

(19) 日本国特許庁(JP)

## (12) 特許公報(B2)

(11) 特許番号

特許第6490710号  
(P6490710)

(45) 発行日 平成31年3月27日(2019.3.27)

(24) 登録日 平成31年3月8日(2019.3.8)

(51) Int.Cl.

F 1

C 12 N 15/62	(2006.01)	C 12 N 15/62	Z N A Z
C 12 N 9/10	(2006.01)	C 12 N 9/10	
C 12 N 1/15	(2006.01)	C 12 N 1/15	
C 12 N 1/19	(2006.01)	C 12 N 1/19	
C 12 N 1/21	(2006.01)	C 12 N 1/21	

請求項の数 18 (全 59 頁) 最終頁に続く

(21) 出願番号 特願2016-562588 (P2016-562588)  
 (86) (22) 出願日 平成27年4月15日 (2015.4.15)  
 (65) 公表番号 特表2017-513479 (P2017-513479A)  
 (43) 公表日 平成29年6月1日 (2017.6.1)  
 (86) 國際出願番号 PCT/US2015/025889  
 (87) 國際公開番号 WO2015/160895  
 (87) 國際公開日 平成27年10月22日 (2015.10.22)  
 審査請求日 平成30年4月13日 (2018.4.13)  
 (31) 優先権主張番号 61/979,871  
 (32) 優先日 平成26年4月15日 (2014.4.15)  
 (33) 優先権主張国 米国(US)  
 (31) 優先権主張番号 62/062,006  
 (32) 優先日 平成26年10月9日 (2014.10.9)  
 (33) 優先権主張国 米国(US)

(73) 特許権者 514202402  
 イラミーナ インコーポレーテッド  
 アメリカ合衆国 カリフォルニア州 92  
 122 サンディエゴ イラミーナ ウェ  
 イ 5200  
 (74) 代理人 100147485  
 弁理士 杉村 憲司  
 (74) 代理人 100181272  
 弁理士 神 紘一郎  
 (74) 代理人 100193437  
 弁理士 高木 義和  
 (72) 発明者 クリストフアン グレックナー  
 ドイツ国 53111 ボン ヴォルフシ  
 ュトラーセ 10エー

最終頁に続く

(54) 【発明の名称】 改善された挿入配列バイアスおよび拡大されたDNAインプット許容差のための修正トランスポザーゼ

## (57) 【特許請求の範囲】

## 【請求項1】

トランスポザーゼ活性を有し、Pro214および/またはAla338の位置に対応する変異を含む配列番号1または配列番号11を含み、

任意に、配列番号1または配列番号11におけるGlu54および/またはMet56および/またはLys212および/またはLeu372と機能的に同等な位置で置換変異を含み、

任意に、Pro214および/またはAla338の位置で前記変異は置換変異である、変異Tn5トランスポザーゼ。

## 【請求項2】

配列番号1または配列番号11におけるPro214、Ala338、およびLys212の位置で変異を含み、

任意に、配列番号1または配列番号11におけるGlu54および/またはMet56および/またはLeu372と機能的に同等な位置で置換変異を含む、請求項1に記載の変異Tn5トランスポザーゼ。

## 【請求項3】

位置Lys212での前記置換変異はアルギニンへの変異を含む、請求項1または2に記載の変異Tn5トランスポザーゼ。

## 【請求項4】

位置Pro214での前記置換変異はアルギニンへの変異を含む、請求項1~3の何れ

かに記載の変異 T n 5 トランスポザーゼ。

【請求項 5】

位置 A 1 a 3 3 8 での前記置換変異はバリンへの変異を含む、請求項 1 ~ 4 の何れかに記載の変異 T n 5 トランスポザーゼ。

【請求項 6】

前記トランスポザーゼはさらに、G 1 y 2 5 1 での置換変異を含み、

任意に、位置 G 1 y 2 5 1 での前記置換変異はアルギニンへの変異を含む、請求項 1 ~ 5 の何れかに記載の変異 T n 5 トランスポザーゼ。

【請求項 7】

ポリペプチド融合ドメインに融合した、請求項 1 ~ 6 の何れかに記載の変異 T n 5 トランスポザーゼを含む、融合タンパク質。10

【請求項 8】

前記ポリペプチド融合ドメインは前記変異 T n 5 トランスポザーゼの N 末端に融合し、20

前記ポリペプチド融合ドメインは前記変異 T n 5 トランスポザーゼの C 末端に融合する、請求項 7 に記載の融合タンパク質。

【請求項 9】

前記ポリペプチド融合ドメインは可溶性を高めるタグを含む、または

前記ポリペプチド融合ドメインは、マルトース結合タンパク質 (MBP) 、伸長因子 Ts (Tsf) 、5 - メチルシトシン結合ドメイン、およびタンパク質 A から選択されるドメインを含む、請求項 8 に記載の融合タンパク質。20

【請求項 10】

請求項 1 ~ 9 の何れかで定義される変異または融合タンパク質 T n 5 トランスポザーゼをコードする、核酸分子。

【請求項 11】

請求項 1 0 に記載の核酸分子を含む、発現ベクター。

【請求項 12】

請求項 1 1 のベクターを含む、宿主細胞。

【請求項 13】

次の構成要素：(i) 請求項 1 ~ 9 の何れか一項に記載の変異または融合タンパク質 T n 5 トランスポザーゼを含むトランスポソーム複合体と；(ii) 標的 DNA とが、相互作用することを可能にするステップを含む、in vitro 転移のための方法。30

【請求項 14】

標的 DNA を配列決定するための方法であって、

(a) 前記標的 DNA をトランスポソーム複合体と共にインキュベーションするステップであって、該トランスポソーム複合体は、

(1) 請求項 1 ~ 7 の何れか一項に記載の変異または融合タンパク質 T n 5 トランスポザーゼと；

(2) 第 1 ポリヌクレオチドとを含み、該第 1 ポリヌクレオチドは、

(i) トランスポゾン端配列を有する 3' 部分と、

(ii) 第 1 配列決定タグドメインを有する第 1 タグとを備え、40

該ステップは、前記標的 DNA を断片化し、そして、前記第 1 ポリヌクレオチドの 3' トランスポゾン端配列を前記断片の 5' 端に移動させる条件下で行われ、

それにより、二本鎖断片が生成され、ここにおいて前記 5' 端は前記第 1 タグでタグ付けされ、前記 5' タグ付き鎖の 3' 端には一本鎖ギャップがある、ステップと；

(b) 第 2 タグが前記 5' タグ付き鎖の 3' 端に付着する条件下で、前記断片を核酸修飾酵素と共にインキュベーションするステップと；

(c) ポリメラーゼおよび前記第 1 ポリヌクレオチドの一部分に対応する増幅プライマーを提供することにより、前記断片を任意に増幅し、それにより、前記 5' 端に前記第 1 タグを、前記 3' 端に第 2 タグを有する、2 つタグ付けされた断片の代表ライブラリを生成するステップと；50

(d) 前記第1配列決定タグドメインに対応する部分を含む第1配列決定プライマーを提供するステップと；

(e) 前記第1配列決定プライマーを伸長し、並行して、2つタグ付けされた断片の前記代表ライプラリの第1配列決定タグドメインに隣接するヌクレオチドの同一性を検出するステップと、を含む方法。

【請求項15】

(1) 請求項1～9の何れか一項に記載の変異または融合タンパク質Tn5トランスポザーゼと；

(2) トランスポゾン端配列を有する3'部分を備えるポリヌクレオチドとを備えたトランスポーム複合体を含む、in vitro転移反応を実行するためのキット。 10

【請求項16】

次の構成要素：(i) トランスポザーゼ酵素を含むトランスポーム複合体と；(ii) 標的DNAと、(iii) Mg<sup>2+</sup>を含むバッファ中の同一酵素と比較して、GCバイアスを縮小するのに十分効果的な濃度でCo<sup>2+</sup>を含む反応バッファとが、相互作用することを可能にするステップを含み、

前記トランスポザーゼ酵素は、請求項1～9の何れか一項に記載の変異または融合タンパク質Tn5トランスポザーゼを含む、in vitro転移のための方法。

【請求項17】

配列番号26を含む、請求項7の融合タンパク質。

【請求項18】

配列番号19のアミノ酸配列を含む、請求項1の変異Tn5トランスポザーゼ。 20

【発明の詳細な説明】

【技術分野】

【0001】

【関連出願】

本出願は、2014年4月15日に出願された米国仮特許出願第61/979,871号、2014年10月9日に出願された同第62/062,006号、2014年11月17日に出願された同第62/080,882号の優先権を主張し、これらはその全体が参照により本明細書に組み込まれる。 30

【背景技術】

【0002】

トランスポザーゼ酵素は、in vitro転移系において有用である。これはゲノムDNAの大量の断片化とタグ付けを可能にし、次世代シーケンシングなどの核酸分析法および增幅法で用いるタグ付きDNA断片のライプラリを、標的DNAから作るのに有用である。改善された特性を有し、元になるサンプルの標的核酸を質的および量的に代表するタグ付きDNA断片を生成する修正トランスポザーゼは、いまだ必要とされている。

【0003】

【配列表】

本出願は、電子フォーマットの配列表と共に出願されている。配列表は、2015年4月13日作成のIP-1198-TW\_Sequencelisting.txtというファイルで提供され、サイズは103Kbである。配列表の電子フォーマットの情報は、その全体が参照により本明細書に組み込まれる。

【発明の概要】

【0004】

本明細書で提示するのは、改善された断片化および核酸サンプルのタグ付けのためのトランスポザーゼ酵素である。本発明者らは、驚くべきことに、挿入配列バイアスを改善し他の多くの関連利点を有する、ある変更トランスポザーゼを特定した。

【0005】

本明細書で提示するのは、野生型Tn5トランスポザーゼと比較し修正した変異Tn5 50

トランスポザーゼである。一部の実施形態において、該変異トランスポザーゼは位置 A s p 2 4 8 で変異を含むことが可能である、ある態様では、位置 A s p 2 4 8 における変異は置換変異である。ある態様では、位置 A s p 2 4 8 における置換変異には、T y r 、T h r 、L y s 、S e r 、L e u 、A l a 、T r p 、P r o 、G l n 、A r g 、P h e 、およびH i s からなる群から選択される残基への変異が含まれ得る。

【 0 0 0 6 】

ある態様では、位置 A s p 2 4 8 での変異は、位置 A s p 2 4 8 の後ろにおける挿入変異である。ある態様では、挿入変異には、位置 A s p 2 4 8 の後ろへの疎水性残基の挿入が含まれ得る。ある態様では、挿入変異には、位置 A s p 2 4 8 の後ろへのバリン残基の挿入が含まれ得る。

10

【 0 0 0 7 】

本明細書では、野生型 T n 5 トランスポザーゼと比較し修正した変異 T n 5 トランスポザーゼであって、位置 A s p 1 1 9 で変異を含む変異トランスポザーゼも提示する。ある態様では、位置 A s p 1 1 9 での変異は置換変異である。ある態様では、位置 A s p 1 1 9 での置換変異には、疎水性残基への変異が含まれ得る。ある態様では、位置 A s p 1 1 9 での置換変異には、親水性残基への変異が含まれ得る。ある態様では、位置 A s p 1 1 9 での置換変異には、L e u 、M e t 、S e r 、A l a 、およびV a l からなる群から選択される残基への変異が含まれ得る。

【 0 0 0 8 】

本明細書では、野生型 T n 5 トランスポザーゼと比較し修正した変異 T n 5 トランスポザーゼであって、位置 T r p 1 2 5 で変異を含む変異トランスポザーゼも提示する。ある態様では、T r p 1 2 5 での変異は置換変異である。ある態様では、位置 T r p 1 2 5 での置換変異には、メチオニン残基への変異が含まれ得る。

20

【 0 0 0 9 】

本明細書では、野生型 T n 5 トランスポザーゼと比較し修正した変異 T n 5 トランスポザーゼであって、位置 L y s 1 2 0 で変異を含む変異 T n 5 トランスポザーゼも提示する。ある態様では、位置 L y s 1 2 0 での変異は置換変異である。ある態様では、位置 L y s 1 2 0 での置換変異には、嵩高い芳香族残基への変異が含まれ得る。ある態様では、位置 L y s 1 2 0 での置換変異には、T y r 、P h e 、T r p 、およびG l u からなる群から選択される残基への変異が含まれ得る。

30

【 0 0 1 0 】

本明細書では、野生型 T n 5 トランスポザーゼと比較し修正した変異 T n 5 トランスポザーゼであって、L y s 2 1 2 および / またはP r o 2 1 4 および / またはA l a 3 3 8 の位置で変異を含む変異トランスポザーゼも提示する。ある態様では、L y s 2 1 2 および / またはP r o 2 1 4 および / またはA l a 3 3 8 の位置での変異は、置換変異である。ある態様では、位置 L y s 2 1 2 での置換変異には、アルギニンへの変異が含まれる。ある態様では、位置 P r o 2 1 4 での置換変異には、アルギニンへの変異が含まれる。ある態様では、位置 A l a 3 3 8 での置換変異には、バリンへの変異が含まれる。一部の実施形態では、トランスポザーゼは、G l y 2 5 1 での置換変異をさらに含むことが可能である。ある態様では、位置 G l y 2 5 1 での置換変異には、アルギニンへの変異が含まれる。

40

【 0 0 1 1 】

本明細書では、野生型 T n 5 トランスポザーゼと比較し修正した変異 T n 5 トランスポザーゼであって、G l u 1 4 6 および / またはG l u 1 9 0 および / またはG l y 2 5 1 の位置で変異を含む変異トランスポザーゼも提示する。ある態様では、G l u 1 4 6 および / またはG l u 1 9 0 および / またはG l y 2 5 1 の位置での変異は、置換変異である。ある態様では、位置 G l u 1 4 6 での置換変異には、グルタミンへの変異が含まれ得る。ある態様では、位置 G l u 1 9 0 での置換変異には、グリシンへの変異が含まれ得る。ある態様では、位置 G l y 2 5 1 での置換変異には、アルギニンへの変異が含まれ得る。

【 0 0 1 2 】

50

配列番号 21 のアミノ酸配列を備える半保存されたドメインへの置換変異を含む改変トランスポザーゼも提示する。該置換変異には、位置 2 における、Trp、Asn、Val、またはLys 以外の任意の残基への変異が含まれる。ある実施形態では、該変異には位置 2 における Met への置換が含まれる。

【0013】

上記実施形態の何れかにおいて、変異 Tn5 トランスポザーゼはさらに、Tn5 トランスポザーゼのアミノ酸配列の Glu54 および / または Met56 および / または Leu372 と機能的に同等な位置で置換変異を含むことが可能である。ある実施形態では、トランスポザーゼは、Tn5 トランスポザーゼのアミノ酸配列の Glu54 Lys および / または Met56 Ala および / または Leu372 Pro に対し相同な置換変異を含む。  
10

【0014】

本明細書では、配列番号 2 ~ 10 および 12 ~ 20 の何れか一つのアミノ酸配列を含む変異 Tn5 トランスポザーゼも提示する。

【0015】

本明細書では、追加ポリペプチドに融合した、上記実施形態の何れかで定義した変異 Tn5 トランスポザーゼを含む融合タンパク質も提示する。一部の実施形態では、トランスポザーゼに融合したポリペプチドドメインは、精製タグ、発現タグ、溶解タグ、またはその組合せを含むことが可能である。一部の実施形態では、トランスポザーゼに融合したポリペプチドドメインは、例えばマルトース結合タンパク質 (Maltose Binding Protein, MBP) を含むことが可能である。一部の実施形態では、トランスポザーゼに融合したポリペプチドドメインは、例えば伸長因子 Ts (Elongation Factor Ts, Ts f) を含むことが可能である。  
20

【0016】

本明細書では、上記実施形態の何れかで定義した変異 Tn5 トランスポザーゼをコードする核酸分子も提示する。本明細書では、上記核酸分子を含む発現ベクターも提示する。本明細書では、上記ベクターを含む宿主細胞も提示する。

【0017】

本明細書では、次の構成要素：(i) 上記実施形態の何れか一つに記載の変異 Tn5 トランスポザーゼを含むトランスポソーム複合体と；(ii) 標的 DNA とが、相互作用することを可能にするステップを含む、in vitro 転移のための方法も提示する。  
30

【0018】

本明細書では、上記の Tn5 トランスポザーゼを利用して標的 DNA を配列決定する方法も提示する。一部の実施形態で、該方法は、(a) 標的 DNA をトランスポソーム複合体と共にインキュベーションするステップであって、該トランスポソーム複合体は、(1) 上記実施形態の何れか一つに記載の変異 Tn5 トランスポザーゼと；(2) 第 1 ポリヌクレオチドとを含み、該第 1 ポリヌクレオチドは、(i) トランスポゾン端配列を有する 3' 部分と、(ii) 第 1 配列決定タグドメインを有する第 1 タグとを備え、該ステップは、前記標的 DNA を断片化し、そして、前記第 1 ポリヌクレオチドの 3' トランスポゾン端配列を前記断片の 5' 端に移動させる条件下で行われ、それにより二本鎖断片が生成され、ここにおいて前記 5' 端は前記第 1 タグでタグ付けされ、前記 5' タグ付き鎖の 3' 端には一本鎖ギャップがある、ステップと；(b) 第 2 タグが前記 5' タグ付き鎖の 3' 端に付着する条件下で、前記断片を核酸修飾酵素と共にインキュベーションするステップと；(c) ポリメラーゼおよび前記第 1 ポリヌクレオチドの一部に対応する増幅プライマーを提供することにより、前記断片を任意に増幅し、それにより、前記 5' 端に前記第 1 タグを、前記 3' 端に第 2 タグを有する 2 つタグ付けされた断片の代表ライプラリを生成するステップと；(d) 前記第 1 配列決定タグドメインに対応する部分を含む第 1 配列決定プライマーを提供するステップと；(e) 前記第 1 配列決定プライマーを伸長し、並行して、2 つタグ付けされた断片の前記代表ライプラリの第 1 配列決定タグドメインに隣接するヌクレオチドの同一性を検出するステップと、を含むことが可能である。  
40  
50

## 【0019】

本明細書では、in vitro転移反応を実行するためのキットも提示する。一部の実施形態において、該キットはトランスポソーム複合体を含むことが可能であり、該トランスポソーム複合体は、(1)上記実施形態の何れか一つに記載の変異Tn5トランスポザーゼと；(2)トランスポゾン端配列を有する3'部分を備えるポリヌクレオチドとを含む。

## 【0020】

1つまたは複数の実施形態の詳細を、添付の図面および以下の記載において説明する。他の特徴、目的、および利点は、該記載および図面、ならびに請求項により明らかになるだろう。

10

## 【図面の簡単な説明】

## 【0021】

【図1A】Tn5トランスポザーゼ(1M U H)、Hermesトランスポザーゼ(2B W 3)、HIVインテグラーーゼ(1 I T G)、Muトランスポザーゼ(1 B C M)、およびMos1トランスポザーゼ(3 H O S)の触媒コアドメインの、構造アライメントを示す図である。図示されたナンバリングは、Tn5トランスポザーゼのアミノ酸残基のナンバリングを表す。

【図1B】Tn5トランスポザーゼ(1M U H、ピンク色)、Hermesトランスポザーゼ(2B W 3、黒色)、HIVインテグラーーゼ(1 I T G、黄褐色)、Muトランスポザーゼ(1 B C M)、およびMos1トランスポザーゼ(3 H O S、黄色)の構造アライメント触媒コアドメインを示す図である。Tn5トランスポザーゼW125位置を、スティック表現(stick representation)で示す。

20

【図2】D248Y変異Tn5トランスポザーゼの改変配列挿入バイアスを、Tn5対照と比較して示すIVCプロットである。

【図3】D119L変異Tn5トランスポザーゼの改変配列挿入バイアスを、Tn5対照と比較して示すIVCプロットである。

【図4】W125M変異Tn5トランスポザーゼの改変配列挿入バイアスを、Tn5対照と比較して示すIVCプロットである。

【図5】i a 2 4 8 V挿入変異Tn5トランスポザーゼの改変配列挿入バイアスを、Tn5対照と比較して示すIVCプロットである。

30

【図6】K120Y Tn5トランスポザーゼ挿入変異体、K120FTn5トランスポザーゼ挿入変異体、およびK120WTn5トランスポザーゼ挿入変異体の改変配列挿入バイアスを、Tn5対照と比較して示すIVCプロットである。

【図7】3つの変異Tn5トランスポザーゼの改変配列挿入バイアスを、Tn5対照と比較して示すIVCプロットである。

【図8】8Aは、3つの変異Tn5トランスポザーゼにより生成したセレウス菌ライプラリにおけるA T / G Cドロップアウトを、Tn5対照と比較して示すグラフである。8Bは、3つの変異Tn5トランスポザーゼにより生成したセレウス菌ライプラリの推定ライプラリサイズを、Tn5対照と比較して示すグラフである。

【図9】9Aは、2つの変異Tn5トランスポザーゼにより生成したライプラリにおけるRapid Capture Enrichment実験でのカバレッジ均一性を、Tn5対照と比較して示すグラフである。9Bは、2つの変異Tn5トランスポザーゼにより生成したライプラリにおけるRapid Capture Enrichment実験での、10X標的カバレッジ、20X標的カバレッジ、および平均標的カバレッジを、Tn5対照と比較して示すグラフである。

40

【図10】10Aは、2つの変異Tn5トランスポザーゼにより生成したライプラリにおけるRapid Capture Enrichment実験での固有リードのフィルタ通過割合およびハイブリッド選択ライプラリサイズを、Tn5対照と比較して示すグラフである。10Bは、2つの変異Tn5トランスポザーゼにより生成したライプラリにおけるRapid Capture Enrichment実験での、10X、20X、および3

50

0 X のカバレッジに及ぶペナルティスコアを、T n 5 対照と比較して示すグラフである。【図11】異なるタグメント化バッファを用いて調製した、T S - T n 5 0 5 9 およびT S - T n 5 のタグメント化DNAライプラリにおける固有分子数の棒グラフを示す図である。

【図12】異なるタグメント化バッファを用いて調製した、T S - T n 5 0 5 9 およびT S - T n 5 のタグメント化DNAライプラリにおけるGCドロップアウト率の棒グラフを示す図である。

【図13】異なるタグメント化バッファを用いて調製した、T S - T n 5 0 5 9 およびT S - T n 5 のタグメント化DNAライプラリにおけるATドロップアウト率の棒グラフを示す図である。

【図14】標準バッファ(TD)配合およびコバルトバッファ(Co)配合を用いて調製したT S - T n 5 0 5 9 ライプラリにおける、断片サイズ分布についてのバイオアナライザトレースのプロットを示す図である。

【図15】コバルト-DMSO(Co-DMSO)バッファ配合、NF2バッファ配合、およびH MWバッファ配合を用いて調製したT S - T n 5 0 5 9 ライプラリにおける、断片サイズ分布についてのバイオアナライザトレースのプロットを示す図である。

【図16】標準バッファ配合(TD)およびコバルトバッファ(Co)を用いて調製したT S - T n 5 ライプラリにおける、断片サイズ分布についてのバイオアナライザトレースのプロットを示す図である。

【図17】コバルト-DMSO(Co-DMSO)バッファ配合、NF2バッファ配合、およびH MWバッファ配合を用いて調製したT S - T n 5 ライプラリにおける、断片サイズ分布についてのバイオアナライザトレースのプロットを示す図である。

【図18】18Aは、T S - T n 5 ライプラリにおける配列含有量のバイアスグラフを示す図である。18Bは、T S - T N 5 - Co ライプラリにおける配列含有量のバイアスグラフを示す図である。18Cは、T S - T n 5 - Co - DMSO ライプラリにおける配列含有量のバイアスグラフを示す図である。18Dは、T S - T n 5 - NF2 ライプラリにおける配列含有量のバイアスグラフを示す図である。

【図19】19Aは、T S - T n 5 0 5 9 ライプラリにおける配列含有量のバイアスグラフを示す図である。19Bは、T S - T N 5 0 5 9 - Co ライプラリにおける配列含有量のバイアスグラフを示す図である。19Cは、T S - T n 5 0 5 9 - Co - DMSO ライプラリにおける配列含有量のバイアスグラフを示す図である。19Dは、T S - T n 5 0 5 9 - NF2 ライプラリにおける配列含有量のバイアスグラフを示す図である。

【図20】異なるタグメント化バッファを用いて調製したMBP-MoS1タグメント化ライプラリにおける、平均リード総数および平均多様性の棒グラフを示す図である。

【図21】MBP-MoS1タグメント化ライプラリにおけるGCおよびATのドロップアウトの棒グラフを示す図である。

【図22】22Aは、Mo s 1 - HEPESライプラリにおける配列含有量のバイアスグラフを示す図である。22Bは、Mo s 1 - HEPES - DMSOライプラリにおける配列含有量のバイアスグラフを示す図である。22Cは、Mo s 1 - HEPES - DMSO - Coライプラリにおける配列含有量のバイアスグラフを示す図である。22Dは、Mo s 1 - HEPES - DMSO - Mnライプラリにおける配列含有量のバイアスグラフを示す図である。

【図23】エクソームの配列決定のためにゲノムDNAライプラリを調製し濃縮する方法の一例のフロー図である。

【図24】24Aは、T S - T n 5 0 5 9 トランスポソームを用いて調製したタグメント化セレウス菌ゲノムDNAライプラリにおける、カバレッジのプロットを示す図である。24Bは、Nextera V2トランスポソームを用いて調製したタグメント化セレウス菌ゲノムDNAライプラリにおける、カバレッジのプロットを示す図である。

【図25】25Aは、T S - T n 5 0 5 9 トランスポソームを用いて調製したタグメント化セレウス菌ゲノムDNAライプラリにおける、ギャップ位置とギャップ長のプロットを

10

20

30

40

50

示す図である。25Bは、Nextera V2トランスポソームを用いて調製したタグメント化セレウス菌ゲノムタグメント化DNAライプラリにおける、ギャップ位置とギャップ長のプロットを示す図である。

【図26】40nM(1X正規化濃度)のTS-Tn5059に正規化した、TDE1(Tn5バージョン-1)およびTS-Tn5を用いて調製したタグメント化ゲノムDNAライプラリにおける、25ngヒトgDNAに対する、断片サイズ分布についてのバイオアナライザトレースパネルを示す図である。

【図27】25ngヒトgDNAを用いる、40nM(1X正規化濃度)のTS-Tn5059に正規化したTDE1(Tn5バージョン-1)およびTS-Tn5を用いて調製したタグメント化ゲノムDNAライプラリにおける、サイズ分布の分析を示す図である。

【図28】ある範囲のDNAインプットを用いて調製したタグメント化ゲノムDNAライプラリの断片サイズ分布についての、バイオアナライザトレースのパネルを示す図である。

【図29】29Aは、第1ユーザが調製し、CoriellヒトDNAを用いたTS-Tn5059タグメント化ライプラリにおける、断片サイズ分布についてのバイオアナライザトレースのプロットを示す図である。29Bは、第2ユーザが調製し、CoriellヒトDNAを用いたTS-Tn5059タグメント化ライプラリにおける、断片サイズ分布についてのバイオアナライザトレースのプロットを示す図である。

【図30】25ng～100ngのgDNAを用い、6X「正規化」濃度のTDE1により調製した、Tn5バージョン1(TDE1)タグメント化ライプラリにおける断片サイズ分布についてのバイオアナライザトレースのプロットを示す図である。

【図31】25ng～100ngのgDNAを用い、6X「正規化」濃度のTS-Tn5により調製した、TS-Tn5タグメント化ライプラリにおける断片サイズ分布についてのバイオアナライザトレースのプロットを示す図である。

【図32】10ng～100ngのgDNAを用い、6X「正規化」濃度のTS-Tn5059により調製した、TS-Tn5059タグメント化ライプラリにおける断片サイズ分布についてのバイオアナライザトレースのプロットを示す図である。

【図33】より広い範囲のgDNA(5ng～500ng)を用い、6X「正規化」濃度のTS-Tn5059により調製した、TS-Tn5059タグメント化ライプラリにおける断片サイズ分布についてのバイオアナライザトレースのプロットを示す図である。

#### 【発明を実施するための形態】

##### 【0022】

DNA配列決定のための一部のサンプル調製方法では、各テンプレートは挿入物の何れかの端にアダプタを含み、そして、DNAまたはRNAを修飾し、所望の修飾反応生成物を精製するために、多くのステップをしばしば必要とする。これらのステップは、典型的には、適応断片をフローセルに付加する前に溶液中で実行し、ここで該適応断片は、プライマーの伸長反応により表面に結合し、これにより、ハイブリダイズされた断片は、表面に共有結合的に付着したプライマーの端にコピーされる。これらの「播種」テンプレートは、その後、数回の増幅サイクルを通し、コピー済みテンプレートのモノクローナルクラスタを発生させる。しかしながら、内容の全体が本明細書に組み込まれる、米国特許出願公開第2010/0120098号明細書に開示されるように、クラスタ形成および配列決定のために用意した溶液中のアダプタ修正テンプレートに、DNAを形質転換するに必要なステップの数は、トランスポザーゼが仲介する断片化およびタグ付け(本明細書ではタグメント化という)を用いることにより最小化することが可能である。例えば、タグメント化は、例えばNextera(登録商標)DNAサンプル調製キット( Illumina社)のワークフロー等で例示されるように、DNAの断片化のために用いることが可能であり、ここではゲノムDNAを、インプットDNAを同時に断片化およびタグ付けする設計トランスポソームにより断片化することにより、断片の端に固有アダプタ配列を含む断片化核酸分子の集団を生成することが可能である。しかしながら、改善された挿入バイアスを呈すトランスポザーゼ酵素が必要である。

10

20

30

40

50

## 【0023】

従って、本明細書では、核酸サンプルの改善された断片化およびタグ付けのためのトランスポザーゼを提示する。本発明者らは驚くべきことに、改善された挿入配列バイアスを呈し、多くの他の関連する利点を有する、ある改変トランスポザーゼを特定した。本明細書に提示する改変トランスポザーゼの一実施形態は、改善された挿入バイアスを呈すトランスポザーゼである。

## 【0024】

本明細書で用いる場合、用語「正規化トランスポソームアクティビティ」は、25 ng g DNAインプット収量において、曲線下の全面積についてバイオアナライザ断片サイズ分布をもたらすトランスポソームの最小濃度を意味する：50 μL反応において、100 ~ 300 bp = 20% ~ 30%；301 ~ 600 bp = 30% ~ 40%；601 ~ 7,000 bp = 30 ~ 40%；100 ~ 7,000 bp = 90%である。この最小濃度を1Xという。

10

## 【0025】

本出願全体で用いる場合、トランスポソーム濃度を正規化アクティビティと区別せずに用いる。加えて、本出願全体で用いる場合、トランスポソーム濃度をトランスポザーゼ濃度と区別せずに用いる。

## 【0026】

本明細書で用いる場合、用語「挿入バイアス」は、挿入部位に対するトランスポザーゼの配列優先性を意味する。例えば、ポリヌクレオチドサンプルにおいてA / T / C / Gのバックグラウンド頻度が等しく分布している場合(25%A、25%T、25%C、25%G)、トランスポザーゼの結合部位または切断部位で他の3つより1つのヌクレオチドが過剰出現することは、該部位における挿入バイアスを映す。挿入バイアスは、当技術分野で既知である多くの方法のうち何れか一つを用いて測定することが可能である。例えば、下記の実施例1で概ね説明するように、挿入部位を配列決定し、挿入部位の各位置における任意の特定のヌクレオチドの相対存在量を比較することが可能である。

20

## 【0027】

「挿入バイアスにおける改善」は、改変トランスポザーゼの結合部位の1つまたは複数の位置における特定の塩基の頻度が低減または増加し、ポリヌクレオチドサンプルの該塩基のバックグラウンド頻度により近づくことを示す。改善とは、非改変トランスポザーゼの該位置における頻度と比較し、該位置における頻度の増加とすることが可能である。あるいは、改善とは、該位置における頻度の減少とすることが可能である。したがって、例えば、ポリヌクレオチドサンプルにおけるTヌクレオチドのバックグラウンド頻度が0.25であり、改変トランスポザーゼがトランスポザーゼ結合部位の特定位置において、Tヌクレオチド頻度を0.25超の頻度から0.25により近い頻度に低減する場合、改変トランスポザーゼは挿入バイアスが改変されている。同様に、例えば、ポリヌクレオチドサンプルのTヌクレオチドのバックグラウンド頻度が0.25であり、改変トランスポザーゼがトランスポザーゼ結合部位の特定位置において、Tヌクレオチド頻度を0.25未満の頻度から0.25により近い頻度に増加させる場合、改変トランスポザーゼは挿入バイアスが改変されている。

30

## 【0028】

挿入バイアスを測定する一方法論は、挿入部位の大規模配列決定により、結合部位の各位置における塩基頻度を挿入部位と比較して測定することである。これは例えばGreen et al. Mobile DNA (2012) 3:3に記載されており、これはその全体が参照により本明細書に組み込まれる。各位置における存在量を表示する典型的なツールは、例えば図2に示すような強度対サイクルの分布プロットである。下記の実施例1に記載するように、トランスポゾンが仲介するタグ化および断片化により生成される断片端は、大規模で配列決定することが可能であり、挿入部位の各位置での塩基の分布頻度を測定して該挿入部位の1つまたは複数の位置のバイアスを検出することが可能である。したがって、例えば図2に示すように、位置(1)において、「G」ヌクレオチドについては0.55、「A」ヌクレオ

40

50

チドについては0.16という頻度の塩基分布は、該位置におけるGに対する著しい優先性と、Aを避けるバイアスを映す。別の例としては、対照的に、図3に示すように、位置(20)での塩基分布は、4塩基のそれぞれについて原則的に0.25であり、これは該位置において配列バイアスがほとんど、または全くないことを映す。

【0029】

本明細書に提示する一部の実施形態では、改変トランスポザーゼ酵素は、挿入部位から1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、または21塩基以上上流または下流に位置する1つまたは複数の部位で、挿入バイアスの縮小をもたらす。一部の実施形態では、改変トランスポザーゼ酵素は、挿入部位から1~15塩基下流に位置する1つまたは複数の部位で、挿入バイアスの縮小をもたらす。一部の実施形態では、改変トランスポザーゼ酵素は、挿入部位から1~15塩基上流に位置する1つまたは複数の部位で、挿入バイアスの縮小をもたらす。

【0030】

本明細書の下記でより詳細に記載するように、発明者らは驚くべきことに、トランスポザーゼのアミノ酸配列のある位置における残基への1つまたは複数の変異が、転移事象中に改善された配列挿入バイアスをもたらすことを発見した。これらの改変トランスポザーゼは、高多様性および低多様性の核酸サンプルのタグメント化において改善された性能を提供し、これは、配列決定される様々な領域において、より優れたカバレッジ均一性とドロップアウトの減少をもたらす。

【0031】

本明細書で用いる場合、用語「DNAインプット許容差」は、ある範囲のインプットDNA量にわたり均一なDNA断片サイズを生成する、トランスポザーゼの能力を意味する。

【0032】

本明細書で用いる場合、伸長因子：TSについての表記は、Ts fと区別せずに用いる。

【0033】

一部の実施形態では、インプットDNAはゲノムDNAである。一部の実施形態では、インプットDNAの範囲は、0.001 μg~1mg、1ng~1mg、1ng~900ng、1ng~500ng、1ng~300ng、1ng~250ng、1ng~100ng、5ng~250ng、または5ng~100ngとすることが可能であり、トランスポザーゼの濃度は、5nM~500nMである。一部の実施形態では、上記範囲のインプットDNAに対するトランスポザーゼの濃度は、およそ、25nM、30nM、35nM、40nM、50nM、60nM、65nM、70nM、75nM、80nM、90nM、95nM、100nM、125nM、130nM、140nM、150nM、175nM、180nM、190nM、200nM、210nM、225nM、230nM、240nM、250nM、260nM、275nM、280nM、290nM、300nM、325nM、350nM、360nM、375nM、380nM、390nM、400nM、425nM、450nM、475nM、または500nMである。一部の実施形態では、上記範囲のインプットDNAに対するトランスポザーゼまたは正規化トランスポーソームの、正規化濃度の濃度は、およそ、0.1X~10X、1X~10X、3X~8X、4X~7Xの範囲から選択する。一部の実施形態では、上記範囲のインプットDNAに対するトランスポザーゼまたは正規化トランスポーソームの正規化濃度は、およそ、0.1X、0.2X、0.3X、0.4X、0.5X、0.6X、0.7X、0.8X、0.9X、1X、1.1X、1.2X、1.3X、1.4X、1.5X、1.6X、1.7X、1.8X、1.9X、2X、2.1X、2.2X、2.3X、2.4X、2.5X、2.6X、2.7X、2.8X、2.9X、3X、3.1X、3.2X、3.3X、3.4X、3.5X、3.6X、3.7X、3.8X、3.9X、4X、4.1X、4.2X、4.3X、4.4X、4.5X、4.6X、4.7X、4.8X、4.9X、5X、5.1X、5.2X、5.3X、5.4X、5.5X、5.6X、5.7X、5.8X、5.9X

10

20

30

40

50

、6X、6.1X、6.2X、6.3X、6.4X、6.5X、6.6X、6.7X、6.8X、6.9X、7X、または7.5X、8X、8.5X、9X、9.5X、10Xである。

【0034】

一部の実施形態では、インプットDNAの量は、1ng、2ng、3ng、4ng、5ng、6ng、7ng、8ng、9ng、10ng、11ng、12ng、13ng、15ng、20ng、25ng、30ng、35ng、40ng、45ng、50ng、55ng、60ng、65ng、70ng、75ng、80ng、85ng、90ng、95ng、100ng、110ng、115ng、120ng、125ng、130ng、135ng、140ng、150ng、155ng、160ng、165ng、170ng、180ng、185ng、190ng、195ng、200ng、210ng、220ng、225ng、230ng、235ng、240ng、245ng、250ng、260ng、270ng、280ng、290ng、300ng、325ng、350ng、375ng、400ng、425ng、450ng、475ng、500ng、525ng、550ng、600ng、650ng、700ng、750ng、800ng、850ng、または900ngである。一部の実施形態では、上記量のインプットDNAに対するトランポザーゼの濃度は、およそ、25nM、30nM、35nM、40nM、50nM、60nM、65nM、70nM、75nM、80nM、90nM、95nM、100nM、125nM、130nM、140nM、150nM、175nM、180nM、190nM、200nM、210nM、225nM、230nM、240nM、250nM、260nM、275nM、280nM、290nM、300nM、325nM、350nM、360nM、375nM、380nM、390nM、400nM、425nM、450nM、475nM、または500nMである。一部の実施形態では、上記量のインプットDNAに対するトランポザーゼまたは正規化トランスポソームの、正規化濃度の濃度は、およそ、0.1X～10X、1X～10X、3X～8X、4X～7Xの範囲から選択する。一部の実施形態では、上記量のインプットDNAに対するトランポザーゼまたは正規化トランスポソームの正規化濃度は、およそ、0.1X、0.2X、0.3X、0.4X、0.5X、0.6X、0.7X、0.8X、0.9X、1X、1.1X、1.2X、1.3X、1.4X、1.5X、1.6X、1.7X、1.8X、1.9X、2X、2.1X、2.2X、2.3X、2.4X、2.5X、2.6X、2.7X、2.8X、2.9X、3X、3.1X、3.2X、3.3X、3.4X、3.5X、3.6X、3.7X、3.8X、3.9X、4X、4.1X、4.2X、4.3X、4.4X、4.5X、4.6X、4.7X、4.8X、4.9X、5X、5.1X、5.2X、5.3X、5.4X、5.5X、5.6X、5.7X、5.8X、5.9X、6X、6.1X、6.2X、6.3X、6.4X、6.5X、6.6X、6.7X、6.8X、6.9X、7X、または7.5X、8X、8.5X、9X、9.5X、10Xである。

【0035】

一部の実施形態では、ng量のインプットDNAに対するトランスポザーゼのnM濃度比率はおよそ、0.5～5、1～5、2～5、2.1～3、または2.1～2.5である。

【0036】

本明細書で用いる場合、用語「ゲノムDNA」は、様々な細胞タンパク質をコードする1つまたは複数の遺伝子を含む細胞に存在する核酸を意味する。一部の実施形態では、ゲノムDNAは原核生物、例えば細菌および古細菌に由来する。一部の実施形態では、ゲノムDNAは真核生物、例えば、ヒト、植物、菌類、アメーバに由来する。

【0037】

本明細書で用いる場合の用語「変異」または「修正」は、野生型の遺伝子または遺伝子産物と比較した場合に配列および/または機能特性において修正(つまり、改変特性)を示す、遺伝子または遺伝子産物を意味する。「変異」または「修正」はまた、野生型の遺伝子または遺伝子産物と比較した場合に配列および/または機能特性における修正(つまり、改変特性)を示す、遺伝子または遺伝子産物を意味する。

10

20

30

40

50

り、改変特性)を示す、特異的なヌクレオチド位置での配列、特定のコドン位置での配列、または特定のアミノ酸位置での配列を意味する。

【0038】

本明細書で用いる場合の「含んで (including)」は、用語「含んで (comprising)」と同じ意味を持つ。

【0039】

本明細書で用いる場合の「約 (およそ)」は、数量的にプラス 10 % またはマイナス 10 % を意味する。

【0040】

下記でさらに詳細に記載するように、本発明者らは驚くべきことに、トランスポザーゼのアミノ酸配列のある位置での残基への 1 つ又は複数の変異は、DNA インプット許容差の拡大をもたらし、結果として、変異トランスポザーゼは野生型トランスポザーゼと比較し、ある範囲のインプット DNA 量にわたって均一な DNA フラグメントサイズを生成することを発見した。一実施形態では、TS - Tn 5059 トランスポザーゼは、他のトランスポザーゼ、例えば TS - Tn 5 および Tn 5 バージョン 1 (TDE1) と比較し、拡大した DNA インプット許容差を呈す。

10

【0041】

一部の実施形態では、TS - Tn 5059 は他のトランスポザーゼと比較し、DNA インプット許容差の拡大を呈すが、該実施形態ではインプット DNA の範囲は 1 ng ~ 20 0 ng のゲノム DNA であり、TS - Tn 5059 の濃度は 100 ~ 300 nM である。一部の実施形態では、TS - Tn 5059 は他のトランスポザーゼと比較し DNA インプット許容差の拡大を呈すが、該実施形態ではインプット DNA の範囲は 5 ng ~ 200 ng のゲノム DNA であり、TS - Tn 5059 の濃度は 100 ~ 250 nM である。一部の実施形態では、TS - Tn 5059 は他のトランスポザーゼと比較し DNA インプット許容差の拡大を呈すが、該実施形態では、インプット DNA の範囲は 5 ng ~ 100 ng のゲノム DNA であり、TS - Tn 5059 の濃度は 240 nM ~ 250 nM である。

20

【0042】

拡大した DNA インプット許容差を有する改善された挿入バイアスは、現在の Next era (登録商標) Rapid Capture protocol (Illumina 社) より、より速くより柔軟なサンプル調製およびエクソーム濃縮プロトコルを提供する。

30

【0043】

本明細書で用いる場合、用語「タグメント化」はトランスポソーム複合体による DNA の修飾を意味し、該トランスポソーム複合体は、トランスポゾン端配列を備えるアダプタと複合体化した、トランスポザーゼ酵素を含む。タグメント化は、DNA の同時断片化と、二重断片の両鎖の 5' 端へのアダプタのライゲーションとをもたらす。

【0044】

本明細書で用いる場合、「トランスポソーム複合体」または「トランスポソーム」は、少なくとも一つのトランスポザーゼ酵素とトランスポザーゼ認識部位からなる。一部のこのようない系において、トランスポザーゼは、トランスポゾン認識部位と機能的複合体を形成することが可能であり、これは転移反応に触媒作用を及ぼすことが可能である。トランスポザーゼは、トランスポザーゼ認識部位に結合し、本明細書でタグメント化と呼ぶプロセスにおいて、該トランスポザーゼ認識部位を標的核酸に挿入することができる。このような挿入事象の一部では、トランスポザーゼ認識部位の一方の鎖を標的核酸に移動させることができる。

40

【0045】

本明細書で示す改変トランスポザーゼ酵素は、トランスポソーム複合体の一部を形成することが可能である。例示的転移複合体には過活動 Tn 5 トランスポザーゼおよび Tn 5 型トランスポザーゼ認識部位が含まれるが、これらに限定されない。過活動 Tn 5 トランスポザーゼには、米国特許第 5 925545 号明細書、同第 5 965443 号明細書、同

50

第7083980号明細書、および同第7608434号明細書、ならびにGoryshin and Reznikoff, J. Biol. Chem., 273:7367 (1998)の開示内容に記載のものも含まれ、これらの各内容はその全体が参照により本明細書に組み込まれる。しかしながら、本明細書で示す改変トランスポザーゼ酵素は、意図した目的のために十分効率的に、ランダムまたはほぼランダムなやり方で、トランスポゾン端をタグ標的核酸に挿入することができる任意の転移系で利用可能であり、そして、提示した方法で用いることが可能であることが理解されよう。

#### 【0046】

例えば、提示する改変トランスポザーゼは、Tn5アミノ酸配列の部位と機能的に同等の位置で少なくとも一つのアミノ酸置換変異を含むことが可能である。図1に例示するように、Tn5に対して相同な領域を本明細書で説明し、他のトランスポザーゼ酵素、例えば、Hermesトランスポザーゼ、HIVインテグラーゼ、Muトランスポザーゼ、およびMos1トランスポザーゼにおける機能的に同等な部位の特定を可能にする。同様に、他のトランスポザーゼ酵素またはインテグラーゼ酵素における機能的に同等の部位は、例えば、Tn5アミノ酸配列の配列アライメントを実行し、保存または半保存された残基またはドメインを特定することにより、当業者に即座に明らかになろう。したがって、本明細書で提供するある実施形態で用いることが可能な転移系には、Tn5の部位と機能的に同等な部位を有する任意の既知のトランスポザーゼが含まれることが理解されよう。例えば、このような系は、MuAトランスポザーゼ、ならびに、R1およびR2の端配列を備えるMuトランスポザーゼ認識部位が含まれ得る (Mizuuchi, K., Cell, 35: 785, 1983; Savilahti, H, et al., EMBO J., 14: 4893, 1995)。

#### 【0047】

本明細書で提供するある実施形態に含まれる転移系のより多くの例には、黄色ブドウ球菌Tn552 (Colegio et al., J. Bacteriol., 183: 2384-8, 2001; Kirby C et al., Mol. Microbiol., 43: 173-86, 2002)、Ty1 (Devine & Boeke, Nucleic Acids Res., 22: 3765-72, 1994および国際公開W95/23875)、トランスポゾンTn7 (Craig, N L, Science, 271: 1512, 1996; Craig, N L, Review in: Curr Top Microbiol Immunol., 204:27-48, 1996)、Tn/OおよびIS10 (Kleckner N, et al., Curr Top Microbiol Immunol., 204:49-82, 1996)、Marinerトランスポザーゼ ((Lampe D J, et al., EMBO J., 15: 5470-9, 1996)、Tc1 (Plasterk R H, Curr. Topics Microbiol. Immunol., 204: 125-43, 1996)、P因子 (Gloor, G B, Methods Mol. Biol., 260: 97-114, 2004), Tn3 (Ichikawa & Ohtsubo, J Biol. Chem. 265:18829-32, 1990)、細菌挿入配列 (Ohtsubo & Sekine, Curr. Top. Microbiol. Immunol. 204: 1-26, 1996)、レトロウイルス (Brown, et al., Proc Natl Acad Sci USA, 86:2525-9, 1989)、ならびにイーストのレトロトランスポゾン (Boeke & Corces, Annu Rev Microbiol. 43:403-34, 1992) が含まれる。上記で引用した参考文献は、その全体が参照により本明細書に組み込まれる。

#### 【0048】

簡潔に言えば、「転移反応」は、1つ又は複数のトランスポゾンを標的核酸のランダムな部位またはほぼランダムな部位に挿入する反応である。転移反応において必須の構成要素は、トランスポザーゼ、および、トランスポゾンのヌクレオチド配列を呈すDNAオリゴヌクレオチドであり、移動するトランスポゾン配列およびその相補体 (つまり、非移動トランスポゾン端配列)、ならびに、機能転移またはトランスポソーム複合体を形成するのに必要な他の構成要素が含まれる。DNAオリゴヌクレオチドはさらに、必要に応じてまたは所望により、追加配列 (例えば、アダプタまたはプライマー配列) を含むことが可能である。

#### 【0049】

核酸の5'および/または3'端に付加するアダプタは、ユニバーサル配列を含むことが可能である。ユニバーサル配列は、2つ以上の核酸分子に共通する、つまり共有されるヌクレオチド配列領域である。任意で、2つ以上の核酸分子はまた、配列の異なる領域を

10

20

30

40

50

有する。したがって、例えば、5'アダプタは同一またはユニバーサルな核酸配列を含むことが可能であり、3'アダプタは同一またはユニバーサルな配列を含むことが可能である。複数の核酸分子の異なるメンバーに存在し得るユニバーサル配列は、ユニバーサル配列に対して相補的な単一ユニバーサルプライマーを用いて多数の異なる配列の複製または増幅を可能にする。

【0050】

トランスポザーゼ変異

したがって、本明細書では、野生型ヒトランスポザーゼと比較し修正した変異トランスポザーゼを提示する。改変トランスポザーゼは、下記の表1に記載の残基に機能的に同等な位置において、少なくとも1つのアミノ酸置換変異を含むことが可能である。表1は、トランスポザーゼ残基における置換変異を記載し、これは改善された挿入バイアスをもたらすことが示されている。表1に記載するように、本明細書で示す置換変異は、本明細書で配列番号1として例示する野生型Tn5トランスポザーゼ、または、他の部位に異なる変異を有するトランスポザーゼといった、任意の機能的トランスポザーゼバックボーンに存在し得、これには、過活動Tn5トランスポザーゼとして知られ、本明細書では配列番号11として例示するトランスポザーゼ配列にあるもの、例えば、米国特許第5925545号明細書、同第5965443号明細書、同第7083980号明細書、および同第7608434号明細書に組み込まれた事項に記載される1つまたは複数の変異が含まれる。

【0051】

10

20

【表 1 - 1】

表 1 - 改善された挿入バイアスをもたらす変異の例

バックボーン	Tn5 変異	配列番号	
Tn5 WT	D248Y D248T D248K D248S D248L D248A D248W D248P D248G D248R D248F D248H	2	10
Tn5 WT	D119L D119M D119S D119A D119V	3	20
Tn5 WT	W125M	4	
Tn5 WT	iaD248	5	
Tn5 WT	K120F K120Y K120E K120W	6	
Tn5 WT	Y、T、K、S、L、A、W、P、G、R、 F、またはHに対しD248 L、M、S、A、またはVに対しD119 W125M K120F	7	30
Tn5 WT	K212R P214R A338V	8	
Tn5 WT	K212R P214R G251R A338V	9	40

【表 1 - 2】

Tn5 WT	E146Q E190G G251R	10
Tn5 過活動	D248Y D248T D248K D248S D248L D248A D248W D248P D248G D248R D248F D248H	10
Tn5 過活動	D119L D119M D119S D119A D119V	13
Tn5 過活動	W125M	14
Tn5 過活動	iaD248	15
Tn5 過活動	K120F K120Y K120E K120W	16
Tn5 過活動	Y、T、K、S、L、A、W、P、G、R、 F、またはHに対しD248 L、M、S、A、またはVに対しD119 W125M K120F	17
Tn5 過活動	K212R P214R A338V	18
Tn5 過活動	K212R P214R G251R A338V	19
Tn5 過活動	E146Q E190G G251R	20

【0052】

当技術分野で理解されるように、上記表に挙げた参照番号は、野生型 Tn5 配列（配列番号 1）のアミノ酸位置を意味する。当業者は、ナンバリングは N 末端の切断、挿入、または融合により変わり得ることを理解しよう。上に挙げたアミノ酸の機能的位置は、位置のナンバリングが変わったとしても同じままだろう。例えば、配列番号 25 に記載の配列

のうち最初の285個のアミノ酸残基は大腸菌TSのN末端融合を含み、その後ろに配列番号11のアミノ酸残基2~476が続く。したがって、例えば配列番号25のPro656は、配列番号11のPro372と機能的に対応する。

【0053】

したがって、ある実施形態では、本明細書で提示する改変トランスポザーゼは、野生型トランスポザーゼと比較し、例えば、Tn5トランスポザーゼのアミノ酸配列のAsp248、Asp119、Trp125、Lys120、Lys212、Pro214、Gly251、Ala338、Glu146、および/またはGlu190と機能的に同等の位置で少なくとも1つのアミノ酸置換変異を含む。

【0054】

一部の実施形態では、変異トランスポザーゼは、位置Asp248で変異を含む。位置Asp248での変異は、例えば、置換変異または挿入変異とすることが可能である。ある実施形態では、変異はAsp以外の任意の残基への置換変異である。ある実施形態では、位置Asp248での置換変異は、Tyr、Thr、Lys、Ser、Leu、Ala、Trp、Pro、Gln、Arg、Phe、およびHisからなる群から選択される残基への変異を含む。

10

【0055】

ある実施形態では、位置Asp248での変異は、位置Asp248の後ろにおける挿入変異である。ある態様では、挿入変異には、Asp248の後ろへの任意の残基の挿入が含まれ得る。ある態様では、挿入変異には、位置Asp248の後ろへの疎水性残基の挿入が含まれ得る。疎水性残基は当業者には既知であり、該疎水性残基には、例えば、Val、Leu、Ile、Phe、Trp、Met、Ala、Tyr、およびCysが含まれる。ある態様では、挿入変異には、位置Asp248の後ろへのバリン残基の挿入が含まれ得る。

20

【0056】

本明細書で提示する一部の実施形態には、野生型Tn5トランスポザーゼと比較し修正した変異Tn5トランスポザーゼを含み、該変異トランスポザーゼは位置Asp119で変異を含む。ある態様では、位置Asp119における変異は置換変異である。ある態様では、位置Asp119における置換変異には、疎水性残基への変異が含まれ得る。疎水性残基は当業者に既知であり、例えば、Val、Leu、Ile、Phe、Trp、Met、Ala、Tyr、およびCysが含まれる。ある態様では、位置Asp119における置換変異には、親水性残基への変異が含まれ得る。親水性残基は当業者に既知であり、例えば、Arg、Lys、Asn、His、Pro、Asp、およびGluが含まれる。ある態様では、位置Asp119での置換変異には、Leu、Met、Ser、Ala、およびValからなる群から選択される残基への変異が含まれ得る。

30

【0057】

本明細書で提示する一部の実施形態には、野生型Tn5トランスポザーゼと比較し修正した変異Tn5トランスポザーゼを含み、該変異トランスポザーゼは位置Trp125で変異を含む。ある態様では、位置Trp125での変異は置換変異である。ある態様では、位置Trp125での置換変異には、メチオニン残基への変異が含まれ得る。

40

【0058】

本明細書に提示する一部の実施形態には、野生型Tn5トランスポザーゼと比較し修正した変異Tn5トランスポザーゼを含み、該変異トランスポザーゼは、位置Lys120で変異を含む。ある態様では、位置Lys120における変異は置換変異である。ある態様では、位置Lys120での置換変異には、嵩高い芳香族残基への変異が含まれ得る。嵩高い芳香族残基として特徴付けられる残基は当業者にとって既知であり、例えば、Phe、Tyr、およびTrpが含まれる。ある態様では、位置Lys120での置換変異に、Tyr、Phe、Trp、およびGluからなる群から選択される残基への変異が含まれ得る。

【0059】

50

本明細書で提示される一部の実施形態には、野生型 T n 5 トランスポザーゼと比較し修正した変異 T n 5 トランスポザーゼを含み、該変異トランスポザーゼは L y s 2 1 2 および / または P r o 2 1 4 および / または A 1 a 3 3 8 の位置で変異を含む。ある態様では、 L y s 2 1 2 および / または P r o 2 1 4 および / または A 1 a 3 3 8 の位置での変異は、置換変異である。ある態様では、位置 L y s 1 2 0 での置換変異にはアルギニンへの変異が含まれる。ある態様では、位置 P r o 2 1 4 での置換変異にはアルギニンへの変異が含まれる。ある態様では、位置 A 1 a 3 3 8 での置換変異にはバリンへの変異が含まれる。一部の実施形態では、トランスポザーゼは G 1 y 2 5 1 で置換変異をさらに含むことが可能である。ある態様では、位置 G 1 y 2 5 1 での置換変異にはアルギニンへの変異が含まれる。

10

#### 【 0 0 6 0 】

本明細書で提示する一部の実施形態は、野生型 T n 5 トランスポザーゼと比較し修正した変異 T N 5 トランスポザーゼを含み、該変異トランスポザーゼは G 1 u 1 4 6 および / または G 1 u 1 9 0 および / または G 1 y 2 5 1 の位置で変異を含む。ある態様では、 G 1 u 1 4 6 および / または G 1 u 1 9 0 および / または G 1 y 2 5 1 の位置での変異は置換変異である。ある態様では、位置 G 1 u 1 4 6 での置換変異にはグルタミンへの変異が含まれ得る。ある態様では、位置 G 1 u 1 9 0 での置換変異にはグリシンへの変異が含まれ得る。ある態様では、位置 G 1 y 2 5 1 での置換変異にはアルギニンへの変異が含まれ得る。

20

#### 【 0 0 6 1 】

上記実施形態の何れかにおいて、変異 T n 5 トランスポザーゼはさらに、 T n 5 トランスポザーゼのアミノ酸配列の G 1 u 5 4 および / または M e t 5 6 および / または L e u 3 7 2 と機能的に同等の位置で置換変異を含むことが可能である。ある実施形態では、トランスポザーゼは、 T n 5 トランスポザーゼのアミノ酸配列の G 1 u 5 4 L y s および / または M e t 5 6 A 1 a および / または L e u 3 7 2 P r o に対し相同な置換変異を含む。

#### 【 0 0 6 2 】

本明細書に提示する一部の実施形態には、配列番号 2 ~ 1 0 および 1 2 ~ 2 0 の何れか一つのアミノ酸配列を含む変異 T n 5 トランスポザーゼが含まれる。

#### 【 0 0 6 3 】

30

本明細書では、半保存されたドメインへの置換変異を含む改変トランスポザーゼも提示する。本明細書で用いる場合、用語「半保存されたドメイン」は、様々なトランスポザーゼおよび / または様々な種のうち、完全に保存された、または、少なくとも部分的に保存された、トランスポザーゼの一部分を意味する。半保存されたドメインは、トランスポザーゼの触媒コアドメインに存在するアミノ酸残基を含む。驚くべきことに、半保存されたドメインにある 1 つまたは複数の残基の変異はトランスポザーゼアクティビティに影響を及ぼし、挿入バイアスにおいて改善をもたらすということを発見した。

#### 【 0 0 6 4 】

一部の実施形態では、半保存されたドメインは配列番号 2 1 に記載される配列を有するアミノ酸を含む。配列番号 2 1 は T n 5 トランスポザーゼのアミノ酸配列の残基 1 2 4 ~ 1 3 3 に対応し、これは本明細書では配列番号 1 と記載する。半保存されたドメインにおける、様々なトランスポザーゼのうち保存を示す構造アライメントを図 1 に記載する。図 1 に示すトランスポザーゼ配列には、 T n 5 トランスポザーゼ ( 1 M U H ) 、 H e r m e s トランスポザーゼ ( 2 B W 3 ) 、 H I V インテグラーゼ ( 1 I T G ) 、 M u トランスポザーゼ ( 1 B C M ) 、および M o s 1 トランスポザーゼ ( 3 H O S ) の触媒コアドメインが含まれる。

40

#### 【 0 0 6 5 】

半保存されたドメインにおける 1 つまたは複数の残基への変異は、驚くべきことに、挿入バイアスの改善をもたらすことが分かっている。例えば、本明細書で提示する改変トランスポザーゼの一部の実施形態では、置換変異には、配列番号 2 1 の位置 2 で T r p 以外

50

の任意の残基への変異が含まれる。ある実施形態では、改変トランスポザーゼは、配列番号 21 の位置 2 で Met への変異を含む。

【 0 0 6 6 】

「機能的に同等な」とは、異なるトランスポザーゼを全面的に用いる研究の場合に、対照トランスポザーゼが、酵素において同一の機能的役割を有する他のトランスポザーゼのアミノ酸位置で起きると考えられるアミノ酸置換を含むことを意味する。例えば、Muトランスポザーゼの位置 288 におけるリシンからメチオニンへの変異 (K288M) は、Tn5トランスポザーゼの位置 125 におけるトリプトファンからメチオニンへの置換 (W125M) と機能的に同等だろう。

【 0 0 6 7 】

2 つ以上の異なるトランスポザーゼにおいて通常機能的に同等な置換変異は、トランスポザーゼのアミノ酸配列における相同なアミノ酸位置において起きる。つまり、用語「機能的に同等な」を本明細書で用いる場合、変異アミノ酸の特定の機能が既知であるか否かに関わらず、所定の変異に「位置的に同等」または「相同」である変異も含む。2 つ以上の異なるトランスポザーゼのアミノ酸配列において、位置的に同等または相同なアミノ酸残基は、配列アライメントおよび / または分子モデリングを基に特定することが可能である。位置的に同等および / または機能的に同等な残基を特定するための配列アライメントおよび分子モデリングの例を図 1 に記載する。したがって、例えば、図 1 に示すように、半保存されたドメインにおける残基は、Tn5トランスポザーゼのアミノ酸配列の位置 124 ~ 133 として特定される。Hermes トランスポザーゼ、HIV インテグラーゼ、Mu トランスポザーゼ、および Mos1 トランスポザーゼにおける対応残基は、垂直に並べた図において特定される。これらは、Tn5 トランスポザーゼのアミノ酸配列の対応残基と機能的に同等であると同時に、位置的に同等であると考えられる。

10

【 0 0 6 8 】

上記の改変トランスポザーゼは追加の置換変異を含むことが可能であり、これは、トランスポザーゼアクティビティの 1 つまたは複数の態様を高めることが知られている。例えば、一部の実施形態では、上記の変異のうち何れかに加え、改変 Tn5 トランスポザーゼは、Tn5 トランスポザーゼのアミノ酸配列の G1u54 および / または Met56 および / または Leu372 と機能的に同等な位置で置換変異をさらに含むことが可能である。当技術分野で既知のように、そして、米国特許第 5 9 2 5 5 4 5 号明細書、同第 5 9 6 5 4 4 3 号明細書、同第 7 0 8 3 9 8 0 号明細書、および同第 7 6 0 8 4 3 4 号明細書の開示内容、ならびに、Goryshin and Reznikoff, J. Biol. Chem., 273:7367 (1998) の開示内容（これらは各々、その全体が参照により組み込まれる）により例示されるように、Tn5 トランスポザーゼのアミノ酸配列の G1u54 および / または Met56 および / または Leu372 と機能的に同等な位置にある 1 つまたは複数の位置において、改善されたアクティビティをもたらす種々の置換変異のうち何れかを行うことが可能である。実施形態では、トランスポザーゼは、Tn5 トランスポザーゼのアミノ酸配列の G1u54 Lys および / または Met56 Ala および / または Leu372 Pro に相同な置換変異を含む。例えば、置換変異には、Tn5 トランスポザーゼのアミノ酸配列の G1u54 Lys および / または Met56 Ala および / または Leu372 Pro に相同な置換変異を含み得る。

20

【 0 0 6 9 】

トランスポザーゼ変異

様々な型の突然変異誘発を本開示で任意に用いて、例えば、上記のトランスポザーゼモデルおよびモデル予測に従って、または、ランダムもしくは半ランダムな変異アプローチを用いて、例えば、トランスポザーゼを修飾して多様体を生成する。通常、任意の利用可能な突然変異誘発手順を、トランスポザーゼ変異体を作成するために用いることが可能である。このような突然変異誘発手順には、対象の 1 つまたは複数のアクティビティ（例えば、改善された挿入バイアス）のために、変異核酸および変異ポリペプチドを選択することが任意で含まれる。用いることが可能な手順には、部位特異的点突然変異誘発、ランダ

30

40

50

ム点突然変異誘発 ( random point mutagenesis ) 、 *in vitro* または *in vivo* の相同組み換え ( DNA シャッフリングおよびコンビナトリアルオーバーラップ PCR ) 、テンプレートを含むウラシルを用いた突然変異誘発、オリゴヌクレオチド特異的突然変異誘発、ホスホロチオエート修正 DNA 突然変異誘発、ギャップ付き二重 DNA を用いた突然変異誘発、点不一致修復、欠損修復宿主株を用いた突然変異誘発、制約選択および制約精製、欠失突然変異誘発、全遺伝子合成による突然変異誘発、縮重 PCR 、二重鎖切断修復、ならびに、当業者に既知の他の多くのものが含まれるが、これらに限定されない。変異用開始トランスポザーゼは本明細書に記載のもののうち何れかとすることが可能であり、これには、例えば、米国特許第 5 9 2 5 5 4 5 号明細書、同第 5 9 6 5 4 4 3 号明細書、同第 7 0 8 3 9 8 0 号明細書、および同第 7 6 0 8 4 3 4 号明細書、ならびに Gorysh in and Reznikoff, *J. Biol. Chem.*, 273:7367 (1998) の開示内容 ( これらはそれぞれ、参考によりその全体が組み込まれる ) で特定されるものといった、利用可能なトランスポザーゼ変異体が含まれる。

#### 【 0 0 7 0 】

任意で、突然変異誘発は、天然発生トランスポザーゼ分子からの既知の情報、または、既知の変更もしくは変異したトランスポザーゼ ( 例えば、上記参考文献に記載される既存の変異トランスポザーゼを用いる ) についての既知の情報、例えば、上記のような、配列、配列比較、物理特性、および / または結晶構造等により誘導することが可能である。しかしながら、別の種類の実施形態では、修飾は、 ( 例えば、標準的または「ファミリー」の DNA シャッフリングのように ( 例えば、 Crameri et al. (1998) "DNA shuffling of a family of genes from diverse species accelerates directed evolution" *Nature* 39 1:288-291 を参照 ) ) 、基本的にランダムとすることが可能である。

#### 【 0 0 7 1 】

変異フォーマットについての追加情報は以下にある : Sambrook et al., *Molecular Cloning--A Laboratory Manual* (3rd Ed.) , Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 2000 ("Sambrook") ; *Current Protocols in Molecular Biology*, F. M. Ausubel et al., eds., *Current Protocols*, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc., (supplemented through 2011) ("Ausubel")) and *PCR Protocols A Guide to Methods and Applications* (Innis et al. eds) Academic Press Inc. San Diego, Calif. (1990) ("Innis")。以下の刊行物およびその中で引用される参考文献は、変異フォーマットについてさらなる詳細を提供する : Arnold, *Protein engineering for unusual environments*, *Current Opinion in Biotechnology* 4:450-455 (1993) ; Bass et al., *Mutant Trp repressors with new DNA-binding specificities*, *Science* 242:240-245 (1988) ; Bordo and Argos (1991) *Suggestions for "Safe" Residue Substitutions in Site-directed Mutagenesis* 217:721-729; Botstein & Shortle, *Strategies and applications of in vitro mutagenesis*, *Science* 229:1193-1201 (1985) ; Carter et al., *Improved oligonucleotide site-directed mutagenesis using M13 vectors*, *Nucl. Acids Res.* 13: 4431-4443 (1985) ; Carter, *Site-directed mutagenesis*, *Biochem. J.* 237:1-7 (1986) ; Carter, *Improved oligonucleotide-directed mutagenesis using M13 vectors*, *Methods in Enzymol.* 154: 382-403 (1987) ; Dale et al., *Oligonucleotide-directed random mutagenesis using the phosphorothioate method*, *Methods Mol. Biol.* 57:369-374 (1996) ; Eghitedarzadeh & Henikoff, *Use of oligonucleotides to generate large deletions*, *Nucl. Acids Res.* 14: 5 115 (1986) ; Fritz et al., *Oligonucleotide-directed construction of mutations: a gapped duplex DNA procedure without enzymatic reactions in vitro*, *Nucl. Acids Res.* 16: 6987-6999 (1988) ; Grundstrom et al., *Oligonucleotide-directed mutagenesis by microscale 'shot-gun' gene synthesis*, *Nucl. Acids Res.* 13: 3305-3316 (1985) ; Hayes (2002) *Combining Computational and Experimental Screening for rapid Optimization of Protein Properties* *PNAS* 99(25) 15926-15931 ; Kunkel, *The efficiency of oligonucleotide directed mutagenesis*, *in Nucleic Acids & Molecular Biol* 50

gy (Eckstein, F. and Lilley, D. M. J. eds., Springer Verlag, Berlin)) (1987) ; Kunkel, Rapid and efficient site-specific mutagenesis without phenotypic selection, Proc. Natl. Acad. Sci. USA 82:488-492 (1985) ; Kunkel et al., Rapid and efficient site-specific mutagenesis without phenotypic selection, Methods in Enzymol. 154, 367-382 (1987) ; Kramer et al., The gapped duplex DNA approach to oligonucleotide-directed mutation construction, Nucl. Acids Res. 12: 9441-9456 (1984) ; Kramer & Fritz Oligonucleotide-directed construction of mutations via gapped duplex DNA, Methods in Enzymol. 154:350-367 (1987) ; Kramer et al., Point Mismatch Repair, Cell 38:879-887 (1984) ; Kramer et al., Improved enzymatic *in vitro* reactions in the gapped duplex DNA approach to oligonucleotide-directed construction of mutations, Nucl. Acids Res. 16: 7207 (1988) ; Ling et al., Approaches to DNA mutagenesis: an overview, Anal Biochem. 254(2): 157-178 (1997) ; Lorimer and Pastan Nucleic Acids Res. 23, 3067-8 (1995) ; Mandecki, Oligonucleotide-directed double-strand break repair in plasmids of *Escherichia coli*: a method for site-specific mutagenesis, Proc. Natl. Acad. Sci. USA, 83:7177-7181(1986) ; Nakamaye & Eckstein, Inhibition of restriction endonuclease Nci I cleavage by phosphorothioate groups and its application to oligonucleotide-directed mutagenesis, Nucl. Acids Res. 14: 9679-9698 (1986) ; Nambiar et al., Total synthesis and cloning of a gene coding for the ribonuclease S protein, Science 223: 1299-1301(1984) ; Sakamara and Khorana, Total synthesis and expression of a gene for the  $\alpha$ -subunit of bovine rod outer segment guanine nucleotide-binding protein (transducin), Nucl. Acids Res. 14: 6361-6372 (1986) ; Sayers et al., Y-T Exonucleases in phosphorothioate-based oligonucleotide-directed mutagenesis, Nucl. Acids Res. 16:791-802 (1988) ; Sayers et al., Strand specific cleavage of phosphorothioate-containing DNA by reaction with restriction endonucleases in the presence of ethidium bromide, (1988) Nucl. Acids Res. 16: 803-814 ; Sieber, et al., Nature Biotechnology, 19:456-460 (2001) ; Smith, *In vitro* mutagenesis, Ann. Rev. Genet. 19:423-462 (1985) ; Methods in Enzymol. 100: 468-500 (1983) ; Methods in Enzymol. 154: 329-350 (1987) ; Stemmer, Nature 370, 389-91(1994) ; Taylor et al., The use of phosphorothioate-modified DNA in restriction enzyme reactions to prepare nicked DNA, Nucl. Acids Res. 13: 8749-8764 (1985) ; Taylor et al., The rapid generation of oligonucleotide-directed mutations at high frequency using phosphorothioate-modified DNA, Nucl. Acids Res. 13: 8765-8787 (1985) ; Wells et al., Importance of hydrogen-bond formation in stabilizing the transition state of subtilisin, Phil. Trans. R. Soc. Lond. A 317: 415-423 (1986) ; Wells et al., Cassette mutagenesis: an efficient method for generation of multiple mutations at defined sites, Gene 34:315-323 (1985) ; Zoller & Smith, Oligonucleotide-directed mutagenesis using M 13-derived vectors: an efficient and general procedure for the production of point mutations in any DNA fragment, Nucleic Acids Res. 10:6487-6500 (1982) ; Zoller & Smith, Oligonucleotide-directed mutagenesis of DNA fragments cloned into M13 vectors, Methods in Enzymol. 100:468-500 (1983) ; Zoller & Smith, Oligonucleotide-directed mutagenesis: a simple method using two oligonucleotide primers and a single-stranded DNA template, Methods in Enzymol. 154:329-350 (1987) ; Clackson et al. (1991) "Making antibody fragments using phage display libraries" Nature 352:624-628 ; Gibbs et al. (2001) "Degenerate oligonucleotide gene shuffling (DOGS): a method for enhancing the frequency of recombination with family shuffling" Gene 271:13-20 ; および、Hiraga and Arnold (2003) "General method for sequence-independent site-directed chimeragenesis: J. Mol. Biol. 330:287-296。上記方法の多くについてのさらなる詳細はMethods in Enzymology Volume 154に見られ、これまた、様々な突然変異誘発法にまつわるトラブルシューティング問題に対する有用な制御について記載する。

## 【0072】

組み換えトランスポザーゼの作成と単離

概して、本明細書に提示するトランスポザーゼをコードする核酸は、クローニング、組み換え、*in vitro*合成、*in vitro*増幅、および/または他の利用可能な方法により作成することが可能である。種々の組み換え方法を、本明細書に提示するトランスポザーゼをコードする発現ベクターを発現するために用いることが可能である。組み換え核酸を作成する方法、発現産物を発現する方法、および発現産物を単離する方法は、当技術分野では周知であり、記載されている。多くの例示的変異および変異の組み合わせ、ならびに所望の変異を設計するための方法を本明細書では記載する。トランスポザーゼの触媒ドメインにおいて変異を作成し選択する方法は、本明細書にあるほか、米国特許第5925545号明細書、同第5965443号明細書、同第7083980号明細書、および同第7608434号明細書で例示されており、これらはその全体が参照により組み込まれる。

## 【0073】

変異、組み換え、および*in vitro*核酸操作の方法（クローニング、発現、およびPCR等を含む）についての追加の有用な参考文献には、Berger and Kimmel, Guide to Molecular Cloning Techniques, Methods in Enzymology volume 152 Academic Press, Inc., San Diego, Calif. (Berger) ; Kaufman et al. (2003) Handbook of Molecular and Cellular Methods in Biology and Medicine Second Edition Ceske (ed) CRC Press (Kaufman) ; The Nucleic Acid Protocols Handbook Ralph Rapley (ed) (2000) Cold Spring Harbor, Humana Press Inc (Rapley) ; Chen et al. (ed) PCR Cloning Protocols, Second Edition (Methods in Molecular Biology, volume 192) Humana Press ; およびin Viljoen et al. (2005) Molecular Diagnostic PCR Handbook Springer, ISBN 1402034032が含まれる。

## 【0074】

加えて、多量のキットが、細胞からプラスミドまたは他の関連核酸を精製するために商業的に利用可能である（例えば、Pharmacia Biotech社によるEasy Prep（登録商標）およびFlexi Prep（登録商標）、Stratagene社によるStrataclean（登録商標）、ならびに、Qiagen社によるQIA prep（登録商標）を参照）。任意の単離および/または精製した核酸をさらに操作して、細胞をトランスフェクトするのに用いる、および/または、関連ベクターに組み込んで発現のために有機体を感染させるなどする他の核酸を、生成することが可能である。典型的なクローニングベクターは、転写ターミネーターおよび翻訳ターミネーター、転写開始配列および翻訳開始配列、ならびに、特定の標的核酸の発現の調節に有用なプロモーターを含む。ベクターは任意で、少なくとも1つの独立ターミネーター配列を含む汎用発現力セットと、真核生物、原核生物、またはその両方において該力セットの複製を可能にする配列（例えば、シャトルベクター）と、原核生物系および真核生物系の両方のための選択マーカーとを含む。ベクターは、原核生物、真核生物、またはその両方における複製および統合に適している。

## 【0075】

一部の実施形態において、本明細書で提示するトランスポザーゼは融合タンパク質として発現する。融合タンパク質は、例えば、トランスポザーゼの可溶性、発現、および/または精製などの特性を高めることができある。本明細書で用いる場合、用語「融合タンパク質」は、天然の単一ポリペプチドには普通は存在しない少なくとも2つのポリペプチドドメインを有する、単一ポリペプチド鎖を意味する。したがって、天然に発生するタンパク質およびその点変異体は、本明細書で用いる「融合タンパク質」ではない。好適には、対象のポリペプチドはペプチド結合を介して少なくとも1つのポリペプチドドメインと融合し、該融合タンパク質はまた、別々のタンパク質に由来するアミノ酸部分間で、アミノ酸の連結領域を含み得る。対象のポリペプチドに融合したポリペプチドドメインは、対象のポリペプチドの可溶性および/または発現を高めることができ、そして、精製タグを

提供して、組み換え融合タンパク質を宿主細胞、培養上清、またはその両方から精製することを可能にすることができる。発現、精製、および／または保存中に可溶性を高めるポリペプチドドメインは当技術分野で周知であり、該ポリペプチドドメインには、Fox, J.D. and Waugh D.S., *E. coli Gene Expression Protocols Methods in Molecular Biology*, (2003) 205:99-117およびHan et al. *FEMS Microbiol. Lett.* (2007) 274:132-138(これらはそれぞれ、その全体が参照により本明細書に組み込まれる)等で例示されるよう、例えはマルトース結合タンパク質(MBP)および伸長因子T<sub>s</sub>(T<sub>s</sub>f)が含まれる。対象のポリペプチドに融合したポリペプチドドメインは、対象のポリペプチドのN末端またはC末端で融合することができる。用語「組み換え」は、例えは、アミノ酸もしくは核酸の単離セグメントを、遺伝子工学技術により化学的に合成する、または、操作することにより、さもなければ分離している2つの配列セグメントを人工的に組み合わせることを意味する。

【0076】

一実施形態では、本発明は修正T<sub>n</sub>5トランスポザーゼおよび伸長因子T<sub>s</sub>(T<sub>s</sub>f)を含むトランスポザーゼ融合タンパク質を提供する。T<sub>s</sub>f-T<sub>n</sub>5融合タンパク質は、融合トランスポザーゼおよびフリートランスポゾン端を含む、機能的二量体トランスポソーム複合体に組み込むことができる。T<sub>s</sub>f-T<sub>n</sub>5融合タンパク質の可溶性および熱安定性は、非融合T<sub>n</sub>5タンパク質と比較し高い。

【0077】

別の実施形態では、本発明は、修正T<sub>n</sub>5トランスポザーゼと、5-メチルシトシンを認識するタンパク質ドメインとを含むトランスポザーゼ融合タンパク質を提供する。5-メチルシトシンを提供する。5-メチルシトシン-T<sub>n</sub>5融合タンパク質は、融合トランスポザーゼおよびフリートランスポゾン端を含む、機能的二量体トランスポソーム複合体に組み込むことができる。5-メチルシトシン結合タンパク質ドメインを用いて、例えは、T<sub>n</sub>5トランスポソーム複合体をゲノムのメチル化領域に標的させることができる。

【0078】

さらに別の実施形態では、本発明は、修正T<sub>n</sub>5トランスポザーゼおよびドメインを結合するタンパク質A抗体を含む、トランスポザーゼ融合タンパク質を提供する。タンパク質A-T<sub>n</sub>5融合タンパク質は、融合トランスポザーゼおよびフリートランスポゾン端を含む機能的二量体トランスポソーム複合体に組み込むことができる。タンパク質Aのドメインを結合する抗体は、例えは、T<sub>n</sub>5トランスポソーム複合体を、ゲノムの抗体結合領域に標的させるのに用いることができる。

【0079】

本発明は、修正T<sub>n</sub>5トランスポザーゼと、伸長因子T<sub>S</sub>(T<sub>s</sub>f)の全てまたは一部とを含む、トランスポザーゼ融合タンパク質を提供する。T<sub>s</sub>fはタンパク質タグであり、これを用いて、細菌の発現系、例えは大腸菌発現系で発現した異種タンパク質の可溶性を高めることができる。異種タンパク質の可溶性を高めるT<sub>s</sub>fの機能は、T<sub>s</sub>fタンパク質に固有の高い折りたたみ効率に起因し得る。タンパク質精製実験において(データ表示なし)、T<sub>s</sub>fをタンパク質T<sub>u</sub>との複合体として精製した。T<sub>u</sub>を複合体中で結合させるには、T<sub>s</sub>fを正しく折りたたむことが必要である、T<sub>s</sub>f-T<sub>u</sub>複合体の精製は、T<sub>s</sub>fが正しく折りたたまれたことを示唆する。例示的な、核酸および大腸菌伸長因子T<sub>S</sub>の対応核酸配列は、それぞれ、配列番号22および23と示す。

【0080】

一例では、T<sub>S</sub>-T<sub>n</sub>5融合タンパク質を、T<sub>S</sub>を過活動T<sub>n</sub>5トランスポザーゼのN末端に融合させることにより構築した。変異T<sub>n</sub>5トランスポザーゼタンパク質とのT<sub>S</sub>-融合の例示的アミノ酸配列は、それぞれ配列番号25および26と示す。配列番号24は、配列番号25のT<sub>S</sub>-変異タンパク質融合をコードする核酸配列に対応する。

【0081】

T<sub>S</sub>の全てまたは一部はN末端またはC末端で融合することが可能だが、当業者は、リンクマー配列をT<sub>S</sub>配列とトランスポザーゼのN末端またはC末端との間に挿入可能である

10

20

30

40

50

ことを理解するだろう。一部の実施形態では、リンカー配列は 1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50 以上のアミノ酸長とすることが可能である。一部の実施形態では、融合タンパク質のトランスポザーゼ部分の 1 つまたは複数のアミノ酸は、削除する、または、リンカー配列と置き換えることができる。一部の実施形態では、融合タンパク質のトランスポザーゼ部分の第一メチオニンは、2 つのアミノ酸、例えば、配列番号 25 および 26 で示される G l y - T h r と置き換えることができる。

## 【0082】

10

TS-Tn5 融合構築物を大腸菌で発現させ、発現、可溶性、および熱安定性について評価した。Tn5 の N 末端への TS の融合は Tn5 の可溶性を高めた。可溶性の向上は、トランスポソーム複合体のロバスト性の増大とタンパク質凝集の低下と関連付けることができる。TS-Tn5 トランスポソームの熱安定性は、非融合 Tn5 トランスポソームの熱安定性と比較し実質的に改善する。一例では、熱による Tn5 の凝集は、非融合 Tn5 対照と比較し、Tsf-Tn5 融合構築物において実質的に縮小する。

## 【0083】

一応用例では、TS-Tn5 融合タンパク質は、次世代配列決定プラットフォーム（例えば、 Illumina 社の GA または HiSeq プラットフォーム）での配列決定向に、ディレクショナル RNA-seq ライブリの構築において用いる。

20

## 【0084】

別の適用例では、TS-Tn5 融合タンパク質を正規化プロセスで用いることができる。別の適用例では、TS 可溶化タグを、他の修正（変異）Tn5 トランスポザーゼ酵素の発現および精製のために用いることができる。

## 【0085】

TS 融合タグに特異的な抗体をプルダウンプロセスに用いて、トランスポソームタグ付き配列を捕捉することができる。一例では、TS 融合タグ抗体はウサギポリクローナルである。

## 【0086】

別の適用例では、TS-Tn5 トランスポソームおよび抗 TS 抗体を、混合トランスポソームプロセスで用いることができる。例えば、トランスポソーム反応を、Tsf-Tn5 トランスポソームおよび Tn5 トランスポソーム（つまり、TS でタグ付けされてない）を用いて実行する。抗 TS 抗体を用いて、Tsf-Tn5 トランスポソームタグ付き配列を特異的にプルダウンする。

30

## 【0087】

本発明は、修正 Tn5 トランスポザーゼと、5-メチルシトシンを認識するタンパク質ドメインとを含む、トランスポザーゼ融合タンパク質を提供する。タンパク質ドメインを結合する 5-メチルシトシンは、例えば、Tn5 トランスポソーム複合体をゲノムのメチル化領域に標的させるために用いることができる。一適用例では、ドメイン-Tn5 トランスポソーム複合体を結合する 5-メチルシトシンを用いて、メチル濃縮し断片化した、そしてタグを付けた（タグメント化）ライブリを、メチル化分析用に生成することができる。

40

## 【0088】

一部の実施形態では、トランスポザーゼに融合したポリペプチドドメインは、タンパク質 A のドメインを結合する抗体を含む。タンパク質 A は、強固な折りたたみ特性のない、比較的小さいコンパクトな分子である。タンパク質 A のドメインを結合する抗体は、例えば、Tn5 トランスポソーム複合体を、ゲノムの抗体結合領域に標的させるために用いることができる。例えば、5-メチルシトシンに特異的な抗体を用いて、ゲノムのメチル化領域に結合する、および、該領域を特定することができる。ゲノムの抗体結合領域は、タンパク質 A-Tn5 融合トランスポソーム複合体を用いて、断片化およびタグ付け（つま

50

り、タグメント化)のために実質的に標的とすることができます。

【0089】

一部の実施形態では、トランスポザーゼに融合したポリペプチドドメインは精製タグを含む。本明細書で用いる場合の用語「精製タグ」は、ポリペプチドの精製または特定に適切な任意のペプチド配列を意味する。精製タグは、該精製タグに対し親和性のある別の部分に特異的に結合する。精製タグに特異的に結合するこのような部分は、通常、アガロースビーズなどの母材または樹脂に付着する。精製タグに特異的に結合する部分には、抗体、他のタンパク質(例えば、タンパク質Aまたはストレプトアビシン)、ニッケルまたはコバルトのイオンまたは樹脂、ビオチン、アミロース、マルトース、およびシクロデキストリンが含まれる。例示的な精製タグには(ヘキサヒスチジンペプチドなどの)ヒスチジン(HIS)タグが含まれ、これはニッケルイオンまたはコバルトイオンなどの金属イオンに結合するだろう。他の例示的精製タグは、mycタグ(EQKLISEEDEL)、連鎖球菌(Strep)タグ(WSHPQFEK)、フラグ(Flag)タグ(DYKDDDDK)、およびV5タグ(GKPIPNPLLGLDST)である。用語「精製タグ」にはまた、「エピトープタグ」、つまり、抗体により特異的に認識されるペプチド配列が含まれる。例示的エピトープタグにはFLAGタグが含まれ、これはモノクローナル抗FLAG抗体により特異的に認識される。抗FLAG抗体により認識されるペプチド配列は、配列DYKDDDDKまたは実質的に同一であるその多様体からなる。一部の実施形態では、トランスポザーゼに融合したポリペプチドドメインは、SUMOタグおよびSTR-EPタグなど、下記の実施例1で例示するような2つ以上のタグを含む。用語「精製タグ」にはまた、実質的に同一の精製タグ多様体が含まれる。本明細書で用いる場合、「実質的に同一の多様体」は、精製タグの誘導体または断片を意味し、これは元の精製タグと比較し、(例えば、アミノ酸の置換、削除、または挿入を介して)修正されているが、精製タグを特異的に認識する部分に特異的に結合するという精製タグの特性は保持する。

【0090】

一部の実施形態では、トランスポザーゼに融合するポリペプチドドメインは発現タグを含む。本明細書で用いる場合、用語「発現タグ」は、第2ポリペプチドに付着することができる、対象の組み換えポリペプチドの可溶性、安定性、および/または発現をサポートするとされる、任意のペプチドまたはポリペプチドを意味する。例示的発現タグには、Fc-タグおよびSUMO-タグが含まれる。原則として、任意のペプチド、ポリペプチド、またはタンパク質は発現タグとして用いることが可能である。

【0091】

例えば、細胞単離および培養についての(例えば、後続核酸単離についての)他の有用な参考文献には、Freshney (1994) *Culture of Animal Cells, a Manual of Basic Technique, third edition*, Wiley-Liss, New York、ならびに、そこで言及された参考文献; Payne et al. (1992) *Plant Cell and Tissue Culture in Liquid Systems* John Wiley & Sons, Inc. New York, N.Y.; Gamborg and Phillips (eds) (1995) *Plant Cell, Tissue and Organ Culture; Fundamental Methods* Springer Lab Manual, Springer-Verlag (Berlin Heidelberg New York)、およびAtlas and Parks (eds) *The Handbook of Microbiological Media* (1993) CRC Press, Boca Raton, Flaが含まれる。

【0092】

本明細書で開示する組み換えトランスポザーゼをコードする核酸はまた、本明細書で提示する実施形態の特徴である。特定のアミノ酸を多数のコドンによりコードすることができる、ある翻訳系(例えば、原核細胞または真核細胞)はしばしばコドンバイアスを呈し、例えば、異なる有機体はしばしば、同一のアミノ酸をコードするいくつかの同義語コドンのうち一つを好む。このように、本明細書で提示する核酸は任意で「コドン最適化」されており、これは、トランスポザーゼを発現するのに用いられる特定の翻訳系に好まれるコドンを含むように核酸を合成することを意味する。例えば、細菌細胞において(または、細菌の特定株であっても)トランスポザーゼを発現することが望ましい場合、トランスポザーゼの効率的な発現のために、該細菌細胞のゲノムに最も多く見受けられるコド

ンを含むように核酸を合成することが可能である。真核細胞においてトランスポザーゼを発現することが望ましい場合、同様の手法を用いることができ、例えば、核酸は該真核細胞に好まれるコドンを含むことが可能である。

#### 【0093】

種々のタンパク質の単離方法および検出方法は既知であり、例えば、本明細書で提示する組み換えトランスポザーゼを発現する細胞の組み換え培養物から、トランスポザーゼを単離するために用いることが可能である。種々のタンパク質単離方法および検出方法は、当技術分野では周知であり、例えば、以下に記載のものを含む：R. Scopes, Protein Purification, Springer-Verlag, N.Y. (1982) ; Deutscher, Methods in Enzymology Vol. 1 10 82: Guide to Protein Purification, Academic Press, Inc. N.Y. (1990) ; Sandana (19 97) Bioseparation of Proteins, Academic Press, Inc. ; Bollag et al. (1996) Protein Methods, 2.sup.nd Edition Wiley-Liss, NY ; Walker (1996) The Protein Protocols Handbook Humana Press, NJ, Harris and Angal (1990) Protein Purification Applications: A Practical Approach IRL Press at Oxford, Oxford, England ; Harris and Angal Protein Purification Methods: A Practical Approach IRL Press at Oxford, Oxford, England ; Scopes (1993) Protein Purification: Principles and Practice 3.sup.rd Edition Springer Verlag, NY ; Janson and Ryden (1998) Protein Purification: Principles, High Resolution Methods and Applications, Second Edition Wiley-VCH, NY ; 20 およびWalker (1998) Protein Protocols on CD-ROM Humana Press, NJ ; ならびにそこに引用される参考文献。タンパク質の精製方法および検出方法に関する更なる詳細は、Satinder Ahuja ed., Handbook of Bioseparations, Academic Press (2000)に見られる。

#### 【0094】

##### 使用方法

本明細書で提示する改変トランスポザーゼは、*in vitro* 転移技法などの配列決定手順において用いることが可能である。簡潔に言うと、*in vitro* 転移は、トランスポソーム複合体と標的DNAを接触させることにより開始することが可能である。本開示のトランスポザーゼとともに用いるのに容易に適応させることが可能な例示的な転移手順および系は、例えば、国際公開第10/048605；米国特許出願公開第2012/0301925号明細書；同第2013/0143774号明細書に記載されており、これらはその全体が参照により本明細書に組み込まれる。

#### 【0095】

例えば、一部の実施形態では、（例えば、次世代配列決定または増幅のテンプレートとして用いるために）対象の任意の dsDNA を含む標的DNAからタグ付きDNA断片のライプラリを精製する方法において、本明細書で提示するトランスポザーゼ酵素を用いることが可能であり、該方法は、少なくとも一つのトランスポザーゼと該トランスポザーゼが転移複合体を共に形成するトランスポゾン端組成物とを用いて、標的DNAを*in vitro* 転移反応でインキュベーションするステップであって、該トランスポゾン端組成物は、標的DNAへの多数の挿入が起きる条件および十分な時間において、（i）移動トランスポゾン端配列、および、任意で、該移動トランスポゾン端配列の追加配列 5' - を呈する移動鎖と、（ii）移動トランスポゾン端配列と相補的な配列を呈す非移動鎖とを含み、それぞれは、移動鎖を含む、または移動鎖からなる第1タグを、標的DNAにおいてヌクレオチドの5' 端につなぎ合わせ、それにより標的DNAの断片化とアニールした5' タグ付きDNA断片集団の生成をもたらし、該集団はそれぞれが5' 端に第1タグを有する、ステップと、その後、5' タグ付きDNA断片の3' 端を第1タグまたは第2タグにつなぎ合わせ、それによりタグ付きDNA断片の、（例えば、タグ付き環状ssDNA断片または5' タグ付きおよび3' タグ付きのDNA断片（つまり、「2つタグ付けされたDNA断片」）の何れかを含む）ライプラリを生成するステップと、を含む。

#### 【0096】

一部の実施形態では、*in vitro* 転移反応で用いる、トランスポザーゼおよびトランスポゾン端組成物の量またはトランスポソーム組成物の量は、50マイクロリットル

10

20

30

40

50

反応につき 50 ナノグラムの標的 DNA に対し約 1 ピコモル～約 25 ピコモルである。本発明の任意の方法の一部の好適な実施形態では、in vitro 転移反応で用いるトランスポザーゼおよびトランスポゾン端組成物の量またはトランスポソーム組成物の量は、50 マイクロリットル反応につき 50 ナノグラムの標的 DNA に対し約 5 ピコモル～約 50 ピコモルである。トランスポザーゼが過活動 Tn5 トランスポザーゼであって、トランスポゾン端組成物が MEDS トランスポゾン端組成物を含むか、または、トランスポソーム組成物が前記過活動 Tn5 トランスポザーゼと、MEDS トランスポゾン端を備えるトランスポゾン端組成物とを含む、本発明の何れかの方法の一部の好適な実施形態では、in vitro 転移反応で用いる前記トランスポザーゼおよびトランスポゾン端組成物の量または前記トランスポソーム組成物の量は、50 マイクロリットル反応につき 50 ナノグラムの標的 DNA に対し約 5 ピコモル～約 25 ピコモルである。トランスポザーゼが過活動 Tn5 トランスポザーゼまたは MuA トランスポザーゼである、本発明の何れかの方法の一部の好適な実施形態では、in vitro 転移反応で用いる、トランスポザーゼおよびトランスポゾン端組成物の最終濃度またはトランスポソーム組成物の最終濃度は、少なくとも 250 nM であり；一部の他の実施形態では、過活動 Tn5 トランスポザーゼまたは MuA トランスポザーゼの最終濃度、および、その各々のトランスポゾン端組成物またはトランスポソーム組成物の最終濃度は、少なくとも 500 nM である。  
10

#### 【0097】

一部の実施形態では、本発明は、エクソーム配列決定のためにゲノム DNA ライブライを調製し濃縮する方法を提供する。様々な実施形態において、本発明の方法は、ゲノムライブライの調製のために、改変 Tn5 トランスポザーゼ、例えば、TS-Tn5059 を用いる。一実施形態では、本発明の方法は、改変トランスポザーゼ、例えば、TS-TN5059 の縮小挿入配列バイアスによって引き起こされる、縮小バイアスを有するゲノムライブライの調製を可能にする。TS-Tn5059 を用いたゲノム DNA のタグメント化は、広い GC / AT 範囲にわたり、ゲノムのより完全なカバレッジをもたらす。  
20

#### 【0098】

別の実施形態では、本発明は、DNA インプット許容差が拡大した改変トランスポザーゼを用いたゲノムライブライの調製方法を提供する。TS-Tn5059 を用いたゲノム DNA のタグメント化は、ある範囲の DNA インプット量にわたって均一な挿入サイズをもたらす。一部の実施形態では、本発明はエクソーム配列決定方法を提供する。  
30

#### 【0099】

##### タグメント化反応条件

本明細書では、タグメント化反応用の、反応条件およびバッファを提示する。一部の実施形態では、二価カチオンをタグメント化反応バッファに含む。特定の実施形態では、二価カチオンは、例えば、Co<sup>2+</sup>、Mn<sup>2+</sup>、Mg<sup>2+</sup>、Cd<sup>2+</sup>、またはCa<sup>2+</sup>とすることが可能である。二価カチオンは、塩化物塩といった任意の適切な塩の形態で、例えば、CoCl<sub>2</sub>、MnCl<sub>2</sub>、MgCl<sub>2</sub>、酢酸マグネシウム、CdCl<sub>2</sub>、またはCaCl<sub>2</sub>として含めることができる。特定の実施形態では、下記の実施例で例示するようにタグメント化バッファは CoCl<sub>2</sub> を含む。実施例 5 の実験証拠により示すように、タグメント化バッファ配合における CoCl<sub>2</sub> の追加は、驚くべきことにタグメント化中の配列バイアスを改善する。  
40

#### 【0100】

ある実施形態では、タグメント化バッファは、ある濃度の二価カチオン、つまり、およそ、0.01 mM、0.02 mM、0.05 mM、0.1 mM、0.2 mM、0.5 mM、1 mM、2 mM、5 mM、8 mM、10 mM、12 mM、15 mM、20 mM、30 mM、40 mM、50 mM、60 mM、70 mM、80 mM、90 mM、100 mM、もしくはそれ以上の濃度の二価カチオン、または、これらのうち任意の値の間の範囲にある濃度の二価カチオン、例えば、0.01 mM～0.05 mM、0.02 mM～0.5 mM、8 mM～12 mM 等の二価カチオンを有することができる。一部の実施形態では、タグメント化バッファは、ある濃度の CoCl<sub>2</sub>、つまり、およそ、0.01 mM、0.02 m  
50

M、0.05 mM、0.1 mM、0.2 mM、0.5 mM、1 mM、2 mM、5 mM、8 mM、10 mM、12 mM、15 mM、20 mM、30 mM、40 mM、50 mM、60 mM、70 mM、80 mM、90 mM、100 mM、もしくはそれ以上の濃度のCoCl<sub>2</sub>、または、これらのうち任意の値の間の範囲にある濃度の二価カチオン、例えば、0.01 mM～0.05 mM、0.02 mM～0.5 mM、8 mM～12 mM等の濃度の二価カチオンを有することができる。一部の実施形態では、タグメント化バッファは、ある濃度のMnCl<sub>2</sub>、つまり、およそ、0.01 mM、0.02 mM、0.05 mM、0.1 mM、0.2 mM、0.5 mM、1 mM、2 mM、5 mM、8 mM、10 mM、12 mM、15 mM、20 mM、30 mM、40 mM、50 mM、60 mM、70 mM、80 mM、90 mM、100 mM、もしくはそれ以上の濃度のMnCl<sub>2</sub>、または、これらのうち任意の値の間の範囲にある濃度の二価カチオン、例えば、0.01 mM～0.05 mM、0.02 mM～0.5 mM、8 mM～12 mM等の濃度の二価カチオンを有することができる。一部の実施形態では、タグメント化バッファは、ある濃度のMgCl<sub>2</sub>、つまり、およそ、0.01 mM、0.02 mM、0.05 mM、0.1 mM、0.2 mM、0.5 mM、1 mM、2 mM、5 mM、8 mM、10 mM、12 mM、15 mM、20 mM、30 mM、40 mM、50 mM、60 mM、70 mM、80 mM、90 mM、100 mM、もしくはそれ以上の濃度のMgCl<sub>2</sub>、または、これらのうち任意の値の間の範囲にある濃度の二価カチオン、例えば、0.01 mM～0.05 mM、0.02 mM～0.5 mM、8 mM～12 mM等の濃度の二価カチオンを有することができる。一部の実施形態では、タグメント化バッファは、ある濃度のCdCl<sub>2</sub>、つまり、およそ、0.01 mM、0.02 mM、0.05 mM、0.1 mM、0.2 mM、0.5 mM、1 mM、2 mM、5 mM、8 mM、10 mM、12 mM、15 mM、20 mM、30 mM、40 mM、50 mM、60 mM、70 mM、80 mM、90 mM、100 mM、もしくはそれ以上の濃度のCdCl<sub>2</sub>、または、これらのうち任意の値の間の範囲にある濃度の二価カチオン、例えば、0.01 mM～0.05 mM、0.02 mM～0.5 mM、8 mM～12 mM等の濃度の二価カチオンを有することができる。一部の実施形態では、タグメント化バッファは、ある濃度のCaCl<sub>2</sub>、つまり、およそ、0.01 mM、0.02 mM、0.05 mM、0.1 mM、0.2 mM、0.5 mM、1 mM、2 mM、5 mM、8 mM、10 mM、12 mM、15 mM、20 mM、30 mM、40 mM、50 mM、60 mM、70 mM、80 mM、90 mM、100 mM、もしくはそれ以上の濃度のCaCl<sub>2</sub>、または、これらのうち任意の値の間の範囲にある濃度の二価カチオン、例えば、0.01 mM～0.05 mM、0.02 mM～0.5 mM、8 mM～12 mM等の濃度の二価カチオンを有することができる。

### 【0101】

一部の実施形態では、トランスポザーゼまたはタグメント化反応によるゲノムDNAの断片化は、25～70、37～65、50～65、または50～60の温度範囲で実行することが可能である。一部の実施形態では、トランスポザーゼまたはタグメント化反応によるゲノムDNA断片化は、37、40、45、50、51、52、53、54、55、56、57、58、59、60、61、62、63、64、または65で実行することが可能である。

### 【0102】

#### 改変トランスポザーゼをコードする核酸

本明細書ではさらに、本明細書で提示する改変トランスポザーゼ酵素をコードする核酸分子を提示する。トランスポザーゼの変異バージョンである任意の所定改変トランスポザーゼであって、該改変トランスポザーゼのアミノ酸配列と、好適には、該トランスポザーゼをコードする野生型ヌクレオチド配列も既知である、改変トランスポザーゼについては、分子生物学の基本原則に従って変異をコードするヌクレオチド配列を得ることが可能である。例えば、Tn5トランスポザーゼをコードする野生型ヌクレオチド配列が既知である場合、1つまたは複数のアミノ酸置換を有する、Tn5の任意の所定変異バージョンをコードするヌクレオチド配列を、標準遺伝子コードを用いて推定することが可能である。

10

20

30

40

50

同様に、ヌクレオチド配列は、他のトランスポザーゼの変異バージョンについても容易に導き出すことが可能である。その後、必要なヌクレオチド配列を有する核酸分子を、当技術分野で既知の標準分子生物学を用いて構築することができる。

【0103】

本明細書で提示する実施形態によると、定義した核酸は、同一の核酸だけでなく、任意のマイナーな塩基バリエーションも含み、特に、保存アミノ酸置換にある縮重コードに従つて同義語コドン（同一のアミノ酸残基を特定する異なるコドン）をもたらす場合、置換を含む。用語「核酸配列」には、塩基バリエーションに関して与えられる任意の単一鎖配列に対し、相補的な配列も含まれる。

【0104】

本明細書に記載の核酸分子はまた、有利には、適切な発現ベクターに含まれ、そこからコードされるトランスポザーゼタンパク質を、適切な宿主において発現することができる。クローン化DNAを、前記細胞をその後形質転換し、そして、形質転換された細胞をその後選択するために適切な発現ベクターに組み込むことは、当業者にとって周知であり、これはSambrook et al. (1989), Molecular cloning: A Laboratory Manual, Cold Spring Harbor Laboratory (これはその全体が参照により組み込まれる)で提示されている。

【0105】

このような発現ベクターには、前記DNA断片の発現に影響を及ぼすことが可能な、プロモーター領域などの調節配列に操作可能に連結する、本明細書に提示の実施形態に記載の核酸を有するベクターが含まれる。用語「操作可能に連結する」とは並列を意味し、ここでは記載した構成要素が、意図したやり方で機能することができる関係にある。このようなベクターを適切な宿主細胞に形質転換して、本明細書に提示の実施形態に記載のタンパク質を発現させることができる。

【0106】

核酸分子は、成熟タンパク質、または、プレタンパク質にリーダー配列をコードし、その後宿主細胞により切断され成熟タンパク質を形成するものを含む、プロ配列を有するタンパク質をコードすることができる。ベクターは、例えば、複製起点とともに提供されるプラスミドベクター、ウイルスベクター、またはファージベクター、および、任意で、前記ヌクレオチドの発現用プロモーター、および、任意で、該プロモーターの調節因子とすることができる。ベクターは、例えば抗生物質耐性遺伝子などの1つまたは複数の選択可能なマーカーを含むことができる。

【0107】

発現に必要な調節エレメントには、RNAポリメラーゼを結合し、リボソーム結合のために転写開始配列と翻訳開始配列の適切なレベルを指示する、プロモーター配列が含まれる。例えば、細菌発現ベクターには、lacプロモーターなどのプロモーター、ならびに、翻訳を開始するためのシャイン・ダルガノ配列および開始コドンAUGが含まれ得る。同様に、真核発現ベクターには、RNAポリメラーゼII用の異種または相同的のプロモーター、下流ポリアデニレーションシグナル、開始コドンAUG、およびリボソームを分離するための停止コドンが含まれ得る。このようなベクターは、商業的に得るか、または、当技術分野で周知の方法により、記載される配列から組み立てることができる。

【0108】

高等真核生物による、トランスポザーゼをコードするDNAの転写は、エンハンサー配列をベクターに含めることにより最適化することができる。エンハンサーは、DNAのcis作用エレメントであり、これは、プロモーターに影響を与えて転写レベルを上げる。ベクターは、概して、選択可能なマーカーに加え、複製起点も含むだろう。

【実施例】

【0109】

(実施例1)

アッセイ法および条件の概略

以下の段落に下記に提示する実施例で用いるアッセイ条件の概略を記載する。

10

20

30

40

50

## 【0110】

ヒトgDNAにおいてWGSに対しTN5を用いたタグメント化

このセクションでは、トランスポザーゼの挿入バイアスをモニタリングするために、下記の実施例で用いるタグメント化アッセイについて記載する。簡潔に言うと、50ngのヒトゲノムDNAを、10mM Tris酢酸、pH 7.6、25mM酢酸マグネシウムにおいて、5μLのTDE1と共に5分間55℃でインキュベーションした。その後、12.5mMのHEPES、pH 7.5、1MのNaCl、50mMのMgCl<sub>2</sub>の反応体積の1/5を加え、続いて、TN5トランスポソームを100nMまで加え、60分間30℃でインキュベーションした。反応物はその後、 Illumina社のNextera手順に記載されているように浄化および増幅し、HiSeq2000機器を用いた配列決定向けに提出した。

## 【0111】

TN5トランスポソームアセンブリ

TN5を、25mMのHEPES、pH 7.6、12.5mMのKCl、18.75mMのNaCl、0.375mMのEDTA、31.75%のグリセロールにおいて、30分間室温で、アニール化トランスポゾンと共に20μMまでインキュベーションした。

## 【0112】

トランスポゾンアセンブリ

トランスポゾンを、10mMのTrisHCl、pH 7.5、50mMのNaCl、1mMのEDTAにおいて、反応物を94℃まで加熱し、それをゆっくりと室温まで冷却することにより、40μMまで別々にアニールした。TN5-ME-Aでは、TN5モザイク端配列A14(TN5MEA)をTN5非移動配列(NTS)にアニールし、TN5-ME-Bでは、TN5端配列B15(TN5MEB)をTN5非移動配列(NTS)にアニールした。これらの配列を以下に示す。

TN5MEA: 5'-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG-3'

TN5MEB: 5'-GTCTCGTGGCTCGGAGATGTGTATAAGAGACAG-3'

TN5 NTS: 5'-CTGTCTCTTATACACATCT-3'

## 【0113】

2.トランスポザーゼのクローニングおよび発現

このセクションでは、下記の実施例で用いる様々なトランスポザーゼ変異体のクローニングおよび発現で用いるアプローチについて記載する。

## 【0114】

突然変異誘発を、標準の部位特異的突然変異誘発方法論を用い、トランスポザーゼに対しバックボーン遺伝子配列をコードする遺伝子において実行した。作成した各変異について、クローン遺伝子配列を配列決定することにより、変異遺伝子の適切な配列を確かめた。

## 【0115】

TN5トランスポザーゼ遺伝子を修正pET11aプラスミドにクローン化した。修正プラスミドは、pASK5plusに由来する精製タグStrepTag-IIと、pET-SUMOに由来するSUMOとを含んだ。発現をBL21(DE3)pLysYコンピテント細胞(New England Biolabs)を用いて実行した。細胞を25℃で0.5OD<sub>600nm</sub>まで成長させ、その後、100μMのIPTGを用いて誘導した。発現を18℃で19時間実行した。細胞ペレットを、プロテアーゼ阻害剤の存在下で、100mMのTrisHCl、pH 8.0、1MのNaCl、1mMのEDTAにおいて、ミクロ流動化剤を用いて溶解した。溶解後、細胞溶解物を、デオキシコール酸を用いて最終濃度が0.1%になるまで30分間インキュベーションした。30000×gで20分間遠心分離する前に、ポリエチレンイミンを0.5%まで加えた。上清を集め、同量の飽和硫酸アンモニウム溶液と混ぜ、その後、氷の上で少なくとも1時間攪拌した。溶液

10

20

30

40

50

をその後20分間30000×gで遠心分離し、100mMのTris、pH8.0、1MのNaCl、1mMのEDTA、および1mMのDTTにおいてペレットを再懸濁した。再懸濁し、ろ過した溶液を、その後、AKTA精製システムを用いてS treptac tinカラムに適応した。カラムを、100mMのTris、pH8.0、1mMのEDTA、1mMのDTT、4MのNaClで洗浄し、その後100mMのTris、pH7.5、1mMのEDTA、1mMのDTT、100mMのNaClで洗浄した。溶出を、100mMのTris、pH7.5、1mMのEDTA、1mMのDTT、100mMのNaCl、および5mM デスチオビオチンを用いて実行した。溶出物を、100mMのTris pH 7.5、100mMのNaCl、0.2mMのEDTA、および2mMのDTTを用いて、ヘパリントラップカラム(heparin trap column)に乗せた。同じバッファを用いて洗浄した後、Tn5多様体を、100mMのTris pH 7.5、1MのNaCl、0.2mMのEDTA、および2mMのDTTを用いて、塩勾配で溶出した。画分を集め、プールし、濃縮化し、そして、グリセロールを加え、-20°での長期保存の前に50%最終濃度を生成した。

## 【0116】

3. 大腸菌ゲノムDNAにおける挿入バイアスのIVC分析

挿入配列バイアスの分析を、IVC-プロットデータ(強度対サイクル)を用いて実行した。このデータは、各DNAトランスポザーゼを用いて生成したDNAライプラリを配列決定した後に、利用可能に生成した。簡潔に言うと、突然変異誘発および発現を上記のように実行した。トランスポザーゼを上記のトランスポゾンAおよびトランスポゾンBを用いてインキュベーションし、大腸菌ゲノムDNAと共にインキュベーションして、DNAライプラリを生成した。各ライプラリを、製造業者の指示書に従い、MiSeq Fast chemistry(Illumina社、カリフォルニア州、サンディエゴ)を行うIllumina社のGenome Analyzerにおいて、少なくとも35サイクルほど配列決定した。Illumina RTAソフトウェアを用いて、各サイクルでベースコード(base call)と強度値を生成した。IVCプロットを生成するため、配列決定リードを大腸菌参照ゲノムに位置合わせし、各サイクルで4つの塩基それぞれの強度(発生)を、位置合わせした全ての配列決定リードの全強度値の画分としてプロットした。

## 【0117】

## (実施例2)

挿入バイアスに対するTn5トランスポザーゼ変異の特定

この実施例は、IVC-プロットデータ(強度対サイクル)を用いた挿入配列バイアスの分析について記載する。このデータは、各DNAトランスポザーゼを用いて生成したDNAライプラリを配列決定した後に利用可能となった。この分析は、安定した結果をもたらすのにいくらかの配列決定リード(20k~40k)のみを必要とし、各Tn5トランスポザーゼ多様体を発現するのに用いる大腸菌細胞溶解物で実行することが可能で、HTS(ハイスループットスクリーニング)目的に適っていた。

## 【0118】

様々な単一アミノ酸置換Tn5多様体についての代表的な結果を、図2~7に記載する。図示する多様体は、例えば、位置248での置換による対称性の喪失(図2)、位置119での置換によるIVCプロットの平坦化(図3)、位置125での置換または位置248の後ろへの挿入による、IVCプロット後半におけるIVCの縮小(図4、5)、および、位置K120で異なる芳香族アミノ酸を使用することによる、9bpから10bpへの複製の増加(図6、対称性が1および9番目のbpから1および10番目のbpへ変化することを示す)を示す。これらの結果は、特定の変異が、wt対照と比較し、改善した挿入バイアスをもたらすことを示す。

## 【0119】

## (実施例3)

細菌gDNAにおける全ゲノム配列決定

10

20

30

40

50

これらの実験は、a) 推定ライプラリサイズ / 多様性およびb) AT / GC - ドロップアウトについて、様々なトランスポザーゼ変異体を比較するために行った。これらの実験は、精製し、アクティビティを正規化したTn5トランスポザーゼ多様体を用いて行った。これらの実験は、500k ~ 1Mの配列決定リード / 実験を必要とする。

【0120】

示した精製Tn5トランスポザーゼ多様体を用いてセレウス菌gDNAにおいてタグメント化を実行することにより、結果を得た。酵素をアクティビティにより正規化し、Nextera（登録商標）キットと共に販売されている商業TDE1酵素のアクティビティに合うように設定した。

【0121】

Tn5001として示す変異体は配列番号11と同じアミノ酸配列を有し、「wt」対照の役割をする。上記表で示すように、Tn5058、Tn5059、およびTn5061として示す変異体は、それぞれ配列番号18、19、および20と同じアミノ酸配列を有する。実験は三回行い、データは、集めたデータの平均と標準偏差を示す。「推定ライプラリサイズ」は、最適複製を用いることなしに計算し、再現可能な結果をもたらす。

10

【0122】

図8Aに示すように、GC - ドロップアウトは低く維持するのに対し、Tn5058およびTn5059のAT - ドロップアウトは、Tn5001と比較し顕著に減少する。同様に、図8Bに示すように、ライプラリサイズは $1.7 \times (Tn5058)$ および $2.2 \times (Tn5059)$ と、著しく増加する。これらの結果は、変異体Tn5058および変異体Tn5059は、野生型トランスポザーゼと比較し、配列挿入バイアスを大きく改善することを示し、下記の実施例4で記載するさらなる実験につながる。

20

【0123】

(実施例4)

ヒトgDNAについてのNextera Rapid Capture Enrichment実験

以下の実験を、実施例3で上記したものと同じ、精製し、アクティビティを正規化したTn5トランスポザーゼ多様体を用いて行った。これらの実験は、典型的には40M ~ 100Mの配列決定リード / 実験を必要とし、配列決定データを分析して、a) 多様性、b) 濃縮、c) カバレッジ、d) カバレッジ均一性、e) ペナルティスコアを比較した。

30

【0124】

Nextera Rapid Capture Exome ( Illumina (登録商標) ) CEXのプールキャプチャープローブ ( pool capture probes ) を製造業者の指示書に従って用い、キャプチャを三重に実行した。図9Aに示すように、示した変異体は、wt対照Tn5001と比較し、カバレッジ均一性において顕著な改善をもたらした。さらに、図9Bに示すように、平均標的カバレッジはより低いにも関わらず、標的カバレッジについて $10 \times$ および $20 \times$ でテストした変異体により、統計的に著しい改善がもたらされた。

【0125】

図10Aに示すように、示した変異体は、固有リード数およびハイブリッド選択ライプラリサイズの増大をもたらした。同様に、図10Bに示すように、変異体は、Tn5001と比較してより低いペナルティスコアをもたらし、これは、カバレッジ深さが $10 \times$ 、 $20 \times$ 、または $30 \times$ に到達するのに必要な、倍強い配列決定である。これらの結果は、テストした変異体が、対照と比較し、より優れた挿入バイアスおよびより均一なカバレッジをもたらすことを示す。

40

【0126】

(実施例5)

Tn5アクティビティに対するタグメント化バッファ組成物の効果

以下の実験を行い、ライプラリアウトプットおよび配列決定測定基準に対する、Tn5タグメント化バッファ組成物の効果および反応条件を特徴付けた。

50

## 【0127】

ライプラリアウトプットおよび配列決定測定基準に対する、Tn5タグメント化バッファ組成物の効果および反応条件を評価するため、Tn5タグメント化DNAライプラリを、セレウス菌ゲノムDNAを用いて構築した。2つの異なるTn5トランスポザーゼを、タグメント化ライプラリ、つまり変異Tn5（「TS-Tn5059」）および対照過活動Tn5（「TS-Tn5」）の構築のために用いた。Tsは融合タグであり、これはTn5およびTn5059のタンパク質の精製のために用いる。Tn5059は、過活動Tn5アミノ酸配列（配列番号11）に関し、4つの追加変異体、K212R、P214R、G251R、およびA338Vを有する。TS-Tn5059は、Tn5059のN末端にTSタグを含む。一部の実施形態では、TS-タグのC末端は、リンカーによりTn5059のN末端に融合することができ、これは、第1メチオニン残基を置換する。一部の実施形態では、リンカーはGly-Thrである。10

## 【0128】

TS-Tn5059を10、40、および80nMの最終濃度で用いた。TS-Tn5を、4、15、および30nMの最終濃度で用いた。TS-Tn5059およびTS-Tn5に対する酵素濃度を（標準バッファ配合を用いて）正規化し、ほぼ同じレベルのタグメント化アクティビティを提供した。つまり、10、40、および80nMのTS-Tn5059は、4、15、および30nMのTn5とそれぞれほぼ同じレベルのアクティビティを有する。各タグメント化ライプラリを、25ngのセレウス菌ゲノムDNAインプットを用いて調製した。セレウス菌のゲノム内容量は約40%のGCと約60%のATである。20

## 【0129】

タグメント化バッファを2×配合で調製した。2×配合とは以下である：標準バッファ（TD；20mMのTris酢酸、pH7.6、10mMの酢酸マグネシウム、および20%ジメチルホルムアミド（DMF））；コバルトバッファ（Co；20mMのTris酢酸、pH7.6、および20mMのCoCl<sub>2</sub>）；コバルト+DMSOバッファ（Co-DMSO；20mMのTris酢酸、pH7.6、20mMのCoCl<sub>2</sub>、および20%ジメチルスルホキシド（DMSO））；高分子量バッファ（HMW；20mMのTris酢酸、pH7.6、および10mMの酢酸マグネシウム）；NF2バッファ（NF2；20mMのTris酢酸、pH7.6、20mMのCoCl<sub>2</sub>、および20%DMF）。CoCl<sub>2</sub>を含むタグメント化バッファは、毎日新しく調製した。各ライプラリについて、50μLの総反応体積中で、20μLのセレウス菌ゲノムDNA（25ng）、25μLの2×タグメント化バッファ、および5μLの酵素（10×Ts-Tn5059または10×Ts-Tn5）を混ぜることにより、タグメント化反応を行った。反応物を55で5分間インキュベーションした。タグメント化反応に続き、サンプルを標準Nextera（登録商標）サンプル調製プロトコルに従って処理した。ライプラリを、MiSeq装置において、 Illumina社のSBS（sequencing-by-synthesis）化学反応を用いて配列決定した。配列決定ランは、V2 MiSeqキットを用いて2×71サイクルだった。各ライプラリにおける断片サイズ分布は、配列決定ランはV2 MiSeqキットを用いて2×71サイクルだった。各ライプラリの断片サイズ分布は、Bioanalyzerで評価した。30

## 【0130】

図11は、異なるタグメント化バッファを用いて調製した、TS-Tn5059およびTS-Tn5のタグメント化DNAライプラリにおける固有分子数についての棒グラフ100を示す。ライプラリ中の固有分子数は、ライプラリの多様性（複雑性）の表れである。グラフの各棒は、タグメント化ライプラリを表す。実験は3回繰り返した（n=3）。対照ライプラリ（つまり、標準タグメント化バッファを用いて調製したライプラリ）は、「酵素-酵素濃度-DNAインプット」により指定する。例えば、棒グラフ100の1番目の棒は「TS-Tn5059-10nM-25ng」とラベル付けされるが、これは、標準バッファ配合において最終濃度が10nMのTS-Tn5059と、25ngのイン40

プットDNAとを用いて調製した対照ライプラリを指定する。修正タグメント化バッファ配合を用いて調製したライプラリは、「酵素 - 酵素濃度 - バッファ添加物 - DNAインプット」により指定する。例えば、棒グラフ100の4番目の棒は「TS - Tn5059 - 10nM - Co - 25ng」とラベル付され、これは、10mMのCoCl<sub>2</sub>を含む修正タグメント化バッファにおいて最終濃度が10nMのTS - Tn5059を用いて調製したライプラリを指定する。データは、10mMのCoCl<sub>2</sub>を含むタグメント化バッファ(つまり、Coバッファ、Co - DMSOバッファ、およびNF2バッファ)を用いて調製した、TS - Tn5059およびTS - Tn5のタグメント化ライプラリは、CoCl<sub>2</sub>の添加なしのバッファ(つまり、標準バッファまたはHMW)において調製したライプラリと比較し、平均多様性がより高い。

10

### 【0131】

図12は、異なるタグメント化バッファを用いて調製した、TS - Tn5059およびTS - Tn5のタグメント化DNAライプラリにおける、GCドロップアウト率についての棒グラフ200を示す。対照ライプラリおよび修正タグメント化バッファを用いて調製したライプラリは、図11に記載するように指定する。GCドロップアウトは、タグメント化ライプラリからドロップした(なくなった)、ゲノム中のGCリッチ領域の割合として定義することができる。データは、標準タグメント化バッファを用いて調製した、TS - Tn5059およびTS - Tn5の対照ライプラリでは、GCドロップアウト率が比較的低いことを示す。データはまた、10mMのCoCl<sub>2</sub>を含むタグメント化バッファ(つまり、Coバッファ、Co - DMSOバッファ、およびNF2バッファ)を用いて調製した、TS - Tn5059およびTS - Tn5のタグメント化バッファは、CoCl<sub>2</sub>の添加なしのバッファ(つまり、標準バッファまたはHMW)で調製したライプラリと比較し、GCドロップアウト率がより高い(つまり、最大約6%)ことを示す。Co含有バッファを用いて調製したライプラリにおけるGCドロップアウトの増加は、TS - Tn5059およびTS - Tn5の濃度の上昇により改善する。例えば、TS - Tn5059 - 10nM - Co - 25ngライプラリにおけるGCドロップアウト率は、TS - Tn5059 - 10nM - 25ng対照ライプラリと比較し、比較的高い。TS - Tn5059の濃度を40nM(つまり、TS - Tn5059 - 40nM - Co - 25ng)および80nM(つまり、TS - Tn5059 - 80nM - Co - 25ng)まで上昇させると、GCドロップアウト率は低下する。

20

### 【0132】

図13は、異なるタグメント化バッファを用いて調製したTS - Tn5059およびTS - Tn5のタグメント化DNAライプラリにおけるATドロップアウト率についての棒グラフ300を示す。対照ライプラリおよび修正タグメント化バッファを用いて調製したライプラリを、図11に記載するように指定する。ATドロップアウトは、タグメント化ライプラリからドロップした(なくなった)、ゲノム中のATリッチ領域の割合として定義することができる。データは、標準タグメント化バッファを用いて調製した、TS - Tn5059およびTS - Tn5の対照ライプラリでは、ある量のATドロップアウト(つまり、それぞれ、約1%~約3%、および約7%~約3%)が観察されることを示す。データはまた、低濃度(つまり、10nM)の酵素と10mMのCoCl<sub>2</sub>を含むタグメント化バッファ(つまり、Coバッファ、Co - DMSOバッファ、およびNF2バッファ)とを用いて調製したTS - Tn5059タグメント化ライプラリは、CoCl<sub>2</sub>の添加なしのバッファ(つまり、標準バッファまたはHMW)において調製したライプラリと比較し、ATドロップアウト率が低いことを示す。同様に、10mMのCoCl<sub>2</sub>を含むタグメント化バッファ(つまり、Coバッファ、Co - DMSOバッファ、およびNF2バッファ)を用いて調製したTS - Tn5ライプラリは、CoCl<sub>2</sub>の添加のないバッファ(つまり、標準バッファまたはHMW)において調製したライプラリと比較し、ATドロップアウト率はより低い。

30

### 【0133】

ここで図12および13を参照すると、タグメント化バッファにCoCl<sub>2</sub>(10nM

40

50

)を添加すること(つまり、Coバッファ、Co-DMSOバッファ、およびNF2バッファ)は、タグメント化ライプラリにおけるGCおよびATのドロップアウト率を「フリップ(flip)」することができる。例えば、TS-Tn5059-10nm-Co-25ngライプラリにおけるGCドロップアウト率(図12)は、TS-Tn5059-10nm-25ng対照ライプラリと比較し比較的高い。一方、TS-Tn5059-10nm-Co-25ngライプラリにおけるATドロップアウト率(図13)は、TS-Tn5059-10nm-25ng対照ライプラリと比較し、比較的低い(または、ない)。

#### 【0134】

図14は、標準バッファ(TD)配合およびコバルトバッファ(Co)配合を用いて調製したTS-Tn5059ライプラリにおける、断片サイズ分布についてのBioanalyzerトレースのプロット400を示す。プロット400は、TS-Tn5059-10nM-Co-25ngライプラリにおける断片サイズ分布の曲線である曲線410、TS-Tn5059-40nM-Co-25ngライプラリにおける断片サイズ分布の曲線である曲線415、TS-Tn5059-80nM-Co-25ngライプラリにおける断片サイズ分布の曲線である曲線420、TS-Tn5059-10nM-TD-25ngライプラリにおける断片サイズ分布の曲線である曲線425、TS-Tn5059-10nM-TD-25ngライプラリにおける断片サイズ分布の曲線である曲線430、および、TS-Tn5059-80nM-TD-25ngライプラリにおける断片サイズ分布の曲線である曲線435を示す。プロット400はまた、塩基対(bp)におけるDNA断片サイズの標準ラダー(standard ladder)である曲線440を示す。ラダー(左から右へ)における断片サイズを表2に示す。

#### 【0135】

#### 【表2】

表2. サイズラダー	
ラダーピーク	サイズ(bp)
1	35
2	50
3	100
4	150
5	200
6	300
7	400
8	500
9	600
10	700
11	1,000
12	2,000
13	3,000
14	7,000
15	10,380

#### 【0136】

データは、タグメント化反応で用いるTS-Tn5059の濃度を10nMから40nMおよび80nMへ増加させることは、断片サイズ分布をより小さい断片サイズにシフトさせることを示す。断片サイズ分布のシフトは、標準バッファ(TD)配合を用いて調製したライプラリにおいてより顕著である。例えば、コバルトバッファ(Co)配合を用いて調製したライプラリにおけるフラグメントサイズ分布は、10nMのTS-Tn505

10

20

30

40

50

9 (曲線410)を用いて調製したライプラリでは約3000bpであり、40nMのTS-Tn5059および80nMのTS-Tn5059(それぞれ、曲線415および曲線420)を用いて調製したライプラリでは約1000~約2000bpである。標準バッファ(TD)配合および80nMのTS-Tn5059(曲線435)を用いて調製したライプラリでは、断片サイズ分布は、約200bp~約1000bpである。

#### 【0137】

図15は、コバルト-DMSO(Co-DMSO)、NF2、およびHMWのバッファ配合を用いて調製したTS-Tn5059における断片サイズ分布についての、Bioanalyzerトレースのプロット500を示す。プロット500は、TS-Tn5059-10nM-Co-DMSO-25ngライプラリにおける断片サイズ分布の曲線である曲線510、TS-Tn5059-10nM-NF2-25ngライプラリにおける断片サイズ分布の曲線である曲線515、TS-Tn5059-10nM-HMW-25ngライプラリにおける断片サイズ分布の曲線である曲線520、TS-Tn5059-40nM-Co-DMSO-25ngライプラリにおける断片サイズ分布の曲線である曲線525、TS-Tn5059-40nM-NF2-25ngライプラリにおける断片サイズ分布の曲線である曲線530、TS-Tn5059-40nM-HMW-25ngライプラリにおける断片サイズ分布の曲線である曲線535、TS-Tn5059-80nM-Co-DMSO-25ngライプラリにおける断片サイズ分布の曲線である曲線540、TS-Tn5059-80nM-NF2-25ngライプラリにおける断片サイズ分布の曲線である曲線545、およびTS-Tn5059-80nM-HMW-25ngライプラリにおける断片サイズ分布の曲線である曲線550を示す。プロット500はまた、図14のプロット400の曲線440を示し、これは、塩基対(bp)単位のDNA断片サイズの標準ラダーである。

#### 【0138】

データは、概して、タグメント化反応で用いるTS-Tn5059の濃度を10nMから40nMおよび80nMへ増加させることは、断片サイズ分布をより小さい断片サイズにシフトさせることを示す。断片サイズ分布のシフトは、CoCl<sub>2</sub>を含まないHMWバッファを用いて調製したライプラリ(例えば、曲線520および曲線535)において、Co-DMSOを用いて調製したライプラリ(例えば、曲線510および曲線525)と比較し、より顕著である。

#### 【0139】

図16は、標準バッファ配合(TD)およびコバルトバッファ(Co)配合を用いて調製したTS-Tn5ライプラリにおける断片サイズ分布についてのBioanalyzerトレースのプロット600を示す。プロット600は、TS-Tn5-4nM-Co-25ngライプラリにおける断片サイズ分布の曲線である曲線610、TS-Tn5-15nM-Co-25ngライプラリにおける断片サイズ分布の曲線である曲線615、TS-Tn5-30nM-Co-25ngライプラリにおける断片サイズ分布の曲線である曲線620、TS-Tn5-4nM-TD-25ngライプラリにおける断片サイズ分布の曲線である曲線625、TS-Tn5-15nM-TD-25ngライプラリにおける断片サイズ分布の曲線である曲線630、および、TS-Tn5-30nM-TD-25ngライプラリにおける断片サイズ分布の曲線である曲線635を示す。プロット600はまた、図14のプロット400の曲線440も示し、これは、塩基対(bp)単位のDNA断片サイズの標準ラダーである。

#### 【0140】

データは、タグメント化反応において用いるTS-Tn5の濃度を4nMから15nMおよび30nMへ増加させることは、断片サイズ分布をより小さい断片サイズにシフトさせることを示す。断片サイズ分布におけるシフトは、標準バッファ(TD)配合を用いて調製したライプラリにおいてより顕著である。この観察結果は、図14のTS-Tn5059ライプラリでの断片サイズ分布と同様である。

#### 【0141】

10

20

30

40

50

図17は、コバルト-DMSO (Co-DMSO)、NF2、およびHMWのバッファ配合を用いて調製したTS-Tn5ライプラリにおける断片サイズ分布についての、Bi o a n a l y z e rトレースのプロット700を示す。プロット700は、TS-Tn5-4nM-Co-DMSO-25ngライプラリにおける断片サイズ分布の曲線である曲線710、TS-Tn5-4nM-NF2-25ngライプラリにおける断片サイズ分布の曲線である曲線715、TS-Tn5-4nM-HMW-25ngライプラリにおける断片サイズ分布の曲線である曲線720、TS-Tn5-15nM-Co-DMSO-25ngライプラリにおける断片サイズ分布の曲線である曲線725、TS-Tn5-15nM-NF2-25ngライプラリにおける断片サイズ分布の曲線である曲線730、TS-Tn5-15nM-HMW-25ngライプラリにおける断片サイズ分布の曲線である曲線735、TS-Tn5-30nM-Co-DMSO-25ngライプラリにおける断片サイズ分布の曲線である曲線740、TS-Tn5-30nM-NF2-25ngライプラリにおける断片サイズ分布の曲線である曲線745、およびTS-Tn5-30nM-HMW-25ngライプラリにおける断片サイズ分布の曲線である曲線750を示す。プロット700はまた、図14のプロット400の曲線400も示し、これは、塩基対(bp)単位のDNA断片サイズの標準ラダーである。

#### 【0142】

データは、タグメント化反応において用いるTS-Tn5の濃度を4nMから15nMおよび30nMへ増加させることは、断片サイズ分布をより小さい断片サイズにシフトさせることを示す。断片サイズ分布におけるシフトは、標準バッファ(TD)配合を用いて調製したライプラリにおいてより顕著である。この観察結果は、図15のTS-Tn5059ライプラリにおける断片サイズ分布と同様である。

#### 【0143】

概して、ここで図14～17を参照すると、10nMのCoCl<sub>2</sub>（例えば、Coバッファ、Co-DMSOバッファ、およびNF2バッファ）を含むタグメント化バッファを用いて調製したTn5059およびTS-Tn5のライプラリにおける断片サイズは、CoCl<sub>2</sub>なしのタグメント化バッファ（つまり、TDおよびHMWバッファ）を用いて調製したTS-Tn5059およびTS-Tn5のライプラリよりも大きい。

#### 【0144】

図18A、18B、18C、18Dはそれぞれ、TS-Tn5ライプラリにおける配列含有量のバイアスグラフ800、TS-Tn5-Coライプラリにおける配列含有量のバイアスグラフ830、TS-Tn5-Co-DMSOライプラリにおける配列含有量のバイアスグラフ840、およびTS-Tn5-NF2ライプラリにおける配列含有量のバイアスグラフ850を示す。バイアスグラフ（つまり、強度対サイクル数(IVC)のグラフ）は、SBSサイクル数の関数として、観察された塩基(A, C, G、またはT)の比率をプロットし、タグメント化中にTn5が有する好適配列コンテキストを示す。

#### 【0145】

バイアスグラフ800、830、840、および850はそれぞれ、サイクル数によるA含有量の曲線である曲線810、サイクル数によるC含有量の曲線である曲線815、サイクル数によるG含有量の曲線である曲線820、およびサイクル数によるT含有量の曲線である曲線825を示す。例えば、図18AのTS-Tn5のライプラリにおいて、塩基Gを表す曲線820は、サイクル1で観察される塩基の約38%がGであることを示し；塩基Tを表す曲線825は、サイクル1で観察される塩基の約15%がTであることを示す。

#### 【0146】

図18Aを参照すると、データは、Tn5配列バイアスが、標準タグメント化バッファ配合を用いて調製したTS-Tn5ライプラリにおけるSBSの、最初の15サイクルあたりで観察されることを示す。約15サイクル後、配列バイアスは徐々に縮小し、A、T、C、およびGの含有量は予想されるゲノム組成物を反映する。セレウス菌については、ゲノムは約40%のGCと約60%のATであり、これはサイクル16または17あたり

からサイクル 3 5までのバイアスグラフで表され、ここで、曲線 8 1 0 (つまり、 A ) および曲線 8 2 5 (つまり、 T ) は約 3 0 % ( A + T ~ 6 0 % ) に収束し；曲線 8 1 5 (つまり、 C ) および曲線 8 2 0 (つまり、 G ) は、約 2 0 % ( C + G ~ 4 0 % ) に収束する。

〔 0 1 4 7 〕

図18B、18C、および18Dを参照すると、データはまた、Tn5配列バイアスが、CoCl<sub>2</sub>を含んだタグメント化バッファを用いて調製したライプラリである、TS-Tn5-Co、TS-Tn5-Co-DMSO、およびTS-Tn5-NF2のライプラリにおいて、SBSの最初の15サイクルあたりで観察されることを示す。この場合もやはり、図18Aに関して記載したように、約15サイクルより後で配列バイアスは徐々に縮小し、A、T、C、およびGの含有量は、予想されるゲノム組成物を反映する。しかしながら、TS-Tn5-Co、TS-Tn5-Co-DMSO、およびTS-Tn5-NF2のライプラリでは、曲線810(つまり、A)および曲線825(つまり、T)は、サイクル10～サイクル15あたりで、予想されるゲノム組成物に向かってシフトし始め、曲線815(つまり、C)および曲線820(つまり、G)は、サイクル10～サイクル15あたりで、予想されるゲノム組成物に向かってシフトし始める。加えて、サイクル2～8の間のバイアスは、図18Aと比べた場合、縮小する。データは、タグメント化バッファ配合におけるCoCl<sub>2</sub>の添加が、タグメント化中にTn5配列を改善することを示す。

〔 0 1 4 8 〕

図19A、19B、19C、および19Dはそれぞれ、TS-Tn5059ライプラリにおける配列含有量のバイアスグラフ900、TS-TN5059-Coライプラリにおける配列含有量のバイアスグラフ930、TS-Tn5059-Co-DMSOライプラリにおける配列含有量のバイアスグラフ940、およびTS-Tn5059-NF2ライプラリにおける配列含有量のバイアスグラフ950を示す。バイアスグラフ900、930、940、および950はそれぞれ、サイクル数によるA含有量の曲線である曲線910、サイクル数によるC含有量の曲線である曲線915、サイクル数によるG含有量の曲線である曲線920、および、サイクル数によるT含有量の曲線である曲線925を示す。

[ 0 1 4 9 ]

図19Aを参照すると、データは、Tn5059配列バイアスが、TS-Tn5059タグメント化ライブラリにおいて、SBSの最初の15サイクルあたりで観察されることを示す。約15サイクルより後で配列バイアスは縮小し、A、T、C、およびGの含有量は、図18Aに関して記載したように、予想されるゲノム組成物を反映する。しかしながら、変異体Tn5059は、図18Aに示すTn5配列バイアスと比較し、配列バイアスの縮小を示す。TS-Tn5059ライブラリでは、曲線910（つまり、A）および曲線925（つまり、T）は、サイクル10～サイクル15あたりで、予想されるゲノム組成物に向かってシフトし始め、曲線915（つまり、C）および曲線920（つまり、G）は、サイクル10～サイクル15あたりで、予想されるゲノム組成物に向かってシフトし始める。

〔 0 1 5 0 〕

図 19 B、19 C、および 19 D を参照すると、データはまた、Tn5059 配列バイアスが、CoCl<sub>2</sub> を含んだタグメント化バッファを用いて調製したライプラリである、TS-Tn5059-Co、TS-Tn5059-Co-DMSO、および Ts-Tn5059-NF2 のライプラリにおいて、SBS の最初の 15 サイクルあたりで観察されることを示す。この場合もやはり、図 18 A に関して記載したように、約 15 サイクルより後で配列バイアスは徐々に縮小し、A、T、C、および G の含有量は、予想されるゲノム組成物を反映する。しかしながら、TS-Tn5059-Co ライプラリでは、曲線 910 (つまり、A) および 曲線 925 (つまり、T) は、サイクル 5 あたりで、予想されるゲノム組成物に向かってシフトし始め、曲線 915 (つまり、C) および 曲線 920 (つ

まり、G)は、サイクル5あたりで、予想されるゲノム組成物に向かってシフトし始める。T S - T n 5 0 5 9 - C o - D M S O、およびT s - T n 5 0 5 9 - N F 2のライプラリでは、曲線910(つまり、A)および曲線925(つまり、T)は、サイクル5よりも前に、予想されるゲノム組成物に向かってシフトし始め、曲線915(つまり、C)および曲線920(つまり、G)は、サイクル5よりも前に、予想されるゲノム組成物に向かってシフトし始める。

**【0151】**

**(実施例6)**

**M o s 1 アクティビティに対するタグメント化バッファ組成物の影響**

以下の実験を行い、ライプラリアウトプットおよび配列決定測定基準に対する、T n 5 10 タグメント化バッファ組成物の効果および反応条件を特徴付けた。

**【0152】**

M o s 1 タグメント化D N Aライプラリを、セレウス菌ゲノムD N Aを用いて構築した。タグメント化ライプラリの構築で用いるM o s 1 トランスポザーゼは、M B P - M o s 1 融合タンパク質だった。マルトース結合タンパク質(M B P)は融合タグであり、これはM o s 1 タンパク質の精製に用いる。M B P - M o s 1 は、1 0 0  $\mu$  Mという最終濃度で用いた。各タグメント化ライプラリは、5 0 n gのセレウス菌ゲノムD N Aインプットを用いて調製した。

**【0153】**

タグメント化配合を2×配合で調製した。2×配合は以下の通りである：標準バッファ( T D ; 2 0 m MのT r i s 酢酸、p H 7 . 6、1 0 m Mの酢酸マグネシウム、および2 0 %のジメチルホルムアミド(D M F)；T D + N a C l ( T D - N a C l ; 2 0 m M のT r i s 酢酸、p H 7 . 6、1 0 m Mの酢酸マグネシウム、2 0 %のD M F、および2 0 0 m MのN a C l )；高分子量バッファ(H M W ; 2 0 m MのT r i s 酢酸、p H 7 . 6、および1 0 m Mの酢酸マグネシウム)；H E P E S ( 5 0 m MのH E P E S、p H 7 . 6、1 0 m Mの酢酸マグネシウム、2 0 %のD M F)；H E P E S - D M S O ( 5 0 m M のH E P E S p H 7 . 6、1 0 m Mの酢酸マグネシウム、および2 0 %のD M S O)；H E P E S - D M S O - C o ( 5 0 m MのH E P E S、p H 7 . 6、2 0 %のD M S O、および2 0 m MのC o C l 2 )；ならびに、H E P E S - D M S O - M n ( 5 0 m MのH E P E S、p H 7 . 6、2 0 %のD M S O、および2 0 m Mのマンガン(M n) )。C o C l 2 を含むタグメント化バッファは、毎日新しく調製した。

**【0154】**

各ライプラリについて、5 0  $\mu$  Lの総反応体積中で、2 0  $\mu$  Lのセレウス菌ゲノムD N A(5 0 n g)、2 5  $\mu$  Lの2×タグメント化バッファ、および5  $\mu$  Lの酵素(1 0 × M B P - M o s 1)を混ぜることにより、タグメント化反応を行った。反応物を3 0 で6 0 分間インキュベーションした。タグメント化反応に続き、サンプルを標準N e x t e r a(登録商標)サンプル調製プロトコルに従って処理した。ライプラリを、M i S e q 装置において、I l l u m i n a社のS B S ( s e q u e n c i n g - b y - s y n t h e s i s )化学反応を用いて配列決定した。配列決定ランは、2×7 1 サイクルだった。

**【0155】**

図20は、異なるタグメント化バッファを用いて調製した、M B P - M o s 1 タグメント化D N Aライプラリにおける平均総リード数および平均多様性についての棒グラフ1000を示す。総リード数とは、フローセルからの総リード数である。多様性とは、ライプラリ中の固有分子数であり、ライプラリの複雑性の指標として用いる。グラフの棒の各対は、タグメント化ライプラリを表す。実験は3回繰り返した(n = 3)。最初の2つのグラフ棒、E Z T n 5 - s t d - b c e r e u s およびN e x t e r a V 2 - 3 0 C は、比較ライプラリであり、これはそれぞれ、T n 5 および標準バッファ配合を用いて、5 5 および3 0 で調製した。タグメント化反応のためにM B P - M o s 1 を用いて調製したライプラリは、「酵素 - 酵素濃度 - バッファ」により指定する。例えば、グラフ棒の3番目の対は「M B P M o s 1 - 1 0 0  $\mu$  M - T D」とラベル付してあり、これは、標準タグ

20

30

40

50

メント化バッファ (T D) において、最終濃度が  $100 \mu M$  の M B P M o s 1 を用いて調製したライプラリを指定する。データは、比較的同一の数または配列決定リードの下、M o s 1 タグメント化により調製したライプラリの多様性に対する、異なるバッファの効果を示す。特に、H E P E S - D M S O - M n バッファは、ライプラリの多様性の増大に寄与する。

#### 【0156】

図21は、M B P - M o s 1 タグメント化ライプラリにおける、G C と A T のドロップアウトについての棒グラフ 1100 を示す。G C および A T のドロップアウトはそれぞれ、タグメント化ライプラリからドロップした（なくなった）、ゲノム中の G C リッチ領域および A T リッチ領域の割合として定義することができる。ライプラリは図20に記載するように指定する。データは、E Z T n 5 および N e x t e r a V 2 (つまり、T n 5 トランスポザーゼ) を用いて調製したライプラリでは、G C ドロップアウトは基本的ないが、A T リッチ領域は、それぞれ約 7 % および約 5 % がタグメント化ライプラリからドロップすることを示す。M B P - M o s 1 および標準タグメント化バッファ (M B P M o s 1 - 100  $\mu M$  - T D) を用いて調製したライプラリでは、A T ドロップアウトは基本的ないが、G C リッチ領域の約 2 % 以下がタグメント化ライプラリからドロップする。M B P - M o s 1 タグメント化ライプラリにおける G C ドロップアウト率は、タグメント化バッファの組成物により影響を受ける。G C ドロップアウト率は、H M W 、H E P E S 、H E P E S - D M S O 、H E P E S - D M S O - C o 、および H E P E S - D M S O - M n バッファを用いて調製した M B P - M o s 1 タグメント化ライプラリにおいて増加する。

#### 【0157】

図22A、22B、22C、および22Dはそれぞれ、M o s 1 - H E P E S ライプラリにおける配列含有量についてのバイアスグラフ 1200 、M o s 1 - H E P E S - D M S O ライプラリにおける配列含有量についてのバイアスグラフ 1230 、M o s 1 - H E P E S - D M S O - C o ライプラリにおける配列含有量についてのバイアスグラフ 1240 、および、M o s 1 - H E P E S - D M S O - M n ライプラリにおける配列含有量についてのバイアスグラフ 1250 を示す。バイアスグラフ 1200 、1230 、1240 、および 1250 はそれぞれ、サイクル数による A 含有量の曲線である曲線 1210 、サイクル数による C 含有量の曲線である曲線 1215 、サイクル数による G 含有量の曲線である曲線 1220 、および、サイクル数による T 含有量の曲線である曲線 1225 を示す。

#### 【0158】

図22Aおよび22Bを参照すると、データは、M o s 1 配列バイアスが、M o s 1 - H E P E S および M o s 1 - H E P E S - D M S O のタグメント化ライプラリにおいて、S B S の最初の数サイクルで観察されることを示す。第1 S B S サイクルでは、T の検出はフローセルを通して約 100 % である。第2 S B S サイクルでは、A の検出はフローセルを通して約 100 % である。約 4 サイクルより後で配列バイアスは縮小し、A 、 T 、 C 、および G の含有量は、予想されるゲノム組成物を反映する。

#### 【0159】

図22Cおよび22Dを参照すると、データはまた、M o s 1 配列バイアスが、マグネシウム (M g) をそれぞれコバルト (C o) またはマンガン (M n) で置き換えたタグメント化バッファを用いて調製したライプラリである、M o s 1 - H E P E S - D M S O - C o および M o s 1 - H E P E S - D M S O - M n のタグメント化ライプラリにおいて、S B S の最初の数サイクルで観察されることを示す。この場合もやはり、約 4 サイクルより後で配列バイアスは縮小し、A 、 T 、 C 、および G の含有量は、予想されるゲノム組成物を反映する。しかしながら、M o s 1 - H E P E S - D M S O - C o および M o s 1 - H E P E S - D M S O - M n のライプラリ、つまり、曲線 1210 (つまり、A) および 曲線 1225 (つまり、T) では、予想されるゲノム組成物へのシフトは、サイクル 1 および サイクル 2 において観察される。予想されるゲノム組成物へのシフトは、M o s 1 - H E P E S - D M S O - M n ライプラリにおいてより顕著である。

10

20

30

40

50

## 【0160】

(実施例7)

TS-Tn5059ライプラリ調製およびエクソーム濃縮プロトコル

一実施形態では、本発明の方法は、Tn5トランスポソームベースのエクソームライプラリの調製および濃縮に対し、合理化ワークフローを提供する。

## 【0161】

図23は、エクソーム配列決定用のゲノムDNAライプラリを調製し濃縮する方法1260の一例のフロー図を示す。方法1260は、TS-Tn5059トランスポソームと、現在のNextera(登録商標) Rapid Captureプロトコルのあるプロセスステップへの修正とを用いて、ある範囲のDNAインプット量にわたる改善されたライプラリ収量および配列決定測定基準を提供する。例えば、方法1260は、「両面」固相可逆固定(solid phase reversible immobilization、SPRI(登録商標))プロトコル(Agencourt AMPure XPビーズ; Beckman Coulter社)を用いて標的DNAを精製し、そして、PCR增幅の前に、第1DNA断片サイズ選択ステップおよび第2DNA断片サイズ選択ステップを提供する。別の例では、前濃縮プロセスを用いて、エクソーム濃縮の前に標的DNAライプラリを濃縮する。方法1260は、以下のステップを含むが、それらに限定されない。

## 【0162】

ステップ1270で、ゲノムDNAをトランスポソームによりタグメント化(タグ付けおよび断片化)する。トランスポソームはゲノムDNAを同時に断片化して端にアダプタ配列を加え、後続のPCRによる增幅を可能にする。一例では、トランスポソームはTS-Tn5059である。タグメント化反応の完了時に、タグメント化停止バッファを反応物に加える。タグメント化停止バッファは、TS-Tn5059トランスポソーム複合物をタグメント化DNAから確実に十分に変性することができるよう、修正することができる(例えば、停止バッファにおけるSDS濃度を、高熱と組み合わせて、0.1%から1.0%へ増加させる)。

## 【0163】

ステップ1275では、第1浄化を行ってトランスポソームからタグメント化DNAを精製し、第1DNA断片サイズ選択ステップを提供する。DNA断片サイズは、DNAに対するSPRIビーズの、体積対体積の比率を変更することにより選択することができる(例えば、1×SPRI=1:1体積SPRI:DNA)。例えば、第1サイズ選択では、DNAに対するSPRIビーズの体積比率を選択し、あるサイズよりも大きいDNA断片を結合させ(つまり、より大きいDNA断片をサンプルから取り除く)、一方で、あるサイズよりも小さいDNA断片は上清に残す。サイズを選択したDNA断片を有する上清を、次の処理のため浄化反応容器に移す。より大きいDNA断片が付いたSPRIビーズは廃棄してよい。一実施形態では、SPRIビーズの濃度は0.8X~1.5Xまで変えることが可能である。一実施形態では、SPRIビーズの濃度は0.8Xである。

## 【0164】

ステップ1280では、第2浄化を行って、あるサイズ範囲のDNA断片をさらに選択する。例えば、DNAに対するSPRIビーズの体積比率を選択して、あるサイズよりも大きいDNA断片を結合させる(つまり、所望のサイズ範囲のDNA断片をSPRIビーズに結合させる)。より小さいDNA断片は上清に残し、廃棄する。結合DNA断片はその後、次の処理のためにSPRIビーズから溶出する。

## 【0165】

任意のステップ1285では、DNA断片サイズ分布を求める。DNA断片サイズ分布は、例えば、Bioanalyzerを用いて求める。

## 【0166】

ステップ1290では、精製タグメント化DNAを限定サイクル(limited-cycle)PCRプログラムを用いて増幅する。PCRステップはまた、次のクラスタ生

10

20

30

40

50

成および配列決定に必要な共通アダプタ (P 5 および P 7) と同様に、インデックス 1 (i 7) およびインデックス 2 (i 5)、ならびに、配列決定を加える。両面 S P R I プロセス (つまり、ステップ 1275 およびステップ 1280) を用いて所望の D N A 断片サイズ範囲を選択することから、所望のサイズ範囲にあるタグメント化 D N A 断片のみ、P C R 増幅で利用可能となる。その結果、ライプラリ収量は著しく増加し、次の配列決定測定基準 (例えば、リード濃縮率) は改善する。

【0167】

ステップ 1295 では、増幅したタグメント化 D N A ライプラリを、ビーズベースの精製プロセスを用いて精製する。

【0168】

任意のステップ 1300 では、P C R 後の D N A 断片サイズ分布を求める。D N A 断片サイズ分布は、例えば、B i o a n a l y z e r を用いて求める。

【0169】

ステップ 1310 では、タグメント化 D N A ライプラリを、次のエクソーム濃縮のためのハイブリダイゼーションの前に予め濃縮する。例えば、タグメント化 D N A ライプラリを約 50  $\mu$  L から約 10  $\mu$  L へ予め濃縮する。タグメント化 D N A ライプラリを予め濃縮することにより、ハイブリダイゼーション速度はより速くなり、ハイブリダイゼーションの時間は減る。

【0170】

ステップ 1320 では、エクソーム濃縮用の第 1 ハイブリダイゼーションを行う。例えば、D N A ライプラリを、対象領域に標的させたビオチン化捕捉プローブと混ぜる。D N A ライプラリを約 95 で約 10 分間変性し、約 58 で約 30 分間プローブにハイブリダイズし、総反応時間を約 40 分とする。

【0171】

ステップ 1325 では、ストレプトアビジンビーズを用いて、対象標的領域にハイブリダイズしたビオチン化プローブを捕捉する。2 つの加熱洗浄手順を用いて、非特異的に結合した D N A をビーズから取り除く。濃縮ライプラリをその後ビーズから溶出し、ハイブリダイゼーションの第 2 ラウンドに備える。

【0172】

ステップ 1330 では、エクソーム濃縮用の第 2 ハイブリダイゼーションを、第 1 ハイブリダイゼーションと同じプローブおよび遮断薬を用いて行う。例えば、ステップ 155 により溶出した D N A ライプラリを約 95 で約 10 分間変性し、約 58 で約 30 分間ハイブリダイズし、総反応時間を約 40 分とする。第 2 ハイブリダイゼーションを用いて、捕捉領域の高特異性を確保する。

【0173】

ステップ 1335 では、ストレプトアビジンビーズを用いて、対象標的領域にハイブリダイズしたビオチン化プローブを捕捉する。2 つの加熱洗浄手順を用いて、非特異的に結合した D N A をビーズから取り除く。エクソーム濃縮ライプラリをその後ビーズから溶出し、配列決定に備えて 10 サイクルの P C R により増幅する。

【0174】

ステップ 1340 では、エクソーム濃縮捕捉サンプル (つまり、エクソーム濃縮 D N A ライプラリ) を、ビーズベースの精製プロトコルを用いて精製する。

【0175】

ステップ 1345 では、エクソーム濃縮 D N A ライプラリを配列決定のために P C R 増幅する。

【0176】

ステップ 1350 では、増幅した濃縮 D N A ライプラリを、ビーズベースの精製プロトコルを用いて任意で精製する。例えば、1 X S P R I ビーズプロトコルを用いて、次のクラスタ増幅および配列決定を妨げる可能性のある不要な産物 (例えば、過剰プライマー) を取り除く。

10

20

30

40

50

## 【0177】

方法100は、約11時間でライプラリ調製とエクソーム濃縮を可能にする。任意のステップ1285および1300を省略する場合、方法1260は約9時間でライプラリ調製とエクソーム濃縮を可能にする。

## 【0178】

(実施例8)

TS-Tn5059挿入バイアス

トランスポザーゼは、タグメント化反応において、ある挿入部位(DNA配列)バイアスを有する場合がある。DNA配列バイアスは、ゲノムのある領域(例えば、GC-リッチ、またはAT-リッチ)をタグメント化ライプラリからドロップさせ得る。例えば、Tn5トランスポザーゼは、ゲノムのGC-リッチ領域に対しあるバイアスを有する。その結果、ゲノムのAT領域はTn5タグメント化ライプラリにおいてドロップする場合がある。より完全なゲノムカバレッジを提供するため、最少の配列バイアスが望ましい。

## 【0179】

ライプラリアウトプットおよび配列決定測定基準に対するTS-Tn5059トランスポソームの影響を評価するため、TS-Tn5059タグメント化DNAライプラリを、全ゲノム配列決定用標準Nextera(登録商標)DNAライプラリ調製キットおよびセレウス菌ゲノムDNAを用いて調製した。TS-Tn5059は最終濃度40nMで用いた。参照対照ライプラリを、25nMのNextera V2トランスポソームという標準反応条件を用いて調製した。ライプラリをsequencing-by-synthesis(SBS)により評価した。

## 【0180】

図24Aは、TS-Tn5059トランスポソームを用いて調製したタグメント化セレウス菌ゲノムDNAライプラリにおけるカバレッジについてのプロット1400を示す。TS-Tn5059トランスポソームは、GC含有量が増えるにつれ、増加するバイアスレベルに対し耐性を示すようになる。図24Bは、Nextera V2トランスポソームを用いて調製したタグメント化セレウス菌ゲノムDNAライプラリにおけるカバレッジについてのプロット1450を示す。図24Bは、GC含有量が増えると、GCリッチ領域のNextera V2カバレッジはバイアスの増加とともに歪み出す。データは、TS-Tn5059を用いて調製したタグメント化DNAライプラリが、Nextera V2を用いて調製したタグメント化ライプラリと比較し、より挿入バイアスが低く、広いGC/AT範囲にわたって、改善されたさらに均一のカバレッジを有することを示す。

## 【0181】

図25Aは、TS-Tn5059トランスポソームを用いて調製したタグメント化セレウス菌ゲノムDNAライプラリにおける、ギャップ位置およびギャップ長についてのプロット1500を示す。図25Bは、Nextera V2トランスポソームを用いて調製したタグメント化セレウス菌ゲノムタグメント化DNAライプラリにおける、ギャップ位置およびギャップ長についてのプロットを示す。TS-Tn5059タグメント化ライプラリにおけるギャップ数は27である。Nextera V2タグメント化ライプラリにおけるギャップ数は208である。データは、TS-Tn5059トランスポソームを用いて調製したタグメント化DNAライプラリが、Nextera V2トランスポソームを用いて調製したタグメント化ライプラリと比較し、よりギャップの少ない、さらに均一なカバレッジを有することを示す。

## 【0182】

(実施例8)

TS-Tn5059DNAインプット許容差

タグメント化ライプラリの調製に、酵素DNA断片化ステップ(例えば、トランスポソーム仲介タグメント化)を用いることから、例えば機械的断片化方法と比較し、DNAインプットに対しより影響を受けやすい場合がある。一実施例では、現在のNextera(登録商標)Rapid Capture Enrichmentプロトコルは、50n

g の総ゲノムDNAインプットに対し最適化されている。より高質量のゲノムDNAインプットは、不完全なタグメント化およびより大きい挿入サイズをもたらす可能性があり、これは、次の濃縮動作に影響を与える。より低質量のゲノムDNAインプットまたは低品質のゲノムDNAは、タグメント化反応において、予想より小さい挿入サイズを生成する可能性がある。より小さい挿入物は、次の浄化ステップ中に失われ、ライプラリ多様性が低下する場合がある。

【0183】

断片（挿入）サイズ分布に対する異なるDNAインプット量の影響を評価するため、様々な酵素濃度で、様々な量のインプットゲノムDNAを用いてTS-Tn5059タグメント化ライプラリを調製し、断片サイズを他トランスポザーゼで得た断片サイズと比較した。該他トランスポザーゼのアクティビティは、40nMのTS-Tn5059および25ngのゲノムDNAインプットのアクティビティに対し正規化する。

【0184】

40nMのTS-Tn5059、正規化TDE1（Tn5バージョン-1）、および正規化TS-Tn5により、25ngのヒトゲノムDNAを用いて生成した断片サイズ分布は、図26および27に示すように類似していた。

【0185】

しかしながら、TS-Tn5059は、より高い酵素濃度で、広範囲のインプットDNA量にわたり、DNAインプット許容差の拡大を示した。図28は、ある範囲のDNAインプットを用いて調製したタグメント化ゲノムDNAライプラリにおける断片サイズ分布についての、Bioanalyzerトレースのパネル1600を示す。240nMのTS-Tn5059タグメント化ライプラリを、タグメント化、1.8XSPRI浄化により調製し、その後Bioanalyzerトレースを続けた。参照対照ライプラリを、現在のNextera（登録商標）Rapid Captureキット（「Nextera」）およびAgilent QXTキット（「Agilent QXT」）を用いて調製した。タグメント化ライプラリを、25ng、50ng、75ng、および100ngのセレウス菌ゲノムDNAを用いて調製した。データは、TS-Tn5059トランスポソームを用いて調製したタグメント化DNAライプラリの断片サイズ分布は、NexteraトランスポソームまたはAgilent QXTトランスポソームを用いて調製したライプラリと比較し、25～100ngのDNAインプット範囲にわたってより一貫していることを示す。DNAインプットの量が25ngから100ngに増加するにつれ、TS-Tnタグメント化ライプラリにおいてタグメント化DNAの収量は増えるが、断片サイズ分布は実質的に同じままである。対照的に、75ngおよび100ngのDNAインプットでは、NexteraおよびAgilent QXTのタグメント化ライプラリは、DNA断片サイズ分布が大きい断片サイズへ実質的にシフトすることを示す。

【0186】

図29Aは、5ngから100ngへ変動するヒトCoriell DNAインプットを用いて第1ユーザが調製した、TS-Tn5059タグメント化ライプラリにおける断片サイズ分布についての、Bioanalyzerトレースのプロット1700を示す。図29Bは、第2ユーザが調製したTS-Tn5059タグメント化ライプラリにおける断片サイズ分布についての、Bioanalyzerトレースのプロット1750を示す。タグメント化ライプラリは、5ng、10ng、25ng、50ng、75ng、および100ngのセレウス菌ゲノムDNAを用いて調製した。図29Aのプロット1700および図29Bのプロット1750は両者とも、5ngのDNAインプットを用いて調製したタグメント化ライプラリにおける断片サイズ分布のライン1710、10ngのDNAインプットを用いて調製したタグメント化ライプラリにおける断片サイズ分布のライン1715、25ngのDNAインプットを用いて調製したタグメント化ライプラリにおける断片サイズ分布のライン1720、50ngのDNAインプットを用いて調製したタグメント化ライプラリにおける断片サイズ分布のライン1725、75ngのDNAインプットを用いて調製したタグメント化ライプラリにおける断片サイズ分布のライン1730、お

10

20

30

40

50

および 100 ng の DNA インプットを用いて調製したタグメント化ライプラリにおける断片サイズ分布のライン 1735 を示す。データは、TS - TN 5059 タグメント化 DNA ライプラリにおける断片サイズ分布が、5 ~ 100 ng の DNA インプット範囲で一貫していることを示す。断片サイズ分布における一貫性は、異なるユーザについて観察される。

【0187】

別の例では、表 3 に、5 ~ 200 ng の DNA インプット範囲にわたる TS - TN 5059 タグメント化 DNA ライプラリでの、ライプラリ挿入サイズの中央値を示す。

【0188】

【表 3】

10

表 3. 5 ng ~ 200 ng の DNA インプットを用いた挿入サイズ中央値									
インプット DNA (ng)	0	5	0	5	00	50	00	平均土標準偏差	
挿入中央値 (bp)	64	69	58	44	71	78	79	75	167 ± 11

【0189】

さらに別の例では、表 4 に、25 ng、50 ng、75 ng、および 100 ng の DNA インプットを用いて調製した TS - TN 5059 タグメント化 DNA についての、ライプラリ挿入サイズとエクソーム濃縮配列決定測定基準を示す。データは、リード濃縮率(%)が約 80 % であることを示す。現在の Nextera (登録商標) Rapid Capture Enrichment プロトコルを用いて調製したタグメント化ライプラリについてのリード濃縮率は、約 60 % である(データ表示なし)。データはまた、15 ng ~ 100 ng の DNA インプット範囲にわたり、一貫した挿入サイズを示す。

20

【0190】

【表 4】

30

表 4. 25 ~ 100 ng の DNA インプットを用いた挿入サイズおよびエクソーム濃縮測定基準				
エクソーム(Picard)測定基準	5ng	0ng	5ng	00ng
リード濃縮率	8	9	0	7
複製率	.6	.7	.9	.4
ゼロカバレッジ率	.6	.9	.2	.2
10x でのエクソームカバレッジ率	1.6	2.4	2	2.4
挿入サイズ 150 ± 25 bp	67	69	68	76
前濃縮ライプラリ quant (ng/μL)	23	40	65	52
				145 ± 15

【0191】

別の例では、表 5 に、TS - TN 5059 タグメント化ライプラリにおける、25 ng ~ 100 ng の DNA インプット範囲にわたる前濃縮ライプラリの収量を示す。

【0192】

40

## 【表5】

表5. 25~100ngのDNAインプットを用いた前濃縮ライプラリ収量				
実験 #	15μLでの収量 (25ng インプット)	15μLでの収量 (50ng インプット)	15μLでの収量 (75ng インプット)	15μLでの収量 (100ng インプット)
実験 #1	1098	1468	2420	2240
実験 #2	1332	2200	2480	1974
実験 #3	1845	2100	2475	2295
実験 #4	1860	2295	2895	2760
実験 #5	1830	2895	1515	1665
実験 #6	1785	2760	1725	1920

## 【0193】

さらに別の例では、表6に、TS-Tn5059タグメント化DNAライプラリについてのエクソーム濃縮配列決定測定基準を示す。50ngのインプットDNAで開始し、ライプラリは、エクソーム濃縮に対し500ng、625ng、および750ngのライプラリDNAインプットを用いて調製した。データは、エクソーム濃縮測定基準は、ある範囲の前濃縮ライプラリインプット量(つまり、500ng~750ng)にわたり一貫していることを示す。

## 【0194】

## 【表6】

表6. 500~750ng 前濃縮ライプラリインプット量に対するエクソーム濃縮測定基準				
マークティング測定基準	00ng	25ng	50ng	平均±標準偏差
リード濃縮率	2.2	1.7	2.3	82.1±0.32
複製率	.1	.3	.9	4.1±0.2
ゼロカバレッジ率	.7	.6	.7	1.67±0.06
10xでのエクソームカバレッジ率	4.3	5.4	5.6	85.5±0.7
HSライプラリサイズ	9M	0M	1M	50M±1M
平均カバレッジ	5.1	8.2	8.6	47.3±1.92
HS 20xペナルティ	.5	.3	.4	6.4±0.1

## 【0195】

さらに別の例では、表7に現在のNextera(登録商標) Rapid Capture Enrichmentハイブリダイゼーションプロトコル(「NRC」)と、図23の方法1260の濃縮ステップ1310~1350とを用いて調製したタグメント化DNAライプラリについての、エクソーム濃縮配列決定測定基準を示す。データは、エクソーム濃縮測定基準が、現在のNextera(登録商標) Rapid Capture Enrichmentプロトコル(「NRC」)においてハイブリダイゼーションプロトコル用いて調製したライプラリと比較し、図23の方法1260を用いて調製したTS-Tn5059タグメント化ライプラリにおいて改善および/または維持されることを示す。

## 【0196】

10

20

30

40

【表7】

キーとなるエクソーム測定基準	N R C ハイブリダイゼーション	方法 1 0 0 ハイブリダイゼーション
リード濃縮（パッドなし）	54.3%	77.7%
リード濃縮（パッド付き）	64.3%	85%
平均カバレッジ	50.3X	52.1X
ゼロ標的ドロップアウト	1.6%	1.9%
複製率（10m i 1. リード）	3.1%	4%
10Xでのカバレッジ	83.3	84.3%
H S ライブラリサイズ	62.3 M	61 M
標的における選択率	76.5%	85.2%
H S 20x ペナルティ	7.4	6.8

## 【0197】

別の実験では、T S - T n 5 0 5 9 は、より高い濃度（6 X に正規化）で、同じ濃度に正規化したT n 5 バージョン - 1 およびT S - T n 5 トランスポザーゼと比較し、DNAインプット許容差の改善を示す。結果を図30～33を示す。6 X 「正規化」濃度のT n 5 バージョン1（図30）およびT S - T n 5 （図31）は、両者とも、g DNAインプットを25～100ngの間で変動させると断片サイズ分布のシフトを示す。対照的に、6 X 正規化濃度のT S - T n 5 0 5 9 は、10～100ngのDNAインプットでは顕著なサイズシフトを示さない（図32）。断片サイズ分布は、200～500ngへg DNAを増加させると、シフトし始める（図33）。T S - T n 5 0 5 9 のDNAインプット許容差の拡大についての結果を下記の表8にまとめた。

## 【0198】

【表8】

表8：50 μ Lの最終反応物における。T S - T n 5 0 5 9 (nM) : g DNA (ng) の比率

TS-Tn5059 体積 (uL)	在庫体積 (nM)	最終反応物 体積 (μL)	反応物の最終濃度 (nM)	gDNA インプット (ng)	50 μ L 反応物中の TS-Tn5059 (nM) : gDNA (ng) の比率	コメント
5	400	50	40	25	1.6	WGS 比率
5	400	50	40	50	0.8	
15	800	50	240(6X)	500	0.48	
15	800	50	240(6X)	400	0.6	
15	800	50	240(6X)	300	0.8	
15	800	50	240(6X)	200	1.2	
15	800	50	240(6X)	100	2.4	観察されるインプット許容差
15	800	50	240(6X)	75	3.2	
15	800	50	240(6X)	50	4.8	
15	800	50	240(6X)	25	9.6	
15	800	50	240(6X)	10	24	
15	800	50	240(6X)	5	48	

## 【0199】

10

20

30

40

50

このように、TS-Tn5059では、2.4 (nM TS-Tn5059 : ng インプットDNA) の比率においてDNAインプット許容差の拡大を示すサイズシフトはなかった。

【 0 2 0 0 】

本出願を通して、様々な刊行物、特許、および／または特許出願を参照した。これらの刊行物の開示内容はその全体が、本出願で参照されることにより本明細書に組み込まれる。

【 0 2 0 1 】

用語「含む」は本明細書では非限定的であることを意図し、言及された要素だけを含むのではなく、任意の追加要素をさらに包含する。

10

【 0 2 0 2 】

多くの実施形態が記載されている。そうは言うものの、様々な修正が可能であることが理解されよう。したがって、他の実施形態も以下の請求項の範囲内にある。

【 1 A 】

71	AEALRKAGAMQTYVLAQEFPELLATEITSL DAKEK KALIGRE KSAVEKDAGASATIDLW DCSPGIWQLDCTHL VEHLDAMQWINGDYLH Mu <sup>2</sup> 7 Mos <sup>2</sup> 7	102	S G KWL IAEN Y TIGQEIA NADS R VNAARYQQQLI
124	WWYHSVLLAEAT. TFRTVGL [HQ. EWW. RNFLLGVTHHNNELRD. L. IIGLSSL. KVLGVAVHAA. . . . . S.G. PA. PKTWFWQDVK. . . . . TR. KILGWRC. TMLCVWWD.Q. . . . . S.G. V. [IYY ELLKPGET. . . . .	103	HAYLQDKL KSLA. N STTKV. AAC KVKEDDPKG. LFL VRDTLE. T. L.
	AAAT. SR. LR[GMSMSN]V. IAVCDR. . . . . EAD. [V KKFLKFLK. L. KFVTDR. . . . . G. AN GRW. P. VKTVDR. . . . . G. S. NFT. G. S. L. S. FMD. VV. TRYG. I. P. EDF. H. ITDNT. RGAA. NKWLTTGGAPNRI. R. F. KVKEDDPKG. L. F. L N. L. N. RA. LQ. RKR. P. EYR. V. [IFLHDNA. P. SH. T. [ARA. . . . .	223	
	AAHNERFVVRSKHPRK. KKFLKFLK. L. KFVTDR. . . . . G. AN GRW. P. VKTVDR. . . . . G. S. NFT. G. S. L. S. FMD. VV. TRYG. I. P. EDF. H. ITDNT. RGAA. NKWLTTGGAPNRI. R. F. KVKEDDPKG. L. F. L N. L. N. RA. LQ. RKR. P. EYR. V. [IFLHDNA. P. SH. T. [ARA. . . . .	342	MEEDPNIERMVISILSIVARLQLRESFTP 372 WWAGI. N. KQEFGINKEL. S. L. S. I. PA. . . . . S. A. SER. I. F. S. L. G. N L. MGAKMHWNSVVA. Y. S. P. D. A. P. S. D. Y. . . . . H. F. A. S. M. G. H. A. E. Q. R AHNERFVVRSKHPRK. KKFLKFLK. L. KFVTDR. . . . . G. AN GRW. P. VKTVDR. . . . . G. S. NFT. G. S. L. S. FMD. VV. TRYG. I. P. EDF. H. ITDNT. RGAA. NKWLTTGGAPNRI. R. F. KVKEDDPKG. L. F. L N. L. N. RA. LQ. RKR. P. EYR. V. [IFLHDNA. P. SH. T. [ARA. . . . .

【 义 1 B 】

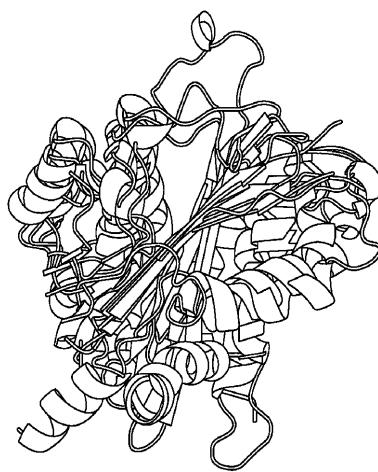
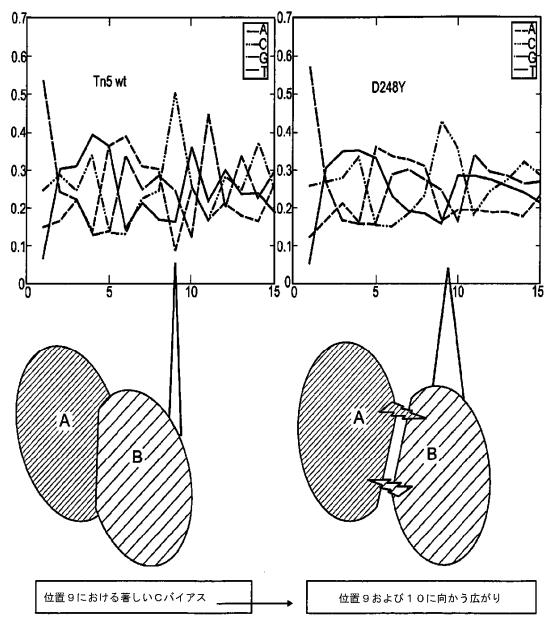
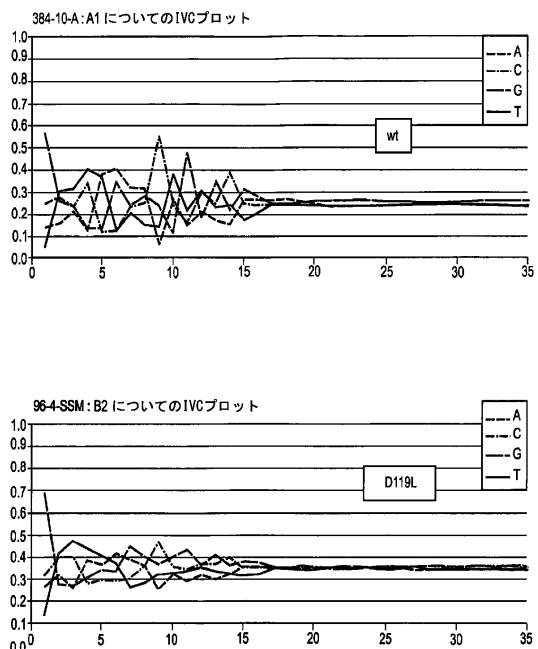


Fig. 1B

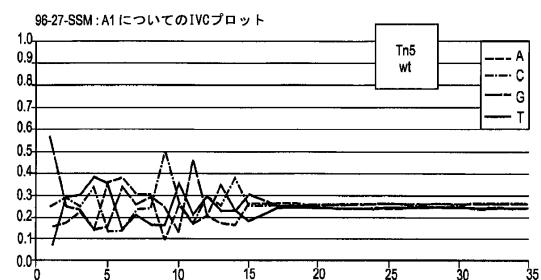
【図2】



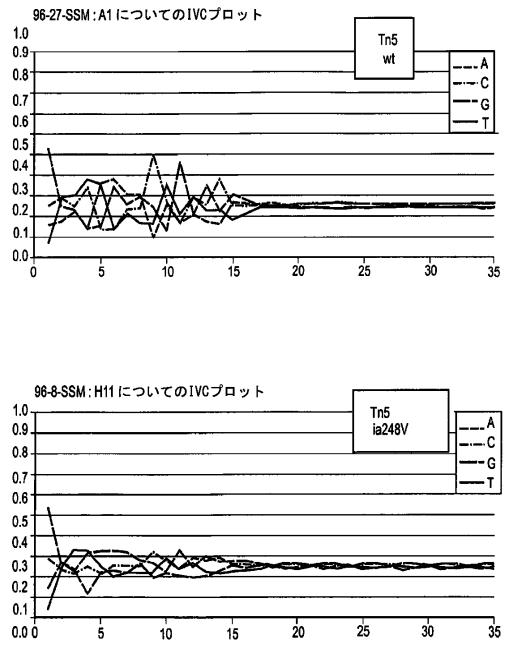
【図3】



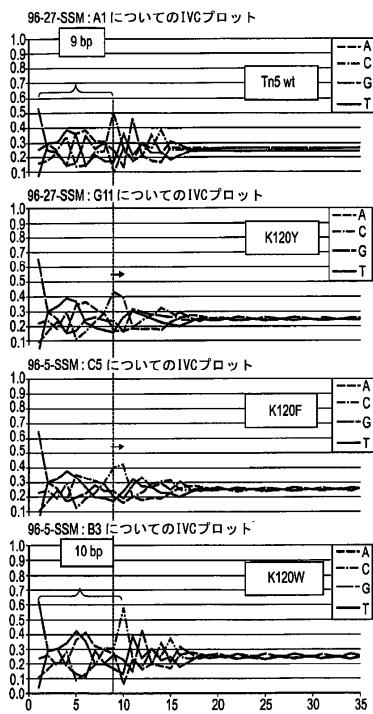
【図4】



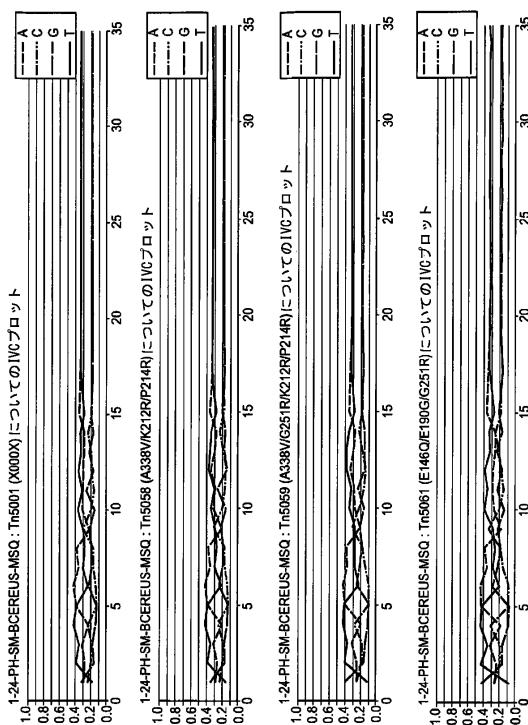
【図5】



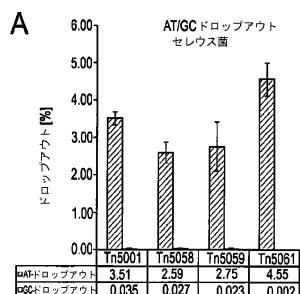
【図6】



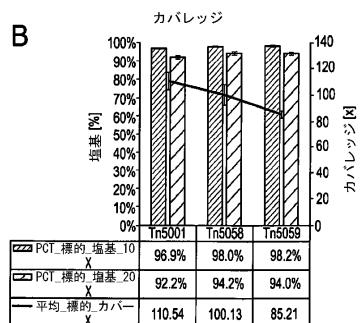
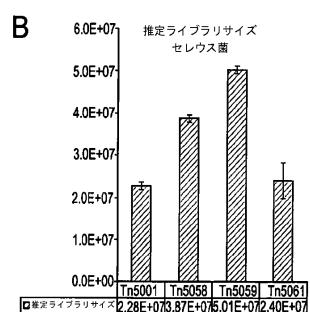
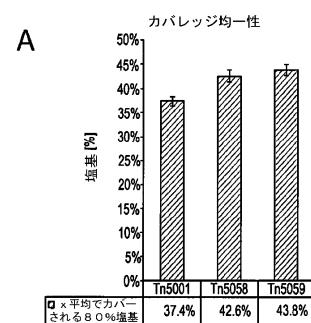
【図7】



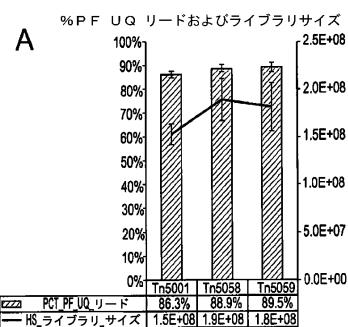
【図8】



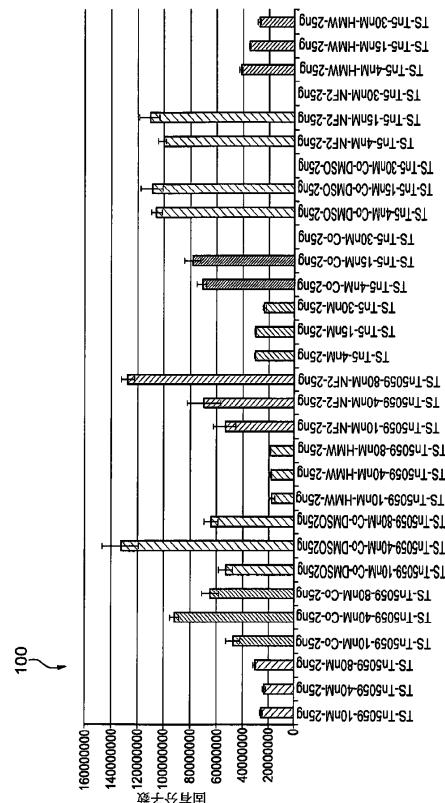
【図9】



【 図 1 0 】



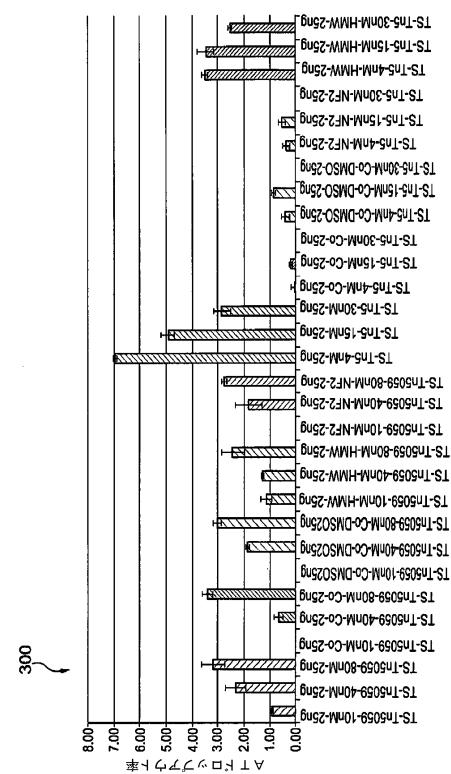
【 図 1 1 】



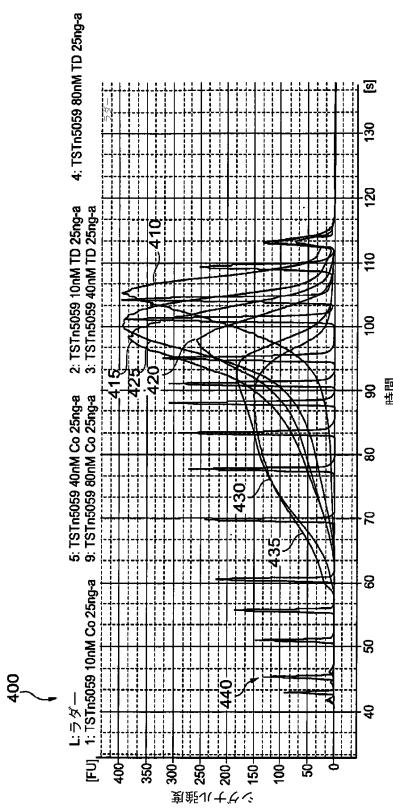
【 义 1 2 】



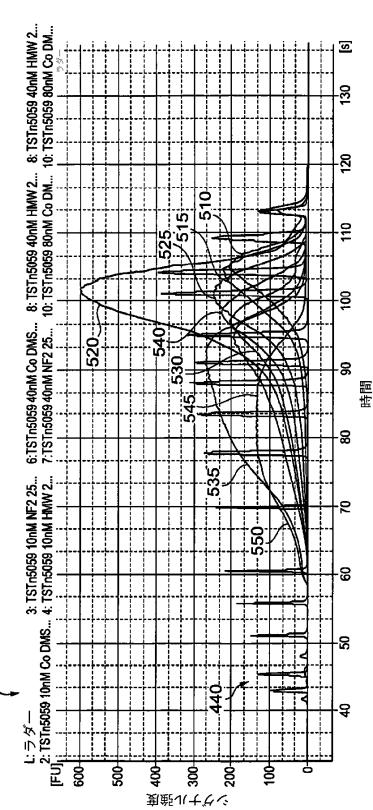
【 13 】



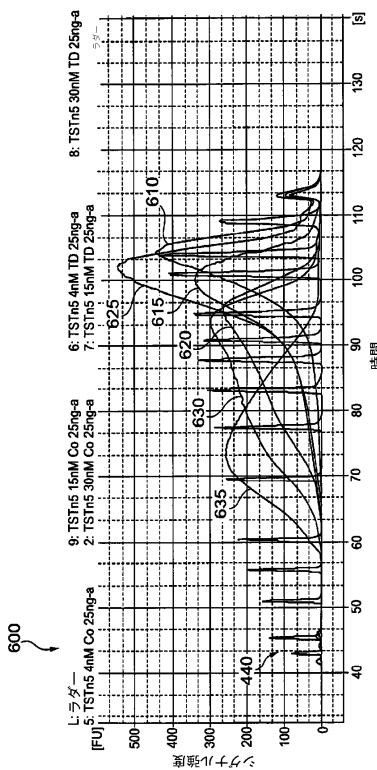
【図14】



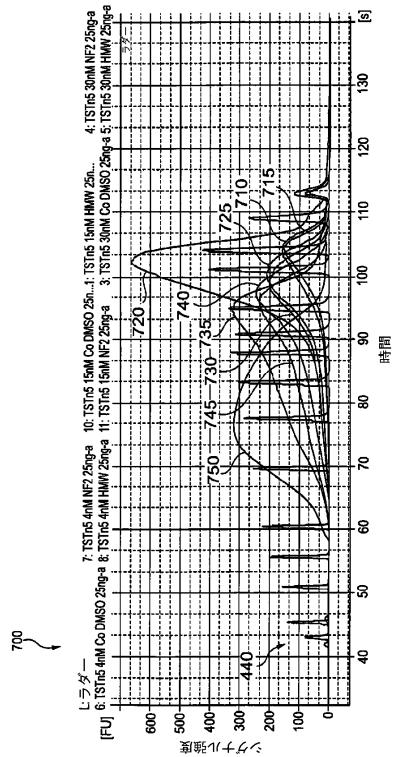
【図15】



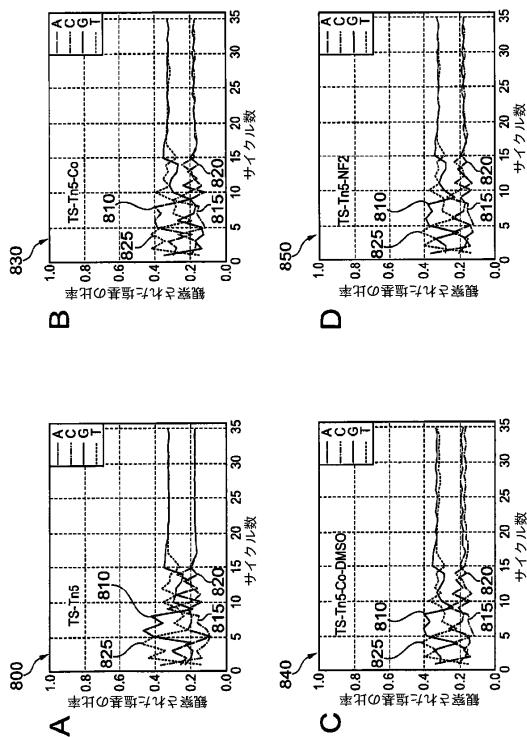
【図16】



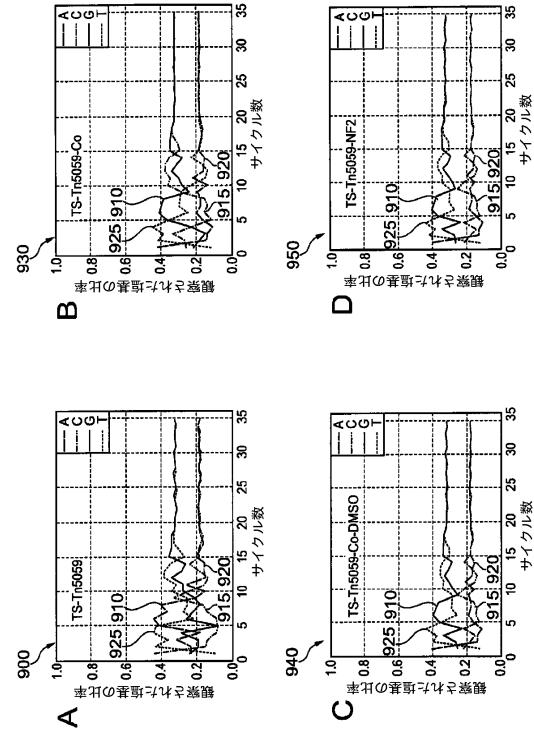
【図17】



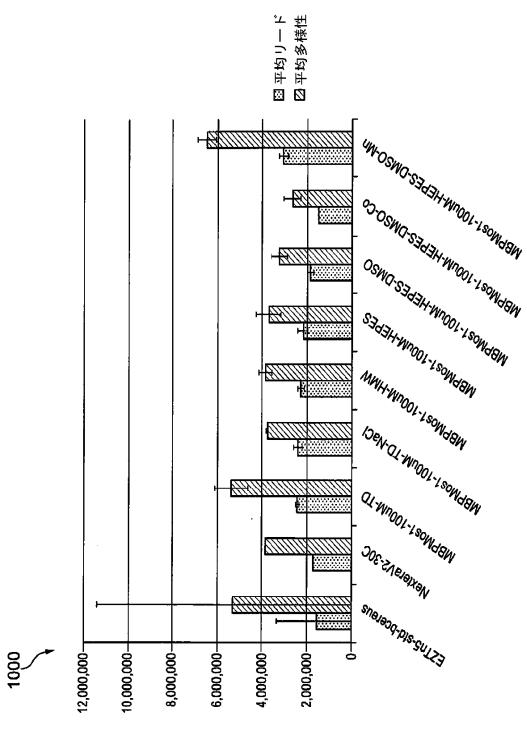
【図18】



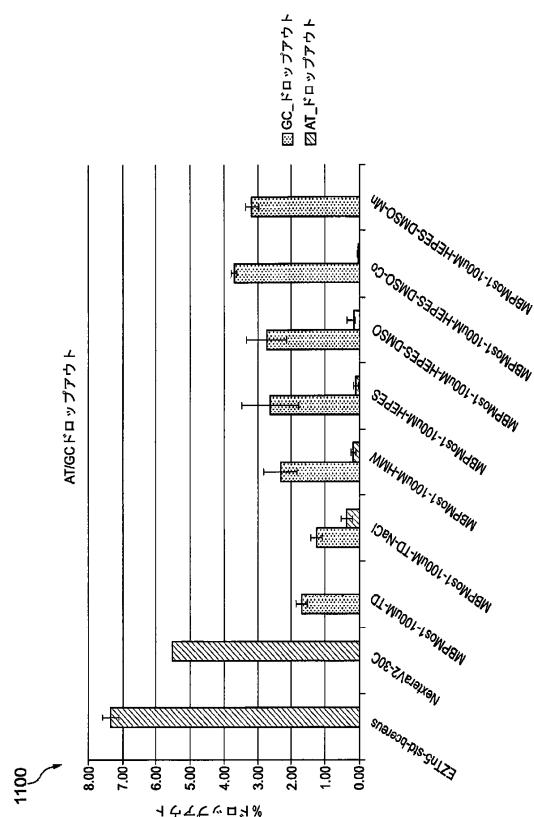
【 図 1 9 】



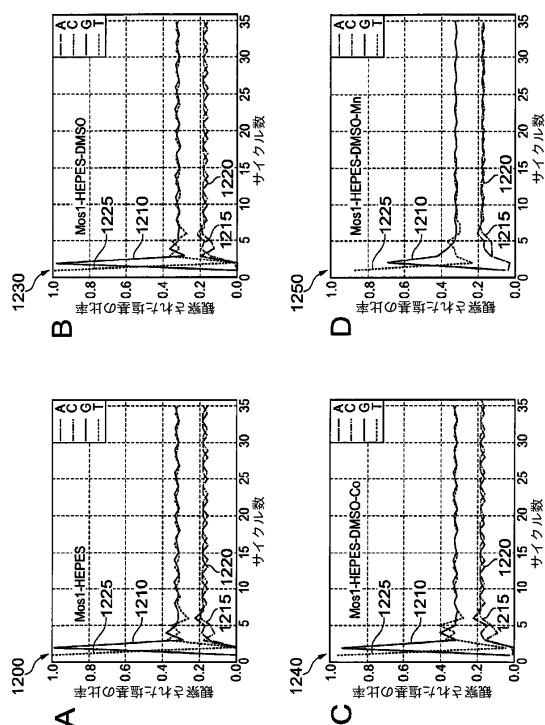
【図20】



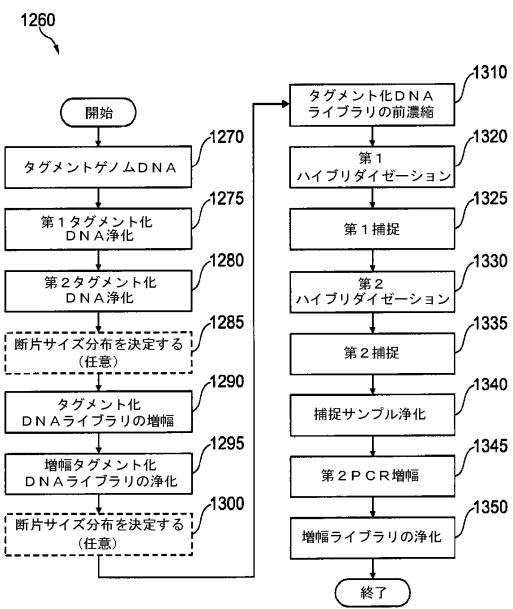
### 【図21】



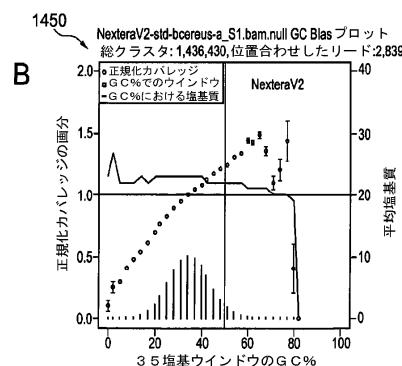
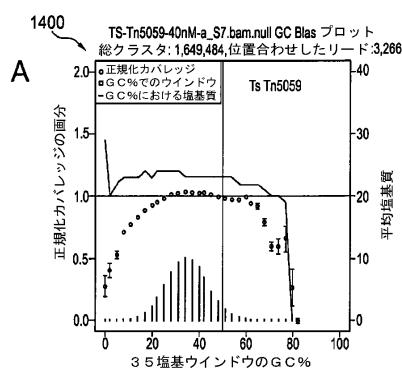
【図22】



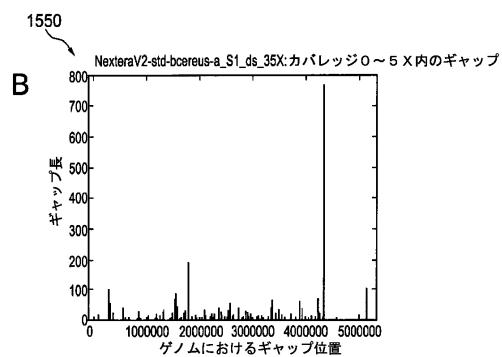
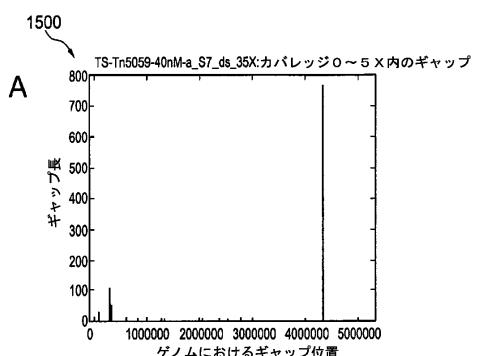
【図23】



【図24】



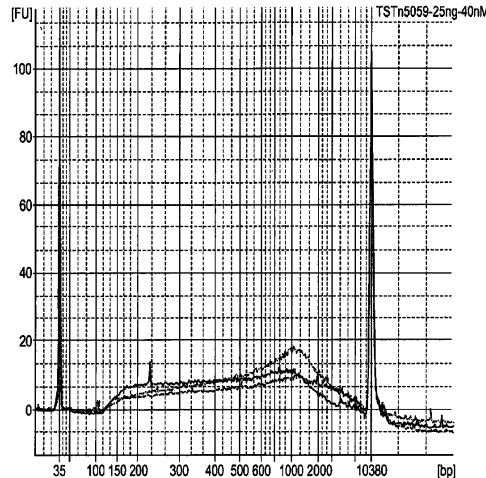
【図25】



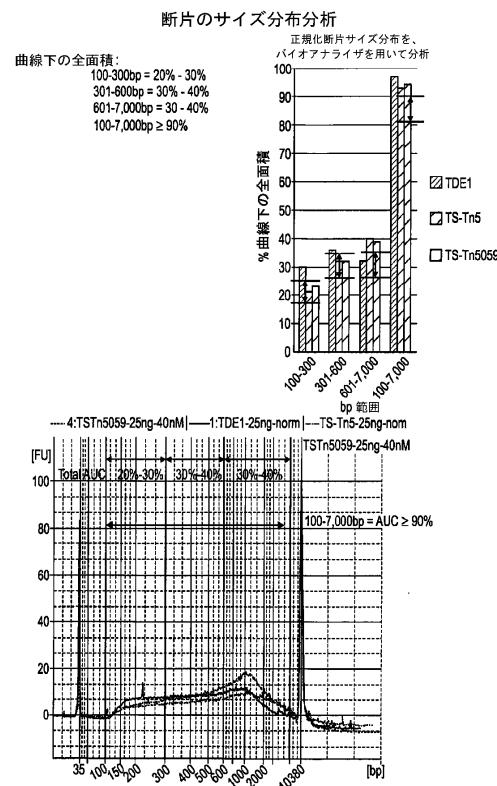
【図26】

40 nMのTS-Tn5059に正規化した、  
TDE1 (Tn5バージョン-1) およびTS-Tn5の、  
25 ngヒトgDNAインプットに対するアクティビティ

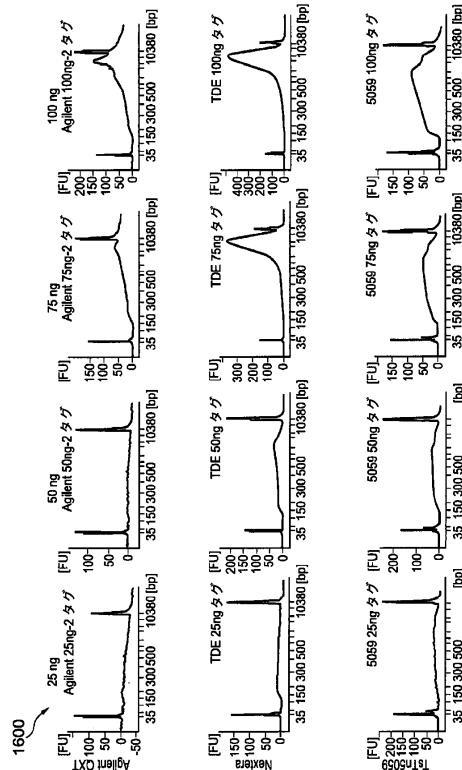
--- 4: TSTn5059-25ng-40nM — 1: TDE1-25ng-norm — 3: TS-Tn5-25ng-norm



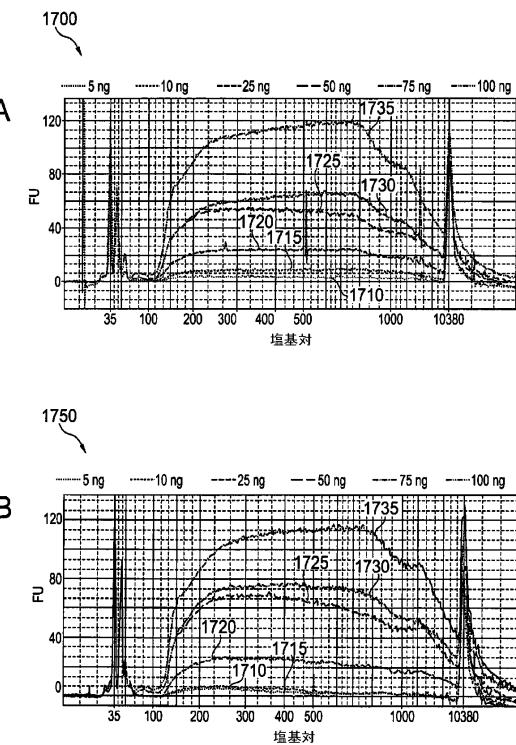
【図27】



【図28】

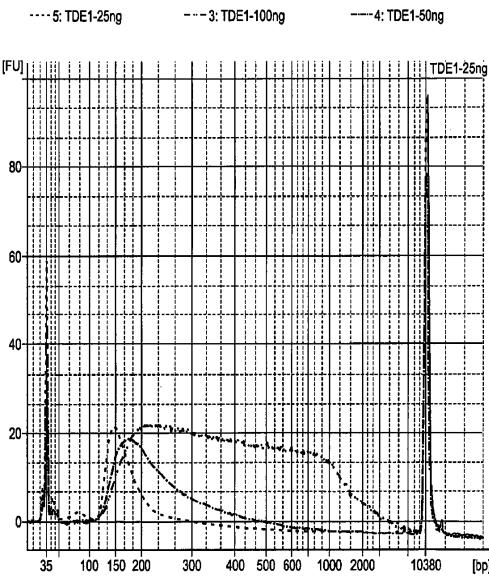


【図29】



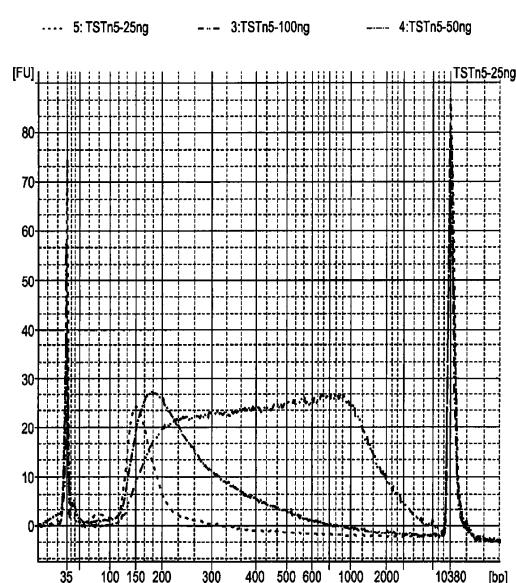
【図30】

25ng～100ngのgDNAを用いた、6x「正規化」濃度のTDE1

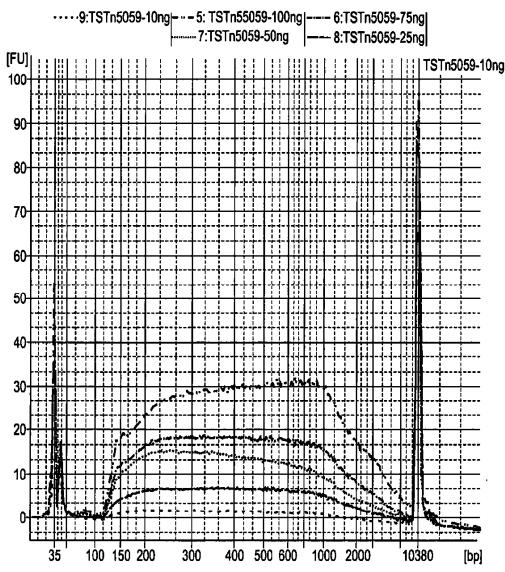


【図31】

25ng～100ngのgDNAを用いた、6x「正規化」濃度のTS-Tn5



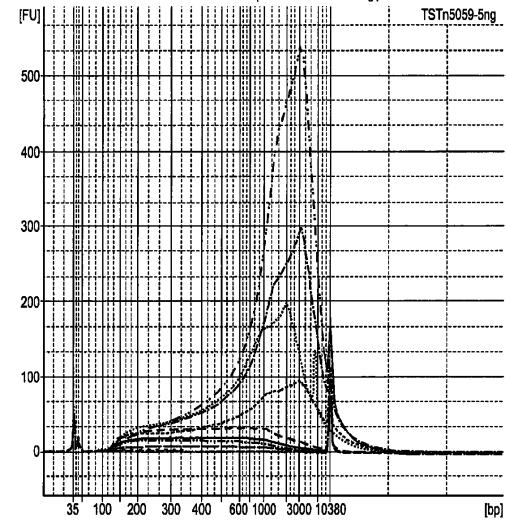
【図32】

10ng～100ngのgDNAを用いた、  
6x「正規化」濃度のTS-Tn5059

【図33】

5ng～500ngのgDNAを用いた、  
6x「正規化」濃度のTS-Tn5059

10: TSTn55059-5ng  
2: TSTn55059-400ng  
4: TSTn55059-200ng  
-1: TSTn55059-500ng  
3: TSTn55059-300ng  
5: TSTn55059-100ng  
6: TSTn55059-75ng  
7: TSTn55059-50ng  
9: TSTn55059-10ng



【配列表】

0006490710000001.app

## フロントページの続き

(51)Int.Cl.	F I
C 12 N 5/10 (2006.01)	C 12 N 5/10
C 12 Q 1/6869 (2018.01)	C 12 Q 1/6869 Z

(31)優先権主張番号 62/080,882

(32)優先日 平成26年11月17日(2014.11.17)

(33)優先権主張国 米国(US)

(72)発明者 アミラリ キア  
アメリカ合衆国 カリフォルニア州 92122 サンディエゴ イラミーナ ウェイ 5200

(72)発明者 エリン ボマティ  
アメリカ合衆国 カリフォルニア州 92122 サンディエゴ イラミーナ ウェイ 5200

(72)発明者 モリー ヘ  
アメリカ合衆国 カリフォルニア州 92122 サンディエゴ イラミーナ ウェイ 5200

(72)発明者 ハイイン リ グリュネンバルト  
アメリカ合衆国 ウィスコンシン州 53719 マディソン リサーチ パーク ブールバード  
5602 スウィート 200

(72)発明者 スコット クエルテン  
アメリカ合衆国 ウィスコンシン州 53719 マディソン リサーチ パーク ブールバード  
5602 スウィート 200

(72)発明者 トリーナ フェイ オソスプラロブ  
アメリカ合衆国 カリフォルニア州 92122 サンディエゴ イラミーナ ウェイ 5200

(72)発明者 ダーリン ハスキンズ  
アメリカ合衆国 ウィスコンシン州 53719 マディソン リサーチ パーク ブールバード  
5602 スウィート 200

(72)発明者 ジョシュア バージェス  
アメリカ合衆国 ウィスコンシン州 53719 マディソン リサーチ パーク ブールバード  
5602 スウィート 200

(72)発明者 アヌバマ カンナ  
アメリカ合衆国 ウィスコンシン州 53719 マディソン リサーチ パーク ブールバード  
5602 スウィート 200

(72)発明者 ダニエル スリングマン  
アメリカ合衆国 ウィスコンシン州 53719 マディソン リサーチ パーク ブールバード  
5602 スウィート 200

(72)発明者 ラメーシュ ヴァイディアナザン  
アメリカ合衆国 ウィスコンシン州 53719 マディソン リサーチ パーク ブールバード  
5602 スウィート 200

審査官 吉森 晃

(56)参考文献 米国特許第7608434(US, B2)  
米国特許第5925545(US, A)  
米国特許第7083980(US, B2)  
国際公開第2004/093645(WO, A2)  
特表2003-505104(JP, A)  
Science, 2000年, Vol.289, p.77-85  
Mol. Microbiol., 2008年, Vol.67, No.3, p.528-540  
JBC, 1998年, Vol.273, No.13, p.7367-7374

JMB, 1997年, Vol.271, p.362-373

FEMS Microbiol. Lett., 2007年, Vol.274, p.132-138

(58)調査した分野(Int.Cl., DB名)

C 12 N 15/00 - 15/90

JST Plus / JMED Plus / JST7580 (JDream III)

Capplus / MEDLINE / EMBASE / BIOSIS (STN)

UniProt / GenSeq