

(19)日本国特許庁(JP)

(12)公表特許公報(A)

(11)公表番号

特表2024-533038

(P2024-533038A)

(43)公表日 令和6年9月12日(2024.9.12)

(51)国際特許分類	F I	テーマコード(参考)
C 1 2 N 9/12 (2006.01)	C 1 2 N 9/12	Z N A 4 B 0 6 5
C 1 2 N 15/54 (2006.01)	C 1 2 N 15/54	
C 1 2 N 15/63 (2006.01)	C 1 2 N 15/63	Z
C 1 2 N 1/15 (2006.01)	C 1 2 N 1/15	
C 1 2 N 1/19 (2006.01)	C 1 2 N 1/19	

審査請求 未請求 予備審査請求 未請求 (全92頁) 最終頁に続く

(21)出願番号	特願2024-506884(P2024-506884)	(71)出願人	522185863 メタゲノミ, インク . アメリカ合衆国 9 4 6 0 8 カリフォル ニア州 エメリービル パーク・アベニュー - 1 5 4 5
(86)(22)出願日	令和4年9月7日(2022.9.7)	(74)代理人	110003797 弁理士法人清原国際特許事務所
(85)翻訳文提出日	令和6年3月22日(2024.3.22)	(72)発明者	トーマス, ブライアン シー . アメリカ合衆国 9 4 6 0 8 カリフォル ニア州 エメリービル パーク・アベニュー - 1 5 4 5
(86)国際出願番号	PCT/US2022/076059	(72)発明者	ブラウン, クリストファー アメリカ合衆国 9 4 6 0 8 カリフォル ニア州 エメリービル パーク・アベニュー - 1 5 4 5
(87)国際公開番号	WO2023/039436		
(87)国際公開日	令和5年3月16日(2023.3.16)		
(31)優先権主張番号	63/241,934		
(32)優先日	令和3年9月8日(2021.9.8)		
(33)優先権主張国・地域又は機関	米国(US)		
(81)指定国・地域	AP(BW,GH,GM,KE,LR,LS,MW,MZ,NA ,RW,SD,SL,ST,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,RU,TJ,TM),EP(AL,A T,BE,BG,CH,CY,CZ,DE,DK,EE,ES,FI,FR ,GB,GR,HR,HU,IE,IS,IT,LT,LU,LV,MC, 最終頁に続く		最終頁に続く

(54)【発明の名称】 カーゴヌクレオチド配列を転位するための系及び方法

(57)【要約】

本開示は、カーゴヌクレオチド配列を標的核酸部位に転位するための系及び方法を提供する。これらの系及び方法は、カーゴヌクレオチド配列を含む第1の二本鎖核酸であって、カーゴヌクレオチド配列がトランスポザゼと相互作用するように構成されている、二本鎖核酸と、トランスポザゼであって、カーゴヌクレオチド配列を標的核酸部位に転位するように構成されている、トランスポザゼと、を含み得る。

【選択図】図2

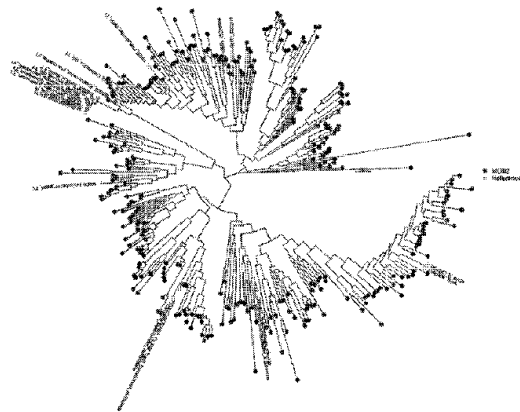


FIG. 2

【特許請求の範囲】

【請求項 1】

操作されたトランスポザーゼ系であって、

(a) カーゴヌクレオチド配列を含む二本鎖核酸であって、前記カーゴヌクレオチド配列が、トランスポザーゼと相互作用するように構成されている、二本鎖核酸と、

(b) トランスポザーゼであって、

(i) 前記カーゴヌクレオチド配列を標的核酸遺伝子座に転位するように構成され、

(i i) 未培養微生物に由来する、トランスポザーゼと、を含む、操作されたトランスポザーゼ系。

【請求項 2】

前記トランスポザーゼが、配列番号 1 ~ 3 4 9 のうちのいずれか 1 つと少なくとも 7 5 % の配列同一性を有する配列を含む、請求項 1 に記載の操作されたトランスポザーゼ系。

【請求項 3】

前記トランスポザーゼが、T n p A トランスポザーゼ又は T n p B トランスポザーゼではない、請求項 1 又は 2 に記載の操作されたトランスポザーゼ系。

【請求項 4】

前記トランスポザーゼが、T n p A トランスポザーゼと 8 0 % 未満の配列同一性を有する、請求項 1 ~ 3 のいずれか一項に記載の操作されたトランスポザーゼ系。

【請求項 5】

前記トランスポザーゼが、T n p B トランスポザーゼと 8 0 % 未満の配列同一性を有する、請求項 1 ~ 4 のいずれか一項に記載の操作されたトランスポザーゼ系。

【請求項 6】

前記トランスポザーゼが、配列番号 1、3、5、7、9、11、13、15、及び 18 ~ 19 のうちのいずれか 1 つと少なくとも約 8 0 %、少なくとも約 8 5 %、少なくとも約 8 6 %、少なくとも約 8 7 %、少なくとも約 8 8 %、少なくとも約 8 9 %、少なくとも約 9 0 %、少なくとも約 9 1 %、少なくとも約 9 2 %、少なくとも約 9 3 %、少なくとも約 9 4 %、少なくとも約 9 5 %、少なくとも約 9 6 %、少なくとも約 9 7 %、少なくとも約 9 8 %、又は少なくとも約 9 9 % の配列同一性を有する、請求項 1 ~ 5 のいずれか一項に記載の操作されたトランスポザーゼ系。

【請求項 7】

前記トランスポザーゼが、触媒チロシン残基を含む、請求項 1 ~ 6 のいずれか一項に記載の操作されたトランスポザーゼ系。

【請求項 8】

前記トランスポザーゼが、サブ末端回文配列を含む左側領域及びサブ末端回文配列を含む右側領域に結合するように構成されている、請求項 1 ~ 7 のいずれか一項に記載の操作されたトランスポザーゼ系。

【請求項 9】

前記トランスポザーゼが、一本鎖デオキシリボ核酸ポリヌクレオチドとして前記カーゴヌクレオチド配列を転位するように構成されている、請求項 1 ~ 8 のいずれか一項に記載の操作されたトランスポザーゼ系。

【請求項 10】

前記トランスポザーゼが、前記トランスポザーゼの N 末端又は C 末端の近位に 1 つ以上の核局在化配列 (N L S) を含む、請求項 1 ~ 9 のいずれか一項に記載の操作されたトランスポザーゼ系。

【請求項 11】

前記 N L S が、配列番号 4 5 5 ~ 4 7 0 からなる群からの配列と少なくとも 8 0 % 同一である配列を含む、請求項 1 ~ 10 のいずれか一項に記載の操作されたトランスポザーゼ系。

【請求項 12】

前記配列同一性が、B L A S T P、C L U S T A L W、M U S C L E、M A F F T、又

10

20

30

40

50

は Smith-Waterman 相同性検索アルゴリズムのパラメータを用いる CLUSTALW によって決定される、請求項 1 ~ 11 のいずれか一項に記載の操作されたトランスポザーゼ系。

【請求項 13】

前記配列同一性が、3 のワード長 (W)、10 の期待値 (E) のパラメータ、及び 11 の存在、1 の延長でギャップコストを設定している BLOSUM62 スコアリングマトリックスを使用し、条件付き組成スコアマトリックス調整を使用した、前記 BLASTP 相同性検索アルゴリズムによって決定される、請求項 12 に記載の操作されたトランスポザーゼ系。

【請求項 14】

操作されたトランスポザーゼ系であって、

(a) カーゴヌクレオチド配列を含む二本鎖核酸であって、前記カーゴヌクレオチド配列が、トランスポザーゼと相互作用するように構成されている、二本鎖核酸と、

(b) トランスポザーゼであって、

(i) 前記カーゴヌクレオチド配列を標的核酸遺伝子座に転位するように構成され、

(ii) 配列番号 1 ~ 349 のうちのいずれか 1 つと少なくとも 75 % の配列同一性を有する配列を含む、トランスポザーゼと、を含む、操作されたトランスポザーゼ系。

【請求項 15】

前記トランスポザーゼが、未培養微生物に由来する、請求項 14 に記載の操作されたトランスポザーゼ系。

【請求項 16】

前記トランスポザーゼが、TnpA トランスポザーゼ又は TnpB トランスポザーゼではない、請求項 14 又は 15 に記載の操作されたトランスポザーゼ系。

【請求項 17】

前記トランスポザーゼが、TnpA トランスポザーゼと 80 % 未満の配列同一性を有する、請求項 14 ~ 16 のいずれか一項に記載の操作されたトランスポザーゼ系。

【請求項 18】

前記トランスポザーゼが、TnpB トランスポザーゼと 80 % 未満の配列同一性を有する、請求項 14 ~ 17 のいずれか一項に記載の操作されたトランスポザーゼ系。

【請求項 19】

前記トランスポザーゼが、配列番号 1、3、5、7、9、11、13、15、及び 18 ~ 19 のうちのいずれか 1 つと少なくとも約 80 %、少なくとも約 85 %、少なくとも約 86 %、少なくとも約 87 %、少なくとも約 88 %、少なくとも約 89 %、少なくとも約 90 %、少なくとも約 91 %、少なくとも約 92 %、少なくとも約 93 %、少なくとも約 94 %、少なくとも約 95 %、少なくとも約 96 %、少なくとも約 97 %、少なくとも約 98 %、又は少なくとも約 99 % の配列同一性を有する、請求項 14 ~ 18 のいずれか一項に記載の操作されたトランスポザーゼ系。

【請求項 20】

前記トランスポザーゼが、触媒チロシン残基を含む、請求項 14 ~ 19 のいずれか一項に記載の操作されたトランスポザーゼ系。

【請求項 21】

前記トランスポザーゼが、サブ末端回文配列を含む左側領域及びサブ末端回文配列を含む右側領域に結合するように構成されている、請求項 14 ~ 20 のいずれか一項に記載の操作されたトランスポザーゼ系。

【請求項 22】

前記トランスポザーゼが、左側認識配列又は右側認識配列に適合する、請求項 14 ~ 20 のいずれか 1 項に記載の操作されたトランスポザーゼ系。

【請求項 23】

前記トランスポザーゼが、一本鎖デオキシリボ核酸ポリヌクレオチドとして前記カーゴヌクレオチド配列を転位するように構成されている、請求項 14 ~ 22 のいずれか一項に

10

20

30

40

50

記載の操作されたトランスポザーゼ系。

【請求項 24】

前記配列同一性が、BLASTP、CLUSTALW、MUSCLE、MAFFT、又はSmith-Waterman 相同性検索アルゴリズムのパラメーターを用いるCLUSTALWによって決定される、請求項14～22のいずれか一項に記載の操作されたトランスポザーゼ系。

【請求項 25】

前記配列同一性が、3のワード長(W)、10の期待値(E)のパラメーター、及び11の存在、1の延長でギャップコストを設定しているBLOSUM62スコアリングマトリックスを使用し、条件付き組成スコアマトリックス調整を使用した、前記BLASTP 10
相同性検索アルゴリズムによって決定される、請求項24に記載の操作されたトランスポ
ザゼ系。

【請求項 26】

請求項1～25のいずれか一項に記載の操作されたトランスポザーゼ系をコードする、
デオキシリボ核酸ポリヌクレオチド。

【請求項 27】

生物における発現に最適化された操作された核酸配列を含む核酸であって、前記核酸が
、トランスポザーゼをコードし、前記トランスポザーゼが、未培養微生物に由来し、前記
生物が、前記未培養微生物ではない、核酸。

【請求項 28】

前記トランスポザーゼが、配列番号1～349のうちのいずれか1つと少なくとも75
%の配列同一性を有するバリエーションを含む、請求項27に記載の核酸。

【請求項 29】

前記トランスポザーゼが、前記トランスポザーゼのN末端又はC末端の近位に1つ以上
の核局在化配列(NLS)をコードする配列を含む、請求項27又は28に記載の核酸。

【請求項 30】

前記NLSが、配列番号455～470から選択される配列を含む、請求項29に記載
の核酸。

【請求項 31】

前記NLSが、配列番号456を含む、請求項29又は30に記載の核酸。

【請求項 32】

前記NLSが、前記トランスポザーゼの前記N末端の近位にある、請求項31に記載の
核酸。

【請求項 33】

前記NLSが、配列番号455を含む、請求項29又は30に記載の核酸。

【請求項 34】

前記NLSが、前記トランスポザーゼの前記C末端の近位にある、請求項33に記載の
核酸。

【請求項 35】

前記生物が、原核生物、細菌、真核生物、真菌、植物、哺乳類、齧歯類、又はヒトであ
る、請求項27～34のいずれか一項に記載の核酸。

【請求項 36】

請求項27～35のいずれか一項に記載の核酸を含む、ベクター。

【請求項 37】

前記トランスポザーゼと複合体を形成するように構成されたカーゴヌクレオチド配列を
コードする核酸を更に含む、請求項36に記載のベクター。

【請求項 38】

前記ベクターが、プラスミド、ミニサークル、CELiD、アデノ随伴ウイルス(AAV)
由来ビリオン、又はレンチウイルスである、請求項36又は37に記載のベクター。

【請求項 39】

10

20

30

40

50

請求項 36 ~ 38 のいずれか一項のいずれか一項に記載のベクターを含む、細胞。

【請求項 40】

請求項 39 に記載の細胞を培養することを含む、トランスポザーゼを製造する方法。

【請求項 41】

カーゴ配列を含む二本鎖デオキシリボ核酸ポリヌクレオチドを結合、ニックング、切断、マーキング、修飾、又は転位する方法であって、

(a) 前記二本鎖デオキシリボ核酸ポリヌクレオチドを、前記カーゴヌクレオチド配列を標的核酸遺伝子座に転位するように構成されたトランスポザーゼと接触させることを含み、

(b) 前記トランスポザーゼが、配列番号 1 ~ 349 のうちのいずれか 1 つと少なくとも 75 % の配列同一性を有する配列を含む、方法。 10

【請求項 42】

前記トランスポザーゼが、未培養微生物に由来する、請求項 41 に記載の方法。

【請求項 43】

前記トランスポザーゼが、TnpA トランスポザーゼ又は TnpB トランスポザーゼではない、請求項 41 又は 42 に記載の方法。

【請求項 44】

前記トランスポザーゼが、TnpA トランスポザーゼと 80 % 未満の配列同一性を有する、請求項 41 ~ 43 のいずれか一項に記載の方法。

【請求項 45】

前記トランスポザーゼが、TnpB トランスポザーゼと 80 % 未満の配列同一性を有する、請求項 41 ~ 44 のいずれか一項に記載の方法。 20

【請求項 46】

前記トランスポザーゼが、配列番号 1、3、5、7、9、11、13、15、及び 18 ~ 19 のうちのいずれか 1 つと少なくとも約 80 %、少なくとも約 85 %、少なくとも約 86 %、少なくとも約 87 %、少なくとも約 88 %、少なくとも約 89 %、少なくとも約 90 %、少なくとも約 91 %、少なくとも約 92 %、少なくとも約 93 %、少なくとも約 94 %、少なくとも約 95 %、少なくとも約 96 %、少なくとも約 97 %、少なくとも約 98 %、少なくとも約 99 %、又は 100 % の配列同一性を有する、請求項 41 から 45 のいずれか一項に記載の方法。 30

【請求項 47】

前記トランスポザーゼが、触媒チロシン残基を含む、請求項 41 ~ 46 のいずれか一項に記載の方法。

【請求項 48】

前記トランスポザーゼが、サブ末端回文配列を含む左側領域及びサブ末端回文配列を含む右側領域に結合するように構成されている、請求項 41 ~ 47 のいずれか一項に記載の方法。

【請求項 49】

前記トランスポザーゼが、左側認識配列又は右側認識配列に適合する、請求項 41 ~ 47 のいずれか一項に記載の方法。 40

【請求項 50】

前記二本鎖デオキシリボ核酸ポリヌクレオチドが、一本鎖デオキシリボ核酸ポリヌクレオチドとして転位される、請求項 41 ~ 49 のいずれか一項に記載の方法。

【請求項 51】

前記二本鎖デオキシリボ核酸ポリヌクレオチドが、真核生物、植物、真菌、哺乳類、齧歯類、又はヒト二本鎖デオキシリボ核酸ポリヌクレオチドである、請求項 41 ~ 50 のいずれか一項に記載の方法。

【請求項 52】

標的核酸遺伝子座を修飾する方法であって、前記方法が、請求項 1 ~ 25 のいずれか一項に記載の操作されたトランスポザーゼ系を前記標的核酸遺伝子座に送達することを含み 40 50

、前記トランスポザーゼが、前記カーゴヌクレオチド配列を前記標的核酸遺伝子座に転位するように構成されており、前記複合体が、前記複合体の前記標的核酸遺伝子座への結合時に、前記複合体が前記標的核酸遺伝子座を修飾するように構成されている、方法。

【請求項 5 3】

前記標的核酸遺伝子座を修飾することが、前記標的核酸遺伝子座を結合、ニッキング、切断、マーキング、修飾、又は転位することを含む、請求項 5 2 に記載の方法。

【請求項 5 4】

前記標的核酸遺伝子座が、デオキシリボ核酸 (DNA) を含む、請求項 5 2 又は 5 3 に記載の方法。

【請求項 5 5】

前記標的核酸遺伝子座が、ゲノム DNA、ウイルス DNA、又は細菌 DNA を含む、請求項 5 4 に記載の方法。

【請求項 5 6】

前記標的核酸遺伝子座が、インピトロである、請求項 5 2 ~ 5 5 のいずれか一項に記載の方法。

【請求項 5 7】

前記標的核酸遺伝子座が、細胞内にある、請求項 5 2 ~ 5 5 のいずれか一項に記載の方法。

【請求項 5 8】

前記細胞が、原核細胞、細菌細胞、真核細胞、真菌細胞、植物細胞、動物細胞、哺乳類細胞、齧歯類細胞、霊長類細胞、ヒト細胞、又は初代細胞である、請求項 5 7 に記載の方法。

【請求項 5 9】

前記細胞が、初代細胞である、請求項 5 7 又は 5 8 に記載の方法。

【請求項 6 0】

前記初代細胞が、T細胞である、請求項 5 9 に記載の方法。

【請求項 6 1】

前記初代細胞が、造血幹細胞 (HSC) である、請求項 5 9 に記載の方法。

【請求項 6 2】

前記操作されたトランスポザーゼ系を前記標的核酸遺伝子座に送達することが、請求項 2 7 ~ 3 5 のいずれか一項に記載の核酸又は請求項 3 6 ~ 3 8 のいずれか一項に記載のベクターを送達することを含む、請求項 5 2 ~ 6 1 のいずれか一項に記載の方法。

【請求項 6 3】

前記操作されたトランスポザーゼ系を前記標的核酸遺伝子座に送達することが、前記トランスポザーゼをコードするオープンリーディングフレームを含む核酸を送達することを含む、請求項 5 2 ~ 6 2 のいずれか一項に記載の方法。

【請求項 6 4】

前記核酸が、前記トランスポザーゼをコードする前記オープンリーディングフレームが作動可能に連結されているプロモーターを含む、請求項 6 3 に記載の方法。

【請求項 6 5】

前記操作されたトランスポザーゼ系を前記標的核酸遺伝子座に送達することが、前記トランスポザーゼをコードする前記オープンリーディングフレームを含有するキャッピングされた mRNA を送達することを含む、請求項 5 2 ~ 6 4 のいずれか一項に記載の方法。

【請求項 6 6】

前記操作されたトランスポザーゼ系を前記標的核酸遺伝子座に送達することが、翻訳されたポリペプチドを送達することを含む、請求項 5 2 ~ 6 5 のいずれか一項に記載の方法。

【請求項 6 7】

前記トランスポザーゼが、前記標的核酸遺伝子座で、又は前記標的核酸遺伝子座の近位で、一本鎖切断又は二本鎖切断を誘導する、請求項 5 2 ~ 6 6 のいずれか一項に記載の方

10

20

30

40

50

法。

【請求項 68】

前記トランスポザーゼが、前記標的遺伝子座内又は前記標的遺伝子座の 5' に互い違いの一本鎖切断を誘導する、請求項 67 に記載の方法。

【請求項 69】

配列番号 1 ~ 349 のうちのいずれか 1 つ又はそのバリエーションと少なくとも 75% の配列同一性を有する異種トランスポザーゼをコードするオープンリーディングフレームを含む、宿主細胞。

【請求項 70】

前記トランスポザーゼが、配列番号 1、3、5、7、9、11、13、15、又は 18 ~ 19 のうちのいずれか 1 つと少なくとも 75% の配列同一性を有する、請求項 69 に記載の宿主細胞。

10

【請求項 71】

前記トランスポザーゼが、配列番号 1、3、5、7、9、11、13、15、又は 18 ~ 19 のうちのいずれか 1 つと少なくとも約 80%、少なくとも約 85%、少なくとも約 86%、少なくとも約 87%、少なくとも約 88%、少なくとも約 89%、少なくとも約 90%、少なくとも約 91%、少なくとも約 92%、少なくとも約 93%、少なくとも約 94%、少なくとも約 95%、少なくとも約 96%、少なくとも約 97%、少なくとも約 98%、少なくとも約 99%、又は 100% の配列同一性を有する、請求項 69 に記載の宿主細胞。

20

【請求項 72】

前記トランスポザーゼが、配列番号 2、4、6、8、10、12、14、又は 17 のうちのいずれか 1 つと少なくとも 75% の配列同一性を有する、請求項 69 に記載の宿主細胞。

【請求項 73】

前記宿主細胞が、E. coli 細胞である、請求項 69 ~ 71 のいずれか一項に記載の宿主細胞。

【請求項 74】

前記 E. coli 細胞が、DE3 リソゲンであるか、又は前記 E. coli 細胞が、BL21 (DE3) 株である、請求項 73 に記載の宿主細胞。

30

【請求項 75】

前記 E. coli 細胞が、ompT lon 遺伝子型を有する、請求項 73 又は 74 に記載の宿主細胞。

【請求項 76】

前記オープンリーディングフレームが、T7 プロモーター配列、T7-lac プロモーター配列、lac プロモーター配列、tac プロモーター配列、trc プロモーター配列、ParaBAD プロモーター配列、PrhaBAD プロモーター配列、T5 プロモーター配列、cspA プロモーター配列、araP_{BAD} プロモーター、ファージラムダからの強い左向きプロモーター (pL プロモーター)、又はそれらの任意の組み合わせに作動可能に連結されている、請求項 69 ~ 75 のいずれか一項に記載の宿主細胞。

40

【請求項 77】

前記オープンリーディングフレームが、前記トランスポザーゼをコードする配列にインフレームで連結された親和性タグをコードする配列を含む、請求項 69 ~ 76 のいずれか一項に記載の宿主細胞。

【請求項 78】

前記親和性タグが、固定化金属親和性クロマトグラフィー (IMAC) タグである、請求項 77 に記載の宿主細胞。

【請求項 79】

前記 IMAC タグが、ポリヒスチジンタグである、請求項 78 に記載の宿主細胞。

【請求項 80】

50

前記親和性タグが、mycタグ、ヒトインフルエンザヘマグルチニン(HA)タグ、マルチース結合タンパク質(MBP)タグ、グルタチオンS-トランスフェラーゼ(GST)タグ、ストレプトアビジンタグ、FLAGタグ、又はそれらの任意の組み合わせである、請求項77に記載の宿主細胞。

【請求項81】

前記親和性タグが、プロテアーゼ切断部位をコードするリンカー配列を介して、前記トランスポザーゼをコードする前記配列にインフレームで連結されている、請求項77~80のいずれか一項に記載の宿主細胞。

【請求項82】

前記プロテアーゼ切断部位が、タバコエッチウイルス(TEV)プロテアーゼ切断部位、PreScission(登録商標)プロテアーゼ切断部位、トロンピン切断部位、第Xa因子切断部位、エンテロキナーゼ切断部位、又はそれらの任意の組み合わせである、請求項81に記載の宿主細胞。

10

【請求項83】

前記オープンリーディングフレームが、前記宿主細胞における発現のためにコドン最適化されている、請求項69~82のいずれか一項に記載の宿主細胞。

【請求項84】

前記オープンリーディングフレームが、ベクター上に提供されている、請求項69~83のいずれか一項に記載の宿主細胞。

【請求項85】

前記オープンリーディングフレームが、前記宿主細胞のゲノムに組み込まれている、請求項69~83のいずれか一項に記載の宿主細胞。

20

【請求項86】

適合する液体培地中に、請求項69~85のいずれか一項に記載の宿主細胞を含む、培養物。

【請求項87】

適合する成長培地中で、請求項69~85のいずれか一項に記載の宿主細胞を培養することを含む、トランスポザーゼを産生する方法。

【請求項88】

追加の化学剤又は増加された量の栄養素を添加することによって、前記トランスポザーゼの発現を誘導することを更に含む、請求項87に記載の方法。

30

【請求項89】

前記追加の化学剤又は増加された量の栄養素が、イソプロピル-D-1-チオガラクトピラノシド(IPTG)又は追加の量のラクトースを含む、請求項88に記載の方法。

【請求項90】

前記培養後に前記宿主細胞を単離することと、前記宿主細胞を溶解してタンパク質抽出物を産生することとを更に含む、請求項87~89のいずれか一項に記載の方法。

【請求項91】

前記タンパク質抽出物をIMAC、又はイオン親和性クロマトグラフィーに供することを更に含む、請求項90に記載の方法。

40

【請求項92】

前記オープンリーディングフレームが、前記トランスポザーゼをコードする配列にインフレームで連結されたIMAC親和性タグをコードする配列を含む、請求項91に記載の方法。

【請求項93】

前記IMAC親和性タグが、プロテアーゼ切断部位をコードするリンカー配列を介して、前記トランスポザーゼをコードする前記配列にインフレームで連結されている、請求項92に記載の方法。

【請求項94】

前記プロテアーゼ切断部位が、タバコエッチウイルス(TEV)プロテアーゼ切断部位

50

、PreScission（登録商標）プロテアーゼ切断部位、トロンピン切断部位、第Xa因子切断部位、エンテロキナーゼ切断部位、又はそれらの任意の組み合わせを含む、請求項93に記載の方法。

【請求項95】

前記プロテアーゼ切断部位に対応するプロテアーゼを前記トランスポザーゼと接触させることによって、前記IMAC親和性タグを切断することを更に含む、請求項93又は94に記載の方法。

【請求項96】

サブラクティブIMAC親和性クロマトグラフィーを実施して、前記トランスポザーゼを含む組成物から前記親和性タグを除去することを更に含む、請求項95に記載の方法

10

【請求項97】

細胞中の遺伝子座を破壊する方法であって、前記細胞に組成物を接触させることを含み、前記組成物が、

(a) カーゴヌクレオチド配列を含む二本鎖核酸であって、前記カーゴヌクレオチド配列が、トランスポザーゼと相互作用するように構成されている、二本鎖核酸と、

(b) トランスポザーゼであって、

(i) 前記カーゴヌクレオチド配列を標的核酸遺伝子座に転位するように構成され、

(ii) 配列番号1～349のうちのいずれか1つと少なくとも75%の配列同一性を有する配列を含み、

20

(iii) 細胞内でTnpAトランスポザーゼと少なくとも同等の転位活性を有する、トランスポザーゼと、を含む、方法。

【請求項98】

前記転位活性が、前記標的核酸遺伝子座を含む細胞に前記トランスポザーゼを導入し、前記細胞内の前記標的核酸遺伝子座の転位を検出することによって、インビトロで測定される、請求項97に記載の方法。

【請求項99】

前記組成物が、20ピコモル(pmol)以下の前記トランスポザーゼを含む、請求項97又は98に記載の方法。

【請求項100】

前記組成物が、1pmol以下の前記トランスポザーゼを含む、請求項99に記載の方法。

30

【請求項101】

操作されたトランスポザーゼ系であって、前記操作されたトランスポザーゼ系が、

(a) カーゴヌクレオチド配列を含む二本鎖核酸であって、前記カーゴヌクレオチド配列が、トランスポザーゼと相互作用するように構成されている、二本鎖核酸と、

(b) トランスポザーゼと、を含み、

(i) 前記トランスポザーゼが、前記カーゴヌクレオチド配列を標的核酸遺伝子座に転位するように構成され、

(ii) 前記二本鎖核酸が、前記カーゴ配列に隣接する隣接配列を含み、前記隣接配列が、配列番号350～454のうちのいずれか1つの少なくとも90個の連続するヌクレオチドと少なくとも約70%の配列同一性を有する、操作されたトランスポザーゼ系。

40

【請求項102】

前記トランスポザーゼが、未培養生物に由来する、請求項101に記載の操作されたトランスポザーゼ系。

【請求項103】

前記トランスポザーゼが、TnpAトランスポザーゼ又はTnpBトランスポザーゼではない、請求項101又は102に記載の操作されたトランスポザーゼ系。

【請求項104】

前記トランスポザーゼが、TnpAトランスポザーゼと80%未満の配列同一性を有す

50

る、請求項 101 ~ 103 のいずれか一項に記載の操作されたトランスポザーゼ系。

【請求項 105】

前記トランスポザーゼが、TnpBトランスポザーゼと80%未満の配列同一性を有する、請求項 101 ~ 104 のいずれか一項に記載の操作されたトランスポザーゼ系。

【請求項 106】

前記トランスポザーゼが、配列番号 1 ~ 349 のうちのいずれか一つと少なくとも75%の配列同一性を有する配列を含む、請求項 101 ~ 105 のいずれか一項に記載の操作されたトランスポザーゼ系。

【請求項 107】

前記トランスポザーゼが、配列番号 1、3、5、7、9、11、13、15、及び18 ~ 19 のうちのいずれか一つと少なくとも約80%、少なくとも約85%、少なくとも約86%、少なくとも約87%、少なくとも約88%、少なくとも約89%、少なくとも約90%、少なくとも約91%、少なくとも約92%、少なくとも約93%、少なくとも約94%、少なくとも約95%、少なくとも約96%、少なくとも約97%、少なくとも約98%、少なくとも約99%、又は100%の配列同一性を有する、請求項 106 に記載の操作されたトランスポザーゼ系。 10

【請求項 108】

前記トランスポザーゼが、触媒チロシン残基を含む、請求項 101 ~ 107 のいずれか一項に記載の操作されたトランスポザーゼ系。

【請求項 109】

前記トランスポザーゼが、サブ末端回文配列を含む左側領域及びサブ末端回文配列を含む右側領域に結合するように構成されている、請求項 101 ~ 108 のいずれか一項に記載の操作されたトランスポザーゼ系。 20

【請求項 110】

前記二本鎖デオキシリボ核酸ポリヌクレオチドが、一本鎖デオキシリボ核酸ポリヌクレオチドとして転位される、請求項 101 ~ 109 のいずれか一項に記載の操作されたトランスポザーゼ系。

【請求項 111】

前記トランスポザーゼが、前記トランスポザーゼのN末端又はC末端の近位に1つ以上の核局在化シグナル(NLS)を含む、請求項 101 ~ 110 のいずれか一項に記載の操作されたトランスポザーゼ系。 30

【請求項 112】

前記1つ以上のNLSのNLSが、配列番号455 ~ 470 からなる群からの配列と少なくとも80%同一である配列を含む、請求項 111 に記載の操作されたトランスポザーゼ系。

【請求項 113】

前記二本鎖デオキシリボ核酸ポリヌクレオチドが、真核生物、植物、真菌、哺乳類、齧歯類、又はヒト二本鎖デオキシリボ核酸ポリヌクレオチドである、請求項 101 ~ 112 のいずれか一項に記載の操作されたトランスポザーゼ系。

【請求項 114】

前記隣接配列が、配列番号350、352、355、356、359、361、362、及び367のうちのいずれか一つの少なくとも90個の連続するヌクレオチドと少なくとも約75%、少なくとも約80%、少なくとも約85%、少なくとも約86%、少なくとも約87%、少なくとも約88%、少なくとも約89%、少なくとも約90%、少なくとも約91%、少なくとも約92%、少なくとも約93%、少なくとも約94%、少なくとも約95%、少なくとも約96%、少なくとも約97%、少なくとも約98%、少なくとも約99%、又は100%の配列同一性を有する、請求項 101 ~ 113 のいずれか一項に記載の操作されたトランスポザーゼ系。 40

【請求項 115】

前記二本鎖核酸が、前記カーゴ配列に隣接する別の隣接配列を含み、前記別の隣接配列 50

が、配列番号 350 ~ 454 のうちのいずれか 1 つの少なくとも 90 個の連続するヌクレオチドと少なくとも約 70 % の配列同一性を有する、請求項 101 ~ 114 のいずれか一項に記載の操作されたトランスポザーゼ系。

【請求項 116】

前記別の隣接配列が、配列番号 351、353、354、357、358、360、363、及び 366 のうちのいずれか 1 つの少なくとも 90 個の連続するヌクレオチドと少なくとも約 75 %、少なくとも約 80 %、少なくとも約 85 %、少なくとも約 86 %、少なくとも約 87 %、少なくとも約 88 %、少なくとも約 89 %、少なくとも約 90 %、少なくとも約 91 %、少なくとも約 92 %、少なくとも約 93 %、少なくとも約 94 %、少なくとも約 95 %、少なくとも約 96 %、少なくとも約 97 %、少なくとも約 98 %、少なくとも約 99 %、又は 100 % の配列同一性を有する、請求項 115 に記載の操作されたトランスポザーゼ系。

10

【請求項 117】

前記隣接配列が、前記カーゴ核酸配列の左端に隣接し、前記別の隣接配列が、前記カーゴ核酸配列の右端に隣接する、請求項 115 又は 116 に記載の操作されたトランスポザーゼ系。

【請求項 118】

前記トランスポザーゼが、前記標的核酸遺伝子座に隣接する挿入モチーフを認識するように構成されている、請求項 101 ~ 117 のいずれか一項に記載の操作されたトランスポザーゼ系。

20

【請求項 119】

前記挿入モチーフが、配列 A A T G A C の少なくとも 3、4、5、又は 6 個の連続するヌクレオチドを含む、請求項 118 に記載の操作されたトランスポザーゼ系。

【請求項 120】

請求項 101 ~ 119 のいずれか一項に記載の操作されたトランスポザーゼ系をコードする、デオキシリボ核酸ポリヌクレオチド。

【請求項 121】

カーゴ配列を含む二本鎖デオキシリボ核酸ポリヌクレオチドを結合、ニックング、切断、マーキング、修飾、又は転位する方法であって、前記方法が、

前記二本鎖デオキシリボ核酸ポリヌクレオチドを、前記カーゴヌクレオチド配列を標的核酸遺伝子座に転位するように構成されたトランスポザーゼと接触させることを含み、

30

前記二本鎖デオキシリボ核酸ポリヌクレオチドが、前記カーゴ配列に隣接する隣接配列を含み、前記隣接配列が、配列番号 350 ~ 454 のうちのいずれか 1 つの少なくとも 90 個の連続するヌクレオチドと少なくとも約 70 % の配列同一性を有する、方法。

【請求項 122】

前記トランスポザーゼが、未培養生物に由来する、請求項 121 に記載の方法。

【請求項 123】

前記トランスポザーゼが、T n p A トランスポザーゼ又は T n p B トランスポザーゼではない、請求項 122 に記載の方法。

【請求項 124】

前記トランスポザーゼが、T n p A トランスポザーゼと 80 % 未満の配列同一性を有する、請求項 121 ~ 123 のいずれか一項に記載の方法。

40

【請求項 125】

前記トランスポザーゼが、T n p B トランスポザーゼと 80 % 未満の配列同一性を有する、請求項 121 ~ 124 のいずれか一項に記載の方法。

【請求項 126】

前記トランスポザーゼが、配列番号 1 ~ 349 のうちのいずれか 1 つと少なくとも 75 % の配列同一性を有する配列を含む、請求項 121 ~ 125 のいずれか一項に記載の方法。

【請求項 127】

50

前記トランスポザーゼが、配列番号 1、3、5、7、9、11、13、15、及び 18 ~ 19 のうちのいずれか 1 つと少なくとも約 80%、少なくとも約 85%、少なくとも約 86%、少なくとも約 87%、少なくとも約 88%、少なくとも約 89%、少なくとも約 90%、少なくとも約 91%、少なくとも約 92%、少なくとも約 93%、少なくとも約 94%、少なくとも約 95%、少なくとも約 96%、少なくとも約 97%、少なくとも約 98%、少なくとも約 99%、又は 100% の配列同一性を有する、請求項 126 に記載の方法。

【請求項 128】

前記トランスポザーゼが、触媒チロシン残基を含む、請求項 121 ~ 127 のいずれか一項に記載の方法。

10

【請求項 129】

前記トランスポザーゼが、サブ末端回文配列を含む左側領域及びサブ末端回文配列を含む右側領域に結合するように構成されている、請求項 121 ~ 128 のいずれか一項に記載の方法。

【請求項 130】

前記トランスポザーゼが、左側認識配列又は右側認識配列に適合する、請求項 121 ~ 129 のいずれか一項に記載の方法。

【請求項 131】

前記二本鎖デオキシリボ核酸ポリヌクレオチドが、一本鎖デオキシリボ核酸ポリヌクレオチドとして転位される、請求項 121 ~ 130 のいずれか一項に記載の方法。

20

【請求項 132】

前記トランスポザーゼが、前記トランスポザーゼの N 末端又は C 末端の近位に 1 つ以上の核局在化シグナル (NLS) を含む、請求項 121 ~ 131 のいずれか一項に記載の方法。

【請求項 133】

前記 1 つ以上の NLS の NLS が、配列番号 455 ~ 470 からなる群からの配列と少なくとも 80% 同一である配列を含む、請求項 121 ~ 132 のいずれか一項に記載の方法。

【請求項 134】

前記二本鎖デオキシリボ核酸ポリヌクレオチドが、真核生物、植物、真菌、哺乳類、齧歯類、又はヒト二本鎖デオキシリボ核酸ポリヌクレオチドである、請求項 121 ~ 133 のいずれか一項に記載の方法。

30

【請求項 135】

前記隣接配列が、配列番号 350、352、355、356、359、361、362、及び 367 のうちのいずれか 1 つの少なくとも 90 個の連続するヌクレオチドと少なくとも約 75%、少なくとも約 80%、少なくとも約 85%、少なくとも約 86%、少なくとも約 87%、少なくとも約 88%、少なくとも約 89%、少なくとも約 90%、少なくとも約 91%、少なくとも約 92%、少なくとも約 93%、少なくとも約 94%、少なくとも約 95%、少なくとも約 96%、少なくとも約 97%、少なくとも約 98%、少なくとも約 99%、又は 100% の配列同一性を有する、請求項 121 ~ 134 のいずれか一項に記載の方法。

40

【請求項 136】

前記二本鎖デオキシリボ核酸ポリヌクレオチドが、前記カーゴ配列に隣接する別の隣接配列を含み、前記別の隣接配列が、配列番号 350 ~ 454 のうちのいずれか 1 つの少なくとも 90 個の連続するヌクレオチドと少なくとも約 70% の配列同一性を有する、請求項 121 ~ 135 のいずれか一項に記載の方法。

【請求項 137】

前記別の隣接配列が、配列番号 351、353、354、357、358、360、363、及び 366 のうちのいずれか 1 つの少なくとも 90 個の連続するヌクレオチドと少なくとも約 75%、少なくとも約 80%、少なくとも約 85%、少なくとも約 86%、少

50

なくとも約 87%、少なくとも約 88%、少なくとも約 89%、少なくとも約 90%、少なくとも約 91%、少なくとも約 92%、少なくとも約 93%、少なくとも約 94%、少なくとも約 95%、少なくとも約 96%、少なくとも約 97%、少なくとも約 98%、少なくとも約 99%、又は 100%の配列同一性を有する、請求項 135 に記載の方法。

【請求項 138】

前記隣接配列が、前記カーゴ核酸配列の左端に隣接し、前記別の隣接配列が、前記カーゴ核酸配列の右端に隣接する、請求項 135 又は 137 に記載の方法。

【請求項 139】

前記トランスポザーゼが、前記標的核酸遺伝子座に隣接する挿入モチーフを認識するように構成されている、請求項 121 ~ 138 のいずれか一項に記載の方法。

10

【請求項 140】

前記挿入モチーフが、配列 A A T G A C の少なくとも 3、4、5、又は 6 個の連続するヌクレオチドを含む、請求項 139 に記載の方法。

【請求項 141】

標的核酸遺伝子座を修飾する方法であって、前記方法が、請求項 101 ~ 119 のいずれか一項に記載の操作されたトランスポザーゼ系を前記標的核酸遺伝子座に送達することを含み、前記トランスポザーゼが、前記カーゴヌクレオチド配列を前記標的核酸遺伝子座に転位するように構成されており、前記複合体が、前記複合体の前記標的核酸遺伝子座への結合時に、前記複合体が前記標的核酸遺伝子座を修飾するように構成されている、方法。

20

【請求項 142】

前記標的核酸遺伝子座を修飾することが、前記標的核酸遺伝子座を結合、ニックング、切断、マーキング、修飾、又は転位することを含む、請求項 141 に記載の方法。

【請求項 143】

前記標的核酸遺伝子座が、デオキシリボ核酸 (DNA) を含む、請求項 141 又は 142 に記載の方法。

【請求項 144】

前記標的核酸遺伝子座が、ゲノム DNA、ウイルス DNA、又は細菌 DNA を含む、請求項 143 に記載の方法。

【請求項 145】

前記標的核酸遺伝子座が、インピトロである、請求項 141 ~ 144 のいずれか一項に記載の方法。

30

【請求項 146】

前記標的核酸遺伝子座が、細胞内にある、請求項 141 ~ 145 のいずれか一項に記載の方法。

【請求項 147】

前記細胞が、原核細胞、細菌細胞、真核細胞、真菌細胞、植物細胞、動物細胞、哺乳類細胞、齧歯類細胞、霊長類細胞、ヒト細胞、又は初代細胞である、請求項 146 に記載の方法。

【請求項 148】

前記細胞が、初代細胞である、請求項 146 又は 147 に記載の方法。

40

【請求項 149】

前記初代細胞が、T細胞である、請求項 148 に記載の方法。

【請求項 150】

前記初代細胞が、造血幹細胞 (HSC) である、請求項 148 に記載の方法。

【請求項 151】

前記操作されたトランスポザーゼ系を前記標的核酸遺伝子座に送達することが、前記トランスポザーゼをコードするオープンリーディングフレームを含む核酸を送達することを含む、請求項 141 ~ 150 のいずれか一項に記載の方法。

【請求項 152】

50

前記核酸が、前記トランスポザーゼをコードする前記オープンリーディングフレームが作動可能に連結されているプロモーターを含む、請求項 1 5 1 に記載の方法。

【請求項 1 5 3】

前記操作されたトランスポザーゼ系を前記標的核酸遺伝子座に送達することが、前記トランスポザーゼをコードする前記オープンリーディングフレームを含有するキャッピングされた mRNA を送達することを含む、請求項 1 5 1 又は 1 5 2 に記載の方法。

【請求項 1 5 4】

前記操作されたトランスポザーゼ系を前記標的核酸遺伝子座に送達することが、翻訳されたポリペプチドを送達することを含む、請求項 1 4 1 ~ 1 5 3 のいずれか一項に記載の方法。

【請求項 1 5 5】

前記トランスポザーゼが、前記標的核酸遺伝子座で、又は前記標的核酸遺伝子座の近位で、一本鎖切断又は二本鎖切断を誘導する、請求項 1 4 1 ~ 1 5 4 のいずれか一項に記載の方法。

【請求項 1 5 6】

前記トランスポザーゼが、前記標的遺伝子座内又は前記標的遺伝子座の 5 ' に互い違いの一本鎖切断を誘導する、請求項 1 5 5 に記載の方法。

【発明の詳細な説明】

【技術分野】

【0 0 0 1】

相互参照

本出願は、2021年9月8日に提出された「SYSTEMS AND METHODS FOR TRANSPOSING CARGO NUCLEOTIDE SEQUENCES」と題された米国仮特許出願第63/241,934号の利益を主張するものであり、その全体が参照により本明細書に組み込まれる。

【背景技術】

【0 0 0 2】

転位因子は、遺伝子の機能及び進化において重要な役割を果たす移動可能なDNA配列である。転位因子はほぼ全ての種類の生命体で見られるが、それらの保有率は生物間で異なり、真核生物ゲノムの大部分は転位因子をコードする（ヒトでは少なくとも45%）。転位因子に関する基礎的研究は1940年代に行われたが、DNA操作及び遺伝子編集の用途におけるそれらの潜在的な有用性が認識されるようになったのは近年のことである。

【0 0 0 3】

配列表

本出願は、XML形式で電子的に提出された配列表を含み、参照によりその全体が本明細書に組み込まれる。2022年9月7日に作成された当該XMLコピーは、55921-733601.xmlと名付けられ、452,421バイトのサイズである。

【発明の概要】

【0 0 0 4】

いくつかの態様では、本開示は、操作されたトランスポザーゼ系を提供し、上記操作されたトランスポザーゼ系は、カーゴヌクレオチド配列を含む二本鎖核酸であって、カーゴヌクレオチド配列がトランスポザーゼと相互作用するように構成されている、二本鎖核酸と、トランスポザーゼであって、カーゴヌクレオチド配列を標的核酸遺伝子座に転位するように構成され、未培養微生物に由来する、トランスポザーゼと、を含む。

【0 0 0 5】

いくつかの実施形態では、トランスポザーゼは、配列番号1~349のうちのいずれか1つと少なくとも75%の配列同一性を有する配列を含む。いくつかの実施形態では、トランスポザーゼは、TnpAトランスポザーゼ又はTnpBトランスポザーゼではない。いくつかの実施形態では、トランスポザーゼは、TnpAトランスポザーゼと80%未満の配列同一性を有する。いくつかの実施形態では、トランスポザーゼは、TnpBトラン

10

20

30

40

50

スポザーゼと80%未満の配列同一性を有する。いくつかの実施形態では、トランススポザーゼは、配列番号1、3、5、7、9、11、13、15、及び18~19のうちのいずれか1つと少なくとも約80%、少なくとも約85%、少なくとも約86%、少なくとも約87%、少なくとも約88%、少なくとも約89%、少なくとも約90%、少なくとも約91%、少なくとも約92%、少なくとも約93%、少なくとも約94%、少なくとも約95%、少なくとも約96%、少なくとも約97%、少なくとも約98%、又は少なくとも約99%の配列同一性を有する。いくつかの実施形態では、トランススポザーゼは、触媒チロシン残基を含む。いくつかの実施形態では、トランススポザーゼは、サブ末端回文配列を含む左側領域及びサブ末端回文配列を含む右側領域に結合するように構成されている。いくつかの実施形態では、トランススポザーゼは、一本鎖デオキシリボ核酸ポリヌクレオチドとしてカーゴヌクレオチド配列を転位するように構成されている。いくつかの実施形態では、トランススポザーゼは、トランススポザーゼのN末端又はC末端の近位に1つ以上の核局在化配列(NLS)を含む。いくつかの実施形態では、NLSは、配列番号455~470からなる群からの配列と少なくとも80%同一である配列を含む。いくつかの実施形態では、配列同一性は、BLASTP、CLUSTALW、MUSCLE、MAFFT、又はSmith-Waterman相溶性検索アルゴリズムのパラメーターを用いるCLUSTALWによって決定される。いくつかの実施形態では、配列同一性は、3のワード長(W)、10の期待値(E)のパラメーター、及び11の存在、1の延長でギャップコストを設定しているBLOSUM62スコアリングマトリックスを使用し、条件付き組成スコアマトリックス調整を使用した、BLASTP相溶性検索アルゴリズムによって決定される。

【0006】

いくつかの態様では、本開示は、操作されたトランススポザーゼ系を提供し、上記操作されたトランススポザーゼ系は、カーゴヌクレオチド配列を含む二本鎖核酸であって、カーゴヌクレオチド配列がトランススポザーゼと相互作用するように構成されている、二本鎖核酸と、トランススポザーゼであって、カーゴヌクレオチド配列を標的核酸遺伝子座に転位するように構成され、配列番号1~349のうちのいずれか1つと少なくとも75%の配列同一性を有する配列を含む、トランススポザーゼと、を含む。

【0007】

いくつかの実施形態では、トランススポザーゼは、未培養微生物に由来する。いくつかの実施形態では、トランススポザーゼは、TnpAトランススポザーゼ又はTnpBトランススポザーゼではない。いくつかの実施形態では、トランススポザーゼは、TnpAトランススポザーゼと80%未満の配列同一性を有する。いくつかの実施形態では、トランススポザーゼは、TnpBトランススポザーゼと80%未満の配列同一性を有する。いくつかの実施形態では、トランススポザーゼは、配列番号1、3、5、7、9、11、13、15、及び18~19のうちのいずれか1つと少なくとも約80%、少なくとも約85%、少なくとも約86%、少なくとも約87%、少なくとも約88%、少なくとも約89%、少なくとも約90%、少なくとも約91%、少なくとも約92%、少なくとも約93%、少なくとも約94%、少なくとも約95%、少なくとも約96%、少なくとも約97%、少なくとも約98%、又は少なくとも約99%の配列同一性を有する。いくつかの実施形態では、トランススポザーゼは、触媒チロシン残基を含む。いくつかの実施形態では、トランススポザーゼは、サブ末端回文配列を含む左側領域及びサブ末端回文配列を含む右側領域に結合するように構成されている。いくつかの実施形態では、トランススポザーゼは、左側認識配列又は右側認識配列に適合する。いくつかの実施形態では、トランススポザーゼは、一本鎖デオキシリボ核酸ポリヌクレオチドとしてカーゴヌクレオチド配列を転位するように構成されている。いくつかの実施形態では、配列同一性は、BLASTP、CLUSTALW、MUSCLE、MAFFT、又はSmith-Waterman相溶性検索アルゴリズムのパラメーターを用いるCLUSTALWによって決定される。いくつかの実施形態では、配列同一性は、3のワード長(W)、10の期待値(E)のパラメーター、及び11の存在、1の延長でギャップコストを設定しているBLOSUM62スコアリングマトリックスを

使用し、条件付き組成スコアマトリックス調整を使用した、BLASTP 相溶性検索アルゴリズムによって決定される。

【0008】

いくつかの態様では、本開示は、本明細書に開示される任意の操作されたトランスポザーゼ系をコードするデオキシリボ核酸ポリヌクレオチドを提供する。

【0009】

いくつかの態様では、本開示は、生物における発現に最適化された操作された核酸配列を含む核酸を提供し、核酸はトランスポザーゼをコードし、トランスポザーゼは未培養微生物に由来し、生物は未培養微生物ではない。

【0010】

いくつかの実施形態では、トランスポザーゼは、配列番号 1 ~ 349 のうちのいずれか 1 つと少なくとも 75 % の配列同一性を有するバリエーションを含む。いくつかの実施形態では、トランスポザーゼは、トランスポザーゼの N 末端又は C 末端の近位に 1 つ以上の核局在化配列 (NLS) をコードする配列を含む。いくつかの実施形態では、NLS は、配列番号 455 ~ 470 から選択される配列を含む。いくつかの実施形態では、NLS は、配列番号 456 を含む。いくつかの実施形態では、NLS は、トランスポザーゼの N 末端の近位にある。いくつかの実施形態では、NLS は、配列番号 455 を含む。いくつかの実施形態では、NLS は、トランスポザーゼの C 末端の近位にある。いくつかの実施形態では、生物は、原核生物、細菌、真核生物、真菌、植物、哺乳類、齧歯類、又はヒトである。

10

20

【0011】

いくつかの態様では、本開示は、本明細書に開示される任意の核酸を含むベクターを提供する。いくつかの実施形態では、核酸は、トランスポザーゼと複合体を形成するように構成されたカーゴヌクレオチド配列をコードする核酸を更に含む。いくつかの実施形態では、ベクターは、プラスミド、ミニサークル、CELLiD、アデノ随伴ウイルス (AAV) 由来ビリオン、又はレンチウイルスである。

【0012】

いくつかの態様では、本開示は、本明細書に開示される任意のベクターを含む細胞を提供する。

【0013】

いくつかの態様では、本開示は、本明細書に開示される任意の細胞を培養することを含む、トランスポザーゼを製造する方法を提供する。

30

【0014】

いくつかの態様では、本開示は、カーゴ配列を含む二本鎖デオキシリボ核酸ポリヌクレオチドを結合、ニックング、切断、マーキング、修飾、又は転位する方法を提供し、上記方法は、二本鎖デオキシリボ核酸ポリヌクレオチドを、カーゴヌクレオチド配列を標的核酸遺伝子座に転位するように構成されたトランスポザーゼと接触させることを含み、トランスポザーゼは、配列番号 1 ~ 349 のうちのいずれか 1 つと少なくとも 75 % の配列同一性を有する配列を含む。

【0015】

いくつかの実施形態では、トランスポザーゼは、未培養微生物に由来する。いくつかの実施形態では、トランスポザーゼは、TnpA トランスポザーゼ又は TnpB トランスポザーゼではない。いくつかの実施形態では、トランスポザーゼは、TnpA トランスポザーゼと 80 % 未満の配列同一性を有する。いくつかの実施形態では、トランスポザーゼは、TnpB トランスポザーゼと 80 % 未満の配列同一性を有する。いくつかの実施形態では、トランスポザーゼは、配列番号 1、3、5、7、9、11、13、15、及び 18 ~ 19 のうちのいずれか 1 つと少なくとも約 80 %、少なくとも約 85 %、少なくとも約 86 %、少なくとも約 87 %、少なくとも約 88 %、少なくとも約 89 %、少なくとも約 90 %、少なくとも約 91 %、少なくとも約 92 %、少なくとも約 93 %、少なくとも約 94 %、少なくとも約 95 %、少なくとも約 96 %、少なくとも約 97 %、少なくとも約 9

40

50

8%、少なくとも約99%、又は100%の配列同一性を有する。いくつかの実施形態では、トランスポザーゼは、触媒チロシン残基を含む。いくつかの実施形態では、トランスポザーゼは、サブ末端回文配列を含む左側領域及びサブ末端回文配列を含む右側領域に結合するように構成されている。いくつかの実施形態では、トランスポザーゼは、左側認識配列又は右側認識配列に適合する。いくつかの実施形態では、二本鎖デオキシリボ核酸ポリヌクレオチドは、一本鎖デオキシリボ核酸ポリヌクレオチドとして転位される。いくつかの実施形態では、二本鎖デオキシリボ核酸ポリヌクレオチドは、真核生物、植物、真菌、哺乳類、齧歯類、又はヒト二本鎖デオキシリボ核酸ポリヌクレオチドである。

【0016】

いくつかの態様では、本開示は、標的核酸遺伝子座を修飾する方法を提供し、上記方法は、本明細書に開示される操作されたトランスポザーゼ系を標的核酸遺伝子座に送達することを含み、トランスポザーゼは、カーゴヌクレオチド配列を標的核酸遺伝子座に転位するように構成されており、複合体は、複合体の標的核酸遺伝子座への結合時に、複合体が標的核酸遺伝子座を修飾するように構成されている。

10

【0017】

いくつかの実施形態では、標的核酸遺伝子座を修飾することは、標的核酸遺伝子座を結合、ニックング、切断、マーキング、修飾、又は転位することを含む。いくつかの実施形態では、標的核酸遺伝子座は、デオキシリボ核酸(DNA)を含む。いくつかの実施形態では、標的核酸遺伝子座は、ゲノムDNA、ウイルスDNA、又は細菌DNAを含む。いくつかの実施形態では、標的核酸遺伝子座は、インピトロである。いくつかの実施形態では、標的核酸遺伝子座は、細胞内にある。いくつかの実施形態では、細胞は、原核細胞、細菌細胞、真核細胞、真菌細胞、植物細胞、動物細胞、哺乳類細胞、齧歯類細胞、霊長類細胞、ヒト細胞、又は初代細胞である。いくつかの実施形態では、細胞は、初代細胞である。いくつかの実施形態では、初代細胞は、T細胞である。いくつかの実施形態では、初代細胞は、造血幹細胞(HSC)である。いくつかの実施形態では、操作されたトランスポザーゼ系を標的核酸遺伝子座に送達することは、本明細書に開示される核酸又は本明細書に開示される任意のベクターを送達することを含む。いくつかの実施形態では、操作されたトランスポザーゼ系を標的核酸遺伝子座に送達することは、トランスポザーゼをコードするオープンリーディングフレームを含む核酸を送達することを含む。いくつかの実施形態では、核酸は、トランスポザーゼをコードするオープンリーディングフレームが作動可能に連結されているプロモーターを含む。いくつかの実施形態では、操作されたトランスポザーゼ系を標的核酸遺伝子座に送達することは、トランスポザーゼをコードするオープンリーディングフレームを含有するキャッピングされたmRNAを送達することを含む。いくつかの実施形態では、操作されたトランスポザーゼ系を標的核酸遺伝子座に送達することは、翻訳されたポリペプチドを送達することを含む。いくつかの実施形態では、トランスポザーゼは、標的核酸遺伝子座で、又は標的核酸遺伝子座の近位で、一本鎖切断又は二本鎖切断を誘導する。いくつかの実施形態では、トランスポザーゼは、標的遺伝子座内又は標的遺伝子座の5'に互い違いの一本鎖切断を誘導する。

20

30

【0018】

いくつかの態様では、本開示は、配列番号1~349のうちのいずれか1つ又はそのバリエーションと少なくとも75%の配列同一性を有する異種トランスポザーゼをコードするオープンリーディングフレームを含む、宿主細胞を提供する。いくつかの実施形態では、トランスポザーゼは、配列番号1、3、5、7、9、11、13、15、又は18~19のうちのいずれか1つと少なくとも75%の配列同一性を有する。いくつかの実施形態では、トランスポザーゼは、配列番号1、3、5、7、9、11、13、15、又は18~19のうちのいずれか1つと少なくとも約80%、少なくとも約85%、少なくとも約86%、少なくとも約87%、少なくとも約88%、少なくとも約89%、少なくとも約90%、少なくとも約91%、少なくとも約92%、少なくとも約93%、少なくとも約94%、少なくとも約95%、少なくとも約96%、少なくとも約97%、少なくとも約98%、少なくとも約99%、又は100%の配列同一性を有する。いくつかの実施形態では

40

50

、トランスポザーゼは、配列番号 2、4、6、8、10、12、14、又は 17 のうちのいずれか 1 つと少なくとも 75 % の配列同一性を有する。いくつかの実施形態では、宿主細胞は、E. coli 細胞である。いくつかの実施形態では、E. coli 細胞は、DE3 リソゲンであるか、又は E. coli 細胞は、BL21 (DE3) 株である。いくつかの実施形態では、E. coli 細胞は、ompT lon 遺伝子型を有する。いくつかの実施形態では、オープンリーディングフレームは、T7 プロモーター配列、T7-lac プロモーター配列、lac プロモーター配列、tac プロモーター配列、trc プロモーター配列、ParaBAD プロモーター配列、PrhaBAD プロモーター配列、T5 プロモーター配列、cspA プロモーター配列、araP BAD プロモーター、ファージラムダからの強い左向きプロモーター (pL プロモーター)、又はそれらの任意の組み合わせに作動可能に連結されている。いくつかの実施形態では、オープンリーディングフレームは、トランスポザーゼをコードする配列にインフレームで連結された親和性タグをコードする配列を含む。いくつかの実施形態では、親和性タグは、固定化金属親和性クロマトグラフィー (IMAC) タグである。いくつかの実施形態では、IMAC タグは、ポリヒスチジンタグである。いくつかの実施形態では、親和性タグは、myc タグ、ヒトインフルエンザヘマグルチニン (HA) タグ、マルトース結合タンパク質 (MBP) タグ、グルタチオン S-トランスフェラーゼ (GST) タグ、ストレプトアビジンタグ、FLAG タグ、又はそれらの任意の組み合わせである。いくつかの実施形態では、親和性タグは、プロテアーゼ切断部位をコードするリンカー配列を介して、トランスポザーゼをコードする配列にインフレームで連結されている。いくつかの実施形態では、プロテアーゼ切断部位は、タバコエッチウイルス (TEV) プロテアーゼ切断部位、PreScission (登録商標) プロテアーゼ切断部位、トロンピン切断部位、第 Xa 因子切断部位、エンテロキナーゼ切断部位、又はそれらの任意の組み合わせである。いくつかの実施形態では、オープンリーディングフレームは、宿主細胞における発現のためにコドン最適化される。いくつかの実施形態では、オープンリーディングフレームは、ベクター上に提供される。いくつかの実施形態では、オープンリーディングフレームは、宿主細胞のゲノムに組み込まれる。

【0019】

いくつかの態様では、本開示は、適合する液体培地中に、本明細書に開示される任意の宿主細胞を含む、培養物を提供する。

【0020】

いくつかの態様では、本開示は、適合する成長培地中で、本明細書に開示される任意の宿主細胞を培養することを含む、トランスポザーゼを産生する方法を提供する。

【0021】

いくつかの実施形態では、方法は、追加の化学剤又は増加された量の栄養素を添加することによって、トランスポザーゼの発現を誘導することを更に含む。いくつかの実施形態では、追加の化学剤又は増加された量の栄養素は、イソプロピル - D - 1 - チオガラクトピラノシド (IPTG) 又は追加の量のラクトースを含む。いくつかの実施形態では、方法は、培養後に宿主細胞を単離することと、宿主細胞を溶解してタンパク質抽出物を産生することとを更に含む。いくつかの実施形態では、方法は、タンパク質抽出物を IMAC、又はイオン親和性クロマトグラフィーに供することを更に含む。いくつかの実施形態では、オープンリーディングフレームは、トランスポザーゼをコードする配列にインフレームで連結された IMAC 親和性タグをコードする配列を含む。いくつかの実施形態では、IMAC 親和性タグは、プロテアーゼ切断部位をコードするリンカー配列を介して、トランスポザーゼをコードする配列にインフレームで連結されている。いくつかの実施形態では、プロテアーゼ切断部位は、タバコエッチウイルス (TEV) プロテアーゼ切断部位、PreScission (登録商標) プロテアーゼ切断部位、トロンピン切断部位、第 Xa 因子切断部位、エンテロキナーゼ切断部位、又はそれらの任意の組み合わせを含む。いくつかの実施形態では、方法は、プロテアーゼ切断部位に対応するプロテアーゼをトランスポザーゼと接触させることによって、IMAC 親和性タグを切断することを更に含む

。いくつかの実施形態では、方法は、サブトラクティブ I M A C 親和性クロマトグラフィーを実施して、トランスポザーゼを含む組成物から親和性タグを除去することを更に含む。

【0022】

いくつかの態様では、本開示は、細胞中の遺伝子座を破壊する方法を提供し、上記方法は、細胞に組成物を接触させることを含み、組成物は、カーゴヌクレオチド配列を含む二本鎖核酸であって、カーゴヌクレオチド配列が、トランスポザーゼと相互作用するように構成されている、二本鎖核酸と、トランスポザーゼであって、カーゴヌクレオチド配列を標的核酸遺伝子座に転位するように構成され、配列番号 1 ~ 349 のうちのいずれか 1 つと少なくとも 75 % の配列同一性を有する配列を含み、細胞内で T n p A トランスポザーゼと少なくとも同等の転位活性を有する、トランスポザーゼと、を含む。

10

【0023】

いくつかの実施形態では、転位活性は、標的核酸遺伝子座を含む細胞にトランスポザーゼを導入し、細胞内の標的核酸遺伝子座の転位を検出することによって、インビトロで測定される。いくつかの実施形態では、組成物は、20ピコモル (p m o l) 以下のトランスポザーゼを含む。いくつかの実施形態では、組成物は、1 p m o l 以下のトランスポザーゼを含む。

【0024】

いくつかの態様では、本開示は、操作されたトランスポザーゼ系を提供し、上記操作されたトランスポザーゼ系は、カーゴヌクレオチド配列を含む二本鎖核酸であって、カーゴヌクレオチド配列がトランスポザーゼと相互作用するように構成されている、二本鎖核酸と、カーゴヌクレオチド配列を標的核酸遺伝子座に転位するように構成されている、トランスポザーゼとを含み、二本鎖核酸は、カーゴ配列に隣接する隣接配列を含み、隣接配列は、配列番号 350 ~ 454 のうちのいずれか 1 つの少なくとも 90 個の連続するヌクレオチドと少なくとも約 70 % の配列同一性を有する。

20

【0025】

いくつかの実施形態では、トランスポザーゼは、未培養生物に由来する。いくつかの実施形態では、トランスポザーゼは、T n p A トランスポザーゼ又は T n p B トランスポザーゼではない。いくつかの実施形態では、トランスポザーゼは、T n p A トランスポザーゼと 80 % 未満の配列同一性を有する。いくつかの実施形態では、トランスポザーゼは、T n p B トランスポザーゼと 80 % 未満の配列同一性を有する。いくつかの実施形態では、トランスポザーゼは、配列番号 1 ~ 349 のうちのいずれか 1 つと少なくとも 75 % の配列同一性を有する配列を含む。いくつかの実施形態では、トランスポザーゼは、配列番号 1、3、5、7、9、11、13、15、及び 18 ~ 19 のうちのいずれか 1 つと少なくとも約 80 %、少なくとも約 85 %、少なくとも約 86 %、少なくとも約 87 %、少なくとも約 88 %、少なくとも約 89 %、少なくとも約 90 %、少なくとも約 91 %、少なくとも約 92 %、少なくとも約 93 %、少なくとも約 94 %、少なくとも約 95 %、少なくとも約 96 %、少なくとも約 97 %、少なくとも約 98 %、少なくとも約 99 %、又は 100 % の配列同一性を有する。いくつかの実施形態では、トランスポザーゼは、触媒チロシン残基を含む。いくつかの実施形態では、トランスポザーゼは、サブ末端回文配列を含む左側領域及びサブ末端回文配列を含む右側領域に結合するように構成されている。いくつかの実施形態では、二本鎖デオキシリボ核酸ポリヌクレオチドは、一本鎖デオキシリボ核酸ポリヌクレオチドとして転位される。いくつかの実施形態では、トランスポザーゼは、トランスポザーゼの N 末端又は C 末端の近位に 1 つ以上の核局在化シグナル (N L S) を含む。いくつかの実施形態では、1 つ以上の N L S の N L S は、配列番号 455 ~ 470 からなる群からの配列と少なくとも 80 % 同一である配列を含む。いくつかの実施形態では、二本鎖デオキシリボ核酸ポリヌクレオチドは、真核生物、植物、真菌、哺乳類、齧歯類、又はヒト二本鎖デオキシリボ核酸ポリヌクレオチドである。いくつかの実施形態では、隣接配列は、配列番号 350、352、355、356、359、361、362、及び 367 のうちのいずれか 1 つの少なくとも 90 個の連続するヌクレオチドと少なく

30

40

50

とも約75%、少なくとも約80%、少なくとも約85%、少なくとも約86%、少なくとも約87%、少なくとも約88%、少なくとも約89%、少なくとも約90%、少なくとも約91%、少なくとも約92%、少なくとも約93%、少なくとも約94%、少なくとも約95%、少なくとも約96%、少なくとも約97%、少なくとも約98%、少なくとも約99%、又は100%の配列同一性を有する。いくつかの実施形態では、二本鎖核酸は、カーゴ配列に隣接する別の隣接配列を含み、上記別の隣接配列は、配列番号350~454のうちのいずれか1つの少なくとも90個の連続するヌクレオチドと少なくとも約70%の配列同一性を有する。いくつかの実施形態では、別の隣接配列は、配列番号351、353、354、357、358、360、363、及び366のうちのいずれか1つの少なくとも90個の連続するヌクレオチドと少なくとも約75%、少なくとも約80%、少なくとも約85%、少なくとも約86%、少なくとも約87%、少なくとも約88%、少なくとも約89%、少なくとも約90%、少なくとも約91%、少なくとも約92%、少なくとも約93%、少なくとも約94%、少なくとも約95%、少なくとも約96%、少なくとも約97%、少なくとも約98%、少なくとも約99%、又は100%の配列同一性を有する。いくつかの実施形態では、隣接配列は、カーゴ核酸配列の左端に隣接し、別の隣接配列は、カーゴ核酸配列の右端に隣接する。いくつかの実施形態では、トランスポザーゼは、標的核酸遺伝子座に隣接する挿入モチーフを認識するように構成されている。いくつかの実施形態では、挿入モチーフは、配列AATGACの少なくとも3、4、5、又は6個の連続するヌクレオチドを含む。

10

【0026】

20

いくつかの態様では、本開示は、本明細書に開示される任意の操作されたトランスポザーゼ系をコードするデオキシリボ核酸ポリヌクレオチドを提供する。

【0027】

いくつかの態様では、本開示は、カーゴ配列を含む二本鎖デオキシリボ核酸ポリヌクレオチドを結合、ニックング、切断、マーキング、修飾、又は転位する方法を提供し、上記方法は、二本鎖デオキシリボ核酸ポリヌクレオチドを、カーゴヌクレオチド配列を標的核酸遺伝子座に転位するように構成されたトランスポザーゼと接触させることを含み、二本鎖デオキシリボ核酸ポリヌクレオチドは、カーゴ配列に隣接する隣接配列を含み、隣接配列は、配列番号350~454のうちのいずれか1つの少なくとも90個の連続するヌクレオチドと少なくとも約70%の配列同一性を有する。

30

【0028】

いくつかの実施形態では、トランスポザーゼは、未培養生物に由来する。いくつかの実施形態では、トランスポザーゼは、TnpAトランスポザーゼ又はTnpBトランスポザーゼではない。いくつかの実施形態では、トランスポザーゼは、TnpAトランスポザーゼと80%未満の配列同一性を有する。いくつかの実施形態では、トランスポザーゼは、TnpBトランスポザーゼと80%未満の配列同一性を有する。いくつかの実施形態では、トランスポザーゼは、配列番号1~349のうちのいずれか1つと少なくとも約75%の配列同一性を有する配列を含む。いくつかの実施形態では、トランスポザーゼは、配列番号1、3、5、7、9、11、13、15、及び18~19のうちのいずれか1つと少なくとも約80%、少なくとも約85%、少なくとも約86%、少なくとも約87%、少なくとも約88%、少なくとも約89%、少なくとも約90%、少なくとも約91%、少なくとも約92%、少なくとも約93%、少なくとも約94%、少なくとも約95%、少なくとも約96%、少なくとも約97%、少なくとも約98%、少なくとも約99%、又は100%の配列同一性を有する。いくつかの実施形態では、トランスポザーゼは、触媒チロシン残基を含む。いくつかの実施形態では、トランスポザーゼは、サブ末端回文配列を含む左側領域及びサブ末端回文配列を含む右側領域に結合するように構成されている。いくつかの実施形態では、トランスポザーゼは、左側認識配列又は右側認識配列に適合する。いくつかの実施形態では、二本鎖デオキシリボ核酸ポリヌクレオチドは、一本鎖デオキシリボ核酸ポリヌクレオチドとして転位される。いくつかの実施形態では、トランスポザーゼは、トランスポザーゼのN末端又はC末端の近位に1つ以上の核局在化シグナル(N

40

50

LS)を含む。いくつかの実施形態では、1つ以上のNLSのNLSは、配列番号455~470からなる群からの配列と少なくとも80%同一である配列を含む。いくつかの実施形態では、二本鎖デオキシリボ核酸ポリヌクレオチドは、真核生物、植物、真菌、哺乳類、齧歯類、又はヒト二本鎖デオキシリボ核酸ポリヌクレオチドである。いくつかの実施形態では、隣接配列は、配列番号350、352、355、356、359、361、362、及び367のうちのいずれか1つの少なくとも90個の連続するヌクレオチドと少なくとも約75%、少なくとも約80%、少なくとも約85%、少なくとも約86%、少なくとも約87%、少なくとも約88%、少なくとも約89%、少なくとも約90%、少なくとも約91%、少なくとも約92%、少なくとも約93%、少なくとも約94%、少なくとも約95%、少なくとも約96%、少なくとも約97%、少なくとも約98%、少なくとも約99%、又は100%の配列同一性を有する。いくつかの実施形態では、二本鎖デオキシリボ核酸ポリヌクレオチドは、カーゴ配列に隣接する別の隣接配列を含み、上記別の隣接配列は、配列番号350~454のうちのいずれか1つの少なくとも90個の連続するヌクレオチドと少なくとも約70%の配列同一性を有する。いくつかの実施形態では、別の隣接配列は、配列番号351、353、354、357、358、360、363、及び366のうちのいずれか1つの少なくとも90個の連続するヌクレオチドと少なくとも約75%、少なくとも約80%、少なくとも約85%、少なくとも約86%、少なくとも約87%、少なくとも約88%、少なくとも約89%、少なくとも約90%、少なくとも約91%、少なくとも約92%、少なくとも約93%、少なくとも約94%、少なくとも約95%、少なくとも約96%、少なくとも約97%、少なくとも約98%、少なくとも約99%、又は100%の配列同一性を有する。いくつかの実施形態では、隣接配列は、カーゴ核酸配列の左端に隣接し、別の隣接配列は、カーゴ核酸配列の右端に隣接する。いくつかの実施形態では、トランスポザーゼは、標的核酸遺伝子座に隣接する挿入モチーフを認識するように構成されている。いくつかの実施形態では、挿入モチーフは、配列AATGACの少なくとも3、4、5、又は6個の連続するヌクレオチドを含む。

【0029】

いくつかの態様では、本開示は、標的核酸遺伝子座を修飾する方法を提供し、上記方法は、本明細書に開示される操作されたトランスポザーゼ系を標的核酸遺伝子座に送達することを含み、トランスポザーゼは、カーゴヌクレオチド配列を標的核酸遺伝子座に転位するように構成されており、複合体は、複合体の標的核酸遺伝子座への結合時に、複合体が標的核酸遺伝子座を修飾するように構成されている。

【0030】

いくつかの実施形態では、標的核酸遺伝子座を修飾することは、標的核酸遺伝子座を結合、ニックング、切断、マーキング、修飾、又は転位することを含む。いくつかの実施形態では、標的核酸遺伝子座は、デオキシリボ核酸(DNA)を含む。いくつかの実施形態では、標的核酸遺伝子座は、ゲノムDNA、ウイルスDNA、又は細菌DNAを含む。いくつかの実施形態では、標的核酸遺伝子座は、インピトロである。いくつかの実施形態では、標的核酸遺伝子座は、細胞内にある。いくつかの実施形態では、細胞は、原核細胞、細菌細胞、真核細胞、真菌細胞、植物細胞、動物細胞、哺乳類細胞、齧歯類細胞、霊長類細胞、ヒト細胞、又は初代細胞である。いくつかの実施形態では、細胞は、初代細胞である。いくつかの実施形態では、初代細胞は、T細胞である。いくつかの実施形態では、初代細胞は、造血幹細胞(HSC)である。いくつかの実施形態では、操作されたトランスポザーゼ系を標的核酸遺伝子座に送達することは、トランスポザーゼをコードするオープンリーディングフレームを含む核酸を送達することを含む。いくつかの実施形態では、核酸は、トランスポザーゼをコードするオープンリーディングフレームが作動可能に連結されているプロモーターを含む。いくつかの実施形態では、操作されたトランスポザーゼ系を標的核酸遺伝子座に送達することは、トランスポザーゼをコードするオープンリーディングフレームを含有するキャッピングされたmRNAを送達することを含む。いくつかの実施形態では、操作されたトランスポザーゼ系を標的核酸遺伝子座に送達することは、翻訳されたポリペプチドを送達することを含む。いくつかの実施形態では、トランスポザー

ゼは、標的核酸遺伝子座で、又は標的核酸遺伝子座の近位で、一本鎖切断又は二本鎖切断を誘導する。いくつかの実施形態では、トランスポザーゼは、標的遺伝子座内又は標的遺伝子座の5'に互い違いの一本鎖切断を誘導する。

【0031】

いくつかの態様では、本開示は、操作されたトランスポザーゼ系を提供し、上記操作されたトランスポザーゼ系は、(a)カーゴヌクレオチド配列を含む二本鎖核酸であって、カーゴヌクレオチド配列がトランスポザーゼと相互作用するように構成されている、二本鎖核酸と、(b)トランスポザーゼであって、(i)カーゴヌクレオチド配列を標的核酸遺伝子座に転位するように構成され、(ii)未培養微生物に由来する、トランスポザーゼと、を含む。いくつかの実施形態では、カーゴヌクレオチド配列は、異種配列である。いくつかの実施形態では、カーゴヌクレオチド配列は、操作された配列である。いくつかの実施形態では、カーゴヌクレオチド配列は、生物中に存在する野生型ゲノム配列ではない。いくつかの実施形態では、トランスポザーゼは、配列番号1~349のうちのいずれか1つと少なくとも75%の配列同一性を有する配列を含む。いくつかの実施形態では、トランスポザーゼは、TnpAトランスポザーゼ又はTnpBトランスポザーゼではない。いくつかの実施形態では、トランスポザーゼは、TnpAトランスポザーゼと80%未満の配列同一性を有する。いくつかの実施形態では、トランスポザーゼは、TnpBトランスポザーゼと80%未満の配列同一性を有する。いくつかの実施形態では、トランスポザーゼは、触媒チロシン残基を含む。いくつかの実施形態では、トランスポザーゼは、サブ末端回文配列を含む左側領域及びサブ末端回文配列を含む右側領域に結合するように構成されている。いくつかの実施形態では、トランスポザーゼは、一本鎖デオキシリボ核酸ポリヌクレオチドとしてカーゴヌクレオチド配列を転位するように構成されている。いくつかの実施形態では、トランスポザーゼは、トランスポザーゼのN末端又はC末端の近位に1つ以上の核局在化配列(NLS)を含む。いくつかの実施形態では、NLSは、配列番号455~470からなる群からの配列と少なくとも80%同一である配列を含む。いくつかの実施形態では、配列同一性は、BLASTP、CLUSTALW、MUSCLE、MAFFT、又はSmith-Waterman相同性検索アルゴリズムのパラメータを用いるCLUSTALWによって決定される。いくつかの実施形態では、配列同一性は、3のワード長(W)、10の期待値(E)のパラメータ、及び11の存在、1の延長でギャップコストを設定しているBLOSUM62スコアリングマトリックスを使用し、条件付き組成スコアマトリックス調整を使用した、BLASTP相同性検索アルゴリズムによって決定される。

【0032】

いくつかの態様では、本開示は、操作されたトランスポザーゼ系を提供し、上記操作されたトランスポザーゼ系は、(a)カーゴヌクレオチド配列を含む二本鎖核酸であって、カーゴヌクレオチド配列がトランスポザーゼと相互作用するように構成されている、二本鎖核酸と、(b)トランスポザーゼであって、(i)カーゴヌクレオチド配列を標的核酸遺伝子座に転位するように構成され、(ii)配列番号1~349のうちのいずれか1つと少なくとも75%の配列同一性を有する配列を含む、トランスポザーゼと、を含む。いくつかの実施形態では、トランスポザーゼは、未培養微生物に由来する。いくつかの実施形態では、トランスポザーゼは、TnpAトランスポザーゼ又はTnpBトランスポザーゼではない。いくつかの実施形態では、トランスポザーゼは、TnpAトランスポザーゼと80%未満の配列同一性を有する。いくつかの実施形態では、トランスポザーゼは、TnpBトランスポザーゼと80%未満の配列同一性を有する。いくつかの実施形態では、トランスポザーゼは、触媒チロシン残基を含む。いくつかの実施形態では、トランスポザーゼは、サブ末端回文配列を含む左側領域及びサブ末端回文配列を含む右側領域に結合するように構成されている。いくつかの実施形態では、トランスポザーゼは、一本鎖デオキシリボ核酸ポリヌクレオチドとしてカーゴヌクレオチド配列を転位するように構成されている。いくつかの実施形態では、配列同一性は、BLASTP、CLUSTALW、MUSCLE、MAFFT、又はSmith-Waterman相同性検索アルゴリズムのパ

10

20

30

40

50

ラメーターを用いるCLUSTALWによって決定される。いくつかの実施形態では、配列同一性は、3のワード長(W)、10の期待値(E)のパラメーター、及び11の存在、1の延長でギャップコストを設定しているBLOSUM62スコアリングマトリックスを使用し、条件付き組成スコアマトリックス調整を使用した、BLASTP相同性検索アルゴリズムによって決定される。

【0033】

いくつかの態様では、本開示は、本明細書に記載される態様又は実施形態のうちのいずれか1つの操作されたトランスポザーゼ系をコードするデオキシリボ核酸ポリヌクレオチドを提供する。

【0034】

いくつかの態様では、本開示は、生物における発現に最適化された操作された核酸配列を含む核酸を提供し、核酸はトランスポザーゼをコードし、トランスポザーゼは未培養微生物に由来し、生物は未培養微生物ではない。いくつかの実施形態では、トランスポザーゼは、配列番号1~349のうちのいずれか1つと少なくとも75%の配列同一性を有するバリエーションを含む。いくつかの実施形態では、トランスポザーゼは、トランスポザーゼのN末端又はC末端の近位に1つ以上の核局在化配列(NLS)をコードする配列を含む。いくつかの実施形態では、NLSは、配列番号455~470から選択される配列を含む。いくつかの実施形態では、NLSは、配列番号456を含む。いくつかの実施形態では、NLSは、トランスポザーゼのN末端の近位にある。いくつかの実施形態では、NLSは、配列番号455を含む。いくつかの実施形態では、NLSは、トランスポザーゼのC末端の近位にある。いくつかの実施形態では、生物は、原核生物、細菌、真核生物、真菌、植物、哺乳類、齧歯類、又はヒトである。

【0035】

いくつかの態様では、本開示は、本明細書に記載される態様又は実施形態のうちのいずれか1つの核酸を含むベクターを提供する。いくつかの実施形態では、ベクターは、トランスポザーゼと複合体を形成するように構成されたカーゴヌクレオチド配列をコードする核酸を更に含む。いくつかの実施形態では、ベクターは、プラスミド、ミニサークル、CELiD、アデノ随伴ウイルス(AAV)由来ビリオン、又はレンチウイルスである。

【0036】

いくつかの態様では、本開示は、本明細書に記載される態様又は実施形態のうちのいずれか1つのうちのいずれか1つのベクターを含む細胞を提供する。

【0037】

いくつかの態様では、本開示は、本明細書に記載される態様又は実施形態のうちのいずれか1つの細胞を培養することを含む、トランスポザーゼを製造する方法を提供する。

【0038】

いくつかの態様では、本開示は、二本鎖デオキシリボ核酸ポリヌクレオチドを結合、ニックング、切断、マーキング、修飾、又は転位する方法を提供し、上記方法は、(a)二本鎖デオキシリボ核酸ポリヌクレオチドを、カーゴヌクレオチド配列を標的核酸遺伝子座に転位するように構成されたトランスポザーゼと接触させることを含み、トランスポザーゼは、配列番号1~349のうちのいずれか1つと少なくとも75%の配列同一性を有する配列を含む。いくつかの実施形態では、トランスポザーゼは、未培養微生物に由来する。いくつかの実施形態では、トランスポザーゼは、TnpAトランスポザーゼ又はTnpBトランスポザーゼではない。いくつかの実施形態では、トランスポザーゼは、TnpAトランスポザーゼと80%未満の配列同一性を有する。いくつかの実施形態では、トランスポザーゼは、TnpBトランスポザーゼと80%未満の配列同一性を有する。いくつかの実施形態では、トランスポザーゼは、触媒チロシン残基を含む。いくつかの実施形態では、トランスポザーゼは、サブ末端回文配列を含む左側領域及びサブ末端回文配列を含む右側領域に結合するように構成されている。いくつかの実施形態では、二本鎖デオキシリボ核酸ポリヌクレオチドは、一本鎖デオキシリボ核酸ポリヌクレオチドとして転位される。いくつかの実施形態では、二本鎖デオキシリボ核酸ポリヌクレオチドは、真核生物、植

10

20

30

40

50

物、真菌、哺乳類、齧歯類、又はヒト二本鎖デオキシリボ核酸ポリヌクレオチドである。

【0039】

いくつかの態様では、本開示は、標的核酸遺伝子座を修飾する方法を提供し、上記方法は、本明細書に記載される態様又は実施形態のうちのいずれか1つの操作されたトランスポザーゼ系を標的核酸遺伝子座に送達することを含み、トランスポザーゼは、カーゴヌクレオチド配列を標的核酸遺伝子座に転位するように構成されており、複合体は、複合体の標的核酸遺伝子座への結合時に、複合体が標的核酸遺伝子座を修飾するように構成されている。いくつかの実施形態では、標的核酸遺伝子座を修飾することは、標的核酸遺伝子座を結合、ニックング、切断、マーキング、修飾、又は転位することを含む。いくつかの実施形態では、標的核酸遺伝子座は、デオキシリボ核酸(DNA)を含む。いくつかの実施形態では、標的核酸遺伝子座は、ゲノムDNA、ウイルスDNA、又は細菌DNAを含む。いくつかの実施形態では、標的核酸遺伝子座は、インピトロである。いくつかの実施形態では、標的核酸遺伝子座は、細胞内にある。いくつかの実施形態では、細胞は、原核細胞、細菌細胞、真核細胞、真菌細胞、植物細胞、動物細胞、哺乳類細胞、齧歯類細胞、霊長類細胞、ヒト細胞、又は初代細胞である。いくつかの実施形態では、細胞は、初代細胞である。いくつかの実施形態では、初代細胞は、T細胞である。いくつかの実施形態では、初代細胞は、造血幹細胞(HSC)である。いくつかの実施形態では、操作されたトランスポザーゼ系を標的核酸遺伝子座に送達することは、本明細書に記載される態様若しくは実施形態のうちのいずれか1つの核酸、又は本明細書に記載される態様若しくは実施形態のうちのいずれか1つのベクターを送達することを含む。いくつかの実施形態では、操作されたトランスポザーゼ系を標的核酸遺伝子座に送達することは、トランスポザーゼをコードするオープンリーディングフレームを含む核酸を送達することを含む。いくつかの実施形態では、核酸は、トランスポザーゼをコードするオープンリーディングフレームが作動可能に連結されているプロモーターを含む。いくつかの実施形態では、操作されたトランスポザーゼ系を標的核酸遺伝子座に送達することは、トランスポザーゼをコードするオープンリーディングフレームを含有するキャッピングされたmRNAを送達することを含む。いくつかの実施形態では、操作されたトランスポザーゼ系を標的核酸遺伝子座に送達することは、翻訳されたポリペプチドを送達することを含む。いくつかの実施形態では、トランスポザーゼは、標的核酸遺伝子座で、又は標的核酸遺伝子座の近位で、一本鎖切断又は二本鎖切断を誘導する。いくつかの実施形態では、トランスポザーゼは、標的遺伝子座内又は標的遺伝子座の5'に互い違いの一本鎖切断を誘導する。

10

20

30

【0040】

いくつかの態様では、本開示は、配列番号1~349のうちのいずれか1つ又はそのバリエーションと少なくとも75%の配列同一性を有する異種トランスポザーゼをコードするオープンリーディングフレームを含む、宿主細胞を提供する。いくつかの実施形態では、トランスポザーゼは、配列番号1、3、5、7、9、11、13、15、又は16のうちのいずれか1つと少なくとも75%の配列同一性を有する。いくつかの実施形態では、トランスポザーゼは、配列番号2、4、6、8、10、12、14、又は17のうちのいずれか1つと少なくとも75%の配列同一性を有する。いくつかの実施形態では、宿主細胞は、E.coli細胞である。いくつかの実施形態では、E.coli細胞は、DE3リソゲンであるか、又はE.coli細胞は、BL21(DE3)株である。いくつかの実施形態では、E.coli細胞は、ompT lon遺伝子型を有する。いくつかの実施形態では、オープンリーディングフレームは、T7プロモーター配列、T7-lacプロモーター配列、lacプロモーター配列、tacプロモーター配列、trcプロモーター配列、ParaBADプロモーター配列、PhaBADプロモーター配列、T5プロモーター配列、cspAプロモーター配列、araPBADプロモーター、ファージラムダからの強い左向きプロモーター(pLプロモーター)、又はそれらの任意の組み合わせに作動可能に連結されている。いくつかの実施形態では、オープンリーディングフレームは、トランスポザーゼをコードする配列にインフレームで連結された親和性タグをコードする配列を含む。いくつかの実施形態では、親和性タグは、固定化金属親和性クロマトグラ

40

50

フィー（IMAC）タグである。いくつかの実施形態では、IMACタグは、ポリヒスチジンタグである。いくつかの実施形態では、親和性タグは、mycタグ、ヒトインフルエンザヘマグルチニン（HA）タグ、マルトース結合タンパク質（MBP）タグ、グルタチオンS-トランスフェラーゼ（GST）タグ、ストレプトアビジンタグ、FLAGタグ、又はそれらの任意の組み合わせである。いくつかの実施形態では、親和性タグは、プロテアーゼ切断部位をコードするリンカー配列を介して、トランスポザーゼをコードする配列にインフレームで連結されている。いくつかの実施形態では、プロテアーゼ切断部位は、タバコエッチウイルス（TEV）プロテアーゼ切断部位、PreScission（登録商標）プロテアーゼ切断部位、トロンピン切断部位、第Xa因子切断部位、エンテロキナーゼ切断部位、又はそれらの任意の組み合わせである。いくつかの実施形態では、オープンリーディングフレームは、宿主細胞における発現のためにコドン最適化される。いくつかの実施形態では、オープンリーディングフレームは、ベクター上に提供される。いくつかの実施形態では、オープンリーディングフレームは、宿主細胞のゲノムに組み込まれている。

10

【0041】

いくつかの態様では、本開示は、適合する液体培地中に、本明細書に記載される態様又は実施形態のうちのいずれか1つの宿主細胞を含む、培養物を提供する。

【0042】

いくつかの態様では、本開示は、適合する成長培地中で、本明細書に記載される態様又は実施形態のうちのいずれか1つの宿主細胞を培養することを含む、トランスポザーゼを産生する方法を提供する。いくつかの実施形態では、方法は、追加の化学剤又は増加された量の栄養素を添加することによって、トランスポザーゼの発現を誘導することを更に含む。いくつかの実施形態では、追加の化学剤又は増加された量の栄養素は、イソプロピル-D-1-チオガラクトピラノシド（IPTG）又は追加の量のラクトースを含む。いくつかの実施形態では、方法は、培養後に宿主細胞を単離することと、宿主細胞を溶解してタンパク質抽出物を産生することとを更に含む。いくつかの実施形態では、方法は、タンパク質抽出物をIMAC、又はイオン親和性クロマトグラフィーに供することを更に含む。いくつかの実施形態では、オープンリーディングフレームは、トランスポザーゼをコードする配列にインフレームで連結されたIMAC親和性タグをコードする配列を含む。いくつかの実施形態では、IMAC親和性タグは、プロテアーゼ切断部位をコードするリンカー配列を介して、トランスポザーゼをコードする配列にインフレームで連結されている。いくつかの実施形態では、プロテアーゼ切断部位は、タバコエッチウイルス（TEV）プロテアーゼ切断部位、PreScission（登録商標）プロテアーゼ切断部位、トロンピン切断部位、第Xa因子切断部位、エンテロキナーゼ切断部位、又はそれらの任意の組み合わせを含む。いくつかの実施形態では、方法は、プロテアーゼ切断部位に対応するプロテアーゼをトランスポザーゼと接触させることによって、IMAC親和性タグを切断することを更に含む。いくつかの実施形態では、方法は、サブトラクティブIMAC親和性クロマトグラフィーを実施して、トランスポザーゼを含む組成物から親和性タグを除去することを更に含む。

20

30

【0043】

いくつかの態様では、本開示は、細胞中の遺伝子座を破壊する方法を提供し、上記方法は、細胞に組成物を接触させることを含み、組成物は、(a)カーゴヌクレオチド配列を含む二本鎖核酸であって、カーゴヌクレオチド配列が、トランスポザーゼと相互作用するように構成されている、二本鎖核酸と、(b)トランスポザーゼであって、(i)カーゴヌクレオチド配列を標的核酸遺伝子座に転位するように構成され、(ii)配列番号1~349のうちのいずれか1つと少なくとも75%の配列同一性を有する配列を含み、(iii)細胞内でTnpAトランスポザーゼと少なくとも同等の転位活性を有する、トランスポザーゼと、を含む。いくつかの実施形態では、転位活性は、標的核酸遺伝子座を含む細胞にトランスポザーゼを導入し、細胞内の標的核酸遺伝子座の転位を検出することによって、インビトロで測定される。いくつかの実施形態では、組成物は、20pmol以

40

50

下のトランスポザーゼを含む。いくつかの実施形態では、組成物は、1 pmol以下のトランスポザーゼを含む。

【0044】

本開示の更なる態様及び利点は、以下の詳細な説明から、当業者に容易に明らかになり、ここで、本開示の例示的な実施形態のみが示され、記載される。理解されるように、本開示は、他の異なる実施形態をすることができ、そのいくつかの詳細は、全て本開示から逸脱することなく、様々な明白な点において変更することができる。したがって、図面及び説明は、本質的に例示とみなされるべきであり、制限としてみなされるべきではない。

【0045】

参照による組み込み

本明細書において言及される全ての刊行物、特許、及び特許出願は、それぞれ個々の刊行物、特許、又は特許出願が、参照により組み込まれるべきことが具体的かつ個別に示されているのと同じ程度に、参照により本明細書に組み込まれる。

【図面の簡単な説明】

【0046】

本発明の新規の特徴は、添付の特許請求の範囲に特記して記載される。本発明の特徴及び利点のより良好な理解は、本発明の原理が利用される例示的な実施形態を記載する以下の詳細な説明、及び添付の図面を参照することによって得られるだろう。

【0047】

【図1A】図1A及び図1Bは、MGトランスポザーゼを示す。図1Aは、チロシン(Y1)トランスポザーゼMG92-1遺伝子座を含むトランスポゾンの組織を示す。MG92-1は、トランスポゾンの5'端でコードされ、その後アクセサリ-転位タンパク質TnpB及び他のカーゴが続く。トランスポゾン端は、16~17bpの直接反復を含有し、それらは、転位活性に関与し得る二次構造を示す。図1Bは、MG Y1トランスポザーゼホモログの複数の配列アライメントを示す。触媒残基HUH及びYは、コンセンサス配列上及びMSA(ボックス)上で強調表示されている。

【図1B】図1A及び図1Bは、MGトランスポザーゼを示す。図1Aは、チロシン(Y1)トランスポザーゼMG92-1遺伝子座を含むトランスポゾンの組織を示す。MG92-1は、トランスポゾンの5'端でコードされ、その後アクセサリ-転位タンパク質TnpB及び他のカーゴが続く。トランスポゾン端は、16~17bpの直接反復を含有し、それらは、転位活性に関与し得る二次構造を示す。図1Bは、MG Y1トランスポザーゼホモログの複数の配列アライメントを示す。触媒残基HUH及びYは、コンセンサス配列上及びMSA(ボックス)上で強調表示されている。

【図2】は、TnpAタンパク質配列の系統樹を示す。この樹は、ここで回収された414の新規TnpA配列(黒い点)及び19の参照TnpA配列(灰色の点)の複数の配列アライメントから構築された。参照配列の標識が含まれた。

【図3】は、例示的な挿入配列IS200/IS605 MG92-28を示す。上のパネル：TnpA様トランスポザーゼ及びその関連TnpB様遺伝子をコードするMG92-28挿入配列のゲノムコンテキスト。両方の遺伝子は、共分散モデルから予測されたLE及びRE(ボックス)に隣接している。下のパネル：LE(左上)及びRE(右下)は、挿入配列の境界を描写している。共分散モデルによって予測された領域は、配列の下に矢印として注釈付けされている。LE及びREの二次構造は、各端について示されている。

【図4】は、PureExpressで発現されたTnpA様タンパク質のウェスタンブロットを示す。レーンは、ラダー、1:HpTnpA、2:HhTpA、3:92-2、4:92-3、5:92-4、6:92-5、7:92-6、8:92-7、9:92-8、10:92-10、11:92-11である。HpTnpA及びHhTpAは、それぞれH.pylori及びH.Heilmanniからの陽性対照である。分子量は、17~23キロダルトン(kDa)の範囲である。

【図5A】は、転位反応のLEのPCR産物を示す。カーゴが指定された対照レーンを除

10

20

30

40

50

き、全ての反応物は、タンパク質及びその対の特定のカーゴを有する。レーンは、1：ラダー、2：HpTnpAカーゴを有する陰性対照NTC、3：92-1、4：92-2、5：92-3、6：92-4、7：92-5、8：92-6、9：92-7、10：92-8、11：92-10、12：92-11、13：HpTnpA、14；HhTnpAである。予想される転位産物は、LEサイズに応じて200～300bpの範囲であり得、矢印でマークされている。92-5における<200bpのバンドは、非特異的プライマー相互作用に関連する。

【図5B】は、転位反応のREのPCR産物を示す。カーゴが指定された対照レーンを除き、全ての反応物は、タンパク質及びその対の特定のカーゴを有する。レーンは、1：HpTnpAカーゴを有するNTC、2：92-1、3：92-2、4：92-3、5：92-4、6：92-5、7：92-6、8：92-7、9：92-8、10：92-10、11：92-11、12：HpTnpA、13；HhTnpA、及び14：ラダーである。予想される転位産物は、REサイズに応じて300～500bpの範囲であり得、矢印でマークされている。8N領域に生じる転位は、隣接配列への転位よりもはるかに弱いバンドを有するため、薄いバンドが予想される。

【図6】は、MG92-3の転位を確認するサンガーシーケンシングデータを示す。クロマトグラムトレースは、カーゴ配列にマッピングされて示され、影付き文字はカーゴと一致する。切断点（矢印）では、トレースは代わりに標的配列（ボックス）上にマッピングされている。標的の分析により、LEと標的との間で共有される配列である挿入モチーフが明らかになる。隣接する非標準塩基相互作用を有する下流のヘアピンを特定することができる。

【図7】は、MG92-3の転位を確認するサンガーシーケンシングデータを示す。クロマトグラムトレースは、カーゴにマッピングされて示され、影付き文字はカーゴと一致する。切断点（矢印）では、トレースは代わりに標的配列（ボックス）上にマッピングされている。標的の分析により、挿入モチーフが明らかになる。推定REにおける切断位置は、REの境界を画定し、これは、TnpA認識及び鎖切断（点線のボックスの差し込み図）を可能にするために、標準ヘアピンに折り畳まれる。

【図8】は、ブレイクポイントを決定するために分析されたカーゴ及び標的配列接続を示すキメラNGSリードの分析を示す。x軸は、カーゴ配列に沿った位置であり、y軸は、その位置で遷移するリードの数である。カーゴ上の2030ntでのブレイクポイントにおける特定されたピークは、サンガーシーケンシングで特定されたブレイクポイントと一致し、LE切断の位置が確認されている。

【図9】は、MG92-4の転位を確認するNGSシーケンシングデータを示す。NGSリードは、標的にマッピングされて示され、薄い影付き文字はカーゴと一致する。切断点（矢印）では、トレースは代わりにカーゴ配列（ボックス）上にマッピングされている。推定REにおける切断位置は、REの境界を画定し、これは、TnpA認識及び鎖切断（点線のボックスの差し込み図）を可能にするために、標準ヘアピンに折り畳まれる。NGSリードヒストグラムは、カーゴ上のこのブレイクポイントに対応するリードの頻度を示す。

【0048】

配列表の簡単な説明

本明細書とともに提出された配列表は、本開示による方法、組成物、及び系で使用するための例示的なポリヌクレオチド配列及びポリペプチド配列を提供する。以下は、その中の配列の例示的な説明である。

MG92

【0049】

配列番号1～349は、MG92転位タンパク質の完全長ペプチド配列を示す。

【0050】

配列番号350～454は、MG92トランスポゾン端の完全長ペプチド配列を示す。

核局在化配列

【0051】

配列番号455～470は、本明細書に記載されるMG92転位タンパク質との使用に好適な核局在化配列(NLS)の完全長ペプチド配列を示す。

【発明を実施するための形態】

【0052】

本発明の様々な実施形態は本明細書に示され、記載されるが、そのような実施形態が、例示の目的でのみ提供されることは、当業者には明らかであろう。多数の変形、変更、及び置換は、本発明から逸脱することなく、当業者にとって想到し得るものである。本明細書に記載される本発明の実施形態に対する様々な代替が用いられ得ることは、理解されるべきである。

10

【0053】

本明細書に開示されるいくつかの方法の実践は、別段の示唆がない限り、免疫学、生化学、化学、分子生物学、微生物学、細胞生物学、ゲノミクス、及び組換えDNAの技術を利用する。例えば、Sambrook and Green, *Molecular Cloning: A Laboratory Manual*, 4th Edition (2012); the series *Current Protocols in Molecular Biology* (F. M. Ausubel, et al. eds.); the series *Methods In Enzymology* (Academic Press, Inc.), *PCR 2: A Practical Approach* (M. J. MacPherson, B. D. Hames and G. R. Taylor eds. (1995)), Harlow and Lane, eds. (1988) *Antibodies, A Laboratory Manual*, and *Culture of Animal Cells: A Manual of Basic Technique and Specialized Applications*, 6th Edition (R. I. Freshney, ed. (2010)) (参照により本明細書に完全に組み込まれる)を参照のこと。

20

【0054】

本明細書で使用される場合、単数形「a」、「an」及び「the」は、文脈が別途明確に示さない限り、複数形も含むことが意図される。更に、用語「含むこと」、「含む」、「有すること」、「有する」、「有する」、又はそのバリエーションが、詳細な説明及び/又は特許請求の範囲のいずれかで使用される限りにおいて、かかる用語は、用語「含むこと」と類似した様式で包含的であることが意図される。

30

【0055】

用語「約」又は「およそ」は、当業者によって決定される特定の値についての許容可能な誤差範囲内であることを意味し、これは、値がどのように測定又は決定されるか、すなわち、測定系の制限に部分的に依存する。例えば、「約」は、当該技術分野の慣行によると、1又は2つ以上の標準偏差内を意味し得る。あるいは、「約」は、所与の値の最大20%、最大15%、最大10%、最大5%、又は最大1%の範囲を意味し得る。

【0056】

本明細書で使用される場合、「細胞」は概して生物学的細胞を指す。細胞は、生きている生物の基本的な構造、機能、及び/又は生物学的単位であり得る。細胞は、1つ以上の細胞を有する任意の生物を起源とし得る。いくつかの非限定的な例としては、原核細胞、真核生物の細胞、細菌細胞、古細菌細胞、単一細胞の真核生物の細胞、原虫細胞、植物由来の細胞(例えば、植物作物、果実、野菜、穀物、大豆、トウモロコシ、トウモロコシ、小麦、種子、トマト、米、キャッサバ、サトウキビ、カボチャ、乾草、ジャガイモ、綿、大麻、タバコ、開花している植物、針葉樹、ジムノスパーム、シダ、ヒカゲノカズラ、ツノゴケ、コケ植物、コケ由来の細胞)、藻類の細胞(例えば、*Botryococcus braunii*、*Chlamydomonas reinhardtii*、*Nannochloropsis gaditana*、*Chlorella pyrenoidosa*、*Sargassum patens* C. Agardhなど)、海藻(例えば、ケルブ

40

50

)、真菌細胞(例えば、酵母細胞、キノコ由来の細胞)、動物細胞、脊椎動物(例えば、フルーツフライ、クニダリアン、エキノデルム、線虫など)由来の細胞、脊椎動物(例えば、魚、両生類、爬虫類、鳥類、哺乳類)由来の細胞、哺乳類(例えば、ブタ、ウシ、ヤギ、ヒツジ、齧歯類、ラット、マウス、非ヒト霊長類、ヒトなど)由来の細胞などが挙げられる。いくつかの場合では、細胞は、天然の生物に由来するものではない(例えば、細胞は、合成的に作製されてもよく、時には人工細胞と呼ばれることがある)。

【0057】

本明細書で使用される場合、用語「ヌクレオチド」は、概して、塩基-糖-リン酸の組み合わせを指す。ヌクレオチドは、合成ヌクレオチドを含んでもよい。ヌクレオチドは、合成ヌクレオチド類似体を含んでもよい。ヌクレオチドは、核酸配列(例えば、デオキシリボ核酸(DNA)及びリボ核酸(RNA))の単量体単位であってもよい。ヌクレオチドという用語は、リボヌクレオシド三リン酸アデノシン三リン酸(ATP)、ウリジン三リン酸(UTP)、シトシン三リン酸(CTP)、グアノシン三リン酸(GTP)及びデオキシリボヌクレオシド三リン酸、例えば、dATP、dCTP、dITP、dUTP、dGTP、dTTP、又はその誘導体を含み得る。かかる誘導体としては、例えば、[S]dATP、7-デアザ-dGTP及び7-デアザ-dATP、並びにそれらを含む核酸分子にヌクレアーゼ耐性を付与するヌクレオチド誘導体を挙げることができる。本明細書で使用される場合、ヌクレオチドという用語は、ジデオキシリボヌクレオシド三リン酸(ddNTP)及びその誘導体を指す場合がある。ジデオキシリボヌクレオシド三リン酸の例としては、ddATP、ddCTP、ddGTP、ddITP、及びddTTPが挙げられるが、これらに限定されない。ヌクレオチドは、光学的に検出可能な部分(例えば、フルオロフォア)を含む部分を使用するなど、非標識又は検出可能に標識されてもよい。標識はまた、量子ドットを用いて行われてもよい。検出可能な標識としては、例えば、放射性同位元素、蛍光標識、化学発光標識、生物発光標識、及び酵素標識を挙げることができる。ヌクレオチドの蛍光標識は、フルオレセイン、5-カルボキシフルオレセイン(FAM)、2'-ジメトキシ-4'-ジクロロ-6-カルボキシフルオレセイン(JOE)、ローダミン、6-カルボキシローダミン(R6G)、N,N,N',N'-テトラメチル-6-カルボキシローダミン(TAMRA)、6-カルボキシ-X-ローダミン(ROX)、4-(4'-ジメチルアミノフェニルアゾ)安息香酸(DABCYL)、カスケードブルー、オレゴングリーン、テキサスレット、シアニン及び5-(2'-アミノエチル)アミノナフタレン-1-スルホン酸(EDANS)を含むが、これらに限定されない。蛍光標識されたヌクレオチドの具体的な例としては、Perkin Elmer、Foster City、Califから入手可能な[R6G]dUTP、[TAMRA]dUTP、[R110]dCTP、[R6G]dCTP、[TAMRA]dCTP、[JOE]ddATP、[R6G]ddATP、[FAM]ddCTP、[R110]ddCTP、[TAMRA]ddGTP、[ROX]ddTTP、[dR6G]ddATP、[dR110]ddCTP、[dTAMRA]ddGTP、及び[dROX]ddTTP; Amersham、Arlington Heights、ILから入手可能なフルオロ結合デオキシヌクレオチド、フルオロ結合Cy3-dCTP、フルオロ結合Cy5-dCTP、フルオロ結合フルオロX-dCTP、フルオロ結合Cy3-dUTP、及びフルオロ結合Cy5-dUTP; Boehringer Mannheim、Indianapolis、Indから入手可能なフルオレセイン-15-dATP、フルオレセイン-12-dUTP、テトラメチル-ローダミン-6-dUTP、IR770-9-dATP、フルオレセイン-12-ddUTP、フルオレセイン-12-UTP、及びフルオレセイン-15-2'-dATP; 並びにMolecular Probes、Eugene、Oregから入手可能な染色体標識ヌクレオチド、BODIPY-FL-14-UTP、BODIPY-FL-4-UTP、BODIPY-TMR-14-UTP、BODIPY-TMR-14-dUTP、BODIPY-TR-14-UTP、BODIPY-TR-14-dUTP、カスケードブルー-7-UTP、カスケードブルー-7-dUTP、フルオレセイン-12-UTP、フルオレセイン-12-dUTP、オレゴング

10

20

30

40

50

リン488-5-dUTP、ローダミングリン-5-UTP、ローダミングリン-5-dUTP、テトラメチルローダミン-6-UTP、テトラメチルローダミン-6-dUTP、テキサスレッド-5-UTP、テキサスレッド-5-dUTP、及びテキサスレッド-12-dUTPを挙げることができる。ヌクレオチドはまた、化学修飾によって標識又はマーキングされてもよい。化学修飾された単一ヌクレオチドは、ビオチン-dNTPであり得る。ビオチン化dNTPのいくつかの非限定的な例としては、ビオチン-dATP（例えば、ビオ-N6-ddATP、ビオチン-14-dATP）、ビオチン-dCTP（例えば、ビオチン-11-dCTP、ビオチン-14-dCTP）、及びビオチン-dUTP（例えば、ビオチン-11-dUTP、ビオチン-16-dUTP、ビオチン-20-dUTP）が挙げられる。

10

【0058】

用語「ポリヌクレオチド」、「オリゴヌクレオチド」、及び「核酸」は、概して、一本鎖、二本鎖、又は多本鎖のいずれかの形態の、デオキシリボヌクレオチド若しくはリボヌクレオチド、又はその類似体のいずれかの、任意の長さのヌクレオチドのポリマー形態を指すように互換的に使用される。ポリヌクレオチドは、細胞にとって外因性又は内因性であってもよい。ポリヌクレオチドは、無細胞環境に存在してもよい。ポリヌクレオチドは、遺伝子又はその断片であってもよい。ポリヌクレオチドは、DNAであってもよい。ポリヌクレオチドは、RNAであってもよい。ポリヌクレオチドは、任意の三次元構造を有してもよく、任意の機能を発揮してもよい。ポリヌクレオチドは、1つ以上の類似体（例えば、改変された骨格、糖、又は核酸塩基）を含んでもよい。存在する場合、ヌクレオチド構造に対する修飾は、ポリマーのアセンブリの前又は後に付与されてもよい。類似体のいくつかの非限定的な例としては、5-プロモウラシル、ペプチド核酸、異種核酸、モルホリノ、ロックド核酸、グリコール核酸、トレオース核酸、ジデオキシヌクレオチド、コーディセピン、7-デアザ-GTP、フルオロフォア（例えば、糖に結合したローダミン又はフルオレセイン）、チオール含有ヌクレオチド、ビオチン結合ヌクレオチド、蛍光塩基類似体、CpGアイランド、メチル-7-グアノシン、メチル化ヌクレオチド、イノシン、チオウリジン、シュードウリジン、ジヒドロウリジン、クエオシン、及びワイオシンが挙げられる。ポリヌクレオチドの非限定的な例としては、遺伝子又は遺伝子断片のコード又は非コード領域、結合分析から定義した遺伝子座（遺伝子座）、エクソン、イントロン、メッセンジャーRNA（mRNA）、トランスファーRNA（tRNA）、リボソームRNA（rRNA）、短い干渉RNA（siRNA）、短いヘアピンRNA（shRNA）、マイクロ-RNA（miRNA）、リボザイム、cDNA、組換えポリヌクレオチド、分岐ポリヌクレオチド、プラスミド、ベクター、任意の配列の単離されたDNA、任意の配列の単離されたRNA、細胞を含まないDNA（cfDNA）及び細胞を含まないRNA（cfRNA）を含む細胞を含まないポリヌクレオチド、核酸プローブ、並びにプライマーが挙げられる。ヌクレオチドの配列は、非ヌクレオチド成分によって中断され得る。

20

30

【0059】

用語「トランスフェクション」又は「トランスフェクトされた」は、概して、非ウイルス又はウイルスベースの方法による細胞内への核酸の導入を指す。核酸分子は、完全なタンパク質又はその機能的部分をコードする遺伝子配列であってもよい。例えば、Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, 18.1-18.88（参照により本明細書に完全に組み込まれる）を参照のこと。

40

【0060】

用語「ペプチド」、「ポリペプチド」、及び「タンパク質」は、本明細書において互換的に使用され、概して、ペプチド結合によって結合された少なくとも2つのアミノ酸残基のポリマーを指す。この用語は、ポリマーの特定の長さを意味しておらず、ペプチドが組換え技術、化学若しくは酵素合成を使用して産生されるか、又は天然に存在するかを暗示又は区別することを意図するものではない。この用語は、天然に存在するアミノ酸ポリマ

50

ー並びに少なくとも1つの修飾アミノ酸を含むアミノ酸ポリマーに適用する。いくつかの実施形態では、ポリマーは、非アミノ酸によって中断されてもよい。この用語は、完全長タンパク質を含む任意の長さのアミノ酸鎖、並びに二次及び/若しくは三次の構造（例えば、ドメイン）を有する又は有さないタンパク質を含む。用語はまた、例えば、ジスルフィド結合形成、グリコシル化、脂質形成、アセチル化、リン酸化、酸化、及び標識成分とのコンジュゲーションなどの任意の他の操作によって修飾されたアミノ酸ポリマーを包含する。本明細書で使用される場合、「アミノ酸」及び「複数のアミノ酸」という用語は、概して、修飾アミノ酸及びアミノ酸類似体を含むが、これに限定されない天然及び非天然アミノ酸を指す。修飾アミノ酸は、天然アミノ酸及び非天然アミノ酸を含んでもよく、これは、アミノ酸上に天然に存在しない基又は化学的部分を含むように化学的に修飾されている。アミノ酸類似体は、アミノ酸誘導体を指す場合がある。用語「アミノ酸」は、D - アミノ酸とL - アミノ酸の両方を含む。

10

【0061】

本明細書で使用される場合、「非天然」は、概して、天然の核酸又はタンパク質では見られない核酸又はポリペプチド配列を指すことができる。非天然は、親和性タグを指してもよい。非天然は、融合物を指してもよい。非天然は、変異、挿入、及び/又は欠失を含む、天然に存在する核酸又はポリペプチド配列を指してもよい。非天然配列は、非天然配列が融合される核酸配列及び/又はポリペプチド配列によっても呈され得る活性（例えば、酵素活性、メチルトランスフェラーゼ活性、アセチルトランスフェラーゼ活性、キナーゼ活性、ユビキチン化活性など）を示し得、かつ/又はコードし得る。非天然核酸又はポリペプチド配列を、遺伝子操作によって天然に生じる核酸及び/又はポリペプチド配列（若しくはそのバリエーション）に連結して、キメラ核酸又はポリペプチドをコードするキメラ核酸及び/又はポリペプチド配列を生成してもよい。

20

【0062】

本明細書で使用される場合、用語「プロモーター」は、概して、遺伝子の転写又は発現を制御し、RNA転写が開始されるヌクレオチドのヌクレオチド又はヌクレオチドの領域に隣接するか、又は重複して位置し得る調節DNA領域を指す。プロモーターは、しばしば転写因子と呼ばれるタンパク質因子に結合する特定のDNA配列を含有してもよく、これは、RNAポリメラーゼのDNAへの結合を促進し、これにより、遺伝子転写をもたらす。「コアプロモーター」とも呼ばれる「基礎プロモーター」は、概して、作動可能に連結されたポリヌクレオチドの転写発現を促進するための全てのベーシックエレメントを含有するプロモーターを指してもよい。いくつかの実施形態では、真核生物の基礎プロモーターは、TATA - ボックス及び/又はCAATボックスを含有する。

30

【0063】

本明細書で使用される場合、用語「発現」は、概して、核酸配列又はポリヌクレオチドがDNA鋳型から（例えば、mRNA又は他のRNA転写物に）転写されるプロセス、及び/又は転写されたmRNAが続いてペプチド、ポリペプチド、又はタンパク質に翻訳されるプロセスを指す。転写物及びコードされたポリペプチドは、「遺伝子産物」と総称され得る。ポリヌクレオチドがゲノムDNAに由来する場合、発現は、真核生物の細胞におけるmRNAのスプライシングを含む。

40

【0064】

本明細書で使用される場合、「作動可能に連結された」、「作動可能な連結」、「作動可能に連結された」、又はその文法的な均等物は、概して、遺伝子エレメント、例えば、プロモーター、エンハンサー、ポリアデニル化配列などの並列化を指し、ここで、エレメントは、それらが予期される様式で作動することを可能にする関係にある。例えば、プロモーター配列及び/又はエンハンサー配列を含み得る、調節エレメントは、調節エレメントが、コード配列の転写を開始するのを助ける場合、コード領域に作動可能に連結される。この機能的関係が維持される限り、調節エレメントとコード領域との間に介在する残基があってもよい。

【0065】

50

本明細書で使用される場合、「ベクター」は、概して、ポリヌクレオチドを含むか、又はポリヌクレオチドと会合する高分子又は高分子の会合を指し、ポリヌクレオチドの細胞への送達を媒介するために使用され得る。ベクターの例としては、プラスミド、ウイルスベクター、リポソーム、及び他の遺伝子送達ビヒクルが挙げられる。ベクターは、概して、標的中の遺伝子の発現を促進するために遺伝子に作動可能に連結された、遺伝子エレメント、例えば、調節エレメントを含む。

【0066】

本明細書で使用される場合、「発現カセット」及び「核酸カセット」は、一緒に発現されるか、又は発現のために作動可能に連結される核酸配列又はエレメントの組み合わせを指すために概して互換的に使用される。いくつかの実施形態では、発現カセットは、調節エレメントと、それらが発現のため作動可能に連結されている遺伝子又は複数の遺伝子との組み合わせを指す。

10

【0067】

DNA又はタンパク質配列の「機能的断片」は、概して、完全長DNA又はタンパク質配列の生物学的活性と実質的に類似した生物学的活性（機能的又は構造的のいずれか）を保持する断片を指す。DNA配列の生物学的活性は、完全長配列に起因する様式で発現に影響を与える能力であり得る。

【0068】

本明細書で使用される場合、「操作された」物体は、概して、物体がヒトの介入によって修飾されたことを示す。非限定的な実施例によれば、核酸は、その配列を、天然では生じない配列に改変することによって修飾されてもよく、核酸は、ライゲーションされた産物が、オリジナルの核酸に存在しない機能を有するように、天然では関連しない核酸にライゲーションすることによって修飾されてもよく、操作された核酸は、天然では存在しない配列を用いてインビトロで合成されてもよく、タンパク質は、天然では存在しない配列にそのアミノ酸配列を変更することによって修飾されてもよく、操作されたタンパク質は、新しい機能又は特性を獲得してもよい。「操作された」系は、少なくとも1つの操作された成分を含む。

20

【0069】

本明細書で使用される場合、「合成」及び「人工」は概して、天然に存在するヒトタンパク質と低い配列同一性（例えば、50%未満の配列同一性、25%未満の配列同一性、10%未満の配列同一性、5%未満の配列同一性、1%未満の配列同一性）を有するタンパク質又はそのドメインを指すために互換的に使用され得る。例えば、VPRドメイン及びVP64ドメインは、合成トランス活性化ドメインである。

30

【0070】

本明細書で使用される場合、用語「転位因子」は、ゲノム内のある位置から別の位置に移動することができる（すなわち、それらは「転位」できる）DNA配列を指す。転位因子は、概して2つのクラスに分けることができる。クラスI転位因子、又は「レトロトランスポゾン」は、RNA中間体の転写及び翻訳を介して転位され、その後、逆転写（逆転写酵素によって媒介されるプロセス）を介してゲノム内にその新しい位置に再び組み込まれる。クラスII転位因子、又は「DNAトランスポゾン」は、両側にトランスポザーゼが隣接する一本鎖又は二本鎖DNAの複合体を介して転位される。この酵素ファミリーの更なる特徴は、例えば、Nature Education 2008, 1(1), 204、及びGenome Biology 2018, 19(199), 1-12に見出すことができ、その各々は参照により本明細書に組み込まれる。

40

【0071】

本明細書で使用される場合、用語「TnpA」は、概して、IS200/IS605細菌挿入配列（「IS」）ファミリーのメンバーに見られるトランスポザーゼを指す。二本鎖DNA中間体を介してDNA転位を実行する他の記録されたISトランスポザーゼとは異なり、TnpAは、一本鎖DNA中間体を介して進行する。TnpAはまた、末端逆位反復ではなく隣接するサブ末端回文配列を含有するという点で、他の記録されたISトラ

50

ンスポザーゼとは異なる。更に、TnpAは、標的部位の重複なしに、特定のATリッチのテトラヌクレオチド又はペンタヌクレオチドの3'を挿入する。最後に、TnpAは、他のISトランスポザーゼの「DDE」スーパーファミリーではなく、酵素のHis-疎水性-His(「HuH」)スーパーファミリーに属する。本明細書で使用される場合、「TnpB」は、概して、IS200/IS605細菌においてTnpAと並んで見出される、記録されていない機能(ただし、転位において調節的役割を果たすと推測される)の酵素を指す。IS200/IS605トランスポザーゼは、「Y1トランスポザーゼ」であり、それらが単一の触媒チロシン残基を含む単一ドメインタンパク質であることを意味する。本明細書で使用される場合、用語「TnpA様」は、概して、TnpAタンパク質と共通する1つ以上の機能的、構造的、生化学的、生物物理学的、又は他の特性若しくは特徴を示すタンパク質を指す。本明細書で使用される場合、用語「TnpB様」は、概して、TnpBタンパク質と共通する1つ以上の機能、構造的、生化学的、生物物理学的、又は他の特性若しくは特徴を示すタンパク質を指す。

10

【0072】

2つ以上の核酸又はポリペプチド配列の文脈における用語「配列同一性」又は「同一性パーセント」は、概して、配列比較アルゴリズムを使用して測定された場合、局所比較ウィンドウ又はグローバル比較ウィンドウにわたって最大の対応について比較及び整列されたとき、同一であるか、又は特定のパーセンテージの、同一であるアミノ酸残基又はヌクレオチドを有する、2つ(例えば、ペアワイズアラインメントにおいて)又はそれ以上(例えば、複数の配列アラインメントにおいて)の配列を指す。ポリペプチド配列に好適な配列比較アルゴリズムとしては、例えば、3のワード長(W)、10の期待値(E)のパラメーター、及び11の存在、1の延長でギャップコストを設定しているBLOSUM62スコアリングマトリックスを使用し、かつ30残基より長いポリペプチド配列についての条件付き組成スコアマトリックス調整を使用したBLASTP; 2のワード長(W)、1000000の期待値(E)のパラメーター、及びオープンギャップに対して9及び30残基より短い配列についての拡張ギャップに対して1でのPAM30スコアリング設定ギャップコストを使用したBLASTP(<https://blast.ncbi.nlm.nih.gov>で入手可能なBLASTにおいてBLASTPについてのデフォルトのパラメーターが存在する); 2の一致、-1のミスマッチ、及び-1のギャップのSmith-Waterman相同性検索アルゴリズムパラメーターを用いたCLUSTALW; デフォルトパラメーターを用いたMUSCLE; 2のリッジー及び1000の最大反復のパラメーターを用いたMAFFT; デフォルトパラメーターを用いたNovafold; デフォルトパラメーターを用いたHMMER hmmlignが挙げられる。

20

30

【0073】

2つ以上の核酸配列又はポリペプチド配列の文脈で、用語「最適に整列された」は、概して、例えば、最も高い又は「最適化された」同一性パーセントのスコアを生成するアラインメントによって決定される、アミノ酸残基又はヌクレオチドの最大対応に整列された2つ(例えば、ペアワイズアラインメントで)又はそれ以上(例えば、複数の配列アラインメントで)の配列を指す。

【0074】

1つ以上の保存的アミノ酸置換を有する本明細書に記載される酵素のうちのいずれかのバリエーションが、本開示に含まれる。こうした保存的置換は、ポリペプチドの三次元構造又は機能を破壊することなく、ポリペプチドのアミノ酸配列においてなされ得る。保存的置換は、アミノ酸を、互いに同様の疎水性、極性、及びR鎖長で置換することによって達成することができる。加えて、又は代わりに、異なる種由来の相同なタンパク質のアラインされた配列を比較することによって、保存的置換は、コードされたタンパク質の基本的な機能を変化させることなく、種間で変異したアミノ酸残基(例えば、非保存残基)を見つけることによって特定され得る。そのような保存的に置換されたバリエーションは、本明細書に記載されるトランスポザーゼタンパク質配列(例えば、本明細書に記載されるMG92ファミリートランスポザーゼ、又は本明細書に記載される任意の他のファミリートランス

40

50

ポザーゼ)のうちのいずれか1つと少なくとも約20%、少なくとも約25%、少なくとも約30%、少なくとも約35%、少なくとも約40%、少なくとも約45%、少なくとも約50%、少なくとも約55%、少なくとも約60%、少なくとも約65%、少なくとも約70%、少なくとも約75%、少なくとも約80%、少なくとも約85%、少なくとも約90%、少なくとも約91%、少なくとも約92%、少なくとも約93%、少なくとも約94%、少なくとも約95%、少なくとも約96%、少なくとも約97%、少なくとも約98%、少なくとも約99%の同一性を有するバリエーションを含んでもよい。いくつかの実施形態では、そのような保存的に置換されたバリエーションは、機能的バリエーションである。そのような機能的バリエーションは、トランスポザーゼの1つ以上の重要な活性部位残基の活性が破壊されないような置換を有する配列を包含することができる。いくつかの実施形態では、本明細書に記載されるタンパク質のうちのいずれかの機能的バリエーションは、図1Bでコールアウトされる保存された残基又は機能的残基のうちの少なくとも1つの置換を欠いている。いくつかの実施形態では、本明細書に記載されるタンパク質のうちのいずれかの機能的バリエーションは、図1Bでコールアウトされる保存された残基又は機能的残基の全ての置換を欠いている。

10

【0075】

また、本開示には、酵素の活性を減少させる又は排除するための1つ以上の触媒残基の置換を有する、本明細書に記載される酵素のうちのいずれかのバリエーション(例えば、活性低下バリエーション)も含まれる。いくつかの実施形態では、本明細書に記載されるタンパク質としての活性低下バリエーションは、図1Bでコールアウトされる少なくとも1つ、少なくとも2つ、又は3つ全ての触媒残基の破壊的置換を含む。

20

【0076】

機能的に類似したアミノ酸を提供する保存的置換表は、様々な参考文献から入手可能である(例えば、Creighton, Proteins: Structures and Molecular Properties (W H Freeman & Co.; 2nd edition (December 1993))を参照のこと)。以下の8つの群はそれぞれ、互いに保存的置換であるアミノ酸を含有する。

- 1) アラニン(A)、グリシン(G)、
- 2) アスパラギン酸(D)、グルタミン酸(E)、
- 3) アスパラギン(N)、グルタミン(Q)、
- 4) アルギニン(R)、リシン(K)、
- 5) イソロイシン(I)、ロイシン(L)、メチオニン(M)、バリン(V)、
- 6) フェニルアラニン(F)、チロシン(Y)、トリプトファン(W)、
- 7) セリン(S)、スレオニン(T)、及び
- 8) システイン(C)、メチオニン(M)。

30

【0077】

概要

固有の機能性及び構造を有する新しい転位因子の発見は、デオキシリボ核酸(DNA)編集技術を更に破壊し、速度、特異性、機能性、及び使いやすさを改善する潜在力を付与する可能性がある。微生物及びまさに多種多様な微生物種における転位因子の予測保有率と比較して、文献には機能的に特徴付けられた転位因子が比較的少ない。これは、実験室条件では、膨大な数の微生物種を容易には培養し得ないことが部分的に理由となっている。多数の微生物種を含有する天然の環境ニッチからのメタゲノムシーケンシングは、記録された新しい転位因子の数を劇的に増加させ、新しいオリゴヌクレオチド編集機能の発見を早める潜在力を付与する可能性がある。

40

【0078】

転位因子は、ゲノム内で位置を変更できるデオキシリボ核酸配列であり、変異の生成又は改善をもたらすことが多い。真核生物では、ゲノムの大部分、及び細胞DNAの質量の大部分が、転位因子に起因する。転位因子は、他の遺伝子を犠牲にして自身を増殖させる「利己的な遺伝子」であるが、様々な重要な機能を果たし、ゲノム進化に重要であること

50

が見出されている。転位因子は、それらの機構に基づいて、クラス I「レトロトランスポゾン」又はクラス II「DNAトランスポゾン」のいずれかに分類される。

【0079】

クラス I 転位因子は、レトロトランスポゾンとも呼ばれ、RNA 中間体を伴う二部分の「コピーアンドペースト」機構に従って機能する。まず、レトロトランスポゾンが転写される。得られた RNA は、その後、逆転写酵素（一般にレトロトランスポゾン自体によってコードされる）によって変換されて DNA に戻され、逆転写されたレトロトランスポゾンは、最終的にインテグラーゼによってゲノム内のその新しい位置に組み込まれる。レトロトランスポゾンは、3つの系列に更に分類される。長い末端反復（「LTR」）を有するレトロトランスポゾンは、逆転写酵素をコードし、反復 DNA の長い鎖に隣接している。長鎖散在反復配列（「LINE」）を有するレトロトランスポゾンは、逆転写酵素をコードし、LTR を欠いており、RNA ポリメラーゼ II によって転写される。短鎖散在反復配列（「SINE」）を有するレトロトランスポゾンは、RNA ポリメラーゼ III によって転写されるが、逆転写酵素を欠いており、代わりに他の転位因子（例えば、LINE）の逆転写機構に依存する。

10

【0080】

クラス II 転位因子は、DNAトランスポゾンとも呼ばれ、RNA 中間体を伴わない機構に従って機能する。多くの DNAトランスポゾンは、トランスポザーゼがトランスポゾンに隣接する末端逆位反復（「TIR」）に結合し、ドナー領域からトランスポゾンを切断し、それをゲノムの標的領域に挿入する「カットアンドペースト」機構を示す。「ヘリトロン」と呼ばれる他のものは、一本鎖 DNA 中間体を伴い、HUH エンドヌクレアーゼ機能及び 5' から 3' へのヘリカーゼ活性を有すると考えられる記録されていないタンパク質によって媒介される「ローリングサークル」機構を示す。まず、DNA の環状鎖がニッキングされて、2つの単一 DNA 鎖が作成される。タンパク質は、ニッキングされた鎖の 5' リン酸に付着したままであり、相補鎖の 3' ヒドロキシル端を露出したままにし、したがって、ポリメラーゼがニッキングされていない鎖を複製することを可能にする。複製が完了すると、新しい鎖は、解離し、それ自体が元の鋳型鎖とともに複製される。更に他の DNAトランスポゾンである「ポリントン」は、「自己合成」機構を経ると理論化されている。転位は、ラケット様構造を形成する一本鎖染色体外ポリントンエレメントのインテグラーゼ切除によって開始される。ポリントンは、DNA ポリメラーゼ B による複製を受け、二本鎖ポリントンは、インテグラーゼによってゲノムへと挿入される。最後に、IS 200 / IS 605 ファミリーのものなどのいくつかの DNAトランスポゾンは、TnpA がドナー遺伝子のラギング鎖鋳型から一本鎖 DNA の一片を（環状「トランスポゾン接続」として）切除し、それを標的遺伝子の複製フォークに再挿入する「ピールアンドペースト」機構を介して進行する。

20

30

【0081】

転位因子は、生物学的ツールとしていくつかの用途を見出したが、記録された転位因子は、可能な生物多様性及び標的可能性の全範囲を包含しておらず、全ての可能な活性を表していない場合がある。ここでは、転位因子について、多数のメタゲノムから数千ものゲノム断片を引き出した。記録された転位因子の多様性は、拡大されている可能性があり、新規な系は、高度に標的化可能で、コンパクトで、かつ正確な遺伝子編集剤へと発展している可能性がある。

40

MG 酵素

【0082】

いくつかの態様では、本開示は、新規なトランスポザーゼを提供する。これらの候補は、1つ以上の新規サブタイプを表していてもよく、いくつかのサブファミリーが特定されてもよい。これらのトランスポザーゼは、長さが約 500 アミノ酸未満である。これらのトランスポザーゼは、送達を単純化する可能性があり、治療用途を拡張する可能性がある。

【0083】

50

いくつかの態様では、本開示は、新規なトランスポザーゼを提供する。そのようなトランスポザーゼは、本明細書に記載される M G 9 2 であってもよい（図 1 A 及び図 1 B を参照のこと）。

【 0 0 8 4 】

一態様では、本開示は、メタゲノムシーケンシングを通して発見された操作されたトランスポザーゼ系を提供する。いくつかの実施形態では、メタゲノムシーケンシングは、試料において行われる。いくつかの実施形態では、試料は、様々な環境から収集され得る。そのような環境は、ヒトマイクロバイオーム、動物マイクロバイオーム、高温環境、低温環境であり得る。そのような環境は、堆積物を含み得る。

【 0 0 8 5 】

一態様では、本開示は、トランスポザーゼを含む操作されたトランスポザーゼ系を提供する。いくつかの実施形態では、トランスポザーゼは、未培養微生物に由来する。トランスポザーゼは、サブ末端回文配列を含む左側領域に結合するように構成されてもよい。トランスポザーゼは、サブ末端回文配列を含む右側領域に結合してもよい。

【 0 0 8 6 】

一態様では、本開示は、トランスポザーゼを含む操作されたトランスポザーゼ系を提供する。いくつかの実施形態では、トランスポザーゼは、配列番号 1 ~ 3 4 9 のうちのいずれか 1 つと少なくとも約 7 0 % の配列同一性を有する。いくつかの実施形態では、トランスポザーゼは、配列番号 1 ~ 3 4 9 のうちのいずれか 1 つと少なくとも約 2 0 %、少なくとも約 2 5 %、少なくとも約 3 0 %、少なくとも約 3 5 %、少なくとも約 4 0 %、少なくとも約 4 5 %、少なくとも約 5 0 %、少なくとも約 5 5 %、少なくとも約 6 0 %、少なくとも約 6 5 %、少なくとも約 7 0 %、少なくとも約 7 5 %、少なくとも約 8 0 %、少なくとも約 8 5 %、少なくとも約 9 0 %、少なくとも約 9 1 %、少なくとも約 9 2 %、少なくとも約 9 3 %、少なくとも約 9 4 %、少なくとも約 9 5 %、少なくとも約 9 6 %、少なくとも約 9 7 %、少なくとも約 9 8 %、又は少なくとも約 9 9 % の同一性を有する。

【 0 0 8 7 】

いくつかの実施形態では、トランスポザーゼは、配列番号 1 ~ 3 4 9 のうちのいずれか 1 つと少なくとも約 2 0 %、少なくとも約 2 5 %、少なくとも約 3 0 %、少なくとも約 3 5 %、少なくとも約 4 0 %、少なくとも約 4 5 %、少なくとも約 5 0 %、少なくとも約 5 5 %、少なくとも約 6 0 %、少なくとも約 6 5 %、少なくとも約 7 0 %、少なくとも約 7 5 %、少なくとも約 8 0 %、少なくとも約 8 5 %、少なくとも約 9 0 %、少なくとも約 9 1 %、少なくとも約 9 2 %、少なくとも約 9 3 %、少なくとも約 9 4 %、少なくとも約 9 5 %、少なくとも約 9 6 %、少なくとも約 9 7 %、少なくとも約 9 8 %、又は少なくとも約 9 9 % の同一性を有するバリエーションを含む。いくつかの実施形態では、トランスポザーゼは、配列番号 1 ~ 3 4 9 のうちのいずれか 1 つと実質的に同一であってもよい。

【 0 0 8 8 】

いくつかの実施形態では、トランスポザーゼは、T n p A 又は T n p B トランスポザーゼではない。いくつかの実施形態では、トランスポザーゼは、T n p A トランスポザーゼと約 9 0 % 未満、約 8 5 % 未満、約 8 0 % 未満、約 7 5 % 未満、約 7 0 % 未満、約 6 5 % 未満、約 6 0 % 未満、約 5 5 % 未満、約 5 0 % 未満、約 4 5 % 未満、約 4 0 % 未満、約 3 5 % 未満、約 3 0 % 未満、約 2 5 % 未満、約 2 0 % 未満、約 1 5 % 未満、約 1 0 % 未満、又は約 5 % 未満の配列同一性を有する。いくつかの実施形態では、トランスポザーゼは、T n p B トランスポザーゼと約 9 0 % 未満、約 8 5 % 未満、約 8 0 % 未満、約 7 5 % 未満、約 7 0 % 未満、約 6 5 % 未満、約 6 0 % 未満、約 5 5 % 未満、約 5 0 % 未満、約 4 5 % 未満、約 4 0 % 未満、約 3 5 % 未満、約 3 0 % 未満、約 2 5 % 未満、約 2 0 % 未満、約 1 5 % 未満、約 1 0 % 未満、又は約 5 % 未満の配列同一性を有する。

【 0 0 8 9 】

いくつかの実施形態では、トランスポザーゼは、触媒チロシン残基を含む。

【 0 0 9 0 】

いくつかの実施形態では、トランスポザーゼは、サブ末端回文配列を含む左側領域に結

10

20

30

40

50

合するように構成されている。いくつかの実施形態では、トランスポザーゼは、サブ末端回文配列を含む右側領域に結合するように構成されている。いくつかの実施形態では、トランスポザーゼは、サブ末端回文配列を含む左側領域及びサブ末端回文配列を含む右側領域に結合するように構成されている。

【 0 0 9 1 】

いくつかの実施形態では、トランスポザーゼは、二本鎖デオキシリボ核酸ポリヌクレオチドとしてカーゴヌクレオチド配列を転位するように構成されている。いくつかの実施形態では、トランスポザーゼは、一本鎖デオキシリボ核酸ポリヌクレオチドとしてカーゴヌクレオチド配列を転位するように構成されている。

【 0 0 9 2 】

いくつかの実施形態では、トランスポザーゼは、真核生物、真菌、植物、哺乳類、又はヒトのゲノムポリヌクレオチド配列と相補的である配列を含む。いくつかの実施形態では、トランスポザーゼは、真核生物のゲノムポリヌクレオチド配列と相補的である配列を含む。いくつかの実施形態では、トランスポザーゼは、真菌のゲノムポリヌクレオチド配列と相補的である配列を含む。いくつかの実施形態では、トランスポザーゼは、植物のゲノムポリヌクレオチド配列と相補的である配列を含む。いくつかの実施形態では、トランスポザーゼは、哺乳類のゲノムポリヌクレオチド配列と相補的である配列を含む。いくつかの実施形態では、トランスポザーゼは、ヒトのゲノムポリヌクレオチド配列と相補的である配列を含む。

【 0 0 9 3 】

いくつかの実施形態では、トランスポザーゼは、1つ以上の核局在化配列 (N L S) を有するバリエーションを含んでもよい。N L S は、トランスポザーゼのN末端又はC末端の近位にあってもよい。N L S は、配列番号 4 5 5 ~ 4 7 0 のうちのいずれか1つ、又は配列番号 4 5 5 ~ 4 7 0 のうちのいずれか1つと少なくとも約 2 0 %、少なくとも約 2 5 %、少なくとも約 3 0 %、少なくとも約 3 5 %、少なくとも約 4 0 %、少なくとも約 4 5 %、少なくとも約 5 0 %、少なくとも約 5 5 %、少なくとも約 6 0 %、少なくとも約 6 5 %、少なくとも約 7 0 %、少なくとも約 7 5 %、少なくとも約 8 0 %、少なくとも約 8 5 %、少なくとも約 9 0 %、少なくとも約 9 1 %、少なくとも約 9 2 %、少なくとも約 9 3 %、少なくとも約 9 4 %、少なくとも約 9 5 %、少なくとも約 9 6 %、少なくとも約 9 7 %、少なくとも約 9 8 %、又は少なくとも約 9 9 % の同一性を有するバリエーションに対して、N末端又はC末端に付加されてもよい。いくつかの実施形態では、N L S は、配列番号 4 5 5 ~ 4 7 0 のうちのいずれか1つと実質的に同一の配列を含んでもよい。いくつかの実施形態では、N L S は、配列番号 4 5 5 と実質的に同一の配列を含んでもよい。いくつかの実施形態では、N L S は、配列番号 4 5 6 と実質的に同一の配列を含んでもよい。

【 0 0 9 4 】

10

20

30

40

50

【表 1】

表 1：本開示のトランスポザーゼとともに使用され得る例示的なNLS配列

供給源	NLS アミノ酸配列	配列番号
SV40	PKKKRKV	455
スクレオプラスミン二分子NLS	KRPAATKKAGQAKKKK	456
c-myc NLS	PAAKRVKLD	457
c-myc NLS	RQRRNELKRSP	458
hRNPA1 M9 NLS	NQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQGGY	459
インポーチン-アルファ IBB ドメイン	RMRIZFKNKGKDTAELRRRRVEVSVELRKAKKDEQILKRRNV	460
筋腫 T タンパク質	VSRKRPRP	461
筋腫 T タンパク質	PPKKARED	462
p53	PQPKKKPL	463
マウス c-abl IV	SALIKKKKKMAP	464
インフルエンザウイルス NS1	DRLRR	465
インフルエンザウイルス NS1	PKQKKRK	466
肝炎ウイルスデルタ抗原	RKLKKKIKKL	467
マウス Mx1 タンパク質	REKKKFLKRR	468
ヒトポリ(ADP-リボース)ポリメラーゼ	KRKGDEVDGVDEVAKKKS	469
ステロイドホルモン受容体(ヒト)グルココルチコイド	RKCLQAGMNLEARKTKK	470

10

20

30

【0095】

いくつかの実施形態では、トランスポザーゼは、配列番号 1、3、5、7、9、11、13、15、若しくは 16 のうちのいずれか 1 つのバリエーション、又はそのバリエーションと少なくとも 70% 同一の配列を含む。いくつかの実施形態では、トランスポザーゼは、配列番号 1、3、5、7、9、11、13、15、若しくは 16 のうちのいずれか 1 つのバリエーション、又はそのバリエーションと少なくとも 75% 同一の配列を含む。いくつかの実施形態では、トランスポザーゼは、配列番号 1、3、5、7、9、11、13、15、若しくは 16 のうちのいずれか 1 つのバリエーション、又はそのバリエーションと少なくとも 80% 同一の配列を含む。いくつかの実施形態では、トランスポザーゼは、配列番号 1、3、5、7、9、11、13、15、若しくは 16 のうちのいずれか 1 つのバリエーション、又はそのバリエーションと少なくとも 85% 同一の配列を含む。いくつかの実施形態では、トランスポザーゼは、配列番号 1、3、5、7、9、11、13、15、若しくは 16 のうちのいずれか 1 つのバリエーション、又はそのバリエーションと少なくとも 90% 同一の配列を含む。いくつかの実施形態では、トランスポザーゼは、配列番号 1、3、5、7、9、11、13、15、若しくは 16 のうちのいずれか 1 つのバリエーション、又はそのバリエーションと少なくとも 95% 同一の配列を含む。

40

50

【0096】

いくつかの実施形態では、トランスポザーゼは、配列番号2、4、6、8、10、12、14、若しくは17のうちのいずれか1つのバリエーション、又はそのバリエーションと少なくとも70%同一の配列を含む。いくつかの実施形態では、トランスポザーゼは、配列番号2、4、6、8、10、12、14、若しくは17のうちのいずれか1つのバリエーション、又はそのバリエーションと少なくとも75%同一の配列を含む。いくつかの実施形態では、トランスポザーゼは、配列番号2、4、6、8、10、12、14、若しくは17のうちのいずれか1つのバリエーション、又はそのバリエーションと少なくとも80%同一の配列を含む。いくつかの実施形態では、トランスポザーゼは、配列番号2、4、6、8、10、12、14、若しくは17のうちのいずれか1つのバリエーション、又はそのバリエーションと少なくとも85%同一の配列を含む。いくつかの実施形態では、トランスポザーゼは、配列番号2、4、6、8、10、12、14、若しくは17のうちのいずれか1つのバリエーション、又はそのバリエーションと少なくとも90%同一の配列を含む。いくつかの実施形態では、トランスポザーゼは、配列番号2、4、6、8、10、12、14、若しくは17のうちのいずれか1つのバリエーション、又はそのバリエーションと少なくとも95%同一の配列を含む。

10

【0097】

いくつかの実施形態では、配列は、BLASTP、CLUSTALW、MUSCLE、若しくはMAFFTアルゴリズム、又はSmith-Waterman相同性検索アルゴリズムパラメータを用いたCLUSTALWアルゴリズムによって決定され得る。配列同一性は、3のワード長(W)、10の期待値(E)のパラメータ、及び11の存在、1の延長でギャップコストを設定しているBLOSUM62スコアリングマトリックスを使用し、条件付き組成スコアマトリックス調整を使用した、BLASTP相同性検索アルゴリズムによって決定され得る。

20

【0098】

一態様では、本開示は、本明細書に記載される操作されたトランスポザーゼ系をコードするデオキシリボ核酸ポリヌクレオチドを提供する。

【0099】

一態様では、本開示は、操作された核酸配列を含む核酸を提供する。いくつかの実施形態では、操作された核酸配列は、生物における発現に最適化されている。いくつかの実施形態では、トランスポザーゼは、未培養微生物に由来する。いくつかの実施形態では、生物は、未培養生物ではない。

30

【0100】

いくつかの実施形態では、トランスポザーゼは、配列番号1~349のうちのいずれか1つと少なくとも約70%の配列同一性を有する。いくつかの実施形態では、トランスポザーゼは、配列番号1~349のうちのいずれか1つと少なくとも約20%、少なくとも約25%、少なくとも約30%、少なくとも約35%、少なくとも約40%、少なくとも約45%、少なくとも約50%、少なくとも約55%、少なくとも約60%、少なくとも約65%、少なくとも約70%、少なくとも約75%、少なくとも約80%、少なくとも約85%、少なくとも約90%、少なくとも約91%、少なくとも約92%、少なくとも約93%、少なくとも約94%、少なくとも約95%、少なくとも約96%、少なくとも約97%、少なくとも約98%、又は少なくとも約99%の同一性を有する。

40

【0101】

いくつかの実施形態では、トランスポザーゼは、配列番号1~349のうちのいずれか1つと少なくとも約20%、少なくとも約25%、少なくとも約30%、少なくとも約35%、少なくとも約40%、少なくとも約45%、少なくとも約50%、少なくとも約55%、少なくとも約60%、少なくとも約65%、少なくとも約70%、少なくとも約75%、少なくとも約80%、少なくとも約85%、少なくとも約90%、少なくとも約91%、少なくとも約92%、少なくとも約93%、少なくとも約94%、少なくとも約95%、少なくとも約96%、少なくとも約97%、少なくとも約98%、又は少なくとも約99%の配列同一性を有するバリエーションを含む。いくつかの実施形態では、トランスポ

50

ザーゼは、配列番号 1 ~ 3 4 9 のうちのいずれか 1 つと実質的に同一であってもよい。

【 0 1 0 2 】

いくつかの実施形態では、トランスポザーゼは、T n p A 又は T n p B トランスポザーゼではない。いくつかの実施形態では、トランスポザーゼは、T n p A トランスポザーゼと約 9 0 % 未満、約 8 5 % 未満、約 8 0 % 未満、約 7 5 % 未満、約 7 0 % 未満、約 6 5 % 未満、約 6 0 % 未満、約 5 5 % 未満、約 5 0 % 未満、約 4 5 % 未満、約 4 0 % 未満、約 3 5 % 未満、約 3 0 % 未満、約 2 5 % 未満、約 2 0 % 未満、約 1 5 % 未満、約 1 0 % 未満、又は約 5 % 未満の配列同一性を有する。いくつかの実施形態では、トランスポザーゼは、T n p B トランスポザーゼと約 9 0 % 未満、約 8 5 % 未満、約 8 0 % 未満、約 7 5 % 未満、約 7 0 % 未満、約 6 5 % 未満、約 6 0 % 未満、約 5 5 % 未満、約 5 0 % 未満、約 4 5 % 未満、約 4 0 % 未満、約 3 5 % 未満、約 3 0 % 未満、約 2 5 % 未満、約 2 0 % 未満、約 1 5 % 未満、約 1 0 % 未満、又は約 5 % 未満の配列同一性を有する。

10

【 0 1 0 3 】

いくつかの実施形態では、トランスポザーゼは、触媒チロシン残基を含む。

【 0 1 0 4 】

いくつかの実施形態では、トランスポザーゼは、サブ末端回文配列を含む左側領域に結合するように構成されている。いくつかの実施形態では、トランスポザーゼは、サブ末端回文配列を含む右側領域に結合するように構成されている。いくつかの実施形態では、トランスポザーゼは、サブ末端回文配列を含む左側領域及びサブ末端回文配列を含む右側領域に結合するように構成されている。

20

【 0 1 0 5 】

いくつかの実施形態では、トランスポザーゼは、二本鎖デオキシリボ核酸ポリヌクレオチドとしてカーゴヌクレオチド配列を転位するように構成されている。いくつかの実施形態では、トランスポザーゼは、一本鎖デオキシリボ核酸ポリヌクレオチドとしてカーゴヌクレオチド配列を転位するように構成されている。

【 0 1 0 6 】

いくつかの実施形態では、トランスポザーゼは、真核生物、真菌、植物、哺乳類、又はヒトのゲノムポリヌクレオチド配列と相補的である配列を含む。いくつかの実施形態では、トランスポザーゼは、真核生物のゲノムポリヌクレオチド配列と相補的である配列を含む。いくつかの実施形態では、トランスポザーゼは、真菌のゲノムポリヌクレオチド配列と相補的である配列を含む。いくつかの実施形態では、トランスポザーゼは、植物のゲノムポリヌクレオチド配列と相補的である配列を含む。いくつかの実施形態では、トランスポザーゼは、哺乳類のゲノムポリヌクレオチド配列と相補的である配列を含む。いくつかの実施形態では、トランスポザーゼは、ヒトのゲノムポリヌクレオチド配列と相補的である配列を含む。

30

【 0 1 0 7 】

いくつかの実施形態では、トランスポザーゼは、1 つ以上の核局在化配列 (N L S) を有するバリエーションを含んでもよい。N L S は、トランスポザーゼの N 末端又は C 末端の近位にあってもよい。N L S は、配列番号 4 5 5 ~ 4 7 0 のうちのいずれか 1 つ、又は配列番号 4 5 5 ~ 4 7 0 のうちのいずれか 1 つと少なくとも約 2 0 % 、少なくとも約 2 5 % 、少なくとも約 3 0 % 、少なくとも約 3 5 % 、少なくとも約 4 0 % 、少なくとも約 4 5 % 、少なくとも約 5 0 % 、少なくとも約 5 5 % 、少なくとも約 6 0 % 、少なくとも約 6 5 % 、少なくとも約 7 0 % 、少なくとも約 7 5 % 、少なくとも約 8 0 % 、少なくとも約 8 5 % 、少なくとも約 9 0 % 、少なくとも約 9 1 % 、少なくとも約 9 2 % 、少なくとも約 9 3 % 、少なくとも約 9 4 % 、少なくとも約 9 5 % 、少なくとも約 9 6 % 、少なくとも約 9 7 % 、少なくとも約 9 8 % 、又は少なくとも約 9 9 % の同一性を有するバリエーションに対して、N 末端又は C 末端に付加されてもよい。いくつかの実施形態では、N L S は、配列番号 4 5 5 ~ 4 7 0 のうちのいずれか 1 つと実質的に同一の配列を含んでもよい。いくつかの実施形態では、N L S は、配列番号 4 5 5 と実質的に同一の配列を含んでもよい。いくつかの実施形態では、N L S は、配列番号 4 5 6 と実質的に同一の配列を含んでもよい。

40

50

【0108】

いくつかの実施形態では、生物は、原核生物である。いくつかの実施形態では、生物は、細菌である。いくつかの実施形態では、生物は、真核生物である。いくつかの実施形態では、生物は、真菌である。いくつかの実施形態では、生物は、植物である。いくつかの実施形態では、生物は、哺乳類である。いくつかの実施形態では、生物は、齧歯類である。いくつかの実施形態では、生物は、ヒトである。

【0109】

一態様では、本開示は、操作されたベクターを提供する。いくつかの実施形態では、操作されたベクターは、トランスポザーゼをコードする核酸配列を含む。いくつかの実施形態では、トランスポザーゼは、未培養微生物に由来する。

10

【0110】

いくつかの実施形態では、操作されたベクターは、本明細書に記載される核酸を含む。いくつかの実施形態では、本明細書に記載される核酸は、本明細書に記載されるデオキシリボ核酸ポリヌクレオチドである。いくつかの実施形態では、ベクターは、プラスミド、ミニサークル、CELLiD、アデノ随伴ウイルス(AAV)由来ビリオン、又はレンチウイルスである。

【0111】

一態様では、本開示は、本明細書に記載されるベクターを含む細胞を提供する。

【0112】

一態様では、本開示は、トランスポザーゼを製造する方法を提供する。いくつかの実施形態では、方法は、細胞を培養することを含む。

20

【0113】

一態様では、本開示は、二本鎖デオキシリボ核酸ポリヌクレオチドを結合、ニックング、切断、マーキング、修飾、又は転位する方法を提供する。方法は、二本鎖デオキシリボ核酸ポリヌクレオチドをトランスポザーゼと接触させることを含み得る。いくつかの実施形態では、トランスポザーゼは、サブ末端回文配列を含む左側領域に結合するように構成されている。いくつかの実施形態では、トランスポザーゼは、サブ末端回文配列を含む右側領域に結合するように構成されている。いくつかの実施形態では、トランスポザーゼは、サブ末端回文配列を含む左側領域及びサブ末端回文配列を含む右側領域に結合するように構成されている。

30

【0114】

いくつかの実施形態では、トランスポザーゼは、TnpAトランスポザーゼ又はTnpBトランスポザーゼではない。いくつかの実施形態では、トランスポザーゼは、TnpAトランスポザーゼと約90%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、又は約5%未満の配列同一性を有する。いくつかの実施形態では、トランスポザーゼは、TnpBトランスポザーゼと約90%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、又は約5%未満の配列同一性を有する。

40

【0115】

いくつかの実施形態では、トランスポザーゼは、触媒チロシン残基を含む。

【0116】

いくつかの実施形態では、トランスポザーゼは、二本鎖デオキシリボ核酸ポリヌクレオチドとしてカーゴヌクレオチド配列を転位するように構成されている。いくつかの実施形態では、トランスポザーゼは、一本鎖デオキシリボ核酸ポリヌクレオチドとしてカーゴヌクレオチド配列を転位するように構成されている。

【0117】

いくつかの実施形態では、トランスポザーゼは、未培養微生物に由来する。いくつかの

50

実施形態では、二本鎖デオキシリボ核酸ポリヌクレオチドは、真核生物、植物、真菌、哺乳類、齧歯類、又はヒト二本鎖デオキシリボ核酸ポリヌクレオチドである。

【0118】

一態様では、本開示は、標的核酸遺伝子座を修飾する方法を提供する。方法は、本明細書に記載される操作されたトランスポザーゼ系を標的核酸遺伝子座に送達することを含み得る。いくつかの実施形態では、複合体は、複合体の標的核酸遺伝子座への結合時に、複合体が標的核酸遺伝子座を修飾するように構成されている。

【0119】

いくつかの実施形態では、標的核酸遺伝子座を修飾することは、標的核酸遺伝子座を結合、ニックング、切断、マーキング、修飾、又は転位することを含む。いくつかの実施形態では、標的核酸遺伝子座は、デオキシリボ核酸(DNA)又はリボ核酸(RNA)を含む。いくつかの実施形態では、標的核酸は、ゲノムDNA、ウイルスDNA、ウイルスRNA、又は細菌DNAを含む。いくつかの実施形態では、標的核酸遺伝子座は、インビトロである。いくつかの実施形態では、標的核酸遺伝子座は、細胞内にある。いくつかの実施形態では、細胞は、原核細胞、細菌細胞、真核細胞、真菌細胞、植物細胞、動物細胞、哺乳類細胞、齧歯類細胞、霊長類細胞、又はヒト細胞である。いくつかの実施形態では、細胞は、初代細胞である。いくつかの実施形態では、初代細胞は、T細胞である。いくつかの実施形態では、初代細胞は、造血幹細胞(HSC)である。

10

【0120】

いくつかの実施形態では、操作されたトランスポザーゼ系の標的核酸遺伝子座への送達は、本明細書に記載される核酸又は本明細書に記載されるベクターを送達することを含む。いくつかの実施形態では、操作されたトランスポザーゼ系の標的核酸遺伝子座への送達は、トランスポザーゼをコードするオープンリーディングフレームを含む核酸を送達することを含む。いくつかの実施形態では、核酸は、プロモーターを含む。いくつかの実施形態では、トランスポザーゼをコードするオープンリーディングフレームは、プロモーターに作動可能に連結されている。

20

【0121】

いくつかの実施形態では、操作されたトランスポザーゼ系の標的核酸遺伝子座への送達は、トランスポザーゼをコードするオープンリーディングフレームを含有するキャッピングされたmRNAを送達することを含む。いくつかの実施形態では、操作されたトランスポザーゼ系の標的核酸遺伝子座への送達は、翻訳されたポリペプチドを送達することを含む。いくつかの実施形態では、操作されたトランスポザーゼ系の標的核酸遺伝子座への送達は、リボ核酸(RNA) polymerase IIIプロモーターに作動可能に連結された操作されたガイドRNAをコードするデオキシリボ核酸(DNA)を送達することを含む。

30

【0122】

いくつかの実施形態では、トランスポザーゼは、標的遺伝子座で、又は標的伝子座の近位で、一本鎖切断又は二本鎖切断を誘導する。いくつかの実施形態では、トランスポザーゼは、標的遺伝子座内又は標的遺伝子座の5'に互い違いの一本鎖切断を誘導する。

【0123】

一態様では、本開示は、異種トランスポザーゼをコードするオープンリーディングフレームを含む宿主細胞を提供する。いくつかの実施形態では、トランスポザーゼは、配列番号1~349のうちのいずれか1つと少なくとも約70%の配列同一性を有する。いくつかの実施形態では、トランスポザーゼは、配列番号1~349のうちのいずれか1つと少なくとも約20%、少なくとも約25%、少なくとも約30%、少なくとも約35%、少なくとも約40%、少なくとも約45%、少なくとも約50%、少なくとも約55%、少なくとも約60%、少なくとも約65%、少なくとも約70%、少なくとも約75%、少なくとも約80%、少なくとも約85%、少なくとも約90%、少なくとも約91%、少なくとも約92%、少なくとも約93%、少なくとも約94%、少なくとも約95%、少なくとも約96%、少なくとも約97%、少なくとも約98%、又は少なくとも約99%の同一性を有する。

40

50

【 0 1 2 4 】

いくつかの実施形態では、トランスポザーゼは、配列番号 1 ~ 3 4 9 のうちのいずれか 1 つと少なくとも約 2 0 %、少なくとも約 2 5 %、少なくとも約 3 0 %、少なくとも約 3 5 %、少なくとも約 4 0 %、少なくとも約 4 5 %、少なくとも約 5 0 %、少なくとも約 5 5 %、少なくとも約 6 0 %、少なくとも約 6 5 %、少なくとも約 7 0 %、少なくとも約 7 5 %、少なくとも約 8 0 %、少なくとも約 8 5 %、少なくとも約 9 0 %、少なくとも約 9 1 %、少なくとも約 9 2 %、少なくとも約 9 3 %、少なくとも約 9 4 %、少なくとも約 9 5 %、少なくとも約 9 6 %、少なくとも約 9 7 %、少なくとも約 9 8 %、又は少なくとも約 9 9 % の同一性を有するバリエーションを含む。いくつかの実施形態では、トランスポザーゼは、配列番号 1 ~ 3 4 9 のうちのいずれか 1 つと実質的に同一であってもよい。

10

【 0 1 2 5 】

いくつかの実施形態では、トランスポザーゼは、T n p A 又は T n p B トランスポザーゼではない。いくつかの実施形態では、トランスポザーゼは、T n p A トランスポザーゼと約 9 0 % 未満、約 8 5 % 未満、約 8 0 % 未満、約 7 5 % 未満、約 7 0 % 未満、約 6 5 % 未満、約 6 0 % 未満、約 5 5 % 未満、約 5 0 % 未満、約 4 5 % 未満、約 4 0 % 未満、約 3 5 % 未満、約 3 0 % 未満、約 2 5 % 未満、約 2 0 % 未満、約 1 5 % 未満、約 1 0 % 未満、又は約 5 % 未満の配列同一性を有する。いくつかの実施形態では、トランスポザーゼは、T n p B トランスポザーゼと約 9 0 % 未満、約 8 5 % 未満、約 8 0 % 未満、約 7 5 % 未満、約 7 0 % 未満、約 6 5 % 未満、約 6 0 % 未満、約 5 5 % 未満、約 5 0 % 未満、約 4 5 % 未満、約 4 0 % 未満、約 3 5 % 未満、約 3 0 % 未満、約 2 5 % 未満、約 2 0 % 未満、約 1 5 % 未満、約 1 0 % 未満、又は約 5 % 未満の配列同一性を有する。

20

【 0 1 2 6 】

いくつかの実施形態では、トランスポザーゼは、触媒チロシン残基を含む。

【 0 1 2 7 】

いくつかの実施形態では、トランスポザーゼは、サブ末端回文配列を含む左側領域に結合するように構成されている。いくつかの実施形態では、トランスポザーゼは、サブ末端回文配列を含む右側領域に結合するように構成されている。いくつかの実施形態では、トランスポザーゼは、サブ末端回文配列を含む左側領域及びサブ末端回文配列を含む右側領域に結合するように構成されている。

【 0 1 2 8 】

いくつかの実施形態では、トランスポザーゼは、二本鎖デオキシリボ核酸ポリヌクレオチドとしてカーゴヌクレオチド配列を転位するように構成されている。いくつかの実施形態では、トランスポザーゼは、一本鎖デオキシリボ核酸ポリヌクレオチドとしてカーゴヌクレオチド配列を転位するように構成されている。

30

【 0 1 2 9 】

いくつかの実施形態では、トランスポザーゼは、配列番号 1、3、5、7、9、11、13、15、若しくは 16 のうちのいずれか 1 つのバリエーション、又はそのバリエーションと少なくとも 7 0 % 同一の配列を含む。いくつかの実施形態では、トランスポザーゼは、配列番号 1、3、5、7、9、11、13、15、若しくは 16 のうちのいずれか 1 つのバリエーション、又はそのバリエーションと少なくとも 7 5 % 同一の配列を含む。いくつかの実施形態では、トランスポザーゼは、配列番号 1、3、5、7、9、11、13、15、若しくは 16 のうちのいずれか 1 つのバリエーション、又はそのバリエーションと少なくとも 8 0 % 同一の配列を含む。いくつかの実施形態では、トランスポザーゼは、配列番号 1、3、5、7、9、11、13、15、若しくは 16 のうちのいずれか 1 つのバリエーション、又はそのバリエーションと少なくとも 8 5 % 同一の配列を含む。いくつかの実施形態では、トランスポザーゼは、配列番号 1、3、5、7、9、11、13、15、若しくは 16 のうちのいずれか 1 つのバリエーション、又はそのバリエーションと少なくとも 9 0 % 同一の配列を含む。いくつかの実施形態では、トランスポザーゼは、配列番号 1、3、5、7、9、11、13、15、若しくは 16 のうちのいずれか 1 つのバリエーション、又はそのバリエーションと少なくとも 9 5 % 同一の配列を含む。

40

50

【0130】

いくつかの実施形態では、トランスポザーゼは、配列番号2、4、6、8、10、12、14、若しくは17のうちのいずれか1つのバリエーション、又はそのバリエーションと少なくとも70%同一の配列を含む。いくつかの実施形態では、トランスポザーゼは、配列番号2、4、6、8、10、12、14、若しくは17のうちのいずれか1つのバリエーション、又はそのバリエーションと少なくとも75%同一の配列を含む。いくつかの実施形態では、トランスポザーゼは、配列番号2、4、6、8、10、12、14、若しくは17のうちのいずれか1つのバリエーション、又はそのバリエーションと少なくとも80%同一の配列を含む。いくつかの実施形態では、トランスポザーゼは、配列番号2、4、6、8、10、12、14、若しくは17のうちのいずれか1つのバリエーション、又はそのバリエーションと少なくとも85%同一の配列を含む。いくつかの実施形態では、トランスポザーゼは、配列番号2、4、6、8、10、12、14、若しくは17のうちのいずれか1つのバリエーション、又はそのバリエーションと少なくとも90%同一の配列を含む。いくつかの実施形態では、トランスポザーゼは、配列番号2、4、6、8、10、12、14、若しくは17のうちのいずれか1つのバリエーション、又はそのバリエーションと少なくとも95%同一の配列を含む。

10

【0131】

いくつかの実施形態では、宿主細胞は、E. coli細胞である。いくつかの実施形態では、E. coli細胞は、DE3リソゲンであるか、又はE. coli細胞は、BL21(DE3)株である。いくつかの実施形態では、E. coli細胞は、ompT10n遺伝子型を有する。

20

【0132】

いくつかの実施形態では、オープンリーディングフレームは、T7プロモーター配列、T7-lacプロモーター配列、lacプロモーター配列、tacプロモーター配列、trcプロモーター配列、ParaBADプロモーター配列、PhaBADプロモーター配列、T5プロモーター配列、cspAプロモーター配列、araPBADプロモーター、ファージラムダからの強い左向きプロモーター(pLプロモーター)、又はそれらの任意の組み合わせに作動可能に連結されている。

【0133】

いくつかの実施形態では、オープンリーディングフレームは、トランスポザーゼをコードする配列にインフレームで連結された親和性タグをコードする配列を含む。いくつかの実施形態では、親和性タグは、固定化金属親和性クロマトグラフィー(IMAC)タグである。いくつかの実施形態では、IMACタグは、ポリヒスチジンタグである。いくつかの実施形態では、親和性タグは、mycタグ、ヒトインフルエンザヘマグルチニン(HA)タグ、マルトース結合タンパク質(MBP)タグ、グルタチオンS-トランスフェラーゼ(GST)タグ、ストレプトアビジンタグ、FLAGタグ、又はそれらの任意の組み合わせである。いくつかの実施形態では、親和性タグは、プロテアーゼ切断部位をコードするリンカー配列を介して、トランスポザーゼをコードする配列にインフレームで連結されている。いくつかの実施形態では、プロテアーゼ切断部位は、タバコエッチウイルス(TEV)プロテアーゼ切断部位、PreScission(登録商標)プロテアーゼ切断部位、トロンピン切断部位、第Xa因子切断部位、エンテロキナーゼ切断部位、又はそれらの任意の組み合わせである。

30

40

【0134】

いくつかの実施形態では、オープンリーディングフレームは、宿主細胞における発現のためにコドン最適化される。いくつかの実施形態では、オープンリーディングフレームは、ベクター上に提供される。いくつかの実施形態では、オープンリーディングフレームは、宿主細胞のゲノムに組み込まれる。

【0135】

一態様では、本開示は、適合する液体培地中に、本明細書に記載される宿主細胞を含む培養物を提供する。

【0136】

50

一態様では、本開示は、適合する成長培地中で、本明細書に記載される宿主細胞を培養することを含む、トランスポザーゼを産生する方法を提供する。いくつかの実施形態では、方法は、追加の化学剤又は増加された量の栄養素を添加することによって、トランスポザーゼの発現を誘導することを更に含む。いくつかの実施形態では、追加の化学剤又は増加された量の栄養素は、イソプロピル - D - 1 - チオガラクトピラノシド (I P T G) 又は追加の量のラクトースを含む。いくつかの実施形態では、方法は、培養後に宿主細胞を単離することと、宿主細胞を溶解してタンパク質抽出物を産生することとを更に含む。いくつかの実施形態では、方法は、タンパク質抽出物を I M A C、又はイオン親和性クロマトグラフィーに供することを更に含む。いくつかの実施形態では、オープンリーディングフレームは、トランスポザーゼをコードする配列にインフレームで連結された I M A C 親和性タグをコードする配列を含む。いくつかの実施形態では、I M A C 親和性タグは、プロテアーゼ切断部位をコードするリンカー配列を介して、トランスポザーゼをコードする配列にインフレームで連結されている。いくつかの実施形態では、プロテアーゼ切断部位は、タバコエッチウイルス (T E V) プロテアーゼ切断部位、P r e S c i s s i o n (登録商標) プロテアーゼ切断部位、トロンピン切断部位、第 X a 因子切断部位、エンテロキナーゼ切断部位、又はそれらの任意の組み合わせを含む。いくつかの実施形態では、方法は、プロテアーゼ切断部位に対応するプロテアーゼをトランスポザーゼと接触させることによって、I M A C 親和性タグを切断することを更に含む。いくつかの実施形態では、方法は、サブトラクティブ I M A C 親和性クロマトグラフィーを実施して、トランスポザーゼを含む組成物から親和性タグを除去することを更に含む。

10

20

【 0 1 3 7 】

一態様では、本開示は、細胞における遺伝子座を破壊する方法を提供する。いくつかの実施形態では、方法は、トランスポザーゼを含む組成物を細胞に接触させることを含む。いくつかの実施形態では、トランスポザーゼは、細胞内で T n p A トランスポザーゼと少なくとも同等の転位活性を有する。いくつかの実施形態では、トランスポザーゼは、配列番号 1 ~ 3 4 9 のうちのいずれか 1 つと少なくとも約 7 0 % の配列同一性を有する。いくつかの実施形態では、トランスポザーゼは、配列番号 1 ~ 3 4 9 のうちのいずれか 1 つと少なくとも約 2 0 %、少なくとも約 2 5 %、少なくとも約 3 0 %、少なくとも約 3 5 %、少なくとも約 4 0 %、少なくとも約 4 5 %、少なくとも約 5 0 %、少なくとも約 5 5 %、少なくとも約 6 0 %、少なくとも約 6 5 %、少なくとも約 7 0 %、少なくとも約 7 5 %、少なくとも約 8 0 %、少なくとも約 8 5 %、少なくとも約 9 0 %、少なくとも約 9 1 %、少なくとも約 9 2 %、少なくとも約 9 3 %、少なくとも約 9 4 %、少なくとも約 9 5 %、少なくとも約 9 6 %、少なくとも約 9 7 %、少なくとも約 9 8 %、又は少なくとも約 9 9 % の同一性を有する。

30

【 0 1 3 8 】

いくつかの実施形態では、トランスポザーゼは、配列番号 1 ~ 3 4 9 のうちのいずれか 1 つと少なくとも約 2 0 %、少なくとも約 2 5 %、少なくとも約 3 0 %、少なくとも約 3 5 %、少なくとも約 4 0 %、少なくとも約 4 5 %、少なくとも約 5 0 %、少なくとも約 5 5 %、少なくとも約 6 0 %、少なくとも約 6 5 %、少なくとも約 7 0 %、少なくとも約 7 5 %、少なくとも約 8 0 %、少なくとも約 8 5 %、少なくとも約 9 0 %、少なくとも約 9 1 %、少なくとも約 9 2 %、少なくとも約 9 3 %、少なくとも約 9 4 %、少なくとも約 9 5 %、少なくとも約 9 6 %、少なくとも約 9 7 %、少なくとも約 9 8 %、又は少なくとも約 9 9 % の同一性を有するパリアントを含む。いくつかの実施形態では、トランスポザーゼは、配列番号 1 ~ 3 4 9 のうちのいずれか 1 つと実質的に同一であってもよい。

40

【 0 1 3 9 】

いくつかの実施形態では、トランスポザーゼは、T n p A 又は T n p B トランスポザーゼではない。いくつかの実施形態では、トランスポザーゼは、T n p A トランスポザーゼと約 9 0 % 未満、約 8 5 % 未満、約 8 0 % 未満、約 7 5 % 未満、約 7 0 % 未満、約 6 5 % 未満、約 6 0 % 未満、約 5 5 % 未満、約 5 0 % 未満、約 4 5 % 未満、約 4 0 % 未満、約 3 5 % 未満、約 3 0 % 未満、約 2 5 % 未満、約 2 0 % 未満、約 1 5 % 未満、約 1 0 % 未満、

50

又は約 5 % 未満の配列同一性を有する。いくつかの実施形態では、トランスポザーゼは、T n p B トランスポザーゼと約 90 % 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、又は約 5 % 未満の配列同一性を有する。

【 0 1 4 0 】

いくつかの実施形態では、トランスポザーゼは、触媒チロシン残基を含む。

【 0 1 4 1 】

いくつかの実施形態では、トランスポザーゼは、サブ末端回文配列を含む左側領域に結合するように構成されている。いくつかの実施形態では、トランスポザーゼは、サブ末端回文配列を含む右側領域に結合するように構成されている。いくつかの実施形態では、トランスポザーゼは、サブ末端回文配列を含む左側領域及びサブ末端回文配列を含む右側領域に結合するように構成されている。

10

【 0 1 4 2 】

いくつかの実施形態では、トランスポザーゼは、二本鎖デオキシリボ核酸ポリヌクレオチドとしてカーゴヌクレオチド配列を転位するように構成されている。いくつかの実施形態では、トランスポザーゼは、一本鎖デオキシリボ核酸ポリヌクレオチドとしてカーゴヌクレオチド配列を転位するように構成されている。

【 0 1 4 3 】

いくつかの実施形態では、トランスポザーゼは、真核生物、真菌、植物、哺乳類、又はヒトのゲノムポリヌクレオチド配列と相補的である配列を含む。いくつかの実施形態では、トランスポザーゼは、真核生物のゲノムポリヌクレオチド配列と相補的である配列を含む。いくつかの実施形態では、トランスポザーゼは、真菌のゲノムポリヌクレオチド配列と相補的である配列を含む。いくつかの実施形態では、トランスポザーゼは、植物のゲノムポリヌクレオチド配列と相補的である配列を含む。いくつかの実施形態では、トランスポザーゼは、哺乳類のゲノムポリヌクレオチド配列と相補的である配列を含む。いくつかの実施形態では、トランスポザーゼは、ヒトのゲノムポリヌクレオチド配列と相補的である配列を含む。

20

【 0 1 4 4 】

いくつかの実施形態では、トランスポザーゼは、1つ以上の核局在化配列 (N L S) を有するバリエーションを含んでもよい。N L S は、トランスポザーゼの N 末端又は C 末端の近位にあってもよい。N L S は、配列番号 455 ~ 470 のうちのいずれか 1 つ、又は配列番号 455 ~ 470 のうちのいずれか 1 つと少なくとも約 20 %、少なくとも約 25 %、少なくとも約 30 %、少なくとも約 35 %、少なくとも約 40 %、少なくとも約 45 %、少なくとも約 50 %、少なくとも約 55 %、少なくとも約 60 %、少なくとも約 65 %、少なくとも約 70 %、少なくとも約 75 %、少なくとも約 80 %、少なくとも約 85 %、少なくとも約 90 %、少なくとも約 91 %、少なくとも約 92 %、少なくとも約 93 %、少なくとも約 94 %、少なくとも約 95 %、少なくとも約 96 %、少なくとも約 97 %、少なくとも約 98 %、又は少なくとも約 99 % の同一性を有するバリエーションに対して、N 末端又は C 末端に付加されてもよい。いくつかの実施形態では、N L S は、配列番号 455 ~ 470 のうちのいずれか 1 つと実質的に同一の配列を含んでもよい。いくつかの実施形態では、N L S は、配列番号 455 と実質的に同一の配列を含んでもよい。いくつかの実施形態では、N L S は、配列番号 456 と実質的に同一の配列を含んでもよい。

30

40

【 0 1 4 5 】

いくつかの実施形態では、転位活性は、標的核酸遺伝子座を含む細胞にトランスポザーゼを導入し、細胞内の標的核酸遺伝子座の転位を検出することによって、インビトロで測定される。いくつかの実施形態では、組成物は、20 ピコモル (p m o l) 以下のトランスポザーゼを含む。いくつかの実施形態では、組成物は、1 p m o l 以下のトランスポザーゼを含む。

【 0 1 4 6 】

50

本開示の系は、例えば、核酸編集（例えば、遺伝子編集）、核酸分子への結合（例えば、配列特異的結合）などの様々な用途に使用され得る。そのような系は、例えば、対象において疾患を引き起こす可能性のある遺伝的に受け継がれた変異に対処する（例えば、除去又は置換する）ため、遺伝子を細胞におけるその機能を確認するために不活性化するため、疾患を引き起こす遺伝子エレメントを検出する診断ツールとして（例えば、逆転写されたウイルスRNA若しくは疾患を引き起こす変異をコードする増幅されたDNA配列の切断を介して）、特定のヌクレオチド配列（例えば、細菌内の抗生物質耐性をコードする配列）を標的とし検出するためのプローブと組み合わせた不活性化酵素として、ウイルスゲノムを標的化することによってウイルスを不活性化するか、若しくは宿主細胞に感染できないようにするため、価値ある低分子、高分子、若しくは二次代謝産物を生じるように生物を操作するための遺伝子を加えるか、若しくは代謝経路を修正するため、進化的選択のための遺伝子駆動エレメントを確立するため、バイオセンサーとして外来低分子及びヌクレオチドによる細胞の摂動を検出するために、使用され得る。

10

【実施例】

【0147】

IUPACの慣例に従って、以下の略語が実施例を通して使用される。

A = アデニン

C = シトシン

G = グアニン

T = チミン

R = アデニン又はグアニン

Y = シトシン又はチミン

S = グアニン又はシトシン

W = アデニン又はチミン

K = グアニン又はチミン

M = アデニン又はシトシン

B = C、G、又はT

D = A、G、又はT

H = A、C、又はT

V = A、C、又はG

20

30

【0148】

実施例 1 - 新しいタンパク質のメタゲノム分析の方法

メタゲノム試料を、堆積物、土壌、及び動物から収集した。デオキシリボ核酸（DNA）を、Zymobiomics DNAミニプレップキットを用いて抽出し、Illumina HiSeq（登録商標）2500で配列決定した。試料を、所有者の同意を得て収集した。公的供給源からの更なる生配列データには、動物マイクロバイオーム、堆積物、土壌、温泉、熱水通気孔、海洋、泥炭、パーマフロスト、及び下水の配列が含まれていた。メタゲノム配列データを、記録されたトランスポザーゼタンパク質配列に基づいて生成された隠れマルコフモデルを使用して検索し、新しいトランスポザーゼを特定した。検索によって特定した新規トランスポザーゼタンパク質を、記録されたタンパク質に対して整列させて、潜在的な活性部位を特定した。このメタゲノムワークフローは、本明細書に記載されるMG92ファミリーの描写をもたらした。

40

【0149】

実施例 2 - トランスポザーゼのMG92ファミリーの発見

実施例1のメタゲノム分析からのデータの分析により、これまでに記述されていない、1つのファミリー（MG92）を含む推定トランスポザーゼ系の新しいクラスターが明らかになった。これらの新しい酵素及びそれらの例示的なサブドメインに対応するタンパク質配列を、配列番号1～349として示す。

【0150】

実施例 3 - インテグラーゼインピトロ活性（予測的）

50

インテグラーゼ活性は、E. coli 溶解物ベースの発現系（例えば、myTXTL、Arbor Biosciences）における発現を介して行うことができる。インビトロ試験に必要な成分は、3つのプラスミド、すなわち、T7プロモーター下のトランスポゾン遺伝子を有する発現プラスミド、標的プラスミド、並びにカーゴ遺伝子（例えば、Tet耐性遺伝子）の周りの転位に必要な左端（LE）及び右端（RE）のDNA配列を含有するドナープラスミドである。溶解物ベースの発現産物、標的DNA、及びドナーDNAをインキュベートして、転位が起こるようにする。転位は、PCRを介して検出される。加えて、転位産物を、T5でタグメントし、NGSを介して配列決定して、転位事象の集団上の挿入部位を決定する。あるいは、インビトロ転位産物を、抗生物質（例えば、Tet）選択下でE. coliへと形質転換することができ、この場合、成長には、転位カーゴがプラスミドへと安定して挿入されることが必要である。単一コロニー又はE. coliの集団のいずれかを配列決定して、挿入部位を決定することができる。

【0151】

組み込み効率は、組み込まれたカーゴを有する標的DNAの実験アウトプットのddPCR又はqPCRを介して測定することができ、同じくddPCRを介して測定される未修飾の標的DNAの量に対して正規化される。

【0152】

このアッセイはまた、溶解物ベースの発現からではなく、精製されたタンパク質成分で行われてもよい。この場合、タンパク質は、T7誘導性プロモーター下でE. coliプロテアーゼ欠損B株で発現され、細胞は超音波処理を用いて溶解され、目的のHisタグ付きタンパク質は、AKTA Avant FPLC (GE Lifescience) 上のHisTrap FF (GE Lifescience) Ni-NTA親和性クロマトグラフィーを用いて精製される。純度は、SDS-PAGE及びInstant Blue Ultrafast (Sigma-Aldrich) クマシー染色アクリルアミドゲル (Bio-Rad) 上で分解されたタンパク質バンドのImageLabソフトウェア (Bio-Rad) における密度測定を用いて決定される。タンパク質を、50 mMのTris-HCl、300 mMのNaCl、1 mMのTCEP、5%のグリセロール、pH 7.5で構成される保存緩衝液中（又は最大安定性について決定された他の緩衝液）で脱塩し、-80 で保存する。精製後、トランスポゾン遺伝子を、反応緩衝液、例えば、26 mMのHEPES pH 7.5、4.2 mMのTRIS pH 8、50 µg/mLのBSA、2 mMのATP、2.1 mMのDTT、0.05 mMのEDTA、0.2 mMのMgCl₂、28 mMのNaCl、21 mMのKCl、1.35%のグリセロール（最終pH 7.5）に15 mMのMgOAc₂を補充したものにおいて、上述の標的DNA及びドナーDNAに添加する。

【0153】

実施例 4 - ゲルシフトを介したトランスポゾン端の検証（予測的）

トランスポゾン端は、電気泳動移動度シフトアッセイ (EMSA) を介してトランスポザゼ結合について試験される。この場合、潜在的なLE又はREは、DNA断片（100 ~ 500 bp）として合成され、FAM標識プライマーを用いたPCRを介してFAMで端標識される。トランスポザゼタンパク質を、インビトロ転写/翻訳系（例えば、PURExpress）で合成する。合成後、1 µLのタンパク質を、結合緩衝液（例えば、20 mMのHEPES pH 7.5、2.5 mMのTris pH 7.5、10 mMのNaCl、0.0625 mMのEDTA、5 mMのTCEP、0.005%のBSA、1 µg/mLのポリ (dI-dC)、及び5%のグリセロール）中の10 µL反応物中の50 nMの標識されたRE又はLEに添加する。結合を30 °で40分間インキュベートし、次いで、2 µLの6Xローディング緩衝液（60 mMのKCl、10 mMのTris pH 7.5、50%グリセロール）を添加する。結合反応物を5% TBEゲル上で分離し、可視化する。トランスポザゼタンパク質の存在下でのLE又はREのシフトは、結合の成功に起因し得、トランスポザゼ活性を示す。このアッセイはまた、トランスポザゼのトランケーション又は変異で、並びにE. coli抽出物又は精製タンパク質を使用

して、実施することができる。

【0154】

実施例 5 - ドナー DNA の切断の検証 (予測的)

トランスポザーゼがドナー DNA の切断に関与することを確認するために、最大 10 bp で分離された RE - LE 接合部を含有する短い (約 140 bp) 断片を、FAM 標識プライマーを用いた PCR を介して、両端で FAM で標識する。標識された DNA 断片をインビトロ転写 / 翻訳トランスポザーゼ産物でインキュベートし、DNA を変性ゲル上で分析する。接合部の各端での切断は、ゲル上で異なる比率で移動する 2 つの標識された一本鎖断片をもたらす。

【0155】

実施例 6 - E. coli におけるインテグラーゼ活性 (予測的)

操作された E. coli 株を、トランスポゾン遺伝子を発現するプラスミドと、組み込みのための左端 (LE) 及び右端 (RE) トランスポゾンモチーフに隣接した選択可能なマーカーを有する温度感受性複製起点を含有するプラスミドとで形質転換する。トランスポザーゼ成分によるドナー ssDNA 優先性を確認するためには、ssDNA プラスミド超らせん形成をドナーとして使用することができる。次いで、これらの遺伝子の発現のために誘導された形質転換体を、プラスミド複製のための制限温度での選択によってゲノム標的へのマーカーの移行についてスクリーニングし、ゲノム内のマーカー組み込みを PCR によって確認する。

【0156】

組み込みは、不偏アプローチを使用してスクリーニングされる。簡潔に述べると、精製された gDNA は、Tn5 でタグメントされ、次いで目的の DNA は、Tn5 タグメンテーション及び選択可能なマーカーに特異的なプライマーを使用して PCR 増幅される。次いで、アンプリコンを NGS シーケンシングのために調製する。得られた配列の分析をトランスポゾン配列からトリミングし、隣接配列をゲノムにマッピングして挿入位置を決定し、挿入比率を決定する。

【0157】

あるいは、42 で欠陥のある DNA ポリメラーゼ I (Pol I) を産生する polA 変異体 E. coli 株 MM383 を使用して、前述のように組み込みを検出する (Brandsmat et al., 1981)。42 での成長後の選択可能なマーカーに対する耐性は、ドナー DNA の染色体への組み込みを示す。ドナーなしの pUC19 プラスミドを、抗生物質選択なしで 42 で 24 時間成長させた後の対照として使用する。

【0158】

選択培地で正常に成長する E. coli 株は、カーゴ耐性遺伝子をコードするドナー DNA を組み込んだものと推定される。抗生物質選択プレートで成長するコロニーは、カーゴの存在について遺伝子型決定され、全ゲノム配列の NGS が実施される。

【0159】

実施例 7 - 哺乳類細胞におけるインテグラーゼ活性 (予測的)

哺乳類細胞における標的化及び切断活性を示すために、トランスポゾンタンパク質の各々を、タンパク質配列のいずれかの末端上の 2 つの NLS ペプチドで精製する。選択可能なネオマイシン耐性マーカー (NeoR)、又は左端 (LE) 及び右端 (RE) モチーフに隣接した蛍光マーカーを含有するプラスミドを合成する。次いで、細胞を、プラスミドでトランスフェクトし、4 ~ 6 時間回収し、その後、トランスポゾンタンパク質でエレクトロポレーションする。ゲノムへの抗生物質耐性組み込みは、G418 - 耐性コロニー数によって定量化され、蛍光マーカーによる陽性転位は、蛍光活性化細胞サイトメトリーによってアッセイされる。共トランスフェクションの 72 時間後、ゲノム DNA が、抽出され、NGS - ライブラリの調製に使用される。組み込み頻度は、Tn5 タグメンテーションによってアッセイされる。

【0160】

実施例 8 - インシリコ分析

10

20

30

40

50

微生物、ウイルス、及び真核生物のゲノムの広範なアセンブリ駆動型メタゲノムデータベースを引き出して、*ssDNA*トランスポザゼ機能を有する予測タンパク質を得た。400を超える予測タンパク質が、挿入配列IS200/IS605のTnpAトランスポザゼに対して有意な e 値($< 1 \times 10^{-5}$)を有した。完全なORFをフィルタリングし、触媒残基(Y1及びHuH)の存在を確認した後、TnpA様タンパク質配列を、パラメーターG-INSI(Mol Biol Evol 30, 772-780(2013))でMAFFTで整列させ、アライメントを使用して、FastTree2で系統樹を推測した(Plos One 5, e9490(2010))。TnpAトランスポザゼの系統的分析により、IS200/IS605挿入配列に関連する新規TnpA様タンパク質配列の高い多様性が明らかになった(図2)。

10

【0161】

挿入配列の左端及び右端(LE及びRE)を予測するために、ISFinderデータベース(<https://www-is.biotoul.fr/>)で利用可能な活性なLE配列及びRE配列から共分散モデルを構築した。具体的には、LE配列及びRE配列の複数の配列アライメント(MSA)を、パラメーターX-INSI(Mol Biol Evol 30, 772-780(2013))でMAFFTで構築し、アライメントの二次構造を、パラメーター-p--aln-stk(Vienna Package)でRNAalifold 2.5.0でMSAから推測した。共分散モデルをInfernalパッケージ(<http://eddylab.org/infernal/>)で構築し、候補TnpAトランスポザゼを含有するゲノム断片を、Infernalコマンド「cmsearch」で共分散モデルを使用して検索した。共分散モデルは、70を超える候補IS200/IS605挿入配列についてLE及びREを予測した(図3)。

20

【0162】

実施例9 - *ssDNA*カーゴの生成

各TnpA様候補は、メタゲノムコンティグで特定された推定の左端(LE)配列及び右端(RE)配列を含む固有のカーゴを有した。これらの推定のLE配列及びRE配列をクローニングして、Gibsonアセンブリを介してカナマイシン(Kan)耐性カーゴ遺伝子に隣接させた。*ssDNA*カーゴを、Phusion HF(NEB)を用いた標準サイクル条件を使用して、フォワードプライマーGTGCGGTAGTAAAGGTTAAATACTGTT及び5'-リン酸修飾リバースプライマーCTATAGTGAGTCGTATTAを用いたLE/RE領域の外側の共通プライマーによるKanカーゴプラスミドのPCRを介して生成した。PCR増幅後、ラムダエキソヌクレアーゼ(NEB)を使用してDNA下部鎖を分解し、残りの上部鎖を、*ssDNA*を精製するために製造業者が推奨する変更を有するDCC-5スピンカラム(Zymo Research)を使用して精製した。一本鎖DNAを、アガロースゲル上でチェックして*dsDNA*の完全な変換を検証し、*ssDNA* Qubitキット(ThermoFisher)によって定量し、20nMの平均濃度を得た。

30

【0163】

実施例10 - TnpAインビトロ発現構築物の設計

インビトロ活性のために、各TnpA様タンパク質遺伝子を、T7プロモーターの制御下でE.coli翻訳のためにコドン最適化されたpET21(+)中で合成し、HAタグを欠く92-1を除き、C末端HAタグ及びHisタグに隣接させた。次いで、TnpA様タンパク質プラスミドを、T7プロモーターの上流及びT7ターミネーターの下流の約150bpに結合するプライマー(プライマーTGGCGAGAAAGGAAGGGAAG及びCCGAAACAAGCGCTCATGAG)を使用して増幅し、SPRIビーズクリーンアップ(MagBio HighPrep)を介して精製して最終鑄型濃度 $> 80 \text{ ng} / \mu\text{L}$ を得た。

40

【0164】

実施例11 - インビトロ転位活性

インビトロ活性のために、TnpA様タンパク質候補を、まず、製造業者が推奨する条

50

件に従って、37 で2時間、8 ng / μ Lの最小鑄型濃度 (PURExpress, NEB) でインビトロ転写 - 翻訳 (IVTT) キットで発現させた。発現をHAタグに対するウェスタンブロットを介して検証したが、このタグを欠く92 - 1は除いた。(図4)。10 μ Lの反応物当たり添加された1 μ LのIVTT産物、平均5 nMのssDNAカーゴ、及び反応緩衝液 (20 mMのHEPES (pH 7.5)、160 mMのNaCl、5 mMのMgCl₂、5 mMのTCEP、20 μ g / mLのBSA、0.5 μ g / mLのポリ-dIdC、及び20%のグリセロール) 中で8 N無作為化配列を含有する50 nMの161 nt「標的」ssDNAを用いて、転位アッセイをセットアップした。対照反応には、Tris緩衝液をPCR鑄型の代わりにIVTTに添加した、IVTTの鑄型なし対照 (NTC) 反応が含まれていた。反応物を37 で1時間インキュベートして、転位を発生させ、次いで反応物を水中で10倍希釈し、PCRを介して転位を検出した。LE接合部は、標的の5'端のフォワードプライマー及びKanカーゴ内のリバースプライマーを介して検出され、RE接合部は、Kanカーゴ内のフォワードプライマー及び標的の3'端のリバースプライマーを介して検出された。PCR産物を、アガロースゲル上で実行して転位を検出し(図5A及び図5B)、サンガー及びNGSシーケンシングを介して配列決定した。標的及びカーゴ配列の両方を含有するキメラリードを分析して、転位の接合部、挿入モチーフ、及びカーゴ上の切断部位を決定した(図6~図9)。

10

【0165】

LE PCR産物については、挿入モチーフは、カーゴと標的との間の重複配列同一性から特定することができる。例えば、MG92 - 3の標的とLEとの間の接合部は、標的及びカーゴの配列がもはや重複しなくなる点として特定される(図6)。挿入モチーフは、転位を伴わない標的DNAの隣接配列の分析を介して特定することができる。8 Nへの挿入の場合、標的モチーフを、REリードではなく、LEリードでのみ曖昧さなしに特定することができる。MG92 - 3については、挿入モチーフを、AATGAC又はその中のヌクレオチドのサブセット、例えばTGACとして特定した(図6~図7)。RE PCR産物については、RE接合部を、カーゴとターゲットへのマッピング間でリードが切り替わるブレイクポイントを介して特定する(図7)。LE接合部及びRE接合部のシーケンシングは、同じ挿入位置を示す。LE接合部は、NGSを介して更に確認され、これはサンガーシーケンシングを介して決定されたLE内の同じ切断点を特定した(図8)。

20

【0166】

これらのデータから、LE境界を以下のように決定することができる。TGAAAACAACAATTTTACCAAGGCCCGCAGGCTCCGTCTATAGCGACAAGCGCTAACTTTGGCTACGCTTGTCTTATAGGCGGGGTTAGT。これは、完全なMG92 - 3 LEのサブセットであり、認識モチーフAATGAC又はその中のヌクレオチドのサブセットに隣接した場合にのみ、MG92 - 3によって認識される。同様に、RE境界を以下のように特定することができる。GTTTGCGCTGTATCTGTGGTTCAGGTATCCACTCCTACCCTAAAGTAGCAAGGCATGAACGAAGAAGTTTATGCGGAGTTTGGAAAGCCCCGTCTATAATTCGCGAAGAAGCGGATTAGGCGGGGAGGGTTTAC、そのいくつか又は全ては、TnpA様タンパク質による認識、切除、及び挿入に必要である。両方の配列は、Cell 132, 208 - 220 (2008) 及びNucleic Acids Res 39, 8503 - 8512 (2011) に記載されるように、TnpA及びTnpA様タンパク質が認識する非標準塩基対形成相互作用に隣接するTnpA様タンパク質認識の予測ヘアピンを含有する(図6~7)。

30

40

【0167】

同様に、MG92 - 4の活性は、NGS検出を介して確認され、サンガーシーケンシングでは検出可能ではない弱いシグナルを有し、RE切断及び挿入を示した(図9)。このシグナルはNGSによってのみ検出可能であったため、これらの結果は、この挿入モチーフが可能であるが、最適な挿入配列ではない可能性があることを示唆している。

【0168】

50

実施例 12 - インビトロ切除アッセイ (予測的)

インビトロ切除活性を決定するために、TnpA様タンパク質候補を、製造業者が推奨する条件に従って、37℃で2時間、8 ng / μLの最小鑄型濃度 (PURExpress, NEB) でインビトロ転写 - 翻訳 (IVTT) キットで発現させた。10 μLの反応物当たり添加された1 μLのIVTT産物、及びTnpA反応緩衝液 (20 mMのHEPES (pH 7.5)、160 mMのNaCl、5 mMのMgCl₂、10 mMのTCEP、20 mg / mLのBSA、0.5 mgのポリ-dIdC、及び20%のグリセロール) 中で37℃で60分間の100 ngのLE-Kan-RE ssDNA (約2.2 kb) を用いて、切除アッセイを設定する。反応を、0.1% SDSを添加し、37℃で更に15分間インキュベーションすることによって終了させる。その後、反応物をRNase処理し、DNAアガロースゲル上で実行して、LE-Kan-RE ssDNAの切除が起こったかどうかを決定する。次いで、切除されたKan配列を、ゲル抽出し、LE及びRE切断モチーフを決定するためのシーケンシングに供する。

10

【0169】

実施例 13 - インビボ切除アッセイ (予測的)

インビボ切除アッセイはまた、一方がLE-Kan-REカーゴ及び他方がTnpAを含有する2つのプラスミドでE. coliを共形質転換することによって実施される。形質転換及び一晩成長させた後、切除は、一晩培養のミニプレップ及びDNAゲル上でKan配列が除去された再開鎖ドナー骨格分子の検出によって決定される。この実験の対照には、単一のプラスミドの形質転換、又はTnpA含有プラスミド及び逆複製起点を有するカーゴプラスミドの両方の形質転換が含まれる。切除されたDNA骨格を、ゲル抽出し、シーケンシングに供して、TnpAトランスポゾンのRE境界及びLE境界を得る。挿入モチーフは、切除された骨格に留まり、シールされた接合部でも特定され得る。

20

【0170】

実施例 14 - 挿入部位特異性の変更 (予測的)

挿入認識部位の操作は、TnpAタンパク質の操作を必要とすることなく、Cell 132, 208 - 220 (2008) によって実証されている。本明細書に記載されるメタゲノミクス由来TnpA様タンパク質によって認識される挿入部位は、挿入部位モチーフに対する配列変異及びLEヘアピン配列に隣接するLE ssDNAにおける塩基対形成パートナーに対する代償的変異を介して修飾される。一連の単一、二重、及び三重配列変異が、挿入部位及びLE配列の合理的に設計された位置に導入される。野生型TnpA様タンパク質による変異挿入部位の認識及び切断を、上述の切除 / 挿入アッセイ及びその後のシーケンシング工程を使用して、野生型LE挿入配列と同時に試験し、活性レベルを比較する。

30

【0171】

実施例 15 - TnpAは、プログラム可能な組み込みのために配列特異的エンドヌクレアーゼとともに使用することができる (予測的)

IS200 / IS605トランスポズンは、特定の標的部位に組み込まれる移動性遺伝子エレメントの一種である。これらのトランスポズンは、それらのコードされたTnpA様トランスポザーゼ、チロシン (Y) トランスポザーゼのファミリーに属する酵素によって動員される (Microbiol Spectr 3, (2015) で論評されている)。IS200 / IS605トランスポズン動員の機構は、TnpA又はTnpA様タンパク質によるその切除、続いて、標的部位が複製フォークでssDNAとしてアクセス可能な場合、宿主複製中に認識された標的部位でのその組み込みを伴う (Cell 142, 398 - 408 (2010))。

40

【0172】

TnpA様タンパク質と共有される標的部位への特定の配列特異的 (例えば、Cas) エンドヌクレアーゼエフェクターのRNAガイド結合能力は、Rループの形成を通してssDNA及び標的部位を利用可能にすることによって、所望のカーゴのTnpA様エフェクター介在性組み込みを補助し得る。具体的には、TnpA様認識可能なLE及びREに

50

隣接した所望のカーゴ（例えば、蛍光マーカ-遺伝子）は、TnpA又はTnpA様エフェクターによってドナー鑄型から切除され、（融合された）配列特異的エンドヌクラーゼの結合によって利用可能になる所望の標的部位（TnpA又はTnpA様タンパク質認識可能モチーフを含有する）へと組み込まれる。配列特異的エンドヌクラーゼは、触媒的に死滅するか、又は低減若しくは改変されたエンドヌクラーゼ（例えば、ニッカーゼ）活性を有するように操作されてもよい。したがって、TnpA様タンパク質は、融合され、操作された（例えば、死滅した又はニッカーゼ）配列特異的エンドヌクラーゼエフェクターによって利用可能にされたTAM依存性標的部位へと所望のカーゴを挿入するように「プログラム」され得る。

【0173】

10

実施例16 - dsDNA中のRループへのTnpA様挿入のインビトロ試験（予測的）

dsDNA中のRループとして生成されたssDNAに挿入するTnpA様タンパク質の能力は、インビトロで特定された活性TnpA様タンパク質、並びにそれらの対応するLE配列及びRE配列を使用して試験することができる。Rループは、IVTT反応で発現されるか、又は精製されたRNPとして添加されるRNA指向性ヌクラーゼ死滅酵素又はニッカーゼなどの配列特異的エンドヌクラーゼを介して生成することができる。TnpA様タンパク質は、標的ssDNAがdsDNA及びRNPによって置き換えられることを除いて、インビトロ挿入アッセイに記載されるように試験される。挿入活性は、LE接合部又はRE接合部のいずれかに隣接するdsDNA標的及びssDNAカーゴ内のプライマーを用いてPCRを介してアッセイされる。挿入部位の最適な位置は、Rループに沿った様々な位置に挿入モチーフを配置して、TnpA様タンパク質による最良のアクセス性を有する部位を決定することによって試験される。ミスマッチDNA鎖がアニーリングされるdsDNA中のssDNAバブルへの挿入も試験することができる。

20

【0174】

30

40

50

【表 2 - 1】

表 2 - 本明細書で言及されるタンパク質配列及び核酸配列

触媒	配列番号	説明	タイプ
MG92 転位タンパク質	1	MG92-1-A 転位タンパク質	タンパク質
MG92 転位タンパク質	2	MG92-1-B 転位タンパク質	タンパク質
MG92 転位タンパク質	3	MG92-2-A 転位タンパク質	タンパク質
MG92 転位タンパク質	4	MG92-2-B 転位タンパク質	タンパク質
MG92 転位タンパク質	5	MG92-3-A 転位タンパク質	タンパク質
MG92 転位タンパク質	6	MG92-3-B 転位タンパク質	タンパク質
MG92 転位タンパク質	7	MG92-4-A 転位タンパク質	タンパク質
MG92 転位タンパク質	8	MG92-4-B 転位タンパク質	タンパク質
MG92 転位タンパク質	9	MG92-5-A 転位タンパク質	タンパク質
MG92 転位タンパク質	10	MG92-5-B 転位タンパク質	タンパク質
MG92 転位タンパク質	11	MG92-6-A 転位タンパク質	タンパク質
MG92 転位タンパク質	12	MG92-6-B 転位タンパク質	タンパク質
MG92 転位タンパク質	13	MG92-7-A 転位タンパク質	タンパク質
MG92 転位タンパク質	14	MG92-7-B 転位タンパク質	タンパク質
MG92 転位タンパク質	15	MG92-8-A 転位タンパク質	タンパク質
MG92 転位タンパク質	16	MG92-9-A 転位タンパク質	タンパク質
MG92 転位タンパク質	17	MG92-9-B 転位タンパク質	タンパク質

10

20

30

40

【 0 1 7 5 】

50

【表 2 - 2】

触媒	配列番号	説明	タイプ
MG92 転位タンパク質	18	MG92-10 転位タンパク質	タンパク質
MG92 転位タンパク質	19	MG92-11 転位タンパク質	タンパク質
MG92 転位タンパク質	20	MG92-12 転位タンパク質	タンパク質
MG92 転位タンパク質	21	MG92-13 転位タンパク質	タンパク質
MG92 転位タンパク質	22	MG92-14 転位タンパク質	タンパク質
MG92 転位タンパク質	23	MG92-15 転位タンパク質	タンパク質
MG92 転位タンパク質	24	MG92-17 転位タンパク質	タンパク質
MG92 転位タンパク質	25	MG92-19 転位タンパク質	タンパク質
MG92 転位タンパク質	26	MG92-20 転位タンパク質	タンパク質
MG92 転位タンパク質	27	MG92-21 転位タンパク質	タンパク質
MG92 転位タンパク質	28	MG92-22 転位タンパク質	タンパク質
MG92 転位タンパク質	29	MG92-23 転位タンパク質	タンパク質
MG92 転位タンパク質	30	MG92-24 転位タンパク質	タンパク質
MG92 転位タンパク質	31	MG92-25 転位タンパク質	タンパク質
MG92 転位タンパク質	32	MG92-26 転位タンパク質	タンパク質
MG92 転位タンパク質	33	MG92-27 転位タンパク質	タンパク質
MG92 転位タンパク質	34	MG92-28 転位タンパク質	タンパク質

10

20

30

40

【 0 1 7 6 】

50

【表 2 - 3】

触媒	配列番号	説明	タイプ
MG92 転位タンパク質	35	MG92-29 転位タンパク質	タンパク質
MG92 転位タンパク質	36	MG92-30 転位タンパク質	タンパク質
MG92 転位タンパク質	37	MG92-31 転位タンパク質	タンパク質
MG92 転位タンパク質	38	MG92-32 転位タンパク質	タンパク質
MG92 転位タンパク質	39	MG92-33 転位タンパク質	タンパク質
MG92 転位タンパク質	40	MG92-34 転位タンパク質	タンパク質
MG92 転位タンパク質	41	MG92-35 転位タンパク質	タンパク質
MG92 転位タンパク質	42	MG92-36 転位タンパク質	タンパク質
MG92 転位タンパク質	43	MG92-37 転位タンパク質	タンパク質
MG92 転位タンパク質	44	MG92-38 転位タンパク質	タンパク質
MG92 転位タンパク質	45	MG92-39 転位タンパク質	タンパク質
MG92 転位タンパク質	46	MG92-40 転位タンパク質	タンパク質
MG92 転位タンパク質	47	MG92-41 転位タンパク質	タンパク質
MG92 転位タンパク質	48	MG92-42 転位タンパク質	タンパク質
MG92 転位タンパク質	49	MG92-43 転位タンパク質	タンパク質
MG92 転位タンパク質	50	MG92-44 転位タンパク質	タンパク質
MG92 転位タンパク質	51	MG92-45 転位タンパク質	タンパク質

10

20

30

40

【 0 1 7 7 】

50

【表 2 - 4】

触媒	配列番号	説明	タイプ
MG92 転位タンパク質	52	MG92-46 転位タンパク質	タンパク質
MG92 転位タンパク質	53	MG92-47 転位タンパク質	タンパク質
MG92 転位タンパク質	54	MG92-48 転位タンパク質	タンパク質
MG92 転位タンパク質	55	MG92-49 転位タンパク質	タンパク質
MG92 転位タンパク質	56	MG92-50 転位タンパク質	タンパク質
MG92 転位タンパク質	57	MG92-51 転位タンパク質	タンパク質
MG92 転位タンパク質	58	MG92-52 転位タンパク質	タンパク質
MG92 転位タンパク質	59	MG92-53 転位タンパク質	タンパク質
MG92 転位タンパク質	60	MG92-54 転位タンパク質	タンパク質
MG92 転位タンパク質	61	MG92-55 転位タンパク質	タンパク質
MG92 転位タンパク質	62	MG92-56 転位タンパク質	タンパク質
MG92 転位タンパク質	63	MG92-57 転位タンパク質	タンパク質
MG92 転位タンパク質	64	MG92-58 転位タンパク質	タンパク質
MG92 転位タンパク質	65	MG92-59 転位タンパク質	タンパク質
MG92 転位タンパク質	66	MG92-60 転位タンパク質	タンパク質
MG92 転位タンパク質	67	MG92-61 転位タンパク質	タンパク質
MG92 転位タンパク質	68	MG92-62 転位タンパク質	タンパク質

10

20

30

40

【 0 1 7 8 】

50

【表 2 - 5】

触媒	配列番号	説明	タイプ
MG92 転位タンパク質	69	MG92-63 転位タンパク質	タンパク質
MG92 転位タンパク質	70	MG92-64 転位タンパク質	タンパク質
MG92 転位タンパク質	71	MG92-65 転位タンパク質	タンパク質
MG92 転位タンパク質	72	MG92-66 転位タンパク質	タンパク質
MG92 転位タンパク質	73	MG92-67 転位タンパク質	タンパク質
MG92 転位タンパク質	74	MG92-68 転位タンパク質	タンパク質
MG92 転位タンパク質	75	MG92-69 転位タンパク質	タンパク質
MG92 転位タンパク質	76	MG92-70 転位タンパク質	タンパク質
MG92 転位タンパク質	77	MG92-71 転位タンパク質	タンパク質
MG92 転位タンパク質	78	MG92-72 転位タンパク質	タンパク質
MG92 転位タンパク質	79	MG92-73 転位タンパク質	タンパク質
MG92 転位タンパク質	80	MG92-74 転位タンパク質	タンパク質
MG92 転位タンパク質	81	MG92-75 転位タンパク質	タンパク質
MG92 転位タンパク質	82	MG92-76 転位タンパク質	タンパク質
MG92 転位タンパク質	83	MG92-77 転位タンパク質	タンパク質
MG92 転位タンパク質	84	MG92-78 転位タンパク質	タンパク質
MG92 転位タンパク質	85	MG92-79 転位タンパク質	タンパク質

10

20

30

40

【 0 1 7 9 】

50

【表 2 - 6】

触媒	配列番号	説明	タイプ
MG92 転位タンパク質	86	MG92-80 転位タンパク質	タンパク質
MG92 転位タンパク質	87	MG92-81 転位タンパク質	タンパク質
MG92 転位タンパク質	88	MG92-82 転位タンパク質	タンパク質
MG92 転位タンパク質	89	MG92-83 転位タンパク質	タンパク質
MG92 転位タンパク質	90	MG92-84 転位タンパク質	タンパク質
MG92 転位タンパク質	91	MG92-85 転位タンパク質	タンパク質
MG92 転位タンパク質	92	MG92-86 転位タンパク質	タンパク質
MG92 転位タンパク質	93	MG92-87 転位タンパク質	タンパク質
MG92 転位タンパク質	94	MG92-88 転位タンパク質	タンパク質
MG92 転位タンパク質	95	MG92-89 転位タンパク質	タンパク質
MG92 転位タンパク質	96	MG92-90 転位タンパク質	タンパク質
MG92 転位タンパク質	97	MG92-91 転位タンパク質	タンパク質
MG92 転位タンパク質	98	MG92-92 転位タンパク質	タンパク質
MG92 転位タンパク質	99	MG92-93 転位タンパク質	タンパク質
MG92 転位タンパク質	100	MG92-94 転位タンパク質	タンパク質
MG92 転位タンパク質	101	MG92-95 転位タンパク質	タンパク質
MG92 転位タンパク質	102	MG92-96 転位タンパク質	タンパク質

10

20

30

40

【 0 1 8 0 】

50

【表 2 - 7】

触媒	配列番号	説明	タイプ
MG92 転位タンパク質	103	MG92-97 転位タンパク質	タンパク質
MG92 転位タンパク質	104	MG92-98 転位タンパク質	タンパク質
MG92 転位タンパク質	105	MG92-99 転位タンパク質	タンパク質
MG92 転位タンパク質	106	MG92-100 転位タンパク質	タンパク質
MG92 転位タンパク質	107	MG92-101 転位タンパク質	タンパク質
MG92 転位タンパク質	108	MG92-102 転位タンパク質	タンパク質
MG92 転位タンパク質	109	MG92-103 転位タンパク質	タンパク質
MG92 転位タンパク質	110	MG92-104 転位タンパク質	タンパク質
MG92 転位タンパク質	111	MG92-105 転位タンパク質	タンパク質
MG92 転位タンパク質	112	MG92-106 転位タンパク質	タンパク質
MG92 転位タンパク質	113	MG92-107 転位タンパク質	タンパク質
MG92 転位タンパク質	114	MG92-108 転位タンパク質	タンパク質
MG92 転位タンパク質	115	MG92-109 転位タンパク質	タンパク質
MG92 転位タンパク質	116	MG92-110 転位タンパク質	タンパク質
MG92 転位タンパク質	117	MG92-111 転位タンパク質	タンパク質
MG92 転位タンパク質	118	MG92-112 転位タンパク質	タンパク質
MG92 転位タンパク質	119	MG92-113 転位タンパク質	タンパク質

10

20

30

40

【 0 1 8 1 】

50

【表 2 - 8】

触媒	配列番号	説明	タイプ
MG92 転位タンパク質	120	MG92-114 転位タンパク質	タンパク質
MG92 転位タンパク質	121	MG92-115 転位タンパク質	タンパク質
MG92 転位タンパク質	122	MG92-116 転位タンパク質	タンパク質
MG92 転位タンパク質	123	MG92-117 転位タンパク質	タンパク質
MG92 転位タンパク質	124	MG92-118 転位タンパク質	タンパク質
MG92 転位タンパク質	125	MG92-119 転位タンパク質	タンパク質
MG92 転位タンパク質	126	MG92-120 転位タンパク質	タンパク質
MG92 転位タンパク質	127	MG92-121 転位タンパク質	タンパク質
MG92 転位タンパク質	128	MG92-122 転位タンパク質	タンパク質
MG92 転位タンパク質	129	MG92-123 転位タンパク質	タンパク質
MG92 転位タンパク質	130	MG92-124 転位タンパク質	タンパク質
MG92 転位タンパク質	131	MG92-125 転位タンパク質	タンパク質
MG92 転位タンパク質	132	MG92-126 転位タンパク質	タンパク質
MG92 転位タンパク質	133	MG92-127 転位タンパク質	タンパク質
MG92 転位タンパク質	134	MG92-128 転位タンパク質	タンパク質
MG92 転位タンパク質	135	MG92-129 転位タンパク質	タンパク質
MG92 転位タンパク質	136	MG92-130 転位タンパク質	タンパク質

10

20

30

40

【 0 1 8 2 】

50

【表 2 - 9】

触媒	配列番号	説明	タイプ
MG92 転位タンパク質	137	MG92-131 転位タンパク質	タンパク質
MG92 転位タンパク質	138	MG92-132 転位タンパク質	タンパク質
MG92 転位タンパク質	139	MG92-133 転位タンパク質	タンパク質
MG92 転位タンパク質	140	MG92-134 転位タンパク質	タンパク質
MG92 転位タンパク質	141	MG92-135 転位タンパク質	タンパク質
MG92 転位タンパク質	142	MG92-136 転位タンパク質	タンパク質
MG92 転位タンパク質	143	MG92-137 転位タンパク質	タンパク質
MG92 転位タンパク質	144	MG92-138 転位タンパク質	タンパク質
MG92 転位タンパク質	145	MG92-139 転位タンパク質	タンパク質
MG92 転位タンパク質	146	MG92-140 転位タンパク質	タンパク質
MG92 転位タンパク質	147	MG92-141 転位タンパク質	タンパク質
MG92 転位タンパク質	148	MG92-142 転位タンパク質	タンパク質
MG92 転位タンパク質	149	MG92-143 転位タンパク質	タンパク質
MG92 転位タンパク質	150	MG92-144 転位タンパク質	タンパク質
MG92 転位タンパク質	151	MG92-145 転位タンパク質	タンパク質
MG92 転位タンパク質	152	MG92-146 転位タンパク質	タンパク質
MG92 転位タンパク質	153	MG92-147 転位タンパク質	タンパク質

10

20

30

40

【 0 1 8 3 】

50

【表 2 - 1 0】

触媒	配列番号	説明	タイプ
MG92 転位タンパク質	154	MG92-148 転位タンパク質	タンパク質
MG92 転位タンパク質	155	MG92-149 転位タンパク質	タンパク質
MG92 転位タンパク質	156	MG92-150 転位タンパク質	タンパク質
MG92 転位タンパク質	157	MG92-151 転位タンパク質	タンパク質
MG92 転位タンパク質	158	MG92-152 転位タンパク質	タンパク質
MG92 転位タンパク質	159	MG92-153 転位タンパク質	タンパク質
MG92 転位タンパク質	160	MG92-154 転位タンパク質	タンパク質
MG92 転位タンパク質	161	MG92-155 転位タンパク質	タンパク質
MG92 転位タンパク質	162	MG92-156 転位タンパク質	タンパク質
MG92 転位タンパク質	163	MG92-157 転位タンパク質	タンパク質
MG92 転位タンパク質	164	MG92-158 転位タンパク質	タンパク質
MG92 転位タンパク質	165	MG92-159 転位タンパク質	タンパク質
MG92 転位タンパク質	166	MG92-160 転位タンパク質	タンパク質
MG92 転位タンパク質	167	MG92-161 転位タンパク質	タンパク質
MG92 転位タンパク質	168	MG92-162 転位タンパク質	タンパク質
MG92 転位タンパク質	169	MG92-163 転位タンパク質	タンパク質
MG92 転位タンパク質	170	MG92-164 転位タンパク質	タンパク質

10

20

30

40

【 0 1 8 4】

50

【表 2 - 1 1】

触媒	配列番号	説明	タイプ
MG92 転位タンパク質	171	MG92-165 転位タンパク質	タンパク質
MG92 転位タンパク質	172	MG92-166 転位タンパク質	タンパク質
MG92 転位タンパク質	173	MG92-167 転位タンパク質	タンパク質
MG92 転位タンパク質	174	MG92-168 転位タンパク質	タンパク質
MG92 転位タンパク質	175	MG92-169 転位タンパク質	タンパク質
MG92 転位タンパク質	176	MG92-170 転位タンパク質	タンパク質
MG92 転位タンパク質	177	MG92-171 転位タンパク質	タンパク質
MG92 転位タンパク質	178	MG92-172 転位タンパク質	タンパク質
MG92 転位タンパク質	179	MG92-173 転位タンパク質	タンパク質
MG92 転位タンパク質	180	MG92-174 転位タンパク質	タンパク質
MG92 転位タンパク質	181	MG92-175 転位タンパク質	タンパク質
MG92 転位タンパク質	182	MG92-176 転位タンパク質	タンパク質
MG92 転位タンパク質	183	MG92-177 転位タンパク質	タンパク質
MG92 転位タンパク質	184	MG92-178 転位タンパク質	タンパク質
MG92 転位タンパク質	185	MG92-179 転位タンパク質	タンパク質
MG92 転位タンパク質	186	MG92-180 転位タンパク質	タンパク質
MG92 転位タンパク質	187	MG92-181 転位タンパク質	タンパク質

10

20

30

40

【 0 1 8 5 】

50

【表 2 - 1 2】

触媒	配列番号	説明	タイプ
MG92 転位タンパク質	188	MG92-182 転位タンパク質	タンパク質
MG92 転位タンパク質	189	MG92-183 転位タンパク質	タンパク質
MG92 転位タンパク質	190	MG92-184 転位タンパク質	タンパク質
MG92 転位タンパク質	191	MG92-185 転位タンパク質	タンパク質
MG92 転位タンパク質	192	MG92-186 転位タンパク質	タンパク質
MG92 転位タンパク質	193	MG92-187 転位タンパク質	タンパク質
MG92 転位タンパク質	194	MG92-188 転位タンパク質	タンパク質
MG92 転位タンパク質	195	MG92-189 転位タンパク質	タンパク質
MG92 転位タンパク質	196	MG92-190 転位タンパク質	タンパク質
MG92 転位タンパク質	197	MG92-191 転位タンパク質	タンパク質
MG92 転位タンパク質	198	MG92-192 転位タンパク質	タンパク質
MG92 転位タンパク質	199	MG92-193 転位タンパク質	タンパク質
MG92 転位タンパク質	200	MG92-194 転位タンパク質	タンパク質
MG92 転位タンパク質	201	MG92-195 転位タンパク質	タンパク質
MG92 転位タンパク質	202	MG92-196 転位タンパク質	タンパク質
MG92 転位タンパク質	203	MG92-197 転位タンパク質	タンパク質
MG92 転位タンパク質	204	MG92-198 転位タンパク質	タンパク質

10

20

30

40

【 0 1 8 6 】

50

【表 2 - 1 3】

触媒	配列番号	説明	タイプ
MG92 転位タンパク質	205	MG92-199 転位タンパク質	タンパク質
MG92 転位タンパク質	206	MG92-200 転位タンパク質	タンパク質
MG92 転位タンパク質	207	MG92-201 転位タンパク質	タンパク質
MG92 転位タンパク質	208	MG92-202 転位タンパク質	タンパク質
MG92 転位タンパク質	209	MG92-203 転位タンパク質	タンパク質
MG92 転位タンパク質	210	MG92-204 転位タンパク質	タンパク質
MG92 転位タンパク質	211	MG92-205 転位タンパク質	タンパク質
MG92 転位タンパク質	212	MG92-206 転位タンパク質	タンパク質
MG92 転位タンパク質	213	MG92-207 転位タンパク質	タンパク質
MG92 転位タンパク質	214	MG92-208 転位タンパク質	タンパク質
MG92 転位タンパク質	215	MG92-209 転位タンパク質	タンパク質
MG92 転位タンパク質	216	MG92-210 転位タンパク質	タンパク質
MG92 転位タンパク質	217	MG92-211 転位タンパク質	タンパク質
MG92 転位タンパク質	218	MG92-212 転位タンパク質	タンパク質
MG92 転位タンパク質	219	MG92-213 転位タンパク質	タンパク質
MG92 転位タンパク質	220	MG92-214 転位タンパク質	タンパク質
MG92 転位タンパク質	221	MG92-215 転位タンパク質	タンパク質

10

20

30

40

【 0 1 8 7 】

50

【表 2 - 1 4】

触媒	配列番号	説明	タイプ
MG92 転位タンパク質	222	MG92-216 転位タンパク質	タンパク質
MG92 転位タンパク質	223	MG92-217 転位タンパク質	タンパク質
MG92 転位タンパク質	224	MG92-218 転位タンパク質	タンパク質
MG92 転位タンパク質	225	MG92-219 転位タンパク質	タンパク質
MG92 転位タンパク質	226	MG92-220 転位タンパク質	タンパク質
MG92 転位タンパク質	227	MG92-221 転位タンパク質	タンパク質
MG92 転位タンパク質	228	MG92-222 転位タンパク質	タンパク質
MG92 転位タンパク質	229	MG92-223 転位タンパク質	タンパク質
MG92 転位タンパク質	230	MG92-224 転位タンパク質	タンパク質
MG92 転位タンパク質	231	MG92-225 転位タンパク質	タンパク質
MG92 転位タンパク質	232	MG92-226 転位タンパク質	タンパク質
MG92 転位タンパク質	233	MG92-227 転位タンパク質	タンパク質
MG92 転位タンパク質	234	MG92-228 転位タンパク質	タンパク質
MG92 転位タンパク質	235	MG92-229 転位タンパク質	タンパク質
MG92 転位タンパク質	236	MG92-230 転位タンパク質	タンパク質
MG92 転位タンパク質	237	MG92-231 転位タンパク質	タンパク質
MG92 転位タンパク質	238	MG92-232 転位タンパク質	タンパク質

10

20

30

40

【 0 1 8 8 】

50

【表 2 - 1 5】

触媒	配列番号	説明	タイプ
MG92 転位タンパク質	239	MG92-233 転位タンパク質	タンパク質
MG92 転位タンパク質	240	MG92-234 転位タンパク質	タンパク質
MG92 転位タンパク質	241	MG92-235 転位タンパク質	タンパク質
MG92 転位タンパク質	242	MG92-236 転位タンパク質	タンパク質
MG92 転位タンパク質	243	MG92-237 転位タンパク質	タンパク質
MG92 転位タンパク質	244	MG92-238 転位タンパク質	タンパク質
MG92 転位タンパク質	245	MG92-239 転位タンパク質	タンパク質
MG92 転位タンパク質	246	MG92-240 転位タンパク質	タンパク質
MG92 転位タンパク質	247	MG92-241 転位タンパク質	タンパク質
MG92 転位タンパク質	248	MG92-242 転位タンパク質	タンパク質
MG92 転位タンパク質	249	MG92-243 転位タンパク質	タンパク質
MG92 転位タンパク質	250	MG92-244 転位タンパク質	タンパク質
MG92 転位タンパク質	251	MG92-245 転位タンパク質	タンパク質
MG92 転位タンパク質	252	MG92-246 転位タンパク質	タンパク質
MG92 転位タンパク質	253	MG92-247 転位タンパク質	タンパク質
MG92 転位タンパク質	254	MG92-248 転位タンパク質	タンパク質
MG92 転位タンパク質	255	MG92-249 転位タンパク質	タンパク質

10

20

30

40

【 0 1 8 9】

50

【表 2 - 1 6】

触媒	配列番号	説明	タイプ
MG92 転位タンパク質	256	MG92-250 転位タンパク質	タンパク質
MG92 転位タンパク質	257	MG92-251 転位タンパク質	タンパク質
MG92 転位タンパク質	258	MG92-252 転位タンパク質	タンパク質
MG92 転位タンパク質	259	MG92-253 転位タンパク質	タンパク質
MG92 転位タンパク質	260	MG92-254 転位タンパク質	タンパク質
MG92 転位タンパク質	261	MG92-255 転位タンパク質	タンパク質
MG92 転位タンパク質	262	MG92-256 転位タンパク質	タンパク質
MG92 転位タンパク質	263	MG92-257 転位タンパク質	タンパク質
MG92 転位タンパク質	264	MG92-258 転位タンパク質	タンパク質
MG92 転位タンパク質	265	MG92-259 転位タンパク質	タンパク質
MG92 転位タンパク質	266	MG92-260 転位タンパク質	タンパク質
MG92 転位タンパク質	267	MG92-261 転位タンパク質	タンパク質
MG92 転位タンパク質	268	MG92-262 転位タンパク質	タンパク質
MG92 転位タンパク質	269	MG92-263 転位タンパク質	タンパク質
MG92 転位タンパク質	270	MG92-264 転位タンパク質	タンパク質
MG92 転位タンパク質	271	MG92-265 転位タンパク質	タンパク質
MG92 転位タンパク質	272	MG92-266 転位タンパク質	タンパク質

10

20

30

40

【 0 1 9 0 】

50

【表 2 - 1 7】

触媒	配列番号	説明	タイプ
MG92 転位タンパク質	273	MG92-267 転位タンパク質	タンパク質
MG92 転位タンパク質	274	MG92-268 転位タンパク質	タンパク質
MG92 転位タンパク質	275	MG92-269 転位タンパク質	タンパク質
MG92 転位タンパク質	276	MG92-270 転位タンパク質	タンパク質
MG92 転位タンパク質	277	MG92-271 転位タンパク質	タンパク質
MG92 転位タンパク質	278	MG92-272 転位タンパク質	タンパク質
MG92 転位タンパク質	279	MG92-273 転位タンパク質	タンパク質
MG92 転位タンパク質	280	MG92-274 転位タンパク質	タンパク質
MG92 転位タンパク質	281	MG92-275 転位タンパク質	タンパク質
MG92 転位タンパク質	282	MG92-276 転位タンパク質	タンパク質
MG92 転位タンパク質	283	MG92-278 転位タンパク質	タンパク質
MG92 転位タンパク質	284	MG92-279 転位タンパク質	タンパク質
MG92 転位タンパク質	285	MG92-280 転位タンパク質	タンパク質
MG92 転位タンパク質	286	MG92-281 転位タンパク質	タンパク質
MG92 転位タンパク質	287	MG92-282 転位タンパク質	タンパク質
MG92 転位タンパク質	288	MG92-283 転位タンパク質	タンパク質
MG92 転位タンパク質	289	MG92-284 転位タンパク質	タンパク質

10

20

30

40

【 0 1 9 1 】

50

【表 2 - 1 8】

触媒	配列番号	説明	タイプ
MG92 転位タンパク質	290	MG92-285 転位タンパク質	タンパク質
MG92 転位タンパク質	291	MG92-286 転位タンパク質	タンパク質
MG92 転位タンパク質	292	MG92-287 転位タンパク質	タンパク質
MG92 転位タンパク質	293	MG92-288 転位タンパク質	タンパク質
MG92 転位タンパク質	294	MG92-290 転位タンパク質	タンパク質
MG92 転位タンパク質	295	MG92-291 転位タンパク質	タンパク質
MG92 転位タンパク質	296	MG92-292 転位タンパク質	タンパク質
MG92 転位タンパク質	297	MG92-293 転位タンパク質	タンパク質
MG92 転位タンパク質	298	MG92-294 転位タンパク質	タンパク質
MG92 転位タンパク質	299	MG92-295 転位タンパク質	タンパク質
MG92 転位タンパク質	300	MG92-296 転位タンパク質	タンパク質
MG92 転位タンパク質	301	MG92-297 転位タンパク質	タンパク質
MG92 転位タンパク質	302	MG92-298 転位タンパク質	タンパク質
MG92 転位タンパク質	303	MG92-299 転位タンパク質	タンパク質
MG92 転位タンパク質	304	MG92-300 転位タンパク質	タンパク質
MG92 転位タンパク質	305	MG92-301 転位タンパク質	タンパク質
MG92 転位タンパク質	306	MG92-302 転位タンパク質	タンパク質

10

20

30

40

【 0 1 9 2】

50

【表 2 - 1 9】

触媒	配列番号	説明	タイプ
MG92 転位タンパク質	307	MG92-303 転位タンパク質	タンパク質
MG92 転位タンパク質	308	MG92-304 転位タンパク質	タンパク質
MG92 転位タンパク質	309	MG92-305 転位タンパク質	タンパク質
MG92 転位タンパク質	310	MG92-306 転位タンパク質	タンパク質
MG92 転位タンパク質	311	MG92-307 転位タンパク質	タンパク質
MG92 転位タンパク質	312	MG92-308 転位タンパク質	タンパク質
MG92 転位タンパク質	313	MG92-309 転位タンパク質	タンパク質
MG92 転位タンパク質	314	MG92-310 転位タンパク質	タンパク質
MG92 転位タンパク質	315	MG92-311 転位タンパク質	タンパク質
MG92 転位タンパク質	316	MG92-312 転位タンパク質	タンパク質
MG92 転位タンパク質	317	MG92-313 転位タンパク質	タンパク質
MG92 転位タンパク質	318	MG92-314 転位タンパク質	タンパク質
MG92 転位タンパク質	319	MG92-315 転位タンパク質	タンパク質
MG92 転位タンパク質	320	MG92-316 転位タンパク質	タンパク質
MG92 転位タンパク質	321	MG92-317 転位タンパク質	タンパク質
MG92 転位タンパク質	322	MG92-318 転位タンパク質	タンパク質
MG92 転位タンパク質	323	MG92-319 転位タンパク質	タンパク質

10

20

30

40

【 0 1 9 3】

50

【表 2 - 2 0】

触媒	配列番号	説明	タイプ
MG92 転位タンパク質	324	MG92-320 転位タンパク質	タンパク質
MG92 転位タンパク質	325	MG92-321 転位タンパク質	タンパク質
MG92 転位タンパク質	326	MG92-322 転位タンパク質	タンパク質
MG92 転位タンパク質	327	MG92-323 転位タンパク質	タンパク質
MG92 転位タンパク質	328	MG92-324 転位タンパク質	タンパク質
MG92 転位タンパク質	329	MG92-325 転位タンパク質	タンパク質
MG92 転位タンパク質	330	MG92-326 転位タンパク質	タンパク質
MG92 転位タンパク質	331	MG92-327 転位タンパク質	タンパク質
MG92 転位タンパク質	332	MG92-328 転位タンパク質	タンパク質
MG92 転位タンパク質	333	MG92-330 転位タンパク質	タンパク質
MG92 転位タンパク質	334	MG92-332 転位タンパク質	タンパク質
MG92 転位タンパク質	335	MG92-334 転位タンパク質	タンパク質
MG92 転位タンパク質	336	MG92-336 転位タンパク質	タンパク質
MG92 転位タンパク質	337	MG92-338 転位タンパク質	タンパク質
MG92 転位タンパク質	338	MG92-340 転位タンパク質	タンパク質
MG92 転位タンパク質	339	MG92-341 転位タンパク質	タンパク質
MG92 転位タンパク質	340	MG92-342 転位タンパク質	タンパク質

10

20

30

40

【 0 1 9 4 】

50

【表 2 - 2 1】

触媒	配列番号	説明	タイプ
MG92 転位タンパク質	341	MG92-343 転位タンパク質	タンパク質
MG92 転位タンパク質	342	MG92-344 転位タンパク質	タンパク質
MG92 転位タンパク質	343	MG92-345 転位タンパク質	タンパク質
MG92 転位タンパク質	344	MG92-346 転位タンパク質	タンパク質
MG92 転位タンパク質	345	MG92-347 転位タンパク質	タンパク質
MG92 転位タンパク質	346	MG92-348 転位タンパク質	タンパク質
MG92 転位タンパク質	347	MG92-349 転位タンパク質	タンパク質
MG92 転位タンパク質	348	MG92-350 転位タンパク質	タンパク質
MG92 転位タンパク質	349	MG92-351 転位タンパク質	タンパク質
MG92 トランスポゾン端	350	MG92-1-A トランスポゾン左端(LE)	ヌクレオチド
MG92 トランスポゾン端	351	MG92-1-A トランスポゾン右端(RE)	ヌクレオチド
MG92 トランスポゾン端	352	MG92-2-A トランスポゾン左端(LE)	ヌクレオチド
MG92 トランスポゾン端	353	MG92-2-A トランスポゾン右端(RE)	ヌクレオチド
MG92 トランスポゾン端	354	MG92-3-A トランスポゾン右端(RE)	ヌクレオチド
MG92 トランスポゾン端	355	MG92-3-A トランスポゾン左端(LE)	ヌクレオチド
MG92 トランスポゾン端	356	MG92-4-A トランスポゾン左端(LE)	ヌクレオチド
MG92 トランスポゾン端	357	MG92-4-A トランスポゾン右端(RE)	ヌクレオチド

10

20

30

40

【 0 1 9 5 】

50

【表 2 - 2 2】

触媒	配列番号	説明	タイプ
MG92 トランスポゾン端	358	MG92-5-A トランスポゾン右端(RE)	ヌクレオチド
MG92 トランスポゾン端	359	MG92-5-A トランスポゾン左端(LE)	ヌクレオチド
MG92 トランスポゾン端	360	MG92-6-A トランスポゾン右端(RE)	ヌクレオチド
MG92 トランスポゾン端	361	MG92-6-A トランスポゾン左端(LE)	ヌクレオチド
MG92 トランスポゾン端	362	MG92-7-A トランスポゾン左端(LE)	ヌクレオチド
MG92 トランスポゾン端	363	MG92-7-A トランスポゾン右端(RE)	ヌクレオチド
MG92 トランスポゾン端	364	MG92-9-A トランスポゾン左端(LE)	ヌクレオチド
MG92 トランスポゾン端	365	MG92-9-A トランスポゾン右端(RE)	ヌクレオチド
MG92 トランスポゾン端	366	MG92-11 トランスポゾン右端(RE)	ヌクレオチド
MG92 トランスポゾン端	367	MG92-11 トランスポゾン左端(LE)	ヌクレオチド
MG92 トランスポゾン端	368	MG92-17 トランスポゾン左端(LE)	ヌクレオチド
MG92 トランスポゾン端	369	MG92-17 トランスポゾン右端(RE)	ヌクレオチド
MG92 トランスポゾン端	370	MG92-20 トランスポゾン左端(LE)	ヌクレオチド
MG92 トランスポゾン端	371	MG92-20 トランスポゾン右端(RE)	ヌクレオチド
MG92 トランスポゾン端	372	MG92-21 トランスポゾン右端(RE)	ヌクレオチド
MG92 トランスポゾン端	373	MG92-21 トランスポゾン左端(LE)	ヌクレオチド
MG92 トランスポゾン端	374	MG92-27 トランスポゾン左端(LE)	ヌクレオチド

10

20

30

40

【 0 1 9 6 】

50

【表 2 - 2 3】

触媒	配列番号	説明	タイプ
MG92 トランスボゾン端	375	MG92-27 トランスボゾン右端(RE)	ヌクレオチド
MG92 トランスボゾン端	376	MG92-28 トランスボゾン右端(RE)	ヌクレオチド
MG92 トランスボゾン端	377	MG92-28 トランスボゾン左端(LE)	ヌクレオチド
MG92 トランスボゾン端	378	MG92-37 トランスボゾン左端(LE)	ヌクレオチド
MG92 トランスボゾン端	379	MG92-37 トランスボゾン右端(RE)	ヌクレオチド
MG92 トランスボゾン端	380	MG92-86 トランスボゾン左端(LE)	ヌクレオチド
MG92 トランスボゾン端	381	MG92-86 トランスボゾン右端(RE)	ヌクレオチド
MG92 トランスボゾン端	382	MG92-136 トランスボゾン右端(RE)	ヌクレオチド
MG92 トランスボゾン端	383	MG92-136 トランスボゾン左端(LE)	ヌクレオチド
MG92 トランスボゾン端	384	MG92-138 トランスボゾン右端(RE)	ヌクレオチド
MG92 トランスボゾン端	385	MG92-138 トランスボゾン左端(LE)	ヌクレオチド
MG92 トランスボゾン端	386	MG92-155、MG92-160 トランスボゾン左端(LE)	ヌクレオチド
MG92 トランスボゾン端	387	MG92-155、MG92-160 トランスボゾン右端(RE)	ヌクレオチド
MG92 トランスボゾン端	388	MG92-157 トランスボゾン右端(RE)	ヌクレオチド
MG92 トランスボゾン末端	389	MG92-157 トランスボゾン左端(LE)	ヌクレオチド
MG92 トランスボゾン端	390	MG92-159 トランスボゾン右端(RE)	ヌクレオチド
MG92 トランスボゾン端	391	MG92-159 トランスボゾン左端(LE)	ヌクレオチド

10

20

30

40

【 0 1 9 7 】

50

【表 2 - 2 4】

触媒	配列番号	説明	タイプ
MG92 トランスポゾン端	392	MG92-162 トランスポゾン右端(RE)	ヌクレオチド
MG92 トランスポゾン端	393	MG92-162 トランスポゾン左端(LE)	ヌクレオチド
MG92 トランスポゾン端	394	MG92-163 トランスポゾン左端(LE)	ヌクレオチド
MG92 トランスポゾン端	395	MG92-163 トランスポゾン右端(RE)	ヌクレオチド
MG92 トランスポゾン端	396	MG92-164 トランスポゾン右端(RE)	ヌクレオチド
MG92 トランスポゾン端	397	MG92-164 トランスポゾン左端(LE)	ヌクレオチド
MG92 トランスポゾン端	398	MG92-165 トランスポゾン右端(RE)	ヌクレオチド
MG92 トランスポゾン端	399	MG92-165 トランスポゾン左端(LE)	ヌクレオチド
MG92 トランスポゾン端	400	MG92-172 トランスポゾン左端(LE)	ヌクレオチド
MG92 トランスポゾン端	401	MG92-172 トランスポゾン右端(RE)	ヌクレオチド
MG92 トランスポゾン端	402	MG92-174 トランスポゾン右端(RE)	ヌクレオチド
MG92 トランスポゾン端	403	MG92-174 トランスポゾン左端(LE)	ヌクレオチド
MG92 トランスポゾン端	404	MG92-177 トランスポゾン左端(LE)	ヌクレオチド
MG92 トランスポゾン端	405	MG92-177 トランスポゾン右端(RE)	ヌクレオチド
MG92 トランスポゾン端	406	MG92-183 トランスポゾン左端(LE)	ヌクレオチド
MG92 トランスポゾン端	407	MG92-183 トランスポゾン右端(RE)	ヌクレオチド
MG92 トランスポゾン端	408	MG92-185 トランスポゾン左端(LE)	ヌクレオチド

10

20

30

40

【 0 1 9 8 】

50

【表 2 - 2 5】

触媒	配列番号	説明	タイプ
MG92 トランスポゾン端	409	MG92-185 トランスポゾン右端(RE)	ヌクレオチド
MG92 トランスポゾン端	410	MG92-187 トランスポゾン左端(LE)	ヌクレオチド
MG92 トランスポゾン端	411	MG92-187 トランスポゾン右端(RE)	ヌクレオチド
MG92 トランスポゾン端	412	MG92-188 トランスポゾン左端(LE)	ヌクレオチド
MG92 トランスポゾン端	413	MG92-188 トランスポゾン右端(RE)	ヌクレオチド
MG92 トランスポゾン端	414	MG92-189 トランスポゾン左端(LE)	ヌクレオチド
MG92 トランスポゾン端	415	MG92-189 トランスポゾン右端(RE)	ヌクレオチド
MG92 トランスポゾン端	416	MG92-196 トランスポゾン左端(LE)	ヌクレオチド
MG92 トランスポゾン端	417	MG92-196 トランスポゾン右端(RE)	ヌクレオチド
MG92 トランスポゾン端	418	MG92-222 トランスポゾン左端(LE)	ヌクレオチド
MG92 トランスポゾン端	419	MG92-222、MG92-266 トランスポゾン右端(RE)	ヌクレオチド
MG92 トランスポゾン端	420	MG92-224 トランスポゾン右端(RE)	ヌクレオチド
MG92 トランスポゾン端	421	MG92-224 トランスポゾン左端(LE)	ヌクレオチド
MG92 トランスポゾン端	422	MG92-226 トランスポゾン右端(RE)	ヌクレオチド
MG92 トランスポゾン端	423	MG92-226 トランスポゾン左端(LE)	ヌクレオチド
MG92 トランスポゾン端	424	MG92-264 トランスポゾン左端(LE)	ヌクレオチド
MG92 トランスポゾン端	425	MG92-264 トランスポゾン右端(RE)	ヌクレオチド

10

20

30

40

【 0 1 9 9 】

50

【表 2 - 2 6】

触媒	配列番号	説明	タイプ
MG92 トランスボゾン端	426	MG92-266 トランスボゾン左端(LE)	ヌクレオチド
MG92 トランスボゾン端	427	MG92-267 トランスボゾン右端(RE)	ヌクレオチド
MG92 トランスボゾン端	428	MG92-267 トランスボゾン左端(LE)	ヌクレオチド
MG92 トランスボゾン端	429	MG92-272 トランスボゾン右端(RE)	ヌクレオチド
MG92 トランスボゾン端	430	MG92-272 トランスボゾン左端(LE)	ヌクレオチド
MG92 トランスボゾン端	431	MG92-274 トランスボゾン右端(RE)	ヌクレオチド
MG92 トランスボゾン端	432	MG92-274 トランスボゾン左端(LE)	ヌクレオチド
MG92 トランスボゾン端	433	MG92-284 トランスボゾン左端(LE)	ヌクレオチド
MG92 トランスボゾン端	434	MG92-284 トランスボゾン右端(RE)	ヌクレオチド
MG92 トランスボゾン端	435	MG92-288 トランスボゾン左端(LE)	ヌクレオチド
MG92 トランスボゾン端	436	MG92-288 トランスボゾン右端(RE)	ヌクレオチド
MG92 トランスボゾン端	437	MG92-291 トランスボゾン左端(LE)	ヌクレオチド
MG92 トランスボゾン端	438	MG92-291 トランスボゾン右端(RE)	ヌクレオチド
MG92 トランスボゾン端	439	MG92-295 トランスボゾン右端(RE)	ヌクレオチド
MG92 トランスボゾン端	440	MG92-295 トランスボゾン左端(LE)	ヌクレオチド
MG92 トランスボゾン端	441	MG92-302 トランスボゾン右端(RE)	ヌクレオチド
MG92 トランスボゾン端	442	MG92-302 トランスボゾン左端(LE)	ヌクレオチド

10

20

30

40

【 0 2 0 0 】

50

【表 2 - 27】

触媒	配列番号	説明	タイプ
MG92 トランスポゾン端	443	MG92-310 トランスポゾン右端(RE)	ヌクレオチド
MG92 トランスポゾン端	444	MG92-310 トランスポゾン左端(LE)	ヌクレオチド
MG92 トランスポゾン端	445	MG92-311 トランスポゾン左端(LE)	ヌクレオチド
MG92 トランスポゾン端	446	MG92-311 トランスポゾン右端(RE)	ヌクレオチド
MG92 トランスポゾン端	447	MG92-312 トランスポゾン右端(RE)	ヌクレオチド
MG92 トランスポゾン端	448	MG92-312 トランスポゾン左端(LE)	ヌクレオチド
MG92 トランスポゾン端	449	MG92-322 トランスポゾン左端(LE)	ヌクレオチド
MG92 トランスポゾン端	450	MG92-322 トランスポゾン右端(RE)	ヌクレオチド
MG92 トランスポゾン端	451	MG92-323 トランスポゾン左端(LE)	ヌクレオチド
MG92 トランスポゾン端	452	MG92-323 トランスポゾン右端(RE)	ヌクレオチド
MG92 トランスポゾン端	453	MG92-344 トランスポゾン左端(LE)	ヌクレオチド
MG92 トランスポゾン端	454	MG92-344 トランスポゾン右端(RE)	ヌクレオチド

10

20

30

【0201】

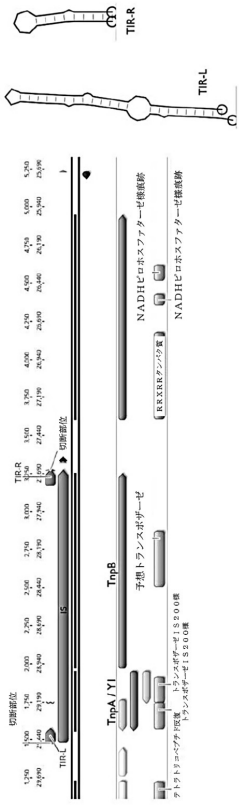
本発明の好ましい実施形態が本明細書に示され、記載されてきたが、そのような実施形態が例示の目的でのみ提供されていることは、当業者には明らかであろう。本発明は、本明細書内で提供される特定の実施例によって限定されることは意図されていない。本発明は前述の説明を参照して記載されているが、本明細書の実施形態の記載及び説明は、限定された意味で解釈されることを意図していない。多数の変形、変更、及び置換は、ここで、本発明から逸脱することなく、当業者にとって生じるであろう。更に、本発明の全ての態様は、様々な条件及び変数に依存する、本明細書に記載される特定の描写、構成又は相対的割合に限定されないことが理解されよう。本明細書に記載される本発明の実施形態に対する様々な代替が、本発明の実施に用いられ得ることは、理解されるべきである。したがって、本発明は、こうした任意の代替、修正、変形、又は均等物も包含することが企図される。以下の特許請求の範囲は本発明の範囲を定義し、これらの特許請求の範囲及びそれらの均等物の範囲内の方法及び構造がそれによって包含されることが意図される。

40

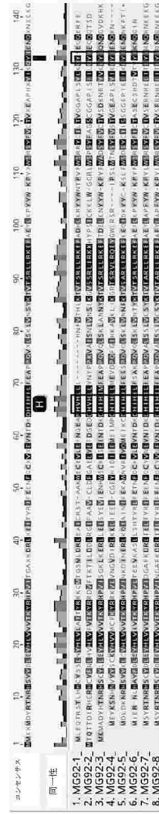
50

【図面】

【図 1 A】



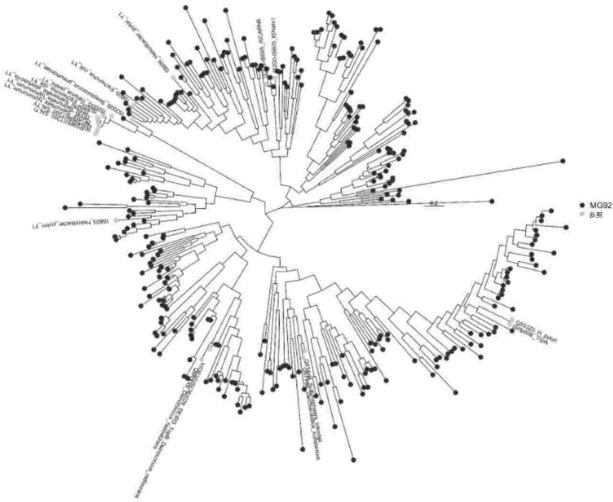
【図 1 B】



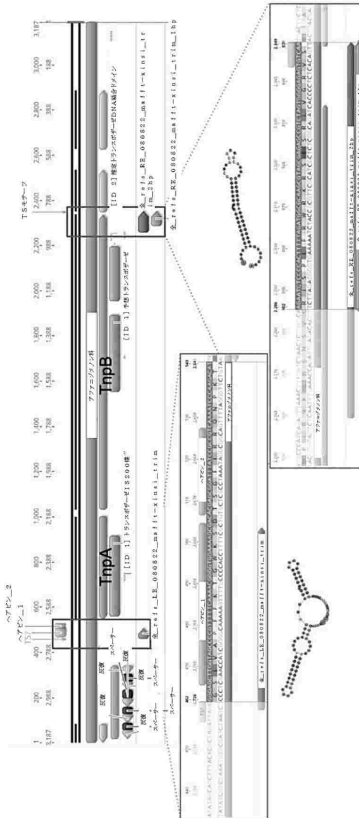
10

20

【図 2】



【図 3】

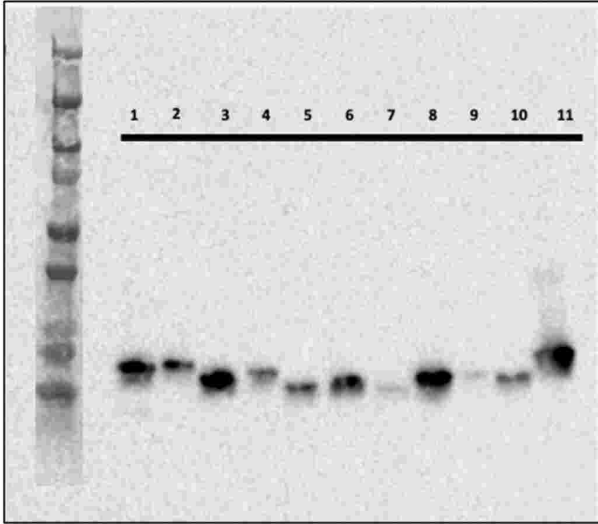


30

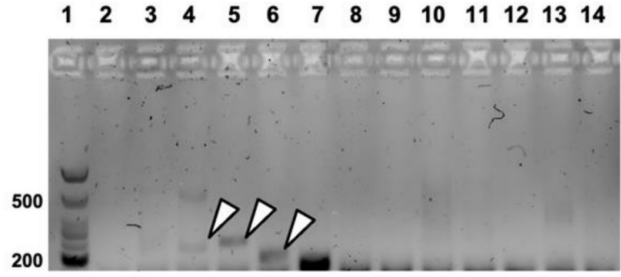
40

50

【 図 4 】

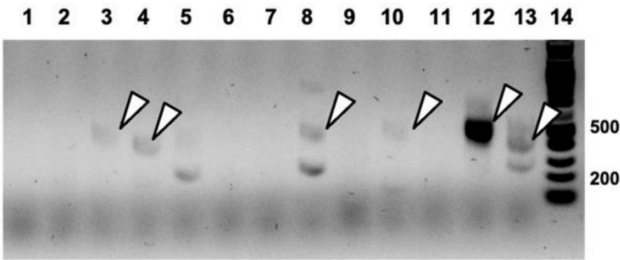


【 図 5 A 】

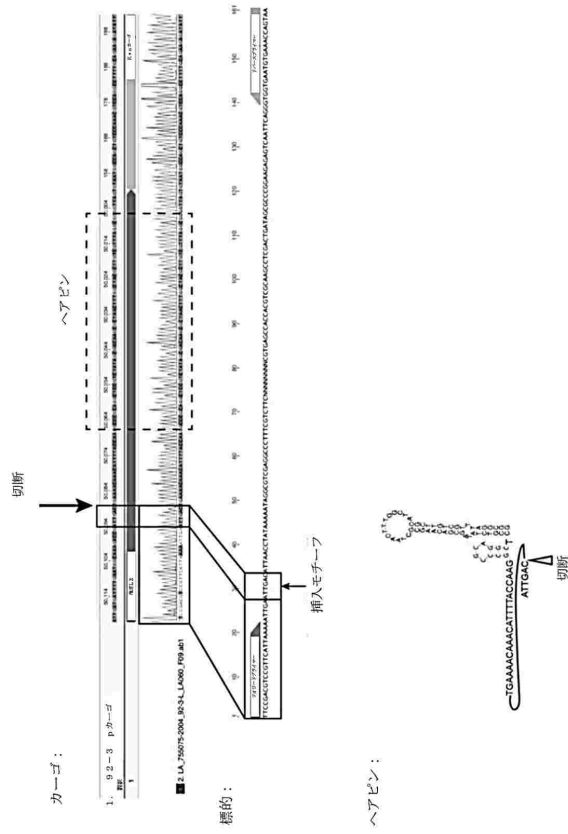


10

【 図 5 B 】



【 図 6 】



20

30

40

50

【配列表】

2024533038000001.xml

10

20

30

40

50

【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		International application No. PCT/US2022/076059
A. CLASSIFICATION OF SUBJECT MATTER C12N 15/62 (2006.01) C12N 15/70 (2006.01) C12N 15/90 (2006.01) C12N 5/10 (2006.01) C12N 9/12 (2006.01)		
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)		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)		
Google Scholar: transposase, IS200/IS605, TpnA, TpnB and like terms. Google Patents: metagenomi, transposase, inventors names GenomeQuest: SEQ ID NOs: 1,3,5,7,9,11,13,15,18,19		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
	Documents are listed in the continuation of Box C	
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C		<input checked="" type="checkbox"/> 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	"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	
"D" document cited by the applicant in the international application	"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	
"E" earlier application or patent but published on or after the international filing date	"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	
"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)	"&" document member of the same patent family	
"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		
Date of the actual completion of the international search 22 December 2022	Date of mailing of the international search report 22 December 2022	
Name and mailing address of the ISA/AU AUSTRALIAN PATENT OFFICE PO BOX 200, WODEN ACT 2606, AUSTRALIA Email address: pct@ipaustralia.gov.au	Authorised officer Dejan Bursac AUSTRALIAN PATENT OFFICE (ISO 9001 Quality Certified Service) Telephone No. +61 2 6283 3197	

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

10

20

30

40

50

INTERNATIONAL SEARCH REPORT		International application No.
C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		PCT/US2022/076059
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	GenBank Accession No. APZ42579 30 January 2017 sequence shares 78.26% identity with SEQ ID NO: 1 and 94.93% identity with SEQ ID NO: 18 whole document	1-5, 7-18, 20-45, 47-69, 73-100, 106, 108-113, 118-120, 126, 128-156 (all in part) and 6, 19, 46, 70-71, 107, 127 (in full)
X	GenBank Accession No. APZ42563 30 January 2017 sequence shares 77.54% identity with SEQ ID NO: 1 and 94.20% identity with SEQ ID NO: 18 whole document	1-5, 7-18, 20-45, 47-69, 73-100, 106, 108-113, 118-120, 126, 128-156 (all in part) and 6, 19, 46, 70-71, 107, 127 (in full)
X	GenBank Accession No. ARU64428 30 May 2017 sequence shares 82.61% identity with SEQ ID NO: 3 whole document	1-5, 7-18, 20-45, 47-69, 73-100, 106, 108-113, 118-120, 126, 128-156 (all in part) and 6, 19, 46, 70-71, 107, 127 (in full)
X	GenBank Accession No. VTP78101 10 May 2019 sequence shares 81.88% sequence identity with SEQ ID NO: 3 whole document	1-5, 7-18, 20-45, 47-69, 73-100, 106, 108-113, 118-120, 126, 128-156 (all in part) and 6, 19, 46, 70-71, 107, 127 (in full)
X	GenBank Accession No. AGQ24374 29 October 2013 sequence shares 81.16% sequence identity with SEQ ID NO: 3 whole document	1-5, 7-18, 20-45, 47-69, 73-100, 106, 108-113, 118-120, 126, 128-156 (all in part) and 6, 19, 46, 70-71, 107, 127 (in full)
X	GenBank Accession No. QUI95803 28 April 2021 sequence shares 80.45% identity with SEQ ID NO: 7 whole document	1-5, 7-18, 20-45, 47-69, 73-100, 106, 108-113, 118-120, 126, 128-156 (all in part) and 6, 19, 46, 70-71, 107, 127 (in full)
	GenBank Accession No. BCL56697 12 December 2020 sequence shares 79.23% identity with SEQ ID NO: 7	

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

10

20

30

40

50

INTERNATIONAL SEARCH REPORT		International application No.
C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		PCT/US2022/076059
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	whole document	1-5, 7-18, 20-45, 47-69, 73-100, 106, 108-113, 118-120, 126, 128-156 (all in part) and 6, 19, 46, 70-71, 107, 127 (in full)
X	GenBank Accession No. BCL57499 12 December 2020 sequence shares 76.92 identity with SEQ ID NO: 7 whole document	1-5, 7-18, 20-45, 47-69, 73-100, 106, 108-113, 118-120, 126, 128-156 (all in part) and 6, 19, 46, 70-71, 107, 127 (in full)
X	D9 GenBank Accession No. AFY32194 11 September 2019 sequence shares 80% sequence identity with SEQ ID NO: 19 whole document	1-5, 7-18, 20-45, 47-69, 73-100, 106, 108-113, 118-120, 126, 128-156 (all in part) and 6, 19, 46, 70-71, 107, 127 (in full)
A	US 20200377945 A1 (PERFECT (CHINA) CO., LTD) 03 December 2020 & sequence listing of US20200377945A1 sequence 12221	
A	Hickman, Alison B., and Fred Dyda. "DNA transposition at work." Chemical reviews 116.20 (2016): 12758-12784.	
A	Barabas, Orsolya, et al. "Mechanism of IS200/IS605 family DNA transposases: activation and transposon-directed target site selection." Cell 132.2 (2008): 208-220.	

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2022/076059

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

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:

- a. forming part of the international application as filed.
- b. furnished subsequent to the international filing date for the purposes of international search (Rule 13*ter*.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.

10

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:

As a result of lack of unity (see below) this opinion has been established on the basis of SEQ ID NOs:
1,3,5,7,9,11,13,15,18,19.

20

30

40

50

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2022/076059

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. <input type="checkbox"/>	Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely: the subject matter listed in Rule 39 on which, under Article 17(2)(a)(i), an international search is not required to be carried out, including	10
2. <input type="checkbox"/>	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:	20
3. <input type="checkbox"/>	Claims Nos.: 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:		
See Supplemental Box for Details		
1. <input type="checkbox"/>	As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.	
2. <input type="checkbox"/>	As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.	30
3. <input type="checkbox"/>	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. <input checked="" type="checkbox"/>	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.: 1-5, 7-18, 20-45, 47-69, 73-100, 106, 108-113, 118-120, 126, 128-156 (all in part) and 6, 19, 46, 70-71, 107, 127 (in full)	
Remark on Protest	<input type="checkbox"/> The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.	40
	<input type="checkbox"/> 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.	
	<input type="checkbox"/> No protest accompanied the payment of additional search fees.	

INTERNATIONAL SEARCH REPORT	International application No. PCT/US2022/076059
Supplemental Box	
<p>Continuation of: Box III</p> <p>This International Application does not comply with the requirements of unity of invention because it does not relate to one invention or to a group of inventions so linked as to form a single general inventive concept.</p> <p>This Authority has found that there are different inventions based on the following features that separate the claims into distinct groups:</p> <p>Inventions 1-349: transposase systems and methods of using transposases systems, comprising transposases defined by each SEQ ID NOs: 1-349. Each transposase sequence is considered to be a separate invention.</p> <p>Inventions 350-454: transposase systems and methods of using transposase systems, comprising flanking sequences defined by each SEQ ID NOs: 350-454. Each flanking sequence is considered to be a separate invention.</p> <p>PCT Rule 13.2, first sentence, states that unity of invention is only fulfilled when there is a technical relationship among the claimed inventions involving one or more of the same or corresponding special technical features. PCT Rule 13.2, second sentence, defines a special technical feature as a feature which makes a contribution over the prior art. When there is no special technical feature common to all the claimed inventions there is no unity of invention.</p> <p>In the above groups of claims, the identified features may have the potential to make a contribution over the prior art but are not common to all the claimed inventions and therefore cannot provide the required technical relationship. Examples 1 and 8 of the present specification teach that these transposases were discovered by applying Hidden Markov Models of known transposases, known catalytic residues and proximity to known IS200/IS605 (specifically TnpA transposase) insertion sequences. Therefore, the only feature common to claimed inventions 1-349 is that they are transposases of a known family. Similarly, the only feature common to inventions 350-454 is that they are nucleotide sequences of transposon ends (para [0058] of the present specification). Transposases and transposon ends are well known in the art and are disclosed in, for example:</p> <p>Barabas, Orsolya, et al. "Mechanism of IS200/IS605 family DNA transposases: activation and transposon-directed target site selection." Cell 132.2 (2008): 208-220.</p> <p>Therefore these common features cannot be a technical feature and thus the requirements for unity of invention are consequently not satisfied a posteriori (in particular see rule 10.56 of the PCT International and preliminary examination guidelines) .The ISA will search and establish an opinion on transposase systems comprising transposases defined by SEQ ID NOs: 1, 3, 5, 7, 9, 11, 13, 15, 18-19, being defined by claims 1-5, 7-18, 20-45, 47-69, 73-100, 106, 108-113, 118-120, 126, 128-156 (all in part) and 6, 19, 46, 70-71,107, 127 (in full).</p>	

10

20

30

40

50

INTERNATIONAL SEARCH REPORT		International application No. PCT/US2022/076059	
Information on patent family members			
This Annex lists known patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.			
Patent Document/s Cited in Search Report		Patent Family Member/s	
Publication Number	Publication Date	Publication Number	Publication Date
US 20200377945 A1	03 December 2020	US 2020377945 A1	03 Dec 2020
		CN 110527717 A	03 Dec 2019
		TW 201934760 A	01 Sep 2019
		TW I707043 B	11 Oct 2020
		WO 2019149247 A1	08 Aug 2019
End of Annex			
Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001. Form PCT/ISA/210 (Family Annex)(July 2019)			

10

20

30

40

50

フロントページの続き

(51)国際特許分類

F I

テーマコード (参考)

C 1 2 N	1/21 (2006.01)	C 1 2 N	1/21		
C 1 2 N	5/10 (2006.01)	C 1 2 N	5/10		
C 1 2 N	15/09 (2006.01)	C 1 2 N	15/09	1 0 0	

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,N
E,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,J
M,JO,JP,KE,KG,KH,KN,KP,KR,KW,KZ,LA,LC,LK,LR,LS,LU,LY,MA,MD,ME,MG,MK,MN,MW,MX,MY
,MZ,NA,NG,NI,NO,NZ,OM,PA,PE,PG,PH,PL,PT,QA,RO,RS,RU,RW,SA,SC,SD,SE,SG,SK,SL,ST,SV,SY,T
H,TJ,TM,TN,TR,TT,TZ,UA,UG,US,UZ,VC,VN,WS,ZA,ZM,ZW

(72)発明者 ゴルツマン, ダニエラ エス. エー.

アメリカ合衆国 9 4 6 0 8 カリフォルニア州 エメリービル パーク・アベニュー 1 5 4 5

(72)発明者 アレクサンダー, リサ

アメリカ合衆国 9 4 6 0 8 カリフォルニア州 エメリービル パーク・アベニュー 1 5 4 5

(72)発明者 ラペリエール, サラ

アメリカ合衆国 9 4 6 0 8 カリフォルニア州 エメリービル パーク・アベニュー 1 5 4 5

F ターム (参考) 4B065 AA01X AA01Y AA57X AA72X AA88X AA90X AB01 AC14 BA02 CA24
CA60