

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第6867045号
(P6867045)

(45) 発行日 令和3年4月28日 (2021.4.28)

(24) 登録日 令和3年4月12日 (2021.4.12)

(51) Int. Cl.

F I

C 1 2 Q 1/6869 (2018.01)

C 1 2 Q 1/6869 Z

G 1 6 B 30/20 (2019.01)

G 1 6 B 30/20

G 1 6 B 25/10 (2019.01)

G 1 6 B 25/10

C 1 2 M 1/00 (2006.01)

C 1 2 M 1/00 A

C 1 2 N 15/11 (2006.01)

C 1 2 N 15/11 Z

請求項の数 39 (全 35 頁)

(21) 出願番号 特願2018-506601 (P2018-506601)
 (86) (22) 出願日 平成28年8月12日 (2016.8.12)
 (65) 公表番号 特表2018-531583 (P2018-531583A)
 (43) 公表日 平成30年11月1日 (2018.11.1)
 (86) 国際出願番号 PCT/CN2016/094802
 (87) 国際公開番号 W02017/025059
 (87) 国際公開日 平成29年2月16日 (2017.2.16)
 審査請求日 令和1年7月25日 (2019.7.25)
 (31) 優先権主張番号 62/204,396
 (32) 優先日 平成27年8月12日 (2015.8.12)
 (33) 優先権主張国・地域又は機関
 米国 (US)

(73) 特許権者 512037244
 ザ チャイニーズ ユニバーシティ オブ
 ホンコン
 中華人民共和国 ホンコン ニュー テリ
 トリーズ シャーティン ピチウ ビルデ
 イング ルーム 301 オフィス オブ
 リサーチ アンド ナレッジ トランス
 ファー サービスイズ
 (74) 代理人 100099759
 弁理士 青木 篤
 (74) 代理人 100123582
 弁理士 三橋 真二
 (74) 代理人 100117019
 弁理士 渡辺 陽一

最終頁に続く

(54) 【発明の名称】 血漿DNAの単分子配列決定

(57) 【特許請求の範囲】

【請求項 1】

対象から得られた生物試料を準備し、ここで、当該生物試料はセルフリーDNA断片を含む体液を含み、

当該生物試料から複数のセルフリーDNA断片を抽出し、

前記複数のセルフリーDNA断片の第1のセットをコンカテマー化して第1のコンカテマーを得、ここで、当該第1のセットのセルフリーDNA断片は、異なった配列を有し、対象のゲノムでの異なった位置由来であり、前記第1のコンカテマーの単分子配列決定を行って、第1の配列を得ること、ここで、当該第1の配列は、当該第1の配列内のサブシーケンスとして、複数のセルフリーDNA断片の第1のセットの異なった配列を含む、
 を含む、方法。

10

【請求項 2】

前記単分子配列決定の実行の一部として、配列決定装置に前記第1のコンカテマーを提供し、

前記配列決定装置を用いて、前記第1のコンカテマーに対応する複数の信号、前記第1のコンカテマーの前記第1の配列に対応する複数の信号を検出すること、
 をさらに含む、請求項1に記載の方法。

【請求項 3】

前記配列決定装置が第1のナノポアを含み、方法は、さらに、
 前記第1のコンカテマーが前記第1のナノポアを通り、

20

前記第 1 のコンカテマーが前記第 1 のナノポアを通るときに第 1 の電気信号を検出すること、ここで、前記第 1 の電気信号は前記第 1 のコンカテマーの前記第 1 の配列に対応し、前記複数の信号は前記第 1 の電気信号を含む、を含む、請求項 2 に記載の方法。

【請求項 4】

前記第 1 の配列を決定するために前記第 1 の電気信号を分析し、

参照ゲノムに対して前記第 1 の配列のサブシーケンスを整列させて、前記複数のセルフリー DNA 断片の前記第 1 のセットそれぞれに対応する断片配列を同定すること、をコンピュータシステムにより実行することをさらに含む、請求項 3 に記載の方法。

【請求項 5】

前記第 1 のコンカテマーが前記第 1 のナノポアを通ることは、

前記第 1 のコンカテマーの第 1 の鎖が前記第 1 のナノポアを通ること、及び、

その後、前記第 1 のコンカテマーの第 2 の鎖が前記第 1 のナノポアを通ること、を含む、請求項 3 に記載の方法。

【請求項 6】

前記第 1 の鎖および前記第 2 の鎖についての前記第 1 の電気信号をコンピュータシステムにより分析して、前記第 1 の配列を決定することをさらに含む、請求項 5 に記載の方法。

【請求項 7】

前記第 1 のナノポアは、基体上の複数のナノポアの 1 つである、請求項 3 に記載の方法。

【請求項 8】

前記複数のセルフリー DNA 断片の第 2 のセットをコンカテマー化して第 2 のコンカテマーを得、

前記第 2 のコンカテマーが前記複数のナノポアの第 2 のナノポアを通り、

前記第 2 のコンカテマーが前記第 2 のナノポアを通るときに複数の第 2 の電気信号を検出すること、ここで、前記第 2 の電気信号は前記第 2 のコンカテマーの第 2 の配列に対応する、

をさらに含む、請求項 7 に記載の方法。

【請求項 9】

前記第 1 のコンカテマーに対して蛍光標識されたヌクレオチドをハイブリダイズし、蛍光シグナルを検出すること、ここで、前記蛍光シグナルはヌクレオチドに対応する、をさらに含む、請求項 2 に記載の方法。

【請求項 10】

参照ゲノムに対して前記第 1 の配列のサブシーケンスを整列させて、前記セルフリー DNA 断片の前記第 1 のセットそれぞれに対応する断片配列を同定することを、コンピュータシステムにより実行することをさらに含む、請求項 1 に記載の方法。

【請求項 11】

前記サブシーケンスの整列に基づいて、前記複数のセルフリー DNA 断片の前記第 1 のセットそれぞれのサイズを決定することをさらに含む、請求項 10 に記載の方法。

【請求項 12】

前記複数のセルフリー DNA 断片の前記第 1 のセットをコンカテマー化することは、前記複数のセルフリー DNA 断片の前記第 1 のセットおよび既知の配列を有する前記複数のセルフリー DNA 断片の第 2 のセットをコンカテマー化することを含み、前記複数のセルフリー DNA 断片の前記第 2 のセットは前記複数のセルフリー DNA 断片の前記第 1 のセットの中に分散され、

前記第 1 の配列のサブシーケンスを整列させることは、前記既知の配列に対してサブシーケンスを整列させて、前記第 1 のコンカテマーにおける前記複数のセルフリー DNA 断片の前記第 2 のセットの位置を同定することを含む、請求項 10 に記載の方法。

【請求項 13】

10

20

30

40

50

前記コンカテマー化は、前記複数のセルフリーＤＮＡ断片の前記第２のセットの１つのセルフリーＤＮＡ断片を、前記複数のセルフリーＤＮＡ断片の前記第１のセットの２つのセルフリーＤＮＡ断片間に配置する、請求項１２に記載の方法。

【請求項１４】

前記第１のサブシーケンスの前に、前記複数のセルフリーＤＮＡ断片の前記第２のセットの既知の配列の１つを同定することに基づいて、前記複数のセルフリーＤＮＡ断片の前記第１のセットの第１のセルフリーＤＮＡ断片に対応する第１のサブシーケンスの開始塩基を、前記コンピュータシステムにより同定し、

前記第１のサブシーケンスの後に、前記複数のセルフリーＤＮＡ断片の前記第２のセットの既知の配列の１つを同定することに基づいて、前記複数のセルフリーＤＮＡ断片の前記第１のセットの前記第１のセルフリーＤＮＡ断片に対応する前記第１のサブシーケンスの終了塩基を、前記コンピュータシステムにより同定すること、
を前記コンピュータシステムにより実行することをさらに含む、請求項１３に記載の方法。

【請求項１５】

前記複数のセルフリーＤＮＡ断片の前記第２のセットの各セルフリーＤＮＡ断片は７ヌクレオチド以下の既知の配列を有する、請求項１２に記載の方法。

【請求項１６】

前記複数のセルフリーＤＮＡ断片の前記第２のセットそれぞれは、同じ既知の配列を有する、請求項１２に記載の方法。

【請求項１７】

前記サブシーケンスの整列は、

前記参照ゲノムに前記第１の配列のスライドウィンドウを整列させることであって、各スライドウィンドウは前記参照ゲノムに整列されたサブシーケンスに対応する、整列させることと、

２つのスライドウィンドウが前記参照ゲノムの異なる領域に整列するとき同定して、前記複数のセルフリーＤＮＡ断片の前記第１のセットの２つのセルフリーＤＮＡ断片の配列間を区別すること、

とを含む、請求項１０に記載の方法。

【請求項１８】

前記参照ゲノムの前記異なる領域に整列させる２つのスライドウィンドウに基づいて、前記複数のセルフリーＤＮＡ断片の前記第１のセットの第１のセルフリーＤＮＡ断片の終了および開始の第１のサブシーケンスを決定することを、前記コンピュータシステムにより実行することをさらに含む、請求項１７に記載の方法。

【請求項１９】

前記参照ゲノムの前記異なる領域は、異なる染色体上の領域を含む、請求項１７に記載の方法。

【請求項２０】

前記２つのスライドウィンドウ間の１つ以上のウィンドウが、前記参照ゲノムに整列していないものとして同定される、請求項１７に記載の方法。

【請求項２１】

前記複数のセルフリーＤＮＡ断片の前記第１のセットの第１のセルフリーＤＮＡ断片のサイズを決定することをさらに含み、前記第１のセルフリーＤＮＡ断片の前記サイズを決定することは、

前記参照ゲノムの１つの領域に整列させる最も長いサブシーケンスの長さを決定することを含む、請求項２０に記載の方法。

【請求項２２】

前記生体試料は、血漿または血清である、請求項１に記載の方法。

【請求項２３】

前記セルフリーＤＮＡ断片の前記第１セットが複数の染色体からランダムに分布されて

10

20

30

40

50

いる、請求項 1 に記載の方法。

【請求項 2 4】

前記複数のセルフリー DNA 断片の第 2 のセットをコンカテマー化して第 2 のコンカテマーを得、ここで、当該第 2 のセットのセルフリー DNA 断片は、前記第 1 のコンカテマーとは異なるセルフリー DNA 断片のコンビネーション又はパーミュテーションを含み、ここで、複数のセルフリー DNA 断片の第 2 のセットをコンカテマー化することは、複数のセルフリー DNA 断片の第 1 のセットをコンカテマー化することと並行して起こる、

前記第 2 のコンカテマーの単分子配列決定を行って、第 2 の配列を得ること、ここで、当該第 2 の配列は、当該第 2 の配列内のサブシーケンスとして、複数のセルフリー DNA 断片の第 2 のセットの異なった配列を含む、

10

をさらに含む、請求項 1 に記載の方法。

【請求項 2 5】

前記第 1 のコンカテマーが、複数のセルフリー DNA 断片の複製された DNA 断片を含まず、前記第 2 のコンカテマーが、複数のセルフリー DNA 断片の複製された DNA 断片を含まない、請求項 2 4 に記載の方法。

【請求項 2 6】

前記第 1 のコンカテマーが、複数のセルフリー DNA 断片の第 1 のセットのセルフリー DNA 断片間に既知の配列を有するスペーサー断片を含まない、請求項 1 に記載の方法。

【請求項 2 7】

第 1 の配列内のサブシーケンスを用いて複数のセルフリー DNA 断片のサイズ分布を決定することを更に含む、請求項 1 に記載の方法。

20

【請求項 2 8】

前記生物試料が、母親セルフリー DNA 断片及び胎児セルフリー DNA 断片を含む母親血漿であり、そこから、複数のセルフリー DNA 断片の少なくとも第 1 のセットが得られる、請求項 1 に記載の方法。

【請求項 2 9】

(1) 第 1 のコンカテマーを生成するセルフリー DNA 断片の第 1 のセットをコンカテマー化し、(2) 第 1 のコンカテマーの単一分子配列決定を行って、第 1 の配列を得ることにより生成された第 1 のコンカテマーの第 1 の配列を準備し、ここで、

前記セルフリー DNA 断片の第 1 のセットは、生物試料から抽出され、

30

前記生物試料は対象から得られ、

前記生物試料は、前記セルフリー DNA 断片を含む体液を含み、

前記第 1 の配列のサブシーケンスを整列させて、前記セルフリー DNA 断片の前記第 1 のセットの各セルフリー DNA 断片に対応する断片配列を同定すること、
をコンピュータシステムにより実行することを含む、方法。

【請求項 3 0】

前記第 1 のコンカテマーは、前記 DNA 断片の前記第 1 のセットと DNA 断片の第 2 のセットとをコンカテマー化することにより生成され、

前記第 1 の配列のサブシーケンスを整列させることは、前記 DNA 断片の前記第 2 のセットにサブシーケンスを整列させることを含む、請求項 2 9 に記載の方法。

40

【請求項 3 1】

前記セルフリー DNA 断片の前記第 2 のセットの各セルフリー DNA 断片は、7ヌクレオチド以下の既知の配列を含む、請求項 3 0 に記載の方法。

【請求項 3 2】

前記第 1 の配列のサブシーケンスを整列させることは、参照ゲノムに前記第 1 の配列のサブシーケンスを整列させることを含む、請求項 2 9 に記載の方法。

【請求項 3 3】

前記サブシーケンスの前記整列は、

前記参照ゲノムに前記第 1 の配列のウィンドウを整列させることと、

前記参照ゲノムの異なる領域に 2 つのウィンドウを整列させるときに同定することとを

50

含む、請求項 3 2 に記載の方法。

【請求項 3 4】

前記 2 つのウィンドウ間の 1 つ以上のウィンドウが前記参照ゲノムに整列していないものとして同定される、請求項 3 3 に記載の方法。

【請求項 3 5】

請求項 2 9 ~ 3 4 のいずれか 1 項に記載の操作を実行するために前記コンピュータシステムを制御するための複数のインストラクションを保存する、コンピュータ読み取り可能媒体。

【請求項 3 6】

請求項 3 5 のコンピュータ読み取り可能媒体と、
前記コンピュータ読み取り可能媒体に保存されるインストラクションを実行するための 1 つ以上のプロセッサと
を含む、システム。

【請求項 3 7】

請求項 2 9 ~ 3 4 のいずれか 1 項を実行するための手段を含む、システム。

【請求項 3 8】

請求項 2 9 ~ 3 4 のいずれか 1 項を実行するために構成される、システム。

【請求項 3 9】

請求項 2 9 ~ 3 4 のいずれか 1 項に記載のステップをそれぞれ実行するモジュールを含む、システム。

【発明の詳細な説明】

【技術分野】

【0001】

関連出願の相互援用

本出願は、2015 年 8 月 12 日に提出された、米国特許仮出願第 62 / 204,396 号の利益を主張するものであり、その内容は、全ての目的のために、参照により本明細書に組み込まれる。

【背景技術】

【0002】

母体血漿 DNA 配列決定による非侵襲的出生前検査 (NIPT) は、現在、胎児染色体異数性のスクリーニングに臨床的に利用可能である (1)。羊水穿刺とは異なり、母体血漿 DNA 配列決定はいかなる流産リスクも引き起こさない。これらの試験は、99% 近くの感度および 99% 近くの特異性を有する (1)。その結果、NIPT の臨床的需要は、2011 年に初めて市販されて以来、大幅に増加している。

【0003】

超並列配列決定は、染色体異数性の NIPT のために現在使用されている実験プロトコルの大部分の中心的構成要素である (2)。高い計装コストのため、これらの試験は現在、参考検査室で実施されている。オックスフォードナノポアテクノロジーは、ナノポアベースの DNA 配列決定プラットフォームを開発した (3)。ナノポア配列決定装置は、設備コストが比較的安く、フットプリントが小さい。各フローセルのコストは 500 ~ 900 米ドルで、48 時間まで複数回使用できる。配列決定速度も比較的速く、各ナノポアから 30 塩基 / 秒を読み取ることができる。このような特徴は、臨床検査室での使用に有利であろう。しかし、現行のナノポア技術は、NIPT に典型的に使用される試料 (例えば、血漿) に対して非効率である。

【発明の概要】

【0004】

概要

実施形態は、比較的小さな DNA 断片で試料を分析する場合の単分子配列決定技術の効率を改善する。例えば、試料中の DNA 断片の濃度を有意に増加させることができ、それにより、より多くの DNA 断片を配列決定装置 (例えば、ナノポア) と相互作用させるこ

10

20

30

40

50

とができる。別の例として、DNA断片をコンカテマーに結合させて、より長い分子を読み取ることができ、それにより、1分子の配列を介して複数のDNA断片（すなわち、試料中にもともとあるDNA断片）を効率的に読み取ることができる。パイオインフォーマティックス手法を用いて、同じコンカテマーの一部である異なるDNA試料を検出することができる。実施形態は、2つの技術を組み合わせることができる。

【0005】

実施形態は、核酸配列を決定する方法を含むこととしてもよい。方法は、複数のDNA断片を受け取ることを含むこととしてもよい。また、方法は、DNA断片の第1のセットをコンカテマー化して、第1のコンカテマーを得ることを含むこととしてもよい。方法は、第1のコンカテマーの単分子配列決定を実行して、第1のコンカテマーの第1の配列を得ることを含む。いくつかの実施形態では、単分子配列決定は、ナノポアを用いて実行され、方法は、第1のコンカテマーが第1のナノポアを通ることを含むこととしてもよい。第1の電気信号は、第1のコンカテマーが第1のナノポアを通るときに検出されることとしてもよい。第1の電気信号は、第1のコンカテマーの第1の配列に対応することとしてもよい。

10

【0006】

別の実施形態は、核酸配列を決定する方法を含むこととしてもよい。方法は、複数のDNA断片を受け取ることを含むこととしてもよい。DNA断片の第1のセットがコンカテマー化されて第1のコンカテマーを得ることとしてもよい。蛍光標識されたヌクレオチドは、コンカテマーへハイブリダイズされることとしてもよい。第1の蛍光シグナルは、特異的ヌクレオチドに対応する第1の蛍光シグナルで検出されることとしてもよい。蛍光標識は、その後、切断され、別の蛍光標識されたヌクレオチドが添加され、プロセスが繰り返されることとしてもよい。

20

【0007】

実施形態は、コンピュータシステムにより実行される方法を含むこととしてもよい。方法は、DNA断片の第1のセットをコンカテマー化することにより生成された第1のコンカテマーの第1の配列を受け取ることを含むこととしてもよい。また、方法は、第1の配列のサブシーケンスを整列させて、DNA断片の第1のセットの各DNA断片に対応する断片配列を同定することを含むこととしてもよい。

【0008】

いくつかの実施形態は、細胞を含まないDNA断片を配列決定する方法を含むこととしてもよい。細胞を含まないDNA断片は、血漿DNA断片を含むこととしてもよい。方法は、複数のDNA断片を含む生体試料を受け取ることを含むこととしてもよい。生体試料は、DNA断片の第1の濃度を有することとしてもよい。方法は、また、DNA断片の第2の濃度を有するように、対応する生体試料を濃縮することを含むこととしてもよい。DNA断片の第2の濃度は、DNA断片の第1の濃度より5倍以上高いこととしてもよい。方法は、さらに、複数のDNA断片が基体上のナノポアを通ることを含むこととしてもよい。複数のDNA断片のそれぞれについて、電気信号は、DNA断片がナノポアを通るときに検出されることとしてもよい。電気信号は、DNA断片の配列に対応することとしてもよい。

30

40

【0009】

実施形態は、細胞を含まないDNA断片の配列決定の方法を含むこととしてもよい。方法は、DNA断片の最初の濃度よりも5倍以上高いDNA断片の第2の濃度を有するように、対応する生体試料を濃縮することを含むこととしてもよい。方法は、さらに、単分子配列決定技術を含むこととしてもよい。DNA断片は、蛍光標識されたヌクレオチドによりハイブリダイズされることとしてもよい。また、方法は、蛍光標識されたヌクレオチドからの信号を、ヌクレオチドに対応する信号で検出することを含むこととしてもよい。蛍光標識されたヌクレオチドは切断されることとしてもよく、プロセスは、追加のヌクレオチドを同定することと、DNA断片の配列を繰り返すこととしてもよい。

【0010】

50

また、実施形態は、複数のインストラクションを保存するコンピュータ読み取り可能媒体を含むコンピュータ製品を含み、本明細書に記載されるDNA配列決定のあらゆる方法の操作を実行することを含むこととしてもよい。いくつかの実施形態は、コンピュータ読み取り可能媒体に保存されるインストラクションを実行するために、1つ以上のプロセッサおよびコンピュータ製品を含む。追加の実施形態は、いずれかの方法を実行するためのシステムを含む。

【0011】

他の実施形態は、本明細書に記載される方法に関連するシステム、ポータブル消費装置、およびコンピュータ読み取り可能媒体を対象とする。

【0012】

本発明の実施形態の性質および利点のより良い理解は、以下の詳細な説明および添付の図面を参照することによって得ることができる。

【図面の簡単な説明】

【0013】

【図1A】図1Aは、本発明の実施形態に係るナノポアデバイスおよび核酸の簡略図を示す。

【図1B】図1Bは、本発明の実施形態によるDNA断片をコンカテマー化するプロセスを示す。

【図2】図2は、本発明の実施形態によるDNA断片とスペーサーDNA断片とのコンカテマー化プロセスを示す。

【図3A】図3Aは、DNA断片をコンカテマー化し、本発明の実施形態による単分子配列決定を使用することによってDNA断片を配列決定する方法の簡略ブロックフロー図を示す。

【図3B】図3Bは、本発明の実施形態によるナノポア配列決定を使用してDNA断片を配列決定する方法の簡略ブロックフロー図を示す。

【図4】図4は、本発明の実施形態によるコンカテマーの配列を分析する方法の簡略ブロックフロー図を示す。

【図5】図5は、本発明の実施形態によるDNA断片の濃度を複数回増加させることによってDNA断片をより効率的に配列決定する方法の簡略ブロックフロー図を示す。

【図6】図6は、本発明の実施形態によるナノポア配列決定装置によって配列決定された血漿DNAプールのサイズ分布を示す。0～500塩基対の範囲の配列決定された血漿DNA断片の頻度分布をプロットする。

【図7】図7は、本発明の実施形態による、ナノポア配列決定およびイルミナ配列決定プラットフォームによって得られたデータからの女性胎児を伴う母体血漿由来の血漿DNAのサイズプロファイルを示す。

【図8】図8は、本発明の実施形態によるマッピング可能なヒトゲノムから予想される分布と比較した、染色体の読み取り分布を示す。ナノポア配列の比例分布は、試料プールごとに各染色体に読み取る。塗りつぶされた灰色のバーは、参照ヒトゲノムhg19のマッピング可能な部分に基づいて、それぞれのヒト染色体に由来するヌクレオチドの割合を表す。残りの色の棒グラフは、血漿DNA試料のそれぞれのヒト染色体に整列した配列決定された読み取りの割合を表す。

【図9】図9は、本発明の実施形態による、ナノポア配列決定および超並列配列決定を使用する癌患者からの血漿DNA配列決定の結果のCircosプロットを示す。

【図10】図10は、本発明の実施形態によるコンカテマー化DNA分子のサイズ分布を示す。

【図11】図11は、本発明の実施形態による、非妊娠女性由来のコンカテマー化セグメントに由来する血漿DNA分子のサイズ分布を示す。

【図12】図12は、本発明の実施形態による非妊娠女性由来の連結血漿DNA由来の整列セグメントのゲノム表現を示す。

【図13】図13は、本発明の実施形態による男性胎児をみごもる妊娠女性由来のコンカ

10

20

30

40

50

テマー化セグメントに由来する血漿DNA分子のサイズ分布を示す。

【図14】図14は、本発明の実施形態による、男性胎児をみごもる妊娠女性由来の濃縮された連結血漿DNA由来の整列セグメントのゲノム表現を示す。

【図15】図15は、本発明の実施形態に従って実行するためのシステムのブロック図を示す。

【図16】図16は、本発明の実施形態によるシステムおよび方法と共に使用可能な例示的なコンピュータシステムのブロック図を示す。

【発明を実施するための形態】

【0014】

用語

10

「組織」は、機能単位として一緒に群を成す細胞群に対応する。1つの組織内に2つ以上のタイプの細胞を見出すことができる。異なるタイプの組織は、異なるタイプの細胞（例えば、肝細胞、肺細胞または血液細胞）からなることができるが、異なる生物（母親または胎児）由来の組織または健常細胞対腫瘍細胞に対応してもよい。

【0015】

「生体試料」とは、対象（例えば、妊婦、癌を有する者、または癌を有すると疑われる者などのヒト、臓器移植レシピエント、または臓器を伴う疾患プロセス（例えば、心筋梗塞の心臓、または脳卒中の脳、または貧血の造血系）を有すると疑われる対象から採取され、関心のある1つまたは複数の核酸分子を含む任意の試料を指す。生体試料は、血液、血漿、血清、尿、膿液、水晶体（例えば精巣）からの流体、膣洗浄液、胸水、腹水、脳脊髄液、唾液、汗、涙、痰、気管支肺胞洗浄液、乳頭からの排出液、身体の異なる部分（例えば、甲状腺、乳房）などからの吸引液などが挙げられる。便試料も使用することができる。様々な実施形態では、細胞を含まないDNA（例えば、遠心分離プロトコルによって得られた血漿試料）が濃縮された生体試料中の大部分のDNAは細胞を含まない、例えば、50%、60%、70%、80%、90%、95%、または99%のDNAは細胞を含まないであり得る。遠心分離プロトコルは、例えば、3,000g×10分を含み、流体部分を得、そして例えば30,000gでさらに10分間遠心分離して残留細胞を除去することができる。試料中の細胞を含まないDNAは、様々な組織の細胞に由来することができ、従って、試料は細胞を含まないDNAの混合物を含むことができる。

20

【0016】

「核酸」は、一本鎖または二本鎖のいずれかの形態のデオキシリボヌクレオチドまたはリボヌクレオチドおよびそのポリマーを意味し得る。この用語は、既知のヌクレオチド類似体または修飾された骨格残基または結合を含む核酸を包含し、それは、合成であり、天然に存在し、天然に存在せず、参照核酸と同様の結合特性を有し、それらは参照ヌクレオチドと同様の態様で代謝される。そのような類似体の例は、ホスホロチオエート、ホスホラミダイト、メチルホスホネート、キラルメチルホスホネート、2-O-メチルリボヌクレオチド、ペプチド-核酸（PNA）を含むが、これらに限定されない。

30

【0017】

特に明記しない限り、特定の核酸配列はまた、その保存的に改変された改変体（例えば、縮重コドン置換）および相補的配列、ならびに明示的に示される配列を暗黙に包含する。具体的には、縮重コドン置換は、1つ以上の選択された（またはすべての）コドンの第3位が混合塩基および/またはデオキシイノシン残基で置換された配列を生成することによって達成され得る。（Batzer et al., Nucleic Acid Res. 19: 5081 (1991) ; Ohtsuka et al., J. Biol. Chem. 260: 2605-2608 (1985) ; Rossolini et al., Mol. Cell. Probes 8: 91-98 (1994)）用語核酸は、遺伝子、cDNA、mRNA、オリゴヌクレオチドおよびポリヌクレオチドと互換的に使用される。

40

【0018】

天然に存在するリボヌクレオチドまたはデオキシリボヌクレオチドモノマーに言及することに加えて、用語「ヌクレオチド」は、文脈上他に明白に示さない限り、そのヌクレオチドが使用される特定の状況（例えば、相補的塩基へのハイブリダイゼーション）に関し

50

て機能的に同等であるその関連する構造的変異体（誘導体および類似体を含む）を指すと理解され得る。

【 0 0 1 9 】

「コンカテマー」は、単分子に結合された別個のDNA断片からなる連続DNA分子である。コンカテマーの別々のDNA断片の様々なものは、同じ配列を有していてもいなくてもよい。コンカテマー中のDNA断片の少なくともいくつかは、異なる配列を有し得る。コンカテマーを作製するために使用される別々のDNA断片は、例えば、DNA断片が細胞を含まないDNA断片である場合、血漿および他の細胞を含まないDNA混合物中に存在し得るように、生体試料中に存在する様々な組織に由来し得る。

【 0 0 2 0 】

「配列読み取り」は、核酸分子の任意の部分または全部から配列決定されたヌクレオチドの列をいう。例えば、読み取られる配列は、生体試料中に存在する核酸断片全体であってもよい。配列読み取りは、単分子配列決定から得ることができる。「単分子配列決定」とは、鋳型DNA分子のクローンコピーからの塩基配列情報を解釈する必要なしに配列読み取りを得るための単一鋳型DNA分子の配列決定を指す。単分子配列決定は、分子全体またはDNA分子の一部のみを配列決定することができる。DNA分子の大部分は配列決定され、例えば、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、または99%以上である。

【 0 0 2 1 】

配列決定の場合、「信号」はある瞬間にとられた測定を指すことができる。このような信号の例には、光信号または電気信号が含まれる。光信号は、例えば、ハイブリダイゼーションが起こる場合に、特定の塩基に対応する特定の色に対応する画像を提供することができる。単一の画像は、配列決定装置のアレイ（例えば、ナノポアまたは他の配列決定容積）であってもよく、それにより、単一の画像が多くの光信号を有する。電気信号は、瞬時に電極を横切る測定値とすることができる。ある期間の電気信号は、DNA分子の配列中に1つ以上の塩基を提供することができる。電気信号は、電流信号または電圧信号を含むことができる。

【 0 0 2 2 】

「ナノポア」は、ナノポア内にある分子の部分の1つまたは複数の特性に基づいてシグナルが検出され得る、分子または分子の一部が配置され得る開口部を指す。ナノポアは、ポリマー、金属または他の固体材料、タンパク質、またはそれらの組み合わせなどの様々な材料から構成することができる。

【 0 0 2 3 】

「分類」とは、試料の特定の特性に関連する任意の数または他の文字をいう。例えば、「+」記号（または「陽性」という語）は、試料が欠失または増幅を有するものとして分類されていることを示すことができる。分類は、バイナリ（例えば、正または負）であってもよく、またはより多くの分類レベル（例えば、1~10または0~1のスケール）を有してもよい。用語「カットオフ」および「閾値」は、操作において使用される所定の数を指す。例えば、カットオフサイズは、断片が除外されるサイズを指す。閾値は、特定の分類が適用される値を上回るか下回る値であってもよい。これらの用語のどちらも、これらのコンテキストのいずれでも使用できる。

【 0 0 2 4 】

本明細書で使用される「染色体異数性」という用語は、二倍体ゲノムの染色体量からの染色体量の変化を意味する。変動は獲得または喪失であることとしてもよい。これは、1つの染色体全体または染色体の領域を含むこととしてもよい。

【 0 0 2 5 】

本明細書で使用する「配列不均衡」または「収差」という用語は、参照量からの臨床的に関連する染色体領域の量における少なくとも1つのカットオフ値によって定義されるような有意な偏差を意味する。配列不均衡は、染色体投与不均衡、対立遺伝子不均衡、突然変異投与不均衡、コピー数不均衡、ハプロタイプ投与不均衡、および他の類似の不均衡を

10

20

30

40

50

含み得る。例として、対立遺伝子の不均衡は、腫瘍が欠失された遺伝子の1つの対立遺伝子または遺伝子の1つの対立遺伝子をそのゲノム中の2つの対立遺伝子の増幅または異なる増幅を有する場合起こり得、これにより、試料中の特定の遺伝子座に不均衡を生じさせる。別の例として、患者は腫瘍サプレッサー遺伝子に遺伝的突然変異を有する可能性がある。次いで、患者は、腫瘍抑制遺伝子の非突然変異対立遺伝子が欠失した腫瘍を発症し得る。したがって、腫瘍内には突然変異の投薬量の不均衡がある。腫瘍がそのDNAを患者の血漿に放出すると、腫瘍DNAは血漿中の患者の構成DNA（正常細胞由来）と混合される。本明細書に記載の方法の使用により、血漿中のこのDNA混合物の突然変異投薬不均衡を検出することができる。収差は、染色体領域の欠失または増幅を含み得る。

【0026】

10

用語「サイズプロファイル」は、一般的に、生体試料中のDNA断片のサイズに関する。サイズプロファイルは、様々なサイズのDNA断片の量の分布を提供するヒストグラムであり得る。1つのサイズプロファイルを別のサイズプロファイルと区別するために、様々な統計的パラメータ（サイズパラメータまたは単なるパラメータとも呼ばれる）を使用することができる。1つのパラメータは、特定のサイズまたは範囲のサイズのDNA断片の、すべてのDNA断片または他のサイズまたは範囲のDNA断片に対する割合である。

【0027】

発明の詳細な説明

実施形態は、約200塩基の比較的短い断片である生体試料（例えば、血漿または血清中）中の細胞を含まないDNA断片に適用される場合、単分子配列決定の効率を改善することができる。血漿DNA断片などの細胞を含まないDNA断片は、小さくまたは短く、典型的には生体試料に低濃度で存在するため、細胞を含まないDNA断片を配列決定するためのナノポアの使用は、正確な結果をもたらすとは考えられない。小さな細孔を通過する小さな断片は、断片がナノポア内にある間に断片を配列決定するのを困難にする。さらに、小さな断片では、予想される配列決定誤差を考慮すると、参照ゲノムへの整列がより困難になるかもしれない。ナノポア配列決定は、約10～15%の配列決定誤差を有し得る。いくつかの実施形態において、効率は、配列決定されるべきより長い分子を生成するためのDNA断片のコンカテマー化によって改善され得る。他の実施形態では、効率は、試料中のDNA断片の濃度を増加させることによって改善することができる。

20

【0028】

30

これらの開発がなければ、血漿DNA上の単分子配列決定は実用的ではないかもしれない。これらのアプローチでは、一連の血漿DNA異常が単分子配列決定によって検出可能であった。実施形態は、出生前診断、癌評価、炎症性疾患管理、自己免疫疾患評価、および外傷などの急性医薬を含む、医学における多くの領域に適用することができる。さらに、メチル化シトシンは、単分子配列決定によって非メチル化シトシンと区別できることが報告されている（12）。我々は以前に、血漿DNAのメチル化プロファイルを検出することにより、胎児DNA画分を決定し、異常な胎盤メチル化プロファイルを持つ妊娠関連疾患を検出し、癌および全身性エリテマトーデスに関連する異常メチル化を検出することができたことを報告した（7、9、11）。従って、実施形態はまた、メチル化検出の用途にも使用することができる。

40

【0029】

I. DNA断片の単分子配列決定

血漿または血清DNAは、ヒト対象の循環系に見出され、細胞死のプロセスの間に自然な代謝回転または病理学的プロセスの一部として放出される細胞を含まない核酸分子である。DNA分子は細胞分解プロセスの一部として循環中に放出されるので、短い断片（<200bp）の形で循環し、低濃度で存在する。健康な対象の血漿DNAの大部分は、血液細胞に由来することが知られている（6）。

【0030】

妊娠中、胎盤は血漿DNA分子を母体循環に与え、非侵襲的な出生前診断のために胎児DNAにアクセスする手段を提供する（7）。腫瘍や癌は細胞回転率が高く、DNAを血

50

漿中に与え、染色体や遺伝子異常が非侵襲的に検出され、癌の診断、モニタリング、スクリーニング、予後診断のための液体生検として役立つ（８、９）。心筋梗塞、脳卒中、肝炎などの炎症状態もまた、炎症を起こした器官からの血漿DNAの付与を増加させる（１０）。全身性エリテマトーデスなどの自己免疫疾患も、血漿DNA異常と関連している（１１）。上記の疾患において、血漿DNAプロファイルは、疾患、染色体または染色体のコピー数異常（コピー数の増加または減少、異数性）、異常なメチル化（過剰メチル化または低メチル化）およびサイズプロファイル異常（余分な量の短いDNA分子または余分な量の長いDNA分子）に関連する単一ヌクレオチド変異体を示し得る。要約すると、循環する細胞を含まないDNA分析は、広範囲の疾患に対する分子診断試験の開発にとって重要な手段となっている。

10

【００３１】

オックスフォードナノポアテクノロジーによって製造されたフォーマットに限定されない単分子配列決定は、各DNA塩基を同定することができる高速度のために、DNA配列決定分析のための魅力的なプラットフォームである。増幅工程が省略されているため、ライブラリ構築のワークフローは単純化されている。これは、数百キロベースまでのDNAの連続した伸長を配列することができる。しかしながら、そのような利点は、血漿DNA分析に容易には適用できない。第一に、血漿DNA分子は、ほとんどが＜２００bpの長さの短い断片である。血漿および血清中のDNA濃度は、組織生検におけるDNA濃度よりも実質的に低い。したがって、血漿DNA試料をナノポア配列決定装置に適用すると、配列決定効率は低かった。言い換えれば、試料中のDNAが希薄であり、血漿DNA分子があまりにも頻繁にナノポアに向かっていることが分かった。血漿DNA分子がナノポアに到達し、その配列長が短いために配列決定されたとしても、配列決定情報はヒトゲノムに関して極めて少量の情報しか提供しなかった。一倍体ヒトゲノムは、サイズが 3.3×10^9 塩基である。さらに、ナノポアを用いた配列決定エラー率は、短いDNA断片の配列決定を、長いDNA断片を配列決定する場合よりも正確さが低いと予測する。したがって、我々は、血漿DNA分子がナノポアおよび他の単分子配列決定装置によって処理され得る頻度またはチャンスを増加させるアプローチを開発することを目的とした。

20

【００３２】

A．ナノポア配列決定

ナノポア配列決定は、DNA塩基がナノポアの使用によって検出される単分子配列決定装置の一形態である。オックスフォードナノポアテクノロジーは、タンパク質細孔、-溶血素を使用している。このような細孔のマトリックスは、膜上に位置するように作製される（５）。タンパク質細孔に加えて、ナノポアは、ナノポアが、窒化ケイ素およびグラフェンなどのケイ素化合物を含む半導体材料で製造される固体状態のナノポアであり得る。各ポアは、電気回路に接続されてもよい。

30

【００３３】

図１Aは、本技術の実施形態によるナノポア１０２およびDNA分子１０４の簡略図を示す。電極１０６および電極１０８は、ナノポアの一部を画定してもよく、またはナノポアの近くに配置されてもよい。電極１０６および電極１０８は、円錐形または三角形の端部を有するものとして示されている。いくつかの実施形態では、電極は異なる形状を有してもよい。例えば、電極は、平坦な端部、半球の端部、または丸みのある端部を有することができる。両方の電極は同じ形状でなくてもよい。例えば、一方の電極は円錐形の端部を有し、他方の電極は平坦であってもよい。電極間の距離は、ナノポアの直径または幅と等しいか、小さいか、または大きくすることができる。電極１０６および電極１０８は、電源１１０に接続することができる。電流１１２は、電極１０６から電極１０８にトンネルすることができる。電源１１０は、複数のナノポアおよびそれぞれの電極の対と電気的に連絡していてもよい。

40

【００３４】

DNA分子１０４がナノポア１０２を通過するとき、電流の変化があり、これは、メータ１１４によって測定することができる。異なるDNA塩基、すなわちA、C、G、Tは

50

、電流変化において異なる大きさを誘発するであろう。各ナノポアの電圧または電流パターンを観察することにより、ナノポア102を通過したDNA分子104の配列を決定することができた。配列決定はDNAの単分子上のDNA塩基を検出するのに十分に感受性である、すなわちDNA配列決定ライブラリの増幅を必要としないので、配列検出の速度は実質的に向上する。

【0035】

しかしながら、欠点は、DNA塩基同定の精度が、イルミナ配列決定のような各DNA分子の増幅クローンからのコンセンサス配列を検出する配列決定技術よりも相対的に劣ることである。しかしながら、2D読み取りが解釈される場合、ベース検出の精度は実質的に改善されることが示されている(3)。ナノポア配列決定のためのDNA試料を調製する場合、各二本鎖DNA分子の一端にヘアピンアダプターを加える。そのようなDNA分子がナノポアに近づくと、二本鎖ははがれ、現在の一本鎖DNA分子の一端は、現在の変化が検出されたときにナノポアを通過する。配列決定がこのシングルエンドの終わりに近づいているとき、ヘアピンによって連結された相補鎖はナノポアを通過し続け、配列決定される。DNA分子の両方の鎖が配列決定される場合、コンセンサス配列が導かれ得、これは2D読み取りと呼ばれる。文献に報告されているように、2D読み取りは、1D読み取り(1本の鎖のみから解釈される配列)より高いベースコール精度を有する。

【0036】

B. 他の単分子配列決定

本実施形態はまた、ナノポアを使用する技術以外の単分子配列決定技術を含み得る。例えば、Helicos単分子配列決定装置(SeqLL)を用いて、DNA分子をガラス表面上にハイブリダイズさせることができる。次いで、蛍光標識されたヌクレオチドを各DNA分子に添加し、画像を捕捉することができる。次いで、蛍光分子を切断して洗い流し、別の蛍光標識ヌクレオチドを添加して、このプロセスを繰り返すことができる。各ヌクレオチドは異なる蛍光標識を有することができ、DNAを配列決定することができる。

【0037】

別の例は、Pacific Biosciences単分子リアルタイム(SMRT)配列決定法を含むことができる。この方法では、DNAポリメラーゼ酵素をゼロモード導波路(ZWM)の底に固定することができる。次いで、ポリメラーゼは、単分子のDNAを捕捉することができる。次いで、蛍光標識されたヌクレオチドをDNA-酵素複合体に組み込むことができる。次いで、検出器が蛍光シグナルを検出し、ベースコールを行うことができる。次いで、蛍光タグを切断し、ZWMから拡散させることができる。このプロセスは、各タイプのヌクレオチドが異なる蛍光標識を有し、DNAの配列決定を可能にするように繰り返してもよい。

【0038】

II. 増殖効率

効率を高めるために、1つの選択肢は、配列決定チャンバまたはフローセルに適用される有限体積内の試料中の遺伝物質の量を増加させるために、血漿DNAプールを増幅することである。しかしながら、このようなプロトコルの採用は、元のDNA鋳型と複製されたDNA断片の全てを配列決定するため、単分子配列決定をもはや実行しないことを意味する。複製されたDNA断片のそのような使用は、元のDNA断片についての定量的情報、例えば染色体全体のような所与のゲノム領域における元のDNA断片のコピー数を得るときにエラーを引き起こす可能性がある。

【0039】

以下のセクションでは、効率を高めるための2つのアプローチについて説明する。1つのアプローチでは、単分子配列決定を実施する前にDNA断片を連結してコンカテマーを形成することができる。別のアプローチでは、DNA断片は、単分子配列決定の前に(例えば、ナノポア配列決定の前に)試料中に高度に濃縮される。

【0040】

A. コンカテマー

いくつかの実施形態では、図 1 B に示すように、DNA 断片（例えば、血漿 DNA 断片）をコンカテマー化してもよい。コンカテマー化は、一連の DNA 断片を互いに結合させることができる。例えば、DNA 断片 152 および DNA 断片 154 は、試料中に存在し得る。DNA 断片を互いに付着させるための最初の手順として、DNA 断片 152 および DNA 断片 154 の末端を、末端にリン酸基を付加して調製して、DNA 断片 156 および DNA 断片 158 を得ることができる。次いで、DNA 断片を平滑末端ライゲーションでリガーゼ酵素によって一緒に連結して、長い DNA 分子、例えばコンカテマー 160 を形成することができる。コンカテマー 160 は、DNA 断片 152 と 154 の組み合わせである新しい分子と考えることができる。コンカテマー 160 は、DNA 断片 152 および 154 に対応するサブシーケンスの組み合わせである配列を有する。

10

【0041】

図 2 は、DNA 断片間にスペーサー断片（「スペーサー」とも呼ばれる）を有するコンカテマーを形成する同様のアプローチを示す。DNA 断片 202 および DNA 断片 204 は、「A-テリング」を受けることができ、Aヌクレオチドを DNA 断片の各鎖の一端に付加して、DNA 断片 206 および DNA 断片 208 を形成することができる。スペーサー DNA 断片 210 は、スペーサー DNA 断片の各鎖の一端に Tヌクレオチドを有する既知の配列を有することができる。スペーサー DNA 断片の既知の配列は、スペーサー断片の末端に付加された A または Tヌクレオチドを含まない、4 ~ 10 塩基対および 10 ~ 20 塩基対を含む 20 塩基対以下であってもよい。スペーサー DNA 断片および血漿 DNA 断片の末端には、他の相補的ヌクレオチドも同様に使用することができる。次いで、スペーサー DNA 断片 210 は、リガーゼ酵素によって DNA 断片 206 および DNA 断片 208 に連結され、コンカテマー 212 として示される長い DNA 分子を形成し得る。スペーサー DNA 断片は、スペーサーではない 2 つの DNA 断片の間に置くことができる。例えば、1 つのスペーサー DNA 断片は、生体試料から抽出された 2 つの DNA 断片の間にあってもよい。コンカテマー 212 は、1 つ以上のスペーサー DNA 断片 210 と共に DNA 断片 206 および 208 に対応する配列の組み合わせである配列を有し得る。

20

【0042】

いくつかの実施形態では、スペーサー断片の既知の配列が、対象に対応する参照ゲノム（例えば、ヒトゲノム）に現れない場合、または特定の回数未満（例えば、2 または 3 未満）で現れる場合、スペーサー断片は、1 つの DNA 断片の末端および別の DNA 断片の開始を示すことができる。コンピュータシステムは、サブシーケンスを、スペーサー断片に使用される 1 つまたは複数の既知の配列の予想されるセットと比較することによって、スペーサー断片の配列を同定することができる。いくつかの実施形態では、スペーサー断片は、1 つの位置での単なる A - T の組み合わせであり得る。そのような実施例では、A および T の使用は、DNA 断片の開始および終了時の識別とは対照的に、主に付着のためであり得る。

30

【0043】

スペーサーを有する方法は、生体試料の DNA 断片に対応するサブシーケンスの開始塩基を同定することを含むことができる。開始塩基の同定は、スペーサー DNA 断片の既知の配列の 1 つを同定することに基づくことができる。サブシーケンスの末端は、サブシーケンスの後のスペーサー DNA 断片の既知の配列の 1 つを同定することに基づいて同定することができる。

40

【0044】

これらのアプローチでは、長い分子（コンカテマー）がナノポアに達すると、コンカテマーが単一の長い分子内にいくつかの血漿 DNA 断片を含むことができるので、多くの血漿 DNA 断片が配列決定され得る。血漿 DNA 分子は天然で二本鎖であり得るので、ヘアピンアダプターをコンカテマーの末端に適用して、2D 読み取りが生成され得るようにする。すなわち、両方の鎖が読み取られる。配列決定の後、コンカテマーに組み込まれた元々の血漿 DNA 分子の単位を同定することができる。いくつかの DNA 断片から長い分子を作製し、長い分子をナノポア内で配列決定することは、同一のいくつかの分子が別々に

50

ナノポアに移動して配列決定されるのを待つよりも効率的である。

【0045】

1. 単分子配列決定における使用

図3Aは、DNA断片をコンカテマー化し、本発明の実施形態による単分子配列決定を使用することによってDNA断片を配列決定する方法300の簡略ブロックフロー図を示す。

【0046】

ブロック302において、方法300は、複数のDNA断片を受け取ることを含むことができる。複数のDNA断片は、生体試料由来の細胞を含まないDNA断片であり得る。生体試料は、血漿または血清であり得る。DNA断片は、例えば、血漿が血液の他の成分から分離されるとき、生体試料の他の成分から分離されてもよい。DNA断片は、本明細書中に記載される任意のDNA断片であり得る。

10

【0047】

ブロック304において、方法300は、第1のコンカテマーを得るためにDNA断片の第1のセットをコンカテマー化することを含むことができる。コンカテマー化は、本明細書中に記載される任意の方法によるものであり得る。第1のコンカテマーは、第1のセット以外のDNA断片を含み得る。したがって、第1の組のDNA断片は、第1のコンカテマーを構成するDNA断片の全てではない可能性がある。第1のコンカテマー中の他のDNA断片は、スペーサーDNA断片を含み得る。

【0048】

ブロック306は、方法300が、第1のコンカテマーの第1の配列を得るために、第1のコンカテマーの単分子配列決定を行うことも含むことができることを示す。単分子配列決定は、ナノポア配列決定を含み得る。第1のコンカテマーは、単分子配列決定の一部として配列決定装置に提供されてもよい。配列決定装置は、ナノポアデバイス、光導波路、コンカテマーがフローセル上にハイブリダイズするように構成されたフローセル、または単一のDNA分子の配列検出が起こる任意の反応細胞または位置であってもよい。フローセルの例は、フローセルの表面に付着したオリゴヌクレオチドを有するフローセルを含むことができ、オリゴヌクレオチドは、第1のコンカテマーとハイブリダイズすることができる。方法300は、配列決定装置を使用して、第1のコンカテマーに対応する複数の信号を検出するステップをさらに含むことができる。複数の信号は、第1のコンカテマーの第1の配列に対応することができる。信号は、蛍光標識または第1のコンカテマーに結合した光学的に検出可能な信号を有する他の標識からの信号を含み得る。蛍光標識からの信号は、光検出器、レーザー、電荷結合素子、ゼロモード導波路、光学的事象の有無を決定することができる他の光学的に敏感な素子、またはそれらの組み合わせによって検出することができる。

20

30

【0049】

2. ナノポアにおける電気信号の検出

図3Bは、本発明の実施形態によるナノポア配列決定を使用してDNA断片を配列決定する方法350の簡略ブロックフロー図を示す。方法350において、配列決定装置はナノポアであり、これは基体上のナノポアのアレイに存在し得る。方法350の態様は、方法300を実施するときに実行することができる。

40

【0050】

ブロック352において、方法350は、複数のDNA断片を受け取ることを含むことができる。複数のDNA断片は、生体試料由来の細胞を含まないDNA断片であり得る。生体試料は、血漿または血清であり得る。DNA断片は、例えば、血漿が血液の他の成分から分離されるとき、生体試料の他の成分から分離されてもよい。

【0051】

DNA断片は、コンカテマー化に使用される容器で受け取ることができる。容器は、バイアルまたはエッペンドルフ(Eppendorf)チューブなどのチューブであってもよい。DNA断片は、容器内でリガーゼ酵素および緩衝液と混合することができる。

50

【 0 0 5 2 】

ブロック 3 5 4 は、方法 3 5 0 が第 1 のコンカテマーを得るために DNA 断片の第 1 のセットをコンカテマー化することも含むことを示す。コンカテマー化は、本明細書中に記載される任意の方法によるものであり得る。第 1 のコンカテマーは、第 1 のセット以外の DNA 断片を含むことができる。したがって、第 1 の組の DNA 断片は、第 1 のコンカテマーを構成する DNA 断片の全てではない可能性がある。

【 0 0 5 3 】

複数のコンカテマー化プロセスを並行して実行することができ、各プロセスは別個のコンカテマーを形成する。様々なコンカテマーは、例えば異なる DNA 断片が異なる長さであり得るため、様々な数の DNA 断片がコンカテマーに組み込まれ得るため、異なる長さであり得る。例えば、1 つのコンカテマーは 3 つの DNA 断片で構成することができ、もう 1 つのコンカテマーは 1 0 0 の DNA 断片で構成することができる。

10

【 0 0 5 4 】

ブロック 3 5 6 において、方法 3 5 0 は、第 1 のナノポアを介して第 1 のコンカテマーを通過させることをさらに含むことができる。第 1 のナノポアは、基体上の複数のナノポアのうちの 1 つであってもよい。第 1 のコンカテマーが第 1 のナノポアを通る工程は、第 1 のコンカテマーの第 1 の鎖がナノポアに通る工程を含み得る。第 1 鎖がナノポアを通過した後、第 1 コンカテマーの第 2 鎖がナノポアを通過することができる。第 1 の鎖および第 2 の鎖は、ヘアピンアダプターによって連結されてもよい。このようにして、両方の鎖を配列決定することができ、これにより、ベースコールの精度を高めることができる。両方のストランドの電気信号は、ベースコールの一部として比較することができる。

20

【 0 0 5 5 】

ブロック 3 5 8 において、第 1 のコンカテマーが第 1 のナノポアを通過する際に、第 1 の電気信号を検出することができる。第 1 の電気信号は、第 1 のコンカテマーの第 1 の配列に対応することができる。第 1 の電気信号は、電流または電圧を含むことができる。第 1 のナノポアを通過する第 1 のコンカテマーが存在しない場合、イオン電流が電極間を通過することができる。コンカテマーなどの生体分子が電極間を通過するとき、生体分子は電極間のイオンまたは電子の通過に影響を及ぼす可能性がある。その結果、電流または電圧が低下する可能性があります。変化の大きさは、電極間の生体分子の部分に関連し得る。例えば、特定のヌクレオチドまたは官能基が電極間を通過するとき、電流または電圧は特定の電気信号シグニチャを有することができる。

30

【 0 0 5 6 】

いくつかの実施形態において、ナノポアは、2 つの電極を含む電気回路の一部であり得る。2 つの電極間の電流は、どのヌクレオチド（塩基）または対応するタグがナノポア内にあるかに基づいて変化し得る。第 1 の電気信号は、回路内の電圧または電流を測定するための任意の適切な技術を用いて検出することができる。

【 0 0 5 7 】

ブロック 3 6 0 において、方法 3 5 0 は、第 1 の配列を決定するために第 1 の電気信号を分析するステップを含むことができる。分析は、電気信号のパターンを、特定の塩基に対応する既知のパターンと比較することを含むことができる。ナノポアのベースコールには、全ての目的のために参照により本明細書に組み込まれる、Schreiber J. and Karplus K., "Analysis of nanopore data using hidden Markov models," Bioinformatics 2015 31: 1897-1903 に記載されているように、隠れマルコフモデルを含む異なるモデルを使用することを含むことができる。分析するステップは、コンピュータシステムによって、第 1 の鎖および第 2 の鎖の第 1 の電気信号を分析して第 1 の配列を決定するステップを含むことができる。例えば、第 1 鎖および第 2 鎖について配列を決定することができ、2 つの配列を互いに比較することができる。配列は相補的でなければならない。相補的でない位置は、例えば、無視または再分析され得る。

40

【 0 0 5 8 】

いくつかの実施形態では、第 1 の電気信号の分析を用いて、第 1 のコンカテマーの様々

50

な部位、例えばCpG部位についてのメチル化分類を決定することができる。メチル化分類は、塩基がメチル化されているかどうか、異常メチル化（過剰メチル化または低メチル化）が存在するかどうか（例えば、CpGアイランドのような領域が異常なメチル化を有するかどうか）、およびコンカテマーがヒドロキシルメチル化されているかどうかを含むことができる。

【0059】

ブロック362において、方法350は、第1の配列のサブシーケンスを整列させて、DNA断片の第1のセットのそれぞれに対応する断片配列を同定することを含み得る。サブシーケンスは、例えばスライディングウィンドウ（sliding window）によって特定されるように、第1の配列の任意の1セットの連続する塩基であり得る。参照ゲノムへのアライメントのミスマッチを可能にする参照ゲノムに対してアライメントを行うことができる。サブシーケンスの整列については、以下でより詳細に説明する。

10

【0060】

コンカテマー化は、第2のコンカテマーを得るためのDNA断片の第2のセットについて、および他のコンカテマーを得るためのDNA断片の他のセットについても行うことができる。コンカテマーの各々は、他のコンカテマーとは異なるDNA断片のコンビネーションまたはパーミュテーションを有することができる。各コンカテマーは、配列決定装置を含むことができる複数のナノポアの第1のナノポアまたは他のナノポアを通過することができる。他のコンカテマーがナノポアを通過する際に、他の電気信号を得ることができる。他の電気信号は、他のコンカテマーのそれぞれの配列に対応することができる。他のコンカテマーを含む詳細は、第1のコンカテマーを含む方法に類似し得る。

20

【0061】

方法350はまた、第1のコンカテマーのDNA断片の第1のセットの各々のサイズを決定することを含むことができる。他のコンカテマーのDNA断片の他のセットのDNA断片のサイズも決定することができる。一例として、DNA断片のサイズは、サブシーケンスを参照ゲノムまたは既知のスペーサー配列と整列させることによって決定することができる。例えば、スペーサー配列が同定され得る場合、DNA断片の長さは、2つのスペーサー配列間の塩基数として同定され得る。スペーサー配列を使用するこのような実施形態では、スペーサー配列がそのような情報を提供することができるので、DNA断片の同定された配列は、それらを同定するために参照ゲノムに整列される必要はない。さらに、DNA断片の配列は、参照ゲノムへのアライメントの後またはその代わりに組み立てることができる。参照ゲノムに整列させる場合、DNA断片のサイズを決定することは、参照ゲノムの1つの領域に整列させる最長サブシーケンスの長さを決定することを含むことができる。

30

【0062】

3. サブシーケンスの整列

図4に示されるように、実施形態は、コンピュータシステムによって実行される方法400を含むことができる。図4は、本発明の実施形態によるコンカテマーの配列を分析する方法の簡略ブロックフロー図を示す。

【0063】

ブロック402において、方法400は、DNA断片の第1のセットをコンカテマー化することによって生成された第1のコンカテマーの第1の配列を受け取ることを含むことができる。いくつかの実施形態において、第1のコンカテマーは、DNA断片の第1のセットおよびDNA断片の第2のセットをコンカテマー化することによって生成され得る。第1の配列は、コンピュータシステムが常駐し得る配列決定装置に常駐することができるベースコールルーチンから受信することができる。別の例として、コンピュータシステムは配列決定装置から分離することができ、第1の配列は、ネットワーク接続を介して、またはリムーバブルメモリ装置を介して受信することができる。コンカテマーは、例えば、本明細書に記載されるような任意のコンカテマーであり得る。

40

【0064】

50

ブロック 404 において、方法 400 はまた、第 1 の配列のサブシーケンスを整列させて、DNA 断片の第 1 のセットの各 DNA 断片に対応する断片配列を同定することを含み得る。いくつかの実施形態では、サブシーケンスを整列させることは、サブシーケンスを DNA 断片の第 2 のセットに整列させることを含み得る。DNA 断片の第 2 セットは、本明細書に記載のスペーサー DNA であってもよい。スペーサー DNA は、15 ~ 20、10 ~ 15 および 5 ~ 10 ヌクレオチドを含む 20 ヌクレオチド以下であってもよく、またはそれに等しい既知の配列であってもよい。

【0065】

いくつかの実施形態では、方法 400 は、第 1 の配列のサブシーケンスを参照ゲノムに整列させることを含み得る。参照ゲノムはヒトゲノムであってもよい。血漿 DNA 分子の元の単位を同定するために、実施形態は長い DNA 配列をウィンドウにわたってヒトゲノムに整列させることができる。例えば、コンカテマーの長い配列からスライディングウィンドウ（例えば、100 ~ 300 塩基）を選択することができ、ウィンドウ配列（サブシーケンス）を参照ゲノムに整列させることができる。参照ゲノムは、参照ヒトゲノムの誘導体であってもよく、例えば、ヒトゲノム配列のサブセット、反復マスクされたゲノム、エクソン、または中程度またはバランスのとれた GC 含量を有するゲノムの一部を含むが、これに限定されない。

【0066】

ウィンドウは、ウィンドウの長さよりも少ない量だけ（例えば、20 ~ 50 塩基分）前進または後退させることができる（スライドさせる）。新しい位置にあるウィンドウは第 2 のウィンドウと見なすことができる。この第 2 のウィンドウのサブシーケンスは、参照ゲノムに整列させることもできる。第 2 のウィンドウのサブシーケンスが、参照ゲノムの前のサブシーケンスと重複する参照ゲノムのサブシーケンスに整列している場合、2 つのサブシーケンスは同じ DNA 断片の一部であると考えられることができる。2 つのスライディングウィンドウを、参照ゲノムの異なる、非連続の、または重複しない領域に整列させる場合、2 つの DNA 断片の配列を区別することができる。

【0067】

サブシーケンスがゲノムに整列しないが、前後のサブシーケンスが整列しない場合、2 つの DNA 断片間の交叉（エッジ）を同定することができる。整列した 2 つの DNA 断片を分析して、交叉の特定の点（例えば、一方の DNA 断片の開始塩基および他方の DNA 断片の終了塩基）を決定することができる。この交叉は、DNA 断片のサブシーケンスの終結または開始であってもよい。このアプローチは、スペーサーがコンカテマーの構築に使用されない場合に特に有用であり得る。他の実施形態では、特定の配列（例えば、特定のバーコードまたはスペーサー）を元の分子の末端に付加することができ、これらの特定の配列は、1 分子の終わりおよび別の分子の開始を示すことができる。

【0068】

したがって、実施形態は、ヒトゲノム上の異なる領域に属する DNA 塩基のストレッチまたはセグメント（一般に長さが数百塩基まで）を見出すことができる。DNA 塩基の各隣接ストレッチまたはセグメントは、1 つの元の血漿 DNA 分子を表すことができる。ヒトゲノムの異なる遠隔部分に整列された隣接する並置された DNA のセグメントまたはストレッチは、コンカテマーに組み立てられた他の血漿 DNA 分子に属し得る。

【0069】

ウィンドウのサイズおよびウィンドウが移動またはスライドされるステップのサイズは、DNA 断片のサイズに対する所望の解像度に基づいて調整することができる。より小さいステップサイズは、計算強度を増加させながら、DNA 断片の決定されたサイズの分解能を増加させることができる。ウィンドウサイズが大きいくほど、ウィンドウサイズよりも小さい DNA 断片は認識されない可能性があるが、ウィンドウサイズが小さければ、ゲノムとの一意の整列が得られないことがある。

【0070】

ウィンドウサイズおよびステップサイズは、動的に調整されてもよい。例えば、大きな

10

20

30

40

50

ステップサイズを用いて潜在的なマッチの領域を絞り込むことができ、マッチをより正確に識別するためにステップサイズを小さくすることができる。サブシーケンスを整列させることは、参照ゲノムの染色体または染色体領域とすることができる。サブシーケンスを整列させることは、配列決定エラーを説明するための多数のまたは頻度のミスマッチを可能にすることを含むことができる。例えば、配列を整列させることにより、約 10 ~ 15 % 以下のミスマッチを可能にすることができる。

【0071】

以下に述べるように、このアプローチは、血漿 DNA 分子の配列決定、ヒト染色体の同定、各染色体のヌクレオチド含量の比例差の検出、およびコンカテマー化前の元の血漿 DNA 試料のサイズプロファイルの決定に有効であった。コンカテマーは、ゲノム全体からの DNA 断片を含むことができる。コンカテマーの DNA 断片は、染色体の全てまたはほとんどの部分からランダムに分布していてもよい。

【0072】

B. 増加した濃度

別の実施形態は、フローセルの試料チャンバにロードされた血漿 DNA ライブラリの濃度を増加させることを含むことができる。血漿 DNA の濃度を増加させることは、通常、ナノポア配列決定の効率を増加させるとは予想されない。ナノポアは、比較的高い配列決定誤差を有し、このような細胞を含まない試料中の低 DNA 濃度と同様に、血漿および他の細胞を含まない試料から分析するための小さい DNA 断片を考慮すると、ナノポアは断片を効果的に配列決定することは期待されない。DNA 断片の濃度の増加は、この問題に対処することは期待されない。しかしながら、抽出された DNA または入力配列決定ライブラリを濃縮することにより、DNA 分子がナノポアまたは他の単分子分析技術に到達する機会が増強される。抽出された DNA を濃縮することは、配列決定技術に適合する体積を得るために必要なレベルを超えて濃縮することを含む。場合によっては、抽出された DNA の濃度が 10 倍以上増加することがある。換言すれば、体積は元の 10 % 未満に低減されてもよい。

【0073】

図 5 は、本発明の実施形態による DNA 断片の濃度を複数倍に増加させることによって DNA 断片をより効率的に配列決定する方法 500 の簡略ブロックフロー図を示す。方法 500 は、様々な単分子配列決定プラットフォームに使用することができる。提供される実施例において、ナノポア配列決定プラットフォームが記載される。

【0074】

ブロック 502 において、方法 500 は、複数の DNA 断片を含む生体試料を受け取ることを含むことができる。生体試料は、出発体積中の DNA 断片の第 1 の濃度を有し得る。生体試料は、例えば、本明細書に記載のように、様々なタイプのものであってもよい。例えば、生体試料は、血漿または血清であり得る。

【0075】

ブロック 504 において、方法 500 はまた、生体試料を濃縮して第 2 の濃度の DNA 断片を有することを含むことができる。種々の例として、DNA 断片の第 2 の濃度は、DNA 断片の第 1 の濃度よりも、5 回以上、6 回以上、7 回以上、8 回以上、9 回以上、10 回以上、50 回以上、100 回以上、500 倍以上、または 1000 倍以上増加させることができる。濃度は、容量あたりまたは質量あたりで測定することができる。

【0076】

濃縮は、当業者には理解されるように、様々な方法で達成され得る。例えば、生体試料の濃縮は、真空乾燥、浸透またはろ過による流体の除去、または当業者に公知の他の濃縮技術によることができる。ろ過または浸出は、遠心分離と組み合わせて、流体をサイズフィルターまたはモレキュラーシーブに通すことができる。流体の一方向の流れを可能にすることができる半透膜も、濃縮に使用することができる。濃縮後の生体試料の容量は、濃度の増加に反比例して減少してもよい。例えば、濃度を 5 倍に増加させると、容量は 5 倍に減少する可能性がある。

【0077】

濃度は、従来のプロセスよりも大幅に増加する可能性がある。いくつかの従来のプロセスでは、少量の血漿DNAが血漿の体積から抽出される。分析用具における反応容積または他の要件を満たすように体積を減少させるために、体積をさらに濃縮してもよい。例えば、従来のプロセスでは、210 μ Lの血漿DNAを4 mLの血漿から抽出することができる。100 μ Lの総反応容量を提供するために、210 μ Lの血漿DNAを85 μ Lに濃縮することができる。従来のプロセスでは、濃度の増加は3倍未満であり、血漿DNAは配列決定精度または正確さを改善するために濃縮されていない。本方法において、濃度の増加は、ナノポアまたは他の配列決定装置を通るDNA断片のより頻繁な通過をもたらす、したがって、検出および分析の改善をもたらす。

10

【0078】

ブロック506において、方法500は、基体上のナノポアを介して複数のDNA断片を通過させるステップをさらに含むことができる。いくつかの実施形態では、方法は、ナノポア以外の単分子配列決定技術を含み得る。例えば、配列決定技術は、Pacific BiosciencesによるSMRT技術またはSeqLLによるHelicos配列決定を含み得る。単分子配列決定技術は、全ての目的のために参照により本明細書に組み込まれる、Eid J. et al, "Real-time DNA sequencing from single polymerase molecules," Science 2009 323: 133-138に記載された技術を含み得る。

【0079】

ブロック508において、複数のDNA断片のそれぞれについて、DNA断片がナノポアを通過するときに電気信号を検出することができる。電気信号は、DNA断片の配列またはサブシーケンスに対応し得る。電気信号は、本明細書に記載された電流または電圧または任意の電気信号を含むことができる。他の配列決定技術が使用される場合、電気信号の代わりに蛍光シグナルが使用され得る。

20

【0080】

ブロック510において、方法500は、電気信号を分析してDNA断片の配列またはサブシーケンスを決定することを含むことができる。方法500は、例えば、整列情報を用いることによって、DNA断片のサイズおよび複数のDNA断片のサイズを決定することを含むことができる。結果として、DNA断片のサイズ分布もまた決定され得る。DNA断片のサイズ分布に基づいて、染色体の相違を決定することもでき、例えば、2010年11月5日出願された米国特許出願第12/940,992号のタイトル "Size-based genomics," ; 2011年11月30日出願された米国特許出願第13/308,473号のタイトル "Detection of genetic or molecular aberrations associated with cancer," ; および2013年3月7日出願された米国特許出願第13/789,553号のタイトル "Size-based analysis of fetal DNA fraction in maternal plasma," に記載される。

30

【0081】

いくつかの実施形態では、電気信号は、第1のコンカテマーのメチル化分類に対応し得る。メチル化の分類には、塩基がメチル化されているか否か、異常メチル化（過剰メチル化または低メチル化）が存在するかどうか、およびコンカテマーがヒドロキシルメチル化されているかが含まれ得る。

40

【0082】

方法は、DNA断片の配列またはサブシーケンスを参照ゲノムと整列させることを含み得る。特に、この整列は、参照ゲノムの特定の染色体または染色体領域とすることができる。

【0083】

同じDNA断片は、同じナノポアを複数回通過し得る。各パスで、電気信号を検出することができる。配列を同定するのを助けるために、異なるパスからの電気信号を比較することができる。DNA断片の濃度を増加させることは、ナノポア以外の単分子配列決定技術と共に使用することができる。

【0084】

50

III. 高めた濃度を用いる実施例

実施例は、血漿DNAを濃縮してナノポア配列決定の効率を高めながら、正確な結果を提供できることを示している。高めた濃度を用いた配列決定について、Cheng S. H. et al., "Noninvasive prenatal testing by nanopore sequencing of maternal plasma DNA: feasibility assessment," Clin. Chem. 61: 10 (2015)にさらに記載されている。

【0085】

A. 材料と方法

血漿試料は、インフォームドコンセントおよび施設承認を得て募集された4グループの集団、すなわち、男性胎児をみごもる3回目の妊娠の妊婦、女性胎児をみごもる3回目の妊娠の妊婦、成人男性および非妊娠女性から得た。EDTA血漿試料を各グループ内にプールして、群当たり少なくとも20mLの血漿を提供した。プールした血漿試料を、QIAamp DSP DNA血液ミニキット (Qiagen, Germany) を用いて抽出した(2)。プールあたり1,050μLの溶出血漿DNAをSpeedvac濃縮装置 (Thermo Fisher Scientific, Waltham, MA) により85μLまで濃縮した。各濃縮血漿DNAプールは、エンドリペアおよびAテーリングモジュール (New England Biolabs, Ipswich, MA) およびゲノムDNA配列決定キット (SQK-MAP-005, Oxford Nanopore Technologies, UK) を用いるDNAライブラリの調製のために完全に消費された。各ライブラリ(150μL)をMinION Flow Cell (v7.3) (Nanopore) に完全にロードし、配列決定した。出力データファイルは、METRICHOPTM (商標) ソフトウェア (Nanopore) を用いてベースコールされた。LAST Genome-Scale Sequence Comparisonソフトウェア (Computational Biology Research Consortium, Japan) を用いて、2D読み取りを抽出し、参照ゲノムhg19に整列させた。

【0086】

各ライブラリが消費されるまで配列決定を行い、6~24時間かかった。読み取りの26.9%~32.5%がベースキャストを通過した。男性胎児をみごもる妊婦、女性胎児をみごもる妊婦、成人男性および非妊娠女性からの血漿プールについては、2D読み取り数はそれぞれ56,844、50,268、35,878および36,167であった。観察された同一性の平均は、参照配列(3)における一致する塩基に整列させた読み取りにおける塩基の割合であり、82.7%(81.4~84.5%)であった。2D読み取りの中で、ユニークなゲノム位置に16.9%(15.6~23.9%)を整列させ、さらに分析した。

【0087】

ヒトゲノムに整列した配列決定された血漿DNA断片は、長さ76~5,776bpの範囲にあり、162bp(155~168bp)でピークに達した(図6)。図6の各グラフは、x軸上の塩基対における配列決定された血漿DNA断片のサイズ、および、配列決定された全血漿DNA断片の百分率としての血漿DNA断片サイズの頻度を示す。グラフは、男性胎児をみごもる母体血漿、女性胎児をみごもる母体血漿、成人男性血漿および非妊娠女性血漿から配列決定されたDNAの結果である。4つの血漿由来のピーク血漿DNAサイズは、イルミナの配列決定プラットフォーム(4)に基づく以前の知見と一致する。長い血漿DNA断片(>1,000bp)の微量(0.06~0.3%)がナノポア配列決定データから観察されたが、他の配列決定プラットフォーム(4)の以前のデータ解析からは観察されなかった。

【0088】

図7は、ナノポア配列決定によって得られ、イルミナ配列決定プラットフォームによって得られたデータからの女性の胎児をみごもる母体血漿由来の血漿DNAのサイズプロファイルを示す。ナノポア配列決定データは、図6の女性胎児をみごもる母体血漿と同じデータである。図7のサイズプロファイルは、ほぼ同じサイズのピークを有する同様の形状を有する。例えば、150bp以下の断片と161bp~170bpの断片のナノポア配列決定のサイズ比は、イルミナ配列決定の1.10と比較して1.21である。これらの結果は、血漿DNAのサイズプロファイルが、ナノポアおよび濃縮血漿DNAを用いて正確に決定され得ることを示した。250~400bpの範囲のピークは、イルミナ配列決

定よりもナノポア配列決定によって得られたデータにおいてより顕著である。このピークは、ジヌクレオソーム由来の細胞を含まないDNAに対応し、ピークの存在は個体間で変化する。さらに、イルミナ配列決定は、このサイズ範囲の配列断片ではあまり効率的ではない。

【0089】

B. 読み取りの分析

図8は、hg19で標識された、マッピング可能なヒトゲノムから予想される分布と比較した、染色体の読み取り分布を示す。染色体はx軸上に列挙されている。y軸上で、各染色体への読み取りの比例分布(ゲノム表現)は、その試料から配列決定され、頻度としておよびパーセンテージとして表された一意的に整列された読み取りの総数に対する各染色体に整列した読み取りの数をカウントすることにより、各試料について計算した。hg19でプロットされているのは、男性胎児をみごもる母体血漿、女性胎児をみごもる母体血漿、男性血漿および非妊娠女性血漿の結果である。全ての4つの血漿DNAプールについての常染色体への読み取り分布は、マッピング可能なヒトゲノムについて予測されたものと同等であった。

【0090】

染色体Xおよび染色体Yの差異が読み取り分布で観察される。X染色体にマッピングされた読み取りの割合は、女性血漿(5.22%)と比較して男性血漿(2.70%)の方が低かった。成人男性(0.30%)血漿DNAプールでは染色体Y配列が検出されたが、妊娠していない女性血漿DNAプールでは検出されなかった。

【0091】

男性胎児をみごもる女性の血漿DNAプールは、染色体Yに整列された読み取りが0.11%であった。以前のデータ(2)と一致して、女性胎児をみごもる女性由来の血漿DNAプール中の染色体Y配列に合わせて0.018%の読み取りが整列された。女性胎児をみごもる女性において、染色体Yに整列した読み取りの存在は、男性のゲノムに整列させることによる既知の誤りの結果であり得る。男性胎児をみごもる母体血漿DNAプールは、女性胎児をみごもる女性よりも約1%少ない染色体X配列を有していた。

【0092】

染色体Xおよび染色体Yに対する読み取りの相対的分布は、予想される結果と同様であることが観察される。男性血漿は、妊娠した女性と妊娠していない女性の両方からの血漿よりも染色体Yが多く見られた。男性血漿は、妊娠した女性および妊娠していない女性の血漿より染色体Xが少なかった。男性血漿には、妊娠していない女性のX染色体の量の約半分が含まれており、このことは、男性は染色体Xを1つとし、女性は2つの染色体Xを有することが予測される。男性胎児をみごもる母体血漿は、女性胎児をみごもる母体血漿および女性血漿よりも染色体Yの読み取り分布が高かった。

【0093】

したがって、胎児DNA配列決定および男性と女性の胎児間の染色体Xの量の差異は、ナノポア配列決定および血漿DNAの濃縮によって検出可能である。男性の胎児の染色体Xの量は、モノソミーXまたはターナー症候群の女性の胎児の染色体Xの量と同等である。従って、この観察は、モノソミーXなどの胎児染色体異数性の非侵襲的検出、またはコピー数異常のためのナノポア配列決定の潜在的実現可能性を示唆する。モノソミーXはゲノム中の1つの染色体コピーの減少を表すので、コピー数の変化の程度は、ゲノム中に1つの染色体コピーが付加されているトリソミーと同等である。したがって、これらのデータはまた、本発明者らのプロトコルが、非侵襲的に、胎児トリソミー21、トリソミー18、トリソミー13および他の胎児染色体異数性の検出に適用され得ることを反映する。これらのデータは、ナノポア配列決定に基づくNIPTおよびポイントオブケアNIPTの実現可能性を示唆している。

【0094】

IV. 高めた濃度を用いる癌配列

循環する細胞を含まないDNAは、癌のリアルタイムモニタリングのための「液体生検

10

20

30

40

50

」として使用することができる。細胞を含まないDNAは、基礎となる腫瘍において見出される遺伝的異常を示し、これは超並列配列決定技術によって検出することができる。癌患者の血漿DNAにおけるこれらの染色体異常はまた、ナノポア配列決定を用いて検出することができる。肝細胞癌（HCC）を有する2人の患者由来の血漿DNA試料を、ナノポア配列決定およびイルミナのプラットフォーム上の超並列配列決定の両方によって分析した。

【0095】

A．材料と方法

手術前にHCCと診断された2人の患者から末梢血をそれぞれ20ミリリットル採取した。血漿を、 $1600 \times g$ で10分間遠心分離し、次いで $16000 \times g$ で10分間遠心分離することにより単離した。QIAamp DSP DNA血液ミニキット（Qiagen）を用いて8mLの血漿からDNAを抽出した。血漿DNAの4分の3をナノポア配列決定に供し、残りのDNAはNextSeq 500（Illumina）によって配列決定した。

【0096】

ナノポア配列決定ライブラリは、エンドリペア/dAテーリングモジュール（NEB）およびナノポア配列決定キット（SQK-NSK007、Oxford Nanopore Technologies）によって調製した。ライブラリをMinion Flow Cell（R9バージョン）に完全にロードし、Minion Mk1B配列決定装置（Nanopore）で配列決定した。出力データファイルは、METRICHO RTMソフトウェア（Nanopore）を用いてベースコールされた。2D読み取りを抽出し、LASTソフトウェアを用いて参照ゲノムhg19に整列した。血漿DNAのイルミナ配列決定は、これまでに記載されているように行った（8）。

【0097】

B．結果

各染色体アーム（ゲノム表現、GR）に対する整列した読み取りの比例分布を各試料について計算した。換言すれば、染色体アーム、pまたはqアームに整列した高品質通過フィルターの読み取りの数は、試料から配列決定された高品質通過フィルター読み取りの割合として表された。次いで、正常個体の血漿DNA試料に対するGRの差を計算した。染色体アームのGRが対照群の平均より3標準偏差上である場合、その領域はコピー数増加を示すと考えられる。染色体アームのGRが対照群の平均よりも3標準偏差低い場合、その領域はコピー数の減少を示すと考えられる。

【0098】

図9は、ナノポア配列決定（外リング）およびイルミナのプラットフォーム（内リング）によるHCCの2つのケース（HOT530およびHOT536）におけるGRの差異を示す。相違する染色体は外部に表示される。分析された領域は染色体のアームであった。染色体ゲインは緑色の棒グラフで表され、それぞれのリングの中心から外向きに伸びている。染色体の損失は、赤色のバーとして表され、それぞれのリングの中心から内側に向かって伸びる。図9に示されるように、ナノポア配列決定の結果は、Illuminaのプラットフォームによって生成された結果とほぼ一致した。配列決定された試料はまた、非癌対象由来の血漿DNAと比較してより長いDNAを有する傾向を示した。この実施例は、増加した濃度法によるナノポア配列決定が、癌患者からの血漿DNAの分析に使用され得ることを示す。

【0099】

V．コンカテマーを用いる実施例

血漿DNA分子は典型的には短く（ $< 200 \text{ bp}$ ）、典型的には、長いDNA分子を配列決定するためにナノポア配列決定が使用され得る。短い血漿DNA分子の配列決定の効率は、コンカテマーと呼ばれる長い分子を構築するために個々の分子を連結または結合することによって改善することができる。

【0100】

A．材料と方法

1．血漿DNAコンカテマーの生成

2人の異なる対象からの試料を試験した。妊娠していない女性被験者および男性胎児をみごもる女性被験者の両方から、末梢静脈血20ミリリットルを採取した。1600×gで10分間の遠心分離後に血漿を回収し、さらに16000×gで10分間遠心分離した。次いで、QIAamp DSP DNA血液ミニキット (Qiagen) を用いて8 mLの血漿からDNAを抽出し、420 μLの血漿DNA量を得た。抽出したDNAをSpeedVac濃縮装置 (Thermo Scientific) で85 μLに濃縮し、NEB Next エンドリペアモジュール (New England Biolabs, NEB) で末端修復した。末端修復されたDNAを、MinElute Reaction Cleanup Kit (Qiagen) を用いて精製し、20 μLの緩衝液EBで溶出した。次に、20 μLの血漿DNAをBlunt/TA Ligase Master Mix (NEB) を加えて25℃で4時間インキュベートすることにより連結し、インキュベーション後にMinElute Reaction Cleanupキットで精製した。

10

【0101】

2. ナノポア配列決定

次いで濃縮DNAを、エンドリペアおよびAテーリングモジュール (NEB) およびゲノムDNA配列決定キット (SQK-MAP-005, Oxford Nanopore Technologies) によるナノポア配列決定ライブラリ調製に使用した。ライブラリをMinion Flow Cell (v7.3) (Nanopore) に完全にロードし、配列決定した。出力データファイルは、METRICHORTM (商標) ソフトウェア (Nanopore) を用いてベースコールされた。2D読み取りが抽出された。

【0102】

3. 整列

20

整列は、LASTソフトウェアを用いて行った。それぞれの可能な開始位置についての最初の一致をみつけた。マッチは、参照ゲノム中の最大回数で起こる最小長のマッチに限定されたか、マッチは特定の所定の長さに制限された。これらの最初のマッチから、これらの最初のマッチよりも長い配列の付加的な整列が実施され、あるギャップスコアを有するものが、配列決定エラーの所望の許容差に基づいて保持された。複数の整列が同じエンドポイントを共有していた場合、最も高いスコアの整列が保持された。このようにして、ヒトゲノムに整列した断片が同定された。LASTで使用するパラメータとルールは、精度と正確さの考慮事項によって変更される可能性がある。

【0103】

B. 結果

30

連結した血漿DNAを、ライブラリが消費されるまで6時間MinIONで配列決定した。

【0104】

1. 非妊娠女性

ベースコールは、2, 234の2D読み取りをもたらし、読み取りの長さは86から8,672bpの範囲であった。図10は、非妊娠女性から配列決定されたコンカテマーのサイズの頻度分布曲線を示す。塩基対におけるコンカテマーのサイズは、x軸上にある。指定されたコンカテマーサイズが配列決定された試料中に存在する頻度 (パーセンテージとして表される) は、y軸上にプロットされる。グラフの可読性を向上させるために、8,672bpの外れ値データポイントの外れ値データポイントは示されていない。配列決定された各試料について、約20~50個の長いDNA分子が検出された。

40

【0105】

配列決定された分子のサイズは、一般に200bp未満である典型的な血漿DNA断片の長さよりもはるかに長い。これらのデータは、血漿DNA断片がコンカテマーとしてうまく組み立てられたことを示している。次に、その読み取りをヒトゲノム (hg19) に整列させた。ヒトゲノム上の分離した領域に属する塩基またはセグメントのストレッチを分離し、1つの血漿DNA断片とみなした。全体として、配列決定された血漿DNA断片である3,801個の一意的にマップされたセグメントが得られ、一意的にマップされたセグメントの80.6%が参照ゲノムと同一の配列を示した。

【0106】

図11は、非妊娠女性の整列セグメントのサイズ分布を示す。これらの整列したセグメ

50

ントのサイズは、78bpから560bpまで変動し、ほとんどが200bp未満であった。ピークサイズは162bpであった。平均サイズは173bpであった。メジアンサイズは162bpであり、イルミナのプラットフォーム上での超並列配列決定による、および高めた濃度のDNA断片を用いたナノポア配列決定による、血漿DNAのサイズに関する以前の観察と一致する。

【0107】

図12は、男性由来の参照ゲノムと比較した整列したセグメントの計算された分布を示す。整列セグメントの各染色体への比例分布(ゲノム表現)を計算した。計算された分布は、すべての常染色体の中でhg19の分布に類似していた。性染色体では、染色体Xのゲノム表現は5.79%であり、これは女性試料について予想される。染色体Yに対するミスアラインメントは見出されなかった。従って、ナノポアを有するコンカテマーを用いる方法は、男性血漿と女性血漿とを区別することができる。

10

【0108】

2. 男性胎児をみごもる女性妊婦

妊娠した女性の男性胎児は、血液試料を採取したときに、38週の妊娠期間と4日間の妊娠期間を有していた。妊娠は正常とみなされた。血漿DNA断片をナノポア配列決定のための処理の前にコンカテマー化した。コンカテマーのサイズは100~13,466bpであり、メジアンサイズは676bpであり、平均サイズは965bpであった。

【0109】

図13は、男性胎児を妊娠した女性のためのコンカテマーの整列したセグメントのサイズプロファイルを示す。サイズプロファイルは、典型的な血漿DNAである。断片のメジアンサイズは196bpであった。ピークサイズは174bpであり、サイズは92~2,934bpであった。断片の約0.4%のみが2000bpより大きいサイズを有していた。この分布は、コンカテマーを用いる方法、高めた濃度、または超並列配列決定によって得られる他のサイズ分布に類似している。例えば、分布は、図6に示すように、男性胎児を有する母体血漿の高めた濃度の方法を用いて見出されたサイズ分布と同様である。

20

【0110】

図14は、男性および妊娠していない女性の参照ゲノムと比較した、男性胎児をみごもる女性妊婦についての整列したセグメントの計算された染色体分布を示す。コンカテマーの分布は、男性および非妊娠女性レベルの間の染色体Xレベルを示す。これは、染色体Xデータが、正常な男性胎児の染色体Xのモノソミーを反映していることを示唆している。さらに、コンカテマーは、胎児由来の染色体Yの証拠を示す。参照ゲノムからの配列決定によって得られた読み取り分布のいくつかの偏差は、参照ゲノムの整列可能な部分のみに分析を限定した結果であり得る。それにもかかわらず、これらの結果によって示されるように、コンカテマーを用いる方法は、男性胎児をみごもる母親血漿と男性または非妊娠女性と区別することができる。これらの方法は、男性胎児をみごもる母体血漿を、女性胎児をみごもる母体血漿と区別するために使用される可能性が高い。

30

【0111】

VI. さらなる実施形態

実施形態1は、複数のDNA断片を受け取ることと、複数のDNA断片の第1のセットをコンカテマー化して第1のコンカテマーを得ることと、前記第1のコンカテマーが前記第1のナノポアを通ることと、前記第1のコンカテマーが前記第1のナノポアを通るときに第1の電気信号を検出することと、前記電気信号は前記第1のコンカテマーの前記第1の配列に対応することを含む方法を含む。

40

【0112】

実施形態2は、実施形態1の方法を含み、さらに、前記第1の配列を決定するために前記第1の電気信号を分析することと、参照ゲノムに対して前記第1の配列のサブシーケンスを整列させて、前記DNA断片の前記第1のセットそれぞれに対応する断片配列を同定することを、コンピュータシステムにより実行することを含む。

【0113】

50

実施形態 3 は、実施形態 1 の方法を含み、前記第 1 のコンカテマーが前記第 1 のナノポアを通ることは、前記第 1 のコンカテマーの第 1 の鎖が前記ナノポアを通り、その後、前記第 1 のコンカテマーの第 2 の鎖が前記ナノポアを通ることを含む。

【0114】

実施形態 4 は、実施形態 3 の方法を含み、さらに、前記第 1 の鎖および前記第 2 の鎖についての前記第 1 の電気信号をコンピュータシステムにより分析して、前記第 1 の配列を決定することを含む。

【0115】

実施形態 5 は、実施形態 1 の方法を含み、複数の DNA 断片は、生体試料由来の細胞を含まない DNA 断片である。

10

【0116】

実施形態 6 は、実施形態 2 の方法を含み、前記生体試料は血漿または血清である。

【0117】

実施形態 7 は、実施形態 1 の方法を含み、前記第 1 のナノポアは基体上の複数のナノポアの 1 つである。

【0118】

実施形態 8 は、実施形態 7 の方法を含み、前記 DNA 断片の第 2 のセットをコンカテマー化して第 2 のコンカテマーを得ることと、前記第 2 のコンカテマーが複数のナノポアの第 2 のナノポアを通り、前記第 2 のコンカテマーが前記第 2 のナノポアを通るときに電気信号を検出し、前記電気信号は前記第 2 のコンカテマーの第 2 の配列に対応することをさらに含む。

20

【0119】

実施形態 9 は、実施形態 7 の方法を含み、前記 DNA 断片の第 2 のセットをコンカテマー化して第 2 のコンカテマーを得ることと、前記第 2 のコンカテマーが前記第 1 のナノポアを通り、前記第 2 のコンカテマーが前記第 1 のナノポアを通るときに電気信号を検出し、前記電気信号は前記第 2 のコンカテマーの第 2 の配列に対応することをさらに含む。

【0120】

実施形態 10 は、複数の DNA 断片を含む生体試料を受け取ることと、前記生体試料を濃縮してより高い濃度の DNA 断片を得ることと、基体上で複数の DNA 断片がナノポアを通ることと、複数の DNA 断片のそれぞれについて、前記 DNA 断片がナノポアを通るときに電気信号を検出することと、前記電気信号が前記 DNA 断片の前記配列に対応することを含む方法を含む。

30

【0121】

実施形態 11 は、DNA 断片の第 1 のセットをコンカテマー化することにより生成される第 1 のコンカテマーの第 1 の配列を受け取ることと、参照ゲノムに前記第 1 の配列のサブシーケンスを整列させて、前記 DNA 断片の前記第 1 のセットそれぞれに対応する断片配列を同定することを、コンピュータシステムにより実行することを含む方法を含む。実施形態 12 は、実施形態 11 の方法を含み、前記サブシーケンスの前記整列は、前記参照ゲノムに前記第 1 の配列のウィンドウを整列させることと、前記参照ゲノムの異なる領域に 2 つのウィンドウを整列させるときに同定することを含む。実施形態 13 は、実施形態 12 の方法を含み、前記 2 つのウィンドウ間の 1 つ以上のウィンドウが前記参照ゲノムに整列されないときに同定される。

40

【0122】

実施形態 14 は、コンピュータシステムを制御するための複数のインストラクションを保存するコンピュータ読み取り可能媒体を含むコンピュータ製品を含み、実施形態 11 ~ 13 のいずれか 1 つの操作を実行することを含む。実施形態 15 は、実施形態 14 のコンピュータ製品と、前記コンピュータ読み取り可能媒体に保存されるインストラクションを実行するための 1 つ以上のプロセッサとを含むシステムを含む。実施形態 16 は、実施形態 11 ~ 13 のいずれか 1 つを実行するための手段を含むシステムを含む。実施形態 17 は、実施形態 11 ~ 13 のいずれか 1 つを実行するように構成されるシステムを含む。実

50

施形態 18 は、実施形態 11 ~ 13 のいずれか実施形態 11 ~ 13 のいずれか 1 つのステップをそれぞれ実行するためモジュールを含むシステムを含む。

【0123】

VII. 実施例の配列決定システム

図 15 は、本発明の実施形態による単分子配列決定を実施するためのシステム 1500 のブロック図を示す。生体試料は、抽出装置 1504 によって患者 1502 から得ることができる。生体試料は、本明細書に記載される任意の体液または任意の生体試料であり得る。抽出装置 1504 は、尿などの試料を採取するためのシリンジ、ランセット、スワブ、または容器またはバイアルを含み得る。抽出装置 1504 は、QIAamp DSP DNA 血液ミニキット (Qiagen, Germany) を含むことができる (2)。生体試料は、細胞を含まない D
NA 断片を含んでいてもよく、これは調製装置 1506 に送られる。調製装置 1506 は、単分子配列決定により効率的な細胞を含まない D NA 断片の形態を生成する装置を含むことができる。例えば、調製装置 1506 の出力は、コンカテマーまたは細胞を含まない D NA 断片の濃縮試料であってもよい。

【0124】

コンカテマーが生成される場合、調製装置 1506 は、浸透またはろ過によって流体を除去する SpeedVac コンセントレータ (Thermo Scientific) または超濃縮器のような濃度を高めるための真空乾燥機を含むことができる。コンカテマーの場合、濃度は 5 倍以上増加しないことがある。調製装置 1506 は、NEBNext エンドリペアモジュール (New England Biolabs, NEB) などの D NA 断片の末端を修復するためのモジュールを含むことができる。さらに、調製装置 1506 は、MinElute Reaction Cleanup Kit (Qiagen) のような末端修復 D NA を精製するためのキットを含むことができる。調製装置 1506 はまた、リガーゼ酵素 (例えば、Blunt/TA Ligase Master Mix (NEB)) を数時間、ある温度 (例えば、25 で 4 時間) でインキュベートし得るインキュベーターを含み得る。調製装置 1506 は、MinElute Reaction Cleanup Kit (Qiagen) を含み得る反応後のコンカテマーを精製するための追加のキットを含み得る。調製装置 1506 はまた、流体を混合して移送することができるロボット液体ハンドラを含むことができる。

【0125】

高めた濃度の細胞を含まない D NA 断片が生成される場合、調製装置 1506 は、真空乾燥機 (例えば Speedvac 濃縮機 (Thermo Fisher Scientific, Waltham, MA)) または浸透またはろ過によって流体を除去する超濃縮機を含むことができる。調製装置 1506 はまた、エンドリペアおよび A テーリングモジュール (New England Biolabs, Ipswich, MA) を含み得る。調製装置 1506 はまた、流体を混合して移送することができるロボット液体ハンドラを含むことができる。

【0126】

調製装置 1506 からのコンカテマーまたは濃縮 D NA 断片は、配列読み取りを得るために単分子配列決定装置 1508 に送られてもよい。単分子配列決定デバイス 1508 はまた、MinION Flow Cell (v7.3) (Nanopore)、SMRT テクノロジー (Pacific Biosciences)、または Helicos Single Molecule 配列決定装置 (SeqLL) のようなデバイスを含み得る。単分子配列決定デバイス 1508 はまた、ゲノム D NA 配列決定キット (SQK-MAP-005, Oxford Nanopore Technologies, UK) または Nanopore Sequencing Kit (SQK-NSK007, Oxford Nanopore Technologies) のようなナノポアに関連するキットを含み得る。

【0127】

単分子配列決定装置 1508 は、コンカテマーまたは濃縮 D NA 断片からの配列読み取りを出力することができる。配列読み取りは、データファイルとして出力されてもよく、コンピュータシステム 1510 によって解析されてもよい。コンピュータシステムは、配列読み取りを分析するソフトウェアを備えた特殊なコンピュータシステムであってもよい。出力データファイルは、METRICHTM (商標) ソフトウェア (Nanopore) を用いてベースコールすることができる。2D 読み取りは、LAST ソフトウェアを使用して抽出し、整列させることができる。コンピュータシステムは、以下に記載されるように、図 16 の

コンピュータシステム 10 であってもよい。

【0128】

V I I I . コンピュータシステム

本明細書で言及されるコンピュータシステムのいずれも、任意の適切な数のサブシステムを利用することができる。コンピュータシステム 10 における、そのようなサブシステムの例を図 16 に示す。いくつかの実施形態では、コンピュータシステムは、単一のコンピュータ装置を含み、サブシステムは、コンピュータ装置の構成要素とすることができる。他の実施形態では、コンピュータシステムは、それぞれが内部構成要素を有するサブシステムである複数のコンピュータ装置を含むことができる。コンピュータシステムは、デスクトップおよびラップトップコンピュータ、タブレット、携帯電話および他のモバイルデバイスを含むことができる。

10

【0129】

図 6 に示すサブシステムは、システムバス 75 を介して相互接続されている。ディスプレイアダプタ 82 に結合されたプリンタ 74、キーボード 78、ストレージデバイス 79、モニタ 76 などの追加のサブシステムが示されている。I/O コントローラ 71 に結合する周辺機器および入力/出力 (I/O) デバイスは、入力/出力 (I/O) ポート 77 のような (例えば、USB、Fire Wire (登録商標)) 当技術分野で公知の任意の数の手段によってコンピュータシステムに接続することができる。例えば、コンピュータシステム 10 を、インターネット、マウス入力装置、またはスキャナなどの広域ネットワークに接続するために、I/O ポート 77 または外部インターフェース 81 (例えば、イーサネット、Wi-Fi など) を使用することができる。システムバス 75 を介する相互接続は、中央プロセッサ 73 が各サブシステムと通信し、システムメモリ 72 または記憶装置 79 (例えば、ハードドライブ、または光ディスクのような固定ディスク) からの複数の命令の実行を制御すること、並びに、サブシステム間の情報の交換を可能にする。システムメモリ 72 および/または記憶装置 79 はコンピュータ読み取り可能媒体を具体化することができる。別のサブシステムは、カメラ、マイクロホン、加速度計などのデータ収集装置 85 である。ここに記載されたデータのいずれも、あるコンポーネントから別のコンポーネントに出力することができ、ユーザに出力することができる。

20

【0130】

コンピュータシステムは、例えば、外部インターフェース 81 によって、内部インターフェースによって、または 1 つの構成要素から別の構成要素に接続および取り外し可能なリムーバブルストレージデバイスを介して互いに接続された複数の同じ構成要素またはサブシステムを含むことができる。いくつかの実施形態では、コンピュータシステム、サブシステム、または装置は、ネットワークを介して通信することができる。そのような場合、1 つのコンピュータはクライアントとみなされ、別のコンピュータはサーバであり、それぞれが同じコンピュータシステムの一部となり得る。クライアントとサーバは、それぞれ複数のシステム、サブシステム、またはコンポーネントを含むことができる。

30

【0131】

実施形態の態様は、ハードウェア (例えば、特定用途向け集積回路またはフィールドプログラマブルゲートアレイ) を使用する制御ロジックの形態で、および/または一般的にプログラム可能なプロセッサを備えたコンピュータソフトウェアをモジュール方式または統合方式で使用して実施することができる。本明細書で 사용되는場合、プロセッサは、単一コアプロセッサ、同一集積チップ上のマルチコアプロセッサ、または単一回路基体上の複数の処理ユニット、またはネットワーク化されたものを含む。本明細書で提供される開示および教示に基づいて、当業者は、ハードウェアおよびハードウェアおよびソフトウェアの組み合わせを使用して本発明の実施形態を実施する他の手法および/または方法を理解し、認識するであろう。

40

【0132】

このアプリケーションに記載されているソフトウェアコンポーネントまたは機能のいずれかは、Java、C、C++、C#、Objective-C、Swift、または例

50

えば、PerlまたはPythonのような従来の技術やオブジェクト指向の技術を使用するスクリプト言語のような、任意の適切なコンピュータ言語を使用してプロセッサによって実行されるソフトウェアコードとして実装されてもよい。ソフトウェアコードは、記憶および/または送信のためにコンピュータ読み取り可能媒体上に一連のインストラクションまたはコマンドとして格納することができる。適切な非一時的なコンピュータ読み取り可能媒体は、ランダムアクセスメモリ(RAM)、読み出し専用メモリ(ROM)、ハードドライブまたはフロッピーディスクなどの磁気媒体、またはコンパクトディスク(CD)またはDVD(デジタル多用途ディスク)、フラッシュメモリなどのような光学媒体を含むことができる。コンピュータ読み取り可能媒体は、そのような記憶装置または伝送装置の任意の組み合わせであってもよい。

10

【0133】

このようなプログラムは、インターネットを含む様々なプロトコルに準拠した有線、光、および/または無線ネットワークを介した伝送に適した搬送波信号を用いて符号化され、伝送されてもよい。このように、コンピュータ読み取り可能媒体は、そのようなプログラムで符号化されたデータ信号を使用して作成されてもよい。プログラムコードでコード化されたコンピュータ読み取り可能媒体は、互換性のあるデバイスと共にパッケージされてもよく、または(例えば、インターネットダウンロードを介して)他のデバイスとは別個に提供されてもよい。このようなコンピュータ読み取り可能媒体は、単一のコンピュータ製品(例えば、ハードドライブ、CD、またはコンピュータシステム全体)上またはその中に存在してもよく、システムまたはネットワーク内の異なるコンピュータ製品上またはその中に存在してもよい。コンピュータシステムは、モニタ、プリンタ、または本明細書で言及された結果のいずれかをユーザに提供するための他の適切なディスプレイを含むことができる。

20

【0134】

本明細書に記載の方法のいずれも、ステップを実行するように構成することができる1つまたは複数のプロセッサを含むコンピュータシステムを用いて全体的にまたは部分的に実行することができる。したがって、実施形態は、可能性として、それぞれのステップまたはそれぞれのステップのグループを実行する異なるコンポーネントを用いて、本明細書に記載の方法のいずれかのステップを実行するように構成されたコンピュータシステムを対象とすることができる。番号付きのステップとして提示されているが、本明細書の方法のステップは、同時に、または異なる順序で実行することができる。さらに、これらのステップの一部は、他の方法の他のステップの一部と共に使用することができる。また、ステップの全部または一部は任意であってもよい。さらに、これらのステップを実行するためのモジュール、ユニット、回路、または他の手段を用いて、方法のいずれのステップも実行することができる。

30

【0135】

特定の実施形態の特定の詳細は、本発明の実施形態の精神および範囲から逸脱することなく、任意の適切な方法で組み合わせることができる。しかしながら、本発明の他の実施形態は、それぞれの個々の態様、またはこれらの個々の態様の特定の組み合わせに関する特定の実施形態に向けられ得る。

40

【0136】

本発明の例示的な実施形態の上記説明は、例示および説明のために提示されたものである。網羅的であることを意図するものでもなく、本発明を記載された厳密な形態に限定することも意図されておらず、上記教示に照らして多くの修正および変形が可能である。

【0137】

「1つ(a)」、「1つ(an)」または「その(the)」の列挙は、特に断らない限り、「1つまたは複数」を意味することを意図している。「または(or)」の使用は、特に反対の指示がない限り、「包含的」であり、「排他的」ではないことを意味することを意図する。「第1の」構成要素への言及は、第2の構成要素が提供されることを必ずしも必要としない。さらに、「第1」または「第2」構成要素への言及は、明示的に述べ

50

られていない限り、参照される構成要素を特定の場所に限定するものではない。

【 0 1 3 8 】

本明細書で言及したすべての特許、特許出願、刊行物および説明は、あらゆる目的のためにその全体が参考として援用される。いずれも先行技術であるとは認められない。

【 0 1 3 9 】

IX. 参考文献:

1. Dondorp W, de Wert G, Bombard Y, Bianchi DW, Bergmann C, Borry P, et al. Non-invasive prenatal testing for aneuploidy and beyond: Challenges of responsible innovation in prenatal screening. *Eur J Hum Genet* 2015.
2. Chiu RWK, Chan KCA, Gao Y, Lau VYM, Zheng W, Leung TY, et al. Noninvasive prenatal diagnosis of fetal chromosomal aneuploidy by massively parallel genomic sequencing of dna in maternal plasma. *Proc Natl Acad Sci USA* 2008 ; 105: 20458-63. 10
3. Jain M, Fiddes IT, Miga KH, Olsen HE, Paten B, Akeson M. Improved data analysis is for the MinION nanopore sequencer. *Nat Methods* 2015 ; 12: 351-6.
4. Lo YMD, Chan KCA, Sun H, Chen EZ, Jiang P, Lun FMF, et al. Maternal plasma DNA sequencing reveals the genome-wide genetic and mutational profile of the fetus. *Sci Transl Med* 2010 ; 2: 61ra91-61ra91.
5. Bayley H. Nanopore sequencing: From imagination to reality. *Clin Chem* 2015 ; 61: 25-31.
6. Zheng YW, Chan KCA, Sun H, Jiang P, Su X, Chen EZ, et al. Nonhematopoietically derived DNA is shorter than hematopoietically derived DNA in plasma: A transplantation model. *Clin Chem* 2012 ; 58: 549-58. 20
7. Lun FMF, Chiu RWK, Sun K, Leung TY, Jiang P, Chan KCA, et al. Noninvasive prenatal methylomic analysis by genomewide bisulfite sequencing of maternal plasma DNA. *Clin Chem* 2013 ; 59: 1583-94.
8. Chan KCA, Jiang P, Zheng YW, Liao GJW, Sun H, Wong J, et al. Cancer genome scanning in plasma: Detection of tumor-associated copy number aberrations, single-nucleotide variants, and tumoral heterogeneity by massively parallel sequencing. *Clin Chem* 2013 ; 59: 211-24.
9. Chan KCA, Jiang P, Chan CWM, Sun K, Wong J, Hui EP, et al. Noninvasive detection of cancer-associated genome-wide hypomethylation and copy number aberrations by plasma DNA bisulfite sequencing. *Proc Natl Acad Sci U S A* 2013 ; 110: 18761-8. 30
10. Chan RWY, Wong J, Chan HLY, Mok TSK, Lo WYW, Lee V, et al. Aberrant concentrations of liver-derived plasma albumin mRNA in liver pathologies. *Clin Chem* 2010 ; 56: 82-9.
11. Chan RWY, Jiang P, Peng X, Tam LS, Liao GJW, Li EK, et al. Plasma DNA aberrations in systemic lupus erythematosus revealed by genomic and methylomic sequencing. *Proc Natl Acad Sci U S A* 2014 ; 111: E5302-11.
12. Schreiber J, Wescoe ZL, Abu-Shumays R, Vivian JT, Baatar B, Karplus K, Akeson M. Error rates for nanopore discrimination among cytosine, methylcytosine, and hydroxymethylcytosine along individual DNA strands. *Proc Natl Acad Sci U S A* 2013 ; 110: 18910-5. 40

【図 1 A】

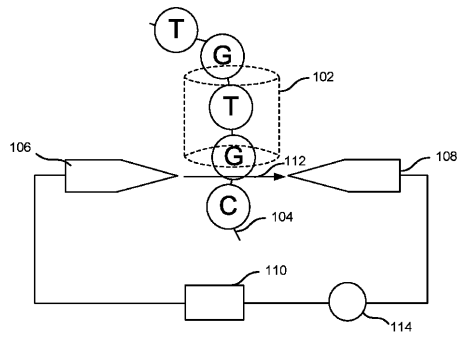


FIG. 1A

【図 1 B】

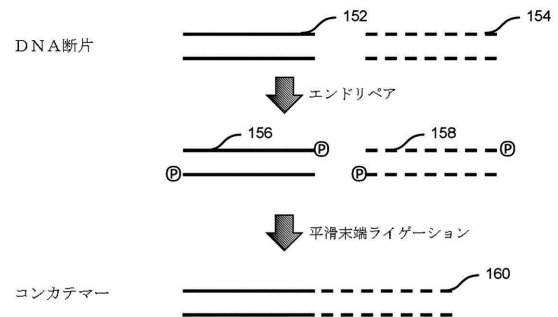


図 1 B

【図 2】

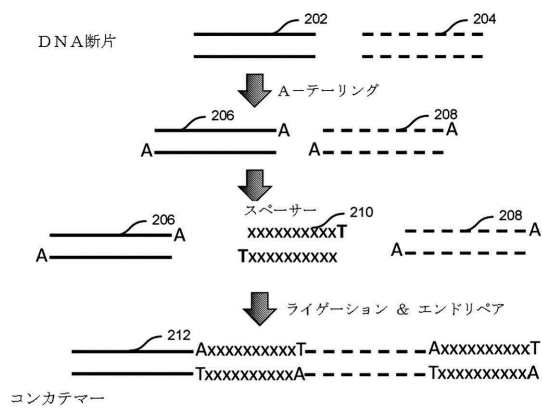


図 2

【図 3 A】

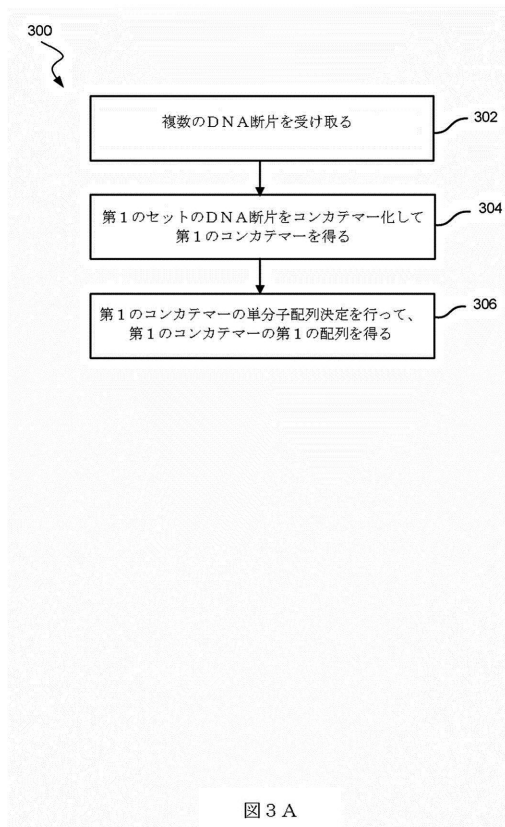


図 3 A

【図 3 B】

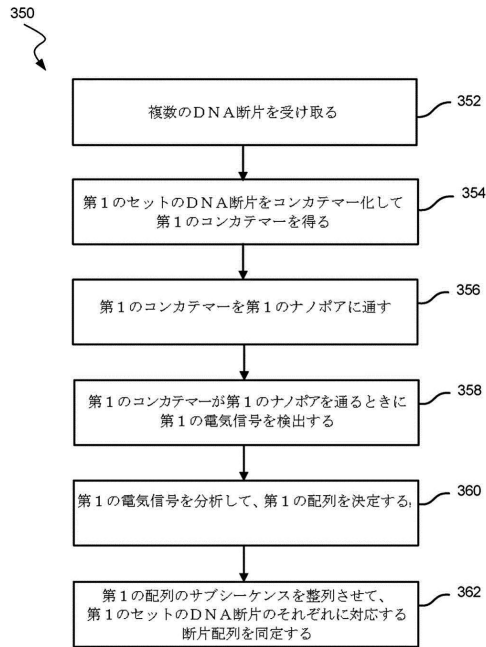


図 3 B

【図 4】

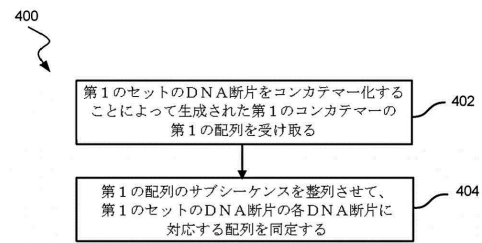


図 4

【図 5】

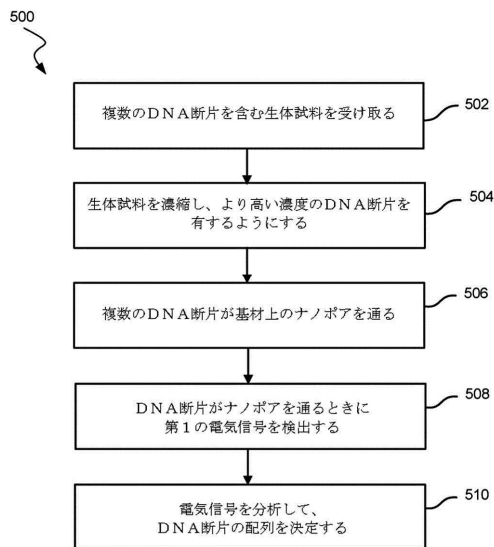


図 5

【図 6】

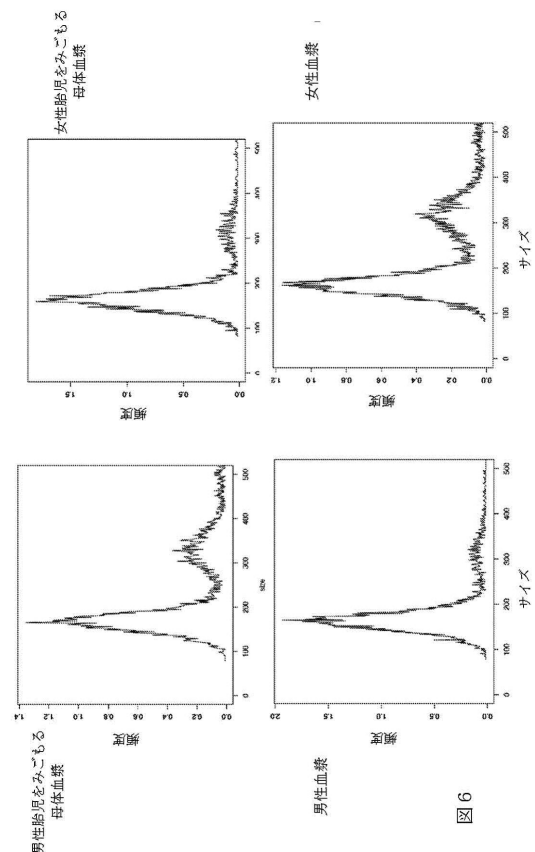


図 6

【図 7】

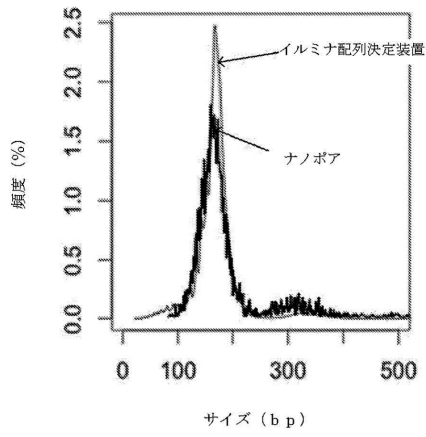


図 7

【図 8】

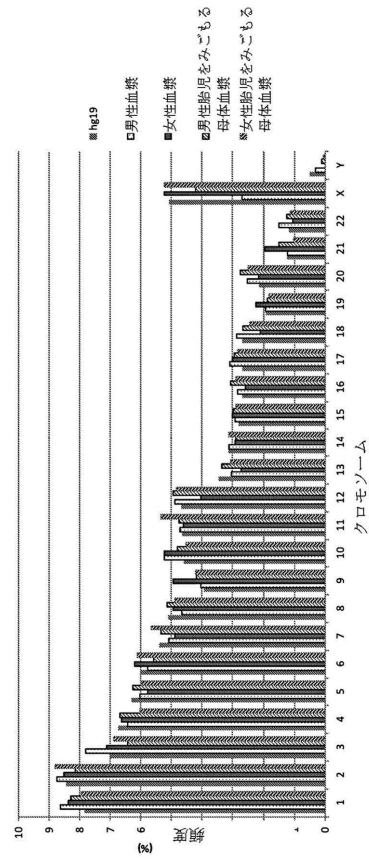


図 8

【図 9】

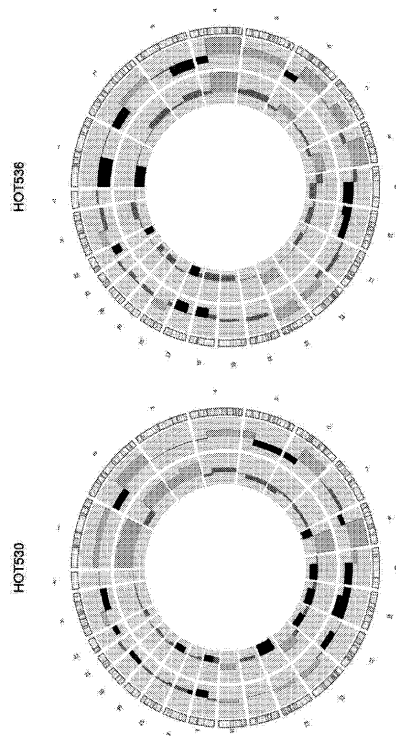


FIG. 9

【図 10】

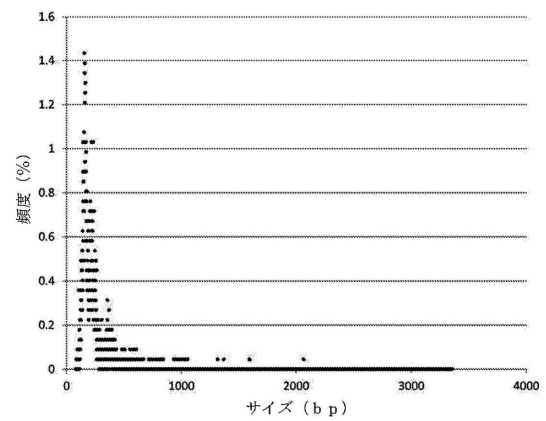


図 10

【図 1 1】

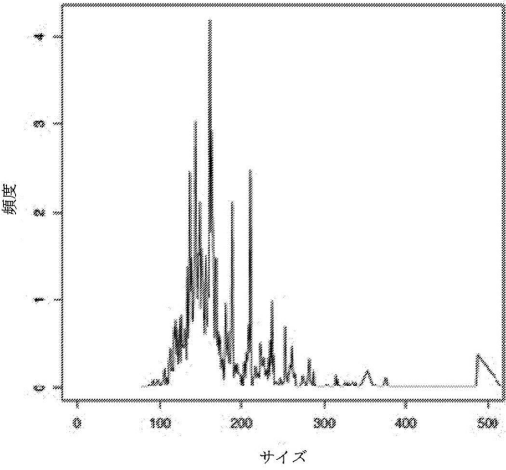


図 1 1

【図 1 2】

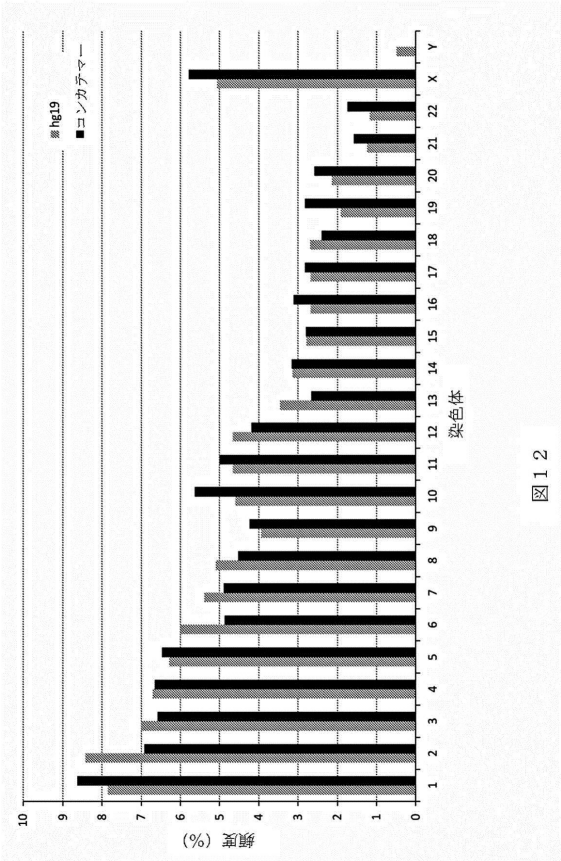


図 1 2

【図 1 3】

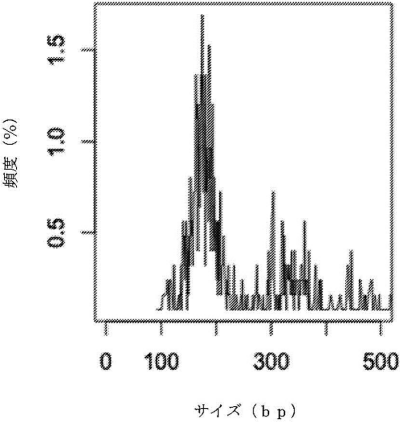


図 1 3

【図 1 4】

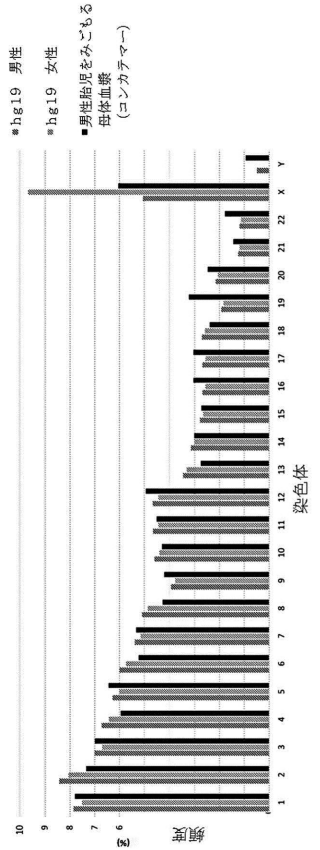
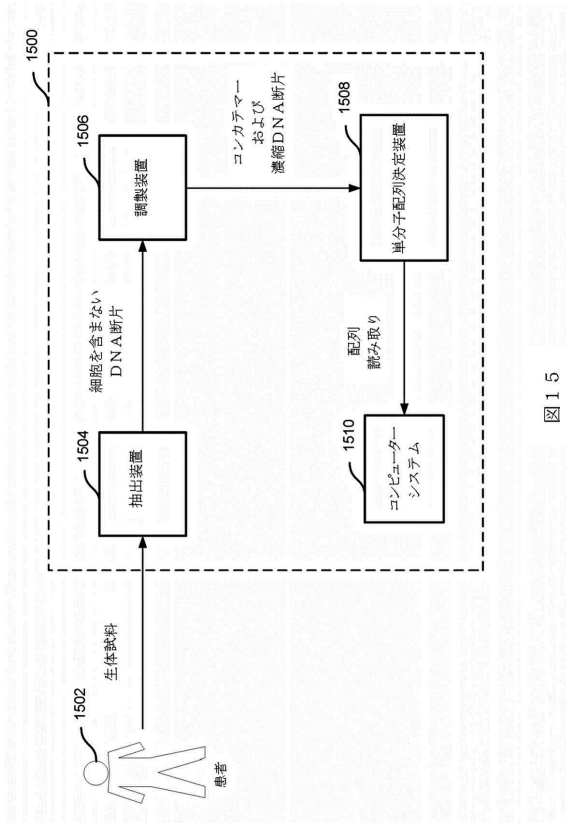
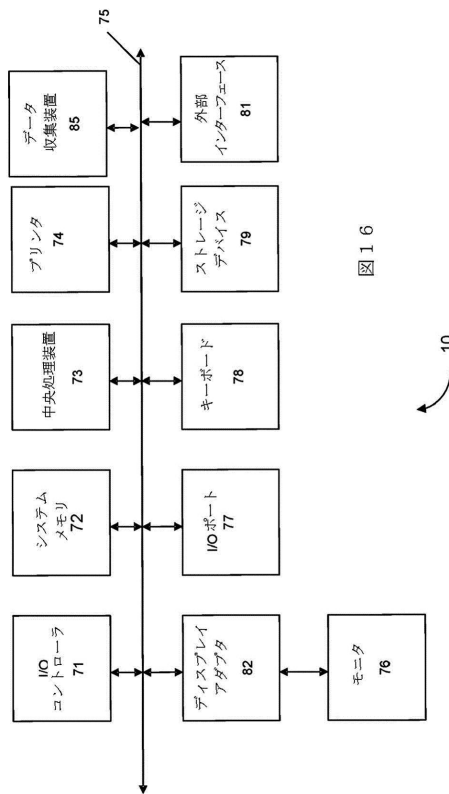


図 1 4

【図 15】



【図 16】



フロントページの続き

(74)代理人 100141977

弁理士 中島 勝

(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 ロ ユク - ミン デニス

中華人民共和国, 香港 999077, カウルーン, ホウマンティン, キング タク ストリート
7, フォース フロア

(72)発明者 チウ ロッサ ウェイ クン

中華人民共和国, 香港 999077, ニュー テリトリーズ, シャティン, マ ロク パス 5
2, ダブル ヘブン, ハウス 31

(72)発明者 チョン スク ハン

中華人民共和国, 香港, 999077, ニュー テリトリーズ, ファンリン, マ シク ロード
1, グリーン コード, タワー 1, 32 / エフ, フラット ビー

審査官 小倉 梢

(56)参考文献 特開2008 - 161056 (JP, A)

特表2015 - 505458 (JP, A)

国際公開第2014 / 033455 (WO, A1)

(58)調査した分野(Int.Cl., DB名)

C12Q 1 / 00 - 1 / 70

G16B 25 / 00 - 25 / 30

G16B 30 / 00 - 30 / 20

JSTPlus / JMEDPlus / JST7580 (JDreamIII)

CAplus / MEDLINE / EMBASE / BIOSIS (STN)

WPIDS / WPIX (STN)