

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6674951号
(P6674951)

(45) 発行日 令和2年4月1日(2020.4.1)

(24) 登録日 令和2年3月11日(2020.3.11)

(51) Int.Cl.

F I

C 1 2 N 15/09 (2006.01)
 C 1 2 Q 1/6874 (2018.01)
 C 1 2 Q 1/6876 (2018.01)
 C 1 2 M 1/34 (2006.01)
 C 1 2 M 1/00 (2006.01)

C 1 2 N 15/09 Z N A Z
 C 1 2 N 15/09 2 0 0
 C 1 2 Q 1/6874 Z
 C 1 2 Q 1/6876 Z
 C 1 2 M 1/34 Z

請求項の数 18 (全 46 頁) 最終頁に続く

(21) 出願番号 特願2017-527209 (P2017-527209)
 (86) (22) 出願日 平成27年11月19日(2015.11.19)
 (65) 公表番号 特表2017-535269 (P2017-535269A)
 (43) 公表日 平成29年11月30日(2017.11.30)
 (86) 国際出願番号 PCT/US2015/061615
 (87) 国際公開番号 W02016/081740
 (87) 国際公開日 平成28年5月26日(2016.5.26)
 審査請求日 平成30年11月16日(2018.11.16)
 (31) 優先権主張番号 62/082,883
 (32) 優先日 平成26年11月21日(2014.11.21)
 (33) 優先権主張国・地域又は機関
 米国 (US)

早期審査対象出願

(73) 特許権者 513322707
 ナノストリング テクノロジーズ, インコ
 ーポレイティド
 アメリカ合衆国, ワシントン 98109
 , シアトル, フェアビュー アベニュー ノ
 ース 530, スイート 2000
 (74) 代理人 100099759
 弁理士 青木 篤
 (74) 代理人 100077517
 弁理士 石田 敬
 (74) 代理人 100087871
 弁理士 福本 積
 (74) 代理人 100087413
 弁理士 古賀 哲次

最終頁に続く

(54) 【発明の名称】 酵素不要及び増幅不要の配列決定

(57) 【特許請求の範囲】

【請求項 1】

標的結合ドメインとバーコードドメインとを含む配列決定プローブであって、

前記標的結合ドメインは、少なくとも8つのヌクレオチドを含み、かつ標的核酸に結合
 することができ、前記標的結合ドメイン中の少なくとも2つのヌクレオチドは、前記標的
 核酸内の任意の塩基に結合することができるユニバーサル塩基であり；

前記バーコードドメインは合成主鎖を含み、前記バーコードドメインは少なくとも6つ
 の付着位置を含み、それぞれの付着位置は少なくとも1つの付着領域を含み、前記付着領
 域は相補的核酸分子にハイブリダイズする少なくとも1つの核酸配列を含み；

前記少なくとも6つの付着位置のそれぞれの付着位置は、標的結合ドメイン中の1つの
 ヌクレオチドに相当し、前記少なくとも6つの付着位置のそれぞれは異なった核酸配列を
 有し、前記少なくとも6つの付着位置のそれぞれの位置の核酸配列は、前記標的結合ドメ
 インによって結合される前記標的核酸中の対応する1つのヌクレオチドの位置及び本体を
 決定し、

前記ユニバーサル塩基は、前記バーコードドメインの位置に対応するヌクレオチドの前
 後に有する、配列決定プローブ。

【請求項 2】

前記合成主鎖が1本鎖DNAを含む、請求項1に記載の配列決定プローブ。

【請求項 3】

前記標的結合ドメインと前記バーコードドメインとの間に2本鎖DNAスペーサーを含

10

20

む、請求項 1 に記載の配列決定プローブ。

【請求項 4】

標的結合ドメイン中のヌクレオチドの数は、バーコードドメイン中の付着位置の数よりも少なくとも 2 つ多い、請求項 1 に記載の配列決定プローブ。

【請求項 5】

前記バーコードドメイン中のそれぞれの位置が、(a) 同じ数の付着領域、(b) 1 つの付着領域、又は (c) 1 超の付着領域、を有する、請求項 1 に記載の配列決定プローブ。

【請求項 6】

それぞれの相補的核酸分子が、

(a) それぞれの付着位置について、検出可能標識物を含み；あるいは

(b) 核酸分子スペーサーを介して第 1 の核酸分子に間接的に連結され；あるいは、

(c) それぞれの付着位置について、8 つのヌクレオチド ~ 20 のヌクレオチドを含み、それぞれの相補的核酸分子が 1 2 つのヌクレオチドを含む、請求項 1 に記載の配列決定プローブ。

【請求項 7】

それぞれの第 1 の核酸分子が少なくとも 1 つ、2 つ、3 つ、4 つ又は 5 つの第 2 の核酸分子とハイブリダイズする、請求項 6 に記載の配列決定プローブ。

【請求項 8】

前記第 2 の核酸分子が少なくとも 1 つの検出可能標識物を含む、請求項 7 に記載の配列決定プローブ。

【請求項 9】

それぞれの第 2 の核酸分子が、少なくとも 1 つ、2 つ、3 つ、4 つ、5 つ、6 つ又は 7 つの、少なくとも 1 つの検出可能標識物を含む第 3 の核酸分子とハイブリダイズする、請求項 7 に記載の配列決定プローブ。

【請求項 10】

前記バーコードドメイン中の 1 つ以上の付着位置が、少なくとも 1 つのフランキング 1 本鎖ポリヌクレオチドに隣接する、請求項 1 に記載の配列決定プローブ。

【請求項 11】

複数の請求項 1 に記載の配列決定プローブを含む配列決定プローブ集団。

【請求項 12】

以下の工程を含む核酸の配列決定方法：

(1) 請求項 1 ~ 10 のいずれか 1 項記載の複数の配列決定プローブを含む第 1 配列決定プローブの少なくとも 1 つの第 1 集団を、基板に固定化された標的核酸にハイブリダイズさせる工程、ここで、前記標的核酸は 1 以上の位置に固定化される；

(2) 少なくとも 6 つの付着位置の第 1 付着位置に、検出可能標識物を含む第 1 の相補的核酸分子、又は検出可能標識物を含む第 1 レポーター複合体の第 1 の相補的核酸分子を結合させる工程；

(3) 結合された第 1 相補的核酸分子の検出可能標識物、又は第 1 レポーター複合体の結合された第 1 の相補的核酸分子の検出可能標識物を検出する工程；

(4) 固定化標的核酸中の第 1 ヌクレオチドの位置及び本体を特定する工程；

(5) 第 1 付着位置に、検出可能標識物を欠く第 1 のハイブリダイズ性核酸分子を結合させ、こうして、検出可能標識物を含む第 1 の相補的核酸分子、又は検出可能標識物を含む第 1 のレポーター複合体の第 1 の相補的核酸分子の結合をはずす工程；

(6) 少なくとも 6 つの付着位置の第 2 付着位置に、検出可能標識物を含む第 2 の相補的核酸分子、又は検出可能標識物を含む第 2 レポーター複合体の第 2 の相補的核酸分子を結合させる工程；

(7) 結合された第 2 の相補的核酸分子の検出可能標識物、又は第 2 のレポーター複合体の結合された第 2 の相補的核酸分子の検出可能標識物を検出する工程；

(8) 固定化標的核酸中の第 2 ヌクレオチドの位置及び本体を特定する工程；

10

20

30

40

50

(9) 少なくとも 6 つの付着位置の各々の付着位置が、検出可能標識物を含む相補的核酸分子、又は検出可能標識物を含むレポーター複合体の相補的核酸分子、及び結合された相補的核酸分子の検出可能標識物、又はレポーター複合体の結合された相補的核酸分子の検出可能標識物によって結合されるまで、工程 (5) ~ (8) を繰り返し、こうして、配列決定プローブの標的結合ドメインによってハイブリダイズされた固定化された標的核酸の少なくとも第 1 領域のための少なくとも 6 つのヌクレオチドの線状順序を特定する工程 ; そして

(10) 固定された標的核酸から第 1 配列決定プローブの少なくとも 1 つの第 1 集団を脱ハイブリダイズする工程。

【請求項 13】

工程 (5) 及び (6) は連続して又は同時に行われる、請求項 12 に記載の方法。

【請求項 14】

第 1 の相補的核酸分子及び検出可能標識物を欠く第 1 のハイブリダイズ性核酸分子が、同一の核酸配列を含む、請求項 12 に記載の方法。

【請求項 15】

検出可能標識物を欠く第 1 のハイブリダイズ核酸分子が、第 1 付着位置に隣接したフランキング 1 本鎖ポリヌクレオチドに相補的な核酸配列を含む、請求項 12 に記載の方法。

【請求項 16】

以下の工程を更に含む請求項 12 に記載の方法 :

(11) 請求項 1 ~ 10 のいずれか 1 項記載の複数の配列決定プローブを含む第 2 配列決定プローブの少なくとも 1 つの第 2 集団を、基板に固定化された標的核酸にハイブリダイズさせる工程、ここで、前記標的核酸は 1 以上の位置の基板に固定化され、第 1 配列決定プローブ及び第 2 配列決定プローブの標的結合ドメインは異なり、第 2 集団の各配列決定プローブは、固定された標的核酸から同じ条件下で脱ハイブリダイズし、及び固定された標的核酸から、第 1 集団中の配列決定プローブとは異なる条件下で脱ハイブリダイズする ;

(12) 少なくとも 6 つの付着位置の第 1 付着位置に、検出可能標識物を含む第 1 の相補的核酸分子、又は検出可能標識物を含む第 1 レポーター複合体の第 1 の相補的核酸分子を結合させる工程 ;

(13) 結合された第 1 相補的核酸分子の検出可能標識物、又は第 1 レポーター複合体の結合された第 1 の相補的核酸分子の検出可能標識物を検出する工程 ;

(14) 固定化標的核酸中の第 1 ヌクレオチドの位置及び本体を特定する工程 ;

(15) 第 1 付着位置に、検出可能標識物を欠く第 1 のハイブリダイズ性核酸分子を結合させ、こうして、検出可能標識物を含む第 1 の相補的核酸分子、又は検出可能標識物を含む第 1 のレポーター複合体の第 1 の相補的核酸分子の結合をはずす工程 ;

(16) 少なくとも 6 つの付着位置の第 2 付着位置に、検出可能標識物を含む第 2 の相補的核酸分子、又は検出可能標識物を含む第 2 レポーター複合体の第 2 の相補的核酸分子を結合させる工程 ;

(17) 結合された第 2 の相補的核酸分子の検出可能標識物、又は第 2 のレポーター複合体の結合された第 2 の相補的核酸分子の検出可能標識物を検出する工程 ;

(18) 固定化標的核酸中の第 2 ヌクレオチドの位置及び本体を特定する工程 ;

(19) 少なくとも 6 つの付着位置の各々の付着位置が、検出可能標識物を含む相補的核酸分子、又は検出可能標識物を含むレポーター複合体の相補的核酸分子、及び結合された相補的核酸分子の検出可能標識物、又はレポーター複合体の結合された相補的核酸分子の検出可能標識物によって結合されるまで、工程 (15) ~ (18) を繰り返し、こうして、配列決定プローブの標的結合ドメインによってハイブリダイズされた固定化された標的核酸の少なくとも第 2 領域のための少なくとも 6 つのヌクレオチドの線状順序を特定する工程 ; そして

(20) 固定された標的核酸から第 2 配列決定プローブの少なくとも 1 つの第 2 集団を脱ハイブリダイズする工程。

10

20

30

40

50

【請求項 17】

固定化された標的核酸の少なくとも第1領域及び少なくとも第2領域におけるヌクレオチドの各特定された線状順序を組み立て、こうして固定化された標的核酸のための配列を特定する、工程を更に含む、請求項16に記載の方法。

【請求項 18】

基板、

請求項1～10のいずれか1項に記載の複数の配列決定プローブ、

少なくとも2つの捕捉プローブ、

配列決定プローブの少なくとも6つの付着位置の1つにハイブリダイズする核酸配列、及び検出可能標識物を含む、検出可能標識物を含む少なくとも6つの相補的核酸分子、

配列決定プローブの少なくとも6つの付着位置の1つにハイブリダイズする核酸配列を含む、検出可能標識物を欠く少なくとも6つの相補的核酸分子、及び

使用説明書

を含む、キット。

【発明の詳細な説明】

【技術分野】

【0001】

(関連出願への相互参照)

本出願は、2014年11月21日に出願された米国仮特許出願第62/082,883号の利益を主張する。前記特許出願の内容は、参照によりその全体が本明細書に組み込まれる。

【0002】

(配列リスト)

本出願は、E F S - W e bを介してA S C I Iフォーマットで提出された配列リストを含み、その全体は参照により本明細書に組み込まれる。2015年11月19日に作成されたこのA S C I Iコピーは、N A T E - 0 2 5 _ S T 2 5 . t x tという名前で、サイズは20,860バイトである。

【背景技術】

【0003】

(発明の背景)

現在、核酸配列決定のための様々な方法、すなわち核酸分子内のヌクレオチドの正確な順序を決定する方法が存在する。現在の方法は、酵素的、例えばPCR及び/又はクローニングにより、核酸を増幅することを必要とする。光検出手段により検出可能な信号を生成するためには、さらなる酵素的重合が必要である。このような増幅及び重合工程は、費用及び/又は時間がかかる。従って当該技術分野では、増幅不要でかつ酵素不要の核酸配列決定の方法が必要とされている。本発明は、これらのニーズに取り組む。

【発明の概要】

【発明が解決しようとする課題】

【0004】

(発明の要約)

本発明は、長い読み取り長さ及び低いエラー率を有する、酵素不要で、増幅不要で、及びライブラリー不要の核酸配列決定法を提供する、配列決定プローブ、方法、キット、及び装置を提供する。更に、本方法、キット、及び装置は、試料から結果への迅速な能力を有する。これらの特徴は、臨床環境での配列決定に特に有用である。

【課題を解決するための手段】

【0005】

本明細書において、標的結合ドメインとバーコードドメインとを含む配列決定プローブが提供される。標的結合ドメイン及びバーコードドメインは、機能的に連結されてもよく、例えば共有結合されていてもよい。配列決定プローブは場合により、標的結合ドメインとバーコードドメインとの間にスペーサーを含む。スペーサーは、適切な機械的特性を有

10

20

30

40

50

する任意のポリマー、例えば1本鎖又は2本鎖DNAスパーサー(1~100ヌクレオチド、例えば2~50ヌクレオチドの)でもよい。2本鎖DNAスパーサーの非限定例には、配列番号25~配列番号29により包含される配列が含まれる。

【0006】

標的結合ドメインは、少なくとも4つ(例えば、4、5、6、7、8、9、10、11、12、又はそれ以上)のヌクレオチドを含み、標的核酸(例えばDNA、RNA、及びPNA)に結合することができる。バーコードドメインは合成主鎖を含み、バーコードドメインは、1つ以上の付着領域を含む少なくとも第1の位置を有する。バーコードドメインは、1、2、3、4、5、6、7、8、9、10、11、12、又はそれ以上の位置を有し、各位置は1つ以上(例えば1~50)の付着領域を有し、各付着領域は、相補的核酸分子(RNA又はDNA)に可逆的に結合可能な少なくとも1つのコピー(すなわち、1~50、例えば10から30コピー)の核酸配列を含む。バーコードドメイン内のある位置が、他の位置よりも多くの付着領域を有することがあり、あるいはバーコードドメイン内の各位置は、同じ数の付着領域を有する。第1付着領域の核酸配列は、標的結合ドメインの第1ヌクレオチドにより結合される標的核酸中の第1ヌクレオチドの位置及び本体を決定し、一方、第2付着領域の核酸配列は、標的結合ドメインの第2ヌクレオチドにより結合される標的核酸中の第2ヌクレオチドの位置及び本体を決定する。同様に、第6付着領域の核酸配列は、標的結合ドメインの第6ヌクレオチドにより結合される標的核酸中の第6ヌクレオチドの位置及び本体を決定する。実施態様では合成主鎖は、多糖、ポリヌクレオチド(例えば、1本鎖又は2本鎖DNA又はRNA)、ペプチド、ペプチド核酸、又はポリペプチドを含む。標的結合ドメイン中のヌクレオチドの数は、バーコードドメイン中の位置の数に等しいか又はそれより多い(例えば、1、2、3、4、又はそれ以上)。バーコードドメインの特定の位置にある各付着領域は、同じ核酸配列の1つのコピー及び/又は同じ核酸配列の複数のコピーを含むことができる。しかしながら付着領域は、たとえ両方の付着領域が同じタイプのヌクレオチド(例えばアデニン、チミン、シトシン、グアニン、ウラシル、及びそれらの類似体)を特定する場合でさえ、付着領域とは異なる核酸配列を、バーコードドメインの異なる位置に含むであろう。付着領域は、合成主鎖中の修飾されたモノマー、例えば修飾されたヌクレオチドに連結されることができ、こうして、主鎖に対して分岐を形成する。付着領域は、合成主鎖のポリヌクレオチド配列の一部であってもよい。1つ以上の付着領域が、少なくとも1つのフランキング1本鎖ポリヌクレオチドに隣接していてもよく、すなわち、付着領域は、5'フランキング1本鎖ポリヌクレオチド及び/又は3'フランキング1本鎖ポリヌクレオチドに機能的に連結され得る。1つ又は2つのフランキング1本鎖ポリヌクレオチドを伴うか又は伴わない付着領域は、検出可能標識物を欠くハイブリダイズ性核酸分子にハイブリダイズされ得る。検出可能標識物を欠くハイブリダイズ性核酸分子は、長さが約4~約20ヌクレオチド、例えば12ヌクレオチド又はそれ以上でもよい。

【0007】

付着領域は、検出可能標識物を含む相補的核酸により結合されていてもよい。各相補的核酸は、検出可能標識物を含み得る。

【0008】

あるいは付着領域は、レポーター複合体の一部である相補的核酸(検出可能標識物を含む)により結合されてもよい。(検出可能標識物を含むか、又はレポーター複合体の)相補的核酸は、長さが約4~約20ヌクレオチド、例えば約8、10、12、及び14ヌクレオチド、又はそれ以上であり得る。レポーター複合体において、相補的核酸は、1次核酸分子に(直接的又は間接的に)連結される。相補的核酸は、1本鎖又は2本鎖の核酸リンカー(例えば、1~100個のヌクレオチドを含むポリヌクレオチド)を介して、1次核酸分子に間接的に連結され得る。1次核酸は、1つ以上の(例えば、1、2、3、4、5、6、7、8、9、10、又はそれ以上の)2次核酸にハイブリダイズされる。各2次核酸は、1つ以上の(例えば、1、2、3、4、5、6、7、8、9、10、又はそれ以上の)3次核酸にハイブリダイズされ; 3次核酸は1つ以上の検出可能標識物を含む。1

つの又は各 2 次核酸は、1 次核酸分子にハイブリダイズせず、かつ 3 次核酸分子にハイブリダイズしない領域（「追加のハンドル」）を含み得る。この領域は、長さが 4 以上（例えば、約 6 ～ 約 40、例えば約 8、10、12、及び 14）のヌクレオチドであってもよい。1 次核酸分子にハイブリダイズせず、かつ 3 次核酸分子にハイブリダイズしない領域は、1 次核酸分子に連結された相補的核酸分子のヌクレオチド配列を含み得る。この領域は、1 次核酸にハイブリダイズするその末端に対して遠位である 2 次核酸の末端近くに位置し得る。相補的核酸のヌクレオチド配列を含む「追加のハンドル」を有することにより、レポーター複合体が配列決定プローブに結合する可能性及び速度が大幅に増加する。本発明の任意の実施態様又は態様において、レポーター複合体が「追加のハンドル」を含む場合、レポーター複合体は、レポーター複合体の相補的核酸を介して又は「追加のハンドル」を介して、配列決定プローブにハイブリダイズすることができる。例えば、「第 1 付着領域・・・第 1 のレポーター複合体の第 1 の相補的核酸分子への結合」という語句は、その明白な意味に従って理解され、「第 1 付着領域・・・第 1 のレポーター複合施設の『追加のハンドル』への結合」を意味するとも理解されるであろう。

【0009】

実施態様において、用語「バーコードドメイン」及び「合成主鎖」は同義語である。

【0010】

本明細書において、本発明の配列決定プローブを用いて核酸配列を決定する方法が提供される。この方法は、(1) 本発明の少なくとも 1 つの配列決定プローブを、基板に固定化された（例えば、1、2、3、4、5、6、7、8、9、10 以上の位置で）標的核酸にハイブリダイズさせる工程と；(2) 第 1 付着領域に、検出可能標識物（例えば蛍光標識物）を有する第 1 の相補的核酸分子（RNA 又は DNA）を、又は検出可能標識物（例えば蛍光標識物）を含む第 1 のレポーター複合体の第 1 の相補的核酸分子を結合させる工程と；(3) 検出可能標識物を検出する工程、及び(4) 固定化された標的核酸中の第 1 ヌクレオチドの位置及び本体を特定する工程を含む。場合により、固定化された標的核酸は、プローブにより結合される前に伸長される。この方法は、更に(5) 第 1 付着領域（1 つ又は 2 つのフランキング 1 本鎖ポリヌクレオチドを含むか又は含まない）に、検出可能標識物を欠く第 1 のハイブリダイズ性核酸分子を接触させ、こうして、検出可能標識物を有する第 1 の相補的核酸分子、又は検出可能標識物を含む第 1 のレポーター複合体の第 1 の相補的核酸分子の結合をはずし、そして、少なくとも、第 1 付着領域に、検出可能標識物を欠く第 1 のハイブリダイズ性核酸を結合させる工程と；(6) 第 2 付着領域に、検出可能標識物を有する第 2 の相補的核酸分子、又は検出可能標識物を含む第 2 のレポーター複合体の相補的核酸分子を結合させる工程と；(7) 検出可能標識物を検出する工程と；(8) 固定化された標的核酸中の、第 2 ヌクレオチドの位置及び本体を特定する工程とを含む。固定化された標的核酸中に有り、標的結合ドメインに対応する各ヌクレオチドが特定されるまで、工程(5)～(8)が繰り返される。工程(5)及び(6)は、同時に又は連続して行うことができる。各（例えば、第 1、第 2、第 3、第 4、第 5、第 6、第 7、第 8、第 9、第 10、又はそれ以上の）相補的核酸分子（検出可能標識物を有する、又はレポーター複合体の一部である）は、検出可能標識物を欠くその対応する（すなわち、第 1、第 2、第 3、第 4、第 5、第 6、第 7、第 8、第 9、第 10、又はそれ以上の）ハイブリダイズ性核酸分子と同じ核酸配列を有する。標的核酸は、標的核酸の第 1 の位置及び/又は第 2 の位置を、第 1 及び/又は第 2 の捕捉プローブと結合させることにより、基板に固定化される。各捕捉プローブは、基板に選択的に結合する親和性タグを含む。第 1 及び/又は第 2 の位置は、標的核酸の末端又はその近傍であってもよい。基板は、当技術分野で公知の任意の固体支持体、例えば被覆スライド及びマイクロ流体デバイス（例えば、ストレプトアビジンで被覆されたもの）であり得る。標的核酸の末端から離れている他の位置は、基板に選択的に結合され得る。核酸は、標的核酸を伸長させるのに十分な力（例えば、重力、流体力、電磁力、流れ伸長力、後退メニスカス技術、及びそれらの組み合わせ）を加えることにより伸長され得る。

【0011】

本明細書では、本発明のプローブの１つの集団又は本発明の複数の集団を用いて、核酸配列を決定する方法が提供される。この方法は、（１）（本発明の）配列決定プローブの第１集団を、基板に固定化された標的核酸にハイブリダイズさせる（第１集団の各配列決定プローブは、ほぼ同じ条件下で、例えばカオトロピック剤、温度、塩濃度、pH、及び流体力学的な力のレベルで、固定化標的核酸から脱ハイブリダイズされる）工程と；（２）各々が検出可能標識物を有する複数の第１の相補的核酸分子を、又は各複合体が検出可能標識物を含む複数の第１のレポーター複合体の複数の第１の相補的核酸分子を、第１集団中の各配列決定プローブ中の第１付着領域に結合させる工程と；（３）検出可能標識物を検出する工程と；（４）第１集団中の配列決定プローブによりハイブリダイズされた固定化標的核酸中の、複数の第１ヌクレオチドの位置及び本体を特定する工程と；（５）第１集団の各配列決定プローブの各第１付着領域に、検出可能標識物を欠く複数の第１のハイブリダイズ性核酸分子を接触させ、こうして、検出可能標識物を有するか又はレポーター複合体の第１の相補的核酸分子の結合をはずし、そして各第１付着領域に、検出可能標識物を欠く第１のハイブリダイズ性核酸分子を結合させる工程と；（６）各々が検出可能標識物を有する複数の第２の相補的核酸分子を、又は各複合体が検出可能標識物を含む複数の第２のレポーター複合体の複数の第２の相補的核酸分子を、第１集団中の各配列決定プローブ中の第２付着領域に結合させる工程と；（７）検出可能標識物を検出する工程と；（８）第１集団中の配列決定プローブによりハイブリダイズされる固定化された標的核酸中の複数の第２ヌクレオチドの位置及び本体を特定する工程とを含む。工程（９）では、固定化された標的核酸中に有り、第１集団中の各配列決定プローブの標的結合ドメインに対応する、各ヌクレオチドが特定されるまで、工程（５）～（８）が繰り返される。工程（５）及び（６）は、同時に又は連続して行うことができる。こうして、配列決定プローブの第１集団中の配列決定プローブの標的結合ドメインによりハイブリダイズされた固定化された標的核酸の領域について、ヌクレオチドの直線的順序が特定される。

【００１２】

実施態様において、プローブの複数の集団（すなわち２つ以上の集団）が使用される場合、この方法は更に、（１０）核酸から、第１集団の各配列決定プローブを脱ハイブリダイズする工程と；（１１）第１集団の各脱ハイブリダイズした配列決定プローブを除去する工程と；（１２）本発明の配列決定プローブの少なくとも第２集団をハイブリダイズする工程とを含み、ここで、第２集団中の各配列決定プローブは、固定化された標的核酸から、ほぼ同じ条件下で脱ハイブリダイズし、固定化された標的核酸から、第１集団中の配列決定プローブとは異なる条件下で脱ハイブリダイズする工程と；（１３）各々が検出可能標識物を有する複数の第１の相補的核酸分子を、又は各複合体が検出可能標識物を含む複数の第１のレポーター複合体の複数の第１の相補的核酸分子を、第２集団の各配列決定プローブ中の第１付着領域に結合させる工程と；（１４）検出可能標識物を検出する工程と；（１５）第２集団中の配列決定プローブによりハイブリダイズされた固定化標的核酸中の、複数の第１ヌクレオチドの位置及び本体を特定する工程と；（１６）第２集団の各配列決定プローブの各第１付着領域に、検出可能標識物を欠く複数の第１のハイブリダイズ性核酸分子を接触させ、こうして第１の相補的核酸分子の結合をはずし（検出可能標識物を有するか、又はレポーター複合体から）、そして各第１付着領域に、検出可能標識物を欠く第１のハイブリダイズ性核酸分子を結合させる工程と；（１７）各々が検出可能標識物を有する複数の第２の相補的核酸分子を、又は各複合体が検出可能標識物を含む複数の第２のレポーター複合体の複数の第２の相補的核酸分子を、第２集団の各配列決定プローブ中の第２付着領域に結合させる工程と；（１８）検出可能標識物を検出する工程と；（１９）第２集団中の配列決定プローブによりハイブリダイズされた固定化された標的核酸中の、複数の第２ヌクレオチドの位置及び本体を特定する工程と；及び（２０）配列決定プローブの第２集団中の配列決定プローブの標的結合ドメインによりハイブリダイズされた固定化された標的核酸の領域について、ヌクレオチドの直線的順序が特定されるまで、工程（１６）～（１９）を繰り返す工程とを含む。工程（１６）と（１７）は、同時に又は連続して行うことができる。

10

20

30

40

50

【 0 0 1 3 】

第 2 集団中の各配列決定プローブは、第 1 集団中の配列決定プローブが標的核酸から脱ハイブリダイズする平均条件とは異なる条件（例えば、より高い温度、より高いレベルのカオトロピック剤、より高い塩濃度、より速い流速、及び異なる pH）で、固定化された標的核酸から脱ハイブリダイズし得る。

【 0 0 1 4 】

しかし、3 つ以上のプローブ集団が使用される場合、2 つの連続した集団中のプローブは異なる条件で脱ハイブリダイズし、非連続集団のプローブが同様の条件で脱ハイブリダイズし得る。一例として、第 1 集団と第 3 集団中のプローブが、同様の条件下で脱ハイブリダイズし得る。ある実施態様において、プローブの連続集団は、より厳しくなる条件（例えば、より高いレベルのカオトロピック剤、塩濃度、及び温度）で脱ハイブリダイズした。温度を例として使用すると、マイクロ流体デバイスでは、プローブの第 1 集団は第 1 の温度ではハイブリダイズしたままであり得るが、第 1 の温度よりも高い第 2 の温度では脱ハイブリダイズし得る。プローブの第 2 集団は第 2 の温度ではハイブリダイズしたままであり得るが、第 2 の温度よりも高い第 3 の温度では脱ハイブリダイズし得る。この例では、最初のプローブ集団のための標的核酸上を流れる溶液（本方法により必要とされる試薬を含む）は、後のプローブ集団のための標的核酸上を流れる溶液よりも低い温度である。

10

【 0 0 1 5 】

ある実施態様において、プローブの集団が使用された後、そのプローブの集団は標的核酸から脱ハイブリダイズされ、同一のプローブ集団の新しいアリコートが使用される。例えば、プローブの第 1 集団がハイブリダイズされ、検出され、そして脱ハイブリダイズされた後、プローブの第 1 集団の次のアリコートがハイブリダイズされる。あるいは、一例として、プローブの第 1 集団は脱ハイブリダイズされ、プローブの第 2 の集団で置き換えることができる。第 2 の集団が検出されて脱ハイブリダイズされると、プローブの第 1 集団の次のアリコートが標的核酸にハイブリダイズされる。従って、次の集団のプローブは、以前に配列決定された標的核酸の領域にハイブリダイズし得る（こうして複製及び／又は確認配列情報を得る）か、又は次の集団のプローブは、以前に配列決定されていなかった標的核酸の領域にハイブリダイズし得る（こうして新しい配列情報を得る）。従って、先読みが（何らかの理由で）不十分であった場合に、及び／又は配列決定読取りから生じる整列の精度を改善する場合に、プローブの集団を再アリコートすることができる。

20

30

【 0 0 1 6 】

同様の条件下でハイブリダイズ及び脱ハイブリダイズするプローブは、同様の長さのそれらの標的結合ドメイン、GC 含量、又は反復塩基の頻度、及びそれらの組み合わせを有し得る。オリゴヌクレオチドの T_m と長さとの関係は、例えば、Sugimoto et al., Biochemistry, 34, 11211-6 に教示されている。

【 0 0 1 7 】

3 つ以上のプローブ集団が使用される場合、配列決定プローブの第 1 及び第 2 の集団について記載されたような工程が、さらなるプローブ集団（例えば、10 ~ 100 ~ 1000 集団）について繰り返される。使用されるプローブの集団の数は、特に限定されないが、標的核酸のサイズ、各集団中のユニークなプローブの数、所望の配列決定プローブ間の重複の程度、及び関心領域へのプローブの濃縮を含む種々の要因に依存する。

40

【 0 0 1 8 】

プローブの集団は、標的核酸の特定の関心領域、例えば突然変異（例えば点突然変異）又は SNP 対立遺伝子を含む領域、に向けられた追加の配列決定プローブを含むことができる。プローブの集団は、標的核酸中の関心の小さい特定の領域に向けられたより少ない配列決定プローブを含むことができる。

【 0 0 1 9 】

配列決定プローブの集団は、配列決定プローブの個別のより小さいプールに区画化することができる。区画化は、配列決定プローブ中の標的結合ドメインの予測される融解温度

50

、及び／又は配列決定プローブ中の標的結合ドメインの配列モチーフに基づくことができる。区画化は、経験則から導かれたルールに基づいてもよい。配列決定プローブの異なるプールは、例えば温度、塩濃度、及び／又は緩衝液含有量などの異なる反応条件を用いて、標的核酸と反応させることができる。区画化は、均一な適用範囲で標的核酸をカバーするために実行し得る。区画化は、既知の適用範囲プロファイルを有する標的核酸をカバーするために実行し得る。

【0020】

配列決定プローブの集団中の標的結合ドメインの長さを短縮させて、標的核酸の特定の領域中のプローブの適用範囲を増加させることができる。配列決定プローブの集団中の標的結合ドメインの長さを延長させて、例えば配列決定装置の分解能限界を超えるように、標的核酸の特定の領域中のプローブの適用範囲を減少させることができる。

10

【0021】

代替的又は追加的に、集団中の配列決定プローブの濃度を高めて、標的核酸の特定の領域におけるプローブの適用範囲を増加させることができる。例えば配列決定装置の分解能限界を超えるように、配列決定プローブの濃度を低下させて、標的核酸の特定の領域中におけるプローブの適用範囲を減少させることができる。

【0022】

核酸配列を決定するための方法は、固定化された標的核酸の各領域についてヌクレオチドの各特定された直線的順序を組み立てる工程を更に含み、こうして、固定化された標的核酸の配列を特定することができる。組立工程は、識別された各ヌクレオチドの線状順序を整列させるようにマイクロプロセッサに指示する、実行可能なプログラムが記憶された非一時的コンピュータ可読記憶媒体を使用し、こうして、核酸の配列が得られる。組み立ては「リアルタイム」で、すなわち全てのデータが収集された後ではなく、配列決定プローブからデータが収集されている間に行われる。

20

【0023】

標的核酸、すなわち配列決定される標的核酸は、約4～1,000,000ヌクレオチドの長さであり得る。標的は、いずれの長さも1,000,000ヌクレオチドを超える、完全な無傷の染色体又はその断片を含み得る。

【0024】

本明細書では、本発明の方法を実施するための装置が提供される。

30

【0025】

本明細書において、本発明の配列決定プローブを含むキット、及び本発明の方法を実施するためのキットが提供される。実施態様においてキットは、捕捉プローブを介して核酸を固定化することができる基板、本発明の複数の配列決定プローブ、少なくとも1つの捕捉プローブ、検出可能標識物を有する少なくとも1つの相補的核酸分子、検出可能標識物を欠く少なくとも1つの相補的核酸分子、及び使用説明書を含む。実施態様においてキットは、約又は少なくとも4096のユニークな配列決定プローブを含む。4096は、可能性のある各六量体の組み合わせ（すなわち、バーコードドメイン中にそれぞれ6つの付着領域を有するプローブについて）を含めるために必要な、ユニークなプローブの最小数である。ここでは、6つの位置について4つのヌクレオチド選択肢（ 4^6 ）があるため、「4096」が達成される。バーコードドメイン内に4つの付着領域を有する1組のプローブについて、256（すなわち 4^4 ）のユニークなプローブが必要になる。標的結合ドメイン中に8個のヌクレオチドを有するプローブのセットについては、 4^8 （すなわち65,536）のユニークなプローブが必要となる。標的結合ドメイン中に10個のヌクレオチドを有するプローブのセットについては、 4^{10} （すなわち1,048,576）のユニークなプローブが必要となる。

40

【0026】

実施態様において、キットは、検出可能標識物を有する約又は少なくとも24の別個の相補的核酸分子、及び検出可能標識物を欠く約又は少なくとも24の異なるハイブリダイズ性核酸分子を含む。相補的核酸は、非限定的な例として、配列番号1～24のうちの1

50

つの配列を有する付着領域に結合することができる。バーコードドメインに含まれ得るさらなる例示的配列は、配列番号42～配列番号81に列記される。実際、ヌクレオチド配列は限定されず、好ましくは、これは既知のヌクレオチド配列との実質的な相同性（例えば50%～99.9%）を欠く。これは、相補的核酸と標的核酸の望ましくないハイブリダイゼーションを回避するのに役立つ。

【0027】

上記態様及び実施態様のいずれも、任意の他の態様又は実施態様と組み合わせることができる。

【0028】

他に定義されない限り、本明細書で使用される全ての技術用語及び科学用語は、本発明が属する技術分野の当業者により一般的に理解されるのと同じ意味を有する。本明細書では、文脈が明確に別に指示しない限り、単数形は複数形も含む。例として、用語「a」、「an」、及び「the」は単数又は複数であると理解され、用語「又は」は包括的であると理解される。例として「ある要素」は1つ以上の要素を意味する。本明細書を通して「含む（comprising）」という単語又は「含む（comprises）」又は「含む（comprising）」などの変形は、記載された要素、整数、若しくは工程、又は要素、整数、若しくは工程の群を含むが、任意の他の要素、整数、若しくは工程、又は要素、整数、若しくは工程の群を排除しないことを意味すると理解される。約は、記載された値の10%、9%、8%、7%、6%、5%、4%、3%、2%、1%、0.5%、0.1%、0.05%、又は0.01%以内であると理解することができる。文脈から特に明白でない限り、本明細書で提供される全ての数値は、用語「約」により修飾される。

【0029】

本明細書に記載されたものと同等の方法及び材料は、本発明の実施又は試験に使用することができるが、適切な方法及び材料は以下に記載される。本明細書で言及される全ての刊行物、特許出願、特許、及び他の参考文献は、その全体が参照により本明細書に組み込まれる。本明細書に引用される参考文献は、本願発明の先行技術であると認めるものではない。矛盾する場合、定義を含む本明細書が支配するものとする。更に、材料、方法、及び実施例は例示的なものに過ぎず、限定することを意図するものではない。本発明の他の特徴及び利点は、以下の詳細な説明及び請求項から明らかであろう。

【0030】

特許又は特許出願ファイルは、カラーで示される少なくとも1つの図面を含む。カラー図面付きの本特許又は特許出願のコピーは、要請及び必要な手数料の納付に応じて、特許庁により提供される。

【0031】

上記及びさらなる特徴は、添付の図面と併せて以下の詳細な説明から明らかになるであろう。

【図面の簡単な説明】

【0032】

【図1】本発明の例示的な配列決定プローブの模式図を示す。

【図2】本発明の例示的な配列決定プローブの模式図を示す。

【図3】本発明の例示的な配列決定プローブの模式図を示す。

【図4】本発明の例示的な配列決定プローブの模式図を示す。

【図5】本発明の例示的な配列決定プローブの模式図を示す。

【図6A】本発明の配列決定プローブの変種を示す概略図である。

【図6B】本発明の配列決定プローブの変種を示す概略図である。

【図6C】本発明の配列決定プローブの変種を示す概略図である。

【図6D】本発明の配列決定プローブの変種を示す概略図である。

【図7】本発明の配列決定プローブの標的結合ドメインの模式図を示す。ドメインは、ユニバーサル塩基を有する0、2、又は4個のヌクレオチドを含む。

【図8A】本発明の配列決定方法の工程を示す。

- 【図 8 B】本発明の配列決定方法の工程を示す。
- 【図 8 C】本発明の配列決定方法の工程を示す。
- 【図 8 D】本発明の配列決定方法の工程を示す。
- 【図 8 E】本発明の配列決定方法の工程を示す。
- 【図 9 A】本発明の配列決定方法の初期工程を示す。
- 【図 9 B】検出可能標識物を含むレポーター複合体の概略図を示す。
- 【図 9 C】それぞれ検出可能標識物を含む複数のレポーター複合体を示す。
- 【図 9 D】図 9 A から始まる配列決定方法のさらなる工程を示す。
- 【図 9 E】図 9 A から始まる配列決定方法のさらなる工程を示す。
- 【図 9 F】図 9 A から始まる配列決定方法のさらなる工程を示す。 10
- 【図 9 G】図 9 A から始まる配列決定方法のさらなる工程を示す。
- 【図 10】図 9 D 及び図 9 E に示される工程の代替的な説明図、及びそこから得られる例示的なデータを示す。示された配列決定プローブの断片は、配列番号 8 2 の配列を有する。
- 【図 11】図 10 に示す方法の変形例を示す。同様に示される配列決定プローブの断片は、配列番号 8 2 の配列を有する。
- 【図 12】本発明の方法を示す。
- 【図 13】本発明の配列決定方法において必要とされる工程を、他の配列決定方法で必要とされる工程と比較する。
- 【図 14】本発明により得られる性能測定値の例を示す。 20
- 【図 15】本発明により得られる性能測定値の例を示す。
- 【図 16】本発明及び様々な他の配列決定方法 / 装置についての、配列決定速度、読み取り回数、及び臨床の有用性を比較する。
- 【図 17】本発明の配列決定方法の低い生データエラー率を示す。示された鋳型配列は、配列番号 8 3 の配列を有する。
- 【図 18】本発明から得られる配列決定データを他の配列決定方法と比較する。
- 【図 19】本発明の配列決定方法の単一塩基特異性を示す。示された鋳型及びプローブ配列（上から下へ）は、配列番号 8 4 ~ 配列番号 8 8 の配列を有する。
- 【図 20 A】本発明のレポーター複合体の種々のデザインを示す。
- 【図 20 B】図 20 A に示されるレポーター複合体から得られた蛍光カウントを示す。 30
- 【図 20 C】本発明のレポーター複合体を構築するための例示的な配合法を示す。
- 【図 21 A】「追加のハンドル」を含むレポーター複合体のデザインを示す。
- 【図 21 B】「追加のハンドル」を有するレポーター複合体から得られた蛍光カウントを示す。
- 【図 22 A】本発明のレポーター複合体の 2 つの例示的なデザインのハイブリダイゼーション速度を示す。
- 【図 22 B】本発明のレポーター複合体の 2 つの例示的なデザインのハイブリダイゼーション速度を示す。
- 【図 23】図 8 ~ 図 12 に示す方法とは異なる方法で使用される本発明の配列決定プローブの概略図を示す。 40
- 【図 24】本発明において有用な消耗型配列決定カードの概略図を示す。
- 【図 25】実施例 3 に記載の 10 量体のミスマッチ検出を示す。示されたヌクレオチド（上から下へ）は、配列番号 8 9 ~ 配列番号 9 9 の配列を有する。
- 【図 26】実施例 3 に記載されるように、標的結合ドメインのサイズに依存するハイブリダイゼーション能力を示す。非常に高いレポーター濃度のためにバックグラウンドは高く、事前の精製はなかった。示されたヌクレオチド（上から下へ）は、配列番号 100 ~ 配列番号 104 の配列を有する。
- 【図 27】単一スポット対完全長レポーターの比較を示す。単一スポットの結果は、ハイブリダイゼーションの速度が全長バーコード（条件 100 nM 標的、30 分のハイブリダイゼーション）よりも 1000 倍大きいことを示す。 50

【発明を実施するための形態】

【0033】

本発明は、長い読み取り長さ及び低いエラー率を有する、酵素不要、増幅不要、及びライブラリー不要の核酸配列決定法を提供する、配列決定プローブ、方法、キット及び装置を提供する。

配列決定プローブ

【0034】

本発明は、標的結合ドメイン及びバーコードドメインを含む配列決定プローブに関する。本発明の配列決定プローブの非限定的な例を図1～6に示す。

【0035】

図1は、本発明の配列決定プローブの概略図を示す。この例示的な配列決定プローブは、6つのヌクレオチドの標的結合ドメインを有し、その各々は、バーコードドメイン（これは1つ以上の付着領域を含む）中の位置に対応する。第1付着領域に注目する。これは、標的結合ドメイン中の第1ヌクレオチドにより結合された標的核酸のヌクレオチドに対応する。バーコードドメインの3番目の位置に注目する。2つの付着領域を含む第5の位置に注目する。バーコードドメイン上の各位置は、複数の添付領域を持つことができる。たとえば、1つの位置は1～50の付着領域を有することができる。バーコードドメイン内の特定の位置は、他の位置よりも多くの付着領域を有することができる（ここでは位置1～4及び6に対して位置5に示されている）。代わりに、バーコードドメイン内の各位置は、同じ数の付着領域を有する（例えば、図2、3、5、及び6参照）。図示されていないが、各付着領域は、相補的核酸分子（RNA又はDNA）に可逆的に結合することができる核酸配列の少なくとも1つ（すなわち1～50、例えば10～30）のコピーを含む。図1では、付着領域は、バーコードドメインを構成する線状ポリヌクレオチド分子と一体化している。

【0036】

図2は、本発明の配列決定プローブの概略図を示す。この例示的な配列決定プローブは、6つのヌクレオチドの標的結合ドメインを有し、その各々はバーコードドメイン中の付着領域に対応する。第1付着領域に注目する。これは、標的結合ドメイン中の第1ヌクレオチドにより結合された標的核酸のヌクレオチドに対応する。バーコードドメインの一部と2つの第4付着領域とを含むバーコードドメイン上の第4の位置を囲んである。2つの6番目の付着領域に注目する。ここで、各位置は2つの付着領域を有するが、バーコードドメイン上の各位置は、1つの付着領域又は複数の付着領域、例えば2～50の付着領域を有することができる。図示されていないが、各付着領域は、相補的核酸分子（RNA又はDNA）に可逆的に結合することができる核酸配列の少なくとも1つ（すなわち1～50、例えば10～30）のコピーを含む。図2においてバーコードドメインは、付着領域が連結された線状ポリヌクレオチド分子であり；付着領域はポリヌクレオチド分子と一体化していない。

【0037】

図3は、本発明の配列決定プローブの別の概略図を示す。この例示的な配列決定プローブは、4つのヌクレオチドの標的結合ドメインを有し、これらの4つのヌクレオチドは、バーコードドメインの4つの位置に対応する。各位置は、3つの連結された付着領域で示されている。

【0038】

図4は、本発明の配列決定プローブの更に別の概略図を示す。この例示的な配列決定プローブは、10ヌクレオチドの標的結合ドメインを有する。しかし、最初の6個のヌクレオチドのみが、バーコードドメイン中の6つの位置に対応する。第7～第10ヌクレオチド（「n1～n4」で示される）は、標的結合ドメインの長さを延長するために付加され、こうして、プローブが標的核酸にハイブリダイズし、ハイブリダイズしたままである可能性に影響を及ぼす。実施態様において、「n」個のヌクレオチドは、バーコードドメイン中の位置に対応するヌクレオチドの前に存在し得る。実施態様において、「n」個のヌ

10

20

30

40

50

クレオチドは、バーコードドメイン中の位置に対応するヌクレオチドの後に存在し得る。図4には4つの「n」ヌクレオチドが示されているが、標的結合ドメインは、5つ以上の「n」ヌクレオチドを含み得る。「n」個のヌクレオチドは、4つの標準塩基のいずれかと塩基対合することができるユニバーサル塩基〔例えば、イノシン、2'-デオキシイノシン（ヒボキサンチンデオキシヌクレオチド）誘導体、ニトロインドール、ニトロアゾール類似体、及び疎水性芳香族非水素結合塩基〕を有することができる。

【0039】

本発明の別の配列決定プローブを図5に示す。ここで「n」個のヌクレオチドは、バーコードドメイン中の位置に対応するヌクレオチドの前と後にある。示された例示的な配列決定プローブは、10ヌクレオチドの標的結合ドメインを有する。しかし、標的結合ドメイン中の3番目～8番目のヌクレオチドのみが、バーコードドメイン中の6つの位置（第1～第6）に対応する。第1、第2、第9、及び第10ヌクレオチド（「n1～n4」で示される）は、標的結合ドメインの長さを延長するために付加されている。図5には、4つの「n」ヌクレオチドが示されている。しかし、標的結合ドメインは、4つより多いか又は少ない「n」ヌクレオチドを含み得る。

【0040】

図6A～6Dは、図1の配列決定プローブの変種を示す。図6Aにおいて、標的結合ドメイン中のヌクレオチドの直線的順序及びバーコードドメイン中の付着領域の直線的順序は、左から右へ進む（図に関し）。図6Bにおいて、標的結合ドメイン中のヌクレオチドの直線的順序及びバーコードドメイン中の付着領域の直線的順序は、右から左へ進む（図に關して）。図6Cにおいて、標的結合ドメイン中のヌクレオチドの直線的順序は、バーコードドメイン中の付着領域の直線的順序に対して逆転している。本発明のいずれのプローブにおいても、標的結合ドメイン中の各ヌクレオチドがバーコードドメイン内の付着ドメインに対応するように、プローブが設計される限りは、標的結合ドメイン中のヌクレオチドとバーコードドメイン内の付着領域のヌクレオチドの厳密な秩序が欠如しているかも知れない。厳密な順序の欠如は図6Dに示されている。本発明のいかなるプローブ（例えば、図1～5に例示されるプローブ）も、図6に示されるようなヌクレオチド及び付着領域の順序を有し得る。

【0041】

標的結合ドメインは、少なくとも4つのヌクレオチド、例えば、少なくとも4、5、6、7、8、9、10、11、12、又はそれ以上のヌクレオチドを有する。好ましい標的結合ドメインはポリヌクレオチドである。標的結合ドメインは、標的核酸に結合することができる。

【0042】

プローブは、合成主鎖に機能的に連結している標的結合ドメインの複数のコピーを含み得る。

【0043】

プローブは、ハイブリダイゼーション及び/又は脱ハイブリダイゼーションの確率及びこれらが生じる速度を制御するように設計することができる。一般に、プローブの T_m が低いほど、プローブは標的核酸に/標的核酸から脱ハイブリダイズすることが、より速かつより高率になる。従って、より低い T_m プローブの使用は、標的核酸に結合したプローブの数を減少させる。

【0044】

標的結合ドメインの長さは、一部は、プローブが標的核酸にハイブリダイズしハイブリダイズしたままである確率に影響を与える。一般に、標的結合ドメインがより長い（ヌクレオチドの数が多い）ほど、相補的配列が標的ヌクレオチド中に存在する確率は低い。反対に、標的結合ドメインが短いほど、相補的配列が標的ヌクレオチド中に存在する確率が高い。例えば、4量体配列が標的核酸中に位置する確率は $1/256$ であり、これに対して6量体配列が標的核酸中に位置する確率は $1/4096$ である。結果的に、より短いプローブのコレクションは、より長いプローブのコレクションと比較した場合、核酸の所定

のストレッチについて、より多くの位置で結合する確率が高い。

【0045】

図7は、10量体の標的結合ドメインを示す。いくつかの実施態様において、標的結合ドメインは、4つの標準ヌクレオチド(A、G、C、及びT)のいずれかと塩基対合する4つのユニバーサル塩基(「U_b」として識別される)を含む。実施態様において、標的結合ドメインは、1~6(例えば2及び4)のユニバーサル塩基を含む。標的結合ドメインは、ユニバーサルヌクレオチドを含まない場合がある。図7は、標的結合ドメイン中の6個の特異的ヌクレオチドを有するプローブの「完全な」集団が4096個のユニークなプローブを必要とし、10個の特異的ヌクレオチドを有するプローブの「完全な」集団が約100万個のユニークなプローブを必要とすることを示す。

10

【0046】

状況に応じて、例えば変異又はSNP対立遺伝子を検出する時、核酸の所定のストレッチにおける読み取りの数を増加させ、こうして、標的核酸又は標的核酸の一部の適用範囲、特に関係する部分の適用範囲を濃縮するために、より短い標的結合ドメインを有するプローブを有することが好ましい。

【0047】

しかしながら、領域内のあまりにも多くのプローブがそれらの検出可能標識物の重複を引き起こし、こうして2つの近傍のプローブの分解能を妨げることがあるため、標的核酸に結合するプローブの数を少なくすることが好ましい。これは以下のように説明される。1つのヌクレオチドは0.34nmの長さであり、配列決定装置の横方向(x-y)空間分解能が約200nmであると仮定すると、配列決定装置の分解能限界は約588塩基対(すなわち、1ヌクレオチド/0.34nm×200nm)である。すなわち、2つのプローブが互いに約588塩基対以内にある場合、上記配列決定装置は、標的核酸にハイブリダイズした2つのプローブからの信号を分解することができ、ない。従って、配列決定装置の分解能に応じて、2つのプローブは、それらの検出可能標識物を別個の「スポット」として分離する前に、約600bpの間隔を置いて配置する必要がある。従って、最適な間隔では、標的核酸の600bpあたり1つのプローブが存在するはずである。標的核酸の分解可能な領域内のハイブリダイズするプローブの数を追跡し、制限し、おそらくは解析し、それに応じてプローブ集団を設計するために、様々なソフトウェアアプローチ(例えば、蛍光強度値と波長依存比の利用)を用いることができる。更に、より離散した信号を提供する検出可能標識物(例えば蛍光標識物)を選択することができる。更に、文献中の方法(例えば、Small and Parthasarathy: "Superresolution localization methods." Annu. Rev. Phys. Chem., 2014; 65:107-25)は、構造照明及び様々な超解像アプローチを記載しており、これらは、配列決定顕微鏡の分解能の上限を数10nmまで低下させる、より高分解能の配列決定装置を使用することにより、より短い標的結合ドメインを有するプローブの使用が可能になる。

20

30

【0048】

上述したように、プローブのT_mを設計することは、標的核酸にハイブリダイズするプローブの数に影響を及ぼす可能性がある。代替的に又は追加的に、集団中の配列決定プローブの濃度を増加させて、標的核酸の特定の領域中のプローブの適用範囲を増加させることができる。例えば、配列決定装置の分解能限界を超えるように、配列決定プローブの濃度を低下させて、標的核酸の特定の領域におけるプローブの適用範囲を減少させることができる。

40

【0049】

「標的核酸」という用語は、本発明のプローブ、方法、及び装置によりその配列が決定される核酸分子(DNA、RNA、又はPNA)を意味するものとする。一般に、用語「標的核酸」、「核酸分子」、「核酸配列」、「核酸」、「核酸断片」、「オリゴヌクレオチド」、及び「ポリヌクレオチド」は互換的に使用され、特に限定されないが、種々の長さを有し得るヌクレオチド(デオキシリボヌクレオチド、又はリボヌクレオチド、又はそれらの類似体)のポリマー形態を含むことが企図される。核酸の非限定的な例には、遺伝

50

子、遺伝子断片、エキソン、イントロン、遺伝子間DNA（特に限定されないが、ヘテロクロマチン性DNAを含む）、メッセンジャーRNA（mRNA）、トランスファーRNA、リボソームRNA、リボザイム、低分子干渉RNA（siRNA）、非コードRNA（ncRNA）、cDNA、組換えポリヌクレオチド、分枝ポリヌクレオチド、プラスミド、ベクター、配列の単離されたDNA、配列の単離されたRNA、核酸プローブ、及びプライマーが含まれる。

【0050】

本方法は、試料、例えば生物由来の試料から得られた核酸分子を、好ましくは変換（又は増幅）工程なしで直接配列決定する。一例として、RNA主体の配列決定のために、本方法は、配列が得られる前にRNA分子のDNA分子への変換（すなわちcDNAの合成による）を必要としない。増幅又は変換が必要でないため、本発明で配列決定された核酸は、核酸が試料中にあるとき又は試料から得られたときに、核酸中に存在する任意のユニークな塩基及び／又はエピジェネティックマーカーを保持するであろう。そのような固有の塩基及び／又はエピジェネティックマーカーは、当技術分野で公知の配列決定法の最中に失われる。

10

【0051】

標的核酸は、任意の試料又は核酸源、例えば、任意の細胞、組織、又は生物から、インビトロで、化学合成機などから得ることができる。標的核酸は、任意の当該技術分野で認識された方法により得ることができる。実施態様において、核酸は臨床的被験体の血液試料から得られる。核酸は、当該分野で周知の方法及びキットを使用して、供給源又は試料から抽出、単離、又は精製することができる。

20

【0052】

標的核酸を含む核酸分子は、当技術分野で公知の任意の手段により断片化することができる。好ましくは断片化は、酵素的手段又は機械的手段により実施される。機械的手段は、超音波処理又は物理的剪断であってもよい。酵素的手段は、ヌクレアーゼ（例えば、デオキシリボヌクレアーゼI（DNase I））又は1つ以上の制限エンドヌクレアーゼによる消化により行われ得る。

【0053】

標的核酸を含む核酸分子がインタクトな染色体である場合、染色体の断片化を避けるための工程が必要である。

30

【0054】

標的核酸は、当技術分野で周知のように、修飾されたヌクレオチドを含む天然又は非天然のヌクレオチドを含むことができる。

【0055】

本発明のプローブは、約20nm～約50nmの全長（標的結合ドメイン、バーコードドメイン、及び任意のドメインを含む）を有し得る。プローブの主鎖は、約120ヌクレオチドを含むポリヌクレオチド分子であり得る。

【0056】

バーコードドメインは合成主鎖を含む。合成主鎖と標的結合ドメインは機能的に連結されており、例えばリンカーを介して共有結合又は結合している。合成主鎖は、多糖類、ポリヌクレオチド、ポリマー、プラスチック、繊維、ペプチド、ペプチド核酸、又はポリペプチドなどの任意の材料を含むことができる。好ましくは、合成主鎖は剛性である。実施態様において、主鎖は、6個のDNA2重らせんの「DNA折り紙」を含む（例えば、Lin et al, “Submicrometre geometrically encoded fluorescent barcodes self-assembled from DNA.” Nature Chemistry; 2012 Oct; 4(10): 832-9 参照）。バーコードは、DNA折り紙タイルで作ることができる（Jungmann et al, “Multiplexed 3D cellular super-resolution imaging with DNA-PAINT and Exchange-PAINT”, Nature Methods, Vol. 11, No. 3, 2014）。

40

【0057】

バーコードドメインは、複数の位置、例えば、1、2、3、4、5、6、7、8、9、

50

10、又はそれ以上の位置を含む。位置の数は、標的結合ドメイン中のヌクレオチドの数より少ないか、等しいか、又はそれ以上でもよい。主鎖ドメイン中の位置の数よりも多くのヌクレオチド、例えば1、2、3、4、5、6、7、8、9、10、又はそれ以上のヌクレオチドを、標的結合ドメイン中に含むことが好ましい。バーコード領域の長さは、上述したように、少なくとも4つの位置に十分なスペースがあれば、制限されない。

【0058】

バーコードドメインにおける各位置は、標的結合ドメイン中のヌクレオチドに対応し、従って標的核酸中のヌクレオチドに対応する。例として、バーコードドメインの第1の位置は標的結合ドメインの第1ヌクレオチドに対応し、バーコードドメインの第6の位置は標的結合ドメインの第6ヌクレオチドに対応する。

10

【0059】

バーコードドメイン内の各位置は、少なくとも1つの付着領域、例えば1～50個又はそれ以上の付着領域を含む。バーコードドメイン内の特定の位置は、他の位置よりも多くの付着領域を有することができる（例えば、第1の位置は3つの付着領域を有し、第2の位置は2つの付着位置を有することができる）。あるいは、バーコードドメイン内の各位置は同じ数の付着領域を有する。各付着領域は、相補的核酸分子（例えばDNA又はRNA）により可逆的に結合されることができる核酸配列の少なくとも1つ（すなわち、1～50、例えば10～30）のコピーを含む。例では、第1付着領域中の核酸配列は、標的結合ドメインの第1ヌクレオチドにより結合される標的核酸中の第1ヌクレオチドの位置及び本体を決定する。各付着領域は、付着領域が合成主鎖から分岐するように、合成主鎖内の修飾モノマー（例えば、修飾ヌクレオチド）に連結されてもよい。実施態様において、付着領域はポリヌクレオチド主鎖と一体化している。すなわち、主鎖は単一のポリヌクレオチドであり、付着領域は単一のポリヌクレオチド配列の一部である。実施態様において、用語「バーコードドメイン」及び「合成主鎖」は同義語である。

20

【0060】

付着領域中の核酸配列は、配列決定プローブの標的結合ドメイン中のヌクレオチドにより結合される標的核酸中のヌクレオチドの位置及び本体を特定する。プローブ中で各付着領域は、ユニークな全体配列を有する。実際、バーコードドメイン上の各位置は、アデニン、チミン/ウラシル、シトシン、及びグアニンの1つに特異的な、4つのヌクレオチドのうちの1つをコードする核酸配列を含む付着領域を有することができる。また、第1の位置の（及び例えばシトシンをコードする）付着領域は、第2の位置の（及び例えばシトシンをコードする）付着領域とは異なる核酸配列を含む。従って、チミンをコードする第1の位置の付着領域の核酸配列に対して、標的結合ドメインの第1ヌクレオチドに対応する標的核酸中のアデニンを特定する相補的核酸分子の結合は存在しない。また、第2の位置の付着領域に対して、標的結合ドメインの第1ヌクレオチドに対応する標的核酸中のアデニンを特定する相補的核酸分子の結合は存在しない。

30

【0061】

バーコードドメイン上の各位置は、1つ以上（最大50、好ましくは10～30）の付着領域を含むことができ、従って、各付着領域は、1つ以上（最大50、好ましくは10～30）の相補的核酸分子に結合することができる。例として、図1のプローブは、2つの付着領域を含む第5の位置を有し、図2のプローブは、6つの付着領域を有する第2の位置を有する。実施態様において、ある位置における付着領域の核酸配列は同一である。従って、それらの付着領域に結合する相補的核酸分子は同一である。別の実施態様において、ある位置における付着領域の核酸配列は同一ではない。従って、それらの付着領域に結合する相補的核酸分子は、同一ではなく、例えばそれぞれ異なる核酸配列及び/又は検出可能標識物を含む。従って、別の実施態様において、付着領域に付着した非同ー核酸分子（例えば、それらの検出可能標識物）の組み合わせは、標的核酸中のヌクレオチドを特定するためのコードを提供する。

40

【0062】

表1は、例示的目的のみで、そのバーコードドメイン内に最大6個の位置を有するプロ

50

ープを配列決定するための付着領域、及びそれに結合する相補的核酸上の検出可能標識物の例示的な配列を提供する。

【 0 0 6 3 】

【 表 1 】

標的結合ドメイン中のヌクレオチド／バーコードドメイン中の位置	ヌクレオチド	付着領域中の核酸配列 (5' → 3')	相補的核酸の検出可能標識物	配列番号
1	A	ATACATCTAG	GFP	1
1	G	GATCTACATA	RFP	2
1	C	TTAGGTAAG	CFP	3
1	U/T	TCTTCATTAC	YFP	4
2	A	ATGAATCTAC	GFP	5
2	G	TCAATGTATG	RFP	6
2	C	AATTGAGTAC	CFP	7
2	U/T	ATGTTAATGG	YFP	8
3	A	AATTAGGATG	GFP	9
3	G	ATAATGGATC	RFP	10
3	C	TAATAAGGTG	CFP	11
3	U/T	TAGTTAGAGC	YFP	12
4	A	ATAGAGAAGG	GFP	13
4	G	TTGATGATAC	RFP	14
4	C	ATAGTGATTC	CFP	15
4	U/T	TATAACGATG	YFP	16
5	A	TTAAGTTAG	GFP	17
5	G	ATACGTTATG	RFP	18
5	C	TGTACTATAG	CFP	19
5	U/T	TTAACAAGTG	YFP	20
6	A	AACTATGTAC	GFP	21
6	G	TAACATGAC	RFP	22
6	C	ACTAATGTTT	CFP	23
6	U/T	TCATTGAATG	YFP	24

【 0 0 6 4 】

表 1 から明らかなように、第 1 付着領域の核酸配列は、配列番号 1 ~ 配列番号 4 のうちの 1 つであってもよく、第 2 の付着の核酸配列は、配列番号 5 ~ 配列番号 8 の 1 つであってもよい。標的核酸中の第 1 ヌクレオチドがアデニンである場合、第 1 付着領域の核酸配列は配列番号 1 の配列を有し、標的核酸中の第 2 ヌクレオチドがアデニンである場合、第 2 付着領域の核酸配列は、配列番号 5 の配列を有するであろう。

【 0 0 6 5 】

実施態様において、相補的核酸分子は検出可能標識物により結合され得る。代替的实施態様において、相補的核酸は、検出可能標識物を含むレポーター複合体と結合される。

【 0 0 6 6 】

相補的核酸のヌクレオチド配列は限定されない。好ましくは、これは既知のヌクレオチド配列と実質的な相同性（例えば 50 % ~ 99.9 %）を欠く；これは、相補的核酸と標的核酸との望ましくないハイブリダイゼーションを回避するのに助ける。

【 0 0 6 7 】

本発明において有用なレポーター複合体の例を図 9 B に示す。この例では、相補的核酸は、第 1 の核酸分子に連結され、次にこれは複数の第 2 の核酸分子にハイブリダイズされ

、その各々は、1つ以上の検出可能標識物が結合した複数の第3の核酸分子にハイブリダイズしている。

【0068】

実施態様において、1次核酸分子は約90ヌクレオチドを含み得る。2次核酸分子は、約87ヌクレオチドを含み得る。3次核酸分子は、約15ヌクレオチドを含み得る。

【0069】

図9Cは、例示的なレポーター複合体の集団を示す。図9Cの左上のパネルには、プローブの付着領域1にハイブリダイズする4つの複合体が含まれる。プローブの標的結合ドメインのヌクレオチド位置1に存在し得る各可能なヌクレオチドについて、1つのタイプのレポーター複合体が存在する。ここで、本発明の配列方法を実施する際に、プローブのレポータードメインの1位が、「青色」の検出可能標識物を有するレポーター複合体により結合されている場合、標的結合ドメイン中の第1ヌクレオチドはアデニンと特定される。あるいは、位置1が、「緑色」の検出可能標識物を有するレポーター複合体により結合されている場合、標的結合ドメイン中の第1ヌクレオチドはチミンとして特定される。

【0070】

レポーター複合体は、様々なデザインのものであり得る。例えば1次核酸分子は、少なくとも1つ（例えば、1、2、3、4、5、6、7、8、9、10、又はそれ以上）の2次核酸分子にハイブリダイズすることができる。各2次核酸分子は、少なくとも1つ（例えば、1、2、3、4、5、6、7、8、9、10、又はそれ以上）の3次核酸分子にハイブリダイズされ得る。例示的なレポーター複合体を図20Aに示す。ここで「4×3」レポーター複合体は、4つの2次核酸分子にハイブリダイズした1つの1次核酸分子（これは相補的核酸分子に連結している）を有し、その各々は、3つの3次核酸分子（それぞれが検出可能標識物を含む）にハイブリダイズしている。この図では、複合体の各相補的核酸は12ヌクレオチド長（「12塩基」）である。しかし、相補的核酸の長さは限定されず、12ヌクレオチド未満又は12ヌクレオチド超であり得る。右下の複合体は、その相補的核酸とその1次核酸分子との間にスペーサー領域を含む。スペーサーは、20～40ヌクレオチドの長さであると特定される。しかし、スペーサーの長さは限定されず、20ヌクレオチドより短くてもよく、又は40ヌクレオチドより長くてもよい。

【0071】

図20Bは、図20Aに示される4つの例示的なレポーター複合体から得られた可変の平均（蛍光）カウントを示す。図20Bでは、10pMのビオチン化標的鋳型をストレプトアビジン被覆フローセル表面上に付着させ、10nMのレポーター複合体をフローセル上に流した。1分間のインキュベーション後、フローセルを洗浄し、フローセルを画像化し、蛍光性特徴をカウントした。

【0072】

実施態様において、レポーター複合体は「予め構築されている」。すなわち、複合体中の各ポリヌクレオチドは、複合体をプローブと接触させる前にハイブリダイズされる。5つの例示的なレポーター複合体を予め構築するための例示的な処方図20Cに示す。

【0073】

図21Aは、2次核酸分子が、3次核酸分子にハイブリダイズせず、1次核酸分子に対して遠位である「追加のハンドル」を有する代替レポーター複合体を示す。この図において、各「追加のハンドル」は12ヌクレオチド長（「12量体（12mer）」）である。しかし、その長さは限定されず、12ヌクレオチド未満又は12ヌクレオチド超であり得る。実施態様において、「追加のハンドル」はそれぞれ、相補的核酸のヌクレオチド配列を含む。従って、レポーター複合体が「追加のハンドル」を含む場合、レポーター複合体は、レポーター複合体の相補的核酸を介して、又は「追加のハンドル」を介して、配列決定プローブにハイブリダイズすることができる。従って、レポーター複合体が配列決定プローブに結合する確率が上昇する。「追加のハンドル」のデザインはまた、ハイブリダイゼーション速度を改善し得る。理論に拘束されるつもりはないが、「追加のハンドル」は本質的に、レポーター複合体の相補的核酸の有効濃度を増加させる。

【 0 0 7 4 】

図 2 1 B は、図 2 0 B について記載された手順を使用して、「追加のハンドル」を有する 5 つの例示的なレポーター複合体から得られた可変の平均（蛍光）カウントを示す。

【 0 0 7 5 】

図 2 2 A 及び 2 2 B は、2 つの例示的なレポーター複合体のハイブリダイゼーション速度及び蛍光強度を示す。約 5 分後までに、総数がプラトーに変わり、加えられたほとんどのレポーター複合体が利用可能な標的を見出したことが示される。

【 0 0 7 6 】

検出可能な部分、標識物、又はレポーターは、蛍光部分、比色部分などの検出可能な部分の直接的又は間接的な結合を含む様々な方法で、相補的核酸に又は 3 次核酸分子に結合させることができる。当業者は、核酸の標識に関する参考文献を参照することができる。蛍光部分の例は、特に限定されないが、黄色蛍光タンパク質（YFP）、緑色蛍光タンパク質（GFP）、シアン蛍光タンパク質（CFP）、赤色蛍光タンパク質（RFP）、ウンベリフェロン、フルオレセイン、フルオレセインイソチオシアネート、ローダミン、ジクロロトリアジニルアミンフルオレセイン、シアニン、塩化ダンシル、フィコシアニン、フィコエリトリンなどが挙げられる。蛍光標識物及びヌクレオチド及び／又はオリゴヌクレオチドへのそれらの結合は、多くの総説に記載されており、Haugland, Handbook of Fluorescent Probes and Research Chemicals, Ninth Edition (Molecular Probes, Inc., Eugene, 2002); Keller and Manak, DNA Probes, 2nd Edition (Stockton Press, New York, 1993); Eckstein, editor, Oligonucleotides and Analogues: A Practical Approach (IRL Press, Oxford, 1991); 及び Wetmur, Critical Reviews in Biochemistry and Molecular Biology, 26:227-259 (1991) がある。本発明に適用可能な具体的な方法論は、以下の文献の例に開示されている：米国特許第 4, 7 5 7, 1 4 1 号；第 5, 1 5 1, 5 0 7；及び第 5, 0 9 1, 5 1 9 号。ある態様において、標識された標的配列の標識物として 1 つ以上の蛍光色素が使用され、これは、例えば米国特許第 5, 1 8 8, 9 3 4 号（4, 7 - ジクロロフルオレセイン色素）；第 5, 3 6 6, 8 6 0 号（スペクトル的に分割可能なローダミン色素）；第 5, 8 4 7, 1 6 2 号（4, 7 - ジクロロローダミン色素）；第 4, 3 1 8, 8 4 6 号（エーテル置換フルオレセイン色素）；第 5, 8 0 0, 9 9 6 号（エネルギー移動色素）；Lee et al., 第 5, 0 6 6, 5 8 0 号（キサンチン色素）；第 5, 6 8 8, 6 4 8 号（エネルギー移動色素）等）に開示されている。標識はまた、以下の特許及び特許公報に開示されているように、量子ドットを用いて行うことができる：米国特許第 6, 3 2 2, 9 0 1 号；第 6, 5 7 6, 2 9 1 号；第 6, 4 2 3, 5 5 1 号；第 6, 2 5 1, 3 0 3 号；第 6, 3 1 9, 4 2 6 号；第 6, 4 2 6, 5 1 3 号；第 6, 4 4 4, 1 4 3 号；第 5, 9 9 0, 4 7 9 号；第 6, 2 0 7, 3 9 2 号；第 2 0 0 2 / 0 0 4 5 0 4 5 号；及び 2 0 0 3 / 0 0 1 7 2 6 4 号。本明細書で使用する用語「蛍光標識物」は、1 つ以上の分子の蛍光吸収及び／又は放出特性を通して情報を伝達する信号伝達部分を含む。このような蛍光特性には、蛍光強度、蛍光寿命、発光スペクトル特性、エネルギー移動などを含む。

【 0 0 7 7 】

ヌクレオチド及び／又はオリゴヌクレオチド配列に容易に組み込まれる市販の蛍光ヌクレオチド類似体としては、特に限定されないが、Cy 3 - dCTP、Cy 3 - dUTP、Cy 5 - dCTP、Cy 5 - dUTP (Amersham Biosciences, Piscataway, NJ)、フルオレセイン - 1 2 - dUTP、テトラメチルローダミン - 6 - dUTP、テキサスレッド（登録商標） - 5 - dUTP、カスケードブルー（登録商標） - 7 - dUTP、BODIPY TMFL - 1 4 - dUTP、BODIPY TMR - 1 4 - dUTP、BODIPY TMTR - 1 4 - dUTP、ローダミングリーン（登録商標）5 - dUTP、OREGON GREEN R（商標）4 8 8 - 5 - dUTP、テキサスレッド（登録商標） - 1 2 - dUTP、BODIPY TM 6 3 0 / 6 5 0 - 1 4 - dUTP、BODIPY TM 6 5 0 / 6 6 5 - 1 4 - dUTP、アレクサフルオール（登録商標）4 8 8 - 5 - dUTP、アレクサフルオール（登録商標）5 3 2 - 5 - dUTP、アレクサフルオール

10

20

30

40

50

(登録商標) 568 - 5 - d U T P、アレクサフルオール(登録商標) 594 - 5 - d U T P、アレクサフルオール(登録商標) 546 - 14 - d U T P、フルオレセイン - 12 - U T P、テトラメチルローダミン - 6 - U T P、テキサスレッド(登録商標) - 5 - U T P、mCherry、カスケードブルー(登録商標) - 7 - U T P、BODIPY T M F L - 14 - U T P、BODIPY T M R - 14 - U T P、BODIPY T M T R - 14 - U T P、ローダミングリーン(登録商標) - 5 - U T P、アレクサフルオール(商標) 488 - 5 - U T P、アレクサフルオール(商標) 546 - 14 - U T P (Molecular Probes, Inc. Eugene, OR) などが挙げられる。あるいは、上記蛍光発色団及び本明細書に記載のものは、例えばホスホロアミダイト又はNH₂S化学を用いてオリゴヌクレオチド合成中に添加することができる。他の蛍光発色団を有するヌクレオチドの特注の合成についてのプロトコルは当該分野で公知である(Henegariu et al. (2000) Nature Biotechnol. 18:345 参照)。2 - アミノプリンは、その合成中にオリゴヌクレオチド配列に直接組み込むことができる蛍光塩基である。核酸はまた、DAPI、YOYO - 1、臭化エチジウム、シアニン色素(例えば、SYBRグリーン)などのインターカレーティング色素を用いて先験的に染色することもできる。

【0078】

合成後の付着に利用可能な他の蛍光発色団には、特に限定されないが、アレクサフルオール(登録商標) 350、アレクサフルオール(登録商標) 405、アレクサフルオール(登録商標) 430、アレクサフルオール(登録商標) 532、アレクサフルオール(登録商標) 546、アレクサフルオール(登録商標) 568、アレクサフルオール(登録商標) 594、アレクサフルオール(商標) 647、BODIPY 493 / 503、BODIPY FL、BODIPY R6G、BODIPY 530 / 550、BODIPY TMR、BODIPY 558 / 568、BODIPY 558 / 568、BODIPY 564 / 570、BODIPY 576 / 589、BODIPY 581 / 591、BODIPY TR、BODIPY 630 / 650、BODIPY 650 / 665、カスケードブルー、カスケードイエロー、ダンシル、リサミンローダミンB、マリーナブルー、オレゴングリーン488、オレゴングリーン514、パシフィックブルー、パシフィックオレンジ、ローダミン6G、ローダミングリーン、ローダミンレッド、テトラメチルローダミン、テキサスレッド(Molecular Probes, Inc., Eugene, ORから入手可能)、Cy2、Cy3、Cy3.5、Cy5、Cy5.5、Cy7(Amersham Biosciences, Piscataway, NJ) などが挙げられる。特に限定されないが、PerCP - Cy5.5、PE - Cy5、PE - Cy5.5、PE - Cy7、PE - テキサスレッド、APC - Cy7、PE - アレキサ色素(610、647、680)、APC - アレキサ色素等を含むFRETタンデム蛍光発色団も使用することができる。

【0079】

金属銀又は金粒子を使用して、蛍光標識されたヌクレオチド及び/又はオリゴヌクレオチド配列からの信号を増強することができる(Lakowicz et al. (2003) BioTechniques 34:62)。

【0080】

オリゴヌクレオチド配列のための他の適切な標識物は、フルオレセイン(FAM、FITC)、ジゴキシゲニン、ジニトロフェノール(DNP)、ダンシル、ビオチン、ブロモデオキシウリジン(BrdU)、ヘキサヒスチジン(6xHis)、ホスホアミノ酸(例えば、P-tyr、P-ser、P-thr)等が挙げられる。ある実施態様において、検出のために以下のハプテン/抗体対が使用され、各抗体は検出可能標識物を用いて誘導体化される: ビオチン/ - ビオチン、ジゴキシゲニン/ - ジゴキシゲニン、ジニトロフェノール(DNP)/ - DNP、5 - カルボキシフルオレセイン(FAM)/ a - FAM。

【0081】

本明細書に記載される検出可能標識物は、スペクトル的に分解可能である。複数の蛍光標識物に関して「スペクトル的に分解可能」とは、標識物の蛍光発光帯が十分に明確であ

10

20

30

40

50

ること、すなわち十分に重複していないことを意味し、それぞれの標識物が付着する分子タグは、それぞれの標識物により生成される蛍光信号に基づいて、標準的な光検出システム、例えば米国特許第4,230,558号;第4,811,218号;など、又はFlow Cytometry: Instrumentation and Data Analysis (Academic Press, New York, 1985) 中の Wheelless et al., pgs. 21-76 に記載されているシステムにより例示されるような、バンドパスフィルター及び光電子増倍管などのシステムを使用して区別することができる。ある態様において、フルオレセイン、ローダミンなどのスペクトル的に分解可能な有機色素は、波長放出最大値が少なくとも20nm離間しており、別の態様において少なくとも40nm離間していることを意味する。別の態様において、キレート化ランタニド化合物、量子ドットなどは、スペクトル分解可能であるとは、波長放出最大値が少なくとも10nm離れており、さらなる態様において、少なくとも15nm離れていることを意味する。

10

配列決定方法

【0082】

本発明は、本発明の配列決定プローブを用いて核酸配列を決定する方法に関する。この方法の例を図8~12に示す。

【0083】

この方法は、本発明の少なくとも1つの配列決定プローブを、基板に固定化された(例えば、1、2、3、4、5、6、7、8、9、10、又はそれ以上の位置で)標的核酸に可逆的にハイブリダイズさせることを含む。

20

【0084】

基板は、当技術分野で公知の任意の固体支持体、例えば標的核酸を固定化することができる、被覆スライド及びマイクロ流体デバイスであり得る。ある実施態様において、基板は、表面、膜、ビーズ、多孔質材料、電極、又はアレイである。標的核酸は、当業者に明らかな任意の基板上に固定化することができる。

【0085】

実施態様において、標的核酸は、標的核酸の一部に相補的なドメインを含む捕捉プローブにより結合される。この部分は、標的核酸の末端であっても末端でなくてもよい。

【0086】

有用な基板の例には、リガンド、抗原、炭水化物、核酸、受容体、レクチン、及び抗体からなる群より選択される結合部分を含むものが挙げられる。捕捉プローブは、基板の結合部分と結合することができる結合部分が含まれる。反応性部分を含む有用な基板の例には、特に限定されないが、エポキシ、アルデヒド、金、ヒドラジド、スルフヒドリル、NH₂-エステル、アミン、チオール、カルボキシレート、マレイミド、ヒドロキシメチルホスフィン、イミドエステル、イソシアネート、ヒドロキシル、ペンタフルオロフェニル-エステル、ソラレン、ピリジルスルフィド、又はビニルスルホン、ポリエチレングリコール(PEG)、ヒドロゲル、又はそれらの混合物を含む表面が挙げられる。そのような表面は、市販品から得ることができるか、又は標準的な技術に従って調製することができる。反応性部分を含む有用な基板の例としては、特に限定されないが、OptArray-DNA-NH₂基(Acceler8)、ネクステリオン(Nexterion)スライドAL(Schott)及びネクステリオンスライドE(Schott)が挙げられる。

30

40

【0087】

実施態様において、捕捉プローブの結合部分はビオチンであり、基板はアビジン(例えばストレプトアビジン)を含む。アビジンを含む有用な基板が市販されており、TBO200(Acceler8)、SAD6、SAD20、SAD100、SAD500、SAD200(Xantec)、スーパーアビジン(Array-It)、ストレプトアビジンスライド(カタログ番号MPC000、Xenopore)、及びSストレプトアビジンスライド(カタログ番号439003、Greiner Bio-one)が挙げられる。

【0088】

実施態様において、捕捉プローブの結合部分はアビジン(例えばストレプトアビジン)

50

であり、基板はビオチンを含む。市販されているビオチンを含む有用な基板には、特に限定されないが、Optiarray-biotin (Accler8)、BD 6、BD 20、BD 100、BD 500、及びBD 2000 (Xantec) が含まれる。

【0089】

実施態様において、捕捉プローブの結合部分は、光活性化により基板に結合することができる反応性部分を含むことができる。基板は光反応性部分を含むことができ、又はナノレポーターの第1の部分は光反応性部分を含むことができる。光反応性部分のいくつかの例としては、アリアルアジド、例えばN((2-ピリジルジチオ)エチル)-4-アジドサリチルアミド；フッ素化アリアルアジド、例えば4-アジド-2,3,5,6-テトラフルオロ安息香酸；ベンゾフェノン主体の試薬、例えば4-ベンゾイル安息香酸のスクシンイミジルエステル；及び5-プロモ-デオキシウリジンが挙げられる。

10

【0090】

実施態様において、捕捉プローブの結合部分は、当業者に明らかな他の結合対を介して基板に固定化することができる。

【0091】

基板に結合した後、標的核酸を伸長させるのに十分な力（例えば、重力、流体力、電磁力「エレクトロストレッチング」、流動延伸、後退メニスカス技術、及びそれらの組み合わせ）を加えることにより、伸長させることができる。

【0092】

標的核酸は、標的核酸の第2の部分に相補的なドメインを含む第2の捕捉プローブにより結合されてもよい。この部分は、標的核酸の末端であっても末端でなくてもよい。第2の捕捉プローブの結合は、標的核酸の伸長の後又は伸長中に、又は伸長されていない標的核酸に対して行うことができる。第2の捕捉プローブは、上記のような結合を有することができる。

20

【0093】

捕捉プローブは、検出可能標識物、すなわち基準スポット点を含むか、又はそれと関連していてもよい。

【0094】

捕捉プローブは、標的核酸を試料から単離することができる。ここでは、捕捉プローブは、標的核酸を含む試料に添加される。捕捉プローブは、標的核酸の領域に相補的な捕捉プローブの領域を介して、標的核酸に結合する。標的核酸が、捕捉プローブの結合部分に結合する部分を含む基板と接触すると、核酸は基板上に固定化される。

30

【0095】

ユーザーが、高断片化試料からできるだけ多くの標的核酸分子を確実に「捕捉」するためには、それぞれが標的核酸の異なる領域に相補的な複数の捕捉プローブを含むことが有用である。例えば、第1のプールが、その5'末端近くの標的核酸の領域に相補的であり、第2のプールが、標的核酸の中央の領域に相補的であり、及び第3のプールがその3'末端の近くで相補的である、3つのプールの捕捉プローブがあり得る。これは、標的核酸あたり「n個の目的の領域」に一般化することができる。この例では、断片化した標的核酸の個々のプールは、ビオチンタグを含むか又はビオチンタグに結合した捕捉プローブに結合した。1/n番目の投入試料（n = 標的核酸中の異なる領域の数）が、各プールチャンバーについて単離される。捕捉プローブは、目的の標的核酸に結合する。次いで、標的核酸は、捕捉プローブのビオチンを介して、基板に接着したアビジン分子に固定化される。場合により標的核酸は、例えば流れ又は静電気力により伸張される。全てのn個のプールは同時に伸長と結合(stretched-and-bound)することができ、又は完全に伸長した分子の数を最大にするために、プール1（これはほとんどの5'領域を捕捉する）を最初に伸長して結合させることができる。次にプール2（これは標的の中間領域を捕捉する）が伸長されて結合される。最後に、プール3を伸長されて結合されることができる。

40

【0096】

必要とされる別個の捕捉プローブの数は、標的核酸断片のサイズに反比例する。換言す

50

れば、高度に断片化された標的核酸には、より多くの捕捉プローブが必要とされる。高度に断片化され分解された標的核酸を有するタイプの試料（例えば、ホルマリン固定パラフィン包埋組織）では、捕捉プローブの複数のプールを含むことが有用であり得る。一方、長い標的核酸断片を有する試料、例えばインビトロで単離された核酸を有する試料では、5'末端に1つの捕捉プローブで十分であり得る。

【0097】

2つまでの捕捉プローブ間、又は1つの捕捉プローブの後で標的核酸の末端の前の標的核酸の領域は、本明細書では「ギャップ」と呼ぶ。このギャップは、本発明の配列決定プローブにより結合されるのに利用可能な標的核酸の一部分である。最小ギャップは標的結合ドメイン長（例えば4～10ヌクレオチド）であり、最大ギャップは全染色体の大部分である。

10

【0098】

固定化された標的核酸を図12に示す。ここで2つの捕捉プローブは、「5'捕捉プローブ」及び「3'捕捉プローブ」として特定される。

【0099】

図8Aは、標的核酸に結合した配列決定プローブの概略図を示す。ここで、標的核酸はチミジン（T）を有する。検出可能標識物を含む相補的核酸又はレポーター複合体の第1のプールが最上部に示され、プールの各メンバーは異なる検出可能標識物（例えば、チミジンは緑色信号により特定される）及び異なるヌクレオチド配列を有する。標的結合ドメイン中の第1ヌクレオチドは、標的核酸中のTに結合する。プローブの第1付着領域は、プローブの標的結合ドメイン中の第1ヌクレオチドがチミジンに結合することを特定する、1つ以上のヌクレオチド配列を含む。従って、チミジンのための相補的核酸のみが、バーコードドメインの第1の位置に結合する。示されるように、検出可能標識物を含むチミジンをコードする第1の相補的核酸又は検出可能標識物を含むレポーター複合体は、プローブのバーコードドメインの位置1の付着領域に結合される。

20

【0100】

相補的核酸又はレポーター複合体のプールの数は、バーコードドメイン中の位置の数と同一である。従って、6つの位置を有するバーコードドメインについては、6つのプールがプローブ上で循環される。

【0101】

あるいは、標的核酸をプローブと接触させる前に、プローブを、その第1の位置で検出可能標識物を含む相補的核酸又はレポーター複合体にハイブリダイズさせることができる。従って、その標的核酸と接触すると、プローブはその第1の位置から検出可能な信号を放出することができ、バーコードドメイン上の第1の位置に向けられた相補的核酸又はレポーター複合体の第1のプールを提供する必要はない。

30

【0102】

図8Bは、図8Aに示す方法を継続する。ここで、バーコードドメインの第1の位置の付着領域に結合したチミジンの第1の相補的核酸（又はレポーター複合体）は、チミジン用であり検出可能標識物を欠く第1のハイブリダイズ性核酸で置換されている。チミジン用であり検出可能標識物を欠く第1のハイブリダイズ性核酸は、検出可能標識物を含むすでに結合している相補的核酸を、又はすでに結合しているレポーター複合体を置換する。こうして、バーコードドメインの位置1はもはや検出可能な信号を放出しない。

40

【0103】

実施態様において、検出可能標識物を含む相補的核酸又はレポーター複合体は、付着領域から除去され得るが、検出可能標識物を欠くハイブリダイズ性核酸では置換され得ない。これは、例えば、カオトロピック剤を添加すること、温度を上昇させること、塩濃度を変えること、pHを調整すること、及び/又は流体力を加えることにより生じ得る。これらの実施態様において、必要な試薬（すなわち、検出可能標識物を欠くハイブリダイズ性核酸）はより少ない。

【0104】

50

図 8 C は、本願発明の方法を継続する。ここで、標的核酸はチミジン (T) に続いてシチジン (C) を有する。相補的核酸の又はレポーター複合体の第 2 のプールが最上部に示され、プールの各メンバーは異なる検出可能標識物及び異なるヌクレオチド配列を有する。更に、第 1 のプールの相補的核酸又はレポーター複合体の相補的核酸のヌクレオチド配列は、第 2 のプールのヌクレオチド配列とは異なる。しかし、塩基特異的検出可能標識物は相補的核酸のプールに共通であり、例えばチミジンは緑色信号により特定される。ここで、標的結合ドメイン中の第 2 ヌクレオチドは、標的核酸中の C に結合する。プローブの第 2 付着領域は、プローブの標的結合ドメイン中の第 2 ヌクレオチドがシチジンに結合することを特定するヌクレオチド配列を有する。従って、検出可能標識物含む相補的核酸又は第 2 のプールからのシチジン用のレポーター複合体のみが、バーコードドメインの第 2 の位置に結合する。示されるように、シチジンをコードする第 2 の相補的核酸又はレポーター複合体は、プローブのバーコードドメインの第 2 の位置で結合される。

10

【 0 1 0 5 】

実施態様において、図 8 C に示される工程は、図 8 B に示される工程に続く。ここで、相補的核酸又はレポーター複合体の第 1 のプール (図 8 A) が、検出可能標識物を欠く第 1 のハイブリダイズ性核酸 (図 8 B) で置換されると、相補的核酸又はレポーター複合体の第 2 のプールが提供される (図 8 C に示すように) 。あるいは、図 8 C に示す工程は、図 8 B に示す工程と同時に起きる。ここで、検出可能標識物を欠く第 1 のハイブリダイズ性核酸 (図 8 B) は、相補的核酸又はレポーター複合体の第 2 のプールと同時に提供される (図 8 C に示すように) 。

20

【 0 1 0 6 】

図 8 D は、図 8 C に示される方法を継続する。ここで、バーコードドメイン上の第 1 ~ 第 5 の位置は、検出可能標識物を含む相補的核酸又はレポーター複合体により結合され、検出可能標識物を欠くハイブリダイズ性核酸で置換されている。バーコードドメインの第 6 の位置は、現在、検出可能標識物を含む相補的核酸又はレポーター複合体により結合されており、これは、標的結合ドメインの第 6 の位置がグアニン (G) に結合していると特定する。

【 0 1 0 7 】

上述したように、検出可能標識物を含む相補的核酸又はレポーター複合体は、付着領域から除去することはできるが、検出可能標識物を欠くハイブリダイズ性核酸で置換することはできない。

30

【 0 1 0 8 】

必要に応じて、検出可能標識物の交換速度を加速する小さな 1 本鎖オリゴヌクレオチドを組み込むことにより、検出可能標識物交換の速度を加速することができる (例えば、*「T o e - H o l d」* プローブ ; 例えば、Seeling et al., "Catalyzed Relaxation of a Metastable DNA Fuel" ; J. Am. Chem. Soc. 2006, 128(37), pp12211-12220 を参照) 。

【 0 1 0 9 】

相補的核酸又はレポーター複合体を、バーコードドメイン上の最終位置 (図 8 D の 6 番目の位置) で置き換えることが可能である。しかし、これは、配列決定プローブを別の配列決定プローブと交換する場合には、必要でないことがある。実際、図 8 D の配列決定プローブは、標的核酸から脱ハイブリダイズして除去され、そして、図 8 E で示されるように、任意の相補的核酸によりまだ結合されていない第 2 の (重複する又は重複していない) 配列決定プローブで置換することができる。図 8 E のプローブは、プローブの第 2 集団に含めることができる。

40

【 0 1 1 0 】

図 8 A ~ 8 E と同様に、図 9 A 及び 9 D ~ 9 G は、本発明の方法の工程を示す。しかし図 9 A 及び 9 D ~ 9 G は、レポーター複合体 (検出可能標識物を含む) が配列決定プローブの付着領域に結合していることを明確に示している。図 9 D 及び 9 E は、レポーター複合体にハイブリダイズしたプローブから放出された蛍光信号を示す。図 9 D 及び 9 E は、標的核酸が「 T - A 」配列を有することを示す。

50

【0111】

図10は、図9D及び図9Eに示される工程を要約したものである。図の上部には、例示的なプローブのヌクレオチド配列が示され、プローブの重要なドメインを特定する。プローブは、その標的結合ドメインとそのバーコードドメインとの間に任意の2本鎖DNAスパーサーを含む。バーコードドメインは、順に、「Flank 1」部分、「AR-1」部分、「AR-1/Flank 2」部分、「AR-2」部分、及び「AR-2/Flank 3」部分を含む。工程1では、「AR-1 Detect」がプローブの「AR-1」部分及び「AR-1/Flank 2」部分にハイブリダイズされる。「AR-1 Detect」は、第1の位置のチミジンをコードする、検出可能標識物を含むレポーター複合体又は相補的核酸に対応する。従って、工程1は図9Dに対応する。工程2では、「Lack 1」がプローブの「Flank 1」部分及び「AR-1」部分にハイブリダイズされる。「Lack 1」は、プローブの第1付着領域（図9Eに第1付着領域を覆う黒色の棒として示す）に特異的である検出可能標識物を欠くハイブリダイズ性核酸に対応する。レポーター複合体又は相補的核酸の5'側である「Flank 1」位置にハイブリダイズさせることにより、ハイブリダイズ性核酸は、レポーター複合体/相補的核酸をプローブからより効率的に置換する。「Flank」部分は、「Toe-Holds」とも呼ばれる。工程3では、「AR-2 Detect」はプローブの「AR-2」及び「AR-2/Flank 3」部分にハイブリダイズされる。「AR-2 Detect」は、第2の位置のグアニンをコードする検出可能標識物を含むレポーター複合体又は相補的核酸に対応する。従って、工程3は図9Eに対応する。この実施態様において、検出可能標識物を欠くハイブリダイズ性核酸/レポーター複合体が逐次提供される。

10

20

【0112】

あるいは、検出可能標識物を欠くハイブリダイズ性核酸及び検出可能標識物を含む相補的核酸/レポーター複合体が同時に提供される。この代替の実施態様を図11に示す。工程2では、「Lack 1」（検出可能な標識物を欠くハイブリダイズ性核酸）が「AR-2 Detect」（位置2のグアニンをコードするレポーター複合体）と共に提供される。この代替の実施態様は、2つの工程を1つにまとめたため、図10に示された実施態様よりも時間的に有効であり得る。

【0113】

図12は、本発明の方法を例示する。ここで、標的核酸は捕捉され、2つの位置で固定化され、こうしてプローブが結合することができる「ギャップ」が生じる。第1集団のプローブが標的核酸上にハイブリダイズされ、検出可能標識物が検出される。最初の工程が、第2集団のプローブ、第3集団のプローブから、及び100を超える集団までのプローブで繰り返される。約100集団のプローブの使用は、標的核酸中の各ヌクレオチドの約5Xの適用範囲を提供する。図12は、1つの視野（FOV）からの信号を検出するのに必要な時間に基づく、読み出し時間の推定速度を提供する。

30

【0114】

標的核酸の長さに沿ったプローブの分布は、検出可能な信号の分解にとって重要である。上述したように、2つの検出可能標識物の分解能限界は約600ヌクレオチドである。好ましくは、プローブの集団中の各配列決定プローブは、互いに600ヌクレオチドよりも近くに結合しない。上述したように、600ヌクレオチドは、典型的な配列決定装置の分解能限界である。この場合、配列決定プローブは単一の読み取りを提供する。これは図12の最も左端の分解能制限スポットに示されている。

40

【0115】

ランダムに、しかし一部は、標的結合ドメインの長さ、プローブの T_m 、及び適用されるプローブの濃度に依存して、集団中の2つの異なる配列決定プローブが互いの600ヌクレオチド以内で結合することが可能である。この場合、順序付けされていない複数の読み取りは、単一の分解能制限スポットから発光する。これは図12の第2の分解能制限スポットに示されている。

【0116】

代替的に又は追加的に、集団中の配列決定プローブの濃度を低下させて、例えば配列決

50

定装置の分解能限界を超えるように、標的核酸の特定の領域におけるプローブの適用範囲を減少させ、こうして分解能制限スポットから単一の読み取りを生成してもよい。

【0117】

図23は、図8～12で使用されたものとは異なる配列決定プローブの概略図を示す。ここで、バーコードドメイン上の各位置は、検出可能標識物を含む相補的核酸又はレポーター複合体により結合される。このように、この例では、相補的核酸を順次置換する必要なく、6つのヌクレオチド配列を読み取ることができる。この配列決定プローブの使用は、記載された方法の多くの工程が省略されるため、配列情報を得るための時間を短縮するであろう。しかしながらこのプローブは、重複しない検出可能標識物から利益を受けると考えられ、例えば蛍光発色団が、重複しない波長の光により励起されるか、又は蛍光発色団が、重複しない波長の光を放射するであろう。

10

【0118】

この方法は更に、固定化された標的核酸の各領域についてヌクレオチドの各特定された直線的順序を組み立て、こうして固定化された標的核酸の配列を特定する工程を含む。この組み立て工程は、実行可能なプログラムが格納された非一時的なコンピュータ可読記憶媒体を使用する。プログラムはマイクロプロセッサに、標的核酸の各領域についてヌクレオチドの各特定された直線的順序を配置し、こうして核酸の配列を得るように指示する。組み立ては、配列決定プローブからデータが収集された後ではなく、「リアルタイム」で、すなわちデータが収集されている間に行われる。

【0119】

20

上記の態様及び実施態様のいずれも、「課題を解決するための手段」及び/又は「発明を実施するための形態」の項で本明細書に開示されている任意の他の態様又は実施態様と組み合わせることができる。

定義：

【0120】

特定の例示的な実施態様において、本明細書で使用される「アニーリング」及び「ハイブリダイゼーション」という用語は、安定した2本鎖の形成を意味するために交換可能に使用される。ある態様において、安定な2本鎖とは、2本鎖の T_m より約5℃低い又は約5℃高い温度、及び低い一価の塩濃度、例えば0.2 M未満、又は0.1 M未満、又は当業者に公知の塩濃度の、厳密な洗浄条件下で、2本鎖構造が破壊されないことを意味する。「完全に一致する」という用語は、2本鎖に関して使用される場合、2本鎖を構成するポリヌクレオチド及び/又はオリゴヌクレオチド鎖が、互いに2本鎖構造を形成して、各鎖中のすべてのヌクレオチドがもう一方の鎖のヌクレオチドとワトソン-クリック塩基対合をすることを意味する。「2本鎖」という用語は、特に限定されないが、使用できるデオキシノシンヌクレオシドなどのヌクレオシド類似体、2-アミノプリン塩基を有するヌクレオシド、PNAなどとの対合を含む。2つのオリゴヌクレオチド間の2本鎖における「ミスマッチ」は、2本鎖中の1対のヌクレオチドがワトソン-クリック結合をしないことを意味する。

30

【0121】

本明細書で使用される「ハイブリダイゼーション条件」という用語は、典型的には、約1 M未満、より一般的は約500 mM未満、更に一般的は約200 mM未満の塩濃度を含む。ハイブリダイゼーション温度は、5℃と低くてもよいが、典型的には22℃より高く、より典型的には約30℃より高くてもよく、しばしば約37℃を超える。ハイブリダイゼーションは、通常厳密性条件下で、例えばプローブがその標的部分配列に特異的にハイブリダイズする条件下で行われる。厳密性条件は、配列に依存性であり、異なる状況では異なる。より長い断片は、特異的ハイブリダイゼーションのためにより高いハイブリダイゼーション温度を必要とすることがある。塩基組成、相補鎖の長さ、有機溶媒の存在、及び塩基のミスマッチの程度を含む他の要因が、ハイブリダイゼーションの厳密性に影響を及ぼす可能性があるため、パラメーターの組み合わせは、いずれか1つの要因のみの絶対的な測定値より重要である。

40

50

【0122】

一般に、厳密性条件は、特定のイオン強度及びpHにおいて、特定の配列のT_mよりも約5℃低くなるように選択される。例示的な厳密性条件としては、pH7.0～8.3及び少なくとも25℃の温度で、少なくとも0.01M～1M以下のNaイオン濃度（又は他の塩）の塩濃度が挙げられる。例えば、5×SSPE（750mM NaCl、50mMリン酸ナトリウム、5mM EDTA、pH7.4）及び25～30℃の温度条件が、対立遺伝子特異的プローブハイブリダイゼーションに適している。厳密性条件については、例えば Sambrook, Fritsche and Maniatis, “Molecular Cloning A Laboratory Manual, 2nd Ed.” Cold Spring Harbor Press (1989) 及び Anderson Nucleic Acid Hybridization, 1st Ed., BIOS Scientific Publishers Limited (1999) を参照されたい。本明細書で使用される「特異的にハイブリダイズする」という用語又は同様の用語は、厳密性条件下で実質的に特異的ヌクレオチド配列への実質的な、分子の結合、2本鎖形成、又はハイブリダイゼーションを指す。

10

【0123】

プローブの特定の位置に関連する検出可能標識物は、1回又は複数回「読み出す」（例えば、その蛍光を検出する）ことができる；「読み出し」は、用語「ベースコール（base call）」と同義であり得る。複数回の読み込みで精度が向上する。単一の元の標的分子に由来する配列情報の連続ストレッチが検出された場合、標的核酸配列は「読み込まれている」。典型的には、これはマルチパスコンセンサス（以下に定義する）により生成される。本明細書で使用される用語「適用範囲（coverage）」又は「適用範囲の深さ」は、標的領域が（離散読み取りを介して）配列決定され、参照配列に整列された回数を指す。読み取り適用範囲は、特定の参照標的配列にマッピングされる読み取りの総数である。塩基適用範囲は、特定のゲノム位置で行われたベースコールの総数である。

20

【0124】

本明細書中で使用される「ハイブ・アンド・セク・サイクル（hybe and seq cycle）」は、特定のプローブ又はプローブ集団上の各付着領域を検出するために必要とされるすべての工程を指す。例えば、標的核酸上の6つの位置を検出することができるプローブの場合、1つの「ハイブ・アンド・セク・サイクル」は、少なくとも、プローブを標的核酸にハイブリダイズさせる工程と、相補的核酸／レポーター複合体をプローブのバーコードドメイン上の6つの位置のそれぞれで付着領域にハイブリダイズさせる工程と、6つの位置のそれぞれに関連する検出可能標識物を検出する工程とを含む。

30

【0125】

「k-量体プローブ」という用語は、本発明のプローブと同義である。

【0126】

離散読み取りからの2つ以上の配列が整列される場合、重複部分を組み合わせることで単一のコンセンサス配列を作製することができる。重複部分が同じ塩基（整列の単一のカラム）を有する位置では、それらの塩基がコンセンサスになる。重複配列の間に不一致がある位置についてコンセンサスを生成するために、様々な規則を使用することができる。単純な多数決規則では、カラム中の最も一般的な塩基をコンセンサスとして使用する。「マルチパスコンセンサス」は、単一の標的分子からの全ての離散プローブ読み取りの整列である。適用されるプローブ集団／ポーリングのサイクルの総数に応じて、単一の標的分子内の各塩基位置を、異なるレベルの冗長性又は重複度で照会することができる。一般に、冗長性はベースコールの信頼水準を高める。

40

【0127】

「生データの正確性（Raw Accuracy）」は、塩基を正しく特定するシステム固有の能力の尺度である。生データの正確性は配列決定技術に依存する。「コンセンサス正確性（Consensus Accuracy）」は、追加の読み取りと統計力を使用して塩基を正しく特定するシステムの能力の尺度である。「特異性」とは、1回の実行当たりの総読み取りのうち、意図した標的にマッピングする読み取りの割合を指す。「均一性」は、標的領域にわたる配列の適用範囲の変動を指し、高い均一性は小さい変動と相関する。この特徴は一般に、すべ

50

ての標的領域にわたる平均適用範囲深度の20%以上によりカバーされる標的領域の割合として報告される。確率的誤差（すなわち、内因性の配列決定化学誤差）は、同じ標的核酸の「マルチパス」配列決定により容易に修正することができる。十分な回数のパスが与えられれば、実質的に「完全なコンセンサス」又は「誤差の無い」配列決定を達成することができる。

本明細書に記載の方法は、本方法を実行し及び／又は結果を記録することができる任意のデバイスを用いて、実行し及び／又は結果を記録することができる。使用できる装置の例には、特に限定されないが、すべてのタイプのコンピュータを含む電子計算装置が含まれる。本明細書に記載の方法がコンピュータで実施され及び／又は記録される場合、本方法の工程を実行するようにコンピュータを構成するために使用できるコンピュータプログラムは、コンピュータプログラムを含むことができる任意のコンピュータ可読媒体に含めることができる。使用可能なコンピューター可読媒体の例には、特に限定されないが、ディスク、CD-ROM、DVD、ROM、RAM、非一時的コンピュータ可読媒体、及び他のメモリ、及びコンピュータ記憶装置が含まれ得る。本方法の工程を実行し、シーケンス情報を組み立て、及び／又は結果を記録するように、コンピュータを構成するために使用できるコンピュータプログラムは、電子ネットワークを介して、例えばインターネット、イントラネット又は他のネットワークを介して提供されてもよい。

【0128】

「消耗型配列決定カード」（図24）は、当技術分野で公知の蛍光イメージング装置に組み込むことができる。様々な特徴を有する任意の蛍光顕微鏡は、この配列決定読み出しを行うことができる。例えば、広視野ランプ、レーザー、LED、多光子、共焦点、又は全内部反射照明を、励起及び／又は検出に使用することができる。蛍光顕微鏡の発光検出チャンネルでは、フィルタ主体又はグレーティング主体のスペクトル分解能（1つ以上のスペクトル分解された発光波長）を有するカメラ（単一又は複数）及び／又は光電子増倍管（単一又は複数）が可能である。標準的コンピュータは、消耗型配列決定カード、カードを流れる試薬、蛍光顕微鏡による検出の全てを制御することができる。

【0129】

配列決定データは、任意の数の標準的な次世代配列決定アセンブラーにより解析することができる（例えば、Wajid and Serpedin, “Review of general algorithmic features for genome assemblers for next generation sequencers” Genomics, proteomics & bioinformatics, 10 (2), 58-73, 2012を参照）。顕微鏡の単一の回折限界領域内で得られた配列決定データは、「局所的に組み立てられ」、回折スポット内の複数の読み取りからコンセンサス配列を生成する。次に複数の回折スポット組み立て読み取りを一緒にマッピングして、標的遺伝子セット全体を表す連続配列が生成されるか、又はゲノム全体の新規組み立てが生成される。

【0130】

本発明に関連するさらなる教示は、以下の1つ以上に記載されている：米国特許第8,148,512号、米国特許第7,473,767号、米国特許第7,919,237号、米国特許第7,941,279号、米国特許第8,415,102号、米国特許第8,492,094号、米国特許第8,519,115号、米国特許第2009/0220978号、米国特許第2009/0299640号、米国特許第2010/0015607号、米国特許第2010/0261026号、米国特許第2011/0086774号、米国特許第2011/0145176号、米国特許第2011/0201515号、米国特許第2011/0229888号、米国特許第2013/0004482号、米国特許第2013/0017971号、米国特許第2013/0178372号、米国特許第2013/0230851号、米国特許第2013/0337444号、米国特許第2013/0345161号、米国特許第2014/005067号、米国特許第2014/0017688号、米国特許第2014/0037620号、米国特許第2014/0087959号、米国特許第2014/0154681号、及び米国特許第2014/0162251号（これらのそれぞれは、その全体が参照により本明細書に組み込まれる）。

【実施例】

【0131】

実施例 1

標的核酸配列を決定する本発明の方法は迅速である

【0132】

図 8 ~ 図 12 に示される本発明の方法中の工程のタイミングが、以下で説明される。

【0133】

本発明の必要な試料調製は最小である。例えば図 13 に示すように、試料中の核酸は、2 時間以下の調製時間後に読み取ることができる。これは、Ion Torrent (AmpliSeq (登録商標)) 又は Illumina (TruSight) 配列決定について必要な時間 (これらは、それぞれ約 12 時間又は 9 時間の調製時間を必要とする) より大幅に短い時間である。

10

【0134】

例示的な実行時間の計算を図 14 に示し、サイクル時間の計算を図 15 に示す。

【0135】

固定化された標的核酸にプローブ集団を結合させるには、約 60 秒が必要である。この反応は、合成主鎖上の標的結合ドメインの複数のコピーを利用することにより促進することができる。マイクロ流体制御性の流体交換装置を用いると、結合していないプローブを洗い流すのに約 0.5 秒が必要である。

【0136】

相補的核酸の第 1 のプール (検出可能標識物を含む) を加え、それらをバーコードドメインの第 1 の位置の付着領域に結合させるには、約 15 秒が必要である。

20

【0137】

各視野 (FOV) は、4 つの異なる色について画像化され、各色が単一の塩基を表す。5' 捕捉プローブ又は 3' 捕捉プローブ (又はその両方) 上に置かれた基準スポットは、2 つの位置の間の 1 本の線の中の光学的バーコードのみ (ギャップ付き標的核酸の存在と一致) を読み取るのに有用であり得る。配列決定プロセスの連続する工程で画像の等しい整列を生成するために、基準スポットを各視野に追加することもできる。4 つの画像はすべて 1 つの FOV で取得することができ、次に、光学読取装置は新しい FOV に移動することができるか、又はすべての FOV を 1 つの色で取り込み次に第 2 の色で再画像化することができる。1 つの FOV は、約 0.5 秒で読み取ることができる。次の FOV に移動するには約 0.5 秒かかる。従って、「n」個の FOV を読む時間は、「n」× 1 秒に等しい。

30

【0138】

検出可能標識物を有する相補的核酸は、加熱、又は検出可能標識物を欠く過剰の相補的核酸による洗浄により、バーコードドメインの第 1 の位置から除去される。必要に応じて、検出可能標識物の交換速度は、検出可能標識物交換の速度を加速する小さな 1 本鎖オリゴヌクレオチドを組み込むことにより、を加速することができる (例えば、「Toe-Hold」プローブ; Seeling et al., "Catalyzed Relaxation of a Metastable DNA Fuel"; J. Am. Chem. Soc. 2006, 128(37), pp 12211-12220 を参照)。FOV は再画像化して、検出可能標識物を有する全ての相補的核酸が除去されたことを確認した後、移動が継続される。これには約 15 秒かかる。バックグラウンド信号レベルに達するまで、この工程を繰り返すことができる。

40

【0139】

上記の工程を繰り返される。

【0140】

読み取りの合計時間は、 m (読み取られた塩基数) \times (15 秒 + n FOV 時間 \times 1 秒 + 15 秒) に等しい。例えば、バーコード領域内の位置の数が 6 で 20 FOV である場合、読み出す時間は $6 \times (30 + 20 + 15)$ 、すなわち 390 秒に等しい。

【0141】

第 1 集団のプローブは、脱ハイブリダイズされる。これには約 60 秒かかる。

50

【 0 1 4 2 】

上記工程は、プローブの第2集団及び続く集団について繰り返される。配列決定プローブの集団が融解温度（ T_m ）により編成されている場合、プローブの各集団は、各塩基が必要な深さ（これはエラー率により決定される）まで覆われることを確実にするために、複数回のハイブリダイゼーションを必要とする。更に、実行中のハイブリダイゼーション読み取りを分析することにより、配列全体が実際に決定される充分前に配列決定されている個々の遺伝子を認識することが可能である。従って、サイクリングは、特定の所望のエラー頻度（又は適用範囲）が満たされるまで繰り返すことができる。

【 0 1 4 3 】

上記タイミングを、いくつかのギャップ付き核酸結合密度推定値と共に使用して、本発明のナノストリング（NSTG）次世代シーケンサー（Nanosttring (NSTG)-Next Generation Sequencer）のスループットを推定することができる。

10

【 0 1 4 4 】

シーケンサの正味のスループットは、以下により与えられる：

塩基占有割合 \times ギャップ長 \times F O V 当たりのギャップの数 \times 光学的バーコード当たりの塩基の数 \div [60 秒（プローブを標的核酸にハイブリダイズさせる） + 0.5 秒（洗浄） + m : バーコードドメイン内の位置 \times (15 秒（相補的核酸を結合する） + $n f o v s \times 1 + 15$ 秒（相補的核酸の結合をはずす）) + 60 秒（標的核酸に対してプローブを脱ハイブリダイズさせる）]

【 0 1 4 5 】

20

従って、ある例では、単一のギャップを有する核酸（図10に示す方法から一緒に加える）の総「サイクル」：

60 秒（プローブを標的核酸にハイブリダイズさせる） + 0.5 秒（洗浄） + m 塩基 \times (15 秒（相補的核酸を結合する） + $n F O V \times 1 + 15$ 秒（相補的核酸の結合をはずす）) + 60 秒（標的核酸に対してプローブを脱ハイブリダイズさせる）。 $m = 6$ 、 $n F O V = 20$ を使用すると、時間 = $60 + 0.5 + 390 + 60 = 510.5$ 秒が得られる。

【 0 1 4 6 】

ギャップ付き核酸領域の占有率 1 %、ギャップ当たり 4000 塩基、及び F O V 当たり 5,000 個のギャップ付き核酸断片、及び $m = 6$ と $n F O V = 20$ （上記のとおり）を仮定すると、

30

$0.01 \times 4000 \times 5000 \times 20 = 510.5$ 秒当たり 4,000,000 の 6 塩基読み取り = 47,012.73 塩基 / 秒

の正味のスループットが得られる。

【 0 1 4 7 】

従って、この例では、24時間の連続測定当たりの正味のスループット = 4.062 ギガ塩基（Gb）/日である。別の推定値は最大 12 Gb / 日。図12を参照。

【 0 1 4 8 】

図14に示されるように、100個の異なる標的核酸（「100ブックス」）を配列決定するのに必要な実行時間は、約 4.6 時間である。1000個の異なる標的核酸（「1000ブックス」）を配列決定するのに必要な実行時間は、約 16 時間である。

40

【 0 1 4 9 】

図16は、本発明及び種々の他の配列決定方法 / 装置の配列決定速度、読み取り数、及び臨床的有用性を比較する。

【 0 1 5 0 】

実施例 2

本発明の方法のエラー率は低い

【 0 1 5 1 】

図17は、末端位置が省略された場合、本発明が約 2.1 % の生のエラー率を有することを示す。

【 0 1 5 2 】

50

本願発明に関して、配列決定に関連するエラー率は、完全に一致した (m + n) 量体と単一塩基のミスマッチ (m - 1 + n) 量体との間の自由エネルギー差に比例する。m + n の合計は標的結合ドメイン中のヌクレオチドの数であり、m はバーコードドメイン中の位置の数を表す。ハイブリダイゼーションの選択性の推定は、方程式を使用して得ることができる (Owczarzy, R. (2005), Biophys. Chem., 117:207-215 及び Integrated DNA Technologies ウェブサイト: ワールドワイドウェブ (www) idtdna.com/analyzer/Applications/Instructions/Default.aspx?AnalyzerDefinitions=true#MismatchMeltTem) :

【 0 1 5 3 】

【数 1】

10

$$\theta = 1 - \left(\frac{K_a([strand2] - [strand1]) - 1}{2K_a[strand2]} + \frac{\sqrt{K_a^2([strand1] - [strand2])^2 + 2K_a([strand1] + [strand2]) + 1}}{2K_a[strand2]} \right)$$

ここで、 K_a は、予測された熱力学的パラメータから得られる結合平衡定数である :

【 0 1 5 4 】

【数 2】

20

$$K_a = \exp \left(\frac{-(\Delta H^\circ - T\Delta S^\circ)}{RT} \right)$$

【 0 1 5 5 】

シートは、特定のハイブリダイゼーション温度で標的にアニーリングされることが予想される、正確な相補体と単一塩基ミスマッチ配列の結合パーセントを表す。T はケルビンで表されるハイブリダイゼーション温度であり、 H° (エンタルピー) 及び S° (エントロピー) は、配列と公開された最近傍の熱力学パラメータから計算される融解パラメータであり、R は理想気体定数 ($1.987 \text{ cal} \cdot \text{K}^{-1} \text{ mol}^{-1}$) であり、[鎖 1 / 2] は、オリゴヌクレオチドのモル濃度であり、-273.15 の定数は、温度をケルビンから摂氏度に変換する。最も正確な最近傍のパラメータは、以下の刊行物から DNA / DNA 塩基対について (Allawi, H., SantaLucia, J. Biochemistry, 36, 10581 参照)、RNA / DNA 塩基対について (Sugimoto et al., Biochemistry, 34, 11211-6 参照)、RNA / RNA 塩基対について (Xia, T. et al., Biochemistry, 37, 14719 参照) 得られた。

30

【 0 1 5 6 】

NSTG シーケンサーから予想される近似エラー率の推定値の例は、以下の通りである。(m + n) = 8 ' 量体について。次の 8 量体バーコードとその 1 塩基のミスマッチを検討する。

5 ' A T C G T A C G 3 ' ,

(配列決定すべき領域)

3 ' T A G C A T G C 5 ' ,

(完璧に一致した光学的バーコードの配列決定)

3 ' T A G T A T G C 5 ' ,

(単一塩基ミスマッチ (G - T) 対を有する光学的バーコードの配列決定)

40

【 0 1 5 7 】

上記方程式に基づいて IDT 計算機を使用すると、

【 0 1 5 8 】

17.4 (完全一致の場合の T_m) で、(50% / 0.3%) は、その配列にハイブリダイズした正しい光学的バーコードと、 T_m での不正なバーコードとの比であり、その

50

配列の推定エラー率 0.6% が得られる。

【0159】

非常に高いGC含量の配列決定計算により以下が得られる：

5'CGCCGGCC3'

(配列決定すべき領域)

3'GCGGCCCGG5'

(完璧に一致した光学的バーコードの配列決定)

3'GCGGACGG5'

(単一塩基ミスマッチ(G-A)対を有する光学的バーコードの配列決定)

【0160】

41.9 (完全一致の場合のT_m)で、(50%/0.3%)は、その配列にハイブリダイズした正しい光学的バーコードと、T_mでの不正なバーコードとの比であり、その配列の推定エラー率 0.8% が得られる。

【0161】

多数の8量体対の試験により、0.2%~1%の範囲のエラー率の分布が得られる。上記計算は使用された条件と同一ではないが、これらの計算は、Pacific Biosciences 及び Oxford Nanopore Technologies (ここで、エラー率は(>>10%)と大きい)のような他の単一分子配列決定技術と比較して、本発明の方法が比較的小さい固有のエラー率を有するという指標を提供する。

【0162】

図18は、本発明の生データの正確さが他の配列決定方法より高いことを証明する。すなわち本発明は、他の配列決定方法に必要とされるパスよりも少ない回数のパスで、単一標的からのコンセンサス配列を提供する。更に本発明は、30回以上のパス後に「完全なコンセンサス」/「誤差無し」配列決定(すなわち、99.9999%/Q60)を得ることができるが、PacBio 配列決定法(例えば)は、70回パス後でもそのようなコンセンサスを達成できない。

【0163】

実施例3

本発明は単一塩基対の分解能を有する

【0164】

図19は、本発明が単一塩基分解能を有し、低いエラー率(具体的なヌクレオチド置換に応じて0%~1.5%の範囲)を有することを示す。

【0165】

バーコードにハイブリダイズされ、通常の ナノストリング(NanoString) 遺伝子発現結合技術を使用してカートリッジの表面に固定化された標的RNAを用いて、さらなる実験を行った(例えば、Geiss et al, "Direct multiplexed measurement of gene expression with color-coded probe pairs"; Nature Biotechnology, 26, 317-325 (2008)を参照)。異なる標的結合ドメイン長及び完全な一致(完全な10量体一致配列に連結されたYGBYGR-2µm光学的バーコード)を有するバーコードがRNA標的にハイブリダイズする能力を測定した(図26)。標的結合ドメインの長さが長いほど、より大きなカウントが得られる。これはまた、10量体標的結合ドメインが、バックグラウンドを超える配列を登録するのに十分であることも示している。個々の単一塩基改変一致の各々は、別の光学的バーコードを用いて合成された。正しくない光学的バーコードに対する正確な光学的バーコードの割合を数えた(図24及び25)。

【0166】

SNPである実際の配列を検出する10量体の能力は、バックグラウンドに対して>15000カウントであり、誤った配列はバックグラウンドに対してせいぜい>400である。正確なプローブが存在する場合、エラー率は実際の配列の3%未満であると予想される。このデータは、(本質的に)最悪の場合であることに注意されたい。6.6キロ塩基の光学的バーコードレポーター(Gen2スタイル)に、10塩基対のハイブリダイゼー

10

20

30

40

50

ション配列のみが付着した。特定の条件の最適化は行われなかった。しかしこのデータは、ナノストリング次世代配列決定アプローチが1塩基対の配列を分割できることを明らかにしている。

【0167】

上記試験で使用された詳細な材料及び方法は以下の通りである：

【0168】

ハイブリダイゼーションプロトコルプローブB + コードセット

- ・25 μ lの要素(194コードセット)を取る
- ・5 μ lプローブB + 相補配列を標的(100 μ M)に添加する
- ・15 μ lのハイブリダイゼーション(Hyb)緩衝液(14.56 \times SSPE 0.18%ツイーン20)を加える

SSPE (150 mM NaCl、NaH₂PO₄ \times H₂O 10 mM、Na₂EDTA 10 mM)

- ・氷上で10分間インキュベートする。
- ・150 μ lのGビーズ(10 mg/mlの40 μ lのGビーズ + 110 μ lの5 \times SSPE 0.1% Tween 20)を加える
- ・室温で10分間インキュベートする
- ・0.1 SSPE 0.1% Tween 20でマグネットコレクタを用いて3回洗浄する

- ・100 μ lの0.1 \times SSPE中で45 で10分間溶出する。

【0169】

標的ハイブリダイゼーションプロトコル(750 mM NaCl)

- ・上記の溶出した試料20 μ lを取る
 - ・10 μ lのハイブリダイゼーション(Hyb)緩衝液を加える
 - ・1 μ lの標的(100 nMビオチン化RNA)を加える
 - ・氷上で30分間インキュベートする
- 15 μ lを取って、ストレプトアビジンスライドに20分間結合させ、Gフックを用いてフローストレッチを行い、nCounterを用いてカウントする

【0170】

材料

要素194コードセット

IDTから購入したオリゴ

SSPE (150 mM NaCl、NaH₂PO₄ \times H₂O 10 mM、Na₂EDTA 10 mM)

ハイブリダイゼーション緩衝液(14.56 \times SSPE 0.18%ツイーン20)

【0171】

表2：12, 11, . . . , 8量体のプローブB配列(配列番号30 ~ 配列番号34)

【表2】

GBRYBG	5	GACTGTACCCACGCGATGACGTTGTCAGAGTCGCATAATCT	3
YRBYRG	5	AGACTGTACCACAAGAATCCCTGCTAGCTGAAGGAGGGTCAAAC	3
YGBYGR	5	GAGACTGTACCCTACGTATATATCCAAGTGGTTATGTCCGACGGC	3
GBRYGB	5	TGAGACTGTACCACCCCTCCAAACGCATTCTTATTGGCAAATGGAA	3
RYGBRG	5	CTGAGACTGTACCCGGAATCGGCATTTGCATTCTTAGGATCTAAA	3

【0172】

表3：標的配列(太字；配列番号35)

10

20

30

40

【表 3】

RNA	5	CAATGTGAGTCTCTTGGTACAGTCTCAGTTAGTCACTCCCTAAG¥ Bio TEG¥	3
-----	---	--	---

【0 1 7 3】

表 4：10 量体ミスマッチのプロープ B 配列（太字；配列番号 36～配列番号 41）

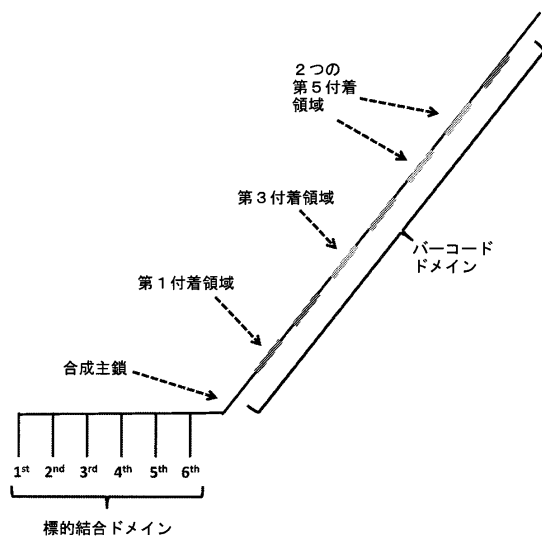
【表 4】

10mermis2A	GAGACAGTACCCTGGTCTAGGTATCTAATTCGTGGGTCGGGTACT
10mermis2C	GAGACCGTACCGCTCATTTTGAACATACGATTGCGATTACGGAAA
10mermis2G	GAGACGGTACCTTAAAGCTATCCACGAATGTCAAAAATGTGGTTT
10mermis1G	GAGAGTGTACCCAATGCTTGCAGTATGTATCCTGATCGTGCGTGC
10mermis1A	GAGAATGTACCCTCATACCAATGTAAAGTATAGTTAAGGCCCTGT
10mermis1T	GAGATTGTACCCTACATATATAGGAAAAGGGAAGGTAGAAGAGCT

10

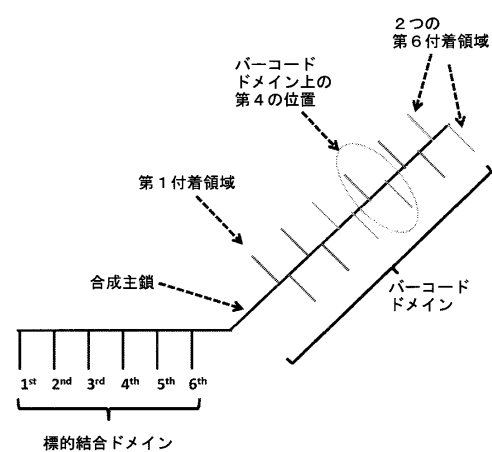
【図 1】

図1



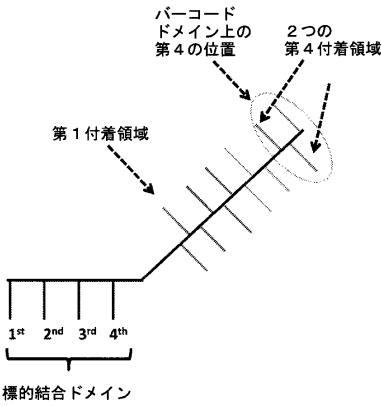
【図 2】

図2



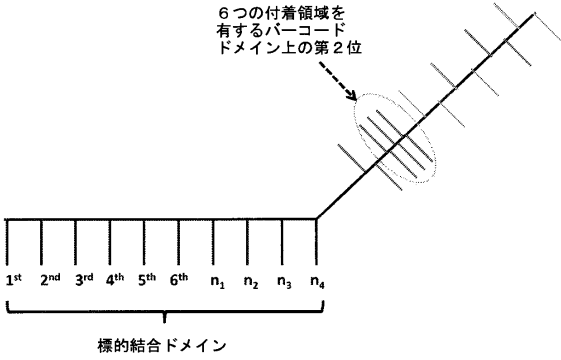
【 図 3 】

図3



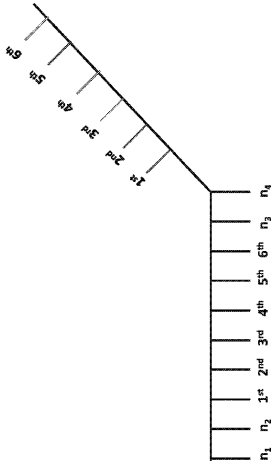
【 図 4 】

図4



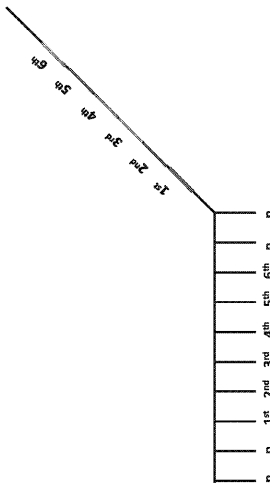
【 図 5 】

Figure 5



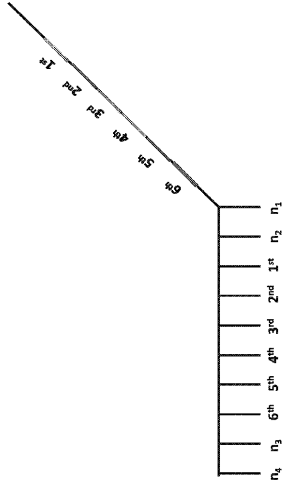
【 図 6 A 】

Figure 6A



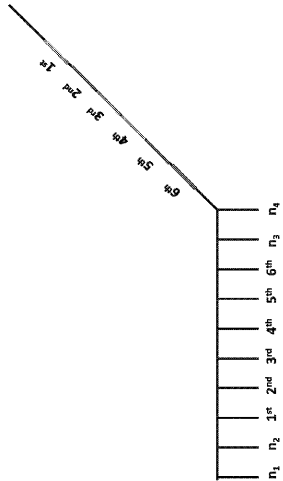
【図 6 B】

Figure 6B



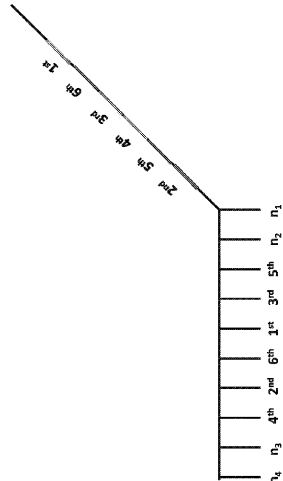
【図 6 C】

Figure 6C



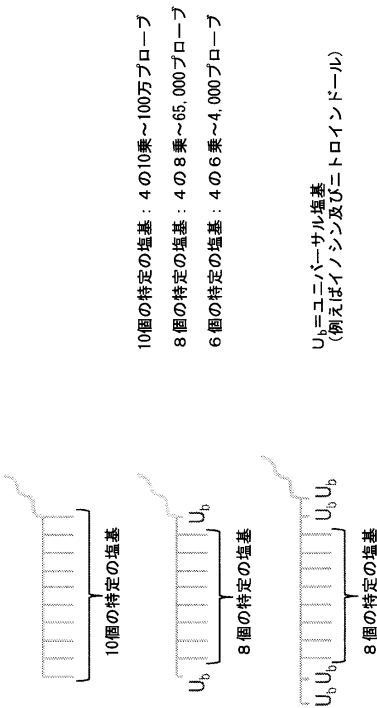
【図 6 D】

Figure 6D



【図 7】

図7



【図 8 A】

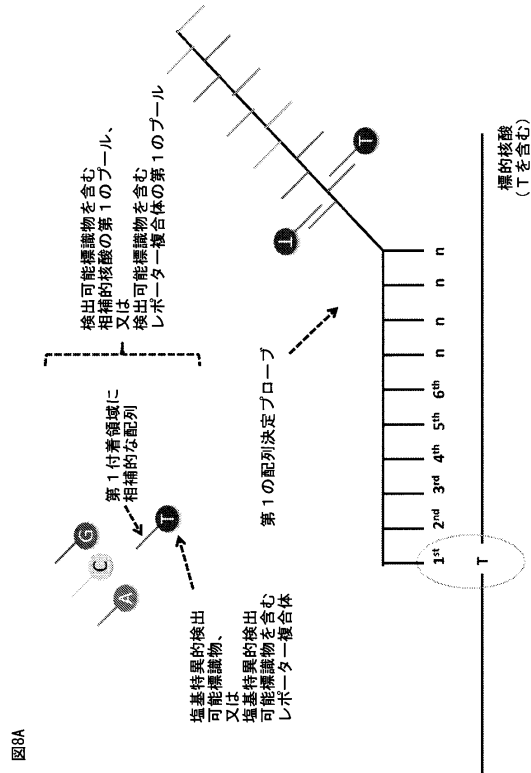


図8A

【図 8 B】

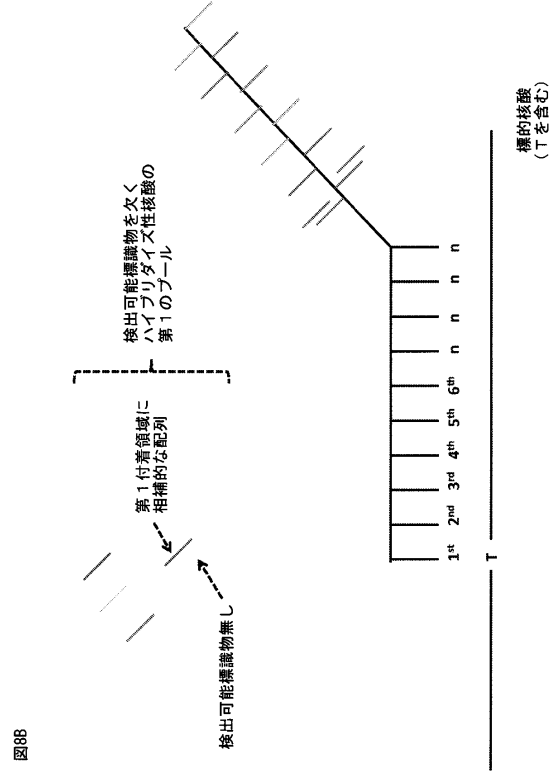


図8B

【図 8 C】

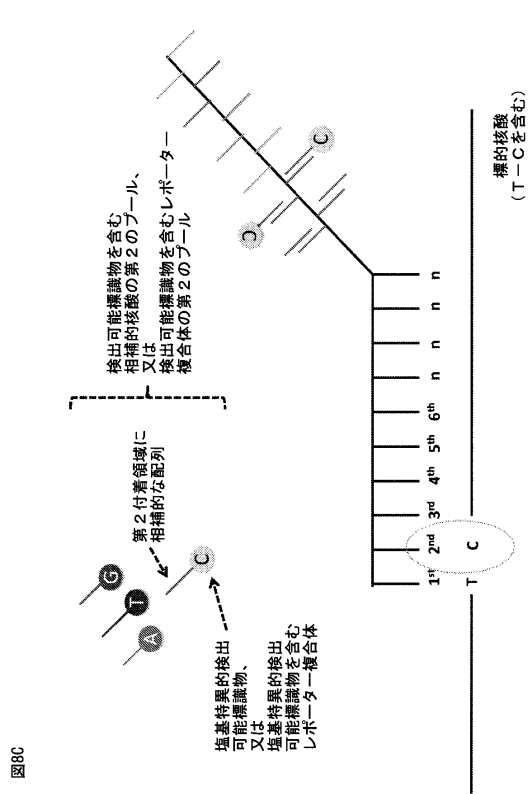


図8C

【図 8 D】

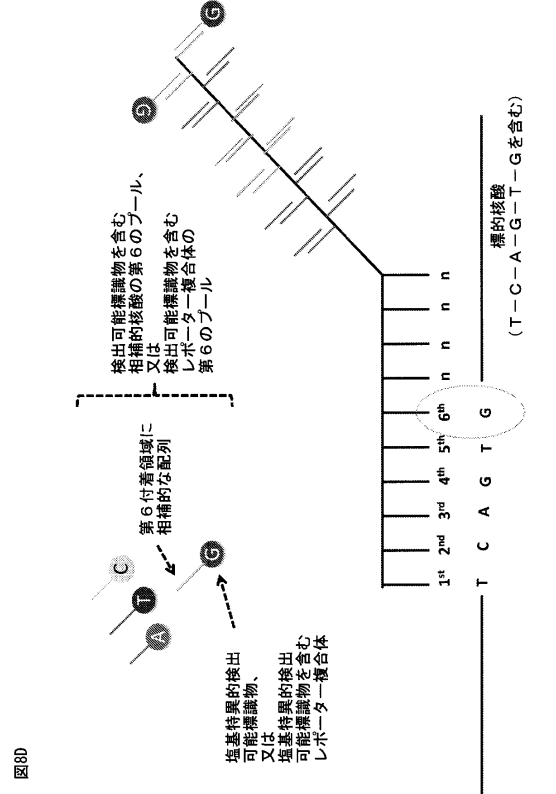
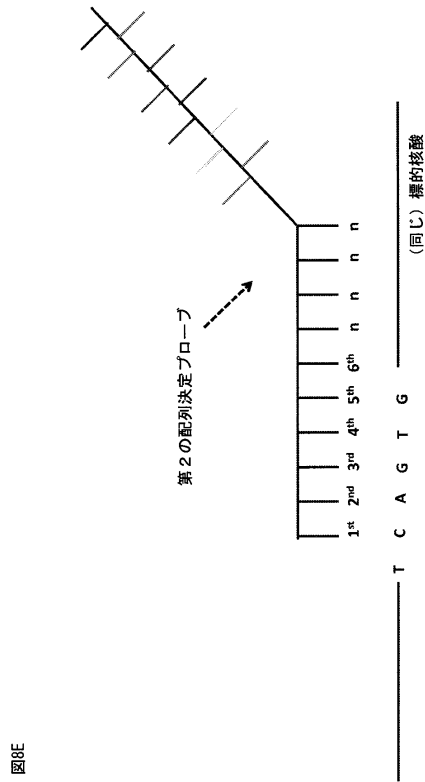
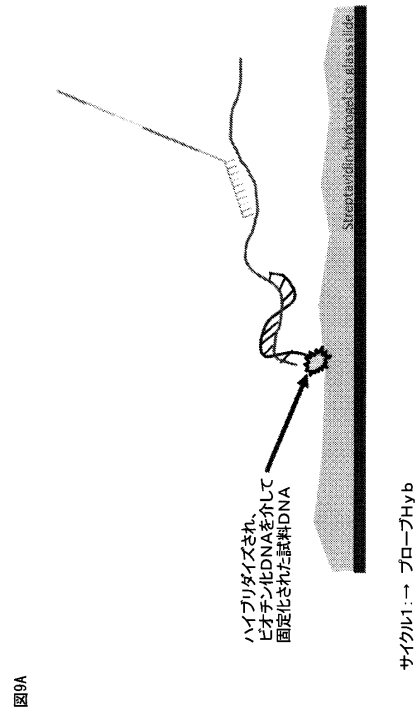


図8D

【図 8 E】

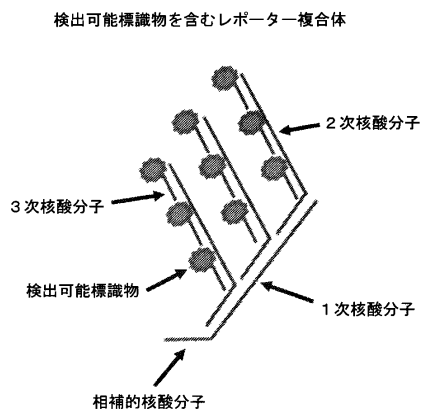


【図 9 A】

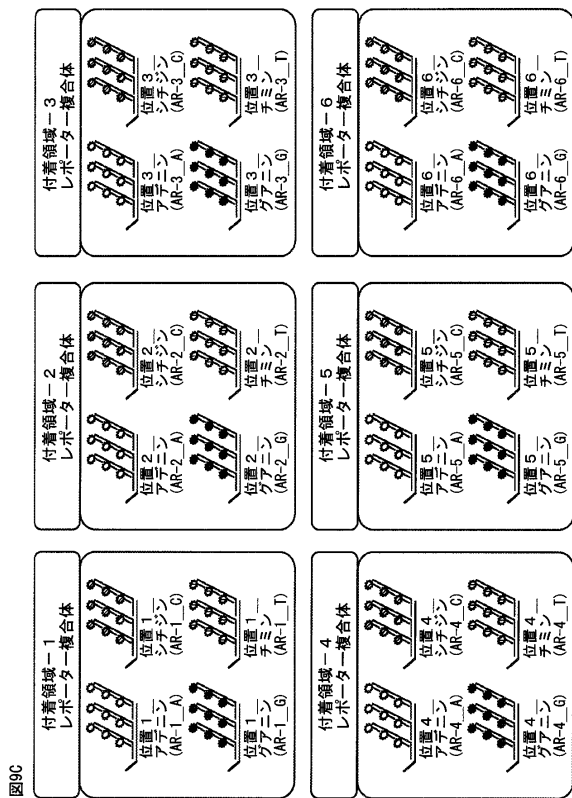


【図 9 B】

図9B

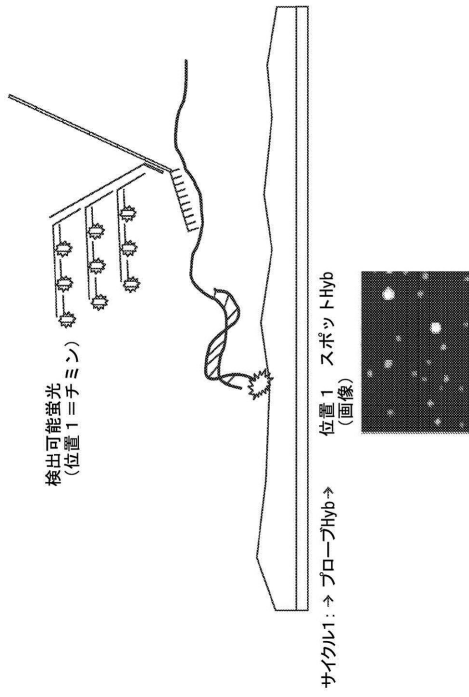


【図 9 C】



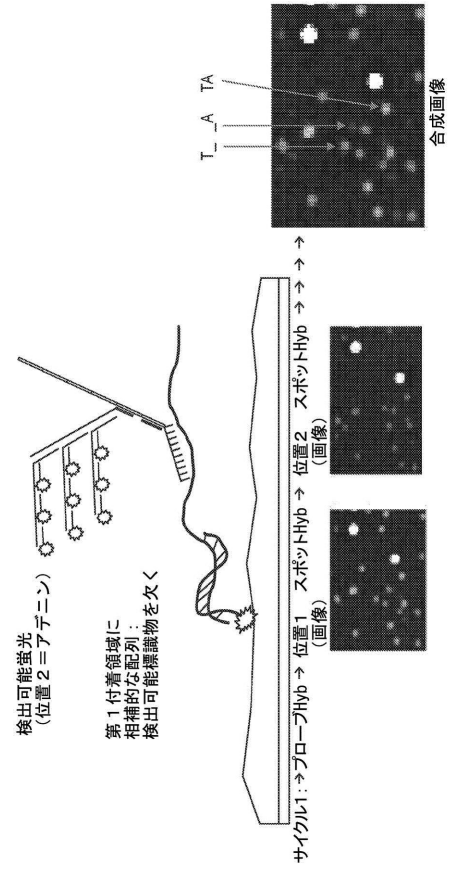
【図 9 D】

図9D



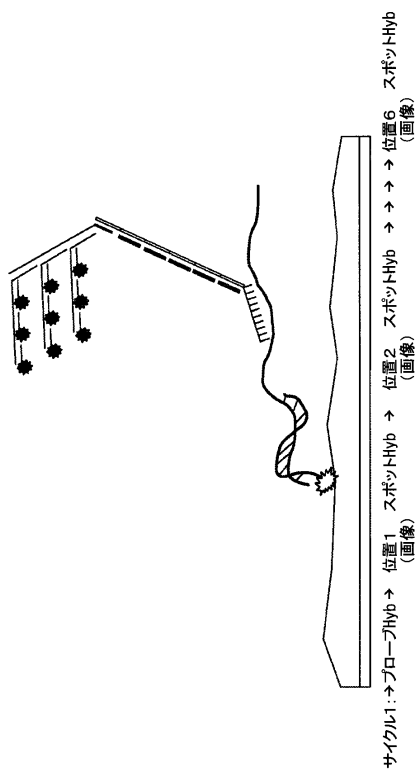
【図 9 E】

図9E



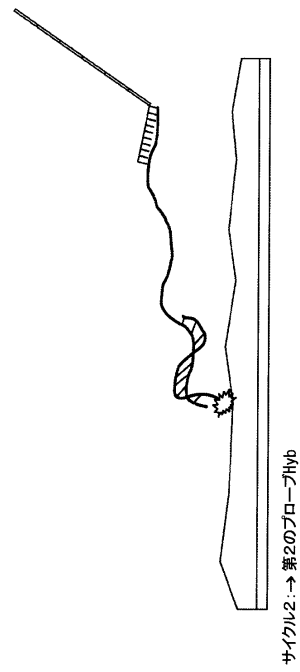
【図 9 F】

図9F

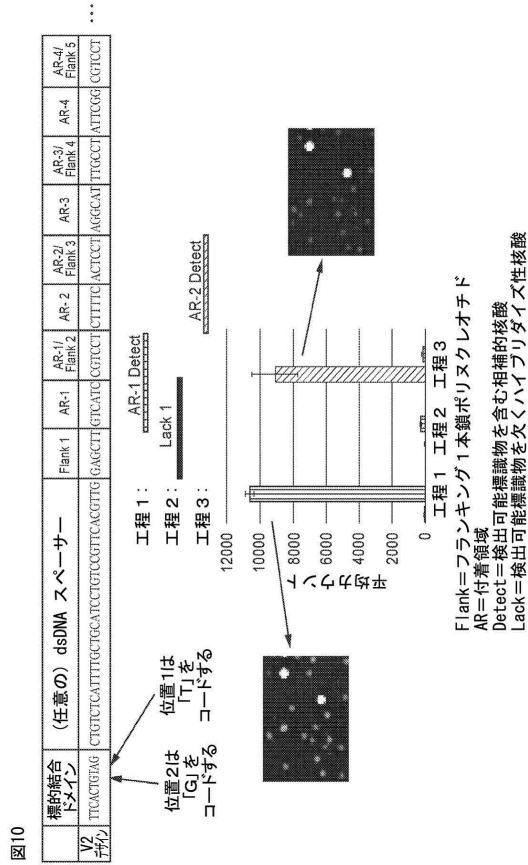


【図 9 G】

図9G

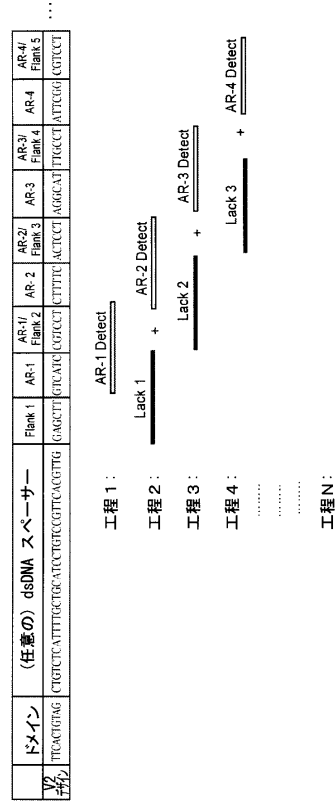


【図 1 0】

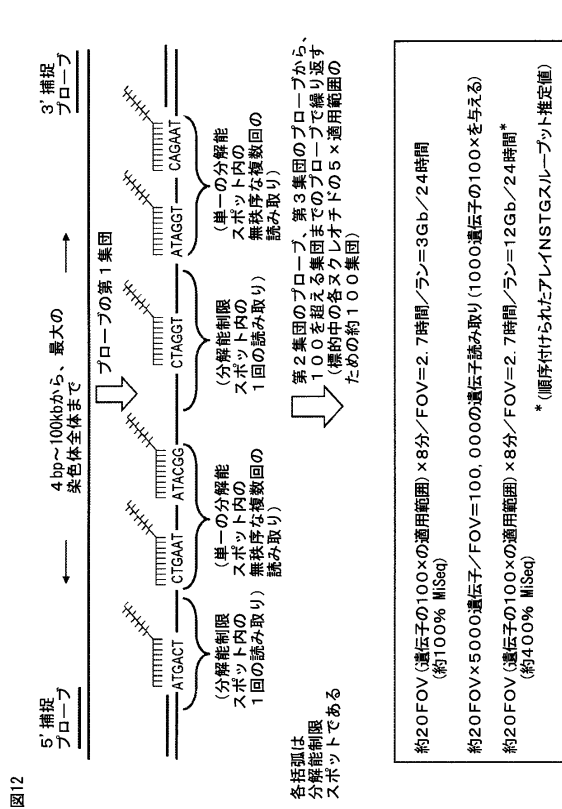


【図 1 1】

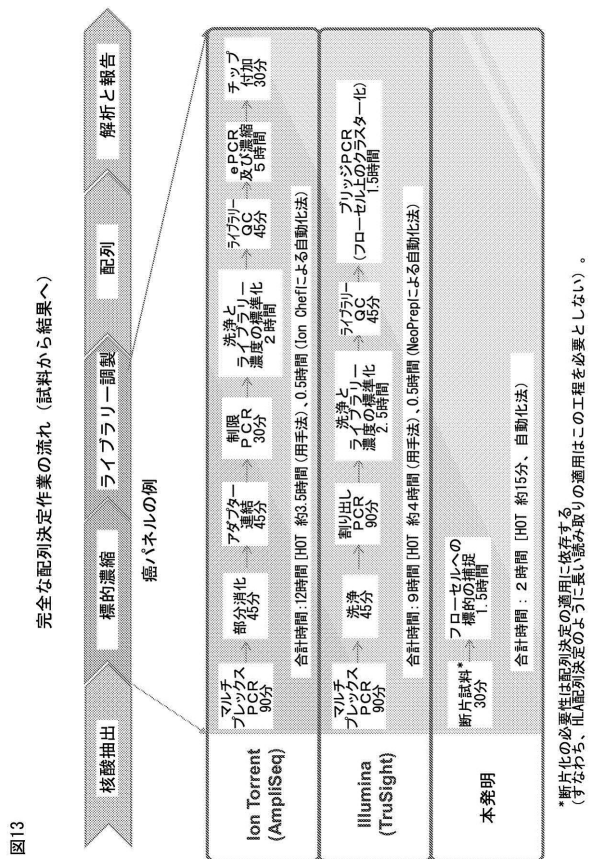
図11



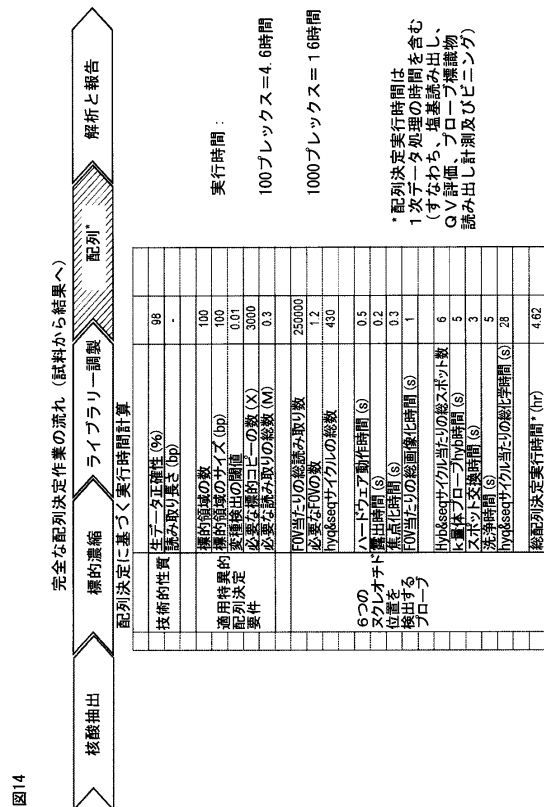
【図 1 2】



【図 1 3】



【 叉 1 4 】



【 ㊦ 1 6 】

ギガ塩基	時間 (hr)	読み取り数	臨床的有用性
HiSeq 2500	11 days	30億	ゲノム配列 癌配列
Next Seq	27hrs	4 億	ゲノム配列 エクソームトランスクリプトーム
Proton PI	2-4	7000万	エクソームトランスクリプトーム
本発明	3 to 12 Gb	50万	種特化 (Dx, dx) + マルチアナライズ+試料から 結果、長い読み取り
MSeq	24hrs	2500万	トランスクリプトーム 種特化 (Dx)
PGM 318	up to 2Gb	400万	種特化
PacBio RS II	2.2Gb/day	5 万	ニッチゲノム、長い読み取り、 高エラー率
PGM 314	0.1Gb	40万	種特化

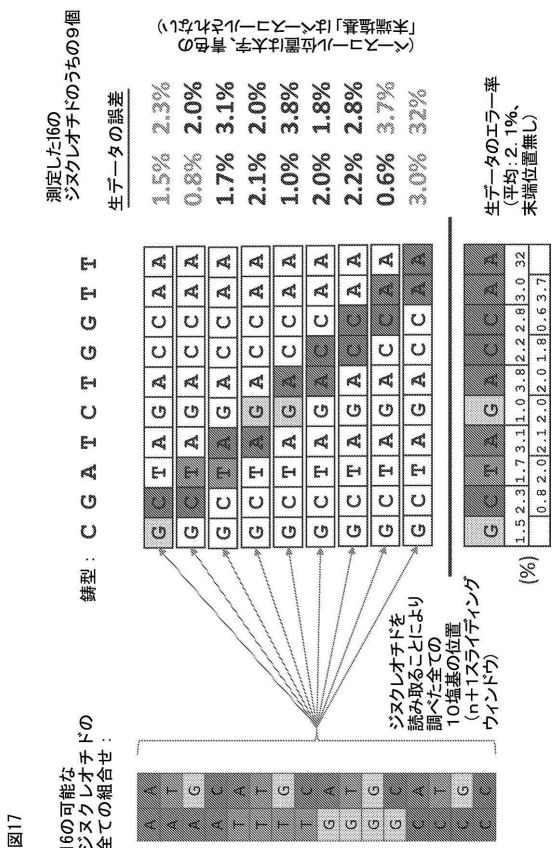
【 叉 1 5 】

图15

プローブ占有に基づくサイクリング計算

フローセル上の単一の分子標的の総数	30000
「ギャップのある」標的の平均長さ (bp)	100
配列決定サイクル当たりのバーコード占有率	0.75
フローセル上の95%超の標的を少なくとも1回「ヒット」するのに必要なバーコードサイクル数	3
位置当たりの色素占有率 (%)	0.8
色素位置の総数	6
配列決定サイクル当たりの1bpのコードされた読み取りの総数	6
配列決定サイクル当たりの2bpのコードされた読み取りの総数	23
配列決定サイクル当たりの3bpのコードされた読み取りの総数	92
配列決定サイクル当たりの4bpのコードされた読み取りの総数	369
配列決定サイクル当たりの5bpのコードされた読み取りの総数	1475
配列決定サイクル当たりの6bpのコードされた読み取りの総数	5898
配列決定サイクル当たりの「使用可能な」読み取りの総数	22499
特定の単一分子標的を約3×適用範囲でカバーするのに必要なバーコードサイクル数	143
特定の単一分子標的を約5×適用範囲でカバーするのに必要なバーコードサイクル数	238
フローセル上の全ての単一分子標的を約3×適用範囲でカバーするのに必要なバーコードサイクル数	430
フローセル上の全ての単一分子標的を約5×適用範囲でカバーするのに必要なバーコードサイクル数	713

【图 17】



【図18】

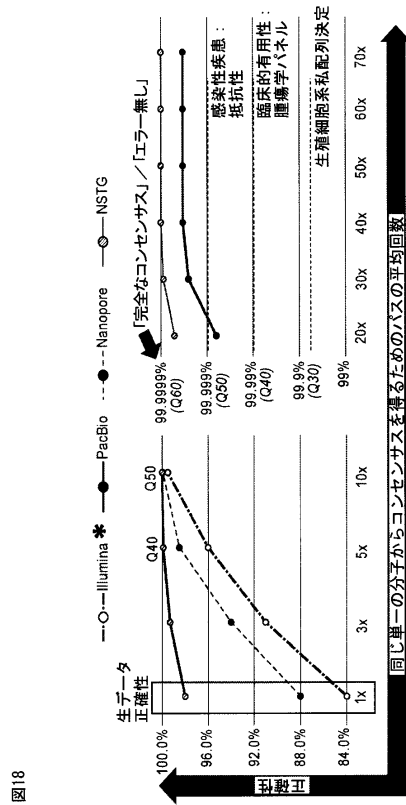


図18

【図20A】

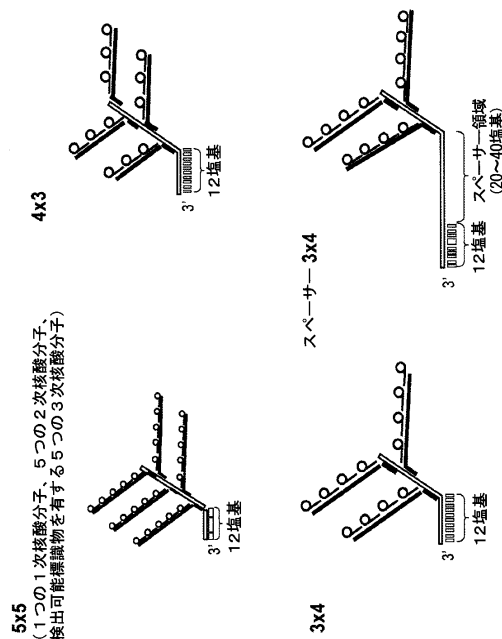


図20A

【図19】

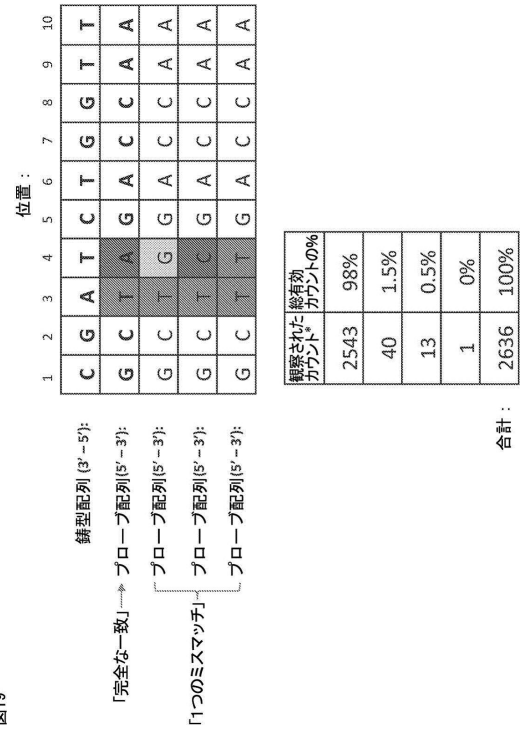
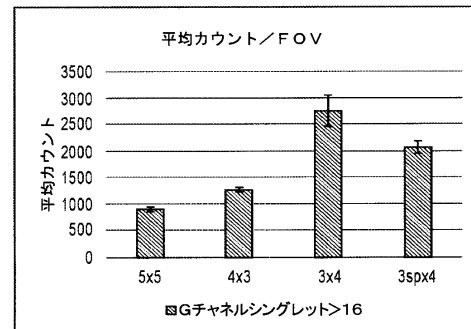


図19

*観察されたカウント値は、バックグラウンドカウントを引いた後のカウントを示す
**全てのハイブリダイゼーションは、緩衝液条件 (1xSSPE/0.1% ツイーン20) 下で行われた

【図20B】

図20B



【図20C】

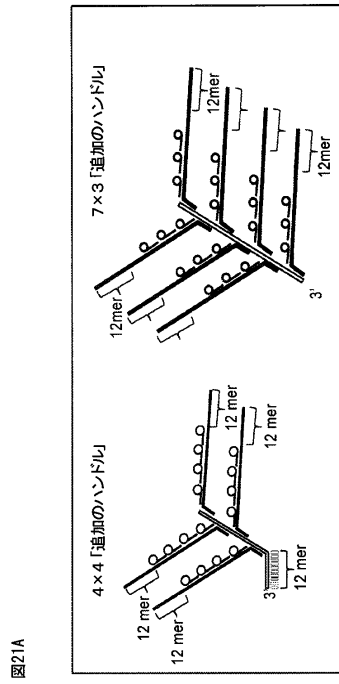
図20C

検出可能標識物を含むレポーター複合体を構築するための処方例

容量 (μ L)	1次核酸 分子 (10 μ M)	2次核酸 分子 (10 μ M)	3次核酸 分子 (100 μ M)	H2O
5x4	1	4.5	2.25	92.25
5x3	1	4.5	1.8	92.7
4x4	1.28	4.5	2.25	91.97
4x3	1.28	4.5	1.8	92.42
3x4	1.8	4.5	2.25	91.45

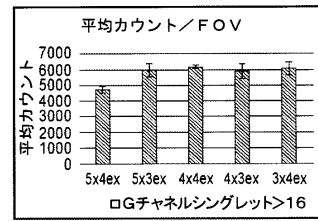
表中の色は、左の図に示された
複合体の図中の色に対応する

【図 2 1 A】

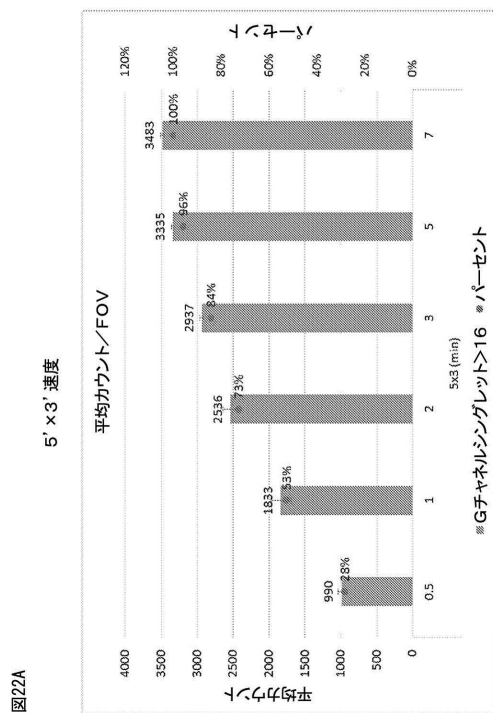


【図 2 1 B】

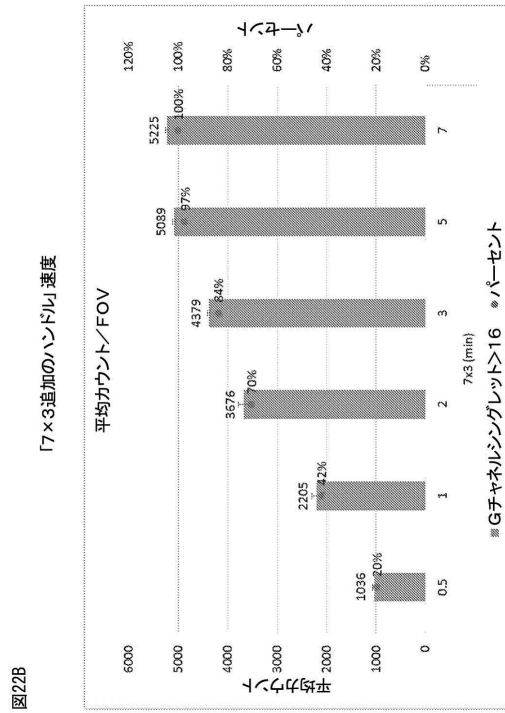
図21B



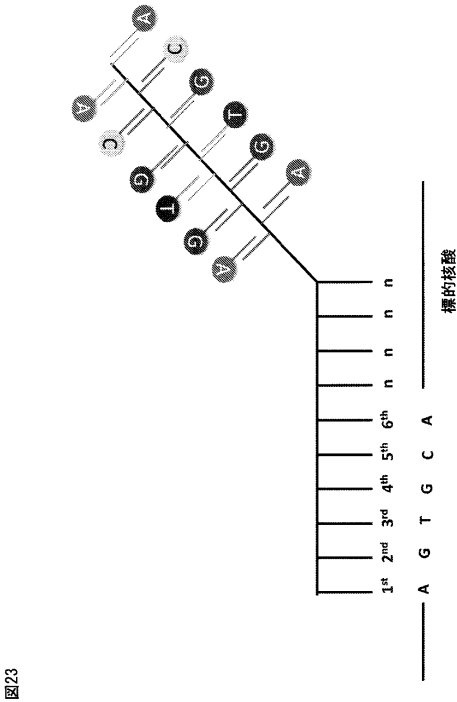
【図 2 2 A】



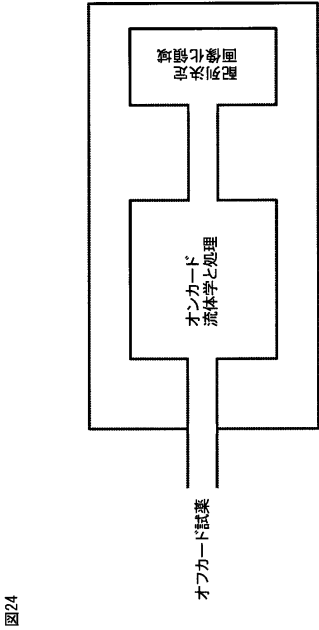
【図 2 2 B】



【 図 2 3 】



【 図 2 4 】



【 図 2 5 】

図25

コード	配列
YGBYGR	GAGACTGTAC - オリジナル
BGYRGB	GAGACAGTAC
YBRYGB	GAGACCGTAC
RGBYGR	GAGACGGTAC
YRBGYR	GAGAGTGTAC
GRYGRB	GAGAATGTAC
RGYBRG	GAGATTGTAC

コード	カウント	配列
YGBYGR	16074	GAGACTGTAC
RGYBRG	1132	GAGATTGTAC
YBRYGB	998	GAGACCGTAC
YRBGYR	784	GAGAGTGTAC
YBRYBG	712	
YBGRBY	699	
YGRYBG	692	
YBRGBR	674	
バックグラウンド 閾値	717	

【 図 2 6 】

図26

DNA	配列	コード	カウント
12 mer	CTGAGACTGTAC	RYGBRG	80845
11 mer	TGAGACTGTAC	GBRYGB	24474
10 mer	GAGACTGTAC	YGBYGR	18339
9 mer	AGACTGTAC	YRBYRG	1667
「12mer」の鏡像	CATGTCAGAGTC	GRBGYR	1473
「11mer」の鏡像	CATGTCAGAGT	BGYRBG	1384
8 mer	GA CTGTAC	GBRYBG	900
		バックグラウンド 閾値	622

【図 27】

図27

完全長レポーター
20 p M完全長 (6 スポット) 3 3 カウント/fov

単一のスポットレポーター 単一のスポット [fM]	カウント/FOV
20,000	飽和
2,000	12000
200	308
20	40
2	11

20 pM NSレポーター \approx 20 fM 単一のスポット

【配列表】

0006674951000001.app

フロントページの続き

(51)Int.Cl. F I
C 1 2 M 1/00 A

(74)代理人 100117019
弁理士 渡辺 陽一

(74)代理人 100150810
弁理士 武居 良太郎

(74)代理人 100134784
弁理士 中村 和美

(72)発明者 ジョーセフ エム・ピーチェム
アメリカ合衆国, オレゴン 97405, ユージーン, リッジライン ドライブ 46

(72)発明者 ルステム ハフィゾフ
アメリカ合衆国, ワシントン 98109, シアトル, マーサー ストリート 1200, アパートメント 212

審査官 藤井 美穂

(56)参考文献 国際公開第2013/055995(WO, A1)
特表2008-512129(JP, A)
特開2004-298082(JP, A)
国際公開第2013/102108(WO, A1)
特開2006-129866(JP, A)
特表2003-517283(JP, A)
米国特許第07745129(US, B1)

(58)調査した分野(Int.Cl., DB名)
C12N 15/00 - 15/90
C12Q 1/00 - 3/00
JSTPlus/JMEDPlus/JST7580(JDreamIII)
PubMed
CPlus/WPIDS/MEDLINE/EMBASE/BIOSIS(STN)