

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号

特許第7217224号

(P7217224)

(45)発行日 令和5年2月2日(2023.2.2)

(24)登録日 令和5年1月25日(2023.1.25)

(51)国際特許分類

F I

C 1 2 Q 1/6869(2018.01)

C 1 2 Q 1/6869

Z

C 1 2 Q 1/6883(2018.01)

C 1 2 Q 1/6883

Z Z N A

C 1 2 Q 1/6886(2018.01)

C 1 2 Q 1/6886

Z

A 6 1 K 45/00 (2006.01)

A 6 1 K 45/00

A 6 1 P 35/00 (2006.01)

A 6 1 P 35/00

請求項の数 57 (全106頁)

(21)出願番号 特願2019-511456(P2019-511456)

(86)(22)出願日 平成29年8月24日(2017.8.24)

(65)公表番号 特表2019-526257(P2019-526257
A)

(43)公表日 令和1年9月19日(2019.9.19)

(86)国際出願番号 PCT/US2017/048434

(87)国際公開番号 WO2018/039463

(87)国際公開日 平成30年3月1日(2018.3.1)

審査請求日 令和2年7月7日(2020.7.7)

(31)優先権主張番号 62/379,593

(32)優先日 平成28年8月25日(2016.8.25)

(33)優先権主張国・地域又は機関

米国(US)

(31)優先権主張番号 62/481,538

(32)優先日 平成29年4月4日(2017.4.4)

最終頁に続く

(73)特許権者 515138791

レゾリューション バイオサイエンス,
インコーポレイテッド
アメリカ合衆国 ワシントン 9 8 0 3 3
, カークランド, カークランド ウェイ
5 5 0, スイート 2 0 0

(74)代理人 100078282

弁理士 山本 秀策

(74)代理人 100113413

弁理士 森下 夏樹

(74)代理人 100181674

弁理士 飯田 貴敏

(74)代理人 100181641

弁理士 石川 大輔

(74)代理人 230113332

最終頁に続く

(54)【発明の名称】 DNA試料中のゲノムコピー変化の検出方法

(57)【特許請求の範囲】

【請求項1】

複数のゲノムDNA断片を含む試験試料からのDNA標的領域上で遺伝子解析を実施する方法であって、

(a) 複数のプールからユニーク試料タグ領域のプールを選択することであって、前記選択されたプールが前記試験試料に対してユニークであること、

(b) 1セットのアダプタに前記試験試料を接触させることにより、ゲノムDNAライブラリを生成することであって、そこで

前記セットのアダプタの各アダプタはユニーク試料タグ領域のプールから選択される1個の試料タグ領域を含み、

前記ゲノムDNAライブラリは複数のDNAライブラリ断片を含み、

各前記DNAライブラリ断片は、アダプタと付着した前記試験試料からのゲノムDNA断片を含む、前記生成すること、

(c) 前記ゲノムDNAライブラリを、前記DNA標的領域に特異的に結着する複数のキャプチャプローブと接触させることにより、前記キャプチャプローブと前記DNA標的領域を含むDNAライブラリ断片との間に複合体を形成すること、ならびに

(d) 前記DNA標的領域を含む前記ゲノムDNA断片の定量的遺伝子解析を実施すること、

を備え、

そこで前記アダプタは増幅領域、試料タグ領域、及びアンカー領域を含むDNAポリヌク

レオチドであり、

そこで前記増幅領域はPCR増幅についてプライマー認識部位として機能することが可能であるポリヌクレオチド配列を含み、

そこで前記試料タグ領域は、前記試験試料を同定するポリヌクレオチド配列を含み、

そこで前記アンカー領域は前記ゲノムDNA断片に付着することが可能なポリヌクレオチド配列を含む、

前記方法。

【請求項2】

前記方法が、一塩基多様体(SNV)、40ヌクレオチド長未満の挿入、40ヌクレオチド長未満のDNA領域の欠失、及び/またはコピー数における変化から選択される、疾患状態を示す遺伝的变化を同定する、請求項1に記載の前記方法。

10

【請求項3】

疾患状態を示す前記遺伝的变化は、コピー数における変化である、請求項2に記載の前記方法。

【請求項4】

前記試験試料は、組織診試料である、請求項1から3のいずれかに記載の前記方法。

【請求項5】

前記組織診試料は、腫瘍、または腫瘍であると疑われる組織から採取された試料である、請求項4に記載の前記方法。

【請求項6】

前記ゲノムDNA断片は、無細胞DNA(cfDNA)または細胞DNAである、請求項1から5のいずれかに記載の前記方法。

20

【請求項7】

前記ゲノムDNA断片は、前記試験試料から単離され、前記試験試料は、羊水、血液、血漿、血清、精液、リンパ液、脳脊髄液、眼液、尿、唾液、便、粘液、及び汗からなる群から選択される生体試料である、請求項6に記載の前記方法。

【請求項8】

前記ゲノムDNA断片は、

(i) 前記試験試料から細胞DNAを単離すること、

(ii) 前記細胞DNAを断片化し、前記ゲノムDNA断片を取得すること、

を備えるステップによって取得される、請求項1から7のいずれかに記載の前記方法。

30

【請求項9】

ステップ(ii)は、前記細胞DNAを少なくとも1個の消化酵素と接触させることによって実施される、請求項8に記載の前記方法。

【請求項10】

ステップ(ii)は、機械的応力を前記細胞DNAへ加えることによって実施される、請求項8に記載の前記方法。

【請求項11】

前記機械的応力は、前記細胞DNAを超音波処理することによって加えられる、請求項10に記載の前記方法。

40

【請求項12】

前記アダプタは、前記ユニークゲノムDNA断片の前記同定を容易にするユニーク分子識別子(UMI)をさらに含む、請求項1から11のいずれかに記載の前記方法。

【請求項13】

前記増幅領域は、10から50の間のヌクレオチド長にある、請求項1から12のうちのいずれか1項に記載の前記方法。

【請求項14】

前記増幅領域は、20から30の間のヌクレオチド長にある、請求項1から13のうちのいずれか1項に記載の前記方法。

【請求項15】

50

前記増幅領域は、25ヌクレオチド長にある、請求項1から14のうちのいずれか1項に記載の前記方法。

【請求項16】

前記試料タグ領域は、5から50の間のヌクレオチド長にある、請求項1から15のうちのいずれか1項に記載の前記方法。

【請求項17】

前記試料タグ領域は、5から15の間のヌクレオチド長にある、請求項16に記載の前記方法。

【請求項18】

前記試料タグ領域は、8ヌクレオチド長にある、請求項16に記載の前記方法。

10

【請求項19】

前記UMIは前記試料タグ領域とUMI増倍管を含み、前記UMI増倍管は、前記試料タグ領域に隣接する、または前記試料タグ領域内に含まれる、請求項12から18のうちのいずれか1項に記載の前記方法。

【請求項20】

前記UMI増倍管は、1から5の間のヌクレオチド長にある、請求項19に記載の前記方法。

【請求項21】

前記UMI増倍管は、3ヌクレオチド長にあり、64個の可能なヌクレオチド配列のうちの1個を含む、請求項19に記載の前記方法。

20

【請求項22】

前記アンカー領域は、1から50の間のヌクレオチド長にある、請求項1から21のうちのいずれか1項に記載の前記方法。

【請求項23】

前記アンカー領域は、5から25の間のヌクレオチド長にある、請求項22に記載の前記方法。

【請求項24】

前記アンカー領域は、10ヌクレオチド長にある、請求項22または23に記載の前記方法。

【請求項25】

30

ステップ(上)は、前記試験試料のゲノムDNA断片の一方の末端に第1のアダプタを付着させることと、前記ゲノムDNA断片のもう一方の末端に第2のアダプタを付着させることを備える、請求項1から24のいずれか1項に記載の前記方法。

【請求項26】

前記ゲノムDNA断片は、前記第1及び第2のアダプタを付着させる前に、末端修復される、請求項25に記載の前記方法。

【請求項27】

前記セットのアダプタの各アダプタの前記増幅領域は、同一のヌクレオチド配列を含む、請求項1から26のいずれかに記載の前記方法。

【請求項28】

40

前記ユニーク試料タグ領域のプールは、2から1,000個の間のユニーク試料タグ領域配列を含む、請求項1から27のいずれかに記載の前記方法。

【請求項29】

前記ユニーク試料タグ領域のプールは、50から500個の間のユニーク試料タグ領域配列を含む、請求項28に記載の前記方法。

【請求項30】

前記ユニーク試料タグ領域のプールは、100から400個の間のユニーク試料タグ領域配列を含む、請求項28に記載の前記方法。

【請求項31】

前記ユニーク試料タグ領域のプールは、200から300個の間のユニーク試料タグ領域

50

配列を含む、請求項 28 に記載の前記方法。

【請求項 32】

前記ユニーク試料タグ領域のプールの各試料タグ領域は、8ヌクレオチド長にある、請求項 28 に記載の前記方法。

【請求項 33】

各ユニーク試料タグ領域配列は、少なくとも2塩基だけいずれかの他のユニーク試料タグ領域配列から分離する、請求項 28 から 32 のいずれかに記載の前記方法。

【請求項 34】

前記ユニーク試料タグ領域のプールは、240個のユニーク試料タグ領域配列を含む、請求項 1 から 33 のいずれか1項に記載の前記方法。

10

【請求項 35】

前記試料タグ領域が、これに付着された前記ゲノムDNA断片を同定するために機能する、請求項 1 から 34 のいずれかに記載の前記方法。

【請求項 36】

前記アダプタは、前記ゲノム断片を同定するためのユニーク配列の数を増加させるUMI増倍管をさらに備える、請求項 1 から 35 のうちのいずれか1項に記載の前記方法。

【請求項 37】

前記セットのアダプタの各アダプタの前記アンカー領域は、4個のヌクレオチド配列のうちの1個を含み、所与の配列の各試料タグ領域は、所与の配列の前記4個のアンカー領域のうちの1個のみに対合される、請求項 1 から 36 のいずれかに記載の前記方法。

20

【請求項 38】

前記セットのアダプタの各アダプタの前記増幅領域は、同一のヌクレオチド配列を含み、前記ユニーク試料タグ領域のプールの各試料タグ領域は、8ヌクレオチド長にあり、各ユニーク試料タグ領域配列は、少なくとも2塩基だけいずれかの他のユニーク試料タグ領域配列から分離し、

前記セットのアダプタの各アダプタは、前記試料タグ領域に隣接する、または前記試料タグ領域内に含まれる前記UMI増倍管を備え、前記セットのアダプタの各アダプタの前記UMI増倍管は、3ヌクレオチド長にあり、所与の配列の前記UMI増倍管は、所与の配列の1個の試料タグ領域に対合され、

前記セットのアダプタの各アダプタの前記アンカー領域は、4個のヌクレオチド配列のうちの1個を含み、所与の配列の各試料タグ領域は、所与の配列の前記4個のアンカー領域のうちの1個のみに対合される、請求項 1 から 37 のいずれかに記載の前記方法。

30

【請求項 39】

前記ゲノムDNA断片と前記第1及び第2のアダプタのそれぞれを付着させる前記ステップは、

(i) アンカー領域の少なくとも一部を含むオリゴヌクレオチドを前記ゲノムDNA断片へ付着させることであって、そこで前記オリゴヌクレオチドはパートナー鎖によって二本鎖にされる5'リン酸化付着鎖を含むDNA二本鎖の部分であり、そこで前記パートナー鎖は化学修飾による前記ゲノムDNA断片への付着からその3'末端で遮断され、そこで前記付着鎖は前記ゲノムDNA断片に付着する、前記付着させること、

40

(ii) 前記付着鎖に付着する前記ゲノムDNA断片を、DNAライゲーションに適切な条件下で、アダプタ配列の全長を符号化するDNAオリゴヌクレオチド、T4ポリヌクレオチドキナーゼ、Taq DNAリガーゼ、及びBstポリメラーゼ全長と接触させること、

を備え、

それらにより、前記第1及び第2のアダプタを前記ゲノムDNA断片に付着させる、請求項 25 から 38 のいずれか1項に記載の前記方法。

【請求項 40】

前記DNA標的領域は、コピー数における変化について解析される、請求項 1 から 39 の

50

いずれかに記載の前記方法。

【請求項 4 1】

前記キャプチャプローブと前記 D N A 標的領域を含む D N A ライブラリ断片との間に形成される前記複合体を精製することをさらに含む、請求項 1 から 4 0 のいずれか 1 項に記載の前記方法。

【請求項 4 2】

前記キャプチャプローブと前記 D N A 標的領域を含む D N A ライブラリ断片との間に形成される前記複合体を精製することをさらに含み、前記 D N A 標的領域を含む前記 D N A ライブラリ断片のプライマー伸長及び/または増幅を実施する、請求項 1 から 4 1 のいずれか 1 項に記載の前記方法。

10

【請求項 4 3】

前記キャプチャプローブと前記 D N A 標的領域を含む D N A ライブラリ断片との間に形成される前記複合体を精製することをさらに含み、前記 D N A 標的領域を含む前記 D N A ライブラリ断片のプライマー伸長及び増幅を実施する、請求項 1 から 4 2 のいずれか 1 項に記載の前記方法。

【請求項 4 4】

ステップ (d) は、複数のシーケンシングリードを生成するために、前記 D N A 標的領域を含む前記 D N A ライブラリ断片の D N A シーケンシングを備える、請求項 1 から 4 3 のいずれか 1 項に記載の前記方法。

【請求項 4 5】

20

前記ゲノム解析は、D N A 関心領域におけるコピー数の変化を測定することを備え、ステップ (d) は、

(i) 前記試験試料に由来する前記ゲノム D N A ライブラリ中に存在する前記関心領域のコピー数を測定すること、及び

(i i) ステップ (i) において測定された前記コピー数を標準試料に由来するゲノム D N A ライブラリ中に存在する前記関心領域のコピー数と比較すること、

を備え、

前記標準試料は、前記 D N A 標的領域の既知のコピー数を有する、

請求項 1 から 4 4 のいずれか 1 項に記載の前記方法。

【請求項 4 6】

30

前記関心領域における前記コピー数を測定することは、複数のシーケンシングリードを生成するために前記 D N A 標的領域を含む前記 D N A ライブラリ断片の D N A シーケンシングを備え、各シーケンシングリードは、ユニーク分子識別素子 (U M I E) を含む、請求項 4 5 に記載の前記方法。

【請求項 4 7】

前記 U M I E は、前記アダプタからのシーケンシング情報、及び前記ゲノム D N A 断片配列の少なくとも一部を含む、請求項 4 6 に記載の前記方法。

【請求項 4 8】

同一 U M I E を含むシーケンシングリードは、ユニークゲノム配列 (U G S) として識別される、請求項 4 7 に記載の前記方法。

40

【請求項 4 9】

前記ゲノム D N A ライブラリと接触する前記キャプチャプローブのそれぞれについて生のゲノム深度 (R G D) を測定することをさらに備える、請求項 4 5 から 4 8 のいずれかに記載の前記方法。

【請求項 5 0】

前記 R G D は、試料複製物群内の各キャプチャプローブ配列と関連する U G S の前記平均数を測定することを備える、請求項 4 9 に記載の前記方法。

【請求項 5 1】

非常にばらつきのある数の U G S と関連するキャプチャプローブは、ノイズの多いプローブとして識別され、さらなる計算から除外される、請求項 5 0 に記載の前記方法。

50

【請求項 5 2】

試料について R G D を計算することをさらに備え、前記試料中のすべてのキャプチャプローブについてすべての R G D の数値平均を計算することを備える、請求項 5 0 に記載の前記方法。

【請求項 5 3】

ノイズの多いプローブについての前記 R G D 値は、試料について R G D を計算する際に含まれない、請求項 5 0 に記載の前記方法。

【請求項 5 4】

前記キャプチャプローブについての前記 R G D は、各キャプチャプローブについての前記 R G D をプローブ特異的な正規化されたリードカウントに変換することによって実験群中のすべての試料にわたり正規化され、前記プローブ特異的な正規化されたリードカウントは、

(i) 試料中の各キャプチャプローブ R G D に正規化定数を乗算することであって、そこで前記正規化定数はいずれかの実数を含む、前記乗算すること、及び

(i i) 前記対応する試料について計算される前記 R G D で (i) の前記積を除算すること、または

(i i i) 1 サブセットのプローブから計算される平均 R G D で (i) の前記積を除算すること、

を備える、請求項 4 9 から 5 3 のいずれかに記載の前記方法。

【請求項 5 5】

前記 1 サブセットのプローブは、1 セットの対照プローブである、請求項 5 4 に記載の前記方法。

【請求項 5 6】

前記プローブ特異的な、正規化されたリードカウントは、コピー数値に変換され、前記コピー数値は、

(i) . 女性に由来する試料において、常染色体及び/または X 連鎖領域を対象とするプローブの前記プローブ特異的な、正規化されたリードカウントに 2 を乗算すること、

(i i) . 男性に由来する試料において、Y 連鎖領域及び/または X 連鎖領域を対象とするプローブの前記プローブ特異的な、正規化されたリードカウントに 1 を乗算すること、

(i i i) . 実験中に、すべての試料にわたり (i) 及び/または (i i) の前記積を平均すること、ならびに

(i v) . (i i i) の前記平均値で (i) 及び/または (i i) の前記積を除算すること、

を備える、請求項 5 5 に記載の前記方法。

【請求項 5 7】

特異的な遺伝子を標的にするすべてのプローブについて前記近似したコピー数値は、平均される、請求項 5 6 に記載の前記方法。

【発明の詳細な説明】

【技術分野】

【0001】

関連出願の参照

本出願は、2016 年 8 月 25 日に出願された米国仮特許出願第 62 / 379,593 号、及び 2017 年 4 月 4 日に出願された米国仮特許出願第 62 / 481,538 号に優先権を主張し、これらのそれぞれは、参照によりそれらの全体が本明細書に援用される。

【0002】

配列表に関する記述

本出願に関連する配列表は、紙媒体の代わりにテキスト形式で提供され、参照により本明細書に援用される。配列表を含むテキストファイルの名称は、CLFK__005__02WO__ST25 である。テキストファイルは、2,238 KB であり、2017 年 8 月 24 日に作成され、EFS - Web を介して電子提出された。

10

20

30

40

50

【 0 0 0 3 】

本発明は、概して、生体試料、たとえば、直接組織診または末梢血液などの定量的遺伝子解析のための組成物及び方法に関する。特に、本発明は、生体試料の、遺伝子特性及び解析と同様に、標的特異的コピー数変化の検出方法に関する。

【背景技術】

【 0 0 0 4 】

全部でないにしてもほとんどの場合、最も一般的なヒトの癌がヒトゲノムの疾患であることは、明らかになりつつある。個人の一生の間に体細胞変異が蓄積し、体細胞変異がひそむ細胞が腫瘍中に発生する可能性を体細胞変異のいくつかが増加させると考えられる。蓄積された変異事象の誤った組み合わせだけで、前癌性増殖は、制御されない増殖を阻止する制約を失った結果、細胞塊が癌になる。癌の原因となるのに必要かつ十分である一連の変異は、「ドライバー変異」と合わせて称されることが多い。最新の、かつ集中的な分子解析から現れた主題の1つは、癌が以前は単一の組織特異的疾患と考えられていたが、実際には特有の分子病理を各有する関連した疾患群であることである。ヒトゲノム計画は、癌のゲノムワイド解析についての基礎を築いた。

10

【 0 0 0 5 】

遺伝子コピー数における変化は、生物学的多様性の基本的なドライバーである。進化の背景に、遺伝子の重複、及び機能の多様性は、種の多様性の良く認識されたドライバーである。ヒト疾患の背景に、体細胞内の遺伝子減少及び遺伝子増幅は、癌などの患部組織の特徴である。ある一定の治療薬は、これらのゲノムの増加及び/または減少変異を有する細胞上に特異的に作用するが、これらのような変異が患部細胞または癌細胞のDNA内にのみ存在することが多く、体の他の細胞中で見出されないために、これらのコピー数多型の同定は困難である。患部組織または細胞が変異したDNAの主な供給源であり、バイオプシーを介してDNAを取得することは、侵襲的で、危険であり、そして不可能であることが多い。瀕死の腫瘍または癌細胞が無細胞DNAまたは循環DNAと称される、それらのDNAの小片を血流中に放出する観察は、血液試料などの、低侵襲性技術によって実施されることが可能である遺伝子検査の開発を可能にした。しかしながら、試料から無細胞DNAを単離することから、DNAのわずかな量のみを取得することができ、全DNAの一部のみは、疾患と関連する変異を保因するであろう。たとえば、癌ゲノムの背景に、診断上著しい腫瘍変異は、有意に50%未満であるマイナーアレル頻度でのみ見出されることが多い。これは、アレル頻度が一般的に約100%、50%または0%である、従来のSNP遺伝子型判定と対照的である。

20

30

【 0 0 0 6 】

したがって、特異的な標的遺伝子座における遺伝子コピー数変化を検出することが可能であるゲノム技術が必要である。

【発明の概要】

【課題を解決するための手段】

【 0 0 0 7 】

cfDNAにおける希少変異を検出する方法は、国際PCT公開第WO2016/028316号に先行して記載されている。しかしながら、これらの技術は、非常にマイナーなアレル頻度で最も希少なコピー数減少を検出するために必要な感度を依然として欠く。本明細書に提供されるのは、標的特異的コピー数変化の検出用の組成物及び方法であり、これらの組成物及び方法は、直接組織診、末梢血液、及び特にcfDNAを含む、いくつかの試料タイプに適用可能である。本明細書に記載されるこれらの組成物及び方法は、全DNAのわずかな一部のみが存在するコピー数における変化を検出するのに十分に高感度である。

40

【 0 0 0 8 】

本発明は、とりわけ、細胞ゲノムDNA（たとえば、組織診試料からの）またはcfDNA（たとえば、血液試料からの）の試料内の、変異性変化、SNP、転座、逆位、欠失、コピー数における変化、または他の遺伝的変異の検出のために有用である組成物及び方

50

法を含む。特に、本発明のこれらの組成物及び方法は、生体試料（たとえば、血液）から全 c f DNA のごく一部中のコピー数多型を検出する際に特に有用である非常に高水準の分解能を提供する。

【 0 0 0 9 】

特定の実施形態は、DNA 標的領域上で試験試料から遺伝子解析を実施する方法に描画され、この方法は、(a) 複数の DNA ライブラリ断片を含むゲノム DNA ライブラリを生成することであって、各 DNA ライブラリ断片は試験試料からのゲノム DNA 断片、及びアダプタを含む、生成すること、(b) DNA 標的領域に特異的に結着する複数のキャプチャプローブとゲノム DNA ライブラリを接触させることにより、DNA 標的領域を含むキャプチャプローブと DNA ライブラリ断片との間に複合体を形成すること、ならびに (c) DNA 標的領域を含むゲノム DNA 断片の定量的遺伝子解析を実施することを備え、そこでアダプタは増幅領域、試料タグ領域、及びアンカー領域を含む DNA ポリヌクレオチドであり、そこで増幅領域は PCR 増幅についてプライマー認識部位として機能することが可能であるポリヌクレオチド配列を含み、そこで試料タグはユニークライブラリ DNA 断片の一致度を符号化し、試験試料の一致度を符号化するポリヌクレオチド配列を含み、そこでアンカー領域は試験試料の一致度を符号化するポリヌクレオチド配列を含み、アンカー領域はゲノム DNA 断片に付着することが可能であり、そこで遺伝子解析は疾患状態を示す遺伝的变化を検出するために実施される。

10

【 0 0 1 0 】

いくつかの実施形態において、疾患状態を示す遺伝的变化は、一塩基多様体 (S N V) 、 4 0 ヌクレオチド長未満の挿入、 4 0 ヌクレオチド長未満の DNA 領域の欠失、及び / またはコピー数における変化から選択される。特定の実施形態において、疾患状態を示す遺伝的变化は、コピー数における変化である。いくつかの実施形態において、試験試料は、組織診である。さまざまな実施形態において、組織診は、腫瘍、または腫瘍であると疑われる組織から採取される。ある一定の実施形態において、ゲノム DNA は、無細胞 DNA (c f DNA) または細胞 DNA である。特定の実施形態において、ゲノム DNA は、試験試料から単離される c f DNA であり、この試験試料は、羊水、血液、血漿、血清、精液、リンパ液、脳脊髄液、眼液、尿、唾液、便、粘液、及び汗からなる群から選択される生体試料である。

20

【 0 0 1 1 】

ある一定の実施形態において、ゲノム DNA 断片は、(i) 試験試料から細胞 DNA を単離すること、及び (i i) この細胞 DNA を断片化してゲノム DNA 断片を取得すること、を備えるステップによって取得される。特定の実施形態において、ステップ (i i) は、細胞 DNA を少なくとも 1 個の消化酵素と接触させることによって実施される。いくつかの実施形態において、ステップ (i i) は、機械的応力を細胞 DNA へ加えることによって実施される。ある一定の実施形態において、この機械的応力は、細胞 DNA を超音波処理することによって加えられる。

30

【 0 0 1 2 】

特定の実施形態において、この試料タグは、ユニークゲノム DNA 断片の同定を容易にするユニーク分子識別子 (U M I) をさらに含む。

40

【 0 0 1 3 】

いくつかの実施形態において、増幅領域は、 1 0 から 5 0 の間のヌクレオチド長にある。特定の実施形態において、増幅領域は、 2 0 から 3 0 の間のヌクレオチド長にある。ある一定の実施形態において、増幅領域は、 2 5 ヌクレオチド長にある。

【 0 0 1 4 】

いくつかの実施形態において、試料タグは、 5 から 5 0 の間のヌクレオチド長にある。特定の実施形態において、試料タグは、 5 から 1 5 の間のヌクレオチド長にある。ある一定の実施形態において、試料タグは、 8 ヌクレオチド長にある。いくつかの実施形態において、U M I 増倍管は、試料タグ領域に隣接する、または試料タグ領域内に含まれる。

【 0 0 1 5 】

50

ある一定の実施形態において、UMI増倍管は、1から5の間のヌクレオチド長にある。特定の実施形態において、UMI増倍管は、3ヌクレオチド長にあり、64個の可能なヌクレオチド配列のうちの1個を含む。

【0016】

いくつかの実施形態において、アンカー領域は、1から50の間のヌクレオチド長にある。特定の実施形態において、アンカー領域は、5から25の間のヌクレオチド長にある。ある一定の実施形態において、アンカー領域は、10ヌクレオチド長にある。

【0017】

本発明の特定の実施形態は、(a)複数のDNAライブラリ断片を含むゲノムDNAライブラリを生成することのステップは、ゲノムDNA断片を複数のアダプタへ付着させることを備える。ある一定の実施形態において、ゲノムDNA断片は、ゲノムDNA断片を複数のアダプタと付着させる前に、末端修復される。特定の実施形態において、複数のアダプタの各アダプタの増幅領域は、同一のヌクレオチド配列を含む。

10

【0018】

ある一定の実施形態において、複数のアダプタの各アダプタの試料タグ領域は、2から1,000個の間のヌクレオチド配列のうちの1個を含む。特定の実施形態において、複数のアダプタの各アダプタの試料タグ領域は、50から500個の間のヌクレオチド配列のうちの1個を含む。さまざまな実施形態において、複数のアダプタの各アダプタの試料タグ領域は、100から400個の間のヌクレオチド配列のうちの1個を含む。いくつかの実施形態において、複数のアダプタの各アダプタの試料タグ領域は、200から300個の間のヌクレオチド配列のうちの1個を含む。ある一定の実施形態において、複数のアダプタの各アダプタの試料タグ領域は、8ヌクレオチド長にある。いくつかの実施形態において、ヌクレオチド配列の各配列は、少なくとも2のハミング距離によって240個のヌクレオチド配列のいずれかの他の配列から分離する。

20

【0019】

特定の実施形態において、複数のアダプタのそれぞれは、試料タグ領域に隣接する、または試料タグ領域内に含まれるUMI増倍管を含む。いくつかの実施形態において、複数のアダプタのそれぞれは、試料タグ領域に隣接するUMI増倍管を含む。ある一定の実施形態において、複数のアダプタの各アダプタのUMI増倍管は、1から5の間のヌクレオチド長にある。いくつかの実施形態において、複数のアダプタの各アダプタのUMI増倍管は、3ヌクレオチド長にある。

30

【0020】

特定の実施形態において、複数のアダプタの各アダプタのアンカータグ領域は、4個のヌクレオチド配列のうちの1個を含み、所与の配列の各試料領域は、所与の配列の4個のアンカー領域のうちの1個のみに対合される。

【0021】

いくつかの実施形態において、複数のアダプタの各アダプタの増幅領域は、同一のヌクレオチド配列を含み、複数のアダプタの各アダプタの試料タグ領域は、8ヌクレオチド長にあり、各試料タグのヌクレオチド配列は、少なくとも2のハミング距離によって複数のアダプタの試料タグのいずれかの他のヌクレオチド配列から分離し、複数のアダプタのそれぞれは、試料タグ領域に隣接する、または試料タグ領域内に含まれるUMI増倍管を備え、複数のアダプタの各アダプタのUMI増倍管は、3ヌクレオチド長にあり、それぞれの可能なヌクレオチド配列のUMI増倍管は、複数のアダプタの各試料タグ領域に対合され、複数のアダプタの各アダプタのアンカータグ領域は、4個のヌクレオチド配列のうちの1個を含み、所与の配列の各試料領域は、所与の配列の4個のアンカー領域のうちの1個のみに対合される。

40

【0022】

本発明の特定の実施形態は、方法について描画し、この方法においてゲノムDNA断片を複数のアダプタと付着させるステップは、(i)アンカー領域の少なくとも一部を含むオリゴヌクレオチドを各ゲノムDNA断片へ付着させることであって、アンカー領域の少

50

なくとも一部を含むオリゴヌクレオチドはパートナー鎖によって二本鎖にされる 5'リン酸化付着鎖を含む DNA 二本鎖であり、パートナー鎖は化学修飾による付着からその 3'末端で遮断され、付着鎖はゲノム DNA 断片に付着する、付着させること、(i i) アンカー領域の少なくとも一部を含むオリゴヌクレオチドに付着するゲノム DNA 断片を、複数のアダプタの各アダプタのヌクレオチド配列についてアダプタ配列の全長を符号化する DNA オリゴヌクレオチドと接触させること、ならびに (i i i) DNA ライゲーションに適切な条件下で、ゲノム DNA 断片、及びアダプタ配列の全長を符号化する DNA オリゴヌクレオチドを、T4 ポリヌクレオチドキナーゼ、Taq DNA リガーゼ及び Bst ポリメラーゼの全長と接触させることを備え、それらにより、複数のアダプタをゲノム DNA 断片に付着させる。いくつかの実施形態において、ゲノム DNA 断片は、cfDNA である。ある一定の実施形態において、DNA 標的領域をコピー数における変化について解析する。

10

【0023】

特定の実施形態において、DNA 標的領域を含むゲノム DNA 断片の定量的遺伝子解析を実施するステップ (c) は、DNA 標的領域を含むキャプチャプローブと DNA ライブラリ断片との間に形成される複合体の精製を有する。ある一定の実施形態において、ステップ (c) は、DNA 標的領域を含むキャプチャプローブと DNA ライブラリ断片との間に形成される複合体の精製を有し、ゲノム DNA ライブラリから関心領域を含む DNA ライブラリ断片のプライマー伸長及び/または増幅を実施する。いくつかの実施形態において、ステップ (c) は、DNA 標的領域を含むキャプチャプローブと DNA ライブラリ断片との間に形成される複合体の精製を有し、ゲノム DNA ライブラリから関心領域を含む DNA ライブラリ断片のプライマー伸長及び増幅を実施する。ある一定の実施形態において、ステップ (c) は、複数のシーケンシングリードを生成するために、DNA 標的領域を含む DNA ライブラリ断片の DNA シーケンシングを備える。

20

【0024】

いくつかの実施形態において、本発明は、方法を描画し、そこでゲノム解析は、DNA 関心領域におけるコピー数の変化を測定することを備え、またそこでステップ (c) は、DNA 標的領域を含むゲノム DNA 断片の定量的遺伝子解析を実施し、試験試料に由来するゲノム DNA ライブラリに存在する関心領域のコピー数を測定すること、そしてそれを標準試料に由来するゲノム DNA ライブラリに存在する関心領域のコピー数と比較することを備え、そこで標準試料は、DNA 標的領域の既知のコピー数を有する。

30

【0025】

いくつかの実施形態において、関心領域においてコピー数を測定することは、複数のシーケンシングリードを生成するために DNA 標的領域を含む DNA ライブラリ断片の DNA シーケンシングを備え、そこで各シーケンシングリードは、ユニーク分子識別素子 (UMIE) を含む。いくつかの実施形態において、UMIE は、アダプタからのシーケンシング情報、及びゲノム DNA 配列の少なくとも一部を含む。いくつかの実施形態において、同一 UMIE を含むシーケンシングリードは、ユニークゲノム配列 (UGS) として識別される。

【0026】

40

いくつかの実施形態において、コピー数を測定する方法は、ゲノム DNA ライブラリと接触するキャプチャプローブのそれぞれについて生のゲノム深度 (RGD) を測定することをさらに備える。いくつかの実施形態において、RGD を測定することは、試料複製物群内の各キャプチャプローブ配列と関連する UGS の平均数を測定することを備える。いくつかの実施形態において、非常にばらつきのある数の UGS と関連するキャプチャプローブは、ノイズの多いプローブとして識別され、さらなる計算から除外される。いくつかの実施形態において、RGD を測定することは、試料について RGD を計算することをさらに備え、この試料中のすべてのキャプチャプローブについてすべての RGD の数値平均を計算することを備える。いくつかの実施形態において、ノイズの多いプローブについての RGD 値は、試料について RGD を計算する際に含まれない。

50

【 0 0 2 7 】

いくつかの実施形態において、キャプチャプローブ用の R G D は、プローブ特異的な正規化されたリードカウントに各キャプチャプローブ用の R G D を変換することによって実験群中のすべての試料にわたり正規化され、これは、(i) 試料中の各キャプチャプローブ R G D に正規化定数を乗算することであって、正規化定数はいずれかの実数を含む、乗算すること、及び (i i) 対応する試料について計算される R G D で (i) の積を除算すること、または (i i i) 1 サブセットのプローブから計算される平均 R G D で (i) の積を除算することを備える。いくつかの実施形態において、1 サブセットのプローブは、1 セットの対照プローブである。

【 0 0 2 8 】

いくつかの実施形態において、プローブ特異的な正規化されたリードカウントは、コピー数値に変換され、これは、(i) 女性に由来する試料において、常染色体及び / または X 連鎖領域を対象とするプローブのプローブ特異的な正規化されたリードカウントに 2 を乗算すること、(i i) 男性に由来する試料において、Y 連鎖領域及び / または X 連鎖領域を対象とするプローブのプローブ特異的な正規化されたリードカウントに 1 を乗算すること、(i i i) 実験において、すべての試料にわたり (i) 及び / または (i i) の積を平均すること、及び (i v) (i i i) の平均値で (i) 及び / または (i i) の積を除算することを備える。いくつかの実施形態において、特異的な遺伝子を標的にするすべてのプローブについて近似したコピー数値を平均する。

【 0 0 2 9 】

いくつかの実施形態において、本発明は、コピー数増加及びコピー数減少の高感度検出方法について描画され、この方法は、(i) キャプチャプローブについて R G D を測定すること、(i i) キャプチャプローブについての R G D をプローブ特異的な正規化されたリードカウントに変換することによって実験群中のすべての試料にわたりキャプチャプローブについての R G D を正規化すること、(i i i) 各プローブ特異的な正規化されたリードカウントについて近似したコピー数値を計算すること、及び (i v) 特異的な遺伝子を標的にするすべてのプローブについて近似したコピー数値を平均することを備える。

【 0 0 3 0 】

いくつかの実施形態において、本発明は、染色体安定性を測定する方法について描画され、この方法は、(i) 1 セットの 1 個以上の染色体安定性プローブを設計し、妥当性確認することであって、これらの染色体安定性プローブはヒト染色体にわたり様に分布する、妥当性確認すること、(i i) 1 つ以上の染色体安定性プローブを使用して患者試料上で標的シーケンシングを実施すること、(i i i) 各染色体プローブについて近似したコピー数値を測定すること、(i v) 患者試料のゲノム表現型を決定することであって、患者試料中の 1 つ以上の染色体プローブについてコピー数値における増減はゲノム不安定性を示す、決定することを備える。

【 0 0 3 1 】

いくつかの実施形態において、本発明は、それを必要とする被験体中の癌を処置する方法について描画され、この被験体は請求項 6 2 の方法に従い、不安定化されたゲノムを含むと識別され、癌を処置する方法は薬学的に有効量の P A R P 阻害剤を投与することを備える。

【 0 0 3 2 】

いくつかの実施形態において、本発明は、方法を描画し、そこでゲノム解析は、DNA 関心領域におけるコピー数の変化を測定することを備え、またそこでステップ (c) は、DNA 標的領域を含むゲノム DNA 断片の定量的遺伝子解析を実施し、試験試料に由来するゲノム DNA ライブラリに存在する関心領域のコピー数を測定すること、そしてそれを標準試料に由来するゲノム DNA ライブラリに存在する関心領域のコピー数と比較することを備え、そこで標準試料は、DNA 標的領域の既知のコピー数を有する。いくつかの実施形態において、関心領域は、遺伝子、または遺伝子の一部である。特定の実施形態では、遺伝子は、疾患と関連する。ある一定の実施形態では、この疾患は、がんである。さま

10

20

30

40

50

ざまな実施形態において、この遺伝子は、BRCA2、ATM、BRCA1、BRIP1、CHEK2、FANCA、HDAC2、及び/またはPALB2である。

【0033】

特定の実施形態は、複数のDNAライブラリ断片を含むゲノムDNAライブラリについて描画され、そこで各DNAライブラリ断片はアダプタ及びゲノムDNA断片を含み、そこでアダプタは増幅領域、試料タグ領域、及びアンカー領域を含むDNAポリヌクレオチドであり、そこで増幅領域はPCR増幅についてプライマー認識部位として機能することが可能であるポリヌクレオチド配列を含み、そこで試料タグはユニークライブラリDNA断片の一致度を符号化し、試験試料の一致度を符号化するポリヌクレオチド配列を含み、そこでアンカー領域は試験試料の一致度を符号化するポリヌクレオチド配列を含み、アンカー領域はゲノムDNA断片に付着することが可能である。いくつかの実施形態において、この試料タグは、ユニーク分子識別子(UMI)をさらに含み、そこでUMIはユニークゲノムDNA断片の同定を容易にする。特定の実施形態において、増幅領域は、10から50の間のヌクレオチド長にある。特定の実施形態において、増幅領域は、25ヌクレオチド長にある。特定の実施形態において、試料タグは、5から50の間のヌクレオチド長にある。ある一定の実施形態において、試料タグは、8ヌクレオチド長にある。いくつかの実施形態において、UMI増倍管は、試料タグ領域に隣接する、または試料タグ領域内に含まれる。特定の実施形態において、UMI増倍管は、1から5の間のヌクレオチド長にある。ある一定の実施形態において、アンカー領域は、1から50の間のヌクレオチド長にある。いくつかの実施形態において、アンカー領域は、10ヌクレオチド長にある。特定の実施形態において、複数のアダプタの各アダプタの増幅領域は、同一のヌクレオチド配列を含む。いくつかの実施形態において、試料タグの各ヌクレオチド配列は、少なくとも2のハミング距離によって試料のヌクレオチド配列のいずれかの他の配列から分離する。ある一定の実施形態において、複数のアダプタのそれぞれは、試料タグ領域に隣接する、または試料タグ領域内に含まれるUMI増倍管を含む。特定の実施形態において、複数のアダプタのそれぞれは、試料タグ領域に隣接するUMI増倍管を含む。いくつかの実施形態において、複数のアダプタの各アダプタのアンカータグ領域は、4個のヌクレオチド配列のうちの1個を含み、所与の配列の各試料領域は、所与の配列の4個のアンカー領域のうちの1個のみに対合される。いくつかの実施形態において、ゲノムDNA断片は、cfDNAである。

【0034】

ある一定の実施形態において、複数のアダプタの各アダプタの増幅領域は、同一のヌクレオチド配列を含み、複数のアダプタの各アダプタの試料タグ領域は、8ヌクレオチド長にあり、複数のアダプタの各アダプタの試料タグ領域は、少なくとも2のハミング距離によって複数のアダプタの試料タグのいずれかの他のヌクレオチド配列から分離するヌクレオチド配列を含み、複数のアダプタのそれぞれは、試料タグ領域に隣接する、または試料タグ領域内に含まれるUMI増倍管を備え、複数のアダプタの各アダプタのUMI増倍管は、3ヌクレオチド長にあり、それぞれの可能なヌクレオチド配列のUMI増倍管は、複数のアダプタの各試料タグ領域に対合され、複数のアダプタの各アダプタのアンカータグ領域は、4個のヌクレオチド配列のうちの1個を含み、所与の配列の各試料領域は、所与の配列の4個のアンカー領域のうちの1個のみに対合される。いくつかの実施形態において、ゲノムDNA断片は、cfDNAである。

【0035】

ある一定の実施形態は、本明細書に記載される1つより多いゲノムライブラリを含む、複数のゲノムDNAライブラリについて描画される。いくつかの実施形態において、複数のゲノムDNAライブラリに属する1つのゲノムDNAライブラリの試料タグ領域の核酸配列は、複数のゲノムDNAライブラリに属する他のゲノムDNAライブラリの試料タグ領域の核酸配列と異なる。特定の実施形態において、複数のゲノムDNAライブラリに属する1つのゲノムDNAライブラリの増幅領域の核酸配列は、複数のゲノムDNAライブラリに属する他のゲノムDNAライブラリの増幅領域の核酸配列に同一である。

【 0 0 3 6 】

ある一定の実施形態は、無細胞DNA (cfDNA) のDNA 標的領域の遺伝子解析方法について描画され、この遺伝子解析方法は、(a) 本明細書に記載されるようなDNA ライブラリを生成すること、(b) DNA 標的領域に特異的に結着する複数のキャプチャプローブとcfDNA ライブラリを接触させることにより、DNA 標的領域を含むキャプチャプローブとDNA ライブラリ断片との間に複合体を形成すること、及び(c) DNA 標的領域を含むcfDNA 断片の定量的遺伝子解析を実施すること、を備え、それらによりDNA 標的領域の遺伝子解析を実施する。

【 0 0 3 7 】

ある一定の実施形態は、被験体における遺伝性疾患を予測する、診断する、または監視する方法を対象とし、この方法は、(a) 試験試料を被験体から取得すること、(b) ゲノムDNA を試験試料から単離すること、(c) 複数のDNA ライブラリ断片を含むDNA ライブラリを生成することであって、各DNA ライブラリ断片は試験試料からのゲノムDNA 断片、及びアダプタを含む、生成すること、(d) DNA 標的領域に特異的に結着する複数のキャプチャプローブとcfDNA ライブラリを接触させることにより、DNA 標的領域を含むキャプチャプローブとDNA ライブラリ断片との間に複合体を形成すること、ならびに(e) cfDNA クローンライブラリにおける遺伝性疾患に関連する1つ以上の標的遺伝子座の定量的遺伝子解析を実施することであって、1つ以上の標的遺伝子座における1つ以上の遺伝子損傷の同定または検出は遺伝性疾患の進行の予後判定、診断、または監視である、実施することを備える。特定の実施形態では、定量的遺伝子解析は、複数のシーケンシングリードを生成するためのDNA シーケンシングを含む。

【 0 0 3 8 】

特定の実施形態は、ユニークゲノムDNA 断片の一致度、及び試験試料の一致度を符号化する1セットのアダプタについて描画され、ゲノムDNA ライブラリを生成する際の使用のために、そこで1セットのアダプタ中の各アダプタは、増幅領域、試料タグ領域、及びアンカー領域を含むDNA ポリヌクレオチドであり、そこで増幅領域は、PCR 増幅についてプライマー認識部位として機能することが可能であるポリヌクレオチド配列を含み、そこで試料タグは、ユニークライブラリDNA 断片の一致度を符号化し、試験試料の一致度を符号化するポリヌクレオチド配列を含み、そこでアンカー領域は、試験試料の一致度を符号化するポリヌクレオチド配列を含み、アンカー領域は、ゲノムDNA 断片に付着することが可能である。いくつかの実施形態において、この試料タグは、ユニーク分子識別子(UMI) をさらに含み、そこでUMI はユニークゲノムDNA 断片の同定を容易にする。さまざまな実施形態において、増幅領域は、10 から50 の間のヌクレオチド長にある。ある一定の実施形態において、増幅領域は、25 ヌクレオチド長にある。特定の実施形態において、試料タグは、5 から50 の間のヌクレオチド長にある。いくつかの実施形態において、試料タグは、8 ヌクレオチド長にある。特定の実施形態において、UMI 増倍管は、試料タグ領域に隣接する、または試料タグ領域内に含まれる。いくつかの実施形態において、UMI 増倍管は、1 から5 の間のヌクレオチド長にある。特定の実施形態において、アンカー領域は、1 から50 の間のヌクレオチド長にある。いくつかの実施形態において、アンカー領域は、10 ヌクレオチド長にある。ある一定の実施形態において、複数のアダプタの各アダプタの増幅領域は、同一のヌクレオチド配列を含む。

【 0 0 3 9 】

いくつかの実施形態において、試料タグの各ヌクレオチド配列は、少なくとも2 のハミング距離によって1セットのアダプタの試料タグのいずれかの他のヌクレオチド配列から分離する。さまざまな実施形態において、複数のアダプタのそれぞれは、試料タグ領域に隣接する、または試料タグ領域内に含まれるUMI 増倍管を含む。特定の実施形態において、複数のアダプタのそれぞれは、試料タグ領域に隣接するUMI 増倍管を含む。

【 0 0 4 0 】

いくつかの実施形態において、複数のアダプタの各アダプタのアンカータグ領域は、4 個のヌクレオチド配列のうちの1 個を含み、所与の配列の各試料領域は、所与の配列の4

10

20

30

40

50

個のアンカー領域のうちの1個のみに対合される。請求項75の1セットのアダプタにおいて、複数のアダプタの各アダプタの増幅領域は、同一のヌクレオチド配列を含み、各アダプタの試料タグ領域は、8ヌクレオチド長にあり、試料タグの各ヌクレオチド配列は、少なくとも2のハミング距離によって1セットのアダプタの試料タグのいずれかの他のヌクレオチド配列から分離し、複数のアダプタのそれぞれは、試料タグ領域に隣接する、または試料タグ領域内に含まれるUMI増倍管を備え、複数のアダプタの各アダプタのUMI増倍管は、3ヌクレオチド長にあり、UMI増倍管は、64個の可能なヌクレオチド配列のうちの1個を含み、それぞれの64個の可能なヌクレオチド配列のUMI増倍管は、複数のアダプタの試料タグ領域のそれぞれに対合され、複数のアダプタの各アダプタのアンカータグ領域は、4個のヌクレオチド配列のうちの1個を含み、所与の配列の各試料領域は、所与の配列の4個のアンカー領域のうちの1個のみに対合される。

10

特定の実施形態では、例えば以下の項目が提供される。

(項目1)

DNA標的領域上で試験試料から遺伝子解析を実施する方法であって、

(a) 複数のDNAライブラリ断片を含むゲノムDNAライブラリを生成することであって、そこで各前記DNAライブラリ断片は前記試験試料からのゲノムDNA断片、及びアダプタを含む、前記生成すること、

(b) DNA標的領域に特異的に結着する複数のキャプチャプローブと前記ゲノムDNAライブラリを接触させることにより、前記DNA標的領域を含む前記キャプチャプローブとDNAライブラリ断片との間に複合体を形成すること、ならびに

20

(c) 前記DNA標的領域を含む前記ゲノムDNA断片の定量的遺伝子解析を実施すること、

を備え、

そこで前記アダプタは増幅領域、試料タグ領域、及びアンカー領域を含むDNAポリヌクレオチドであり、

そこで前記増幅領域はPCR増幅についてプライマー認識部位として機能することが可能であるポリヌクレオチド配列を含み、

そこで前記試料タグは前記ユニークライブラリDNA断片の一致度を符号化し、前記試験試料の一致度を符号化するポリヌクレオチド配列を含み、

そこで前記アンカー領域は前記試験試料の前記一致度を符号化するポリヌクレオチド配列を含み、そこで前記アンカー領域は前記ゲノムDNA断片に付着することが可能であり、そこで前記遺伝子解析は疾患状態を示す遺伝的变化を検出するために実施される、

30

前記方法。

(項目2)

疾患状態を示す前記遺伝的变化は、一塩基多様体(SNV)、40ヌクレオチド長未満の挿入、40ヌクレオチド長未満のDNA領域の欠失、及び/またはコピー数における変化から選択される、項目1に記載の前記方法。

(項目3)

疾患状態を示す前記遺伝的变化は、コピー数における変化である、項目1に記載の前記方法。

40

(項目4)

前記試験試料は、組織診である、項目1から3のいずれかに記載の前記方法。

(項目5)

前記組織診は、腫瘍、または腫瘍であると疑われる組織から採取される、項目4に記載の前記方法。

(項目6)

前記ゲノムDNAは、無細胞DNA(cfDNA)または細胞DNAである、項目1から3のいずれかに記載の前記方法。

(項目7)

前記ゲノムDNAは、前記試験試料から単離されるcfDNAであり、前記試験試料は、

50

羊水、血液、血漿、血清、精液、リンパ液、脳脊髄液、眼液、尿、唾液、便、粘液、及び汗からなる群から選択される生体試料である、項目 6 に記載の前記方法。

(項目 8)

前記ゲノム DNA 断片は、

(i) 前記試験試料から細胞 DNA を単離すること、

(i i) 前記細胞 DNA を断片化し、前記ゲノム DNA 断片を取得すること、

を備えるステップによって取得される、項目 1 から 5 のいずれかに記載の前記方法。

(項目 9)

ステップ (i i) は、前記細胞 DNA を少なくとも 1 個の消化酵素と接触させることによって実施される、項目 8 に記載の前記方法。

(項目 10)

ステップ (i i) は、機械的応力を前記細胞 DNA へ加えることによって実施される、項目 8 に記載の前記方法。

(項目 11)

前記機械的応力は、前記細胞 DNA を超音波処理することによって加えられる、項目 10 に記載の前記方法。

(項目 12)

前記試料タグは、前記ユニークゲノム DNA 断片の前記同定を容易にするユニーク分子識別子 (U M I) をさらに含む、先行項目のいずれかに記載の前記方法。

(項目 13)

前記増幅領域は、10 から 50 の間のヌクレオチド長にある、先行項目のうちのいずれか 1 項に記載の前記方法。

(項目 14)

前記増幅領域は、20 から 30 の間のヌクレオチド長にある、先行項目のうちのいずれか 1 項のいずれかに記載の前記方法。

(項目 15)

前記増幅領域は、25 ヌクレオチド長にある、先行項目のうちのいずれか 1 項のいずれかに記載の前記方法。

(項目 16)

前記試料タグは、5 から 50 の間のヌクレオチド長にある、先行項目のうちのいずれか 1 項に記載の前記方法。

(項目 17)

前記試料タグは、5 から 15 の間のヌクレオチド長にある、項目 16 に記載の前記方法。

(項目 18)

前記試料タグは、8 ヌクレオチド長にある、項目 16 に記載の前記方法。

(項目 19)

前記 U M I 増倍管は、前記試料タグ領域に隣接する、または前記試料タグ領域内に含まれる、項目 12 から 18 のうちのいずれか 1 項に記載の前記方法。

(項目 20)

前記 U M I 増倍管は、1 から 5 の間のヌクレオチド長にある、項目 19 に記載の前記方法。

(項目 21)

前記 U M I 増倍管は、3 ヌクレオチド長にあり、6.4 個の可能なヌクレオチド配列のうちの 1 個を含む、項目 19 に記載の前記方法。

(項目 22)

前記アンカー領域は、1 から 50 の間のヌクレオチド長にある、先行項目のうちのいずれか 1 項に記載の前記方法。

(項目 23)

前記アンカー領域は、5 から 25 の間のヌクレオチド長にある、項目 22 に記載の前記方法。

(項目 24)

10

20

30

40

50

前記アンカー領域は、10ヌクレオチド長にある、項目22または23に記載の前記方法。
(項目25)
ステップ(a)は、前記ゲノムDNA断片を複数のアダプタへ付着させることを備える、
先行項目のいずれか1項に記載の前記方法。
(項目26)
前記ゲノムDNA断片は、前記ゲノムDNA断片を複数のアダプタと付着させる前に、末
端修復される、項目25に記載の前記方法。
(項目27)
前記複数のアダプタの各アダプタの前記増幅領域は、同一のヌクレオチド配列を含む、項
目25に記載の前記方法。
(項目28)
前記複数のアダプタの各アダプタの前記試料タグ領域は、2から1,000個の間のヌク
レオチド配列のうちの1個を含む、項目26または27に記載の前記方法。
(項目29)
前記複数のアダプタの各アダプタの前記試料タグ領域は、50から500個の間のヌクレ
オチド配列のうちの1個を含む、項目28に記載の前記方法。
(項目30)
前記複数のアダプタの各アダプタの前記試料タグ領域は、100から400個の間のヌク
レオチド配列のうちの1個を含む、項目28に記載の前記方法。
(項目31)
前記複数のアダプタの各アダプタの前記試料タグ領域は、200から300個の間のヌク
レオチド配列のうちの1個を含む、項目28に記載の前記方法。
(項目32)
前記複数のアダプタの各アダプタの前記試料タグ領域は、8ヌクレオチド長にある、項目
28に記載の前記方法。
(項目33)
前記ヌクレオチド配列の各配列は、少なくとも2のハミング距離によって前記240個の
ヌクレオチド配列のいずれかの他の配列から分離する、項目28から32のいずれかに記
載の前記方法。
(項目34)
前記複数のアダプタのそれぞれは、前記試料タグ領域に隣接する、または前記試料タグ領
域内に含まれるUMI増倍管を備える、項目26から33のうちのいずれか1項に記載の前記
方法。
(項目35)
前記複数のアダプタのそれぞれは、前記試料タグ領域に隣接するUMI増倍管を備える、
項目26から34のうちのいずれか1項に記載の前記方法。
(項目36)
前記複数のアダプタの各アダプタの前記UMI増倍管は、1から5の間のヌクレオチド長
にある、項目34または35に記載の前記方法。
(項目37)
前記複数のアダプタの各アダプタの前記UMI増倍管は、3ヌクレオチド長にある、項目
36に記載の前記方法。
(項目38)
前記複数のアダプタの各アダプタの前記アンカータグ領域は、4個のヌクレオチド配列の
うちの1個を含み、所与の配列の各試料領域は、所与の配列の前記4個のアンカー領域の
うちの1個のみに対合される、項目26から37に記載の前記方法。
(項目39)
前記複数のアダプタの各アダプタの前記増幅領域は、同一のヌクレオチド配列を含み、
前記複数のアダプタの各アダプタの前記試料タグ領域は、8ヌクレオチド長にあり、各試
料タグの前記ヌクレオチド配列は、少なくとも2のハミング距離によって前記複数のアダ

10

20

30

40

50

プタの前記試料タグのいずれかの他のヌクレオチド配列から分離し、
前記複数のアダプタのそれぞれは、前記試料タグ領域に隣接する、または前記試料タグ領域内に含まれるUMI増倍管を備え、前記複数のアダプタの各アダプタの前記UMI増倍管は、3ヌクレオチド長にあり、それぞれの前記可能なヌクレオチド配列の前記UMI増倍管は、前記複数のアダプタの各試料タグ領域に対合され、
前記複数のアダプタの各アダプタの前記アンカータグ領域は、4個のヌクレオチド配列のうちの1個を含み、所与の配列の各試料領域は、所与の配列の前記4個のアンカー領域のうちの1個のみに対合される、
項目25または26に記載の前記方法。

(項目40)

前記ゲノムDNA断片を複数のアダプタと付着させる前記ステップは、
(i) アンカー領域の少なくとも一部を含むオリゴヌクレオチドを各ゲノムDNA断片へ付着させることであって、そこでアンカー領域の少なくとも一部を含む前記オリゴヌクレオチドはパートナー鎖によって二本鎖にされる5'リン酸化付着鎖を含むDNA二本鎖であり、そこで前記パートナー鎖は化学修飾による付着からその3'末端で遮断され、そこで前記付着鎖は前記ゲノムDNA断片に付着する、前記付着させること、

(ii) 前記複数のアダプタの各アダプタのヌクレオチド配列についてアダプタ配列の全長を符号化するDNAオリゴヌクレオチドと、前記アンカー領域の少なくとも一部を含む前記オリゴヌクレオチドに付着する前記ゲノムDNA断片を接触させること、ならびに

(iii) DNAライゲーションに適切な条件下でT4ポリヌクレオチドキナーゼ、Taq DNAリガーゼ、及びBstポリメラーゼ全長と、前記ゲノムDNA断片、及び前記アダプタ配列の全長を符号化する前記DNAオリゴヌクレオチドを接触させること、
を備え、

それらにより、前記複数のアダプタを前記ゲノムDNA断片に付着させる、項目25から39のいずれか1項に記載の前記方法。

(項目41)

前記ゲノムDNA断片は、cfDNAである、項目25から40のいずれかに記載の前記方法。

(項目42)

前記DNA標的領域は、コピー数における変化について解析される、項目25から41のいずれかに記載の前記方法。

(項目43)

ステップ(c)は、前記DNA標的領域を含む前記キャプチャプローブとDNAライブラリ断片との間に形成される前記複合体の精製を備える、先行項目のいずれか1項に記載の前記方法。

(項目44)

ステップ(c)は、前記DNA標的領域を含む前記キャプチャプローブとDNAライブラリ断片との間に形成される前記複合体の精製を備え、前記ゲノムDNAライブラリから前記関心領域を含む前記DNAライブラリ断片のプライマー伸長及び/または増幅を実施する、先行項目のいずれか1項に記載の前記方法。

(項目45)

ステップ(c)は、前記DNA標的領域を含む前記キャプチャプローブとDNAライブラリ断片との間に形成される前記複合体の精製を備え、前記ゲノムDNAライブラリから前記関心領域を含む前記DNAライブラリ断片のプライマー伸長及び増幅を実施する、先行項目のいずれか1項に記載の前記方法。

(項目46)

ステップ(c)は、複数のシーケンシングリードを生成するために、前記DNA標的領域を含む前記DNAライブラリ断片のDNAシーケンシングを備える、先行項目のいずれか1項に記載の前記方法。

(項目47)

前記ゲノム解析は、DNA 関心領域におけるコピー数の変化を測定することを備え、ステップ(c)は、

(i) 前記試験試料に由来する前記ゲノムDNAライブラリ中に存在する前記関心領域のコピー数を測定すること、及び

(ii) ステップ(i)において測定された前記コピー数を標準試料に由来する前記ゲノムDNAライブラリ中に存在する前記関心領域のコピー数と比較すること、

を備え、

前記標準試料は、前記DNA標的領域の既知のコピー数を有する、

先行項目のいずれか1項に記載の前記方法。

(項目48)

前記関心領域における前記コピー数を測定することは、複数のシーケンシングリードを生成するために前記DNA標的領域を含む前記DNAライブラリ断片のDNAシーケンシングを備え、各シーケンシングリードは、ユニーク分子識別素子(UMIE)を含む、項目47に記載の前記方法。

(項目49)

前記UMIEは、前記アダプタからのシーケンシング情報、及び前記ゲノムDNA配列の少なくとも一部を含む、項目48に記載の前記方法。

(項目50)

同一UMIEを含むシーケンシングリードは、ユニークゲノム配列(UGS)として識別される、項目49に記載の前記方法。

(項目51)

前記ゲノムDNAライブラリと接触する前記キャプチャプローブのそれぞれについて生のゲノム深度(RGD)を測定することをさらに備える、項目47から50のいずれかに記載の前記方法。

(項目52)

前記RGDは、試料複製物群内の各キャプチャプローブ配列と関連するUGSの前記平均数を測定することを備える、項目51に記載の前記方法。

(項目53)

非常にばらつきのある数のUGSと関連するキャプチャプローブは、ノイズの多いプローブとして識別され、さらなる計算から除外される、項目52に記載の前記方法。

(項目54)

試料についてRGDを計算することをさらに備え、前記試料中のすべてのキャプチャプローブについてすべてのRGDの数値平均を計算することを備える、項目52に記載の前記方法。

(項目55)

ノイズの多いプローブについての前記RGD値は、試料についてRGDを計算する際に含まれない、項目52に記載の前記方法。

(項目56)

前記キャプチャプローブについての前記RGDは、各キャプチャプローブについての前記RGDをプローブ特異的な正規化されたリードカウントに変換することによって実験群中のすべての試料にわたり正規化され、

前記プローブ特異的な正規化されたリードカウントは、

(i) 試料中の各キャプチャプローブRGDに正規化定数を乗算することであって、そこで前記正規化定数はいずれかの実数を含む、前記乗算すること、及び

(ii) 前記対応する試料について計算される前記RGDで(i)の前記積を除算すること、または

(iii) 1サブセットのプローブから計算される平均RGDで(i)の前記積を除算すること、

を備える、項目51から55のいずれかに記載の前記方法。

(項目57)

10

20

30

40

50

前記1サブセットのプロープは、1セットの対照プロープである、項目56に記載の前記方法。

(項目58)

前記プロープ特異的な、正規化されたリードカウントは、コピー数値に変換され、前記コピー数値は、

(i)、女性に由来する試料において、常染色体及び/またはX連鎖領域を対象とするプロープの前記プロープ特異的な、正規化されたリードカウントに2を乗算すること、

(ii)、男性に由来する試料において、Y連鎖領域及び/またはX連鎖領域を対象とするプロープの前記プロープ特異的な、正規化されたリードカウントに1を乗算すること、

(iii)、実験中に、すべての試料にわたり(i)及び/または(ii)の前記積を平均すること、ならびに

(iv)、(iii)の前記平均値で(i)及び/または(ii)の前記積を除算すること、

を備える、項目57に記載の前記方法。

(項目59)

特異的な遺伝子を標的にするすべてのプロープについて前記近似したコピー数値は、平均される、項目58に記載の前記方法。

(項目60)

コピー数増加及びコピー数減少の高感度検出方法であって、

(i)キャプチャプロープについてRGDを測定すること、

(ii)前記キャプチャプロープについての前記RGDをプロープ特異的な、正規化されたリードカウントに変換することによって、実験群中にすべての試料にわたり前記キャプチャプロープについての前記RGDを正規化すること、

(iii)各プロープ特異的な、正規化されたリードカウントについて近似したコピー数値を計算すること、及び

(iv)特異的な遺伝子を標的にするすべてのプロープについて前記近似したコピー数値を平均すること、

を備える、前記方法。

(項目61)

染色体安定性を測定する方法であって、

(i)1セットの1つ以上の染色体安定性プロープを設計し、妥当性確認することであって、前記染色体安定性プロープはヒト染色体にわたり均一に分布する、前記妥当性確認すること、

(ii)前記1つ以上の染色体安定性プロープを使用して患者試料上で標的シーケンシングを実施すること、

(iii)各染色体プロープについて近似したコピー数値を決定すること、

(iv)患者試料のゲノム表現型を決定することであって、そこで前記患者試料中の1つ以上の染色体プロープについて前記コピー数値における増減はゲノム不安定性を示すこと、

を備える、前記方法。

(項目62)

それを必要とする被験体中の癌を処置する方法であって、

前記被験体は項目61の前記方法に従い、不安定化されたゲノムを含むと識別され、

前記癌を処置する前記方法は薬学的に有効量のPARP阻害剤を投与する、

ことを備える、前記方法。

(項目63)

前記関心領域は、遺伝子、または前記遺伝子の一部である、先行項目のいずれか1項に記載の前記方法。

(項目64)

前記遺伝子は、疾患と関連する、項目63に記載の前記方法。

(項目65)

10

20

30

40

50

前記疾患は、癌である、項目 6 4 に記載の前記方法。

(項目 6 6)

前記遺伝子は、BRCA2、ATM、BRCA1、BRIP1、CHEK2、FANCA、HDAC2、及び/またはPALB2である、項目 6 3 に記載の前記方法。

(項目 6 7)

複数のDNAライブラリ断片を含むゲノムDNAライブラリであって、

前記DNAライブラリ断片のそれぞれはアダプタ及びゲノムDNA断片を含み、

前記アダプタは増幅領域、試料タグ領域、及びアンカー領域を含むDNAポリヌクレオチドであり、

前記増幅領域はPCR増幅についてプライマー認識部位として機能することが可能であるポリヌクレオチド配列を含み、

前記試料タグは前記ユニークライブラリDNA断片の一致度を符号化し、前記試験試料の一致度を符号化するポリヌクレオチド配列を含み、

前記アンカー領域は前記試験試料の前記一致度を符号化するポリヌクレオチド配列を含み、前記アンカー領域は前記ゲノムDNA断片に付着することが可能である、

前記ゲノムDNAライブラリ。

(項目 6 8)

前記試料タグは、ユニーク分子識別子(UMI)をさらに含み、前記UMIは、前記ユニークゲノムDNA断片の前記同定を容易にする、項目 6 7 に記載の前記ゲノムDNAライブラリ。

(項目 6 9)

前記増幅領域は、10から50の間のヌクレオチド長にある、項目 6 7 または 6 8 に記載の前記ゲノムDNAライブラリ。

(項目 7 0)

前記増幅領域は、25ヌクレオチド長にある、項目 6 9 に記載の前記ゲノムDNAライブラリ。

(項目 7 1)

前記試料タグは、5から50の間のヌクレオチド長にある、先行項目のうちのいずれか1項に記載の前記ゲノムDNAライブラリ。

(項目 7 2)

前記試料タグは、8ヌクレオチド長にある、項目 7 1 に記載の前記ゲノムDNAライブラリ。

(項目 7 3)

前記UMI増倍管は、前記試料タグ領域に隣接する、または前記試料タグ領域内に含まれる、項目 6 7 から 7 2 のうちのいずれか1項に記載の前記ゲノムDNAライブラリ。

(項目 7 4)

前記UMI増倍管は、1から5の間のヌクレオチド長にある、項目 7 3 に記載の前記ゲノムDNAライブラリ。

(項目 7 5)

前記アンカー領域は、1から50の間のヌクレオチド長にある、先行項目のうちのいずれか1項に記載の前記ゲノムDNAライブラリ。

(項目 7 6)

前記アンカー領域は、10ヌクレオチド長にある、項目 7 5 に記載の前記ゲノムDNAライブラリ。

(項目 7 7)

前記複数のアダプタの各アダプタの前記増幅領域は、同一のヌクレオチド配列を含む、項目 6 7 から 7 6 のいずれか1項に記載の前記ゲノムDNAライブラリ。

(項目 7 8)

前記試料タグの各ヌクレオチド配列は、少なくとも2のハミング距離によって前記試料の前記ヌクレオチド配列のいずれかの他の配列から分離する、項目 6 7 から 7 7 のいずれか

10

20

30

40

50

1 項に記載の前記ゲノム DNA ライブラリ。

(項目 7 9)

前記複数のアダプタのそれぞれは、前記試料タグ領域に隣接する、または前記試料タグ領域内に含まれる U M I 増倍管を備える、項目 6 7 から 7 8 のいずれか 1 項に記載の前記ゲノム DNA ライブラリ。

(項目 8 0)

前記複数のアダプタのそれぞれは、前記試料タグ領域に隣接する U M I 増倍管を備える、項目 6 7 から 7 8 のいずれか 1 項に記載の前記ゲノム DNA ライブラリ。

(項目 8 1)

前記複数のアダプタの各アダプタの前記アンカータグ領域は、4 個のヌクレオチド配列のうちの 1 個を含み、所与の配列の各試料領域は、所与の配列の前記 4 個のアンカー領域のうちの 1 個のみに対合される、項目 6 7 から 7 8 のいずれか 1 項に記載の前記ゲノム DNA ライブラリ。

10

(項目 8 2)

前記複数のアダプタの各アダプタの前記増幅領域は、同一のヌクレオチド配列を含み、前記複数のアダプタの各アダプタの前記試料タグ領域は、8 ヌクレオチド長にあり、前記複数のアダプタの各アダプタの前記試料タグ領域は、少なくとも 2 のハミング距離によって前記複数のアダプタの前記試料タグのいずれかの他のヌクレオチド配列から分離し、前記複数のアダプタのそれぞれは、前記試料タグ領域に隣接する、または前記試料タグ領域内に含まれる U M I 増倍管を備え、前記複数のアダプタの各アダプタの前記 U M I 増倍管は、3 ヌクレオチド長にあり、それぞれの前記可能なヌクレオチド配列の前記 U M I 増倍管は、前記複数のアダプタの前記試料タグ領域のそれぞれに対合され、前記複数のアダプタの各アダプタの前記アンカータグ領域は、4 個のヌクレオチド配列のうちの 1 個を含み、

20

所与の配列の各試料領域は、所与の配列の前記 4 個のアンカー領域のうちの 1 個のみに対合される、

項目 6 7 に記載の前記ゲノム DNA ライブラリ。

(項目 8 3)

前記ゲノム DNA 断片は、c f DNA である、項目 6 7 から 8 2 のいずれかに記載の前記ゲノム DNA ライブラリ。

30

(項目 8 4)

項目 6 7 から 8 3 のいずれか 1 項に従い、1 個より多いゲノムライブラリを含む、複数のゲノム DNA ライブラリ。

(項目 8 5)

前記複数のゲノム DNA ライブラリに属する 1 個のゲノム DNA ライブラリの前記試料タグ領域の前記核酸配列は、前記複数のゲノム DNA ライブラリに属する他のゲノム DNA ライブラリの前記試料タグ領域の前記核酸配列と異なる、項目 8 4 に記載の前記複数のゲノム DNA ライブラリ。

(項目 8 6)

前記複数のゲノム DNA ライブラリに属する 1 個のゲノム DNA ライブラリの前記増幅領域の前記核酸配列は、前記複数のゲノム DNA ライブラリに属する他のゲノム DNA ライブラリの前記増幅領域の前記核酸配列に同一である、項目 8 4 または 8 5 に記載の前記複数のゲノム DNA ライブラリ。

40

(項目 8 7)

無細胞 DNA (c f DNA) の DNA 標的領域の遺伝子解析方法であって、

(a) 項目 6 7 から 8 6 のいずれかに記載の前記 DNA ライブラリを生成すること、

(b) DNA 標的領域に特異的に結着する複数のキャプチャプローブと前記 c f DNA ライブラリを接触させることにより、前記 DNA 標的領域を含む前記キャプチャプローブと DNA ライブラリ断片との間に複合体を形成すること、及び

(c) 前記 DNA 標的領域を含む前記 c f DNA 断片の定量的遺伝子解析を実施すること、

50

を備え

それらにより、前記DNA標的領域の遺伝子解析を実施する、
前記遺伝子解析方法。

(項目88)

被験体における遺伝性疾患を予測する、診断する、または監視する方法であって、

(a) 前記被験体から試験試料を得ることと、

(b) 前記試験試料からゲノムDNAを単離することと、

(c) 複数のDNAライブラリ断片を含むDNAライブラリを生成することであって、前記DNAライブラリ断片のそれぞれは前記試験試料からのゲノムDNA断片、及びアダプタを含む、前記生成することと、

(d) DNA標的領域に特異的に結着する複数のキャプチャプローブと前記cfDNAライブラリを接触させることにより、前記DNA標的領域を含む前記キャプチャプローブとDNAライブラリ断片との間に複合体を形成することと、

(e) 前記cfDNAクローンライブラリにおける前記遺伝性疾患に関連する1つ以上の標的遺伝子座の定量的遺伝子解析を実施することであって、前記1つ以上の標的遺伝子座における1つ以上の遺伝子損傷の前記同定または検出は前記遺伝性疾患の前記進行の予後判定、診断、または監視である、前記実施することと、

を含む前記方法。

(項目89)

前記定量的遺伝子解析は、複数のシーケンシングリードを生成するためのDNAシーケンシングを含む、項目87または88に記載の前記方法。

(項目90)

ユニークゲノムDNA断片の一致度、及び試験試料の一致度を符号化する1セットのアダプタであって、ゲノムDNAライブラリを生成する際の使用のために、

前記1セットのアダプタ中の各アダプタは増幅領域、試料タグ領域、及びアンカー領域を含むDNAポリヌクレオチドであり、

前記増幅領域はPCR増幅についてプライマー認識部位として機能することが可能であるポリヌクレオチド配列を含み、

前記試料タグは前記ユニークライブラリDNA断片の前記一致度を符号化し、前記試験試料の前記一致度を符号化するポリヌクレオチド配列を含み、

前記アンカー領域は前記試験試料の前記一致度を符号化するポリヌクレオチド配列を含み、前記アンカー領域は前記ゲノムDNA断片に付着することが可能である、

前記1セットのアダプタ。

(項目91)

前記試料タグは、ユニーク分子識別子(UMI)をさらに含み、前記UMIは、前記ユニークゲノムDNA断片の前記同定を容易にする、項目90に記載の前記1セットのアダプタ。

(項目92)

前記増幅領域は、10から50の間のヌクレオチド長にある、項目90または91に記載の前記1セットのアダプタ。

(項目93)

前記増幅領域は、25ヌクレオチド長にある、項目90から92のいずれか1項に記載の前記1セットのアダプタ。

(項目94)

前記試料タグは、5から50の間のヌクレオチド長にある、先行項目のいずれか1項に記載の前記1セットのアダプタ。

(項目95)

前記試料タグは、8ヌクレオチド長にある、項目94に記載の前記1セットのアダプタ。

(項目96)

前記UMI増倍管は、前記試料タグ領域に隣接する、または前記試料タグ領域内に含まれ

10

20

30

40

50

る、項目 9 0 から 9 5 のいずれか 1 項に記載の前記 1 セットのアダプタ。

(項目 9 7)

前記 U M I 増倍管は、1 から 5 の間のヌクレオチド長にある、項目 9 6 に記載の前記 1 セットのアダプタ。

(項目 9 8)

前記アンカー領域は、1 から 5 0 の間のヌクレオチド長にある、項目 9 0 から 9 7 に記載の前記 1 セットのアダプタ。

(項目 9 9)

前記アンカー領域は、1 0 ヌクレオチド長にある、項目 9 8 に記載の前記 1 セットのアダプタ。

(項目 1 0 0)

前記複数のアダプタの各アダプタの前記増幅領域は、同一のヌクレオチド配列を含む、項目 9 0 から 9 9 のいずれかに記載の前記 1 セットのアダプタ。

(項目 1 0 1)

前記試料タグの各ヌクレオチド配列は、少なくとも 2 のハミング距離によって前記 1 セットのアダプタの前記試料タグのいずれかの他のヌクレオチド配列から分離する、項目 1 0 0 に記載の前記 1 セットのアダプタ。

(項目 1 0 2)

前記複数のアダプタのそれぞれは、前記試料タグ領域に隣接する、または前記試料タグ領域内に含まれる U M I 増倍管を備える、項目 9 0 から 1 0 1 のいずれか 1 項に記載の前記 1 セットのアダプタ。

(項目 1 0 3)

前記複数のアダプタのそれぞれは、前記試料タグ領域に隣接する U M I 増倍管を備える、項目 9 0 から 1 0 1 のいずれか 1 項に記載の前記 1 セットのアダプタ。

(項目 1 0 4)

前記複数のアダプタの各アダプタの前記アンカータグ領域は、4 個のヌクレオチド配列のうちの 1 個を含み、所与の配列の各試料領域は、所与の配列の前記 4 個のアンカー領域のうちの 1 個のみに対合される、項目 1 0 3 に記載の前記 1 セットのアダプタ。

(項目 1 0 5)

前記複数のアダプタの各アダプタの前記増幅領域は、同一のヌクレオチド配列を含み、各アダプタの前記試料タグ領域は、8 ヌクレオチド長にあり、前記試料タグの各ヌクレオチド配列は、少なくとも 2 のハミング距離によって前記 1 セットのアダプタの前記試料タグのいずれかの他のヌクレオチド配列から分離し、

前記複数のアダプタのそれぞれは、前記試料タグ領域に隣接する、または前記試料タグ領域内に含まれる U M I 増倍管を備え、前記複数のアダプタの各アダプタの前記 U M I 増倍管は、3 ヌクレオチド長にあり、前記 U M I 増倍管は、6 4 個の可能なヌクレオチド配列のうちの 1 個を含み、それぞれの前記 6 4 個の可能なヌクレオチド配列の前記 U M I 増倍管は、前記複数のアダプタの前記試料タグ領域のそれぞれに対合され、

前記複数のアダプタの各アダプタの前記アンカータグ領域は、4 個のヌクレオチド配列のうちの 1 個を含み、

所与の配列の各試料領域は、所与の配列の前記 4 個のアンカー領域のうちの 1 個のみに対合される、

項目 9 0 から 1 0 4 のいずれか 1 項に記載の前記 1 セットのアダプタ。

【図面の簡単な説明】

【 0 0 4 1 】

【図 1】コピー数減少 (C N L) アッセイのフレームワークを示す。各遺伝子 (縦列) は、本明細書に陰影によって表される特徴的なユニークリード値を示す。各試料 (横列) を同一の遺伝子パネルにわたり調べる。

【図 2】C N L アッセイシグナルのドライバーを示す図解である。

【図 3】無細胞 D N A (c f D N A) 上で実施される例示的な C N L アッセイのステップ

10

20

30

40

50

を説明する図解を示す。

【図 4 - 1】例示的な第一生成アダプタ（図 4 A 及び図 4 B）、及び本発明のアダプタ（図 4 C から図 4 E）の図解を示す。図 4 A は、第一生成アダプタ設計を示す。図 4 B は、第一生成アダプタ中で 249 個の可能な配列タグの採取があり、これらの各配列タグは単一のアンカー配列に付着した 5 ヌクレオチド（ n t）長にあったことを示す。図 4 C は、第二生成アダプタの図を示す。図 4 D は、4 セットの 8 m e r タグ配列からなり、各セットが 60 個のメンバーを含む、単一の試料に適用される例示的な 1 セットのアダプタを示す。60 個のタグの各セットは、4 個のアンカー配列のうちの 1 個に特異的である。図 4 E は、47 n t のアダプタの例示的な D N A 配列を示す。

【図 4 - 2】同上。

10

【図 5】図 5 A 及び図 5 B は、試料タグ内の U M I 増倍管の位置を移動させることによりユニーク試料タグの数が増加することが可能であることを説明する図を示す。

【図 6】C N L アッセイについてのゲノムライブラリを構築するプロセスを説明する図を示す。図 6 A は、10 n t のアンカー配列をゲノム断片の 3' 末端に付着させるステップを示す。図 6 B は、ゲノムアダプタの全長を最初のアンカー配列にアニールするステップを示す。

【図 7】D N A インプットを C N L ライブラリ中に示す。左に示されるマーカーサイズ（ $b p$ ）を有するアガロースゲル画像を示す。

【図 8】図 8 A ~ 図 8 C は、C N L 解析によって決定されるように 8 個の試料にわたる測定された遺伝子コピーの従来の箱ひげ図を示す。

20

【図 9】図 9 A 及び図 9 B は、断片化されたゲノム試料についての C N L 測定値における有意な標準からの偏差を定量化する $L o g_{10}$ の P 値プロットを示す。S N P 百分率は、A T M 及び B R C A 2 試料中に存在する希少なヘテロ接合 S N P のマイナーアレル頻度を最上部に示す。

【図 10】図 10 A 及び図 10 B は、断片化されたゲノム D N A によって添加される $c f$ D N A 試料についての C N L 測定値における有意な標準からの偏差を定量化する $L o g_{10}$ の P 値プロットを示す。S N P 百分率は、A T M 及び B R C A 2 試料中に存在する希少なヘテロ接合 S N P のマイナーアレル頻度を最上部に示す。

【図 11】標的ハイブリッドキャプチャプラットフォームを示す。図 11 A は、普遍的なシングルプライマー P C R 増幅配列、試料多重化タグ、及びユニーク分子識別子をすべてのゲノムクローンへ提供する追加のアダプタ配列によってゲノムライブラリへの $c f$ D N A の変換を示す。図 11 B は、標的特異的キャプチャプローブ及びプライマー伸長によってハイブリダイズされる変性し増幅されたゲノムを示す。図 11 C は、非対称ペアエンドシーケンシングの概略図を示す。図 11 D は、一般的な標的キャプチャシーケンスランから 377, 711, 020 個の I l l u m i n a N e x t S e q リードについてマッピング統計を示す。98.5% のリードをそれらの意図された標的へマッピングする。重複排除後、20.40% のリード（77, 053, 048）は、ユニークゲノムクローンに由来する。

30

【図 12 A】図 12 A ~ 図 12 H は、プール 1 - 3 からのオリゴヌクレオチドのアダプタの配列を示す。

40

【図 12 B】同上。

【図 12 C】同上。

【図 12 D】同上。

【図 12 E】同上。

【図 12 F】同上。

【図 12 G】同上。

【図 12 H】同上。

【図 13 A】図 13 A ~ 図 13 H は、プール 4 - 6 からのオリゴヌクレオチドのアダプタの配列を示す。

【図 13 B】同上。

50

【図 1 3 C】同上。

【図 1 3 D】同上。

【図 1 3 E】同上。

【図 1 3 F】同上。

【図 1 3 G】同上。

【図 1 3 H】同上。

【図 1 4 A】図 1 4 A ~ 図 1 4 I は、プール 7 - 9 からのオリゴヌクレオチドのアダプタの配列を示す。

【図 1 4 B】同上。

【図 1 4 C】同上。

【図 1 4 D】同上。

【図 1 4 E】同上。

【図 1 4 F】同上。

【図 1 4 G】同上。

【図 1 4 H】同上。

【図 1 4 I】同上。

【図 1 5 A】図 1 5 A ~ 図 1 5 H は、プール 1 0 - 1 2 からのオリゴヌクレオチドのアダプタの配列を示す。

【図 1 5 B】同上。

【図 1 5 C - 1】同上。

【図 1 5 C - 2】同上。

【図 1 5 D】同上。

【図 1 5 E】同上。

【図 1 5 F】同上。

【図 1 5 G】同上。

【図 1 6 A】図 1 6 A ~ 図 1 6 H は、プール 1 3 - 1 5 からのオリゴヌクレオチドのアダプタの配列を示す。

【図 1 6 B】同上。

【図 1 6 C】同上。

【図 1 6 D】同上。

【図 1 6 E】同上。

【図 1 6 F】同上。

【図 1 6 G】同上。

【図 1 6 H】同上。

【図 1 7 A】図 1 7 A ~ 図 1 7 H は、プール 1 6 - 1 8 からのオリゴヌクレオチドのアダプタの配列を示す。

【図 1 7 B】同上。

【図 1 7 C】同上。

【図 1 7 D】同上。

【図 1 7 E】同上。

【図 1 7 F】同上。

【図 1 7 G】同上。

【図 1 7 H】同上。

【図 1 8 A】図 1 8 A ~ 図 1 8 H は、プール 1 9 - 2 1 からのオリゴヌクレオチドのアダプタの配列を示す。

【図 1 8 B】同上。

【図 1 8 C】同上。

【図 1 8 D】同上。

【図 1 8 E】同上。

【図 1 8 F】同上。

10

20

30

40

50

【図 1 8 G】同上。

【図 1 8 H】同上。

【図 1 9 A】図 1 9 A ~ 図 1 9 H は、プール 2 2 - 2 4 からのオリゴヌクレオチドのアダプタの配列を示す。

【図 1 9 B】同上。

【図 1 9 C】同上。

【図 1 9 D】同上。

【図 1 9 E】同上。

【図 1 9 F】同上。

【図 1 9 G】同上。

【図 1 9 H】同上。

10

【図 2 0 A】図 2 0 A ~ 図 2 0 H は、プール 2 5 - 2 7 からのオリゴヌクレオチドのアダプタの配列を示す。

【図 2 0 B】同上。

【図 2 0 C】同上。

【図 2 0 D】同上。

【図 2 0 E】同上。

【図 2 0 F】同上。

【図 2 0 G】同上。

【図 2 0 H】同上。

20

【図 2 1 A】図 2 1 A ~ 図 2 1 H は、プール 2 8 - 3 0 からのオリゴヌクレオチドのアダプタの配列を示す。

【図 2 1 B】同上。

【図 2 1 C】同上。

【図 2 1 D】同上。

【図 2 1 E】同上。

【図 2 1 F】同上。

【図 2 1 G】同上。

【図 2 1 H】同上。

【図 2 2 A】図 2 2 A ~ 図 2 2 H は、プール 3 1 - 3 2 からのオリゴヌクレオチドのアダプタの配列を示す。

30

【図 2 2 B】同上。

【図 2 2 C】同上。

【図 2 2 D】同上。

【図 2 2 E】同上。

【図 2 2 F】同上。

【図 2 2 G】同上。

【図 2 2 H】同上。

【図 2 3】図 2 3 A ~ 図 2 3 C は、TP53 遺伝子の標的シーケンシングを示す。図 2 3 A は、キャプチャプローブの BedFile 表示を図示する。図 2 3 B は、0 から 8 0 0 0 のスケールのユニークリード上で各塩基位置におけるカバレッジ深度を示す。図 2 3 C は、既知の TP53 スプライス変異体の UCS C 遺伝子モデル表示を示す。より厚い矩形領域は、TP53 符号化タンパク質についてアミノ酸コード化領域を表す。

40

【図 2 4】図 2 4 A ~ 図 2 4 C は、16 個の試料にわたり、単一のプローブ、TP53 r 1 0 _ 1 について生の正規化されたユニークリード密度を示す。図 2 4 A は、「重複排除」によって冗長リードの除去後に 16 個の独立した試料に対してプローブ TP53 r 1 0 _ 1 によって生のユニークリードキャプチャ数を示す。図 2 4 B は、16 個の試料すべてについて 2 5 9 6 個のキャプチャプローブにわたりユニークリードの網羅的平均を示す。図 2 4 C は、16 個の試料 にわたる正規化されたユニークリード深度を示す ([プローブ TP53 r 1 0 _ 1 からの試料 n ユニークリード × 定数 ÷ 試料 n からの網羅的平均ユニ

50

ークリード/プローブ]として計算された)。

【図25】プローブ間の有意な平均深度変動にもかかわらず、いずれかの所与のTP53プローブ内の16個の試料のすべてについて正規化されたユニークリードカウントの一般的な一貫性を示す。16個の試料すべてについて正規化されたユニークリードカウントを狭い間隔の棒グラフの「柱」として示し、TP53を標的にする45個のプローブすべてについて結果を示す。「ノイズの多い」カウント挙動を示す2個のプローブを矢印で強調表示する。これらのようなプローブからのカウントは、後続のコピー数解析における外れ値として現れることが多い。

【図26】2596個のプローブの幅広のパネルにわたり正規化されたプローブごとのユニークリードカウントの試料間の一貫性を図示する。3個の代表的な試料からの散布図を示す。各点は、異なるプローブを表す。x軸は、16個の試料にわたる1個のプローブあたりの正規化された平均ユニークリード深度である。y軸は、3個の異なる個々の試料について1個のプローブあたりの正規化されユニークリード深度である。一貫したプローブごとのユニークリードカウントは、染色体コピー変動の定量的解析を支援する。

【図27】図27A～図27Cは、健康な女性及び男性ドナーからのcfDNA、ならびに進行期前立腺癌患者からのcfDNAのコピー数解析を図示する。図27Aは、健康な女性ドナーからのcfDNAの解析を示す。x軸は、22個すべての常染色体からの領域を標的にする一連の対照プローブ、X連鎖AR遺伝子を標的にする一連のプローブ、及びTP53遺伝子のコード化領域を標的にする一連のプローブである。y軸は、各プローブについて計算された倍数性を示す。この近似は、倍数性が既知である一連の対照試料に対して観察されたユニークリードカウントを正規化することによって各プローブについて計算される([試料__Zのプローブ__Yについてユニークリードカウント] × 2 ÷ [複数の対照試料についてのプローブ__Yに対する平均ユニークリードカウント])。図27Bは、X連鎖AR遺伝子が健康な男性における単数体コピー数を示すことを説明する。図27Cは、進行期前立腺癌患者からのcfDNAのコピー数解析を図示し、対照プローブ、AR遺伝子の増幅、及びTP53遺伝子の減少にわたり非常に有意な異数性の根拠を示す。

【図28】対照試料に関連する前立腺患者のcfDNAライブラリの全ゲノム異数性解析を示す。239個の各対照プローブについて近似した倍数性を染色体によって選別し、示す。患者の染色体の2個のプローブは、一貫したコピー減少を示し、大部分の染色体の5個のプローブは、コピー増加を示す。全てではないが多くの患者対照プローブについて、近似した倍数性の有意な偏差がみられる。

【図29】コピー数減少の検出の解析の妥当性確認を示す。不死化株NA02718(単一アレル ATM)からのゲノムDNA、及びNA09596(単一アレル BRCA2)からのゲノムDNAは、16%でNA12878からの「ゴールドスタンダード」ゲノムDNA中に添加され、8%の両アレル欠失のマイナーアレル頻度に相当するものをもたらした。標的シーケンシング及びCNV解析後に、プローブごとの倍数性を2個の標的遺伝子について平均した。比較のために、2個の攪乱のない対照遺伝子、BRIP1及びHDAC2を示す。

【発明を実施するための形態】

【0042】

A. 概要

本発明は、とりわけ、細胞ゲノムDNA(たとえば、組織診試料からの)またはcfDNA(たとえば、血液試料からの)の試料内の、変異性変化、SNP、転座、逆位、欠失、コピー数における変化、または他の遺伝的変異の検出のために有用である組成物及び方法を含む。本発明のこれらの組成物及び方法は、絶妙な分解能によって生体試料(たとえば、血液)からcfDNA中のコピー数多型を検出する、信じられないほど困難な検出をする際に特に有用である。特に、本発明のいくつかの実施形態は、アダプタへ付着するゲノムDNA断片から作製されるゲノムDNAライブラリを生成し、DNA標的領域を複数のキャプチャプローブによってキャプチャし、このDNA標的領域を含むDNAライブラリ断片を単離し、DNA標的領域の定量的遺伝子解析を実施することによりDNA標的領

10

20

30

40

50

域のコピー数を測定することによって、試験試料からDNA標的領域のコピー数を検出する方法について描画される。本明細書に記載されるアダプタは、試料の一致度またはゲノムDNAの供給源と同様に、配列されている個々のDNA断片の同定を可能にする。

【0043】

本発明は、直接組織診及び末梢血を含むがこれらに限定されない、いくつかの試料タイプに適用可能である標的に特異的なコピー数変化の検出のための組成物及び方法を部分的に企図する。癌ゲノミクス、特に、固形腫瘍の解析のための無細胞DNA(cfDNA)アッセイとの関連で、腫瘍DNA量は、DNA全体のごくわずかな割合であることが多い。さらにコピー数減少は、ゲノムDNAアッセイ、特にコピー数変化が試料、たとえば、cfDNAアッセイからの全ゲノムDNAの一部にのみ存在することができるゲノムDNAアッセイ中で検出することが困難である。たとえば、癌患者から抽出される無細胞DNAの大部分は、正常な供給源に由来し、二倍体コピー数を有する(男性被験体中のX連鎖遺伝子を除く)。癌患者、たとえば、血漿から抽出される2%の循環DNAが腫瘍に由来する患者などにおいて、腫瘍に由来するDNAの割合は、低いマイナーアレル頻度を有することが多い。腫瘍抑制遺伝子(たとえば、乳癌中のBRCA1)の1コピーの減少は、検出可能なゲノム断片が存在しないマイナーアレル頻度が1%であることを意味する。この状況において、操作されるコピー数減少アッセイは、100コピー(正常)から99コピー(ヘテロ接合性遺伝子消失)の間で識別することが可能でなければならない。したがって、特定の実施形態は、本発明の方法及び組成物がcfDNAとの関連でもマイナーアレル頻度でコピー数における変化を検出するのに十分な分解能によってコピー数変化の検出を可能にすることを企図する。

10

20

【0044】

この水準の識別を達成するために、本発明は、新規の試料アダプタ設計を提供する。本発明のこれらのアダプタは、(i)アダプタ全体に均一な性能、(ii)多数のユニーク分子識別子(UMI)、(iii)高効率の付着、及び(iv)試料多重化への適応能を含む、成功したコピー数減少のアッセイ性能について重要な特徴を有するように設計される。たとえば、本発明のアダプタは、以下を提供する。

【0045】

アダプタ全体に均一な性能：生物情報学的解析は、試料内プローブ性能及び試料間プローブ性能に目を向けることが多い。したがって、試料にわたるアダプタプール間のいずれかの性能変動がCNL解析に必要とされるわずかな変動を検出する能力に悪影響を与えると思定される。本発明において、この性能の均一性は、固定試料タグ領域(試料及びゲノム断片の両方を識別するように機能する)が各プールについてランダムに選択されながら、各試料タグプールにすべて見受けられる複数のアンカータグと、ゲノム断片を識別するためにユニーク試料タグ配列を増加させるUMI増倍管とを含むことによって達成される。

30

【0046】

多数のユニーク分子識別子(UMI)：これらのアダプタが分子生物学的観点から機能的に当量でなければならず、それらは、ユニークゲノム断片の同定を増強させる非常に多数のユニーク配列のタグ(10,000)を含まなければならない。これに関連して、「増強させる」ことによって、二本鎖DNAを切断したゲノム配列中の位置に対応する断片化部位の特定の対を各ゲノムクローン断片が含むことを意味する。各クローンが異なる切断部位を含む可能性が高いため、この切断部位を使用して、ユニークゲノムクローンを区別する。しかしながら、数千個の独立したクローンを含むライブラリにおいて、ユニークに由来する断片は、全く同一の切断部位を含むことが多い。同一の切断部位を共有するゲノムクローン(すなわち、断片)は、同一の試料に由来する他のクローン配列に関して、ユニークまたは冗長のいずれか一方として分類されることが可能である。非常に多様性である配列タグを導入するアダプタを付着させることにより、同一の切断部位を共有する異なるゲノムクローンは、ユニークとして識別される可能性がさらに高い。このシステムにおいて、UMIは、UMI増倍管と試料タグ領域の組み合わせによって作製される。UMI及び切断部位の組み合わせは、ユニーク分子識別子素子(UMIE)を作製し、

40

50

このUMIEは、冗長リードまたはユニークリードとして配列リードの分類を容易にする。特定の実施形態は、UMI増倍管がより長い配列、またはより短い配列を含み、全体的なUMI複雑性を増す、または低下させることが可能であることを企図する。

【0047】

高効率の付着：アダプタは、ゲノム断片に高効率で付着しなければならない。ほとんどの腫瘍学用途において、利用可能な細胞DNAまたはcfDNAの量は、限られるため、ゲノムライブラリクローンへのこれらのゲノム断片の変換は、非常に効率的でなければならない。これを達成するために、本発明のいくつかの態様において、本明細書に記載されるアダプタシステムは、約25%から約50%以上のゲノムインプット断片をゲノムライブラリクローンに変換する。

10

【0048】

試料多重化への適応能：一般に、アダプタの異なるセットのプールがなければならず、そこでセットの各ユニークアダプタを異なる試料に付着させる。同時に、アダプタのセットの各メンバーは、セット中のすべての他のメンバーへ本質的に同一挙動（配列カウント観点から）を有さなければならない。これを達成するために、いくつかの実施形態において、試料タグ領域は、いずれかの他の可能な試料タグ組み合わせ間で2のハミング距離を有し、リードが誤った試料に擬似的に割り当てられる可能性を減らす。いくつかの実施形態において、各セットのアダプタをプール中で分断し、特異的なアンカー領域と対合させ、試料多重化解除におけるエラーの可能性のさらなる減少を可能にする。たとえば、2のハミング距離を有する8merのタグにおいて、可能な配列の全数は、16,384個である。

20

【0049】

特定の実施形態において、アダプタのオリゴヌクレオチドの所定のプールを提供する。これらのような所定のプールを使用して、単一の試料を表す。すなわち、X個のアダプタのオリゴヌクレオチドの各プールにおける各アダプタ配列（上記の所与の例において16,384個）は、他の試料を識別するために使用されるあらゆる他のプール中の各アダプタ配列と異なる。当業者は、アダプタのオリゴヌクレオチドについて可能である別個の所定のプール数が試料タグ及び/またはUMI増倍管の長さに依存することを認識するであろう。

【0050】

30

したがって、ある一定の実施形態において、これらのアダプタは、1個の配列を含む、すなわち、試料タグ及び隣接する及び/または包含されたUMI増倍管は、両方の試料を表し、または識別し、遺伝子断片をユニークに識別する。これは、配列を識別するためにランダムに生成されたタグと、多重化を可能にするインデックスを付ける別々のバーコードまたはシーケンサとを使用する、当該技術分野において用いられる現在のシステムに対して全く対照的である。

【0051】

図3において、標的に特異的なコピー数変化を試料から取得されるDNA内で検出するために例示的な実施形態を示す。図3は、cfDNAからDNAライブラリを生成するが、この例示的な手順は、他の供給源、たとえば、断片化された細胞DNAなどからのDNAによって使用されることが可能である。図3に示されるように、cfDNAを捕集する（最上部のパネル）。つぎに、本発明のゲノムライブラリアダプタ（灰色の丸）をゲノムDNAにコンジュゲートすることによって、ゲノムライブラリをcfDNAから生成する。関心のゲノム領域を認識するキャプチャプローブ（黒色の丸）によってゲノムDNA断片を捕らえる。関心のゲノムDNAを配列決定し、関心のゲノムDNAのコピー減少解析及び/または特徴についてデータ解析を実施する。

40

【0052】

反対のことが明示されない限り、本発明における特定の実施形態の実施には、化学、生化学、有機化学、分子生物学、微生物学、組換えDNA技術、遺伝学、免疫学、及び細胞生物学における当業者の技能範囲内の従来的な方法が用いられることになり、これらの多

50

くを、例示の目的で後に説明する。このような技法は、文献で十分に説明されている。たとえば、Sambrook, et al., Molecular Cloning: A Laboratory Manual (3rd Edition, 2001); Sambrook, et al., Molecular Cloning: A Laboratory Manual (2nd Edition, 1989); Maniatis et al., Molecular Cloning: A Laboratory Manual (1982); Ausubel et al., Current Protocols in Molecular Biology (John Wiley and Sons, updated July 2008); Short Protocols in Molecular Biology: A Compendium of Methods from Current Protocols in Molecular Biology, Greene Pub. Associates and Wiley-Interscience; Glover, DNA Cloning: A Practical Approach, vol. I&II (IRL Press, Oxford, 1985); Anand, Techniques for the Analysis of Complex Genomes, (Academic Press, New York, 1992); Transcription and Translation (B. Hames & S. Higgins, Eds., 1984); Perbal, A Practical Guide to Molecular Cloning (1984); 及び Harlow and Lane, Antibodies, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1998) を参照する。

10

20

B. 定義

【0053】

別途定義されない限り、本明細書で使用する技術的用語及び科学的用語は、本発明が属する技術分野の当業者によって一般的に理解されている意味と同じ意味を有する。本発明の実施または試験の際に、本明細書に記載の方法及び材料に類似したまたは同等の、任意の方法及び材料を使用することは可能であるが、好ましい実施形態の組成物、方法、及び材料について説明する。本発明において、以下の用語は下記のように定義する。

【0054】

冠詞「a」、「an」、及び「the」は、本明細書では、冠詞の文法的対象の1つまたは複数（すなわち、少なくとも1つ）を指すように使用している。例として、「(an)要素」は、1つの要素または1つを超える要素を意味する。

30

【0055】

選択肢（例えば、「または(or)」）の使用は、選択肢のうちの1つ、両方、またはその任意の組合せのいずれかを意味するものと理解されたい。

【0056】

「及び/または(and/or)」という用語は、選択肢のうちの1つまたは両方のいずれかを意味するものと理解されたい。

【0057】

本明細書で使用する「約(about)」または「およそ(approximately)」という用語は、参照の量、レベル、値、数、頻度、パーセンテージ、寸法、サイズ、量、重量、または長さに対し、15%、10%、9%、8%、7%、6%、5%、4%、3%、2%、または1%の分変動する量、レベル、値、数、頻度、パーセンテージ、寸法、サイズ、量、重量、または長さを指す。一実施形態では、「約(about)」または「およそ(approximately)」という用語は、参照の量、レベル、値、数、頻度、パーセンテージ、寸法、サイズ、量、重量、または長さについて、±15%、±10%、±9%、±8%、±7%、±6%、±5%、±4%、±3%、±2%、または±1%の量、レベル、値、数、頻度、パーセンテージ、寸法、サイズ、量、重量、または長さの範囲を指す。

40

【0058】

50

本明細書全体において、文脈上他の意味が要求されない限り、「含む (comprise、comprises、comprising)」という語は、述べられるステップ若しくは要素またはステップ若しくは要素の群を含むことを意味し、ただし任意の他のステップ若しくは要素またはステップ若しくは要素の群を除外することを意味するものではないことを理解されたい。特定の実施形態では、「含む (include)」、「有する (has)」、「含有する (contain)」、及び「含む (comprise)」は同義語として使用される。

【0059】

「～からなる (consisting of)」は、「～からなる (consisting of)」という語句の後に来るものすべてを含み、それに限定されることを意味する。したがって、「～からなる (consisting of)」という語句は、挙げられる要素が必要または必須であり、かつ他の要素は存在しない場合があることを示す。

【0060】

「本質的に～からなる (consisting essentially of)」は、このフレーズの後に挙げられる任意の要素を含み、他の要素については、挙げられる要素の開示内容で指定された活性または作用を妨げず、寄与もしない要素に限定することを意味する。したがって、「本質的に～からなる (consisting essentially of)」というフレーズは、挙げられる要素は必要または必須であるが、他の要素については任意選択ではなく、挙げられる要素の活性または作用に影響を及ぼすかどうかに応じて、他の要素が存在する場合もあれば存在しない場合もあることを示す。

【0061】

本明細書全体における「一実施形態 (one embodiment)」、「一実施形態 (an embodiment)」、「特定の実施形態」、「関連した実施形態」、「ある特定の実施形態 (a certain embodiment)」、「追加的な実施形態」、若しくは「さらなる実施形態」、またはこれらの組合せに対する言及は、実施形態に関連して記載されている特定の特徵、構造、または特色が本発明の少なくとも1つの実施形態に含まれていることを意味する。したがって、本明細書全体の様々な箇所で出現する上記フレーズが、必ずしも全て同じ実施形態を言及しているわけではない。さらに、特定の特徵、構造、または特色は、1つ以上の実施形態における任意の好適な方法で組み合わせることができる。

【0062】

本明細書で使用する「単離した (isolated)」という用語は、天然の状態において通常は付随する構成要素を実質的または本質的に含まない材料を意味する。特定の実施形態では、「取得した (obtained)」または「導き出した (derived)」という用語は、「単離した」の同義語として使用される。

【0063】

本明細書で使用する「DNA」という用語は、デオキシリボ核酸を指す。様々な実施形態では、DNAという用語は、ゲノムDNA、組換えDNA、合成DNA、またはcDNAを指す。一実施形態では、DNAはゲノムDNAまたはcDNAを指す。特定の実施形態において、DNAは「標的領域」を含む。本明細書で企図されるDNAライブラリは、ゲノムDNAライブラリ、及びRNAから構築するcDNAライブラリ、例えば、RNA発現ライブラリなどが含まれる。様々な実施形態では、DNAライブラリは1つ以上の追加的なDNA配列及び/またはタグを含む。

【0064】

用語、「標的遺伝子座」及び「DNA標的領域」は、本明細書に互換的に使用され、DNA配列内の関心領域を指す。様々な実施形態では、標的遺伝子座に対し標的化遺伝子解析を実施する。特定の実施形態では、DNA標的領域は、特定の遺伝子様相、遺伝子状態、遺伝性疾患、胎児検査、遺伝子モザイク、父子鑑定、薬物治療に対する応答の予測、医学的状态の診断若しくは監視、マイクロバイームプロファイリング、病原体スクリーニング、または臓器移植監視に関連する遺伝子の領域である。さらに実施形態において、D

10

20

30

40

50

N A 標的領域は、特定のヒト染色体、たとえば、特定の常染色体若しくはX連鎖染色体、またはその領域（たとえば、ユニーク染色体領域）などに関連するDNA配列である。

【0065】

本明細書で使用する「循環DNA」、「循環セフリーDNA」、及び「セフリーDNA」という用語は、しばしば互換的に使用され、細胞外DNAであるDNA、細胞から押し出されたDNA、またはネクロシス性細胞若しくはアポトーシス性細胞から放出されたDNAを指す。この用語は、「細胞ゲノムDNA」または「細胞DNA」と対照的に使用されることが多く、本明細書に互換的に使用され、ゲノムDNAを指し、このゲノムDNAは、細胞（すなわち、ヌクレアーゼ）内に含まれ、また細胞の完全性にある、またはその他の方法により細胞の完全性を破壊することによって、本明細書に記載されるこれらのような分子生物学的技術に唯一到達可能である。

10

【0066】

本明細書で使用する「対象」、「個体」、または「患者」には、本明細書で企図する組成物により検出または識別され得る状態の症状を示す任意の動物が含まれる。好適な対象としては、実験動物（例えば、マウス、ラット、ウサギ、またはモルモット）、家畜（例えば、ウマ、ウシ、ヒツジ、ブタ）、飼育動物、すなわちペット（例えば、ネコまたはイヌ）が挙げられる。特定の実施形態では、対象は哺乳類である。あるいくつかの実施形態では、対象は非ヒト霊長類であり、好ましい実施形態では、対象はヒトである。

【0067】

本明細書に使用されるように、用語「対合された」は、異なるポリヌクレオチド配列を含むDNAの2個の異なるポリヌクレオチド配列または領域に関連して使用されるときに、異なるポリヌクレオチド配列を含むDNAの2個の異なるポリヌクレオチド配列または領域が同一のポリヌクレオチド上に存在することを意味する。たとえば、DNAの特定の試料タグ領域がDNAの特定の増幅領域に対合されると言われる場合に、試料タグ領域及び増幅タグが同一のDNAポリヌクレオチド分子上に存在することを意味する。

20

【0068】

C. コピー数解析方法

様々な実施形態では、DNA標的領域のDNAのコピー数解析方法を提供する。ある一定の実施形態において、コピー数解析は、ゲノムDNA断片及びアダプタを各含むDNAライブラリ断片のゲノムDNAライブラリを生成し、これらのDNA標的領域を含むDNAライブラリ断片を単離し、DNA標的領域の定量的遺伝子解析を行うことによって実施される。「定量的遺伝子解析」は、DNA変異、SNP、転座、欠失、及びコピー数多型（CNV）を含むが、これらに限定されない、DNA（たとえば、遺伝子、遺伝子座、関心の標的領域など）における変化を定量化することが可能である、いずれかの分子生物学的技術によって実施される解析を意味する。ある一定の実施形態では、定量的遺伝子解析をシーケンシング、たとえば、次世代シーケンシングによって実施する。

30

【0069】

次世代DNAシーケンシング（NGS）は、2つの診断用途に最適である。最初は、莫大な規模のDNA配列の決定である。本発明に関連して、この能力は、効果的な治療決定を導く、稀な、実施可能な変異体についての検索を可能にする。つぎは、遺伝子コピー数をカウントすることである。数百万個の独立した配列のアウトプットは、ゲノムワイドスケールで遺伝子コピー数の正確な測定を可能にすることができる。胎児トリソミー用の母体血試料からの非侵襲的な出生前検査の出現は、この能力に対する証明である。インプットがゲノムDNAよりもむしろRNA（cDNA）であるが、RNAseq、すなわち、NGSを使用する遺伝子発現プロファイリング技術は、別の例である。現在のキャプチャ方法の比較は、Samorodnitsky et al. J Mol Diagn. 2015 Jan; 17(1): 64-75、に記載される。

40

【0070】

本発明は、NGSカウント能力を標的ハイブリッドキャプチャ方法の範囲に拡張する。本明細書に記載されるこれらの方法は、それらが以下の4つの品質を有するために、少な

50

くとも部分的にコピー数多型の検出に効果的である。

(a) 本発明のこれらの方法は、ユニーククローンと冗長クローンとを区別する。増幅されたゲノムDNAライブラリ断片のNGSシーケンシングは、複数の個々のNGSリードをもたらし、これらのNGSリードは、特異的なヒトゲノム配列に結合されるアダプタ符号化配列情報を各含む。これらの素子は、すべてのクローンの一致度を定義する。キャプチャされたゲノム領域をPCRによって増幅するため、その後のNGS解析で同一のクローンに数回遭遇することは、珍しくはない。単一クローン化及びキャプチャプロセスに由来するリード群は、「冗長リード」と称される。2個以上の冗長リードは、ユニーク分子識別素子(UMIE)によって提供されるシーケンシング情報に基づき冗長リードとして識別される。UMIEは、アダプタタグからの配列情報、及びゲノムDNA配列の開始の組み合わせを指す。同一UMIEを含む2個以上のリードを冗長リードとして識別する。冗長リードを合わせてグループ化し、単一の代表的なコンセンサス配列を冗長リードのファミリーから組み立てる。このコンセンサス配列は、「ユニークリード」または「ユニークゲノム配列」(UGS)と称される。各ユニークリードは、元のDNA検体から分離したクローンを表す。冗長クローンファミリーを識別し、グループ化するプロセス、及びこのファミリーを表す単一のユニークリードを生成するプロセスを、「重複排除」と定義する。これらのアダプタは、ゲノムライブラリを作製するために使用され、ユニーク試料タグ情報(1個のアダプタあたり15,360個のコード)の非常に深いレパートリーをもつ。それぞれのキャプチャされたゲノムクローン(キャプチャプローブに関して>100の異なる位置に及ぶことが可能である)の正確なマッピング座標と併せて適用するとき、ゲノムライブラリに生成された後に、標的特異的キャプチャプローブによって取得される各ユニーククローンは、同一のキャプチャ環境を包含するすべての他のユニーククローンと区別可能である非常に高い尤度を有する。ユニーククローンを冗長クローンと区別する能力は、本明細書に記載される方法に中心となるものである。

(b) ゲノムライブラリを作製するために使用されるアダプタは、アダプタ間のばらつきをコピー数カウントにおいて生じることなく試料多重化を可能にする。コピー数決定の重要な基礎は、単一のシーケンシングラン内ですべて処理された1セットの試料の同時解析である。これにより、正及び負の対照を臨床試料に加えて含むことが可能である。以前のアダプタ設計繰り返しに関する主な課題は、同一の対照試料間の遺伝子コピーカウントにおけるわずかなシフトを誘導し、実際には、血液に基づく、固形腫瘍遺伝子型判定アッセイに臨床的に有益であるには高すぎたシグナル対ノイズの不確実性閾値を設定した。本発明は、この課題を克服し、シングルコピー遺伝子減少が2%のマイナーアレル頻度で検出可能であるように、シグナル対ノイズ閾値を実質的に下げる。この改善されたシグナル認識は、本発明の方法が腫瘍DNAアッセイを循環させる際に有意な臨床的有用性を有することを可能にする。

(c) 本明細書に使用される独自開発の標的ハイブリッドキャプチャ方法は、すべての標的にわたり非常に均一な「標的上」リードカバレッジを作製しなければならない。本明細書に記載されるものなどの、ユニークゲノム断片のカウントに頼り、コピー数を推定する方法は、すべての可能なユニーク断片に遭遇する観点から近飽和を達成しなければならない。オーバーサンプリング、すなわち、最終的に遭遇するユニークリード数より多いシーケンシングリードを集めることによってのみ近飽和を達成する。実用的、スケーラブル、及び経済的であるために、標的上リードの<10倍のオーバーサンプリング、及び好ましくは標的上リードの<4倍のオーバーサンプリングが>90%のユニーク標的上リードをすべての標的遺伝子座でキャプチャするように、標的ハイブリッドキャプチャライブラリ中のユニークリードは、十分な均一性を示さなければならない。

(d) 標的ハイブリッドキャプチャ方法(米国特許公開第2014-0274731号を参照する)は、高い標的上捕捉率を有さなければならない。実用的、スケーラブル、及び経済的であるために、換言すれば、この分野における他の技術と比較して本開示の顕著な特徴であるために、この方法は、>90%、好ましくは>95%の標的上リードを達成しなければならない。標的上マッピング率が95%を超えながら、標的上リードの4から

10

20

30

40

50

10 倍のオーバーサンプリングについての要件、及びオーバーサンプリング全体についての要件は、同一のものである。

【0071】

いくつかの実施形態では、試料中に存在するDNA標的領域のコピー数を定量的遺伝子解析によって決定する。いくつかの実施形態において、試料中に存在するDNA標的領域のコピー量を比較すること、またそれを既知のコピー数を有する1つ以上の試料中に存在するDNA標的領域の量と比較することによって、DNA標的領域のコピー数を決定する。

【0072】

特定の実施形態は、本明細書に記載される組成物及び方法がゲノムDNAの試料中のコピー数における変化を検出するために特に有益であることを企図し、そこで試料中の全ゲノムDNAの一部のみがコピー数における変化を有する。たとえば、有意な腫瘍変異は、試料、たとえば、無細胞DNAの試料などの中に存在する可能性があり、すなわち、アレル頻度が一般的に約100%、50%または0%である従来のSNP遺伝子型判定と比較して、有意に50%未満で（たとえば、0.1%から>20%の範囲に）あるマイナーアレル頻度に存在する。当業者は、本発明の組成物及び方法が一塩基多様体（SNV）、短い（たとえば、40塩基対（bp）未満の）挿入、及び欠失（インデル）を含む変異の他のタイプと、発癌遺伝子融合を含むゲノム再編成とを検出する際にまた有用であることを認識するであろう。

【0073】

ある一定の実施形態において、本明細書に記載される本発明の組成物及び/または方法は、1つ以上のDNA標的領域のコピー数における変化を検出する、識別する、観察する、及び/または明らかにするために有用であり、これらを行うことができ、これらを行うことに適しており、及び/またはこれらを行うことができ、このコピー数における変化は、試料からの全ゲノムDNAの、約20%未満、約19%未満、約18%未満、約17%未満、約16%未満、約15%未満、約14%未満、約13%未満、約12%未満、約11%未満、約10%未満、約9%未満、約8%未満、約7%未満、約6%未満、約5%未満、約4%未満、約3%未満、約2%未満、約1%未満、約0.5%未満、約0.2%未満、または約0.1%未満に存在する。いくつかの実施形態において、本発明の方法は、試料からの全ゲノムDNAの、約0.01%から約100%、約0.01%から約50%、及び/または約0.1%から約20%の間に存在する1つ以上のDNA標的領域のコピー数における変化を検出する、識別する、観察する、及び/または明らかにするために有用であり、これらを行うことが可能であり、これらを行うのに適しており、及び/またはこれらを行うことができる。

【0074】

特定の実施形態を図1に図示される概念的なフレームワークによって表す。図1において、各遺伝子を縦列によって表し、各患者試料を横列として表す。いずれかの所与のゲノムDNA試料内に、各個々の遺伝子について計数される断片数は、ある程度のばらつきを有し、そしてその関心のいずれかの所与のDNA領域、たとえば、遺伝子などについて、他の試料中のDNA標的領域へ正規化されたカウントに関して有意な断片カウント偏差としてコピー数における攪乱を検出する。このようなアッセイは、試料内で遺伝子ごとの断片カウントプロファイルが再現性のあることを要求し、また試料ごとのカウントプロファイルは同等性が高いことを要求する。アッセイ要件の両方は、良好なシグナル対ノイズカウント識別を要求する。

【0075】

いくつかの実施形態は、図2に示されるように、シグナル対ノイズ比を増すことに寄与するアッセイ要素がゲノムインプット、プローブ数、及びシーケンシング深度であることを企図する。

【0076】

特定の実施形態では、cfDNAの遺伝子解析方法は、cfDNAライブラリを生成及び増幅すること、cfDNAライブラリにおけるゲノム当量数を決定すること、ならびに

10

20

30

40

50

1つ以上のゲノム標的遺伝子座の定量的遺伝子解析を実施することを含む。

【0077】

特定の実施形態は、本明細書に記載される方法及び組成物のいずれかが、ゲノムDNA、たとえば、細胞DNAまたはcfDNAなどを使用して、遺伝子状態、遺伝子条件、遺伝性疾患、遺伝子モザイク現象、胎児診断、父子鑑定、マイクロバイームプロファイリング、病原体スクリーニング、及び臓器移植監視を効率的に解析する、検出する、診断する、及び/または監視するための使用に効果的であり、そこで試料中の全ゲノムDNAのすべて、または一部のみが関心特徴、たとえば、遺伝子損傷、変異、一塩基多様体(SNV)などを含むことを企図する。いくつかの実施形態において、関心特徴は、疾患または状態と関連する遺伝子特徴である。たとえば、有意な腫瘍変異は、試料、たとえば、cfDNAの試料などの中に存在する可能性があり、アレル頻度が一般的に約100%、50%または0%である従来のSNP遺伝子型判定と比較して、有意に50%未満で(たとえば、0.1%から>20%の範囲に)あるマイナーアレル頻度に存在する。

10

【0078】

ある一定の実施形態において、本明細書に記載される本発明の組成物及び/または方法は、1つ以上のDNA標的領域の遺伝子損傷を検出する、識別する、観察する、及び/または明らかにするために有用であり、これらを行うことが可能であり、これらを行うことに適し、及び/またはこれらを行うことができ、1つ以上のDNA標的領域のこの遺伝子損傷は、試料からの全ゲノムDNAの、約20%未満、約19%未満、約18%未満、約17%未満、約16%未満、約15%未満、約14%未満、約13%未満、約12%未満、約11%未満、約10%未満、約9%未満、約8%未満、約7%未満、約6%未満、約5%未満、約4%未満、約3%未満、約2%未満、約1%未満、約0.5%未満、約0.2%未満、または約0.1%未満に存在する。いくつかの実施形態において、本発明の方法は、試料からの全ゲノムDNAの、約0.01%から約100%、約0.01%から約50%、及び/または約0.1%から約20%の間に存在する1つ以上のDNA標的領域の遺伝子損傷を検出する、識別する、観察する、及び/または明らかにするために有用である、これらを行うことが可能である、これらを行うのに適している、及び/またはこれらを行うことができる。

20

【0079】

1. DNAライブラリの生成

30

特定の実施形態において、本明細書に企図される遺伝子解析方法は、末端修復DNAを生成するために1つ以上の末端修復酵素によって処置するcfDNAまたは断片化細胞ゲノムDNAを含むDNAライブラリを生成すること、及びDNAライブラリを生成するために末端修復DNAの各末端に1つ以上のアダプタを付着させることを備える。

【0080】

ゲノムDNA

特定の実施形態において、本明細書で企図される方法及び組成物は、ゲノムDNAをアナライトとして使用して、コピー数における変化を効率的に解析、検出、診断、及び/または監視するように設計される。ある一定の実施形態において、コピー数解析は、試験試料、たとえば、組織診などの生体試料などから得られるゲノムDNAからゲノムDNAライブラリを生成することによって実施される。ある一定の実施形態において、ゲノムDNAは、循環または無細胞DNAである。いくつかの実施形態において、ゲノムDNAは、細胞ゲノムDNAである。

40

【0081】

ある一定の実施形態において、骨髄、食道、胃、十二指腸、直腸、結腸、回腸、膵臓、肺、肝臓、前立腺、脳、神経、髄膜組織、腎組織、子宮内膜組織、子宮頸部組織、乳房、リンパ節、筋肉、及び皮膚を含むが、これらに限定されない、組織試料、または組織から得られる組織診からゲノムDNAを取得する。ある一定の実施形態において、組織試料は、腫瘍、または疑わしい腫瘍の組織診である。特定の実施形態において、腫瘍は、癌性である、または癌性であると疑われる。特定の実施形態において、組織試料は、癌細胞、ま

50

たは癌性であると疑われる細胞を含む。

【 0 0 8 2 】

細胞からの、または細胞から構成される生体組織からのゲノムDNAを精製する方法は、当該技術分野において周知であり、当業者は、組織、及び組織を取得する条件に応じて、最適な手順または市販のキットを認識するであろう。いくつかの実施形態は、組織からの細胞DNAを精製することにより、たとえば、化学的方法及び物理的方法によって、細胞DNAを曝露するために細胞破碎または細胞溶解を必要とすることを企図し、これらの化学的方法及び物理的方法は、組織試料を配合する、粉碎する、または超音波処理すること、細胞溶解にも機能する界面活性剤またはサーファクタントを加えることによって、膜脂質を除去し、任意選択で、たとえば、プロテアーゼを加えることなどによって、タンパク質を除去すること、たとえば、RNaseを加えることなどによって、RNAを除去すること、ならびにたとえば、細胞溶解ステップ中に使用される界面活性剤、タンパク質、塩類及び試薬類からの、DNA精製などである。たとえば、エタノールまたはイソプロパノールによる、沈殿によって、またフェノール-クロロホルム抽出によって、DNA精製を実施することができる。

10

【 0 0 8 3 】

特定の実施形態において、本明細書に記載されるようにゲノムDNAライブラリを取得する、生成する、作製する、形成する、及び/または産生する前に、またはこれらをする間に、組織及び/または細胞から取得される細胞DNAを断片化する。当業者は、DNA断片化についていくつかの適切な技術があることを理解し、次世代シーケンシングを有するが、これに限定されない、DNAシーケンシングについてゲノムDNAライブラリを生成するために、細胞DNAを断片化するために適切な技術を認識し、識別することが可能である。ある一定の実施形態は、物理的断片化、酵素的断片化、及び化学的せん断を有するが、これらに限定されない、方法によってライブラリを生成するために、適切な、及び/または十分な長さの断片に細胞DNAを断片化することが可能であることを企図する。

20

【 0 0 8 4 】

物理的断片化は、音響学的せん断、超音波処理、及び流体力学的せん断を有するが、これらに限定されないことが可能である。いくつかの実施形態において、物理的断片化によって細胞DNAを断片化する。特定の実施形態では、音響学的せん断または超音波処理によって、細胞DNAを断片化する。特定の実施形態は、音響学的せん断及び超音波処理が細胞DNAをせん断するために使用される共通の物理的方法であることを企図する。Covaris（登録商標）機器（Woburn, MA）は、DNAを100から5kbpで破壊するための音響装置である。またCovarisは、メイトペアライブラリについて、試料を6から20kbで処理するチューブ（gTube）を製造する。Bioruptor（登録商標）（Denville, NJ）は、クロマチン、DNAをせん断し、組織を攪乱するために利用される超音波処理装置である。わずかな量のDNAを150から1kbの長さでせん断することが可能である。Digilab（Marlborough, MA）からのHydroShearは、DNAをせん断するために流体力を利用する。またNebulizer（Life Tech, Grand Island, NY）を使用して、圧縮空気をを用いて液体を微粒化し、DNAを100から3kb断片に数秒でせん断することが可能である。噴霧化は、低コストであるが、このプロセスは、元の試料からの細胞DNAの約30%の減少を引き起こすことがある。ある一定の実施形態において、超音波処理によって、細胞DNAを断片化する。

30

40

【 0 0 8 5 】

酵素的断片化は、制限エンドヌクレアーゼ、たとえば、DNase Iによる処置、または非特異的ヌクレアーゼによる処理を含むが、これらに限定されないことが可能である。いくつかの実施形態において、酵素的断片化によって細胞DNAを断片化する。特定の実施形態において、制限エンドヌクレアーゼによる処置によって、細胞DNAを断片化する。いくつかの実施形態において、非特異的ヌクレアーゼによる処置によって、細胞DNAを断片化する。ある一定の実施形態において、トランスポサーゼによる処置によって、

50

細胞DNAを断片化する。ある一定の実施形態は、細胞DNAを小片にせん断する酵素的な方法がDNAse I、マルトース結合タンパク質(MBP)-T7 Endo I及び非特異的ヌクレアーゼの組み合わせ、*Vibrio vulnificus*(Vvn) New England Biolabs(Ipswich, MA)のFragmentase、及びNextera tagmentation technology(Illumina, San Diego, CA)を含むことを企図する。非特異的ヌクレアーゼ及びT7 Endo Iの組み合わせは、相乗的に機能し、非特異的ニック及びカウタニックを産生し、ニック部位から8ヌクレオチド以下を解離する断片を生成する。タグメンテーションは、アダプタを二本鎖DNA上で同時に断片化し、挿入するトランスポサナーゼを使用する。

【0086】

10

化学的断片化は、加温及び二価の金属カチオンによる処置を有することが可能である。いくつかの実施形態において、化学的断片化によってゲノムDNAを断片化する。特定の実施形態は、ゲノムDNAに対向するように長いRNA断片の粉碎のために化学的せん断をより一般的に使用することを企図する。化学的断片化は、二価の金属カチオン(マグネシウムまたは亜鉛)によってDNAの加温消化を通して一般的に実施される。インキュベーション時間を増加させる、または減少させることによって、DNA断片の長さを調整することが可能である。

【0087】

特定の実施形態において、本明細書で企図される方法及び組成物は、無細胞DNA(cfDNA)をアナライトに使用して、コピー数における変化を効率的に解析、検出、診断、及び/または監視するように設計される。cfDNAのサイズ分布は、約150bpから約180bp断片の範囲である。cfDNAの断片化は、エンドヌクレアーゼ及び/またはエキソヌクレアーゼ活性の結果であり得るものであり、この断片化により、正確で信頼性の高い口バストなcfDNA解析に対する困難な課題が提示される。cfDNA解析に関するもう1つの課題は、血流における半減期が約15分程度と短いことである。いかなる特定の理論にも拘泥されることは望まないが、本発明は、その一部として、cfDNAの解析が「リキッドバイオプシー」に類似するものであり、現在の生体プロセスのリアルタイムスナップショットであることを企図している。

20

【0088】

さらに、cfDNAは、細胞内に見出されず、また生物学的流体及び便試料を含むが、これらに限定されない、複数の適切な供給源から得ることができるため、解析される組織への直接アクセスなどの、次世代シーケンシング解析を困難にする既存の制限を受けない。

30

【0089】

特定の実施形態においてcfDNAを単離するための好適な供与源となる生物学的流体の例示的な例は、以下に限定するものではないが、羊水、血液、血漿、血清、精液、リンパ液、脳脊髄液、眼液、尿、唾液、粘液、及び汗が挙げられる。特定の実施形態では、生物学的流体は血液または血漿である。

【0090】

ある一定の実施形態では、市販されているキット及び当業者に知られた他の方法を使用して、患者の生物学的流体から、または過去に取得し、任意選択により安定化させた生物学的試料(例えば、凍結及び/または酵素キレート剤(EDTA、EGTA、または2価カチオンに特異的な他のキレート剤を含むが、これに限定されない)による安定化)から、直接cfDNAを単離することができる。

40

【0091】

(a) 末端修復cfDNAの生成

特定の実施形態において、ゲノムDNAライブラリを生成することは、単離されたcfDNAまたは断片化細胞DNAの末端修復を備える。断片化cfDNAまたは細胞DNAを末端修復酵素によって処理し、平滑末端、5'-オーバーハング、または3'-オーバーハングによって、末端修復cfDNAを生成する。いくつかの実施形態において、末端修復酵素は、たとえば産生することが可能である。いくつかの実施形態において、末端修復

50

c f D N Aまたは細胞D N Aは、平滑末端を含有する。いくつかの実施形態では、末端修復細胞D N Aまたはc f D N Aは、平滑末端を含有するように処理される。いくつかの実施形態では、末端修復c f D N Aまたは細胞D N Aの平滑末端は、単一塩基対オーバーハングを含有するように、さらに修飾される。いくつかの実施形態では、平滑末端を含む末端修復c f D N Aまたは細胞D N Aは、アデニン（A）/チミン（T）オーバーハングを含有するようにさらに処理されることが可能である。いくつかの実施形態では、平滑末端を含む末端修復c f D N Aまたは細胞D N Aは、単一塩基対オーバーハングとしてアデニン（A）/チミン（T）オーバーハングを含有するように、さらに処理されることが可能である。いくつかの実施形態では、末端修復c f D N Aまたは細胞D N Aは、非鋳型化3'オーバーハングを有する。いくつかの実施形態では、末端修復c f D N Aまたは細胞D N Aは、3'オーバーハングを含むように処理される。一部の実施形態では、末端修復c f D N Aまたは細胞D N Aは、3'オーバーハングを含有するように末端トランスフェラーゼ（T d T）で処理される。一部の実施形態では、T d TによりGテールを加えてもよい。いくつかの実施形態では、末端修復c f D N Aまたは細胞D N Aは、オーバーハング末端を含有するように、任意の公知の制限酵素(例えば、S a u 3 A酵素など)による部分的消化を用いて処理される。

10

【0092】

(b) 末端修復c f D N Aへのアダプタ分子の付着

特定の実施形態において、c f D N Aライブラリを生成することは、1つ以上のアダプタを末端修復c f D N Aの各末端へ付着させることを備える。本発明は、部分的に、c f D N Aライブラリ中の多数のゲノム当量を収容するように設計されるアダプタモジュールを企図する。アダプタモジュールは、c f D N Aライブラリに存在するゲノム当量数を測定するように構成され、伸長によって、シーケンシングアッセイの感度を使用して、配列変異を識別した。

20

【0093】

本明細書に使用されるように、用語「アダプタ」及び「アダプタモジュール」は、互換的に使用され、また増幅領域、試料タグ領域、及びアンカー領域の少なくとも3つの要素を含むポリヌクレオチドを指す。特定の実施形態において、アダプタは、増幅領域、試料タグ領域、及びアンカー領域を含む。いくつかの実施形態において、このアダプタは、ユニーク分子識別子（U M I）をも含む。特定の実施形態では、このアダプタは、1つ以上の増幅領域、1つ以上の試料タグ領域、1つ以上のU M I、及び/または1つ以上のアンカー領域を含む。いくつかの実施形態において、このアダプタは、5'から3'の順に、増幅領域、試料タグ領域、U M I、及びアンカー領域を含む。特定の実施形態において、このアダプタは、5'から3'の順に、増幅領域、試料タグ領域、U M I、及びアンカー領域を含む。ある一定の実施形態において、U M Iは、試料タグ領域内に含まれ、このアダプタは、5'から3'の順に、増幅領域、組み込まれた試料タグ/U M I領域、及びアンカー領域を含む。

30

【0094】

本明細書に使用されるように、用語「増幅領域」は、P C R増幅についてプライマー認識部位として機能することが可能であるポリヌクレオチド配列を含むアダプタ分子の要素を指す。特定の実施形態では、アダプタは、ゲノムD N Aライブラリのシングルプライマー増幅について1つ以上のプライマー認識配列を含有する増幅領域を含む。いくつかの実施形態では、増幅領域は、ゲノムD N Aライブラリのシングルプライマー増幅について、1個、2個、3個、4個、5個、6個、7個、8個、9個、10個、または10個超のプライマー認識配列を含む。

40

【0095】

いくつかの実施形態において、増幅領域は、約5から50の間のヌクレオチド長、10から45の間のヌクレオチド長、15から40の間のヌクレオチド長、または20から30の間のヌクレオチド長にある。いくつかの実施形態では、増幅領域は、10個のヌクレオチド、11個のヌクレオチド、12個のヌクレオチド、13個のヌクレオチド、14個

50

のヌクレオチド、15個のヌクレオチド、16個のヌクレオチド、17個のヌクレオチド、約18個のヌクレオチド、19個のヌクレオチド、20個のヌクレオチド、21個のヌクレオチド、22個のヌクレオチド、23個のヌクレオチド、24個のヌクレオチド、25個のヌクレオチド、26個のヌクレオチド、27個のヌクレオチド、28個のヌクレオチド、29個のヌクレオチド、30個のヌクレオチド、31個のヌクレオチド、32個のヌクレオチド、33個のヌクレオチド、34個のヌクレオチド、35個のヌクレオチド、36個のヌクレオチド、37個のヌクレオチド、38個のヌクレオチド、39個のヌクレオチド、または40個以上のヌクレオチドである。特定の実施形態において、増幅領域は、25ヌクレオチド長にある。

【0096】

本明細書に使用されるように、用語「試料タグ」または「試料タグ領域」は、互換的に使用され、それが由来した試料と同様に特定のDNA断片をユニークに識別するポリヌクレオチド配列を含むアダプタ素子を指す。

【0097】

ある一定の実施形態において、試料タグ領域は、約3から50の間のヌクレオチド長、約3から25の間のヌクレオチド長、または約5から15の間のヌクレオチド長にある。いくつかの実施形態では、試料タグ領域は、3ヌクレオチド長、4ヌクレオチド長、5ヌクレオチド長、6ヌクレオチド長、7ヌクレオチド長、8ヌクレオチド長、9ヌクレオチド長、10ヌクレオチド長、約11ヌクレオチド長、12ヌクレオチド長、13ヌクレオチド長、14ヌクレオチド長、15ヌクレオチド長、16ヌクレオチド長、17ヌクレオチド長、18ヌクレオチド長、19ヌクレオチド長、または20ヌクレオチド長以上にある。

【0098】

ある一定の実施形態では、アダプタは、UMI増倍管を含み、そこでUMI増倍管は、少なくとも1ヌクレオチド長、少なくとも2ヌクレオチド長、少なくとも3ヌクレオチド長、少なくとも4ヌクレオチド長、少なくとも5ヌクレオチド長、少なくとも6ヌクレオチド長、少なくとも7ヌクレオチド長、少なくとも8ヌクレオチド長、少なくとも9ヌクレオチド長、または少なくとも10ヌクレオチド長にある。

【0099】

ある一定の実施形態において、UMI増倍管の各ヌクレオチド位置は、アデニン、グアニン、シトシン、またはチミンのいずれかを含むことが可能である。したがって、いくつかの実施形態において、 n 個のヌクレオチドを含むUMI増倍管は、 n^4 個の可能なヌクレオチド配列のいずれかを含むことが可能である。いくつかの実施形態において、UMI増倍管は、1ヌクレオチド長にあり、4個の可能な配列のうちの1個を含む。いくつかの実施形態において、UMI増倍管は、2ヌクレオチド長にあり、16個の可能な配列のうちの1個を含む。いくつかの実施形態において、UMI増倍管は、3ヌクレオチド長にあり、64個の可能な配列のうちの1個を含む。いくつかの実施形態において、UMI増倍管は、4ヌクレオチド長にあり、256個の可能な配列のうちの1個を含む。いくつかの実施形態において、UMI増倍管は、5ヌクレオチド長にあり、1,024個の可能な配列のうちの1個を含む。いくつかの実施形態において、UMI増倍管は、6ヌクレオチド長にあり、4,096個の可能な配列のうちの1個を含む。いくつかの実施形態において、UMI増倍管は、7ヌクレオチド長にあり、16,384個の可能な配列のうちの1個を含む。いくつかの実施形態において、UMI増倍管は、8ヌクレオチド長にあり、65,536個の可能な配列のうちの1個を含む。いくつかの実施形態において、UMI増倍管は、9ヌクレオチド長にあり、262,144個の可能な配列のうちの1個を含む。いくつかの実施形態において、UMI増倍管は、10以上のヌクレオチド長にあり、1,048,576個以上の可能な配列のうちの1個を含む。

【0100】

特定の実施形態において、アダプタは、UMI増倍管を含み、そこでUMI増倍管は、試料タグ領域(図5A)に隣接する、またはこの試料タグ領域内に含まれる。試料タグに

10

20

30

40

50

隣接する、または試料タグ内に含まれるUMI増倍管の例示的な例を図5Bに示す。図5Bにおいて、隣接するUMI増倍管（縦列の最上部及び最下部）、または試料タグ（中間の7縦列）内に組み込まれるUMI増倍管を含む、8-mer試料タグ領域を示す。いくつかの実施形態において、そのアダプタは、8ヌクレオチド長にある試料タグ、及び3ヌクレオチド長にあり、64個の可能な配列のうちの1個を含むUMI増倍管を含み、そこでUMI増倍管は、試料タグ領域に隣接する、またはこの試料タグ領域内に含まれる。いくつかの実施形態において、同一のプロセスは、アダプタ全長をゲノム断片の他の末端に付着させる。

【0101】

特定の実施形態では、アダプタモジュールは、1つ以上のアンカー配列を含む。本明細書に使用されるように、「アンカー領域」及び「アンカー配列」は、互換的に使用され、パートナーオリゴヌクレオチドをハイブリダイズするヌクレオチド配列を指す。いくつかの実施形態では、アンカー領域は、以下の3つの特性を有する。（1）各アンカー配列は、伸長内の各部位に4個の可能なDNA塩基のそれぞれを合わせて表す2個以上のアンカー配列のファミリーの部分であり、この特徴、平衡した塩基表現は、特定の実施形態においてシーケンシングリードにおける適切な塩基呼び出しを校正するために有用である。（2）各アンカー配列は、4個の可能な塩基のうちの2個のみから構成され、そしてこれらは、いずれかの数の、また等しい数のA+C、または等しい数のG+Tであるように特異的に選択され、2個の塩基のみから形成されるアンカー配列は、アンカー配列が適切なアダプタ機能を妨げる二次構造の形成に関与する可能性を減少させる。そして（3）各アンカー配列は、等しい数のA+CまたはG+Tから構成され、各アンカー配列は、1セットの4個中のすべての他のアンカー配列と同一の融解温度及び二本鎖安定性を大まかに共有する。

【0102】

いくつかの実施形態において、アンカー配列は、1から50の間のヌクレオチド長にある。いくつかの実施形態において、アンカー配列は、4から40の間のヌクレオチド長にある。ある一定の実施形態において、アンカー領域は、5から25の間のヌクレオチド長にある。特定の実施形態では、アンカー領域は、少なくとも4ヌクレオチド長、少なくとも6ヌクレオチド長、少なくとも8ヌクレオチド長、少なくとも10ヌクレオチド長、少なくとも12ヌクレオチド長、少なくとも14ヌクレオチド長、または少なくとも16ヌクレオチド長にある。特定の実施形態において、アンカー領域は、10ヌクレオチド長にある。

【0103】

特定の実施形態では、付着ステップは、「タグ付き」ゲノムDNAライブラリを生成するためにアダプタモジュールを末端修復cfDNAまたは細胞DNAに付着させる/ライゲーションすることを備える。いくつかの実施形態では、単一アダプタモジュールを用いる。いくつかの実施形態では、2種、3種、4種、または5種のアダプタモジュールを用いる。いくつかの実施形態では、同一配列のアダプタモジュールを、断片化した末端修復DNAの各末端に付着させる。

【0104】

いくつかの実施形態では、複数種のアダプタを末端修復細胞または無細胞ゲノムDNA断片に付着させる。これらの複数種のアダプタのそれぞれは、cfDNAまたは細胞DNAライブラリの増幅について1個以上の増幅領域と、cfDNAまたは細胞ゲノムDNA断片の同定、及び個々の試料の同定について1個以上の試料タグ領域と、DNAシーケンシングについて1個以上の配列とを含むことができる。

【0105】

いくつかの実施形態において、複数種のアダプタは、試料の末端修復細胞または無細胞ゲノムDNA断片に付着し、複数種のアダプタは、同一のヌクレオチド配列の増幅領域をすべて含む。

【0106】

10

20

30

40

50

ある一定の実施形態において、試料からのゲノムDNAを、複数のアダプタと付着させ、これらの複数のアダプタは、他の試料からのゲノムDNA断片に付着するアダプタ中の試料タグ領域の他の配列とすべて異なる試料タグ配列を含む。

【0107】

特定の実施形態において、複数種のアダプタを試料からの末端修復細胞または無細胞ゲノムDNA断片に付着させ、これらの複数のアダプタは、2から10,000個の間のヌクレオチド配列のうちの1個、5から5,000個の間のヌクレオチド配列のうちの1個、25から1,000個の間のヌクレオチド配列のうちの1個、50から500個の間のヌクレオチド配列のうちの1個、100から400個の間のヌクレオチド配列のうちの1個、または200から300個の間のヌクレオチド配列のうちの1個を含有する1個以上の試料タグ領域をすべて含む。いくつかの実施形態において、各アダプタの試料タグ領域は、8ヌクレオチド長にあり、複数のアダプタの各試料タグ領域は、240個のヌクレオチド配列のうちの1個を含む。

10

【0108】

ある一定の実施形態において、複数種のアダプタは、試料からの末端修復細胞または無細胞ゲノムDNA断片に付着し、複数種のアダプタの試料タグ領域は、1、2、3、4または4を上回るハミング距離によって互いに異なるヌクレオチド配列を含む。特定の実施形態では、ハミング距離は、2である。

【0109】

特定の実施形態において、試料のゲノムDNA断片に付着する複数のアダプタの試料タグ領域は、8ヌクレオチド長にあり、2のハミング距離によって互いに異なる240個のヌクレオチド配列のうちの1個を含む。

20

【0110】

ある一定の実施形態において、試料タグ領域は、個々のゲノムDNA断片を識別するように、また個々の試料、すなわち、ゲノムライブラリ供給源を識別するように機能する。たとえば、試料に付着する複数のアダプタの試料タグが240個の可能な配列のうちの1個を含むときに、各試料は、240個の可能なタグのうちの1個を含むと識別され、各試料は、2のハミング距離（一方のタグを他方のタグに変化させるために2塩基変化を必要とすることを意味する）によっていずれかの他の試料から分離する、240個のタグの1セットを受容する。これらの同一のタグを使用し、クローン多様性を列挙するので、それらは、配列タグとしても機能する、すなわち、ゲノムDNA断片を識別する。可能な配列タグの多様性をさらに増すために、UMI増倍管を加えることができる。たとえば、3塩基の64個の可能な組み合わせからなる3個のヌクレオチドを含むアダプタ領域にUMI増倍管を加えることが可能である。加えて、複数のアダプタは、1個を上回るアンカー配列を含むことが可能である。たとえば、複数のアダプタは、同時に使用される4個の異なるアンカー配列を含むことができる。また、エラーを少なくするために試料多重化解除中に、これらのアンカー配列を使用することができる。

30

【0111】

図4は、第一生成アダプタ（図4A及び図4B）と本発明のアダプタ（図4Cから図4E）との間の例示的な比較を示す。図4A及び図4Bは、40ntの長さであり、別々のPCR増幅配列、配列タグ、及び試料タグからなる第一生成アダプタの例を示す。ここで、試料は、固定配列（配列タグ）によって識別され、この固定配列は、試料からDNAライブラリを生成するために使用されるすべてのアダプタ上に存在する。個々のゲノム断片は、別々の、また別個の配列（配列タグ）によって識別される。図4Cから図4Eは、本発明からのアダプタの例示的な例を示す。示される例示的なアダプタは、47ヌクレオチド長にあり、配列タグは、試料タグと混合する。3塩基の64個の可能な組み合わせからなる、追加の3ntの配列、UMI増倍管がある。10ntのアンカー配列は、4個の異なる別個の配列のうちの1個である。

40

【0112】

したがって、例示的な例（図4Cから図4Eを参照する）において、単一の試料に関連

50

して使用される 1 セットのアダプタは、240 個の試料タグ配列を含み、これらの試料タグ配列は、4 セットの試料タグ配列に分割され、各セットは、60 個のタグ（各ヌクレオチド、A、C、T 及び G について 1 個）を含むことが可能である。したがって、60 個のタグの各セットは、4 個のアンカー配列のうちの 1 個に特異的である。合計で、1 試料あたり 240 個の可能な試料タグ配置のプールが可能である。具体的に、この事態において、これらの 240 個の試料タグ配列を 60 個の配列の 4 セットに分割し、この各セットは、特異的なアンカー領域を対象とする。したがって、試料 ID は、8 個のヌクレオチド試料タグからの配列情報だけでなく、関連したアンカー配列情報をも含む。加えて、リード内の配列の位置を固定するので、試料タグ及びアンカー配列は、下流の考察用に封入体フィルタを通すため、シーケンシングリード内に固定位置を有さなければならない。さらに、UMI 増倍管の封入体は、配列タグ多様性を 240 から $240 \times 64 = 15,360$ 個の可能な配列タグまで増加させる。

10

【0113】

本明細書に企図される 1 個以上のアダプタの付着を当業者に知られている方法によって実施することができる。特定の実施形態において、本明細書に企図される 1 個以上のアダプタを、平滑末端を含む末端修復 c f DNA に付着させる。ある一定の実施形態において、本明細書に企図される 1 個以上のアダプタは、用いられる付着方法に適切である相補的な末端を含む末端修復 c f DNA に付着する。ある一定の実施形態において、本明細書企図される 1 個以上のアダプタは、3' オーバーハングを含む末端修復 c f DNA に付着する。

20

【0114】

いくつかの実施形態において、ゲノム DNA 断片を複数のアダプタに付着させることは、末端修復 c f DNA または細胞 DNA 断片を、アンカー領域の少なくとも一部を含むオリゴヌクレオチドに付着させるステップを備える。いくつかの実施形態において、オリゴヌクレオチドは、全アンカー領域を含む。特定の実施形態において、オリゴヌクレオチドは、パートナー鎖によって二本鎖にされる 5' リン酸化付着鎖を含む DNA 二本鎖であり、そこでパートナー鎖は、化学修飾による付着からその 3' 末端で遮断され、そこで付着鎖は、ゲノム DNA 断片に付着する。ある一定の実施形態において、つぎにアンカー領域の少なくとも一部と付着する DNA 断片を、DNA オリゴヌクレオチドによってアニールし、アダプタ配列の全長を符号化する。特定の実施形態において、1 個以上のポリヌクレオチドキナーゼ、1 個以上の DNA リガーゼ、及び / または 1 個以上の DNA ポリメラーゼを、ゲノム DNA 断片、及び DNA オリゴヌクレオチドに加え、アダプタ配列の全長を符号化する。いくつかの実施形態において、ポリヌクレオチドキナーゼは、T4 ポリヌクレオチドキナーゼである。いくつかの実施形態において、DNA リガーゼは、Taq DNA リガーゼである。ある一定の実施形態において、DNA ポリメラーゼは、Taq ポリメラーゼである。特定の実施形態において、DNA ポリメラーゼは、Bst ポリメラーゼ全長である。

30

【0115】

図 6 は、複数のアダプタを修復された DNA 断片の 3' 末端に付着させる例示的な方法を示す。第一ステップにおいて、アンカー配列をゲノム断片の 3' 末端に付着させる。このステップにおいて、アンカー部分は、DNA 二本鎖であり、この DNA 二本鎖において、10 ヌクレオチドの 5' リン酸化「付着鎖」は、8 ヌクレオチド「パートナー鎖」によって二本鎖にされ、これらのヌクレオチドパートナー鎖は、化学修飾による付着からその 3' 末端で遮断される。アンカー二本鎖は、リン酸化 / 遮断末端上の平滑末端であるので、平滑末端ゲノム断片に付着することが可能である。つぎのステップにおいて、全アダプタ配列を符号化するオリゴヌクレオチドのプールを最初のアンカー配列にアニールする。T4 ポリヌクレオチドキナーゼ、Taq DNA リガーゼ、及び Bst ポリメラーゼの全長の混合作用は、上部の鎖について図示されるようにライゲーションを介してこのオリゴヌクレオチドを付着させ、最初のアンカー配列を DNA 重合によって下部の鎖上に伸長させ、アダプタ配列の全長を完成させる。同一のプロセスを使用して、アダプタ全長をゲノム断片の

40

50

5'末端に付着させることができる。

【0116】

2. DNAライブラリ増幅

特定の実施形態において、本明細書に企図される遺伝子解析方法は、ゲノムDNAライブラリ、たとえば、細胞DNAライブラリまたはcfDNAライブラリなどの増幅を有し、1個のDNAクローンライブラリ、若しくは複数のDNAクローンの1個のライブラリ、たとえば、1個のcfDNAクローンライブラリ、若しくは複数のcfDNAクローンの1個のライブラリ、または1個の細胞DNAクローンライブラリ、若しくは複数の細胞DNAクローンの1個のライブラリなどを生成する。DNAライブラリの各分子は、末端修復DNA断片の各末端に付着するアダプタを含み、各アダプタは、1つ以上の増幅領域を含む。いくつかの実施形態において、異なるアダプタを末端修復cfDNAの異なる末端に付着させる。特定の実施形態において、異なるアダプタを末端修復細胞DNAの異なる末端に付着させる。

10

【0117】

いくつかの実施形態において、同じアダプタをDNA断片の両方の末端に付着させる。末端修復DNAの両方の末端への同一アダプタの付着により、シングルプライマー配列によるPCR増幅が可能である。特定の実施形態において、アダプタが付着したcfDNAライブラリの一部を、シングルプライマー配列駆動増幅による標準的なPCR手法を用いて増幅することになる。一実施形態では、単一プライマー配列は、約25ヌクレオチドであり、任意選択により、標準的なイオン強度条件下で55の推定T_mを有する。

20

【0118】

特定の実施形態において、ピコグラム単位の初期ゲノムDNAライブラリ、たとえば、細胞DNAライブラリまたはcfDNAライブラリを、マイクログラム単位のDNAクローンに増幅、つまり10,000倍に増幅する。増幅産物の量は、当該技術分野で公知の方法(例えば、Qubit 2.0またはNanodrop装置での定量化)を用いて測定されることができる。

【0119】

3. ゲノム当量数の決定

様々な実施形態において、ゲノムDNAの遺伝子解析方法は、DNAクローンライブラリにおけるゲノム当量数を決定することを備える。本明細書で使用する「ゲノム当量」という用語は、各ライブラリにおけるゲノムコピー数を指す。本明細書で企図する組成物及び方法が遭遇する重要な課題は、遺伝子配列における希少な遺伝子突然変異または相違を検出し解析するのに十分なアッセイ感度を達成することである。試料ごとに基づいてアッセイ感度の値を決定するために、シーケンシングライブラリ内に存在するゲノム当量数を測定することにより、各試料内に存在する異なる別々の配列の数を測定する。感度を確立するため、ゲノム当量数を試料ライブラリごとに測定しなければならない。

30

【0120】

ゲノム当量数は、qPCRアッセイにより、またはシーケンシング実施後のバイオフィォマティクスに基づくカウントにより決定されることができる。臨床試料のプロセスフローでは、ゲノム当量のqPCR測定は、DNAライブラリ、たとえば、cfDNAライブラリまたはゲノムDNAライブラリなどにQCステップとして使用される。配列解析の前にアッセイ感度の期待値を確立させ、試料に対応するDNAクローンライブラリがゲノム当量の要求深度を欠く場合は、その試料を解析から除外することが可能になる。最終的には、バイオフィォマティクスに基づくゲノム当量カウントも使用してゲノム当量を特定する。そのため、それぞれの所与のDNAクローンライブラリに対するアッセイ感度及び偽陰性推定値も特定される。

40

【0121】

実験的qPCRアッセイ及び統計的カウントアッセイは十分に相互関係を有するべきである。シーケンシングによってDNAクローンライブラリにおける配列深度が明らかにならない場合は、DNAクローンライブラリの再処理及び/または追加的なシーケンシング

50

が必要となり得る。

【0122】

一実施形態では、細胞DNAまたはcfDNAクローンライブラリにおけるゲノム当量を定量的PCR(qPCR)アッセイを用いて決定する。特定の実施形態では、既知濃度の標準的なライブラリを使用して標準曲線を構築し、この得られた標準曲線にqPCRアッセイを適合させ、適合度からゲノム当量の値を導き出す。本発明者らは、ゲノム中の共通配列、たとえば、反復配列などへ特異的にハイブリダイズする一方のプライマーと、アダプタ中のプライマー結着部位へ結着する他方のプライマーとを含むqPCR「反復に基づく」アッセイがアダプタの特異的なプライマー(cfDNAクローンの両方の末端上に存在する)だけを使用する方法と比較してゲノム当量中で8倍増加を測定したことを発見した。反復に基づくアッセイにより測定されるゲノム当量数は、より一貫したライブラリ間パフォーマンスをもたらし、またゲノム当量のqPCR推定値とシーケンシング実行時にバイオフィォマティクスのカウントしたタグ当量との間のアライメントをより良好なものにする。

10

【0123】

本明細書で企図する反復に基づくゲノム当量アッセイでの使用に好適な、反復の例示的な例としては、以下に限定するものではないが、短散在型核内反復配列(SINE)(例えば、Alu反復)、長散在型反復配列(LINE)(例えば、LINE1、LINE2、LINE3)、マイクロサテライト反復配列(例えば、短いタンDEM反復(STR)、単純反復配列(SSR))、及び哺乳類散在型反復(MIR)が挙げられる。

20

【0124】

一実施形態では、反復はAlu反復である。

【0125】

4. 定量的遺伝子解析

様々な実施形態では、ゲノムDNA、たとえば、ゲノム細胞またはcfDNAなどの遺伝子解析方法は、DNAライブラリクローンにおける1つ以上の標的遺伝子座の定量的遺伝子解析を含む。定量的遺伝子解析は、以下のステップの1つ以上、または全てを含む：標的遺伝子座を含むDNAクローンのキャプチャ；キャプチャした標的遺伝子座の増幅；キャプチャし増幅した標的遺伝子座のシーケンシング；及び得られたシークエンスリードのバイオフィォマティクス解析。本明細書に使用されるように、用語「DNAライブラリクローン」は、DNAライブラリ断片を指し、そこでアダプタ及びゲノムDNA断片の組み合わせは、ユニークDNA配列(たとえば、その別のDNAライブラリクローンと区別されることが可能であるDNA配列)をもたらす。

30

【0126】

(a) 標的遺伝子座のキャプチャ

本発明は、その一部として、より大きなプローブの効率及び信頼性を保持するように設計され、ただしより小さなDNA断片を含むゲノムDNAライブラリ、たとえば、cfDNAクローンライブラリなどにおける情報価値のない配列の生成を最小限にとどめる、キャプチャプローブモジュールを企図する。本明細書に使用されるような「キャプチャプローブ」または「キャプチャプローブモジュール」は、互換的に使用され、キャプチャプローブ配列及びテール配列を含むポリヌクレオチドを指す。特定の実施形態では、キャプチャプローブモジュール配列またはその一部は、1つ以上のシーケンシングプライマーのためのプライマー結合部位として働く。

40

【0127】

特定の実施形態では、キャプチャプローブモジュールはキャプチャプローブを含む。本明細書で使用する「キャプチャプローブ」とは、特異的なDNA標的領域とハイブリダイズすることができる領域を指す。いくつかの実施形態において、キャプチャプローブは、細胞DNAから構築されるゲノムDNAライブラリによって使用される。特定の実施形態において、キャプチャプローブは、cfDNAから構築されるゲノムDNAライブラリによって使用される。cfDNAの平均サイズが約150から約170bpであり、また

50

高度に断片化されるため、ある一定の実施形態は、対象となるDNA標的領域を調べるための高密度及び比較的短いキャプチャプローブの使用を含む本明細書で企図される組成物及び方法を対象とする。いくつかの実施形態において、キャプチャプローブは、すべての染色体セグメントにわたり均一な密度で分布する、DNA標的領域にハイブリダイズすることが可能である。1セットのこれらのようなキャプチャプローブは、本明細書において「染色体安定性プローブ」と称される。染色体コピー数（たとえば、染色体倍数性）のゲノムワイド測定を提供するために、これらの染色体安定性プローブを使用して、コピー数多型をゲノムワイドスケール上で調べる。

【0128】

高密度キャプチャプローブの使用に関する特定の懸念の1つは、概して、キャプチャプローブが特異的な「配列規則」を用いて設計されていることである。たとえば、冗長配列の領域、または極端に塩基組成物に偏る領域は、キャプチャプローブを設計する際に一般的に排除される。しかし、発明者らは、キャプチャプローブの設計規則が柔軟性に欠けることで、プローブの性能に対する実質的影響が生じるわけではないことを発見した。対照的に、位置的な制約により厳密に選ばれたキャプチャプローブは、標的上の配列情報を提供し、標的外のマッピング不可能なリードキャプチャはごくわずかしき示さず、少数の例外を除いて均一で有用で標的上にリードをもたらす。その上、近接したプローブ間隔での高冗長性は、時として低性能となるキャプチャプローブを補って余りあるものである。

【0129】

特定の実施形態では、1つの標的領域を複数のキャプチャプローブが標的とし、任意の2つ以上のキャプチャプローブは、当該標的領域に対し、互いの10ヌクレオチド以内、互いの15ヌクレオチド以内、互いの20ヌクレオチド以内、互いの25ヌクレオチド以内、互いの30ヌクレオチド以内、互いの35ヌクレオチド以内、互いの40ヌクレオチド以内、互いの45ヌクレオチド以内、若しくは互いの50ヌクレオチド以内、またはそれより多いヌクレオチド以内、さらにこれらの間にある全てのヌクレオチド長で結合するように設計される。

【0130】

一実施形態では、キャプチャプローブは、約25ヌクレオチド、約26ヌクレオチド、約27ヌクレオチド、約28ヌクレオチド、約29ヌクレオチド、約30ヌクレオチド、約31ヌクレオチド、約32ヌクレオチド、約33ヌクレオチド、約34ヌクレオチド、約35ヌクレオチド、約36ヌクレオチド、約37ヌクレオチド、約38ヌクレオチド、約39ヌクレオチド、約40ヌクレオチド、約41ヌクレオチド、約42ヌクレオチド、約43ヌクレオチド、約44ヌクレオチド、または約45ヌクレオチドである。

【0131】

一実施形態では、キャプチャプローブは、約100ヌクレオチド、約200ヌクレオチド、約300ヌクレオチド、約400ヌクレオチド、または約100ヌクレオチドである。別の実施形態では、キャプチャプローブは、約100ヌクレオチド～約500ヌクレオチド、約200ヌクレオチド～約500ヌクレオチド、約300ヌクレオチド～約500ヌクレオチド、若しくは約400ヌクレオチド～約500ヌクレオチド、またはこれらの間にある任意の範囲である。

【0132】

特定の実施形態では、キャプチャプローブは60ヌクレオチドである。別の実施形態では、キャプチャプローブは、実質的に60ヌクレオチドより小さいが、同じDNA標的領域を標的とする60ヌクレオチドのキャプチャプローブと同等に、同様に、またはより良好にハイブリダイズする。あるいくつかの実施形態では、キャプチャプローブは40ヌクレオチドである。

【0133】

あるいくつかの実施形態では、キャプチャプローブモジュールはテール配列を含む。本明細書で使用する「テール配列」という用語は、キャプチャプローブモジュールの5'末端におけるポリヌクレオチドを指し、これは特定の実施形態においてプライマー結合部位と

10

20

30

40

50

して働き得る。特定の実施形態では、シーケンシングプライマーは、テール領域内のプライマー結合部位に結合する。

【0134】

特定の実施形態では、テール配列は、約5～約100ヌクレオチド、約10～約100ヌクレオチド、約5～約75ヌクレオチド、約5～約50ヌクレオチド、約5～約25ヌクレオチド、または約5～約20ヌクレオチドである。あるいくつかの実施形態では、第3の領域は、約10～約50ヌクレオチド、約15～約40ヌクレオチド、約20～約30ヌクレオチド、若しくは約20ヌクレオチド、または間にある任意の数のヌクレオチドである。

【0135】

特定の実施形態では、テール配列は、約30ヌクレオチド、約31ヌクレオチド、約32ヌクレオチド、約33ヌクレオチド、約34ヌクレオチド、約35ヌクレオチド、約36ヌクレオチド、約37ヌクレオチド、約38ヌクレオチド、約39ヌクレオチド、または約40ヌクレオチドである。

【0136】

様々な実施形態では、キャプチャプローブモジュールは、キャプチャプローブとハイブリダイズするタグ付き及び/または増幅したゲノムDNAライブラリ（たとえば、細胞またはcfDNAライブラリ）における1つ以上のキャプチャした断片の単離及び/または精製を可能にする、特定のメンバーの結合対を含む。特定の実施形態では、キャプチャプローブモジュールは、ピオチンまたは他の好適なハプテン（例えば、ジニトロフェノール、ジゴキシゲニン）に共役する。

【0137】

様々な実施形態では、キャプチャプローブモジュールを、タグ付きの、任意選択により増幅したDNAライブラリとハイブリダイズして、複合体を形成する。一部の実施形態では、多機能のキャプチャプローブモジュールは、DNAライブラリにおける特異的なゲノム標的領域と実質的にハイブリダイズする。

【0138】

ハイブリダイゼーションまたはハイブリダイズ条件には、2つのヌクレオチド配列が安定した複合体を形成する（例えば、タグ付きDNAライブラリ及びキャプチャプローブモジュールが安定したタグ付きDNAライブラリ-キャプチャプローブモジュール複合体を形成する）任意の反応条件が含まれ得る。これらのような反応条件は、当技術分野において周知であり、当業者であれば、これらのような条件は、例えばアニーリング温度を低くし、長さの短いキャプチャプローブを用いるというように、必要に応じて、かつ本発明の範囲内で改変することができることを理解することになる。キャプチャプローブ複合体の第2の領域が、タグ付きDNAライブラリの領域に対し、100%、99%、98%、97%、96%、95%、94%、93%、92%、91%、90%、89%、88%、85%、80%、75%、または70%の配列一致度、相同性、または相補性を示す場合、実質的なハイブリダイゼーションが起こり得る。

【0139】

特定の実施形態では、キャプチャプローブは約40ヌクレオチドであり、約44～約47の最適アニーリング温度を有する。

【0140】

あるいくつかの実施形態では、本明細書で企図する方法は、タグ付きcfDNAライブラリ-キャプチャプローブモジュール複合体を単離することを含む。特定の実施形態では、DNA複合体を単離する方法は当業者に周知であり、当業者が適切とみなす任意の方法を本発明の方法で用いることができる（Ausubel et al., Current Protocols in Molecular Biology, 2007-2012）。特定の実施形態では、ピオチン-ストレプトアビジン単離技術を用いて複合体を単離する。

【0141】

10

20

30

40

50

特定の実施形態では、単離したタグ付きDNAライブラリ断片 - キャプチャプローブモジュール複合体からの1本鎖の3'末端除去を企図している。あるいくつかの実施形態では、当該方法は、1本鎖の3'末端を除去するための、単離したタグ付きDNAライブラリ - 多機能キャプチャプローブモジュール複合体の3' - 5'エキソヌクレアーゼ酵素処理を含む。

【0142】

あるいくつかの他の実施形態では、当該方法は、単離したタグ付きDNAライブラリ断片を鋳型として利用して多機能キャプチャプローブの5' - 3' DNAポリメラーゼ伸長を実施することを含む。

【0143】

ある一定の他の実施形態において、これらの方法は、5' FLAPエンドヌクレアーゼ、DNA重合、及びDNAリガーゼによるニック閉鎖の協調作用を通じて、ハイブリッドキャプチャプローブ - 単離したタグ付きDNA標的分子、たとえば、タグ付きcfDNA標的分子またはタグ付き細胞DNA標的分子などを創出することを含む。

【0144】

単離したタグ付きDNAライブラリ - 多機能キャプチャプローブモジュール複合体の3' - 5'エキソヌクレアーゼ酵素処理用に様々な酵素を用いることができる。3' - 5'エキソヌクレアーゼ酵素活性を示し、特定の実施形態で使用するすることができる好適な酵素の例示的例は、限定するものではないが、T4またはエキソヌクレアーゼI、III、Vを含む (Shevellev IV, Hubbscher U., Nat Rev Mol Cell Biol. 3 (5): 364 - 76 (2002) も参照)。特定の実施形態において、3' - 5'エキソヌクレアーゼ活性を備える酵素は、T4ポリメラーゼである。特定の実施形態では、3' - 5'エキソヌクレアーゼ酵素活性を示し、プライマー鋳型伸長が可能な酵素を使用することができ、これには、例えばT4またはエキソヌクレアーゼI、III、V (同上) が含まれる。

【0145】

一部の実施形態では、本明細書で企図する方法は、上記または本明細書の他の箇所で論じられている3' - 5'エキソヌクレアーゼ酵素処理をした複合体に、シーケンシング及び/またはPCRを実施することを含む。特定の実施形態では、ハイブリッド核酸分子を生成するために、キャプチャプローブ分子のテール部分をコピーする。一実施形態では、生成したハイブリッド核酸分子は、キャプチャプローブモジュールへハイブリダイズ可能な標的領域と、キャプチャプローブモジュールテール配列の相補配列とを含む。

【0146】

特定の実施形態では、遺伝子解析は、a) 1個以上のキャプチャプローブモジュールを複数のゲノムDNAライブラリクローン中の1個以上の標的遺伝子座にハイブリダイズして、1個以上のキャプチャプローブモジュール - DNAライブラリクローン複合体を形成すること; b) a) からの1個以上のキャプチャプローブモジュール - DNAライブラリクローン複合体を単離すること; c) ステップb) からの1個以上の単離したキャプチャプローブモジュール - DNAライブラリクローン複合体を酵素処理すること; d) c) からの酵素処理した複合体にPCRを実施することであって、増幅したハイブリッド核酸分子を生成するためにキャプチャプローブ分子のテール部分をコピーし、増幅したハイブリッド核酸分子がキャプチャプローブにハイブリダイズ可能な標的ゲノム遺伝子座位内の標的配列と、キャプチャプローブモジュールテール配列の相補配列とを含む、実施すること; 及びe) d) からの増幅したハイブリッド核酸分子上に定量的遺伝子解析を実施すること、を含む。

【0147】

特定の実施形態では、特異的な標的遺伝子座のコピー数を決定する方法が企図され、この方法は、a) 1個以上のキャプチャプローブモジュールを複数のDNAライブラリクローンにおける1個以上の標的遺伝子座にハイブリダイズして、1個以上のキャプチャプローブモジュール - DNAライブラリクローン複合体を形成すること; b) a) から1個以

10

20

30

40

50

上のキャプチャプローブモジュール - DNAライブラリクローン複合体を単離すること；
 c) ステップ b) からの 1 個以上の単離したキャプチャプローブモジュール - DNAライ
 ブラリクローン複合体を酵素処理すること； d) c) からの酵素処理した複合体上に P C
 R を実施することであって、増幅したハイブリッド核酸分子を生成するためにキャプチャ
 プローブ分子のテール部分をコピーし、増幅したハイブリッド核酸分子がキャプチャプ
 ローブにハイブリダイズ可能な標的遺伝子座内の標的配列と、キャプチャプローブモジ
 ュールテール配列の相補配列とを含む、実施すること； e) d) における増幅したハイブリ
 ッド核酸分子の P C R 増幅を実施すること；及び f) e) における P C R 反応を定量化す
 ることであって、この定量化によって、特異的な標的領域のコピー数を決定することがで
 ける、定量化することを備える。

10

【 0 1 4 8 】

一実施形態では、ステップ c) の酵素処理は、3' - 5' エキソヌクレアーゼ活性を有す
 る酵素を用いて、b) からの 1 個以上のキャプチャプローブモジュール - DNAライブラ
 リクローン複合体上に 3' - 5' エキソヌクレアーゼ酵素処理を実施して 1 本鎖の 3' 末端を
 除去すること；5' F L A P エンドヌクレアーゼ、DNA 重合、及び DNA リガーゼによる
 ニック閉鎖の協調作用を通じて、1 個以上のハイブリッドキャプチャプローブ - c f D N
 A ライブラリクローン分子を創出すること；または、複合体内の単離した DNA クローン
 を鋳型として用いてキャプチャプローブの 5' - 3' DNA ポリメラーゼ伸長を実施するこ
 と、を含む。

【 0 1 4 9 】

20

一実施形態では、ステップ c) の酵素処理は、複合体内の単離した DNA クローンを鋳
 型として用いてキャプチャプローブの 5' - 3' DNA ポリメラーゼ伸長を実施することを
 含む。

【 0 1 5 0 】

特定の実施形態では、当業者に周知である任意の標準的な P C R 反応条件を用いて P C
 R を実施することができる。ある一定の実施形態において、e) における P C R 反応は、
 2 個の P C R プライマーを用いる。一実施形態では、e) における P C R 反応は、標的遺
 伝子座内の反復にハイブリダイズする第 1 の P C R プライマーを用いる。特定の実施形態
 では、e) における P C R 反応は、標的遺伝子座 / テール接合部におけるハイブリッド核
 酸分子にハイブリダイズする第 2 の P C R プライマーを用いる。あるいくつかの実施形態
 では、e) における P C R 反応は、標的遺伝子座にハイブリダイズする第 1 の P C R プ
 ライマーを用い、第 2 の P C R プライマーは、標的遺伝子座 / テール接合部における増幅
 したハイブリッド核酸分子にハイブリダイズする。特定の実施形態では、プライマーの少
 なくとも 1 個以上のヌクレオチドは標的遺伝子座にハイブリダイズし、プライマーの少
 なくとも 1 個以上のヌクレオチドはテール配列にハイブリダイズするように、第 2 のプ
 ライマーは、標的遺伝子座 / テール接合部にハイブリダイズする。

30

【 0 1 5 1 】

あるいくつかの実施形態では、ステップ e) から取得する増幅したハイブリッド核酸分
 子を配列決定し、配列を水平方向にアライメントを取る。すなわち、参照配列に対してで
 はなく、互いに対しアライメントを取る。特定の実施形態では、a) から e) のステップ
 を 1 個以上のキャプチャプローブモジュールで 1 回以上反復する。キャプチャプローブモ
 ジュールは、同じであっても異なってもよく、標的遺伝子座のいずれかの c f D N A
 鎖を標的にするように設計され得る。一部の実施形態では、キャプチャプローブが異なる
 場合、これらのキャプチャプローブは、タグ付き c f D N A クローンライブラリにおける
 標的遺伝子座内のオーバーラップまたは隣接する標的配列にハイブリダイズする。一実施
 形態では、高密度キャプチャプローブ計画を使用し、この計画で複数のキャプチャプ
 ローブは、標的遺伝子座にハイブリダイズし、複数のキャプチャプローブのそれぞれは、
 タグ付き DNA クローンライブラリ中の標的遺伝子座にハイブリダイズする任意の他のキャ
 プチャプローブから約 5、10、15、20、25、30、35、40、45、50、100、200 b p 以上（間にある全ての距離を含む）内の標的遺伝子座にハイブリダイズす

40

50

る。

【 0 1 5 2 】

一部の実施形態において、この方法は、標的遺伝子座ごとに2個のキャプチャプローブモジュールを用いて実施されることができ、この一方が標的領域の上流にある「ワトソン」鎖（非コード鎖または鋳型鎖）にハイブリダイズし、一方が標的領域の下流にある「クリック」鎖（コード鎖または非鋳型鎖）にハイブリダイズする。

【 0 1 5 3 】

特定の実施形態では、本明細書で企図する方法は、標的遺伝子座ごとに、任意の組合せでワトソン鎖またはクリック鎖にハイブリダイズする任意の数のキャプチャプローブモジュール、例えば、2、3、4、5、6、7、8、9、または10、またはそれ以上のキャプチャプローブモジュールを用いて、さらに複数回実施することができる。一部の実施形態では、取得した配列は、複数の相違のいずれかを識別するために相互にアライメントを取ることができる。

【 0 1 5 4 】

あるいくつかの実施形態では、複数の標的遺伝子座を、1つ以上のキャプチャプローブモジュールを用いた1回の反応で、例えば、100、200、300、400、500、600、700、800、900、1000、1500、2000、2500、3000、3500、4000、4500、5000、10000、50000、100000、500000またはそれ以上調べる。

【 0 1 5 5 】

（ b ）シーケンシング

特定の実施形態では、定量的遺伝子解析は、本明細書の上記の他の箇所で論じているように複数のハイブリッド核酸分子を配列決定して、複数のユニークシーケンシングリードを取得するのに十分なシーケンシング深度を生成することを含む。用語「ユニークリード」または「ユニークゲノム配列」（UGS）は、本明細書に互換的に使用され、個々の冗長リードを合わせて「ファミリー」にグループ化することによって識別される。これらの冗長リードは、個々のUMIEを共有する（たとえば、同一リードコード、及び同一DNA配列の開始位置をゲノム配列内で共有する）配列リードであり、単一の付着事象に由来するため、増幅由来の互いの「同胞」である。冗長リードのファミリーを代表する単一のコンセンサスをユニークリードまたはUGSとして進める。それぞれのユニークリードまたはUGSは、ユニーク付着事象とみなされる。特定のキャプチャプローブに対応するユニークリードの合計は、その特定のキャプチャプローブについての「生のゲノム深度」（RGD）と称される。各キャプチャプローブは、ファミリーにグループ化することにより全リードから計算的に抽出される1セットのユニークリードをもたらす。次に所与の試料についてのユニークリード（たとえば、試料についての生のゲノム深度）を、プローブごとに観察される全てのユニークリードの平均値として計算する。ユニークリードは、それぞれがユニークゲノムDNAクローンから導き出されなければならないため、重要である。それぞれのユニークリードは、一倍体当量のゲノムDNAのインプット及び解析を表す。ユニークリードの和は、解析した一倍体ゲノムの和である。そして解析したゲノムの数は、シーケンシングアッセイの感度を定義する。非限定的な例として、ユニークリードの平均カウントが100ゲノム当量である場合、この特定のアッセイは100のうちの1つの変異リード、すなわち1%を検出することができる感度を有する。これを下回るいかなる観察も正当ではない。

【 0 1 5 6 】

明らかなコピー数の変化（たとえば、ノイズの多いプローブの例）がある場合は、試料平均値の計算に使用するデータセットから除外する。本明細書において、「ノイズの多いプローブ」は、大規模な1セットの同一の試料中のユニークリードの非常にばらつきのある数（たとえば、12から16個の試料複製物中のユニークリードの非常にばらつきのある数）をキャプチャするプローブを指す。いくつかの実施形態において、ノイズの多いプローブと関連するユニークリード数は、50%以上までの試料についてのユニークリード

10

20

30

40

50

の平均数と比較して増加する。いくつかの実施形態において、ノイズの多いプローブと関連するユニークリード数は、50%以上までの試料についてのユニークリードの平均数と比較して減少する。いくつかの実施形態において、特定の解析において使用される、約2%から約4%のプローブをノイズの多いプローブとして識別し、計算から除外し、所与の試料についてユニークリードの平均数を決定する。

【0157】

いくつかの実施形態において、シーケンシングリードを「標的上リード」または「標的外リード」のいずれか一方として識別する。標的上リードは、ゲノムライブラリを作製するために使用されるキャプチャプローブの近くにマッピングするゲノムDNA配列を含む。いくつかの実施形態において、各ゲノム配列を特異的なキャプチャプローブへ物理的に結合し、ゲノムセグメント及びキャプチャプローブの両方の配列を統合情報として決定し、開始座標が対応するキャプチャプローブの3'末端の400bp内に、またより一般的に200bp内に、マッピングされるいずれかのゲノム配列として標的上リードを定義する。キャプチャプローブに関して、500の塩基対の位置に基準ゲノムにアライメントを取るゲノム配列を含む（及び完全に異なる染色体にマッピングすることがさらに多い）ものとして標的外リードを定義する。

【0158】

特定の実施形態では、定量的遺伝子解析は複数の試料に由来するハイブリッド核酸分子の多重シーケンシングを含む。

【0159】

様々な実施形態では、定量的遺伝子解析は、それぞれが第1のDNA配列及び第2のDNA配列を含む、1つ以上または複数のタグ付きDNAライブラリのクローンを取得することであって、第1のDNA配列が標的とする遺伝子座内に配列を含み、第2のDNA配列がキャプチャプローブ配列を含む、取得すること；1つ以上のクローンに対し、対合した末端のシーケンシング反応を実施し1つ以上のシーケンシングリードを取得すること、または、1つ以上のクローンに対し、約100、200、300、400、500またはそれ以上のヌクレオチドを上回る単一の長いシーケンシングリードを取得するシーケンシング反応を実施することであって、リードが第1のDNA配列及び第2のDNA配列の両方を識別するのに十分である、実施すること；ならびに、1つ以上のクローンのシーケンシングリードをシーケンシングリードのプローブ配列に従って順序付けまたはクラスター化すること、を含む。

【0160】

(c) バイオインフォマティクス解析

様々な実施形態では、定量的遺伝子解析は、シーケンシングリードのバイオインフォマティクス解析を含む。バイオインフォマティクス解析により、シーケンシングのための組成物または方法が不在の場合に実施されるいかなる純粋な精神的解析も除外される。あるいくつかの実施形態では、バイオインフォマティクス解析には、以下に限定するものではないが、配列アライメント、ゲノム当量解析、一塩基多様体(SNV)解析、遺伝子コピー数多型(CNV)解析、染色体コピー数の測定、及び遺伝子損傷の検出が含まれる。特定の実施形態では、バイオインフォマティクス解析は、cfDNAクローンライブラリにおいて解析するゲノム当量数の定量化、標的遺伝子座の遺伝子状態の検出、標的遺伝子座における遺伝子損傷の検出、及び標的遺伝子座内のコピー数変動の測定に有用である。

【0161】

配列アライメントは、配列リードと1つ以上のヒト参照DNA配列との間で実施されることができる。特定の実施形態では、シーケンシングアライメントは、標的遺伝子座における遺伝子損傷の検出に使用することができ、遺伝子損傷の検出には、以下に限定するものではないが、ヌクレオチドのトランジション若しくはトランスバージョン、ヌクレオチドの挿入若しくは欠失、ゲノム再編成、コピー数の変化、または遺伝子融合の検出が含まれる。原因または予後の指標である遺伝子損傷の検出は、特定の遺伝子の状態または疾患の診断、予後判定、処置、及び/または監視に有用であり得る。

10

20

30

40

50

【 0 1 6 2 】

本明細書には、参照配列に対するアライメントの必要なしに実施されることができる配列アライメント解析の方法も企図されており、本明細書ではこの方法を水平配列解析と呼ぶ。このような解析は、本明細書で企図する方法または他の任意の方法によって生成した任意の配列に対し実施されることができる。特定の実施形態では、この配列解析は、本明細書で企図する方法によって取得したリードに対し、配列アライメントを実施することを含む。

【 0 1 6 3 】

一実施形態では、c f DNA クローンライブラリにおけるゲノム当量は、シーケンシングの実施後にバイオインフォマティクスに基づくカウントを用いて決定される。各シーケンシングリードは、特定のキャプチャプローブと関連し、各キャプチャプローブに割り当てられたリードの集合体は、群にパースされる。群内において、個々のリードのセットは、ゲノム配列内で同じリードコード、及び同じDNA配列の開始位置を共有する。これらの個々のリードは、「ファミリー」にグループ化され、このファミリーを代表する単一のコンセンサスは、「ユニークリード」として進められる。ファミリーを構成した個々のリードのすべては、単一の付着事象に由来するため、これらは、互いに増幅由来の「同胞」である。それぞれのユニークリードは、ユニーク付着事象とみなされ、ユニークリードの和は、解析されるゲノム当量数に相当するとみなされる。

10

【 0 1 6 4 】

ユニーククローンの数が、可能な配列組合せの総数に近づくにつれ、確率は、独立した事象により同じコード及び開始部位が創出されること、またこれらの独立した事象が不適切な形で1つのファミリーにグループ化されることを示す。最終結果は、解析ゲノム当量より少ない見積もりとなり、希少な変異リードは、同じ識別子を有する野生型リードと重複するという理由からシーケンシングエラーとして廃棄される場合がある。

20

【 0 1 6 5 】

特定の実施形態では、c f DNA クローンライブラリの正確な解析を提供するため、解析ゲノム当量数は、可能なユニーククローン数の約 $1/10$ 、約 $1/12$ 、約 $1/14$ 、約 $1/16$ 、約 $1/18$ 、約 $1/20$ 、約 $1/25$ またはそれ以下である。上記で概説されている手順が例示的なものに過ぎず、非限定的であることを理解されたい。

【 0 1 6 6 】

一部の実施形態では、解析するゲノム当量数を増加させる必要があり得る。ゲノム当量の深度を拡大するため、少なくとも2つの解決策が企図されている。第1の解決策は、試料ごとに2つ以上のアダプタセットを使用することである。アダプタを組み合わせることにより、可能なクローンの総数を乗法的に拡大し、そのためゲノムインプットの無理のない限界を拡大することが可能である。第2の解決策は、リードコードを1つ、2つ、3つ、4つ、若しくは5つ、またはそれ以上の塩基の分拡大することである。他の全てのリードコードから少なくとも2塩基の分異なる可能なリードコードの数は、 $4^{(n-1)}$ のスケールであり、式中、 n はリードコード内の塩基の数である。したがって、非限定的な例において、リードコードが5ヌクレオチドであれば $4^{(5-1)} = 256$ であるから、追加的な塩基を含めることで、利用可能なレパートリーは追加塩基1つにつき4倍拡大する。

30

40

【 0 1 6 7 】

一実施形態では、定量的遺伝子解析は、希少な一塩基多様体 (SNV) を識別するための、シーケンシングリードのバイオインフォマティクス解析を含む。

【 0 1 6 8 】

次世代シーケンシングは、大まかに 0.02 から 0.02% の固有のエラー率を有し、これは $1/200$ から $1/500$ のいずれかの塩基呼び出しが不正確であることを意味する。これより低い頻度 (例えば、 $1/1000$ 配列当たり1の頻度) で発生する変異体及び他の突然変異を検出するため、分子アノテーション計画を発動する必要がある。非限定的な例として、標的配列キャプチャ技術を用いて 5000 個のユニーク分子を解析すれば、 $> 50,000$ リードの十分なシーケンシング深度で、 5000 ユニークリードの集合 (各

50

ユニークリードは、いずれもが同じリードコードを所持しているリードの「ファミリー」に属する)が生成されることになる。ファミリー内で発生するSNVは、希少変異体としての候補である。この同じ変異体が2つ以上のファミリーで観察される場合、開始試料内に存在する希少変異体としての非常に有力な候補となる。これに対し、複数ファミリー内で散発的に生じる変異体は、シーケンシングエラーである可能性があり、1つ及び1つのみのファミリー内で生じる変異体は、希少であるかまたは生体外で生じる塩基変更(例えば、DNA塩基の酸化またはPCRにより導入されたエラー)の結果である。

【0169】

一実施形態では、SNVを検出する方法は、アッセイにおける所望の標的感度の10倍以上のゲノムインプット(ゲノムまたはゲノム当量)を導入することを含む。非限定的な一例では、所望の感度が2%(100のうち2)である場合、実験的標的は2000ゲノムのインプットである。

10

【0170】

特定の実施形態では、シーケンシングデータのバイオインフォマティクス解析は、遺伝子の様相、状態または疾患、遺伝子モザイク、胎児検査、父子鑑定、薬物治療に対する応答の予測、医学的状态の診断または監視、マイクロバイームプロファイリング、病原体スクリーニング、及び臓器移植の監視に関連するSNVの検出または識別に使用される。

【0171】

様々な実施形態では、コピー数決定解析の方法であって、それぞれが第1のDNA配列及び第2のDNA配列を含む、1つ以上または複数のクローンを取得することを含み、第1のDNA配列が標的とする遺伝子座内に配列を含み、第2のDNA配列がキャプチャプローブ配列を含む、方法が提供される。関係する実施形態では、1つ以上のクローンに対し、対合した末端のシーケンシング反応を実施し、1つ以上のシーケンシングリードを取得する。別の実施形態では、1個以上のクローン上でシーケンシング反応を実施し、約100個を上回るヌクレオチドの単一の長いシーケンシングリードを取得し、このリードは、第一DNA配列及び第二DNA配列の両方を識別するのに十分である。1つ以上のクローンのシーケンシングリードは、シーケンシングリードのプローブ配列に従って順序付けまたはクラスター化することができる。

20

【0172】

コピー数解析には、以下に限定するものではないが、所与のゲノムDNA試料で生じる特定の遺伝子または突然変異のコピー数を調べる解析が含まれ、またさらに、所与の遺伝子のコピー数、または所与の試料中の配列相違数の定量的決定も含まれ得る。特定の実施形態では、コピー数解析は、遺伝子の様相、状態または疾患、胎児検査、遺伝子モザイク、父子鑑定、薬物治療に対する応答の予測、医学的状态の診断または監視、マイクロバイームプロファイリング、病原体スクリーニング、及び臓器移植の監視に関連する遺伝子増幅の検出または識別に使用される。

30

【0173】

いくつかの実施形態において、コピー数解析を使用して、染色体不安定性を測定する。これらのような実施形態において、染色体安定性プローブを含むキャプチャプローブセットを使用して、コピー数多型をすべての染色体セットにわたり均一な密度で決定する。コピー数解析を各染色体安定性プローブについて実施し、その後、染色体安定性プローブをそれらの染色体標的に従い順序付けする。これは、ゲノムにわたるコピー数減少または増加の視覚化を可能にし、染色体安定性の測定として機能することが可能である。

40

【0174】

特定の実施形態では、シーケンシングデータのバイオインフォマティクス解析は、標的遺伝子座における1個以上の配列または遺伝子損傷の検出または識別に使用されることができ、これらの検出または識別には、以下に限定するものではないが、ヌクレオチドのトランジション若しくはトランスバージョン、ヌクレオチドの挿入若しくは欠失、ゲノム再編成、コピー数の変化、または遺伝子融合の検出が含まれる。原因または予後の指標である遺伝子損傷の検出は、特定の遺伝子の状態または疾患の診断、予後判定、処置、及びノ

50

または監視に有用であり得る。一実施形態では、遺伝子損傷は、遺伝子の様相、状態または疾患、胎児検査、遺伝子モザイク、父子鑑定、薬物治療に対する応答の予測、医学的状态の診断または監視、マイクロバイームプロファイリング、病原体スクリーニング、及び臓器移植の監視に関連する。

【0175】

D．定量的CNLアッセイの臨床応用

様々な実施形態では、本発明は、関心領域における、変異性変化、SNP、転座、逆位、欠失、コピー数における変化、または他の遺伝的変異を検出することによって、被験体における状態または疾患を検出する、識別する、予測する、診断する、または監視する方法を企図する。

10

【0176】

E．定量的遺伝子解析の臨床応用

様々な実施形態では、本発明は、被験体における状態または疾患を検出、識別、予測、診断、または監視する方法を企図している。

【0177】

特定の実施形態では、被験体における遺伝子の様相、状態、または疾患を検出、識別、予測、診断、または監視する方法は、DNAクローンライブラリにおける1つ以上の標的遺伝子座の定量的遺伝子解析を実施して、1つ以上の標的遺伝子座での配列における変化を検出する、または識別することを含む。いくつかの実施形態において、この変化は、コピー数における変化である。

20

【0178】

1つの実施形態において、遺伝子状態、条件または疾患を検出する、識別する、予測する、診断する、または監視する方法は、細胞DNAまたはcfDNAを被験体の生体試料から単離する、または取得すること、細胞DNAまたはcfDNAを1個以上の末端修復酵素によって処置し、末端修復DNAを生成すること、1個以上のアダプタを末端修復DNAの各末端に付着させて、ゲノムDNAライブラリを生成すること、DNAライブラリを増幅して、DNAクローンライブラリを生成すること、DNAクローンライブラリ中の遺伝子当量数を決定すること、及び1個以上の標的遺伝子座の定量的遺伝子解析をDNAクローンライブラリ中で実施して、1つ以上の標的遺伝子座の、配列における変化、たとえば、SNP、転座、逆位、欠失、またはコピー数における変化などを検出する、または識別することを備える。

30

【0179】

特定の実施形態において、遺伝性疾患、遺伝子モザイク現象、胎児検査、父子鑑定、薬物治療に対する応答の予測、医学的状态の診断または監視、マイクロバイームプロファイリング、病原体スクリーニング、及び臓器移植の監視からなる群から選択される、遺伝子様相、または遺伝子の状態若しくは疾患を検出、識別、予測、診断、または監視する方法は、ゲノムDNAを被験体の生体試料から単離する、または取得すること、DNAを1個以上の末端修復酵素によって処置して、末端修復DNAを生成すること、1個以上のアダプタを末端修復DNAの各末端に付着させて、ゲノムDNAライブラリを生成すること、ゲノムDNAライブラリを増幅して、DNAクローンライブラリを生成すること、ゲノム当量数をDNAクローンライブラリ中で決定すること、及び1個以上の標的遺伝子座の定量的遺伝子解析をDNAクローンライブラリ中で実施して、1個以上の標的遺伝子座で配列におけるヌクレオチドのトランジション若しくはトランスバージョン、ヌクレオチドの挿入若しくは欠失、ゲノム再編成、コピー数における変化、または遺伝子融合を検出する、または識別することを備える。

40

【0180】

本明細書に企図される組成物及び方法によって検出される、識別される、予測される、診断される、または監視されることが可能である遺伝性疾患の例示的な例は、アルツハイマー病(APOE1)、シャルコー・マリー・トゥース病、レーベル遺伝性視神経症(LHON)、アンジェルマン症候群(UBE3A、ユビキチン-タンパク質リガーゼE3A

50

)、プラダー・ウィリー症候群(染色体15中の領域)、サラセミア(HBB、 α -グロビン)、ゴーシェ病(1型)(GBA、グルコセレブロシダーゼ)、嚢胞性線維症(CFTR、上皮クロライドチャネル)、鎌状赤血球症(HBB、 α -グロビン)、テイ・サックス病(HEXA、ヘキソサミニダーゼA)、フェニルケトン尿症(PAH、フェニルアラニンヒドロリアーゼ)、家族性高コレステロール血症(LDLR、低密度リボタンパク質受容体)、成人型嚢胞腎(PKD1、ポリシスチン)、ハンチントン病(HDD、ハンチンチン)、神経線維腫症I型(NF1、NF1腫瘍抑制遺伝子)、筋緊張性ジストロフィー(DM、ミオトニン)、結節性硬化症(TSC1、ツベリン)、軟骨形成不全(FGFR3、線維芽細胞増殖因子受容体)、脆弱X染色体症候群(FMR1、RNA結合タンパク質)、デュシェンヌ型筋ジストロフィー(DMD、ジストロフィン)、血友病A(F8C、血液凝固第VIII因子)、レッシュ・ナイハン症候群(HPR1、ヒポキサンチンアニンリボシルトランスフェラーゼ1)、及び副腎白質ジストロフィー(ABCD1)を有するが、これらに限定されない。

【0181】

本明細書で企図される組成物及び方法によって検出される、識別される、予測される、診断される、または監視されることが可能である癌の例示的な例は、B細胞癌、例えば、多発性骨髄腫、黒色腫、乳癌、肺癌(非小細胞肺癌腫またはNSCLCなど)、気管支癌、結腸直腸癌、前立腺癌、膵臓癌、胃癌(stomach cancer)、卵巣癌、膀胱癌、脳または中枢神経系癌、末梢神経系癌、食道癌、子宮頸癌、子宮または子宮内膜癌、口腔または咽頭の癌、肝臓癌、腎臓癌、精巣癌、胆道癌、小腸または虫垂癌、唾液腺癌、甲状腺癌(thyroid gland cancer)、副腎癌、骨肉腫、軟骨肉腫、血液組織の癌、腺癌腫、炎症性筋線維芽細胞腫、消化管間質腫瘍(GIST)、結腸癌、多発性骨髄腫(MM)、骨髄異形成症候群(MDS)、骨髄増殖性障害(MPD)、急性リンパ球性白血病(ALL)、急性骨髄球性白血病(AML)、慢性骨髄球性白血病(CML)、慢性リンパ球性白血病(CLL)、真性赤血球増加症、ホジキンリンパ腫、非ホジキンリンパ腫(NHL)、軟部組織肉腫、線維肉腫、粘液肉腫、脂肪肉腫、骨原性肉腫、脊索腫、血管肉腫、内皮肉腫(endotheliosarcoma)、リンパ管肉腫、リンパ管内皮肉腫(lymphangioendotheliosarcoma)、滑膜腫、中皮腫、ユーイング腫瘍、平滑筋肉腫、横紋筋肉腫、扁平上皮細胞癌腫、基底細胞癌腫、腺癌腫、汗腺癌腫、脂腺癌腫、乳頭癌腫、乳頭腺癌腫、髄様癌腫、気管支癌腫、腎細胞癌腫、肝細胞癌、胆管癌腫、絨毛癌腫、セミノーマ、胎生期癌腫、ウィルムス腫瘍、膀胱癌腫、上皮癌腫、神経膠腫、星状細胞腫、髄芽細胞腫、頭蓋咽頭腫、上衣腫、松果体腫、血管芽細胞腫、聴神経腫、乏突起神経膠腫、髄膜腫、神経芽細胞腫、網膜芽細胞腫、濾胞性リンパ腫、びまん性大細胞型B細胞リンパ腫、マントル細胞リンパ腫、肝細胞癌腫、甲状腺癌(thyroid cancer)、胃癌(gastric cancer)、頭頸部癌、小細胞癌、本態性血小板血症、特発性骨髄化生、好酸球増加症候群、全身性肥満細胞症、一般的な過好酸球増加症、慢性好酸球性白血病、神経内分泌癌、カルチノイド腫瘍などを含むが、これらに限定されない。

【0182】

1つの実施形態において、遺伝子損傷は、Cosmicデータベースで注釈がつけられた損傷(損傷及び配列データはオンラインで利用可能であり、CosmicウェブサイトのCancer Gene Census sectionからダウンロードされることが可能である)、またはCancer Genome Atlasで注釈がつけられた損傷(損傷及び配列データはオンラインで利用可能であり、Cancer Genome Atlasウェブサイトからダウンロードされることが可能である)である。

【0183】

本明細書で企図する組成物及び方法により検出、識別、予測、診断、または監視することができるがんに関連する1つ以上の遺伝子損傷を有する遺伝子の例示的な例としては、以下に限定するものではないが、ABCB1、ABCC2、ABCC4、ABCG2、ABL1、ABL2、AKT1、AKT2、AKT3、ALDH4A1、ALK、APC、

AR、ARAF、ARFRP1、ARID1A、ATM、ATR、AURKA、AURKB、BCL2、BCL2A1、BCL2L1、BCL2L2、BCL6、BRAF、BRCA1、BRCA2、Clorf144、CARD11、CBL、CCND1、CCND2、CCND3、CCNE1、CDH1、CDH2、CDH20、CDH5、CDK4、CDK6、CDK8、CDKN2A、CDKN2B、CDKN2C、CEBPA、CHEK1、CHEK2、CRKL、CRLF2、CTNNB1、CYP1B1、CYP2C19、CYP2C8、CYP2D6、CYP3A4、CYP3A5、DNMT3A、DOT1L、DPYD、EGFR、EPHA3、EPHA5、EPHA6、EPHA7、EPHB1、EPHB4、EPHB6、EPHX1、ERBB2、ERBB3、ERBB4、ERCC2、ERG、ESR1、ESR2、ETV1、ETV4、ETV5、ETV6、EWSR1、EZH2、FANCA、FBXW7、FCGR3A、FGFR1、FGFR2、FGFR3、FGFR4、FLT1、FLT3、FLT4、FOXP4、GATA1、GNA11、GNAQ、GNAS、GPR124、GSTP1、GUCY1A2、HOXA3、HRAS、HSP90AA1、IDH1、IDH2、IGF1R、IGF2R、IKBKE、IKZF1、INHBA、IRS2、ITPA、JAK1、JAK2、JAK3、JUN、KDR、KIT、KRAS、LRP1B、LRP2、LTK、MAN1B1、MAP2K1、MAP2K2、MAP2K4、MCL1、MDM2、MDM4、MEN1、MET、MITF、MLH1、MLL、MPL、MRE11A、MSH2、MSH6、MTHFR、MTOR、MUTYH、MYC、MYCL1、MYCN、NF1、NF2、NKX2-1、NOTCH1、NPM1、NQO1、NRAS、NRP2、NTRK1、NTRK3、PAK3、PAX5、PDGFRA、PDGFRB、PIK3CA、PIK3R1、PKHD1、PLCG1、PRKDC、PTCH1、PTEN、PTPN11、PTPRD、RAF1、RARA、RB1、RET、RICTOR、RPTOR、RUNX1、SLC19A1、SLC22A2、SLCO1B3、SMAD2、SMAD3、SMAD4、SMARCA4、SMARCB1、SMO、SOD2、SOX10、SOX2、SRC、STK11、SULT1A1、TBX22、TET2、TGFB2、TMPRSS2、TNFRSF14、TOP1、TP53、TPMT、TSC1、TSC2、TYMS、UGT1A1、UMPS、USP9X、VHL、及びWT1が挙げられる。

10

20

【0184】

特定の実施形態では、遺伝子損傷は、ヌクレオチドのトランジション若しくはトランスポージョン、ヌクレオチドの挿入若しくは欠失、ゲノム再編成、コピー数の変化、または遺伝子融合を含む。

30

【0185】

一実施形態では、遺伝子損傷は、ALK遺伝子の3'コード領域を別の遺伝子に融合させる遺伝子融合である。

【0186】

一実施形態では、遺伝子損傷は、ALK遺伝子の3'コード領域をEML4遺伝子に融合させる遺伝子融合である。

【0187】

本明細書で企図される組成物及び方法によって検出される、識別される、予測される、診断される、または監視されることが可能である胎児検査に適している条件の例示的な例は、ダウン症候群(トリソミー21)、エドワード症候群(トリソミー18)、パトー症候群(トリソミー13)、クラインフェルター症候群(XXY)、トリプルX症候群、XYY症候群、トリソミー8、トリソミー16、ターナー症候群(XO)、ロバートソン転座、ディジョージ症候群、及びウォルフ・ヒルショルン症候群を含むが、これらに限定されない。

40

【0188】

本明細書で企図される組成物及び方法によって検出される、識別される、予測される、診断される、または監視されることが可能である父子鑑定に適しているアレルの例示的な例は、D20S1082、D6S474、D12ATA63、D22S1045、D10

50

S 1 2 4 8、D 1 S 1 6 7 7、D 1 1 S 4 4 6 3、D 4 S 2 3 6 4、D 9 S 1 1 2 2、D 2 S 1 7 7 6、D 1 0 S 1 4 2 5、D 3 S 3 0 5 3、D 5 S 2 5 0 0、D 1 S 1 6 2 7、D 3 S 4 5 2 9、D 2 S 4 4 1、D 1 7 S 9 7 4、D 6 S 1 0 1 7、D 4 S 2 4 0 8、D 9 S 2 1 5 7、アメロゲニン、D 1 7 S 1 3 0 1、D 1 G A T A 1 1 3、D 1 8 S 8 5 3、D 2 0 S 4 8 2、及びD 1 4 S 1 4 3 4のうちの16個以上を含むが、これらに限定されない。

【0189】

本明細書で企図される組成物及び方法によって検出される、識別される、予測される、診断される、または監視されることが可能である薬物治療に対する応答を予測するのに適している遺伝子の例示的な例は、以下の遺伝子、A B C B 1 (A T P - 結着力セット、サブファミリーB (M D R / T A P)、メンバー1)、A C E (酵素を変換するアンジオテンシンI)、A D H 1 A (アルコール脱水素酵素1A (クラスI)、アルファポリペプチド)、A D H 1 B (アルコール脱水素酵素1B (クラスI)、ベータポリペプチド)、A D H 1 C (アルコール脱水素酵素1C (クラスI)、ガンマポリペプチド)、A D R B 1 (アドレナリン作動性、ベータ-1-、受容体)、A D R B 2 (アドレナリン作動性、ベータ-2-、受容体、表面)、A H R (アリアル炭化水素受容体)、A L D H 1 A 1 (アルデヒド脱水素酵素1ファミリー、メンバーA1)、A L O X 5 (アラキドン酸5-リポキシゲナーゼ)、B R C A 1 (乳癌1、早期発症型)、C O M T (カテコール-O-メチルトランスフェラーゼ)、C Y P 2 A 6 (チトクロムP450、ファミリー2、サブファミリーA、ポリペプチド6)、C Y P 2 B 6 (チトクロムP450、ファミリー2、サブファミリーB、ポリペプチド6)、C Y P 2 C 9 (チトクロムP450、ファミリー2、サブファミリーC、ポリペプチド9)、C Y P 2 C 1 9 (チトクロムP450、ファミリー2、サブファミリーC、ポリペプチド19)、C Y P 2 D 6 (チトクロムP450、ファミリー2、サブファミリーD、ポリペプチド6)、C Y P 2 J 2 (チトクロムP450、ファミリー2、サブファミリーJ、ポリペプチド2)、C Y P 3 A 4 (チトクロムP450、ファミリー3、サブファミリーA、ポリペプチド4)、C Y P 3 A 5 (チトクロムP450、ファミリー3、サブファミリーA、ポリペプチド5)、D P Y D (ジヒドロピリミジン脱水素酵素)、D R D 2 (ドーパミン受容体D2)、F 5 (凝固因子V)、G S T P 1 (グルタチオンS-トランスフェラーゼパイ)、H M G C R (3-ヒドロキシ-3-メチルグルタリル-コエンザイムA還元酵素)、K C N H 2 (電位依存性カリウムチャネル、サブファミリーH (e a g 関連の)、メンバー2) K C N J 1 1 (内向き整流性カリウムチャネル、サブファミリーJ、メンバー11)、M T H F R (5, 10-メチレンテトラヒドロ葉酸還元酵素 (N A D P H))、N Q O 1 (N A D (P) H脱水素酵素、キノン1)、P 2 R Y 1 (プリン受容体P2Y、G-タンパク質共役型、1)、P 2 R Y 1 2 (プリン受容体P2Y、G-タンパク質共役型、12)、P T G I S (プロスタグランジンI2 (プロスタサイクリン) シンターゼ)、S C N 5 A (ナトリウムチャネル、電位依存性、V型、アルファ (Q T 延長症候群3))、S L C 1 9 A 1 (溶質輸送体ファミリー19 (葉酸トランスポーター)、メンバー1)、S L C O 1 B 1 (溶質輸送体有機アニオン輸送体ファミリー、メンバー1B1)、S U L T 1 A 1 (スルホトランスフェラーゼファミリー、サイトゾル、1A、フェノールが好ましい、メンバー1)、T P M T (チオプリンS-メチルトランスフェラーゼ)、T Y M S (チミジル酸シンターゼ)、U G T 1 A 1 (U D P グルクロノシルトランスフェラーゼ1ファミリー、ポリペプチドA1)、V D R (ビタミンD (1, 25-ジヒドロキシビタミンD3) 受容体)、V K O R C 1 (ビタミンKエポキシドレダクターゼ複合体、サブユニット1)のうちの1個以上を含むが、これらに限定されない。

【0190】

本明細書で企図する組成物及び方法により検出、識別、予測、診断、または監視することができる医学的状態の例示的な例としては、以下に限定するものではないが、脳卒中、一過性脳虚血発作、外傷性脳損傷、心疾患、心臓発作、アンギナ、粥状動脈硬化、及び高血圧が挙げられる。

10

20

30

40

50

【0191】

本明細書で企図する組成物及び方法によりスクリーニングすることができる病原体の例示的な例としては、以下に限定するものではないが、細菌、真菌、及びウイルスが挙げられる。

【0192】

本明細書で企図する組成物及び方法によりスクリーニングすることができる細菌種の例示的な例としては、以下に限定するものではないが、*Mycobacterium*種、*Pneumococcus*種、*Escherichia*種、*Campylobacter*種、*Corynebacterium*種、*Clostridium*種、*Streptococcus*種、*Staphylococcus*種、*Pseudomonas*種、*Shigella*種、*Treponema*種、または*Salmonella*種が挙げられる。

10

【0193】

本明細書で企図する組成物及び方法によりスクリーニングすることができる真菌種の例示的な例としては、以下に限定するものではないが、*Aspergillus*種、*Blastomyces*種、*Candida*種、*Coccidioides*種、*Cryptococcus*種、皮膚糸状菌、*Tinea*種、*Trichophyton*種、*Microsporum*種、*Fusarium*種、*Histoplasma*種、*Mucoromycotina*種、*Pneumocystis*種、*Sporothrix*種、*Exserophilum*種、または*Cladosporium*種が挙げられる。

【0194】

20

本明細書に企図される組成物及び方法によりスクリーニングされることができるウイルスの例示的な例としては、インフルエンザA（たとえば、H1N1、H1N2、H3N2及びH5N1（鳥インフルエンザ）などの）、インフルエンザB、インフルエンザCウイルス、肝炎Aウイルス、肝炎Bウイルス、肝炎Cウイルス、肝炎Dウイルス、肝炎Eウイルス、ロタウイルス、ノーウォークウイルス群のうちのいずれかのウイルス、腸管アデノウイルス、パルボウイルス、デング熱ウイルス、サル痘、モノネガウイルス目、リッサウイルス（たとえば、狂犬病、ラゴスコウモリウイルス、モコラウイルス、ドゥベンヘイジウイルス、ヨーロッパコウモリウイルス1&2、及びオーストラリアコウモリウイルス）、エフェメロウイルス、ベジクロウイルス、水疱性口内炎ウイルス（VSV）、ヘルペスウイルス属（たとえば、単純ヘルペスウイルス1及び2型、水痘帯状発疹、サイトメガロウイルス、エプスタイン・パールウイルス（EBV）、ヒトヘルペスウイルス（HHV）、ヒトヘルペスウイルス6型及び8型）、モロニー Maus 白血病ウイルス（M-MuLV）、モロニー Maus 肉腫ウイルス（MoMSV）、ハーベイ Maus 肉腫ウイルス（HaMuSV）、マウス乳癌ウイルス（MuMTV）、テナガザル白血病ウイルス（GaLV）、ネコ白血病ウイルス（FLV）、スプーマウイルス、フレンド Maus 白血病ウイルス、マウス幹細胞ウイルス（MSCV）及びラウス肉腫ウイルス（RSV）、HIV（HIV1型、及びHIV2型を含むヒト免疫不全ウイルス）、ビスナ・マエディウイルス（VMV）、ヤギ関節炎脳炎ウイルス（CAEV）、ウマ伝染性貧血ウイルス（EIAV）、ネコ免疫不全ウイルス（FIV）、ウシ免疫不全ウイルス（BIV）、及びサル免疫不全ウイルス（SIV）、パピローマウイルス、マウスガンマヘルペスウイルス、アレナウイルス科（たとえば、アルゼンチン出血熱ウイルス、ポリビア出血熱ウイルス、サビア関連出血熱ウイルス、ベネズエラ出血熱ウイルス、ラッサ熱ウイルス、マチュポウイルス、リンパ球性脈絡髄膜炎ウイルス（LCMV））、ブニヤウイルス科（Bunyaviridae）（たとえば、クリミア・コンゴ出血熱ウイルス、ハンタウイルス、腎症候性出血熱を引き起こすウイルス、リフトバレー熱ウイルス）、フィロウイルス科（フィロウイルス）（エボラ出血熱及びマールブルグ出血熱を含む）、フラビウイルス科（キャサヌール森林病ウイルス、オムスク出血熱ウイルス、ダニ媒介脳炎を引き起こすウイルスを含む）及びパラミクソウイルス科（ヘンドラウイルス及びニパウイルスを含む）、大痘瘡及び小痘瘡（痘瘡）、アルファウイルス属（たとえば、ベネズエラウマ脳炎ウイルス、東部ウマ脳炎ウイルス、西部ウマ脳炎ウイルス）、SARS関連コロナウイルス（SARS-CoV）

30

40

50

、西ナイルウイルス、ならびにいずれかの脳炎 (e n c e p h a l i t i s) を引き起こすウイルスが挙げられるが、これらに限定されない。

【 0 1 9 5 】

本明細書で企図される組成物及び方法によって検出される、識別される、予測される、診断される、または監視されることが可能である移植患者における臓器移植を監視するのに適している遺伝子の例示的な例は、以下の遺伝子、H L A - A、H L A - B、H L A - C、H L A - D R、H L A - D P、及びH L A - D Qのうちの1つ以上を含むが、これらに限定されない。

【 0 1 9 6 】

特定の実施形態では、パイオインフォマティクス解析は、c f D N Aクローンライブラリにおいて解析されるゲノム当量数の定量化；標的遺伝子座における遺伝子変異体の検出；標的遺伝子座における突然変異の検出；標的遺伝子座における遺伝子融合の検出、または標的遺伝子座におけるコピー数変動の測定のために使用される。

10

【 0 1 9 7 】

F . コンパニオン診断

さまざまな実施形態において、遺伝性疾患についてのコンパニオン診断を提供し、このコンパニオン診断は、ゲノムD N Aを被験体の生体試料から単離する、または取得すること、D N Aを1個以上の末端修復酵素によって処置して、末端修復D N Aを生成すること、1個以上のアダプタを末端修復D N Aの各末端へ付着させて、D N Aライブラリを生成すること、D N Aライブラリを増幅して、D N Aクローンライブラリを生成すること、D N Aクローンライブラリ中のゲノム当量数を決定すること、及びD N Aクローンライブラリ中の遺伝性疾患と関連する1個以上のバイオマーカーの定量的遺伝子解析を実施することを備え、そこで1個以上のバイオマーカーのうちの少なくとも1個の検出、またはこれらの検出の失敗は、被験体を遺伝性疾患について処置するかどうかを示す。いくつかの実施形態では、このD N Aはc f D N Aである。特定の実施形態では、このD N Aは細胞D N Aである。

20

【 0 1 9 8 】

本明細書で使用する「コンパニオン診断」という用語は、特定の抗がん療法に結びついた診断検査を指す。特定の実施形態では、これらの診断方法は、生体試料に関連するバイオマーカーにおける遺伝子損傷の検出を含むことにより、患者を抗がん療法で治療すべきかそうでないかの迅速な識別が可能になる。

30

【 0 1 9 9 】

抗がん療法としては、以下に限定するものではないが、手術、放射線照射、化学療法、抗がん剤、及び免疫調節剤が挙げられる。

【 0 2 0 0 】

抗がん薬の例示的な例は、アルキル化剤（たとえば、チオテパ及びシクロホスファミド (C Y T O X A N (商 標))、アルキルスルホン酸塩類（たとえば、ブスルファン、インプロスルファン及びピボスルファン）、アジリジン類（たとえば、ベンゾドパ、カルボコン、メツレドパ、及びウレドパ）、エチレンイミン類及びメチラメラミン類（たとえば、アルトレタミン、トリエチレンメラミン、トリエチレンホスホルアミド、トリエチレンチオホスファオルアミド、及びトリメチルオロメラミンレジュームを含む）、ナイトロジェンマスタード類（たとえば、クロラムブシル、クロルナファジン、コロホスファミド、エストラムスチン、イホスファミド、メクロレタミン、メクロレタミンオキシド塩酸塩、メルファラン、ノブエンピキン、フェネステリン、ブレドニムスチン、トロホスファミド、ウラシルマスタード）、ニトロソウレア類（たとえば、カルムスチン、クロロゾトシン、ホテムスチン、ロムスチン、ニムスチン、ラニムスチン）、抗生物質類（たとえば、アクラシノマイシン類、アクチノマイシン、オースラマイシン、アザセリン、プレオマイシン類、カクチノマイシン、カリチアマイシン、カラピシン、カルミノマイシン、カルチノフィリン、クロモマイシン類、ダクチノマイシン、ダウノルピシン、デトルピシン、6 - ジアゾ - 5 - オキソ - L - ノルロイシン、ドキシソルピシンとそのペグ化製剤類、エピルビ

40

50

シン、エソルピシン、イダルピシン、マルセロマイシン、マイトマイシン類、ミコフェノール酸、ノガラマイシン、オリボマイシン類、ペプロマイシン、ポトフィロマイシン、ピューロマイシン、クエラマイシン、ロドルピシン、ストレプトニグリン、ストレプトゾシン、ツベルシジン、ウベニメクス、ジノスタチン、ゾルピシン)、代謝拮抗剤類(たとえば、メトトレキサート及び5-フルオロウラシル(5-FU))、葉酸類似体(たとえば、デノプテリン、メトトレキサート、プテロプテリン、トリメトトレキサート)、プリン類似体(たとえば、フルダラビン、6-メルカプトプリン、チアミプリン、チオグアニン)、ピリミジン類似体(たとえば、アンシタビン、アザシチジン、6-アザウリジン、カルモフル、シタラビン、ジデオキシウリジン、ドキシフルリジン、エノシタビン、フロクスウリジン、5-FU)、アンドロゲン類(たとえば、カルステロン、プロピオン酸ドロモスタノロン、エピチオスタノール、メピチオスタン、テストラクトン)、抗副腎剤類(たとえば、アミノグルテチミド、ミトタン、トリロスタン)、葉酸補充薬(たとえば、フォリン酸(frolic acid))、アセグラトン、アルドホスファミド配糖体、アミノレプリン酸、アムサクリン、ベストラブシル、ピサントレン、エダトラキサート、デフォファミン、デメコルシン、ジアジクオン、エルホルミチン、エリプチニウムアセテート、エトグルシド、硝酸ガリウム、ヒドロキシ尿素、レンチナン、ロニダミン、ミトグアゾン、ミトキサントロン、モピダモール、ニトラクリン、ペントスタチン、フェナメット、ピラルピシン、ポドフィリン酸、2-エチルヒドラジド、プロカルバジン、PSK(登録商標)、ラゾキサ、シゾフィラン、スピロゲルマニウム、テヌアゾン酸、トリアジコン、2,2',2''-トリクロロトリエチルアミン、ウレタン、ビンデシン、ダカルバジン、マンノムスチン、ミトブロニトール、ミトラクトール、ピポプロマン、ガシトシン、アラピノシド(「Ara-C」)、シクロホスファミド、チオテパ、タキソイド類(たとえば、パクリタキセル(TAXOL(登録商標)、Bristol-Myers Squibb Oncology, Princeton, N.J.)及びドキシタキセル(TAXOTERE(登録商標)、Rhne-Poulenc Rorer, Antony, France))、クロラムブシル、ゲムシタビン、6-チオグアニン、メルカプトプリン、メトトレキサート、白金類似体(たとえば、シスプラチン及びカルボプラチン)、ビンブラスチン、白金、エトポシド(VP-16)、イホスファミド、マイトマイシンC、ミトキサントロン、ピンクリスチン、ビノレルビン、ナベルビン、ノバントロン、テニボシド、アミノプテリン、ゼローダ、イバンドロネート、CPT-11、トポイソメラーゼ阻害剤RFS2000、ジフルオロメチロマイシン(difluoromethylomithine)(DMFO)、レチノイン酸誘導体(たとえば、Targretin(商標)(ベキサロテン)、Panretin(商標)(アリトレチノイン))、ONTAK(商標)(デニロイキンジフチトクス)、エスペラミシン類、カペシタビン、ならびに上記いずれかの薬学的に許容される塩類、酸類、または誘導体を含むが、これらに限定されない。この定義にさらに含まれるものとして、がんに対するホルモン作用を調節または阻害するように作用する抗ホルモン剤、例えば、抗エストロゲン剤(例えば、タモキシフェン、ラロキシフェン、アロマターゼ阻害4(5)-イミダゾール類、4-ヒドロキシタモキシフェン、トリオキシフェン、ケオキシフェン、LY117018、オナプリストン、及びトレミフェン(Fareston)を含む)、ならびに抗アンドロゲン剤(例えば、フルタミド、ニルタミド、ピカルタミド、ロイプロリド、及びゴセレリン)、ならびに上記のいずれかの薬学的に許容される塩類、酸類、または誘導体がある。

【0201】

免疫調節剤の例示的な例としては、以下に限定するものではないが、シクロスポリン、タクロリムス、トレスペリムス、ピメクロリムス、シロリムス、ベロリムス、ラフルニムス、ラキニモド及びイミキモド、ならびにこれらの類似体、誘導体、塩類、イオン類、及び複合体が挙げられる。

【0202】

いくつかの実施形態では、抗がん薬は、ポリ-ADPリボースポリメラーゼ(PARP)阻害剤を含むことができる。PARP阻害剤の例示的な例は、オラパリブ(AZD-2

10

20

30

40

50

281)、ルカパリブ(AG014699またはPF-01367338)、ニラパリブ(MK-4827)、タラゾパリブ(BMN-673)、ベリパリブ(ABT-888)、CEP-9722、E7016、BGB-290、3-アミノベンズアミドを含むが、これらに限定されない。

【0203】

本明細書で引用された全ての刊行物、特許出願及び発行された特許は、個々の刊行物、特許出願、または発行された特許が明確に及び個々に示され、参照により援用されているかのように、参照により本明細書に援用される。特に、国際PCT公開第WO2016/028316号の全内容は、参照により具体的に援用される。

【0204】

上記の発明を、明快に理解するための例示及び例を通じてある程度詳細に説明してきたが、本発明の教示を踏まえた上で、添付の請求項の趣旨または範囲から逸脱することなく、本発明にあるいくつかの変更及び修正を施してもよいことは、当業者には直ちに明らかになることである。以下の実施例は、単に例として提供されるものであり、限定として提供されるものではない。当業者であれば、本質的に同様な結果をもたらすために変更または修正ができるであろう様々な重要性の低いパラメーターがあることを容易に認識することになる。

【実施例】

【0205】

実施例1：断片化ゲノムDNAの配合物を含む試料のコピー数解析

断片化された野生型ヒトgDNA試料中に添加されたATMまたはBRCA2の不活化ヒト試料に由来するDNAを含んだ、断片化ゲノムDNAの配合物を細心に生成した。この試料タイプの利点は、組成を慎重に制御することが可能であり、試料の入手可能性が本質的に無制限であることである。

【0206】

野生型、ヒト女性ゲノムDNAを、健康な志願者によって提供された全血試料から精製した。ATM遺伝子(NA09596、ATM)を網羅するヘテロ接合性欠失を宿す不死化細胞からゲノムDNAを単離し、BRCA2(NA02718、BRCA2)のヘテロ接合性欠失を有する別々の試料をCoriellリポジトリから取得した。重要なことに、これらの試料は、ゲノムの残部にわたりほかの点では正常な倍数性を有するようにみえた。ATM試料は、男性ドナーに由来したため、X連鎖AR遺伝子についてのコピー数において、ヘミ接合性でもあった。無細胞DNA(cfDNA)を健康な女性または男性由来のドナー血漿試料から取得した。ライブラリ構築のために、ゲノムDNAを、Covaris機器に関して200bpの設定で超音波処理し、つぎにさらに「両側」DNAビーズ精製を使用して、サイズを選択した。ライブラリインプットDNA試料を図7に示す。

【0207】

断片化したcfDNA試料の適切な組み合わせを配合して、百分率で定義し、末端修復し、そしてゲノムライブラリに変換した。約500ngの各ライブラリを8個の試料セットに混合し、2304個のDNAプローブを含んだ前立腺プローブプールのコピー数減少(CNL)にハイブリダイズした。試料処理後、8個の試料の各セットを、6千万リード/試料に対応する、Illumina NextSeq NGS機器上で、約4億8千万個のパスフィルタリードの深度へ配列決定した。おおよそ95%のリードは、妥当な試料IDタグを含み、ヒト基準ゲノムにアライメントを取り、そしてこれらのうちの、約98%を意図された標的遺伝子座へマッピングした。シーケンシング深度全体を、1個のプローブにつき、1個のインプットゲノムあたりのリード数として測定し(平均ゲノム深度(2500)によって除算し、またプローブカウント(2400)によって除算する、標的上リード(6千万)として計算した)、1個のプローブにつき、1個のインプットゲノムあたり約10リードであった。コピー数減少解析のグラフ表示を図1に示す。コピー数攪乱を矢印で強調表示した。(試料1、女性DNA中の5%の男性DNA、試料2、女性DN

10

20

30

40

50

A 中の 5 % の A T M D N A (男性) 、 試料 3 、 女性 D N A 中の 5 % の B R C A 2 D N A (女性) 、 試料 4 、 純粋な女性 D N A) 。

【 0 2 0 8 】

C N L 呼び出し側は、冗長リードを識別し、これらを単一のコンセンサスリード中に縮合させ、つぎに各プローブ位置で定量化する。この情報を遺伝子ごとのコピー数平均値にさらに縮合した。最後に、各 C N L 測定で検出された偏差に、統計学的有意性を割り当て、これを $\log_{10} P$ 値の統計的有意性としてグラフで示す。

【 0 2 0 9 】

図 8 は、断片化され、配合されたゲノムライブラリ中の A R 遺伝子 (図 8 B) 及び A T M 遺伝子 (図 8 C) についてコピー数決定の箱ひげ図を示す。 A T M 試料が男性であるため、A R 遺伝子 (X 連鎖、ヘミ接合性) 及び A T M 遺伝子の両方は、C N L 挙動を示した。予測されたように、測定されたコピー変動の大きさは、中程度であった。図 9 B に示される統計解析は、観察されたコピー増減が統計的に有意であったことを実証する。さらに、均一なコピー特性を示すと予測された残りの遺伝子中で、ほとんど有意な増減は観察されなかった。これらの値は、さまざまなゲノム配合物について予測された頻度と良く関連していた。図 1 0 は、異性から c f D N A の微量なスパイクイン、または断片化 g D N A の微量な追加のいずれか一方を有する、主に c f D N A であった試料中で、統計的に有意なコピー増減も容易に観察されたことを示す。これらの値は、さまざまなゲノム配合物について予測された頻度と良く関連していた。断片化された g D N A 、及び c f D N A の両方に関してみられた、これらの結果は、同等であったことにより、アッセイの完全性を実証し、この完全性が臨床試料に翻訳されることを示した。

【 0 2 1 0 】

これらのデータは、遺伝子コピー数において 2 % のマイナーアレル頻度に下がった、わずかな変化を検出するアッセイシステムの能力を実証する。提示された実証例の焦点は、コピー数減少にあるが、技術は、染色体アーム重複及び局所的増幅を通して発生する遺伝子コピーにおける増加を含む、コピー数増加の検出に等しく良く適している。このアッセイは、S N V 、インデル及び遺伝子融合 (染色体再編成) を含む、他の種類のゲノム変異体を検出する能力をさらに保持する。重要なことには、これらのデータは、血漿に由来するゲノム D N A に、また組織などの他の供給源及び他の身体供給源に由来するゲノム D N A に、この方法を適用することが可能であることを実証する。

【 0 2 1 1 】

実施例 2 : 健康なドナー及び癌患者からの c f D N A のコピー数解析

以下の例は、ゲノムライブラリ構築中に、またハイブリダイゼーション処理後に、追加される分子特徴を使用して、コピー数解析を生成する方式を示す。Q i a g e n C i r c u l a t i n g N u c l e i c A c i d s E x t r a c t i o n キット (Q i a g e n , H i l d e n , G e r m a n y) を用いて、16 個の健康なドナー、及び 1 個の去勢抵抗性前立腺癌の患者の血漿から D N A を抽出した。Q u b i t 蛍光光度計 (T h e r m o F i s h e r , W a l t h a m , M A) 、及び対応する h s D N A 定量化キットを用いて、二本鎖 D N A の収量を定量化した。ゲル電気泳動法を 2 % のアガロースゲル上に用いて、サイズ規格として P C R マーカー (N e w E n g l a n d B i o l a b s , I p s w i c h , M A) によって、サイズ解析を実施した。試料からの c f D N A の収量によって、約 4 0 から 1 0 0 n g の c f D N A をライブラリ構築に使用した。

【 0 2 1 2 】

ライブラリ構築の基本的な特徴を図 1 1 A から 1 1 C に示す。最初に c f D N A を脱リン酸化した後に、二段階プロセスで修復し、末端を平滑末端化した。短い、1 0 n t のアンカー配列は、リン酸化ライゲーション鎖及び不活性パートナー鎖からなり、つぎに c f D N A にライゲーションされた。4 個のアンカー配列のセットを作製するために用いられた 8 個のオリゴヌクレオチドを表 1 に示す。

10

20

30

40

【表 1】

表 1：ライゲーションアンカーオリゴヌクレオチド

オリゴ I D	核酸配列	配列番号：
パートナー鎖 o ライゲーション鎖オリゴ o _ 1 6 - 1	G T A T G C C [3 - d A - Q] *	1
パートナー鎖 o ライゲーション鎖オリゴ o _ 1 6 - 2	A G C G T T A [3 - d C - Q] *	2
パートナー鎖 o ライゲーション鎖オリゴ o _ 1 6 - 3	T C G A C A T [3 - d G - Q] *	3
パートナー鎖 o ライゲーション鎖オリゴ o _ 1 6 - 4	C A T C A G G [3 - d T - Q] *	4
ライゲーション鎖オリゴ _ 1 6 - 1	/ 5 P h o s / T G G C A T A C G T **	5
ライゲーション鎖オリゴ _ 1 6 - 2	/ 5 P h o s / G T A A C G C T A G **	6
ライゲーション鎖オリゴ _ 1 6 - 3	/ 5 P h o s / C A T G T C G A T C **	7
ライゲーション鎖オリゴ _ 1 6 - 4	/ 5 P h o s / A C C T G A T G C A **	8

* [3 - d (A , C , G , または T) - Q] は、修飾塩基を示し、この中で、ヒドロキシ基は、リボース環の 2' 位上に存在する。

** / 5 P h o s / は、5' 位の塩基へ 5' リン酸基の化学付加を示す。

【 0 2 1 3 】

アダプタ構造は、アンカー配列にアニールしたアダプタ配列全長の付加によって完成した。32セットのアダプタ配列を、240個のメンバーから各構成し、図12から図22に示す。これらのアダプタをcfDNAに付着させ、ポリヌクレオチドキナーゼ、DNAポリメラーゼ及びDNAリガーゼの協調作用を通して伸長させ、ゲノムライブラリを生成した。シーケンシング前の品質管理ステップとして、得られたゲノムライブラリをqPCRによってカバレッジの深度について定量化した。つぎにゲノムライブラリを増幅し、特異的な遺伝子を標的にするプローブセットへハイブリダイズした（図11B）。ハイブリダイゼーション後に、プローブのプライマー伸長を用いて、キャプチャされたゲノム配列をコピーし、情報を付着したアダプタ中で符号化した（図11C）。標準的な次世代解析ソフトウェアを用いたシーケンシング解析後の例を図11Dに示す。この解析をシーケンシングラン上で実施し、このシーケンシングランは、32個の試料（28個の癌患者試料、及び4個の野生型対照）を含み、それは、シーケンシングリードの全体的な分布を表示する。

【 0 2 1 4 】

本明細書に記載される標的ハイブリッドキャプチャプラットフォームの中心となる機能は、それが複数の種類のゲノム情報を提供することである。キャプチャプローブの1つの必須機能は、標的領域にわたり深いカバレッジ深度で変異検出を提供することである。この機能をキャプチャプローブの配列構成、密度、及び配置によって管理し、TP53遺伝子に関して図23に示す（TP53プローブ配列を以下の表2に示す）。等しい有意性の、標的ハイブリッドキャプチャプラットフォームアッセイを有意な変異が検出されなかった領域中の等しいカバレッジ深度からの読み出しを生成した。これらのデータは、有害な変異が検出されなかった場合に統計的有意性を加えるため、医師及び患者に重要である。

10

20

30

40

50

【表 2 - 1】

表 2 : TP53 プローブ

名称	配列	配列番号 :
TP5 3_1	GGCACAGACCCCTCTCACTCATG TGATGTTCATCTCTCCTCC	7689
TP5 3_2	ATGGGGGGTGGGAGGCTGTCTCAGT GGGGAAACAAGAAGTGGAG	7690
TP5 3_3	GTCAGTCTGAGTCAGGCCCTTC TGTCCTTGAAACATGAGTTT	7691
TP5 3_4	CCTGAAGTCCAAAAAGGGTCTCAG TCTACCTCCCGCCATAAA	7692
TP5 3_5	TCATGCTGGATCCCCACTTTTC CTCTTGTCAGCAGCCAGAC	7693
TP5 3_6	GTTGGGGGTGGGGGTGGTGGGCC TGCCCTTCCCAATGGATCC	7694
TP5 3_7	CAGTTTCCATAGGTCTGAAAAT GTTTCCTGACTCAGAGGG	7695
TP5 3_8	CTGCCATGGAGGAGCCGCAGTC AGATCCTAGCGTCGAGCC	7696
TP5 3_9	GCAGAGACCTGTGGGAAGCGAA AATTCCATGGGACTGACT	7697
TP5 3_10	CTGGGGGGGCTGGGGGGGCTGAGG ACCTGGTTCCTCTGACTGC	7698
TP5 3_11	GCAGGGGGGATACGGCCAGGCAT TGAAGTCTCATGGAAGCC	7699
TP5 3_12	GTGGCCCCCTGCACCAGCAGCTC CTACACCGGCGGCCCTG	7700
TP5 3_13	GGGGGGAGCAGCCTCTGGCATT CTGGGAGCTTCATCTGGA	7701
TP5 3_14	CCGTGCAAGTCACAGACTTGGC TGTCCTCAGAATGCAAGAA	7702
TP5 3_15	CCCCGGACGATATTGAACAATG GTTCACTGAAGACCCAGG	7703
TP5 3_16	CCAGAAAACCTACCAGGGCAGC TACGGTTTCCGTCTGGGC	7704
TP5 3_17	TAGGTTTTCTGGGAAGGGACAG AAGATGACAGGGGCCAGG	7705
TP5 3_18	TGCTTTTATCTGTTCACTTGTGC CCTGACTTTCAACTCTGT	7706

10

20

30

40

50

【表 2 - 2】

TP 5 3_1 9	CCTGGGGCAACCAGCCCTGTCGT CTCTCCAGCCCCAGCTGC	7 7 0 7
TP 5 3_2 0	TTTGCCAACTGGCCAAGACCTG CCCTGTGCAGCTGTGGGT	7 7 0 8
TP 5 3_2 1	CCATCGCTATCTGAGCAGCGCT CATGGTGGGGGCGAGCGCC	7 7 0 9
TP 5 3_2 2	GCCATCTACAAGCAGTCACAGC ACATGACGGAGGTTGTGA	7 7 1 0
TP 5 3_2 3	CATGGCGCGGACGCGGGTGCCG GGCGGGGGTGTGGAATCA	7 7 1 1
TP 5 3_2 4	CCAGGGTCCCCAGGCCTCTGAT TCCTCACTGATTGCTCTT	7 7 1 2
TP 5 3_2 5	GAGGGCCACTGACAACCACCCT TAACCCCTCCTCCCAGAG	7 7 1 3
TP 5 3_2 6	CCTCAGGCGGCTCATAGGGCAC CACCACACTATGTGCGAAA	7 7 1 4
TP 5 3_2 7	AGGAAATTTGCGTGTGGAGTAT TTGGATGACAGAAACACT	7 7 1 5
TP 5 3_2 8	CTTGCCACAGGTCTCCCCAAGG CGCACTGGCCTCATCTTG	7 7 1 6
TP 5 3_2 9	GAGGCAAGCAGAGGCTGGGGCA CAGCAGGCCAGTGTGCAG	7 7 1 7
TP 5 3_3 0	CCTGGAGTCTTCCAGTGTGATG ATGGTGAGGATGGGCCTC	7 7 1 8
TP 5 3_3 1	ACTACATGTGTAAACAGTTCCTG CATGGGCGGCATGAACCG	7 7 1 9
TP 5 3_3 2	GGACAGGTAGGACCTGATTTCC TTACTGCCTCTTGCTTCT	7 7 2 0
TP 5 3_3 3	CTGCACCCTTGGTCTCCTCCAC CGCTTCTTGTCTGCTTG	7 7 2 1
TP 5 3_3 4	TCTCTTTTCCTATCCTGAGTAG TGGTAATCTACTGGGACG	7 7 2 2

10

20

30

40

50

【表 2 - 3】

TP 5 3_3 5	CCTCGCTTAGTGCTCCCTGGGG GCAGCTCGTGGTGAGGCT	7 7 2 3
TP 5 3_3 6	GACCGGCGCACAGAGGAAGAGA ATCTCCGCAAGAAAGGGG	7 7 2 4
TP 5 3_3 7	TCTCCCAGGACAGGCGACAAACA CGCACCTCAAAGCTGTTC	7 7 2 5
TP 5 3_3 8	TGCCTCAGATTCACTTTTATCA CCTTTCTCTTGCCCTCTTTC	7 7 2 6
TP 5 3_3 9	GGCATTTTTGAGTGTTAGACTGG AAACTTTTCCACTTGATAA	7 7 2 7
TP 5 3_4 0	CCTGAAGGGTGAAATATTCTCC ATCCAGTGGTTTCTTCTT	7 7 2 8
TP 5 3_4 1	CCTAGCACTGCCCAACAACACC AGCTCCTCTCCCCAGCCA	7 7 2 9
TP 5 3_4 2	CATCTTTTAACTCAGGTACTGT GTATATACTTACTTCTCC	7 7 3 0
TP 5 3_4 3	ATGGCTTTTCCAACCTAGGAAGG CAGGGGAGTAGGGCCAGG	7 7 3 1
TP 5 3_4 4	CCTGGAGTGAGCCCTGCTCCCC CCTGGCTCCTTCCCAGCC	7 7 3 2
TP 5 3_4 5	TCCGAGAGCTGAATGAGGCCTT GGAAC TCAAGGATGCCCA	7 7 3 3

【0 2 1 5】

キャプチャされたゲノム配列（図 1 1 C）のキャプチャプローブの連鎖は、各プローブ位置におけるゲノム深度の測定をも容易にした。実験に使用されたすべてのキャプチャプローブと関連するユニークリード数を測定した（図 2 4）。図 2 4 に示されるデータを、16 個の健康なドナーの c f D N A 試料を解析したシーケンシングランから導出した。TP 5 3 遺伝子中の 1 箇所のプローブ位置において各試料中で遭遇したユニークリードの深度を計算した（図 2 4 に示された生のユニークリードカウント）。ユニークリードの広範な試料間分布に反映されたように、各試料は、ユニークライブラリ深度を有した。また実験において、2596 個すべてのキャプチャプローブにわたるユニークリード深度の網羅的平均を計算した（図 2 4 B）。有意には、すべてのプローブについて測定された網羅的なユニークリード深度によって図 2 4 C に表された、単一のプローブ部位で観察されたりード深度の正規化は、正規化されたユニークリードの均一な密度を明らかにした。これらのデータは、試料間から解析のために選択された特定のプローブのキャプチャ性能が均一であり、それぞれの個々のライブラリのゲノム深度に比例したことを示す。

【 0 2 1 6 】

この同一の正規化機能を図 2 3 に示される 4 5 個の T P 5 3 - 特異的プローブに適用した (図 2 5 に示される正規化データ)。図 2 3 は、T P 5 3 コード領域のシーケンシング深度へのすべてのプローブの凝集寄与を示すが、図 2 5 は、個々の各プローブによって取得された正規化深度を示す。個々の各プローブによって取得された正規化深度は、いずれかの所与のプローブについて試料間で一般的に一貫していたが、一方のプローブを他方と比較したときにいくぶんばらつきがあった。いくつかの要因は、プローブ間で観察された正規化後のキャプチャ深度における差を支配し、最も有意なのは、互いに関するプローブの配置であり、またゲノム反復領域に対するプローブの近接性であった。ほとんどのプローブは、均一なキャプチャ挙動を示すが、キャプチャ性能が一貫していなかった 2 個のプローブを図 2 5 に矢印によって強調表示する。しかしながら、これらのデータは、これらのようなプローブが稀であり、容易に識別されることを示す。またこのようなものとして、それらを下流側コピー数解析から除外することが可能である。

10

【 0 2 1 7 】

均一なキャプチャ性能は、図 2 5 中に 4 5 個の T P 5 3 標的プローブによって示され、本明細書に記載される標的ハイブリッドキャプチャプラットフォームの一般的な特徴である。図 2 6 において、この実験中に、特性を明らかにされた 1 6 個すべての正常な c f D N A ライブラリについて、2 5 9 6 個のキャプチャプローブの 1 つのパネル中の各プローブについて平均キャプチャ深度を計算した。つぎに散布図解析を使用して、平均値を 3 個の代表的な試料と個々に比較した。各点は、異なるプローブを表し、グラフ上の各点の位置は、x 軸上の平均値と y 軸上の個々の試料との比較である。大多数のプローブの緊密な対角分布は、ほとんどのプローブの高い相関関係のユニークリードキャプチャ性能を反映した (3 つのグラフすべてについて、 R^2 相関関係 0 . 9 5)。重要なことには、プローブごとのシーケンシング深度の一貫性は、コピー数測定における標的ハイブリッドキャプチャプラットフォームの使用をサポートする。

20

【 0 2 1 8 】

コピー数に関して、プローブデータの最も直接的な処理は、常染色体中で「2」の二倍体平均値に発生する調整されたゲノム深度値をさらに正規化することである。X連鎖性の遺伝子座について女性に発生するプローブ値についても同様である。正常な男性中のX連鎖領域及びY連鎖領域について、平均されたコピー値を「1」へ適宜に設定する。上記に詳細に考察される、この数値変換を、染色体対照プローブ (表 3 で、2 2 個すべての常染色体上で選択された遺伝子座を標的とする 2 3 9 個のプローブ) のセット、X連鎖 A R 遺伝子を標的とする 1 9 9 個のプローブのセット、及び 4 5 個の T P 5 3 - 特異的プローブに適用した (図 2 7 A 及び 2 7 B)。各点は、個々のプローブについての値を表す。稀な「ノイズの多い」プローブを除き、領域中の大多数の個々のプローブカウントは、約「2」であった二倍体を含む値であると予測された。健康な男性中の A R 遺伝子についてプローブは、予測された「1」に近い平均値に関して増減した。

30

40

50

【表 3 - 1】

表 3 : 染色体対照プローブ

名称	配列	配列番号 :
Chr__1 __1	GTGTCTCGGCAACCACTCTTCA CCAATATCACAGTGGACA	7 7 3 4
Chr__1 __2	ATCCAAGGGGAGGAGATCAGTG CCCCATTTGTATCGCAC	7 7 3 5
Chr__1 __3	ACTTACTGAAGCAAGAACCTCA TCAAGCTGCCTCCCACCA	7 7 3 6
Chr__1 __4	AGTTTGTGATCCTCCTGTGGGC AACCTCAGCAGTCTGGTT	7 7 3 7
Chr__1 __5	GGAGAGCGGAGCTGCTCAGAGC TTGGCCAGGTTCTAAGTG	7 7 3 8
Chr__1 __6	GACTGTGGCAATGAGGCAGCTA AGTGTTTCACCAACTTCT	7 7 3 9
Chr__1 __7	GGTGTATTTTGTACAACGGTGGA CCCAGACACTGGAGTCAT	7 7 4 0
Chr__1 __8	GTTGGTCTATTCTTGCGGTTGT AAAAGTGGCCCAGAGTGA	7 7 4 1
Chr__1 __9	GTGAGCCTTCTCTCACCATTTCT GTCCAAAATAGCAGCCCT	7 7 4 2
Chr__1 __10	CAGCCTAGATATGATTCTCTCAC TACCCTGTTCCATGGTTC	7 7 4 3
Chr__1 __11	AAAGAATGTGTTGGCTCATGAT CAGACTTGAGCACTTGGG	7 7 4 4
Chr__1 __12	CCTAGGCTGTTGCTGCTGGACC TGTTTGTGCTTTCATCACA	7 7 4 5
Chr__2 __1	CAGTTGACCCTTTCAGCCACAGG GGTTTGAACCTTTGAAGGA	7 7 4 6
Chr__2 __2	AGGACCTGAGTATGCACGTTTT GGTATACTGGGTAGGGGT	7 7 4 7
Chr__2 __3	TATCAGCTGGGATGGTCCGGTC AGCAGCATTACCTGT	7 7 4 8
Chr__2 __4	TGCCTGCTCAGCCCAGATTTCA GTCATGCTGGCCATAAAC	7 7 4 9
Chr__2 __5	CTGGGGGGGTGAGGTTTGAGGTT TGAGTGTGGGATGTGAGG	7 7 5 0
Chr__2 __6	CCAGCTTTTTTCAGAAGCTGGGA AAGTAATAACCCGTGTTG	7 7 5 1
Chr__2 __7	CCCAGCGCCCGTGGCTTTGGCT CCTCAGTCCCATTAAAT	7 7 5 2
Chr__2 __8	TATACCACCAAGTCTACCTACT GCCTGCACATGCTATGGC	7 7 5 3
Chr__2 __9	GGTCAATCCGGCACTACTGGTT GTCCAAAGGGAGGTTACT	7 7 5 4
Chr__2 __10	AATCAAACATCAGGACCGCCCA CAGCACAGGTCAATGAAC	7 7 5 5
Chr__2 __11	GTGTCTCCTGGAGGTGCATGGG TGGTTTTGAACCTTCATTG	7 7 5 6

10

20

30

40

【表 3 - 2】

C h r _ 2 _ 1 2	G A C C C A T G T A A G G G G T T G G G T T A T G T T C T C C T T T T G C C C A	7 7 5 7
C h r _ 2 _ 1 3	T C A C T G A C A T G C G A A G C T G G G A A C G A G A A A A T G C A C A T C C	7 7 5 8
C h r _ 2 _ 1 4	T C C T A C A G T G C T T A G G G A T G A A T C T G G C A A A G A A G G A T G C	7 7 5 9
C h r _ 2 _ 1 5	G A A A G C A G T C C T T A C C A C A A G A A G A C C C C G A T G T G G T G G T	7 7 6 0
C h r _ 2 _ 1 6	A T T G C T C A C T G G C T G G C T T G C A T T T G G T A T G C G A T T G G G A	7 7 6 1
C h r _ 2 _ 1 7	G T C C C T G G G A C C A T C T G T G C A T T G T T C T T G T A A C T G G A A A	7 7 6 2
C h r _ 2 _ 1 8	G A C C G A A T G G C G A A C G C A G T G A A T A G A T C A G G A G G G A A A A	7 7 6 3
C h r _ 3 _ 1	G A A G G A A T G G A G T G G A A C A G A T A G G G G T G A G G G A A T A A C G	7 7 6 4
C h r _ 3 _ 2	C C A C T G C C A T C C T C A G A G G G A G A T T C A C A A G T C T C A C A A T	7 7 6 5
C h r _ 3 _ 3	A T C C A G G C T T C A T G T T C A A A T G C A A T G G C C C T T G C C C C A T	7 7 6 6
C h r _ 3 _ 4	A A A T T T C C C C T G G C T C C C T A C T G C T T T G C A G G C C A A G T A A	7 7 6 7
C h r _ 3 _ 5	A C C T T A A A G A C G G G C C C A C A T C T C T T T G G A T G G G A T T A G G	7 7 6 8
C h r _ 3 _ 6	G G G C T T C G G T T T T G G C G A A G G T G C T C A C A A T C T T G A T A T C	7 7 6 9
C h r _ 3 _ 7	T G A G C T G T C C T T C A T G C C T G C A T T T C C C A T G T C T G T C T T C	7 7 7 0
C h r _ 3 _ 8	A T C T T T A T C C A G G G C T A C C A G T G G T G G G T C C A A A A T G A C T	7 7 7 1
C h r _ 3 _ 9	T A C A G G T G A A G G A T G T C A A C G A G T T T G C T C C C A C C T T C A A	7 7 7 2
C h r _ 3 _ 1 0	G C T G T T G T G A C G G A G G G C A A G A T C T A T G A C A G C A T T C T G C	7 7 7 3
C h r _ 3 _ 1 1	A A T G A A G G G G A T T C A A G C C T T G C C A C C G A C T T A C A G G A A G	7 7 7 4
C h r _ 3 _ 1 2	T G T G A G C G T A C T T T C T C C C C C A G G T T G A A G A G G A A T G A G T	7 7 7 5
C h r _ 4 _ 1	A T T C C A A G T C C A G G T C C C A A A T C T A T C A G T A C C G G C T G G C	7 7 7 6
C h r _ 4 _ 2	G A C A C A G A G T G C A T G A A G A C C G T T C A A A T A T G T C A G G G A C	7 7 7 7
C h r _ 4 _ 3	C A T G A G T C C T T C T A T G A C T C C C T C T C A G A C A T G C A G G A A G	7 7 7 8
C h r _ 4 _ 4	T T T T T A G G A G A C A G G T A C C C A C T G T C T G G T G A C G A G G A C T	7 7 7 9
C h r _ 4 _ 5	C C T T C T G T T G A G T C G C T A G G A G A T G C C T C A G T T C A A C A A T	7 7 8 0

10

20

30

40

【表 3 - 3】

C h r _ 4 _ 6	G A C A G A A A C T T C A T A C C C A A G A G C T G C T T T T C T C A G C T G G A	7 7 8 1
C h r _ 4 _ 7	C A G G C A A C T T T T G G C A A G A C C A A G T C A G C C T T C T C A T C T C T	7 7 8 2
C h r _ 4 _ 8	C C C T T G C T A C C A T C A C T G T T G T C A T C T G T G C T T T G C A T T C C	7 7 8 3
C h r _ 5 _ 1	A G G T C T C A C T C C A A C T G C C C C T G T A T T A G A G C T A G G C T G C	7 7 8 4
C h r _ 5 _ 2	G A A A C C A T G C G G G A T T C A T C T T T G T C A G A G T G G A G C G G C A	7 7 8 5
C h r _ 5 _ 3	T A T G A A A T T A G G C G G T G G T T G G A C G T G A C T G T G T G T T G A C	7 7 8 6
C h r _ 5 _ 4	T G A A A C T T G C A T G A C A T A C T G C G G C T G C C C A T T C A C T A G G	7 7 8 7
C h r _ 5 _ 5	T G C T T C T T G T T T A T A A C T C C C C T G G C C A C C A T C T C G G G C T	7 7 8 8
C h r _ 5 _ 6	A T T C C C T C T C A T T T G T G G T T G G T G G C T G G A T A T C T G T T C C	7 7 8 9
C h r _ 5 _ 7	A G C A T C A G C A T T T C C C T G T G G A C T T A C C T C T C T C A G T A G T	7 7 9 0
C h r _ 5 _ 8	A A A A T T T A A A G G T C G G C G G T A A G G C T G A A A G C C A A C A G G C	7 7 9 1
C h r _ 5 _ 9	G A G T G T G T C G G T C A G A A G G A A C A C C T G A G A A A C C G C T T T A	7 7 9 2
C h r _ 5 _ 1 0	C A T A G C A A A T A C C T G T C G C T G A G C C A G G A G T A A A G T C T G G	7 7 9 3
C h r _ 5 _ 1 1	A A G A G G C T C T G A G C T C T T G A T A G A G G T T A C A T G G G G A G C A	7 7 9 4
C h r _ 5 _ 1 2	G G A G A C A A C T T A G G A G G T T A T C T A G A C C A T T C C C G C C T T C	7 7 9 5
C h r _ 5 _ 1 3	G T G T T T C C T C C C A G C A T G C A C T T T G T G G C T G C C T T T C T T T	7 7 9 6
C h r _ 5 _ 1 4	T G G C T T G T G T A G C G T G T T T C A T T T T G G A A C C T T G G A G C C G	7 7 9 7
C h r _ 5 _ 1 5	G A C A C C T C T G G T G C A G T T T T G A G G C T G G C C G G G A A G G G A T	7 7 9 8
C h r _ 5 _ 1 6	G T T T C A G A T C T T G C A A T G G G A G G G A T C G A C T C G G C C C T T T	7 7 9 9
C h r _ 5 _ 1 7	T G C C T A A A T C A G A A A T G G G C T A C T T C C C T T G G C C A C A T C C	7 8 0 0
C h r _ 5 _ 1 8	C A A T C T A C C A C C T C A A G G T T C A C G C G T G G A T T C T A C A C C T	7 8 0 1
C h r _ 6 _ 1	G A G T T T T T C T T T C A G G T A G T C T G A G A T G G C C C G C A C C A A G	7 8 0 2
C h r _ 6 _ 2	T A C T A T A A A G A A G G C A C C T C T A G G C T T G G C A A G C A C A C G T	7 8 0 3
C h r _ 6 _ 3	G G C A G A T T C G A T G G G A C T T T A G A C A C T T G C T T T G C T C C C T	7 8 0 4

10

20

30

40

50

【表 3 - 4】

C h r _ 6 _ 4	C A A A T G T C C C C A T G C A A A C A T G T C C C G C A C T G T G T G G T A A	7 8 0 5
C h r _ 6 _ 5	A C A T G T G T A A T C T T C T T C T C C T A G G G C G G C A G A A C T C A T G	7 8 0 6
C h r _ 6 _ 6	C C C G A G G A A A G C T C C T C T T T G C T G A C T G T A A T G T A C T G C A	7 8 0 7
C h r _ 6 _ 7	G A G G A C A G C A T T C G C A T A T C A G G T C G A A A T T T C T C C G C G A	7 8 0 8
C h r _ 6 _ 8	G T C C A G C T T T C A T C C T T G A T C C T G C T A C T C T A G G C T C T C C	7 8 0 9
C h r _ 6 _ 9	A C T G A T G G T G T T C A C T T G C A C C A T C A G G T C T G A T G G A G G A	7 8 1 0
C h r _ 6 _ 1 0	A A T T G G T T C A C A A A G C G T C G G G T G A T C C A G T A A C A G T C G A	7 8 1 1
C h r _ 6 _ 1 1	C A G A A C T C T G C T C T A A C G C C A A G C C T T C A A T A T G T C T T C G	7 8 1 2
C h r _ 7 _ 1	C A A T T C T T A C C A T C C A C A A A A T G G A T C C A G A C A A C T G T T C	7 8 1 3
C h r _ 7 _ 2	A C T A C A C C T C A G A T A T A T T T C T T C A T G A A G A C C T C A C A G T	7 8 1 4
C h r _ 7 _ 3	T G C T A T A G A C G C A C A A A C G A C C G C G A G C C A C A A A T C A A G C	7 8 1 5
C h r _ 7 _ 4	C C A T G A C T T A T G T G C A G C T T G C G C A T C C A G G G G T A G A T C T	7 8 1 6
C h r _ 7 _ 5	A G G A G T T G G T G G C T A A A C C G C T G A C T T T T C T A T T G C A G A C	7 8 1 7
C h r _ 7 _ 6	G A A A T A T A A C A G G A C C A G A A G T G G C T C G C A G G A G A C T C A T	7 8 1 8
C h r _ 7 _ 7	T A G C C A G A C A G A A G G C G G A C A C T G A T G A T A C C T C A A G A C T	7 8 1 9
C h r _ 7 _ 8	G T T T G C C A C C A G C G A A G A G A G C C A T C C T G G T A G A A T T G G A	7 8 2 0
C h r _ 7 _ 9	G G A G A T A T G C A C T T G C C C T T T G G T A A T C C T G C T C C T T C T G	7 8 2 1
C h r _ 7 _ 1 0	A A A A C T A A C C A G T A A G T A C A G G G A G G G A C C G A G A G G C A T C	7 8 2 2
C h r _ 7 _ 1 1	A A G A A C A C C A G T C C A T A A A G A C G C A T G T C C G G T G A T G C C T	7 8 2 3
C h r _ 7 _ 1 2	A A T C T G T T T A G A C T G A G C A A C T G T G C C A G C A G A G G G A C C T	7 8 2 4
C h r _ 8 _ 1	A A G A T G G C G A A G G T C T C A G A G C T T T A C G A T G T C A C T T G G G	7 8 2 5
C h r _ 8 _ 2	C C A T G C C T G C C A G C T G A T A A G A T T T G G T T A C C T T T C C A T G	7 8 2 6
C h r _ 8 _ 3	G C T G C A A G A A A G C G T A A G A T T G C C A T T C G A A A A G C C C A G G	7 8 2 7
C h r _ 8 _ 4	A T G C A G G A G T A C A A T G T G G G C A T G T C C A C C C T C T A C G A C A	7 8 2 8

10

20

30

40

50

【表 3 - 5】

C h r _ 8 _ 5	AGAAACGGCTTTTGCTGTCTTCCG GCAAACCTATGGTTCTGA	7 8 2 9
C h r _ 8 _ 6	TGGCTTTTGGCGCTTTAAGGCCA GACACGGCATTTAAAAAGC	7 8 3 0
C h r _ 8 _ 7	GCAGGCAGAGAAAGATGGCTTT AGAAACCTCTTCCCCACC	7 8 3 1
C h r _ 8 _ 8	TCAGCTGTGGCCATTGGTGGAT CTCATCCTTAGTACTAGT	7 8 3 2
C h r _ 8 _ 9	CCATGGTTCTGTGAGACTGGTA GAAAGCACAGACCCCTTA	7 8 3 3
C h r _ 9 _ 1	AATGTGCTTATCACTCGTGATG GGGTCCTGAAGCTGGCAG	7 8 3 4
C h r _ 9 _ 2	AGGGTCTCATTTTAAAGACAGCT TGATTTGAGGGTGAGGGG	7 8 3 5
C h r _ 9 _ 3	CAGTTGCAAACCATACTTCCTT CAGCCCAGTCCTGTCTAT	7 8 3 6
C h r _ 9 _ 4	GTCTAAGGGCATCTTACCTCCA AGAACTGCTTGAGGCGTA	7 8 3 7
C h r _ 9 _ 5	TACCTAGGGAATGACCACTAAG CACCATCTCCGTCCTCT	7 8 3 8
C h r _ 9 _ 6	GGAAGAGAGGAGGGTCATCCAG TCAGTTTTTGACAGGAATCT	7 8 3 9
C h r _ 9 _ 7	TGCTGTCAGTGTCGGAAGAAACC TACCTGCGTTTTCTTAGAA	7 8 4 0
C h r _ 9 _ 8	CATCATACCTATGGCATAGCCA TCAGGGCACTGCAGTTTG	7 8 4 1
C h r _ 9 _ 9	TATATCTCACGTGACCGAGGAT GGGTCTGTGGGCATTACACA	7 8 4 2
C h r _ 9 _ 10	GAAATGGCCATCTATAGGTGGG AACCACCTCCAGTGTCACA	7 8 4 3
C h r _ 1 0 _ 1	GGAAACCTTTTCAGTCTCTACTA GAAGCGCGGAGAGAACTC	7 8 4 4
C h r _ 1 0 _ 2	TCTGGCCCGGCATTCATTTAAGG CCTAAGGATGAAGGCGGT	7 8 4 5
C h r _ 1 0 _ 3	AGATACCCTATCGTTCCTTATC TCAGCGAAACAACCTCCCC	7 8 4 6
C h r _ 1 0 _ 4	CGCAACTCCTCCAGATCGCAGT GGTGCTTCTTCACTTTCA	7 8 4 7
C h r _ 1 0 _ 5	TGATTCCATGGTTGCCCGTATA CTCCATAAGGCGGTACTT	7 8 4 8
C h r _ 1 0 _ 6	ATACCATATCCGGCTTGTTAG GAGGAGGTATTACAGGGG	7 8 4 9
C h r _ 1 0 _ 7	GTACCTGTTAAACCCAGACGCAA TTCCTCCACAGTACACAG	7 8 5 0
C h r _ 1 1 _ 1	ATGTGACACTTGTCATCCAGGGA GGTCACCATCTGTGTATG	7 8 5 1
C h r _ 1 1 _ 2	CTAGGTCCTGAAGAGGTGGCAA GGAACCAGGACAGAACAT	7 8 5 2

10

20

30

40

50

【表 3 - 6】

C h r _ 1 1 _ 3	TCTGTCATTGGGTGACGCCATCT AGACTCTTGGCTTTGGGA	7 8 5 3
C h r _ 1 1 _ 4	AAGGTATAGAGCTGGGCGGCTT TCCTCGTTATAGGTGGAG	7 8 5 4
C h r _ 1 1 _ 5	CTCCTACGTAGCCGGGTAGAAA CTTATGGCAGAAGTCAGG	7 8 5 5
C h r _ 1 1 _ 6	TGGATTCCCAGGGTTAATTGTG ACCCATTGCAGGAAGGTG	7 8 5 6
C h r _ 1 1 _ 7	AATGCTGTCTCTACTATGGTCTG TACCTGTCCCAGAGGTGG	7 8 5 7
C h r _ 1 1 _ 8	GTGCACCTGGAGAGCATACAGG GCACTGACTTGTAGATCA	7 8 5 8
C h r _ 1 1 _ 9	TTCCATCTCGCATAACCTGCCC CTAAACTCTTCTCGGTTC	7 8 5 9
C h r _ 1 1 _ 10	ATGAAGGCCTGCTTTGAGTTAT CAGATAGGAAGGGGCCAG	7 8 6 0
C h r _ 1 1 _ 11	AGGTCATGTCCCCTTTTGGCT GAACCTAGTTTTTGCCCAA	7 8 6 1
C h r _ 1 2 _ 1	CTGCATTCTCCATGAGTAGAGT ACGAGCCTCATGTTGGTA	7 8 6 2
C h r _ 1 2 _ 2	AAGGCTGTCTTCCACCAACTGGG TAGGTGTGGATCAAGACC	7 8 6 3
C h r _ 1 2 _ 3	CTGACTTTGGGTGTTGGGGAGTC GGTGGTCCCTTCTTCCATT	7 8 6 4
C h r _ 1 2 _ 4	ACTGCAGAGGACCAGACTGGGA AAACAACGATATGGCAGG	7 8 6 5
C h r _ 1 2 _ 5	CCTGGCTTAGAAAGTCTGGCCGG TCCTTCTTTCAGCTTCTTA	7 8 6 6
C h r _ 1 2 _ 6	AATCTCAGAAAGAGTTCCCTGGG ACCATGGCAAATGGTGGC	7 8 6 7
C h r _ 1 2 _ 7	ACATTATATCCGGTCCAGGAAT ATCTGGCTCAGGCTGGGT	7 8 6 8
C h r _ 1 2 _ 8	AAGCACAGGAAATGTGCCCTCAC ACGACTTTCACATGCCCTT	7 8 6 9
C h r _ 1 2 _ 9	GGGGGCTTTGCGGGAAGAGGGG ACTAAACAACCCCTTCTGT	7 8 7 0
C h r _ 1 2 _ 10	AAAAGAAATGCGATCAGCGCAA CCCATCCGGTGTGGCGCT	7 8 7 1
C h r _ 1 2 _ 11	GGCAGTGGTACCATGACATACT TAGCAGAGATGGACTACA	7 8 7 2
C h r _ 1 3 _ 1	ATTTCCCATGCGAGAGGTAGCT TGCCCAGGCTGTTGGATA	7 8 7 3
C h r _ 1 3 _ 2	TTCCATGCCGAGTCCTGATGGA AACTAGCACTGAAAGACC	7 8 7 4
C h r _ 1 3 _ 3	TCACGGGAGCTTCCCTTCACTGA GTTCTGCGAATCTGAAGC	7 8 7 5
C h r _ 1 3 _ 4	TTTCCAGAGATGAAGCACTACC CAGTCTTACCCAAGTTCG	7 8 7 6

10

20

30

40

50

【表 3 - 7】

C h r _ 1 3 _ 5	CCACCGAGAACAGTGATGAAGG ACTTAAAGTGAGAGATGG	7 8 7 7
C h r _ 1 3 _ 6	GTTCACCTCGTCGGTTTTTTCACC AACACAGACTAGCCTCA	7 8 7 8
C h r _ 1 3 _ 7	ACGCAGCTGTGTTGAGTGCACA GGAAGCTCTTAGGGTTAA	7 8 7 9
C h r _ 1 3 _ 8	TCTCAGTGAAACAGAGGGCTCAC TGAGAGGACTTTGAATAC	7 8 8 0
C h r _ 1 3 _ 9	ATGGCACAGGCCACATACTGGA ATGAATGACGGGCTTCAT	7 8 8 1
C h r _ 1 3 _ 10	TGCTGCTTGATGGTGGCATCAC TGTCCCCTCATTCATGA	7 8 8 2
C h r _ 1 4 _ 1	GGACACATGTGGACAGTGTGAA ACCTCAGAACACTAACCC	7 8 8 3
C h r _ 1 4 _ 2	AAGTTCTTATCCTTAGGGACCC AGCGGAGACCTTG GTTCT	7 8 8 4
C h r _ 1 4 _ 3	CGACGATGCCTGGGAATAGGAT CCATGGGATTGATGAGAA	7 8 8 5
C h r _ 1 4 _ 4	GGGAGCCATGAAGATTTCTCCC AGCTCCTGAGGAAC TTTG	7 8 8 6
C h r _ 1 4 _ 5	TCTGGTCCCTCAAGTCCTCAGCT GTAGAAGTTCTCATTTGCG	7 8 8 7
C h r _ 1 4 _ 6	TGCCAACCCCTGGAAACTGGCTT GTGTGTCCACAACAGAAA	7 8 8 8
C h r _ 1 5 _ 1	TAGGTGACAGCACTGTCCTTTC CCTGCCATTTG CAGGGAA	7 8 8 9
C h r _ 1 5 _ 2	TTCTTCTAGATGGCAGACATTG TTGAGGCCTCCCGTACCT	7 8 9 0
C h r _ 1 5 _ 3	AGAGAGCTGCGAGACAAGACTT GGAGTGCGACAAGATTTTC	7 8 9 1
C h r _ 1 5 _ 4	TTCAATCAGGTACTCCGAGTTTC CCTTGAGAGGCCAAAAGGA	7 8 9 2
C h r _ 1 5 _ 5	AGGAATATGGGGTCCATCTGAG ACTCGCAAGTGATGATAC	7 8 9 3
C h r _ 1 5 _ 6	GATCTCCAGGACCAGCTCTCAG AAATGCACGATGAACTGG	7 8 9 4
C h r _ 1 5 _ 7	ACAGTGTGATGGAGCAGCAGTC CAAGTTTCATCCTCCAAGA	7 8 9 5
C h r _ 1 5 _ 8	AAGATGACAGGATCCAGGAAAC AAGACGCATGGGCCAGAA	7 8 9 6
C h r _ 1 5 _ 9	AAAGAGTGGGTCTGTTAATAAT CAGGCCGAGACCACAGC	7 8 9 7
C h r _ 1 5 _ 10	CACCCTTGTTCTGTGGCCCTTGC TTGGTAAACTGGTATCCA	7 8 9 8
C h r _ 1 5 _ 11	CCCAAGTATGGGTGAGGATGCT AGAAAATGCCACATAATG	7 8 9 9
C h r _ 1 5 _ 12	AAGACTGTCATTGGTAGGTCAT GATCCTTGGCAGCATGAC	7 9 0 0

10

20

30

40

50

【表 3 - 8】

C h r _ 1 6 _ 1	GTGGGGACGGTTCATTATCAGCT TTCTGGACACACAGACAG	7 9 0 1
C h r _ 1 6 _ 2	TGAGAGGCCCAAAGAATATCAGT TGACTCTGGATCAGGGGC	7 9 0 2
C h r _ 1 6 _ 3	GAGGCTTTTTTAGGGCAGCGAGA AAACGGGGAACCTTCATTCC	7 9 0 3
C h r _ 1 6 _ 4	AGGACTTCTCTGGACCTGTGCC TCAACTACTCACCTGGAT	7 9 0 4
C h r _ 1 6 _ 5	TGGCCACAAATGTTGCCCTCCAG CTGCTCAATGTTCTCCAA	7 9 0 5
C h r _ 1 6 _ 6	CTGGCATTGGTGAGTAATAGGA GCCAGACGGGTCTGTGTT	7 9 0 6
C h r _ 1 6 _ 7	ATACTTACCTGCACGAGAATGA GTTTGGAGCGCAAGGGGG	7 9 0 7
C h r _ 1 6 _ 8	TTCCCCCAGAGACTCTGTCCAC TATGGACATTAAAAATGTG	7 9 0 8
C h r _ 1 6 _ 9	GTGCTACCCTCCTCCCTTCAGG TTATGTGGTCCAGGCTTT	7 9 0 9
C h r _ 1 6 _ 10	TAAGTGGAACAACATTCCCTTC ATTATAGCCCTTCGTGGG	7 9 1 0
C h r _ 1 6 _ 11	GCAACGTCAACAACACTACTACGT GCACAAGCGCCTCTACTG	7 9 1 1
C h r _ 1 7 _ 1	GCGGATGTCGTTATGGGACAGG TACAAGTAGATAAGTTGC	7 9 1 2
C h r _ 1 7 _ 2	GTGGTCAACCATCTCTTCAAACC ATTTGGACTGGGCCTGGT	7 9 1 3
C h r _ 1 7 _ 3	AAGCCAAGGAGTTCTGAGAGAG CTTAGCTAAGTTCTTCGC	7 9 1 4
C h r _ 1 7 _ 4	TTTTTTTAGTACCCCAGTG TGTA AGACCAACTGAGGGTGGC	7 9 1 5
C h r _ 1 7 _ 5	GTTGTTCATTGGGGCTATAGACA TAAGCACCTTCCGGAATC	7 9 1 6
C h r _ 1 7 _ 6	CTGAGTGTCGAGGGGGAAGATA TTGGTGAAAGACCTGTTCT	7 9 1 7
C h r _ 1 7 _ 7	GTCAGACCCTGTCTCTCGTCTCC TTTACCTTGTCTCGATTT	7 9 1 8
C h r _ 1 7 _ 8	TAAACTATGCTCGCCACCACTC AGCACTCACCTCTTGGGC	7 9 1 9
C h r _ 1 7 _ 9	GGCAACTTCCTGAGACAGATCG GTAAAAACAACCCCTTCT	7 9 2 0
C h r _ 1 7 _ 10	TCAACTGTATTTTCATCAGAGAG ATGTGGCTTTCCCAAGACA	7 9 2 1
C h r _ 1 7 _ 11	GTTTCCCTCATGTTCCCCCAAG TTCTGTGTCAGGTGAAGCTG	7 9 2 2
C h r _ 1 8 _ 1	TTAACCCATCTCTACCCGTCCT GTGTCAAGAACGGAGGCT	7 9 2 3
C h r _ 1 8 _ 2	CTGCCCAAAATAGAAACCGAGG TTCTCCGTGACCTACATC	7 9 2 4

10

20

30

40

50

【表 3 - 9】

C h r _ 1 8 _ 3	TTCCTTTGTCAGTAACAGCGGGA ACATGAAGCCGCCACTCT	7 9 2 5
C h r _ 1 8 _ 4	TGGTTTGGCCAGTTCAGACACCC AGCCAAATTGCCCTCTCA	7 9 2 6
C h r _ 1 8 _ 5	TAGTGCAGCTGGCTTTTGAGCCT GTTCCCGAATGTTTCAGAT	7 9 2 7
C h r _ 1 8 _ 6	AGGGTAATAGCACCAAGCTCTA GTCTACCCACCTCTCTGA	7 9 2 8
C h r _ 1 8 _ 7	CCGCATCTCTGGAGTAGGAATT GATCAGCCACCATATGGG	7 9 2 9
C h r _ 1 8 _ 8	CTATGAGCATACTGGGGAGGGA AACCTCTAAGCGGAACTT	7 9 3 0
C h r _ 1 8 _ 9	AAAAACCTGCAGGAAGGAGACC TGAATGCAACTGTGGGTC	7 9 3 1
C h r _ 1 8 _ 10	CAGGTGCTCCAAACCTTCCAGT CTATGTTGTAGATTGCAG	7 9 3 2
C h r _ 1 8 _ 11	GCCATACTAACCTACTTCTCCT TGAAGCTCTTGGCCCATC	7 9 3 3
C h r _ 1 9 _ 1	ACTGTGAGATAGCCCTCATCAT CTTCAACAGCGCCAACCG	7 9 3 4
C h r _ 1 9 _ 2	AGATACACGGTTCACAGACGCCA TGTGTTGTGGCTTCTGCA	7 9 3 5
C h r _ 1 9 _ 3	CACATCCTCTCACCTTTTCCGA AGGTTGTCAGCTCCTTCTC	7 9 3 6
C h r _ 1 9 _ 4	TCTGTCTCACCGGTCCCTTCAT TCCTAGGCAACTGTAGAT	7 9 3 7
C h r _ 1 9 _ 5	ATATCATGGTCTGTATCCCCCA GGTACCTTGACACAGGCC	7 9 3 8
C h r _ 1 9 _ 6	CTCTCCGCCTTTCTTTAGACCT GAGCATGCAGAATTCCGA	7 9 3 9
C h r _ 1 9 _ 7	AAGGCATTTTAAATGGGACAGCG TCCCATGCGTGACTTCTC	7 9 4 0
C h r _ 1 9 _ 8	TCTTTCTAACAGACGAACAGCC TACACCTACAACCCCGAG	7 9 4 1
C h r _ 1 9 _ 9	GTCCCAGCCCCAAAAGCATCTTG GGTAAGGATTTGGGATCA	7 9 4 2
C h r _ 1 9 _ 10	GTTGTTCTGGGGCCAGTGTTAGT TGCTCACATGTCCTGTCT	7 9 4 3
C h r _ 1 9 _ 11	AACATGCCTCTTAGTCTCTGGGC CATACCTTAGCCTTGTGC	7 9 4 4
C h r _ 2 0 _ 1	TAACCTCCAAAAGAGGTACCCA TTGGCGCTCAACCGAATT	7 9 4 5
C h r _ 2 0 _ 2	CTATATCTCCGACTATGCCTTC TTGGGCACTGCACTGCTG	7 9 4 6
C h r _ 2 0 _ 3	TCTAGATGGAAGCTGTATCCAA GGATGCTCCGGAATGTTG	7 9 4 7
C h r _ 2 0 _ 4	ATCTTCTCTGCCTGCCGCACTA GCTTCTTGGTGACTTCTC	7 9 4 8

10

20

30

40

50

【表 3 - 1 0】

C h r _ 2 0 _ 5	A T C G A G T T G T C G A G C C C C A T G A T T C G A C A C C A A G A T C C C A	7 9 4 9
C h r _ 2 0 _ 6	A G G T G C T T G T T T T A C T C T C T C C A G G T G A T G A T G C C A G G G A	7 9 5 0
C h r _ 2 0 _ 7	G T G C A C T G T C A G A T C T T G G A A A C G G C C A A A G G A T T T T T C C	7 9 5 1
C h r _ 2 0 _ 8	C A T T T T T G C A G G A G G C T G C T A A T T A A G G C T G A G G G C C A T C A	7 9 5 2
C h r _ 2 0 _ 9	T C A A T G G T A G A C T G G A G T A C C T T G C C A G G G C A G A G A A A A A	7 9 5 3
C h r _ 2 0 _ 1 0	C T C C T C C A G G A G C T G G C A G C A T C A A G A C C C C A C T T C G C T T	7 9 5 4
C h r _ 2 1 _ 1	A A A T A A T A G C A G G C G T T G A G A T G T C C C T T C C C C A G C A C T C	7 9 5 5
C h r _ 2 1 _ 2	A A G T C T G A C A G C A T C T G C T T G A A C T G A G G C A C A G T G A T G G	7 9 5 6
C h r _ 2 1 _ 3	A T T C G T G A T G G C G C T C A T T T C C A T A A A G G A C G A C A G G T C A	7 9 5 7
C h r _ 2 1 _ 4	G A A G A G T G A A T T C C C G C T T C T G C G C C A A C A T T C T G T T T C C	7 9 5 8
C h r _ 2 1 _ 5	A C A G G T G A A G T C T T T G C G T G C C T C C C T G T T G G A C T C A A A T	7 9 5 9
C h r _ 2 1 _ 6	T A A T G A T A T T C T G G C A C A A G G A G C A G A G C C C C T C T T C T T C	7 9 6 0
C h r _ 2 1 _ 7	A G A C C C A G C C T A C C T G C A T G A T C T C T T G T A C A G C T T T G C A	7 9 6 1
C h r _ 2 1 _ 8	T C A T G G A A C A T G G G C C T T G C A A A G G G G T C A A G A T C A C A A C	7 9 6 2
C h r _ 2 1 _ 9	G T C A A A A A G G T C C A A T C A G C T A G A G A C T A G G C C A G A C C C A	7 9 6 3
C h r _ 2 2 _ 1	T G T G A C C A C C C T A A A G G G A G G G C A G A A G C C G A G T C A C C C T	7 9 6 4
C h r _ 2 2 _ 2	A C G C C T C C A C C T G C T G C T A G G A C T C C C C T C C C A A A C A A A G	7 9 6 5
C h r _ 2 2 _ 3	C A C A G T C T A G A C C C T G A T G G G C G A T C T C A G T A G T G C T G T T	7 9 6 6
C h r _ 2 2 _ 4	C C T A T C A A C G T G C A A G T G G G A T T T G T C T C C A C T G G C T T T C	7 9 6 7
C h r _ 2 2 _ 5	G A A A A T C A T T C C C C A T T C T G C A G G A T C C G T T C C C C T G G C A	7 9 6 8
C h r _ 2 2 _ 6	A G T G G G A C A T A C C A A C T T G A T G A G G C A G T T G T G C G A G T T C	7 9 6 9
C h r _ 2 2 _ 7	G T A A A C A G C T G T C T T C T T A C C C T A C A G A T C A T T G G G C A G G	7 9 7 0
C h r _ 2 2 _ 8	C A G A A G G A T A C T A G A A T G G A A T G T C C T G C G T G A C G A A A G C	7 9 7 1
C h r _ 2 2 _ 9	A G T T C A C A T C T G A T T C T C C T A T G G C T G C T A G G C T C C A G G A	7 9 7 2

10

20

30

40

【 0 2 1 9】

有意には、去勢抵抗性前立腺癌の患者の血液の血漿部分から収集された c f D N A に対して、健常試料を使用した正規化対照と同一の解析を適用したときに、3つの突出した特徴が現れた（図 2 7 C）。第一に、すべての対照プローブは、ノイズの多いカウント挙動を示した。第二に、すべての A R プローブにわたるカウントは、有意に「1」の正常な値から約「5」の増幅された値へ上昇した。A R 遺伝子の増幅は、進行前立腺癌において一貫して観察された。第三に、T P 5 3 プローブカウントは、より密接にクラスター化し、「2」の所期の値よりはるかに「1」に近い平均値を有した。これは、腫瘍組織に由来した循環 D N A 部分におけるコピー数減少によって、T P 5 3 の一方または両方のアレルの

50

不活性化を反映した可能性がある。

【0220】

これらのデータは、本発明の方法が3つの重要な核型分析の態様を有することを示した。すなわち、本明細書に記載されるこれらの方法は、一般染色体異数性、特異的な標的遺伝子のコピー増加、及び同一の特異的な標的遺伝子におけるコピー減少を検出する。さらにこれらの結果は、3つすべてのこれらのゲノム異常が癌に発生することが多いので、本明細書に記載される方法及びプラットフォームが精密治療の使用を導くことが可能であることを示す。

【0221】

健常対照（茶色の点）と比較した、去勢抵抗性前立腺癌の患者試料（青色の点）についての一般染色体異数性を測定した（図28）。この解析において、実験に使用された239個すべての対照プローブについてほぼ倍数性を、それらの染色体標的に従い順序付けした。いくつかの染色体（たとえば、染色体1及び染色体22）について、患者試料と対照試料との間で同様の倍数性の「2」の値を観察した。他の事例において、これら2個の試料間の偏差を観察した。これらの実験によって提供される、ゲノム倍数性全体についての情報の程度を、使用された対照プローブの数及び密度によって制限した。しかしながら、これらのデータは、均一な密度で染色体セグメントをすべて覆うより高密度なプローブパネルを、本発明の追加のユニーク特徴と併せて使用することが可能であることを示す。これらのような解析は、より高い分解能の、染色体コピー数のゲノムワイド測定を提供するであろう。

【0222】

これらのデータは、精密治療についてのガイドとして本発明の可能性をさらに強調する。たとえば、相同組換え修復においてゲノム欠損を有する腫瘍は、非常に不安定化した染色体倍数性を示すことが多く、またこれらのような腫瘍を含む患者は、PARP酵素複合体の阻害剤について良い候補である（Popova et al., Genome Biol. 2009; 10(11): R128 参照）。腫瘍の遺伝子型を探すほとんどのシーケンシングアッセイと異なり、本明細書に記載されるこれらのアッセイは、腫瘍の表現型を駆動する原因の変異が標的解析から隠されている場合でも、シーケンシングを使用して、この腫瘍の表現型として不安定化した染色体倍数性を検出する。

【0223】

固形腫瘍から脱落するDNA中の遺伝子減少を検出する能力は、特に有意である。腫瘍抑制遺伝子の変異及び欠失は、がんゲノム中で頻繁な事象であり、さらに腫瘍抑制遺伝子の生殖系列喪失に関する個体は、高齢期にがんにかかることに対してユニークに脆弱である。液体バイオプシーコピー数減少（CNL）アッセイの診断値は、その感度に正比例する。本明細書に記載される本発明についての検出の下限値を決定するために、実施例1に記載される不死化株を「ゲノム・イン・ア・ボトル」基準の細胞株、NA12878中で系統的に希釈した。一方の株は、ATMのシングルコピー欠失（単一アレルの減少）を有し、他方の株は、BRCA2のシングルコピー欠失を有する。実験は、純粋なNA12878の4個の対照試料、及び16%の各単一アレル欠失株を含有する8個のスパイクイン試料を含んだ（図29）。報告目的のために、これは、両アレル減少の8%のマイナーアレル頻度に対応する。特異的な遺伝子を標的とするすべてのプローブについての平均値、及び2個の追加の、欠失のない対照遺伝子を図29に示す。ATM及びBRCA2のコピー減少をスパイクイン試料のみに限定した。このデータの追加の計算上の処理は、2%のマイナーアレル頻度まで下げる、両アレル欠失の確信的なコピー減少呼び出しを明らかにした。この感度は、標準的な血液に基づく遺伝子型判定アッセイにおけるコピー減少呼び出しを常に有するため、本発明が特別な配慮を必要としないことを示した。

【0224】

これらのデータは、標的ゲノム遺伝子座のコピー数増加及びコピー数減少の両方を含む、コピー数の解析のためのプローブ特異的ゲノムキャプチャデータの使用を実証する。加えて、本明細書に記載される本発明は、単一のヌクレオチド変異体、単一のヌクレオチド

10

20

30

40

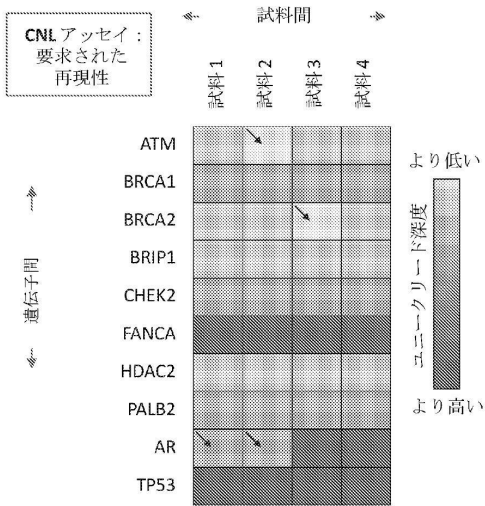
50

からなん千もの塩基対に及ぶ挿入及び欠失、ならびに異常な変異プロセスによる染色体再編成（PCT公開第WO2016/028316号、及び米国特許公開第2014-0274731号を参照）に起因する遺伝子融合を検出するために、感度の高い能力をもつことを示している。これらの変異プロセスのすべては、新生物がんへの正常な組織の形質転換に寄与する可能性があり、精密治療が出現し続けるなら、これらの疾患ゲノム特性の正確な診断は、次第に精密医薬の不可欠な機能となるであろう。

【図面】

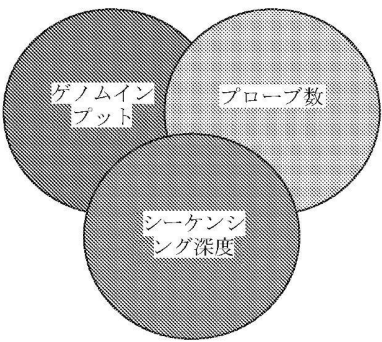
【図1】

【図1】



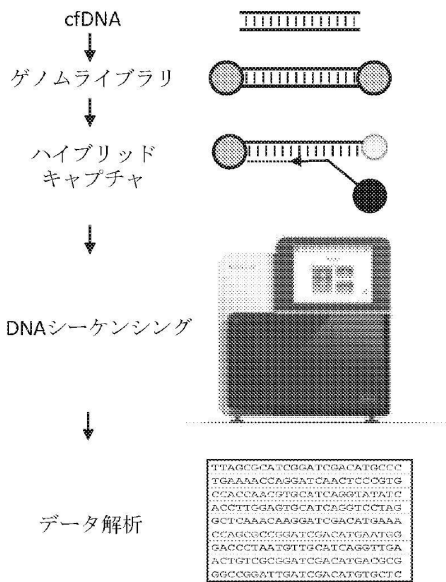
【図2】

【図2】



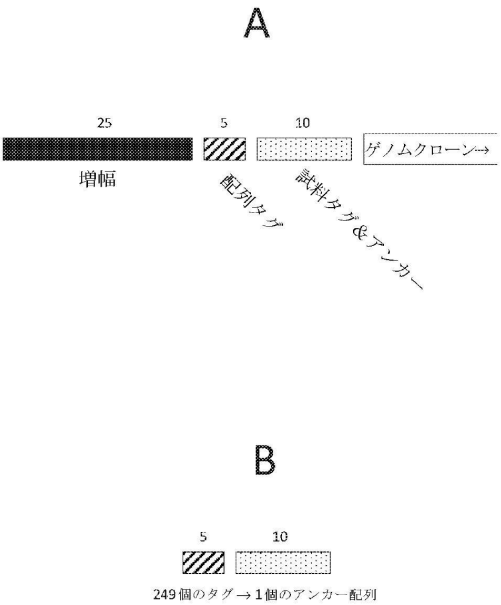
【図3】

【図3】



【図4-1】

【図4-1】



10

20

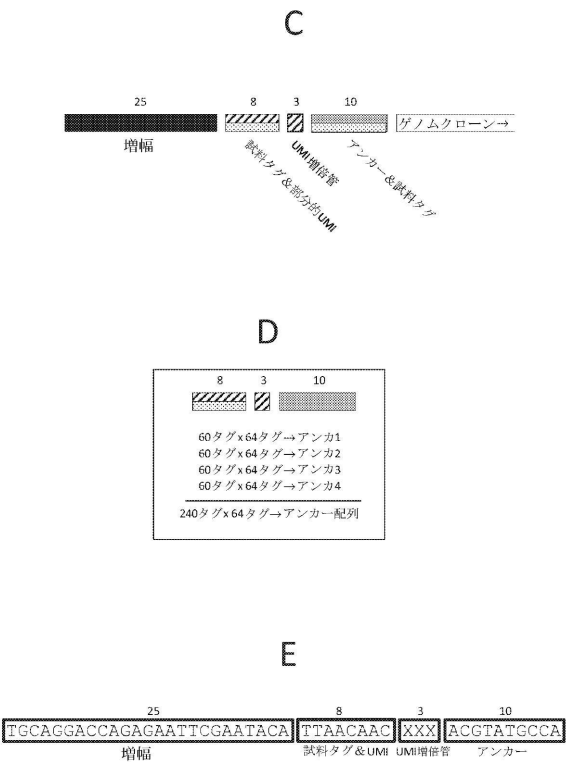
30

40

50

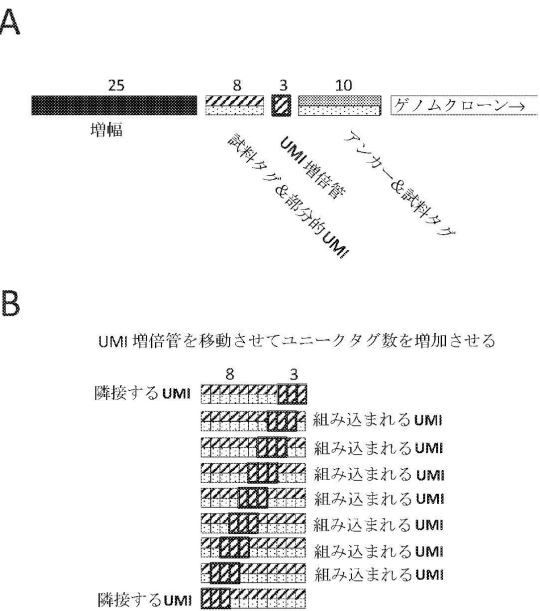
【図 4 - 2】

【図 4-2】



【図 5】

【図 5】

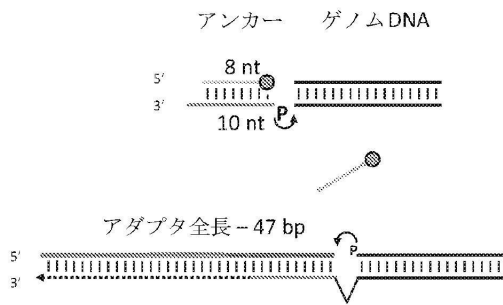


10

20

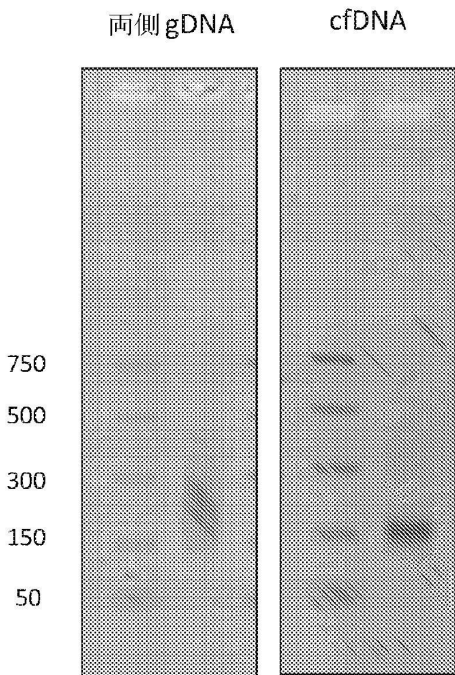
【図 6】

【図 6】



【図 7】

【図 7】



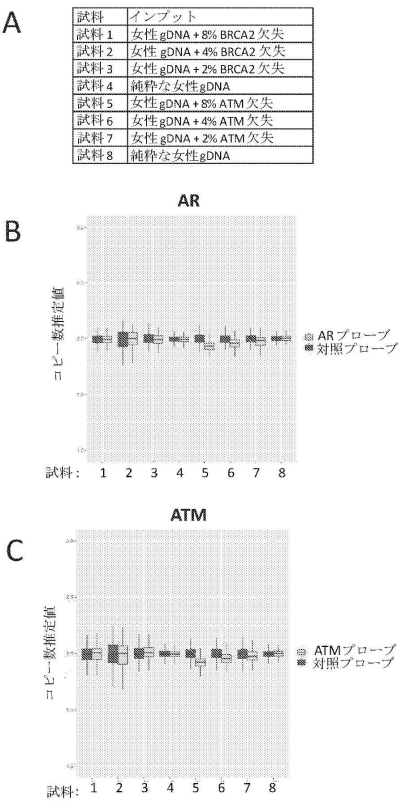
30

40

50

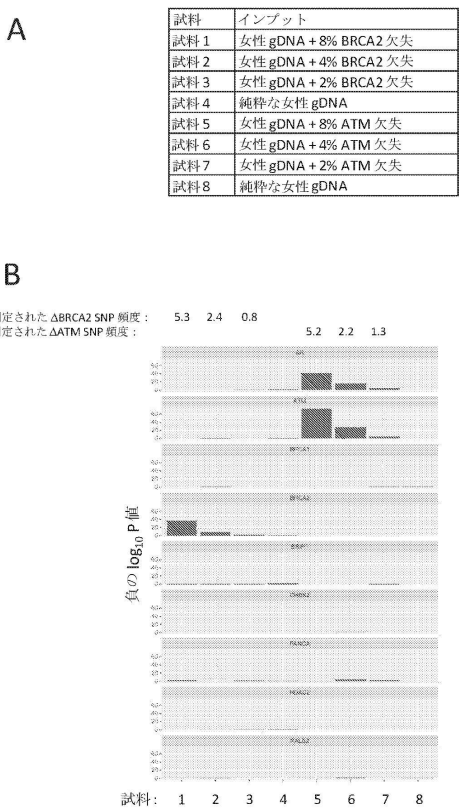
【図 8】

【図 8】



【図 9】

【図 9】

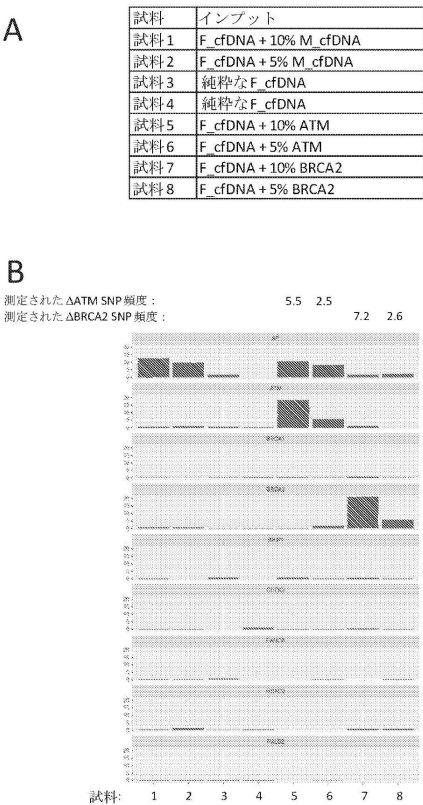


10

20

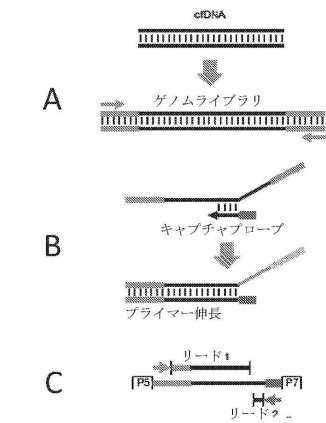
【図 10】

【図 10】

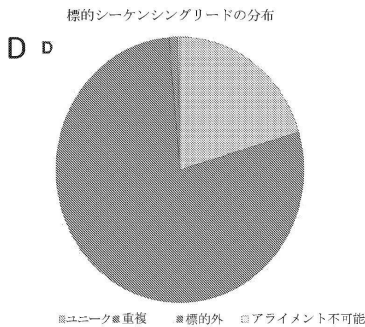


【図 11】

【図 11】



30



40

50

【図 1 2 B】

【図 12B】

[illegible]

【図 1 2 D】

【图 1-2D】

[illegible]

【圖 1 2 F】

【図 1 2 F】

[illegible][illegible]

【 図 1 2 H 】

【図 1 2 H】

[illegible]

グループ1	配列番号:	グループ2	配列番号:	グループ3	配列番号:
TCGACGACGACGAAATTCGAATA	720	TCGACGACGACGAAATTCGAATA	480	TCGACGACGACGAAATTCGAATA	720
CATGCTCTACGNNNTGCATCAGGT		CACACGACGAGNNNTGCATCAGGT		CATATCAGCAGCAGNNNTGCATCAGGT	
TCGACGACGACGAAATTCGAATA	721	TCGACGACGACGAAATTCGAATA	481	TCGACGACGACGAAATTCGAATA	721
CATTCGCTACGNNNTGCATCAGGT		CATATCTCATGNNNTGCATCAGGT		CAGTTTATCGNNNTGCATCAGGT	
TCGACGACGACGAAATTCGAATA	722	TCGACGACGACGAAATTCGAATA	482	TCGACGACGACGAAATTCGAATA	722
CAGAGCTCGANNNTGCATCAGGT		CAGATTTTCAGNNNTGCATCAGGT		CATTAAATCGCAGNNNTGCATCAGGT	
TCGACGACGACGAAATTCGAATA	723	TCGACGACGACGAAATTCGAATA	483	TCGACGACGACGAAATTCGAATA	723
CAGTCACTCTGNNNTGCATCAGGT		CAGCTAAACGCGNNNTGCATCAGGT		CAGTCACTCTGNNNTGCATCAGGT	
TCGACGACGACGAAATTCGAATA	724	TCGACGACGACGAAATTCGAATA	484	TCGACGACGACGAAATTCGAATA	724
CATTCGCTCTGNNNTGCATCAGGT		CATATCTAGAGNNNTGCATCAGGT		CAGAGATCTCGNNNTGCATCAGGT	
TCGACGACGACGAAATTCGAATA	725	TCGACGACGACGAAATTCGAATA	485	TCGACGACGACGAAATTCGAATA	725
CAGACCTAGCGNNNTGCATCAGGT		CAGCAGAACTGNNNTGCATCAGGT		CAGAGATCAAGNNNTGCATCAGGT	
TCGACGACGACGAAATTCGAATA	726	TCGACGACGACGAAATTCGAATA	486	TCGACGACGACGAAATTCGAATA	726
CAGCGAGCTCGNNNTGCATCAGGT		CAGCAAGTTATGNNNTGCATCAGGT		CAGANTTCAGNNNTGCATCAGGT	
TCGACGACGACGAAATTCGAATA	727	TCGACGACGACGAAATTCGAATA	487	TCGACGACGACGAAATTCGAATA	727
CAGAGATCTCGNNNTGCATCAGGT		CAGAGCTGAGNNNTGCATCAGGT		CAGCTAGATCTGNNNTGCATCAGGT	
TCGACGACGACGAAATTCGAATA	728	TCGACGACGACGAAATTCGAATA	488	TCGACGACGACGAAATTCGAATA	728
CATATCTGAGCAGNNNTGCATCAGGT		CATCTGCTGCTGNNNTGCATCAGGT		CATTTTCGCGNNNTGCATCAGGT	

【図 1 3 B】

【図 1 3 B】

ブルー4	配列番号:	ブルー5	配列番号:	ブルー6	配列番号:
TCGAGGACGAGAAATCGAATA	761	TCGAGGACGAGAAATCGAATA	1001	TCGAGGACGAGAAATCGAATA	1241
CAGATGCGNNNAGATGTGCGA		CAGACAGCGNNNAGATGTGCGA		CATGAGATTAANNAGATGTGCGA	
TCGAGGACGAGAAATCGAATA	762	TCGAGGACGAGAAATCGAATA	1002	TCGAGGACGAGAAATCGAATA	1242
CATCTATCGTNNNAGATGTGCGA		CAGCTCTATCGTNNNAGATGTGCGA		CATCTCTCTCTNNNAGATGTGCGA	
TCGAGGACGAGAAATCGAATA	763	TCGAGGACGAGAAATCGAATA	1003	TCGAGGACGAGAAATCGAATA	1243
CAGATGCGTNNNAGATGTGCGA		CAGACGCGGNNNAGATGTGCGA		CAGCGCGGNNNAGATGTGCGA	
TCGAGGACGAGAAATCGAATA	764	TCGAGGACGAGAAATCGAATA	1004	TCGAGGACGAGAAATCGAATA	1244
CAGAGATCGTNNNAGATGTGCGA		CAGATCTCTGNNNAGATGTGCGA		TCGAGGACGAGAAATCGAATA	
TCGAGGACGAGAAATCGAATA	765	TCGAGGACGAGAAATCGAATA	1005	TCGAGGACGAGAAATCGAATA	1245
CAGAGATCTAANNNAGATGTGCGA		CATTACTCTGNNNAGATGTGCGA		CAGGCGATAGANNNAGATGTGCGA	
TCGAGGACGAGAAATCGAATA	766	TCGAGGACGAGAAATCGAATA	1006	TCGAGGACGAGAAATCGAATA	1246
CAGATGCGTNNNAGATGTGCGA		CAGATGCGTNNNAGATGTGCGA		CAGGCGCGGNNNAGATGTGCGA	
TCGAGGACGAGAAATCGAATA	767	TCGAGGACGAGAAATCGAATA	1007	TCGAGGACGAGAAATCGAATA	1247
CAGATCTCTGNNNAGATGTGCGA		CAGCAGCTCTGNNNAGATGTGCGA		CAGCTTAATCTANNNAGATGTGCGA	
TCGAGGACGAGAAATCGAATA	768	TCGAGGACGAGAAATCGAATA	1008	TCGAGGACGAGAAATCGAATA	1248
CAGATCTCTGNNNAGATGTGCGA		CAGATGCGTNNNAGATGTGCGA		CATATTCAGTNNNAGATGTGCGA	
TCGAGGACGAGAAATCGAATA	769	TCGAGGACGAGAAATCGAATA	1009	TCGAGGACGAGAAATCGAATA	1249
CAGATCTCTGNNNAGATGTGCGA		CAGATCTCTGNNNAGATGTGCGA		CAGATCTCTGNNNAGATGTGCGA	
TCGAGGACGAGAAATCGAATA	770	TCGAGGACGAGAAATCGAATA	1010	TCGAGGACGAGAAATCGAATA	1250
CATCTCTGCGGNNAGATGTGCGA		CATCTCTGCGGNNAGATGTGCGA		CAATCTCTGNNNAGATGTGCGA	
TCGAGGACGAGAAATCGAATA	771	TCGAGGACGAGAAATCGAATA	1011	TCGAGGACGAGAAATCGAATA	1251
CATCTCTGCGGNNAGATGTGCGA		CAGATCTCTGNNNAGATGTGCGA		CATCTGCGGNNNAGATGTGCGA	
TCGAGGACGAGAAATCGAATA	772	TCGAGGACGAGAAATCGAATA	1012	TCGAGGACGAGAAATCGAATA	1252
CAGATGATCTAANNNAGATGTGCGA		CATCTCTGCGGNNAGATGTGCGA		CAGATGCGGNNNAGATGTGCGA	
TCGAGGACGAGAAATCGAATA	773	TCGAGGACGAGAAATCGAATA	1013	TCGAGGACGAGAAATCGAATA	1253
CAGACGCTCTGNNNAGATGTGCGA		CAGCTCTGAGNNNAGATGTGCGA		CAGCTCTCTANNNAGATGTGCGA	
TCGAGGACGAGAAATCGAATA	774	TCGAGGACGAGAAATCGAATA	1014	TCGAGGACGAGAAATCGAATA	1254
CAGCTCTCTGNNNAGATGTGCGA		CAGCTCTCTGNNNAGATGTGCGA		CAGCTCTCTGNNNAGATGTGCGA	
TCGAGGACGAGAAATCGAATA	775	TCGAGGACGAGAAATCGAATA	1015	TCGAGGACGAGAAATCGAATA	1255
CAGCTCTCTGNNNAGATGTGCGA		CAGCTCTCTGNNNAGATGTGCGA		CATCTCTAGGNNNAGATGTGCGA	
TCGAGGACGAGAAATCGAATA	776	TCGAGGACGAGAAATCGAATA	1016	TCGAGGACGAGAAATCGAATA	1256
CAGCTCTCTGNNNAGATGTGCGA		CAGCTCTCTGNNNAGATGTGCGA		CAGCTCTCTGNNNAGATGTGCGA	
TCGAGGACGAGAAATCGAATA	777	TCGAGGACGAGAAATCGAATA	1017	TCGAGGACGAGAAATCGAATA	1257
CAGATCTCTGNNNAGATGTGCGA		CATCTCTGNNNAGATGTGCGA		CATCTCTGNNNAGATGTGCGA	
TCGAGGACGAGAAATCGAATA	778	TCGAGGACGAGAAATCGAATA	1018	TCGAGGACGAGAAATCGAATA	1258
CAGATCTCTGNNNAGATGTGCGA		CAGATCTCTGNNNAGATGTGCGA		CAGCTCTGNNNAGATGTGCGA	
TCGAGGACGAGAAATCGAATA	779	TCGAGGACGAGAAATCGAATA	1019	TCGAGGACGAGAAATCGAATA	1259
CAGATCTCTGNNNAGATGTGCGA		CAGATCTCTGNNNAGATGTGCGA		CAGATCTCTGNNNAGATGTGCGA	
TCGAGGACGAGAAATCGAATA	780	TCGAGGACGAGAAATCGAATA	1020	TCGAGGACGAGAAATCGAATA	1260
CATCTCTGCGGNNAGATGTGCGA		CATCTCTGCGGNNAGATGTGCGA		CAGATCTGNNNAGATGTGCGA	
TCGAGGACGAGAAATCGAATA	781	TCGAGGACGAGAAATCGAATA	1021	TCGAGGACGAGAAATCGAATA	1261
CAGATCTCTGNNNAGATGTGCGA		CATCTCTGCGGNNAGATGTGCGA		CATCTCTGNNNAGATGTGCGA	
TCGAGGACGAGAAATCGAATA	782	TCGAGGACGAGAAATCGAATA	1022	TCGAGGACGAGAAATCGAATA	1262
CAGCTCTCTGNNNAGATGTGCGA		CAGCTCTGNNNAGATGTGCGA		CATCTCTAANNNAGATGTGCGA	
TCGAGGACGAGAAATCGAATA	783	TCGAGGACGAGAAATCGAATA	1023	TCGAGGACGAGAAATCGAATA	1263
CAGCTCTCTGNNNAGATGTGCGA		CATCTCTCTGNNNAGATGTGCGA		CATCTCTGNNNAGATGTGCGA	
TCGAGGACGAGAAATCGAATA	784	TCGAGGACGAGAAATCGAATA	1024	TCGAGGACGAGAAATCGAATA	1264
CAGCTCTCTGNNNAGATGTGCGA		CAGCTCTCTGNNNAGATGTGCGA		TCGAGGACGAGAAATCGAATA	
TCGAGGACGAGAAATCGAATA	785	TCGAGGACGAGAAATCGAATA	1025	TCGAGGACGAGAAATCGAATA	1265
CAGCTCTCTGNNNAGATGTGCGA		CATCTCTCTGNNNAGATGTGCGA		CATCTCTAANNNAGATGTGCGA	
TCGAGGACGAGAAATCGAATA	786	TCGAGGACGAGAAATCGAATA	1026	TCGAGGACGAGAAATCGAATA	1266
CAGCTCTCTGNNNAGATGTGCGA		CATCTCTCTGNNNAGATGTGCGA		CATCTCTCTGNNNAGATGTGCGA	
TCGAGGACGAGAAATCGAATA	787	TCGAGGACGAGAAATCGAATA	1027	TCGAGGACGAGAAATCGAATA	1267
CAGATCTCTGNNNAGATGTGCGA		CAGCTCTGCGGNNAGATGTGCGA		CAGCAGCTCTGNNNAGATGTGCGA	
TCGAGGACGAGAAATCGAATA	788	TCGAGGACGAGAAATCGAATA	1028	TCGAGGACGAGAAATCGAATA	1268
CAGATCTCTGNNNAGATGTGCGA		CATTAATCTCTGNNNAGATGTGCGA		CATTAATCTGNNNAGATGTGCGA	
TCGAGGACGAGAAATCGAATA	789	TCGAGGACGAGAAATCGAATA	1029	TCGAGGACGAGAAATCGAATA	1269
CAGATCTCTGNNNAGATGTGCGA		CAGATCTCTGNNNAGATGTGCGA		CAGATCTCTGNNNAGATGTGCGA	
TCGAGGACGAGAAATCGAATA	790	TCGAGGACGAGAAATCGAATA	1030	TCGAGGACGAGAAATCGAATA	1270
CAGCGGCTGCTNNNCTGAGGGTAC		CAGCTCTCTGNNNAGATGTGCGA		CATCTCTCTGNNNAGATGTGCGA	
TCGAGGACGAGAAATCGAATA	791	TCGAGGACGAGAAATCGAATA	1031	TCGAGGACGAGAAATCGAATA	1271
CATCTCTCTGNNNAGATGTGCGA		CATCTCTCTGNNNAGATGTGCGA		CAAGACGAGGNNNAGATGTGCGA	
TCGAGGACGAGAAATCGAATA	792	TCGAGGACGAGAAATCGAATA	1032	TCGAGGACGAGAAATCGAATA	1272

【圖 1 3 D】

【図 13 D】

[illegible]

【 図 1 3 F 】

【图 13 F】

[illegible]

【図 13 H】

【図 1 3 H】

グループ	配列番号:	グループ	配列番号:	グループ	配列番号:
TCGAGGCGACGAAGATTTCGATA CATGAGGAGTNNBNTGCATCAAGG	953	TCGAGGCGACGAAGATTTCGATA CACCGGAGTNNBNTGCATCAAGG	1193	TCGAGGACGACGAAGATTTCGATA CAAGGAGTTCNNBNTGCATCAAGG	1433
TCGAGGACGACGAAGATTTCGATA CATGAGGAGTNNBNTGCATCAAGG	954	TCGAGGACGACGAAGATTTCGATA CAATTCGAGTNNBNTGCATCAAGG	1194	TCGAGGACGACGAAGATTTCGATA CAATTCGAGTNNBNTGCATCAAGG	1434
TCGAGGACGACGAAGATTTCGATA CATGACGAGTNNBNTGCATCAAGG	955	TCGAGGACGACGAAGATTTCGATA CAATTTCAGGNNBNTGCATCAAGG	1195	TCGAGGACGACGAAGATTTCGATA CAATTCGAGTNNBNTGCATCAAGG	1435
TCGAGGACGACGAAGATTTCGATA CAAGGTCGCGNNBNTGCATCAAGG	956	TCGAGGACGACGAAGATTTCGATA CAATTCGAGTNNBNTGCATCAAGG	1196	TCGAGGACGACGAAGATTTCGATA CAATTCGAGTNNBNTGCATCAAGG	1436
TCGAGGACGACGAAGATTTCGATA CAATTCGAGTNNBNTGCATCAAGG	957	TCGAGGACGACGAAGATTTCGATA CAATTCGAGTNNBNTGCATCAAGG	1197	TCGAGGACGACGAAGATTTCGATA CAATTCGAGTNNBNTGCATCAAGG	1437
TCGAGGACGACGAAGATTTCGATA CAATTCGAGTNNBNTGCATCAAGG	958	TCGAGGACGACGAAGATTTCGATA CAATTCGAGTNNBNTGCATCAAGG	1198	TCGAGGACGACGAAGATTTCGATA CAATTCGAGTNNBNTGCATCAAGG	1438
TCGAGGACGACGAAGATTTCGATA CAATTCGAGTNNBNTGCATCAAGG	959	TCGAGGACGACGAAGATTTCGATA CAATTCGAGTNNBNTGCATCAAGG	1199	TCGAGGACGACGAAGATTTCGATA CAATTCGAGTNNBNTGCATCAAGG	1439
TCGAGGACGACGAAGATTTCGATA CATGTCGAGTNNBNTGCATCAAGG	960	TCGAGGACGACGAAGATTTCGATA CAATTCGAGTNNBNTGCATCAAGG	1200	TCGAGGACGACGAAGATTTCGATA CAATTCGAGTNNBNTGCATCAAGG	1440
TCGAGGACGACGAAGATTTCGATA CATGTCGAGTNNBNTGCATCAAGG	961	TCGAGGACGACGAAGATTTCGATA CATGTCGAGTNNBNTGCATCAAGG	1201	TCGAGGACGACGAAGATTTCGATA CATGTCGAGTNNBNTGCATCAAGG	1441
TCGAGGACGACGAAGATTTCGATA CATGTCGAGTNNBNTGCATCAAGG	962	TCGAGGACGACGAAGATTTCGATA CATGTCGAGTNNBNTGCATCAAGG	1202	TCGAGGACGACGAAGATTTCGATA CATGTCGAGTNNBNTGCATCAAGG	1442
TCGAGGACGACGAAGATTTCGATA CATGTCGAGTNNBNTGCATCAAGG	963	TCGAGGACGACGAAGATTTCGATA CATGTCGAGTNNBNTGCATCAAGG	1203	TCGAGGACGACGAAGATTTCGATA CATGTCGAGTNNBNTGCATCAAGG	1443
TCGAGGACGACGAAGATTTCGATA CATGTCGAGTNNBNTGCATCAAGG	964	TCGAGGACGACGAAGATTTCGATA CATGTCGAGTNNBNTGCATCAAGG	1204	TCGAGGACGACGAAGATTTCGATA CATGTCGAGTNNBNTGCATCAAGG	1444
TCGAGGACGACGAAGATTTCGATA CATGTCGAGTNNBNTGCATCAAGG	965	TCGAGGACGACGAAGATTTCGATA CATGTCGAGTNNBNTGCATCAAGG	1205	TCGAGGACGACGAAGATTTCGATA CATGTCGAGTNNBNTGCATCAAGG	1445
TCGAGGACGACGAAGATTTCGATA CATGTCGAGTNNBNTGCATCAAGG	966	TCGAGGACGACGAAGATTTCGATA CATGTCGAGTNNBNTGCATCAAGG	1206	TCGAGGACGACGAAGATTTCGATA CATGTCGAGTNNBNTGCATCAAGG	1446
TCGAGGACGACGAAGATTTCGATA CATGTCGAGTNNBNTGCATCAAGG	967	TCGAGGACGACGAAGATTTCGATA CATGTCGAGTNNBNTGCATCAAGG	1207	TCGAGGACGACGAAGATTTCGATA CATGTCGAGTNNBNTGCATCAAGG	1447
TCGAGGACGACGAAGATTTCGATA CATGTCGAGTNNBNTGCATCAAGG	968	TCGAGGACGACGAAGATTTCGATA CATGTCGAGTNNBNTGCATCAAGG	1208	TCGAGGACGACGAAGATTTCGATA CATGTCGAGTNNBNTGCATCAAGG	1448

【 図 1 4 F 】

【図 1 4 F】

[illegible]

【図 1 4 H】

【図 14H】

[illegible]

【図 1 4 I】

【図 1 4 I】

ルール7	配列番号:	ルール8	配列番号:	ルール9	配列番号:
TCGAGGACAGAGATTGGAATA	1489	TCGAGGACAGAGATTGGAATA	1490	TCGAGGACAGAGATTGGAATA	1491
CAAGTGTGNNNTGCAAGGT		CAAGTGTGNNNTGCAAGGT		CAAGTGTGNNNTGCAAGGT	
TCGAGGACAGAGATTGGAATA	1492	TCGAGGACAGAGATTGGAATA	1493	TCGAGGACAGAGATTGGAATA	1494
CAAGTGTGNNNTGCAAGGT		CAAGTGTGNNNTGCAAGGT		CAAGTGTGNNNTGCAAGGT	
TCGAGGACAGAGATTGGAATA	1495	TCGAGGACAGAGATTGGAATA	1496	TCGAGGACAGAGATTGGAATA	1497
CAAGTGTGNNNTGCAAGGT		CAAGTGTGNNNTGCAAGGT		CAAGTGTGNNNTGCAAGGT	
TCGAGGACAGAGATTGGAATA	1498	TCGAGGACAGAGATTGGAATA	1499	TCGAGGACAGAGATTGGAATA	1500
CAAGTGTGNNNTGCAAGGT		CAAGTGTGNNNTGCAAGGT		CAAGTGTGNNNTGCAAGGT	

【図 1 5 A】

【図 1 5 A】

ルール10	配列番号:	ルール11	配列番号:	ルール12	配列番号:
TCGAGGACAGAGATTGGAATA	1501	TCGAGGACAGAGATTGGAATA	1502	TCGAGGACAGAGATTGGAATA	1503
CAAGTGTGNNNTGCAAGGT		CAAGTGTGNNNTGCAAGGT		CAAGTGTGNNNTGCAAGGT	
TCGAGGACAGAGATTGGAATA	1504	TCGAGGACAGAGATTGGAATA	1505	TCGAGGACAGAGATTGGAATA	1506
CAAGTGTGNNNTGCAAGGT		CAAGTGTGNNNTGCAAGGT		CAAGTGTGNNNTGCAAGGT	
TCGAGGACAGAGATTGGAATA	1507	TCGAGGACAGAGATTGGAATA	1508	TCGAGGACAGAGATTGGAATA	1509
CAAGTGTGNNNTGCAAGGT		CAAGTGTGNNNTGCAAGGT		CAAGTGTGNNNTGCAAGGT	
TCGAGGACAGAGATTGGAATA	1510	TCGAGGACAGAGATTGGAATA	1511	TCGAGGACAGAGATTGGAATA	1512
CAAGTGTGNNNTGCAAGGT		CAAGTGTGNNNTGCAAGGT		CAAGTGTGNNNTGCAAGGT	

10

【図 1 5 B】

【図 1 5 B】

ルール10	配列番号:	ルール11	配列番号:	ルール12	配列番号:
TCGAGGACAGAGATTGGAATA	1513	TCGAGGACAGAGATTGGAATA	1514	TCGAGGACAGAGATTGGAATA	1515
CAAGTGTGNNNTGCAAGGT		CAAGTGTGNNNTGCAAGGT		CAAGTGTGNNNTGCAAGGT	
TCGAGGACAGAGATTGGAATA	1516	TCGAGGACAGAGATTGGAATA	1517	TCGAGGACAGAGATTGGAATA	1518
CAAGTGTGNNNTGCAAGGT		CAAGTGTGNNNTGCAAGGT		CAAGTGTGNNNTGCAAGGT	
TCGAGGACAGAGATTGGAATA	1519	TCGAGGACAGAGATTGGAATA	1520	TCGAGGACAGAGATTGGAATA	1521
CAAGTGTGNNNTGCAAGGT		CAAGTGTGNNNTGCAAGGT		CAAGTGTGNNNTGCAAGGT	
TCGAGGACAGAGATTGGAATA	1522	TCGAGGACAGAGATTGGAATA	1523	TCGAGGACAGAGATTGGAATA	1524
CAAGTGTGNNNTGCAAGGT		CAAGTGTGNNNTGCAAGGT		CAAGTGTGNNNTGCAAGGT	

【図 1 5 C - 1】

【図 1 5 C】

ルール10	配列番号:	ルール11	配列番号:	ルール12	配列番号:
TCGAGGACAGAGATTGGAATA	1525	TCGAGGACAGAGATTGGAATA	1526	TCGAGGACAGAGATTGGAATA	1527
CAAGTGTGNNNTGCAAGGT		CAAGTGTGNNNTGCAAGGT		CAAGTGTGNNNTGCAAGGT	
TCGAGGACAGAGATTGGAATA	1528	TCGAGGACAGAGATTGGAATA	1529	TCGAGGACAGAGATTGGAATA	1530
CAAGTGTGNNNTGCAAGGT		CAAGTGTGNNNTGCAAGGT		CAAGTGTGNNNTGCAAGGT	
TCGAGGACAGAGATTGGAATA	1531	TCGAGGACAGAGATTGGAATA	1532	TCGAGGACAGAGATTGGAATA	1533
CAAGTGTGNNNTGCAAGGT		CAAGTGTGNNNTGCAAGGT		CAAGTGTGNNNTGCAAGGT	
TCGAGGACAGAGATTGGAATA	1534	TCGAGGACAGAGATTGGAATA	1535	TCGAGGACAGAGATTGGAATA	1536
CAAGTGTGNNNTGCAAGGT		CAAGTGTGNNNTGCAAGGT		CAAGTGTGNNNTGCAAGGT	

30

40

50

【 図 1 5 C - 2 】

【図 15 C】

[illegible]

【図 1 5 D】

【図 15 D】

[illegible]

【 図 1 5 E 】

【図 15 E】

[illegible]

【 図 1 5 F 】

【図 15 F】

[illegible]

【図15G】

【図15G】

配列番号	配列番号	配列番号	配列番号
2393	2394	2395	2396
2397	2398	2399	2400
2401	2402	2403	2404
2405	2406	2407	2408

【図16A】

【図16A】

配列番号	配列番号	配列番号	配列番号
2409	2410	2411	2412
2413	2414	2415	2416
2417	2418	2419	2420
2422	2423	2424	2425
2427	2428	2429	2430
2433	2434	2435	2436
2439	2440	2441	2442
2445	2446	2447	2448
2451	2452	2453	2454
2457	2458	2459	2460
2463	2464	2465	2466
2471	2472	2473	2474
2477	2478	2479	2480
2483	2484	2485	2486
2491	2492	2493	2494
2497	2498	2499	2500
2503	2504	2505	2506
2509	2510	2511	2512
2513	2514	2515	2516
2519	2520	2521	2522

10

20

【図16B】

【図16B】

配列番号	配列番号	配列番号	配列番号
2523	2524	2525	2526
2529	2530	2531	2532
2533	2534	2535	2536
2539	2540	2541	2542
2543	2544	2545	2546
2549	2550	2551	2552
2553	2554	2555	2556
2559	2560	2561	2562
2563	2564	2565	2566
2569	2570	2571	2572
2573	2574	2575	2576
2579	2580	2581	2582
2583	2584	2585	2586
2591	2592	2593	2594
2597	2598	2599	2600
2603	2604	2605	2606
2609	2610	2611	2612
2613	2614	2615	2616
2619	2620	2621	2622
2623	2624	2625	2626
2629	2630	2631	2632
2633	2634	2635	2636
2639	2640	2641	2642
2643	2644	2645	2646
2649	2650	2651	2652
2653	2654	2655	2656
2659	2660	2661	2662
2663	2664	2665	2666
2669	2670	2671	2672
2673	2674	2675	2676
2679	2680	2681	2682
2683	2684	2685	2686
2691	2692	2693	2694
2697	2698	2699	2700
2703	2704	2705	2706
2709	2710	2711	2712
2713	2714	2715	2716
2719	2720	2721	2722
2723	2724	2725	2726
2729	2730	2731	2732
2733	2734	2735	2736
2739	2740	2741	2742
2743	2744	2745	2746
2749	2750	2751	2752
2753	2754	2755	2756
2759	2760	2761	2762
2763	2764	2765	2766
2769	2770	2771	2772
2773	2774	2775	2776
2779	2780	2781	2782
2783	2784	2785	2786
2791	2792	2793	2794
2797	2798	2799	2800
2803	2804	2805	2806
2809	2810	2811	2812
2813	2814	2815	2816
2819	2820	2821	2822
2823	2824	2825	2826
2829	2830	2831	2832
2833	2834	2835	2836
2839	2840	2841	2842
2843	2844	2845	2846
2849	2850	2851	2852
2853	2854	2855	2856
2859	2860	2861	2862
2863	2864	2865	2866
2869	2870	2871	2872
2873	2874	2875	2876
2879	2880	2881	2882
2883	2884	2885	2886
2891	2892	2893	2894
2897	2898	2899	2900
2903	2904	2905	2906
2909	2910	2911	2912
2913	2914	2915	2916
2919	2920	2921	2922

30

40

【図16C】

【図16C】

配列番号	配列番号	配列番号	配列番号
2923	2924	2925	2926
2929	2930	2931	2932
2933	2934	2935	2936
2939	2940	2941	2942
2943	2944	2945	2946
2949	2950	2951	2952
2953	2954	2955	2956
2959	2960	2961	2962
2963	2964	2965	2966
2969	2970	2971	2972
2973	2974	2975	2976
2979	2980	2981	2982
2983	2984	2985	2986
2991	2992	2993	2994
2997	2998	2999	3000
3003	3004	3005	3006
3009	3010	3011	3012
3013	3014	3015	3016
3019	3020	3021	3022
3023	3024	3025	3026
3029	3030	3031	3032
3033	3034	3035	3036
3039	3040	3041	3042
3043	3044	3045	3046
3049	3050	3051	3052
3053	3054	3055	3056
3059	3060	3061	3062
3063	3064	3065	3066
3069	3070	3071	3072
3073	3074	3075	3076
3079	3080	3081	3082
3083	3084	3085	3086
3091	3092	3093	3094
3097	3098	3099	3100
3103	3104	3105	3106
3109	3110	3111	3112
3113	3114	3115	3116
3119	3120	3121	3122
3123	3124	3125	3126
3129	3130	3131	3132
3133	3134	3135	3136
3139	3140	3141	3142
3143	3144	3145	3146
3149	3150	3151	3152
3153	3154	3155	3156
3159	3160	3161	3162

50

JP 7217224 B2 2023.2.2

JP 7217224 B2 2023.2.2

JP 7217224 B2 2023.2.2

JP 7217224 B2 2023.2.2

JP 7217224 B2 2023.2.2

JP 7217224 B2 2023.2.2

【図 17 A】

【図 17 A】

[illegible]

10

20

【 図 1 7 C 】

【図 17 C】

[illegible]

30

40

【図 17 E】

【図 17 E】

配列番号	配列番号	配列番号	配列番号
7340	7341	7342	7343
7344	7345	7346	7347
7348	7349	7350	7351
7352	7353	7354	7355
7356	7357	7358	7359
7360	7361	7362	7363
7364	7365	7366	7367
7368	7369	7370	7371
7372	7373	7374	7375
7376	7377	7378	7379
7380	7381	7382	7383
7384	7385	7386	7387
7388	7389	7390	7391
7392	7393	7394	7395
7396	7397	7398	7399
7400	7401	7402	7403
7404	7405	7406	7407
7408	7409	7410	7411
7412	7413	7414	7415
7416	7417	7418	7419
7420	7421	7422	7423
7424	7425	7426	7427
7428	7429	7430	7431
7432	7433	7434	7435
7436	7437	7438	7439
7440	7441	7442	7443
7444	7445	7446	7447
7448	7449	7450	7451
7452	7453	7454	7455
7456	7457	7458	7459
7460	7461	7462	7463
7464	7465	7466	7467
7468	7469	7470	7471
7472	7473	7474	7475
7476	7477	7478	7479
7480	7481	7482	7483
7484	7485	7486	7487
7488	7489	7490	7491
7492	7493	7494	7495
7496	7497	7498	7499

【 図 1 7 G 】

【図 17 G】

[illegible]

【図 17H】

【図 17H】

ルール16	配列番号:	ルール17	配列番号:	ルール18	配列番号:
TCGAGGACGAGAAATTCGATACAG AGGATGANNNGGATCAGGT	4339	TCGAGGACGAGAAATTCGATACAG TGTTCGNNNTGATCAGGT	4340	TCGAGGACGAGAAATTCGATACAG TAGCTTCGNNNTGATCAGGT	4341
TCGAGGACGAGAAATTCGATACAG AGGATGANNNGGATCAGGT	4342	TCGAGGACGAGAAATTCGATACAG TGTTCGNNNTGATCAGGT	4343	TCGAGGACGAGAAATTCGATACAG TAGCTTCGNNNTGATCAGGT	4344
TCGAGGACGAGAAATTCGATACAG AGGATGANNNGGATCAGGT	4345	TCGAGGACGAGAAATTCGATACAG TGTTCGNNNTGATCAGGT	4346	TCGAGGACGAGAAATTCGATACAG TAGCTTCGNNNTGATCAGGT	4347
TCGAGGACGAGAAATTCGATACAG AGGATGANNNGGATCAGGT	4348	TCGAGGACGAGAAATTCGATACAG TGTTCGNNNTGATCAGGT	4349	TCGAGGACGAGAAATTCGATACAG TAGCTTCGNNNTGATCAGGT	4350

【図 18A】

【図 18A】

ルール19	配列番号:	ルール20	配列番号:	ルール21	配列番号:
TCGAGGACGAGAAATTCGATACAG AGGATGANNNGGATCAGGT	4351	TCGAGGACGAGAAATTCGATACAG TGTTCGNNNTGATCAGGT	4352	TCGAGGACGAGAAATTCGATACAG TAGCTTCGNNNTGATCAGGT	4353
TCGAGGACGAGAAATTCGATACAG AGGATGANNNGGATCAGGT	4354	TCGAGGACGAGAAATTCGATACAG TGTTCGNNNTGATCAGGT	4355	TCGAGGACGAGAAATTCGATACAG TAGCTTCGNNNTGATCAGGT	4356
TCGAGGACGAGAAATTCGATACAG AGGATGANNNGGATCAGGT	4357	TCGAGGACGAGAAATTCGATACAG TGTTCGNNNTGATCAGGT	4358	TCGAGGACGAGAAATTCGATACAG TAGCTTCGNNNTGATCAGGT	4359
TCGAGGACGAGAAATTCGATACAG AGGATGANNNGGATCAGGT	4360	TCGAGGACGAGAAATTCGATACAG TGTTCGNNNTGATCAGGT	4361	TCGAGGACGAGAAATTCGATACAG TAGCTTCGNNNTGATCAGGT	4362

10

【図 18B】

【図 18B】

ルール19	配列番号:	ルール20	配列番号:	ルール21	配列番号:
TCGAGGACGAGAAATTCGATACAG AGGATGANNNGGATCAGGT	4363	TCGAGGACGAGAAATTCGATACAG TGTTCGNNNTGATCAGGT	4364	TCGAGGACGAGAAATTCGATACAG TAGCTTCGNNNTGATCAGGT	4365
TCGAGGACGAGAAATTCGATACAG AGGATGANNNGGATCAGGT	4366	TCGAGGACGAGAAATTCGATACAG TGTTCGNNNTGATCAGGT	4367	TCGAGGACGAGAAATTCGATACAG TAGCTTCGNNNTGATCAGGT	4368
TCGAGGACGAGAAATTCGATACAG AGGATGANNNGGATCAGGT	4369	TCGAGGACGAGAAATTCGATACAG TGTTCGNNNTGATCAGGT	4370	TCGAGGACGAGAAATTCGATACAG TAGCTTCGNNNTGATCAGGT	4371
TCGAGGACGAGAAATTCGATACAG AGGATGANNNGGATCAGGT	4372	TCGAGGACGAGAAATTCGATACAG TGTTCGNNNTGATCAGGT	4373	TCGAGGACGAGAAATTCGATACAG TAGCTTCGNNNTGATCAGGT	4374

【図 18C】

【図 18C】

ルール19	配列番号:	ルール20	配列番号:	ルール21	配列番号:
TCGAGGACGAGAAATTCGATACAG AGGATGANNNGGATCAGGT	4385	TCGAGGACGAGAAATTCGATACAG TGTTCGNNNTGATCAGGT	4386	TCGAGGACGAGAAATTCGATACAG TAGCTTCGNNNTGATCAGGT	4387
TCGAGGACGAGAAATTCGATACAG AGGATGANNNGGATCAGGT	4388	TCGAGGACGAGAAATTCGATACAG TGTTCGNNNTGATCAGGT	4389	TCGAGGACGAGAAATTCGATACAG TAGCTTCGNNNTGATCAGGT	4390
TCGAGGACGAGAAATTCGATACAG AGGATGANNNGGATCAGGT	4391	TCGAGGACGAGAAATTCGATACAG TGTTCGNNNTGATCAGGT	4392	TCGAGGACGAGAAATTCGATACAG TAGCTTCGNNNTGATCAGGT	4393
TCGAGGACGAGAAATTCGATACAG AGGATGANNNGGATCAGGT	4394	TCGAGGACGAGAAATTCGATACAG TGTTCGNNNTGATCAGGT	4395	TCGAGGACGAGAAATTCGATACAG TAGCTTCGNNNTGATCAGGT	4396

30

40

50

【 図 1 8 E 】

【図 18 E】

[illegible]

【 図 1 8 G 】

【図 18 G】

グループ19	配列番号:	グループ20	配列番号:	グループ21	配列番号:
TCGAGACAGCAAGATTTCGATAT	4327	TCGAGACAGCAAGATTTCGATAT	4767	TCGAGACAGCAAGATTTCGATAT	5007
CACTACGACGAGGTTTCATGAGT	4328	CACTACGACGAGGTTTCATGAGT	4768	CACTACGACGAGGTTTCATGAGT	5008
CATTGCGAGGNNNTTCATCGAGT	4329	CATTGCGAGGNNNTTCATCGAGT	4769	CATTGCGAGGNNNTTCATCGAGT	5009
TCGAGACAGCAAGATTTCGATAT	4330	TCGAGACAGCAAGATTTCGATAT	4770	TCGAGACAGCAAGATTTCGATAT	5010
CAAGTCGACGAGGTTTCATGAGT	4331	CAAGTCGACGAGGTTTCATGAGT	4771	CAAGTCGACGAGGTTTCATGAGT	5011
CAGAGTCGAGGNNNTTCATCGAGT	4332	CAGAGTCGAGGNNNTTCATCGAGT	4772	CAGAGTCGAGGNNNTTCATCGAGT	5012
TCGAGACAGCAAGATTTCGATAT	4333	TCGAGACAGCAAGATTTCGATAT	4773	TCGAGACAGCAAGATTTCGATAT	5013
CAGTCGACGAGGTTTCATGAGT	4334	CAGTCGACGAGGTTTCATGAGT	4774	CAGTCGACGAGGTTTCATGAGT	5014
CAGAGTCGAGGNNNTTCATCGAGT	4335	CAGAGTCGAGGNNNTTCATCGAGT	4775	CAGAGTCGAGGNNNTTCATCGAGT	5015
TCGAGACAGCAAGATTTCGATAT	4336	TCGAGACAGCAAGATTTCGATAT	4776	TCGAGACAGCAAGATTTCGATAT	5016
CAAGTCGACGAGGTTTCATGAGT	4337	CAAGTCGACGAGGTTTCATGAGT	4777	CAAGTCGACGAGGTTTCATGAGT	5017
CAGAGTCGAGGNNNTTCATCGAGT	4338	CAGAGTCGAGGNNNTTCATCGAGT	4778	CAGAGTCGAGGNNNTTCATCGAGT	5018
TCGAGACAGCAAGATTTCGATAT	4339	TCGAGACAGCAAGATTTCGATAT	4779	TCGAGACAGCAAGATTTCGATAT	5019
CAAGTCGACGAGGTTTCATGAGT	4340	CAAGTCGACGAGGTTTCATGAGT	4780	CAAGTCGACGAGGTTTCATGAGT	5020
CAGAGTCGAGGNNNTTCATCGAGT	4341	CAGAGTCGAGGNNNTTCATCGAGT	4781	CAGAGTCGAGGNNNTTCATCGAGT	5021
TCGAGACAGCAAGATTTCGATAT	4342	TCGAGACAGCAAGATTTCGATAT	4782	TCGAGACAGCAAGATTTCGATAT	5022
CAAGTCGACGAGGTTTCATGAGT	4343	CAAGTCGACGAGGTTTCATGAGT	4783	CAAGTCGACGAGGTTTCATGAGT	5023
CAGAGTCGAGGNNNTTCATCGAGT	4344	CAGAGTCGAGGNNNTTCATCGAGT	4784	CAGAGTCGAGGNNNTTCATCGAGT	5024
TCGAGACAGCAAGATTTCGATAT	4345	TCGAGACAGCAAGATTTCGATAT	4785	TCGAGACAGCAAGATTTCGATAT	5025
CAAGTCGACGAGGTTTCATGAGT	4346	CAAGTCGACGAGGTTTCATGAGT	4786	CAAGTCGACGAGGTTTCATGAGT	5026
CAGAGTCGAGGNNNTTCATCGAGT	4347	CAGAGTCGAGGNNNTTCATCGAGT	4787	CAGAGTCGAGGNNNTTCATCGAGT	5027
TCGAGACAGCAAGATTTCGATAT	4348	TCGAGACAGCAAGATTTCGATAT	4788	TCGAGACAGCAAGATTTCGATAT	5028
CAAGTCGACGAGGTTTCATGAGT	4349	CAAGTCGACGAGGTTTCATGAGT	4789	CAAGTCGACGAGGTTTCATGAGT	5029
CAGAGTCGAGGNNNTTCATCGAGT	4350	CAGAGTCGAGGNNNTTCATCGAGT	4790	CAGAGTCGAGGNNNTTCATCGAGT	5030
TCGAGACAGCAAGATTTCGATAT	4351	TCGAGACAGCAAGATTTCGATAT	4791	TCGAGACAGCAAGATTTCGATAT	5031
CAAGTCGACGAGGTTTCATGAGT	4352	CAAGTCGACGAGGTTTCATGAGT	4792	CAAGTCGACGAGGTTTCATGAGT	5032
CAGAGTCGAGGNNNTTCATCGAGT	4353	CAGAGTCGAGGNNNTTCATCGAGT	4793	CAGAGTCGAGGNNNTTCATCGAGT	5033
TCGAGACAGCAAGATTTCGATAT	4354	TCGAGACAGCAAGATTTCGATAT	4794	TCGAGACAGCAAGATTTCGATAT	5034
CAAGTCGACGAGGTTTCATGAGT	4355	CAAGTCGACGAGGTTTCATGAGT	4795	CAAGTCGACGAGGTTTCATGAGT	5035
CAGAGTCGAGGNNNTTCATCGAGT	4356	CAGAGTCGAGGNNNTTCATCGAGT	4796	CAGAGTCGAGGNNNTTCATCGAGT	5036
TCGAGACAGCAAGATTTCGATAT	4357	TCGAGACAGCAAGATTTCGATAT	4797	TCGAGACAGCAAGATTTCGATAT	5037
CAAGTCGACGAGGTTTCATGAGT	4358	CAAGTCGACGAGGTTTCATGAGT	4798	CAAGTCGACGAGGTTTCATGAGT	5038
CAGAGTCGAGGNNNTTCATCGAGT	4359	CAGAGTCGAGGNNNTTCATCGAGT	4799	CAGAGTCGAGGNNNTTCATCGAGT	5039

【 図 1 9 A 】

【図 19 A】

[illegible]

【 図 1 9 C 】

【图 19C】

[illegible]

【図 19H】

【図 19H】

ルール22	配列番号:	ルール23	配列番号:	ルール24	配列番号:
TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5280	TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5281	TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5282
TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5283	TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5284	TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5285
TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5286	TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5287	TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5288
TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5289	TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5290	TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5291

【図 20A】

【図 20A】

ルール25	配列番号:	ルール26	配列番号:	ルール27	配列番号:
TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5292	TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5293	TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5294
TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5295	TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5296	TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5297
TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5298	TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5299	TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5300
TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5301	TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5302	TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5303

10

【図 20B】

【図 20B】

ルール25	配列番号:	ルール26	配列番号:	ルール27	配列番号:
TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5304	TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5305	TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5306
TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5307	TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5308	TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5309
TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5310	TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5311	TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5312
TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5313	TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5314	TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5315

【図 20C】

【図 20C】

ルール25	配列番号:	ルール26	配列番号:	ルール27	配列番号:
TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5316	TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5317	TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5318
TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5319	TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5320	TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5321
TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5322	TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5323	TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5324
TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5325	TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5326	TCGAGGACACAGAAATTCGAATACAT CACTGCTTNNINAGTATGCCA	5327

20

30

40

50

20

20

[illegible]

40

40

[illegible]

【図 2 1 A】

【図 2 1 A】

[illegible]

【圖 2 1 C】

【図 2 1 C】

[illegible]

【 図 2 1 D 】

【図 2 1 D】

[illegible]

【 図 2 1 E 】

【図 2 1 E】

[illegible]

【 図 2 1 F 】

【図 2 1 F】

プール 29	配列番号	プール 29	配列番号	プール 30	配列番号
TCGACGACGAGAAATTGCGATG	6454	TCGACGACGAGAAATTGCGATG	6894	TCGACGACGAGAAATTGCGATG	7134
CATACAGATCGGNNNGTCAGCGT		CATGAGCTCTNNNGTCAGCGT	6895	CAAAATCGGNNNGTCAGCGT	7135
TCGACGACGAGAAATTGCGATG	6455	TCGACGACGAGAAATTGCGATG	6896	TCGACGACGAGAAATTGCGATG	7136
CAGGCTGACGANNNGTCAGCGT		CAAGTCTGACGANNNGTCAGCGT	6897	CAAACTATCTNNNGTCAGCGT	7137
TCGACGACGAGAAATTGCGATG	6456	TCGACGACGAGAAATTGCGATG	6898	TCGACGACGAGAAATTGCGATG	7138
TCGACGACGAGAAATTGCGATG	6457	TCGACGACGAGAAATTGCGATG	6899	TCGACGACGAGAAATTGCGATG	7139
TCGACGACGAGAAATTGCGATG	6458	TCGACGACGAGAAATTGCGATG	6900	TCGACGACGAGAAATTGCGATG	7140
TCGACGACGAGAAATTGCGATG	6459	TCGACGACGAGAAATTGCGATG	6901	TCGACGACGAGAAATTGCGATG	7141
TCGACGACGAGAAATTGCGATG	6460	TCGACGACGAGAAATTGCGATG	6902	TCGACGACGAGAAATTGCGATG	7142
TCGACGACGAGAAATTGCGATG	6461	TCGACGACGAGAAATTGCGATG	6903	TCGACGACGAGAAATTGCGATG	7143
TCGACGACGAGAAATTGCGATG	6462	TCGACGACGAGAAATTGCGATG	6904	TCGACGACGAGAAATTGCGATG	7144
TCGACGACGAGAAATTGCGATG	6463	TCGACGACGAGAAATTGCGATG	6905	TCGACGACGAGAAATTGCGATG	7145
TCGACGACGAGAAATTGCGATG	6464	TCGACGACGAGAAATTGCGATG	6906	TCGACGACGAGAAATTGCGATG	7146
TCGACGACGAGAAATTGCGATG	6465	TCGACGACGAGAAATTGCGATG	6907	TCGACGACGAGAAATTGCGATG	7147
TCGACGACGAGAAATTGCGATG	6466	TCGACGACGAGAAATTGCGATG	6908	TCGACGACGAGAAATTGCGATG	7148
TCGACGACGAGAAATTGCGATG	6467	TCGACGACGAGAAATTGCGATG	6909	TCGACGACGAGAAATTGCGATG	7149
TCGACGACGAGAAATTGCGATG	6468	TCGACGACGAGAAATTGCGATG	6910	TCGACGACGAGAAATTGCGATG	7150
TCGACGACGAGAAATTGCGATG	6469	TCGACGACGAGAAATTGCGATG	6911	TCGACGACGAGAAATTGCGATG	7151
TCGACGACGAGAAATTGCGATG	6470	TCGACGACGAGAAATTGCGATG	6912	TCGACGACGAGAAATTGCGATG	7152
TCGACGACGAGAAATTGCGATG	6471	TCGACGACGAGAAATTGCGATG	6913	TCGACGACGAGAAATTGCGATG	7153
TCGACGACGAGAAATTGCGATG	6472	TCGACGACGAGAAATTGCGATG	6914	TCGACGACGAGAAATTGCGATG	7154
TCGACGACGAGAAATTGCGATG	6473	TCGACGACGAGAAATTGCGATG	6915	TCGACGACGAGAAATTGCGATG	7155
TCGACGACGAGAAATTGCGATG	6474	TCGACGACGAGAAATTGCGATG	6916	TCGACGACGAGAAATTGCGATG	7156
TCGACGACGAGAAATTGCGATG	6475	TCGACGACGAGAAATTGCGATG	6917	TCGACGACGAGAAATTGCGATG	7157
TCGACGACGAGAAATTGCGATG	6476	TCGACGACGAGAAATTGCGATG	6918	TCGACGACGAGAAATTGCGATG	7158
TCGACGACGAGAAATTGCGATG	6477	TCGACGACGAGAAATTGCGATG	6919	TCGACGACGAGAAATTGCGATG	7159
TCGACGACGAGAAATTGCGATG	6478	TCGACGACGAGAAATTGCGATG	6920	TCGACGACGAGAAATTGCGATG	7160
TCGACGACGAGAAATTGCGATG	6479	TCGACGACGAGAAATTGCGATG	6921	TCGACGACGAGAAATTGCGATG	7161
TCGACGACGAGAAATTGCGATG	6480	TCGACGACGAGAAATTGCGATG	6922	TCGACGACGAGAAATTGCGATG	7162
TCGACGACGAGAAATTGCGATG	6481	TCGACGACGAGAAATTGCGATG	6923	TCGACGACGAGAAATTGCGATG	7163
TCGACGACGAGAAATTGCGATG	6482	TCGACGACGAGAAATTGCGATG	6924	TCGACGACGAGAAATTGCGATG	7164
TCGACGACGAGAAATTGCGATG	6483	TCGACGACGAGAAATTGCGATG	6925	TCGACGACGAGAAATTGCGATG	7165
TCGACGACGAGAAATTGCGATG	6484	TCGACGACGAGAAATTGCGATG	6926	TCGACGACGAGAAATTGCGATG	7166
TCGACGACGAGAAATTGCGATG	6485	TCGACGACGAGAAATTGCGATG	6927	TCGACGACGAGAAATTGCGATG	7167
TCGACGACGAGAAATTGCGATG	6486	TCGACGACGAGAAATTGCGATG	6928	TCGACGACGAGAAATTGCGATG	7168

【 ㇴ 2 1 G 】

【図 2 1 G】

[illegible]

【図 2 2 A】

【図 2 2 A】

[illegible]

【 ㊦ 2 2 C 】

【図 2 2 C】

[illegible]

【 図 2 2 E 】

【图 2 2 E】

配列番号	配列番号	配列番号	
TGGAGGACAGAGGATTCGGAAT	7308	TGGAGGACAGAGGATTCGGAAT	7348
ATATCGTCGNNNNTCAGGTTAC		CAATCTCGGCGNNNATCGGTTAC	
TGGAGGACAGAGGATTCGGAAT	7309	TGGAGGACAGAGGATTCGGAAT	7349
ATATCTCGTCGNNNNTCAGGTTAC		CAATGTCCTGNNNNTCAGGTTAC	
TGGAGGACAGAGGATTCGGAAT	7310	TGGAGGACAGAGGATTCGGAAT	7350
ATAGGAGCGGNNNNTCAGGTTAC		CAAGACATCGGNNNNTCAGGTTAC	
TGGAGGACAGAGGATTCGGAAT	7311	TGGAGGACAGAGGATTCGGAAT	7351
ATACCTGTCGNNNNTCAGGTTAC		CAGCGGACATCGNNNNTCAGGTTAC	
TGGAGGACAGAGGATTCGGAAT	7312	TGGAGGACAGAGGATTCGGAAT	7352
ATATCTCGTCGNNNNTCAGGTTAC		CAGCAATGTCGNNNNTCAGGTTAC	
TGGAGGACAGAGGATTCGGAAT	7313	TGGAGGACAGAGGATTCGGAAT	7353
ATAGTACCGGNNNNTCAGGTTAC		CAGCGAGGACGNNNNTCAGGTTAC	
TGGAGGACAGAGGATTCGGAAT	7314	TGGAGGACAGAGGATTCGGAAT	7354
ATAGGACATCGNNNNTCAGGTTAC		CAAGACATCGNNNNTCAGGTTAC	
TGGAGGACAGAGGATTCGGAAT	7315	TGGAGGACAGAGGATTCGGAAT	7355
ATACCTGTCGNNNNTCAGGTTAC		CAGCATCTCGNNNNTCAGGTTAC	
TGGAGGACAGAGGATTCGGAAT	7316	TGGAGGACAGAGGATTCGGAAT	7356
ATAGGACATCGNNNNTCAGGTTAC		CAGCATCTCGNNNNTCAGGTTAC	
TGGAGGACAGAGGATTCGGAAT	7317	TGGAGGACAGAGGATTCGGAAT	7357
ATGTCGTCGNNNNTCAGGTTAC		CAGTATCTCGNNNNTCAGGTTAC	
TGGAGGACAGAGGATTCGGAAT	7318	TGGAGGACAGAGGATTCGGAAT	7358
ATAGGACATCGNNNNTCAGGTTAC		CAGGCTGACGNNNNTCAGGTTAC	
TGGAGGACAGAGGATTCGGAAT	7319	TGGAGGACAGAGGATTCGGAAT	7359
ATAGATTCGNNNNTCAGGTTAC		CAGCTGTCGNNNNTCAGGTTAC	
TGGAGGACAGAGGATTCGGAAT	7320	TGGAGGACAGAGGATTCGGAAT	7360
ATATCTCGTCGNNNNTCAGGTTAC		CAGCGAGGACGNNNNTCAGGTTAC	
TGGAGGACAGAGGATTCGGAAT	7321	TGGAGGACAGAGGATTCGGAAT	7361
ATATCTCGTCGNNNNTCAGGTTAC		CAGGTAAAGATTCNNNNTCAGGTTAC	
TGGAGGACAGAGGATTCGGAAT	7322	TGGAGGACAGAGGATTCGGAAT	7362
ATAGATTCGNNNNTCAGGTTAC		CAGTATCTCGNNNNTCAGGTTAC	
TGGAGGACAGAGGATTCGGAAT	7323	TGGAGGACAGAGGATTCGGAAT	7363
ATCTCGTCGNNNNTCAGGTTAC		CAGGAAAAATGNNNNTCAGGTTAC	
TGGAGGACAGAGGATTCGGAAT	7324	TGGAGGACAGAGGATTCGGAAT	7364
ATAACTCGGCGNNNNTCAGGTTAC		CAGGCTGGTGGNNNNTCAGGTTAC	
TGGAGGACAGAGGATTCGGAAT	7325	TGGAGGACAGAGGATTCGGAAT	7365
ATAGGACATCGNNNNTCAGGTTAC		CAGGCTGACGNNNNTCAGGTTAC	
TGGAGGACAGAGGATTCGGAAT	7326	TGGAGGACAGAGGATTCGGAAT	7366
ATCATATCGNNNNTCAGGTTAC		CAGCTCTCGGNNNNTCAGGTTAC	
TGGAGGACAGAGGATTCGGAAT	7327	TGGAGGACAGAGGATTCGGAAT	7367
ATATCTCGTCGNNNNTCAGGTTAC		CATCTATCGGNNNNTCAGGTTAC	
TGGAGGACAGAGGATTCGGAAT	7328	TGGAGGACAGAGGATTCGGAAT	7368
ATATTAATTCGNNNNTCAGGTTAC		CAAGGCGGCTGNNNNTCAGGTTAC	
TGGAGGACAGAGGATTCGGAAT	7329	TGGAGGACAGAGGATTCGGAAT	7369
ATATCTCGTCGNNNNTCAGGTTAC		CAGGCTGACGNNNNTCAGGTTAC	
TGGAGGACAGAGGATTCGGAAT	7330	TGGAGGACAGAGGATTCGGAAT	7370
ATATGCGGCGNNNNTCAGGTTAC		CATGCTCGGCGNNNNTCAGGTTAC	
TGGAGGACAGAGGATTCGGAAT	7331	TGGAGGACAGAGGATTCGGAAT	7371
ATAGGCGGCGNNNNTCAGGTTAC		CATCGGCGGNNNNTCAGGTTAC	
TGGAGGACAGAGGATTCGGAAT	7332	TGGAGGACAGAGGATTCGGAAT	7372
ATAGCTAAAGNNNNTCAGGTTAC		CATCATCTCGGNNNNTCAGGTTAC	
TGGAGGACAGAGGATTCGGAAT	7333	TGGAGGACAGAGGATTCGGAAT	7373
ATACCTCGGCGNNNNTCAGGTTAC		CAGTCACTCGGNNNNTCAGGTTAC	
TGGAGGACAGAGGATTCGGAAT	7334	TGGAGGACAGAGGATTCGGAAT	7374
ATACCTATGTCGNNNNTCAGGTTAC		CAGCGGTCGNNNNTCAGGTTAC	
TGGAGGACAGAGGATTCGGAAT	7335	TGGAGGACAGAGGATTCGGAAT	7375
ATATGCTGTCGNNNNTCAGGTTAC		CATATCTCGGNNNNTCAGGTTAC	
TGGAGGACAGAGGATTCGGAAT	7336	TGGAGGACAGAGGATTCGGAAT	7376
ATAGCTCGGAGNNNNTCAGGTTAC		CAGGCGGAGNNNNTCAGGTTAC	
TGGAGGACAGAGGATTCGGAAT	7337	TGGAGGACAGAGGATTCGGAAT	7377
ATAGGACATCGNNNNTCAGGTTAC		CAGGACGACGNNNNTCAGGTTAC	
TGGAGGACAGAGGATTCGGAAT	7338	TGGAGGACAGAGGATTCGGAAT	7378
ATAGGACATCGNNNNTCAGGTTAC		CAGCTCTCGGNNNNTCAGGTTAC	
TGGAGGACAGAGGATTCGGAAT	7339	TGGAGGACAGAGGATTCGGAAT	7379
ATCGGATTCGNNNNTCAGGTTAC		CAGATTCGCTGNNNNTCAGGTTAC	
TGGAGGACAGAGGATTCGGAAT	7340	TGGAGGACAGAGGATTCGGAAT	7380
ATATCTCGTCGNNNNTCAGGTTAC		CAGTATCTCGNNNNTCAGGTTAC	

【 図 2 2 G 】

【図 2 2 G】

グループ31	配列番号	グループ32	配列番号
TGCGAGCCAGCAAGATTCGAAATAC	7614	TGCGAGCCAGCAAGATTCGAAATAC	7614
AATAAATCCNNINNGATCGATG		CATTCGGATCCNNINNGATCGATG	
TGCGAGCCAGCAAGATTCGAAATAC	7315	TGCGAGCCAGCAAGATTCGAAATAC	7615
ATTCTGGTCGNNINNGATCGATG		CATTCGAGCCAGNNINNGATCGATG	
TGCGAGCCAGCAAGATTCGAAATAC	7316	TGCGAGCCAGCAAGATTCGAAATAC	7616
ATGTCAGCATCCNNINNGATCGATG		CATGAGCATTCNNINNGATCGATG	
TGCGAGCCAGCAAGATTCGAAATAC	7317	TGCGAGCCAGCAAGATTCGAAATAC	7617
AGTATGTCGNNINNGATCGATG		CATGTCAGCCAGCAAGATTCGAAATAC	
TGCGAGCCAGCAAGATTCGAAATAC	7318	TGCGAGCCAGCAAGATTCGAAATAC	7618
ACGCCATCATCCNNINNGATCGATG		CAGCGGCTGCNNINNGATCGATG	
TGCGAGCCAGCAAGATTCGAAATAC	7319	TGCGAGCCAGCAAGATTCGAAATAC	7619
AGTATGTCGNNINNGATCGATG		TGCGAGCCAGCAAGATTCGAAATAC	
TGCGAGCCAGCAAGATTCGAAATAC	7380	TGCGAGCCAGCAAGATTCGAAATAC	7620
AAATATCAATGATTCGATCGATG		CAGATCTCCANNINNGATCGATG	
TGCGAGCCAGCAAGATTCGAAATAC	7381	TGCGAGCCAGCAAGATTCGAAATAC	7621
AGTATGTCGNNINNGATCGATG		CAGCATGTCAGNNINNGATCGATG	
TGCGAGCCAGCAAGATTCGAAATAC	7382	TGCGAGCCAGCAAGATTCGAAATAC	7622
AGTCTGACATCCNNINNGATCGATG		CGATCTGTCGNNINNGATCGATG	
TGCGAGCCAGCAAGATTCGAAATAC	7383	TGCGAGCCAGCAAGATTCGAAATAC	7623
AGTTCTAGTCGNNINNGATCGATG		TGCGAGCCAGCAAGATTCGAAATAC	
TGCGAGCCAGCAAGATTCGAAATAC	7384	TGCGAGCCAGCAAGATTCGAAATAC	7624
ATCTCCAGAGCCNNINNGATCGATG		CAGAGCCATCCNNINNGATCGATG	
TGCGAGCCAGCAAGATTCGAAATAC	7385	TGCGAGCCAGCAAGATTCGAAATAC	7625
AAATATCCNNINNGATCGATG		CAGCGGCTGCNNINNGATCGATG	
TGCGAGCCAGCAAGATTCGAAATAC	7386	TGCGAGCCAGCAAGATTCGAAATAC	7626
AAATGTCGNNINNGATCGATG		CAGATTCGNNINNGATCGATG	
TGCGAGCCAGCAAGATTCGAAATAC	7387	TGCGAGCCAGCAAGATTCGAAATAC	7627
AGATTTCTCCNNINNGATCGATG		CATGACGACATCCNNINNGATCGATG	
TGCGAGCCAGCAAGATTCGAAATAC	7388	TGCGAGCCAGCAAGATTCGAAATAC	7628
AGATGTCGNNINNGATCGATG		CATGACGATTCNNINNGATCGATG	
TGCGAGCCAGCAAGATTCGAAATAC	7389	TGCGAGCCAGCAAGATTCGAAATAC	7629
ATATCAACGACNNINNGATCGATG		CAGAAATCTGCNNINNGATCGATG	
TGCGAGCCAGCAAGATTCGAAATAC	7390	TGCGAGCCAGCAAGATTCGAAATAC	7630
AGGTCGACGNNINNGATCGATG		CATCTGCTGTCNNINNGATCGATG	
TGCGAGCCAGCAAGATTCGAAATAC	7391	TGCGAGCCAGCAAGATTCGAAATAC	7631
AGATATGACNNINNGATCGATG		CATCTGGAATCCNNINNGATCGATG	
TGCGAGCCAGCAAGATTCGAAATAC	7392	TGCGAGCCAGCAAGATTCGAAATAC	7632
AAATATCCNNINNGATCGATG		CACGAGCTGCNNINNGATCGATG	
TGCGAGCCAGCAAGATTCGAAATAC	7393	TGCGAGCCAGCAAGATTCGAAATAC	7633
ATATCAATATCCNNINNGATCGATG		CAGTCTGTCANNINNGATCGATG	
TGCGAGCCAGCAAGATTCGAAATAC	7394	TGCGAGCCAGCAAGATTCGAAATAC	7634
AGACGACGNNINNGATCGATG		CACATATCTCCNNINNGATCGATG	
TGCGAGCCAGCAAGATTCGAAATAC	7395	TGCGAGCCAGCAAGATTCGAAATAC	7635
AGTATGTCGNNINNGATCGATG		CATGAGCCAGCAAGATTCGAAATAC	
TGCGAGCCAGCAAGATTCGAAATAC	7396	TGCGAGCCAGCAAGATTCGAAATAC	7636
ACGCGACGNNINNGATCGATG		CATATATCTCCNNINNGATCGATG	
TGCGAGCCAGCAAGATTCGAAATAC	7397	TGCGAGCCAGCAAGATTCGAAATAC	7637
AGGTCGACGNNINNGATCGATG		CAGGCTGTCNNINNGATCGATG	
TGCGAGCCAGCAAGATTCGAAATAC	7398	TGCGAGCCAGCAAGATTCGAAATAC	7638
AGCTTATCTCCNNINNGATCGATG		CAGCTATGACNNINNGATCGATG	
TGCGAGCCAGCAAGATTCGAAATAC	7400	TGCGAGCCAGCAAGATTCGAAATAC	7639
AGGCTGTCGNNINNGATCGATG		CATGTCAGCCAGCAAGATTCGAAATAC	
TGCGAGCCAGCAAGATTCGAAATAC	7401	TGCGAGCCAGCAAGATTCGAAATAC	7640
ACGCGACGNNINNGATCGATG		CAGAGGAGCAGNNINNGATCGATG	
TGCGAGCCAGCAAGATTCGAAATAC	7402	TGCGAGCCAGCAAGATTCGAAATAC	7641
AGTATGTCGNNINNGATCGATG		CATATGATCTCCNNINNGATCGATG	
TGCGAGCCAGCAAGATTCGAAATAC	7403	TGCGAGCCAGCAAGATTCGAAATAC	7642
AGTTCTGATCCNNINNGATCGATG		CATCTCTCATCCNNINNGATCGATG	
TGCGAGCCAGCAAGATTCGAAATAC	7404	TGCGAGCCAGCAAGATTCGAAATAC	7643
AGTATGTCGNNINNGATCGATG		CAGGATATCCNNINNGATCGATG	
TGCGAGCCAGCAAGATTCGAAATAC	7405	TGCGAGCCAGCAAGATTCGAAATAC	7644
AGTCTGTCNNINNGATCGATG		CATGCTGTCNNINNGATCGATG	
TGCGAGCCAGCAAGATTCGAAATAC	7406	TGCGAGCCAGCAAGATTCGAAATAC	7646
AGTATGTCGNNINNGATCGATG			

プール31	配列番号:	プール32	配列番号:
TCGCAGACACGAGAATTCGATATC	1341	TCGCAGACACGAGAATTCGATATC	7591
ATTCTTCGCGNNATCGATCGT		CACTGTGCTGNNATCGATCGT	
TCGCAGACACGAGAATTCGATATC	7342	TCGCAGACACGAGAATTCGATATC	7592
CACTGTGCTGNNATCGATCGT		CACGACGCGCGNNATCGATCGT	
TCGCAGACACGAGAATTCGATATC	1343	TCGCAGACACGAGAATTCGATATC	7593
CACTGTGCTGNNATCGATCGT		CACCTATGCTGNNATCGATCGT	
TCGCAGACACGAGAATTCGATATC	1344	TCGCAGACACGAGAATTCGATATC	7594
ACATGATTCGNNATCGATCGT		TCGCAGACACGAGAATTCGATATC	
TCGCAGACACGAGAATTCGATATC	1345	TCGCAGACACGAGAATTCGATATC	7595
ACACCGCTGNNATCGATCGT		CACCTTCCTGNNATCGATCGT	
TCGCAGACACGAGAATTCGATATC	1346	TCGCAGACACGAGAATTCGATATC	7596
TCGCAGACACGAGAATTCGATATC		TCGCAGACACGAGAATTCGATATC	
TCGCAGACACGAGAATTCGATATC	1347	TCGCAGACACGAGAATTCGATATC	7597
ACCTGATGCGNNATCGATCGT		CACACGACGAGAATTCGATATC	
TCGCAGACACGAGAATTCGATATC	1348	TCGCAGACACGAGAATTCGATATC	7598
TCCTGCTGNNATCGATCGT		TCGCAGACACGAGAATTCGATATC	
TCGCAGACACGAGAATTCGATATC	1349	TCGCAGACACGAGAATTCGATATC	7599
ACAGCTATGNNATCGATCGT		CATGCGCTACGNNATCGATCGT	
TCGCAGACACGAGAATTCGATATC	1350	TCGCAGACACGAGAATTCGATATC	7590
ACCTCATGNNATCGATCGT		CATCGCTGCTGNNATCGATCGT	
TCGCAGACACGAGAATTCGATATC	1351	TCGCAGACACGAGAATTCGATATC	7591
ACATATATGNNATCGATCGT		TCGCAGACACGAGAATTCGATATC	
TCGCAGACACGAGAATTCGATATC	1352	TCGCAGACACGAGAATTCGATATC	7592
TCGCAGACACGAGAATTCGATATC		CACCTACGCTGNNATCGATCGT	
TCGCAGACACGAGAATTCGATATC	1353	TCGCAGACACGAGAATTCGATATC	7593
CAATATGCTGNNATCGATCGT		TCGCAGACACGAGAATTCGATATC	
TCGCAGACACGAGAATTCGATATC	1354	TCGCAGACACGAGAATTCGATATC	7594
ATGCTGCTGCTGNNATCGATCGT		CACCTTCGCGNNATCGATCGT	
TCGCAGACACGAGAATTCGATATC	1355	TCGCAGACACGAGAATTCGATATC	7595
TCCTGCTGNNATCGATCGT		CATCTGCTGNNATCGATCGT	
TCGCAGACACGAGAATTCGATATC	1356	TCGCAGACACGAGAATTCGATATC	7596
ACCGACGCTGNNATCGATCGT		CACGACGCTATGNNATCGATCGT	
TCGCAGACACGAGAATTCGATATC	1357	TCGCAGACACGAGAATTCGATATC	7597
TCGCAGACACGAGAATTCGATATC		CACATATGCTGNNATCGATCGT	
TCGCAGACACGAGAATTCGATATC	1358	TCGCAGACACGAGAATTCGATATC	7598
CGAGCACTGNNATCGATCGT		TCGCAGACACGAGAATTCGATATC	
TCGCAGACACGAGAATTCGATATC	1359	TCGCAGACACGAGAATTCGATATC	7599
ACATGACGCTGNNATCGATCGT		CACGATGACGNNATCGATCGT	
TCGCAGACACGAGAATTCGATATC	1360	TCGCAGACACGAGAATTCGATATC	7600
AGGTCACGCGNNATCGATCGT		CACGCTGCTGNNATCGATCGT	
TCGCAGACACGAGAATTCGATATC	1361	TCGCAGACACGAGAATTCGATATC	7601
CAAGACGCGNNATCGATCGT		CATCGCGCTGNNATCGATCGT	
TCGCAGACACGAGAATTCGATATC	1362	TCGCAGACACGAGAATTCGATATC	7602
TCCTGCTGNNATCGATCGT		TCGCAGACACGAGAATTCGATATC	
TCGCAGACACGAGAATTCGATATC	1363	TCGCAGACACGAGAATTCGATATC	7603
ACCGACGCTGNNATCGATCGT		CATCGCTCTGNNATCGATCGT	
TCGCAGACACGAGAATTCGATATC	1364	TCGCAGACACGAGAATTCGATATC	7604
TCCTGCTGNNATCGATCGT		TCGCAGACACGAGAATTCGATATC	
TCGCAGACACGAGAATTCGATATC	1365	TCGCAGACACGAGAATTCGATATC	7605
ACCTGATGCTGNNATCGATCGT		TCGCAGACACGAGAATTCGATATC	
TCGCAGACACGAGAATTCGATATC	1366	TCGCAGACACGAGAATTCGATATC	7606
TCCTGCTGNNATCGATCGT		CACGCTGCTGNNATCGATCGT	
TCGCAGACACGAGAATTCGATATC	1367	TCGCAGACACGAGAATTCGATATC	7607
ATATCTGAGNNATCGATCGT		CAGCTTGGAGNNATCGATCGT	
TCGCAGACACGAGAATTCGATATC	1368	TCGCAGACACGAGAATTCGATATC	7608
ATATCATGCTGNNATCGATCGT		TCGCAGACACGAGAATTCGATATC	
TCGCAGACACGAGAATTCGATATC	1369	TCGCAGACACGAGAATTCGATATC	7609
ACATCATGCGNNATCGATCGT		TCGCAGACACGAGAATTCGATATC	
TCGCAGACACGAGAATTCGATATC	1370	TCGCAGACACGAGAATTCGATATC	7610
ATCATGCGNNATCGATCGT		CATATGCGCGNNATCGATCGT	
TCGCAGACACGAGAATTCGATATC	1371	TCGCAGACACGAGAATTCGATATC	7611
TCCTGCTGNNATCGATCGT		CATCTGCTGNNATCGATCGT	
TCGCAGACACGAGAATTCGATATC	1372	TCGCAGACACGAGAATTCGATATC	7612
ATATCTGCTGNNATCGATCGT		CATATGCTGCTGNNATCGATCGT	
TCGCAGACACGAGAATTCGATATC	1373	TCGCAGACACGAGAATTCGATATC	7613
TCCTGCTGNNATCGATCGT		CATATGCTGCTGNNATCGATCGT	

10

20

30

40

50

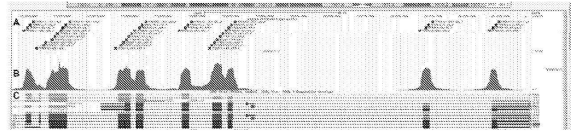
【図 2 2 H】

【図 2 2 H】

プローブ31	配列番号:	プローブ32	配列番号:
TGCAGGACCCAGAGAATTGGAATAC	7440	TGCAGGACCCAGAGAATTGGAATA	7680
AGGTCGGTCNNNTGCATCAGGT		CATTTCGCCANNNITGCATCAGGT	
TGCAGGACCCAGAGAATTGGAATAC	7441	TGCAGGACCCAGAGAATTGGAATA	7681
AAATATCGTANNNTGCATCAGGT		CAGACCATTCNNNTGCATCAGGT	
TGCAGGACCCAGAGAATTGGAATAC	7442	TGCAGGACCCAGAGAATTGGAATA	7682
AACCTCTCANNNTGCATCAGGT		CAGCCGCTAGNNNTGCATCAGGT	
TGCAGGACCCAGAGAATTGGAATAC	7443	TGCAGGACCCAGAGAATTGGAATA	7683
AGTGGTCNNNTGCATCAGGT		CATTAGCTGNNNTGCATCAGGT	
TGCAGGACCCAGAGAATTGGAATAC	7444	TGCAGGACCCAGAGAATTGGAATA	7684
ACTCGCCGNNNTGCATCAGGT		CATCTGATTTHNNITGCATCAGGT	
TGCAGGACCCAGAGAATTGGAATAC	7445	TGCAGGACCCAGAGAATTGGAATA	7685
ACAGGCTGNNNTGCATCAGGT		CATCATAGGNNNTGCATCAGGT	
TGCAGGACCCAGAGAATTGGAATAC	7446	TGCAGGACCCAGAGAATTGGAATA	7686
ATCGCTTACNNNTGCATCAGGT		CATGGAGGTANNNTGCATCAGGT	
TGCAGGACCCAGAGAATTGGAATAC	7447	TGCAGGACCCAGAGAATTGGAATA	7687
ACGGATTGNNNTGCATCAGGT		CAGTCTCTACNNNTGCATCAGGT	
TGCAGGACCCAGAGAATTGGAATAC	7448	TGCAGGACCCAGAGAATTGGAATA	7688
ACAAGCTGNNNTGCATCAGGT		CAGGCTGGTNNNTGCATCAGGT	

【図 2 3】

【図 2 3】

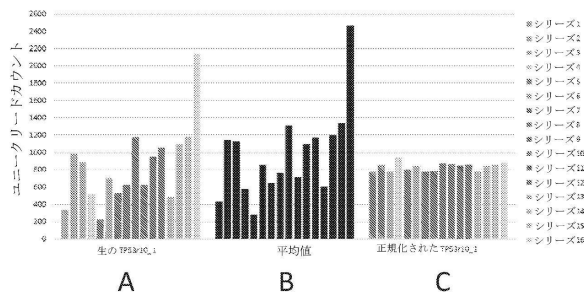


10

【図 2 4】

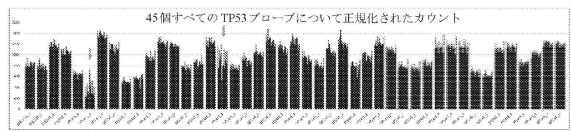
【図 2 4】

試料平均値全体に対するプローブ TP53r10.1 のユニークリードカウントの正規化



【図 2 5】

【図 2 5】



20

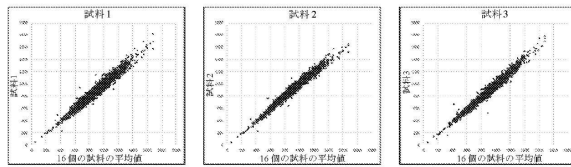
30

40

50

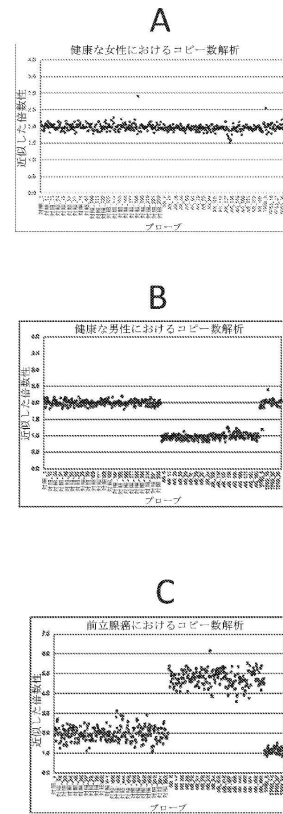
【図 26】

【図 26】



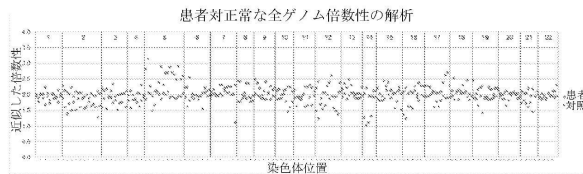
【図 27】

【図 27】



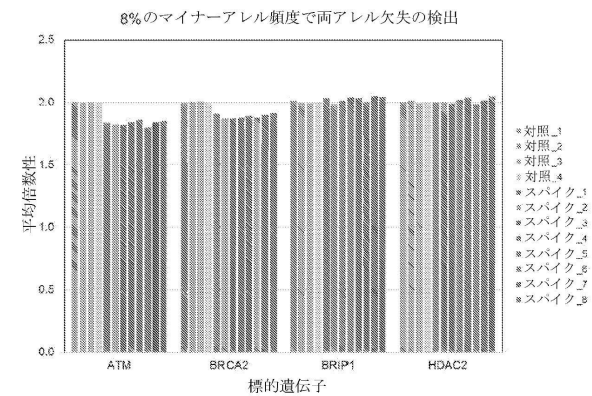
【図 28】

【図 28】



【図 29】

【図 29】



10

20

30

40

50

【配列表】

0007217224000001.app

10

20

30

40

50

フロントページの続き

(33)優先権主張国・地域又は機関

米国(US)

前置審査

弁護士 山本 健策

(72)発明者 レイモンド, クリストファー ケー.

アメリカ合衆国 ワシントン 98033, カークランド, カークランド ウェイ 550, スイート 200, レゾリューション バイオサイエンス, インコーポレイテッド 気付

(72)発明者 リム, リー ピー.

アメリカ合衆国 ワシントン 98033, カークランド, カークランド ウェイ 550, スイート 200, レゾリューション バイオサイエンス, インコーポレイテッド 気付

(72)発明者 ヘルナンデス, ジェニファー

アメリカ合衆国 ワシントン 98033, カークランド, カークランド ウェイ 550, スイート 200, レゾリューション バイオサイエンス, インコーポレイテッド 気付

審査官 田中 晴絵

(56)参考文献 国際公開第2016/028316(WO, A1)

米国特許出願公開第2014/0100792(US, A1)

米国特許出願公開第2014/0242581(US, A1)

米国特許出願公開第2015/0111757(US, A1)

国際公開第2016/022833(WO, A1)

国際公開第2014/055790(WO, A2)

(58)調査した分野 (Int.Cl., DB名)

C12Q 1/68 - 1/6897

C12N 15/00 - 15/90

C A p l u s / M E D L I N E / E M B A S E / B I O S I S (S T N)