

(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

C 1 2 Q	1/6869(2018.01)	F I	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 98033
 , カークランド, カークランド ウェイ
 550, スイート 200
 (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ップチャプローブと接触させることにより、前記キャッ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、

こ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、

そこで前記增幅領域は 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 D N A 断片に付着することができる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 (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 D N A 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変化を同定する、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変化は、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 D N A 断片は、無細胞 D N A (c f D N A) または細胞 D N A 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 D N A 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 D N A 断片は、

(i) 前記試験試料から細胞 D N A を単離すること、

(i i) 前記細胞 D N A を断片化し、前記ゲノム D N A 断片を取得すること、 30

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。

【請求項 9】

ステップ (i i) は、前記細胞 D N A 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 (i i) は、機械的応力を前記細胞 D N A 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 D N A を超音波処理することによって加えられる、請求項 1 0 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 D N A 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、請求項 1 から 1 1 のいずれかに記載の前記方法。

【請求項 13】

前記増幅領域は、 1 0 から 5 0 の間のヌクレオチド長にある、請求項 1 から 1 2 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、 2 0 から 3 0 の間のヌクレオチド長にある、請求項 1 から 1 3 のうちのいずれか 1 項に記載の前記方法。

【請求項 15】

10

20

30

40

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】

ステップ(b)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。

30

【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】

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。

40

【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 2 8 に記載の前記方法。

【請求項 3 2】

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 2 8 に記載の前記方法。

【請求項 3 3】

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 2 8 から 3 2 のいずれかに記載の前記方法。

【請求項 3 4】

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 3 3 のいずれか 1 項に記載の前記方法。

10

【請求項 3 5】

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 3 4 のいずれかに記載の前記方法。

【請求項 3 6】

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 3 5 のうちのいずれか 1 項に記載の前記方法。

20

【請求項 3 7】

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 個のみに対応される、請求項 1 から 3 6 のいずれかに記載の前記方法。

【請求項 3 8】

前記セットのアダプタの各アダプタの前記增幅領域は、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 增倍管を備え、前記セットのアダプタの各アダプタの前記 UMI 增倍管は、3 ヌクレオチド長にあり、所与の配列の前記 UMI 增倍管は、所与の配列の 1 個の試料タグ領域に対応され、

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 個のみに対応される、

30

請求項 1 から 3 7 のいずれかに記載の前記方法。

【請求項 3 9】

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 2 5 から 3 8 のいずれか 1 項に記載の前記方法。

【請求項 4 0】

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 3 9 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DNA標的領域を含むDNAライプ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40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DNA標的領域を含むDNAライプ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DNA標的領域を含む前記DNAライプ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41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DNA標的領域を含むDNAライプ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DNA標的領域を含む前記DNAライプ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42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DNA標的領域を含む前記DNAライプラリ断片のDNAシーケンシングを備える、請求項1から43のいずれか1項に記載の前記方法。

20

【請求項 4 5】

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(d)は、

(i) 前記試験試料に由来する前記ゲノムDNAライプ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DNAライプ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、

請求項1から44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DNA標的領域を含む前記DNAライプラリ断片の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(U M I E)を含む、請求項45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DNA断片配列の少なくとも一部を含む、請求項46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47に記載の前記方法。

40

【請求項 4 9】

前記ゲノムDNAライプ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45から48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49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50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ー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】

前記キャプチャープローブについての前記 R G D は、各キャプチャー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、

10

(i) 試料中の各キャプチャープローブ 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。

20

【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

30

(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】****【0 0 0 1】**

40

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 6 2 / 3 7 9 , 5 9 3 号、及び 2017 年 4 月 4 日に出願された米国仮特許出願第 6 2 / 4 8 1 , 5 3 8 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。

【0 0 0 2】**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、C L F K _ 0 0 5 _ 0 2 W O _ S T 2 5 である。テキストファイルは、2 , 2 3 8 K B であり、2017年8月24日に作成され、E F S - W e b を介して電子提出された。

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、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

【 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。

30

【発明の概要】**【課題を解決するための手段】****【 0 0 0 7 】**

c f DNA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c f DNA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に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c f DNA(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 D N A 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、D N A 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D N A ライブラリ断片を含むゲノムD N A ライブラリを生成することであって、各D N A ライブラリ断片は試験試料からのゲノムD N A 断片、及びアダプタを含む、生成すること、(b) D N A 標的領域に特異的に結着する複数のキャップチャープローブとゲノムD N A ライブラリを接触させることにより、D N A 標的領域を含むキャップチャープローブとD N A 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) D N A 標的領域を含むゲノムD N A 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D N A ポリヌクレオチドであり、そこで増幅領域は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み、そこで試料タグはユニークライブラリD N A 断片の一一致度を符号化し、試験試料の一一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D N A 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D N A 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D N A は、無細胞D N A (c f D N A) または細胞D N A である。特定の実施形態において、ゲノムD N A は、試験試料から単離されるc f D N A 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D N A 断片は、(i) 試験試料から細胞D N A を単離すること、及び(i i) この細胞D N A を断片化してゲノムD N A 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D N A 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D N A 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D N A 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D N A 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ライプラリ断片を含むゲノムDNAライプ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個のみに対応される。

【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応され、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個のみに対応される。

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、T 4 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。

【 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 ライブライア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 ライブライア断片との間に形成される複合体の精製を有し、ゲノム DNA ライブライアから関心領域を含む DNA ライブライア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 ライブライア断片との間に形成される複合体の精製を有し、ゲノム DNA ライブライアから関心領域を含む DNA ライブライア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 ライブライア断片の DNA シーケンシングを備える。

【 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 ライブライア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 ライブライア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。

【 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 ライブライア断片の 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 (UME) を含む。いくつかの実施形態において、UME 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 UME 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 】

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 ライブライア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。

10

20

30

40

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一様に分布する、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、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、またそこでステップ(c)は、D N A 標的領域を含むゲノム D N A 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 D N A 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 D N A 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、D N A 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、B R C A 2、A T M、B R C A 1、B R I P 1、C H E K 2、F A N C A、H D A C 2、及び／またはP A L B 2である。

【 0 0 3 3 】

特定の実施形態は、複数のD N Aライブラリ断片を含むゲノムD N Aライブラリについて描画され、そこで各D N Aライブラリ断片はアダプタ及びゲノムD N A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D N Aポリヌクレオチドであり、そこで増幅領域は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み、そこで試料タグはユニークライブラリD N A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D N A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(U M I)をさらに含み、そこでU M IはユニークゲノムD N A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、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 5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5 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、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。特定の実施形態において、U M I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5 0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、1 0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U M I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U M I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個のみに対応される。いくつかの実施形態において、ゲノムD N A断片は、c f D N Aである。

【 0 0 3 4 】

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U M I増倍管を備え、複数のアダプタの各アダプタのU M I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U M I増倍管は、複数のアダプタの各試料タグ領域に対応され、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個のみに対応される。いくつかの実施形態において、ゲノムD N A断片は、c f D N Aである。

【 0 0 3 5 】

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D N A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D N Aライブラリに属する1つのゲノムD N Aライブラリの試料タグ領域の核酸配列は、複数のゲノムD N Aライブラリに属する他のゲノムD N A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D N Aライブラリに属する1つのゲノムD N Aライブラリの増幅領域の核酸配列は、複数のゲノムD N Aライブラリに属する他のゲノムD N Aライブラリの増幅領域の核酸配列に同一である。

10

20

30

40

50

【 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ープローブと 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ープローブと 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ープローブと 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ープローブと 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。

10

20

30

40

【 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、增幅領域、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、そこで增幅領域は、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 增倍管は、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 增倍管を含む。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 增倍管を含む。

【 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

50

個のアンカー領域のうちの 1 個のみに対応される。請求項 7 5 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 U M I 増倍管を備え、複数のアダプタの各アダプタの U M I 増倍管は、3 ヌクレオチド長にあり、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 6 4 個の可能なヌクレオチド配列の U M I 增倍管は、複数のアダプタの試料タグ領域のそれぞれに対応され、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 個のみに対応される。

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)

D N A 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 D N A ライブライ断片を含むゲノム D N A ライブライを生成することであつて、そこで各前記 D N A ライブライ断片は前記試験試料からのゲノム D N A 断片、及びアダプタを含む、前記生成すること、

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

(c) 前記 D N A 標的領域を含む前記ゲノム D N A 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 D N A ポリヌクレオチドであり、

そこで前記増幅領域は 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み、

そこで前記試料タグは前記ユニークリブライ D N A 断片の一一致度を符号化し、前記試験試料の一一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 D N A 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、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 (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 D N A 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。

(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 D N A は、無細胞 D N A (c f D N A) または細胞 D N A 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 D N A は、前記試験試料から単離される c f D N A であり、前記試験試料は、

10

20

30

40

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
汎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、

(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、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 (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
つて実施される、項目 8 に記載の前記方法。

10

(項目 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、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 (UMI) をさらに含む、先行項目のいずれかに記載の前記方法。

20

(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。

30

(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 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 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 UMI 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。

40

(項目 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、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)

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個のアンカーレジアのうちの1個のみに対応される、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ヌクレオチド長になり、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

プラの前記試料タグのいずれかの他のス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スクレオチド長にあり、それぞれの前記可能なス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応される。

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個のス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個のみに対応される。

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ス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ス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ス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オリゴスクレオチドと、前記アンカー領域の少なくとも一部を含む前記オリゴス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ポリス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オリゴス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ープローブと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ープローブと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/または增幅を実施する、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ープローブと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增幅を実施する、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)

10

20

30

40

50

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ライプ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ライプ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ライプラリ断片の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ライプ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ブは、1 セットの対照プローブである、項目 5 6 に記載の前記方法。

(項目 5 8)

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、

(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、

(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 5 7 に記載の前記方法。

(項目 5 9)

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、項目 5 8 に記載の前記方法。

(項目 6 0)

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) キャプチャプローブについて R G D を測定すること、

(ii) 前記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、実験群中にすべての試料にわたり前記キャプチャプローブについての前記 R G D を正規化すること、

(iii) 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

(iv) 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を平均すること、

を備える、前記方法。

(項目 6 1)

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ブを設計し、妥当性確認することであって、前記染色体安定性プローブ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、

(iii) 各染色体プローブ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ブ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 6 2)

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 6 1 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 P A R P 阻害剤を投与する、

ことを備える、前記方法。

(項目 6 3)

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 6 4)

前記遺伝子は、疾患と関連する、項目 6 3 に記載の前記方法。

(項目 6 5)

10

20

30

40

50

前記疾患は、癌である、項目64に記載の前記方法。

(項目66)

前記遺伝子は、B R C A 2、A T M、B R C A 1、B R I P 1、C H E K 2、F A N C A、H D A C 2、及び／またはP A L B 2である、項目63に記載の前記方法。

(項目67)

複数のD N Aライプラリ断片を含むゲノムD N Aライプラリであって、

前記D N Aライプラリ断片のそれぞれはアダプタ及びゲノムD N A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D N Aポリヌクレオチドであり、

前記増幅領域は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み、

前記試料タグは前記ユニークライプラリD N A断片の一致度を符号化し、前記試験試料の一一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D N A断片に付着することが可能である、

前記ゲノムD N Aライプラリ。

(項目68)

前記試料タグは、ユニーク分子識別子(U M I)をさらに含み、前記U M Iは、前記ユニークゲノムD N A断片の前記同定を容易にする、項目67に記載の前記ゲノムD N Aライプラリ。

(項目69)

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67または68に記載の前記ゲノムD N Aライプラリ。

(項目70)

前記増幅領域は、25ヌクレオチド長にある、項目69に記載の前記ゲノムD N Aライプラリ。

(項目71)

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D N Aライプラリ。

(項目72)

前記試料タグは、8ヌクレオチド長にある、項目71に記載の前記ゲノムD N Aライプラリ。

(項目73)

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67から72のうちのいずれか1項に記載の前記ゲノムD N Aライプラリ。

(項目74)

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73に記載の前記ゲノムD N Aライプラリ。

(項目75)

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D N Aライプラリ。

(項目76)

前記アンカー領域は、10ヌクレオチド長にある、項目75に記載の前記ゲノムD N Aライプラリ。

(項目77)

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67から76のいずれか1項に記載の前記ゲノムD N Aライプラリ。

(項目78)

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67から77のいずれか

10

20

30

40

50

1項に記載の前記ゲノムDNAライプラリ。

(項目79)

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67から78のいずれか1項に記載の前記ゲノムDNAライプラリ。

(項目80)

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67から78のいずれか1項に記載の前記ゲノムDNAライプラリ。

(項目81)

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個のみに対応される、項目67から78のいずれか1項に記載の前記ゲノムDNAライプラリ。

10

(項目82)

前記複数のアダプタの各アダプタの前記增幅領域は、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つ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前記試料タグ領域のそれぞれに対応され、

20

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個のみに対応される、

項目67に記載の前記ゲノムDNAライプラリ。

(項目83)

前記ゲノムDNA断片は、cfDNAである、項目67から82のいずれかに記載の前記ゲノムDNAライプラリ。

30

(項目84)

項目67から83のいずれか1項に従い、1個より多いゲノムライプラリを含む、複数のゲノムDNAライプラリ。

(項目85)

前記複数のゲノムDNAライプラリに属する1個のゲノムDNAライプ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ライプラリに属する他のゲノムDNAライプラリの前記試料タグ領域の前記核酸配列と異なる、項目84に記載の前記複数のゲノムDNAライプラリ。

(項目86)

前記複数のゲノムDNAライプラリに属する1個のゲノムDNAライプラリの前記增幅領域の前記核酸配列は、前記複数のゲノムDNAライプラリに属する他のゲノムDNAライプラリの前記增幅領域の前記核酸配列に同一である、項目84または85に記載の前記複数のゲノムDNAライプラリ。

40

(項目87)

無細胞DNA(cfDNA)のDNA標的領域の遺伝子解析方法であって、

(a) 項目67から86のいずれかに記載の前記DNAライプラリを生成すること、

(b) DNA標的領域に特異的に結合する複数のキャップチャープローブと前記cfDNAライプ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ップチャープローブとDNAライプ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cfDNA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ライプラリ断片を含むDNAライプラリを生成することであって、前記DNAライプ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合する複数のキャプチャープローブと前記cfDNAライプ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ープローブとDNAライプラリ断片との間に複合体を形成することと、
(e) 前記cfDNAクローンライプ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ライプ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プラリ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。

10

(項目 1 0 0)

前記複数のアダプタの各アダプタの前記增幅領域は、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。

20

(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 增倍管を備える、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 個のみに対応される、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增幅領域は、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、

30

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応され、

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、

40

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 個のみに対応される、

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

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 (nt) 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 8mer 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、47nt のアダプタの例示的な DNA 配列を示す。

【図4 - 2】同上。

【図5】図5 A 及び図5 B は、試料タグ内の UMI 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】CNL アッセイについてのゲノムライプラリを構築するプロセスを説明する図を示す。図6 A は、10nt 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】DNA インプットを CNL ライプラリ中に示す。左に示されるマーカーサイズ (bp) 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、CNL 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。

【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 CNL 測定値における有意な標準からの偏差を定量化する Log₁₀ の P 値プロットを示す。SNP 百分率は、ATM 及び BRCA2 試料中に存在する希少なヘテロ接合 SNP 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 DNA によって添加される cf DNA 試料についての CNL 測定値における有意な標準からの偏差を定量化する Log₁₀ の P 値プロットを示す。SNP 百分率は、ATM 及び BRCA2 試料中に存在する希少なヘテロ接合 SNP 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 PCR 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ニングへ提供する追加のアダプタ配列によってゲノムライプラリへの cf DNA の変換を示す。図11 B は、標的特異的キャプチャー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ーシーケンスランから 377,711,020 個の Illumina NextSeq 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ニングに由来する。

【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。

【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。

10

20

30

40

50

【図13C】同上。

【図13D】同上。

【図13E】同上。

【図13F】同上。

【図13G】同上。

【図13H】同上。

【図14A】図14A～図14Iは、プール7-9からのオリゴスクレオチドのアダプタの配列を示す。

【図14B】同上。

【図14C】同上。

【図14D】同上。

【図14E】同上。

【図14F】同上。

【図14G】同上。

【図14H】同上。

【図14I】同上。

【図15A】図15A～図15Hは、プール10-12からのオリゴスクレオチドのアダプタの配列を示す。

【図15B】同上。

【図15C-1】同上。

【図15C-2】同上。

【図15D】同上。

【図15E】同上。

【図15F】同上。

【図15G】同上。

【図16A】図16A～図16Hは、プール13-15からのオリゴスクレオチドのアダプタの配列を示す。

【図16B】同上。

【図16C】同上。

【図16D】同上。

【図16E】同上。

【図16F】同上。

【図16G】同上。

【図16H】同上。

【図17A】図17A～図17Hは、プール16-18からのオリゴスクレオチドのアダプタの配列を示す。

【図17B】同上。

【図17C】同上。

【図17D】同上。

【図17E】同上。

【図17F】同上。

【図17G】同上。

【図17H】同上。

【図18A】図18A～図18Hは、プール19-21からのオリゴスクレオチドのアダプタの配列を示す。

【図18B】同上。

【図18C】同上。

【図18D】同上。

【図18E】同上。

【図18F】同上。

10

20

30

40

50

【図18G】同上。

【図18H】同上。

【図19A】図19A～図19Hは、プール22-24からのオリゴスクレオチドのアダプタの配列を示す。

【図19B】同上。

【図19C】同上。

【図19D】同上。

【図19E】同上。

【図19F】同上。

【図19G】同上。

【図19H】同上。

【図20A】図20A～図20Hは、プール25-27からのオリゴスクレオチドのアダプタの配列を示す。

【図20B】同上。

【図20C】同上。

【図20D】同上。

【図20E】同上。

【図20F】同上。

【図20G】同上。

【図20H】同上。

【図21A】図21A～図21Hは、プール28-30からのオリゴスクレオチドのアダプタの配列を示す。

【図21B】同上。

【図21C】同上。

【図21D】同上。

【図21E】同上。

【図21F】同上。

【図21G】同上。

【図21H】同上。

【図22A】図22A～図22Hは、プール31-32からのオリゴスクレオチドのアダプタの配列を示す。

【図22B】同上。

【図22C】同上。

【図22D】同上。

【図22E】同上。

【図22F】同上。

【図22G】同上。

【図22H】同上。

【図23】図23A～図23Cは、TP53遺伝子の標的シーケンシングを示す。図23Aは、キャプチャプローブのBedFile表示を図示する。図23Bは、0から8000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23Cは、既知のTP53スプライス変異体のUCSC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。

【図24】図24A～図24Cは、16個の試料にわたり、単一のプローブ、TP53r10_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24A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TP53r10_1によって生のユニークリードキャプチャ数を示す。図24Bは、16個の試料すべてについて2596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24C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TP53r10_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

10

20

30

40

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 T P 5 3 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、T P 5 3 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 27 A ~ 図 27 C は、健康な女性及び男性ドナーからの c f D N A、ならびに進行期前立腺癌患者からの c f D N A のコピー数解析を図示する。図 27 A は、健康な女性ドナーからの c f D N A 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 A R 遺伝子を標的にする一連のプローブ、及び T P 5 3 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 27 B は、X 連鎖 A R 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 27 C は、進行期前立腺癌患者からの c f D N A のコピー数解析を図示し、対照プローブ、A R 遺伝子の增幅、及び T P 5 3 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 c f D N A ライブ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 N A 0 2 7 1 8 (単一アレル A T M) からのゲノム D N A、及び N A 0 9 5 9 6 (単一アレル B R C A 2) からのゲノム D N A は、16 % で N A 1 2 8 7 8 からの「ゴールドスタンダード」ゲノム D N A 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 C N V 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 個の搅乱のない対照遺伝子、B R I P 1 及び H D A C 2 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 D N A (たとえば、組織診試料からの) または c f D N A (たとえば、血液試料からの) の試料内の、変異性変化、S N P、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 c f D N A 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 D N A 断片から作製されるゲノム D N A ライブリを生成し、D N A 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 D N A 標的領域を含む D N A ライブリ断片を単離し、D N A 標的領域の定量的遺伝子解析を実施することにより D N A 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、固体腫瘍の解析のための無細胞DNA(*c f DNA*)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、*c f DNA*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*c f DNA*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

【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。

20

【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想定される。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ニング断片が含むことを意味する。各クローニング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ニングを区別する。しかしながら、数千個の独立したクローニング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ニング(すなわち、断片)は、同一の試料に由来する他のクローニング配列に関して、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ニングは、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c f D N A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内で分断し、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8mehr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。

30

【0050】

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、c f DNA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、c f DNA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c f DNAから生成する。関心のゲノム領域を認識するキャップチャプローブ（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。

【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。

【0058】

10

20

30

40

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。

10

【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。

20

【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徴、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、特定の特徴、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。

30

【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。

40

【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、医学的状態の診断若しくは監視、マイクロバイオ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

50

DNA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）などと関連するDNA配列である。

【0065】

本明細書で使用する「循環DNA」、「循環セルフリーDNA」、及び「セルフリー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、またはネクロー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、又クレアーゼ）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ライプラリ断片のゲノムDNAライプラリを生成し、これらのDNA標的領域を含むDNAライプラリ断片を単離し、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ライプラリ断片の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。單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、ゲノムライプ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(キャプチャプロープに関して>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に、ゲノムライプラリに生成された後に、標的特異的キャプチャプロープ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。
10

(b) ゲノムライプ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数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、固体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。
20

(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プ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。
40

(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%を超えるながら、標的上リードの4から
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。

10

【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、これらをすることができ、これらをすることに適しており、及び/またはこれらをすることができ、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、これらをすることが可能であり、これらをするのに適しており、及び/またはこれらのことをすることができる。

20

【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搅乱を検出する。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。

30

【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。

40

【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

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、マイクロバイオ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、これらをすることが可能であり、これらをすることに適し、及び/またはこれらをすることができ、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、これらをすることが可能である、これらをするのに適している、及び/またはこれらをすることができる。

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、脾臓、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。

【 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ライプラリを取得する、生成する、作製する、形成する、及び／または產生する前に、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ライプ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、方法によってライプ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。

【 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5kb bpで破壊するための音響装置である。またCovarisは、メイトペアライプラリについて、試料を6から20kbで処理するチューブ(gTube)を製造する。Bioptron(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。Digitalab(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。

【 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、トランスポンサーによる処置によって、

10

20

30

40

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ンタニックを产生し、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ーザーを使用する。

【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(cf DNA)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。cf DNAのサイズ分布は、約150 bpから約180 bp断片の範囲である。cf DNA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ロバストなcf DNA解析に対する困難な課題が提示される。cf DNA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、cf DNAの解析が「リキッドバイオпси」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】

30

さらに、cf DNA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。

【0089】

特定の実施形態においてcf DNA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】

40

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cf DNAを単離することができる。

【0091】

(a) 末端修復cf DNA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cf DNAまたは断片化細胞DNAの末端修復を備える。断片化cf DNA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cf DNA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、たとえば产生することができる。いくつかの実施形態において、末端修復

50

c f DNA または細胞DNA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DNAまたはc f DNA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 DNAまたは細胞DNA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 DNAまたは細胞DNA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 DNAまたは細胞DNA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 DNAまたは細胞DNAは、非鑄型化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 DNAまたは細胞DNA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 DNAまたは細胞DNA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(TdT)で処理される。一部の実施形態では、TdT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 DNAまたは細胞DNA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、Sau3A酵素など)による部分的消化を用いて処理される。

【0092】

(b) 末端修復c f DNAへのアダプタ分子の付着

特定の実施形態において、c f DNA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 DNA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 DNA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 DNA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。

【0093】

本明細書に使用されるように、用語「アダプタ」と「アダプタモジュール」は、互換的に使用され、また增幅領域、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(UMI)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UMI、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、UMI、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、UMI、及びアンカー領域を含む。ある一定の実施形態において、UMI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/UMI領域、及びアンカー領域を含む。

【0094】

本明細書に使用されるように、用語「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むアダプタ分子の元素を指す。特定の実施形態では、アダプタは、ゲノムDNA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DNA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。

【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

10

20

30

40

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】

10

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。

20

【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】

30

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。

40

【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

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-merr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較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。

10

【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。

20

【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ライプラリを生成するため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、單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。

30

【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ライプ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。

40

【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】

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、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）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の長さにあり、別々のPCR增幅配列、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ライブル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

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はなく、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、 U M I 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。

【 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 D N A 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 D N A 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 D N A に付着する。
。

【 0114 】

いくつかの実施形態において、ゲノム D N A 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 D N A または細胞 D N A 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 D N A 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 D N A 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 D N A 断片を、 D N A 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 D N A リガーゼ、及び / または 1 個以上の D N A ポリメラーゼを、ゲノム D N A 断片、及び D N A 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、 T 4 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、 D N A リガーゼは、 T a q D N A リガーゼである。ある一定の実施形態において、 D N A ポリメラーゼは、 T a q ポリメラーゼである。特定の実施形態において、 D N A ポリメラーゼは、 B s t ポリメラーゼ全長である。

【 0115 】

図 6 は、複数のアダプタを修復された D N A 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、 D N A 二本鎖であり、この D N A 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。アンカーニ本鎖は、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。 T 4 ポリヌクレオチドキナーゼ、 T a q D N A リガーゼ、及び B s t 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 D N A 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

5'末端に付着させることができる。

【0116】

2. DNAライプラリ増幅

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ライプラリ、たとえば、細胞DNAライプラリまたはc f DNAライプラリなどの増幅を有し、1個のDNAクローンライプラリ、若しくは複数のDNAクローンの1個のライプラリ、たとえば、1個のc f DNAクローンライプラリ、若しくは複数のc f DNAクローンの1個のライプラリ、または1個の細胞DNAクローンライプラリ、若しくは複数の細胞DNAクローンの1個のライプラリなどを生成する。DNAライプラリの各分子は、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c f DNA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c f DNAライプラリの一部を、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Tm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ライプラリ、たとえば、細胞DNAライプラリまたはc f DNAライプラリを、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。

20

【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クローンライプ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、各ライプ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プ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プ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インフォマティクスに基づくカウントにより決定することができる。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ライプラリ、たとえば、c f DNAライプラリまたはゲノムDNAライプラリなどに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クローンライプ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、バイオインフ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クローンライプ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クローンライプラリにおける配列深度が明らかにならない場合は、DNAクローンライプラリの再処理及び/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c f DNA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合部位へ結合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(c f DNA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インフォマティクス的に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、短いタンデム反復(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c f DNA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インフォマティクス解析。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、ただしそりより小さなDNA断片を含むゲノムDNAライブラリ、たとえば、c f DNA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、c f 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。c f DNA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か示さず、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。

10

【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。

20

【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。

30

【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。

40

【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

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、キャプチャープローブモジュールは、キャプチャー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、キャプチャープローブモジュールは、ビ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。

10

【0137】

様々な実施形態では、キャプチャー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ー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ープローブモジュールが安定したタグ付きDNAライブラリ・キャプチャー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ー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。キャプチャー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。

20

【0139】

特定の実施形態では、キャプチャープローブは約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。

30

【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・キャプチャー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、ビオチン・ストレプトアビシン単離技術を用いて複合体を単離する。

【0141】

40

50

特定の実施形態では、単離したタグ付きDNAライプラリ断片 - キャップチャッ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ライプラリ - 多機能キャップチャッ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ライプラリ断片を鋳型として利用して多機能キャップチャップローブの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' 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、ハイブリッドキャップチャップローブ - 単離したタグ付きDNA標的分子、たとえば、タグ付きc f DNA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ライプラリ - 多機能キャップチャッ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 I V, Hubscher 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、キャップチャッ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、キャップチャッ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、キャップチャッ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個以上のキャップチャップローブモジュールを複数のゲノムDNAライプラリクローン中の1個以上の標的遺伝子座にハイブリダイズして、1個以上のキャップチャップローブモジュール - DNAライプラリクローン複合体を形成すること；b) a)からの1個以上のキャップチャップローブモジュール - DNAライプ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ップチャップローブモジュール - DNAライプ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、増幅したハイブリッド核酸分子を生成するためにキャップチャッ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ップチャッ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、キャップチャッ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個以上のキャップチャップローブモジュールを複数のDNAライプ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個以上のキャップチャップローブモジュール - DNAライプ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ライプ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ライプ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 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 PCR 増幅を実施すること；及び f) e) における PCR 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きる、定量化することを備える。

10

【 0148 】

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ライプ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 - cf DNA ライプ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 クローンを鑄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。

20

【 0149 】

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 クローンを鑄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。

【 0150 】

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 PCR 反応条件を用いて PCR 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 PCR 反応は、2 個の PCR プライマーを用いる。一実施形態では、e) における PCR 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 PCR プライマーを用いる。特定の実施形態では、e) における PCR 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 PCR プライマーを用いる。あるいはいくつかの実施形態では、e) における PCR 反応は、標的遺伝子座にハイブリダイズする第 1 の PCR プライマーを用い、第 2 の PCR 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

【 0151 】

あるいは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いてもよく、標的遺伝子座のいずれかの cf DNA 鎮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 cf DNA クローンライプ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 クローンライプ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 bp 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。

【0152】

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。

【0153】

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。

10

【0154】

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。

20

【0155】

(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、增幅由来の互いの「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。

30

【0156】

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

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またはそれ以上のス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、ス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、ス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す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することができる。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 D N A 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、特定のキャプチャープローブと関連し、各キャプチャー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D N A 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 D N A 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 / 1 0 、約1 / 1 2 、約1 / 1 4 、約1 / 1 6 、約1 / 1 8 、約1 / 2 0 、約1 / 2 5 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 】

30

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。

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(S N V)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 . 0 2 から0 . 0 2 %の固有のエラー率を有し、これは1 / 2 0 0 から1 / 5 0 0 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 0 0 0 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5 0 0 0 個のユニーク分子を解析すれば、> 5 0 , 0 0 0 リードの十分なシーケンシング深度で、5 0 0 0 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、医学的状態の診断または監視、マイクロバイオ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配列がキャプチャー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、医学的状態の診断または監視、マイクロバイオ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ー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マutation、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/

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、医学的状態の診断または監視、マイクロバイオ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 C N L 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、S N P 、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、D N A 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 D N A または c f 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、細胞 D N A または c f D N A を 1 個以上の末端修復酵素によって処置し、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 の各末端に付着させて、ゲノム 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 D N A 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、S N P 、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、医学的状態の診断または監視、マイクロバイオ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 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 の各末端に付着させて、ゲノム D N A ライブラリを生成すること、ゲノム当量数を D N A 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 D N A 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 (A P O E 1) 、シャルコー・マリー・トゥース病、レーベル遺伝性視神経症 (L H O N) 、アンジェルマン症候群 (U B E 3 A 、ユビキチン - タンパク質リガーゼ E 3 A

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、-グロビン)、ゴーシェ病(1型)(GBA、グルコセレブロシダーゼ)、囊胞性線維症(CFTR、上皮クロライドチャネル)、鎌状赤血球症(HBB、-グロビン)、ティ・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、低密度リポタンパク質受容体)、成人型囊胞腎(PKD1、ポリシスチン)、ハンチントン病(HDD、ハンチントン)、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(Hprt1、ヒポキサンチングアニンリボシリルransフェラーゼ1)、及び副腎白質ジストロフィー(ABC D1)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、肺臓癌、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)、真性赤血球增加症、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(endothelial sarcoma)、リンパ管肉腫、リンパ管内皮肉腫(lymphangiopericytoma)、滑膜腫、中皮腫、ユーリング腫瘍、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、ABC B1、ABC C2、ABC C4、ABC G2、ABL 1、ABL 2、AKT 1、AKT 2、AKT 3、ALDH4A1、ALK、APC、

10

20

30

40

50

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、ERC2、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、RUX1、SLC19A1、SLC22A2、SLCO1B3、SMAD2、SMAD3、SMAD4、SMARCA4、SMARCB1、SMO、SOD2、SOX10、SOX2、SRC、STK11、SULT1A1、TBX22、TET2、TGFBR2、TMPRSS2、TNFRSF14、TOP1、TP53、TPMT、TSC1、TSC2、TYMS、UGT1A1、UMPS、USP9X、VHL、及びWT1が挙げられる。
【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、または遺伝子融合を含む。

【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。

【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

10

20

30

40

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 1 6 個以上を含むが、これらに限定されない。

【 0 1 8 9 】

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 1 A (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 1 B (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 1 C (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 (アリール炭化水素受容体) 、 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-A 1) 、 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 P 4 5 0 、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 P 4 5 0 、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 P 4 5 0 、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 P 4 5 0 、ファミリー-2 、サブファミリー-C 、ポリペプチド 1 9) 、 C Y P 2 D 6 (チトクロム P 4 5 0 、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 P 4 5 0 、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 P 4 5 0 、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 P 4 5 0 、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 D 2) 、 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-1 1) 、 M T H F R (5 , 1 0 - 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 P 2 Y 、 G - タンパク質共役型、 1) 、 P 2 R Y 1 2 (プリン受容体 P 2 Y 、 G - タンパク質共役型、 1 2) 、 P T G I S (プロスタグランジン I 2 (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-1 9 (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-1 B 1) 、 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、 1 A 、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 グルクロノシリルトランスフェラーゼ 1 ファミリー、ポリペプチド A 1) 、 V D R (ビタミン D (1 , 2 5 - ジヒドロキシビタミン D 3) 受容体) 、 V K O R C 1 (ビタミン K 工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。

【 0 1 9 0 】

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、*Blastomycetes*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】

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あ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、水痘帶状発疹、サイトメガロウイルス、エプスタイン・バールウイルス（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）、モロニーマウス白血病ウイルス（M-MuLV）、モロニーマウス肉腫ウイルス（MoMSV）、ハーベイマウス肉腫ウイルス（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、フレンドマウス白血病ウイルス、マウス幹細胞ウイルス（MSCV）及びラウス肉腫ウイルス（RSV）、HIV（HIV-1型、及びHIV-2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、ボ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、ハントウイルス、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含む）及びパラミクソウイルス科（ Hendra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）4050

、西ナイルウイルス、ならびにいずれかの脳炎（encephalitis）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。

【0195】

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、HLA-A、HLA-B、HLA-C、HLA-DR、HLA-DP、及びHLA-DQのうちの1つ以上を含むが、これらに限定されない。

【0196】

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

【0197】

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の各末端へ付着させて、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ゲノム当量数を決定すること、及びDNA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DNAはcfDNAである。特定の実施形態では、このDNAは細胞DNAである。

20

【0198】

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

【0199】

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。

【0200】

抗がん薬の例示的な例は、アルキル化剤（たとえば、チオテバ及びシクロホスファミド（CYTOXAN（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ビポスルファン）、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）、ナイトロジエンマスターード類（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、ノブエンビキン、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、ブレオマイシン類、カクチノマイシン、カリチアマイシン、カラビシン、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、ダウノルビシン、デトルビシン、6-ジアゾ-5-オキソ-L-ノルロイシン、ドキソルビシンとそのペグ化製剤類、エピルビ

40

50

シン、エソルビシン、イダルビシン、マルセロマイシン、マイトイマイシン類、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、ロドルビシン、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、ゾルビシン)、代謝拮抗剤類(たとえば、メトレキサート及び5-フルオロウラシル(5-FU))、葉酸類似体(たとえば、デノブテリン、メトレキサート、ブロブテリン、トリメトレキサート)、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、プロピオニ酸ドロモスターロン、エピチオスタノール、メビチオスタン、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、フオリン酸(folinic acid))、アセグラトン、アルドホスファミド配糖体、アミノレブリン酸、アムサクリン、ベストラブシル、ビサントレン、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tron、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、ピラルビシン、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)、ラゾキサン、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、ピポブロマン、ガシトシン、アラビノシド('Ara-C')、シクロホスファミド、チオテバ、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.)及びドキセタキセル(TAXOTERE(登録商標)、Rhône-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)、ビンプラスチン、白金、エトポシド(VP-16)、イホスファミド、マイトイマイシンC、ミトキサンtron、ビンクリスチン、ビノレルビン、ナベルビン、ノバントロン、テニポシド、アミノブテリン、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、ならびに上記いづ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、ビカルタミド、ロイプロリド、及びゴセレリン)、ならびに上記のいづ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、シクロスボリン、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の不死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みえた。ATM試料は、男性ドナーに由來したため、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。ライプラリ構築のために、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。ライプラリインプット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、そしてゲノムライプラリに変換した。約500ngの各ライプラリを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 NG机器上で、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。コピー数搅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ATM DNA(男性)、試料3、女性DNA中の5%のBRCA2 DNA(女性)、試料4、純粋な女性DNA)。

【0208】

CNL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、つぎに各プロープ位置で定量化する。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CNL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。

【0209】

図8は、断片化され、配合されたゲノムライプラリ中のAR遺伝子(図8B)及びATM遺伝子(図8C)についてコピー数決定の箱ひげ図を示す。ATM試料が男性であるため、AR遺伝子(X連鎖、ヘミ接合性)及びATM遺伝子の両方は、CNL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9B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相關していた。図10は、異性からcfDNAの微量なスパイクイン、または断片化gDNAの微量な追加のいずれか一方を有する、主にcfDNA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相關していた。断片化されたgDNA、及びcfDNA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。

10

20

30

【0210】

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增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、SNV、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DNA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DNA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。

【0211】

実施例2：健康なドナー及び癌患者からのcfDNAのコピー数解析

以下の例は、ゲノムライプラリ構築中に、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。Qiaegen Circulating Nucleic Acids Extraction Kit(Qiaegen, Hilden, Germany)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DNAを抽出した。Qubit蛍光光度計(The Fisher, Waltham, MA)、及び対応するhsDNA定量化キットを用いて、二本鎖DNA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PCRマーカー(New England Biolabs, Ipswich, MA)によって、サイズ解析を実施した。試料からのcfDNAの収量によって、約40から100ngのcfDNAをライプラリ構築に使用した。

40

【0212】

ライプラリ構築の基本的な特徴を図11Aから11Cに示す。最初にcfDNA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、10nt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cfDNA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。

50

【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ID	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_16-1	G T A T G C C [3-dA-Q] *	1
パートナー鎖oライゲーション鎖オリゴo_16-2	A G C G T T A [3-dC-Q] *	2
パートナー鎖oライゲーション鎖オリゴo_16-3	T C G A C A T [3-dG-Q] *	3
パートナー鎖oライゲーション鎖オリゴo_16-4	C A T C A G G [3-dT-Q] *	4
ライゲーション鎖オリゴ_16-1	/5 Phos/TGG C AT ACG T**	5
ライゲーション鎖オリゴ_16-2	/5 Phos/GTA A CG CTA G**	6
ライゲーション鎖オリゴ_16-3	/5 Phos/CAT G TC GAT C**	7
ライゲーション鎖オリゴ_16-4	/5 Phos/ACC T GA TGC 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 Phos/ は、5'位の塩基へ5'リン酸基の化学付加を示す。

【0213】

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、ゲノムライプラリを生成した。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プラリをqPCRによってカバレッジの深度について定量化した。つぎにゲノムライプラリを増幅し、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。

【0214】

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つ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つ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 : TP5 3プローブ

名称	配列	配列番号 :
TP5 3_1	G G C A C A G A C C C T C T C A C T C A T G T G A T G T C A T C T C T C C T C C	7 6 8 9
TP5 3_2	A T G G G G G T G G G A G G C T G T C A G T G G G G A A C A A G A A G T G G A G	7 6 9 0
TP5 3_3	G T C A G T C T G A G T C A G G C C C T T C T G T C T T G A A C A T G A G T T T	7 6 9 1
TP5 3_4	C C T G A A G T C C A A A A A G G G T C A G T C T A C C T C C C G C C A T A A A	7 6 9 2
TP5 3_5	T C A T G C T G G A T C C C C A C T T T T C C T C T T G C A G C A G C C A G A C	7 6 9 3
TP5 3_6	G T T G G G G T G G G G G T G G T G G G C C T G C C C T T C C A A T G G A T C C	7 6 9 4
TP5 3_7	C A G T T T C C A T A G G T C T G A A A A T G T T T C C T G A C T C A G A G G G	7 6 9 5
TP5 3_8	C T G C C A T G G A G G A G C C G C A G T C A G A T C C T A G C G T C G A G C C	7 6 9 6
TP5 3_9	G C A G A G A C C T G T G G G A A G C G A A A A T T C C A T G G G A C T G A C T	7 6 9 7
TP5 3_10	C T G G G G G G C T G G G G G G C T G A G G A C C T G G T C C T C T G A C T G C	7 6 9 8
TP5 3_11	G C A G G G G G A T A C G G C C A G G C A T T G A A G T C T C A T G G A A G C C	7 6 9 9
TP5 3_12	G T G G C C C C T G C A C C A G C A G C T C C T A C A C C G G C G G C C C T G	7 7 0 0
TP5 3_13	G G G G G G A G C A G C C T C T G G C A T T C T G G G A G C T T C A T C T G G A	7 7 0 1
TP5 3_14	C C G T G C A A G T C A C A G A C T T G G C T G T C C C A G A A T G C A A G A A	7 7 0 2
TP5 3_15	C C C C G G A C G A T A T T G A A C A A T G G T T C A C T G A A G A C C C A G G	7 7 0 3
TP5 3_16	C C A G A A A A C C T A C C A G G G C A G C T A C G G T T T C C G T C T G G G C	7 7 0 4
TP5 3_17	T A G G T T T C T G G G A A G G G A C A G A A G A T G A C A G G G G C C A G G	7 7 0 5
TP5 3_18	T G C T T T A T C T G T T C A C T T G T G C C C T G A C T T T C A A C T C T G T	7 7 0 6

10

20

30

40

50

【表 2 - 2】

TP 5 3_1 9	CCTGGGCAACCAGCCCTGTCGT CTCTCCAGCCCCAGCTGC	7707
TP 5 3_2 0	TTTGCCAACTGGCCAAGACCTG CCCTGTGCAGCTGTGGGT	7708
TP 5 3_2 1	CCATCGCTATCTGAGCAGCGCT CATGGTGGGGCAGCGCC	7709
TP 5 3_2 2	GCCATCTACAAGCAGTCACAGC ACATGACGGAGGTTGTGA	7710
TP 5 3_2 3	CATGGCGCGGACGCGGGTGCCG GGCGGGGGTGTGGAATCA	7711
TP 5 3_2 4	CCAGGGTCCCCAGGCCCTCTGAT TCCTCACTGATTGCTCTT	7712
TP 5 3_2 5	GAGGGCCACTGACAACCACCT TAACCCCTCCTCCCAGAG	7713
TP 5 3_2 6	CCTCAGGCGGCTCATAGGGCAC CACCAACTATGTCGAAA	7714
TP 5 3_2 7	AGGAAATTGCGTGTGGAGTAT TTGGATGACAGAAACACT	7715
TP 5 3_2 8	CTTGCACAGGTCTCCCCAAGG CGCACTGGCCTCATCTTG	7716
TP 5 3_2 9	GAGGCAAGCAGAGGCTGGGGCA CAGCAGGCCAGTGTGCAAG	7717
TP 5 3_3 0	CCTGGAGTCTTCCAGTGTGATG ATGGTGAGGATGGGCCCTC	7718
TP 5 3_3 1	ACTACATGTTAACAGTTCTG CATGGGCGGCCATGAACCG	7719
TP 5 3_3 2	GGACAGGTAGGACCTGATTTCC TTACTGCCTCTTGCTTCT	7720
TP 5 3_3 3	CTGCACCCCTTGGTCTCCCTCCAC CGCTTCTTGTCCTGCTTG	7721
TP 5 3_3 4	TCTCTTTCCATCCTGAGTAG TGGTAATCTACTGGGACG	7722

10

20

30

40

【表 2 - 3】

TP 5 3_3 5	C C T C G C T T A G T G C T C C C T G G G G G C A G C T C G T G G T G A G G C T	7 7 2 3
TP 5 3_3 6	G A C C G G C G C A C A G A G G A A G A G A A T C T C C G C A A G A A A G G G G	7 7 2 4
TP 5 3_3 7	T C T C C C A G G A C A G G C A C A A A C A C G C A C C T C A A A G C T G T T C	7 7 2 5
TP 5 3_3 8	T G C C T C A G A T T C A C T T T T A T C A C C T T T C C T T G C C T C T T T C	7 7 2 6
TP 5 3_3 9	G G C A T T T T G A G T G T T A G A C T G G A A A C T T T C C A C T T G A T A A	7 7 2 7
TP 5 3_4 0	C C T G A A G G G T G A A A T A T T C T C C A T C C A G T G G T T T C T T C T T	7 7 2 8
TP 5 3_4 1	C C T A G C A C T G C C C A A C A A C A C C A G C T C C T C T C C C C A G C C A	7 7 2 9
TP 5 3_4 2	C A T C T T T A A C T C A G G T A C T G T G T A T A T A C T T A C T T C T C C	7 7 3 0
TP 5 3_4 3	A T G G C T T T C C A A C C T A G G A A G G C A G G G G A G T A G G G C C A G G	7 7 3 1
TP 5 3_4 4	C C T G G A G T G A G C C C T G C T C C C C C C T G G C T C C T T C C C C A G C C	7 7 3 2
TP 5 3_4 5	T C C G A G A G C T G A A T G A G G C C T T G G A A C T C A A G G A T G C C C A	7 7 3 3

【0215】

キャプチャされたゲノム配列（図 11C）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 24）。図 24 に示されるデータを、16 個の健康なドナーの c f DNA 試料を解析したシーケンシングランから導出した。TP 53 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 24 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、ユニークリード深度を有した。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 24B）。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 24C 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リのゲノム深度に比例したことを示す。

10

20

30

40

50

【 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。

【 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。

【 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。

10

20

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	G T G T C T C G G C A A C C A C T C T T C A C C A A T A T C A C A G T G G A C A	7 7 3 4
Chr_1 _2	A T C C A A G G G G A G G A G A T C A G T G C C C C T A T T G T A T C G C A C	7 7 3 5
Chr_1 _3	A C T T A C T G A A G C A A G A A C C T C A T C A A G C T G C C T C C C A C C A	7 7 3 6
Chr_1 _4	A G T T T G T G A T C C T C C T G T G G G C A A C C T C A G C A G T C T G G T T	7 7 3 7
Chr_1 _5	G G A G A G C G G G A G C T G C T C A G A G C T T G G C C A G G T T C T A A G T G	7 7 3 8
Chr_1 _6	G A C T G T G G C A A T G A G G G C A G C T A A G T G G T T C A C C A A C T T C T	7 7 3 9
Chr_1 _7	G G T G T A T T T G A C A A C G G T G G A C C C A G A C A C T G G G A G T C A T	7 7 4 0
Chr_1 _8	G T T G G T C T A T T C T T G C G G T T G T A A A A G T G G C C C A G A G T G A	7 7 4 1
Chr_1 _9	G T G A G C C T T C T C A C C A T T C T G T C C A A A A T A G C A G C C C T	7 7 4 2
Chr_1 _10	C A G C C T A G A T A T G A T T C C T C A C T A C C C T G T T C C A T G G T T C	7 7 4 3
Chr_1 _11	A A A G A A T G T G T T G G C T C A T G A T C A G A C T T G A G C A C T T G G G	7 7 4 4
Chr_1 _12	C C T A G G G C T G T T G C T G C T G G A C C T G T T T G T G C T T C A T C A C A	7 7 4 5
Chr_2 _1	C A G T T G A C C C T T C A G C C A C A G G G G T T T G A A C T T T G A A G G A	7 7 4 6
Chr_2 _2	A G G A C C T G A G T A T G C A C G T T T T G G T A T A C T G G G T A G G G G T	7 7 4 7
Chr_2 _3	T A T C A G C T G G G A T G G T C C G G T C A G C A G C A T T A C C C T G T T T	7 7 4 8
Chr_2 _4	T G C C T G C T C A G C C C A G A T T T C A G T C A T G C T G G C C A T A A A C	7 7 4 9
Chr_2 _5	C T G G G G G G T G A G G T T T G A G G T T T G A G T G T G G G A T G T G A G G	7 7 5 0
Chr_2 _6	C C A G C T T T T C A G A A G C T G G G A A A G T A A T A A C C C G T G T T G	7 7 5 1
Chr_2 _7	C C C A G C G C C C G T G G C T T T G G C T C C T C A G T C C C A T T T A A A T	7 7 5 2
Chr_2 _8	T A T A C C A C C A A G T C T A C C T A C T G C C T G C A C A T G C T A T G G C	7 7 5 3
Chr_2 _9	G G T C A A T C C G G C A C T A C T G G T T G T C C A A A G G G A G G T T A C T	7 7 5 4
Chr_2 _10	A A T C A A A C A T C A G G A C C G C C C A C A G C A C A G G T C A A T G A A C	7 7 5 5
Chr_2 _11	G T G T C T C C T G G A G G T G C A T G G G T G G T T T G A A C T T C A T T G	7 7 5 6

10

20

30

40

50

【表3 - 2】

Chr_2 _12	GACCCATGTAAGGGGTTGGGTT ATGTTCTCCTTTGCCCA	7757
Chr_2 _13	TCACTGACATGCGAAGCTGGGA ACGAGAAAATGCACATCC	7758
Chr_2 _14	TCCTACAGTGCTTAGGGATGAA TCTGGCAAAGAAGGATGC	7759
Chr_2 _15	GAAAGCAGTCCTTACCAACAAGA AGACCCCGATGTGGTGGT	7760
Chr_2 _16	ATTGCTCACTGGCTGGCTTGCA TTTGGTATGCGATTGGGA	7761
Chr_2 _17	GTCCCTGGGACCATCTGTCAT TGTTCCTGTAACGGAAA	7762
Chr_2 _18	GACCGAATGGCGAACCGCAGTGA ATAGATCAGGAGGGAAAAA	7763
Chr_3 _1	GAAGGAATGGAGTGGAACAGAT AGGGGTGAGGGAATAACG	7764
Chr_3 _2	CCACTGCCATCCTCAGAGGGAG ATTCACAAGTCTCACAAAT	7765
Chr_3 _3	ATCCAGGCTTCATGTTCAAATG CAATGGCCCTTGCCCCAT	7766
Chr_3 _4	AAATTTCCTGGCTCCCTACT GCTTTGCAGGCCAAGTAA	7767
Chr_3 _5	ACCTTAAAGACGGGCCACATC TCCTTGGATGGGATTAGG	7768
Chr_3 _6	GGGCTTCGGTTTGGCGAAGGT GCTCACAAATCTTGATATC	7769
Chr_3 _7	TGAGCTGTCCTTCATGCCCTGCA TTTCCCAGTCTGTCTTC	7770
Chr_3 _8	ATCTTATCCAGGGCTACCACT GGTGGGTCCAAAATGACT	7771
Chr_3 _9	TACAGGTGAAGGATGTCACAGA GTTTGCTCCCACCTCAA	7772
Chr_3 _10	GCTGTTGTGACGGAGGGCAAGA TCTATGACAGCATCTGC	7773
Chr_3 _11	AATGAAGGGGATTCAAGCCTTG CCACCGACTTACAGGAAG	7774
Chr_3 _12	TGTGAGCGTACTTCTCCCCCA GGTTGAAGAGGAATGAGT	7775
Chr_4 _1	ATTCCAAGTCCAGGTCCCAAAT CTATCAGTACCGGGCTGGC	7776
Chr_4 _2	GACACAGAGTGCATGAAGACCG TTCAAATATGTCAGGGAC	7777
Chr_4 _3	CATGAGTCCTCTATGACTCCC TCTCAGACATGCAGGAAG	7778
Chr_4 _4	TTTTTAGGAGACAGGTACCCAC TGTCTGGTGAAGGAGACT	7779
Chr_4 _5	CCCTCTGTTGAGTCGGCTAGGAG ATGCCTCAGTCAACAAAT	7780

10

20

30

40

50

【表3 - 3】

Chr_4 _6	GACAGAAACTTCATACCCAAGA GCTGCTTCTCAGCTGGA	7781
Chr_4 _7	CAGGCCAACTTGGCAAGACCAA GTCAAGCCTCTCATCTCT	7782
Chr_4 _8	CCCTTGCTACCACATCACTGTTGT CATCTGTGCTTGCATTCC	7783
Chr_5 _1	AGGTCTCACTCCAAC TGCCCCCT GTATTAGAGCTAGGCTGC	7784
Chr_5 _2	GAAACCATGCGGGATTCACTCTT TGTCAAGAGTGGAGCGGGCA	7785
Chr_5 _3	TATGAAATTAGGCGGGTGGTTGG ACGTGACTGTGTGTTGAC	7786
Chr_5 _4	TGAAACCTTGCATGACATACTGC GGCTGCCCATTCACTAGG	7787
Chr_5 _5	TGCTTCTTGTTTATAACTCCCC TGGCCACCATCTCGGGCT	7788
Chr_5 _6	ATTCCCCTCTCATTTGTGGTTGG TGGCTGGATATCTGTTCC	7789
Chr_5 _7	AGCATTACAGCATTTCCTGTGGA CTTACCTCTCAGTAGT	7790
Chr_5 _8	AAAATTAAAGGTGGCGGTAA GGCTGAAAGCCAACAGGC	7791
Chr_5 _9	GAGTGTGTCGGTCAGAAGGAAC ACCTGAGAAACCGCTTTA	7792
Chr_5 _10	CATAGCAAATACCTGTCGCTGA GCCAGGAGTAAGTCTGG	7793
Chr_5 _11	AAGAGGCTCTGAGCTCTTGATA GAGGTTACATGGGGAGCA	7794
Chr_5 _12	GGAGACAACCTAGGAGGTTATC TAGACCATTCCCGCCTTC	7795
Chr_5 _13	GTGTTTCCCTCCCAGCATGCACT TTGTGGCTGCCTTCTTT	7796
Chr_5 _14	TGGCTTGTGTAGCGTGTTCAT TTTGGAACCTTGGAGCCG	7797
Chr_5 _15	GACACCTCTGGTGCAGTTTGAA GGCTGGCCGGGAAGGGAT	7798
Chr_5 _16	GTTTCAGATCTTGCACATGGGAG GGATCGACTCGGCCCTTT	7799
Chr_5 _17	TGCCTAAATCAGAAATGGGCTA CTTCCCTTGGCCACATCC	7800
Chr_5 _18	CAATCTACCAACCTCAAGGTTCA CGCGTGGATTCTACACCT	7801
Chr_6 _1	GAGTTTTCTTTCAAGGTAGTCT GAGATGGCCCCGCACCAAG	7802
Chr_6 _2	TACTATAAAGAAGGCACCTCTA GGCTTGGCAAGCACACGT	7803
Chr_6 _3	GGCAGATTGATGGGACTTTAG ACACTTGCTTGTCCCT	7804

10

20

30

40

50

【表3 - 4】

Chr_6 _4	CAAATGTCCCCCATGC 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	7805
Chr_6 _5	A C A T G T G T A A T C T T C T T C C T A G G G C G G C A G A A C T C A T G	7806
Chr_6 _6	C C C G A G G A A A G C T C C T C T T G C T G A C T G T A A T G T A C T G C A	7807
Chr_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	7808
Chr_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 G C T C T C C	7809
Chr_6 _9	A C T G A T G 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	7810
Chr_6 _10	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	7811
Chr_6 _11	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	7812
Chr_7 _1	C A A T T C T T A C C A T C C A C A A A T G G A T C C A G A C A A C T G T T C	7813
Chr_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	7814
Chr_7 _3	T G C T A T A G A C G G C A C A A A C G A C C G C G A G C C A C A A T C A A G C	7815
Chr_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	7816
Chr_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	7817
Chr_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	7818
Chr_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	7819
Chr_7 _8	G T T T G C C A C C A G C G A A G A G A G A G C C A T C C T G G G T A G A A T T G G A	7820
Chr_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	7821
Chr_7 _10	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 G C A T C	7822
Chr_7 _11	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	7823
Chr_7 _12	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	7824
Chr_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	7825
Chr_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	7826
Chr_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	7827
Chr_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	7828

10

20

30

40

50

【表3 - 5】

Chr_8 _5	AGAACGGCTTGCCTGCTTCGG GCAACCTATGGTTCTGA	7829
Chr_8 _6	TGGCTTGGCGCTTTAAGGCCA GACACGGCATTAAGAAC	7830
Chr_8 _7	GCAAGGCAGAGAAAGATGGCTT AGAAACCTCTTCCCCACC	7831
Chr_8 _8	TCAGCTGTGGCCATTGGTGGAT CTCATCCTTAGTACTAGT	7832
Chr_8 _9	CCATGGTTCTGTGAGACTGGTA GAAAGCACAGACCCCTTA	7833
Chr_9 _1	AATGTGCTTATCACTCGTGATG GGGTCCCTGAAGCTGGCAG	7834
Chr_9 _2	AGGGTCTCATTTAACAGACAGCT TGATTTGAGGGTGAGGGGG	7835
Chr_9 _3	CAGTTGCAAACCATACTTCCTT CAGCCCAGTCCTGCTAT	7836
Chr_9 _4	GTCTAAGGGCATCTTACCTCCA AGAACTGCTTGAGGCGTA	7837
Chr_9 _5	TACCTAGGGAATGACCACTAAG CACCATCTCCGTCACTCT	7838
Chr_9 _6	GGAAGAGAGGAGGGTCATCCAG TCAGTTTGCAAGGAATCT	7839
Chr_9 _7	TGCTGCAGTGTGCGGAAGAACCC TACCTGCGTTCTTAGAA	7840
Chr_9 _8	CATCATACCTATGGCATAGCCA TCAGGGCACTGCAGTTG	7841
Chr_9 _9	TATATCTCACGTGACCGAGGAT GGGTCGTGGGCATTACACA	7842
Chr_9 _10	GAAATGGCCATCTATAGGTGGG AACCACTCCAGTGTACACA	7843
Chr_10 _1	GGAAACCTTTCAGTCTCTACTA GAAGCGCGGAGAGAACTC	7844
Chr_10 _2	TCTGGCCGGCATTCAATTAAAGG CCTAAGGATGAAGGGGGT	7845
Chr_10 _3	AGATACCCATATCGTTCTTATC TCAGCGAAACAACTCCCC	7846
Chr_10 _4	CGCAACTCCTCCAGATCGCAGT GGTGCTTCTTCACTTTCA	7847
Chr_10 _5	TGATTCCATGGTTGCCGTATA CTCCATAAGGGGGTACTT	7848
Chr_10 _6	ATACCATATCCGGCTTGGTTAG GAGGAGGTATTACAGGGGG	7849
Chr_10 _7	GTACCTGTTAACCCAGACGCAA TTCCCTCCACAGTACACAG	7850
Chr_10 _1	ATG TGACACTTGCATCCAGGG GGTCACCATCTGTGTATG	7851
Chr_10 _2	CTAGGTCCCTGAAGAGGTGGCAA GGAACCAGGACAGAACAT	7852

10

20

30

40

50

【表3 - 6】

Chr_1 1_3	TCTGTCATTGGTGACGCCATCT AGACTCTTGGCTTTGGGA	7853
Chr_1 1_4	AAGGTATAGAGCTGGGGGGCTT TCCTCGTTATAAGGTGGAG	7854
Chr_1 1_5	CTCCTACGTAGCCGGGTAGAAA CTTATGGCAGAACAGTCAGG	7855
Chr_1 1_6	TGGATTCCCAGGGTTAATTGTG ACCCAT TGCAGGAAGGTG	7856
Chr_1 1_7	AATGCTGTCCTACTATGGTCTG TACCTGTCAGGAGGGTGG	7857
Chr_1 1_8	GTGCACCTGGAGAGCATACAGG GCACTGACTTGTAGATCA	7858
Chr_1 1_9	TTCCATCTCGCATAACCTGCC CTAAACCTCTCTCGGTTTC	7859
Chr_1 1_10	ATGAAGGCCCTGCTTTGAGTTAT CAGATAGGAAGGGGCCAG	7860
Chr_1 1_11	AGGTCATGTCAGGCTTTGGCT GAACCTAGTTTGCCCCAA	7861
Chr_1 2_1	CTGCATTCTCCATGAGTAGAGT ACGAGCCTCATGTTGGTA	7862
Chr_1 2_2	AAGGCTGTCTTCACCAACTGGG TAGGTGTGGATCAAGACC	7863
Chr_1 2_3	CTGACTTTGGTGTTGGGAGTC GGTGGTCCTTCTTCCATT	7864
Chr_1 2_4	ACTGCAGAGGACCAAGACTGGG AAACAAACGATATGGCAGG	7865
Chr_1 2_5	CCTGGCTTAGAACAGTCTGGCCGG TCCTTCTTCAGCTTCTTA	7866
Chr_1 2_6	AATCTCAGAAAGAGTTCCCTGG ACCATGGCAAATGGTGGC	7867
Chr_1 2_7	ACATTATATCCGGTCCAGGAAT ATCTGGCTCAGGCTGGGT	7868
Chr_1 2_8	AAGCACAGGAATGTGCCTCAC ACGACTTCACATGCCCTT	7869
Chr_1 2_9	GGGGGCTTTGCGGGAAAGAGGGG ACTAAACAACCCCTCTGT	7870
Chr_1 2_10	AAAAGAAATGCGATCAGCGCAA CCCATCCGGTGTGGCGCT	7871
Chr_1 2_11	GGCAGTGGTACCATGACATACT TAGCAGAGATGGACTACA	7872
Chr_1 3_1	ATTTCCCATGCGAGAGGTAGCT TGCCCCAGGCTGTTGGATA	7873
Chr_1 3_2	TTCCATGCCGAGTCCTGATGGA AACTAGCACTGAAAGACC	7874
Chr_1 3_3	TCACGGGAGCTTCCTTCAGTGA GTTCTGCGAATCTGAAGC	7875
Chr_1 3_4	TTTCCAGAGATGAAGCACTACC CAGTCTTACCCAAGTTCG	7876

10

20

30

40

50

【表3 - 7】

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

10

20

30

40

50

【表3 - 8】

Chr_1 6_1	G T G G G G A C G G T C A T T A T C A G C T T T C T G G A C A C A C A G A C A G	7 9 0 1
Chr_1 6_2	T G A G A G G C C A A A G A A T A T C A G T T G A C T C T G G A T C A G G G G C	7 9 0 2
Chr_1 6_3	G A G G C T T T T A G G G C A G C G A G A A A A C G G G A A C T T C A T T C C	7 9 0 3
Chr_1 6_4	A G G A C T T C T C T G G A C C T G T G C C T C A A C T A C T C A C C T G G A T	7 9 0 4
Chr_1 6_5	T G G C C A C A A A T G T T G C C T C C A G C T G C T C A A T G T T C T C C A A	7 9 0 5
Chr_1 6_6	C T G G C A T T G G T G A G T A A T A G G A G C C A G A C G G G T C T G T G T T	7 9 0 6
Chr_1 6_7	A T A C T T A C C T G C A C G A G A A T G A G T T T G G A G C G C A A G G G G G	7 9 0 7
Chr_1 6_8	T T C C C C C A G A G A C T C T G T C C A C T A T G G A C A T T A A A T G T G	7 9 0 8
Chr_1 6_9	G T G C T A C C C T C C T C C C T T C A G G T T A T G T G G T C C A G G C T T	7 9 0 9
Chr_1 6_10	T A A G T G G A A C A A C A T T C C C T T C A T T A T A G G C C C T T C G T G G G	7 9 1 0
Chr_1 6_11	G C A A C G T C A A C A A C T A C T A C G T G C A C A A G C G C C T C T A C T G	7 9 1 1
Chr_1 7_1	G C G G A T G T C G T T A T G G G A C A G G T A C A A G T A G A T A A G T T G C	7 9 1 2
Chr_1 7_2	G T G G T C A C C A T C T C T T C A A A C C A T T T G G A C T G G G C T G G T	7 9 1 3
Chr_1 7_3	A A G C C A A G G A G T T C T G A G A G A G C T T A G C T A A G T T C T T C G C	7 9 1 4
Chr_1 7_4	T T T T T T A G T A C C C C A G T G T G T A A G A C C A A C T G A G G G T G G C	7 9 1 5
Chr_1 7_5	G T T G T C A T T G G G G C T A T A G A C A T A A G C A C C T T C C G G A A T C	7 9 1 6
Chr_1 7_6	C T G A G T G T G C G A G G G G A A G A T A T T G G T G A A G A C C T G T T C T	7 9 1 7
Chr_1 7_7	G T C A G A C C C T G T C C T C G T C T C C T T T A C C T T G T C T C G A T T T	7 9 1 8
Chr_1 7_8	T A A A C T A T G C T C G C C A C C A C T C A G C A C T C A C C T C T T G G G C	7 9 1 9
Chr_1 7_9	G G C A A C T T C C T G A G A C A G A T C G G T A A A A A C A A C C C C T T C T	7 9 2 0
Chr_1 7_10	T C A A C T G T A T T C A T C A G A G A G A T G T G G C T T T C C C A G A C A	7 9 2 1
Chr_1 7_11	G T T T C C C T C A T G T T C C C C C A G G T T C T G T C A G G T G A A G C T G	7 9 2 2
Chr_1 8_1	T T A A C C C A T C T C T A C C C G T C C T G T G T C A A G A A C G G A G G C T	7 9 2 3
Chr_1 8_2	C T G C C C A A A A T A G A A A C C G A G G T T C T C C G T G A C C T A C A T C	7 9 2 4

10

20

30

40

50

【表3 - 9】

Chr_1 8_3	T T C C T T T G C A G T A A C A G C G G G A A C A T G A A G C C G C C A C T C T	7 9 2 5
Chr_1 8_4	T G G T T G C C A G T T C A G A C A C C C A G C C A A A T T G C C C T C T C A	7 9 2 6
Chr_1 8_5	T A G T G C A G C T G G C T T T G A G C C T G T T C C C G A A T G T T C A G A T	7 9 2 7
Chr_1 8_6	A G G G T A A T A G C A C C A A G C T C T A G T C T A C C C A C C T C T C T G A	7 9 2 8
Chr_1 8_7	C C G C A T C T C T G G A G T A G G A A T T G A T C A G C C A C C A T A T G G G	7 9 2 9
Chr_1 8_8	C T A T G A G C A T A C T G G G G A G G G A A A C C T C T A A G C G G A A C T T	7 9 3 0
Chr_1 8_9	A A A A A C C T G C A G G A A G G G A G C C T G A A T G C A A C T G T G G G T C	7 9 3 1
Chr_1 8_10	C A G G T G C T C C A A A C C T T C C A G T C T A T G T T G T A G A T T G C A G	7 9 3 2
Chr_1 8_11	G C C A T A C T A A C C T A C T T C T C C T T G A A G C T C T T G G C C C A T C	7 9 3 3
Chr_1 9_1	A C T G T G A G A T A G C C C T C A T C A T C T T C A A C A G C G G C C A A C C G	7 9 3 4
Chr_1 9_2	A G A T A C A C G G T C A C A G A C G C C A T G T G T T G T G G C T T C T G C A	7 9 3 5
Chr_1 9_3	C A C A T C C T C T C A C C T T T C C G A A G G T T G C A G C T C C T T C T C	7 9 3 6
Chr_1 9_4	T C T G T C T C A C C G G T C C C T T C A T T C C T A G G C A A C T G T A G A T	7 9 3 7
Chr_1 9_5	A T A T C A T G G T C T G T A T C C C C C A G G T A C C T T G A C A C A G G G C C	7 9 3 8
Chr_1 9_6	C T C T C C G C C T T C T T T A G A C C T G A G C A T G C A G A A T T C C G A	7 9 3 9
Chr_1 9_7	A A G G C A T T T A A A T G G G A C A G C G T C C C A T G C G T G A C T T C T C	7 9 4 0
Chr_1 9_8	T C T T T C T A A C A G A C G G A A C A G C C T A C A C C T A C A A C C C C G A G	7 9 4 1
Chr_1 9_9	G T C C C A G C C C A A A A G C A T C T T G G G T A A G G A T T T G G G A T C A	7 9 4 2
Chr_1 9_10	G T T G T T C T G G G C C A G T G T T A G T T G C T C A C A T G T C C T G T C T	7 9 4 3
Chr_1 9_11	A A C A T G C C C T C T T A G T C C T G G G C C A T A C C T T A G C C T T G T G C	7 9 4 4
Chr_2 0_1	T A A C C T C C A A A A G A G G T A C C C A T T G G C G C T C A A C C G A A T T	7 9 4 5
Chr_2 0_2	C T A T A T C T C C G A C T A T G C C T T C T T G G G C A C T G C A C T G C T G	7 9 4 6
Chr_2 0_3	T C T A G A T G G A A G C T G T A T C C A A G G A T G C T C C G G A A T G T T G	7 9 4 7
Chr_2 0_4	A T C T T C T C T G C C T G C C G C A C T A G C T T C T T G G T G A C T T C T C	7 9 4 8

10

20

30

40

50

【表3-10】

Chr_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
Chr_2 0_6	A G G T G C T T G 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
Chr_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
Chr_2 0_8	C A T T T G C A G G A G 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
Chr_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
Chr_2 0_10	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
Chr_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
Chr_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
Chr_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
Chr_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
Chr_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
Chr_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
Chr_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
Chr_2 1_8	T C A T G G A A C A T T G G G 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
Chr_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
Chr_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
Chr_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
Chr_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
Chr_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
Chr_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
Chr_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
Chr_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
Chr_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
Chr_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

【0219】

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27C）。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」の所期の値よりもはるかに「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

10

20

30

40

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。

10

【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。

20

【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（单一アレルの減少）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%の各单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。

30

【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、单一のヌクレオチド変異体、单一のヌクレオチド

40

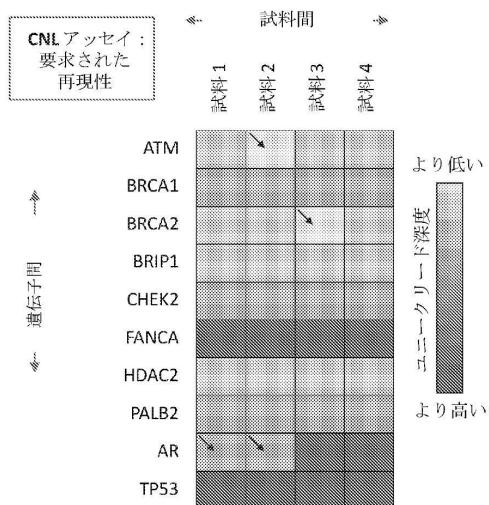
50

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WO 2016 / 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】

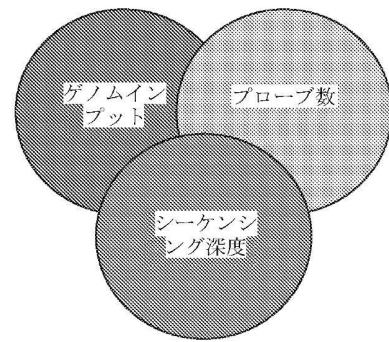
【図 1】



10

【図 2】

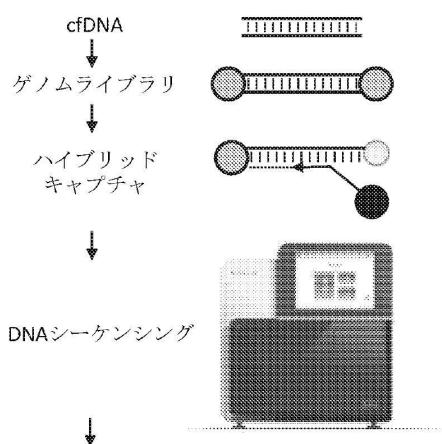
【図 2】



20

【図 3】

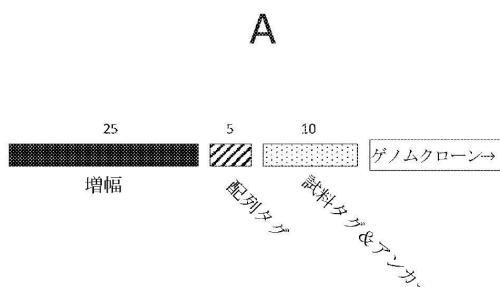
【図 3】



30

【図 4 - 1】

【図 4-1】



40

データ解析

```

TTAGGGCATCGGATCGACATGCC
TGAAACCAAGGATCAACTCCCGTG
GCCACACCTTGATCAGGATATTC
ACCTTGAGTGATCACTGGTCTAG
GCTCAAACAAGGATCAGATGAAA
CCAGCCCGGATTCGACATGAGATG
GACCTTAATGTTGCAACAGTTGA
ACTGTGGGGAATGAGAATGAGCGG
GGCGGGATTGATGCAACATGTC

```

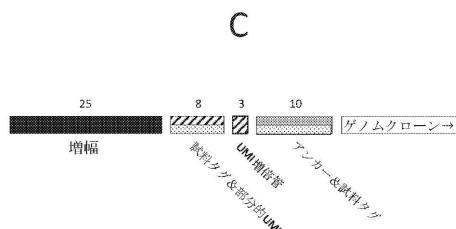
5 10

249個のタグ→1個のアンカ配列

50

【図4-2】

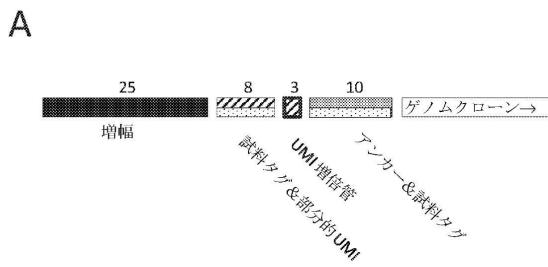
【図4-2】



C

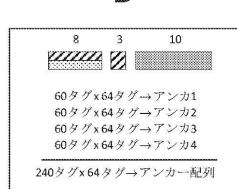
【図5】

【図5】



A

B



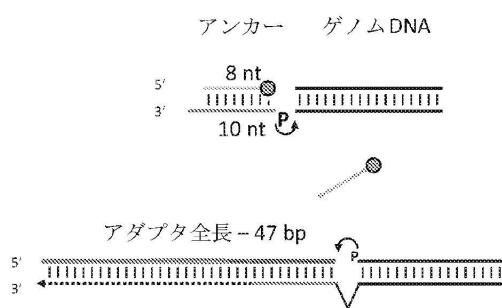
D



E

【図6】

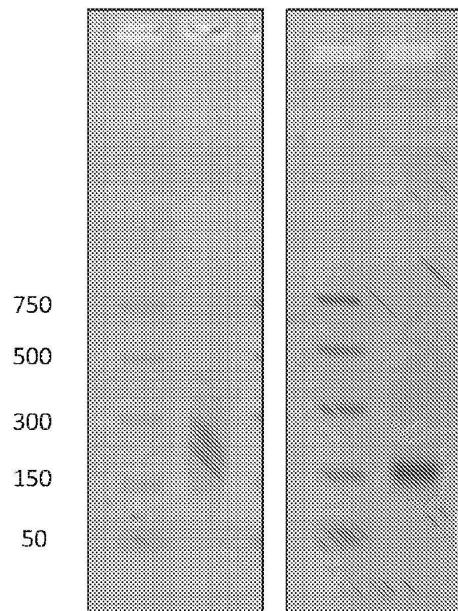
【図6】



【図7】

【図7】

両側gDNA cfDNA



10

20

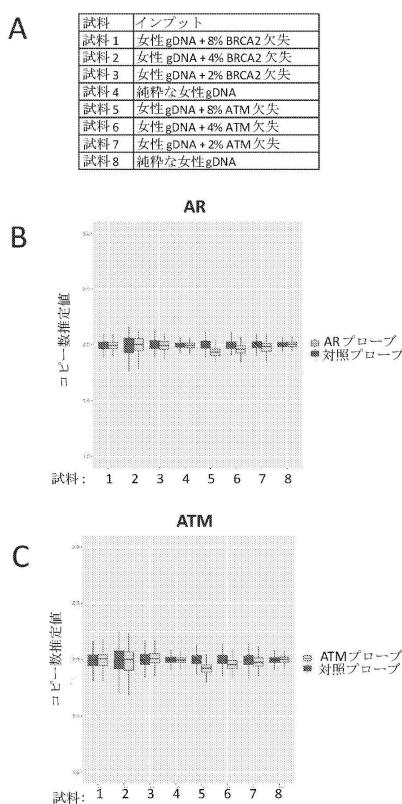
30

40

50

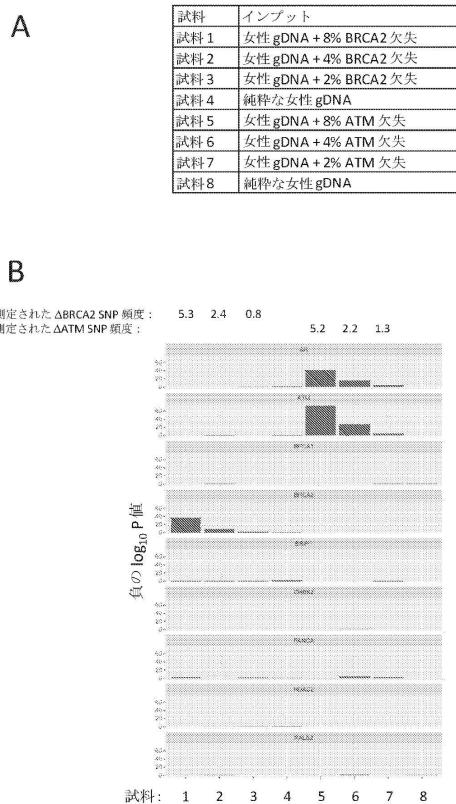
【図 8】

【図 8】



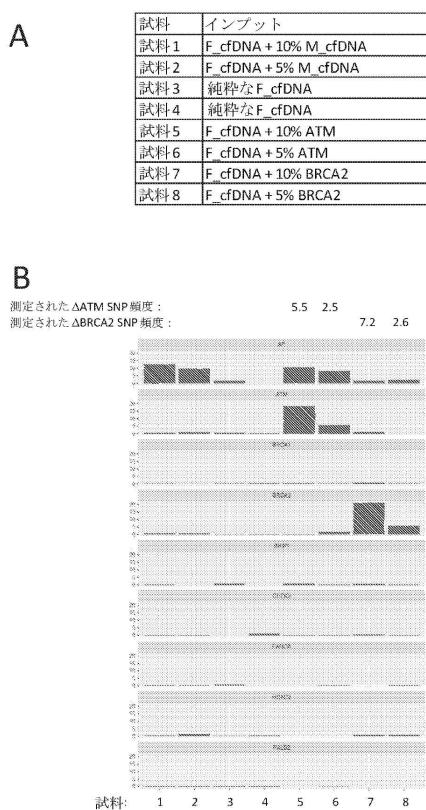
【図 9】

【図 9】



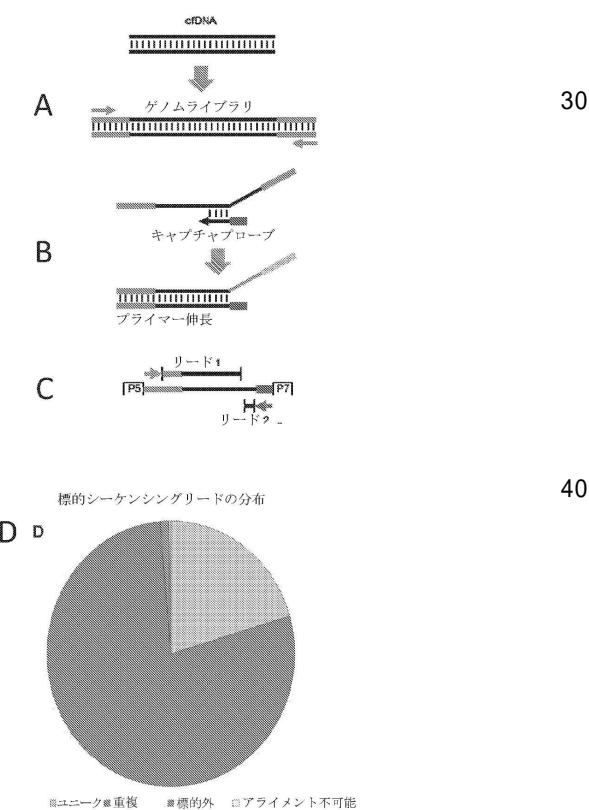
【図 10】

【図 10】



【図 11】

【図 11】



【図12E】

【図12E】

【図12F】

【図12F】

【図12G】

【図12G】

【図12H】

【図12H】

プール1	配列番号:	プール2	配列番号:	プール3	配列番号:
TGCGAGACGAGAAATTCGATA CATCTTGAGNNNTGCATCAAGT	240	TGCGAGACGAGAAATTCGATA CACCGCGAAGNNNTGCATCAAGT	480	TGCGAGACGAGAAATTCGATA CATATCACCGNNNTGCATCAAGT	720
TGCGAGACGAGAAATTCGATA CTGGCGTACAGNNNTGCATCAAGT	241	TGCGAGACGAGAAATTCGATA CACTATCTTNNNTGCATCAAGT	481	TGCGAGACGAGAAATTCGATA CACCTTATGCAGNNNTGCATCAAGT	721
TGCGAGACGAGAAATTCGATA CAAACTGAGNNNTGCATCAAGT	242	TGCGAGACGAGAAATTCGATA CACAACTGAGNNNTGCATCAAGT	482	TGCGAGACGAGAAATTCGATA CATTAATGAGNNNTGCATCAAGT	722
TGCGAGACGAGAAATTCGATA TGCGAGACGAGAAATTCGATA	243	TGCGAGACGAGAAATTCGATA CACCTTAACTGAGNNNTGCATCAAGT	483	TGCGAGACGAGAAATTCGATA CAGACTCTGAGNNNTGCATCAAGT	723
TGCGAGACGAGAAATTCGATA CATTTGTATCCNNNTGCATCAAGT	244	TGCGAGACGAGAAATTCGATA CATTTGTAGAGNNNTGCATCAAGT	484	TGCGAGACGAGAAATTCGATA CGAGAGTCCTNNNTGCATCAAGT	724
TGCGAGACGAGAAATTCGATA CAAACTGAGNNNTGCATCAAGT	245	TGCGAGACGAGAAATTCGATA CAAACTGAGNNNTGCATCAAGT	485	TGCGAGACGAGAAATTCGATA CGAGAGTCCTNNNTGCATCAAGT	725
TGCGAGACGAGAAATTCGATA CGAGAGTCCTNNNTGCATCAAGT	246	TGCGAGACGAGAAATTCGATA CACAGAGTCCTNNNTGCATCAAGT	486	TGCGAGACGAGAAATTCGATA CGAGATGAGCAGNNNTGCATCAAGT	726
TGCGAGACGAGAAATTCGATA CAAACTGAGNNNTGCATCAAGT	247	TGCGAGACGAGAAATTCGATA CAAACTGAGNNNTGCATCAAGT	487	TGCGAGACGAGAAATTCGATA CAACTAGATCTNNNTGCATCAAGT	727
TGCGAGACGAGAAATTCGATA CTTGTGCGTACAGNNNTGCATCAAGT	248	TGCGAGACGAGAAATTCGATA CTTGTGCGTACAGNNNTGCATCAAGT	488	TGCGAGACGAGAAATTCGATA CATTTTGCGTACAGNNNTGCATCAAGT	728

【図13A】

【図13A】

【図13B】

【図13B】

【図13C】

【图13C】

【図13D】

【図13D】

(13 E)

[図13E]

【 义 1 3 F 】

[図1-3F]

ペール4	配列番号:	ペール5	配列番号:	ペール6	配列番号:
TGAGCGACGAGAACTGCAAT	889	TGAGCGACGAGAACTGCAAT	1128	TGAGCGACGAGAACTGCAAT	1369
TGAGCGACGAGAACTGCAAT	890	CAGCGACGAGAACTGCAAT	1129	TCAGCGACGAGAACTGCAAT	1370
TGAGCGACGAGAACTGCAAT	891	TGAGCGACGAGAACTGCAAT	1130	TGAGCGACGAGAACTGCAAT	1370
TGAGCGACGAGAACTGCAAT	892	TGAGCGACGAGAACTGCAAT	1131	TGAGCGACGAGAACTGCAAT	1371
TGAGCGACGAGAACTGCAAT	893	TGAGCGACGAGAACTGCAAT	1132	TGAGCGACGAGAACTGCAAT	1372
TGAGCGACGAGAACTGCAAT	894	TGAGCGACGAGAACTGCAAT	1133	TGAGCGACGAGAACTGCAAT	1373
TGAGCGACGAGAACTGCAAT	895	TGAGCGACGAGAACTGCAAT	1134	TGAGCGACGAGAACTGCAAT	1374
TGAGCGACGAGAACTGCAAT	896	TGAGCGACGAGAACTGCAAT	1135	TGAGCGACGAGAACTGCAAT	1375
TGAGCGACGAGAACTGCAAT	897	TGAGCGACGAGAACTGCAAT	1136	TGAGCGACGAGAACTGCAAT	1376
TGAGCGACGAGAACTGCAAT	898	TGAGCGACGAGAACTGCAAT	1137	TGAGCGACGAGAACTGCAAT	1377
TGAGCGACGAGAACTGCAAT	899	TGAGCGACGAGAACTGCAAT	1138	TGAGCGACGAGAACTGCAAT	1378
TGAGCGACGAGAACTGCAAT	900	TGAGCGACGAGAACTGCAAT	1139	TGAGCGACGAGAACTGCAAT	1379
TGAGCGACGAGAACTGCAAT	901	TGAGCGACGAGAACTGCAAT	1140	TGAGCGACGAGAACTGCAAT	1380
TGAGCGACGAGAACTGCAAT	902	TGAGCGACGAGAACTGCAAT	1141	TGAGCGACGAGAACTGCAAT	1381
TGAGCGACGAGAACTGCAAT	903	TGAGCGACGAGAACTGCAAT	1142	TGAGCGACGAGAACTGCAAT	1382
TGAGCGACGAGAACTGCAAT	904	TGAGCGACGAGAACTGCAAT	1143	TGAGCGACGAGAACTGCAAT	1383
TGAGCGACGAGAACTGCAAT	905	TGAGCGACGAGAACTGCAAT	1144	TGAGCGACGAGAACTGCAAT	1384
TGAGCGACGAGAACTGCAAT	906	TGAGCGACGAGAACTGCAAT	1145	TGAGCGACGAGAACTGCAAT	1385
TGAGCGACGAGAACTGCAAT	907	TGAGCGACGAGAACTGCAAT	1146	TGAGCGACGAGAACTGCAAT	1386
TGAGCGACGAGAACTGCAAT	908	TGAGCGACGAGAACTGCAAT	1147	TGAGCGACGAGAACTGCAAT	1387
TGAGCGACGAGAACTGCAAT	909	TGAGCGACGAGAACTGCAAT	1148	TGAGCGACGAGAACTGCAAT	1388
TGAGCGACGAGAACTGCAAT	910	TGAGCGACGAGAACTGCAAT	1149	TGAGCGACGAGAACTGCAAT	1389
TGAGCGACGAGAACTGCAAT	911	TGAGCGACGAGAACTGCAAT	1150	TGAGCGACGAGAACTGCAAT	1390
TGAGCGACGAGAACTGCAAT	912	TGAGCGACGAGAACTGCAAT	1151	TGAGCGACGAGAACTGCAAT	1391
TGAGCGACGAGAACTGCAAT	913	TGAGCGACGAGAACTGCAAT	1152	TGAGCGACGAGAACTGCAAT	1392
TGAGCGACGAGAACTGCAAT	914	TGAGCGACGAGAACTGCAAT	1153	TGAGCGACGAGAACTGCAAT	1393
TGAGCGACGAGAACTGCAAT	915	TGAGCGACGAGAACTGCAAT	1154	TGAGCGACGAGAACTGCAAT	1394
TGAGCGACGAGAACTGCAAT	916	TGAGCGACGAGAACTGCAAT	1155	TGAGCGACGAGAACTGCAAT	1395
TGAGCGACGAGAACTGCAAT	917	TGAGCGACGAGAACTGCAAT	1156	TGAGCGACGAGAACTGCAAT	1396
TGAGCGACGAGAACTGCAAT	918	TGAGCGACGAGAACTGCAAT	1157	TGAGCGACGAGAACTGCAAT	1397
TGAGCGACGAGAACTGCAAT	919	TGAGCGACGAGAACTGCAAT	1158	TGAGCGACGAGAACTGCAAT	1398
TGAGCGACGAGAACTGCAAT	920	TGAGCGACGAGAACTGCAAT	1159	TGAGCGACGAGAACTGCAAT	1399
TGAGCGACGAGAACTGCAAT	921	TGAGCGACGAGAACTGCAAT	1160	TGAGCGACGAGAACTGCAAT	1400

【図13G】

[図 1 3 G]

【図13H】

【図1.3H】

配列番号:	配列番号:	配列番号:	
ブルー4	ブルー5	ブルー6	
TGCAAGCCAGAAGATCGATA CATGAGCTTGNNTGCAAGG	953	TGCAAGCCAGAAGATCGATA CAACGGCCTGGNNNTGCAAGG	1193
TGCAAGCCAGAAGATCGATA CATGAGCTTGNNTGCAAGG	954	TGCAAGCCAGAAGATCGATA CACTAACTGGNNNTGCAAGG	1194
TGCAAGCCAGAAGATCGATA CATGAGCTTGNNTGCAAGG	955	TGCAAGCCAGAAGATCGATA CACTAACTGGNNNTGCAAGG	1195
TGCAAGCCAGAAGATCGATA CATGAGCTTGNNTGCAAGG	956	TGCAAGCCAGAAGATCGATA CACTAACTGGNNNTGCAAGG	1196
TGCAAGCCAGAAGATCGATA CATGAGCTTGNNTGCAAGG	957	TGCAAGCCAGAAGATCGATA CACTAACTGGNNNTGCAAGG	1197
TGCAAGCCAGAAGATCGATA CATGAGCTTGNNTGCAAGG	958	TGCAAGCCAGAAGATCGATA CACTAACTGGNNNTGCAAGG	1198
TGCAAGCCAGAAGATCGATA CATGAGCTTGNNTGCAAGG	959	TGCAAGCCAGAAGATCGATA CACTAACTGGNNNTGCAAGG	1199
TGCAAGCCAGAAGATCGATA CATGAGCTTGNNTGCAAGG	960	TGCAAGCCAGAAGATCGATA CACTAACTGGNNNTGCAAGG	1200
TGCAAGCCAGAAGATCGATA CATGAGCTTGNNTGCAAGG	961	TGCAAGCCAGAAGATCGATA CACTAACTGGNNNTGCAAGG	1201
TGCAAGCCAGAAGATCGATA CATGAGCTTGNNTGCAAGG	962	TGCAAGCCAGAAGATCGATA CACTAACTGGNNNTGCAAGG	1202
TGCAAGCCAGAAGATCGATA CATGAGCTTGNNTGCAAGG	963	TGCAAGCCAGAAGATCGATA CACTAACTGGNNNTGCAAGG	1203
TGCAAGCCAGAAGATCGATA CATGAGCTTGNNTGCAAGG	964	TGCAAGCCAGAAGATCGATA CACTAACTGGNNNTGCAAGG	1204
TGCAAGCCAGAAGATCGATA CATGAGCTTGNNTGCAAGG	965	TGCAAGCCAGAAGATCGATA CACTAACTGGNNNTGCAAGG	1205
TGCAAGCCAGAAGATCGATA CATGAGCTTGNNTGCAAGG	966	TGCAAGCCAGAAGATCGATA CACTAACTGGNNNTGCAAGG	1206
TGCAAGCCAGAAGATCGATA CATGAGCTTGNNTGCAAGG	967	TGCAAGCCAGAAGATCGATA CACTAACTGGNNNTGCAAGG	1207
TGCAAGCCAGAAGATCGATA CATGAGCTTGNNTGCAAGG	968	TGCAAGCCAGAAGATCGATA CACTAACTGGNNNTGCAAGG	1208

【図14A】

【図14A】

【図14B】

【図14B】

グループ	細胞型番号	グループB	細胞型番号	グループC	細胞型番号
TGCGACGCCAGAAGATGATA	1477	TGCGAGGACGAGAAGATGATA	1717	TGCGACGCCAGAAGATGATA	1951
GAAAGCCTGGAAGATGATA		CACTCGCTTGANNNAGATGCGA		CACATGAGAGAGNNAGATGCGA	
TGCGACGCCAGAAGATGATA	1478	TGCGAGGACGAGAAGATGATA	1718	TGCGACGCCAGAAGATGATA	1958
GAAAGCCTGGAAGATGATA		CAACAGCTTCTNNNAGATGCGA		CACTGTATTCANNNAGATGCGA	
TGCGACGCCAGAAGATGATA	1479	TGCGAGGACGAGAAGATGATA	1719	TGCGACGCCAGAAGATGATA	1959
GAAAGCCTGGAAGATGATA		CAGCTGATTAACNNNAGATGCGA		CACAGCAGGAGNNAGATGCGA	
TGCGACGCCAGAAGATGATA	1480	TGCGAGGACGAGAAGATGATA	1720	TGCGACGCCAGAAGATGATA	1960
GAAAGCCTGGAAGATGATA		CACTGGCTTGTNNNAGATGCGA		CAACAGCAGGAGNNAGATGCGA	
TGCGACGCCAGAAGATGATA	1481	TGCGAGGACGAGAAGATGATA	1721	TGCGACGCCAGAAGATGATA	1961
GAAAGCCTGGAAGATGATA		CACTTATGTTNNNAGATGCGA		CAACAGCAGGAGNNAGATGCGA	
TGCGACGCCAGAAGATGATA	1482	TGCGAGGACGAGAAGATGATA	1722	TGCGACGCCAGAAGATGATA	1962
GAAAGCCTGGAAGATGATA		CAAGCGACCTNNNAGATGCGA		CACTGTATTCANNNAGATGCGA	
TGCGACGCCAGAAGATGATA	1483	TGCGAGGACGAGAAGATGATA	1723	TGCGACGCCAGAAGATGATA	1963
GAAAGCCTGGAAGATGATA		CATTGCTTCTNNNAGATGCGA		CAACAGCAGGAGNNAGATGCGA	
TGCGACGCCAGAAGATGATA	1484	TGCGAGGACGAGAAGATGATA	1724	TGCGACGCCAGAAGATGATA	1964
GATGATTTTGTTNNNAGATGCGA		CATTTTGTTGATNNNAGATGCGA		CAACAGCAGGAGNNAGATGCGA	
TGCGACGCCAGAAGATGATA	1485	TGCGAGGACGAGAAGATGATA	1725	TGCGACGCCAGAAGATGATA	1965
GAAAGCCTGGAAGATGATA		CACTGAGCTTNNNAGATGCGA		CAACAGCAGGAGNNAGATGCGA	
TGCGACGCCAGAAGATGATA	1486	TGCGAGGACGAGAAGATGATA	1726	TGCGACGCCAGAAGATGATA	1966
GAAAGCCTGGAAGATGATA		CAACAGCTTCTNNNAGATGCGA		CAACAGCAGGAGNNAGATGCGA	
TGCGACGCCAGAAGATGATA	1487	TGCGAGGACGAGAAGATGATA	1727	TGCGACGCCAGAAGATGATA	1967
GAAAGCCTGGAAGATGATA		CAAGATTTTNNNAGATGCGA		CACTGGTTTCTNNNAGATGCGA	
TGCGACGCCAGAAGATGATA	1488	TGCGAGGACGAGAAGATGATA	1728	TGCGACGCCAGAAGATGATA	1968
GAAAGCCTGGAAGATGATA		CAACGGCTTCTNNNAGATGCGA		CAACAGCAGGAGNNAGATGCGA	
TGCGACGCCAGAAGATGATA	1489	TGCGAGGACGAGAAGATGATA	1729	TGCGACGCCAGAAGATGATA	1969
GAAAGCCTGGAAGATGATA		CACTAGCTTCTNNNAGATGCGA		CAACAGCAGGAGNNAGATGCGA	
TGCGACGCCAGAAGATGATA	1490	TGCGAGGACGAGAAGATGATA	1730	TGCGACGCCAGAAGATGATA	1970
GAAAGCCTGGAAGATGATA		CACTGAGCTTNNNAGATGCGA		CAACAGCAGGAGNNAGATGCGA	
TGCGACGCCAGAAGATGATA	1491	TGCGAGGACGAGAAGATGATA	1731	TGCGACGCCAGAAGATGATA	1971
GAAAGCCTGGAAGATGATA		CACTAGCTTCTNNNAGATGCGA		CACTGTATTCANNNAGATGCGA	
TGCGACGCCAGAAGATGATA	1492	TGCGAGGACGAGAAGATGATA	1732	TGCGACGCCAGAAGATGATA	1972
GAAAGCCTGGAAGATGATA		CAAGATACCTNNNAGATGCGA		CAACAGCAGGAGNNAGATGCGA	
TGCGACGCCAGAAGATGATA	1493	TGCGAGGACGAGAAGATGATA	1733	TGCGACGCCAGAAGATGATA	1973
GAAAGCCTGGAAGATGATA		CACTGATTTTNNNAGATGCGA		CACTGGCTTCTNNNAGATGCGA	
TGCGACGCCAGAAGATGATA	1494	TGCGAGGACGAGAAGATGATA	1734	TGCGACGCCAGAAGATGATA	1974
GAAAGCCTGGAAGATGATA		CACTATGTTCTNNNAGATGCGA		CACTGTATTCANNNAGATGCGA	
TGCGACGCCAGAAGATGATA	1495	TGCGAGGACGAGAAGATGATA	1735	TGCGACGCCAGAAGATGATA	1975
GAAAGCCTGGAAGATGATA		CACTGAGCTTNNNAGATGCGA		CAACAGCAGGAGNNAGATGCGA	
TGCGACGCCAGAAGATGATA	1496	TGCGAGGACGAGAAGATGATA	1736	TGCGACGCCAGAAGATGATA	1976
GAAAGCCTGGAAGATGATA		CACTTCTTCTNNNAGATGCGA		CACTGGCTTCTNNNAGATGCGA	
TGCGACGCCAGAAGATGATA	1497	TGCGAGGACGAGAAGATGATA	1737	TGCGACGCCAGAAGATGATA	1977
GAAAGCCTGGAAGATGATA		CACTCCCTTCTNNNAGATGCGA		CACTGGCTTCTNNNAGATGCGA	
TGCGACGCCAGAAGATGATA	1498	TGCGAGGACGAGAAGATGATA	1738	TGCGACGCCAGAAGATGATA	1978
GAAAGCCTGGAAGATGATA		CACTAGCTTCTNNNAGATGCGA		CACTGTATTCANNNAGATGCGA	
TGCGACGCCAGAAGATGATA	1499	TGCGAGGACGAGAAGATGATA	1739	TGCGACGCCAGAAGATGATA	1979
GAAAGCCTGGAAGATGATA		CACTGAGCTTNNNAGATGCGA		CACTGTCTTCTNNNAGATGCGA	
TGCGACGCCAGAAGATGATA	1500	TGCGAGGACGAGAAGATGATA	1740	TGCGACGCCAGAAGATGATA	1980
GAAAGCCTGGAAGATGATA		CACTGTATTCANNNAGATGCGA		CACTGGCTTCTNNNAGATGCGA	
TGCGACGCCAGAAGATGATA	1501	TGCGAGGACGAGAAGATGATA	1741	TGCGACGCCAGAAGATGATA	1981
GAAAGCCTGGAAGATGATA		CACTGGCTTCTNNNAGATGCGA		CACTGGCTTCTNNNAGATGCGA	
TGCGACGCCAGAAGATGATA	1502	TGCGAGGACGAGAAGATGATA	1742	TGCGACGCCAGAAGATGATA	1982
GAAAGCCTGGAAGATGATA		CACTATTCCTNNNAGATGCGA		CACTGTATTCANNNAGATGCGA	
TGCGACGCCAGAAGATGATA	1503	TGCGAGGACGAGAAGATGATA	1743	TGCGACGCCAGAAGATGATA	1983
GAAAGCCTGGAAGATGATA		CAACAGCTTCTNNNAGATGCGA		CAACAGCAGGAGNNAGATGCGA	
TGCGACGCCAGAAGATGATA	1504	TGCGAGGACGAGAAGATGATA	1744	TGCGACGCCAGAAGATGATA	1984
GAAAGCCTGGAAGATGATA		CACTTCTGCTTCTNNNAGATGCGA		CACTGGCTTCTNNNAGATGCGA	
TGCGACGCCAGAAGATGATA	1505	TGCGAGGACGAGAAGATGATA	1745	TGCGACGCCAGAAGATGATA	1985

【図14C】

【図14C】

【 义 1 4 D 】

【図14D】

ブルト	配列番号	ブルトB	配列番号	ブルトC	配列番号
TGCAAGGCCAGAAGATGGATA	1535	TGCGAGGCCAGAAGATGGATA	1775	TGCGAGGCCAGAAGATGGATA	2019
GACGAGCCAGAAGATGGATA	1536	CATCGAGCCAGAAGATGGATA	1776	CATCGAGCCAGAAGATGGATA	2016
CGACGAGCCAGAAGATGGATA	1536	CGACGAGCCAGAAGATGGATA	1776	CGACGAGCCAGAAGATGGATA	2016
CGACGAGCCAGAAGATGGATA	1537	TGCGAGGCCAGAAGATGGATA	1777	TGCGAGGCCAGAAGATGGATA	2017
CGACGAGCCAGAAGATGGATA	1537	CGTACAGCCAGCNCNTGGCTTC	1777	CGACGAGCCAGAAGATGGATA	2017
CGACGAGCCAGAAGATGGATA	1538	TGCGAGGCCAGAAGATGGATA	1778	TGCGAGGCCAGAAGATGGATA	2018
GACGAGCCAGAAGATGGATA	1539	CAGTGAACTGGTACGGCTTC	1779	CAGTGAACTGGTACGGCTTC	2019
CGACGAGCCAGAAGATGGATA	1539	TGCGAGGCCAGAAGATGGATA	1779	TGCGAGGCCAGAAGATGGATA	2019
CGACGAGCCAGAAGATGGATA	1540	TGCGAGGCCAGAAGATGGATA	1780	TGCGAGGCCAGAAGATGGATA	2020
GACGAGCCAGAAGATGGATA	1541	CGACGAGCCAGAAGATGGATA	1781	CGACGAGCCAGAAGATGGATA	2021
CGACGAGCCAGAAGATGGATA	1542	CGACGAGCCAGAAGATGGATA	1782	CGACGAGCCAGAAGATGGATA	2022
CGACGAGCCAGAAGATGGATA	1543	CACCGACTCAGNNCTGGCTTC	1782	CGACGAGCCAGAAGATGGATA	2022
CGACGAGCCAGAAGATGGATA	1543	TGCGAGGCCAGAAGATGGATA	1783	TGCGAGGCCAGAAGATGGATA	2023
CGACGAGCCAGAAGATGGATA	1544	CGACGAGCCAGAAGATGGATA	1784	CGACGAGCCAGAAGATGGATA	2023
CGACGAGCCAGAAGATGGATA	1545	CGACGAGCCAGAAGATGGATA	1785	CGACGAGCCAGAAGATGGATA	2023
CGACGAGCCAGAAGATGGATA	1546	CGACGAGCCAGAAGATGGATA	1786	CGACGAGCCAGAAGATGGATA	2023
CGACGAGCCAGAAGATGGATA	1547	CGACGAGCCAGAAGATGGATA	1787	CGACGAGCCAGAAGATGGATA	2023
CGACGAGCCAGAAGATGGATA	1547	CGACGAGCCAGAAGATGGATA	1787	CGACGAGCCAGAAGATGGATA	2023
CGACGAGCCAGAAGATGGATA	1548	TGCGAGGCCAGAAGATGGATA	1788	TGCGAGGCCAGAAGATGGATA	2023
CGACGAGCCAGAAGATGGATA	1548	CGACGAGCCAGAAGATGGATA	1788	CGACGAGCCAGAAGATGGATA	2023
CGACGAGCCAGAAGATGGATA	1549	TGCGAGGCCAGAAGATGGATA	1789	TGCGAGGCCAGAAGATGGATA	2023
CGACGAGCCAGAAGATGGATA	1550	TGCGAGGCCAGAAGATGGATA	1790	TGCGAGGCCAGAAGATGGATA	2030
CGACGAGCCAGAAGATGGATA	1551	CGACGAGCCAGAAGATGGATA	1791	CGACGAGCCAGAAGATGGATA	2031
CGACGAGCCAGAAGATGGATA	1552	CGACGAGCCAGAAGATGGATA	1792	CGACGAGCCAGAAGATGGATA	2032
CGACGAGCCAGAAGATGGATA	1552	TGCGAGGCCAGAAGATGGATA	1792	TGCGAGGCCAGAAGATGGATA	2032
CGACGAGCCAGAAGATGGATA	1553	CGACGAGCCAGAAGATGGATA	1793	CGACGAGCCAGAAGATGGATA	2033
CGACGAGCCAGAAGATGGATA	1554	CGACGAGCCAGAAGATGGATA	1794	CGACGAGCCAGAAGATGGATA	2034
CGACGAGCCAGAAGATGGATA	1555	CGACGAGCCAGAAGATGGATA	1795	CGACGAGCCAGAAGATGGATA	2035
CGACGAGCCAGAAGATGGATA	1556	CGACGAGCCAGAAGATGGATA	1796	CGACGAGCCAGAAGATGGATA	2036
CGACGAGCCAGAAGATGGATA	1557	CGACGAGCCAGAAGATGGATA	1797	CGACGAGCCAGAAGATGGATA	2037
CGACGAGCCAGAAGATGGATA	1558	TGCGAGGCCAGAAGATGGATA	1798	TGCGAGGCCAGAAGATGGATA	2038
CGACGAGCCAGAAGATGGATA	1559	TGCGAGGCCAGAAGATGGATA	1799	TGCGAGGCCAGAAGATGGATA	2039
CGACGAGCCAGAAGATGGATA	1560	TGCGAGGCCAGAAGATGGATA	1800	TGCGAGGCCAGAAGATGGATA	2040

【図14E】

【図14E】

【図 1-4 F】

【図14F】

ゲートル	配列番号	ゲートル	配列番号	ゲートル	配列番号
TGCGAGGACGAGAATGGATA	1593	TGCGAGGACGAGAATGGATA	1833	TGCGAGGACGAGAATGGATA	2073
CAAGCCTGTCTTNNNGAACTGATA		CAAGCCTGTCTTNNNGAACTGATA		CAGTAGCAGGCGNNNGAACGATG	
TGCGAGGACGAGAATGGATA	1594	TGCGAGGACGAGAATGGATA	1834	TGCGAGGACGAGAATGGATA	2074
CAACCCGGGAACTTNNNGAACTGATA		CAACCCGGGAACTTNNNGAACTGATA		CAGAATTCAGGCGNNNGAACGATG	
TGCGAGGACGAGAATGGATA	1595	TGCGAGGACGAGAATGGATA	1835	TGCGAGGACGAGAATGGATA	2075
CAACCCGGGAACTTNNNGAACTGATA		CAACCCGGGAACTTNNNGAACTGATA		CAGAACGCCCTNNNGAACGATG	
TGCGAGGACGAGAATGGATA	1596	TGCGAGGACGAGAATGGATA	1836	TGCGAGGACGAGAATGGATA	2076
CAACCCGGGAACTTNNNGAACTGATA		CAACCCGGGAACTTNNNGAACTGATA		CAGAACGCCCTNNNGAACGATG	
TGCGAGGACGAGAATGGATA	1597	TGCGAGGACGAGAATGGATA	1837	TGCGAGGACGAGAATGGATA	2077
CAATGGATCTGGNNNGAACGATG		CAATGGATCTGGNNNGAACGATG		CAGTCCTTCTTNNNGAACGATG	
TGCGAGGACGAGAATGGATA	1598	TGCGAGGACGAGAATGGATA	1838	TGCGAGGACGAGAATGGATA	2078
CAAGCAGGCGNNNGAACGATG		CAAGCAGGCGNNNGAACGATG		CABARCTGCCTNNNGAACGATG	
TGCGAGGACGAGAATGGATA	1599	TGCGAGGACGAGAATGGATA	1839	TGCGAGGACGAGAATGGATA	2079
CAAGCAGGCGNNNGAACGATG		CAAGCAGGCGNNNGAACGATG		CAGAACGCCCTNNNGAACGATG	
TGCGAGGACGAGAATGGATA	1600	TGCGAGGACGAGAATGGATA	1840	TGCGAGGACGAGAATGGATA	2080
CAAGCAGGCGNNNGAACGATG		CAAGCAGGCGNNNGAACGATG		CAGAACGCCCTNNNGAACGATG	
TGCGAGGACGAGAATGGATA	1601	TGCGAGGACGAGAATGGATA	1841	TGCGAGGACGAGAATGGATA	2081
CAAGCAGGCGNNNGAACGATG		CAAGCAGGCGNNNGAACGATG		CAGTCCTTCTTNNNGAACGATG	
TGCGAGGACGAGAATGGATA	1602	TGCGAGGACGAGAATGGATA	1842	TGCGAGGACGAGAATGGATA	2082
CAACCCGGGAACTTNNNGAACTGATA		CAACCCGGGAACTTNNNGAACTGATA		CAGAACGCCCTNNNGAACGATG	
TGCGAGGACGAGAATGGATA	1603	TGCGAGGACGAGAATGGATA	1843	TGCGAGGACGAGAATGGATA	2083
CAATGGATCTGGNNNGAACGATG		CAATGGATCTGGNNNGAACGATG		CAGTAATCCGGNNNGAACGATG	
TGCGAGGACGAGAATGGATA	1604	TGCGAGGACGAGAATGGATA	1844	TGCGAGGACGAGAATGGATA	2084
CAAGCAGGCGNNNGAACGATG		CAAGCAGGCGNNNGAACGATG		CAGAACGCCCTNNNGAACGATG	
TGCGAGGACGAGAATGGATA	1605	TGCGAGGACGAGAATGGATA	1845	TGCGAGGACGAGAATGGATA	2085
CAACCCGGGAACTTNNNGAACTGATA		CAACCCGGGAACTTNNNGAACTGATA		CATAGATGGCTTNNNGAACGATG	
TGCGAGGACGAGAATGGATA	1606	TGCGAGGACGAGAATGGATA	1846	TGCGAGGACGAGAATGGATA	2086
CAATGGATCTGGNNNGAACGATG		CAATGGATCTGGNNNGAACGATG		CAGAACGCCCTNNNGAACGATG	
TGCGAGGACGAGAATGGATA	1607	TGCGAGGACGAGAATGGATA	1847	TGCGAGGACGAGAATGGATA	2087
CACTGGCTGTNNNGAACGATG		CACTGGCTGTNNNGAACGATG		CATGGCAGCGNNNGAACGATG	
TGCGAGGACGAGAATGGATA	1608	TGCGAGGACGAGAATGGATA	1848	TGCGAGGACGAGAATGGATA	2088
CAAGCAGGCGNNNGAACGATG		CAAGCAGGCGNNNGAACGATG		CAGTTTGCAGTNNNGAACGATG	
TGCGAGGACGAGAATGGATA	1609	TGCGAGGACGAGAATGGATA	1849	TGCGAGGACGAGAATGGATA	2089
CAAGCAGGCGNNNGAACGATG		CAAGCAGGCGNNNGAACGATG		CAGAACGCCCTNNNGAACGATG	
TGCGAGGACGAGAATGGATA	1610	TGCGAGGACGAGAATGGATA	1850	TGCGAGGACGAGAATGGATA	2090
CACTTCTCCTTNNNGAACGATG		CACTTCTCCTTNNNGAACGATG		CACCCAGCCCTNNNGAACGATG	
TGCGAGGACGAGAATGGATA	1611	TGCGAGGACGAGAATGGATA	1851	TGCGAGGACGAGAATGGATA	2091
CACTCTCTCTTNNNGAACGATG		CACTCTCTCTTNNNGAACGATG		CATGGAGAGCTTNNNGAACGATG	
TGCGAGGACGAGAATGGATA	1612	TGCGAGGACGAGAATGGATA	1852	TGCGAGGACGAGAATGGATA	2092
CAATTCAGGGAACTTNNNGAACTGATA		CAATTCAGGGAACTTNNNGAACTGATA		CATTTGAGAGCTTNNNGAACGATG	
TGCGAGGACGAGAATGGATA	1613	TGCGAGGACGAGAATGGATA	1853	TGCGAGGACGAGAATGGATA	2093
CAATCATATCTTNNNGAACGATG		CAATCATATCTTNNNGAACGATG		CAGCTTTCAGTNNNGAACGATG	
TGCGAGGACGAGAATGGATA	1614	TGCGAGGACGAGAATGGATA	1854	TGCGAGGACGAGAATGGATA	2094
CAATCATATCTTNNNGAACGATG		CAATCATATCTTNNNGAACGATG		CAGACAGCTTNNNGAACGATG	
TGCGAGGACGAGAATGGATA	1615	TGCGAGGACGAGAATGGATA	1855	TGCGAGGACGAGAATGGATA	2095
CACTCTCTCTTNNNGAACGATG		CACTCTCTCTTNNNGAACGATG		CAGACGGCTTNNNGAACGATG	
TGCGAGGACGAGAATGGATA	1616	TGCGAGGACGAGAATGGATA	1856	TGCGAGGACGAGAATGGATA	2096
CAAGCAGGCGNNNGAACGATG		CAAGCAGGCGNNNGAACGATG		CATGGAGCTGTNNNGAACGATG	
TGCGAGGACGAGAATGGATA	1617	TGCGAGGACGAGAATGGATA	1857	TGCGAGGACGAGAATGGATA	2097
CAAGCAGGCGNNNGAACGATG		CAAGCAGGCGNNNGAACGATG		CACACAGCGNNNGAACGATG	
TGCGAGGACGAGAATGGATA	1618	TGCGAGGACGAGAATGGATA	1858	TGCGAGGACGAGAATGGATA	2098
CAAGCAGGCGNNNGAACGATG		CAAGCAGGCGNNNGAACGATG		CAGAACGCCCTNNNGAACGATG	
TGCGAGGACGAGAATGGATA	1619	TGCGAGGACGAGAATGGATA	1859	TGCGAGGACGAGAATGGATA	2099
CAAGCAGGCGNNNGAACGATG		CAAGCAGGCGNNNGAACGATG		CATTTGAGAGCTTNNNGAACGATG	
TGCGAGGACGAGAATGGATA	1620	TGCGAGGACGAGAATGGATA	1860	TGCGAGGACGAGAATGGATA	2100
CAACTCTCTCTTNNNGAACGATG		CAACTCTCTCTTNNNGAACGATG		CACCTAGCTTNNNGAACGATG	
TGCGAGGACGAGAATGGATA	1621	TGCGAGGACGAGAATGGATA	1861	TGCGAGGACGAGAATGGATA	2101
CAAGCAGGCGNNNGAACGATG		CAAGCAGGCGNNNGAACGATG		CACACAGCGNNNGAACGATG	

【図14G】

【図14G】

【図 1-4-H】

[図1-4H]

【図14-I】

【図14-1】

グループ	配列名	指紋:	グループ	配列名	指紋:	グループ	配列名	指紋:
	TOGAGCGGCAGCAATTGGATATA	1680		TGAGGACAGAACGAAATTGCTGATTA	1921		TGCGAGGACAGAACGAAATTGCTGATTA	2160
	AAGGCTTGCTTNNNTGTCATGAGG			CACCTTGATCCTNNNTGTCATGAGG			CATCTAGGAGGNNNTGTCATGAGG	
1	TGAGGACAGAACGAAATTGCTGATTA	1681		TGAGGACAGAACGAAATTGCTGATTA	1921		TGCGAGGACAGAACGAAATTGCTGATTA	2161
	GCGACGAGCTTGTGTCATGAGG			CACCTTGATCCTNNNTGTCATGAGG			CATCTAGGAGGNNNTGTCATGAGG	
	TGAGGACAGAACGAAATTGCTGATTA	1682		TGAGGACAGAACGAAATTGCTGATTA	1922		TGCGAGGACAGAACGAAATTGCTGATTA	2162
	CATTCAGCGAANNNTGTCATGAGG			CACCTTGATCCTNNNTGTCATGAGG			CACCTGCTTGTGTCATGAGG	
	TGAGGACAGAACGAAATTGCTGATTA	1683		TGAGGACAGAACGAAATTGCTGATTA	1923		TGCGAGGACAGAACGAAATTGCTGATTA	2163
	CACCTTGATCCTNNNTGTCATGAGG			CACCTTGATCCTNNNTGTCATGAGG			CATCTAGGAGGNNNTGTCATGAGG	
	TGAGGACAGAACGAAATTGCTGATTA	1684		TGAGGACAGAACGAAATTGCTGATTA	1924		TGCGAGGACAGAACGAAATTGCTGATTA	2164
	GCGACATCTTNNNTGTCATGAGG			CATGGTTTCTNNNTGTCATGAGG			CACGCGGTCAGCCTNNNTGTCATGAGG	
	TGAGGACAGAACGAAATTGCTGATTA	1685		TGAGGACAGAACGAAATTGCTGATTA	1925		TGCGAGGACAGAACGAAATTGCTGATTA	2165
	CACCTTGATCCTNNNTGTCATGAGG			GAGGTTGTTNNNTGTCATGAGG			GTTGTTGTTNNNTGTCATGAGG	
	TGAGGACAGAACGAAATTGCTGATTA	1686		TGAGGACAGAACGAAATTGCTGATTA	1926		TGCGAGGACAGAACGAAATTGCTGATTA	2166
	CATTCAGCGAANNNTGTCATGAGG			CACCTTGATCCTNNNTGTCATGAGG			CACCTGCTTGTGTCATGAGG	
	TGAGGACAGAACGAAATTGCTGATTA	1687		TGAGGACAGAACGAAATTGCTGATTA	1927		TGCGAGGACAGAACGAAATTGCTGATTA	2167
	CATTCAGCGAANNNTGTCATGAGG			CACCTTGATCCTNNNTGTCATGAGG			CAAAACACTCCNNNTGTCATGAGG	
	TGAGGACAGAACGAAATTGCTGATTA	1688		TGAGGACAGAACGAAATTGCTGATTA	1928		TGCGAGGACAGAACGAAATTGCTGATTA	2168
	AAGGCTTGCTTNNNTGTCATGAGG			CACGCTGTTGTTNNNTGTCATGAGG			CAGAGAAACACCCNNNTGTCATGAGG	

【図 15A】

【図15A】

【図15B】

【図15B】

【図15C-1】

【図15C】

【図 15C-2】

【図15C】

【図 15D】

【図15D】

【図15E】

【图 15E】

【図 15 F】

【図15F】

【図15G】

【図15G】

【図16A】

【図16A】

【 図 1 6 B 】

【図16B】

【図16C】

【図16C】

【図16D】

【図16D】

【図16E】

【図16E】

【図16F】

【図16F】

【図16G】

【図16G】

【図16H】

【図16H】

【図17A】

【図17A】

(四 17 B)

【図17B】

【図 1 7 C】

【図17C】

【図17D】

【図17D】

【図17E】

【图 17E】

【図17F】

【図17F】

【図17H】

【図17H】

プール16	配列番号	プール17	配列番号	プール18	配列番号	プール19	配列番号	プール20	配列番号	プール21	配列番号
TGCAAGGCCAAGAAATGCTAATCAG	3835	TGCAAGGCCAAGAAATGCTAATCAG	1075	TGCAAGGCCAAGAAATGCTAATCAG	4319	TGCAAGGCCAAGAAATGCTAATCAG	4329	TGCAAGGCCAAGAAATGCTAATCAG	4569	TGCAAGGCCAAGAAATGCTAATCAG	4809
TGCAAGGCCAAGAAATGCTAATCAG	3840	TGCAAGGCCAAGAAATGCTAATCAG	4080	TGCAAGGCCAAGAAATGCTAATCAG	4320	TGCAAGGCCAAGAAATGCTAATCAG	4330	TGCAAGGCCAAGAAATGCTAATCAG	4579	TGCAAGGCCAAGAAATGCTAATCAG	4810
TGCAAGGCCAAGAAATGCTAATCAG	3841	TGCAAGGCCAAGAAATGCTAATCAG	4081	TGCAAGGCCAAGAAATGCTAATCAG	4321	TGCAAGGCCAAGAAATGCTAATCAG	4331	TGCAAGGCCAAGAAATGCTAATCAG	4571	TGCAAGGCCAAGAAATGCTAATCAG	4811
TGCAAGGCCAAGAAATGCTAATCAG	3842	TGCAAGGCCAAGAAATGCTAATCAG	4092	TGCAAGGCCAAGAAATGCTAATCAG	4322	TGCAAGGCCAAGAAATGCTAATCAG	4332	TGCAAGGCCAAGAAATGCTAATCAG	4572	TGCAAGGCCAAGAAATGCTAATCAG	4812
TGCAAGGCCAAGAAATGCTAATCAG	3843	TGCAAGGCCAAGAAATGCTAATCAG	4083	TGCAAGGCCAAGAAATGCTAATCAG	4323	TGCAAGGCCAAGAAATGCTAATCAG	4333	TGCAAGGCCAAGAAATGCTAATCAG	4573	TGCAAGGCCAAGAAATGCTAATCAG	4813
TGCAAGGCCAAGAAATGCTAATCAG	3844	TGCAAGGCCAAGAAATGCTAATCAG	4084	TGCAAGGCCAAGAAATGCTAATCAG	4324	TGCAAGGCCAAGAAATGCTAATCAG	4334	TGCAAGGCCAAGAAATGCTAATCAG	4574	TGCAAGGCCAAGAAATGCTAATCAG	4814
TGCAAGGCCAAGAAATGCTAATCAG	3845	TGCAAGGCCAAGAAATGCTAATCAG	4085	TGCAAGGCCAAGAAATGCTAATCAG	4325	TGCAAGGCCAAGAAATGCTAATCAG	4335	TGCAAGGCCAAGAAATGCTAATCAG	4575	TGCAAGGCCAAGAAATGCTAATCAG	4815
TGCAAGGCCAAGAAATGCTAATCAG	3846	TGCAAGGCCAAGAAATGCTAATCAG	4086	TGCAAGGCCAAGAAATGCTAATCAG	4326	TGCAAGGCCAAGAAATGCTAATCAG	4336	TGCAAGGCCAAGAAATGCTAATCAG	4576	TGCAAGGCCAAGAAATGCTAATCAG	4816
TGCAAGGCCAAGAAATGCTAATCAG	3847	TGCAAGGCCAAGAAATGCTAATCAG	4087	TGCAAGGCCAAGAAATGCTAATCAG	4327	TGCAAGGCCAAGAAATGCTAATCAG	4337	TGCAAGGCCAAGAAATGCTAATCAG	4577	TGCAAGGCCAAGAAATGCTAATCAG	4817
TGCAAGGCCAAGAAATGCTAATCAG	3848	TGCAAGGCCAAGAAATGCTAATCAG	4088	TGCAAGGCCAAGAAATGCTAATCAG	4328	TGCAAGGCCAAGAAATGCTAATCAG	4338	TGCAAGGCCAAGAAATGCTAATCAG	4578	TGCAAGGCCAAGAAATGCTAATCAG	4818

【図 18 A】

【図18A】

【図18B】

【図18B】

【図 1 8 C】

【図18C】

【図18D】

【図18D】

グルー19	配列番号	グルー20	配列番号	グルー21	配列番号
TGCGAGCCGGACCTGGATGATA	4408	CGGAGACGAGAACGATTGATA	4408	TGCAGGACGAGAACGATTGATA	4910
CAGCTTGAGTCAANNCATGGCTAC	4409	CAACAGTCATCANNCTTACGGTAC	4409	GAGCAGTGTGCGNNCTTACGGTAC	4910
TGAGGACGGACGGAACTGGATA	4429	CGGGAGACGAGAACGATTGATA	4469	TGGAGGACGGACGAAGATTGATA	4909
CAGCTTGAGTCAANNCATGGCTAC	4430	CAACAGTCATCANNCTTACGGTAC	4470	CAATTTGAGGACGGACGAAGATTGATA	4910
CAGCTTGAGTCAANNCATGGCTAC	4431	CAACAGTCATCANNCTTACGGTAC	4471	TGGAGGACGGACGAAGATTGATA	4911
CAAGATGGACGGACGGAACTGGATA	4432	CGGGAGACGAGAACGATTGATA	4472	TGGAGGACGGACGAAGATTGATA	4912
CAACAGTCATCANNCTTACGGTAC	4433	CAACAGTCATCANNCTTACGGTAC	4473	TGGAGGACGGACGAAGATTGATA	4913
TGAGGACGGACGGAACTGGATA	4434	CGGGAGACGAGAACGATTGATA	4474	TGGAGGACGGACGAAGATTGATA	4914
CAGCTTGAGTCAANNCATGGCTAC	4435	CAACAGTCATCANNCTTACGGTAC	4475	TGGAGGACGGACGAAGATTGATA	4915
CAACAGTCATCANNCTTACGGTAC	4436	CGGGAGACGAGAACGATTGATA	4476	TGGAGGACGGACGAAGATTGATA	4916
TGAGGACGGACGGAACTGGATA	4437	CGGGAGACGAGAACGATTGATA	4477	TGGAGGACGGACGAAGATTGATA	4917
CAGCTTGAGTCAANNCATGGCTAC	4438	CAACAGTCATCANNCTTACGGTAC	4478	TGGAGGACGGACGAAGATTGATA	4918
TGAGGACGGACGGAACTGGATA	4439	CGGGAGACGAGAACGATTGATA	4479	TGGAGGACGGACGAAGATTGATA	4919
CAGCTTGAGTCAANNCATGGCTAC	4440	CAACAGTCATCANNCTTACGGTAC	4480	TGGAGGACGGACGAAGATTGATA	4920
TGAGGACGGACGGAACTGGATA	4441	CGGGAGACGAGAACGATTGATA	4481	TGGAGGACGGACGAAGATTGATA	4921
CAACAGTCATCANNCTTACGGTAC	4442	CAACAGTCATCANNCTTACGGTAC	4482	TGGAGGACGGACGAAGATTGATA	4922
TGAGGACGGACGGAACTGGATA	4443	CGGGAGACGAGAACGATTGATA	4483	TGGAGGACGGACGAAGATTGATA	4923
CAGCTTGAGTCAANNCATGGCTAC	4444	CAACAGTCATCANNCTTACGGTAC	4484	TGGAGGACGGACGAAGATTGATA	4924
TGAGGACGGACGGAACTGGATA	4445	CGGGAGACGAGAACGATTGATA	4485	TGGAGGACGGACGAAGATTGATA	4925
CAGCTTGAGTCAANNCATGGCTAC	4446	CAACAGTCATCANNCTTACGGTAC	4486	TGGAGGACGGACGAAGATTGATA	4926
TGAGGACGGACGGAACTGGATA	4447	CGGGAGACGAGAACGATTGATA	4487	TGGAGGACGGACGAAGATTGATA	4927
CAACAGTCATCANNCTTACGGTAC	4448	CAACAGTCATCANNCTTACGGTAC	4488	TGGAGGACGGACGAAGATTGATA	4928
TGAGGACGGACGGAACTGGATA	4449	CGGGAGACGAGAACGATTGATA	4489	TGGAGGACGGACGAAGATTGATA	4929
CAGCTTGAGTCAANNCATGGCTAC	4450	CAACAGTCATCANNCTTACGGTAC	4490	TGGAGGACGGACGAAGATTGATA	4930
TGAGGACGGACGGAACTGGATA	4451	CGGGAGACGAGAACGATTGATA	4491	TGGAGGACGGACGAAGATTGATA	4931
CAGCTTGAGTCAANNCATGGCTAC	4452	CAACAGTCATCANNCTTACGGTAC	4492	TGGAGGACGGACGAAGATTGATA	4932
TGAGGACGGACGGAACTGGATA	4453	CGGGAGACGAGAACGATTGATA	4493	TGGAGGACGGACGAAGATTGATA	4933
CAGCTTGAGTCAANNCATGGCTAC	4454	CAACAGTCATCANNCTTACGGTAC	4494	TGGAGGACGGACGAAGATTGATA	4934
TGAGGACGGACGGAACTGGATA	4455	CGGGAGACGAGAACGATTGATA	4495	TGGAGGACGGACGAAGATTGATA	4935
CAGCTTGAGTCAANNCATGGCTAC	4456	CAACAGTCATCANNCTTACGGTAC	4496	TGGAGGACGGACGAAGATTGATA	4936
TGAGGACGGACGGAACTGGATA	4457	CGGGAGACGAGAACGATTGATA	4497	TGGAGGACGGACGAAGATTGATA	4937
CAACAGTCATCANNCTTACGGTAC	4458	CAACAGTCATCANNCTTACGGTAC	4498	TGGAGGACGGACGAAGATTGATA	4938
TGAGGACGGACGGAACTGGATA	4459	CGGGAGACGAGAACGATTGATA	4499	TGGAGGACGGACGAAGATTGATA	4939
CAGCTTGAGTCAANNCATGGCTAC	4460	CAACAGTCATCANNCTTACGGTAC	4500	TGGAGGACGGACGAAGATTGATA	4940

【図18E】

【図1.8E】

【 図 1 8 F 】

【図18F】

【図18G】

【図18G】

【図18H】

【図18H】

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
TGCGACGACGAACTTGGCAATTA	4960	TGCAGACGCCAGAAATTGCGAT	4989	TGCAGACGCCAGAAATTGCGAT	5042
GAGCTTTGCGGTTTCGACGAGT		CACAGAATCCNNHTGATCAGGT		CATGCGGACGCCAGAAATTGCGAT	
TGCGACGACGAACTTGGCAATTA	4961	TGCAGACGCCAGAAATTGCGAT	4980	TGCAGACGCCAGAAATTGCGAT	5041
GAGCTTTGCGGTTTCGACGAGT		CACAGAATCCNNHTGATCAGGT		CATGCGGACGCCAGAAATTGCGAT	
TGCGACGACGAACTTGGCAATTA	4962	TGCAGACGCCAGAAATTGCGAT	4982	TGCAGACGCCAGAAATTGCGAT	5042
CATACATCGGTTGCGGTTTCGACGAGT		CACAGAATCCNNHTGATCAGGT		CACCGCGGACGCCAGAAATTGCGAT	
TGCGACGACGAACTTGGCAATTA	4963	TGCAGACGCCAGAAATTGCGAT	4983	TGCAGACGCCAGAAATTGCGAT	5043
GAGCTTTGCGGTTTCGACGAGT		CACAGAATCCNNHTGATCAGGT		CACCGCGGACGCCAGAAATTGCGAT	
TGCGACGACGAACTTGGCAATTA	4964	TGCAGACGCCAGAAATTGCGAT	4984	TGCAGACGCCAGAAATTGCGAT	5044
GAGCTTTGCGGTTTCGACGAGT		CACAGAATCCNNHTGATCAGGT		CACCGCGGACGCCAGAAATTGCGAT	
TGCGACGACGAACTTGGCAATTA	4965	TGCAGACGCCAGAAATTGCGAT	4985	TGCAGACGCCAGAAATTGCGAT	5045
GAGCTTTGCGGTTTCGACGAGT		CACAGAATCCNNHTGATCAGGT		CACCGCGGACGCCAGAAATTGCGAT	
TGCGACGACGAACTTGGCAATTA	4966	TGCAGACGCCAGAAATTGCGAT	4986	TGCAGACGCCAGAAATTGCGAT	5046
GAGCTTTGCGGTTTCGACGAGT		CACAGAATCCNNHTGATCAGGT		CACCGCGGACGCCAGAAATTGCGAT	
TGCGACGACGAACTTGGCAATTA	4967	TGCAGACGCCAGAAATTGCGAT	4987	TGCAGACGCCAGAAATTGCGAT	5047
GAGCTTTGCGGTTTCGACGAGT		CACAGAATCCNNHTGATCAGGT		CACCGCGGACGCCAGAAATTGCGAT	
TGCGACGACGAACTTGGCAATTA	4968	TGCAGACGCCAGAAATTGCGAT	4988	TGCAGACGCCAGAAATTGCGAT	5048
GAGCTTTGCGGTTTCGACGAGT		CACAGAATCCNNHTGATCAGGT		CACCGCGGACGCCAGAAATTGCGAT	

【図 19A】

【図19A】

【図19B】

【図19B】

ブール'22	配列番号	ブール'23	配列番号	ブール'24	
TCCAGGACCGAGAACTGAA	5082	TGGAGAACGAGAAATTGTGAA	5322	TGGAGAACGAGAAACTGGATA	5562
CATACGCTTCAAGGAACTGAA	5083	TGGAGAACGAGAAATTGTGAA	5323	CAGCAGCGGAGAACTGGATA	5563
CAATCTGCTTCAAGGAACTGAA	5083	TGGAGAACGAGAAATTGTGAA	5323	CAGCAGCGGAGAACTGGATA	5563
TGGAGAACGAGAACTGAA	5084	TGGAGAACGAGAAATTGTGAA	5324	TGGAGAACGAGAACTGGATA	5564
CATACGCTTCAAGGAACTGAA	5085	TGGAGAACGAGAAATTGTGAA	5325	CATGGAGAACGAGAACTGGATA	5565
CAATCTGCTTCAAGGAACTGAA	5085	TGGAGAACGAGAAATTGTGAA	5325	CATGGAGAACGAGAACTGGATA	5565
TGGAGAACGAGAACTGAA	5086	TGGAGAACGAGAAATTGTGAA	5326	TGGAGAACGAGAACTGGATA	5566
CATACGCTTCAAGGAACTGAA	5087	TGGAGAACGAGAAATTGTGAA	5327	CAGCAGCGGAGAACTGGATA	5567
CAACTTGAGGAACTGAA	5087	TGGAGAACGAGAAATTGTGAA	5327	CGTCGAGAACGAGAACTGGATA	5567
TGGAGAACGAGAACTGAA	5088	TGGAGAACGAGAAATTGTGAA	5328	TGGAGAACGAGAACTGGATA	5568
CAACATGGTTCAAGGAACTGAA	5089	TGGAGAACGAGAAATTGTGAA	5329	TGGAGAACGAGAACTGGATA	5569
CAATCTGCTTCAAGGAACTGAA	5089	TGGAGAACGAGAAATTGTGAA	5329	TGGAGAACGAGAACTGGATA	5569
TGGAGAACGAGAACTGAA	5090	TGGAGAACGAGAAATTGTGAA	5330	TGGAGAACGAGAACTGGATA	5570
CATATGCTTCAAGGAACTGAA	5090	TGGAGAACGAGAAATTGTGAA	5330	CATCTGAGAACGAGAACTGGATA	5570
TGGAGAACGAGAACTGAA	5091	TGGAGAACGAGAAATTGTGAA	5331	TGGAGAACGAGAACTGGATA	5571
CATACGCTTCAAGGAACTGAA	5092	TGGAGAACGAGAAATTGTGAA	5332	TGGAGAACGAGAACTGGATA	5572
TGGAGAACGAGAACTGAA	5093	TGGAGAACGAGAAATTGTGAA	5333	TGGAGAACGAGAACTGGATA	5573
CATACGCTTCAAGGAACTGAA	5093	TGGAGAACGAGAAATTGTGAA	5333	TGGAGAACGAGAACTGGATA	5573
CACTGGTGTGAGGAACTGAA	5094	TGGAGAACGAGAAATTGTGAA	5334	TGGAGAACGAGAACTGGATA	5574
CAACATGGTTCAAGGAACTGAA	5094	TGGAGAACGAGAAATTGTGAA	5334	TGGAGAACGAGAACTGGATA	5574
TGGAGAACGAGAACTGAA	5095	TGGAGAACGAGAAATTGTGAA	5335	TGGAGAACGAGAACTGGATA	5575
CATACGCTTCAAGGAACTGAA	5095	TGGAGAACGAGAAATTGTGAA	5335	CATGTCGAGAACGAGAACTGGATA	5575
CAACATGGTTCAAGGAACTGAA	5096	TGGAGAACGAGAAATTGTGAA	5336	TGGAGAACGAGAACTGGATA	5576
TGGAGAACGAGAACTGAA	5097	TGGAGAACGAGAAATTGTGAA	5337	CATATGCTTCAAGGAACTGGATA	5577
CATACGCTTCAAGGAACTGAA	5098	TGGAGAACGAGAAATTGTGAA	5338	TGGAGAACGAGAACTGGATA	5578
CAACATGGTTCAAGGAACTGAA	5098	TGGAGAACGAGAAATTGTGAA	5338	CATCTGAGAACGAGAACTGGATA	5578
TGGAGAACGAGAACTGAA	5099	TGGAGAACGAGAAATTGTGAA	5339	TGGAGAACGAGAACTGGATA	5579
CATACGCTTCAAGGAACTGAA	5099	TGGAGAACGAGAAATTGTGAA	5339	CATCTGAGAACGAGAACTGGATA	5579
CACTGGTGTGAGGAACTGAA	5100	TGGAGAACGAGAAATTGTGAA	5340	TGGAGAACGAGAACTGGATA	5580
CAACATGGTTCAAGGAACTGAA	5100	TGGAGAACGAGAAATTGTGAA	5340	TGGAGAACGAGAACTGGATA	5580
TGGAGAACGAGAACTGAA	5101	TGGAGAACGAGAAATTGTGAA	5341	TGGAGAACGAGAACTGGATA	5581
CATACGCTTCAAGGAACTGAA	5102	TGGAGAACGAGAAATTGTGAA	5342	TGGAGAACGAGAACTGGATA	5582
CACTGGTGTGAGGAACTGAA	5102	TGGAGAACGAGAAATTGTGAA	5342	TGGAGAACGAGAACTGGATA	5582
TGGAGAACGAGAACTGAA	5103	TGGAGAACGAGAAATTGTGAA	5343	TGGAGAACGAGAACTGGATA	5583
CAACATGGTTCAAGGAACTGAA	5103	TGGAGAACGAGAAATTGTGAA	5343	TGGAGAACGAGAACTGGATA	5583
TGGAGAACGAGAACTGAA	5104	TGGAGAACGAGAAATTGTGAA	5344	TGGAGAACGAGAACTGGATA	5584
CATACGCTTCAAGGAACTGAA	5104	TGGAGAACGAGAAATTGTGAA	5344	TGGAGAACGAGAACTGGATA	5584
CACTGGTGTGAGGAACTGAA	5105	TGGAGAACGAGAAATTGTGAA	5345	TGGAGAACGAGAACTGGATA	5585
CAACATGGTTCAAGGAACTGAA	5105	TGGAGAACGAGAAATTGTGAA	5345	TGGAGAACGAGAACTGGATA	5585
TGGAGAACGAGAACTGAA	5106	TGGAGAACGAGAAATTGTGAA	5346	TGGAGAACGAGAACTGGATA	5586
CATACGCTTCAAGGAACTGAA	5107	TGGAGAACGAGAAATTGTGAA	5347	CATATGCTTCAAGGAACTGGATA	5587
CACTGGTGTGAGGAACTGAA	5107	TGGAGAACGAGAAATTGTGAA	5347	CATATGCTTCAAGGAACTGGATA	5587
TGGAGAACGAGAACTGAA	5108	TGGAGAACGAGAAATTGTGAA	5348	TGGAGAACGAGAACTGGATA	5588
CAACATGGTTCAAGGAACTGAA	5108	TGGAGAACGAGAAATTGTGAA	5348	TGGAGAACGAGAACTGGATA	5588
TGGAGAACGAGAACTGAA	5109	TGGAGAACGAGAAATTGTGAA	5349	TGGAGAACGAGAACTGGATA	5589
CACTGGTGTGAGGAACTGAA	5109	TGGAGAACGAGAAATTGTGAA	5349	TGGAGAACGAGAACTGGATA	5589
TGGAGAACGAGAACTGAA	5110	TGGAGAACGAGAAATTGTGAA	5350	TGGAGAACGAGAACTGGATA	5590
CATACGCTTCAAGGAACTGAA	5110	TGGAGAACGAGAAATTGTGAA	5350	CACGACGAGAACTGGATA	5590
CACTGGTGTGAGGAACTGAA	5111	TGGAGAACGAGAAATTGTGAA	5351	CATCTGAGAACGAGAACTGGATA	5591
TGGAGAACGAGAACTGAA	5112	TGGAGAACGAGAAATTGTGAA	5352	CGTCGAGAACGAGAACTGGATA	5592
CAACATGGTTCAAGGAACTGAA	5112	TGGAGAACGAGAAATTGTGAA	5352	TGGAGAACGAGAACTGGATA	5592
TGGAGAACGAGAACTGAA	5113	TGGAGAACGAGAAATTGTGAA	5353	TGGAGAACGAGAACTGGATA	5593
CAACATGGTTCAAGGAACTGAA	5113	TGGAGAACGAGAAATTGTGAA	5353	TGGAGAACGAGAACTGGATA	5593
TGGAGAACGAGAACTGAA	5114	TGGAGAACGAGAAATTGTGAA	5354	TGGAGAACGAGAACTGGATA	5594

【図 19c】

[図19C]

【図19D】

【図19D】

【 四 19 E 】

【图 19E】

【図19F】

[图 1.9 F]

【 19G 】

[図19G]

【図19H】

【図19H】

ブール22	配列番号:	ブール23	配列番号:	ブール24	配列番号:
TGCGAGCCGAGAATGCCATA	5261	TGCGAGCCGAGAATGCCATA	5320	TGCGAGCCGAGAATGCCATA	5760
CATCTGCAGCCTTGGATCAGG		CATCGGGCACTTGGATCAGG		CACAAAGGTCTTNNNGTCATCAGG	
TGCGAGCAGAACATTGCGATA	5281	TGCGAGCAGAACATTGCGATA	5521	TGCGAGCAGAACATTGCGATA	5761
CACCGCATCTTGGATCAGG		CACCGCATCTTGGATCAGG		TGCGAGCAGAACATTGCGATA	
TGCGAGCAGAACATTGCGATA	5282	TGCGAGCAGAACATTGCGATA	5522	TGCGAGCAGAACATTGCGATA	5762
CGACACTGGATCTGGATCAGG		CGACACTGGATCTGGATCAGG		CACACTGGATCTGGATCAGG	
TGCGAGCAGAACATTGCGATA	5283	TGCGAGCAGAACATTGCGATA	5523	TGCGAGCAGAACATTGCGATA	5763
CATCTGCAGCCTTGGATCAGG		CATCGGGCACTTGGATCAGG		CACGCTGGCACTTGGATCAGG	
TGCGAGCAGAACATTGCGATA	5284	TGCGAGCAGAACATTGCGATA	5524	TGCGAGCAGAACATTGCGATA	5764
CGACACTGGATCTGGATCAGG		CGACACTGGATCTGGATCAGG		TGCGAGCAGAACATTGCGATA	
TGCGAGCAGAACATTGCGATA	5285	TGCGAGCAGAACATTGCGATA	5525	TGCGAGCAGAACATTGCGATA	5765
CACCGCATCTTGGATCAGG		CACCGCATCTTGGATCAGG		CACGCGGTCTTNNNGTCATCAGG	
TGCGAGCAGAACATTGCGATA	5286	TGCGAGCAGAACATTGCGATA	5526	TGCGAGCAGAACATTGCGATA	5766
CACCGCATCTTGGATCAGG		CACCGCATCTTGGATCAGG		TGCGAGCAGAACATTGCGATA	
TGCGAGCAGAACATTGCGATA	5287	TGCGAGCAGAACATTGCGATA	5527	TGCGAGCAGAACATTGCGATA	5767
CATCTGCAGCCTTGGATCAGG		CATCGGGCACTTGGATCAGG		CACCTCTGGATCTGGATCAGG	
TGCGAGCAGAACATTGCGATA	5288	TGCGAGCAGAACATTGCGATA	5528	TGCGAGCAGAACATTGCGATA	5768

【図20A】

【図20A】

【図20B】

【図20B】

【图20c】

【図20C】

【図20D】

【図20D】

【図20E】

【図20E】

【 図 2 0 F 】

【図20F】

【図20H】

【図20H】

【図21A】

【図21A】

ペアル28	配列番号:	ペアル29	配列番号:	ペアル30	配列番号:
TGCGAGCAAGCAAGATGATA	6485	TGCGAGACAGAGAATGATA	6723	TGCGAGACAGAGAATGATA	6963
GACATTATGNNAACTATGCGA		CAGCTTACGNNNAACTATGCGA		CATATGTTCTNNNACTATGCGA	
TGCGAGACAGAATGATA	6490	TGCGAGACAGAATGATA	6730	TGCGAGACAGAATGATA	6970
GACATTATGNNAACTATGCGA		CAGCTTACGNNNAACTATGCGA		CACGGACAGCAGAATGATA	
TGCGAGACAGAATGATA	6491	TGCGAGACAGAATGATA	6731	TGCGAGACAGAATGATA	6971
GACATTATGNNAACTATGCGA		CAGCTTACGNNNAACTATGCGA		CAGTGTGCGNNNAACTATGCGA	
TGCGAGACAGAATGATA	6492	TGCGAGACAGAATGATA	6732	TGCGAGACAGAATGATA	6972
GACATTATGNNAACTATGCGA		CAGCTTACGNNNAACTATGCGA		CGGACAGCAGAATGATA	
TGCGAGACAGAATGATA	6493	TGCGAGACAGAATGATA	6733	TGCGAGACAGAATGATA	6973
GACATTATGNNAACTATGCGA		CAGCTTACGNNNAACTATGCGA		GAATACCCGNNNAACTATGCGA	
TGCGAGACAGAATGATA	6494	TGCGAGACAGAATGATA	6734	TGCGAGACAGAATGATA	6974
GACATTATGNNAACTATGCGA		CAGCTTACGNNNAACTATGCGA		CACGACACAGAATGATA	
TGCGAGACAGAATGATA	6495	TGCGAGACAGAATGATA	6735	TGCGAGACAGAATGATA	6975
GACATTATGNNAACTATGCGA		CAGCTTACGNNNAACTATGCGA		CGGACAGCAGAATGATA	
TGCGAGACAGAATGATA	6496	TGCGAGACAGAATGATA	6736	TGCGAGACAGAATGATA	6976
GACATTATGNNAACTATGCGA		CAGCTTACGNNNAACTATGCGA		CAGTGTGCGNNNAACTATGCGA	
TGCGAGACAGAATGATA	6497	TGCGAGACAGAATGATA	6737	TGCGAGACAGAATGATA	6977
GACATTATGNNAACTATGCGA		CAGCTTACGNNNAACTATGCGA		CGGACAGCAGAATGATA	
TGCGAGACAGAATGATA	6498	TGCGAGACAGAATGATA	6738	TGCGAGACAGAATGATA	6978
GACATTATGNNAACTATGCGA		CAGCTTACGNNNAACTATGCGA		CAAAATTCGNNNAACTATGCGA	
TGCGAGACAGAATGATA	6499	TGCGAGACAGAATGATA	6739	TGCGAGACAGAATGATA	6979
GACATTATGNNAACTATGCGA		CAGCTTACGNNNAACTATGCGA		CGGACAGCAGAATGATA	
TGCGAGACAGAATGATA	6500	TGCGAGACAGAATGATA	6740	TGCGAGACAGAATGATA	6980
GACATTATGNNAACTATGCGA		CAGCTTACGNNNAACTATGCGA		GAATCTTAATGNNNAACTATGCGA	
TGCGAGACAGAATGATA	6501	TGCGAGACAGAATGATA	6741	TGCGAGACAGAATGATA	6981
GACATTATGNNAACTATGCGA		CAGCTTACGNNNAACTATGCGA		CGAACGACAGAATGATA	
TGCGAGACAGAATGATA	6502	TGCGAGACAGAATGATA	6742	TGCGAGACAGAATGATA	6982
GACATTATGNNAACTATGCGA		CAGCTTACGNNNAACTATGCGA		CATACTTGCTNNNAACTATGCGA	
TGCGAGACAGAATGATA	6503	TGCGAGACAGAATGATA	6743	TGCGAGACAGAATGATA	6983
GACATTATGNNAACTATGCGA		CAGCTTACGNNNAACTATGCGA		CGATCATGNNNAACTATGCGA	
TGCGAGACAGAATGATA	6504	TGCGAGACAGAATGATA	6744	TGCGAGACAGAATGATA	6984
GACATTATGNNAACTATGCGA		CAGCTTACGNNNAACTATGCGA		CGATCTTGCTNNNAACTATGCGA	
TGCGAGACAGAATGATA	6505	TGCGAGACAGAATGATA	6745	TGCGAGACAGAATGATA	6985
GACATTATGNNAACTATGCGA		CAGCTTACGNNNAACTATGCGA		GAACGACAGCAGAATGATA	
TGCGAGACAGAATGATA	6506	TGCGAGACAGAATGATA	6746	TGCGAGACAGAATGATA	6986
GACATTATGNNAACTATGCGA		CAGCTTACGNNNAACTATGCGA		CGGACAGCAGAATGATA	
TGCGAGACAGAATGATA	6507	TGCGAGACAGAATGATA	6747	TGCGAGACAGAATGATA	6987
GACATTATGNNAACTATGCGA		CAGCTTACGNNNAACTATGCGA		CACCTCTACGNNNAACTATGCGA	
TGCGAGACAGAATGATA	6508	TGCGAGACAGAATGATA	6748	TGCGAGACAGAATGATA	6988
GACATTATGNNAACTATGCGA		CAGCTTACGNNNAACTATGCGA		CGACGGCGAGCAGAATGATA	
TGCGAGACAGAATGATA	6509	TGCGAGACAGAATGATA	6749	TGCGAGACAGAATGATA	6989
GACATTATGNNAACTATGCGA		CAGCTTACGNNNAACTATGCGA		CGAGCTGCTNNNAACTATGCGA	
TGCGAGACAGAATGATA	6510	TGCGAGACAGAATGATA	6750	TGCGAGACAGAATGATA	6990
GACATTATGNNAACTATGCGA		CAGCTTACGNNNAACTATGCGA		CGACCATGNNNAACTATGCGA	
TGCGAGACAGAATGATA	6511	TGCGAGACAGAATGATA	6751	TGCGAGACAGAATGATA	6991
GACATTATGNNAACTATGCGA		CAGCTTACGNNNAACTATGCGA		CGATGCTGCTNNNAACTATGCGA	
TGCGAGACAGAATGATA	6512	TGCGAGACAGAATGATA	6752	TGCGAGACAGAATGATA	6992
GACATTATGNNAACTATGCGA		CAGCTTACGNNNAACTATGCGA		CGACGGCGAGCAGAATGATA	
TGCGAGACAGAATGATA	6513	TGCGAGACAGAATGATA	6753	TGCGAGACAGAATGATA	6993
GACATTATGNNAACTATGCGA		CAGCTTACGNNNAACTATGCGA		CGGACAGCAGAATGATA	
TGCGAGACAGAATGATA	6514	TGCGAGACAGAATGATA	6754	TGCGAGACAGAATGATA	6994
GACATTATGNNAACTATGCGA		CAGCTTACGNNNAACTATGCGA		GAATGCTGCTNNNAACTATGCGA	
TGCGAGACAGAATGATA	6515	TGCGAGACAGAATGATA	6755	TGCGAGACAGAATGATA	6995
GACATTATGNNAACTATGCGA		CAGCTTACGNNNAACTATGCGA		CGACGCTGCTNNNAACTATGCGA	
TGCGAGACAGAATGATA	6516	TGCGAGACAGAATGATA	6756	TGCGAGACAGAATGATA	6996
GACATTATGNNAACTATGCGA		CAGCTTACGNNNAACTATGCGA		GAATCTGCTGCTNNNAACTATGCGA	
TGCGAGACAGAATGATA	6517	TGCGAGACAGAATGATA	6757	TGCGAGACAGAATGATA	6997
GACATTATGNNAACTATGCGA		CAGCTTACGNNNAACTATGCGA		CGACGGCGAGCAGAATGATA	
TGCGAGACAGAATGATA	6518	TGCGAGACAGAATGATA	6758	TGCGAGACAGAATGATA	6998
GACATTATGNNAACTATGCGA		CAGCTTACGNNNAACTATGCGA		CGAGCTGCTGCTNNNAACTATGCGA	
TGCGAGACAGAATGATA	6519	TGCGAGACAGAATGATA	6759	TGCGAGACAGAATGATA	6999
GACATTATGNNAACTATGCGA		CAGCTTACGNNNAACTATGCGA		GAATCTGCTGCTNNNAACTATGCGA	
TGCGAGACAGAATGATA	6520	TGCGAGACAGAATGATA	6760	TGCGAGACAGAATGATA	7000
GACATTATGNNAACTATGCGA		CAGCTTACGNNNAACTATGCGA		CGACTCTGCTGCTNNNAACTATGCGA	
TGCGAGACAGAATGATA	6521	TGCGAGACAGAATGATA	6761	TGCGAGACAGAATGATA	7001
GACATTATGNNAACTATGCGA		CAGCTTACGNNNAACTATGCGA		CAATACGCTGCTNNNAACTATGCGA	

【図21B】

【図21B】

【図21C】

【図21C】

〔 叴 2 1 H 〕

【図21H】

ゲノム番号	ゲノム番号	ゲノム番号	ゲノム番号
TGCGAGACGAGGAATCTGGATA	6720	TGCGAGACGAGGAATCTGGATA	6760
CATGATTTCATTTGGATCAAGGT		CATGAGAACAGAAATTCTGGATA	
TGCGAGACGAGGAATCTGGATA	6721	TGCGAGACGAGGAATCTGGATA	6861
CATGATTTCATTTGGATCAAGGT		CATGAGAACAGAAATTCTGGATA	
TGCGAGACGAGGAATCTGGATA	6722	TGCGAGACGAGGAATCTGGATA	6862
CATGATTTCATTTGGATCAAGGT		CATGAGAACAGAAATTCTGGATA	
TGCGAGACGAGGAATCTGGATA	6723	TGCGAGACGAGGAATCTGGATA	6863
CATGATTTCATTTGGATCAAGGT		CATGAGAACAGAAATTCTGGATA	
TGCGAGACGAGGAATCTGGATA	6724	TGCGAGACGAGGAATCTGGATA	6864
CATGATTTCATTTGGATCAAGGT		CATGAGAACAGAAATTCTGGATA	
TGCGAGACGAGGAATCTGGATA	6725	TGCGAGACGAGGAATCTGGATA	6865
CATGATTTCATTTGGATCAAGGT		CATGAGAACAGAAATTCTGGATA	
TGCGAGACGAGGAATCTGGATA	6726	TGCGAGACGAGGAATCTGGATA	6866
CATGATTTCATTTGGATCAAGGT		CATGAGAACAGAAATTCTGGATA	
TGCGAGACGAGGAATCTGGATA	6727	TGCGAGACGAGGAATCTGGATA	6867
CATGATTTCATTTGGATCAAGGT		CATGAGAACAGAAATTCTGGATA	
TGCGAGACGAGGAATCTGGATA	6728	TGCGAGACGAGGAATCTGGATA	6868
CATGATTTCATTTGGATCAAGGT		CATGAGAACAGAAATTCTGGATA	

【 2 2 A 】

【図22A】

【 図 2 2 B 】

【図22B】

ペール31	配列番号	ペール32	配列番号
TGCGAGCCAGAGAAATTGGATAC	7242	TGAGACGACGGAGAATTGGATAA	7482
AAGGCCAGCAGNNAGATGGCGCA		CAGAACCTTGCTTGNNAAGCTTGC	
TTTCAGCGACGAGATTGGATAC	7243	TGAGACGACGGAGAATTGGATAA	7483
TGCGAGCCAGAGAAATTGGATAC	7244	CAGAACCTTGCTTGNNAAGCTTGC	7484
ACCGCTTAATNNNNAGATGGCGCA		TGAGACGACGGAGAATTGGATAA	
TTTCAGCGACGAGATTGGATAC	7245	CAGAACCTTGCTTGNNAAGCTTGC	7485
ATTCAGCGACGAGATTGGATAC	7246	TGAGACGACGGAGAATTGGATAA	7486
TGAGCGGCGAGGAATTGGATAC	7247	CATCCCTTGCTTGNNAAGCTTGC	
TGAGCGGCGAGGAATTGGATAC	7248	TGAGACGACGGAGAATTGGATAA	7487
ATTCAGCGACGAGATTGGATAC	7249	CATCCCTTGCTTGNNAAGCTTGC	
ACCGCTTAATNNNNAGATGGCGCA		TGAGACGACGGAGAATTGGATAA	
TTTCAGCGACGAGATTGGATAC	7250	CAGAACCTTGCTTGNNAAGCTTGC	7489
TGAGCGGCGAGGAATTGGATAC	7251	TGAGACGACGGAGAATTGGATAA	7490
ATTCAGCGACGAGATTGGATAC	7252	CAGAACCTTGCTTGNNAAGCTTGC	7491
TGAGCGGCGAGGAATTGGATAC	7253	TGAGACGACGGAGAATTGGATAA	7492
ACCGCTTAATNNNNAGATGGCGCA		CAGAACCTTGCTTGNNAAGCTTGC	
TTTCAGCGACGAGATTGGATAC	7254	TGAGACGACGGAGAATTGGATAA	7494
ATTCAGCGACGAGATTGGATAC	7255	CATCCCTTGCTTGNNAAGCTTGC	
TGAGCGGCGAGGAATTGGATAC	7256	TGAGACGACGGAGAATTGGATAA	7495
ATTCAGCGACGAGATTGGATAC	7257	CAGAACCTTGCTTGNNAAGCTTGC	
TGAGCGGCGAGGAATTGGATAC	7258	TGAGACGACGGAGAATTGGATAA	7498
ACCGCTTAATNNNNAGATGGCGCA		CAGAACCTTGCTTGNNAAGCTTGC	
TTTCAGCGACGAGATTGGATAC	7259	TGAGACGACGGAGAATTGGATAA	7499
ATTCAGCGACGAGATTGGATAC	7260	CATCCCTTGCTTGNNAAGCTTGC	
TGAGCGGCGAGGAATTGGATAC	7261	TGAGACGACGGAGAATTGGATAA	7500
ATTCAGCGACGAGATTGGATAC	7262	CAGAACCTTGCTTGNNAAGCTTGC	
TGAGCGGCGAGGAATTGGATAC	7263	TGAGACGACGGAGAATTGGATAA	7502
ACCGCTTAATNNNNAGATGGCGCA		CAGAACCTTGCTTGNNAAGCTTGC	
TTTCAGCGACGAGATTGGATAC	7264	TGAGACGACGGAGAATTGGATAA	7504
ATTCAGCGACGAGATTGGATAC	7265	CACACGGCGAGNNAGCTTGC	
TGAGCGGCGAGGAATTGGATAC	7266	TGAGACGACGGAGAATTGGATAA	7505
ATTCAGCGACGAGATTGGATAC	7267	CATCCCTTGCTTGNNAAGCTTGC	
TGAGCGGCGAGGAATTGGATAC	7268	TGAGACGACGGAGAATTGGATAA	7506
ACCGCTTAATNNNNAGATGGCGCA		CAGAACCTTGCTTGNNAAGCTTGC	
TTTCAGCGACGAGATTGGATAC	7269	TGAGACGACGGAGAATTGGATAA	7507
ATTCAGCGACGAGATTGGATAC	7270	CATCCCTTGCTTGNNAAGCTTGC	
TGAGCGGCGAGGAATTGGATAC	7271	TGAGACGACGGAGAATTGGATAA	7511
ATTCAGCGACGAGATTGGATAC	7272	CATCCCTTGCTTGNNAAGCTTGC	
TGAGCGGCGAGGAATTGGATAC	7273	TGAGACGACGGAGAATTGGATAA	7512
ACCGCTTAATNNNNAGATGGCGCA		CAGAACCTTGCTTGNNAAGCTTGC	
TTTCAGCGACGAGATTGGATAC	7274	TGAGACGACGGAGAATTGGATAA	7513
ATTCAGCGACGAGATTGGATAC	7275	CACACGGCGAGNNAGCTTGC	
TGAGCGGCGAGGAATTGGATAC	7276	TGAGACGACGGAGAATTGGATAA	7515
ATTCAGCGACGAGATTGGATAC	7277	CAGAACCTTGCTTGNNAAGCTTGC	
TGAGCGGCGAGGAATTGGATAC	7278	TGAGACGACGGAGAATTGGATAA	7516
ACCGCTTAATNNNNAGATGGCGCA		CAGAACCTTGCTTGNNAAGCTTGC	
TTTCAGCGACGAGATTGGATAC	7279	TGAGACGACGGAGAATTGGATAA	7517

【図22C】

[図22C]

グルー31	配列番号	グルー32	配列番号
TGCGACGAGCAAGATTCGAATT	7215	TGCGACGAGCAAGATTCGAATT	7515
AAGCTTGTTCGGTTTGATGTTTC		CGACGAGCAAGATTCGAATT	
TGCGACGAGCAAGATTCGAATT	7216	TGCGACGAGCAAGATTCGAATT	7516
ACGGCGCTCCNNCTAGCGTTAC		CGACGAGCAAGATTCGAATT	
TGCGACGAGCAAGATTCGAATT	7217	TGCGACGAGCAAGATTCGAATT	7517
ACGGCGCTCCNNCTAGCGTTAC		CGACGAGCAAGATTCGAATT	
TGCGACGAGCAAGATTCGAATT	7218	TGCGACGAGCAAGATTCGAATT	7518
ACGGCGCTCCNNCTAGCGTTAC		CGACGAGCAAGATTCGAATT	
TGCGACGAGCAAGATTCGAATT	7219	TGCGACGAGCAAGATTCGAATT	7519
ATGGACCGTTCNNCTAGCGTTAC		CGACGAGCAAGATTCGAATT	
TGCGACGAGCAAGATTCGAATT	7286	TGCGACGAGCAAGATTCGAATT	7520
AAGCTTGTTCGGTTTGATGTTTC		CGACGAGCAAGATTCGAATT	
TGCGACGAGCAAGATTCGAATT	7281	TGCGACGAGCAAGATTCGAATT	7521
ACGGCTCTCCNNCTAGCGTTAC		CGACGAGCAAGATTCGAATT	
TGCGACGAGCAAGATTCGAATT	7282	TGCGACGAGCAAGATTCGAATT	7522
ACGGCTCTCCNNCTAGCGTTAC		CGACGAGCAAGATTCGAATT	
TGCGACGAGCAAGATTCGAATT	7283	TGCGACGAGCAAGATTCGAATT	7523
ATGGACCGTTCNNCTAGCGTTAC		CGACGAGCAAGATTCGAATT	
TGCGACGAGCAAGATTCGAATT	7284	TGCGACGAGCAAGATTCGAATT	7524
ATTCRACAATCCTNNCTAGCGTTAC		CGACGAGCAAGATTCGAATT	
TGCGACGAGCAAGATTCGAATT	7285	TGCGACGAGCAAGATTCGAATT	7525
ACGGCTCTCCNNCTAGCGTTAC		CGACGAGCAAGATTCGAATT	
TGCGACGAGCAAGATTCGAATT	7286	TGCGACGAGCAAGATTCGAATT	7526
ACGGCTCTCCNNCTAGCGTTAC		CGACGAGCAAGATTCGAATT	
TGCGACGAGCAAGATTCGAATT	7287	TGCGACGAGCAAGATTCGAATT	7527
AAAGCTTGTTCGGTTTGATGTTTC		CGACGAGCAAGATTCGAATT	
TGCGACGAGCAAGATTCGAATT	7288	TGCGACGAGCAAGATTCGAATT	7528
AAATGTTTCTTCTTCTTCTTCTTCT		CGACGAGCAAGATTCGAATT	
TGCGACGAGCAAGATTCGAATT	7289	TGCGACGAGCAAGATTCGAATT	7529
ATCTCCGGCTTCTTCTTCTTCTTCT		CGACGAGCAAGATTCGAATT	
TGCGACGAGCAAGATTCGAATT	7290	TGCGACGAGCAAGATTCGAATT	7530
ACGGCTCTCCNNCTAGCGTTAC		CGACGAGCAAGATTCGAATT	
TGCGACGAGCAAGATTCGAATT	7291	TGCGACGAGCAAGATTCGAATT	7531
AAAGCTTGTTCGGTTTGATGTTTC		CGACGAGCAAGATTCGAATT	
TGCGACGAGCAAGATTCGAATT	7292	TGCGACGAGCAAGATTCGAATT	7532
ACGGCTCTCCNNCTAGCGTTAC		CGACGAGCAAGATTCGAATT	
TGCGACGAGCAAGATTCGAATT	7293	TGCGACGAGCAAGATTCGAATT	7533
AAATAATGCTTCTTCTTCTTCTTCT		CGACGAGCAAGATTCGAATT	
TGCGACGAGCAAGATTCGAATT	7294	TGCGACGAGCAAGATTCGAATT	7534
ATCCCGCTTCTTCTTCTTCTTCT		CGACGAGCAAGATTCGAATT	
TGCGACGAGCAAGATTCGAATT	7295	TGCGACGAGCAAGATTCGAATT	7535
ACGGCTCTCCNNCTAGCGTTAC		CGACGAGCAAGATTCGAATT	
TGCGACGAGCAAGATTCGAATT	7296	TGCGACGAGCAAGATTCGAATT	7536
ACGGCTCTCCNNCTAGCGTTAC		CGACGAGCAAGATTCGAATT	
TGCGACGAGCAAGATTCGAATT	7297	TGCGACGAGCAAGATTCGAATT	7537
AAAGCTTGTTCGGTTTGATGTTTC		CGACGAGCAAGATTCGAATT	
TGCGACGAGCAAGATTCGAATT	7298	TGCGACGAGCAAGATTCGAATT	7538
ACGGCTCTCCNNCTAGCGTTAC		CGACGAGCAAGATTCGAATT	
TGCGACGAGCAAGATTCGAATT	7299	TGCGACGAGCAAGATTCGAATT	7539
ACGGCTCTCCNNCTAGCGTTAC		CGACGAGCAAGATTCGAATT	
TGCGACGAGCAAGATTCGAATT	7300	TGCGACGAGCAAGATTCGAATT	7540
ACGGCTCTCCNNCTAGCGTTAC		CGACGAGCAAGATTCGAATT	
TGCGACGAGCAAGATTCGAATT	7301	TGCGACGAGCAAGATTCGAATT	7541
ACGGCTCTCCNNCTAGCGTTAC		CGACGAGCAAGATTCGAATT	
TGCGACGAGCAAGATTCGAATT	7302	TGCGACGAGCAAGATTCGAATT	7542
ACGGCTCTCCNNCTAGCGTTAC		CGACGAGCAAGATTCGAATT	
TGCGACGAGCAAGATTCGAATT	7303	TGCGACGAGCAAGATTCGAATT	7543
ACGGCTCTCCNNCTAGCGTTAC		CGACGAGCAAGATTCGAATT	
TGCGACGAGCAAGATTCGAATT	7304	TGCGACGAGCAAGATTCGAATT	7544
ACGGCTCTCCNNCTAGCGTTAC		CGACGAGCAAGATTCGAATT	
TGCGACGAGCAAGATTCGAATT	7305	TGCGACGAGCAAGATTCGAATT	7545
ACGGCTCTCCNNCTAGCGTTAC		CGACGAGCAAGATTCGAATT	
TGCGACGAGCAAGATTCGAATT	7306	TGCGACGAGCAAGATTCGAATT	7546
ACGGCTCTCCNNCTAGCGTTAC		CGACGAGCAAGATTCGAATT	
TGCGACGAGCAAGATTCGAATT	7307	TGCGACGAGCAAGATTCGAATT	7547
ACGGCTCTCCNNCTAGCGTTAC		CGACGAGCAAGATTCGAATT	

【図22D】

【図22D】

ゲル番号	配列番号	ゲル番号	配列番号
TGAGCAGGACGAGAATGGTACAT	7306	TGAGCAGGACGAGAATGGTACAT	7548
AAATCCGTGCGGAGCTTGTCAT		CATGTCGCGGAGCTTGTCAT	
TGAGCAGGACGAGAATGGTACAT	7309	TGAGCAGGACGAGAATGGTACAT	7549
ATTACTCCGCGGAGCTTGTCAT		CATGTCGCGGAGCTTGTCAT	
TGAGCAGGACGAGAATGGTACAT	7310	TGAGCAGGACGAGAATGGTACAT	7550
AGAGGCGGAGCTTGTCAT		CAGAGCAGGAGCTTGTCAT	
TGAGCAGGACGAGAATGGTACAT	7311	TGAGCAGGACGAGAATGGTACAT	7551
AGAGGCGGAGCTTGTCAT		CAGAGCAGGAGCTTGTCAT	
TGAGCAGGACGAGAATGGTACAT	7312	TGAGCAGGACGAGAATGGTACAT	7552
AAATACTGCCTTNNNAGCTGGTC		CATGAGCTGTTGNNNGAGCTGGTC	
TGAGCAGGACGAGAATGGTACAT	7313	TGAGCAGGACGAGAATGGTACAT	7553
AGATGAGCAGNNAGCTGGTC		CAGAGCAGGAGCTTGTCAT	
TGAGCAGGACGAGAATGGTACAT	7314	TGAGCAGGACGAGAATGGTACAT	7554
AGAGGCGGAGCTTGTCAT		CAGAGCAGGAGCTTGTCAT	
TGAGCAGGACGAGAATGGTACAT	7315	TGAGCAGGACGAGAATGGTACAT	7555
AAATCCCTTNNNAGCTGGTC		CAGAGCAGGAGCTTGTCAT	
TGAGCAGGACGAGAATGGTACAT	7316	TGAGCAGGACGAGAATGGTACAT	7556
ATTAATGAGCAGNNAGCTGGTC		CAGAGCAGGAGCTTGTCAT	
TGAGCAGGACGAGAATGGTACAT	7317	TGAGCAGGACGAGAATGGTACAT	7557
AGAGGCGGAGCTTGTCAT		CAGAGCAGGAGCTTGTCAT	
TGAGCAGGACGAGAATGGTACAT	7318	TGAGCAGGACGAGAATGGTACAT	7558
AGACATGCGGAGCTTGTCAT		CAGAGCAGGAGCTTGTCAT	
TGAGCAGGACGAGAATGGTACAT	7319	TGAGCAGGACGAGAATGGTACAT	7559
RAAGAGGCGGAGCTTGTCAT		CAGAGCAGGAGCTTGTCAT	
TGAGCAGGACGAGAATGGTACAT	7320	TGAGCAGGACGAGAATGGTACAT	7560
ATTAATGAGCAGNNAGCTGGTC		CAGAGCAGGAGCTTGTCAT	
TGAGCAGGACGAGAATGGTACAT	7321	TGAGCAGGACGAGAATGGTACAT	7561
ATTGGCGGAGCTTGTCAT		CAGAGCAGGAGCTTGTCAT	
TGAGCAGGACGAGAATGGTACAT	7322	TGAGCAGGACGAGAATGGTACAT	7562
ACAGATGCGGAGCTTGTCAT		CATGATTCAGNNNGAGCTGGTC	
TGAGCAGGACGAGAATGGTACAT	7323	TGAGCAGGACGAGAATGGTACAT	7563
AGAGGCGGAGCTTGTCAT		CAGAGCAGGAGCTTGTCAT	
TGAGCAGGACGAGAATGGTACAT	7324	TGAGCAGGACGAGAATGGTACAT	7564
AAATCCCTTNNNAGCTGGTC		CAGCTGTTGNNNGAGCTGGTC	
TGAGCAGGACGAGAATGGTACAT	7325	TGAGCAGGACGAGAATGGTACAT	7565
ACCTGAGGCGGAGCTTGTCAT		CAGAGCAGGAGCTTGTCAT	
TGAGCAGGACGAGAATGGTACAT	7326	TGAGCAGGACGAGAATGGTACAT	7566
AGAGGCGGAGCTTGTCAT		CAGAGCAGGAGCTTGTCAT	
TGAGCAGGACGAGAATGGTACAT	7327	TGAGCAGGACGAGAATGGTACAT	7567
AAATTCTGCTTNNNAGCTGGTC		CATGAGCTGTTGNNNGAGCTGGTC	
TGAGCAGGACGAGAATGGTACAT	7328	TGAGCAGGACGAGAATGGTACAT	7568
HATAAACTGCGGAGCTTGTCAT		CAGAGCAGGAGCTTGTCAT	
TGAGCAGGACGAGAATGGTACAT	7329	TGAGCAGGACGAGAATGGTACAT	7569
AGAGGCGGAGCTTGTCAT		CATGAGCTGTTGNNNGAGCTGGTC	
TGAGCAGGACGAGAATGGTACAT	7330	TGAGCAGGACGAGAATGGTACAT	7570
ATTCAGGCGGAGCTTGTCAT		CAGAGCAGGAGCTTGTCAT	
TGAGCAGGACGAGAATGGTACAT	7331	TGAGCAGGACGAGAATGGTACAT	7571
AGAGGCGGAGCTTGTCAT		CATGAGCTGTTGNNNGAGCTGGTC	
TGAGCAGGACGAGAATGGTACAT	7332	TGAGCAGGACGAGAATGGTACAT	7572
ACACATGCGGAGCTTGTCAT		CAGAGCAGGAGCTTGTCAT	
TGAGCAGGACGAGAATGGTACAT	7333	TGAGCAGGACGAGAATGGTACAT	7573
ACCTGAGGCGGAGCTTGTCAT		CAGAGCAGGAGCTTGTCAT	
TGAGCAGGACGAGAATGGTACAT	7334	TGAGCAGGACGAGAATGGTACAT	7574
ACCTGAGGCGGAGCTTGTCAT		CAGCGGTGAGNNNGAGCTGGTC	
TGAGCAGGACGAGAATGGTACAT	7335	TGAGCAGGACGAGAATGGTACAT	7575
AGAGGCGGAGCTTGTCAT		CAGAGCAGGAGCTTGTCAT	
TGAGCAGGACGAGAATGGTACAT	7336	TGAGCAGGACGAGAATGGTACAT	7576
AGATGAGCAGNNNGAGCTGGTC		CAGAGCAGGAGCTTGTCAT	
TGAGCAGGACGAGAATGGTACAT	7337	TGAGCAGGACGAGAATGGTACAT	7577
ACACCCGCGGAGCTTGTCAT		CAGAGCAGGAGCTTGTCAT	
TGAGCAGGACGAGAATGGTACAT	7338	TGAGCAGGACGAGAATGGTACAT	7578
AGAGGCGGAGCTTGTCAT		CAGAGCAGGAGCTTGTCAT	
TGAGCAGGACGAGAATGGTACAT	7339	TGAGCAGGACGAGAATGGTACAT	7579
ATTCGCGGAGNNNGAGCTGGTC		CAGAGCAGGAGCTTGTCAT	
TGAGCAGGACGAGAATGGTACAT	7340	TGAGCAGGACGAGAATGGTACAT	7580
AGAGGCGGAGCTTGTCAT		CATGAGCTGTTGNNNGAGCTGGTC	

【 図 2 2 E 】

[図2.2.E]

ブルー31	配列番号	ブルー32	配列番号
TCCAGGACCAAGCAGAAATTCGAAATAC	7341	TGGAGAACAGAAGAATTGCGAATA	7381
ATTCGGCAGGNNNGTGCGACATG		CAGCGTTCAGGNNNGTGCGACATG	
TCCAGGACCAAGCAGAAATTGGATACATC	7342	TGGAGAACAGAAGAATTGCGAATA	7382
ATTCGGCAGGNNNGTGCGACATG		CAGCGTTCAGGNNNGTGCGACATG	
TCCAGGACCAAGCAGAAATTGGATACATC	7343	TGGAGAACAGAAGAATTGCGAATA	7383
ACCCCTATGTTTGGACGACATG		CAGCGTTCAGGNNNGTGCGACATG	
TCCAGGACCAAGCAGAAATTGGATACATC	7344	TGGAGAACAGAAGAATTGCGAATA	7384
AGATGTAGTANNNGTGCGACATG		CAGCGTTCAGGNNNGTGCGACATG	
TCCAGGACCAAGCAGAAATTGGATACATC	7345	TGGAGAACAGAAGAATTGCGAATA	7385
XNNTTGGACGACATG		CAGCGTTCAGGNNNGTGCGACATG	
TCCAGGACCAAGCAGAAATTGGATACATC	7346	TGGAGAACAGAAGAATTGCGAATA	7386
ATGGTCCGGNNNGTGCGACATG		CAGCGTTCAGGNNNGTGCGACATG	
TCCAGGACCAAGCAGAAATTGGATACATC	7347	TGGAGAACAGAAGAATTGCGAATA	7387'
AGTCGGCAGGNNNGTGCGACATG		CAGCGTTCAGGNNNGTGCGACATG	
TCCAGGACCAAGCAGAAATTGGATACATC	7348	TGGAGAACAGAAGAATTGCGAATA	7388
AGTCGGCAGGNNNGTGCGACATG		CAGCGTTCAGGNNNGTGCGACATG	
TGGAGAACAGAAGAATTGGATACATC	7349	TGGAGAACAGAAGAATTGCGAATA	7389
AGTCGATTAATTGCGACATG		CAGCGTTCAGGNNNGTGCGACATG	
TCCAGGACCAAGCAGAAATTGGATACATC	7350	TGGAGAACAGAAGAATTGCGAATA	7390
AGTCGATTAATTGCGACATG		CAGCGTTCAGGNNNGTGCGACATG	
TCCAGGACCAAGCAGAAATTGGATACATC	7351	TGGAGAACAGAAGAATTGCGAATA	7391
AGTCGATTAATTGCGACATG		CAGCGTTCAGGNNNGTGCGACATG	
TCCAGGACCAAGCAGAAATTGGATACATC	7352	TGGAGAACAGAAGAATTGCGAATA	7392
AAAATCAGGNNNGTGCGACATG		CAGCGTTCAGGNNNGTGCGACATG	
TCCAGGACCAAGCAGAAATTGGATACATC	7353	TGGAGAACAGAAGAATTGCGAATA	7393
AAAATCAGGNNNGTGCGACATG		CAGCGTTCAGGNNNGTGCGACATG	
TCCAGGACCAAGCAGAAATTGGATACATC	7354	TGGAGAACAGAAGAATTGCGAATA	7394
ATGGCTTCTTNNNGTGCGACATG		CAGCGTTCAGGNNNGTGCGACATG	
TCCAGGACCAAGCAGAAATTGGATACATC	7355	TGGAGAACAGAAGAATTGCGAATA	7395
ATTACCGAGNNNGTGCGACATG		CAGCGTTCAGGNNNGTGCGACATG	
TCCAGGACCAAGCAGAAATTGGATACATC	7356	TGGAGAACAGAAGAATTGCGAATA	7396
ACCCCTATGTTTGGACGACATG		CAGCGTTCAGGNNNGTGCGACATG	
TCCAGGACCAAGCAGAAATTGGATACATC	7357	TGGAGAACAGAAGAATTGCGAATA	7397
ACCCCTATGTTTGGACGACATG		CAGCGTTCAGGNNNGTGCGACATG	
TCCAGGACCAAGCAGAAATTGGATACATC	7358	TGGAGAACAGAAGAATTGCGAATA	7398
AGTCGACCCNNNGTGCGACATG		CAGCGTTCAGGNNNGTGCGACATG	
TCCAGGACCAAGCAGAAATTGGATACATC	7359	TGGAGAACAGAAGAATTGCGAATA	7399
AGTCGACCCNNNGTGCGACATG		CAGCGTTCAGGNNNGTGCGACATG	
TCCAGGACCAAGCAGAAATTGGATACATC	7360	TGGAGAACAGAAGAATTGCGAATA	7400
AGTCGACCCNNNGTGCGACATG		CAGCGTTCAGGNNNGTGCGACATG	
TCCAGGACCAAGCAGAAATTGGATACATC	7361	TGGAGAACAGAAGAATTGCGAATA	7401
AGTCGACCCNNNGTGCGACATG		CAGCGTTCAGGNNNGTGCGACATG	
TCCAGGACCAAGCAGAAATTGGATACATC	7362	TGGAGAACAGAAGAATTGCGAATA	7402
AGTCGACCCNNNGTGCGACATG		CAGCGTTCAGGNNNGTGCGACATG	
TCCAGGACCAAGCAGAAATTGGATACATC	7363	TGGAGAACAGAAGAATTGCGAATA	7403
AGTCGACCCNNNGTGCGACATG		CAGCGTTCAGGNNNGTGCGACATG	
TCCAGGACCAAGCAGAAATTGGATACATC	7364	TGGAGAACAGAAGAATTGCGAATA	7404
AGTCGGTGNNGTGCGACATG		CAGCGTTCAGGNNNGTGCGACATG	
TCCAGGACCAAGCAGAAATTGGATACATC	7365	TGGAGAACAGAAGAATTGCGAATA	7405
ACTCTCTTCTTNNNGTGCGACATG		CAGCGTTCAGGNNNGTGCGACATG	
TCCAGGACCAAGCAGAAATTGGATACATC	7366	TGGAGAACAGAAGAATTGCGAATA	7406
AGGCAGCAGNNNGTGCGACATG		CAGCGTTCAGGNNNGTGCGACATG	
TCCAGGACCAAGCAGAAATTGGATACATC	7367	TGGAGAACAGAAGAATTGCGAATA	7407
ATTCGATTAATTGCGACATG		CAGCGTTCAGGNNNGTGCGACATG	
TCCAGGACCAAGCAGAAATTGGATACATC	7368	TGGAGAACAGAAGAATTGCGAATA	7408
AGTCGGCAGGNNNGTGCGACATG		CAGCGTTCAGGNNNGTGCGACATG	
TCCAGGACCAAGCAGAAATTGGATACATC	7369	TGGAGAACAGAAGAATTGCGAATA	7409
ACCCATCAGGNNNGTGCGACATG		CAGCGTTCAGGNNNGTGCGACATG	
TCCAGGACCAAGCAGAAATTGGATACATC	7370	TGGAGAACAGAAGAATTGCGAATA	7410
ACCCATCAGGNNNGTGCGACATG		CAGCGTTCAGGNNNGTGCGACATG	
TCCAGGACCAAGCAGAAATTGGATACATC	7371	TGGAGAACAGAAGAATTGCGAATA	7411
ACCCATCAGGNNNGTGCGACATG		CAGCGTTCAGGNNNGTGCGACATG	
TCCAGGACCAAGCAGAAATTGGATACATC	7372	TGGAGAACAGAAGAATTGCGAATA	7412
ATTCGATTAATTGCGACATG		CAGCGTTCAGGNNNGTGCGACATG	
TGGTCCGGNNNGTGCGACATG		CAGCGTTCAGGNNNGTGCGACATG	
TGGTCCGGNNNGTGCGACATG		CAGCGTTCAGGNNNGTGCGACATG	

【図22F】

【図22F】

ブルー31	配列番号	ブルー32	配列番号
TTCAGGCGGAGGAGATGGATAAC	7374	TGCAGGCGGAGGAGATGGATAAC	7614
ATAAACCTCCTTNNGGATGCGATC		CATCCGCTATTINNGATGCGATC	
TTCAGGCGGAGGAGATGGATAAC	7375	TGCAGGCGGAGGAGATGGATAAC	7615
ATTCCTGCTTNNGGATGCGATC		CATGGCCCTTNNGGATGCGATC	
TTCAGGCGGAGGAGATGGATAAC	7376	TGCAGGCGGAGGAGATGGATAAC	7616
AGTGACCACTTNNGGATGCGATC		CATGJATGTTTINNGATGCGATC	
TTCAGGCGGAGGAGATGGATAAC	7377	TGCAGGCGGAGGAGATGGATAAC	7617
ACCCCTTCTTNNGGATGCGATC		CATGJATGTTTINNGATGCGATC	
TTCAGGCGGAGGAGATGGATAAC	7378	TGCAGGCGGAGGAGATGGATAAC	7618
ACCCATCTTNNGGATGCGATC		CAGGGCTTCTTNNGGATGCGATC	
TTCAGGCGGAGGAGATGGATAAC	7379	TGCAGGCGGAGGAGATGGATAAC	7619
ACCCATCTTNNGGATGCGATC		CAGGGCTTCTTNNGGATGCGATC	
TTCAGGCGGAGGAGATGGATAAC	7380	TGCAGGCGGAGGAGATGGATAAC	7620
AAATAACCTTNNGGATGCGATC		CAACTCTCCCTNNGGATGCGATC	
TTCAGGCGGAGGAGATGGATAAC	7381	TGCAGGCGGAGGAGATGGATAAC	7621
AAATACTGCTTNNGGATGCGATC		CAATAGAATCTTNNGGATGCGATC	
TTCAGGCGGAGGAGATGGATAAC	7382	TGCAGGCGGAGGAGATGGATAAC	7622
ACCCATCTTNNGGATGCGATC		CAATAGAATCTTNNGGATGCGATC	
TTCAGGCGGAGGAGATGGATAAC	7383	TGCAGGCGGAGGAGATGGATAAC	7623
AGTTGCTCTTNNGGATGCGATC		CAATGCTCAGAABNNNGATGCGATC	
TTCAGGCGGAGGAGATGGATAAC	7384	TGCAGGCGGAGGAGATGGATAAC	7624
ATTCCTGCTTNNGGATGCGATC		CAANACGTCATCCTNNGGATGCGATC	
TTCAGGCGGAGGAGATGGATAAC	7385	TGCAGGCGGAGGAGATGGATAAC	7625
ACCCATCTTNNGGATGCGATC		CAANACGTCATCCTNNGGATGCGATC	
TTCAGGCGGAGGAGATGGATAAC	7386	TGCAGGCGGAGGAGATGGATAAC	7626
AAATGGCTCTTNNGGATGCGATC		CAAACTGCAACCTNNGGATGCGATC	
TTCAGGCGGAGGAGATGGATAAC	7387	TGCAGGCGGAGGAGATGGATAAC	7627
AGTGACCACTTNNGGATGCGATC		CAATGACCTTNNGGATGCGATC	
TTCAGGCGGAGGAGATGGATAAC	7388	TGCAGGCGGAGGAGATGGATAAC	7628
ACCCATCTTNNGGATGCGATC		CAATGCTTATTINNGATGCGATC	
TTCAGGCGGAGGAGATGGATAAC	7389	TGCAGGCGGAGGAGATGGATAAC	7629
ATACCAACGNNCTTCATGCGATC		CAAACTGCAACCTNNGGATGCGATC	
TTCAGGCGGAGGAGATGGATAAC	7390	TGCAGGCGGAGGAGATGGATAAC	7630
ACCCATCTTNNGGATGCGATC		CAATGACCTTNNGGATGCGATC	
TTCAGGCGGAGGAGATGGATAAC	7391	TGCAGGCGGAGGAGATGGATAAC	7631
AGTGACCACTTNNGGATGCGATC		CACTCGTGAAGGNNNGATGCGATC	
TTCAGGCGGAGGAGATGGATAAC	7392	TGCAGGCGGAGGAGATGGATAAC	7632
AAAACCTCCTTNNGGATGCGATC		CAGCGGAGGNNNGATGCGATC	
TTCAGGCGGAGGAGATGGATAAC	7393	TGCAGGCGGAGGAGATGGATAAC	7633
ACCCATCTTNNGGATGCGATC		CAATGCTTATTINNGATGCGATC	
TTCAGGCGGAGGAGATGGATAAC	7394	TGCAGGCGGAGGAGATGGATAAC	7634
AGAGCGACGNNTTCATGCGATC		CAACACGTCATCCTNNGGATGCGATC	
TTCAGGCGGAGGAGATGGATAAC	7395	TGCAGGCGGAGGAGATGGATAAC	7635
ATTCCTGCTTNNGGATGCGATC		CAATACGCTTATTINNGATGCGATC	
TTCAGGCGGAGGAGATGGATAAC	7396	TGCAGGCGGAGGAGATGGATAAC	7636
ACCCATCTTNNGGATGCGATC		CAATACGCTTATTINNGATGCGATC	
TTCAGGCGGAGGAGATGGATAAC	7397	TGCAGGCGGAGGAGATGGATAAC	7637
ACCCAGGCTTNNGGATGCGATC		CAGGAGTCTCAGTNNNGATGCGATC	
TTCAGGCGGAGGAGATGGATAAC	7398	TGCAGGCGGAGGAGATGGATAAC	7638
ATTCCTGCTTNNGGATGCGATC		CAGGAGTCTCAGTNNNGATGCGATC	
TTCAGGCGGAGGAGATGGATAAC	7399	TGCAGGCGGAGGAGATGGATAAC	7639
ACCCATCTTNNGGATGCGATC		CACTCGTGAAGGNNNGATGCGATC	
TTCAGGCGGAGGAGATGGATAAC	7400	TGCAGGCGGAGGAGATGGATAAC	7640
AGTGACCACTTNNGGATGCGATC		CAACACGTCATCCTNNGGATGCGATC	
TTCAGGCGGAGGAGATGGATAAC	7401	TGCAGGCGGAGGAGATGGATAAC	7641
ACCCATCTTNNGGATGCGATC		CAATGACCTTNNGGATGCGATC	
TTCAGGCGGAGGAGATGGATAAC	7402	TGCAGGCGGAGGAGATGGATAAC	7642
ATTCCTGCTTNNGGATGCGATC		CACTCCCTATTINNGATGCGATC	
TTCAGGCGGAGGAGATGGATAAC	7403	TGCAGGCGGAGGAGATGGATAAC	7643
ACCCATCTTNNGGATGCGATC		CAGGAGTCTCAGTNNNGATGCGATC	
TTCAGGCGGAGGAGATGGATAAC	7404	TGCAGGCGGAGGAGATGGATAAC	7644
AGTGACCACTTNNGGATGCGATC		CACTCGTGAAGGNNNGATGCGATC	
TTCAGGCGGAGGAGATGGATAAC	7405	TGCAGGCGGAGGAGATGGATAAC	7645
ACCCATCTTNNGGATGCGATC		CAACACGTCATCCTNNGGATGCGATC	
TTCAGGCGGAGGAGATGGATAAC	7406	TGCAGGCGGAGGAGATGGATAAC	7646

【図22G】

【図22G】

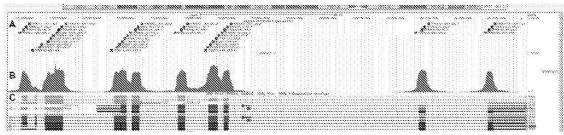
【図 2 2 H】

【図 2 2 H】

プール33	配列番号:	プール32	配列番号:
TGCAGGACCAAGAATTGGAATAC	7440	TGCAAGRCAGAGAAATTGAAATA	7680
AAGCTCGGTTCNNNTGCACTCAGGT		CATTTCGCCNNNTGCRTCAGGT	
TCCAGGACCAAGAATTGGAATAC	7441	TGCGGGACCAAGAGAAATTGGAATA	7681
AATATCCTTCAANNNTGCACTCAGGT		CAGACCATCAGNNNTGCACTCAGGT	
TGCAGGACCAAGAATTGGAATAC	7442	TGCAAGACAGAGAAATTGGAATA	7682
AACCTTCTCANNNTGCACTCAGGT		CACCGCCTAGNNNTGCRTCAGGT	
TGCAGGACCAAGAATTGGAATAC	7443	TGCAAGGNCAAGAGAAATTGGAATA	7683
AATCGGTTCNNNTGCACTCAGGT		CATTAGCTCAGNNNTGCRTCAGGT	
TGCAGGACCAAGAATTGGAATAC	7444	TGCAAGGCACAGAGAAATTGGAATA	7684
AATCGCCCGNNNTGCACTCAGGT		CATCCTGATTCNNNTGCRTCAGGT	
TGCAGGACCAAGAATTGGAATAC	7445	TGCAAGGRCAGAGAAATTGGAATA	7685
ACAGGCTGNNNTGCACTCAGGT		CATCATAGGNNNTGCRTCAGGT	
TGCAGGACCAAGAATTGGAATAC	7446	TGCAAGGACCAAGAGAAATTGGAATA	7686
AATCGGTTCNNNTGCACTCAGGT		CATGGAGCTAGNNNTGCRTCAGGT	
TGCAGGACCAAGAATTGGAATAC	7447	TGCAAGGCCAGAGAAATTGGAATA	7687
ACGGATCTGNNNTGCACTCAGGT		CAGCTCTACTAGNNNTGCRTCAGGT	
TGCAGGACCAAGAATTGGAATAC	7448	TGCAAGGACCAAGAGAAATTGGAATA	7688
ACGAGCTGTNNNTGCACTCAGGT		CAGGCTGGTGGNNNTGCRTCAGGT	

【図 2 3】

【図 2 3】

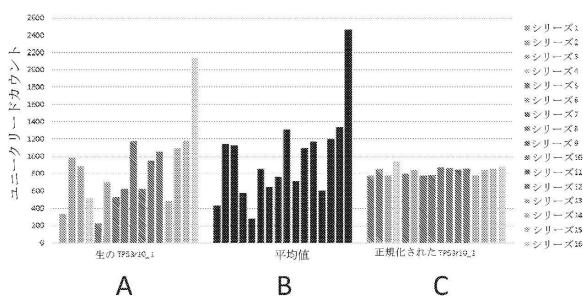


10

【図 2 4】

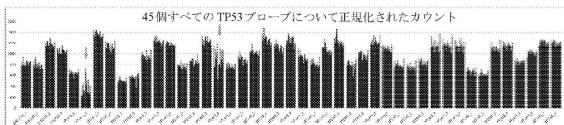
【図 2 4】

試料平均値全体に対するプローブ TP53r10_1 のユニーカリードカウントの正規化



【図 2 5】

【図 2 5】



20

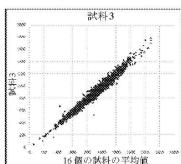
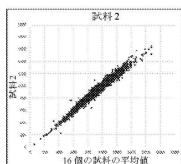
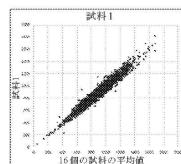
30

40

50

【図 2 6】

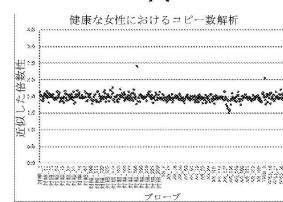
【図 2 6】



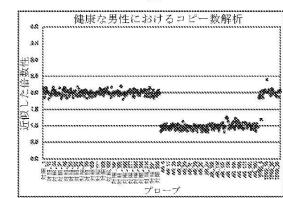
【図 2 7】

【図 2 7】

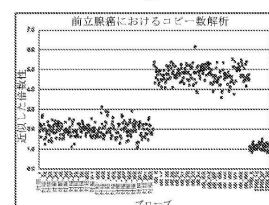
A



B

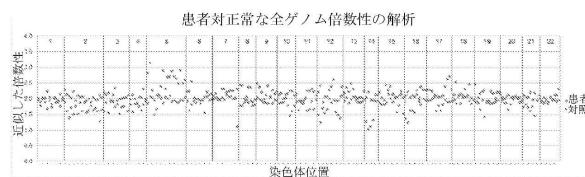


C



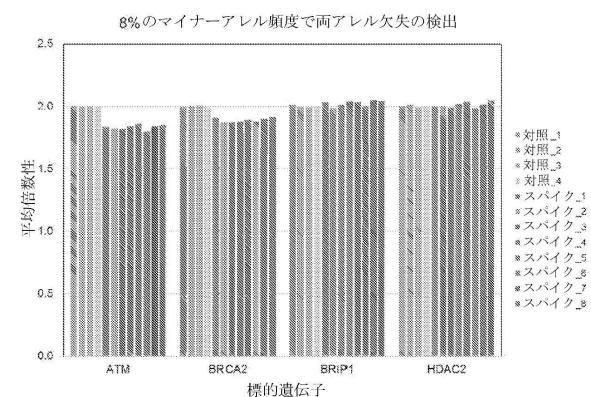
【図 2 8】

【図 2 8】



【図 2 9】

【図 2 9】



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)