

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 512 448**

51 Int. Cl.:

G06F 19/18 (2011.01)

C12Q 1/68 (2006.01)

G06F 19/22 (2011.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **29.06.2011 E 11863253 (8)**

97 Fecha y número de publicación de la concesión europea: **27.08.2014 EP 2561103**

54 Título: **Detección no invasiva de anomalías genéticas fetales**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
24.10.2014

73 Titular/es:

BGI DIAGNOSIS CO., LTD. (100.0%)
2F, 3F, 11 Buildings, Beishan Industrial Zone,
Beishan Road 146, Yantian District, Shenzhen
Guangdong 518083, CN

72 Inventor/es:

JIANG, FUMAN;
CHEN, HUIFEI;
CHAI, XIANGHUA;
YUAN, YUYING;
ZHANG, XIUQING y
CHEN, FANG

74 Agente/Representante:

LAZCANO GAINZA, Jesús

ES 2 512 448 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

DESCRIPCIÓN

Detección no invasiva de anomalías genéticas fetales

Campo técnico

5 La invención se relaciona con métodos no invasivos para la detección de anomalías genéticas fetales, más particularmente aneuploidía cromosómica, mediante secuenciación de ADN de muestras de sangre periférica tomadas de mujeres embarazadas que contienen ADN tanto materno como fetal. Más particularmente, esta invención se relaciona con el análisis de datos que tiene en cuenta el sesgo de GC introducido por amplificación y secuenciación de muestras de ADN a través de bases sobre relaciones preestablecidas entre el contenido de GC y la profundidad de cobertura para un cromosoma en ausencia de aneuploidía y en el tamaño de fragmentación de ADN seleccionado como se discute en más detalle más adelante. Puede utilizarse análisis estadístico para comparar la profundidad de cobertura calculada para un cromosoma en una muestra con respecto a la profundidad de cobertura esperada de acuerdo con tal relación establecida con lo cual puede lograrse la detección de aneuploidía cromosómica fetal.

Técnica antecedente

15 Los métodos de diagnóstico prenatales convencionales con procedimientos invasivos, tales como muestreo de villos coriónicos y amniocentesis, conllevan riesgos potenciales tanto para los fetos como para las madres. La selección no invasiva de aneuploidía fetal utilizando marcadores en el suero materno y ultrasonido está disponible, pero tiene sensibilidad y especificidad limitadas (Kagan, et al., Human Reproduction (2008) 23:1968-1975; Malone, et al., N Engl J Med (2005) 353:2001-2011).

20 Estudios recientes han demostrado que la detección no invasiva de aneuploidía fetal mediante secuenciación paralela masiva de moléculas de ADN en el plasma de mujeres embarazadas es factible. El ADN fetal ha sido detectado y cuantificado en plasma y suero maternos (Lo, et al., Lancet (1997) 350:485-487; Lo, et al., Am. J. hum. Genet. (1998) 62:768-775). Se presentan múltiples tipos de células fetales en la circulación materna, incluyendo granulocitos, linfocitos, glóbulos rojos nucleados y células de trofoblastos fetales (Pertl y Bianchi, Obstetrics and Gynecology (2001) 98:483-490). El ADN fetal puede ser detectado en el suero en la séptima semana de gestación, y se incrementa con el término del embarazo. El ADN fetal presente en el suero y en el plasma materno es comparable a la concentración de ADN obtenida de los protocolos de aislamiento de células fetales.

25 El ADN fetal en circulación ha sido utilizado para determinar el sexo del feto (Lo, et al., Am. J. hum. Genet. (1998) 62:768-775). También se ha detectado el genotipo D de Rhesus fetal utilizando ADN fetal. Sin embargo, el diagnóstico y las aplicaciones clínicas del ADN fetal en circulación están limitados a genes que están presentes en el feto pero no en la madre (Pertl y Bianchi, Obstetrics and Gynecology (2001) 98:483-490). Así, existe una necesidad por un método no invasivo que pueda determinar la secuencia del ADN fetal y proveer un diagnóstico definitivo de las anomalías cromosómicas en un feto.

30 El descubrimiento de las células fetales y de los ácidos nucleicos fetales libres de células en material sanguíneo en las últimas décadas y la aplicación de la secuenciación biológica de alto rendimiento de ADN libre de células en plasma materno lo hace disponible para detectar pequeños cambios en la representación de cromosomas aportados por un feto aneuploide en una muestra de plasma materno. Se ha alcanzado la detección no invasiva de embarazos por trisomía 13, 18 y 21.

35 Sin embargo, como lo muestran algunos estudios, el sesgo de GC introducido por la amplificación y secuenciación ha puesto un límite práctico sobre la sensibilidad de la detección de aneuploidía. El sesgo de GC podría ser introducido durante la preparación de la muestra y el proceso de secuenciación, bajo diferentes condiciones tales como composición de reactivos, densidad y temperatura de la aglomeración, lo cual lleva a un muestreo diferencial de las moléculas de ADN con diferente composición de GC y variación significativa en los datos de secuenciación para los cromosomas que son ricos en GC o pobres en GC.

40 Para mejorar la sensibilidad, se han desarrollado protocolos para la eliminación del efecto del sesgo de GC. Fan y Quake desarrollaron un método para eliminar computacionalmente el sesgo de GC aplicando peso a cada densidad de GC con base en el contenido GC genómico local, para mejorar el número de lecturas mapeadas en cada conjunto basura multiplicando el peso correspondiente (Fan y Quake PLoS ONE (2010) 5: e10439). Sin embargo, el método tiene dificultades al tratar con trastornos en cromosomas sexuales, especialmente trastornos relevantes del cromosoma Y por la razón de que el proceso puede producir una ligera distorsión de datos la cual interfiere con la precisión de la detección.

45 Aquí, hemos descubierto un método que hace uso del sesgo de GC con el fin de obtener una alta sensibilidad en la detección de aneuploidía fetal así como para evitar la distorsión de los datos. Este método permite la definición de parámetros usados para pruebas estadísticas de acuerdo con el contenido de GC. Además, los inventores introdujeron la fracción fetal estimada en el diagnóstico mediante una hipótesis binaria la cual muestra sensibilidad y

especificidad más altas. Nuestro método también muestra que debe ser posible incrementar la sensibilidad de la detección no invasiva de anomalías genéticas fetales para fijar la precisión para muestras maternas que contienen una baja fracción de ADN fetal secuenciando más fragmentos de polinucleótidos. El remuestreo del plasma materno en las últimas semanas de la gestación también puede incrementar la sensibilidad del diagnóstico.

5 Resumen de la invención

La presente invención está dirigida a métodos para la detección no invasiva de aneuploidía cromosómica fetal en los cuales se lleva a cabo una secuenciación a gran escala de nucleótidos a partir de muestras de sangre periférica materna. En un aspecto, la presente invención provee un método implementado por ordenador para determinar una anomalía genética fetal la cual es una aneuploidía cromosómica o aneuploidía parcial, método que comprende:

10 (a) obtener la información de secuencia de múltiples fragmentos de polinucleótidos a partir de una muestra, siendo dicha muestra una muestra de sangre periférica derivada de un sujeto femenino en embarazo y que contiene ADN tanto materno como fetal;

15 (b) asignar dichos fragmentos a cromosomas con base en dicha información de secuencia comparando dichos fragmentos con las lecturas únicas de referencia del mismo tamaño para cada uno de dichos cromosomas, en donde las lecturas únicas de referencia son fragmentos de un cromosoma que tiene una secuencia única la cual puede ser asignada de manera no ambigua a una localización cromosómica individual con base en una secuencia genómica de referencia;

20 (c) determinar la profundidad de cobertura y el contenido de GC de un cromosoma con base en la información de secuencia para aquellos fragmentos que han sido asignados a las lecturas únicas de referencia de dicho cromosoma en la etapa (b), en donde la profundidad de cobertura es la relación entre el número de fragmentos asignados de manera única a dicho cromosoma y el número de lecturas únicas de referencia para dicho cromosoma del mismo tamaño del fragmento con base en dicha secuencia genómica humana de referencia;

25 (d) determinar la profundidad de cobertura ajustada de dicho cromosoma utilizando dicho contenido de GC de dicho cromosoma y la relación establecida entre la profundidad de cobertura y el contenido de GC para dicho cromosoma en la ausencia de aneuploidía, en donde dicha relación establecida ha sido determinada por un método que comprende las etapas de:

30 (i) obtener información de secuencia de múltiples fragmentos de polinucleótidos que cubren dicho cromosoma a partir de una pluralidad de muestras de sangre periférica euploide que contienen ADN genómico, en donde el tamaño del fragmento es el mismo que el tamaño del fragmento de los fragmentos de polinucleótido múltiples de la etapa (a) anterior;

(ii) asignar dichos fragmentos a cromosomas con base en dicha información de secuencia como en la etapa (b) más arriba;

(iii) determinar la profundidad de cobertura y el contenido de GC de dicho cromosoma con base en dicha información de secuencia para cada muestra euploide como en la etapa (c) más arriba; y

35 (iv) utilizar la profundidad de cobertura y el contenido de GC determinados para cada muestra en la etapa (iii) para determinar la relación entre la profundidad de cobertura y el contenido de GC de dicho cromosoma en la ausencia de aneuploidía; y

(e) comparar dicha profundidad de cobertura ajustada con la profundidad de cobertura de dicho cromosoma determinada en la etapa (c), en donde una diferencia entre ellos indica aneuploidía cromosómica fetal.

40 En una realización los fragmentos de polinucleótidos varían desde aproximadamente 10 hasta aproximadamente 1000 bp en longitud. En otra realización, los fragmentos de polinucleótidos varían desde aproximadamente 15 hasta aproximadamente 500 bp en longitud. En aún otra realización los fragmentos de polinucleótidos varían desde aproximadamente 20 hasta aproximadamente 200 bp en longitud. En todavía otra realización los fragmentos de polinucleótidos varían desde aproximadamente 25 hasta aproximadamente 100 bp en longitud. En una realización
45 adicional los fragmentos de polinucleótidos tienen aproximadamente 35 bp de longitud.

En una realización, la información de secuencia es obtenida por secuenciación genómica paralela. En otra realización la asignación del fragmento a los cromosomas se hace comparando la secuencia de los fragmentos con una secuencia genómica humana de referencia. La secuencia genómica humana de referencia puede ser cualquier construcción de genoma humano adecuada y/o publicada, tal como hg18 o hg19. Se descartan los fragmentos que
50 se asignan a más de un cromosoma o no se asignan a ningún cromosoma.

Como se indicó anteriormente, la profundidad de cobertura de un cromosoma para el propósito del método de la invención es la relación entre el número de fragmentos que se asignan únicamente al cromosoma y el número de

lecturas únicas de referencia del cromosoma. En otra realización, la profundidad de la cobertura está normalizada. En todavía otra realización, la normalización se calcula contra la cobertura de todos los otros autosomas. En aún otra realización, la normalización se calcula contra la cobertura de todos los otros cromosomas.

En una realización, la relación está en la fórmula:

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

5
 en donde $f(GC_{i,j})$ representa la función de la relación entre la profundidad de cobertura normalizada y el correspondiente contenido de GC de la muestra i , cromosoma j , $\varepsilon_{i,j}$ representa el residual de la muestra i , cromosoma j . En otras realizaciones, la relación entre la profundidad de cobertura y el contenido de GC se calcula mediante regresión polinómica local. En algunas realizaciones, la relación puede ser una relación lineal no fuerte. En algunas realizaciones, la relación se determina mediante algoritmo Loess.

En algunas realizaciones, el método comprende adicionalmente calcular la profundidad de cobertura ajustada de acuerdo con la fórmula:

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

15
 En algunas realizaciones, el método comprende adicionalmente calcular la variación estándar de acuerdo con la fórmula:

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

en donde ns representa el número de muestras de referencia.

En algunas realizaciones, el método comprende adicionalmente calcular la estadística t de Student de acuerdo con la fórmula:

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

20
 En una realización, el contenido de GC en un cromosoma es el contenido de GC promedio de todos los fragmentos que están asignados al cromosoma para el propósito de la etapa (c) del método de la invención como se definió más arriba. El contenido de GC de un fragmento puede ser calculado dividiendo el número de nucleótidos G/C en el fragmento por el número total de nucleótidos del fragmento. En otra realización, el contenido de GC de un cromosoma es el contenido de GC agregado de las lecturas únicas de referencia del cromosoma.

Al menos 2, 5, 10, 20, 50, 100, 200, 500 o 1000 muestras euploides pueden ser utilizadas con el fin de determinar la relación entre la profundidad de cobertura y el contenido de GC para un cromosoma en la ausencia de aneuploidía. El cromosoma puede ser cualquiera de los cromosomas 1, 2, ..., 22, X o Y.

30
 [0019] Las muestras para este propósito pueden ser de sujetos femeninos en estado de embarazo. Las muestras pueden ser de sujetos masculinos. Alternativamente, las muestras pueden ser tanto de sujetos femeninos en estado de embarazo como de sujetos masculinos. Las muestras son muestras de sangre periférica.

En algunas realizaciones, un método de la invención para determinar una aneuploidía cromosómica fetal comprende adicionalmente la determinación del género del feto. El género del feto puede ser determinado de acuerdo con la fórmula:

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

en donde $cr.a_{i,x}$ y $cr.a_{i,y}$ están normalizadas con respecto a la profundidad de cobertura de los cromosomas X y Y, respectivamente.

5 En algunas realizaciones, el método comprende adicionalmente estimar la fracción fetal. La fracción fetal puede ser calculada de acuerdo con la fórmula:

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

10 en donde $\hat{c}r_{i,Yf} = f(GC_{i,Yf})$ es la profundidad de cobertura ajustada calculada a partir de la profundidad de cobertura del cromosoma Y y el correspondiente contenido de GC de muestras de mujeres embarazadas con un feto hembra, $\hat{c}r_{i,Ym} = f(GC_{i,Ym})$ se refiere a la profundidad de cobertura ajustada calculada a partir de la relación de la profundidad de cobertura del cromosoma Y y el contenido correspondiente de GC de sujetos masculinos. Alternativamente, la fracción fetal puede ser calculada de acuerdo con la fórmula:

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

15 en donde $\hat{c}r_{i,Xf} = f(GC_{i,Xf})$ es la profundidad de cobertura ajustada calculada a partir de la relación de la profundidad de cobertura del cromosoma X y el contenido correspondiente de GC de muestras de mujeres embarazadas con un feto hembra, $\hat{c}r_{i,Xm} = f(GC_{i,Xm})$ se refiere a la profundidad de cobertura ajustada calculada a partir de la relación de la profundidad de la cobertura del cromosoma X y el contenido correspondiente de GC de muestras de sujetos masculinos. Adicionalmente, la fracción fetal puede ser calculada de acuerdo con la fórmula:

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

20 en donde $\hat{c}r_{i,Xf} = f(GC_{i,Xf})$ es la profundidad de cobertura ajustada calculada a partir de la relación de la profundidad de cobertura del cromosoma X y el contenido correspondiente de GC de muestras de mujeres embarazadas con un feto hembra, $\hat{c}r_{i,Yf} = f(GC_{i,Yf})$ se refiere a la profundidad de cobertura ajustada calculada a partir de la relación de la profundidad de cobertura del cromosoma Y y el contenido correspondiente de GC de muestras de mujeres embarazadas con un feto hembra, $\hat{c}r_{i,Xm} = f(GC_{i,Xm})$ se refiere a la profundidad de cobertura ajustada calculada a partir de la relación de la profundidad de cobertura de cromosoma X y el contenido correspondiente de GC de muestras de sujetos masculinos, $\hat{c}r_{i,Ym} = f(GC_{i,Ym})$ se refiere a la profundidad de cobertura ajustada calculada a partir de la relación de la profundidad de la cobertura del cromosoma Y y el contenido correspondiente de GC de sujetos masculinos.

30 La aneuploidía fetal que va a ser detectada puede ser un trastorno para un autosoma seleccionado del grupo consistente de la trisomía 13, 18 y 21. En aún otra realización, la aneuploidía fetal que va a ser detectada puede ser un trastorno para un cromosoma sexual seleccionado del grupo consistente de XO, XXX, XXY y XYY. En algunas realizaciones, la comparación de dicha profundidad de cobertura ajustada con dicha profundidad de cobertura del cromosoma se lleva a cabo mediante una prueba de hipótesis estadística, en donde una hipótesis es que el feto es euploide (H0) y la otra hipótesis es que el feto es aneuploide (H1). Puede calcularse una estadística a partir de ambas hipótesis. En algunas realizaciones, se calcula la estadística t de Student para H0 y H1 de acuerdo con la fórmula: $t1_{ij} = (cr_{i,j} - \hat{c}r_{i,j}) / std_j$ y $t2_{ij} = (cr_{i,j} - \hat{c}r_{i,j}(1 + fxy_{ij}/2)) / std_j$, respectivamente, en donde fxy es la fracción fetal. En algunas realizaciones, la relación probable logarítmica de t1 y t2 se calcula de acuerdo con la fórmula: $L_{ij} = \log(p(t1_{ij}, \text{grado}|D)) / \log(p(t2_{ij}, \text{grado}|T))$, en donde grado se refiere a un grado de distribución t, D se refiere a diploidía, T se refiere a trisomía, y p (T1_{ij}, grado|*), * = D, T representa la densidad de probabilidad condicional dada a un grado de distribución t.

En una realización, el género fetal es femenino, y la estadística t de Student se calcula de acuerdo con la fórmula $t_{1,X} = (cr_{i,X} - \hat{c}r_{i,Xf}) / std_{Xf}$, en donde $\hat{c}r_{i,Xf} = f(GC_{i,Xf})$ es la profundidad de cobertura ajustada calculada a partir de la relación de la profundidad de la cobertura del cromosoma X y el contenido correspondiente de GC de muestras a partir de mujeres embarazadas con un feto femenino. En algunas realizaciones, $|t1| > 3.13$ indica que el feto puede ser XXX o XO. En algunas realizaciones, $|t1| > 5$ indica que el feto es XXX o XO.

En otra realización, el género del feto es masculino, y la estadística t de Student se calcula de acuerdo con la fórmula: $t_{2,Y} = (cr_{i,X} - (1 - fy_i / 2) \cdot \hat{c}r_{i,Xf}) / std_{Xf}$, en donde $\hat{c}r_{i,Xf} = f(GC_{i,Xf})$ es la profundidad de cobertura ajustada calculada a partir de la relación de la profundidad de la cobertura del cromosoma X y el contenido correspondiente de GC de muestras a partir de mujeres embarazadas con un feto femenino. En algunas realizaciones, $|t2| > 3.13$ indica que el feto puede ser XXY o XYY. En algunas realizaciones, $|t2| > 5$ indica que el feto es XXY o XYY.

En otro aspecto, se provee aquí un medio legible por ordenador que comprende una pluralidad de instrucciones adaptadas para ejecutar un método de la invención para el diagnóstico prenatal de una anomalía genética fetal cuando dichas instrucciones son suplementadas con información de secuencia de fragmentos múltiples de polinucleótidos a partir de una muestra de sangre periférica derivada de un sujeto femenino en embarazo y que contiene ADN tanto materno como fetal. El medio legible por ordenador: b) asignará dichos fragmentos de polinucleótidos a cromosomas con base en dicha información de secuencia; c) determinará la profundidad de cobertura y el contenido de GC de un cromosoma con base en la información de secuencia para aquellos fragmentos asignados únicamente a dicho cromosoma; d) determinará la profundidad de cobertura ajustada de dicho cromosoma utilizando dicho contenido de GC de dicho cromosoma y establecerá una relación entre la profundidad de cobertura y el contenido de GC de dicho cromosoma; y e) comparará dicha profundidad de cobertura con dicha profundidad de cobertura de dicho cromosoma, en donde una diferencia entre ellos indica aneuploidía cromosómica fetal.

En todavía otro aspecto, se provee aquí un sistema que comprende medios adaptados para llevar a cabo un método de la invención. Tal sistema comprende: a) medios para obtener información de secuencia de múltiples fragmentos de polinucleótidos a partir de una muestra; y b) un medio legible por ordenador como se describió anteriormente.

Breve descripción de los dibujos

La Figura 1 muestra un proceso esquemático para calcular la profundidad de cobertura y el contenido de GC utilizando información de secuencia de fragmentos de polinucleótidos.

La Figura 2 ilustra la correlación entre profundidad de cobertura normalizada-contenido de GC establecida utilizando datos a partir de 300 casos de referencia. La profundidad de cobertura normalizada para cada caso se representa gráficamente contra el contenido de GC secuenciado correspondiente. Las cruces denotan casos con feto femenino euploide, los cuadrados denotan casos con feto masculino euploide. Las líneas sólidas son las líneas de ajuste de la profundidad de cobertura y el contenido de GC.

La Figura 3 ilustra la tendencia entre la profundidad de cobertura normalizada y el correspondiente contenido de GC disponiendo los cromosomas con su contenido de GC ascendente inherente. El contenido de GC ascendente inherente de cada cromosoma se refiere aquí al contenido de GC promedio de etiquetas secuenciadas de ese cromosoma a partir de 300 casos de referencia.

La Figura 4 muestra diferentes composiciones de clase GC para cada cromosoma. El contenido de GC de cada lectura de 35 bp de las lecturas únicas de referencia se calculó para cada cromosoma, el contenido de GC fue clasificado en 36 niveles y el porcentaje de cada nivel fue calculado como la composición en GC de cada cromosoma. Los cromosomas fueron entonces representados gráficamente mediante el mapa de calor y aglomerados jerárquicamente.

La Figura 5 demuestra que la desviación dla secuenciación introduce la correlación mostrada en la Figura 2 mediante la simulación manual del proceso de preferencia del secuenciador.

La Figura 6 representa gráficamente la variación estándar contra el número total de fragmentos de polinucleótidos secuenciados. En 150 muestras, la varianza estándar ajustada de cada cromosoma muestra una relación lineal con el recíproco de la raíz cuadrada del número de lecturas únicas.

La Figura 7 muestra una gráfica Q-Q del residual de cada cromosoma calculado mediante la Fórmula 3. Se muestra una relación lineal con una distribución normal.

La Figura 8 muestra el histograma de profundidad de cobertura del cromosoma Y. Hay dos picos lo cual implica que el género de los casos puede ser distinguido por la profundidad de cobertura del cromosoma Y. La curva es la distribución de la profundidad de cobertura relativa del cromosoma Y estimada por la estimación de la densidad núcleo con el núcleo Gaussiano.

- La Figura 9 muestra un diagrama del proceso para el diagnóstico de 903 muestras de prueba para anomalía cromosómica fetal.
- La Figura 10 muestra el resultado de la aneuploidía: casos de trisomía 13, 18, 21 y XO, XXY, XYY y casos normales. La Figura 10A muestra la gráfica de profundidad de cobertura normalizada versus contenido de GC de los cromosomas 13, 18 y 21. La Figura 10B muestra las gráficas de los cromosomas X y Y. Los círculos representan profundidad de cobertura relativa normal de fetos femeninos con contenido de GC, los puntos representan fetos masculinos normales. La línea sólida es la línea de ajuste de la cobertura relativa y el contenido de GC, la línea de rayas es el absoluto del valor t_1 , las líneas punteadas son el absoluto del valor t_2 y las líneas de puntos-rayas son el absoluto del valor t_3 .
- La Figura 11 compara los valores de confianza de diferentes metodologías de diagnóstico.
- La Figura 12 muestra la relación entre la fracción de ADN fetal y la edad gestacional. La fracción del ADN fetal en plasma materno se correlaciona con la edad gestacional. La fracción de ADN fetal fue estimada por X y Y juntos. Hay una correlación significativa estadísticamente entre la fracción de ADN fetal promedio y la edad gestacional ($P < 0.001$). Nótese que el valor R^2 representa el cuadrado del coeficiente de correlación pequeño. La fracción mínima es 3.49%.
- La Figura 13 muestra la relación entre la varianza estándar con el número de casos requeridos para la detección. Las varianzas estándar calculadas por la fórmula 5 de cada cromosoma varían con diferentes números de muestras. La varianza estándar se hace estable cuando el número de muestras es superior a 100.
- La Figura 14 muestra el número estimado de lecturas únicas para la detección de aneuploidía fetal en plasma libre de células como función de la fracción de ADN fetal. Los estimados están basados en el nivel de valor t de confianza no inferior a 3 para aneuploidía de cromosomas 13, 18, 21, y X, incluso Y (a partir de la relación entre X y Y) teniendo cada uno diferente longitud. A medida que la fracción de ADN fetal disminuye, el número total de secuencias por biolística requeridas se incrementa. Con un resultado de secuenciación de 4 millones de lecturas de secuencia por canal en la celda de flujo, la trisomía 21 puede ser detectada si el 3.5% del ADN libre de células es fetal. La aneuploidía del cromosoma X no fue detectada fácilmente cuando la fracción y el número de lecturas únicas fueron pequeños, tal como 4% y 5 millones de lecturas. Cromosomas diferentes requieren diferentes niveles de fracción de ADN fetal y un número de lecturas único, el cual puede ser causado por la estructura de GC del cromosoma.
- La Figura 15 muestra una gráfica de contornos de sensibilidad mapeada mediante volúmenes de datos y edad gestacional (semanas) para la detección de la trisomía del cromosoma 13 para fetos femeninos, para cada semana gestacional y cada punto de volumen de datos.
- La Figura 16 muestra una gráfica de contornos de sensibilidad mapeada por volumen de datos y edad gestacional (semanas) para la detección de trisomía del cromosoma 18 para fetos femeninos, para cada semana gestacional y cada punto de volumen de datos.
- La Figura 17 muestra una gráfica de contornos de sensibilidad mapeada por volumen de datos y edad gestacional (semanas) para la detección de la trisomía del cromosoma 21 para fetos femeninos, para cada semana gestacional y cada punto de volumen de datos.
- La Figura 18 muestra una gráfica de contornos de sensibilidad mapeada por volumen de datos y edad gestacional (semanas) para la detección de trisomía del cromosoma X para fetos femeninos, para cada semana gestacional y cada punto de volumen de datos.
- La Figura 19 muestra una gráfica de contornos de sensibilidad mapeada por volumen de datos y edad gestacional (semanas) para la detección de la trisomía del cromosoma 13 de un sujeto masculino. Para cada semana gestacional y cada punto de volumen de datos, calculamos su distribución empírica de la fracción de ADN fetal y la varianza estándar para cada volumen de datos primeramente, y comparando la fracción estimada por XY o Y calculamos entonces la sensibilidad de cada tipo de aneuploidía.
- La Figura 20 muestra una gráfica de contornos de sensibilidad mapeada por volumen de datos y edad gestacional (semanas) para la detección de trisomía del cromosoma 18 de un sujeto masculino.
- La Figura 21 muestra una gráfica de contornos de sensibilidad mapeada por volumen de datos y edad gestacional (semanas) para la detección de trisomía del cromosoma 21 de un sujeto masculino.
- Descripción detallada de la invención
- La presente invención está dirigida a métodos para detección no invasiva de anomalías genéticas fetales mediante una secuenciación a gran escala de fragmentos de polinucleótidos a partir de una muestra de sangre

periférica materna. Como se indicó anteriormente, se proveen métodos que tienen en cuenta el sesgo de GC de los resultados de secuenciación que surgen de la diferencia en el contenido de GC de un cromosoma con base en la relación entre la profundidad de cobertura de un cromosoma y el contenido correspondiente de GC. De acuerdo con lo anterior, se describe aquí un método para ajustar computacionalmente los parámetros de referencia que están siendo usados en un cálculo de t de Student con contenidos de GC pesando localmente la regresión polinómica para ajustar la profundidad de cobertura de un cromosoma en muestras contra el contenido de GC de los fragmentos de polinucleótido.

También se provee aquí un método para determinar la anormalidad cromosómica de un feto tal como se describió anteriormente involucrando análisis estadístico mediante el uso de una prueba de hipótesis estadística. Además, se describen métodos para calcular estándares para control de calidad de los datos (DQC) útiles en la determinación de la cantidad de muestras clínicas necesarias para un cierto nivel de significado estadístico.

I. Definiciones

A menos que se defina de otra manera, todos los términos técnicos y científicos utilizados aquí tienen el mismo significado que se entiende comúnmente por parte de una persona de experiencia normal en la técnica a la cual pertenece esta invención. Todas las patentes, solicitudes, solicitudes publicadas y otras publicaciones referidas aquí se incorporan como referencia en su totalidad. Si una definición fijada en esta sección es contraria o de alguna manera inconsistente con una definición fijada en las patentes, solicitudes, solicitudes publicadas y otras publicaciones que se incorporan aquí como referencia, la definición fijada en esta sección prevalece sobre la definición que se incorpora aquí como referencia.

Tal como se utilizan aquí, las formas singulares "un/una", y "el, la, los, las" incluyen referencias plurales a menos que se indique otra cosa. Por ejemplo, "un" dímero incluye uno o más dímeros.

El término "anormalidad cromosómica" se refiere a una desviación entre la estructura del cromosoma sujeto y un cromosoma homólogo normal. El término "normal" se refiere al cariotipo predominante o patrón de bandas encontrado en individuos saludables de una especie particular. Una anormalidad cromosómica puede ser numérica o estructural, e incluye pero no se limita a aneuploide, poliploide, inversión, una trisomía, una monosomía, duplicación, eliminación, eliminación de una parte de un cromosoma, adición, adición de una parte de un cromosoma, inserción, un fragmento de un cromosoma, una región de un cromosoma, reordenamiento cromosómico, y translocación. Una anormalidad cromosómica puede estar correlacionada con la presencia de una condición patológica o con una predisposición a desarrollar una condición patológica. Tal como se define aquí, un polimorfismo de nucleótido individual ("SNP") no es una anormalidad cromosómica.

La monosomía X (XO, ausencia de un cromosoma X completo) es el tipo más común del síndrome de Turner, que se presenta en 1 de 2500 hasta 1 a 3000 niñas nacidas vivas (Sybert and McCauley N Engl J Med (2004) 351:1227-1238). El síndrome XXY es una condición en la cual los sujetos masculinos humanos tienen un cromosoma X extra, existiendo en apenas 1 de cada 1000 sujetos masculinos (Bock, Understanding Klinefelter Syndrome: A Guide for XXY Males and Their Families. NIH Pub. No. 93-3202 (1993)). El síndrome XYY es una aneuploidía de los cromosomas sexuales en los cuales un sujeto masculino humano recibe un cromosoma Y extra, dando un total de 47 cromosomas en vez de los más usuales 46, afectando a 1 de cada 1000 nacimientos masculinos, mientras que potencialmente lleva a infertilidad masculina (Aksglaede, et al., J Clin Endocrinol Metab (2008) 93:169-176).

El síndrome de Turner abarca varias condiciones, de las cuales la monosomía X (XO, ausencia de un cromosoma sexual completo, el cuerpo de Barr) es la más común. Los sujetos femeninos típicos tienen dos cromosomas X, pero en el síndrome de Turner, uno de estos cromosomas sexuales falta. Presentándose en 1 de 2000 a 1 de 5000 sujetos femeninos fenotípicos, el síndrome se manifiesta así mismo de numerosas maneras. El síndrome de Klinefelter es una condición en la cual los sujetos masculinos humanos tienen un cromosoma X extra. En humanos, el síndrome de Klinefelter es el trastorno de cromosomas sexuales más común y la segunda condición más común causada por la presencia de cromosomas extra. La condición existe en apenas 1 de cada 1000 sujetos masculinos. El síndrome XYY es una aneuploidía de los cromosomas sexuales en los cuales un sujeto masculino humano recibe un cromosoma Y extra, dando un total de 47 cromosomas en vez de los más usuales 46. Esto produce un cariotipo XYY 47. Esta condición usualmente es asintomática y afecta a 1 de cada 1000 nacimientos masculinos, mientras que potencialmente lleva a infertilidad masculina.

La trisomía 13 (síndrome de Patau), la trisomía 18 (síndrome de Edward) y la trisomía 21 (síndrome de Down) son las trisomías autosómicas más clínicamente importantes y cómo detectarlas siempre ha sido un tópico candente. La detección de la aberración cromosómica fetal anterior tiene gran significado en el diagnóstico prenatal (Ostler, Diseases of the eye and skin: a color atlas. Lippincott Williams & Wilkins. pp. 72. ISBN 9780781749992 (2004); Driscoll and Gross N Engl J Med (2009) 360: 2556-2562; Kagan, et al., Human Reproduction (2008) 23:1968-1975).

El término " lecturas únicas de referencia" se refiere a fragmentos de un cromosoma que tienen una secuencia única. Por lo tanto, tales fragmentos pueden ser asignados de manera no ambigua a una localización cromosómica

individual. Las lecturas únicas de referencia de un cromosoma pueden ser construidas con base en una secuencia genómica de referencia publicada, tal como hg18 o hg19.

5 Los términos "polinucleótido", "oligonucleótido", "ácido nucleico" y "molécula de ácido nucleico" se utilizan de manera intercambiable aquí para referirse a una forma polimérica de nucleótidos de cualquier longitud, y pueden comprender
 10 ribonucleótidos, desoxirribonucleótidos, análogos de los mismos o mezclas de los mismos. Este término se refiere solamente a la estructura primaria de la molécula. Así, el término incluye ácido desoxirribonucleico ("ADN") de cadena triple, doble y sencilla, así como ácido ribonucleico ("ARN") de cadena triple, doble y sencilla. También incluye formas modificadas, por ejemplo por alquilación, y/o por recubrimiento y no modificadas del polinucleótido. Más particularmente, los términos "polinucleótido", "oligonucleótido", "ácido nucleico" y "molécula de ácido nucleico"
 15 incluyen polidesoxirribonucleótidos (que contienen 2-desoxi-D-ribosa), polirribonucleótidos (que contienen D-ribosa), incluyendo ARNt, ARNr, ARNh, y ARNm, bien sea divididos o no divididos, y cualquier otro tipo de polinucleótido el cual es un N- o C-glicósido de una base purínica o pirimidínica, y otros polímeros que contienen esqueletos no nucleotídicos, por ejemplo, poliamida (por ejemplo, ácidos nucleicos peptídicos ("APNs")) y polímeros polimorfolino (comercialmente disponibles en Anti-Virals, Inc., Corvallis, OR., como NeuGene®) y otros polímeros de ácidos
 20 nucleicos sintéticos específicos para la secuencia considerando que los polímeros contienen nucleobases en una configuración que permite el apareamiento de bases y el apilamiento de bases, tal como se encuentra en el ADN y el ARN. Así, estos términos incluyen, por ejemplo, 3'-desoxi-2',5'-DNA, oligodesoxirribonucleótido N3' a P5' fosforamidatos, ARN sustituido con 2'-O-alquilo, híbridos entre ADN y ARN, o entre APN y ADN o ARN, y también incluyen tipos conocidos de modificaciones, por ejemplo, etiquetas, alquilación, "caperuzas", sustitución de uno o más de los nucleótidos con un análogo, modificaciones de internucleótidos tales como, por ejemplo, aquellas con enlaces no cargados (por ejemplo, metil fosfonatos, fosfotriésteres, fosforamidatos, carbamatos, etc.), con enlaces cargados negativamente (por ejemplo, fosforotioatos, fosforoditioatos, etc.), y con enlaces cargados positivamente (por ejemplo, aminoalquilfosforamidatos, aminoalquilfosfotriésteres), aquellos que contienen unidades estructurales sobran-
 25 tes, tales como, por ejemplo, proteínas (incluyendo enzimas (por ejemplo, nucleasas), toxinas, anticuerpos, péptidos de señal, poli-L-lisina, etc.), aquellos con intercaladores (por ejemplo, acridina, psoralen, etc.), aquellos que contienen quelatos (de, por ejemplo, metales, metales radioactivos, boro, metales oxidantes, etc.), aquellos que contienen alquiladores, aquellos con enlaces modificados (por ejemplo, ácidos nucleicos alfa anoméricos, etc.), así como formas no modificadas del polinucleótido u oligonucleótido.

30 "Secuenciación paralela masiva" significa técnicas para secuenciar millones de fragmentos de ácidos nucleicos, por ejemplo, utilizando el enlace de ADN genómico fragmentado aleatoriamente a una superficie plana transparente ópticamente y amplificación en fase sólida para crear una celda de flujo de secuenciación de alta densidad con millones de aglomeraciones, conteniendo cada una ~ 1,000 copias de plantilla por centímetro cuadrado. Estas plantillas son secuenciadas que utilizan la tecnología de secuenciación por síntesis de ADN de cuatro colores. Véanse los productos ofrecidos por Illumina, Inc., San Diego, California. La secuenciación usada actualmente se
 35 lleva a cabo preferiblemente sin una preamplificación o etapa de clonación, pero puede ser combinada con métodos basados en amplificación en un chip microfluído que tiene cámaras de reacción para secuenciación tanto por PCR como microscópica con base en plantillas. Solo aproximadamente 30 bp de información de secuencia aleatoria son necesarios para identificar una secuencia como perteneciente a un cromosoma humano específico. Secuencias más largas pueden identificar de manera única objetivos más particulares. En el caso presente, se obtuvo un gran
 40 número de lecturas de 35 bp. Se encontrará una descripción adicional de un método de secuenciación en paralelo masivo en Rogers y Ventner, Nature (2005) 437:326-327.

Se entiende que aspectos y realizaciones de la invención descrita aquí incluyen aspectos y realizaciones "consistentes en" y/o "consistentes esencialmente de".

45 Otros objetivos, ventajas y características de la presente invención serán evidentes a partir de la siguiente especificación tomada en conjunción con los dibujos acompañantes.

II. Establecimiento de una relación entre profundidad de cobertura y contenido de GC

50 Como se indicó anteriormente, el establecimiento de una relación entre la profundidad de cobertura y el contenido de GC de un cromosoma, comprenderá: obtener información de secuencia de múltiples fragmentos de polinucleótidos que cubren dicho cromosoma a partir de más de una muestra; asignar dichos fragmentos a cromosomas basados en dicha información de secuencia; determinar la profundidad de cobertura y el contenido de GC de dicho cromosoma con base en dicha información de secuencia para cada muestra; y determinar la relación entre la profundidad de cobertura y el contenido de GC de dicho cromosoma.

55 Para calcular la profundidad de cobertura y el contenido de GC de una localización cromosómica, la información de secuencia de los fragmentos de polinucleótidos se obtiene secuenciando el ADN de plantilla obtenido de una muestra de sangre periférica. En una realización, el ADN de plantilla contiene tanto el ADN materno como el ADN fetal. En otra realización, el ADN de plantilla se obtiene a partir de sangre de un sujeto femenino en embarazo. La sangre puede ser recolectada utilizando cualquier técnica estándar para extracción de sangre incluyendo pero no limitándose a la venipuntura. Por ejemplo, la sangre puede ser extraída de una vena del interior del codo o del dorso de la mano. Las muestras de sangre pueden ser recolectadas de un sujeto femenino en embarazo en cualquier

momento durante la gestación del feto. Por ejemplo, las muestras de sangre pueden ser recolectadas de sujetos femeninos humanos a 1-4, 4-8, 8-12, 12-16, 16-20, 20-24, 24-28, 28-32, 32-36, 36-40, o 40-44 semanas de gestación del feto, y preferiblemente entre 8-28 semanas de la gestación del feto.

- 5 Los fragmentos de polinucleótidos son asignados a una localización cromosómica basada en la información de secuencia. Se utiliza una secuencia genómica de referencia para obtener las lecturas únicas de referencia. Tal como se utiliza aquí, el término "lecturas únicas de referencia" se refiere a todos los fragmentos de polinucleótidos únicos que han sido asignados a una localización genómica específica con base en una secuencia genómica de referencia. En algunas realizaciones, las lecturas únicas de referencia tienen la misma longitud de, por ejemplo, aproximadamente 10, 12, 15, 20, 25, 30, 35, 40, 50, 100, 200, 300, 500, o 1000 bp. En algunas otras realizaciones, pueden utilizarse construcciones de genoma humano hg18 o hg 19 como secuencia genómica de referencia. Una localización cromosómica puede ser una ventana contigua sobre un cromosoma que tiene una longitud de aproximadamente 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, 10000 o más kb. Una localización cromosómica también puede ser un cromosoma individual.
- 10
- 15 Tal como se utiliza aquí, el término "profundidad de cobertura" se refiere a la relación entre el número de fragmentos que se asigna a una localización cromosómica y el número de lecturas únicas de referencia de la localización cromosómica utilizando la siguiente fórmula:

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

- 20 en donde $n_{i,j}$ es el número de lecturas de secuencia única mapeadas al cromosoma j en la muestra i; $C_{i,j}$ es la profundidad de cobertura en el cromosoma j en la muestra i; N_j es el número de Lecturas Únicas de Referencia en el cromosoma j.

- 25 Los fragmentos de polinucleótidos que no se asignan a una única localización cromosómica individual o se asignan a localizaciones cromosómicas múltiples son descartados. En algunas realizaciones, la profundidad de cobertura es normalizada, con base en la profundidad de cobertura de otra localización cromosómica, otro cromosoma, promedio de todos los otros autosomas, promedio de todos los otros cromosomas, o promedio de todos los cromosomas. En algunas realizaciones, la profundidad de cobertura promedio de 22 autosomas se utiliza como constante de normalización para tener en cuenta las diferencias en el número total de lecturas de secuencia obtenidas para diferentes muestras:

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

- 30 en donde $c_{i,j}$ representa la profundidad de cobertura relativa del cromosoma j en la muestra i. Desde este punto en adelante, "profundidad de cobertura relativa" para cada cromosoma se refiere al valor normalizado y se utiliza para comparar diferentes muestras y para análisis subsecuente.

- 35 El contenido de GC de una localización cromosómica puede calcularse por el porcentaje de GC promedio de una localización cromosómica con base en las lecturas de referencia únicas en la localización cromosómica, o en los fragmentos de polinucleótido secuenciados que se asignan a la localización del cromosoma. El contenido de GC de un cromosoma puede ser calculado usando la siguiente fórmula:

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

en donde i representa la muestra i, j representa el cromosoma j, $NGC_{i,j}$ representa el número de bases de ADN G y C y $BASE_{i,j}$ representa el número de bases de ADN en el cromosoma j en la muestra i.

- 40 Para establecer una relación entre la profundidad de cobertura del contenido de GC de una localización cromosómica, el cálculo puede estar basado en la información de la secuencia de los fragmentos de polinucleótido obtenida a al menos 1, 2, 5, 10, 20, 50, 100, 200, 500 o 1000 muestras.

En algunos casos, la relación entre la profundidad de cobertura y el contenido de GC es una relación lineal no fuerte. Pueden utilizarse el algoritmo de Loess, o una regresión polinómica pesada localmente, para establecer las

relaciones no lineales (correlaciones) entre pares de valores, tales como entre la profundidad de cobertura y el contenido de GC.

III. Determinación de una anomalía genética

5 Como se anotó previamente, un método para determinar una anomalía genética fetal de acuerdo con la invención, comprende: a) obtener información de secuencia de fragmentos múltiples de polinucleótidos a partir de una muestra de sangre periférica la cual es derivada de un sujeto femenino en embarazo y contiene ADN tanto materno como fetal; b) asignar dichos fragmentos a cromosomas con base en dicha información de secuencia; c) determinar la profundidad de cobertura y el contenido de GC de un cromosoma con base en la información de secuencia para aquellos fragmentos que se asignan únicamente a dicho cromosoma en la etapa (b); d) determinar la
10 profundidad de cobertura ajustada de dicho cromosoma utilizando dicho contenido de GC de dicho cromosoma y relaciones establecidas entre la profundidad de cobertura y el contenido de GC para dicho cromosoma en la ausencia de aneuploidía; y e) comparar dicha profundidad de cobertura ajustada con la profundidad de cobertura de dicho cromosoma determinada en la etapa (c), en donde una diferencia entre ellas indica una anomalía genética fetal la cual es una aneuploidía cromosómica.

15 Los métodos de la invención son especialmente útiles para la detección de aneuploidía, poliploidía, monosomía, trisomía, trisomía 21, trisomía 13, trisomía 14, trisomía 15, trisomía 16, trisomía 18, trisomía 22, triploidía, tetraploidía y anomalías cromosómicas sexuales incluyendo XO, XXY, XYY, y XXX. También se puede focalizar en ciertas regiones dentro del genoma humano de acuerdo con los presentes métodos con el fin de identificar monosomías parciales y trisomías parciales. Por ejemplo, los métodos pueden involucrar el análisis de datos de secuencias en una "ventana" cromosómica deslizante, tal como regiones contiguas, no superpuestas de 50 Kb esparcidas a través de un cromosoma. Se han reportado, entre otras, las trisomías parciales de 13q, 8p (8p23.1), 7q, distal 6p, 5p, 3q (3q25.1), 2q, 1q (1q42.1 y 1q21-qter), Xp parcial y monosomía 4q35.1. Por ejemplo, duplicaciones parciales del brazo largo del cromosoma 18 pueden dar como resultado el síndrome de Edwards en el caso de una duplicación de 18q21.1-qter (Mewar, et al., Am J Hum Genet. (1993) 53:1269-78).

25 En algunas realizaciones, la fracción fetal se estima con base en la información de secuencia obtenida para los fragmentos de polinucleótido de una muestra. La profundidad de cobertura, y el contenido de GC, de los cromosomas X y Y pueden utilizarse para la estimación de la fracción fetal. En algunas realizaciones, el género del feto se determina con base en la información de secuencia obtenida para los fragmentos de polinucleótido de una muestra. La profundidad de cobertura, y el contenido de GC, de los cromosomas X y Y pueden ser utilizados para
30 determinar el género del feto.

En algunas realizaciones, la comparación de dicha profundidad de cobertura ajustada a dicha profundidad de cobertura del cromosoma se lleva a cabo mediante una prueba de hipótesis estadística, en donde una hipótesis es que el feto es euploide (H0) y la otra hipótesis es que el feto es aneuploide (H1). En algunas realizaciones, la estadística t de Student se calcula tanto para la hipótesis t1 como para la t2, respectivamente. En algunas
35 realizaciones, se calcula la relación logarítmica probable de t1 y t2. En algunas realizaciones, una relación logarítmica probable de > 1 indica trisomía del feto.

IV. Medio legible por ordenador y sistema para diagnóstico de una anomalía genética fetal

En otro aspecto, se provee aquí un medio legible en ordenador que comprende una pluralidad de instrucciones adaptadas para ejecutar un método de la invención para diagnóstico prenatal de una anomalía genética fetal cuando dichas instrucciones son suplementadas con información de secuencia obtenida como en la etapa (a) del método.
40

En todavía otro aspecto, como se indicó previamente, también se provee aquí un sistema que comprende medios adaptados para ejecutar un método de la invención. Tal sistema comprende: a) medios para obtener información de secuencia a partir de fragmentos de polinucleótidos adecuados; y b) un medio legible en ordenador como se describió anteriormente.
45

Será evidente para los experimentados en la técnica que puede usarse un cierto número de métodos de secuenciación y variaciones diferentes. En una realización, la secuenciación se hace utilizando secuenciación en paralelo masiva. La secuenciación en paralelo masiva, tal como el factible sobre la plataforma 454 (Roche) (Margulies, et al., Nature (2005) 437:376-380), Illumina Genome Analyzer (o plataforma Solexa™) o SOLiD System (Applied Biosystems) o la tecnología de secuenciación Helicos True Single Molecule DNA (Harris, et al., Science (2008) 320:106-109), la tecnología de molécula individual, tiempo real (SMRT™) de Pacific Biosciences y la secuenciación en nanoporo (Soni and Meller, Clin Chem (2007) 53:1996-2001), permiten la secuenciación de muchas moléculas de ácido nucleico aisladas a partir de un espécimen en altos órdenes de multiplexión en una forma paralela (Dear, Brief Funct Genomic Proteomic (2003) 1:397-416). Cada una de estas plataformas secuencia moléculas individuales expandidas por clonación o incluso no amplificadas de fragmentos de ácido nucleico. Los equipos de secuenciación disponibles comercialmente pueden ser utilizados en la obtención de la información de secuencia de los fragmentos de polinucleótidos.
50
55

V. Ejemplos

Los siguientes ejemplos se ofrecen con el fin de ilustrar pero no de limitar la invención.

Ejemplo 1. Análisis de factores que afectan la sensibilidad de la detección: sesgo de GC y género

5 En la Figura 1 se ilustra un marco de procedimiento esquemático para calcular la profundidad de cobertura y el contenido de GC. Hemos utilizado un software para producir las lecturas únicas de referencia haciendo una incisión en las secuencias de referencia hg18 en 1-mer (1-mer aquí es una lectura que es descompuesta artificialmente a partir de la referencia de secuencia humana con la misma longitud "1" con las lecturas de secuenciación de muestras) y recolectamos aquellas 1-mer "únicas" como nuestras lecturas únicas de referencia. En segundo lugar, mapeamos nuestras lecturas de muestra secuenciada a las lecturas únicas de referencia para cada cromosoma. En 10 tercer lugar, eliminamos las que quedaban por fuera aplicando el método de corte quintil atípico para obtener un conjunto de datos claro. Finalmente, contamos la profundidad de cobertura de cada cromosoma para cada muestra y el contenido de GC de las lecturas únicas secuenciadas mapeado en cada cromosoma para cada muestra.

Con el fin de investigar cómo el contenido de GC afecta nuestros datos, escogimos 300 casos euploides con resultado de cariotipo y dispersamos su profundidad de cobertura y el contenido de GC relacionado de las lecturas 15 secuenciadas en una gráfica, la cual mostró una fuerte correlación entre ellos, y este fenómeno no fue reportado previamente (Figura 2). En la Figura 2, la profundidad de cobertura se correlacionaba fuertemente con el contenido de GC, y mostraba una tendencia obviamente hacia abajo en algunos cromosomas tales como 4, 13, etc., mientras que tendía hacia arriba en otros cromosomas tales como 19, 22, etc. Todos los cromosomas fueron dispuestos en orden ascendente en cuanto a su contenido de GC inherente y hay presente una tendencia hacia abajo en los 20 cromosomas del grupo con contenido inferior de GC mientras que hay una tendencia hacia arriba en cromosomas del grupo con contenido mayor de GC como se muestra en la Figura 3. Puede interpretarse que si los fragmentos de polinucleótidos que están siendo secuenciados para una muestra tienen un contenido de GC más alto que la otra muestra, la profundidad de cobertura que representa esta muestra caerá en comparación con la de otra muestra en los cromosomas del grupo con contenido más bajo de GC mientras se eleva en los cromosomas del grupo con contenido más alto de GC.

La posible explicación para tal tendencia cambiante diferente entre cromosomas con diferente contenido de GC es la diferencia en la composición del contenido de GC en diferentes cromosomas mostrada en la Figura 4 en combinación con el sesgo de GC introducido en el proceso de secuenciación. El contenido de GC de cada lectura 30 única de referencia de 35-mer para cada cromosoma se utiliza para clasificar el contenido de GC en 36 niveles. El porcentaje de cada nivel así como la composición en GC de cada cromosoma fueron calculados y luego usados para dibujar el mapa de calor con el software Heatmap2. Tomando el cromosoma 13 como ejemplo, su parte grande consiste de segmentos de secuencia con contenido inferior de GC pero una pequeña parte del mismo consiste de segmentos de secuencia con contenido más alto de GC. Si las condiciones durante el proceso de secuenciación o de PCR favorecen la secuenciación de aquellos segmentos con contenido de GC más alto, entonces una parte 35 relativamente grande del cromosoma 13 con contenido de GC bajo será difícil de ser secuenciada con el resultado de que la profundidad de cobertura en este cromosoma 13 de la muestra se hace más baja. En comparación, en un grupo con contenido más grande de GC tal como el cromosoma 19, la profundidad de cobertura en este cromosoma 19 de la muestra se hace más alto para que una parte grande del cromosoma 19 fuera de contenido de GC más alto la cual es preferida por el secuenciador. No importa en qué cromosoma, los segmentos pobres en GC y ricos en GC fueron difíciles de secuenciar pero la influencia introducida por el sesgo de GC fue diferente en diferentes cromosomas con diferente composición en contenido de GC. Cada cromosoma de referencia fue dividido en porciones basura de 1 KB, y se calculó el contenido de cada lectura de referencia única en la porción basura. Se calcula el contenido de GC de cada porción basura en el intervalo apropiado de [0.3, 0.6] dividido por el tamaño de la etapa de 0.001, y la cobertura relativa en cada intervalo. La Figura 5 muestra la representación gráfica de la 40 cobertura relativa y el contenido de GC para cada cromosoma.

La influencia del género fetal sobre los datos fue analizada utilizando una prueba t de dos muestras independientes. No se encontró una diferencia significativa entre autosomas excepto apenas para los cromosomas sexuales en el mismo contenido de GC, pero hay una diferencia obvia en el porcentaje de UR entre sujetos femeninos y masculinos (Chiu et al., (2008) Proc Natl Acad Sci USA 105:20458-20463), implicando que no hay necesidad de distinguir entre el género fetal cuando se detecta un autosoma aneuploide, pero es necesario distinguir el género fetal primeramente, cuando se detecta un cromosoma aneuploide tal como XO, XYY, etc. 50

Ejemplo 2. Modelo estadístico

Utilizando este fenómeno discutido anteriormente, tratamos de utilizar polinomios locales para ajustar la relación entre la profundidad de la cobertura y el correspondiente contenido de GC. La profundidad de cobertura consiste de una función de GC y un residual de la distribución normal como sigue: 55

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

en donde $f(GC_{i,j})$ representa la función de la relación entre la profundidad de cobertura y el correspondiente contenido de GC de la muestra i , cromosoma j , $\varepsilon_{i,j}$ representa el residual de la muestra i , cromosoma j .

5 Hay una relación lineal no fuerte entre la profundidad de cobertura y el correspondiente contenido de GC de tal manera que aplicamos el algoritmo de Loess para ajustar la profundidad de cobertura con el correspondiente contenido de GC, a partir del cual calculamos un valor importante para nuestro modelo, esto es, la profundidad de cobertura ajustada:

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

10 Con la profundidad de cobertura ajustada, se calcularon la varianza estándar y el t de Student de acuerdo con las siguientes Fórmula 6 y Fórmula 7:

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

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

Ejemplo 3. Estimación de la fracción fetal

15 Por la razón de que la fracción fetal es muy importante para nuestra detección estimamos así la fracción fetal antes del procedimiento de prueba. Como hemos mencionado antes, hemos secuenciado 19 adultos masculinos, cuando se comparó su profundidad de cobertura con la de los casos que portaban fetos femeninos, encontrando que la profundidad de cobertura masculina del cromosoma X es casi 1/2 veces la de los femeninos, y la profundidad de cobertura de los masculinos del cromosoma Y es casi 0.5 mayor que la de los femeninos. Así podemos estimar la fracción fetal dependiente de la profundidad de cobertura del cromosoma X y Y como la Fórmula 8, Fórmula 9 y Fórmula 10, considerando la correlación de GC también:

$$fy_i = (cr_{i,y} - \hat{cr}_{i,yf}) / (\hat{cr}_{i,ym} - \hat{cr}_{i,yf}) \quad (8)$$

$$fx_i = (cr_{i,x} - \hat{cr}_{i,xf}) / (\hat{cr}_{i,xm} - \hat{cr}_{i,xf}) \quad (9)$$

$$fxy_i = \arg \min_{\varepsilon \in (0,1)} \left(\frac{(\hat{cr}_{i,xf} \cdot (1-\varepsilon) + \hat{cr}_{i,ym} \cdot \varepsilon - cr_{i,x})^2}{(\hat{\sigma}_{x,f} \cdot (1-\varepsilon))^2 + (\hat{\sigma}_{x,m} \cdot \varepsilon)^2} + \frac{(\hat{cr}_{i,yf} \cdot (1-\varepsilon) + \hat{cr}_{i,ym} \cdot \varepsilon - cr_{i,y})^2}{(\hat{\sigma}_{y,f} \cdot (1-\varepsilon))^2 + (\hat{\sigma}_{y,m} \cdot \varepsilon)^2} \right) \quad (10)$$

20 en donde $\hat{cr}_{i,xf} = f(GC_{i,xf})$ es la profundidad de cobertura ajustada por la correlación de regresión de la profundidad de cobertura del cromosoma X y el contenido correspondiente de GC de los casos con feto femenino, $\hat{cr}_{i,yf} = f(GC_{i,yf})$ se refiere a la profundidad de cobertura ajustada por la correlación de regresión de la profundidad de cobertura del cromosoma Y y el contenido correspondiente de GC de los casos con feto femenino, $\hat{cr}_{i,ym} = f(GC_{i,ym})$ se refiere a la profundidad de cobertura ajustada por la correlación de regresión de la profundidad de cobertura del cromosoma X y el contenido correspondiente de GC de adultos masculinos, $\hat{cr}_{i,ym} = f(GC_{i,ym})$ se refiere a la profundidad de la cobertura ajustada por la correlación de regresión de la profundidad de cobertura del cromosoma Y y el contenido correspondiente de GC de adultos masculinos. Para la computación simplemente, se da que $\hat{\sigma}_{x,f}$ y $\hat{\sigma}_{x,m}$ son iguales y $\hat{\sigma}_{y,f}$ y $\hat{\sigma}_{y,m}$ son iguales.

Ejemplo 4. Cálculo del residual de cada cromosoma

La Figura 6 muestra que la variación estándar (véase Fórmula 3) para cada cromosoma bajo un cierto número total de lecturas únicas es influenciada por el número de casos participantes de la referencia. La variación estándar escasamente se incrementa cuando el número de casos seleccionados fue más de 150 bajo la condición de que 1.7 millones de número de lecturas únicas totales fueron secuenciadas para cada caso. Sin embargo, la variación estándar fue diferente para diferentes cromosomas. Después de considerar el sesgo de GC, nuestro método tiene una variación estándar moderada para el cromosoma 13 (0.0063), el cromosoma 18 (0.0066) y el cromosoma 21 (0.0072). La variación estándar del cromosoma X es mayor que en los cromosomas antes mencionados los cuales requerirían más estrategias para dar una detección anormal exacta.

La Figura 7 muestra la gráfica Q-Q, en donde el residual es compilado a la distribución normal lo cual implica que el cálculo t de Student es razonable.

Ejemplo 5. Distinción del género fetal

Para descubrir el trastorno del cromosoma sexual, es mejor distinguir el género fetal. Existieron dos picos obvios cuando investigamos la distribución de frecuencia de la profundidad de cobertura del cromosoma Y en nuestros 300 casos, lo cual nos dio una pista para distinguir el género por la profundidad de cobertura del cromosoma Y. Los casos con profundidad de cobertura menor de 0.04 pueden verse como portadores de fetos femeninos, mientras que más de 0.051 son vistos como portadores de fetos masculinos, entre 0.04 y 0.051 son vistos como de género incierto como la Figura 8. Para estos géneros dudosos y casos de aneuploidía, se utilizó la regresión logarítmica para predecir su género como la Fórmula 11 (Fan, et al., Proc Natl Acad Sci USA (2008) 42:16266-16271):

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

en donde $cr.a_{i,x}$ y $cr.a_{i,y}$ son una cobertura relativa normalizada de X y Y, respectivamente.

En comparación con el resultado del cariotipo, nuestro método para la distinción del género fetal funcionó bastante bien en nuestros 300 casos de referencia con 100% de exactitud, mientras que hubo una equivocación en un caso cuando se ejecutó en nuestro conjunto de 901 casos y la profundidad de cobertura del cromosoma Y de este caso erróneo estaba entre 0.04 y 0.051.

Ejemplo 6. Rendimiento diagnóstico de la metodología de prueba t en correlación con GC

Reclutamiento de muestra

Se reclutaron 903 participantes de manera prospectiva del Hospital del Pueblo de Shenzhen y del Materno de Shenzhen y del centro de servicio de cuidado infantil con sus resultados de cariotipos. Las aprobaciones fueron obtenidas de los comités de revisión institucionales de cada sitio de reclutamiento y todos los participantes dieron consentimiento escrito informado. Se registraron las edades de las madres y las semanas de gestación en el momento del muestreo de sangre. Los 903 casos incluyeron 2 casos de trisomía 13, 15 casos de trisomía 18, 16 casos de trisomía 21, 3 casos XO, 2 casos XXY y 1 caso XYY. Su distribución de resultados de cariotipo se muestra en la Figura 9.

Secuenciación de ADN de plasma materno

Se recolectó sangre venosa periférica (5 mililitros) de cada mujer embarazada participante en tubos de EDTA y se centrifugó a 1,600 g durante 10 minutos en 4 horas. El plasma fue transferido a tubos de microcentrífuga y se recentrifugó a 16,000 g durante 10 minutos para retirar células residuales. El plasma libre de células fue almacenado a 80°C hasta la extracción del ADN. Cada muestra de plasma fue congelada y descongelada sólo una vez.

Para una secuenciación genómica paralela masiva, todo el ADN extraído de 600 µl de plasma materno fue utilizado para la construcción de una biblioteca de ADN de acuerdo con un protocolo modificado de Illumina. En resumen, el reapareamiento en extremos de los fragmentos de ADN de plasma materno fue llevado a cabo utilizando T4 ADN polimerasa, polimerasa Klenow™, y T4 polinucleótido quinasa. Se ligaron adaptadores comercialmente disponibles (Illumina) a los fragmentos de ADN después de la adición de residuos A terminales. El ADN ligado al adaptador fue amplificado adicionalmente utilizando un PCR de 17 ciclos con cebadores multiplex estándar. Se utilizó un kit de 60 ml de Agencourt AMPure™ (Beckman) para la purificación de los productos de PCR. La distribución de tamaño de las bibliotecas de secuenciación fue analizada con un kit de DNA 1000 en el 2100 Bioanalyzer™ (Agilent) y se cuantificó con PCR en tiempo real. Las bibliotecas de secuenciación con diferentes índices fueron reunidas en una

sola en cantidades iguales antes de la estación de aglomeración en el Illumina GA II™ (secuenciación de extremo individual)

Se secuenciaron muestras de 19 sujetos masculinos euploides para el análisis subsecuente para la estimación de la fracción de ADN fetal. Una nueva metodología de prueba t de correlación de GC ha sido desarrollada por nosotros para el diagnóstico de trisomía 13, trisomía 18, trisomía 21 y anomalías del cromosoma sexual. Comparamos este nuevo método con otros dos métodos mencionados más abajo en términos de rendimiento de diagnóstico.

Ejemplo 7. Detección de aneuploidía fetal, tal como trisomía 13, 18 y 21

Para determinar si el número de copias de un cromosoma dentro de un caso de un paciente desviado del normal, la profundidad de cobertura de un cromosoma fue comparada con la de todos los otros casos de referencia. Todo el estudio previo tenía sólo una hipótesis nula. Introdujimos hipótesis binarias por primera vez utilizando una doble hipótesis nula. Una hipótesis nula (H0: el feto es euploide) fue la suposición de que la profundidad de cobertura media de la distribución de caso del paciente y la profundidad de cobertura media de toda la distribución de referencia normal eran iguales, lo que significa que el caso del paciente era euploide si se aceptaba esta hipótesis nula. Utilizando la prueba t de Student, puede calcularse t1 como la Fórmula 12:

$$t1_{i,j} = (cr_{i,j} - \hat{cr}_{i,j}) / std_j \quad (12)$$

La otra hipótesis nula (H1: el feto es aneuploide) fue que la profundidad de cobertura media de la distribución del caso del paciente con apenas una fracción fetal era igual a la profundidad de cobertura media de la distribución de casos aneuploides con la misma fracción fetal, lo que significa que este caso del paciente es aneuploide si se aceptara esta hipótesis nula. La estadística t de Student, t2 fue calculada como la Fórmula 13:

$$t2_{i,j} = (cr_{i,j} - \hat{cr}_{i,j}(1 + fxy_{i,j} / 2)) / std_j \quad (13)$$

|t1|>3 y |t2|<3 indicarían un caso aneuploide en la mayoría de las instancias especialmente cuando las distribuciones entre los casos euploides y los casos aneuploides fueron completamente discriminadas, mientras que en otra condición tal como la precisión insuficiente o fracción fetal insuficiente y así sucesivamente, |t1| puede ser menor que 3 pero el feto era anormal. t1 y t2 combinadas pueden ayudarnos a tomar una decisión más exacta, y entonces empleamos la relación de probabilidad Log de t1 y t2 como en la Fórmula 14:

$$L_{i,j} = \log(p(t1_{i,j}, \text{grado} | D)) / \log(p(t2_{i,j}, \text{grado} | T)) \quad (14)$$

en donde L_{i,j} es la relación de probabilidad Log. Si la relación era superior a 1, podríamos inferir que el feto puede tener trisomía.

Pero para casos con fetos femeninos, es difícil para nosotros estimar su fracción fetal de tal forma que es imposible computarla. Sin embargo, podemos dar un valor de referencia (RV) de una fracción del 7% de acuerdo con la distribución empírica de la fracción fetal.

Se investigaron 903 casos, 866 de los cuales portaban fetos euploides dentro de los cuales 300 casos habían sido seleccionados aleatoriamente para desarrollar la metodología t de Student de correlación de GC. Además, participaron de nuestro estudio 2 casos de trisomía 13, 12 de trisomía 18, 16 de trisomía 21, 4 de XO (consistentes de 3 casos XO y 1 caso de quimera 45, xo/46, xx (27:23)), 2 XXY y 1 XYY. Después del alineamiento, obtuvimos una media de 1.7 millones de lecturas alineadas únicas de datos (SD = 306185) por caso sin equivocaciones. Utilizando nuestra recién desarrollada prueba t de Student de correlación de GC, todos los casos T13 (2 de 2) fueron identificados exitosamente, mientras que 901 de los 901 casos sin trisomía 13 fueron clasificados correctamente (Figura 10A). La sensibilidad y especificidad de esta metodología fueron 100% y 100% (Tabla 1).

Para la trisomía 18, 12 de los 12 casos de trisomía 18 y 888 de 891 sin trisomía 18 pudieron ser identificados correctamente (Figura 10A). La sensibilidad y especificidad de esta metodología fue de 100% y 99.66% respectivamente. Para la trisomía 21, 16 de 16 casos de trisomía 21 y 16 de 16 casos sin trisomía 21 también pudieron ser detectados correctamente (Figura 10A). La sensibilidad y especificidad de esta metodología fueron 100% y 100%, respectivamente.

Ejemplo 8. Detección de XO, XXX, XXY, XYY

Más arriba hemos considerado la detección de trisomía para autosomas, trastornos para el cromosoma sexual tales como XO, XXX, XXY y XYY que pueden ser detectados también por nuestro método.

Primeramente, el género fue confirmado por distinción de género. Si se confirmaba que un caso de prueba portaba feto femenino, fue necesario calcular el valor t de Student $t_{1,X} = (cr_{i,X} - \hat{cr}_{i,Xf}) / std_{Xf}$ para detección de XXX o XO, en donde $\hat{cr}_{i,Xf}$ y std_{Xf} son las mismas que en la Fórmula 10; si t_1 es mayor que 3.13 o menor que -3.13, este caso puede ser XXX o XO. Pero considerando que la precisión estaba limitada por la gran desviación de profundidad de cobertura para el cromosoma X, muestreamos el plasma de nuevo y se repitió el experimento para hacer una decisión más creíble cuando $|t_1| < 5$ incluso aunque $|t_1| > 3.13$. $|t_1| > 5$ fue confirmado como aneuploide en este caso. Todos nuestros procesos de detección se basaron en la premisa de que los datos satisfacían el control de calidad estándar.

Si se confirmó que la muestra de prueba portaba feto masculino, la fracción de ADN fetal fue estimada primero por Y y X. Entre tanto, pudimos extrapolar la profundidad de cobertura ajustada para el cromosoma X con la fracción de ADN fetal estimada solamente por la profundidad de la cobertura del cromosoma Y y puede calcularse $t_2 = (cr_{i,X} - (1 - f_{Yi}/2) \cdot \hat{cr}_{i,Xf}) / std_{Xf}$. Si t_2 es demasiado grande (mayor de 5) o demasiado pequeño (menor de -5) el feto puede ser XXY o XYY. Además, la brecha entre las fracciones fetales estimadas por X y Y independientemente proveerá información para detectar trastornos acerca de los cromosomas sexuales.

En la detección de XO, se detectaron 3 de cada 4 casos de XO, y el caso que no pudo ser identificado fue un caso de quimera (Figura 10B). La sensibilidad y especificidad para esta metodología fueron de 75% (100% si se descarta el caso de la quimera) y 99.55% respectivamente. Para casos XXY, los dos 2 casos fueron identificados exitosamente, mientras que 901 de 901 casos que no eran XXY fueron clasificados correctamente (Figura 10B) con la sensibilidad de 100% y especificidad de 100%. Para el caso XYY, lo identificamos correctamente (Figura 10B) y la sensibilidad y especificidad fueron de 100% y 100% respectivamente.

Para evaluar si nuestra nueva metodología tenía ventajas en comparación con otras dos metodologías reportadas, marcador z y marcador z con corrección de GC, implementamos todas estas tres metodologías para analizar nuestros 900 casos y los mismos 300 casos como referencia fijados para todas estas metodologías. La precisión de la medición siempre estuvo incorporada en el valor de confianza (CV). En nuestra investigación, el CV de la metodología estándar de marcador z es mayor que las otras metodologías en cromosomas 18 y 21 interesantes clínicamente (Figura 11) llevando a una rata de sensibilidad menor para trisomía 18 y 21 (Tabla 1).

Tabla 1 Comparación de sensibilidad y especificidad de diferentes métodos

	Diagnóstico (# de casos)	Metodología de marcador z estándar		Metodología de marcador z con corrección de GC		Nuestra metodología con prueba t de correlación de GC	
		Sensibilidad (%)	Especificidad (%)	Sensibilidad (%)	Especificidad (%)	Sensibilidad (%)	Especificidad (%)
Autosoma	Trisomía 13 (2)	50%	99.89	100%	100%	100%	100%
	Trisomía 18 (12)	91.67%	100%	100%	99.89%	100%	99.96%
	Trisomía 21 (16)	93.75%	100%	100%	100%	100%	100%
Heterosoma	XO (3 XO, 1 XO/XX quimera)	no disponible	no disponible	no disponible	no disponible	75%	100%
	XXY (1)	no disponible	no disponible	no disponible	no disponible	100%	100%
	XXY (2)	no disponible	no disponible	no disponible	no disponible	100%	100%

Hipótesis binaria	no disponible	no disponible	disponible
-------------------	---------------	---------------	------------

5 Para la metodología de marcador z con corrección de GC, el valor CV de cromosoma 13 es 0.0066 con una tasa de sensibilidad de 100% y una tasa de especificidad de 100%. Para la novedosa metodología t de Student de correlación de GC discutida aquí, el valor de CV del cromosoma 13 es 0.0063 y con 100% de tasa de sensibilidad y 100% de tasa de especificidad. En el cromosoma 18, los CV de estas dos metodologías fueron 0.0062 y 0.0066, respectivamente, ambos con tasas de sensibilidad y especificidad de 100% que para ellos fueron 99.89% y 99.96%, respectivamente. El comportamiento fue similar al comparar los CV de estas dos metodologías para el cromosoma 21: 0.0088 y 0.0072, respectivamente. Ambos dieron como resultado la misma tasa de sensibilidad de 100% en nuestro estudio de conjunto de casos pequeños y alcanzó la misma tasa de especificidad de 100%. Y estos dos métodos se comportaron mejor que la metodología estándar de marcador z. No solamente nuestra nueva metodología fue desarrollada con correlación de GC comparable con la metodología de corrección de GC con buen rendimiento, sino que también tenía otra ventaja en la detección de anomalías del cromosoma sexual tales como XO, XXY y XYY. Nuestros datos muestran que cuando se emplea una metodología de corrección de GC habría dificultad en distinguir el género de los fetos por desviación de los datos que representan los cromosomas sexuales introducidos cuando se enmienda el número de etiquetas de secuencia multiplicándolas por el factor de peso de tal forma que la detección del trastorno de los cromosomas sexuales parecería difícil.

Ejemplo 9. Comportamiento teórico de la metodología de prueba t con correlación de GC en consideración del tamaño de datos, semanas de gestación y fracción de ADN fetal

20 Medir la aneuploidía sigue siendo un reto por la alta señal de fondo del ADN materno (Fan, et al., Proc Natl Acad Sci USA (2008) 42:16266-16271) y la fracción de ADN fetal pequeña arbitraria fue el factor restrictivo más significativo para la detección aneuploide por la metodología de secuenciación de genoma paralelo masiva (MPGS) hasta hoy en día. Sin embargo, no hubo una gran complicación en la determinación de la fracción de ADN fetal mínima clínicamente antes de la detección de MPGS especialmente para fetos femeninos mientras que la única pista clínica relacionada con la implicación de la fracción del ADN fetal era las semanas de gestación. Se reporta que hay una correlación estadísticamente significativa entre la fracción de ADN fetal y la edad gestacional previamente (Lo, et al., Am. J. Human Genet. (1998) 62:768-775). En nuestro estudio, para investigar la relación entre la fracción de ADN fetal estimada y la edad gestacional, representamos gráficamente en la Figura 12 la fracción de ADN fetal de todos aquellos casos participantes con fetos masculinos (en total 427 casos) con referencia a la Fórmula de estimación 10. La fracción de ADN fetal estimada para cada muestra se correlaciona con la edad gestacional (P menor de 0.0001). También mostró que aunque en la edad gestacional 20, hubo 4 de 65 casos con fracción de ADN fetal menor de 5%, lo cual afectaría adversamente la exactitud de la detección. Para evaluar el método de estimación de la fracción fetal seleccionamos más casos distribuidos jerárquicamente en fracción fetal estimada, y luego la Q-PCR ayudó a calcular otra fracción fetal relativa. Obtuvimos entonces una curva estándar de correlación que mostraba una fuerte correlación entre ellas lo cual demostró que la estimación de la fracción fetal por nuestro método es creíble.

35 Entre tanto, la profundidad de secuenciación (el número total de lecturas únicas) fue otro factor significativo que afectó la precisión de la detección aneuploide incorporándolos en el valor de la variación estándar. La variación estándar para cada cromosoma empleado en nuestra metodología correlacionada con GC pudo ser fijado bajo un cierto nivel de profundidad de secuenciación cuando el número de caso de referencia alcanzó 150 (Figura 13). Para investigar cómo la profundidad de secuenciación influye en la variación estándar para cada cromosoma, secuenciamos 150 casos no solamente en nuestro nivel de 1.7 millones sino en otro nivel de profundidad de secuenciación alcanzando el número de lecturas únicas totales 5 millones (SD = 1.7 millones). Dependiendo de estos dos conjuntos, encontramos que la varianza estándar es lineal con el recíproco de la raíz cuadrada del número de lecturas únicas totales (véase descripción de la Figura 6).

45 Para una fracción de ADN fetal dada, pudimos estimar el número de lecturas únicas totales requeridas en nuestro método para detectar la variación del número de copias de cromosoma a partir de normal a t1 igual 3 (Figura 14). Mostró que cuanto menor era la fracción de ADN fetal, mayor era la profundidad de secuenciación requerida. En nuestro conjunto de 1.7 millones de lecturas únicas, nuestra metodología es capaz de detectar fetos aneuploides para el cromosoma 13 y X con una fracción de ADN fetal de más de 4.5% y fetos aneuploides para el cromosoma 21 y 18 de más de 4%; mientras que en nuestro conjunto de referencia de 5 millones, nuestra metodología fue capaz de detectar trisomía 18 y trisomía 21 incluso con una fracción de ADN fetal de aproximadamente 3%. Si queremos identificar fetos anormales en cromosoma X tales como XXX o XO con una fracción fetal de aproximadamente 4%, el número único total requerido en esos casos y en casos de referencia correspondientes debería alcanzar 5 millones. Si el ADN fetal es menor de 3.5%, el requerimiento de profundidad de secuenciación estaría más allá de 20M. Y si la fracción fetal de ADN fuera inferior la detección sería no creíble y difícil, de manera que propusimos otra estrategia, esto es, deberíamos remuestrear el plasma de la mujer embarazada, rehacer nuestro experimento y reanalizar los datos cuando la edad gestacional fuese mayor sobre la gran probabilidad de que la fracción de ADN fetal se incrementaría con el incremento de la edad gestacional. Y esta estrategia también podría ser aplicada a muestras de las que se sospecha que tienen una fracción de ADN fetal pequeña.

Si bien nuestra metodología se comporta bien no es persuasiva sin un gran conjunto de casos anormales. Para estimar la sensibilidad de esta metodología de t Student con correlación de GC aplicada por nosotros, publicamos la sensibilidad teórica considerando diferentes edades gestacionales y diferentes profundidades de secuenciación.

Calculamos la sensibilidad teórica de la aneuploidía con las siguientes etapas. Primeramente, aplicamos un análisis

5 de regresión para ajustar la fracción de ADN fetal con la edad gestacional $\hat{fr}_j = f(gsa_i)$, donde \hat{fr}_j es la media de ajuste de la fracción de ADN fetal en su edad gestacional gsa_i , y estimamos la distribución de fracción de ADN fetal empleando una estimación de densidad de núcleo Gaussiano (Birke, (2008) Journal of Statistical Planning and Inference 139:2851-2862) con referencia principalmente a la fracción de ADN fetal estimada distribuida en las

10 semanas gestacionales 19 y 20 antes de extrapolar la distribución de fracción de ADN fetal en las otras semanas de acuerdo con la relación entre la fracción de ADN fetal y la edad gestacional $\hat{pd}_j(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - X_i}{h}\right) + (fr_j - \frac{fr_{19} + fr_{20}}{2})$, donde \hat{pd}_j es la densidad de probabilidad de ajuste de la fracción de ADN fetal en la edad gestacional, donde X es los datos de las semanas gestacionales 19 y 20 (figura 12). En

15 segundo lugar, estimamos la varianza estándar de acuerdo con el número de lecturas únicas totales como lo mencionamos antes $\hat{\sigma} = f(tuqn)$ donde tuqn es el número de lecturas únicas totales. Finalmente, para calcular la sensibilidad en cada edad gestacional a un cierto nivel de profundidad de secuenciación de acuerdo con la distribución de la fracción de ADN fetal y la varianza estándar estimada en cada profundidad de secuenciación, computamos la densidad de probabilidad de negativos falsos en cada fracción de ADN fetal (aquí, supusimos que la fluctuación de la fracción de ADN fetal se distribuía normalmente) y luego las integramos para obtener una

20 $FNR(fr, semana) = \int_0^1 pd(fr) \int_0^{\alpha/\sigma_j=3} \exp(x - fr)^2 / 2 \cdot \sigma_j^2 dx dfr$ donde j es el cromosoma j. Con facilidad, la sensibilidad teórica en una cierta profundidad de secuenciación en esta edad gestacional se calcula como 1-FNR. Las figuras 15-21 muestran las gráficas resultantes de nuestros cálculos. La t de Student mayor de 3 se fijó para

25 identificar fetos femeninos aneuploides mientras que para fetos masculinos, cuando se computa la densidad de probabilidad de negativos falsos en cada fracción, se empleó una probabilidad logarítmica mayor que 1 como el valor crítico que mencionamos en la hipótesis binaria que ayudaba a alcanzar una sensibilidad más alta en comparación con los femeninos.

Sin embargo, nuestra inferencia es relativamente conservadora por la razón de que es difícil obtener una distribución infinitamente aproximada a la distribución real de la fracción de ADN fetal junto con la edad gestacional especialmente en una edad gestacional corta en un muestreo a escala pequeña.

30 Referencias

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*, London, UK: Chapman & Hall/CRC.
- 5 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* 42:16266-16271.
12. Melanie Birke. (2008) Shape constrained kernel density estimation. *Journal of Statistical Planning and Inference* Volume 139, Issue 8, 1 August 2009, Pages 2851-2862.
13. Lo et al., *Lancet* 350:485-487 (1997).
- 10 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 Sep. 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 December; 53(6):1269-78.
- 15 18. Margulies et al., (2005) *Nature* 437:376-380.
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.

REIVINDICACIONES

1. Un método implementado por ordenador para determinar una anomalía genética fetal la cual es una aneuploidía cromosómica, método que comprende:
- 5 (a) obtener la información de secuencia de múltiples fragmentos de polinucleótidos a partir de una muestra, siendo dicha muestra una muestra de sangre periférica derivada de un sujeto femenino en embarazo y que contiene ADN tanto materno como fetal;
- (b) asignar dichos fragmentos a cromosomas con base en dicha información de secuencia comparando dichos fragmentos con las lecturas únicas de referencia del mismo tamaño para cada uno de dichos cromosomas, en donde las lecturas únicas de referencia son fragmentos de un cromosoma que tiene una secuencia única la cual puede ser asignada de manera no ambigua a una única localización cromosómica individual con base en una secuencia genómica de referencia;
- 10 (c) determinar la profundidad de cobertura y el contenido de GC de un cromosoma con base en la información de secuencia para aquellos fragmentos que han sido asignados a las lecturas únicas de referencia de dicho cromosoma en la etapa (b), en donde la profundidad de cobertura es la relación entre el número de fragmentos asignados de manera única a dicho cromosoma y el número de lecturas únicas de referencia para dicho cromosoma del mismo tamaño del fragmento con base en dicha secuencia genómica humana de referencia;
- 15 (d) determinar la profundidad de cobertura ajustada de dicho cromosoma utilizando dicho contenido de GC de dicho cromosoma y la relación establecida entre la profundidad de cobertura y el contenido de GC para dicho cromosoma en la ausencia de aneuploidía, en donde dicha relación establecida ha sido determinada por un método que comprende las etapas de:
- 20 (i) obtener información de secuencia de múltiples fragmentos de polinucleótidos que cubren dicho cromosoma a partir de una pluralidad de muestras de sangre periférica euploide que contienen ADN genómico, en donde el tamaño del fragmento es el mismo que el tamaño del fragmento de los fragmentos de polinucleótido múltiples de la etapa (a) anterior;
- 25 (ii) asignar dichos fragmentos a cromosomas con base en dicha información de secuencia como en la etapa (b) más arriba;
- (iii) determinar la profundidad de cobertura y el contenido de GC de dicho cromosoma con base en dicha información de secuencia para cada muestra euploide como en la etapa (c) más arriba; y
- 30 (iv) utilizar la profundidad de cobertura y el contenido de GC determinados para cada muestra en la etapa (iii) para determinar la relación entre la profundidad de cobertura y el contenido de GC de dicho cromosoma en la ausencia de aneuploidía; y
- (e) comparar dicha profundidad de cobertura ajustada con la profundidad de cobertura de dicho cromosoma determinada en la etapa (c), en donde una diferencia entre ellos indica aneuploidía cromosómica fetal.
- 35 2. El método de la reivindicación 1, en donde la etapa (a) comprende adicionalmente obtener información de secuencia de múltiples fragmentos de polinucleótidos a partir de muestras diferentes y la profundidad de cobertura es normalizada para tener en cuenta las diferencias en el número total de lecturas de secuencia obtenidas para diferentes muestras, por ejemplo contra la profundidad de cobertura promedio de otro cromosoma, preferiblemente contra la profundidad de cobertura promedio de todos los otros autosomas o contra la profundidad de cobertura promedio de todos los otros cromosomas.
- 40 3. El método de acuerdo con la reivindicación 1 o reivindicación 2, en donde el contenido de GC del cromosoma se determina como el contenido de GC promedio de todos los fragmentos que se asignan a dicho cromosoma para el propósito de la etapa (c).
- 45 4. El método de acuerdo con la reivindicación 2 que comprende adicionalmente determinar el género fetal, por ejemplo de acuerdo con la fórmula

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

donde $cr_{i,x}$ y $cr_{i,y}$ son la profundidad de cobertura relativa de los cromosomas X y Y de la muestra i , respectivamente.

5. El método de acuerdo con la reivindicación 2 que comprende adicionalmente estimar la fracción fetal, en donde la fracción fetal es calculada con base en la profundidad de la cobertura del cromosoma X y/o Y determinada como en la etapa (c) de la reivindicación 1 de acuerdo con una fórmula seleccionada de:

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

10 donde $cr_{i,y}$ es la profundidad de cobertura del cromosoma Y de la muestra i , $\hat{cr}_{i,yf} = f(GC_{i,yf})$ es la profundidad de cobertura ajustada calculada a partir de la relación de la profundidad de la cobertura del cromosoma Y y el contenido correspondiente de GC de las muestras de mujeres en embarazo con un feto femenino, y $\hat{cr}_{i,ym} = f(GC_{i,ym})$ se refiere a la profundidad de la cobertura ajustada calculada a partir de la relación de la profundidad de la cobertura del cromosoma Y y el contenido correspondiente de GC de sujetos masculinos; o

$$(ii) \quad fx_i = (cr_{i,x} - \hat{cr}_{i,xf}) / (\hat{cr}_{i,xm} - \hat{cr}_{i,xf}),$$

15 en donde $cr_{i,x}$ es la profundidad de cobertura del cromosoma X de la muestra i , $\hat{cr}_{i,xf} = f(GC_{i,xf})$ es la profundidad de la cobertura ajustada calculada a partir de la relación de la profundidad de cobertura del cromosoma X y el contenido correspondiente de GC de las muestras de mujeres embarazadas con un feto femenino, $\hat{cr}_{i,xm} = f(GC_{i,xm})$ se refiere a la profundidad de la cobertura ajustada calculada a partir de la relación de la profundidad de cobertura del cromosoma X y el contenido correspondiente de GC de muestras de sujetos masculinos; o

$$(iii) \quad fxy_i = \arg \min_{\varepsilon \in (0,1)} \left(\frac{(\hat{cr}_{i,yf} \cdot (1-\varepsilon) + \hat{cr}_{i,xm} \cdot \varepsilon - cr_{i,x})^2}{(\hat{\sigma}_{x,f} \cdot (1-\varepsilon))^2 + (\hat{\sigma}_{x,m} \cdot \varepsilon)^2} + \frac{(\hat{cr}_{i,yf} \cdot (1-\varepsilon) + \hat{cr}_{i,ym} \cdot \varepsilon - cr_{i,x})^2}{(\hat{\sigma}_{y,f} \cdot (1-\varepsilon))^2 + (\hat{\sigma}_{y,m} \cdot \varepsilon)^2} \right),$$

20 en donde $\hat{cr}_{i,xf} = f(GC_{i,xf})$ es la profundidad de la cobertura ajustada calculada a partir de la relación de la profundidad de cobertura del cromosoma X y el contenido correspondiente de GC de muestras de mujeres embarazadas con un feto femenino, $\hat{cr}_{i,yf} = f(GC_{i,yf})$ se refiere a la profundidad de la cobertura ajustada calculada a partir de la relación de la profundidad de la cobertura del cromosoma Y y el contenido correspondiente de GC de muestras de mujeres embarazadas con un feto femenino, $\hat{cr}_{i,xm} = f(GC_{i,xm})$ se refiere a la profundidad de la cobertura ajustada calculada a partir de la relación de la profundidad de cobertura del cromosoma X y el contenido correspondiente de GC de

25 muestras de sujetos masculinos, y $\hat{cr}_{i,ym} = f(GC_{i,ym})$ se refiere a la profundidad de cobertura ajustada calculada a partir de la relación de la profundidad de cobertura del cromosoma Y y el contenido correspondiente de GC de sujetos masculinos.

6. El método de acuerdo con la reivindicación 2, en donde la comparación de dicha profundidad de cobertura ajustada con la profundidad de cobertura del cromosoma determinada en la etapa (c) se lleva a cabo mediante una prueba de hipótesis estadística, en donde una hipótesis es que el feto es euploide (H0) y la otra hipótesis es que el feto exhibe aneuploidía para dicho cromosoma (H1).

7. El método de la reivindicación 6, en donde la estadística t de Student se calcula para ambas hipótesis.

8. El método de la reivindicación 7, en donde la estadística t de Student se calcula para H0 y H1 de acuerdo con las fórmulas

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

35 y

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

respectivamente, en donde f_{xy} es la fracción fetal,

i es el índice de muestra,

j es el número de cromosomas,

5
$$cr_{i,j} = f(GC_{i,j}) + \varepsilon_{i,j}, j = 1, 2, \dots, 22, X, Y$$

en donde $f(GC_{i,j})$ representa la función para la relación entre la profundidad de cobertura normalizada y el contenido correspondiente de GC de la muestra i , cromosoma j , $\varepsilon_{i,j}$ representa el residual de la muestra i , cromosoma j ,

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

y representa la profundidad de cobertura ajustada, y

10 std_j es la variación estándar de acuerdo con la fórmula

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

en donde ns representa el número de muestras de referencia.

15 9. El método de la reivindicación 8, en donde la relación de probabilidad logarítmica de t_1 y t_2 se calcula de acuerdo con la fórmula: $L_{i,j} = \log(p(t_{1_{i,j}}, \text{grado}|D)) / \log(p(t_{2_{i,j}}, \text{grado}|T))$, en donde $L_{i,j}$ es la relación de probabilidad logarítmica, en donde grado se refiere a un grado de distribución t , D se refiere a Diploidía, T se refiere a Trisomía, y $p(t_{i,j}, \text{grado}|*)$, $*=D, T$ representa la densidad de probabilidad condicional dada a un grado de distribución t , si la relación es más grande que 1 infiriéndose que el feto exhibe trisomía para dicho cromosoma.

10. El método de acuerdo con una cualquiera de las reivindicaciones 1 a 9 para uso en la determinación de aneuploidía en autosoma fetal.

20 11. El método de la reivindicación 10, en donde la aneuploidía fetal se selecciona del grupo consistente de trisomía 13, 18 y 21.

12. El método de acuerdo con la reivindicación 4 para uso en la determinación de aneuploidía de cromosoma sexual, tal como aneuploidía de cromosoma sexual seleccionada del grupo consistente de XO, XXX, XXY y XYY.

25 13. Un medio legible por ordenador que comprende una pluralidad de instrucciones adaptadas para ejecutar el método de una cualquiera de las reivindicaciones 1 a 12 cuando dichas instrucciones son suplementadas con información de secuencia obtenida como en la etapa (a) de la reivindicación 1.

14. Un sistema que comprende medios adaptados para ejecutar el método de una cualquiera de las reivindicaciones 1 a 12.

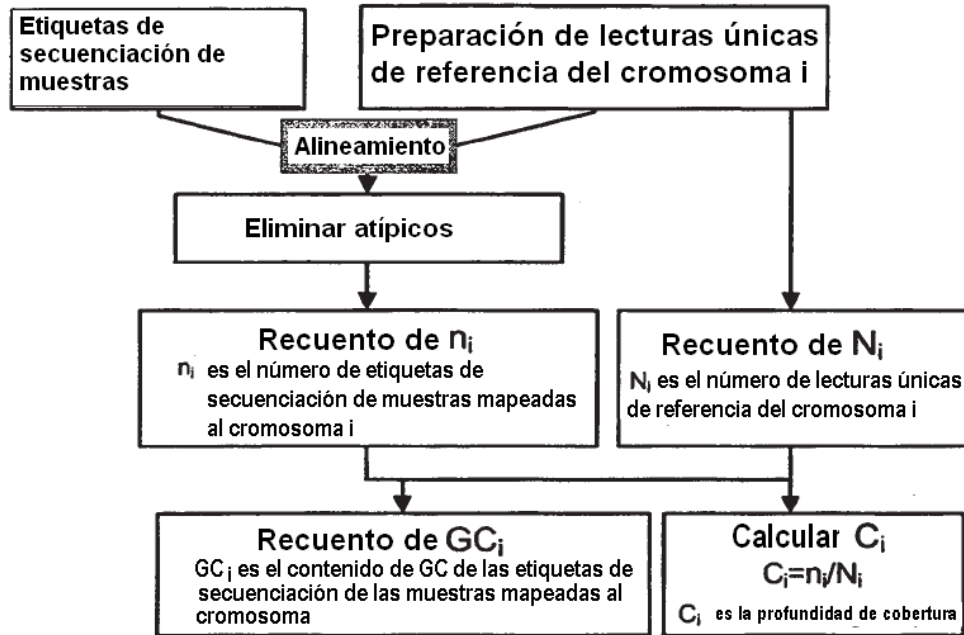


Figura 1

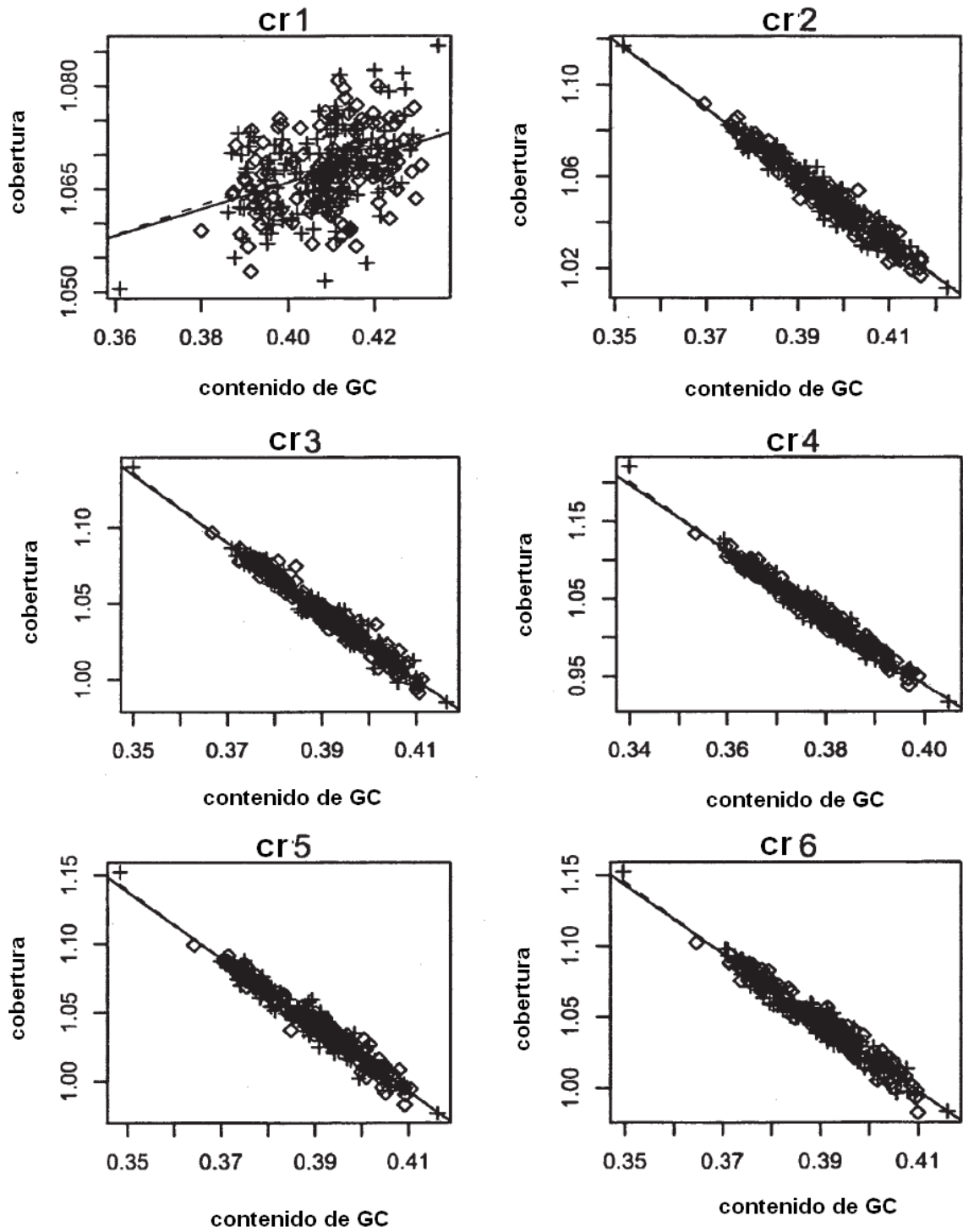


Figura 2
(página 1 de 4)

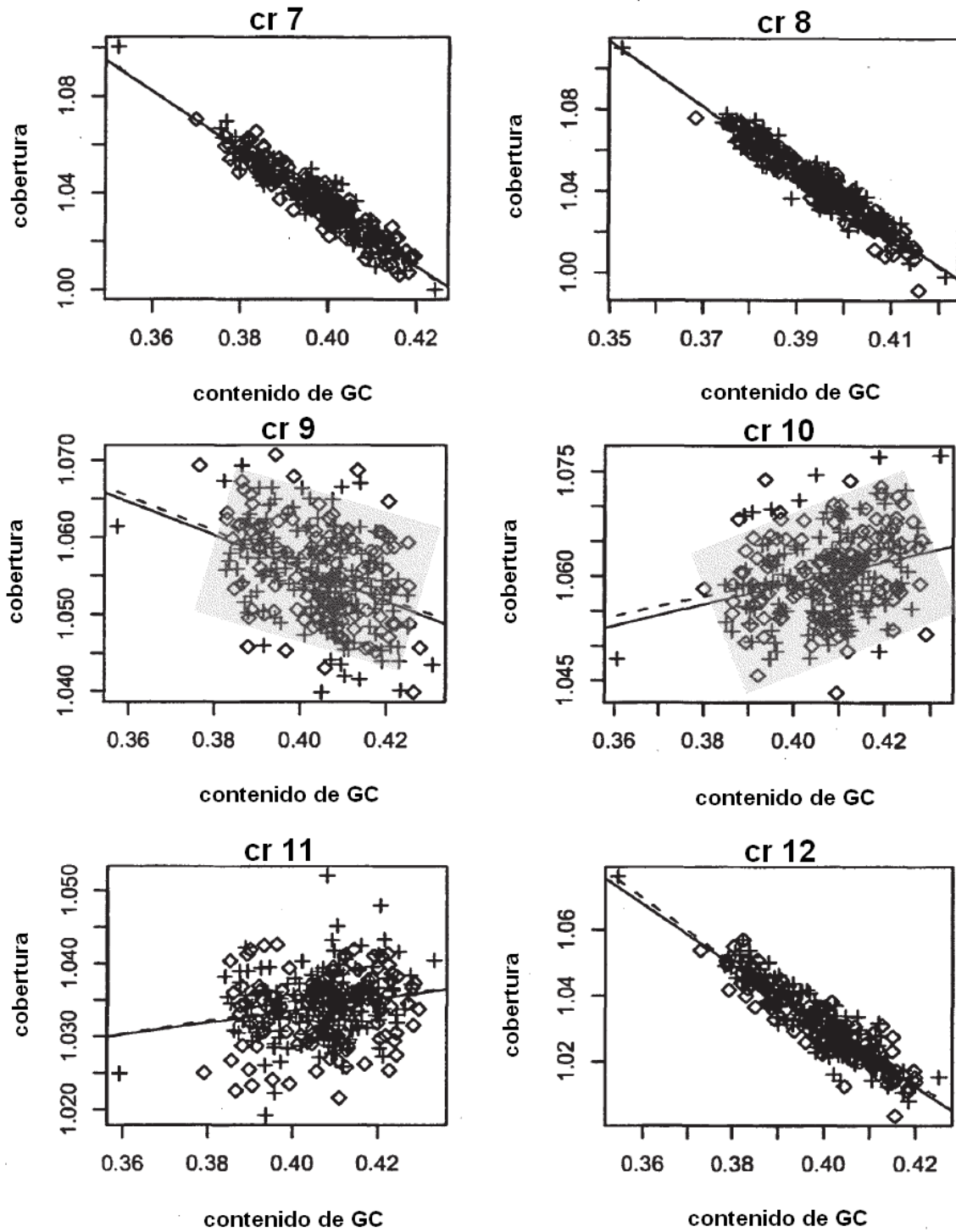


Figura 2
(página 2 de 4)

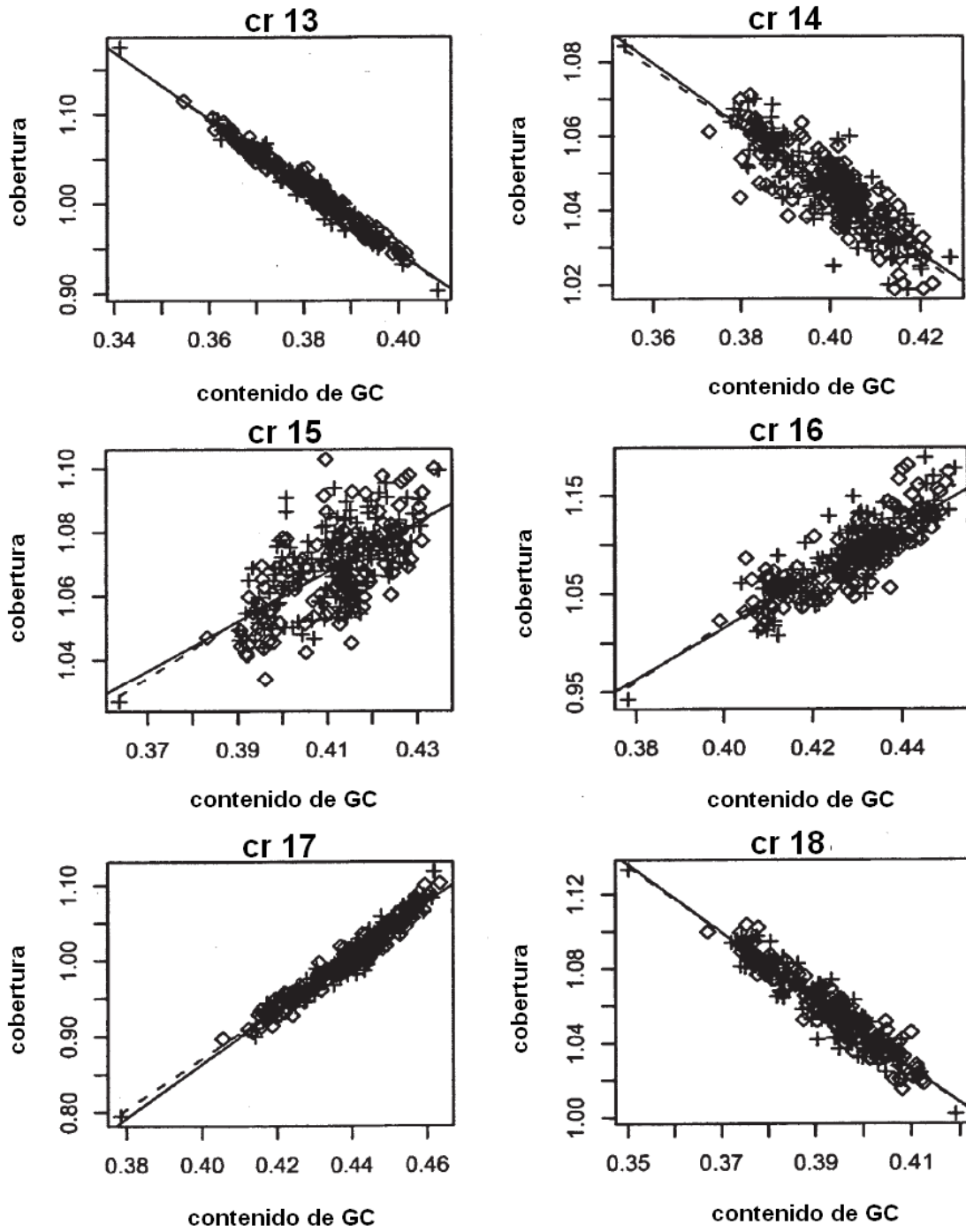


Figura 2
(página 3 de 4)

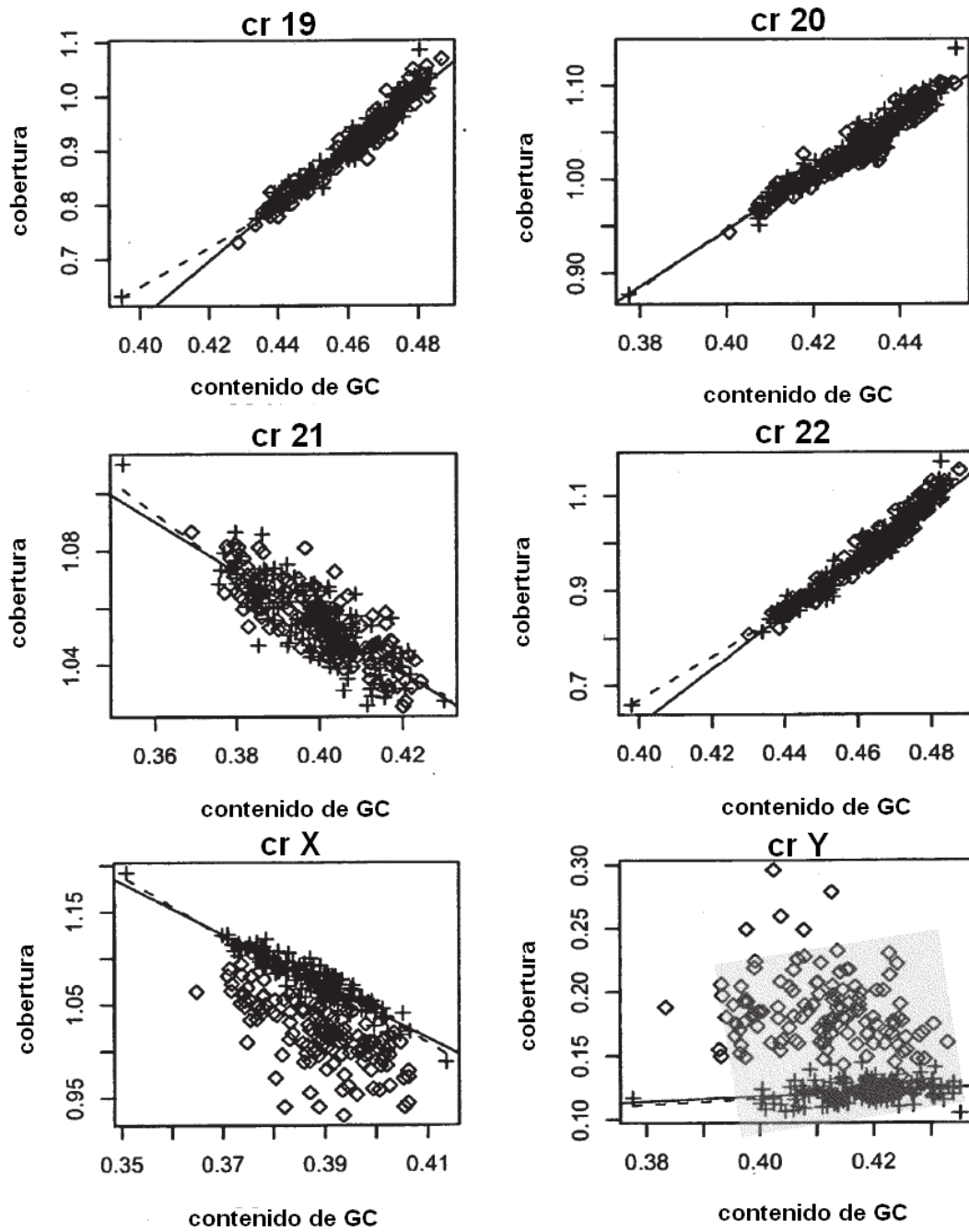
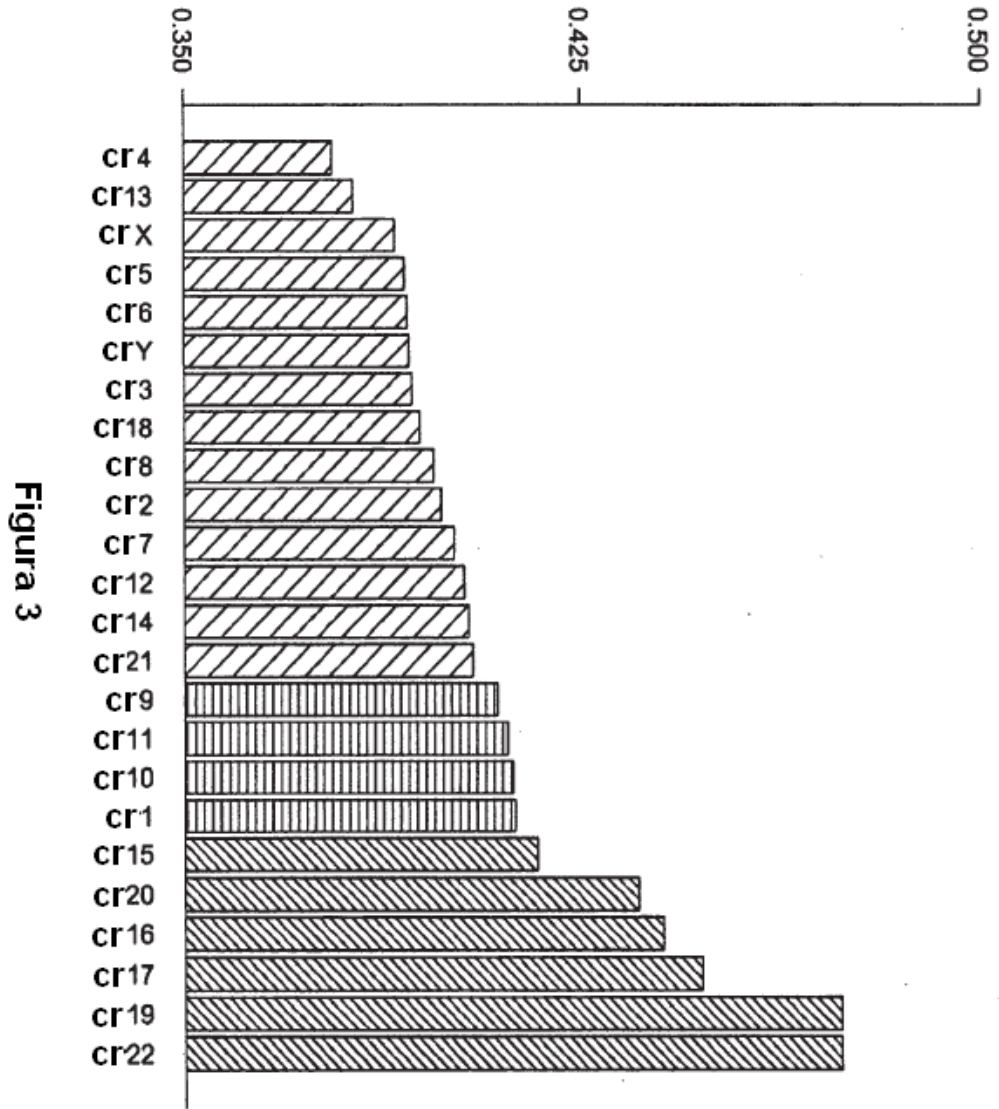


Figura 2
(página 4 de 4)



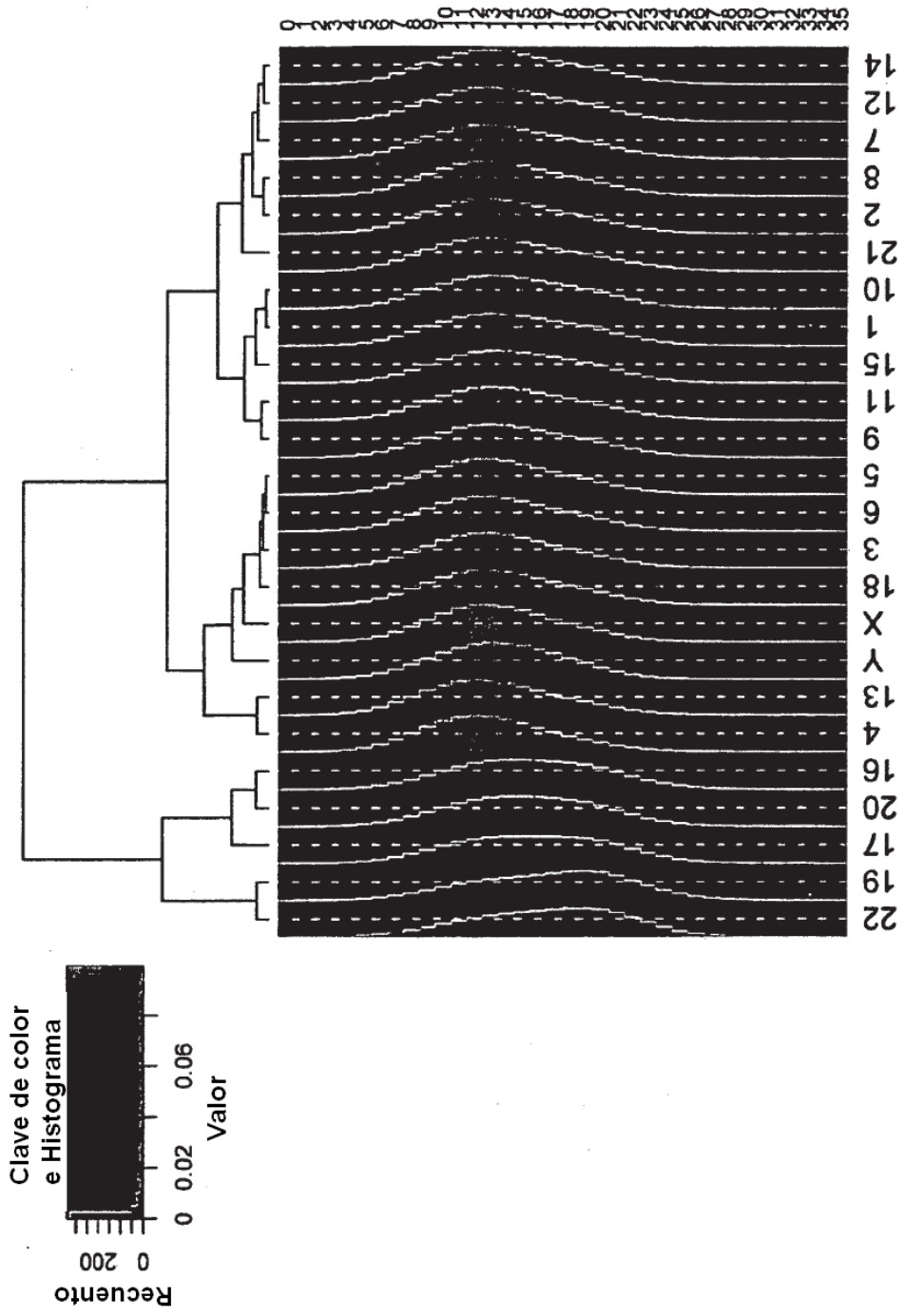


Figura 4

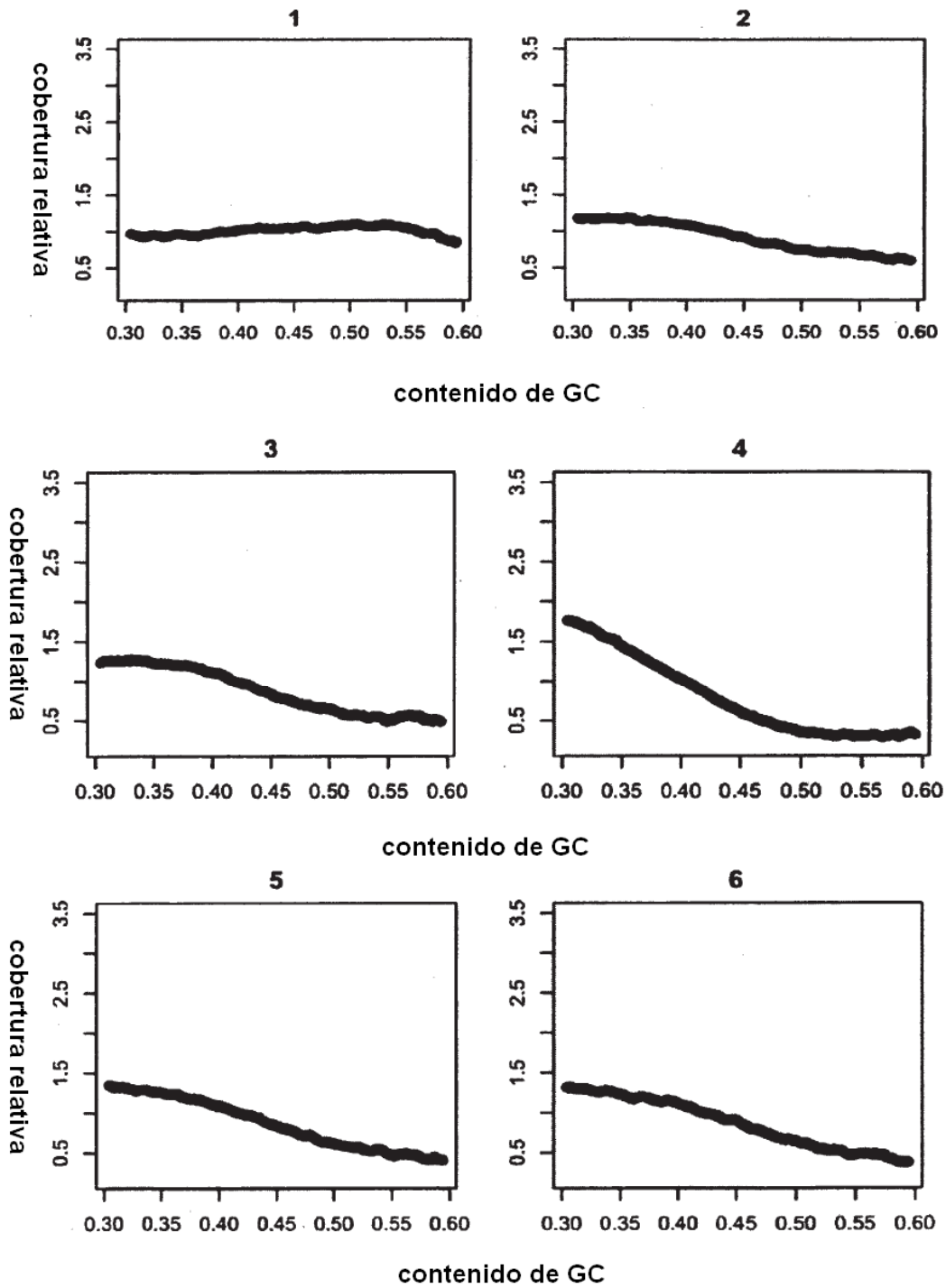


Figura 5
(página 1 de 4)

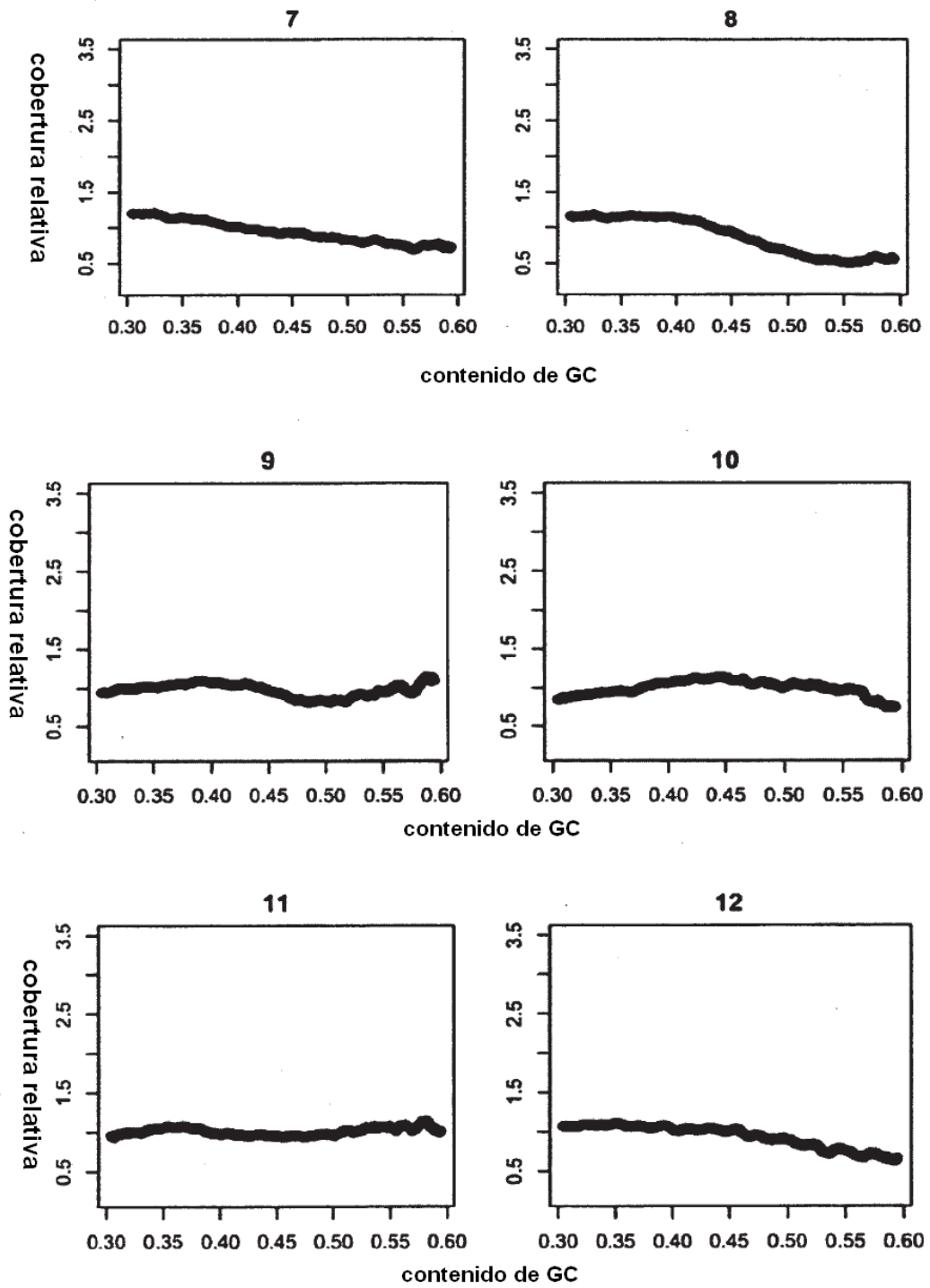


Figura 5
(página 2 de 4)

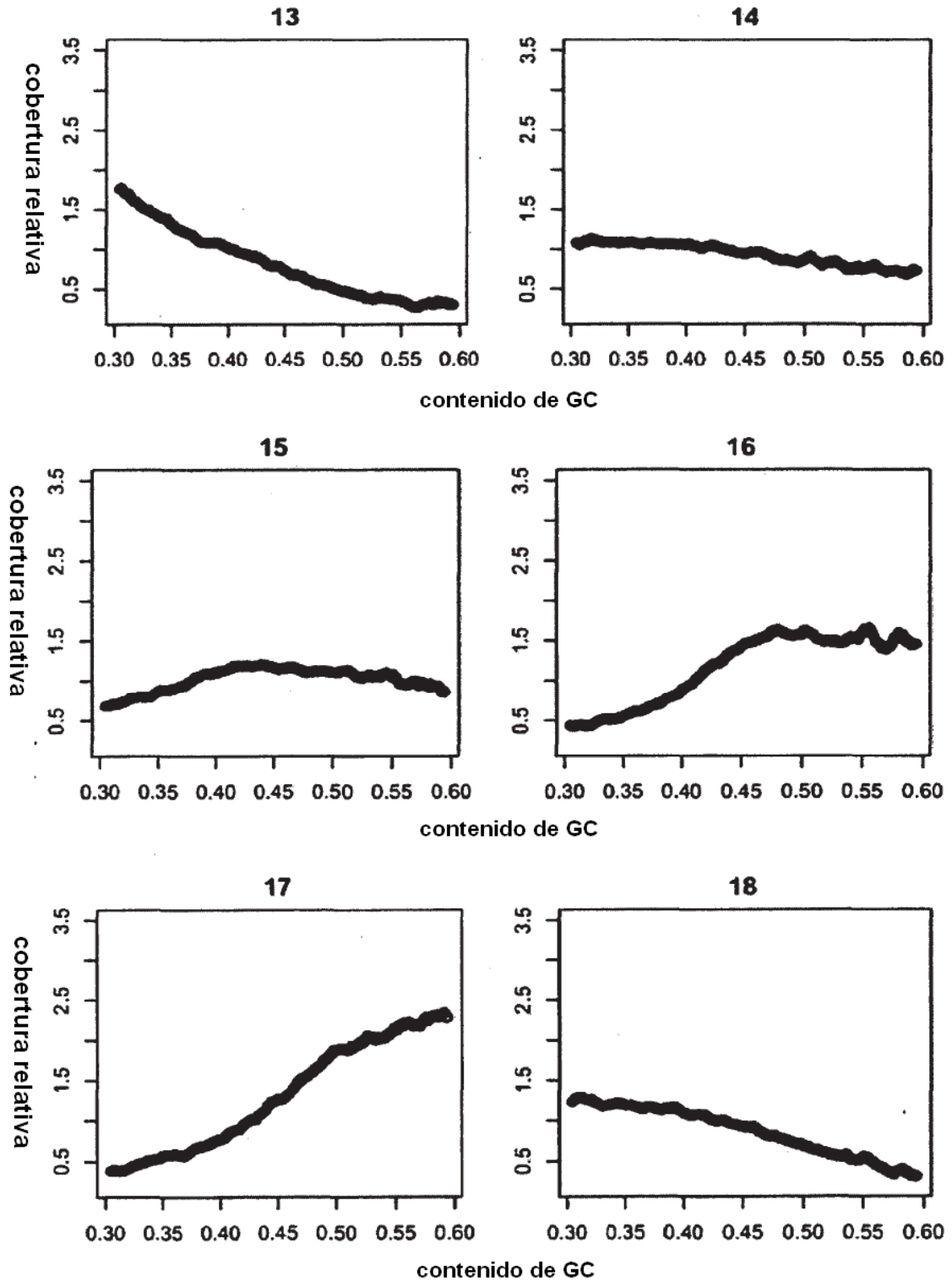


Figura 5
(página 3 de 4)

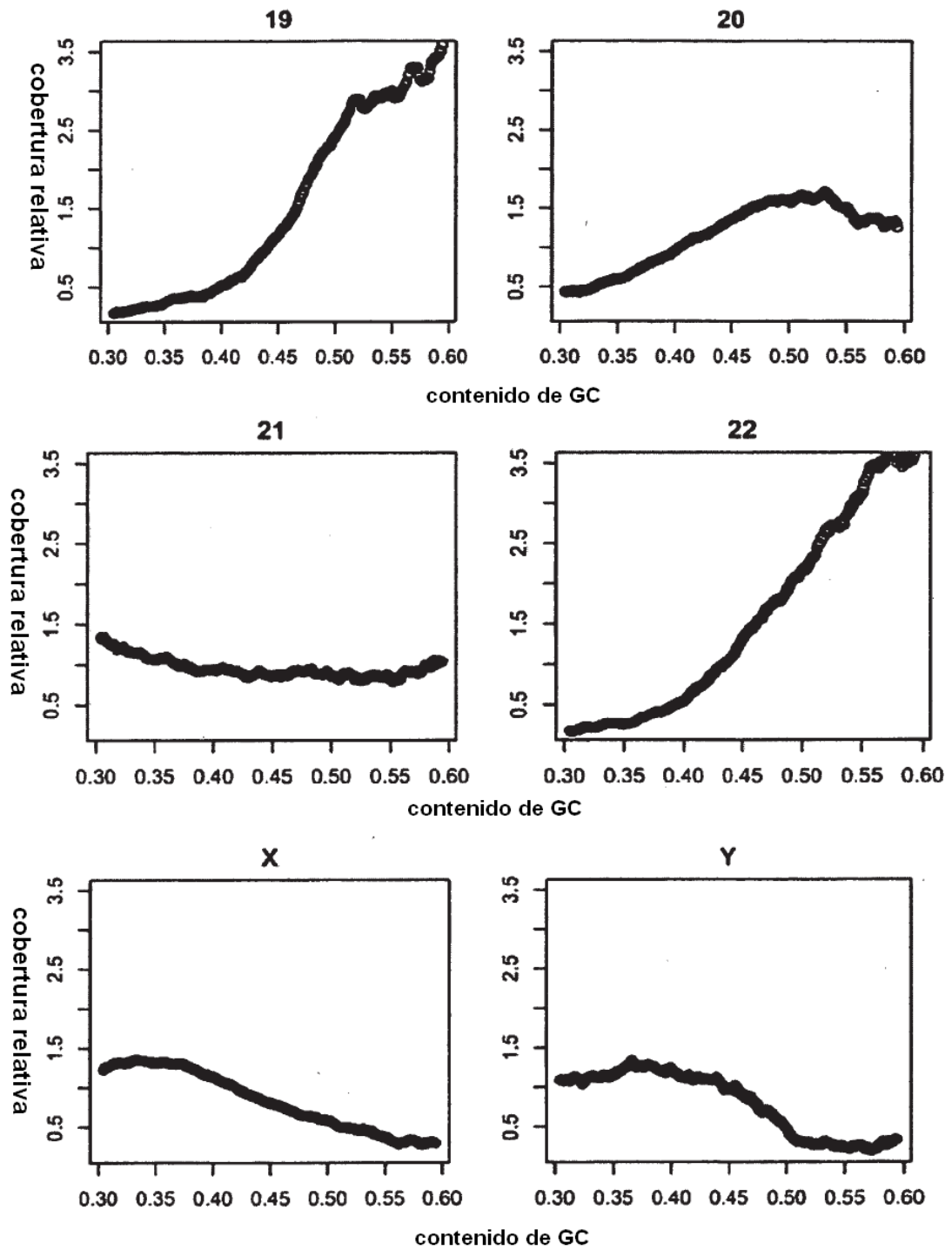


Figura 5
(página 4 de 4)

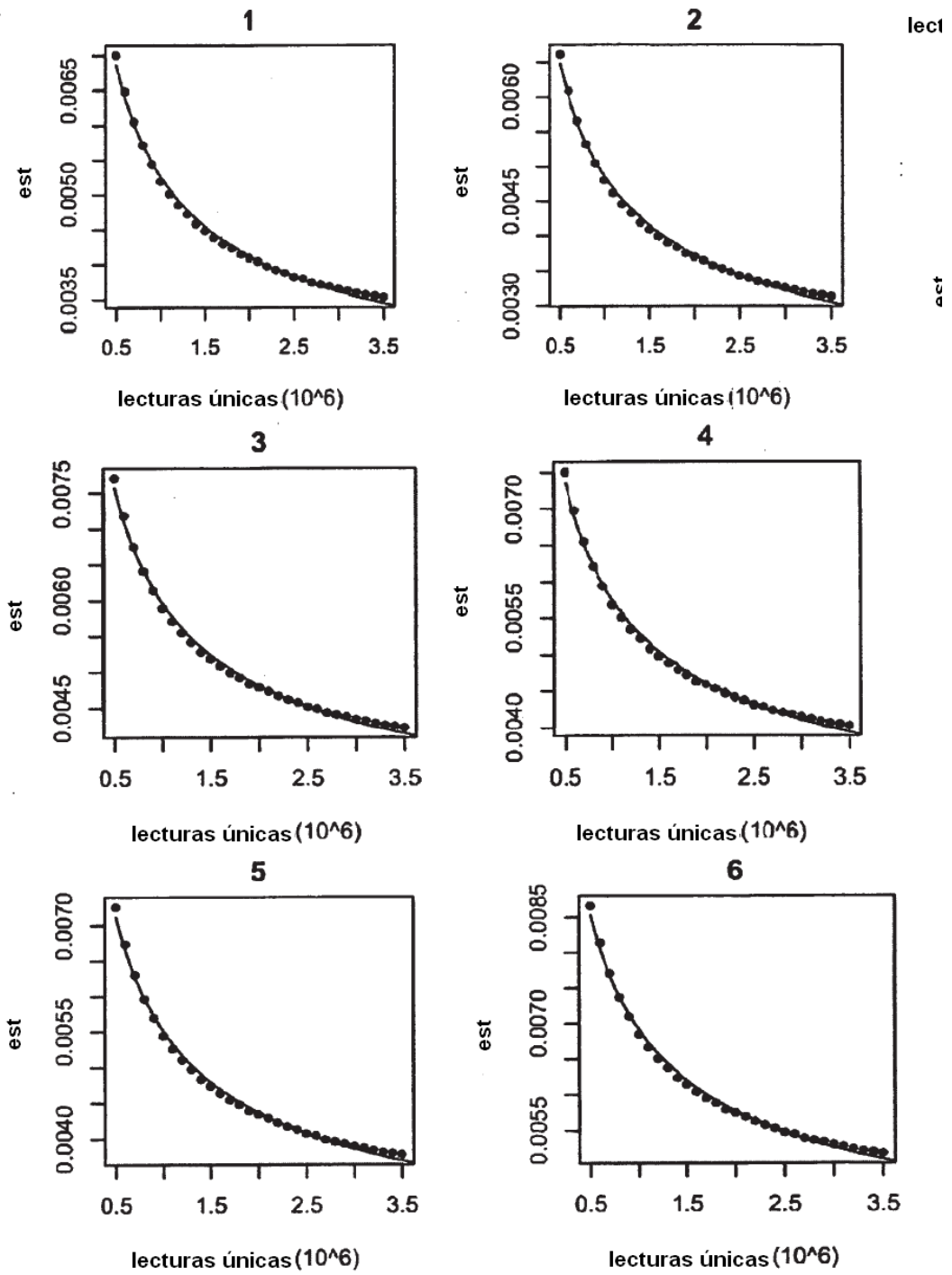


Figura 6
(página 1 de 4)

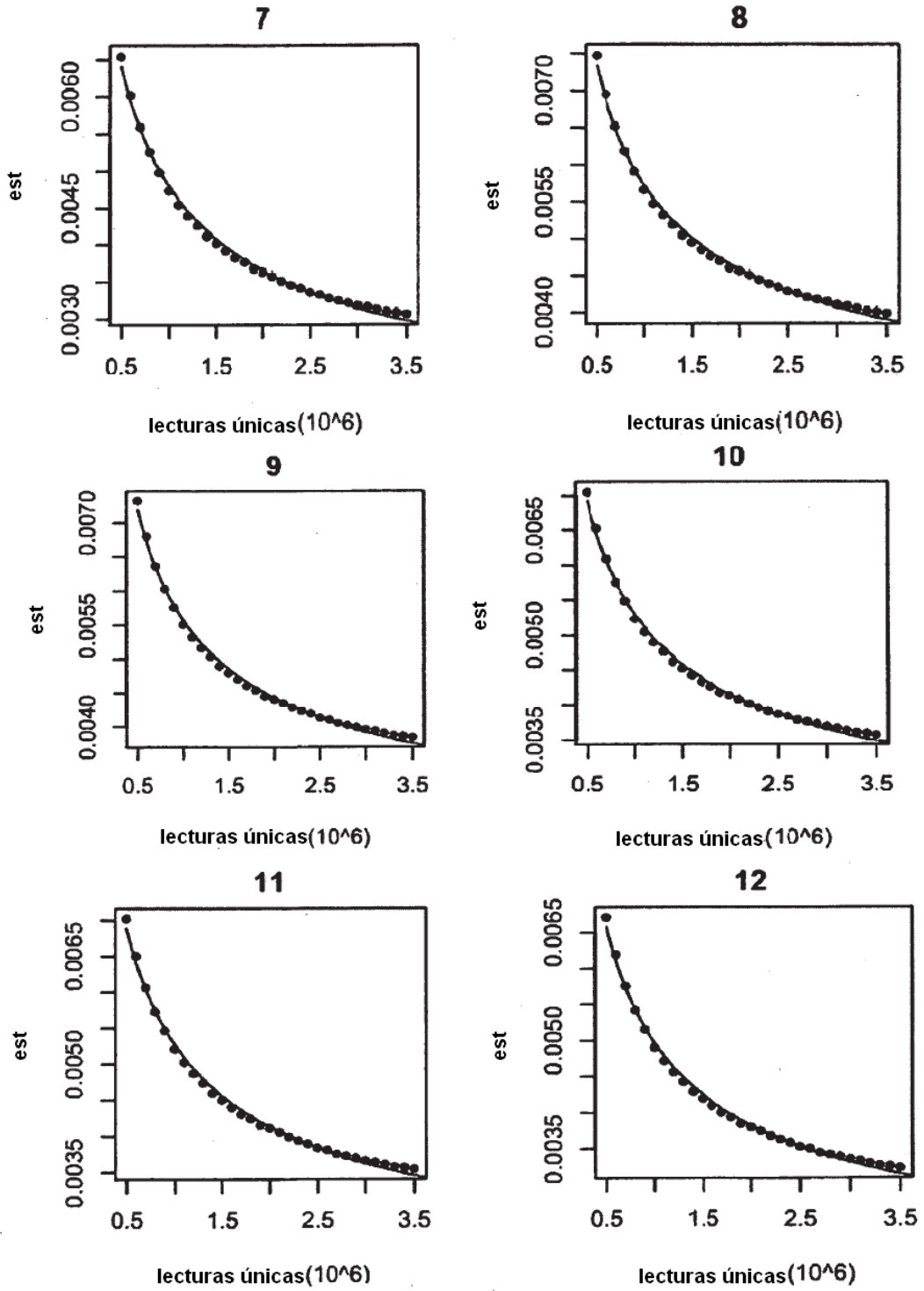


Figura 6
(página 2 de 4)

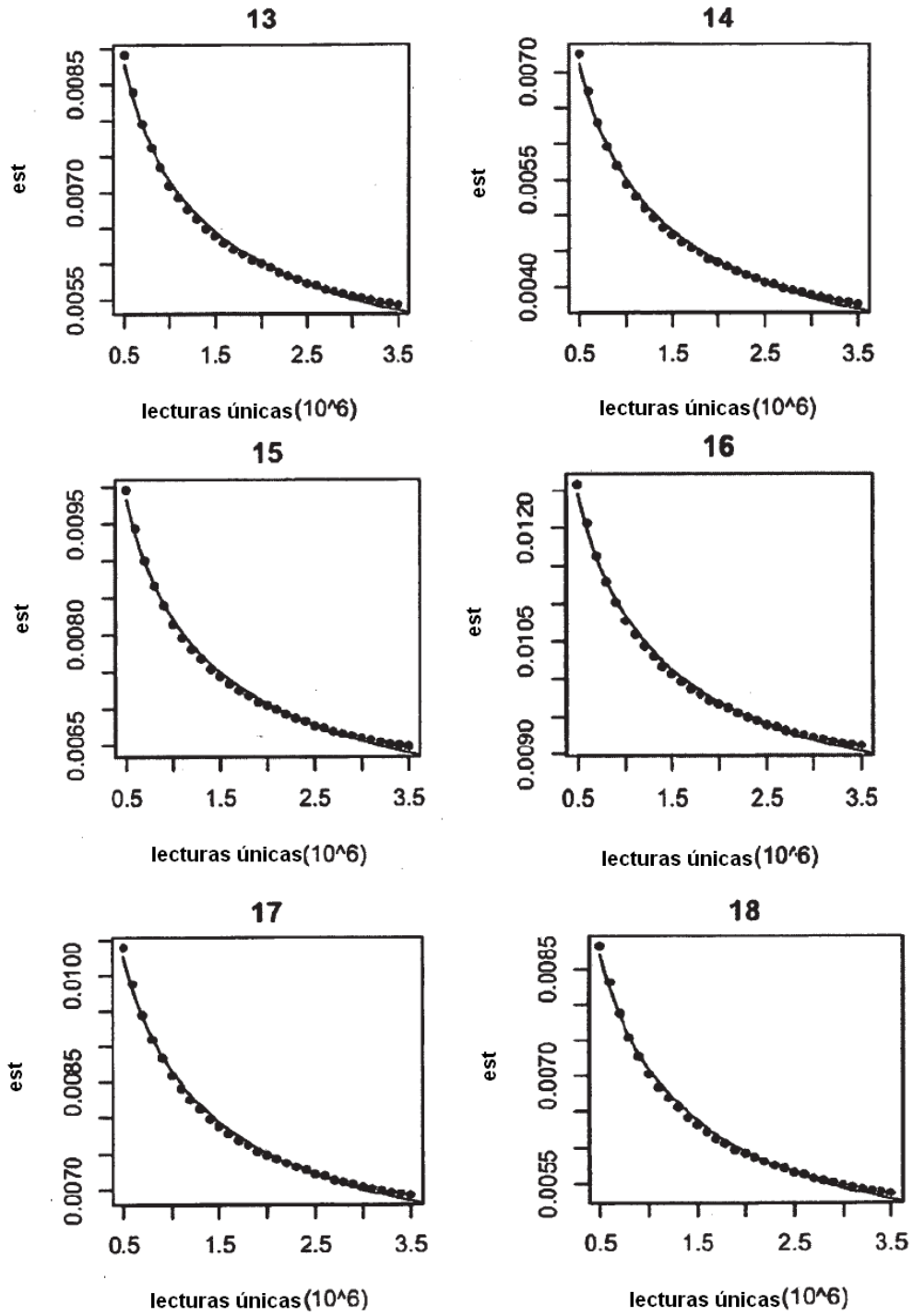


Figura 6
(página 3 de 4)

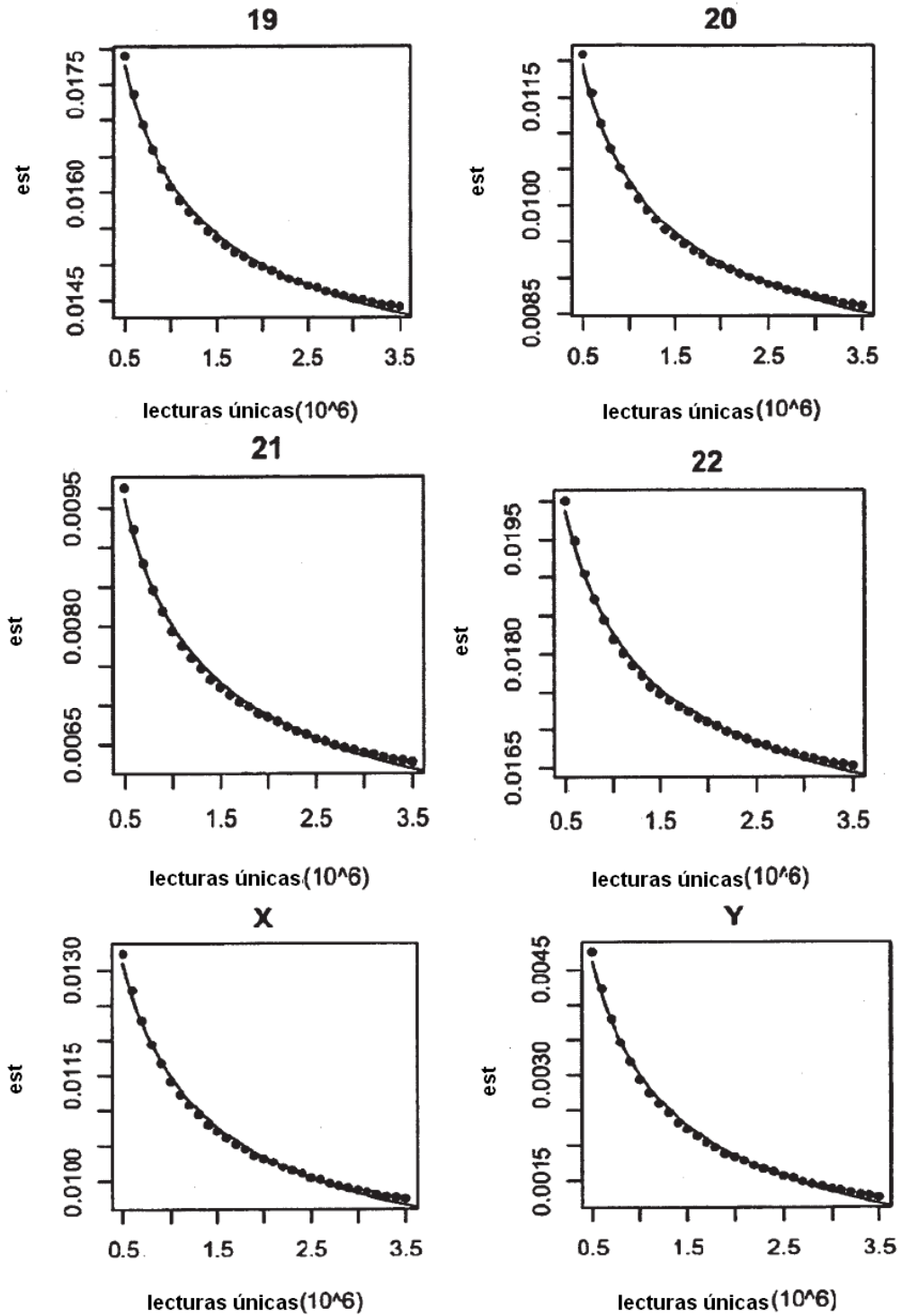


Figura 6
(página 4 de 4)

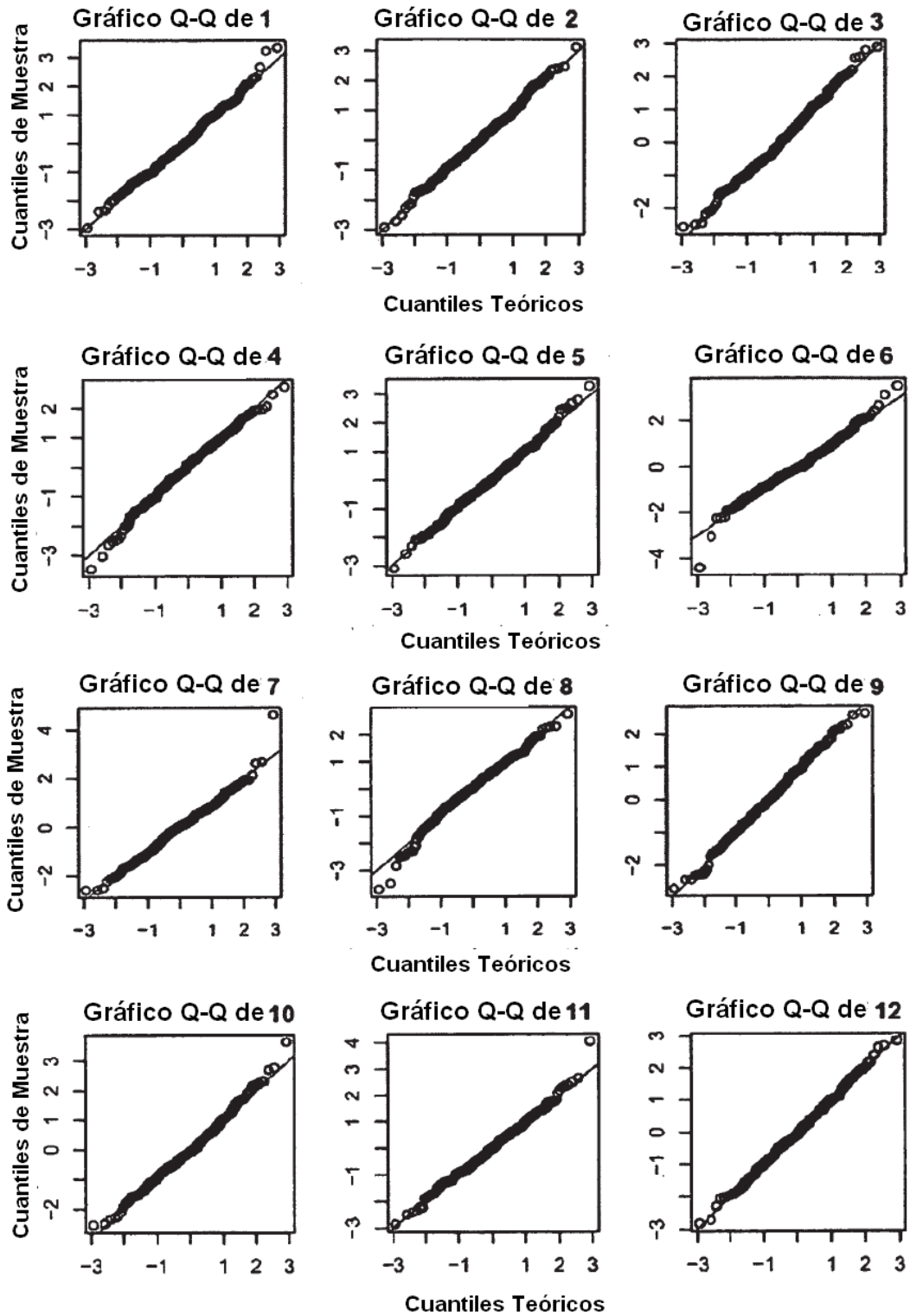


Figura 7
(página 1 de 2)

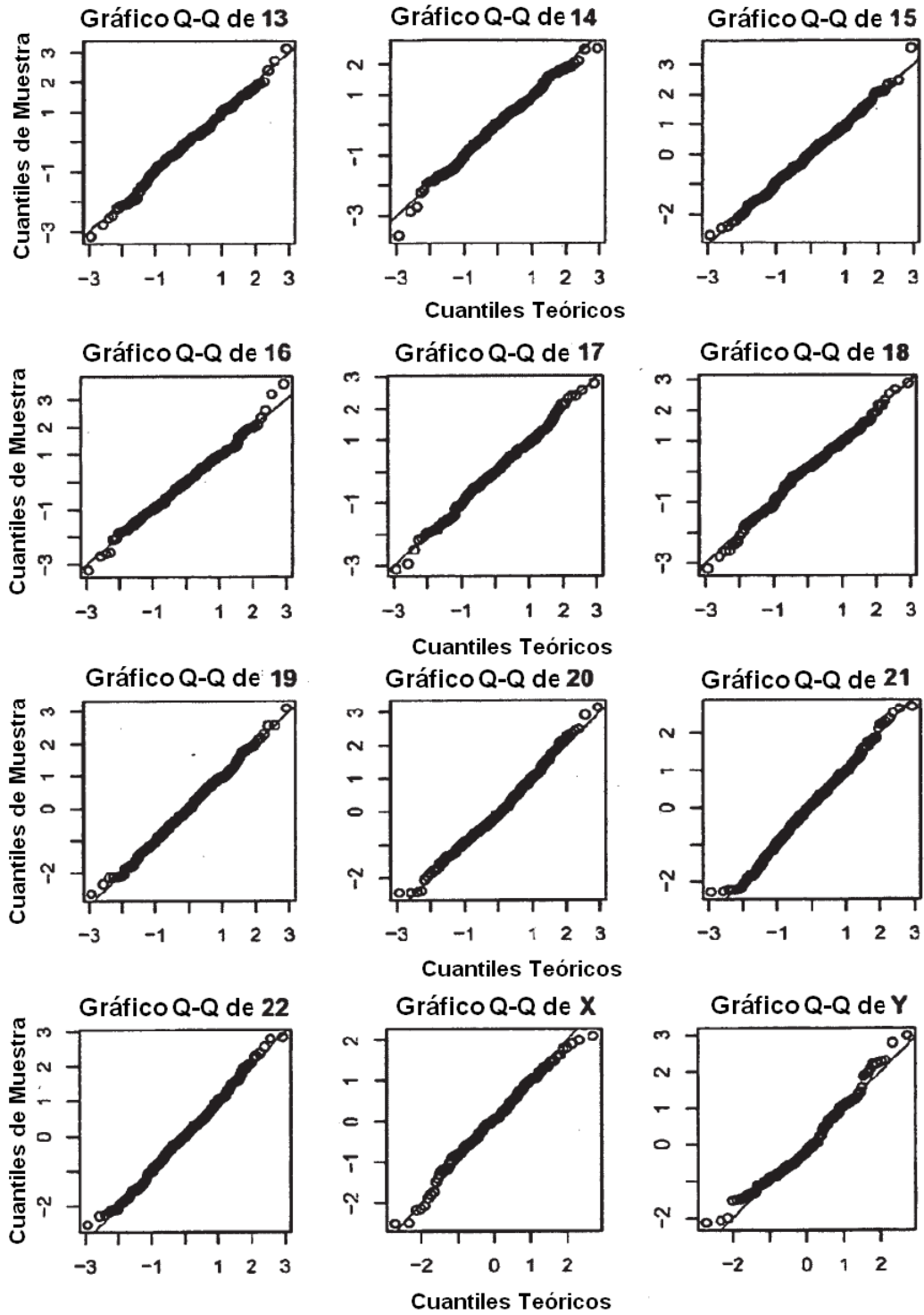


Figura 7
(página 2 de 2)

Histograma de cobertura relativa del cr Y

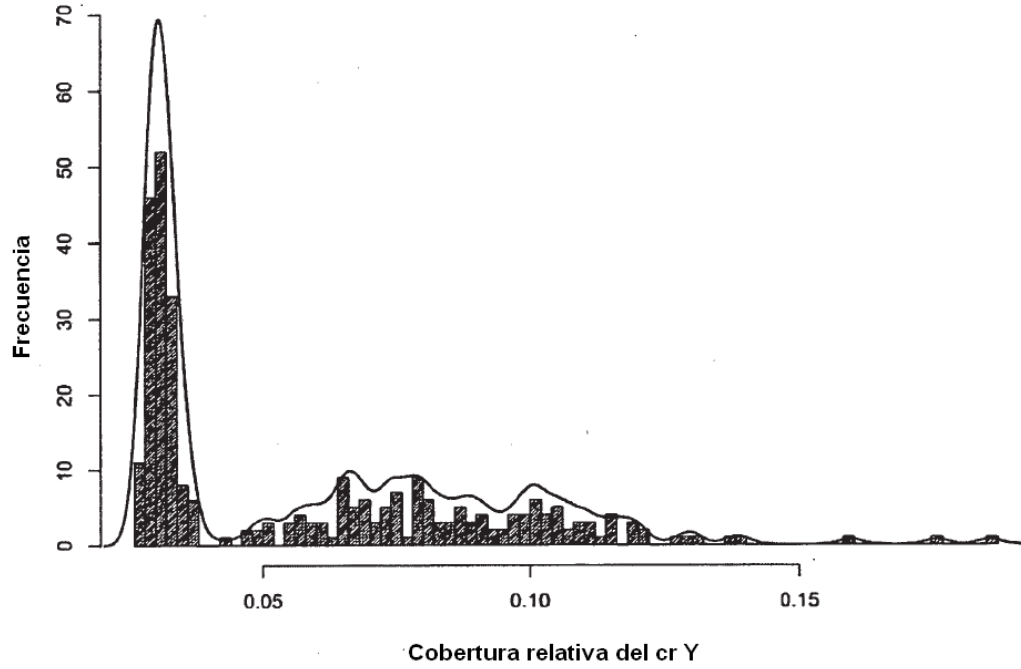


Figura 8

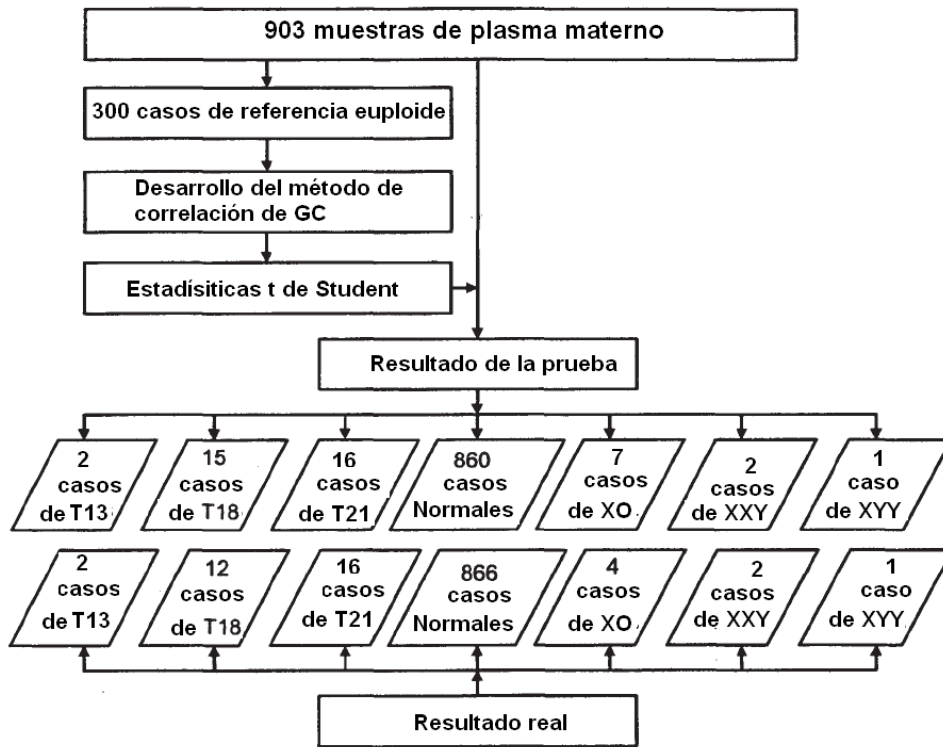


Figura 9

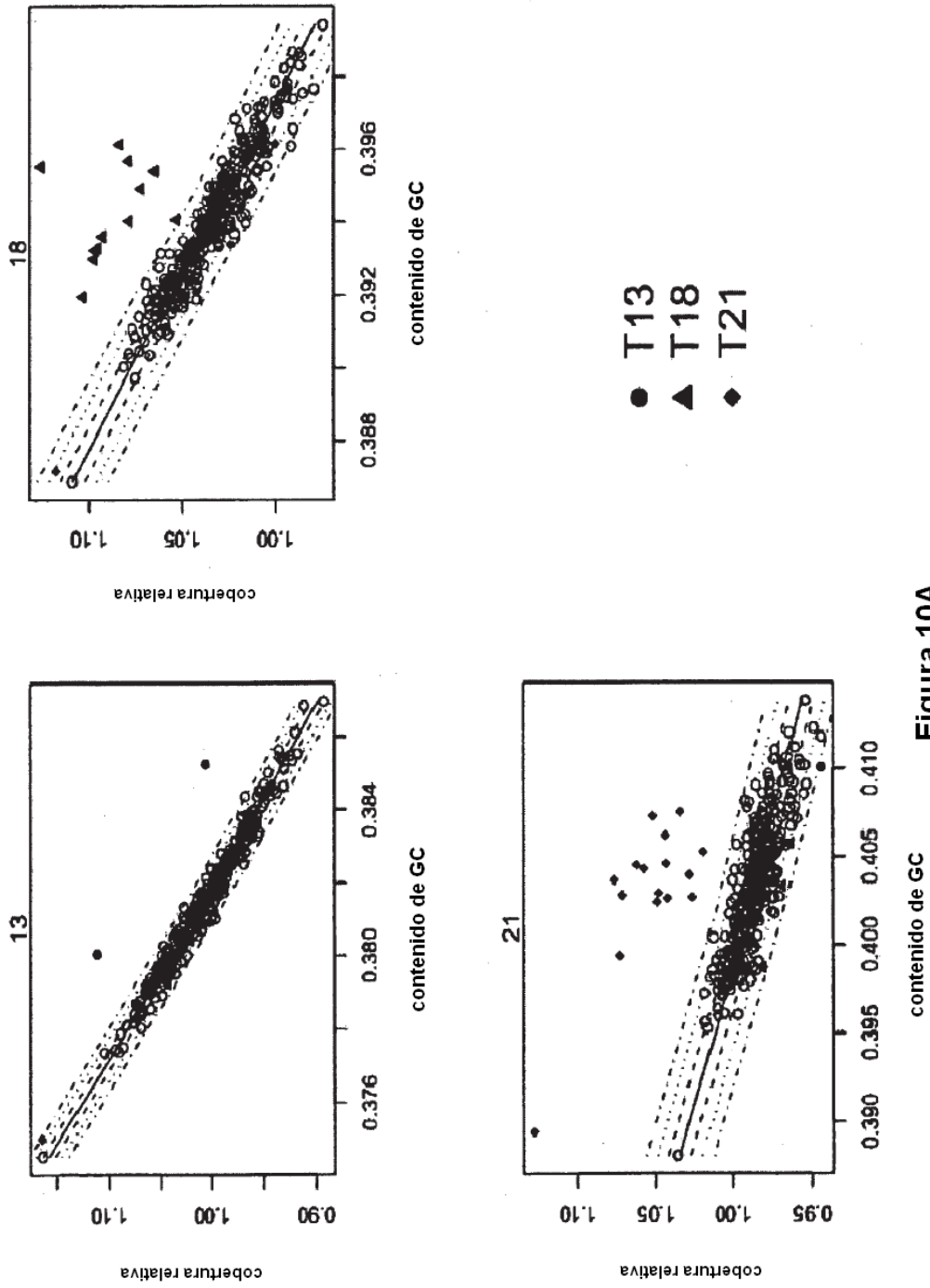


Figura 10A

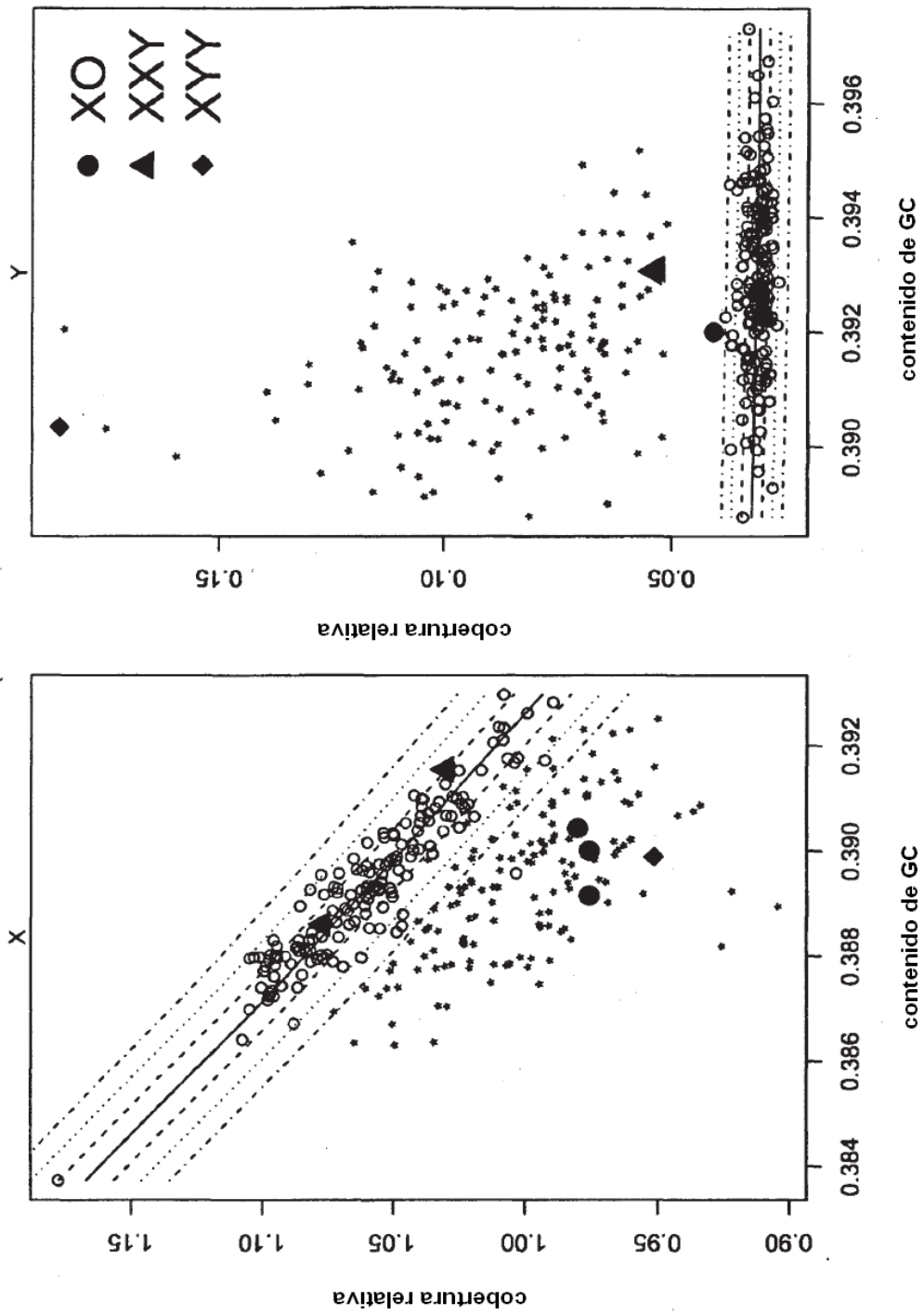


Figura 10B

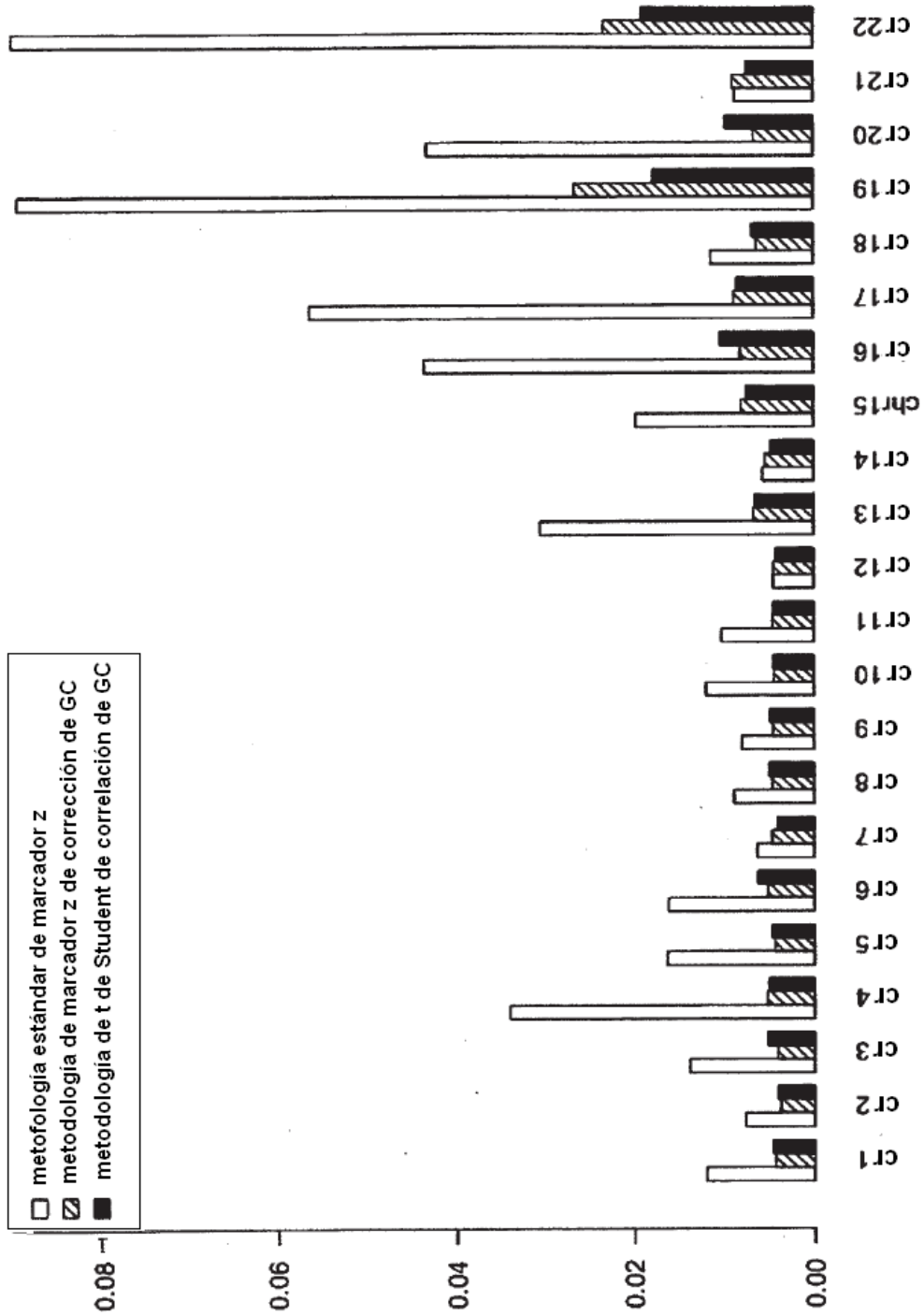


Figura 11

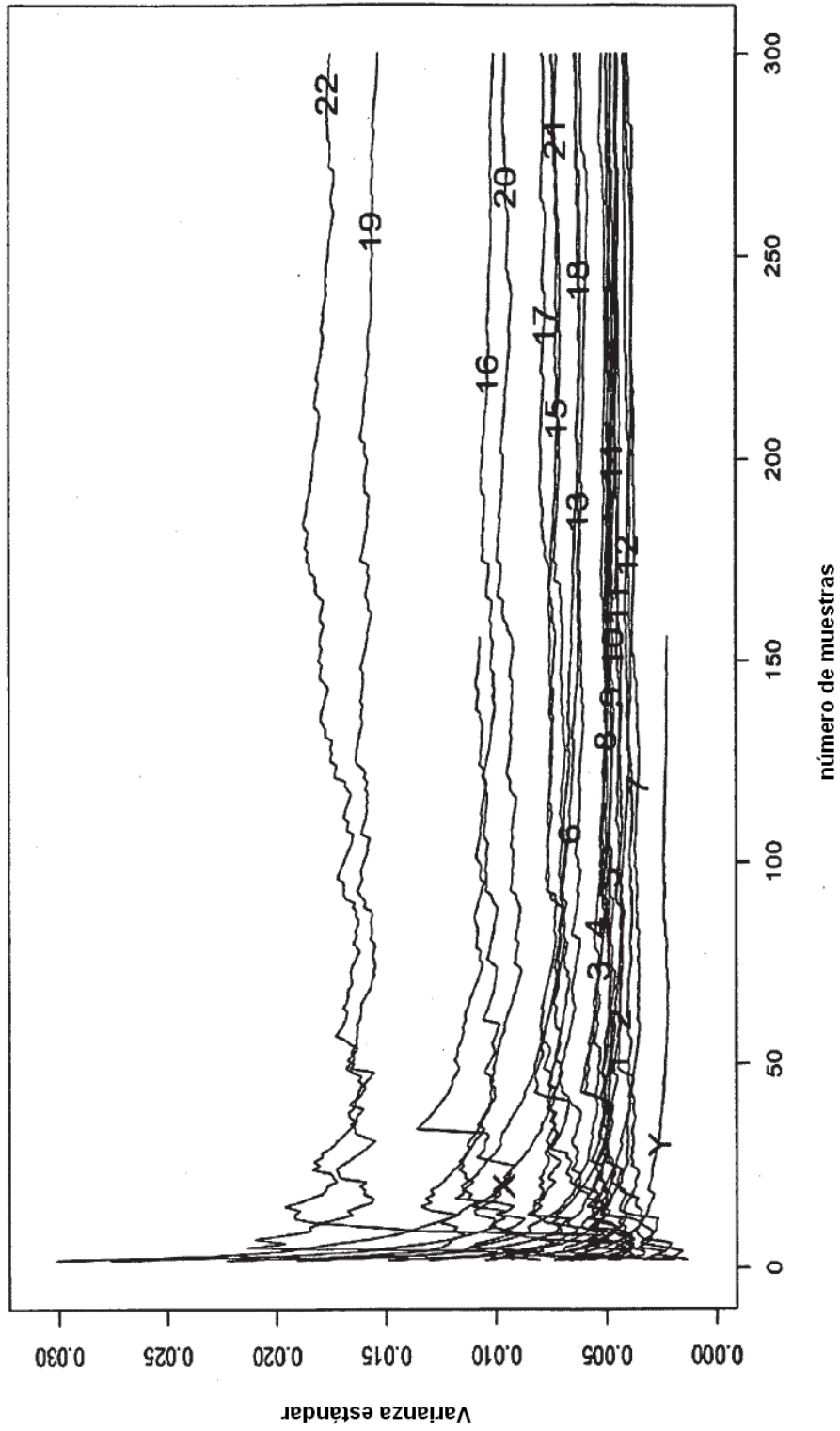
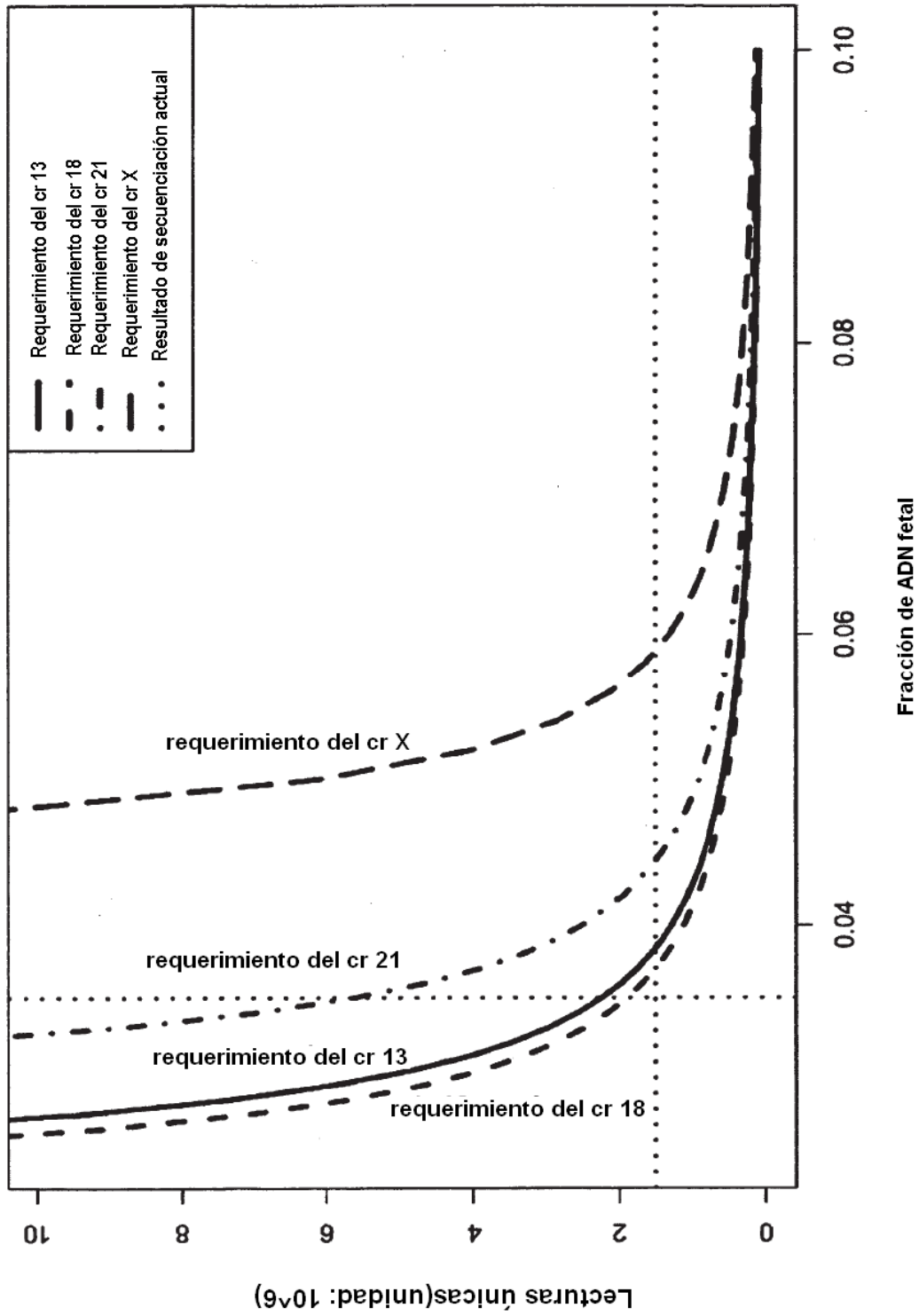


Figura 13



Fracción de ADN fetal

Figura 14

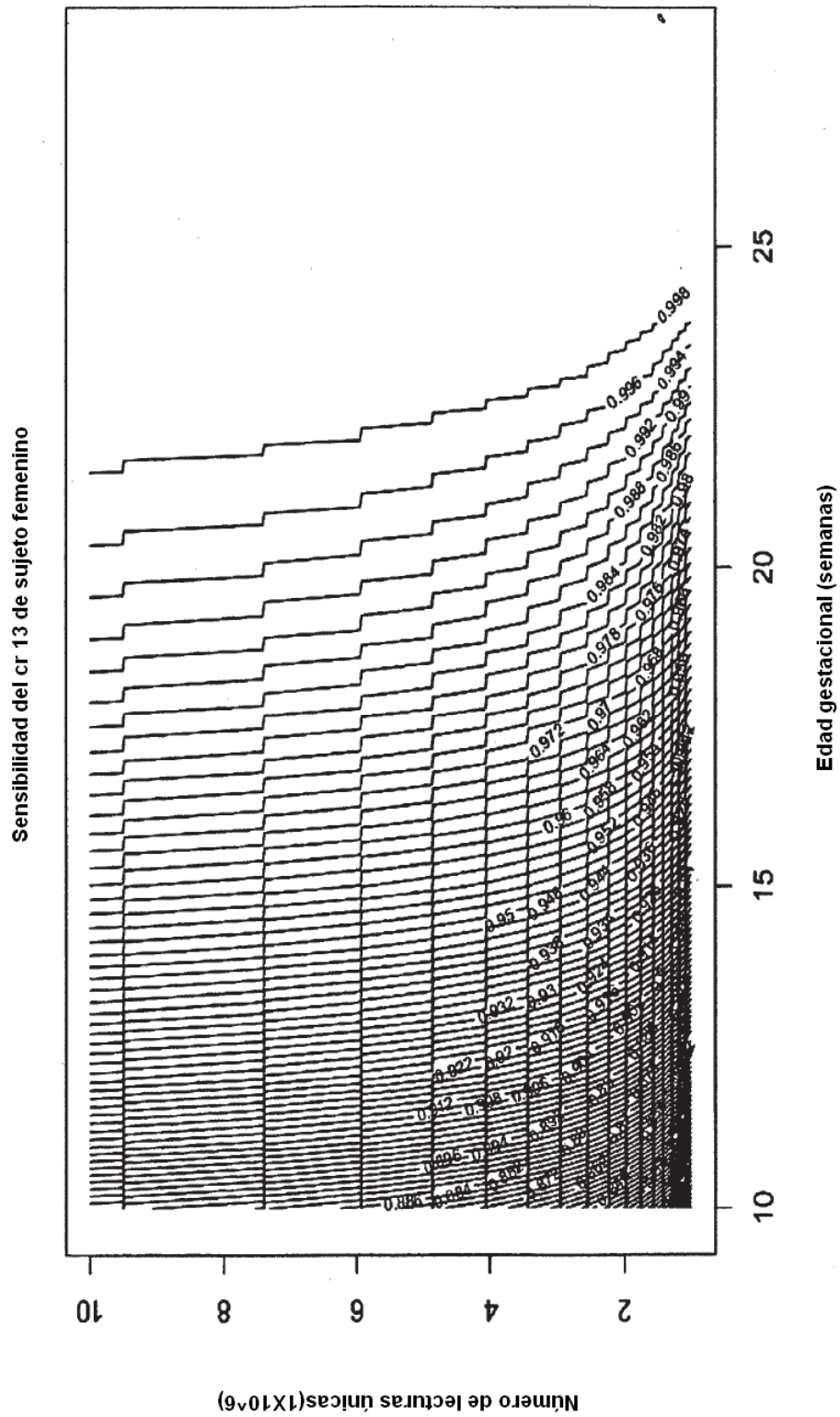


Figura 15

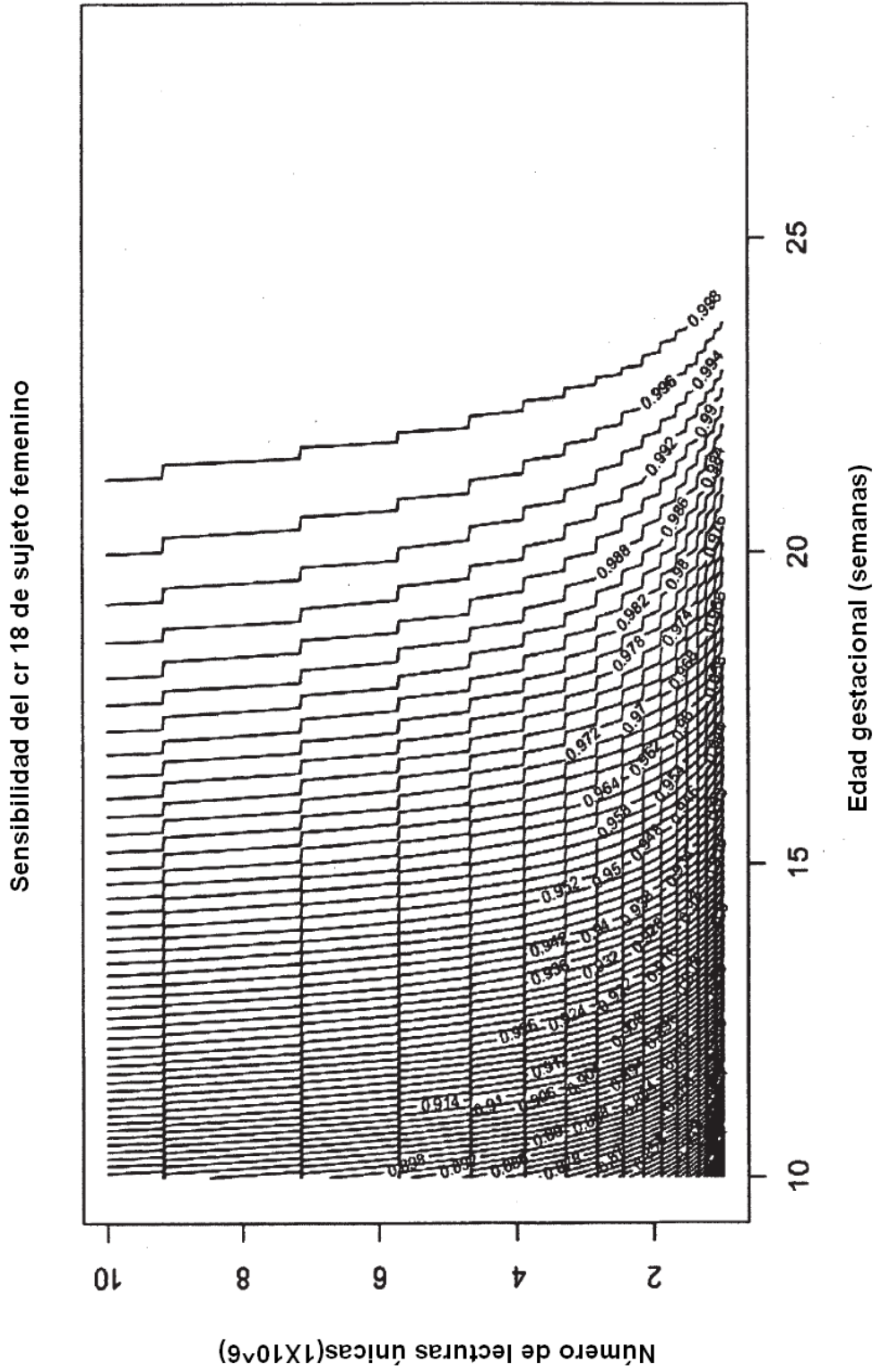
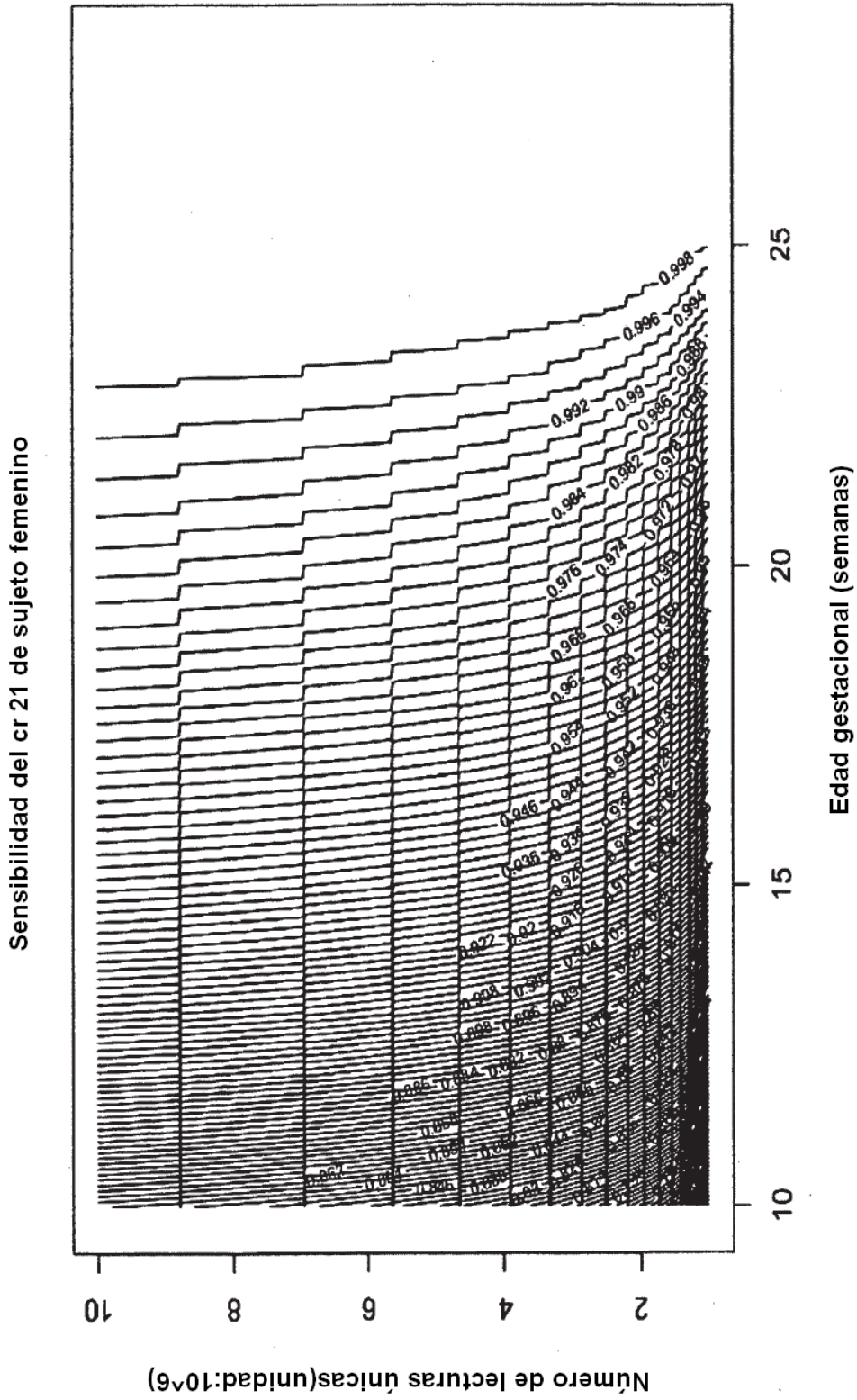
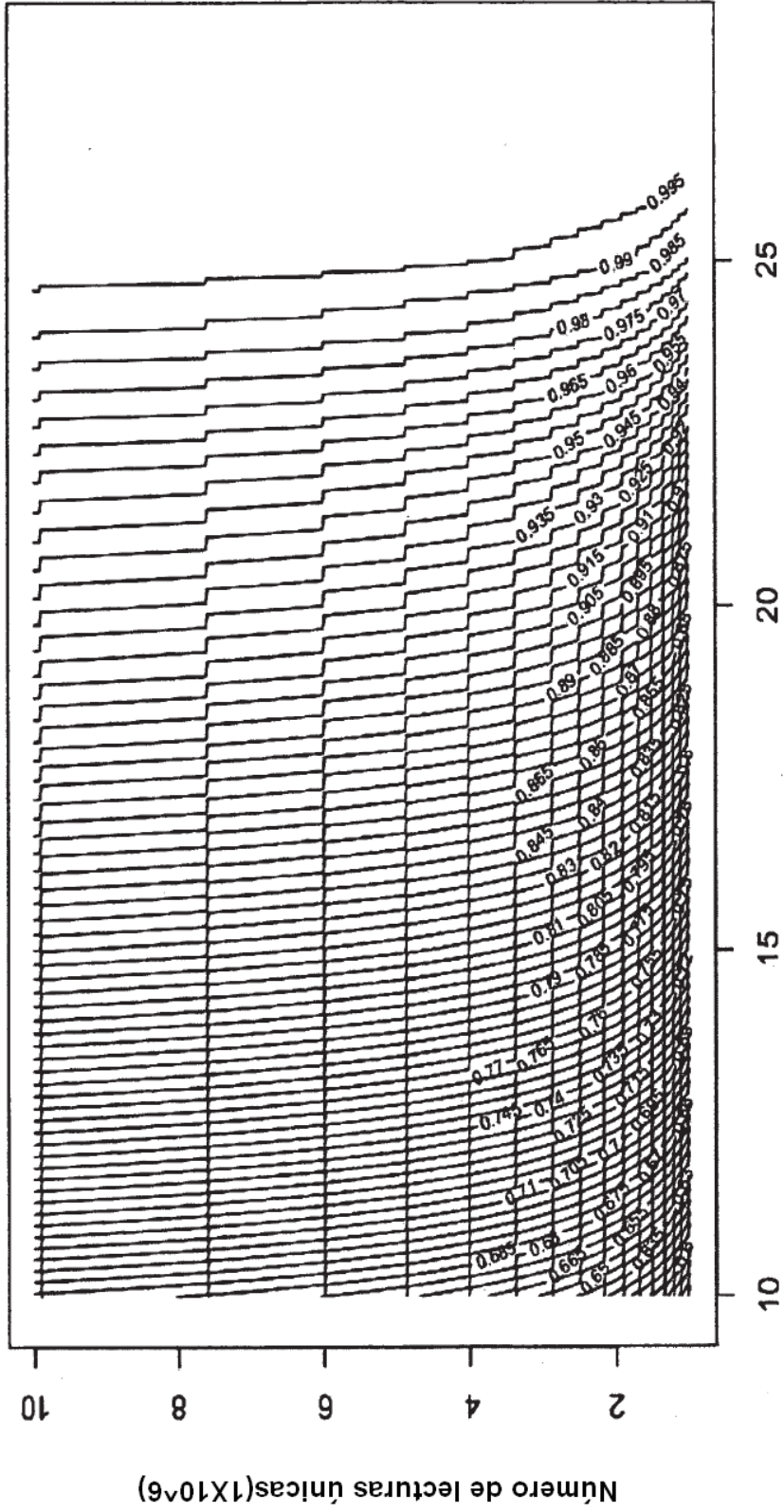


Figura 16



Sensibilidad del cr X: XO, XXX, XXY, etc.



Edad gestacional (semanas)

Figura 18

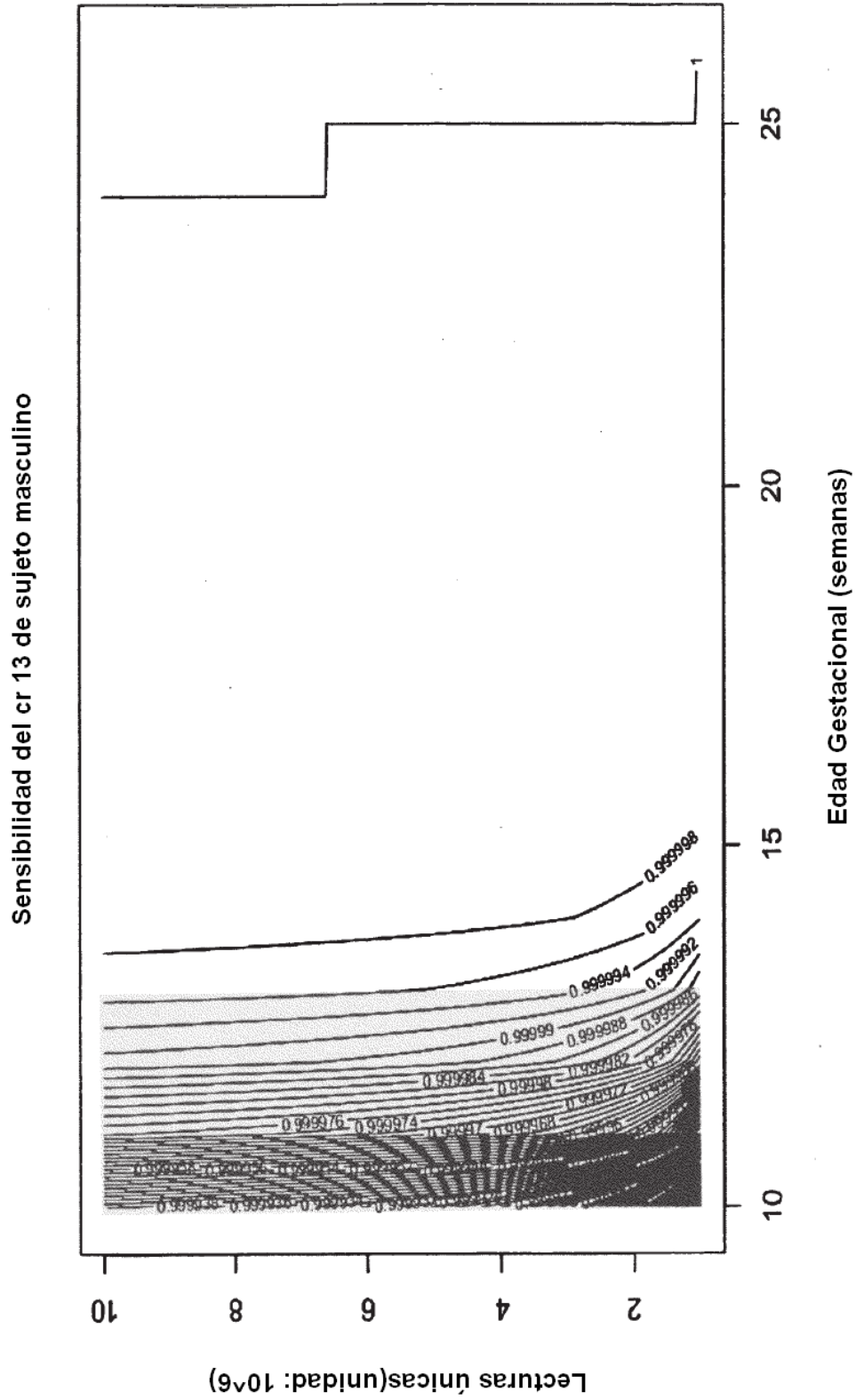


Figura 19

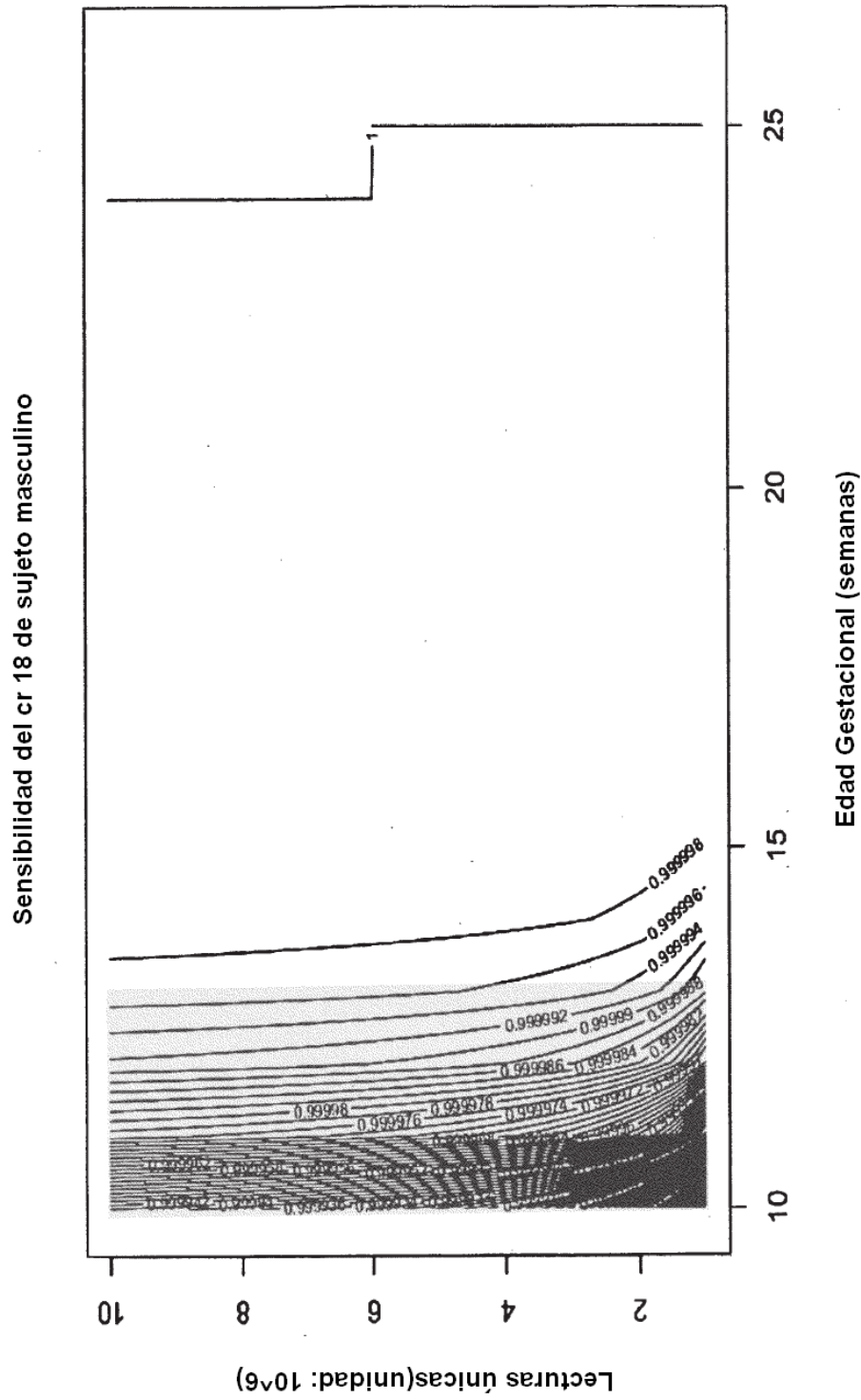


Figura 20

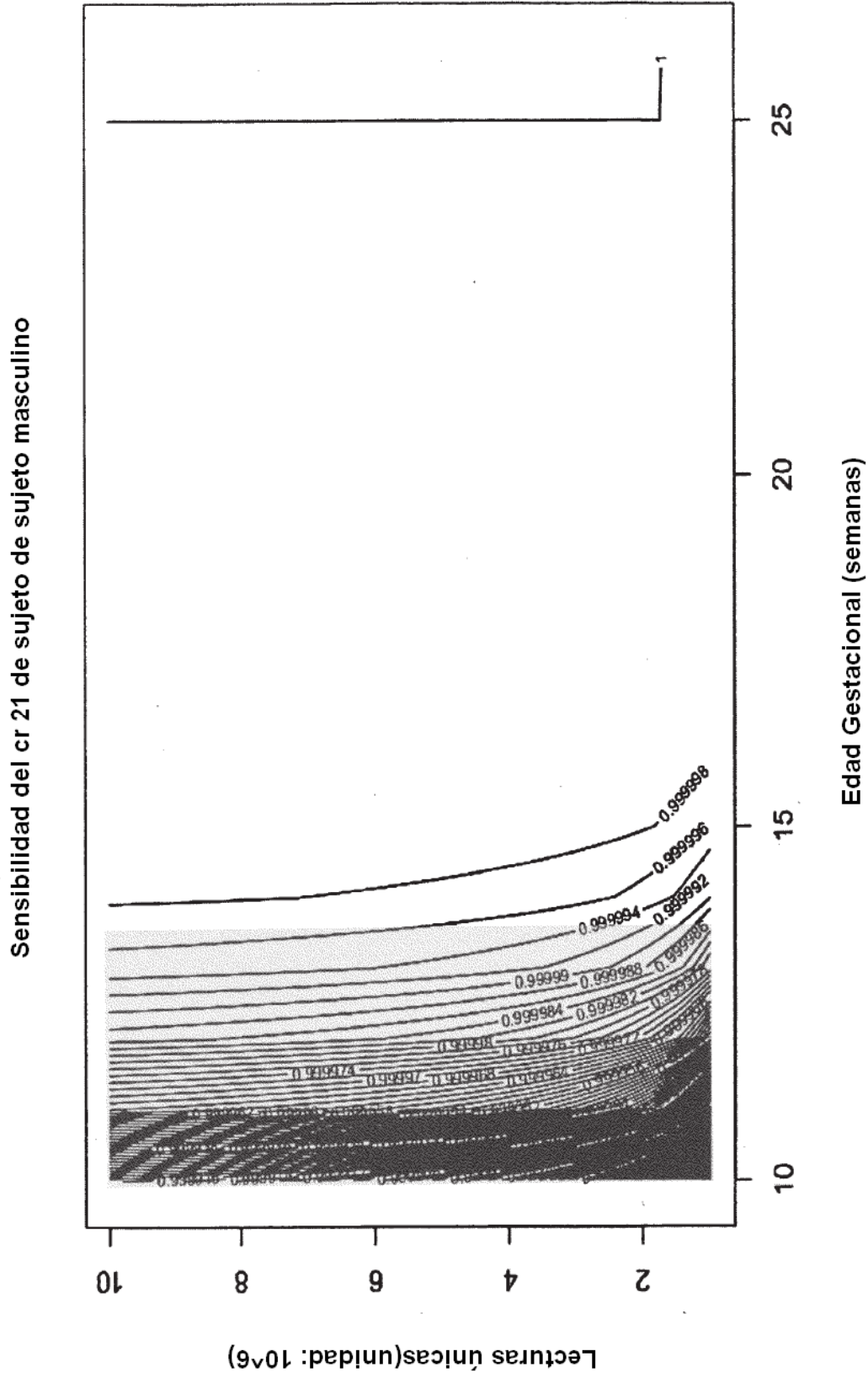


Figura 21