

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2018-75013

(P2018-75013A)

(43) 公開日 平成30年5月17日(2018.5.17)

(51) Int.Cl.	F I	テーマコード (参考)
<b>C 1 2 Q 1/68</b> (2018.01)	C 1 2 Q 1/68	Z 4 B 0 6 3
<b>G O 1 N 33/53</b> (2006.01)	G O 1 N 33/53	M
<b>G O 6 F 19/22</b> (2011.01)	G O 6 F 19/22	
<b>C 1 2 N 15/09</b> (2006.01)	C 1 2 N 15/00	A

審査請求 有 請求項の数 13 O L (全 25 頁)

(21) 出願番号	特願2017-225951 (P2017-225951)	(71) 出願人	515177387 ザ チャイニーズ ユニバーシティ オブ ホンコン 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
(22) 出願日	平成29年11月24日 (2017.11.24)		
(62) 分割の表示	特願2015-551935 (P2015-551935) の分割	(74) 代理人	100099759 弁理士 青木 篤
原出願日	平成26年1月9日 (2014.1.9)	(74) 代理人	100087871 弁理士 福本 積
(31) 優先権主張番号	61/751, 213	(74) 代理人	100087413 弁理士 古賀 哲次
(32) 優先日	平成25年1月10日 (2013.1.10)	(74) 代理人	100117019 弁理士 渡辺 陽一
(33) 優先権主張国	米国 (US)	(74) 代理人	100150810 弁理士 武居 良太郎
(31) 優先権主張番号	13/837, 776		
(32) 優先日	平成25年3月15日 (2013.3.15)		
(33) 優先権主張国	米国 (US)		

最終頁に続く

(54) 【発明の名称】 母体血漿の無侵襲的出生前分子核型分析

(57) 【要約】

【課題】 胎児ゲノム中に微小増幅又は微小欠失が存在するか否かを判断する、方法、システム及び装置を提供すること。

【解決手段】 胎児ゲノム中の微小増幅又は微小欠失を検出する方法、システム及び装置が本明細書に開示されている。一部の実施形態において、方法は、生体試料の複数のDNA断片のそれぞれについて配列タグを得ること、配列タグのゲノム位置を決定すること、複数のゲノム領域のそれぞれにおけるDNAの密度が異常に高いか又は低いかを判断すること、異常に高い密度を有した連続するゲノム領域の集合を微小増幅と識別すること、異常に低い密度を有した連続するゲノム領域の集合を微小欠失と識別すること、を含む。生体試料は、胎児を宿した女性対象から無侵襲的に得られる血液試料であり得る。

【選択図】 図 1

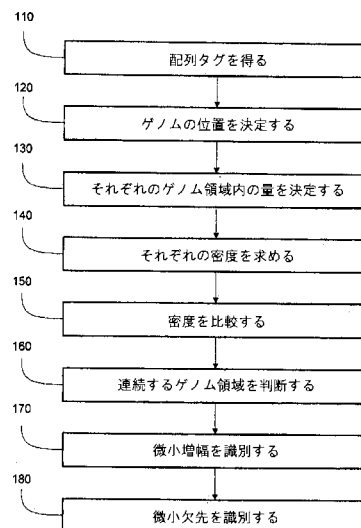


FIG. 1

## 【特許請求の範囲】

## 【請求項 1】

胎児を宿した女性対象から得られる生体試料を分析することによって、前記胎児のゲノム中の微小増幅又は微小欠失を識別する方法であって、前記生体試料には、前記胎児及び前記女性対象からの無細胞DNAが含まれ、前記方法は、

前記生体試料の複数のDNA断片のそれぞれについて1以上の配列タグを得ることと、  
複数のゲノム領域のそれぞれについて、前記配列タグのゲノム位置を決定することと、  
コンピュータシステムによって、前記ゲノム領域内にゲノム位置を有する配列タグから、  
前記ゲノム領域内のDNA断片のそれぞれの量を決定することと、  
前記それぞれの量を標準化して、それぞれの密度を求めることと、

該それぞれの密度と基準密度を比較して、前記それぞれの密度が、前記基準密度と統計的に異なるか否かを識別することと、

前記基準密度と統計的に異なるそれぞれの密度を有すると識別されたいずれかの前記ゲノム領域が、前記基準密度と統計的に異なるそれぞれの密度を有すると識別された別のゲノム領域と連続しているか否かを判断することと、

前記基準密度よりも統計的に高いそれぞれの密度を有すると識別された少なくともN (Nは2以上の整数である。)の第1のゲノム領域が連続している場合、該第1の連続するゲノム領域を、微小増幅と識別することと、

前記基準密度よりも統計的に低いそれぞれの密度を有すると識別された少なくともNの第2のゲノム領域が連続している場合、該第2の連続するゲノム領域を、微小欠失と識別すること、

とを含む、方法。

## 【請求項 2】

前記生体試料が、母体血液、血漿、血清、尿又は唾液である、請求項1に記載の方法。

## 【請求項 3】

前記配列タグが、超並列配列決定によって得られる、請求項1に記載の方法。

## 【請求項 4】

Nが3である、請求項1に記載の方法。

## 【請求項 5】

前記ゲノム領域がオーバーラップしていない、請求項1に記載の方法。

## 【請求項 6】

前記ゲノム領域が隣接している、請求項1に記載の方法。

## 【請求項 7】

前記ゲノム領域が等しいサイズである、請求項1に記載の方法。

## 【請求項 8】

それぞれのゲノム領域のサイズが約1Mbである、請求項7に記載の方法。

## 【請求項 9】

前記ゲノム領域のそれぞれの密度が、前記ゲノム領域のDNA断片のそれぞれの量を、複数のゲノム領域のDNA断片の総量で割ることによって求められる、請求項1に記載の方法。

## 【請求項 10】

前記ゲノム領域のそれぞれの密度が、前記ゲノム領域のDNA断片のそれぞれの量と等しい、請求項1に記載の方法。

## 【請求項 11】

前記ゲノム領域の基準密度が、前記ゲノム領域において微小増幅又は微小欠失を示さない1以上の他の生体試料から決定される、複数のそれぞれの密度の平均又は中央値である、請求項1に記載の方法。

## 【請求項 12】

前記ゲノム領域の基準密度が、他のゲノム領域について求められる、複数のそれぞれの密度の平均又は中央値である、請求項1に記載の方法。

10

20

30

40

50

## 【請求項 13】

微小増幅又は微小欠失が母性遺伝されているか否かを判断するステップを更に含む、請求項 1 に記載の方法。

## 【請求項 14】

前記微小増幅又は前記微小欠失に相当する前記連続するゲノム領域のそれぞれにおいて、前記それぞれの密度と前記基準密度との間の差が特定のカットオフを超える場合、微小増幅又は微小欠失が母性遺伝されていると判断され、前記特定のカットオフが、前記それぞれの密度及び前記基準密度が統計的に異なるか否かを判断するのに用いられるカットオフよりも大きい、請求項 13 に記載の方法。

## 【請求項 15】

複数の指示を記憶する非一時的コンピュータ可読媒体を備えたコンピュータ製品であって、該コンピュータ製品は、実行される場合、前記胎児を宿した女性対象から得られる生体試料を分析することによって、胎児のゲノム中の微小増幅又は微小欠失を識別するように、コンピュータシステムを制御し、前記生体試料には、前記胎児及び前記女性対象からの無細胞 DNA が含まれ、該方法は、

前記生体試料の複数の DNA 断片のそれぞれについて 1 以上の配列タグを得ることと、  
 複数のゲノム領域のそれぞれについて、前記配列タグのゲノム位置を決定することと、  
 前記ゲノム領域内にゲノム位置を有する配列タグから、前記ゲノム領域内の DNA 断片のそれぞれの量を決定することと、

前記それぞれの量を標準化して、それぞれの密度を求めることと、

該それぞれの密度と基準密度を比較して、前記それぞれの密度が、前記基準密度と統計的に異なるか否かを識別することと、

前記基準密度と統計的に異なるそれぞれの密度を有すると識別されたいずれかの前記ゲノム領域が、前記基準密度と統計的に異なるそれぞれの密度を有すると識別された別のゲノム領域と連続しているか否かを判断することと、

前記基準密度よりも統計的に高いそれぞれの密度を有すると識別された少なくとも N (N は 2 以上の整数である。) の第 1 のゲノム領域が連続している場合、該第 1 の連続するゲノム領域を、微小増幅と識別することと、

前記基準密度よりも統計的に低いそれぞれの密度を有すると識別された少なくとも N の第 2 のゲノム領域が連続している場合、該第 2 の連続するゲノム領域を、微小欠失と識別することと、

とを含む、コンピュータ製品。

## 【請求項 16】

N が 3 である、請求項 15 に記載のコンピュータ製品。

## 【請求項 17】

前記ゲノム領域がオーバーラップしていない、請求項 15 に記載のコンピュータ製品。

## 【請求項 18】

前記ゲノム領域が隣接している、請求項 15 に記載のコンピュータ製品。

## 【請求項 19】

前記ゲノム領域が等しいサイズである、請求項 15 に記載のコンピュータ製品。

## 【請求項 20】

前記ゲノム領域のそれぞれの密度が、前記ゲノム領域の DNA 断片のそれぞれの量を、複数のゲノム領域の DNA 断片の総量で割ることによって求められる、請求項 15 に記載のコンピュータ製品。

## 【発明の詳細な説明】

## 【背景技術】

## 【0001】

## 関連出願の相互参照

本出願は、2013年1月10日に提出された「母体血漿の無侵襲的出生前分子核型分析」という名称の米国仮特許出願第 61/751,213 号、及び、2013年3月15

10

20

30

40

50

日に出願された「母体血漿の無侵襲的出生前分子核型分析」という名称の米国特許出願第 13 / 837, 776 号に対する優先権を主張する。それぞれの優先権主張出願は、すべての目的上、その全体において参照により本明細書に援用される。

#### 【0002】

本出願は、Loらによって2008年7月23日に出願された「超並列ゲノム配列決定を用いる胎児染色体異数性の診断(Diagnosing Fetal Chromosomal Aneuploidy Using Massively Parallel Genomic Sequencing)」という名称の共有米国特許出願第12 / 178, 181号に関連しており、この開示は、その全体において参照により本明細書に援用される。

#### 【0003】

##### 背景

母体血漿中の胎児DNAの存在によって、無侵襲的出生前検査を刺激する可能性が開かれている[1, 2]。近年、出生前検査の目的で循環胎児DNAを分析する超並列配列決定(MPS)の利用への関心が高い。このため、胎児のトリソミー21、13、18、及び選択性染色体異数性は、母体血漿DNAにおいてMPSを用いて検出され[3-7]、臨床サービスに急速に導入されている。

#### 【0004】

染色体全体に關与するコピー数の変化による異常とは別に、母体血漿のMPSに基づく分析は、サブ染色体の欠失又は複製を検出するのに十分な感度があり得るか否かについて評価することが重要である。この点に関して、Petersらは、妊娠35週目において得られる母体血漿試料中の染色体12上の4.2 Mbの欠失検出について報告している[8]。Jensenらは、妊娠19週目及び20週目において2人の妊婦から得られる母体血漿試料中の染色体22上の3 Mbの欠失検出について報告している[9]。欠失領域とは別に、Petersらは、染色体12上の別の領域とともに、染色体14上の20 Mbの非オーバーラップ4 Mb領域に関して統計分析も行っている[8]。一方で、Jensenらは、染色体22上の欠失領域に関する統計分析に焦点を置いているにすぎない[9]。したがって、Petersら及びJensenらによって示されているデータから、このアプローチが、微小欠失(microlletion)若しくは微小重複(microduplication)のゲノムワイド調査に十分ロバストであるか、又は、実際に、胎児核型の無侵襲的決定に十分ロバストであるか否かは明らかではない。

#### 【0005】

Loらは、胎児の一塩基多型(SNP)が、母体血漿DNA配列決定を用いて、ゲノムワイド規模により遺伝子型を決定できることを報告している[10]。特に、これらの研究者は、母親から胎児に遺伝する単一遺伝子障害のSNP対立遺伝子及び突然変異が、相対的ハプロタイプ用量分析(haplotype dosage analysis)と呼ばれる方法によって解明できることを実証している[10]。Fanらは、相対的ハプロタイプ用量分析のロバスト性を確認し、また、母親から胎児に遺伝する~2.85 Mbの欠失を検出するのに、このアプローチを用いている[11]。この方法を無侵襲的出生前核型分析の臨床的实施に用いることに対しては、2つの懸念がある。第一に、この方法は、母体のハプロタイプ決定を行う必要があり、このハプロタイプ決定は、追加の分析手順[12, 13]又は家系分析が必要である。第二に、この方法を用いて、デノボサブ染色体の欠失又は重複を検出できたか否かは不明である。

#### 【発明の概要】

#### 【0006】

##### 簡単な概要

胎児ゲノム中の微小増幅(microamplification)又は微小欠失を検出する方法、システム及び装置が本明細書に開示されている。一部の実施形態において、方法は、生体試料の複数のDNA断片のそれぞれについて配列タグを得ることと、配列タグのゲノム位置を決定することと、複数のゲノム領域のそれぞれにおけるDNAの密度が異常に高いか又は低いかを判断することと、異常に高い密度を有した連続するゲノム領域の集合を微小増幅と識

10

20

30

40

50

別することと、異常に低い密度を有した連続するゲノム領域の集合を微小欠失と識別することと、を含む。生体試料は、胎児を宿した女性対象から無侵襲的に得られる血液試料であり得る。

【図面の簡単な説明】

【0007】

【図1】図1は、胎児ゲノム中の微小増幅又は微小欠失を識別する方法を示すフローチャートである。

【図2】図2は、母体血漿中のゲノムにおいて検出されたコピー数異常のCircosプロットである。内側から外側において、事例01～06である。染色体の表意文字（最も外側のリング）は、時計回り方向に、長腕末端(qter)に対し短腕末端(pter)に配向される。それぞれのバーは、1 Mbのウィンドウを表す。血漿の増加又は低減を表す、3つ以上の連続する1 Mbのピン(bin)を有する領域をそれぞれ、緑色及び赤色のバーで示す。赤色の矢印は、これらの異常な領域上のおおよその染色体の位置を強調する。

【図3】図3は、母体血漿中で検出されたコピー数異常を示す。それぞれの事例のコピー数異常を示す染色体を示す。ゲノムの位置は、x軸上に示され、zスコアは、y軸上にプロットされている。それぞれの垂直バーは、1 Mbのピンを表す。血漿中の増加又は低減を表す、3つ以上の連続する1 Mbのピンを有する領域をそれぞれ、緑色及び赤色のバーで示す。

【図4】図4は、母体血漿中で検出されたコピー数異常を示す。それぞれの事例のコピー数異常を示す染色体を示す。ゲノムの位置は、x軸上に示され、zスコアは、y軸上にプロットされている。それぞれの垂直バーは、1 Mbのピンを表す。血漿中の増加又は低減を表す、3つ以上の連続する1 Mbのピンを有する領域をそれぞれ、緑色及び赤色のバーで示す。

【図5】図5は、微小欠失/微小重複による影響を受けた領域のゲノム表現の変化によって推定された胎児DNAの割合、及び母体血漿の染色体のY配列の割合を示す表である。(a)染色体Y(chr Y)アプローチは、男性胎児による事例にのみ適用される。(b)事例05では、母親も異常を有していたため、母体血漿中の影響を受けた領域のゲノム表現を用いて、胎児DNAの割合を決定することができなかった。(c)前者及び後者の図はそれぞれ、染色体3上の微小重複及び染色体4上の微小欠失から推定された、胎児DNAの割合を表す。

【図6】図6は、3 Mbの微小欠失/微小重複を検出する診断感度を示す。異常を検出する診断感度は、胎児DNAの割合に対してプロットされている。コンピュータシミュレーション解析は、計1億5000万の血漿DNA分子を分析したものと仮定して行われた。

【図7】図7は、種々の診断解像度、及び胎児DNAの割合を5%と仮定して診断感度を達成するために、配列決定及び配列させるのに必要な分子数を示す表である。この理論解析では、診断特異性は、3つの連続するピンが、同じ方向に参照の平均から(過剰表現(overrepresentation)又は過少表現(underrepresentation)のいずれかにおいて)3SDを超えるゲノム表現を有するという基準に基づいて、すべての事例において99.9%を超える。

【図8】図8は、実施例において論じる試料についての情報を示す表である。

【図9】図9は、本発明の実施形態に係るシステム及び方法とともに使用可能な例示的なコンピュータシステム900のブロック図を示す。

【発明を実施するための形態】

【0008】

#### 詳細な説明

本発明の実施形態は、胎児ゲノム中に微小増幅又は微小欠失が存在するか否かを判断する、方法、システム及び装置を提供する。簡潔にいうと、このような判断は、生体試料を得て、複数のゲノム領域のそれぞれに由来する試料中のゲノムDNAの量を定量することによって、行うことができる。それぞれのゲノム領域の量は、適切に標準化して、その領

10

20

30

40

50

域の密度（すなわち、それぞれの密度）を求め、基準密度と比較することができる。それぞれの密度と基準密度との間の統計的有意差は、ゲノム領域若しくは複数にわたるゲノム領域内の微小増幅又は微小欠失の存在を示し得る。偽陽性を回避するために、かかる統計的有意差が少なくとも2つの連続するゲノム領域のそれぞれに存在する場合、微小増幅又は微小欠失が識別される。

#### 【0009】

##### I. 緒言

実施形態は、遺伝子のコピー数又は染色体領域の相違を検出するのに用いることができる。微小増幅（微小重複とも呼ばれる。）から生じる異常に高いコピー数の遺伝子は、これらの遺伝子の過剰発現又は病的発現(pathological expression)を生じることがあり、これらは、癌などの疾患をもたらす。逆に、微小欠失から生じる低いコピー数の遺伝子は、発現の減少、又は生物学的（例えば、酵素）機能の損失を生じることがある。したがって、異常なコピー数の検出は、胎児の出生前又は出生後に直面し得る疾患の早期警告を提供することができる。

10

#### 【0010】

生体試料は、胎児を宿した女性対象から無侵襲的に得ることができる。試料は、血液、血漿、血清、尿又は唾液を含み得る。血液は、無細胞DNA断片を含有し、妊娠した対象では、これらの断片の一部は胎児由来である。DNA断片は、配列タグを得るように、例えば超並列配列決定法を用いることによって配列決定することができる。合成による配列決定プラットフォーム(sequencing-by-synthesis platform)（例えば、イルミナ社(Illumina) / ソレクサ社(Solexa)）、ライゲーションによる配列決定プラットフォーム(sequencing-by-ligation platform)（例えば、ライフテクノロジー(Life Technology) SOLiDプラットフォーム）、又は単一分子配列決定プラットフォーム（例えば、ヘリコス社(Helicos)若しくはパシフィックバイオサイエンス社(Pacific Biosciences)）を含む、いずれかの超並列配列決定プラットフォームを用いることができる。それぞれの配列タグは、これが発生したDNA断片の配列のすべて又は一部を含有し、断片の由来のゲノム領域を決定するように、参照ゲノム配列とともに配列させることができる。

20

#### 【0011】

「ピン」とも呼ばれるゲノム領域は、参照ゲノム配列を分割することによって表現することができる。この領域は、染色体の連続する配列部分に相当する。好適な実施形態において、それぞれの領域は、1つの染色体に関連付けられ、複数の染色体にわたらない。一部の実施形態において、領域は、1 Mbなどの等しいサイズである。ゲノム領域のサイズは、本明細書に開示されている方法の解像度と、統計的确实性を有した微小増幅及び微小欠失を識別するのに必要なDNA断片の数を決定する。

30

#### 【0012】

##### II. 方法

図1は、胎児を宿した女性対象から得られる生体試料を分析することによって、胎児のゲノム中の微小増幅又は微小欠失を識別する、方法100のフローチャートであり、生体試料には、胎児及び女性対象からの無細胞DNAが含まれる。

#### 【0013】

ステップ110において、生体試料中の複数のDNA断片のそれぞれについて1以上の配列タグが得られる。配列タグは、当該技術分野において公知のいずれかの方法、例えば、サンガー(Sanger)配列決定又は超並列配列決定を用いて、DNA断片を配列決定することによって得ることができる。それぞれのタグは、シークエンス「リード」(sequencing 'read')と呼ばれ、これが発生したDNA断片の一部のすべてに相当することもある。例えば、タグは、断片の一端又は断片内部の配列を含有していてもよい。

40

#### 【0014】

DNA断片は、生体試料から単離され、いずれかの公知の方法を用いて、配列決定のために調製することができる。例えば、断片は、例えばポリメラーゼ連鎖反応(PCR)を用いて、配列決定前にコピーされてもよいし、又は、用いられる配列決定法に適切なアダ

50

ブター分子若しくは「バーコード」配列にライゲーションされてもよい。また、断片は、イルミナ配列決定法及びこれに類似する方法において行われるように、ブリッジ増幅(bridge amplification)のクローンクラスタを発生させるのに用いてもよい。配列決定のために調製されるDNA断片の集合は、「シークエンスライブラリ」と称することができる。

#### 【0015】

一部の実施形態において、配列タグは、DNA断片のコピーが両端から2つの方向において配列決定される、ペアエンド配列決定(paired-end sequencing)に従って発生する。2つの方向の配列データの比較によって、断片、具体的には断片の末端の配列を照合することができる。ペアエンド配列決定から得られる配列タグの生の数値は、試料中のDNA断片数の推定値を提供する。この数は、試料のサイズ及び性質によって、100万、1000万、1億、10億及び100億の間で変化し得る。一部の実施形態において、多重配列タグが同一の配列(すなわち、両端で同一の配列)を表す場合、重複タグは廃棄されるか、又は更なる分析から除外することができる。

#### 【0016】

ステップ120において、配列タグのゲノムの位置が決定される。これは、最初に、当該技術分野において公知の標準的方法を用いて、それぞれの配列タグと、非反復ヒト参照ゲノム(NCBI構築36.1/hg18)などの参照配列を配列させることによって行われる。一部の実施形態において、同じ染色体に配列させた両端を有するタグのみが、更なる分析のために維持される。また、例えば、参照配列と非常に多くの不一致が存在する場合、又は、タグが所定のサイズの範囲内でない場合、タグを破棄してもよい。そして、上述したように、その配列に基づいて、それぞれのタグは、参照配列内のゲノム領域又は「ピン」に割り当てられる。

#### 【0017】

ステップ130において、複数のゲノム領域のそれぞれについて、ゲノム領域内のDNA断片のそれぞれの量は、ゲノム領域内にゲノム位置を有する配列タグから決定される。ゲノム領域のDNA断片のそれぞれの量は、領域に割り当てられた配列タグに含有するデータの全体、又はその一部から算出することができるパラメータである。これらのデータには、例えば、タグの数及び長さ、並びにこれらの間のオーバーラップの量が含まれる。量パラメータは、配列タグを生成するのに用いられる方法の知識により算出し、同一のDNA断片に由来し得る複数のタグの存在などのこの方法の結果を考慮する。ゲノム領域内のそれぞれの量は、例えば、試料中に存在している及びゲノム領域に由来すると推測される断片の数、又はゲノム領域からのDNAの全質量であり得る。好適な実施形態において、ゲノム領域内のDNA断片のそれぞれの量は、試料中のその領域のDNA量の定量的尺度である。

#### 【0018】

ステップ140において、複数のゲノム領域のそれぞれについて、それぞれの量は標準化され、それぞれの密度が求められる。標準化は、多くの方法により、例えば、ゲノム領域のそれぞれの量を、ゲノム領域が発生する染色体のそれぞれの量の合計、又は全ゲノムのそれぞれの量の合計で割ることにより行うことができる。標準化の出力、それぞれの密度によって、比較される種々のゲノム領域からの、例えば、種々の染色体又は試料からのそれぞれの量が許容される。それぞれの密度には、ゲノム領域に由来する試料中のDNA断片数の割合などの多くの異なる解釈があり得る。一部の実施形態において、それぞれの量を直接比較することができる場合、標準化は不要であり、それぞれの密度は、それぞれの量と等しくなり得る。

#### 【0019】

ステップ150において、複数のゲノム領域のそれぞれについて、それぞれの密度は、基準密度と比較され、それぞれの密度が、基準密度と統計的に異なるか否かが識別される。ゲノム領域の基準密度は、試料のデータの一部又は全部を用いて求めることができる。基準密度は、例えば、染色体又は全ゲノムのそれぞれの密度の平均と等しくなり得る。また、基準密度は、他の試料のデータを用いて算出することもできる。例えば、試料は、そ

10

20

30

40

50

れぞれ胎児を宿した複数の女性対象から得られてもよく、また、ステップ110～140は、同様に、それぞれの試料において行ってもよい。このように、特定のゲノム領域の基準密度は、すべての試料からのその領域、又は試料の部分集合のそれぞれの密度の平均である。一部の実施形態において、独立したデータは、女性対象が、それぞれ、ゲノムの微小増幅又は微小欠失を欠く胎児を宿しており、このため、基準密度が微小増幅又は微小欠失の欠如を反映するように算出できることを実証するように存在し得る。

#### 【0020】

上述したことを考慮すると、基準密度は、(どのように算出されるかによって)複数若しくはすべてのゲノム領域において同じであるか、又はそれぞれのゲノム領域ごとに異なる値であり得る。

#### 【0021】

それぞれの密度が基準密度と統計的に異なるか否か(すなわち、差が統計的に有意であるか否か)の識別は、基準密度を算出するのに用いられる、それぞれの密度の分布についての知識を必要とし得る。例えば、1つのゲノム領域の基準密度がそれぞれの密度の集合の平均である場合、その集合の標準偏差を算出することができる。(例えば、関心対象の試料の)そのゲノム領域の1つのそれぞれの密度、基準密度及び標準偏差を用いて、統計的検定を行うことができる。かかる統計的検定は、Z検定又はスチューデントのt検定であり、それぞれの密度が基準値と同じ分布から導き出される確率を提供することができる。この確率が閾値を下回る場合、それぞれの密度と基準密度との間の差は、統計的に有意であると考えることができる。または、差がカットオフを超える場合、例えば、特定の倍数の標準偏差を超える場合、差は、統計的に有意であると考えることができる。

#### 【0022】

それぞれの密度と基準密度の比較、及びこれらの間の統計的な差の識別は、すべてのゲノム領域において同じ基準を用いてなされることが好ましい。このステップにおいて、統計的な差がゲノム領域に存在する場合、それぞれの密度が基準密度よりも高いか又は低いかについて留意するように注意する必要がある。

#### 【0023】

平均及び標準偏差に加えて、それぞれの密度の集合についての他のパラメータを算出し、基準密度を決定し、それぞれの密度と基準密度が異なるか否かを判断することができる。これらのパラメータには、中央値、最頻値、パーセンタイル、分散、スキュー、尖度及び他のものが含まれるが、これらに限定されない。Z検定及びt検定に加えて、これらの目的上、他の統計的検定、例えば、入力情報として前述のパラメータを用いた検定を利用することができる。

#### 【0024】

ステップ160において、基準密度と統計的に異なるそれぞれの密度を有すると識別されたゲノム領域のいずれかが、他のそのように識別されたゲノム領域と連続しているか否かが判断される。施術者の関心に適合するように、この判断は、染色体の一部、染色体全体、複数の染色体又は全ゲノムに対して行うことができる。ここで、ゲノム領域が参照ゲノム配列の連続部分に相当する場合、ゲノム領域は連続している。好適な実施形態において、ゲノム領域が同じ染色体に相当する場合、ゲノム領域は連続し得るにすぎない。ここで、関心対象は、主に、それぞれの密度がそれぞれの領域の基準密度と統計的に異なり、差のすべてが同じ方向にある、連続するゲノム領域の集合である。すなわち、かかる集合内では、それぞれの密度すべてが、基準密度よりも高い(すなわち、統計的に高い)か、又は、すべてが、基準密度よりも低い(すなわち、統計的に低い)。ゲノム領域のかかる集合の例は、3つの連続する領域であって、3つのすべての領域において、基準密度が統計的に異なり、基準密度よりも高い(すなわち、統計的に高い)領域である。基準密度よりも統計的に高いか又は低いそれぞれの密度を有した連続するゲノム領域はそれぞれ、微小増幅又は微小欠失と一致している。集合内の連続するゲノム領域の数が大きいほど、大きな微小増幅又は微小欠失と一致している。

#### 【0025】

10

20

30

40

50

ステップ170において、第1の連続するゲノム領域は、基準密度よりも統計的に高いそれぞれの密度を有すると識別された少なくともN（Nは2以上の整数である。）の第1のゲノム領域が連続している場合、微小増幅と識別される。

【0026】

ステップ180において、第2のゲノム領域は、基準密度よりも統計的に低いそれぞれの密度を有すると識別された少なくともN（Nは2以上の整数である。）の第2のゲノム領域が連続している場合、微小欠失と識別される。

【0027】

#### III. 異常領域の判断

本発明を用いてゲノム微小増幅及び微小欠失を検出するために、参照ゲノム配列は、ゲノム領域又は「ピン」に分割される。試料からのDNA断片は、配列に従ってそれぞれの領域に関連付けられ、それぞれの領域内のDNAの密度が決定される。異常に高いか又は低い密度を有する領域は、「異常」と考えられ、微小増幅又は微小欠失に相当し得る。異常なゲノム領域を識別し、これらが微小増幅又は微小欠失に相当するか否かを判断するための基準及び手順を、以下に説明する。

【0028】

#### A. ピン及び量

本明細書に用いられるゲノム領域のサイズは、施術者が所望するような、及び用いられる配列決定法に適切な種々のサイズとすることができる。小さな領域によって、異常な密度の高い解析が可能になるが、統計的确实性を有する異常な密度を識別するのに多くのDNA断片（したがって、大きな試料）が必要である。逆に、大きな領域は、不十分な解析を提供するが、少ないDNA断片が必要であるにすぎない。好適な実施形態において、等しいサイズのゲノム領域を用いて、染色体及びゲノム全体における密度の比較及び標準化が可能になる。例えば、ゲノム領域のサイズは、100 kb、200 kb、500 kb、1 Mb、2 Mb、5 Mb又は10 Mbのオーダーであってもよい。一部の実施形態において、ゲノム領域は、オーバーラップしていない及び/又は（以下に論じるように）隣接しているが、場合によっては、例えば、領域の端部付近で生じる配列タグの分析を簡素化するために、領域間にオーバーラップ又はギャップがあってもよい。

【0029】

DNA断片は、当該技術分野において公知の超並列配列決定法を用いて配列決定することができる。一部の実施形態において、配列決定は、イルミナの合成による配列決定法を用いて行われる。この方法は、配列決定されるDNA断片のGC含量によって、一貫性のない効率を有する配列タグを生成することが知られている。したがって、数百から数千の塩基対に及ぶ規模の可変レベルのGC含量を有する、ゲノム領域について決定されたDNAのそれぞれの量を補正し、この配列決定結果を考慮することが望ましい場合がある。1 Mbのゲノム領域を用いる場合、それぞれの染色体は、100 kbのピンに最初に分割することができ、局所重み付け散布図スムージング(locally weighted scatterplot smoothing) (LOESS)を行い、配列タグの数のGC関連バイアスを補正することができる[20]。その後、100 kbのピンは、1 Mbのゲノム領域に融合することができ、その結果、GC補正密度は、すべてのその後の算出に用いることができる。

【0030】

一実施形態において、ゲノム領域（例えば、1 Mbのピン）におけるDNAの密度は、ゲノム表現( $GR_{x,y}$ )（式中、 $x$ 及び $y$ は、領域のゲノム座標の開始及び終点を示す。）として算出することができる。それぞれの領域から生じる配列リード（又はタグ）の数は、 $RC_{x,y}$ であり、 $GR_{x,y}$ は、次式を用いて算出される[20]。

【0031】

10

20

30

40

【数 1】

$$GR_{x-y} = \frac{RC_{x-y}}{RC_{total}}$$

【0032】

式中、 $RC_{total}$  は、リードの総数である。

【0033】

$RC_{total}$  で割ることは、ゲノム領域内の DNA のそれぞれの量を標準化してそれぞれの密度を求める例の一つである。方法の他の実施において、比は決定されず、また、 $RC_{x-y}$  の値は、ゲノム領域間で直接比較される。 $RC_{total}$  が試料において同じになるように制御される場合、これを行うことができる。

10

【0034】

B. 基準との比較

上述したように、ゲノム領域のそれぞれの密度を、基準密度と比較することができる。基準密度は、複数の種々の方法により、例えば、複数のゲノム領域上のそれぞれの密度を平均するか、又は複数の試料から得られる同じゲノム領域のそれぞれの密度を平均することにより求めることができる。ゲノム領域のそれぞれの密度が異常であるか否かを判断するために、パラメータは、それぞれの密度及び基準密度を用いて算出され、その後、パラメータはカットオフと比較される。パラメータの例としては、2つの値の差、差の絶対値、又は比率が挙げられる。これらの値は、適切に扱うことができ、例えば、スカラーを掛けてパラメータを求めることができ、カットオフとの有意な比較が可能になる。

20

【0035】

C. カットオフ

一部の実施形態において、それぞれの密度及び基準密度から算出したパラメータが、カットオフを超えているか否かについては、それぞれの密度が異常である（統計的に異なる）か否かを示し得る。パラメータの表示は、それぞれの密度が異常に高いか又は低いかなを示し、微小増幅又は微小欠失をそれぞれ示唆し得る（が、決定的ではない）。例えば、パラメータが、それぞれの密度と基準密度との間の単純な差である場合、異常に高く低いそれぞれの密度はそれぞれ、パラメータの正負の表示に相当する。したがって、カットオフは、正又は負であってもよく、また、パラメータがカットオフを超えることは、正のカットオフを超えるか、又は負のカットオフを下回ることを意味すると理解することができる。同様に、パラメータが比率である場合において、特定の値よりも大きいか、又はその値の逆数よりも小さいと、カットオフを超え得る。カットオフは、施術者によって所望されるように、選択することができる。例えば、これは任意であってもよいし、又は、異常な密度の判断が、所望レベルの統計的确实性によって行われることを保証するように選択されてもよい。

30

【0036】

一実施形態において、パラメータは、z 検定によりカットオフと比較される。z 検定は、統計的検定であり、分布の幅の点から、数が分布の平均からどれくらい離れているかの尺度である、z スコアを生成する。ここで、それぞれの密度と基準密度を比較するのに用いられるパラメータは、基準密度を算出するのに用いられる値の標準偏差で割ったこれらの密度の間の差である。

40

【0037】

デモンストレーションとして、正常な胎児核型を有する 8 人の妊婦（単生児）対象から生体試料を得て、それぞれの試料を用い、個々の 1 Mb ゲノム領域のそれぞれの密度を算出した。所定のゲノム領域  $x-y$  において、試料からのそれぞれの密度を平均し（すなわち、平均を算出し）、それぞれの密度  $mean GR_{x-y, reference}$  を求め、それぞれの密度の標準偏差  $SD_{x-y, reference}$  を算出した。その後、関心対象の別の妊婦対象が

50

ら試験試料を得て、試験試料のデータを用い、領域  $x$   $y$  のそれぞれの密度 ( $GR_{x-y}$   $_{test}$ ) を算出した。このそれぞれの密度の  $z$  スコア  $z$   $score_{GR_{x-y}}$  を次のように算出した。

【0038】

【数2】

$$z-score_{GR_{x-y}} = \frac{GR_{x-y_{test}} - meanGR_{x-y_{reference}}}{SD_{x-y_{reference}}}$$

10

【0039】

その後、 $z$  スコアを、カットオフ、例えば3と比較した。+3よりも大きいか又は-3よりも小さい $z$  スコアはそれぞれ、試験試料からのそれぞれの密度が、異常に高いか又は低いことを示していた。種々の染色体間の系統的な試料間の偏差を最小にするために、中央値の補正をそれぞれの染色体について行った。したがって、特定の染色体に相当するすべてのゲノム領域の中央ゲノム表現を、ベースラインとして用いた。特定の染色体上に位置するすべての領域については、このベースライン値との差を、 $z$  スコアの算出に用いた。

【0040】

$z$  スコア、又はゲノム領域における異常な密度のいずれかの他の尺度は、試料中の胎児DNAの豊富度に対し敏感であり得る。当該技術分野において知られているように、妊婦血液試料中の相当な割合の無細胞DNAは、胎児由来である。この割合は、妊娠及び他の要因の時間によって、1%未満から20%超の範囲であることが観察されている。本明細書に記載されている方法において、配列タグは、試料中の母体及び胎児DNAの両方から発生する。微小増幅又は微小欠失が、母体ではなく、胎児のゲノムに存在する場合、配列タグを発生させ密度を決定するのに用いられる少数のDNAが胎児であるので、比較的小さな異常が、対応するゲノム領域のDNA密度において観察される。逆に、微小増幅又は微小欠失が、母体由来(すなわち、母性遺伝)である場合、微小増幅又は微小欠失は、胎児と母親の両方のゲノム、したがって、試料中の実質的にすべてのDNAに存在する。したがって、DNAの密度の高い異常が観察される。(微小増幅又は微小欠失は、母体ゲノム中に存在する可能性は低い、胎児ゲノムでは低くない。)

20

30

【0041】

したがって、第2のカットオフは、ゲノム領域におけるDNAのそれぞれの密度の異常が、母性遺伝されているか否かを判断するのに用いることができる。第2のカットオフは、通常、胎児ゲノムの異常を識別するのに用いられるカットオフよりも大きく、基準密度からのそれぞれの密度の大きな又は統計的に有意な逸脱を必要とする。例えば、 $z$  検定を用いて、それぞれの密度と基準密度を比較する場合、母性遺伝の異常を識別するカットオフとして機能する $z$  スコアは、母体ゲノムには存在しない異常を識別するのに数回用いることができる。一部の実施形態において、第2のカットオフは、10、20又はそれ以上の $z$  スコアに相当する。

40

【0042】

母性遺伝した微小増幅又は微小欠失から生じるものなどのDNAの密度の高い異常は、低い異常よりも同じレベルの確実性によって識別するのに、より少ないDNA及び少ない配列タグを必要とする。したがって、以下に論じるように、胎児のゲノム異常を検出するための本明細書に開示されている方法の感度は、試料中の胎児DNAの割合(胎児%)によって異なる。異常なゲノム領域における配列タグの過小表現又は過剰表現の程度は、その領域の胎児DNAの割合と直線的に相関している[4]。一部の実施形態において、胎児DNAの割合は、次式を用いて算出される。

【0043】

## 【数 3】

$$\text{胎児\%} = \frac{(GR_{x-y_{\text{test}}} - \text{mean}GR_{x-y_{\text{reference}}}) \times 2}{\text{mean}GR_{x-y_{\text{reference}}}} \times 100\%$$

## 【0044】

## D. 基準密度

上述したように、特定のゲノム領域におけるDNAのそれぞれの密度は、基準密度と比較され、それぞれの密度が異常であるか否かが判断される。基準密度は、1つの試料又は複数の試料から得たデータを用いて種々の多くの方法により算出することができる。1つの試料の場合、基準密度は、例えば、ゲノム領域の集合のそれぞれの密度の平均とすることができる。この集合は、染色体の一部若しくは全部、複数の染色体又は全ゲノムに相当し得る。複数の試料の場合、それぞれの試料は、通常、異なる個体から得られる。ゲノム領域の基準密度は、その領域又は領域の集合のそれぞれの密度であり、個々において平均することができる。一部の実施形態において、生体試料は、微小増幅又は微小欠失の欠如を反映した基準密度を確立するために、公知の正常な胎児核型を有する妊婦対象から得られる。そして、基準密度は、未知の胎児核型を有する試験対象のそれぞれの密度と比較することができる。一部の実施形態において、胎児の複数の（又はすべての）核型の個々は未知であり、また、基準密度は、これらの個々の試料中に存在し得る異常についての予備知識なしで算出される。

## 【0045】

基準密度は、種々のゲノム領域において同じであっても又は異なってもよい。例えば、同じ値は、染色体上又はゲノム内の全ゲノム領域に用いてもよい。または、異なる値は、それぞれのゲノム領域の基準密度に割り当ててもよい。基準密度は、例えばそれぞれの密度を平均することによって、それぞれの密度から算出する必要はなく、例えば、ゲノム領域の相対的なサイズを反映するだけでよい。上述したように、基準密度は、種々の染色体又はゲノム領域の配列タグの不均一な発生などの結果を補正することができる。

## 【0046】

## E. 偽陽性の回避

ゲノム領域のそれぞれの密度が異常である（基準密度よりも静的に高いか又は低い）と判明した後、ゲノム領域が、少なくとも1つの他の異常なゲノム領域と連続している場合、微小増幅又は微小欠失を識別することができる。上述したように、行われるかかる識別については、それぞれの密度は、連続する領域に対して同じ方向の基準密度から外れる必要がある。したがって、微小増幅は、それぞれの密度がそれぞれの領域において異常に高い連続するゲノム領域に相当し、また、微小欠失は、それぞれの密度が異常に低い連続するゲノム領域に相当する。

## 【0047】

ゲノム領域の基準密度とそれぞれの密度を比較するのに用いる方法に依存して、実際に微小増幅又は微小欠失が存在しない場合、領域がDNAの異常な密度を有する、有意な確率が存在し得る。したがって、多くのゲノム領域が確立され、これらのそれぞれの密度が評価される場合、一部の領域は、単に偶然の異常と考えることができる。微小増幅又は微小欠失を識別する連続する異常領域は、偽陽性の可能性、すなわち、かかる事象を誤差と識別する可能性を低減する。また、この要件は、ゲノム領域のサイズの倍数（例えば、2又は3）で方法の解像限界も設定する。例えば、2以上の連続する領域が、微小増幅又は微小欠失を識別するのに必要とされる場合、1 Mbのゲノム領域の解像限界は2 Mbである。

## 【0048】

本明細書に記載の方法を用いて識別される微小増幅又は微小欠失は、当該技術分野において公知の他の方法を用いて検証することができる。かかる他の方法としては、例えば、

羊水穿刺及び臍帯穿刺が挙げられ、かかる他の方法は、侵襲性であるか、又は流産の危険性を含み得る。複数の方法を用いる異常なコピー数の検証は、偽陽性の可能性を更に低減することができる。

【0049】

本明細書では、2以上のゲノム領域が、染色体に沿って、又は、ゲノム参照配列において、連続する配列位置を占める場合、これらのゲノム領域は連続すると考えられ、これらのゲノム領域の間には、他のゲノム領域が存在しない。一部の実施形態において、連続する領域の対が順番に互いに直接隣接し、このため、隣接していると考えられる。2つの領域（例えば、少数の塩基対）間の配列ギャップは、これらの領域が隣接することを妨げるが、ギャップが本明細書に開示されている方法下における別のゲノム領域として扱われない場合には、これらの領域が連続することを妨げない。一部の実施形態において、ゲノム領域はオーバーラップしていないが、オーバーラップ領域を代わりに用いることができ、連続させることができる。施術者は、ゲノムの特定の領域に集中し、データ分析を簡素化するために、又は他の理由により、隣接していないか又はオーバーラップするゲノム領域を確立することを希望し得る。

10

【0050】

IV. シミュレーション解析

微小欠失又は微小重複検出の感度及び特異性は、試料中の胎児%、配列決定及び配列させた血漿DNA分子の数、並びに異常のサイズを含む種々のパラメータによる影響を受ける可能性がある。したがって、コンピュータシミュレーション解析は、1) 既存の配列決定の深度によって微小欠失/微小重複サイズ（例えば、~3 Mbのサイズ）の検出感度と、2) 特定の胎児%（例えば、5%）における特定の感度（例えば、95%/99%）を達成するように解析するのに必要な分子数を決定するために行うことができる。

20

【0051】

それぞれのシミュレーション解析では、全ゲノム（3,000 Mb）は、所望の解像度、例えば3 Mbに従って、同じサイズのピンに分割することができる。一部の実施形態において、サブ染色体異常の検出のために、3つの連続するピンは、同じ方向に対照群の平均から3を超える標準偏差のゲノム表現（過剰表現又は過少表現のいずれか）を有することが必要である。したがって、ピンのサイズは、所望の診断解像度の1/3と等しい。例えば、1 Mbのピンサイズが、3 Mbの異常を検出するのに必要である。微小欠失/微小重複によってカバーされる3つのピンは、少数の胎児DNAの原因となる異常なゲノム表現を有すると仮定できる。血漿中の、影響を受けた領域内のピンの範囲内に入る総分子の予測割合（E）は、次のように算出することができる。

30

【0052】

【数4】

$$E = \left(1 + \frac{f}{2} \times d\right) \times \frac{1}{T}$$

【0053】

40

式中、fは、血漿中の胎児DNAの割合であり、

dは、異常の染色体数の変化であり（dは、微小欠失において-1、及び微小重複において+1と等しい。）、

Tは、全ゲノムのピンの総数である。

【0054】

シミュレーション、例えば、1,000の正常な事例、及び1,000の影響を受けた事例は、上で算出されたように、予測血漿表現による血漿DNAの二項分布を仮定して行うことができる。胎児%、ピンサイズ、及び解析される分子の総数は、所望の目的を達成するように変更することができる。シミュレーションは、R ([www.r-project.org/](http://www.r-project.org/)) の `rbino` 関数を用いて行うことができる。

50

## 【実施例】

## 【0055】

## V. 実施例

本実施例は、ヒト胎児及び母体DNA中の微小増幅及び微小欠失を識別するために提供される。

## 【0056】

## A. データ分析のフレームワーク

イルミナHiSeq 2000シーケンサにおけるフローセルの1つのレーンを用いて、6つの試験事例及び8つの対照の母体血漿試料をそれぞれ分析した。平均2億1100万（範囲：1億7700万～2億3600万）のDNA断片を、それぞれの血漿DNA試料から配列決定した。かかる配列決定によって、平均1億4400万（範囲：9600万～1億8000万）の配列と、半数体ヒトゲノムの4.81倍に相当する事例ごとの非重複配列決定リードが生じた。

10

## 【0057】

血漿核型を得るために、全ゲノムを2,687の1 Mbピンに分割した。試験試料のそれぞれの1 Mbのピンについてのゲノム表現を、対照群と比較した。正常なゲノム表現を有する領域では、すべての1 Mbのピンのzスコアの予想分布は、ゼロ付近であった。基準区間を、+3から-3のzスコアと定義した。かかる基準区間によって、ピンのおよそ0.3%が、統計的に、この区間外と偶然なった。2,687のピンを分析したところ、平均で、8つのピンが偶然基準区間外になると予想された。したがって、偽陽性とするのを低減するために、追加基準は、3つの連続する1 Mbのピンが、基準区間外の及び同じ方向にxスコアを示した場合にのみ、コピー数異常とすることが含まれていた。

20

## 【0058】

## B. サブ染色体のコピー数異常の検出

それぞれの事例の全ゲノムにおける1 Mbのピンのzスコアを、Circosプロット[21]を用いてプロットした(図2)。試験試料では、94.9%～98.7%の1 Mbのピンが正常な表現を示した。3つの連続するピンが同じ異常を示した場合にのみコピー数異常とする上述の基準によって、偽陽性のないすべての事例において、コピー数異常が正確に識別された。

30

## 【0059】

図3及び4は、それぞれの事例におけるコピー数異常を示す染色体のすべての1 Mbのピンのzスコアを示している。事例01、02及び03では、過少表現が、染色体22のq腕上の3つの連続する1 Mbのピンにおいて検出された。これらは、デノボ22q11.2微小欠失を有する3つの事例であった。事例04及び05では、過剰表現は、染色体22q上の3つの連続する1 Mbのピンにおいて検出された。事例04は、2.4 Mbのデノボ22q11.2微小重複を有する事例であった。事例05は、同じ領域において母性遺伝微小重複を有する事例であった。事例05では、母親自身が微小重複を保有していたので、異常を母体血漿中で容易に検出することができた。このことは、3つの連続するピンの非常に高いzスコア値(範囲、39.7～71.7)によって裏付けられていた。胎児の無侵襲的出生前検査の更なる調査は、SNPに基づく方法、すなわち、相対的突然変異用量又は相対的ハプロタイプ用量分析[10、11、22]の使用により、進めることができた。事例06では、5つの連続する1 Mbのピンが、染色体3のq腕上の過剰表現によって検出され、31の連続する1 Mbのピンが、染色体4のq腕上の過少表現によって検出され、これらは、3q上の5 Mbの重複、及び4q上の31 Mbの欠失に相当していた。すべての事例では、検出されたコピー数異常は、アレイCGH、FISH及び/又はQF-PCRによって確認されたものと同等のサイズを有していた。事例05では、母親が保有していた微小重複が、アレイCGHによって確認された。場合06では、母親が保有していた均衡型転座が、完全核型分析によって確認された。

40

## 【0060】

50

### C . 胎児DNAの割合

過少表現又は過剰表現を示す領域のDNA配列を用いて、母体血漿中の胎児%を推定した(図5)。このアプローチは、この方法を用いて、男性胎児を保有する3つの事例(すなわち、事例02、03及び04)では染色体Yに基づく方法[4]を用いて算出された胎児%を比較することによって検証された。胎児%の値は、2つの方法の間で非常に一致していた(図5)。胎児のデノボコピー数異常を有する5つの事例では、胎児%は、9.2%から17.8%の範囲であった。事例05では、微小重複のゲノム表現によって推定された胎児%は、96.7%であり、このことは、循環DNAのほとんどすべてが、この変化を有することを示唆していた。この結果は、母親が異常を保有しているという事実と一致している。

10

#### 【0061】

### D . 診断感度のシミュレーション解析

コンピュータシミュレーションを、ショットガンMPS系無侵襲的出生前分子核型分析(shotgun MPS-based noninvasive prenatal molecular karyotyping)の診断感度を決定するために用いた(図6)。~1億5000万のリードの既存の配列決定深度によって、胎児%が5%である場合、3 Mbの染色体異常を検出する診断感度は、およそ96%であった。胎児%が6%に達すると、感度は99%まで増加する。小さいサイズの染色体異常を検出するために、多くの血漿DNA分子を分析する必要があった。図7は、3つの連続するピンの基準を用いて、95%/99%の感度とともに、3 Mb、2 Mb及び1 Mbの診断解像度を達成するのに必要な血漿DNA分子の数を示している。95%の診断感度を達成するために、それぞれのピンは、およそ42,000の分子を分析する必要があった。したがって、95%の診断感度で2 Mb及び1 Mbの微小欠失/微小重複を検出するのに分析される必要がある血漿DNA分子の総数はそれぞれ、1億9200万及び3億8000万であった。99%の診断感度を達成するために、分析される必要がある分子の総数は、2つの異なる解像度でそれぞれ、2億4000万及び4億8000万であった。

20

#### 【0062】

### E . 考察

本研究では、胎児の染色体微小欠失及び微小重複の無侵襲的出生前検出についての実行可能性は、ゲノムワイドレベルにおいて、3 Mbの解像度で実証された。5つの事例において、染色体3q、4q若しくは22qに關与する胎児由来のサブ染色体の欠失又は重複が検出された。6番目の事例では、非常に高いzスコアによって証明されるように、染色体22qの母体由来微小重複が検出された。実際に、事例04及び06は、胎児の微小重複が、母体血漿から無侵襲的に検出されたことを初めて示す。これらの結果は、1つ又は少数のゲノム領域におけるコピー数異常の試験に主に焦点を置いている、Petersら[8]及びJensenら[9]による先の報告と比較して重要な一步を示している。本明細書に示されているデータは、ゲノムワイド規模でサブ染色体コピー数異常を検出するのに、換言すれば、胎児分子核型を得るのにショットガンMPSを用いることができることを明確に実証している。

30

#### 【0063】

3つの研究事例において、母体血漿試料を侵襲的手順の後に採取した。これらの事例における胎児DNAの割合は、9.2~17.4%であり、この割合は、侵襲的手順の前に採取した試料についてのChiuら[4]によって先に観察された範囲内である。同様に、ほとんどの試験試料が妊娠20週目以降に採取されたが、これらの事例の胎児DNAの割合も、妊娠初期に採取した試料と大部分が重複している。それにもかかわらず、将来の予想される、妊娠第1及び第2期初期での侵襲的手順の前に採取した試料を用いた大規模施設研究において、これらの結果を検証するのに有用であった。

40

#### 【0064】

分析的に、診断アルゴリズムは、一部の実施形態において、サブ染色体のコピー数異常を検出するために、全zスコアが+3を超えるか又は-3未満の3つの連続するピンを必

50

要とする。このアルゴリズムは、およそ3 Mbの連続ストレッチにわたって検出可能なコピー数異常を必要とする。実際に、データによって示されているように、このアルゴリズムは、2.4 Mbのコピー数異常を検出することができた(事例04及び05)。

#### 【0065】

かかる診断解像度を達成するのに行われる配列決定の深度は、トリソミー試験に必要なものよりもはるかに高かった。したがって、それぞれの事例において、配列決定を、イルミナHiSeq 2000シーケンサの1つのレーンにより行い、トリソミー試験の少なくとも1つの商用供給者によって行われるものと同じ配列決定プラットフォームを用いる12plexショットガン配列決定と比較した。最新の配列決定の深度と、その結果生じる3 Mbの診断解像度において、最新のプロトコルは、公知の病原性コピー数変異体[23]のおよそ20%を網羅することができた。2億4000万及び4億8000万血漿DNA分子が、99%の感度で、2 Mb及び1 Mbの診断解像度を拡張するように、配列決定し配列させる必要があることが上で予測されていた。これらの診断解像度では、母体血漿DNAのショットガンMPSは、公知の病原性コピー数変異体[23]のそれぞれおよそ50%及び80%を網羅すると予想された。超並列シーケンサの処理能力の継続的な増加、及び配列決定コストの同時低減によって、かかる配列決定深度に関連するコストは、数年後に、医療提供者に許容可能なレベルに達する可能性がある。このアプローチによって必要とされる配列決定の量は、試料ごとに何十億もの配列決定リードを用いて行われる、先に報告されている胎児由来単一ヌクレオチド変異検出法[10]に対し、既に大幅に低減されている。コストの更なる低減は、母体血漿から胎児由来単一ヌクレオチド変異検出のために達成されているもの[24, 25]と同様に、病原性コピー数変異体を保有するゲノム領域の標的配列決定によって達成することができた。そして、単一分子配列決定の出現も、配列決定される分子のゲノム表現を歪め得る、必要とされない増幅過程[26]として、このアプローチの診断精度を更に向上させると予想された。

10

20

#### 【0066】

要約すると、母体血漿DNAのショットガンMPSによって無侵襲的出生前分子核型を得ることが可能なことが実証される。この方法は、胎児のデノボコピー数の変化、不均衡型転座及び母体コピー数の変化を検出することができる。これらの結果は、無侵襲的出生前診断の診断範囲を更に拡大した。結論として、母体血漿DNAのMPS分析に基づく方法は、全染色体の異数性の出生前検出[3-7]、サブ染色体のコピー数変化、及び単一遺伝子障害の胎児の突然変異[10]のために開発された。無侵襲的試験このアレイは、最初の事例において、胎児ゲノム及び染色体異常のスクリーニングに適用することができた。無侵襲的母体血漿DNA試験によって明らかにされた異常は、従来の侵襲的出生前検査によって更に確認することができた。大規模なプロスペクティブ研究による検証において、無侵襲的母体血漿DNA配列決定は、胎児ゲノム及び染色体異常の大きな範囲の出生前評価を提供し、安全な出生前評価を提供できることが想定される。

30

#### 【0067】

##### F. 材料及び方法

倫理的声明。本研究は、香港共同中文大学(the Joint Chinese University of Hong Kong) 病院当局新界東部クラスタ臨床研究倫理委員会(Hospital Authority New Territories East Cluster Clinical Research Ethics Committee)によって承認された。妊婦は、香港のプリンス・オブ・ウェールズ病院(the Prince of Wales Hospital)、クォンワウ病院(the Kwong Wah Hospital)、及びトゥサンユク病院(the Tsan Yuk Hospital)、並びにソウルのアサン医療センター(the Asan Medical Center)からの書面によるインフォームドコンセントによって募集した。

40

#### 【0068】

試料採取。事例01、02及び03では、母体末梢血試料を、侵襲的手順の後に、EDTAを収容したチューブに集めた(表1)。事例04、05及び06では、母体末梢血試料を、いずれかの侵襲的手順を行う前に採取した。母体血液試料を、妊娠12-3/7週から28-4/7週において採取した(図8)。

50

## 【0069】

6つの試験試料のうち、3つの胎児デノボ22q11.2微小欠失の事例(事例01、02及び03)、1つの胎児のデノボ22q11.2微小重複(2.4 Mb)事例(事例04)、及び1つの母性遺伝22q11.2微小重複(2.4 Mb)の事例(事例05)であった。また、母親がt(3;4)(q29;q32)の均衡型転座を有する1つの事例(事例06)もあり、胎児は、3q29微小重複及び4q32.1 q35.2欠失を有することが分かった。完全な核型分析を行い、さらに、胎児核型を、アレイ比較ゲノムハイブリダイゼーション(アレイCGH)[16]、蛍光インサイチュハイブリダイゼーション(FISH)、又は定量的蛍光PCR(QF-PCR)とFISHの組合せによって確認した。

10

## 【0070】

試料の処理及びDNA抽出。末梢血試料を、4℃で10分間、1600gで遠心分離し、また、血漿部分を、4℃で10分間、16000gで遠心した[17]。無細胞DNAを、先に記載されているように[3]、QIAamp DSP DNA血液ミニキット(キアゲン社(Qiagen)製)を用いて、1.8~8.4 mLの母体血漿から抽出した。抽出された血漿DNAを、先に記載されているように[18]、レプチン(LEP)遺伝子を標的とするリアルタイムPCRアッセイにより定量した。

## 【0071】

血漿DNA配列決定。血漿DNAのシーケン斯拉イブラリを、先に記載されているように[19]、ペアエンド配列決定試料調製キット(イルミナ社製)によって調製した。抽出された血漿DNAの13~20 ngをライブラリの調製に用いた。アダプター結合血漿DNAを、12サイクルPCRによって濃縮した。クラスタ生成を、TruSeq PEクラスタ発生キットv3(イルミナ社製)を備えたcBotクローン増幅システム(イルミナ社製)において行った。それぞれのライブラリ(試験試料と参照試料の両方)を、50 bp x 2のペアエンド形式のHiSeq 2000配列決定システム(イルミナ社製)において、フローセルの1つのレーンによって配列決定した。

20

## 【0072】

配列アラインメント及びフィルタリング。ショート・オリゴヌクレオチド・アライメント・プログラム2(SOAP2)(<http://http://soap.genomics.org.cn/>)を用いて、ペアエンドリードを、非反復マスクヒト参照ゲノム(NCBI構築36.1/hg18)に配列させた。最大2つのヌクレオチドのミスマッチが、ペアエンドリードのそれぞれのメンバーにおいて許容された。両端を有するペアエンドリードのみを、正確な方向で同じ染色体に配列させ、600 bp以下の挿入サイズを下流分析に含めた。また、ヒトゲノムにおいて同一の開始位置及び終了位置を示すペアエンドリードとして定義される重複リードを除去した。

30

## 【0073】

## G. 実施例の要約

胎児DNAは、妊婦の血漿中に存在する。母体血漿DNAの超並列配列決定を用いて、無侵襲的に、胎児トリソミー21、18、13、及び選択性染色体異数性が検出された。超並列配列決定を用いて、母体血漿から胎児の微小欠失の検出について記載された事例報告がされている。しかしながら、これらの先の報告は、多型依存性であるか、又は1つ若しくは少数のゲノムの選択部分に限定された統計分析が用いられている。本実施例において、手順は、母体血漿DNAの超並列配列決定により、全ゲノムにおいて3 Mbの解像度で無侵襲的出生前核型分析を行うことが報告された。この方法を用いて、6つの事例から得られた血漿を分析した。5つの事例では、胎児の微小重複又は微小欠失が、母体血漿から成功裡に検出された。胎児の微小重複を有する2つの事例は、母体血漿からのかかる変化の最初の無侵襲的出生前検出を示した。他の事例では、血漿DNA配列決定の結果は、妊娠した母親が、微小重複のキャリアであることが一致した。無侵襲的出生前核型分析の診断解像度を2 Mb及び1 Mbまで向上させるために配列決定し配列させる必要がある、血漿DNA分子の数を決定するのにシミュレーション解析を行った。結論として、超

40

50

並列配列決定により母体血漿から無侵襲的出生前分子核型分析は実施可能であり、無侵襲的出生前検査の診断範囲を高めた。

【0074】

VI. コンピュータシステム

本明細書に言及されているコンピュータシステムはいずれも、いずれかの適切な数のサブシステムを利用することができる。かかるサブシステムの例を、コンピュータ機器900により図9に示している。一部の実施形態において、コンピュータシステムには、サブシステムがコンピュータ機器の構成要素であり得る、単一のコンピュータ機器が含まれる。他の実施形態において、コンピュータシステムには、それぞれサブシステムである複数のコンピュータ機器と、内部要素とが含まれ得る。

10

【0075】

図9に示されているサブシステムは、システムバス975を介して相互連結されている。アダプター982を表示するように連結されている、プリンタ974、キーボード978、記憶装置979、モニター976などの追加のサブシステム、及び他のものが示されている。入力/出力(I/O)コントローラ971に連結されている、周辺機器及びI/O装置は、シリアルポート977などのいずれかの数の当該技術分野において公知の手段によって、コンピュータシステムに接続することができる。例えば、シリアルポート977、外部インターフェース981(例えば、イーサネット(Ethernet)(登録商標)、Wi-Fiなど)を用いて、コンピュータシステム900を、インターネットなどのワイドエリアネットワーク、マウス入力装置又はスキャナに接続することができる。システムバス975を介する相互接続によって、中央処理装置973は、それぞれのサブシステムと通信し、システムメモリ972又は記憶装置979(例えば、ハードドライブ若しくは光ディスクなどの固定ディスク)からの命令の実行とともに、サブシステム間の情報交換も制御することができる。システムメモリ972及び/又は記憶装置979は、コンピュータ可読媒体を具現化することができる。本明細書に記載されているいずれかのデータは、一方の要素から他方の要素に出力することができ、また、ユーザに出力することができる。

20

【0076】

コンピュータシステムには、複数の同一構成要素又はサブシステムが含まれ、これらは、例えば、外部インターフェース981によって又は内部インターフェースによって互いに接続されている。一部の実施形態において、コンピュータシステム、サブシステム又は機器は、ネットワーク上で通信することができる。かかる例では、1つのコンピュータは、クライアント及び別のコンピュータをサーバと考えることができ、それぞれが同じコンピュータシステムの一部とすることができる。クライアント及びサーバにはそれぞれ、複数のシステム、サブシステム又は構成要素が含まれ得る。

30

【0077】

本発明のいずれかの実施形態は、ハードウェア(例えば、特定用途向け集積回路、若しくはフィールド・プログラマブル・ゲート・アレイ)、及び/又は、一般にモジュールによってか若しくは一体的に一般にプログラム可能なプロセッサを備えたコンピュータソフトウェアを用いて、制御論理の形態で実施できることを理解する必要がある。本明細書において、ユーザとしてのプロセッサには、同一の集積チップ上にマルチコアプロセッサ、又は単一の回路基板若しくはネットワーク上に複数の処理ユニットが含まれる。本明細書に提供される開示及び教示に基づいて、当業者であれば、ハードウェア、及びハードウェアとソフトウェアの組合せを用いて本発明の実施形態を実施する他の手段及び/又は方法を理解及び認識する。

40

【0078】

本出願に記載されているソフトウェア要素又は機能のいずれかは、いずれかの適切なコンピュータ言語、例えば、Java(登録商標)、C++又はPerlを用いてプロセッサによって実行されるソフトウェアコードとして、例えば、従来技術又はオブジェクト指向技術を用いて実行することができる。ソフトウェアコードは、記憶及び/又は送信するためのコンピュータ可読媒体上の一連の命令又はコマンドとして記憶することができ、適

50

切な媒体には、ランダム・アクセス・メモリ（RAM）、読出し専用メモリ（ROM）、ハードドライブ若しくはフロッピー（登録商標）ディスクなどの磁気媒体、コンパクトディスク（CD）若しくはDVD（デジタル・バーサタイル・ディスク）などの光媒体、フラッシュメモリなどが含まれる。コンピュータ可読媒体は、かかる記憶装置又は送信装置のいずれかの組合せであってもよい。

【0079】

また、かかるプログラムは、コード化され、インターネットを含む、種々のプロトコルに一致する有線ネットワーク、光ネットワーク及び/又は無線ネットワークを介した送信に適したキャリア信号を用いて送信することもできる。このため、本発明の実施形態に係るコンピュータ可読媒体は、かかるプログラムによりコード化されたデータ信号を用いて生成することができる。コンピュータ可読媒体は、互換性のある装置とともにパッケージ化されるか、又は（例えば、インターネットダウンロードを介して）他の装置とは別に設けることができる。かかるコンピュータ可読媒体は、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、若しくはコンピュータシステム全体）上に又は内に存在し、また、システム若しくはネットワーク内の種々のコンピュータプログラム製品上に又は内に存在し得る。コンピュータシステムには、ユーザに本明細書に言及されている結果のいずれかを提供する、モニター、プリンタ又は他の適切なディスプレイが含まれ得る。

10

【0080】

本明細書に記載されている方法のいずれかは、ステップを実行するように構成可能な1以上のプロセッサを含むコンピュータシステムによって、完全に又は部分的に行うことができる。したがって、実施形態は、潜在的に、それぞれのステップ又はそれぞれのステップ群を実行する種々の要素によって、本明細書に記載されているいずれかの方法のステップを実行するように構成されたコンピュータシステムに導くことができる。符号を付けたステップとして示しているが、本明細書による方法のステップは、同時に又は異なる順序で行うことができる。加えて、これらのステップの一部は、他の方法の他のステップの一部とともに用いることができる。また、ステップのすべて又は一部は任意であってもよい。さらに、いずれかの方法のいずれかのステップを、これらのステップを実行するモジュール、回路又は他の手段によって行うことができる。

20

【0081】

特定の実施形態の具体的な詳細は、本発明の実施形態の趣旨及び範囲から逸脱することなく、いずれかの適切な方法により組み合わせることができる。ただし、本発明の他の実施形態は、それぞれの個々の態様、又はこれらの個々の態様の特定の組合せに関連する、特定の実施形態に関するものであってもよい。

30

【0082】

本発明の例示的な実施形態についての上述した説明は、例示及び説明の目的で示されている。本発明を、記載されている正確な形態に徹底するか又は限定するものではなく、上述の教示に照らして、多くの変更及び変形が可能である。本発明の原理とその実用化について最良に説明し、これによって、他の当業者が、種々の実施形態により、及び意図される特定の用途に適した種々の変更により本発明を最良に利用することができるように、実施形態が選択され記載されている。

40

【0083】

「a」、「an」又は「the」の記載は、特に明記しない限り、「1以上」を意味するものとする。

【0084】

本明細書に言及されているすべての特許、特許出願、刊行物及び説明は、すべての目的上、その全体において参照により援用される。いずれも、従来技術であると認めるものではない。

【0085】

参考文献

50

1. Lo YMD, Corbetta N, Chamberlain PF, Rai V, Sargent IL, et al. (1997) Presence of fetal DNA in maternal plasma and serum. *Lancet* 350: 485-487.
2. Lo YMD, Chiu RWK (2012) Genomic analysis of fetal nucleic acids in maternal blood. *Annu Rev Genomics Hum Genet* 13: 285-306.
3. Chiu RWK, Chan KCA, Gao Y, Lau VYM, Zheng W, et al. (2008) Noninvasive prenatal diagnosis of fetal chromosomal aneuploidy by massively parallel genomic sequencing of DNA in maternal plasma. *Proc Natl Acad Sci U S A* 105: 20458-20463.
4. Chiu RWK, Akolekar R, Zheng YWL, Leung TY, Sun H, et al. (2011) Non-invasive prenatal assessment of trisomy 21 by multiplexed maternal plasma DNA sequencing: large scale validity study. *BMJ* 342: c7401. 10
5. Palomaki GE, Kloza EM, Lambert-Messerlian GM, Haddow JE, Neveux LM, et al. (2011) DNA sequencing of maternal plasma to detect Down syndrome: An international clinical validation study. *Genet Med* 13: 913-920.
6. Palomaki GE, Deciu C, Kloza EM, Lambert-Messerlian GM, Haddow JE, et al. (2012) DNA sequencing of maternal plasma reliably identifies trisomy 18 and trisomy 13 as well as Down syndrome: an international collaborative study. *Genet Med* 14: 296-305.
7. Bianchi DW, Platt LD, Goldberg JD, Abuhamad AZ, Sehnert AJ, et al. (2012) Genome-wide fetal aneuploidy detection by maternal plasma DNA sequencing. *Obstet Gynecol* 119: 890-901. 20
8. Peters D, Chu T, Yatsenko SA, Hendrix N, Hogge WA, et al. (2011) Noninvasive prenatal diagnosis of a fetal microdeletion syndrome. *N Engl J Med* 365: 1847-1848.
9. Jensen TJ, Dzakula Z, Deciu C, van den Boom D, Ehrich M (2012) Detection of microdeletion 22q11.2 in a fetus by next-generation sequencing of maternal plasma. *Clin Chem* 58: 1148-1151.
10. Lo YMD, Chan KCA, Sun H, Chen EZ, Jiang P, et al. (2010) Maternal plasma DNA sequencing reveals the genome-wide genetic and mutational profile of the fetus. *Sci Transl Med* 2: 61ra91.
11. Fan HC, Gu W, Wang J, Blumenfeld YJ, El-Sayed YY, et al. (2012) Non-invasive prenatal measurement of the fetal genome. *Nature* 487: 320-324. 30
12. Fan HC, Wang J, Potanina A, Quake SR (2011) Whole-genome molecular haplotyping of single cells. *Nat Biotechnol* 29: 51-57.
13. Peters BA, Kermani BG, Sparks AB, Alferov O, Hong P, et al. (2012) Accurate whole-genome sequencing and haplotyping from 10 to 20 human cells. *Nature* 487: 190-195.
14. Chan KCA, Jiang P, Zheng YWL, Liao GJW, Sun H, et al. (2012) Cancer genome scanning in plasma: detection of tumor-associated copy number aberrations, single-nucleotide variants, and tumoral heterogeneity by massively parallel sequencing. *Clin Chem* 59: 211-224. 40
15. Swanton C (2012) Plasma-derived tumor DNA analysis at whole-genome resolution. *Clin Chem*. 59: 6-8.
16. Leung TY, Vogel I, Lau TK, Chong W, Hyett JA, et al. (2011) Identification of submicroscopic chromosomal aberrations in fetuses with increased nuchal translucency and apparently normal karyotype. *Ultrasound Obstet Gynecol* 38: 314-319.
17. Chiu RWK, Poon LLM, Lau TK, Leung TN, Wong EM, et al. (2001) Effects of blood-processing protocols on fetal and total DNA quantification in maternal plasma. *Clin Chem* 47: 1607-1613.
18. Tsui NBY, Jiang P, Chow KCK, Su X, Leung TY, et al. (2012) High resolution size analysis of fetal DNA in the urine of pregnant women by paired-end massive 50

ly parallel sequencing. PLoS One 7: e48319.

19. Zheng YWL, Chan KCA, Sun H, Jiang P, Su X, et al. (2012) Nonhematopoietically derived DNA is shorter than hematopoietically derived DNA in plasma: a transplantation model. Clin Chem 58: 549-558.

20. Chen EZ, Chiu RWK, Sun H, Akolekar R, Chan KCA, et al. (2011) Noninvasive prenatal diagnosis of fetal trisomy 18 and trisomy 13 by maternal plasma DNA sequencing. PLoS One 6: e21791.

21. Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, et al. (2009) Circos: an information aesthetic for comparative genomics. Genome Res 19: 1639-1645.

22. Lun FMF, Tsui NBY, Chan KCA, Leung TY, Lau TK, et al. (2008) Noninvasive prenatal diagnosis of monogenic diseases by digital size selection and relative mutation dosage on DNA in maternal plasma. Proc Natl Acad Sci U S A 105: 19920-19925.

23. Wapner RJ, Martin CL, Levy B, Ballif BC, Eng CM, et al. (2012) Chromosomal microarray versus karyotyping for prenatal diagnosis. N Engl J Med 367: 2175-2184.

24. Liao GJW, Lun FMF, Zheng YWL, Chan KCA, Leung TY, et al. (2011) Targeted massively parallel sequencing of maternal plasma DNA permits efficient and unbiased detection of fetal alleles. Clin Chem 57: 92-101.

25. Lam KWG, Jiang P, Liao GJW, Chan KCA, Leung TY, et al. (2012) Noninvasive prenatal diagnosis of monogenic diseases by targeted massively parallel sequencing of maternal plasma: application to beta-thalassemia. Clin Chem 58: 1467-1475.

26. van den Oever JM, Balkassmi S, Verweij EJ, van Iterson M, Adama van Scheltema PN, et al. (2012) Single molecule sequencing of free DNA from maternal plasma for noninvasive trisomy 21 detection. Clin Chem 58: 699-706.

10

20

【 図 1 】

【 図 2 】



FIG. 1

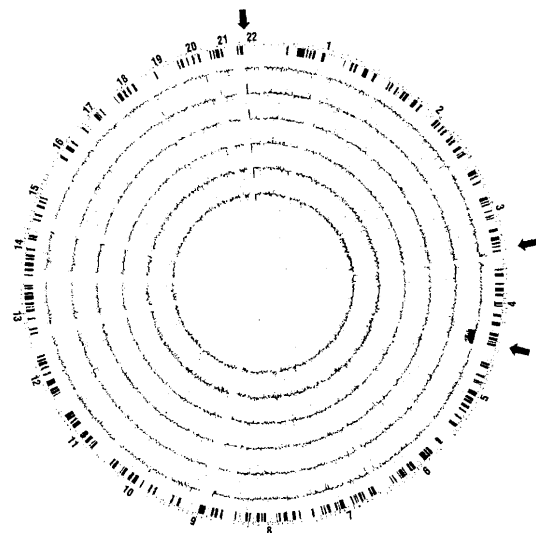
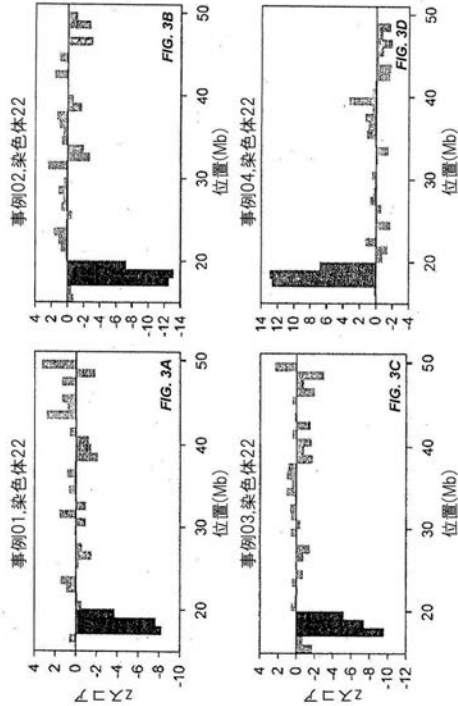


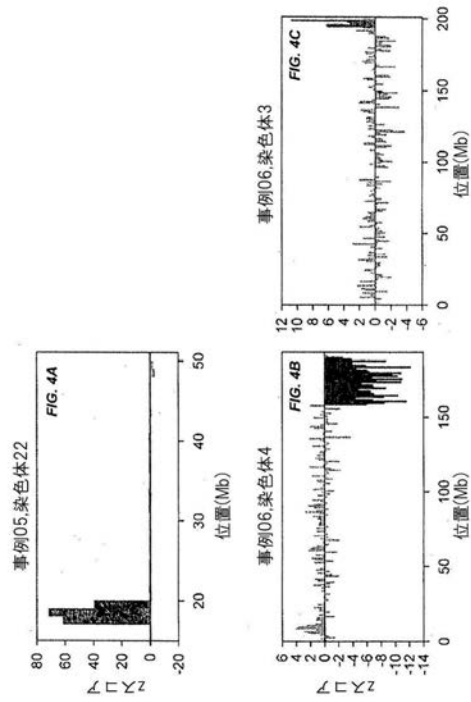
FIG. 2

100

【 図 3 】



【 図 4 】



【 図 5 】

事例	胎児の性別	胎児のDNAの割合	
		影響を受けた染色体領域のゲノム表現によるもの	染色体Y(chr Y)アプローチによるもの
01	F	10.5%	-
02	M	17.4%	21.5%
03	M	9.2%	13.7%
04	M	17.8%	20.3%
05 <sup>b</sup>	F	-	-
06	F	10.9% / 13.4% <sup>c</sup>	-

FIG. 5

【 図 7 】

診断感度		診断解像度		
		3 Mb	2 Mb	1 Mb
95%	それぞれのビンにおいて必要な分子数	42,000	42,000	42,000
	全ゲノムのビンの総数	3,000	4,500	9,000
	全ゲノムに必要な分子の総数	1億2500万	1億9200万	3億8000万
99%	それぞれのビンにおいて必要な分子数	53,000	53,000	53,000
	全ゲノムのビンの総数	3,000	4,500	9,000
	全ゲノムに必要な分子の総数	1億6000万	2億4000万	4億8000万

FIG. 7

【 図 6 】

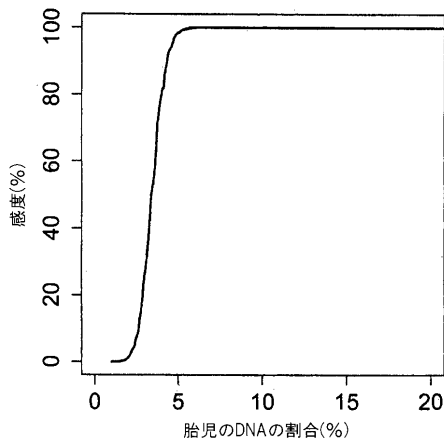


FIG. 6

【 図 8 】

事例番号	血漿採取時の妊娠期間(週)	侵襲的手順に関する血漿採取	侵襲の手順	染色体異常	核型を確認するのに用いられる方法
01	24 1/7	侵襲後	臍帯穿刺	22q11.2 微小欠失	FISH
02	28 4/7	侵襲後	臍帯穿刺	22q11.2 微小欠失	FISH
03	22 5/7	侵襲後	羊水穿刺 (Aminocentesis)	22q11.2 微小欠失	QF-PCR 及び FISH
04	12 3/7	侵襲前	絨毛膜絨毛採取	22q11.2 微小重複 (2.4Mb)	アレイ CGH
05	20 2/7	侵襲前	羊水穿刺	22q11.2 微小重複 (2.4Mb)	アレイ CGH
06	21 4/7	侵襲前	羊水穿刺	3q29 微小重複 (5.1Mb); 4q32.1-q35.2 微小欠失 (32.9Mb)	アレイ CGH

FIG. 8

【 図 9 】

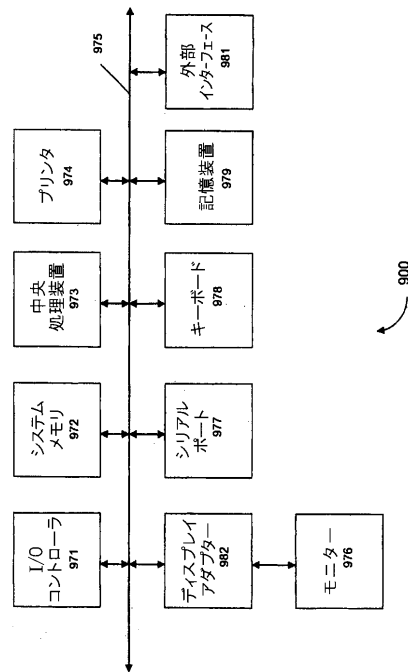


FIG. 9

【 手続補正書 】

【 提出日 】平成29年12月25日 (2017.12.25)

【 手続補正 1 】

【 補正対象書類名 】特許請求の範囲

【 補正対象項目名 】全文

【 補正方法 】変更

【 補正の内容 】

【 特許請求の範囲 】

【 請求項 1 】

胎児を宿した女性対象から得られる生体試料を分析することによって、前記胎児のゲノム中の微小増幅又は微小欠失を識別する方法であって、前記生体試料には、前記胎児及び前記女性対象からの無細胞DNAが含まれ、前記方法は、

特定の診断感度のために分析されるDNA断片の最小数、及び特定の決定に相当する微小増幅又は微小欠失のサイズを決定することと、

前記生体試料の複数のDNA断片のそれぞれについて1以上の配列タグを得ることと、ここで、複数のDNA断片は、特定の診断感度及び特定の決定のためのDNA断片の少なくとも最小数である、

前記配列タグのゲノム位置を決定することと、

複数のゲノム領域のそれぞれについて、コンピュータシステムによって、前記ゲノム領域内にゲノム位置を有する配列タグから、前記ゲノム領域内のDNA断片のそれぞれの量を決定することと、

前記それぞれの量を標準化して、それぞれの密度を求めることと、

該それぞれの密度と基準密度を比較して、前記それぞれの密度が、前記基準密度と統計的に異なるか否かを識別することと、

それぞれの密度が基準密度よりも統計的に高い場合に、ゲノム領域を微小増幅と同定す

ることと、

それぞれの密度が基準密度よりも統計的に低い場合に、ゲノム領域を微小欠失と同定すること、

とを含む、方法。

【請求項 2】

生体試料中の胎児 DNA の割合を測定し、  
DNA 断片の最小数を決定するために測定した胎児 DNA の割合を使用すること  
を更に含む、請求項 1 に記載の方法。

【請求項 3】

生物試料中の胎児 DNA の割合を推定し、及び  
DNA 断片の最小数を決定するために推定した胎児 DNA の割合を使用すること  
を更に含む、請求項 1 に記載の方法。

【請求項 4】

最小数がゲノム全体について特定される、請求項 1 ~ 3 のいずれか 1 項に記載の方法。

【請求項 5】

分析される DNA 断片の最小数が複数のゲノム領域のそれぞれのゲノム領域について DNA 断片数から計算される、請求項 1 ~ 3 のいずれか 1 項に記載の方法。

【請求項 6】

95% の診断感度について分析される DNA 断片の最小数が、複数のゲノム領域の各々のゲノム領域について、少なくとも 42,000 である、請求項 1 ~ 3 のいずれか 1 項に記載の方法。

【請求項 7】

99% の診断感度について分析される DNA 断片の最小数が、複数のゲノム領域の各々のゲノム領域について、少なくとも 53,000 である、請求項 5 に記載の方法。

【請求項 8】

それぞれのゲノム領域のサイズが約 1 Mb、約 2 Mb、又は約 3 Mb である、請求項 1 に記載の方法。

【請求項 9】

前記ゲノム領域が等しいサイズである、請求項 1 ~ 8 のいずれか 1 項に記載の方法。

【請求項 10】

請求項 1 ~ 9 のいずれか 1 項に記載の方法を実行するためのコンピュータ装置を制御する教示を含むコンピュータプログラム。

【請求項 11】

請求項 10 に記載のコンピュータプログラムを含むコンピュータ可読保存媒体。

【請求項 12】

請求項 10 に記載のコンピュータプログラムを含むコンピュータ装置。

【請求項 13】

プロセッサによって実行される場合に、請求項 1 ~ 9 のいずれか 1 項に記載の方法を実行するコンピュータプログラムコードを含むデータパケット。

## フロントページの続き

- (74)代理人 100134784  
弁理士 中村 和美
- (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, マー ロク パス 5  
2, ダブル ヘブン, ハウス 31
- (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, カウルーン, ホマンティン, キング タク 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, カウルーン, メイ フォー スゥン チュウエン, 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 チャン 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 62 - 66, クウオン ラム コート, ブロック ビー, フラット 7, 1/エフ
- (72)発明者 ユイ チュク イン ジャンディ  
中華人民共和国, 香港特別行政区, ニュー テリトリーズ, マー オン シャーン, ウー カイ  
シャア ロード 8, ダブル コーブ, ブロック 5, 20/エフ
- Fターム(参考) 4B063 QA08 QA13 QA19 QQ03 QQ42 QQ52 QR08 QR32 QR35 QS34  
QX02

专利名称(译)	母体血浆の无侵袭的出生前分子核型分析		
公开(公告)号	<a href="#">JP2018075013A</a>	公开(公告)日	2018-05-17
申请号	JP2017225951	申请日	2017-11-24
[标]申请(专利权)人(译)	香港中文大学		
申请(专利权)人(译)	香港大学中国		
[标]发明人	チウワイクン ロッサ ローユクミン デニス チャンクワン チー チャンペイヨン ユイチュクイン ジャンディ		
发明人	チウ ワイ クン ロッサ ロー ユク ミン デニス チャン クワン チー チャン ペイヨン ユイ チュク イン ジャンディ		
IPC分类号	C12Q1/68 G01N33/53 G06F19/22 C12N15/09		
CPC分类号	G16B20/00 C12Q1/6827 G16B20/10 G16B30/10		
FI分类号	C12Q1/68.Z G01N33/53.M G06F19/22 C12N15/00.A G16B30/00		
F-TERM分类号	4B063/QA08 4B063/QA13 4B063/QA19 4B063/QQ03 4B063/QQ42 4B063/QQ52 4B063/QR08 4B063/QR32 4B063/QR35 4B063/QS34 4B063/QX02		
代理人(译)	青木 笃 渡边洋一 武井良太郎 喀米·金加繆拉		
优先权	61/751213 2013-01-10 US 13/837776 2013-03-15 US		
外部链接	<a href="#">Espacenet</a>		

摘要(译)

要解决的问题：提供用于确定胎儿基因组中是否存在微量扩增或微缺失的方法，系统和装置。解决方案：本说明书中公开了用于检测胎儿基因组中的微放大或微缺失的方法，系统和装置。在一些实施方案中，该方法包括获得生物样品的多个DNA片段中的每一个的序列标签，确定序列标签的基因组位置，确定多个基因组区域中的每一个中的DNA密度是否异常高或者低，将具有异常高密度的一组连续基因组区域鉴定为微量扩增，将具有异常低密度的一组连续基因组区域鉴定为微缺失。生物样品可以从具有胎儿的女性受试者非侵入性地获得的血液样品。图1：图1

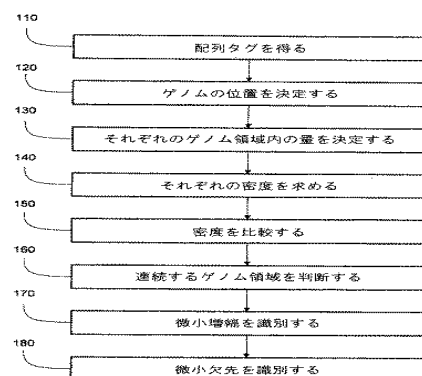


FIG. 1