

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第6652512号
(P6652512)

(45) 発行日 令和2年2月26日 (2020.2.26)

(24) 登録日 令和2年1月27日 (2020.1.27)

(51) Int. Cl.	F I
C 1 2 N 15/10 (2006.01)	C 1 2 N 15/10 Z
C 4 O B 50/06 (2006.01)	C 4 O B 50/06
C 1 2 Q 1/6869 (2018.01)	C 1 2 Q 1/6869 Z

請求項の数 35 (全 42 頁)

(21) 出願番号	特願2016-575734 (P2016-575734)	(73) 特許権者	514202402
(86) (22) 出願日	平成27年6月26日 (2015.6.26)		イラミーナ インコーポレーテッド
(65) 公表番号	特表2017-520580 (P2017-520580A)		アメリカ合衆国 カリフォルニア州 92
(43) 公表日	平成29年7月27日 (2017.7.27)		122 サンディエゴ イラミーナ ウェ
(86) 国際出願番号	PCT/US2015/038050		イ 5200
(87) 国際公開番号	W02016/003814	(74) 代理人	100147485
(87) 国際公開日	平成28年1月7日 (2016.1.7)		弁理士 杉村 憲司
審査請求日	平成30年6月22日 (2018.6.22)	(74) 代理人	100181272
(31) 優先権主張番号	62/019,209		弁理士 神 絃一郎
(32) 優先日	平成26年6月30日 (2014.6.30)	(74) 代理人	100193437
(33) 優先権主張国・地域又は機関	米国 (US)		弁理士 高木 義和
		(72) 発明者	フランク ジュー ステーマス
			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

最終頁に続く

(54) 【発明の名称】 片側転移を用いた方法および組成物

(57) 【特許請求の範囲】

【請求項 1】

二本鎖標的核酸からシーケンシングライブラリを調製する方法であって、

(a) 複数のトランスポソームを提供するステップであって、各トランスポソームモノマーはトランスポザーゼとトランスポゾン核酸とを含み、前記トランスポソームは前記二本鎖標的核酸の一方の鎖にのみニックを入れるように構成されている、提供ステップと、

(b) 前記標的核酸をトランスポソームに接触させて、その結果、前記標的核酸の複数の部位で前記標的核酸にニックを入れ、少なくとも1つの前記ニック入り標的核酸の少なくとも1つに単一トランスポゾン核酸を付加して転移核酸を生成することによりシーケンシング用の修飾核酸ライブラリを得る接触ステップと、を含む方法。

10

【請求項 2】

前記方法は、バーコードを有するシーケンシングライブラリを調製する方法であって、

前記提供ステップでは、(a) 各トランスポソームはトランスポザーゼと、認識配列を含むトランスポゾン核酸とを含み、

前記接触ステップは、(b) 前記トランスポゾン核酸を前記標的核酸の鎖に挿入するステップであって、

(i) 前記標的核酸を前記トランスポソームに接触させ、その結果、複数の部位で前記標的核酸にニックを入れ、前記ニック入り鎖のニック部位の一方の側に単一トランスポゾン核酸を付加して転移核酸を生成し、

(i i) 前記ニック入り鎖のニック部位のもう一方の側に付加した前記単一トランス

20

ポゾン核酸をライゲーションすることによりシーケンシング用の修飾核酸ライブラリを得るステップを含む、請求項 1 に記載の方法。

【請求項 3】

標的 DNA の近接情報を得るための方法であって、

(a) 複数のトランスポソームを提供するステップであって、各トランスポソームモノマーはトランスポザーゼとトランスポゾン核酸とを含み、前記トランスポソームは二本鎖標的核酸の一方の鎖にのみニックを入れるように構成されている、ステップと、

(b) 前記標的 DNA を前記トランスポソームに接触させて、その結果、前記標的核酸の複数の部位で前記標的 DNA にニックを入れるステップと、

(c) 前記標的 DNA 配列に 1 つまたは複数の認識配列を追加または挿入して処置済み標的 DNA を生成するステップと、

(d) 前記処置済み標的 DNA をシーケンシングするステップと、

(e) 共通の特性を有する、前記標的 DNA 配列または認識配列を特定することにより近接情報を得るステップと、を含む方法。

【請求項 4】

標的 DNA の近接情報を取得するための方法であって、

(a) 複数のトランスポソームを提供するステップであって、各トランスポソームモノマーはトランスポザーゼと、認識配列を含むトランスポゾン核酸とを含み、前記トランスポソームは二本鎖標的核酸の一方の鎖にのみニックを入れるように構成されている、ステップと、

(b) 前記トランスポゾン核酸を前記標的核酸の鎖に挿入するステップであって、

(i) 前記標的核酸を前記トランスポソームに接触させ、その結果、複数の部位で前記標的核酸にニックを入れ、前記ニック入り鎖のニック部位の一方の側に単一トランスポゾン核酸を付加し、

(i i) 前記ニック入り鎖のニック部位のもう一方の側に前記付加した単一トランスポゾン核酸をライゲーションすることにより修飾核酸を得るステップと、

(c) 前記修飾核酸を増幅することにより、挿入した認識配列を含む複数の核酸を得るステップと、

(d) 処置済み前記標的 DNA をシーケンシングするステップと、

(e) 共通の特性を有する、前記標的 DNA 配列または認識配列を特定することにより近接情報を取得するステップと、を含む方法。

【請求項 5】

前記認識配列はバーコードであり、少なくとも 1 つのトランスポゾン核酸の前記バーコードは異なり、

任意に前記認識配列はバーコードであり前記トランスポゾン核酸の前記バーコードは同じではない、請求項 3 または 4 に記載の方法。

【請求項 6】

修飾核酸を表面上に捕捉するステップをさらに含む、請求項 1 ~ 5 のいずれか一項に記載の方法。

【請求項 7】

前記表面は複数の捕捉プローブを含む、請求項 6 に記載の方法。

【請求項 8】

前記捕捉プローブは核酸を含む、請求項 7 に記載の方法。

【請求項 9】

前記修飾核酸を前記捕捉プローブにハイブリダイズするステップを含み、

任意に前記修飾核酸および前記捕捉プローブはそれぞれ親和性部を含み、

任意に前記親和性部はピオチン、アビジン、およびストレプトアビジンからなる群から選択される、請求項 8 に記載の方法。

【請求項 10】

前記表面は、 mm^2 あたり少なくともおよそ 100,000 の捕捉核酸を含む、請求項 6 ~ 9 のい

10

20

30

40

50

ずれか一項に記載の方法。

【請求項 1 1】

ステップ (b) において標的核酸に接触させたトランスポソームを表面に付加することにより前記修飾核酸を前記表面上に捕捉する、請求項 1 ~ 1 0 のいずれか一項に記載の方法。

【請求項 1 2】

表面上の捕捉核酸をシーケンシングするステップを更に含む、請求項 6 ~ 1 1 のいずれか一項に記載の方法。

【請求項 1 3】

標的核酸配列の線形表現において 2 つの捕捉核酸から得た配列情報の近接度は、前記表面上の捕捉核酸の近接度を示す、請求項 6 ~ 1 2 のいずれか一項に記載の方法。

10

【請求項 1 4】

前記表面上で互いの近接度がより高い捕捉核酸は、近接度のより低い捕捉核酸と比較し、前記標的核酸配列の表現における近接度がより高い配列を含む、請求項 1 3 に記載の方法。

【請求項 1 5】

前記標的核酸配列の表現はハプロタイプ表現を含む、請求項 1 3 および 1 4 のいずれか一項に記載の方法。

【請求項 1 6】

前記トランスポソームは、片側トランスポザゼ活性を含む、請求項 1 ~ 1 5 のいずれか一項に記載の方法。

20

【請求項 1 7】

前記トランスポソームはトランスポザゼ活性を欠くモノマーサブユニットを含む、請求項 1 6 に記載の方法。

【請求項 1 8】

前記トランスポソームは、共有結合的に連結したモノマーサブユニットを含む、請求項 1 6 に記載の方法。

【請求項 1 9】

前記トランスポザゼの四次構造はモノマー状である、請求項 1 6 ~ 1 8 のいずれか一項に記載の方法。

30

【請求項 2 0】

前記トランスポザゼはダイマーを形成する機能を欠く、請求項 1 6 に記載の方法。

【請求項 2 1】

前記トランスポザゼは、Mu、Mu E 3 9 2 Q、T n 5、高活性 T n 5、T n 5 多様体、V i b h a r、R A G、および T n 5 5 2 からなる群から選択される、請求項 1 ~ 2 0 のいずれか一項に記載の方法。

【請求項 2 2】

前記 1 つまたは複数のトランスポゾン核酸は非機能的である、請求項 1 ~ 2 1 のいずれか一項に記載の方法。

【請求項 2 3】

前記非機能的トランスポゾン核酸の 3' 末端は、ジデオキシ基、スパーサ基、アミン基、アルキル基、アリール基、ホスフェート基、チオール基、逆ヌクレオチド、アジド基、硫酸基、およびピオチン基からなる群から選択される、請求項 2 2 に記載の方法。

40

【請求項 2 4】

前記複数のトランスポソームは、前記トランスポザゼを、機能的トランスポゾン核酸と非機能的トランスポゾン核酸とに接触させることにより調製する、請求項 2 2 に記載の方法。

【請求項 2 5】

非機能的トランスポゾン核酸を含むトランスポゾン核酸の、機能的トランスポゾン核酸に対する比率は 1 : 1 以上、例えば 1 0 : 1 以上である、請求項 2 2 に記載の方法。

50

【請求項 26】

(i) 前記ニック入り標的核酸にDNAポリメラーゼを提供するステップと、
 (ii) 前記捕捉鎖を鋳型として用いて前記標的核酸の3'を伸長させるステップと、
 (iii) 任意に前記伸長した核酸を増幅するステップと、をさらに含む、請求項3に記載の方法。

【請求項 27】

前記伸長した核酸を増幅するステップは、アンカー部位、シーケンシングプライマー部位、増幅プライマー部位、バーコード、およびレポータータグからなる群から選択される配列を含むテイル化増幅プライマーを用いる、請求項26に記載の方法。

【請求項 28】

捕捉核酸を増幅するステップを含み、
 任意に前記捕捉核酸を増幅するステップはブリッジ増幅を有する、請求項6～27のいずれか一項に記載の方法。

【請求項 29】

表面は複数の捕捉プローブを含み、
 任意に前記捕捉プローブは核酸を含む、請求項6～28のいずれか一項に記載の方法。

【請求項 30】

前記トランスポゾン核酸は、アンカー部位、バーコード、シーケンシングプライマー部位、増幅プライマー部位、およびレポータータグからなる群から選択される配列を含む、請求項1～29のいずれか一項に記載の方法。

【請求項 31】

少なくとも1つのトランスポソームが2つのトランスポゾン核酸を含み、
 任意に前記2つのトランスポゾン核酸は異なる配列を含む、請求項1～30のいずれか一項に記載の方法。

【請求項 32】

前記複数のトランスポソームは、少なくとも2つの異なるトランスポゾン核酸を含む、請求項1～31のいずれか一項に記載の方法。

【請求項 33】

前記複数のトランスポソームは、トランスポソームを形成する機能は有するが転移する機能は欠く、少なくとも1つのトランスポザーゼを含む、請求項1～32のいずれか一項に記載の方法。

【請求項 34】

前記標的核酸は、ゲノムDNA、ゲノムDNA断片、およびcDNAからなる群から選択される、請求項1～33のいずれか一項に記載の方法。

【請求項 35】

前記トランスポゾンは第1バーコードセットを含み、
 前記第1バーコードセットを転移中に前記標的核酸に導入して、第1バーコードセットを含む転移標的核酸を生成し、
 前記転移標的核酸をプールして、転移標的核酸の第1プールを生成し、
 第2バーコードセットを前記転移標的核酸の第1プールに導入して、第1および第2のバーコードセットを含む標的核酸を生成し、
 第1および第2のバーコードセットを含む標的核酸をプールして転移標的核酸の第2プールを生成し、
 任意に追加バーコードを導入するステップと、プールしてバーコード化標的核酸ライブラリを生成するステップを繰り返す、
 組み合わせバーコーディングをさらに含む請求項1～34のいずれか一項に記載の方法。

【発明の詳細な説明】

【関連出願】

【0001】

本出願は、2014年6月30日に出願された米国仮特許出願公開第62/019209号の優先権を主張し、これはその全体が参照により本明細書に組み込まれる。

【技術分野】

【0002】

本明細書で提供する実施形態は、次世代シーケンシングのための方法および組成物に関する。一部の実施形態には、片(one sided)転移ともいう片側(one-sided)転移を用いて標的核酸から鋳型ライブラリを調製するステップと、鋳型ライブラリをシーケンシングするステップと、近接情報を得るステップを含む。

【背景技術】

【0003】

いくつかの次世代シーケンシング技術を、ゲノムの全配列を早く経済的に決定するために利用可能である。典型的には、鋳型核酸のライブラリをシーケンシングの前に標的ゲノムDNAサンプルから調製する。サンプルの調製には通常、より大きいDNA鎖を切断して次世代シーケンシング技術により適するより小さいDNA断片にする、DNA断片化ステップが含まれる。しばしばアダプターをDNA断片の末端に付加するが、これはアダプターライゲーション後のDNA末端修復か、または、つい最近ではトランスポソーム系(transposome system)を用いることにより達成することが可能である。トランスポザゼとトランスポゾン核酸との複合体であるトランスポソームの使用によりゲノムの断片化および断片のアダプターライゲーションが同時に可能になり、それによりライブラリ調製を単純化することができる。しかしながら、ゲノムDNAの断片化は、近接度、フェージング(phasing)、およびハプロタイプに関する個々の核酸分子の情報の損失を招き得る。そのため、代替のライブラリ調製方法に対する必要性が存在する。

【発明の概要】

【0004】

本明細書の一部の実施形態に記載するのは、片側転移のための方法である。本出願の発明者らは、驚くべきことに、片側転移を行うことにより二本鎖標的DNAの一方の鎖にのみニックが入り、そのような転移の後の標的DNAはトランスポソームを取り除いた後でも無傷のままであることを発見した。従って、標的DNAの近接度は転移事象の後でも維持される。一部の実施形態では、近接情報を得るために片側転移のための方法を用いることが可能である。一部の実施形態では、シーケンシングライブラリを調製するために片側転移のための方法を用いることが可能である。一部の実施形態では、フェージング情報またはハプロタイプ情報を確定するために片側転移のための方法を用いることが可能である。

【0005】

一部の実施形態では、二本鎖標的DNAの一方の鎖のみにニックを入れ、トランスポソームモノマーのトランスポゾンの移動鎖1つのみをそのニック入り標的DNAに移動させるようにトランスポソームダイマーを構成する。一部の実施形態では、トランスポソームダイマーの1つのモノマーユニットは、片側転移を引き起こす転移はできない。一部の実施形態では、トランスポソームダイマーの1つのトランスポザゼはトランスポゾンに結合することによりトランスポソーム複合体を形成することができるが、標的DNAにニックを入れることはできない。

【0006】

一部の実施形態ではトランスポゾンは機能的であり、その結果、トランスポゾンをトランスポザゼに接触させることによりトランスポソームを形成し、トランスポゾン配列を標的核酸に移動させることが可能である。一部の実施形態ではトランスポゾンは非機能的であり、その結果、トランスポゾンをトランスポザゼに接触させることによりトランスポソームは形成されるが、トランスポゾン配列を標的核酸に移動することはできない。一部の実施形態では、移動鎖の3'-末端は、標的核酸の5'-末端に対する求核攻撃が不可能な3'-終端核酸を含む。一部の実施形態では、認識配列の3'-末端はブロックされる。一部の実施形態では、ブロック化認識配列の3'-末端は、3'-終端ジデオキシヌクレオチド、アミ

10

20

30

40

50

ノ基、アルキル基、アリール基、チオール基、硫酸基、逆ヌクレオチド、アジド基、またはピオチンを含む。

【0007】

一部の実施形態では、トランスポザーゼはトランスポソームを形成することはできるが、標的DNAにニックを入れることはできない。一部の実施形態では、トランスポザーゼは1つまたは複数のアミノ酸修飾を含み、その結果、それはトランスポソームを形成することはできるが、標的DNAにニックを入れることはできない。

【0008】

一部の実施形態では、トランスポソームが効率的にダイマーを形成することができないようにトランスポソーム複合体を構成する。一部の実施形態では、トランスポソームが全くダイマーを形成することができないように、トランスポソーム複合体を構成する。一部の実施形態では、トランスポソームモノマーは二本鎖DNAの一方の鎖においてのみニックを形成し、トランスポソンの移動鎖をニック入り標的DNAに移動させる。

【0009】

一部の実施形態では、標的DNAの二本の鎖の転移に対する抵抗差(differential resistance)を活用することにより、片側転移を行う。一部の実施形態では、標的DNAの一方の鎖は、転移に対し抵抗性を持つ修飾塩基または修飾ホスホジエステル結合を含む。標的DNAの二本の鎖で転移に対して抵抗差がある標的DNAを、トランスポソームにさらすことで片側転移を引き起こす。一部の実施形態では、標的核酸は、cDNAの一方の鎖が修飾塩基および/または修飾ホスホジエステル結合を含み、その結果、その鎖が転移に対し抵抗性を持つ、二本鎖cDNAである。一部の実施形態では、標的核酸は、一方の鎖が部分的または全体的に転移に対し抵抗性を持つように修飾されている二本鎖ゲノムDNAである。

【0010】

例示的な片側転移スキームを図14に示す。一部の実施形態では、一本鎖核酸鋳型で開始し(実線)、元の鋳型鎖よりも転移に対し抵抗差を持つ相補鎖(点線)を合成する。すると、通常のトランスポソン複合体(例えば、活性で、ブロックされていない)を用いても片側転移が起きる。一例では、新たに合成した鎖は転移に対しより高い抵抗性を持つ。別の例では、元の鋳型が高い抵抗性を持ち、合成された鎖はそれ自身への転移が可能である。この実施形態では、抵抗性のより低い鎖がライブラリ要素を形成し、これはより抵抗性の高い鎖に近接して保持される。

【0011】

出願人は驚くべきことに、片側転移を実行した後、トランスポソームのトランスポザーゼを取り除いた後でさえ、二本鎖標的核酸が近接情報を失わずに無傷のままであることを発見した。一部の実施形態では、SDS、尿素、プロテアーゼ、または熱で処置することにより、転移後に、転移した標的核酸からトランスポザーゼを取り除く。従って、片側転移は、配列情報、近接情報、フェージング情報、およびハプロタイプ情報を確定するうえで有利であり得る。近接情報は広範囲にわたるハプロタイプ分析能を提供し得る。ハプロタイピングは、稀なアレルと、遺伝子再配列、遺伝子重複などの構造的多様体のフェージングを可能にする。

【0012】

一部の実施形態では、片側転移は、片側転移を介して第1バーコードセットを付加し、第2バーコードセットを後続の増幅により付加する、組み合わせバーコーディングと連動させることが可能である。

【0013】

一部の実施形態では、第1バーコードセットを転移中に標的核酸に導入し、第1バーコードセットを含む転移標的核酸を生成する。転移標的核酸をプールして転移標的核酸の第1プールを生成する。第2バーコードセットを転移標的核酸の第1プールに導入して、第1および第2のバーコードセットを含む標的核酸を生成する。第2バーコードセットは、後続の増幅、ライゲーション、または追加の転移のいずれかにより導入することができる

10

20

30

40

50

。一部の実施形態では、第1バーコードセットと第2バーコードセットは異なる。第1および第2のバーコードセットを含む標的核酸をプールして転移標的核酸の第2プールを生成する。オプションとして、追加バーコードを導入するステップと、プールしてバーコード化標的核酸ライブラリを生成するステップは繰り返すことができる。

【0014】

一部の実施形態では、単一細胞に由来する核酸の配列情報または近接情報を確定するために片側転移を用いることが可能である。核酸は単一細胞のmRNAから生成した、ゲノム核酸またはcDNAとすることが可能である。一部の実施形態では、第1バーコードセットを単一細胞に由来する核酸に導入することができ、これは単一細胞の識別子として機能する。一部の実施形態では、第1バーコードセットを単一細胞に由来する核酸に導入した後、バーコード化核酸をプールし、それをさらに、後続の増幅、ライゲーション、または、追加バーコードの導入を含むもしくは含まない追加の転移により処理することが可能である。

【0015】

本明細書に提供する方法および組成物の一部の実施形態には、二本鎖標的核酸からシーケンシングライブラリを調製する方法であって、(a)複数のトランスポソームを提供するステップであって、各トランスポソームはトランスポザゼとトランスポゾン核酸とを含み、前記トランスポソームは前記標的核酸の一方の鎖にのみニックを入れてトランスポゾンを転移させるように構成されている、ステップと、(b)前記標的核酸を前記トランスポソームに接触させて、その結果、前記標的核酸の複数の部位で前記標的核酸にニックを入れ、前記ニック入り標的核酸にトランスポゾン核酸を付加することによりシーケンシング用の修飾核酸ライブラリを得るステップと、を含む方法を含む。

【0016】

一部の実施形態は、二本鎖標的核酸からシーケンシングライブラリを調製する方法であって、(a)複数のトランスポソームを提供するステップであって、各トランスポソームはトランスポザゼとトランスポゾン核酸とを含み、前記トランスポソームは前記標的核酸の一方の鎖にのみニックを入れてトランスポゾンを転移させるように構成されている、ステップと、(b)前記標的核酸をトランスポソームに接触させて、その結果、前記標的核酸の複数の部位で前記標的核酸にニックを入れ、前記ニック入り核酸にトランスポゾン核酸を付加するステップと、(c)前記トランスポゾン核酸にプライマーをハイブリダイズし、前記ハイブリダイズしたプライマーを伸長することによりシーケンシング用の修飾核酸ライブラリを得るステップと、を含む方法を含む。片側転移を用いたライブラリ調製の例示的なスキームを図13に示す。

【0017】

一部の実施形態は、標的DNAの近接情報を取得するための方法を含む。該方法は、(a)複数のトランスポソームを提供するステップであって、各トランスポソームモノマーはトランスポザゼとトランスポゾン核酸とを含み、前記トランスポソームは二本鎖標的核酸の一方の鎖にのみニックを入れるように構成されている、ステップと、(b)前記標的DNAを前記トランスポソームに接触させて、その結果、前記標的核酸の複数の部位で前記標的DNAにニックを入れるステップと、(c)前記標的DNA配列に1つまたは複数の認識配列を追加または挿入して処置済み標的DNAを生成するステップと、(d)前記処置済み標的DNAをシーケンシングするステップと、(e)共通の特性を有する、前記標的DNA配列または認識配列を特定することにより近接情報を取得するステップと、を含む。

【0018】

一部の実施形態には、標的DNAの近接情報を取得するための方法を含む。該方法は、(a)複数のトランスポソームを提供するステップであって、各トランスポソームモノマーはトランスポザゼと、認識配列を含むトランスポゾン核酸とを含み、前記トランスポソームは二本鎖標的核酸の一方の鎖にのみニックを入れるように構成されている、ステップと、(b)前記トランスポゾン核酸を前記標的核酸の鎖に挿入するステップであって、

(i) 前記標的核酸を前記トランスポソームに接触させ、その結果、複数の部位で前記標的核酸にニックを入れ、前記ニック入り鎖のニック部位の一方の側に単一トランスポゾン核酸を付加し、(i i) 前記ニック入り鎖のニック部位のもう一方の側に前記付加した単一トランスポゾン核酸をライゲーションすることにより修飾核酸を得るステップと、(c) 前記修飾核酸を増幅することにより、挿入した認識配列を含む複数の核酸を得るステップと、(d) 前記処置済み標的DNAをシーケンシングするステップと、(e) 共通の特性を有する、前記標的DNA配列または認識配列を特定することにより近接情報を取得するステップとを含む。

【 0 0 1 9 】

一部の実施形態は、修飾核酸を表面上に捕捉するステップも含む。

10

【 0 0 2 0 】

一部の実施形態では、ステップ(b)で標的核酸に接触させたトランスポソームを表面に付加することにより、修飾核酸を表面上に捕捉する。

【 0 0 2 1 】

一部の実施形態は、捕捉した核酸を表面上でシーケンシングするステップも含む。

【 0 0 2 2 】

一部の実施形態では、2つの捕捉核酸から得た配列情報の、標的核酸配列の線形表現における近接度は、表面上の捕捉核酸の近接度を示唆する。

【 0 0 2 3 】

一部の実施形態では、表面上で互いの近接度がより高い捕捉核酸は、近接度のより低い捕捉核酸と比較し、標的核酸配列の表現における近接度がより高い配列を含む。

20

【 0 0 2 4 】

一部の実施形態では、標的核酸配列の表現にはハプロタイプ表現が含まれる。一部の実施形態では、標的核酸配列の表現には規則正しい短リードが含まれる。

【 0 0 2 5 】

一部の実施形態では、トランスポザーゼは片側トランスポザーゼ活性を含む。

【 0 0 2 6 】

一部の実施形態では、トランスポザーゼはトランスポザーゼ活性を欠いたモノマーサブユニットを含む。一部の実施形態では、トランスポザーゼは共有結合的に連結したモノマーサブユニットを含む。一部の実施形態では、トランスポザーゼの四次構造はモノマー状である。一部の実施形態では、トランスポザーゼはダイマーを形成する機能を欠く。

30

【 0 0 2 7 】

一部の実施形態では、トランスポザーゼは、Mu、Mu E 3 9 2 Q、T n 5、高活性T n 5 (Goryshin and Reznikoff, J. Biol. Chem., 273:7367 (1998))、E Z - T n 5 (商標)トランスポザーゼ(ウィスコンシン州マディソンのE p i c e n t r e B i o t e c h n o l o g i e s社)、T n 5の多様体、R A G、T n 7、T n 1 0、V i b h a rトランスポザーゼ、およびT n 5 5 2からなる群から選択する。アミノ酸の置換、挿入、欠失、および/または他のタンパク質もしくはペプチドとの融合があるようなT n 5トランスポザーゼの多様体は、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、同第7 6 0 8 4 3 4号明細書、および米国特許出願公開第1 4 / 6 8 6 9 6 1号明細書に開示されている。前記特許および特許出願公開はその全体が参照により本明細書に組み込まれる。一部の実施形態では、T n 5トランスポザーゼは、米国特許出願公開第1 4 / 6 8 6 9 6 1号明細書に開示されているように、野生型タンパク質に対し、54、56、372、212、214、251、および338の位置で1つまたは複数の置換を含む。一部の実施形態では、T n 5野生型タンパク質またはその多様体は、融合ポリペプチドをさらに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T sを含むことが可能である。本段落で言及した参考文献はそれぞれ、その全体が参照により本明細書に組み込まれる。

40

【 0 0 2 8 】

一部の実施形態では、トランスポゾン核酸はブロックされている。一部の実施形態では

50

、トランスポゾンの移動鎖の3'-末端はブロックされている。一部の実施形態では、ブロック化トランスポゾン核酸の3'-末端は、ジデオキシ基、スパーサ基、アミン基、アジド基、ホスフェート基、アルキル基、逆ヌクレオチド、およびピオチン基からなる群から選択される。一部の実施形態では、トランスポゾン配列は、塩基の置換、追加、または該トランスポゾン配列からの塩基の欠失により変えることが可能である。

【0029】

一部の実施形態では、トランスポザーゼを機能的トランスポゾン核酸および非機能的トランスポゾン核酸に接触させることにより、複数のトランスポソームを調製する。一部の実施形態では、非機能的トランスポゾンとしては、ブロック化トランスポゾンが挙げられる。一部の実施形態では、非機能的トランスポゾン核酸を含むトランスポゾン核酸の、機能的トランスポゾン核酸に対する比率は、1:1以上である。一部の実施形態では、非機能的トランスポゾン核酸を含むトランスポゾン核酸の、機能的トランスポゾン核酸に対する比率は、1:2、1:3、1:5、1:10、1:20、1:30、1:40、1:50、1:75、1:100、2:1、3:1、4:1、5:1、6:1、7:1、8:1、9:1、10:1、20:1、30:1、40:1、50:1、60:1、70:1、80:1、90:1、または100:1とすることが可能である。

【0030】

一部の実施形態には、伸長した核酸を増幅するステップも含む。一部の実施形態では、伸長した核酸の増幅には、アンカー部位、シーケンシングプライマー部位、増幅プライマー部位、およびレポータータグからなる群から選択される配列を含む、テイル化 (tailed) 増幅プライマーを用いる。

【0031】

一部の実施形態には、捕捉した核酸を増幅するステップも含む。一部の実施形態では、捕捉した核酸の増幅にはブリッジ増幅を含む。

【0032】

一部の実施形態では、表面は複数の捕捉プローブを含む。一部の実施形態では、捕捉プローブは核酸を含む。一部の実施形態では、修飾核酸を捕捉プローブにハイブリダイズするステップも含む。

【0033】

一部の実施形態では、修飾核酸と捕捉プローブはそれぞれ親和性部を含む。一部の実施形態では、親和性部は結合対のメンバとすることが可能である。場合によっては、修飾核酸は結合対の第1メンバを含み、捕捉プローブは結合対の第2メンバを含むことができる。場合によっては、捕捉プローブは固体表面に固定することができ、修飾核酸は結合対の第1メンバを含むことができ、捕捉プローブは結合対の第2メンバを含むことができる。このような場合、結合対の第1メンバおよび第2メンバを結合させることで、修飾核酸は固体表面に固定される。結合対の例としては、ピオチン - アビジン、ピオチン - ストレプトアビジン、ピオチン - ニュートラアビジン、リガンド - 受容体、ホルモン - 受容体、レクチン - 糖タンパク質、および抗原 - 抗体が挙げられるが、これらに限定されない。

【0034】

一部の実施形態は、修飾核酸の親和性部と捕捉プローブの親和性部を結合するステップも含む。

【0035】

一部の実施形態では、トランスポゾン核酸は、アンカー部位、バーコード、シーケンシングプライマー部位、増幅プライマー部位、ユニークな分子インデックス、およびレポータータグからなる群から選択される配列を含む。

【0036】

一部の実施形態では、少なくとも1つのトランスポソームが2つのトランスポゾン核酸を含む。

【0037】

一部の実施形態では、2つのトランスポゾン核酸は異なる配列を有する。

【0038】

一部の実施形態では、複数のトランスポソームは少なくとも2つの異なるトランスポゾン核酸を含む。

【0039】

一部の実施形態では、標的核酸は、DNAおよびRNAからなる群から選択される。一部の実施形態では、標的核酸は、ゲノムDNAおよびcDNAからなる群から選択される。一部の実施形態では、標的核酸はゲノムDNAである。

【0040】

一部の実施形態では、表面は、ビーズ、スライド、フローセル、チャネル、ディップスティック、およびウェルからなる群から選択される基材上にある。

【0041】

一部の実施形態では、表面は mm^2 あたり少なくともおよそ10,000の捕捉核酸を含む。一部の実施形態では、表面は mm^2 あたり少なくともおよそ100,000の捕捉核酸を含む。一部の実施形態では、表面は mm^2 あたり少なくともおよそ、1,000,000、1,500,000、2,000,000、3,000,000、5,000,000、10,000,000、15,000,000、20,000,000、30,000,000、40,000,000、50,000,000、60,000,000、70,000,000、80,000,000、90,000,000、100,000,000、150,000,000、200,000,000、300,000,000、350,000,000、400,000,000、450,000,000、500,000,000、550,000,000、600,000,000、650,000,000、700,000,000、750,000,000、800,000,000、850,000,000、900,000,000、950,000,000、1000,000,000、1200,000,000、1300,000,000、1400,000,000、1500,000,000、1600,000,000、1700,000,000、1800,000,000、1900,000,000、2000,000,000、3000,000,000、4000,000,000、5000,000,000、6000,000,000、7000,000,000、8000,000,000、9000,000,000、10,000,000,000以上の捕捉核酸を含む。

【0042】

一部の実施形態は、前記方法のうち任意の1つにより調製したシーケンシングライブラリを含む。

【0043】

本明細書で提供する方法および組成物の一部の実施形態は、二本鎖標的核酸に由来するバーコードを有するシーケンシングライブラリを調製する方法であって、(a)複数のトランスポソームを提供するステップであって、各トランスポソームはトランスポザーゼと、バーコードを含むトランスポゾン核酸とを含む、ステップと、(b)前記トランスポゾン核酸を前記標的核酸の鎖に挿入するステップであって、(i)前記標的核酸を前記トランスポソームに接触させ、その結果、複数の部位で前記標的核酸にニックを入れ、前記ニック入り鎖のニック部位の一方の側に単一トランスポゾン核酸を付加し、(ii)前記ニック入り鎖のニック部位のもう一方の側に前記付加した単一トランスポゾン核酸をライゲーションすることにより修飾核酸を得るステップと、を含む方法を含む。

【0044】

一部の実施形態は、(c)修飾標的核酸を表面上に捕捉するステップも含む。

【0045】

一部の実施形態は、二本鎖標的核酸に由来するバーコードを有するシーケンシングライブラリを調製する方法であって、(a)複数のトランスポソームを提供するステップであって、各トランスポソームはトランスポザーゼと、バーコードを含むトランスポゾン核酸とを含む、ステップと、(b)前記トランスポゾン核酸を前記標的核酸の鎖に挿入するステップであって、(i)前記標的核酸を前記トランスポソームに接触させ、その結果、複数の部位で前記標的核酸にニックを入れ、前記ニック入り鎖のニック部位の一方の側に単一トランスポゾン核酸を付加し、(ii)前記ニック入り鎖のニック部位のもう一方の側に前記付加した単一トランスポゾン核酸をライゲーションすることにより修飾核酸を得るステップと、(c)前記修飾核酸を増幅することにより、挿入バーコードを含む複数の核酸を得るステップと、を含む方法を含む。

【0046】

一部の実施形態では、修飾標的核酸を表面上に捕捉するステップも含む。

【0047】

一部の実施形態では、ステップ (b) で標的核酸に接触させたトランスポソームを表面に付加することにより修飾核酸を表面上に捕捉する。

【 0 0 4 8 】

一部の実施形態は、捕捉した核酸をシーケンシングするステップも含む。

【 0 0 4 9 】

一部の実施形態では、2つの捕捉核酸から得た配列情報の、標的核酸配列の線形表現の近接度は、表面上の捕捉核酸の近接度を示唆する。

【 0 0 5 0 】

一部の実施形態では、表面上で互いの近接度がより高い捕捉核酸は、近接度のより低い捕捉核酸と比較し、標的核酸配列の表現における近接度がより高い配列を含む。

10

【 0 0 5 1 】

一部の実施形態では、標的核酸配列の表現にはハプロタイプ表現が含まれる。

【 0 0 5 2 】

一部の実施形態では、少なくとも1つのトランスポゾン核酸のバーコードは異なる。

【 0 0 5 3 】

一部の実施形態では、トランスポゾン核酸のバーコードは同じではない。

【 0 0 5 4 】

一部の実施形態は、配列中で共通するバーコードの存在に従って核酸配列をアライメントして、標的核酸表現を生成するステップも含む。

【 0 0 5 5 】

20

一部の実施形態では、トランスポザーゼは片側トランスポザーゼ活性を含む。

【 0 0 5 6 】

一部の実施形態では、トランスポザーゼはトランスポザーゼ活性を欠いたモノマーサブユニットを含む。

【 0 0 5 7 】

一部の実施形態では、トランスポザーゼは共有結合的に連結したモノマーサブユニットを含む。一部の実施形態では、トランスポザーゼの四次構造はモノマー状である。一部の実施形態では、トランスポザーゼはダイマーを形成する機能を欠く。

【 0 0 5 8 】

一部の実施形態では、トランスポザーゼは、Mu、Mu E 3 9 2 Q、T n 5、高活性 T n 5、E Z - T n 5 (商 標)、T n 5 の多様体、R A G、T n 7、T n 1 0、T n 5 5 2、および V i b h a r トランスポザーゼからなる群から選択する。

30

【 0 0 5 9 】

一部の実施形態では、トランスポゾン核酸はブロックされている。

【 0 0 6 0 】

一部の実施形態では、ブロック化トランスポゾン核酸の3'-末端は、ジデオキシ基、スパーサ基、アミノ基、アジド基、アルキル基、アリール基、逆ヌクレオチド、チオホスフェート基、およびピオチン基からなる群から選択される。

【 0 0 6 1 】

一部の実施形態では、トランスポザーゼを、非機能的トランスポゾン核酸と機能的トランスポゾン核酸とに接触させることにより複数のトランスポソームを調製する。一部の実施形態では、非機能的トランスポゾンはブロック化3'-末端を含む。一部の実施形態では、トランスポザーゼを、ブロック化トランスポゾン核酸と非ブロック化トランスポゾン核酸とに接触させることにより複数のトランスポソームを調製する。一部の実施形態では、ブロック化トランスポゾン核酸を含むトランスポゾン核酸の非ブロック化トランスポゾン核酸に対する比率は、1:1以上である。一部の実施形態では、ブロック化トランスポゾン核酸を含むトランスポゾン核酸の、非ブロック化トランスポゾン核酸に対する比率は、1:2、1:3、1:5、1:10、1:20、1:30、1:40、1:50、1:75、1:100、2:1、3:1、4:1、5:1、6:1、7:1、8:1、9:1、10:1、20:1、30:1、40:1、50:1、60:1、70:1、80:1、90:1、または100:1とすることが可能である。

40

50

【 0 0 6 2 】

一部の実施形態は、増幅アダプターを標的核酸に付加するステップも含む。一部の実施形態では、増幅アダプターは、アンカー部位、シーケンシングプライマー部位、増幅プライマー部位、およびレポータータグからなる群から選択される配列を含む。

【 0 0 6 3 】

一部の実施形態は、捕捉核酸を増幅するステップも含む。一部の実施形態では、捕捉核酸の増幅にはブリッジ増幅を含む。

【 0 0 6 4 】

一部の実施形態では、表面は複数の捕捉プローブを含む。一部の実施形態では、捕捉プローブは核酸を含む。一部の実施形態では、捕捉プローブはそれぞれ親和性部を含む。一部の実施形態では、親和性部はピオチン、アビジン、ストレプトアビジン、およびリコンビナーゼからなる群から選択される。

10

【 0 0 6 5 】

一部の実施形態では、トランスポゾン核酸は、アンカー部位、シーケンシングプライマー部位、増幅プライマー部位、ユニークな分子インデックス、およびレポータータグからなる群から選択される配列を含む。

【 0 0 6 6 】

一部の実施形態では、少なくとも1つのトランスポソームは、2つのトランスポゾン核酸を含む。一部の実施形態では、2つのトランスポゾン核酸は異なる配列を有する。

【 0 0 6 7 】

一部の実施形態では、複数のトランスポソームは少なくとも2つの異なるトランスポゾン核酸を含む。

20

【 0 0 6 8 】

一部の実施形態では、標的核酸は、ゲノムDNAのDNA断片およびcDNAからなる群からなる群から選択される。一部の実施形態では、標的核酸はゲノムDNAおよびcDNAからなる群から選択される。一部の実施形態では、標的核酸はゲノムDNAである。

【 0 0 6 9 】

一部の実施形態では、表面は、ビーズ、スライド、フローセル、チャネル、ディップスティック、およびウェルからなる群から選択される基材上にある。

【 0 0 7 0 】

一部の実施形態では、表面は mm^2 あたり少なくともおよそ10,000の捕捉核酸を含む。一部の実施形態では、表面は mm^2 あたり少なくともおよそ100,000の捕捉核酸を含む。一部の実施形態では、表面は mm^2 あたり少なくともおよそ1,000,000の捕捉核酸を含む。一部の実施形態では、表面は mm^2 あたり少なくともおよそ、1,000,000、1,500,000、2,000,000、3,000,000、5,000,000、10,000,000、15,000,000、20,000,000、30,000,000、40,000,000、50,000,000、60,000,000、70,000,000、80,000,000、90,000,000、100,000,000、150,000,000、200,000,000、300,000,000、350,000,000、400,000,000、450,000,000、500,000,000、550,000,000、600,000,000、650,000,000、700,000,000、750,000,000、800,000,000、850,000,000、900,000,000、950,000,000、1000,000,000、1200,000,000、1300,000,000、1400,000,000、1500,000,000、1600,000,000、1700,000,000、1800,000,000、1900,000,000、2000,000,000、3000,000,000、4000,000,000、5000,000,000、6000,000,000、7000,000,000、8000,000,000、9000,000,000、10,000,000,000以上の捕捉核酸を含む。

30

40

【 0 0 7 1 】

一部の実施形態は、前記方法のうち任意の1つにより調製した、バーコードを含むシーケンシングライブラリを含む。

【 0 0 7 2 】

一部の実施形態では、トランスポザーゼでの処置後、または、後続の増幅後、1つまたは複数の認識配列をニック入り標的核酸に追加または挿入することができる。1つまたは複数の認識配列としては、限定されるわけではないが、ニックング部位における、バーコード配列、プライマー配列、またはアダプターDNA配列が挙げられ、これは、標的隣接

50

関係、コンパートメント関係、または距離空間関係がユニークであるとして核酸断片をタグ付けする。

【0073】

タグ付けした後、上記の、共通の特性を有する認識配列を特定することにより近接情報を得るシーケンシングプラットフォームを用いて、ショットガン核酸分子をシーケンシングすることができる。一部の実施形態では、共通の特性とは同一または相補的なバーコード配列である。例えば、隣接する元のリード配列は、共通のバーコード配列を介して特定することができ、または、リードは、同一の標的DNAセグメントに由来する共通のコンパートメント特異的バーコードに基づいて、コンパートメントごとに定義することができる。他の実施形態では、共通の特性とは、共通するまたは制約付きの物理的な位置であり、これはフローセル上の1つまたは複数のx、y座標により指定することができる。「制約付き」の物理的位置は、近い物理的位置、同一の物理的位置、実質的に同一の物理的位置、または、相対的な物理的座標が、核酸断片が由来する標的核酸配列上の相対配列座標と相関する、2つ以上の物理的位置の組を意味することができる。例えば、ロングレンジ近接に関する方法では、シーケンシングフローセル表面上で引き延ばしたHMWゲノムDNAへのin situ転移をアダプター配列を用いて実行して、アダプター配列、ハイブリダイズしたDNA断片、またはその組み合わせの制約付き物理的位置（つまり、物理的に連結したシーケンシング鋳型が固定される相対的座標）の特定により、距離空間関係を得る。該方法は、ショートレンジ、ミッドレンジ、およびロングレンジの近接情報の取得のために用いることが可能である。

【0074】

一部の実施形態では、片側転移を組み合わせバーコーディングと組み合わせることが可能である。片側転移要素の使用は、プロセス中に関連ライブラリ要素をまとめておくための任意の追加メカニズムを必要とすることなく、組み合わせバーコーディングを可能にする。片側転移を組み合わせバーコーディングと組み合わせる例示的なスキームを図15に示す。

【図面の簡単な説明】

【0075】

【図1】トランスポゾン核酸を含むトランスポソームの集合に標的核酸を接触させる、例示的な実施形態を示す図である。

【図2】異なるトランスポゾン核酸を含むトランスポソームの集合に標的核酸を接触させ、異なるトランスポゾン核酸を標的核酸の鎖の異なるニック部位に付加する、例示的な実施形態を示す図である。

【図3】修飾標的核酸を線形増幅により増幅して、ある増幅産物を得る、例示的な実施形態を示す図である。

【図4】トランスポソームはダイマートランスポザーゼを含み、トランスポゾン核酸は、モザイク要素(ME)を含む2つのトランスポゾン要素を含み、該MEの1つは3'末端においてジデオキシ基でブロックされている、例示的な実施形態を示す図である。

【図5】異なるバーコードを含むトランスポソームの集合を標的核酸に接触させ；トランスポゾン核酸をニック部位の一方の側に付加し、もう一方の付加されていないトランスポゾン核酸末端をライゲーションによりニック部位のもう一方の側に付加し；修飾標的核酸を全ゲノム増幅(WGA)により増幅する、例示的な実施形態を示す図である。

【図6】ゲノムDNAを、トランスポソームなし(アンプリコン)、(1)トランスポザーゼと、3'ビオチン基でブロックされたトランスポゾン核酸とを含むトランスポソーム(3' Bio)、(2)トランスポザーゼと、3'スぺーサ基でブロックされたトランスポゾン核酸とを含むトランスポソーム(3'スぺーサ)、または(3)トランスポザーゼと、非ブロック化トランスポゾン核酸とを含むトランスポソーム(TDE1)で処置した結果を描く図である。

【図7】挿入物頻度が100bpである500bpリードおよび挿入物頻度が50bpである300bpリードの、名目カバレッジ深度と平均合成リード長のグラフを示す図であ

10

20

30

40

50

る。

【図 8 A】y-型アダプタートランスポゾンを用いた片側転移を示す図である。

【図 8 B】y-型アダプタートランスポゾンを用いた片側転移を示す図である。

【図 9】片側転移反応からのサンプルを搭載したアガロースゲルの写真を示す図である。

【図 10】 $n + 1$ および $n - 1$ のトランスポゾン多様体を用いた転移反応ランからのサンプルを搭載したアガロースゲルの写真を示す図である。

【図 11】活性および不活性のトランスポソーム混合物で実行した転移反応の、ダイアグラムを示す図である。

【図 12】標的核酸にニックを入れ、オリゴヌクレオチドアダプターをライゲーションする例示的なスキームを示す図である。

【図 13】片側転移を用いたライブラリ調製の例示的なスキームを示す図である。

【図 14】DNA の 2 つの鎖の転移に対する抵抗差を活用することによる片側転移の例示的なスキームを示す図である。

【図 15】組み合わせバーコーディングと連動する片側転移の例示的なスキームである。片側転移を用いて、片面 (single-sided) 産物同士を、外部メカニズムを必要とすることなく近接に維持することが可能である。ユニークだが識別できない分子 (A、B、および C) を一緒に収容する。それらはランダムに別の反応に分け、モジュールバーコードを加える。各ステップにおける別個の反応数はモジュール数よりも少ないが、反応経路は各分子でユニークな傾向があり、それぞれにおいてユニークなバーコードの組み合わせを生む。

【発明を実施するための形態】

【0076】

本明細書で提供する実施形態は、次世代シーケンシングのための方法および組成物に関する。一部の実施形態は、片側転移を用いた標的核酸からの鋳型ライブラリの調製と、鋳型ライブラリのシーケンシングとを含む。一部の実施形態では、片側転移は、二本鎖核酸の鎖にニックを入れるトランスポザーゼと、トランスポゾン核酸を、ニック入り鎖のニック部位の一方の側に付加するステップとを含む。有利なことに、片側転移は、両側転移 (例えば *Nextera* (商標)) と比較し、二本鎖標的核酸を断片化しない。そのため、ゲノム DNA などのある標的核酸では、近接情報、ハプロタイプ情報、および/またはフェージング情報が維持され得る。

【0077】

本明細書で提供する方法および組成物の一部の実施形態は、片側トランスポザーゼ活性を有するトランスポソームと、該トランスポソームを用いてシーケンシングライブラリを調製するステップと、該ライブラリをシーケンシングするステップとを含む。一部の実施形態では、トランスポソームは、片側トランスポザーゼ活性を有するトランスポザーゼを含むことが可能である。一部の実施形態では、トランスポソームは、二本鎖標的核酸の両鎖にトランスポゾンを挿入することを阻害するブロッキング基を有し得る、トランスポゾン核酸を含むことが可能である。トランスポザーゼはまた、レトロトランスポゾンに由来するインテグラーゼ、およびレトロウイルストランスポザーゼを含む。例示的なトランスポザーゼとしては、Mu、Tn10、Tn5、および高活性 Tn5 (Goryshin and Reznikoff, J. Biol. Chem., 273:7367 (1998)) が挙げられるが、これらに限定されない。本明細書で提供する方法および組成物の一部で有用なトランスポザーゼの実施形態としては、米国特許出願公開第 2010/0120098 号明細書 (これはその全体が参照により本明細書に組み込まれる) に開示されているものが挙げられる。トランスポザーゼとトランスポゾン要素のより多くの実施形態としては、高活性 Tn5 トランスポザーゼおよび Tn5 型トランスポザーゼ要素 (Goryshin and Reznikoff, J. Biol. Chem., 273:7367 (1998)) (これはその全体が参照により本明細書に組み込まれる) 、MuA トランスポザーゼ、および R1 末端配列と R2 末端配列とを含む Mu トランスポザーゼ要素 (Mizuuchi, Cell, 35: 785, (1983) および Savilanti, et al., EMBO J., 14: 4893, 15 (1995)) (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) が挙げられる。高活性

10

20

30

40

50

Tn5トランスポザーゼ（例えば、ウィスコンシン州マディソン、Epicentre Biotechnologies社のEZ-Tn5（商標）トランスポザーゼ）と複合体を形成する例示的なトランスポザーゼ要素は、国際公開第2012/061832号、米国特許出願公開第2012/0208724号明細書、同第2012/0208705号明細書、および国際公開第2014/018423号（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）に記載されている。本明細書で提供する方法および組成物の一部で有用なトランスポザーゼおよびトランスポゾン配列のより多くの実施形態としては、黄色ブドウ球菌Tn552 (Colegio et al., J. Bacteriol., 183: 2384-8 (2001); Kirby et al., Mol. Microbiol., 43: 173-86 (2002))、Ty1 (Devine & Boeke, Nucleic Acids Res., 22: 3765-72 (1994)および国際公開第95/23875号)、トランスポゾンTn7 (Craig, Science 271: 1512 (1996); Craig, Curr Top Microbiol Immunol., 204:27-48 (1996))、Tn/OおよびIS10 (Kleckner et al., Curr Top Microbiol Immunol., 204:49-82 (1996))、Marinerトランスポザーゼ (Lampe et al., EMBO J., 15: 5470-9, (1996))、Tel (Plasterk, Curr Topics Microbiol. Immunol., 204: 125-43, (1996))、P因子 (Gloor, Methods Mol. Biol., 260: 97-114, (2004))、Tn3 (Ichikawa & Ohtsubo, J Biol. Chem. 265: 18829-32, (1990))、バクテリア挿入配列 (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))、イーストのレトロトランスポゾン (Boeke & Corces, Annu Rev Microbiol. 43:403-34, (1989))が挙げられる。より多くの例としては、IS5、Tn10、Tn903、IS911、およびトランスポザーゼファミリー酵素の改変バージョンが挙げられる (Zhang et al., PLoS Genet. 5:e1000689. Epub 2009 Oct 16、およびWilson et al. Microbiol. Methods 71:332-5 (2007))。より多くの例としては、MuAトランスポザーゼ（例えば、Rasila TS, et al., (2012) PLoS ONE 7(5):e37922. doi:10.1371/journal.pone.0037922を参照）が挙げられる。本明細書で提供する方法および組成物の一部の実施形態で有用なトランスポゾンの例は、Leschziner, A.E., et al., (1998) P.N.A.S. 95:7345-7350、およびHaapa S., et al., (1999) N.A. Res. 27:2777-2784に記載されている（これらはそれぞれ、その全体が参照により組み込まれる）。アミノ酸の置換、挿入、欠失、および/または他のタンパク質もしくはペプチドとの融合を有するようなTn5トランスポザーゼの多様体は、米国特許第5925545号明細書、同第5965443号明細書、同第7083980号明細書、同第7608434号明細書、および米国特許出願公開第14/686961号明細書に開示されている。前記特許および特許出願はその全体が参照により本明細書に組み込まれる。一部の実施形態では、Tn5トランスポザーゼは、米国特許出願公開第14/686961号明細書に開示されているように、野生型タンパク質に対し、54、56、372、212、214、251、および338の位置で1つまたは複数の置換を含む。一部の実施形態では、Tn5野生型タンパク質またはその多様体は、融合ポリペプチドをさらに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を含むことが可能である、本段落で言及した参考文献はそれぞれ、その全体が参照により本明細書に組み込まれる。

【0078】

一部の実施形態では、二本鎖標的核酸を複数のトランスポソームに接触させて、その結果、標的核酸の鎖にニックを入れ、ニック入り標的核酸鎖のニック部位の一方の側にトランスポゾン核酸を付加して、修飾標的核酸を得る。一部の実施形態では、トランスポソームを溶液中で標的核酸に接触させる。本実施形態では、修飾標的核酸を溶液中で生成し、それに続いて表面上に捕捉することが可能である。代わりに、トランスポソームと標的核酸の接触は、表面上で行うことも可能である。接触を行う前に、トランスポソームまたは標的核酸を表面上に付加することが可能である。トランスポソームと標的核酸の表面上での接触から生じた修飾標的核酸は、表面上に捕捉したままにすることが可能であり、または、該修飾標的核酸は表面から解放することが可能である。

【0079】

一部の実施形態では、捕捉した核酸をシーケンシングする。一部の実施形態では、2つの捕捉核酸から得た配列情報の、標的核酸配列の線形表現における近接度は、表面上の捕捉核酸の近接度を示唆する。一部の実施形態では、表面上で互いの近接度がより高い捕捉核酸は、近接度のより低い捕捉核酸と比較し、標的核酸配列の表現における近接度がより高い配列を含む。一部の実施形態では、標的核酸配列の表現にはハプロタイプまたはアセンブリ表現が含まれる。

【0080】

本明細書で提供する方法および組成物の一部の実施形態はまた、シーケンシングした標的核酸断片のde novoアセンブリにおいて片側転移を使用するステップを含む。一部の実施形態では、ランドマークを標的核酸に挿入し、それを、シーケンシングした標的核酸断片のアセンブリで用いて標的核酸配列の表現を生成することが可能である。一部の実施形態では、重複断片は共通の挿入ランドマークを含み得る。ランドマークの使用は、高反復配列を含む標的核酸で特に有利である。また、一部の実施形態では、参照配列を必要としない。

【0081】

一部の実施形態では、片側トランスポザーゼ活性と、異なるバーコードを含むトランスポゾン核酸とを有するトランスポソームの集合に標的核酸を接触させることにより、ランドマークを標的核酸に挿入する。一部の実施形態では、トランスポゾン核酸を、片側転移およびその後のライゲーションにより標的核酸の一本鎖に挿入する。一部の実施形態では、トランスポザーゼは標的核酸の鎖にニックを入れ、ニック入り標的核酸の一方の鎖のニック部位にトランスポゾン核酸を付加し、トランスポゾン核酸のもう一方の末端は、ニック入り標的核酸のニック部位のもう一方の側にライゲーションすることにより、ループを含む一方の鎖に挿入を有する修飾二本鎖標的核酸を得る。一部の実施形態では、修飾核酸を増幅してシーケンシングすることが可能である。一部の実施形態では、修飾核酸を表面に付加することが可能である。一部の実施形態では、付加は、一本鎖結合タンパク質、つまり一本鎖ループを結合させるタンパク質、例えばリコンビナーゼを介して行うことが可能である。

【0082】

一部の実施形態では、標的核酸は転移なしに修飾される。一部の実施形態では、標的核酸はランダムに、ニックングエンドヌクレアーゼ、例えば、アメリカ合衆国マサチューセッツ州のNew England Biolabs社によるニックングエンドヌクレアーゼ、または制限エンドヌクレアーゼを用いていることが可能である。例示的な制限エンドヌクレアーゼとして、EcoRI、EcoRII、BamHI、HindIII、TaqI、NotIが挙げられるが、これらに限定されない。制限エンドヌクレアーゼの他の例は、New England Biolabs社のカタログで見つけられる。オプションとして、3'または5'のエキソヌクレアーゼ活性を有する酵素で、例えば、Exonuclease IまたはExonuclease IIまたはExonuclease IIIで、ギャップを伸長させることが可能である。オリゴヌクレオチドアダプターを、標的核酸のニック入り末端にライゲーションすることが可能である。一部の実施形態では、オリゴヌクレオチドアダプターは、シーケンシングプライマー部位および増幅プライマー部位などのプライマー結合部位を含むことが可能であり、追加配列はまた、切断部位、ユニークな分子インデックス、アンカー部位、レポータータグ、およびバーコードを含むことが可能である。従って、標的核酸にニックを入れることおよび1つまたは複数のアダプターをライゲーションすることは、標的核酸を断片化なしに無傷のままとする。標的核酸にニックを入れることと、オリゴヌクレオチドをライゲーションすることの例示的なスキームを図12に示す。

【0083】

本明細書で用いる場合、「核酸」には、連結し合った少なくとも2つのヌクレオチドモノマーを含む。例としては、ゲノムDNAもしくはcDNAなどのDNA、mRNA、sRNA、もしくはrRNAなどのRNA、または、DNAおよびRNAのハイブリッドが

10

20

30

40

50

挙げられるが、これらに限定されない。下記の例および本明細書のその他の箇所から明らかのように、核酸は、天然に発生した核酸構造または非天然に発生した核酸アナログ構造を有し得る。核酸はホスホジエステル結合を含み得る。しかしながら、一部の実施形態では、核酸は他の種類のバックボーンを有してよく、例えば、ホスホルアミド、ホスホロチオエート、ホスホロジチオエート、0-メチルホスホロアミダイト、およびペプチド核酸の、バックボーンおよび連鎖が挙げられる。核酸は、陽性バックボーン、非イオン性バックボーン、および非リボース系バックボーンを有することが可能である。核酸はまた、1つまたは複数の炭素環式糖質を含むことができる。本明細書の方法または組成物で用いる核酸は、定めるように、一本鎖、または代わりに二本鎖とすることができる。一部の実施形態では、核酸は、例えば二又(forked)アダプターに示されるように、二本鎖配列および一本鎖配列両方の一部分を含むことが可能である。核酸は、デオキシリボヌクレオチドおよびリボヌクレオチドの任意の組み合わせ、ならびに、ウラシル、アデニン、チミン、シトシン、グアニン、イノシン、キサンチン、ヒポキサンチン、イソシトシン、イソグアニン、および、(3-ニトロピロールなどの)ニトロピロールそして(5-ニトロインドールなどの)ニトロインドールなどといった塩基アナログを含む、塩基の任意の組み合わせを含み得る。一部の実施形態では、核酸は少なくとも1つの無差別(promiscuous)塩基を含み得る。無差別塩基は、1超の異なる種類の塩基と塩基対になることが可能であり、例えば、ゲノムDNAサンプルなどの複合体核酸サンプルでのランダムなハイブリダイゼーションに用いる、オリゴヌクレオチドプライマーまたはオリゴヌクレオチド挿入に含まれる場合に有用であり得る。無差別塩基の例としては、アデニン、チミンまたはシトシンと対になり得るイノシンが挙げられる。他の例としては、ヒポキサンチン、5-ニトロインドール、アシル化5-ニトロインドール、4-ニトロピラゾール、4-ニトロイミダゾールおよび3-ニトロピロールが挙げられる。少なくとも2つ、3つ、4つ以上の種類の塩基と塩基対になることが可能な無差別塩基を用いることが可能である。

【0084】

本明細書で用いる場合、「ヌクレオチド配列」には、核酸ポリマーのヌクレオチドモノマーの並び順と種類を含む。ヌクレオチド配列は、核酸分子の特徴であり、例えば、描写、イメージ、電子媒体、一連のシンボル、一連の数、一連の文字、一連の色などを含む種々のフォーマットのいずれかで表現することが可能である。情報は、例えば、単一ヌクレオチド分析、より高度な分析(例えば、ヌクレオチドサブユニットの分子構造を示す)、または、より低度の分析(ハプロタイプブロックなどの染色体領域を示す)で表現することが可能である。一連の「A」、「T」、「G」、および「C」の文字は、DNAの周知の配列表現であり、これは、単一ヌクレオチド分析においてDNA分子の実際の配列と互いに関連付けることが可能である。一続きにおいて「T」を「U」に置き換えることを除き、同様の表現がRNAでも用いられる。

【0085】

本明細書で用いる場合、「異なる」という用語は、核酸に関連して用いる際は、核酸が互いに同一ではないヌクレオチド配列を有することを意味する、2つ以上の核酸は、その全長に沿って異なるヌクレオチド配列を有することが可能である。代わりに、2つ以上の核酸は、その長さの実質的部分に沿って異なる核酸配列を有することが可能である。例えば、2つ以上の核酸は、2つ以上の分子で異なる標的ヌクレオチド配列部分を有する一方で、2つ以上の分子で同一のユニバーサルな配列部分も有することも可能である。ユニバーサル配列は核酸末端で、または、コピー、検出、もしくは増幅する予定の核酸領域の側面で生じ得る。

【0086】

本明細書で用いる場合、「ハプロタイプ」には、個人が両親の一人から受け継いだ1超の座におけるアレルの組を含む。ハプロタイプには、染色体の全てまたは一部からの2つ以上の座が含まれ得る。アレルには例えば、一塩基多型(SNP)、縦列型反復配列(STR)、遺伝子配列、染色体挿入、染色体欠失などが含まれる。「フェーズド(phased)アレル」という用語は、特定の染色体による特定のアレルの分布またはその一部分を意味

10

20

30

40

50

する。従って、2つのアレルの「フェーズ (phase)」は、1つまたは複数の染色体の2つ以上のアレルの相対位置の特徴付けまたは表現を意味し得る。

【0087】

本明細書で用いる場合、核酸の「ニック」は、2つの鎖のうち一方のみが切断バックボーン構造を含む、二本鎖核酸の領域を意味する。従って、「ニックング (nicking)」とは、二本鎖核酸の領域内で一方の核酸鎖のみの共有結合構造を切断する行為を意味する。前記領域は概して、二本鎖核酸の一部分のみである。該部分には、例えば、多くて、5塩基対、10塩基対、25塩基対、50塩基対、100塩基対、200塩基対、300塩基対、400塩基対、500塩基対、1000塩基対が含まれ得る。該領域には、二本鎖核酸のより大きい部分またはより小さい部分が含まれ得る。例えば、上記で例示した上限の代わりに、またはそれに加えて、ニックを入れる核酸の一部分の下限は、オプションとして、少なくとも500塩基対、400塩基対、300塩基対、200塩基対、100塩基対、50塩基対、25塩基対、10塩基対以下とすることが可能である。上記の範囲で挙げた値は、核酸領域の集合の全メンバの最大サイズまたは最小サイズを定義するか、または、代わりに、前記領域を有する核酸の集合の平均を意味し得る。二本鎖核酸では両方の鎖にニックを入れることが可能であり、第1ニックは第1領域で生じ、第2ニックは第2領域で生じることが理解されよう。該して、2つのニック入り領域の効果的な連結性は、核酸がハイブリダイズされた二本鎖形態のままであるという条件下で維持され得る。対照的に、二本鎖核酸の同一の領域における両鎖の切断は、切断部位の側面に位置する核酸領域の間の効果的な連結性の消失を招き得る。

【0088】

本明細書で用いる場合、「表面」という用語は、ガス状流体または液状流体などの周囲の流体と直接接触している固体支持体またはゲル物質の、部分または層を意味することを意図する。表面は、ガス、液体、ゲル、ポリマー、有機ポリマー、類似のもしくは異なる物質の第2表面、金属、またはコートなど別の物質に接触してよい。表面またはその領域は、実質的に平坦とすることが可能である。表面は、ウェル、ピット、チャネル、リッジ、隆起領域、ペグ、またはポストなどの表面特徴を有することが可能である。有孔基材の場合、表面は、流体が基材に接触している孔に位置することができる。例えば、表面はゲルの孔内にあり、ここで付加部は孔に入った流体と相互作用する。従って、表面の「上」にある部分は、ゲルなどの有孔物質の孔に位置することができる。

【0089】

本明細書で用いる場合、「固体支持体」という用語は、水性液体に溶けない剛性基材を意味する。基材は無孔または有孔とすることが可能である。基材はオプションとして、液体を（例えば有孔性のために）吸収することができるが、典型的には、十分に剛性であるため基材は液体を吸収した場合も実質的には膨張せず、液体が乾燥により除かれた場合も実質的には収縮しない。無孔固体支持体は、概して液体またはガスを通さない。例示的な固体支持体としては、ガラス、修正したまたは機能性をもたせたガラス、プラスチック（アクリル、ポリスチレン、およびスチレンと他の材料とのコポリマー、ポリプロピレン、ポリエチレン、ポリブチレン、ポリウレタン、テフロン（登録商標）、環状オレフィン、ポリイミドなど）、ナイロン、セラミックス、樹脂、ゼオノア（登録商標）、シリカ、もしくはシリコンおよび修飾シリコンを含むシリカ系物質、炭素、金属、無機ガラス、光ファイバー束、ならびにポリマーが挙げられるが、これらに限定されない。一部の実施形態で特に有用な固体支持体は、フローセル装置内に位置する。

【0090】

本明細書で用いる場合、「ゲル物質」という用語は、液体およびガスを通す半剛性の基材を意味することを意図する。典型的には、ゲル物質は液体を吸収した場合に膨張し、液体が乾燥により除かれた場合は収縮することが可能である。例示的なゲルとしては、アガロースなどのコロイド構造、ゼラチンなどのポリマーメッシュ構造、または、ポリアクリルアミド、SFA（例えば、米国特許出願公開第2011/0059865号明細書（これは参照により本明細書に組み込まれる）を参照）、もしくはPAZAM（例えば、米国

仮特許出願公開第61/753833号明細書（これは参照により本明細書に組み込まれる）を参照）などの架橋ポリマー構造を有するものが挙げられるが、これらに限定されない。

【0091】

本明細書で用いる場合、「付加された(attached)」という用語は、拡散による分離を防ぐ力により連結されていることを意味することを意図する。該用語は、事実上共有結合的または非共有結合的な連結を含み得る。例えば核酸は、結合の鎖を生成する1つまたは複数の共有結合的結合により、表面上に共有結合的に付加させることが可能である。非共有結合的付加は、2つのものの間（例えば、核酸と表面の間）の少なくとも1つの結合が共有結合的結合ではない場合に起きる。非共有結合的結合の例としては、例えば、水素結合、イオン性結合、ファンデルワールス力、または疎水結合などが挙げられる。

10

【0092】

本明細書で用いる場合、「近接情報」という用語は、共通の情報に基づいた、2つ以上のDNA断片間の空間関係を意味する。情報の共通点は、隣接関係、コンパートメント関係、または距離空間関係に関するものとすることができる。これらの関係に関する情報は、その結果、DNA断片に由来するシーケンスリードの階層的なアセンブリまたはマッピングを容易にする。この近接情報は、このようなアセンブリまたはマッピングの効率性および正確性を改善する。これは、既存のショットガンシーケンシングに関連して用いる従来のアセンブリまたはマッピングの方法が、個別シーケンスリードが由来する2つ以上のDNA断片の間の空間関係に関係する場合において、個別シーケンスリードの相対的ゲノム源または座標を考慮しないためである。そのため、本明細書に記載の実施形態では、近接情報を得る方法は、隣接する空間関係を確定するショートレンジ近接方法、コンパートメント空間関係を確定するミッドレンジ近接方法、または、距離空間関係を確定するロングレンジ近接方法により行うことができる。これらの方法は、DNA配列のアセンブリまたはマッピングの正確性および質を改善し、本明細書に記載のものなどの任意のシーケンシング法とともに用いることができる。

20

【0093】

一部の実施形態では、このステップは、標的DNA配列に由来するショットガン核酸分子ライブラリの生成をもたらす。代替の実施形態では、断片化または挿入でさえ、以下に記載するようにYアダプターアプローチにより行うことができる。1つまたは複数のトランスポザーゼ分子は可溶性フリートランスポザーゼとすることができ、または、表面結合認識配列と関連付けることができる。

30

【0094】

本明細書で用いる場合、「バーコード」という用語は、標的核酸配列に特有で、該標的核酸配列とは全く無関係の核酸配列を意味する。該してバーコードは、1つまたは複数の特定の核酸を特定するのに用いることが可能な、1つまたは複数のヌクレオチド配列を含み得る。バーコードは、人工配列とするか、または、かつて並置されたDNA断片の末端にある同一の側面ゲノムDNA配列(g-コード)などの、転移中に生成される天然発生配列とすることが可能である。バーコードは、数なくともおよそ、1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20以上の連続ヌクレオチドを含み得る。一部の実施形態では、バーコードは少なくともおよそ、10、20、30、40、50、60、70、80、90、100以上の連続ヌクレオチドを含む。一部の実施形態では、バーコードを含む核酸集合のバーコードの少なくとも一部分が異なる。一部の実施形態では、バーコードの少なくともおよそ、10%、20%、30%、40%、50%、60%、70%、80%、90%、95%、99%が異なる。より多くのこのような実施形態では、バーコードの全てが異なる。バーコードを含む核酸集合における異なる様々なバーコードが、ランダムまたは非ランダムに発生し得る。

40

【0095】

一部の実施形態では、トランスポゾン配列は少なくとも1つのバーコードを含む。2つの非連続トランスポゾン配列を含むトランスポソームなどの一部の実施形態では、第1トランスポゾン配列は第1バーコードを含み、第2トランスポゾン配列は第2バーコードを

50

含む。一部の実施形態では、トランスポゾン配列は、第1バーコード配列と第2バーコード配列とを含むバーコードを含む。前記実施形態の一部では、第1バーコード配列は特定することが可能であり、または、第2バーコード配列と対になるように設計することが可能である。例えば、既知の第1バーコード配列は、互いに対になることが知られている複数の第1バーコード配列および第2バーコード配列を含む参照表を用いて、既知の第2バーコード配列と対にできることが知られている。

【0096】

別の例では、第1バーコード配列は第2バーコード配列と同じ配列を含み得る。別の例では、第1バーコード配列は第2バーコード配列の逆相補体を含み得る。一部の実施形態では、第1バーコード配列および第2バーコード配列は異なる。第1バーコード配列と第2バーコード配列はb i -コードを含み得る。

10

【0097】

本明細書に記載する組成物および方法の一部の実施形態では、バーコードを鋳型核酸の調製において用いる。理解されるように、莫大な数の利用可能なバーコードは、各鋳型核酸分子がユニークな識別性を備えることを可能にする。鋳型核酸混合物の各分子のユニークな識別性は、いくつかの用途で利用することが可能である。例えば、ユニークに特定された分子を適用して、多数の染色体を有するサンプル、ゲノム、細胞、細胞型、細胞病状態、ならびに、ハプロタイプシーケンシング、親アレル識別、メタゲノムシーケンシング、およびゲノムのサンプルシーケンシングなどの種類において、個別核酸分子を特定することが可能である。

20

【0098】

一部の実施形態では、標的核酸全体にわたる複数のユニークなバーコードを転移中に挿入することができる。一部の実施形態では、各バーコードは第1バーコード配列と第2バーコード配列を含み、その間には断片部位が配置されている。第1バーコード配列および第2バーコード配列は特定するか、または、互いに対になるように設計することが可能である。対にすることは有益であるため、第1バーコードは第2バーコードと関連付ける。有利なことに、対にしたバーコード配列を用いて、鋳型核酸ライブラリからのシーケンシングデータをアセンブルすることが可能である。例えば、第1バーコード配列を含む第1鋳型核酸と、第1と対になった第2バーコード配列を含む第2鋳型核酸とを特定することは、第1鋳型核酸および第2鋳型核酸が、標的核酸配列の表現において互いに隣接する配列を表すことを示す。このような方法を用いて、参照ゲノムを必要とすることなく、de novoで標的核酸配列表現をアセンブルすることが可能である。

30

【0099】

本明細書で用いる場合、「少なくとも一部分」という用語および/またはその文法的同等物は、全体量のうち任意の部分を意味し得る。例えば、「少なくとも一部分」は、全体量のうち少なくともおよそ、1%、2%、3%、4%、5%、6%、7%、8%、9%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、99%、99.9%、または100%を意味する。

【0100】

本明細書で用いる場合、「およそ」という用語は+/-10%を意味する。

40

【0101】

標的核酸

本明細書で提供する方法および組成物の一部の実施形態は、標的核酸を含む。一部の実施形態では、標的核酸は二本鎖核酸を含む。一部の実施形態では、標的核酸はゲノムDNAまたはcDNAを含む。一部の実施形態ではミトコンドリアDNAまたはクロロプラストDNAを用いる。一部の実施形態では標的核酸は、RNA、または、mRNAもしくはcDNAなどのその誘導体を含む。本明細書に記載の一部の実施形態は、1つのコピー（つまり、単一分子）に存在する、または代わりに多数のコピー（つまり、同一の配列を有する核酸分子のアンサンブル）に存在する、単一の標的核酸種を利用することが可能である。他の実施形態は、複数の異なる標的核酸種（例えば、前記複数に存在する異なるヌク

50

レオチド配列を有する核酸分子)を利用することが可能である。従って、複数の標的核酸は、複数の同一の標的核酸、一部の標的核酸が同じである複数の異なる標的核酸、または、全ての標的核酸が異なる複数の標的核酸を含むことが可能である。標的核酸は、単一生物から得た核酸分子から、または、1超の生物を含む源から得た核酸分子の集合から調製することができる。標的核酸は、単一細胞から、単一生物の多数細胞、組織、もしくは体液から、同一種のいくつかの生物の細胞、組織、もしくは体液から、または、環境サンプルなど、メタゲノムサンプルと同様の多数の種からとすることが可能である。核酸分子源としては、例えばオルガネラ、細胞、組織、器官または生物が挙げられるが、これらに限定されない。

【0102】

一部の実施形態では、標的核酸をトランスポソームに接触させ、その結果、トランスポゾン核酸を標的核酸に挿入または付加して修飾核酸を提供する。一部の実施形態では、修飾核酸は、伸長、増幅、およびライゲーションなど、さらに操作することができる。

【0103】

トランスポソーム

本明細書で提供する方法および組成物の一部の実施形態にはトランスポソームを含む。一部の実施形態では、トランスポソームは1つまたは複数のトランスポゾン核酸に結合したトランスポザーゼを含む。一部の実施形態では、トランスポソームは片側トランスポザーゼ活性を含み、これには、二本鎖核酸の鎖にニックを入れることと、ニック入り鎖のニック部位の一方の側にトランスポゾン核酸を付加することが含まれる。

【0104】

一部の実施形態では、片側トランスポザーゼ活性を有するトランスポソームには、片側トランスポザーゼ活性を有するある種類のトランスポザーゼを含むトランスポソームが含まれる。一部の実施形態では、野生型トランスポザーゼは片側トランスポザーゼ活性を有し、または、片側トランスポザーゼ活性を有するように修飾される。片側トランスポザーゼ活性を有する、または片側トランスポザーゼ活性を有するように修飾し得るトランスポザーゼの例としては、Mu、Mu E392Q、Tn5、RAG、高活性Tn5、Tn5多様体、Vibhar、およびTn552が挙げられる(Leschziner, A.E., et al., (1998) P.N.A.S. 95:7345-7350;およびHaapa S., et al., (1999) N.A. Res. 27:2777-2784 (これらはそれぞれその全体が参照により組み込まれる))。片側トランスポザーゼ活性を有する、または、片側トランスポザーゼ活性を有するように修飾し得るトランスポザーゼのより多くの例を、本明細書で列記する。一部の実施形態では、片側トランスポザーゼ活性を有するトランスポソームは、単一モノマーとトランスポゾン核酸を含む。一部の実施形態では、トランスポザーゼは、ダイマーを形成する機能を欠くように修飾することができる。一部の実施形態では、片側トランスポザーゼ活性を有するトランスポソームは、モノマーの一つがトランスポザーゼ活性を欠いたダイマーを含む。一部の実施形態では、ダイマーのモノマーサブユニットは、共有結合的に連結することができる。

【0105】

一部の実施形態では、片側トランスポザーゼ活性を有するトランスポソームは、ブロック化トランスポゾン核酸を含む。一部の実施形態では、ブロック化トランスポゾン核酸は、ニック入り二本鎖核酸の鎖へ付加することを妨げられる。ブロック化トランスポゾン核酸は、トランスポゾン核酸の3'末端で、該トランスポゾン核酸の別の核酸への付加を阻害するブロッキング基を含むことが可能である。一部の実施形態では、ブロッキング基としてジデオキシ基、スパーサ基、およびビオチン基が挙げられる。一部の実施形態では、一部の実施形態では、片側トランスポザーゼ活性を有するトランスポソームの集合は、ブロック化トランスポゾン核酸と非ブロック化トランスポゾン核酸にトランスポザーゼを接触させることにより調製することが可能である。非ブロック化トランスポゾン核酸には、ブロッキング基を欠くトランスポゾン核酸が含まれる。一部の実施形態では、ブロック化トランスポゾン核酸と非ブロック化トランスポゾンとを有するトランスポザーゼダイマーを備えるトランスポソームを含む集合を得る。一部の実施形態では、トランスポザーゼダイ

10

20

30

40

50

マーには、2つのブロック化トランスポゾン核酸が含まれる。一部の実施形態では、トランスポザーゼダイマーには、2つの非ブロック化トランスポゾン核酸が含まれる。一部の実施形態では、集合における異なる種類のダイマーの割合は、非ブロック化トランスポゾン核酸に対する比率が様々であるブロック化トランスポゾン核酸にトランスポザーゼを接触させることにより、操作することが可能である。一部の実施形態では、ブロック化トランスポゾン核酸の非ブロック化トランスポゾン核酸に対する比率は、1:1、5:1、10:1、50:1、100:1、200:1、500:1、1000:1以上、または前記比率の間の任意の範囲である。

【0106】

他の有用なトランスポゾン核酸は、標準のトランスポゾンよりも短いまたは長いものである。例えば、3'末端の移動鎖を（例えば、1つまたは複数の塩基を取り除くことにより）短くして、移動反応が起きないようにすることが可能である。同様に3'末端を長くすることにより同様に阻止することが可能である。

【0107】

片側転移のための方法はまた、転移後のライブラリの挿入物サイズを調節するために用いることが可能である。このようなアプローチの利点は、挿入物サイズの長さを、活性複合物と不活性複合物の比率により決めることができることであり、これは多くの場合、トランスポソームおよび核酸のインキュベーションまたは濃縮の時間よりも制御することが容易である。

【0108】

活性トランスポソームモノマーサブユニットと不活性トランスポソームサブユニットを有するトランスポソームダイマーを調製し、転移反応が始まる前後で核酸に加えることが可能である。例えば、反応は Mg^{2+} を加えることにより開始することが可能である。特定の実施形態では、3種のトランスポソームダイマー（つまり、活性：活性ダイマー、不活性：不活性ダイマー、活性：不活性ダイマー）の混合物を含む集合を形成することが可能である。異なる活性を有する集合は、組み合わせで混合物を形成する活性トランスポソームサブユニットと不活性トランスポソームサブユニットの、比率を変えることにより調製することが可能である。比率は、核酸サンプルを混合物で処置する際に生成される平均挿入物サイズに影響を与えるように選択することが可能である。挿入物サイズに対する同様の制御は、活性：活性および不活性：不活性という2つのトランスポソーム種のみを有するトランスポソーム種混合物を用いることで達成することが可能である。図11の図で示すように、標的DNAに結合することはできるが標的を転移させることはできない不活性：不活性種は、スペーサとして作用しよう。言い換えると、不活性：不活性ダイマーは、その他の場合は活性：活性ダイマーに結合し、転移されるであろう部位をめぐって争う。トランスポザーゼ反応混合物に混入した不活性：不活性ダイマーの量の機械的滴定を用いて、混合物により生成される平均断片サイズを制御することが可能である。これらの方法は例えば、相対的に時間非依存的で制御可能であるといった、いくつかの利点を有する。従来の転移反応（例えば、Illumina社（カリフォルニア州サンディエゴ）によるNextera（登録商標）Sample Preparation法）は、所望の平均サイズの断片を生成する転移反応を得るために、反応時間の慎重な制御を必要とする。一方、片側転移の現在の方法は、上記のように時間に左右されにくく実行することが可能である。具体的には、活性モノマーサブユニットの不活性モノマーサブユニットに対する比率を選択して、ライブラリの断片サイズを決定することが可能である。

【0109】

一部の実施形態では、片側トランスポザーゼ活性を有するトランスポソームを、表面に付加することが可能である。トランスポソームはトランスポザーゼを介して、またはトランスポゾン核酸を介して付加することが可能である。例えば、トランスポザーゼは表面上に共有結合的または非共有結合的に付加することが可能である。代わりに、または加えて、トランスポゾン核酸は表面上に共有結合的または非共有結合的に付加することが可能である。有用な付加物、表面、およびその調製と使用のための関連する方法は、本明細書および米国特許出願公開第13/790220号明細書（これは参照により本明細書に組

10

20

30

40

50

み込まれる)にさらに詳細に記載されている。

【0110】

一部の実施形態では、トランスポザーゼには、トランスポゾン要素またはトランスポザーゼ要素を含むトランスポゾン核酸と機能的複合体を形成し、標的核酸へのトランスポゾン核酸の挿入または転移に触媒作用を及ぼして修飾核酸を提供することができる酵素を含む。例えば、*in vitro*転移反応において、トランスポゾン核酸を標的DNAに挿入して修飾DNAを提供する。一部の実施形態では、トランスポザーゼには、トランスポゾン要素またはトランスポザーゼ要素を含むトランスポゾン核酸と機能的複合体を形成し、標的核酸への片側転移に触媒作用を及ぼして修飾核酸を提供することができる酵素を含む。

【0111】

一部の実施形態では、トランスポザーゼによるトランスポゾン核酸の挿入または付加は、標的核酸のランダムまたは実質的にランダムな部位とすることが可能である。トランスポザーゼには、レトロトランスポゾンに由来するインテグラーゼ、またはレトロウイルストランスポザーゼも含む。本明細書で提供する方法および組成物の一部で有用なトランスポザーゼの実施形態としては、米国特許出願公開第2010/0120098号明細書(これはその全体が参照により本明細書に組み込まれる)で開示されているものが挙げられる。トランスポザーゼおよびトランスポゾン要素のより多くの実施形態としては、高活性Tn5トランスポザーゼおよびTn5型トランスポザーゼ要素(Goryshin and Reznikoff, *J. Biol. Chem.*, 273:7367 (1998)(これはその全体が参照により本明細書に組み込まれる))、MuAトランスポザーゼ、およびR1末端配列とR2末端配列とを含むMuトランスポザーゼ要素(Mizuuchi, *Cell*, 35: 785, (1983)およびSavilahti, et al., *EMBO J.*, 14: 4893, 15 (1995)(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))が挙げられる。高活性Tn5トランスポザーゼ(例えば、ウィスコンシン州マディソン、Epicentre Biotechnologies社のEZ-Tn5(商標)トランスポザーゼ)と複合体を形成する例示的なトランスポザーゼ要素は、国際公開第2012/061832号、米国特許出願公開第2012/0208724号明細書、同第2012/0208705号明細書、および国際公開第2014/018423号(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)に記載されている。本明細書で提供する方法および組成物の一部で有用なトランスポザーゼおよびトランスポゾン核酸のより多くの実施形態としては、黄色ブドウ球菌Tn552(Colegio et al., *J. Bacteriol.*, 183: 2384-8 (2001); Kirby et al., *Mol. Microbiol.*, 43: 173-86 (2002))、Ty1(Devine & Boeke, *Nucleic Acids Res.*, 22: 3765-72 (1994)および国際公開第95/23875号)、トランスポゾンTn7(Craig, *Science* 271: 1512 (1996); Craig, *Curr Top Microbiol Immunol.*, 204:27-48 (1996))、Tn/OおよびIS10(Kleckner et al., *Curr Top Microbiol Immunol.*, 204:49-82 (1996))、Marinerトランスポザーゼ(Lampe et al., *EMBO J.*, 15: 5470-9, (1996))、Tel(Plasterk, *Curro Topics Microbiol. Immunol.*, 204: 125-43, (1996))、P因子(Gloor, *Method s Mol. Biol.*, 260: 97-114, (2004))、Mos-1トランスポザーゼ(Richardson et al., *EMBO Journal* 25:1324-1334 (2006))、Tn3(Ichikawa & Ohtsubo, *J Biol. Chem.* 265: 18829-32, (1990))、バクテリア挿入配列(Ohtsubo & Sekine, *Curro 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, (1989))が挙げられる。より多くの例としては、IS5、Tn10、Tn903、IS911、およびトランスポザーゼファミリー酵素の改変バージョンが挙げられる(Zhang et al., *PLoS Genet.* 5:e1000689. Epub 2009 Oct 16、およびWilson et al. *Microbiol. Methods* 71:332-5 (2007))。より多くの例としては、MuAトランスポザーゼ(例えば、Rasila TS, et al., (2012) *PLoS ONE* 7(5): e37922. doi:10.1371/journal.pone.0037922を参照)が挙げられる。本段落で言及した参考文献はそれぞれ、その全体が参照により本明細書に組み込まれる。

【0112】

一部の実施形態では、トランスポゾン核酸は二本鎖核酸を含む。トランスポゾン要素には、トランスポザゼまたはインテグラーゼ酵素とトランスポソームを形成する核酸配列を含む、核酸分子またはその一部が含まれる。一部の実施形態では、トランスポゾン要素は、転移反応においてトランスポザゼと機能的複合体を形成することができる。トランスポゾン要素の例を本明細書では提供し、該例としては、19-bp外側末端(「OE」)トランスポゾン末端、内側末端(「IE」)トランスポゾン末端、野生型もしくは変異体Tn5トランスポザゼなどにより認識される「モザイク末端」(「ME」)トランスポゾン末端、または、R1トランスポゾン末端およびR2トランスポゾン末端(例えば、米国特許出願公開第2010/0120098号明細書(これはその全体が参照により本明細書に組み込まれる)を参照)が含まれ得る。トランスポゾン要素には、*in vitro*転移反応においてトランスポザゼまたはインテグラーゼ酵素とともに機能的複合体を形成するのに適切な、任意の核酸または核酸アナログが含まれ得る。例えば、トランスポゾン末端は、DNA、RNA、修飾塩基、非天然塩基、修飾バックボーンを含み得、一方または両方の鎖でニックを含み得る。

【0113】

一部の実施形態では、トランスポゾン核酸は、トランスポゾン要素および追加配列を含み得る。一部の実施形態では、追加配列を転移反応において標的核酸に挿入または付加することが可能である。追加配列は、シーケンシングプライマー部位および増幅プライマー部位などのプライマー結合部位を含むことが可能であり、追加配列はまた、切断部位、ユニークな分子インデックス、アンカー部位、レポータータグ、およびバーコードを含むことが可能である。

【0114】

一部の実施形態では、プライマー結合部位は、シーケンシング反応において核酸にアニールするためのシーケンシングプライマー用配列を含むことが可能である。一部の実施形態では、プライマー結合部位は、増幅反応または他の伸長反応において、核酸にアニールするためのプライマー用配列を含むことが可能である。

【0115】

一部の実施形態では、切断部位には、断片化され得るトランスポゾン核酸の部位が含まれる。例えば、切断部位を含むトランスポゾン核酸は、標的核酸に挿入することが可能であり、修飾核酸はその後挿入切断部位で断片化することが可能である。一部の実施形態では、切断部位には、制限酵素認識配列および/または制限酵素切断部位が含まれる。一部の実施形態では、切断部位は、その他の場合はデオキシリボヌクレオチドを含み得る核酸において少なくとも1つのリボヌクレオチドを含み得、RNAseで切断することができる。デオキシリボヌクレオチドとリボヌクレオチドの間のホスホジエステル結合を選択的に切断できる化学切断剤を用いることが可能であり、例えば、希土類金属イオンなどの金属イオン(例えば、La³⁺、特にTm³⁺、Yb³⁺、もしくはLu³⁺、Fe(3)、またはCu(3))か、または高pHへの露呈が挙げられる。一部の実施形態では、切断部位には、ニックアーゼ、つまり、二本鎖核酸の特定の領域の一鎖を切断するニックングエンドヌクレアーゼ用の、1つまたは複数の認識配列が含まれ得る。従って、断片化部位は、第1ニックアーゼ認識配列と、オプションとして第2ニックアーゼ認識配列を含むことが可能である。第1ニックアーゼ認識配列および第2ニックアーゼ認識配列は、互いに同一でも、互いに異なってよい。一部の実施形態では、切断部位は、無塩基部位を含む1つまたは複数のヌクレオチドアナログを含み得、ポリアミン、N,N'-ジメチルエチレンジアミン(DMEM)などのある化学薬剤の存在下において、断片部位での切断を可能にする(例えば、米国特許出願公開第2010/0022403号明細書(これはその全体が参照により本明細書に組み込まれる)を参照)。一部の実施形態では、無塩基部位は、切断部位内のウラシルヌクレオチドの修飾により、例えば、ウラシルDNAグリコシラーゼ(UDG)酵素を用いて作られ得る。無塩基部位を含むポリヌクレオチド鎖は、その後、エンドヌクレアーゼ(例えば、Endo IVエンドヌクレアーゼ、APリアーゼ、FPGグリコシラーゼ/APリアーゼ、Endo VII Iグリコシラーゼ/APリアーゼ)、熱、

10

20

30

40

50

またはアルカリでの処置により無塩基部位で切断することができる。無塩基部位はまた、デオキシウリジン以外のヌクレオチドアナログで生成され得、エンドヌクレアーゼ、熱、またはアルカリでの処置による類似の方法で切断することができる。例えば、8-オキソグアニンは、FPGグリコシラーゼにさらすことにより無塩基部位に変換することが可能である。デオキシイノシンは、AlkAグリコシラーゼにさらすことにより無塩基部位に変換することが可能である。このように生成された無塩基部位は、その後、典型的にはEndo IVまたはAPリナーゼなどの適切なエンドヌクレアーゼで処置することにより切断することができる（例えば、米国特許出願公開第2011/0014657号明細書（これはその全体が参照により本明細書に組み込まれる）を参照）。別の例では、切断部位はジオール結合を含み得、これは過ヨウ素酸（例えば、過ヨウ素酸ナトリウム）での処置による切断を可能にする。別の例では、切断部位はジスルフィド基を含み得、これは、化学還元剤、例えばトリス(2-カルボキシエチル)ホスフェートハイドロクロライド(TCEP)での切断を可能にする。一部の実施形態では、切断部位は光切断可能部を含み得る。光化学切断は、光エネルギーを利用して共有結合的結合を切断する種々の方法のいずれかにより行うことが可能である。光化学切断のための部位は、ホスホラミダイト[4-(4,4'-ジメトキシトリチルオキシ)ブチルアミドメチル]-1-(2-ニトロフェニル)-エチル]-2-シアノエチル-(N,N'-ジイソプロピル)-ホスホラミダイト（アメリカ合衆国バージニア州のスターリングのGlen Research社、Cat No. 10-4913-XX）など、核酸中の非ヌクレオチド化学部により提供され得る。

【0116】

一部の実施形態では、トランスポゾン核酸はアンカー部位を含み得る。一部の実施形態では、アンカー部位は、捕捉プローブに特異的に結合することが可能な配列を含み得る。一部の実施形態では、アンカー部位は、核酸を含む捕捉プローブに対し相補的な、および/または実質的に相補的な配列を含む。一部の実施形態では、アンカー部位は、対応する受容体またはリガンドを含む捕捉プローブを結合する、リガンドまたは受容体を含み得る。言い換えると、アンカー部位および捕捉プローブは、リガンド/受容体の対を含むことが可能である。一部の実施形態では、リガンドまたは受容体は、修飾ヌクレオチドを介してトランスポゾン核酸のアンカー部位と関連付けることが可能である。リガンドおよび受容体の例としては、それぞれストレプトアビジンまたはニッケルを結合することが可能なビオチンまたはpolyHisが挙げられる。他の例としては、当技術分野で既知のリガンドおよびその受容体の対、例えば、限定されるわけではないが、2-イミノビオチン、デスチオビオチン、NeutrAvidin（分子プローブ、オレゴン州のEugene社）、CaptAvidin（分子プローブ）などが含まれる、アビジン-ビオチン、ストレプトアビジン-ビオチン、およびビオチン、ストレプトアビジン、またはアビジンの誘導体；マルトース-マルトース結合タンパク質(MBP)、カルシウム-カルシウム結合タンパク質/ペプチド(CBP)が含まれる、結合タンパク質/ペプチド；c-MYC、HA、VSV-G、HSV、V5、およびFLAG Tag（商標）などのエピトープタグを含む抗原-抗体、およびそれに対応する抗エピトープ抗体；ジニトロフェニルおよびジゴキシゲニンなどのハプテンおよびそれに対応する抗体；アプタマーおよびそれに対応する標的；対応する固定化金属イオンアフィニティークロマトグラフィー(IMAC)物質と抗ポリ-His抗体を含む、ポリ-Hisタグ（例えば、ペンタ-Hisおよびヘキサ-His）とその結合パートナー；フルオロフォアおよび抗フルオロフォア抗体；核酸鎖およびその相補的な鎖などが挙げられる。

【0117】

一部の実施形態では、トランスポゾン核酸はレポータータグを含むことが可能である。有用なレポータータグには、当技術分野で既知である種々の特定可能なタグ、標識、または基のいずれかが含まれる。ある実施形態では、レポータータグはシグナルを発することが可能である。シグナルの例としては、蛍光性、化学発光性、生物発光性、燐光性、放射性、比色性、または電気化学発光性のものが挙げられる。例示的なレポータータグには、フルオロフォア、放射性同位体、色原体、酵素、エピトープタグを含む抗原、量子ドット

などの半導体ナノクリスタル、重金属、色素、燐光性基、化学発光基、電気化学的検出部、結合タンパク質、蛍光体、希土類キレート、遷移金属キレート、近赤外染料、電気化学発光標識、ならびに、質量タグ、電荷タグ、および同位体などの質量分析計対応リポータータグが挙げられる。本明細書に記載する方法および組成物で用いることができるより多くのレポータータグとしては、蛍光色素（例えば、フルオレセインイソチオシアネート、Texas red（登録商標）、およびローダミン等）などのスペクトル標識；放射標識（例えば、 ^3H 、 ^{125}I 、 ^{35}S 、 ^{14}C 、 ^{32}P 、 ^{33}P など）；酵素（例えば、ホースラディッシュペルオキシダーゼ、アルカリホスファターゼなど）；コロイド状金、着色ガラス、または着色プラスチック（例えば、ポリスチレン、ポリプロピレン、ラテックス等）などのスペクトル比色標識；ビーズ；磁気標識；電気標識；熱標識；ならびに質量タグが挙げられる。

10

【0118】

一部の実施形態では、トランスポゾン核酸はバーコードを含むことが可能である。一部の実施形態では、トランスポソームの集合は、同一のバーコード、1つまたは複数の異なるバーコードを含むトランスポゾン核酸を含むことが可能であるか、または、各トランスポゾン核酸は異なるバーコードを含むことが可能である。一部の実施形態では、標的核酸に挿入または付加したバーコードを用いて、標的核酸を特定することが可能である。一部の実施形態では、バーコードを用いて、標的核酸への挿入事象を特定することが可能である。一部の実施形態では、トランスポソーム集合の各トランスポソームは、標的核酸の挿入部位を特定するのに用いることが可能な、異なるバーコードを有するトランスポゾン核酸を含む。一部の実施形態では、バーコードを用いて、例えば、バーコードが切断部位の両側に位置する切断部位での断片化後に、挿入部位を特定することが可能である。例示的なバーコード、ならびにその調製および使用のための方法は、国際公開第2012/061832号、米国特許出願公開第2012/0208724号明細書、同第2012/0208705号明細書、およびPCT/US2013/031023に記載されており、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。一部の実施形態では、標的核酸に挿入したバーコードは、標的核酸配列の表現を得るための、断片化配列の後続アライメントにおけるランドマークとして有用であり得る。一部の実施形態では、共通のバーコードを含む断片は重複配列を有すると認められる。

20

【0119】

一部の実施形態では、トランスポゾン核酸は、互いに連結する2つのトランスポゾン要素を含むことが可能である。リンカーは挿入物に含まれ得、その結果、第1トランスポゾン要素は第2トランスポゾン要素と連続する。特に有用な挿入物は、国際公開第2012/061832号、米国特許出願公開第2012/0208724号明細書、同第2012/0208705号明細書、およびPCT/US2013/031023（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）に記載されている、「ループ化」複合体を形成するものである。このような構造では、連続トランスポゾン要素を有する単一挿入物は2つのトランスポザーゼサブユニットに結合し、「ループ化」複合体を形成する。一部の実施形態では、トランスポゾン核酸は、ブロッキング基を含むことが可能である。

30

40

【0120】

基材

本明細書で提供する方法および組成物の一部の実施形態には、表面を有する基材の使用を含む。有用な基材としては、例えば固体支持体およびゲルが挙げられる。一部の実施形態では、表面は核酸を結合する。一部の実施形態では、表面は、ワトソン-クリック相補性を介して核酸を表面に結合する複数の捕捉プローブを含む。一部の実施形態では、捕捉プローブはアンカータグを結合する。一部の実施形態では、捕捉プローブおよびアンカータグはそれぞれ核酸を含む。一部の実施形態では、捕捉プローブおよびアンカータグは、本明細書で提供する受容体またはリガンドなど、例えば、ビオチン、アビジン、HisD、ニッケル、抗体、および抗原といった、互いに特異的に結合する小分子基を含む。

50

【0121】

基材は、二次元または三次元とすることが可能であり、平面表面とする（例えば、ガラススライド）、または形付けることが可能である。有用な物質としては、ガラス（例えば、CPG（controlled pore glass））、石英、プラスチック（ポリスチレン（低度架橋ポリスチレンおよび高度架橋ポリスチレン）、ポリカーボネート、ポリプロピレン、ポリ（メチルメタクリレート）など）、アクリルコポリマー、ポリアミド、シリコン、金属（例えば、アルカンチオレート誘導体化金）、セルロース、ナイロン、ラテックス、デキストラン、ゲルマトリックス（例えば、シリカゲル）、ポリアクロレイン、または複合材が挙げられる。適切な三次元固体支持体としては、例えば、球、微粒子、ビーズ、膜、スライド、プレート、マイクロ加工チップ、管（例えば、毛細管）、マイクロウェル、マイクロ流体装置、チャンネル、フィルター、または、核酸もしくは他の捕捉プローブをつなぎ止めるのに適切な任意の他の構造体が挙げられる。固体支持体としては、核酸の集合またはプライマーまたは他の捕捉プローブを含む領域を有することが可能な、平坦なマイクロアレイまたはマトリックスが挙げられる。例としては、ヌクレオシド誘導体化CPGおよびポリスチレンスライド；誘導体化磁気スライド；ポリスチレングリコールをグラフト化したポリスチレンなどが挙げられる。

10

【0122】

種々の組成物および、その組成物を作成し使用する関連方法を用いて、核酸などの捕捉プローブを基材の表面に付加する、つなぎ止める、または固定することが可能である。付加は、表面への直接的または非直接的な結合を介して達成することが可能である。結合は、共有結合的連結によりなすことが可能である（例えば、Joos et al. (1997) Analytical Biochemistry, 247:96-101; Oroskar et al. (1996) Clin. Chem., 42:1547-1555; およびKhandjian (1986) Mol. Bio. Rep., 11:107-11（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）を参照）。好適な付加は、核酸の終端ヌクレオチドを、表面と一体化したエポキシドに直接的にアミン結合することである。結合は、非共有結合的連結を介してもなすことが可能である。例えば、ピオチン - ストレプトアビジン（Taylor et al. (1991) 1. Phys. D: Appl. Phys., 24:1443（これはその全体が参照により本明細書に組み込まれる））および抗ジゴキシゲニン付きジゴキシゲニン（Smith et al., Science, 253: 1122 (1992)（これはその全体が参照により本明細書に組み込まれる））が、核酸を表面につなぎ止めるための一般的なツールである。核酸の表面への付加は、ビーズ、粒子、またはゲルなどの中間構造体を介して行うことが可能である。核酸のゲルを介したアレイへの付加は、Illumina社（カリフォルニア州サンディエゴ）から商業的に入手可能な、または、米国特許出願公開第2010/10111768号明細書；同第2012/0270305号明細書；および国際公開第05/065814号（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）に記載されている、フローセルにより例証される。

20

30

【0123】

一部の実施形態では、基材は、連続の、または一体型の表面を有することが可能である。従って、核酸断片を空間的にランダムな位置に付加することが可能であり、最も近くに隣接する断片（または、最も近くに隣接する断片由来のクラスター）の間の距離は可変であろう。結果として生じるアレイは、可変またはランダムな空間的特徴パターンを有し得る。一部の実施形態では、本明細書に記載の方法で用いる基材としては、反復パターンで存在する捕捉プローブのアレイが挙げられる。一部のこのような実施形態では、捕捉プローブは、核酸が付加することのできる位置を提供する。一部の実施形態では、反復パターンは、六角形パターン、直線状線形パターン、格子パターン、鏡映対称を有するパターン、または回転対称を有するパターンなどである。修飾核酸を付加する相手である捕捉プローブは、それぞれおよそ、 1mm^2 、 $500\text{ }\mu\text{m}^2$ 、 $100\text{ }\mu\text{m}^2$ 、 $25\text{ }\mu\text{m}^2$ 、 $10\text{ }\mu\text{m}^2$ 、 $5\text{ }\mu\text{m}^2$ 、 $1\text{ }\mu\text{m}^2$ 、 500nm^2 、または 100nm^2 以下の、または、前記値のうち任意の2つで定義される範囲の領域を有することが可能である。代わりに、または加えて、各特徴はおよそ、 100nm^2 、 250nm^2 、 500nm^2 、 $1\text{ }\mu\text{m}^2$ 、 $2.5\text{ }\mu\text{m}^2$ 、 $5\text{ }\mu\text{m}^2$ 、 $10\text{ }\mu\text{m}^2$ 、 $100\text{ }\mu\text{m}^2$ 、または $500\text{ }\mu\text{m}^2$ 以上の、または、前記

40

50

値のうち任意の2つで定義される範囲の領域を有することが可能である。アレイ（パターン化されていても、空間的にランダムでも）上での断片の増幅により生じる核酸のクラスタまたはコロニーは、同様に、上記範囲または上記で例示したものから選択した上限および下限に挟まれた範囲の領域を有することが可能である。

【0124】

一部の実施形態では、表面における核酸、捕捉プローブ、または捕捉核酸などの特徴の濃度は、少なくとも、1000特徴/mm²、10000特徴/mm²、100000特徴/mm²、1000000特徴/mm²、または前記値の間の任意の範囲とすることが可能である。一部の実施形態では、表面における核酸、捕捉プローブ、または捕捉核酸などの特徴の濃度は、少なくとも、1000特徴/μm²、10000特徴/μm²、100000特徴/μm²、1000,000特徴/μm²、2000,000特徴/μm²、3000,000特徴/μm²、4000,000特徴/μm²、5000,000特徴/μm²、6000,000特徴/μm²、7000,000特徴/μm²、8000,000特徴/μm²、9000,000特徴/μm²、10,000,000特徴/μm²、20,000,000特徴/μm²、50,000,000特徴/μm²、100,000,000特徴/μm²、または前記値の間の任意の範囲とすることが可能である。

10

【0125】

いくつかの商業的に入手可能なシーケンシングプラットフォームは、配列検出ステップ中の検出試薬（例えば、454 Life Sciences社（スイス国バーゼルのRoche社の子会社）から入手可能なプラットフォームのピロリン酸、またはIon Torrent社（カリフォルニア州カールスバッドのLife Technologies社の子会社）から入手可能なプラットフォームのプラトン）の拡散に対し防壁を提供するウェルを有する基材を利用する。

20

【0126】

本明細書で提供する一部の実施形態には、標的核酸、修飾核酸、またはその断片の一部を増幅するステップを含む。当技術分野で既知である任意の適切な増幅手法を用いることが可能である。一部の実施形態では、核酸断片は基材において、または基材上で増幅する。例えば、一部の実施形態では、核酸断片は、米国特許第5641658号明細書；米国特許出願公開第2002/0055100号明細書；米国特許第7115400号明細書；米国特許出願公開第2004/0096853号明細書；10米国特許出願公開第2004/0002090号明細書；同第2007/0128624号明細書；および同第2008/0009420号明細書（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）の開示により例示されるブリッジ増幅手法を用いて増幅する。

30

【0127】

ブリッジ増幅法は、固定化核酸分子のクラスタ（または「コロニー」）を含むアレイを形成するために、基材において、または基材上に増幅産物を固定することを可能にする。このようなアレイ上の各クラスタまたはコロニーは、複数の同一固定化ポリヌクレオチド鎖および複数の同一固定化相補的ポリヌクレオチド鎖から形成される。そのように形成したアレイは、本明細書では「クラスタ化アレイ」ということが可能である。固相増幅反応の産物が固定化ポリヌクレオチド鎖および固定化相補的鎖のアニール化対により形成され、両鎖が好適には共有結合的付加を介して5'末端で固体支持体に固定されている場合、前記産物はいわゆる「ブリッジ化」構造体である。ブリッジ増幅手法は、固定化核酸鋳型を用いて固定化アンプリコンを生成する方法の例である。他の適切な手法を用いて、本明細書で提供する方法に従って生成した固定化核酸断片から固定化アンプリコンを生成することも可能である。例えば、1つまたは複数のクラスタまたはコロニーは、各増幅プライマー対のプライマーの一方が固定されても、両方が固定されても、固相PCR、固相MDA、固相RCAなどを介して形成することが可能である。

40

【0128】

本明細書に記載する、または当技術分野で一般的に知られる増幅手法のいずれも、ユニバーサルプライマーまたは標的的特異的プライマーとともに利用して固定化DNA断片を増幅することが可能である。適切な増幅法には、ポリメラーゼ連鎖反応（PCR）、鎖置換増幅（SDA）、転写介在増幅（TMA）、および、例えば、米国特許出願公開第800

50

3354号明細書（これは、その全体が参照により本明細書に組み込まれる）に記載されている核酸配列に基づく増幅（N A S B A）が挙げられるが、これらに限定されるわけではない。上記増幅方法を用いて、1つまたは複数の対象核酸を増幅することが可能である。例えば、PCR、マルチプレックスPCR、SDA、TMA、およびN A S B Aなどを利用して、固定化核酸断片を増幅することが可能である。一部の実施形態では、対象の核酸を特異的に対象とするプライマーが増幅反応に含まれる。

【0129】

核酸増幅のための他の適切な方法には、オリゴヌクレオチドの伸長およびライゲーション、ローリングサークル増幅（R C A）（Lizardi et al., Nat. Genet. 19:225-232 (1998)（これはその全体が参照により本明細書に組み込まれる））、ならびにオリゴヌクレオチドライゲーションアッセイ（O L A）（例えば、米国特許第7582420号明細書、同第5185243号明細書、同第5679524号明細書、および同第5573907号明細書；欧州特許第0320308号明細書；同第0336731号明細書；同第0439182号明細書；国際公開第90101069号；同第89/12696号；および同第89109835号（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）を参照）が含まれる。これらの増幅手法は、固定化核酸断片を増幅するように設計することが可能であることが理解されよう。例えば、一部の実施形態では、増幅法には、ライゲーションプローブ増幅、または対象の核酸を特異的に対象とするプライマーを含むオリゴヌクレオチドライゲーションアッセイ（O L A）反応が含まれる。一部の実施形態では、増幅法には、対象の核酸を特異的に対象とするプライマーを含む、プライマー伸長 - ライゲーション反応が含まれる。対象の核酸を増幅するために特異的に設計することが可能なプライマー伸長とライゲーションプライマーの非限定的例として、増幅には、G o l d e n G a t e（登録商標）アッセイ（カリフォルニア州サンディエゴのI l l u m i n a社）、または米国特許第7582420号明細書および同第7611869号明細書（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）に記載の1つまたは複数のアッセイで用いるプライマーを含めることが可能である。

【0130】

等温増幅技法を、本開示の方法で用いることが可能である。例示的な等温増幅法としては、例えばDean et al., Proc. Natl. Acad. Sci. USA 99:5261-66 (2002)により例示されるM D A（Multiple Displacement Amplification）、または、例えば米国特許第6214587号明細書により例示される等温鎖置換核酸増幅が挙げられるが、これらに限定されない（前掲参考文献は、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）。本開示で用いることが可能な他のPCRに基づかない方法としては、例えば、Walker et al., Molecular Methods for Virus Detection, Academic Press, Inc., 1995；米国特許第5455166号明細書、同第5130238号明細書、およびWalker et al., Nucl. Acids Res. 20:1691-96 (1992)に記載の鎖置換増幅（S D A）、または、例えばLage et al., Genome Research 13:294-307 (2003)に記載の多分岐（hyperbranched）鎖置換増幅などが挙げられる（前掲参考文献は、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）。

【0131】

増幅反応、増幅条件、および増幅成分についての追加の記載は、米国特許第7670810号明細書（これはその全体が参照により本明細書に組み込まれる）に記載されている。他の有用な等温増幅技法としては、T w i s t D x社（英国ケンブリッジ）によりT w i s t A m p（商標）キットとして商業的に販売されているものなど、リコンビナーゼ促進増幅技法が挙げられる。リコンビナーゼ促進増幅試薬の有用な成分および反応条件は、米国特許第5223414号明細書および同第7399590号明細書（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）に記載されている。ヘリカーゼ依存性増幅も、例えばXu et al. EMBO Rep 5:795-800 (2004)（これはその全体が参照により本明細書に組み込まれる）に記載されるように、用いることが可能である。

【0132】

一部の実施形態では、再播種 (re-seeding) ステップを実行することが望ましい場合がある。例えば、修飾核酸断片を表面領域内のある位置で捕捉し、それを1つまたは複数のサイクルの増幅プロセスで複製することが可能であり、元の断片および/またはそのアンプリコンは該位置から解放することが可能であり、解放した核酸は同一領域の他の位置で捕捉することが可能であり、そして、新たに捕捉した核酸は増幅することが可能である。特定の例では、表面に播種した断片について単一サイクルのブリッジ増幅を実行することが可能であり、表面からの解放にあたり元の鋳型断片を洗浄する代わりに、鋳型断片を、それがもともと播種されていた位置に近い新しい位置で、表面に再播種することが可能である。ブリッジ増幅の後続ラウンドは、元の播種位置と再播種位置の両方でのクラスタ成長を可能にしよう。このような方法を用いて、複製コロニーを表面領域で生成して、技術的複製 (technical replicate) を提供することが可能である。技術的複製についての配列分析は、エラーチェックという利点を提供することが可能である。例えば、(技術的複製と認められる) 近いクラスタのサブセットのみで起きる観察配列多様体は増幅エラーと認めることが可能であり、ここにおいて、特定の断片の技術的複製と認められる全クラスタで発生する配列多様体は、真の多様体である可能性が高い。

10

【0133】

シーケンシング核酸

本明細書に記載する方法の一部の実施形態には、標的核酸に由来する断片をシーケンシングするステップを含むことが可能である。一例は、合成によるシーケンシング (SBS) である。SBSでは、核酸鋳型 (例えば、標的核酸の断片またはそのアンプリコン) に沿った核酸プライマーの伸長をモニタリングして、鋳型のヌクレオチド配列を決定する。プライマーは、上記の挿入物に存在するプライミング部位にハイブリダイズすることが可能である。基本となる化学プロセスは (例えば、ポリメラーゼ酵素により触媒される) 重合とすることが可能である。ポリメラーゼに基づく特定のSBS実施形態では、蛍光で標識化したヌクレオチドを鋳型に依存するやり方でプライマーに加え (それによりプライマーを伸長し)、その結果、プライマーに加えたヌクレオチドの並び順と型の検出を用いて鋳型配列を決定する。本明細書に記載のステップを用いてアレイの異なる位置に付加した複数の異なる核酸断片を、異なる鋳型で起きる事象をアレイにおけるその位置により識別することが可能な条件下で、SBS技法に供することが可能である。

20

【0134】

一部の実施形態では、本開示の方法により生成し、サイクル中に試薬の反復送達を含むSBSまたは他の検出技法に供する核酸断片のアレイを収容するのに都合のよいフォーマットを、フローセルが提供する。本明細書で用いる場合、「フローセル」は、1つまたは複数の流体試薬を流すことが可能な表面を有するチャンバを含む。概して、フローセルは、流体の流れを容易にする入口開口および出口開口を有するだろう。本開示の方法で容易に用いることが可能なフローセル、ならびに関連する流体系および検出プラットフォームは、例えば、Bentley et al., Nature 456:53-59 (2008)、国際公開第04/018497号; 米国特許第7057026号明細書; 国際公開第91/06678号; 同第071123744号; 米国特許第7329492号明細書; 同第7211414号明細書; 同第7315019号明細書; 同第7405281号明細書、および米国特許出願公開第2008/0108082号明細書 (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) に記載されている。特定の実施形態では、ゲルはフローセルの内面に存在し、該ゲルは、本明細書に記載する1つまたは複数の組成物が付加し、および/または、本明細書に記載する1つまたは複数の方法ステップを行う場所である基材を提供する。

30

40

【0135】

一部の実施形態では、第1SBSサイクルを開始するため、1つまたは複数の標識ヌクレオチド、DNAポリメラーゼなどを、核酸断片のアレイを収容するフローセルに流すことが可能である。(例えば、核酸断片に付加した挿入物に位置するプライミング部位にプライマーをハイブリダイズすることによる) プライマー伸長により、標識ヌクレオチドの取り込みが起きるアレイの部位を、検出することが可能である。オプションとして、ヌク

50

レオチドはさらに、ヌクレオチドをプライマーにいったん加えたらさらなるプライマー伸長を終了させる、可逆終了特性を備えることが可能である。例えば、可逆終了部を有するヌクレオチドアナログをプライマーに加えることが可能であり、その結果、非ブロック剤を送達して該部を取り除くまで、それに続く伸長は起き得ない。従って、可逆終了を用いる実施形態では、非ブロック剤を（検出を行う前後で）フローセルに送達することが可能である。洗浄を、種々の送達ステップの間で行うことが可能である。サイクルはその後、「n」回繰り返してnヌクレオチド分プライマーを伸長することにより、長さ「n」の配列を検出することが可能である。本開示の方法により生成されるアレイで用いるのに容易に適應させることが可能な、例示的なSBS手順、流体系、および検出プラットフォームは、例えば、Bentley et al., Nature 456:53-59 (2008)、国際公開第04/018497号；米国特許第7057026号明細書；国際公開第91/06678号；同第071123744号；米国特許第7329492号明細書；同第7211414号明細書；同第7315019号明細書；同第7405281号明細書、および米国特許出願公開第2008/0108082号明細書に記載されており、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。

【0136】

一部の実施形態では、パイロシーケンシングなど、周期性反応を用いるほかのシーケンシングの手順を用いることが可能である。パイロシーケンシングは、特定のヌクレオチドが新生核酸鎖に取り込まれる際の無機ピロリン酸（PPi）の放出を検出する（Ronaghi, et al., 1996, Analytical Biochemistry 242(1), 84-9; Ronaghi, 2001, Genome Res. 11(1), 3-11; Ronaghi et al., 1998, Science 281(5375), 363; 米国特許第6210891号明細書；同第6258568号明細書、および同第6274320号明細書（これらはそれぞれ参照により本明細書に組み込まれる））。パイロシーケンシングでは、放出されたPPiを、ATPスルフィラーゼによりアデノシン三リン酸（ATP）に変換することにより検出することが可能であり、生成されたATP量はルシフェラーゼ生成光子を介して検出することが可能である。従って、シーケンシング反応は発光検出系によりモニタリングすることが可能である。蛍光に基づく検出系で用いる励起放射線源は、パイロシーケンシングの手順には必要ではない。パイロシーケンシングを本開示の方法に適用するために用いることが可能な有用な流体系、検出器および手順は、例えば、国際公開第2012058096号、米国特許出願公開第2005/0191698号明細書、米国特許第7595883号明細書、および同第7244559号明細書に記載されており、これらはそれぞれ、その全体が参照によって本明細書に組み込まれる。ライゲーションによるシーケンシング反応も有用であり、該反応としては、例えばShendure et al. Science 309:1728-1732 (2005)；米国特許第5599675号明細書；および同第5750341号明細書（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）に記載されているものが挙げられる。一部の実施形態は、例えば、Bains et al., Journal of Theoretical Biology 135(3),303-7 (1988)；Drmanac et al., Nature Biotechnology 16,54-58 (1998)；Fodor et al., Science 251(4995), 767-773 (1995)；および国際公開第1989110977号（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）に記載されているような、ハイブリダイゼーションによるシーケンシングの手順を含み得る。

【0137】

一部の実施形態、ライゲーションによるシーケンシングおよびハイブリダイゼーションによるシーケンシングの手順では、アレイのある部位に存在する標的核酸断片（またはそのアンプリコン）を、オリゴヌクレオチド送達と検出の反復サイクルに供する。本明細書または本明細書で言及する参考文献に記載のSBS法のための流体系は、ライゲーションによるシーケンシングまたはハイブリダイゼーションによるシーケンシングの手順での試薬の送達に、容易に適應させることが可能である。典型的には、オリゴヌクレオチドは蛍光的に標識化し、本明細書または本明細書で言及する参考文献においてSBSに関連して記載されるものに類似する蛍光検出器を用いて検出することが可能である。

【 0 1 3 8 】

一部の実施形態では、DNAポリメラーゼ活性をリアルタイムでモニタリングするステップを含む方法を利用することが可能である。例えば、ヌクレオチドの取り込みは、フルオロフォア担持ポリメラーゼと γ -ホスフェート標識ヌクレオチドの間の蛍光共鳴エネルギー移動 (FRET) 相互作用を介して、またはZMW (zeromode waveguides) を用いて検出することが可能である。FRETに基づくシーケンシング用の技法および試薬は、例えばLevene et al. Science 299, 682-686 (2003); Lundquist et al. Opt. Lett. 33, 1026-1028 (2008); およびKorlach et al. Proc. Natl. Acad. Sci. USA 105, 1176-1181 (2008)に記載されており、これらの開示内容はその全体が参照により本明細書に組み込まれる。

10

【 0 1 3 9 】

一部のSBS実施形態には、ヌクレオチドが伸長産物に取り込まれる際に放出される光子の検出を含む。例えば、放出された光子の検出に基づくシーケンシングは、Ion Torrent社(コネチカット州ギルフォードのLife Technologies社の子会社)より商業的に入手可能な電気検出器および関連技法、または、米国特許出願公開第2009/10026082号明細書、同第2009/10127589号明細書、同第2010/10137143号明細書、もしくは同第2010/10282617号明細書(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)に記載されているシーケンシング法およびシーケンシングシステムを用いることが可能である。

【 0 1 4 0 】

一部の実施形態では、本方法のシーケンシングステップは、Deamer & Akesson Trends Biotechnol. 18, 147-151 (2000); Deamer & Branton, Acc. Chem. Res. 35:817-825 (2002); およびLi et al., Nat. Mater. 2:611-615 (2003) (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)に記載されているものなどの、ナノポアシーケンシング技法を含むことが可能である。このような実施形態では、標的核酸断片をナノポアに通過させる。ナノポアは、合成ポアまたは α -溶血素などの生物学的膜たんぱく質とすることが可能である。標的核酸がナノポアを通過する際、ポアの電気伝導度の変動を測定することにより各塩基対を特定することが可能である(米国特許第7001792号明細書; Soni & Meller Clin. Chem. 53, 1996-2001 (2007); Healy, Nanomed. 2:459-481 (2007); およびCockroft et al., 1. Am. Chem. Soc. 130:818-820 (2008) (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))。一部の実施形態では、個々のナノポアの位置は、本明細書で例示するアレイ上の部位または特徴に類似する。ナノポアの互いの近接度は、それらが読み取る断片配列の近接度と相関し、例えば、その断片を、それらが由来するより大きい配列にアセンブリすることを容易にする。

20

30

【 0 1 4 1 】

一部の実施形態では、本明細書に記載のシーケンシングステップは、マルチプレックスフォーマットで有利に実行することが可能であり、その結果、多数の異なる標的核酸を同時に操作する。特定の実施形態では、異なる標的核酸を、一般的な反応容器または特定の基材の表面で処置することが可能である。これは、シーケンシング試薬を都合よく送達すること、未反応試薬の除去、および取り込み事象を多重に検出することを可能にする。表面結合標的核酸またはその断片を用いる実施形態では、標的核酸または断片は、アレイ形式とすることが可能である。アレイ形式では、標的核酸の断片は、典型的には、空間的に識別可能な方法で、例えば、本明細書に記載の付加技法を用いて、表面に結合させることが可能である。アレイが各部位(特徴ともいう)で標的核酸断片の単一コピーを含むか、または、同一の配列を有する多数のコピーが各部位つまり特徴で存在することが可能である。ブリッジ増幅またはエマルジョンPCRなどの増幅法により多数のコピーを生成することが可能である。

40

【 0 1 4 2 】

核酸の調製とシーケンシング

本明細書で提供する組成物および方法の一部の実施形態には、標的核酸からシーケンシ

50

ングライブラリを調製するステップを含む。一部の実施形態には、調製したライブラリをシーケンシングするステップも含む。一部の実施形態では、各トランスポソームがトランスポザゼおよびトランスポゾン核酸を含む、トランスポソームの集合を標的核酸に接触させる。接触は基材において、または基材上で、または代わりに溶液中で行うことが可能である。トランスポソームは、複数の部位で標的核酸にニックを入れ、ニック入り鎖のニック部位の一方の側に単一トランスポゾン核酸を付加するように、片側トランスポザゼ活性を含むことが可能である。一部の実施形態では、付加したトランスポゾン核酸のそれぞれにプライマーをハイブリダイズし、伸長させて、一本鎖修飾核酸の集合を得ることが可能である。一部の実施形態では、伸長させた核酸を増幅することが可能である。一部の実施形態では、伸長させたおよび/または増幅した核酸、つまり修飾核酸を、シーケンシングのために表面上に捕捉することが可能である。一部の実施形態には、捕捉した核酸をシーケンシングするステップを含む。

10

【0143】

図1には、トランスポゾン核酸を含むトランスポソームの集合に標的核酸を接触させる、例示的な実施形態を描く。複数の部位で標的核酸にニックを入れ、ニック入り標的核酸の一方の鎖にあるニック部位の一方の側にトランスポゾン核酸を付加する。プライマーを、付加したトランスポゾン核酸にハイブリダイズして、伸長した核酸の集合を提供する。一部の実施形態では、伸長した核酸を増幅することが可能である。一部の実施形態では、伸長した核酸はシーケンシングライブラリのための鋳型を提供する。

【0144】

20

一部の実施形態では、片側トランスポザゼ活性を有するトランスポソームは、片側トランスポザゼ活性を有するトランスポザゼを含む。一部の実施形態では、トランスポソームはブロック化トランスポゾン核酸を含む。標的核酸からライブラリを調製しシーケンシングする方法および組成物に有用なトランスポソームを、本明細書で提供する。一部の実施形態では、トランスポゾン核酸はアンカー部位、バーコード、シーケンシングプライマー部位、増幅プライマー部位、および/またはレポータータグを含む。図2には、異なるトランスポゾン核酸を含むトランスポソームの集合を標的核酸に接触させ、異なるトランスポゾン核酸を、標的核酸の鎖の異なるニック部位に付加する例示的な実施形態を描く。一部の実施形態では、異なるトランスポゾン核酸は、異なるアンカー部位、異なるバーコード、異なるシーケンシングプライマー部位、異なる増幅プライマー部位、および/または異なるレポータータグを含むことが可能である。

30

【0145】

一部の実施形態では、伸長した核酸を増幅する。一部の実施形態では、増幅はテイル化増幅プライマーを用いる。テイル化プライマーは追加の末端配列を含み得、その結果、該追加配列は増幅産物に含まれる。一部の実施形態では、増幅プライマーには、アンカー部位、シーケンシングプライマー部位、増幅プライマー部位、およびレポータータグが含まれ得る。図3には、修飾標的核酸を線形増幅により増幅して、ある増幅産物を得る、例示的な実施形態を描く。

【0146】

図4には、トランスポソームがダイマートランスポザゼを含み、トランスポゾン核酸は、モザイク要素(ME)を含む2つのトランスポゾン要素を含み、該MEの1つは3'末端においてジデオキシ基でブロックされている、例示的な実施形態を描く。一部の実施形態では、トランスポゾン核酸は、2つのトランスポゾン要素の間に切断可能なリンカーを含む。トランスポゾン核酸は切断することが可能であり、トランスポゾン核酸の非ブロック化断片は、ニック入り標的核酸鎖のニック部位に付加することが可能である。

40

【0147】

一部の実施形態では、修飾核酸を表面上に捕捉する。一部の実施形態では、表面は複数の捕捉プローブを含む。一部の実施形態では、捕捉プローブは核酸を含む。一部の実施形態では、捕捉プローブは修飾核酸に特異的にハイブリダイズする。一部の実施形態では、捕捉プローブは修飾核酸の親和性部に結合する親和性部を含む。一部の実施形態では、捕

50

捉核酸を、例えばブリッジ増幅により増幅する。一部の実施形態では、捕捉核酸を表面上でシーケンシングする。

【0148】

本明細書で提供する方法および組成物の一部の実施形態には、バーコードを含むシーケンシングライブラリを調製することを含む。一部の実施形態には、このようなライブラリをシーケンシングすることを含む。一部の実施形態では、バーコードは、シーケンシングした標的核酸断片のアライメントで有用なランドマークを提供する。一部の実施形態では、片側転移およびライゲーションにより、トランスポゾン核酸を標的核酸の単一鎖に挿入する。修飾標的核酸を増幅し、断片をシーケンシングする。重複する断片は共通の挿入を含み得、これは、標的核酸配列の表現を生成するための、シーケンシングした断片のアライメントにおいて有用である。図5には、異なるバーコードを含むトランスポソーム集合を標的核酸に接触させ、トランスポゾン核酸をニック部位の一方の側に付加し、付加されていないもう一方のトランスポゾン核酸末端をライゲーションによりニック部位のもう一方の側に付加し、修飾標的核酸を全ゲノム増幅(WGA)により増幅する、例示的な実施形態を描く。

10

【0149】

一部の実施形態では、片側トランスポザーゼ活性を有するトランスポソームの集合を、標的核酸に接触させる。トランスポソームは、標的核酸の鎖に挿入されるトランスポゾン核酸を含む。このような実施形態で有用なトランスポソームを本明細書に記載する。一部の実施形態では、標的核酸をトランスポソームに接触させ、その結果、複数の部位で標的核酸にニックを入れ、ニック入り鎖のニック部位の一方の側に単一トランスポゾン核酸を付加し、ニック入り鎖のニック部位のもう一方の側に付加した単一トランスポゾン核酸をライゲーションすることにより、トランスポゾン核酸を二本鎖標的核酸の単一鎖に挿入する。一部の実施形態では、リガーゼとして非相同末端接合リガーゼが挙げられる。一部の実施形態では、リガーゼとしてリガーゼI Vが挙げられる。一部の実施形態では、修飾核酸を増幅する。一部の実施形態では、修飾核酸を表面上に捕捉する。一部の実施形態では、修飾核酸をシーケンシングする。一部の実施形態では、修飾核酸の配列を、重複配列で共通するバーコードの存在に従ってアライメントを行う。一部の実施形態には、本明細書で提供する方法により調製した、バーコードを含むシーケンシングライブラリを含む。

20

【0150】

ハプロタイプ情報の取得

ゲノムDNAなどの標的核酸は複数のハプロタイプを含み得る。例えば、ヒトゲノムDNAは2セットのDNA分子を含み、各セットは母系配列と父系配列の異なる組み合わせを有する。本明細書で提供する一部の実施形態は、単一の核酸分子の断片またはそのコピーから配列情報を得るのに有用である。

30

【0151】

一部の実施形態では、基材上にある断片の物理的な近接度は維持される。一部の実施形態では、線形標的核酸の配列において互いの近接度がより高い断片配列は、線形標的核酸の配列において互いの近接度がより低い断片配列と比較し、表面上での互いの物理的近接度はより高い。ある断片の物理的近接度は種々の方法により保持することが可能である。

40

【0152】

一部の実施形態では、片側転移は標的核酸を断片化しない。一部の実施形態では、片側トランスポザーゼ活性を有するトランスポソームに標的核酸を接触させて、修飾核酸を得ることが可能である。一部の実施形態では、修飾核酸を表面に接触させることが可能である。一部の実施形態では、トランスポゾン核酸はアンカータグを含み、その結果、捕捉プローブを含む表面上に修飾配列を捕捉することが可能である。一部の実施形態では、修飾核酸を表面との接触中に断片化することが可能である。一部の実施形態では、修飾核酸を表面に近い位置で断片化することが可能である。一部の実施形態では、修飾核酸を表面上でシーケンシングすることが可能である。

【0153】

50

一部の実施形態では、ハプロタイプ情報を得る方法には、表面上の近接位置について求めた相補的な配列を比較して配列エラーを特定するステップが含まれる。一部の実施形態では、表面上の任意の2つの断片種の相対近接度は、2つの断片から得られる配列情報のアライメントに有用な情報を提供し得る。具体的には、表面上の任意の2つの所定断片に由来するクラスタ間の距離は、国際公開第2012/025250（これはその全体が参照により本明細書に組み込まれる）でより詳細に記載されるように、該2つのクラスタが同一の標的ポリヌクレオチド分子に由来する可能性と正の相関を持ち得る。

【0154】

例として、一部の実施形態では、フローセルの表面上に捕捉した長い核酸分子に由来する断片は、フローセルの表面全体に一系列になって（例えば、断片化または増幅の前に核酸を引き延ばした場合）、または表面上で塊になって発生する。さらに、その後、固定化核酸の物理地図を生成することが可能である。従って物理地図は、固定化核酸を増幅した後のクラスタの物理的關係と相関を持つ。具体的には、国際公開第2012/025250号の組み込まれた資料に記載されているように、物理地図を用いて、任意の2つのクラスタから得た配列データが関連している可能性を計算する。

【0155】

一部の実施形態では、表面を撮像して表面全体の固定化核酸分子の位置を確定することにより物理地図を生成する。一部の実施形態では、造影剤を固定支持体に加えて、造影剤からのシグナルを検出することにより固定化核酸を撮像する。一部の実施形態では、造影剤は検出可能な標識である。適切な検出可能標識としては、プロトン、ハプテン、放射性核種、酵素、蛍光標識、化学発光標識、および/または発色剤が挙げられるが、これらに限定されない。例えば、一部の実施形態では、造影剤はインターカレーティング色素または非インターカレーティングDNA結合剤である。限定されることはないが、米国特許出願公開第2012/0282617号明細書（これはその全体が参照により本明細書に組み込まれる）に記載されているものなど、当技術分野で既知である任意の適切なインターカレーティング色素または非インターカレーティングDNA結合剤を用いることが可能である。

【0156】

ある実施形態では、複数の修飾核酸分子を複数のナノチャネルを含むフローセルに流す。本明細書で用いる場合、ナノチャネルという用語は、長い線形核酸分子を引き延ばす細いチャネルを意味する。一部の実施形態では、鎖の数は、1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、もしくは1000以下の、各ナノチャネルに沿って引き延ばした長い個別の核酸鎖か、または、前記値のうち任意の2つにより定義された範囲である。一部の実施形態では、個々のナノチャネルは、長い個別の標的核酸鎖が多数のナノチャネルと相互作用しないようにする物理的バリアにより分離する。一部の実施形態では、固体支持体は少なくとも、10、50、100、200、500、1000、3000、5000、10000、30000、50000、80000、または100000のナノチャネル、または前記値のうち任意の2つにより定義される範囲を含む。

【0157】

一部の実施形態では、いったん修飾核酸をチャネルに沿って引き延ばしたら、該核酸を切断する。オプションとして、結果として生じる断片を増幅してチャネルの表面に沿ってクラスタを形成することが可能である。その後、例えば、これらのチャネルの1つの長さに沿ってクラスタを辿ることにより近接マッピングを実行することが可能である。例として、マッピングされた固定化断片化産物を中に有する1000以上のナノチャネルを備えたフローセルを用いて、「位置決めされた」短いリードを用いて生物のゲノムをシーケンシングすることが可能である。一部の実施形態では、ナノチャネル内のマッピングされた固定化断片化産物を用いてハプロタイプを分析することが可能である。一部の実施形態では、ナノチャネル内のマッピングされた固定化断片化産物を用いて、フェージング問題を解決することが可能である。

【 0 1 5 8 】

一部の実施形態では、片側転移を用いて、人工DNAをgDNAに挿入する。一部の実施形態では、人工DNAをゲノムDNA（または他の核酸）の反復領域に挿入して反復領域をユニークなものにする。反復領域を、例えば上記のようなシーケンシング技法により分析して、反復の数を数えるか、または、反復領域と関連するゲノムDNAの他の配列の位置を確認することが可能である。別の例では、片側転移により挿入した人工DNAは、二本鎖核酸の上下鎖を異ならせる。従って、挿入法の産物は、例えば上記のようなシーケンシング技法により分析して、一方の鎖をもう一方と識別することが可能である。これはさらに、ゲノムDNA（または他の二本鎖核酸）の復元において、上下鎖の単独アセンブリを可能にする。

10

【実施例】

【 0 1 5 9 】

実施例 1 - ブロック化トランスポゾン核酸

標的DNAを、トランスポソームなし（アンプリコン）、（1）トランスポザーゼと、3'ビオチン基でブロックされたトランスポゾン核酸とを含むトランスポソーム（3' Bio）、（2）トランスポザーゼと、3'スパーサ基でブロックされたトランスポゾン核酸とを含むトランスポソーム（3'スパーサ）、または（3）トランスポザーゼと、非ブロック化トランスポゾン核酸とを含むトランスポソーム（TDE1）で処置した。図6には、ブロック化トランスポゾン核酸を含むトランスポソーム（3' Bioおよび3'スパーサ）では転移が起きないという結果を描く。

20

【 0 1 6 0 】

実施例 2 - ランドマーク挿入とアセンブリのモデル

12bpのランダムな配列を含むランドマークを標的DNAに挿入した。DNAをシーケンシングし、配列断片をde novoアセンブリした。図7には、挿入物頻度が100bpである500bpリードおよび挿入物頻度が50bpである300bpリードの、名目カバレッジ深度と平均合成リード長のグラフを示す。6~7kbが50Xカバレッジでde novoにアセンブリ可能であることが示された。

【 0 1 6 1 】

実施例 3 - グリセロールありおよびなしでの片側転移

標的DNA（pUC19）をトランスポソームでインキュベーションし、DNA産物を1.2%ゲル上で分離させた。10サンプルを実行した：最初の5サンプルはグリセロールを含まず、次の5サンプルは65%のグリセロールを含んだ。5サンプルの各セットを、異なる濃度のトランスポソームの滴定としてセットアップした。トランスポソームはトランスポザーゼと非ブロック化トランスポゾン核酸とからなる（TDE1）。着色したゲルの写真を図9に示す。ゲルには、非カット（uncut）pUC19（pUC19のみ）、線形化（linearized）pUC19（EcoRI）、および一本鎖ニック入り（single strand nicked）pUC19（NbsrDI）という対照も搭載した。「グリセロールなし」のサンプルを搭載したゲルのレーンで示されるように、TDEの濃度の増加は、片側転移（つまり、ニック入り）産物と、両側転移（つまり、線形）産物の増加をもたらした。比較すると、65%グリセロールの存在下で行った反応は、TDE1の増加とともに片側転移産物の量が増加することを示したが、TDE1の濃度を増加させても両側転移産物の増加はほとんどまたは全く見られなかった。

30

40

【 0 1 6 2 】

実施例 4 - 移動鎖のトランスポゾン核酸の長さの変更が転移を阻害する

この例では、一ヌクレオチドの減算（ $n - 1$ ）または一ヌクレオチドの追加（ $n + 1$ ）によるトランスポゾンの移動鎖の長さの変更が、転移の効率を下げることを示す。

【 0 1 6 3 】

トランスポソームを3' $n - 1$ および3' $n + 1$ のMETSトランスポゾンで形成し、室温で一晩、0、1%、5%、50%、90%、99%、または100%のTDE1にハイブリダイズした。結果として生じたトランスポソームをその後、1kbのアンプリコンとともに一晩室温で反

50

応させ、続いてSDSで処置し、その後、TBEゲル上で分離させた。図10は、分子量ラダーに沿った反応産物と、トランスポザーゼ酵素なしの対照サンプルとを搭載したTBEゲルを示す。驚くべきことに、一晚インキュベーションしても、標的DNAの大部分はまだ各サンプル中に存在し、 $n-1$ および $n+1$ のトランスポゾンが転移に対し阻害効果を有することを示した。さらに、阻害効果は、 $n-1$ および $n+1$ のトランスポゾンのパーセンテージの増加と関連した。

【0164】

本明細書で用いる場合、「含んで (comprising)」という用語は、「含んで (including)」、「含有して (containing)」、または「～を特徴とする (characterized by)」と同義であり、包括的つまり非限定 (open-ended) であり、追加の言及されていない要素または方法ステップを除外しない。

10

【0165】

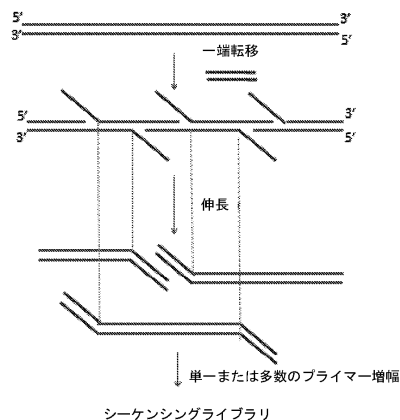
上記記載は、本発明のいくつかの方法および物質を開示する。本発明は、方法および物質の修飾、ならびに、組み立て方法および装置の変更を許す。このような修飾は、本明細書で開示する発明のこの開示内容または実践についての考慮から、当業者には明白になる。その結果として、本発明は本明細書で開示する特定の実施形態に限定されることを意図するものではなく、本発明の真の範囲および趣旨内にある修飾および代替を全て包含することを意図する。

【0166】

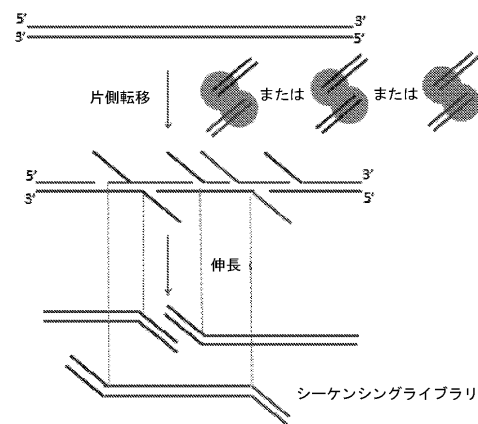
本明細書で言及する全ての参考文献には、限定されるわけではないが、公開出願、非公開出願、特許、および参考文献が含まれ、これらはその全体が参照により本明細書に組み込まれ、それにより本明細書の一部とする。参照により組み込まれる刊行物および特許または特許出願が本明細書に含む開示内容と矛盾する限り、本明細書は任意のこのような矛盾物に取って代わるおよび/または優先されることを意図する。

20

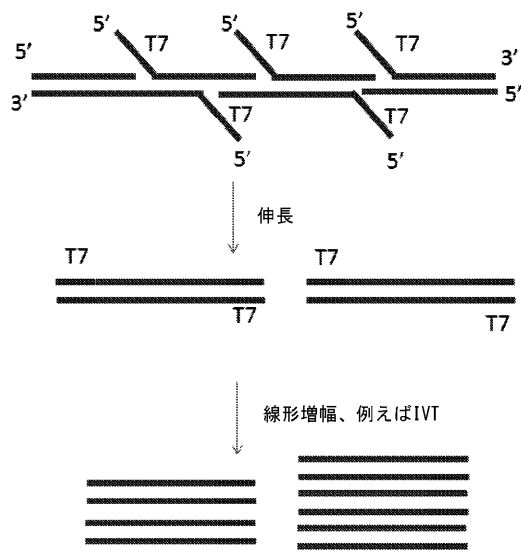
【図1】



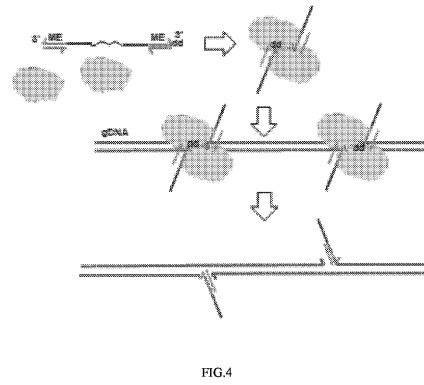
【図2】



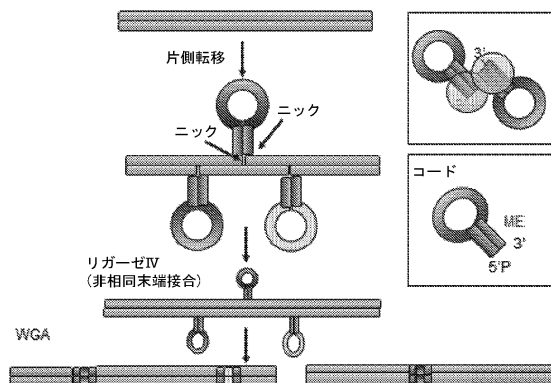
【図 3】



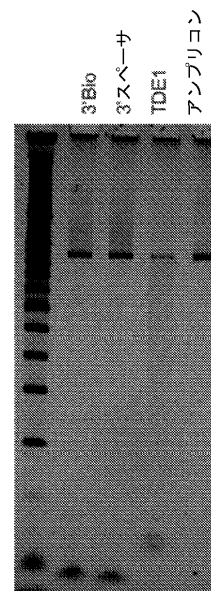
【図 4】



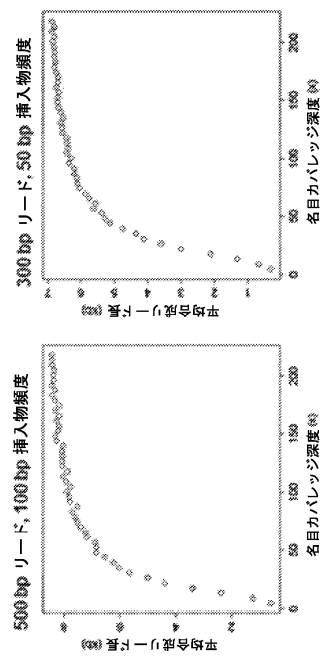
【図 5】



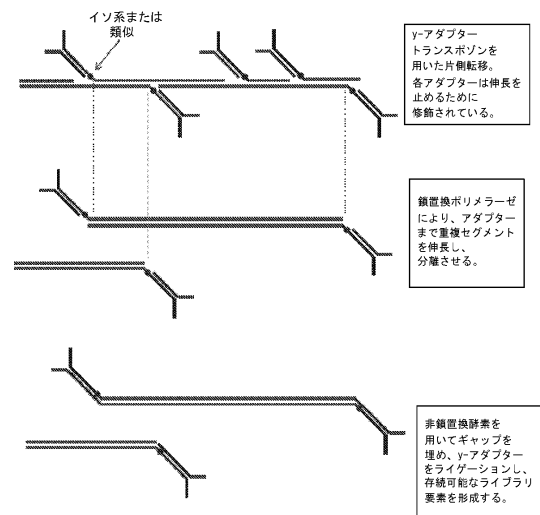
【図 6】



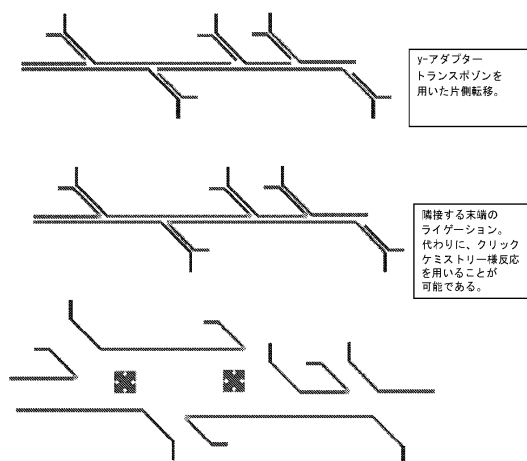
【図 7】



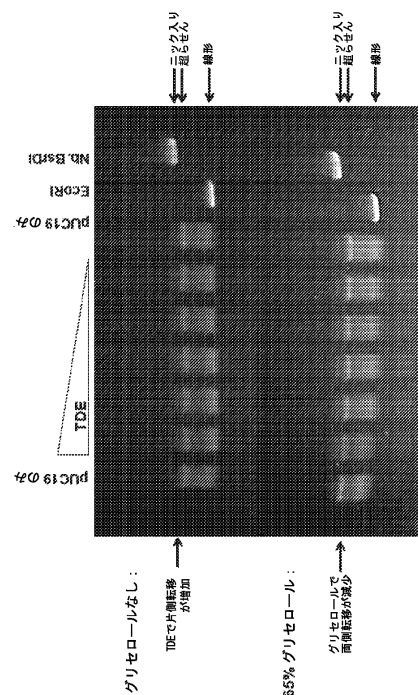
【図 8 A】



【図 8 B】



【図 9】



【図 14】

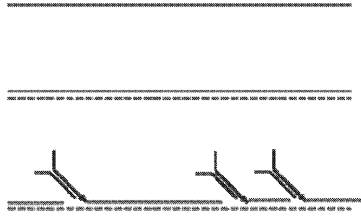
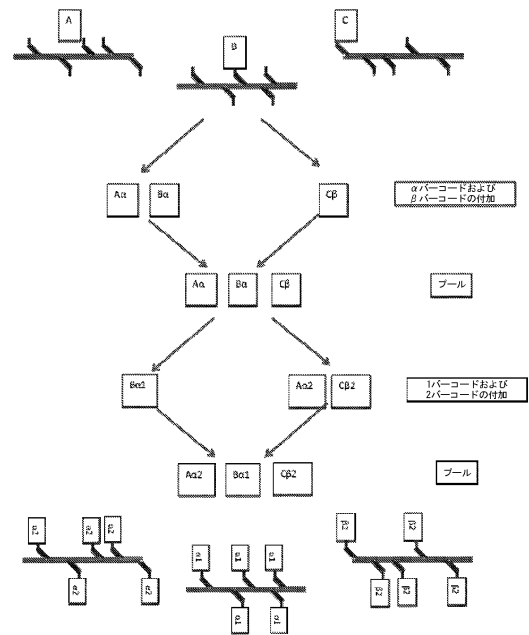


FIG. 14

【図 15】



フロントページの続き

- (72)発明者 ジェフリー エス フィッシャー
アメリカ合衆国 カリフォルニア州 9 2 1 2 2 サンディエゴ イラミーナ ウェイ 5 2 0 0
- (72)発明者 ケビン エル ガンダーソン
アメリカ合衆国 カリフォルニア州 9 2 1 2 2 サンディエゴ イラミーナ ウェイ 5 2 0 0
- (72)発明者 ササン アミニ
アメリカ合衆国 カリフォルニア州 9 4 0 6 1 レッドウッド シティ ポプラー アベニュー
6 2 6
- (72)発明者 クリスチアン グロックナー
ドイツ国 5 3 1 1 1 ボン ヴォルフシュトラッセ 1 0 エイ

審査官 池上 文緒

- (56)参考文献 特表 2 0 1 2 - 5 0 6 7 0 4 (J P , A)
国際公開第 2 0 1 3 / 1 3 1 9 6 2 (W O , A 1)
国際公開第 2 0 1 2 / 1 0 3 5 4 5 (W O , A 1)

- (58)調査した分野(Int.Cl. , DB名)
C 1 2 N 1 5 / 0 0 - 1 5 / 9 0
C 4 0 B 5 0 / 0 6
C 1 2 Q 1 / 6 8 6 9
J S T P l u s / J M E D P l u s / J S T 7 5 8 0 (J D r e a m I I I)
C A p l u s / M E D L I N E / E M B A S E / B I O S I S (S T N)
W P I D S / W P I X (S T N)