

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2021-526857

(P2021-526857A)

(43) 公表日 令和3年10月11日(2021.10.11)

(51) Int.Cl.	F I	テーマコード (参考)
C 1 2 Q 1/6809 (2018.01)	C 1 2 Q 1/6809 Z	4 B O 6 3
C 1 2 Q 1/6844 (2018.01)	C 1 2 Q 1/6844 Z	
C 4 O B 20/04 (2006.01)	C 4 O B 20/04	
C 1 2 Q 1/6869 (2018.01)	C 1 2 Q 1/6869 Z	

審査請求 未請求 予備審査請求 未請求 (全 35 頁)

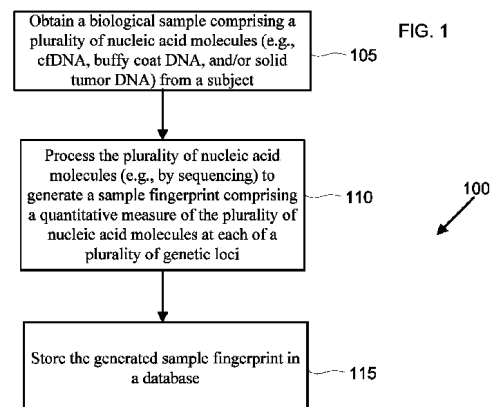
(21) 出願番号	特願2021-518049 (P2021-518049)	(71) 出願人	520478530 レクセント バイオ, インコーポレイテッド アメリカ合衆国 マサチューセッツ 02141, ケンブリッジ, セカンド ストリート 150
(86) (22) 出願日	令和1年6月6日 (2019.6.6)	(74) 代理人	100078282 弁理士 山本 秀策
(85) 翻訳文提出日	令和3年2月3日 (2021.2.3)	(74) 代理人	100113413 弁理士 森下 夏樹
(86) 国際出願番号	PCT/US2019/035871	(74) 代理人	100181674 弁理士 飯田 貴敏
(87) 国際公開番号	W02019/236906	(74) 代理人	100181641 弁理士 石川 大輔
(87) 国際公開日	令和1年12月12日 (2019.12.12)		
(31) 優先権主張番号	62/681, 642		
(32) 優先日	平成30年6月6日 (2018.6.6)		
(33) 優先権主張国・地域又は機関	米国 (US)		

最終頁に続く

(54) 【発明の名称】 生体試料のフィンガープリンティングのための方法

(57) 【要約】

本開示は、試料不一致を同定するための方法であって、被験体から第1の複数の核酸分子を含む第1の生体試料を得るステップ；第1の複数の核酸分子を処理して、複数の遺伝子座の各々における第1の複数の核酸分子の定量的尺度を含む第1の試料フィンガープリントを生成するステップであって、複数の遺伝子座が、常染色体一塩基多型 (SNP) を含む、ステップ；被験体から第2の複数の核酸分子を含む第2の生体試料を得るステップ；第2の複数の核酸分子を処理して、複数の遺伝子座の各々における第2の複数の核酸分子の定量的尺度を含む第2の試料フィンガープリントを生成するステップ；第1の試料フィンガープリントと第2の試料フィンガープリントとの間の差異を決定するステップ；および差異が所定の基準を満たす場合に、試料不一致を同定するステップを含む、方法を提供する。



【特許請求の範囲】

【請求項 1】

試料不一致を同定するための方法であって、

被験体から第 1 の複数の核酸分子を含む第 1 の生体試料を得るステップ；

コンピュータによって前記第 1 の複数の核酸分子を処理して、複数の遺伝子座の各々における前記第 1 の複数の核酸分子の定量的尺度を含む第 1 の試料フィンガープリントを生成するステップであって、前記複数の遺伝子座が、常染色体一塩基多型（SNP）を含む、ステップ；

前記被験体から第 2 の複数の核酸分子を含む第 2 の生体試料を得るステップ；

コンピュータによって前記第 2 の複数の核酸分子を処理して、前記複数の遺伝子座の各々における前記第 2 の複数の核酸分子の定量的尺度を含む第 2 の試料フィンガープリントを生成するステップ；

前記第 1 の試料フィンガープリントと前記第 2 の試料フィンガープリントとの間の差異を決定するステップ；および

前記第 1 の試料フィンガープリントと前記第 2 の試料フィンガープリントとの間の前記差異が所定の閾値を超える場合に、前記試料不一致を同定するステップを含み、

前記第 1 の複数の核酸分子の前記定量的尺度が、前記第 1 の複数の核酸分子の 12 個以下の独立した測定値を含む、方法。

【請求項 2】

試料不一致を同定するための方法であって、

被験体から第 1 の複数の核酸分子を含む第 1 の生体試料を得るステップ；

コンピュータによって前記第 1 の複数の核酸分子を処理して、複数の遺伝子座の各々における前記第 1 の複数の核酸分子の定量的尺度を含む第 1 の試料フィンガープリントを生成するステップであって、前記複数の遺伝子座が、常染色体一塩基多型（SNP）を含む、ステップ；

前記被験体から第 2 の複数の核酸分子を含む第 2 の生体試料を得るステップ；

コンピュータによって前記第 2 の複数の核酸分子を処理して、前記複数の遺伝子座の各々における前記第 2 の複数の核酸分子の定量的尺度を含む第 2 の試料フィンガープリントを生成するステップ；

前記第 1 の試料フィンガープリントと前記第 2 の試料フィンガープリントとの間の差異を決定するステップ；および

前記第 1 の試料フィンガープリントと前記第 2 の試料フィンガープリントとの間の前記差異が所定の閾値を超える場合に、前記試料不一致を同定するステップを含み、

前記常染色体一塩基多型が、単純な一塩基多型を含む、方法。

【請求項 3】

試料不一致を同定するための方法であって、

被験体から第 1 の複数の核酸分子を含む第 1 の生体試料を得るステップ；

コンピュータによって前記第 1 の複数の核酸分子を処理して、複数の遺伝子座の各々における前記第 1 の複数の核酸分子の定量的尺度を含む第 1 の試料フィンガープリントを生成するステップであって、前記複数の遺伝子座が、常染色体一塩基多型（SNP）を含む、ステップ；

前記被験体から第 2 の複数の核酸分子を含む第 2 の生体試料を得るステップ；

コンピュータによって前記第 2 の複数の核酸分子を処理して、前記複数の遺伝子座の各々における前記第 2 の複数の核酸分子の定量的尺度を含む第 2 の試料フィンガープリントを生成するステップ；

前記第 1 の試料フィンガープリントと前記第 2 の試料フィンガープリントとの間の差異を決定するステップ；および

前記第 1 の試料フィンガープリントと前記第 2 の試料フィンガープリントとの間の前記

差異が所定の閾値を超える場合に、前記試料不一致を同定するステップを含み、

前記常染色体一塩基多型が、所定の閾値を超えるマイナー対立遺伝子割合を有する、方法。

【請求項 4】

前記常染色体一塩基多型が、約 7.5 % を超えるマイナー対立遺伝子割合を有する、請求項 3 に記載の方法。

【請求項 5】

前記第 1 の複数の核酸分子および前記第 2 の複数の核酸分子が、無細胞 DNA (cfDNA) を含む、請求項 1 から 4 のいずれかに記載の方法。

10

【請求項 6】

前記第 1 の複数の核酸分子および前記第 2 の複数の核酸分子が、パフィーコート DNA を含む、請求項 1 から 4 のいずれかに記載の方法。

【請求項 7】

前記第 1 の複数の核酸分子および前記第 2 の複数の核酸分子が、固形腫瘍 DNA を含む、請求項 1 から 4 のいずれかに記載の方法。

【請求項 8】

前記第 2 の生体試料が、前記第 1 の生体試料を得るステップの後の時点で前記被験体から得られる、請求項 1 から 4 のいずれかに記載の方法。

【請求項 9】

前記第 1 の複数の核酸分子を処理するステップが、前記第 1 の複数の核酸分子を配列決定して、第 1 の複数の配列決定読み取りを生成する工程を含み、前記第 2 の複数の核酸分子を処理するステップが、前記第 2 の複数の核酸分子を配列決定して、第 2 の複数の配列決定読み取りを生成する工程を含む、請求項 1 から 4 のいずれかに記載の方法。

20

【請求項 10】

前記配列決定する工程が、全ゲノム配列決定 (WGS) を含む、請求項 9 に記載の方法。

【請求項 11】

前記配列決定する工程が、約 $10 \times$ 以下の深度で実施される、請求項 10 に記載の方法。

30

【請求項 12】

前記配列決定する工程が、約 $8 \times$ 以下の深度で実施される、請求項 10 に記載の方法。

【請求項 13】

前記配列決定する工程が、約 $6 \times$ 以下の深度で実施される、請求項 10 に記載の方法。

【請求項 14】

前記第 1 の複数の核酸分子の前記定量的尺度が、前記複数の遺伝子座の各々における前記第 1 の複数の核酸分子のカバレッジを含み、前記第 2 の複数の核酸分子の前記定量的尺度が、前記複数の遺伝子座の各々における前記第 2 の複数の核酸分子のカバレッジを含む、請求項 9 に記載の方法。

【請求項 15】

前記第 1 の複数の核酸分子を処理するステップが、前記第 1 の複数の核酸分子の結合測定を実施する工程を含み、前記第 2 の複数の核酸分子を処理するステップが、前記第 2 の複数の核酸分子の結合測定を実施する工程を含む、請求項 1 から 4 のいずれかに記載の方法。

40

【請求項 16】

前記複数の遺伝子座の各々における前記第 1 の複数の核酸分子の前記定量的尺度が、前記遺伝子座を含有する前記第 1 の複数の核酸分子の数を含み、前記複数の遺伝子座の各々における前記第 2 の複数の核酸分子の前記定量的尺度が、前記遺伝子座を含有する前記第 2 の複数の核酸分子の数を含む、請求項 15 に記載の方法。

【請求項 17】

50

前記複数の遺伝子座の少なくとも一部分について、前記第1の複数の核酸分子および／または前記第2の複数の核酸分子を富化するステップをさらに含む、請求項1から16のいずれかに記載の方法。

【請求項18】

前記富化が、前記第1の複数の核酸分子および／または前記第2の複数の核酸分子の少なくとも一部分を増幅する工程を含む、請求項17に記載の方法。

【請求項19】

前記増幅が、選択的増幅を含む、請求項18に記載の方法。

【請求項20】

前記増幅が、普遍的増幅を含む、請求項18に記載の方法。

10

【請求項21】

前記富化が、前記第1の複数の核酸分子および／または前記第2の複数の核酸分子の少なくとも一部分を選択的に単離する工程を含む、請求項17に記載の方法。

【請求項22】

前記複数の遺伝子座が、少なくとも約50の別個の常染色体一塩基多型（SNP）を含む、請求項1から4のいずれかに記載の方法。

【請求項23】

前記複数の遺伝子座が、少なくとも約100の別個の常染色体一塩基多型（SNP）を含む、請求項1から4のいずれかに記載の方法。

【請求項24】

20

前記第1の試料フィンガープリントを生成するステップが、前記被験体から第3の複数の核酸分子を含む第3の生体試料を得る工程、および前記第3の複数の核酸分子を処理して、第2の複数の遺伝子座の各々における前記第3の複数の核酸分子の定量的尺度を得る工程をさらに含む、前記第2の複数の遺伝子座が、常染色体一塩基多型（SNP）を含み、前記第2の試料フィンガープリントを生成するステップが、前記被験体から第4の複数の核酸分子を含む第4の生体試料を得る工程、および前記第4の複数の核酸分子を処理して、前記第2の複数の遺伝子座の各々における前記第4の複数の核酸分子の定量的尺度を得る工程をさらに含む、請求項1から4のいずれかに記載の方法。

【請求項25】

前記第3の複数の核酸分子および前記第4の複数の核酸分子が、無細胞DNA（cfDNA）を含む、請求項24に記載の方法。

30

【請求項26】

前記第3の複数の核酸分子および前記第4の複数の核酸分子が、バフィーコートDNAを含む、請求項24に記載の方法。

【請求項27】

前記第3の複数の核酸分子および前記第4の複数の核酸分子が、固形腫瘍DNAを含む、請求項24に記載の方法。

【請求項28】

前記第1の試料フィンガープリントを生成するステップが、前記被験体から第5の複数の核酸分子を含む第5の生体試料を得る工程、および前記第5の複数の核酸分子を処理して、第3の複数の遺伝子座の各々における前記第5の複数の核酸分子の定量的尺度を得る工程をさらに含む、前記第3の複数の遺伝子座が、常染色体一塩基多型（SNP）を含み、前記第2の試料フィンガープリントを生成するステップが、前記被験体から第6の複数の核酸分子を含む第6の生体試料を得る工程、および前記第6の複数の核酸分子を処理して、前記第3の複数の遺伝子座の各々における前記第6の複数の核酸分子の定量的尺度を得る工程をさらに含む、請求項24に記載の方法。

40

【請求項29】

前記第3の複数の核酸分子および前記第4の複数の核酸分子が、無細胞DNA（cfDNA）を含む、請求項28に記載の方法。

【請求項30】

50

前記第 3 の複数の核酸分子および前記第 4 の複数の核酸分子が、パフィーコート DNA を含む、請求項 28 に記載の方法。

【請求項 31】

前記第 3 の複数の核酸分子および前記第 4 の複数の核酸分子が、固形腫瘍 DNA を含む、請求項 28 に記載の方法。

【請求項 32】

少なくとも約 90 % の感度で前記試料不一致を同定するステップを含む、請求項 1 から 31 のいずれかに記載の方法。

【請求項 33】

少なくとも約 90 % の特異度で前記試料不一致を同定するステップを含む、請求項 1 から 31 のいずれかに記載の方法。

10

【請求項 34】

少なくとも約 90 % の陽性的中率 (PPV) で前記試料不一致を同定するステップを含む、請求項 1 から 31 のいずれかに記載の方法。

【請求項 35】

少なくとも約 90 % の陰性的中率 (NPV) で前記試料不一致を同定するステップを含む、請求項 1 から 31 のいずれかに記載の方法。

【請求項 36】

少なくとも約 0.90 の曲線下面積 (AUC) で前記試料不一致を同定するステップを含む、請求項 1 から 31 のいずれかに記載の方法。

20

【請求項 37】

所定の基準が、前記差異が、遺伝子型類似性における所定の閾値よりも大きい差異を含むことである、請求項 1 から 31 のいずれかに記載の方法。

【請求項 38】

前記所定の閾値が約 0.8 である、請求項 37 に記載の方法。

【請求項 39】

同定された前記試料不一致に基づいて、前記第 2 の生体試料をさらなるアッセイから排除するステップをさらに含む、請求項 1 から 38 のいずれかに記載の方法。

【請求項 40】

前記第 1 の試料フィンガープリントと前記第 2 の試料フィンガープリントとの間の前記差異が所定の基準を満たさない場合に、試料一致を同定するステップをさらに含む、請求項 1 から 4 のいずれかに記載の方法。

30

【請求項 41】

少なくとも約 90 % の感度で前記試料一致を同定するステップを含む、請求項 40 に記載の方法。

【請求項 42】

少なくとも約 90 % の特異度で前記試料一致を同定するステップを含む、請求項 40 に記載の方法。

【請求項 43】

少なくとも約 90 % の陽性的中率 (PPV) で前記試料一致を同定するステップを含む、請求項 40 に記載の方法。

40

【請求項 44】

少なくとも約 90 % の陰性的中率 (NPV) で前記試料一致を同定するステップを含む、請求項 40 に記載の方法。

【請求項 45】

少なくとも約 0.90 の曲線下面積 (AUC) で前記試料一致を同定するステップを含む、請求項 40 に記載の方法。

【請求項 46】

同定された前記試料一致に基づいて、前記第 2 の生体試料をさらなるアッセイに供するステップをさらに含む、請求項 40 から 45 のいずれかに記載の方法。

50

【請求項 47】

同定された前記試料一致に基づいて、データベース中に前記第2の試料フィンガープリントを記憶するステップ、および必要に応じて前記データベース中に前記第1の試料フィンガープリントを記憶するステップをさらに含む、請求項40から45のいずれかに記載の方法。

【請求項 48】

1つまたは複数のコンピュータプロセッサによる実行の際に試料不一致を同定するための方法を実装する機械実行可能コードを含む非一過性コンピュータ可読媒体であって、前記方法が、

複数の遺伝子座の各々における第1の生体試料の第1の複数の核酸分子の定量的尺度を含む第1の試料フィンガープリントの情報を受け取るステップであって、前記複数の遺伝子座が、常染色体一塩基多型（SNP）を含み、前記第1の複数の核酸分子の前記定量的尺度が、前記複数の核酸分子の12個以下の独立した測定値を含む、ステップ；

前記複数の遺伝子座の各々における第2の生体試料の第2の複数の核酸分子の定量的尺度を含む第2の試料フィンガープリントの情報を受け取るステップであって、前記第2の生体試料が、被験体から得られる、ステップ；

前記第1の試料フィンガープリントと前記第2の試料フィンガープリントとの間の差異を決定するステップ；および

前記第1の試料フィンガープリントと前記第2の試料フィンガープリントとの間の前記差異が所定の基準を満たす場合に、前記試料不一致を同定するステップを含む、非一過性コンピュータ可読媒体。

【発明の詳細な説明】**【技術分野】****【0001】**

相互参照

本出願は、参照により本明細書に完全に組み込まれる、2018年6月6日出願の米国仮特許出願第62/681,642号、表題「METHODS FOR FINGERPRINTING OF BIOLOGICAL SAMPLES」の利益を主張する。

【背景技術】**【0002】**

被験体から得られた生体試料の収集およびアッセイは、臨床および実験室プロセスを通じた試料同一性の信頼性のある維持に関する課題にしばしば遭遇し得る。例えば、生体試料は、実験室または臨床設定において不注意にしばしば取り違えられ得、それによって、検出されず修正されないままだと、潜在的に不正確な臨床結果を生じる。

【発明の概要】**【発明が解決しようとする課題】****【0003】**

遺伝子座のパネルを使用して生体試料をフィンガープリンティングするための方法は、所望の感度、特異度または精度で遺伝情報を得るために、十分に深いカバレッジ（coverage）を要求し得る。例えば、深いカバレッジが、異なる試料から生成されたフィンガープリント間を識別するために十分に高いシグナル・ノイズ比（SNR）のために要求され得る。かかる試料は、（例えば、2つの異なる時点において同じ被験体から得られた）長期的試料であり得る。ローパス配列決定を使用して処理される長期的試料は、以下に関する課題に遭遇し得る：（1）異なる時点からの一致する試料について修正を行うこと、および（2）いずれか1つの位置における比較的低い読み取りカバレッジにもかかわらず試料フィンガープリンティングに適切な遺伝子座のパネルを同定すること。

【課題を解決するための手段】**【0004】**

生体試料のフィンガープリントを生成し比較するための方法およびシステムが提供される。試料フィンガープリントは、1つまたは複数の時点の各々において被験体から得られ

10

20

30

40

50

た生体試料由来の核酸分子の1つまたは複数のセットを配列決定することによって生成され得る。試料フィンガープリントのペアワイズ比較は、試料不一致（例えば、2つの試料が異なる被験体から得られたこと）または試料一致（例えば、2つの試料が同じ被験体から得られたこと）が、試料フィンガープリントが生成された2つの生体試料間に存在するかどうかを決定するために実施され得る。

【0005】

ある態様では、本開示は、試料不一致を同定するための方法であって、被験体から第1の複数の核酸分子を含む第1の生体試料を得るステップ；コンピュータによって第1の複数の核酸分子を処理して、複数の遺伝子座の各々における第1の複数の核酸分子の定量的尺度を含む第1の試料フィンガープリントを生成するステップであって、複数の遺伝子座が、常染色体一塩基多型（SNP）を含む、ステップ；被験体から第2の複数の核酸分子を含む第2の生体試料を得るステップ；コンピュータによって第2の複数の核酸分子を処理して、複数の遺伝子座の各々における第2の複数の核酸分子の定量的尺度を含む第2の試料フィンガープリントを生成するステップ；第1の試料フィンガープリントと第2の試料フィンガープリントとの間の差異を決定するステップ；および第1の試料フィンガープリントと第2の試料フィンガープリントとの間の差異が所定の閾値を超える場合に、試料不一致を同定するステップを含む、方法を提供する。さらに、この態様では、第1の複数の核酸分子の定量的尺度は、第1の複数の核酸分子の12個以下の独立した尺度を含む。

【0006】

別の態様では、本開示は、試料不一致を同定するための方法であって、被験体から第1の複数の核酸分子を含む第1の生体試料を得るステップ；コンピュータによって第1の複数の核酸分子を処理して、複数の遺伝子座の各々における第1の複数の核酸分子の定量的尺度を含む第1の試料フィンガープリントを生成するステップであって、複数の遺伝子座が、常染色体一塩基多型（SNP）を含む、ステップ；被験体から第2の複数の核酸分子を含む第2の生体試料を得るステップ；コンピュータによって第2の複数の核酸分子を処理して、複数の遺伝子座の各々における第2の複数の核酸分子の定量的尺度を含む第2の試料フィンガープリントを生成するステップ；第1の試料フィンガープリントと第2の試料フィンガープリントとの間の差異を決定するステップ；および第1の試料フィンガープリントと第2の試料フィンガープリントとの間の差異が所定の閾値を超える場合に、試料不一致を同定するステップを含む、方法を提供する。さらに、この態様では、常染色体一塩基多型は、単純な一塩基多型を含む。

【0007】

別の態様では、本開示は、試料不一致を同定するための方法であって、被験体から第1の複数の核酸分子を含む第1の生体試料を得るステップ；コンピュータによって第1の複数の核酸分子を処理して、複数の遺伝子座の各々における第1の複数の核酸分子の定量的尺度を含む第1の試料フィンガープリントを生成するステップであって、複数の遺伝子座が、常染色体一塩基多型（SNP）を含む、ステップ；被験体から第2の複数の核酸分子を含む第2の生体試料を得るステップ；コンピュータによって第2の複数の核酸分子を処理して、複数の遺伝子座の各々における第2の複数の核酸分子の定量的尺度を含む第2の試料フィンガープリントを生成するステップ；第1の試料フィンガープリントと第2の試料フィンガープリントとの間の差異を決定するステップ；および第1の試料フィンガープリントと第2の試料フィンガープリントとの間の差異が所定の閾値を超える場合に、試料不一致を同定するステップを含む、方法を提供する。さらに、この態様では、常染色体一塩基多型は、所定の閾値を超えるマイナー対立遺伝子割合を有する。常染色体一塩基多型が特定の閾値を超えるマイナー対立遺伝子割合を有する一部の実施形態では、常染色体一塩基多型は、約7.5%を超えるマイナー対立遺伝子割合を有する。

【0008】

一部の実施形態では、第1の複数の核酸分子および第2の複数の核酸分子は、無細胞DNA（cfDNA）を含む。一部の実施形態では、第1の複数の核酸分子および第2の複数の核酸分子は、バフィーコートDNAを含む。一部の実施形態では、第1の複数の核酸

分子および第 2 の複数の核酸分子は、固形腫瘍 DNA を含む。

【0009】

一部の実施形態では、第 2 の生体試料は、第 1 の生体試料を得るステップの後の時点で被験体から得られる。一部の実施形態では、第 1 の複数の核酸分子を処理するステップは、第 1 の複数の核酸分子を配列決定して、第 1 の複数の配列決定読み取りを生成する工程を含み、第 2 の複数の核酸分子を処理するステップは、第 2 の複数の核酸分子を配列決定して、第 2 の複数の配列決定読み取りを生成する工程を含む。

【0010】

一部の実施形態では、配列決定する工程は、全ゲノム配列決定 (WGS) を含む。一部の実施形態では、配列決定する工程は、約 $10\times$ 以下の深度で実施される。一部の実施形態では、配列決定する工程は、約 $8\times$ 以下の深度で実施される。一部の実施形態では、配列決定する工程は、約 $6\times$ 以下の深度で実施される。一部の実施形態では、第 1 の複数の核酸分子の定量的尺度は、複数の遺伝子座の各々における第 1 の複数の核酸分子のカバレッジを含み、第 2 の複数の核酸分子の定量的尺度は、複数の遺伝子座の各々における第 2 の複数の核酸分子のカバレッジを含む。

【0011】

一部の実施形態では、第 1 の複数の核酸分子を処理するステップは、第 1 の複数の核酸分子の結合測定を実施する工程を含み、第 2 の複数の核酸分子を処理するステップは、第 2 の複数の核酸分子の結合測定を実施する工程を含む。一部の実施形態では、複数の遺伝子座の各々における第 1 の複数の核酸分子の定量的尺度は、遺伝子座を含有する第 1 の複数の核酸分子の数を含み、複数の遺伝子座の各々における第 2 の複数の核酸分子の定量的尺度は、遺伝子座を含有する第 2 の複数の核酸分子の数を含む。

【0012】

一部の実施形態では、この方法は、複数の遺伝子座の少なくとも一部分について、第 1 の複数の核酸分子および／または第 2 の複数の核酸分子を富化するステップをさらに含む。一部の実施形態では、富化は、第 1 の複数の核酸分子および／または第 2 の複数の核酸分子の少なくとも一部分を増幅する工程を含む。一部の実施形態では、増幅は、選択的増幅を含む。一部の実施形態では、増幅は、普遍的増幅を含む。一部の実施形態では、富化は、第 1 の複数の核酸分子および／または第 2 の複数の核酸分子の少なくとも一部分を選択的に単離する工程を含む。

【0013】

一部の実施形態では、複数の遺伝子座は、少なくとも約 50 の別個の常染色体一塩基多型 (SNP) を含む。一部の実施形態では、複数の遺伝子座は、少なくとも約 100 の別個の常染色体一塩基多型 (SNP) を含む。

【0014】

一部の実施形態では、第 1 の試料フィンガープリントを生成するステップは、被験体から第 3 の複数の核酸分子を含む第 3 の生体試料を得る工程、および第 3 の複数の核酸分子を処理して、第 2 の複数の遺伝子座の各々における第 3 の複数の核酸分子の定量的尺度を得る工程をさらに含み、第 2 の複数の遺伝子座は、常染色体一塩基多型 (SNP) を含み、第 2 の試料フィンガープリントを生成するステップが、被験体から第 4 の複数の核酸分子を含む第 4 の生体試料を得る工程、および第 4 の複数の核酸分子を処理して、第 2 の複数の遺伝子座の各々における第 4 の複数の核酸分子の定量的尺度を得る工程をさらに含む。

【0015】

一部の実施形態では、第 3 の複数の核酸分子および第 4 の複数の核酸分子は、無細胞 DNA (cfDNA) を含む。一部の実施形態では、第 3 の複数の核酸分子および第 4 の複数の核酸分子は、パフィーコート DNA を含む。一部の実施形態では、第 3 の複数の核酸分子および第 4 の複数の核酸分子は、固形腫瘍 DNA を含む。一部の実施形態では、第 1 の試料フィンガープリントを生成するステップは、被験体から第 5 の複数の核酸分子を含む第 5 の生体試料を得る工程、および第 5 の複数の核酸分子を処理して、第 3 の複数の遺

伝子座の各々における第5の複数の核酸分子の定量的尺度を得る工程をさらに含み、第3の複数の遺伝子座は、常染色体一塩基多型（SNP）を含み、第2の試料フィンガープリントを生成するステップは、被験体から第6の複数の核酸分子を含む第6の生体試料を得る工程、および第6の複数の核酸分子を処理して、第3の複数の遺伝子座の各々における第6の複数の核酸分子の定量的尺度を得る工程をさらに含む。

【0016】

一部の実施形態では、第3の複数の核酸分子および第4の複数の核酸分子は、無細胞DNA（cfDNA）を含む。一部の実施形態では、第3の複数の核酸分子および第4の複数の核酸分子は、パフィーコートDNAを含む。一部の実施形態では、第3の複数の核酸分子および第4の複数の核酸分子は、固形腫瘍DNAを含む。

10

【0017】

一部の実施形態では、この方法は、少なくとも約90%の感度で試料不一致を同定するステップを含む。一部の実施形態では、試料不一致を同定するステップは、少なくとも約95%の感度で実施される。一部の実施形態では、この方法は、少なくとも約99%の感度で試料不一致を同定するステップを含む。

【0018】

一部の実施形態では、この方法は、少なくとも約90%の特異度で試料不一致を同定するステップを含む。一部の実施形態では、この方法は、少なくとも約95%の特異度で試料不一致を同定するステップを含む。一部の実施形態では、この方法は、少なくとも約99%の特異度で試料不一致を同定するステップを含む。

20

【0019】

一部の実施形態では、この方法は、少なくとも約90%の陽性的中率（positive predictive value）（PPV）で試料不一致を同定するステップを含む。一部の実施形態では、この方法は、少なくとも約95%の陽性的中率（PPV）で試料不一致を同定するステップを含む。一部の実施形態では、この方法は、少なくとも約99%の陽性的中率（PPV）で試料不一致を同定するステップを含む。

【0020】

一部の実施形態では、この方法は、少なくとも約90%の陰性的中率（negative predictive value）（NPV）で試料不一致を同定するステップを含む。一部の実施形態では、この方法は、少なくとも約95%の陰性的中率（NPV）で試料不一致を同定するステップを含む。一部の実施形態では、この方法は、少なくとも約99%の陰性的中率（NPV）で試料不一致を同定するステップを含む。

30

【0021】

一部の実施形態では、この方法は、少なくとも約0.90の曲線下面積（AUC）で試料不一致を同定するステップを含む。一部の実施形態では、この方法は、少なくとも約0.95の曲線下面積（AUC）で試料不一致を同定するステップを含む。一部の実施形態では、この方法は、少なくとも約0.99の曲線下面積（AUC）で試料不一致を同定するステップを含む。

【0022】

一部の実施形態では、所定の基準は、差異が、遺伝子型類似性における所定の閾値よりも大きい差異を含むことである。一部の実施形態では、所定の閾値は、約0.8である。

40

【0023】

一部の実施形態では、この方法は、同定された試料不一致に基づいて、第2の生体試料をさらなるアッセイから排除するステップをさらに含む。

【0024】

一部の実施形態では、この方法は、第1の試料フィンガープリントと第2の試料フィンガープリントとの間の差異が所定の基準を満たさない場合に、試料一致を同定するステップをさらに含む。

【0025】

一部の実施形態では、この方法は、少なくとも約90%の感度で試料一致を同定するス

50

テップを含む。一部の実施形態では、この方法は、少なくとも約 95 % の感度で試料一致を同定するステップを含む。一部の実施形態では、この方法は、少なくとも約 99 % の感度で試料一致を同定するステップを含む。

【0026】

一部の実施形態では、この方法は、少なくとも約 90 % の特異度で試料一致を同定するステップを含む。一部の実施形態では、この方法は、少なくとも約 95 % の特異度で試料一致を同定するステップを含む。一部の実施形態では、この方法は、少なくとも約 99 % の特異度で試料一致を同定するステップを含む。

【0027】

一部の実施形態では、この方法は、少なくとも約 90 % の陽性的中率 (PPV) で試料一致を同定するステップを含む。一部の実施形態では、この方法は、少なくとも約 95 % の陽性的中率 (PPV) で試料一致を同定するステップを含む。一部の実施形態では、この方法は、少なくとも約 99 % の陽性的中率 (PPV) で試料一致を同定するステップを含む。

10

【0028】

一部の実施形態では、この方法は、少なくとも約 90 % の陰性的中率 (NPV) で試料一致を同定するステップを含む。一部の実施形態では、この方法は、少なくとも約 95 % の陰性的中率 (NPV) で試料一致を同定するステップを含む。一部の実施形態では、この方法は、少なくとも約 99 % の陰性的中率 (NPV) で試料一致を同定するステップを含む。

20

【0029】

一部の実施形態では、この方法は、少なくとも約 0.90 の曲線下面積 (AUC) で試料一致を同定するステップを含む。一部の実施形態では、この方法は、少なくとも約 0.95 の曲線下面積 (AUC) で試料一致を同定するステップを含む。一部の実施形態では、この方法は、少なくとも約 0.99 の曲線下面積 (AUC) で試料一致を同定するステップを含む。

【0030】

一部の実施形態では、この方法は、同定された試料一致に基づいて、第 2 の生体試料をさらなるアッセイに供するステップをさらに含む。一部の実施形態では、この方法は、同定された試料一致に基づいて、データベース中に第 2 の試料フィンガープリントを記憶するステップ、および必要に応じてデータベース中に第 1 の試料フィンガープリントを記憶するステップをさらに含む。

30

【0031】

別の態様では、本開示は、1 つまたは複数のコンピュータプロセッサによる実行の際に試料不一致を同定するための方法を実装する機械実行可能コードを含む非一過性コンピュータ可読媒体であって、この方法が、複数の遺伝子座の各々における第 1 の生体試料の第 1 の複数の核酸分子の定量的尺度を含む第 1 の試料フィンガープリントの情報を受け取るステップであって、複数の遺伝子座が、常染色体一塩基多型 (SNP) を含み、第 1 の複数の核酸分子の定量的尺度が、複数の核酸分子の 12 個以下の独立した尺度を含む、ステップ；複数の遺伝子座の各々における第 2 の生体試料の第 2 の複数の核酸分子の定量的尺度を含む第 2 の試料フィンガープリントの情報を受け取るステップであって、第 2 の生体試料が、被験体から得られる、ステップ；第 1 の試料フィンガープリントと第 2 の試料フィンガープリントとの間の差異を決定するステップ；および第 1 の試料フィンガープリントと第 2 の試料フィンガープリントとの間の差異が所定の基準を満たす場合に、試料不一致を同定するステップを含む、非一過性コンピュータ可読媒体を提供する。

40

【0032】

別の態様では、本開示は、試料不一致を同定するためのコンピュータに実装された方法であって、第 1 の複数の核酸分子 (例えば、被験体から得られた第 1 の生体試料由来) を処理して、複数の遺伝子座の各々における第 1 の複数の核酸分子の定量的尺度を含む第 1 の試料フィンガープリントを生成するステップであって、複数の遺伝子座が、常染色体一

50

塩基多型（SNP）を含む、ステップ；第2の複数の核酸分子（例えば、被験体から得られた第2の生体試料由来）を処理して、複数の遺伝子座の各々における第2の複数の核酸分子の定量的尺度を含む第2の試料フィンガープリントを生成するステップ；第1の試料フィンガープリントと第2の試料フィンガープリントとの間の差異を決定するステップ；および第1の試料フィンガープリントと第2の試料フィンガープリントとの間の差異が所定の閾値を超える場合に、試料不一致を同定するステップを含み、第1の複数の核酸分子の定量的尺度が、第1の複数の核酸分子の12個以下の独立した尺度を含む、方法を提供する。

【0033】

別の態様では、本開示は、試料不一致を同定するためのコンピュータに実装された方法であって、第1の複数の核酸分子（例えば、被験体から得られた第1の生体試料由来）を処理して、複数の遺伝子座の各々における第1の複数の核酸分子の定量的尺度を含む第1の試料フィンガープリントを生成するステップであって、複数の遺伝子座が、常染色体一塩基多型（SNP）を含む、ステップ；第2の複数の核酸分子（例えば、被験体から得られた第2の生体試料由来）を処理して、複数の遺伝子座の各々における第2の複数の核酸分子の定量的尺度を含む第2の試料フィンガープリントを生成するステップ；第1の試料フィンガープリントと第2の試料フィンガープリントとの間の差異を決定するステップ；および第1の試料フィンガープリントと第2の試料フィンガープリントとの間の差異が所定の閾値を超える場合に、試料不一致を同定するステップを含み、常染色体一塩基多型が、単純な一塩基多型を含む、方法を提供する。

【0034】

別の態様では、本開示は、試料不一致を同定するためのコンピュータに実装された方法であって、第1の複数の核酸分子（例えば、被験体から得られた第1の生体試料由来）を処理して、複数の遺伝子座の各々における第1の複数の核酸分子の定量的尺度を含む第1の試料フィンガープリントを生成するステップであって、複数の遺伝子座が、常染色体一塩基多型（SNP）を含む、ステップ；第2の複数の核酸分子（例えば、被験体から得られた第2の生体試料由来）を処理して、複数の遺伝子座の各々における第2の複数の核酸分子の定量的尺度を含む第2の試料フィンガープリントを生成するステップ；第1の試料フィンガープリントと第2の試料フィンガープリントとの間の差異を決定するステップ；および第1の試料フィンガープリントと第2の試料フィンガープリントとの間の差異が所定の閾値を超える場合に、試料不一致を同定するステップを含み、常染色体一塩基多型が、所定の閾値を超えるマイナー対立遺伝子割合を有する、方法を提供する。

【0035】

別の態様では、本開示は、少なくとも1つの電子プロセッサによって実行される場合に少なくとも以下：第1の複数の核酸分子（例えば、被験体から得られた第1の生体試料由来）を処理して、複数の遺伝子座の各々における第1の複数の核酸分子の定量的尺度を含む第1の試料フィンガープリントを生成するステップであって、複数の遺伝子座が、常染色体一塩基多型（SNP）を含む、ステップ；第2の複数の核酸分子（例えば、被験体から得られた第2の生体試料由来）を処理して、複数の遺伝子座の各々における第2の複数の核酸分子の定量的尺度を含む第2の試料フィンガープリントを生成するステップ；第1の試料フィンガープリントと第2の試料フィンガープリントとの間の差異を決定するステップ；および第1の試料フィンガープリントと第2の試料フィンガープリントとの間の差異が所定の閾値を超える場合に、試料不一致を同定するステップを実施する非一過性コンピュータ実行可能命令を含むコンピュータ可読媒体を含む、またはそれにアクセスすることが可能なコントローラーを含むシステムであって、第1の複数の核酸分子の定量的尺度が、第1の複数の核酸分子の12個以下の独立した尺度を含む、システムを提供する。

【0036】

別の態様では、本開示は、少なくとも1つの電子プロセッサによって実行される場合に少なくとも以下：第1の複数の核酸分子（例えば、被験体から得られた第1の生体試料由来）を処理して、複数の遺伝子座の各々における第1の複数の核酸分子の定量的尺度を

含む第1の試料フィンガープリントを生成するステップであって、複数の遺伝子座が、常染色体一塩基多型（SNP）を含む、ステップ；第2の複数の核酸分子（例えば、被験体から得られた第2の生体試料由来）を処理して、複数の遺伝子座の各々における第2の複数の核酸分子の定量的尺度を含む第2の試料フィンガープリントを生成するステップ；第1の試料フィンガープリントと第2の試料フィンガープリントとの間の差異を決定するステップ；および第1の試料フィンガープリントと第2の試料フィンガープリントとの間の差異が所定の閾値を超える場合に、試料不一致を同定するステップを実施する非一過性コンピュータ実行可能命令を含むコンピュータ可読媒体を含む、またはそれにアクセスすることが可能なコントローラーを含むシステムであって、常染色体一塩基多型が、単純な一塩基多型を含む、システムを提供する。

10

【0037】

別の態様では、本開示は、少なくとも1つの電子プロセッサによって実行される場合に少なくとも以下：第1の複数の核酸分子（例えば、被験体から得られた第1の生体試料由来）を処理して、複数の遺伝子座の各々における第1の複数の核酸分子の定量的尺度を含む第1の試料フィンガープリントを生成するステップであって、複数の遺伝子座が、常染色体一塩基多型（SNP）を含む、ステップ；第2の複数の核酸分子（例えば、被験体から得られた第2の生体試料由来）を処理して、複数の遺伝子座の各々における第2の複数の核酸分子の定量的尺度を含む第2の試料フィンガープリントを生成するステップ；第1の試料フィンガープリントと第2の試料フィンガープリントとの間の差異を決定するステップ；および第1の試料フィンガープリントと第2の試料フィンガープリントとの間の差異が所定の閾値を超える場合に、試料不一致を同定するステップを実施する非一過性コンピュータ実行可能命令を含むコンピュータ可読媒体を含む、またはそれにアクセスすることが可能なコントローラーを含むシステムであって、常染色体一塩基多型が、所定の閾値を超えるマイナー対立遺伝子割合を有する、システムを提供する。

20

【0038】

別の態様では、本開示は、試料不一致を同定するためのコンピュータに実装された方法であって、複数の遺伝子座の各々における第1の複数の核酸分子（例えば、被験体から得られた第1の生体試料由来）の定量的尺度を含む第1の試料フィンガープリントを得るステップであって、複数の遺伝子座が、常染色体一塩基多型（SNP）を含む、ステップ；複数の遺伝子座の各々における第2の複数の核酸分子（例えば、被験体から得られた第2の生体試料由来）の定量的尺度を含む第2の試料フィンガープリントを得るステップ；第1の試料フィンガープリントと第2の試料フィンガープリントとの間の差異を決定するステップ；および第1の試料フィンガープリントと第2の試料フィンガープリントとの間の差異が所定の閾値を超える場合に、試料不一致を同定するステップを含み、第1の複数の核酸分子の定量的尺度が、第1の複数の核酸分子の12個以下の独立した尺度を含む、方法を提供する。

30

【0039】

別の態様では、本開示は、試料不一致を同定するためのコンピュータに実装された方法であって、複数の遺伝子座の各々における第1の複数の核酸分子（例えば、被験体から得られた第1の生体試料由来）の定量的尺度を含む第1の試料フィンガープリントを得るステップであって、複数の遺伝子座が、常染色体一塩基多型（SNP）を含む、ステップ；複数の遺伝子座の各々における第2の複数の核酸分子（例えば、被験体から得られた第2の生体試料由来）の定量的尺度を含む第2の試料フィンガープリントを得るステップ；第1の試料フィンガープリントと第2の試料フィンガープリントとの間の差異を決定するステップ；および第1の試料フィンガープリントと第2の試料フィンガープリントとの間の差異が所定の閾値を超える場合に、試料不一致を同定するステップを含み、常染色体一塩基多型が、単純な一塩基多型を含む、方法を提供する。

40

【0040】

別の態様では、本開示は、試料不一致を同定するためのコンピュータに実装された方法であって、複数の遺伝子座の各々における第1の複数の核酸分子（例えば、被験体から得

50

られた第1の生体試料由来)の定量的尺度を含む第1の試料フィンガープリントを得るステップであって、複数の遺伝子座が、常染色体一塩基多型(SNP)を含む、ステップ；複数の遺伝子座の各々における第2の複数の核酸分子(例えば、被験体から得られた第2の生体試料由来)の定量的尺度を含む第2の試料フィンガープリントを得るステップ；第1の試料フィンガープリントと第2の試料フィンガープリントとの間の差異を決定するステップ；および第1の試料フィンガープリントと第2の試料フィンガープリントとの間の差異が所定の閾値を超える場合に、試料不一致を同定するステップを含み、常染色体一塩基多型が、所定の閾値を超えるマイナー対立遺伝子割合を有する、方法を提供する。

【0041】

別の態様では、本開示は、少なくとも1つの電子プロセッサによって実行される場合に少なくとも以下：複数の遺伝子座の各々における第1の複数の核酸分子(例えば、被験体から得られた第1の生体試料由来)の定量的尺度を含む第1の試料フィンガープリントを得るステップであって、複数の遺伝子座が、常染色体一塩基多型(SNP)を含む、ステップ；複数の遺伝子座の各々における第2の複数の核酸分子(例えば、被験体から得られた第2の生体試料由来)の定量的尺度を含む第2の試料フィンガープリントを得るステップ；第1の試料フィンガープリントと第2の試料フィンガープリントとの間の差異を決定するステップ；および第1の試料フィンガープリントと第2の試料フィンガープリントとの間の差異が所定の閾値を超える場合に、試料不一致を同定するステップを実施する非一過性コンピュータ実行可能命令を含むコンピュータ可読媒体を含む、またはそれにアクセスすることが可能なコントローラーを含むシステムであって、第1の複数の核酸分子の定量的尺度が、第1の複数の核酸分子の12個以下の独立した尺度を含む、システムを提供する。

【0042】

別の態様では、本開示は、少なくとも1つの電子プロセッサによって実行される場合に少なくとも以下：複数の遺伝子座の各々における第1の複数の核酸分子(例えば、被験体から得られた第1の生体試料由来)の定量的尺度を含む第1の試料フィンガープリントを得るステップであって、複数の遺伝子座が、常染色体一塩基多型(SNP)を含む、ステップ；複数の遺伝子座の各々における第2の複数の核酸分子(例えば、被験体から得られた第2の生体試料由来)の定量的尺度を含む第2の試料フィンガープリントを得るステップ；第1の試料フィンガープリントと第2の試料フィンガープリントとの間の差異を決定するステップ；および第1の試料フィンガープリントと第2の試料フィンガープリントとの間の差異が所定の閾値を超える場合に、試料不一致を同定するステップを実施する非一過性コンピュータ実行可能命令を含むコンピュータ可読媒体を含む、またはそれにアクセスすることが可能なコントローラーを含むシステムであって、常染色体一塩基多型が、単純な一塩基多型を含む、システムを提供する。

【0043】

別の態様では、本開示は、少なくとも1つの電子プロセッサによって実行される場合に少なくとも以下：複数の遺伝子座の各々における第1の複数の核酸分子(例えば、被験体から得られた第1の生体試料由来)の定量的尺度を含む第1の試料フィンガープリントを得るステップであって、複数の遺伝子座が、常染色体一塩基多型(SNP)を含む、ステップ；複数の遺伝子座の各々における第2の複数の核酸分子(例えば、被験体から得られた第2の生体試料由来)の定量的尺度を含む第2の試料フィンガープリントを得るステップ；第1の試料フィンガープリントと第2の試料フィンガープリントとの間の差異を決定するステップ；および第1の試料フィンガープリントと第2の試料フィンガープリントとの間の差異が所定の閾値を超える場合に、試料不一致を同定するステップを実施する非一過性コンピュータ実行可能命令を含むコンピュータ可読媒体を含む、またはそれにアクセスすることが可能なコントローラーを含むシステムであって、常染色体一塩基多型が、所定の閾値を超えるマイナー対立遺伝子割合を有する、システムを提供する。

【0044】

本開示のさらなる態様および利点は、本開示の例示的な実施形態のみが示され記載され

10

20

30

40

50

た以下の詳細な説明から当業者に容易に明らかとなる。理解されるように、本開示は、他の実施形態および異なる実施形態が可能であり、その一部の詳細は、全て本開示から逸脱することなしに、種々の明らかな事項において改変が可能である。したがって、図面および説明は、事実上、限定的ではなく例示的とみなすべきである。

【0045】

参照による組み込み

本明細書中で言及される全ての刊行物、特許および特許出願は、各個々の刊行物、特許または特許出願が参照により本明細書に組み込まれると具体的かつ個々に示されるのと同じ程度まで、参照により本明細書に組み込まれる。参照により組み込まれる刊行物および特許または特許出願が、本明細書中に含まれる開示と矛盾する範囲について、本明細書が、任意のかかる矛盾する資料に優先する (supercede) および／または優先する (take precedence over) 意図である。

【0046】

本発明のいくつかの新規の特色は、添付の特許請求の範囲に詳細に示される。本発明の特色および利点のより良い理解は、本発明の原理が利用される例示的な実施形態を示す以下の発明を実施するための形態、および添付の図面（本明細書で「図 (Figure)」および「図 (FIG.)」とも）を参照して得られる。

【図面の簡単な説明】

【0047】

【図1】図1は、一部の実施形態に従う、生体試料のフィンガープリンティングのための方法の一例を示す。

【0048】

【図2】図2は、一部の実施形態に従う、第1の生体試料および第2の生体試料をフィンガープリンティングすることに基づいて試料不一致を同定するための方法の一例を示す。

【0049】

【図3】図3は、複数のアッセイされた生体試料から生成された試料フィンガープリントの比較の完全な可視化を示す。対角線に沿った濃い黒線は、取り違えなかった全ての試料を示す（例えば、試料一致）。対角線から離れた要素は、異なる被験体から得られたと推定される試料と類似しすぎている試料を示す（例えば、潜在的試料不一致）。

【0050】

【図4】図4は、明確な内部試料不一致（例えば、試料取り違え）の一例を示し、このとき、2つの異なる被験体から得られた多数の生体試料に対して、アッセイの比較の可視化を実施した。対角線上の「分割された」四角の隣の対角線から離れたバーは、これらの2つの試料が差し替えられたことを示す (BLIB00366およびBLIB00367)。

【0051】

【図5】図5は、明確な試料不一致（例えば、試料取り違え）および分解できない試料食い違いの一例の画像を示す。第1の患者 (ID番号4181) および第2の患者 (ID番号4175) から得られた組織試料を取り違えた。第3の患者 (ID番号4161) についてのcfDNA試料の1つは、第3の患者 (ID番号4161) 由来であると推定される他の試料を含む他のあらゆる試料とも一致しない。したがって、この試料は、さらなるアッセイおよび処理から排除した。

【0052】

【図6】図6は、同じまたは異なる被験体（例えば、患者または人間）由来の試料の対の間の期待された遺伝子型類似性を示すプロットを示す。このプロットは、同じ人間から得られた試料と異なる人間から得られた試料との間を識別または差別化するために適切な閾値がどのように同定されるかを示している。取り違えられたと疑われる試料および低いカバレッジ（低い数の遺伝子型比較をもたらす）を有する試料を排除することによって潜在的試料不一致が説明された後で、分布は完全に分離される。したがって、閾値決定は、0.8の遺伝子型類似性において実施され得る。

10

20

30

40

50

【0053】

【図7】図7は、複数のアッセイされたDNA試料についてのジェンダーコールの比較を示す。X読み取りはX軸上に示され、Y読み取りはY軸上に示される。青色の試料は、男性被験体から得られたと推定され、赤色の試料は、女性被験体から得られたと推定され、灰色の試料は、かかる情報が入手不能であった。閾値ラインを十分に上回る位置にあるデータポイントの第1のセットは、男性としてコールされ、閾値ラインを十分に下回る位置にあるデータポイントの第2のセットは、女性としてコールされる。このプロットは、試料不一致として同定された（例えば、取り違えられたと同定された）試料に対応する、閾値ラインの下に位置する少数の青色のデータポイントおよび閾値の上に位置する少数の赤色のデータポイントを示している。閾値ラインのちょうど上にあるデータポイントは、X染色体の大部分が重複したがん患者から得られた。

10

【0054】

【図8】図8は、本明細書で提供される方法を実装するようにプログラミングされるまたは他の方法で構成されるコンピュータシステムを示す。

【発明を実施するための形態】

【0055】

用語「核酸」または「ポリヌクレオチド」は、本明細書で使用される場合、一般に、1つまたは複数の核酸サブユニットまたはヌクレオチドを含む分子を指す。核酸は、アデノシン（A）、シトシン（C）、グアニン（G）、チミン（T）およびウラシル（U）、またはそれらのバリエーションから選択される1つまたは複数のヌクレオチドを含み得る。ヌクレオチドは、一般に、1つのヌクレオシドおよび少なくとも1、2、3、4、5、6、7、8、9、10個またはそれよりも多いリン酸（ PO_3 ）基を含む。ヌクレオチドは、1つの核酸塩基、1つの五炭糖（リボースまたはデオキシリボースのいずれか）および1つまたは複数のリン酸基を、個々にまたは組み合わせて含み得る。

20

【0056】

リボヌクレオチドは、糖がリボースであるヌクレオチドである。デオキシリボヌクレオチドは、糖がデオキシリボースであるヌクレオチドである。ヌクレオチドは、ヌクレオシドーリン酸またはヌクレオシドポリリン酸であり得る。ヌクレオチドは、検出可能なタグ、例えば、発光タグまたはマーカー（例えば、フルオロフォア）を含む、例えば、デオキシアデノシン三リン酸（dATP）、デオキシシチジン三リン酸（dCTP）、デオキシグアニン三リン酸（dGTP）、ウリジン三リン酸（dUTP）およびデオキシチミジン三リン酸（dTTP）のデオキシリボヌクレオシド三リン酸（dNTP）から選択され得るdNTPなどのデオキシリボヌクレオシドポリリン酸であり得る。ヌクレオチドには、成長中の核酸鎖中に取り込まれ得る任意のサブユニットが含まれ得る。かかるサブユニットは、A、C、G、TもしくはU、あるいは1つもしくは複数の相補的A、C、G、TもしくはUに対して特異的な、またはプリン（即ち、AもしくはG、またはそれらのバリエーション）もしくはピリミジン（即ち、C、TもしくはU、またはそれらのバリエーション）に対して相補的な、任意の他のサブユニットであり得る。一部の例では、核酸は、デオキシリボ核酸（DNA）、リボ核酸（RNA）、またはそれらの誘導体もしくはバリエーションである。核酸は、一本鎖または二本鎖であり得る。核酸分子は、線状、曲線状もしくは環状、またはそれらの任意の組合せであり得る。

30

40

【0057】

用語「核酸分子」、「核酸配列」、「核酸断片」、「オリゴヌクレオチド」および「ポリヌクレオチド」は、本明細書で使用される場合、一般に、デオキシリボヌクレオチドもしくはリボヌクレオチド（RNA）のいずれか、またはそれらのアナログなどの、種々の長さを有し得るポリヌクレオチドを指す。核酸分子は、少なくとも約5塩基、10塩基、20塩基、30塩基、40塩基、50塩基、60塩基、70塩基、80塩基、90、100塩基、110塩基、120塩基、130塩基、140塩基、150塩基、160塩基、170塩基、180塩基、190塩基、200塩基、300塩基、400塩基、500塩基、1キロ塩基（kb）、2kb、3kb、4kb、5kb、10kbもしくは50k

50

bの長さを有し得、または上述の値の任意の2つの間の任意の数の塩基を有し得る。オリゴヌクレオチドは、典型的には、4つのヌクレオチド塩基：アデニン（A）；シトシン（C）；グアニン（G）；およびチミン（T）（ポリヌクレオチドがRNAである場合、チミン（T）の代わりにウラシル（U））の特定の配列から構成される。したがって、用語「核酸分子」、「核酸配列」、「核酸断片」、「オリゴヌクレオチド」および「ポリヌクレオチド」は、ポリヌクレオチド分子のアルファベット表示であることが、少なくとも一部意図される。あるいは、これらの用語は、ポリヌクレオチド分子自体に適用され得る。このアルファベット表示は、中央処理装置を有するコンピュータ中のデータベース中に入力され得、ならびに／またはバイオインフォマティクス適用、例えば、機能ゲノミクスおよび相同性検索のために使用され得る。オリゴヌクレオチドは、1つまたは複数の非標準ヌクレオチド、ヌクレオチドアナログおよび／または改変ヌクレオチドを含み得る。

10

【0058】

用語「試料」は、本明細書で使用される場合、一般に、生体試料を指す。生体試料の例には、核酸分子、アミノ酸、ポリペプチド、タンパク質、炭水化物、脂肪またはウイルスが含まれる。一例では、生体試料は、1つまたは複数の核酸分子を含む核酸試料である。核酸分子は、無細胞または無細胞核酸分子、例えば、無細胞DNA（cfDNA）または無細胞RNA（cfRNA）であり得る。核酸分子は、バフィーコート核酸分子、例えば、バフィーコートDNAであり得る。核酸分子は、ヒト、哺乳動物、非ヒト哺乳動物、類人猿、サル、チンパンジー、爬虫類、両生類またはトリ供給源を含む種々の供給源に由来し得る。さらに、試料は、血液、血清、血漿、硝子体、痰、尿、涙、汗、唾液、精液、粘膜排出物、粘液、髄液、羊水、リンパ液などが含まれるがこれらに限定されない、無細胞配列を含有する種々の動物体液から抽出され得る。無細胞ポリヌクレオチド（例えば、cfDNA）は、起源が胎仔であり得（妊娠被験体から採取した体液を介して）、または被験体自身の組織に由来し得る。

20

【0059】

用語「被験体」は、本明細書で使用される場合、一般に、処理または分析を受けている生体試料を有する個体を指す。被験体は、動物または植物であり得る。被験体は、哺乳動物、例えば、ヒト、イヌ、ネコ、ウマ、ブタまたはげっ歯類であり得る。被験体は、例えば、疾患、例えば、1つもしくは複数のがん、1つもしくは複数の感染性疾患、1つもしくは複数の遺伝性障害、または1つもしくは複数の腫瘍、あるいはそれらの任意の組合せを有する、またはそれを有すると疑われる、患者であり得る。1つまたは複数の腫瘍を有するまたはそれを有すると疑われる被験体について、腫瘍は、1つまたは複数の型のものであり得る。

30

【0060】

用語「全血」は、本明細書で使用される場合、一般に、（例えば、遠心分離によって）副成分へと分離されていない血液試料を指す。全血の血液試料は、cfDNAおよび／または生殖系列DNAを含有し得る。全血DNA（これは、cfDNAおよび／または生殖系列DNAを含有し得る）は、血液試料から抽出され得る。全血DNA配列決定読み取り（これは、cfDNA配列決定読み取りおよび／または生殖系列DNA配列決定読み取りを含有し得る）は、全血DNAから抽出され得る。

40

【0061】

被験体から得られた生体試料の収集およびアッセイは、臨床および実験室プロセスを通じた試料同一性の信頼性のある維持に関する課題にしばしば遭遇し得る。例えば、生体試料は、実験室または臨床設定において不注意にしばしば取り違えられ得、それによって、検出されず修正されないままだと、潜在的に不正確な臨床結果を生じる。

【0062】

遺伝子座のパネルを使用して生体試料をフィンガープリンティングするための方法は、所望の感度、特異度または精度で遺伝情報を得るために、十分に深いカバレッジを要求し得る。例えば、深いカバレッジが、異なる試料から生成されたフィンガープリント間を識別するために十分なシグナル・ノイズ（SNR）比のために要求され得る。かかる試料は

50

、例えば、2つの異なる時点において同じ被験体から得られた長期的試料であり得る。ローパス配列決定を使用して処理される長期的試料は、以下に関する課題に遭遇し得る：（１）異なる時点からの一致する試料について修正を行うこと、および（２）いずれか１つの位置における比較的低い読み取りカバレッジにもかかわらず試料フィンガープリンティングに適切な遺伝子座のパネルを同定すること。

【0063】

生体試料のフィンガープリントを生成し比較するための方法およびシステムが提供される。試料フィンガープリントは、１つまたは複数の時点の各々において被験体から得られた生体試料由来の核酸分子の１つまたは複数のセットを配列決定することによって生成され得る。試料フィンガープリントのペアワイズ比較は、試料不一致（例えば、2つの試料が異なる被験体から得られたこと）または試料一致（例えば、2つの試料が同じ被験体から得られたこと）が、試料フィンガープリントが生成された2つの生体試料間に存在するかどうかを決定するために実施され得る。

10

【0064】

ある態様では、本開示は、試料フィンガープリントを生成するための方法であって、被験体から複数の核酸分子を含む生体試料を得るステップ；および複数の核酸分子を処理して、複数の遺伝子座の各々における複数の核酸分子の定量的尺度を含む試料フィンガープリントを生成するステップであって、複数の遺伝子座が、常染色体一塩基多型（SNP）を含む、ステップを含む方法を提供する。生成された試料フィンガープリントは、データベース中に記憶され得る。

20

【0065】

別の態様では、本開示は、試料不一致を同定するための方法であって、被験体から第１の複数の核酸分子を含む第１の生体試料を得るステップ；第１の複数の核酸分子を処理して、複数の遺伝子座の各々における第１の複数の核酸分子の定量的尺度を含む第１の試料フィンガープリントを生成するステップであって、複数の遺伝子座が、常染色体一塩基多型（SNP）を含む、ステップ；被験体から第２の複数の核酸分子を含む第２の生体試料を得るステップ；第２の複数の核酸分子を処理して、複数の遺伝子座の各々における第２の複数の核酸分子の定量的尺度を含む第２の試料フィンガープリントを生成するステップ；第１の試料フィンガープリントと第２の試料フィンガープリントとの間の差異を決定するステップ；および第１の試料フィンガープリントと第２の試料フィンガープリントとの間の差異が所定の基準を満たす場合に、試料不一致を同定するステップを含む方法を提供する。

30

【0066】

図１は、一部の実施形態に従う、生体試料の試料フィンガープリントを生成するための方法の一例を示す。試料フィンガープリントを生成するための方法は、被験体から複数の核酸分子を含む生体試料を得るステップを含み得る。一部の実施形態では、複数の核酸分子は、複数の無細胞DNA（cfDNA）分子、複数のパフィーコートDNA分子、複数の固形腫瘍DNA分子、またはそれらの組合せを含み得る（操作１０５にある通り）。

【0067】

試料フィンガープリントを生成するための方法は、複数の核酸分子を処理して、複数の遺伝子座の各々における複数の核酸分子の定量的尺度を含む試料フィンガープリントを生成するステップを含み得る。一部の実施形態では、複数の核酸分子を処理するステップは、複数の核酸分子を配列決定して、複数の遺伝子座の各々における配列決定読み取りを生成する工程を含む（操作１１０にある通り）。

40

【0068】

一部の実施形態では、複数の遺伝子座は、複数の別個の常染色体SNPを含み得る。一部の例では、分析される複数の遺伝子座は、約１００よりも多い遺伝子座を含み得る。一部の例では、分析される複数の遺伝子座は、約２００よりも多い遺伝子座、約３００よりも多い遺伝子座、約５００よりも多い遺伝子座、約１，０００よりも多い遺伝子座、約１，５００よりも多い遺伝子座、約２，０００よりも多い遺伝子座、約２，５００よりも多

50

い遺伝子座、約 3, 000 よりも多い遺伝子座、約 3, 500 よりも多い遺伝子座、約 4, 000 よりも多い遺伝子座、約 4, 500 よりも多い遺伝子座、約 5, 000 よりも多い遺伝子座、または約 5, 500 よりも多い遺伝子座を含み得る。一部の例では、別個の常染色体 SNP を有する遺伝子座には、db SNP などの公的データベース中に含まれる第 1 染色体上に位置するアノテーションされた SNP である rs 2839 が含まれ得る。一部の例では、試料フィンガープリントプロファイルの一部としての使用に適切な別個の常染色体 SNP、例えば、rs 2839 は、例えば、品質および信頼性標準を満たす SNP をコールするために、品質基準に基づいて公知の SNP のデータベースをフィルタリングすること、またはヒト参加者の大きいセットからのゲノムデータの大きいデータセットを分析することによって、同定され得る。

10

【0069】

一部の実施形態では、SNP、例えば、個人のゲノムを固有に同定できる SNP が、ある特定の基準についてフィルタリングされ得る。SNP のかかるセットは、2 つの個体が（例えば、試料フィンガープリントについて）同じゲノムプロファイルを有する極度に小さい可能性を集会的にもたらし得る。例えば、主要五大陸の集団に及ぶ報告された対立遺伝子頻度を有する SNP（例えば、1000 ゲノムプロジェクトおよび ExAC Consortium からの）は、試料フィンガープリントプロファイル中への包含についてさらに分析される候補 SNP として機能し得る。別の例として、被験体の ABO 血液型を予測するために使用され得る SNP が使用され得る。別の例として、被験体の性別を予測するために使用され得る SNP が使用され得る。SNP を選択する方法は、例えば、Du et al. (“A SNP panel and online tool for checking genotype concordance through comparing QR codes”, PLOS One, 2017) および Hu et al. (“Evaluating information content of SNPs for sample-tagging in re-sequencing projects”, Scientific Reports, 2015) によって記載される通りであり得、これらは各々、その全体がこれにより参照により本明細書に組み込まれる。

20

【0070】

一部の例では、SNP は、常染色体 SNP を選択するためにフィルタリングされ得る。一部の例では、SNP は、単純な SNP を選択するためにフィルタリングされ得る。単純な SNP は、挿入も欠失も有さない 2 つの対立遺伝子のみを有する SNP を含み得る。単純な SNP は、単一塩基変化のみを有し得る。一部の例では、SNP は、低い参照 SNP ID (rs 番号) で、db SNP 中でアノテーションされ得る。これらの rs 番号は、データベースへの提出の時点で順次割り当てられる。一部の例では、より小さい rs 番号を有するより早期の提出は、より少ない技術的アーチファクトを有し得る。一部の例では、SNP は、ある特定の閾値よりも大きいマイナー対立遺伝子割合を有するようにフィルタリングされ得る。一部の例では、SNP は、約 1 % よりも大きい、約 1.5 % よりも大きい、約 2 % よりも大きい、約 2.5 % よりも大きい、約 3 % よりも大きい、約 3.5 % よりも大きい、約 4 % よりも大きい、約 4.5 % よりも大きい、約 5 % よりも大きい、約 5.5 % よりも大きい、約 6 % よりも大きい、約 6.5 % よりも大きい、約 7 % よりも大きい、約 7.5 % よりも大きい、約 8 % よりも大きい、8.5 % よりも大きい、約 9 % よりも大きい、約 9.5 % よりも大きい、または約 10 % よりも大きいマイナー対立遺伝子割合を有するようにフィルタリングされ得る。

30

40

【0071】

一部の実施形態では、試料フィンガープリントを生成するための方法は、データベース中に、生成された試料フィンガープリントを記憶するステップをさらに含み得る（操作 115 にある通り）。

【0072】

例えば、配列決定読み取りは、任意の適切な配列決定方法を使用して、核酸分子から生

50

成され得る。配列決定方法は、第一世代配列決定方法、例えば、Maxam-GilbertもしくはSanger配列決定、またはハイスループット配列決定（例えば、次世代配列決定、即ちNGS）方法であり得る。ハイスループット配列決定方法は、少なくとも約10,000、100,000、100万、1000万、1億、10億またはそれよりも多いポリヌクレオチド分子を同時に（または実質的に同時に）配列決定し得る。配列決定方法には、以下が含まれ得るがこれらに限定されない：パイロシーケンシング、合成による配列決定、単一分子配列決定、nanopore配列決定、半導体配列決定、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、Digital Gene Expression (Helicos)、超並列配列決定、例えば、Helicos、Clonal Single Molecule Array (Solexa/Illumina)、PacBio、SOLiD、Ion TorrentまたはNanoporeプラットフォームを使用する配列決定。

【0073】

一部の実施形態では、配列決定する工程は、全ゲノム配列決定 (WGS) を含む。配列決定する工程は、被験体から得られた生体試料から試料フィンガープリントを生成するのに十分な、または2つの試料フィンガープリント間の差異に基づいて、所望のパフォーマンス（例えば、精度、感度、特異度、陽性的中率 (PPV)、陰性的中率 (NPV)、または受信者動作特性 (ROC) の曲線下面積 (AUC)）で試料不一致もしくは試料一致を同定するのに十分な深度で実施され得る。一部の実施形態では、配列決定する工程は、例えば、約12×以下、約11×以下、約10×以下、約9×以下、約8×以下、約7×以下、約6×以下、約5×以下、約4×以下、約3×以下、約2×以下、または約1×以下の深度で、「ローパス」様式で実施される。

【0074】

一部の実施形態では、被験体から得られた生体試料から試料フィンガープリントを生成するステップは、配列決定読み取りを参照ゲノムにアラインメントさせる工程を含み得る。参照ゲノムは、ゲノム（例えば、ヒトゲノム）の少なくとも一部分を含み得る。参照ゲノムは、ゲノム全体（例えば、ヒトゲノム全体）を含み得る。参照ゲノムは、ゲノムのコードおよび／または非コードゲノム領域に対応する複数のゲノム領域を含むデータベースを含み得る。データベースは、ゲノムのコードおよび／または非コードゲノム領域に対応する複数のゲノム領域、例えば、一塩基多型 (SNP)、一塩基バリエーション (SNV)、コピー数バリエーション (CNV)、挿入または欠失 (インデル)、融合遺伝子および反復エレメントを含み得る。アラインメントは、Burrows-Wheelerアルゴリズムまたは他のアラインメントアルゴリズムを使用して実施され得る。

【0075】

一部の実施形態では、被験体から得られた生体試料から試料フィンガープリントを生成するステップは、複数の遺伝子座の各々について、配列決定読み取りの定量的尺度を生成する工程を含み得る。配列決定読み取りの定量的尺度、例えば、所与の遺伝子座とアラインメントされる配列決定読み取りのカウントが生成され得る。

【0076】

一部の実施形態では、被験体から得られた生体試料から試料フィンガープリントを生成するための方法は、1つまたは複数のDNA試料（例えば、cfDNA、バフィーコートDNAおよび／または固形腫瘍DNA）の各々について、複数のSNPの各々における塩基コール（例えば、一部の塩基についての不確かなコールを含む）を生成するステップを含み得る。塩基コールは、例えば、GATKまたは他のSNPコーリングパッケージを使用して生成され得る。

【0077】

一部の実施形態では、被験体から得られた生体試料から生成された試料フィンガープリントは、被験体から得られた1つまたは複数の生体試料のセットを表すために、データベース中に記憶され得る。生体試料のセットは、1つまたは複数の時点において収集された1つまたは複数の型のDNA試料（例えば、cfDNA、バフィーコートDNAおよび／

10

20

30

40

50

または固形腫瘍DNA)を表し得る。データベース中に記憶された試料フィンガープリントは、約1ギガバイト(GB)以下、約500メガバイト(MB)以下、約100MB以下、約50MB以下、約10MB以下、約5MB以下、約1MB以下、約500キロバイト(KB)以下、約250KB以下、または約100KB以下のデータサイズを有し得る。

【0078】

一部の実施形態では、複数のSNPは、ゲノム全体に広がった良好に挙動するSNPの非常に大きいセットであり得る。SNPの各々は、非常に高くなくてもよいいくらかの情報内容を提供し得る。複数のSNPは、常染色体SNPであり得る。複数のSNPは、テロメアに緊密には近接していない位置にあり得る。複数のSNPは、ある特定の日付前の生成を示すIDで、dbSNP中でアノテーションされ得る。複数のSNPは、2つの対立遺伝子のみで、約1%よりも大きいマイナー対立遺伝子割合(MAF)を有し得る。一部の実施形態では、複数のSNPは、2つの対立遺伝子のみで、約1%、1.5%、2%、2.5%、3%、3.5%、4%、4.5%、5%、5.5%、6%、6.5%、7%、7.5%、8%、8.5%、9%、9.5%、10%、10.5%、11%、11.5%、12%、12.5%、13%、13.5%、14%、14.5%、15%、15.5%、16%、16.5%、17%、17.5%、18%、18.5%、19%、19.5%、20%、20.5%、21%、21.5%、22%、22.5%、23%、23.5%、24%、24.5%、25%、25.5%、26%、26.5%、27%、27.5%、28%、28.5%、29%、29.5%、30%、30.5%、31%、31.5%、32%、32.5%、33%、33.5%、34%、34.5%、35%、35.5%、36%、36.5%、37%、37.5%、38%、38.5%、39%、39.5%、40%、40.5%、41%、41.5%、42%、42.5%、43%、43.5%、44%、44.5%、45%、または45%超よりも大きい、マイナー対立遺伝子割合(MAF)を有し得る。

【0079】

図2は、一部の実施形態に従う、第1の生体試料および第2の生体試料をフィンガープリンティングすることに基づいて試料不一致を同定するための方法の一例を示す。一部の実施形態では、被験体から得られた生体試料から試料フィンガープリントを生成するための方法は、ベースライン時点および1つまたは複数の引き続く時点において、無細胞DNA(cfDNA)試料、バフィーコートDNA試料および/または固形腫瘍DNA試料を収集するステップを含み得る。同じベースライン時点においてまたはそのあたりにおいて被験体から得られたDNA試料の各セットは、被験体についての、ベースライン時点に対応するベースライン試料フィンガープリントを生成するために処理され得る。同じ引き続く時点においてまたはそのあたりにおいて被験体から得られたDNA試料の各セットは、被験体についての、引き続く時点に対応する引き続く試料フィンガープリントを生成するために処理され得る。

【0080】

例えば、第1の複数の核酸分子を含む第1の生体試料が、被験体から得られ得る(操作205にある通り)。第1の複数の核酸分子は、複数の遺伝子座の各々における第1の複数の核酸分子の定量的尺度を含む第1の試料フィンガープリントを生成するために処理され得る(操作210にある通り)。一部の実施形態では、複数の遺伝子座は、常染色体一塩基多型(SNP)を含む。次に、第2の複数の核酸分子を含む第2の生体試料が、被験体から得られ得る(操作215にある通り)。第2の複数の核酸分子は、複数の遺伝子座の各々における第2の複数の核酸分子の定量的尺度を含む第2の試料フィンガープリントを生成するために処理され得る(操作220にある通り)。次に、第1の試料フィンガープリントと第2の試料フィンガープリントとの間の差異が、決定され得る(操作225にある通り)。次に、差異が所定の基準を満たす場合に、試料不一致が同定され得る(操作230にある通り)。

【0081】

一部の実施形態では、被験体から得られた生体試料から複数の試料フィンガープリントが生成された後に、試料フィンガープリントは、試料フィンガープリントの配列データのペアワイズ比較を生成するために処理され得る。試料フィンガープリントの配列データのペアワイズ比較は、以下を確実にするために実施され得る：（a）同じ被験体（人間）由来であると推定される試料の全ての対が、実際に同じ被験体（人間）由来であること、（b）異なる被験体（人）由来であると推定される試料の全ての対が、実際に異なる被験体（人）由来であること、ならびに（c）全ての試料が、試料が得られる被験体の性別からの期待値に従うX染色体読み取りおよびY染色体読み取りを有すること。例えば、2つの試料間のペアワイズ比較は、第1の試料のフィンガープリント（cfDNA、パフィーコートDNAおよび／または固形腫瘍DNAをアッセイすることによって得られた定量的尺度を使用する）を、第2の試料のフィンガープリント（第1の試料フィンガープリントにおいて入手可能なDNAと同じ型のDNAをアッセイすることによって得られた定量的尺度を使用する）と比較することによって実施され得る。例えば、かかる定量的尺度は、核酸分子を配列決定することまたは核酸分子の結合測定を実施することによって生成され得る。

10

【0082】

試料フィンガープリントの配列データのペアワイズ比較を実施することは、コールの精度における所望の信頼度を有するように、両方の試料において十分な数の読み取りが存在するSNPコールの各々を比較することによって、遺伝子型類似性の定量的尺度を生成することを含み得る。所与のSNPについて、読み取りの数は、所与のSNPについての所定の閾値よりも大きい場合に、十分であると判断され得る。かかる所定の閾値は、（例えば、既知のSNP状態を有する患者についての）患者データの分析に基づいて、各SNPについて同定され得る。例えば、各SNPについての所定の閾値は、ホモ接合性コールよりもヘテロ接合性コールについて確信的なコールを行うために必要なより低い数の読み取りを考慮に入れることに基づいて決定され得る。

20

【0083】

試料フィンガープリントの配列データのペアワイズ比較を実施することは、2つの試料フィンガープリント間で同一である遺伝子型コールの割合に少なくとも一部基づいて、2つの試料が同じ被験体（人間）由来である（例えば、試料一致）または同じ被験体（人間）由来でない（例えば、試料不一致）と同定することを含み得る。例えば、2つの試料フィンガープリント間で同一である遺伝子型コールの割合は、試料不一致または試料一致を同定するために、所定の閾値と比較され得る。所定の閾値は、複数の被験体から生成された多数の試料フィンガープリントから集約された大量のデータを分析すること、および所望のパフォーマンス（例えば、精度、感度、特異度、陽性的中率（PPV）、陰性的中率（NPV）、または受信者動作特性（ROC）の曲線下面積（AUC））を最適化する所定の閾値を選択することによって、生成され得る。

30

【0084】

試料フィンガープリントの配列データのペアワイズ比較を実施することは、被験体（人間）によって群分けされる試料の全ての対について遺伝子型類似性のヒートマップを生成することを含み得る。これらの可視化では、内部試料取り違い（例えば、ユーザーの実験室設定において生じる試料不一致）は、対角線の端にある明るい四角とカップリングした、対角線から離れた暗い四角として示され得る。外部試料取り違い（例えば、診療所または他の試料収集場所において生じる試料不一致）は、対角線上の四角中の明るい「ギャップ」として示され得る。この可視化を補助するために、ヒートマップの生成は、取り違いられたと疑われる試料のセットに限定され得る。

40

【0085】

試料フィンガープリントの配列データのペアワイズ比較を実施することは、X染色体読み取りおよびY染色体読み取りの比較を含み得る。例えば、X染色体読み取りおよびY染色体読み取りの比較は、異なる性別の試料間の試料取り違い（試料不一致）を検出するために実施され得る。Y読み取り（例えば、Y性染色体への配列読み取りマッピング）のX

50

読み取り（例えば、X 性染色体への配列読み取りマッピング）に対する比が決定され得る。Y 読み取りの X 読み取りに対する比（Y/X 読み取り比）は、男性被験体および女性被験体中に存在する Y/X 比の公知の分布と比較され得る。各試料は、生成された Y/X 読み取り比に基づいて、男性または女性または曖昧と分類され得る。

【0086】

試料の性別分類は、性別分類のパフォーマンス測定基準（例えば、感度、特異度、陽性的中率、陰性的中率または曲線下面積）を決定するために、被験体の既知の性別と比較され得る。例えば、曖昧分類は、腫瘍が男性において X 染色体の増幅された部分を有する試料を分析することから生成され得、それによって、罹患していない男性集団における Y/X 読み取り比よりもはるかに低い Y/X 読み取り比を生じる。試料の性別分類が被験体の（患者の）既知の性別と一致しない場合、試料は、取り換えられたと具体的に疑われる。かかる結果は、試料の性別分類のための方法中に供給され得、その曖昧さをなくし得、取り換えが起きた場所（例えば、実験室設定または臨床設定）の指標を提供し得る。

【0087】

取り換えられた試料の同定情報（例えば、試料不一致または試料一致）ならびに X 染色体および Y 染色体を分析することに基づく性別不一致の同定情報は、検出された試料取り換えが生じた正確な状況を明らかにするために、近接試料（例えば、試料処理におけるある特定のステップにおいて隣同士であった試料）の記録を含有するデータベースと比較され得る。多くの場合、かかる比較は、試料同定情報をそれらの正確な被験体に再度割り当てることによる、同定された試料不一致の修正を可能にする。一部の 경우에는、同定された試料不一致の修正は、例えば、試料フィンガープリントが、アッセイされた他のあらゆる試料とも一致しない場合などには、可能でない場合がある。かかる状況は、外部パートナーから間違った試料を送られること、またはまだアッセイされていない試料との試料取り換えによって引き起こされ得る。かかる場合には、かかる不確定な試料は、データベースにおいてマークされ得、さらなる分析から排除され得る。

【0088】

一部の実施形態では、第 1 の複数の核酸分子を処理するステップは、第 1 の複数の核酸分子の結合測定を実施する工程を含み、第 2 の複数の核酸分子を処理するステップは、第 2 の複数の核酸分子の結合測定を実施する工程を含む。一部の実施形態では、複数の遺伝子座の各々における第 1 の複数の核酸分子の定量的尺度は、遺伝子座を含有する第 1 の複数の核酸分子の数を含み、複数の遺伝子座の各々における第 2 の複数の核酸分子の定量的尺度は、遺伝子座を含有する第 2 の複数の核酸分子の数を含む。例えば、結合測定は、複数の核酸分子中の複数の SNP の少なくとも一部分に対して選択的なプローブを使用して複数の核酸分子をアッセイすることによって得られ得る。一部の実施形態では、プローブは、複数の SNP の核酸配列との配列相補性を有する核酸分子である。一部の実施形態では、プローブは、プライマーまたは富化配列である核酸分子である。一部の実施形態では、アッセイは、アレイハイブリダイゼーションもしくはポリメラーゼ連鎖反応（PCR）、または核酸配列決定の使用を含む。

【0089】

一部の実施形態では、この方法は、複数の SNP の少なくとも一部分について、複数の核酸分子を富化するステップをさらに含む。一部の実施形態では、富化は、複数の核酸分子を増幅する工程を含む。例えば、複数の核酸分子は、選択的増幅（例えば、複数の SNP の核酸配列との配列相補性を有する核酸分子を含むプライマーまたはプローブのセットを使用することによる）によって増幅され得る。あるいはまたは組み合わせて、複数の核酸分子は、普遍的増幅（例えば、ユニバーサルプライマーを使用することによる）によって増幅され得る。一部の実施形態では、富化は、複数の核酸分子の少なくとも一部分を選択的に単離する工程を含む。

【0090】

複数の遺伝子座は、少なくとも約 10 の別個の常染色体一塩基多型（SNP）、少なくとも約 50 の別個の常染色体 SNP、少なくとも約 100 の別個の常染色体 SNP、少な

くとも約500の別個の常染色体SNP、少なくとも約1000の別個の常染色体SNP、少なくとも約5000の別個の常染色体SNP、少なくとも約10000の別個の常染色体SNP、少なくとも約50000の別個の常染色体SNP、少なくとも約100000の別個の常染色体SNP、少なくとも約500000の別個の常染色体SNP、少なくとも約100万の別個の常染色体SNP、少なくとも約200万の別個の常染色体SNP、少なくとも約300万の別個の常染色体SNP、少なくとも約400万の別個の常染色体SNP、少なくとも約500万の別個の常染色体SNP、少なくとも約1000万の別個の常染色体SNP、または約1000万よりも多い別個の常染色体SNPを含み得る。

【0091】

一部の実施形態では、試料不一致を同定するステップは、少なくとも約10%、少なくとも約20%、少なくとも約30%、少なくとも約40%、少なくとも約50%、少なくとも約60%、少なくとも約70%、少なくとも約80%、少なくとも約90%、少なくとも約95%、少なくとも約96%、少なくとも約97%、少なくとも約98%、少なくとも約99%、少なくとも約99.5%、少なくとも約99.6%、少なくとも約99.7%、少なくとも約99.8%、少なくとも約99.9%、少なくとも約99.99%または少なくとも約99.999%の感度で実施される。試料不一致を同定するステップの感度は、本開示の方法を使用して同定されると期待される試料不一致のパーセンテージとして測定または見積もりされ得る。感度は、ある特定の数の別個の遺伝子座（例えば、常染色体SNP）にわたり十分なカバレッジを得ており、試料品質の問題（例えば、部分的混入、例えば、試料混合）が存在しないという仮定の下で、測定または見積もりされ得る。

10

20

【0092】

一部の実施形態では、試料不一致を同定するステップは、少なくとも約10%、少なくとも約20%、少なくとも約30%、少なくとも約40%、少なくとも約50%、少なくとも約60%、少なくとも約70%、少なくとも約80%、少なくとも約90%、少なくとも約95%、少なくとも約96%、少なくとも約97%、少なくとも約98%、少なくとも約99%、少なくとも約99.5%、少なくとも約99.6%、少なくとも約99.7%、少なくとも約99.8%、少なくとも約99.9%、少なくとも約99.99%または少なくとも約99.999%の特異度で実施される。試料不一致を同定するステップの特異度は、本開示の方法を使用して同定されると期待される、不一致でない試料（例えば、試料一致）のパーセンテージとして測定または見積もりされ得る。特異度は、ある特定の数の別個の遺伝子座（例えば、常染色体SNP）にわたり十分なカバレッジを得ており、試料品質の問題（例えば、部分的混入、例えば、試料混合）が存在しないという仮定の下で、測定または見積もりされ得る。

30

【0093】

一部の実施形態では、試料不一致を同定するステップは、少なくとも約10%、少なくとも約20%、少なくとも約30%、少なくとも約40%、少なくとも約50%、少なくとも約60%、少なくとも約70%、少なくとも約80%、少なくとも約90%、少なくとも約95%、少なくとも約96%、少なくとも約97%、少なくとも約98%、少なくとも約99%、少なくとも約99.5%、少なくとも約99.6%、少なくとも約99.7%、少なくとも約99.8%、少なくとも約99.9%、少なくとも約99.99%または少なくとも約99.999%の陽性的中率（PPV）で実施される。試料不一致を同定するステップのPPVは、本開示の方法を使用して同定された試料不一致が真陽性である（例えば、試料の対が、その方法が試料のその対を不一致と同定したとしたときに、互いに真に不一致である）可能性として測定または見積もりされ得る。PPVは、ある特定の数の別個の遺伝子座（例えば、常染色体SNP）にわたり十分なカバレッジを得ており、試料品質の問題（例えば、部分的混入、例えば、試料混合）が存在しないという仮定の下で、測定または見積もりされ得る。

40

【0094】

一部の実施形態では、試料不一致を同定するステップは、少なくとも約10%、少なく

50

とも約20%、少なくとも約30%、少なくとも約40%、少なくとも約50%、少なくとも約60%、少なくとも約70%、少なくとも約80%、少なくとも約90%、少なくとも約95%、少なくとも約96%、少なくとも約97%、少なくとも約98%、少なくとも約99%、少なくとも約99.5%、少なくとも約99.6%、少なくとも約99.7%、少なくとも約99.8%、少なくとも約99.9%、少なくとも約99.99%または少なくとも約99.999%の陰性的中率(NPV)で実施される。試料不一致を同定するステップのNPVは、本開示の方法を使用して不一致でないと同定された試料(例えば、試料一致)が真陰性である(例えば、試料の対が、その方法が試料のその対を不一致でないと同定したとしたときに、互いに真に不一致でない)可能性として測定または見積もりされ得る。NPVは、ある特定の数の別個の遺伝子座(例えば、常染色体SNP)にわたり十分なカバレッジを得ており、試料品質の問題(例えば、部分的混入、例えば、試料混合)が存在しないという仮定の下で、測定または見積もりされ得る。

10

【0095】

一部の実施形態では、試料不一致を同定するステップは、少なくとも約0.5、少なくとも約0.6、少なくとも約0.7、少なくとも約0.75、少なくとも約0.8、少なくとも約0.85、少なくとも約0.9、少なくとも約0.95、少なくとも約0.96、少なくとも約0.97、少なくとも約0.98、少なくとも約0.99、少なくとも約0.995、少なくとも約0.996、少なくとも約0.997、少なくとも約0.998、少なくとも約0.999、少なくとも約0.9999または少なくとも約0.99999の、受信者動作特性(ROC)の曲線下面積(AUC)で実施される。

20

【0096】

一部の実施形態では、この方法は、第1の試料フィンガープリントと第2の試料フィンガープリントとの間の差異が所定の基準を満たさない場合に、試料一致を同定するステップをさらに含む。

【0097】

一部の実施形態では、試料一致を同定するステップは、少なくとも約10%、少なくとも約20%、少なくとも約30%、少なくとも約40%、少なくとも約50%、少なくとも約60%、少なくとも約70%、少なくとも約80%、少なくとも約90%、少なくとも約95%、少なくとも約96%、少なくとも約97%、少なくとも約98%、少なくとも約99%、少なくとも約99.5%、少なくとも約99.6%、少なくとも約99.7%、少なくとも約99.8%、少なくとも約99.9%、少なくとも約99.99%または少なくとも約99.999%の感度で実施される。試料一致を同定するステップの感度は、本開示の方法を使用して同定されると期待される試料一致のパーセンテージとして測定または見積もりされ得る。感度は、ある特定の数の別個の遺伝子座(例えば、常染色体SNP)にわたり十分なカバレッジを得ており、試料品質の問題(例えば、部分的混入、例えば、試料混合)が存在しないという仮定の下で、測定または見積もりされ得る。

30

【0098】

一部の実施形態では、試料一致を同定するステップは、少なくとも約10%、少なくとも約20%、少なくとも約30%、少なくとも約40%、少なくとも約50%、少なくとも約60%、少なくとも約70%、少なくとも約80%、少なくとも約90%、少なくとも約95%、少なくとも約96%、少なくとも約97%、少なくとも約98%、少なくとも約99%、少なくとも約99.5%、少なくとも約99.6%、少なくとも約99.7%、少なくとも約99.8%、少なくとも約99.9%、少なくとも約99.99%または少なくとも約99.999%の特異度で実施される。試料一致を同定するステップの特異度は、本開示の方法を使用して同定されると期待される、一致でない試料(例えば、試料不一致)のパーセンテージとして測定または見積もりされ得る。特異度は、ある特定の数の別個の遺伝子座(例えば、常染色体SNP)にわたり十分なカバレッジを得ており、試料品質の問題(例えば、部分的混入、例えば、試料混合)が存在しないという仮定の下で、測定または見積もりされ得る。

40

【0099】

50

一部の実施形態では、試料一致を同定するステップは、少なくとも約10%、少なくとも約20%、少なくとも約30%、少なくとも約40%、少なくとも約50%、少なくとも約60%、少なくとも約70%、少なくとも約80%、少なくとも約90%、少なくとも約95%、少なくとも約96%、少なくとも約97%、少なくとも約98%、少なくとも約99%、少なくとも約99.5%、少なくとも約99.6%、少なくとも約99.7%、少なくとも約99.8%、少なくとも約99.9%、少なくとも約99.99%または少なくとも約99.999%の陽性的中率（PPV）で実施される。試料一致を同定するステップのPPVは、本開示の方法を使用して同定された試料一致が真陽性である（例えば、試料の対が、その方法が試料のその対を一致と同定したとしたときに、互いに真に一致である）可能性として測定または見積もりされ得る。PPVは、ある特定の数の別個の遺伝子座（例えば、常染色体SNP）にわたり十分なカバレッジを得ており、試料品質の問題（例えば、部分的混入、例えば、試料混合）が存在しないという仮定の下で、測定または見積もりされ得る。

10

【0100】

一部の実施形態では、試料一致を同定するステップは、少なくとも約10%、少なくとも約20%、少なくとも約30%、少なくとも約40%、少なくとも約50%、少なくとも約60%、少なくとも約70%、少なくとも約80%、少なくとも約90%、少なくとも約95%、少なくとも約96%、少なくとも約97%、少なくとも約98%、少なくとも約99%、少なくとも約99.5%、少なくとも約99.6%、少なくとも約99.7%、少なくとも約99.8%、少なくとも約99.9%、少なくとも約99.99%または少なくとも約99.999%の陰性的中率（NPV）で実施される。試料一致を同定するステップのNPVは、本開示の方法を使用して一致でないと同定された試料（例えば、試料不一致）が真陰性である（例えば、試料の対が、その方法が試料のその対を一致でないと同定したとしたときに、互いに真に一致でない）可能性として測定または見積もりされ得る。NPVは、ある特定の数の別個の遺伝子座（例えば、常染色体SNP）にわたり十分なカバレッジを得ており、試料品質の問題（例えば、部分的混入、例えば、試料混合）が存在しないという仮定の下で、測定または見積もりされ得る。

20

【0101】

一部の実施形態では、試料一致を同定するステップは、少なくとも約0.5、少なくとも約0.6、少なくとも約0.7、少なくとも約0.75、少なくとも約0.8、少なくとも約0.85、少なくとも約0.9、少なくとも約0.95、少なくとも約0.96、少なくとも約0.97、少なくとも約0.98、少なくとも約0.99、少なくとも約0.995、少なくとも約0.996、少なくとも約0.997、少なくとも約0.998、少なくとも約0.999、少なくとも約0.9999または少なくとも約0.99999の、受信者動作特性（ROC）の曲線下面積（AUC）で実施される。

30

【0102】

一部の実施形態では、試料不一致を同定する方法は、第1の試料フィンガープリントと第2の試料フィンガープリントとの間の差異が所定の基準を満たすかどうかを決定するステップをさらに含む。所定の閾値は、1つまたは複数の対照被験体由来の1つまたは複数の試料から試料フィンガープリントを生成し、対照試料の変動性（同じ被験体内の、および（例えば、異なる性別の）異なる被験体間の変動性）に基づいて適切な所定の閾値を同定することによって生成され得る。

40

【0103】

所定の閾値は、試料不一致および／または試料一致を同定するステップの所望の感度、特異度、陽性的中率（PPV）、陰性的中率（NPV）または精度に基づいて調整され得る。例えば、所定の閾値は、試料不一致を同定するステップの高い感度が所望される場合、より低くなるように調整され得る。あるいは、所定の閾値は、試料不一致を同定するステップの高い特異度が所望される場合、より高くなるように調整され得る。所定の閾値は、対照被験体から得られた対照試料の受信者動作特性（ROC）の曲線下面積（AUC）を最大化するように調整され得る。所定の閾値は、試料不一致および／または試料一致を

50

同定するステップにおける偽陽性（F P）と偽陰性（F N）との間の所望のバランスを達成するように調整され得る。

【0104】

図3は、複数のアッセイされた生体試料から生成された試料フィンガープリントの比較の完全な可視化を示す。対角線に沿った濃い黒線は、取り違えなかった全ての試料を示す（例えば、試料一致）。例えば、かかる試料一致は、同じ患者に真に属すると同定される一致する患者同定情報（例えば、I D 番号、生年月日、性別など）を有する試料の対に対応し得る。対角線から離れた要素は、異なる被験体から得られたと推定される試料と類似しすぎている試料を示す。例えば、かかる試料不一致は、異なる患者から得られた可能性が高いと同定される一致する患者同定情報（例えば、I D 番号、生年月日、性別など）を有する試料の対に対応し得る（例えば、潜在的試料取り違え）。同定された試料不一致の場合、不一致の試料フィンガープリントは、試料不一致を同定および修正することを試みるために、不一致の患者同定情報（例えば、I D 番号、生年月日、性別など）を有する、データベース中に記憶された他の試料フィンガープリント（他の患者に属すると称される）と比較され得る。試料不一致は、試料フィンガープリントと関連する患者同定情報を、データベース中に見出される場合にそれらの正確な正体に一致するように交換するまたは更新することによって、修正され得る。不一致の試料の正確な正体が決定できない場合（例えば、データベース中に見出されない場合）、不一致の試料は、さらなるアッセイおよび処理から排除するようにマークされ得る。

【0105】

図4は、明確な内部試料不一致（例えば、試料取り違え）の一例を示し、このとき、2つの異なる被験体から得られた多数の生体試料に対して、アッセイの比較の可視化を実施した。対角線上の「分割された」四角の隣の対角線から離れたバーは、これらの2つの試料が差し替えられたことを示す（B L I B 0 0 3 6 6 および B L I B 0 0 3 6 7）。試料不一致は、試料フィンガープリントの対と関連する患者同定情報を、データベース中に見出されたのでそれらの正確な正体に一致するように交換するまたは更新することによって、修正され得る。

【0106】

図5は、明確な試料不一致（例えば、試料取り違え）および分解できない試料食い違いの一例の画像を示す。第1の患者（I D 番号 4 1 8 1）および第2の患者（I D 番号 4 1 7 5）から得られた組織試料を取り違えた。第3の患者（I D 番号 4 1 6 1）についての c f D N A 試料の1つは、第3の患者（I D 番号 4 1 6 1）由来であると推定される他の試料を含む他のあらゆる試料とも一致しない。第3の患者（I D 番号 4 1 6 1）について不一致の試料（試料食い違いを有する）の正確な正体が決定できない場合（例えば、データベース中に見出されなかった場合）、不一致の試料は、さらなるアッセイおよび処理から排除するようにマークされ得る。

【0107】

図6は、同じまたは異なる被験体（例えば、患者または人間）由来の試料の対の間の期待された遺伝子型類似性を示すプロットを示す。このプロットは、同じ人間から得られた試料と異なる人間から得られた試料との間を識別または差別化するために適切な閾値がどのように同定されるかを示している。取り違えられたと疑われる試料および低いカバレッジ（低い数の遺伝子型比較をもたらす）を有する試料を排除することによって潜在的試料不一致が説明された後で、分布は完全に分離される。

【0108】

例えば、取り違えられたと疑われる試料を排除することによって、同じ人間由来の試料の対の間の期待された遺伝子型類似性の分布は、（1番目の柱から3番目の柱へと）上方にシフトする。低いカバレッジ（低い数の遺伝子型比較をもたらす）を有する試料をさらに排除することによって、同じ人間由来の試料の対の間の期待された遺伝子型類似性の分布は、（3番目の柱から5番目の柱へと）上方にさらにシフトする。同様に、取り違えられたと疑われる試料を排除することによって、異なる人間由来の試料の対の間の期待され

た遺伝子型類似性の分布は、（２番目の柱から４番目の柱へと）下方にシフトする。低いカバレッジ（低い数の遺伝子型比較をもたらす）を有する試料をさらに排除することによって、異なる人間由来の試料の対の間の期待された遺伝子型類似性の分布は、（４番目の柱から６番目の柱へと）下方にさらにシフトする。したがって、この例では、同じ人間由来の試料の場合（取り違いおよび低いカバレッジを排除する）（５番目の柱）と異なる人間由来の試料の場合（取り違いおよび低いカバレッジを排除する）（６番目の柱）との間の閾値決定は、０．８の遺伝子型類似性において正確に実施され得る。異なる被験体から得られた試料フィンガープリントと比較して、同じ被験体から得られた試料フィンガープリントの類似性測定基準間には良好な分離が存在するので、遺伝子型類似性についての様々な可能なカットオフ値（所定の基準）が、試料一致および／または試料不一致を正確に決定するために使用され得る。所定の基準は、例えば、異なるが関連する被験体から得られた試料を分析する場合、偽陽性一致コールの確率を回避または最小化するために、比較的高い値に設定され得る。

10

【０１０９】

試料不一致を決定するための所定の基準は、２つの試料フィンガープリント間の遺伝子型類似性における差異が所定の閾値よりも大きいことであり得る。かかる所定の閾値は、例えば、少なくとも約０．０５、少なくとも約０．１、少なくとも約０．１５、少なくとも約０．２、少なくとも約０．２５、少なくとも約０．３、少なくとも約０．３５、少なくとも約０．４、少なくとも約０．４５、少なくとも約０．５、少なくとも約０．５５、少なくとも約０．６、少なくとも約０．６５、少なくとも約０．７、少なくとも約０．７５、少なくとも約０．８、少なくとも約０．８５または少なくとも約０．９の、遺伝子型類似性における差異であり得る。

20

【０１１０】

同様に、試料一致を決定するための所定の基準は、２つの試料フィンガープリント間の遺伝子型類似性における差異が所定の閾値以下であることであり得る。かかる所定の閾値は、例えば、約０．０５以下、約０．１以下、約０．１５以下、約０．２以下、約０．２５以下、約０．３以下、約０．３５以下、約０．４以下、約０．４５以下、約０．５以下、約０．５５以下、約０．６以下、約０．６５以下、約０．７以下、約０．７５以下、約０．８以下、約０．８５以下または約０．９以下の、遺伝子型類似性における差異であり得る。

30

【０１１１】

図７は、複数のアッセイされたDNA試料についてのジェンダーコールの比較を示す。X読み取りはX軸上に示され、Y読み取りはY軸上に示される。青色の試料は、男性被験体から得られたと推定され、赤色の試料は、女性被験体から得られたと推定され、灰色の試料は、かかる情報が入手不能であった。閾値ラインを十分に上回る位置にあるデータポイントの第１のセットは、男性としてコールされ、閾値ラインを十分に下回る位置にあるデータポイントの第２のセットは、女性としてコールされる。このプロットは、試料不一致として同定された（例えば、取り違いられたと同定された）試料に対応する、閾値ラインの下に位置する少数の青色のデータポイントおよび閾値の上に位置する少数の赤色のデータポイントを示している。閾値ラインのちょうど上にあるデータポイントは、X染色体の大部分が重複したがん患者から得られた。

40

【０１１２】

コンピュータシステム

本開示は、本開示の方法を実装するようにプログラミングされるコンピュータシステムを提供する。図８は、例えば、核酸分子を処理して、複数の遺伝子座の各々における核酸分子の定量的尺度を含む試料フィンガープリントを生成し、２つの試料フィンガープリント間の差異を決定し、２つの試料フィンガープリント間の差異が所定の基準を満たす場合に、試料不一致を同定するように、プログラミングされるまたは他の方法で構成されるコンピュータシステム８０１を示す。コンピュータシステム８０１は、本開示の分析、計算および生成の種々の態様、例えば、核酸分子を処理して、複数の遺伝子座の各々における

50

核酸分子の定量的尺度を含む試料フィンガープリントを生成するステップ、2つの試料フィンガープリント間の差異を決定するステップ、および2つの試料フィンガープリント間の差異が所定の基準を満たす場合に、試料不一致を同定するステップなどを調節できる。コンピュータシステム801は、ユーザーの電子デバイス、または電子デバイスに関してリモートに位置するコンピュータシステムであり得る。電子デバイスは、モバイル電子デバイスであり得る。

【0113】

コンピュータシステム801は、シングルコアもしく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であり得る、中央処理装置(CPU、本明細書で「プロセッサ」および「コンピュータプロセッサ」とも呼ばれる)805を含む。コンピュータシステム801は、メモリまたはメモリロケーション810(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)、電子記憶装置815(例えば、ハードディスク)、1つまたは複数の他のシステムとの通信のための通信インターフェース820(例えば、ネットワークアダプタ)、ならびに周辺デバイス825、例えば、キャッシュ、他のメモリ、データ記憶および/または電子ディスプレイアダプタもまた含む。メモリ810、記憶装置815、インターフェース820および周辺デバイス825は、通信バス(実線)(例えば、マザーボード)を介して、CPU805と通信する。記憶装置815は、データを記憶するためのデータ記憶装置(またはデータリポジトリ)であり得る。コンピュータシステム801は、通信インターフェース820に補助されて、コンピュータネットワーク(「ネットワーク」)830に動作可能に連結され得る。ネットワーク830は、インターネット、インターネットおよび/もしくはエクストラネット、またはインターネットと通信するイントラネットおよび/もしくはエクストラネットであり得る。ネットワーク830は、一部の場合では、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830は、分散型コンピューティング、例えば、クラウドコンピューティングを可能にできる、1つまたは複数のコンピュータサーバーを含み得る。例えば、1つまたは複数のコンピュータサーバーは、ネットワーク830を経由したクラウドコンピューティング(「クラウド」)が、本開示の分析、計算および生成の種々の態様、例えば、核酸分子を処理して、複数の遺伝子座の各々における核酸分子の定量的尺度を含む試料フィンガープリントを生成するステップ、2つの試料フィンガープリント間の差異を決定するステップ、および2つの試料フィンガープリント間の差異が所定の基準を満たす場合に、試料不一致を同定するステップなどを実施するのを可能にし得る。かかるクラウドコンピューティングは、クラウドコンピューティングプラットフォーム、例えば、Amazon Web Services(AWS)、Microsoft Azure、Google Cloud PlatformおよびIBMクラウドなどによって提供され得る。一部の場合では、コンピュータシステム801に補助されたネットワーク830は、コンピュータシステム801に連結されたデバイスがクライアントまたはサーバーとして挙動することを可能にし得るピアツーピアネットワークを実装し得る。

【0114】

CPU805は、プログラムまたはソフトウェアで具体化され得る機械可読命令のシーケンスを実行できる。命令は、メモリロケーション、例えば、メモリ810中に記憶され得る。命令は、CPU805に向けられ得、これは引き続いて、本開示の方法を実装するようにCPU805をプログラミングし得るまたは他の方法で構成し得る。CPU805によって実施される操作の例には、フェッチ、デコード、実行およびライトバックが含まれ得る。

【0115】

CPU805は、回路、例えば、集積回路の一部であり得る。システム801の1つまたは複数の他のコンポーネントは、回路中に含まれ得る。一部の場合では、回路は、特定用途向け集積回路(ASIC)である。

【0116】

記憶装置815は、ファイル、例えば、ドライバ、ライブラリおよび保存されたプログ

ラムを記憶できる。記憶装置 815 は、ユーザーデータ、例えば、ユーザーの好みおよびユーザープログラムを記憶できる。コンピュータシステム 801 は、一部の 경우에는、例えば、イントラネットまたはインターネットを介してコンピュータシステム 801 と通信するリモートサーバー上に位置する、コンピュータシステム 801 に対して外部の 1 つまたは複数のさらなるデータ記憶装置を含み得る。

【0117】

コンピュータシステム 801 は、ネットワーク 830 を介して、1 つまたは複数のリモートコンピュータシステムと通信できる。例えば、コンピュータシステム 801 は、ユーザー（例えば、医師、看護師、介護者、患者または被験体）のリモートコンピュータシステムと通信できる。リモートコンピュータシステムの例には、パーソナルコンピュータ（例えば、ポータブル PC）、スレートもしくはタブレット PC（例えば、Apple（登録商標）iPad（登録商標）、Samsung（登録商標）Galaxy Tab）、電話、スマートフォン（例えば、Apple（登録商標）iPhone（登録商標）、Android 有効化デバイス、Blackberry（登録商標））または携帯情報端末が含まれる。ユーザーは、ネットワーク 830 を介してコンピュータシステム 801 にアクセスできる。

【0118】

本明細書に記載される方法は、コンピュータシステム 801 の電子記憶場所上、例えば、メモリ 810 または電子記憶装置 815 上などに記憶された機械（例えば、コンピュータプロセッサ）実行可能コードによって実装され得る。機械実行可能コードまたは機械可読コードは、ソフトウェアの形態で提供され得る。使用の間に、コードは、プロセッサ 805 によって実行され得る。一部の 경우에는、コードは、記憶装置 815 から検索され得、プロセッサ 805 による即座のアクセスのために、メモリ 810 上に記憶され得る。一部の状況では、電子記憶装置 815 は、除外され得、機械実行可能命令は、メモリ 810 上に記憶される。

【0119】

コードは、コードを実行するように適応されたプロセッサを有する機械での使用のために事前コンパイルおよび構成され得、またはランタイムの間にコンパイルされ得る。コードは、事前コンパイル方式またはアズコンパイルド（*as-compiled*）方式でコードの実行を可能にするように選択され得るプログラミング言語で供給され得る。

【0120】

本明細書で提供されるシステムおよび方法の態様、例えば、コンピュータシステム 801 は、プログラミングで具体化され得る。テクノロジーの種々の態様は、典型的には、ある型の機械可読媒体上に保持されるまたは具体化される機械（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の、「製品」または「製造品」とあると考えられ得る。機械実行可能コードは、電子記憶装置、例えば、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上に記憶され得る。「記憶」型媒体には、ソフトウェアプログラミングのためにいつでも非一過性記憶を提供し得る、コンピュータ、プロセッサなどの有形メモリのうちいずれかもしくは全て、またはその関連モジュール、例えば、種々の半導体メモリ、テープドライブ、ディスクドライブなどが含まれ得る。ソフトウェアの全てまたは部分は、時折、インターネットまたは種々の他の電気通信ネットワークを介して通信され得る。例えば、かかる通信は、1 つの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、例えば、管理サーバー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ーのコンピュータプラットフォームへの、ソフトウェアのロードを可能にし得る。したがって、ソフトウェア要素を運び得る別の型の媒体には、例えば、ローカルデバイス間の物理インターフェースを横断して、有線および光学地上通信線ネットワークを介して、ならびに種々のエアリンク（*air-link*）を通じて使用される、光波、電波および電磁波が含まれる。かかる波を伝える物理要素、例えば、有線または無線リンク、光リンクなどもまた、ソフトウェアを運ぶ媒体とみなされ得る。本明細書で使用される場合、非一過性

有形「記憶」媒体に限定されない限り、コンピュータまたは機械「可読媒体」などの用語は、実行のために命令をプロセッサに提供することに関与する任意の媒体を指す。

【0121】

したがって、機械可読媒体、例えば、コンピュータ実行可能コードは、有形記憶媒体、搬送波媒体または物理伝送媒体が含まれるがこれらに限定されない多くの形態をとり得る。非揮発性記憶媒体には、例えば、図面に示される、例えば、データベースなどを実装するために使用され得る任意のコンピュータ（複数可）など中の記憶デバイスのいずれかなどの、光ディスクまたは磁気ディスクが含まれる。揮発性記憶媒体には、ダイナミックメモリ、例えば、かかるコンピュータプラットフォームのメインメモリが含まれる。有形伝送媒体には、コンピュータシステム内のバスを構成するワイヤを含む、同軸ケーブル；銅線および光ファイバーが含まれる。搬送波伝送媒体は、電気もしくは電磁気信号、または音波もしくは光波、例えば、無線周波数（RF）および赤外線（IR）データ通信の間に生成されるものの形態をとり得る。したがって、一般的な形態のコンピュータ可読媒体には、例えば、以下が含まれる：フロッピー（登録商標）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、任意の他の磁気媒体、CD-ROM、DVDもしくはDVD-ROM、任意の他の光媒体、パンチカード紙テープ、孔のパターンを用いた任意の他の物理記憶媒体、RAM、ROM、PROMおよびEPROM、FLASH（登録商標）-EPROM、任意の他のメモリチップもしくはカートリッジ、データもしくは命令を伝達する搬送波、かかる搬送波を伝達するケーブルもしくはリンク、またはコンピュータがプログラミングコードおよび／もしくはデータを読み取ることができる任意の他の媒体。これらの形態のコンピュータ可読媒体の多くは、実行のために1つまたは複数の命令の1つまたは複数の配列をプロセッサに伝えることに関与し得る。

【0122】

コンピュータシステム801は、例えば、複数の遺伝子座の各々における核酸分子の定量的尺度を含む生成された試料フィンガープリント、2つの試料フィンガープリント間の決定された差異、および同定された試料不一致を提供するためのユーザーインターフェース（UI）840を含む電子ディスプレイ835を含み得る、またはそれと通信し得る。UIの例には、限定なしに、グラフィカルユーザーインターフェース（GUI）およびウェブベースのユーザーインターフェースが含まれる。

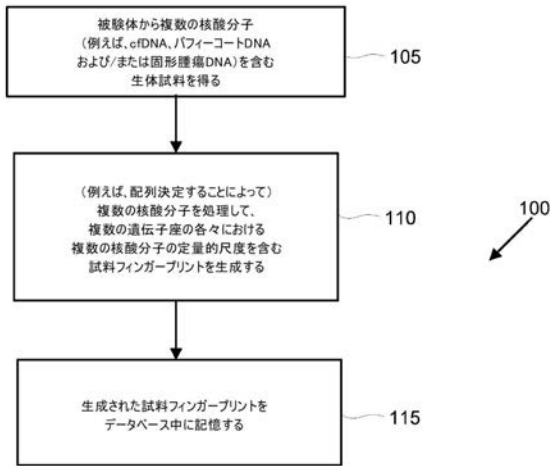
【0123】

本開示の方法およびシステムは、1つまたは複数のアルゴリズムによって実装され得る。アルゴリズムは、中央処理装置805による実行の際に、ソフトウェアによって実装され得る。アルゴリズムは、例えば、核酸分子を処理して、複数の遺伝子座の各々における核酸分子の定量的尺度を含む試料フィンガープリントを生成し、2つの試料フィンガープリント間の差異を決定し、2つの試料フィンガープリント間の差異が所定の基準を満たす場合に、試料不一致を同定することができる。

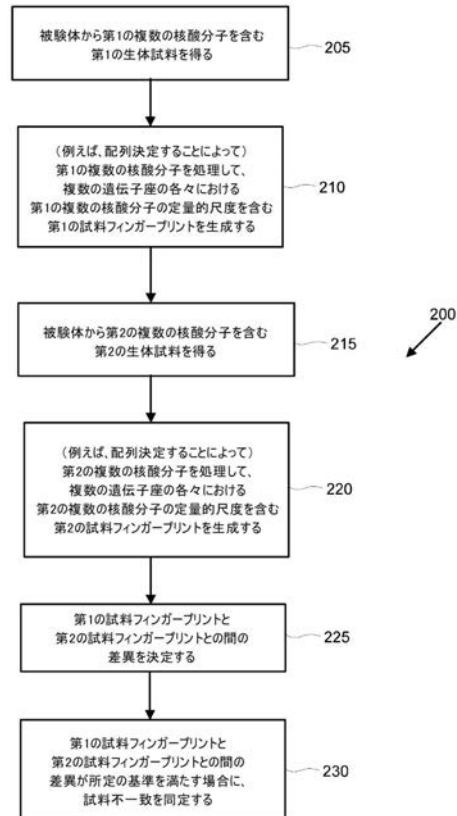
【0124】

本発明の好ましい実施形態が本明細書に示され記載されてきたが、かかる実施形態が例として提供されているに過ぎないことが、当業者に明らかである。本発明は、本明細書内に提供される具体的な例によって限定されないことが意図される。本発明は、上述の明細書を参照して記載されてきたが、本明細書の実施形態の記載および説明は、限定的な意味で解釈されることを意味しない。多数のバリエーション、変化および置換が、本発明から逸脱することなしに、当業者に直ちに明らかである。さらに、本発明の全ての態様が、種々の条件および変数に依存する本明細書に示される具体的な描写、構成または相対的割合に限定されないことを理解すべきである。本明細書に記載される本発明の実施形態に対する種々の代替法が、本発明を実施する際に使用され得ることを理解すべきである。したがって、本発明は、任意のかかる代替法、改変、バリエーションまたは均等物も当然カバーすることが意図される。以下の特許請求の範囲が本発明の範囲を定義すること、ならびにこれらの特許請求の範囲およびその均等物の範囲内の方法および構造がそれによってカバーされることが意図される。

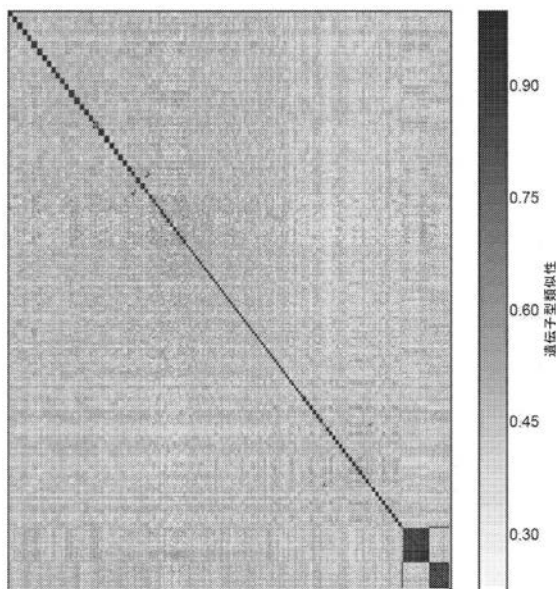
【図 1】



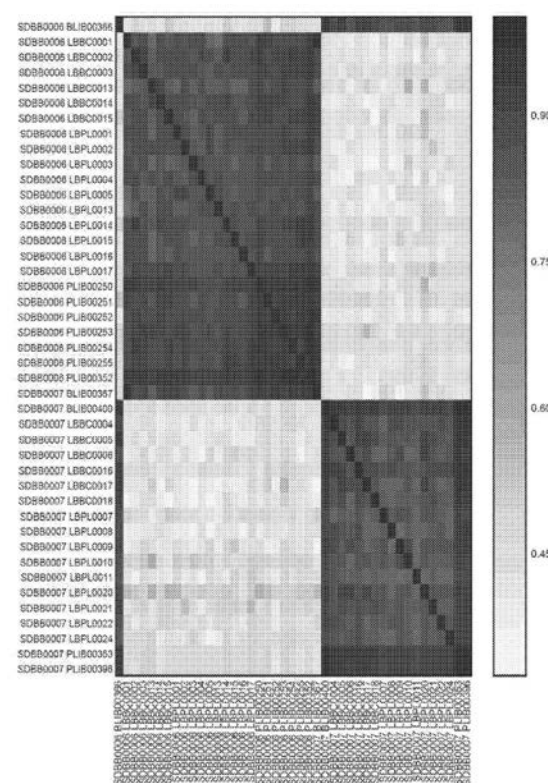
【図 2】



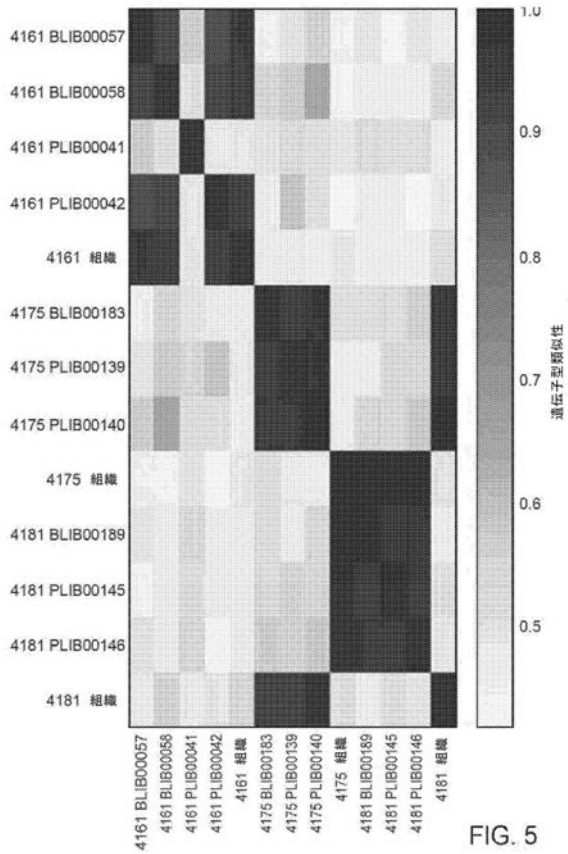
【図 3】



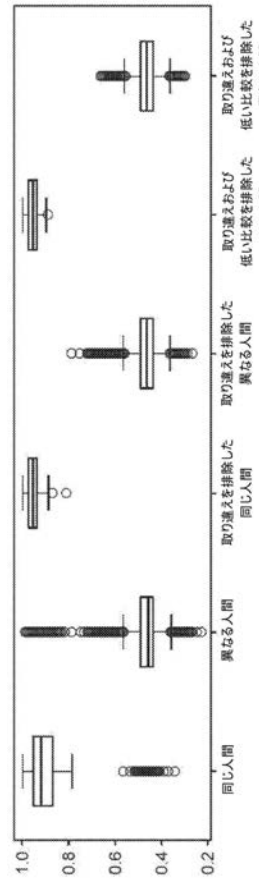
【図 4】



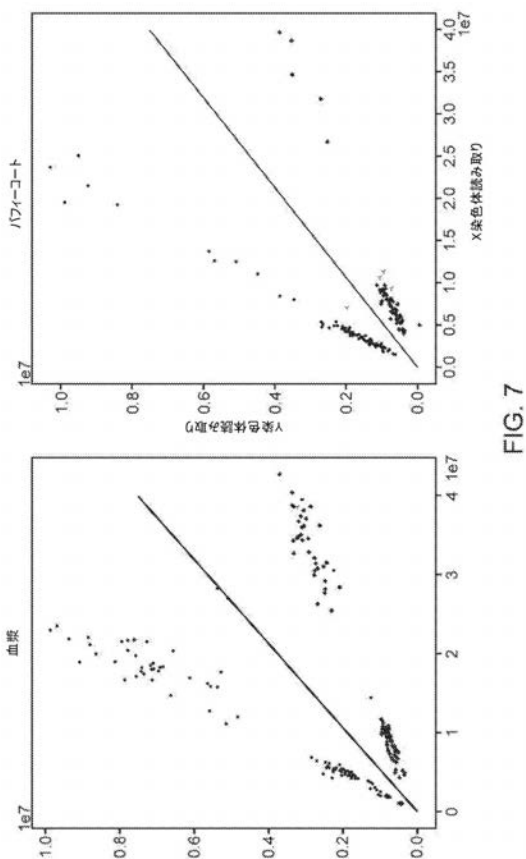
【図 5】



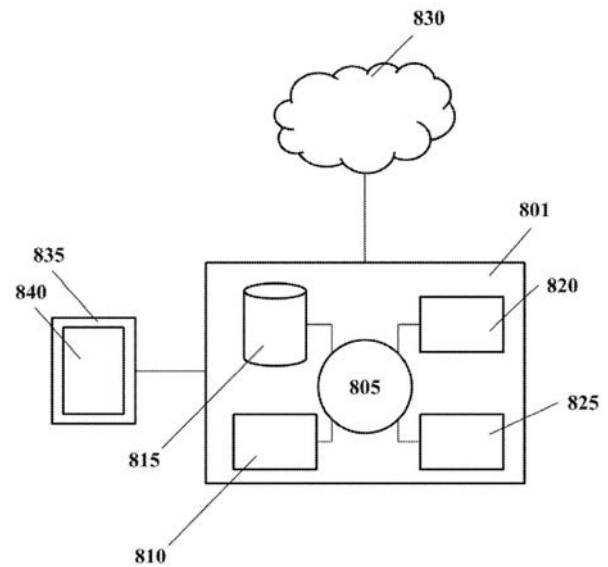
【図 6】



【図 7】



【図 8】



【国際調査報告】

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US19/35871

A. CLASSIFICATION OF SUBJECT MATTER

IPC - C40B 30/00; C12Q 1/68, 1/6827, 1/6883 (2019.01)

CPC - C12Q 1/68, 1/6827, 1/6883, 1/6886

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

See Search History document

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

See Search History document

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

See Search History document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X — Y	YOUSSEFI et al. A SNP Panel For Identification Of DNA And RNA Specimens . BMC Genomics. 25 January 2018, 19:30; pp 1-12; abstract; page 2, column 2, third paragraph – page 3, column 1, first paragraph; page 4, column 1, first paragraph; page 6; column 2, third paragraph; page 9, column 2, second paragraph Figure 1. DOI 10.1186/s12864-018-4482-7	1-4, 9/3, 10/9/3, 11/10/9/3, 12/10/9/3, 13/10/9/3, 14/9/3, 22/3, 23/3, 24/3, 28/24/3, 40/3, 41/40/3, 42/40/3, 43/40/3, 44/40/3, 45/40/3, 48 — 5/3, 6/3, 7/3, 8/3, 15/3, 16/15/3, 25/24/3, 26/24/3, 27/24/3, 29/28/24/3, 30/28/24/3, 31/28/24/3
Y	US 2011/0230358 A1 (RAVA) 02 March 2016; abstract; paragraphs [0008], [0055], [0150], [0167]	5/3, 6/3, 7/3, 25/24/3, 26/24/3, 27/24/3, 29/28/24/3, 30/28/24/3, 31/28/24/3
Y	US 2016/0060701 A1 (THE REGENTS OF THE UNIVERSITY OF COLORADO) 03 March 2016; paragraphs [0026], [0045], [0048], [0073], [0186]	8/3, 15/3, 16/15/3

☐ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"D" document cited by the applicant in the international application

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

02 August 2019 (02.08.2019)

Date of mailing of the international search report

27 AUG 2019

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents
P.O. Box 1450, Alexandria, Virginia 22313-1450
Facsimile No. 571-273-8300

Authorized officer

Shane Thomas

Telephone No. PCT Helpdesk: 571-272-4300

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US19/35871

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☒ Claims Nos.: 17-21, 32-39, 46-47
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

フロントページの続き

(81)指定国・地域 AP(BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), EA(AM, AZ, BY, KG, KZ, RU, TJ, TM), EP(AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT

(特許庁注：以下のものは登録商標)

1. ANDROID

(74)代理人 230113332

弁護士 山本 健策

- (72)発明者 ロバートソン, アレクサンダー デ ヨング
アメリカ合衆国 カリフォルニア 94103, サンフランシスコ, ブラナン ストリート
1000, スイート 402
- (72)発明者 スリヴァス, ロヒト カンナパン
アメリカ合衆国 カリフォルニア 94103, サンフランシスコ, ブラナン ストリート
1000, スイート 402
- (72)発明者 ウィルソン, ティモシー ジョゼフ
アメリカ合衆国 カリフォルニア 94103, サンフランシスコ, ブラナン ストリート
1000, スイート 402
- (72)発明者 ピーターマン, ニール
アメリカ合衆国 カリフォルニア 94103, サンフランシスコ, ブラナン ストリート
1000, スイート 402
- (72)発明者 ランバート, ニコル ジャシンダ
アメリカ合衆国 カリフォルニア 94103, サンフランシスコ, ブラナン ストリート
1000, スイート 402
- (72)発明者 テツカン, ハルク
アメリカ合衆国 カリフォルニア 94103, サンフランシスコ, ブラナン ストリート
1000, スイート 402

Fターム(参考) 4B063 QA01 QA13 QQ42 QQ52 QS24 QS28 QS36 QX02