

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 986 657**

51 Int. Cl.:

**G01N 33/68**

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **10.03.2014** **E 19206123 (2)**

97 Fecha y número de publicación de la concesión europea: **24.07.2024** **EP 3660512**

54 Título: **Biomarcadores de la enfermedad de hígado graso no alcohólico (EHGNA) y de la esteatohepatitis no alcohólica (EHNA) y usos de los mismos**

30 Prioridad:

**15.03.2013 US 201361787967 P**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**12.11.2024**

73 Titular/es:

**SOMALOGIC OPERATING CO., INC. (100.0%)**  
**2945 Wilderness Place**  
**Boulder, CO 80301, US**

72 Inventor/es:

**NIKRAD, MALTI;**  
**FIELD, STUART G. y**  
**WILLIAMS, STEPHEN ALARIC**

74 Agente/Representante:

**VALLEJO LÓPEZ, Juan Pedro**

ES 2 986 657 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

Biomarcadores de la enfermedad de hígado graso no alcohólico (EHGNA) y de la esteatohepatitis no alcohólica (EHNA) y usos de los mismos

**Campo de la invención**

La presente solicitud se refiere, en general, a la detección de biomarcadores y a la caracterización de la enfermedad del hígado graso no alcohólico (EHGNA), por ejemplo, para identificar sujetos con esteatosis y esteatohepatitis no alcohólica (EHNA). En diversas realizaciones, la invención se refiere a uno o más biomarcadores y reactivos para caracterizar la EHGNA y la EHNA en un individuo.

**Antecedentes**

La siguiente descripción proporciona un resumen de información y no es una admisión de que cualquier información o publicaciones proporcionadas referidas en el presente documento es de la técnica anterior a la presente solicitud.

La enfermedad del hígado graso no alcohólico (HGNA) se define como la presencia de esteatosis hepática, con o sin inflamación y fibrosis, en ausencia de antecedentes de alcoholismo. La EHGNA se subdivide en hígado graso no alcohólico (HGNA) y esteatohepatitis no alcohólica (EHNA). En el HGNA, la esteatosis hepática está presente sin evidencia de inflamación significativa, mientras que en la EHNA, la esteatosis hepática se asocia con inflamación hepática que puede ser histológicamente indistinguible de la esteatohepatitis alcohólica.

La EHGNA se ha convertido en una epidemia en todo el mundo, y es la principal causa de enfermedad hepática en Norteamérica, como resultado de la creciente frecuencia de la obesidad. Sin embargo, los datos exactos basados en la población sobre la incidencia del HGNA y la EHNA son escasos, en parte, porque el diagnóstico requiere documentación histopatológica. Los principales factores de riesgo para la EHGNA son la obesidad central, la diabetes mellitus de tipo 2, los niveles altos de triglicéridos (grasa) en sangre y la presión arterial alta. En EE. UU., la EHGNA está presente en el 20-40 % de la población y la EHNA está presente en aproximadamente el 25 % de la población obesa. Entre el diez y el veinte por ciento de los pacientes con EHNA desarrollan cirrosis, y entre el 4 % y el 27 % de ellos desarrollan cáncer hepático.

La mayoría de las personas con EHNA no tienen síntomas. Algunas pueden presentar dolor en el cuadrante superior derecho, hepatomegalia o síntomas inespecíficos, tales como malestar abdominal, debilidad, fatiga o malestar general. Un médico o profesional de enfermería puede sospechar de la presencia de EHNA a partir de los resultados de los análisis de sangre rutinarios. En la EHGNA, las enzimas hepáticas aspartato aminotransferasa (AST) y alanina aminotransferasa (ALT) suelen ser altas.

El patrón de referencia actual para confirmar la EHNA es una evaluación histológica de la biopsia hepática, que es costosa, invasiva y puede causar dolor, hemorragia e incluso la muerte.

Un simple análisis de sangre que identificara y distinguiera las diversas etapas de la EHGNA y EHNA (y, por lo tanto, redujera la necesidad de realizar una biopsia hepática) sería altamente deseable.

Bell et al, Hepatology, 51 (1), 111-120 (2010) describieron la proteómica sérica y el descubrimiento de biomarcadores en todo el espectro de la EHGNA.

**Sumario**

La invención se expone en el conjunto de reivindicaciones anexas.

En el presente documento se describen, pero no se reivindican, métodos de determinación de si un sujeto tiene la enfermedad de hígado graso no alcohólico (EHGNA). En el presente documento se describen métodos de identificación de sujetos con esteatosis y métodos de determinación de la gravedad de la esteatosis.

En las realizaciones reivindicadas de la invención, se proporcionan métodos de determinación de si un sujeto tiene esteatohepatitis no alcohólica (EHNA). En algunas realizaciones, se proporcionan métodos de identificación de sujetos con EHNA. En algunas realizaciones, se proporcionan métodos de distinción de sujetos con EHNA de sujetos con esteatosis. En algunas realizaciones, se proporcionan métodos de determinación de la gravedad de la EHNA.

En el presente documento también se describen, pero no se reivindican, métodos de determinación de si un sujeto tiene la enfermedad de hígado graso no alcohólico (EHGNA). Los métodos pueden comprender detectar el nivel de al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho, al menos nueve, al menos diez u once biomarcadores seleccionados de ACY, SHBG, CTSZ, MET, GSN, LGALS3BP, PLAT, CHL1, SERPINC1, SIGLEC7 y SIGLEC14, en una muestra del sujeto. Los métodos

pueden comprender detectar el nivel de al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete u ocho biomarcadores seleccionados de ACY, SHBG, CTSZ, MET, GSN, LGALS3BP, CHL1, SERPINC1. Los métodos pueden comprender detectar el nivel de ACY y, opcionalmente, uno o más de SHBG, CTSZ, MET, GSN, LGALS3BP, PLAT, CHL1, SERPINC1, SIGLEC7 y SIGLEC14, en una muestra del sujeto. Los métodos pueden comprender detectar el nivel de SHBG y, opcionalmente, uno o más de ACY, CTSZ, MET, GSN, LGALS3BP, PLAT, CHL1, SERPINC1, SIGLEC7 y SIGLEC14, en una muestra del sujeto. Los métodos pueden comprender detectar el nivel de ACY y SHBG, y, opcionalmente, uno o más de CTSZ, MET, GSN, LGALS3BP, PLAT, CHL1, SERPINC1, SIGLEC7 y SIGLEC14, en una muestra del sujeto. Un nivel de al menos un biomarcador seleccionado entre ACY, CTSZ, LGALS3BP, SIGLEC7, SIGLEC14 y PLAT que es superior a un nivel de control del respectivo biomarcador puede indicar que el sujeto tiene EHNA. Un nivel de al menos un biomarcador seleccionado entre SHBG, MET, GSN, CHL1 y SERPINC1 que es inferior a un nivel de control del respectivo biomarcador puede indicar que el sujeto tiene EHNA. El método puede comprender determinar si un sujeto tiene esteatosis. La esteatosis puede ser esteatosis leve, moderada o grave. El método puede comprender detectar al menos uno, al menos dos o tres biomarcadores seleccionados entre ACY, SHBG y SIGLEC14. El método puede comprender determinar si un sujeto tiene esteatohepatitis no alcohólica (EHNA). La EHNA puede ser EHNA de fase 1, 2, 3 o 4. El método puede comprender detectar al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis o siete biomarcadores seleccionados de ACY, SHBG, CTSZ, MET, GSN, LGALS3BP y SIGLEC7. El método puede comprender detectar al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho o nueve biomarcadores seleccionados de ACY, SHBG, CTSZ, MET, GSN, LGALS3BP, PLAT, CHL1 y SERPINC1.

En realizaciones reivindicadas de la invención, se proporcionan métodos de determinación de si un sujeto con esteatosis no alcohólica tiene esteatohepatitis no alcohólica (EHNA). El método comprende detectar el nivel de al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho o nueve biomarcadores seleccionados entre C7, PPID, IGFBP3, SIGLEC14, AIMP1, TOP1, COLEC11, CA6 y STX1A en una muestra del sujeto. El método comprende detectar el nivel de C7 y opcionalmente de uno o más de COLEC11, PPID, IGFBP3, SIGLEC14, AIMP1, TOP1, CA6 y STX1A en una muestra del sujeto. En algunas realizaciones, un método comprende detectar el nivel de C7, COLEC11, PPID y IGFBP3 y opcionalmente de uno o más de SIGLEC14, AIMP1, TOP1, CA6 y STX1A en una muestra del sujeto. El nivel de al menos un biomarcador seleccionado de C7, COLEC11 y SIGLEC14 que es superior al nivel de control del respectivo biomarcador, indica que el sujeto tiene EHNA, en donde el al menos un biomarcador es C7. En algunas realizaciones, un nivel de al menos un biomarcador seleccionado entre IGFBP3, AIMP1, TOP1, CA6, PPID y STX1A que es inferior al nivel de control del respectivo biomarcador, indica que el sujeto tiene EHNA. En algunas realizaciones, el sujeto tiene esteatosis leve, moderada o grave. En algunas realizaciones, la EHNA es EHNA de fase 1, 2, 3 o 4. En algunas realizaciones, la EHNA es EHNA de fase 2, 3 o 4. En algunas realizaciones, el método comprende detectar los niveles de C7 y de al menos uno, al menos dos o tres biomarcadores seleccionados entre COLEC11, PPID e IGFBP3. En algunas realizaciones, el método comprende detectar los niveles de C7 y de al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco o seis biomarcadores seleccionados entre SIGLEC14, AIMP1, TOP1, COLEC11, CA6 y STX1A.

En algunas realizaciones, se proporcionan métodos de control de un sujeto con esteatosis no alcohólica en cuanto al desarrollo de EHNA. En algunas realizaciones, un método comprende detectar el nivel de al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho o al menos nueve biomarcadores seleccionados entre C7, PPID, IGFBP3, SIGLEC14, AIMP1, TOP1, COLEC11, CA6 y STX1A en una muestra del sujeto. En algunas realizaciones, un nivel de al menos un biomarcador seleccionado entre C7, COLEC11 y SIGLEC14 que es superior al nivel de control del respectivo biomarcador indica que el sujeto tiene EHNA, en donde el al menos un biomarcador es C7. En algunas realizaciones, un nivel de al menos un biomarcador seleccionado entre IGFBP3, AIMP1, TOP1, CA6, PPID y STX1A que es inferior al nivel de control del respectivo biomarcador indica que el sujeto tiene EHNA. En algunas realizaciones, el sujeto tiene esteatosis leve, moderada o grave. En algunas realizaciones, la EHNA es EHNA de fase 1, 2, 3 o 4. En algunas realizaciones, la EHNA es EHNA de fase 2, 3 o 4. En algunas realizaciones, el método comprende detectar los niveles de al menos uno, al menos dos, al menos tres o cuatro biomarcadores seleccionados entre C7, COLEC11, PPID e IGFBP3. En algunas realizaciones, el método comprende detectar los niveles de C7 y de al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco o seis biomarcadores seleccionados entre SIGLEC14, AIMP1, TOP1, COLEC11, CA6 y STX1A.

En el presente documento también se describe, pero no se reivindican, métodos de determinación de si un sujeto tiene EHNA, en donde el método comprende detectar el nivel de al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho, al menos nueve o al menos diez biomarcadores seleccionados entre los biomarcadores de la Tabla 7. El método puede comprender formar un grupo de biomarcadores que tenga N biomarcadores proteicos de los biomarcadores proteicos enumerados en la Tabla 7 y detectar el nivel de cada uno de los N biomarcadores proteicos del grupo en una muestra del sujeto. N puede ser de 1 a 25, de 2 a 25, de 3 a 25, de 4 a 25 o de 5 a 25. N puede ser de 1 a 10, de 2 a 10, de 3 a 10, de 4 a 10 o de 5 a 10. Al menos uno de los N biomarcadores proteicos puede seleccionarse entre ACY y SHBG. Dos de los N biomarcadores proteicos son ACY y SHBG. Un nivel de un biomarcador de la Tabla 7 distinto de SHBG y GSN que es superior al nivel de control del respectivo biomarcador puede indicar que el sujeto tiene EHNA. Un nivel de al

menos un biomarcador seleccionado entre SHBG y GSN que es inferior al nivel de control del respectivo biomarcador puede indicar que el sujeto tiene EHNA. El método puede comprender determinar si un sujeto tiene EHNA.

- 5 En el presente documento también se describen, pero no se reivindican, métodos de determinación de si un sujeto tiene EHNA, en donde el método comprende detectar el nivel de al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete u ocho biomarcadores seleccionados entre los biomarcadores de la Tabla 9. El método puede comprender formar un grupo de biomarcadores que tenga N biomarcadores proteicos de los biomarcadores proteicos enumerados en la Tabla 9 y detectar el nivel de cada uno de los N biomarcadores proteicos del grupo en una muestra del sujeto. N puede ser de 1 a 8, de 2 a 8, de 3 a 8, de 4 a 8 o de 5 a 8. Al menos uno de los N biomarcadores proteicos puede ser COLEC11. Al menos uno de los N biomarcadores proteicos puede seleccionarse entre C7, COLEC11, PPID e IGFBP3. Un nivel de un biomarcador de la Tabla 9 distinto de IGFBP3 y GPC3 que es superior al nivel de control del respectivo biomarcador puede indicar que el sujeto tiene EHNA. Un nivel de al menos un biomarcador seleccionado entre IGFBP3 y GPC3 que es inferior al nivel de control del respectivo biomarcador puede indicar que el sujeto tiene EHNA.

En cualquiera de las realizaciones descritas en el presente documento, el sujeto puede estar en riesgo de desarrollar EHNA. En cualquiera de las realizaciones descritas en el presente documento, el sujeto puede tener una morbilidad asociada a EHNA seleccionada entre obesidad, obesidad abdominal, síndrome metabólico, enfermedad cardiovascular y diabetes. En cualquiera de las realizaciones descritas en el presente documento, el sujeto puede ser obeso.

En cualquiera de las realizaciones descritas en el presente documento, al menos un biomarcador puede ser un biomarcador proteico. En cualquiera de las realizaciones descritas en el presente documento, cada biomarcador puede ser un biomarcador proteico. En algunas realizaciones, un método comprende poner en contacto biomarcadores de la muestra del sujeto con un conjunto de reactivos de captura de biomarcadores, en el que cada reactivo de captura de biomarcador del conjunto de reactivos de captura de biomarcadores se une específicamente a un biomarcador diferente que se detecta. En algunas realizaciones, cada reactivo de captura de biomarcador es un anticuerpo o un aptámero. En algunas realizaciones, cada reactivo de captura de biomarcador es un aptámero. En algunas realizaciones, al menos un aptámero es un aptámero de velocidad de disociación lenta. En algunas realizaciones, al menos un aptámero de velocidad de disociación lenta comprende al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho, al menos nueve o al menos 10 nucleótidos con modificaciones. En algunas realizaciones, cada aptámero de velocidad de disociación lenta se une a su proteína diana con una velocidad de disociación ( $t^{1/2}$ ) de  $\geq 30$  minutos,  $\geq 60$  minutos,  $\geq 90$  minutos,  $\geq 120$  minutos,  $\geq 150$  minutos,  $\geq 180$  minutos,  $\geq 210$  minutos o  $\geq 240$  minutos.

En cualquiera de las realizaciones descritas en el presente documento, la muestra es una muestra de sangre. En cualquiera de las realizaciones descritas en el presente documento, la muestra puede seleccionarse entre una muestra de suero y una muestra de plasma.

En cualquiera de las realizaciones descritas en el presente documento, si el sujeto tiene EHNA o EHNA, se puede recomendar al sujeto una pauta seleccionada entre la pérdida de peso, el control del azúcar en sangre y el no consumo de alcohol. En cualquiera de las realizaciones descritas en el presente documento, si el sujeto tiene EHNA o EHNA, se puede recomendar al sujeto una cirugía de derivación gástrica. En cualquiera de las realizaciones descritas en el presente documento, si el sujeto tiene EHNA o EHNA, al sujeto se le puede recetar al menos un agente terapéutico seleccionado entre pioglitazona, vitamina E y metformina.

En algunas realizaciones, un método descrito en el presente documento tiene el fin de determinar una prima de seguro médico o una prima de seguro de vida. En algunas realizaciones, un método comprende además determinar una prima de seguro médico o una prima de seguro de vida. En algunas realizaciones, un método descrito en el presente documento comprende además usar la información procedente del método para predecir y/o gestionar la utilización de recursos médicos.

En el presente documento también se describen, pero no se reivindican, kits. Un kit puede comprender al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho, al menos nueve, al menos diez u once aptámeros que se unen específicamente a una proteína diana seleccionada entre ACY, SHBG, CTSZ, MET, GSN, LGALS3BP, PLAT, CHL1, SERPINC1, SIGLEC7 y SIGLEC14. Un kit puede comprender al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete u ocho aptámeros que se unen específicamente a una proteína diana seleccionada entre ACY, SHBG, CTSZ, MET, GSN, LGALS3BP, CHL1, SERPINC1. Un kit puede comprender un aptámero que se une específicamente a ACY y, opcionalmente, uno o más aptámeros que se unen específicamente a una o más proteínas diana seleccionadas entre SHBG, CTSZ, MET, GSN, LGALS3BP, PLAT, CHL1, SERPINC1, SIGLEC7 y SIGLEC14. Un kit puede comprender un aptámero que se une específicamente a SHBG y, opcionalmente, uno o más aptámeros que se unen específicamente a una o más proteínas diana seleccionadas entre ACY, CTSZ, MET, GSN, LGALS3BP, PLAT, CHL1, SERPINC1, SIGLEC7 y SIGLEC14. Un kit puede comprender un aptámero que se une específicamente a ACY y un aptámero que se une específicamente a SHBG y, opcionalmente, uno o más aptámeros

que se unen específicamente a una o más proteínas diana seleccionadas entre CTSZ, MET, GSN, LGALS3BP, PLAT, CHL1, SERPINC1, SIGLEC7 y SIGLEC14.

En el presente documento también se describen, pero no se reivindican, kits que comprenden al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho, al menos nueve o al menos diez aptámeros que se unen específicamente a una proteína diana seleccionada entre las proteínas de la Tabla 7. Un kit puede comprender N aptámeros, en donde cada aptámero se une específicamente a un biomarcador proteico enumerado en la Tabla 7. N puede ser de 1 a 25, de 2 a 25, de 3 a 25, de 4 a 25 o de 5 a 25. N puede ser de 1 a 10, de 2 a 10, de 3 a 10, de 4 a 10, de 5 a 10. Al menos uno de los N aptámeros puede unirse específicamente a un biomarcador proteico seleccionado entre ACY y SHBG. Uno de los N aptámeros puede unirse específicamente a ACY y uno de los N aptámeros puede unirse específicamente a SHBG. Cada aptámero puede unirse específicamente a una proteína diana diferente.

Un kit puede comprender al menos uno, al menos dos o tres aptámeros que se unen específicamente a una proteína diana seleccionada entre ACY, SHBG y SIGLEC14. Un kit puede comprender al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis o siete aptámeros que se unen específicamente a una proteína diana seleccionada entre ACY, SHBG, CTSZ, MET, GSN, LGALS3BP y SIGLEC7. Un kit puede comprender al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho o nueve aptámeros que se unen específicamente a una proteína diana seleccionada entre ACY, SHBG, CTSZ, MET, GSN, LGALS3BP, PLAT, CHL1 y SERPINC1.

Un kit puede comprender al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho, al menos nueve o al menos diez aptámeros que se unen específicamente a una proteína diana seleccionada entre C7, PPID, IGFBP3, SIGLEC14, AIMP1, TOP1, COLEC11, CA6, STX1A, GPC5, HIPK3, IGFBP7 e IL3RA. Un kit puede comprender al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho o nueve aptámeros que se unen específicamente a una proteína diana seleccionada entre C7, PPID, IGFBP3, SIGLEC14, AIMP1, TOP1, COLEC11, CA6 y STX1A. Un kit puede comprender al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete u ocho aptámeros que se unen específicamente a una proteína diana seleccionada entre C7, PPID, IGFBP3, COLEC11, GPC5, HIPK3, IGFBP7 e IL3RA. Cada aptámero puede unirse a una proteína diana diferente.

Un kit puede comprender al menos uno, al menos dos, al menos tres o cuatro aptámeros que se unen específicamente a una proteína diana seleccionada entre C7, COLEC11, PPID e IGFBP3. Un kit puede comprender al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco o seis aptámeros que se unen específicamente a una proteína diana seleccionada entre SIGLEC14, AIMP1, TOP1, COLEC11, CA6 y STX1A.

Un kit puede comprender al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete u ocho aptámeros que se unen específicamente a una proteína diana seleccionada entre las proteínas de la Tabla 9. Un kit puede comprender N aptámeros, en donde cada aptámero se une específicamente a un biomarcador proteico enumerado en la Tabla 9. N puede ser de 1 a 8, de 2 a 8, de 3 a 8, de 4 a 8 o de 5 a 8. Al menos uno de los N aptámeros puede unirse específicamente a COLEC11. Al menos uno de los N aptámeros puede unirse específicamente a un biomarcador proteico seleccionado entre C7, COLEC11, PPID e IGFBP3. Cuatro de los N aptámeros puede unirse específicamente a un biomarcador proteico diferente seleccionado entre C7, COLEC11, PPID e IGFBP3. Cada aptámero puede unirse específicamente a una proteína diana diferente.

En cualquiera de las realizaciones descritas en el presente documento, al menos un aptámero puede ser un aptámero de velocidad de disociación lenta. En cualquiera de las realizaciones descritas en el presente documento, cada aptámero puede ser un aptámero de velocidad de disociación lenta. En algunas realizaciones, al menos un aptámero de velocidad de disociación lenta comprende al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho, al menos nueve o al menos 10 nucleótidos con modificaciones hidrófobas. En algunas realizaciones, cada aptámero de velocidad de disociación lenta se une a su proteína diana con una velocidad de disociación ( $t^{1/2}$ ) de  $\geq 30$  minutos,  $\geq 60$  minutos,  $\geq 90$  minutos,  $\geq 120$  minutos,  $\geq 150$  minutos,  $\geq 180$  minutos,  $\geq 210$  minutos o  $\geq 240$  minutos.

En el presente documento también se describen, pero no se reivindican, composiciones. Una composición puede comprender proteínas de una muestra de un sujeto y al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho, al menos nueve, al menos diez u once aptámeros que se unen específicamente a una proteína diana seleccionada entre ACY, SHBG, CTSZ, MET, GSN, LGALS3BP, PLAT, CHL1, SERPINC1, SIGLEC7 y SIGLEC14. Una composición puede comprender proteínas de una muestra de un sujeto y al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete u ocho aptámeros que se unen específicamente a una proteína diana seleccionada entre ACY, SHBG, CTSZ, MET, GSN, LGALS3BP, CHL1, SERPINC1. Una composición puede comprender proteínas de una muestra de un sujeto y al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho, al menos nueve o al menos diez aptámeros que se unen específicamente a una proteína diana seleccionada entre las proteínas de la Tabla 7. Una composición puede comprender proteínas de una muestra de un sujeto y N aptámeros, en donde que cada aptámero se une específicamente a un biomarcador proteico enumerado

en la Tabla 7. N puede ser de 1 a 25, de 2 a 25, de 3 a 25, de 4 a 25 o de 5 a 25. N puede ser de 1 a 10, de 2 a 10, de 3 a 10, de 4 a 10 o de 5 a 10. Al menos uno de los N aptámeros puede unirse específicamente a un biomarcador proteico seleccionado entre ACY y SHBG. Uno de los N aptámeros puede unirse específicamente a ACY y uno de los N aptámeros puede unirse específicamente a SHBG. Cada aptámero puede unirse específicamente a una

proteína diana diferente.

Una composición puede comprender proteínas de una muestra de un sujeto y al menos uno, al menos dos o tres aptámeros que se unen específicamente a una proteína diana seleccionada entre ACY, SHBG y SIGLEC14. Una composición puede comprender proteínas de una muestra de un sujeto y al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis o siete aptámeros que se unen específicamente a una proteína diana seleccionada entre ACY, SHBG, CTSZ, MET, GSN, LGALS3BP y SIGLEC7. Una composición puede comprender proteínas de una muestra de un sujeto y al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho o nueve aptámeros que se unen específicamente a una proteína diana seleccionada entre ACY, SHBG, CTSZ, MET, GSN, LGALS3BP, PLAT, CHL1 y SERPINC1.

Una composición puede comprender proteínas de una muestra de un sujeto y al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho, al menos nueve o al menos diez aptámeros que se unen específicamente a una proteína diana seleccionada entre C7, PPID, IGFBP3, SIGLEC14, AIMP1, TOP1, COLEC11, CA6, STX1A, GPC5, HIPK3, IGFBP7 e IL3RA. Una composición puede comprender proteínas de una muestra de un sujeto y al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho o nueve aptámeros que se unen específicamente a una proteína diana seleccionada entre C7, PPID, IGFBP3, SIGLEC14, AIMP1, TOP1, COLEC11, CA6 y STX1A. Una composición puede comprender proteínas de una muestra de un sujeto y al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete u ocho aptámeros que se unen específicamente a una proteína diana seleccionada entre C7, PPID, IGFBP3, COLEC11, GPC5, HIPK3, IGFBP7 e IL3RA. Cada aptámero puede unirse a una proteína diana diferente.

Una composición puede comprender proteínas de una muestra de un sujeto y al menos uno, al menos dos, al menos tres o cuatro aptámeros que se unen específicamente a una proteína diana seleccionada entre C7, COLEC11, PPID e IGFBP3. Una composición puede comprender proteínas de una muestra de un sujeto y al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco o seis aptámeros que se unen específicamente a una proteína diana seleccionada entre SIGLEC14, AIMP1, TOP1, COLEC11, CA6 y STX1A.

Una composición puede comprender proteínas de una muestra de un sujeto y al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete u ocho aptámeros que se unen específicamente a una proteína diana seleccionada entre las proteínas de la Tabla 9. Una composición puede comprender proteínas de una muestra de un sujeto y N aptámeros, en donde cada aptámero se une específicamente a un biomarcador proteico enumerado en la Tabla 9. N puede ser de 1 a 8, de 2 a 8, de 3 a 8, de 4 a 8 o de 5 a 8. Al menos uno de los N aptámeros puede unirse específicamente a COLEC11. Al menos uno de los N aptámeros puede unirse específicamente a un biomarcador proteico seleccionado entre C7, COLEC11, PPID e IGFBP3. Cuatro de N aptámeros puede unirse específicamente a cada biomarcador proteico diferente seleccionado entre C7, COLEC11, PPID e IGFBP3. Cada aptámero puede unirse específicamente a una proteína diana diferente.

En cualquiera de las realizaciones descritas en el presente documento, la muestra es una muestra de sangre. En cualquiera de las realizaciones descritas en el presente documento, la muestra se puede seleccionar de una muestra de suero y una muestra de plasma.

En cualquiera de las realizaciones descritas en el presente documento, al menos un aptámero puede ser un aptámero de velocidad de disociación lenta. En cualquiera de las realizaciones descritas en el presente documento, cada aptámero puede ser un aptámero de velocidad de disociación lenta. En algunas realizaciones, al menos un aptámero de velocidad de disociación lenta comprende al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho, al menos nueve o al menos 10 nucleótidos con modificaciones hidrófobas. En algunas realizaciones, cada aptámero de velocidad de disociación lenta se une a su proteína diana con una velocidad de disociación ( $t^{1/2}$ ) de  $\geq 30$  minutos,  $\geq 60$  minutos,  $\geq 90$  minutos,  $\geq 120$  minutos,  $\geq 150$  minutos,  $\geq 180$  minutos,  $\geq 210$  minutos o  $\geq 240$  minutos.

En cualquiera de las realizaciones descritas en el presente documento, cada biomarcador puede ser un biomarcador proteico. En cualquiera de las realizaciones descritas en el presente documento, el método puede comprender poner en contacto biomarcadores de la muestra del sujeto con un conjunto de reactivos de captura de biomarcadores, en las que cada reactivo de captura de biomarcador del conjunto de reactivos de captura de biomarcadores se une específicamente a un biomarcador que se está detectando. En algunas realizaciones, cada reactivo de captura de biomarcador del conjunto de reactivos de captura de biomarcadores se une específicamente a un biomarcador diferente que se está detectando. En cualquiera de las realizaciones descritas en el presente documento, cada reactivo de captura de biomarcador puede ser un anticuerpo o un aptámero. En cualquiera de las realizaciones descritas en el presente documento, cada reactivo de captura de biomarcador puede ser un aptámero. En cualquiera de las realizaciones descritas en el presente documento, al menos un aptámero puede ser un aptámero de velocidad

de disociación lenta. En cualquiera de las realizaciones descritas en el presente documento, al menos un aptámero de velocidad de disociación lenta puede comprender al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho, al menos nueve o al menos 10 nucleótidos con modificaciones. En algunas realizaciones, las modificaciones son modificaciones hidrófobas. En algunas realizaciones, las modificaciones son modificaciones de bases hidrófobas. En algunas realizaciones, una o más de las modificaciones se pueden seleccionar de las modificaciones mostradas en la Figura 11. En algunas realizaciones, cada aptámero de velocidad de disociación lenta se une a su proteína diana con una velocidad de disociación ( $t^{1/2}$ ) de  $\geq 30$  minutos,  $\geq 60$  minutos,  $\geq 90$  minutos,  $\geq 120$  minutos,  $\geq 150$  minutos,  $\geq 180$  minutos,  $\geq 210$  minutos o  $\geq 240$  minutos.

En cualquiera de las realizaciones descritas en el presente documento, la muestra es una muestra de sangre. En algunas realizaciones, la muestra de sangre se selecciona entre una muestra de suero y una muestra de plasma.

En cualquiera de las composiciones descritas en el presente documento, la muestra en una composición puede ser una muestra de sangre. La muestra de sangre puede seleccionarse entre una muestra de suero y una muestra de plasma.

Un kit o una composición, descrito en el presente documento, pero no reivindicado, puede comprender al menos un aptámero que sea un aptámero de velocidad de disociación lenta. Cada aptámero de un kit o composición puede ser un aptámero de velocidad de disociación lenta. Al menos un aptámero de velocidad de disociación lenta puede comprender al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho, al menos nueve o al menos 10 nucleótidos con modificaciones. Al menos un nucleótido con una modificación puede ser un nucleótido con una modificación de base hidrófoba. Cada nucleótido con una modificación puede ser un nucleótido con una modificación de base hidrófoba. Cada modificación de base hidrófoba puede seleccionarse independientemente de la modificación de la Figura 11. Cada aptámero de velocidad de disociación lenta de un kit puede unirse a su proteína diana con una velocidad de disociación ( $t^{1/2}$ ) de  $\geq 30$  minutos,  $\geq 60$  minutos,  $\geq 90$  minutos,  $\geq 120$  minutos,  $\geq 150$  minutos,  $\geq 180$  minutos,  $\geq 210$  minutos o  $\geq 240$  minutos.

#### Breve descripción de las figuras

La FIG. 1 muestra trayectorias de selección de la estabilidad del clasificador de esteatosis, como se describe en el Ejemplo 2.

La FIG. 2 muestra una curva ROC de un clasificador de nueve biomarcadores para la esteatosis, como se describe en el Ejemplo 2.

La FIG. 3 muestra el voto por clase para el clasificador de bosque aleatorio de nueve marcadores para la esteatosis, como se describe en el Ejemplo 2.

La FIG. 4 muestra las funciones de distribución acumulativa para cada uno de los biomarcadores en el clasificador de nueve marcadores para la esteatosis, como se describe en el Ejemplo 2.

La FIG. 5 muestra trayectorias de selección de la estabilidad del clasificador de la EHNA (fibrosis), como se describe en el Ejemplo 2.

La FIG. 6 muestra una curva ROC de un clasificador de cuatro biomarcadores para la EHNA (fibrosis), como se describe en el Ejemplo 2.

La FIG. 7 muestra diagramas de cajas para el clasificador de cuatro marcadores para la EHNA (fibrosis) en cada uno de los grupos de sujetos, como se describe en el Ejemplo 2.

La FIG. 8 muestra las funciones de distribución acumulativa para cada uno de los biomarcadores en el clasificador de cuatro marcadores para la EHNA (fibrosis), como se describe en el Ejemplo 2.

La FIG. 9 ilustra un sistema informático ilustrativo no limitante para su uso con diversos métodos implementados por ordenador descritos en el presente documento.

La FIG. 10 ilustra un ensayo de aptámero ilustrativo no limitante que se puede usar para detectar uno o más biomarcadores en una muestra biológica.

La FIG. 11 muestra ciertas pirimidinas modificadas ilustrativas que pueden incorporarse a aptámeros, tales como aptámeros de velocidad de disociación lenta.

La FIG. 12 muestra diagramas de cajas para el clasificador de 8 marcadores para la esteatosis en cada uno de los grupos de sujetos, como se describe en el Ejemplo 5.

La FIG. 13 muestra diagramas de cajas para el clasificador de 8 marcadores para la fibrosis en cada uno de los grupos de sujetos, como se describe en el Ejemplo 5.

La FIG. 14 muestra las funciones de distribución acumulativa para cada uno de los biomarcadores en el clasificador de cuatro marcadores para la EHNA (fibrosis), como se describe en el Ejemplo 5.

La FIG. 15 muestra diagramas de cajas para el clasificador de 8 marcadores para la esteatosis para las muestras con ocultación después del desenmascaramiento, de acuerdo con su grupo de muestras real, como se describe en el Ejemplo 6.

La FIG. 16 muestra curvas ROC para el rendimiento del clasificador de la esteatosis de 8 marcadores para el conjunto del descubrimiento y para el conjunto de validación con ocultación, como se describe en el Ejemplo 6.

La FIG. 17 muestra diagramas de cajas para el clasificador de 8 marcadores para la fibrosis para las muestras con ocultación después del desenmascaramiento, de acuerdo con su grupo de muestras real, como se describe en el Ejemplo 6.

La FIG. 18 muestra curvas ROC para el rendimiento del clasificador de la fibrosis de 8 marcadores para el conjunto del descubrimiento y para los conjuntos de validación con ocultación, como se describe en el Ejemplo 6.

Las FIG. 19A y B muestran distribuciones de sensibilidad y especificidad de 2.500 iteraciones de arranque del conjunto de verificación del 20 % de retención, usando (A) el clasificador de la esteatosis de 8 marcadores y (B) el clasificador de la fibrosis de 8 marcadores, como se describe en el Ejemplo 6.

La FIG. 20 muestra diagramas de cajas para el clasificador de la esteatosis de 8 marcadores en cada una de las muestras pediátricas de acuerdo con su grupo de muestras real, como se describe en el Ejemplo 7.

## Descripción detallada

Si bien la invención se describirá junto con ciertas realizaciones representativas, se entenderá que la invención está definida por las reivindicaciones, y que no se limita a dichas realizaciones.

Un experto en la materia reconocerá que es posible usar muchos métodos y materiales similares o equivalentes a los descritos en el presente documento en la práctica de la presente invención. La presente invención no se limita de ningún modo a los métodos y materiales descritos.

A menos que se defina lo contrario, los términos técnicos y científicos usados en el presente documento tienen el significado comúnmente entendido por un experto habitual en la materia a la que pertenece la presente invención. Aunque, en la práctica de la invención, se puede usar cualquier método, dispositivo y material similar o equivalente a los descritos en el presente documento, en el presente documento, se describen ciertos métodos, dispositivos y materiales.

Como se usan en la presente solicitud, incluyendo las reivindicaciones adjuntas, las formas en singular "un", "el" y "ella" incluyen el plural, a menos que el contexto indique claramente lo contrario, y pueden usarse indistintamente con "al menos uno/a" y "uno/a o más". Por lo tanto, la referencia a "un aptámero" incluye mezclas de aptámeros, la referencia a "una sonda" incluye mezclas de sondas, y similares.

Como se usan en el presente documento, los términos y las expresiones "comprende", "que comprende", "incluye", "que incluye", "contiene", "que contiene" y cualquier variación de los mismos, están destinados a cubrir una inclusión no exclusiva, de modo que un proceso, método, producto por proceso o composición de materia que comprende, incluye o contiene un elemento o una lista de elementos puede incluir otros elementos no enumerados expresamente.

La presente solicitud incluye biomarcadores, métodos, dispositivos, reactivos, sistemas y kits para determinar si un sujeto tiene EHNA. La presente solicitud también incluye biomarcadores, métodos, dispositivos, reactivos, sistemas y kits para determinar si un sujeto tiene EHNA. En el presente documento también se describen biomarcadores, métodos, dispositivos, reactivos, sistemas y kits para determinar si un sujeto con EHNA tiene EHNA.

En el presente documento se describen, pero no se reivindican, uno o más biomarcadores para su uso bien solos o en diversas combinaciones para determinar si un sujeto tiene EHNA. Como se describe en detalle a continuación, los ejemplos incluyen los biomarcadores proporcionados en las Tablas 3, 4, 6, 7, 8 y 9, que se identificaron usando un ensayo multiplexado basado en aptámeros. La Tabla 3 enumera nueve biomarcadores que son útiles para distinguir muestras obtenidas de individuos obesos normales de muestras de individuos con EHNA. La Tabla 8 enumera ocho biomarcadores que son útiles para distinguir muestras obtenidas de individuos obesos normales de muestras de individuos con EHNA. La Tabla 4 enumera cuatro biomarcadores que son útiles para distinguir muestras obtenidas de individuos con esteatosis de muestras de individuos con EHNA en fase 2, 3 y 4. La Tabla 9 enumera ocho biomarcadores que son útiles para distinguir muestras obtenidas de individuos con esteatosis de muestras de individuos con EHNA en fase 2, 3 y 4. Las Tablas 6 y 7 enumeran biomarcadores adicionales que se pueden usar en cualquier combinación entre sí y/o con los biomarcadores de las Tablas 3, 4, 8 y/o 9. Un subconjunto de biomarcadores de las Tablas 3, 4, 6, 7, 8 y 9 puede combinarse en un grupo que se muestra en la Tabla 5.

En el presente documento se describen, pero no se reivindican, uno o más biomarcadores, para su uso solos o en diversas combinaciones para determinar si un sujeto tiene esteatosis. El sujeto puede ser obeso. Como se describe en detalle a continuación, los ejemplos incluyen los biomarcadores proporcionados en la Tabla 3 y/o la Tabla 8, que se identificaron usando un ensayo multiplexado basado en aptámeros. La Tabla 3 enumera cuatro biomarcadores que son útiles para distinguir muestras obtenidas de individuos obesos de muestras de individuos con esteatosis. La Tabla 8 enumera ocho biomarcadores que son útiles para distinguir muestras obtenidas de individuos obesos de muestras de individuos con esteatosis. Además, se pueden usar uno o más de los biomarcadores en la Tabla 3 y/o la Tabla 8 en combinación con uno o más biomarcadores de la Tabla 4 y/o Tabla 6 y/o Tabla 7 y/o Tabla 8 y/o Tabla 9, con o sin uno o más biomarcadores no enumerados en ninguna de las Tablas 3, 4, 6, 7, 8 o 9 en un método descrito en el presente documento.

En algunas realizaciones, se proporcionan uno o más biomarcadores para su uso solos o en diversas

combinaciones para determinar si un sujeto tiene EHNA de cualquier fase, de acuerdo con el método de las reivindicaciones anexas. En algunas realizaciones, se proporcionan uno o más biomarcadores para su uso solos o en diversas combinaciones para determinar si un sujeto tiene EHNA en fase 2, 3 o 4. En las realizaciones reivindicadas, ya se sabe que el sujeto tiene esteatosis. Como se describe en detalle a continuación, las realizaciones ilustrativas incluyen los biomarcadores proporcionados en la Tabla 4 y/o la Tabla 9, que se identificaron usando un ensayo multiplexado basado en aptámeros. La Tabla 4 enumera cuatro biomarcadores que son útiles para distinguir muestras obtenidas de individuos con esteatosis de muestras de individuos con EHNA. La Tabla 9 enumera ocho biomarcadores que son útiles para distinguir muestras obtenidas de individuos con esteatosis de muestras de individuos con EHNA. Además, se pueden usar uno o más de los biomarcadores de la Tabla 4 y/o la Tabla 9 en combinación con uno o más biomarcadores de la Tabla 3 y/o Tabla 6 y/o Tabla 7 y/o Tabla 8, con o sin uno o más biomarcadores no enumerados en ninguna de las Tablas 3, 4, 6, 7, 8 o 9 en un método descrito en el presente documento.

El número y la identidad de los biomarcadores de un grupo pueden seleccionarse basándose en la sensibilidad y especificidad para la combinación en particular de valores de biomarcadores. Los términos "sensibilidad" y "especificidad" se usan en el presente documento con respecto a la capacidad de clasificar correctamente a un individuo, basándose en uno o más niveles de biomarcadores detectados en una muestra biológica, como aquel que tiene la enfermedad o no la tiene. Los términos "sensibilidad" y "especificidad" se pueden usar en el presente documento con respecto a la capacidad de clasificar correctamente a un individuo, basándose en uno o más niveles de biomarcadores detectados en una muestra biológica, como aquel que tiene esteatosis o que no tiene esteatosis. En dichas realizaciones, "sensibilidad" indica el rendimiento del/de los biomarcador/es con respecto a la clasificación correcta de individuos con esteatosis. La "especificidad" indica el rendimiento del/de los biomarcador/es con respecto a la clasificación correcta de los individuos que no tienen esteatosis. Por ejemplo, un 85 % de especificidad y un 90 % de sensibilidad para un grupo de biomarcadores usado para analizar un conjunto de muestras de control (tales como muestras de individuos o sujetos sanos que no tienen esteatosis) y muestras de ensayo (tales como muestras de individuos con esteatosis) indica que el 85 % de las muestras de control fueron clasificadas correctamente como muestras de control por el grupo, y que el 90 % de las muestras de ensayo fueron clasificadas correctamente como muestras de ensayo por el grupo.

Los términos "sensibilidad" y "especificidad" pueden usarse en el presente documento con respecto a la capacidad de clasificar correctamente a un individuo, basándose en uno o más niveles de biomarcadores detectados en una muestra biológica, como aquel que tiene EHNA (o EHNA en fase 2, 3 o 4) o aquel que tiene esteatosis. "Sensibilidad" indica el rendimiento del/de los biomarcador/es con respecto a clasificar correctamente a los individuos con EHNA (o EHNA en fase 2, 3 o 4). La "especificidad" indica el rendimiento del/de los biomarcador/es con respecto a la clasificación correcta de los individuos que no tienen EHNA (o que no tienen EHNA en fase 2, 3 o 4). Por ejemplo, un 85 % de especificidad y un 90 % de sensibilidad para un grupo de biomarcadores usados para analizar un conjunto de muestras de control (tales como muestras de individuos con esteatosis) y muestras de ensayo (tales como muestras de individuos con EHNA o EHNA en fase 2, 3 o 4) indica que el 85 % de las muestras de control fueron clasificadas correctamente como muestras de control por el grupo, y que el 90 % de las muestras de ensayo fueron clasificadas correctamente como muestras de ensayo por el grupo.

En algunas realizaciones, el rendimiento global de un grupo de uno o más biomarcadores está representado por el valor de área bajo la curva (ABC). El valor de ABC se deriva de la curva de característica operativa del receptor (ROC), que se ilustra en el presente documento. La curva ROC es la gráfica de la tasa de positivos verdaderos (sensibilidad) de un ensayo frente a la tasa de falsos positivos (especificidad 1) del ensayo. La expresión "área bajo la curva" o "ABC" se refiere al área bajo la curva de una curva de característica operativa del receptor (ROC), siendo ambas bien conocidas en la técnica. Las medidas de ABC son útiles para comparar la exactitud de un clasificador en todo el intervalo de datos. Los clasificadores con un valor de ABC mayor tienen una mayor capacidad para clasificar correctamente las incógnitas entre dos grupos de interés (por ejemplo, los individuos normales y los individuos con EHNA, o los individuos con esteatosis y los individuos con EHNA). Las curvas ROC son útiles para representar el rendimiento de una determinada característica (por ejemplo, cualquiera de los biomarcadores descritos en la presente y/o cualquier elemento de información biomédica adicional) para distinguir entre dos poblaciones. Por lo general, los datos de características de toda la población se ordenan en orden ascendente según el valor de una sola característica. Luego, para cada valor de esa característica, se calculan las tasas de verdaderos positivos y de falsos positivos para los datos. La tasa de verdaderos positivos se determina contando el número de casos por encima del valor de esa característica y luego dividiendo entre el número total de casos. La tasa de falsos positivos se determina contando el número de controles por encima del valor de esa característica y luego dividiendo entre el número total de controles. Aunque esta definición se refiere a escenarios en los que una característica es elevada en los casos en comparación con los controles, esta definición también se aplica a escenarios en los que una característica es menor en los casos en comparación con los controles (en dicho escenario, se contarían las muestras por debajo del valor de esa característica). Las curvas ROC se pueden generar para una sola característica, así como para otros resultados individuales, por ejemplo, se puede combinar matemáticamente una combinación de dos o más características (por ejemplo, sumarse, restarse, multiplicarse, etc.) para proporcionar un solo valor de suma, y este único valor de suma se puede representar en una curva ROC. Además, en una curva ROC, se puede representar cualquier combinación de múltiples características, en la que la combinación deriva un solo resultado.

En el presente documento también se describen, pero no se reivindican, métodos que comprenden detectar el nivel de al menos un biomarcador enumerado en la Tabla 3, 4, 6, 7, 8 y/o 9 en una muestra de un sujeto para determinar si un sujeto tiene EHNA. En algunos de dichos métodos, el método comprende poner en contacto la muestra o una parte de la muestra del sujeto con al menos un reactivo de captura, en las que cada reactivo de captura se une específicamente a un biomarcador cuyos niveles se están detectando. El método puede comprender poner en contacto la muestra, o proteínas de la muestra, con al menos un aptámero, en donde cada aptámero se une específicamente a un biomarcador cuyos niveles se están detectando.

En el presente documento también se describe, pero no se reivindica, un método que comprende detectar el nivel de al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho, al menos nueve, al menos diez u once biomarcadores seleccionados entre ACY, SHBG, CTSZ, MET, GSN, LGALS3BP, PLAT, CHL1, SERPINC1, SIGLEC7 y SIGLEC14 en una muestra de un sujeto, en las que un nivel de al menos un biomarcador seleccionado entre ACY, CTSZ, LGALS3BP, SIGLEC7, SIGLEC14 y PLAT que es superior al nivel de control del respectivo biomarcador y/o un nivel de al menos un biomarcador seleccionado entre SHBG, MET, GSN, CHL1 y SERPINC1 que es inferior al nivel de control del respectivo biomarcador indica que el sujeto tiene EHNA. Un método puede comprender detectar el nivel de al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete u ocho biomarcadores seleccionados entre ACY, SHBG, CTSZ, MET, GSN, LGALS3BP, CHL1, SERPINC1 en una muestra de un sujeto, en las que un nivel de al menos un biomarcador seleccionado entre ACY, CTSZ y LGALS3BP que es superior al nivel de control del respectivo biomarcador y/o un nivel de al menos un biomarcador seleccionado entre SHBG, MET, GSN, CHL1 y SERPINC1 que es inferior al nivel de control del respectivo biomarcador indica que el sujeto tiene EHNA. El sujeto puede ser un sujeto obeso. El método puede comprender determinar si el sujeto tiene esteatosis y/o determinar si la esteatosis es leve, moderada o grave. En algunos de dichos métodos, el método comprende detectar el nivel de al menos uno, al menos dos o tres biomarcadores seleccionados entre ACY, SHBG y SIGLEC14. El método puede comprender determinar si el sujeto tiene EHNA, tal como EHNA en fase 1, 2, 3 o 4. En algunas de dichas realizaciones, el método comprende detectar el nivel de al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis o siete biomarcadores seleccionados entre ACY, SHBG, CTSZ, MET, GSN, LGALS3BP y SIGLEC7. Un método puede comprender determinar si un sujeto tiene EHNA, que puede ser bien esteatosis o EHNA. En algunos de dichos métodos, el método comprende detectar los niveles de al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho o nueve biomarcadores seleccionados entre ACY, SHBG, CTSZ, MET, GSN, LGALS3BP, PLAT, CHL1 y SERPINC1.

En el presente documento también se describe, pero no se reivindica, un método que comprende detectar el nivel de al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho, al menos nueve o al menos diez biomarcadores seleccionados entre los biomarcadores de la Tabla 7. Un nivel de un biomarcador de la Tabla 7 distinto de SHBG y GSN que es superior al nivel de control del respectivo biomarcador y/o un nivel de al menos un biomarcador seleccionado entre SHBG y GSN que es inferior al nivel de control del respectivo biomarcador puede indicar que el sujeto tiene EHNA.

Los biomarcadores identificados en el presente documento proporcionan una serie de opciones para los subconjuntos o grupos de biomarcadores que se pueden usar para identificar de manera eficaz la EHNA. La selección del número apropiado de dichos biomarcadores puede depender de la combinación específica de los biomarcadores seleccionados. Además, en cualquiera de los métodos descritos en el presente documento, excepto cuando se indique explícitamente, un grupo de biomarcadores puede comprender biomarcadores adicionales que no se muestran en las Tablas 3, 4, 6, 7, 8 o 9.

En algunas realizaciones descritas, pero no reivindicadas, en el presente documento, un método comprende detectar el nivel de al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho, al menos nueve o al menos diez biomarcadores seleccionados entre C7, PPID, IGFBP3, SIGLEC14, AIMP1, TOP1, COLEC11, CA6, STX1A, GPC5, HIPK3, IGFBP7 e IL3RA en una muestra del sujeto, en las que un nivel de al menos un biomarcador seleccionado entre C7, COLEC11, HIPK3, IGFBP7, IL3RA y SIGLEC14 que es superior al nivel de control del respectivo biomarcador y/o un nivel de al menos un biomarcador seleccionado entre IGFBP3, AIMP1, TOP1, CA6, PPID, GPC3, IGFBP3 y STX1A que es inferior al nivel de control del respectivo biomarcador indica que el sujeto tiene EHNA. En realizaciones reivindicadas, un método comprende detectar el nivel de al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho o nueve biomarcadores seleccionados entre C7, PPID, IGFBP3, SIGLEC14, AIMP1, TOP1, COLEC11, CA6 y STX1A en una muestra de un sujeto, en las que un nivel de al menos un biomarcador seleccionado entre C7, COLEC11 y SIGLEC14 que es superior al nivel de control del respectivo biomarcador y/o un nivel de al menos un biomarcador seleccionado entre IGFBP3, AIMP1, TOP1, CA6, PPID y STX1A que es inferior al nivel de control del respectivo biomarcador indica que el sujeto tiene EHNA, en donde al menos un biomarcador es C7. En algunas otras realizaciones desveladas, pero no reivindicadas, en el presente documento, un método comprende detectar el nivel de al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete u ocho biomarcadores seleccionados entre C7, PPID, IGFBP3, COLEC11, GPC5, HIPK3, IGFBP7 e IL3RA en una muestra del sujeto, en las que un nivel de al menos un biomarcador seleccionado entre C7, COLEC11, HIPK3, IGFBP7 e IL3RA que es superior al nivel de control del respectivo biomarcador indica que el sujeto tiene EHNA y/o

un nivel de al menos un biomarcador seleccionado entre IGFBP3, PPID y GPC3 que es inferior al nivel de control del respectivo biomarcador indica que el sujeto tiene EHNA. En algunas realizaciones, el sujeto es obeso. En las realizaciones reivindicadas, ya se ha determinado que el sujeto tiene esteatosis, que puede ser esteatosis leve, moderada o grave. En algunas realizaciones, la EHNA es EHNA de fase 1, 2, 3 o 4. En algunas realizaciones, la EHNA es EHNA de fase 2, 3 o 4. En algunas de dichas realizaciones, el método comprende detectar los niveles de C7 y de al menos uno, al menos dos o al menos tres biomarcadores seleccionados entre COLEC11, PPID e IGFBP3.

En algunas realizaciones, un método comprende detectar el nivel de al menos un biomarcador enumerado en la Tabla 3, 4, 6, 7, 8 y/o 9 en una muestra de un sujeto para determinar si un sujeto tiene EHNA o EHNA en fase 2, 3 o 4. En algunas de dichas realizaciones, el método comprende poner en contacto la muestra o una parte de la muestra del sujeto con al menos un reactivo de captura, en las que cada reactivo de captura se une específicamente a un biomarcador cuyos niveles se están detectando. En algunas realizaciones, el método comprende poner en contacto la muestra, o proteínas de la muestra, con al menos un aptámero, en las que cada aptámero se une específicamente a un biomarcador cuyos niveles se están detectando.

Los biomarcadores identificados en el presente documento proporcionan una serie de opciones para los subconjuntos o grupos de biomarcadores que se pueden usar para identificar de manera eficaz la EHNA o EHNA en fase 2, 3 o 4. La selección del número apropiado de dichos biomarcadores puede depender de la combinación específica de los biomarcadores seleccionados. Además, en cualquiera de los métodos descritos en el presente documento, excepto cuando se indique explícitamente, un grupo de biomarcadores puede comprender biomarcadores adicionales que no se muestran en las Tablas 3, 4, 6, 7, 8 o 9.

Como se usa en el presente documento, "enfermedad de hígado graso no alcohólico" o "EHGNA" se refiere a una afección en la que se deposita grasa en el hígado (esteatosis hepática), con o sin inflamación y fibrosis, en ausencia de un consumo excesivo de alcohol.

Como se usa en el presente documento, "esteatosis" y "esteatosis no alcohólica" se usan indistintamente, e incluyen esteatosis leve, moderada y grave, sin inflamación ni fibrosis, en ausencia de un consumo excesivo de alcohol. La Tabla 1 muestra una clasificación ilustrativa de esteatosis leve, moderada y grave.

Como se usa en el presente documento, "esteatohepatitis no alcohólica" o "EHNA" se refiere a EHGNA en la que hay inflamación y/o fibrosis en el hígado. La EHNA se puede dividir en cuatro fases. Los métodos ilustrativos para determinar la fase de la EHNA se describen, por ejemplo, en Kleiner *et al.*, 2005, *Hepatology*, 41(6):1313-1321 y Brunt *et al.*, 2007, *Modern Pathol.*, 20: S40-S48. La Tabla 1 muestra la clasificación ilustrativa de la EHNA de fase 1, 2, 3 y 4.

Como se usa en el presente documento, "obeso" con referencia a un sujeto se refiere a un sujeto con un IMC de 30 o superior.

"Muestra biológica", "muestra" y "muestra de ensayo" se usan indistintamente en el presente documento para referirse a cualquier material, fluido biológico, tejido o célula, que se obtiene o deriva de otro modo de un individuo. Esto incluye sangre (incluyendo sangre entera, leucocitos, células mononucleares de sangre periférica, capa leucocítica, plasma y suero), esputo, lágrimas, mucus, lavados nasales, aspirado nasal, orina, saliva, lavados peritoneales, ascitos, líquido quístico, líquido glandular, líquido linfático, aspirado bronquial, líquido sinovial, aspirado de las articulaciones, secreciones de órganos, células, un extracto celular y líquido cefalorraquídeo. Esto también incluye fracciones separadas experimentalmente de todo lo anterior. Por ejemplo, una muestra de sangre puede fraccionarse en suero, plasma o en fracciones que contengan determinados tipos de células sanguíneas, tales como glóbulos rojos o glóbulos blancos (leucocitos). En algunas realizaciones, una muestra puede ser una combinación de muestras de un individuo, tal como una combinación de una muestra tisular y líquida. La expresión "muestra biológica" también incluye materiales que contienen material sólido homogenizado, tal como de una muestra de heces, una muestra tisular o una biopsia tisular, por ejemplo. La expresión "muestra biológica" también incluye materiales procedentes de un cultivo tisular o celular. Puede emplearse cualquier método adecuado para obtener una muestra biológica; incluyendo los métodos ilustrativos, por ejemplo, flebotomía, hisopo (por ejemplo, hisopo bucal), y un procedimiento de biopsia por aspirado de aguja fina. Los tejidos ilustrativos susceptibles de ser aspirados con aguja fina incluyen ganglio linfático, pulmón, tiroides, mama, páncreas e hígado. Las muestras también pueden recogerse, por ejemplo, por microdissección (por ejemplo, microdissección por captura de láser (LCM) o microdissección por láser (LMD)), lavado de vejiga, frotis (por ejemplo, un frotis de PAP) o lavado ductal. Una "muestra biológica" obtenida o procedente de un individuo incluye cualquier muestra que se haya procesado de cualquier manera adecuada tras su obtención del individuo.

Además, en algunas realizaciones, una muestra biológica puede derivarse tomando muestras biológicas de un número de individuos y combinándolas, o agrupando una alícuota de la muestra biológica de cada individuo. La muestra combinada se puede tratar como se describe en el presente documento para una muestra de un solo individuo y, por ejemplo, si se establece un mal pronóstico en la muestra agrupada, entonces cada muestra biológica individual puede volver a analizarse para determinar qué individuo/s tienen esteatosis y/o EHNA.

"Diana", "molécula diana" y "analito" se usan indistintamente en el presente documento para hacer referencia a cualquier molécula de interés que pueda estar presente en una muestra biológica. Una "molécula de interés" incluye cualquier variación menor de una determinada molécula, tal como, en el caso de una proteína, por ejemplo variaciones menores en secuencia de aminoácidos, formación de enlaces disulfuro, glicosilación, lipidación, acetilación, fosforilación o cualquier otra manipulación o modificación, tal como conjugación con un componente de marcaje, que no altere sustancialmente la identidad de la molécula. Una "molécula diana", "diana" o "analito" se refieren a un conjunto de copias de un tipo o especie de molécula o estructura multimolecular. Las "moléculas diana", las "dianas" y los "analitos" se refieren a más de un tipo o especie de molécula o estructura multimolecular. Las moléculas diana ilustrativas incluyen proteínas, polipéptidos, ácidos nucleicos, hidratos de carbono, lípidos, polisacáridos, glicoproteínas, hormonas, receptores, antígenos, anticuerpos, aficuerpos, miméticos de anticuerpos, virus, patógenos, sustancias tóxicas, sustratos, metabolitos, análogos de estado de transición, cofactores, inhibidores, fármacos, colorantes, nutrientes, factores de crecimiento, células, tejidos y cualquier fragmento o parte de cualquiera de los anteriores. En algunas realizaciones, una molécula diana es una proteína, en cuyo caso, la molécula diana puede denominarse una "proteína diana".

Como se usa en el presente documento, un "agente de captura" o "reactivo de captura" se refiere a una molécula que es capaz de unirse específicamente a un biomarcador. Un "reactivo de captura de proteína diana" se refiere a una molécula que es capaz de unirse específicamente a una proteína diana. Los reactivos de captura ilustrativos no limitantes incluyen aptámeros, anticuerpos, adnectinas, anquirinas, otros miméticos de anticuerpos y otros armazones proteicos, autoanticuerpos, quimeras, moléculas pequeñas, ácidos nucleicos, lectinas, receptores de unión al ligando, polímeros impresos, avímeros, peptidomiméticos, receptores de hormonas, receptores de citocinas, receptores sintéticos, y modificaciones y fragmentos de cualquiera de los reactivos de captura anteriormente mencionados. En algunas realizaciones, un reactivo de captura se selecciona entre un aptámero y un anticuerpo.

El término "anticuerpo" se refiere a anticuerpos de longitud completa de cualquier especie, y fragmentos y derivados de dichos anticuerpos, incluyendo fragmentos Fab, fragmentos F(ab')<sub>2</sub>, anticuerpos monocatenarios, fragmentos Fv y fragmentos Fv monocatenarios. El término "anticuerpo" también se refiere a anticuerpos derivados sintéticamente, tales como anticuerpos y fragmentos derivados de presentación en fagos, aficuerpos, nanoanticuerpos, etc.

Como se usa en el presente documento, "marcador" y "biomarcador" se usan indistintamente para hacer referencia a una molécula diana que indica o que es una señal de un proceso normal o anómalo en un individuo o de una enfermedad u otra afección en un individuo. Más concretamente, un "marcador" o "biomarcador" es un parámetro anatómico, fisiológico, bioquímico o molecular asociado con la presencia de un estado o proceso fisiológico específico, bien normal o anómalo, y, si es anómalo, bien crónico o agudo. Los biomarcadores son detectables y medibles mediante diversos métodos incluyendo ensayos de laboratorio y generación de imágenes médicas. En algunas realizaciones, un biomarcador es una proteína diana.

Como se usa en el presente documento, "nivel de biomarcador" y "nivel" se refieren a una medición que se realiza usando cualquier método analítico para detectar el biomarcador en una muestra biológica y que indica la presencia, ausencia, cantidad absoluta o concentración, cantidad relativa o concentración, título, un nivel, un nivel de expresión, una relación de niveles medidos, o similares, de, para o correspondientes al biomarcador en la muestra biológica. La naturaleza exacta del "valor" o "nivel" depende del diseño específico y de los componentes del método analítico empleado en particular para detectar el biomarcador.

Un "nivel de control" de una molécula diana se refiere al nivel de la molécula diana en el mismo tipo de muestra de un individuo que no tiene la enfermedad o afección, o de un individuo que no se sospecha que tenga la enfermedad o afección. No es necesario determinar un "nivel de control" de una molécula diana cada vez que se lleven a cabo los métodos actuales, y puede ser un nivel previamente determinado que se use como referencia o umbral para determinar si el nivel de una muestra en particular es superior o inferior a un nivel normal. En algunas realizaciones, un nivel de control en un método descrito en el presente documento es el nivel que se ha observado en uno o más sujetos sin EHNA. En algunas realizaciones, un nivel de control en un método descrito en el presente documento es el nivel que se ha observado en uno o más sujetos con EHNA, pero no con EHNA. En algunas realizaciones, un nivel de control en un método descrito en el presente documento es el nivel promedio o medio, opcionalmente más o menos una variación estadística, que se ha observado en una pluralidad de sujetos normales, o sujetos con EHNA, pero no con EHNA.

Como se usa en el presente documento, "individuo" y "sujeto" se usan indistintamente para referirse a un sujeto de ensayo o paciente. El individuo puede ser un mamífero o no mamífero. En diversas realizaciones, el individuo es un mamífero. Un individuo mamífero puede ser humano o no humano. En diversas realizaciones, el individuo es un ser humano. Un individuo sano o normal es un individuo en el que la enfermedad o afección de interés (tal como EHNA) no es detectable mediante métodos de diagnóstico convencionales.

"Diagnosticar", "diagnosticando", "diagnóstico" y sus variaciones se refieren a la detección, determinación o reconocimiento de un estado o condición de salud de un individuo basándose en uno o más signos, síntomas, datos u otra información pertinente a ese individuo. El estado de salud de un individuo puede diagnosticarse como

sano/normal (es decir, un diagnóstico de la ausencia de una enfermedad o afección) o diagnosticarse como enfermo/anómalo (es decir, un diagnóstico de la presencia, o una evaluación de las características, de una enfermedad o afección). Los términos "diagnosticar", "diagnosticando", "diagnóstico", etc., engloban, con respecto a una determinada enfermedad o afección, la detección inicial de la enfermedad; la caracterización o clasificación de la enfermedad; la detección de la progresión, remisión o recurrencia de la enfermedad; y la detección de la respuesta de la enfermedad tras la administración de un tratamiento o terapia al individuo. El diagnóstico de EHGNA incluye distinguir a los individuos que tienen EHGNA de los individuos que no la tienen. El diagnóstico de EHNA incluye distinguir a los individuos que tienen EHNA de los individuos que tienen esteatosis en el hígado, pero no EHNA, y de los individuos sin enfermedad hepática.

"Pronosticar", "pronosticando", "pronóstico" y las variaciones de los mismos se refieren a la predicción de un curso futuro de una enfermedad o afección en un individuo que tiene la enfermedad o afección (por ejemplo, predecir la supervivencia del paciente), y dichos términos engloban la evaluación de la respuesta de la enfermedad tras la administración de un tratamiento o terapia al individuo.

"Evaluar", "evaluando", "evaluación" y sus variaciones engloban tanto "diagnosticar" como "pronosticar", y también engloban determinaciones o predicciones sobre el curso futuro de una enfermedad o afección en un individuo que no tiene la enfermedad, así como las determinaciones o predicciones sobre la probabilidad de que una enfermedad o afección recurra en un individuo quien aparentemente ha sido curado de la enfermedad. El término "evaluar" también engloba la evaluación de la respuesta de un individuo a una terapia, tal como, por ejemplo, predecir si es probable que un individuo responda favorablemente a un agente terapéutico o si es poco probable que responda a un agente terapéutico (o experimentará toxicidad u otros efectos secundarios no deseados, por ejemplo), seleccionando un agente terapéutico para la administración a un individuo, o controlar o determinar la respuesta de un individuo a una terapia que se ha administrado al individuo. Por lo tanto, "evaluar" la EHGNA puede incluir, por ejemplo, cualquiera de los siguientes: pronosticar el curso futuro de la EHGNA en un individuo; predecir si la EHGNA progresará a EHNA; predecir si una determinada etapa de la EHNA progresará a una etapa más alta de EHNA; etc.

Como se usa en el presente documento, "detectar" o "determinar" con respecto a un nivel de biomarcador incluye el uso tanto del instrumento usado para observar y registrar una señal correspondiente a un nivel de biomarcador y el/los material/es requerido/s para generar esa señal. En diversas realizaciones, el nivel se detecta usando cualquier método adecuado, incluyendo fluorescencia, quimioluminiscencia, resonancia de plasmón superficial, ondas acústicas superficiales, espectrometría de masas, espectroscopía de infrarrojos, espectroscopía Raman, microscopía de fuerza atómica, microscopía de túnel de barrido, métodos de detección electroquímica, resonancia magnética nuclear, puntos cuánticos y similares.

Como se usa en el presente documento, un "sujeto con EHGNA" se refiere a un sujeto que ha sido diagnosticado de EHGNA. En algunas realizaciones, se sospecha de EHGNA durante un chequeo de rutina, el control del síndrome metabólico y la obesidad, o el control de los posibles efectos secundarios de los fármacos (por ejemplo, agentes reductores del colesterol o esteroides). En algún caso, las enzimas hepáticas tales como AST y ALT son altas. En algunas realizaciones, un sujeto es diagnosticado tras la generación de imágenes abdominales o torácicas, ultrasonidos hepáticos o imágenes de resonancia magnética. En algunas realizaciones, se han descartado otras afecciones tales como el consumo excesivo de alcohol, la hepatitis C y la enfermedad de Wilson antes de un diagnóstico de EHGNA. En algunas realizaciones, un sujeto ha sido diagnosticado tras una biopsia hepática.

Como se usa en el presente documento, un "sujeto con esteatosis" y un "sujeto con esteatosis no alcohólica" se usan indistintamente, y se refieren a un sujeto que ha sido diagnosticado de esteatosis. En algunas realizaciones, la esteatosis se diagnostica mediante un método descrito anteriormente para la EHGNA en general.

Como se usa en el presente documento, un "sujeto con EHNA" se refiere a un sujeto que ha sido diagnosticado de EHNA. En algunas realizaciones, la EHNA se diagnostica mediante un método descrito anteriormente para la EHGNA en general. En algunas realizaciones, se diagnostica fibrosis avanzada en un paciente con EHGNA, por ejemplo, de acuerdo con Gambino R, *et. al. Annals of Medicine* 2011; 43(8):617-49.

Como se usa en el presente documento, un "sujeto con riesgo de desarrollar EHGNA" se refiere a un sujeto con una o más morbilidades asociadas a la EHGNA, tales como obesidad, obesidad abdominal, síndrome metabólico, enfermedad cardiovascular y diabetes.

Como se usa en el presente documento, un "sujeto en riesgo de desarrollar esteatosis" se refiere a un sujeto que no ha sido diagnosticado de esteatosis, pero que tiene una o más morbilidades asociadas a la EHGNA, tales como obesidad, obesidad abdominal, síndrome metabólico, enfermedad cardiovascular y diabetes.

Como se usa en el presente documento, un "sujeto con riesgo de desarrollar EHNA" se refiere a un sujeto con esteatosis que sigue teniendo una o más morbilidades asociadas a la EHGNA, tales como obesidad, obesidad abdominal, síndrome metabólico, enfermedad cardiovascular y diabetes.

"Soporte sólido" se refiere en el presente documento a cualquier sustrato que tenga una superficie a la que las

moléculas puedan unirse, directa o indirectamente, a través de enlaces covalentes o no covalentes. Un "soporte sólido" puede tener varios formatos físicos, que pueden incluir, por ejemplo, una membrana; un chip (por ejemplo, un chip de proteína); un portaobjetos (por ejemplo, un portaobjetos de vidrio o un cubreobjetos); una columna; una partícula hueca, sólida, semisólida, que contiene poros o cavidades, tal como, por ejemplo, una perla; un gel; una fibra, incluyendo un material de fibra óptica; una matriz; y un receptáculo de muestras. Los receptáculos ilustrativos de muestras incluyen pocillos de muestras, tubos, capilares, viales y cualquier otro recipiente, ranura o hendidura capaz de contener una muestra. Un receptáculo de muestras puede estar contenido en una plataforma de múltiples muestras, tal como una placa de microtitulación, un portaobjetos, dispositivo microfluidico y similares. Un soporte puede componerse de un material natural o sintético, un material orgánico o inorgánico. La composición del soporte sólido sobre el que se unen los reactivos de captura en general depende del método de unión (por ejemplo, unión covalente). Otros receptáculos ilustrativos incluyen microgotas y emulsiones de aceite/acuosas controladas en microfluidos o a granel dentro de las que pueden producirse ensayos y manipulaciones relacionadas. Los soportes sólidos adecuados incluyen, por ejemplo, plásticos, resinas, polisacáridos, materiales basados en sílice o sílice, vidrio funcionalizado, silicio modificado, carbono, metales, vidrios inorgánicos, membranas, nylon, fibras naturales (tales como, por ejemplo, seda, lana y algodón), polímeros y similares. El material que compone el soporte sólido puede incluir grupos reactivos tales como, por ejemplo, grupos carboxi, amino o hidroxilo, que se usan para la unión de los reactivos de captura. Los soportes sólidos poliméricos pueden incluir, por ejemplo, poliestireno, tetraftalato de polietilenglicol, acetato de polivinilo, cloruro de polivinilo, polivinilpirrolidona, poliácridonitrilo, polimetilmetacrilato, politetrafluoroetileno, caucho de butilo, caucho de estirenobutadieno, caucho natural, polietileno, polipropileno, (poli)tetrafluoroetileno, (poli)vinilidenofluoruro, policarbonato y polimetilpenteno. Las partículas de soporte sólido adecuadas que pueden usarse incluyen, por ejemplo, partículas codificadas, tales como partículas codificadas de tipo Luminex®, partículas magnéticas y partículas de vidrio.

#### Usos ilustrativos de los biomarcadores

En diversos ejemplos descritos, pero no reivindicados, en el presente documento, se proporcionan métodos para determinar si un sujeto tiene EHGNA. En el presente documento también se describen, pero no se reivindican, métodos para determinar si un sujeto tiene esteatosis, que puede ser una esteatosis leve, moderada o grave. En diversas realizaciones, se proporcionan métodos para determinar si un sujeto tiene EHNA, que puede ser EHNA en fase 1, 2, 3 o 4, o que puede ser EHNA en fase 2, 3 o 4. En las realizaciones reivindicadas, se proporcionan métodos para determinar si un sujeto con esteatosis tiene EHNA, que puede ser EHNA en fase 1, 2, 3 o 4, o que puede ser EHNA en fase 2, 3 o 4. Los métodos comprenden detectar uno o más niveles de biomarcadores correspondientes a uno o más biomarcadores que están presentes en la circulación de un individuo, tal como en suero o plasma, mediante cualquier número de métodos analíticos, incluyendo cualquiera de los métodos analíticos descritos en el presente documento. Estos biomarcadores están, por ejemplo, presentes a diferentes niveles en individuos con EHGNA en comparación con individuos normales (pudiendo ser los individuos normales individuos obesos). En algunas realizaciones, los biomarcadores están presentes a diferentes niveles en individuos con EHNA (tales como EHNA en fase 1, 2, 3 o 4, o EHNA en fase 2, 3 o 4) en comparación con los individuos normales (pudiendo ser los individuos normales individuos obesos). En algunas realizaciones, los biomarcadores están presentes en diferentes niveles en individuos con EHNA (tales como EHNA en fase 1, 2, 3 o 4, o EHNA en fase 2, 3 o 4) en comparación con los sujetos con esteatosis, que puede ser esteatosis leve, moderada o grave.

La detección de los niveles diferenciales de un biomarcador en un individuo puede usarse, por ejemplo, para permitir la determinación de si un individuo tiene EHGNA (que puede ser esteatosis o EHNA), o si un individuo con esteatosis ha desarrollado EHNA. Cualquiera de los biomarcadores descritos en el presente documento puede usarse para controlar individuos (tales como individuos obesos) en cuanto al desarrollo de EHGNA, o para controlar individuos con esteatosis en cuanto al desarrollo de EHNA.

Como ejemplo de la manera en que cualquiera de los biomarcadores descritos en el presente documento puede usarse para determinar si un sujeto tiene EHGNA, los niveles de uno o más de los biomarcadores descritos en un individuo que no ha sido diagnosticado de EHGNA, pero que tiene una o más morbilidades asociadas con la EHGNA pueden indicar que el individuo ha desarrollado EHGNA en una etapa más precoz que la que se determinaría usando una prueba invasiva tal como la biopsia hepática. Debido a que los métodos actuales no son invasivos, se pueden usar para controlar individuos en riesgo de desarrollar EHGNA (tales como, por ejemplo, individuos obesos). Al detectar la EHGNA en una etapa más precoz, la intervención médica puede ser más eficaz. Dicha intervención médica puede incluir, pero sin limitación, pérdida de peso, control del azúcar en sangre y abstinencia de alcohol. En algunas realizaciones, pueden usarse agentes terapéuticos, tales como pioglitazona, vitamina E y/o metformina. Véase, por ejemplo, Sanyal *et al.*, 2010, *NEJM*, 362: 1675-1685. En algunos casos, dicha intervención temprana puede retrasar o prevenir la insuficiencia hepática y la necesidad de un trasplante de hígado.

De manera similar, como un ejemplo adicional de la manera en que los biomarcadores descritos en el presente documento pueden usarse para determinar si un sujeto que tiene esteatosis está desarrollando EHNA, los niveles de uno o más de los biomarcadores descritos en un individuo con esteatosis pueden indicar que el individuo está desarrollando EHNA. Debido a que los métodos actuales no son invasivos, las personas con esteatosis pueden ser controladas en cuanto al desarrollo de EHNA. Al detectar la EHNA en una etapa anterior, la intervención médica puede ser más eficaz. Dicha intervención médica puede incluir, pero sin limitación, pérdida de peso, control del

azúcar en sangre y abstinencia de alcohol. En algunas realizaciones, pueden usarse agentes terapéuticos, tales como pioglitazona, vitamina E y/o metformina. Véase, por ejemplo, Sanyal *et al.*, 2010, *NEJM*, 362: 1675-1685. En algunos casos, dicha intervención precoz puede retrasar o prevenir la insuficiencia hepática y la necesidad de un trasplante de hígado.

- 5 Además, en algunas realizaciones, un nivel de expresión diferencial de uno o más de los biomarcadores en un individuo a lo largo del tiempo puede ser indicativo de la respuesta del individuo a un determinado régimen terapéutico. En algunas realizaciones, los cambios en la expresión de uno o más de los biomarcadores durante el control de seguimiento pueden indicar que un determinado tratamiento es eficaz o pueden sugerir que el régimen
- 10 terapéutico se debe alterar de alguna manera, tal como controlando de forma más agresiva el azúcar en sangre, buscar más agresivamente la pérdida de peso, etc. En algunas realizaciones, un nivel de expresión constante de uno o más de los biomarcadores en un individuo a lo largo del tiempo puede ser indicativo de que la esteatosis de un individuo no está empeorando, o no se está desarrollando en EHNA.
- 15 Además de analizar los niveles de biomarcadores como un ensayo de diagnóstico independiente, los niveles de biomarcadores también pueden realizarse junto con la determinación de polimorfismos de un solo nucleótido (SNP) u otras lesiones o variabilidad genéticas que son indicativas de un mayor riesgo de susceptibilidad de la enfermedad. (Véase, por ejemplo, Amos *et al.*, *Nature Genetics* 40, 616-622 (2009)).
- 20 Además de analizar los niveles de biomarcadores como un ensayo de diagnóstico independiente, los niveles de biomarcadores también se pueden realizar junto con otros métodos de detección de la EHGNA, tales como la detección de un hígado agrandado, el análisis de sangre (por ejemplo, para detectar elevaciones en ciertas enzimas hepáticas, como ALT y/o AST), los ultrasonidos abdominales y la biopsia hepática. En algunos casos, los métodos que usan los biomarcadores descritos en el presente documento pueden facilitar la justificación médica y económica
- 25 para implementar tratamientos más agresivos para la EHGNA o la EHNA, la detección de seguimiento más frecuente, etc. Los biomarcadores también se pueden usar para comenzar el tratamiento en individuos con riesgo de desarrollar EHGNA, pero que no han sido diagnosticados de esteatosis, si el ensayo de diagnóstico indica que es probable que se desarrolle la enfermedad.
- 30 Además de analizarse los niveles de biomarcadores en combinación con otros métodos de diagnóstico de la EHGNA, la información con respecto a los biomarcadores también se puede evaluar junto con otros tipos de datos, en particular, los datos que indican el riesgo individual de EHGNA. Estos diversos datos pueden evaluarse mediante métodos automáticos, tales como un programa/software informático, que pueden incorporarse a un ordenador u otro aparato/dispositivo.

### 35 **Detección y determinación de biomarcadores y niveles de biomarcadores**

- Un nivel de biomarcador para los biomarcadores descritos en el presente documento puede detectarse usando cualquiera de varios métodos analíticos conocidos. En una realización, se detecta un nivel de biomarcador usando
- 40 un reactivo de captura. En diversas realizaciones, el reactivo de captura puede exponerse al biomarcador en solución o puede exponerse al biomarcador mientras el reactivo de captura está inmovilizado en un soporte sólido. En otras realizaciones, el reactivo de captura contiene una característica que es reactiva con una característica secundaria en un soporte sólido. En estas realizaciones, el reactivo de captura puede exponerse al biomarcador en solución, y después, se puede usar la característica del reactivo de captura junto con la característica secundaria
- 45 sobre el soporte sólido para inmovilizar el biomarcador sobre el soporte sólido. El reactivo de captura se selecciona según el tipo de análisis que se realizará. Los reactivos de captura incluyen, pero sin limitación, aptámeros, anticuerpos, adnectinas, anquirinas, otros miméticos de anticuerpos y otros armazones proteicos, autoanticuerpos, quimeras, moléculas pequeñas, fragmentos F(ab')<sub>2</sub>, fragmentos de anticuerpos monocatenarios, fragmentos Fv, fragmentos Fv monocatenarios, ácidos nucleicos, lectinas, receptores de unión a ligandos, aficuerpos,
- 50 nanoanticuerpos, polímeros impresos, avímeros, peptidomiméticos, receptores de hormonas, receptores de citocinas y receptores sintéticos, y modificaciones y fragmentos de estos.

En algunas realizaciones, se detecta un nivel de biomarcador usando un complejo de biomarcador/reactivo de captura.

- 55 En algunas realizaciones, el nivel de biomarcador se deriva del complejo de biomarcador/reactivo de captura y se detecta indirectamente, tal como, por ejemplo, como resultado de una reacción que es posterior a la interacción del biomarcador/reactivo de captura, pero que depende de la formación del complejo biomarcador/reactivo de captura.
- 60 En algunas realizaciones, el nivel de biomarcador se detecta directamente del biomarcador en una muestra biológica.

- En algunas realizaciones, los biomarcadores se detectan usando un formato multiplexado que permite la detección simultánea de dos o más biomarcadores en una muestra biológica. En algunas realizaciones del formato multiplexado, los reactivos de captura se inmovilizan, directa o indirectamente, de forma covalente o no covalente, en ubicaciones diferenciadas sobre un soporte sólido. En algunas realizaciones, un formato multiplexado usa
- 65

soportes sólidos diferenciados en los que cada soporte sólido tiene un reactivo de captura único asociado con ese soporte sólido, tal como, por ejemplo, puntos cuánticos. En algunas realizaciones, se usa un dispositivo individual para la detección de cada uno de los biomarcadores múltiples que se va a detectar en una muestra biológica. Los dispositivos individuales se pueden configurar para permitir que cada biomarcador de la muestra biológica se procese simultáneamente. Por ejemplo, se puede usar una placa de microtitulación de manera que cada pocillo de la placa se use para analizar uno o más de los múltiples biomarcadores que se van a detectar en una muestra biológica.

En una o más de las realizaciones anteriores, se puede usar un marcador fluorescente para marcar un componente del complejo de biomarcador/reactivo de captura para permitir la detección del nivel de biomarcador. En diversas realizaciones, el marcador fluorescente puede conjugarse con un reactivo de captura específico de cualquiera de los biomarcadores descritos en el presente documento usando técnicas conocidas, y el marcador fluorescente se puede usar luego para detectar el nivel de biomarcador correspondiente. Los marcadores fluorescentes adecuados incluyen quelatos de tierras raras, fluoresceína y sus derivados, rodamina y sus derivados, dansilo, alofococianina, PBXL-3, Qdot 605, lisamina, ficoeritrina, Rojo Texas y otros compuestos similares.

En algunas realizaciones, el marcador fluorescente es una molécula de colorante fluorescente. En algunas realizaciones, la molécula de colorante fluorescente incluye al menos un sistema de anillo de indolio sustituido en el que el sustituyente en el tercer átomo de carbono del anillo de indolio contiene un grupo químicamente reactivo o una sustancia conjugada. En algunas realizaciones, la molécula de colorante incluye una molécula de AlexaFluor, tal como, por ejemplo, AlexaFluor 488, AlexaFluor 532, AlexaFluor 647, AlexaFluor 680 o AlexaFluor 700. En algunas realizaciones, la molécula de colorante incluye un primer tipo y un segundo tipo de molécula de colorante, tal como, por ejemplo, dos moléculas AlexaFluor diferentes. En algunas realizaciones, la molécula de colorante incluye un primer tipo y un segundo tipo de molécula de colorante, y las dos moléculas de colorante tienen diferentes espectros de emisión.

La fluorescencia se puede medir con diversa instrumentación compatible con una amplia selección de formatos de ensayo. Por ejemplo, los espectrofluorímetros se han diseñado para analizar placas de microtitulación, portaobjetos de microscopio, matrices impresas, cubetas, etc. Véase "Principles of Fluorescence Spectroscopy", de J. R. Lakowicz, Springer Science + Business Media, Inc., 2004. Véase "Bioluminescence & Chemiluminescence: Progress & Current Applications"; Philip E. Stanley y Larry J. Kricka editors, World Scientific Publishing Company, enero de 2002.

En una o más realizaciones, se puede usar, opcionalmente, un marcador de quimioluminiscencia para marcar un componente del complejo de biomarcador/captura para permitir la detección de un nivel de biomarcador. Los materiales quimioluminiscentes adecuados incluyen cloruro de oxalilo, Rodamina 6G, Ru(bipy)<sub>3</sub><sup>2+</sup>, TMAE (tetraquis(dimetilamino)etileno), Pirogalol (1,2,3-trihidroxibenceno), Lucigenina, peroxioxalatos, Ariloxalatos, ésteres de Acridinio, dioxetanos y otros.

En algunas realizaciones, el método de detección incluye una combinación de enzima/sustrato que genera una señal detectable que corresponde al valor de biomarcador. En general, la enzima cataliza una alteración química del sustrato cromogénico que puede medirse usando diversas técnicas, incluyendo espectrofotometría, fluorescencia y quimioluminiscencia. Las enzimas adecuadas incluyen, por ejemplo, luciferasas, luciferina, malato deshidrogenasa, ureasa, peroxidasa de rábano picante (HRPO), fosfatasa alcalina, beta-galactosidasa, glucoamilasa, lisozima, glucosa oxidasa, galactosa oxidasa y glucosa-6-fosfato deshidrogenasa, uricasa, xantina oxidasa, lactoperoxidasa, microperoxidasa y similares.

En algunas realizaciones, el método de detección puede ser una combinación de fluorescencia, quimioluminiscencia, radionúclido o combinaciones de enzima/sustrato que generan una señal medible. En algunas realizaciones, la señalización multimodal podría tener características únicas y ventajosas en formatos de ensayo de biomarcadores.

En algunas realizaciones, los niveles de biomarcadores para los biomarcadores descritos en el presente documento pueden detectarse usando métodos analíticos conocidos incluyendo ensayos de aptámeros individuales, ensayos multiplexados basados en aptámeros, inmunoensayos individuales o múltiples, perfil de expresión de ARNm, perfil de expresión de miARN, análisis espectrométrico de masas, métodos histológicos/citológicos, etc. como se describe más adelante.

#### **Determinación de niveles de biomarcadores usando ensayos basados en aptámeros**

Los ensayos dirigidos a la detección y cuantificación de moléculas fisiológicamente significativas en muestras biológicas y otras muestras son herramientas importantes en la investigación científica y en el campo de la atención sanitaria. Una clase de dichos ensayos implica el uso de una micromatriz que incluye uno o más aptámeros inmovilizados en un soporte sólido. Los aptámeros son capaces cada uno de unirse con una molécula diana de una manera altamente específica y con afinidad muy alta. Véase, por ejemplo, la patente de EE.UU. n.º 5.475.096 titulada "Nucleic Acid Ligands"; véase también, por ejemplo, la patente de EE.UU. n.º 6.242.246, la patente de EE.UU. n.º 6.458.543 y la patente de EE.UU. n.º 6.503.715, tituladas "Nucleic Acid Ligand Diagnostic Biochip". Una

vez que la micromatriz se pone en contacto con una muestra, los aptámeros se unen con sus respectivas moléculas diana presentes en la muestra y, de este modo, permiten una determinación de un nivel de biomarcador correspondiente a un biomarcador.

5 Como se usa en el presente documento, un "aptámero" se refiere a un ácido nucleico que tiene una afinidad de unión específica hacia una molécula diana. Se reconoce que las interacciones de afinidad son cuestión del grado, sin embargo, en el presente contexto, la "afinidad de unión específica" de un aptámero hacia su diana significa que el aptámero se une con su diana en general con un grado mucho mayor de afinidad que con el que se une con otros componentes en una muestra de ensayo. Un "aptámero" es un conjunto de copias de un tipo o una especie de  
10 molécula de ácido nucleico que tiene una determinada secuencia de nucleótidos. Un aptámero puede incluir cualquier número adecuado de nucleótidos, incluyendo cualquier número de nucleótidos modificados químicamente. "Aptámeros" se refiere a más de uno de dicho conjunto de moléculas. Diferentes aptámeros pueden tener el mismo número o diferentes números de nucleótidos. Los aptámeros pueden ser ADN o ARN o ácidos nucleicos químicamente modificados, y pueden ser monocatenarios, bicatenarios o contener regiones bicatenarias, y pueden  
15 incluir estructuras de mayor orden. Un aptámero también puede ser un fotoaptámero, en el que se incluye un grupo funcional fotorreactivo o químicamente reactivo en el aptámero para permitir que se una covalentemente con su diana correspondiente. Cualquiera de los métodos de aptámeros desvelados en el presente documento puede incluir el uso de dos o más aptámeros que se unen específicamente con la misma molécula diana. Como se describe mejor más adelante, un aptámero puede incluir un marcador. Si un aptámero incluye un marcador, no es necesario que  
20 todas las copias del aptámero tengan el mismo marcador. Además, si diferentes aptámeros incluyen cada uno un marcador, estos diferentes aptámeros pueden tener el mismo marcador o un marcador diferente.

Un aptámero puede identificarse usando cualquier método conocido, incluyendo el proceso de SELEX. Una vez  
25 identificado, un aptámero puede prepararse o sintetizarse de acuerdo con cualquier método conocido, incluyendo métodos sintéticos químicos y métodos sintéticos enzimáticos.

El término "SELEX" y la expresión "proceso SELEX" se usan indistintamente en el presente documento para hacer referencia, en general, a una combinación de (1) la selección de aptámeros que interaccionan con una molécula  
30 diana de una manera deseable, por ejemplo, la unión con alta afinidad a una proteína; con (2) la amplificación de los ácidos nucleicos seleccionados. El proceso SELEX se puede usar para identificar aptámeros con alta afinidad hacia una diana o un biomarcador específico.

SELEX incluye, en general, preparar una mezcla candidata de ácidos nucleicos, unir la mezcla candidata con la  
35 molécula diana deseada para formar un complejo de afinidad, separar los complejos de afinidad de los ácidos nucleicos candidatos no unidos, separar y aislar el ácido nucleico del complejo de afinidad, purificar el ácido nucleico e identificar una secuencia de aptámeros específica. El proceso puede incluir múltiples series para perfeccionar la afinidad del aptámero seleccionado. El proceso puede incluir etapas de amplificación en uno o más puntos en el proceso. Véase, por ejemplo, la patente de EE.UU. n.º 5.475.096, titulada "Nucleic Acid Ligands". El proceso SELEX se puede usar para generar un aptámero que se una covalentemente con su diana, así como un aptámero que no se  
40 una covalentemente con su diana. Véase, por ejemplo, la patente de EE.UU. n.º 5.705.337, titulada "Systematic Evolution of Nucleic Acid Ligands by Exponential Enrichment: Chemi-SELEX".

El proceso SELEX se puede usar para identificar aptámeros de alta afinidad que contengan nucleótidos modificados que confieran características mejoradas al aptámero, tales como, por ejemplo, mejor estabilidad *in vivo* o mejores  
45 características de administración. Los ejemplos de dichas modificaciones incluyen sustituciones químicas en las posiciones de ribosa y/o fosfato y/o base. Los aptámeros identificados con el proceso SELEX que contienen nucleótidos modificados se describen en la patente de los EE.UU. n.º 5.660.985, titulada "High Affinity Nucleic Acid Ligands Containing Modified Nucleotides", que describe oligonucleótidos que contienen derivados de nucleótidos modificados químicamente en las posiciones 5' y 2' de las pirimidinas. La patente de EE.UU. n.º 5.580.737, véase  
50 *supra*, describe aptámeros altamente específicos que contienen uno o más nucleótidos modificados con 2'-amino (2'-NH<sub>2</sub>), 2'-fluoro (2'-F) y/o 2'-O-metilo (2'-OMe). Véase también la publicación de solicitud de patente de EE.UU. n.º 2009/0098549, titulada "SELEX and PHOTOSELEX", que describe bibliotecas de ácido nucleico que tienen propiedades físicas y químicas expandidas, y su uso en SELEX y photoSELEX.

SELEX también puede usarse para identificar aptámeros que tengan características de velocidad de disociación deseables. Véase la publicación de EE.UU. n.º US 2009/0004667, titulada "Method for Generating Aptamers with Improved Off-Rates", que describe métodos de SELEX mejorados para generar aptámeros que pueden unirse con  
55 moléculas diana. Se describen métodos para producir aptámeros y fotoaptámeros que tengan velocidades de disociación más lentas de sus respectivas moléculas diana. Los métodos implican poner en contacto la mezcla candidata con la molécula diana, permitir que se produzca la formación de complejos de ácido nucleico-diana, y realizar un proceso de enriquecimiento de velocidad de disociación lenta en el que los complejos de ácido nucleico-diana con velocidades de disociación rápidas se disociarán y no se volverán a formar, mientras que los complejos con velocidades de disociación lentas permanecerán intactos. Además, los métodos incluyen el uso de nucleótidos modificados en la producción de mezclas candidatas de ácido nucleico para generar aptámeros con mejor  
60 rendimiento de velocidad de disociación. Los nucleótidos modificados ilustrativos no limitantes incluyen, por ejemplo, las pirimidinas modificadas que se muestran en la Figura 11. En algunas realizaciones, un aptámero comprende al

menos un nucleótido con una modificación, tal como la modificación de una base. En algunas realizaciones, un aptámero comprende al menos un nucleótido con una modificación hidrófoba, tal como la modificación de una base hidrófoba, que permite contactos hidrófobos con una proteína diana. Dichos contactos hidrófobos, en algunas realizaciones, contribuyen a una mayor afinidad y/o a una velocidad de disociación más lenta por el aptámero. Los ejemplos de nucleótidos no limitantes con modificaciones hidrófobas se muestran en la Figura 11. En algunas realizaciones, un aptámero comprende al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho, al menos nueve o al menos 10 nucleótidos con modificaciones hidrófobas, en el que cada modificación hidrófoba puede ser igual o diferente de las otras. En algunas realizaciones, se pueden seleccionar al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho, al menos nueve o al menos 10 modificaciones hidrófobas de un aptámero independientemente de las modificaciones hidrófobas que se muestran en la Figura 11.

En algunas realizaciones, un aptámero de velocidad de disociación lenta (incluyendo los aptámeros que comprenden al menos uno nucleótido con una modificación hidrófoba) tiene una velocidad de disociación ( $t^{1/2}$ ) de  $\geq 30$  minutos,  $\geq 60$  minutos,  $\geq 90$  minutos,  $\geq 120$  minutos,  $\geq 150$  minutos,  $\geq 180$  minutos,  $\geq 210$  minutos o  $\geq 240$  minutos.

En algunas realizaciones, un ensayo emplea aptámeros que incluyen grupos funcionales fotorreactivos que permiten que los aptámeros se unan covalentemente o se "fotorreticulen" sus moléculas diana. Véase, por ejemplo, la patente de EE.UU. n.º 6.544.776 titulada "Nucleic Acid Ligand Diagnostic Biochip". Estos aptámeros fotorreactivos también se denominan fotoaptámeros. Véase, por ejemplo, la patente de EE.UU. n.º 5.763.177, la patente de EE.UU. n.º 6.001.577 y la patente de EE.UU. n.º 6.291.184, cada una de las cuales se titula "Systematic Evolution of Nucleic Acid Ligands by Exponential Enrichment: Photoselection of Nucleic Acid Ligands and Solution SELEX"; véase también, por ejemplo, la patente de EE.UU. n.º 6.458.539, titulada "Photoselection of Nucleic Acid Ligands". Tras ponerse en contacto la micromatriz con la muestra y de que los fotoaptámeros hayan tenido la oportunidad de unirse con sus moléculas diana, los fotoaptámeros se fotoactivan, y el soporte sólido se lava para retirar cualquier molécula unida de forma inespecífica. Pueden usarse condiciones de lavado rigurosas, ya que las moléculas diana que se unen a los fotoaptámeros, en general, no se retiran, debido a los enlaces covalentes creados por el/los grupo/s funcional/es fotoactivado/s en los fotoaptámeros. De esta manera, el ensayo permite la detección de un nivel de biomarcador correspondiente a un biomarcador en la muestra de ensayo.

En algunos formatos de ensayo, los aptámeros se inmovilizan en el soporte sólido antes de ponerse en contacto con la muestra. En ciertas circunstancias, sin embargo, la inmovilización de los aptámeros antes de su contacto con la muestra puede no proporcionar un ensayo óptimo. Por ejemplo, la preinmovilización de los aptámeros puede dar lugar a una mezcla ineficaz de los aptámeros con las moléculas diana en la superficie del soporte sólido, quizás conduciendo a tiempos de reacción largos y, por lo tanto, períodos de incubación ampliados para permitir la unión eficaz de los aptámeros con sus moléculas diana. Además, cuando se emplean fotoaptámeros en el ensayo y dependiendo del material utilizado como soporte sólido, el soporte sólido puede tender a dispersar o absorber la luz usada para efectuar la formación de enlaces covalentes entre los fotoaptámeros y sus moléculas diana. Además, dependiendo del método empleado, la detección de moléculas diana unidas a sus aptámeros puede estar sujeta a imprecisión, ya que la superficie del soporte sólido también puede exponerse a y verse afectada por cualquier agente de marcaje que se use. Finalmente, la inmovilización de los aptámeros en el soporte sólido, en general, implica una etapa de preparación de aptámeros (es decir, la inmovilización) antes de la exposición de los aptámeros a la muestra, y esta etapa de preparación puede afectar a la actividad o funcionalidad de los aptámeros.

También se han descrito ensayos de aptámeros que permiten que un aptámero capture su diana en solución y después emplear etapas de separación que se diseñan para retirar componentes específicos de la mezcla de aptámero-diana antes de la detección (véase la publicación de EE.UU. n.º 2009/0042206 titulada "Multiplexed Analyses of Test Samples"). Los métodos de ensayo de aptámeros descritos permiten la detección y cuantificación de una diana no de ácido nucleico (por ejemplo, una diana proteica) en una muestra de ensayo detectando y cuantificando un ácido nucleico (es decir, un aptámero). Los métodos descritos crean un sustituto de ácido nucleico (es decir, el aptámero) para detectar y cuantificar una diana no de ácido nucleico, permitiendo de este modo que se aplique la amplia variedad de tecnologías de ácido nucleico, incluyendo la amplificación, a una serie más amplia de dianas deseadas, incluyendo dianas proteicas.

Se pueden construir aptámeros para facilitar la separación de los componentes de ensayo de un complejo de biomarcadores de aptámero (o complejo covalente de biomarcadores de fotoaptámero) y permitir el aislamiento del aptámero para la detección y/o cuantificación. En una realización, estas construcciones pueden incluir un elemento escindible o liberable dentro de la secuencia del aptámero. En otras realizaciones, se puede introducir funcionalidad adicional en el aptámero, por ejemplo, un componente marcado o detectable, un componente espaciador, o un marcador de unión o elemento de inmovilización específico. Por ejemplo, el aptámero puede incluir un marcador conectado con el aptámero mediante un resto escindible, un marcador, un componente espaciador que separa el marcador y el resto escindible. En una realización, un elemento escindible es un enlazador fotoescindible. El enlazador fotoescindible puede unirse con un resto de biotina y una sección espaciadora, puede incluir un grupo NHS para la derivatización de aminas, y puede usarse para introducir un grupo de biotina en un aptámero, permitiendo de este modo la liberación del aptámero posteriormente en un método de ensayo.

Los ensayos homogéneos, realizados con todos los componentes de ensayo en solución, no requieren la separación de muestra y reactivos antes de la detección de señal. Estos métodos son rápidos y fáciles de usar. Estos métodos generan señal basándose en un reactivo de captura molecular o unión que reacciona con su diana específica. En algunas realizaciones de los métodos descritos en el presente documento, los reactivos de captura molecular comprenden un aptámero o un anticuerpo o similar, y la diana específica puede ser un biomarcador mostrado en la Tabla 3, 4, 6, 7, 8 y/o 9.

En algunas realizaciones, un método de generación de señal aprovecha el cambio de señal de anisotropía debido a la interacción de un reactivo de captura marcado con fluoróforo con su diana biomarcadora específica. Cuando el reactivo de captura marcado reacciona con su diana, el aumento de peso molecular hace que el movimiento rotacional del fluoróforo unido con el complejo se haga mucho más lento cambiando el valor de anisotropía. Controlando el cambio de anisotropía, se pueden usar uniones para medir de forma cuantitativa los biomarcadores en soluciones. Otros métodos incluyen ensayos de polarización de fluorescencia, métodos de balizas moleculares, inactivación de la fluorescencia resuelta en el tiempo, quimioluminiscencia, transferencia de energía por resonancia de fluorescencia y similares.

Un ensayo de aptámero basado en solución ilustrativo que se puede usar para detectar un nivel de biomarcador en una muestra biológica incluye lo siguiente: (a) preparar una mezcla poniendo en contacto la muestra biológica con un aptámero que incluye un primer marcador y que tiene una afinidad específica por el biomarcador, en el que se forma un complejo de afinidad de aptámero cuando el biomarcador está presente en la muestra; (b) exponer la muestra a un primer soporte sólido que incluye un primer elemento de captura, y permitir que el primer marcador se asocie con el primer elemento de captura, (c) retirar cualquier componente de la mezcla no asociado con el primer soporte sólido; (d) unir un segundo marcador con el componente de biomarcador del complejo de afinidad de aptámero; (e) liberar el complejo de afinidad de aptámero del primer soporte sólido; (f) exponer el complejo de afinidad de aptámero liberado a un segundo soporte sólido que incluye un segundo elemento de captura y permitir que el segundo marcador se asocie con el segundo elemento de captura; (g) retirar cualquier aptámero que no forme complejo de la mezcla dividiendo el aptámero que no forma complejo del complejo de afinidad del aptámero; (h) eluir el aptámero del soporte sólido; e (i) detectar el biomarcador detectando el componente de aptámero del complejo de afinidad de aptámero.

En el Ejemplo 7, se describe un método ilustrativo no limitante para detectar biomarcadores en una muestra biológica usando aptámeros. Véase también Kraemer *et al.*, PLoS One 6(10): e26332.

### **Determinación de niveles de biomarcadores usando inmunoensayos**

Los métodos de inmunoensayos se basan en la reacción de un anticuerpo hacia su diana o analito correspondiente, y pueden detectar el analito en una muestra dependiendo del formato de ensayo específico. Para mejorar la especificidad y sensibilidad de un método de ensayo basado en la inmunorreactividad, se suelen usar anticuerpos monoclonales y fragmentos de los mismos debido a su reconocimiento de epítopos específicos. Los anticuerpos policlonales también se han usado satisfactoriamente en diversos inmunoensayos debido a su mayor afinidad hacia la diana en comparación con anticuerpos monoclonales. Se han diseñado inmunoensayos para su uso con una amplia selección de matrices de muestras biológicas. Se han diseñado formatos de inmunoensayos para proporcionar resultados cualitativos, semicuantitativos y cuantitativos.

Se generan resultados cuantitativos mediante el uso de una curva patrón creada con concentraciones conocidas del analito específico que se va a detectar. La respuesta o señal de una muestra desconocida se representa en la curva patrón, y se establece una cantidad o un nivel correspondiente a la diana en la muestra desconocida.

Se han diseñado numerosos formatos de inmunoensayo. ELISA o EIA pueden ser cuantitativos para la detección de un analito. Este método basado en la unión de un marcador con el analito o el anticuerpo y el componente marcador incluye, BIE directa o indirectamente, una enzima. Los ensayos de ELISA pueden formatearse para la detección directa, indirecta, competitiva o de tipo sándwich del analito. Otros métodos se basan en marcadores tales como, por ejemplo, radioisótopos ( $I^{125}$ ) o fluorescencia. Las técnicas adicionales incluyen, por ejemplo, aglutinación, nefelometría, turbidimetría, transferencia Western, inmunoprecipitación, inmunocitoquímica, inmunohistoquímica, citometría de flujo, ensayo Luminex y otros (véase "ImmunoAssay: A Practical Guide", editado por Brian Law, publicado por Taylor y Francis, Ltd., edición 2005).

Los formatos de ensayo ilustrativos incluyen ensayo inmunoabsorbente ligado a enzimas (ELISA), radioinmunoensayo, fluorescencia, quimioluminiscencia e inmunoensayos de transferencia de energía por resonancia de fluorescencia (FRET) o FRET resuelta en el tiempo (TR-FRET). Los ejemplos de procedimientos de detección de biomarcadores incluyen la inmunoprecipitación de biomarcadores seguida de métodos cuantitativos que permiten la diferenciación del nivel de péptido y del tamaño, tales como electroforesis en gel, electroforesis capilar, electrocromatografía planar y similares.

Los métodos de detección y/o cuantificación de un marcador detectable o material generador de señal dependen de

la naturaleza del marcador. Los productos de reacciones catalizadas por enzimas apropiadas (cuando el marcador detectable sea una enzima; véase anteriormente) pueden ser, sin limitación, fluorescentes, luminiscentes o radiactivos, o pueden absorber luz visible o ultravioleta. Los ejemplos de detectores adecuados para detectar dichos marcadores detectables incluyen, sin limitación, película de rayos X, contadores de radiactividad, contadores de centelleo, espectrofotómetros, colorímetros, fluorímetros, luminómetros y densitómetros.

Cualquiera de los métodos de detección puede realizarse en cualquier formato que permita cualquier preparación, procesamiento y análisis de las reacciones adecuados. Este puede ser, por ejemplo, en placas de ensayos de múltiples pocillos (por ejemplo, 96 pocillos o 386 pocillos) o usando cualquier matriz o micromatriz adecuada. Pueden prepararse soluciones madre para diversos agentes manual o robóticamente, y todo el pipeteo, la dilución, el mezclado, la distribución, el lavado, la incubación, la lectura de muestras, la recogida de datos y el análisis posteriores puede realizarse robóticamente usando software de análisis disponible en el mercado, robótica e instrumentación de detección capaz de detectar un marcador detectable.

#### **Determinación de niveles de biomarcadores usando perfil de expresión génica**

La medición del ARNm en una muestra biológica se puede usar, en algunas realizaciones, como un sustituto para la detección del nivel de la proteína correspondiente en la muestra biológica. Por lo tanto, en algunas realizaciones, un biomarcador o grupo de biomarcadores descritos en el presente documento también puede detectarse mediante la detección del ARN apropiado.

En algunas realizaciones, los niveles de expresión de ARNm se miden por reacción en cadena de la polimerasa cuantitativa de transcripción inversa (RT-PCR seguida de qPCR). La RT-PCR se usa para crear un ADNc a partir del ARNm. El ADNc se puede usar en un ensayo de qPCR para producir fluorescencia a medida que el proceso de amplificación del ADN progresa. Por comparación con una curva patrón, la qPCR puede producir una medición absoluta tal como el número de copias de ARNm por célula. Las transferencias Northern, las micromatrices, los ensayos de Invader y la RT-PCR combinada con electroforesis capilar se han usado todos para medir los niveles de expresión de ARNm en una muestra. Véase "Gene Expression Profiling: Methods and Protocols", Richard A. Shinkets, editor, Humana Press, 2004.

#### **Detección de biomarcadores usando tecnologías de generación de imágenes moleculares *in vivo***

En algunas realizaciones, se puede usar un biomarcador descrito en el presente documento en ensayos de generación de imágenes moleculares. Por ejemplo, se puede acoplar un agente de generación de imágenes a un reactivo de captura, que se puede usar para detectar el biomarcador *in vivo*.

Las tecnologías de generación de imágenes *in vivo* proporcionan métodos no invasivos de determinación del estado de una determinada enfermedad en el cuerpo de un individuo. Por ejemplo, se pueden ver partes enteras del cuerpo, o incluso todo el cuerpo, como una imagen tridimensional, proporcionando de este modo información valiosa relativa a la morfología y a las estructuras del cuerpo. Dichas tecnologías pueden combinarse con la detección de los biomarcadores descritos en el presente documento para proporcionar información relativa al biomarcador *in vivo*.

El uso de tecnologías de generación de imágenes moleculares *in vivo* se expande debido a diversos avances en la tecnología. Estos avances incluyen el desarrollo de nuevos agentes de contraste o marcadores, tales como radiomarcadores y/o marcadores fluorescentes, que pueden proporcionar fuertes señales dentro del cuerpo; y el desarrollo de nueva tecnología de generación de imágenes potente, que puede detectar y analizar estas señales desde fuera del cuerpo, con suficiente sensibilidad y precisión para proporcionar información útil. El agente de contraste puede visualizarse en un sistema de generación de imágenes apropiado, proporcionando de este modo una imagen de la parte o las partes del cuerpo en las que se encuentra el agente de contraste. El agente de contraste puede unirse o asociarse con un reactivo de captura tal como un aptámero o un anticuerpo, por ejemplo y/o con un péptido o una proteína, o un oligonucleótido (por ejemplo, para la detección de la expresión génica), o un complejo que contiene cualquiera de estos con una o más macromoléculas y/u otras formas de partículas.

El agente de contraste también puede presentar un átomo radiactivo que es útil en la generación de imágenes. Los átomos radiactivos adecuados incluyen tecnecio-99m o yodo-123 para estudios escintigráficos. Otras fracciones fácilmente detectables incluyen, por ejemplo, marcadores de espín para la generación de imágenes por resonancia magnética (MRI) tales como, por ejemplo, yodo-123 de nuevo, yodo-131, indio-111, flúor-19, carbono-13, nitrógeno-15, oxígeno-17, gadolinio, manganeso o hierro. Dichos marcadores se conocen bien en la técnica y podrían ser fácilmente seleccionados por un experto habitual en la materia.

Las técnicas de generación de imágenes convencionales incluyen, pero sin limitación, la generación de imágenes por resonancia magnética, exploración de tomografía computarizada, tomografía de emisión de positrones (PET), tomografía computarizada de emisión de un único fotón (SPECT) y similares. Para la generación de imágenes *in vivo* de diagnóstico, el tipo de instrumento de detección disponible es un factor importante en la selección de un agente de contraste dado, tal como un radionúclido dado y el biomarcador en particular que se usa para dirigir (proteína, ARNm y similares). El radionúclido escogido normalmente tiene un tipo de degradación que es detectable

por un tipo de instrumento dado. Además, cuando se selecciona un radionúclido para el diagnóstico *in vivo*, su semivida debería ser suficientemente larga para permitir la detección en el momento de captación máxima por el tejido diana, pero suficientemente corta para que se reduzca al mínimo la radiación perjudicial del hospedador.

- 5 Las técnicas de generación de imágenes ilustrativas incluyen, pero sin limitación, PET y SPECT, que son técnicas de generación de imágenes en las que un radionúclido se administra sintética o localmente a un individuo. La captación posterior del radiotrazador se mide a lo largo del tiempo y se usa para obtener información acerca del tejido diana y del biomarcador. Debido a las emisiones de alta energía (rayos gamma) de los isótopos específicos empleados, y la sensibilidad y sofisticación de los instrumentos usados para detectarlos, la distribución  
10 bidimensional de la radiactividad puede inferirse desde fuera del cuerpo.

- Los núclidos emisores de positrones usados normalmente en PET incluyen, por ejemplo, carbono-11, nitrógeno-13, oxígeno-15 y flúor-18. Se usan isótopos que se degradan por captura de electrones y/o emisión gamma en SPECT e incluyen, por ejemplo yodo-123 y tecnecio-99m. Un método ilustrativo para marcar aminoácidos con tecnecio-99m es la reducción de ion pertecnetato en presencia de un precursor quelante para formar el complejo de tecnecio-99m-precursor lábil, que, a su vez, reacciona con el grupo de unión a metales de un péptido quimiotáctico modificado  
15 bifuncionalmente para formar un conjugado de tecnecio-99m-péptido quimiotáctico.

- Los anticuerpos se usan frecuentemente para dichos métodos de diagnóstico de generación de imágenes *in vivo*. La preparación y el uso de anticuerpos para diagnóstico *in vivo* se conocen bien en la técnica. De forma similar, pueden usarse aptámeros para dichos métodos de diagnóstico de generación de imágenes *in vivo*. Por ejemplo, un aptámero que se usó para identificar un determinado biomarcador descrito en el presente documento puede marcarse de forma apropiada e inyectarse en un individuo para detectar el biomarcador *in vivo*. El marcador usado se seleccionará de acuerdo con la modalidad de generación de imágenes que se vaya a usar, como se ha descrito  
20 previamente. Los agentes de generación de imágenes dirigidos a aptámeros podrían tener características únicas y ventajosas en relación con la penetración tisular, distribución tisular, cinética, eliminación, potencia y selectividad en comparación con otros agentes de generación de imágenes.

- Dichas técnicas también pueden realizarse opcionalmente con oligonucleótidos marcados, por ejemplo, para la detección de la expresión génica mediante la generación de imágenes con oligonucleótidos antisentido. Estos métodos se usan para la hibridación *in situ*, por ejemplo, con moléculas fluorescentes o radionúclidos como el marcador. Otros métodos para la detección de la expresión génica incluyen, por ejemplo, la detección de la actividad de un gen indicador.  
30

- Otro tipo general de tecnología de generación de imágenes es la generación de imágenes óptica, en la que se detectan señales fluorescentes dentro del sujeto mediante un dispositivo óptico que es externo al sujeto. Estas señales pueden deberse a la fluorescencia real y/o a la bioluminiscencia. Las mejoras en la sensibilidad de los dispositivos de detección óptica han aumentado la utilidad de la generación de imágenes ópticas para ensayos de diagnóstico *in vivo*.  
40

Para una revisión de otras técnicas, véase N. Blow, *Nature Methods*, 6, 465-469, 2009.

#### **Determinación de los biomarcadores usando métodos de histología/citología**

- 45 Los biomarcadores descritos en el presente documento se pueden detectar en varias muestras de tejidos usando métodos histológicos o citológicos. Por ejemplo, se pueden usar biopsias endo y transbronquiales, aspirados de aguja fina, agujas de grabación y biopsias centrales para histología. Se puede usar el lavado y cepillado bronquial, la aspiración pleural y el esputo, para citología. Cualquiera de los biomarcadores identificados en el presente documento se puede usar para teñir una muestra como un indicio de enfermedad.  
50

- Uno o más reactivos de captura específicos del/de los biomarcador/es correspondiente/s pueden usarse en una evaluación citológica de una muestra, y pueden incluir uno o más de los siguientes: recoger una muestra celular, fijar la muestra celular, deshidratar, clarificar, inmovilizar la muestra celular en un portaobjetos de microscopio, permeabilizar la muestra celular, tratar para la recuperación de analitos, teñir, destañir, lavar, bloquear y hacer reaccionar con uno o más reactivos de captura en una solución tamponada. En otra realización, la muestra celular se produce a partir de un bloque celular.  
55

- Uno o más reactivos de captura específicos de los biomarcadores correspondientes se usan en una evaluación histológica de una muestra tisular, y pueden incluir uno o más de los siguientes: recoger una muestra tisular, fijar la muestra tisular, deshidratar, clarificar, inmovilizar la muestra tisular en un portaobjetos de microscopio, permeabilizar la muestra tisular, tratar para la recuperación de analitos, teñir, destañir, lavar, bloquear, rehidratar, y hacer reaccionar con uno o varios reactivos de captura en una solución tamponada. En otra realización, la fijación y la deshidratación se reemplazan con congelación.  
60

- 65 El uno o más aptámeros específicos del/de los biomarcador/es correspondiente/s, pueden hacerse reaccionar con la muestra histológica o citológica y pueden actuar como la diana de ácido nucleico en un método de amplificación de

ácido nucleico. Los métodos de amplificación de ácidos nucleicos adecuados incluyen, por ejemplo, PCR, q-beta replicasa, amplificación por círculo rodante, desplazamiento de cadena, amplificación dependiente de helicasa, amplificación isotérmica mediada por bucle, reacción en cadena de la ligasa, y amplificación por círculo rodante ayudado por restricción y circularización.

El uno o más reactivos de captura específicos de los biomarcadores correspondientes para su uso en la evaluación histológica o citológica pueden mezclarse en una solución tamponada que puede incluir cualquiera de los siguientes: materiales de bloqueo, competidores, detergentes, estabilizantes, vehículo de ácido nucleico, materiales polianiónicos, etc.

Un "protocolo de citología" incluye, en general, la recogida de muestras, la fijación de muestras, la inmovilización de muestras y la tinción. La "preparación celular" puede incluir varias etapas de procesamiento después de recogida de muestras, incluyendo el uso de uno más aptámeros para la tinción de las células preparadas.

## **Determinación de los niveles de biomarcadores usando métodos de espectrometría de masas**

Se puede usar varias configuraciones de espectrómetros de masas para detectar los niveles de biomarcadores. Están disponibles varios tipos de espectrómetros de masas o pueden producirse con diversas configuraciones. En general, un espectrómetro de masas tiene los siguientes componentes principales: una entrada de muestras, una fuente de iones, un analizador de masas, un detector, un sistema de vacío y un sistema de control de instrumentos y un sistema de datos. La diferencia en la entrada de muestras, la fuente de iones y el analizador de masas, en general, define el tipo de instrumento y sus capacidades. Por ejemplo, una entrada puede ser una fuente de cromatografía líquida en columna de capilar o puede ser una sonda directa o fase tal como la usada en desorción por láser asistida por matriz. Las fuentes de iones comunes son, por ejemplo, electropulverización, incluyendo nanopulverización y micropulverización o desorción láser asistida por matriz. Los analizadores de masa comunes incluyen un filtro de masas cuadrupolo, un analizador de masas de trampa iónica y un analizador de masas de tiempo de vuelo. Se conocen bien en la técnica métodos de espectrometría de masas adicionales (véase Burlingame *et al. Anal. Chem.* 70:647 R-716R (1998); Kinter y Sherman, Nueva York (2000)).

Pueden detectarse biomarcadores proteicos y niveles de biomarcadores, y medirse mediante cualquiera de los siguientes: espectrometría de masas por ionización de electropulverización (ESI-MS), ESI-MS/MS, ESI-MS/(MS)<sub>n</sub>, espectrometría de masas por ionización de desorción de láser asistido por matriz tiempo de vuelo (MALDI-TOF-MS), espectrometría de masas de desorción/ionización por láser potenciada en superficie tiempo de vuelo (SELDI-TOF-MS), desorción/ionización en silicio (DIOS), espectrometría de masas de iones secundarios (SIMS), cuadrupolo tiempo de vuelo (Q-TOF), tecnología de tandem tiempo de vuelo (TOF/TOF), la denominada TOF/TOF ultraflex III, espectrometría de masas por ionización química a presión atmosférica (APCI-MS), APCI-MS/MS, APCI-MS)<sub>N</sub>, espectrometría de masas por fotoionización a presión atmosférica (APPI-MS), APPI-MS/MS y APPI-MS)<sub>N</sub>, espectrometría de masas cuadrupolo, espectrometría de masas transformada de Fourier (FTMS), espectrometría de masas cuantitativa y espectrometría de masas de trampa iónica.

Se usan estrategias de preparación de muestras para marcar y enriquecer muestras antes de la caracterización espectroscópica de masas de biomarcadores proteicos y la determinación de niveles de biomarcador. Los métodos de marcaje incluyen, pero sin limitación, marcador isobárico de cuantificación relativa y absoluta (iTRAQ) y marcaje de isótopos estables con aminoácidos en cultivo celular (SILAC). Los reactivos de captura usados para enriquecer selectivamente muestras con respecto a biomarcadores proteicos candidatos antes del análisis espectroscópico de masas incluyen, pero sin limitación, aptámeros, anticuerpos, sondas de ácido nucleico, quimeras, moléculas pequeñas, un fragmento F(ab')<sub>2</sub>, un fragmento de anticuerpo monocatenario, un fragmento Fv, un fragmento Fv monocatenario, un ácido nucleico, una lectina, un receptor de unión a ligando, anticuerpos, nanocuerpos, anquirinas, anticuerpos de dominio, armazones de anticuerpo alternativos (por ejemplo diacuerpos, etc.), polímeros grabados, avimeros, peptidomiméticos, peptoides, ácidos nucleicos peptídicos, ácido nucleico de treosa, un receptor de hormonas, un receptor de citocinas y receptores sintéticos, y modificaciones y fragmentos de los mismos.

Los ensayos anteriores permiten la detección de niveles de biomarcadores que son útiles en métodos descritos en el presente documento, en los que los métodos comprenden detectar, en una muestra biológica de un individuo, al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho o al menos nueve biomarcadores seleccionados entre los biomarcadores de la Tablas 3, 4, 6, 7, 8 y 9. Los métodos comprenden detectar los niveles de uno o más biomarcadores seleccionados de cualquiera de los grupos de biomarcadores descritos en el presente documento, tales como los grupos mostrados en la Tabla 5 y los subconjuntos de biomarcadores mostrados en las Tablas 3, 4, 6, 7, 8 y 9. Por lo tanto, aunque algunos de los biomarcadores descritos pueden ser útiles solos para detectar EHNA y/o EHNA, en el presente documento también se describen métodos para el agrupamiento de múltiples biomarcadores y subconjuntos de los biomarcadores para formar grupos de dos o más biomarcadores. De acuerdo con cualquiera de los métodos descritos en el presente documento, pueden detectarse y clasificarse individualmente niveles de biomarcadores o pueden detectarse y clasificarse colectivamente, como por ejemplo en un formato de ensayo multiplexado.

## **Clasificación de los biomarcadores y cálculo de puntuaciones de enfermedad**

En algunas realizaciones, una "identificación genética" de biomarcador para un ensayo de diagnóstico dado contiene un conjunto de biomarcadores, teniendo cada biomarcador niveles característicos en las poblaciones de interés. Niveles característicos, en algunas realizaciones, puede referirse a la media de los niveles de biomarcador para los individuos de un determinado grupo. Un método de diagnóstico descrito en el presente documento se puede usar para asignar una muestra desconocida de un individuo a uno de dos grupos, bien EHNA o normal. Un método de diagnóstico descrito en el presente documento se puede usar para asignar una muestra desconocida de un individuo a uno de dos grupos, bien EHNA o EHGNA. Un método de diagnóstico descrito en el presente documento se puede usar para asignar una muestra desconocida de un individuo a uno de tres grupos, EHGNA sin EHNA y EHNA.

La asignación de la muestra a uno de dos o más grupos se conoce como clasificación, y el procedimiento usado para conseguir esta asignación se conoce como clasificador o un método de clasificación. Los métodos de clasificación también pueden denominarse métodos de puntuación. Hay muchos métodos de clasificación que pueden usarse para construir un clasificador de diagnóstico a partir de un conjunto de niveles de biomarcadores. En algunos casos, los métodos de clasificación se realizan usando técnicas de aprendizaje supervisadas en las que se recoge un conjunto de datos usando muestras obtenidas de individuos dentro de dos (o más, para múltiples estados de clasificación) grupos distintos que se desean distinguir. Ya que la clase (grupo o población) a la que pertenece cada muestra se conoce previamente para cada muestra, el método de clasificación puede entrenarse para dar la respuesta de clasificación deseada. También es posible usar técnicas de aprendizaje no supervisadas para producir un clasificador de diagnóstico.

Los enfoques comunes para desarrollar clasificadores de diagnóstico incluyen árboles de decisión; envasado + refuerzo + bosques; aprendizaje basado en reglas de inferencia; Ventanas de Parzen; modelos lineales; logística; métodos de redes neuronales; agrupamiento no supervisado; medias K; ascenso/descenso jerárquico; aprendizaje semisupervisado; métodos prototipo; vecino más cercano; estimación de densidad de Kernel; máquinas de vector de soporte; modelos de Markov oculto; aprendizaje de Boltzmann; y pueden combinarse clasificadores bien de forma sencilla o de formas que reduzcan al mínimo las funciones objetivo concretas. Para una revisión, véase, por ejemplo, "Pattern Classification", R.O. Duda, *et al.*, editores, John Wiley & Sons, 2ª edición, 2001; véase también, "The Elements of Statistical Learning - Data Mining, Inference and Prediction", T. Hastie, *et al.*, editores, Springer Science+Business Media, LLC, 2ª edición, 2009.

Para producir un clasificador usando técnicas de aprendizaje supervisadas, se obtiene un conjunto de muestras denominado datos de entrenamiento. En el contexto de ensayos de diagnóstico, los datos de entrenamiento incluyen muestras de los distintos grupos (clases) a los que se asignarán posteriormente muestras desconocidas. Por ejemplo, las muestras recogidas de individuos de una población de control e individuos de una población de una enfermedad determinada pueden constituir datos de entrenamiento para desarrollar un clasificador que puede clasificar muestras desconocidas (o, más particularmente, los individuos de los que se obtuvieron las muestras) bien con la enfermedad o sin la enfermedad. El desarrollo del clasificador de los datos de entrenamiento se conoce como entrenar al clasificador. Los detalles específicos de entrenamiento de clasificadores dependen de la naturaleza de la técnica de aprendizaje supervisada. El entrenamiento de un clasificador Bayesiano independiente es un ejemplo de dicha técnica de aprendizaje supervisado (véase, por ejemplo, "Pattern Classification", R. O. Duda, *et al.*, editors, John Wiley & Sons, 2ª edición, 2001; véase también, "The Elements of Statistical Learning - Data Mining, Inference and Prediction", T. Hastie, *et al.*, editores, Springer Science+Business Media, LLC, 2ª edición, 2009). El entrenamiento de un clasificador Bayesiano independiente se describe, por ejemplo, en las publicaciones de EE.UU. n.º 2012/0101002 y 2012/0077695.

Ya que normalmente hay muchos más niveles de posibles biomarcadores que muestras en un conjunto de entrenamiento, debe tenerse cuidado para evitar el sobreajuste. El sobreajuste se produce cuando un modelo estadístico describe error aleatorio o ruido en lugar de la relación subyacente. El sobreajuste puede evitarse de varias maneras, incluyendo, por ejemplo, mediante la limitación del número de marcadores usado en el desarrollo del clasificador, suponiendo que las respuestas de marcadores sean independientes entre sí, limitando la complejidad del modelo estadístico subyacente empleado, y asegurando que el modelo estadístico subyacente se ajuste a los datos.

Un ejemplo ilustrativo del desarrollo de un ensayo de diagnóstico en el que se usa un conjunto de biomarcadores incluye la aplicación de un clasificador de Bayes independiente, un clasificador probabilístico sencillo basado en el teorema de Bayes con tratamiento independiente estricto de los biomarcadores. Cada biomarcador se describe por una función de densidad de probabilidad dependiente de la clase (pdf) para los valores de UFR medidos o los valores del log de UFR (unidades de fluorescencia relativas) en cada clase. Se supone que las pdf unidas para el conjunto de marcadores en una clase es el producto de las pdf dependientes de clase individuales para cada biomarcador. El entrenamiento de un clasificador de Bayes independiente en este contexto supone asignar parámetros ("parametrización") para caracterizar las pdf dependientes de clase. Puede usarse cualquier modelo subyacente para las pdf dependientes de clase, pero el modelo debería ajustarse en general a los datos observados en el conjunto de entrenamiento.

El rendimiento del clasificador de Bayes independiente depende del número y de la calidad de los biomarcadores

usados para construir y entrenar al clasificador. Un único biomarcador funcionará de acuerdo con su distancia KS (Kolmogorov-Smirnov). La adición de biomarcadores posteriores con buenas distancias KS ( $> 0,3$ , por ejemplo) mejorará, en general, el rendimiento de la clasificación si los biomarcadores añadidos posteriormente son independientes del primer biomarcador. Usando la sensibilidad más la especificidad como puntuación del clasificador, se pueden generar muchos clasificadores de puntuación alta con una variación de un algoritmo voraz. (Un algoritmo voraz es cualquier algoritmo que siga la metaheurística de resolución de problemas de realizar la elección localmente óptima en cada fase con la esperanza de encontrar el óptimo global).

Otra forma de representar el rendimiento del clasificador es a través de una característica operativa del receptor (ROC), o simplemente una curva ROC o una gráfica ROC. La ROC es una representación gráfica de la sensibilidad, o tasa de verdaderos positivos, frente a la tasa de falsos positivos ( $1 - \text{especificidad}$  o  $1 - \text{tasa de verdaderos negativos}$ ), para un sistema clasificador binario, ya que su umbral de discriminación varía. La ROC también se puede representar de manera equivalente mediante la representación de la fracción de verdaderos positivos de los positivos (TPR = tasa positiva real) frente a la fracción de falsos positivos de los negativos (FPR = tasa de falsos positivos). También se conoce como curva de característica operativa relativa, porque es una comparación de dos características operativas (TPR y FPR) a medida que cambia el criterio. El área bajo la curva ROC (ABC) se usa comúnmente como una medida resumida de la precisión del diagnóstico. Puede tomar valores de 0,0 a 1,0. La ABC tiene una propiedad estadística importante: el ABC de un clasificador es equivalente a la probabilidad de que el clasificador clasifique una instancia positiva escogida al azar superior a una instancia negativa escogida al azar (Fawcett T., 2006. "An introduction to ROC analysis. Pattern Recognition Letters". 27: 861-874). Es equivalente a la prueba de rangos de Wilcoxon (Hanley, J. A., McNeil, B. J., 1982. "The meaning and use of the area under a receiver operating characteristic (ROC) curve". *Radiology* 143, 29-36).

Los ejemplos descritos, pero no reivindicados, en el presente documento, usan cualquier número de biomarcadores enumerados en las Tablas 3, 4, 6, 7, 8 y 9 en diversas combinaciones para producir pruebas de diagnóstico para identificar individuos con EHNA. Los biomarcadores enumerados en las Tablas 3, 4, 6, 7, 8 y 9 se pueden combinar de muchas maneras para producir clasificadores. Los grupos de biomarcadores pueden comprender diferentes conjuntos de biomarcadores dependiendo de un criterio de rendimiento de diagnóstico específico que se seleccione. Por ejemplo, ciertas combinaciones de biomarcadores pueden producir pruebas que son más sensibles (o más específicas) que otras combinaciones. Se escoge un grupo de biomarcadores para identificar individuos con EHNA de los grupos en la Tabla 5.

Los ejemplos descritos, pero no reivindicados, en el presente documento, usan cualquier número de biomarcadores enumerados en las Tablas 3, 4, 6, 7, 8 y 9 en diversas combinaciones para producir pruebas de diagnóstico para identificar individuos con esteatosis. Los biomarcadores enumerados en las Tablas 3, 4, 6, 7, 8 y 9 se pueden combinar de muchas maneras para producir clasificadores. Los grupos de biomarcadores pueden comprender diferentes conjuntos de biomarcadores dependiendo de un criterio de rendimiento de diagnóstico específico que se seleccione. Por ejemplo, ciertas combinaciones de biomarcadores pueden producir pruebas que son más sensibles (o más específicas) que otras combinaciones. Un grupo de biomarcadores para identificar individuos con esteatosis puede seleccionarse de los grupos de la Tabla 5. Un grupo de biomarcadores para identificar individuos con esteatosis puede comprender los biomarcadores de la Tabla 3. Un grupo de biomarcadores para identificar individuos con esteatosis puede comprender los biomarcadores de la Tabla 8.

Los biomarcadores enumerados en las Tablas 3, 4, 6, 7, 8 y 9, en diversas combinaciones, pueden producir pruebas diagnósticas para identificar a individuos con EHNA. Los biomarcadores enumerados en las Tablas 3, 4, 6, 7, 8 y 9 pueden combinarse de muchas maneras para producir clasificadores. Los grupos de biomarcadores pueden estar compuestos por diferentes conjuntos de biomarcadores en función de un criterio de rendimiento diagnóstico específico que se seleccione. Por ejemplo, ciertas combinaciones de biomarcadores pueden producir pruebas más sensibles (o más específicas) que otras combinaciones. Un grupo de biomarcadores para identificar individuos con EHNA puede seleccionarse de entre los grupos de la Tabla 5. Un grupo de biomarcadores para identificar individuos con EHNA puede comprender los biomarcadores de la Tabla 4. Un grupo de biomarcadores para identificar individuos con EHNA puede comprender los biomarcadores de la Tabla 9.

Los ejemplos descritos, pero no reivindicados, en el presente documento, usan cualquier número de biomarcadores enumerados en las Tablas 3, 4, 6, 7, 8 y 9 en diversas combinaciones para producir pruebas de diagnóstico para identificar individuos con EHNA, esteatosis y/o EHNA. Los biomarcadores enumerados en las Tablas 3, 4, 6, 7, 8 y 9 se pueden combinar de muchas maneras para producir clasificadores. Los grupos de biomarcadores pueden comprender diferentes conjuntos de biomarcadores dependiendo de un criterio de rendimiento de diagnóstico específico que se selecciona. Por ejemplo, ciertas combinaciones de biomarcadores pueden producir pruebas que son más sensibles (o más específicas) que otras combinaciones. Un grupo de biomarcadores para identificar individuos con EHNA, esteatosis y/o EHNA puede seleccionarse de los grupos de la Tabla 5.

Una vez que se define un grupo para incluir un determinado conjunto de biomarcadores de las Tablas 3, 4, 6, 7, 8 y 9, y se construye un clasificador a partir de un conjunto de datos de entrenamiento, se completan los parámetros de prueba de diagnóstico. Puede procesarse una muestra biológica en uno o más ensayos para producir los niveles de biomarcadores cuantitativos relevantes usados para la clasificación. Los niveles de biomarcadores medidos se usan

como entrada para el método de clasificación que arroja una clasificación y una puntuación opcional para la muestra que refleja la confianza de la asignación de la clase.

En algunas realizaciones, opcionalmente, se diluye una muestra biológica y se procesa en un ensayo multiplexado basado en aptámeros, y los datos se evalúan de la siguiente manera. En primer lugar, los datos del ensayo se normalizan y, opcionalmente, se calibran, y los niveles de biomarcadores resultantes se usan como entrada para un esquema de clasificación de Bayes. En segundo lugar, se calcula la relación de probabilidad logarítmica para cada biomarcador medido individualmente y luego se suma para producir una puntuación de clasificación final, que también se conoce como puntuación de diagnóstico. Se puede indicar la asignación resultante, así como la puntuación de clasificación general. En algunas realizaciones, también se pueden indicar los factores de riesgo de probabilidad logarítmica individuales calculados para cada nivel de biomarcador.

### Kits

Puede detectarse cualquier combinación de los biomarcadores descritos en el presente documento usando un kit adecuado, tal como para su uso en la realización de los métodos desvelados en el presente documento. Además, cualquier kit puede contener uno o más marcadores detectables como los descritos en el presente documento, tales como una fracción fluorescente, etc.

Un kit puede incluir (a) uno o más reactivos de captura (tales como, por ejemplo, al menos un aptámero o anticuerpo) para detectar uno o más biomarcadores en una muestra biológica y, opcionalmente, (b) uno o más productos de programas informáticos o de software para predecir si el individuo del que se obtuvo la muestra biológica tiene EHGNA, esteatosis y/o EHNA (tal como EHNA en fase 1, 2, 3 o 4 o EHNA en fase 2, 3 o 4). Como alternativa, en lugar de uno o más productos de programas informáticos, puede proporcionarse una o más instrucciones para que un ser humano realice manualmente las etapas anteriores.

Un kit puede comprender un soporte sólido, un reactivo de captura y un material generador de señal. El kit también puede incluir instrucciones para usar los dispositivos y reactivos, manejar la muestra y analizar los datos. Además, el kit puede usarse con un sistema informático o software para analizar e informar el resultado del análisis de la muestra biológica.

Los kits también pueden contener uno o más reactivos (por ejemplo, tampones de solubilización, detergentes, lavados o tampones) para procesar una muestra biológica. Cualquiera de los kits descritos en el presente documento también puede incluir, por ejemplo, tampones, agentes de bloqueo, materiales de matriz de espectrometría de masas, agentes de captura de anticuerpos, muestras de control positivo, muestras de control negativo, software e información tales como protocolos, directrices y datos de referencia.

Pueden proporcionarse kits para el análisis de la EHGNA y/o EHNA, en las que los kits comprenden cebadores de PCR para uno o más biomarcadores descritos en el presente documento. Un kit puede incluir además instrucciones para su uso y la correlación de los biomarcadores con el pronóstico de la EHGNA y/o EHNA. Un kit puede incluir una matriz de ADN que contiene el complemento de uno o más de los biomarcadores descritos en el presente documento, reactivos y/o enzimas para amplificar o aislar ADN de muestras. Los kits pueden incluir reactivos para la PCR en tiempo real, por ejemplo, sondas de TaqMan y/o cebadores, y enzimas.

Por ejemplo, un kit puede comprender (a) reactivos que comprenden al menos un reactivo de captura para determinar el nivel de uno o más biomarcadores de una muestra de prueba, y opcionalmente (b) uno o más algoritmos o programas informáticos para realizar las etapas de comparar la cantidad de cada biomarcador cuantificada en la muestra de ensayo a uno o más puntos de corte predeterminados. Un algoritmo o programa informático puede asignar una puntuación para cada biomarcador cuantificado basándose en dicha comparación y pueden combinarse las puntuaciones asignadas para cada biomarcador cuantificado para obtener una puntuación total. Además, un algoritmo o programa informático puede comparar la puntuación total con una puntuación predeterminada, y usar la comparación para determinar si el individuo tiene EHGNA, esteatosis y/o EHNA. Como alternativa, en lugar de uno o más algoritmos o programas informáticos, se puede proporcionar una o más instrucciones para realizar manualmente las etapas anteriores por parte de un ser humano.

### Métodos y programas informáticos

Una vez que se selecciona un biomarcador o grupo de biomarcadores, un método para evaluar la EHGNA en un individuo puede comprender lo siguiente: 1) recoger u obtener de otro modo una muestra biológica; 2) realizar un método analítico para detectar y medir el biomarcador o los biomarcadores en el grupo de la muestra biológica; e 3) informar los resultados de los niveles de biomarcadores. Los resultados de los niveles de biomarcadores pueden informarse cualitativa y no cuantitativamente, como, por ejemplo, un diagnóstico propuesto ("EHGNA", "esteatosis", "EHNA", "EHNA en fase 2, 3 o 4", etc.) o simplemente un resultado positivo/negativo donde "positivo" y "negativo" están definidos. Un método para evaluar la EHGNA en un individuo puede comprender lo siguiente: 1) recoger u obtener de otro modo una muestra biológica; 2) realizar un método analítico para detectar y medir el biomarcador o los biomarcadores en el grupo de la muestra biológica; 3) realizar cualquier normalización o estandarización de

datos; 4) calcular cada nivel de biomarcador; y 5) informar los resultados de los niveles de biomarcadores. Los niveles de biomarcador pueden combinarse de alguna manera y se informa de un solo valor para los niveles de biomarcador combinados. En este enfoque, el valor informado puede ser un solo número determinado a partir de la suma de todos los cálculos de biomarcadores, que se compara con un valor umbral preestablecido que es una indicación de la presencia o ausencia de enfermedad. O la puntuación de diagnóstico puede ser una serie de barras que representan un valor de biomarcador y el patrón de las respuestas se puede comparar con un patrón preestablecido para la determinación de la presencia o ausencia de la enfermedad.

Al menos algunas realizaciones de los métodos descritos en el presente documento pueden implementarse con el uso de un ordenador. En la Figura 9, se muestra un ejemplo de un sistema informático 100. Con referencia a la Figura 9, el sistema 100 se muestra compuesto de elementos de hardware que están conectados eléctricamente a través del bus 108, que incluye un procesador 101, dispositivo de entrada 102, dispositivo de salida 103, dispositivo de almacenamiento 104, lector de medios de almacenamiento legible por ordenador 105a, sistema de comunicaciones 106, aceleración de procesamiento (por ejemplo, DSP o procesadores de propósito especial) 107 y memoria 109. El lector de medios de almacenamiento legible por ordenador 105a está conectado además a medios de almacenamiento 105b legibles por ordenador, representando la combinación de forma completa dispositivos de almacenamiento remotos, locales, fijos y/o extraíbles, además de medios de almacenamiento, memoria, etc. para contener de forma temporal y/o permanente información legible por ordenador, que puede incluir el dispositivo de almacenamiento 104, la memoria 109 y/o cualquier otro recurso de dicho sistema accesible 100. El sistema 100 también comprende elementos de software (que se muestran ubicados actualmente dentro de la memoria de trabajo 191) que incluyen un sistema operativo 192 y otro código 193, tales como programas, datos y similares.

Con respecto a la Figura 9, el sistema 100 tiene una gran flexibilidad y capacidad de configuración. Por lo tanto, por ejemplo, se puede usar una sola arquitectura para implementar uno o más servidores que se pueden configurar además de acuerdo con los protocolos, variaciones de protocolo, extensiones, etc., que se desean en la actualidad. Sin embargo, será evidente para los expertos en la materia que las realizaciones se pueden utilizar bien de acuerdo con requisitos de aplicación más específicos. Por ejemplo, uno o más elementos del sistema podrían implementarse como subelementos dentro de un componente del sistema 100 (por ejemplo, dentro del sistema de comunicaciones 106). También se podría utilizar hardware personalizado y/o se podrían implementar elementos particulares en el hardware, software o ambos. Además, aunque puede emplearse la conexión a otros dispositivos informáticos, tales como dispositivos de entrada/salida de red (no mostrados), debe entenderse que también se puede usar cableado, inalámbrico, módem y/u otra conexión o conexiones a otros dispositivos informáticos.

En un aspecto, el sistema puede comprender una base de datos que contenga características de biomarcadores característicos de la EHGNA y/o EHNA. Los datos de biomarcadores (o información de biomarcadores) se pueden utilizar como una entrada al ordenador para usarlos como parte de un método implementado por ordenador. Los datos de biomarcadores pueden incluir los datos como se describe en el presente documento.

En un aspecto, el sistema comprende además uno o más dispositivos para proporcionar datos de entrada al uno o más procesadores.

El sistema comprende además una memoria para almacenar un conjunto de datos de elementos de datos clasificados.

En otro aspecto, el dispositivo para proporcionar datos de entrada comprende un detector para detectar la característica del elemento de datos, por ejemplo, tal como un espectrómetro de masas o un lector de chips de genes.

El sistema puede comprender además un sistema de tratamiento de bases de datos. Las solicitudes o consultas de los usuarios pueden formatearse en un lenguaje apropiado que entienda el sistema de tratamiento de la base de datos que procesa la consulta para extraer la información relevante de la base de datos de los conjuntos de capacitación.

El sistema puede conectarse a una red a la que estén conectados un servidor de red y uno o más clientes. La red puede ser una red de área local (LAN) o una red de área amplia (WAN), como se conoce en la técnica. Preferentemente, el servidor incluye el hardware necesario para ejecutar los productos del programa informático (por ejemplo, el software) para acceder a los datos de la base de datos para procesar las solicitudes del usuario.

El sistema puede incluir un sistema operativo (por ejemplo, UNIX® o Linux) para ejecutar instrucciones desde un sistema de tratamiento de bases de datos. En un aspecto, el sistema operativo puede operar en una red de comunicaciones global, tal como Internet, y utilizar un servidor de red de comunicaciones global para conectarse a dicha red.

El sistema puede incluir uno o más dispositivos que comprendan una interfaz de visualización gráfica que comprenda elementos de interfaz tales como botones, menús desplegados, barras de desplazamiento, campos para introducir texto y similares como se encuentran habitualmente en interfaces de usuario gráficas conocidas en la

técnica. Las peticiones introducidas en una interfaz de usuario pueden transmitirse a un programa de aplicación en el sistema para formatear con el fin de buscar información relevante en una o más de las bases de datos del sistema. Las peticiones o consultas introducidas por un usuario pueden construirse en cualquier lenguaje de base de datos adecuado.

La interfaz gráfica de usuario puede generarse por un código de interfaz gráfica de usuario como parte del sistema operativo y puede usarse para introducir datos y/o para presentar datos introducidos. El resultado de los datos procesados puede presentarse en la interfaz, imprimirse en una impresora en comunicación con el sistema, guardarse en un dispositivo de memoria y/o transmitirse por la red o puede proporcionarse en forma del medio legible por ordenador.

El sistema puede estar en comunicación con un dispositivo de entrada para proporcionar datos relativos a elementos de datos al sistema (por ejemplo, niveles de expresión). En un aspecto, el dispositivo de entrada puede incluir un sistema de perfil de expresión génica incluyendo, por ejemplo, un espectrómetro de masas, una microplaca génica o lector de matrices, y similares.

Los métodos y aparatos para analizar la información de biomarcadores de acuerdo con diversas realizaciones pueden implementarse de cualquier manera adecuada, por ejemplo, usando un programa informático que actúe en un sistema informático. Puede usarse un sistema informático convencional que comprenda un procesador y una memoria de acceso aleatorio, tal como un servidor de aplicación accesible de forma remota, servidor de red, ordenador personal o estación de trabajo. Los componentes de sistemas informáticos adicionales pueden incluir dispositivos de memoria o sistemas de almacenamiento de información, tales como un sistema de almacenamiento en masa y una interfaz de usuario, por ejemplo, un monitor convencional, teclado y dispositivo de seguimiento. El sistema informático puede ser un sistema independiente o parte de una red de ordenadores que incluya un servidor y una o más bases de datos.

El sistema de análisis de biomarcadores puede proporcionar funciones y operaciones para completar el análisis de datos, tales como la recogida de datos, el procesamiento, el análisis, la presentación y/o el diagnóstico. Por ejemplo, en una realización, el sistema informático puede ejecutar el programa informático que puede recibir, almacenar, buscar, analizar y presentar información relativa a los biomarcadores. El programa informático puede comprender múltiples módulos que realizan diversas funciones u operaciones, tales como un módulo de procesamiento para procesar datos en bruto y generar datos complementarios, y un módulo de análisis para analizar datos en bruto y datos complementarios para generar un estado y/o diagnóstico de la enfermedad. La identificación de la EHNA, esteatosis y/o EHNA puede comprender la generación o recopilación de cualquier otra información, incluyendo información biomédica adicional, con respecto al estado del individuo en relación con la enfermedad, la identificación de si se desean ensayos adicionales o la evaluación de otro modo del estado de salud del individuo.

Algunos métodos descritos en el presente documento pueden implementarse para incluir un producto de programa informático. Un producto de programa informático puede incluir un medio legible por ordenador que tenga un código de programa legible por ordenador incorporado en el medio para hacer que un programa de aplicación se ejecute en un ordenador con una base de datos.

Como se usa en el presente documento, un "producto de programa informático" se refiere a un conjunto organizado de instrucciones en forma de afirmaciones en lenguaje natural o de programación que están contenidas en un medio físico de cualquier naturaleza (por ejemplo, escrito, electrónico, magnético, óptico u otro), y que se puede usar con un ordenador u otro sistema automatizado de procesamiento de datos. Dichas afirmaciones en lenguaje de programación, cuando son ejecutadas por un ordenador o sistema de procesamiento de datos, hacen que el ordenador o el sistema de procesamiento de datos actúen de acuerdo con el contenido particular de las afirmaciones. Los productos de programas informáticos incluyen, sin limitación: programas en código fuente y objeto y/o bibliotecas de prueba o de datos integradas en un medio legible por ordenador. Además, el producto de programa informático que permite que un sistema informático o un dispositivo de procesamiento de datos actúe de formas preseleccionadas puede proporcionarse de una serie de formas que incluyen, pero sin limitación, código fuente original, código de ensamblaje, código de objeto, lenguaje de máquina, versiones encriptadas o comprimidas de lo anterior y todos los equivalentes.

En un aspecto, en el presente documento se describe, pero no se reivindica, un producto de programa informático para indicar si un individuo tiene EHNA, si un individuo tiene esteatosis y/o si un individuo tiene EHNA (como EHNA en fase 1, 2, 3 o 4, o EHNA en fase 2, 3 o 4). El producto de programa informático incluye un código legible por ordenador que incorpora código de programa ejecutable por un procesador de un dispositivo o sistema informático, comprendiendo el código de programa: código que recupera datos atribuidos a una muestra biológica de un individuo, donde los datos comprenden niveles de biomarcadores que corresponden a uno o más de los biomarcadores descritos en el presente documento, y un código que ejecuta un método de clasificación que indica el estado de la EHNA, esteatosis y/o EHNA del individuo en función de los niveles de biomarcadores.

Aunque se han descrito diversas realizaciones como métodos o aparatos, debe entenderse que los métodos pueden implementarse a través de un código acoplado con un ordenador, por ejemplo, un código residente en un ordenador

o accesible por el ordenador. Por ejemplo, se podrían utilizar software y bases de datos para implementar muchos de los métodos mencionados anteriormente. Por tanto, además de los métodos realizados por hardware, también se observa que estos métodos se pueden realizar mediante el uso de un artículo de fabricación compuesto de un medio que se pueda usar por ordenador que tenga un código de programa legible por ordenador incorporado en el mismo, lo que provoca la habilitación de las funciones desveladas en la presente descripción. Además, los métodos pueden incorporarse como código almacenado en una memoria legible por ordenador de prácticamente cualquier tipo que incluye, sin limitación, RAM, ROM, medios magnéticos, medios ópticos o medios magneto-ópticos. Incluso más en general, los métodos podrían implementarse en software, o en hardware, o cualquier combinación de los mismos, incluyendo, pero sin limitación, software que se ejecuta en un procesador de propósito general, microcódigo, arreglos lógicos programables (PLA) o circuitos integrados específicos de la aplicación (ASIC).

También se prevé que los métodos podrían realizarse como señales de ordenador incorporadas en una onda portadora, así como también señales (por ejemplo, eléctricas y ópticas) propagadas a través de un medio de transmisión. Por lo tanto, los diversos tipos de información tratados anteriormente podrían formatearse en una estructura, tal como una estructura de datos, y transmitirse como una señal eléctrica a través de un medio de transmisión o almacenarse en un medio legible por ordenador.

### Métodos de tratamiento

En algunas realizaciones, tras la determinación de que un sujeto tiene EHGNA, esteatosis o EHNA, el sujeto se somete a un régimen terapéutico para retrasar o prevenir el empeoramiento de la enfermedad. Los regímenes terapéuticos ilustrativos no limitantes para la EHGNA, esteatosis y/o EHNA incluyen pérdida de peso, control del azúcar en sangre y abstinencia del alcohol. En algunas realizaciones, a un sujeto se le administra un agente terapéutico tal como pioglitazona, vitamina E y/o metformina. Véase, por ejemplo, Sanyal *et al.*, 2010, NEJM, 362: 1675-1685. En algunas realizaciones, un sujeto se somete a cirugía de derivación gástrica (o similar), por ejemplo, para acelerar la pérdida de peso.

En algunas realizaciones, se proporcionan métodos de control de la EHGNA. En algunas realizaciones, los presentes métodos de determinación de si un sujeto tiene EHGNA se llevan a cabo en un punto temporal 0. En algunas realizaciones, el método se lleva a cabo de nuevo en un punto temporal 1, y opcionalmente, un punto temporal 2, y opcionalmente, un punto temporal 3, etc., con el fin de controlar la progresión de la EHGNA en el sujeto. En algunas realizaciones, se usan biomarcadores diferentes en diferentes puntos temporales, dependiendo del estado actual de la enfermedad del individuo y/o dependiendo de la velocidad a la que se cree o se predice que progresará la enfermedad.

### Otros métodos

En algunas realizaciones, los biomarcadores y los métodos descritos en el presente documento se usan para determinar una prima de seguro médico y/o una prima de seguro de vida. En algunas realizaciones, los resultados de los métodos descritos en el presente documento se usan para determinar una prima de seguro médico y/o una prima de seguro de vida. En algunos casos, una organización que proporciona un seguro médico o seguro de vida solicita u obtiene información de otro modo sobre el estado de la EHGNA o EHNA de un sujeto, y usa esa información para determinar una prima de seguro médico o de seguro de vida apropiada para el sujeto. En algunas realizaciones, la prueba es solicitada y pagada por la organización que proporciona el seguro médico o seguro de vida.

En algunas realizaciones, los biomarcadores y métodos descritos en el presente documento se usan para predecir y/o gestionar la utilización de recursos médicos. En algunas de dichas realizaciones, los métodos no se llevan a cabo con el fin de dicha predicción, pero la información obtenida del método se usa en dicha predicción y/o gestión de la utilización de recursos médicos. Por ejemplo, un centro de pruebas u hospital puede recopilar información de los presentes métodos para muchos temas con el fin de predecir y/o gestionar la utilización de recursos médicos en un determinado centro o en una zona geográfica en particular.

### Ejemplos

Los siguientes ejemplos se proporcionan solo con fines ilustrativos y no pretenden limitar el alcance de la solicitud tal como se define en las reivindicaciones adjuntas. Las técnicas de biología molecular habituales descritas en los siguientes ejemplos se pueden llevar a cabo como se describe en manuales de laboratorio convencionales, tales como Sambrook *et al.*, "Molecular Cloning: A Laboratory Manual", 3ª. ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., (2001).

#### Ejemplo 1. Sujetos de estudio de la EHGNA

Las muestras usadas para identificar biomarcadores procedían de Geisinger Health. Se tomaron muestras de suero y se realizaron biopsias hepáticas en 576 pacientes obesos antes de someterles a cirugía bariátrica para perder peso.

Las muestras se recogieron en tubos de suero rojos superiores y se procesaron según el protocolo; en resumen, se dejó coagular una muestra durante 30 minutos a temperatura ambiente, y luego se centrifugó a 1300 xg durante 10 minutos y se eliminó la capa superior, y se almacenó a -80 °C. Las muestras se descongelaron una vez para preparar alícuotas y una vez para el ensayo.

Para identificar biomarcadores que distingan sujetos con EHNA de sujetos obesos normales, y biomarcadores que distinguirán sujetos con EHNA de sujetos con esteatosis en el hígado, se dividieron los sujetos del estudio en tres niveles normales de esteatosis hepática (esteatosis leve, moderada y grave) y cuatro fases de EHNA según los resultados de la biopsia hepática, usando el método de clasificación de Brunt (Brunt *et al.*, 2007, *Modern Pathol.*, 20: S40- S48). Los grupos se subclasificaron como se muestra en la Tabla 1.

Tabla 1: Subclasificación de los grupos de fases de esteatosis y EHNA

| Grupo                      | Esteatosis | Inflamación | Abombamiento | Fibrosis |
|----------------------------|------------|-------------|--------------|----------|
| Controles de obesos        | 0          | 0           | 0            | 0        |
| Esteatosis leve            | 1          | 0 o 1       | 0            | 0        |
| Esteatosis moderada        | 2          | 0, 1 o 2    | 0            | 0        |
| Esteatosis grave           | 3          | 0, 1 o 2    | 0            | 0        |
| EHNA en fase 1             | 1, 2 o 3   | 0, 1 o 2    | 1            | 0 o 1    |
| EHNA en fase 2             | 1, 2 o 3   | 0, 1 o 2    | 2            | 2        |
| EHNA en fase 3 (en puente) | 1, 2 o 3   | 0, 1 o 2    | 2            | 3        |
| EHNA en fase 4 (cirrosis)  | 1, 2 o 3   | 0, 1 o 2    | 2            | 4        |

#### 15 Datos demográficos de los sujetos

En la Tabla 2, se muestran ciertas características para los individuos de cada uno de los grupos analizados anteriormente.

Tabla 2: Datos demográficos de los sujetos

| Grupo               | n.º de mujeres | n.º de varones | Total | n.º con ocultación | Edad en la biopsia, edad media | IMC, media kg/m² | LDL, media mg/dl | % con diabetes |
|---------------------|----------------|----------------|-------|--------------------|--------------------------------|------------------|------------------|----------------|
| Control             | 111            | 14             | 125   | 41                 | 44,59                          | 46,8             | 107,86           | 28 (22,4 %)    |
| Esteatosis leve     | 44             | 7              | 51    | 16                 | 45,75                          | 46,6             | 112,12           | 15 (29,4 %)    |
| Esteatosis moderada | 25             | 7              | 32    | 10                 | 44,5                           | 44,77            | 103,29           | 12 (37,5 %)    |
| Esteatosis grave    | 45             | 9              | 54    | 17                 | 47,11                          | 47,75            | 101,11           | 23 (42,6 %)    |
| EHNA Fase 1         | 93             | 20             | 113   | 37                 | 46,69                          | 47,29            | 103,37           | 58 (51,3 %)    |
| EHNA Fase 2         | 27             | 11             | 38    | 12                 | 48,52                          | 48,9             | 109,74           | 22 (57,9 %)    |
| EHNA Fase 3         | 12             | 8              | 20    | 0                  | 50,05                          | 46,27            | 101,06           | 17(85 %)       |
| EHNA Fase 4         | 6              | 4              | 10    | 0                  | 48,7                           | 47,22            | 101,9            | 5 (50 %)       |
| Total               | 363            | 80             | 443   |                    |                                |                  |                  |                |

Como se muestra en la Tabla 2, se determinaron la edad, el índice de masa corporal y los niveles de LDL para los sujetos, y se encontró que estaban equilibrados en todos los grupos.

**Ejemplo 2. Ensayo multiplexado basado en aptámeros para la identificación de biomarcadores**

Se evaluó la calidad de las muestras normales, EHNA y EHNA de los grupos mencionados anteriormente mediante la comparación de las distribuciones de biomarcadores asociados con la manipulación de las muestras, como la cizalla, lisis celular y activación del complemento en los casos y controles. La calidad de las muestras fue buena y no hubo sesgo de control de los casos.

Se usó un ensayo multiplexado basado en aptámeros para analizar las muestras y los controles para identificar biomarcadores predictivos de la EHNA y EHNA. El análisis multiplexado usado en este experimento incluía aptámeros para detectar 1.129 proteínas en sangre a partir de pequeños volúmenes de muestra (~65 µl de suero o plasma), con límites de detección bajos (mediana de 1 pM), ~7 registros de rango dinámico y ~5 % de la mediana del coeficiente de variación. El ensayo multiplexado basado en aptámeros se describe, por ejemplo, en Gold *et al.* (2010) "Aptamer-Based Multiplexed Proteomic Technology for Biomarker Discovery". PLoS ONE 5(12): e15004; y publicaciones de EE.UU. n.º 2012/0101002 y 2012/0077695.

La selección de la estabilidad toma muchos subconjuntos de la mitad de los datos y realiza la selección de biomarcadores usando el clasificador de lazo, que es un modelo de regresión logística regularizado. Véase, por ejemplo, Meinshausen *et al.*, 2010, J. Royal Statistical Soc: Series B (Statistical Methodology), 72: 417-473. La vía de selección para un solo biomarcador es la proporción de estos subconjuntos para los que ese biomarcador fue seleccionado por el modelo de lazo en un intervalo de lambda. Lambda es un parámetro de ajuste que determina la cantidad de biomarcadores seleccionados por el lazo. La máxima probabilidad de selección en un intervalo de valores lambda es la medida final usada para seleccionar un conjunto de biomarcadores.

Los biomarcadores candidatos se identificaron mediante selección de la estabilidad, que luego se usaron para generar el modelo clasificador de bosque aleatorio. Véase, por ejemplo, Shi *et al.*, J. Comput. Graph. Stat. 15(1): 118-138 (2006). En resumen, un predictor de bosque aleatorio es un conjunto de predictores de árbol de clasificación individual. Véase, por ejemplo, Breiman, *Machine Learning*, 45(1): 5-32 (2001). Para cada observación, cada árbol individual vota por una clase, y el bosque predice la clase que tiene la pluralidad de votos. El usuario especifica el número de variables seleccionadas al azar (mtry) para buscar la mejor división en cada nodo. El índice de Gini se usa como el criterio de división. Véase, por ejemplo, Breiman *et al.*, "Classification and Regression Trees", Chapman y Hall, Nueva York, 1984. El árbol más grande posible se cultiva y no se poda. El nodo raíz de cada árbol del bosque contiene una muestra de arranque de los datos originales como el conjunto de entrenamiento. Las observaciones que no están en el conjunto de entrenamiento, aproximadamente 1/3 del conjunto de datos original, se conocen como observaciones fuera de la bolsa (OOB). Se puede llegar a las predicciones de OOB de la siguiente manera: para un caso en los datos originales, se predice el resultado por voto de pluralidad que implique solo a aquellos árboles que no contengan el caso en su correspondiente muestra de arranque. Al contrastar estas predicciones OOB con los resultados del conjunto de entrenamiento, se puede llegar a una estimación de la tasa de error de la predicción, que se conoce como la tasa de error OOB.

El análisis univariado se realizó usando la prueba no paramétrica de Kolmogorov-Smirnov (estadísticas KS), que cuantifica la distancia entre la función de distribución acumulada de cada aptámero para dos distribuciones de referencia designadas caso (esteatosis leve, moderada y grave y/o EHNA 1-4) y control (obeso normal). El rendimiento del clasificador de bosque aleatorio depende del número y de la calidad de los biomarcadores usados para construir y entrenar al clasificador. Un solo biomarcador funcionará de acuerdo con su distancia KS y su valor de PCA (análisis del componente principal) como se ilustra en el presente documento. Si una métrica de rendimiento de clasificador se define como la suma de la sensibilidad (fracción de verdaderos positivos,  $f_{TP}$ ) y especificidad (uno menos la fracción de falsos positivos,  $1-f_{FP}$ ), un clasificador perfecto tendrá una puntuación de dos y un clasificador aleatorio, como media, tendrá una puntuación de uno. Usando la definición de la distancia KS, ese nivel  $x^*$  que maximiza la diferencia en las funciones cdf (función de distribución acumulada) puede encontrarse resolviendo:

$$\frac{\partial KS}{\partial x} = \frac{\partial(cdf_c(x) - cdf_d(x))}{\partial x} = 0$$

para la que  $x$  que conduce a  $p(x^* | c) = p(x^* | d)$ , es decir, la distancia KS aparece cuando las pdf (funciones de densidad de la probabilidad) dependientes de clase se cruzan. La sustitución del nivel de  $x^*$  en la expresión para la distancia KS produce la siguiente definición para KS:

$$\begin{aligned} KS &= cdf_c(x^*) - cdf_d(x^*) \\ &= \int_{-\infty}^{x^*} p(x|c)dx - \int_{-\infty}^{x^*} p(x|d)dx \end{aligned}$$

$$\begin{aligned}
 &= 1 - \int_{x^*}^{\infty} p(x|c)dx - \int_{-\infty}^{x^*} p(x|d)dx \\
 &= 1 - f_{FP} - f_{FN},
 \end{aligned}$$

la distancia KS es uno menos la fracción total de errores usando un ensayo con un punto de corte en  $x^*$ , esencialmente un clasificador Bayesiano de un solo analito. Dado que los inventores definen una puntuación de *sensibilidad + especificidad* = 2 -  $f_{FP}$  -  $f_{FN}$ , combinando la definición anterior de la distancia KS se ve que *sensibilidad + especificidad* = 1 + KS. Se seleccionan biomarcadores con una estadística que es inherentemente adecuada para construir clasificadores.

La adición de biomarcadores posteriores con buenas distancias KS (> 0,3, por ejemplo) mejorará, en general, el rendimiento de la clasificación si los biomarcadores añadidos posteriormente son independientes del primer biomarcador. Usando la sensibilidad más la especificidad como una puntuación de clasificador, se pueden generar muchos clasificadores de alta puntuación.

#### A. Clasificador de la esteatosis

Basándose en las clasificaciones de los sujetos, se supone que todos los grupos de esteatosis, así como de EHNA en fase 1-4, tienen grasa en las células hepáticas. Se desarrolló un clasificador (esteatosis o grasa en el hígado) al comparar los sujetos normales obesos con todos los sujetos con EHNA.

Se suministraron los marcadores escogidos por selección de estabilidad (véase la Figura 1) a un algoritmo de árbol aleatorio para generar un modelo. La curva ROC resultante se proporciona (véase a continuación).

En la Figura 2, se muestra una curva ROC para un clasificador de nueve marcadores para la EHNA (esteatosis). El área bajo la curva (ABC) fue de  $0,90 \pm 0,03$ . La sensibilidad fue del 92 % y la especificidad fue del 63 %, con un punto de corte de 0,5.

Se representó la puntuación de probabilidad, es decir, Prob(Esteatosis), del modelo de cada clasificador para cada individuo en todos los grupos para evaluar si se podría usar como un modelo de gravedad/control además de la decisión binaria sobre la que se construyó (Figura 3). El diagrama muestra una clara diferenciación entre ausencia de esteatosis y esteatosis, y el voto de probabilidad aumenta con el nivel de esteatosis. Los sujetos con EHNA en todas las fases tienen esteatosis grave.

La Figura 4 muestra las funciones de distribución acumulativa (CDF) para los 9 biomarcadores en el clasificador.

La Tabla 3 muestra los biomarcadores en el clasificador de 9 marcadores. La Tabla 3 también proporciona un nombre alternativo para ciertos biomarcadores, el nombre del gen y el número de acceso de UniProt para cada biomarcador, y si el biomarcador está presente a niveles superiores o inferiores en la población con EHNA, en comparación con la población normal.

Tabla 3: Clasificador de nueve biomarcadores para EHNA

| Biomarcador/Aliasas                                                                             | Nombre del gen | UniProt | Nivel de biomarcador superior/inferior en toda la población de EHNA |
|-------------------------------------------------------------------------------------------------|----------------|---------|---------------------------------------------------------------------|
| Aminoacilasa-1                                                                                  | ACY1           | Q03154  | Superior                                                            |
| Globulina de unión a hormonas sexuales                                                          | SHBG           | P04278  | Inferior                                                            |
| Catepsina Z, Catepsina P, Catepsina X                                                           | CTSZ           | Q9UBR2  | Superior                                                            |
| c-met, receptor de factor de crecimiento de hepatocitos, tirosina quinasa del proto-oncogén Met | MET            | P08581  | Inferior                                                            |
| Gelsolina                                                                                       | GSN            | P06396  | Inferior                                                            |
| Proteína de unión a galectina-3, proteína de unión 3 soluble de unión a lectina galactosidasa   | LGALS3BP       | Q08380  | Superior                                                            |

(continuación)

| Biomarcador/Aliasas                                           | Nombre del gen | UniProt | Nivel de biomarcador superior/inferior en toda la población de EHGNA |
|---------------------------------------------------------------|----------------|---------|----------------------------------------------------------------------|
| Activador de plasminógeno de tipo tisular, tPA                | PLAT           | P00750  | Superior                                                             |
| Proteína de tipo molécula de adhesión a células neuronales L1 | CHL1           | O00533  | Inferior                                                             |
| Antitrombina III                                              | SERPINC1       | P01008  | Inferior                                                             |

La Figura 3 muestra diagramas de cajas para el clasificador de nueve biomarcadores en cada uno de los grupos de sujetos (de izquierda a derecha: normal, esteatosis leve, esteatosis moderada, esteatosis grave, EHNA1, EHNA2, EHNA3, EHNA4). La línea negra de dentro de cada caja representa la mediana (o percentil 50) de los puntos de datos, y la propia caja representa el intervalo intercuartílico (IQR), el área que abarca los puntos de datos del percentil 25 al 75. Los filamentos se extienden para cubrir puntos de datos dentro de 1.5 x IQR de la parte superior e inferior de la caja.

#### B. Clasificador de la EHNA (Fibrosis)

Todos los sujetos con EHNA tienen alguna forma de inflamación y abombamiento asociados con la fibrosis. Por lo tanto, se compararon todos los grupos de esteatosis con la EHNA en fase 2, 3 y 4. Para garantizar la identificación de los biomarcadores de fibrosis verdaderos, se excluyó el grupo de EHNA en fase 1.

Se suministraron los marcadores escogidos por selección de estabilidad (Figura 5) a un algoritmo de árbol aleatorio para generar un modelo. La ROC resultante se proporciona a continuación.

En la Figura 6, se muestra una curva ROC para el clasificador de cuatro marcadores para EHNA en fase 2, 3 y 4 (fibrosis). El área bajo la curva (ABC) fue  $0,82 \pm 0,07$ , con una sensibilidad del 62 % y una especificidad del 92 % en un punto de corte de 0,5.

Se representó la puntuación de probabilidad, es decir, Prob(Esteatosis), del modelo de cada clasificador para cada individuo en todos los grupos para evaluar si se podría usar como un modelo de gravedad/control además de la decisión binaria sobre la que se construyó (Figura 7). El diagrama muestra una clara diferenciación entre ausencia de esteatosis y esteatosis, y el voto de probabilidad aumenta con el nivel de esteatosis. Los sujetos con EHNA en todas las fases tienen esteatosis grave.

La Tabla 4 muestra los biomarcadores en el clasificador de 4 marcadores. La Tabla 4 también proporciona un nombre alternativo para ciertos biomarcadores, el nombre del gen y el número de acceso de UniProt para cada biomarcador, y si el biomarcador está presente a niveles superiores o inferiores en las poblaciones con EHNA en fase 2, 3 y 4, en comparación con todas las poblaciones de EHGNA.

Tabla 4: Clasificador de cuatro biomarcadores para EHNA en fases 2, 3 y 4 frente a la esteatosis (EHGNA)

| Biomarcador/Aliasas                                           | Nombre del gen | UniProt | Nivel de biomarcador superior/inferior en la población con fibrosis (EHNA) |
|---------------------------------------------------------------|----------------|---------|----------------------------------------------------------------------------|
| Complemento C7                                                | C7             | P10643  | Superior                                                                   |
| Colectina renal 1                                             | COLEC11        | Q9BWP8  | Superior                                                                   |
| Peptidilpropilo isomerasa D                                   | PPID           | Q08752  | Inferior                                                                   |
| Proteína de unión al factor de crecimiento de tipo insulina 3 | IGFBP3         | P17936  | Inferior                                                                   |

La Figura 7 muestra diagramas de cajas para el clasificador de cuatro biomarcadores en cada uno de los grupos de sujetos (de izquierda a derecha: normal, esteatosis leve, esteatosis moderada, esteatosis grave, EHNA1, EHNA2, EHNA3, EHNA4). La línea negra dentro de cada cuadro representa la mediana (o percentil 50) de los puntos de datos, y el propio recuadro representa el intervalo intercuartílico (IQR), el área que abarca los puntos de datos del percentil 25 al 75. Los filamentos se extienden para cubrir puntos de datos dentro de IQR x 1,5 de la parte superior e inferior de la caja.

La Figura 8 muestra las CDF para los 4 biomarcadores en el clasificador.

#### Ejemplo 3. Biomarcadores y clasificadores adicionales para EHGNA y/o EHNA

La selección de la estabilidad toma muchos subconjuntos de la mitad de los datos y realiza la selección de biomarcadores usando el clasificador de lazo, que es un modelo de regresión logística regularizado. Véase, por ejemplo, Meinshausen *et al.*, 2010, *J. Royal Statistical Soc: Series B* (Statistical Methodology), 72: 417-473. La vía de selección para un solo biomarcador es la proporción de estos subconjuntos para los que ese biomarcador fue seleccionado por el modelo de lazo en un intervalo de  $\lambda$ .  $\lambda$  es un parámetro de ajuste que determina la cantidad de biomarcadores seleccionados por el lazo. La máxima probabilidad de selección en un intervalo de valores  $\lambda$  es la medida final usada para seleccionar un conjunto de biomarcadores.

Usando el método de selección de la estabilidad, se definieron clasificadores adicionales para distinguir diversos grupos de individuos. Los clasificadores (incluyendo los clasificadores analizados anteriormente) se muestran en la Tabla 5. Los marcadores de las comparaciones 2 y 5 se usaron para construir un clasificador de árbol aleatorio para la esteatosis (EHNA) y la fibrosis (EHNA), como se ha analizado anteriormente.

Tabla 5. Clasificadores obtenidos usando la selección de la estabilidad

|                                                           | Comparación                                                | Sensibilidad | Especificidad | Marcadores                                                |
|-----------------------------------------------------------|------------------------------------------------------------|--------------|---------------|-----------------------------------------------------------|
| 1                                                         | Control frente a EHNA en fase 1 a 4                        | 0,8641304    | 0,8196721     | ACY, SHBG, LGALS3BP, SIGLEC7, CTSZ, MET, GSN              |
| 2                                                         | Control frente a (todas las esteatosis + EHNA en fase 1-4) | 0,8635015    | 0,7452830     | ACY, SHBG, CTSZ, MET, GSN, LGALS3BP, PLAT, CHL1, SERPINC1 |
| 3                                                         | Control frente a todas las esteatosis                      | 0,7664234    | 0,7440000     | ACY, SHBG, SIGLEC14                                       |
| 4                                                         | Todas las esteatosis frente a EHNA 1                       | 0,4901961    | 0,5743243     | TOP1, SIGLEC14                                            |
| 5                                                         | Todas las esteatosis frente a EHNA en fase 2-4             | 0,7543860    | 0,8310811     | C7, COLEC11, PPID, IGFBP3                                 |
| 6                                                         | Todas las esteatosis frente a EHNA en fase 1-4             | 0,6734694    | 0,5983607     | SIGLEC14, AIMP1, TOP1, COLEC11, CA6, STX1A                |
| *Todas las esteatosis: esteatosis leve, moderada y grave. |                                                            |              |               |                                                           |

La comparación 1 de la Tabla 5 muestra un clasificador de 7 marcadores que distingue a los sujetos de control de los que tienen EHNA en fase 1 a 4 con un 86,4 % de sensibilidad y un 82 % de especificidad. La comparación 3 muestra un clasificador de 3 marcadores que distingue a los sujetos de control de todas las esteatosis (leve, moderada y grave) con 76,6 % de sensibilidad y 74,4 % de especificidad.

En la Tabla 6, se muestra información adicional sobre los biomarcadores enumerados en la Tabla 5 que no están en las Tablas 3 y 4 anteriores.

Tabla 6: Biomarcadores adicionales para EHNA y/o EHNA

| Biomarcador/Alias                                              | Nombre del gen | UniProt | Nivel de biomarcador superior/inferior en EHNA frente a esteatosis frente a los controles de obesos |
|----------------------------------------------------------------|----------------|---------|-----------------------------------------------------------------------------------------------------|
| Lectina 7 de tipo Ig de unión al ácido siálico                 | SIGLEC 7       | Q9Y286  | Superior                                                                                            |
| Siglec-14                                                      | SIGLEC14       | Q08ET2  | Superior                                                                                            |
| Topoisomerasa I                                                | TOP1           | P11387  | Inferior                                                                                            |
| Polipéptido de activación de monocitos endoteliales 2 (EMAP-2) | AIMP1          | Q12904  | Inferior                                                                                            |
| Anhidrasa carbónica 6                                          | CA6            | Q16623  | Inferior                                                                                            |
| Sintaxina 1 <sup>a</sup>                                       | STX1A          | Q9Y286  | Inferior                                                                                            |

En la Tabla 7, se muestra los 25 biomarcadores principales por distancia KS univariante para el grupo de control frente a EHNA en fase 1 a 4. Estos biomarcadores y las combinaciones de estos biomarcadores se pueden usar para separar sujetos de control (tales como sujetos obesos) de sujetos con EHNA y/o para separar los sujetos de control de los sujetos con esteatosis.

Tabla 7: 25 biomarcadores principales

| Biomarcador        | Distancia KS con signo | UniProt                    | Nivel de biomarcador superior/inferior en EHNA frente a los controles de obesos |
|--------------------|------------------------|----------------------------|---------------------------------------------------------------------------------|
| ACY1               | 0,61843                | Q03154                     | Superior                                                                        |
| THBS2              | 0,48071                | P35442                     | Superior                                                                        |
| LGALS3BP           | 0,44946                | Q08380                     | Superior                                                                        |
| KYNU               | 0,44146                | Q16719                     | Superior                                                                        |
| COLEC11            | 0,43708                | Q9BWP8                     | Superior                                                                        |
| CTS2               | 0,42166                | Q9UBR2                     | Superior                                                                        |
| IL19               | 0,41688                | Q9UHD0                     | Superior                                                                        |
| POR                | 0,41021                | P16435                     | Superior                                                                        |
| INS                | 0,40888                | P01308                     | Superior                                                                        |
| SHBG               | -0,39253               | P04278                     | Inferior                                                                        |
| GPT                | 0,37459                | Q9H3H5                     | Superior                                                                        |
| GNS                | 0,37136                | P15586                     | Superior                                                                        |
| RET                | 0,37043                | P07949                     | Superior                                                                        |
| AFM                | 0,36986                | P43652                     | Superior                                                                        |
| SELE               | 0,36239                | P16581                     | Superior                                                                        |
| CD163              | 0,36146                | Q86VB7                     | Superior                                                                        |
| ENPP7              | 0,35518                | Q6UWV6                     | Superior                                                                        |
| IGFBP7             | 0,35063                | Q16270                     | Superior                                                                        |
| GSN                | -0,34851               | P06396                     | Inferior                                                                        |
| SIGLEC7            | 0,34564                | Q9Y286                     | Superior                                                                        |
| LAMA1,LAMB 1,LAMC1 | 0,33401                | P25391<br>P07942<br>P11047 | Superior                                                                        |
| IL18R1             | 0,31973                | Q13478                     | Superior                                                                        |
| FN1                | 0,31898                | P02751                     | Superior                                                                        |
| TGFBI              | 0,31823                | Q15582                     | Superior                                                                        |
| AKR1A1             | 0,31421                | P14550                     | Superior                                                                        |

**Ejemplo 4: Detección de biomarcadores ilustrativa con el uso de aptámeros**

- 5 Un método ilustrativo de detección de uno o más biomarcadores en una muestra se describe, por ejemplo, en Kraemer *et al.*, PLoS Uno 6(10): e26332, y se describe a continuación. Se describen tres métodos diferentes de cuantificación: hibridación basada en micromatrices, un método basado en perlas Luminex y qPCR.

**Reactivos**

- 10 HEPES, NaCl, KCl, EDTA, EGTA, MgCl<sub>2</sub> y Tween-20 pueden adquirirse, por ejemplo, en Fisher Biosciences. La sal sódica de sulfato de dextrano (DxSO<sub>4</sub>), nominalmente de peso molecular de 8.000, se puede adquirir, por ejemplo, en AIC, y se somete a diálisis frente a agua desionizada durante al menos 20 horas con un intercambio. La ADN polimerasa KOD EX se puede adquirir, por ejemplo, en VWR. El cloruro de tetrametilamonio y CAPSO se pueden
- 15 adquirir, por ejemplo, en Sigma-Aldrich, y la estreptavidina-ficoeritrina (SAPE) se puede adquirir, por ejemplo, en Moss Inc. El clorhidrato de 4-(2-aminoetil)-bencenosulfonilfluoruro (AEBSF) se puede adquirir, por ejemplo, en Gold Biotechnology. Las placas de 96 pocillos recubiertas con estreptavidina se pueden adquirir, por ejemplo, en Thermo Scientific (placas recubiertas con estreptavidina HBC de Pierce, transparentes, de 96 pocillos, número de producto 15500 o 15501). NHS-PEO4-biotina se puede adquirir, por ejemplo, en Thermo Scientific (EZ-Link HS-PEO4-Biotina,

número de producto 21329), disuelto en DMSO anhidro, y se puede almacenar congelado en alícuotas de un solo uso. IL-8, MIP-4, Lipocalina-2, RANTES, MMP-7 y MMP-9 se pueden adquirir, por ejemplo, en R&D Systems. Resistina y MCP-1 se pueden adquirir, por ejemplo, en PeproTech, y tPA se puede adquirir, por ejemplo, en VWR.

## 5 Ácidos nucleicos

Los oligodesoxinucleótidos convencionales (incluyendo los sustituidos con amino y biotina) se pueden adquirir, por ejemplo, en Integrated DNA Technologies (IDT). El bloque Z es un oligodesoxinucleótido monocatenario de secuencia 5'- (AC-BnBn)7-AC-3', en el que Bn indica un resto de desoxiuridina sustituido con bencilo. El bloque Z puede sintetizarse usando química de fosforamidita convencional. Los reactivos de captura de aptámeros también se pueden sintetizar mediante química de fosforamidita convencional, y se pueden purificar, por ejemplo, en una columna PRP-3 de 21,5 x 75 mm, que funciona a 80 °C en un sistema 2767 de autopurificación de Waters (o sistema semiautomático serie 600 de Waters), usando, por ejemplo, un calentador TL-600 o TL-150 Timberline y un gradiente de bicarbonato de trietilamonio (TEAB)/ACN para eluir el producto. La detección se realiza a 260 nm, y las fracciones se recogen a través del máximo principal antes de agrupar las mejores fracciones.

### Tampones

El tampón SB18 se compone de HEPES 40 mM, NaCl 101 mM, KCl 5 mM, MgCl<sub>2</sub> 5 mM y Tween 20 al 0,05 % (v/v) ajustado a pH 7,5 con NaOH. El tampón SB17 es SB18 suplementado con EDTA trisódico 1 mM. El tampón PB1 se compone de HEPES 10 mM, NaCl 101 mM, KCl 5 mM, MgCl<sub>2</sub> 5 mM, EDTA trisódico 1 mM y Tween 20 al 0,05 % (v/v) ajustado a pH 7,5 con NaOH. El tampón de elución CAPSO consiste en CAPSO 100 mM a pH 10,0 y NaCl 1 M. El tampón de neutralización contiene HEPES 500 mM, HCl 500 mM y Tween-20 al 0,05 % (v/v). El tampón de hibridación Agilent es una formulación patentada que se suministra como parte de un kit (kit de hibridación Oligo aCGH/ChIP sobre chip). El tampón de lavado Agilent 1 es una formulación patentada (tampón de lavado Oligo aCGH/ChIP sobre chip 1, Agilent). El tampón de lavado Agilent 2 es una formulación patentada (tampón de lavado Oligo aCGH/ChIP sobre chip 2, Agilent). La solución de hibridación de TMAC consiste en cloruro de tetrametilamonio 4,5 M, EDTA trisódico 6 mM, Tris-HCl 75 mM (pH 8,0) y Sarcosilo al 0,15 % (v/v). El tampón KOD (10 veces concentrado) consiste en Tris-HCl 1200 mM, MgSO<sub>4</sub> 15 mM, KCl 100 mM, 60 mM (H<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, Triton-X 100 AL 1 % v/v y 1 mg/ml de BSA.

### Preparación de las muestras

Se descongela el suero (almacenado a -80 °C en alícuotas de 100 µl) en un baño de agua a 25 °C durante 10 minutos, y luego se almacena en hielo antes de la dilución de la muestra. Las muestras se mezclan mediante agitación suave durante 8 segundos. Se prepara una solución de muestra de suero al 6 % mediante dilución en 0,94 x SB17 suplementado con MgCl<sub>2</sub> 0,6 mM, EGTA trisódico 1 mM, AEBSF 0,8 mM y Bloque Z 2 µM. Se diluye una parte de la solución madre de suero al 6 % con factor de dilución de 10 en SB 17 para crear una solución madre de suero del 0,6 %. En algunas realizaciones, se usan soluciones madre del 6 % y 0,6 % para detectar analitos de abundancia alta y baja, respectivamente.

### Preparación de placa de reactivo de captura (aptámero) y estreptavidina

Los aptámeros se agrupan en 2 mezclas de acuerdo con la abundancia relativa de sus analitos afines (o biomarcadores). Las concentraciones madre son 4 nM para cada aptámero, y la concentración final de cada aptámero es de 0,5 nM. Se diluyen las mezclas madre de aptámero con factor de dilución de 4 en tampón SB17, se calientan a 95 °C durante 5 min y se enfrían a 37 °C durante un período de 15 minutos antes de su uso. Este ciclo de desnaturalización-renaturalización está destinado a normalizar las distribuciones de los confórmers de los aptámeros y, de este modo, garantizar la actividad reproducible del aptámero a pesar de las historias variables. Se lavan dos veces las placas de estreptavidina con 150 µl de tampón PB1 antes de su uso.

### Equilibrado y captura de la placa

Se combinan las mezclas de aptámero x 2 enfriadas por calor (55 µl) con un volumen igual de diluciones de suero al 6 % o 0,6 %, produciendo mezclas de equilibrado que contienen suero al 3 % y 0,3 %. Se sellan las placas con una estera de sellado de silicona (estera de sellado de silicona Axymat, VWR) y se incuban durante 1,5 h a 37 °C. A continuación, se transfieren las mezclas de equilibrado a los pocillos de una placa de estreptavidina de 96 pocillos lavada y se incuban más en un termomezclador Eppendorf ajustado a 37 °C, con agitación a 800 rpm, durante dos horas.

### Ensayo manual

A menos que se especifique lo contrario, se elimina el líquido por descarga, seguido de dos toques sobre toallas de papel en capas. Los volúmenes de lavado son de 150 µl, y todas las incubaciones de agitación se realizan en un termomezclador Eppendorf ajustado a 25 °C, 800 rpm. Se retiran las mezclas de equilibrado mediante pipeteo, y se lavan las placas dos veces durante 1 minuto con tampón PB1 suplementado con sulfato de dextrano 1 mM y biotina

500  $\mu$ M, luego 4 veces durante 15 segundos con tampón PB1. Se añade una solución recién preparada de NHS-PEO4-biotina 1 mM en tampón PB1 (150  $\mu$ l/pocillo), y se incuban las placas durante 5 minutos con agitación. Se retira la solución de NHS-biotina y se lavan las placas 3 veces con tampón PB1 suplementado con glicina 20 mM y 3 veces con tampón PB1. Luego se añaden ochenta y cinco  $\mu$ l de tampón PB1 suplementado con  $\text{DxSO}_4$  1 mM a cada pocillo, y se irradian las placas bajo una lámpara UV BlackRay (longitud de onda nominal 365 nm) a una distancia de 5 cm durante 20 minutos con agitación. Se transfieren las muestras a una placa recubierta con estreptavidina lavada recién preparada, o un pocillo sin usar de la placa de estreptavidina lavada existente, combinando mezclas de dilución de muestras alta y baja en un único pocillo. Las muestras se incuban a temperatura ambiente con agitación durante 10 minutos. Se retira el material no adsorbido y se lavan las placas 8 veces durante 15 segundos cada una con tampón PB1 suplementado con glicerol al 30 %. Se lavan las placas una vez con tampón PB1. Los aptámeros se eluyen durante 5 minutos a temperatura ambiente con 100  $\mu$ l de tampón de elución CAPSO. Se transfieren 90  $\mu$ l de la fracción eluída a una placa HybAid de 96 pocillos y se añaden 10  $\mu$ l de tampón de neutralización.

#### Ensayo semiautomatizado

Se colocan las placas de estreptavidina que portan mezclas de equilibrio adsorbidas en la plataforma de un lavador de placas BioTek EL406, que se programa para realizar las siguientes etapas: Se retira el material no adsorbido por aspiración y se lavan los pocillos 4 veces con 300  $\mu$ l de tampón PB1 suplementado con sulfato de dextrano 1 mM y biotina 500  $\mu$ M. Se lavan los pocillos 3 veces con 300  $\mu$ l de tampón PB1. Se añaden ciento cincuenta  $\mu$ l de una solución recién preparada (a partir de una solución madre 100 mM en DMSO) de NHS-PEO4-biotina 1 mM en tampón PB1. Las placas se incuban durante 5 minutos con agitación. Se aspira el líquido y se lavan los pocillos 8 veces con 300  $\mu$ l de tampón PB1 suplementado con glicina 10 mM. Se añaden cien  $\mu$ l de tampón PB1 suplementado con sulfato de dextrano 1 mM. Tras estas etapas automatizadas, se retiran las placas del lavador de placas y se colocan en un termoagitador montado bajo una fuente de luz UV (BlackRay, longitud de onda nominal 365 nm) a una distancia de 5 cm durante 20 minutos. El termoagitador se configura a 800 rpm y 25 °C. Después de 20 minutos de irradiación, se transfieren las muestras manualmente a una placa de estreptavidina lavada y recién preparada (o a un pocillo sin usar de la placa lavada existente). En este punto, se combinan en un solo pocillo mezclas de reacción de alta abundancia (mezcla de suero al 3 % + aptámeros al 3 %) y de baja abundancia (mezcla de suero al 3 % + aptámeros al 3 %). Se coloca esta placa "Catch-2" en la plataforma del lavador de placas BioTek EL406, que está programado para realizar las siguientes etapas: la placa se incuba durante 10 minutos con agitación. Se aspira el líquido y se lavan los pocillos 21 veces con 300  $\mu$ l de tampón PB1 suplementado con glicerol al 30 %. Se lavan los pocillos 5 veces con 300  $\mu$ l de tampón PB1, y se aspira el lavado final. Se añaden cien  $\mu$ l de tampón de elución CAPSO, y se eluyen los aptámeros durante 5 minutos con agitación. Tras estas etapas automatizadas, se retira la placa de la plataforma del lavador de placas y se transfieren 90 alícuotas de las muestras manualmente a los pocillos de una placa HybAid de 96 pocillos que contiene 10  $\mu$ l de tampón de neutralización.

#### Hibridación a micromatrices Agilent 8x 15k personalizadas

Se transfieren 24  $\mu$ l de la fracción eluída neutralizada a una nueva placa de 96 pocillos y se añaden 6  $\mu$ l de Agilent Block x 10 (kit de hibridación Oligo aCGH/ChIP sobre chip, Gran volumen, Agilent 5188-5380), que contiene un conjunto de controles de hibridación compuesto de 10 aptámeros Cy3 a cada pocillo. Se añaden treinta  $\mu$ l de tampón de hibridación Agilent x 2 a cada muestra y se mezcla. Se pipetea manualmente cuarenta  $\mu$ l de la solución de hibridación resultante en cada "pocillo" del portaobjetos de la junta de hibridación (portaobjetos de junta de hibridación, formato de micromatriz 8 por portaobjetos, Agilent). Se disponen los portaobjetos de las micromatrices Agilent personalizadas, que portan 10 sondas por matriz complementaria a la región aleatoria de 40 nucleótidos de cada aptámero con un enlazador 20x dT, en los portaobjetos de la junta de acuerdo con el protocolo del fabricante. Se sujeta la pieza (Kit de cámara de hibridación – habilitado por SureHyb, Agilent) y se incuba durante 19 horas a 60 °C mientras gira a 20 rpm.

#### Lavado posterior a la hibridación

Se disponen aproximadamente 400 ml de tampón de lavado Agilent 1 en cada una de dos placas de tinción de vidrio separadas. Se desmontan los portaobjetos (no más de dos a la vez) y se separan mientras se sumergen en el tampón de lavado 1, luego se transfieren a un portaobjetos de una segunda placa de tinción que también contiene tampón de lavado 1. Se incuban los portaobjetos durante 5 minutos más en tampón de lavado 1 con agitación. Se transfieren los portaobjetos a tampón de lavado 2 preequilibrado a 37 °C y se incuban durante 5 minutos con agitación. Se transfieren los portaobjetos a un cuarto de la placa de tinción que contiene acetónitrilo, y se incuban durante 5 minutos con agitación.

#### Generación de imágenes de micromatriz

Se generan imágenes de los portaobjetos de las micromatrices con un sistema de escáner de micromatrices Agilent G2565CA, usando el canal Cy3 a una resolución de 5  $\mu$ m una configuración PMT del 100 %, y la opción de XRD habilitada a 0,05. Las imágenes TIFF resultantes se procesan usando el software de extracción de características Agilent versión 10.5.1.1 con el protocolo GE1\_105\_Dec08. Los datos primarios de Agilent están disponibles como Información complementaria (Figura S6).

Diseño de la sonda Lumindex

Las sondas inmovilizadas en perlas tienen 40 desoxinucleótidos complementarios al extremo 3' de la región aleatoria de 40 nucleótidos del aptámero diana. La región complementaria del aptámero se acopla a microesferas Lumindex a través de un enlazador de hexaetilenglicol (HEG) portador del extremo amino 5'. Los desoxioligonucleótidos de detección biotinilados comprenden 17-21 desoxinucleótidos complementarios a la región de cebador 5' de los aptámeros diana. Las fracciones de biotina se unen a los extremos 3' de los oligos de detección.

Acoplamiento de sondas a microesferas Lumindex

Se acoplan las sondas a microesferas Lumplex Microplex esencialmente según las instrucciones del fabricante, pero con las siguientes modificaciones: las cantidades de oligonucleótidos amino terminales son de 0,08 nmol por  $2,5 \times 10^6$  microesferas, y la segunda adición de EDC es de 5  $\mu$ l a 10 mg/ml. Las reacciones de acoplamiento se realizan en un termoagitador Eppendorf ajustado a 25 °C y 600 rpm.

Hibridación de microesferas

Se someten las soluciones madre de microesferas (aproximadamente 40.000 microesferas/ $\mu$ l) a movimientos vorticiales y a tratamiento de ultrasonidos en un limpiador ultrasónico Health Sonics (Modelo: T1.9C) durante 60 segundos para suspender las microesferas. Se diluyen las microesferas suspendidas a 2000 microesferas por reacción en soluciones de hibridación TMAC x 1,5 y se mezclan mediante agitación vorticial y tratamiento de ultrasonidos. Se transfieren treinta y tres  $\mu$ l por reacción de la mezcla de perlas a una placa HybAid de 96 pocillos. Se añaden siete  $\mu$ l de solución madre de oligonucleótidos de detección biotinilados 15 nM en tampón TE x 1 a cada reacción y se mezclan. Se añaden diez  $\mu$ l de muestra de ensayo neutralizada y se sella la placa con un sello de estera de silicona. Primero se incuba la placa a 96 °C durante 5 minutos y se incuba a 50 °C sin agitación durante la noche en un horno de hibridación convencional. Se humedece previamente una placa de filtro (Dura pore, número de pieza Millipore MSBVN1250, tamaño de poro de 1,2  $\mu$ m) con 75  $\mu$ l de solución de hibridación TMAC x 1 complementada con BSA al 0,5 % (p/v). Se transfiere el volumen de muestra completo de la reacción de hibridación a la placa filtrante. Se enjuaga la placa de hibridación con 75  $\mu$ l de solución de hibridación TMAC x 1 que contiene BSA al 0,5 % y se transfiere cualquier material restante a la placa filtrante. Se filtran las muestras al vacío lentamente, con de 150  $\mu$ l de tampón al vacío durante aproximadamente 8 segundos. Se lava la placa filtrante una vez con 75  $\mu$ l de solución de hibridación de TMAC que contiene BSA al 0,5 % y se vuelven a suspender las microesferas de la placa filtrante en 75  $\mu$ l de solución de hibridación TMAC x 1 que contiene BSA al 0,5 %. Se protege la placa filtrante de la luz y se incuba en un termomezclador Eppendorf R durante 5 minutos a 1000 rpm. A continuación, se lava la placa filtrante una vez con 75  $\mu$ l de solución de hibridación TMAC x 1 que contiene BSA al 0,5 %. Se añaden 75  $\mu$ l de 10  $\mu$ g/ml de estreptavidina-ficoeritrina (SAPE-100, MOSS, Inc.) en solución de hibridación TMAC x 1 a cada reacción y se incuban en un termomezclador Eppendorf R a 25 °C a 1.000 rpm durante 60 minutos. Se lava la placa filtrante una vez con 75  $\mu$ l de 1x solución de hibridación de TMAC que contiene BSA al 0,5 % y se vuelven a suspender las microesferas de la placa filtrante en 75  $\mu$ l de solución de hibridación TMAC x 1 que contiene BSA al 0,5 %. A continuación, se incuba la placa filtrante protegida de la luz en un termomezclador Eppendorf R durante 5 minutos a 1000 rpm. A continuación, se lava la placa filtrante una vez con 75  $\mu$ l de solución de hibridación TMAC x 1 que contiene BSA al 0,5 %. Se vuelven a suspender las microesferas en 75  $\mu$ l solución de hibridación TMAC x 1 complementada con BSA al 0,5 %, y se analizan en un instrumento Lumindex 100 que ejecuta el software XPonent 3.0. Se cuentan al menos 100 microesferas por tipo de perlas, bajo una alta calibración de PMT y un ajuste de discriminador de doblete de 7.500 a 18.000.

Lectura de qPCR

Se preparan curvas patrón para qPCR se en agua que varía de 108 a 102 copias con diluciones de factor de dilución de 10 y un control sin molde. Se diluyen las muestras de ensayo neutralizadas 40 veces en H<sub>2</sub>O desionizada. La mezcla maestra de qPCR se prepara a una concentración final del doble (tampón KOD x 2, mezcla de dNTP 400  $\mu$ M, mezcla de cebador directo e inverso 400 nM, SYBR Green I x 2 y 0,5 U de KOD EX). Se añaden diez  $\mu$ l de mezcla maestra qPCR x 2 a 10  $\mu$ l de muestra de ensayo diluida. Se ejecuta la qPCR en un BioRad MyiQ iCycler con 2 minutos a 96 °C, seguidos de 40 ciclos de 96 °C durante 5 segundos y 72 °C durante 30 segundos.

**Ejemplo 5. Análisis adicional de EHGNA y EHNA**

Se sometieron los datos del ensayo multiplexado basado en aptámeros descrito en el Ejemplo 2 para las muestras descritas en el Ejemplo 1 a un análisis adicional, usando un modelo de regresión logística.

La selección de la estabilidad toma muchos subconjuntos de la mitad de los datos y realiza la selección de biomarcadores usando el clasificador de lazo, que es un modelo de regresión logística regularizado. Véase, por ejemplo, Meinshausen *et al.*, 2010, *J. Royal Statistical Soc: Series B* (Statistical Methodology), 72: 417-473. La vía de selección para un solo biomarcador es la proporción de estos subconjuntos para los que ese biomarcador fue seleccionado por el modelo de lazo en un intervalo de lambda. Lambda es un parámetro de ajuste que determina la

cantidad de biomarcadores seleccionados por el lazo. La máxima probabilidad de selección en un intervalo de valores lambda es la medida final usada para seleccionar un conjunto de biomarcadores.

### C. Clasificador de la esteatosis

5 Basándose en las clasificaciones de los sujetos, se supone que todos los grupos de esteatosis, así como de EHNA en fase 1-4, tienen grasa en las células hepáticas. Se desarrolló un clasificador (esteatosis o grasa en el hígado) al comparar los sujetos normales obesos con todos los sujetos con EHGNA y EHNA.

10 Los marcadores se ajustaron a los modelos de regresión logística para generar clasificadores.

La Tabla 8 muestra los biomarcadores en el clasificador de la esteatosis de 8 marcadores (EHGNA). La Tabla 8 también muestra los coeficientes de regresión logística para el modelo y el intervalo de confianza del 95 % para la estimación del coeficiente. El clasificador de 8 marcadores contiene todos los marcadores del clasificador de 9  
15 marcadores como se muestra en la Tabla 3, a excepción de PLAT.

Tabla 8: Clasificador de 8 biomarcadores para EHGNA

| Biomarcador | Nivel de biomarcador superior/inferior en toda la población de EHGNA | Coeficiente | Intervalo de confianza al 95 % bajo | Intervalo de confianza del 95 % alto |
|-------------|----------------------------------------------------------------------|-------------|-------------------------------------|--------------------------------------|
| Intercept   |                                                                      | 23,682      | 0,0205235                           | 48,665007                            |
| ACY1        | Superior                                                             | 2,205       | 1,479981                            | 3,028281                             |
| SHBG        | Inferior                                                             | -1,219      | -1,96066                            | -0,52774                             |
| CTSZ        | Superior                                                             | 3,070       | 1,607459                            | 4,636245                             |
| MET         | Inferior                                                             | -1,900      | -3,84342                            | -0,01765                             |
| GSN         | Inferior                                                             | -2,014      | -4,08462                            | 0,024845                             |
| LGALS3BP    | Superior                                                             | 1,814       | 0,58276                             | 3,132223                             |
| CHL1        | Inferior                                                             | -2,491      | -4,46361                            | -0,59133                             |
| SERPINC1    | Inferior                                                             | -1,502      | -2,96705                            | -0,13184                             |

20 La Figura 12 muestra diagramas de cajas para el clasificador de 8 biomarcadores en cada conjunto de sujetos con esteatosis y EHNA. La línea negra dentro de cada cuadro representa la mediana (o percentil 50) de los puntos de datos, y el propio recuadro representa el intervalo intercuartílico (IQR), el área que abarca los puntos de datos del percentil 25 al 75. Los filamentos se extienden para cubrir puntos de datos dentro de 1.5 x IQR de la parte superior e inferior de la caja. El diagrama muestra una clara diferenciación entre ausencia de esteatosis y esteatosis, y el voto  
25 de probabilidad aumenta con el nivel de esteatosis. Los sujetos con EHNA en todas las fases tienen esteatosis grave.

### D. Clasificador de la EHNA (Fibrosis)

30 Todos los sujetos con EHNA tienen alguna forma de inflamación y abombamiento asociados con la fibrosis. Para este análisis, se compararon todos los grupos de esteatosis leve y moderada con la EHNA en fase 2, 3 y 4. Para garantizar la identificación de los biomarcadores de fibrosis verdaderos, se excluyó el grupo de EHNA en fase 1. Los sujetos con esteatosis grave se omitieron en este análisis, porque el muestreo de biopsia en esos sujetos podría haber perdido algo de fibrosis.

35 Los marcadores se ajustaron a los modelos de regresión logística para generar clasificadores para generar un modelo.

La Tabla 9 muestra los biomarcadores en el clasificador de la fibrosis (EHNA) de 8 marcadores. La Tabla 9 también muestra los coeficientes de regresión logística para el modelo y el intervalo de confianza del 95 % para la estimación  
40 del coeficiente. El clasificador de 8 marcadores contiene todos los marcadores del clasificador de 4 marcadores como se muestra en la Tabla 4, más cuatro marcadores adicionales.

Tabla 9: Clasificador de ocho biomarcadores para EHNA en fases 2, 3 y 4 frente a la esteatosis leve a moderada (EHGNA)

| Biomarcador/<br>Alias | Nivel de biomarcador<br>superior/inferior en la<br>población con fibrosis (EHNA) | Coefficiente | Intervalo de<br>confianza del 95 %<br>bajo | Intervalo de<br>confianza del 95 %<br>alto |
|-----------------------|----------------------------------------------------------------------------------|--------------|--------------------------------------------|--------------------------------------------|
| Intercept             |                                                                                  | -58,307      | -117,01013                                 | -8,491203                                  |
| C7                    | Superior                                                                         | 2,711339     | 0,016014                                   | 5,569299                                   |
| COLEC11               | Superior                                                                         | 0,578295     | -1,01116                                   | 2,184399                                   |
| PPID                  | Inferior                                                                         | -1,47222     | -3,09367                                   | 0,024429                                   |
| IGFBP3                | Inferior                                                                         | -1,88591     | -4,25566                                   | 0,245891                                   |
| GPC5                  | Inferior                                                                         | -2,24248     | -3,99262                                   | -0,74934                                   |
| HIPK3                 | Superior                                                                         | 1,176491     | 0,286007                                   | 2,806758                                   |
| IGFBP7                | Superior                                                                         | 4,363192     | 1,274937                                   | 8,250915                                   |
| IL3RA                 | Superior                                                                         | 3,106253     | 0,455772                                   | 6,208888                                   |

La Figura 13 muestra diagramas de cajas para el clasificador de 8 biomarcadores de la fibrosis en cada uno de los grupos de sujetos (de izquierda a derecha: normal, esteatosis leve, esteatosis moderada, esteatosis grave, EHNA1, EHNA2, EHNA3, EHNA4). La línea negra dentro de cada cuadro representa la mediana (o percentil 50) de los puntos de datos, y el propio recuadro representa el intervalo intercuartílico (IQR), el área que abarca los puntos de datos del percentil 25 al 75. Los filamentos se extienden para cubrir puntos de datos dentro de 1.5 x IQR de la parte superior e inferior de la caja.

Los cuatro biomarcadores adicionales del clasificador de la fibrosis de la Tabla 9 se muestran en la Tabla 10, junto con nombres alternativos, nombres de genes y el número de acceso UniProt de cada biomarcador.

Tabla 10. Biomarcadores adicionales del clasificador de 8 biomarcadores de la fibrosis

| Biomarcador/Alias                                           | Nombre del<br>gen | UniProt                |
|-------------------------------------------------------------|-------------------|------------------------|
| Glipicano 5                                                 | GPC5              | <a href="#">Q8N158</a> |
| Proteína quinasa 3 de interacción con el homeodominio       | HIPK3             | <a href="#">Q9H422</a> |
| Proteína de unión al factor de crecimiento de tipo insulina | IGFBP7            | <a href="#">Q16270</a> |
| 7                                                           |                   |                        |
| Receptor alfa de la interleucina 3                          | IL3RA             | <a href="#">P26951</a> |

La Figura 14 muestra las funciones de distribución acumulativa para los cuatro marcadores adicionales del clasificador de ocho biomarcadores para la fibrosis.

#### Ejemplo 6. Desenmascaramiento de muestras con ocultación

Como se muestra en la Tabla 2, se ocultó un conjunto de muestras de cada uno de los controles, esteatosis leve, esteatosis moderada, esteatosis grave, EHNA en fase 1 y EHNA en fase 2. Se determinó la probabilidad de que cada una de las muestras con ocultación fuera de un sujeto con esteatosis usando el clasificador de 8 marcadores que se muestra en la Tabla 8. La Figura 15 muestra diagramas de cajas para las muestras con ocultación tras el desenmascaramiento, de acuerdo con su grupo de muestra real.

La Figura 16 muestra el rendimiento del clasificador de la esteatosis de 8 marcadores para el conjunto de descubrimiento y para el conjunto de validación con ocultación. La curva superior es la curva ROC para el conjunto de descubrimiento, que tiene un área bajo la curva (ABC) de  $0,927 \pm 0,03$  y una sensibilidad del 92,8 % y una especificidad del 73,3 % en un punto de corte de 0,5. La curva inferior es la curva ROC para el conjunto de validación con ocultación, que tiene una ABC de  $0,889 \pm 0,06$  y una sensibilidad del 88 % y una especificidad del 65,8 % en un punto de corte de 0,5.

La Tabla 10 muestra el rendimiento del clasificador de la esteatosis de 8 marcadores para el conjunto de validación con ocultación para identificar un sujeto como positivo o negativo para la esteatosis.

Tabla 10: Rendimiento del clasificador de la esteatosis de 8 marcadores para el conjunto de validación con ocultación

| Verdad/ensayo | Neg. | Pos. | Total |
|---------------|------|------|-------|
| Neg.          | 27   | 14   | 41    |
| Pos.          | 11   | 81   | 92    |
| Total         | 38   | 95   | 133   |

Se determinó la probabilidad de que cada una de las muestras con ocultación fuera de un sujeto con fibrosis usando el clasificador de 8 marcadores que se muestra en la Tabla 9. La Figura 17 muestra diagramas de cajas para las muestras con ocultación tras el desenmascaramiento, de acuerdo con su grupo de muestra real.

La Figura 18 muestra el rendimiento del clasificador de la fibrosis de 8 marcadores para el conjunto de descubrimiento y para dos conjuntos de validación con ocultación diferentes. El segundo conjunto de validación incluye el primer conjunto de validación más el grupo normal (los individuos negativos para la enfermedad). Debido a que el modelo no fue entrenado con el grupo normal, es independiente de esas muestras, pudiendo revelar falsos positivos. La curva superior (roja) es la curva ROC para el conjunto de descubrimiento (control (esteatosis leve + moderada) frente a la EHNA en fase 2, 3 y 4), que tiene un área bajo la curva (ABC) de  $0,929 \pm 0,04$  y una sensibilidad del 78,8 % y especificidad del 89,9 % en un punto de corte de 0,5. La curva media (azul) es la curva ROC para el primer conjunto de validación con ocultación, que tiene una ABC de  $0,885 \pm 0,11$  y una sensibilidad del 66,7 % y una especificidad del 84,6 % a un valor de corte de 0,5. La curva inferior (azul verdoso) es la curva ROC para el segundo conjunto de validación con ocultación, que tiene una ABC de  $0,891 \pm 0,08$  y una sensibilidad del 66,7 % y una especificidad del 89,5 % a un valor de corte de 0,5.

Las Tablas 11 y 12 muestran el rendimiento del clasificador de la fibrosis de 8 marcadores para los dos conjuntos de validación con ocultación para identificar los sujetos como positivos o negativos para la fibrosis.

Tabla 11: Rendimiento del clasificador de la fibrosis de 8 marcadores para el primer conjunto de validación con ocultación

| Verdad/ensayo | Neg. | Pos. | Total |
|---------------|------|------|-------|
| Neg.          | 22   | 4    | 26    |
| Pos.          | 4    | 8    | 12    |
| Total         | 26   | 12   | 38    |

Tabla 12: Rendimiento del clasificador de la fibrosis de 8 marcadores para segundo conjunto de validación con ocultación

| Verdad/ensayo | Neg. | Pos. | Total |
|---------------|------|------|-------|
| Neg.          | 60   | 7    | 67    |
| Pos.          | 4    | 8    | 12    |
| Total         | 64   | 15   | 79    |

La Figura 19 muestra las distribuciones de la sensibilidad y especificidad de 2.500 iteraciones de arranque del conjunto de verificación de retención del 20 %, usando (A) el clasificador de la esteatosis de 8 marcadores y (B) el clasificador de la fibrosis de 8 marcadores. Se muestran los intervalos de confianza empíricos del 95 %, así como las estimaciones de entrenamiento y validación, que se encuentran dentro de los intervalos de confianza del 95 %.

#### Ejemplo 7. Rendimiento del clasificador en sujetos pediátricos

Se obtuvieron noventa muestras de suero de niños del Departamento de Gastroenterología y Nutrición Pediátrica del Medical College of Wisconsin. Los sujetos que aportaron muestras se dividieron en cuatro grupos: controles no obesos (N = 45), obesos con pruebas de función hepática normal (N = 20), obesos con pruebas de función hepática elevada (N = 7) y sujetos diagnosticados con EHNA (N = 18). Las muestras se analizaron usando el clasificador de la esteatosis de 8 marcadores que se muestra en la Tabla 8 y el clasificador de la fibrosis de 8 marcadores que se muestra en la Tabla 9.

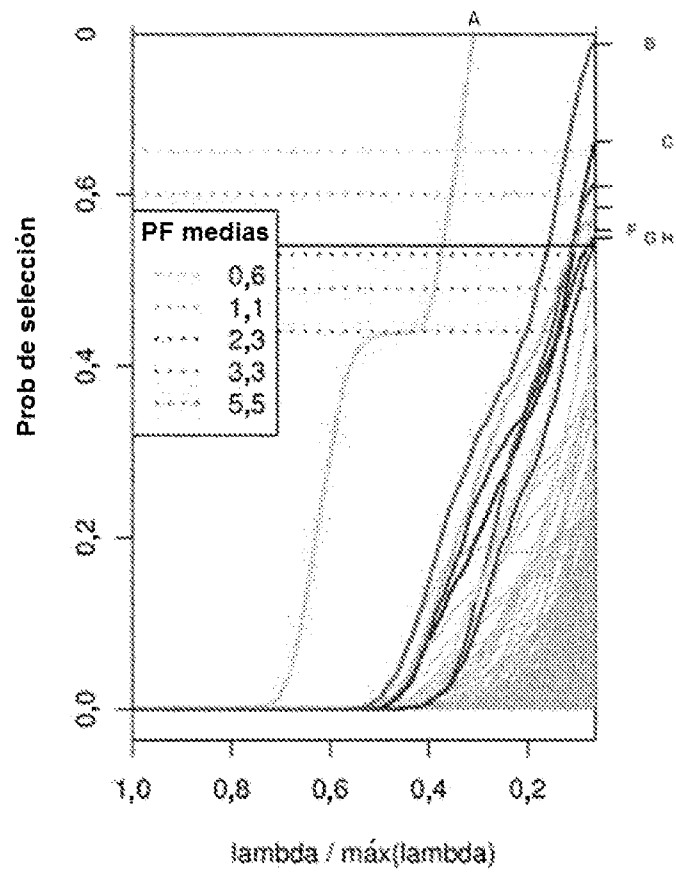
Se determinó la probabilidad de que cada una de las muestras pediátricas fuera de un sujeto con esteatosis usando el clasificador de 8 marcadores que se muestra en la Tabla 8. La Figura 20 muestra diagramas de cajas para las muestras pediátricas de acuerdo con su grupo de muestra real. El clasificador de esteatosis identificó con precisión sujetos pediátricos con EHNA.

A diferencia del clasificador de la esteatosis, el clasificador de la fibrosis de 8 marcadores no identificó con eficacia a los sujetos pediátricos con fibrosis de este experimento. (Datos no mostrados).

- 5 Las realizaciones y los ejemplos anteriores son meramente ilustrativos. Ninguna realización, ejemplo ni elemento en particular de una realización o ejemplo en particular debe interpretarse como un elemento o característica fundamental, necesaria o esencial de cualquiera de las reivindicaciones. Se pueden realizar diversas alteraciones, modificaciones, sustituciones y otras variaciones en las realizaciones desveladas sin apartarse del alcance de la presente invención, que se define mediante las reivindicaciones adjuntas. La memoria descriptiva, que incluye las
- 10 figuras y los ejemplos, se debe considerar de una manera ilustrativa, en lugar de restrictiva, y todas las modificaciones y sustituciones de este tipo pretenden incluirse dentro del alcance de la solicitud. Las etapas enumeradas en cualquiera de las reivindicaciones de métodos o procesos pueden ejecutarse en cualquier orden viable y no se limitan a un orden presentado en cualquiera de las realizaciones, los ejemplos o las reivindicaciones. Además, en cualquiera de los métodos mencionados anteriormente, se pueden excluir específicamente uno o más
- 15 biomarcadores enumerados específicamente como un biomarcador individual o como un biomarcador de cualquier grupo.

## REIVINDICACIONES

1. Un método de determinación de si un sujeto con esteatosis no alcohólica tiene esteatohepatitis no alcohólica (EHNA), que comprende detectar el nivel de al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho o nueve biomarcadores seleccionados entre Complemento C7 (C7), PPID, IGFBP3, SIGLEC14, AIMP1, TOP1, COLEC11, CA6 y STX1A en una muestra del sujeto, en donde un nivel de al menos un biomarcador seleccionado entre C7, COLEC11 y SIGLEC14 que es superior a un nivel de control del biomarcador respectivo, y/o un nivel de al menos un biomarcador seleccionado entre IGFBP3, AIMP1, TOP1, CA6, PPID y STX1A que es inferior a un nivel de control del biomarcador respectivo, indica que el sujeto tiene EHNA, en donde al menos un biomarcador es C7, y en donde la muestra es una muestra de sangre, plasma o suero.
2. El método de la reivindicación 1, en el que el sujeto tiene esteatosis leve, moderada o grave.
3. El método de una cualquiera de las reivindicaciones 1 a 2, en el que la EHNA es EHNA de fase 1, 2, 3 o 4.
4. El método de la reivindicación 3, en el que la EHNA es EHNA de fase 2, 3 o 4.
5. El método de una cualquiera de las reivindicaciones 1 a 4, en donde el método comprende detectar el nivel de C7 y al menos uno, al menos dos, o al menos tres biomarcadores seleccionados entre COLEC11, PPID e IGFBP3.
6. El método de una cualquiera de las reivindicaciones anteriores, en el que el sujeto es obeso.
7. El método de una cualquiera de las reivindicaciones anteriores, en el que cada biomarcador es un biomarcador proteico.
8. El método de una cualquiera de las reivindicaciones anteriores, en donde el método comprende poner en contacto biomarcadores de la muestra del sujeto con un conjunto de reactivos de captura de biomarcadores, en donde cada reactivo de captura de biomarcadores del conjunto de reactivos de captura de biomarcadores se une específicamente a un biomarcador diferente que se está detectando.
9. El método de la reivindicación 8, en el que cada reactivo de captura de biomarcadores es un anticuerpo o un aptámero.
10. El método de la reivindicación 9, en el que al menos un aptámero es un aptámero de velocidad de disociación lenta.
11. El método de la reivindicación 10, en el que al menos un aptámero de velocidad de disociación lenta comprende al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho, al menos nueve o al menos 10 nucleótidos con modificaciones.
12. El método de la reivindicación 10 o de la reivindicación 11, en el que cada aptámero de velocidad de disociación lenta se une a su proteína diana con una velocidad de disociación ( $t_{1/2}$ ) de  $\geq 30$  minutos,  $\geq 60$  minutos,  $\geq 90$  minutos,  $\geq 120$  minutos,  $\geq 150$  minutos,  $\geq 180$  minutos,  $\geq 210$  minutos o  $\geq 240$  minutos.



|   | Diana    |
|---|----------|
| A | ACY1     |
| B | SHBG     |
| C | CTSZ     |
| D | MET      |
| E | GSN      |
| F | LGALS3BP |
| G | PLAT     |
| H | CHL1     |
| I | SERPINC1 |

**FIG. 1**

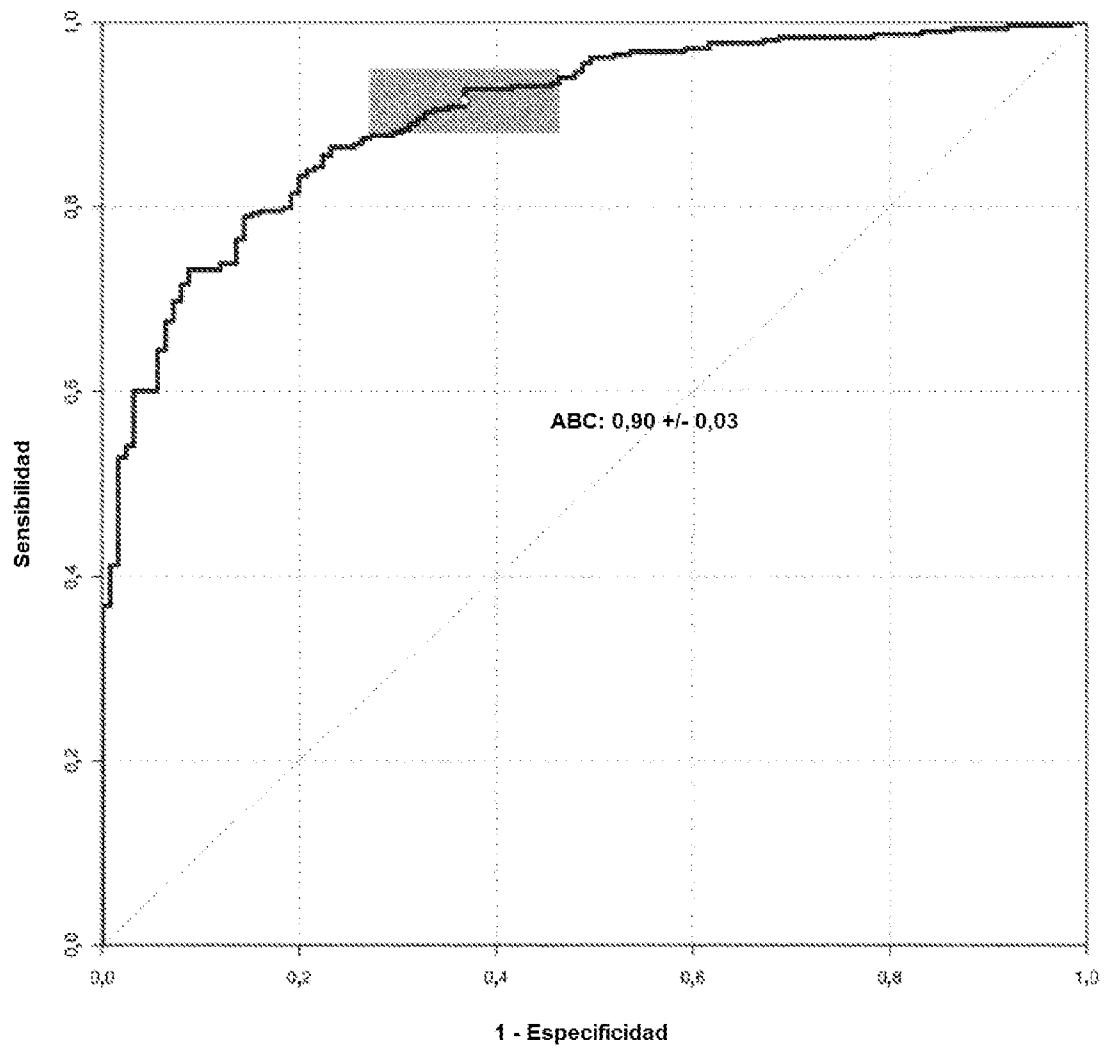


**FIG. 2**

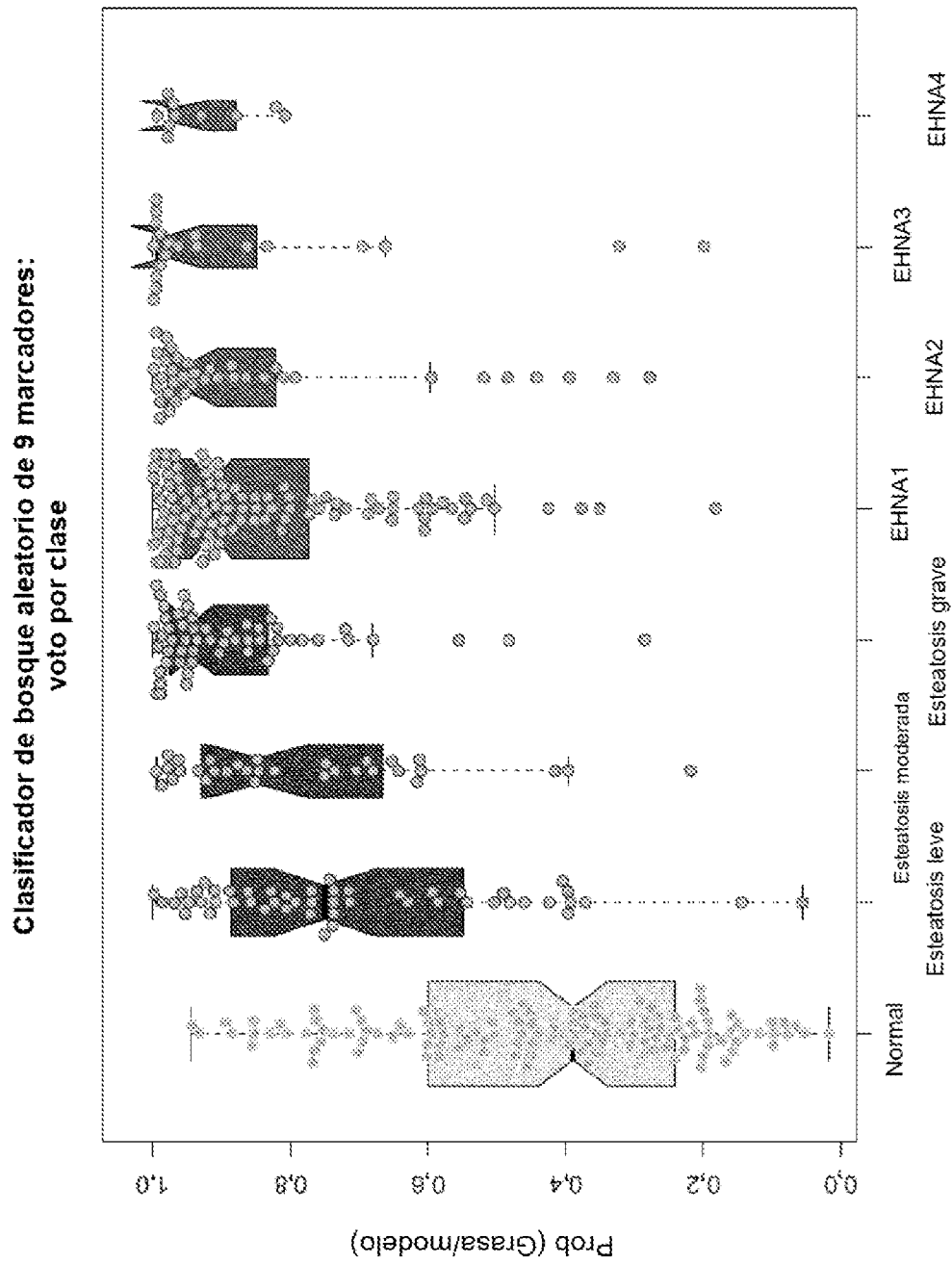


**FIG. 3**

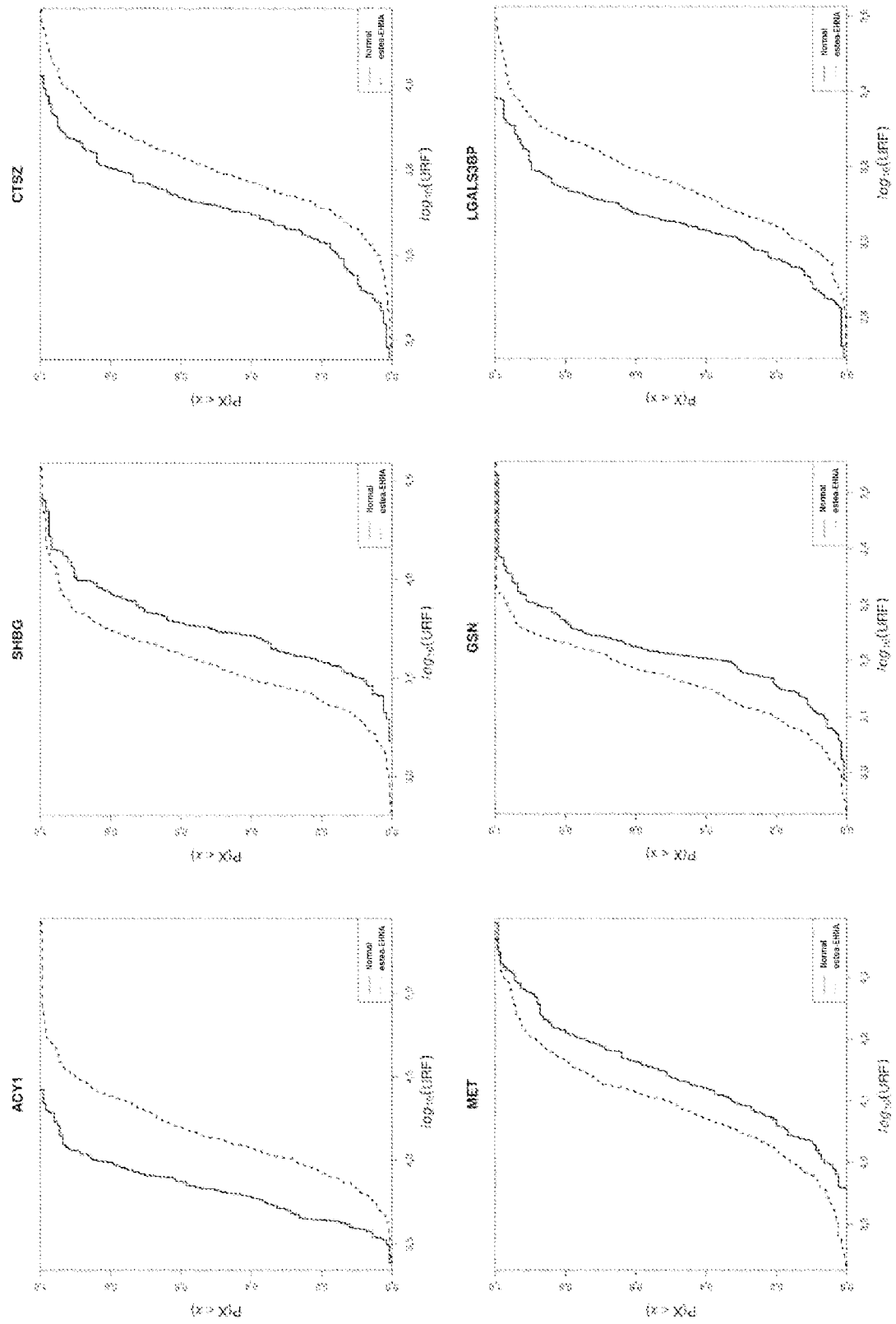


**FIG. 4**

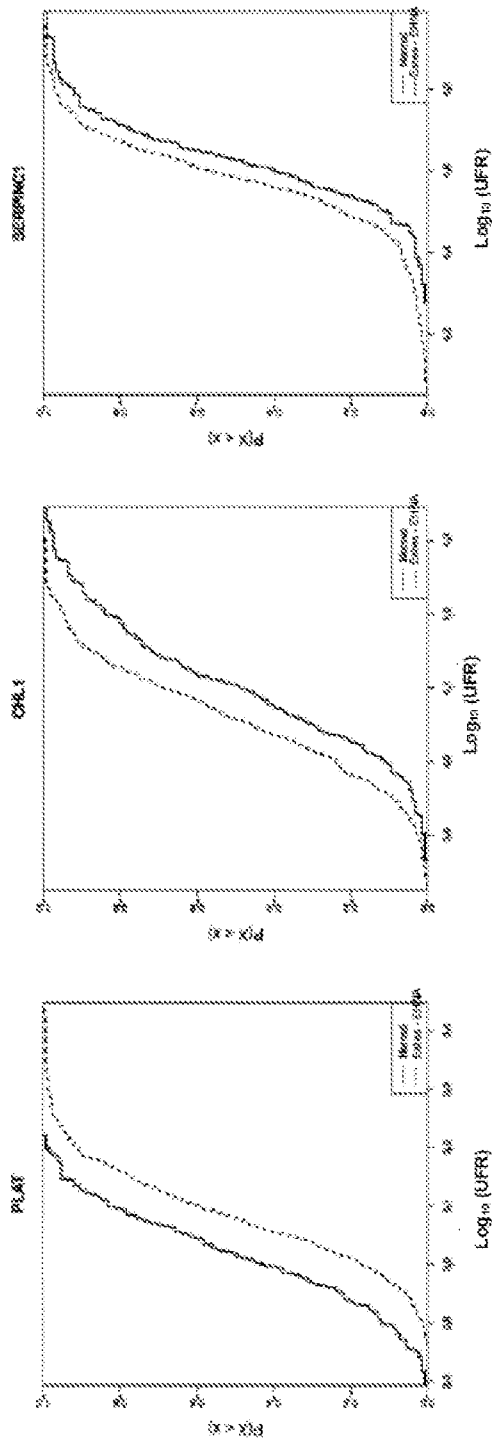
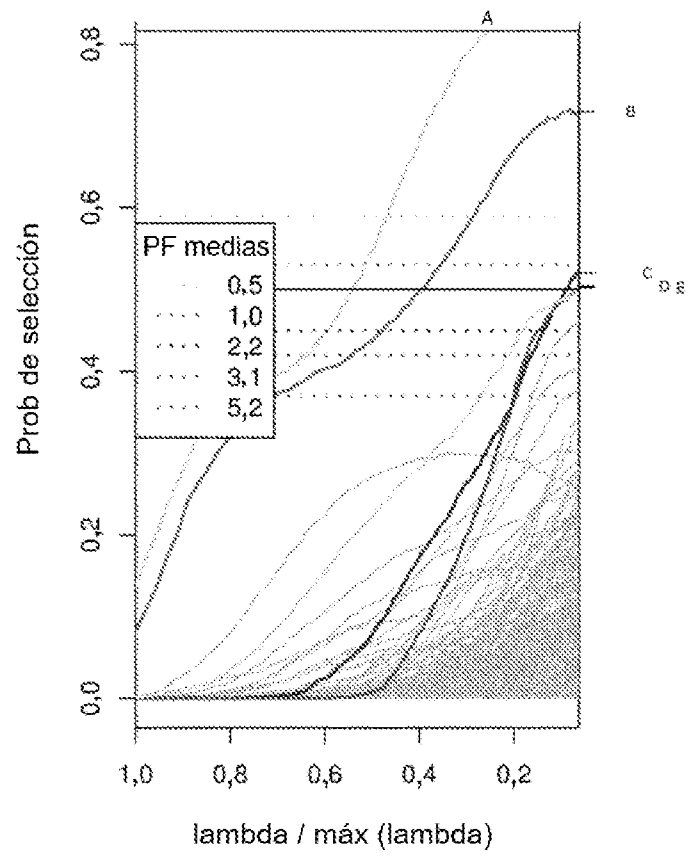


FIG. 4 (cont.)



|   | Diana   |
|---|---------|
| A | C7      |
| B | COLEC11 |
| C | PPID    |
| D | IGFBP3  |

**FIG. 5**

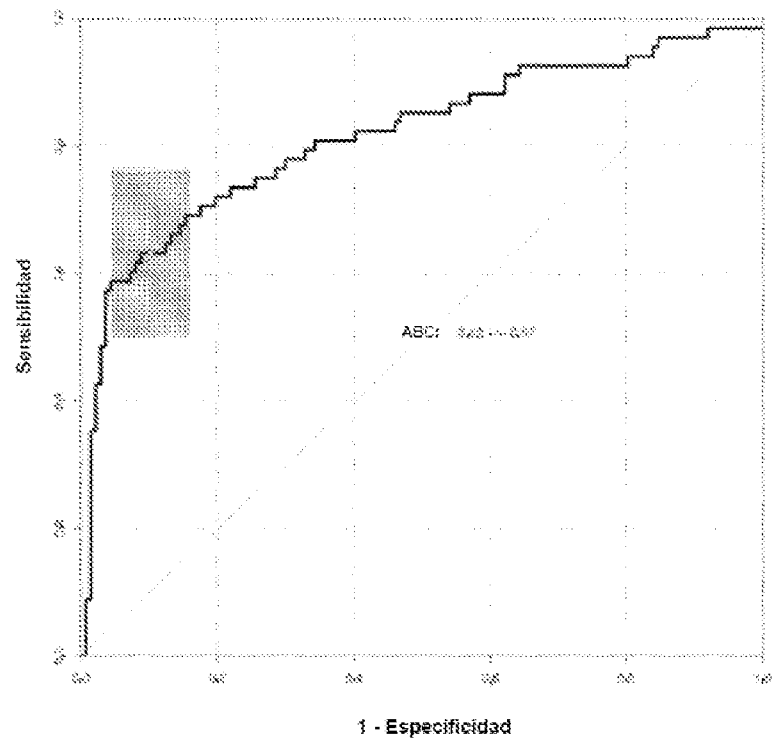


**FIG. 6**

Clasificador de bosque aleatorio de 4 marcadores:  
voto por clase

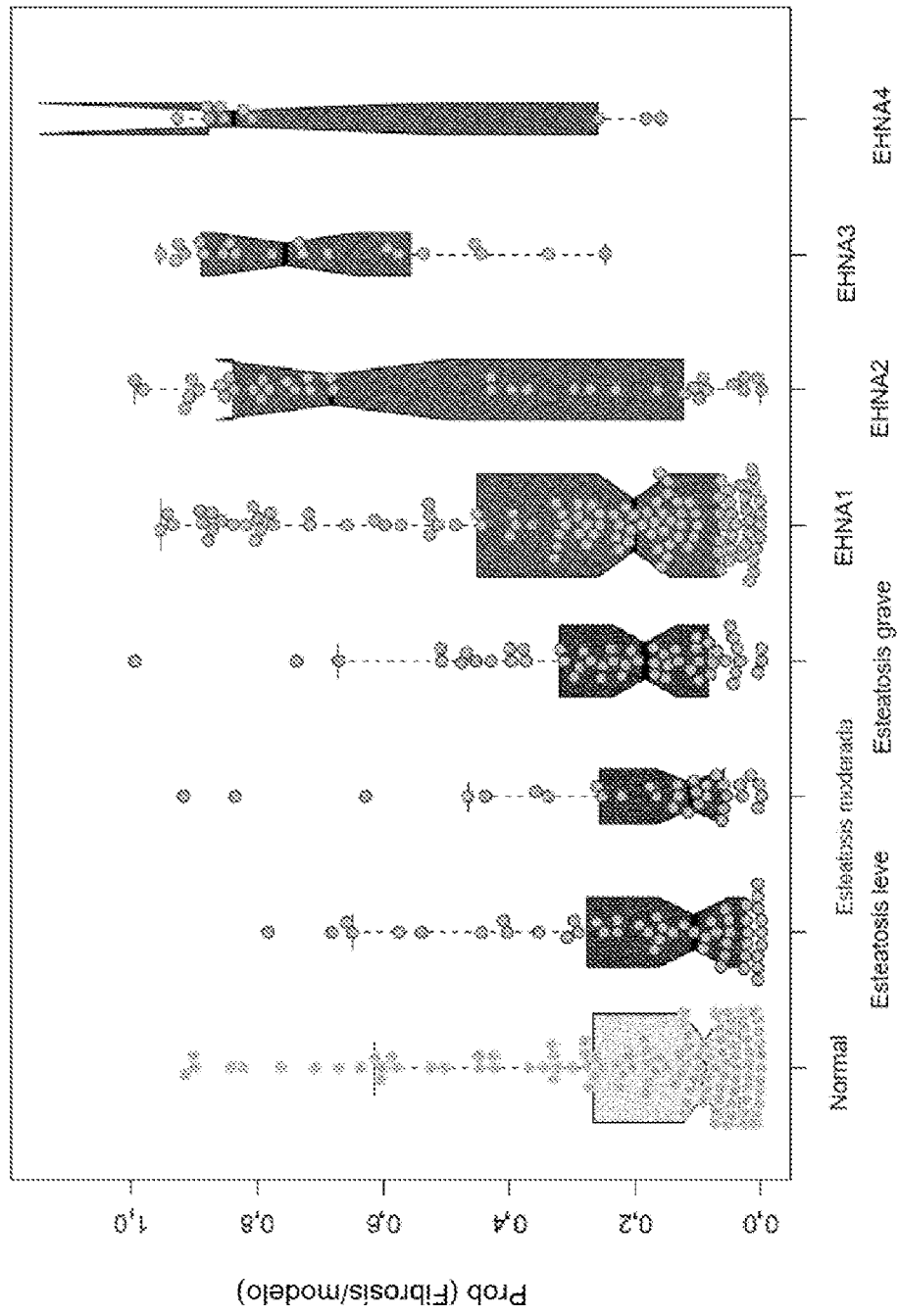


**FIG. 7**

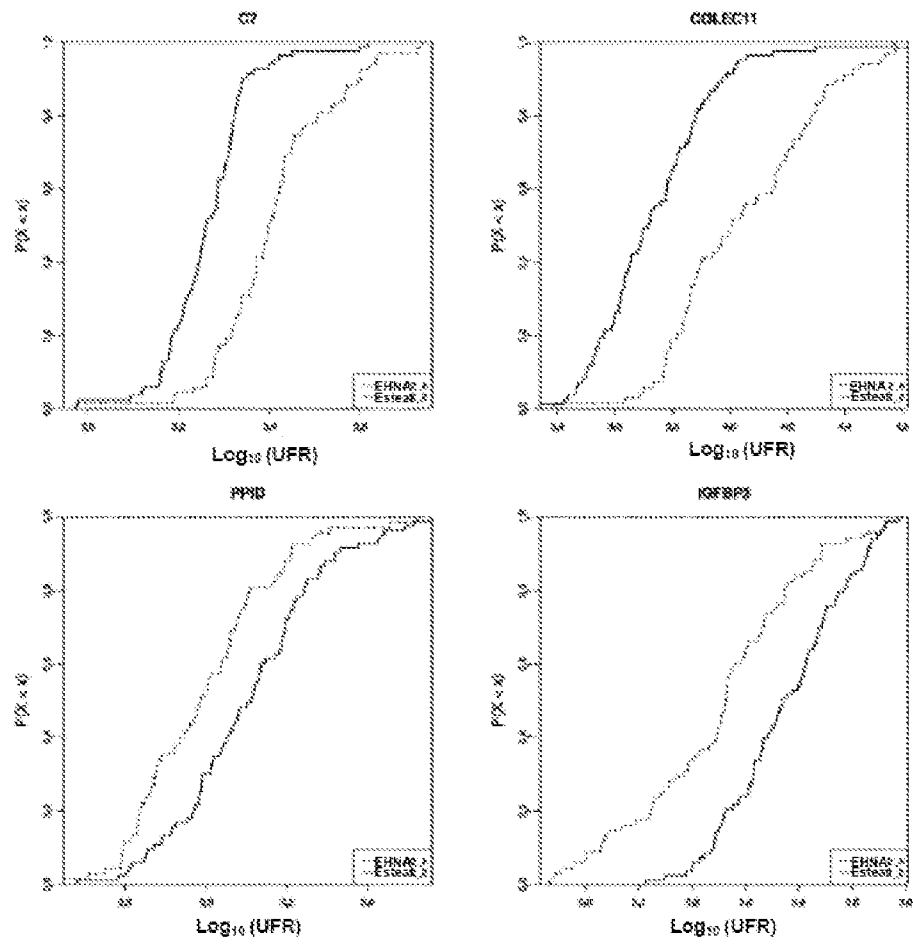


**FIG. 8**

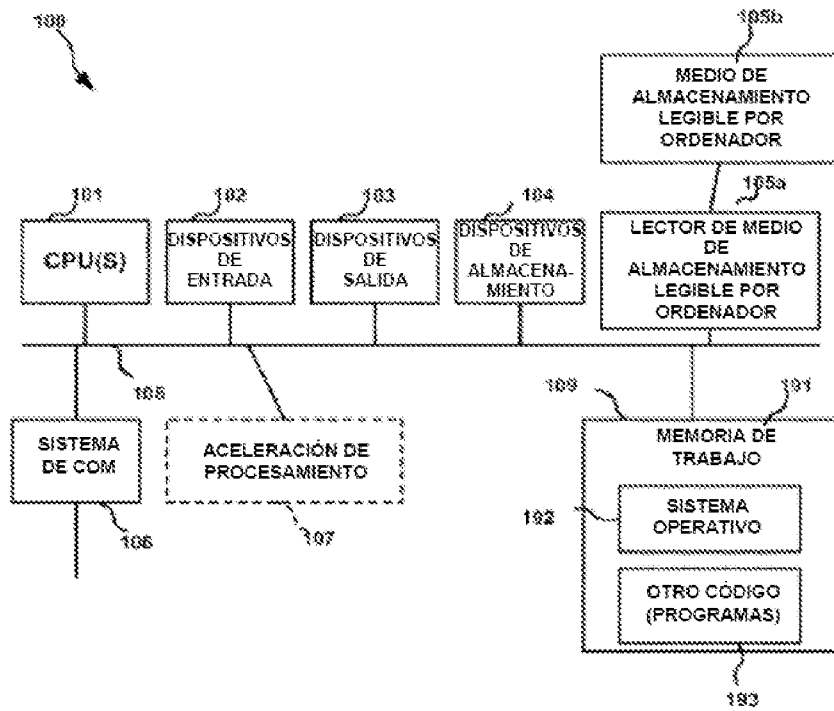


FIG. 9

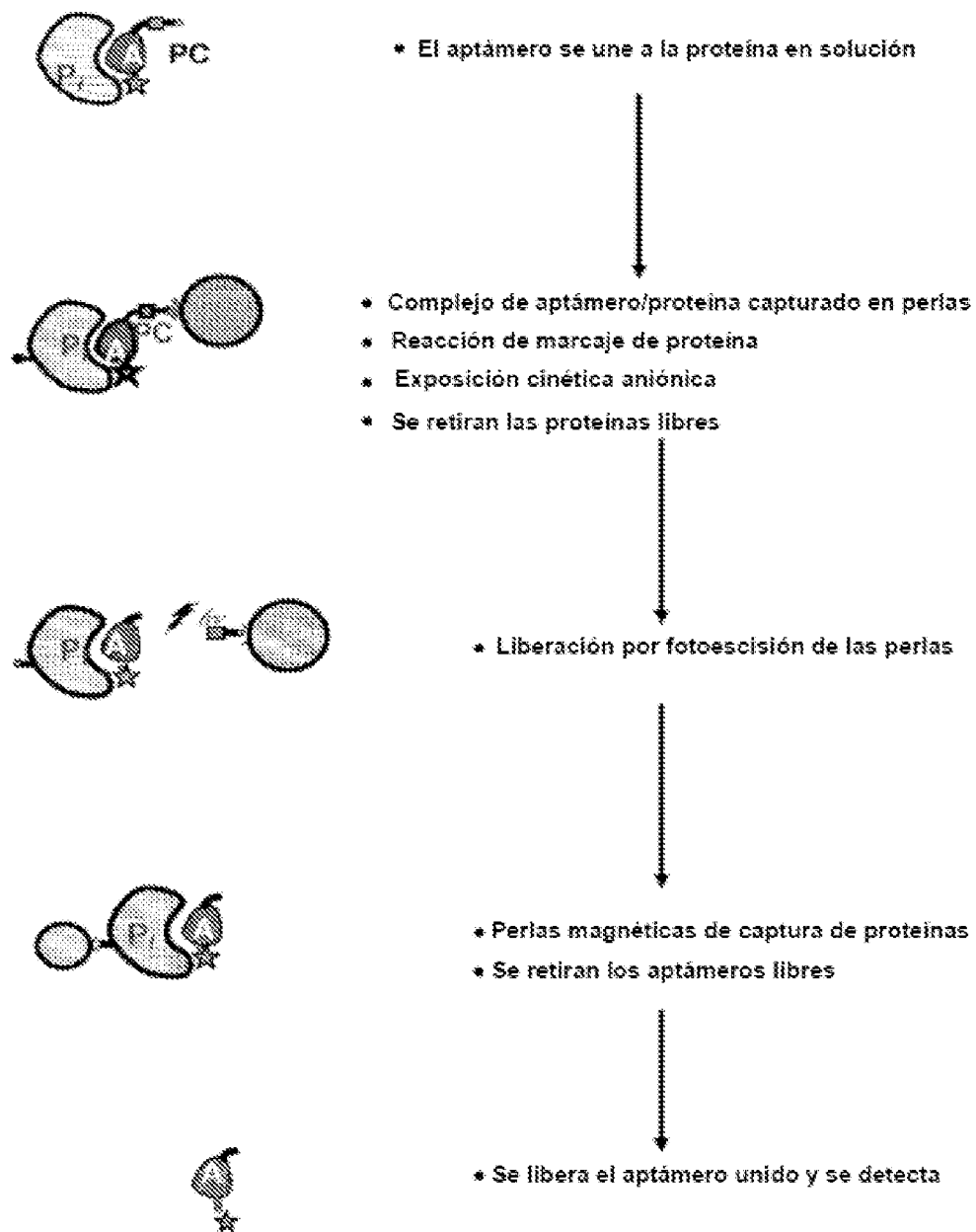


FIG. 10

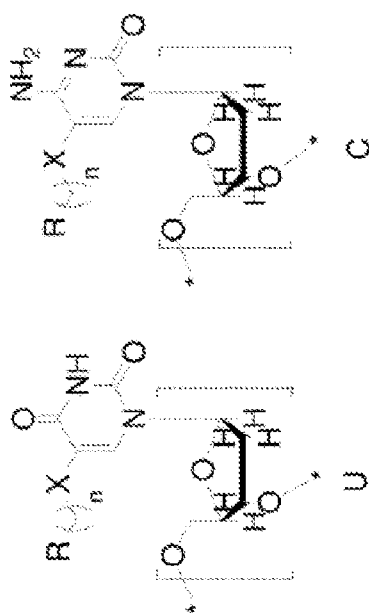
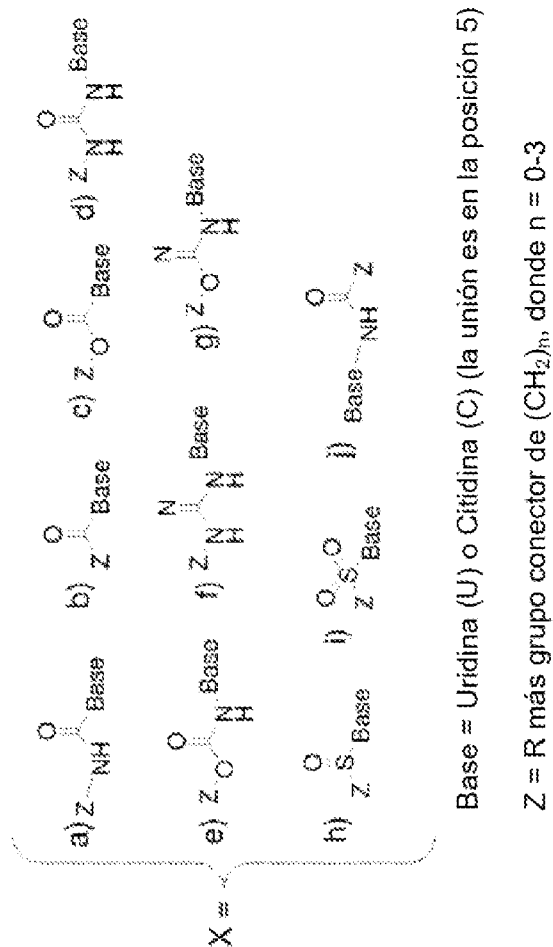


FIG. 11

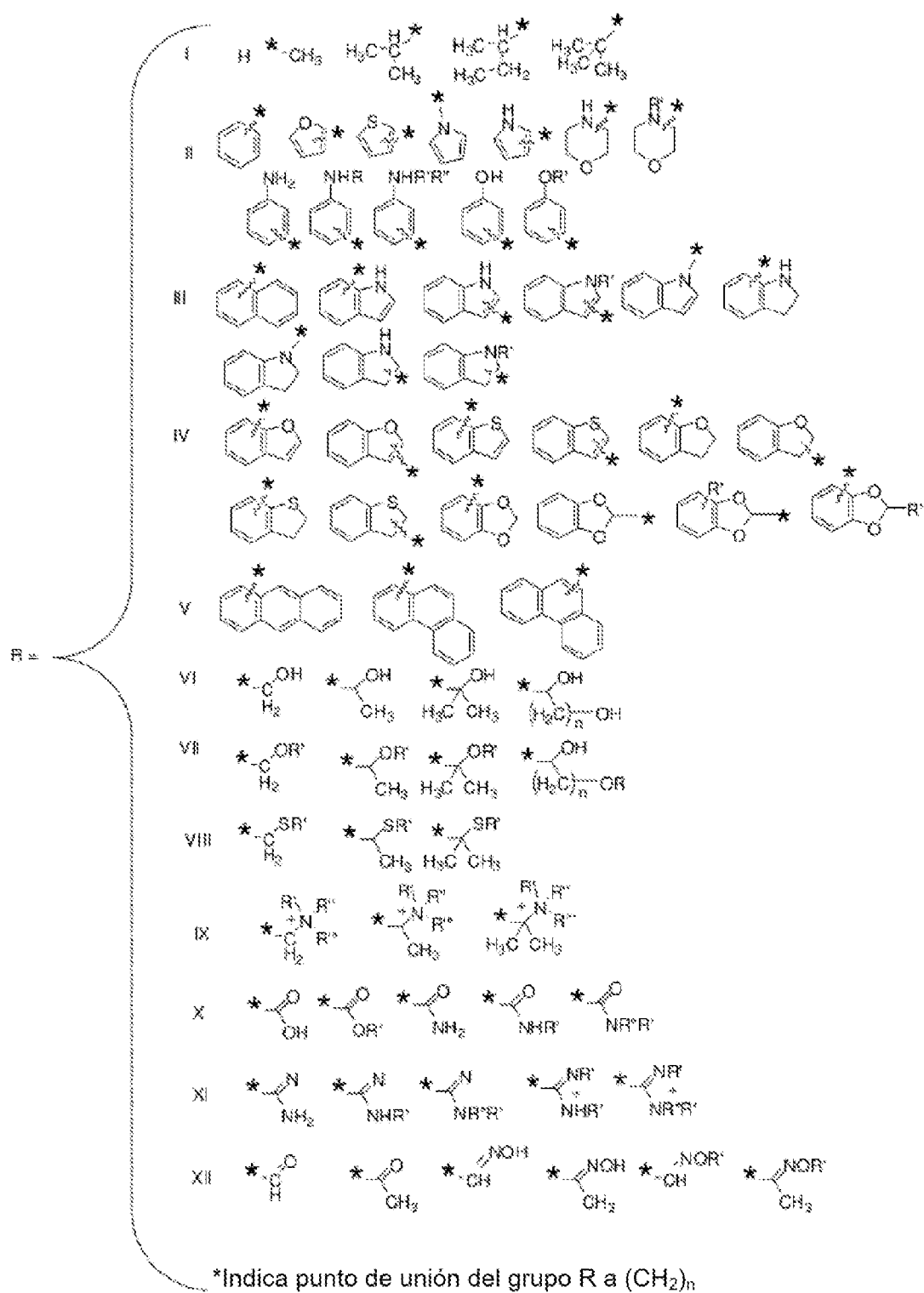


FIG. 11 (cont.)

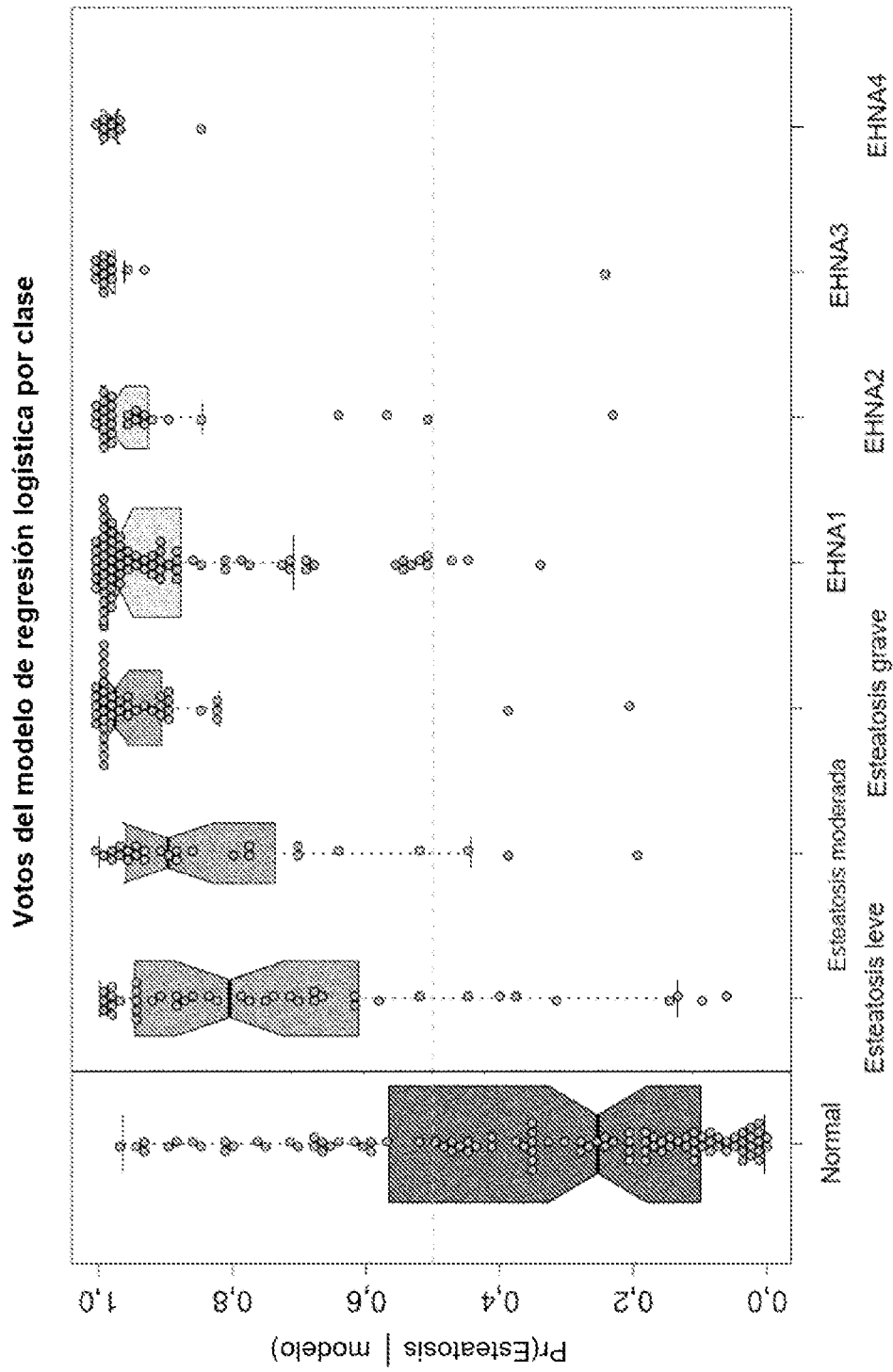


FIG. 12

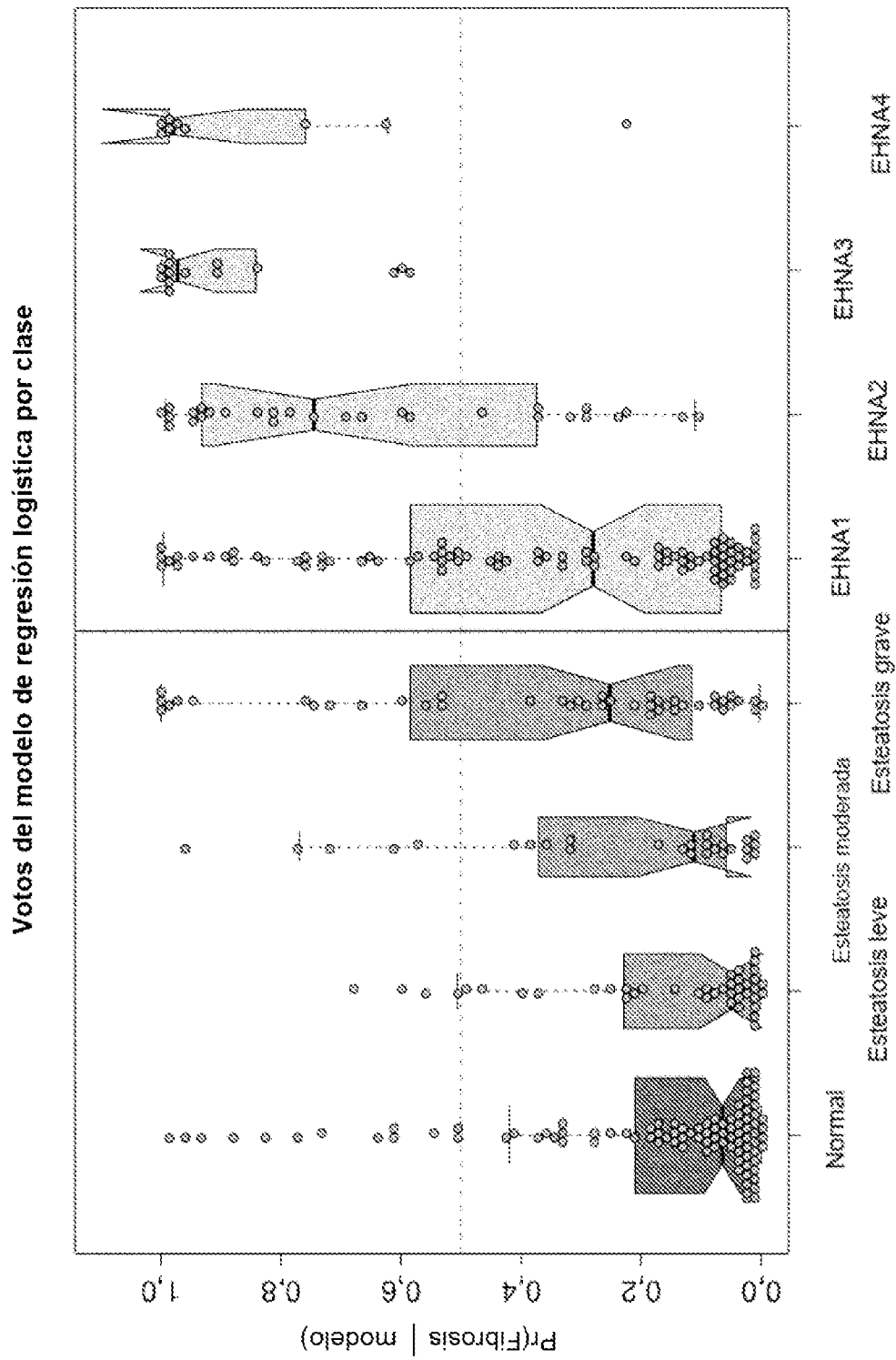


FIG. 13

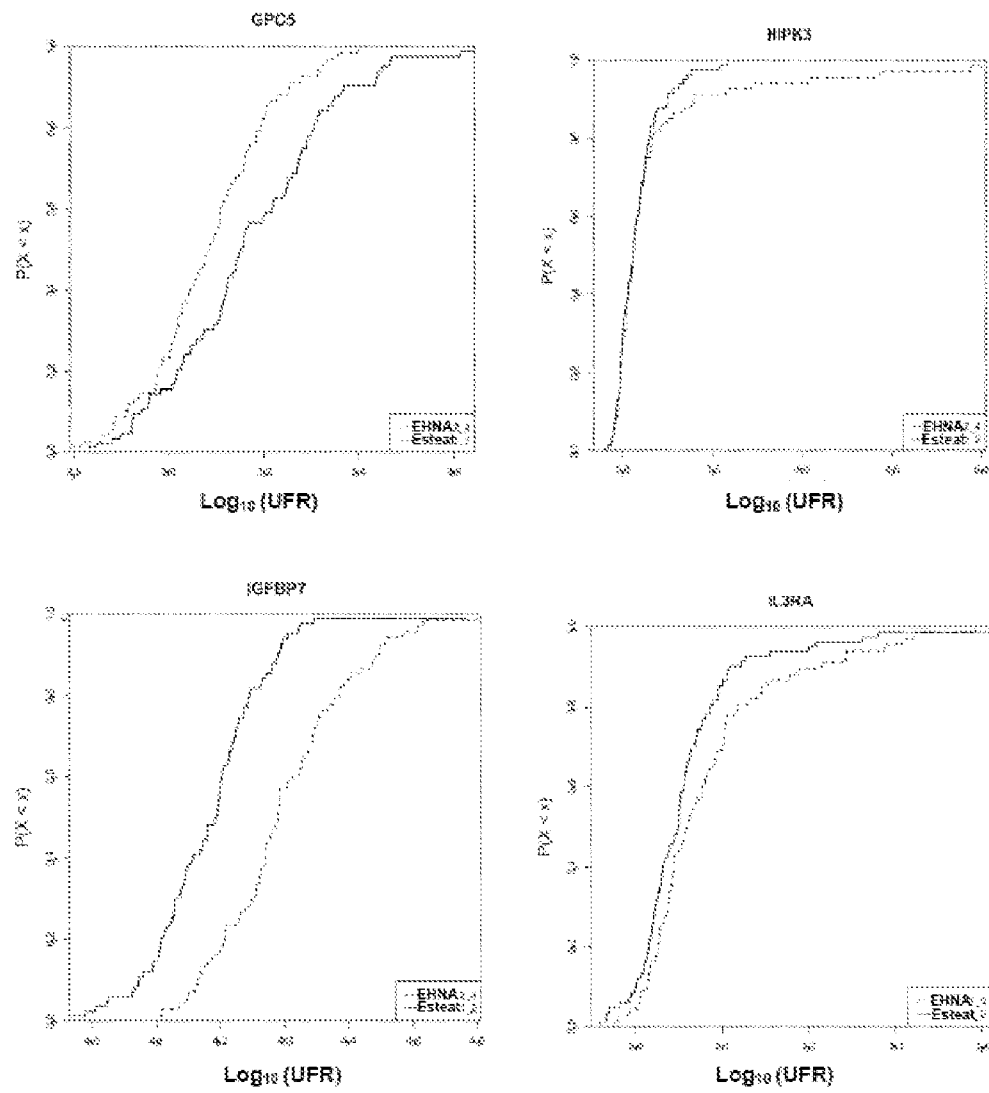


**FIG. 14**

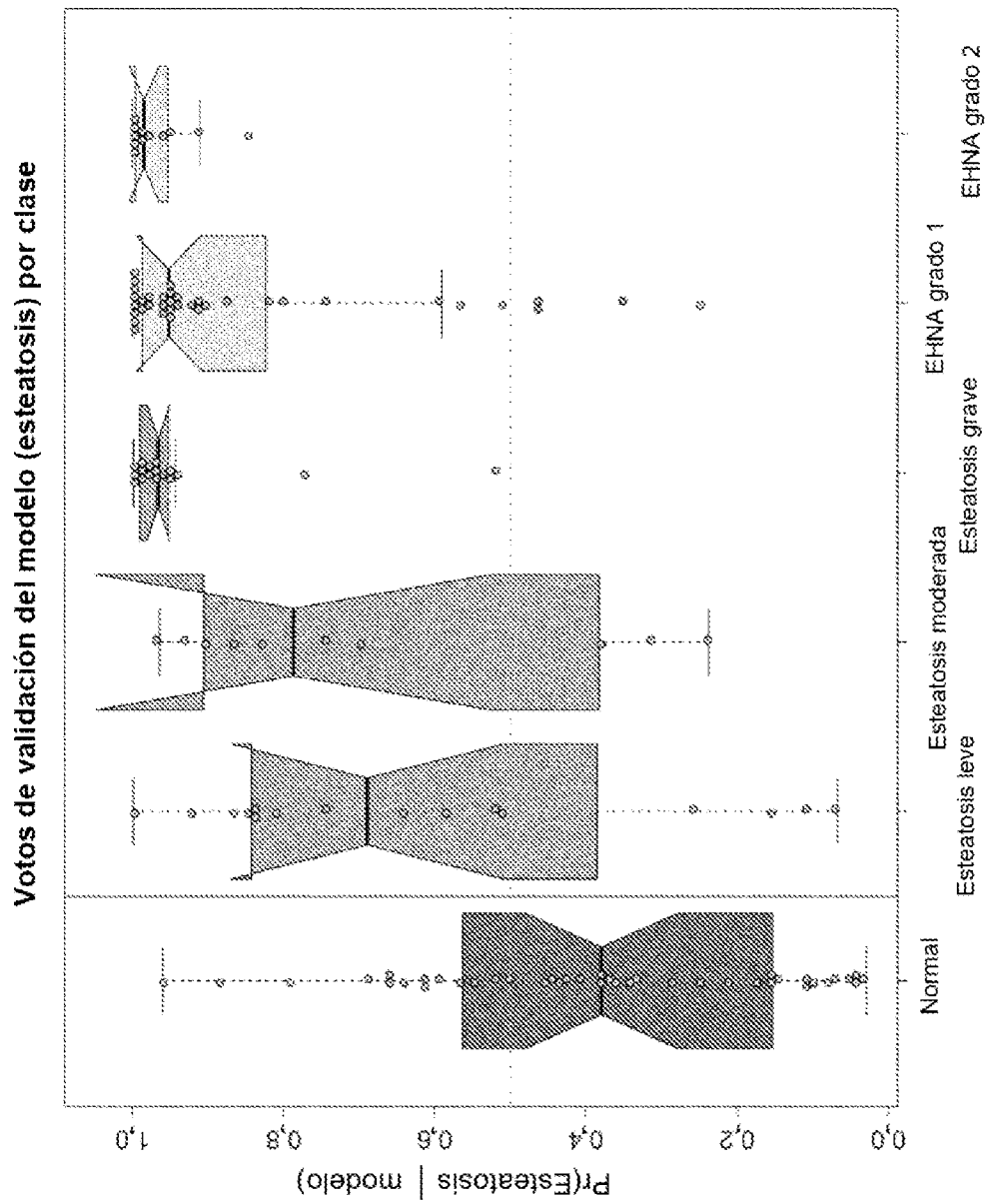


**FIG. 15**

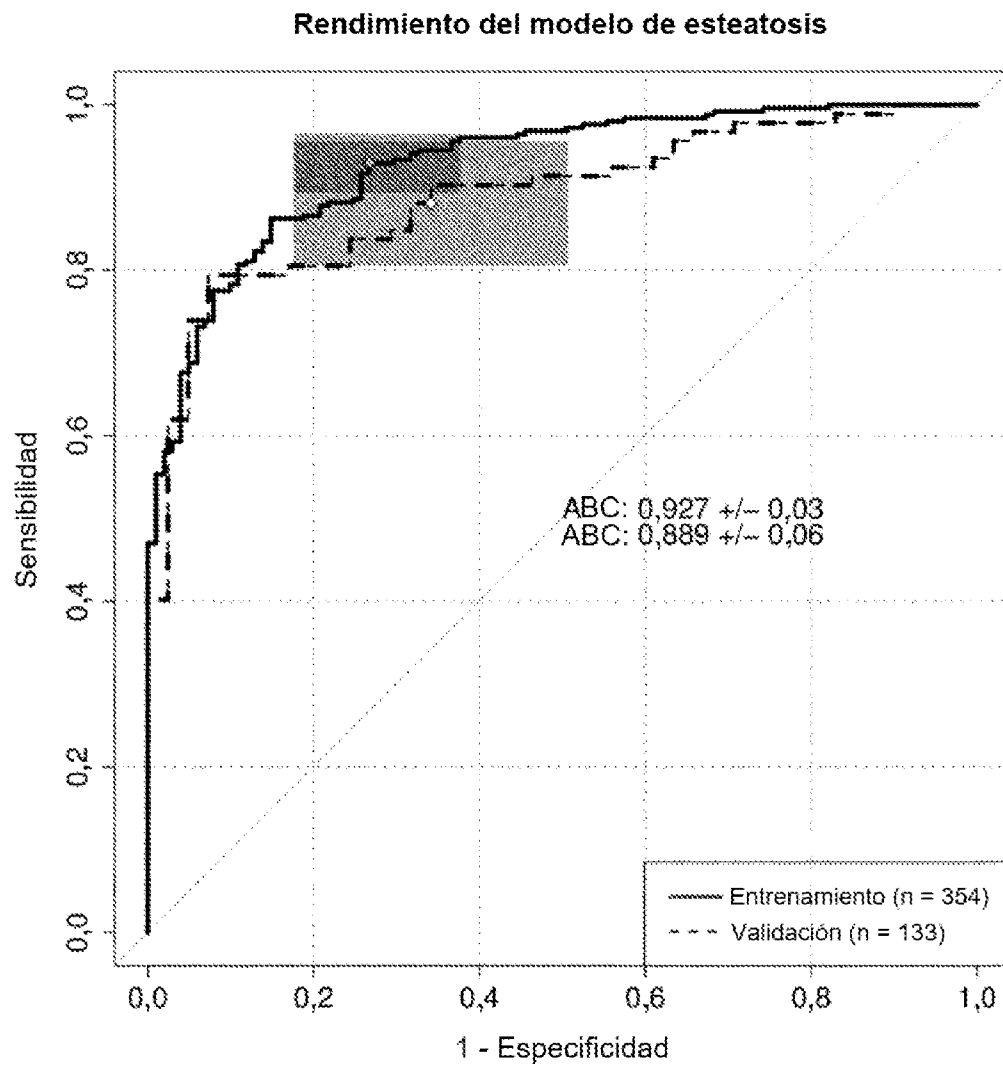


**FIG. 16**

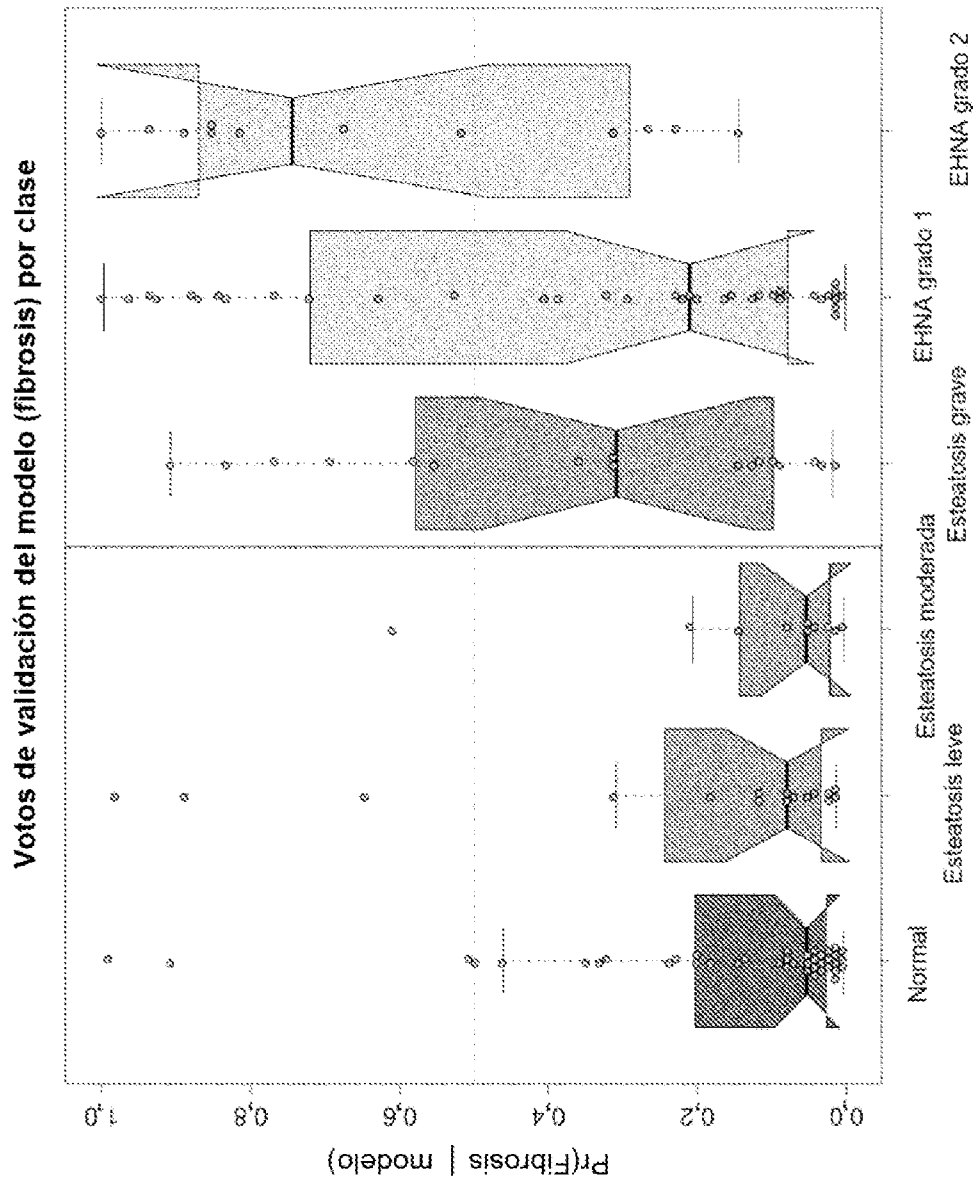


**FIG. 17**

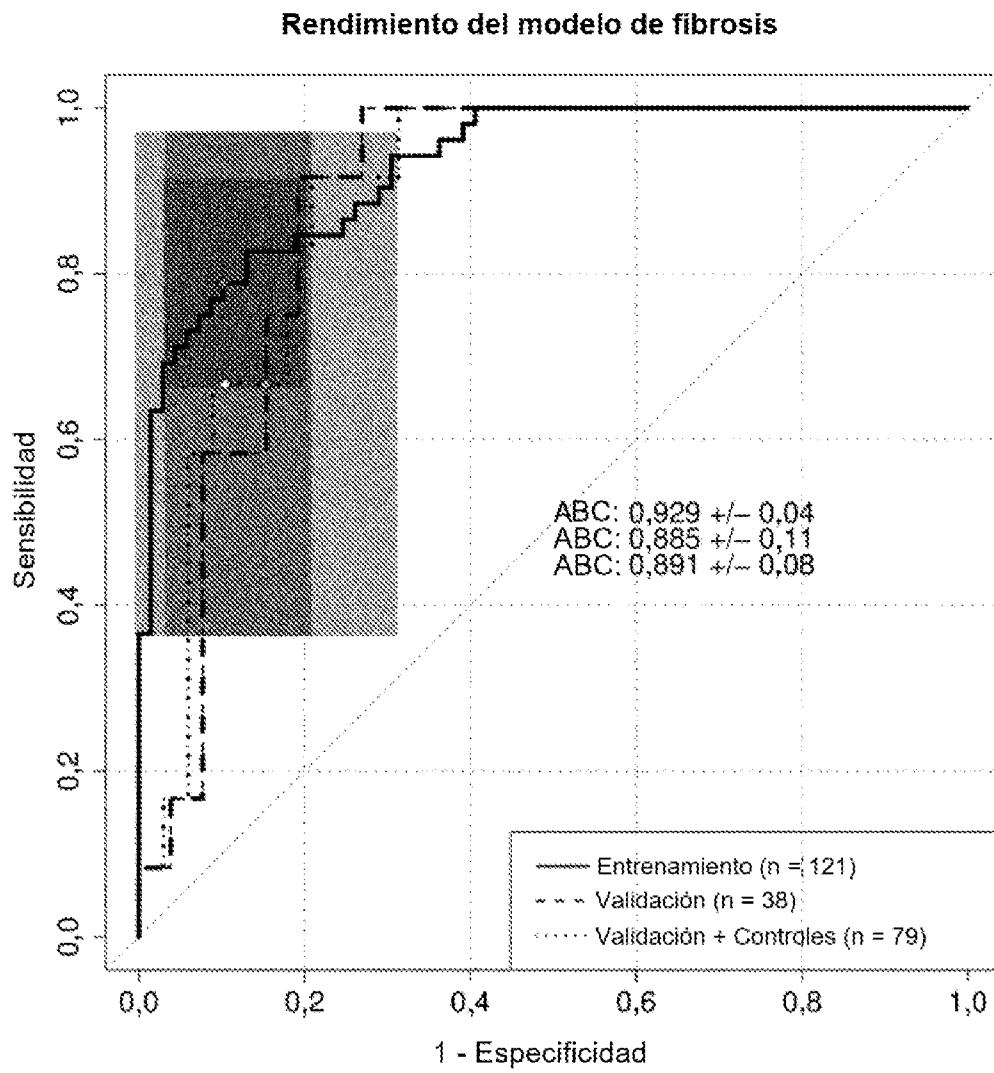


**FIG. 18**

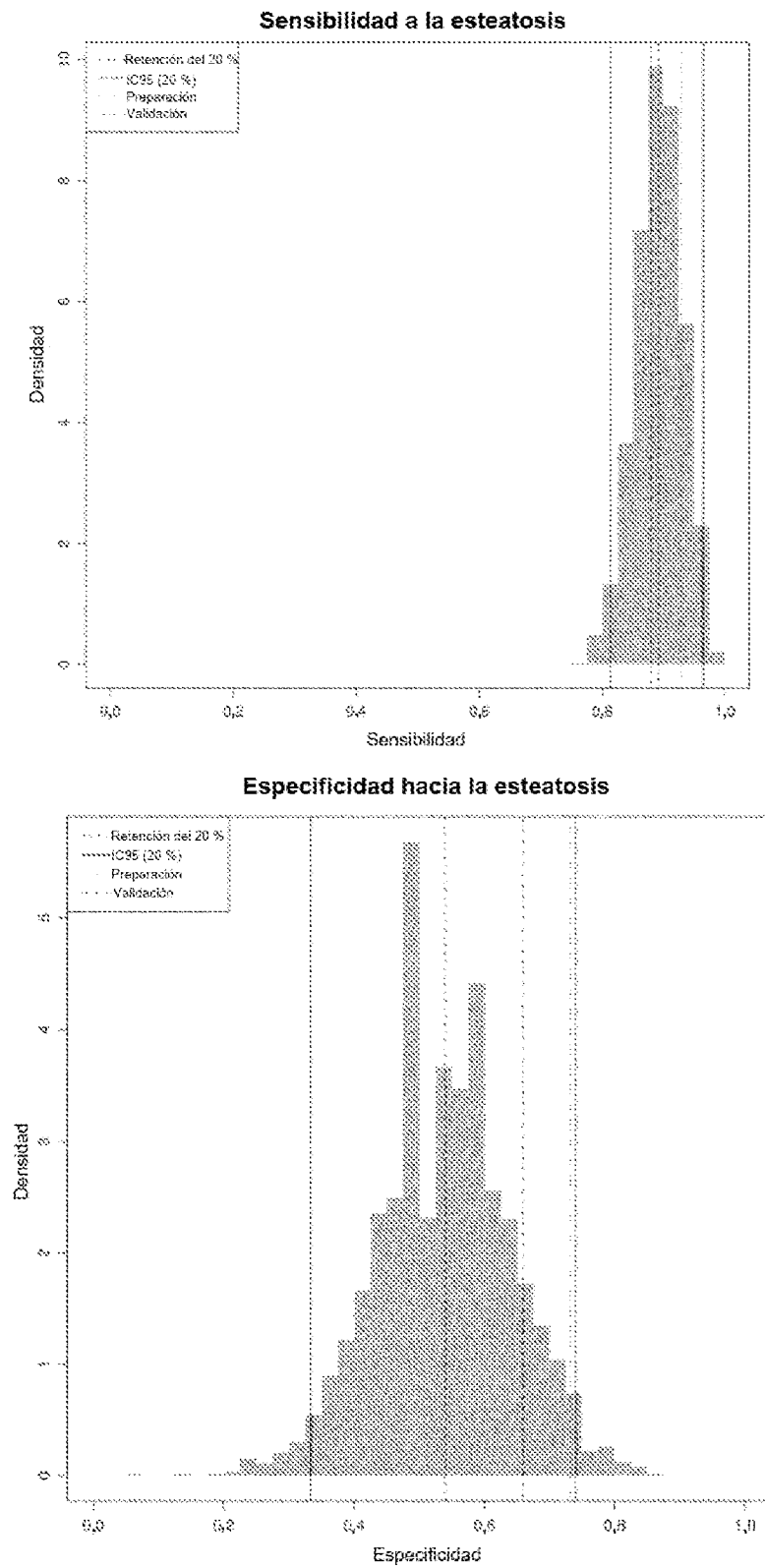


**FIG. 19A**

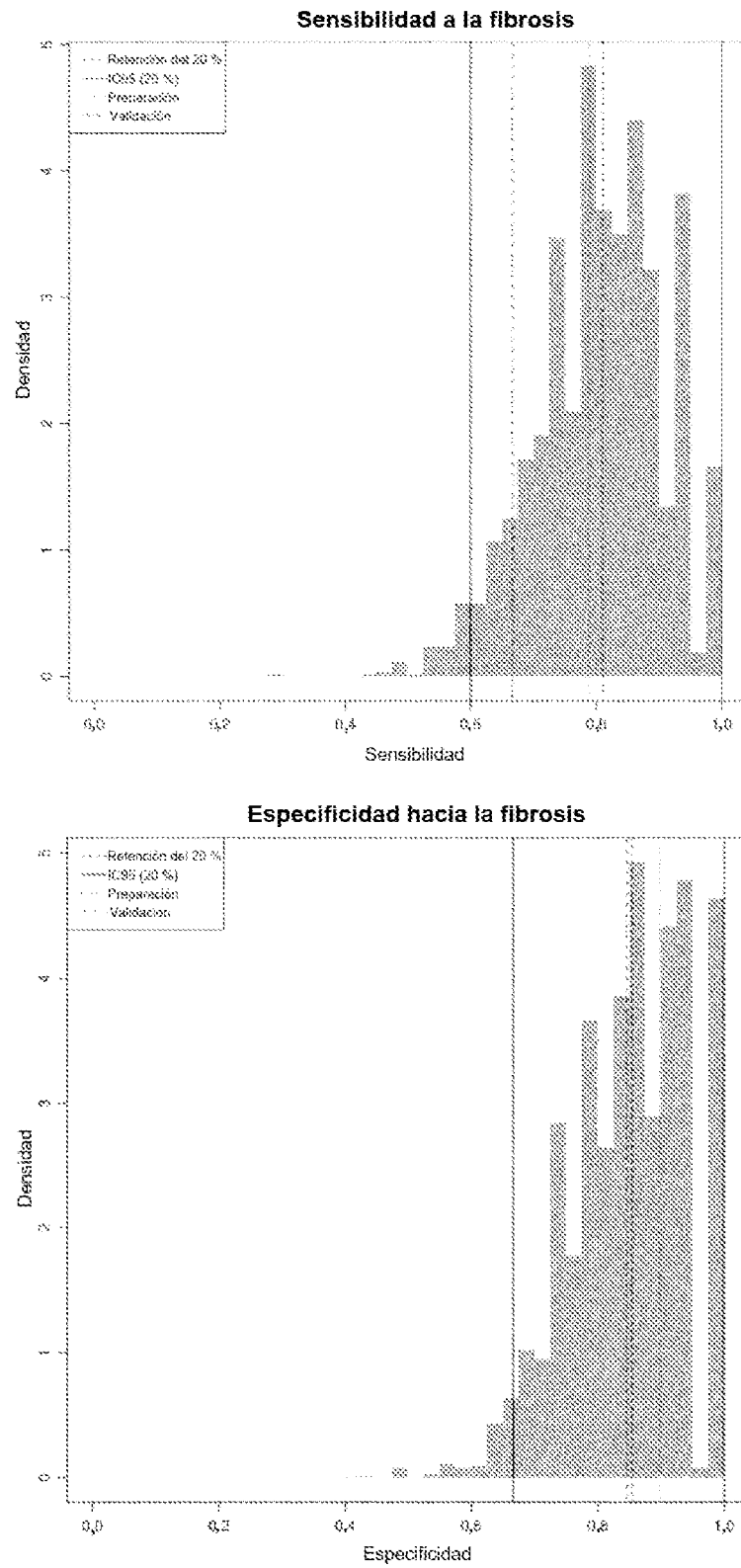


**FIG. 19B**

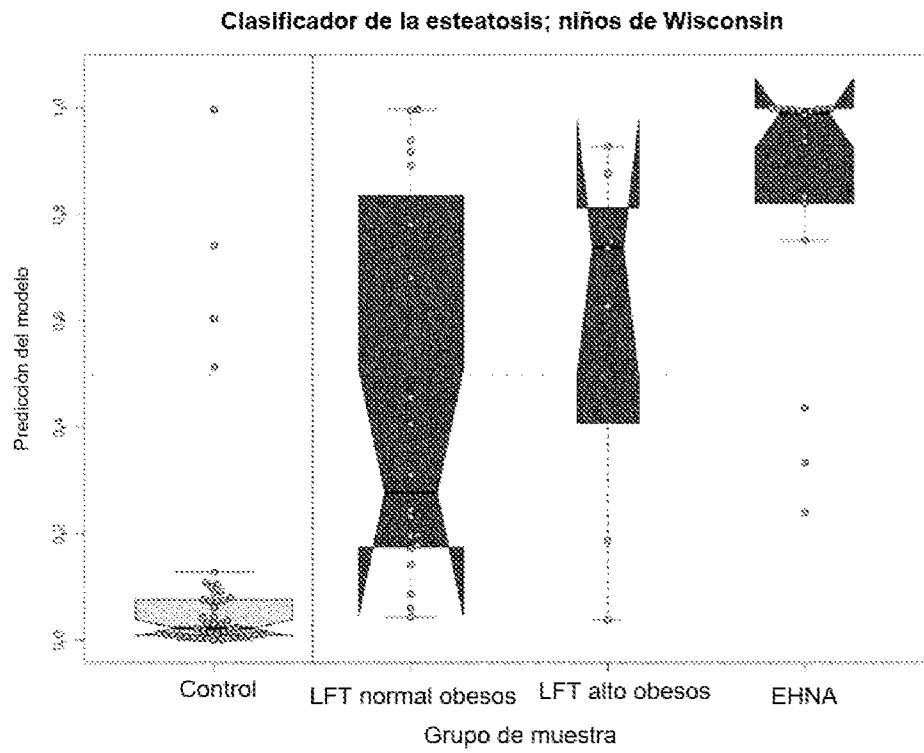


**FIG. 20**