



República Federativa do Brasil  
Ministério da Economia  
Instituto Nacional da Propriedade Industrial

**(11) BR 112012033760-2 B1**



**(22) Data do Depósito: 29/06/2011**

**(45) Data de Concessão: 17/11/2020**

---

**(54) Título:** DETECÇÃO NÃO INVASIVA DE ANORMALIDADES GENÉTICAS FETAIS

**(51) Int.Cl.:** C12Q 1/68.

**(73) Titular(es):** BGI HEALTH SERVICE CO., LTD.

**(72) Inventor(es):** FUMAN JIANG; HUIFEI CHEN; XIANGHUA CHAI; YUNYING YUAN; XIUQING ZHANG; FANG CHEN.

**(86) Pedido PCT:** PCT CN2011001070 de 29/06/2011

**(87) Publicação PCT:** WO 2013/000100 de 03/01/2013

**(85) Data do Início da Fase Nacional:** 28/12/2012

**(57) Resumo:** DETECÇÃO NÃO INVASIVA DE ANORMALIDADE GENÉTICAS FETAIS A presente invenção é direcionada a métodos para detecção não invasiva de anormalidades genéticas fetais através de sequenciamento em larga escala de nucleotídeos da amostra biológica materna. Métodos também são fornecidos para remover a tendência GC dos resultados do sequenciamento devido a diferença no conteúdo GC de um cromossoma. A presente invenção não somente realizada a detecção com mais precisão, como também representa um método para detecção de aneuploidia fetal incluindo distúrbios dos cromossomos sexuais, tais como X0, XXX, XXY e XYY, etc.

## “DETECÇÃO NÃO INVASIVA DE ANORMALIDADES GENÉTICAS FETAIS”

### Campo técnico

A invenção está relacionada a métodos não invasivos para a detecção de anormalidades genéticas fetais através da sequência de DNA de amostras de mulheres grávidas.

5 Mais particularmente, essa invenção está relacionada à análise de dados para obtenção da amostras de DNA. Essa invenção também está relacionada à análise estatística para detectar anormalidades genéticas fetais, tais como anormalidades cromossômicas, incluindo aneuploidia.

### Antecedentes da invenção

10 Métodos de diagnóstico pré-natal convencionais com procedimentos invasivos, tais como amostra do vilo coriônico e amniocentese, oferecem riscos potenciais tanto para a mãe como para o feto. Pesquisas não invasivas de aneuploidia fetal, utilizando marcadores do soro materno e ultrassom, estão disponíveis, mas apresentam sensibilidade e especificidade limitadas (Kagan, et al., *Human Reproduction* (2008) 23: 1968-1975; Malone, et al., *N Engl J Med* (2005) 353:2001-2011).

Estudos recentes demonstraram que a detecção não invasiva da aneuploidia fetal através do sequenciamento paralelo das moléculas de DNA no plasma de mulheres grávidas é viável. O DNA fetal foi detectado e quantificado no plasma e soro materno. (Lo, et al., *Lancet* (1997) 350:485-487; Lo, et al., *Am. J. hum. Genet.* (1998) 62:768-775). Tipos de células fetais múltiplas ocorrem na circulação materna, incluindo granulócitos fetais, linfócitos, hemácias nucleadas e células do trofoblasto (Pertl and Bianchi, *Obstetrics and Gynecology* (2001) 98:483-490). O DNA fetal pode ser detectado no soro na 17ª semana de gestação e aumenta com o término da gravidez. O DNA fetal presente no soro materno e no plasma é comparável à concentração do DNA obtido de protocolos de isolamento de células fetais.

25 O DNA fetal circulante foi utilizado para determinar o sexo do feto (Lo, et al., *Am. J. hum. Genet.* (1998) 62:768-775). O genótipo Rh D fetal foi detectado utilizando DNA fetal. Entretanto, o diagnóstico e as aplicações clínicas do DNA fetal circulante estão limitados aos genes presentes no feto e não na mãe (Pertl e Bianchi, *Obstetrics and Gynecology* (2001) 98:483-490). Logo, existe uma necessidade de um método não invasivo que possa determinar a sequência do DNA fetal, e oferecer um diagnóstico definitivo de anormalidades cromossômicas em um feto.

A descoberta de células e ácidos nucleicos fetais no sangue materno nas últimas décadas e a aplicação de sequenciamento aleatório de alto desempenho do DNA livre de células do plasma materno tornou possível detectar pequenas alterações na representação de cromossomos produzidas por um feto aneuploide em uma amostra do plasma materno. A detecção não invasiva da trissomia do cromossomo 13, 18 e 21 pode ser obtida.

Entretanto, como alguns estudos demonstram, a tendência GC introduzida pela

amplificação e sequenciamento estabeleceu um limite prático na sensibilidade da detecção da aneuploidia. A tendência GC pode ser introduzida durante a preparação da amostra e do processo de sequenciamento, sob condições diferentes, tais como composição do reagente, densidade e temperatura do cluster, produzindo uma amostragem diferencial de moléculas de DNA com composição de GC diferente e variação significativa nos dados de seqüenciamento, que podem ser ricos ou pobres em GC.

Para melhorar a sensibilidade, foram desenvolvidos protocolos para remoção do efeito da tendência GC. Fan e Quake desenvolveram um método para isolar a tendência GC estatisticamente utilizando peso para cada densidade GC, baseado no conteúdo genômico do GC local, para melhorar o número de leituras mapeadas em cada bin, multiplicando o peso correspondente ((Fan e Quake *PLoS ONE* (2010) 5:e10439). Entretanto, o método apresentou dificuldade com relação aos distúrbios dos cromossomos sexuais, especialmente os relevantes relacionados ao cromossomo Y, em virtude da possibilidade do processo causar uma leve distorção dos dados, podendo interferir na precisão da detecção.

Na presente invenção, é descrito um método para isolar estatisticamente a tendência GC para obter uma maior sensibilidade na detecção da anormalidade genética fetal, assim como evitar uma distorção dos dados. Esse método define parâmetros utilizados para testes estatísticos de acordo com o conteúdo de GC. Além disso, foi introduzida a fração fetal estimada no diagnóstico por hipótese binária, apresentando maior sensibilidade e especificidade. O método da presente invenção também demonstrou que é possível aumentar a sensibilidade da detecção não invasiva de anormalidade genética fetal definindo com precisão a amostra materna contendo uma pequena fração de DNA fetal, através do sequenciamento de mais fragmentos de polinucleotídeos. Uma nova amostra do plasma materno nas últimas semanas de gestação pode aumentar a sensibilidade do diagnóstico.

#### Resumo da invenção

A presente invenção é direcionada a métodos de detecção não invasiva de anormalidades genéticas fetais através do sequenciamento em larga escala dos nucleotídeos da amostra biológica materna. Métodos também são fornecidos para isolar a tendência GC dos resultados do sequenciamento devido a diferentes conteúdos GC de um cromossomo.

Logo, em um aspecto, é fornecido um método para estabelecer uma relação entre a profundidade da cobertura e o conteúdo de um cromossomo, cujo método compreende: obtenção da informação da sequência de fragmentos de polinucleotídeos múltiplos que compõem o citado cromossomo de mais de uma amostra; atribuição dos citados fragmentos aos cromossomos, baseado na informação da citada sequência; cálculo da profundidade da cobertura e do conteúdo de GC do citado cromossomo, de acordo com a informação da citada sequência para cada amostra; e determinação da relação entre a profundidade da cobertura e o conteúdo GC do citado cromossomo.

Em uma modalidade, os fragmentos de polinucleotídeos variam de cerca de 10 a cerca de 100 bp de comprimento. Em outra modalidade, os fragmentos de polinucleotídeos variam de cerca de 15 a cerca de 500 bp de comprimento. Ainda em outra modalidade, os fragmentos de polinucleotídeos variam de cerca de 20 a cerca de 200 bp de comprimento.

5 Ainda em outra modalidade, os fragmentos de polinucleotídeos variam de cerca de 25 a cerca de 100 bp de comprimento. Em outra modalidade, os fragmentos de polinucleotídeos apresentam cerca de 35 bp de comprimento.

Em uma modalidade, a informação da sequência é obtida pelo sequenciamento genômico paralelo. Em outra modalidade, a determinação do fragmento em relação aos cromossomos é feita por comparação da sequência dos fragmentos com uma sequência genômica humana de referência. A sequência genômica humana de referência pode ser qualquer configuração do genoma humano publicado e/ou adequado, tal como hg18 ou hg19. Os fragmentos que correspondem a mais de um cromossomo, ou não correspondem a qualquer cromossomo podem ser desconsiderados.

15 Em uma modalidade, a profundidade da cobertura de um cromossomo é a proporção entre o número de fragmentos que correspondem ao cromossomo, e o seu número de leituras de referência. Em outra modalidade, a profundidade da cobertura é normalizada. Ainda em outra modalidade, a normalização é calculada em relação à cobertura de todos os outros autossomos. Em outra modalidade, a normalização é calculada em relação à cobertura de todos os outros cromossomos.

Em uma modalidade, a relação está na fórmula:

$$cr_{i,j} = f(GC_{i,j}) + \varepsilon_{i,j}, j = 1, 2, \dots, 22, X, Y$$

onde  $f(GC_{ij})$  representa a função da relação entre a profundidade da cobertura normalizada e o conteúdo GC correspondente da amostra  $i$ , cromossomo  $j$ ,  $\varepsilon_{i,j}$  representa o resíduo da amostra  $i$ , cromossomo  $j$ . Em algumas modalidades, a relação entre profundidade da cobertura e o conteúdo GC é calculada pela regressão polinomial ponderada local. Em algumas modalidades, a relação pode ser uma relação linear não resistente. Em algumas modalidades, a relação é determinada por algoritmo Loess.

Em algumas modalidades, o método também compreende o cálculo ajustado da profundidade da cobertura, de acordo com a fórmula:

$$\hat{cr}_{i,j} = f(GC_{i,j}), j = 1, 2, \dots, 22, X, Y$$

Em algumas modalidades, o método também compreende o cálculo da variação padrão de acordo com a fórmula:

$$std_j = \sqrt{\sum_i (cr_{i,j} - \hat{cr}_{i,j})^2 / (ns - 1)}, j = 1, 2, \dots, 22, X, Y$$

35 onde  $ns$  representa o número de amostras de referência.

Em algumas modalidades, o método também compreende o cálculo do teste estatístico t de student, de acordo com a fórmula:

$$t_{i,j} = (cr_{i,j} - \hat{cr}_{i,j}) / std_j, j = 1, 2, \dots, 22, X, Y$$

Em uma modalidade, o conteúdo GC de um cromossomo é o conteúdo GC médio de todos os fragmentos que pertencem ao cromossomo. O conteúdo GC de um fragmento pode ser calculado dividindo o número de nucleotídeos G/C no fragmento pelo número total de nucleotídeos do fragmento. Em outra modalidade, o conteúdo GC de um cromossomo é o conteúdo GC agregado das leituras de referência do cromossomo.

Em algumas modalidades, no mínimo 2, 5, 10, 20, 50, 100, 200, 500 ou 1000 amostras são utilizadas. Em algumas modalidades, o cromossomo é o 1, 2, ..., 22, X ou Y.

Em uma modalidade, as amostras são de mulheres grávidas. Em outras modalidades, as amostras são de homens. Ainda em outra modalidade, as amostras são tanto de mulheres como de homens.

Em algumas modalidades, as amostras são amostras biológicas. Em algumas modalidades, são provenientes do sangue periférico.

É fornecido também um método para determinar uma anormalidade genética fetal, cujo método compreende: a) obtenção de informação da sequência de fragmentos de polinucleotídeos múltiplos de uma amostra; b) correlação dos citados fragmentos aos cromossomos baseados na citada informação da sequência; c) cálculo da profundidade da cobertura e conteúdo GC de um cromossomo baseado na citada informação da sequência; d) cálculo da profundidade da cobertura ajustada do citado cromossomo, utilizando o conteúdo GC do citado cromossomo, e uma relação estabelecida entre a profundidade da cobertura e o conteúdo GC do citado cromossomo; e e) comparação da citada profundidade da cobertura ajustada com a profundidade da cobertura do citado cromossomo, onde uma diferença entre eles indica uma anormalidade genética fetal.

Em algumas modalidades, o método também compreende a etapa f) determinação do sexo fetal. O sexo fetal pode ser determinado de acordo com a fórmula:

$$\text{logit}(p_i) = \ln\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 cr.a_{i,x} + \beta_2 cr.a_{i,y}$$

onde  $cr.a_{i,x}$  e  $cr.a_{i,y}$  são cobertura relativa normalizada dos cromossomos X e Y, respectivamente.

Em algumas modalidades, o método também compreende a etapa g) estimativa da fração fetal. A fração fetal pode ser calculada de acordo com a fórmula:

$$fy_i = (cr_{i,y} - \hat{cr}_{i,yf}) / (\hat{cr}_{i,ym} - \hat{cr}_{i,yf})$$

onde  $\hat{cr}_{i,yf} = f(GC_{i,yf})$  é a profundidade ajustada da cobertura calculada a partir da relação da profundidade do cromossomo Y e conteúdo GC correspondente de amostras de

mulheres grávidas com um feto do sexo feminino,  $\hat{c}r_{i,Xf} = f(GC_{i,Xf})$  se refere à profundidade de cobertura calculada a partir da relação da profundidade do cromossomo Y e o conteúdo GC correspondente de homens. De forma alternada, a fração fetal pode ser calculada de acordo com a fórmula:

$$5 \quad f\hat{x}_i = (cr_{i,X} - \hat{c}r_{i,Xf}) / (\hat{c}r_{i,Xm} - \hat{c}r_{i,Xf})$$

onde  $\hat{c}r_{i,Xf} = f(GC_{i,Xf})$  é a profundidade ajustada de cobertura calculada a partir da relação da profundidade do cromossomo X e conteúdo GC correspondente de amostras de mulheres grávidas com um feto do sexo feminino,  $\hat{c}r_{i,Xf} = f(GC_{i,Xf})$  se refere à profundidade de cobertura calculada a partir da relação da profundidade do cromossomo X e o conteúdo GC correspondente de homens. Logo, a fração fetal pode ser calculada de acordo com a fórmula:

$$10 \quad fxy_i = \arg \min_{\varepsilon \in (0,1)} \left( \frac{(\hat{c}r_{i,Xf} \cdot (1-\varepsilon) + \hat{c}r_{i,Xm} \cdot \varepsilon - cr_{i,X})^2}{(\hat{\sigma}_{X,f} \cdot (1-\varepsilon))^2 + (\hat{\sigma}_{X,m} \cdot \varepsilon)^2} + \frac{(\hat{c}r_{i,Yf} \cdot (1-\varepsilon) + \hat{c}r_{i,Ym} \cdot \varepsilon - cr_{i,X})^2}{(\hat{\sigma}_{Y,f} \cdot (1-\varepsilon))^2 + (\hat{\sigma}_{Y,m} \cdot \varepsilon)^2} \right)$$

onde  $\hat{c}r_{i,Xf} = f(GC_{i,Xf})$  é a profundidade ajustada de cobertura calculada a partir da relação da profundidade do cromossomo X e do conteúdo GC correspondente de amostras de mulheres grávidas com um feto do sexo feminino,  $\hat{c}r_{i,Xf} = f(GC_{i,Xf})$  se refere à profundidade de cobertura calculada a partir da relação da profundidade do cromossomo Y e o conteúdo GC correspondente de amostras de mulheres grávidas com um feto do sexo feminino,  $\hat{c}r_{i,Xm} = f(GC_{i,Xm})$  se refere à profundidade da cobertura ajustada calculada a partir da relação da profundidade do cromossomo X e conteúdo GC correspondente de amostras de homens,  $\hat{c}r_{i,Ym} = f(GC_{i,Ym})$  se refere à profundidade da cobertura calculado a partir da relação da profundidade do cromossomo Y e conteúdo GC correspondente de homens.

Em uma modalidade, a anormalidade genética é uma anormalidade cromossômica. Em outra modalidade, a anormalidade genética é a aneuploidia. Ainda em outra modalidade, a aneuploidia fetal é um distúrbio de um autossoma selecionado do grupo que consiste da trissomia do cromossomo 13, 18 e 21. Ainda em outra modalidade, a aneuploidia fetal é um distúrbio para um cromossomo sexual selecionado do grupo que consiste de XO, XXX, XXY e XYY.

Em algumas modalidades, a comparação da citada profundidade de cobertura ajustada com a citada profundidade do cromossomo é conduzida por um teste estatístico de hipóteses, onde uma hipótese é que o feto é euplóide (H0) e a outra hipótese é que o feto é aneuplóide (H1). Uma estatística pode ser calculada para ambas as hipóteses. Em algumas modalidades, a estatística t de student é calculada para H0 e H1 de acordo com a fórmula:

$t1_{i,j} = (cr_{i,j} - \hat{cr}_{i,j}) / std_j$  e  $t2_{i,j} = (cr_{i,j} - \hat{cr}_{i,j}(1 + fxy_{i,j} / 2)) / std_j$ , respectivamente, onde  $fxy$  é a fração fetal. Em algumas modalidades a proporção da probabilidade logarítmica de t1 e t2 é calculada de acordo com a fórmula:  $L_{i,j} = \log(p(t1_{i,j}, degree | D)) / \log(p(t2_{i,j}, degree | T))$ , onde nível (degree) se refere à distribuição t, D se refere à diploidia, T se refere à trissomia, e  $p(t1_{i,j}, degree | *)$ ,  $*=D, T$  representa densidade de probabilidade condicional fornecida a um nível de distribuição t.

Em uma modalidade, o sexo do feto é feminino, e o teste estatístico t de student é calculado de acordo com a fórmula:  $t1_{i,x} = (cr_{i,x} - \hat{cr}_{i,xf}) / std_{xf}$ , onde  $\hat{cr}_{i,xf} = f(GC_{i,xf})$  é a profundidade de cobertura ajustada calculada a partir da relação da profundidade de cobertura do cromossomo X e o conteúdo GC correspondente de amostras de mulheres grávidas com um feto do sexo feminino. Em algumas modalidades,  $|t1| > 3,13$  indica que o feto pode ser XXX ou X0. Em algumas modalidades,  $|t1| > 5$  indica que o feto é XXX ou X0.

Em outra modalidade, o sexo do feto é masculino, e o teste estatístico t de student é calculado de acordo com a fórmula:  $t2_i = (cr_{i,x} - (1 - fy_i / 2) \cdot \hat{cr}_{i,xf}) / std_{xf}$  onde  $\hat{cr}_{i,xf} = f(GC_{i,xf})$  é a profundidade de cobertura ajustada calculada a partir da relação da profundidade de cobertura do cromossomo X e o conteúdo GC correspondente de amostras de mulheres grávidas com um feto do sexo feminino. Em algumas modalidades,  $|t2| > 3,13$  indica que o feto pode ser XXY ou XYY. Em algumas modalidades,  $|t2| > 5$  indica que o feto é XXY ou XYY.

É fornecido também um método para determinar uma anormalidade genética fetal, cujo método compreende: a) obtenção de informação da sequência de fragmentos de polinucleotídeos múltiplos de um cromossomo de mais de uma amostra normal; b) correlação dos citados fragmentos aos cromossomos baseados na citada informação da sequência; c) cálculo da profundidade da cobertura e conteúdo GC do citado cromossomo baseado na citada informação de sequência das citadas amostras normais; d) determinação da relação entre a profundidade de cobertura e conteúdo GC do citado cromossomo; e) obtenção da informação da sequência de fragmentos de polinucleotídeos múltiplos de uma amostra biológica; f) correlação dos citados fragmentos aos cromossomos, baseada na citada informação da sequência da citada amostra biológica; g) cálculo da profundidade da cobertura e o conteúdo GC do citado cromossomo, baseado na informação da citada sequência das citadas amostras biológicas; h) cálculo da profundidade de cobertura ajustada do citado cromossomo utilizando o citado conteúdo GC do citado cromossomo e a citada relação entre a profundidade da cobertura e o conteúdo GC do citado cromossomo; e i) comparação da citada profundidade da cobertura ajustada com a profundidade da cobertura do citado cromossomo, onde uma diferença entre eles indica uma anormalidade genética fetal.

Em outro aspecto, é fornecido na presente invenção um meio legível por computador compreendendo uma pluralidade de instruções para realizar um diagnóstico pré-natal de uma anormalidade genética fetal, que compreende as etapas de: a) recebimento da informação da sequência de fragmentos de polinucleotídeos múltiplos de uma amostra; b) correlação dos citados fragmentos de polinucleotídeos aos cromossomos baseada na citada informação da sequência; c) cálculo da profundidade de cobertura e conteúdo GC de um cromossomo baseado na citada informação da sequência; d) cálculo da profundidade de cobertura ajustada do citado cromossomo utilizando o citado conteúdo GC do citado cromossomo e uma relação estabilizada entre a profundidade de cobertura e conteúdo GC do citado cromossomo; e e) comparação da citada profundidade de cobertura ajustada em relação à citada profundidade de cobertura do citado cromossomo, onde uma diferença entre eles indica uma anormalidade genética.

Ainda em outro aspecto, é fornecido na presente invenção um sistema para determinação da anormalidade genética fetal, cujo método compreende: a) meio para obtenção da informação da sequência de fragmentos múltiplos de polinucleotídeos de uma amostra; e b) meio legível por computador compreendendo uma pluralidade de instruções para realizar um diagnóstico pré-natal de uma anormalidade genética fetal. Em algumas modalidades, o sistema também compreende uma amostra biológica obtida de uma mulher grávida, onde a amostra biológica inclui fragmentos de polinucleotídeos múltiplos.

#### Breve descrição das figuras

A figura 1 apresenta um esquema para cálculo da profundidade e conteúdo GC utilizando a informação sequencial de fragmentos de polinucleotídeos.

A figura 2 ilustra a correlação estabelecida entre profundidade de cobertura normalizada –conteúdo GC utilizando dados de 300 casos de referência. A profundidade de cobertura normalizada para cada caso é plotada em relação ao conteúdo GC sequenciado correspondente. As cruzes significam casos com fetos femininos euplóides, quadrados significam casos com fetos masculinos euplóides. A linha extensa é a linha de ajuste entre a profundidade de cobertura e o conteúdo GC.

A figura 3 ilustra a tendência entre a profundidade de cobertura normalizada e o conteúdo GC correspondente pelo arranjo de cromossomo com seu conteúdo GC inerente ascendente. O conteúdo GC ascendente inerente de cada cromossomo se refere ao conteúdo GC médio de etiquetas de sequências de cromossomos de 300 casos de referência.

A figura 4 demonstra composições diferentes da classe GC para cada cromossomo. O conteúdo GC de cada leitura de 35 bp das leituras de referência foi calculado para cada cromossomo, o conteúdo GC foi classificado em 36 níveis, e a porcentagem de cada nível foi calculada como a composição GC de cada cromossomo. Os cromossomos foram mapeados com calor, colocados em um gráfico e agrupados hierarquicamente.

A figura 5 demonstra a tendência do sequenciamento introduzindo a correlação demonstrada na figura 2 por simulação manual do processo de preferência.

A figura 6 demonstra a plotagem da variação padrão em relação ao número total de fragmentos de polinucleotídeos. Em 150 amostras, a variância padrão ajustada de cada cromossomo apresenta relação linear com a raiz quadrada do número de leituras únicas.

A figura 7 demonstra a plotagem Q-Q de resíduos de cada cromossomo calculados pela fórmula 3. Uma relação linear é demonstrada com uma distribuição normal.

A figura 8 demonstra o histograma da profundidade de cobertura do cromossomo Y. Existem dois picos que indicam que o gênero dos casos pode ser distinguido pela profundidade da cobertura do cromossomo Y. A curva é a distribuição da profundidade de cobertura do cromossomo Y calculada pela estimativa da densidade com Kernel gaussiano.

A figura 9 apresenta um diagrama do processo para diagnóstico de 903 amostras teste para anormalidade cromossômica.

A figura 10 demonstra o resultado da aneuploidia: casos de trissomia do 13, 18, 21 e XO, XXY, XYY e casos de normalidade. A figura 10A apresenta as marcações da profundidade de cobertura normalizadas em relação ao conteúdo GC dos cromossomos 13, 18 e 21. A figura 10B apresenta as marcações dos cromossomos X e Y. Os círculos representam as marcações da profundidade da cobertura com o conteúdo GC dos fetos femininos normais, e os pontos pontilhados representam os fetos masculinos normais. A linha contínua é a linha de ajuste da cobertura relativa e do conteúdo GC, as linhas tracejadas apresentam o valor t absoluto igual a 1, as linhas pontilhadas apresentam valor t igual a 2, e as linhas tracejadas com pontilhados: valor absoluto de t igual a 3.

A figura 11 compara o valor da confiança de diferentes tentativas de diagnóstico.

A figura 12 demonstra a relação entre a fração do DNA fetal e a idade gestacional. A fração do DNA fetal no plasma materno está relacionada com a idade gestacional. A fração do DNA fetal foi calculada pelos cromossomos X e Y. Existe uma correlação estatisticamente significativa entre a fração do DNA fetal média e a idade gestacional ( $P < 0,001$ ). Observar que o valor de  $R^2$  que representa o quadrado do coeficiente de correlação é pequeno. A fração mínima é de 3,49%.

A figura 13 apresenta a relação entre a variância padrão e o número de casos necessários para detecção. As variâncias padronizadas computadas pela fórmula 5 para cada cromossomo podem variar de acordo com números diferentes de amostras. A variância padrão se torna estável quando o número de amostras é maior do que 100.

A figura 14 demonstra o número estimado de leituras únicas para a detecção de aneuploidia fetal no plasma sem células como uma função da fração do DNA fetal. As estimativas são baseadas no nível de confiança do valor t maior do que 3 para aneuploidia dos cromossomos 13, 18, 21 e X, mesmo Y (da relação entre X e Y), cada um apresentando um

comprimento diferente. A medida que a fração do DNA fetal diminui, o número total de sequências aleatórias necessárias aumenta. Com um desempenho de sequenciamento de 4 milhões de leituras de sequência por canal na célula de fluxo, a trissomia do cromossomo 21 pode ser detectada se 3,5% do DNA livre de células for fetal. A aneuploidia do cromossomo X não é detectada facilmente quando a fração e o número de leituras únicas forem pequenas, tal como 4% e 5 milhões de leituras. Cromossomos diferentes exigem níveis diferentes da fração do DNA fetal e número de leituras único, o que pode ser causado pela estrutura GC do cromossomo.

5 A figura 15 apresenta um gráfico de contorno da sensibilidade mapeada por volume de dados e idade gestacional (semanas) para detecção da trissomia do cromossomo 13 de fetos femininos, para cada semana gestacional e cada ponto do volume de dados.

A figura 16 apresenta um gráfico de contorno da sensibilidade mapeada por volume de dados e idade gestacional (semanas) para detecção da trissomia do cromossomo 18 de fetos femininos, para cada semana gestacional e cada ponto do volume de dados.

15 A figura 17 apresenta um gráfico de contorno da sensibilidade mapeada por volume de dados e idade gestacional (semanas) para detecção da trissomia do cromossomo 21 de fetos femininos, para cada semana gestacional e cada ponto do volume de dados.

A figura 18 apresenta um gráfico de contorno da sensibilidade mapeada por volume de dados e idade gestacional (semanas) para detecção da trissomia do cromossomo X de fetos femininos, para cada semana gestacional e cada ponto do volume de dados.

20 A figura 19 apresenta um gráfico de contorno da sensibilidade mapeada por volume de dados e idade gestacional (semanas) para detecção da trissomia do cromossomo 13 de fetos masculinos. Para cada semana gestacional e cada ponto do volume de dados, foi computada sua distribuição empírica da fração de DNA fetal e variância padrão para o volume de dados fornecido inicialmente, e comparando a fração estimada por XY ou Y, foi computada a sensibilidade de cada tipo de aneuploidia.

A figura 20 apresenta um gráfico de contorno da sensibilidade mapeada por volume de dados e idade gestacional (semanas) para detecção da trissomia do cromossomo 18 de fetos masculinos.

30 A figura 21 apresenta um gráfico de contorno da sensibilidade mapeada por volume de dados e idade gestacional (semanas) para detecção da trissomia do cromossomo 21 de fetos masculinos.

#### Descrição detalhada da invenção

35 A presente invenção é direcionada a métodos de detecção não invasiva de anormalidades genéticas fetais através do sequenciamento em larga escala de fragmentos de polinucleotídeos a partir de amostras biológicas maternas. Métodos também são fornecidos para isolar a tendência GC dos resultados do sequenciamento em virtude da diferença no

conteúdo GC de um cromossomo, baseada na relação entre a profundidade de cobertura de um cromossomo e seu conteúdo GC. Desta forma, é fornecido na presente invenção um método para ajustar por computador os parâmetros de referência que são utilizados no cálculo do teste t de student com os conteúdos GC, por regressão polinomial local para ajustar a profundidade da cobertura de um cromossomo de cada amostra em relação ao conteúdo GC dos fragmentos de polinucleotídeos.

Um método também é fornecido para determinar a anormalidade genética de um feto por análise estatística, utilizando um teste estatístico hipotético. Além disso, métodos são fornecidos para calcular padrões do controle de qualidade (DQC) úteis na determinação da quantidade de amostras clínicas necessárias para um nível de significância estatística.

### Definições

A menos que definido ao contrário, todos os termos técnicos e científicos utilizados na presente invenção são compreendidos pelos versados na técnica a qual essa invenção pertence. Todas as patentes, aplicações, pedidos publicados e outras publicações mencionadas na presente invenção são incorporadas como referência em sua totalidade. Se uma definição descrita nessa seção for contrária ou de outra forma inconsistente com uma definição descrita nas patentes, pedidos, pedidos publicados e outras publicações que são incorporadas na presente invenção como referência, a definição descrita nessa seção prevalecerá sobre as outras.

Conforme utilizado na presente invenção, as formas no singular “um”, “uma”, e “a/o” incluem as formas no plural, a menos que indicado ao contrário. Por exemplo, “um” dímero inclui um ou mais dímeros.

O termo “anormalidade cromossômica” se refere a um desvio entre a estrutura do cromossomo do indivíduo e um cromossomo homólogo normal. O termo “normal” se refere ao cariótipo predominante ou padrão encontrado em indivíduos sadios de uma espécie em particular. Uma anormalidade cromossômica pode ser numérica ou estrutural, e inclui, mas não está limitada a aneuploidia, poliploidia, inversão, trissomia, monossomia, duplicação, deleção, deleção de uma parte de um cromossomo, adição, adição de uma parte de um cromossomo, inserção, um fragmento de um cromossomo, uma região de um cromossomo, um rearranjo cromossômico, e translocação. Uma anormalidade cromossômica pode estar correlacionada com a presença de uma condição patológica ou com uma predisposição para desenvolver uma condição patológica. Conforme definido na presente invenção, um polimorfismo único dos nucleotídeos (“SNP”) não é uma anormalidade cromossômica.

A monossomia do cromossomo X (X0, ausência de um cromossomo X) é o tipo mais comum da síndrome de Turner, ocorrendo em 1 entre 2500, e 1 entre 3000 meninas que nascem vivas (Sybert e McCauley *N Engl J Med* (2004) 351:1227-1238). A síndrome XXY é uma condição onde homens apresentam um cromossomo X extra, ocorrendo em 1

entre cada 1000 homens (Bock, Understanding Klinefelter Syndrome: A Guide for XXY Males and their families. NIH Pub. No. 93-3202 (1993)). A síndrome XYY é uma aneuploidia do cromossomo sexual onde o homem recebe um cromossomo Y extra, recebendo um total de 47 cromossomos no lugar dos 46 usuais, afetando 1 entre 1000 nascimentos, ocasionando a infertilidade masculina (Aksglæde, *et al.*, *J Clin Endocrinol Metab* (2008) 93: 169-176).

A síndrome de Turner abrange diversas condições, onde a monossomia do cromossomo X (X0, ausência de um cromossomo sexual, o corpo de Barr) é a mais comum. Mulheres normais apresentam os dois cromossomos X, mas na Síndrome de Turner, um destes cromossomos se perde. Ocorrendo em 1 entre 2000 para 1 em 5000 mulheres fenotípicas, a síndrome se manifesta de várias formas. A síndrome de Klinefelter é uma condição onde homens apresentam um cromossomo X a mais. Em humanos, a síndrome de Klinefelter é o distúrbio sexual cromossômico mais comum, e a segunda condição mais comum causada pela presença de cromossomos extras. A condição ocorre em 1 entre cada 1000 homens. A síndrome XYY é uma aneuploidia dos cromossomos sexuais onde um homem recebe um cromossomo Y extra, totalizando 47 cromossomos no lugar de 46. Isso produz um cariótipo 47, XYY. Essa condição é geralmente assintomática e afeta 1 entre 1000 homens podendo ocasionar infertilidade masculina.

A trissomia do 13 (síndrome de Patau), a trissomia do 18 (síndrome de Edward) e a trissomia do 21 (síndrome de Down) são as trissomias autossômicas clinicamente mais importantes, e detectá-las é sempre um aspecto importante. A detecção das aberrações cromossômicas fetais descritas acima é de grande importância no diagnóstico pré-natal (Ostler, Diseases of the eye and skin: a color atlas. Lippincott Williams & Wilkins. pp. 72. ISBN 9780781749992 (2004); Driscoll e Gross *N Engl J Med* (2009) 360: 2556-2562; Kagan, *et al.*, *Human Reproduction* (2008) 23: 1968-1975).

O termo “leitura de referência única” se refere a fragmentos de um cromossomo que apresentam uma sequência única. Logo, tais fragmentos podem ser atribuídos a uma localização única do cromossomo. Leituras de referência únicas de um cromossomo podem ser construídas baseadas em uma sequência genômica de referência publicada, tal como hg18 ou hg19.

Os termos “polinucleotídeos”, “oligonucleotídeos”, “ácido nucléico” e “molécula de ácido nucléico” são utilizados intercambialmente na presente invenção se referindo a uma forma polimérica de nucleotídeos de qualquer comprimento, e podem compreender ribonucleotídeos, desoxirribonucleotídeos, seus análogos, ou suas misturas. Esse termo se refere somente à estrutura primária da molécula. Logo, o termo inclui um ácido desoxirribonucléico de fita simples, dupla ou tripla (“DNA”), assim como um ácido ribonucléico de fita simples, dupla ou tripla (“RNA”). Pode incluir também os modificados, por exemplo, por alquilação, e/ou por encapsulamento, e formas não modificadas do polinucleotídeo. Mais particularmen-

te, os termos “polinucleotídeo”, “oligonucleotídeo”, “ácido nucléico” e “molécula de ácido nucléico” incluem polidesoxirribonucleotídeos (contendo 2-desoxi-D-ribose), polirribonucleotídeos (contendo D-ribose), incluindo tRNA, rRNA, hRNA e mRNA, com splicing (remoção de introns) ou não, qualquer outro tipo de polinucleotídeo, que é um N- ou C-glicosídeo de uma base purina ou pirimidina, e outros polímeros contendo bases nucleotídicas normais, por exemplo, poliamida (por exemplo, ácidos nucléicos peptídicos (“PNAs”)) e polímeros polimorfolinos (disponíveis comercialmente da Anti-Virals, Inc., Corvallis, OR., como NeuGene®), e outros polímeros de ácidos nucléicos com sequência sintética específica, com os polímeros apresentando nucleobases em uma configuração que permita o pareamento e empilhamento das bases, tal como é encontrada no DNA e RNA. Logo, estes termos incluem, por exemplo, 3’ – desoxi – 2’,5’ – DNA, N3’ a P5’ fosforamidatos de oligodesoxirribonucleotídeo, RNA 2’ – O – alquila substituído, híbrido entre DNA e RNA ou entre PNAs e DNA ou RNA, e também incluem tipos de modificações conhecidas, por exemplo, marcações, alquilação, “cápsulas”, substituição de um ou mais nucleotídeos por um análogo, modificações entre os nucleotídeos tais como, por exemplo, as com ligações sem carga (por exemplo, fosfonatos de metila, fosfotriésteres, fosforamidatos, carbamatos, etc.), com ligações carregadas negativamente (por exemplo, fosforotioatos, fosforoditioatos, etc.) e com ligações carregadas positivamente (por exemplo, aminoalquilfosforamidatos, aminoalquilfosfotriésteres), as contendo metades pendentes, tais como, por exemplo, proteínas (incluindo enzimas (por exemplo, nucleases), toxinas, anticorpos, peptídeos de sinal, poli-L-lisina, etc), as com agentes intercaladores (por exemplo, acridina, psoraleno, etc.), as contendo quelatos (por exemplo, de metais, metais radioativos, boro, metais oxidativos, etc.), as contendo agentes de alquilação, as com ligações modificadas (por exemplo, ácidos nucléicos alfa anoméricos, etc.), assim como as formas não modificadas do polinucleotídeo ou oligonucleotídeo.

“Sequenciamento paralelo maciço” significa técnicas para sequenciamento de milhões de fragmentos de ácidos nucléicos, por exemplo, utilizando a ligação de DNA genômico fragmentado aleatoriamente a uma superfície ópticamente transparente, planar, e amplificação de fase sólida para criar um sequenciamento de fluxo celular de alta densidade com milhões de grupos, cada um contendo ~ 1.000 cópias de modelo por  $\text{cm}^2$ . Esses modelos são sequenciados utilizando uma tecnologia de sequenciamento por síntese de DNA com quatro cores. Ver produtos oferecidos pela Illumina, Inc., San Diego, Califórnia. A sequência utilizada é de preferência a realizada sem uma etapa de pré-amplificação ou clonagem, mas pode ser combinada com métodos a base de amplificação em um chip microfluído, que apresenta câmaras de reação tanto para PCR como para sequenciamento com base de modelo microscópico. Somente cerca de 30 bp de informação de sequência randômica é necessária para identificar uma sequência como pertencente a um cromossomo humano

específico. Sequências mais longas podem identificar alvos mais particulares. Na presente invenção, um grande número de leituras de 35 bp foram obtidas. Outra descrição de um método de sequenciamento paralelo maciço é encontrada em Rogers e Ventner, *Nature* (2005) 437: 326-327.

5 Conforme utilizado na presente invenção, “amostra biológica” se refere a qualquer amostra obtida de uma fonte viral ou viva, ou de outra fonte de macromoléculas e biomoléculas, e inclui qualquer tipo celular ou tecido de um indivíduo na qual o ácido nucléico, proteína ou outra macromolécula pode ser obtida. A amostra biológica pode ser uma amostra obtida diretamente de uma fonte biológica ou de uma amostra que é processada. Por exemplo, ácidos nucléicos isolados que são amplificados constituem uma amostra biológica. Amostras biológicas incluem, mas não estão limitadas a fluidos corporais, tais como sangue, plasma, soro, fluido cerebrospinal, fluido sinovial, urina e suor, amostras teciduais e de órgãos de animais e plantas e amostras processadas derivadas.

15 É compreendido que aspectos e modalidades da invenção descritas incluem “consistindo” e/ou “consistindo essencialmente de” aspectos e modalidades.

Outros objetos, vantagens e aspectos da presente invenção se tornarão claras a partir das especificações descritas abaixo juntamente com as figuras em anexo.

## II. Estabelecendo uma relação entre profundidade de cobertura e conteúdo GC

20 É fornecido um método para estabelecer uma relação entre a profundidade de cobertura e o conteúdo GC de um cromossomo, compreendendo: obtenção da informação da sequência de fragmentos múltiplos de polinucleotídeos do citado cromossomo de mais de uma amostra; correlação dos citados fragmentos aos cromossomos baseados na citada informação da sequência; cálculo da profundidade de cobertura e conteúdo GC do citado cromossomo baseado na citada informação da sequência para cada amostra; e determinação da relação entre a profundidade de cobertura e o conteúdo GC do citado cromossomo. As etapas de operação podem ser realizadas em qualquer ordem. Em algumas modalidades, o método pode ser realizado na seguinte ordem: a) obtenção de informação da sequência de fragmentos de polinucleotídeos múltiplos que abrangem o citado cromossomo baseada na citada informação da sequência; b) correlação dos citados fragmentos aos cromossomos baseados na citada informação da sequência; c) cálculo da profundidade da cobertura e conteúdo GC de um cromossomo baseado na citada informação de sequência para cada amostra; e d) determinação da relação entre a profundidade de cobertura e o conteúdo GC do citado cromossomo.

35 Para calcular a profundidade de cobertura e o conteúdo GC de uma localização do cromossomo, a informação da sequência de fragmentos de polinucleotídeos é obtida pelo sequenciamento do modelo de DNA obtido de uma amostra. Em uma modalidade, o modelo de DNA contém tanto DNA materno como fetal. Em outra modalidade, o modelo de DNA é

obtido do sangue de uma mulher grávida. O sangue pode ser coletado utilizando qualquer técnica padrão para sua retirada incluindo, mas não limitada a venopunção. Por exemplo, o sangue pode ser retirado de uma veia da parte interna do cotovelo ou da parte dorsal da mão. As amostras podem ser coletadas de uma mulher grávida a qualquer momento durante a gestação. Por exemplo, amostras de sangue podem ser coletadas de mulheres com 1-4, 4-8, 8-12, 12-16, 16-20, 20-24, 24-28, 28-32, 32-36, 36-40 ou 40-44 semanas de gestação, de preferência entre 8-28 semanas de gestação.

Os fragmentos de polinucleotídeos estão relacionados a uma localização do cromossomo baseado na informação da sequência. Uma sequência genômica de referência é utilizada para obter as leituras únicas de referência. Conforme utilizado na presente invenção, o termo “leituras únicas de referência” se refere a todos os fragmentos de polinucleotídeos relacionados a um local genômico específico, baseado em uma sequência genômica de referência. Em algumas modalidades, as leituras únicas de referência apresentam o mesmo comprimento de, por exemplo, cerca de 10, 12, 15, 20, 25, 30, 35, 40, 50, 100, 200, 300, 500 ou 1000 bp. Em outras modalidades, as construções do genoma humano hg18 ou hg 19 podem ser utilizadas como a sequência genômica de referência. Uma localização do cromossomo pode ser uma janela contínua em um cromossomo que apresenta um comprimento de cerca de 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10.000 ou mais Kb. Uma localização do cromossomo pode ser um único cromossomo.

Conforme utilizado na presente invenção, o termo “profundidade de cobertura” se refere à proporção entre o número de fragmentos que está relacionado à localização do cromossomo, e o número de leituras únicas de referência da localização do cromossomo utilizando a seguinte fórmula:

$$C_{i,j} = n_{i,j} / N_j, j = 1, 2, \dots, 22, X, Y \quad (1)$$

onde  $n_{i,j}$  é o número de leituras de sequência únicas mapeadas em relação ao cromossomo  $j$  na amostra  $i$ ;  $C_{i,j}$  é a profundidade de cobertura no cromossomo  $j$  na amostra  $i$ ;  $N_j$  é o número de leituras de referência únicas no cromossomo  $j$ .

Em algumas modalidades, os fragmentos dos polinucleotídeos que não estão correlacionados a uma localização única do cromossomo ou a vários locais são descartados. Em algumas modalidades, a profundidade da cobertura é normalizada, baseado na profundidade da cobertura de outra localização do cromossomo, de outro cromossomo, da média de todos os outros autossomas, da média de todos os outros cromossomos, ou da média de todos os cromossomos. Em algumas modalidades, a profundidade de cobertura média de 22 autossomas é utilizada como uma constante de normalização para relacionar as diferenças no número total de leituras de sequência obtidas para amostras diferentes:

$$cr_{i,j} = C_{i,j} / \left( \sum_{j=1}^{22} C_{i,j} / 22 \right), j = 1, 2, \dots, 22, X, Y \quad (2)$$

onde  $cr_{i,j}$  representa a profundidade de cobertura relativa do cromossomo  $j$  na amostra  $i$ . A partir deste momento, o termo “profundidade da cobertura relativa” para cada cromossomo, se refere ao valor normalizado, e é utilizado para comparar amostras diferentes e análise subsequente.

O conteúdo GC de uma localização do cromossomo pode ser calculado pela porcentagem média de GC de uma localização do cromossomo, baseado nas leituras de referência únicas na localização do cromossomo, ou nos fragmentos de polinucleotídeos sequenciados que correspondem à localização do cromossomo. O conteúdo GC de um cromossomo pode ser calculado utilizando a seguinte fórmula:

$$GC_{i,j} = NGC_{i,j} / BASE_{i,j} \quad (3)$$

$$GC_{i,chrj} = \frac{\text{no. GC}_i}{\text{no. BASE}_i}$$

onde  $i$  representa amostra  $i$ ,  $j$  representa cromossomo  $j$ ,  $NGC_{i,j}$  representa o número de bases G e C do DNA e  $BASE_{i,j}$  representa o número de bases de DNA no cromossomo  $j$  na amostra  $i$ .

A profundidade de cobertura e o conteúdo GC podem ser baseados na informação da sequência de fragmentos de polinucleotídeos obtidos de uma única amostra ou de múltiplas amostras. Para estabelecer a relação entre a profundidade da cobertura e o conteúdo GC da localização de um cromossomo, o cálculo pode ser baseado na informação da sequência de fragmentos de polinucleotídeos obtidos de no mínimo 1, 2, 5, 10, 20, 50, 100, 200, 500 ou 1000 amostras.

Em algumas modalidades, a relação entre a profundidade de cobertura e o conteúdo GC é uma relação linear instável. O algoritmo Loess, ou a regressão polinomial ponderada local pode ser utilizada para avaliar as relações não lineares (correlações) entre os pares de valores, tais como profundidade de cobertura e conteúdo GC.

### 25 III. Determinando uma anormalidade genética fetal

É fornecido um método para determinar a anormalidade genética fetal, cujo método compreende: a) obtenção da informação da sequência de fragmentos múltiplos de polinucleotídeos de uma amostra; b) correlação dos citados fragmentos aos cromossomos baseados na citada informação da sequência; c) cálculo da profundidade da cobertura e conteúdo GC de um cromossomo baseado na citada informação de sequência para cada amostra; d) cálculo da profundidade de cobertura ajustada do citado cromossomo utilizando o citado conteúdo GC do citado cromossomo e uma relação estabelecida entre a profundidade de cobertura e o conteúdo GC do citado cromossomo; e e) comparação da citada profundidade de cobertura ajustada à citada profundidade de cobertura do citado cromossomo, onde uma

diferença entre elas indica uma anormalidade genética fetal.

Os métodos podem ser utilizados para detectar anormalidades cromossômicas fetais, e são especialmente úteis na detecção de aneuploidia, poliploidia, monossomia, trissomia, trissomia do 21, trissomia do 13, trissomia do 14, trissomia do 15, trissomia do 16, trissomia do 18, trissomia do 22, triploidia, tetraploidia, e anormalidades do cromossomo sexual incluindo X0, XXY, XYY e XXX. Podem também se concentrar em certas regiões dentro do genoma humano de acordo com os métodos da presente invenção para identificar monossomias e trissomias parciais. Por exemplo, os métodos podem envolver a análise de dados sequenciais em uma “janela” de uma porção cromossômica definida, tal como regiões de 50 kb contínuas, não sobrepostas espalhadas pelo cromossomo. Trissomias parciais de 13q, 8p (8p23,1), 7q, 6p distal, 5p, 3q (3q 25,1), 2q, 1q (1q42,1 e 1q21-qter), monossomia parcial Xpand 4q 35,1 foram relatadas, entre outras. Por exemplo, duplicações parciais do longo braço do cromossomo 18 podem causar a síndrome de Edwards, no caso da duplicação do 18q21,1-qter (Mewar, *et al.*, *Am J Hum Genet.* (1993) 53:1269-78).

Em algumas modalidades, a fração fetal é estimada baseada na informação obtida da sequência dos fragmentos de polinucleotídeos de uma amostra. A profundidade de cobertura e o conteúdo GC do cromossomo X e Y podem ser utilizados para calcular a fração fetal. Em algumas modalidades, o sexo do feto é determinado baseado na informação da sequência obtida para os fragmentos de polinucleotídeos de uma amostra. A profundidade da cobertura e o conteúdo GC do cromossomo X e Y podem ser utilizados para determinar o sexo do feto.

Em algumas modalidades, a comparação da citada profundidade de cobertura ajustada em relação à citada profundidade de cobertura do cromossomo é realizada através de um teste estatístico de hipóteses, onde uma hipótese é que o feto é euplóide (H0) e a outra hipótese é que o feto é aneuplóide (H1). Em algumas modalidades, o teste estatístico de student é calculado para ambas as hipóteses como t1 e t2, respectivamente. Em algumas modalidades, a proporção da probabilidade logarítmica de t1 e t2 é calculada. Em algumas modalidades, uma proporção da probabilidade logarítmica >1 indica trissomia do feto.

#### IV. Meio de leitura legível por computador e sistema para diagnóstico de uma anormalidade genética fetal

Em outro aspecto, é fornecido na presente invenção um meio de leitura legível por computador compreendendo uma pluralidade de instruções para realizar diagnóstico pré-natal de uma anormalidade genética fetal, que compreende as etapas de: a) recebimento da citada informação da sequência; b) correlação dos citados fragmentos aos cromossomos baseados na citada informação da sequência; c) cálculo da profundidade da cobertura e conteúdo GC de um cromossomo baseado na citada informação de sequência para cada amostra; d) cálculo da profundidade de cobertura ajustada do citado cromossomo utilizando

o citado conteúdo GC do citado cromossomo e uma relação estabelecida entre a profundidade de cobertura e o conteúdo GC do citado cromossomo; e e) comparação da citada profundidade de cobertura ajustada à citada profundidade de cobertura do citado cromossomo, onde uma diferença entre elas indica uma anormalidade genética fetal.

5 Ainda em outro aspecto, um sistema é fornecido na presente invenção para determinar a aneuploidia fetal, cujo método compreende: a) meio para obtenção de informação da sequência dos citados fragmentos de polinucleotídeos; e b) meio de leitura legível por computador compreendendo uma pluralidade de instruções para realizar um diagnóstico pré-natal de uma anormalidade genética fetal. Em algumas modalidades, o sistema também  
10 compreende uma amostra biológica obtida de uma mulher grávida, com fragmentos de polinucleotídeos.

Será compreendido pelos versados na técnica que um número de métodos diferentes de sequenciamento pode ser utilizado. Em uma modalidade, o sequenciamento é realizado utilizando o paralelo maciço. O sequenciamento paralelo maciço, tal como o obtido na  
15 plataforma 454 (Roche) (Margulies, *et al.*, *Nature* (2005) 437:376-380), no Analisador Genome Illumina (ou plataforma Solexa™) ou Sistema SOLiD (Applied Biosystems) ou tecnologia de sequenciamento de uma molécula de DNA Helicos True (Harris, *et al.*, *Science* (2008) 320:106-109), tecnologia de tempo real (SMRT™) de molécula simples, da Pacific Bioscience, e por sequenciamento nanoporo (Soni e Meller, *Clin Chem* (2007) 53: 1996-2001), per-  
20 mite o sequenciamento de muitas moléculas de ácido nucléico isolado de uma amostra com altos níveis de multiplexação na forma paralela (Dear, *Brief Funct Genomic Proteomic* (2003) 1:397-416). Cada uma destas sequências de plataformas se expande ou ainda moléculas simples não ampliadas de fragmentos de ácidos nucléicos. Equipamentos de sequenciamento disponíveis comercialmente podem ser utilizados na obtenção da informação da  
25 sequência dos fragmentos dos polinucleotídeos.

#### V. Exemplos

Os seguintes exemplos são fornecidos para ilustrar, e não limitar a invenção.

##### Exemplo 1. Análise dos fatores que afetam a sensibilidade da detecção: tendência GC e gênero

30 Um esquema do procedimento para calcular a profundidade de cobertura e conteúdo GC é ilustrado na figura 1. Foi utilizado um software para produzir as leituras de referência únicas através da quebra das sequências de referência hg18 em 1-mer (1-mer é uma leitura que é decomposta artificialmente da referência da sequência humana com o mesmo comprimento "1" das leituras de sequenciamento da amostra), e definindo essas unidades de  
35 repetição 1-mer como leituras de referência únicas. Segundo, mapeamos nossas leituras da amostra sequenciadas de acordo com as leituras de referência únicas de cada cromossomo. Terceiro, deletamos os valores atípicos (outlier) aplicando método de corte da quinta parte

para obtenção de uma configuração de dados claros. Finalmente, foi avaliada a contagem da profundidade de cobertura de cada cromossomo para cada amostra e o conteúdo GC das leituras de sequência única mapeadas para cada cromossomo de cada amostra.

Para pesquisar como o conteúdo GC afeta os dados, foram escolhidos 300 casos de cariótipos de euploidia, e os dados das leituras sequenciadas de suas profundidades de cobertura e dos conteúdos GC foram distribuídos em um gráfico, demonstrando uma intensa correlação entre eles, e esse fenômeno não foi relatado anteriormente (figura 2). Na figura 2, a profundidade de cobertura está intensamente relacionada ao conteúdo GC, e demonstra uma direção descendente de alguns cromossomos, tais como 4,13, etc., enquanto demonstra uma direção ascendente de outros cromossomos, tais como 19, 22, etc. Todos os cromossomos foram arranjados em ordem crescente de acordo com seu conteúdo GC, e uma tendência descendente está presente em cromossomos com conteúdo GC menor, enquanto outro grupo de cromossomos com maior conteúdo GC apresenta uma tendência ascendente, conforme demonstrado na figura 3. Pode ser interpretado que se os fragmentos de polinucleotídeos que estão sendo sequenciados para uma amostra apresentam um conteúdo GC maior do que a outra amostra, a profundidade de cobertura que representa esta amostra irá cair em comparação com a da outra amostra com conteúdo GC menor, enquanto aumenta nos cromossomos com conteúdo GC maior.

A possível explicação para tal tendência de alteração entre cromossomos com diferentes conteúdos de GC é as diferenças na composição do conteúdo GC nos diferentes cromossomos demonstrados na figura 4 combinados com a tendência CG introduzida no processo de sequenciamento. O conteúdo GC de cada leitura de referência única com 35mer para cada cromossomo é utilizado para classificar o conteúdo GC em 36 níveis. A porcentagem de cada nível como a composição GC de cada cromossomo foi calculada, e então utilizada para elaborar o mapa por aquecimento com o software Heatmap2. Utilizando o cromossomo 13 como exemplo, a sua maior parte consiste de segmentos de sequências com conteúdo GC menor, mas a sua parte menor consiste de segmentos de sequências com conteúdo GC maior. Se as condições durante o sequenciamento ou processo PCR forem favoráveis aos segmentos com um conteúdo GC maior, então uma parte maior do cromossomo 13 com baixo conteúdo de GC apresentará dificuldade no sequenciamento, ocasionando a redução da profundidade de cobertura no cromossomo 13 desta amostra. Em comparação, em um grupo com conteúdo GC maior, tal como o cromossomo 19, a profundidade da cobertura na amostra do cromossomo 19 se torna maior nessa parte com o maior conteúdo de GC, facilitando o sequenciamento. Não importa o cromossomo, os segmentos com conteúdo GC maior ou menor apresentam dificuldade de sequenciamento, mas a influência introduzida pela tendência GC foi diferente para os diversos cromossomos com composições GC diferentes. Cada cromossomo de referência foi dividido em bins de 1KB, e

o conteúdo GC de cada leitura de referência no bin foi calculado. O conteúdo GC de cada bin no intervalo próprio [0,3, 0,6], dividido pelo tamanho da etapa de 0,001, e o valor da cobertura relativa em cada intervalo é calculado. A figura 5 apresenta marcações da profundidade de cobertura e conteúdo GC para cada cromossomo.

5 A influência do sexo do feto nos dados foi analisada utilizando um teste t independente de duas amostras. Nenhuma diferença foi encontrada entre os autossomas exceto nos cromossomos sexuais com aproximadamente o mesmo conteúdo GC, mas existe uma diferença óbvia na UR% entre homens e mulheres (Chiu et al., (2008) *Proc Natl Acad Sci USA* 105: 20458-20463), sugerindo que não existe necessidade de distinguir o sexo do feto  
10 para detectar aneuploidia autossômica, mas existe a necessidade de distinguir o sexo inicialmente, para detectar aneuploidia do cromossomo sexual, tal como X0, XYY, etc.

#### Exemplo 2 – Modelo estatístico

Baseado no fenômeno discutido acima, foi utilizada regressão polinomial local para ajustar a relação entre a profundidade de cobertura e o conteúdo GC correspondente. A  
15 profundidade de cobertura consiste de uma função de GC e um residual da distribuição normal conforme abaixo:

$$cr_{i,j} = f(GC_{i,j}) + \varepsilon_{i,j}, j = 1, 2, \dots, 22, X, Y \quad (4)$$

onde  $f(GC_{i,j})$  representa a função para a relação entre a profundidade de cobertura e o conteúdo GC correspondente da amostra  $i$ , cromossomo  $j$ ,  $\varepsilon_{i,j}$  representa o residual de  
20 amostra  $i$ , cromossomo  $j$ .

Não existe uma relação linear forte entre a profundidade de cobertura e o conteúdo GC correspondente, logo, é aplicado algoritmo Loess para ajustar a profundidade de cobertura com o conteúdo GC correspondente, do qual é calculado um importante valor para o modelo da invenção, que é a profundidade de cobertura ajustada.

$$25 \quad \hat{cr}_{i,j} = f(GC_{i,j}), j = 1, 2, \dots, 22, X, Y \quad (5)$$

Com a profundidade de cobertura ajustada, a variância padrão e o teste t de student foram calculados de acordo com as fórmulas 6 e 7:

$$std_j = \sqrt{\sum_i (cr_{i,j} - \hat{cr}_{i,j})^2 / (ns - 1)}, j = 1, 2, \dots, 22, X, Y \quad (6)$$

$$t1_{i,j} = (cr_{i,j} - \hat{cr}_{i,j}) / std_j, j = 1, 2, \dots, 22, X, Y \quad (7)$$

#### Exemplo 3. Cálculo da fração fetal

30 Em virtude da fração fetal ser muito importante para a detecção da presente invenção, ela foi calculada antes do procedimento do teste. Conforme mencionado anteriormente, foram sequenciados 19 homens adultos, e quando suas profundidades de cobertura foram comparadas com casos de fetos femininos, foi descoberto que a profundidade de cobertura

do cromossomo X dos homens é quase metade das mulheres, e a profundidade de cobertura do cromossomo Y dos homens é quase 0,5 maior do que das mulheres. Então, podemos calcular a fração fetal dependendo da profundidade de cobertura do cromossomo X e Y conforme as fórmulas 8, 9 e 10, considerando a correlação GC:

$$5 \quad fy_i = (cr_{i,Y} - \hat{c}r_{i,Yf}) / (\hat{c}r_{i,Ym} - \hat{c}r_{i,Yf}) \quad (8)$$

$$fx_i = (cr_{i,X} - \hat{c}r_{i,Xf}) / (\hat{c}r_{i,Xm} - \hat{c}r_{i,Xf}) \quad (9)$$

$$fxy_i = \arg \min_{\varepsilon \in (0,1)} \left( \frac{(\hat{c}r_{i,Xf} \cdot (1-\varepsilon) + \hat{c}r_{i,Xm} \cdot \varepsilon - cr_{i,X})^2}{(\hat{\sigma}_{X,f} \cdot (1-\varepsilon))^2 + (\hat{\sigma}_{X,m} \cdot \varepsilon)^2} + \frac{(\hat{c}r_{i,Yf} \cdot (1-\varepsilon) + \hat{c}r_{i,Ym} \cdot \varepsilon - cr_{i,Y})^2}{(\hat{\sigma}_{Y,f} \cdot (1-\varepsilon))^2 + (\hat{\sigma}_{Y,m} \cdot \varepsilon)^2} \right) \quad (10)$$

onde  $\hat{c}r_{i,Xf} = f(GC_{i,Xf})$  é a profundidade de cobertura ajustada pela correlação de regressão da profundidade de cobertura do cromossomo X e o conteúdo GC correspondente de casos com fetos femininos,  $\hat{c}r_{i,Yf} = f(GC_{i,Yf})$  se refere à profundidade de cobertura ajustada pela correlação da regressão da profundidade de cobertura do cromossomo Y e o conteúdo GC correspondente de casos com fetos femininos,  $\hat{c}r_{i,Xm} = f(GC_{i,Xm})$  se refere à profundidade de cobertura ajustada pela correlação da regressão da profundidade de cobertura do cromossomo X e conteúdo CG correspondente de homens,  $\hat{c}r_{i,Ym} = f(GC_{i,Ym})$  se refere à profundidade de cobertura ajustada pela correlação de regressão da profundidade de cobertura do cromossomo Y e conteúdo GC correspondente de homens adultos. Somente para cálculo,  $\hat{\sigma}_{X,f}$  e  $\hat{\sigma}_{X,m}$  são iguais e  $\hat{\sigma}_{Y,f}$  e  $\hat{\sigma}_{Y,m}$  são iguais.

#### Exemplo 4. Cálculo do residual de cada cromossomo

A figura 6 demonstra que a variação padrão (ver fórmula 3) para cada cromossomo com um certo número total de leituras únicas é influenciada pelos números de referência de casos participantes. A variação padrão apenas aumenta quando o número de casos selecionados foi maior do que 150, com a condição de que 1,7 milhões de leituras únicas foram sequenciadas para cada caso. Entretanto, a variação padrão foi diferente para os diversos cromossomos. Após considerar a tendência GC, o método da presente invenção apresentou uma variação padrão moderada para o cromossomo 13 (0,0063), para o cromossomo 18 (0,0066) e para o cromossomo 21 (0,0072). A variação padrão do cromossomo X é maior do que a dos cromossomos mencionados acima, podendo exigir a realização de mais estratégias para detecção precisa de anormalidades.

A figura 7 apresenta a marcação Q-Q, onde o residual é compilado para distribuição normal, considerando o cálculo do teste t de student razoável.

#### Exemplo 5. A distinção do sexo do feto

Para descobrir anormalidades do cromossomo sexual, é melhor distinguir o sexo fetal. Existem dois picos óbvios quando a distribuição da frequência da profundidade de cober-

tura do cromossomo Y é investigada nos 300 casos da presente invenção, sugerindo a distinção do sexo pela profundidade de cobertura do cromossomo Y. Casos com profundidade de cobertura menor do que 0,04 podem estar relacionados a fetos femininos, enquanto maiores do que 0,051 podem indicar fetos masculinos, porém, não é possível distinguir o sexo com valores entre 0,04 e 0,051, de acordo com a figura 8. Para estes casos de aneuploidia e determinação duvidosa do sexo, a regressão logística foi utilizada para prever o sexo de acordo com a fórmula 11 (Fan, *et al.*, *Proc Natl Acad Sci USA* (2008) 42: 16266-16271):

$$\text{logit}(p_i) = \ln\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 cr.a_{i,x} + \beta_2 cr.a_{i,y} \quad (11)$$

onde  $cr.a_{i,x}$  e  $cr.a_{i,y}$  são normalizados em relação à cobertura de X e Y, respectivamente.

Em comparação com o resultado do cariótipo, o método da presente invenção para distinguir o sexo fetal foi bem sucedido nos 300 casos de referência com 100% de precisão, porém, apresentou um erro na realização do 901º caso, e o valor da profundidade do cromossomo Y deste caso permaneceu entre 0,04 e 0,051.

#### 15 Exemplo 6. Desempenho do diagnóstico da tentativa do test t da correlação GC Convocação das amostras

Novecentos e três (903) participantes foram recrutados do Hospital Shenzhen People e Shenzhen Maternal e centro de serviço de atendimento a crianças com seus resultados de cariótipos. As aprovações foram obtidas de grupos de revisões institucionais de cada local de convocação, e todos os participantes deram seu consentimento por escrito. As idades da mãe e idades gestacionais nas amostras sanguíneas foram registradas. Os 903 casos incluíram 2 casos de trissomia do 13, 15 casos de trissomia do 18, 16 casos de trissomia do 21, 3 casos de X0, 2 casos de XXY e 1 caso de XYY. A distribuição dos resultados dos cariótipos é demonstrada na figura 9.

#### 25 Sequenciamento do DNA plasmático materno

Sangue venoso periférico (5 mL) foi coletado de cada participante grávida em tubos de EDTA, e centrifugado a 1.600 g/10 minutos durante 4 horas. O plasma foi transferido para tubos de microcentrifugação e centrifugado novamente a 16.000 g/10 minutos, para remover células residuais. O plasma sem células foi armazenado à temperatura de 80 °C até a extração do DNA. Cada amostra de plasma foi congelada e descongelada somente uma vez.

Para sequenciamento genômico paralelo maciço, todo o DNA extraído da quantidade de 600 µL do plasma materno foi utilizado para construção do banco de DNA, de acordo com um protocolo modificado da Illumina. Resumidamente, um reparo nas extremidades de fragmentos do DNA plasmático materno foi realizado utilizando T4 DNA polimerase, Klenow™ polimerase e T4 polinucleotídeo quinase. Adaptadores disponíveis comercialmente

(Illumina) foram ligados aos fragmentos de DNA após a adição de resíduos A terminais. O DNA ligado ao adaptador foi então adicionalmente ampliado utilizando um PCR de 17 ciclos com primers múltiplos padrões. O kit Agencourt AMPure™ de 60 mL (Beckman) foi utilizado para a purificação dos produtos PCR. A distribuição do tamanho dos bancos de sequenciamento foi analisada com um kit de DNA 1000 em Bioanalyzer™ 2100 (Agilent) e quantificada com PCR de tempo real. Os bancos de sequenciamento com índices diferentes foram então reunidos por quantidades semelhantes antes da estação de agrupamento em Illumina GA II™ (sequenciamento de extremidade única).

Amostras euplóides de 19 homens foram sequenciadas para análise subsequente da fração de DNA fetal. Um novo teste t GC correlacionado foi desenvolvido na invenção para o diagnóstico da trissomia do 13, trissomia do 21, e anormalidades do cromossomo sexual. Este método foi comparado com outros dois métodos mencionados abaixo sobre o diagnóstico do desempenho.

#### Exemplo 7. Detecção da aneuploidia fetal tal como trissomia do 13, 18 e 21

Para determinar se o número de cópias de um cromossomo dentro de um caso de um paciente desviou do normal, a profundidade da cobertura de um cromossomo foi comparada com as de todas as outras dos casos de referência. Todos os estudos anteriores apresentaram somente uma hipótese nula. Uma hipótese nula ( $H_0$ : o feto é euplóide) estipulava que a profundidade de cobertura média da distribuição de casos de pacientes e a profundidade de cobertura média de toda distribuição de referência normal eram iguais, o que significava que o paciente era euplóide se essa hipótese nula fosse aceita. Utilizando o teste t de Student,  $t_1$  pode ser calculado pela fórmula 12:

$$t_{1,i,j} = (cr_{i,j} - \hat{cr}_{i,j}) / std_j \quad (12)$$

As outras hipóteses nulas ( $H_1$ : o feto é aneuplóide) estipulavam que a profundidade de cobertura média da distribuição de casos de pacientes com uma fração fetal imprecisa era igual à profundidade de cobertura média da distribuição de casos aneuplóides com a mesma fração fetal, o que significava que esse paciente era aneuplóide se essa hipótese nula fosse aceita. Utilizando o teste t de Student,  $t_2$  foi calculado pela fórmula 13:

$$t_{2,i,j} = (cr_{i,j} - \hat{cr}_{i,j}(1 + f_{xy_{i,j}} / 2)) / std_j \quad (13)$$

$|T_1| > 3$  e  $|t_2| < 3$  podem indicar um caso de aneuploidia especialmente quando as distribuições entre os casos de euploidia e aneuploidia forem completamente discriminadas, embora em outra condição, tal como na precisão insuficiente ou fração fetal insuficiente e assim por diante,  $|t_1|$  pode ser menor do que 3, mas o feto é anormal.  $T_1$  e  $t_2$  combinados podem ajudar na obtenção de uma decisão mais precisa, logo, foi empregada a proporção de probabilidade logarítmica de  $t_1$  e  $t_2$  de acordo com a fórmula 14:

$$L_{i,j} = \log(p(t_{1,i,j}, \text{degree} | D)) / \log(p(t_{2,i,j}, \text{degree} | T)) \quad (14)$$

onde  $L_{i,j}$  é a proporção de probabilidade logarítmica. Se a proporção for maior que 1, podemos concluir que o feto pode apresentar trissomia.

Porém, em casos de fetos femininos, é difícil avaliar sua fração fetal, logo é impossível calcular. Entretanto, pode ser fornecido um Valor de referência (RV) de 7%, de acordo com a distribuição empírica da fração fetal.

5 Novecentos e três (903) casos forem pesquisados, dentre eles 866 continham fetos euplóides, dos quais 300 casos foram selecionados aleatoriamente para a realização do teste t correlacionado com GC. Além disso, 2 casos de trissomia do 13, 12 casos de trissomia do 18, 16 casos de trissomia do 21, 4X (consistindo de 3 casos de X0, 1 caso de quimera 10 ra 45, xo/46, xx (27:23), e 2 casos de XXY e 1 de XYY participaram do estudo. Após o alinhamento, foi obtida uma média de 1,7 milhões de dados (SD = 306185) de leituras alinhadas únicas por caso, sem erros. Utilizando o teste t de Student correlacionado com GC, da presente invenção, todos os casos de T13 (2 entre 2) foram identificados com sucesso, enquanto 901 entre 901 casos de não ocorrência da trissomia do 13 foram corretamente classificados (figura 10A). A sensibilidade e especificidade deste teste foram de 100% e 100% (tabela 1).

Para os casos de trissomia do 18, 12 entre 12 casos de trissomia do 18, e 888 entre 891 casos de não ocorrência da trissomia do 18 puderam ser corretamente identificados (figura 10A). A sensibilidade e especificidade deste teste foram de 100% e 99,66%, respectivamente. Para trissomia do 21, 16 entre 16 casos de trissomia do 21, e 16 entre 16 casos de não ocorrência da trissomia do 21 puderam ser também corretamente detectados (figura 10A). A sensibilidade e especificidade deste teste foram de 100% e 100% respectivamente.

#### Exemplo 8. Detecção de X0, XXX, XXY, XYY

25 Consideramos que a detecção da trissomia para autossomos, e para os distúrbios dos cromossomos sexuais, tais como X0, XXX, XXY e XYY, também pode ser realizada pelo método da presente invenção.

Inicialmente, o sexo foi confirmado pela distinção de sexo. Se um caso foi confirmado para um feto feminino, o valor de t1 do teste de student  $t_{1,i,x} = (cr_{i,x} - \hat{cr}_{i,x}) / std_{x_f}$  foi necessário para cálculo da detecção de XXX ou X0, onde  $\hat{cr}_{i,x}$  e  $std_{x_f}$  são os mesmos da fórmula 10; se t1 é maior do que 3,13, esse caso pode ser XXX ou X0. Mas considerando que a precisão foi limitada pela grande variação da profundidade de cobertura para o cromossomo X, foi coletada uma nova amostra de plasma e repetido o experimento para obter uma decisão mais precisa quando  $|t1| < 5$ , ou ainda  $|t1| > 3,13$ . O valor de  $|t1| > 5$  foi confirmado como sendo aneuplóide nesse caso. Todos os processos de detecção partem da premissa 35 que os dados satisfazem o controle de qualidade padrão.

Se a amostra teste foi confirmada por apresentar fetos masculinos, a fração do DNA fetal foi estimada inicialmente por Y e X. Entretanto, podemos extrapolar a profundidade de cobertura ajustada para o cromossomo X com a fração de DNA fetal calculada somente pela profundidade de cobertura do cromossomo Y, e  $t_2$  pode ser calculado.

5  $t_{2i} = (cr_{i,x} - (1 - fy_i / 2) \cdot \hat{c}r_{i,x_f}) / std_{x_f}$ . Se  $t_2$  apresenta um valor muito alto (maior do que 5) ou muito pequeno (menor do que -5), o feto pode ser XXY ou XYY. Além disso, o intervalo entre as frações fetais estimadas por X e Y independentemente pode oferecer informação para a detecção de distúrbios dos cromossomos sexuais.

Na detecção de X0, três de 4 casos de X0 foram detectados, e o caso não foi detectado como um caso de quimera (figura 10B). A sensibilidade e especificidade deste teste foram de 75% (100% se não for considerado o caso de quimera) e 99,5% respectivamente. Para casos de XXY, todos os dois casos foram identificados com sucesso, enquanto 901 de 901 casos de não XXY foram corretamente classificados (figura 10B) com 100% de sensibilidade 100% de especificidade. Para o caso XYY, foi identificado corretamente (Figura 10B), e a sensibilidade e especificidade foram de 100% e 100% respectivamente.

Para avaliar se o novo teste da invenção apresenta vantagens quando comparado com outros dois testes, z-score e z-score com correção para GC, foram realizados os três testes para analisar os 900 casos, com os mesmos 300 casos de referência. A precisão de uma medição foi expressa no valor da confiança (CV). Em uma pesquisa da invenção, o valor da confiança (CV) do teste z-score padrão é maior do que o de outros testes de interesse clínico do cromossomo 18 e 21 (figura 11), gerando uma taxa de sensibilidade menor para a trissomia do 18 e 21 (tabela 1).

Tabela 1. Comparação da sensibilidade e especificidade de métodos diferentes

	Diagnóstico (# de casos)	Teste z-score padrão		Teste z-score com correção GC		Teste da invenção com teste t correlacionado com GC	
		Sensibilidade (%)	Especificidade (%)	Sensib. (%)	Espec. (%)	Sensib. (%)	Espec. (%)
Autossomas	Trissomia 13 (2)	50%	99,89	100%	100%	100%	100%
	Trissomia 18 (12)	91,67%	100%	100%	99,89%	100%	99,96%
	Trissomia 21 não (16)	93,75%	100%	100%	100%	100%	100%
heterosoma	X0 (3 X0, 1 quimera X0/XX)	Não disponível	Não disponível	Não disponível	Não disponível	75%	100%
	XXY (1)	Não disponível	Não disponível	Não disponível	Não disponível	100%	100%
	XXY (2)	Não disponível	Não disponível	Não disponível	Não disponível	100%	100%

			pon.		
Hipótese binária	Não disponível		Não disponível		Não disponível

Para o teste z-score com correlação GC, o valor de CV do cromossomo 13 é de 0,0066 com 100% de taxa de sensibilidade e 100% de taxa de especificidade. Para o novo teste t de Student com correlação GC discutido na presente invenção, o valor de CV do cromossomo 13 é de 0,0063, com 100% de taxa de sensibilidade e 100% de taxa de especificidade. No cromossomo 18, o CV destes dois testes foram de 0,0062 e 0,0066, respectivamente, ambos com taxas de 100% de sensibilidade e 100% e taxas de especificidade de 99,89% e 99,96% respectivamente. O desempenho foi similar quando comparando o CV destes dois testes para o cromossomo 21: 0,0088 e 0,0072, respectivamente. Ambos apresentaram a mesma taxa de sensibilidade de 100% no estudo de pequenos casos e obtiveram a mesma taxa de especificidade de 100%. E estes dois métodos apresentaram um melhor desempenho do que o teste z-score padrão. O novo teste desenvolvido na invenção com correção GC não foi somente comparável ao teste de correção GC com bom desempenho, mas também apresentou outra vantagem na detecção das anormalidades relacionadas aos cromossomos sexuais, tais como X0, XXY e XYY. Os dados demonstram que quando o teste de correção GC é utilizado, pode ser difícil distinguir o sexo dos fetos pelo desvio de dados que representam os cromossomos sexuais, introduzindo o número de etiquetas de sequências expressas multiplicado pelo fator peso, dificultando a detecção do distúrbio do cromossomo sexual.

#### Exemplo 9. Desempenho teórico do teste t correlacionado com GC em relação ao tamanho dos dados, semanas de gestação e fração do DNA fetal

A medição de aneuplóides permanece um desafio devido a alta concentração de DNA materno (Fan *et al.*, *Proc Natl Acad Sci USA* (2008) 42:16266 – 16271), e uma pequena fração de DNA fetal é o fator de restrição mais significativo para detecção de aneuploidia através do sequenciamento do genoma paralelo maciço (MPGS). Entretanto, não existe nenhum grande avanço na determinação clínica da fração mínima do DNA fetal anterior à detecção por MPGS, especialmente em fetos femininos, embora as semanas gestacionais sejam o único sinal clínico relacionado à fração de DNA fetal. Foi relatado, anteriormente, que existe uma correlação estatisticamente significativa entre a fração de DNA fetal e a idade gestacional (Lo, *et al.*, *Am. J. Human Genet.* (1998) 62: 768-775). No estudo da invenção, para pesquisar a relação entre a fração de DNA fetal estimada e a idade gestacional, foi marcada na figura 12, a fração de DNA fetal de todos os casos participantes com fetos masculinos (total de 427 casos), calculados pela fórmula 10. A fração do DNA fetal estimada para cada amostra está correlacionada com a idade gestacional (P menor do que 0,0001). Foi demonstrado também que mesmo na idade gestacional 20, ocorreram 4 entre 65 casos com fração de DNA fetal menor do que 5%, o que poderia afetar a precisão da detecção.

Para avaliar o método de cálculo da fração fetal, foram selecionados alguns casos distribuídos na fração fetal estimada, e então Q-PCR ajudou a calcular outra fração fetal relativa. Então, foi obtida uma curva padrão demonstrando uma forte correlação entre eles, e que o cálculo da fração fetal pelo método da presente invenção é confiável.

5           Entretanto, a profundidade de sequenciamento (o número de leituras totais únicas) foi outro fator significativo que afetou a precisão da detecção da aneuploidia empregado no valor da variação padrão. A variação padrão para cada cromossomo utilizado no teste relacionado ao GC pode ser fixada de acordo com certo nível de profundidade de sequenciamento, quando o número de casos de referência atinge o valor de 150 (figura 13). Para investigar como a profundidade de sequenciamento influencia a variação padrão para cada cromossomo, cento e cinquenta (150) casos foram sequenciados não somente no nível de 10 1,7 milhões, mas em outro nível com o número total de leituras únicas atingindo 5 milhões (SD = 1,7 milhões). Dependendo destas duas configurações, foi descoberto que a variância padrão é linear com o inverso da raiz quadrada do número total de leituras únicas demonstrada na figura 6.

15           Para uma determinada fração de DNA fetal, podemos estimar o número total de leituras únicas, de acordo com o método da presente invenção, para detectar o desvio do número de cópias do cromossomo do normal, com  $t_1$  igual a 3 (figura 14). Isto demonstra que quanto menor a fração de DNA, maior a profundidade de sequenciamento necessária. Com a configuração de 1,7 milhões de leituras únicas, o teste é capaz de detectar fetos com aneuploidia no cromossomo 13 e X com fração de DNA fetal maior do que 4,5%, e fetos com aneuploidia no cromossomo 21 e 18 com mais de 4%; enquanto na configuração com referência de 5 milhões, o teste foi capaz de detectar trissomia no cromossomo 18 e 21, ainda com a fração de DNA fetal de cerca de 3%. Para a identificação de fetos com anormalidades no cromossomo X, tais como XXX ou X0, com a fração fetal de cerca de 4%, o número total de leituras únicas necessário nestes casos e nos casos correspondentes deve atingir 5 milhões. Se o DNA fetal for menor do que 3,5%, a exigência da profundidade de sequenciamento ultrapassará 20M. E se a fração do DNA fetal for menor, a detecção será difícil e não confiável, logo, é proposta outra estratégia, como por exemplo, coletar uma nova amostra do plasma materno, refazer o experimento e reanalisar os dados quando a gestação estiver mais avançada, ou seja, quando a fração do DNA se tornar maior com o aumento da idade gestacional. E essa estratégia também pode ser aplicada em amostras que apresentem uma pequena fração de DNA fetal.

35           Mesmo que os testes apresentem um bom desempenho, não são convincentes sem um grande número de casos de anormalidades. Para calcular a sensibilidade do método da presente invenção da correlação do teste t de student com GC, foi publicada a sensibilidade teórica considerando diferentes idades gestacionais e diferentes profundidades de

sequenciamento.

Foi calculada a sensibilidade teórica da aneuploidia através das seguintes etapas. Primeiro, foi aplicada uma análise da regressão para ajustar o DNA fetal à idade gestacional

$\hat{fr}_i = f(gsa_i)$ , onde  $\hat{fr}_i$  é a média de ajuste da fração do DNA fetal em sua idade gestacional,  $gsa_i$  e cálculo da distribuição da fração de DNA fetal empregando o cálculo da densidade de kernel Gaussian (Birke, (2008) *Journal of Statistical Planning and Inference* 19: 2851-2862), principalmente em relação à fração do DNA fetal estimada, distribuída em 19 e 20 semanas gestacionais antes de extrapolar a distribuição da fração de DNA fetal nas outras semanas, de acordo com a relação entre a fração de DNA fetal e a idade gestacional

5  $\hat{pd}_i(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - X_i}{h}\right) + (fr_i - \frac{fr_{19} + fr_{20}}{2})$ , onde  $\hat{pd}_i$  é a densidade da probabilidade ajustada da fração de DNA fetal em sua idade gestacional, onde X é a data de 19 e 20 semanas gestacionais (figura 12). Segundo, foi calculada a variância padrão de acordo com os números de leituras gestacionais únicas, conforme mencionado anteriormente  $\hat{\sigma} = f(tuqn)$ , onde  $tuqn$  é o número total de leituras únicas. Finalmente, para calcular a sensibilidade em cada

15 idade gestacional em certo nível de profundidade de sequenciamento, foi computada a densidade da probabilidade do falso negativo em cada fração do DNA fetal (na presente invenção, com flutuação da fração de DNA fetal normalmente distribuída), e então com a sua integração para obter uma taxa de falso negativo (FNR) em uma idade gestacional com todos os níveis de fração de DNA fetal  $FNR(fr, week) = \int_0^1 pd(fr) \int_0^{x/\sigma_j=3} \exp(x - fr)^2 / 2 \cdot \sigma_j^2 dx dfr$ ,

20 onde j é o cromossomo j. A sensibilidade teórica em certa profundidade de sequenciamento nesta idade gestacional é facilmente calculada como 1-FNR. As figuras 15-21 demonstraram as marcações dos resultados dos cálculos. O teste t de Student maior do que 3 foi configurado para identificar fetos femininos aneuplóides, enquanto para os fetos masculinos, durante a computação da densidade da probabilidade de falso negativo em cada fração, um logaritmo maior do que 1 foi empregado como o valor crítico, conforme mencionado na hipótese binária, ajudando na obtenção de uma alta sensibilidade em comparação com os fetos femininos.

25 Entretanto, nossa conclusão é relativamente conservadora pelo fato de que é difícil obter uma distribuição aproximada a distribuição real da fração de DNA fetal juntamente com a idade gestacional, especialmente no início da gestação com escala menor da amostra.

### Referências

1. Virginia P. Sybert, Elizabeth McCauley (2004). Turner's Syndrome., *N Engl J Med* (2004) 351:1227-1238.

2. Robert Bock (1993). Understanding Klinefelter Syndrome: A Guide for XXY Males and Their Families. NIH Pub. No. 93-3202 August 1993
3. Aksglaede, Lise; Skakkebaek, Niels E.; Juul, Anders (January 2008). "Abnormal sex chromosome constitution and longitudinal growth: serum levels of insulin-like growth factor (IGF)-I, IGF binding protein-3, luteinizing hormone, and testosterone in 109 males with 47,XXY, 47,XYY, or sex-determining region of the Y chromosome (SRY)-positive 46,XX karyotypes". *J Clin Endocrinol Metab* 93 (1): 169–176. doi:10.1210/jc.2007-1426.PMID 17940117
4. H. Bruce Ostler (2004). Diseases of the eye and skin: a color atlas. Lippincott Williams & Wilkins. pp. 72. ISBN 9780781749992.
5. Driscoll DA, Gross S (2009) Clinical practice. Prenatal screening for aneuploidy. *N Engl J Med* 360: 2556-2562.
6. Karl O.Kagan, Dave Wright, Catalina Valencia etc (2008). Screening for trisomies 21, 18 and 13 by maternal age, fetal nuchal translucency, fetal heart rate, free b-hCG and pregnancy-associated plasma protein-A. Human Reproduction Vol.23, No.9 pp. 1968–1975, 2008 doi:10.1093/humrep/den224.
7. Malone FD, et al. (2005) First-trimester or second-trimester screening, or both, for Down's syndrome. *N Engl J Med* 353:2001–2011.
8. Fan HC, Quake SR (2010) Sensitivity of Noninvasive Prenatal Detection of Fetal Aneuploidy from Maternal Plasma Using Shotgun Sequencing Is Limited Only by Counting Statistics. *PLoS ONE* 5(5): e10439. doi:10.1371/journal.pone.0010439.
9. Chiu RW, Chan KC, Gao Y, Lau VY, 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 USA* 105: 20458–20463.
10. McCullagh, P. and Nelder, J.~A. (1989), Generalized Linear Models, Londres, Inglaterra: Chapman & Hall/CRC.
11. Fan HC, Blumenfeld YJ, et al. (2008) Noninvasive diagnosis of fetal aneuploidy by shotgun sequencing DNA from maternal blood. *Proc Natl Acad Sci USA* 105:16266–16271.
12. Melanie Birke. (2008) Shape constrained kernel density estimation. *Journal of Statistical Planning and Inference* Volume 139, Issue 8, 1 de agosto de 2009, Páginas 2851-2862.
13. Lo et al., *Lancet* 350:485-487 (1997).
14. Lo et al., *Am. J. hum. Genet.* 62:768-775 (1998).
15. Pertl and Bianchi, *Obstetrics and Gynecology* 98:483-490 (2001).
16. Rogers and Ventner, "Genomics: Massively parallel sequencing," *Nature*, 437, 326-327 (15 de setembro. 2005).

17. Mewar et al., "Clinical and molecular evaluation of four patients with partial duplications of the long arm of chromosome 18," *Am J Hum Genet.* 1993 Dezembro; 53(6):1269-78.
18. Margulies et al., (2005) *Nature* 437:376-380.
- 5 19. Harris et al., (2008) *Science*, 320:106-109.
20. Soni and Meller, (2007) *Clin Chem* 53:1996-2001.
21. Dear, (2003) *Brief Funct Genomic Proteomic* 1:397-416.

## REIVINDICAÇÕES

1. Método implementado por computador para determinar uma anormalidade genética fetal a qual é uma aneuploidia cromossômica, **CARACTERIZADO** pelo fato de que compreende:

5           (a) obtenção de informação de sequência de fragmentos de polinucleotídeos múltiplos de uma amostra, a dita amostra sendo uma amostra de sangue periférico derivada de uma fêmea grávida e contendo DNA materno e fetal;

          (b) atribuir os ditos fragmentos aos cromossomos baseado na dita informação de sequência, pela comparação dos ditos fragmentos com as Leituras Únicas de Referência do mesmo tamanho para cada um dos ditos cromossomos, em que as Leituras Únicas de Referência são fragmentos de um cromossomo que tem uma sequência única que pode ser inequivocamente atribuída a uma única localização cromossômica com base em uma sequência genômica de referência;

10           (c) determinar a profundidade da cobertura e do conteúdo de GC de um cromossomo com base na informação de sequência para os fragmentos que foram atribuídos a Leituras Únicas de Referência do dito cromossomo na etapa (b),

          em que a profundidade de cobertura é a razão entre o número de fragmentos atribuídos unicamente ao dito cromossomo e o número de Leituras Únicas de Referência para o dito cromossomo do mesmo tamanho de fragmento com base na dita sequência genômica humana de referência;

20           (d) determinar a profundidade de cobertura ajustada do dito cromossomo utilizando o referido conteúdo de GC do dito cromossomo e estabelecer uma relação entre a profundidade de cobertura e o conteúdo de GC do dito cromossomo na ausência de aneuploidia, em que a dita relação estabelecida foi determinada por um método que compreende as etapas de:

          (i) obtenção de informação de sequência de múltiplos fragmentos polinucleotídicos cobrindo o dito cromossomo a partir de uma pluralidade de amostras de sangue periférico euplóide contendo DNA genômico, em que o tamanho do fragmento é o mesmo que o tamanho do fragmento dos fragmentos de polinucleotídeos múltiplos da etapa (a) acima;

30           (ii) atribuir os ditos fragmentos a cromossomos com base na dita informação de sequência como na etapa (b) acima;

          (iii) determinar a profundidade de cobertura e o conteúdo de GC do dito cromossomo com base na dita informação de sequência para cada amostra euplóide como na etapa (c) acima; e

35           (iv) usar a profundidade de cobertura e o conteúdo de GC determinado para cada amostra na etapa (iii) para determinar a relação entre a profundidade de cobertura e o conteúdo de GC do dito cromossomo na ausência de aneuploidia; e

(e) comparar a dita profundidade de cobertura ajustada com a profundidade de cobertura do dito cromossomo determinado na etapa (c), em que uma diferença entre eles indica aneuploidia cromossômica fetal.

2. Método, de acordo com a reivindicação 1, **CHARACTERIZADO** pelo fato de que a etapa (a) compreende ainda a obtenção de informação de sequência de múltiplos fragmentos polinucleotídicos de múltiplas amostras diferentes e a profundidade de cobertura é normalizada para se levar em conta as diferenças no número total de sequências lidas obtidas para amostras diferentes, por exemplo contra a profundidade média de cobertura de outro cromossomo, de preferência contra a profundidade média de cobertura de todos os outros autossomos ou contra a profundidade média de cobertura de todos os outros cromossomos.

3. Método, de acordo com a reivindicação 1 ou 2, **CHARACTERIZADO** pelo fato de que o conteúdo de GC do cromossomo é determinado como o conteúdo médio de GC de todos os fragmentos que atribuem ao dito cromossomo para o propósito da etapa (c).

4. Método, de acordo com a reivindicação 2, **CHARACTERIZADO** pelo fato de que compreende ainda determinar o sexo fetal, por exemplo de acordo com a fórmula:

$$\text{logit}(p_i) = \ln\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 cr.a_{i,x} + \beta_2 cr.a_{i,y}$$

em que  $cr.a_{i,x}$  e  $cr.a_{i,y}$  são a profundidade de cobertura relativa dos cromossomos X e Y da amostra  $i$ , respectivamente.

5. Método, de acordo com a reivindicação 2, **CHARACTERIZADO** pelo fato de que compreende ainda estimar a fração fetal, em que a fração fetal é calculada com base na profundidade de cobertura do cromossomo X e/ou Y, determinada conforme definido na etapa (c) da reivindicação 1, de acordo com uma fórmula selecionada de:

$$(i) \quad f_{y_i} = (cr_{i,Y} - \hat{c}r_{i,Yf}) / (\hat{c}r_{i,Ym} - \hat{c}r_{i,Yf}),$$

em que  $cr_{i,Y}$  é a profundidade de cobertura do cromossomo Y da amostra  $i$ ,  $\hat{c}r_{i,Yf} = f(GC_{i,Yf})$  é a profundidade de cobertura ajustada calculada a partir da relação entre a profundidade de cobertura do cromossomo Y e o conteúdo de GC correspondente de amostras de mulheres grávidas com um feto feminino, e  $\hat{c}r_{i,Ym} = f(GC_{i,Ym})$  se refere à profundidade de cobertura ajustada calculada a partir da relação entre a profundidade de cobertura do cromossomo Y e o conteúdo de GC correspondente de indivíduos do sexo masculino; ou

$$(ii) \quad f_{x_i} = (cr_{i,X} - \hat{c}r_{i,Xf}) / (\hat{c}r_{i,Xm} - \hat{c}r_{i,Xf}),$$

em que  $cr_{i,X}$  é a profundidade de cobertura do cromossomo X da amostra  $i$ ,  $\hat{cr}_{i,Yf} = f(GC_{i,Yf})$  é a profundidade de cobertura ajustada calculada a partir da relação entre a profundidade de cobertura do cromossomo X e o conteúdo de GC correspondente de amostras de mulheres grávidas com um feto feminino,  $\hat{cr}_{i,Xm} = f(GC_{i,Xm})$  se refere à profundidade de cobertura ajustada calculada a partir da relação entre a profundidade de cobertura do cromossomo X e o conteúdo de GC correspondente de amostras de indivíduos do sexo masculino; ou

$$(iii) \quad f_{xy} = \arg \min_{\varepsilon \in (0,1)} \left( \frac{(\hat{cr}_{i,Yf} \cdot (1-\varepsilon) + \hat{cr}_{i,Xm} \cdot \varepsilon - cr_{i,X})^2}{(\hat{\sigma}_{X,f} \cdot (1-\varepsilon))^2 + (\hat{\sigma}_{X,m} \cdot \varepsilon)^2} + \frac{(\hat{cr}_{i,Yf} \cdot (1-\varepsilon) + \hat{cr}_{i,Ym} \cdot \varepsilon - cr_{i,X})^2}{(\hat{\sigma}_{Y,f} \cdot (1-\varepsilon))^2 + (\hat{\sigma}_{Y,m} \cdot \varepsilon)^2} \right),$$

em que  $\hat{cr}_{i,Yf} = f(GC_{i,Yf})$  é a profundidade de cobertura ajustada calculada a partir da relação entre a profundidade de cobertura do cromossomo X e o conteúdo de GC correspondente de amostras de mulheres grávidas com um feto feminino,  $\hat{cr}_{i,Yf} = f(GC_{i,Yf})$  se refere à profundidade de cobertura ajustada calculada a partir da relação entre a profundidade de cobertura do cromossomo Y e o conteúdo de GC correspondente de amostras de mulheres grávidas com um feto feminino,  $\hat{cr}_{i,Xm} = f(GC_{i,Xm})$  se refere à profundidade de cobertura ajustada calculada a partir da relação entre a profundidade de cobertura do cromossomo X e o conteúdo de GC correspondente de amostras de indivíduos do sexo masculino, e  $\hat{cr}_{i,Ym} = f(GC_{i,Ym})$  se refere à profundidade de cobertura ajustada calculada a partir da relação entre a profundidade de cobertura do cromossomo Y e o conteúdo de GC correspondente de indivíduos do sexo masculino.

6. Método, de acordo com a reivindicação 2, **CARACTERIZADO** pelo fato de que a comparação da dita profundidade de cobertura ajustada com a profundidade de cobertura do cromossomo determinada na etapa (c) é conduzida por um teste de hipótese estatística, em que uma hipótese é de que o feto seja euploide (H0) e a outra hipótese é que o feto exiba aneuploidia para o dito cromossomo (H1).

7. Método, de acordo com a reivindicação 6, **CARACTERIZADO** pelo fato de que o teste estatístico t de Student é calculado para ambas as hipóteses.

8. Método, de acordo com a reivindicação 7, **CARACTERIZADO** pelo fato de que o teste estatístico t de Student é calculado para H0 e H1 de acordo com as fórmulas:

$$t1_{i,j} = (cr_{i,j} - \hat{cr}_{i,j}) / std_j \quad e$$

$$t2_{i,j} = (cr_{i,j} - \hat{cr}_{i,j}(1 + f_{xy_{i,j}}/2)) / std_j,$$

respectivamente, em que  $f_{xy}$  é a fração fetal,

$i$  é o índice de amostra,

$j$  é o número do cromossomo,

$$cr_{i,j} = f(GC_{i,j}) + \varepsilon_{i,j}, j = 1, 2, \dots, 22, X, Y,$$

em que  $f(GC_{i,j})$  representa a função para a relação entre a profundidade de cobertura normalizada e o conteúdo de GC correspondente da amostra  $i$ , cromossomo  $j$ ,  $\varepsilon_{i,j}$  representa

o resíduo da amostra  $i$ , cromossomo  $j$ ,

$$\hat{cr}_{i,j} = f(GC_{i,j}), j = 1, 2, \dots, 22, X, Y,$$

e representa a profundidade de cobertura ajustada, e

$std_j$  é a variação padrão de acordo com a fórmula:

$$std_j = \sqrt{\sum_i (cr_{i,j} - \hat{cr}_{i,j})^2 / (ns - 1)}, j = 1, 2, \dots, 22, X, Y,$$

em que  $ns$  representa o número de amostras de referência.

9. Método, de acordo com a reivindicação 8, **CARACTERIZADO** pelo fato de que a proporção da probabilidade logarítmica de  $t1$  e  $t2$  é calculada de acordo com a fórmula:

$L_{i,j} = \log(p(t1_{i,j}, degree | D)) / \log(p(t2_{i,j}, degree | T))$ , em que  $L_{i,j}$  é a proporção da probabilidade logarítmica, em que *degree* se refere ao nível da distribuição  $t$ ,  $D$  se refere à

Diploidia,  $T$  se refere à Trissomia, e  $p(t1_{i,j}, degree | *)$ ,  $*=D, T$  representa a densidade de probabilidade condicional fornecida a um nível de distribuição  $t$ , se a proporção for maior que 1, o feto sendo inferido para exibir trissomia do dito cromossomo.

10. Método, de acordo com qualquer uma das reivindicações 1 a 9, **CARACTERIZADO** pelo fato de que é para uso na determinação de uma aneuploidia autossômica fetal.

11. Método, de acordo com a reivindicação 10, **CARACTERIZADO** pelo fato de que a aneuploidia fetal é selecionada do grupo que consiste nas trissomias 13, 18 e 21.

12. Método, de acordo com a reivindicação 4, **CARACTERIZADO** pelo fato de que é para uso na determinação de uma aneuploidia do cromossomo sexual, tal como uma aneuploidia do cromossomo sexual selecionada do grupo que consiste em XO, XXX, XXY e XYY.

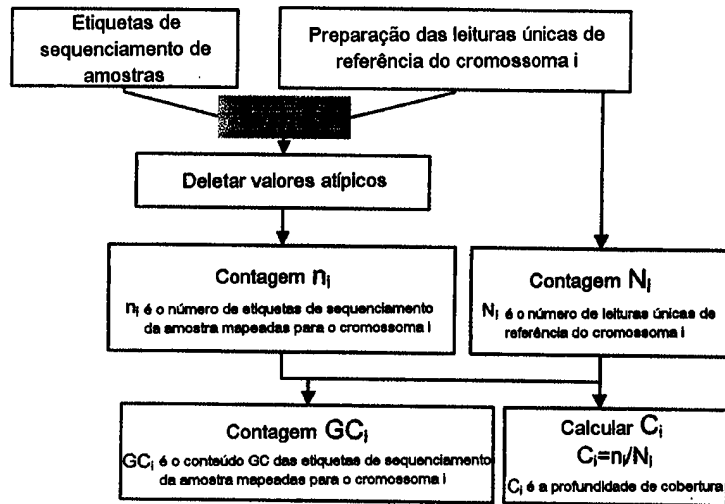


Figura 1

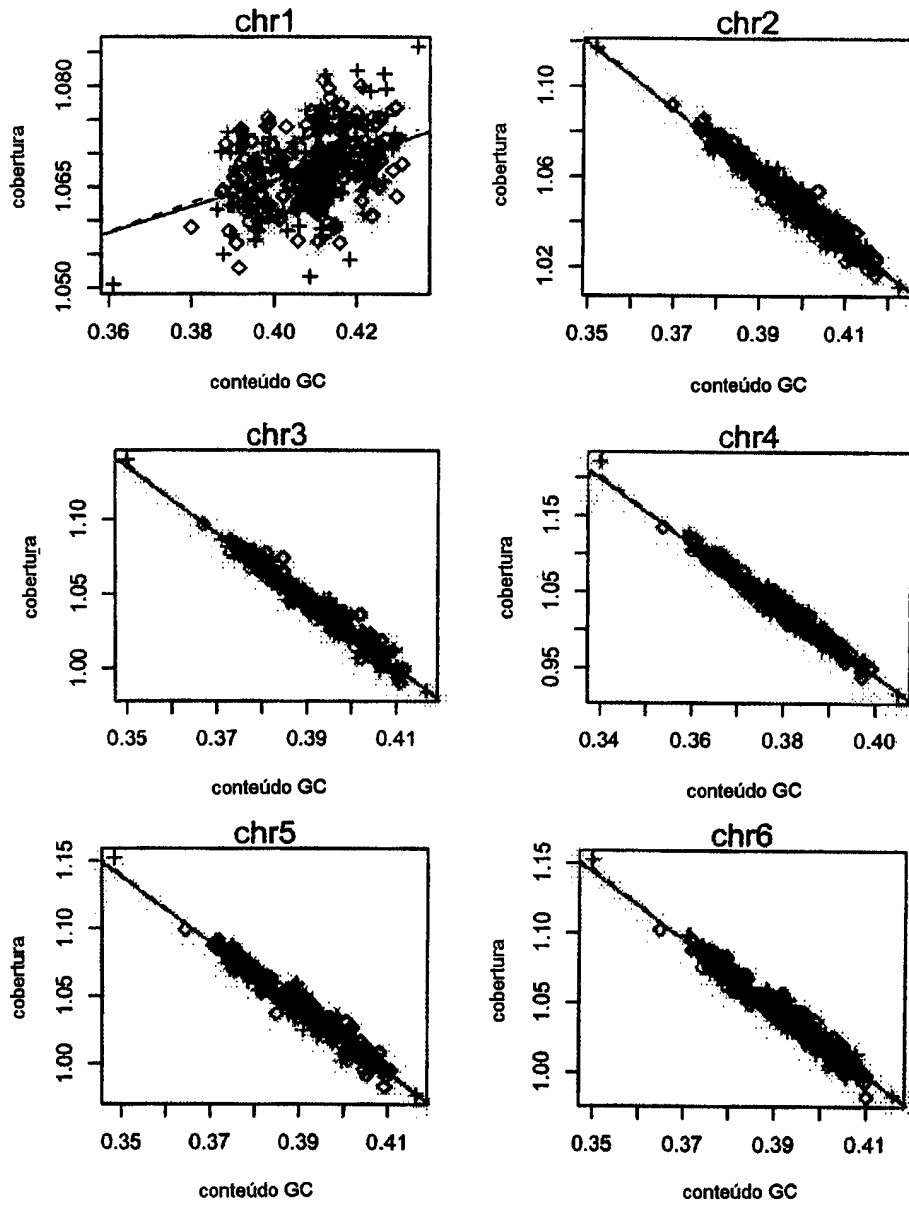


Figura 2

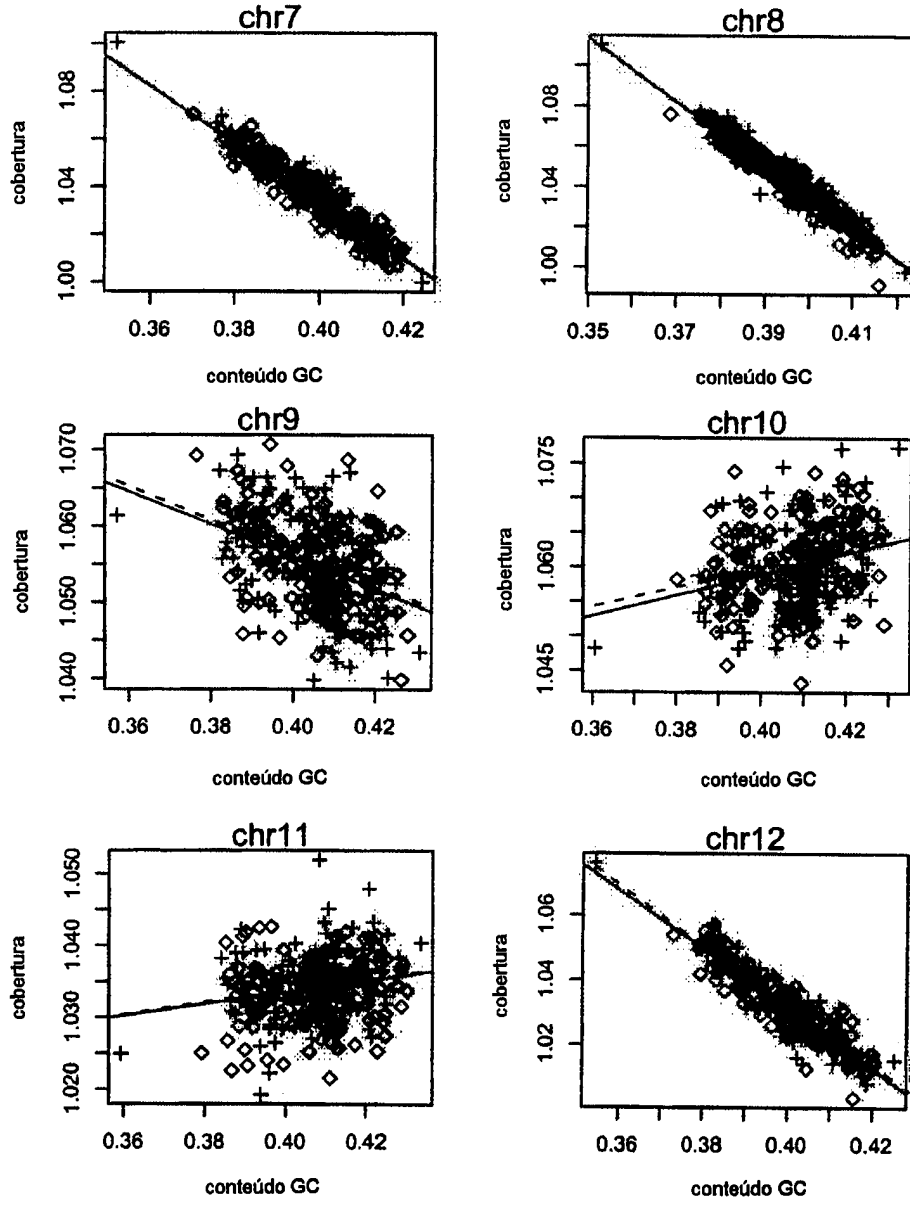


Figura 2

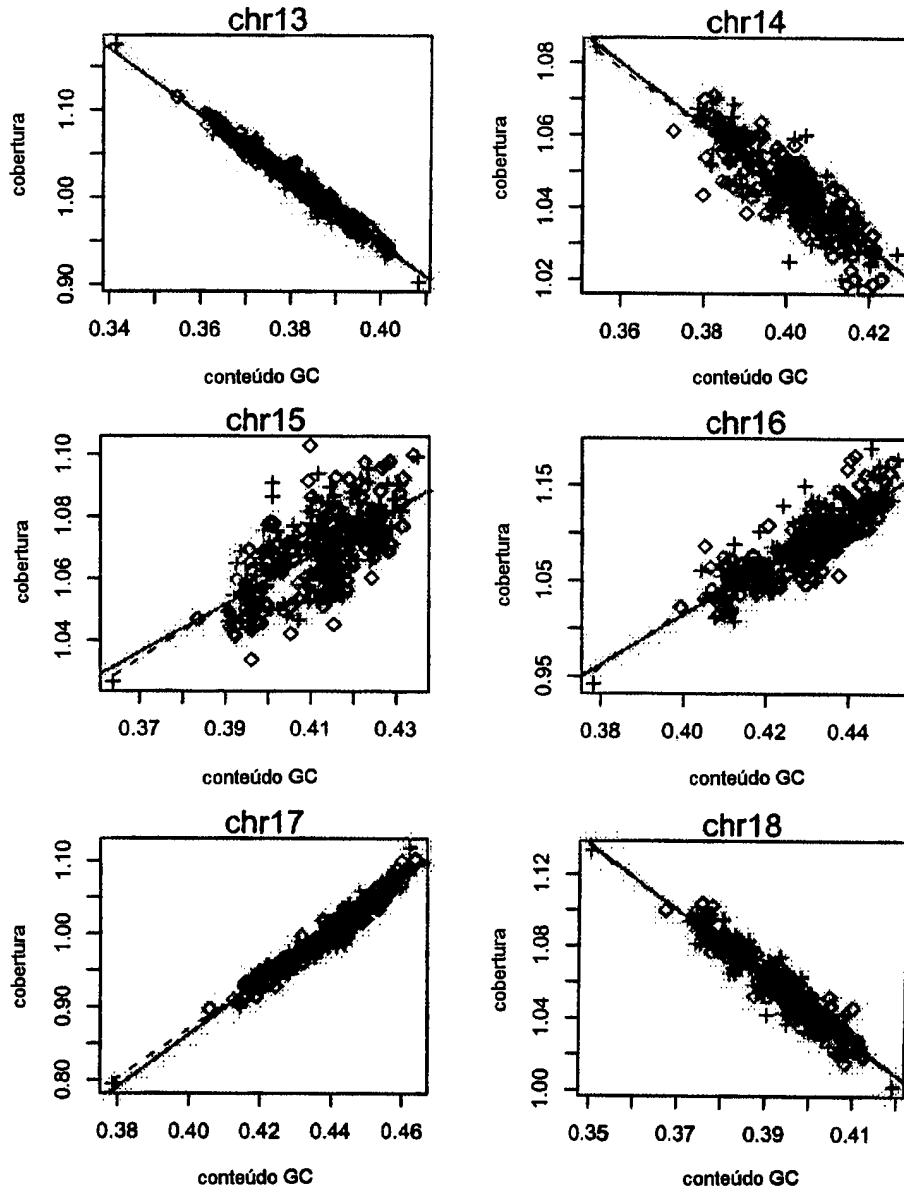


Figura 2

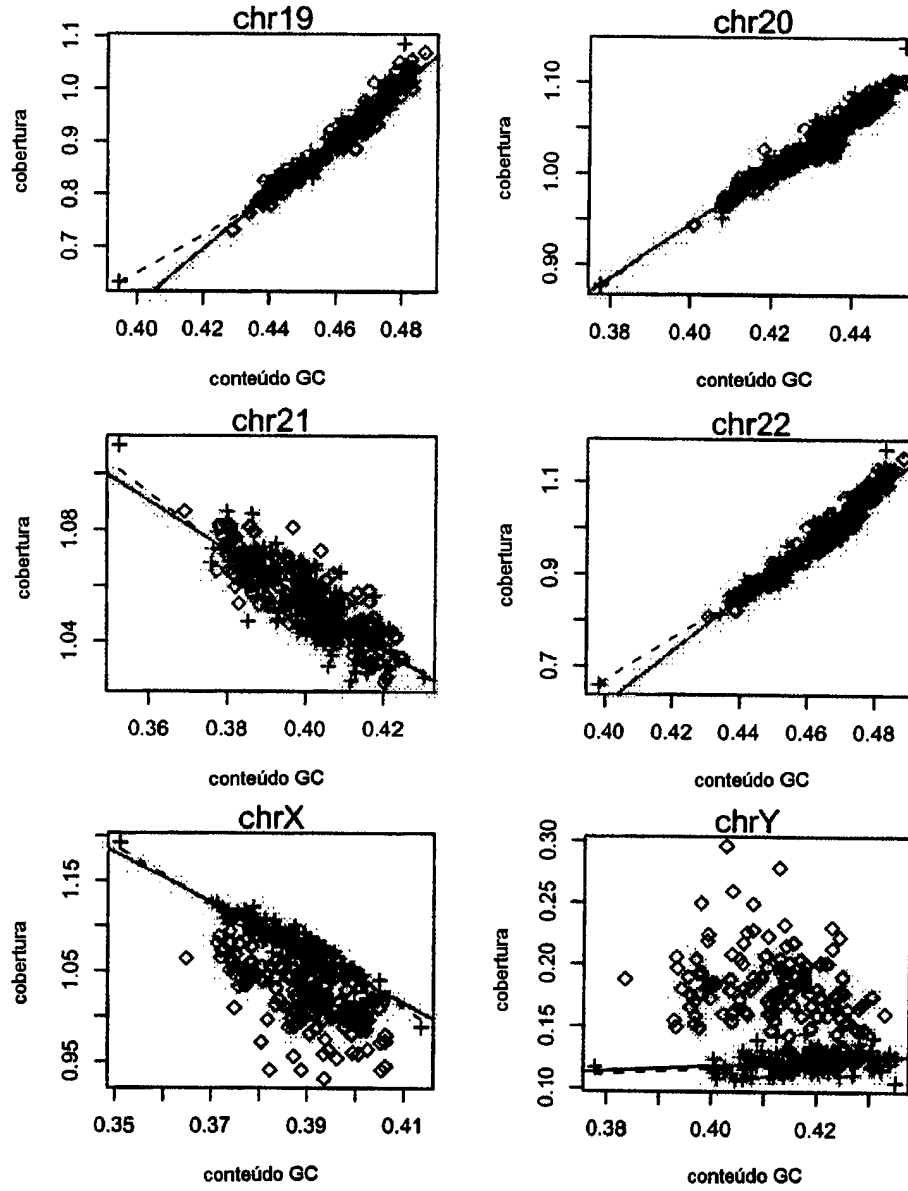


Figura 2

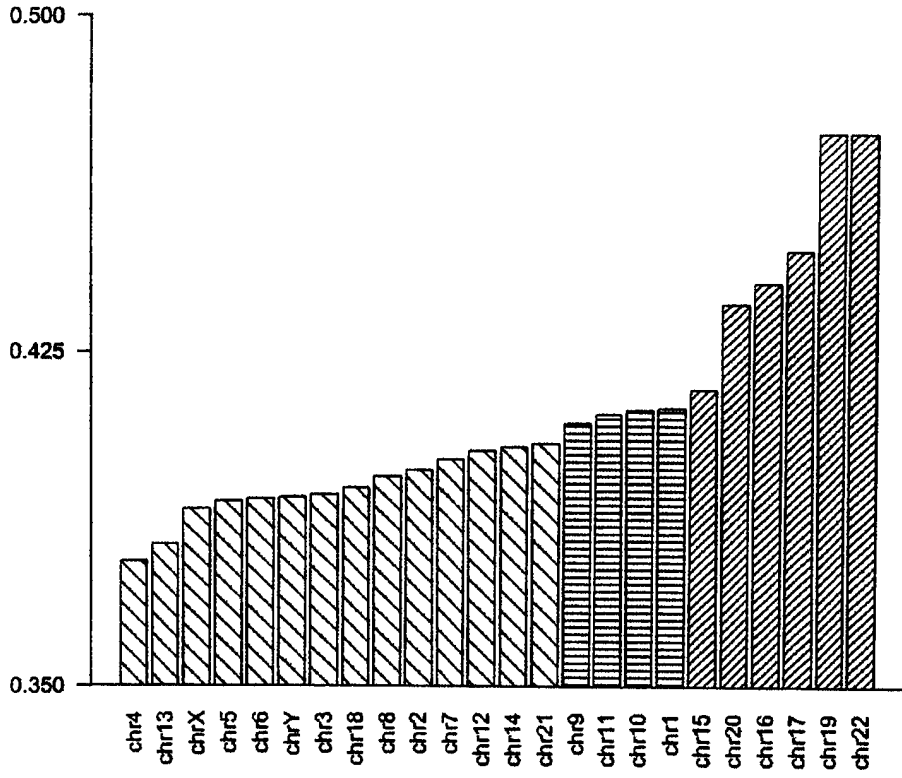
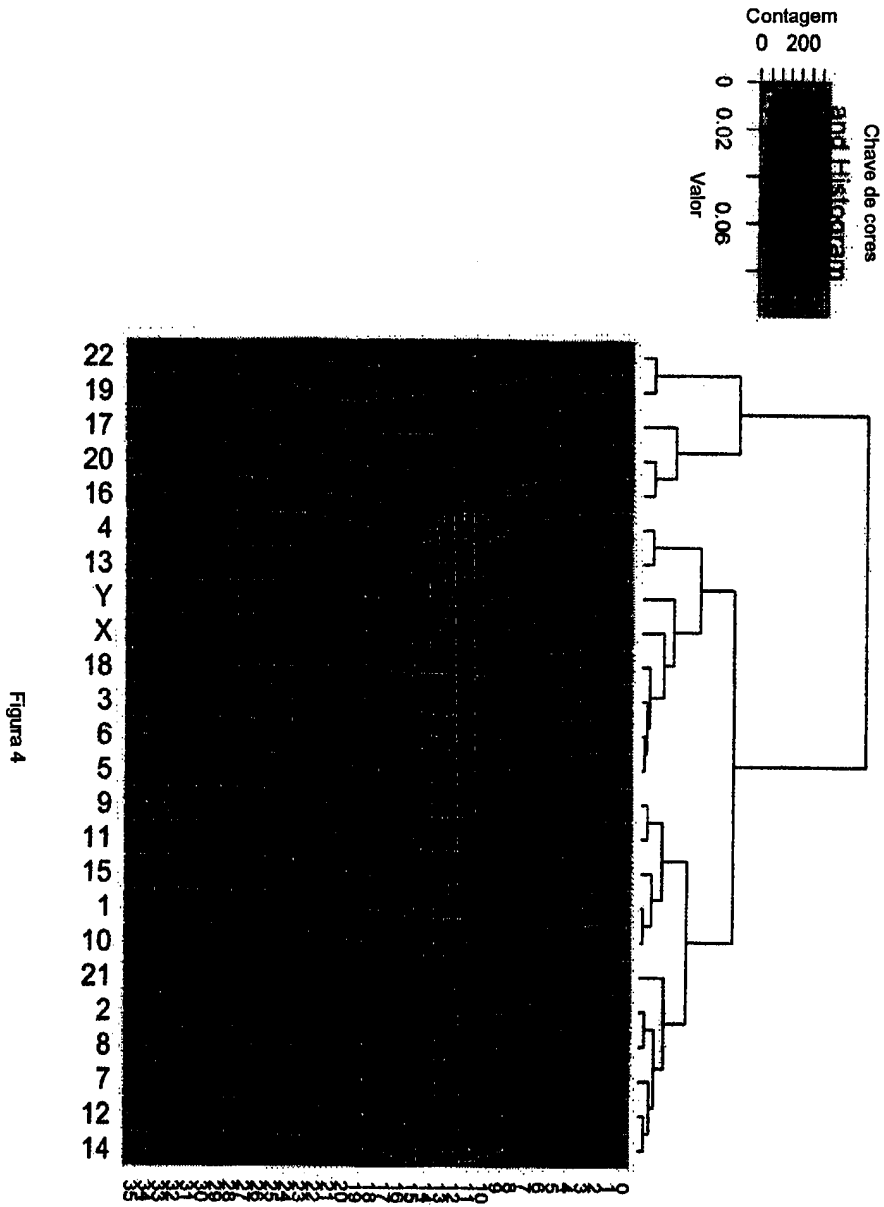


Figura 3



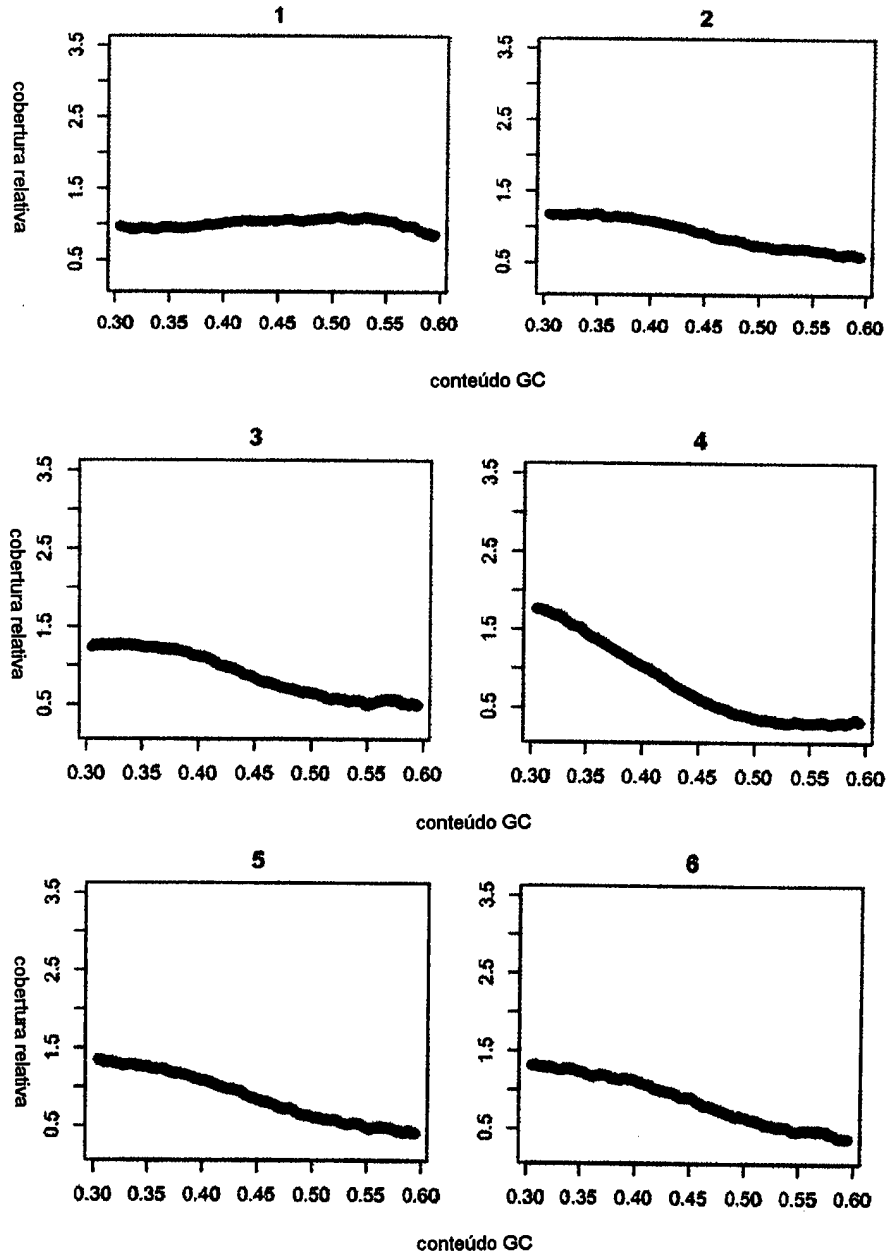


Figura 5

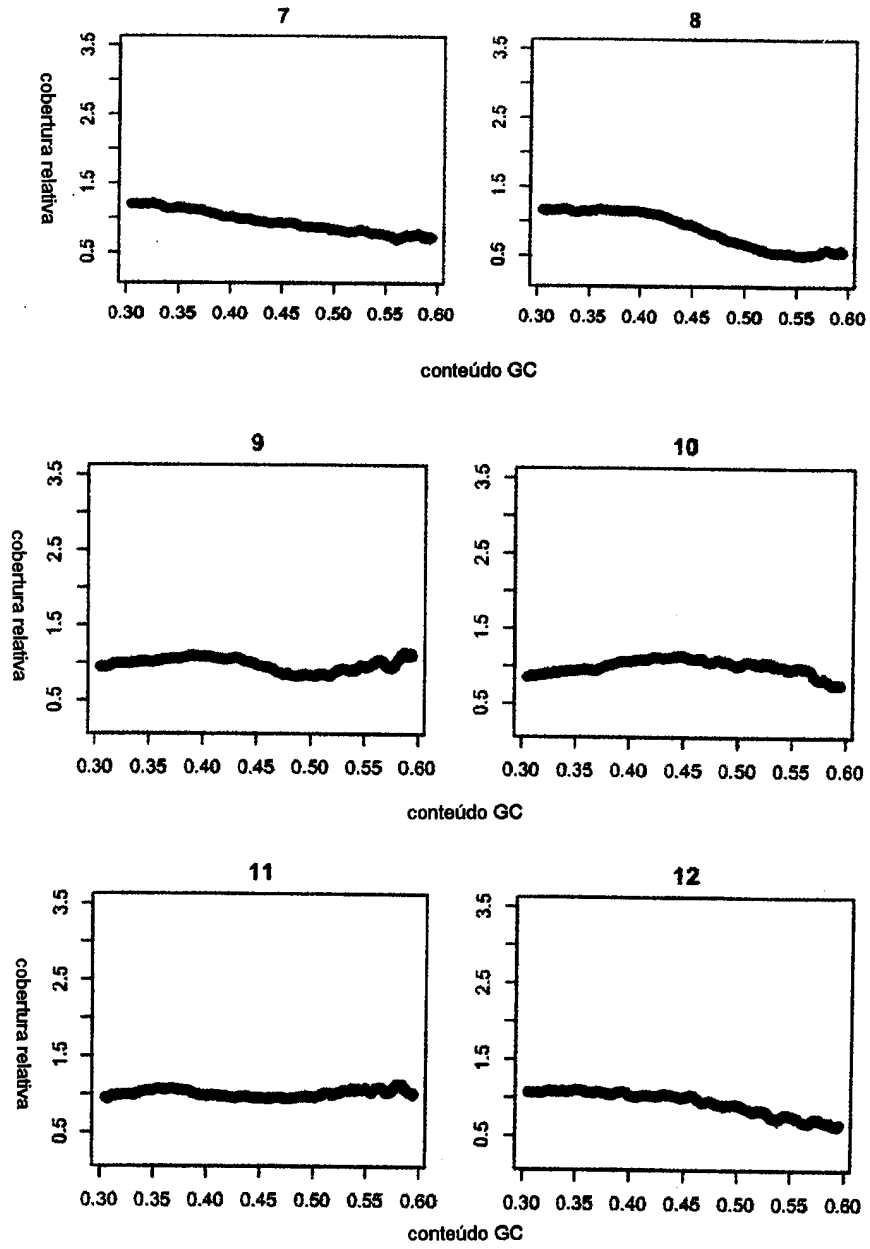


Figura 5

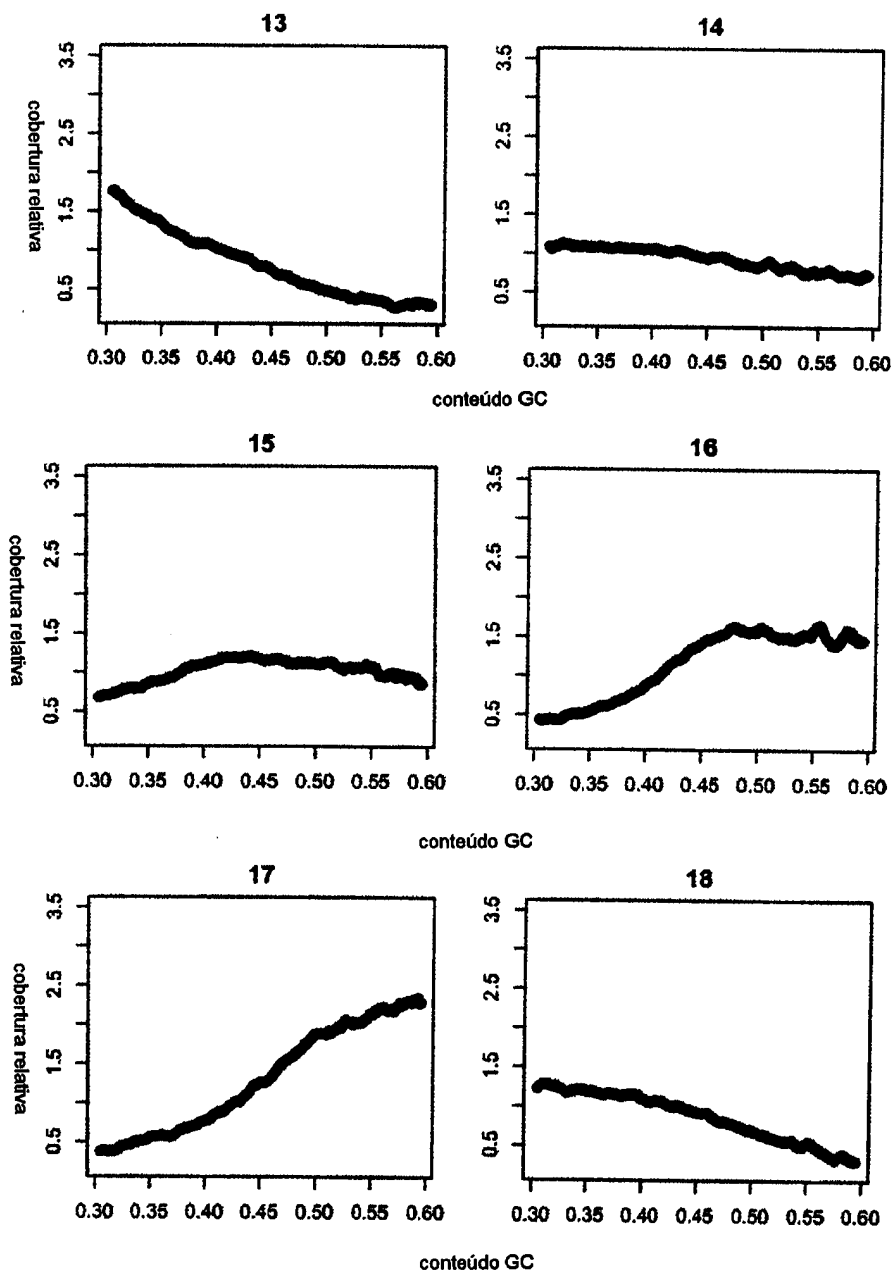


Figura 5

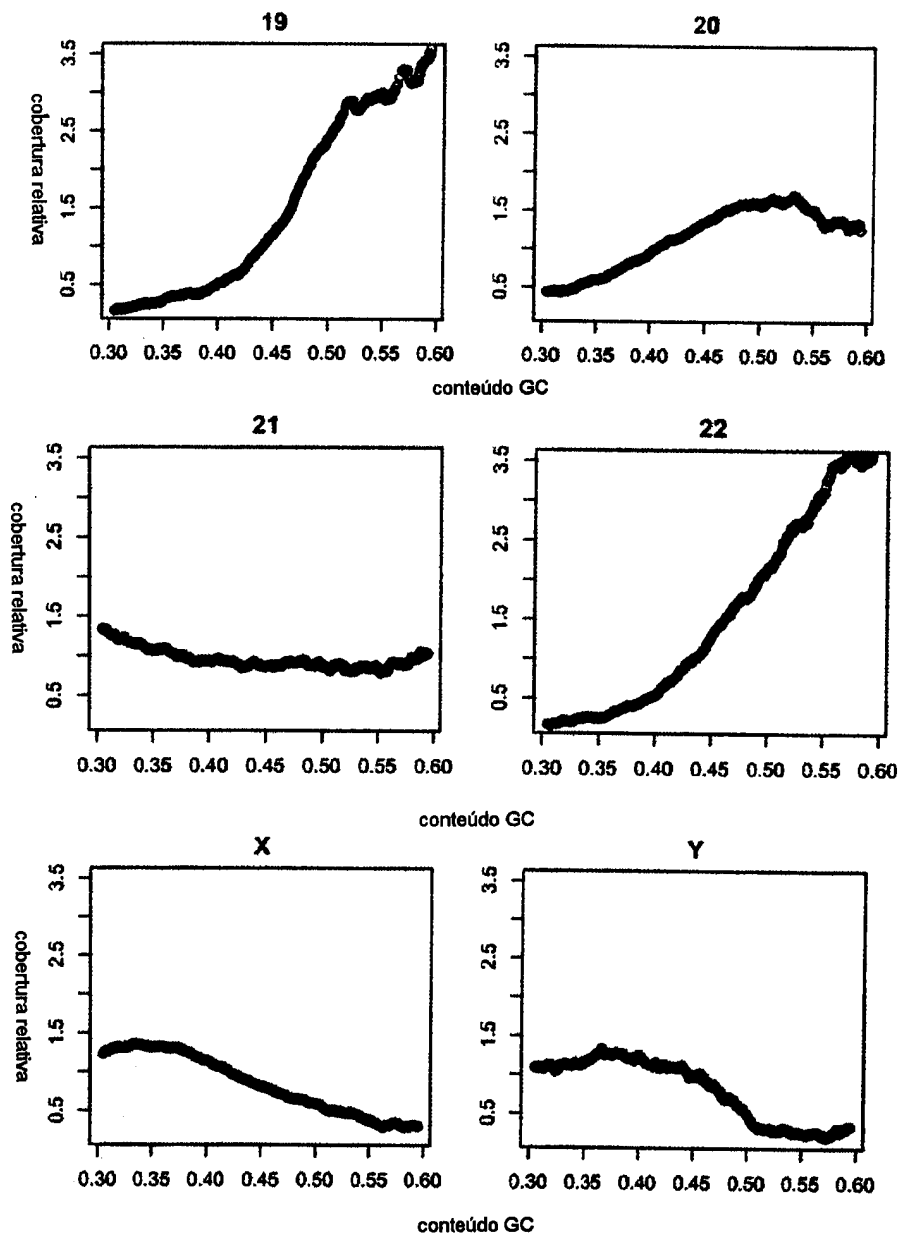


Figura 5

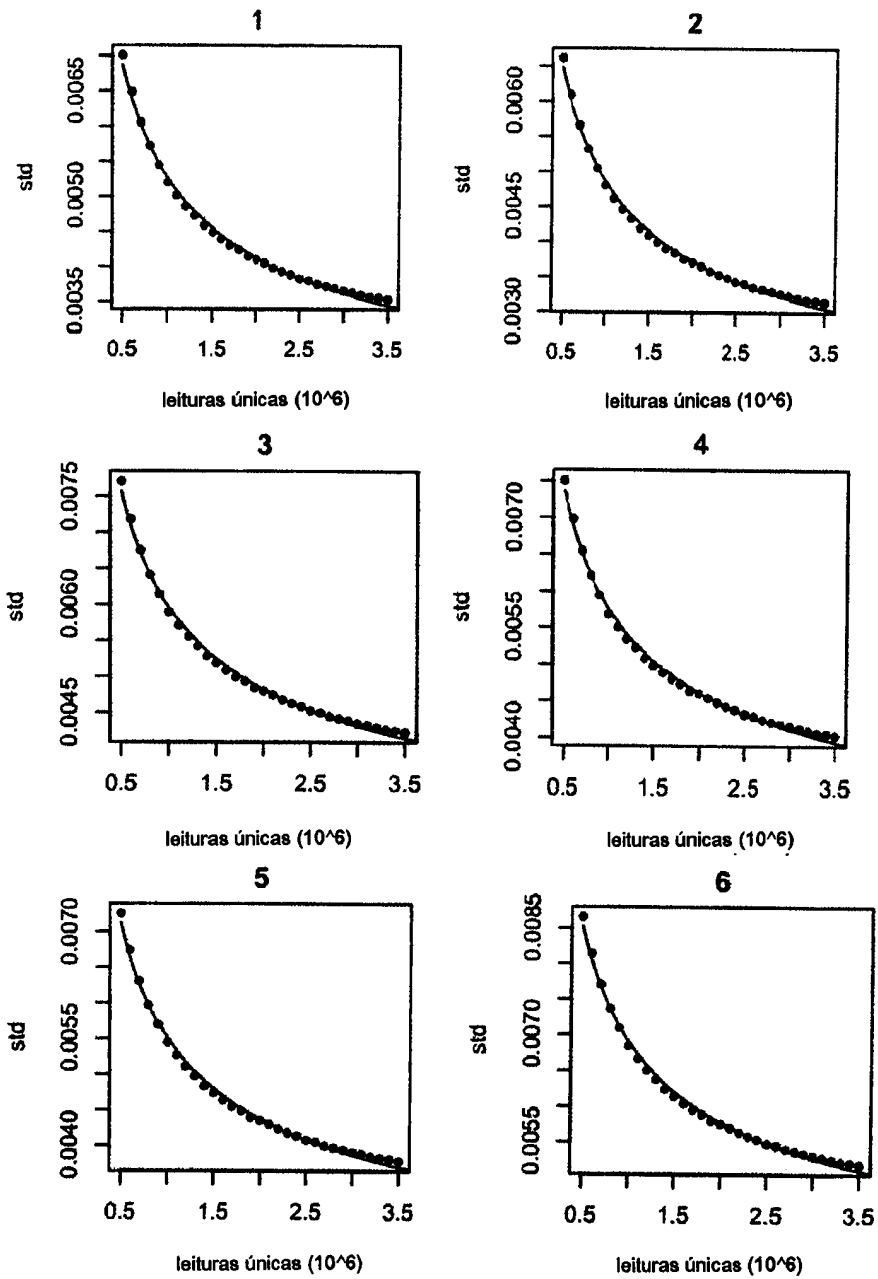


Figura 6

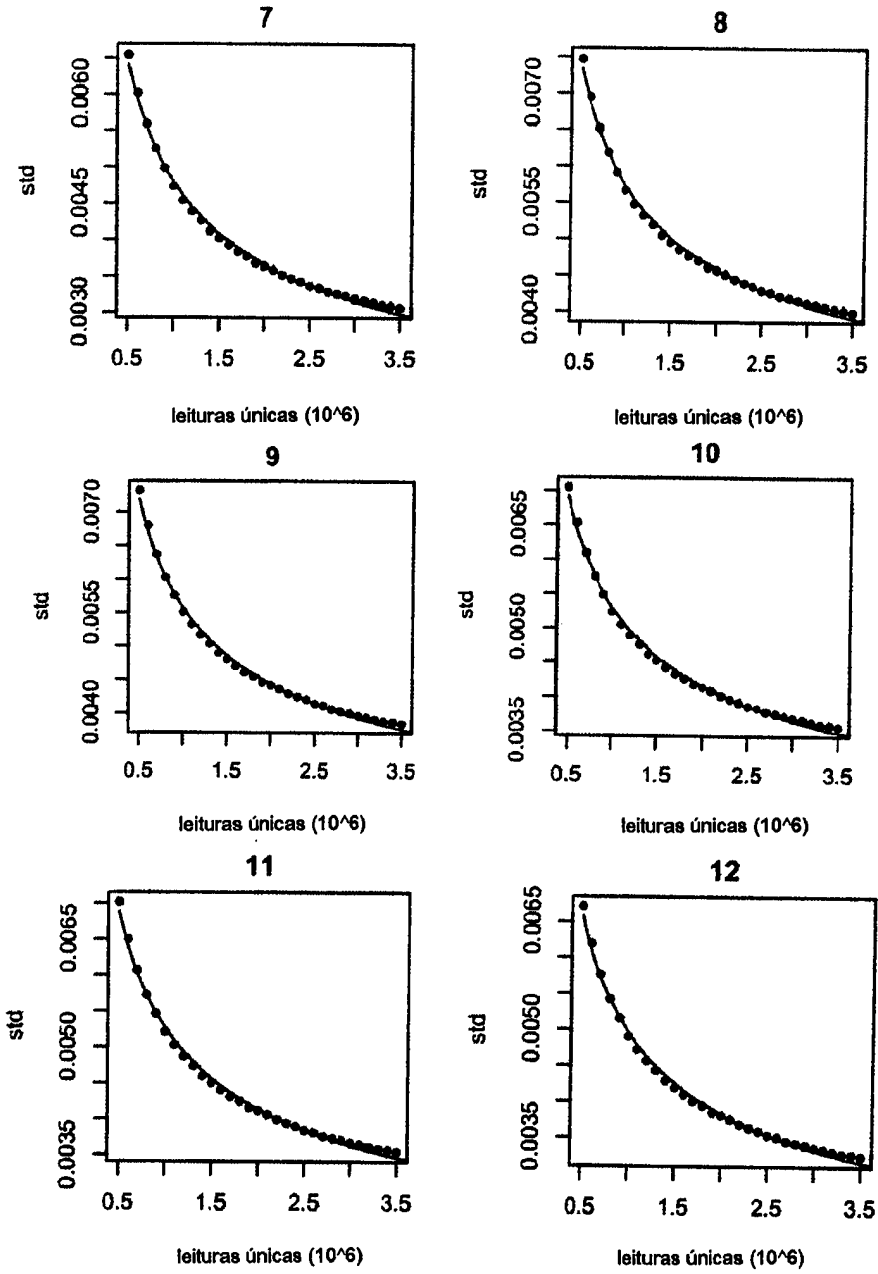


Figura 6

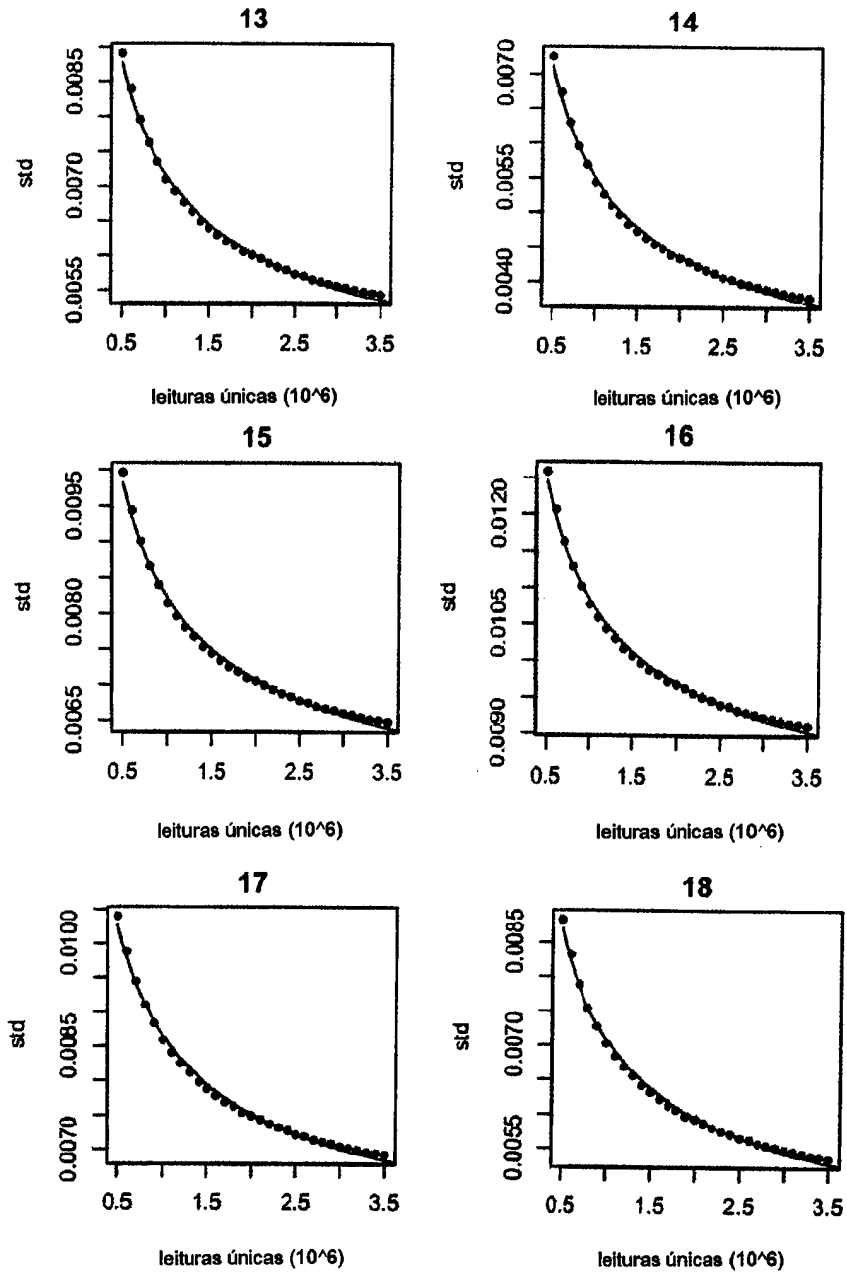


Figura 6

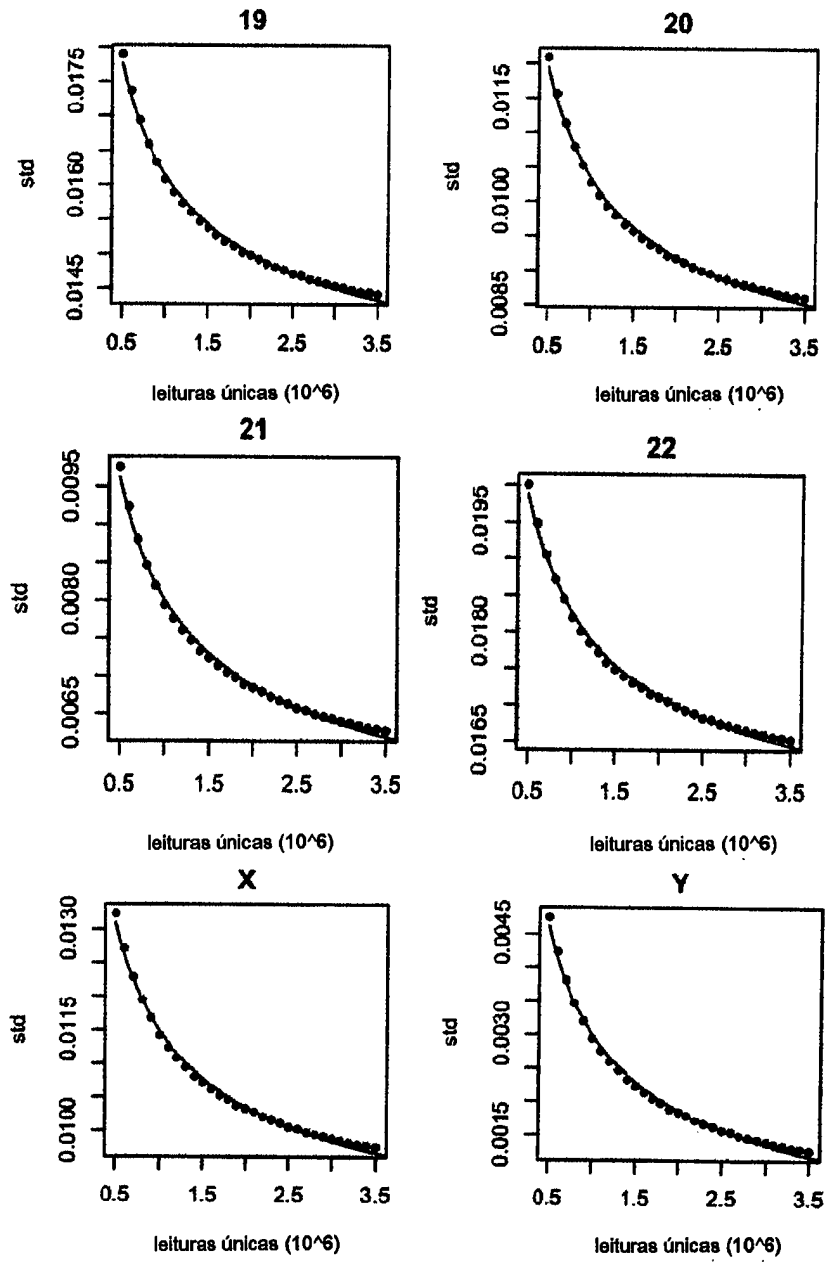


Figura 6

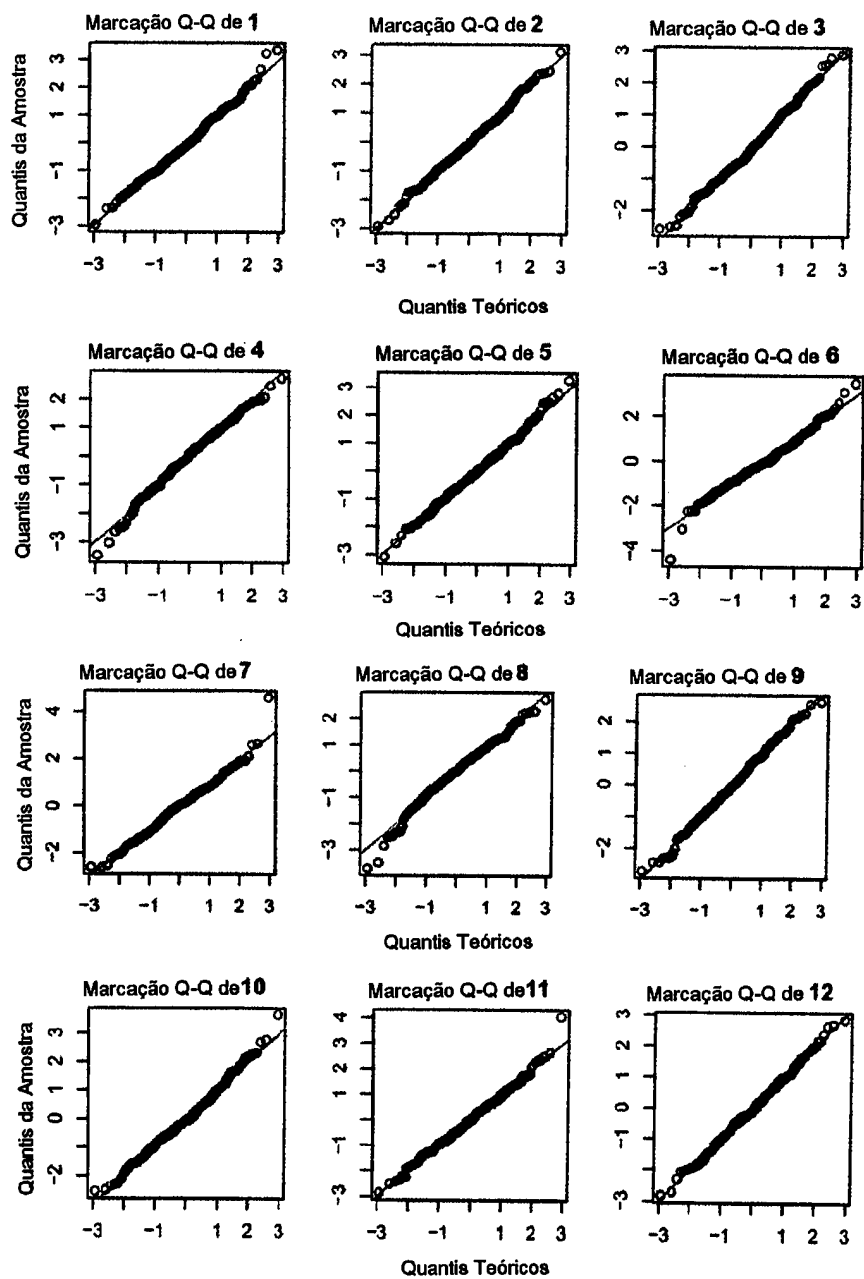


Figura 7

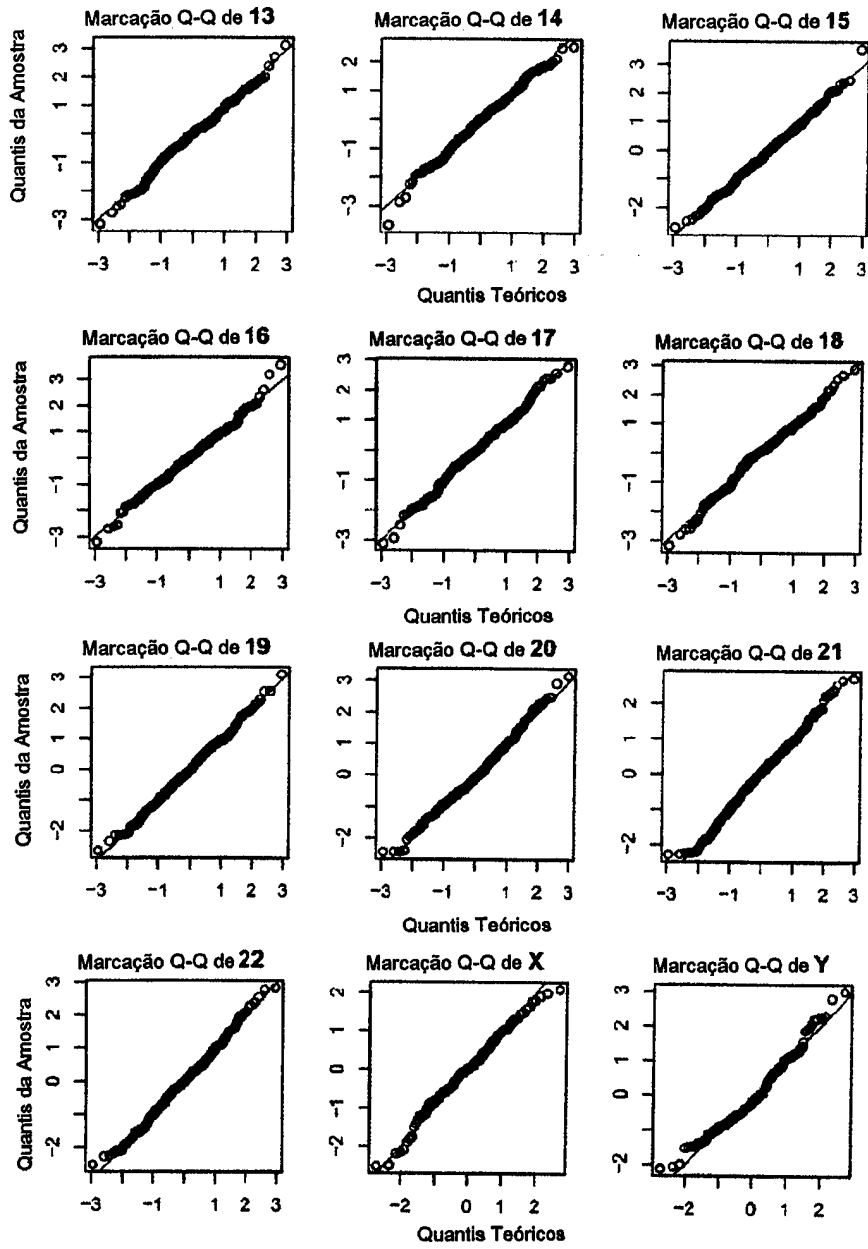


Figura 7

Histograma de cr relativo de chrY

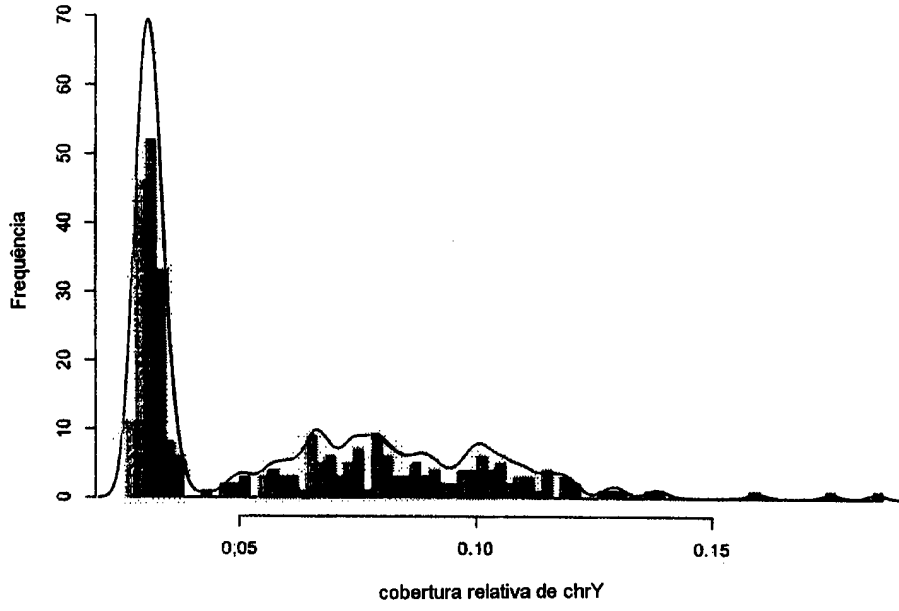


Figura 8

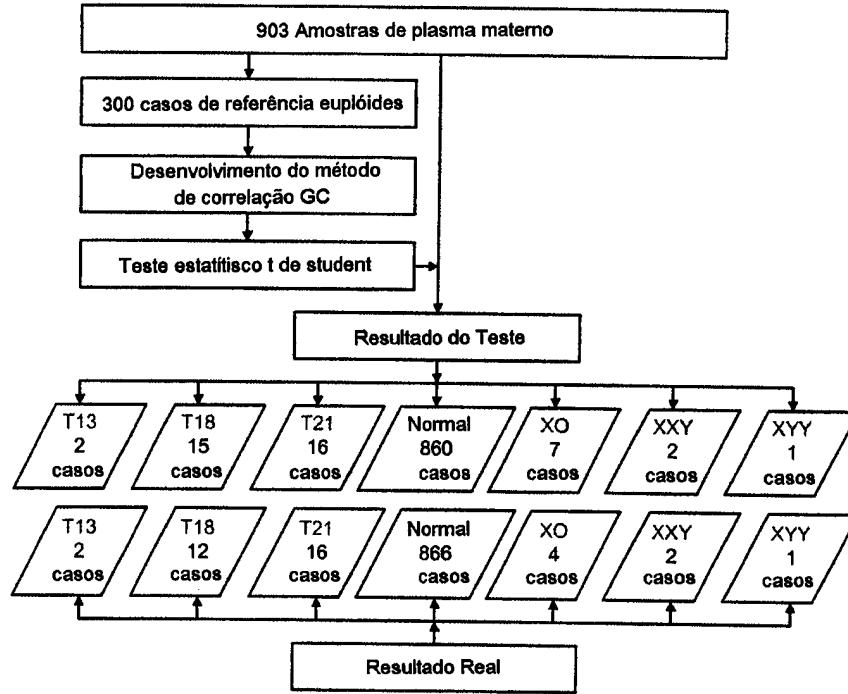


Figura 9

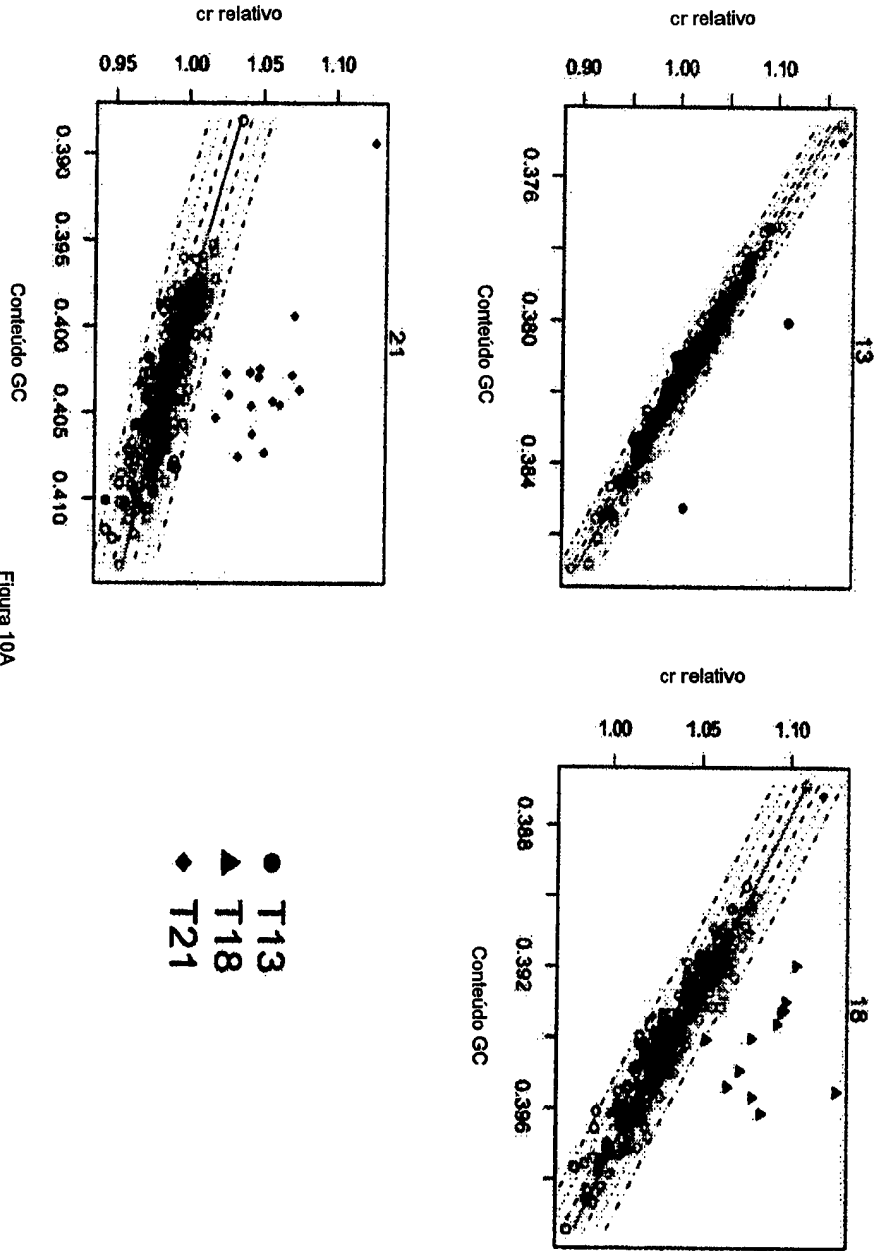


Figura 10A

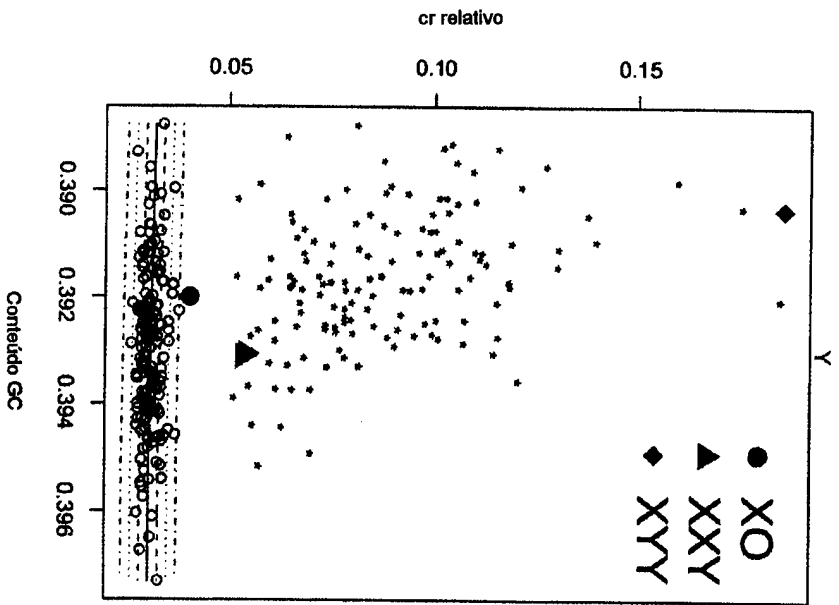
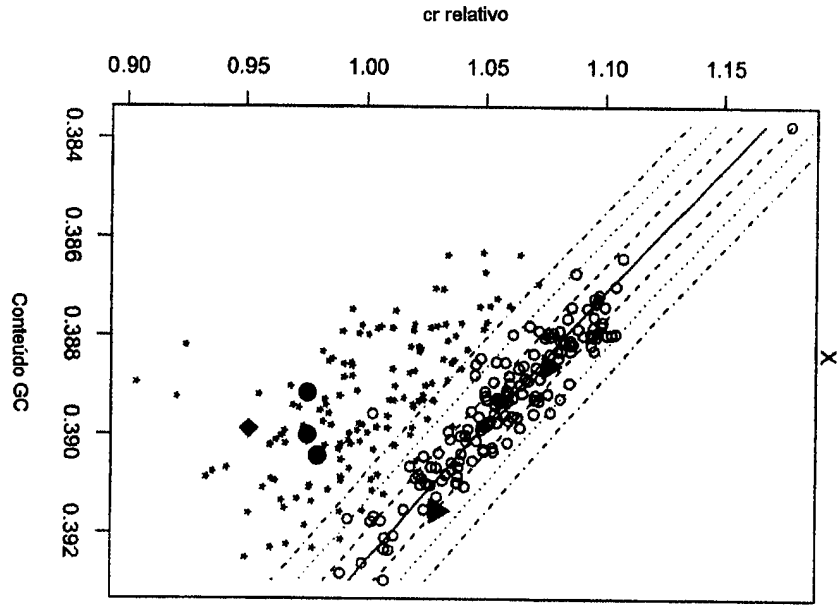


Figura 10B

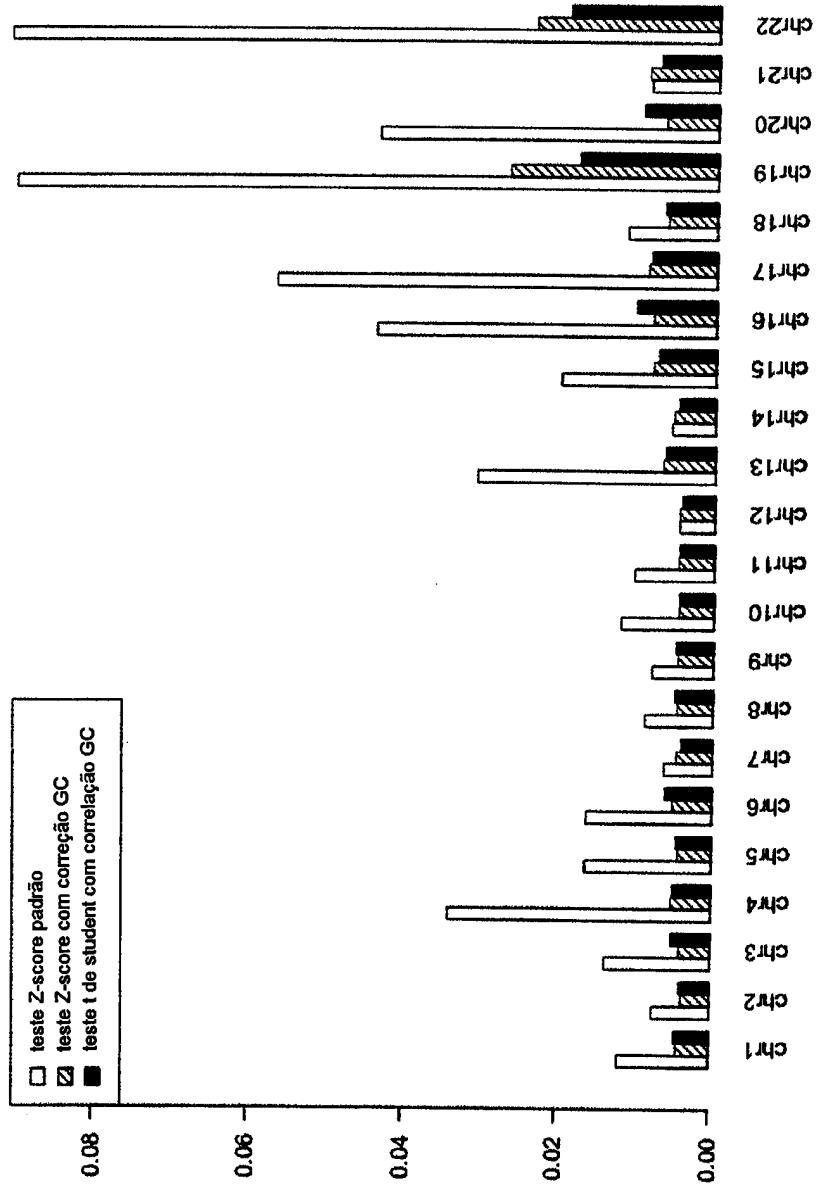


Figura 11

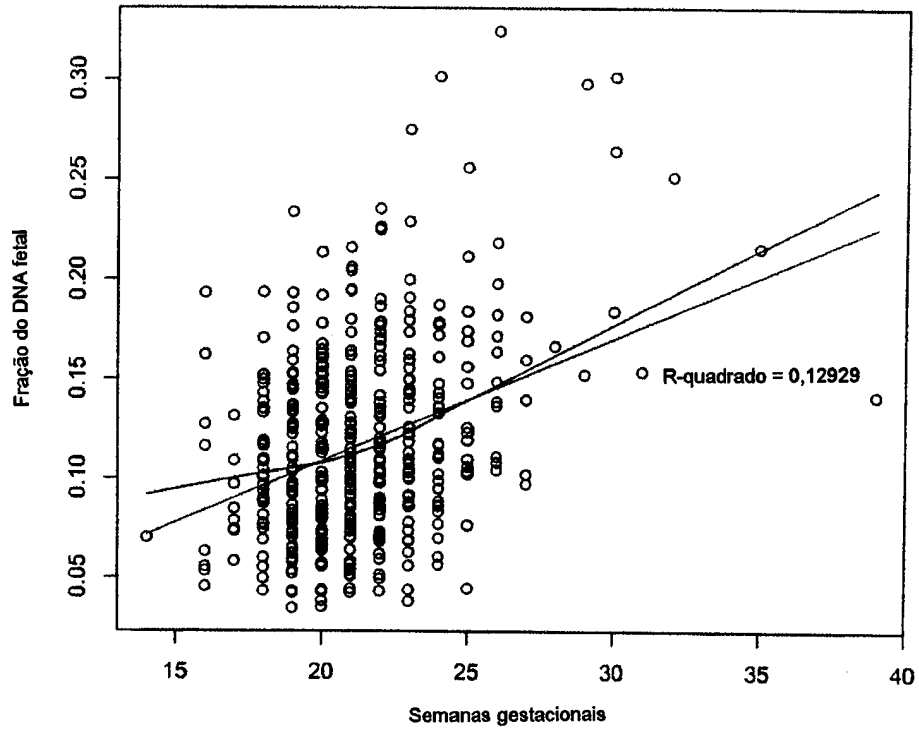


Figura 12

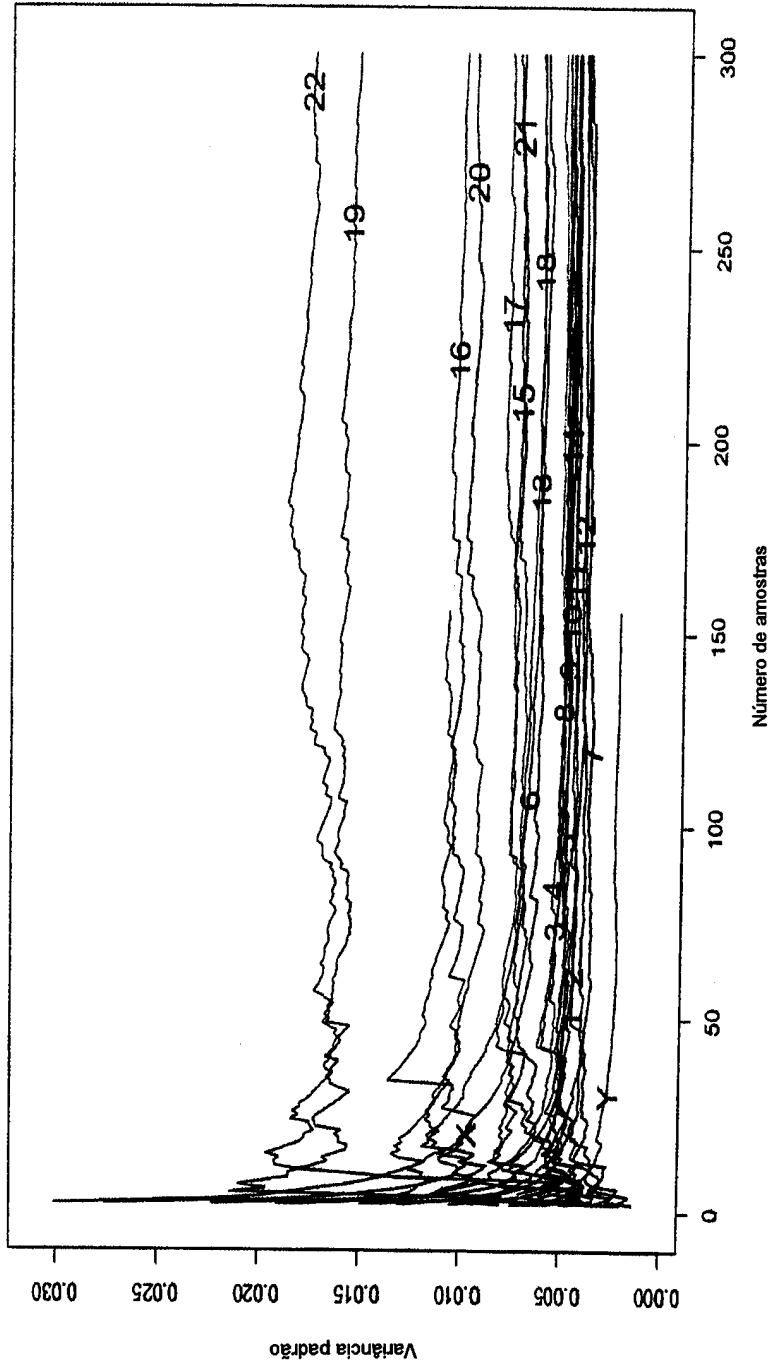
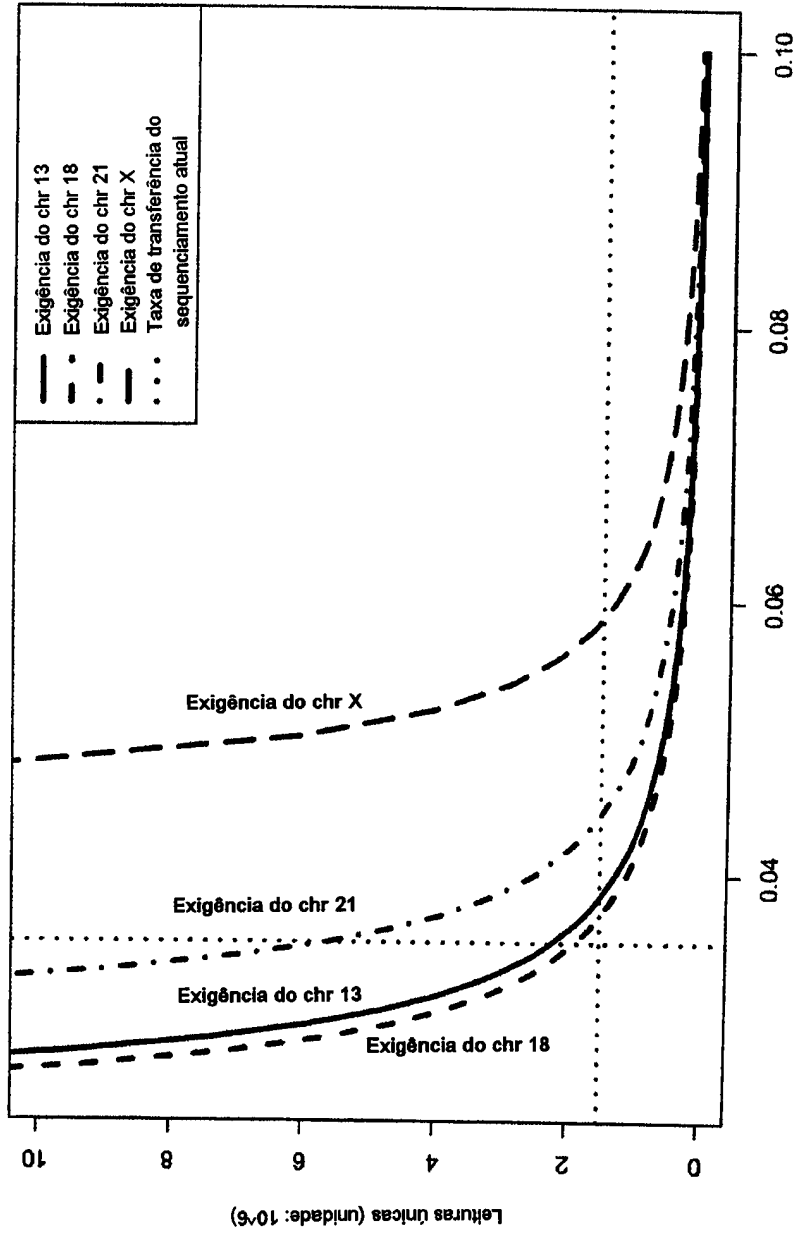


Figura 13



Fração de DNA fetal  
Figura 14

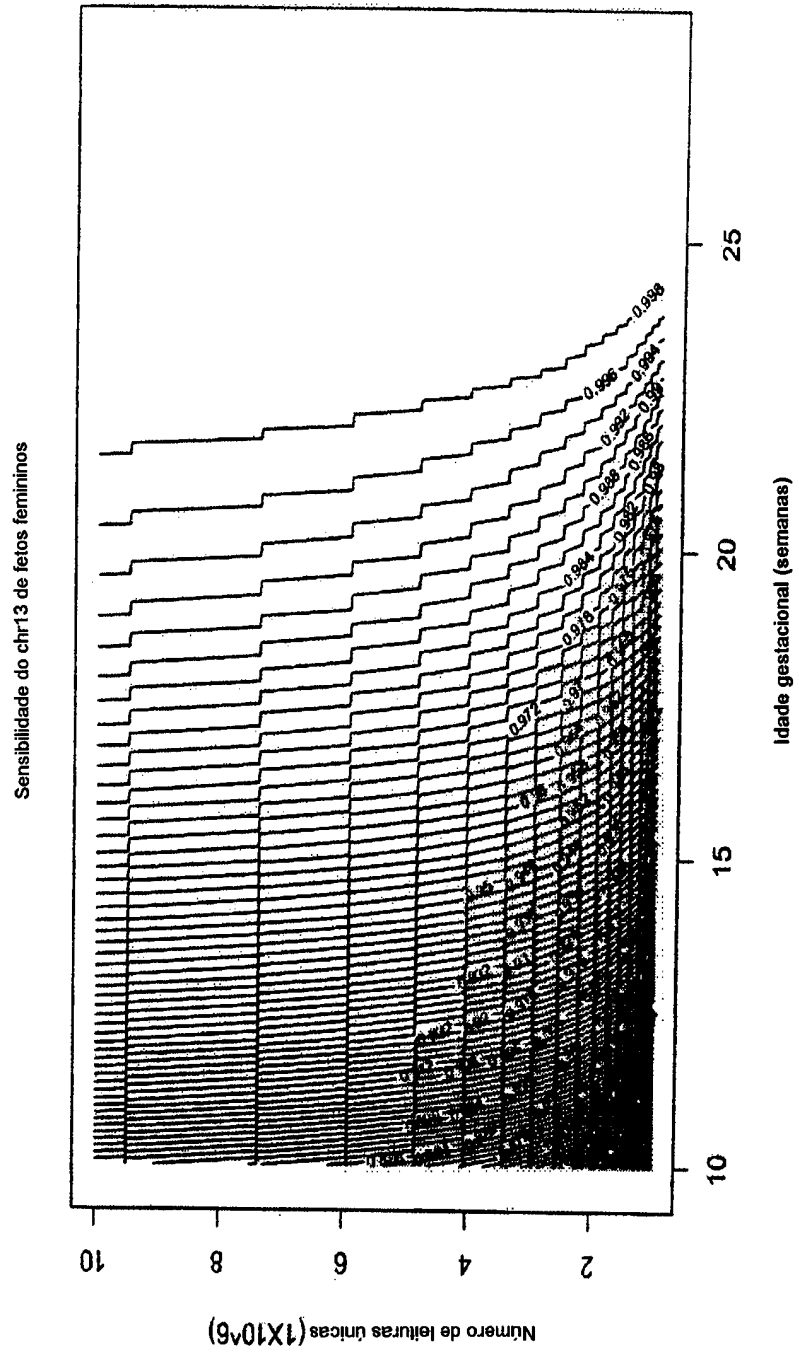


Figura 15

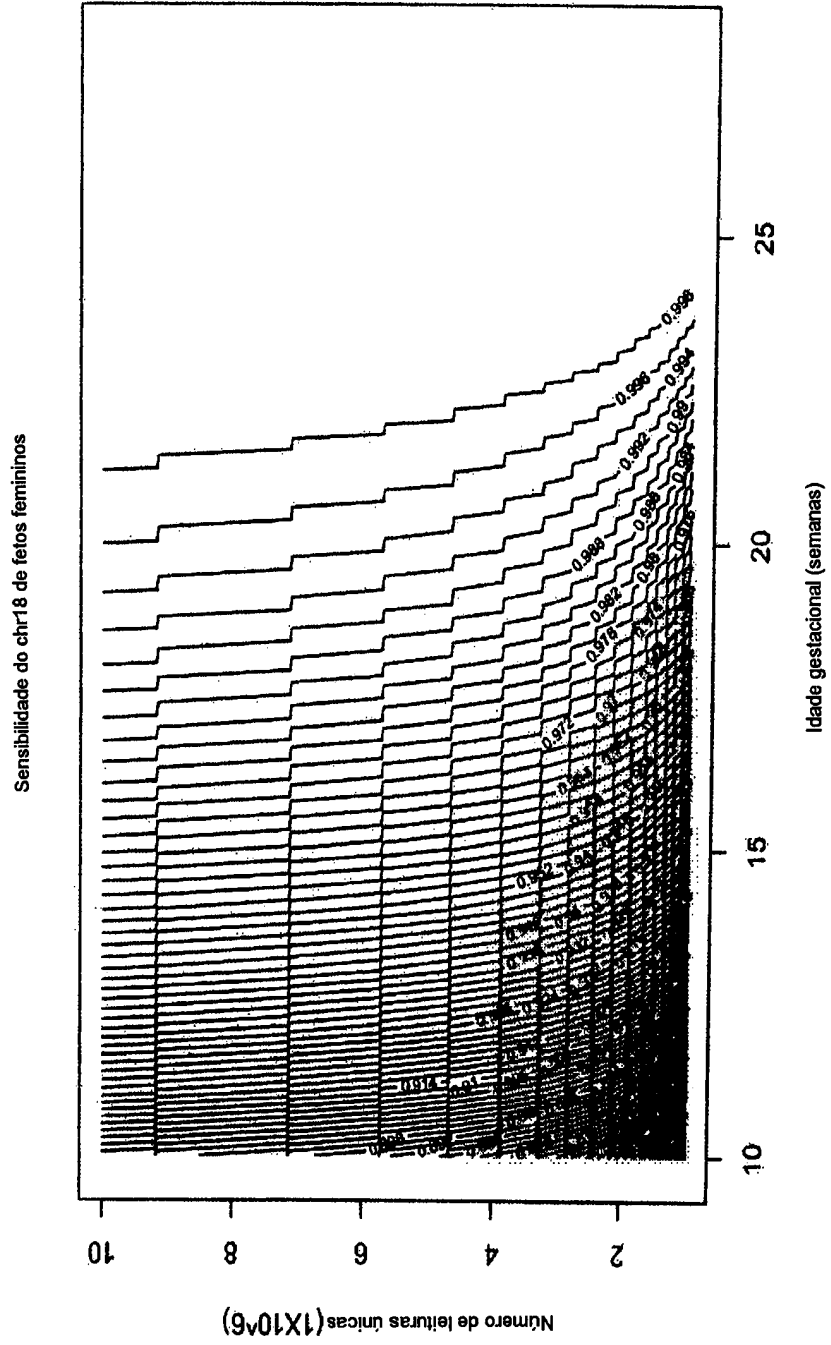
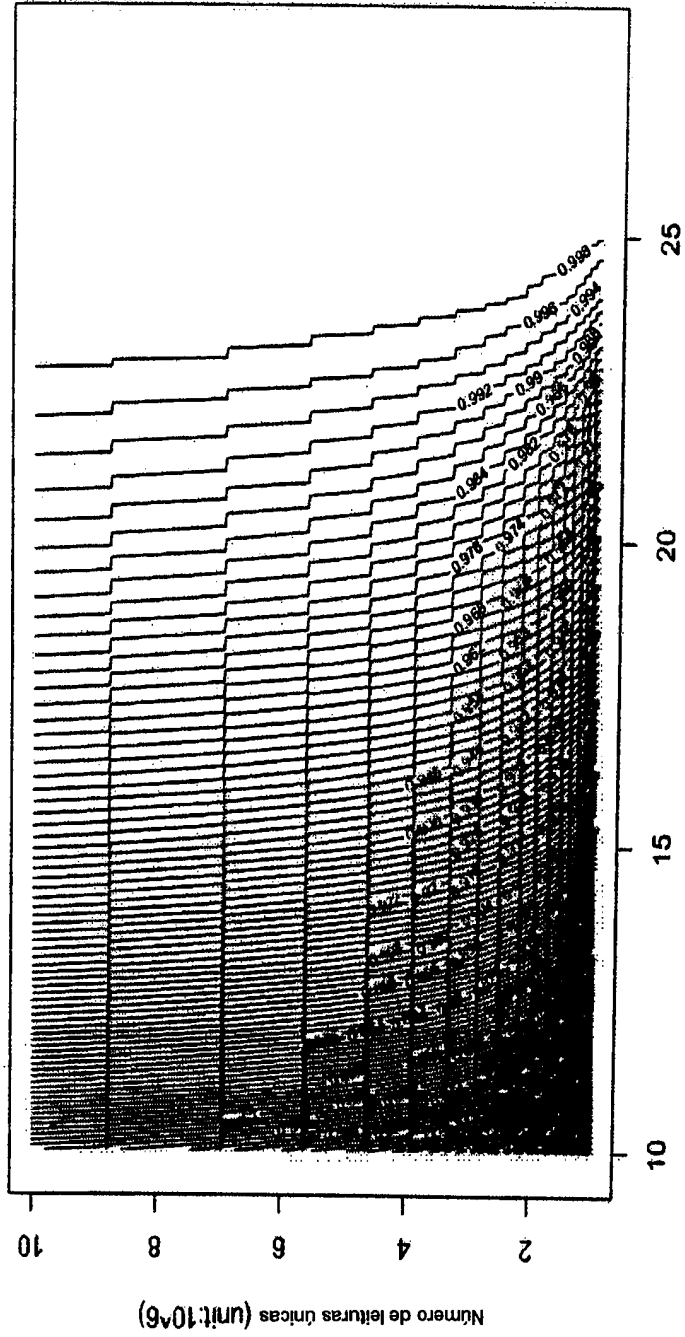


Figura 16

Sensibilidade do chr21 de fetos femininos



Idade gestacional (semanas)

Figura 17

Sensibilidade do chrX: XO, XXX, XXY, etc.

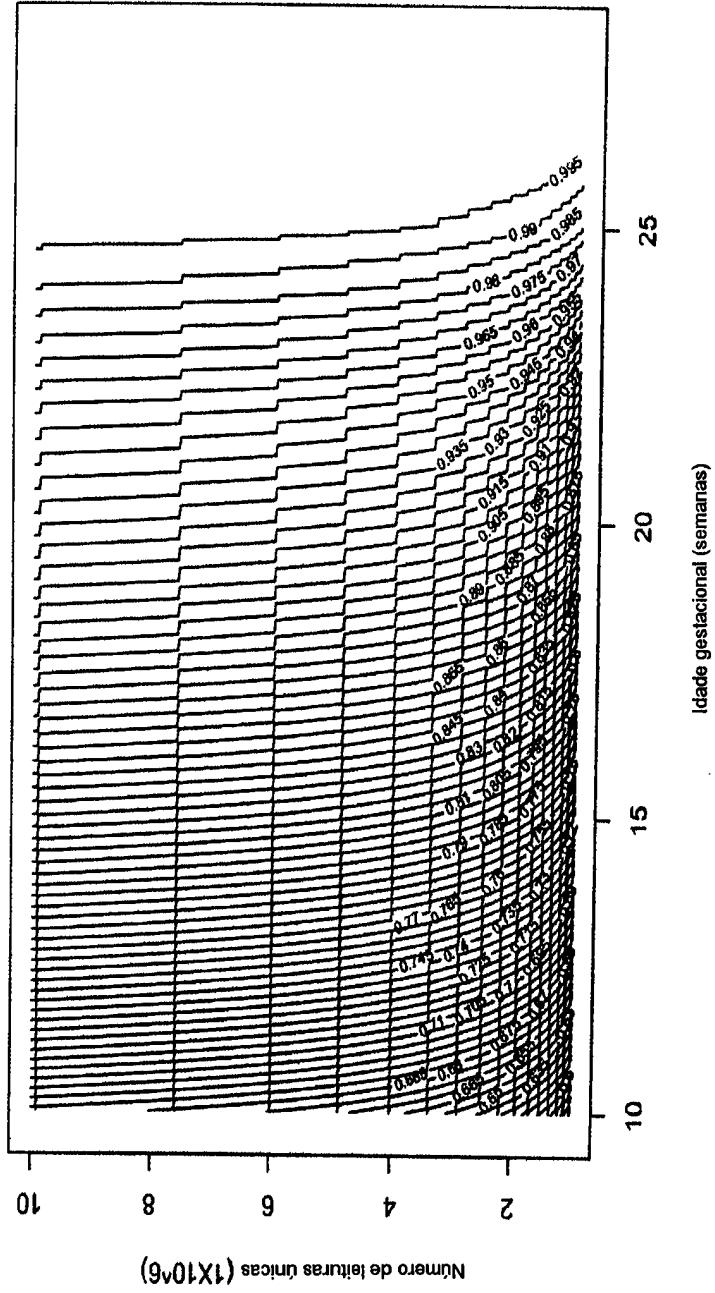


Figura 18





