

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号
特許第6073382号
(P6073382)

(45) 発行日 平成29年2月1日 (2017.2.1)

(24) 登録日 平成29年1月13日 (2017.1.13)

(51) Int. Cl.

F I

C 1 2 Q 1/68 (2006.01)

C 1 2 Q 1/68 A

C 1 2 N 15/00 (2006.01)

C 1 2 N 15/00 Z

請求項の数 15 (全 36 頁)

| | | | |
|---------------|-------------------------------|-----------|--|
| (21) 出願番号 | 特願2014-560451 (P2014-560451) | (73) 特許権者 | 513154234 |
| (86) (22) 出願日 | 平成25年3月8日 (2013.3.8) | | ザ チャイニーズ ユニバーシティー オ ブ ホンコン |
| (65) 公表番号 | 特表2015-510757 (P2015-510757A) | | 中華人民共和国 香港, ニューテリトリ ーズ, シャンティン, ザ ノーレッジ トランスファー オフィス |
| (43) 公表日 | 平成27年4月13日 (2015.4.13) | (74) 代理人 | 100099759 |
| (86) 国際出願番号 | PCT/IB2013/000312 | | 弁理士 青木 篤 |
| (87) 国際公開番号 | W02013/132305 | (74) 代理人 | 100077517 |
| (87) 国際公開日 | 平成25年9月12日 (2013.9.12) | | 弁理士 石田 敬 |
| 審査請求日 | 平成27年3月17日 (2015.3.17) | (74) 代理人 | 100087871 |
| (31) 優先権主張番号 | 61/608, 623 | | 弁理士 福本 積 |
| (32) 優先日 | 平成24年3月8日 (2012.3.8) | (74) 代理人 | 100087413 |
| (33) 優先権主張国 | 米国 (US) | | 弁理士 古賀 哲次 |
| (31) 優先権主張番号 | 61/621, 451 | | |
| (32) 優先日 | 平成24年4月6日 (2012.4.6) | | |
| (33) 優先権主張国 | 米国 (US) | | |

最終頁に続く

(54) 【発明の名称】 癌のレベルの区分に対するDNA断片のサイズに基づく解析

(57) 【特許請求の範囲】

【請求項 1】

生物の生体試料を解析する方法であって、当該生体試料は、正常細胞由来のDNA断片と癌に関連しているおそれがある細胞由来のDNA断片を含み、ここで少なくともDNA断片の一部は、生体試料における無細胞DNAであり、当該方法は：

生体試料由来の複数のDNA断片の各々について；
該DNA断片のシーケンシングにより得られた1つまたは複数の読み取り配列を受け取り、ここで該1つまたは複数の読み取り配列は該DNA断片の両端を含む；
該1つまたは複数の読み取り配列を参照ゲノムにアラインさせ、該DNA断片の両端についてアラインする位置を得る；および
アラインした位置を用いて該DNA断片のサイズを決定する；
複数のサイズのうちの各サイズについて；
サイズに対応する生体試料由来の複数のDNA断片の第1セットの量を、第一セットについてアラインされた位置から決定されたサイズを用いて決定し；
第1パラメータの第1値を複数のサイズのDNA断片の量に基づいて計算し、ここで第1パラメータは、生体試料におけるDNA断片のサイズプロファイルの統計的尺度を提供し；
第1値を基準値と比較し；そして、
その比較に基づいて、生物の癌のレベルの区分を決定する；
ことを含む、方法。

【請求項 2】

DNA断片の第1セットは、生物のゲノムの1つまたは複数の所定領域に対応する、請求項1に記載の方法。

【請求項 3】

生体試料からDNA断片の追加のセットを同定し、ここで当該DNA断片の追加のセットの各セットは、前記生物のゲノムの異なる所定領域に対応し；

異なる所定領域に対応する該DNA断片の追加のセットの各々のセットについて；

複数のサイズに対応するDNA断片の量を決定し；

第1パラメータのサイズ値を計算し；

該サイズ値を、それぞれの基準値と比較し；そして

該サイズ値がそれぞれの基準値と統計学的に異なる所定領域を同定する；

ことをさらに含む、請求項2に記載の方法。

10

【請求項 4】

同定した所定領域を利用して、1つまたは複数の可能性のある癌の種類を決定する、ここで可能性のある癌の種類は同定された所定領域に関連する；

ことをさらに含む、請求項3に記載の方法。

【請求項 5】

生物における癌のレベルの区分を比較に基づいて決定することは；

対応するサイズ値がそれぞれの基準値と統計学的に異なる同定された所定領域の数を決定し；そして、

その数を領域の閾値の数と比較して、生物における癌のレベルの区分を決定する；

ことを含む、請求項3に記載の方法。

20

【請求項 6】

それぞれの基準値は、前記DNA断片の追加のセットの各々の第1パラメータのサイズ値の少なくとも2つについて異なる、請求項3に記載の方法。

【請求項 7】

決定する区分は、腫瘍のサイズにまたは腫瘍の数に対応する、請求項1に記載の方法。

【請求項 8】

生体試料は治療後の生物から得られ、そして基準値は治療前の該生物から採取した試料から決定された第1パラメータの値に対応する、請求項1に記載の方法。

30

【請求項 9】

基準値は、生物が癌を有しないと推測される場合の試料から決定される第1パラメータの値に対応する、請求項1に記載の方法。

【請求項 10】

基準値は、1つまたは複数の健康な生物から得られた1つまたは複数の生体試料から確立される、請求項1に記載の方法。

【請求項 11】

区分は、生物が癌を有していない、または生物が癌を有しているということである、請求項1に記載の方法。

【請求項 12】

区分は、癌のレベルが低下した、または癌のレベルが増加したということである、請求項1に記載の方法。

40

【請求項 13】

生物はヒトである、請求項1に記載の方法。

【請求項 14】

前記複数のDNA断片の第1のセットは、ランダムに選択される、請求項1に記載の方法。

【請求項 15】

実行される場合、請求項1～14のいずれか1項に記載の方法の操作を実行するコンピュータシステムを制御する複数の命令を記憶したコンピュータ可読媒体。

50

【発明の詳細な説明】

【技術分野】

【0001】

関連出願の相互参照

本出願は、2012年3月8日に出願された「SIZE-BASED ANALYSIS OF FETAL DNA FRACTION IN MATERNAL PLASMA」と題する米国仮特許出願第61/608,623号、および2012年4月6日に出願された「SIZE-BASED ANALYSIS OF FETAL DNA FRACTION IN MATERNAL PLASMA」と題する米国仮特許出願第61/621,451号の優先権の利益を主張する非仮特許出願であり、それらの全内容は、

10

【背景技術】

【0002】

母親血漿中における無細胞胎児DNAの発見が、非侵襲的な出生前診断の新たな可能性を切り開いてきた(Lo YMD et al. Lancet 1997; 350: 485-487)。分画胎児DNA濃度の平均/中央値は、約3%~10%であることが報告されている(Lo YMD et al. Am J Hum Genet 1998; 62: 768-775; Lun FMF et al. Clin Chem 2008; 54: 1664-1672)。分画胎児DNA濃度は、母体血漿DNAを用いた非侵襲的な出生前診断テストの性能に影響する重要なパラメータである。例えば、胎児染色体の異数体(例えば、トリソミー21、トリソミー18、またはトリソミー13)についての非侵襲的な出生前診断では、分画胎児DNA濃度が高いほど、母体血漿中における異数染色体由来のDNA配列がより過剰に発現する。実際に、母体血漿中の分画胎児DNA濃度を2分の1にすると、異数性を検出するためにカウントするのが必要な分子数は4倍になる(Lo YMD et al. Proc Natl Acad Sci USA 2007; 104: 13116-13121)。

20

【0003】

ランダムな大量並列シーケンシングによる胎児トリソミーの非侵襲的な出生前検出では、試料の分画胎児DNA濃度が、ロバストな検出を達成するために必要なシーケンシングの量に影響を与える(Fan HC and Quake SR. PLoS One 2010; 5: e10439)。実際、多くのグループでは、品質管理工程を含め、この工程で分画胎児DNA濃度をまず測定し、最小分画胎児DNA濃度より多くを含む試料のみが診断結果を出すのに適格とした(Palomaki GE et al. Gene Med 2011; 13: 913-920)。他のグループでは、特定の母体血漿試料が異数性妊娠由来であるリスクを推定するための診断アルゴリズムに、分画胎児DNA濃度を含めた(Sparks AB et al. Am J Obstet Gynecol 2012; 206: 319.e1-9)。

30

【0004】

異数性の検出に加えて、分画胎児DNA濃度は、例えば、異常ヘモグロビン症(Lun FMF et al. Proc Natl Acad Sci USA 2008; 105: 19920-19925)や血友病(Tsui NBY et al. Blood 2011; 117: 3684-3691)などの単一遺伝子疾患を検出するための母性血漿DNAを用いた非侵襲的な出生前診断テストにも、同様に影響を与える。また、分画胎児DNA濃度は、胎児の全ゲノムシーケンシングに加え、胎児の遺伝および突然変異についてのゲノムマップを構築するために行う必要があるシーケンシングの深さにも影響を与える(Lo YMD et al. Sci Transl Med 2010; 2: 61ra91 and U.S. Patent Application 2011/0105353)。

40

【0005】

分画胎児DNA濃度を測定するための多くの方法が記載されている。一つのアプローチ

50

は、母系ゲノムには存在しない胎児特異的な父系配列の濃度を測定することである。そのような配列の例として、男児が有するY染色体上にある配列、およびRh(D)陰性の妊婦が宿しているRh(D)陽性胎児が有するRHD遺伝子由来の配列が挙げられる。また、母体および胎児の両方に存在する配列を用いて、全母体血漿DNAを測定することができる。それから、分画胎児DNA濃度に達するために、全母体血漿DNAの濃度を超える胎児特異的な父系配列の濃度の比率を計算できる。

【0006】

使用可能な配列の別の例として、一塩基多型の使用が挙げられる(Lo YMD et al. Sci Transl Med 2010; 2: 61ra91)。分画胎児DNA濃度を測定するための遺伝子マーカーを使用することの欠点は、どの遺伝子マーカーセットも単体では全ての胎児-母体ペアについての情報を提供できないことである。使用可能な更に別の方法として、母体血漿における胎児または胎盤特異的なDNAメチル化パターンを示すDNA配列を使用することが挙げられる(Nygren AO et al. Clin Chem 2010; 56: 1627-1635)。DNAメチル化マーカーを使用することについての潜在的な欠点は、DNAメチル化のレベルは個体間で差があるということである。更に、DNAメチル化マーカーの検出に使用される方法は、通常複雑であり、メチル化感受性制限酵素による消化(Chan KCA et al. Clin Chem 2008; 52: 2211-2218)または亜硫酸水素塩転換(Chim SSC et al. Proc Natl Acad Sci USA 2005; 102: 14753-14758)またはメチル化DNA免疫沈降(MeDIP)(Papageorgiou EA et al. Nat Med 2011; 17: 510-513)を利用することなどが含まれる。

【0007】

分画胎児DNA濃度は重要な値なので、その値を決定するための追加的な方法およびシステムを有することが望ましい。

【発明の概要】

【0008】

実施形態は、生体試料由来のDNAの混合物における臨床的に関連するDNAの分画濃度を、様々なサイズのDNA断片の量に基づいて推定するための方法およびシステムを提供し得る。例えば、母体の血漿における胎児DNAまたは患者の血漿における腫瘍DNAの分画濃度が測定できる。DNA断片のサイズは、胎児DNAの割合および腫瘍DNAの割合と相関していることが示された。(例えば、校正関数としての)校正データ点は、サイズパラメータの値と臨床的に関連するDNAの分画濃度との間の対応関係を示す。ある試料について、サイズパラメータの第1値は、試料におけるDNA断片のサイズから決定できる。第1値を校正データ点と比較することにより、臨床的に関連するDNAの分画濃度が推定できる。

【0009】

一実施形態によれば、方法は、生体試料における臨床的に関連するDNAの分画濃度を推定するものであり、ここで生体試料は臨床的に関連するDNAおよびその他のDNAを含む。複数のサイズのうちの各サイズについて、そのサイズに対応する生体試料由来の複数のDNA断片の量を測定する。コンピュータシステムは、第1パラメータの第1値を複数のサイズのDNA断片の量に基づいて計算する。第1パラメータは、生体試料におけるDNA断片のサイズプロファイルの統計的尺度を提供する。1つまたは複数の第1校正データ点が得られる。各第1校正データ点は、第1パラメータの校正值に対応する臨床的に関連するDNAの分画濃度を特定する。1つまたは複数の校正データ点は、複数の校正試料から決定される。第1値を少なくとも1つの校正データ点の校正值と比較する。生体試料における臨床的に関連するDNAの分画濃度を当該比較に基づいて推定する。

【0010】

別の実施形態によれば、方法は、生物の生体試料を解析するものである。生体試料は、正常細胞由来のDNAと、癌に関連しているおそれがある細胞由来のDNAを含む。少な

10

20

30

40

50

くともDNAの一部は、生体試料における無細胞DNAである。複数のサイズのうちの各サイズについて、そのサイズに対応する生体試料由来の複数のDNA断片の量を測定する。コンピュータシステムは、第1パラメータの第1値を複数のサイズのDNA断片の量に基づいて計算する。第1パラメータは、生体試料におけるDNA断片のサイズプロファイルの統計的尺度を提供する。第1値を基準値と比較する。生物の癌のレベルの区分を当該比較に基づいて決定する。

【0011】

他の実施形態は、本明細書に記載の方法に関連するシステム、携帯可能な消費者用装置、およびコンピュータ可読媒体に関する。

【0012】

本発明の性質および利点は、以下の詳細な説明および添付の図面を参照してより良く理解され得る。

【図面の簡単な説明】

【0013】

【図1】図1は、本発明の実施形態に従った、母体血漿中の循環無細胞DNAのサイズ分布についてのプロット100を示す。

【0014】

【図2】図2Aは、本発明の実施形態に従った、異なる分画胎児DNA濃度を有する2つの母体血漿試料における胎児DNA（妊娠第1期）のサイズ分布についてのプロット200を示す。図2Bは、本発明の実施形態に従った、異なる分画胎児DNA濃度を有する2つの母体血漿試料における胎児DNA（妊娠第2期）のサイズ分布についてのプロット250を示す。

【0015】

【図3】図3は、本発明の実施形態に従った、生体試料における臨床的に関連するDNAの分画濃度を推定する方法を示す方法300のフローチャートである。

【0016】

【図4】図4は、本発明の実施形態に従った、電気泳動を利用して得られる母体血漿DNAのサイズ分布（エレクトロフェログラム）を示すプロット400である。

【0017】

【図5】図5Aは、本発明の実施形態に従った、母体血漿における胎児DNAの割合が様々な試料について150bp以下のDNA断片の割合を示すプロット500である。図5Bは、 $(CF(\text{サイズ } 150) / \text{サイズ}(163 \sim 169))$ と表する150bp以下のDNA断片と163bp～169bpのDNAとの量のサイズ比を示すプロット550である。

【0018】

【図6】図6Aは、 $(\text{サイズ}(140 \sim 146) / \text{サイズ}(163 \sim 169))$ と表する140bp～146bpのDNA断片と163bp～169bpのDNAとの量のサイズ比を示すプロット600である。図6Bは、 $(\text{サイズ}(140 \sim 154) / \text{サイズ}(163 \sim 169))$ と表する140bp～154bpのDNA断片と163bp～169bpのDNAとの量のサイズ比を示すプロット650である。

【0019】

【図7】図7は、 $(\text{サイズ}(100 \sim 150) / \text{サイズ}(163 \sim 169))$ と表する100bp～150bpのDNA断片と163bp～169bpのDNAとの量のサイズ比を示すプロット700である。

【0020】

【図8】図8は、本発明の実施形態に従った、母体血漿における胎児DNAの割合が様々な試料について150bp以下のDNA断片の割合を示すプロット800である。

【0021】

【図9】図9Aは、 $(CF(\text{サイズ } 150) / \text{サイズ}(163 \sim 169))$ と表する150bp以下のDNA断片と163bp～169bpのDNAとの量のサイズ比を

10

20

30

40

50

示すプロット900である。図9Bは、(サイズ(140~146)/サイズ(163~169))と表する140bp~146bpのDNA断片と163bp~169bpのDNAとの量のサイズ比を示すプロット950である。

【0022】

【図10】図10Aは、(サイズ(140~154)/サイズ(163~169))と表する140bp~154bpのDNA断片と163bp~169bpのDNAとの量のサイズ比を示すプロット1000である。図10Bは、(サイズ(100~150)/サイズ(163~169))と表する100bp~150bpのDNA断片と163bp~169bpのDNAとの量のサイズ比を示すプロット1005である。

【0023】

【図11】図11は、本発明の実施形態に従った、反復要素のサイズについて胎児DNAの割合に対しプロットしたサイズ比を示すプロットである。

【0024】

【図12】図12Aは、本発明の実施形態に従った、サイズ比を決定するために使用できるエレクトロフェログラム1200である。図12Bは、本発明の実施形態に従った、母体血漿における胎児DNAの割合が様々な試料について200bp~267bpのDNA断片と290bp~294bpのDNAとの量のサイズ比を示すプロット1250である。

【0025】

【図13】図13は、本発明の実施形態に従った、校正試料から得た測定値より校正データ点を決定するための方法1300のフローチャートである。

【0026】

【図14】図14Aは、本発明の実施形態に従った、トレーニングセットについての胎児DNAの分画濃度に対するサイズ比のプロット1400である。図14Bは、本発明の実施形態に従った、胎児特異的配列を用いて測定した分画濃度に対する図14Aの線形関数1410から導き出した(推定した)分画濃度のプロット1450である。

【0027】

【図15】図15Aは、本発明の実施形態に従った、腫瘍切除前後の2人の肝細胞癌(HCC)患者の血漿における腫瘍DNAの割合が様々な試料について150bp以下のDNA断片の割合を示すプロット1500である。図15Bは、腫瘍切除前後の2人のHCC患者について、(CFサイズ(サイズ150)/サイズ(163~169))と表する150bp以下のDNA断片と163bp~169bpのDNAとの量のサイズ比を示すプロット1550である。

【0028】

【図16】図16Aは、腫瘍切除前後の2人のHCC患者について、(サイズ(140~146)/サイズ(163~169))と表する140bp~146bpのDNA断片と163bp~169bpのDNAとの量のサイズ比を示すプロット1600である。図16Bは、腫瘍切除前後の2人のHCC患者について、(サイズ(140~154)/サイズ(163~169))と表する140bp~154bpのDNA断片と163bp~169bpのDNAとの量のサイズ比を示すプロット1650である。

【0029】

【図17】図17は、腫瘍切除前後の2人のHCC患者について、(サイズ(100~150)/サイズ(163~169))と表する100bp~150bpのDNA断片と163bp~169bpのDNAとの量のサイズ比を示すプロット1700である。

【0030】

【図18】図18Aは、腫瘍切除前後のHCC患者について、150bp以下のDNA断片の割合を示すプロット1800である。図18Bは、腫瘍切除前後のHCC患者について、(CFサイズ(サイズ150)/サイズ(163~169))と表する150bp以下のDNA断片と163bp~169bpのDNAとの量のサイズ比を示すプロット1850である。

10

20

30

40

50

【0031】

【図19】図19Aは、腫瘍切除前後のHCC患者について、(サイズ(140~146)/サイズ(163~169))と表する140bp~146bpのDNA断片と163bp~169bpのDNAとの量のサイズ比を示すプロット1900である。図19Bは、腫瘍切除前後のHCC患者について、(サイズ(140~154)/サイズ(163~169))と表する140bp~154bpのDNA断片と163bp~169bpのDNAとの量のサイズ比を示すプロット1950である。

【0032】

【図20】図20は、腫瘍切除前後のHCC患者について、(サイズ(100~150)/サイズ(163~169))と表する100bp~150bpのDNA断片と163bp~169bpのDNAとの量のサイズ比を示すプロット2000である。

10

【0033】

【図21】図21は、本発明の実施形態に従った、生物の生体試料を解析して癌のレベルの区分を決定するための方法2100を示すフローチャートである。

【0034】

【図22】図22は、様々な種類の癌で見られる一般的な染色体異常のいくつかを示す表2200である。

【0035】

【図23】図23は、本発明の実施形態に係るシステムおよび方法に使用可能な例としてコンピュータシステム2300の概略図を示す。

20

【発明を実施するための形態】

【0036】

定義

本明細書で使用する場合、用語「生体試料」とは、被験者(例えば、妊婦などのヒト)から採取され、対象とする1つまたは複数の核酸分子を含有するいずれかの試料を指す。例として、血漿、唾液、胸水、汗、腹水、胆汁、尿、血清、腓液、糞便および子宮頸部スメア試料が挙げられる。生体試料は、ヒト、動物、または他の適切な生物から得ることができる。「校正試料」は、生体試料に対応するもので、生体試料の臨床的に関連するDNA画分が、例えば、臨床的に関連するDNAに特異的なアレルを使用する方法等の校正方法によって分かるまたは決定されるものである。臨床的に関連するDNAの例としては、母体の血漿における胎児DNAまたは患者の血漿における腫瘍DNAが挙げられる。

30

【0037】

本明細書で使用する場合、用語「座位(locus)」又はその複数形の「座位(loci)」とは、ゲノム間で多様性を有する任意の長さのヌクレオチド(又は塩基対)の位置又はアドレスのことである。用語「読み取り配列」とは、核酸分子、例えば、DNA断片の全部または一部から得られる配列を指す。一実施形態では、断片の一方の末端のみがシーケンシングされる。代わりに、断片の両方の末端(例えば、各末端から約30bp)をシーケンシングして2つの読み取り配列を生成してもよい。その後、この一対の読み取り配列を参照ゲノムにアラインして、断片の長さが分かるようにしてもよい。更に別の実施形態では、直鎖状DNA断片を、例えば、ライゲーションにより環状化してもよく、ライゲーション部位にまたがる部分をシーケンシングしてもよい。

40

【0038】

用語「ユニバーサルシーケンシング」とは、アダプターを断片の末端に加え、シーケンシング用のプライマーをアダプターに結合させるシーケンシングのことを指す。従って、任意の断片を同じプライマーを用いてシーケンシングすることができ、シーケンシングがランダムにできる。

【0039】

分画胎児DNA濃度という用語は、胎児DNAの割合および胎児DNA画分という用語と交換可能に使用され、生体試料(例えば、母体血漿または血清試料)中に存在する胎児から由来した胎児DNA分子の割合を指す(Lo YMD et al. Am J H

50

um Genet 1998; 62: 768 - 775; Lun FMF et al. Clin Chem 2008; 54: 1664 - 1672)。同様に、分画腫瘍DNA濃度という用語は、腫瘍DNAの割合および腫瘍DNA画分という用語と交換可能に使用され、生体試料中に存在する腫瘍DNA分子の割合を指す。

【0040】

用語「サイズプロファイル」は、一般に、生体試料中におけるDNA断片のサイズに関するものである。サイズプロファイルは、様々なサイズのDNA断片の量の分布を示すヒストグラムであってもよい。様々な統計パラメータ（サイズパラメータまたは単にパラメータとも呼ばれる）が、あるサイズプロファイルと別のサイズプロファイルとを区別するために使用できる。パラメータの1つとして、全DNA断片に対するまたは他のサイズ範囲のDNA断片に対する、特定のサイズまたはサイズ範囲のDNA断片の割合がある。

10

【0041】

「臨床的に関連する」DNAの例としては、母体の血漿における胎児DNAおよび患者の血漿における腫瘍DNAが挙げられる。別の例として、移植患者の血漿における移植片に関連するDNAの量の測定値が挙げられる。更なる例として、被験者の血漿における造血DNAと非造血DNAとの相対量の測定値が挙げられる。この後者の実施形態は、造血および/または非造血組織が関連する病理学的プロセスまたは傷害を検出またはモニタリングまたは予知するのに使用できる。

【0042】

「校正データ点 (calibration data point)」は、「校正值 (calibration value)」および対象とするDNA（すなわち、臨床的に関連するDNA）について測定された分画濃度あるいは既知の分画濃度を含む。校正值とは、臨床的に関連するDNAの分画濃度が既知である校正試料について決定されたサイズパラメータの値である。校正データ点は、様々な方法で定義でき、例えば、離散点として、または校正関数（検量線 (calibration curve)）もしくは検量面 (calibration surface) とも呼ばれる）として定義できる。

20

【0043】

用語「癌のレベル」とは、癌が存在するか否か、癌の段階、腫瘍のサイズ、ある染色体領域において欠失または増幅がいくつあるか（例えば、2重または3重）、および/または癌の重症度についてのその他の尺度を指すことがある。癌のレベルは数字またはその他の記号であってもよい。レベルはゼロの場合もある。癌のレベルには、欠失または増幅に関連する前悪性または前癌状態も含まれる。

30

【0044】

詳細な説明

母体血漿中に存在する無細胞胎児DNA分子は、一般的に、母親由来の分子よりも短いことが知られている (Chan KCA et al. Clin Chem 2004; 50: 88 - 92; Lo YMD et al. Sci Transl Med 2010; 2: 61ra91)。胎児DNAの存在により、母体血漿DNAの全体的なサイズ分布がシフトし、そのシフトの程度が胎児DNAの分画濃度と関連する。母体血漿DNAのサイズプロファイルの特定の値を測定することにより、実施形態で母体血漿中の分画胎児DNA濃度を得ることができる。

40

【0045】

また、非侵襲的な出生前診断における用途とは別に、実施形態は、癌の検出、移植、および医学的なモニタリングをするのに有用な可能性のある生物の液中にある様々なサイズの臨床的に有用な核酸種の分画濃度を測定するのに使用できる。以前から、癌患者の血漿中では、腫瘍由来のDNAは非癌由来のDNAよりも短いことが示されている (Diehl F et al. Proc Natl Acad Sci USA 2005; 102: 16368 - 16373)。移植関連では、造血由来のDNAは、非造血DNAよりも短いことが示されている (Zheng YW et al. Clin Chem 2012; 58: 549 - 558)。例えば、患者がドナーから肝臓を受けると、その血

50

漿中では肝臓（成人では非造血器官）由来のDNAが造血由来のDNAよりも短くなる（Zheng YW et al. Clin Chem 2012; 58: 549 - 558）。同様に、心筋梗塞または脳卒中の患者では、損傷を受けた非造血器官（すなわち、それぞれ心臓および脳）から放出されたDNAにより、血漿DNAのサイズプロファイルは短いサイズ範囲にシフトする結果となることが期待されるであろう。

【0046】

I. サイズ分布

実施形態を実証するために、例えば、ペアエンド大量並列シーケンシングあるいは（例えば、パイオアナライザーを使用した）電気泳動により、サイズプロファイルを測定することができることを以下の実施例で示す。パイオアナライザーを用いた電気泳動は迅速かつ比較的安価な手順であるため、後者の例が特に有用である。これにより、血漿DNA試料を比較的高価なシーケンシングプロセスに供する前に、この解析を品質管理手段として迅速に実行できるだろう。

【0047】

図1は、本発明の実施形態に従った、母体血漿中の循環無細胞DNAのサイズ分布についてのプロット100を示す。サイズ分布は、DNA断片のサイズを測定し、その後、様々なサイズ、例えば、50塩基～約220塩基の範囲内のDNA断片の数を数えることで得ることができる。プロット100は、2つの分布を示す。分布110は、母体血漿試料における全DNA断片についてのものであり、分布120は、胎児由来のDNAのみについてのものである。横軸は、DNA断片の塩基対（bp）のサイズである。縦軸は、測定したDNA断片の割合である。

【0048】

図1では、母体血漿における胎児由来のDNAのサイズ分布は、母親由来のものよりも短いことが示される（Chan KC et al. Clin Chem 2004; 50: 88 - 92）。近年、我々は、妊婦における胎児特異的DNAおよび（主に母親由来の）全DNAについて解像度の高いサイズ分布を決定するために、ペアエンド超並列配列解析を使用してきた。我々は、DNAにおける2つの種の間の主な違いは、胎児由来のDNAだと166bpのDNA断片の分画が減少し、150bp以下の短いDNAの割合が増加することであることを示した（Lo YM et al. Sci Transl Med 2010; 2: 61ra91）。

【0049】

ここで、我々は、母体血漿試料（生体試料の一例）における全DNA断片のサイズ分布の解析をすることが、母体血漿における胎児DNAの分画濃度を測定するのにいかに有用であるかを概説する。母体血漿における胎児DNAの分画濃度が増加すると、全DNAについての全体的なサイズ分布が短くなる。一実施形態では、約144bpのDNA断片と約166bpのDNA断片との相対量（パラメータの一例）が、胎児DNAの分画濃度を反映させるのに使用できる。別の実施形態では、サイズプロファイルに関する他のパラメータまたはパラメータの組み合わせが血漿DNAのサイズ分布を反映させるのに使用できる。

【0050】

図2Aは、本発明の実施形態に従った、異なる分画胎児DNA濃度を有する2つの母体血漿試料における胎児DNA（妊娠第1期）のサイズ分布についてのプロット200を示す。これら二人の妊婦は、いずれも男児を宿していた。分画胎児DNA濃度は、全シーケンシングされたDNA断片におけるY染色体由来の配列の割合から決定した。両方の試料は、妊娠第1期の妊婦から採取した。ケース338（実線、分画胎児DNA濃度は10%）は、ケース263（点線、分画胎児DNA濃度は20%）よりも低い分画胎児DNA濃度を有していた。ケース263と比較すると、ケース338は、166bpにおけるピークは高かったが、150bp未満のサイズのピークは低かった。言い換えると、ケース263では150bp未満のDNA断片がより多く、一方ケース338では約166bpの断片がより多かった。これらの観察は、短いDNAとおおよび長いDNAとの相対量が分画

胎児DNA濃度に相関するという仮説と一致する。

【0051】

図2Bは、本発明の実施形態に従った、異なる分画胎児DNA濃度を有する2つの母体血漿試料における胎児DNA（妊娠第2期）のサイズ分布についてのプロット250を示す。両方の試料は、妊娠第2期の妊婦から採取した。これら二人の妊婦は、いずれも男児を宿していた。分画胎児DNA濃度は、全シーケンシングされたDNA断片におけるY染色体由来の配列の割合から決定した。先の実施例と同様に、ケース5415（点線、分画胎児DNA濃度は19%と高め）は、150bp未満のサイズのピークが高く、一方ケース5166（実線、分画胎児DNA濃度は12%と低め）では166bpにおけるピークが高かった。

10

【0052】

分画胎児DNA濃度の値に対する異なる値のサイズパラメータの相関を、以下のデータプロットに示す。さらに、腫瘍DNAの断片のサイズは、腫瘍DNA断片および正常細胞由来のDNA断片を有する試料における腫瘍DNA断片の割合に相関する。したがって、腫瘍断片のサイズも、試料における腫瘍断片の割合を決定するのに使用できる。

【0053】

II. 方法

DNA断片のサイズは分画濃度（割合とも呼ばれる）に相関しているので、実施形態では、この相関関係を利用して試料における特定の種類のDNA（例えば、胎児DNAまたは腫瘍由来のDNA）の分画濃度を決定できる。特定の種類のDNAは、臨床的に関連するDNAであり、その分画濃度が推定対象である。したがって、この方法は、測定したDNA断片のサイズに基づいて、生体試料における臨床的に関連するDNAの分画濃度を推定することができる。

20

【0054】

図3は、本発明の実施形態に従った、生体試料における臨床的に関連するDNAの分画濃度を推定する方法を示す方法300のフローチャートである。生体試料は、臨床的に関連するDNAおよびその他のDNAを含む。生体試料は、患者、例えば、胎児を宿している妊婦の被験体から得ることができる。別の実施形態では、患者は腫瘍を有してもよく、または有すると疑われる者であってもよい。一実施形態では、生体試料は、DNA断片のサイズを決定するのに利用可能な測定データ（例えば、読み取り配列）を出力する装置、例えば、シーケンシング装置に受けることができる。方法300は、本明細書に記載される他の方法と同様、全部または一部をコンピュータシステムによって実行してもよい。

30

【0055】

ブロック310では、様々なサイズに対応するDNA断片の量を測定する。複数のサイズのうちの各サイズについて、そのサイズに対応する生体試料由来の複数のDNA断片の量を測定することができる。例えば、140塩基の長さを有するDNA断片の数を測定できる。この量をヒストグラムとして保存してもよい。一実施形態では、生体試料由来の複数の核酸のうちの各核酸のサイズを測定する。この測定は個別に行ってもよく（例えば、一分子のシーケンシングにより）、またはグループ単位で行ってもよい（例えば、電気泳動により）。サイズは、ある範囲に対応することもある。したがって、量は、サイズが特定の範囲内にあるDNA断片であってもよい。

40

【0056】

複数のDNA断片を、ランダムに選択またはゲノムの1つまたは複数の所定の領域から優先的に選択できる。例えば、上述したように、標的の濃縮を行ってもよい。別の実施形態では、DNA断片は、ランダムに（例えば、ユニバーサルシーケンシングを用いて）シーケンシングでき、得られた読み取り配列を、被験体（例えば、参照ヒトゲノム）に対応するゲノムにアラインさせることができる。そして、読み取り配列が1つまたは複数の所定の領域にアラインするDNA断片のみをサイズを決定するのに使用できる。

【0057】

様々な実施形態では、サイズは、質量、長さ、または他の適切なサイズ測定値であり得

50

る。本明細書に記載のように、測定は様々な方法で行うことができる。例えば、ペアエンドシーケンシングおよびDNA断片のアライメントを行ってもよく、または電気泳動を利用してもよい。生体試料の正確なサイズプロファイル提供するために、統計学的に有意なDNA断片数を測定してもよい。統計学的に有意なDNA断片数の例として、100, 000超; 1, 000, 000超; 2, 000, 000超、または他の適切な値が挙げられるが、これは求める精度に依存しうる。

【0058】

一実施形態では、ペアエンドシーケンシングまたは電気泳動などの物理的測定から得られたデータを、コンピュータで受信し、解析して、DNA断片のサイズの測定を達成してもよい。例えば、ペアエンドシーケンシングによる読み取り配列を（例えば、アラインメントにより）解析してサイズを決定してもよい。別の例として、電気泳動から得られた電気泳動図を解析してサイズを決定してもよい。一実施形態では、DNA断片の解析には、シーケンシングまたはDNA断片を電気泳動にかけるといった実際のプロセスを含むが、別の実施形態では、得られたデータの解析を行うのみであってもよい。

【0059】

ブロック320では、第1パラメータの第1値を複数のサイズのDNA断片の量に基づいて計算する。一態様では、第1パラメータは、生体試料におけるDNA断片のサイズプロファイルの統計的尺度（例えば、ヒストグラム）を提供する。このパラメータは、複数のDNA断片のサイズから決定されるため、サイズパラメータと呼ぶことがある。

【0060】

第1パラメータは、様々な形態をとり得る。このようなパラメータが、特定のサイズのDNA断片の数を全断片数で除した値であり、これはヒストグラム（特定のサイズの断片の絶対的または相対的な数を提供する任意のデータ構造）から得ることができる。別の例として、パラメータは、特定のサイズまたは特定の範囲の断片の数を、別のサイズまたは範囲の断片数で除した値であることもある。除算することで、異なる試料について、解析対象のDNA断片数の違いを考慮した正規化として機能し得る。正規化は、各試料について同じ数のDNA断片を解析することによって達成できるが、これは解析対象の全断片数で除することでも同じ結果が得られる。他のパラメータの例を、本明細書に記載する。

【0061】

ブロック330では、1つまたは複数の第1校正データ点を得る。各第1校正データ点は、第1パラメータの特定の値（校正値）に対応する臨床的に関連するDNAの分画濃度を特定できる。分画濃度は、特定の濃度または濃度範囲として特定できる。校正値は、複数の校正試料から決定された第1パラメータ（すなわち、特定のサイズパラメータ）の値に対応しうる。校正データ点は、既知の分画濃度を有する校正試料から決定でき、分画濃度は、本明細書に記載の様々な技術によって測定できる。少なくともいくつかの校正試料は異なる分画濃度を有するが、いくつかの校正試料は同じ分画濃度を有していてもよい。

【0062】

様々な実施形態では、1つまたは複数の校正点は、1つの離散点として、離散点のセットとして、関数として、1つの離散点と関数として、あるいは他の離散的または連続的な値のセットの組み合わせとして、定義してもよい。例として、校正データ点は、特定の分画濃度の試料について、サイズパラメータ（例えば、特定のサイズまたはサイズ範囲内の断片数）の1つの校正値から決定できる。複数のヒストグラムを、各校正試料についての別のヒストグラムとともに使用してもよく、そのうちのいくつかの校正試料は同じ分画濃度を有してもよい。

【0063】

一実施形態では、同じ分画濃度の複数の試料から得た同じサイズパラメータについての測定値を組み合わせ、特定の分画濃度についての校正データ点を決定できる。例えば、サイズパラメータの値の平均を同じ分画濃度の試料のサイズデータから得て、特定の校正データ点を決定（または校正データ点に対応する範囲を提供）できる。別の実施形態では、同じ校正値を有する複数のデータ点を使用して、平均分画濃度を決定できる。

【 0 0 6 4 】

一実施態様では、多くの校正試料についてDNA断片のサイズを測定する。同じサイズパラメータの校正值を各校正試料について測定する。各校正試料についてのサイズパラメータを、試料の既知の分画濃度に対しプロットしてもよい。関数当てはめにより新しい試料の分画濃度を決定するのに使用できる校正データ点を定義する関数をプロットのデータ点に当てはめてもよい。

【 0 0 6 5 】

ブロック340では、第1値を少なくとも1つの校正データ点の校正值と比較する。比較は様々な方法で行うことができる。例えば、比較は、第1値が校正值よりも高いか低いかということであってもよい。比較は、(校正データ点からなる)検量線と比較することを含み得るので、比較により、第1パラメータの第1値を有する曲線上の点を同定することができる。例えば、(新しい試料中におけるDNAサイズの測定値から決定される)第1パラメータの計算値 X は、 F が校正関数(曲線)である関数 $F(X)$ に代入するのに使用できる。 $F(X)$ の算出値は、分画濃度である。誤差範囲は、各 X 値について異なってもよく、これにより $F(X)$ の算出値の範囲が設けられる。

10

【 0 0 6 6 】

ステップ350では、生体試料における臨床的に関連するDNAの分画濃度を、比較に基づいて推定する。一実施形態では、第1パラメータの第1値が校正閾値より大きいまたは小さいか否かを決定し、これにより当該試料について推定した分画濃度が校正閾値に対応する分画濃度より大きいまたは小さいか否かを決定する。例えば、生体試料について計算した第1値 X_1 が校正值 X_c より大きい場合、この生体試料の分画濃度 FC_1 は X_c に対応する分画濃度 FC_c より大きいと決定できる。この比較は、例えば、胎児の異数性のテストなど他のテストを行うのに生体試料中に十分な分画濃度があるか否かを決定するのに使用できる。この大きい小さいという関係は、パラメータをいかに定義するか依存し得る。このような実施形態では、1つの校正データ点のみが必要であってもよい。

20

【 0 0 6 7 】

別の実施形態では、比較は、校正関数に第1値を代入することによって達成される。校正関数は、第1値に対応する曲線上の点を同定することによって、第1値を効果的に校正值と比較することができる。推定した分画濃度を、校正関数の算出値として設ける。

【 0 0 6 8 】

一実施形態では、生体試料についての複数のパラメータの値を決定できる。例えば、その生体試料におけるDNA断片のサイズプロファイルが異なる統計的尺度に対応する第2パラメータについての第2値を決定できる。第2値は、DNA断片の同じサイズ尺度、または異なるサイズ尺度を用いて決定できる。各パラメータは、異なる検量線に対応しうる。一実施態様では、異なる値を、独立して異なる検量線と比較して複数の推定分画濃度を求めることができ、これらを平均してもよく、または算出範囲を設けるために用いてもよい。

30

【 0 0 6 9 】

別の実施態様では、異なる値のパラメータを分画濃度を算出する単一の校正関数に効果的に代入できる多次元検量線を用いてもよい。単一の校正関数は、校正試料から得られた全てのデータ点を関数当てはめすることにより得られる。したがって、一実施形態では、第1校正データ点および第2校正データ点を多次元曲線上の点とすることができ、この曲線の場合、比較は、第1値および1つまたは複数の第2値に対応する座標を有する多次元の点を同定することを含む。

40

【 0 0 7 0 】

III. サイズの決定

血漿DNAのサイズ分布は、例えば、これらに限定されないが、リアルタイムPCR、電気泳動、および質量分析を利用して決定できる。様々な実施形態では、測定するサイズは、長さ、分子量、または長さもしくは分子量に比例する測定パラメータ、例えばエレクトロフェログラムにおける移動度や電気泳動もしくは質量分析計において一定距離を移動

50

するのにかかる時間などである。別の例では、DNAを、結合した染料の量が、DNA分子の長さに比例するように、インターカレーター性蛍光色素、例えば臭化エチジウム又はSYBRグリーンで染色できる。UV光が試料に照射されたとき発する蛍光の強度によって結合した色素の量を決定できる。サイズを測定するためのいくつかの例およびその結果のデータを以下に記載する。

【0071】

A．シーケンシングを利用した第1胎児試料セット

表1は、例として、胎児DNA画分に関する試料情報およびシーケンシング解析を示す。血漿試料は、それぞれ1人の男児を宿している80人の妊婦から採取した。80人の妊婦のうち、39人は正倍数体の胎児を宿しており、18人はトリソミー21(T21)の胎児を宿しており、10人はトリソミー18(T18)の胎児を宿しており、13人はトリソミー13(T13)の胎児を宿していた。血漿DNAのサイズ分布は、ペアエンド大量並列シーケンシングを用いて決定した。母体血漿DNAのシーケンシングライブラリーは、6塩基のバーコードを3回のプライマーPCR増幅により各血漿試料のDNA分子へ導入した以外は従前に記載(Lo YM et al. Sci Transl Med 2010; 2: 61ra91)のように構築した。

【0072】

2つの試料を1シーケンシングレーン(すなわち2プレックスシーケンシング)に導入した。別の実施形態では、2個より多くの試料、例えば、6個、または12個、または20個、または20個を超える試料を1シーケンシングレーンに導入できる。全ライブラリーは、36bp×2PEフォーマットを用いたGenome Analyzer IIX(Illumina)によりシーケンシングした。更に7サイクルのシーケンシングを行い、それぞれシーケンシングした血漿DNA分子上のインデックス配列をデコードした。36bpの読み取り配列は、Short Oligonucleotide Alignment Program 2(SOAP2)(soap.genomics.org.cn)を用いてリピート領域をマスクしないヒト参照ゲノム(Hg18)(genome.ucsc.edu)にアラインした。個々のメンバーのペアエンド(PE)読み取り配列は、フローセル上の同じクラスタ位置でシーケンシングし、そして、ヒトゲノムの単一の位置に正しい方向で特異的にアラインした。ここで、ヌクレオチドのミスマッチは見られなかった。他の実施形態では、アラインメントは、特異的でなくてもよく、ミスマッチも許容され得る。

【0073】

600bp以下の挿入サイズを示したPE読み取り配列のみを解析用に回収した。これらの基準によると、これらの実験で解析した血漿DNA断片のサイズは、36bp~600bpの範囲にあった。それぞれシーケンシングしたDNA断片のサイズを、シーケンシングした断片の各末端における最も外側のヌクレオチドの座標から推定した。

【0074】

【表1】

| 症例の種類 | 症例数 | 妊娠期間(週) 中央値(範囲) | ペアエンド読み取り 配列の数(100万) 中央値(範囲) | 胎児DNA画分(%) 中央値(範囲) |
|-------|-----|--------------------|------------------------------------|-----------------------|
| 正倍数体 | 39 | 13.2 (11.3-15.1) | 4.7 (1.8-12.0) | 15.7 (5.9-25.7) |
| T21 | 18 | 13.0 (12.1-17.9) | 5.2 (2.5-8.9) | 13.8 (7.4-27.2) |
| T18 | 10 | 13.3 (12.1-14.2) | 4.9 (3.6-6.2) | 7.2 (4.8-16.7) |
| T13 | 13 | 12.4 (11.5-16.4) | 5.3 (2.7-7.7) | 7.5 (3.2-14.1) |
| All | 80 | 13.1 (11.3-17.9) | 4.9 (1.8-12.0) | 13.7 (3.2-27.2) |

10

20

30

40

50

【0075】

表1は、様々な異数性状態の試料についてのデータを示す。データには、症例数、妊娠期間の中央値と範囲、ペアエンド読み取り配列の数の中央値と範囲、胎児DNA画分の中央値と範囲が含まれる。

【0076】

母体血漿試料における胎児DNAの分画濃度は、先に記載のようにY染色体にアラインする配列の量から導き出した(Chiu RW et al. BMJ 2011; 342: c7401)。この技術は、校正方法の一例である。したがって、表1で測定された胎児DNA画分を校正データ点に使用して、新しい試料における胎児DNA画分を推定できる。表1のデータを収集するために使用される試料を、校正試料とすることができる。

10

【0077】

B. 標的シーケンシングを利用した第2胎児試料セット

表2は、試料情報および本発明の実施形態に従って標的を濃縮した母体血漿DNAを示す。血漿試料は、それぞれ1人の男児を宿している48人の妊婦から採取した。48人の妊婦のうち、21人は正倍数体の胎児を宿しており、17人はトリソミー21(T21)の胎児を宿しており、9人はトリソミー18(T18)の胎児を宿しており、1人はトリソミー13(T13)の胎児を宿していた。これらのデータは、以下の実施例と共に、標的技術を用いることができる実施形態を示すものである。血漿DNAのサイズ分布は、ペアエンド大量並列シーケンシングを用いて決定できる。他の実施形態では、血漿DNAのサイズ分布は、例えば、これらに限定されないが、リアルタイムPCR、電気泳動、および質量分析を利用して決定できる。

20

【0078】

高倍率(high-fold)シーケンシングで標的領域をカバーするために、一実施形態ではAgilent SureSelect Target Enrichment Systemを用いて、chr7(0.9Mb領域)、chr13(1.1Mb領域)、chr18(1.2Mb領域)およびchr21(1.3Mb領域)のDNA分子を捕捉するプローブを設計した。このプローブ設計において、chr7、chr13、chr18、およびchr21のダウン症候群に関する重要領域(21q22.1-q22.3)にあるエキソンをまず標的領域として選択した。chr13、chr18およびchr21は、chr7よりもエキソン領域が少ないので、chr13、chr18、およびchr21のダウン症候群に関する重要領域には追加の非エキソン領域を導入し、上記4つの染色体間の標的領域の全長のバランスをとった。選択した非エキソン領域は、120bpの長さで、特異的にマッピング可能で、GC含量が0.5に近く、そして標的染色体にわたって均一に分布していた。

30

【0079】

上記エキソンおよび非エキソン領域の全ての座標をプローブ設計用にAgilent eArray platformに供した。500ngの各母体血漿DNAライブラリーを、捕捉プローブと共に65℃で24時間インキュベートした。ハイブリダイゼーション後、標的DNA分子を溶離し、製造業者の説明書に従い12サイクルのPCRによって増幅した。標的を濃縮したライブラリーは、50bp x 2PEフォーマットを用いたGA IIX(Illumina)によりシーケンシングした。更に7サイクルのシーケンシングを行い、それぞれシーケンシングした血漿DNA分子上のインデックス配列をデコードした。50bpの読み取り配列は、Short Oligonucleotide Alignment Program 2(SOAP2)(soap.genomics.org.cn)を用いてリピーター領域をマスクしないヒト参照ゲノム(Hg18)(genome.ucsc.edu)にアラインした。個々のメンバーのPE読み取り配列は、フローセル上の同じクラスター位置でシーケンシングし、そして、ヒトゲノムの単一の位置に正しい方向で特異的にアラインした。2つのミスマッチが許容された。シーケンシングライブラリーの複雑さは、標的の濃縮後に大幅に減少した。

40

【0080】

50

600bp以下の挿入サイズを示したPE読み取り配列のみを解析用に回収した。これらの基準によると、この試験で解析した血漿DNA断片のサイズは、36bp～600bpの範囲にあった。それぞれシーケンシングしたDNA断片のサイズを、シーケンシングした断片の各末端における最も外側のヌクレオチドの座標から推定した。母体血漿試料における胎児DNAの分画濃度は、胎児特異的アレルを有する断片とそれぞれの母体と同じアレルを有する断片との比率より推定した。

【0081】

【表2】

| 症例の種類 | 症例数 | 妊娠期間(週) 中央値 (範囲) | ペアエンド読み取り 配列の数 (100万) 中央値 (範囲) | 胎児DNA画分 (%) 中央値 (範囲) |
|-------|-----|---------------------|--------------------------------------|-------------------------|
| 正倍数体 | 21 | 13.0 (12.0-13.3) | 2.2 (1.7-3.0) | 13.5 (8.4-22.0) |
| T21 | 17 | 13.6 (12.6-20.9) | 2.1 (1.5-2.7) | 15.4 (8.7-22.7) |
| T18 | 9 | 12.7 (11.9-13.7) | 1.9 (1.7-3.1) | 10.5 (7.2-16.3) |
| T13 | 1 | 13 | 1.6 | 9.2 |
| All | 48 | 13.1 (11.9-20.9) | 2.1 (1.5-3.1) | 13.4 (7.2-22.7) |

10

【0082】

表2は、様々な異数性状態の試料について標的をシーケンシングしたデータを示す。

【0083】

C. 胎児サンプルの電気泳動

血漿DNAのサイズ分布の解析は、大量並列シーケンシングの利用に加え、電気泳動により達成される。電気泳動では、断片が媒体を移動する時間を測定する。粒子のサイズが異なると、媒体を移動するのにかかる時間が異なる。したがって、一実施形態では、母体血漿DNAのシーケンシングライブラリーのマイクロ流体電気泳動により、母体血漿DNAのサイズ分布を決定できる。

【0084】

図4は、本発明の実施形態に従った、電気泳動を利用して得られる母体血漿DNAのサイズ分布を示すプロット400（エレクトロフェログラム）である。マイクロ流体電気泳動は、Agilent 2100 Bioanalyzerを用いて行った。2つの試料のシーケンシングライブラリーのエレクトロフェログラムをプロット400に示す。X軸は、DNAがセンサに到達するまでにかかる時間を示し、DNA断片のサイズに対応する。Y軸は、特定の時間にセンサを通過するDNA断片の蛍光単位（FU）を表す。

20

【0085】

DNA断片がセンサに到達するまでにかかる時間は、DNA断片のサイズと正相関する。バイオアナライザーは、自動的に、被験試料がかかった時間を既知の長さのDNA断片の混合物（すなわち、DNAラダー）がかかった時間と比較することにより、時間を断片のサイズに変換できる。続けて、DNAシーケンシングライブラリーを大量並列シーケンシングを利用してシーケンシングし、そしてY染色体配列の分画を用いてこれらの試料の分画胎児DNA濃度を決定した。

30

【0086】

プロット400において、実線410は、分画胎児DNA濃度が8.3%の試料UK92797を示し、破線420は、分画胎児DNA濃度が20.3%の試料UK94884を示す。UK92797試料と比較して、試料UK94884（分画胎児DNAが多い試料）は、200bp～267bpのDNAサイズに対応する63秒～73秒の電気泳動時間の間（領域A）に比較的多いDNA量を有し、292bpまでのDNAサイズに対応する73秒の電気泳動時間（領域B）以内に比較的低いDNA量を有していた。

40

50

【 0 0 8 7 】

製造業者のプロトコル従い、合計サイズが122bpのDNAアダプターおよびプライマーセットを、シーケンシングライブラリー構築のため血漿DNAに導入した。したがって、領域Aは、約78bp~145bpの血漿DNA断片に対応し、領域Bは、約170bpの血漿DNA断片に対応する。そのように導き出すことは、DNAライブラリーの構築のための異なるプロトコルに応用できる。例えば、Illumina シングルリードシーケンシングライブラリーの調製では合計サイズ92bpのアダプター/プライマーを導入するが、このサイズは、標準的なペアエンドシーケンシングライブラリーの調製では119bpになる。

【 0 0 8 8 】

別の実施形態では、血漿DNAは、当業者に知られている全ゲノム増幅システム、例えば、Rubicon Genomics PlasmaPlex WGA kit (www.rubicongenomics.com/products) により増幅できる。増幅産物は、その後パイオアナライザーによって解析できる。さらに別の実施形態では、増幅産物は、例えば、Caliper (www.caliperls.com/products/labchip-systems) の電気泳動システムにより解析できる。さらに別の実施形態では、血漿DNAのサイズ分布は、例えば、ナノポアベースのシーケンサー(例えば、Oxford Nanopore Technologies (www.nanoporetech.com)) より入手) またはHelico DNA シーケンサー(www.helicosbio.com) を用いて増幅せずに直接解析できる。

【 0 0 8 9 】

I V . サイズパラメータ

上述のように様々なパラメータが、生体試料におけるDNA断片のサイズプロファイルの統計的尺度となる。パラメータは、解析するDNA断片の全サイズのまたは一部のみを使用して決定してもよい。一実施形態では、パラメータは、短いDNA断片と長いDNA断片の相対量であり、ここで短いDNAおよび長いDNAは、特定のサイズまたはサイズ範囲に対応しうる。

【 0 0 9 0 】

母体血漿DNAの全体的なサイズ分布が分画胎児DNA濃度を反映させるのに使用できるかを調べるために、異なるパラメータを使用して短いDNAおよび長いDNAの相対量を定量し、これらのパラメータと分画胎児DNA濃度との間の相関を決定した。これらの調査の結果を以下のセクションで示す。短いDNAの相対量を反映するために、例示として我々が用いたパラメータは、以下を含む：

i . CF (サイズ 150) と表する150bp以下のDNA断片の割合。CFは累積度数(cumulative frequency)を指す。よって、CF (サイズ 150) は150bp以下の断片の累積度数を指す。

ii . (CF (サイズ 150) / サイズ(163~169)) と表する150bp以下のDNA断片と163bp~169bpのDNAとの量の比。

iii . (サイズ(140~146) / サイズ(163~169)) と表する140bp~146bpのDNA断片と163bp~169bpのDNAとの量の比。

iv . (サイズ(140~154) / サイズ(163~169)) と表する140bp~154bpのDNA断片と163bp~169bpのDNAとの量の比。

v . (サイズ(100~150) / サイズ(163~169)) と表する100bp~150bpのDNA断片と163bp~169bpのDNAとの量の比。

【 0 0 9 1 】

パラメータの他の例として、ヒストグラムの頻度数がある。一実施形態では、複数のパラメータを使用してよい。例えば、各パラメータの値が異なる割合を示すことがあるので、これらの割合の平均を決定してもよい。別の実施形態では、各パラメータが、多次元校正関数の異なる次元に対応し、ここで、新しい試料のパラメータの値は、対応する多次

10

20

30

40

50

元面上の座標に対応する。

【 0 0 9 2 】

V . サイズと分画濃度との相関

シーケンシングを利用した2つの試料セットを様々なサイズパラメータと分画濃度との相関を説明するのに使用する。リピート要素のサイズの解析も行う。また、電気泳動のデータも、サイズパラメータと分画濃度との相関を示す。

【 0 0 9 3 】

A . 第1試料セット

図5Aは、本発明の実施形態に従った、母体血漿における胎児DNAの割合が様々な試料について、150bp以下のDNA断片の割合を示すプロット500である。80個の母体血漿試料について分画胎児DNA濃度に対する150bp以下のDNAの割合をプロットした。正倍数体試料を黒丸で示す。トリソミー13(T13)試料を白三角で示す。トリソミー18(T18)試料を白菱形で示し、トリソミー21(T21)試料を白逆三角で示す。

10

【 0 0 9 4 】

全試料について分画胎児DNA濃度と150bp以下のDNAとの割合の間に正の相関があった(ピアソン相関係数=0.787)。サイズパラメータと分画胎児DNA濃度との間の正の相関は、胎児の染色体の状態が異なる試料間で一貫しているようだ。これらの結果は、サイズパラメータの解析が母体血漿試料における分画胎児DNA濃度を推定するために有用であることを示唆している。したがって、図5のデータ点を方法300の校正データ点として使用できる。その後、新たな試料についてのパラメータCF(サイズ150)が30であると判定された場合、胎児DNAの割合は約7%~16%の間であると推定できる。図5のデータ点も、導き出し示した生データ点に当てはまる校正関数を決定するために使用できる。

20

【 0 0 9 5 】

図5Bは、(CFサイズ(サイズ150)/サイズ(163~169))と表する150bp以下のDNA断片と163bp~169bpのDNAとの量のサイズ比を示すプロット550である。80個の母体血漿試料について分画胎児DNA濃度に対するCF(サイズ150)/サイズ(163~169)比をプロットした。全試料について分画胎児DNA濃度とCF(サイズ150)/サイズ(163~169)比の間に正の相関があった(ピアソン相関係数=0.815)。サイズパラメータと分画胎児DNA濃度との間の正の相関は、胎児の染色体の状態が異なる試料間で一貫している。

30

【 0 0 9 6 】

図6Aは、(サイズ(140~146)/サイズ(163~169))と表する140bp~146bpのDNA断片と163bp~169bpのDNAとの量のサイズ比を示すプロット600である。80個の母体血漿試料について分画胎児DNA濃度に対するサイズ(140~146)/サイズ(163~169)比をプロットした。全試料について分画胎児DNA濃度とサイズ(140~146)/サイズ(163~169)比の間に正の相関があった(ピアソン相関係数=0.808)。サイズパラメータと分画胎児DNA濃度との間の正の相関は、胎児の染色体の倍数性の状態が異なる試料間で一貫している。

40

【 0 0 9 7 】

図6Bは、(サイズ(140~154)/サイズ(163~169))と表する140bp~154bpのDNA断片と163bp~169bpのDNAとの量のサイズ比を示すプロット650である。80個の母体血漿試料について分画胎児DNA濃度に対するサイズ(140~154)/サイズ(163~169)比をプロットした。全試料について分画胎児DNA濃度とサイズ(140~154)/サイズ(163~169)比の間に正の相関があった(ピアソン相関係数=0.802)。サイズパラメータと分画胎児DNA濃度との間の正の相関は、胎児の染色体の倍数性の状態が異なる試料間で一貫しているようだ。

【 0 0 9 8 】

50

図7は、(サイズ(100~150)/サイズ(163~169))と表する100bp~150bpのDNA断片と163bp~169bpのDNAとの量のサイズ比を示すプロット700である。80個の母体血漿試料について分画胎児DNA濃度に対するサイズ(100~150)/サイズ(163~169)比をプロットした。全試料について分画胎児DNA濃度とサイズ(100~150)/サイズ(163~169)比の間に正の相関があった(ピアソン相関係数=0.831)。サイズパラメータと分画胎児DNA濃度との間の正の相関は、胎児の染色体の状態が異なる試料間で一貫している。

【0099】

B. 第2試料セット

図8は、本発明の実施形態に従った、母体血漿における胎児DNAの割合が様々な試料について150bp以下のDNA断片の割合を示すプロット800である。標的を濃縮後に48個の超並列ペアエンドシーケンシングした母体血漿試料について分画胎児DNA濃度に対する150bp以下のDNAの割合をプロットした。正倍数体試料を黒丸で示す。トリソミー13(T13)試料を白三角で示す。トリソミー18(T18)試料を白菱形で示し、トリソミー21(T21)試料を白逆三角で示す。全試料について分画胎児DNA濃度と150bp以下のDNAとの割合の間に正の相関があった(ピアソン相関係数=0.816)。サイズパラメータと分画胎児DNA濃度との間の正の相関は、胎児の染色体の状態が異なる試料間で一貫している。これらの結果は、サイズパラメータの解析が母体血漿試料における分画胎児DNA濃度を推定するために有用であることを示唆している。

【0100】

図9Aは、(CFサイズ(サイズ150)/サイズ(163~169))と表する150bp以下のDNA断片と163bp~169bpのDNAとの量のサイズ比を示すプロット900である。48個の母体血漿試料について分画胎児DNA濃度に対するCF(サイズ150)/サイズ(163~169)比をプロットした。全試料について分画胎児DNA濃度とCF(サイズ150)/サイズ(163~169)比の間に正の相関があった(ピアソン相関係数=0.776)。サイズパラメータと分画胎児DNA濃度との間の正の相関は、胎児の染色体の状態が異なる試料間で一貫している。

【0101】

図9Bは、(サイズ(140~146)/サイズ(163~169))と表する140bp~146bpのDNA断片と163bp~169bpのDNAとの量のサイズ比を示すプロット950である。48個の母体血漿試料について分画胎児DNA濃度に対するサイズ(140~146)/サイズ(163~169)比をプロットした。全試料について分画胎児DNA濃度とサイズ(140~146)/サイズ(163~169)比の間に正の相関があった(ピアソン相関係数=0.790)。サイズパラメータと分画胎児DNA濃度との間の正の相関は、胎児の染色体の状態が異なる試料間で一貫している。

【0102】

図10Aは、(サイズ(140~154)/サイズ(163~169))と表する140bp~154bpのDNA断片と163bp~169bpのDNAとの量のサイズ比を示すプロット1000である。48個の母体血漿試料について分画胎児DNA濃度に対するサイズ(140~154)/サイズ(163~169)比をプロットした。全試料について分画胎児DNA濃度とサイズ(140~154)/サイズ(163~169)比の間に正の相関があった(ピアソン相関係数=0.793)。サイズパラメータと分画胎児DNA濃度との間の正の相関は、胎児の染色体の状態が異なる試料間で一貫している。

【0103】

図10Bは、(サイズ(100~150)/サイズ(163~169))と表する100bp~150bpのDNA断片と163bp~169bpのDNAとの量のサイズ比を示すプロット1005である。48個の母体血漿試料について分画胎児DNA濃度に対するサイズ(100~150)/サイズ(163~169)比をプロットした。全試料について分画胎児DNA濃度とサイズ(100~150)/サイズ(163~169)比の間

に正の相関があった（ピアソン相関係数 = 0.798）。サイズパラメータと分画胎児 DNA 濃度との間の正の相関は、胎児の染色体の倍数性の状態が異なる試料間で一貫している。

【0104】

C. 反復要素

上述の通り我々は、母体血漿における全てのマッピング可能な DNA 断片のサイズが分画胎児 DNA 濃度と相関することを実証した。このセクションでは、ゲノム中の反復要素のサイズ解析も、血漿における分画胎児 DNA 濃度の推定に使用できるか否かを調べる。本実施例で、我々は、ゲノムの A1u 反復にマッピングした DNA 断片のサイズ分布を解析した。

10

【0105】

図 11 は、本発明の実施形態に従った、反復要素のサイズについて胎児 DNA の割合に対しプロットしたサイズ比を示すプロットである。この例では、100 bp ~ 150 bp の DNA 断片と 163 bp ~ 169 bp の DNA との量の比（サイズ（100 ~ 150）/ サイズ（163 ~ 169））を、胎児 DNA の割合に対するサイズ分布の変化を反映させるのに用いる。サイズ比と分画胎児 DNA 濃度の間に正の相関があった（ピアソン相関係数 = 0.829）。この結果は、反復要素のサイズ分析もまた、母体試料の胎児 DNA 分画濃度を決定するのに使用できることを示唆している。

【0106】

大量並列シーケンシングの使用に加えて、他の方法、例えば、PCR、リアルタイム PCR、電気泳動、および質量分析を利用して母体血漿における反復要素（例えば、A1u 反復）のサイズ分布を決定できる。一実施形態では、母体血漿試料における DNA は、リンカーにライゲーションできる。その後、PCR を、A1u 配列に特異的な 1 のプライマーおよびリンカーに特異的な他のプライマーを用いて行うことができる。PCR に続いて、PCR 産物のサイズを、例えば、電気泳動、質量分析、または大量並列シーケンシングによって解析できる。これにより、母体血漿の A1u 反復由来の配列のサイズを読み取ることができるだろう。この戦略は、他の標的配列または配列ファミリーにも使用できる。更に、PCR に続き、別の A1u 特異的プライマーを用いるネステッド（nested）PCR を行ってもよく、ここでは同じリンカー特異的プライマーまたはリンカー中のネステッドプライマーのいずれかと組み合わせる。かかるネステッド PCR は、対象とする配列（この場合、A1u 配列）を増幅させる特異性を増加させるという利点を有する。

20

30

【0107】

反復要素を利用する利点の 1 つは、反復要素のコピー数は比較的多いので、解析が容易なことである。例えば、増幅のサイクルが少なくてもよい場合がある。また、コピー数が多いと、解析精度が潜在的に高くなる。考えられる欠点は、個体間でコピー数が異なる反復要素のクラスもあるということである。

【0108】

D. 電気泳動

図 12A は、本発明の実施形態に従った、サイズ比を決定するために使用できるエレクトロフェログラム 1200 である。解析した全ての DNA ライブラリーについて、約 292 bp に鋭いピークがあり、これに続き 300 bp ~ 400 bp の範囲の二次ピークがあった。あるサイズ範囲の曲線下面積は、その領域由来の DNA 断片の相対量を表すことができるので、領域 A（200 bp ~ 267 bp）の面積と領域 B（290 bp ~ 294 bp）の面積の比を用いて、短い断片 DNA および長い DNA 断片の相対量を定量化する。まず、手動で蛍光単位（FU）のベースラインを 0 に調整し、次に、選択領域の面積を求めた。

40

【0109】

図 12B は、本発明の実施形態に従った、母体血漿における胎児 DNA の割合が様々な試料について 200 bp ~ 267 bp の DNA 断片と 290 bp ~ 294 bp の DNA との量のサイズ比（すなわち、エレクトロフェログラムに示された領域 A と B の面積比）を

50

示すプロット 1 2 5 0 である。2 9 2 b p のピークが F U 値 6 . 1 と低い T 1 3 のケースが 1 つあったが、それ以外のケースは全て 2 0 以上の F U 値を示した。この低い F U 値は面積測定が不正確になるだろうので、解析から除外した。領域 A と B の面積比を、他の 7 9 個の母体血漿試料全てについての分画胎児 D N A 濃度に対してプロットする。これらの試料について分画胎児 D N A 濃度と A と B の面積比との間に正の相関があった（ピアソン相関係数 = 0 . 7 2 3 ）。

【 0 1 1 0 】

V I . 校正データ点の決定

上述したように、校正データ点は、様々な方法で定義できる。さらに、校正データ点は、様々な方法で得ることができる。例えば、校正データ点は、単純に、あるパラメータの一連の校正值として対応する分画濃度とともにメモリから読み出すものであってもよい。また、校正データ点を定義する校正関数をメモリから読み出すこともできる（例えば、所定の関数形式を有する線形または非線形関数）。いくつかの実施形態では、校正データ点は、校正試料から測定されたデータから計算できる。

【 0 1 1 1 】

A . 方法

図 1 3 は、本発明の実施形態に従った、校正試料から得た測定値より校正データ点を決定するための方法 1 3 0 0 のフローチャートである。校正試料には、臨床的に関連する D N A およびその他の D N A が含まれる。

【 0 1 1 2 】

ブロック 1 3 1 0 では、複数の校正試料を受けとる。校正試料は、本明細書に記載のように得られる。各試料は、別個の実験、またはある分子についてどの試料由来なのかを同定するいくつかの同定手段（例えば、バーコードで D N A 断片をタグ付けする）により別個に解析できる。例えば、校正試料を、例えば、シーケンシング装置などの D N A 断片のサイズを決定するために用いることができる測定データ（例えば、読み取り配列）を出力する装置で受け取ってもよく、または電気泳動装置で受け取ってもよい。

【 0 1 1 3 】

ブロック 1 3 2 0 では、臨床的に関連する D N A の分画濃度を、複数の校正試料のそれぞれについて測定する。胎児 D N A 濃度を測定する様々な実施形態では、父系遺伝配列又は胎児特異的エピジェネティックマーカーを使用してもよい。例えば、父系遺伝アレルは、妊婦のゲノムには存在しないので、母体血漿における分画胎児 D N A 濃度に比例する割合で検出できる。胎児特異的エピジェネティックマーカーには、母体血漿における胎児または胎盤特異的 D N A メチル化パターンを示す D N A 配列を含みうる。

【 0 1 1 4 】

ブロック 1 3 3 0 では、それぞれの校正試料由来の D N A 断片の量を、様々なサイズについて測定する。サイズは、本明細書に記載のように測定できる。校正試料のサイズプロファイルに関するデータを取得するために、サイズをカウントやプロットしてもよく、ヒストグラムを作成するのに使用してもよく、または他の分類手段に供してもよい。

【 0 1 1 5 】

ブロック 1 3 4 0 では、あるパラメータについての校正值を複数のサイズの D N A 断片の量に基づいて計算する。校正值は、各校正試料について計算できる。一実施形態では、各校正值について同じパラメータを使用する。しかし、本明細書に記載の複数のパラメータを使用できる実施形態もある。例えば、1 5 0 塩基未満の D N A 断片の累積分画をパラメータとして使用でき、異なる分画濃度を有する試料は異なる校正值を有するであろう。各試料について校正值および測定された分画濃度を含む当該試料についての校正データ点を決定してもよい。これらの校正データ点は、方法 3 0 0 で使用でき、あるいは最終校正データ点（例えば、関数当てはめにより定義される）を決定するのに使用できる。

【 0 1 1 6 】

ブロック 1 3 5 0 では、複数の分画濃度にわたり校正值を近似する関数を決定する。例えば、線形関数は、分画濃度の関数として校正值に当てはまるだろう。線形関数は、方法

10

20

30

40

50

300で使用される校正データ点を定義できる。

【0117】

いくつかの実施形態では、複数のパラメータについての校正值を、各試料について計算できる。ある試料についての校正值は、分画濃度とともにデータ点を提供できる多次元座標（各次元は、各パラメータについてのものである）を定義できる。従って、一実施態様では、多次元関数は、全ての多次元データ点に当てはめることができる。したがって、異なる値のパラメータを単一の校正関数に効果的に代入して分画濃度を算出できる多次元検量線を使用できる。そして、単一の校正関数を、校正試料から得られた全てのデータ点の関数当てはめにより作成できる。

【0118】

B．腫瘍DNA濃度の測定

上述のように、実施形態はまた、生体試料における腫瘍DNAの濃度にも適用できる。腫瘍DNAの分画濃度の決定に関する例を次に示す。

【0119】

我々は、肝細胞癌（HCC）に罹患している2人の患者から外科的腫瘍切除の前後の血漿試料を採取した。サイズ解析は、ペアエンド（PE）大量並列シーケンシングを用いて行った。母体血漿DNAのシーケンシングライブラリーは、従前に記載のように構築した（Lo YM et al. Sci Transl Med 2010; 2: 61ra91）。全ライブリは、HiSeq 2000（Illumina）50bp×2PEフォーマットを用いてシーケンシングした。50bpの読み取り配列は、Short Oligonucleotide Alignment Program 2（SOAP2）（soap.genomics.org.cn）を用いて、リピート領域をマスクしないヒト参照ゲノム（Hg18）（http://genome.ucsc.edu）にアラインした。それぞれシーケンシングした断片のサイズを、アラインした断片の各末端における最も外側のヌクレオチドの座標から推定した。

【0120】

我々は、Affymetrix SNP6.0 microarray systemを使用して、HCC患者の血液細胞および腫瘍試料から抽出したDNAの遺伝子型を決定した。それぞれのケースで、腫瘍組織においてヘテロ接合性の消失（LOH）を示す領域を、Affymetrix Genotyping Console v4.0を用いてSNP座位の異なるアレルの強度に基づき同定した。腫瘍由来のDNAの分画濃度（F）は、式： $F = (A - B) / A \times 100\%$ （式中、Aは、LOH領域においてヘテロ接合SNPが欠失していないアレルを有する読み取り配列の数であり、Bは、LOH領域においてヘテロ接合SNPが欠失しているアレルを有する読み取り配列の数である）を用い、LOH領域で欠失および欠失していないアレルを有する配列の量の差から推定した。表3は、その結果を示す。

【0121】

【表3】

| ケース番号 | 採取時期 | 読み取り配列の数 | 血漿における腫瘍DNAの分画濃度（%） |
|-------|-------|----------|---------------------|
| 1 | 腫瘍切除前 | 448M | 51.60 |
| | 腫瘍切除後 | 486M | 0.90 |
| 2 | 腫瘍切除前 | 479M | 5.60 |
| | 腫瘍切除後 | 542M | 0.90 |

【0122】

表3は、血漿試料におけるシーケンシング情報および腫瘍DNAの分画濃度の測定値を示す。

【 0 1 2 3 】

別の実施形態では、重複を示す座位を使用することができる。例えば、腫瘍では、2つのホモ接合染色体のうちの1つについて1コピー分増加しアレルが重複しているようにあることがある。よって、1つまたは複数のヘテロ接合の座位において非重複アレル（例えば、SNP）を有する読み取り配列の第1量Aおよびヘテロ接合の座位において重複アレルを有する読み取り配列の第2量Bを決定できる。臨床的に関連するDNAの分画濃度Fは、第1量と第2量の比として $(B - A) / A$ 比を用いて計算できる。

【 0 1 2 4 】

別の実施形態では、1つまたは複数のホモ接合の座位を使用してもよい。例えば、患者がホモ接合であり、1ヌクレオチドの変異が腫瘍組織内に存在する場合、1つまたは複数の座位を同定できる。その後、1つまたは複数のホモ接合の座位において野生型アレルを有する読み取り配列の第1量Aを決定できる。そして、1つまたは複数のホモ接合の座位において変異型アレルを有する読み取り配列の第2量Bを決定できる。臨床的に関連するDNAの分画濃度Fは、第1量と第2量の比として $2B / (A + B)$ 比を用いて計算できる。

【 0 1 2 5 】

C. データ点に対する関数当てはめ (Functional Fit) の例

校正試料から決定されるパラメータ値に対し関数当てはめを実施する例をここで説明する。それぞれ1人の男児を宿している80人の妊婦から採取した血漿試料を解析した。80人の妊婦のうち、39人は正倍数体の胎児を宿しており、13人はトリソミー13 (T13) の胎児を宿しており、10人はトリソミー18 (T18) の胎児を宿しており、18人はトリソミー21 (T21) の胎児を宿していた。これらの妊婦の妊娠期間の中間値は13週と1日であった。DNAを血漿試料から抽出し、先に記載したIllumina HiSeq 2000プラットフォーム (Zheng YW et al. Clin Chem. 2012; 58: 549 - 58) を用いてシーケンシングした。ただし、シーケンシングは8-plexフォーマットで行った。各DNA分子について、両端からそれぞれから50ヌクレオチドをシーケンシングし、参照ゲノム (hg18) にアラインした。

【 0 1 2 6 】

次に、それぞれシーケンシングした分子のサイズを、両端の最も外側のヌクレオチドの座標から導き出した。各試料について、中央値が1110万の断片をシーケンシングし、参照ゲノムに特異的にアラインした。サイズが100bp ~ 150bpのDNA分子の割合をサイズが163bp ~ 169bpのDNA分子の割合で除した比を計算し、この比をサイズ比と呼ぶ。80人の妊婦は、全員男児を宿していたので、Y染色体に特異的にアラインした読み取り配列の割合を、各血漿DNA試料における胎児DNAの分画濃度を決定するのに用いた。

【 0 1 2 7 】

試料は、ランダムに2セット、すなわちトレーニングセットとバリデーションセットに分けた。線形回帰を使用し、トレーニングセット内の試料に基づいて、分画胎児DNA濃度とサイズ比との関係を確認した。その後、サイズ比から、線形回帰式を用いてバリデーションセットについての分画胎児DNA濃度を導き出した。バリデーションは、次のセクションで説明する。

【 0 1 2 8 】

図14Aは、本発明の実施形態に従った、トレーニングセットについての胎児DNAの分画濃度に対するサイズ比のプロット1400である。上述のように、サイズ比は、サイズが100bp ~ 150bpのDNA分子の割合をサイズが163bp ~ 169bpのDNA分子の割合で除することで計算する。データ点1405で示すように、サイズ比を、胎児のDNAの分画濃度に対してプロットする。白丸は、正倍数体のケースを表す。黒塗りのマークは、異数性のケースを表す (T13は黒四角、T18は黒丸、T21は黒三角)。線形回帰線1410は、データ点の関数当てはめによる結果である。関数当てはめは

、任意の適切な技術、例えば、最小二乗法により行うことができる。この線 1 4 1 0 は、トレーニングセット以外の他の試料について測定されたパラメータの値を推定するのに使用できる。線 1 4 1 0 の各部分を校正データ点と考えることができる。

【 0 1 2 9 】

V I I . 校正データ点との比較

上述したように、校正データ点を用いて、臨床的に関連する D N A の分画濃度を決定できる。例えば、図 1 4 A の生データ点 1 4 0 5 を用いて、ある校正値に対する分画 D N A 濃度の範囲を設けることができる（図 1 4 A でラベルしたサイズ比）。この範囲は、分画濃度が閾値量より多いか否かを決定するのに使用できる。範囲の代わりに、特定のサイズ比における分画濃度の平均値を使用してもよい。例えば、新しい試料についてのサイズ比として 1 . 3 という測定値に対応する分画濃度は、1 . 3 における 2 つのデータ点から計算した平均濃度として決定できる。一実施形態では、関数当てはめ（例えば、線 1 4 1 0 ）を使用してもよい。

10

【 0 1 3 0 】

図 1 4 B は、本発明の実施形態に従った、胎児特異的配列を用いて測定した分画濃度に対する図 1 4 A の線形関数 1 4 1 0 から導き出した（推定した）分画濃度のプロット 1 4 5 0 である。トレーニングセットのデータに基づいて決定された回帰式（すなわち、線 1 4 1 0 ）を用いて、バリデーション試料について決定されるサイズ比を、バリデーションセットの試料についての胎児 D N A の分画濃度を導き出すのに使用した。測定された分画濃度は、血漿 D N A 試料における Y 染色体配列の割合（すなわち、Y 染色体にアラインする読み取り配列の割合）に対応する。

20

【 0 1 3 1 】

線 1 4 6 0 は、2 セットの値の間で完全な相関関係を表す。データ点 1 4 5 5 のずれは、推定がどの程度正確であることを示し、点が線 1 4 6 0 上にあると点は完全に正確であることを示す。本明細書で述べたように、求める試験は、単に、生体試料において臨床的に関連する D N A が十分な割合で存在するか否かを決定するためのものである。推定は完全に正確である必要はない。白丸は、正倍数体のケースを表す。黒塗りのマークは、異数性のケースを表す（T 1 3 は黒四角、T 1 8 は黒丸、T 2 1 は黒三角）。サイズ比から導き出した胎児 D N A 濃度と Y 染色体配列の割合から測定した分画胎児 D N A 濃度間の差の中央値は 2 . 1 % であった。試料の 9 0 % で、その差は 4 . 9 % 未満であった。

30

【 0 1 3 2 】

異なる倍数性状態を有する試料を、キャリブレーションセットとバリデーションセットの両方で使用した。図 1 4 A に示すように、サイズ比と分画胎児 D N A 濃度間の関係は倍数性状態が異なる試料間で一貫していた。その結果、分画胎児 D N A 濃度は、図 1 4 B に示すように、試料の倍数性状態を事前に知ることなく、試料のサイズ比から導き出すことができる。1 つの検量線を、異なる倍数性状態を有する試料について使用した。つまり、我々は、分画胎児 D N A 濃度を決定するための実施形態を適用する前に、試料の倍数性状態を知る必要がない。

【 0 1 3 3 】

V I I I . 癌

40

本明細書に記載のように、実施形態は、生体試料における腫瘍 D N A の分画濃度を推定するのに使用できる。胎児の例と同様に、例えば、サイズパラメータの値と測定した分画濃度との間の相関関係を示すデータ点に関数（例えば、線形関数）を当てはめることにより、校正試料を用いて相関データ点を決定できる。

【 0 1 3 4 】

A . サイズと腫瘍 D N A 濃度の相関

図 1 5 A は、本発明の実施形態に従った、腫瘍切除前後の 2 人の H C C 患者の血漿における腫瘍 D N A の割合が様々な試料について 1 5 0 b p 以下の D N A 断片の割合を示すプロット 1 5 0 0 である。2 人の H C C 患者について、腫瘍切除の前（黒丸）および後（白丸）の分画腫瘍 D N A 濃度に対する 1 5 0 b p 以下の D N A の割合をプロットした。2 つ

50

の白丸の位置は互いに近接している（ほぼ重なっている）。これらの結果は、サイズパラメータの解析がHCC患者の血漿試料における分画腫瘍DNA濃度を推定するのに有用であることを示唆している。腫瘍切除後の150bp以下の分画腫瘍DNA濃度およびDNA断片の割合の両方で減少がある。黒丸1505は、腫瘍DNAの割合がかなり低い試料に対応し、これは、腫瘍のサイズが小さいことに関連する。言い換えれば、大きな腫瘍を有する患者は、短いDNAの割合が高く、これは、小さい腫瘍を有する患者と比較してCF（150bp）の割合が高いことで反映される。

【0135】

図15Bは、腫瘍切除前後の2人のHCC患者について、（CFサイズ（サイズ150）/サイズ（163～169））と表する150bp以下のDNA断片と163bp～169bpのDNAとの量のサイズ比を示すプロット1550である。CF（サイズ150）/サイズ（163～169）比を2人のHCC患者について、腫瘍切除の前（黒丸）および後（白丸）の分画腫瘍DNA濃度に対しプロットした。2つの白丸の位置は互いに近接している。腫瘍切除後は、分画腫瘍DNA濃度とサイズ比の両方の減少があった。

10

【0136】

図16Aは、腫瘍切除前後の2人のHCC患者について、（サイズ（140～146）/サイズ（163～169））と表する140bp～146bpのDNA断片と163bp～169bpのDNAとの量のサイズ比を示すプロット1600である。サイズ（140～146）/サイズ（163～169）比を2人のHCC患者について、腫瘍切除の前（黒丸）および後（白丸）の分画腫瘍DNA濃度に対しプロットした。腫瘍切除後は、分画腫瘍DNA濃度とサイズ比の両方の減少があった。

20

【0137】

図16Bは、腫瘍切除前後の2人のHCC患者について、（サイズ（140～154）/サイズ（163～169））と表する140bp～154bpのDNA断片と163bp～169bpのDNAとの量のサイズ比を示すプロット1650である。サイズ（140～154）/サイズ（163～169）比を2人のHCC患者について、腫瘍切除の前（黒丸）および後（白丸）の分画腫瘍DNA濃度に対しプロットした。腫瘍切除後は、分画腫瘍DNA濃度とサイズ比の両方の減少があった。

【0138】

図17は、腫瘍切除前後の2人のHCC患者について、（サイズ（100～150）/サイズ（163～169））と表する100bp～150bpのDNA断片と163bp～169bpのDNAとの量のサイズ比を示すプロット1700である。サイズ（100～150）/サイズ（163～169）比を2人のHCC患者について、腫瘍切除の前（黒丸）および後（白丸）の分画腫瘍DNA濃度に対しプロットした。腫瘍切除後は、分画腫瘍DNA濃度とサイズ比の両方の減少があった。

30

【0139】

B．治療によるサイズの減少

図18Aは、腫瘍切除前後のHCC患者について、150bp以下のDNA断片の割合を示すプロット1800である。同じ癌患者由来の一对の試料は、破線でつながっている同一の記号で表す。一般的に、腫瘍切除後の癌患者の血漿DNAにおける150bp以下のDNAの割合の減少がある。

40

【0140】

治療前と治療後についての割合の値の差分は、腫瘍の存在とサイズパラメータの値との相関関係を示している。治療前と治療後についての割合の値の差分を利用して、例えば、その割合が閾値を下回ると成功であることを示す場合、閾値に対する割合を比較することにより、治療の成功度を決定できる。別の例では、治療前と治療後の間の差分を閾値と比較できる。

【0141】

割合（または他のサイズパラメータの値）を用いて腫瘍の発生を検出できる。例えば、サイズパラメータのベースライン値を決定できる。そして、その後、サイズパラメータの

50

値を再度測定できる。サイズパラメータの値が有意な変化を示した場合、その患者は、腫瘍を有しているリスクが高いとすることができる。割合があまり変わらない(すなわち、治療後の値が同じ)図18Aが示すように、サイズパラメータの値が個人間であまり変わらない場合、同じベースライン値を他の患者に適用できる。したがって、ベースライン値を、患者ごとに取る必要がない。

【0142】

図18Bは、腫瘍切除前後のHCC患者について、(CFサイズ(サイズ150)/サイズ(163~169))と表する150bp以下のDNA断片と163bp~169bpのDNAとの量のサイズ比を示すプロット1850である。同じ癌患者由来の一对の試料は、破線でつながっている同一の記号で表す。腫瘍切除後は、両者のケースともサイズ比の減少があった。

10

【0143】

図19Aは、腫瘍切除前後のHCC患者について、(サイズ(140~146)/サイズ(163~169))と表する140bp~146bpのDNA断片と163bp~169bpのDNAとの量のサイズ比を示すプロット1900である。同じ癌患者由来の一对の試料は、破線でつながっている同一の記号で表す。腫瘍切除後は、両者のケースともサイズ比の減少があった。

【0144】

図19Bは、腫瘍切除前後のHCC患者について、(サイズ(140~154)/サイズ(163~169))と表する140bp~154bpのDNA断片と163bp~169bpのDNAとの量のサイズ比を示すプロット1950である。同じ癌患者由来の一对の試料は、破線でつながっている同一の記号で表す。腫瘍切除後は、両者のケースともサイズ比の減少があった。

20

【0145】

図20は、腫瘍切除前後のHCC患者について、(サイズ(100~150)/サイズ(163~169))と表する100bp~150bpのDNA断片と163bp~169bpのDNAとの量のサイズ比を示すプロット2000である。同じ癌患者由来の一对の試料は、破線でつながっている同一の記号で表す。腫瘍切除後は、両者のケースともサイズ比の減少があった。

【0146】

30

C. 方法

図21は、本発明の実施形態に従った、生物の生体試料を解析して癌のレベルの区分を決定するための方法2100を示すフローチャートである。方法2100は、生物(例えば、ヒト)の生体試料を解析できる。生物試料は、正常細胞由来のDNAと、癌に関連しているおそれがある細胞由来のDNAを含む。少なくともDNAの一部は、生体試料における無細胞DNAである。方法300および1300の態様を方法2100の実施形態に適用できる。

【0147】

ブロック2110では、様々なサイズに対応するDNA断片の量を測定する。方法300に記載のように、複数のサイズのうちの各サイズについて、そのサイズに対応する生体試料由来の複数のDNA断片の量を測定することができる。複数のDNA断片を、ランダムに選択またはゲノムの1つまたは複数の所定の領域から優先的に選択できる。例えば、上述したように、標的の濃縮を行ってもよく、あるいはゲノムの特定の領域由来の読み取り配列を選択してもよい。

40

【0148】

ブロック2120では、第1パラメータの第1値を複数のサイズのDNA断片の量に基づいて計算する。一態様では、第1パラメータは、生体試料におけるDNA断片のサイズプロファイルの統計的尺度(例えば、ヒストグラム)を提供する。このパラメータは、複数のDNA断片のサイズから決定されるため、サイズパラメータと呼ぶことができる。パラメータの例を本明細書に記載する。本明細書に記載のように複数のパラメータを用いて

50

もよい。

【0149】

ブロック2130では、第1値を基準値と比較する。基準値の例として、正常値および正常値から特定の距離（例えば、標準偏差の単位内）にあるカットオフ値を含む。基準値は、（例えば、健康であることがわかっている）同じ生物由来の異なる試料から測定してもよい。よって、基準値は、生物が癌を有しないと推測される試料から決定される第1パラメータの値に対応し得る。一実施形態では、生体試料は、治療後の生物から得られ、基準値は、治療前に採取した試料から決定された第1パラメータの値に対応する（例えば、上記に例示されている）。基準値は、他の健康な生物の試料から決定してもよい。

【0150】

ブロック2140では、生物の癌のレベルの区分を、比較に基づいて決定する。様々な実施形態では、区分は、数値、文字、または他の任意の指標であり得る。区分することで、早期に癌に関する確率またはその他のスコアに関し *y e s* または *n o* という2値の結果を提供することができる。確率およびその他のスコアは、絶対値であってもよく、あるいはその生物の従前の区分に対する相対値であってもよい。一実施形態では、区分は、生物が癌を有していない、または癌のレベルが低下したということである。別の実施形態では、区分は、生物が癌を有している、または癌のレベルが増加したということである。

【0151】

本明細書に記載のように、癌のレベルには、癌の存在、癌の段階、または腫瘍のサイズを含み得る。例えば、第1値を超えたか否か（例えば、第1パラメータを如何に定義するかによるが、その値よりも大きいまたは小さいということ）を利用して、癌が存在するかどうか、または少なくともその可能性（例えば、パーセンテージ尤度）を判断できる。さらに、閾値を超える程度が可能性の増大を示すようにして、複数の閾値を利用してもよい。さらに、上記の程度は、異なる癌のレベル、例えば、腫瘍の数または大きさに対応してもよい。したがって、実施形態は、生物の癌のレベルの診断、段階付け、予知、または進行状況のモニターが可能である。

【0152】

D．特定の領域のサイズ分布の決定

他の実施形態と同様に、DNA断片の第1セットが、生物のゲノムにおける1つまたは複数の所定領域に対応しうる。よって、サイズ解析は、選択領域、例えば、特定の染色体、染色体の腕、または同じ長さ、例えば1Mbの複数領域（ピン）、についても実施できる。例えば、対象の種類で一般的に変化する領域に集中できる。図22の表2200は、様々な種類の癌で見られる一般的な染色体異常のいくつかを示す。増加は、特定のセグメント内の1つまたは複数のコピー数の追加を伴う染色体の増幅を指し、欠失は、特定のセグメント内のホモ接合染色体の片方または両方の欠失を指す。

【0153】

一実施形態では、DNA断片の追加のセットを、生体試料から同定してもよい。DNA断片の各セットは、表2200で特定した領域などの異なる所定領域に対応しうる。癌に関連しない領域もまた、例えば、基準値を決定するために使用できる。本明細書に記載のように、様々なサイズに対応するDNA断片の量を決定でき、DNA断片の追加のセットそれぞれについてのパラメータのサイズ値を決定できる。したがって、DNA断片のセットとゲノム領域との間に1対1の対応が存在するようなゲノム領域についてそれぞれの異なるサイズ値を決定できる。

【0154】

それぞれのサイズ値を、それぞれの基準値と比較できる。対応するサイズ値がそれぞれの基準値と統計的に差異がある所定領域を同定できる。基準値が正常値である場合、統計的な差異は、カットオフのサイズ値と比較することによって決定できる（例えば、推定または測定した統計的分布に基づくカットオフ値が特定の数の正常値からの標準偏差である場合）。それぞれの基準値は、異なる領域に対し同一または異なってもよい。例えば、異なる領域は、サイズごとに異なる正常値を有することがある。

10

20

30

40

50

【 0 1 5 5 】

一実施形態では、基準値と統計学的に差異がある領域の数を利用して、区分を決定してもよい。従って、対応するサイズの値がそれぞれの基準値と統計学的に差異がある所定領域を同定する数を決定できる。その数を、領域の閾値数と比較して、生物における癌のレベルの区分を決定できる。閾値数は、正常試料内および癌試料内の差に基づいて決定できる。

【 0 1 5 6 】

表 2 2 0 0 で示すように、異なる癌はゲノムの異なる部分に関連する。したがって、どの領域が統計学的に差異があるかということを利用して、可能性のある 1 つまたは複数の癌の種類が同定された領域に関連する場合、その可能性のある癌の種類を決定できる。例えば、染色体セグメント 7 p 由来の DNA 断片のサイズ値が正常値よりも有意に低いと分かった場合（例えば、カットオフ値の決定による）、その区分が癌の存在を示すなら結腸直腸癌が可能性のある癌として同定できる。染色体セグメント 7 p のサイズ値を区分を決定するただ一つの指標として用いてもよいし、あるいは複数の領域を用いてもよいことに留意されたい。一実施形態では、区分が全体として癌であることを示す場合にのみ、染色体セグメント 7 p のサイズ値を用いて結腸直腸癌を可能性のある癌として同定する。

【 0 1 5 7 】

I X . コンピュータシステム

本明細書に記載のいずれのコンピュータシステムに関しても、任意の適切な数のサブシステムを利用することが可能である。このようなサブシステムの例を、図 2 3 のコンピュータ装置 2 3 0 0 に示す。いくつかの実施形態では、コンピュータシステムは、ただ 1 つのコンピュータ装置を含み、この場合、サブシステムがコンピュータ装置の構成要素でありうる。他の実施形態では、コンピュータシステムは、複数のコンピュータ装置を含み、それぞれが内部構成要素を含むサブシステムでありうる。

【 0 1 5 8 】

図 2 3 に示すサブシステムがシステムバス 2 3 7 5 を介して相互接続される。例として、プリンター 2 3 7 4、キーボード 2 3 7 8、固定ディスク 2 3 7 9、ディスプレイアダプター 2 3 8 2 に接続されたモニター 2 3 7 6、等の追加のサブシステム、が示されている。I / O 制御装置 2 3 7 1 に接続された周辺装置および入 / 出力（I / O）装置を、当技術分野で既知の任意の数のシリアルポート 2 3 7 7、等の手段によりコンピュータシステムに接続できる。例えば、シリアルポート 2 3 7 7 または外部インターフェイス 2 3 8 1（例えば、イーサネット（登録商標）、Wi-Fi 等）を使って、コンピュータシステム 2 3 0 0 をインターネット等の広域ネットワーク、マウス入力装置、またはスキャナーに接続できる。システムバス 2 3 7 5 を介した相互接続は、中央処理装置 2 3 7 3 のそれぞれのサブシステムとの通信、およびシステムメモリー 2 3 7 2 または固定ディスク 2 3 7 9 からの命令の実行の制御、ならびにサブシステム間の情報交換の制御を可能とする。システムメモリー 2 3 7 2 および / または固定ディスク 2 3 7 9 は、コンピュータ可読媒体を組み入れることができる。本明細書に記載のいずれのデータも、1 つの装置から別の装置に出力でき、また、ユーザーに対しても出力できる。

【 0 1 5 9 】

コンピュータシステムは、例えば、外部インターフェイス 2 3 8 1 または内部インターフェイスにより接続された複数の同じ装置またはサブシステムを含むことができる。いくつかの実施形態では、コンピュータシステム、サブシステム、または装置は、ネットワークを介して通信できる。このような場合は、1 つのコンピュータをクライアント、別のコンピュータをサーバーと見なすことができ、この場合、それぞれは、同じコンピュータシステムの一部となりうる。クライアントおよびサーバーは、それぞれ複数のシステム、サブシステム、または装置を含んでもよい。

【 0 1 6 0 】

本発明のいずれの実施形態も、モジュールまたは集積方式で一般的にプログラム可能なプロセッサを有するハードウェア（例えば、特定の目的の集積回路またはフィールドプロ

10

20

30

40

50

グラム可能なゲートアレイ)および/またはコンピュータソフトウェアを使った制御論理の形で実装できることと理解すべきである。本明細書における開示および教示に基づいて、当業者なら、ハードウェアおよびハードウェアとソフトウェアの組み合わせを使って、本発明の実施形態を実施する他の手段、および/または方法が分かり、理解するであろう。

【0161】

本出願に記載のいずれのソフトウェア成分または機能も、例えば、Java(登録商標)、C++またはPerl等のいずれかの適切なコンピュータ言語を使い、例えば、従来の、またはオブジェクト指向技術を使って、プロセッサにより実行されるソフトウェアコードとして実装可能である。ソフトウェアコードは、保存および/または送信用のコンピュータ可読媒体上の一連の指令または命令として保存できる。適切な媒体には、ランダムアクセスメモリ(RAM)、読取り専用メモリ(ROM)、ハードドライブもしくはフロッピー(登録商標)ディスク、等の磁気メディア、またはコンパクトディスク(CD)もしくはDVD(デジタル多用途ディスク)等の光学メディア、フラッシュメモリ、等が含まれる。コンピュータ可読媒体は、このような保存または送信装置のいずれかの組み合わせであってもよい。

【0162】

このようなプログラムは、また、インターネットを含む種々のプロトコルに準拠した有線、光学、および/または無線ネットワークを介した送信に合わせたキャリアシグナルを使って、コード化および送信できる。従って、本発明の実施形態に係るコンピュータ可読媒体は、このようなプログラムでコードされたデータシグナルを使って作成できる。プログラムコードでコードされたコンピュータ可読媒体は、互換性のある装置と一緒にパッケージされても、または他の装置から別々に(例えば、インターネットのダウンロードにより)提供されてもよい。このようなコンピュータ可読媒体はいずれも、単一のコンピュータプログラム製品(例えば、ハードドライブ、CD、または全体コンピュータシステム)上にあっても、内部にあってもよく、また、システムまたはネットワーク内の異なるコンピュータプログラム製品上にあっても、内部にあってもよい。コンピュータシステムは、モニター、プリンター、または本明細書記載のいずれかの結果をユーザーに提供するための他の適切なディスプレイを含んでもよい。

【0163】

本明細書に記載の任意の方法は、全体的に、または部分的に、ステップを実行するように構成できる、1つまたは複数のプロセッサを含むコンピュータシステムを使って実施できる。従って、実施形態は、本明細書に記載の任意の方法のステップを実行するように構成され、場合によっては、それぞれのステップまたはそれぞれのステップグループを実行する異なる装置を備えた、コンピュータシステムに関するものであってもよい。本明細書の方法のステップは、番号を付したステップとして示したが、同時にまたは異なる順序で行うことができる。さらに、これらのステップの一部は、他の方法による他のステップの一部と一緒に使用できる。また、全部または一部のステップを任意に選択できる。さらに、任意の方法の任意のステップを、モジュール、回路、またはこれらのステップを実行する他の手段と一緒に行うことができる。

【0164】

特定の実施形態の特定の詳細は、本発明の実施形態の趣旨および範囲を逸脱することなく、いずれかの適切な方法で組み合わせることができる。しかし、本発明の他の実施形態は、それぞれの個別態様、またはこれらの個別態様の特定の組み合わせ、に関連する具体的な実施形態に関するものであってもよい。

【0165】

本発明の代表的な実施形態の上記記載は、例示および説明の目的で提示するものである。本発明を記載された通りの形態として網羅的または限定的とする意図はなく、上記教示を考慮した多くの修正および変形が可能である。実施形態は、本発明の原理およびその実際の適用を最良に説明し、それにより、種々な実施形態で、特定の意図された用途に適合

10

20

30

40

50

するように種々の変形を用いて、他の当業者が本発明を最良に利用できるように選択および記載した。

【0166】

特に逆の指示がなければ、「a」、「an」または「the」の記述は、「1つまたは複数の」を意味するという意図である。

【0167】

全ての特許、特許出願、出版物、および上述の記載は、それらの全内容があらゆる目的において参照により本明細書に援用される。これらのいずれも先行技術であることを認めるものではない。

【図1】

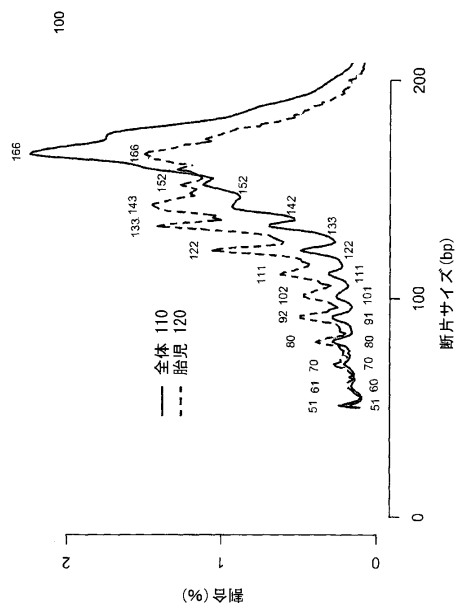


FIG. 1

【図2】

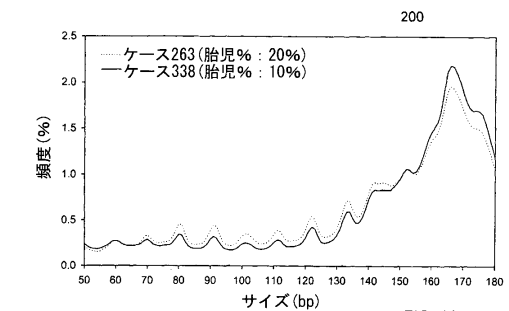


FIG. 2A

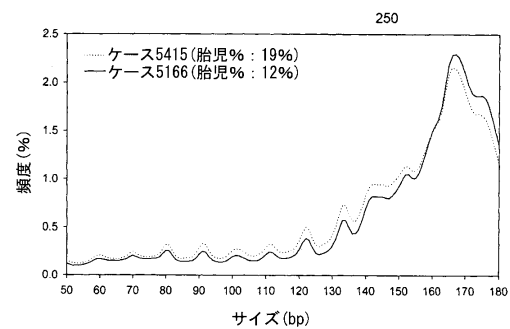


FIG. 2B

【図3】

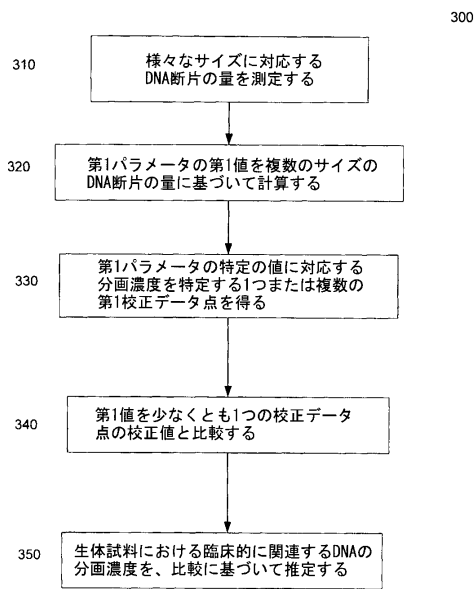
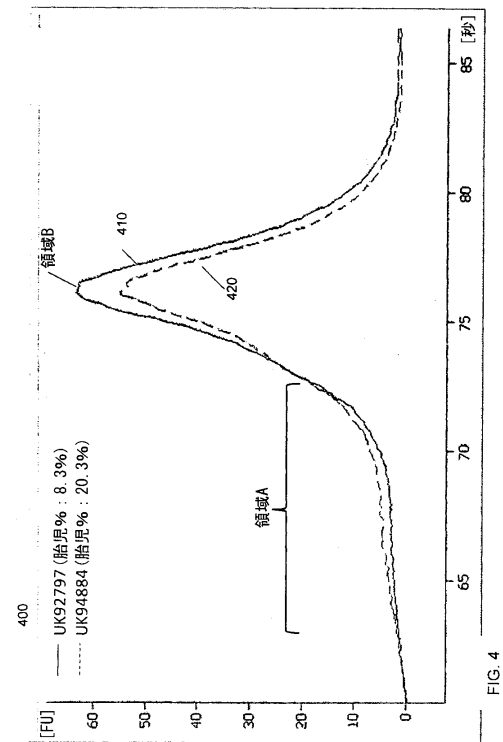
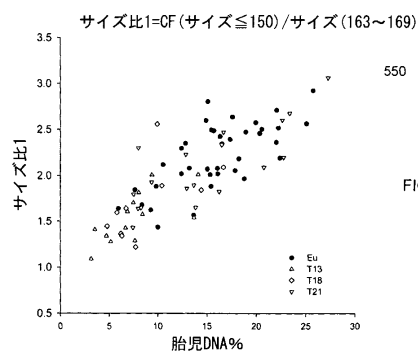
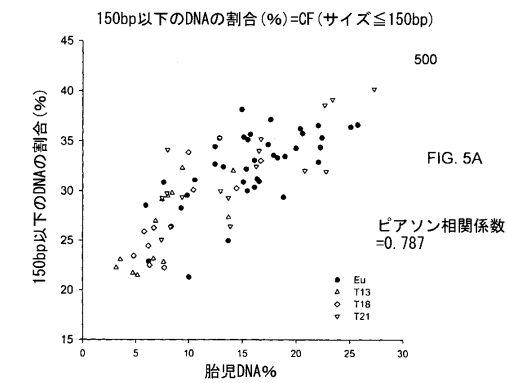


FIG. 3

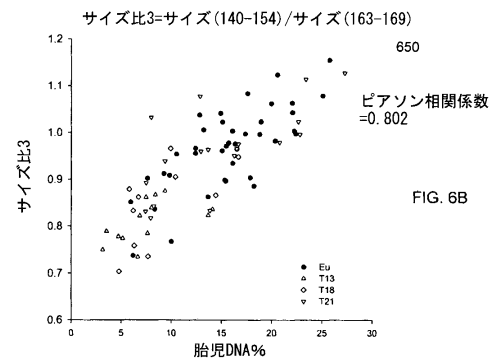
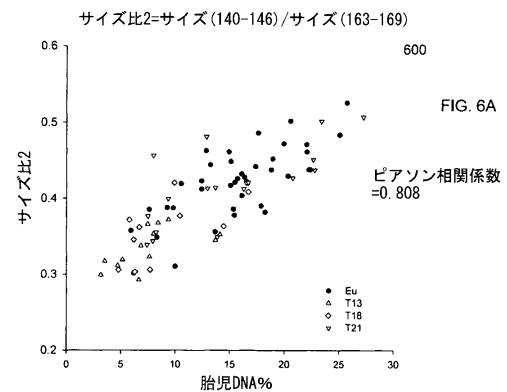
【図4】



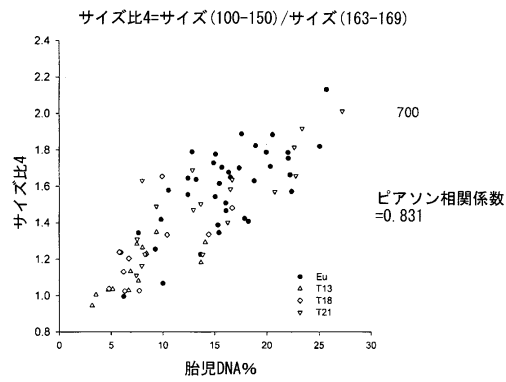
【図5】



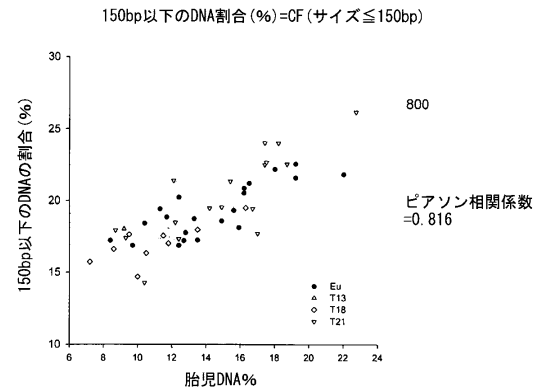
【図6】



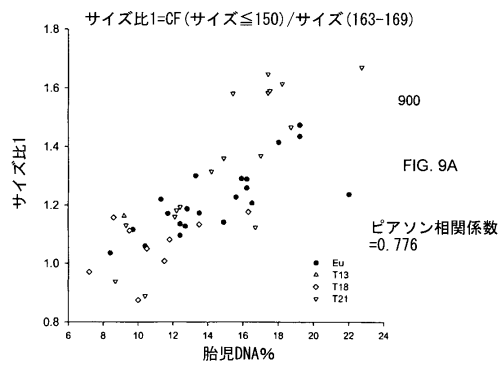
【図 7】



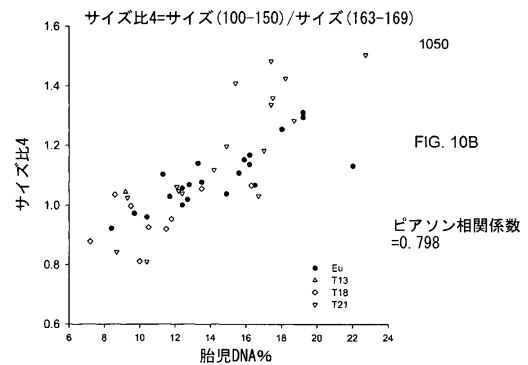
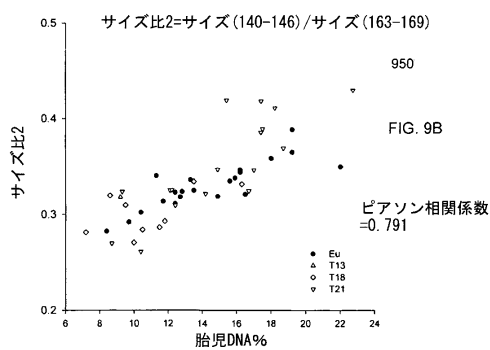
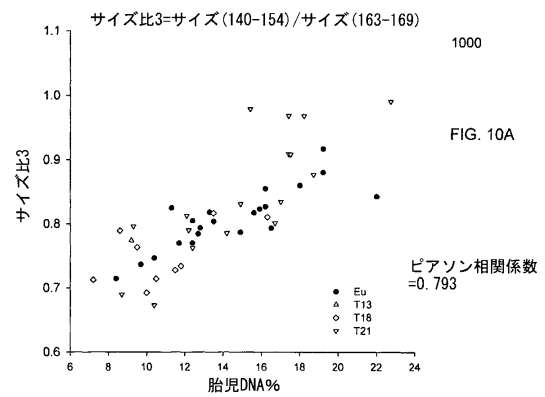
【図 8】



【図 9】



【図 10】



【図 1 1】

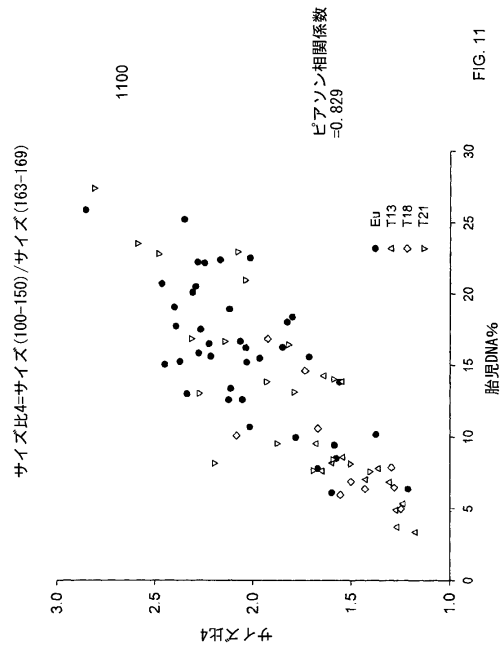


FIG. 11

【図 1 2】

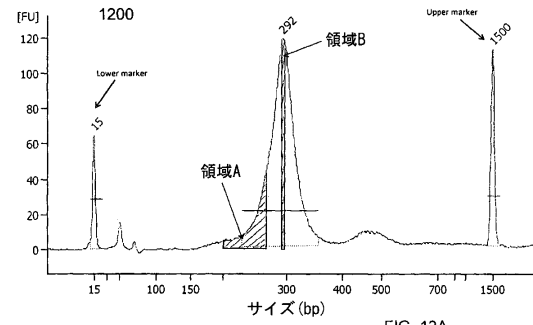


FIG. 12A

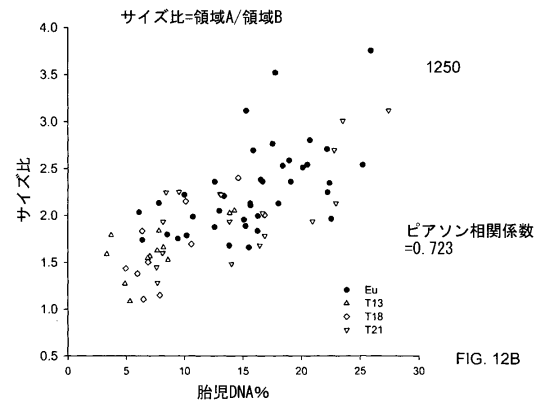


FIG. 12B

【図 1 3】

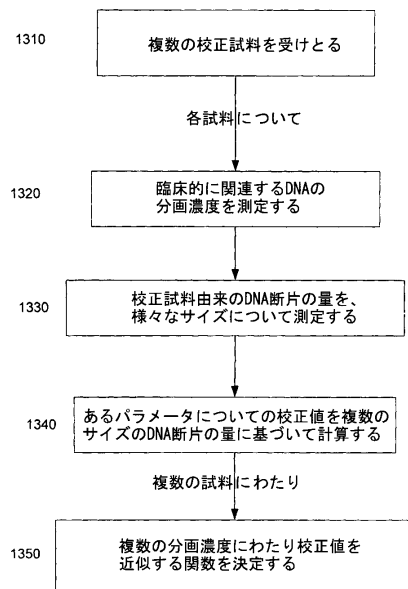


FIG. 13

【図 1 4】

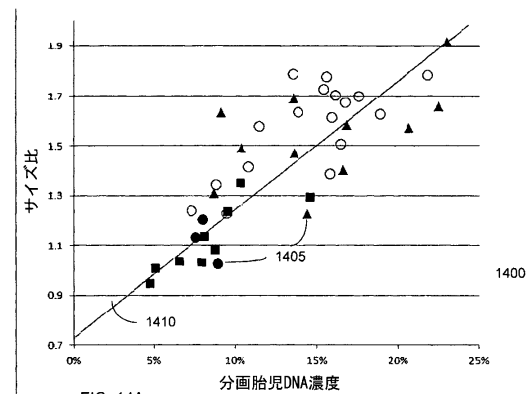


FIG. 14A

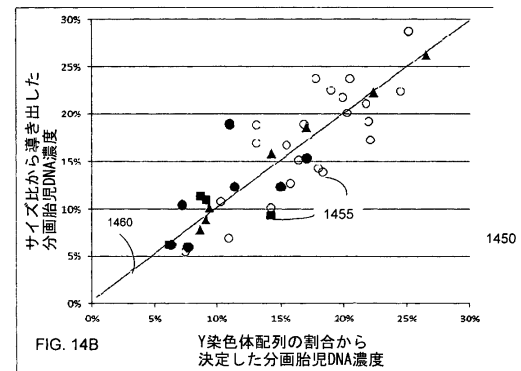
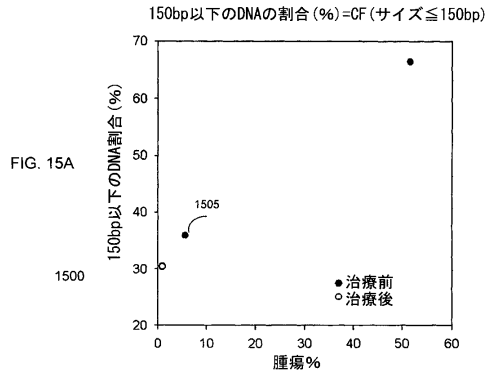
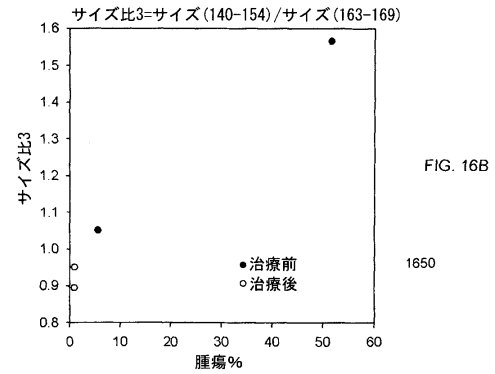
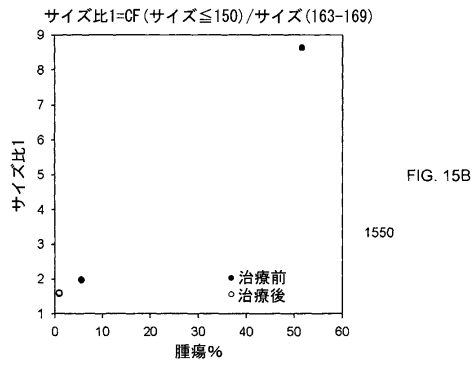
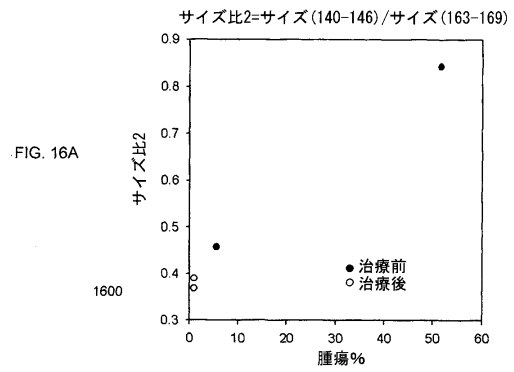


FIG. 14B

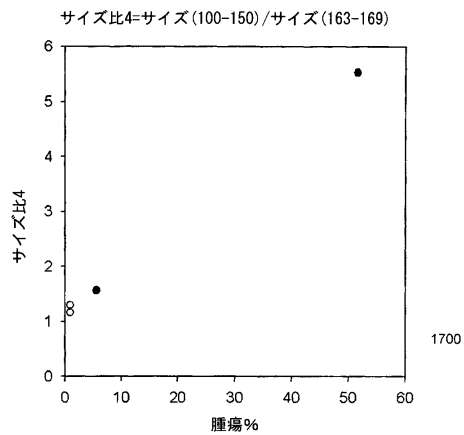
【図 15】



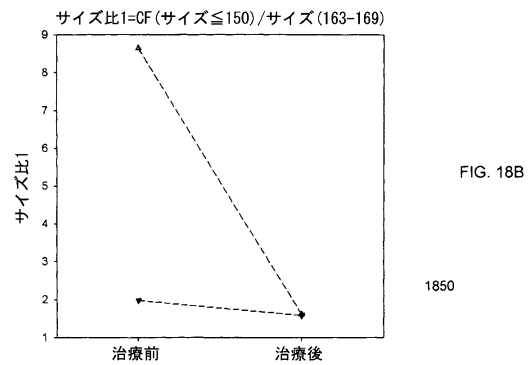
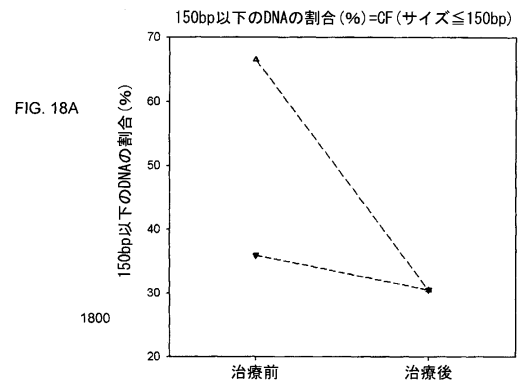
【図 16】



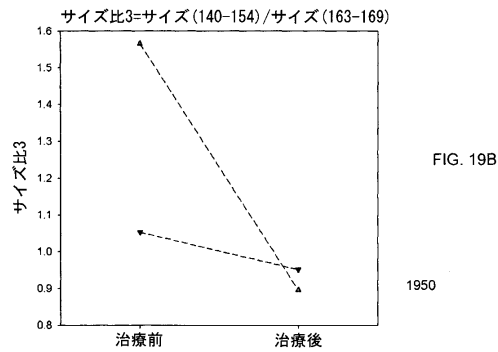
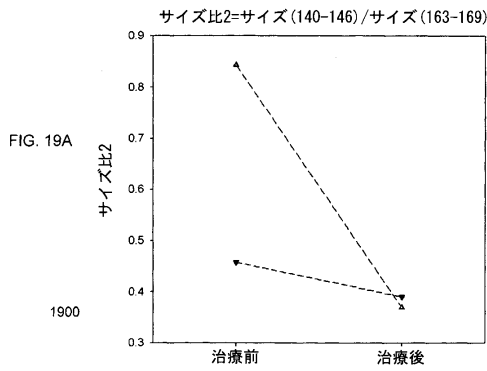
【図 17】



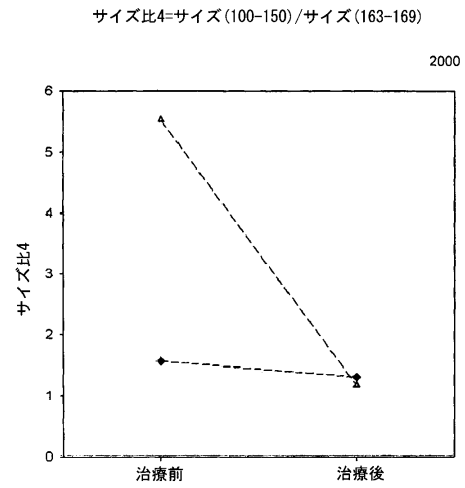
【図 18】



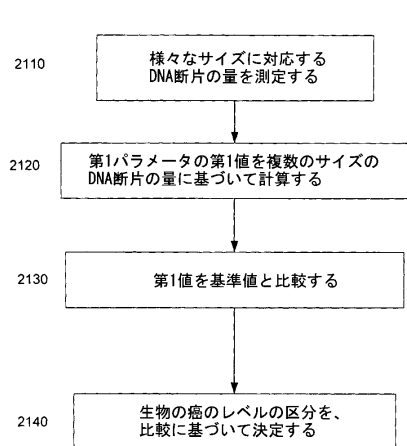
【図 19】



【図 20】



【図 21】



【図 22】

FIG. 22

| | 増加 | 欠失 | 参考文献 |
|-------|--|--------------------------------|---|
| 結腸直腸癌 | 7p, 7q, 8q, 11q, 13q, および 20q, | 5q, 8p, 17p, 18p, 18q および 20p, | (Nakao et al. Carcinogenesis 2004;25: 1345-1357.) (Tsafir et al. Cancer Res 2006; 66: 2129-2137) |
| 乳癌 | 1q, 6p, 8q, 11q, 16p, 17q, 19, および 20q | 6q, 13q, 16q, 17p, および 22q | (Tirkkonen et al. Gene Chromosome Canc 1998; 21: 177-184) (Richard et al. Int J Cancer 2000; 89: 305-310) (Pinkel et al. Nat Genet 1998; 20: 207-211) (Persson et al. Gene Chromosome Canc 1999; 25: 115-122) (Nishizaki et al. Int J Cancer 1997; 74: 513 - 517) |
| 肺癌 | 1q, 3q, 5p, および 8q | 3p, 6q, 8p, 9p, 13q, および 17p | (Berrieman et al. Brit J Cancer 2004; 90: 900-905) (Luk et al. Cancer Genet Cytogen 2001; 125: 87 - 99) (Petersen et al. Cancer Res 1997; 57: 2331-2335) (Pei et al. Gene Chromosome Canc 2001; 31: 282-287) |
| HCC | 1q, 8q, 17q および 20q | 4q, 6q, 8p, 13q, 16q および 17p | (Kusano et al. Cancer 2002; 94: 746-751) (Laurent-Puig et al. Gastroenterology 2001; 120: 1763-1773) (Moizadeh et al. Brit J Cancer 2005; 92: 935-941) |
| 卵巣癌 | 20q, 3q, 1q, 8q, 12p, 11q, および 17q | Xp, 18q, 4q, 9p, および 13q | (Taetle et al. Gene Chromosome Canc 1999; 25: 290-300) (Schraml et al. Am J Pathol 2003; 163: 985 - 992) (Sonoda et al. Gene Chromosome Canc 1997; 20: 320-328) |

【図 23】

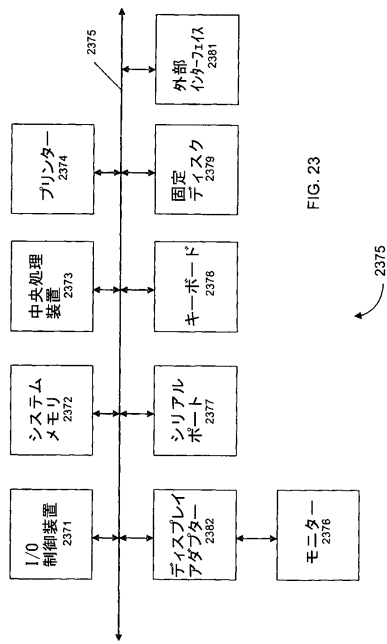


FIG. 23

フロントページの続き

- (74)代理人 100117019
弁理士 渡辺 陽一
- (74)代理人 100141977
弁理士 中島 勝
- (74)代理人 100196977
弁理士 上原 路子
- (72)発明者 ロー ユク ミン デニス
中華人民共和国, 香港特別行政区, クーロン, ホマンティン, キング タック ストリート 7,
フォース フロア
- (72)発明者 チャン クワン チー
中華人民共和国, 香港特別行政区, クーロン, メイ フー スン チェン, ブロードウェイ スト
リート, ブロック 34, フラット エー, 13 / エフ
- (72)発明者 ジュヨン ウェンリ
中華人民共和国, 香港特別行政区, ニュー テリトリーズ, ピージーエイチ2 シャーティン, ア
ール2601, シーノオーザ チャイニーズ ユニバーシティ オブ ホンコン
- (72)発明者 ジアーン ペイヨン
中華人民共和国, 香港特別行政区, ニュー テリトリーズ, シャーティン, シウ ルク ユエン
ロード ノス 62-66, クワン ラム コート, ファースト フロア オブ ブロック ビー
, フラット 7
- (72)発明者 リャオ ジアウェイ
中華人民共和国, 香港特別行政区, ニュー テリトリーズ, タイ ポ マーケット, ヤット ンガ
コート, ヤット ヤン ハウス, フラット 16, 10 / エフ
- (72)発明者 チウ ワイ クン ロッサ
中華人民共和国, 香港特別行政区, ニュー テリトリーズ, シャーティン, マ ロック パス 5
2, ダブル ヘイブン, ハウス 31

審査官 中村 勇介

- (56)参考文献 欧州特許出願公開第02426217 (EP, A1)
国際公開第2010/112316 (WO, A1)
国際公開第2009/019455 (WO, A1)
PLoS ONE, 2011年 9月, Vol.6, No.9, e23418 p.1-10
The Journal of Urology, 2009年 1月, Vol.181, p.363-371
J. AM. CHEM. SOC., 2010年, Vol.132, p.5793-5798

(58)調査した分野(Int.Cl., DB名)

C12Q 1/00 - 3/00

C12N15/00

JSTPlus/JMEDPlus/JST7580(JDreamIII)

CAplus/MEDLINE/EMBASE/BIOSIS(STN)