

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7315326号
(P7315326)

(45)発行日 令和5年7月26日(2023.7.26)

(24)登録日 令和5年7月18日(2023.7.18)

(51)国際特許分類

C 1 2 M	1/00 (2006.01)	C 1 2 M	1/00	A Z N A
C 1 2 N	15/09 (2006.01)	C 1 2 N	15/09	2 0 0
C 1 2 Q	1/6837(2018.01)	C 1 2 Q	1/6837	Z
C 1 2 Q	1/6844(2018.01)	C 1 2 Q	1/6844	Z

請求項の数 9 (全41頁)

(21)出願番号 特願2018-526506(P2018-526506)
 (86)(22)出願日 平成28年11月15日(2016.11.15)
 (65)公表番号 特表2018-533965(P2018-533965)
 A)
 (43)公表日 平成30年11月22日(2018.11.22)
 (86)国際出願番号 PCT/US2016/062090
 (87)国際公開番号 WO2017/087416
 (87)国際公開日 平成29年5月26日(2017.5.26)
 審査請求日 令和1年11月14日(2019.11.14)
 審判番号 不服2021-13121(P2021-13121/J
 1)
 審判請求日 令和3年9月29日(2021.9.29)
 (31)優先権主張番号 62/306,597
 (32)優先日 平成28年3月10日(2016.3.10)
 (33)優先権主張国・地域又は機関 最終頁に続く

(73)特許権者 518173090
 リヴィア バイオセンサーズ, エルエル
 シー
 アメリカ合衆国 4 8 0 0 9 ミシガン州
 バーミンガム アスペン・ロード 5 6 7
 (74)代理人 100082072
 弁理士 清原 義博
 (74)復代理人 100208292
 弁理士 清原 直己
 (72)発明者 ケイン, ロバート, チャールズ
 アメリカ合衆国 9 2 1 3 0 カリフォル
 ニア州 サンディエゴ メサ・ノルテ・ド
 ライブ 6 5 2 0
 (72)発明者 シエン, リチャード
 アメリカ合衆国 9 2 0 6 7 カリフォル
 最終頁に続く

(54)【発明の名称】 遺伝子サンプルを識別且つ区別するためのシステムと方法

(57)【特許請求の範囲】

【請求項1】

バイオチップシステムであって、該バイオチップシステムは、
 光学的解像度を有する光検出器であって、該光検出器は前記光学的解像度以下のシグナル
 を検出できない、光検出器と、

前記光検出器に光学的に接続されたバイオチップを含み、

前記バイオチップは、

(i) 第1の被験体に特異的な特徴を標的とする第1の複数の同一のプローブと、
 (ii) 第2の被験体に特異的な特徴を標的とする第2の複数の同一のプローブと、を含
 み、

前記第1の複数の同一のプローブと前記第2の複数の同一のプローブは前記バイオチップ
 の表面に固定されてなり、

前記第1の複数の同一のプローブと前記第2の複数の同一のプローブは互いに異なり、
 前記第1の複数の同一のプローブ及び前記第2の複数の同一のプローブに含まれるプロー
 ブは、単一では前記光検出器の有する前記光学解像度未満またはほぼ等しい領域に包含さ
 れ、

ここで、前記光検出器が、

i) 前記第1の複数の同一プローブに関連するシグナルを統合して前記光学解像度により
 検出される要素とし、及び

ii) 前記第2の複数の同一プローブに関連するシグナルを統合して前記光学解像度によ

り検出される要素とし、
検出する、
バイオチップシステム。

【請求項 2】

前記第1の被験体に特異的な特徴、及び、前記第2の被験体に特異的な特徴とは、異なるものである、ことを特徴とする請求項1に記載のバイオチップシステム。

【請求項 3】

前記第1の被験体に特異的な特徴、及び、前記第2の被験体に特異的な特徴とは、それぞれ、核酸配列を含む、ことを特徴とする請求項1に記載のバイオチップシステム。

【請求項 4】

前記第1の被験体に特異的な特徴又は前記第2の被験体に特異的な特徴は、細胞型を示す、ことを特徴とする請求項1に記載のバイオチップシステム。

10

【請求項 5】

前記第1の被験体に特異的な特徴又は前記第2の被験体に特異的な特徴は、有機体の型を示す、ことを特徴とする請求項1に記載のバイオチップシステム。

【請求項 6】

前記第1の被験体に特異的な特徴又は前記第2の被験体に特異的な特徴は、種を示す、ことを特徴とする請求項1に記載のバイオチップシステム。

20

【請求項 7】

前記第1の被験体に特異的な特徴又は前記第2の被験体に特異的な特徴は、種の個々のメンバーを示す、ことを特徴とする請求項1に記載のバイオチップシステム。

【請求項 8】

前記第1の被験体に特異的な特徴又は前記第2の被験体に特異的な特徴は、抗生物質耐性の形質を示す、ことを特徴とする請求項1に記載のバイオチップシステム。

【請求項 9】

前記第1の複数の同一のプローブ及び前記第2の複数の同一のプローブは、核酸である、ことを特徴とする請求項1に記載のバイオチップシステム。

【発明の詳細な説明】

【技術分野】

【0001】

30

相互参照

本出願は、2015年1月16日出願の米国仮特許出願第61/256,049号、及び2016年3月10日出願の米国仮特許出願第62/306,597号の利益を請求するものであり、その各々は、全ての目的のために全体を参照することで本明細書に組み込まれる。

【背景技術】

【0002】

DNAマイクロアレイ（又はバイオチップ）は、標的核酸の存在に対しサンプルをプローブで探すために頻繁に使用される。マイクロアレイは、固体支持体に固定されたプローブのアレイを含んでいる。プローブのアレイは、プローブのクラスターとして構築することができ、各々が個々に対処可能である。各クラスターは複数のプローブを含むことができ、各プローブは各クラスター中の他のプローブと同一であり、各々が同じ標的核酸配列に結合することができる。サンプルがマイクロアレイにハイブリダイズされた後、プローブに結合される標的核酸の存在が判定され得る。マイクロアレイには、コスト効率が良く、サンプル中の何千から何百万もの配列の存在を判定することができるという点で非常にスケーラブルであり、且つ他の同様に調整した（scaled）手法よりも早い時間での返答を提供するという利点を提供することができる。

40

【0003】

しかし、マイクロアレイは、DNA配列決定などの他の技術と比較して低い感度及び特異性を示す場合がある。プローブは、対象の標的に対する特異性を確保するように注意深

50

く設計されなければならない。このことは、環境サンプルなどの複合サンプル、即ち遺伝物質の1より多くのソースを含むものを調べる時に、問題となり得る。更に、高レベルの遺伝的類似性を共有する2以上の被験体を区別する能力は、マイクロアレイ技術を使用して区別することが困難となる場合がある。本明細書中の方法とシステムは、複合サンプルにおける被験体の識別と区別のための新たなマイクロアレイ又はバイオチップを記載する。

【発明の概要】

【0004】

幾つかの態様において、複数の被験体に特異的な特徴を含むバイオチップが開示され、ここで、被験体に特異的な特徴はそれぞれ複数の異なるプローブを含み、複数の異なるプローブは、サンプル中で複数の被験体から1つの被験体を区別することが可能な標的に結合することができる。幾つかの実施形態において、被験体は細胞型である。幾つかの実施形態において、サンプル中の被験体は異なる細胞型である。幾つかの実施形態において、被験体は有機体である。幾つかの実施形態において、サンプル中の被験体は異なる有機体である。幾つかの実施形態において、プローブは核酸である。幾つかの実施形態において、標的はゲノム領域である。

10

【0005】

幾つかの態様において、以下の工程を含む方法が開示される：(a)複数の被験体を含むサンプルを得る工程；(b)サンプルから核酸を抽出して断片化する工程；(c)抽出して断片化した核酸をバイオチップへとハイブリダイズする工程であって、バイオチップは、複数の固有のプローブを含む、被験体に特異的な特徴を含んでいる、工程；(d)ハイブリダイズしたプローブを持つ被験体に特異的な特徴を識別するためにバイオチップを画像化する工程；及び(e)バイオチップを使用して識別される複数の被験体から被験体を列挙するレポートを提供する工程。幾つかの実施形態において、核酸は増幅されない。幾つかの実施形態において、バイオチップには100を超える被験体に特異的な特徴がある。幾つかの実施形態において、被験体は細胞型である。幾つかの実施形態において、被験体は有機体である。幾つかの実施形態において、被験体は細菌である。幾つかの実施形態において、被験体は遺伝子である。幾つかの実施形態において、被験体は保存領域である。幾つかの実施形態において、被験体は、病原性、ビルレンス、又は抗生物質耐性に関連した領域である。

20

【0006】

30

1つの態様において、1つ以上のセットのプローブを含むバイオチップが提供され、ここで、1つ以上のセットのプローブの各セットは複数のプローブを含み、複数のプローブの各々は1つ以上の被験体に特異的な特徴を含み、1つ以上のセットのプローブの各セットは、複数の異なる被験体の1人の異なる被験体からの標的核酸に結合する。場合によっては、1セットのプローブ内の複数のプローブの各々は同一である。場合によっては、1セットのプローブ内の複数のプローブの各々は異なっている。場合によっては、複数のプローブの各セットは複数の固有のプローブを含む。場合によっては、複数のプローブの各セットは、複数の固有のプローブの予め定めた平均表示を含む。場合によっては、複数の固有のプローブの平均表示は、プローブの1つ以上のセットの各セット内のプローブの総数を制限することにより、予め定めた比率で前記複数の固有のプローブを混合することにより、或いはその両方の組み合わせにより、制御される。場合によっては、前記1つ以上のセットのプローブの各々は、約2-1000の固有のプローブを含む。場合によっては、平均表示は、前記プローブのセット内に前記複数の固有のプローブの各々の約2-100の表示を含む。場合によっては、プローブの各セット内の被験体に特異的な特徴は同一である。場合によっては、前記1つ以上のセットのプローブの各セットは、異なる被験体に特異的な特徴を含む。場合によっては、前記1つ以上のセットのプローブの各セットは、個々に対応可能である。場合によっては、1セットのプローブ内の前記複数のプローブの各々は、前記標的核酸上に存在する同一の核酸配列に相補的である。場合によっては、1セットのプローブ内の前記複数のプローブの各々は、前記標的核酸上に存在する異なる核酸配列に相補的である。場合によっては、前記1つ以上のセットのプローブの各セッ

40

50

トは、被験体のゲノムの固有領域に相補的である。場合によっては、前記被験体のゲノムの固有領域は、異なる被験体のゲノムには表わされない。場合によっては、複数の異なる被験体は複数の異なる細胞型を含む。場合によっては、前記1つ以上のセットのプローブの各セットは、前記複数の異なる細胞型の1つの異なる細胞型からの標的核酸に結合する。場合によっては、複数の異なる被験体は複数の異なる有機体を含む。場合によっては、複数の異なる被験体は複数の異なる個体を含む。場合によっては、複数の異なる被験体は複数の異なる株を含む。場合によっては、複数の異なる被験体は複数の異なる遺伝子を含む。場合によっては、複数の異なる被験体は複数の異なるゲノム領域を含む。場合によっては、前記1つ以上のセットのプローブの各セットは、前記複数の異なる有機体の1つの異なる有機体からの標的核酸に結合する。場合によっては、複数のプローブは核酸分子を含む。場合によっては、複数のプローブは固体支持体に固定される。場合によっては、固体支持体はビーズである。場合によっては、被験体に特異的な特徴は1つ以上の遺伝子特徴を含む。場合によっては、1つ以上の遺伝子特徴は、以下から成る群から選択される：種を表わすゲノム、種の中の株を表わすゲノム、クロマチン、染色体、染色体遺伝子座、染色体物質、対立遺伝子、遺伝子、遺伝子クラスター、遺伝子座、遺伝的多型、遺伝子突然変異、スクレオチド、一塩基多型(SNP)、制限酵素断片長多型(RFLP)、タンデム反復数(variable tandem repeat)(VTR)、コピー数多型(CNV)、マイクロサテライト配列、遺伝子マーカー、配列マーカー、配列でタグ付けした部位(STR)、プラスミド、転写ユニット、転写産物、遺伝子発現状態、保存領域、病原性アイランド、及びそれらの任意の組み合わせ。場合によっては、1つ以上のセットのプローブは、100より多くのセットのプローブを含む。場合によっては、前記1つ以上のセットのプローブの各々は、約50-1000のプローブを含む。

【0007】

1つの態様において、以下を含む方法が提供される：a)複数の異なる被験体に由来する複数の核酸を含むサンプルを提供する工程であって、前記複数の核酸は、前記複数の異なる被験体の少なくとも2つからの少なくとも1つの標的核酸を含む、工程；b)前記複数の核酸をバイオチップにハイブリダイズする工程であって、前記バイオチップは1つ以上のセットのプローブを含み、前記1つ以上のセットのプローブの各セットは複数のプローブを含み、前記複数のプローブの各々は1つ以上の被験体に特異的な特徴を含み、前記1つ以上のセットのプローブの各セットは、前記複数の異なる被験体のうち少なくとも2つからの前記少なくとも1つの標的核酸に結合する、工程；c)前記複数のプローブの1つのプローブへの前記少なくとも1つの標的核酸の結合に関連したシグナルを検出する工程；及びd)前記サンプル中の前記少なくとも1つの標的核酸の存在に基づいて前記複数の異なる被験体を識別する工程。場合によっては、前記方法は、工程a)の前に、前記複数の異なる被験体から前記複数の核酸を抽出する工程を更に含む。場合によっては、前記方法は、工程b)の前に、前記複数の核酸を断片化する工程を更に含む。場合によっては、前記方法は、工程b)の前に、前記複数の核酸を増幅する工程を更に含む。場合によっては、複数の核酸は増幅されない。場合によっては、前記方法は、前記複数の異なる被験体を識別する1つ以上のレポートを提供する工程を更に含む。場合によっては、複数の異なる被験体は複数の異なる細胞型を含む。場合によっては、複数の異なる被験体は複数の異なる有機体を含む。場合によっては、1セットのプローブ内の前記複数のプローブの各々は同一である。場合によっては、1セットのプローブ内の前記複数のプローブの各々は異なっている。場合によっては、前記複数のプローブの各セットは複数の固有のプローブを含む。場合によっては、前記複数のプローブの各セットは、前記複数の固有のプローブの平均表示を含む。場合によっては、前記複数の固有のプローブの平均表示は、前記1つ以上のセットのプローブの各セット内のプローブの総数を制限することにより、予め定めた比率で前記複数の固有のプローブを混合することにより、或いはその両方の組み合わせにより、制御される。場合によっては、前記1つ以上のセットのプローブの各々は、約2-1000の固有のプローブを含む。場合によっては、平均表示は、前記プローブのセット内に前記複数の固有のプローブの各々の約2-1000の表示を含む。場合によっては

10

20

30

40

50

、プローブの各セット内の被験体に特異的な特徴は同一である。場合によっては、前記1つ以上のセットのプローブの各セットは、異なる被験体に特異的な特徴を含む。場合によっては、前記1つ以上のセットのプローブの各セットは、個々に対処可能である。場合によっては、1セットのプローブ内の前記複数のプローブの各々は、前記標的核酸上に存在する同一の核酸配列に相補的である。場合によっては、1セットのプローブ内の前記複数のプローブの各々は、前記標的核酸上に存在する異なる核酸配列に相補的である。場合によっては、前記1つ以上のセットのプローブの各セットは、被験体のゲノムの固有領域に相補的である。場合によっては、被験体のゲノムの固有領域は、異なる被験体のゲノムには表わされない。場合によっては、複数のプローブは核酸分子を含む。場合によっては、複数のプローブは固体支持体に固定される。場合によっては、固体支持体はビーズである。場合によっては、1つ以上の被験体に特異的な特徴は1つ以上の遺伝子特徴を含む。場合によっては、1つ以上の遺伝子特徴は、以下から成る群から選択される：種のゲノム、株のゲノム、クロマチン、染色体、染色体遺伝子座、染色体物質、対立遺伝子、遺伝子、遺伝子クラスター、遺伝子座、遺伝的多型、遺伝子突然変異、ヌクレオチド、一塩基多型（S N P）、制限酵素断片長多型（R F L P）、タンデム反復数（V T R）、コピー数多型（C N V）、マイクロサテライト配列、遺伝子マーカー、配列マーカー、配列でタグ付けした部位（S T S）、プラスミド、転写ユニット、転写産物、遺伝子発現状態、保存領域、病原性アイランド、及びそれらの任意の組み合わせ。場合によっては、1つ以上のセットのプローブは、100より多くのセットのプローブを含む。場合によっては、前記1つ以上のセットのプローブの各々は、約50 - 1000のプローブを含む。場合によっては、少なくとも1つの標的核酸は、検出可能な標識で標識される。場合によっては、検出可能な標識は蛍光染料を含む。場合によっては、前記複数のプローブの第1のプローブは第1の被験体に特異的な特徴を含み、前記複数のプローブの第2のプローブは第2の被験体に特異的な特徴を含み、前記第1のプローブと前記第2のプローブは、前記少なくとも1つの標的核酸にハイブリダイズする。

【0008】

別の態様において、標識された核酸断片を產生するための方法が提供され、該方法は、以下を含む：(a)二本鎖である標的核酸を提供する工程；(b)前記標的核酸を、(i)トランスポゾンと(ii)標識で標識されるオリゴヌクレオチドとを含むトランスポソームと接触させる工程；及び(c)前記トランスポソームで、前記標的核酸から核酸断片を產生する工程であって、前記核酸断片は二本鎖であり、且つ(i)前記標的核酸の一部と(ii)前記標識とを含む、工程。請求項60に記載の方法は、標識された一本鎖断片をもたらすために前記核酸断片を変性させる工程を更に含む。場合によっては、前記方法は、前記標識された一本鎖断片をアレイにハイブリダイズする工程を更に含む。場合によっては、前記核酸断片は前記オリゴヌクレオチドの少なくとも一部を更に含む。場合によっては、前記標識は蛍光標識である。場合によっては、前記核酸断片は5'末端にて前記標識を含む。

【0009】

1つの態様において、第1の鎖と第2の鎖とを含む、標的核酸の二本鎖断片；前記第1の鎖に共有結合された第1のオリゴヌクレオチド；前記第1のオリゴヌクレオチドに結合される第1の標識；前記第2の鎖に共有結合された第2のオリゴヌクレオチド；及び前記第2のオリゴヌクレオチドに結合される第2の標識、を含む組成物が提供される。場合によっては、前記第1のオリゴヌクレオチドは、前記第1の鎖の5'末端にて前記第1の鎖に共有結合され、前記第2のオリゴヌクレオチドは、前記第2の鎖の5'末端にて前記第2の鎖に共有結合される。場合によっては、前記第1の標識と第2の標識は蛍光標識を含む。

【0010】

1つの態様において、光学的分解能を特徴とする光検出器；及び前記光検出器に光学的に接続されたバイオチップを含むバイオチップシステムが提供され、前記バイオチップは、(i)第1の複数の同一のプローブを含む第1の特徴、及び(ii)第2の複数の同一のプローブを含む第2の特徴を含み、前記第1の複数の同一のプローブと前記第2の複数

10

20

30

40

50

の同一のプローブは、互いに異なっており；ここで、前記第1の特徴と前記第2の特徴は、前記光学的分解能未満の領域或いはそれに略等しい領域に包含される。場合によっては、前記光学的分解能は、前記光検出器のピクセルサイズにより判定される。場合によっては、前記第1の複数の同一のプローブは被験体の第1の被験体に特異的な特徴を標的とし、前記第2の複数の同一のプローブは前記被験体の第2の被験体に特異的な特徴を標的とする。場合によっては、前記第1の被験体に特異的な特徴及び前記第2の被験体に特異的な特徴は、異なるものである。場合によっては、前記第1の被験体に特異的な特徴及び前記第2の被験体に特異的な特徴はそれぞれ、核酸配列を含む。場合によっては、前記第1の被験体に特異的な特徴又は前記第2の被験体に特異的な特徴は、細胞型を示す。場合によっては、前記第1の被験体に特異的な特徴又は前記第2の被験体に特異的な特徴は、有機体の型を示す。場合によっては、前記第1の被験体に特異的な特徴又は前記第2の被験体に特異的な特徴は、種を示す。場合によっては、前記第1の被験体に特異的な特徴又は前記第2の被験体に特異的な特徴は、種の個々のメンバーを示す。場合によっては、前記第1の被験体に特異的な特徴又は前記第2の被験体に特異的な特徴は、抵抗の形質 (resistance trait) を示す。場合によっては、前記第1の複数の同一のプローブ及び前記第2の複数の同一のプローブは、核酸である。

【0011】

1つの態様において、(i) 第1のプローブと第2のプローブとを含む第1の特徴、及び(ii) 前記第1のプローブと前記第2のプローブとを含む第2の特徴を含むバイオチップ、を含むデバイスが提供され、前記第1のプローブは前記第2のプローブとは異なる。場合によっては、前記第1のプローブは被験体の第1の被験体に特異的な特徴を標的とし、前記第2のプローブは前記被験体の第2の被験体に特異的な特徴を標的とする。場合によっては、前記第1の被験体に特異的な特徴及び前記第2の被験体に特異的な特徴は、異なるものである。場合によっては、前記第1の被験体に特異的な特徴及び前記第2の被験体に特異的な特徴は、核酸配列を含む。場合によっては、前記第1の被験体に特異的な特徴又は前記第2の被験体に特異的な特徴は、細胞型を示す。場合によっては、前記第1の被験体に特異的な特徴又は前記第2の被験体に特異的な特徴は、有機体の型を示す。場合によっては、前記第1の被験体に特異的な特徴又は前記第2の被験体に特異的な特徴は、種を示す。場合によっては、前記第1の被験体に特異的な特徴又は前記第2の被験体に特異的な特徴は、種の個々のメンバーを示す。場合によっては、前記第1の被験体に特異的な特徴又は前記第2の被験体に特異的な特徴は、抵抗の形質を示す。場合によっては、前記第1の複数の同一のプローブ及び前記第2の複数の同一のプローブは、核酸である。場合によっては、前記第1の特徴から産生されたシグナルは、前記第1のプローブと前記第2のプローブの両方の標的の存在を示す。場合によっては、前記第1の特徴及び前記第2の特徴から産生されたシグナルは、信頼値における特異的な種、株、遺伝子、又はゲノム特徴の存在を示す。

【図面の簡単な説明】

【0012】

【図1】バイオチップシステムの実施形態を例示する。図1のAは4つの典型的な特徴を示し、各特徴は、サンプル中に同一のプローブ、及び1つの被験体からの4つの結合されていない標識された標的を含む。図1のBは、サンプル中に4つのプローブと、1つの被験体からの4つの結合されていない標識された標的とを含む、被験体に特異的な特徴を示す。図1のCは、4つの異なる特徴に結合された4つの標的を示す。図1のDは、1つの被験体に特異的な特徴に結合された4つの標的を示す。図1のCを図1のDと比較すると、複数の被験体標的に配向された複数の異なるプローブを使用した時に1つの特徴に生じ得るシグナル増幅が実証される。図1のEは、1つの異なる特徴内での固有のプローブの順序づけたプール (pooling) を含む特徴を示す。図1のFは、特徴の中に固有のプローブのランダムプールを含む特徴を示す。

【図2】1つの特徴ごとに結合するフルオロフォアの数の増大により得ることができる相対的なシグナルを示す。

【図3】本明細書に開示される方法を実行するのに適切な、典型的なコンピュータシステムを示す。

【図4】M 1 3 m p 1 8 ファージベクトル配列上で識別された22の固有領域を示す。

【図5】本明細書に記載される方法を使用してM 1 3 m p 1 8 ファージベクトル配列に対して設計された固有のプロープの例を示す。

【図6】トランスポソーム複合体を使用した二本鎖DNAの一工程の(one - step)断片化及び標識化の典型的な図を示す。

【図7 A】別個の捕捉配列及び検出配列でのハイブリダイゼーションを介した標的核酸のハイブリダイゼーション及び検出を示す。

【図7 B】別個の捕捉配列及び検出配列でのハイブリダイゼーションを介した標的核酸のFRETシグナル検出を示す。

【図7 C】複数のハイブリダイゼーション配列を介した、アレイへの標的核酸のハイブリダイゼーションを表す。

【図8】フィルター基板上のオリゴヌクレオチドアレイに濃縮された核酸の典型的な概略図を示す。

【図9 A】典型的なコンピュータシステムを表す。

【図9 B】コンピュータシステムの典型的構成を表す。

【図9 C】コンピュータシステムの典型的なネットワークを表す。

【図9 D】典型的なマルチプロセッサコンピュータシステムを表す。

【図10】非ウイルスの結核に関するアッセイの結果を示す。

【発明を実施するための形態】

【0013】

定義

以下の用語は、当業者によるこれらの用語の理解に加えて、本明細書において使用される用語の意味を例示するために説明される。本明細書および特許請求の範囲で使用されるように、「a」、「an」、および「the」は、その内容が他に明確に示していない限り、複数の参照を含む。例えば、用語「細胞」は複数の細胞を含み、その混合物も含む。

【0014】

明細書において使用されるように、用語「エピゲノム」は、DNAメチル化および染色質再編成、あるいは再形成などの配列レベルに反映されない遺伝物質への変更、あるいは遺伝物質のタンパク質発現を指す。「トランスクリプトーム」は、特定の環境条件の下で、有機体によって合成された遺伝子転写産物(mRNA)の全体を指す。トランスクリプトームデータセットには、対象の遺伝子発現の活性化あるいは非活性化に関する質および量的情報が、制限なく含まれる。トランスクリプトームはさらに、microRNA、piwiRNA、構造RNA、タンパク質と結合するRNA、テロメラーゼRNA、およびトランスポゾンRNAを含む、タンパク質をコードしないRNA転写産物(非コードRNAあるいはncRNA)を含む。「エクソーム」は、エクソソムによって形成されたゲノムの部分、すなわち転写された時に成熟RNA内に残存する配列を指す。「マイクロバイオーム」は、通常は微生物を起原とするが、種に関わらず、生体サンプル内のゲノム全体を指す。

【0015】

明細書において使用されるように、用語「遺伝的特徴」は、任意のゲノム、遺伝子型、ハプロタイプ、染色質、染色体、染色体座、染色体物質、デオキシリボ核酸(DNA)、対立遺伝子、遺伝子、遺伝子集団、遺伝子座、遺伝的多形、遺伝子突然変異、遺伝子突然変異比率、ヌクレオチド、ヌクレオチド塩基対、一塩基多型(SNP)、制限酵素断片長多型(RFLP)、タンデム反復数(VTR)、コピー数多型(CNV)、マイクロサテライト配列、遺伝子マーカー、配列マーカー、配列でタグ付けした部位(STS)、プラスミド、転写ユニット、転写産物、遺伝子発現レベル、遺伝子発現(例えば転写)状態、リボ核酸(RNA)、相補的DNA(cDNA)、保存領域、および病原性アイランドを指し、上記のいずれかに関連するヌクレオチド配列およびコードされたアミノ酸配列を含

10

20

30

40

50

む。後成的な特徴は、体細胞分裂中に遺伝され得る、および時には生殖細胞系伝達において遺伝され得るよう遺伝子発現に影響を与えるが、DNA塩基配列へと突然変異せず、したがって基本的に可逆性であり、限定されないがDNAヌクレオチドのメチル化およびクロマチン結合性ヒストンタンパク質のアセチル化を含む、遺伝物質、すなわちすべてのゲノム、媒介体、およびプラスミドDNAと染色質等の任意の特徴である。したがって、明細書において使用されるように、遺伝子配列データは、ヌクレオチド配列、デオキシリボ核酸(DNA)配列、およびリボ核酸(RNA)配列を、制限なく含み得る。

【0016】

本明細書において使用される用語「被験体に特異的な特徴」は、ある被験体を他から区別することができる任意の特徴あるいは属性を指し得る。場合によっては、被験体に特異的な特徴は遺伝的特徴である。上記のように、遺伝的特徴は、被験体から単離した核酸上で存在することができる。場合によっては、被験体に特異的な特徴は、一連の機能を区別する1つ、あるいは複数の特徴に関連し得る。これは、例えば、単一の遺伝子、複数の遺伝子、既知のエピゲノムの機能を備えるプロモーター領域等のゲノム領域を標的にするために設計されたプローブによって達成され得る。被験体に特異的な特徴は、バイオチップ上のプローブとして表すことができる。被験体に特異的な特徴を表すプローブは、被験体から得られた1つ以上の標的核酸配列に結合することが可能でありうる。場合によっては、被験体に特異的な特徴は、各々が被験体を他から区別可能な、複数の非同一のプローブを含む。場合によっては、細菌株などの特定の被験体は、標的株に特有の、標的株を含む種に特有の、標的株に存在する保存領域に含まれる、あるいは標的株に含まれる病原性アイランドを認識する特徴を含む、バイオチップ上の1つまたは複数の特徴により区別可能である。場合によっては、単純に病原性アイランドを識別するために有益で有り得、なぜならこれは被験体がより多くの試験を必要とすることを示し得るからである。

10

【0017】

用語「アセンブリ」は、シーケンサーまたは質量分析計によって產生された配列の連続が、最初の配列の連続を再構成する目的で互いに融合され、全ての配列の連続一式が抽出される、任意の計算プロセスで有り得る。いくつかの例では、アセンブリは個々の有機体からである。いくつかの例では、多数の個体がアセンブリを作成するために使用され得る。いくつかの例では、アセンブリは、基準配列を使用することなく、新たに作成される。いくつかの例では、アセンブリは基準配列を用いて作成される。基準配列は同じ種からのゲノムで有り得る。基準ゲノムは、密接に関係した種からのゲノムで有り得る。

20

【0018】

本明細書で使用される用語「被験体」は、一般に遺伝物質の特定の供給源を指す。被験体は生物学的存在で有り得る。生物学的存在は、例えば、バクテリア、ウイルス、真菌類、および原生動物を含む、植物、動物、あるいは微生物であり得る。被験体は、器官、組織あるいは細胞で有り得る。被験体は、インビオで得られる、あるいはインビトロで培養され得る。被験体は細胞株で有り得る。被験体は培養で繁殖され得る。被験体は疾病細胞で有り得る。被験体は癌細胞で有り得る。被験体は哺乳動物で有り得る。哺乳動物はヒトで有り得る。被験体は、遺伝物質の特定の供給源の個別の代表を意味する場合がある(例えば、被験体は特定の個人あるいは特定の菌株で有り得る)。代替的に、被験体は、遺伝物質の特定の供給源の一種を一般的に代表するもので有りえ、例えば、被験体は単一の種の任意の、そして全ての構成員で有り得る。被験体はさらに、例えばサンプルが完全なゲノムを含んでいない場合、ゲノムの一部で有り得る。

30

【0019】

「サンプル」あるいは「核酸サンプル」は、核酸を含む、あるいは含んでいると推定される任意の物質を指し得る。サンプルは被験体から得られた生体サンプルで有り得る。核酸は、RNA、DNA、例えばゲノムDNA、ミトコンドリアDNA、ウイルスDNA、合成DNAなどのDNA、あるいはRNAから逆転写されたcDNAで有り得る。核酸サンプル中の核酸は、ハイブリダイズされたプライマーを拡張するためのテンプレートとなり得る。幾つかの場合には、生体サンプルは液状サンプルである。液状サンプルは、例え

40

50

ば全血、血漿、血清、腹水、精液、脳脊髄液、汗、尿、涙、唾液、頬側サンプル、窓洗浄水あるいは器官洗浄水で有り得る。液状サンプルは、本質的に無細胞の液状サンプル（例えば血漿、血清、汗、尿、涙等）で有り得る。その他の場合において、生体サンプルは、例えば糞便、毛、爪、例えば腫瘍生検などの組織生検によって得られた試料等の、固体の生体サンプルで有り得る。サンプルはさらにインビトロ細胞培養成分（限定されないが、細胞培養培地中の細胞の成長によってもたらされる調整培地、組み換え細胞、および細胞構成体を含む）を含み得る。サンプルは、癌細胞を含む、あるいは癌細胞由来で有り得る。サンプルはマイクロバイオームを含み得る。

【 0 0 2 0 】

本明細書において使用される「複合サンプル」は、2つ以上の被験体を含む、あるいは2つ以上の被験体からの物質（例えば核酸）を含むサンプルを指す。複合サンプルは、2つ以上の被験体からの遺伝物質を含むことができる。複合サンプルは、2つ以上の被験体からの核酸分子を含むことができる。複合サンプルは、バクテリア、ウイルス、菌類などの2つ以上の株からの核酸を含むことができる。複合サンプルは2つ以上の分解可能な被験体（つまり互いに区別可能な2つ以上の被験体）を含むことができる。場合によっては、複合サンプルは環境から得ることができる。例えば、複合サンプルは、空気サンプル、土または泥のサンプル、あるいは水サンプル（例えば川、湖、海洋、廃水等）で有り得る。環境サンプルは、バクテリア、ウイルス、原生動物、ソウ類、菌類およびその他同種のものの1つ以上の種を含むことができる。

【 0 0 2 1 】

「ヌクレオチド」は核酸を形成可能な生体分子で有り得る。ヌクレオチドは、既知のプリンおよびピリミジン塩基だけでなく、修飾された他の複素環式塩基も含む部分を有し得る。そのような修飾は、メチル化されたプリンまたはピリミジン、アシル化されたプリンまたはピリミジン、アルキル化されたリボース、あるいは他の複素環を含む。加えて、用語「ヌクレオチド」は、ハプテン、ビオチン、あるいは蛍光標識を含み、従来型のリボースおよびデオキシリボース糖だけでなく他の糖も同様に含み得る部分を含む。修飾されたヌクレオシドあるいはヌクレオチドは、さらに、糖部分の修飾を含み、例えばそこでは、ヒドロキシル基の1つ以上がハロゲン原子または脂肪族基で取り換えられているか、エーテル、アミン等として官能化される。

【 0 0 2 2 】

「ヌクレオチド」はさらにロックド核酸（LNA）あるいは架橋された核酸（BNA）を含み得る。BNAとLNAは、2'酸素および4'炭素を結合する架橋で、リボース部分が修飾された修飾リボヌクレオチドを一般的に指す。一般に架橋は、しばしばA型二本鎖に見られる3' - endo (North) 立体構造中のリボースを「ロック」する。用語「ロックド核酸」（LNA）は一般に、2' - O原子を4' - C原子に結合するメチレン架橋でリボース環が「ロックされる」BNAのクラスを指す。DNAとRNAに現われる6つの共通の核酸塩基（T、C、G、A、UおよびmC）を含むLNAヌクレオシドは、標準的なワトソン - クリック型塩基対規則に従って、それらの相補的なヌクレオシドで塩基対を形成することができる。したがって、BNAおよびLNAヌクレオチドは、望まれればいつでも、オリゴヌクレオチド中のDNAまたはRNA塩基と混合可能である。ロックドリボース立体構造は塩基の積み重ねと脊椎の事前の組織化を増強する。塩基の積み重ねと脊椎の事前の組織化は、熱的安定性（例えば、増大したTm）および二重螺旋の識別力の増加をもたらすことができる。LNAは、他の核酸では識別することができない状況下での、単一塩基のミスマッチを識別することができる。

【 0 0 2 3 】

用語「ポリヌクレオチド」、「核酸」、「ヌクレオチド」および「オリゴヌクレオチド」は、区別なく使用することができる。それらは、任意の長さのヌクレオチド、デオキシリボヌクレオチドまたはリボヌクレオチドのいずれか、またはそのアナログの高分子成形体を指し得る。ポリヌクレオチドは任意の三次元構造を有し得、任意の未知または既知の機能を行い得る。下記はポリヌクレオチドの非限定的な例である：遺伝子または遺伝子断

10

20

30

40

50

片のコードあるいは非コード領域、連鎖解析から限定された座 (loci / locus) 、エクソン、イントロン、メッセンジャー RNA (mRNA) 、転移 RNA 、リボソーム RNA 、リボザイム、cDNA 、組み換えポリヌクレオチド、分岐ポリヌクレオチド、プラスミド、ベクター、任意の配列の分離した DNA 、任意の配列の分離した RNA 、核酸プローブ、およびプライマー。ポリヌクレオチドは、メチル化されたヌクレオチドおよびヌクレオチドアナログ等の修飾されたヌクレオチドを含み得る。ヌクレオチド構造に対する修飾が存在する場合、それはポリマーのアセンブリの前または後で与えられ得る。ヌクレオチドの配列は非ヌクレオチド成分によって妨げられ得る。ポリヌクレオチドは、標識要素との結合等によって、重合の後にさらに修飾され得る。

【0024】

10

「変異体」は、核酸配列 (例えば遺伝子) の正常な配列の変質で有り得る。いくつかの例では、遺伝子型とそれに対応する表現型は変異体に関連する。他の例では、変異体の既知の作用はない。変異体は SNP で有り得る。変異体は SNV で有り得る。変異体は複数のヌクレオチドの挿入で有り得る。変異体は複数のヌクレオチドの欠失で有り得る。変異体は突然変異で有り得る。変異体はコピー数多型で有り得る。変異体は構造変異種で有り得る。変異体は、集団中の 2 つ以上の個体間の核酸偏差で有り得る。

【0025】

本明細書で使用される用語「標的ポリヌクレオチド」あるいは「標的核酸」は、一般に研究対象のポリヌクレオチドを指す。ある場合には、標的ポリヌクレオチドは、研究対象の 1 つ以上の配列を含む。標的ポリヌクレオチドは、例えばゲノム配列を含むことができる。標的ポリヌクレオチドはその存在、量、および / またはヌクレオチド配列、あるいはこれらにおける変化が判定されることが望まれる、標的配列を含むことができる。標的ポリヌクレオチドは、ゲノムの非コード領域を含むことができる。

20

【0026】

用語「ゲノム」は、生体の遺伝的相補体を指し得る。そして用語「ゲノムデータ」および「ゲノムデータセット」は、生体の染色体、遺伝子あるいは DNA の配列情報を含む。

【0027】

本明細書において使用される用語「ゲノムデータ」は、以下の 1 つ以上で有り得るデータを指す：ミトコンドリア、細胞、卵子と精子を含み、組織、新生物、腫瘍、器官、有機体、微生物、ウイルス、個体、あるいは無細胞 DNA の、 1 つ以上のゲノムまたはエクソン配列、あるいは 1 つ以上の任意の組み合わせまたは混合、およびさらに限定されないが、核酸配列情報、遺伝子型情報、遺伝子発現情報、遺伝子データ、DNA メチル化、アセチル化、あるいは類似の DNA 修飾データを含む後成の情報、RNA 転写、スプライシング、編集あるいは加工情報、あるいは医療、健康あるいは表現型のデータ、または栄養素、食事あるいは環境の状態、あるいは露出情報、あるいは任意の微生物、ウイルス、細胞、組織、新生物、腫瘍、器官、臓器系、無細胞サンプル (例えば血清または培地) についての他の属性データ、個別あるいはひとかたまりのサンプル、または個体を含む。したがって、本明細書において使用される用語「ゲノムの配列」は、ゲノムに発生する配列を指す。RNA がゲノムから転写されるため、この用語は、有機体の核ゲノムに存在する配列を含み、そのようなゲノムから転写された RNA (例えば mRNA) の cDNA コピーにある配列も同様に含む。「ゲノム配列」はさらに、細胞質上またはミトコンドリア中に発生する配列で有り得る。

30

【0028】

40

用語「判定すること (determining) 」、「測定すること (measuring) 」、「評価すること (evaluating) 」、「検出すること (assessing) 」、「アッセイすること (assaying) 」、及び「分析すること (analyzing) 」は、測定の任意の形態を指し、本明細書で互換的に使用され得る。およびある要素の存在の有無を判定することを含み得る。このような用語は、定量的及び / 又は定質的両方の判定を含み得る。評価は相対的、又は絶対的な場合がある。「～の存在を検出すること」は、存在の有無を判定することと共に、存在するものの量を判定することを含み

50

得る。

【 0 0 2 9 】

用語「ゲノム断片」は、本明細書において使用されるように、例えば、ヒト、サル、ラット、魚あるいは昆虫や植物等の、動物または植物のゲノムの、ゲノム領域を指し得る。ゲノム断片はアダプターライゲートされることもあるが、されないこともある。ゲノム断片は、アダプターライゲートされ得る（その場合、少なくとも分子の5'末端に、断片の1つあるいは両方の末端にライゲートされたアダプターを有する）、あるいはアダプター以外でライゲートされることもある。

【 0 0 3 0 】

本明細書において使用される用語「バーコード」は、一般的には、アッセイについての情報をコード可能なヌクレオチド配列を指す。幾つかの例では、バーコードは唯一無二である。バーコード配列は、調べられた対立遺伝子の正体、標的ポリヌクレオチドあるいはゲノム遺伝子座の正体、サンプルの正体、被験体、あるいはそれらの任意の組み合わせに関する情報をコードすることができる。バーコード配列は、プライマー、リポータープローブ、あるいはその両方の一部で有り得る。バーコード配列は、オリゴヌクレオチドの5' - 末端または3' - 末端ありえ、あるいはオリゴヌクレオチドの任意の領域に位置し得る。バーコード配列は、例えば研究下のサンプルに存在しない配列などの、非自然的に発生したもので有り得る。他の例では、自然発生する配列は、バーコードあるいはバーコード配列の一部として使用することができる。いくつかの例では、核酸が結合されている結合点は、バーコードとなり得る。いくつかの例では、配列決定アダプターは、バーコードあるいはバーコードの一部となり得る。いくつかの例では、バーコードは、標的分子、例えば対象のゲノム配列、を超過している。いくつかの例では、バーコードは、ランダムあるいは半ランダムに標的分子に関連付けられる。いくつかの例では、バーコードは、設計によって標的分子に関連付けられている。

10

20

30

40

【 0 0 3 1 】

用語「突然変異」は、本明細書において使用されるように、一般にゲノムのヌクレオチド配列の変更を指す。突然変異は、DNAの大きな部位を含み得る（例えばコピー数多型）。突然変異は染色体全体を含み得る（例えば異数性）。突然変異は、DNAの小さな部位を含み得る。DNAの小さな部位を含む突然変異の例は、例えば、点変異あるいは一塩基多型、多数の一塩基多型、挿入（例えば、座における1つ以上のヌクレオチドの挿入）、多数のヌクレオチドの変更、欠失（例えば、座における1つ以上のヌクレオチドの欠失）、反転（例えば、1つ以上のヌクレオチド配列の反転）を含む。

【 0 0 3 2 】

本明細書において使用される用語「遺伝子座」は、遺伝子、ヌクレオチド、あるいは染色体上の配列の位置を指し得る。本明細書において使用される遺伝子座の「対立遺伝子」は、遺伝子座におけるヌクレオチドあるいは配列の代替的な形態を指し得る。「野生型対立遺伝子」は一般的に、被験体集団において最も高い頻度を有する対立遺伝子を指す。「野生型」対立遺伝子は一般的に疾病に関係しない。「変異遺伝子」は一般的に、より低い頻度の「野生型対立遺伝子」を有する対立遺伝子を指し、および疾病に関連付けられ得る。「変異遺伝子」は疾病に関連付けられる必要のないこともある。用語「調べられた対立遺伝子」は、一般的に、アッセイで検知するように設計された対立遺伝子を指す。

【 0 0 3 3 】

本明細書において使用される用語「一塩基多型」あるいは「SNP」は、ある配列内の単一のヌクレオチドの置換によってもたらされる、一種のゲノム配列変種を一般的に指す。「SNP対立遺伝子」あるいは「SNPの対立遺伝子」は、一般的に特定の座におけるSNPの代替的な形態を指す。用語「調べられたSNP対立遺伝子」は、一般的に、アッセイで検知するように設計されたSNP対立遺伝子を指す。

【 0 0 3 4 】

本明細書は、サンプル中の1つ以上の被験体を識別する、あるいは被験体についての重要な特徴、例えば、病原性、毒性、あるいは抗生物質耐性などを識別する能力を有する、

50

新規性のあるバイオチップのための方法およびシステムを開示する。バイオチップは、被験体に特異的な1つ以上の特徴を含む複数のプローブを含むことができる。本明細書において使用される用語「被験体に特異的な特徴」は、ある被験体を他から区別および識別可能な複数のプローブを指す。本発明のいくつかの態様では、被験体に特異的な特徴は、複合サンプル中にある被験体を識別するために利用することができる。複合サンプルは、1つより多くの被験体（すなわち、2つ以上の被験体）からの物質を含む、生物学的あるいは他の任意のサンプルでありうる。場合によっては、被験体は、ウイルス、細菌、原生動物、真菌、およびその他同種のもの等の有機体である。その他の場合において、被験体はそれら由来の組織、器官あるいは細胞である。組織、器官あるいは細胞は、ヒトなどの動物に由来する場合がある。複合サンプルは複数の細胞型を含むことができる。場合によっては、複合サンプルは、腫瘍生検などの組織生検によって得られた試料を含むことができる。いくつかの例においては、複合サンプルは、微生物の2つ以上の株（例えば、バクテリア、ウイルス、真菌など）を含む。他の例においては、複合サンプルは、2つ以上の微生物の種を含む。場合によっては、複合サンプルは、2つ以上の被験体からの核酸などの物質を含む。被験体に特異的な特徴は、複合サンプル中にある1つ以上の被験体の正体を判定するために使用することができる。本明細書における方法とシステムは、任意の1つのタイプの複合サンプルに限定されない。重要な態様は、複合サンプルが少なくとも1つの区別可能な特徴をもつ複数の被験体を含むということである。

【0035】

複合サンプルは核酸の混合物を含み得る。核酸は1つを超える被験体に由来する場合がある。核酸のサンプルを生成するどんな方法も本開示によって許容可能である。場合によっては、生体細胞を含む複合サンプルが得られ、生体細胞をその後、溶解させて細胞から核酸を放出する。核酸を物理的な方法でも生体細胞から放出することができる。他の場合には、無細胞の核酸が得られる。無細胞の核酸は、ヒトまたは動物から、例えば、血液から得ることができる。無細胞の核酸は環境からも得ることができ、例えば、核酸は有機体から環境へと放出される。無細胞の核酸は、例えば、ウイルスのカプシド、あるいは胞子内に含まれる病原体に由来する場合がある。

【0036】

複合サンプル内の核酸は標的核酸配列を含むことができる。標的核酸配列は1つの被験体を別の被験体と区別する核酸配列であり得る。例えば、標的核酸配列は、被験体Bで見られない被験体Aの複数のゲノムの配列であり得る。こうした標的核酸配列は、被験体Aと被験体Bを含む複合サンプル中で被験体Aの存在を特定するために利用できる。同様に、標的核酸配列は、被験体Aで見られない被験体Bの複数のゲノムの配列であり得る。こうした標的核酸配列は、被験体Aと被験体Bを含む複合サンプル中で被験体Bの存在を特定するために利用できる。場合によっては、バイオチップは、被験体Aを被験体Bと識別することができる（つまり、単に被験体Aを認識するプローブを有する）か、あるいは被験体Bを被験体Aと識別することができる（つまり、単に被験体Bを認識するプローブを有する）か、あるいは被験体Aと被験体Bの両方を識別することができる（つまり、被験体Aを認識するプローブと被験体Bを認識するプローブを有する）。

【0037】

場合によっては、本明細書の方法とシステムは、2、3、4、5、6、7、8、9、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、100、200、300、400、500、600、700、800、900、1000、10000、あるいは10000を超える被験体を区別することができる。場合によっては、バイオチップは、2、3、4、5、6、7、8、9、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、100、200、300、400、500、600、700、800、900、1000、10000、あるいは10000を超える被験体に特異的な特徴を含む。場合によっては、本明細書の方法とシステムは、少なくとも2、3、4、5、6、7、8、9、10、15、20、25、30、35、40、45、50、55

10

20

30

40

50

、 6 0 、 6 5 、 7 0 、 7 5 、 8 0 、 8 5 、 9 0 、 9 5 、 1 0 0 、 2 0 0 、 3 0 0 、 4 0 0
 、 5 0 0 、 6 0 0 、 7 0 0 、 8 0 0 、 9 0 0 、 1 0 0 0 、 1 0 0 0 0 、 あるいは 1 0 0 0
 0 を超える被験体を区別することができる。場合によっては、バイオチップは、少なくとも 2 、 3 、 4 、 5 、 6 、 7 、 8 、 9 、 1 0 、 1 5 、 2 0 、 2 5 、 3 0 、 3 5 、 4 0 、 4 5
 、 5 0 、 5 5 、 6 0 、 6 5 、 7 0 、 7 5 、 8 0 、 8 5 、 9 0 、 9 5 、 1 0 0 、 2 0 0 、 3
 0 0 、 4 0 0 、 5 0 0 、 6 0 0 、 7 0 0 、 8 0 0 、 9 0 0 、 1 0 0 0 、 1 0 0 0 0 、 あるいは 1 0 0 0 0 を超える被験体に特異的な特徴を含む。

【 0 0 3 8 】

標的核酸配列は、複合サンプル内に含まれる核酸上に存在する 1 つ以上の核酸配列であり得る。標的配列は、サンプル内のその遺伝的補体に結合する（例えば、化学的に結合する）ように設計可能である。1 つ以上の核酸配列は 1 つの他の核酸配列とは区別可能であり、それによって、サンプル内の核酸の起原を解決する能力を与える。例えば、複合サンプルは 2 つを超える被験体を含み得る。個々の被験体はそれぞれ核酸を含むことがあり、したがって、複合サンプルは個々の被験体からの核酸を含むことがある。場合によっては、本明細書の方法とシステムは、サンプル中に存在する個々の被験体を特定するために使用される。例えば、被験体 A と被験体 B を含むサンプルを例に取る。サンプルは、被験体 A と被験体 B の両方に由来する核酸を含み得る。核酸は、被験体 A と被験体 B を、あるいはその逆を区別する少なくとも 1 つの標的核酸配列を含み得る。本明細書の方法とシステムは少なくとも 1 つの標的核酸配列を特定するために使用可能である。この情報は複合サンプルが被験体 A と被験体 B の両方を含んでいたと断定するために使用可能である。

10

【 0 0 3 9 】

標的核酸

標的核酸配列は、1 つの被験体を別の被験体と識別するか、あるいは抗生物質耐性または病原性などの標的の属性を区別するあらゆる核酸配列であり得る。場合によっては、複合サンプル中に存在する 1 つを超える被験体は、実質的に同一であり、かつ標準的なマイクロアレイ技術を駆使しても解決するのが難しいこともあるゲノムを有する。場合によっては、1 つを超える被験体は、9 0 % 、 9 1 % 、 9 2 % 、 9 3 % 、 9 4 % 、 9 5 % 、 9 6 % 、 9 7 % 、 9 8 % 、 9 9 % 、 9 9 . 1 % 、 9 9 . 2 % 、 9 9 . 3 % 、 9 9 . 4 % 、 9 9 . 5 % 、 9 9 . 6 % 、 9 9 . 7 % 、 9 9 . 8 % 、 9 9 . 9 % 、 9 9 . 9 9 % 、 9 9 . 9 9 9 % 同一であるゲノムを有する。場合によっては、1 つを超える被験体は、微生物の 1 つ以上の異なる菌株、例えば、細菌、ウイルス、真菌などの 1 つ以上の菌株である。

20

【 0 0 4 0 】

標的核酸配列は 1 つを超える遺伝的特徴を含み得る。1 つを超える遺伝的特徴は、1 つの被験体を別の被験体と区別することができる。遺伝的特徴は、ゲノム、遺伝子型、ハプロタイプ、染色質、染色体、染色体座、染色体材料、対立遺伝子、遺伝子、遺伝子群、遺伝子座、遺伝的多型、遺伝子突然変異、一塩基多型 (S N P) 、制限断片長多型 (R F L P) 、可変タンデムリピート (V T R) 、コピー数多型 (C N V) 、マイクロサテライト配列、遺伝子マーカー、配列マーカー、配列でタグ付けされた部位 (S T S) 、プラスミド、転写単位、転写生成物、遺伝子発現レベル、遺伝子発現状態を含み得る。標的核酸配列は任意の既知の遺伝的特徴も本質的に含むことができる。

30

【 0 0 4 1 】

標的核酸はデオキシリボ核酸 (D N A) あるいはリボ核酸 (R N A) を含み得る。 D N A はゲノム D N A または c D N A であり得る。 c D N A は当該技術分野の当業者に知られているような R N A の逆転写によって生成可能である。標的核酸は一本鎖または二本鎖であり得る。場合によっては、標的核酸は修飾可能である。核酸修飾は当該技術分野で知られているものを含むことができ、標的核酸は本質的にどんな修飾も含み得る。有用な修飾は、限定されないが、ビオチンまたはジゴキシゲニンのようなアンカーリガンドと同様に、放射標識と蛍光標識を含む。修飾は標的の内部に、あるいは 5 ' または 3 ' のいずれかの末端に置くことができる。標的修飾は、ライゲーションあるいはポリメラーゼによる伸長などの化学反応または酵素反応のいずれかによって、合成後に行うことができる。

40

50

【0042】

標的核酸の長さは変わることがある。標的核酸は、数十から数百、または数千もの塩基対、あるいは数万または数十万の塩基対のサイズで変動し得る。いくつかの例において、標的核酸は、約10、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、1000、1500、2000、2500、3000、3500、4000、4500、5000、5500、6000、6500、7000、7500、8000、8500、9000、9500、あるいは10000以上の塩基対の長さである。いくつかの例において、標的核酸は、少なくとも約10、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、1000、1500、2000、2500、3000、3500、4000、4500、5000、5500、6000、6500、7000、7500、8000、8500、9000、9500、あるいは10000以上の塩基対の長さである。いくつかの例において、標的核酸は、せいぜい約10、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、1000、1500、2000、2500、3000、3500、4000、4500、5000、5500、6000、6500、7000、7500、8000、8500、9000、9500、あるいは10000の塩基対長さである。

10

【0043】

バイオチップへの適用の前に、標的核酸は任意の数のサンプル調製工程を経験することがある。こうした工程は、当業者に知られている任意の数の断片化、增幅、修飾、あるいは調製の工程を含み得る。

20

【0044】

標的核酸は化学溶解、超音波処理、均質化などを含む任意の技術によって生体サンプルから放出可能である。標的核酸は、さらなる処理工程の前に、当該技術分野で知られている（例えば、細胞残屑、汚染物質、あるいは他の材料を除去するために）任意の数の精製工程を経験することがある。

【0045】

場合によっては、標的核酸は、バイオチップへの適用前に標識可能である。場合によっては、標的核酸はバイオチップへの適用後に標識可能である。標的核酸は多重標識で標識可能である。核酸ラベルは核酸の検出を可能にする任意のタグであり得る。放射標識、フルオロフォア、染料、ビオチン、酵素（例えば、ワサビペルオキシダーゼ（H R P）、アルカリフェオヌクレオターゼ（A P）などを含む任意の数の標識を用いることができる。標的核酸は5'末端、3'末端、あるいはその両方で標識可能である。場合によっては、標的核酸に身体標識される。末端デオキシヌクレオチジルトランスフェラーゼ（T d T）、T 4 RNAリガーゼ、T 4 ポリヌクレオチドキナーゼ（P N K）、DNAポリメラーゼ、RNAポリメラーゼなどの酵素的技術；あるいは、過ヨウ素酸酸化、5'リン酸塩の1-エチル-3-[3-ジメチルアミノプロピル]カルボジイミド塩酸塩（E D C）活性化、あるいは化学的な無作為標識（例えば、光反応性標識システム、K r e a t e c h D i a g n o s t i c s から市販されているU n i v e r s a l L i n k a g e S y s t e m）などの化学的技術を含む、核酸を標識する任意の方法が使用可能である。場合によっては、いかなる標識も必要とされず、標的の結合は、陽子の放出、表面上の化学組成の変化、光路の屈折率の変化、あるいは、ハイブリダイゼーション事象の直接的な電気的な検出を介して検知可能である。

30

【0046】

標的核酸は染料またはステインで標識可能である。核酸を標識するのに適切な染料は当該技術分野で知られているものを含み得る。染料は蛍光染料であり得る。場合によっては、緩衝液はC y 3である。場合によっては、緩衝液はC y 5である。

40

【0047】

標的核酸は5'末端、3'末端で標識可能であるか、あるいは身体標識可能である。どの

50

方法を使用するべきかに関する決定は、必要とされる標識化の程度と、標識が立体障害を引き起こしてプローブとの相互作用を防ぐことがあり得るかに部分的に依存することがある。

【0048】

場合によっては、核酸標識は核酸分子の全体にわたって無作為に組込まれる（つまり、身体標識される）。様々な方法を用いて標的核酸に身体標識することができる。身体標識プロトコルは、標識されたヌクレオチドを標的核酸に組み込むための酵素の使用を含み得る。場合によっては、身体標識された核酸は標準的なポリメラーゼ連鎖反応（P C R）方法で生成される。この方法は2つの目的に役立つことができる：1）成長している核酸鎖へ標識されたヌクレオチドを無作為に取り込むこと；2）鑄型核酸を増幅させること。この方法は、標的に特異的なプライマーまたは無作為のプライマーの使用を含み得る。場合によっては、標的核酸はバイオチップへの適用前にP C Rによって増幅される。

10

【0049】

場合によっては、標識されたヌクレオチドは無作為のプライマー伸長によって無作為に組み込まれる。この例において、複数の無作為のプライマー（例えば、ヘキサヌクレオチド）を用いて、一本鎖D N A鑄型上で第1のD N A合成を無作為に刺激する。標識されたヌクレオチドのD N A合成と無作為な取り込みは、D N AポリメラーゼIあるいはD N AポリメラーゼIのクレノウ断片の使用を含み得る。場合によっては、標識は、例えば、二本鎖D N A標識プロトコルを使用することにより、標的核酸のプローブへのハイブリダイゼーションの後に生じることがある。

20

【0050】

他の場合では、標識されたヌクレオチドはローリングサークル増幅によって無作為に組み込まれる。この方法は、標的核酸分子が環状である（例えば、プラスミド、バクテリオファージの環状ゲノム、ウイロイドの環状R N Aゲノムなど）ときに、とりわけ良く適していることがある。ローリングサークル増幅では、ニックは、不連続鎖と連続鎖を形成する環状の核酸分子の一本鎖で生成される。環状のベクターの連続的な鎖は等温の増幅反応を用いて増幅される。場合によっては、ローリングサークル増幅は、高い鎖置換活性を呈する29D N Aポリメラーゼを使用する。

【0051】

場合によっては、標的核酸はバイオチップへのハイブリダイゼーション前には増幅されない。

30

【0052】

場合によっては、標的核酸はバイオチップへの適用前にせん断されるか、または断片化される。せん断方法は当該技術分野で知られており、超音波処理、針せん断、フレンチプレッシャーセル（French pressure cell）の通過、ポイント-シンク（point-sink）せん断、音響せん断、制限消化、断片化酵素、あるいはトランスポソーム媒介性の断片化を含み得る。場合によっては、標的核酸はせん断または断片化の前に標識される。この方法は、標識法が例えば、ローリングサークル増幅を含む場合、適切となることがある。他の場合では、標的核酸のせん断は標識の前に生じる。

【0053】

トランスポソームを媒介とする断片化を利用して、同時に、断片を生成し、かつ検出のためにその断片を標識することができる。トランスポーザ（Tn5など）は、合成D N A配列を切断して、これを他のD N A分子の5'末端へ共有結合することができる。例えば、5'標識されたフルオロフォアを合成D N Aに結合することにより、D N Aを同時に断片化して標識することができる。変性後のこの断片化および標識された（例えば、蛍光標識された）D N Aは、アレイへのハイブリダイゼーションの準備を整わせることができる。例えば、図6は、標識（603）（例えば、蛍光標識）を伴うオリゴヌクレオチドを含むトランスポソーム複合体（602）と相互作用する二本鎖D N A（d s D N A）（601）の典型的な概略図を示す。インキュベーション（604）の後、標識（605）を備えた断片化したd s D N Aは、トランスポソーム複合体によって生成される。その後、

40

50

変性(606)を用いて、ハイブリダイゼーション用の標識された一本鎖DNA(ssDNA)(607)を生成することができる。トランスポソーム媒介性の断片化を利用して、標識とトランスポソームからの合成DNAの小片も含有する、二本鎖または一本鎖(例えば、変性後)の断片化したDNAを生成することができる。こうした技術は、2工程プロセスの断片化と標識と比較して、収率と効率を増加させることができる。

【0054】

核酸などのサンプル材料は濃縮および/または精製可能である。このことはサンプル材料の解析を助けることがある。例えば、核酸を通過させないが、イオン、タンパク質、および他の細胞残屑を通過させる膜(例えば、ジエチルアミノエチル(DEAE)セルロース紙などの分子量遮断膜(molecular weight cutoff membrane))を用いて、核酸を濃縮することができる。この膜上でオリゴヌクレオチド捕捉アレイをスポットして固定することで、捕捉プローブの近くの標的の濃度を増加させて標的を捕捉プローブに持ち込み、アレイへのハイブリダイゼーションを加速させることができる。電場または流量の方向は一時的に反対にするか、あるいはパルス化することで、表面に平行な面の流れを助長して、ハイブリダイゼーション率を改善し、ハイブリダイゼーション時間を短くすることができる。例えば、図8は、フィルター基質(801)上のオリゴヌクレオチドアレイに近づけたサンプルDNA(800)の例を示す。電場が適用されるとき(負(802)正(803))、サンプルDNAは電気泳動にさらされ、アレイに集まる。代替的に、電流を(例えば、単純な分子量遮断フィルター膜を用いて)簡単に逆にすると、核酸を膜から溶液へ移動させることができ、これを用いて、例えば、シリカ、プラスチック、ガラス、あるいは別の基質上に固定されたアレイにハイブリダイズすることができる。電気泳動運動の代わりに、あるいは電気泳動運動に加えて、膜に対して核酸を移動させるために流量を適用することにより濃縮を行うこともできる。

10

20

30

【0055】

酸化的損傷の遊離ラジカルあるいは他の供給源の核酸への濃縮が減少するように、電極(例えば、図8の(802)と(803))を互いに一定の間隔で配置することができる。この種の設計は、核酸が、例えば濃縮工程の間に経験する酸化的損傷の量を減らすことができる。

【0056】

拡散に加えて様々な手段によって、核酸をバイオチップあるいは他のアレイ表面に近接させることができる。上に議論されるように、電気泳動および/または流量を用いて、アレイ表面あるいはアレイ表面の近くに核酸を濃縮することができる。他の手法も使用できる。例えば、アレイ表面は、その表面の全体または一部にわたって(例えばプローブ特徴における)疎水性の界面化学を備えることができ、標的核酸は疎水性の部分でタグ付けされ、核酸は表面の疎水性領域をエネルギー的に優先するようになる。別の例において、標的核酸は磁粉でタグ付け可能であり、磁場を用いてアレイ表面へ標的核酸を近づけることができる。

40

【0057】

体積排除(volume-excluding)化合物を用いても、サンプルDNAなどのサンプル材料を有効に濃縮することができる。体積排除剤(volume excluding reagent)は、占有された液体体積からサンプル材料を除外するために使用可能であり、それによって、残りの液体体積にサンプル材料を濃縮することができる。このメカニズムは、基質に対するサンプル核酸のハイブリダイゼーションなどのサンプル材料の捕捉あるいは結合を加速するのを助けることができる。例えば、体積排除剤はハイブリダイゼーション動力学を改善するためにハイブリダイゼーション緩衝剤に含めることができる。体積排除剤は、限定されないが、デキストラン硫酸塩、フィコール、およびポリエチレングリコールを含む、例えば、ビーズまたはポリマーであり得る。体積排除剤は高分子量ポリマーであり得る。体積排除剤は、例えば、核酸の体積排除剤への結合を減らすために負に帯電させることができる。

【0058】

50

プローブ

本明細書で開示されたバイオチップは、表面に分布する複数のプローブを有する。場合によっては、複数のプローブはバイオチップの表面に固定される。場合によっては、表面は固形物である。他の場合では、表面は半固形物である。場合によっては、表面はガラスまたはシリコンである。複数のプローブは界面化学を用いて表面へ固定可能である。

【 0 0 5 9 】

1つの非限定的な例において、複数のプローブは、1 - エチル - 3 - [3 - ジメチルアミノプロピル] カルボジイミド塩酸塩 (E D C) 化学を用いて、バイオチップの表面へ固定される。この方法では、バイオチップ上のカルボキシル基は E D C で活性化される。活性化されたカルボキシル基は第一級アミン基と反応して、安定したアミド結合を形成することができる。この例において、バイオチップはビーズ、場合によっては、シリカあるいはガラスビーズであり得る。複数のプローブは、5' あるいは 3' の末端でアミノ修飾される。アミノ修飾の非限定的な例としては、5' - アミノアリル - d U T P 、5 - プロパルギルアミノ - d C T P 、N⁶ - 6 - アミノヘキシル - d A T P 、および 7 - D e a z a - 7 - プロパルギルアミノ - d A T P を含む 1 つ以上のアミノ修飾されたヌクレオチドが挙げられる。プローブは、この方法を使用して、5' または 3' 末端のいずれかでバイオチップに固定可能である。場合によっては、2 工程方法が利用される：1) E D C による活性化と、その後の、2) 効率を改善するか、あるいは乾燥して安定している（アミン反応性）中間体を作製するための、N - ヒドロキシコハク酸イミド (N H S) による処置。いくつかの例では、2 工程の E D C 処置は固定効率を改善するために使用される。この例において、E D C の第 1 の濃度をバイオチップに適用可能であり、その後に、E D C の後の第 2 の濃度を適用可能である。場合によっては、E D C の第 1 の濃度は E D C の第 2 の濃度よりも低い。場合によっては、この 2 工程の E D C 処置はプローブ固定の効率を改善する。

【 0 0 6 0 】

バイオチップは複数のプローブを含みうる。複数のプローブは、バイオチップの表面上に分布されうる。バイオチップは、単一の固体として物理的に結合される必要のない複数の表面を含みうる。プローブを含む表面は例えばビーズまたは一連のビーズでありうる。ビーズは同一でありうる。ビーズはマイクロビーズでありうる。ビーズは個別に溶解可能でありうる。ビーズはビーズに特異的なバーコードを含みうる。ビーズはビーズに特異的な標識を含みうる。ビーズはビーズに特異的な結合部位を含みうる。

【 0 0 6 1 】

マイクロアレイの活性領域またはハイブリダイゼーション領域を縮小することにより、マイクロアレイの経済問題、および必要とされるハイブリダイゼーション時間を改善することができる。この場合、小さなマイクロアレイを製作することでコストを省き、かつそのより小さな活性領域によって、より濃縮されたサンプルを使用できるようになり、ハイブリダイゼーション活性を促進せらるだろう。

【 0 0 6 2 】

プローブは、核酸サンプル中に存在する標的配列を持つワトソン - クリック型塩基対の形成を可能とするオリゴヌクレオチドでありうる。プローブの長さは変動しうる。いくつかの例では、遺伝的特徴内のプローブは、長さが 20 %、10 %、5 %、または 1 % 未満変動する。いくつかの例では、プローブは同じ長さである。プローブは、数十から数百もしくは数千の塩基対、または、数万または数十万の塩基対のサイズにさえ変動しうる。場合によっては、プローブは、約 20 塩基長、約 25 塩基長、約 30 塩基長、約 35 塩基長、約 40 塩基長、約 45 塩基長、約 50 塩基長、約 55 塩基長、約 60 塩基長、約 65 塩基長、約 70 塩基長、約 75 塩基長、約 80 塩基長、約 85 塩基長、約 90 塩基長、約 95 塩基長、約 100 塩基長、約 110 塩基長、約 120 塩基長、約 130 塩基長、約 140 塩基長、約 150 塩基長、約 200 塩基長、約 250 塩基長、約 300 塩基長、約 350 塩基長、約 400 塩基長、約 450 塩基長、約 500 塩基長、約 600 塩基長、約 700 塩基長、約 800 塩基長、約 900 塩基長、約 1000 塩基長、または 1000 塩基長以上である。

10

20

30

40

50

【0063】

プローブは、バイオチップの表面上に、かつ被験体に特異的な特徴へと分布しうる。被験体に特異的な特徴は複数のプローブを含みうる。いくつかの例では、被験体に特異的な特徴は、10、100、1,000、10,000、または100,000を超える個々のプローブを含む。被験体に特異的な特徴は複数の同一のプローブを含みうる。他の例では、被験体に特異的な特徴は複数のプールされた同一でないプローブを含みうる。同一でないプローブは、様々な領域における標的核酸に結合しうる。プローブは非重複領域における標的に結合しうる。場合によっては、同一でないプローブは重複配列を有する。被験体に特異的な特徴は少なくとも10、100、1,000、10,000、または100,000以上の同一でないプローブを含みうる。いくつかの例では、バイオチップは10、100、1,000、10,000、100,000、1,000,000、10,000,000、または1,000,000,000を超える個々の被験体に特異的な特徴を含む。

【0064】

被験体に特異的な特徴は、分離したスポットまたはクラスタなどの中に、個々に対処可能（例えば、検知のために個々に対処可能）であるように、そのような方法でバイオチップ上に分布されうる。被験体に特異的な特徴に対応する複数のプローブは、プローブの1つ以上のセットへと配置されうる。プローブの各セット内では、複数のプローブは同一である可能性があり、または互いとは異なる可能性もある。各セット内では、複数のプローブはそれぞれ被験体に特異的な特徴を含みうる。各セット内の複数のプローブは、別のものから1つの被験体を判別する1つ以上の被験体に特異的な特徴を含みうる。場合によっては、被験体に特異的な特徴は、円形、正方形、または長方形の領域などのアレイ上のスポットまたは領域でありうる。場合によっては、被験体に特異的な特徴はビーズでありうる。場合によっては、被験体に特異的な特徴は特徴特異的タグによって標識された一連のプローブでありうる。特徴特異的タグは、例えば、特徴特異的バーコード、または特徴特異的標識のための結合部位でありうる。いくつかの例では、特徴は複製特徴を有する。いくつかの例では、複製特徴は同一である。いくつかの例では、複製特徴は、同じ標的ポリヌクレオチドを特定するように設計される。いくつかの例では、複製特徴は、同じゲノムを特定するように設計される。いくつかの例では、複製特徴は、種内のいかなる株も特定するように設計される。場合によっては、複製特徴は、個体を特定するように設計される。

【0065】

図1のA - Fは、典型的なバイオチップシステムを例示する。図1のAは、同一のプローブを含む各特徴を有する4つの典型的な特徴、およびサンプル中の単一の被験体からの4つの未結合の標識化された標的を図示する。したがって、プローブに対して標的を結合することにより、各特徴からの1単位のシグナルを結果としてもたらすだろう。図1のBは、サンプル中の単一の被験体からの4つのプローブおよび4つの未結合の標識化標的を備えた、被験体に特異的な特徴を図示する。したがって、プローブに対して標的を結合させることにより、結果として1つの特徴から4単位のシグナルをもたらすだろう。図1のCは、各特徴からの1単位のシグナルを結果としてもたらす、4つの異なる特徴に結合されている4つの標的を図示する。図1のDは、1つの特徴からの4単位のシグナルを結果としてもたらす、単一の被験体に特異的な特徴に結合されている4つの標的を図示する。図1のCと図1のDとの比較は、多数の被験体標的を対象とする複数の異なるプローブを使用するときに单一の特徴に存在しうるシグナル増幅を実証する。図1のEは、図1のAおよび図1のCにおいて示されるもののようなアレイ性能を結果として生じる、別個の特徴における固有のプローブの順序付けられたプールによる特徴を図示する。図1のFは、図1のBおよび図1のDにおいて示されるもののような配置を結果として生じる、特徴中の固有のプローブをランダムにプールする特徴を図示する。

【0066】

プローブセットまたは被験体に特異的な特徴内の多数の固有のプローブは、検出システムの解像度よりも小さい、または同等の大きさの領域内に位置しうる。多数の、固有な、

10

20

30

40

50

順序付けられたプローブによって囲まれた領域は、検出システムの解像度よりも小さくなリうる、もしくは検出システムの解像度と同等になりうる、または全ての固有のプローブによって囲まれた領域は、セット中のランダムに順序付けられた固有のプローブの少なくとも2つによって囲まれた領域が検出システムの解像度とほぼ同等である、もしくはそれ未満である限り、より大きくなりうる。この場合、多数の固有のプローブまたは特徴からのシグナルは、1または少數のピクセル、または他の解像度要素において収集されるか統合されうる。このような取り組みは、単一の特徴へとプールする同一でないプローブに類似した結果を達成することができる。

【0067】

参考のために、マイクロアレイ光学的検出で使用される画像化システムは、ピクセルまたは解像度要素当たり、1マイクロメーター(μm)から $5\mu\text{m}$ の解像度を有しうる。典型的には、直径または長さが $5\mu\text{m}$ である光学的に検出されたマイクロアレイ特徴は、1 μm から $5\mu\text{m}$ の光学的解像度が可能な光学システムを用いて画像化されるだろう。別の例は、ビーズの中心から中心までの間隔が $2\mu\text{m}$ 離れた $1\mu\text{m}$ 径のビーズで構成されるマイクロアレイであるだろう。このアレイは、 $0.5\mu\text{m}$ から $1\mu\text{m}$ の解像度光学システムをもちいて画像化されてもよい。

10

【0068】

一実施例では、個々の固有のプローブ(例えばDNA断片)はプローブセットとして基板に堆積させる前にプールされ、個々の固有のプローブは、固有のプローブとの間の平均距離、例えばプローブ中心からプローブ中心まで 10nm から数 10nm 程度で基板に対して付着しうる。このプローブセットを含む単一の特徴の大きさは、例えば直径が約 $1\mu\text{m}$ 、 $2\mu\text{m}$ 、 $3\mu\text{m}$ 、 $4\mu\text{m}$ または $5\mu\text{m}$ でありうる。その後、例えば約 $1\mu\text{m}$ の解像力を持つ画像化システムは、1ピクセルもしくは最大 25ピクセル 、または他の解像度要素内の多数の個々のプローブからのシグナルを収集する、または統合することができる。

20

【0069】

いくつかの設計では、特徴は情報を含まない特徴との間に間隔を開けて配置されるか、またはその特徴はシグナルがない特徴との間にいかなる領域もないよう、特徴の市松模様などの境界に接触して順序付けられうる。例えば、他の特徴に直接隣接して配置される $5\mu\text{m}$ 正方形の特徴の場合には、特徴を区別するために $1\mu\text{m}$ または $2\mu\text{m}$ の解像度が必要とされることもある。一方、 $5\mu\text{m}$ のビーズが、他の $5\mu\text{m}$ のビーズとともに、中心から中心までの距離が $15\mu\text{m}$ の間隔で配置された場合、その後、 $5\mu\text{m}$ 、または場合によっては $10\mu\text{m}$ の解像度の画像化システムが、マイクロアレイ上の異なる特徴とシグナルとを区別するのに適切でありうる。

30

【0070】

別の例では、多数の同一のプローブは、画像化システムの解像度よりも小さな大きさの第1の特徴において共にグループ化され、そしてこの特徴は、特徴内のプローブが同一であるが、第1の特徴中のプローブとは異なる、同じ被験体を標的とする他の特徴に隣接して位置決めされる。全てのプローブが検出システムの解像力によって画定されたおおよその領域以下の領域内において囲まれる場合、その後、検出器は、すべてのプローブからのシグナルを、単一のピクセルまたは解像度要素へと統合することができる。画像化システムの解像度近くまたはそれ未満の大きさの領域中の同一のプローブ型の群をプールすることによって、プローブのランダムにプールされたセットと同じ恩恵を受けることができる。

40

【0071】

バイオチップは多数の被験体に特異的な特徴を含みうる。いくつかの例では、バイオチップは 10 、 100 、 $1,000$ 、 $10,000$ 、または $100,000$ を超える被験体に特異的な特徴を含む。場合によっては、多数の被験体に特異的な特徴はプローブの多数のセットへと配置され、ここでプローブの各セットは異なる被験体を識別する。

【0072】

プローブは標的を結合することができる可能性がある。プローブは標的に相補的になり

50

うる。プローブは標的に対する親和性を有しうる。プローブは3つ全ての組み合わせでありうる。異なる被験体において対象とされた特徴は異なるプローブを含みうる。いくつかの例では、非複製特徴は別の特徴を備えるプローブも共有しない。いくつかの例では、非複製特徴は、別の特徴を備える0.1%、1%、5%、または10%未満のそのプローブを共有する。

【0073】

場合によっては、プローブの各セットは、固有のプローブ型の平均的な表現を有する。場合によっては、固有のプローブ型の平均的な表現は、約2、3、4、5、6、7、8、9、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、100、200、300、400、500、600、700、800、900、1000またはそれ以上である。

10

【0074】

セット内のプローブ断片の総数、およびセット内の单一で固有のプローブ型の平均的な表現は、プローブセットの特異性およびプローブセットのダイナミックレンジを制御するために制御されうる。一実施例では、プローブの総数がおよそ1000に限定され、かつセット中の固有のプローブ型の数が250である場合、そのとき、個々のプローブ型の平均的な表現は約4であろう。第1の例では、4つの被験体細胞のDNAがサンプル内に存在し、4つの被験体細胞からのゲノムDNAの約90%がサンプル調製プロセス後にサンプル内にとどまり、実質的にアレイに結合する場合、シグナル強度は最大強度の約95%になるだろう。同様に、2つの被験体細胞からのDNAが存在し、処理され、95%の効率でハイブリダイズされた場合、シグナル強度は最大強度の約50%でありうる。第2の例では、被験体細胞がサンプル中に存在しないが、しかしながら、非被験体有機体が1000を超える細胞のサンプル中に非常に豊富であり、非被験体有機体からのDNAは单一のプローブ型とマッチする単一の領域を含む場合、単一のマッチしたプローブ型から生成されたシグナルは、平均して最大シグナルの約1%未満である。

20

【0075】

これに対して、上記の同じ2つの例を考慮すると、固有で個々のプローブ型の数が10にセットされ、プローブ断片の総数が10,000である場合は、平均プローブ表現は約1000だろう。このシナリオでは、上記例Aを考慮すると、シグナル強度は最大シグナルの1%未満となり、場合によっては偽陰性判定を結果としてもたらすだろう。同じシナリオであるが、上記例Bを考慮すると、シグナル強度は最大シグナルの約10%未満となり、起こりうる偽陽性判定を結果としてもたらすだろう。検出されうる標的細胞からの検出の下限は、その後システムの特異性、すなわちセット内のプローブの1つ以上における結合エラーが原因である偽陽性コールを拒否する能力を増加させるように折衷されうる。

30

【0076】

いくつかの態様では、被験体に特異的な特徴中のプローブの総数は制御される。被験体に特異的な特徴中のプローブの総数を制御する方法は、限定されることなく、合計特徴サイズを制御する工程と、被験体に特異的な特徴内の個々のプローブ間隔を制御する工程と、を含む。

40

【0077】

いくつかの態様では、プローブは、偽陽性を検出するように設計されうる。例えば、プローブは、プローブセットを設計することにより設計することができ、ここで、プローブセット内の個々のプローブは、他のプローブセット中の個々のプローブにミスマッチされた1つ以上の塩基を考慮して設計されている。場合によっては、プローブセットは相補的である。その他の場合において、プローブセットは相補的ではない。別の例では、プローブセットは、多数の類似した株を有する被験体有機体を探すように設計されうる。この例において、標的でないが、1つ以上の個々の固有の特性を備えた標的のゲノムに非常に近いゲノムを有する株内に含まれる個々の配列を検出するために、プローブセットが添加されうる。

50

【0078】

プローブセットの設計は、起こり得る偽陽性の挙動についてより多く学ぶために、経験的スクリーニング工程も含みうる。この工程は、自然界で発見されうる、かつハイブリダイゼーション活性において含まれる非被験体材料の一部でありうる、様々な型のゲノムDNAに対する対照を添加する工程に基づきうる。これらのプローブセットは、別個にまたはセットとしてスクリーニングされうる。

【0079】

バイオチップ上の特徴（例えばアレイスポット）は、それぞれ異なることもあり、限定されることなく、異なる数のプローブ、異なる型のプローブ、異なる被験体に特異的な特徴、固有のプローブの異なる平均的な表現、などを含む。非限定的な一実施例では、バイオチップは多数の被験体に特異的な特徴のアレイを含みうる。各被験体に特異的な特徴は、プローブのセットを含むことができ、各セットが複数の固有のプローブを含む。いくつかの特徴（またはスポット）では、プローブのセットは、アレイの感度を測定するために使用することができる、標的被験体のための固有のプローブの低い平均的な表現を含みうる。同じアレイ上の他の特徴では、プローブのセットは、サンプル中の標的核酸の存在量を測定するために使用することができる、同じ標的被験体のための固有のプローブの高い平均的な表現を含みうる。特徴の感度および／または存在量の任意の数は、バイオチップ上で設計されうる。

10

【0080】

そのバイオチップをもちいて、（例えば2つの異なるマイクロバイオームのサンプルからの）異なるソースのハイブリダイゼーション強度を連続的に比較するために、多数の調査実験をすることができる。例えば、バイオチップは多数の連続的なハイブリダイゼーション反応、および随意に多数の読み取り反応（read reaction）を受けることができ、各反応は、異なるプローブセットの特性に対して最適化されている。

20

【0081】

プローブは、ポリヌクレオチド（または標的）を標的とするようにハイブリダイズするために、設計されうる。標的ポリヌクレオチドはゲノム配列でありうる。標的ポリヌクレオチドはゲノムの非コード領域でありうる。標的是ゲノム特徴でありうる。標的はミトコンドリアDNAでありうる。標的はプラスミドからのDNAでありうる。標的は変異体でありうる。標的は、ゲノムの保存領域、または病原性に関連する領域でありうる。

30

【0082】

標的は、別のゲノムからの判別可能なゲノムの領域でありうる。標的は、被験体の集団内の被験体に特有のポリヌクレオチドでありうる。標的は、そのマイクロバイオーム内に表現された特定の種とは別個の、マイクロバイオーム内のポリヌクレオチドでありうる。

【0083】

プローブ設計のための典型的なプロトコルは以下のとおりである。第一に、長さ基準（例えば35塩基）をプローブセットに対して選択する。第二に、選択された長さのk-merのセットは、被験体のゲノムを通じて連続して進むことにより、標的ゲノムから作り出される。第三に、k-merを、同じ種類の他のゲノムと比較する（例えばBLASTによって）。第四に、ヒト、細菌、ウイルス等などの公的に入手可能な全ての他のゲノム（すなわち同じ種類ではない）と、k-merを比較する（例えばBLASTによって）。場合によっては、第三および第四の工程は、いくらかの種類（例えばE. coli）に対しては非常に小さなセットしか結果としてもたらすことができないが、ともに実行することができる。第五に、固有の候補のk-merの一覧表を作る。さらに、これらの固有のプローブの中央の塩基を各直交塩基（orthogonal base）に変更することができ（1つのk-merが3つのミスマッチk-merを結果としてもたらす）、ミスマッチk-merを公的に入手可能な全ての他のゲノムと（例えばBLASTによって）、および／または同じ種類の他のゲノムと比較することができる。第六に、自己相補性（self complementarity）（すなわち、プローブがそれ自体に結合するかどうか）に関して候補をテストする。第七に、融解温度を一本鎖の自由エネルギーに基

40

50

づいて評価する。第八に、*k* - m e r を、固有性の値（例えば固有の配列のパーセント）に基づいてランク付ける。さらに、*k* - m e r は、G C 含量などの他の特徴に関してフィルタリングされうる。例えば、場合によっては、< 60 % の G C 含量を有する *k* - m e r のみが含まれる。第九に、候補を、同じ種類の選択されたゲノムに対してハイブリダイズすることにより経験に基づいてテストする。第十に、これらの結果に基づいて、最終候補のプールを選択する。

【0084】

プローブは、既知の核酸配列に対して相補的になるように設計することができる。場合によっては、被験体のポリヌクレオチドは、ポリヌクレオチドの配列を判定するために、プローブを設計する前に配列することができる。一旦ポリヌクレオチドが配列されたならば、プローブは被験体のポリヌクレオチドを標的とするように設計することができる。場合によっては、被験体のポリヌクレオチドは、被験体のゲノム中で発見される配列を含む。この例では、被験体のゲノムは配列され、プローブはゲノム内の標的ポリヌクレオチドに対して設計されうる。いくつかの例では、標的のリストは、2人以上の被験体間の非重複ゲノム領域を判定することにより生成されうる。場合によっては、標的はアセンブリを比較することにより識別される。配列決定法は、キャピラリー配列決定、次世代配列決定、サンガー配列決定、合成による配列決定、単一分子ナノポア配列決定、ライゲーションによる配列決定、ハイブリダイゼーションによる配列決定、ナノポアの流動を制限することによる配列決定 (sequencing by nanopore current restriction)、またはそれらの組み合わせを含みうる。合成による配列決定は、可逆的な転写終結配列決定、前進的な単一分子配列決定、連続的なヌクレオチドフロー配列決定 (sequential nucleotide flow sequencing)、またはそれらの組み合わせを含みうる。連続的なヌクレオチドフロー配列決定は、パイロシークエンシング、pH媒介配列決定)、半導体配列決定、またはそれらの組み合わせを含みうる。1つ以上の配列決定反応の実施は、非標的配列決定 (untargeted sequencing) (例えばゲノム全体の配列決定)、または標的配列決定 (targeted sequencing) (例えばエクソーム配列決定) を含む。

【0085】

配列決定法は、マクサム - ギルバート法 (Maxim-Gilbert)、連鎖停止反応、またはハイスループットシステムを含みうる。代替的にまたは加えて、配列決定法は、Helioscope (商標) 単一分子配列決定、ナノポアDNA配列決定、Lynx Therapeuticsの大規模並列シグネチャー配列決定 (MPSS)、454パイロシークエンシング、単一分子リアルタイム (RNA P) 配列決定、Illumina (Solexa) 配列決定、SOLID配列決定、Ion Torrent (商標)、イオン半導体配列決定、単一分子SMRT (商標) 配列決定、ポロニー配列決定、DNAナノポール配列決定、Visigen Biotechnologiesアプローチ、またはそれらの組み合わせを含みうる。代替的にまたは加えて、配列決定法は、限定されないが、Illuminaによって提供されたGenome Analyzer IIx, HiSeq, NextSeq、およびMiSeqと、Pacific Biosciences (California) およびSolexa Sequencerによって提供された、Pacific BiosystemsなどのSingle Molecule Real Time (SMRT (商標)) の技術と、Helicos Inc. (Cambridge, MA) によって提供されたHelicoscope (商標) シークエンサーなどのTrue Single Molecule Sequencing (tSMS (商標)) の技術と、Genia Technologies, Inc. によって開発されたナノポアベースの配列決定プラットフォームと、Oxford Nanopore MinIONと、を含む1つ以上の配列決定プラットフォームを含む。

【0086】

場合によっては、配列または遺伝子発現のデータベースをクエリして、標的対象の既知の核酸配列を識別する。配列または遺伝子発現のデータベースの非限定的な例は、NCB

10

20

30

40

50

IのGenBank、the European Molecular Biology Laboratory (EMBL)、the DNA DataBank of Japan (DDBJ)、ENSEMBL、the Ashbya Genome Database (AGD)、BioCyc、CleanEx、CYGD、Dictybase、EchoBase、EcoGene、euHCVdb、Evotrace、FlyBase、GeneCards、GeneDB、GeneFarm、GenoList、Gramene、HGNC、HInv-DB、HOGENOM、KEGG、MaizeGDB、MEROPS、MGD、NMPDR、NCBI Nucleotide db、NCBI RefSeq、PANTHER、PCCDB、Peroxibase、Pfam、PhosphositePlus、PlasmoDB、PptaseDB、Pseudocap、RGD、SGD、TAIR、TIGR/SCVI、UniGene、VectorBase、WormBase、およびZ-FINを含む。

【0087】

<方法>

いくつかの態様において、本明細書に記載された方法は、複数の異なる被験体を含むサンプルを提供する工程を含む。場合によっては、サンプルは、複数の異なる被験体に由来する複数の核酸を含む。場合によっては、複数の核酸は、複数の様々な被験体の少なくとも2つ以上から得られる少なくとも1つの標的核酸を含む。

【0088】

いくつかの態様において、方法は、複数の被験体から核酸を抽出する工程をさらに含む。いくつかの態様において、方法はまた、複数の被験体から抽出された核酸を断片化する工程を含む。バイオチップへの適用に先立って、追加の処理工程を核酸に対して行うことができる。場合によっては、核酸を、バイオチップへのハイブリダイゼーションに先立つて修正できる。例えば、本明細書に記載される通り、核酸を標識することができる。さらには代替的に、例えば、捕捉プローブの使用、または增幅工程によって、標的核酸を富化することができる。他の例では、非標的核酸は、ハイブリダイゼーションの前にサンプルから枯渇させることができる。

【0089】

さらなる態様では、方法はバイオチップへ複数の核酸をハイブリダイズさせる工程を含む。本明細書に記述されるようにバイオチップは設計することができる。ハイブリダイゼーション後、方法は、任意の数の洗浄工程をさらに含みうる。例えば、バイオチップ上のプローブへ核酸をハイブリダイズさせた後に、ハイブリダイズされない核酸を除去するために、例えば緩衝液または洗浄液でバイオチップを1回以上洗浄することができる。ハイブリダイズされない核酸は、さらなる工程のために、廃棄または収集されうる。

【0090】

いくつかの態様において、ハイブリダイズした核酸は、例えば、電気伝導度、容量または抵抗の変化によって検出することができる。場合によっては、バイオチップは画像化される。いくつかの例において、検出に先だってバイオチップに読み取り緩衝液を加えられる。読み取り緩衝液は、ハイブリダイズした核酸に適用すると検出可能なシグナルを生成する試薬を含みうる。他の例において、核酸分子は検出可能なように標識される。

【0091】

いくつかの態様において、追加の特異性のある工程を追加することができる。一例において、ハイブリダイゼーション後のライゲーション工程は、単一塩基のミスマッチを区別できる。特異性はまた、捕捉、検出、またはその両方のための追加のハイブリダイゼーション工程を加えることによって増大させることができる。例えば、図7Aは、捕捉配列(702)、および、隣接するまたは近隣の検出配列(703)を含む標的核酸(701)を示す。捕捉オリゴヌクレオチド(705)(例えば、捕捉配列に相補的な配列)を有するアレイ基板(704)を用いて、標的核酸にハイブリダイズしそれを捕捉(707)することができる。検出オリゴヌクレオチド(706)(例えば、蛍光標識されたもの)を用いて、標的核酸の検出配列にハイブリダイズ(708)させて、標的核酸の検出を可能

10

20

30

40

50

にすることができる。検出オリゴヌクレオチドは、非標識標的核酸を有する遊離溶液中に存在し得る。標的核酸への検出オリゴヌクレオチドのハイブリダイゼーションは、アレイ上で標的核酸の捕捉の前、最中または後に行うことができる。検出オリゴヌクレオチドは、完全に（またはほぼ完全に）相補的配列にハイブリダイズすることができる。検出は、迅速にハイブリダイズすることができるよう、比較的高い濃度で存在することができる。捕捉配列および検出された配列は近くにあるので、これらの2つの配列がDNAの小さな断片上で同時発生する可能性は低くなり、したがって、検出されたシグナルが非特異的である可能性が低くなる。

【0092】

場合によっては、捕捉配列および検出配列は、それらが標的核酸の別々の断片に位置する可能性を減少させるために十分に密接に配置される。

10

【0093】

捕捉配列および検出された配列が隣接している場合、エネルギー移動色素の組合せ（例えばFRET）の使用によりバックグラウンドを減少させることができる。一例では、図7Bに示されるように、ドナー色素は捕捉オリゴ（715）の3'末端上にあり、アクセプター色素は検出オリゴヌクレオチド（716）上の5'末端にある。捕捉配列（712）および検出配列（713）を含む標的核酸（711）は、アレイ基板（714）上で捕捉することができる。捕捉オリゴヌクレオチド（717）および検出オリゴヌクレオチド（718）の両方の標的核酸へのハイブリダイゼーションは、ドナーおよびアクセプター色素を互いにFRET距離にすることができる。一旦FRET距離内に入ると、励起光（719）はFRETドナーを励起することができ、次に共鳴エネルギー移動を介してFRETアクセプターを励起し、FRETシグナル（720）の生成を可能にする。ドナー色素およびアクセプター色素の位置および配置は変更することができる。例えば、ドナー色素は検出オリゴヌクレオチド上にあることができ、アクセプター色素は捕捉オリゴヌクレオチド上にあることができる。色素は、捕捉オリゴヌクレオチド以外の位置でアレイ表面に結合することができる。これらの考えは、同じ特徴の多数の捕捉遺伝子座に対して拡大可能である（scalable）。

20

【0094】

複数の捕捉および/または検出配列を单一の標的核酸に用いることができる。例えば、図7Cは、3つの捕捉配列（732）、（733）、（734）を含む標的核酸（731）の典型的な概略を示す。アレイ基板（735）は同様に、1つの特徴内に3つの捕捉オリゴヌクレオチド（736）、（737）、（738）を含む。ハイブリダイゼーション（739）の後、捕捉配列のそれぞれは、対応する捕捉オリゴヌクレオチドにハイブリダイズされる。標的核酸をアレイに結合されたままで保つために、全ての捕捉配列よりも少ない配列のハイブリダイゼーションが不十分であるように、局所条件（例えば、緩衝液組成、温度、pH）を構成することができ、分析の特異性を高める。複数の捕捉配列は、正しい標的配列が存在する場合には協調的に作用することができるが、相互作用が非特異的である場合には独立して作用する。ある領域が捕捉されると、他の隣接する領域は迅速に捕捉され、（例えば、捕捉配列の局所的な高濃度のために）除去するのが比較的難しい。

30

【0095】

同様に、複数の検出配列および対応する検出オリゴヌクレオチドを、すべての検出オリゴヌクレオチドの存在が陽性シグナルのために必要とされるように利用されうる。一例では、各検出オリゴヌクレオチドは異なる発光波長を有し、陽性シグナルを記録するため、異なる発光波長ごとにシグナルが検出される。別の例において、FRET対の検出オリゴヌクレオチドを使用することができ、標的核酸上の検出配列へのハイブリダイゼーションにより、それらを互いにFRET距離内にすることができる。異なる検出オリゴヌクレオチドを使用して異なる特性を認識することができる。例えば、一つの検出オリゴヌクレオチドを用いて被験体の同一性（例えば、種、株または個体）を示すことができ、一方、別の検出オリゴヌクレオチドを用いて、被験体の遺伝子、変異、または他の特徴（例えば、抗生物質耐性、ビルレンス）を示すことができる。

40

50

【0096】

いくつかの態様において、元のサンプル中に存在する被験体の同一性は、サンプル中の核酸の存在の検出に基づいて決定され得る。場合によっては、特定の被験体の標的が、特定の株を検出するように設計することができる。場合によっては、標的は、対象の種についての、または、病原性に関連する代表的な保存された領域（複数可）などの他の方法において固有である対象内に含まれる他の領域についての、プローブを含み得る。場合によっては、標的は、特定の人などの、特定の個体を差別化することができるプローブを含むことができる。個々のプローブセットは一意に特定の個体を識別することができる。他の場合では、個体の識別は固有のものではないことがあり、固有性のレベルに基づいて有用なコール信頼水準（call confidence level）を提供することができる。いくつかの実施形態では、個体コールの信頼水準（confidence level of an individual call）は、約50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、96%、96%、98%、99%、99.9%、99.99%、または、99.999%である可能性がある。いくつかの実施形態では、個体コールの信頼水準は、少なくとも、約50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、96%、96%、98%、99%、99.9%、99.99%、または、99.999%である可能性がある。いくつかの実施形態では、個体コールの信頼水準は、多くとも、約50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、96%、96%、98%、99%、99.9%、99.99%、または、99.999%である可能性がある。

10

20

【0097】

場合によっては、結果がレポートで提供される。そのレポートは、元のサンプル中の複数の被験体から識別された被験体を列挙することができる。被験体が、株特異的特徴、種特異的特徴、および被験体内に存在する保存された領域からなる特徴、などの複数の特徴によって表される場合、これらの他の特徴も検出されたかどうかをレポートで列挙することができる。次いで、最終的な陽性コールは、被験体に連結された特徴が全て検出されたか、全く検出されなかったか、一部が検出されたかどうかに基づいて計算された信頼水準に基づきうる。

【0098】

いくつかの態様において、方法は、検出の後に十分なサンプルを保存する能力を含む。これは、例えば、サンプル調製プロセスの時点でサンプルをAサンプルおよびBサンプルに分けることによって達成することができる。Aサンプルは、完全なサンプル調製プロセスを進めることができ、Bサンプルは、後の処理のためにリザーバに流用することができる。別の例では、サンプルのハイブリダイズされない部分は、例えば、ハイブリダイズされない核酸を除去するためにバイオチップを洗浄することによって、ハイブリダイゼーションプロセス後にリザーバに転換され得る。さらに別の例では、ハイブリダイズされた核酸は、後でクエリするために、脱ハイブリダイズ（de-hybridized）し、リザーバに転換されうる。

30

【0099】

本開示の技術は、工程の一部または全部について自動化した操作を使用して実施することができる。例えば、場合によっては、ユーザーが行う唯一の工程は、サンプルの充填であり、流体ハンドリング、アッセイ、検出、および結果の報告などの他のすべての工程は、自動的に行われる。他の場合では、サンプルの充填でさえ自動的に行うことができる。例えば、ラボラトリ自動化機器または環境サンプリング機器を使用して、分析のためのデバイスにサンプルを提供することができる。

40

【0100】

<検出>

本開示のいくつかの態様において、バイオチップ上のプローブへの標的核酸の結合が検出される。検出は、当業者に公知の任意の方法を包含し得る。場合によっては、検出は、標的核酸分子、プローブ、またはその両方に存在する検出可能な標識を検出することを含

50

む。他の場合では、検出は、標的核酸分子とプローブとの相互作用に基づいて生成されるシグナルを検出することを含む。

【0101】

場合によっては、シグナルは蛍光共鳴エネルギー移動 (FRET) に基づくシグナルである。この例において、標的核酸分子およびプローブの両方が1つ以上の蛍光標識で標識される。1つ以上の蛍光標識は、1つ以上のFRET対であり得る。1つ以上のFRET対は、少なくとも1つのFRETドナーおよび少なくとも1つのFRETアクセプターを含むことができる。場合によっては、FRETドナーは標的核酸分子に付着し、FRETアクセプターはプローブに付着する。他の場合では、FRETアクセプターは標的核酸分子に付着し、FRETアクセプターはプローブに付着する。FRETドナーおよびアクセプターは、標的核酸分子およびプローブのいずれかの末端 (3'または5') に付着し得る。場合によっては、FRETドナーはCy3であり、FRETアクセプターはCy5である。FRET対の他の非限定的な例には、FITC/TRITC、EGFP/Cy3、CFP/YFP、およびEGFP/YFPが含まれる。

10

【0102】

他の場合では、検出は、標的核酸上に存在する検出可能な標識を検出することを含む。この例において、検出可能に標識された標的核酸分子がプローブに結合されている場合にのみ、シグナルを検出することができる。場合によっては、標的核酸分子はCy5で5'標識される。場合によっては、非標識ハイブリダイゼーション結合は、例えば、インターフェロメトリックオシレータ上の表面状態の差を検出すること、光路中の屈折率差、または走査型電子顕微鏡法 (SEM) 技術を用いた直接検出によって検出することができる。

20

【0103】

光検出が利用される場合、バイオチップは、CMOSカメラなどの光検出器の表面にあり得るか、または、バイオチップは外部光学システムによってプローブで探されるフローセル内にあり得る。光学的蛍光検出の場合、蛍光共焦点顕微鏡法を含む従来の蛍光光学的検出アーキテクチャを用いることができる。プローブがCMOS検出器に直接固定されている場合、CMOS検出器は、システムの励起光を遮断し、選択された色素からの光を通過させる、CMOS検出器とプローブとの間の層を有することができる。

【0104】

蛍光標識に加えて、他の標識を使用することができる。標識は、それ自体が検出可能であることができ、または他の検出可能な種の結合を可能にすることができる。典型的な標識には、限定されないが、フルオロフォア、ナノ粒子（例えば金ナノ粒子）、量子ドット、放射性標識、磁性粒子、バーコード（例えば核酸バーコード）、活性部位、結合部位、FRET可能標識、疎水性種、親水性種、抗体、アプタマーを含む。それ自身では検出できない標識は、その後、検出可能な種と接触させることができる。例えば、標的核酸をビオチンで標識し、続いて検出のためにストレプトアビシン結合フルオロフォアに結合させることができる。別の例では、標的核酸を核酸バーコードで標識することができ；続いて、核酸バーコード配列を増幅して検出することができる。検出様式は、限定されないが、光検出 (FRET、蛍光寿命および他の光学的特性を含む)、電気的検出、磁気的検出、放射性標識検出、配列決定、サイズ検出（例えば、電気泳動分離による）、表面プラズモン共鳴 (SPR)、ラマン分光法、および質量分析法を含む。

30

【0105】

特徴間の相対的なシグナルは、特徴ごとに結合する標識の数によって判定することができる。例えば、図2に示されるように、異なる数のフルオロフォアが特徴ごとに結合することができ、結果として相対輝度が異なる。図2は、特徴あたり1フルオロフォア（上）、特徴あたり10フルオロフォア（中）、および特徴あたり50フルオロフォア（下）の相対輝度を示す。

40

【0106】

<コンピュータシステム>

本開示のシステムは1つ以上のコンピュータシステムを含むことができる。本開示の技

50

術およびデバイスは、操作、自動化、サンプル処理、データ処理、データの送信、分析、結果の提示、および他の機能のためのコンピュータシステムを使用することができる。図3は、データを受信し、サンプル中の被験体の存在または非存在を識別するなど、本開示の方法を実施するようにプログラムされた、または他の形で構成された、コンピュータシステム(301)を示す。コンピュータシステム(301)は、單一コアまたはマルチコアプロセッサ、または並列処理のための複数のプロセッサとすることができる中央処理装置(CPU、本明細書では「プロセッサ」および「コンピュータプロセッサ」)(305)を含む。コンピュータシステム(301)はまた、メモリ(310)(例えば、ランダムアクセスメモリ、リードオンリーメモリ、フラッシュメモリ)と、電子ストレージユニット(315)(例えば、ハードディスク)と、一つ以上の他のコンピュータシステムと通信するための通信インターフェース(320)(例えば、ネットワークアダプタ)と、キャッシュ、他のメモリ、データストレージおよび/または電子ディスプレイアダプタなどの周辺装置(125)と、を含む。メモリ(310)、ストレージユニット(315)、インターフェース(320)および周辺装置(325)は、マザーボードなどの通信バス(実線)を介してCPU(305)と通信する。ストレージユニット(315)は、データを保存するためのデータストレージユニット(またはデータリポジトリ)であってもよい。コンピュータシステム(301)は、通信インターフェース(320)の助けを借りてコンピュータネットワーク(「ネットワーク」)(330)に動作可能に結合される。ネットワーク(330)は、インターネット、インターネットおよび/またはエクストラネット、またはインターネットと通信しているイントラネットおよび/またはエクストラネット、であることができる。ネットワーク(330)は、場合によっては、電気通信および/またはデータネットワークである。ネットワーク(330)は、クラウドコンピューティングのような分散コンピューティングを可能にすることができる1つ以上のコンピュータサーバを含むことができる。ネットワーク(330)は、場合によっては、コンピュータシステム(301)の助けを借りて、コンピュータシステム(301)に連結されたデバイスがクライアントまたはサーバとして動作することを可能にするピアツーピアナットワークを実装することができる。コンピュータシステムは、物理的にデバイスの近くにある必要はなく；有線または無線のモダリティを介してデバイスと通信することができる。

【0107】

コンピュータシステム(301)は、処理システム(335)と通信することができる。処理システム(335)は、1つ以上の標的核酸配列の存在を識別する、または複数の被験体をレポートで分類するなど、本明細書に開示される方法を実施するように構成することができる。処理システム(335)は、ネットワーク(330)を通じて、または直接(例えば、有線、無線の)接続によって、コンピュータシステム(301と)通信することができる。処理システム(335)は、核酸配列分析などの分析のために構成することができる。

【0108】

本明細書で記載される方法およびシステムは、例えばメモリ(310)または電子ストレージユニット(315)などのコンピュータシステム(301)の電子ストレージ位置に保存されたマシン(またはコンピュータプロセッサ)実行可能コード(またはソフトウェア)によって実行されうる。使用中に、コードはプロセッサ(305)によって実行することができる。いくつかの例において、コードは、ストレージユニット(315)から検索され、プロセッサ(305)による容易なアクセスのためにメモリ(310)上に保存されうる。いくつかの状況では、電子ストレージユニット(315)を排除することができ、マシン実行可能命令はメモリ(310)に保存される。

【0109】

コードは、コードを実行するように適合されたプロセッサを有するマシンと共に使用するためにはコンパイルされかつ構成されることができ、または、実行時にコンパイルされることができ、または実行中に解釈することができる。コードは、プレコンパイルされ

10

20

30

40

50

た方法か、コンパイルされたままの方法か、または、解釈された方法でコードを実行できるように選択できるプログラミング言語で提供され得る。

【 0 1 1 0 】

本明細書に提供されたシステムと方法の態様はプログラミングで具体化することができる。この技術の様々な態様は、マシン可読媒体の一種に搭載されているか、または組み込まれたマシン（またはプロセッサ）実行可能コードおよび／または関連データの形態で、典型的には「製品」または「製造品目」とみなすことができる。マシン実行可能コードは、メモリ（例えば、リードオンリーメモリ、ランダムアクセスメモリ、フラッシュメモリ）またはハードディスクなどの電子ストレージユニットに保存することができる。「ストレージ」タイプの媒体は、様々な半導体メモリ、テープドライブ、ディスクドライブなどの、コンピュータ、プロセッサなどの有形のメモリ、またはそれらの関連するモジュールのいずれかまたはすべてを含むことができ、ソフトウェアプログラミングのためにいつでも非一時的なストレージを提供し得る。ソフトウェアの全部または一部は、時には、インターネットまたは様々な他の電気通信ネットワークを介して通信され得る。このような通信は、例えば、管理サーバまたはホストコンピュータからアプリケーションサーバのコンピュータプラットフォームへといった、あるコンピュータまたはプロセッサから、別のコンピュータまたはプロセッサへの、ソフトウェアのローディングを可能にすることができる。したがって、ソフトウェア要素を保持できる別のタイプの媒体には、ローカルデバイス間の有線および光陸上通信線ネットワーク、および様々なエアリンクを介して使用されるような、光、電気および電磁波が含まれる。例えば有線または無線リンク、光リンクなどの、このような波を運ぶ物理的要素も、ソフトウェアを担持する媒体とみなすことができる。本明細書に使用されるように、非一時的な有形「ストレージ」媒体に限定されない限り、コンピュータまたはマシンの「可読媒体」などの用語は、実行のためにプロセッサに命令を提供することに関与する任意の媒体を指す。

10

20

30

40

【 0 1 1 1 】

従って、コンピュータ実行可能コードなどのマシン可読媒体は、有形ストレージ媒体、搬送波媒体または物理的伝送媒体を含むが、これに限定されない、多くの形態をとることができる。不揮発性ストレージ媒体には、例えば、データベースなどを実装するために使用することができるような、任意のコンピュータなどにおけるストレージデバイスのいずれかなどの光学ディスクまたは磁気ディスクが含まれる。揮発性ストレージ媒体は、そのようなコンピュータプラットフォームのメインメモリのような動的メモリを含む。有形の伝送媒体には、同軸ケーブルが含まれ；コンピュータシステム内のバスを備えるワイヤを含む銅線および光ファイバを含む。搬送波送信媒体は、電気または電磁信号、または無線周波数（R F）および赤外線（I R）データ通信中に生成されるような音響波または光波の形態をとることができる。したがって、コンピュータ可読媒体の共通の形式には、例えば：フロッピーディスク、フレキシブルディスク、ハードディスク、磁気テープ、他の任意の磁気媒体、C D - R O M、D V D またはD V D - R O M、任意の他の光学媒体、パンチカード紙テープ、穴のパターンを有する他の任意の物理的ストレージ媒体、R A M、R O M、P R O M およびE P R O M、F L A S H - E P R O M、他の任意のメモリチップまたはカートリッジ、データまたは命令を搬送する搬送波、そのような搬送波を搬送するケーブルまたはリンク、または、コンピュータがプログラミングコードおよび／またはデータを読み取ることができる他の任意の媒体が、含まれる。これらの形式のコンピュータ可読媒体の多くは、実行のために1つ以上の命令の1つ以上のシーケンスをプロセッサに送ることに関わりうる。

【 0 1 1 2 】

コンピュータシステム（301）は、例えば、本開示の方法によって分析され得る遺伝的変異体のカスタマイズ可能なメニューを提供するためのユーザーインターフェース（U I）を含む電子ディスプレイを含むか、またはそれと通信することができる。U Iの例としては、限定しないが、グラフィカルユーザーインターフェース（G U I）およびウェブベースのユーザーインターフェースを含む。

50

【0113】

幾つかの場合では、コンピュータシステム(301)は、ユーザーに視覚情報を提供するディスプレイを含む。幾つかの場合では、ディスプレイは、陰極線管(CRT)である。幾つかの場合では、ディスプレイは、液晶ディスプレイ(LCD)である。さらなる例では、ディスプレイは、薄膜トランジスタ液晶ディスプレイ(TFT-LCD)である。幾つかの場合では、ディスプレイは、有機発光ダイオード(OLED)ディスプレイである。様々な更なる例では、OLEDディスプレイは、パッシブマトリックスOLED(PMOLED)またはアクティブマトリックスOLED(AMOLED)のディスプレイである。幾つかの場合では、ディスプレイは、プラズマディスプレイである。他の場合では、ディスプレイは、ビデオプロジェクターである。またさらなる場合では、ディスプレイは、本明細書に開示されるものなどのデバイスの組み合わせである。ディスプレイは、本明細書に記載される方法によって生成されるようなエンドユーザーに1つ以上の生物医学的なレポートを提供することができる。

【0114】

幾つかの場合では、コンピュータシステム(301)は、ユーザーから情報を受信する入力デバイスを含む。幾つかの実施形態では、入力デバイスは、キーボードである。幾つかの実施形態では、入力デバイスは、限定しない例として、マウス、トラックボール、トラックパッド、ジョイスティック、ゲームコントローラ、またはスタイルスを含む、ポインティングデバイスである。幾つかの場合では、入力デバイスは、タッチスクリーンまたはマルタッチスクリーンである。他の場合では、入力デバイスは、声または他の音入力を捕らえるマイクロфонである。他の場合では、入力デバイスは、動作または視覚の入力を捕らえるビデオカメラである。またさらなる例では、入力デバイスは、本明細書に開示されるものなどのデバイスの組み合わせである。

【0115】

コンピュータシステム(301)は、1つ以上のデータベースを含むことができるか、またはそれに動作可能に連結され得る。データベースは、ゲノム、プロテオミクス、薬理ゲノミクス、生物医学、および科学のデータベースを含むことができる。データベースは、公に利用可能なデータベースであり得る。代替的に、またはさらに、データベースは、プロプライエタリ・データベースを含むことができる。データベースは、市販のデータベースであり得る。データベースは、限定されないが、MendelDB、PharmGKB、Varimed、Regulome、curated BreakSeq junctions、Online Mendelian Inheritance in Man(OMIM)、Human Genome Mutation Database(HGMD)、NCBI dbSNP、NCBI RefSeq、GENCODE、GO(gene ontology)、およびKyoto Encyclopedia of Genes and Genomes(KEGG)を含む。

【0116】

データは、データのユーザーと同じ国を含む地理的位置において生成及び/又は送信され得る。データは、例えば、1か国の地理的位置から生成及び/又は送信され、データのユーザーは異なる国に存在し得る。幾つかの場合では、本開示のシステムによってアクセスされたデータは、複数の地理的位置の1つからユーザーに送信され得る。データは、例えば、ネットワーク、安全なネットワーク、不安定なネットワーク、インターネット、またはインターネットによって複数の地理的位置の中から送受信され得る。

【0117】

トータルシステムは、各々が個別モジュールを有する3つの主要なコンポーネントを使用することができる様々な方法で設計され得る。コンポーネントは、ユーザーインターフェース、ハードウェアプラットフォーム、および消耗部品(consumable)を含むことができる。一例では、ユーザーインターフェースは、システムとの直接的なインタラクションを可能にするハードウェアに組み込まれ、消耗部品は、本開示の方法を実行するのに必要な試薬を含むことができる。システムインテリジェンスのすべてを含む又は全

く含まない遠隔の、無線接続されたユーザーインターフェース、自動化またはヒューマンインターラクションを介して接続された1つ以上の複数のハードウェアコンポーネント、およびすべての試薬およびアッセイに存在する生物学的情報をプロセシングおよび報告のためにコンピュータに転送され得るデジタルまたはアナログの情報に変換するのに必要とされる検出システムのすべて又は一部を含む完全に内包された消耗部品を使用して、より複雑なアーキテクチャが設計され得る。システムは、限定しないが、システムによって連続的に又は並列して処理される多数のサンプルを制御するスケジューリングモジュール、およびシステムが、生物的制御を使用する内部機能のチェックまたは試験を介して適切に動作していることを確かなものとする品質保証モジュールを含む、複数の追加のモジュールを含むことができる。

10

【0118】

図9Aに例示されるコンピュータシステム(900)は、媒体(906)及び/又は固定媒体(907)を有しているサーバー(905)に随意に接続され得るネットワークポート(903)からの命令を読み取ることができる論理装置として理解され得る。図9Aに示されるシステムなどのシステムは、CPU(901)、を含むことができる、ディスクドライブ(902)、キーボード及び/又はマウスなどの随意の入力デバイス(908)、および随意のモニター(904)を含むことができる。データ通信は、示された通信媒体を通って、ローカル位置またはリモート位置でのサーバーへと達成され得る。通信媒体は、データを送信及び/又は受信するあらゆる手段を含むことができる。例えば、通信媒体は、ネットワーク接続、無線接続またはインターネット接続であり得る。そのような接続は、ワールドワイドウェブ上の通信を提供することができる。本開示に関連するデータは、図9Aに例示されるような当事者(909)による受信及び/又はレビューのためのそのようなネットワークまたは接続を介して送信され得ることが想定される。

20

【0119】

図9Bは、本開示の例となる実施形態に関連して使用され得るコンピュータシステム(910)の第1の例となるアーキテクチャを例示するブロック図である。図9Bで示されるように、例となるコンピュータシステムは、命令を処理するためのプロセッサ(911)を含むことができる。プロセッサの限定しない例としては、以下が挙げられる：Intel Xeon(商標)プロセッサ、AMD Opteron(商標)プロセッサ、Samsung 32-bit RISC ARM 1176JZ(F)-S v1.0(商標)プロセッサ、RM Cortex-A8 Samsung S5PC100(商標)ARM Cortex-A8 Apple A4(商標)プロセッサ、Marvell PXA 930(商標)プロセッサ、または機能的に同等なプロセッサ。実行の複数のスレッドは、並列処理に使用され得る。幾つかの実施形態では、單一コンピュータシステム、クラスターであろうと、または複数のコンピュータ、携帯電話、及び/又はパーソナルデータアシスタントデバイスを含むネットワーク上でシステムにわたって分配されようと、複数のコアを備える複数のプロセッサが使用され得る。図9Bで例示されるように、最近プロセッサ(911)によって使用されている、または頻繁に使用される命令またはデータのための高速メモリを提供するために、高速キャッシュ(912)が、プロセッサ(911)に接続されるか、またはそこに組み込まれ得る。プロセッサ(911)は、プロセッサバス(914)によってノースブリッジ(913)に接続される。ノースブリッジ(913)は、メモリーバス(916)によってランダムアクセスメモリ(RAM)(915)に接続され、プロセッサ(911)によってRAM(915)に対するアクセスを管理する。ノースブリッジ(913)はまた、チップセットバス(918)によってサウスブリッジ(917)に接続される。サウスブリッジ(917)は、順に、周辺バス(919)に接続される。周辺バスは、例えば、PCI、PCI-X、PCI Express、または他の周辺バスであり得る。ノースブリッジおよびサウスブリッジは、しばしば、プロセッサチップセットと呼ばれ、プロセッサ、RAMと、周辺バス(919)上の周辺コンポーネントの間のデータ転送を管理する。幾つかの代替的なアーキテクチャでは、ノースブリッジの機能性は、別のノースブリッジチップを使用する代わりに、プロセッサに組み込まれ得る。幾つ

30

40

50

かの実施形態では、システム(910)は、周辺バス(919)に付けられたアクセラレータカード(922)を含むことができる。アクセラレータは、フィールドプログラマブルゲートアレイ(FPGA)または特定のプロセシングを加速するための他のハードウェアを含むことができる。例えば、アクセラレータは、適応データの再構成のために、または拡張された設定処理に使用される代数式を評価するために使用され得る。ソフトウェアおよびデータは、外部ストレージ(923)中に保存され、プロセッサによる使用のために、RAM(915)及び/又はキャッシュ(912)にロードされ得る。システム(910)は、システムリソースを管理するためのオペレーティングシステムを含み、オペレーティングシステムの限定しない例としては、以下が挙げられる：Linux(登録商標)、Windows(商標)、MAC OS(商標)、BlackBerry OS(商標)、iOS(商標)、および他の機能的に同等なオペレーティングシステムの他に、本開示の例となる実施形態に従ってデータの記憶および最適化を管理するためのオペレーティングシステム上で実行するアプリケーションソフトウェア。本実施例では、システム(910)はまた、ネットワーク接続ストレージ(NAS)、および並列処理の分配に使用され得る他のコンピュータシステムなどの、外部ストレージにネットワークインターフェースを提供するための周辺バスに接続されたネットワークインターフェースカード(NIC)(920)および(921)を含むことができる。

【0120】

図9Cは、複数のコンピュータシステム(931)、および(932)、複数の携帯電話およびパーソナルデータアシスタント(933)、およびネットワーク接続ストレージ(NAS)(934)、および(935)を備えるネットワーク(930)を示す図である、例となる実施形態では、システム(931)、(932)および(933)は、データの記憶を管理し、ネットワーク接続ストレージ(NAS)(934)および(935)に保存されたデータのためのデータアクセスを最適化することができる。数学的モデルが、データに使用され、コンピュータシステム(931)および(932)、および携帯電話およびパーソナルデータアシスタントのシステム(933)にわたる分配された並列処理を使用して評価され得る。コンピュータシステム(931)および(932)、および携帯電話およびパーソナルデータアシスタントのシステム(933)はまた、ネットワーク接続ストレージ(NAS)(934)および(935)に保存されたデータの適応データの再構成に並列処理を提供することができる。図9Cは、単に一例を例示するものであり、種々様々な他のコンピューターのキテクチャーおよびシステムが、本開示の様々な実施形態とともに使用され得る。例えば、並列処理を提供するためにブレードサーバーを使用することができる。並列処理を提供するために、プロセッサブレードがバックプレーンで接続され得る。ストレージも、バックプレーンに接続され得るか、または別のネットワークインターフェースを介してネットワーク接続ストレージ(NAS)として接続され得る。幾つかの例となる実施形態では、プロセッサは、別のメモリ空間を維持することができ、ネットワークインターフェース、バックプレーン、または他のプロセッサによる並列処理のための他のコネクターを介してデータを送信することができる。他の実施形態では、プロセッサの幾つかまたはすべては、共有の仮想アドレスメモリ空間を使用することができる。

【0121】

図9Dは、例となる実施形態による共有の仮想アドレスメモリ空間を使用して、マルチプロセッサコンピュータシステム(940)のブロック図である。システムは、共有メモリサブシステム(942)にアクセスすることができる複数のプロセッサ(941a) - (941f)を含む。システムは、複数のプログラム可能なハードウェアメモリアルゴリズムプロセッサ(MAP)(943a) - (943f)を共有メモリサブシステム(942)に組み込む。MAP(943a) - (943f)はそれぞれ、メモリ(944a) - (944f)および1つ以上のフィールドプログラマブルゲートアレイ(FPGA)(945a) - (945f)を含むことができる。MAPは、設定可能な機能ユニットを提供し、特定のアルゴリズムまたはアルゴリズムの部分が、それぞれのプロセッサと密に協働

10

20

30

40

50

した処理のために F P G A (9 4 5 a) - (9 4 5 f) に提供され得る。例えば、 M A P は、データモデルに関する代数式を評価する及び例となる実施形態において適応データの再構成を実行するために使用され得る。本実施例では、 M A P はそれぞれ、これらの目的のためのプロセッサのすべてによって地球規模で利用可能である。1つの構成では、 M A P はそれぞれ、関連するメモリ (9 4 4 a) - (9 4 4 f) にアクセスするために直接メモリアクセス (D M A) を使用することができ、これによって、それぞれのマイクロプロセッサ (9 4 1 a) - (9 4 1 f) とは無関係に、およびそれらとは非同期的にタスクを実行することが可能になる。この構成では、 M A P は、アルゴリズムのパイプライン処理および並列実行のために別の M A P に結果を直接供給することができる。

【 0 1 2 2 】

上記のコンピュータアーキテクチャおよびシステムは、単に例であり、汎用のプロセッサ、コプロセッサ、 F P G A および他のプログラマブルロジックデバイス、システムオンチップ (S O C) 、特定用途向け集積回路 (A S I C) 、並びに他の処理素子および論理素子のあらゆる組み合わせを使用するシステムを含む、種々様々な他のコンピュータ、携帯電話、パーソナルデータアシスタントのアーキテクチャおよびシステムが、例となる実施形態に関連して使用され得る。幾つかの実施形態では、コンピュータシステムのすべてまたは一部は、ソフトウェアまたはハードウェアにおいて実施され得る。ランダムアクセスメモリ、ハードドライブ、フラッシュメモリ、テープドライブ、ディスクアレイ、ネットワーク接続ストレージ (N A S) 、他のローカルまたは配信されたデータストレージデバイスおよびシステムを含む、あらゆる種類のデータストレージ媒体も、例となる実施形態に関連して使用され得る。

【 0 1 2 3 】

例となる実施形態では、コンピュータシステムは、上記または他のコンピュータアーキテクチャおよびシステムのいずれかの上で実行するソフトウェアモジュールを使用して実施され得る。他の実施形態では、システムの機能は、ファームウェア、図 9 D で参照されるようなフィールドプログラマブルゲートアレイ (F P G A) 、などのプログラマブルロジックデバイス、システムオンチップ (S O C) 、特定用途向け集積回路 (A S I C) 、並びに他の処理素子および論理素子において部分的または完全に実施され得る。例えば、セットプロセッサ (S e t P r o c e s s o r) およびオプティマイザー (O p t i m i z e r) は、図 9 B に例示されるアクセラレータカード (9 2 2) などのハードウェアアクセラレータカードの使用によってハードウェアアクセラレーションで実施され得る。

【 0 1 2 4 】

< レポート >

方法およびシステムはレポートの作成を提供し、ここでレポートは、複合サンプルに存在する1つ以上の被験体を特定することができる。代替的に、レポートは、マイクロアレイ上に含まれるすべての特徴の読み出しに関する詳細情報を提供することができる。レポートは、本明細書に記載される方法の結果がエンドユーザーに中継される技術であり得る。レポートは、スクリーンまたは電子ディスプレイに表示され得るか、または例えば1枚の紙に印刷され得る。幾つかの場合では、レポートはネットワークを介して送信される。幾つかの場合では、ネットワークはインターネットである。幾つかの場合では、レポートは手動で作成され得る。他の場合では、レポートは自動的に作成され得る。幾つかの場合では、レポートはリアルタイムで作成され得る。幾つかの場合では、レポートは、モバイルデバイス、スマートフォン、タブレットまたは別のネットワーク使用可能な装置に提供され得る。

【 実施例 】

【 0 1 2 5 】

実施例 1 . 主題の特異的なプローブの生成

サンプルを得る。サンプルは複数の被験体を含む。特定されるべき各被験体のゲノムが得られる。被験体のゲノムの非重複領域が特定される。非重複領域に特異的なプローブが設計される。

10

20

30

40

50

【0126】

実施例2. 主題の特異的な特徴の構築

主題の特異的な特徴を含むバイオチップを構築する。特徴はそれぞれ、個々の被験体に特異的な複数のプローブを含む。

【0127】

実施例3. バイオチップを使用する被験体の存在に関する分析

多くのタイプの被験体を含む試験サンプルを得る。サンプルからのDNAをひとまとめに得る。増幅なしで、DNAをバイオチップにハイブリダイズする。複数の標的がバイオチップの表面上のプローブに結合する。十分な数のプローブが特徴内で結合されると、シグナルが検出可能であり、被験体に特異的な特徴は陽性(positive)と呼ばれる。陽性の特徴は、サンプル中の被験体の存在を暗示している。幾つかの場合では、陽性シグナルは、特異的な有機体または種の存在を示す。別の場合では、陽性シグナルは、対象の特異的な遺伝子または形質の存在を示す。

10

【0128】

実施例4. 2段階EDCプロトコルを使用するプローブの固定化の改善

2段階EDC(1-エチル-3-(3-ジメチルアミノプロピル)カルボジイミド塩酸塩)プロトコルを利用して、シリカビーズへのプローブの固定化を改善した。シリカビーズを、低濃度のEDCで処理し、洗浄し、続いてより高濃度のEDCで処理した。下記の表1は、プローブの固定化の効率に対する異なるEDC濃度の効果を実証している。

20

【0129】

【表1】

表1. 2段階EDCプロトコルは、シリカビーズへのプローブの固定化を改善する。

名称	EDC 濃度 (mM)	rxn内の #ビーズ	加えた Cy5 (nmole)	結合し た Cy5 (nmole)	非結合の Cy5 (nmole)	収率	1ビーズ 当たりの #プロー ブ	プローブ 間隔(Å)
ED C-0	3.2/6.4	2.39×10^9	1.12	0.20	0.89	18%	49,900	79.4
ED C-1	32/64	1.19×10^9	0.56	0.43	0.62	41%	107,900	54.0
ED C-2	64/128	1.19×10^9	0.56	0.31	0.55	36%	78,300	63.4
ED C-3	96/192	1.19×10^9	0.56	0.22	0.53	29%	55,200	75.4

30

【0130】

実施例5. M13mp8配列に対するプローブ設計

プローブを、M13バクテリオファージに由来するM13mp8ファージベクターに対して設計した。簡単には、M13mp8配列を、GenBank Viruses, Bacteria, and Humanのデータベースおよび「天然の」M13バクテリオファージ配列に対してクエリした。M13mp8が22の固有の領域(例えば図4を参照)および380の固有の35-merを有することが判定された。これらの配列を使用して、M13mp8を複合サンプルと区別することができる10のプローブを生成した。プローブを、様々なGC含量/TmおよびヘアピンTmを有するように設計した。幾つかの場合では、プローブを以下を含むように修正した: 1. 修飾なし; 2. アミノ修飾(5'); 3. アミノ修飾(5') + Cy5(3')。幾つかの場合では、標的核酸を、修飾なし又は

40

50

C y 3 (5 ') 修飾のいずれかで生成した。図 5 は、本明細書に提供される方法を使用して設計されたプローブの例を示す。

【 0 1 3 1 】

実施例 6 . 結核サンプルの分析

ウイルス (T B V) および非ウイルス (T B A) の結核サンプルの 2 つのカテゴリーを、ビーズベースのプローブを使用して分析した。プローブを、異なるプールに入れて選択し、様々な多重のプローブを生成した。プローブを、その後、 1 ミクロロンのビーズ上に置いた。その後、これらのビーズを、 T B 株の特異的な標的 (T B V) およびプローブが結合するべきでない非特異的な標的 (T B A) を使用してハイブリダイズした。

【 0 1 3 2 】

アッセイを、以下のプロトコルに従って実行した。ハイブリダイゼーション緩衝液および洗浄緩衝液は、表 2 に記載されている。 10 mg / mL のビーズ 4 μ L を、 1 X ハイブリダイゼーション緩衝液とともに 200 μ L の全容積中に希釈した。その後、ビーズ溶液を超音波処理し (Bransoni 2510 の超音波処理器において 1 分間) 、 10 μ L のビーズ溶液を、最終的なハイブリダイゼーション溶液に加えて、 0.1 mg / mL の終末濃度にした。最終的なハイブリダイゼーション溶液は、 1 X ハイブリダイゼーション緩衝液中に 20 μ L の標識された DNA およびビーズを含んだ。 10 μ L の DNA を、 1 X ハイブリダイゼーション緩衝液中の等価な望ましい細胞で最終的なハイブリダイゼーション溶液に加えた。サンプルを混合し、遠心沈殿させた。その後、温度を、 5 分間最大 95 まで上昇させ、その後、毎分 2 の速度で 42 まで低下させた。その後、サンプルを、再び遠心沈殿させ、箔で覆い、下垂 (nutate) 反応を、 42 で一晩 (約 16 時間) 実行した。その後、サンプルを、 100 μ L の 1 X ハイブリダイゼーション緩衝液中で 2 回洗浄し、各洗浄工程で 80 μ L 除去し、各々の再懸濁後にボルテックスした。その後、洗浄後の残りの最終的な 20 μ L を、ボルテックスおよび超音波処理し、総量をフローセルに加えた。フローセルにおけるサンプルを、 10 ~ 15 分間インキュベートした。その後、レーンをそれぞれ、 150 μ L (3 X 50 μ L) の 1 X 洗浄緩衝液で洗浄した。その後、結果を、顕微鏡での観察によって収集した。平均シグナルおよび平均バックグラウンドを、ハイブリダイゼーション反応当たりの少なくとも 30 のビーズに対して測定した。

【 0 1 3 3 】

10

20

30

40

50

【表 2】

表2. ハイブリダイゼーションおよび洗浄の緩衝液。

1X ハイブリダイゼーション緩衝液 – 每日新鮮に作る			
	ストック	1X ハイブリダイゼーション緩衝液	1X 量 (μL)
SSC 緩衝液 (x)	20	5	500
SDS (%)	10	0.1	20
ホルムアミド、 脱イオン化(%)	100	50	1000
H ₂ O			480
合計			2000

1X 洗浄緩衝液 – 室温で維持し、再利用する			
	ストック	1X 洗浄緩衝液	1X 量 (μL)
SSC 緩衝液 (x)	20	0.1	100
SDS (%)	10	0.1	200
H ₂ O			19700
合計			20000

【0134】

図10は、実験からの結果を示す。X軸は、固有のプローブの数を表わす、プローブの煩雑性 (plexity) (1、4、9、および12)を示し、y軸は、バックグラウンドより上の平均シグナルを示す。特異的な標的シグナル (TBV、右)は、特異的な標的に対するプレックス (plex) 因子の各々の増加に応じて増加する。さらに、非特異的なシグナル (TBA、左)は、プレックス因子増加と横ばい状態である (flat)。

【0135】

本発明は、1つ以上の好ましい実施形態の観点から記載されており、当然のことながら、明確に述べられていない限り、多くの同等物、代替案、変化および修正が可能であり、本発明の範囲内にある。

10

20

30

40

50

【図面】

【図 1 A】

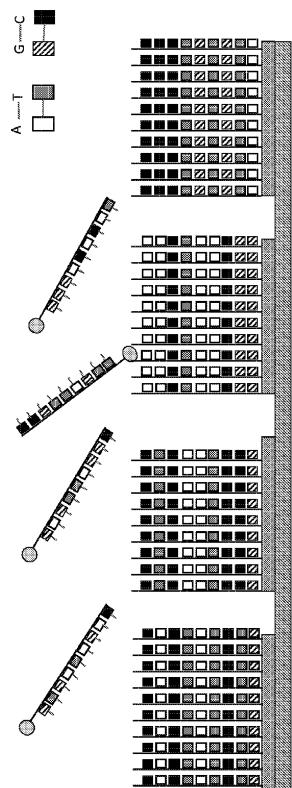


FIG. 1A

【図 1 B】

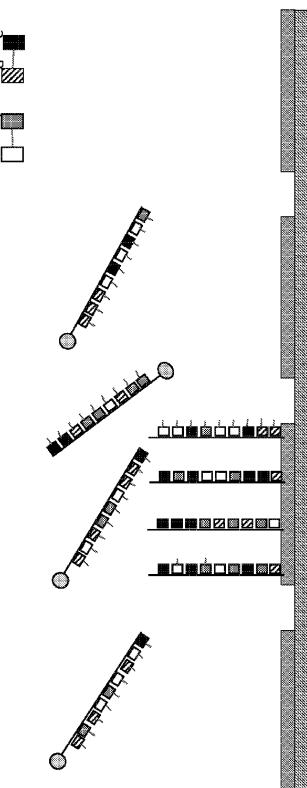


FIG. 1B

10

20

30

40

50

【図 1 C】

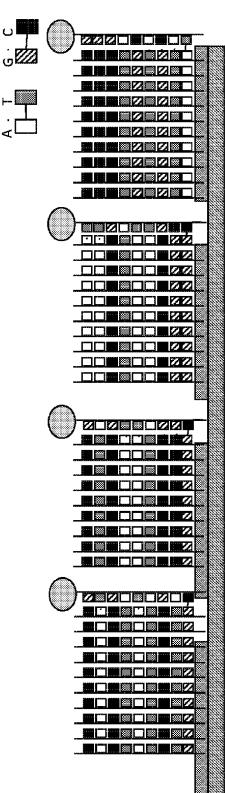


FIG. 1C

【図 1 D】

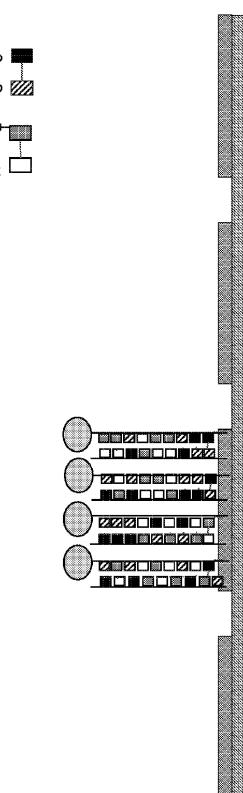


FIG. 1D

【図 1 E】

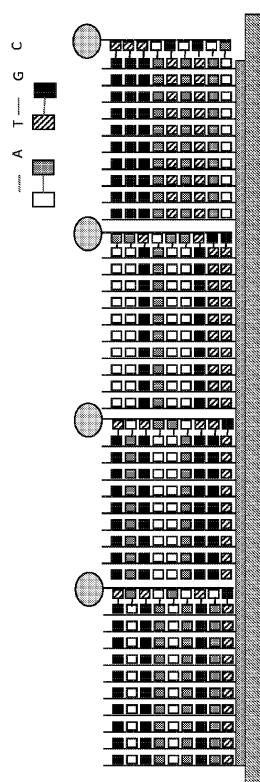


FIG. 1E

【図 1 F】

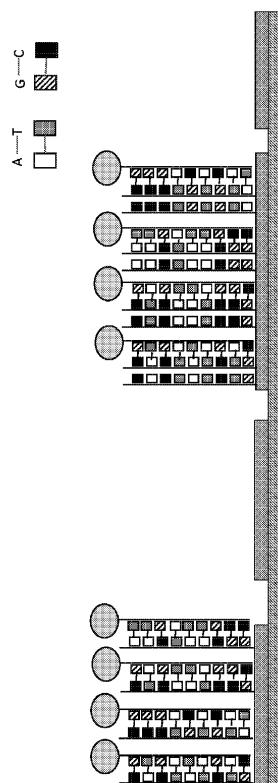


FIG. 1F

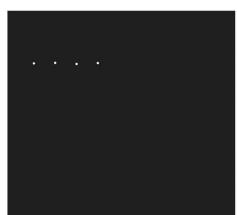
10

20

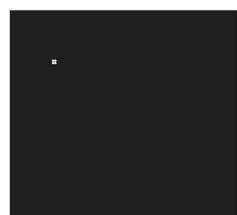
30

40

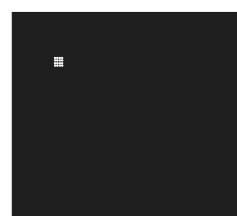
【図 2】



相対輝度 = 1 のフルオロフォア/特徴



相対輝度 = 10 のフルオロフォア/特徴



相対輝度 = 50 のフルオロフォア/特徴

【図 3】

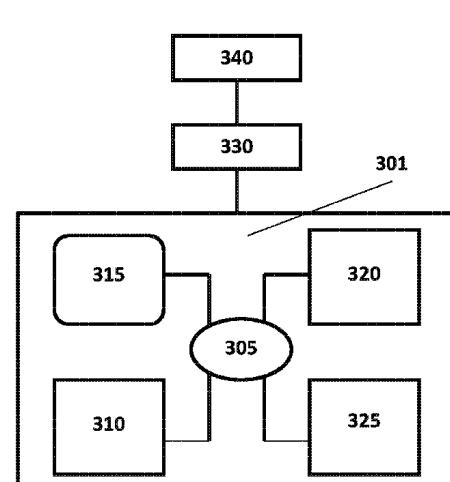
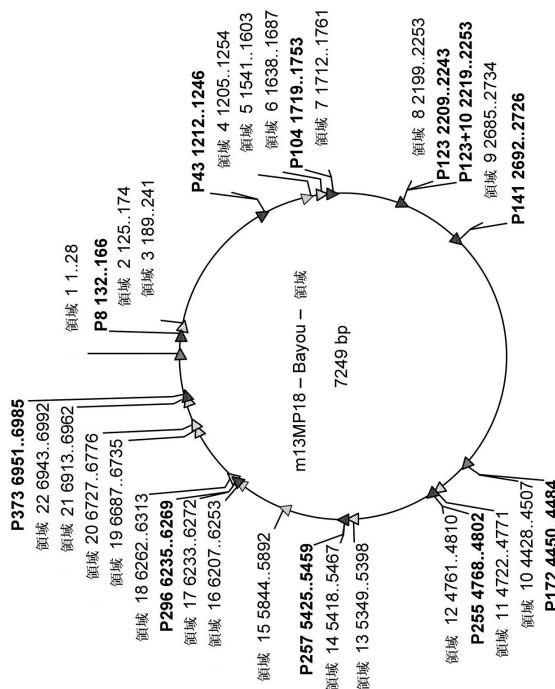


FIG. 3

50

【図 4】



【図 5】

プローブの名稱 / 配列	Gc 含量	Tm	ハイビンクTm	開始	スケール	特異	5'修飾	3'修飾
M13P96 5'-GAACTCGGGGAACTGGGGAACTCTTAACTGAC NtP -4213-5'-GAACTCGGGAACTGGGGAACTCTTAACTGAC	65.7	69.8	63.5	250mmole	PAGE	5'-アツコ修飾NtP-Al2A		
M13P96A 5'-GAACTCGGGAACTGGGGAACTCTTAACTGAC NtP -4213-5'-GAACTCGGGAACTGGGGAACTCTTAACTGAC	65.7	69.8	63.5	1umole	IPG/PAGE	5'-アツコ修飾NtP-Al2A	3'G,N	
M13P96C 5'-GAACTCGGGAACTGGGGAACTCTTAACTGAC O4-5'-GTCGTCCTTAACTGGGGAACTCTTAACTGAC	65.7	69.8	63.5	250mmole	PAGE	5'-アツコ修飾NtP-Al2A		
M13P96C3 5'-GAACTCGGGAACTGGGGAACTCTTAACTGAC O4-5'-GTCGTCCTTAACTGGGGAACTCTTAACTGAC	65.7	69.8	63.5	1umole	IPG/PAGE	5'-アツコ修飾NtP-Al2A	3'G,N	
M13P14A NtP -4213-5'-GAACTCGGGAACTGGGGAACTCTTAACTGAC NtP -4213-5'-GAACTCGGGAACTGGGGAACTCTTAACTGAC	51.4	67	59.6	54.5	1umole	IPG/PAGE	5'-アツコ修飾NtP-Al2A	
M13P13A 5'-GAACTCGGGAACTGGGGAACTCTTAACTGAC NtP -4213-5'-GAACTCGGGAACTGGGGAACTCTTAACTGAC	45.7	61.9	55.1	69.5	250mmole	PAGE	5'-アツコ修飾NtP-Al2A	
M13P17A 5'-GAACTCGGGAACTGGGGAACTCTTAACTGAC NtP -4213-5'-GAACTCGGGAACTGGGGAACTCTTAACTGAC	45.7	61.9	55.1	69.5	1umole	IPG/PAGE	5'-アツコ修飾NtP-Al2A	3'G,N
M13P17C 5'-GAACTCGGGAACTGGGGAACTCTTAACTGAC O4-5'-GTCGTCCTTAACTGGGGAACTCTTAACTGAC	45.7	61.9	55.1	69.5	250mmole	PAGE	5'-アツコ修飾NtP-Al2A	3'G,N
M13P21A NtP -4213-5'-GAACTCGGGAACTGGGGAACTCTTAACTGAC NtP -4213-5'-GAACTCGGGAACTGGGGAACTCTTAACTGAC	40	61.3	57	69.8	1umole	IPG/PAGE	5'-アツコ修飾NtP-Al2A	
M13P22A 5'-GCTTAACTGAACTTATTTTGATGAA NtP -4213-5'-GCTTAACTGAACTTATTTTGATGAA	25.7	55.6	16.1	20.9	250mmole	PAGE	5'-アツコ修飾NtP-Al2A	
M13P22A 5'-GCTTAACTGAACTTATTTTGATGAA NtP -4213-5'-GCTTAACTGAACTTATTTTGATGAA	25.7	55.6	16.1	20.9	1umole	IPG/PAGE	5'-アツコ修飾NtP-Al2A	3'G,N
M13P22C 5'-GCTTAACTGAACTTATTTTGATGAA O4-5'-TGAATCTTAACTGAACTTATTTTGATGAA	25.7	55.6	16.1	20.9	250mmole	PAGE	5'-アツコ修飾NtP-Al2A	3'G,N
M13P22C3 5'-GCTTAACTGAACTTATTTTGATGAA O4-5'-TGAATCTTAACTGAACTTATTTTGATGAA	25.7	55.6	16.1	20.9	1umole	IPG/PAGE	5'-アツコ修飾NtP-Al2A	3'G,N
M13P23 5'-GCTTAACTGAACTTATTTTGATGAA NtP -4213-5'-GCTTAACTGAACTTATTTTGATGAA	25.7	55.6	16.1	20.9	1umole	IPG/PAGE	5'-アツコ修飾NtP-Al2A	
M13P23A 5'-GCTTAACTGAACTTATTTTGATGAA NtP -4213-5'-GCTTAACTGAACTTATTTTGATGAA	34.3	59.7	44.1	132	1umole	IPG/PAGE	5'-アツコ修飾NtP-Al2A	
M13P23A 5'-GCTTAACTGAACTTATTTTGATGAA NtP -4213-5'-GCTTAACTGAACTTATTTTGATGAA	31.1	59.2	26.7	112	1umole	IPG/PAGE	5'-アツコ修飾NtP-Al2A	
M13P23A 5'-GCTTAACTGAACTTATTTTGATGAA NtP -4213-5'-GCTTAACTGAACTTATTTTGATGAA	34.3	57.8	41.5	4450	1umole	IPG/PAGE	5'-アツコ修飾NtP-Al2A	

【図 6】

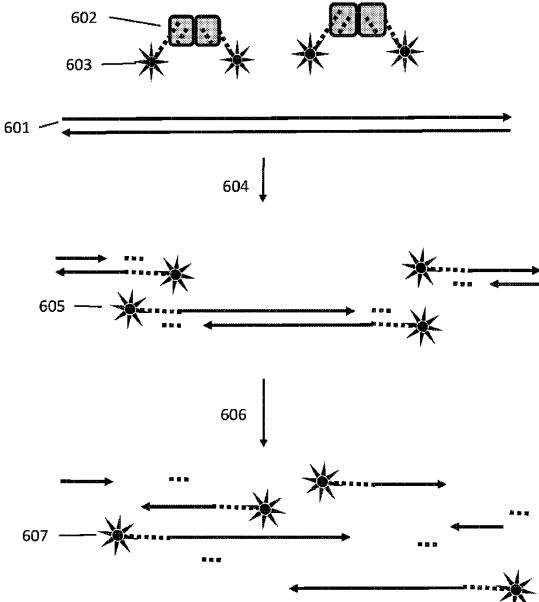


FIG. 6

【図 7 A】

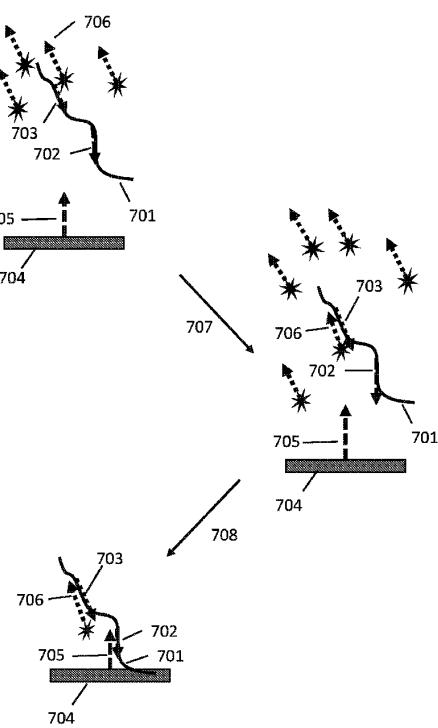


FIG. 7A

【図 7 B】

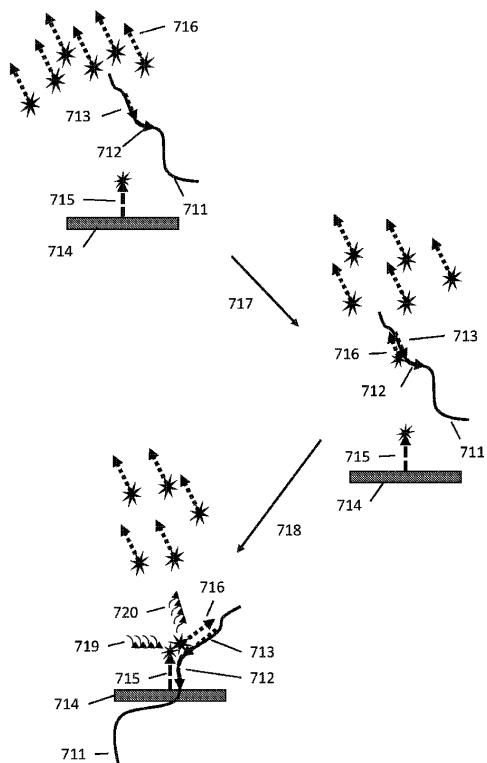


FIG. 7B

【図 7 C】

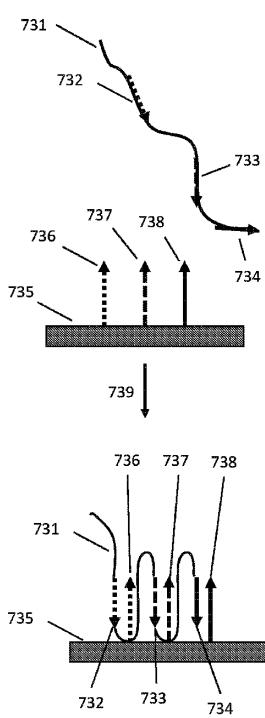


FIG. 7C

10

20

30

40

【図 8】

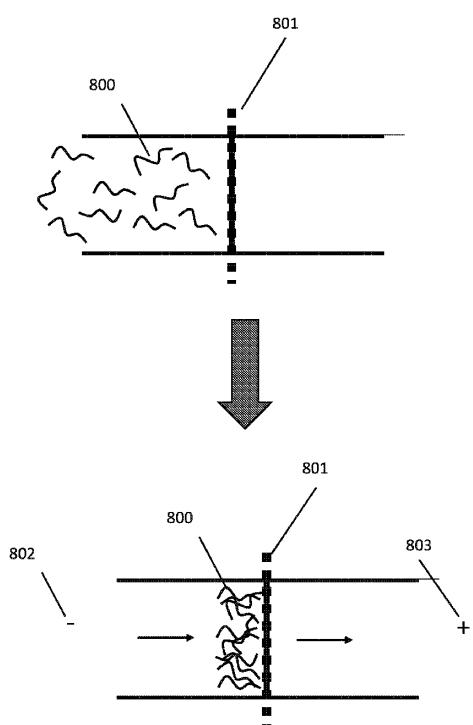
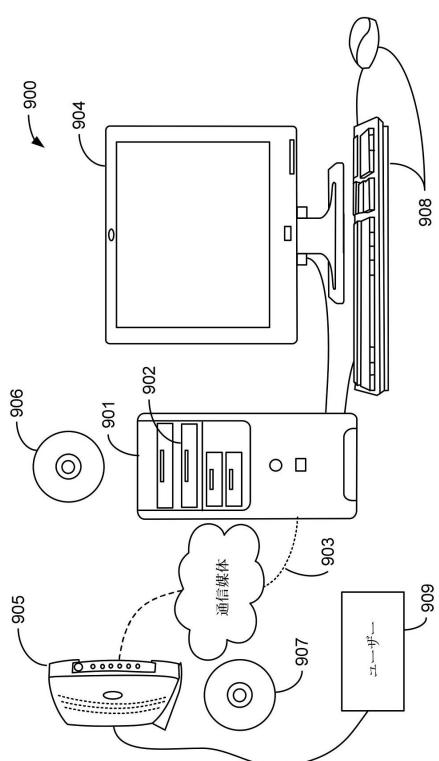


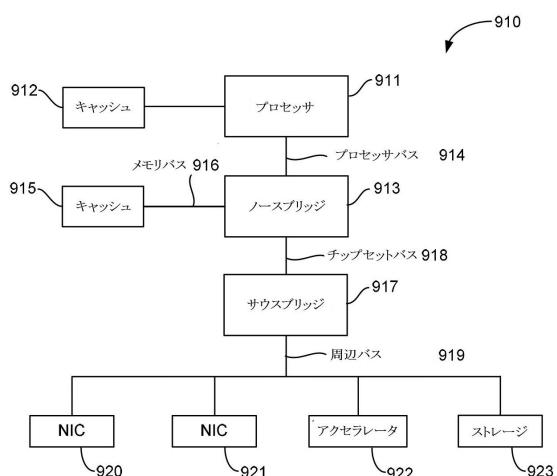
FIG. 8

【図 9 A】

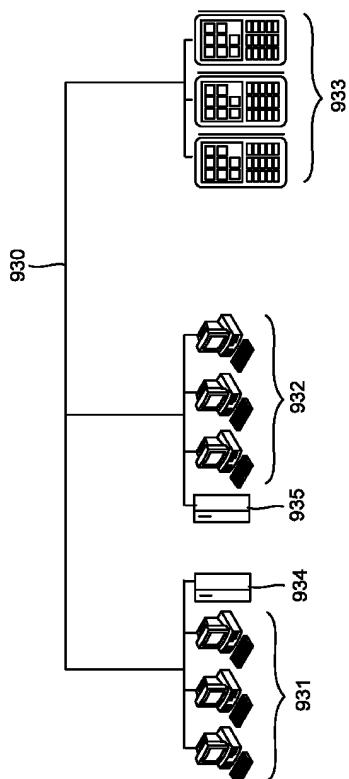


50

【図 9 B】



【図 9 C】

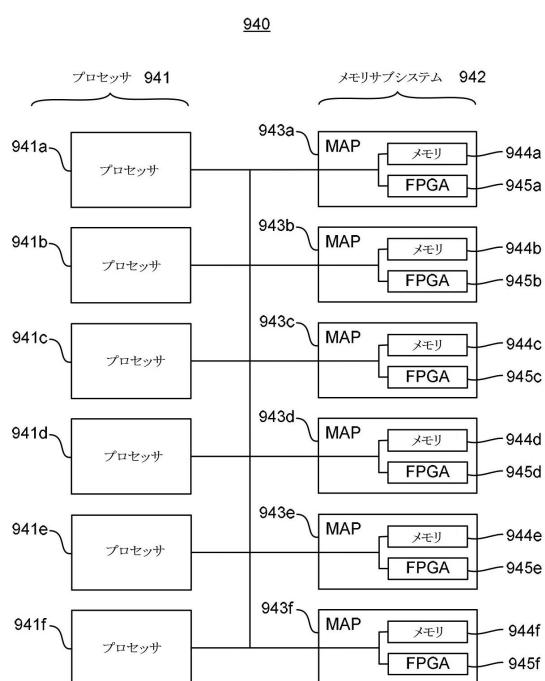


10

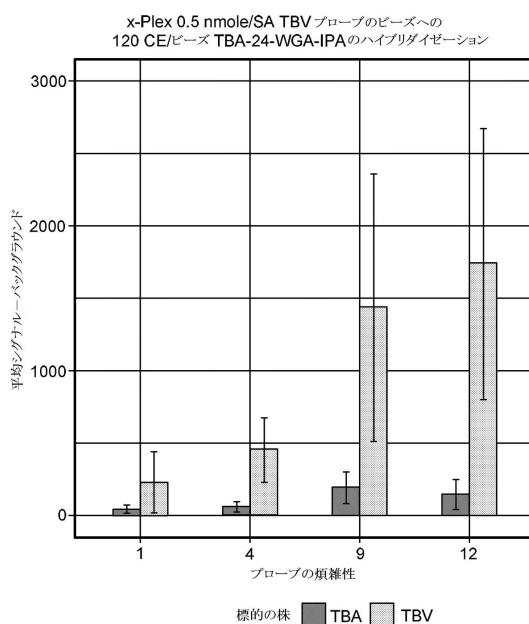
FIG. 9C

20

【図 9 D】



【図 10】



30

40

50

フロントページの続き

米国(US)

(31)優先権主張番号 62/256,049

(32)優先日 平成27年11月16日(2015.11.16)

(33)優先権主張国・地域又は機関

米国(US)

ニア州 ランチョ・サンタフェ ピーオー ボックス 1388

合議体

審判長 長井 啓子

審判官 福井 悟

審判官 飯室 里美

(56)参考文献 特表2005-509127 (JP, A)

特表2001-514906 (JP, A)

Anal Biochem. 2005, 342, pp. 59-68

(58)調査した分野 (Int.Cl., DB名)

C12N15/00-15/90

JST7580 / JSTPlus / JMEDPlus (JDream3)

PubMed

BIOSIS / CAPLUS / MEDLINE / WPIDS (STN)