

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7532455号
(P7532455)

(45)発行日 令和6年8月13日(2024.8.13)

(24)登録日 令和6年8月2日(2024.8.2)

(51)国際特許分類		F I	
C 1 2 Q	1/6874(2018.01)	C 1 2 Q	1/6874 Z
C 4 0 B	40/06 (2006.01)	C 4 0 B	40/06
C 1 2 N	15/10 (2006.01)	C 1 2 N	15/10 Z

請求項の数 20 外国語出願 (全66頁)

(21)出願番号	特願2022-129859(P2022-129859)	(73)特許権者	502279294
(22)出願日	令和4年8月17日(2022.8.17)		イルミナ ケンブリッジ リミテッド
(62)分割の表示	特願2020-204004(P2020-204004))の分割		英国シービー21・6ディエフ、ケンブリッジシャー、ケンブリッジ、グレート・アビントン、グランタ・パーク19番
原出願日	平成27年10月16日(2015.10.16)	(74)代理人	100147485
(65)公開番号	特開2022-172158(P2022-172158A)		弁理士 杉村 憲司
(43)公開日	令和4年11月15日(2022.11.15)	(72)発明者	フランク ジェイ スティーマーズ
審査請求日	令和4年9月14日(2022.9.14)		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
(31)優先権主張番号	62/065,544	(72)発明者	ケヴィン エル ガンダーソン
(32)優先日	平成26年10月17日(2014.10.17)		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
(33)優先権主張国・地域又は機関	米国(US)		
(31)優先権主張番号	62/157,396		
(32)優先日	平成27年5月5日(2015.5.5)		
	最終頁に続く		最終頁に続く

(54)【発明の名称】 連続性を維持した転位

(57)【特許請求の範囲】
【請求項1】

(a) それぞれその上に複数のトランスポソーム複合体を固定化している複数の固体支持体を提供するステップであって、前記トランスポソーム複合体の各々は2つ以上のトランスポソームの単量体単位を含み、前記トランスポソームの単量体単位が互いに結合してトランスポソーム複合体を形成しており、前記トランスポソームの各単量体単位が、ポリヌクレオチドに結合したトランスポザーゼを含み、前記ポリヌクレオチドは、

(i) トランスポゾン末端配列を含む3'部分と、
(ii) 第1のバーコード配列を含む第1のアダプターとを含む、ステップと、

(b) 二本鎖の標的核酸を、前記二本鎖の標的核酸が前記トランスポソーム複合体により二本鎖の標的核酸フラグメントにフラグメント化され、前記ポリヌクレオチドの前記トランスポゾン末端配列が前記二本鎖の標的核酸フラグメントの少なくとも1つの鎖の5'末端に転移される条件下で、前記複数の固体支持体に適用し、それにより、少なくとも1つの鎖が前記第1のバーコード配列で5'タグ化されている二本鎖の標的核酸フラグメントの固定化ライブラリーを作製するステップと、を含む、タグ化核酸フラグメントの固定化ライブラリーを調製する方法。

【請求項2】

(a) それぞれその上に複数のトランスポソーム複合体を固定化している複数の固体支持体を提供するステップであって、前記トランスポソーム複合体の各々は2つ以上のトラ

ンスポソームの単量体単位を含み、前記トランスポソームの単量体単位が互いに結合してトランスポソーム複合体を形成しており、前記トランスポソームの各単量体単位が、ポリヌクレオチドに結合したトランスポザーゼを含み、前記ポリヌクレオチドは、

(i) トランスポゾン末端配列を含む3'部分と、

(ii) 第1のバーコード配列を含む第1のアダプターと

を含む、ステップと、

(b) 二本鎖の標的核酸を、前記標的核酸の連続性を維持しながら、前記二本鎖の標的核酸が前記トランスポソーム複合体により二本鎖の標的核酸フラグメントにフラグメント化され、前記ポリヌクレオチドの前記トランスポゾン末端配列が前記二本鎖の標的核酸フラグメントの少なくとも1つの鎖の5'末端に転移される条件下で、前記複数の固体支持体に適用し、それにより、少なくとも1つの鎖が前記第1のバーコード配列で5'タグ化されている二本鎖の標的核酸フラグメントの固定化ライブラリーを作製するステップであり、ある固体支持体上の前記第1のバーコード配列は同じ核酸配列を含み、前記複数の固体支持体中のある固体支持体からの第1のバーコードの核酸配列は、前記複数の固体支持体中の他の固体支持体からの第1のバーコードの核酸配列とは異なる、ステップと、

(c) 前記標的核酸のフラグメントの配列及び関連した前記第1のバーコード配列を決定するステップと、

(d) 前記第1のバーコード配列を識別することにより、前記標的核酸の連続性情報を決定するステップと

を含む、標的核酸配列の連続性情報を決定する方法。

【請求項3】

トランスポソーム単量体単位の前記トランスポザーゼが、同じ前記トランスポソーム複合体の別のトランスポソーム単量体単位の別のトランスポザーゼに結合している、請求項1又は2に記載の方法。

【請求項4】

トランスポソーム単量体単位の前記トランスポゾン末端配列が、同じ前記トランスポソーム複合体の別のトランスポソーム単量体単位の別のトランスポゾン末端配列に結合している、請求項1又は2に記載の方法。

【請求項5】

前記トランスポソーム複合体中の前記トランスポザーゼのサブユニットが、共有及び非共有結合により互いに結合している、請求項1又は2に記載の方法。

【請求項6】

前記トランスポソーム単量体単位の前記トランスポザーゼが、トランスポゾンが加わってトランスポソーム複合体が形成される前に互いに結合する、請求項1又は2に記載の方法。

【請求項7】

前記トランスポソーム単量体単位の前記トランスポザーゼが、トランスポソーム複合体の形成後に互いに結合する、請求項1又は2に記載の方法。

【請求項8】

前記第1のアダプターが、第1のプライマー結合配列をさらに含む、請求項1又は2に記載の方法。

【請求項9】

標的核酸配列の前記連続性情報が、ハプロタイプ情報を示す、請求項2に記載の方法。

【請求項10】

標的核酸配列の前記連続性情報が、欠損、転位、染色体間遺伝子融合、重複、及びパラログからなる群から選択されるゲノム変異を示す、請求項2～9のいずれか一項に記載の方法。

【請求項11】

前記バーコードを含む標的核酸フラグメントが、前記標的核酸フラグメントの配列を決定する前に増幅される、請求項2～10のいずれか一項に記載の方法。

【請求項 1 2】

前記標的核酸をトランスポソーム複合体と接触させる前に、前記標的核酸をプレフラグメント化するステップをさらに含む、請求項 1 ～ 1 1 のいずれか一項に記載の方法。

【請求項 1 3】

前記標的核酸が単一の細胞から得られる、請求項 1 ～ 1 2 のいずれか一項に記載の方法。

【請求項 1 4】

前記標的核酸が単一の細胞小器官から得られる、請求項 1 ～ 1 3 のいずれか一項に記載の方法。

【請求項 1 5】

前記標的核酸がゲノム DNA を含む、請求項 1 ～ 1 4 のいずれか一項に記載の方法。

【請求項 1 6】

前記標的核酸が他の核酸に架橋する、請求項 1 ～ 1 5 のいずれか一項に記載の方法。

【請求項 1 7】

前記標的核酸が無細胞腫瘍 DNA を含む、請求項 1 ～ 1 6 のいずれか一項に記載の方法。

【請求項 1 8】

前記標的核酸が c DNA を含む、請求項 1 ～ 1 7 のいずれか一項に記載の方法。

【請求項 1 9】

前記固体支持体がビーズを含む、請求項 1 ～ 1 8 のいずれか一項に記載の方法。

【請求項 2 0】

(a) それぞれその上に複数のトランスポソーム複合体を固定化している複数の固体支持体であって、前記トランスポソーム複合体の各々は 2 つ以上のトランスポソームの単量体単位を含み、前記トランスポソームの単量体単位が互いに結合してトランスポソーム複合体を形成しており、前記トランスポソームの各単量体単位が、ポリヌクレオチドに結合したトランスポザーゼを含み、前記ポリヌクレオチドは、

(i) トランスポゾン末端配列を含む 3' 部分と、

(ii) 第 1 のバーコード配列を含む第 1 のアダプターとを含む、複数の固体支持体と、

(b) 前記固体支持体に固定化された二本鎖の標的核酸であって、前記二本鎖の標的核酸は前記トランスポソーム複合体により二本鎖の標的核酸フラグメントにフラグメント化されており、前記ポリヌクレオチドの前記トランスポゾン末端配列が前記二本鎖の標的核酸フラグメントの少なくとも 1 つの鎖の 5' 末端に転移され、少なくとも 1 つの鎖が前記第 1 のバーコード配列で 5' タグ化されている、二本鎖の標的核酸と、を含む、標的核酸のバーコード化核酸フラグメントのライブラリーを調製するための組成物。

【発明の詳細な説明】

【技術分野】

【0001】

関連出願

本願は、その全体が参照により本明細書に組み込まれる、2014 年 10 月 17 日出願の米国仮特許出願第 62/065,544 号及び 2015 年 5 月 5 日出願の米国仮特許出願第 62/157,396 号の優先権を主張する。

【0002】

本発明の実施形態は、核酸のシーケンシングに関する。具体的には、本明細書で提供する方法及び組成物の実施形態は、核酸鋳型の調製及び核酸鋳型からの配列データの取得に関する。

【背景技術】

【0003】

生体試料中に存在する特定の核酸配列の検出は、例えば、微生物の同定及び分類、感染症の診断、遺伝子異常の検出及び特徴付け、癌に関連した遺伝子変化の同定、疾患への遺伝的感受性の研究、並びに様々な種類の治療に対する反応の測定を行う方法として用いら

10

20

30

40

50

れてきた。生体試料中の特定の核酸配列を検出する一般的な技術は、核酸シーケンシングである。

【0004】

核酸シーケンシング法は、Maxam及びGilbertが用いた化学分解法、及びSangerが用いた鎖伸長法から著しく発展してきた。今日、核酸の並列処理を全て1回のシーケンシングランで行うことを可能にする、いくつかのシーケンシング法が用いられている。従って、1回のシーケンシングランから生成される情報は、膨大なものとなる可能性がある。

【発明の概要】

【課題を解決するための手段】

【0005】

1つの態様において、本明細書は、標的核酸のバーコード化DNAフラグメントのライブラリーを調製する方法について記載する。当該方法は、標的核酸を複数のトランスポソーム複合体と接触させることを含むものであり、各トランスポソーム複合体は、トランスポゾン及びトランスポザゼを含み、トランスポゾンは転移鎖及び非転移鎖を含む。トランスポソーム複合体の少なくとも1つのトランスポゾンは、相補的な捕捉配列にハイブリダイズすることが可能なアダプター配列を含む。標的核酸を複数のフラグメントにフラグメント化し、標的核酸の連続性 (contiguity) を維持しながらフラグメントの少なくとも1つの鎖の5'末端に複数の転移鎖を挿入する。標的核酸の複数のフラグメントを複数の固体支持体と接触させる。複数の固体支持体の各々は複数の固定化オリゴヌクレオチドを含み、各オリゴヌクレオチドは相補的な捕捉配列及び第1のバーコード配列を含み、複数の固体支持体中の各固体支持体からの第1のバーコード配列は、複数の固体支持体中の他の固体支持体からの第1のバーコード配列とは異なる。バーコード配列の情報を標的核酸のフラグメントに転移させ、それにより、同じ標的核酸の少なくとも2つのフラグメントが同一のバーコード情報を受け取るように、少なくとも1つの鎖が第1のバーコード配列で5'タグ化された、二本鎖フラグメントの固定化ライブラリーを作製する。

【0006】

1つの態様において、本明細書は、標的核酸配列の連続性情報を決定する方法について記載する。当該方法は、標的核酸を複数のトランスポソーム複合体と接触させることを含むものであり、各トランスポソーム複合体は、トランスポゾン及びトランスポザゼを含み、トランスポゾンは転移鎖及び非転移鎖を含み、トランスポソーム複合体のトランスポゾンの少なくとも1つは、相補的な捕捉配列にハイブリダイズすることが可能なアダプター配列を含む。標的核酸を複数のフラグメントにフラグメント化し、標的核酸の連続性を維持しながら複数の転移鎖を複数のフラグメントに挿入する。標的核酸の複数のフラグメントを複数の固体支持体と接触させる。複数の固体支持体の各々は複数の固定化オリゴヌクレオチドを含み、各オリゴヌクレオチドは相補的な捕捉配列及び第1のバーコード配列を含み、複数の固体支持体中の各固体支持体からの第1のバーコード配列は、複数の固体支持体中の他の固体支持体からの第1のバーコード配列とは異なる。バーコード配列情報を、同じ標的核酸の少なくとも2つのフラグメントが同一のバーコード情報を受け取るように、標的核酸フラグメントに転移させる。標的核酸フラグメントの配列及びバーコード配列を決定する。標的核酸の連続性情報を、バーコード配列を識別することにより決定する。いくつかの実施形態において、トランスポソーム複合体のトランスポザゼを転位 (transposition) 後に除去し、続いて、トランスポゾンのアダプター配列を相補的な捕捉配列にハイブリダイズさせる。いくつかの実施形態において、トランスポザゼをSDS処理により除去する。いくつかの実施形態において、トランスポザゼをタンパク質分解酵素処理により除去する。

【0007】

1つの態様において、本明細書は、標的核酸配列のフェージング情報及びメチル化状態を同時に測定する方法について記載する。当該方法は、標的核酸を複数のトランスポソーム複合体と接触させることを含むものであり、各トランスポソーム複合体は、トランスポ

10

20

30

40

50

ゾン及びトランスポザーゼを含み、トランスポゾンに転移鎖及び非転移鎖を含み、トランスポソーム複合体のトランスポゾンの少なくとも1つは、相補的な捕捉配列にハイブリダイズすることが可能なアダプター配列を含む。標的核酸を複数のフラグメントにフラグメント化し、標的核酸の連続性を維持しながら複数の転移鎖を標的核酸フラグメントに挿入する。標的核酸の複数のフラグメントを複数の固体支持体と接触させ、複数の固体支持体の各々は複数の固定化オリゴヌクレオチドを含み、各オリゴヌクレオチドは相補的な捕捉配列及び第1のバーコード配列を含み、複数の固体支持体中の各固体支持体からの第1のバーコード配列は、複数の固体支持体中の他の固体支持体からの第1のバーコード配列とは異なる。バーコード配列情報を、同じ標的核酸の少なくとも2つのフラグメントが同一のバーコード情報を受け取るように、標的核酸フラグメントに転移させる。バーコードを含む標的核酸フラグメントを亜硫酸水素塩処理に付し、それにより、バーコードを含む亜硫酸水素塩処理標的核酸フラグメントを生成する。亜硫酸水素塩処理標的核酸フラグメントの配列及びバーコード配列を決定する。標的核酸の連続性情報を、バーコード配列を識別することにより決定する。

10

【0008】

1つの態様において、本明細書は、タグ化DNAフラグメントの固定化ライブラリーを調製する方法について記載する。当該方法は、それに固定化されたトランスポソーム複合体を有する複数の固体支持体を提供することを含み、トランスポソーム複合体は多量体であり、同じトランスポソーム複合体のトランスポソーム単量体単位は互いに結合し、トランスポソーム単量体単位は第1のポリヌクレオチドに結合したトランスポザーゼを含み、第1のポリヌクレオチドは、(i) トランスポゾン末端配列を含む3'部分、及び(ii) 第1のバーコードを含む第1のアダプターを含む。標的DNAを、標的DNAがトランスポソーム複合体によりフラグメント化され、第1のポリヌクレオチドの3'トランスポゾン末端配列がフラグメントの少なくとも1つの鎖の5'末端に転移する条件下で、複数の固体支持体に適用する。それにより、少なくとも1つの鎖が第1のバーコードで5'タグ付けされた二本鎖フラグメントの固定化ライブラリーを作製する。

20

【0009】

1つの態様において、本明細書は、標的核酸のメチル化状態を決定するためのシーケンシングライブラリーを調製する方法について記載する。当該方法は、標的核酸を2つ又はそれ以上のフラグメントにフラグメント化することを含む。第1の共通アダプター配列を標的核酸のフラグメントの5'末端に組み込み、アダプター配列は、第1のプライマー結合配列及び親和性部分を含み、親和性部分は、結合ペアの1つのメンバーに存在する。標的核酸フラグメントを変性させる。標的核酸フラグメントを固体支持体に固定化する。固体支持体は、結合ペアの他のメンバーを含み、標的核酸の固定化は、結合ペアの結合により行う。固定化標的核酸フラグメントを亜硫酸水素塩処理に付す。第2の共通アダプター配列を亜硫酸水素塩処理した固定化標的核酸フラグメントに組み込み、第2の共通アダプターは、第2のプライマー結合部位を含む。固体支持体に固定化された亜硫酸水素塩処理固定化標的核酸フラグメントを増幅し、それにより、標的核酸のメチル化状態を決定するためのシーケンシングライブラリーを作製する。

30

【0010】

1つの態様において、本明細書は、標的核酸のメチル化状態を決定するためのシーケンシングライブラリーを調製する方法について記載する。当該方法は、それに固定化された固定化トランスポソーム複合体を含む複数の固体支持体を提供することを含む。トランスポソーム複合体は、トランスポゾン及びトランスポザーゼを含み、トランスポゾンは、転移鎖及び非転移鎖を含む。転移鎖は、(i) トランスポザーゼ認識配列を含む3'末端の第1の部分及び(ii) 第1のアダプター配列及び結合ペアの第1のメンバーを含む5'から第1の部分に位置する第2の部分を含む。結合ペアの第1のメンバーは、固体支持体上で結合ペアの第2のメンバーに結合し、それにより、トランスポゾンを固体支持体に固定化する。第1のアダプターはまた、第1のプライマー結合配列を含む。非転移鎖は、(i) トランスポザーゼ認識配列を含む5'末端の第1の部分及び(ii) 3'末端の末

40

50

端ヌクレオチドがブロックされている第2のアダプター配列を含む3'から第1の部分に位置する第2の部分を含む。第2のアダプターはまた、第2のプライマー結合配列を含む。標的核酸を、固定化トランスポソーム複合体を含む複数の固体支持体と接触させる。標的核酸を複数のフラグメントにフラグメント化し、複数の転移鎖をフラグメントの少なくとも1つの鎖の5'末端に挿入し、それにより、標的核酸フラグメントを固体支持体に固定化する。フラグメント化された標的核酸の3'末端を、DNAポリメラーゼで伸長させる。非転移鎖を、フラグメント化された標的核酸の3'末端にライゲートする。固定化標的核酸フラグメントを亜硫酸水素塩処理に付す。亜硫酸水素塩処理中に損傷を受けた固定化標的核酸フラグメントの3'末端を、固定化標的核酸フラグメントの3'末端がホモポリマーテイルを含むように、DNAポリメラーゼを用いて伸長させる。第2のアダプター配列を、亜硫酸水素塩処理中に損傷を受けた固定化標的核酸フラグメントの3'末端に導入する。固体支持体に固定化された亜硫酸水素塩処理標的核酸フラグメントを、第1及び第2のプライマーを用いて増幅し、それにより、標的核酸のメチル化状態を決定するためのシーケンシングライブラリーを作製する。

【0011】

1つの態様において、本明細書は、標的核酸のメチル化状態を決定するためのシーケンシングライブラリーを調製する方法について記載する。当該方法は、標的核酸をトランスポソーム複合体と接触させることを含むものであり、トランスポソーム複合体は、トランスポゾン及びトランスポザーゼを含む。トランスポゾンは、転移鎖及び非転移鎖を含む。転移鎖は、(i) トランスポザーゼ認識配列を含む3'末端の第1の部分及び(ii) 第1のアダプター配列及び結合ペアの第1のメンバーを含む5'から第1の部分に位置する第2の部分を含み、結合ペアの第1のメンバーは、結合ペアの第2のメンバーに結合する。非転移鎖は、(i) トランスポザーゼ認識配列を含む5'末端の第1の部分及び(ii) 3'末端の末端ヌクレオチドがブロックされている第2のアダプター配列を含む3'から第1の部分に位置する第2の部分を含み、第2のアダプターは、第2のプライマー結合配列を含む。標的核酸を複数のフラグメントにフラグメント化し、複数の転移鎖をフラグメントの少なくとも1つの鎖の5'末端に挿入し、それにより、標的核酸フラグメントを固体支持体に固定化する。トランスポゾン末端を含む標的核酸フラグメントを結合ペアの第2のメンバーを含む複数の固体支持体に接触させ、結合ペアの第1のメンバーと結合ペアの第2のメンバーとの結合により、標的核酸を固体支持体に固定化する。フラグメント化された標的核酸の3'末端を、DNAポリメラーゼで伸長させる。非転移鎖を、フラグメント化された標的核酸の3'末端にライゲートする。固定化標的核酸フラグメントを亜硫酸水素塩処理に付す。亜硫酸水素塩処理中に損傷を受けた固定化標的核酸フラグメントの3'末端を、固定化標的核酸フラグメントの3'末端がホモポリマーテイルを含むように、DNAポリメラーゼを用いて伸長させる。第2のアダプター配列を、亜硫酸水素塩処理中に損傷を受けた固定化標的核酸フラグメントの3'末端に導入する。固体支持体に固定化された亜硫酸水素塩処理標的核酸フラグメントを、第1及び第2のプライマーを用いて増幅し、それにより、標的核酸のメチル化状態を決定するためのシーケンシングライブラリーを作製する。

【0012】

いくつかの実施形態において、第2のアダプターの3'末端の末端ヌクレオチドを、ジデオキシヌクレオチド、リン酸基、チオリン酸基、及びアジド基からなる群から選択される1つのメンバーによりブロックする。

【0013】

いくつかの実施形態において、親和性部分を結合ペアのメンバーとすることができる。いくつかのケースにおいて、修飾された核酸は、結合ペアの第1のメンバーを含んでいてもよく、捕捉プローブは、結合ペアの第2のメンバーを含んでいてもよい。いくつかのケースにおいて、捕捉プローブを固体表面に固定化してもよく、修飾された核酸は、結合ペアの第1のメンバーを含んでいてもよく、捕捉プローブは、結合ペアの第2のメンバーを含んでいてもよい。そのようなケースでは、結合ペアの第1及び第2のメンバーの結合に

より、修飾された標的核酸を固体表面に固定化する。結合ペアの例としては、これに限定されないが、ビオチン-アビジン、ビオチン-ストレプトアビジン、ビオチン-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、オリゴヌクレオチド-相補的オリゴヌクレオチド、及び抗原-抗体が挙げられる。

【0014】

いくつかの実施形態において、第1の共通アダプター配列を、片側 (one-sided) 転位により標的核酸の5'末端フラグメントに組み込む。いくつかの実施形態において、第1の共通アダプター配列を、ライゲーションにより標的核酸の5'末端フラグメントに組み込む。いくつかの実施形態において、第2の共通アダプター配列を亜硫酸水素塩処理固定化標的核酸フラグメントに組み込むステップは、(i) 固定化標的核酸フラグメントの3'末端を、ターミナルトランスフェラーゼを用いて伸長し、ホモポリマーテイルを含ませるステップ、(ii) 一本鎖ホモポリマー部分を含むオリゴヌクレオチドと、第2の共通アダプター配列を含む二本鎖部分とをハイブリダイズさせるステップであって、一本鎖ホモポリマー部分がホモポリマーテイルに相補的である、ステップ、及び(iii) 第2の共通アダプター配列を固定化標的核酸フラグメントにライゲートし、それにより、第2の共通アダプター配列を亜硫酸水素塩処理固定化標的核酸フラグメントに組み込むステップを含む。

【0015】

いくつかの実施形態において、標的核酸は、単一の細胞に由来する。いくつかの実施形態において、標的核酸は、単一の細胞小器官に由来する。いくつかの実施形態において、標的核酸は、ゲノムDNAである。いくつかの実施形態において、標的核酸は、他の核酸と架橋する。いくつかの実施形態において、標的核酸は、ホルマリン固定パラフィン包埋 (FFPE: formalin fixed paraffin embedded) サンプルに由来する。いくつかの実施形態において、標的核酸は、タンパク質と架橋する。いくつかの実施形態において、標的核酸は、DNAと架橋する。いくつかの実施形態において、標的核酸は、ヒストンに保護されたDNAである。いくつかの実施形態において、ヒストンを標的核酸から除去する。いくつかの実施形態において、標的核酸は、無細胞腫瘍DNAである。いくつかの実施形態において、無細胞腫瘍DNAは、胎盤液から得る。いくつかの実施形態において、無細胞腫瘍DNAは、血漿から得る。いくつかの実施形態において、血漿は、血漿用採取ゾーンを有する膜分離装置を用いて全血から採取する。いくつかの実施形態において、血漿用採取ゾーンは、固体支持体に固定化されたトランスポソーム複合体を含む。いくつかの実施形態において、標的核酸は、cDNAである。いくつかの実施形態において、固体支持体は、ビーズである。いくつかの実施形態において、複数の固体支持体は、複数のビーズであり、複数のビーズは、様々なサイズである。

【0016】

いくつかの実施形態において、単一のバーコード配列が、各個々の固体支持体上の複数の固定化オリゴヌクレオチドに存在する。いくつかの実施形態において、異なるバーコード配列が、各個々の固体支持体上の複数の固定化オリゴヌクレオチドに存在する。いくつかの実施形態において、標的核酸フラグメントへのバーコード配列情報の転移は、ライゲーションによる。いくつかの実施形態において、標的核酸フラグメントへのバーコード配列情報の転移は、ポリメラーゼ伸長による。いくつかの実施形態において、標的核酸フラグメントへのバーコード配列情報の転移は、ライゲーション及びポリメラーゼ伸長の両方による。いくつかの実施形態において、ポリメラーゼ伸長は、ライゲートした固定化オリゴヌクレオチドを鋳型として用い、非ライゲートトランスポゾン鎖の3'末端をDNAポリメラーゼで伸長させることによる。いくつかの実施形態において、アダプター配列の少なくとも一部は、第2のバーコード配列をさらに含む。

【0017】

いくつかの実施形態において、トランスポソーム複合体は、多量体であり、各単量体単位のトランスポソンのアダプター配列は、同じトランスポソーム複合体の他の単量体単位とは異なる。いくつかの実施形態において、アダプター配列は、第1のプライマー結合配

10

20

30

40

50

列をさらに含む。いくつかの実施形態において、第1のプライマー結合部位は、捕捉配列又は捕捉配列の相補体に対して、配列相同性を持たない。いくつかの実施形態において、固体支持体上の固定化オリゴヌクレオチドは、第2のプライマー結合配列をさらに含む。

【0018】

いくつかの実施形態において、トランスポソーム複合体は、多量体であり、トランスポソーム単量体単位は、同じトランスポソーム複合体内で互いに結合する。いくつかの実施形態において、トランスポソーム単量体単位のトランスポザーゼは、同じトランスポソーム複合体の別のトランスポソーム単量体単位のトランスポザーゼに結合する。いくつかの実施形態において、トランスポソーム単量体単位のトランスポゾン、同じトランスポソーム複合体の別のトランスポソーム単量体単位のトランスポゾンに結合する。いくつかの実施形態において、トランスポソーム単量体単位のトランスポザーゼは、同じトランスポソーム複合体の別のトランスポソーム単量体単位のトランスポザーゼに共有結合により結合する。いくつかの実施形態において、1つの単量体単位のトランスポザーゼは、同じトランスポソーム複合体の別のトランスポソーム単量体単位のトランスポザーゼにジスルフィド結合により結合する。いくつかの実施形態において、トランスポソーム単量体単位のトランスポゾンは、同じトランスポソーム複合体の別のトランスポソーム単量体単位のトランスポゾンに共有結合により結合する。

【0019】

いくつかの実施形態において、標的核酸配列の連続性情報は、ハプロタイプ情報を示す。いくつかの実施形態において、標的核酸配列の連続性情報は、ゲノム変異を示す。いくつかの実施形態において、ゲノム変異は、欠損、転位 (translocations)、染色体間の遺伝子融合、重複、及びパラログからなる群から選択される。いくつかの実施形態において、固体支持体に固定化されたオリゴヌクレオチドは、部分的二本鎖領域及び部分的一本鎖領域を含む。いくつかの実施形態において、オリゴヌクレオチドの部分的一本鎖領域は、第2のバーコード配列及び第2のプライマー結合配列を含む。いくつかの実施形態において、バーコードを含む標的核酸フラグメントを、標的核酸フラグメントの配列を決定する前に増幅する。いくつかの実施形態において、後続の増幅を、標的核酸フラグメントの配列を決定する前に、単一の反応区画で行う。いくつかの実施形態において、第3のバーコード配列を、増幅中に標的核酸フラグメントに導入する。

【0020】

いくつかの実施形態において、上記方法は、バーコードを含む標的核酸フラグメントを、複数の第1のセットの反応区画から、バーコードを含む標的核酸フラグメントのプールにまとめるステップ、バーコードを含む標的核酸フラグメントの前記プールを複数の第2のセットの反応区画に再分配するステップ、及び標的核酸フラグメントを第2のセットの反応区画内でシーケンシング前に増幅することにより、第3のバーコードを標的核酸フラグメントに導入するステップをさらに含んでもよい。

【0021】

いくつかの実施形態において、上記方法は、標的核酸をトランスポソーム複合体と接触させる前にプレフラグメント化するステップをさらに含んでもよい。いくつかの実施形態において、標的核酸のプレフラグメント化は、超音波処理及び制限消化からなる群から選択される方法により行われる。

【図面の簡単な説明】

【0022】

【図1】トランスポソームをビーズ表面に結合させる方法の一例を示すフローチャートである。

【図2】図1の方法のステップを絵で示す図である。

【図3】ビーズ表面上でのタグメント化工程の一例を示す模式図である。

【図4】DNA収量の一例を、図3のビーズベースタグメント化工程からのクラスター数の観点から示すデータ表である。

【図5】図3のビーズベースタグメント化工程の再現性の別例を、均一なサイズの観点か

10

20

30

40

50

ら示すデータ表である。

【図 6】図 6 A 及び 6 B は、それぞれ、図 5 のインデックス付きサンプルのプール 1 の挿入サイズのプロット及びプール 2 の挿入サイズのプロットを示す図である。

【図 7】リードの合計数及び図 5 に記載の実験のためにアラインしたリードの割合の再現性を示す棒グラフである。

【図 8】図 8 A、8 B、及び 8 C は、それぞれ、エクソーム濃縮アッセイにおける、コントロールライブラリーでの挿入サイズのプロット、ビーズベースタグメント化ライブラリーでの挿入サイズのプロット、及びサマリーデータ表を示す図である。

【図 9】図 9 A、9 B、及び 9 C は、それぞれ、エクソーム濃縮アッセイにおける、dupsp PF 画分の棒グラフ、selected bases 画分の棒グラフ、及び PCT usable bases on target の棒グラフを示す図である。

10

【図 10】ビーズ表面上でのトランスポソーム複合体の形成方法の一例を示すフローチャートである。

【図 11】図 10 の方法のステップを絵で示す図である。

【図 12】図 10 の方法のステップを絵で示す図である。

【図 13】図 10 の方法のステップを絵で示す図である。

【図 14】図 13 に示すトランスポソーム被覆ビーズを用いたタグメント化工程の模式図である。

【図 15】固体支持体上でのトランスポソーム形成の例示的なスキームを示す図である。

【図 16】固有インデックスを有する連結 (contiguously linked) ライブラリー作製の例示的なスキームを示す図である。

20

【図 17】固有インデックスを有する連結ライブラリー作製の例示的なスキームを示す図である。

【図 18】CPT-DNA がビーズに巻き付いている、単一のクローンインデックス付きビーズ上での単一の CPT-DNA の捕捉を表わす図である。

【図 19】CPT-DNA がビーズに巻き付いている、単一のクローンインデックス付きビーズ上での単一の CPT-DNA の捕捉を表わす図である。

【図 20】固体表面に固定化した Y-アダプターを、ライゲーション及びギャップ充填により標的 DNA に結合させる例示的なスキームを示す図である。

【図 21】CPT-DNA と固体支持体上の固定化オリゴヌクレオチドとのライゲーションの間に、前記 Y-アダプターを作製する例示的なスキームを示す図である。

30

【図 22】サイズ排除クロマトグラフィーによる連結ライブラリーからの遊離トランスポソームの除去を示す、アガロースゲル電気泳動を表わす図である。

【図 23】特定の DNA フラグメントのショットガン配列ライブラリーを生成する例示的なスキームを示す図である。

【図 24】クローンインデックス付きシーケンシングライブラリーからの配列情報をアセンブルする例示的なスキームを示す図である。

【図 25】ビーズ上の捕捉プローブ密度の最適化の結果を示す図である。

【図 26】分子内ハイブリダイゼーションによるビーズ上での CPT-DNA のインデックス付きシーケンシングライブラリーの調製の実現性を試験した結果を示す図である。

40

【図 27】クローンインデックス化の実現性を試験した結果を示す図である。

【図 28】タグメント化後の鋳型核酸に対して隣接して (neighbouring) アラインされたリードの島内 (intra) 及び島間 (inter) の特定の距離に対する、シーケンシングリードの頻度を示すグラフである。

【図 29】図 29 A 及び 29 B は、固体支持体上の連続性情報を引き出す例示的なアプローチを示す図である。

【図 30】単一の反応容器 (ワンポット) におけるインデックス付きクローンビーズ転位の概略図及び転位結果を示す図である。

【図 31】単一の反応容器 (ワンポット) におけるインデックス付きクローンビーズ転位の概略図及び転位結果を示す図である。

50

【図 3 2】5'又は3'ビオチン化オリゴヌクレオチドを用いたビーズ上でのクローントランスポソームの作製を示す概略図である。

【図 3 3】ビーズ上のトランスポソームに対するライブラリーサイズを示す図である。

【図 3 4】挿入サイズに対するトランスポソーム表面密度の影響を示す図である。

【図 3 5】サイズ分布に対するインプット DNA の影響を示す図である。

【図 3 6】ビーズベース及び溶液ベースのタグメント化反応を用いた島のサイズ及び分布を示す図である。

【図 3 7】それぞれ固有インデックスを受け取った、いくつかの個々の DNA 分子のクローンインデックス化を示す図である。

【図 3 8】全血から血漿を分離する装置の略図である。

【図 3 9】血漿を分離する装置及び分離された血漿のその後の使用を示す略図である。

【図 4 0】血漿を分離する装置及び分離された血漿のその後の使用を示す略図である。

【図 4 1】ゲノムの特定の領域を濃縮することによる標的フェージングの例示的なスキームを示す図である。

【図 4 2】エクソン間の SNP を用いたエクソームフェージングの例示的なスキームを示す図である。

【図 4 3】フェージング及びメチル化の同時検出の例示的なスキームを示す図である。

【図 4 4】フェージング及びメチル化の同時検出の別の例示的なスキームを示す図である。

【図 4 5】単一アッセイにおいて、種々のサイズのクローンインデックス付きビーズを用いて種々のサイズのライブラリーを生成する、例示的なスキームを示す図である。

【図 4 6】異なる長さスケールのライブラリーで遺伝的変異を決定する例示的なスキームを示す図である。

【図 4 7 A】染色体 1 における 60 kb ヘテロ接合欠損の検出結果を示す図である。

【図 4 7 B】染色体 1 における 60 kb ヘテロ接合欠損の検出結果を示す図である。

【図 4 8】本願の方法を用いた遺伝子融合検出の結果を示す図である。

【図 4 9】本願の方法を用いた遺伝子欠損検出の結果を示す図である。

【図 5 0】亜硫酸水素塩変換前後の ME 配列を示す図である。

【図 5 1】亜硫酸水素塩変換効率の最適化の結果を示す図である。

【図 5 2】亜硫酸水素塩変換後の結果を IVC プロット（各塩基当たりの強度対サイクル）で示す図である。

【図 5 3】BSC の後の PCR 後のインデックス付き結合ライブラリーのアガロースゲル電気泳動の画像を示す図である。

【図 5 4】サイズ選択をしていない濃縮前の全ゲノムインデックス付き結合 CPT-seq ライブラリーのバイオアナライザートレースを示す図である。

【図 5 5】濃縮後のライブラリーのアガロースゲル分析を示す図である。

【図 5 6】標的化ハプロタイピングを染色体の HLA 領域に適用した結果を示す図である。

【図 5 7】ME 交換 (swapping) のいくつかの考え得るメカニズムを示す図である。

【図 5 8】ME 交換 (swapping) のいくつかの考え得るメカニズムを示す図である。

【図 5 9】Cys で置換することができる例示的なアミノ酸残基 Asp 468、Tyr 407、Asp 461、Lys 459、Ser 458、Gly 462、Ala 466、Met 470 を有する Tn5 トランスポザーゼの一部を示す図である。

【図 6 0】システイン残基が 2 つの単量体単位間でジスルフィド結合を形成できるようにした、S458C、K459C、及び A466C のアミノ酸置換を有する Tn5 トランスポザーゼの一部を示す図である。

【図 6 1】アミン被覆ナノ粒子を用いた二量体トランスポザーゼ (dTnp) ナノ粒子 (NP) バイオコンジュゲート (dTnp-NP) の作製及び使用の例示的なスキームを示す図である。

【図 6 2】トランスポソーム二量体とアミン被覆固体支持体とのコンジュゲーションの例

10

20

30

40

50

示的なスキームを示す図である。

【図 6 3】トランスポゾン末端が結合した M u トランスポソーム複合体を示す図である。

【図 6 4】疑似遺伝子のアセンブリ／フェージングのためのインデックス付き結合リードの略図、及びより短いフラグメントを用いた疑似遺伝子の変異識別の利点を示す図である。

【図 6 5】交換された (s w a p p e d) インデックスの割合 (%) として示す、4 つの別個の実験からのインデックス交換 (e x c h a n g e) のプロットを示す図である。

【図 6 6】T s - T n 5 滴定のフラグメントサイズの、A g i l e n t B i o A n a l y z e r による分析を示す図である。

【図 6 7】亜硫酸水素塩処理後の破損したライブラリーエレメントを回復させる酵素法を用いて、E p i - C P T S e q プロトコルの DNA 収量を改善する、例示的なスキームを示す図である。

10

【図 6 8】図 6 8 A ~ 6 8 C は、亜硫酸水素塩処理後の破損したライブラリーエレメントを回復させる酵素法を用いて、E p i - C P T S e q プロトコルの DNA 収量を改善する、いくつかの例示的なスキームを示す図である。

【図 6 9】ランダムプライマー伸長を用いた鋳型救出 (r e s c u e) の例示的なスキームを示す図である。

【図 7 0】亜硫酸水素ナトリウム変換中の DNA ライブラリーのフラグメント化を示す図である。左のパネルは、磁性ビーズ上にタグメント化された DNA の一部を亜硫酸水素変換している最中のフラグメント化を示す。右のパネルは、C P T - s e q 及び E p i - C P T - s e q (M e - C P T - s e q) ライブラリーのバイオアナライザートレースを示す。

20

【図 7 1】T d T 媒介 s s DNA ライゲーション反応の例示的なスキーム及び結果を示す図である。

【図 7 2】ライブラリーに結合した亜硫酸水素ナトリウム変換ビーズの T d T 媒介回復のスキーム及び結果を示す図である。左のパネルは、T d T 媒介ライゲーション反応を用いた損傷亜硫酸水素塩変換 DNA ライブラリーの救出の操作フローを示す。DNA ライブラリー救出実験の結果を右のパネルに示す。

【図 7 3】メチルー C P T - s e q アッセイの結果を示す図である。

【図 7 4】DNA のビーズベース亜硫酸水素塩変換の例示的なスキームを示す図である。

【図 7 5】図 7 5 A 及び 7 5 B は、亜硫酸水素塩変換効率の最適化の結果を示す図である。

30

【発明を実施するための形態】

【0023】

本発明の実施形態は、核酸のシーケンシングに関する。具体的には、本明細書で提供する方法及び組成物の実施形態は、核酸鋳型の調製及び核酸鋳型からの配列データの取得に関する。

【0024】

1 つの態様において、本発明は、タグメント化 (フラグメント化及びタグ化) した標的核酸ライブラリーを構築するために、固体支持体上で標的核酸をタグメント化する方法に関する。1 つの実施形態において、固体支持体は、ビーズである。1 つの実施形態において、標的核酸は、DNA である。

40

【0025】

1 つの態様において、本発明は、標的核酸の連続性情報を引き出すことが可能な、固体支持体、トランスポザーゼに基づく方法の、方法及び組成物に関する。いくつかの実施形態において、組成物及び方法は、アセンブリ／フェージング情報を引き出すことができる。

【0026】

1 つの態様において、本発明は、連結した転位標的核酸を固体支持体上に捕捉することにより、連続性情報を引き出す方法及び組成物に関する。

【0027】

1 つの態様において、本明細書に開示する方法及び組成物は、ゲノム変異の分析に関す

50

る。例示的なゲノム変異としては、これに限定されないが、欠損、染色体間転位、重複、パラログ、染色体間遺伝子融合が挙げられる。いくつかの実施形態において、本明細書に開示する方法及び組成物は、ゲノム変異のフェージング情報の決定に関する。

【0028】

1つの態様において、本明細書に開示する方法及び組成物は、標的核酸の特定の領域のフェージングに関する。1つの実施形態において、標的核酸は、DNAである。1つの実施形態において、標的核酸は、ゲノムDNAである。いくつかの実施形態において、標的核酸は、RNAである。いくつかの実施形態において、RNAは、mRNAである。いくつかの実施形態において、標的核酸は、相補的DNA (cDNA: complementary DNA) である。いくつかの実施形態において、標的核酸は、単一の細胞に由来する。いくつかの実施形態において、標的核酸は、循環腫瘍細胞に由来する。いくつかの実施形態において、標的核酸は、無細胞DNAである。いくつかの実施形態において、標的核酸は、ホルマリン固定パラフィン包埋組織サンプルに由来する。いくつかの実施形態において、標的核酸は、架橋標的核酸である。いくつかの実施形態において、標的核酸は、タンパク質に架橋する。いくつかの実施形態において、標的核酸は、核酸に架橋する。いくつかの実施形態において、標的核酸は、ヒストンに保護されたDNAである。いくつかの実施形態において、ヒストンに保護されたDNAを、ヒストンに対する抗体を用いて細胞溶解物から沈殿させ、ヒストンを除去する。

【0029】

いくつかの態様において、インデックス付きライブラリーは、クローンインデックス付きビーズを用いて、標的核酸から作製する。いくつかの実施形態において、タグメント化標的核酸は、トランスポザゼが標的DNAに結合したままで、クローンインデックス付きビーズを用いて捕捉することができる。いくつかの実施形態において、特異的な捕捉プローブを用いて、標的核酸中の目的の特異的な領域を捕捉する。標的核酸の捕捉された領域は、種々のストリンジェンシーで洗浄し、任意選択的に増幅し、その後シーケンシングすることができる。いくつかの実施形態において、捕捉プローブは、ビオチン化してもよい。ビオチン化捕捉プローブがインデックス付き標的核酸の特異的な領域にハイブリダイズした複合体は、ストレプトアビジンビーズを用いて分離することができる。標的フェージングの例示的なスキームを図41に示す。

【0030】

いくつかの態様において、本明細書に開示する組成物及び方法は、エクソームのフェージングに用いることができる。いくつかの実施形態において、エクソン、プロモーターを濃縮することができる。マーカー、例えば、エクソン領域間のヘテロ接合SNPは、特にエクソン間の距離が大きい場合に、エクソームのフェージングに役立つ可能性がある。例示的なエクソームのフェージングを図42に示す。いくつかの実施形態において、インデックス付き結合リードは、隣接しているエクソンのヘテロ接合SNPに同時に及ぶ（カバーする）ことができない。従って、2つ又はそれ以上のエクソンをフェージングすることは困難である。本明細書に開示する組成物及び方法はまた、エクソン間のヘテロ接合SNPを濃縮し、例えば、エクソン1をSNP1に、SNP2をエクソン2にフェージングする。従って、SNP1の使用を通じて、エクソン1及びエクソン2を、図42に示すようにフェージングすることができる。

【0031】

1つの態様において、本明細書に開示する組成物及び方法は、フェージング及び同時のメチル化の検出に用いることができる。亜硫酸水素塩変換 (BSC: bisulfite conversion) によるメチル化検出は、BSC反応がDNAに対して厳しい (harsh) ものであり、DNAをフラグメント化し、それにより連続性／フェージング情報を除去するため、困難である。また、本願に開示する方法は、従来のBSCアプローチで必要とされるのとは対照的に、更なる精製ステップを必要としないことにより、収率を改善するため、さらに利点がある。

【0032】

1つの態様において、本明細書に開示する組成物及び方法を用いて、異なるサイズのライブラリーを1つのアッセイで調製することができる。いくつかの実施形態において、異なるサイズのクローンインデックス付きビーズを用いて、異なるサイズのライブラリーを調製することができる。図1は、トランスポザーゼをビーズ表面に結合させる方法100の一例のフローチャートを示す。トランスポザーゼは、トランスポゾンオリゴヌクレオチド、トランスポザーゼ、及び固相に加えてもよい任意の化学物質を用いて、ビーズ表面に結合させてもよい。1つの例において、トランスポソームを、ビオチン-ストレプトアビジン結合複合体を介してビーズ表面に結合させる。方法100は、これに限定されないが、以下のステップを含む。

10

【0033】

1つの実施形態において、トランスポゾンは、シーケンシングプライマー結合部位を含んでいてもよい。シーケンス結合部位の例示的な配列としては、これに限定されないが、AATGATACGGCGACCAACCGAGATCTACAC（P5配列）及びCAAGCAGAAAGACGGCATACGAGAT（P7配列）が挙げられる。いくつかの実施形態において、トランスポゾンをビオチン化してもよい。

【0034】

図1のステップ110において、P5及びP7ビオチン化トランスポゾンを生成する。トランスポゾンはまた、1つ又はそれ以上のインデックス配列（固有の識別子）を含んでいてもよい。例示的なインデックス配列としては、これに限定されないが、TAGATCGC、CTCTCTAT、TATCCTCT、AGAGTAGA、GTAAGGAG、ACTGCATA、AAGGAGTA、CTAAGCCTが挙げられる。別の例では、P5トランスポゾンのみ又はP7トランスポゾンのみをビオチン化する。さらに別の例では、トランスポゾンは、モザイク末端（ME:mosaic end）配列のみ、又はME配列に加えてP5及びP7配列ではない追加の配列を含む。この例では、P5及びP7配列は、後続のPCR増幅ステップで付加する。

20

【0035】

図1のステップ115において、トランスポソームをアセンブルする。アセンブルしたトランスポソームは、P5及びP7トランスポソームの混合物である。P5及びP7トランスポソームの混合物は、図11及び12と関連してより詳細に説明する。

30

【0036】

図1のステップ120において、P5／P7トランスポソーム混合物を、ビーズ表面に結合させる。この例では、ビーズは、ストレプトアビジン被覆ビーズであり、トランスポソームを、ビオチン-ストレプトアビジン結合複合体を介してビーズ表面に結合させる。ビーズは、種々のサイズであってもよい。1つの例において、ビーズは、2.8 µmビーズであってもよい。別の例において、1 µmビーズであってもよい。1 µmビーズの懸濁液（例えば、1 µL）は、体積当たり大きな表面積をトランスポソーム結合にもたらす。トランスポソーム結合に用いることができる表面積により、反応当たりのタグメント化生成物の数が増加する。

【0037】

40

図2は、図1の方法100のステップ110、115、及び120を絵で示す。この例では、トランスポゾンを二本鎖で示す。別の例（図示しない）では、ヘアピン等の別の構造、即ち、二本鎖を形成することができる自己相補的な領域を有する単一のオリゴヌクレオチドを用いてもよい。

【0038】

方法100のステップ110において、複数のビオチン化P5トランスポゾン210a及び複数のP7トランスポゾン210bを生成する。P5トランスポゾン210a及びP7トランスポゾン210bをビオチン化する。

【0039】

方法100のステップ115において、P5トランスポゾン210a及びP7トランス

50

ポゾン 210b をトランスポザーゼ Tn5 215 と混合し、複数のアセンブルしたトランスポソーム 220 を形成する。

【0040】

方法 100 のステップ 120 において、トランスポソーム 220 を、ビーズ 225 に結合させる。ビーズ 225 は、ストレプトアビジン被覆ビーズである。トランスポソーム 220 を、ビオチン-ストレプトアビジン結合複合体を介してビーズ 225 に結合させる。

【0041】

1 つの実施形態において、トランスポソームの混合物を、図 10、11、12、及び 13 に示すように、ビーズ表面等の固体支持体上で形成してもよい。この例では、P5 及び P7 オリゴヌクレオチドを、トランスポソーム複合体のアセンブリ前に、初めにビーズ表面に結合させる。

【0042】

図 3 は、ビーズ表面上におけるタグメント化工程 300 の例の模式図を示す。工程 300 において、トランスポソーム 220 が結合した図 2 のビーズ 225 を示す。DNA 310 の溶液をビーズ 225 の懸濁液に加える。DNA 310 がトランスポソーム 220 に接触すると、DNA がタグメント化（フラグメント化及びタグ化）され、トランスポソーム 220 を介してビーズ 225 に結合する。結合及びタグメント化された DNA 310 を PCR 増幅して、溶液中（ビーズを含まない）で増幅産物 315 のプールを生成してもよい。増幅産物 315 は、フローセル 320 の表面に転移させてもよい。クラスター生成プロトコル（例えば、ブリッジ増幅プロトコル、又はクラスター生成に使用することができる任意のその他の増幅プロトコル）を用いて、複数のクラスター 325 をフローセル 320 の表面上に生成してもよい。クラスター 325 は、タグメント化 DNA 310 のクローン増幅産物である。これでクラスター 325 は、シーケンシングプロトコルの次のステップ用に準備できたことになる。

【0043】

別の実施形態において、トランスポソームは、マイクロ遠心チューブの壁等の任意の固体表面に結合してもよい。

【0044】

ビーズ表面上にトランスポソーム複合体の混合物を形成する別の実施形態において、オリゴヌクレオチドは、トランスポソームのアセンブリ前に、初めにビーズ表面に結合させる。図 10 は、ビーズ表面上でトランスポソーム複合体を形成する方法 1000 の一例のフローチャートを示す。方法 1000 は、これに限定されないが、以下のステップを含む。

【0045】

ステップ 1010 において、P5 及び P7 オリゴヌクレオチドを、ビーズ表面に結合させる。1 つの例において、P5 及び P7 オリゴヌクレオチドは、ビオチン化し、ビーズは、ストレプトアビジン被覆ビーズである。このステップはまた、図 11 の模式図 1100 に絵で示す。ここで、図 11 を参照すると、P5 オリゴヌクレオチド 1110 及び P7 オリゴヌクレオチド 1115 は、ビーズ 1120 の表面に結合する。この例においては、1 つの P5 オリゴヌクレオチド 1110 及び 1 つの P7 オリゴヌクレオチド 1115 が、ビーズ 1120 の表面に結合しているが、任意の数の P5 オリゴヌクレオチド 1110 及び／又は P7 オリゴヌクレオチド 1115 が、複数のビーズ 1120 の表面に結合してもよい。1 つの例において、P5 オリゴヌクレオチド 1110 は、P5 プライマー配列、インデックス配列（固有の識別子）、リード 1 シーケンシングプライマー配列、及びモザイク末端（ME）配列を含む。この例において、P7 オリゴヌクレオチド 1115 は、P7 プライマー配列、インデックス配列（固有の識別子）、リード 2 シーケンシングプライマー配列、及び ME 配列を含む。別の例（図示せず）において、インデックス配列は、P5 オリゴヌクレオチド 1110 のみに存在する。さらに別の例（図示せず）において、インデックス配列は、P7 オリゴヌクレオチド 1115 のみに存在する。さらに別の例（図示せず）において、インデックス配列は、P5 オリゴヌクレオチド 1110 及び P7 オリ

10

20

30

40

50

ゴヌクレオチド 1 1 1 5 のいずれにも存在しない。

【0046】

ステップ 1015 において、相補的モザイク末端 (ME') オリゴヌクレオチドを、ビーズ結合 P5 及び P7 オリゴヌクレオチドにハイブリダイズさせる。このステップはまた、図 12 の模式図 1200 に絵で示す。ここで、図 12 を参照すると、相補的 ME 配列 (ME') 1125 は、P5 オリゴヌクレオチド 1110 及び P7 オリゴヌクレオチド 1115 にハイブリダイズする。相補的 ME 配列 (ME') 1125 (例えば、相補的 ME 配列 (ME') 1125a 及び相補的 ME 配列 (ME') 1125b) は、P5 オリゴヌクレオチド 1110 及び P7 オリゴヌクレオチド 1115 の ME 配列にそれぞれハイブリダイズする。相補的 ME 配列 (ME') 1125 は、典型的には、約 15 塩基長であり、5' 末端でリン酸化されている。

10

【0047】

ステップ 1020 において、トランスポザーゼ酵素をビーズ結合オリゴヌクレオチドに添加し、ビーズ結合トランスポソーム複合体の混合物を形成する。このステップはまた、図 13 の模式図 1300 に絵で示す。ここで図 13 を参照すると、トランスポザーゼ酵素が添加されて、複数のトランスポソーム複合体 1310 を形成する。この例において、トランスポソーム複合体 1310 は、トランスポザーゼ酵素、2つの表面結合オリゴヌクレオチド配列、及びそれらにハイブリダイズした相補的 ME 配列 (ME') 1125 を含む二本鎖構造である。例えば、トランスポソーム複合体 1310a は、相補的 ME 配列 (ME') 1125 にハイブリダイズした P5 オリゴヌクレオチド 1110 及び相補的 ME 配列 (ME') 1125 にハイブリダイズした P7 オリゴヌクレオチド 1115 (即ち、P5:P7) を含み、トランスポソーム複合体 1310b は、相補的 ME 配列 (ME') 1125 にハイブリダイズした 2つの P5 オリゴヌクレオチド 1110 (即ち、P5:P5) 含み、トランスポソーム複合体 1310c は、相補的 ME 配列 (ME') 1125 にハイブリダイズした 2つの P7 オリゴヌクレオチド 1115 (即ち、P7:P7) 含む。P5:P5、P7:P7、及び P5:P7 トランスポソーム複合体の割合は、例えば、25:25:50 であってもよい。

20

【0048】

図 14 は、図 13 のトランスポソーム被覆ビーズ 1120 を用いたタグメント化工程の例示的な模式図 1400 を示す。この例において、トランスポソーム複合体 1310 を有するビーズ 1120 を、タグメント化バッファー中の DNA 1410 の溶液に加え、タグメント化を生じさせ、DNA をビーズ 1120 の表面にトランスポソーム 1310 を介して結合させる。DNA 1410 の連続したタグメント化により、トランスポソーム 1310 間に複数のブリッジ分子 1415 が生じる。ブリッジ分子 1415 の長さは、ビーズ 1120 の表面におけるトランスポソーム複合体 1310 の密度に依存する可能性がある。1つの例において、ビーズ 1120 の表面上のトランスポソーム複合体 1310 の密度は、図 10 の方法 1000 のステップ 1010 においてビーズ 1120 の表面に結合する P5 及び P7 オリゴヌクレオチドの量を変化させることにより調整してもよい。別の例において、ビーズ 1120 の表面上のトランスポソーム複合体 1310 の密度は、図 10 の方法 1000 のステップ 1015 において、P5 及び P7 オリゴヌクレオチドにハイブリダイズする相補的 ME 配列 (ME') の量を変化させることにより、調整してもよい。さらに別の例において、ビーズ 1120 の表面上のトランスポソーム複合体 1310 の密度は、図 1 の方法 1000 のステップ 1020 において加えるトランスポザーゼ酵素の量を変化させることにより、調整してもよい。

30

【0049】

ブリッジ分子 1415 の長さは、タグメント化反応で用いられた、トランスポソーム複合体 1310 が結合したビーズ 1120 の量に依存しない。同様に、タグメント化反応においてより多い又はより少ない DNA 1410 を加えることは、最終的なタグメント化産物のサイズを変えないが、反応の収率に影響を与える可能性がある。

40

【0050】

50

1つの例において、ビーズ1120は、常磁性ビーズである。この例では、タグメント化反応の精製は、ビーズ1120を磁石で固定化し、洗浄することにより容易に行うことができる。従って、タグメント化及びその後のPCR増幅を、単一の反応区画（「ワンポット（one-pot）」）での反応で実施してもよい。

【0051】

1つの態様において、本発明は、固体支持体上で標的核酸の連続性情報を引き出すことが可能なトランスポザーゼに基づく方法の、方法及び組成物に関する。いくつかの実施形態において、組成物及び方法は、アセンブリ／フェーズ情報を引き出すことができる。1つの実施形態において、固体支持体は、ビーズである。1つの実施形態において、標的核酸は、DNAである。1つの実施形態において、標的核酸は、ゲノムDNAである。いくつかの実施形態において、標的核酸は、RNAである。いくつかの実施形態において、RNAは、mRNAである。いくつかの実施形態において、標的核酸は、相補的DNA（cDNA）である。

【0052】

いくつかの実施形態において、トランスポゾン、ビーズ等の固体支持体に二量体として固定化し、その後トランスポザーゼをトランスポゾンに結合してトランスポソームを形成してもよい。

【0053】

いくつかの実施形態において、固相化トランスポゾン及びトランスポザーゼの添加による、固相でのトランスポソームの形成に特に関連して、2つのトランスポゾンを、固体支持体において互いにごく近接して（好ましくは、一定の距離で）固定化してもよい。このアプローチには、いくつかの利点がある。1つ目としては、好ましくは、2つのトランスポゾンがトランスポソームを効率良く形成するのに最適なリンカー長及び方向で、2つのトランスポゾンが、常に同時に固定化されることになる。2つ目としては、トランスポソームの形成効率が、トランスポゾン密度の関数とはならないであろうことである。2つのトランスポゾンが、トランスポソームを形成するのに適切な方向及び両者間の距離で、常に利用できることになる。3つ目としては、表面上のランダムな固定化トランスポゾンにより、トランスポゾン間に種々の距離が形成され、それにより、1つの画分のみがトランスポソームを効率良く形成するのに最適な方向及び距離を有する。結果として、全てのトランスポゾンがトランスポソームに変換するのではなく、固相化非複合体化トランスポゾンが存在することになる。これらのトランスポゾンは、ME部分が二本鎖DNAであるため、転位の標的となり易い。これにより、転位効率の低下をもたらし、望ましくない副産物を形成する可能性がある。従って、続けて使用して、タグメント化及びシーケンシングを通じて連続性情報を導き出すことができる固体支持体上に、トランスポソームを調製してもよい。例示的なスキームを図15に示す。いくつかの実施形態において、トランスポゾンは、化学的結合以外の手法で固体支持体に固定化してもよい。固体支持体上にトランスポゾンを固定化する例示的な方法としては、これに限定されないが、ストレプトアビジン-ビオチン、マルトース-マルトース結合タンパク質、抗原-抗体、DNA-DNA又はDNA-RNAハイブリダイゼーション等の親和結合が挙げられる。

【0054】

いくつかの実施形態において、トランスポソームは、プレアセンブルした後、固体支持体に固定化することができる。いくつかの実施形態において、トランスポゾンは、固有のインデックス、バーコード、及び増幅プライマー結合部位を含む。トランスポザーゼを、トランスポゾンを含む溶液に添加し、固体支持体上に固定化することが可能なトランスポソーム二量体を形成することができる。1つの実施形態において、各セットが固定化トランスポゾンに由来する同じインデックスを有し、それによりインデックス付きビーズを生成する、複数のビーズセットを生成することができる。図29Aに示すように、標的核酸を、インデックス付きビーズの各セットに添加することができる。

【0055】

いくつかの実施形態において、標的核酸は、インデックス付きビーズの各セットに添加

することができ、タグメント化及びその後のPCR増幅を別々に行ってもよい。

【0056】

いくつかの実施形態において、標的核酸、インデックス付きビーズ、及びトランスポソームは、多くの液滴が、1つのビーズと1つ又はそれ以上のDNA分子及び十分なトランスポソームとを含むように、液滴内で組み合わせることができる。

【0057】

いくつかの実施形態において、インデックス付きビーズをプールすることができ、プールに標的核酸を加えることができ、タグメント化及びその後のPCR増幅を単一の反応区画（「ワンポット」）で行ってもよい。

【0058】

1つの態様において、本発明は、連結転位標的核酸を固体支持体上に捕捉することにより、連続性情報を引き出す方法及び組成物に関する。いくつかの実施形態において、連続性保存転位（CPT: contiguity preserving transposition）をDNA上で行うが、DNAは、無傷（intact）のままであり（CPT-DNA）、従って、連結ライブラリーを形成する。連続性情報は、トランスポザゼを用いて標的核酸に隣接した鋳型核酸フラグメントの関連性（association）を維持することにより、保存することができる。CPT-DNAは、固体支持体、例えば、ビーズに固定化された、固有のインデックス又はバーコードを有する相補的オリゴヌクレオチドのハイブリダイゼーションにより捕捉することができる（図29B）。いくつかの実施形態において、固体支持体に固定化されたオリゴヌクレオチドは、バーコードに加えて、プライマー結合部位、固有分子インデックス（UMI: unique molecular indices）をさらに含んでもよい。

【0059】

有利なことに、このようにトランスポソームを用いて、フラグメント化された核酸の物理的近接性を保つことにより、同じ起源の分子、例えば、染色体からのフラグメント化された核酸が、同じ固有のバーコード及びインデックス情報を固体支持体に固定化されたオリゴヌクレオチドから受け取る可能性が増える。これにより、固有のバーコードを有する連結シーケンシングライブラリーが得られることになる。連結シーケンシングライブラリーをシーケンスして、連続配列情報を引き出すことができる。

【0060】

図16及び17は、固有のバーコード又はインデックスを有する連結ライブラリーを製作する本発明の上記態様の、例示的な実施形態の模式図を示す。例示的な方法は、CPT-DNAと、固有のインデックス及びバーコードを有する固体支持体上の固定化オリゴヌクレオチドとのライゲーション、及び鎖置換PCRを活用して、シーケンシングライブラリーを生成する。1つの実施形態において、クローンインデックス付きビーズは、ランダム又は特定のプライマー及びインデックス等の固定化DNA配列で生成してもよい。連結ライブラリーは、固定化オリゴヌクレオチドへのハイブリダイゼーションとその後のライゲーションにより、クローンインデックス付きビーズ上に捕捉することができる。分子内ハイブリダイゼーション捕捉は、分子間ハイブリダイゼーションよりもはるかに速いため、連続転位ライブラリーが、ビーズに「巻き付く（wrap around）」ことになる。図18及び19は、クローンインデックス付きビーズ上でのCPT-DNAの捕捉及び連続性情報の保存を示す。鎖置換PCRは、クローンビーズインデックス情報を個々の分子に転移することができる。従って、各連結ライブラリーは、特異的にインデックス付けされることになる。

【0061】

いくつかの実施形態において、固体支持体に固定化されたオリゴヌクレオチドは、一方の鎖は固体支持体に固定化され、もう一方の鎖は、固定化された鎖に部分的に相補的であることによりY-アダプターとなるような、部分的二本鎖構造を含むことができる。いくつかの実施形態において、固体支持体に固定化されたY-アダプターは、ライゲーション及びギャップ充填により連結タグメント化DNAに結合する。図20に示す。

10

20

30

40

50

【0062】

いくつかの実施形態において、Y-アダプターは、ビーズ等の固体支持体上のプローブ／インデックスを用いた、CPT-DNAのハイブリダイゼーション捕捉を通じて、形成される。図21は、このようなY-アダプターを作製する例示的なスキームを示す。これらのY-アダプターを使用することにより、潜在的に各フラグメントがシーケンシングライブラリーになる可能性があることを確かにする。これにより、シーケンシング当たりの適用範囲（coverage）が増加する。

【0063】

いくつかの実施形態において、遊離トランスポソームを、CPT-DNAから分離してもよい。いくつかの実施形態において、遊離トランスポソームの分離は、サイズ排除クロマトグラフィーによる。1つの実施形態において、分離は、MicroSpin S-400 HR Columns（ペンシルバニア州ピッツバーグ、GE Healthcare Life Sciences社）により行ってもよい。図22は、遊離トランスポソームから分離したCPT-DNAのアガロースゲル電気泳動を示す。

【0064】

ハイブリダイゼーションを通じた連結転位標的核酸の固体支持体上への捕捉には、いくつかの特有の利点がある。1つ目としては、方法は、ハイブリダイゼーションに基づくものであり、転位に基づくものではない。分子内ハイブリダイゼーション率>>分子間ハイブリダイゼーション率である。従って、単一の標的DNA分子の連続転位ライブラリーが固有インデックス付きビーズに巻き付く可能性は、2つ又はそれ以上の異なる単一の標的DNA分子が固有インデックス付きビーズに巻き付くのに比べてはるかに高い。2つ目としては、DNAの転位及び転位したDNAのバーコード化は、2つの別個のステップで生じる。3つ目としては、ビーズ上の活性化トランスポソームのアセンブリ及び固体表面上のトランスポソームの表面密度の最適化に関連した課題を、回避することができる。4つ目としては、自己転位産物をカラム精製により除去することができる。5つ目としては、連結転位DNAがギャップを含むため、DNAがより柔軟であり、それ故、トランスポソームをビーズに固定化する方法に比べて、転位密度（挿入サイズ）への負荷が少ない。6つ目としては、方法に、組み合わせ（combinatorial）バーコードスキームを用いることができる。7つ目としては、インデックス付きオリゴをビーズに共有結合させるのが容易である。従って、インデックス交換の可能性が少ない。8つ目としては、タグメント化及びその後のPCR増幅を多重化してもよく、単一反応区画（「ワンポット」）反応で行うことができるため、各インデックス配列に対して個々の反応を行う必要がなくなる。

【0065】

いくつかの実施形態において、転位の間に、複数の固有のバーコードを標的核酸全体にわたって挿入してもよい。いくつかの実施形態において、各バーコードは、間にフラグメント化部位が配置された第1のバーコード配列及び第2のバーコード配列を含む。第1のバーコード配列及び第2のバーコード配列は、互いにペアとなるように同定又は設計することができる。第1のバーコードと第2のバーコードとが関連するようにペアを作ることには、有益（informative）である可能性がある。有利なことに、ペアとなったバーコード配列を用いて、シーケンシングデータを鋳型核酸のライブラリーからアセンブルすることができる。例えば、第1のバーコード配列を含む第1の鋳型核酸、及び第1のバーコード配列とペアとなる第2のバーコード配列を含む第2の鋳型核酸を同定することは、第1及び第2の鋳型核酸が、標的核酸の配列表示において互いに隣接した配列を表すことを意味する。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、標的核酸の配列表示をデノボで（新たに）アセンブルすることができる。

【0066】

1つの態様において、本発明は、特定のDNAフラグメントのショットガン配列ライブラリーを生成する方法及び組成物に関する。

【0067】

1つの実施形態において、クローンビーズインデックス付きビーズを、固定化オリゴヌクレオチド配列：ランダム又は特異的なプライマー及び固有のインデックスで生成する。標的核酸を、クローンインデックス付きビーズに加える。いくつかの実施形態において、標的核酸は、DNAである。1つの実施形態において、標的DNAを変性させる。標的DNAは、固体表面（例えば、ビーズ）に固定化された固有のインデックスを含むプライマーにハイブリダイズし、続いて、同じインデックスを有する別のプライマーにハイブリダイズする。ビーズ上のプライマーは、DNAを増幅させる。1つ又はそれ以上の更なる増幅ラウンドを行ってもよい。1つの実施形態において、増幅は、3'ランダムn量体配列を有するビーズ固定化プライマーを用いて、全ゲノム増幅により行ってもよい。好ましい実施形態において、ランダムn量体は、増幅中のプライマー-プライマー相互作用を防ぐために、疑似相補的塩基（2-チオチミン、2-アミノdA、N4-エチルシトシン等）を含む（Hoshika, S; Chen, F; Leal, NA; Benner, SA, Angew. Chem. Int. Ed. 49 (32) 5554-5557 (2010)）。図23は、特定のDNAフラグメントのショットガン配列ライブラリーを生成する例示的なスキームを示す。クローンインデックス付きシーケンシングライブラリー及び増幅産物のライブラリーを生成することができる。1つの実施形態において、このようなライブラリーは、転位により生成することができる。インデックス情報を指針として用いることにより、クローンインデックス付きライブラリーの配列情報を用いて、連続性情報をアセンブルすることができる。図24は、クローンインデックス付きシーケンシングライブラリーから配列情報をアセンブルする例示的なスキームを示す。

【0068】

上記実施形態の方法には、いくつかの利点がある。ビーズ上での分子内増幅は、ビーズ間増幅よりもはるかに速い。従って、ビーズ上の生成物は、同じインデックスを有することになる。特定のDNAフラグメントのショットガンライブラリーを、作製することができる。ランダムプライマーは、ランダムな場所で鋳型を増幅するため、同じインデックスを有するショットガンライブラリーを特定の分子から生成することができ、インデックス付き配列を用いて配列情報をアセンブルすることができる。上記実施形態の方法の大きな利点は、反応を単一の反応（ワンポット反応）で多重化することができ、多くの個別のウェルを用いる必要がなくなることである。多くのインデックス付きクローンビーズを調製することができ、そのため、多くの異なるフラグメントを固有にラベルすることができ、同じゲノム領域に対して親対立遺伝子を識別することができる。多数のインデックスを用いることにより、父親のDNAコピー及び母親のDNAコピーが同じゲノム領域に対して同じインデックスを受け取る可能性は低い。当該方法は、内（intra）反応が間（inter）反応よりもはるかに速いという事実を利用するものであり、ビーズは、大きな物理的区画において実質的な仕切りを基本的に作り出す。

【0069】

本発明の全ての上記態様のうちのいくつかの実施形態において、方法を、無細胞DNA（cfDNA: cell free DNA）アッセイにおいてcfDNAに用いてもよい。いくつかの実施形態において、cfDNAは、血漿、胎盤液から得る。

【0070】

1つの実施形態において、血漿は、膜ベースの沈降支援血漿分離器を用いて無希釈の全血から得ることができる（Liu et al. Anal Chem. 2013 Nov 5; 85 (21): 10463-70）。1つの実施形態において、血漿分離器における血漿採取ゾーンは、トランスポソームを含む固体支持体を含んでもよい。トランスポソームを含む固体支持体は、血漿が全血から分離される時に、単離された血漿からcfDNAを捕捉してもよく、cfDNAの濃縮及び／又はDNAのタグメント化を行うことができる。いくつかの実施形態において、タグメント化は、固有のバーコードをさらに導入することにより、続いて分離（demultiplexing）を、ライブラリープールのシーケンシング後に行うことを可能にするであろう。

【0071】

いくつかの実施形態において、分離器の採取ゾーンは、PCRマスターミックス（プライマー、ヌクレオチド、バッファー、金属）及びポリメラーゼを含んでもよい。1つの実施形態において、マスターミックスは、血漿が分離器から出てくる時に再構成されるように、乾燥状態であってもよい。いくつかの実施形態において、プライマーは、ランダムプライマーである。いくつかの実施形態において、プライマーは、特定の遺伝子に対する特異的プライマーであってもよい。cfDNAのPCR増幅の結果として、分離された血漿から直接ライブラリーを生成することになる。

【0072】

いくつかの実施形態において、分離器の採取ゾーンは、RT-PCRマスターミックス（プライマー、ヌクレオチド、バッファー、金属）、逆転写酵素、及びポリメラーゼを含んでもよい。いくつかの実施形態において、プライマーは、ランダムプライマー又はオリゴdTプライマーである。いくつかの実施形態において、プライマーは、特定の遺伝子に対する特異的プライマーであってもよい。得られたcDNAは、シーケンシングに用いることができる。或いは、cDNAは、配列ライブラリー調製のために、固体支持体に固定化されたトランスポソームで処理してもよい。

【0073】

いくつかの実施形態において、血漿分離器は、バーコード（1D又は2Dバーコード）を含んでもよい。いくつかの実施形態において、分離装置は、採血装置を有していてもよい。これにより、血液を血漿分離器及びライブラリー調製装置に直接送ることになる。いくつかの実施形態において、装置は、下流配列分析器を有していてもよい。いくつかの実施形態において、配列分析器は、単回使用シーケンサーである。いくつかの実施形態において、シーケンサーは、まとめてシーケンシングする前に、サンプルの列を作ることができる。或いは、シーケンサーは、サンプルがシーケンシング領域に送達される、ランダムアクセス機能を有していてもよい。

【0074】

いくつかの実施形態において、血漿用採取ゾーンは、無細胞DNAが濃縮されるように、シリカ基質を含んでもよい。

【0075】

フェージング及びメチル化の同時検出

5-メチルシトシン（5-Me-C）及び5-ヒドロキシメチルシトシン（5-ヒドロキシ-C）は、エピジェネティック（Epi）修飾としても知られ、細胞代謝、分化、及び癌増殖において重要な役割を果たす。本願の発明者らは、驚くべきことに、且つ予想外にも、フェージング及び同時のメチル化の検出が、本願の方法及び組成物を用いて可能であることを見出した。本願の方法は、ビーズ上でのCPTシーケンシング（CPT-seq）（インデックス付き連結ライブラリー）とDNAメチル化検出とを組み合わせることを可能にするものである。例えば、ビーズ上に生成された個々のライブラリーを亜硫酸水素塩で処理して、非メチル化CをUに変換するが、メチル化CはUに変換しないことにより、5-Me-Cの検出を可能にすることができる。ヘテロ接合SNPを用いた更なるフェージング分析を通じて、Epi-メチル化-フェージングブロックを複数のメガ塩基領域で確立することができる。

【0076】

いくつかの実施形態において、分析されるDNAのサイズは、約100塩基～約複数メガ塩基まで可能である。いくつかの実施形態において、分析されるDNAのサイズは、約100塩基、200塩基、300塩基、400塩基、500塩基、600塩基、700塩基、800塩基、900塩基、1000塩基、1200塩基、1300塩基、1500塩基、2000塩基、3000塩基、3500塩基、4000塩基、4500塩基、5000塩基、5500塩基、6000塩基、6500塩基、7000塩基、7500塩基、8000塩基、8500塩基、9000塩基、9500塩基、10,000塩基、10,500塩基、11,000塩基、11,500塩基、12,000塩基、12,500塩基、13,000塩基、14,000塩基、14,500塩基、15,000塩基、15,

5 0 0 塩基、1 6 , 0 0 0 塩基、1 6 , 5 0 0 塩基、1 7 , 0 0 0 塩基、1 7 , 5 0 0 塩基、1 8 , 0 0 0 塩基、1 8 , 5 0 0 塩基、1 9 , 0 0 0 塩基、1 9 , 5 0 0 塩基、2 0 , 0 0 0 塩基、2 0 , 5 0 0 塩基、2 1 , 0 0 0 塩基、2 1 , 5 0 0 塩基、2 2 , 0 0 0 塩基、2 2 , 5 0 0 塩基、2 3 , 0 0 0 塩基、2 3 , 5 0 0 塩基、2 4 , 0 0 0 塩基、2 4 , 5 0 0 塩基、2 5 , 0 0 0 塩基、2 5 , 5 0 0 塩基、2 6 , 0 0 0 塩基、2 6 , 5 0 0 塩基、2 7 , 0 0 0 塩基、2 7 , 5 0 0 塩基、2 8 , 0 0 0 塩基、2 8 , 5 0 0 塩基、2 9 , 5 0 0 塩基、3 0 , 0 0 0 塩基、3 0 , 5 0 0 塩基、3 1 , 0 0 0 塩基、3 1 , 5 0 0 塩基、3 2 , 0 0 0 塩基、3 3 , 0 0 0 塩基、3 4 , 0 0 0 塩基、3 5 , 0 0 0 塩基、3 6 , 0 0 0 塩基、3 7 , 0 0 0 塩基、3 8 , 0 0 0 塩基、3 9 , 0 0 0 塩基、4 0 , 0 0 0 塩基、4 2 , 0 0 0 塩基、4 5 , 0 0 0 塩基、5 0 , 0 0 0 塩基、5 5 , 0 0 0 塩基、6 0 , 0 0 0 塩基、6 5 , 0 0 0 塩基、7 0 , 0 0 0 塩基、7 5 , 0 0 0 塩基、8 0 , 0 0 0 塩基、8 5 , 0 0 0 塩基、9 0 , 0 0 0 塩基、9 5 , 0 0 0 塩基、1 0 0 , 0 0 0 塩基、1 1 0 , 0 0 0 塩基、1 2 0 , 0 0 0 塩基、1 3 0 , 0 0 0 塩基、1 4 0 , 0 0 0 塩基、1 5 0 , 0 0 0 塩基、1 6 0 , 0 0 0 塩基、1 7 0 , 0 0 0 塩基、1 8 0 , 0 0 0 塩基、2 0 0 , 0 0 0 塩基、2 2 5 , 0 0 0 塩基、2 5 0 , 0 0 0 塩基、3 0 0 , 0 0 0 塩基、3 5 0 , 0 0 0 塩基、4 0 0 , 0 0 0 塩基、4 5 0 , 0 0 0 塩基、5 0 0 , 0 0 0 塩基、5 5 0 , 0 0 0 塩基、6 0 0 , 0 0 0 塩基、6 5 0 , 0 0 0 塩基、7 0 0 , 0 0 0 塩基、7 5 0 , 0 0 0 塩基、8 0 0 , 0 0 0 塩基、8 5 0 , 0 0 0 塩基、9 0 0 , 0 0 0 塩基、1 , 0 0 0 , 0 0 0 塩基、1 , 2 5 0 , 0 0 0 塩基、1 , 5 0 0 , 0 0 0 塩基、2 , 0 0 0 , 0 0 0 塩基、2 , 5 0 0 , 0 0 0 塩基、3 , 0 0 0 , 0 0 0 塩基、4 , 0 0 0 , 0 0 0 塩基、5 , 0 0 0 , 0 0 0 塩基、6 , 0 0 0 , 0 0 0 塩基、7 , 0 0 0 , 0 0 0 塩基、8 , 0 0 0 , 0 0 0 塩基、9 , 0 0 0 , 0 0 0 塩基、1 0 , 0 0 0 , 0 0 0 塩基、1 5 , 0 0 0 , 0 0 0 塩基、2 0 , 0 0 0 , 0 0 0 塩基、3 0 , 0 0 0 , 0 0 0 塩基、4 0 , 0 0 0 , 0 0 0 塩基、5 0 , 0 0 0 , 0 0 0 塩基、7 5 , 0 0 0 , 0 0 0 塩基、1 0 0 , 0 0 0 , 0 0 0 塩基、又はそれ以上である。

【0077】

5-ヒドロキシーC、DNA酸化生成物、DNAアルキル化生成物、ヒストン末端(histone-foot)プリンティング等の他のEpi修飾はまた、本願に開示する方法及び組成物を用いてフェージングの中で分析することもできる。

【0078】

いくつかの実施形態において、DNAは、初めに、固体支持体上でインデックス付き結合ライブラリーに変換される。元のDNAよりもはるかに小さい個々のインデックス付きライブラリーは、個々のライブラリーがより小さいため、フラグメント化され難い。インデックス付きライブラリーの小画分が消失したとしても、フェージング情報は、インデックス付きDNA分子の全長にわたって依然として維持される。例えば、100kbの分子の場合、従来の亜硫酸水素塩変換(BSC)では半分にフラグメント化され、連続性はもはや50kbに制限される。本明細書に開示する方法では、100kbのライブラリーは、初めにインデックス化され、個々のライブラリーの画分が消失しても、連続性は、依然として~100kbである(全ライブラリーがDNA分子の一端から消失する稀な事態を除いて)。また、本明細書に開示する方法は、従来の亜硫酸水素塩変換アプローチでは必要とされるのとは対照的に、更なる精製ステップが必要とされないことにより、収率が上がるため、さらに利点を有する。本明細書に開示する方法では、ビーズは、亜硫酸水素塩変換の後に洗浄するだけである。さらに、DNAが固相に結合しているままで、DNA(インデックス付きライブラリー)の最小限の消失及び少ない手間でバッファー交換を容易に行うことができる。

【0079】

フェージング及びメチル化の同時検出の例示的なスキームを図43に示す。操作フローは、ビーズ上でのDNAのタグメント化、9塩基対の反復領域のギャップ充填ライゲーシオン、SDSによるTn5の除去、及びビーズ上の個々のライブラリーの亜硫酸水素塩変換からなる。隣接する相補的ライブラリーが再アニールしないことを確実にするために、

亜硫酸水素塩変換を変性条件下で行い、それにより、亜硫酸水素塩変換効率を低下させる。B S Cは、非メチル化CをUに変換し、メチル化Cは、変換されない。

【0080】

図44は、フェージング及びメチル化の同時検出の別の例示的なスキームを示す。転位後にシーケンシングライブラリーを調製した後、一本鎖鋳型を調製するために、ギャップ充填ライゲーションしたライブラリーの画分を分解する。一本鎖鋳型は、既に鋳型が、ライブラリーの消失を低減すること又は亜硫酸水素塩変換効率を改善することができる一本鎖であるため、亜硫酸水素塩変換に対してより穏やかな条件を必要とする。1つの実施形態において、3'チオ保護トランスポゾン（Exo抵抗性）及び非保護トランスポゾンの混合物を、同じビーズ上で用いる。酵素、例えば、Exo Iを用いて、非チオ保護ライブラリーを分解し、それらを一本鎖ライブラリーに変換することができる。チオ保護トランスポゾン：非保護トランスポゾンが50：50である混合物を用いることにより、ライブラリーの50%を一本鎖ライブラリーに変換し（50%では、ライブラリーの1つのトランスポゾンは保護されており、1つのトランスポゾン（相補鎖）は保護されていない）、25%は変換せず（両方のトランスポゾンがチオ保護されている）、25%は両方とも変換してライブラリー全体を除去する（両方のトランスポゾンが保護されていない）。

【0081】

ストレプトアビジン磁性ビーズ等の固相に結合したDNAの亜硫酸水素塩変換を行う上での1つの課題は、DNAが結合したビーズを、亜硫酸水素ナトリウムにより高温で長時間処理することで、DNA及びビーズの両方に損傷を与えることである。DNA損傷の回復を助けるため、キャリアDNA（即ち、ラムダDNA）を亜硫酸水素塩処理前に反応混合物に加える。キャリアDNAが存在しても、当初のDNAの約80%が消失することが予測された。結果として、CPTSeq連続性ブロックは、従来のCPTSeqプロトコルよりも少ないメンバーを有する。

【0082】

従って、本明細書において、Epi-CPTSeqプロトコルのDNA収量を改善するためのいくつかのストラテジーを提案する。第1のストラテジーは、ストレプトアビジンビーズにトランスポソーム複合体をより密集させることで、ライブラリー挿入サイズを小さくすることに依拠する。ライブラリーサイズを小さくすることにより、より少ない割合のライブラリーエレメントが亜硫酸水素塩処理により分解される。

【0083】

Epi-CPTSeqプロトコルのDNA収量を改善するための第2のストラテジーは、破損したライブラリーエレメントを酵素により回復させることである。回復ストラテジーの目的は、ライブラリー増幅に必要な3'共通配列を、亜硫酸水素塩処理中に3'部分が分解及び消失したビーズ結合ライブラリーエレメントに再び加えることである。3'共通配列を加えた後は、これらのエレメントをPCR増幅及びシーケンシングすることができる。図67及び68は、このストラテジーの例示的なスキームを示す。二本鎖CPTSeqライブラリーエレメントを変性し、亜硫酸水素塩変換する（上段）。亜硫酸水素塩変換の間に、DNA鎖の1つが損傷を受け（中段）、3'末端上のPCR共通配列が消失する。鋳型救出ストラテジーにより、PCR増幅に必要な3'共通配列（緑色）を回復させる（下段）。1つの例において、3'リン酸化アテニューエーターオリゴ、即ち、シーケンシングアダプターとそれに続くオリゴdTストレッチを含む配列の存在下で、ターミナルトランスフェラーゼを使用する（図68A）。簡単に言えば、Td Tが、10～15 dAのストレッチを、アテニューエーターオリゴのオリゴdT部分にアニールする破損ライブラリーエレメントの3'末端に、加える。このDNAハイブリッドの形成により、Td T反応が停止し、破損ライブラリーエレメントの3'末端をDNAポリメラーゼにより結果的に伸長させるための鋳型をもたらす。

【0084】

別の操作フロー（図68B）では、Td Tテーリング反応を、一本鎖オリゴdT部分及び5'リン酸化二本鎖シーケンシングアダプター部分を有する部分的二本鎖アテニューエ

10

20

30

40

50

ーターオリゴの存在下で行う。T d T 反応の終結時、最後に付加された d A と 5' リン酸化アテニュエーターオリゴとの間のニックを、DNA リガーゼで封止する。

【0085】

記載した操作フローのどちらも、近年開発された制御可能な T d T テーリング反応に依拠しており、米国特許出願公開第 2015/0087027 号明細書に記載されている。共通シーケンシングアダプターはまた、近年導入された MMLV RT の ssDNA 鑄型スイッチング活性により、破損ライブラリーエレメントの 3' 末端に付加することができる。つまり、MMLV RT 及び鑄型スイッチオリゴ (TS オリゴ: template switch oligo) を、損傷 DNA に加える (図 68C)。この反応の最初のステップにおいて、逆転写酵素は、一本鎖 DNA フラグメントの 3' 末端にいくつかの追加のヌクレオチドを付加し、これらの塩基は、TS オリゴの 1 つの 3' 末端に存在するオリゴ (N) 配列とペアを作る。次いで、逆転写酵素の鑄型スイッチング活性により、アニールされた共通プライマーの配列を BSC 破損ライブラリーエレメントの 3' 末端に付加し、共通シーケンシングプライマーによる PCR で増幅される能力を回復させる。

【0086】

第 3 の戦略の一部として、Epicycle 社の EpiGenome キット「ポスト亜硫酸水素塩変換」ライブラリー構築法を用いて、亜硫酸水素塩変換中に 3' 末端の共通配列を消失したライブラリーエレメントを救出することができる。図 69 に示すように、このライブラリー救出法は、共通配列及びそれに続く短いストレッチのランダム配列を有する 3' リン酸化オリゴを利用する。これらの短いランダム配列は、亜硫酸水素塩処理一本鎖 DNA にハイブリダイズし、続いて共通配列が、DNA ポリメラーゼにより破損ライブラリー鎖に複製される。

【0087】

図 74 は、ビーズ上での亜硫酸水素塩シーケンシング法を改良する第 4 の戦略を示す。捕捉タグを含む第 1 の共通配列を、DNA の 5' 末端に共有結合させる。第 1 の共通配列は、片側転位 (図示する)、アダプターライゲーション、又は米国特許出願公開第 2015/0087027 号明細書に記載されたターミナルトランスフェラーゼ (T d T) アダプターライゲーションを含む、種々の方法を用いて DNA に結合させることができる。

【0088】

次に、DNA を変性させ (例えば、高熱でのインキュベーション)、固体支持体に結合させる。ビオチンを CS1 上の捕捉タグとして使用する場合は、例えば、DNA は、ストレプトアビジン磁性ビーズ (図示する) を用いて結合させることができる。固体支持体に結合してしまえば、バッファー交換を容易に行うことができる。

【0089】

次のステップにおいて、ssDNA の亜硫酸水素塩変換を行う。一本鎖の形では、DNA は、亜硫酸水素塩変換に容易に利用できるはずであり、Promega 社の Methyl Edge BSC キットの改良版を用いて変換効率 95% まで観察した (図 75)。

【0090】

亜硫酸水素塩変換後、第 2 の共通配列を、固体支持体に結合した ssDNA の 3' 末端に共有結合させる。オリゴを ssDNA に共有結合させるためのいくつかの方法を、上述してきた。T d T アテニュエーター／アダプターライゲーション法を用いて、>95% のライゲーション効率を達成した。結果として、提案したメチルシーケンシング (MethylSeq) 操作フローを用いた最終ライブラリーの収量は、既存の方法よりも高くなるはずである。

【0091】

最終ステップにおいて、PCR を行ってライブラリーを増幅し、ライブラリーを固体支持体から除去する。PCR プライマーは、シーケンシングアダプター等の追加の共通配列を、MethylSeq ライブラリーの末端に付加するようにデザインすることができる。

10

20

30

40

50

【0092】

単一アッセイにおける様々なサイズのライブラリーの調製

ゲノムのアセンブリの精度は、様々な長さスケールの技術の使用次第である。例えば、ショットガン（数百bp）－メイトペア（～3Kb）から－Hi-C（Mbスケール）は全て、アセンブリ及びコンティグ長を経時的に改良する方法である。課題は、これを行うために多重アッセイが必要であり、多層アプローチを扱い難く且つ費用が掛かるものとしていることである。本明細書に開示する組成物及び方法は、単一のアッセイで複数の長さスケールに対応することができる。

【0093】

いくつかの実施形態において、ライブラリー調製を、サイズの異なる固体支持体、例えば、ビーズを用いて単一アッセイで達成することができる。各ビーズのサイズは、ビーズの物理的サイズがライブラリーサイズを決定し、特定のライブラリーサイズ又はサイズ範囲をもたらすことになる。様々なサイズのビーズは全て、ライブラリーに転移する固有のクローンインデックスを有する。従って、様々なサイズのライブラリーが、固有にインデックス付けされた異なる各ライブラリースケール長で生成される。様々な長さスケールのライブラリーを同じ物理的区画で同時に調製するため、コストが減り、操作フロー全体を改善する。いくつかの実施形態において、各特定の固体支持体サイズ、例えば、ビーズサイズは、固有のインデックスを受け取る。いくつかの他の実施形態において、同じ固体支持体サイズ、例えば、同じビーズサイズの複数の異なるインデックスもまた調製することで、複数のDNA分子をそのサイズ範囲に対してインデックス区分することができる。図45は、単一アッセイで、様々なサイズのクローンインデックス付きビーズを用いて様々なサイズのライブラリーを生成する例示的なスキームを示す。

【0094】

いくつかの実施形態において、生成されるライブラリーのサイズは、約50塩基、75塩基、100塩基、150塩基、200塩基、250塩基、300塩基、350塩基、400塩基、500塩基、600塩基、700塩基、800塩基、900塩基、1000塩基、1200塩基、1300塩基、1500塩基、2000塩基、3000塩基、3500塩基、4000塩基、4500塩基、5000塩基、5500塩基、6000塩基、6500塩基、7000塩基、7500塩基、8000塩基、8500塩基、9000塩基、9500塩基、10,000塩基、10,500塩基、11,000塩基、11,500塩基、12,000塩基、12,500塩基、13,000塩基、14,000塩基、14,500塩基、15,000塩基、15,500塩基、16,000塩基、16,500塩基、17,000塩基、17,500塩基、18,000塩基、18,500塩基、19,000塩基、19,500塩基、20,000塩基、20,500塩基、21,000塩基、21,500塩基、22,000塩基、22,500塩基、23,000塩基、23,500塩基、24,000塩基、24,500塩基、25,000塩基、25,500塩基、26,000塩基、26,500塩基、27,000塩基、27,500塩基、28,000塩基、28,500塩基、29,500塩基、30,000塩基、30,500塩基、31,000塩基、31,500塩基、32,000塩基、33,000塩基、34,000塩基、35,000塩基、36,000塩基、37,000塩基、38,000塩基、39,000塩基、40,000塩基、42,000塩基、45,000塩基、50,000塩基、55,000塩基、60,000塩基、65,000塩基、70,000塩基、75,000塩基、80,000塩基、85,000塩基、90,000塩基、95,000塩基、100,000塩基、110,000塩基、120,000塩基、130,000塩基、140,000塩基、150,000塩基、160,000塩基、170,000塩基、180,000塩基、200,000塩基、225,000塩基、250,000塩基、300,000塩基、350,000塩基、400,000塩基、450,000塩基、500,000塩基、550,000塩基、600,000塩基、650,000塩基、700,000塩基、750,000塩基、800,000塩基、850,000塩基、900,000塩基、1,000,000塩基、1,

10

20

30

40

50

250,000塩基、1,500,000塩基、2,000,000塩基、2,500,000塩基、3,000,000塩基、4,000,000塩基、5,000,000塩基、6,000,000塩基、7,000,000塩基、8,000,000塩基、9,000,000塩基、10,000,000塩基、15,000,000塩基、20,000,000塩基、30,000,000塩基、40,000,000塩基、50,000,000塩基、75,000,000塩基、100,000,000塩基、又はそれ以上である。

【0095】

いくつかの実施形態において、上述の複数の長さスケールのライブラリーを、1つの大きな長さスケールを有する代わりに、疑似遺伝子、パラログ等のアセンブリに用いることができる。いくつかの実施形態において、複数の長さスケールのライブラリーを、単一アッセイで同時に調製する。利点は、少なくとも1つの長さスケールが、疑似遺伝子又は遺伝子のみを有し、両方は有さない固有領域に結合することである。従って、この長さスケールで検出される変異は、遺伝子又は疑似遺伝子のどちらかに変異を一意的に決めることができる。同じことがコピー数の変異、パラログ等にも当てはまる。アセンブリの長所は、異なる長さスケールの使用である。本明細書に開示する方法を用いることにより、異なる長さスケールのインデックス付き結合ライブラリーを、異なる長さスケールのための個々の異なるライブラリー調製を行わずに、単一のアッセイで生成することができる。図46は、異なる長さスケールのライブラリーで遺伝子変異を決定する例示的なスキームを示す。

【0096】

遺伝子変異の分析

本明細書に開示する組成物及び方法は、遺伝子変異の分析に関する。例示的な遺伝子変異としては、これに限定されないが、欠損、染色体間転位、重複、パラログ、染色体間遺伝子融合が挙げられる。いくつかの実施形態において、本明細書に開示する組成物及び方法は、遺伝子変異のフェージング情報の決定に関する。以下の表は、例示的な染色体間遺伝子融合を示す。

【0097】

【表1】

表1：染色体間遺伝子融合

染色体間									
BP1 染色体	BP1 開始	BP1 終了	長さ	BP2 染色体	BP2 開始	BP2 終了	長さ	Supp.	CPT検出
chr2	3939448	3935237	4219	chr12	124494434	124500374	6030	100	可
chr2	11349818	11364248	6220	chr21	12143482	12149318	6034	104	可
chr3	73160440	73182742	4302	chr17	41379324	41384337	5013	80	可
chr3	73165440	73162742	4302	chr17	41367704	41403394	6190	132	可
chr3	73365674	73396286	13712	chr30	26196677	26213307	16230	271	可
chr3	87146382	87348214	2002	chr29	81828873	81828137	6264	54	不可
chr3	111971673	111275046	3472	chr8	128528232	128538379	9547	138	可
chr5	48603586	48602710	1762	chr22	38769007	38719736	6728	64	不可
chr7	81708161	81734472	6312	chr10	60900343	60904322	4079	86	可
chr7	100936524	100940377	3853	chr11	114417647	114425198	7258	45	可
chr7	111939829	111866168	8340	chr12	108196690	108203183	3683	85	可
chr8	28431882	28436443	4561	chr17	78768768	78762846	4298	51	不可
chr9	12824462	12827476	3028	chr22	29085133	29087395	2262	13	可
chr9	98381389	9838424	6085	chr13	19443778	19450674	6098	29	可
chr9	101473686	101479687	6001	chr22	32575696	32579748	4052	87	可
chr9	109772276	109777707	5431	chr29	28873630	28878949	5319	70	可
chr9	121298728	121356890	6862	chr11	11990608	11992376	21487	178	可
chr12	58987510	58991001	4291	chr15	29959272	29959958	3968	25	可
chr13	158177158	158182388	5230	chr19	12896441	12894258	7817	72	可
chr13	63618689	63624163	5484	chr17	21660385	21670256	3678	115	可

【0098】

表 2 は、染色体 1 における例示的な欠損を示す。

【0099】

【表 2】

表 2：染色体 1 における例示的な欠損

染色体	欠損				サイズ (bp)		CPTシグナル
	BP1 開始	BP1 終了	BP2 開始	BP2 終了			
chr1	1,218,882	12,234,834	3,452	13,255,777	12,036,895	35,38	有り
chr1	8,559,198	9,635,541	3,356	9,597,237	9,000,038	38,62	有り
chr1	14,433,353	14,435,290	3,654	14,437,703	14,441,375	38,63	有り
chr1	25,154,326	25,158,976	4,443	25,158,154	25,155,093	35,79	有り
chr1	25,455,596	25,458,338	2,940	25,453,739	25,455,542	47,53	有り
chr1	34,989,513	34,991,181	5,863	34,992,339	34,995,952	43,23	有り
chr1	55,827,433	55,831,036	3,596	55,834,939	55,838,738	38,49	無し
chr1	63,732,441	63,735,334	2,893	63,732,26	63,732,091	38,65	??
chr1	71,234,178	71,237,339	3,163	71,236,06	71,240,648	41,43	有り
chr1	73,453,444	73,454,45	3,001	73,454,939	73,455,955	38,66	??
chr1	81,403,458	81,404,472	4,004	81,405,503	81,413,953	50,13	有り
chr1	83,455,558	83,458,370	3,811	83,477,134	83,482,174	49,60	有り
chr1	84,513,313	84,517,907	4,594	84,524,953	84,528,435	38,05	有り
chr1	84,708,940	84,711,988	3,048	84,715,225	84,719,553	38,34	有り
chr1	97,557,768	97,560,911	3,142	97,563,327	97,572,164	28,77	無し
chr1	1,08E+08	1,08E+08	4,894	1,08E+08	1,08E+08	37,23	有り
chr1	1,07E+08	1,07E+08	5,095	1,07E+08	1,07E+08	51,41	??
chr1	1,09E+08	1,09E+08	2,959	1,09E+08	1,09E+08	40,18	有り
chr1	1,1E+08	1,1E+08	5,059	1,1E+08	1,1E+08	44,70	??
chr1	1,13E+08	1,13E+08	5,237	1,13E+08	1,13E+08	45,28	有り
chr1	1,15E+08	1,15E+08	4,594	1,15E+08	1,15E+08	45,31	??
chr1	1,45E+08	1,45E+08	5,056	1,45E+08	1,45E+08	47,31	有り
chr1	1,59E+08	1,59E+08	4,541	1,59E+08	1,59E+08	38,61	有り(二重欠損)
chr1	1,79E+08	1,79E+08	4,585	1,79E+08	1,79E+08	51,91	有り
chr1	1,79E+08	1,79E+08	4,630	1,79E+08	1,79E+08	38,17	有り
chr1	1,85E+08	1,85E+08	4,475	1,85E+08	1,85E+08	33,16	有り
chr1	1,97E+08	1,97E+08	4,323	1,97E+08	1,97E+08	38,94	無し
chr1	2,06E+08	2,06E+08	4,995	2,06E+08	2,06E+08	52,15	有り
chr1	2,06E+08	2,06E+08	3,536	2,06E+08	2,06E+08	49,64	有り
chr1	2,1E+08	2,1E+08	3,973	2,1E+08	2,1E+08	29,43	有り(二重欠損)
chr1	2,3E+08	2,3E+08	3,828	2,3E+08	2,3E+08	30,79	有り
chr1	2,32E+08	2,32E+08	5,031	2,32E+08	2,32E+08	37,58	??
chr1	2,38E+08	2,38E+08	3,548	2,38E+08	2,38E+08	39,03	有り
chr1	2,4E+08	2,4E+08	3,457	2,4E+08	2,4E+08	38,73	有り

【0100】

いくつかの実施形態において、標的核酸は、標的核酸をトランスポソームに曝す前にフラグメント化することができる。例示的なフラグメント化法としては、これに限定されないが、超音波処理、機械的剪断、及び制限消化が挙げられる。タグメント化（フラグメント化及びタグ化）前の標的核酸のフラグメント化は、疑似遺伝子（例えば、CYP2D6）のアセンブリ／フェージングに有利である。インデックス付き結合リードの長い島（＞30kb）は、図64に示すように、疑似遺伝子A及びA'に及ぶであろう。高い配列相同性のため、どの変異が遺伝子A及び遺伝子A'に属するのかを決定することが課題となるであろう。より短い変異は、固有の周囲配列で疑似遺伝子の1つの変異に結合するであろう。そのようなより短い島は、タグメント化前に標的核酸をフラグメント化することにより達成することができる。

【0101】

結合トランスポソーム

いくつかの実施形態において、トランスポザーゼは、トランスポソーム複合体中で多量体であり、例えば、二量体、四量体等をトランスポソーム複合体中で形成する。本願の発明者らは、驚くべきことに、且つ予想外にも、多量体トランスポソーム複合体中の単量体トランスポザーゼを結合させること、又は多量体トランスポソーム複合体中のトランスポソーム単量体のトランスポゾン末端を結合させることには、いくつかの利点があることを見出した。1つ目としては、トランスポザーゼ又はトランスポゾンの結合は、より安定した複合体をもたらし、大きな画分が活性化状態にある。2つ目としては、より低い濃度のトランスポソームを、転位反応によるフラグメント化において用いることができる可能性がある。3つ目としては、結合により、トランスポソーム複合体のモザイク末端(ME)の交換が少なくなり、それにより、バーコード又はアダプター分子の混合が少なくなる。そのようなME末端の交換は、複合体がバラバラになり、再編成する場合、又は、トランスポソームがストレプトアビジン／ビオチンにより固体支持体上に固定化され、ストレプトアビジン／ビオチン相互作用が壊れて再編成する場合、又は、コンタミネーションの可能性がある場合に、生じる可能性がある。本願の発明者らは、種々の反応条件下で、ME末端の重大な交換(swap又はexchange)があることに気付いた。いくつかの実施形態において、交換は、15%にまで達する可能性がある。交換は、高塩濃度バッファーで顕著であり、グルタミン酸バッファーでは低下する。図57及び58は、ME交換のいくつかの考え得るメカニズムを示す。

10

20

【0102】

いくつかの実施形態において、トランスポソーム複合体中のトランスポザーゼのサブユニットは、共有及び非共有手段により互いに結合させることができる。いくつかの実施形態において、トランスポザーゼ単量体は、トランスポソーム複合体を形成する前に(トランスポゾンが加わる前に)結合させることができる。いくつかの実施形態において、トランスポザーゼ単量体は、トランスポソームの形成後に結合させることができる。

【0103】

いくつかの実施形態において、天然アミノ酸残基を多量体界面でシステイン(Cys)アミノ酸で置換し、ジスルフィド結合の形成を促進していてもよい。例えば、Tn5トランスポザーゼにおいて、Asp468、Tyr407、Asp461、Lys459、Ser458、Gly462、Ala466、Met470をCysで置換して、単量体サブユニット間のジスルフィド結合を促進していてもよい。図59及び60に示す。Mos-1トランスポザーゼに関して、システインで置換することができる例示的なアミノ酸としては、これに限定されないが、Leu21、Leu32、Ala35、His20、Phe17、Phe36、Ile16、Thr13、Arg12、Gln10、Glu9が挙げられる。図61に示す。いくつかの実施形態において、システインで置換されたアミノ酸残基を有する修飾トランスポザーゼは、マレイミド又はピリジルジチオール反応基を用いた化学架橋剤を用いて、互いに化学的に架橋させることができる。例示的な化学架橋剤は、Pierce Protein Biology/ThermoFisher Scientific社(米国ニューヨーク州グランドアイランド)から市販されている。

30

40

【0104】

いくつかの実施形態において、トランスポソーム多量体複合体は、固体支持体に共有結合させることができる。例示的な固体支持体としては、これに限定されないが、ナノ粒子、ビーズ、フローセル表面、カラムマトリックスが挙げられる。いくつかの実施形態において、固体表面は、アミン基で被覆されていてもよい。システインで置換されたアミノ酸残基を有する修飾トランスポザーゼは、そのようなアミン基に対し、アミン-スルフヒドリル架橋剤(即ち、スクシンイミジル-4-(N-マレイミドメチル)シクロヘキサン-1-カルボキシレート(SMCC))を用いて、化学的に架橋させることができる。例示的なスキームを図62に示す。いくつかの実施形態において、マレイミド-PEG-ビオチン架橋剤を用いて、dTnpをストレプトアビジン被覆固体表面に結合させてもよい。

50

【0105】

いくつかの実施形態において、トランスポザーゼ遺伝子は、単一のポリペプチドで多量体タンパク質を発現するように、修飾することができる。例えば、Tn5又はMos-1遺伝子は、単一のポリペプチドで、2つのTn5又はMos-1タンパク質を発現するように修飾することができる。同様に、Muトランスポザーゼ遺伝子は、単一のポリペプチドで、4つのmuトランスポザーゼユニットをコードするように修飾することができる。

【0106】

いくつかの実施形態において、トランスポソーム単量体単位のトランスポゾン末端を結合させて、結合トランスポソーム多量体複合体を形成することができる。トランスポゾン末端を結合させることにより、プライマー部位を挿入することができ、シーケンシングプライマー、増幅プライマー、又は任意のロール (role) DNAが、標的DNAをフラグメント化することなくgDNAに働くことができる。そのような機能性の挿入は、情報が無傷分子から抽出する必要があるか又はサブサンプリングが重要である、ハプロタイプアッセイ又は結合部タグ化アッセイにおいて、利点である。いくつかの実施形態において、Muトランスポソームのトランスポゾン末端は、「ループ状の」Muトランスポザーゼ／トランスポゾン構造に結合させることができる。Muは四量体であるため、これに制限されないが、R2UJ及び／又はR1UJをR2J及び／又はR1Jに結合させることにより、種々の構造が可能である。これらの構造において、R2UJ及びR1UJは、R2J及びR1Jと、それぞれ結合することができる又は結合しない。図63は、トランスポゾン末端が結合したMuトランスポソーム複合体を示す。いくつかの実施形態において、Tn5のトランスポゾン末端又はMos-1トランスポソームのトランスポゾン末端を、結合させることができる。

【0107】

本明細書で用いる場合、用語「トランスポゾン」は、in vitro転位反応において機能するトランスポザーゼ又はインテグラーゼ酵素と複合体を形成するのに必要なヌクレオチド配列（「トランスポゾン末端配列」）のみを示す、二本鎖DNAを意味する。トランスポゾンは、トランスポゾンを認識し、且つ結合するトランスポザーゼ又はインテグラーゼとともに、「複合体」又は「シナプス複合体」又は「トランスポソーム複合体」又は「トランスポソーム組成物」を形成する。複合体は、in vitro転位反応において一緒にインキュベートする標的DNAに、トランスポゾンを挿入又は転位することが可能である。トランスポゾンは、「転移トランスポゾン配列」すなわち「転移鎖」、及び「非転移トランスポゾン配列」すなわち「非転移鎖」からなる2つの相補的配列を示す。例えば、in vitro転位反応において活性がある機能亢進性Tn5トランスポザーゼ（例えば、EZ-Tn5（商標）トランスポザーゼ、米国ウィスコンシン州マディソン、EPICENTRE Biotechnologies社）とともに複合体を形成する1つのトランスポゾンは、以下の「転移トランスポゾン配列」を示す転移鎖：

5' AGATGTGTATAAGAGACAG 3'

及び以下の「非転移トランスポゾン配列」を示す非転移鎖：

5' CTGTCTCTTATACACATCT 3'

を含む。

【0108】

転移鎖の3'末端は、in vitro転位反応において、標的DNAに結合又は転移する。非転移鎖は、転移トランスポゾン末端配列に相補的なトランスポゾン配列を示し、in vitro転位反応において、標的DNAに結合又は転移しない。いくつかの実施形態において、トランスポゾン配列は、以下の配列うちの1つ又はそれ以上を含んでもよい：バーコード、アダプター配列、タグ配列、プライマー結合配列、捕捉配列、固有分子識別 (UMI: unique molecular identifier) 配列。

【0109】

本明細書で用いる場合、用語「アダプター」は、バーコード、プライマー結合配列、捕捉配列、捕捉配列に相補的な配列、固有分子識別 (UMI) 配列、親和性部分、制限部位

を含むことができる核酸配列を意味する。

【0110】

本明細書で用いる場合、用語「連続性情報」は、共有情報に基づいた、2つ又はそれ以上のDNAフラグメント間の空間的関係を指す。情報の共有態様は、隣接関係、区画関係、及び距離空間関係に関するものとして行うことができる。これらの関係に関する情報は、順に、DNAフラグメントに由来する配列リードの階層的なアセンブリ又はマッピングを容易にする。この連続性情報は、そのようなアセンブリ又はマッピングの効率及び精度を改善する。なぜなら、従来のショットガンシーケンシングと関連して用いられる伝統的なアセンブリ又はマッピング法では、個々の配列リードが由来した2つ又はそれ以上のDNAフラグメントの空間的関係に関して、個々の配列リードの相対的なゲノム起源又は座標を考慮に入れないからである。従って、本明細書に記載する実施形態によれば、連続性情報を捕捉する方法を、隣接空間関係を決定する短距離連続性法、区画空間関係を決定する中距離連続性法、又は距離空間関係を決定する長距離連続性法により行ってもよい。これらの方法は、DNA配列アセンブリ又はマッピングの精度及び質を高める。また、これらの方法は、上述のシーケンシング法等の任意のシーケンシング法とともに用いてもよい。

10

【0111】

連続性情報は、個々の配列リードが由来した2つ又はそれ以上のDNAフラグメントの空間的関係に関して、個々の配列リードの相対的なゲノム起源又は座標を含む。いくつかの実施形態において、連続性情報は、非重複配列リードからの配列情報を含む。

20

【0112】

いくつかの実施形態において、標的核酸配列の連続性情報は、ハプロタイプ情報を示す。いくつかの実施形態において、標的核酸配列の連続性情報は、ゲノム変異を示す。

【0113】

本明細書で用いる場合、用語「標的核酸の連続性の維持」は、核酸のフラグメント化との関連において、同じ標的核酸からのフラグメントの核酸配列の順番を維持することを意味する。

【0114】

本明細書で用いる場合、用語「少なくとも一部」及び／又はその文法的等価物は、全量のうちの任意の分量を指すことができる。例えば、「少なくとも一部」は、全量の少なくとも約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%、又は100%を指すことができる。

30

【0115】

本明細書で用いる場合、用語「約」は、±10%を意味する。

【0116】

本明細書で用いる場合、用語「シーケンシングリード」及び／又はその文法的等価物は、ポリマー中の単量体の順序を示すシグナルを得るために行われる物理的又は化学的ステップの繰り返し工程を指すことができる。シグナルは、単一の単量体解像度又はより低い解像度で単量体の順序を示すことができる。特定の実施形態において、ステップを、核酸標的に対して開始し、核酸標的中の塩基の順序を示すシグナルを得るために行うことができる。工程は、典型的な完了まで行うことができ、典型的な完了とは、通常、工程からのシグナルが、合理的なレベルの確実性で、標的の塩基をそれ以上区別することができない時点までと定義される。所望する場合、例えば、所望の配列情報量が得られるまで等、より早く完了することができる。シーケンシングリードは、単一の標的核酸分子に対して、又は同じ配列を有する標的核酸分子群に対して同時に、又は異なる配列を有する標的核酸群に対して同時に行うことができる。いくつかの実施形態において、シーケンシングリードは、シグナルが、シグナル取得が開始された1つ又はそれ以上の標的核酸分子からそれ以上は得られない時点で、終了する。例えば、シーケンシングリードは、固相基

40

50

質上に存在する1つ又はそれ以上の標的核酸分子に対して開始し、1つ又はそれ以上の標的核酸分子を基質から除去した時点で終了させることができる。或いは、シーケンシングは、シーケンシングランを開始した時に基質上に存在していた標的核酸の検出を中止することにより、終了することができる。例示的なシーケンシング方法は、その全体が参照により本明細書に組み込まれる、米国特許第9,029,103号明細書に記載されている。

【0117】

本明細書で用いる場合、用語「シーケンシング表示」及び／又はその文法的等価物は、ポリマー中の単量体単位の順序及び種類を示す情報を指すことができる。例えば、情報は、核酸中のヌクレオチドの順序及び種類を示すことができる。情報は、例えば、描写、画像、電子メディア、一連の記号、一連の数字、一連の文字、一連の色等を含む、任意の種々の形式であってよい。情報は、単一の単量体解像度又はより低い解像度であってもよい。例示的なポリマーは、ヌクレオチド単位を有するDNA又はRNA等の核酸である。一連の「A」、「T」、「G」、及び「C」文字は、単一のヌクレオチド解像度でDNA分子の実際の配列と相関性があるDNAに対する周知の配列表示である。その他の例示的なポリマーは、アミノ酸単位を有するタンパク質、及び糖単位を有する多糖類である。

【0118】

固体支持体

本明細書全体を通じて、固体支持体及び固体表面は、交換可能に用いられる。いくつかの実施形態において、固体支持体又はその表面は、管又は容器の内表面又は外表面等、非平面である。いくつかの実施形態において、固体支持体は、ミクロスフェア又はビーズを含む。本明細書において、「ミクロスフェア」又は「ビーズ」又は「粒子」又はその文法的等価物は、小さな分散粒子を意味する。適切なビーズ組成物としては、これに限定されないが、プラスチック、セラミック、ガラス、ポリスチレン、メチルスチレン、アクリルポリマー、常磁性物質、トリアゾル (thoriosaol)、カーボングラファイト、二酸化チタン、ラテックス、又は、セファロース等の架橋デキストラン、セルロース、ナイロン、架橋ミセル、及びテフロンが挙げられ、同様に、固形支持体として本明細書で概説する任意の他の材料を全て使用することができる。Bangs Laboratories社 (インディアナ州フィッシャーズ) の「ミクロスフェア検出ガイド (Microsphere Detection Guide)」は、役立つガイドである。特定の実施形態において、ミクロスフェアは、磁性ミクロスフェア又はビーズである。いくつかの実施形態において、ビーズは、色分けされていてもよい。例えば、Luminex社 (テキサス州オースティン) のMicroPlex (登録商標) ミクロスフェアを用いてもよい。

【0119】

ビーズは球状である必要はなく、不規則粒子を用いてもよい。代わりに又は加えて、ビーズは多孔質であってもよい。ビーズのサイズは、直径で、ナノメートル即ち約10nmから、ミリメートル即ち1mmに及び、約0.2ミクロン～約200ミクロンのビーズが好ましく、約0.5～約5ミクロンのビーズが特に好ましいが、いくつかの実施形態において、より小さい又はより大きいビーズを用いてもよい。いくつかの実施形態において、ビーズは、直径が約0.1µm、0.2µm、0.3µm、0.4µm、0.5µm、0.6µm、0.7µm、0.8µm、0.9µm、1µm、1.5µm、2µm、2.5µm、2.8µm、3µm、3.5µm、4µm、4.5µm、5µm、5.5µm、6µm、6.5µm、7µm、7.5µm、8µm、8.5µm、9µm、9.5µm、10µm、10.5µm、15µm、20µm、25µm、30µm、35µm、40µm、45µm、50µm、55µm、60µm、65µm、70µm、75µm、80µm、85µm、90µm、95µm、100µm、150µm、又は200µmであってもよい。

【0120】

トランスポソーム

「トランスポソーム」は、インテグラーゼ又はトランスポザーゼ等の組み込み（インテグレーション）酵素、及びトランスポザーゼ認識部位等の組み込み認識部位を含む核酸を含む。本明細書で提供する実施形態において、トランスポザーゼは、転位反応を触媒することが可能なトランスポザーゼ認識部位とともに機能的複合体を形成することができる。トランスポザーゼは、「タグメント化」と称することもある工程において、トランスポザーゼ認識部位に結合し、トランスポザーゼ認識部位を標的核酸に挿入する可能性がある。いくつかのそのような組み込み事象において、トランスポザーゼ認識部位の1つの鎖が、標的核酸に転移してもよい。1つの例において、トランスポソームは、2つのサブユニットを含む二量体トランスポザーゼ、及び2つの非連続トランスポゾン配列を含む。別の例において、トランスポソームは、2つのサブユニットを含む二量体トランスポザーゼを含むトランスポザーゼ、及び連続トランスポゾン配列を含む。

10

【0121】

いくつかの実施形態は、機能亢進性Tn5トランスポザーゼ及びTn5型トランスポザーゼ認識部位（Goryshin and Reznikoff, J. Biol. Chem., 273:7367 (1998)）、又は、MuAトランスポザーゼ、及びR1及びR2末端配列を含むMuトランスポザーゼ認識部位（Mizuuchi, K., Cell, 35:785, 1983; Savilahti, H, et al., EMBO J., 14:4893, 1995）の使用を含めることができる。機能亢進性Tn5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認識部位（例えば、EZ-Tn5（商標）トランスポザーゼ、ウィスコンシン州マディソン、Epicentre Biotechnologies社）は、以下の19塩基の転移鎖（時に、「M」又は「ME」）及び非転移鎖：それぞれ、5'AGATGTGTATAAGAGACAG3'、5'CTGTCTCTTATACACATCT3'を含む。ME配列はまた、当業者により最適化されて、使用されてもよい。

20

【0122】

本明細書で提供する組成物及び方法の特定の実施形態とともに使用することができる転位システムの更なる例としては、黄色ブドウ球菌（*Staphylococcus aureus*）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、及び国際公開第95/23875号）、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及びIS10（Kleckner N, et al., Curr Top Microbiol Immunol., 204:49-82, 1996）、マリナートランスポザーゼ（Lampe D J, et al., EMBO J., 15:5470-9, 1996）、Tc1（Plasterk R H, Curr. Topics Microbiol. Immunol., 204:125-43, 1996）、P元素（Glor, 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）、及び酵母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。更なる例としては、IS5、Tn10、Tn903、IS911、カタバミ（Sleeping Beauty）、SPIN、hAT、ピギーバック（PiggyBac）、ハーミス（Hermes）、Tcバスター（TcBuster）、Aeバスター1（AeBuster1）、Tol2、及び改変型トランスポザーゼファミリー

30

40

50

酵素 (Zhang et al., (2009) PLoS Genet. 5:e1000689. Epub 2009 Oct 16; Wilson C. et al (2007) J. Microbiol. Methods 71:332-5) が挙げられる。

【0123】

本明細書で提供する方法及び組成物とともに使用することができるインテグラーゼの更なる例としては、レトロウイルスインテグラーゼ、及びレトロウイルスインテグラーゼに対するインテグラーゼ認識配列が挙げられ、例えば、HIV-1、HIV-2、SIV、PFV-1、RSVのインテグラーゼが挙げられる。

【0124】

バーコード

一般に、バーコードは、1つ又はそれ以上の特定の核酸を同定するために使用することができる1つ又はそれ以上のヌクレオチド配列を含むことができる。バーコードは、人工配列であってもよく、或いは、転位の際に生成する自然発生配列、例えば、以前に並置されたDNAフラグメントの末端にある同一のフランキングゲノムDNA配列 (gコード) 等であってもよい。いくつかの実施形態において、バーコードは、標的核酸配列にはない人工配列であり、1つ又はそれ以上の標的核酸配列を同定するのに用いることができる。

【0125】

バーコードは、少なくとも約1個、2個、3個、4個、5個、6個、7個、8個、9個、10個、11個、12個、13個、14個、15個、16個、17個、18個、19個、20個、又はそれ以上の連続したヌクレオチドを含んでいてもよい。いくつかの実施形態において、バーコードは、少なくとも約10個、20個、30個、40個、50個、60個、70個、80個、90個、100個、又はそれ以上の連続したヌクレオチドを含む。いくつかの実施形態において、バーコードを含む核酸群におけるバーコードの少なくとも一部は、異なっている。いくつかの実施形態において、バーコードの少なくとも約10%、20%、30%、40%、50%、60%、70%、80%、90%、95%、99%が、異なっている。更なるそのような実施形態において、バーコードの全てが異なっている。バーコードを含む核酸群において、異なるバーコードの多様性は、ランダムに又は非ランダムに生成することができる。

【0126】

いくつかの実施形態において、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等、いくつかの実施形態において、第1のトランスポゾン配列は、第1のバーコードを含み、第2のトランスポゾン配列は、第2のバーコードを含む。いくつかの実施形態において、トランスポゾン配列は、第1のバーコード配列及び第2のバーコード配列を含むバーコードを含む。前述の実施形態のいくつかにおいて、第1のバーコード配列は、第2のバーコード配列とペアになるように同定又は設計することができる。例えば、互いにペアとなることが知られている複数の第1及び第2のバーコード配列を含む参照表を用いて、既知の第1のバーコード配列が既知の第2のバーコード配列とペアになることを知ることができる。

【0127】

別の例において、第1のバーコード配列は、第2のバーコード配列と同じ配列を含んでいてもよい。別の例において、第1のバーコード配列は、第2のバーコード配列の逆相補配列を含んでいてもよい。いくつかの実施形態において、第1のバーコード配列及び第2のバーコード配列は、異なる。第1及び第2のバーコード配列は、バイコード (bi-code) を含んでいてもよい。

【0128】

本明細書に記載する組成物及び方法のいくつかの実施形態において、バーコードは、鋳型核酸の調製に用いる。当然のことながら、膨大な数の利用可能なバーコードにより、各鋳型核酸分子は、固有の識別を含むことができる。鋳型核酸の混合物における各分子の固有の識別は、いくつかの用途に用いることができる。例えば、固有に識別された分子は、例えば、ハプロタイプシーケンシング、親対立遺伝子の識別、メタゲノムシーケンシ

10

20

30

40

50

ング、及びゲノムのサンプルシーケンシングにおいて、複数の染色体を有するサンプル、ゲノム、細胞、細胞型、細胞の病状、及び種における、個々の核酸分子の同定に応用することができる。

【0129】

例示的なバーコード配列としては、これに限定されないが、T A T A G C C T、A T A G A G G C、C C T A T C C T、G G C T C T G A、A G G C G A A G、T A A T C T T A、C A G G A C G T、及びG T A C T G A Cが挙げられる。

【0130】

プライマー部位

いくつかの実施形態において、トランスポゾン配列は、「シーケンシングアダプター」又は「シーケンシングアダプター部位」、言い換えれば、プライマーにハイブリダイズすることができる1つ又はそれ以上の部位を含む領域を含むことができる。いくつかの実施形態において、トランスポゾン配列は、増幅及びシーケンシング等に有用な少なくとも第1のプライマー部位を含むことができる。配列結合部位の例示的な配列としては、これに限定されないが、A A T G A T A C G G C G A C C A C C G A G A T C T A C A C (P5配列)及びC A A G C A G A A G A C G G C A T A C G A G A T (P7配列)が挙げられる。

【0131】

標的核酸

標的核酸としては、任意の目的の核酸を挙げることができる。標的核酸としては、DNA、RNA、ペプチド核酸、モルフォリノ核酸、ロックド核酸、グリコール核酸、トレオース核酸、核酸の混合サンプル、倍数性DNA（即ち、植物DNA）、これらの混合物、及びこれらのハイブリッドを挙げることができる。好ましい実施形態において、ゲノムDNA又はその増幅コピーを、標的核酸として用いる。別の好ましい実施形態において、cDNA、ミトコンドリアDNA、又は葉緑体DNAを用いる。いくつかの実施形態において、標的核酸は、mRNAである。

【0132】

いくつかの実施形態において、標的核酸は、単一の細胞由来、又は単一の細胞の画分由来である。いくつかの実施形態において、標的核酸は、単一の細胞小器官由来である。例示的な単一の細胞小器官としては、これに限定されないが、単一の核、単一のミトコンドリア、及び単一のリボソームが挙げられる。いくつかの実施形態において、標的核酸は、ホルマリン固定パラフィン包埋（FFPE）サンプルに由来する。いくつかの実施形態において、標的核酸は、架橋核酸である。いくつかの実施形態において、標的核酸は、タンパク質と架橋する。いくつかの実施形態において、標的核酸は、架橋DNAである。いくつかの実施形態において、標的核酸は、ヒストンに保護されたDNAである。いくつかの実施形態において、ヒストンは、標的核酸から除去される。いくつかの実施形態において、標的核酸は、ヌクレオソーム由来である。いくつかの実施形態において、標的核酸は、核タンパク質が除去されたヌクレオソーム由来である。

【0133】

標的核酸は、任意のヌクレオチド配列を含むことができる。いくつかの実施形態において、標的核酸は、ホモポリマー配列を含む。標的核酸はまた、繰り返し配列を含むことができる。繰り返し配列は、例えば、2ヌクレオチド、5ヌクレオチド、10ヌクレオチド、20ヌクレオチド、30ヌクレオチド、40ヌクレオチド、50ヌクレオチド、100ヌクレオチド、250ヌクレオチド、500ヌクレオチド、1000ヌクレオチド、又はそれ以上を含めた任意の種々の長さとすることができる。繰り返し配列は、連続又は非連続して、例えば、2回、3回、4回、5回、6回、7回、8回、9回、10回、15回、20回、又はそれ以上を含めた、任意の種々の回数の繰り返しとすることができる。

【0134】

本明細書に記載するいくつかの実施形態は、単一の標的核酸を使用することができる。他の実施形態は、複数の標的核酸を使用することができる。そのような実施形態において

10

20

30

40

50

、複数の標的核酸としては、複数の同じ標的核酸、いくつかの標的核酸は同じである複数の異なる標的核酸、又は全ての標的核酸が異なる複数の標的核酸を挙げることができる。複数の標的核酸を使用する実施形態は、例えば、1つ又はそれ以上のチャンバー又はアレイ表面で試薬が標的核酸に同時に送られるように、多重型式で行うことができる。いくつかの実施形態において、複数の標的核酸としては、実質的に全ての特定の生命体のゲノムを挙げることができる。複数の標的核酸としては、例えば、ゲノムの少なくとも約1%、5%、10%、25%、50%、75%、80%、85%、90%、95%、又は99%を含めた、特定の生命体のゲノムの少なくとも一部を挙げることができる。特定の実施形態において、当該一部は、ゲノムの最大で約1%、5%、10%、25%、50%、75%、80%、85%、90%、95%、又は99%である上限を有することができる。

10

【0135】

標的核酸は、あらゆる供給源から得ることができる。例えば、標的核酸は、単一の生命体から得られる核酸分子、又は1つ又はそれ以上の生命体を含む自然源から得られる核酸分子群から調製してもよい。核酸分子の供給源としては、これに限定されないが、細胞小器官、細胞、組織、臓器、又は生命体が挙げられる。標的核酸分子の供給源として用いられてもよい細胞は、原核細胞（バクテリア細胞、例えば、大腸菌（*Escherichia*）、バチルス（*Bacillus*）、セラチア（*Serratia*）、サルモネラ（*Salmonella*）、ブドウ球菌（*Staphylococcus*）、連鎖球菌（*Streptococcus*）、クロストリジウム（*Clostridium*）、クラミジア（*Chlamydia*）、ナイセリア（*Neisseria*）、トレポネマ（*Treponema*）、マイコプラズマ（*Mycoplasma*）、ボレリア（*Borrelia*）、レジオネラ（*Legionella*）、シュードモナス（*Pseudomonas*）、マイコバクテリウム（*Mycobacterium*）、ヘリコバクター（*Helicobacter*）、エルウィニア（*Erwinia*）、アグロバクテリウム（*Agrobacterium*）、根粒菌（*Rhizobium*）、及び *Streptomyces* genera）；クレン古細菌門（*crenarchaeota*）、ナノ古細菌門（*nanoearchaeota*）、又はユリ古細菌門（*euryarchaeotia*）等の古細菌細胞；又は真菌（例えば、酵母）、植物、原生動物や他の寄生生物、及び動物（昆虫（例えば、ショウジョウバエ（*Drosophila* spp.））、線形動物（例えば、線虫（*Caenorhabditis elegans*））、及び哺乳動物（例えば、ラット、マウス、サル、非ヒト霊長類、及びヒト）を含む）等の真核細胞であってもよい。標的核酸及び鋳型核酸は、当該技術分野でよく知られている種々の方法を用いて、目的とする特定の配列を濃縮することができる。そのような方法の例は、その全体が参照により本明細書に組み込まれる国際公開第2012/108864号において提供されている。いくつかの実施形態において、核酸は、鋳型ライブラリーの調製方法の間に、さらに濃縮させてもよい。例えば、核酸は、トランスポソームの挿入前、トランスポソームの挿入後、及び／又は核酸の増幅後に、特定の配列について濃縮させてもよい。

20

30

【0136】

また、いくつかの実施形態において、標的核酸及び／又は鋳型核酸は、高度に精製することができる。例えば、核酸は、本明細書で提供する方法で用いる前に、混入物を少なくとも約70%、80%、90%、95%、96%、97%、98%、99%、又は100%含まないものとするることができる。いくつかの実施形態において、当該技術分野で知られている標的核酸の質及びサイズを維持する方法を用いることは有益であり、例えば、標的DNAの単離及び／又は直接転位を、アガロースプラグを用いて行ってもよい。転位はまた、細胞群、ライセート、及び未精製DNAを用いて、細胞で直接行うことができる。

40

【0137】

いくつかの実施形態において、標的核酸は、生体サンプル又は患者サンプルから得てもよい。本明細書で用いる場合、用語「生体サンプル」又は「患者サンプル」は、組織及び体液等のサンプルを含む。「体液」としては、これに限定されないが、血液、血清、血漿、唾液、脳脊髄液、胸膜液、涙液、乳管液（*lactal duct fluid*）、リ

50

ンパ液、喀痰、尿、羊水、及び精液が挙げられる。サンプルは、「無細胞」である体液を含んでいてもよい。「無細胞体液」は、約1%（質量／質量）未満の全細胞物質を含む。血漿又は血清は、無細胞体液の例である。サンプルは、天然又は合成由来（即ち、無細胞となるように調製された細胞サンプル）の標本を含んでいてもよい。

【0138】

上記に開示した方法のいくつかの実施形態において、標的核酸は、標的核酸をトランスポソームに曝す前に、（例えば、超音波処理、制限消化、他の機械的手段により）フラグメント化することができる。

【0139】

本明細書で用いる場合、用語「血漿」は、血液で見られる無細胞液体を指す。「血漿」は、当技術分野で知られている方法（例えば、遠心分離及びろ過等）により、血液から全細胞物質を除去することにより、血液から得てもよい。

【0140】

特段の定めがない限り、用語「a」又は「an」は、本明細書全体を通じて、「1つ又はそれ以上」を意味する。

【0141】

用語「例えば（for example）」、「例えば（e.g.）」、「等（such as）」、「含む（include）」、「含めて（including）」、又はそれらの変形が本明細書で用いられる場合、これらの用語は、限定の用語であるとは見なされないものであり、「これに限定されるものではないが」又は「限定されずに」を意味すると解釈されるものである。

【0142】

以下の実施例は、説明のための実施例を提供するものであり、決して本明細書で提供する発明を制限するものではない。

【実施例】

【0143】

実施例1 ビーズベースタグメント化工程からのDNAクラスターの収量

図3のビーズベースタグメント化工程からのDNAクラスターの収量を評価し、図4の表に示した。本実施例において、50 ng、250 ng、1000 ngのヒトNA12878 DNAを、同じバッチのタグメント化ビーズ（2.8 µmビーズ）を用いてタグメント化した。第2の50 ng分量のNA12878 DNAを、第2のバッチのタグメント化ビーズ（完全リピート：2.8 µmビーズ）を用いてタグメント化した。ビーズ結合タグメント化DNAサンプルをPCR増幅し、精製した。一定分量（5.4 µL）の各精製PCR産物（未定量）を270倍希釈し、約50 pMのストックサンプル溶液を作製した。各サンプルに対し、50 pMストック溶液を15 pM、19 pM、21 pM、及び24 pMに希釈した。希釈したサンプルを、クラスター生成及びシーケンシングのためにフローセルにロードした。データによれば、同じ希釈液（～50 pM）から始めて、クラスター数は、同じセットのビーズを用いた3つの異なるインプットレベル（即ち、50 ng、250 ng、1000 ng）に対して、100～114%の間であることが分かる。50 ng完全リピートでのクラスター数（異なるバッチのビーズを用いた）は、81%であった。異なる希釈液（15 pM、19 pM、21 pM、及び24 pM）は、約10%以内の同数のクラスターを産生する。データによれば、ビーズが収量を大きく制御し、収量は、異なるDNAインプット及び異なるリピートで再現性があることが分かる。

【0144】

実施例2 ビーズベースタグメント化工程の再現性

図3のビーズベースタグメント化工程の再現性を図5に示す。本実施例において、「同じ」トランスポソーム密度で作製した6種類の異なるインデックス付きビーズ（インデックス1～6；2.8 µmビーズ）調製物を、50 ng及び500 ngのインプットNA12878 DNAを用いたタグメント化DNAの調製に用いた。タグメント化DNAをPCR増幅し、精製した。2つのHiSeqレーン用に、12種類の精製PCR産物をプー

10

20

30

40

50

ルして、6種類ずつの2つの混合物（プール1及びプール2）にした。各プールは、1レーン当たり3～50 ng及び3～500 ngのサンプルを含む。データ表500は、各インデックス付きサンプルの挿入サイズ中央値及び挿入サイズ平均値を示す。

【0145】

実施例3 プール1の挿入サイズ及びプール2の挿入サイズ

図5のインデックス付きサンプルのプール1の挿入サイズ及びプール2の挿入サイズを、それぞれ図6A（プロット600）及び図6B（プロット650）に示す。データはまた、挿入サイズが、6種類の異なるインデックス付きビーズ調製物の間で均一であることを示す。ビーズベースタグメント化は、挿入サイズ及びDNA収量を制御するメカニズムをもたらす。

【0146】

実施例4 リードの合計数の再現性

図5に記載の実験に関して、リードの合計数及びアラインされたリードの割合の再現性を図7（棒グラフ700）に示す。両方のインプット（50 ng及び500 ng）で、リードの合計数は、同じインデックス付きビーズ調製物に対して同様である。6種類のインデックス付きビーズ調製物のうちの4種類（インデックス1、2、3、及び6）で、極めて近い収量を示し、インデックス付きビーズ調製物4及び5では、インデックス配列によるものである可能性がある、幾分のばらつきが見られた。

【0147】

1つの応用において、ビーズベースタグメント化工程は、タグメント化ステップを含むエクソーム濃縮アッセイ、例えば、Illumina社のNextera（登録商標）急速捕捉濃縮プロトコルに用いてもよい。現在のエクソーム濃縮アッセイ（即ち、Illumina社のNextera（登録商標）急速捕捉濃縮プロトコル）では、溶液ベースのタグメント化（Nextera）をゲノムDNAのフラグメント化に用いる。その後、遺伝子特異的プライマーを用いて、目的とする特異的遺伝子フラグメントをプルダウンする。2回の濃縮サイクルを行った後、プルダウンしたフラグメントをPCR及びシーケンシングにより濃縮する。

【0148】

エクソーム濃縮アッセイにおけるビーズベースタグメント化工程の使用を評価するため、ヒトNA12878 DNAを、25 ng、50 ng、100 ng、150 ng、200 ng、及び500 ngのインプットDNAを用いてタグメント化した。コントロールライブラリー（NA00536）を、標準プロトコルに従って、50 ngのインプットDNAから調製した。各DNAインプットは、異なるインデックス（固有識別子）を有した。標準方法に合わせるため、且つ、十分な量のフラグメントがプルダウン用に存在することを確実にするために、濃縮ポリメラーゼマスターミックス（EPM：enhanced polymerase master mix）を用いた10サイクルのPCRを行った。増幅プロトコルは、72℃で3分、98℃で30秒、続いて98℃で10秒を10サイクル、65℃で30秒、及び72℃で1分とした。その後、サンプルを10℃で保持した。次に、サンプルを、エクソーム濃縮プルダウン工程及びシーケンシングを通じて処理した。

【0149】

実施例5 エクソーム濃縮アッセイにおけるコントロールライブラリー及びビーズベースタグメント化ライブラリーの挿入サイズ

図8A、8B、及び8Cは、それぞれ、エクソーム濃縮アッセイにおける、コントロールライブラリーでの挿入サイズのプロット800、ビーズベースタグメント化ライブラリーでの挿入サイズのプロット820、及びサマリーデータ表840を示す。データによれば、ビーズベースタグメント化ライブラリーは、コントロールライブラリーに比べて、広い挿入サイズ分布を有するが、挿入サイズは、サンプルのDNAインプットに関係なく極めて近いことが分かる。

【0150】

10

20

30

40

50

実施例6 リード配列の質

図9A、9B、及び9Cは、それぞれ、図8A、8B、及び8Cのエクソーム濃縮アッセイにおける、フィルターを通過した複製物(dups PF:duplicates passing filters)の割合の棒グラフ900、PCT selected basesの棒グラフ920、及びPCT usable bases on targetの棒グラフ940を示す。図9Aを参照すると、dups PFの割合(%)は、いくつかのリードがフローセルの他の部分で複製されているかを示す尺度である。この数値は、全てのクラスターが結果に対して有益なデータをもたらすことを確実にするためには、理想的には、(ここで示すように)低くなるものである。

【0151】

図9Bは、PCT selected basesを示し、濃縮工程の間に濃縮されていたはずの目的の部位に又はその近くに配列するリードの割合の尺度である。理想的には、この数値は、濃縮工程の成功を反映して1に近くなるものである。また、この数値は、濃縮されるべきではないリードが工程を終えていないことを示す。

【0152】

図9Cは、PCT usable bases on targetを示し、濃縮された領域内で目的とする特定の塩基上に実際に配列しているリードの割合の尺度である。理想的には、全ての濃縮リードが濃縮されたリード内の目的とする塩基上に配列するものであるが、タグメント化のランダム性及び様々な挿入長さのために、目的とする領域上で配列し終えていないリードが濃縮される可能性がある。

【0153】

2つの技術を使用して挿入サイズ分布を最適化してもよい。一例を挙げれば、SPRIクリーンアップを用いて小さ過ぎる又は大き過ぎるフラグメントを除去してもよい。SPRIクリーンアップは、サイズ及び所望の沈澱又は非沈澱DNAの保持(即ち、第1ステップは、所望のサイズよりも大きいDNAのみを沈澱させ、可溶性小さいフラグメントを保持する)に基づいた選択的DNA沈澱により、所望のサイズよりも大きい又は小さいフラグメントを除去する工程である。その後、小さいフラグメントをさらに沈澱させ、この時、望まない極めて小さいフラグメント(まだ溶液中にある)を除去し、沈澱したDNAを保持し、洗浄した後、再可溶化して所望のサイズ範囲のDNAを得る。別の例を挙げれば、ビーズ表面上の活性化トランスポソームのスペーシングを用いて、挿入サイズ分布を制御してもよい。例えば、ビーズ表面上のギャップを不活性トランスポソーム(例えば、不活性トランスポゾンを含むトランスポソーム)で充填してもよい。

【0154】

ビーズベースのタグメント化工程の連続性を評価した。表3は、インデックスを共有する1000bpウィンドウ内で0回、1回、2回、又は3回のリードが起きた回数を示す。ビーズを9種類の異なるインデックス付きトランスポソームで生成し、少量のヒトDNAのタグメント化に用いた。リードを生成させ、アラインし、同じインデックスを共有する1000bp又は10Kbウィンドウ内のリードの数について分析した。インデックスを共有する小さいウィンドウ内のリードの中には、偶然生成するものがあるが、これが何回起こる可能性があるかという予測を、表3及び表4の「ランダム」列に示す。「ビーズ」列の数は、インデックスを共有する1000bp(表3)又は10Kp(表4)ウィンドウの実際の数値を示す。表3及び表4に示すように、同じインデックスが1000bp又は10Kpウィンドウ内で見つかった実際の回数は、ランダムケースでの予測よりも顕著に多い。「0」枠は、特定の1000bpウィンドウがそれにマッピングするインデックスリードを有さない全ての回数値を示す。数値は、極少量のヒトゲノムのみがシーケンシングされ、大半のウィンドウがそれらにアラインされるリードを有さないため、ここでは最も大きい。「1」は、ただ1つのリードが1000bp(又は10Kp)ウィンドウにマップする回数であり、「2」は、1000bp(又は10Kp)ウィンドウ内で2つのリードがインデックスを共有する回数である、等である。このデータは、1400超のケースにおいて、同じピースのDNA(10Kp超)が、約15000回のタグメン

10

20

30

40

50

ト化事象の中で、同じビーズにより、少なくとも2回から5回までタグメント化されていることを示唆している。フラグメントは、インデックスを共有しているため、それらが偶然にそこに存在する可能性は低く、同じビーズに由来している。

【0155】

【表3】

表3 インデックスを共有する1000bpウィンドウにおけるリード数				
	0	1	2	3
ビーズ	25913666	15220	305	7
ランダム	25913334	15855	9	0

10

【0156】

表4は、インデックスを共有する10Kpウィンドウ内のリード数（5個まで）を示す。

【0157】

【表4】

表4 インデックスを共有する10Kpウィンドウにおけるリード数						
	0	1	2	3	4	5
ビーズ	2578669	12683	1267	169	28	3
ランダム	2577012	15742	64	1	0	0

20

【0158】

実施例7 CPT-DNAからの遊離トランスポソームの分離

転位の後、CPT-DNA及び遊離トランスポソームを含む反応混合物を、Sepha-cryl S-400及びSepha-cryl S-200サイズ排除クロマトグラフィーを用いたカラムクロマトグラフィーにかけた。図22に示す。CPT-DNAは、NC P DNAと表示する。

【0159】

実施例8 ビーズ上の捕捉プローブ密度の最適化

捕捉プローブA7及びB7の密度を、1µmビーズ上で最適化し、結果を図25に示した。レーン1（A7）及びレーン3（B7）は、高いプローブ密度を有し、レーン2（A7）及びレーン4（B7）は、1µmビーズ当たり推定10,000～100,000のプローブ密度を有した。標的分子に対する捕捉プローブのライゲーション産物を、アガロースゲルで評価した。ビーズ当たり約10,000～100,000のプローブ密度は、より高いプローブ密度よりも良好なライゲーション効率を有した。

【0160】

実施例9 分子内ハイブリダイゼーションによるビーズ上でのCPT-DNAのインデックス付きシーケンシングライブラリーの調製の実現性試験

トランスポソームを、ビーズ上のA7及びB7捕捉配列に相補的なA7'及びB7'捕捉配列を有するトランスポゾンと機能亢進性Tn5トランスポザゼとを混合することにより、調製した。高分子量ゲノムDNAをトランスポソームと混合し、CPT-DNAを生成する。それとは別に、ビーズを、固定化オリゴヌクレオチド：P5-A7、P7-B7、又はP5-A7+P7-B7で調製する。ここで、P5及びP7は、プライマー結合配列であり、A7及びB7は、それぞれA7'及びB7'配列に相補的な捕捉配列である。P5-A7単独、P7-B7単独、P5-A7+P7-B7、又はP5-A7ビーズ及びP7-B7ビーズの混合物を含むビーズを、CPT-DNAで処理し、反応混合物にリガーゼを添加して、固定化オリゴの転位DNAに対するハイブリダイゼーションの効率を決定した。結果を図26に示す。シーケンシングライブラリーは、アガロースゲル上で

30

40

50

高分子量バンドにより示されるように、P 5 - A 7 及び P 7 - B 7 がともに 1 つのビーズ上に固定化されている場合（レーン 4）のみで作製される。結果は、高効率の分子内ハイブリダイゼーションを示し、また、分子内ハイブリダイゼーションによるビーズ上での C P T - D N A のインデックス付きシーケンシングライブラリーの調製の実現性を証明した。

【 0 1 6 1 】

実施例 1 0 クローンインデックス化の実現性試験

いくつかのトランスポソームセットを調製した。1 つのセットにおいて、機能亢進性 T n 5 トランスポザーゼを、5 ' ビオチンを有するトランスポゾン配列 T n p 1 と混合し、トランスポソーム 1 を調製する。別のセットにおいて、5 ' ビオチンを有する固有インデックス 2 を有する T n p 2 で、トランスポソーム 2 を調製する。別のセットにおいて、トランスポソーム 3 の調製のため、機能亢進性 T n 5 トランスポザーゼを、5 ' ビオチンを有するトランスポゾン配列 T n p 3 と混合する。別のセットにおいて、固有インデックス 4 及び 5 ' ビオチンを有する T n p 4 でトランスポソーム 4 を調製する。トランスポソーム 1 及び 2、並びにトランスポソーム 3 及び 4 を、それぞれ別々にストレプトアビジンビーズと混合し、ビーズセット 1 及びビーズセット 2 を生成する。次に、2 つのセットのビーズを混合し、ゲノム D N A 及びタグメント化バッファーとともにインキュベートして、ゲノム D N A のタグメント化を促進する。この後、タグメント化配列の P C R 増幅を行う。増幅した D N A をシーケンシングし、インデックス配列の挿入を分析する。タグメント化がビーズに限定される場合、大多数のフラグメントは、T n p 1 / T n p 2 及び T n p 3 / T n p 4 インデックスでコードされることになる。分子内ハイブリダイゼーションがある場合には、フラグメントは、T n p 1 / T n p 4、T n p 2 / T n p 3、T n p 1 / T n p 3、及び T n p 2 / T n p 4 インデックスでコードされることになる。5 サイクル及び 1 0 サイクルの P C R 後のシーケンシング結果を図 2 7 に示した。コントロールは、混合され、ビーズ上に固定化された、4 種類全てのトランスポゾンに有する。結果は、大多数の配列が T n p 1 / T n p 2 又は T n p 3 / T n p 4 インデックスを有することを示し、クローンインデックス化が実現可能であることを示している。コントロールは、インデックスを区別しないことを示す。

【 0 1 6 2 】

実施例 1 1 単一反応におけるインデックス付きクローンビーズ転位

9 6 種類のインデックス付きトランスポソームビーズを調製する。個々のインデックス付きトランスポソームは、5 ' 末端の T n 5 モザイク末端配列（M E）及びインデックス配列を有するオリゴヌクレオチドを含むトランスポゾンを混合して、調製した。個々のインデックス付きトランスポソームを、ストレプトアビジン-ビオチン相互作用によりビーズ上に固定化した。ビーズ上のトランスポソームを洗浄し、ビーズ上の 9 6 種類全ての個々にインデックス化されたトランスポソームをプールした。M E 配列に相補的でインデックス配列を有するオリゴヌクレオチドを、固定化オリゴヌクレオチドにアニールし、固有のインデックスを有するトランスポゾンを作製した。9 6 種類のクローンインデックス付きトランスポソームビーズセットを混ぜ合わせ、高分子量（H M W : h i g h m o l e c u l a r w e i g h t）ゲノム D N A とともに、N e x t e r a タグメント化バッファーの存在下、単一のチューブでインキュベートした。

【 0 1 6 3 】

ビーズを洗浄し、反応混合物を 0 . 1 % S D S で処理することにより、トランスポザーゼを除去する。タグメント化 D N A をインデックス付きプライマーで増幅し、P E H i S e q フローセル v 2 で、T r u e S e q v 3 クラスターキットを用いてシーケンシングし、シーケンシングデータを分析する。

【 0 1 6 4 】

リードのクラスター、すなわち島を観察する。各配列に対するリード間の最近隣距離のプロットは、主要なピーク、1 つはクラスター内からのもの（近位）ともう 1 つはクラスター間からのもの（遠位）、を基本的に示す。方法及び結果の模式図を、図 3 0 及び 3 1

に示す。島のサイズは、約3～10 kbの範囲である。カバーされた塩基の割合は、約5%～10%である。ゲノムDNAの挿入サイズは、約200～300塩基である。

【0165】

実施例12 ビーズ上のトランスポソームに対するライブラリーサイズ

初めに、ME'配列を有する第1のオリゴヌクレオチド、ME-バーコード-P5/P7配列を有する第2のオリゴヌクレオチド、及びTn5トランスポザーゼを混合することにより、トランスポソームを溶液中にアセンブルした。第1のセットにおいて、ME'配列を有する第1のオリゴヌクレオチドを、3'末端でビオチン化する。第2のケースにおいて、ME-バーコード-P5/P7配列を有するオリゴヌクレオチドを、5'末端でビオチン化する。種々の濃度(10 nM、50 nM、及び200 nM)の得られた各トランスポソームセットに対し、ストレプトアビジンビーズを添加して、トランスポソームがストレプトアビジンビーズに固定化されるようにする。ビーズを洗浄し、HMWゲノムDNAを加え、タグメント化を行う。いくつかのケースでは、タグメント化DNAを0.1% SDSで処理し、他のケースでは、タグメント化DNAを処理しない。タグメント化DNAを5～8サイクルでPCR増幅し、シーケンシングする。模式図を図32に示す。

【0166】

図33に示すように、SDS処理により増幅効率及びシーケンシングの質が改善される。3'ビオチンを有するオリゴヌクレオチドは、トランスポソームに対してより良いライブラリーサイズを有する。

【0167】

図34は、挿入サイズに対するトランスポソーム表面密度の影響を示す。5'ビオチンを有するトランスポソームは、より小さいサイズのライブラリー及びより多くの自己挿入副産物を示す。

【0168】

実施例13 インプットDNAの滴定

種々の量の標的HMW DNAを、50 mMのTn5：トランスポソン密度を有するクロンインデックス付きビーズに加え、37℃で15分間若しくは60分間、又は室温で60分間インキュベートした。トランスポソームは、3'ビオチンを有するオリゴヌクレオチドを含んだ。タグメント化を行い、反応混合物を0.1% SDSで処理し、PCR増幅させた。増幅したDNAをシーケンシングした。図35は、サイズ分布に対するインプットDNAの影響を示す。10 pgのインプットDNAによる反応は、最小シグナルを示した。サイズ分布パターンは、20 pg、40 pg、及び200 pgのDNAインプットで同様であった。

【0169】

実施例14 溶液ベース及びビーズベースの方法を用いた島のサイズ及び分布

溶液ベース及びビーズベースの方法を用いて、島のサイズ及び分布を比較した。溶液ベースアプローチにおいて、トランスポソンにそれぞれ固有のインデックスを有する96種類のトランスポソームを、96穴プレートにアセンブルする。HMWゲノムDNAを添加し、タグメント化反応を行う。反応生成物を0.1% SDSで処理し、PCR増幅させる。増幅したDNAをシーケンシングした。

【0170】

ビーズベースアプローチにおいて、トランスポソンにそれぞれ固有インデックスを有する96種類のトランスポソームを、96穴プレートにアセンブルした。オリゴヌクレオチドは、3'末端ビオチンを含んだ。ストレプトアビジンビーズを96穴プレートの各々に添加し、トランスポソームがストレプトアビジンビーズに固定化されるようにインキュベートする。ビーズをそれぞれ洗浄してプールし、HMWゲノムDNAを添加し、タグメント化反応を単一反応容器(ワンポット)内で行う。反応生成物を0.1% SDSで処理し、PCR増幅させる。増幅産物をシーケンシングした。

【0171】

ネガティブコントロールでは、最初に、それぞれ固有のインデックスを有する96種類

10

20

30

40

50

全てのトランスポゾン配列を混合する。オリゴヌクレオチドは、3'末端ビオチンを含んだ。トランスポソームを個々の混合インデックス付きトランスポゾンから調製する。ストレプトアビジンビーズを混合物に添加する。HMWゲノムDNAを添加し、タグメント化反応を行う。反応生成物を0.1% SDSで処理し、PCR増幅させる。増幅産物をシーケンシングした。

【0172】

島内リードの数を島のサイズに対してプロットする。図36に示す結果は、島（近位リード）が、溶液ベースの方法と同様に、ワンポットクローンインデックス付きビーズで観られることを示している。インデックス付きトランスポソームをトランスポソーム形成前に混合した場合、島（近位リード）は観られなかった。トランスポソーム形成前にトランス

10

【0173】

実施例15 CPT-seqを用いた構造変異の分析

60 kb ヘテロ接合欠損の検出

シーケンシングデータをfastqファイルとして抽出し、分離（デマルチプレックス）工程を行って、各バーコードに対する個々のfastqファイルを生成する。CPTシーケンシングからのfastqファイルを、インデックスに従って分離し、重複を除去した参照ゲノムにアラインする。スキャンウィンドウ内の任意のリードを示すインデックスの数を記録する、5 kb / 1 kb ウィンドウにより、染色体をスキャンする。統計的には、ヘテロ接合欠損領域のため、隣接領域と比較して、半量のDNAしかライブラリー生成に利用できない。従って、インデックスの数も隣接領域の約半分となるべきである。NA12878 chr1の60 kb ヘテロ接合欠損を、9216個のインデックス付きCPTシーケンシングデータから5 kb ウィンドウにスキャンすることにより、図47 A及び47 Bに示す。

20

【0174】

遺伝子融合の検出

CPTシーケンシングからのfastqファイルを、インデックスに従って分離し、重複を除去した参照ゲノムにアラインする。染色体を2 kb ウィンドウにスキャンする。各2 kb ウィンドウは、36864ベクターであり、固有インデックスからのリードがこの2 kb ウィンドウでいくつか見つかったかを各エレメントが記録する。ゲノムにわたり、2 kb ウィンドウペア（X、Y）毎に、重み付けジャカード（weighted-Jaccard）インデックスを計算する。このインデックスは、事実上、サンプルの（X、Y）間の距離を示す。これらのインデックスを、図48に示すヒートマップとして表示する。各データポイントは、2 kb スキャンウィンドウのペアを表わし、左上の四角は、ともに領域1からのX、Yであり、右下は、ともに領域2からのX、Yであり、右上は、領域1から領域2にわたる領域からのX、Yである。遺伝子融合シグナルは、このケースでは中央の横線として示される。

30

【0175】

欠損の検出

CPTシーケンシングからのfastqファイルを、インデックスに従って分離し、重複を除去した参照ゲノムにアラインする。染色体を1 kb ウィンドウにスキャンする。図49は、遺伝子欠損の検出結果を示す。

40

【0176】

実施例16 フェージング及びメチル化検出

亜硫酸水素塩変換効率の最適化

ビーズ上のインデックス付き結合CPT-seqライブラリーに対して、ME（モザイクエレメント領域）及びgDNA領域で、変換を評価した。Promega社のMethylEdge亜硫酸水素塩変換システムを最適化して、効率を改善した。

【0177】

50

【表 5】

条件	DNA	ビーズ	B S C 処理
1	10ng	無し	6 0℃で1時間/ 0.3M NaOH
2		有り	6 0℃で1時間/ 0.3M NaOH
3			6 0℃で1時間/ 1M NaOH
4			6 5℃で1時間/ 0.3M NaOH

10

【0178】

ME配列を分析して、亜硫酸水素塩変換処理の効率を決定した。図50に示す。ビーズに付着したインデックス付き結合ライブラリーのうち、95%が亜硫酸水素塩変換（BSC）した。亜硫酸水素塩条件間で、同様のPCR収量が観察された。>より厳しい亜硫酸水素塩処理でも、ライブラリーを分解しないように思われた。図51に示す。ビーズ上のインデックス付き結合ライブラリーの約95%でBSCが観察された。BSCを改良（C>U）するために研究した変数は、温度及びNaOH濃度（変性）であった。60℃及び1MのNaOH、又は℃及び0.3MのNaOHで良い結果となった。

20

【0179】

ビーズライブラリーでのBSC変換CPT-seqのシーケンシング後、期待したシーケンシングリード構造が観察された。塩基の割合（%）の計測値を、IVCプロットとともに図52に示す。

【0180】

図53は、亜硫酸水素塩変換後のPCRの後の、インデックス付き結合ライブラリーのアガロースゲル電気泳動の画像を示す。200～500bpの期待したサイズ範囲のライブラリーが観察された。DNA無しでの反応では、インデックス付き結合ライブラリーを産生しない。

30

【0181】

実施例17 標的フェージング

全ゲノムインデックス付き結合CPT-seqライブラリーを濃縮した。図54は、サイズ選択をしない濃縮前の全ゲノムインデックス付き結合CPT-seqライブラリーのバイオアナライザートレースを示す。図55は、濃縮後のライブラリーのアガロースゲル分析を示す。

【0182】

HLA領域に対する濃縮統計を以下に示す。

40

【0183】

50

【表 6】

サンプル ID :	C3
サンプル名 :	H L A プローブ
パッドサイズ :	150
標的参照 (targeted reference) の全長 :	5062748
P F リードの合計 :	2516
Q 3 0 の割合 (%) :	94.90%
アラインされたリードの合計 :	2498
アラインされたリードの割合 (%) :	99.40%
標的のアラインされたリード :	840
リードの濃縮 :	30.80%
重複したリードペアの割合 (%) :	12.70%
フラグメント長の中央値 :	195

【 0 1 8 4 】

図 5 6 は、染色体の H L A 領域に対して標的化ハプロタイピングを適用した結果を示す。全ゲノムインデックス付き結合リードライブラリーの濃縮の図解を左に示す。小さいバーは、各々インデックス付き短ライブラリーを示す。インデックス付きライブラリーのクラスターは、単一ビーズ上で同じインデックスでクローンインデックス付けされた領域である「島」であり、従って、ゲノムスケール上でリードの近位性（「島」性）を示す。標的領域におけるライブラリーの濃縮（国際公開第 2 0 1 2 / 1 0 8 8 6 4 号「核酸の選択的濃縮（Selective enrichment of nucleic acids）」を参照）を右に示す。リードは、H L A 領域で濃縮される。さらに、リードがインデックスで選別され、ゲノムにアラインされる場合、リードは、インデックス付き結合リードから連続性情報が維持されていることを示す「島」構造を再び示す。

【 0 1 8 5 】

実施例 1 8 インデックス交換

トランスポソーム複合体のモザイク末端（ME）の交換を評価するため、異なるインデックスを有するビーズを調製した。混合後、ライブラリーをシーケンシングし、各ライブラリーのインデックスをレポートすることにより、インデックス交換を決定した。「交換（swapped）」の割合（%）を、 $(D4 + D5 + E3 + E5 + f4) / (全96種の合計)$ で計算した。図 6 5 に示す。

【 0 1 8 6 】

実施例 1 9 トランスポソーム複合体をストレプトアビジンビーズにより密集させること

によるライブラリー挿入サイズの縮小

ストレプトアビジン磁性ビーズを、1 倍、6 倍、及び 12 倍濃度の T s T n 5 トランスポソーム複合体とともにロードした。各ビーズタイプに対して、E p i - C P T S e q プロトコルを実施した。分析のため、最終 P C R 産物を A g i l e n t B i o A n a l y z e r 上にロードした。図に示す。E p i - C P T S e q ライブラリーフラグメントは、比較的小さく、ビーズ上に T s T n 5 を多くロードするほど多く産生した。

【0187】

実施例 20 亜硫酸水素塩変換中の DNA ライブラリーのフラグメント化

亜硫酸水素塩変換後、DNA は損傷を受け、結果として P C R 増幅に必要な共通配列 (C S 2) が減少する。DNA フラグメント C P T S e q 及び E p i - C P T S e q (M e - C P T S e q) ライブラリーを、B i o A n a l y z e r で分析した。亜硫酸水素塩変換中の DNA 損傷により、E p i - C P T S e q ライブラリーは、図 70 に示すように、C P T S e q ライブラリーと比較して、5 倍低い収量であり、小さいライブラリーサイズ分布を有する。

【0188】

実施例 21 T d T 媒介 s s DNA ライゲーション反応

ターミナルトランスフェラーゼ (T d T) 媒介ライゲーションによる DNA 末端修復の実現性を試験した。簡単には、5 p m o l の s s DNA 鋳型を、T d T (10 / 50 U)、アテニューエーター／アダプター二本鎖 (0 / 15 / 25 p m o l)、及び DNA リガーゼ (0 / 10 U) とともに、37 °C で 15 分間インキュベートした。伸長／ライゲーションの DNA 産物を、T B E - U r e a ゲルで分析し、結果を図 71 に示した。全ての反応成分を添加した結果、アダプター分子のほぼ完全なライゲーションが行われた (レーン 5 ~ 8)。

【0189】

ターミナルトランスフェラーゼ (T d T) 媒介ライゲーションによる DNA 末端修復の実現性を、亜硫酸水素ナトリウム変換ビーズ結合ライブラリーに対して試験した。図 72 に示す。簡単には、DNA をビーズ上でタグメント化し (最初の 2 レーン)、P r o m e g a 社の M e t h y l E d g e 亜硫酸水素塩変換キットで処理し (レーン 3 及び 4)、DNA 救出プロトコルを行った (レーン 5 及び 6)。救出反応後の DNA ライブラリーの収量及びサイズは、明らかに増加している。また、挿入トランスポゾン (S I) の存在量も増加しており、アダプター分子の効率的なライゲーションを示している。

【0190】

メチル - C P T S e q アッセイの結果を図 73 に示す。

10

20

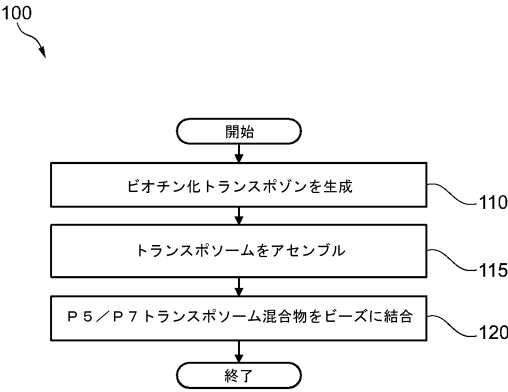
30

40

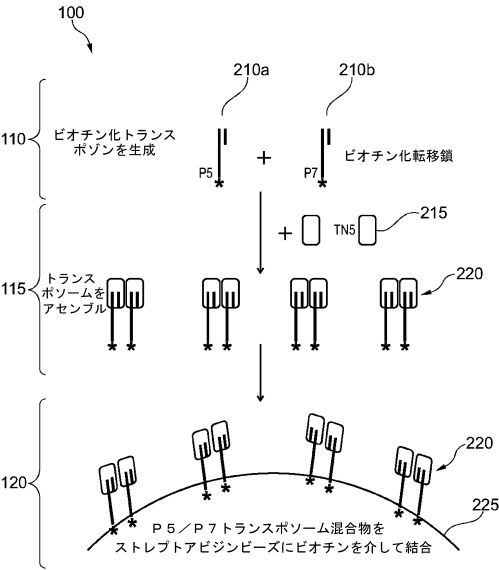
50

【図面】

【図 1】



【図 2】



10

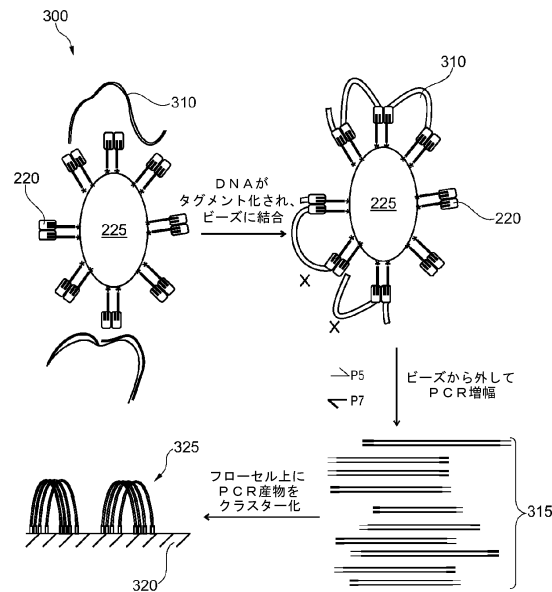
20

30

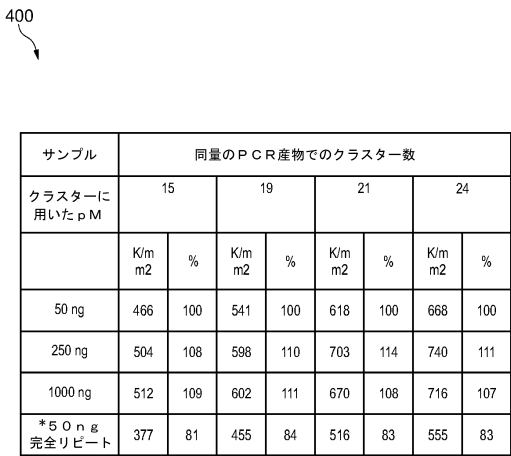
40

50

【図 3】



【図 4】



10

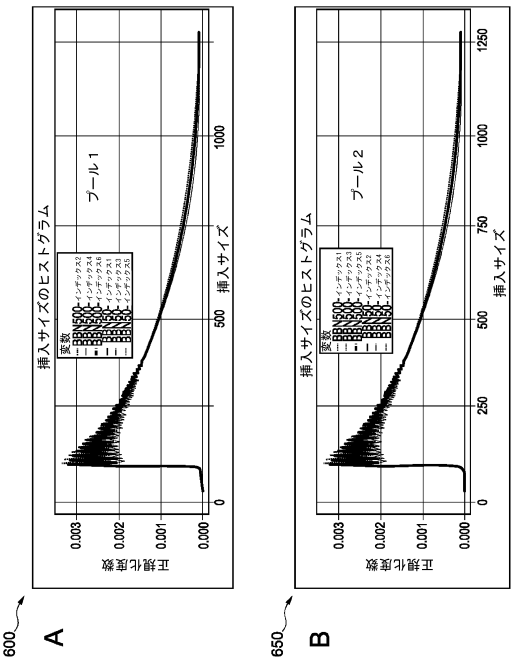
20

【図 5】

500

	挿入サイズ中央値	挿入サイズ平均値
BBN500-インデックス1	341	409.78
BBN500-インデックス2	344	415.89
BBN500-インデックス3	337	405.97
BBN500-インデックス4	346	419.91
BBN500-インデックス5	330	400.35
BBN500-インデックス6	327	398.65
BBN500-インデックス7	321	398.71
BBN500-インデックス8	328	400.75
BBN500-インデックス9	327	397.62
BBN500-インデックス10	328	399.88
BBN500-インデックス11	324	399.35
BBN500-インデックス12	331	401.50

【図 6】

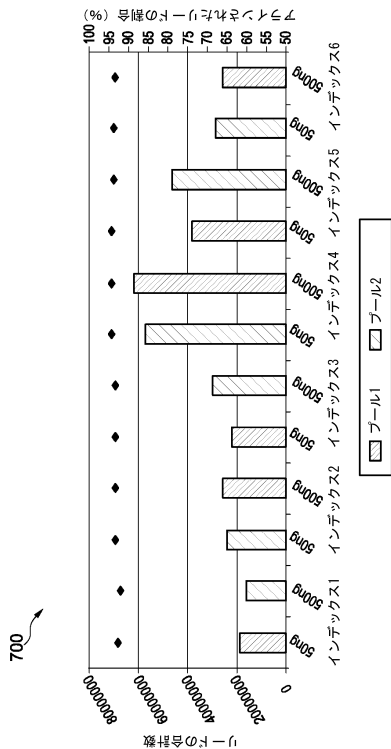


30

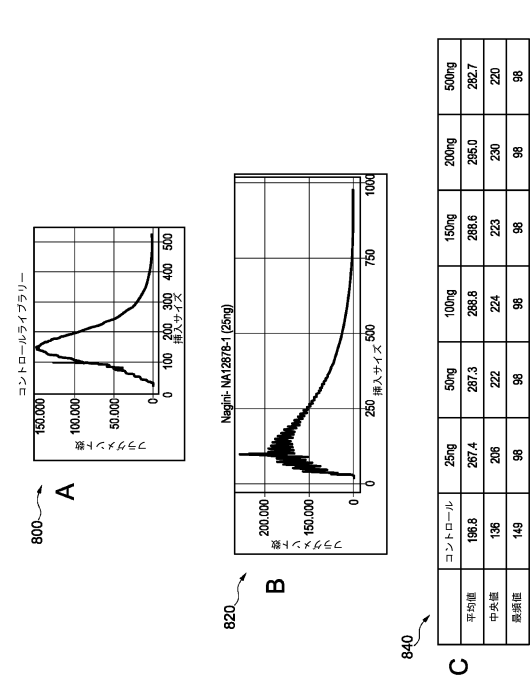
40

50

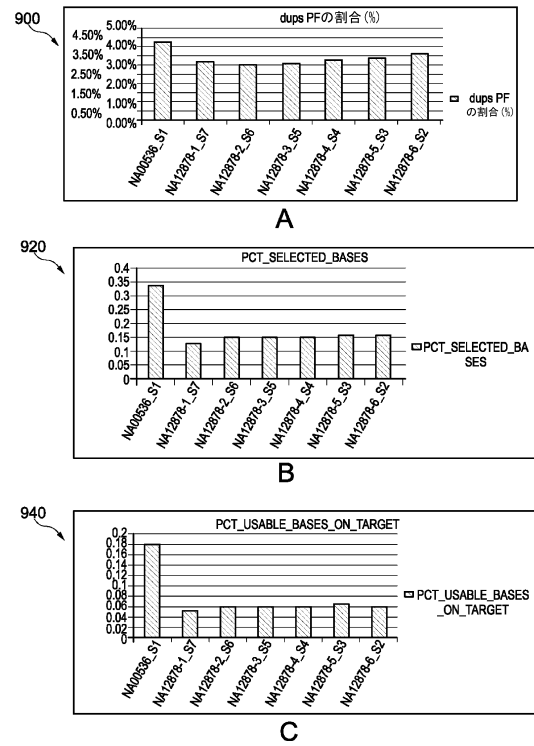
【図 7】



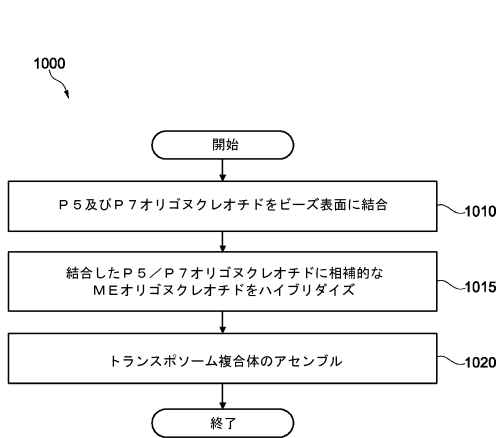
【図 8】



【図 9】



【図 10】



10

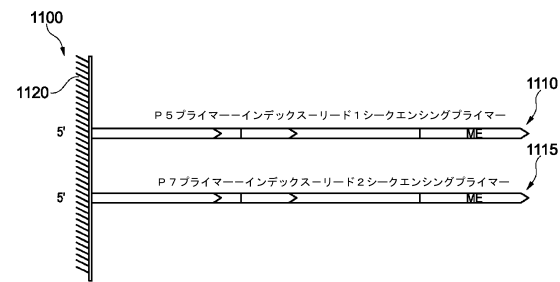
20

30

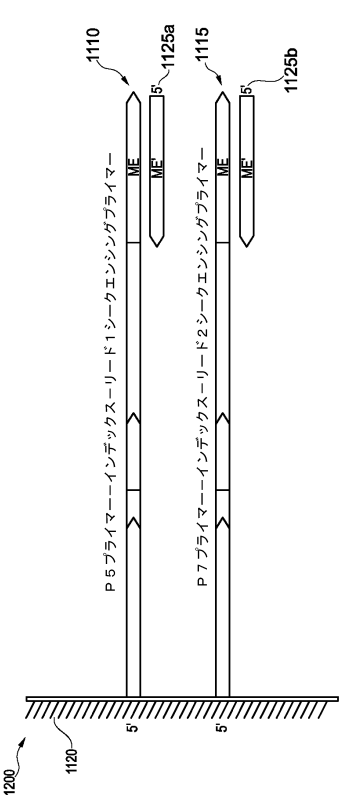
40

50

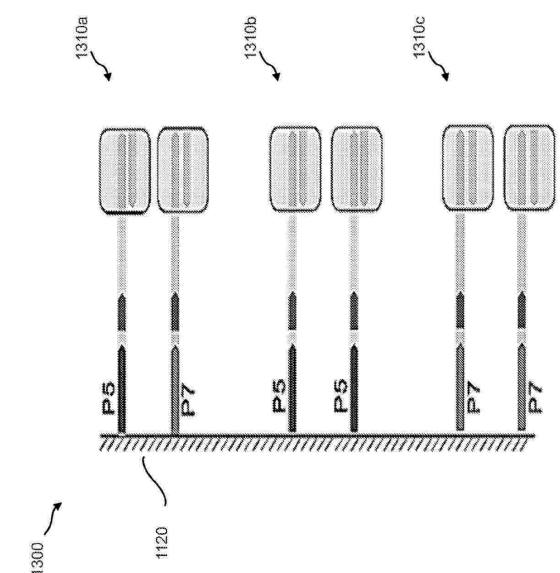
【図 1 1】



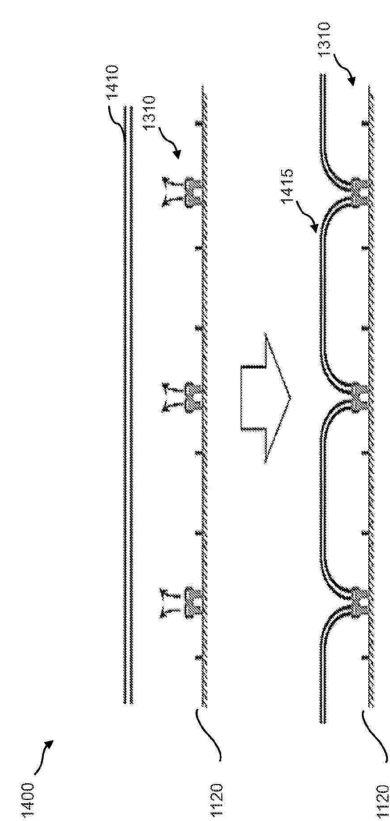
【図 1 2】



【図 1 3】



【図 1 4】



10

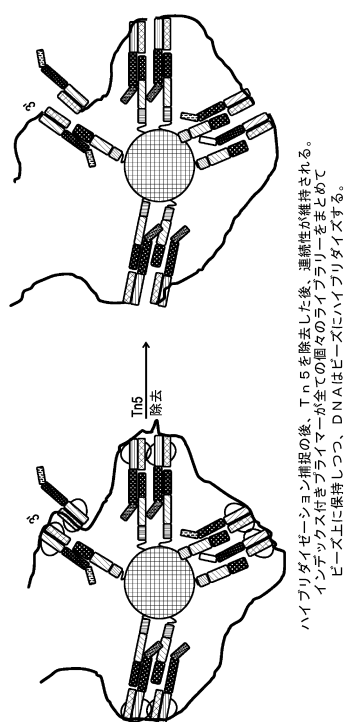
20

30

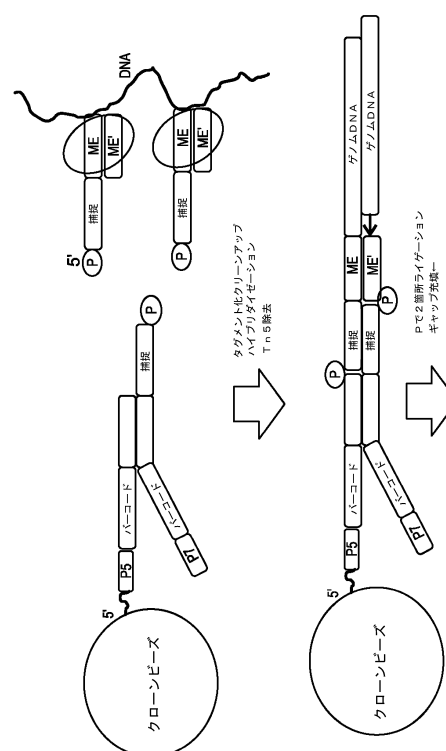
40

50

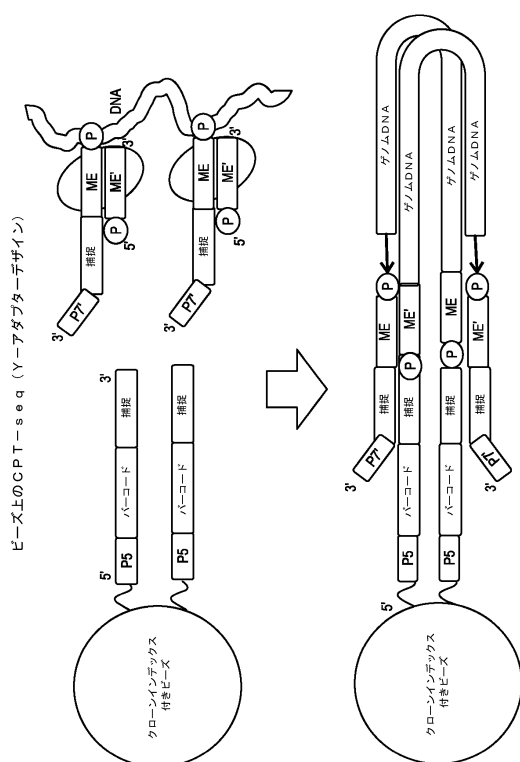
【図 19】



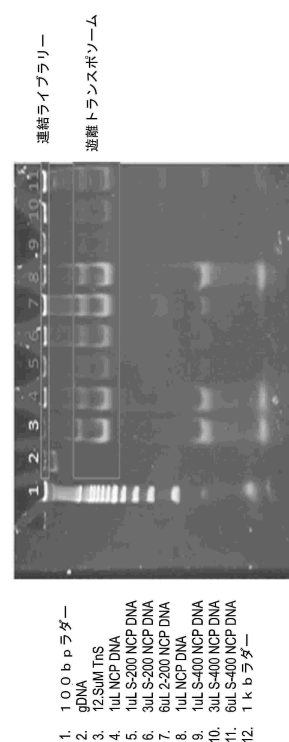
【図 20】



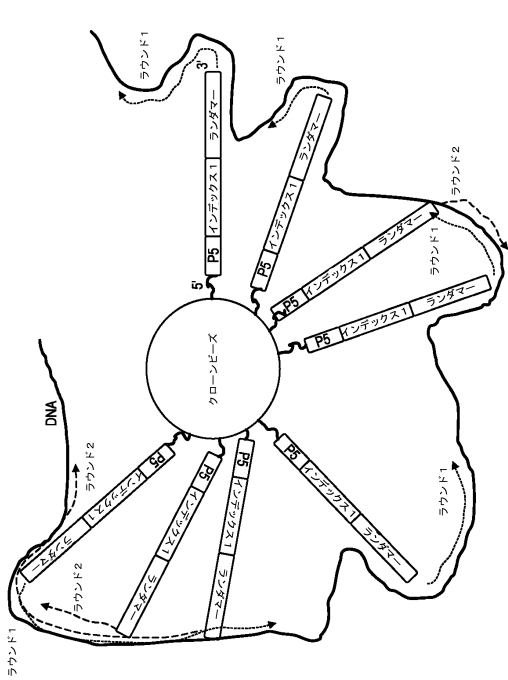
【図 2 1】



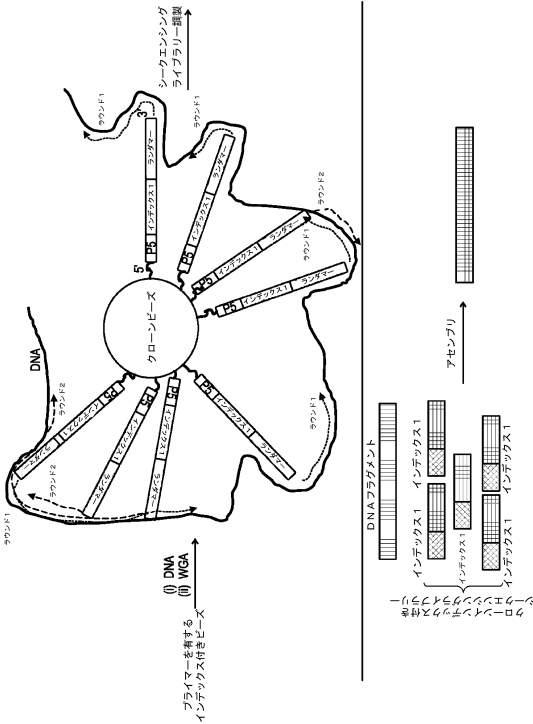
【図 22】



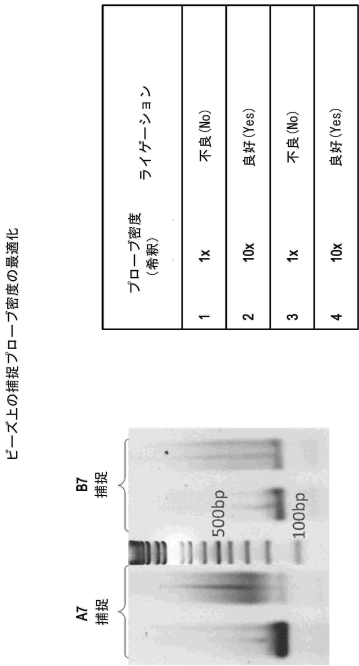
【図 2 3】



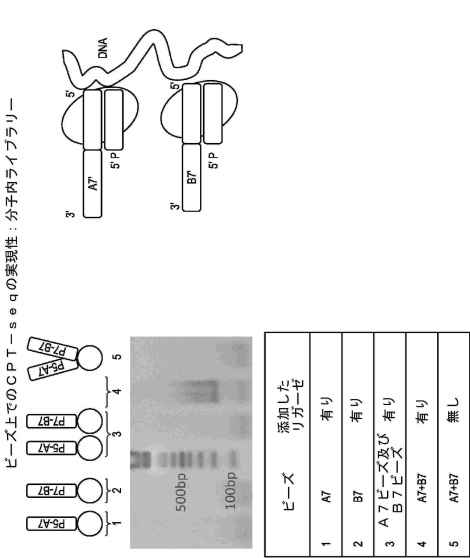
【図 2 4】



【図 2 5】



【図 2 6】



10

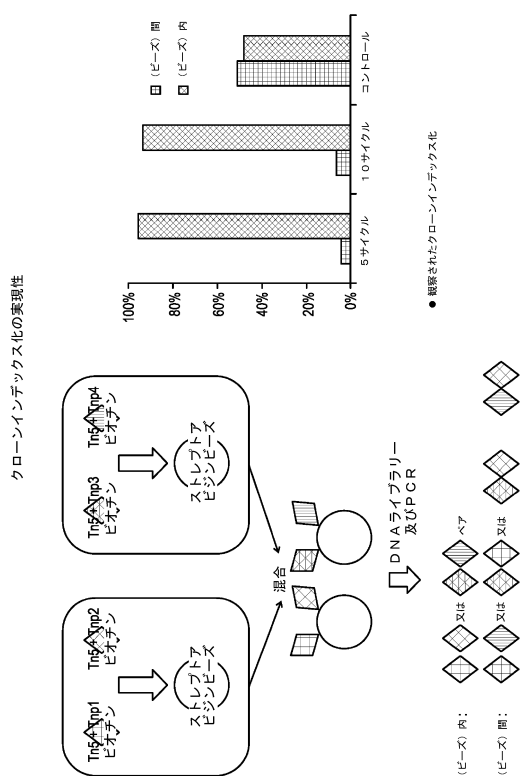
20

30

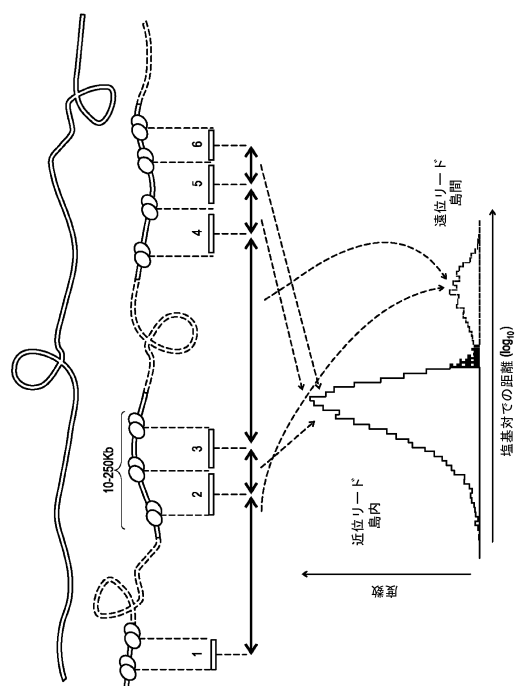
40

50

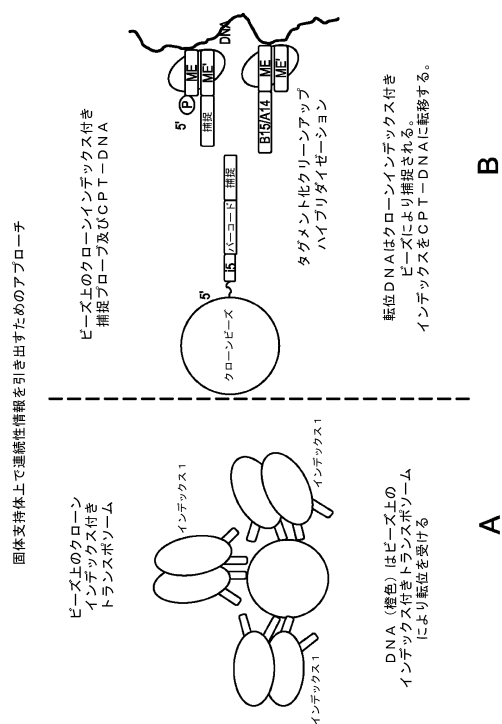
【図 27】



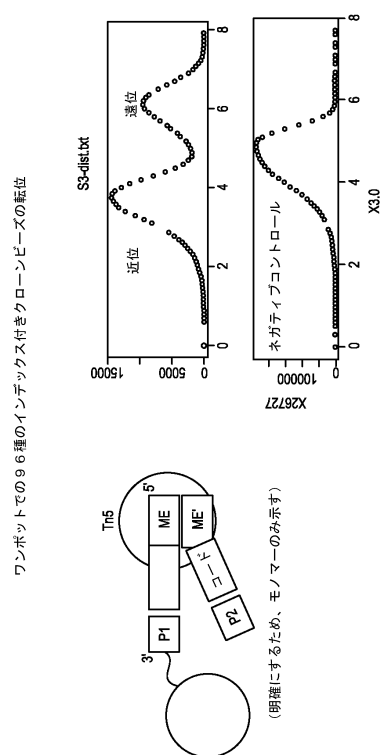
【図 28】



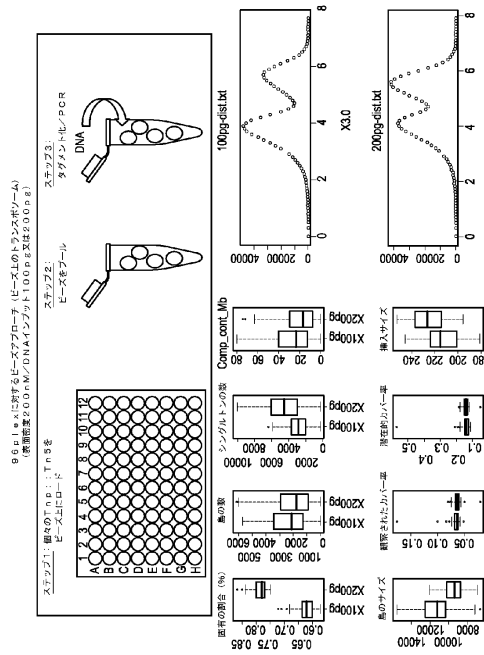
【圖 29】



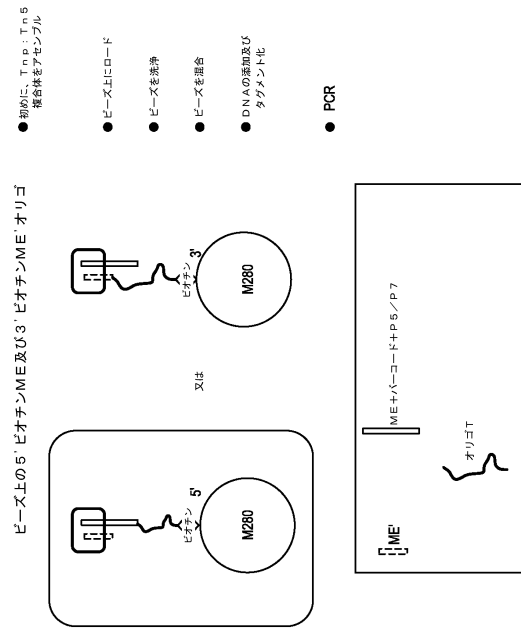
【図 30】



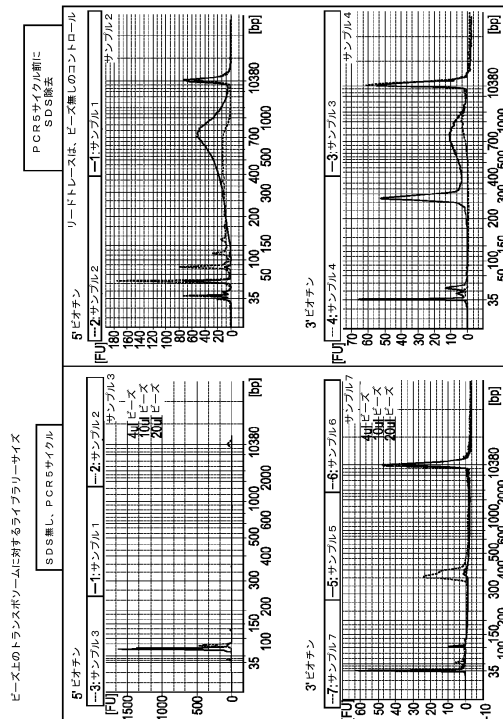
【図 3 1】



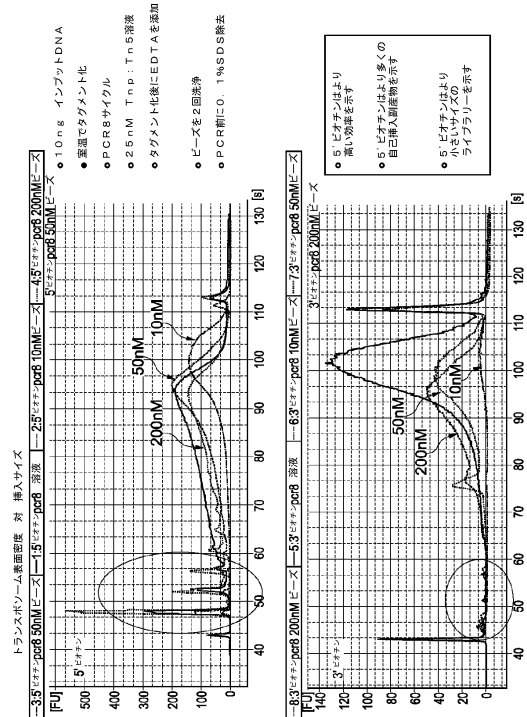
【図 3 2】



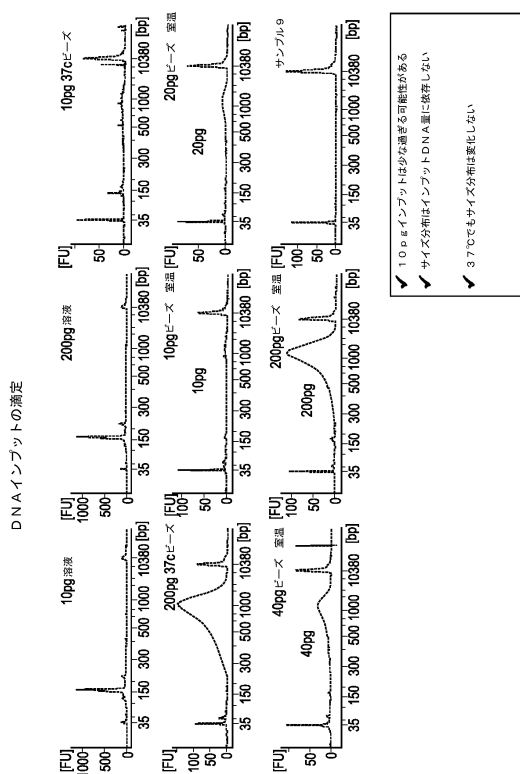
【図 3 3】



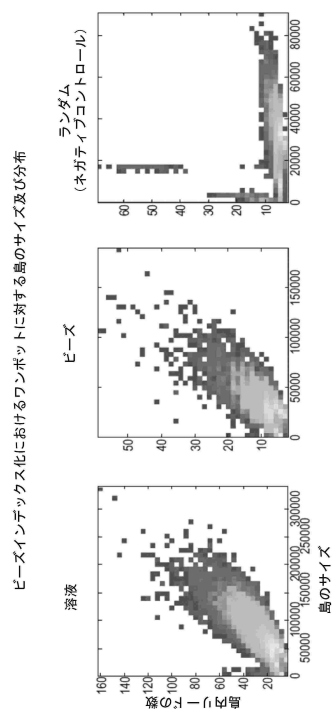
【図 3 4】



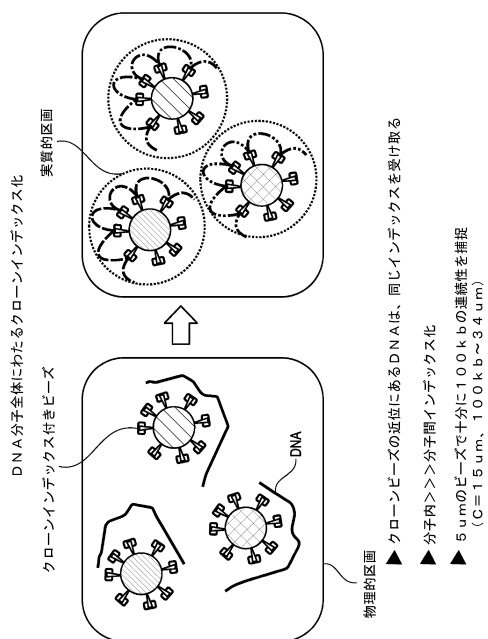
【図 3 5】



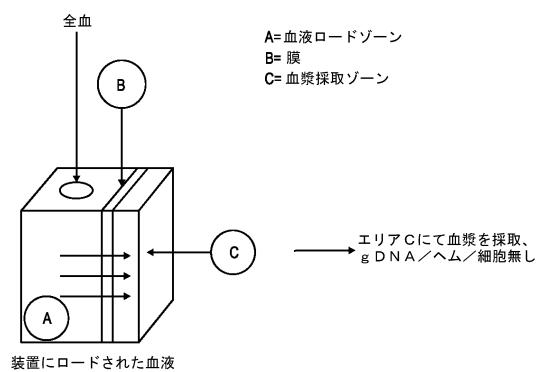
【図 3 6】



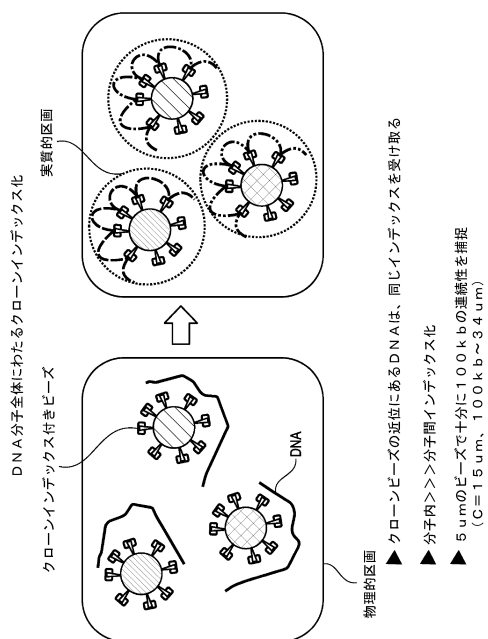
【図 3 7】



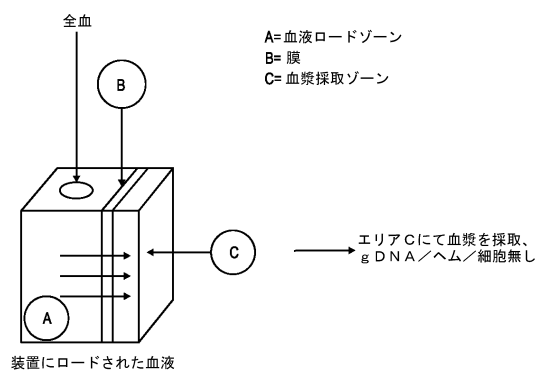
【図 38】



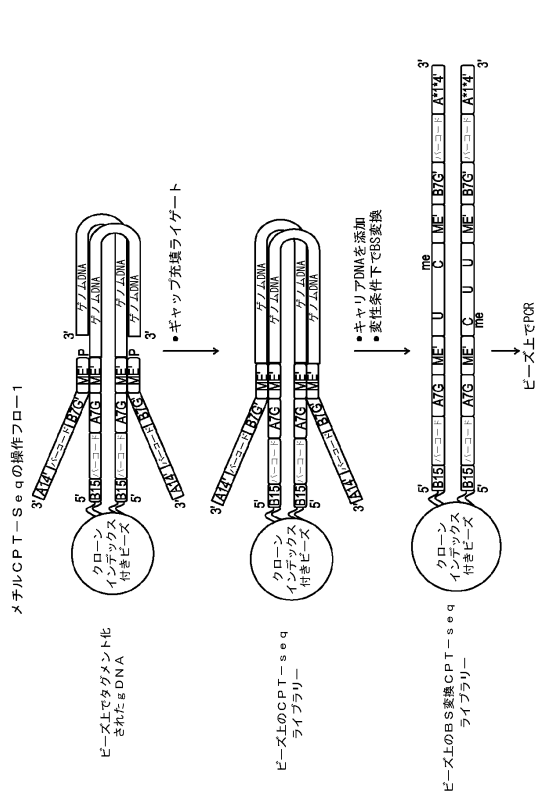
【図 3 7】



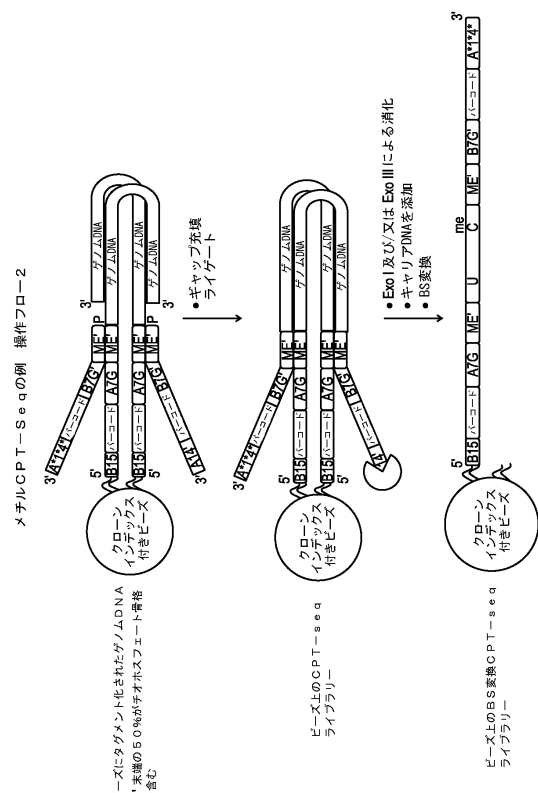
【図 38】



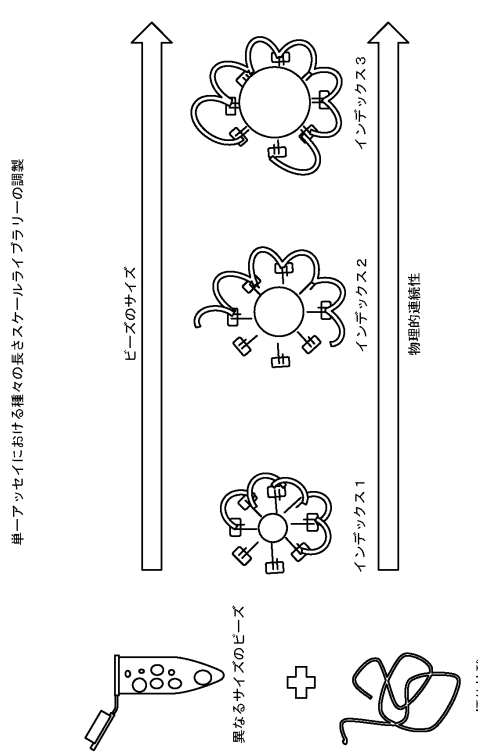
【図 4 3】



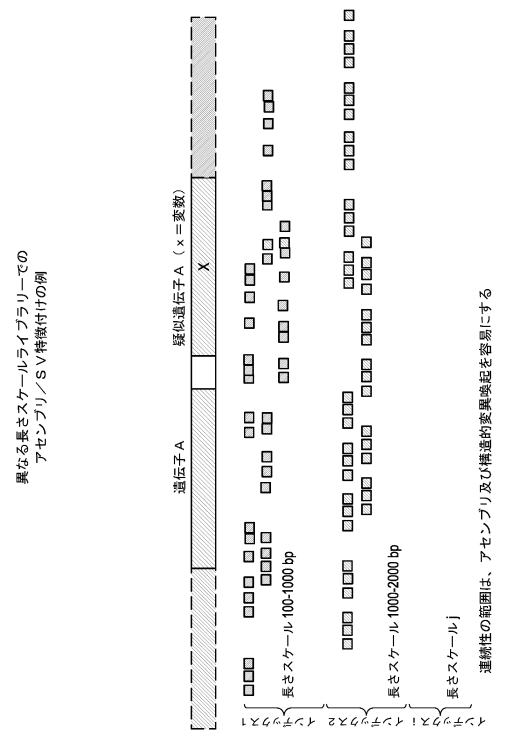
【図 4 4】



【図 4 5】



【図 4 6】



10

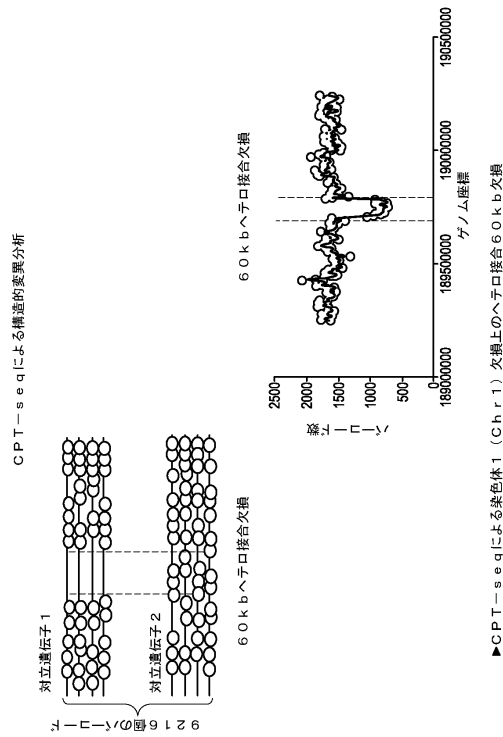
20

30

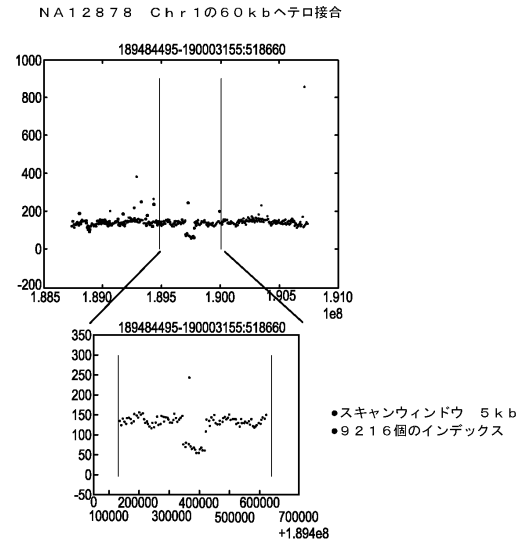
40

50

【図 47 A】



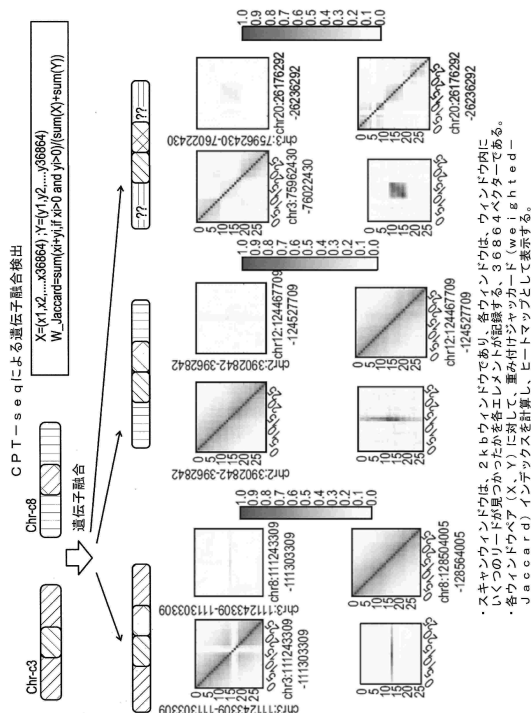
【図 47 B】



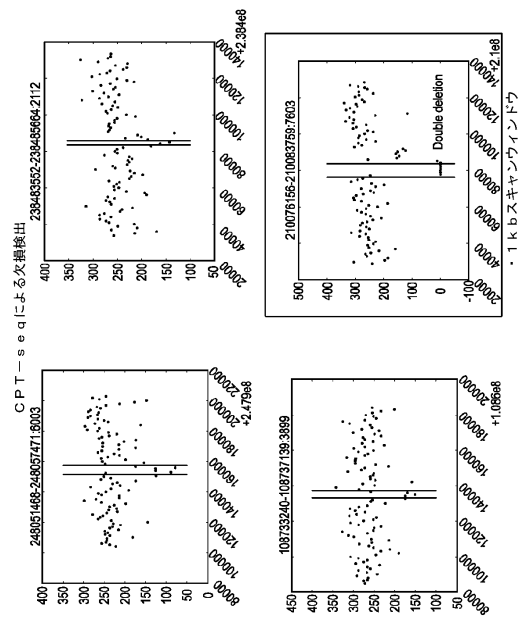
10

20

【図 48】



【図 49】



30

40

50

【図 50】

亜硫酸水素塩変換効率の最適化

・Promega社のMethyEdge亜硫酸酸素塩変換システムを最適化し、変換効率を改善

条件	DNA	ヒーズ	B S C処理
1	10ng	無し	60℃で1時間 / 0.3M NaOH
2			60℃で1時間 / 0.3M NaOH
3		有り	60℃で1時間 / 1M NaOH
4			65℃で1時間 / 0.3M NaOH

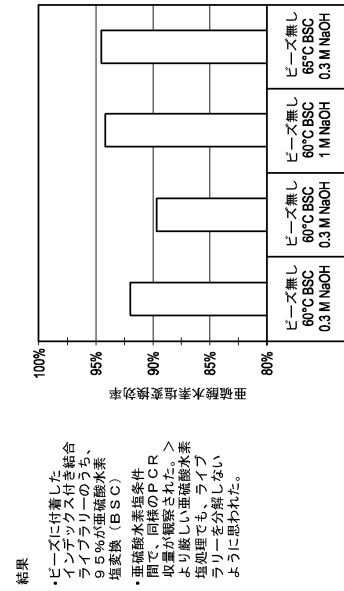
- ・亜硫酸水素塩変換処理の効率を決定するために分析したME配列

未変換ME	GATGTGTATAAGAGACAG
BSC ME	AATATATATAAAAAACAA

P7	B15'	バ-コード	A7G'	ME'	メノDNA	ME	B7G	バ-コード	P5
P7	B15'	バ-コード	A7G	ME	メノDNA	ME'	B7G'	バ-コード	P5'

【図 5 1】

亜硫酸水素塩変換効率の最適化

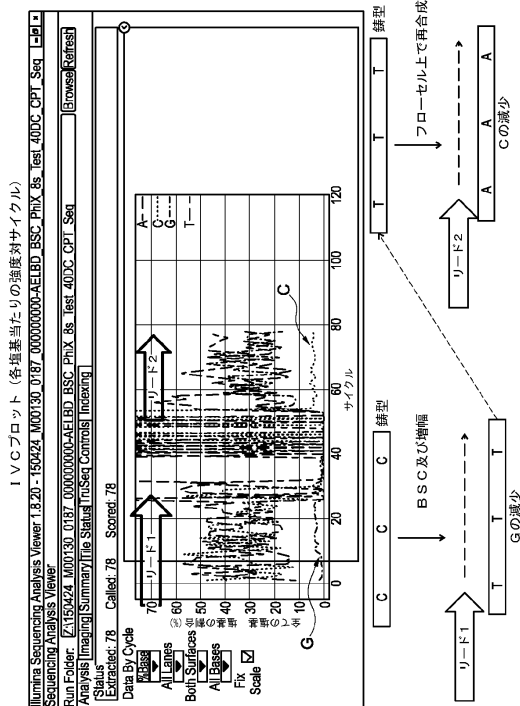


結果

- ・ビーズに付着したインデックス付き結合ライブラリーのうち、95%が亜硫酸水素塩変換(RSC)

・亜硫酸水素塩条件間で、同様のPCR収量が観察された。>より厳しい亜硫酸水素塩処理でも、ライブラリーを分解しないように思われた。

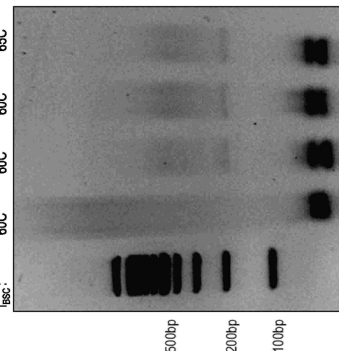
【図 5 2】



【圖 5 3】

亞硫酸水素塩変換効率の最適化

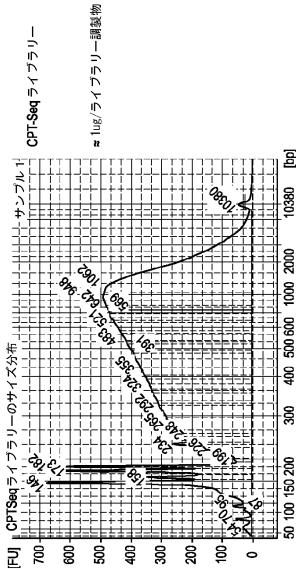
ビーズ:	無し	有り	有り	有り
[NaOH]:	0.3M	0.3M	1M	0.3M
T _{Res} ¹ :	60C	60C	60C	65C



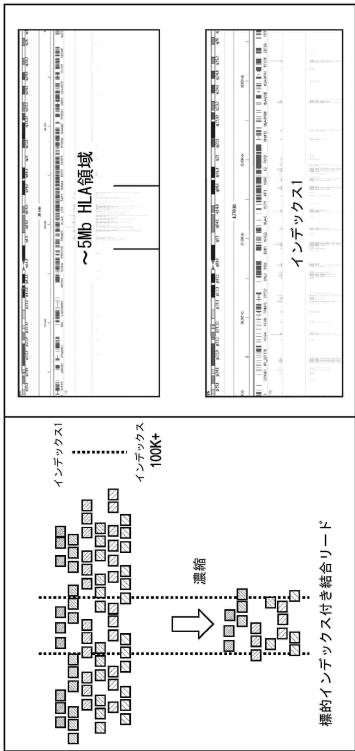
- 亜硫酸水素塩変換後のインデックス付き結合ライブラリー
- 亜硫酸水素塩変換効率～95%

【図 5 4】

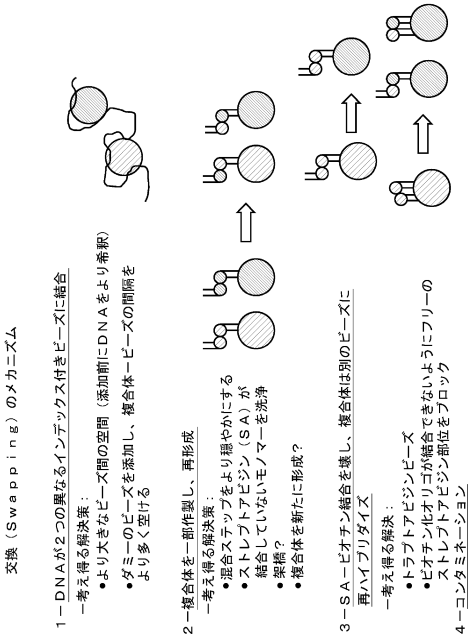
CPT-Seq ライブラリーの濃縮									
条件	ライブラリー	フロー-ブセット	【フロー-ブ】 pH/各フロー-ブ	【ハイブリダイゼーションランフ タイム】	濃縮ハイブリ ライブラリー	溶出	洗浄	SMB	
1	1.1ug CPTSeq	DNA Dumi A, B, K	100pM	25uL	95 > 68 (QN)	ET1	標準	3x	
2	1.1ug CPTSeq	EEX	2.5uL	25uL	95 > 68 (QN)	ET1	標準	3x	
3	1.1ug CPTSeq	RNA Dumi A, B, K	100pM	25uL	95 > 65 (QN)	RNase H	標準	3x	



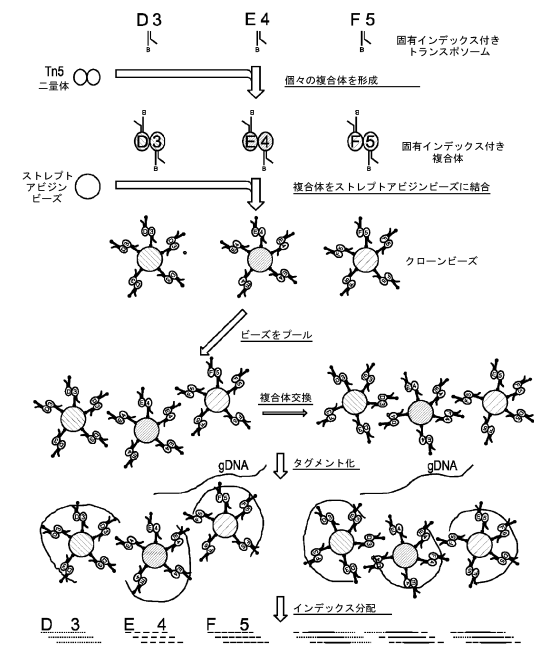
【図 5 6】



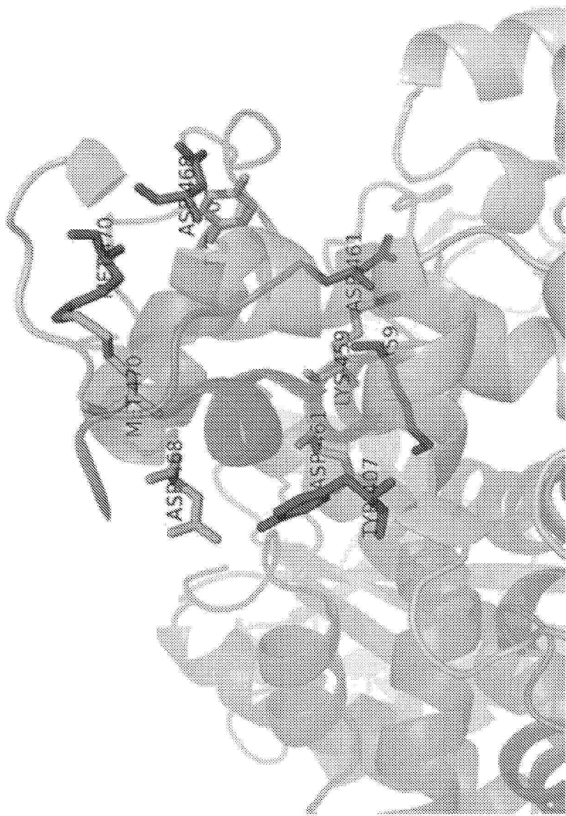
【図 5 7】



【図 5 8】



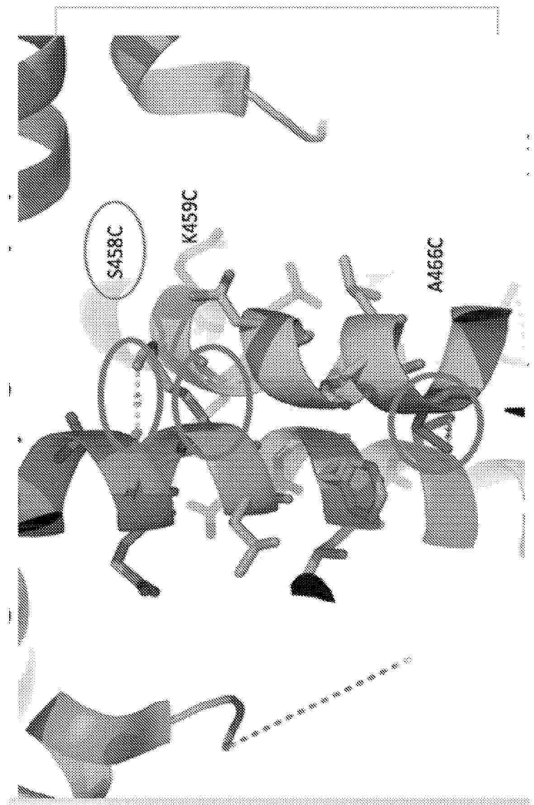
【図 5 9】



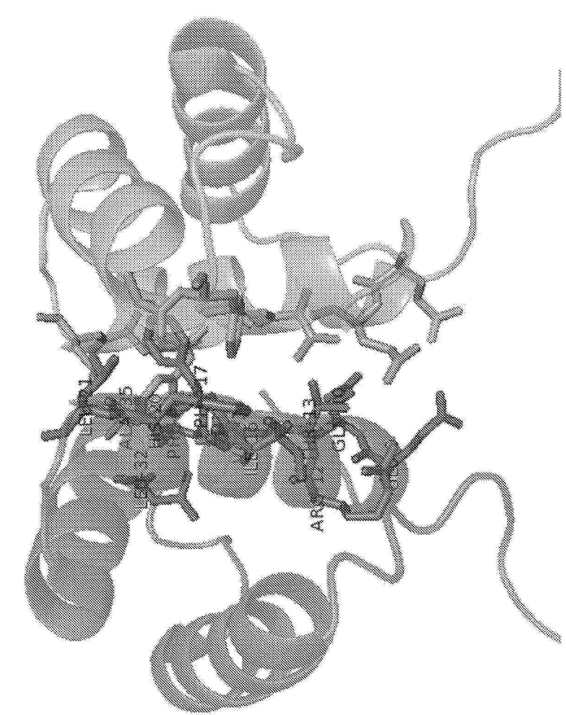
10

20

【図 6 0】



【図 6 1】

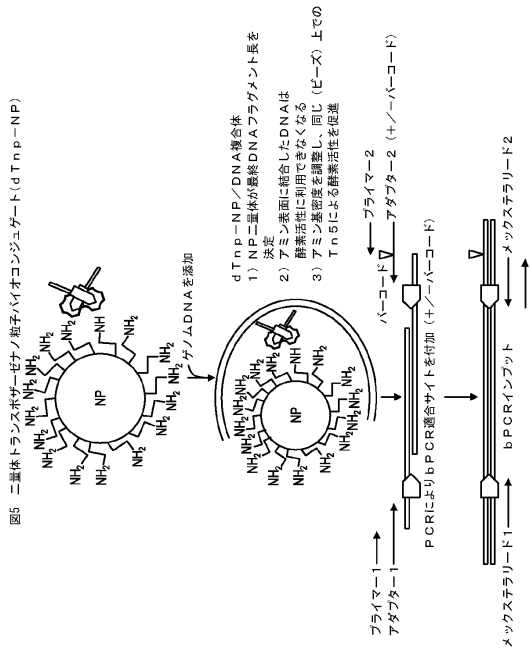


30

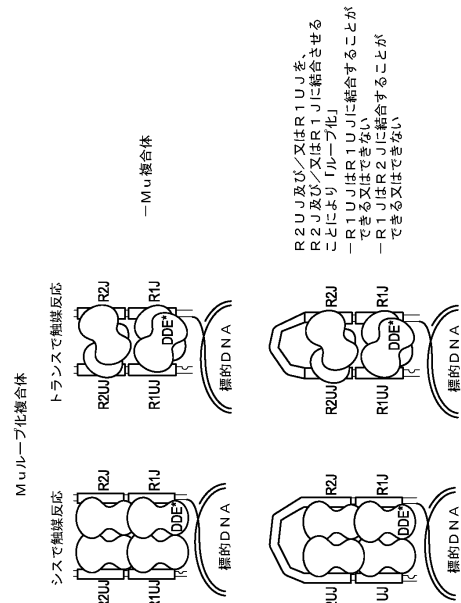
40

50

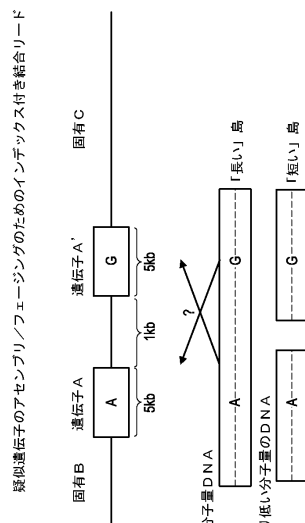
【図 6 2】



【図 6 3】



【図 6 4】



- ・例、CYP2D6
 ・「ゲンテツナ付き結合リードの長い島 (>20 kb) は、
 ・類似遺伝子、遺伝子A及びA'まで及ぶであろう
 ・長い (h i g h l y) 配列のため、その変異が遺伝子A又は遺伝子A'に
 ・及ぶことができない
 ・より短い島は、類似遺伝子の1つにおける2つの変異のみを固有の
 ・周囲配列に結合するであろう
 ・→ これらは、CPT-Jensen シンジンの組換えDNAをせん断することにより変異
 ・されたDNA及びJensen シンジンの組換えDNAを用いることによって選別
 ・することができ

【図 6 5】

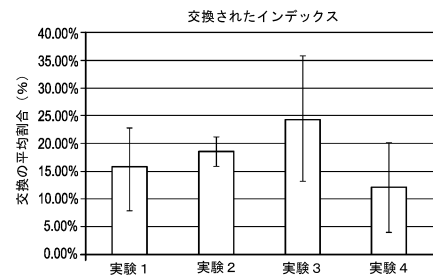
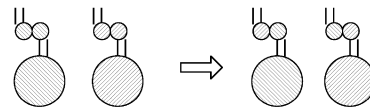
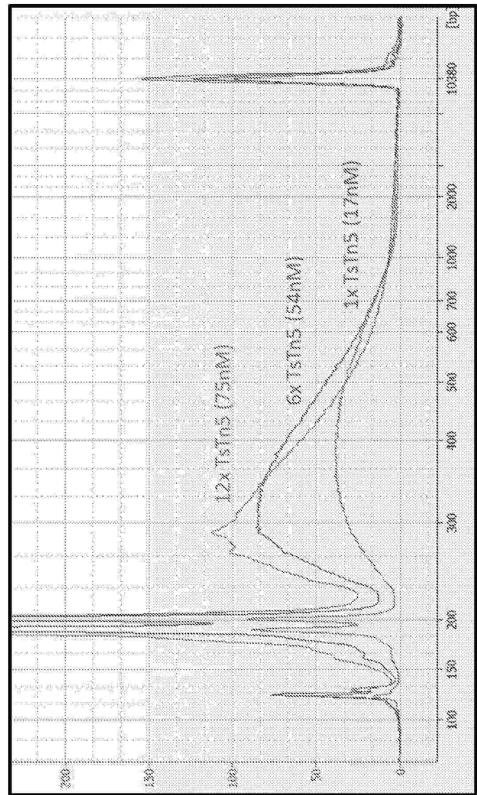


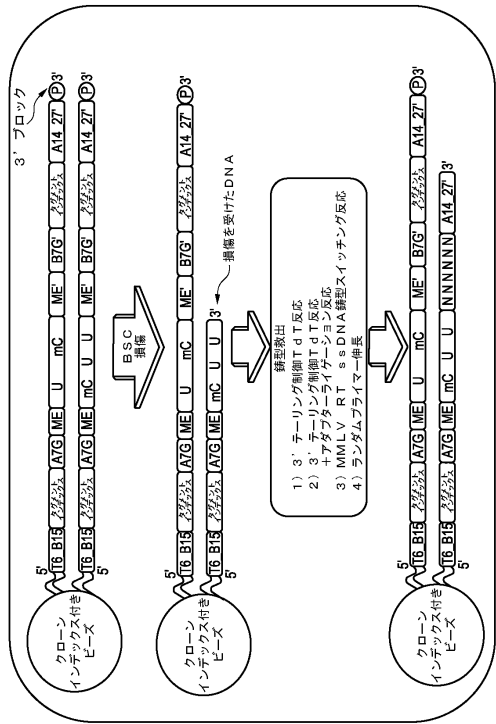
図1：インデックス交換。異なるインデックスを有するビーズを調製した。混合後、インデックス交換をライブラリーのシーケンシング及び各ライブラリーのインデックスを記録することにより決定した。「交換した (swapped)」割合 (%) を $(D4 + D5 + E3 + E5 + f4) / (\text{全96種の合計})$ により計算し、図に記録した。



【図 6 6】



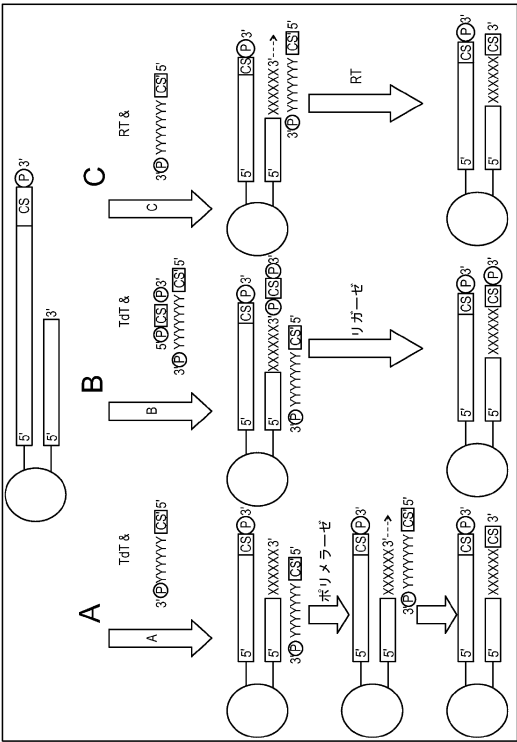
【図 6 7】



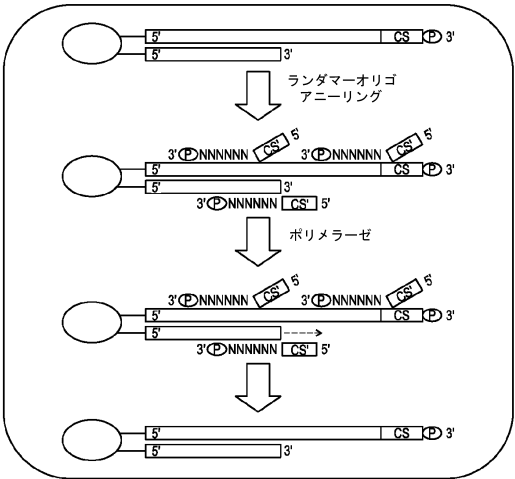
10

20

【図 6 8】



【図 6 9】

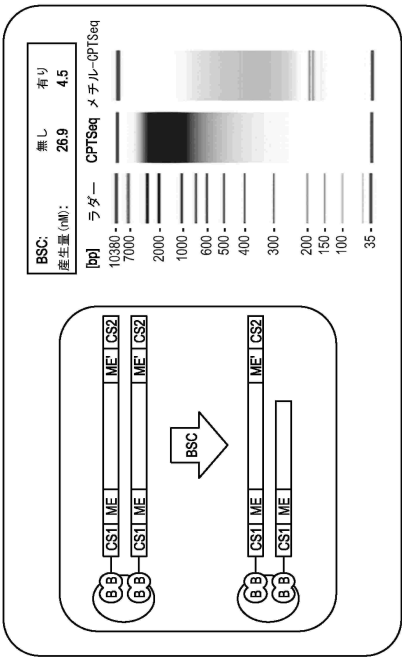


30

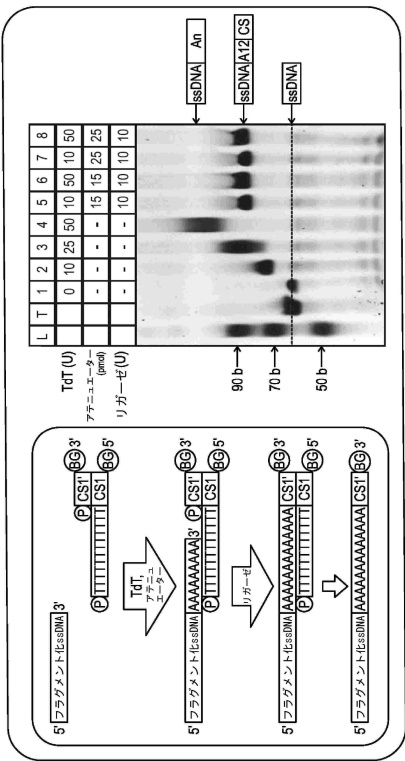
40

50

【図 7 0】



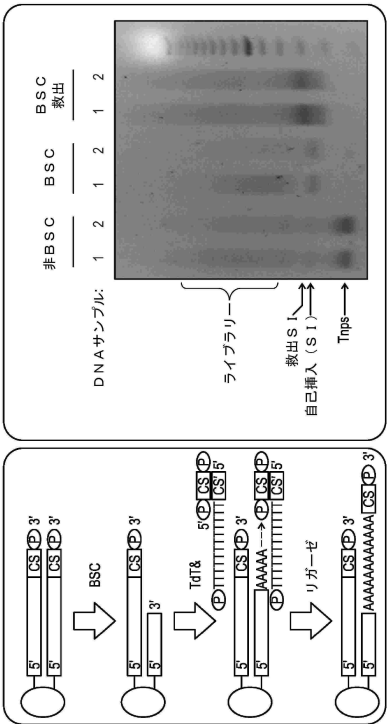
【図 7 1】



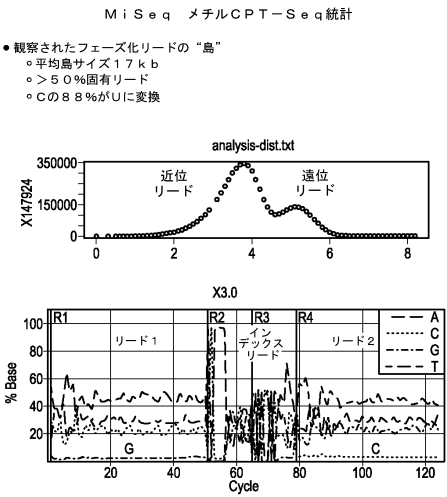
10

20

【図 7 2】



【図 7 3】



30

40

50

フロントページの続き

(33)優先権主張国・地域又は機関

米国(US)

(31)優先権主張番号 62/242,880

(32)優先日 平成27年10月16日(2015.10.16)

(33)優先権主張国・地域又は機関

米国(US)

(72)発明者 ファン チャン

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 ジェyson リチャード ベトリー

イギリス国 エセックス州 シービー1 0 1 エックスエル ニア サフロン ウォールデン リトル
チェスターフォード

(72)発明者 ニール アントニー ゴームリー

イギリス国 エセックス州 シービー1 0 1 エックスエル ニア サフロン ウォールデン リトル
チェスターフォード

(72)発明者 ウーター メウールマン

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 ジャクリーン ウィアー

イギリス国 エセックス州 シービー1 0 1 エックスエル ニア サフロン ウォールデン リトル
チェスターフォード

(72)発明者 アヴグスタ イオアノウ

イギリス国 エセックス州 シービー1 0 1 エックスエル ニア サフロン ウォールデン リトル
チェスターフォード

(72)発明者 ガレス ジェンキンス

イギリス国 エセックス州 シービー1 0 1 エックスエル ニア サフロン ウォールデン リトル
チェスターフォード

(72)発明者 ロザモンド ジャクソン

イギリス国 エセックス州 シービー1 0 1 エックスエル ニア サフロン ウォールデン リトル
チェスターフォード

(72)発明者 ナタリー モレル

イギリス国 エセックス州 シービー1 0 1 エックスエル ニア サフロン ウォールデン リトル
チェスターフォード

(72)発明者 ドミトリー ケイ ポコロック

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 スティーブン ジェイ ノルベルク

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 モーリー ホー

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 アミラーリ キア

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 イゴール ガルイシン

アメリカ合衆国 ウィスコンシン州 5 3 7 1 9 マディソン リサーチ パーク ブルバード 5 6 0
2 スイート 2 0 0

(72)発明者 リゴー パントーハ

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

審査官 松村 真里

(56)参考文献 国際公開第2014/108810 (WO, A2)

特開2013-150611 (JP, A)

特表2016-508715 (JP, A)

特表2014-506788 (JP, A)

特表 2013-535986 (JP, A)

米国特許出願公開第 2014/0235506 (US, A1)

国際公開第 2013/131962 (WO, A1)

国際公開第 2014/124338 (WO, A1)

米国特許出願公開第 2012/0053063 (US, A1)

(58)調査した分野 (Int.Cl., DB名)

C12Q

C12N

C40B

JSTPlus/JMEDPlus/JST7580 (JDreamIII)

CAPLUS/MEDLINE/EMBASE/BIOSIS (STN)