGCCCAGUACAUGACCUUAUGGGACUUU
CCUACUUGGCAGUACAUCUACGUAAUAGUCAUCGCUAAUACCAUGGUCGAGGUGA
GCCCCACGUUCUGCUUCACUCUCCCCAUCUCCCCCCCCUCCCCACCCCCAAUUUU
GUAUUUAUUUAUUUUUUAAUUAUUUUGUGCAGCGAUUGGGGGCGGGGGGGGGGG
GGGGCGCGCGCCAGGCGGGGCGGGGCGGGGCGAGGGGCGGGGCGGGGCGAGGCGG
AGAGGUGCGGCGGCAGCCAAUCAGAGCGGCGCGCUCCGAAAGUUUCCUUUUAUGG
CGAGGCGGCGGCGGCGGCGGCCCUAUA AAAAGCGAAGCGCGGCGGGCGGGAGU
CGCUGCGACGCUGCCUUCGCCCCGUGCCCCGCUCCGCGCGCCUCGCGCCGCCCG
CCCCGGCUCUGACUGACCGCGUACUCCACAGGUGAGCGGGCGGGACGGCCCUU
CUCCUCCGGGCGUAAUUAAGCUGAGCAAGAGGUAAGGGUUUAAGGGAUGGUUGGU
UGGUGGGGUAAUUAUGUUUAUUAACCUGGAGCACCUGCCUGAAAUACUUUUUU
UCAGGUUGGCGUACGgccaaccAUGCAAUUCGAGCUGUCCACUUGCUCUCCUGUG
CCUGCUGCGCUUUUGUUUCAGUGCUACCCGCAGAUACUACCUGGGAGCCGUGGAG
CUUUCUUGGGACUAUAUGCAGAGCGAUUCUGGGCGAGCUGCCAGUCGACGCCAGAU
UUCCACCUAGAGUGCCUAAGAGUUUCCCCUUAACACCUCUGUGGUGUAACAAGAA
AACCUGUUUGUGGAGUUCACCGACCACCUUGUUAUAUUCGCUAAGCCUCGGCCU
CCUUGGAUGGGCCUGCUGGGCCCCACCAUUCAGGCUGAAGUGUACGACACCGUGG
UGAUCACCCUGAAGAACAUGGCCUCCACCCUGUCUCUCUGCACGCCGUCGGCGU
GUCGUACUGGAAGGCCAGCGAGGGCGCUGAAUACGACGACCAGACCAGCCAAAGA
GAGAAGGAGGAUGAUAAAGGUGUUCCAGGUGGCAGCCACACCUACGUGUGGCAGG
UGCUGAAAGAGAACGGCCCAAUGGCCUCAGACCCUCUGUGCCUGACAUAACAGCUA
UUUAUCCCACGUUGACCUGGUGAAGGACCUGAACAGCGGACUGAUCGGCGCCUG
CUGGUGUGCAGAGAAGGCUCCUGGCCAAGGAAAAGACACAGACACUACACAAGU
UCAUCUUAUCUGUUUGCCGUGUUCGACGAGGGCAAGAGCUGGCACAGCGAGACCAA
GAAUAGCCUGAUGCAAGACAGAGAUGCCGCCUCCGCUAGAGCGUGGCCUAAAUG
CACACCGUGAACGGCUACGUCAACAGAUUCUUGCCCGGCCUGAUUGGCUGUCACC
GGAAGUCCGUGUACUGGCACGUGAUCGGAAUGGGCACAAACCCCGGAGGUGCACUC
AAUCUUCUGGAAGGCCACACGUUCCUGGUGAGAAACACCGGCAGGCCUCGCUC
GAGAUUCUUUUUAUCACAUUUCUGACCGCCAGACCUUGCUCUAUGGACCUGGGCC
AGUUUCUGCUUUUCUGCCACAUCAGCUCCACACAGCACGACGGCAUGGAGGCUUA
CGUGAAGGUGGAUUCUUGUCCUGAGGAGCCUCAGCUGAGAAUGAAGAAUAAUGAG
GAAGCCGAAGAUUACGACGACGACCUGACGGACAGCGAGAUGGAUGUUGUGAGGU
UUGACGAUGAUAAUAGCCCUUCCUUUAUCCAGAUCCGGUCAGUGGCGAAGAAACA
CCCUAAGACAUGGGUGCACUACAUCGCCGCGUGAGGAGGAGGAUUGGGACUACGCC
CCCCUGGUGCUGGCCCCUGACGACCGGAGCUACAAUUCUCAGUACCUGAACACG

10

20

30

40

50

GCCCUCAAAGAAUCGGCAGAAAGUACAAGAAGGUGAGAUUCAUGGCUUAUACAGA
UGAAACAUAUUAAAACCAGAGAAGCUAUCCAGCAUGAGAGCGGAAUCCUGGGACCU
CUGCUGUACGGCGAGGUGGGCGACACCCUGCUGAUCAUCUUUAAGAACCAGGCCA
GCAGACCUUACAAUAUCUACCCUCACGGCAUUACAGACGUGAGACCUCUGUACAG
CCGGCGGCUGCCCAAGGGCGUGAAGCACCCUGAAGGACUUCCCAAUCCUGCCUGGA
GAGAUUCUUAAGUACAAGUGGACCGUGACAGUGGAAGAUGGGCCCUACCAAGUCUG
ACCCCCGGUGCCUGACCCGGUACUACUCGAGUUUUGUGAACAUGGAGCGCGACCU
GGCCUCUGGACUCAUCGGCCCAUUGCUGAUUCUGCUACAAAGAGUCCGUGGACCAG
AGAGGCAACCAGAUCAUGAGCGACAAGCGUAACGUGAUCCUGUUCAGCGUGUUUG
AUGAGAAUAGAAGCUGGUAUCUGACCGAGAACAUCAGCGGUUCCUCCCCAACCC
UGCCGGCGUGCAGCUGGAGGACCCUGAGUUCAGGCCAGCAACAUCAUGCACUCU
AUUAACGGCUACGUGUUCGACUCCUGCAGCUAAGCGUUUGCCUGCACGAGGUGG
CUUACUGGUACAUCUGAGCAUCGGCGCCAGACCGACUUCUGUCCGUGUUCUU
CUCAGGCUACACAUAUCAAACAUAAGAUGGUGUAUGAGGACACCCUGACCCUGUUC
CCCUUUUCCGGAGAGACCGUGUUAUGAGCAUGGAGAACCUGGGCCUUUGGAUCC
UGGGCUGUCACAAUAGCGACUUCGGGAACCGGGGAUGACCGCCUGCUGAAAGU
CAGCUCUUGCGACAAGAACACCGGGCGACUACUACGAGGACAGCUAUGAGGACAUC
UCCGCCUACCUGCUGAGCAAGAACAACGCCAUCGAGCCUCGGUCGUUCAGCCAGA
ACCCCCCUGUGCUGAAGAGACACCAGAGGGGAGAUACCCGGACAACCCUGCAGAG
CGAUCAGGAGGAGAU CGAUUAUGAUGACACUAUCAGCGUGGAGAUGAAGAAAGAG
GACUUCGACAUCUACGACGAGGACGAAAACCAAAGCCUCGGAGCUUCCAGAAGA
AAACCAGACACUAUUUCAUCGCCGCUGUGGAGAGACUGUGGGACUACGGCAUGAG
CUCCUCUCCCCACGUGCUGCGGAACCGGGCCCAGAGCGGAUCUGUUCCUCAGUUC
AAGAAAGUCGUGUUC CAAGAGUUCACCGACGGCUCUUUUACCCAGCCUCUGUAUA
GAGGCGAAUUAACGAACACCUGGGCCUGCUGGGGCCCUUACAUCAGAGCAGAAGU
UGAGGAUAACAUC AUGGUGACCUUCAGAAAUACAGGCCUCUCGGCCCUACUCUUUC
UACUCCUCGCUCAUCAGCUACGAGGAGGACCAGAGGCAAGGCGCCGAACCCAGAA
AGAACUUCGUGAAGCCCAACGAAACCAAGACUUAUUUUGGAAGGUGCAACACCA
CAUGGGCCCUACCAAAGACGAGUUUGAUUGCAAGGCCUGGGCCUACUUCAGCGAU
GUGGACCUGGAAAAGGAUGUCCAUAAGCGGCCUGAUUGGGCCCUUGCUGGUGUGCC
ACACAAACACACUCAAUCCUGCCACGGGAGACAGGUGACCGUGCAGGAGUUCGC
CCUGUUCUUCACCAUCUUCGACGAGACAAAGAGCUGGUACUUCACAGAAAACAUG
GAAAGAAAUUGCAGAGCCCCUUGCAACAUAUCAGAUUGGAGGAUCCACCCUUAAGG
AAAAUUAAGAUUCCACGCCAUCAACGGAUACAUAUGGACACACUGCCAGGACU
GGUUAUGGGCCAGGAUCAGAGAAUCCGGUGGUACCUGCUCUCUAUGGGCUCAAAC
GAGAAUAUCCACUCCAUCCACUUAAGCGGCCACGUGUUCACCGUGAGAAAGAAAG
AGGAGUACAAGAUGGGCCUGUACAAUCUGUACCCCGGCGUGUUCGAGACAGUGGA
AAUGCUGCCAAGCAAGGCCGGGAUCUGGCGGGUGGAUGUCUUAUUGGCGAGCAC
CUGCACGCCGGAUAGAGCACACUGUUUCUGGUUAUACAGCAACAAGUGCCAGACAC
CCUGGGCAUGGGCUUCUGGUCACAUCAGAGAUUUUCAGAUACAGCCUCUGGUCA
GUACGGACAGUGGGCUCCCAAGCUGGCCAGACUGCACUACUCCGGCUCCAUCAAC
GCCUGGAGCACAAAAGAGCCUUUCAGCUGGAUCAAGGUCGAUCUGCUGGCUCCAA
UGAUCAUCCACGGAAUCAAGACCCAGGGCGCCCGGCAGAAAUUCUCGAGCCUCUA
CAUCUCGCAGUUAUAUAUGUACAGCCUGGACGGAAAGAAGUGGCAGACCUAC
AGAGGAAAUAGUACCGGCACCCUGAUGGUGUUCUUCGGCAAUGUGGACAGCUCAG
GAAUUAAGCACAAACAUCUCAAUCCUCCAUAUAAGCCAGAUACAUCAGACUUA
UCCUACCCACUACUCUAUCAGAAGCACGCUGAGAAUGGAGUGGAUGGGAUUGUGAC
CUGAACAGCUGCAGCAUGCCUCUUGGCAUGGAAUCUAAAGCCAUCUCUGAUGCCC
AGAUCACCGCUUCUAGUUAUCUACCAACAUGUUCGCCACAUGGUCCCCUAGCAA
GGCCAGACUGCACCUGCAGGGCAGGUCUAAUGCCUGGCGGCCUCAGGUGAACAAAC

10

20

30

40

50

CCCAAGGAAUGGCUGCAGGUCGAUUUCCAGAAAACCAUGAAAGUGACCGGAGUGA
CCACCCAGGGCGUGAAAAGCCUGCUGACCAGCAUGUACGUGAAGGAGUUCCUGAU
CUCUAGCUGCCAGGAUGGCCACCAGUGGACCCUGUUCUUCAGAACGGCAAAGUG
AAAGUGUUCCAAGGCAACCAGGACAGCUUCACCCAGUCGUGAACAGCCUGGAUC
CACCUCUCCUUAACCAGAUACCUGAGAAUCCAUCCCCAGUCCUGGGUCCACCAGAU
CGCCCUGCGGAUGGAGGUGCUGGGCGUGCGAGGCCCAAGACCUGUACUAAGCUUGA
UCCAGACAUGAUAAAGAUACAUAUGAUGAGUUUGGACAAACCACAACUAGAAUGCAG
UGAAAAAAAUAGCUUUAUUUGUGAAAAUUUGUGAUGCUAUUAGCUUUAUUUGUAACC
AUUAUAAGCUGCAAUAAACAAGUUGUGGAAUUGUGAGCGCUCACAAUUCCACAGC
GGCCGCUGAGGUAGAUAAUCUUGUAUAGUGGGGGGGGAUCUCAUGUACCGGGUU
UCUUUUUAUUUGAUUUUCAAUAAAACAGACGGUAGCUAGGUUCGCAAGGCAGCCAC
AAGCCAAAGAUAGGUAGGGUGCUCAUAGUGAGUAGGGACAGUGCCUUUUGAUUCA
CAACGCGUCAAUACCAUCUGACACGGAUACCCUUAACCGGACUUGUCAUGAUCUCC
CAGACUUGUCCAAGGUGGACGGGCCACCUUUAACUUAACCCGGAAAAGGAACAUAU
AUUAAUUAUAUGUGUUCGGAAAAUAGCAAAAAAAAAAAAAAAAAAAAAA

10

525. TCA5_CBhBsi_F8N6mutant_EFS_TdT_GeFo3 RNA 配列

gggagacgggucaaccgcguaggagcggugaccgGAAUUCGGCGGGAGUAAACUAUGACUC
UCUUAAGGAGUCAUAGAGCCAGAACCUCUCGUGGUCCCGCUGGGGCACAGGGGAUU
AAUUUUUCUGUGGCAAAUUGACUGGCUUCAGAGAGCGUUUUUCGAAGUGGACU
GUGUGACUGCGUUCSCCCCUUAGUUGCUAUAUCCGCUUCGAUUAACAUCUCACCU
CGACGUUAAGAUAUUUCUAGAAGGUCGACCAGAUGUCCGAGGUCGACCAGUUG
UCCGUGUGGAAUUGUGAGCGCUCACAAUUCCACACGUUACAUAACUACGGUAAA
UGGCCCGCCUGGCUGACCGCCCAACGACCCCGCCCAUUGACGUCAAUAGUAACG
CCAAUAGGGACUUAUCCAUAUGACGUCAAUGGGUGGAGUAUUUACGGUAAACUGCCC
ACUUGGCAGUACAUAAGUGUAUCAUAUGCCAAGUACGCCCCCUAUUGACGUCAA
UGACGGUAAAUGGCCCGCCUGGCAUUGUGCCCAUACAUGACCUUAUGGGACUUU
CCUACUUGGCAGUACAUCUACGUAAUAGUCAUCGCUAAUACCAUGGUCGAGGUGA
GCCCCACGUUCUGCUUCACUCUCCCCAUCUCCCCCCCCUCCCCACCCCCAAUUUU
GUAUUUAUUUAUUUUUUAAUUAUUUUUGUGCAGCGAUUGGGGGCGGGGGGGGGGGG
GGGGCGCGCGCCAGGCGGGGCGGGGCGGGGCGAGGGGCGGGGCGGGGCGAGGCGG
AGAGGUGCGGCGGCAGCCAAUCAGAGCGGCGCGCUCCGAAAGUUUCCUUUAUGG
CGAGGCGGCGGCGGCGGCGGCCUUAUAAAAGCGAAGCGCGCGGCGGGCGGGAGU
CGCUGCGACGCUGCCUUCGCCCGUGCCCCGCUCCGCCGCGCCUCGCGCCGCCCG
CCCCGGCUCUGACUGACCGCGUUAUCUCCACAGGUGAGCGGGCGGGACGGCCCUU
CUCCUCCGGGCUGUAAUUAAGCUGAGCAAGAGGUAAAGGUUUAAGGGAUGGUUGGU
UGGUGGGGUAAUUAUGUUUAUUAUACCUGGAGCACCUGCCUGAAAUACUUUUUU
UCAGGUUGGCGUACGgccaaccAUGCAAUUCGAGCUGUCCACUUGCUUCUUCUGUG
CCUGCUGCGCUUUUGUUUCAGUGCUACCCGCAGAUACUACCUGGGAGCCGUGGAG
CUUUCUUGGGACUAUAUGCAGAGCGAUUCUGGGCGAGCUGCCAGUCGACGCCAGAU
UUCCACCUAGAGUGCCUAAGAGUUUCCCCUUAACACCUCUGUGGUGUACAAGAA
AACCUGUUUGUGGAGUUCACCGACCACCUGUUAUAUACGCUAAGCCUCGGCCU
CCUUGGAUGGGCCUGCUGGGCCCCACCAUUCAGGCUGAAGUGUACGACACCGUGG
UGAUCACCCUGAAGAACAUGGCCUCCACCCUGUCUCUCUGCACGCCGUCGGCGU
GUCGUACUGGAAGGCCAGCGAGGGCGCUGAAUACGACGACCAGACCAGCCAAAGA
GAGAAGGAGGAUGAUAAAGGUGUUCCAGGUGGCAGCCACACCUACGUGUGGCAGG
UGCUGAAAGAGAACGGCCCAAUGGCCUCAGACCCUCUGUGCCUGACAUAACAGCUA
UUUAUCCACAGUUGACCUGGUGAAGGACCUGAACAGCGGACUGAUCGGCGCCUG
CUGGUGUGCAGAGAAGGCUCUCCUGGCCAAGGAAAAGACACAGACACUACACAAGU
UCAUCUUAUCUGUUUGCCGUGUUCGACGAGGGCAAGAGCUGGCACAGCGAGACCAA
GAAUAGCCUGAUGCAAGACAGAGAUGCCGCCUCCGCUAGAGCGUGGCCUAAAUG

20

30

40

50

CACACCGUGAACGGCUACGUCAACAGAUCUCUGCCCCGGCCUGAUUGGCUGUCACC
GGAAGUCCGUGUACUGGCACGUGAUCGGAAUGGGCACAACCCCGGAGGUGCACUC
AAUCUUCCUGGAAGGCCACACGUUCCUGGUGAGAAACCACCGGCAGGCCUCGCUC
GAGAUCUCCCCUAUCACAUUUCUGACCGCCCCAGACCUUGCUCUAUGGACCUGGGCC
AGUUUCUGCUUUUCUGCCACAUCAGCUCCCAACAGCACGACGGCAUGGAGGCCUUA
CGUGAAGGUGGAUUCUUGUCCUGAGGAGCCUCAGCUGAGAAUGAAGAAUAAUGAG
GAAGCCGAAGAUUACGACGACGACCUGACGGACAGCGAGAUGGAUGUUGUGAGGU
UUGACGAUGAUAAUAGCCCUUCCUUUAUCCAGAUCCGGUCAGUGGCGAAGAAACA
CCCUAAGACAUGGGUGCACUACAUCGCCGCUGAGGAGGAGGAUUGGGACUACGCC
CCCCUGGUGCUGGCCCCUGACGACCGGAGCUACAAAUCUCAGUACCUGAACAACG
GCCCUCAAAGAAUCGGCAGAAAGUACAAGAAGGUGAGAUUCAUGGCCUUAUACAGA
UGAAACAUUUAAAACAGAGAAGCUAUCCAGCAUGAGAGCGGAAUCCUGGGACCU
CUGCUGUACGGCGAGGUGGGCGACACCCUGCUGAUCAUCUUUAAGAACCAGGCCA
GCAGACCUUACAUAUCUACCCUCACGGCAUUACAGACGUGAGACCUCUGUACAG
CCGGCGGCUGCCCAAGGGCGUGAAGCACCUGAAGGACUUCCAAUCCUGCCUGGA
GAGAUUCUACAAGUACAAGUGGACCGUGACAGUGGAAGAUGGGCCUACCAAGUCUG
ACCCCCGGUGCCUGACCCGGUACUACUCGAGUUUUGUGAACAUGGAGCGCGACCU
GGCCUCUGGACUCAUCGGCCCAUUGCUGAUCUGCUACAAAGAGUCCGUGGACCAG
AGAGGCAACCAGAUCAUGAGCGACAAGCGUAACGUGAUCCUGUUCAGCGUGUUUG
AUGAGAAUAGAAGCUGGUUUCUGACCGAGAACAUCAGCGGUUCCUCCCCAACCC
UGCCGGCGUGCAGCUGGAGGACCCUGAGUUCAGGCCAGCAACAUCAUGCACUCU
AUUAACGGCUACGUGUUCGACUCCUGCAGCUAAGCGUUUGCCUGCACGAGGUGG
CUUACUGGUACAUCUGAGCAUCGGCGCCAGACCGACUUCUGUCCGUGUUCUU
CUCAGGCUACACAUAUCAAACAUAAGAUGGUGUAUGAGGACACCCUGACCCUGUUC
CCCUUUUCCGGAGAGACCGUGUUCAUGAGCAUGGAGAACCUGGCCUUUGGAUCC
UGGGCUGUCACAUAAGCGACUUCGGAACCGGGGAUGACCGCCCUGCUGAAAGU
CAGCUCUUGCGACAAGAACACCGGCGACUACUACGAGGACAGCUAUGAGGACAUC
UCCGCCUACCUGCUGAGCAAGAACAACGCCAUUCGAGCCUCGGUCGUUCAGCCAGA
ACCCCCUGUGCUGAAGAGACACCAGAGGGGAGAUACCCGGACAACCCUGCAGAG
CGAUCAGGAGGAGAUUCGAUUAUGAUGACACUAUCAGCGUGGAGAUGAAGAAAGAG
GACUUCGACAUCUACGACGAGGACGAAAACCAAAGCCUCGGAGCUUCCAGAAGA
AAACCAGACACUAUUUCAUCGCCGCUGUGGAGAGACUGUGGGACUACGGCAUGAG
CUCCUCUCCCCACGUGCUGCGGAACCGGGCCCAGAGCGGAUCUGUUCUCAGUUC
AAGAAAGUCGUGUUCCAAGAGUUCACCGACGGCUCUUUUACCCAGCCUCUGUAUA
GAGGCGAAUUAACGAACACCUGGGCCUGCUGGGGCCUUACAUCAGAGCAGAAGU
UGAGGAUAACAUAUGGUGACCUUCAGAAAUCAGGCCUCUCGGCCCUACUCUUUC
UACUCCUCGCUCAUCAGCUACGAGGAGGACCAGAGGCAAGGCGCCGAACCCAGAA
AGAACUUCGUGAAGCCCAACGAAACCAAGACUUAUUUGGAAGGUGCAACACCA
CAUGGCCCCUACCAAAAGACGAGUUUGAUUGCAAGGCCUGGGCCUACUUCAGCGAU
GUGGACCUGGAAAAGGAUGUCCAUAAGCGGCCUGAUUGGGCCCUUGCUGGUGUGCC
ACACAAACACACUCAAUCCUGCCACGGGAGACAGGUGACCGUGCAGGAGUUCGC
CCUGUUCUUCACCAUCUUCGACGAGACAAAGAGCUGGUACUUCACAGAAAACAUG
GAAAGAAAUUGCAGAGCCCCUUGCAACAUAUCAGAUGGAGGAUCCACCUUCAAGG
AAAAUUACAGAUUCCACGCCAUCAACGGAUACAUAUGGACACACUGCCAGGACU
GGUUAUGGGCCAGGAUCAGAGAAUCCGGUGGUACCUGCUCUCUAUGGGCUCAAAC
GAGAAUAUCCACUCCAUCCACUUUAGCGGCCACGUGUUCACCGUGAGAAAGAAAG
AGGAGUACAAGAUGGGCCUGUACAAUCUGUACCCCGGCGUGUUCGAGACAGUGGA
AAUGCUGCCAAGCAAGGCCGGGAUCUGGCGGGUGGAUGUCUUAUUGGCGAGCAC
CUGCACGCCGGAUGAGCACACUGUUUCUGGUUAUACAGCAACAAGUGCCAGACAC
CCUGGGCAUGGCUUCUGGUCACAUCAGAGAUUUUCAGAUCACAGCCUCUGGUCA

10

20

30

40

50

GUACGGACAGUGGGGCUCCCAAGCUGGGCCAGACUGCACUACUCCGGGCUCCAUCAAC
GCCUGGAGCACAAAAGAGCCUUUCAGCUGGAUCAAGGUCGAUCUGCUGGGCUCCAA
UGAUCAUCCACGGAAUCAAGACCCAGGGCGCCCGGCAGAAAUUCUCGAGCCUCUA
CAUCUCGCAGUUCAUUUAUCAUGUACAGCCUGGACGGAAAGAAGUGGCAGACCUAC
AGAGGAAAUAGUACCGGCACCCUGAUGGUGUUCUUCGGCAAUGUGGACAGCUCAG
GAAUUAAGCACAAACAUUCUCAAUCCUCCAUAUCAUAGCCAGAUACAUCAGACUUCA
UCCUACCCACUACUCUAUCAGAAGCACGCUGAGAAUGGAGUGGAUGGGGAUGUGAC
CUGAACAGCUGCAGCAUGCCUCUUGGCAUGGAAUCUAAAGCCAUCUCUGAUGCCC
AGAUCACCGCUUCUAGUUACUUCACCAACAUGUUCGCCACAUGGUCCCCUAGCAA
GGCCAGACUGCACCUGCAGGGCAGGUCUAAUGCCUGGGCGGCCUCAGGGUGAACAAAC
CCCAAGGAAUGGGCUGCAGGUCGAUUUCCAGAAAACCAUGAAAGUGACCGGAGUGA
CCACCCAGGGCGUGAAAAGCCUGCUGACCAGCAUGUACGUGAAGGAGUUCUGAU
CUCUAGCUCUCCAGGAUGGGCCACCAGUGGACCCUGUUCUUCAGAACGGCAAAGUG
AAAGUGUUCCAAGGCAACCAGGACAGCUUCACCCAGUCGUGAACAGCCUGGAUC
CACCUCUCCUUAACCAGAUACCUGAGAAUCCAUCCCAGUCCUGGGUCCACCAGAU
CGCCCUGCGGAUGGAGGUGCUGGGCUGCGAGGCCCAAGACCUGUACUAAGCUUGA
UCCAGACAUGAUAAAGAUACAUGAUGAGUUUGGACAAACCACAACUAGAAUGCAG
UGAAAAAAAUUGC UUUAUUUGUGAAAUUUGUGAUGCUAUUUGC UUUAUUUGUAACC
AUUAUAAGCUGCAAUAAACAAGUUGUGGAAUUGUGAGCGCUCACAAUCCACAGC
GGCCGCGgggcagagcgcacaucgcccacaguccccgagaaguuggggggagggggucggcaauu
gaaccggugccuagagaagguggcgcgggguaaacugggaaagugaugucguguacuggcucc
gccuuuuucccgagggugggggagaaccguauuaaagugcaguagucgccgugaacguucuuu
uucgcaacggguuugccgccagaacacaggcguacggccaccaugguguccaaaggcgaggaag
ugaucaaggaaauuaugagauucaaggucagaauuggaaggaucaugaacggccacgaguucg
agaucgaggggagaaggagaaggcagaccuacgagggcacccaaaccgcuaaaauuaaaagugac
aaaggggcggaccucugccguucgccugggauauccuguccccccaguucauguacggcagcaag
gccuacgugaagcaccucugacauccccgauuacaaaagcugucuuuuccugagggcuua
aguggggagagagugaugaacuucgaggacggagggccugugacgguuacccaagacagcagccu
gcaagauggaacaacugaucuacaaagugaagaugcggggcacuaauuuuccuccagacggcccc
gugaugcagaagaaaaccaugggucgggagggccucuaacagaacggcuguaccccagagacggag
ugcugaaggggcgagaauaccaggcacugaagcugaagauggcggacacuaccugguggagu
ucaagaccuuuacauuggccaagaagccugugcagcugccuggcuacuacuacguggacaccaa
gcuggacaucaccucccacaauagaggauuauaccaucguggaacaguacgagagaagcgagggc
agacaccaccuguuuccugggccacggcacaggcuccaccggcucuggcagcagcggcacagccag
cagcagggacaacaacauggccgugaucuaaggaguucaugcgguucaaggugcggauuggaagg
caguaugaacggccaugaguucgaaaucgagggcggaaggcggaaggcagaccuuacgagggcaca
cagacagcuaaagcugaaggugaccaaggaggagggccucugccauucgccugggacauccugagcc
cucaguuuauanguauggcagcaaggccuacgugaagcacccccgccgacauccccgacuacaagaa
gcugagcuucccugagggguuucaaaugggaaagggugaugaacuucgaggacggcggacuggu
gacagugacccaggacagcagccugcaggauggcaccucgaucuaauaaggugaagaugagaggc
accaacuuccaccugacgggcccgugaugcagaaaaaaaccauggggaugggagggccagcaccg
agcggcuguacccuagagauggcguucugaaggggcgagauccaccaggcccugaagcugaagga
cggcgggccacuaccugguggaguucaagaccaucuaauaggccaagaaaccugugcagcugccu
ggauacuacuacguggacacaaagcuggacauaccucucacaaacgaggacuaauaccaucgucg
agcaguacgaacgcagcgaaggcagacaccaucguuccuguaacgguaugggacgagcuguacaa
guaaugauuaauuaaaauaaaagaucuuuanguucaaunagaucuguguguuugguuuuugugu
gugugggaaugugagcgcucacaaauccacagcggccgcUGAGGUAGAUAUAUCUUUGUAU
AGUGGGGGGGGAUCUCAUGUACCGGGUUUCUUAUUAUUGAUUUUCAUAUAAACA
GACGGUAGCUAGGUUCGCAAGGCAGCCACAAGCCAAAGAUAGGUAGGGUGCUCAU
AGUGAGUAGGGACAGUGCCUUUUGAUUACACAACGCGUCAAUACCAUCUGACACGG

10

20

30

40

50

AUACCCUUACCGGACUUGUCAUGAUCUCCCAGACUUGUCCAAGGUGGACGGGGCCA
CCUUUACUUAACCCGGAAAAGGAACAUAUAUUAUUAUUAUGUGUUCGGAAAUAAG
CAAAAAAAAAAAAAAAAAAAAAAAAAA

526. TCARZ_CBhBsi_GFP_GeFo3_R4A22

gggagacgggucaaccgcuaggagcgguagaccggAAUUCGGCGGGAGUAACUAUGACUC
UCUUAAGGAGUCAUAGAGCCAGAACCUCUCGUGGUCCCGCUGGGGCACAGGGGAUU
AAUUUUUCUGUGGCAAAUUGACUGGCUCUAGAAGGUCGACCAGAUGUCCGAGGU
CGACCAGUUGUCCGUGUGGAAUUGUGAGCGCUCACAAUCCACACGUUACAUAAC
UUACGGUAAAUGGCCCGCCUGGCUGACCGCCCAACGACCCCCGCCCAUUGACGUC
AAUAGUAAACGCCAAUAGGGACUUUCCAUUGACGUCAAUGGGUGGAGUAUUUACGG
UAAACUGCCCACUUGGCAGUACAUAAGUGUAUCAUAUGCCAAGUACGCCCCCUA
UUGACGUCAAUGACGGUAAAUGGCCCGCCUGGC AUUGUGCCCAGUACAUGACCUU
AUGGGACUUUCCUACUUGGCAGUACAUCUACGUAAUAGUCAUCGCUAUUACCAUG
GUCGAGGUGAGCCCCACGUUCUGCUUCACUCUCCCCAUUCCCCCCCCCUCCCCAC
CCCCAAUUUUUGUAUUUAUUUAUUUUUAAUUAUUUUUGUGCAGCGAUGGGGGGCGG
GGGGGGGGGGGGGGGGGGCGCGCGCCAGGCGGGGGCGGGGCGGGGCGAGGGGGCGGGGCGG
GGCGAGGCGGAGAGGUGCGGGCGGCAGCCAAUCAGAGCGGCGCGCUCCGAAAGUUU
CCUUUUUAUGGCGAGGCGGCGGCGGCGGCGGCCCUAUA AAAAGCGAAGCGCGCGGC
GGGCGGGAGUCGCUUGCGACGCUUGCCUUCGCCCCGUGCCCCGCUCCGCCGCCGCCUC
GCGCCGCCCGCCCCGGCUCUGACUGACCGCGUUAUCUCCACAGGUGAGCGGGCGG
GACGGCCCUUCUCCUCCGGGCGUGAAUUAAGCUGAGCAAGAGGUAAGGGUUUAAGG
GAUGGUUGGUUGGUGGGGUAUUAAUGUUUAAUUAACUGGAGCACCUGCCUGAAA
UCACUUUUUUUCAGGUUGGCGUACGGCCACCAUGGUGAGCAAGGGCGAGGAGCUG
UUCACCGGGGUGGUGCCCAUCCUGGUCGAGCUGGACGGCGACGUAAACGGCCACA
AGUUCAGCGUGUCCGGCGAGGGCGAGGGCGAUGCCACCUACGGCAAGCUGACCCU
GAAGUUCAUCUGCACCACCGGCAAGCUGCCCGUGCCCUUGGCCACCCUCGUGACC
ACCCUGACCUACGGCGUGCAGUGCUUCAGCCGCUACCCCGACCACAUGAAGCAGC
ACGACUUCUUAAGUCCGCCAUGCCC GAAGGCUACGUCCAGGAGCGCACCAUCUU
CUUCAAGGACGACGGCAACUACAAGACCCGCGCCGAGGUGAAGUUCGAGGGGCGAC
ACCCUGGUGAACC GCAUCGAGCUGAAGGGCAUCGACUUAAGGAGGACGGCAACA
UCCUGGGGCAACAAGCUGGAGUACAACUACAACAGCCACAACGUCUAUAUCAUGGC
CGACAAGCAGAAGAACGGCAUCAAGGUGAACUUAAGA UCCGCCACAACAUCGAG
GACGGCAGCGUGCAGCUCGCCGACCACUACCAGCAGAACACCCCCAU CGGCGACGG
CCCCGUGCUGCUGCCCCGACAACCACUACCUGAGCACCCAGUCCGCCCUAGCAAAAG
ACCCCAACGAGAAGCGCGAUCACAUGGUCCUGCUGGAGUUCGUGACCGCCGCCGG
GAUCACUCUCGGCAUGGACGAGCUGUACAAGUAAAAGCUUGAUCCAGACAUGAUA
AGAUACA UUGAUGAGUUUGGACAAACCACAACUAGAAUGCAGUGAAAAAAUUGCU
UUAUUUGUGAAA UUGUGAUGCUAUUGCUUUUAUUUGUAACCAUUAUAAGCUGCA
AUAACAAGUUGUGGAAUUGUGAGCGCUCACAAUUCACAGCGGCCGCGUGAGGUA
GAUAAUCUUUGUAUAGUGGGGGGGGAUCUCAUGUACCGGGUUUCUUUUAUUUGA
UUUUCAAUAAAACAGACGGUAGCUAGGUUCGCAAGGCAGCCACAAGCCAAAGUA
GGUAGGGUGCUCAUAGUGAGUAGGGACAGUGCCUUUUGAUUCACAACGCGUCAAU
ACCAUCUGACACGGAUACCCUUAACGGACUUGUCAUGAUCUCCAGACUUGUCCA
AGGUGGACGGGCCACCUUUA CUUAACCCGGAAAAGGAACAUAUAUUAUUAUUAUG
UGUUCGGAAAUAAGCAAAAAAAAAAAAAAAAAAAAAAAAAA

527. TCARZ_CBh_GFP_GeFo3

gggagacgggucaaccgcuaggagcgguagaccggAAUUCGGCGGGAGUAACUAUGACUCU
CUUAAGGAGUCAUAGAGCCAGAACCUCUCGUGGUCCCGCUGGGGCACAGGGGAUUA
AUUUUUUCUGUGGCAAAUUGACUGGCUCUAGAAGGUCGACCAGAUGUCCGAGGU
GACCAGUUGUCCGUGUGGAAUUGUGAGCGCUCACAAUUCACACACGUUACAUAACU

10

20

30

40

50

UACGGUAAAUGGCCCGCCUGGCUGACCGCCCAACGACCCCCGCCCAUUGACGUCA
AUAGUAAACGCCAAUAGGGACUUUCCAUUGACGUCAAUGGGUGGAGUAUUUACGGU
AAACUGCCCACUUGGCAGUACAUCAAGUGUAUCAUAUGCCAAGUACGCCCCCUAU
UGACGUCAAUGACGGUAAAUGGCCCGCCUGGCAGUUGUGCCCAGUACAUGACCUUA
UGGGACUUUCCUACUUGGCAGUACAUCUACGUAAUAGUCAUCGCUAAUACCAUGG
UCGAGGUGAGCCCCACGUUCUGCUUCACUCUCCCCAUUCUCCCCCCCCUCCCCACCC
CCAAUUUUUGUAUUUAUUUAUUUUUAAUUAUUUUGUGCAGCGAUGGGGGCGGGG
GGGGGGGGGGGGCGCGCGCCAGGCGGGGCGGGGCGGGGCGAGGGGCGGGGCGGGG
CGAGGCGGAGAGGUGCGGCGGCAGCCAAUCAGAGCGGCGCGCUCCGAAAGUUUCC
UUUUUAUGGCGAGGCGGCGGCGGCGGCGGCCUAUAAAAAGCGAAGCGCGCGGGCGG
GCGGGAGUCGCGUGCGACGCGUGCCUUCGCCCCGUGCCCCGCUCCGCCGCCGCCUCGC
GCCGCCCGCCCCGGCUCUGACUGACCGCGUACUCCACAGGUGAGCGGGCGGGA
CGGCCCUCUCCUCCGGGCGUGAAUUAAGCUGAGCAAGAGGUAAAGGGUUUAAGGGA
UGGUUGGUUGGUGGGGUAAUUAUGUUAAUUAACUGGAGCACCUGCCUGAAAU
ACUUUUUUUUCAGGUUGGCGUACGGCCACCAUGGUGAGCAAGGGCGAGGAGCUGUU
CACCGGGGUGGUGCCCAUCCUGGUCGAGCUGGACGGCGACGUAAACGGCCACAAG
UUCAGCGUGUCCGGCGAGGGCGAGGGCGAUGCCACCUACGGCAAGCUGACCCUGA
AGUUCAUCUGCACCAACCGGCAAGCUGCCCGUGCCUGGCCACCCUCGUGACCAC
CCUGACCUACGGCGUGCAGUGCUUCAGCCGCUACCCCGACCACAUGAAGCAGCAC
GACUUCUUAAGUCCGCCAUGCCCCGAAGGCUACGUCCAGGAGCGCACCAUCUUCU
UCAAGGACGACGGCAACUACAAGACCCGCGCCGAGGUGAAGUUCGAGGGCGACAC
CCUGGUGAACC GCAUCGAGCUGAAGGGCAUCGACUUAAGGAGGACGGCAACAUC
CUGGGGCGACAAGCUGGAGUACAACUACAACAGCCACAACGUCUAUAUCAUGGCCG
ACAAGCAGAAGAACGGCAUCAAGGUGAACUUAAGAUCGCGCCACAACAUCGAGGA
CGGCAGCGUGCAGCUCGCCGACCACUACCAGCAGAACACCCCCAUCCGGCGACGGCC
CCGUGCUGCUGCCCGACAACCACUACCUGAGCACCCAGUCCGCCCUAGAGCAAAGAC
CCCAACGAGAAGCGCGAUCACAUGGUCCUGCUGGAGUUCGUGACCGCCGCCGGGA
UCACUCUCGGCAUGGACGAGCUGUACAAGUAAAAGCUUGAUCCAGACAUGAUAAAG
AUACAUAUGAUGAGUUUGGACAAACCACAACUAGAAUGCAGUGAAAAAAAUGCUUU
AUUUGUGAAAAUUUGUGAUGCUAUUGCUUUUAUUUGUAACCAUUAUAAGCUGCAAU
AAACAAGUUGUGGAAUUGUGAGCGCUCACAAUUCCACAGCGGCCGCUGAGGUAGA
UAAUCUUUGUAUAGUGGGGGGGGAUCUCAUGUACCGGGUUUCUUUAUUUGAUU
UUCAUAAAACAGACGGUAGCUAGGUUCGCAAGGCAGCCACAAGCCAAAGAUAGG
UAGGGUGCUCAUAGUGAGUAGGGACAGUGCCUUUUGAUUCACAACGCGUCAAUAC
CAUCUGACACGGAUACCCUUAACCGGACUUGUCAUGAUCUCCAGACUUGUCCAAG
GUGGACGGGCCACCUUUAACUUAACCCGGAAAAGGAACAUAUAUUAUUAUGUG
UUCGGAAAAUAGCAAAAAAAAAAAAAAAAAAAAAA

10

20

30

40

50

528. TCARZ_CMV*_GFP_GeFo3

ggguaaacGGCGGGAGUAACUAUGACUCUCUUAAGGAGUCAUAGAGCCAGAACCUC
CUCGUGGUCCCCGUGGGGCACAGGGAUUAUUUUUCUGUGGCAAAUUGACUGGCU
CUAGACGUUACAUAACUUAACGGUAAAUGGCCCGCCUGGCUGACCGCCCAACGACC
CCCGCCCAUUGACGUCAAUAAUGACGUUUGUCCCAUAGUAACGCCAAUAGGGAC
UUUCCAUUGACGUCAAUGGGUGGAGUAUUUACGGUAAACUGCCCACUUGGCAGUA
CAUCAAGUGUAUCAUAUGCCAAGUACGCCCCCUAUUGACGUCAAUGACGGUAAAU
GGCCCGCCUGGCAUUAUGCCCAGUACAUGACCUUAUGGGACUUUCCUACUUGGCA
GUACAUCUACGUAAUAGUCAUCGCUAAUACCAUGGUGAUGCGGUUUUGGCAGUAC
AUCAAUGGGGCGUGGAUAGGGGUUUGACUCACGGGGAUUUCCAAGUCUCCACCCCA
UUGACGUCAAUGGGAGUUUGUUUUGGCACCAAAAUCAACGGGACUUUCCAAAUG
UCGUAAACAACUCCGCCCAUUGACGCAAAUGGGCGGUAGGCGUGUACGGUGGGAG
GUCUAUAUAAGCAGAGCUCGUUUAGUGAACCGCUUGUCUCGCUCCGGGGAACGCU

CGGAAACUCCCGGCCGCCGCCACCCGCGUCUGUUCUGUUAACACAAGGGGAAGAAAA
GCCGUGCCGCACUCCGAGUGUGUACGGCCACCAUGGUGAGCAAGGGCGAGGAG
CUGUUCACCGGGGUGGUGCCCAUCCUGGUCGAGCUGGACGGCGACGUAAACGGCC
ACAAGUUCAGCGUGUCCGGCGAGGGCGAGGGCGAUGCCACCUACGGCAAGCUGAC
CCUGAAGUUCAUCUGCACCAACCGGCAAGCUGCCCGUGCCCUGGGCCACCCUCGUG
ACCACCCUGACCUACGGCGUGCAGUGCUUCAGCCGCUACCCCGACCAACAUGAAGC
AGCACGACUUCUUAAGUCCGCCAUGCCCGAAGGCUACGUCCAGGAGCGCACCAU
CUUCUUAAGGACGACGGCAACUACAAGACCCGCGCCGAGGUGAAGUUCGAGGGC
GACACCCUGGUGAACCGCCAUCGAGCUGAAGGGCAUCGACUUAAGGAGGACGGCA
ACAUCCUGGGGACACAAGCUGGAGUACAACUACAACAGCCACAACGUCUAUAUCAU
GGCCGACAAGCAGAAGAACGGCAUCAAGGUGAACUUAAGAUCGCGCCACAACAUC
GAGGACGGCAGCGUGCAGCUCGCCGACCACUACCAGCAGAACACCCCCAUCGGCGA
CGGCCCCGUGCUGCUGCCCGACAACCACUACCUGAGCACCCAGUCCGCCCUGAGCA
AAGACCCCAACGAGAAGCGCGAUCACAUGGUCCUGCUGGAGUUCGUGACCGCCGC
CGGGAUCACUCUCGGCAUGGACGAGCUGUACAAGUAAAAUAAAAGAUCUUUAUGU
UCAUUAGAUCUGUGUGUUGGUUUUUGUGUGGCUGAGGUAGAUAAUCUUUGUAUA
GUGGGGGGGGAUCUCAUGUACCGGGUUCUUUUUAUUUGAUUUUCAAUAAAACAG
ACGGUAGCUAGGUUCGCAAGGCAGCCACAAGCCAAAGAUAGGUAGGGUGCUCAUA
GUGAGUAGGGACAGUGCCUUUUGAUUCACAACGCGUCAAUACCAUCUGACACGGA
UACCCUUACCGGACUUGUCAUGAUCUCCAGACUUGUCCAAGGUGGACGGGCCAC
CUUUACUUAACCCGGAAAAGGAACAUAUAUUAUUAUUGUGUUCGGAAAUAAGC
AAAAAAAAAAAAAAAAAAAAA

10

20

30

40

50

529. TCARZ_SV40*_GFP_GeFo3

ggguaaacGGCGGGGAGUAAUAUGACUCUCUUAAGGAGUCAUAGAGCCAGAACCUC
CUCGUGGUCCCGCUGGGCACAGGGAUUAUUUUUCUGUGGCAAAUUUGACUGGCU
CUAGAGCUGUGGAAUGUGUGUCAGUUAGGGUGUGGAAAGUCCCCAGGCUCCCCAG
CAGGCAGAAGUAUGCAAAGCAUGCAUCUCAAUUAGUCAGCAACCAGGUGUGGAAA
GUCCCCAGGCUCCCCAGCAGGCAGAAGUAUGCAAAGCAUGCAUCUCAAUUAGUCA
GCAACCAUAGUCCCGCCCCUAACUCCGCCCAUCCCGCCCCUAACUCCGCCCAGUUC
CGCCCAUUCUCCGCCCAUGGCUGACUAAUUGCGUCUAUAUAUGCAGAGGCCGAG
GCCGCCUCGGCCUCUGAGCUAUUCCAGAAGUAGUGAGGAGGCCGUACGGCCACCA
UGGUGAGCAAGGGCGAGGAGCUGUUCACCGGGGUGGUGCCCAUCCUGGUCGAGCU
GGACGGCGACGUAAACGGCCACAAGUUCAGCGUGUCCGGCGAGGGGCGAGGGGCGAU
GCCACCUACGGCAAGCUGACCCUGAAGUUCAUCUGCACCAACCGGCAAGCUGCCCG
UGCCCUGGGCCACCCUCGUGACCACCCUGACCUACGGCGUGCAGUGCUUCAGCCG
CUACCCCGACCAACAUGAAGCAGCACGACUUCUUAAGUCCGCCAUGCCCGAAGGC
UACGUCCAGGAGCGCACCAUCUUCUUAAGGACGACGGCAACUACAAGACCCGCG
CCGAGGUGAAGUUCGAGGGCGACACCCUGGUGAACCAGCAUCGAGCUGAAGGGCAU
CGACUUAAGGAGGACGGCAACAUCUGGGGCAACAAGCUGGAGUACAACUACAAC
AGCCACAACGUCUAUAUCAUGGCCGACAAGCAGAAGAACGGCAUCAAGGUGAACU
UCAAGAUCGCGCCACAACAUCGAGGACGGCAGCGUGCAGCUCGCCGACCACUACCA
GCAGAACACCCCAUCGGCGACGGCCCCGUGCUGCUGCCCGACAACCACUACCUGA
GCACCCAGUCCGCCUGAGCAAAGACCCCAACGAGAAGCGCGAUCACAUGGUCCU
GCUGGAGUUCGUGACCGCCGCCGGGAUCACUCUCGGCAUGGACGAGCUGUACAAG
UAAAUAUAAAGAUCUUUAUGUUCAUUAGAUCUGUGUGUUGGUUUUUGUGUGGCU
GAGGUAGAUAAUCUUUGUAUAGUGGGGGGGGAUCUCAUGUACCGGGUUCUUUU
AUUUGAUUUUCAAUAAAACAGACGGUAGCUAGGUUCGCAAGGCAGCCACAAGCCA
AAGAUAGGUAGGGUGCUCAUAGUGAGUAGGGACAGUGCCUUUUGAUUCACAACGC
GUCAAUACCAUCUGACACGGAUACCCUUACCGGACUUGUCAUGAUCUCCAGACU
UGUCCAAGGUGGACGGGCCACCUUACUUAACCCGGAAAAGGAACAUAUAUUAU

UAUAUGUGUUCGGAAAAUAGCAAAAAAAAAAAAAAAAAAAAAA

530. TCA5_CBh_mCherry_GeFo3 gggagacggucaaccgcguaggagcggugaccgga
auucGGCGGGGAGUAAUAUGACUCUCUUAAGGAGUCAUAGAGCCAGAACCUCUC
GUGGUCCCCGCUGGGGCACAGGGAUUAUUUUUCUGUGGCCAAAUUUGACUGGCUUCA
GAGAGCGUUUUUCGAAGUGGACUGUGUGACUGCGUUCCCCCCUUAGUUGCUAUUAU
CCGCUUCGAUUAACAUCUCACCUCGACGUUAUAGAUAUUCUAGAAGGUCGACC
AGAUGUCCGAGGUCGACCAGUUGUCCGUGUGGAUUGUGAGCGCUCACAAUCCA
CACGUUACAUAACUUAACGGUAAAUGGCCCGCCUGGCUGACCGCCCAACGACCCCC
GCCCAUUGACGUCAAUAGUAACGCCAAUAGGGACUUCUCCAUGACGUCAAUGGGU
GGAGUAUUUACGGUAAACUGCCCACUUGGCAGUACAUAAGUGUAUCAUAUGCCA
AGUACGCCCCCUAUUGACGUCAAUGACGGUAAAUGGCCCGCCUGGCAUUGUGCCC
AGUACAUGACCUUAUGGGACUUCUACUUGGCAGUACAUCUACGUAAUAGUCAU
CGCUAAUACCAUGGUCGAGGUGAGCCCCACGUUCUGCUUCACUCUCCCCAUCUCC
CCCCCUCUCCCCACCCCCAAUUUUGUAUUUAUUUAUUUUUAAUUAUUUUGUGCA
GCGAUGGGGGGCGGGGGGGGGGGGGGGGGGGCGCGCGCCAGGCGGGGGCGGGGGCGGGGGCG
AGGGGGCGGGGGCGGGGGCGAGGCGGAGAGGUGCGGCGGCAGCCAAUCAGAGCGGGCGC
GCUCCGAAAGUUUCCUUUAUGGCGAGGCGGCGGCGGCGGCGGCGGCCCCUAUAAAAAG
CGAAGCGCGCGGGCGGGCGGGGAGUCGCGUGCGACGCGUGCCUUCGCCCCGUGCCCCGC
UCCGCCGCGCGCCUCGCGCGCCGCCCGCCCCGGCUCUGACUGACCGCGUUAUCUCCACA
GGUGAGCGGGCGGGACGGCCCUUCUCCUCCGGGCUGUAAUUAAGCUGAGCAAGAGG
UAAGGGUUUAAGGGAUGGUUGGUUGGUGGGGUAAUUAUGUUUAUUAUACCUGGAG
CACCUGCCUGAAAUCACUUUUUUUCAGGUUGGCGUACGGCCACCAUGGUGAGCAA
GGGCGAGGAGGAUAAUAUGGCCAUCAUCAAGGAGUUAUGCGCUUCAAGGUGCAC
AUGGAGGGCUCGUGAACGGCCACGAGUUCGAGAUCGAGGGGCGAGGGGCGAGGGCC
GCCCCUACGAGGGCACCCAGACCGCCAAGCUGAAGGUGACCAAGGGUGGCCCCCUG
CCCUUCGCCUGGGACAUCUCCUGUCCCCUCAGUUCAUGUACGGCUCCAAGGCCUACG
UGAAGCACCCCGCCGACAUCCCCGACUACUUGAAGCUGUCCUUCCCCGAGGGCUU
CAAGUGGGGAGCGCGUGAUGAACUUCGAGGACGGCGGGCGUGGUGACCGUGACCCAG
GACUCCUCCUGCAGGACGGCGAGUUAUCUACAAGGUGAAGCUGCGCGGCACCA
ACUUCCCCUCCGACGGCCCCGUAAUGCAGAAGAAAACCAUGGGCUGGGAGGGCCUC
CUCCGAGCGGAUGUACCCCGAGGACGGCGCCCUGAAGGGGCGAGAUCAAGCAGAGG
CUGAAGCUGAAGGACGGCGGCCACUACGACGCGUGAGGUCAAGACCACCUACAAGG
CCAAGAAGCCCGUGCAGCUGCCCGGCGCCUACAACGUACAACAUAAGUUGGACAU
CACCUCACACAACGAGGACUACACCAUCGUGGGAACAGUACGAACGCGCCGAGGGCC
GCCACUCCACCGGCGGGCAUGGACGAGCUGUACAAGUAGAAGCUUGAUCCAGACAU
GAUAAGAUACAUAUGAUGAGUUUGGACAAACCACAACUAGAAUGCAGUGAAAAAAA
UGC UUUAUUUGUGAAAAUUUGUGAUGCUAUUGCUUUUAUUUGUAACCAUUAUAAGC
UGCAAUAAACAAGUUGUGGAAUUGUGAGCGCUCACAAUCCACAGCGGCCGCUGA
GGUAGAUAAUCUUUGUAUAGUGGGGGGGGGAUCUCAUGUACCGGGUUUCUUUUUAU
UUGAUUUUCAAUAAAACAGACGGUAGCUAGGUUCGCAAGGCAGCCACAAGCCAAA
GAUAGGUAGGGUGCUCAUAGUGAGUAGGGACAGUGCCUUUUGAUUCACAACGCGU
CAAUACCAUCUGACACGGAUACCCUUAACCGGACUUGUCAUGAUCUCCAGACUUG
UCCAAGGUGGACGGGCCACCUUUAACUUAACCCGGAAAAGGAACAUAUAUUAUUA
UAUGUGUUCGGAAAAUAGCAAAAAAAAAAAAAAAAAAAAAA

531. TCARZ_CMV*_mCherry_GeFo3

ggguaaacGGCGGGGAGUAAUAUGACUCUCUUAAGGAGUCAUAGAGCCAGAACCUC
CUCGUGGUCCCCGCUGGGGCACAGGGAUUAUUUUUCUGUGGCCAAAUUUGACUGGCU
CUAGACGUUACAUAACUUAACGGUAAAUGGCCCGCCUGGCUGACCGCCCAACGACC
CCCGCCCAUUGACGUCAAUAUGACGUUAUGUUCUCCAUGUAACGCCAAUAGGGAC
UUUCCAUGACGUCAAUGGGUGGAGUAUUUACGGUAAACUGCCCACUUGGCAGUA

10

20

30

40

50

CAUCAAGUGUAUCAUAUGCCAAGUACGCCCCCUAUUGACGUCAAUGACGGUAAAU
GGCCCGCCUGGCAUUAUGCCCAGUACAUGACCUUAUGGGACUUUCCUACUUGGCA
GUACAUCUACGUAAUAGUCAUCGCUAAUACCAUGGUGAUGCGGUUUUGGCAGUAC
AUCAAUGGGGCGUGGAUAGGGGUUUGACUCACGGGGAUUUCCAAGUCUCCACCCCA
UUGACGUCAAUGGGAGUUUGUUUUGGCACCAAAAUCAACGGGACUUUCCAAAAUG
UCGUAAACAACUCCGCCCCAUUGACGCAAAUGGGCGGUAGGCGUGUACGGUGGGAG
GUCUAUAUAAGCAGAGCUCGUUUAGUGAACCGCUUGUCUCGCUCCGGGGAACGCU
CGGAAACUCCCGGCCGCCGCCACCCGCGUCUGUUCUGUUACACAAGGGAAGAAAA
GCCGCUUGCCGCACUCCGAGUGUCGUACGGCCACCAUGGUGAGCAAAGGCGAAGAA
GAUAACAUGGCGAUUAUUAAGAAUUAUGCGCUUUAAGUGCAUAUGGAAGGCA
GCGUGAACGGCCAUGAAUUAUGAAAUUGAAGGCGAAGGCGAAGGCCGCCCGUAUGA
AGGCACCCAGACCGCGAAACUGAAAGUGACCAAAGGCGGCCCGCUGCCGUUUGCG
UGGGAUAAUUCUGAGCCCGCAGUUUAUGUAUGGCAGCAAAGCGUAUGUGAAACAUC
CGGCGGAUAUUCGGAUUAUCUGAAACUGAGCUUUCGGAAAGGCUUUAUAUGGGA
ACGCGUGAUGAACUUUGAAGAUGGCGGCGUGGUGACCGUGACCCAGGAUAGCAGC
CUGCAGGAUGGCGAAUUAUUAUAAGUGAAACUGCGCGGCACCAACUUUCCGA
GCGAUGGCCCGGUGAUGCAGAAAAAAACCAUGGGCUGGGAAGCGAGCAGCGAACG
CAUGUAUCCGGAAGAUGGCGCGCUGAAAGGCGAAAUUAACAGCGCCUGAAACUG
AAAGAUGGCGGCCAUUAUGAUGCGGAAGUGAAAACCAACCUAUAAGCGAAAAAAC
CGGUGCAGCUGCCGGGCGCGUAUAACGUGAACAUUAACUGGAUAUUACCAGCCA
UAACGAAGAUUAUACCAUUGUGGAACAGUAUGAACGCGCGGAAGGCCGCCAUAGC
ACCGGCGGCAUGGAUGAACUGUAUAAAUAGUGAAAGCUUAAUAAAAGAUCUUUAU
GUUCAUUAAGAUCUGUGUGUUUGGUUUUUGUGUGGCGGCCGCUAGGGUAGAUAAUC
UUUGUAUAGUGGGGGGGGAUCUCAUGUACCGGGUUCUUUUAUUUGAUUUUCAA
UAAAACAGACGGUAGCUAGGUUCGCAAGGCAGCCACAAGCCAAAGAUAGGUAGGG
UGCUCUAUAGUGAGUAGGGACAGUGCCUUUUGAUUACACAACGCGUCAAUACCAUCU
GACACGGAUACCCUUACCGGACUUGUCAUGAUCUCCAGACUUGUCCAAGGUGGA
CGGGCCACCUUUACUUAACCCGGAAAAGGAACAUAUAUUAUUAUUGUGUUCGG
AAAAUAGCAAAAAAAAAAAAAAAAAAAAAAA

10

20

30

40

50

532. TCARZ_SV40*_mCherry_GeFo3

ggguaaacGGCGGGAGUAACUAUGACUCUCUUAAGGAGUCAUAGAGCCAGAACCUC
CUCGUGGUCCCGCUGGGCACAGGGAUUAUUUUUCUGUGGCAAAUUGACUGGCU
CUAGAGCUGUGGAAUGUGUGUCAGUUAGGGUGUGGAAAGUCCCCAGGCUCCCCAG
CAGGCAGAAGUAUGCAAAGCAUGCAUCUCAAUUAUGUCAGCAACCAGGUGUGGAAA
GUCCCCAGGCUCCCCAGCAGGCAGAAGUAUGCAAAGCAUGCAUCUCAAUUAUGUCA
GCAACCAUAGUCCCGCCCCUAACUCCGCCCAUCCCGCCCCUAACUCCGCCCAGUUC
CGCCCAUUCUCCGCCCAUGGCUGACUAAUUGCGUCUAUAUAUGCAGAGGCCGAG
GCCGCCUCGGCCUCUGAGCUAUUCCAGAAGUAGUGAGGAGGCCGUACGGCCACCA
UGGUGAGCAAAGGCGAAGAAGAUAAUAUGGCGAUUAUUAAGAAUUAUGCGCUU
UAAAGUGCAUAUGGAAGGCAGCGUGAACGGCCAUGAAUUAUGAAAUUGAAGGCGAA
GGCGAAGGCCGCCGUAUGAAGGCACCCAGACCGCGAAACUGAAAGUGACCAAAG
GCGGCCCGCUGCCGUUUGCGUGGGGAUAUUCUGAGCCCGCAGUUUAUGUAUGGCAG
CAAAGCGUAUGUGAAACAUCGCGGGAUAUUCGGAUUAUCUGAAACUGAGCUUU
CCGGAAGGCUUUAUAUGGGAACGCGUGAUGAACUUUGAAGAUGGCGGCGUGGUGA
CCGUGACCCAGGAUAGCAGCCUGCAGGAUGGCGAAUUAUUAUUAUAAAGUGAAACU
GCGCGGCACCAACUUUCCGAGCGAUGGCCCGGUGAUGCAGAAAAAAACCAUGGGC
UGGGAAGCGAGCAGCGAACGCAUGUAUCCGGAAGAUGGCGCGCUGAAAGGCGAAA
UUAACAGCGCCUGAAACUGAAAGAUGGCGGCCAUUAUGAUGCGGAAGUGAAAC
CACCUAUAAGCGAAAAAACCGGUGCAGCUGCCGGGCGCGUAUAACGUGAACAUU
AAACUGGAUAUUACCAGCCAUAACGAAGAUUAUACCAUUGUGGAACAGUAUGAAC

GCGCGGAAGGCCGCCAUAGCACCGGCGGCAUGGAUGAACUGUAUAAAUAAGUGAAA
GCUUAAUAAAAGAUCUUUAUGUUCAUUAGAUCUGUGUGUUGGUUUUUGUGUGGC
GGCCGCUAGGGUAGAUAAUCUUUGUAUAGUGGGGGGGGAUCUCAUGUACCGGGUU
UCUUUUUAUUUGAUUUUCAAUAAAACAGACGGUAGCUAGGUUCGCAAGGCAGCCAC
AAGCCAAAGAUAGGUAGGGUGCUCAUAGUGAGUAGGGACAGUGCCUUUUGAUUCA
CAACGCGUCAAUACCAUCUGACACGGAUACCCUUAACCGGACUUGUCAUGAUCUCC
CAGACUUGUCCAAGGUGGACGGGGCCACCUUUAACUUAACCCGGAAAAGGAACAUAU
AUUAAUUAUAUGUGUUCGGAAAAUAGCAAAAAAAAAAAAAAAAAAAAAA

533. TCARZ_SV40*_GFP_minPA_CMV*_mCherry_SV40LPA_GeFo3

ggguaaacGGCGGGAGUAACUAUGACUCUCUUAAGGAGUCAUAGAGCCAGAACCUC
CUCGUGGUCCCCGUGGGGCACAGGGAUUAUUUUUCUGUGGCAAAUUUGACUGGCU
CUAGACUGUGGGAUGUGUGUCAGUUAGGGUGUGGAAAGUCCCCAGGCUCCCCAGC
AGGCAGAAGUAUGCAAAGCAUGCAUCUCAAUUAGUCAGCAACCAGGUGUGGAAAG
UCCCCAGGCUCCCCAGCAGGCAGAAGUAUGCAAAGCAUGCAUCUCAAUUAGUCAG
CAACCAUAGUCCCGCCCCUAACUCCGCCCAUCCCGCCCCUAACUCCGCCCAGUUC
GCCCAUUCUCCGCCCCAUGGCUGACUAAUUGCGUCUAUAUAUGCAGAGGCCGAGG
CCGCCUCGGCCUCUGAGCUAUUCCAGAAGUAGUGAGGAGGCCGUACGGCCACCAU
GGUGAGCAAGGGCGAGGAGCUGUUCACCGGGGUGGUGCCCAUCCUGGUCGAGCUG
GACGGCGACGUAAACGGCCACAAGUUCAGCGUGUCCGGCGAGGGGCGAGGGCGAUG
CCACCUACGGCAAGCUGACCCUGAAGUUCAUCUGCACCACCGGCAAGCUGCCCGU
GCCUGGCCCAACCUCGUGACCACCCUGACCUACGGCGUGCAGUGCUUCAGCCGC
UACCCCGACCACAUGAAGCAGCACGACUUCUUAAGUCCGCCAUGCCCGAAGGCU
ACGUCCAGGAGCGCACCAUCUUCUUAAGGACGACGGCAACUACAAGACCCGCGC
CGAGGUGAAGUUCGAGGGCGACACCCUGGUGAACCGCCAUCGAGCUGAAGGGCAUC
GACUUCAAGGAGGACGGCAACAUCUGGGGCACAAGCUGGAGUACAACUACAACA
GCCACAACGUCUAUAUCAUGGCCGACAAGCAGAAGAACGGCAUCAAGGUGAACUU
CAAGAUCCGCCACAACAUCGAGGACGGCAGCGUGCAGCUCGCCGACCACUACCAGC
AGAACACCCCCAUCGGCGACGGCCCCGUGCUGCUGCCCGACAACCACUACCUGAGC
ACCCAGUCCGCCCUGAGCAAAGACCCCAACGAGAAGCGCGAUCACAUGGUCCUGC
UGGAGUUCGUGACCGCCGCGGGGAUCACUCUCGGCAUGGACGAGCUGUACAAGUA
AAAUAAAAGAUCUUUAUGUUCAUUAGAUCUGUGUGUUGGUUUUUGUGUGGGCGUU
ACAUAAUUAACGGUAAAUGGCCCGCCUGGCUGACCGCCCAACGACCCCCGCCCAU
UGACGUCAAUAUAGACGUUAGUUCUCCAUAGUAACGCCAAUAGGGACUUUCCAUUG
ACGUCAAUGGGUGGAGUAUUUACGGUAAACUGCCCACUUGGCAGUACAUAAGUG
UAUCAUAUGCCAAGUACGCCCCCUAUUGACGUCAAUGACGGUAAAUGGCCCGCCU
GGCAUUAUGCCCAGUACAUGACCUUAUGGGACUUAUCCUACUUGGCAGUACAUCUA
CGUAUUAGUCAUCGCUAUUACCAUGGUGAUGCGGUUUUGGCAGUACAUAUUGGG
CGUGGAUAGGGGUUUGACUCACGGGGAUUCCAAGUCUCCACCCCAUUGACGUCA
AUGGGAGUUUGUUUUGGCACCAAAAUCAACGGGACUUAUCCAAAUGUCGUAAACAA
CUCCGCCCAUUGACGCAAUUGGGCGGUAGGCGUGUACGGUGGGAGGUCUAUAUA
AGCAGAGCUCGUUUAGUGAACCVCUUGUCUCGCUCCGGGGAAACGUCGGAAACUC
CCGGCCGCGCCACCCGCGUCUGUUCUGUUAACAAAGGGAAGAAAAGCCGCUGCC
GCACUCCGAGUGUCGUACGGCCACCAUGGUGAGCAAAGGCGAAGAAGUAACAUG
GCGAUUAUUAAAGAAUUUAUGCGCUUUAAGUGCAUAUGGAAGGCAGCGUGAACG
GCCAUGAAUUUGAAAUUGAAGGCGAAGGCGAAGGCCGCCCGUAUGAAGGCACCCA
GACCGCGAAACUGAAAGUGACCAAAGGCGGCCCGCUGCCGUUUGCGUGGGGAUAUU
CUGAGCCCGCAGUUUAUGUAUGGCAGCAAAGCGUAUGUGAAACAUCGGCGGGAUA
UUCCGGAUUAUCUGAAACUGAGCUUUCGGGAAGGCUUUAUAAUGGGAACGCGUGAU
GAACUUUGAAGAUGGCGGCGUGGUGACCGUGACCCAGGAUAGCAGCCUGCAGGAU
GGCGAAUUUAUUUAUAAAGUGAAACUGCGCGGCACCAACUUCGAGCGAUGGCC

10

20

30

40

50

CGGUGAUGCAGAAAAAAACCAUGGGCUGGGAAGCGAGCAGCGAACGCAUGUAUCC
GGAAGAUGGCGCGCUGAAAGGCGAAAUUAAACAGCGCCUGAAACUGAAAGAUGGC
GGCCAUAUGAUGCAGGAAGUGAAAACCAACCUAUAAGCGAAAAAACGGUGCAGC
UGCCGGGCGCGUAUAACGUGAACAUUAAACUGGAUAUUACCAGCCAUAACGAAGA
UUAUACCAUUGUGGAACAGUAUGAACGCGCGGGAAGGCCGCCAUAGCACCGGCGGC
AUGGAUGAACUGUAUAAAUAGUGAAAGCUUAAUAAAAGAUUUUAUGUUAUA
GAUCUGUGUGUUGGUUUUUGUGUGGCGGCCGCGGUAGAUAAUCUUGUAUAGUG
GGGGGGGAUCUCAUGUACCGGGUUCUUAUUAUUGAUUUUCAUAAAACAGACG
GUAGCUAGGUUCGCAAGGCAGCCACAAGCCAAAGAUAGGUAGGGUGCUCAUAGUG
AGUAGGGACAGUGCCUUUGAUUCACAACGCGUCAAUACCAUCUGACACGGAUAC
CCUUAACGGACUUGUCAUGAUCUCCAGACUUGUCCAAGGUGGACGGGCCACCUU
UACUUAACCCGGAAGGAACAUAUAUUAUUAUUGUGUUCGGAAAAUAGCAA
AAAAAAAAAAAAAAAAAAAA

10

534. TCARZ_CBhBsi_GFP_GeFo217_R4A22

gggagacgggucaaccgcuaggagcgguagaccgGAAUUCGGCGGGAGUAACUAUGACUC
UCUUAAGGAGUCAUAGAGCCAGAACCUCUCGUGGUCCCGCUGGGCACAGGGAUU
AAUUUUUCUGUGGCAAAUUGACUGGCUCUAGAAGGUCGACCAGAUGUCCGAGGU
CGACCAGUUGUCCGUGUGGAAUUGUGAGCGCUCACAAUCCACACGUUACAUAAC
UUACGGUAAAUGGCCCGCCUGGCUGACCGCCCAACGACCCCCGCCCAUUGACGUC
AAUAGUAACGCCAAUAGGGACUUUCCAUUGACGUCAAUGGGUGGAGUAUUUACGG
UAAACUGCCCACUUGGCAGUACAUAAGUGUAUCAUAUGCCAAGUACGCCCCCUA
UUGACGUCAAUGACGGUAAAUGGCCCGCCUGGCAUUGUGCCCAGUACAUGACCUU
AUGGGACUUUCCUACUUGGCAGUACAUCUACGUAAUAGUCAUCGCUAUUACCAUG
GUCGAGGUGAGCCCCACGUUCUGCUUCACUCUCCCCAUCUCCCCCCCCUCCCCAC
CCCCAAUUUUUGUAUUUAUUUAUUUUUAAUUAUUUUUGUGCAGCGAUGGGGGCGG
GGGGGGGGGGGGGGCGCGCGCCAGGCGGGGCGGGGCGGGGCGAGGGGCGGGGCGG
GGCGAGGCGGAGAGGUGCGGCGGCAGCCAAUCAGAGCGGCGCGCUCCGAAAGUUU
CCUUUUUAUGGCGAGGCGGCGGCGGCGGCGGCCCUAUA AAAAGCGAAGCGCGCGG
GGGCGGGAGUCGCGUGCGACGCGUGCCUUCGCCCCGUGCCCCGCUCCGCCGCCGCCUC
GCGCCGCCCGCCCCGGCUCUGACUGACCGCGUUAUCUCCACAGGUGAGCGGGCGG
GACGGCCCUUCUCCUCCGGGCGUGAAUUAAGCUGAGCAAGAGGUAAAGGUUUAAAG
GAUGGUUGGUUGGUGGGGUAUUAAUGUUUAAUUAACUGGAGCACCUGCCUGAAA
UCACUUUUUUUCAGGUUGGCGUACGGCCACCAUGGUGAGCAAGGGCGAGGAGCUG
UUCACCGGGGUGGUGCCCAUCCUGGUCGAGCUGGACGGCGACGUAAACGGCCACA
AGUUCAGCGUGUCCGGCGAGGGCGAGGGCGAUGCCACCUACGGCAAGCUGACCCU
GAAGUUAUCUGCACCAACGGCAAGCUGCCCGUGCCCGUGGCCACCCUCGUGACC
ACCCUGACCUACGGCGUGCAGUGCUUCAGCCGCUACCCCGACCAUAUGAAGCAGC
ACGACUUCUUAAGUCCGCCAUGCCCGAAGGCUACGUCCAGGAGCGCACCAUCUU
CUUCAAGGACGACGGCAACUACAAGACCCGCGCCGAGGUGAAGUUCGAGGGCGAC
ACCCUGGUGAACC GCAUCGAGCUGAAGGGCAUCGACUUAAGGAGGACGGCAACA
UCCUGGGGCAACAAGCUGGAGUACAACUACAACAGCCACAACGUCUAUAUCAUGGC
CGACAAGCAGAAGAAGCGCAUCAAGGUGAACUUAAGAUCGCCACAACAUCGAG
GACGGCAGCGUGCAGCUCGCCGACCAUACCAGCAGAACACCCCCAUUCGGCGACGG
CCCCGUGCUGCUGCCCGACAACCACUACCUGAGCACCCAGUCCGCCUGAGCAAAG
ACCCCAACGAGAAGCGCGAUCACAUGGUCCUGCUGGAGUUCGUGACCGCCGCCGG
GAUCACUCUCGGCAUGGACGAGCUGUACAAGUAAAAGCUUGAUCCAGACAUGAUA
AGAUACAUAUGAUGAGUUUGGACAAACCACAACUAGAAUGCAGUGAAAAAAUUGCU
UUAUUUGUGAAAUAUUGUGAUGCUAUUGCUUUAUUUGUAACCAUUAUAAGCUGCA
AUAAACAAGUUGUGGAAUUGUGAGCGCUCACAAUUCACAGCGGCCGCGUGAGACG
GUAGCUAGGUUCGCAAGGCAGCCACAAGCCAAAGAUAGGUAGGGUGCUCAUAGUG

20

30

40

50

AGUAGGGACAGUGCCUUUUGAUUCACAACGCGUCAAUACCAUCUGACACGGAUAC
CCUUACCGGACUUGUCAUGAUCUCCCAGACUUGUCCAAGGUGGACGGGCCACCUU
UACUUAACCCGGAAAAGGAACAUAUAUAAUUAUAUGUGUUCGGAAAAUAGCAAA
AAAAAAAAAAAAAAAAAAAA

535. TCARZ_CBhBsi_GFP_GeFo98_R4A22

gggagacgggucaaccgcguaggagcggugaccgGAAUUCGGCGGGAGUAACUAUGACUC
UCUUAAGGAGUCAUAGAGCCAGAACCUCUCGUGGUCCCGCUGGGCACAGGGGAUU
AAUUUUUCUGUGGC AAAUUUGACUGGCUCUAGAAGGUCGACCAGAUGUCCGAGGU
CGACCAGUUGUCCGUGUGGAAUUGUGAGCGCUCACAAUUCCACACGUUACAUAAC
UUACGGUAAAUGGCCCGCCUGGCUGACCGCCCAACGACCCCCGCCCAUUGACGUC
AAUAGUAACGCCAAUAGGGACUUUCCAUUGACGUCAAUGGGUGGAGUAUUUACGG
UAAACUGCCCAUUGGCAGUACAUAAGUGUAUCAUAUGCCAAGUACGCCCCCUA
UUGACGUCAAUGACGGUAAAUGGCCCGCCUGGCAUUGUGCCCAGUACAUGACCUU
AUGGGACUUUCCUACUUGGCAGUACAUCUACGUAAUAGUCAUCGCUAAUACCAUG
GUCGAGGUGAGCCCCACGUUCUGCUUCACUCUCCCCAUCUCCCCCCCCUCCCCAC
CCCCAAUUUUUGUAUUUAUUUAUUUUUAAUUAUUUUGUGCAGCGAUGGGGGCGG
GGGGGGGGGGGGGGCGCGCGCCAGGCGGGGCGGGGCGGGGCGAGGGGCGGGGCGG
GGCGAGGCGGAGAGGUGCGGCGGCAGCCAAUCAGAGCGGCGCGCUCCGAAAGUUU
CCUUUUUAUGGCGAGGCGGCGGCGGCGGCGGCCCUAUA AAAAGCGAAGCGCGCGG
GGGCGGGAGUCGCGUGCGACGCGUGCCUUCGCCCCGUGCCCCGCUCCGCCGCCGCU
GCGCCGCCCGCCCCGGCUCUGACUGACCGCGUUAUCUCCACAGGUGAGCGGGCGG
GACGGCCCUUCUCCUCCGGGCUGUAAUAGCUGAGCAAGAGGUAAGGGUUUAAGG
GAUGGUUGGUUGGUGGGGUAAUUA AUGUUUA AUUACCUGGAGCACCUGCCUGAAA
UCACUUUUUUUCAGGUUGGCGUACGGCCACCAUGGUGAGCAAGGGCGAGGAGCUG
UUCACCGGGGUGGUGCCCAUCCUGGUCGAGCUGGACGGCGACGUAAACGGCCACA
AGUUCAGCGUGUCCGGCGAGGGCGAGGGCGAUGCCACCUACGGCAAGCUGACCCU
GAAGUUCAUCUGCACCACCGGCAAGCUGCCCGUGCCCUGGCCACCCUCGUGACC
ACCCUGACCUACGGCGUGCAGUGCUUCAGCCGCUACCCCGACCAUAUGAAGCAGC
ACGACUUCUUAAGUCCGCCAUGCCCGAAGGCUACGUCCAGGAGCGCACCAUCUU
CUUCAAGGACGACGGCAACUACAAGACCCGCGCCGAGGUGAAGUUCGAGGGCGAC
ACCCUGGUGAACC GCAUCGAGCUGAAGGGCAUCGACUUAAGGAGGACGGCAACA
UCCUGGGGCACAAGCUGGAGUACAACUACAACAGCCACAACGUCUAUAUCAUGGC
CGACAAGCAGAAGAACGGCAUCAAGGUGAACUUAAGA UCCGCCACAACAUCGAG
GACGGCAGCGUGCAGCUCGCCGACCACUACCAGCAGAACACCCCCAUUGGCGACGG
CCCCGUGCUGCUGCCCGACAACCACUACCUGAGCACCCAGUCCGCCUGAGCAAAG
ACCCCAACGAGAAGCGCGAUCACAUGGUCCUGCUGGAGUUCGUGACCGCCGCCGG
GAUCACUCUCGGCAUGGACGAGCUGUACAAGUAAAAGCUUGAUCCAGACAUGAUA
AGAUACAUAUGAUGAUUUGGACAAACCACAACUAGAAUGCAGUGAAAAAAUAGCU
UUAUUUGUGAAA AUUGUGAUGCUAUUGCUUUAUUUGUAACCAUUAUAAGCUGCA
AUAAACAAGUUGUGGAAUUGUGAGCGCUCACAAUUCCACAGCGGCCGCUGACCGG
ACUUGUCAUGAUCUCCCAGACUUGUCCAAGGUGGACGGGCCACCUUUA CUUAACC
CGGAAAAGGAACAUAUAUUAUUAUUAUGUGUUCGGAAAAUAGCAAAAAAAAAA
AAAAAAAAAA

536. TCARZ_CBhBsi_GFP_GeFo68_R4A22

gggagacgggucaaccgcguaggagcggugaccgGAAUUCGGCGGGAGUAACUAUGACUC
UCUUAAGGAGUCAUAGAGCCAGAACCUCUCGUGGUCCCGCUGGGCACAGGGGAUU
AAUUUUUCUGUGGC AAAUUUGACUGGCUCUAGAAGGUCGACCAGAUGUCCGAGGU
CGACCAGUUGUCCGUGUGGAAUUGUGAGCGCUCACAAUUCCACACGUUACAUAAC
UUACGGUAAAUGGCCCGCCUGGCUGACCGCCCAACGACCCCCGCCCAUUGACGUC
AAUAGUAACGCCAAUAGGGACUUUCCAUUGACGUCAAUGGGUGGAGUAUUUACGG

UAAACUGCCCACUUGGCAGUACAUCAAGUGUAUCAUAUGCCAAGUACGCCCCCUA
UUGACGUCAAUGACGGUAAAUGGCCCGCCUGGCAUUGUGCCCAGUACAUGACCUU
AUGGGACUUUCCUACUUGGCAGUACAUCUACGUAAUAGUCAUCGCUAAUACCAUG
GUCGAGGUGAGCCCCACGUUCUGCUUCACUCUCCCCAUCUCCCCCCCCUCCCCAC
CCCCAAUUUUUGUAUUUAUUUAUUUUUAAUUAUUUUUGUGCAGCGAUGGGGGGCGG
GGGGGGGGGGGGGGGGGGCGCGCGCCAGGCGGGGGCGGGGGCGAGGGGGCGGGGGCGG
GGCGAGGCGGAGAGGUGCGGGCGGCAGCCAAUCAGAGCGGGCGCGCUCCGAAAGUUU
CCUUUUUAUGGCGAGGCGGCGGGCGGGCGGGCCCUAUA AAAAGCGAAGCGCGCGGC
GGGCGGGAGUCGCGUGCGACGCGUGCCUUCGCCCCGUGCCCCGCUCCGCCGCCGCCUC
GCGCCGCCCGCCCCGGCUCUGACUGACCGCGUUACUCCCACAGGUGAGCGGGGCGG
GACGGCCCUUCUCCUCCGGGCGUGAAUUAGCUGAGCAAGAGGUAAGGGUUUAAGG
GAUGGUUGGUUGGUGGGGUAAUUA AUGUUAAUUAUACCUGGAGCACCUGCCUGAAA
UCACUUUUUUUCAGGUUGGCGUACGGCCACCAUGGUGAGCAAGGGCGAGGAGCUG
UUCACCGGGGUGGUGCCCAUCCUGGUCGAGCUGGACGGCGACGUAAACGGCCACA
AGUUCAGCGUGUCCGGCGAGGGCGAGGGCGAUGCCACCUACGGCAAGCUGACCCU
GAAGUUCAUCUGCACCACCGGCAAGCUGCCCGUGCCUGGGCCACCCUCGUGACC
ACCCUGACCUACGGCGUGCAGUGCUUCAGCCGCUACCCCGACCACAUGAAGCAGC
ACGACUUCUUAAGUCCGCCAUGCCC GAAGGCUACGUCCAGGAGCGCACCAUCUU
CUUCAAGGACGACGGCAACUACAAGACCCGCGCCGAGGUGAAGUUCGAGGGGCGAC
ACCCUGGUGAACC GCAUCGAGCUGAAGGGCAUCGACUUAAGGAGGACGGCAACA
UCCUGGGGGCACAAGCUGGAGUACAACUACAACAGCCACAACGUCUAUAUCAUGGC
CGACAAGCAGAAGAACGGCAUCAAGGUGAACUUAAGA UCCGCCACAACAUCGAG
GACGGCAGCGUGCAGCUCGCCGACCACUACCAGCAGAACACCCCCAUCGGCGACGG
CCCCGUGCUGCUGCCCGACAACCACUACCUGAGCACCCAGUCCGCCCCUGAGCAAAG
ACCCCAACGAGAAGCGCGAUCACAUGGUCCUGCUGGAGUUCGUGACCGCCGCCGG
GAUCACUCUCGGCAUGGACGAGCUGUACAAGUAAAAGCUUGAUCCAGACAUGAUA
AGAUACA UUGAUGAGUUUGGACAAACCACAACUAGAAUGCAGUGAAAAAAAUGCU
UUAUUUGUGAAA UUGUGAUGCUAUUGCUUUUAUUUGUAACCAUUAUAAGCUGCA
AUAACAAGUUGUGGAAUUGUGAGCGCUCACAAU UCCACAGCGGCCGCGUGACAAG
GUGGACGGGGCCACCUUUA CUUAACCCGGAAAAGGAACAUUAUAUUAUUAUGUG
UUCGGAAAAUAGCAAAAAAAAAAAAAAAAAAAAAA

10

20

30

537. AGGTCGACCAGATGTCCGAGGTGACCCAGTTGTCCG

538. TAGCaaaaaaaaaaaaaaaaaaaaa

539. CATCATGGATTAGGATCGGAAGACCCCCG

540. GTACGCCGGCGAAATTGGATCAGTAGATG

541. GAGAAACAGACGGGCCTGATCTACACCC)

542. CTATCTGAACCGAAGTTCCGCAACGCCTACGTAC

543. CACTGCGTGTGGTCAAGTTTTCCTAGCATGCACG

544. GATGTTATGCCAAGACAGCAAGCAAATGTTTTGAACCAAACG

545. TTGAGGCGAGTCACCACTCGCTTTCCGG

546. GTGTCCGTACGGGGACGACATCCGAGTG

547. CTCCTGACCAACTAGCTCACTGACTAATTTTAAAC

548. CCACTTATTCTACACCTCTCATGTCTTTCACCG

549. CTAGCAGCCGACTTAGAACTGGTGCGG

550. CTTGCTCTTCGGAATCCATGTCCATAGC

551. CCAGGGGAATCCGACTGTTTAATTAAAACAAAGC

552. GCGACTCGCATCACTGACTTTAATTGGTTG

553. CGCGCUUCGUAAUUCAGAAtt

554. UUCUGAAAUACGAAGCGCGtg

555. UGACCUCUCCAAACCCUCUtt

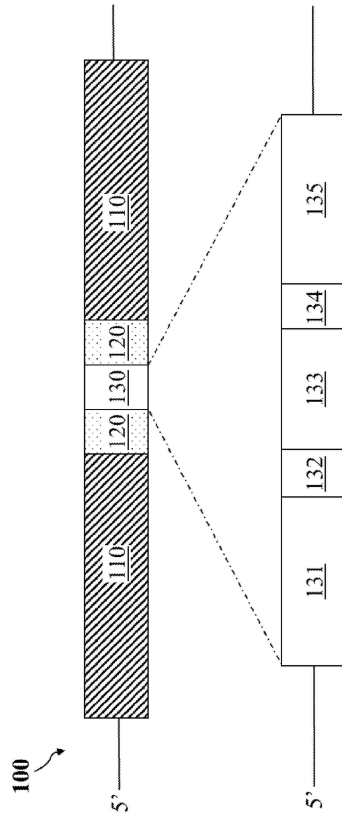
40

50

556. AGAGGGUUUGGAGAGGUCAtg
 557. GGGAGCACCUCAAUCCUAAtt
 558. UUAGGAUUCAGGUGCUCCCgg
 559. GGAUAUUACUUUCGUGUAAtt
 560. UUACACGAAAGUAAUAUCCaa
 561. CGUCGAUUCCCAGAUCUUAAtt
 562. UAAGAUCUGGGAAUCGACGaa
 563. GAAUCGCAAGGAUAUGAUAtt
 564. UAUCAUAUCCUUGCGAUUctc

【図面】

【図 1】



【図 2】



10

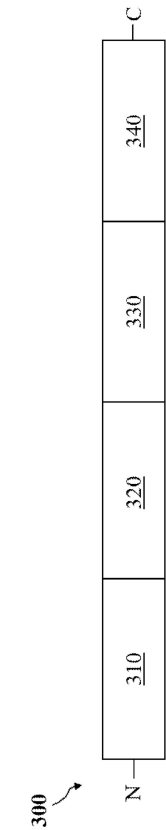
20

30

40

50

【図 3】



【図 4】

DNA結合ドメイン (DB) の供給源				RNA結合ドメイン (RB) の供給源				逆転写酵素ドメイン (RT) の供給源				エンドヌクレアーゼ ドメイン (EN) の供給源			
A1	A2	A3	B1	A1	A2	A3	B1	A1	A2	A3	B1	A1	A2	A3	B1
B2	B3	C1	C2	B2	B3	C1	C2	B2	B3	C1	C2	B2	B3	C1	C2
C3	C4	C5	D1	C3	C4	C5	D1	C3	C4	C5	D1	C3	C4	C5	D1
D2	E1	E2	E3	D2	E1	E2	E3	D2	E1	E2	E3	D2	E1	E2	E3
E4				E4				E4				E4			

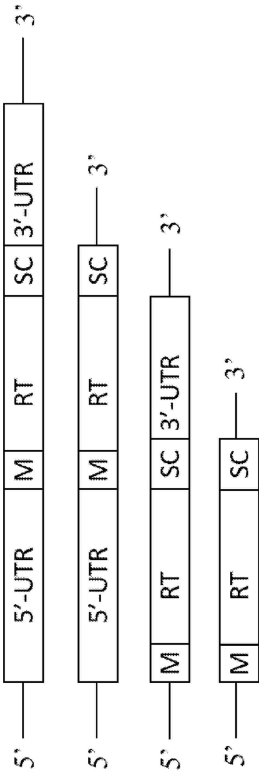
キメラRTタンパク質ドメインの例

DB	RB	RT	EN
A1	A1	A1	A1
A2	A2	A2	A2
B2	B2	B2	B2
A3	E1	A3	B2
E4	B3	C5	B3
D1	E4	B3	C5
A3	B2	C2	D2
C5	B2	E3	D2

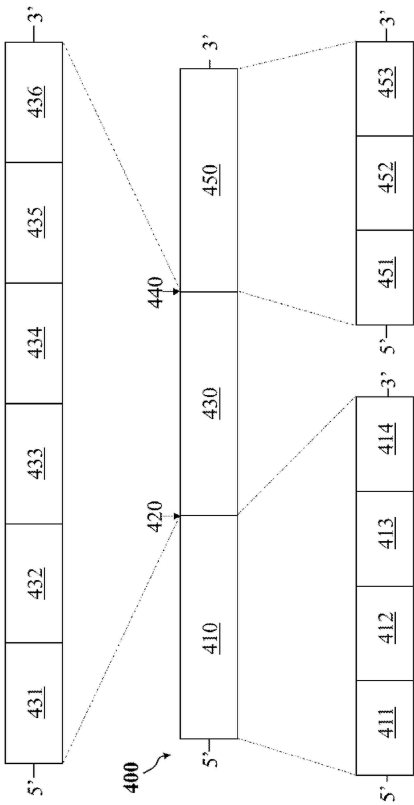
生物種の凡例

A1=Z. albicollis
A2=T. guttata
A3=T. guttatus
B1=P. pungitius
B2=O. latipes
B3=G. aculeatus
C1=N. vitripennis
C2=D. melanogaster
C3=T. castaneum
C4=B. mori
C5=D. simulans
D1=L. cousei
D2=T. cancriformis
E1=H. magnipapillata
E2=L. polyphemus
E3=A. vaga
E4=C. intestinalis

【図 5】



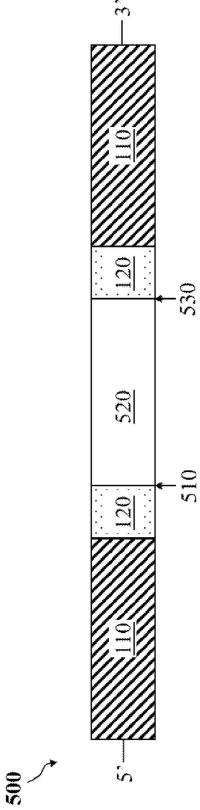
【図 6】



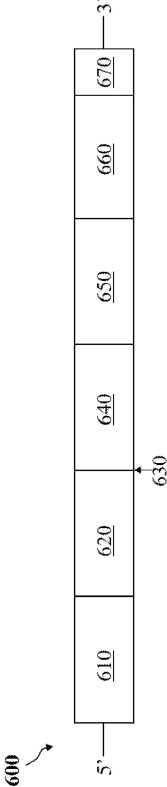
【図 7】



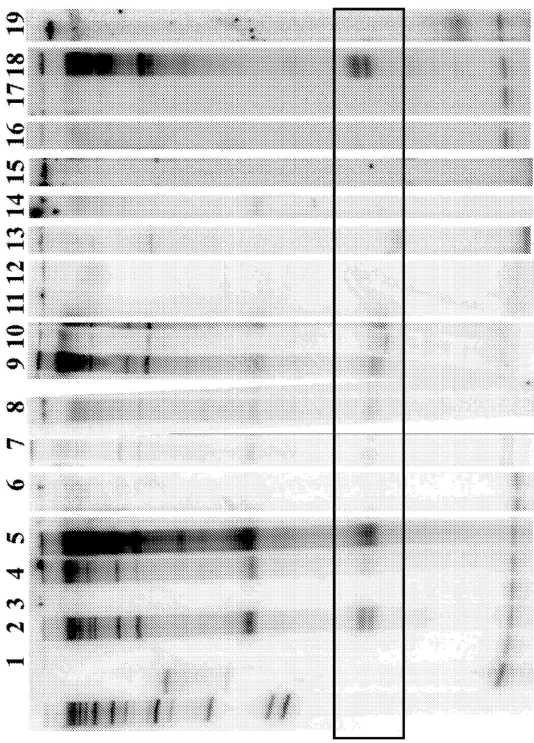
【図 8】



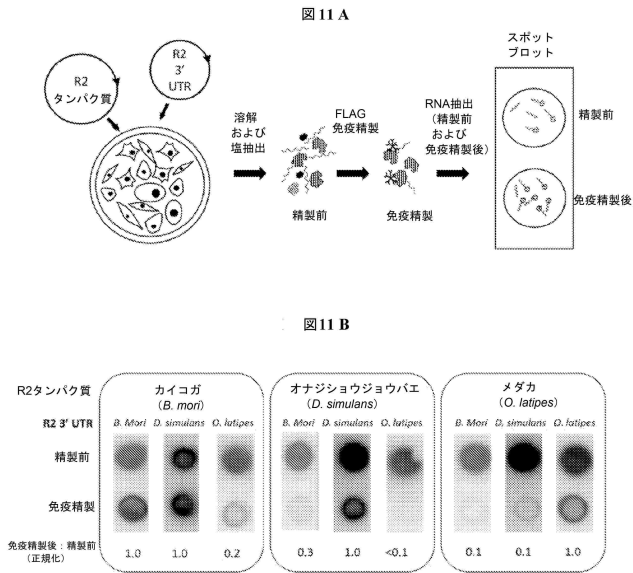
【図 9】



【図 10】



【図 1 1 A - B】



【図 1 2 A - B】

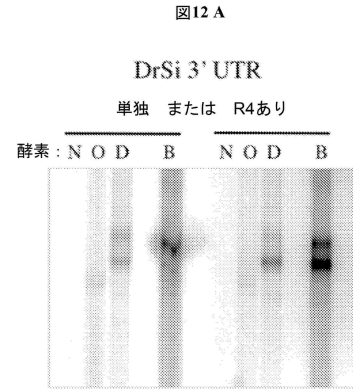
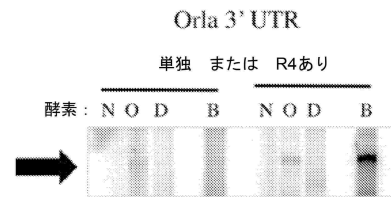
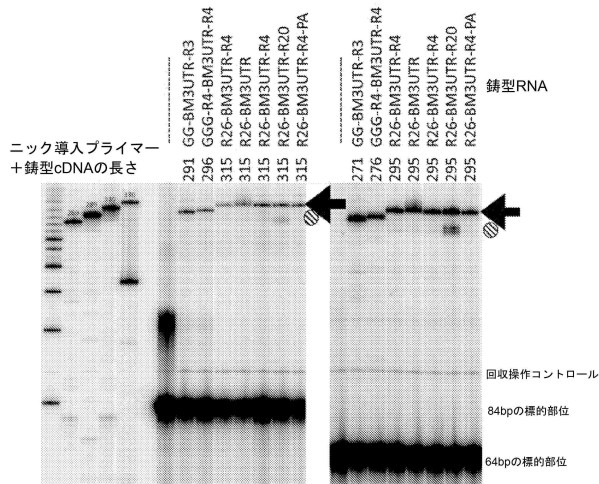


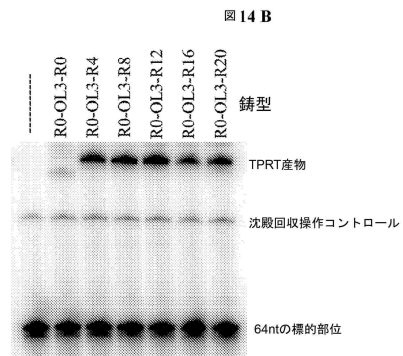
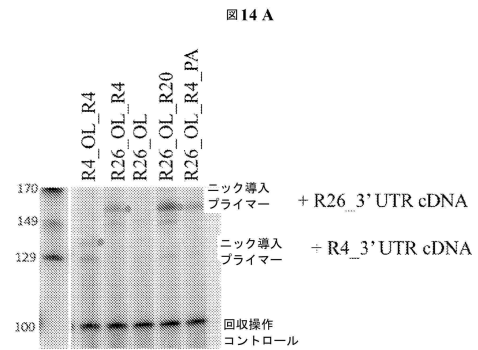
図 12 B



【図 1 3】



【図 1 4 A - B】



10

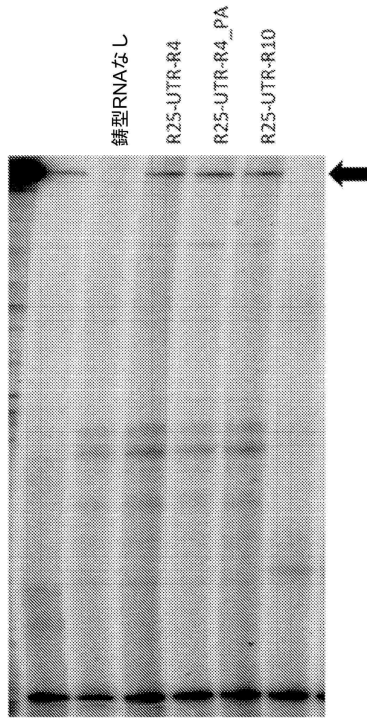
20

30

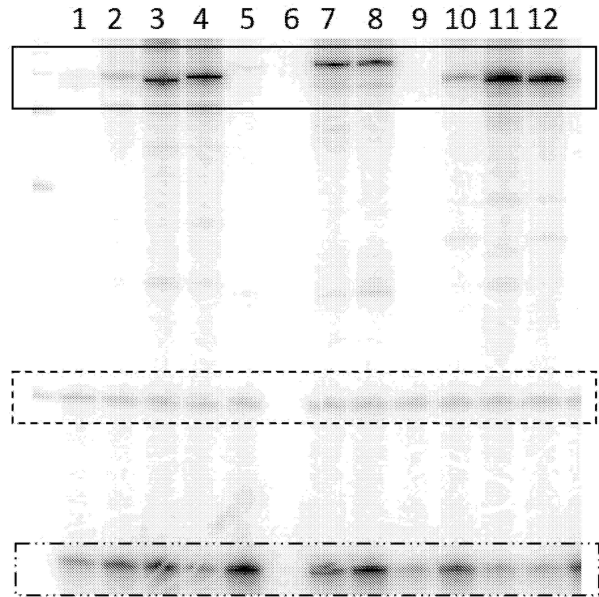
40

50

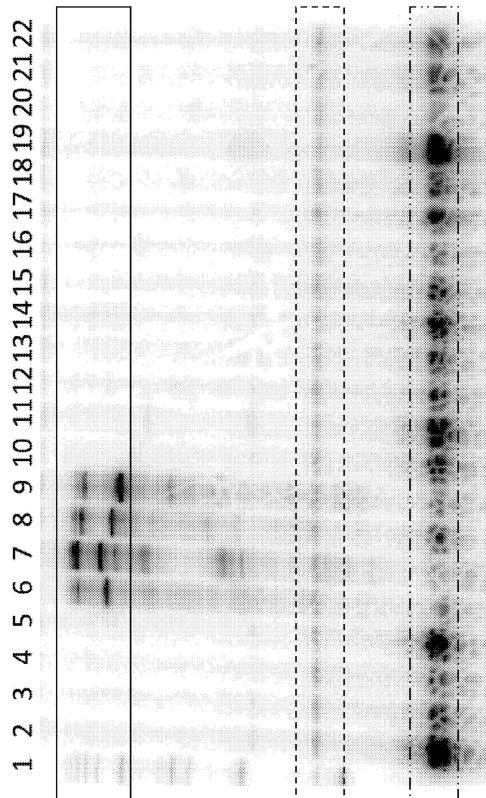
【図 15】



【図 16】



【図 17】



【図 18 A - B】

図 18 A

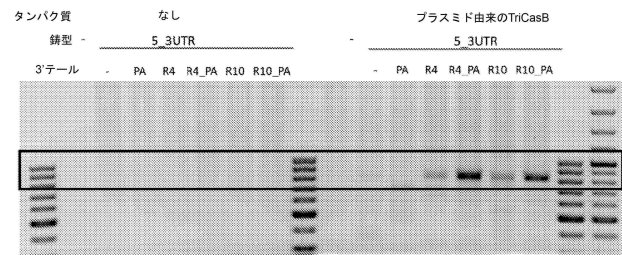
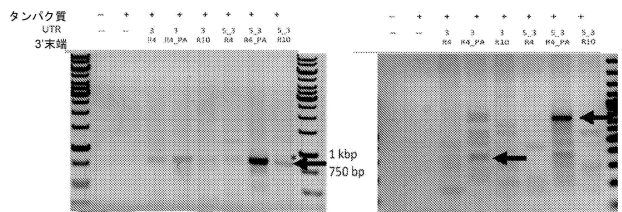


図 18 B



10

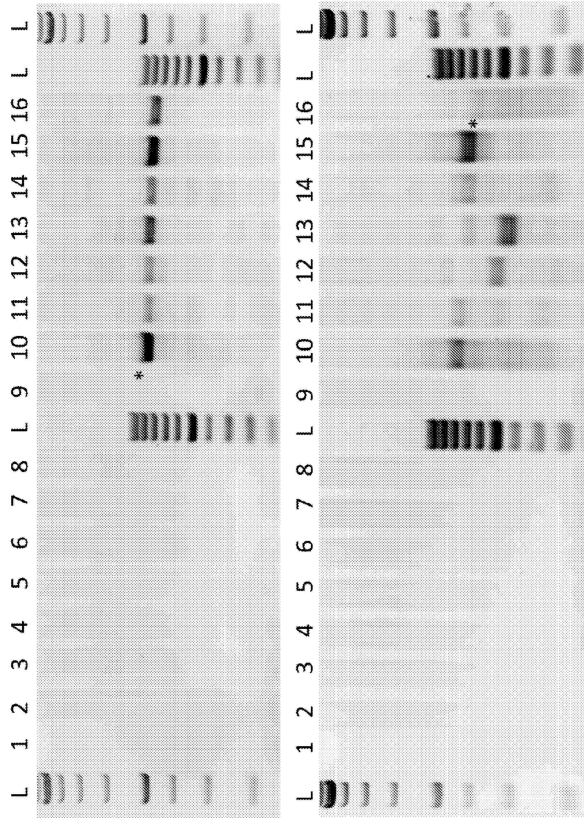
20

30

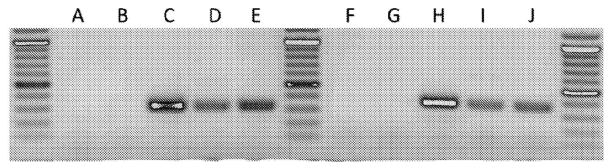
40

50

【図 19】



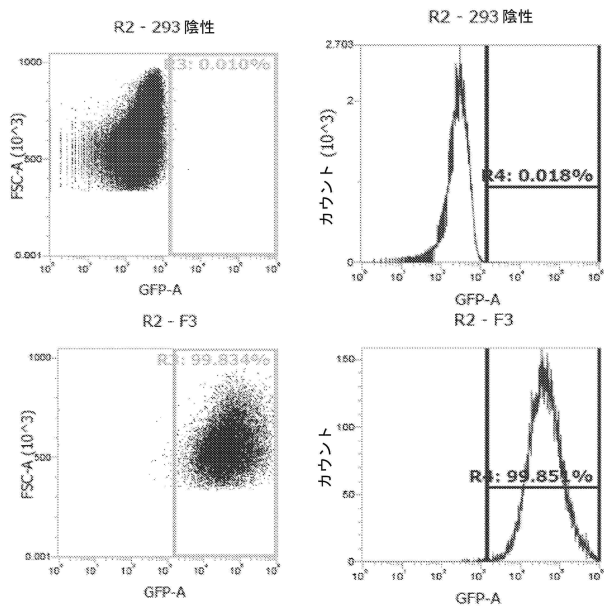
【図 20】



10

20

【図 21】



30

40

【配列表】

2025517630000001.xml

50

【手続補正書】

【提出日】令和7年1月7日(2025.1.7)

【手続補正1】

【補正対象書類名】明細書

【補正対象項目名】配列表

【補正方法】追加

【補正の内容】

【配列表】

2025517630000001.xml

10

20

30

40

50

【国際調査報告】

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US23/66470

A. CLASSIFICATION OF SUBJECT MATTER

IPC - INV. C12N 9/12; C12N 15/11; C12N 15/90 (2023.01)
ADD. A61K 48/00 (2023.01)
CPC - INV. C12N 9/1276; A61K 48/0091; C12N 15/11; C12N 15/90

ADD. A61K 48/00; C12N 2800/22; C12N 2800/40

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)
See Search History document

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

Electronic database consulted during the international search (name of database and, where practicable, search terms used)
See Search History document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	SUN, "Investigating the Mechanism of LINE-1 Retrotransposition in Human Cells" 1-24. Department of Basic Medical Sciences New York University. Web. 01 May 2019; Pages 4 and 9; DOI: https://www.proquest.com/openview/a7d158abaaf4023fb8cff38c6a08499e/1?pq-origsite=gscho lar&cbl=18750&dis=y	61-62 1-3
A	WO 2021/233975 A1 (COMMISSARIAT à L'ENERGIE ATOMIQUE ET AUX ENERGIES ALTERNATIVES) 25 November 2021; Page 2	61-62
A	US 2012/0045471 A1 (HALLER AURELIA) 23 February 2012; paragraph [0382]	1-3

☐ 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

02 January 2024 (02.01.2024)

Date of mailing of the international search report

FEB 23 2024

Name and mailing address of the ISA/

Mall Stop PCT, Attn: ISA/US, Commissioner for Patents
P.O. Box 1450, Alexandria, Virginia 22313-1450
Facsimile No. 571-273-8300

Authorized officer

Shane Thomas

Telephone No. PCT Helpdesk: 571-272-4300

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US23/66470

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.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
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:
3. ☒ Claims Nos.: 4-60, 63-75
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US23/66470

Box No. 1 Nucleotide and/or amino acid sequence(s) (Continuation of item 1.e 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.

b. ☒ furnished subsequent to the international filing date for the purposes of international search (Rule 13ter.1(a)),
☒ accompanied by a statement to the effect that the sequence listing does not go beyond the disclosure in the international application as filed.

2. ☐ With regard to any nucleotide and/or amino acid sequence disclosed in the international application, this report has been established to the extent that a meaningful search could be carried out without a WIPO Standard ST.26 compliant sequence listing.

3. Additional comments:

An "Invitation to Furnish Nucleotide and/or Amino Acid Sequence Listing and to Pay, Where Applicable, Late Furnishing Fee" (Form PCT/ISA/225) was mailed on 17 October 2023 (17.10.2023). The late furnishing fee was not paid (PCT Rule 13ter.1(c) and (d)), therefore, the sequence listing cannot be used for search. Pursuant to PCT Rule 13ter.1(d), the claims have been carried out only to the extent possible without a sequence listing.

フロントページの続き

(51)国際特許分類

A 6 1 K 31/7115(2006.01)
A 6 1 K 9/14 (2006.01)

F I

A 6 1 K 31/7115
A 6 1 K 9/14

テーマコード (参考)

,MC,ME,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,CV,CZ,DE,DJ,DK,DM,DO,DZ,EC,EE,EG,ES,FI,GB,GD,GE,GH,GM,GT,HN,HR,HU,ID,IL,IN,IQ,IR,IS,IT,JM,JO,JP,KE,KG,KH,KN,KP,KR,KW,KZ,LA,LC,LK,LR,LS,LU,LY,MA,MD,MG,MK,MN,MU,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

(特許庁注：以下のものは登録商標)

1. T R I T O N

(74)代理人 100127568

弁理士 酒井 善典

(74)代理人 100171402

弁理士 上田 ▲茂▼

(74)代理人 100213779

弁理士 小川 有佳子

(74)代理人 100077012

弁理士 岩谷 龍

(72)発明者 コリンズ, キャスリーン

アメリカ合衆国 9 4 7 0 4 カリフォルニア州, バークレイ, シャタック アヴェニュー 2 1 5 0
 スイート 5 1 0

(72)発明者 ジャン, シャオズ

アメリカ合衆国 9 4 7 0 4 カリフォルニア州, バークレイ, シャタック アヴェニュー 2 1 5 0
 スイート 5 1 0

(72)発明者 ヴァン トリーク, ブリアナ

アメリカ合衆国 9 4 7 0 4 カリフォルニア州, バークレイ, シャタック アヴェニュー 2 1 5 0
 スイート 5 1 0

(72)発明者 アプトン, ヘザー イー.

アメリカ合衆国 9 4 7 0 4 カリフォルニア州, バークレイ, シャタック アヴェニュー 2 1 5 0
 スイート 5 1 0

(72)発明者 パーム, サラ

アメリカ合衆国 9 4 7 0 4 カリフォルニア州, バークレイ, シャタック アヴェニュー 2 1 5 0
 スイート 5 1 0

(72)発明者 マッキンタイヤー, ジェレミー

アメリカ合衆国 9 4 7 0 4 カリフォルニア州, バークレイ, シャタック アヴェニュー 2 1 5 0
 スイート 5 1 0

F ターム (参考) 4C076 AA29 AA95 CC37 EE51 FF70

4C084 AA13 NA14 ZC411 ZC412

4C086 AA01 AA02 AA03 EA16 MA01 MA04 NA05 NA14 ZC41