

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 635 429**

51 Int. Cl.:

C12Q 1/68 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **13.11.2008 E 13151356 (6)**

97 Fecha y número de publicación de la concesión europea: **31.05.2017 EP 2584050**

54 Título: **Biomarcadores diagnósticos de la diabetes**

30 Prioridad:

13.11.2007 US 987540 P

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

03.10.2017

73 Titular/es:

**JANSSEN DIAGNOSTICS, LLC (100.0%)
700 US Highway 202
Raritan, NJ 08869, US**

72 Inventor/es:

**PALMA, JOHN F.;
BACKUS, JOHN W.;
ZHANG, YI;
WANG, YIXIN;
YU, JACK X.;
VENER, TATIANA;
DERECHO, CARLO y
LEE, DONGHO**

74 Agente/Representante:

IZQUIERDO BLANCO, María Alicia

ES 2 635 429 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).

Biomarcadores diagnósticos de la diabetes**DESCRIPCIÓN**

5

Campo de la solicitud

La solicitud se refiere al campo del diagnóstico médico y describe procedimientos y kits para el diagnóstico en el punto de atención de un estado de enfermedad diabética en un paciente.

10

Antecedentes de la invención

La diabetes es un grupo de enfermedades que se caracterizan por altos niveles de glucosa en la sangre como resultado de defectos en la producción de insulina, en la acción de la insulina, o en ambos. Si no se trata, puede causar muchas complicaciones serias a corto plazo, incluyendo síntomas de hipoglucemia, cetoacidosis o coma hiperosmolar no cetótico. A largo plazo, se sabe que la diabetes contribuye a un mayor riesgo de arteriosclerosis, insuficiencia renal crónica, daño a la retina (incluyendo ceguera), daño neural y daño microvascular.

15

Se prevé que la propagación epidémica de diabetes en el mundo en desarrollo tenga un profundo impacto en el sistema de salud de Estados Unidos. En un estudio reciente realizado por el Centro para el Control y la Prevención de Enfermedades se indica que la incidencia de nuevos casos de diabetes en Estados Unidos casi se ha duplicado en los últimos 10 años. A fecha de 2007, al menos 57 millones de personas en los Estados Unidos tienen pre-diabetes. Junto con los casi 24 millones que ya tienen diabetes, esto coloca a más del 25 % de la población de EE.UU. en riesgo de sufrir nuevas complicaciones de esta enfermedad. Según la Asociación Americana de Diabetes, el coste estimado de la diabetes en Estados Unidos en 2007 ascendió a 174 mil millones de dólares, con unos costes médicos directos cercanos a los 116 mil millones de dólares.

20

25

Aunque la etiología de la diabetes parece ser de naturaleza multifactorial, el aumento de la evidencia experimental sugiere que la aparición de obesidad, especialmente obesidad abdominal, altera la homeostasis inmune y metabólica y, en última instancia, conduce a una amplia respuesta inflamatoria. La producción de citocinas inflamatorias en el tejido adiposo, tal como el TNF alfa, altera la regulación de la respuesta inmunitaria y la capacidad de una célula para responder a la insulina. La detección de una alteración en los perfiles transcripcionales de las células inmunitarias circulantes, tales como monocitos y macrófagos, proporciona, por lo tanto, una vía conveniente para diagnosticar la enfermedad y monitorizar su progresión antes incluso de que aparezcan los signos más evidentes de intolerancia a la glucosa.

30

35

Por las razones expuestas, existe una necesidad insatisfecha de ensayos diagnósticos rápidos y precisos para el diagnóstico y la monitorización de pacientes con riesgo de desarrollar diabetes. En particular, existe una necesidad insatisfecha de ensayos de diagnóstico en un formato de kit que pueda usarse fácilmente en un centro de punto de atención para el cribado de rutina de diabetes de inicio temprano en los pacientes.

40

El documento WO 2007/128884 se refiere a un procedimiento de diagnóstico de la diabetes de tipo 2. El documento EP1538219 se refiere a procedimientos que se supone que pueden utilizarse para diagnosticar o detectar diabetes mellitus de tipo 1 (DM1) en un sujeto mediante el ensayo de un marcador para la DM1. El documento WO 2008/074963 Se refiere a un procedimiento *in vitro* para el diagnóstico precoz de diabetes a partir de una muestra biológica de un paciente.

45

Sumario de la solicitud

Se describen procedimientos para la determinación de perfiles de expresión de la firma génica que son diagnósticos de estados de enfermedad pre-diabéticos y diabéticos. La divulgación se refiere además a kits de diagnóstico que comprenden reactivos para la medición rápida de perfiles de expresión de firma génica en la muestra de sangre de un paciente. El formato del kit es rentable y cómodo para su uso en una instalación de punto de atención.

50

En una realización de la divulgación, se describe un procedimiento para el diagnóstico de la diabetes mellitus en un paciente, comprendiendo el procedimiento las etapas de (a) proporcionar una muestra de ensayo tomada de un paciente, (b) medir el perfil de expresión génica de una firma génica que comprende un gen seleccionado del grupo de los genes TOP1, CD24 y STAP1, (c) comparar el perfil de expresión génica con un perfil de expresión génica de diagnóstico de la firma génica, (d) determinar un estado patológico diabético en el paciente basado al menos en parte de una coincidencia sustancial entre el perfil de expresión génica y el perfil de expresión génica de diagnóstico y (e) mostrar la determinación a un profesional médico.

55

60

La etapa de determinación puede ejecutarse mediante un sistema informático que ejecute uno o más algoritmos seleccionados del grupo de combinación lineal de señales de expresión génica, modelo de regresión lineal, modelo de regresión logística, modelo de análisis de discriminación lineal (LDA), el modelo de vecino más próximo y el análisis de predicción de microarrays (PAM). La etapa de determinación también puede incluir el análisis del perfil de

65

enfermedad metabólica del paciente.

5 La firma génica puede incluir dos genes cuales quiera genes seleccionados del grupo de los genes TOP1, CD24 y STAP1 o los tres genes. En una realización, la firma génica puede incluir uno o más genes seleccionados del grupo de genes TOP1, CD24 y STAP1 y uno o más genes seleccionados de los genes enumerados en las TABLAS 1 o 6.

El paciente puede tener un IMC normal. El estado de enfermedad diabética puede ser un estado de enfermedad pre-diabética o un estado de enfermedad de diabetes de tipo 2.

10 La muestra de ensayo puede ser una muestra de sangre o una muestra de ensayo que contiene PBMC o células CD11c⁺ o CD11b⁺ o Emr⁺ o [CD11b⁺ CD11c⁺] o [Emr⁺ CD11b⁺] o [Emr⁺ CD11c⁺] o [Emr⁺ CD11b⁺ CD11c⁺] monocitos CD14⁺.

15 La medición puede implicar PCR en tiempo real, un ensayo inmunoquímico o una hibridación específica de oligonucleótidos.

20 En una realización de la divulgación, se describe un procedimiento para el diagnóstico de la diabetes mellitus en un paciente, comprendiendo el procedimiento las etapas de (a) proporcionar una muestra de ensayo tomada de un paciente, (b) medir el perfil de expresión génica de una firma génica que comprende los genes TULP4, AA741300, ESCO1, EIF5B, ACTR2, WNK1, COCH, SON, TPR y NOG, (c) comparar el perfil de expresión génica con un perfil de expresión génica de diagnóstico de la firma génica, (d) determinar un estado patológico diabético en el paciente basado al menos en parte de una coincidencia sustancial entre el perfil de expresión génica y el perfil de expresión génica de diagnóstico y (e) mostrar la determinación a un profesional médico.

25 La etapa de determinación puede ejecutarse mediante un sistema informático que ejecute uno o más algoritmos seleccionados del grupo de combinación lineal de señales de expresión génica, modelo de regresión lineal, modelo de regresión logística, modelo de análisis de discriminación lineal (LDA), el modelo de vecino más próximo y el análisis de predicción de microarrays (PAM). La etapa de determinación también puede incluir el análisis del perfil de enfermedad metabólica del paciente.

30 En un aspecto, la firma génica incluye los genes TOP1, CD24 y STAP1 además de los genes TULP4, AA741300, ESCO1, EIF5B, ACTR2, WNK1, COCH, SON, TPR y NOG. En otro aspecto, la firma génica incluye los genes TULP4, AA741300, ESCO1, EIF5B, ACTR2, WNK1, COCH, SON, TPR y NOG y uno o más genes seleccionados de los genes enumerados en los TABLAS 1 o 6.

35 El paciente puede tener un IMC normal. El estado de enfermedad diabética puede ser un estado de enfermedad pre-diabética o un estado de enfermedad de diabetes de tipo 2.

40 La muestra de ensayo puede ser una muestra de sangre o una muestra de ensayo que contiene PBMC o células CD11c⁺ o CD11b⁺ o Emr⁺ o [CD11b⁺ CD11c⁺] o [Emr⁺ CD11b⁺] o [Emr⁺ CD11c⁺] o [Emr⁺ CD11b⁺ CD11c⁺] monocitos CD14⁺.

45 La medición puede implicar PCR en tiempo real, un ensayo inmunoquímico o una hibridación específica de oligonucleótidos.

50 En una realización de la divulgación, se describe un procedimiento para el diagnóstico de la diabetes mellitus en un paciente, comprendiendo el procedimiento las etapas de (a) proporcionar una muestra de ensayo tomada de un paciente, (b) medir el perfil de expresión génica de una firma génica que comprende los genes TCF7L2 y CLC, (c) comparar el perfil de expresión génica con un perfil de expresión génica de diagnóstico de la firma génica, (d) determinar un estado patológico diabético en el paciente basado al menos en parte de una coincidencia sustancial entre el perfil de expresión génica y el perfil de expresión génica de diagnóstico y (e) mostrar la determinación a un profesional médico.

55 La etapa de determinación puede ejecutarse mediante un sistema informático que ejecute uno o más algoritmos seleccionados del grupo de combinación lineal de señales de expresión génica, modelo de regresión lineal, modelo de regresión logística, modelo de análisis de discriminación lineal (LDA), el modelo de vecino más próximo y el análisis de predicción de microarrays (PAM). La etapa de determinación también puede incluir el análisis del perfil de enfermedad metabólica del paciente.

60 La firma génica puede incluir el gen TCF7L2 o el gen CLC. En un aspecto, la firma génica incluye una o más variantes del gen TCF7L2 o CLC. En un aspecto, la firma génica incluye el gen TCF7L2 o CLC y uno o más genes seleccionados de los genes enumerados en las TABLAS 1 o 6.

65 El paciente puede tener un IMC normal. El estado de enfermedad diabética puede ser un estado de enfermedad pre-diabética o un estado de enfermedad de diabetes de tipo 2.

La muestra de ensayo puede ser una muestra de sangre o una muestra de ensayo que contiene PBMC o células CD11c⁺ o CD11b⁺ o Emr⁺ o [CD11b⁺ CD11c⁺] o [Emr⁺ CD11b⁺] o [Emr⁺ CD11c⁺] o [Emr⁺ CD11b⁺ CD11c⁺]o monocitos CD14⁺.

- 5 La medición puede implicar PCR en tiempo real, un ensayo inmunoquímico o una hibridación específica de oligonucleótidos.

10 En una realización de la divulgación, se describe un procedimiento para el diagnóstico de un cambio en estado de enfermedad diabética de un paciente, que comprende las etapas de (a) proporcionar una primera muestra de ensayo tomada de un paciente en un primer punto de tiempo, (b) medir el primer perfil de expresión de una firma génica que comprende un gen seleccionado del grupo de los genes TOP 1, CD24 y STAP1 en la primera muestra de ensayo, (c) comparar una segunda muestra de ensayo tomada del paciente en un segundo punto de tiempo, (d) medir un segundo perfil de expresión de la firma génica en la segunda muestra de ensayo, (e) comparar el primer perfil de expresión con el segundo perfil de expresión, (f) determinar un cambio en el estado de enfermedad diabética en el paciente basado, al menos en parte, en una diferencia sustancial entre el primer perfil de expresión génica y el segundo perfil de expresión génica, y (g) mostrar la determinación a un profesional médico.

15 En un aspecto, la etapa de determinación se ejecuta mediante un sistema informático que ejecute uno o más algoritmos seleccionados del grupo de combinación lineal de señales de expresión génica, modelo de regresión lineal, modelo de regresión logística, modelo de análisis de discriminación lineal (LDA), el modelo de vecino más próximo y el análisis de predicción de microarrays (PAM). En otro aspecto, la etapa de determinación también incluye un análisis del perfil de enfermedad metabólica del paciente.

20 La firma génica puede incluir dos genes cuales quiera genes seleccionados del grupo de los genes TOP1, CD24 y STAP1. En un aspecto, la firma génica incluye los genes TOP1, CD24 y STAP1. En otro aspecto, la firma génica incluye un gen seleccionado del grupo de los genes TOP1, CD24 y STAP1 y uno o más genes seleccionados de los genes enumerados en las TABLAS 1 o 6.

25 En un aspecto, el periodo de tiempo entre el primer punto de tiempo y el segundo punto de tiempo es de 0 a 2 años o de ¼ a 2 años o de ½ a 2 años o de 2 a 5 años, o de 5 a 10 años o más.

30 Un cambio en el estado de enfermedad diabética puede ser indicativo de una progresión hacia un estado de enfermedad pre-diabética o un estado de enfermedad de diabetes de tipo II. En un aspecto, el paciente en el primer punto de tiempo tiene un IMC normal.

35 La primera y segunda muestras de ensayo pueden ser muestras de sangre. En un aspecto, las primera y segunda muestras de ensayo pueden ser una muestra de ensayo que contiene PBMC o células CD11c⁺ o CD11b⁺ o Emr⁺ o [CD11b⁺ CD11c⁺] o [Emr⁺ CD11b⁺] o [Emr⁺ CD11c⁺] o [Emr⁺ CD11b⁺ CD11c⁺]o monocitos CD14⁺.

40 La medición puede implicar PCR en tiempo real, un ensayo inmunoquímico o una hibridación específica de oligonucleótidos.

45 En una realización, se describe un procedimiento para el diagnóstico de un cambio en estado de enfermedad diabética de un paciente, que comprende las etapas de (a) proporcionar una primera muestra de ensayo tomada de un paciente en un primer punto de tiempo, (b) medir el primer perfil de expresión de una firma génica que comprende un gen seleccionado del grupo de los genes TULP4, AA741300, ESCO1, EIF5B, ACTR2, WNK1, COCH, SON, TPR y NOG en la primera muestra de ensayo, (c) comparar una segunda muestra de ensayo tomada del paciente en un segundo punto de tiempo, (d) medir un segundo perfil de expresión de la firma génica en la segunda muestra de ensayo, (e) comparar el primer perfil de expresión con el segundo perfil de expresión, (f) determinar un cambio en el estado de enfermedad diabética en el paciente basado, al menos en parte, en una diferencia sustancial entre el primer perfil de expresión génica y el segundo perfil de expresión génica, y (g) mostrar la determinación a un profesional médico.

50 En un aspecto, la etapa de determinación se ejecuta mediante un sistema informático que ejecute uno o más algoritmos seleccionados del grupo de combinación lineal de señales de expresión génica, modelo de regresión lineal, modelo de regresión logística, modelo de análisis de discriminación lineal (LDA), el modelo de vecino más próximo y el análisis de predicción de microarrays (PAM). En otro aspecto, la etapa de determinación también incluye un análisis del perfil de enfermedad metabólica del paciente.

55 En otro aspecto, la firma génica incluye los genes TULP4, AA741300, ESCO1, EIF5B, ACTR2, WNK1, COCH, SON, TPR y NOG y uno o más genes seleccionados de los genes enumerados en los TABLAS 1 o 6.

60 En un aspecto, el periodo de tiempo entre el primer punto de tiempo y el segundo punto de tiempo es de 0 a 2 años o de ¼ a 2 años o de ½ a 2 años o de 2 a 5 años, o de 5 a 10 años o más.

65 Un cambio en el estado de enfermedad diabética puede ser indicativo de una progresión hacia un estado de

enfermedad pre-diabética o un estado de enfermedad de diabetes de tipo II. En un aspecto, el paciente en el primer punto de tiempo tiene un IMC normal.

5 La primera y segunda muestras de ensayo pueden ser muestras de sangre. En un aspecto, las primera y segunda muestras de ensayo pueden ser una muestra de ensayo que contiene PBMC o células CD11c⁺ o CD11b⁺ o Emr⁺ o [CD11b⁺ CD11c⁺] o [Emr⁺ CD11b⁺] o [Emr⁺ CD11c⁺] o [Emr⁺ CD11b⁺ CD11c⁺]o monocitos CD14⁺.

10 La medición puede implicar PCR en tiempo real, un ensayo inmunoquímico o una hibridación específica de oligonucleótidos.

15 En otra realización de la divulgación se describe un kit para evaluar la susceptibilidad de un paciente a la diabetes en el que se realiza la evaluación con un aparato de ensayo. El kit incluye (a) reactivos para recoger una muestra de ensayo de un paciente; y (b) reactivos para medir el perfil de expresión de una firma génica que comprende los genes TCF7L2 y CLC o variantes de los mismos en la muestra de ensayo de un paciente.

Los reactivos en las etapas (a) y (b) son suficientes para una pluralidad de pruebas. Los reactivos para recoger una muestra de ensayo de un paciente pueden envasarse en recipientes estériles.

20 La firma génica puede incluir uno o más de los genes seleccionados del grupo de genes TCF7L2 y CLC y uno o más genes seleccionados de la lista de genes de las TABLAS 1 o 6.

La muestra de ensayo puede ser una muestra de sangre.

25 El kit puede incluir también reactivos para el aislamiento de PBMC o reactivos para el aislamiento de células CD11c⁺ o CD11b⁺ o Emr⁺ o [CD11b⁺ CD11c⁺] o [Emr⁺ CD11b⁺] o [Emr⁺ CD11c⁺] o [Emr⁺ CD11b⁺ CD11c⁺] o reactivos para el aislamiento de monocitos CD14⁺. Los reactivos para medir el perfil de expresión de una firma génica pueden ser reactivos de PCR en tiempo real, reactivos de ensayo inmunoquímico o para hibridación de oligonucleótidos específicos.

30 En otra realización de la divulgación se describe un kit para evaluar la susceptibilidad de un paciente a la diabetes en el que se realiza la evaluación con un aparato de ensayo. El kit incluye (a) reactivos para recoger una muestra de ensayo de un paciente; y (b) reactivos para medir el perfil de expresión de una firma génica que comprende los genes TOP 1, CD24 y STAP1 o variantes de los mismos en la muestra de ensayo de un paciente.

35 La firma génica puede incluir dos genes cualesquiera genes seleccionados del grupo de los genes TOP1, CD24 y STAP1. En otro aspecto, la firma génica incluye uno o más de los genes seleccionados del grupo de genes TOP1, CD24 y STAP1. En otro aspecto, la firma génica incluye uno o más genes seleccionados del grupo de genes TOP1, CD24 y STAP1 y uno o más genes seleccionados de la lista de genes de las TABLAS 1 o 6.

40 Los reactivos en las etapas (a) y (b) son suficientes para una pluralidad de pruebas. Los reactivos para recoger una muestra de ensayo de un paciente pueden envasarse en recipientes estériles.

La muestra de ensayo puede ser una muestra de sangre.

45 El kit puede incluir también reactivos para el aislamiento de PBMC o reactivos para el aislamiento de células CD11c⁺ o CD11b⁺ o Emr⁺ o [CD11b⁺ CD11c⁺] o [Emr⁺ CD11b⁺] o [Emr⁺ CD11c⁺] o [Emr⁺ CD11b⁺ CD11c⁺] o reactivos para el aislamiento de monocitos CD14⁺. Los reactivos para medir el perfil de expresión de una firma génica pueden ser reactivos de PCR en tiempo real, reactivos de ensayo inmunoquímico o para hibridación de oligonucleótidos específicos.

50 En otra realización se describe un kit para evaluar la susceptibilidad de un paciente a la diabetes en el que se realiza la evaluación con un aparato de ensayo. El kit incluye (a) reactivos para recoger una muestra de ensayo de un paciente; y (b) reactivos para medir el perfil de expresión de una firma génica de los genes TULP4, AA741300, ESCO1, EIF5B, ACTR2, WNK1, COCH, SON, TPR y NOG en la muestra de ensayo de un paciente.

55 Los reactivos en las etapas (a) y (b) son suficientes para una pluralidad de pruebas. Los reactivos para recoger una muestra de ensayo de un paciente pueden envasarse en recipientes estériles.

60 La firma génica también puede incluir los genes TULP4, AA741300, ESCO1, EIF5B, ACTR2, WNK1, COCH, SON, TPR y NOG y uno o más genes seleccionados de la lista de genes de las TABLAS 1 o 6.

La muestra de ensayo puede ser una muestra de sangre.

65 El kit puede incluir también reactivos para el aislamiento de PBMC o reactivos para el aislamiento de células CD11c⁺ o CD11b⁺ o Emr⁺ o [CD11b⁺ CD11c⁺] o [Emr⁺ CD11b⁺] o [Emr⁺ CD11c⁺] o [Emr⁺ CD11b⁺ CD11c⁺] o reactivos para el aislamiento de monocitos CD14⁺. Los reactivos para medir el perfil de expresión de una firma génica pueden ser

reactivos de PCR en tiempo real, se definen en las reivindicaciones, o reactivos de ensayo inmunoquímico.

Debe entenderse que esta solicitud no está limitada a las realizaciones divulgadas en este Sumario y está destinada a cubrir modificaciones y variaciones que están dentro del alcance de las reivindicaciones.

5 Las realizaciones descritas anteriormente tienen muchas ventajas, incluyendo firmas genéticas nuevas para el diagnóstico precoz de un estado de enfermedad pre-diabética y la monitorización de pacientes que están en riesgo de desarrollar diabetes o que ya han adquirido la enfermedad. La divulgación describe también kits con reactivos e instrucciones para la prueba rentable y rápida de muestras de sangre por parte del personal médico en una
10 instalación de punto de atención.

Breve descripción de los dibujos

15 La figura 1 representa un análisis de curva ROC del gen CLC en comparación con OGTT de acuerdo con una primera realización;
la figura 2A representa un análisis de curva ROC de TCF7L2 conjunto 1 en comparación con OGTT de acuerdo con una segunda realización;
la figura 2B representa un análisis de curva ROC de TCF7L2 conjunto 1 en comparación con OGTT frente a FPG de acuerdo con una tercera realización;
20 la figura 3 muestra un análisis de curva ROC del gen CDKN1C de acuerdo con una cuarta realización;
la figura 4 muestra un análisis ROC de la firma de 3 genes en comparación con OGTT de acuerdo con una quinta realización;
la figura 4B muestra un análisis ROC de la firma de 3 genes en comparación con FPG frente a OGTT de acuerdo con una sexta realización;
25 la figura 4C representa un gráfico de barras de la expresión media de la firma de 3 genes de acuerdo con una séptima realización;
la figura 5A muestra un análisis ROC de la firma de 10 genes en comparación con OGTT de acuerdo con una octava realización;
la figura 5B muestra un análisis ROC de la firma de 10 genes en comparación con FPG frente a OGTT de
30 acuerdo con una novena realización; y
la figura 5C representa un gráfico de barras de la expresión media de la firma de 10 genes de acuerdo con una décima realización.

Descripción detallada

35 A menos que se defina otra cosa, todos los términos técnicos y científicos usados en el presente documento tienen el mismo significado que un experto en la técnica. Las siguientes definiciones se proporcionan para ayudar a interpretar la divulgación y las reivindicaciones de la presente solicitud. En el caso de que una definición en esta sección no sea consistente con las definiciones en otra parte, la definición establecida en esta sección tendrá
40 prioridad.

Además, la práctica de la invención emplea, a menos que se indique lo contrario, técnicas moleculares biológicas e inmunológicas convencionales dentro de experiencia en la materia. Tales técnicas son bien conocidas por el experto en la materia y se explican completamente en la bibliografía. Véase, por ejemplo, Colignan, Dunn, Ploegh, Speicher
45 y Wingfield "Current protocols in Protein Science" (1999–2008) Volumen I y II, incluyendo todos los suplementos (John Wiley & Sons Inc.); y Bailey, J. E. y Ollis, D. F., Biochemical Engineering Fundamentals, McGraw–Hill Book Company, NY, 1986; Ausubel, et al., ed., Current Protocols in Molecular Biology, John Wiley & Sons, Inc., NY, N.Y. (1987–2008), incluidos todos los suplementos; Sambrook, et al., Molecular Cloning: A Laboratory Manual, 2ª Edición, Cold Spring Harbor, N.Y. (1989); y Harlow y Lane, Antibodies, a Laboratory Manual, Cold Spring Harbor, N.Y. (1989). El análisis ROC se revisa en "An Introduction to ROC Analysis" by Tom Fawcett, Pattern Recognition Letters
50 27 (2006) 861–874.

Como se usa en el presente documento, "Diabetes Mellitus" se refiere a cualquier enfermedad caracterizada por una alta concentración de glucosa en sangre (hiperglucemia). La diabetes mellitus se diagnostica demostrando uno
55 cualquiera de los siguientes: un nivel de glucosa en plasma en ayunas de 126 mg/dl (7,0 mmol/l) o una glucosa en plasma igual o superior a 200 mg/dl (11.1 mmol/l) dos horas después de una carga de glucosa oral de 75 g como en una prueba de tolerancia a la glucosa o síntomas de hiperglucemia y glucosa en plasma ocasional en o por encima de 200 mg/dL (11,1 mmol).

60 Como se usa en el presente documento, la diabetes se refiere a la "diabetes tipo 1" también conocida como diabetes infantil, diabetes juvenil y diabetes insulino dependiente (IDDM) o "diabetes tipo 2", también conocida como diabetes de adultos, diabetes relacionada con la obesidad y diabetes no insulino dependiente (NIDDM) u otras formas de diabetes incluyen diabetes gestacional, diabetes tipo 1 resistente a la insulina (o "diabetes doble"), diabetes autoinmune latente de adultos (o LADA) y diabetes de la madurez de los jóvenes MODY), que es un grupo de varios
65 trastornos de un solo gen (monogénicos) con antecedentes familiares fuertes que se presentan como diabetes de tipo 2 antes de los 30 años de edad.

5 Como se usa en el presente documento, un "estado de enfermedad diabética" se refiere a un estado de enfermedad pre-diabética, estados de enfermedad diabética intermedia caracterizados por estados de la enfermedad más avanzados que el estado de enfermedad pre-diabética y a estados de enfermedad característicos de la diabetes franca como se define en el presente documento, Incluyendo diabetes tipo I o II.

10 Tal como se utiliza en el presente documento, un "estado de enfermedad pre-diabética" es aquel en el que un paciente tiene un nivel de glucosa en ayunas alterado y una tolerancia a la glucosa deteriorada. Una glucosa en ayunas alterada se define como un nivel de glucosa en sangre de 100 a 125 mg/dl (6,1 y 7,0 mmol/l), es decir, una glucosa en ayunas alterada. Se considera que los pacientes con glucosa en plasma igual o superior a 140 mg/dL o 7,8 mmol/l, pero no más de 200, dos horas después de una carga oral de glucosa de 75 g, tienen una tolerancia alterada a la glucosa.

15 Tal como se usa en este documento, un "profesional médico" es un médico o un técnico médico entrenado o un profesional de enfermería en una instalación en el punto de atención.

20 Una instalación en el "punto de atención" puede estar en un lugar de hospitalización, tal como en un hospital o en un centro de atención ambulatoria, tal como la consulta del médico o una clínica ambulatoria. En una realización, el ensayo de diagnóstico puede distribuirse como kit comercial a consumidores junto con instrumentos para el análisis del perfil de expresión de la firma génica en una muestra de sangre. En otra realización, el kit comercial puede combinarse con instrumentos y reactivos para la monitorización de los niveles de glucosa en sangre.

25 El término "nivel de glucosa en sangre" se refiere a la concentración de glucosa en la sangre. El nivel normal de glucosa en sangre (euglucemia) es de aproximadamente 120 mg/dl. Este valor fluctúa tanto como 30 mg/dl en los no diabéticos.

30 El trastorno de "hiperglucemia" (nivel alto de azúcar en la sangre) es un trastorno en la que el nivel de glucosa en la sangre es demasiado alto. Típicamente, la hiperglucemia ocurre cuando el nivel de glucosa en sangre se eleva por encima de 180 mg/dl.

35 Como se usa en el presente documento, una "muestra de ensayo" es cualquier muestra biológica de un paciente que contiene células que expresan diferencialmente genes en respuesta a un estado de enfermedad diabética. La muestra biológica puede ser cualquier material biológico aislado de un mamífero atópico o no atópico, tal como un ser humano, incluyendo un componente celular de la sangre, la médula ósea, el plasma, el suero, la linfa, el líquido cefalorraquídeo u otras secreciones, tales como lágrimas, saliva, o leche; muestras de biopsia de tejidos o órganos; o células cultivadas. Preferentemente, la muestra biológica es una muestra celular que puede obtenerse de un paciente con una intervención mínima. En una realización preferida, una muestra de ensayo es una muestra de sangre o una preparación de PBMC (células mononucleares de sangre periférica) o monocitos CD14 + o células CD11b + o CD11c + o Emr⁺.

40 El mamífero puede ser un ser humano o puede ser un animal doméstico o de compañía o zoológico. Aunque se contempla particularmente, las herramientas de diagnóstico descritas en el presente documento son adecuadas para su uso en el tratamiento médico de seres humanos, son también aplicables al tratamiento veterinario, incluyendo el tratamiento de animales de compañía, tales como perros y gatos, y animales domésticos tales como caballos, ganado bovino y ovejas o animales de zoológico, tales como primates no humanos, felinos, cánidos, bóvidos y ungulados.

45 Como se usa en el presente documento, la expresión "expresión génica" se refiere al proceso de convertir la información genética codificada en un gen en ARN (por ejemplo, ARNm, ARNr, ARNt o ARNsn) mediante la "transcripción" del gen (es decir, mediante la acción enzimática de una ARN polimerasa) y para genes de codificación de proteínas, en la proteína mediante la "traducción" del ARNm.

50 "Perfil de expresión génica" se refiere a niveles de expresión identificados de al menos un polinucleótido o proteína expresada en una muestra biológica.

55 El término "cebador", tal como se utiliza en el presente documento, se refiere a un oligonucleótido de origen natural (por ejemplo, como un fragmento de restricción) o que se produce sintéticamente, que es capaz de actuar como un punto de iniciación de la síntesis de un producto de extensión del cebador que es complementario a una hebra de ácido nucleico (molde o secuencia diana) cuando está en condiciones adecuadas (por ejemplo, tampón, sal, temperatura y pH) en presencia de nucleótidos y un agente para la polimerización de ácidos nucleicos, tal como polimerasa dependiente de ADN o dependiente de ARN. Un cebador debe ser lo bastante largo como para cebear la síntesis de los productos de extensión en presencia de un agente para la polimerización. Un cebador típico contiene al menos aproximadamente 10 nucleótidos de longitud de una secuencia sustancialmente complementaria u homóloga a la secuencia diana, pero se prefieren cebadores algo más largos. Normalmente, los cebadores contienen aproximadamente 15-26 nucleótidos.

Como se usa en el presente documento, una "firma génica" se refiere a un patrón de expresión génica de un conjunto seleccionado de genes que proporciona un identificador único de una muestra biológica. Una firma génica es un diagnóstico de un estado de enfermedad diabética si el patrón de expresión génica del conjunto seleccionado de genes es una concordancia sustancial con una firma génica en una muestra de referencia tomada de un paciente con un estado de enfermedad diabética. Para los fines de esta solicitud, una "firma génica" puede ser una combinación predeterminada de secuencias de ácido nucleico o polipéptido (si los genes son genes que codifican proteínas). Las firmas génicas pueden comprender genes de función desconocida o genes sin marcos de lectura abiertos que incluyen, pero sin limitaciones, ARNr, ARNUsn, microARN o ARNt.

Como se usa en el presente documento, un "perfil de expresión génica de diagnóstico" se refiere al perfil de expresión génica de una firma génica en una muestra biológica tomada de un paciente diagnosticado con un estado de enfermedad particular. El estado de enfermedad puede ser un estado de enfermedad diabética o un estado de enfermedad no diabética. Una "coincidencia sustancial" entre un perfil de expresión génica de ensayo del paciente y un perfil de expresión génica de diagnóstico característico de un estado de enfermedad diabética indica que el paciente tiene un estado de enfermedad diabética. Como alternativa, una "coincidencia sustancial" entre un perfil de expresión génica de ensayo del paciente y un perfil de expresión génica de diagnóstico característico de un estado de enfermedad diabética indica que el paciente no tiene un estado de enfermedad diabética.

Como se usa en el presente documento, una "variante" de un gen significa secuencias génicas que son al menos aproximadamente el 75 % idénticas a ellas, más preferentemente al menos aproximadamente 85 % idénticas y, lo más preferentemente, al menos 90 % idénticas y, aún más preferentemente, al menos aproximadamente 95- 99 % idénticas cuando estas secuencias de ADN se comparan con una secuencia de ácido nucleico del gen de tipo salvaje prevalente. En una realización, una variante de un gen es un gen con una o más alteraciones en la secuencia de ADN del gen que incluye, pero sin limitaciones, mutaciones puntuales o polimorfismos de un solo nucleótido, deleciones, inserciones, reordenamientos, mutaciones de corte y empalme del sitio de donante o aceptor y alteraciones génicas características de un pseudogén. A lo largo de esta solicitud, un gen incluye implícitamente formas tanto de tipo salvaje como variantes del gen como se define en el presente documento.

Tal como se utiliza en el presente documento, una "coincidencia sustancial" se refiere a la comparación del perfil de expresión génica de una firma génica en una muestra de ensayo con el perfil de expresión génica de la firma génica en una muestra de referencia tomada de un paciente con un estado patológico definido. Los perfiles de expresión "sustancialmente coincidentes" si la expresión de la firma génica en la muestra de ensayo y la muestra de referencia están sustancialmente en los mismos niveles, es decir, no hay diferencias estadísticamente significativas entre las muestras después de la normalización de las muestras. En una realización, el intervalo de confianza de los perfiles de expresión sustancialmente coincidentes es de al menos aproximadamente el 50 % o de aproximadamente el 50 % a aproximadamente el 75 % o de aproximadamente el 75 % a aproximadamente el 80 % o de aproximadamente el 80 % a aproximadamente el 85 % o de aproximadamente el 85 % a aproximadamente el 90 % o de aproximadamente el 90 % a aproximadamente el 95 %. En una realización preferida, el intervalo de confianza de perfiles de expresión sustancialmente coincidentes es de aproximadamente 95 % a aproximadamente 100 %. En otra realización preferida, el intervalo de confianza de perfiles de expresión sustancialmente coincidentes es cualquier número entre aproximadamente 95 % a aproximadamente 100 %. En otra realización preferida, el intervalo de confianza de perfiles de expresión sustancialmente coincidente es aproximadamente 95 % o aproximadamente 96 % o aproximadamente 97 % o aproximadamente 98 % o aproximadamente 99 % o aproximadamente 99,9 %.

Tal como se utiliza en el presente documento, una "diferencia sustancial" se refiere a la diferencia en el perfil de expresión génica de un gen en un punto de tiempo con el perfil de expresión génica de la misma firma génica en un segundo punto de tiempo. Los perfiles de expresión son "sustancialmente diferentes" si la expresión de la firma génica en el primer y segundo puntos de tiempo están en diferentes niveles, es decir, existe una diferencia estadísticamente significativa entre las muestras después de la normalización de las muestras. En una realización, los perfiles de expresión son "sustancialmente diferentes" si la expresión de la firma génica en el primer y segundo puntos de tiempo está fuera del intervalo de confianza calculado. En una realización, el intervalo de confianza de los perfiles de expresión sustancialmente diferentes es menor de aproximadamente 50 % o menor de aproximadamente 75 % o menor de aproximadamente 80 % o menor de aproximadamente 85 % o menor de aproximadamente 90 % o menor de aproximadamente 95 %.

Un intervalo de confianza IC del 95 % es igual a $AUC + 1,96 \times \text{error estándar de la AUC}$, donde AUC es el área bajo la curva ROC.

Tal como se usa en el presente documento, "ROC" se refiere a una característica operativa del receptor, o simplemente curva ROC, que es un gráfico de la sensibilidad frente a (1-especificidad) para un sistema clasificador binario a medida que varía su umbral de discriminación.

Como se usa en el presente documento, los términos "diagnóstico" o "diagnosticar" se refiere al procedimiento para distinguir un estado de enfermedad diabética de otro estado de enfermedad diabética o determinar si un estado de enfermedad diabética está presente en un paciente (atópico) con respecto al estado normal o "no diabético" (no atópico), y/o determinar la naturaleza de un estado de enfermedad diabética.

Tal como se utiliza en el presente documento, la "determinación de un estado de enfermedad diabética" se refiere a una integración de toda la información que es útil para diagnosticar un paciente con un estado o afección de enfermedad diabética y/o clasificar la enfermedad. Esta información incluye, pero sin limitaciones, antecedentes familiares, datos de genética humana, IMC, actividad física, perfil de enfermedad metabólica y los resultados de un análisis estadístico de los perfiles de expresión de una o más firmas génicas una muestra de ensayo tomada de un paciente. En el entorno del punto de atención, esta información es analizada y mostrada por un sistema informático que tiene un software de análisis de datos apropiado. La integración de los datos clínicos proporciona al médico asistente la información necesaria para determinar si el paciente tiene una afección diabética, información relacionada con la naturaleza o la clasificación de la diabetes, así como información relacionada con el pronóstico y/o información útil para seleccionar un tratamiento apropiado. En una realización, los ensayos de diagnóstico descritos en el presente documento proporcionan al profesional médico una determinación de la eficacia del tratamiento médico prescrito.

Como se usa en el presente documento, un "perfil de enfermedad metabólica" se refiere a cualquier número de medidas metabólicas estándar y otros factores de riesgo que pueden ser diagnósticos de un estado de enfermedad diabética incluyendo, pero sin limitaciones, glucosa en plasma en ayunas, insulina, proinsulina, péptido c, Insulina intacta, IMC, circunferencia de cintura, GLP-1, adiponectina, PAI-1, hemoglobina A1c, HDL, LDL, VLDL, triglicéridos, ácidos grasos libres. El perfil de enfermedad metabólica se puede utilizar para generar un modelo superior para la equivalencia de la clasificación a OGTT de 2 horas.

Una prueba de tolerancia a la glucosa es la administración de glucosa para determinar la rapidez con que se elimina de la sangre. La prueba suele usarse para detectar diabetes, resistencia a la insulina y, a veces, hipoglucemia reactiva. La glucosa generalmente se administra por vía oral, por lo que la prueba habitual es técnicamente una prueba oral de tolerancia a la glucosa (OGTT).

La prueba de glucosa en plasma en ayunas (FPG) es una prueba del metabolismo de los hidratos de carbono que mide los niveles de glucosa en plasma o sangre después de ayunar. El ayuno estimula la liberación de la hormona glucagón, que a su vez aumenta los niveles de glucosa en plasma. En personas sin diabetes, el cuerpo producirá y procesará la insulina para contrarrestar el aumento de los niveles de glucosa. En personas con diabetes, esto no sucede y los niveles de glucosa analizados se mantendrán altos.

Tal como se utiliza en el presente documento, el índice de masa corporal (IMC), o índice Quetelet, es una medida estadística que compara el peso y la altura de una persona. Debido a su facilidad de medición y cálculo, es la herramienta diagnóstica más utilizada para identificar la obesidad.

Como se usa en el presente documento, NGT significa tolerancia normal a la glucosa, IGT significa tolerancia alterada a la glucosa y T2D significa diabetes de tipo 2.

Como se usa en el presente documento, CD11c⁺, CD11b⁺ Y Emr⁺ son marcadores de la superficie celular de monocitos/macrófagos humanos y células mieloides y sus precursores. En ratones, los marcadores de superficie de celular de monocitos/macrófagos y células mieloides más usados son F4/80 y CD11b, aunque se ha informado que los anticuerpos F4/80 y CD11b reaccionan con los eosinófilos y las células dendríticas y NK y otros subtipos de células T y B, respectivamente (Nguyen, et al. (2007) J Biol Chem 282, 35279–35292; Patsouris, et al. (2008) Cell Metab. 8, 301–309). El gen F4/80 en ratón es el ortólogo del gen Emr1 humano. El ortólogo humano para el gen CD11c de ratón es ITGAX, también denominado integrina, alfa X (subunidad 4 del receptor 3 del componente del complemento), SLEB6, OTTHUMP00000163299; leu M5, subunidad alfa; antígeno de la superficie de leucocitos p150,95, subunidad alfa; antígeno de la membrana mieloides, subunidad alfa; cadena alfa de la integrina p150 95 integrina (Cromosoma: 16; Localización: 16p11.2 Anotación: Cromosoma 16, NC_000016.8 (31274010..31301819) MIM: 151510, ID del gen:3687). El ortólogo humano para el gen CD11b de ratón es ITGAM o integrina, alfa M (subunidad 3 del receptor 3 del componente del complemento) también denominada CD11B, CR3A, MAC-1, MAC1A, MGC117044, MO1A, SLEB6, polipéptido alfa del antígeno del macrófagos; subunidad alfa-M del receptor de la adherencia de neutrófilos (cromosoma: 16; Localización: 16p11.2 Cromosoma 16, NC_000016.8 (31178789.31251714), MIM: 120980, ID del gen 3684). Las células CD11c⁺, CD11b⁺ y Emr⁺ y CD14⁺ se pueden purificar de PBMC mediante selección positiva usando el kit de aislamiento celular de sangre humana (StemCell Technologies). La pureza de las poblaciones de células aisladas (> 85 %) se confirma después mediante tinción con citometría de flujo de anticuerpos conjugados con fluorescencia al marcador de la superficie celular apropiada (BioLegend).

Tal como se usa en el presente documento, "PCR en tiempo real" se refiere a la reacción en cadena de la polimerasa en tiempo real, también llamada reacción en cadena de la polimerasa en tiempo real cuantitativa (Q-PCR/qPCR) o reacción cinética en cadena de la polimerasa. La PCR en tiempo real es una técnica de laboratorio basada en la reacción en cadena de la polimerasa, que se utiliza para amplificar y cuantificar simultáneamente una molécula de ADN dirigida. Permite la detección y cuantificación (como número absoluto de copias o cantidad relativa cuando se normaliza a la entrada de ADN o genes normalizadores adicionales) de una secuencia específica en una muestra de ADN.

Como se usa en el presente documento, un ensayo inmunoquímico es un ensayo bioquímico que mide la concentración de una sustancia en un extracto celular usando la reacción de un anticuerpo o anticuerpos contra su antígeno. En esta divulgación, el antígeno es una proteína expresada por uno cualquiera de los genes codificantes de la proteína que comprende una firma génica. En una realización preferida, el ensayo inmunoquímico es un ensayo de inmunosorción ligado a enzimas (ELISA).

Tal como se utiliza en el presente documento, "hibridación de oligonucleótidos específicos" se refiere a la hibridación entre las secuencias de la sonda sobre un soporte sólido, tal como un chip y secuencias de ADNc generadas a partir de transcripciones dentro de la muestra de ensayo del paciente. Si las dos secuencias de ácido nucleico son sustancialmente complementarias, se produce hibridación que es directamente proporcional a la cantidad de secuencias de ADNc en la muestra de ensayo. A continuación, la detección de la hibridación se consigue utilizando técnicas bien conocidas en la materia. Numerosos factores influyen en la eficacia y la selectividad de la hibridación de dos ácidos nucleicos, por ejemplo, un miembro de ácido nucleico en una matriz, con una secuencia de ácido nucleico diana. Estos factores incluyen la longitud del miembro de ácido nucleico, la secuencia y/o la composición de nucleótidos, la temperatura de hibridación, la composición del tampón y el potencial de impedimento estérico en la región a la que se requiere que el miembro de ácido nucleico se hibride. Existe una correlación positiva entre la longitud del miembro de ácido nucleico y tanto la eficiencia como la precisión con la que un miembro de ácido nucleico se hibrida a una secuencia diana. En particular, las secuencias más largas tienen una temperatura de fusión (TM) más alta que las más cortas y es menos probable que se repitan dentro de una secuencia diana dada, minimizando de este modo la hibridación promiscua. La temperatura de hibridación varía inversamente con la eficacia de hibridación del miembro de ácido nucleico, al igual que la concentración de disolventes orgánicos, por ejemplo, formamida, que podrían incluirse en una mezcla de hibridación, mientras que los aumentos en la concentración de sales facilitan la unión. Bajo condiciones de hibridación rigurosas, los ácidos nucleicos más largos se hibridan más eficientemente que los más cortos, que son suficientes en condiciones más permisivas.

Tal como se utiliza en el presente documento, el término "anticuerpo" incluye anticuerpos tanto policlonales como monoclonales; y puede ser una molécula intacta, un fragmento del mismo (tal como fragmentos Fv, Fd, Fab, Fab' and F(ab)2 o multímeros o agregados de moléculas intactas y/o fragmentos; y puede encontrarse en la naturaleza o producirse, por ejemplo, mediante inmunización, síntesis o ingeniería genética.

Como se usan en el presente documento, todas las referencias a las sondas en las Tablas 1, 7, 8, 9, 10A y 10B se refieren a conjuntos de sondas representados en el GeneChip Human Genome U133 Plus 2.0 Array.

La siguiente descripción se refiere a ciertas realizaciones de la solicitud y a una metodología particular para diagnosticar un estado de enfermedad diabética en un paciente. En particular, la solicitud divulga una serie de genes, incluyendo algunos que no se había considerado previamente que estaban asociados con un estado de enfermedad diabética, se expresan diferencialmente en células mononucleares de sangre periférica (PMBC) de pacientes que tienen un estado de enfermedad diabética o pre-diabética en comparación con los pacientes que no tienen un estado de enfermedad diabética.

En una realización, los genes que se expresan diferencialmente en PBMC de NGT y T2D se identifican usando análisis de microarrays. Las transcripciones de PBMC de pacientes con NGT y T2D (en este ejemplo, una cohorte de 107 pacientes) se seleccionaron inicialmente usando el chip Affymetrix Human Genome HG-U133Plus2, de acuerdo con las instrucciones del fabricante. Se seleccionaron aproximadamente 200 genes expresados diferencialmente que tenían una tasa de falsos descubrimientos, FDR < 20 %, cambio > 1,7 entre NGT y T2D usando el programa de análisis de la significancia de Microarray (SAM) (véase la TABLA 1).

Los procedimientos de clasificación de la diabetes se describen a continuación determinando la expresión diferencial de diferentes combinaciones de genes de susceptibilidad a la diabetes identificados en la pantalla de microarrays inicial (véase la TABLA 12A). La Tabla 12A también incluye los números de acceso a Genbank de cada uno de los genes seleccionados.

La expresión génica de los genes de susceptibilidad a la diabetes se puede medir en una muestra biológica usando una serie de técnicas diferentes. Por ejemplo, la identificación del ARNm de los genes asociados a la diabetes dentro de una mezcla de diversos ARNm se lleva a cabo convenientemente mediante el uso de la reacción en cadena de la polimerasa-transcriptasa inversa (RT-PCR) y una sonda de hibridación oligonucleotídica que está marcada con un resto detectable.

Primero se extrae una muestra de ensayo de un paciente. Para obtener ARN de alta calidad es necesario minimizar la actividad de la ARNasa liberada durante la lisis celular. Esto se logra normalmente utilizando procedimientos de aislamiento que alteran los tejidos e inactivan o inhiben las ARNasas simultáneamente. Para los especímenes bajos en ribonucleasa endógena, los protocolos de aislamiento usan habitualmente tampones de extracción que contienen detergentes para solubilizar las membranas e inhibidores de la ARNasa, tales como inhibidores de ribonucleasa placentaria o complejos de vanadil-ribonucleósido. El aislamiento de ARN de muestras más exigentes, tales como tejidos intactos o células ricas en ribonucleasa endógena, requiere un enfoque más agresivo. En estos casos, el

tejido o las células se homogeneizan rápidamente en un potente desnaturalizante de proteínas (usualmente isotiocianato de guanidinio), para inactivar de forma irreversible las nucleasas y solubilizar las membranas celulares. Si una muestra de tejido no puede homogeneizarse rápidamente, debe congelarse rápidamente mediante inmersión en nitrógeno líquido y almacenarse a -80 °C. Las muestras congeladas de esta manera nunca deben descongelarse antes del aislamiento de ARN o el ARN se degradará rápidamente mediante la ARNasa liberada durante la lisis celular que se produce durante la congelación. El tejido debe sumergirse en una piscina de nitrógeno líquido y molerse hasta un polvo fino usando mortero y mano de mortero. Una vez pulverizado, el tejido todavía congelado se homogeneiza en tampón de extracción de ARN. Actualmente, una serie de kits para el aislamiento de ARN están disponibles comercialmente (Ambion, Quiagen).

Como es bien conocido en la técnica, primero se genera el ADNc mediante transcripción inversa de una primera cadena de ADNc a partir de un ARNm molde usando una ADN polimerasa dependiente de ARN y un cebador. Las transcriptasas inversas útiles de acuerdo con la solicitud incluyen, pero sin limitaciones, transcriptasas inversas de VIH, HTLV-1, HTLV-II, FeLV, FIV, SIV, AMV, MMTV, MoMuLV y otros retrovirus (para revisiones véase, por ejemplo, Levin, 1997, Cell 88:5-8; Verma, 1977, Biochim. Biophys. Acta 473:1-38; Wu et al., 1975, CRC Crit. Rev. Biochem. 3:289-347). Más recientemente, actualmente hay disponibles comercialmente una serie de kits para reacciones de RT-PCR usando transcriptasa inversa termoestable, por ejemplo el kit GeneAmp® Thermo Stable rTth Reverse Transcriptase RNA PCR (Applied Biosystems).

La "reacción en cadena de polimerasa" o "PCR", tal como se usa en el presente documento, se refiere, generalmente, a un procedimiento para la amplificación de una secuencia de nucleótidos deseada *in vitro*, como se describe en las patentes de Estados Unidos n.º 4.683.202, 4.683.195, 4.800.159 y 4.965.188. La reacción de PCR implica una serie repetitiva de ciclos de temperatura y se realiza típicamente en un volumen de 10-100 µl. La mezcla de reacción comprende dNTP (cada uno de los cuatro desoxinucleótidos dATP, dCTP, dGTP y dTTP), cebadores, tampones, ADN polimerasa y molde de ácido nucleico. La reacción de PCR comprende proporcionar un conjunto de cebadores polinucleotídicos en los que un primer cebador contiene una secuencia complementaria a una región en una cadena de la secuencia del molde de ácido nucleico e inicia la síntesis de una cadena de ADN complementaria, y un segundo cebador contiene una secuencia complementaria a una región en una segunda hebra de la secuencia de ácido nucleico diana e inicia la síntesis de una cadena de ADN complementaria y se amplifica la secuencia del molde de ácido nucleico empleando una polimerasa de ácido nucleico como agente polimerizante dependiente del molde en condiciones que son permisivas para las etapas de ciclos de PCR de (i) hibridación de cebadores requerida para la amplificación a una secuencia de ácido nucleico diana contenida dentro de la secuencia molde, (ii) extensión de los cebadores en la que la polimerasa de ácido nucleico sintetiza un producto de extensión del cebador.

Otros procedimientos de amplificación incluyen, pero sin limitaciones, reacción en cadena de la ligasa (LCR), amplificación basada en un polinucleótido específico (NSBA).

Los expertos en la materia pueden diseñar y sintetizar fácilmente cebadores para la región de ácido nucleico de interés. Se apreciará que los cebadores adecuados para su uso con la solicitud se pueden diseñar usando cualquier procedimiento adecuado. La selección del cebador para la PCR se describe, por ejemplo, en la patente de Estados Unidos n.º 6.898.531, presentada el 24 de mayo de 2005, titulada "Algoritmos para la selección de pares de cebadores" y el documento de Estados Unidos con n.º de serie 10/236.480, presentado el 5 de septiembre de 2002; para la PCR de corto alcance, el documento de Estados Unidos con n.º de serie 10/341.832, presentado el 14 de enero de 2003 proporciona orientación con respecto a la selección del cebador. También hay programas disponibles públicamente, tales como "Oligo", LASERGENE®, primer 5 (disponible en la página web de la empresa Premier Biosoft) y primer3 (disponible en la página web del Instituto Whitehead Institute for Biomedical Research, Cambridge, Mass, EE.UU.). El diseño del cebador se basa en una serie de parámetros, tales como la temperatura óptima de fusión (Tm) para las condiciones de hibridación que se van a utilizar y la longitud deseada de la sonda oligonucleotídica. Además, el diseño de oligonucleótidos intenta minimizar las estructuras secundarias potenciales que una molécula podría contener, tales como estructuras de horquilla y dímeros entre sondas, siendo el objetivo maximizar la disponibilidad de la sonda resultante para la hibridación. En una realización preferida, los cebadores usados en el procedimiento de PCR serán complementarios a las secuencias de nucleótidos dentro del molde de ADNc y, preferentemente, sobre los límites de exón-intrón.

En una realización, la reacción de PCR puede usar cebadores de PCR anidada.

En una realización, se puede incluir un marcador detectable en una reacción de amplificación. Los marcadores adecuados incluyen fluorocromos, por ejemplo isotiocianato de fluoresceína (FITC), rodamina, rojo Texas, ficoeritrina, alofocianina, 6-carboxifluorexceína (6-FAM), 2',7'-dimetoxi-4',5'-dicloro-6-carboxifluoresceína (JOE) 6-carboxi-2',4',7',4,7-hexaclorofluoresceína (HEX), 5-carboxifluoresceína (5-FAM) o N, N, N', N'- tetrametil-6-carboxirrodamina (TAMRA), marcadores radiactivos, por ejemplo ³²P, ³⁵S, ³H; así como otros. El marcador puede ser un sistema de dos etapas, en el que el ADN amplificado está conjugado a biotina, haptenos o similares, que tienen una pareja de unión de alta afinidad, por ejemplo, avidina, anticuerpos específicos, etc., en la que la pareja de unión está conjugada a un marcador detectable. El marcador puede conjugarse con uno o ambos de los cebadores. Como alternativa, el grupo de nucleótidos utilizado en la amplificación está marcado, de manera que incorpora el marcador

en el producto de amplificación.

En una realización particularmente preferida, la solicitud utiliza un sistema combinado de sonda de hibridación y PCR para aprovechar los sistemas de ensayo, tales como el uso de sondas FRET como se divulga en las patentes de Estados Unidos n.º 6.140.054 y 6.174.670. En una de sus configuraciones más simples, el enfoque FRET o "transferencia de energía de resonancia fluorescente" emplea dos oligonucleótidos que se unen a sitios adyacentes en la misma cadena del ácido nucleico que se está amplificando. Un oligonucleótido está marcado con un fluoróforo donante que absorbe luz a una primera longitud de onda y emite luz en respuesta, y el segundo está marcado con un fluoróforo aceptor que es capaz de brillar en respuesta a la luz emitida del primer donante (pero no sustancialmente por la fuente de luz que excita el primer donante y cuya emisión se puede distinguir de la del primer fluoróforo). En esta configuración, el segundo o fluoróforo aceptor muestra un aumento sustancial de la fluorescencia cuando está muy cerca del primero o fluoróforo donante, tal como ocurre cuando los dos oligonucleótidos se aproximan mucho cuando se hibridan con sitios adyacentes en el ácido nucleico que se está Amplificado, por ejemplo en la fase de hibridación de la PCR, formando un complejo fluorogénico. A medida que se acumula más ácido nucleico que se está amplificando, se puede formar más complejo fluorogénico y hay un aumento en la fluorescencia de la sonda aceptora, y esta se puede medir. Por lo tanto, el procedimiento permite la detección de la cantidad de producto a medida que se está formando.

Los expertos en la técnica apreciarán también que la detección de la amplificación puede llevarse a cabo usando numerosos medios en la materia, por ejemplo utilizando sondas de hibridación TaqMan™ en la reacción de PCR y medición de la fluorescencia específica para los ácidos nucleicos diana una vez que se ha producido suficiente amplificación. La PCR en tiempo real TaqMan mide la acumulación de un producto a través del fluoróforo durante las etapas exponenciales de la PCR, en lugar de en el punto final como en la PCR convencional. El aumento exponencial del producto se utiliza para determinar el ciclo de umbral, CT, es decir, el número de ciclos de PCR en los que se detecta un incremento exponencial significativo de la fluorescencia, y que está directamente correlacionado con el número de copias del molde de ADN presentes en la reacción.

Aunque los expertos en la técnica conocerán que existen otros sistemas similares de amplificación/detección de ácidos nucleicos cuantitativos "en tiempo real" y homogéneos, tales como los basados en el enfoque TaqMan (véanse las patentes de Estados Unidos n.º 5.538.848 y 5.691.146), ensayos de polarización de fluorescencia (por ejemplo, Gibson et al., 1997, Clin Chem., 43: 1336–1341), y en ensayo Invader (por ejemplo, Agarwal et al., Diagn Mol Pathol 2000 September; 9(3): 158–164; Ryan D et al, Mol Diagn 1999 June; 4(2): 135–144). Tales sistemas también serían adaptables para su uso en la aplicación descrita, permitiendo la monitorización en tiempo real de la amplificación de ácido nucleico.

En otra realización de la solicitud, las matrices o microchips se fabrican de modo que contengan una matriz de loci que contienen cada uno un oligonucleótido de secuencia conocida. En esta divulgación, cada locus contiene un exceso molar de oligómeros sintéticos inmovilizados seleccionados sintetizados para que contengan secuencias complementarias para porciones deseadas de un gen de susceptibilidad a la diabetes. Las transcripciones de los genes de susceptibilidad a la diabetes presentes en PBMC se amplifican mediante RT-PCR y se marcan, como se describe en el presente documento. A continuación, los oligómeros en los microchips se hibridan con los ácidos nucleicos amplificados del gen de susceptibilidad a la diabetes amplificados por RT-PCR. La hibridación se produce en condiciones rigurosas para asegurar que, durante la hibridación, solo se producirán coincidencias perfectas o casi perfectas entre la secuencia incluida en el microchip y la secuencia diana. La fluorescencia resultante en cada locus es proporcional al nivel de expresión de uno o más genes de susceptibilidad a la diabetes en las PBMC.

En otras realizaciones de la solicitud, se determinan perfiles de expresión de firma génica de genes codificantes de proteínas usando técnicas bien conocidas en la materia de inmunoquímica incluyendo, por ejemplo, ensayos de unión basados en anticuerpos, tales como ELISA o radioinmunoensayos o conjuntos de proteínas que contienen anticuerpos dirigidos a los productos proteicos de genes dentro de una firma predeterminada como se define en el presente documento.

Los perfiles de expresión de los genes TCF7L2 y CLC se analizaron en células mononucleares de sangre periférica de pacientes normales tolerantes a la glucosa y diabéticos de tipo 2.

El gen de la proteína cristalina de Charcot-Leyden humana se expresa principalmente en eosinófilos. CLC está regulado por disminución secuencialmente en PBMC de NGT a IGT a T2D. Las intensidades de señal medias de su expresión en microarrays de la cohorte de 107 pacientes se enumeran en la TABLA 2 siguiente. El análisis de la característica operativa del receptor (ROC) demostró que el nivel de expresión génica de CLC se puede usar para separar los NGT de los IGT/T2D.

TABLA 2: Expresión del gen CLC en PBMC aislados de NGT, IGT y T2D

NGT	IGT	T2D
1504	900	410

El rendimiento del gen CLC en la predicción del estado clínico se examinó más a fondo utilizando un análisis de la característica operativa del receptor (ROC). Una curva ROC muestra la relación entre sensibilidad y especificidad. Es decir, un aumento en la sensibilidad se acompañará de una disminución de la especificidad. Cuanto más cerca sigue la curva el eje izquierdo y luego el borde superior del espacio ROC, más precisa será la prueba. A la inversa, cuanto más cercana esté la curva de la diagonal de 45 grados del gráfico ROC, menor será la precisión de la prueba. El área bajo la ROC es una medida de la precisión de la prueba. La precisión de la prueba depende de lo bien que la prueba separe el grupo que se está probando en aquellos con y sin la enfermedad en cuestión. Un área bajo la curva (denominada "AUC") de 1 representa una prueba perfecta, mientras que un área de 0,5 representa una prueba menos útil. Por lo tanto, los genes y los procedimientos de diagnóstico preferidos de la presente solicitud tienen una AUC superior a 0,50, los ensayos más preferidos tienen una AUC superior a 0,60, las pruebas más preferidas tienen una AUC superior a 0,70.

El área bajo curva (AUC) se calculó como una medida del rendimiento del gen CLC en la predicción del estado del paciente. El análisis de la característica operativa del receptor (ROC) de los datos del gen CLC demostró que el nivel de expresión génica de CLC se puede usar para separar los NGT de los IGT/T2D (véase la figura 1).

Las variantes genéticas en el gen que codifica el factor de transcripción 7 de tipo 2 (TCF7L2) se han asociado fuertemente con un riesgo de desarrollar diabetes tipo 2 y función alterada de la insulina de las células (véase el documento US 2006/0286588). Los estudios de asociación de todo el genoma implican que los SNP dentro de TCF7L2 dan la puntuación más alta del riesgo durante la para la predicción de la progresión de la diabetes de tipo 2 en comparación con los SNP en otros genes marcadores, incluyendo CDKAL1, CDKN2A/2B, FTO, IGF2BP2 y SLC30A8 (intervalo de puntuaciones de riesgo 1,12-1,20). El TCF7L2 se expresa ampliamente y se sabe que este factor de transcripción responde a señales de desarrollo de miembros de la familia de proteínas Wnt. Los estudios funcionales y genéticos apuntan a un papel crucial para el TCF7L2 en el desarrollo de la expresión génica intestinal y proglucagón en las células enteroendocrinas.

Para determinar si el TCF7L2 y el gen CLC son marcadores diagnósticos de la diabetes, individualmente o en combinación, se reclutaron 180 sujetos de la población alemana en asociación con el Institute for Clinical Research & Development (IKFE), en Mainz, Alemania. Se obtuvieron las aprobaciones apropiadas de la IRB antes de la recolección de la muestra del paciente. Los criterios de inclusión consistieron en pacientes de entre 18-75 años de edad y un índice de masa corporal (IMC) ≥ 30 que no tenían diagnóstico previo de diabetes y la capacidad legal y capacidad para comprender la naturaleza y la extensión del estudio clínico y los procedimientos requeridos. Los criterios de exclusión consistieron en donaciones de sangre en los últimos 30 días, mujeres con diabetes mellitus insulino dependiente, lactantes o embarazadas, o mujeres que desean quedar embarazadas durante el estudio, mujeres sexualmente activas que no utilizan anticonceptivos, antecedentes de alergias graves/múltiples, abuso de drogas o alcohol y falta de cumplimiento de los requisitos del estudio. Todas las mediciones clínicas, incluyendo los resultados de la prueba oral de tolerancia a la glucosa de 75 g (OGTT) se obtuvieron usando procedimientos estándar.

Las muestras de sangre se extrajeron mediante punción venosa en tubos de CPT (BD Biosciences). Las PBMC se aislaron de acuerdo con el protocolo del fabricante y el sedimento celular final se resuspendió en 1 ml de Trizol (Invitrogen) y se almacenó a -80°C . Posteriormente, el ARN total se purificó usando el protocolo del fabricante y se resuspendió en ddH₂O tratada con DEP. La cuantificación y la calidad del ARN se realizó utilizando el espectrofotómetro ND-1000 (NanoDrop) y se reconfirmó mediante cuantificación espectrofotométrica con el kit RiboGreen (Molecular Probes). La calidad de los moldes de ARN se midió utilizando el Bioanalyzer 2100 (Agilent Technologies).

La síntesis de ADNc de primera cadena se realizó usando 200 ng de ARN total de cada muestra de PBMC de pacientes usando el kit de transcripción inversa de ADNc de alta capacidad (Applied Biosystems). Después, la mezcla de reacción se diluyó 10 veces con ddH₂O y se usaron 4 μl como molde en una reacción de PCR Taqman de 10 μl en el sistema de detección de secuencias ABI Prism 7900HT. Los componentes de reacción consistieron en 2X mezcla maestra para PCR Taqman (Applied Biosystems), 0,9 μM de cada cebador y 0,25 μM de sonda fluorescente marcada (Biosearch Technologies). Las secuencias para los conjuntos de cebador/sonda utilizados en el ensayo de RT-PCR Taqman se presentan en la **Tabla 3**.

Las condiciones de ciclación para la etapa de transcripción inversa fueron las siguientes:

	Etapa 1	Etapa 2	Etapa 3	Etapa 4
Temp. ($^{\circ}\text{C}$)	25	37	85	4
Tiempo	10 min.	120 min.	5 s	

Tabla 3. Secuencias de la sonda y el cebador usadas en los ensayos Taqman para TCF7L2, CLC y ACTINA

Conjunto de marcadores #	Secuencia 5'>3'	Nombre de la secuencia	Comentarios
Conjunto 1 de TCF7L2	ACCTGAGCGCTCCTAAGAAATG	TCF7L2 DL F1	NO sobre la unión
	AGGGCCGCACCAGTTATTC	TCF7L2 DL R1	NO sobre la unión
	FAM-AGCGCGCTTTGGCCTTGATCAAC-BHQ1	TCF7L2 DL Pro1	NO sobre la unión
Conjunto 2 de TCF7L2	CGTCGACTTCTTGTTACATTCC	TCF7L2 DL F2	NO sobre la unión
	CACGACGCTAAAGCTATTCTAAAGAC	TCF7L2 DL R2	NO sobre la unión
	FAM- CAGCCGCTGTCGCTCGTCACC-BHQ1	TCF7L2 DL Pro2	NO sobre la unión
Conjunto 3 de TCF7L2	GAAAGCGCGCCATCAAC	TCF7L2 1564 U18	Sobre la unión
	CAGCTCGTAGTATTTGCTTGCT	TCF7L2 1644 L23	Sobre la unión
	FAM-TCCTTGGGCGGAGGTGGCATG-BHQ1	TCF7L2 1586 P21	Sobre la unión
Conjunto 1 de CLC	GCTACCCGTGCCATACACAGA	CLC 85 U21	Sobre la unión
	GCAGATATGGTTCATTCAAGAAACA	CLC 185 L25	Sobre la unión
	FAM-TTCTACTGTGACAATCAAAGGGCGACCA-BHQ1	CLC 127 P28	Sobre la unión
ACTINA	CCTGGCACCCAGCACAAT	B-actina-1F	
	GCCGATCCACACGGAGTACTT	B-actina-1R	
	FAM-ATCAAGATCATTGCTCCTCCTGAGCGC-BHQ1	B-actina P	Control interno

Se realizó RT-PCR cuantitativa en tiempo real mediante el ensayo de Taqman usando dos combinaciones diferentes de cebador/sonda específicas para TCF7L2 y un conjunto cebador/sonda para CLC, sobre ARN aislado de PBMC de pacientes individuales. El perfil de termociclado para la etapa de PCR fue el siguiente, 95 °C durante 10 minutos, seguido de 40 ciclos de 95 °C (15 s) y 60 °C (1 min). Los valores de Ct se calcularon a partir de los datos brutos utilizando el software SDS versión 2.1 (Applied Biosystems), con un umbral establecido de 0,2 a 0,3. La reproducibilidad de ejecución-ejecución mediante la correlación de Pearson fue $R^2 = 0,96-0,98$ para los marcadores mencionados anteriormente. El valor de Ct delta (umbral de ciclo) se calculó restando el Ct del gen doméstico de -actina del Ct del marcador de interés, por ejemplo, TCF7L2 (Ct TCF7L2-Ct actina). El valor de $[2^{-(\text{delta Ct})} \times 1000]$ se usó para representar la expresión de TCF7L2 con respecto a la -actina. El resultado de la OGTT se utilizó como el verdadero estado clínico. Se utilizó la prueba T de Student para determinar la significación estadística entre los niveles de expresión de marcadores génicos usando valores normalizados de Ct.

Según la medición de la OGTT a 2 horas y las guías actuales de la ADA, de los 180 pacientes incluidos en el estudio, 104 pacientes fueron clasificados como NGT (tolerancia normal a la glucosa), 49 pacientes como IGT (tolerancia alterada a la glucosa) y 27 pacientes se consideraron T2D (diabetes de tipo 2). Debido a que se diagnosticó a los sujetos con T2S con diabetes por primera vez en este estudio, se desconocía la duración de la enfermedad para cada paciente y la duración del mantenimiento de las PBMC en un microambiente hiperglucémico.

El análisis de la prueba T para los niveles de expresión de TCF7L2 y CLC normalizados a -actina para cada paciente, basado en los valores de los conjuntos cebador/sonda y separados por la tolerancia a la glucosa, se muestra en la Tabla 4. Los grupos de pacientes NGT e IGT más T2D tuvieron una diferencia estadísticamente significativa entre los niveles de expresión mediante la prueba T de Student, con valores de $p = 0,004$; $0,021$ y $0,022$ para TCF7L2 conjunto 1, conjunto 3 y CLC, respectivamente. Estos resultados indican expresión diferencial de los genes TCF7L2 y CLC en PBMC de pacientes pre-diabéticos (IGT) o pacientes pre-diabéticos y T2D combinados en comparación con NGT.

TABLA 4: Diferencia estadística en los niveles de expresión de TCF7L2 y CLC mediante T de Student

Muestra n	Prueba t	Ct de B-actina	TCF7L2_1 normalizado	TCF7L2_2 normalizado	TCF7L2_3 normalizado	CLC normalizado
104/49	NGT/IGT	0,451	0,005	0,075	0,008	0,013
104/76	NGT/(IGT+T2D)	0,277	0,004	0,093	0,021	0,022
Diferenciación estadísticamente significativa						

10 A continuación, se evaluó el rendimiento de los ensayos TCF7L2 y CLC Taqman como herramienta de diagnóstico para la clasificación de pacientes como normales o pre-diabéticos/diabéticos, en comparación con la OGTT de 2 horas. Se generaron curvas de las características operativas del receptor (ROC) para cada valor de Ct delta normalizado del conjunto de cebador/sonda TCF7L2 y CLC (Tabla 5, Figuras 2A y 2B). Los valores de la AUC para los ensayos de PCR de TCF7L2 conjunto 1 y CLC fueron 0,63 y 0,61, respectivamente. En comparación con la clasificación de OGTT a 2 horas, la expresión de TCF7L2 conjunto 1 por los PBMC puede clasificar correctamente a un paciente como normal o pre-diabético/diabético con una AUC de 0,73 cuando se utiliza junto con la prueba de FPG (figuras 2A y 2B). CLC no tenía un valor aditivo para TCF7L2 conjunto 1 y no se consideró para el algoritmo de diagnóstico. Además, la exclusión de 14 pacientes que tenían FPG 126mg/dL (también considerados diabéticos) no cambió el rendimiento del ensayo.

20 Tabla 5: Valores de AUC de la curva ROC para cada conjunto de sonda-cebador marcador y en combinación con los valores de FPG

N.º de conjunto de marcadores	Valor de AUC en ROC	
	Marcador vs. OGTT	Marcador + FPG vs. OGTT
Conjunto 1 de TCF7L2	0,63	0,73
Conjunto 2 de TCF7L2	0,59	ND
Conjunto 3 de TCF7L2	0,61	0,72
Conjunto 1 de CLC	0,60	0,69

En una realización, cada uno de los genes seleccionados en el análisis de microarrays (véase la TABLA 1) puede combinarse con el rendimiento de TCF7L2 conjunto 1 para coincidir más estrechamente con el resultado de OGTT de 2 horas.

En otra realización, los genes que están fuertemente asociados con un riesgo de diabetes tipo 2 (véase la TABLA 6) también pueden combinarse con el rendimiento de TCF7L2 conjunto 1 para coincidir más estrechamente con el resultado de la OGTT a 2 horas. En otra realización, los genes de la Tabla 6 en combinación con uno o más genes de la TABLA 1 pueden analizarse como se describe en el presente documento para firmas génicas que son diagnósticas de un estado de enfermedad diabética.

TABLA 6:

Símbolo del gen			
NOTCH2	IGF2BP2	LGR5	CDKN2A-2B
THADA	WFS1	FTO	HHEX-IDE
PPARG	KCNJ11	JAZF11	CDC123
ADAMTS9	TSPAN8	SLC30A8	CAMK1D

En otra realización, la expresión de CDKN1C, un miembro de la familia CIP/KIP también se expresó diferencialmente en PBMC de NGT y T2D.

La familia CIP/KIP consiste en tres miembros, CDKN2A, CDKN2B y CDKN1C. Los tres miembros pueden inhibir la actividad de CDK4, que desempeña un papel central en la regulación del ciclo celular de los mamíferos. La

replicación de las células de los islotes desempeña un papel esencial en el mantenimiento de la homeostasis en masa de las células. Se sabe que CDK4 tiene un papel importante en la regulación del peso corporal y la proliferación de las células pancreáticas. En los ratones, la pérdida del gen CDK4 dio lugar a diabetes deficiente en insulina debido a la reducción de la masa de las células, mientras que la activación de CDK4 causó hiperplasia de las células de los islotes. Recientemente, los estudios de asociación genómica de la diabetes tipo 2 han revelado que la variación de nucleótidos cerca de los genes CDKN2A y CDKN2B se asocia con el riesgo de diabetes tipo 2. Además, la sobreexpresión de CDKN2A conduce a una disminución de la proliferación de los islotes en ratones envejecidos y la sobreexpresión de CDKN2B está relacionada con la hipoplasia de los islotes y diabetes en modelos murinos. CDKN1C es un gen de expresión materna localizado en el cromosoma 11p15.5 y está implicado en la patogenia del síndrome de Beckwith-Wiedemann (BWS), un trastorno caracterizado por hipoglucemia hiperinsulinémica neonatal, así como por sobrecrecimiento pre y posnatal. Estudios recientes también demostraron que el CDKN1C está regulado por disminución por la insulina y que las variantes de CDKN1C pueden estar asociadas con el aumento del peso al nacer en pacientes con diabetes tipo 2. Además de regular el ciclo celular, la familia CIP/KIP desempeña un papel importante en otros procesos biológicos, tales como la apoptosis, la regulación de la transcripción, la diferenciación y la migración celular. Se analizó la expresión de los tres genes en la cohorte de 107 pacientes. Sólo CDKN1C mostró expresión diferencial entre NGT, IGT y T2D (véase la TABLA 7). Hay 5 sondas que se expresan en PBMC para CDKN1C en el HG-U133Plus2 GeneChip. Cada uno de ellos mostró expresión diferencial entre NGT e IGT/T2D (TABLA 7). El análisis de ROC mostró que los niveles de expresión de las 5 sondas se pueden usar para separar los NGT de los T2D (Figura 3).

TABLA 7: Expresión del gen CDKN1C en PBMC aislados de NGT, IGT y T2D

Gen	Sonda	Expresión en PBMC	Media_NGT	Media_IGT	Media_T2D
CDKN1C	213182_x_at	Sí	935	1178	1784
	213183_s_at	Sí	531	712	624
	213348_at	Sí	2648	3246	3957
	216894_x_at	Sí	797	1030	1439
	219534_x_at	Sí	1092	1356	1973

Un experto en la materia reconocerá que las realizaciones descritas proporcionan una premisa para investigar las firmas génicas como herramienta de diagnóstico de la diabetes. Para investigar los procesos biológicos subyacentes entre sujetos normales y pacientes de pre-diabetes y diabetes, se realizó el análisis de la vía. A saber, las sondas en el chip HG-U133Plus2 se mapearon con Gene Ontology Biological Process (GOBP) como se describe en Yu et al. BMC Cancer 7:182 (2007). Dado que los genes con expresión muy baja tienden a tener mayores variaciones, los genes cuya intensidad media es inferior a 200 en el conjunto de datos se eliminaron del análisis de la vía. Como resultado, se retuvieron 21247 sondas. Para identificar las vías que tienen una asociación significativa con el desarrollo de pre-diabetes o diabetes, el programa de prueba global se realizó mediante la comparación de NGT frente a IGT, NGT frente a T2D o NGT frente a IGT + T2D. Se identificaron las vías que tienen al menos 10 sondas y un valor p significativo ($p < 0,05$) para cada comparación. Hubo 3 vías que tuvieron una asociación consistente con los resultados del paciente a través de las tres comparaciones. Son activación de células B (GO0042113), respuesta inmunitaria humoral (GO0006959) y desenrollamiento del ADN durante la replicación (GO0006268). Entre las 3 vías, la activación de células B y la respuesta inmunitaria humoral tienen una asociación dominante negativa con la diabetes (menor expresión en IGT/T2D), mientras que el desenrollamiento del ADN durante la replicación tiene una asociación positiva con la diabetes (mayor expresión en IGT/T2D).

Para construir una firma génica basada en la vía de las 3 vías clave, los genes con una $p < 0,05$ se agruparon y ordenaron en función de su significación estadística (puntuación z de la prueba global). Si un gen tiene más de una sonda en la lista y sus comportamientos eran consistentes, se mantuvo el que tenía la mayor significación, Si un gen tiene más de una sonda en la lista y sus comportamientos eran opuestos, se eliminaron todas las sondas para este gen. Como resultado, se obtuvieron 14 genes únicos (VÉASE LA TABLA 8 más adelante).

TABLA 8: Genes significativos de la vía

	PSID	Símbolo del gen	Título del gen
5	208900_s_at	TOP1	topoisomerasa (ADN) I
	216379_x_at	CD24	Antígeno CD24 (antígeno 4 del clúster del carcinoma de pulmón microcítico)
	222430_s_at	YTHDF2	Miembro del da la familia de dominio YTH
10	1554343_a_at	BRDG1	Señalización 1 aguas debajo de BCR
	228592_at	MS4A1	Miembro 1, subfamilia A, dominios 4 transmembranales
	216894_x_at	CDKN1C	Inhibidor 1C de la cinasa dependiente de ciclina (p57, Kip2)
	1558662_s_at	BANK1	Proteína armazón de células B con repeticiones 1 de anquirina
	205267_at	POU2AF1	Dominio POU, clase 2, factor de asociación 1
15	205859_at	LY86	Anfígeno 86 de linfocitos
	221969_at	PAX5	Gen 5 caja pareada (activador específico del linaje de las células B)
	207655_s_at	BLNK	Enlazador de células B
20	206126_at	BLR1	Receptor 1 del linfoma de Burkitt, proteína de unión a GTP (quimioquina (motivo C-X-C) receptor 5)
	206983_at	CCR6	Receptor 6 de quimiocina (motivo C-C)
	204946_s_at	TOP3A	topoisomerasa (DNA) III alfa
	214252_s_at	CLN5	Lipofuscinosis ceroide, neuronal 5

25 Para construir una firma utilizando genes con una variación relativamente alta, se retuvieron 10 genes con una CV > 0,25. Para determinar el número óptimo de genes para una firma, la combinación de 2-10 genes principales se examinó en el conjunto de datos. El resultado indicó que los 3 genes principales dieron el mejor rendimiento en la predicción de los resultados de los pacientes. Los 3 genes, TOP1, CD24 y STAP1 se muestran a continuación en la TABLA 9.

30

TABLA 9: Expresión de 3 genes en PBMC aislados de NGT, IGT y T2D

3 genes principales del análisis de la vía			
Sonda	Símbolo del gen	Título del gen	
208900_s_at	TOP1	topoisomerasa (ADN) I	
216379_x_at	CD24	Antígeno CD24 (antígeno 4 de clúster de carcinoma de pulmón microcítico)	
1554343_a_at	STAP1	Miembro1 de la familia de adaptador de la transducción de la señal	

40

La expresión media de los 3 genes principales en subgrupos

Sonda	Gen	Media_NGT	Media_IGT	Media_T2D
208900_s_at	TOP1	868	1145	1418
216379_x_at	CD24	1767	1274	1194
1554343_a_at	STAP1	373	283	265

50

El análisis ROC de la firma de 3 genes en la cohorte de 107 pacientes (figuras 4A y 4B) demuestra que esta firma puede separar los NGT de los IGT/T2D. En la figura 4C se muestra un histograma que representa la expresión media de los genes.

55

Para eliminar los genes no informativos, solo se retuvieron los genes que tenían 10 o más llamadas de presencia en la cohorte. A continuación, la cohorte de 107 pacientes se dividió en un grupo de formación de 54 pacientes y un grupo de ensayo de 53 pacientes. Basado en la clasificación de la OGTT, hay 28 NGT, 17 IGT y 9 DT2 en el grupo de formación, mientras que hay 29 NGT, 16 IGT y 8 DT2 en el grupo de ensayo. Para identificar los genes que tienen expresión diferencial entre pacientes NGT y IGT + T2D, se realizó el programa de análisis significativo de Microarray (SAM). Los genes se seleccionaron si la tasa de falsos descubrimientos (FDR) era inferior al 20 %. Como resultado, se seleccionaron 235 genes. Para restringir aún más la lista de genes, se mantuvieron los genes con cambios mayores de 1,5 entre los dos grupos y la intensidad promedio del gen en el conjunto de datos es mayor de 200. Como resultado, se obtuvieron 17 sondas. Entre ellos, 4 eran sondas que representan el gen de la hemoglobina. Teniendo en cuenta que la hemoglobina tiene una expresión extremadamente alta en los glóbulos rojos, se retiraron las 4 sondas para eliminar la posible contaminación. Para determinar el número óptimo de genes

65

como firma, el rendimiento de la combinación de los principales genes se examinó de 2 a 13 en el conjunto de formación. El resultado indicó que los 10 genes principales daban el mejor rendimiento según el área bajo curva (AUC) (véase la Tabla 10).

5 **TABLA 10A: Expresión de 10 genes en PBMC aislados de NGT, IGT y T2D**

Sonda	Símbolo	Título
239742_at	TULP4	Proteína 4 de tipo Tubby
244450_at	AA741300	Débilmente similar a la SECUENCIA SX DE LA SUBFAMILIA TO ALU8_HUMAN ALU
235216_at	ESCO1	Homólogo 1 del establecimiento de cohesión
201026_at	EIF5B	Factor 5B de iniciación de la traducción en eucariotas
200727_s_at	ACTR2	Homólogo 2 de la proteína relacionada con la actina ARP2
211993_at	WNK1	Proteína cinasa 1 deficiente en WNK lisina
205229_s_at	COCH	Homólogo C del factor de coagulación, cochlina
201085_s_at	SON	Proteína de unión a ADN SON
1557227_s_at	TPR	Región promotora translocada (al oncogén MET activado)
231798_at	NOG	Noggin

20 **TABLA 10B: La expresión media de los 10 genes principales en subgrupos**

Sonda	Gen	Media_NGT	Media_IGT	Media_T2D
239742_at	TULP4	514	659	702
244450_at	AA741300	674	461	482
235216_at	ESCO1	199	262	351
201026_at	EIF5B	330	440	500
200727_s_at	ACTR2	2153	2751	3590
211993_at	WNK1	397	505	625
205229_s_at	COCH	330	231	250
201085_s_at	SON	3300	4103	4900
1557227_s_at	TPR	378	445	616
231798_at	NOG	515	430	302

35 Para evaluar aún más la firma génica, se determinaron los resultados del paciente en el conjunto de prueba. También se examinó la predicción de la pre-diabetes y la diabetes utilizando niveles plasmáticos de glucosa en ayunas (FPG). Para investigar el efecto complementario entre la firma génica y los niveles de FPG, la combinación de estos dos predictores se utilizó para predecir los resultados del paciente. En la figura 5 se representa una comparación de los análisis de ROC entre el uso de FPG, o firma de 10 genes, o la combinación de FPG y la firma de 10 genes en el conjunto de prueba. Demuestra que la firma de 10 genes puede separar de forma independiente los NTG de los IGT/T2D y la FPG y la forma de 10 genes son complementarios para una mejor predicción (véanse las figuras 5A y 5B). Las señales de expresión media de los 10 genes en la cohorte de 107 pacientes se muestran en la tabla y el gráfico de barras de la figura 5C.

45 El análisis estadístico de los datos clínicos identificó una firma de 3 genes y 10 genes que se expresan diferencialmente en NGT y T2D.

50 En otra realización, se describe un ensayo diagnóstico para la clasificación del punto de atención de pacientes normales frente a pre-diabéticos/diabéticos o para la predicción de la progresión a prediabetes/diabetes durante un periodo de tiempo definido, por ejemplo de ½ a 2 años o de 2 a 5 años, o de 5 a 10 años o más.

55 Como alternativa, los perfiles de expresión génica se determinan mediante la detección de la proteína codificada por el ARNm, por ejemplo utilizando ELISA o matriz proteómica. En la técnica se conocen bien todos estos procedimientos.

60 Divulgación en el presente documento también proporciona un formato de kit que comprende una unidad de paquete que tiene uno o más reactivos para el diagnóstico de un estado de enfermedad diabética en un paciente. El kit también puede contener uno o más de los siguientes elementos: tampones, instrucciones y controles positivos o negativos. Los kits pueden incluir recipientes de reactivos mezclados juntos en proporciones adecuadas para llevar a cabo los procedimientos descritos en el presente documento. Los recipientes de reactivo contienen, preferentemente, reactivos en cantidades unitarias que obvian las etapas de medición cuando se realizan los procedimientos objeto.

65 El kit puede incluir agujas y tubos/recipientes estériles para la recogida de sangre de un paciente. Los tubos de recogida contendrán típicamente ciertos aditivos, por ejemplo heparina, para inhibir la coagulación de la sangre.

Los kits también pueden contener reactivos para la medición de un perfil de expresión de la firma génica en la muestra de un paciente. Como se divulga en el presente documento, los perfiles de expresión de las firmas génicas pueden medirse mediante una diversidad de medios conocidos en la materia, incluyendo ensayos de RT-PCR, ensayos basados en oligonucleótidos usando microchips o ensayos basados en proteínas, tales como ensayos ELISA.

En una realización preferida, los perfiles de expresión de la firma génica se miden mediante RT-PCR en tiempo real.

En una realización de la solicitud, el kit comprende cebadores de la amplificación y detección de perfiles de expresión de la firma génica en la muestra de sangre de un paciente. Los cebadores pueden tener una secuencia que es complementaria a una cualquiera de los genes de susceptibilidad a la diabetes, como se define en el presente documento, incluyendo los TOP1, CD24, STAP1, TULP4, AA741300, ESCO1, EIF5B, ACTR2, WNK1, COCH, SON, TPR, NOG o uno cualquiera de los genes enumerados en las Tablas 1 o 6.

Ejemplos de secuencias de cebadores utilizados para la RT-PCR en tiempo real de los genes de susceptibilidad a la diabetes se divulgan en las Tablas 12B y 12C.

En una realización preferida, los reactivos del kit están diseñados para funcionar con el instrumento 7500 Fast Dx Real-Time PCR de Applied Biosystems, que es una tecnología basada en PCR que fue aprobada por la Oficina de Diagnóstico In Vitro de la FDA (FDA-OIVD).

En otra realización más, el kit incluye un microchip que comprende un conjunto de sondas de hibridación para las firmas de 3 genes (TOP1, CD24 y STAP1) o de 10 genes (TULP4, AA741300, ESCO1, EIF5B, ACTR2, WNK1, COCH, SON, TPR y NOG). En otro aspecto, los microchips pueden comprender además una matriz de una o más sondas de hibridación para uno o más de los genes enumerados en las Tablas 1 o 6.

En una realización preferida, los microchips están diseñados para funcionar con la tecnología Affymetrix GeneChipDx que puede medir, en paralelo, la expresión génica de 1 a más de 55.000 ARNm. FDA-OIVD esta plataforma para su uso con el producto AmpliChip P450 de Roche Molecular Diagnostics y el Pathwork Diagnostics Tissue of Origin test.

TABLA 1

Sonda	Símbolo del gen	Título del gen
218659_at	ASXL2	Peinetas sexuales adicionales tipo 2 (Drosophila)
230528_s_at	MGC2752	Proteína hipotética MGC2752
211921_x_at	PTMA	Protimosina, alfa (secuencia génica 28) /// protimosina, alfa (secuencia génica 28)
209102_s_at	HBP1	factor de transcripción 1 HMG-box
239946_at	KIAA0922	Proteína KIAA0922
226741_at	TMEM85	Proteína transmembrana 85
239742_at	TULP4	Proteína 4 de tipo Tubby
202844_s_at	RALBP1	Proteína 1 de unión a rala
237768_x_at	TAF15	ARN polimerasa II TAF15, factor asociado a la proteína de unión a la caja TATA (TBP), 68kDa
202373_s_at	RAB3GAP2	Subunidad 2 de la proteína activadora de la GTPasa RAB3 (no catalítica)
223413_s_at	LYAR	Proteína hipotética FLJ20425
222371_at	PIAS1	Inhibidor proteico de de STAT activado, 1
244450_at	MAK	cinasa asociada a células germinales masculinas
201024_x_at	EIF5B	Factor 5B de iniciación de la traducción en eucariotas
202615_at	GNAQ	Proteína de unión a nucleótidos de guanina (proteína G), polipéptido q
222621_at	DNAJC1	Homólogo DnaJ (Hsp40), subfamilia C, miembro 1
212774_at	ZNF238	Proteína 238 del dedo de cinc
238883_at	THRAP2	Proteína asociada al receptor de la hormona tiroidea 2
223130_s_at	MYLIP	Proteína de interacción reguladora de la cadena ligera de miosina
225445_at	—	Locus transcrito
235601_at	MAP2K5	Proteína cinasa cinasa 5 activada por mitógenos
209258_s_at	CSPG6	Proteoglicano condroitina sulfato 6 (bamacan)
1557238_s_at	SETD5	Dominio SET que contiene 5
202927_at	PIN1	Proteína 1 de interacción con NIMA (Peptidil-prolilo cis/trans isomerasa)
1568618_a_at	GALNT1	UDP-N-acetil-alfa-D-galactosamina: polipéptido Nacetilgalactosaminiltransferasa 1 (GalNAc-T1)

ES 2 635 429 T3

	222417_s_at	SNX5	Nexina 5 de clasificación
	208836_at	ATP1B3	ATPasa, transportador de Na +/K +, polipéptido beta 3
	202738_s_at	PHKB	Fosforilasa cinasa, beta
	224872_at	KIAA1463	Proteína KIAA1463
	235200_at	ZNF561	Proteína 561 del dedo de cinc
	235216_at	ESCO1	Homólogo 1 del establecimiento de cohesión (<i>S. cerevisiae</i>)
5	201026_at	EIF5B	Factor 5B de iniciación de la traducción en eucariotas
	208095_s_at	SRP72	Partícula de reconocimiento de señal 72 kDa
	244457_at	ITPR2	Familia con similitud de secuencia 20, miembro C
	216563_at	ANKRD12	Dominio 12 de repetición de anquirina
	211983_x_at	ACTG1	Actina, gamma 1
10	227854_at	FANCL	Anemia de Fanconi, grupo de complementación L
	1552343_s_at	PDE7A	Fosfodiesterasa 7A
	221548_s_at	ILKAP	Serina/treonina fosfatasa 2C asociada a la cinasa unida a integrina
	215772_x_at	SUCLG2	succinato-CoA ligasa, formador de ADP, subunidad beta
	229010_at	CBL	Secuencia de transformación retroviral ecotrópica Cas-Br-M (murina)
15	226879_at	MGC15619	Proteína hipotética MGC15619
	1556451_at	BACH2	Homología 1 BTB y CNC, factor básico de transcripción de la cremallera de leucina 2
	225490_at	ARID2	Dominio 2 interactivo rico en AT (ARID, de tipo RFX)
	214055_x_at	BAT2D1	dominio BAT2 que contiene 1
20	32069_at	N4BP1	Proteína 1 de unión a Nedd4
	235457_at	MAML2	De tipo mastermind 2 (<i>Drosophila</i>)
	217985_s_at	BAZ1A	Dominio de bromo adyacente al dominio del dedo de cinc, 1A
	229399_at	C10orf118	Marco de lectura abierto 118 del cromosoma 10
	208994_s_at	PPIG	Peptidil-prolil isomerasa G (ciclofilina G)
25	202656_s_at	SERTAD2	Dominio SERTA que contiene 2
	241917_at	FCHSD2	FCH y dominios SH3 dobles 2
	238807_at	ANKRD46	Dominio 46 de repetición de anquirina
	204415_at	G1P3	Interferón, proteína alfa-inducible (clon IFI-6-16)
	240176_at	LOC391426	Similar a ENSANGP00000004103
30	233284_at	—	—
	232583_at	—	—
	200772_x_at	PTMA	protimosina, alfa (secuencia génica 28)
	239721_at	UBE2H	Enzima de conjugación de ubiquitina E2H (homólogo de UBC8, levadura)
	218607_s_at	SDAD1	Dominio SDA1 que contiene 1
35	204160_s_at	ENPP4	Ectonucleótido pirofosfatasa/fosfodiesterasa 4 (función putativa)
	212991_at	FBXO9	Proteína 9 de caja F
	227770_at	VPS4A	Clasificación de proteínas vacuolares 4A (levadura)
	222111_at	FAM63B	Familia con similitud de secuencia 63, miembro B
	1558604_a_at	—	MRNA; clon CD 43T7
40	205229_s_at	COCH	Homólogo C del factor de coagulación, cochlina (<i>Limulus polyphemus</i>)
	219130_at	FLJ10287	Proteína hipotética FLJ10287
	241262_at	—	—
	202412_s_at	USP1	Peptidasa 1 específica de ubiquitina
	225092_at	RABEP1	Rabaptina, proteína efectora 1 de unión a RAB GTPasa
45	200905_x_at	HLA-E	Complejo mayor de histocompatibilidad, clase I, E
	201010_s_at	TXNIP	Proteína de interacción con tioredoxina
	221607_x_at	ACTG1	Actina, gamma 1
	201085_s_at	SON	Proteína de unión a ADN SON
	214723_x_at	KIAA1641	KIAA1641
50	201565_s_at	ID2	Inhibidor 2 de la unión al ADN, proteína hélice-bucle-hélice negativa dominante
	201861_s_at	LRRFIP1	Proteína 1 de interacción rica en leucina (en FLII)
	207785_s_at	RBPSUH	Recombinación de proteína de unión recombinante de sin pelo (<i>Drosophila</i>)
	230415_at	—	—
55	236620_at	RIF1	Homólogo del factor de interacción con RAP1 (levadura)
	206363_at	MAF	Homólogo del oncogén de fibrosarcoma musculoesquelético v-maf (aviar)
	1558748_at	NAPE-PLD	Fosfolipasa D hidrolizante de N-azil-fosfatidiletanolamina
	223101_s_at	ARPC5L	Complejo 2/3 de proteína relacionada con actina 2/3, similar a la subunidad 5
60	236370_at	SMURF1	Ubiquitina E3 proteína ligasa 1 específica de SMAD
	200702_s_at	DDX24	Polipéptido 24 de caja DEAD (Asp-Glu-Ala-Asp)
	1557227_s_at	TPR	Región promotora translocada (al oncogén MET activado)
	220934_s_at	MGC3196	Proteína hipotética MGC3196
65	233333_x_at	AVIL	Advillina
	231798_at	NOG	Noggin
	228986_at	OSBPL8	Proteína de tipo 8 de unión a oxisterol

	243303_at	ECHDC1	Dominio de enoil coenzima A hidratasa que contiene 1
	225266_at	ZNF652	Proteína 652 del dedo de cinc
	220072_at	CSPP1	Proteína 1 asociada al centrosoma y el polo del huso
5	234196_at	TMCC3	Familia de dominios transmembrana y superenrollados 3
	222616_s_at	USP16	Peptidasa 16 específica de ubiquitina
	201274_at	PSMA5	Subunidad de proteasoma (prosoma, macropaina), tipo alfa, 5
	238714_at	RAB12	RAB12, miembro de la familia del oncogén RAS
	204563_at	SELL	Selectina L (molécula de adhesión de linfocitos 1)
10	1557239_at	BBX	Homólogo Bobby sox (Drosophila)
	232510_s_at	DPP3	dipeptidilpeptidasa 3
	235653_s_at	THAP6	Dominio THAP que contiene 6
	200727_s_at	ACTR2	Homólogo 2 de la proteína relacionada con la actina ARP2 (levadura)
	221564_at	HRMT1L1	HMT1 hnRNP metiltransferasa de tipo 1 (S. cerevisiae)
15	211993_at	WNK1	Proteína cinasa 1 deficiente en lisina WNK /// Proteína cinasa 1 deficiente en lisina WNK
	201114_x_at	PSMA7	Subunidad de proteasoma (prosoma, macropaina), tipo alfa, 7
	233089_at	QRSL1	Glutaminil-tRNA sintasa 1 de tipo (hidrolizante de glutamina)

20 **Gen Descripción (funciones, dominios)**

25 **TCL7L2** El producto del gen TCL7L2 es un factor de transcripción que contiene la caja del grupo de alta movilidad implicado en la homeostasis de la glucemia. Las proteínas del grupo de alta movilidad (HMG o HMGB) son de componentes no histonas de peso molecular relativamente bajo de la cromatina. HMG1 (también d HMG-T en el pescado) y HMG2 son dos proteínas altamente relacionadas que se unen, preferentemen monocatenario y desenrollan el ADN bicatenario. Aunque no tienen especificidad de secuencia, tienen una e por el ADN doblado o distorsionado, y doblan el ADN lineal. HMG1 y HMG2 contienen dos dominios de ca unión a ADN (A y B) que muestran diferencias estructurales y funcionales, y tienen un dominio C-terminal rico en residuos de ácido aspártico y glutámico. La cola ácida modula la afinidad de las cajas HMG en HMG1 y 2 para diversas dianas de ADN, HMG1 y 2 parecen desempeñar papeles arquitectónicos import: ensamblaje de complejos de nucleoproteínas en diversos procesos biológicos, por ejemplo la recombinación iniciación de la transcripción y la reparación del ADN.

35 **CLC** La proteína codificada por este gen es una lisofosfolipasa expresada en eosinófilos y basófilos. la isofosfa glicerofosfocolina y un ácido graso libre. Esta proteína puede poseer actividades de unión a carbohidratos estructural y funcionalmente relacionada con la familia de galectinas de proteínas de unión a beta-galactós estar asociada con inflamación y algunas leucemias mieloides. Las galectinas (anteriormente S-lectina exclusivamente a beta-galactósidos como la lactosa. No requieren iones metálicos para la actividad. Las ga encuentran predominantemente, pero no exclusivamente en los mamíferos. Su función no está clara. Están por el desarrollo y pueden estar involucrados en la diferenciación, la regulación celular y la construcción de t

40 **CDKN1C** La proteína codificada por este gen es un inhibidor potente de unión estrecha de varios complejos de ciclin: un regulador negativo de la proliferación celular. Las mutaciones en este gen están implicadas en esporádicos en el síndrome de Beckwith–Wiedemann, lo que sugiere que este gen es un candidato a supres Para este gen se han encontrado tres variantes del transcrito que codifican dos isoformas diferentes.

50

55

60

65

Gen	Descripción (funciones, dominios)
5 10 15 20 25	<p>TULP4</p> <p>La proteína 4 de tipo Tubby contiene los dominios WD40 y SOCS. Las repeticiones WD-40 (también conocidas como repeticiones de WD o beta-transducina) son motivos cortos de ~40 aminoácidos, que a menudo terminan en un dipéptido Trp-Asp (W-D). Las proteínas que contienen WD tienen de 4 a 16 unidades repetitivas, todas las cuales se cree que forman una estructura de hélice beta circularizada. Las proteínas de repetición de WD son una familia grande que se encuentra en todos los eucariotas y están implicadas en diversas funciones que van desde la transducción de señales y la regulación de la transcripción hasta el control del ciclo celular y la apoptosis. La función común subyacente de todas las proteínas de repetición WD es la coordinación de ensamblajes complejos de multiproteínas, donde las unidades repetitivas sirven como armazón rígido para interacciones de proteínas. La especificidad de las proteínas está determinada por las secuencias fuera de las repeticiones mismas. Ejemplos de tales complejos son las proteínas G (la subunidad beta es una beta-hélice), el factor de transcripción TAFII y la ubiquitina ligasa E3.</p> <p>La caja SOCS se identificó primero en las proteínas que contienen el dominio SH2 de la familia de supresores de la señalización de citocinas (SOCS) pero más tarde se encontró en: la familia WSB (proteínas que contienen repeticiones WD—4D con una caja SOCS), la familia SSB (proteínas que contienen el dominio SPRY con una caja SOCS), la familia ASB (proteínas que contienen la repetición de anquirina con una caja SOCS) y GTPasas de tipo ras y ras.</p> <p>La caja SOCS encontrada en estas proteínas es un dominio carboxi-terminal de aproximadamente 50 aminoácidos compuesto por dos bloques de residuos bien conservados separados por entre 2 y 10 residuos no conservados. La región conservada C-terminal es una secuencia rica en L/P de función desconocida, mientras que la región conservada N-terminal es una caja BC de consenso, que se une al complejo Elongina BC. Se ha propuesto que esta asociación podría acoplar proteínas unidas a la ubiquitinación o compartimentos proteasómicos.</p>
30	<p>AA741300 Proteína desconocida (Nueva proteína)</p>
35 40	<p>ESCO1</p> <p>El establecimiento del homólogo 1 de cohesión 1 (ESCO1) pertenece a una familia conservada de acetiltransferasas que participan en la cohesión de la cromátida hermana.</p>
45 50 55 60 65	<p>EIFSB</p> <p>El inicio exacto de la traducción en eucariotas es complejo y requiere muchos factores, algunos de los cuales están compuestos de múltiples subunidades.. El proceso es más sencillo en procariotas, que tienen sólo tres factores de iniciación (IF1, IF2, IF3). Dos de estos factores se conservan en eucariotas: el homólogo de IF1 es eIF1A y el homólogo de IF2 es eIF5B. Este gen codifica eIF5B. Los factores eIF1A y eIF5B interactúan en el ribosoma junto con otros factores de iniciación y GTP para posicionar el ARNt de metionina de iniciación en el codón de iniciación del ARNm, de manera que la traducción se inicia con precisión.</p>

Gen	Descripción (funciones, dominios)
5	<p>ACTR2</p> <p>Se sabe que el homólogo de proteína 2 relacionado con la actina ARP2 (ACTR2) es un constituyente principal del complejo ARP2/3. Este complejo se encuentra en la superficie celular y es esencial para la forma y la motilidad celular a través del conjunto de actina lamellipodial y protrusión. Para este gen se han encontrado dos variantes del transcrito que codifican isoformas diferentes.</p>
10	<p>WNK1</p> <p>El gen WNK1 codifica una serina-treonina cinasa citoplásmica expresada en la nefrona distal. Las proteínas cinasas son un grupo de enzimas que poseen una subunidad catalítica que transfiere el fosfato gamma de los nucleótidos trifosfatos (a menudo ATP) a uno o más residuos de aminoácidos en una cadena lateral del sustrato proteínico, dando como resultado un cambio conformacional que afecta la función de la proteína. Las enzimas se dividen en dos clases amplias, caracterizadas con respecto a la especificidad del sustrato: serina/treonina específica y tirosina específica.</p> <p>La función de proteína cinasa se ha conservado evolutivamente de Escherichia coli a Homo sapiens. Las proteínas cinasas desempeñan un papel en una multitud de procesos celulares, incluyendo división, proliferación, apoptosis y diferenciación. La fosforilación normalmente da como resultado un cambio funcional de la proteína diana cambiando la actividad enzimática, la localización celular o la asociación con otras proteínas.</p> <p>Las subunidades catalíticas de las proteínas cinasas están altamente conservadas y se han resuelto varias estructuras, dando lugar a cribados grandes para desarrollar inhibidores específicos de la cinasa para el tratamiento de una serie de enfermedades. Las proteínas cinasas eucariotas son enzimas que pertenecen a una familia muy extensa de proteínas que comparten un núcleo catalítico conservado común con ambas serina/treonina y tirosina proteína cinasas. Hay un número de regiones conservadas en el dominio catalítico de las proteínas cinasas. En el extremo N-terminal del dominio catalítico hay un tramo de residuos rico en glicina en la proximidad de un residuo de lisina, que se ha demostrado que está implicado en la unión al ATP. En la parte central del dominio catalítico hay un residuo de ácido aspártico conservado que es importante para la actividad catalítica de la enzima.</p>
20	<p>WNK1</p> <p>Las subunidades catalíticas de las proteínas cinasas están altamente conservadas y se han resuelto varias estructuras, dando lugar a cribados grandes para desarrollar inhibidores específicos de la cinasa para el tratamiento de una serie de enfermedades. Las proteínas cinasas eucariotas son enzimas que pertenecen a una familia muy extensa de proteínas que comparten un núcleo catalítico conservado común con ambas serina/treonina y tirosina proteína cinasas. Hay un número de regiones conservadas en el dominio catalítico de las proteínas cinasas. En el extremo N-terminal del dominio catalítico hay un tramo de residuos rico en glicina en la proximidad de un residuo de lisina, que se ha demostrado que está implicado en la unión al ATP. En la parte central del dominio catalítico hay un residuo de ácido aspártico conservado que es importante para la actividad catalítica de la enzima.</p>
25	<p>WNK1</p> <p>Las subunidades catalíticas de las proteínas cinasas están altamente conservadas y se han resuelto varias estructuras, dando lugar a cribados grandes para desarrollar inhibidores específicos de la cinasa para el tratamiento de una serie de enfermedades. Las proteínas cinasas eucariotas son enzimas que pertenecen a una familia muy extensa de proteínas que comparten un núcleo catalítico conservado común con ambas serina/treonina y tirosina proteína cinasas. Hay un número de regiones conservadas en el dominio catalítico de las proteínas cinasas. En el extremo N-terminal del dominio catalítico hay un tramo de residuos rico en glicina en la proximidad de un residuo de lisina, que se ha demostrado que está implicado en la unión al ATP. En la parte central del dominio catalítico hay un residuo de ácido aspártico conservado que es importante para la actividad catalítica de la enzima.</p>
30	<p>WNK1</p> <p>Las subunidades catalíticas de las proteínas cinasas están altamente conservadas y se han resuelto varias estructuras, dando lugar a cribados grandes para desarrollar inhibidores específicos de la cinasa para el tratamiento de una serie de enfermedades. Las proteínas cinasas eucariotas son enzimas que pertenecen a una familia muy extensa de proteínas que comparten un núcleo catalítico conservado común con ambas serina/treonina y tirosina proteína cinasas. Hay un número de regiones conservadas en el dominio catalítico de las proteínas cinasas. En el extremo N-terminal del dominio catalítico hay un tramo de residuos rico en glicina en la proximidad de un residuo de lisina, que se ha demostrado que está implicado en la unión al ATP. En la parte central del dominio catalítico hay un residuo de ácido aspártico conservado que es importante para la actividad catalítica de la enzima.</p>
35	<p>COC</p> <p>La proteína codificada por este gen está altamente conservada en seres humanos, ratones y pollos, mostrando una identidad de aminoácidos del 94 % y 79 % de las secuencias de ser humano a ratón y de pollo, respectivamente. La hibridación a este gen se detectó en células en forma de huso situadas a lo largo de las fibras nerviosas entre el ganglio auditivo y el epitelio sensorial. Estas células acompañan a las neuritas en la habenula perforata, la abertura a través de la cual las neuritas se extienden para inervar las células ciliadas. Este y el patrón de expresión de este gen en el oído interno de pollo son paralelos a los finales histológicos de los depósitos acidófilos, en consonancia con la sustancia H con mucopolisacáridos, en los huesos temporales de pacientes de DFNA9 (sordera 9 neurosensorial no sindrómica autonómica dominante). Las mutaciones que causan DFNA9 se han notificado en este gen. El corte y empalme alternativo da como resultado múltiples variantes de transcripción que codifican la misma proteína. Se han descrito variantes de corte y empalme adicionales que codifican distintas isoformas, pero no se ha demostrado su validez biológica. La proteína contiene un dominio VWA en proteínas eucariotas extracelulares que median la adhesión a través de sitios de adhesión dependientes de iones metálicos (II/IIDAS). Recientemente se han identificado dominios VWA intracelulares y homólogos en procariotas. Los dominios VINA propuestos en las subunidades beta de la integrina han sido recientemente justificados usando procedimientos basados en la secuencia.</p>
40	<p>COC</p> <p>La proteína codificada por este gen está altamente conservada en seres humanos, ratones y pollos, mostrando una identidad de aminoácidos del 94 % y 79 % de las secuencias de ser humano a ratón y de pollo, respectivamente. La hibridación a este gen se detectó en células en forma de huso situadas a lo largo de las fibras nerviosas entre el ganglio auditivo y el epitelio sensorial. Estas células acompañan a las neuritas en la habenula perforata, la abertura a través de la cual las neuritas se extienden para inervar las células ciliadas. Este y el patrón de expresión de este gen en el oído interno de pollo son paralelos a los finales histológicos de los depósitos acidófilos, en consonancia con la sustancia H con mucopolisacáridos, en los huesos temporales de pacientes de DFNA9 (sordera 9 neurosensorial no sindrómica autonómica dominante). Las mutaciones que causan DFNA9 se han notificado en este gen. El corte y empalme alternativo da como resultado múltiples variantes de transcripción que codifican la misma proteína. Se han descrito variantes de corte y empalme adicionales que codifican distintas isoformas, pero no se ha demostrado su validez biológica. La proteína contiene un dominio VWA en proteínas eucariotas extracelulares que median la adhesión a través de sitios de adhesión dependientes de iones metálicos (II/IIDAS). Recientemente se han identificado dominios VWA intracelulares y homólogos en procariotas. Los dominios VINA propuestos en las subunidades beta de la integrina han sido recientemente justificados usando procedimientos basados en la secuencia.</p>
45	<p>COC</p> <p>La proteína codificada por este gen está altamente conservada en seres humanos, ratones y pollos, mostrando una identidad de aminoácidos del 94 % y 79 % de las secuencias de ser humano a ratón y de pollo, respectivamente. La hibridación a este gen se detectó en células en forma de huso situadas a lo largo de las fibras nerviosas entre el ganglio auditivo y el epitelio sensorial. Estas células acompañan a las neuritas en la habenula perforata, la abertura a través de la cual las neuritas se extienden para inervar las células ciliadas. Este y el patrón de expresión de este gen en el oído interno de pollo son paralelos a los finales histológicos de los depósitos acidófilos, en consonancia con la sustancia H con mucopolisacáridos, en los huesos temporales de pacientes de DFNA9 (sordera 9 neurosensorial no sindrómica autonómica dominante). Las mutaciones que causan DFNA9 se han notificado en este gen. El corte y empalme alternativo da como resultado múltiples variantes de transcripción que codifican la misma proteína. Se han descrito variantes de corte y empalme adicionales que codifican distintas isoformas, pero no se ha demostrado su validez biológica. La proteína contiene un dominio VWA en proteínas eucariotas extracelulares que median la adhesión a través de sitios de adhesión dependientes de iones metálicos (II/IIDAS). Recientemente se han identificado dominios VWA intracelulares y homólogos en procariotas. Los dominios VINA propuestos en las subunidades beta de la integrina han sido recientemente justificados usando procedimientos basados en la secuencia.</p>
50	<p>COC</p> <p>La proteína codificada por este gen está altamente conservada en seres humanos, ratones y pollos, mostrando una identidad de aminoácidos del 94 % y 79 % de las secuencias de ser humano a ratón y de pollo, respectivamente. La hibridación a este gen se detectó en células en forma de huso situadas a lo largo de las fibras nerviosas entre el ganglio auditivo y el epitelio sensorial. Estas células acompañan a las neuritas en la habenula perforata, la abertura a través de la cual las neuritas se extienden para inervar las células ciliadas. Este y el patrón de expresión de este gen en el oído interno de pollo son paralelos a los finales histológicos de los depósitos acidófilos, en consonancia con la sustancia H con mucopolisacáridos, en los huesos temporales de pacientes de DFNA9 (sordera 9 neurosensorial no sindrómica autonómica dominante). Las mutaciones que causan DFNA9 se han notificado en este gen. El corte y empalme alternativo da como resultado múltiples variantes de transcripción que codifican la misma proteína. Se han descrito variantes de corte y empalme adicionales que codifican distintas isoformas, pero no se ha demostrado su validez biológica. La proteína contiene un dominio VWA en proteínas eucariotas extracelulares que median la adhesión a través de sitios de adhesión dependientes de iones metálicos (II/IIDAS). Recientemente se han identificado dominios VWA intracelulares y homólogos en procariotas. Los dominios VINA propuestos en las subunidades beta de la integrina han sido recientemente justificados usando procedimientos basados en la secuencia.</p>
55	<p>COC</p> <p>La proteína codificada por este gen está altamente conservada en seres humanos, ratones y pollos, mostrando una identidad de aminoácidos del 94 % y 79 % de las secuencias de ser humano a ratón y de pollo, respectivamente. La hibridación a este gen se detectó en células en forma de huso situadas a lo largo de las fibras nerviosas entre el ganglio auditivo y el epitelio sensorial. Estas células acompañan a las neuritas en la habenula perforata, la abertura a través de la cual las neuritas se extienden para inervar las células ciliadas. Este y el patrón de expresión de este gen en el oído interno de pollo son paralelos a los finales histológicos de los depósitos acidófilos, en consonancia con la sustancia H con mucopolisacáridos, en los huesos temporales de pacientes de DFNA9 (sordera 9 neurosensorial no sindrómica autonómica dominante). Las mutaciones que causan DFNA9 se han notificado en este gen. El corte y empalme alternativo da como resultado múltiples variantes de transcripción que codifican la misma proteína. Se han descrito variantes de corte y empalme adicionales que codifican distintas isoformas, pero no se ha demostrado su validez biológica. La proteína contiene un dominio VWA en proteínas eucariotas extracelulares que median la adhesión a través de sitios de adhesión dependientes de iones metálicos (II/IIDAS). Recientemente se han identificado dominios VWA intracelulares y homólogos en procariotas. Los dominios VINA propuestos en las subunidades beta de la integrina han sido recientemente justificados usando procedimientos basados en la secuencia.</p>
60	<p>COC</p> <p>La proteína codificada por este gen está altamente conservada en seres humanos, ratones y pollos, mostrando una identidad de aminoácidos del 94 % y 79 % de las secuencias de ser humano a ratón y de pollo, respectivamente. La hibridación a este gen se detectó en células en forma de huso situadas a lo largo de las fibras nerviosas entre el ganglio auditivo y el epitelio sensorial. Estas células acompañan a las neuritas en la habenula perforata, la abertura a través de la cual las neuritas se extienden para inervar las células ciliadas. Este y el patrón de expresión de este gen en el oído interno de pollo son paralelos a los finales histológicos de los depósitos acidófilos, en consonancia con la sustancia H con mucopolisacáridos, en los huesos temporales de pacientes de DFNA9 (sordera 9 neurosensorial no sindrómica autonómica dominante). Las mutaciones que causan DFNA9 se han notificado en este gen. El corte y empalme alternativo da como resultado múltiples variantes de transcripción que codifican la misma proteína. Se han descrito variantes de corte y empalme adicionales que codifican distintas isoformas, pero no se ha demostrado su validez biológica. La proteína contiene un dominio VWA en proteínas eucariotas extracelulares que median la adhesión a través de sitios de adhesión dependientes de iones metálicos (II/IIDAS). Recientemente se han identificado dominios VWA intracelulares y homólogos en procariotas. Los dominios VINA propuestos en las subunidades beta de la integrina han sido recientemente justificados usando procedimientos basados en la secuencia.</p>
65	<p>COC</p> <p>La proteína codificada por este gen está altamente conservada en seres humanos, ratones y pollos, mostrando una identidad de aminoácidos del 94 % y 79 % de las secuencias de ser humano a ratón y de pollo, respectivamente. La hibridación a este gen se detectó en células en forma de huso situadas a lo largo de las fibras nerviosas entre el ganglio auditivo y el epitelio sensorial. Estas células acompañan a las neuritas en la habenula perforata, la abertura a través de la cual las neuritas se extienden para inervar las células ciliadas. Este y el patrón de expresión de este gen en el oído interno de pollo son paralelos a los finales histológicos de los depósitos acidófilos, en consonancia con la sustancia H con mucopolisacáridos, en los huesos temporales de pacientes de DFNA9 (sordera 9 neurosensorial no sindrómica autonómica dominante). Las mutaciones que causan DFNA9 se han notificado en este gen. El corte y empalme alternativo da como resultado múltiples variantes de transcripción que codifican la misma proteína. Se han descrito variantes de corte y empalme adicionales que codifican distintas isoformas, pero no se ha demostrado su validez biológica. La proteína contiene un dominio VWA en proteínas eucariotas extracelulares que median la adhesión a través de sitios de adhesión dependientes de iones metálicos (II/IIDAS). Recientemente se han identificado dominios VWA intracelulares y homólogos en procariotas. Los dominios VINA propuestos en las subunidades beta de la integrina han sido recientemente justificados usando procedimientos basados en la secuencia.</p>

5	Gen	<p>Descripción (funciones, dominios)</p> <p>La proteína codificada por este gen se une a una secuencia de ADN específica aguas arriba de la secuencia reguladora aguas arriba del promotor del núcleo y el segundo potenciador del virus de la hepatitis B humano (HBV). A través de esta unión, reprime la actividad promotora del núcleo de HBV, la transcripción de genes de HBV y la producción de viriones de HBV. La proteína muestra similitudes de secuencias con otras proteínas estructurales de unión al ADN tales como gallina, oncoproteínas de la familia MYC y la oncoproteína MOS. También puede estar implicado en la protección de células frente a la apoptosis y en el corte y empalme pre-ARNm. Se han descrito varias variantes de transcripción que codifican diferentes isoformas para este gen, pero solo se ha determinado la naturaleza de longitud completa de solo dos de ellas. Los miembros de esta familia pertenecen a la superfamilia del colágeno. Los colágenos son generalmente proteínas estructurales extracelulares implicadas en la formación de la estructura del tejido conectivo. La secuencia es, predominantemente, repeticiones de G-X-Y y las cadenas polipeptídicas forman una triple hélice, La primera posición de la repetición es glicina, la segunda y tercera posiciones pueden ser cualquier residuo pero frecuentemente son prolina e hidroxiprolina. Los colágenos son modificados después de la traducción por la hidroxilasa de prolina para formar los residuos de hidroxiprolina. La hidroxilación defectuosa es la causa del escorbuto. Algunos miembros de la superfamilia del colágeno no están involucrados en la estructura del tejido conjuntivo, pero comparten la misma estructura helicoidal triple.</p>
10	SON	
15		
20		
25		<p>Este gen codifica una proteína en espiral de gran tamaño que forma filamentos intranucleares unidos a la superficie interna de complejos de poros nucleares (NPC). La proteína interacciona directamente con varios componentes del NPC. Se requiere para la exportación nuclear de ARNm y algunas proteínas. Se producen fusiones oncogénicas del extremo 5' de este gen con varios genes de cinasa diferentes en algunas neoplasias. Los filamentos intermedios (IF) son proteínas que son componentes primordiales del citoesqueleto y la envoltura nuclear. Generalmente forman estructuras filamentosas de 8 a 14 nm de ancho. Las proteínas IF son miembros de una familia multigénica muy grande de proteínas que se ha subdividido en cinco subgrupos principales: Tipo I: citoqueratinas ácidas. Tipo II: citoqueratinas básicas. Tipo III: Vimentina, desmina, proteína ácida fibrilar glial (GFAP), periferina y plasticina. Tipo IV: Neurofilamentos L, H y M. alfa-internexina y nestina. Tipo V: : laminas nucleares A, B1, B2 y C. Todas las proteínas IF son estructuralmente similares en que consisten en: un dominio de varilla central que comprende unos 300 a 350 residuos que está dispuesto en alfa-hélice enrollada en espiral, con al menos dos Interrupciones características cortas; un dominio no helicoidal N-terminal (cabeza) de longitud variable; y un dominio C-terminal (cola) que también es no helicoidal y que muestra una variación de longitud extrema entre diferentes proteínas IF. Aunque las proteínas IF están evolutiva y estructuralmente relacionadas, tienen homologías de secuencia limitadas, excepto en varias regiones del dominio de varilla. Esta entrada representa el dominio de varilla central encontrado en las proteínas IF.</p>
30	TPR	
35		
40	Gen	<p>Descripción (funciones, dominios)</p> <p>El polipéptido secretado, codificado por este gen, se une e inactiva miembros de las proteínas de señalización de la superfamilia del factor de crecimiento transformante beta (TGF-beta), tales como la proteína 4-morfogénica ósea [BMP4]. Al difundirse a través de matrices extracelulares más eficientemente que los miembros de la superfamilia TGF-beta, esta proteína puede tener un papel principal en la creación de gradientes morfogénicos. La proteína parece tener un efecto pleiotrópico tanto en el desarrollo como en etapas posteriores. Originalmente se aisló de <i>Xenopus</i> basado en su capacidad para restaurar el eje normal del cuerpo dorsal-ventral en embriones que habían sido ventralizados artificialmente por tratamiento UV. Los resultados de la inactivación en ratón de los ortólogos sugieren que está involucrado en numerosos procesos de desarrollo, como la fusión del tubo neural y la formación de la articulación. Recientemente, se identificaron varias mutaciones dominantes de NOG humana en familias no relacionadas con sinfalangismo proximal (SYIV11) y síndrome de sinostosis múltiple [SYNS1]; tanto SYM1 como SYNS1 tienen múltiples fusión de las articulaciones como su característica principal y el mapa 10 de la misma región (17q22) que este gen. Todas estas mutaciones alteraron los residuos de aminoácidos conservados evolutivamente. La secuencia de aminoácidos de este gen humano es altamente homóloga a la de <i>Xenopus</i>, rata y ratón. Esta familia se compone de las proteínas eucarióticas Noggin. Noggin es una glicoproteína que enlaza las proteínas morfogénicas óseas (BMP) de forma selectiva y, cuando se añade a los osteoblastos, se opone a los efectos de las BMP. Se ha encontrado que el noggin detiene la diferenciación de las células del estroma, evitando la maduración celular.</p>
45	NOG	
50		
55		
60		
65		

TABLA 12A

Símbolo del gen	Número de acceso en GENBANK®
CDKN1C	gi 169790897 ref NM_000076.2
TCF7L2	gi 170014695 ref NM_030756.3
CLC	gi 20357558 ref NM_001828.4
WFS1	NM_006005
TSPAN8	NM_004616
THADA	NM_022065
TCF7L2	NM_030756
SLC30A8	NM_173851
PPARG	NM_138712
NOTCH2	NM_024408
LGR5	NM_003667
KCNJ11	NM_000525
JAZF1	NM_175061
IGF2BP2	NM_001007225
HHEX-IDE	NM_002729, NM_004969
FTO	NM_001080432
CDKN2B	NM_078487
CDKN2A	NM_058195

Firma de 3 genes

Firma de 3 genes

Símbolo del gen	Número de acceso en GenBank
TOP1	NM_003286
CD24	NM_013230
STAP1	NM_012108

Firma de 10 genes

Símbolo del gen Número de acceso en GenBank

TULP4	NM_020245
AA741300	AA741300
ESCO1	NM_052911
EIF5B	NM_015904
ACTR2	NM_001005386
WNK1	NM_018979
COCH	NM_004086
SON	NM_138927
TPR	NM_003292
NOG	NM_005450

TABLA 12B

Firma de 3 genes				
Símbolo del gen	Acceso	Secuencia del cebador superior	Secuencia de la sonda	Secuencia del cebador inferior
TOP1	NM_003286	CCCTGTACTTCATCGA CAAGC	AGCAGCAGCCCACA GTGT	AGAGCAGGCAATGAAAAGGAGGAAG
CD24	NM_013230	GCCAGGGCAATGATG AATG	CTCAATATGGATAAT CAAGAGTTGCT	TCTACCCCAGATCCAAGCAGCCT
STAP1	NM_012108	TGAAAAGAAGTGTGCG AAATTC	CACTTTCTGTGTTCT CTGTCTTCAG	CCTTGTTTTGCCGAAAGAGGAAGTACA

STAP1-331F22	TGAAAAGAAGTGTGCGAAATTC
STAP1-407R25	CACTTTCTGTGTTCTCTGTCTTCAG
STAP1-355P27	CCTTGTTTTGCCGAAAGAGGAAGTACA
CD24_996_U19	GCCAGGGCAATGATGAATG
CD24_1069_L26	CTCAATATGGATAATCAAGAGTTGCT
CD24_1019_P24	TCTACCCCAGATCCAAGCAGCCT
TOP1_1679_U22	CCCTGTACTTCATCGACAAGC
TOP1_1762_L18	AGCAGCAGCCCACAGTGT
TOP1_1708_P26	AGAGCAGGCAATGAAAAGGAGGAAG

TABLA 12C

Firma de 10 genes			
Símbolo del gen	Secuencia del cebador superior	Secuencia de la sonda	Secuencia del cebador inferior
TULP4	GAAGAGTGTGTGTCTATG TGCATT TAAA	CAAGTTGCTCCAT CTGATTCTTA AATT	CACATTCACACGGGAAGACAGGCT CA
AA741300	No disponible		
ESCO1	CTAAACGGCAGCACAAAA GGA	CATGTCTTATGGCT AACACGTTT CTT	TGCAAACCAACAGACTCAGCAAAC AAGG
EIF5B	CAGCCAAGGCATCAAGAT CA	GAGCGCCATTGAC AAGCAAT	TCATCCTTGGTGCTGTCTTCGCTCT TGTT
ACTR2	CATTCAACTCCAGGACAT GGAA	TCCCCAAGACACC AGAATAAAA CT	AGGCCTCTCTCTGCCCTTTGACTG GA
WNK1	GCATGCTTGAGATGGCTA CATC	TGGTCACGCGACG GTAGAT	TCCTTACTCGGAGTGCCAAAATGCT GC
COCH	CCATTTAGGCAAATAAGC ACTCCTT	GCCTCAGCAGTGT TTTTAACAAA G	AAGCCGCTGCCTTCTGGTTACAATT TACA
SON	GCTCTGCTCAGCCCTAAA GAAA	TCCTCAATATTGGC AGAAAATCC T	CCTCCCCCTCCTAAAGAGACACTG CCTG
TPR	CTGCCCAAGTCTGTCCAG AAC	CCTGACTGTGGGA CAACCTCTT	ATCAGCAATCCGAGATCGATGGCC T
NOG	CACCCGGACACTTGATCG AT	GTTTCATTGAAAACC CTCGCTAG A	ACCGCCTCCAACCAGTCCACCAC

Símbolo del gen	Secuencias de cebadores	Símbolo del gen	Secuencias de cebadores
TULP4- F1	GAAGAGTGTGTGTCTATGTGCATT TAAA	COCH- F1	CCATTTAGGCAAATAAGCACTCCTT
TULP4- R1	CAAGTTGCTCCATCTGATTCTTAAA TT	COCH- R1	GCCTCAGCAGTGTTTTTAACAAAAG
TULP4- Pro1	CACATTCACACGGGAAGACAGGCT CA	COCH- Pro1	AAGCCGCTGCCTTCTGGTTACAATTT ACA
ESCO1- F	CTAAACGGCAGCACAAAAGGA	SON-F1	GCTCTGCTCAGCCCTAAAGAAA
ESCO1- R	CATGTCTTATGGCTAACACGTTTCT T	SON- R1	TCCTCAATATTGGCAGAAAATCCT
ESCO1- Pro2	TGCAAACCAACAGACTCAGCAAAC AAGG	SON- Pro1	CCTCCCCCTCCTAAAGAGACACTGC CTG
NOG- F1	CACCCGGACACTTGATCGAT	EIF5B- F1	CAGCCAAGGCATCAAGATCA
NOG- R1	GTTTCATTGAAAACCCTCGCTAGA	EIF5B- R1	GAGCGCCATTGACAAGCAAT
NOG- Pro1	ACCGCCTCCAACCAGTCCACCAC	EIF5B- Pro1	TCATCCTTGGTGCTGTCTTCGCTCTT GTT
WNK1- F1	GCATGCTTGAGATGGCTACATC	ACTR2- F1	CATTCAACTCCAGGACATGGAA
WNK1- R1	TGGTCACGCGACGGTAGAT	ACTR2- R1	TCCCCAAGACACCAGAATAAACT
WNK1- Pro1	TCCTTACTCGGAGTGCCAAAATGC TGC	ACTR2- Pro1	AGGCCTCTCTCTGCCCTTTGACTGG A
		TPR-F1	CTGCCCAAGTCTGTCCAGAAC
		TPR-R1	CCTGACTGTGGGACAACCTCTT
		TPR- Pro1	ATCAGCAATCCGAGATCGATGGCCT

LISTA DE SECUENCIA

[0143]

- <110> Veridex, LLC
- <120> DIAGNÓSTICO DE MARCADORES DE DIABETES
- 5 <130> P060769EP
- <140> EP__
- <141> 2008-11-13
- <150> PCT/US2008/083424
- <151> 2008-11-13
- 10 <150> US 60/987,540
- <151> 2007-11-13
- <160> 87
- <170> PatentIn version 3.5
- <210> 1
- 15 <211> 21
- <212> DNA
- <213> Homo sapiens
- <400> 1
- ccctgtactt catcgacaag c 21
- 20 <210> 2
- <211> 19
- <212> DNA
- <213> Homo sapiens
- <400> 2
- 25 gccagggcaa tgatgaatg 19
- <210> 3
- <211> 22
- <212> DNA
- <213> Homo sapiens
- 30 <400> 3
- tgaaaagaac tgtgcgaaat tc 22
- <210> 4
- <211> 18
- <212> DNA
- 35 <213> Homo sapiens
- <400> 4
- agcagcagcc cacagtgt 18
- <210> 5
- <211> 26
- 40 <212> DNA
- <213> Homo sapiens
- <400> 5
- ctcaatatgg ataatcaaga gttgct 26
- <210> 6
- 45 <211> 25
- <212> DNA
- <213> Homo sapiens
- <400> 6
- cactttctgt gttctctgtc ttcag 25
- 50 <210> 7
- <211> 25
- <212> DNA
- <213> Homo sapiens
- <400> 7
- 55 agagcaggca atgaaaagga ggaag 25
- <210> 8
- <211> 24
- <212> DNA
- <213> Homo sapiens
- 60 <400> 8
- tctaccccca gatccaagca gcct 24
- <210> 9
- <211> 27
- <212> DNA
- 65 <213> Homo sapiens
- <400> 9

cctgttttg ccgaaagagg aagtaca 27
 <210> 10
 <211> 22
 <212> DNA
 5 <213> Homo sapiens
 <400> 10
 tgaaaagaac tgtcgaaat tc 22
 <210> 11
 <211> 25
 10 <212> DNA
 <213> Homo sapiens
 <400> 11
 cactttctgt gttctctgtc ttcag 25
 <210> 12
 15 <211> 27
 <212> DNA
 <213> Homo sapiens
 <400> 12
 cctgttttg ccgaaagagg aagtaca 27
 20 <210> 13
 <211> 19
 <212> DNA
 <213> Homo sapiens
 <400> 13
 25 gccagggcaa tgatgaatg 19
 <210> 14
 <211> 26
 <212> DNA
 <213> Homo sapiens
 30 <400> 14
 ctcaatatgg ataatcaaga gttgct 26
 <210> 15
 <211> 24
 <212> DNA
 35 <213> Homo sapiens
 <400> 15
 tctaccccca gatccaagca gcct 24
 <210> 16
 <211> 21
 40 <212> DNA
 <213> Homo sapiens
 <400> 16
 ccctgtactt catcgacaag c 21
 <210> 17
 45 <211> 18
 <212> DNA
 <213> Homo sapiens
 <400> 17
 agcagcagcc cacagtgt 18
 50 <210> 18
 <211> 25
 <212> DNA
 <213> Homo sapiens
 <400> 18
 55 agagcaggca atgaaaagga ggaag 25
 <210> 19
 <211> 28
 <212> DNA
 <213> Homo sapiens
 60 <400> 19
 gaagagtgtg tgtctatgtg cattaaa 28
 <211> 21
 <212> DNA
 65 <213> Homo sapiens
 <400> 20

ctaaacggca gcacaaaagg a 21
 <210> 21
 <211> 20
 <212> DNA
 5 <213> Homo sapiens
 <400> 21
 cagccaaggc atcaagatca 20
 <210> 22
 <211> 22
 10 <212> DNA
 <213> Homo sapiens
 <400> 22
 cattcaactc caggacatgg aa 22
 <210> 23
 15 <211> 22
 <212> DNA
 <213> Homo sapiens
 <400> 23
 gcatgcttga gatggctaca tc 22
 20 <210> 24
 <211> 25
 <212> DNA
 <213> Homo sapiens
 <400> 24
 25 ccatttaggc aaataagcac tcctt 25
 <210> 25
 <211> 22
 <212> DNA
 <213> Homo sapiens
 30 <400> 25
 gctctgctca gccctaaaga aa 22
 <210> 26
 <211> 21
 <212> DNA
 35 <213> Homo sapiens
 <400> 26
 ctgcccaagt ctgtccagaa c 21
 <210> 27
 <211> 20
 40 <212> DNA
 <213> Homo sapiens
 <400> 27
 caccgggaca ctgatcgat 20
 <210> 28
 45 <211> 27
 <212> DNA
 <213> Homo sapiens
 <400> 28
 caagtgctc catctgattc ttaaatt 27
 50 <210> 29
 <211> 26
 <212> DNA
 <213> Homo sapiens
 <400> 29
 55 catgtcttat ggctaacacg ttctt 26
 <210> 30
 <211> 20
 <212> DNA
 <213> Homo sapiens
 60 <400> 30
 gagcgccatt gacaagcaat 20
 <210> 31
 <211> 24
 <212> DNA
 65 <213> Homo sapiens
 <400> 31

tccccaagac accagaataa aact 24
 <210> 32
 <211> 19
 <212> DNA
 5 <213> Homo sapiens
 <400> 32
 tggtcacgcg acggtagat 19
 <210> 33
 <211> 24
 10 <212> DNA
 <213> Homo sapiens
 <400> 33
 gcctcagcag tgttttaac aaag 24
 <210> 34
 15 <211> 24
 <212> DNA
 <213> Homo sapiens
 tcctcaatat tggcagaaaa tcct 24
 <210> 35
 20 <211> 22
 <212> DNA
 <213> Homo sapiens
 <400> 35
 cctgactgtg ggacaacctc tt 22
 25 <210> 36
 <211> 23
 <212> DNA
 <213> Homo sapiens
 <400> 36
 30 gttcattgaa aaccctcgct aga 23
 <210> 37
 <211> 26
 <212> DNA
 <213> Homo sapiens
 35 <400> 37
 cacattcaca cgggaagaca ggctca 26
 <210> 38
 <211> 28
 <212> DNA
 40 <213> Homo sapiens
 <400> 38
 tgcaaaccaa cagactcagc aaacaagg 28
 <210> 39
 <211> 29
 45 <212> DNA
 <213> Homo sapiens
 <400> 39
 tcatccttgg tgctgtcttc gctctgtt 29
 <210> 40
 50 <211> 26
 <212> DNA
 <213> Homo sapiens
 <400> 40
 aggcctctct ctgcccttg actgga 26
 55 <210> 41
 <211> 27
 <212> DNA
 <213> Homo sapiens
 <400> 41
 60 tccttactcg gagtgccaaa atgctgc 27
 <210> 42
 <211> 29
 <212> DNA
 <213> Homo sapiens
 65 <400> 42
 aagccgctgc ctctggta caattaca 29

<210> 43
 <211> 28
 <212> DNA
 <213> Homo sapiens
 5 <400> 43
 cctccccctc ctaaagagac actgcctg 28
 <210> 44
 <211> 25
 <212> DNA
 10 <213> Homo sapiens
 <400> 44
 atcagcaatc cgagatcgat ggcct 25
 <210> 45
 <211> 24
 15 <212> DNA
 <213> Homo sapiens
 <400> 45
 accgcctcca accagttcca ccac 24
 <210> 46
 20 <211> 28
 <212> DNA
 <213> Homo sapiens
 <400> 46
 gaagagtgtg tgtctatgtg catttaa 28
 25 <210> 47
 <211> 27
 <212> DNA
 <213> Homo sapiens
 <400> 47
 30 caagttgctc catctgattc ttaaatt 27
 <210> 48
 <211> 26
 <212> DNA
 <213> Homo sapiens
 35 <400> 48
 cacattcaca cgggaagaca ggctca 26
 <210> 49
 <211> 21
 <212> DNA
 40 <213> Homo sapiens
 <400> 49
 ctaaacggca gcacaaaagg a 21
 <210> 50
 <211> 26
 45 <212> DNA
 <213> Homo sapiens
 <400> 50
 catgtcttat ggctaacacg tttctt 26
 <210> 51
 50 <211> 28
 <212> DNA
 <213> Homo sapiens
 <400> 51
 tgcaaaccaa cagactcagc aaacaagg 28
 55 <210> 52
 <211> 20
 <212> DNA
 <213> Homo sapiens
 <400> 52
 60 caccgggaca ctgatcgat 20
 <210> 53
 <211> 23
 <212> DNA
 <213> Homo sapiens
 65 <400> 53
 gttcattgaa aaccctcgct aga 23

<210> 54
 <211> 24
 <212> DNA
 <213> Homo sapiens
 5 <400> 54
 accgcctcca accagttcca ccac 24
 <210> 55
 <211> 22
 <212> DNA
 10 <213> Homo sapiens
 <400> 55
 gcatgcttga gatggctaca tc 22
 <210> 56
 <211> 19
 15 <212> DNA
 <213> Homo sapiens
 <400> 56
 tggtcacgcg acggtagat 19
 <210> 57
 20 <211> 27
 <212> DNA
 <213> Homo sapiens
 <400> 57
 tccttactcg gagtgccaaa atgctgc 27
 25 <210> 58
 <211> 25
 <212> DNA
 <213> Homo sapiens
 <400> 58
 30 ccatttaggc aaataagcac tcctt 25
 <210> 59
 <211> 24
 <212> DNA
 <213> Homo sapiens
 35 <400> 59
 gcctcagcag tgttttaac aaag 24
 <210> 60
 <211> 29
 <212> DNA
 40 <213> Homo sapiens
 <400> 60
 aagccgctgc ctctggta caattaca 29
 <210> 61
 <211> 22
 45 <212> DNA
 <213> Homo sapiens
 <400> 61
 gctctgctca gccctaaaga aa 22
 <210> 62
 50 <211> 24
 <212> DNA
 <213> Homo sapiens
 <400> 62
 tcctcaatat tggcagaaaa tcct 24
 55 <210> 63
 <211> 28
 <212> DNA
 <213> Homo sapiens
 <400> 63
 60 cctccccctc ctaaagagac actgcctg 28
 <210> 64
 <211> 20
 <212> DNA
 <213> Homo sapiens
 65 <400> 64
 cagccaaggc atcaagatca 20

<210> 65
 <211> 20
 <212> DNA
 <213> Homo sapiens
 5 <400> 65
 gagcgccatt gacaagcaat 20
 <210> 66
 <211> 29
 <212> DNA
 10 <213> Homo sapiens
 <400> 66
 tcacacctgg tgctgtcttc gctctgtt 29
 <210> 67
 <211> 22
 15 <212> DNA
 <213> Homo sapiens
 <400> 67
 cattcaactc caggacatgg aa 22
 <210> 68
 20 <211> 24
 <212> DNA
 <213> Homo sapiens
 <400> 68
 tccccagac accagaataa aact 24
 25 <210> 69
 <211> 26
 <212> DNA
 <213> Homo sapiens
 <400> 69
 30 aggcctctct ctgcccttg actgga 26
 <210> 70
 <211> 21
 <212> DNA
 <213> Homo sapiens
 35 <400> 70
 ctgcccaagt ctgtccagaa c 21
 <210> 71
 <211> 22
 <212> DNA
 40 <213> Homo sapiens
 <400> 71
 cctgactgtg ggacaacctc tt 22
 <210> 72
 <211> 25
 45 <212> DNA
 <213> Homo sapiens
 <400> 72
 atcagcaatc cgagatcgat ggcct 25
 <210> 73
 50 <211> 22
 <212> DNA
 <213> Homo sapiens
 <400> 73
 acctgagcgc tctaagaaa tg 22
 55 <210> 74
 <211> 19
 <212> DNA
 <213> Homo sapiens
 <400> 74
 60 agggccgcac cagttattc 19
 <210> 75
 <211> 23
 <212> DNA
 <213> Artificial Sequence
 65 <220>
 <223> Taqman probe

<220>
 <221> misc_feature
 <222> (1)..(1)
 <223> 6-carboxyfluorescein (FAM)-dATP
 5 <220>
 <221> misc_feature
 <222> (23) .. (23)
 <223> dCMP - BHQ1 (Black Hole Quencher-1)
 <400> 75
 10 ngcgcgcttt ggccttgatc aan 23
 <210> 76
 <211> 23
 <212> DNA
 <213> Homo sapiens
 15 <400> 76
 cgtcgacttc ttggtacat tcc 23
 <210> 77
 <211> 26
 <212> DNA
 20 <213> Homo sapiens
 <400> 77
 cacgacgcta aagctattct aaagac 26
 <210> 78
 <211> 21
 25 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Taqman probe
 <220>
 30 <221> misc_feature
 <222> (1)..(1)
 <223> 6-carboxyfluorescein (FAM)-dCTP
 <220>
 <221> misc_feature
 35 <222> (21) .. (21)
 <223> dCMP - BHQ1 (Black Hole Quencher-1)
 <400> 78
 nagccgctgt cgctcgtcac n 21
 <210> 79
 40 <211> 18
 <212> DNA
 <213> Homo sapiens
 <400> 79
 gaaagcgcg ccatcaac 18
 45 <210> 80
 <211> 23
 <212> DNA
 <213> Homo sapiens
 <400> 80
 50 cagctcgtag tatttcgctt get 23
 <210> 81
 <211> 21
 <212> DNA
 <213> Artificial Sequence
 55 <220>
 <223> Taqman probe
 <220>
 <221> misc_feature
 <222> (1)..(1)
 60 <223> 6-carboxyfluorescein (FAM)-dTTP
 <220>
 <221> misc_feature
 <222> (21) .. (21)
 <223> dGMP - BHQ1 (Black Hole Quencher-1)
 65 <400> 81
 nccttgggcg gaggtggcat n 21

<210> 82
 <211> 21
 <212> DNA
 <213> Homo sapiens
 5 <400> 82
 gctaccCGtg ccatacacag a 21
 <210> 83
 <211> 25
 <212> DNA
 10 <213> Homo sapiens
 <400> 83
 gcagatatgg ttcatcaag aaaca 25
 <210> 84
 <211> 28
 15 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Taqman probe
 <220>
 20 <221> misc_feature
 <222> (1)..(1)
 <223> 6-carboxyfluorescein (FAM)-dTTP
 <220>
 <221> misc_feature
 25 <222> (28) .. (28)
 <223> dAMP - BHQ1 (Black Hole Quencher-1)
 <400> 84
 ntctactgtg acaatcaaag ggcgaccn 28
 <210> 85
 30 <211> 18
 <212> DNA
 <213> Homo sapiens
 <400> 85
 cctggcacc agcacaat 18
 35 <210> 86
 <211> 21
 <212> DNA
 <213> Homo sapiens
 <400> 86
 40 gccgatccac acggagtact t 21
 <210> 87
 <211> 27
 <212> DNA
 <213> Artificial Sequence
 45 <220>
 <223> Taqman probe
 <220>
 <221> misc_feature
 <222> (1)..(1)
 50 <223> 6-carboxyfluorescein (FAM)-dATP
 <220>
 <221> misc_feature
 <222> (27)..(27)
 <223> dCMP - BHQ1 (Black Hole Quencher-1)
 55 <400> 87
 ntcaagatca ttgctctcc tgagcgn 27

 60

 65

Reivindicaciones

1. Un procedimiento para diagnosticar diabetes mellitus en un paciente, que comprende las etapas de:
 - 5 a. medir el perfil de expresión génica de una firma génica que comprende los genes TULP4, AA741300, ESCO1, EIF5B, ACTR2, WNK1, COCH, SON, TPR y NOG en una muestra de ensayo tomada de un paciente; y
 - b. comparar dicho perfil de expresión génica con un perfil de expresión de genes de diagnóstico de dicha firma génica;
 - 10 c. determinar un estado de enfermedad diabética en dicho paciente basado al menos en parte en una coincidencia sustancial entre dicho perfil de expresión de genes y dicho perfil de expresión de genes de diagnóstico; y
 - d. mostrar dicha determinación a un profesional médico.

2. Un procedimiento para diagnosticar un cambio en el estado de enfermedad diabética de un paciente, que comprende las etapas de:
 - 15 a. medir un primer perfil de expresión génica de una firma génica en una primera muestra de ensayo tomada en un primer punto de tiempo, comprendiendo dicha firma génica los genes TULP4, AA741300, ESCO1, EIF5B, ACTR2, WNK1, COCH, SON, TPR y NOG;
 - 20 b. medir un segundo perfil de expresión de dicha firma génica en una segunda muestra de ensayo tomada en un segundo punto de tiempo;
 - c. comparar dicho primer perfil de expresión con dicho segundo perfil de expresión;
 - d. determinar un cambio en el estado de enfermedad diabética en dicho paciente basado al menos en parte en una diferencia sustancial entre dicho primer perfil de expresión de genes y dicho segundo perfil de expresión de genes de diagnóstico; y
 - 25 e. mostrar dicha determinación a un profesional médico.

3. El procedimiento de acuerdo con la reivindicación 1 o 2, en el que dicha etapa de determinación se ejecuta mediante un sistema informático, ejecutando dicho sistema informático uno o más algoritmos seleccionados del grupo que consiste en una combinación lineal de señales de expresión génica, modelo de regresión lineal, modelo de regresión logística, modelo de análisis de discriminación lineal (LDA), el modelo de vecino más próximo y el análisis de predicción de microarrays (PAM).

4. El procedimiento de acuerdo con una cualquiera de las reivindicaciones 1 a 3, en el que dicha etapa de determinación comprende además un análisis del perfil de enfermedad metabólica del paciente.

5. El procedimiento de acuerdo con una cualquiera de las reivindicaciones 1 a 4, en el que dicho estado de enfermedad diabética es un estado de enfermedad pre-diabética o un estado de enfermedad de diabetes Tipo 2.

6. El procedimiento de acuerdo con una cualquiera de las reivindicaciones 1 a 5, en el que dicha muestra de ensayo es una muestra de sangre.

7. El procedimiento de acuerdo con una cualquiera de las reivindicaciones 1 a 6, en el que dicha muestra de ensayo comprende PBMC o células CD11c⁺ o CD11b⁺ o Emr⁺ o [CD11b⁺CD11c⁺] o [Emr⁺ CD11b⁺] o [Emr⁺CD11c⁺] o [Emr⁺CD11b⁺CD11c⁺] o monocitos CD14⁺.

8. El procedimiento de acuerdo con una cualquiera de las reivindicaciones 1 a 7, en el que dicha etapa de medición implica PCR en tiempo real, un ensayo inmunoquímico o hibridación de oligonucleótidos específicos.

9. El uso de un kit para evaluar la susceptibilidad de un paciente a la diabetes en el que se realiza la evaluación con un aparato de ensayo, comprendiendo el kit:
 - a. reactivos para recoger una muestra de ensayo de un paciente; y
 - 55 b. reactivos para medir el perfil de expresión génica de una firma génica que comprende los genes TULP4, AA741300, ESCO1, EIF5B, ACTR2, WNK1, COCH, SON, TPR y NOG en una muestra de ensayo de un paciente, en el que los reactivos son para: (i) PCR en tiempo real, en la que dichos reactivos comprenden cebadores para la amplificación y detección del perfil de expresión génica o (ii) un ensayo inmunohistoquímico, en el que dichos reactivos comprenden anticuerpos para la detección de las proteínas TULP4, AA741300, ESCO1, EIF5B, ACTR2, WNK1, COCH, SON, TPR y NOG; y envasado de los mismos.
 - 60

10. El uso de la reivindicación 9, en el que dichos reactivos para recoger una muestra de ensayo son reactivos para recoger una muestra de sangre.

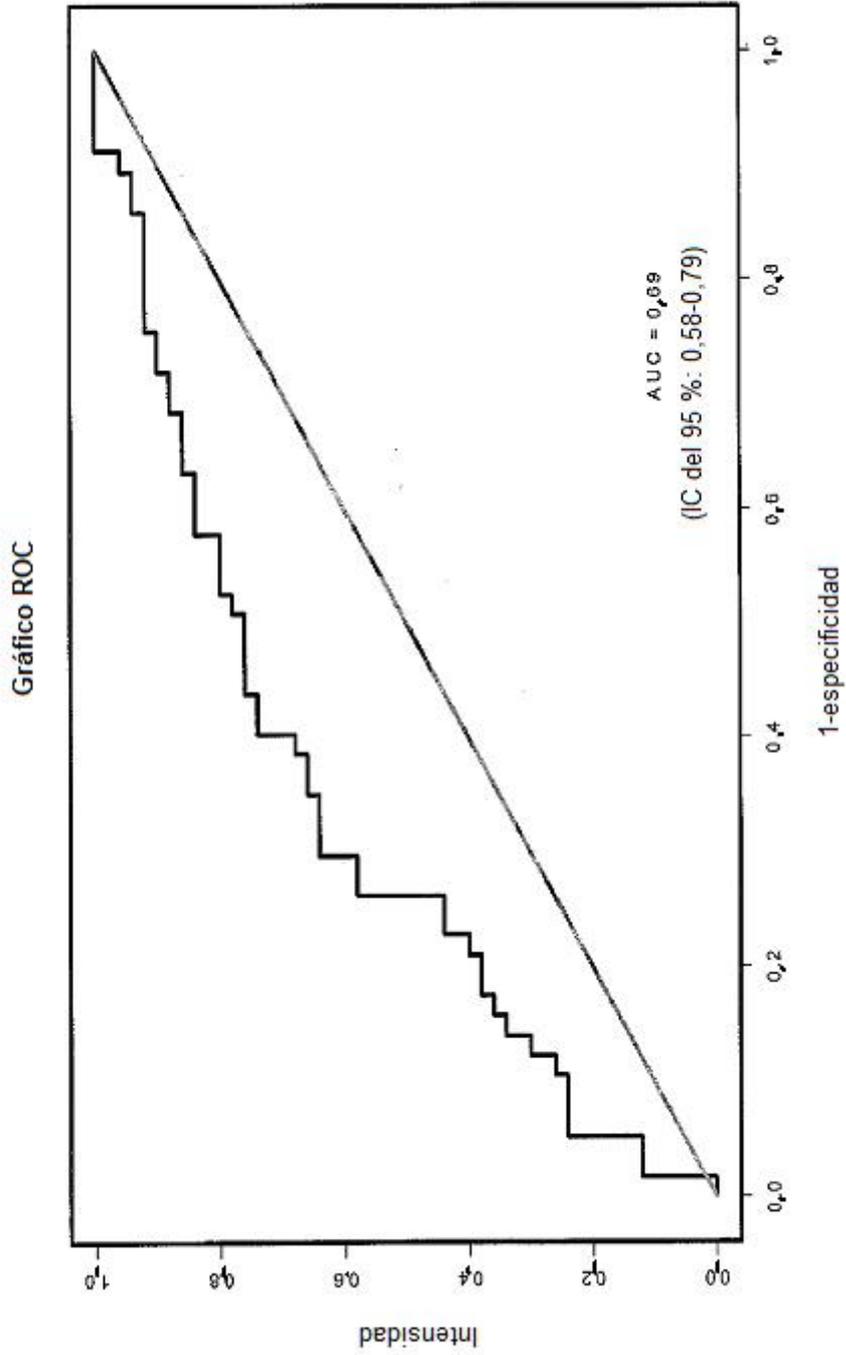


FIGURA 1:

FIGURA 2A:

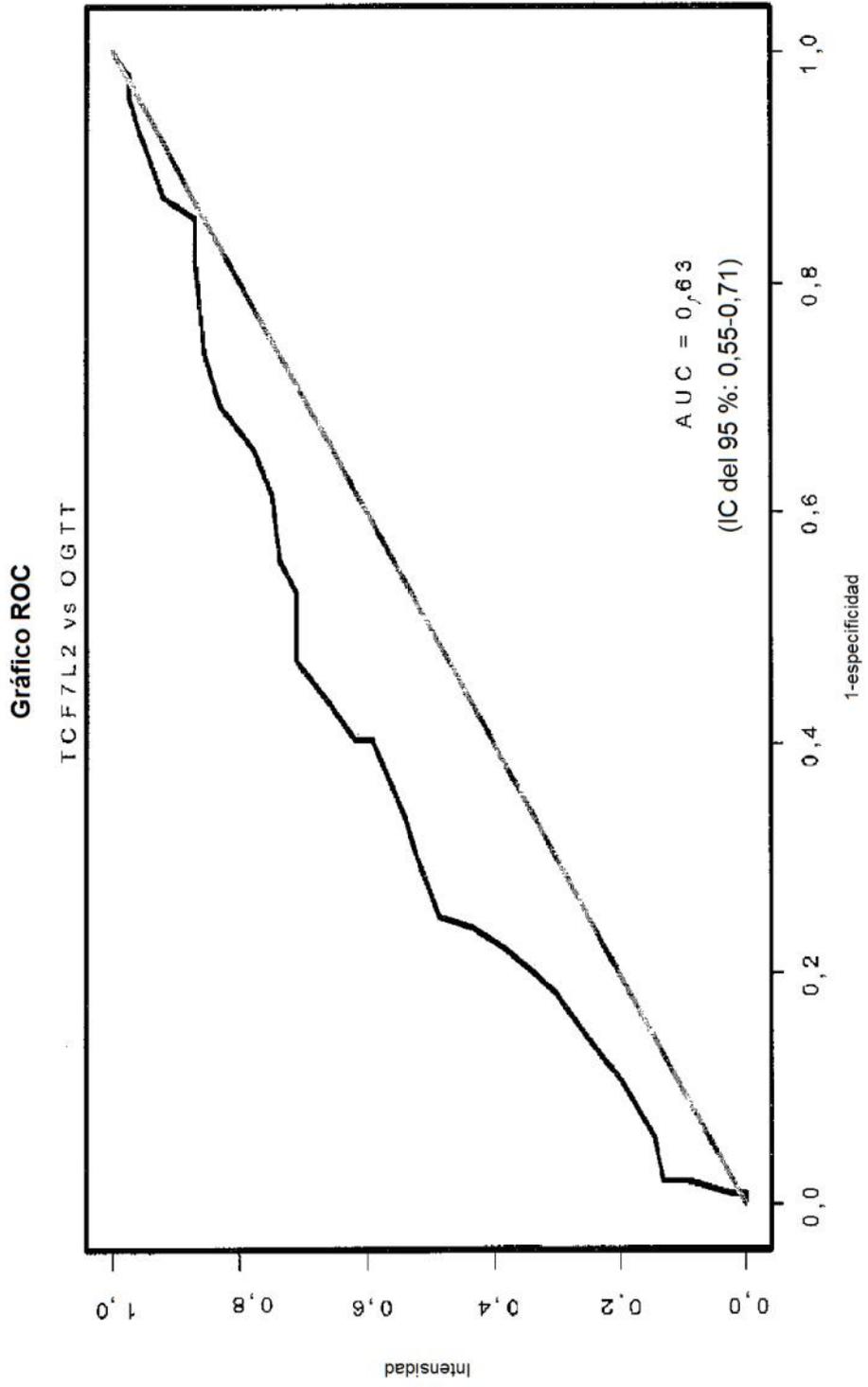


FIGURE 2B:

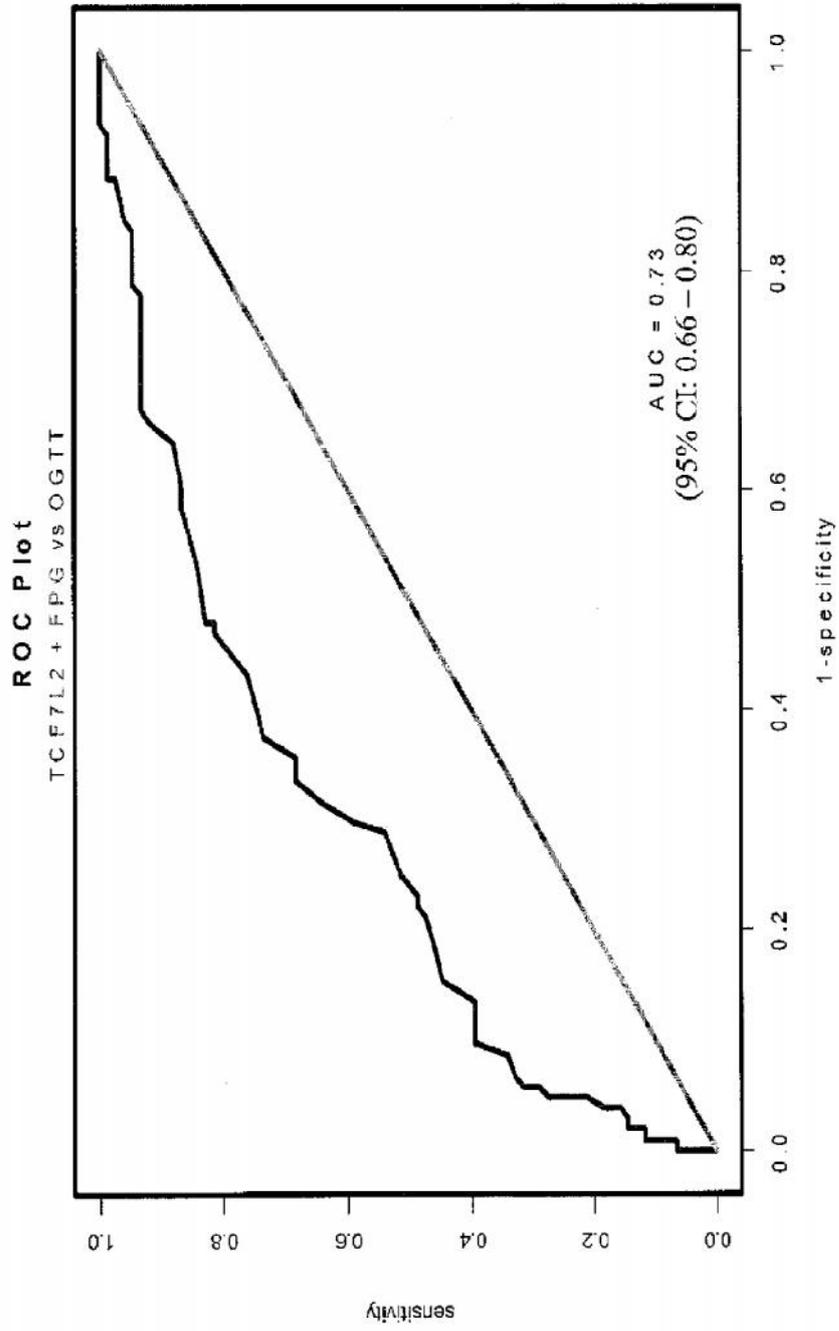


FIGURA 3:

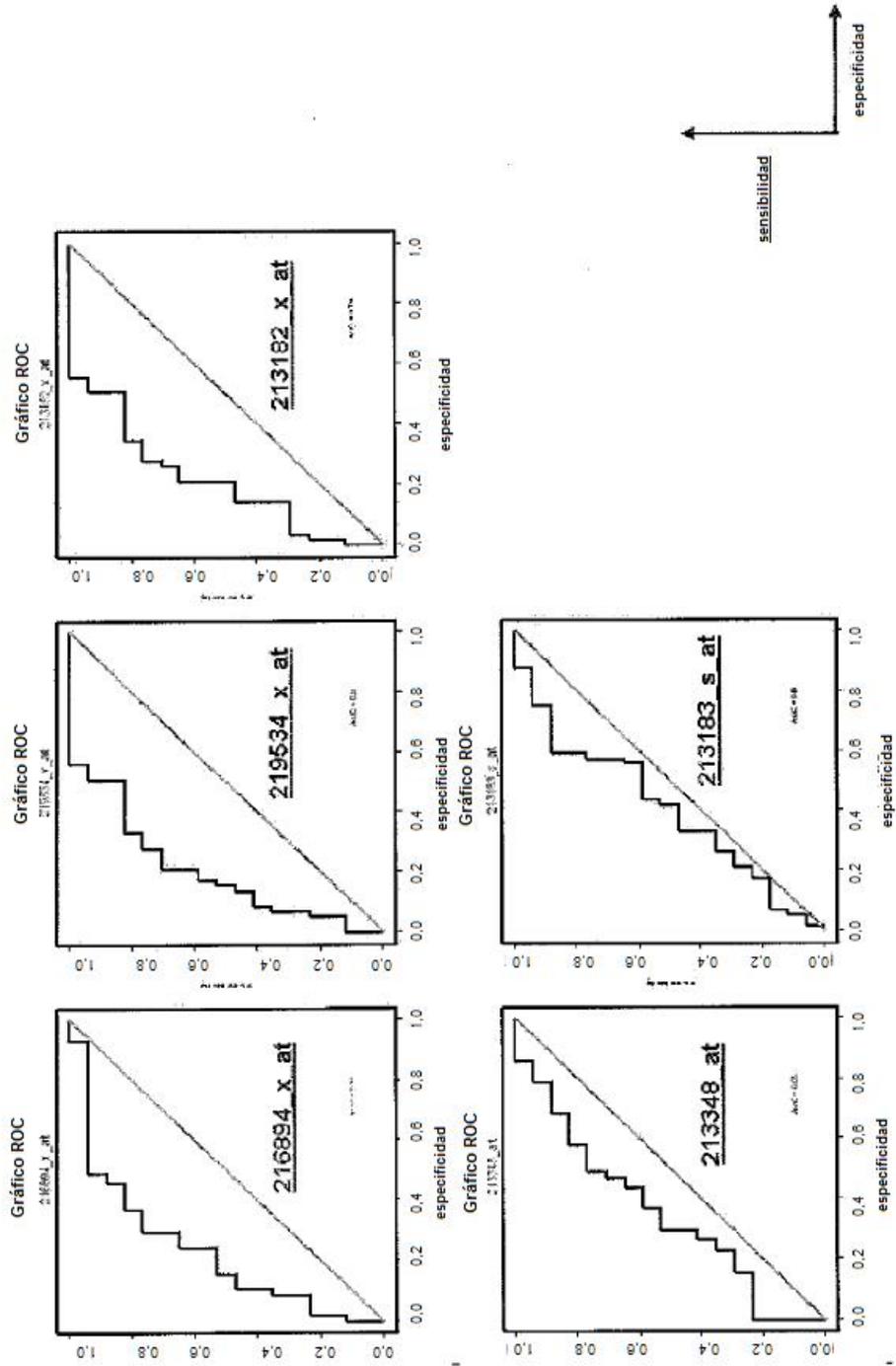


FIGURA 4A:

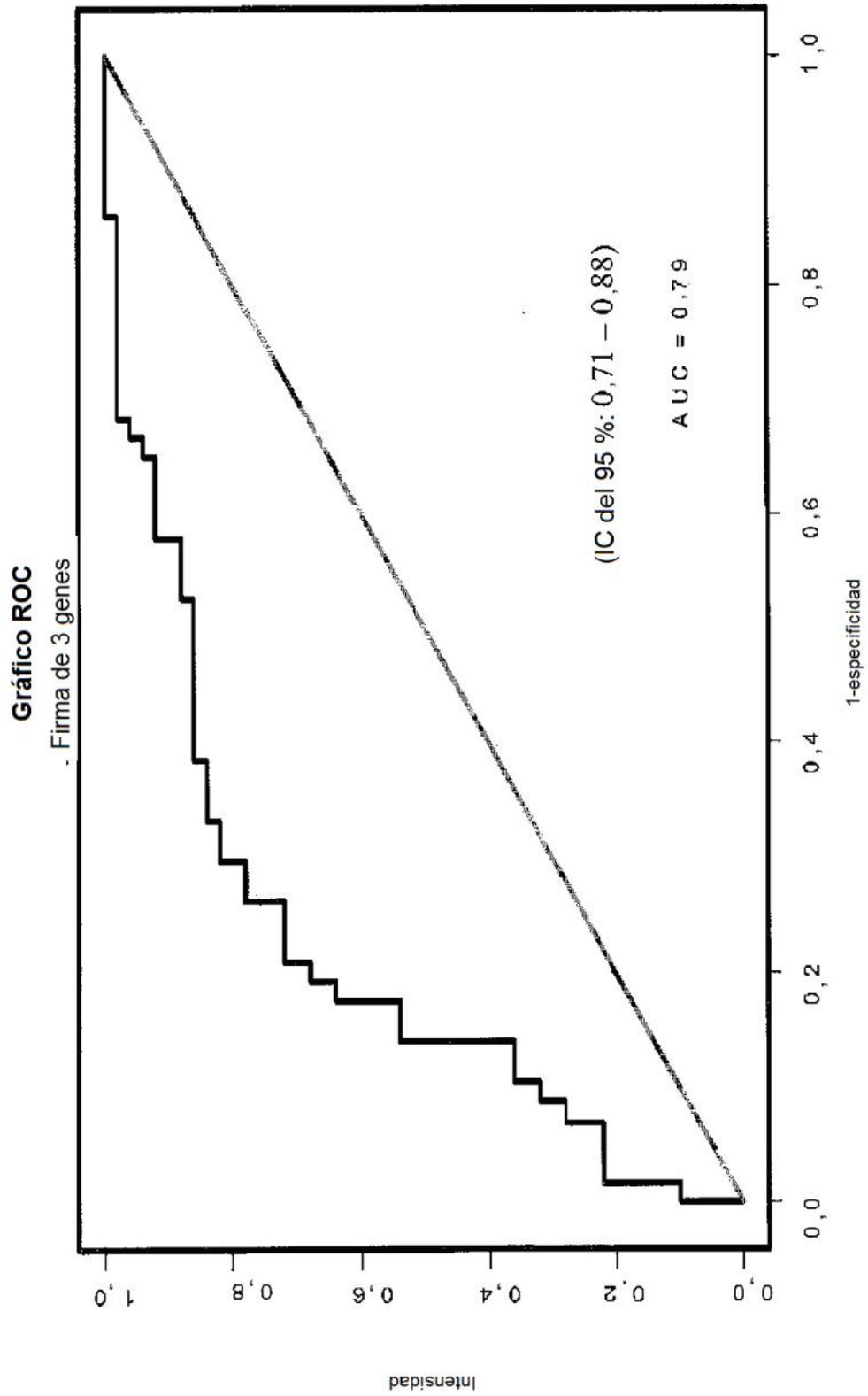


FIGURA 4B:

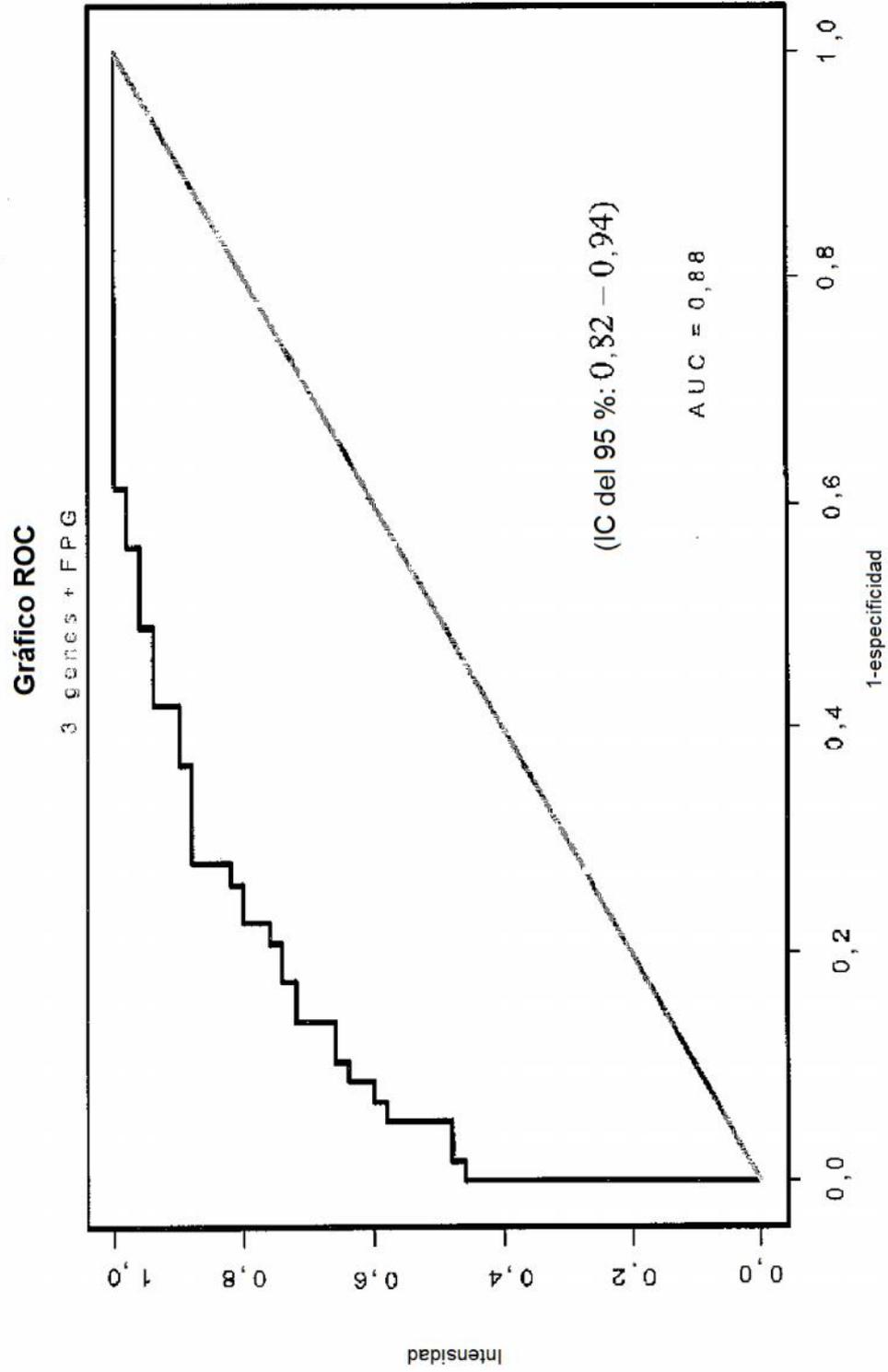


FIGURA 4C:

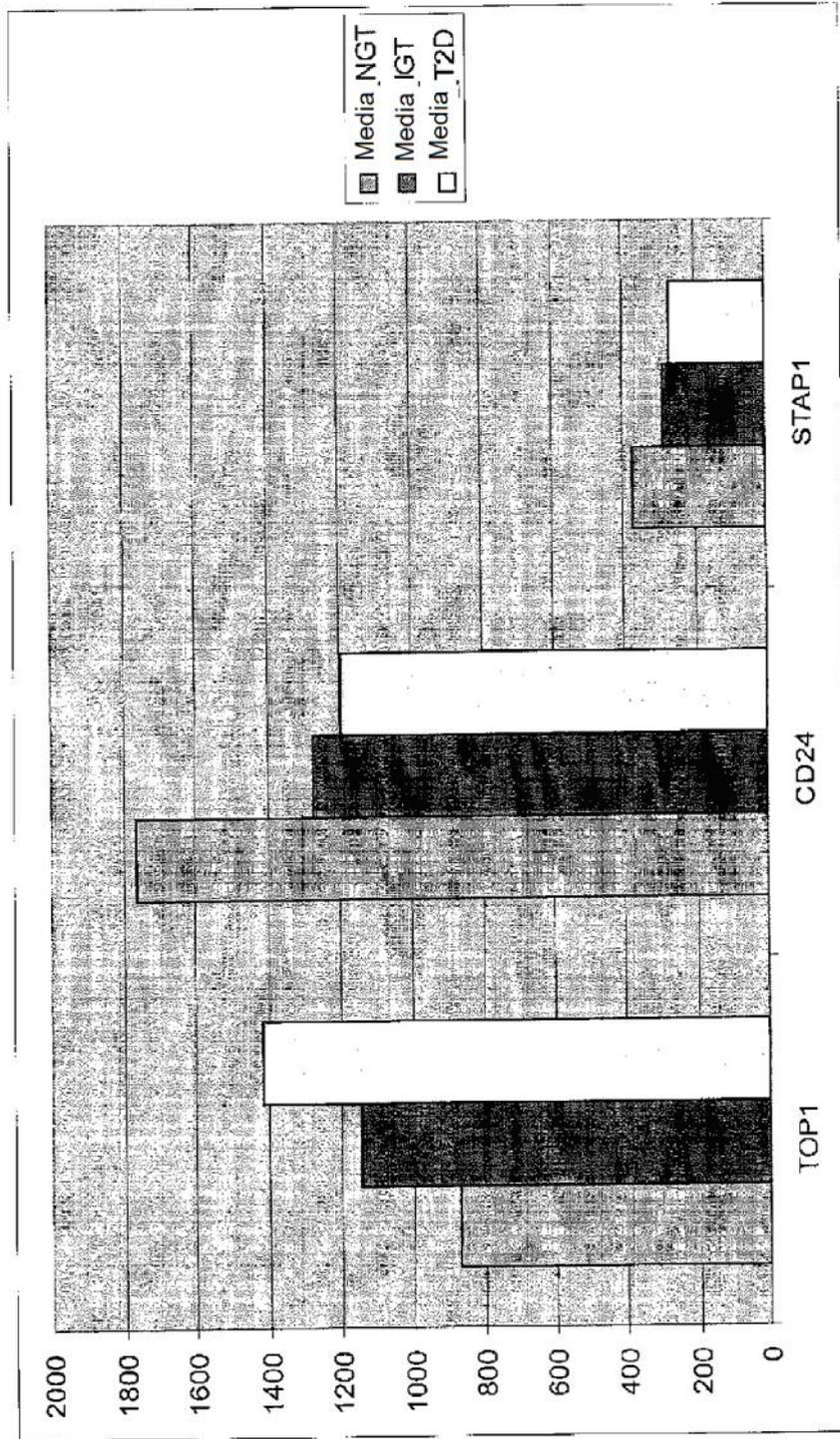
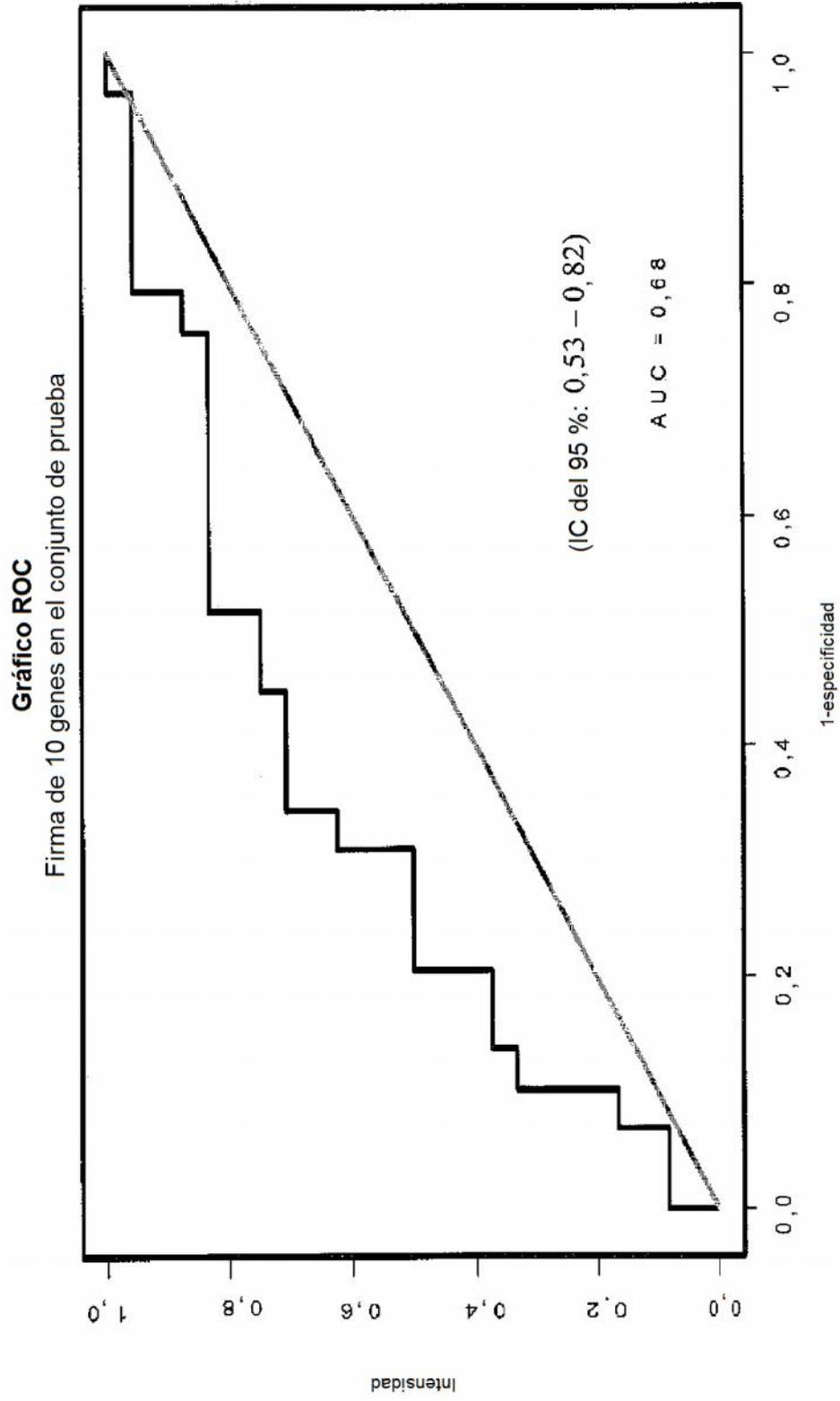


FIGURA 5A:



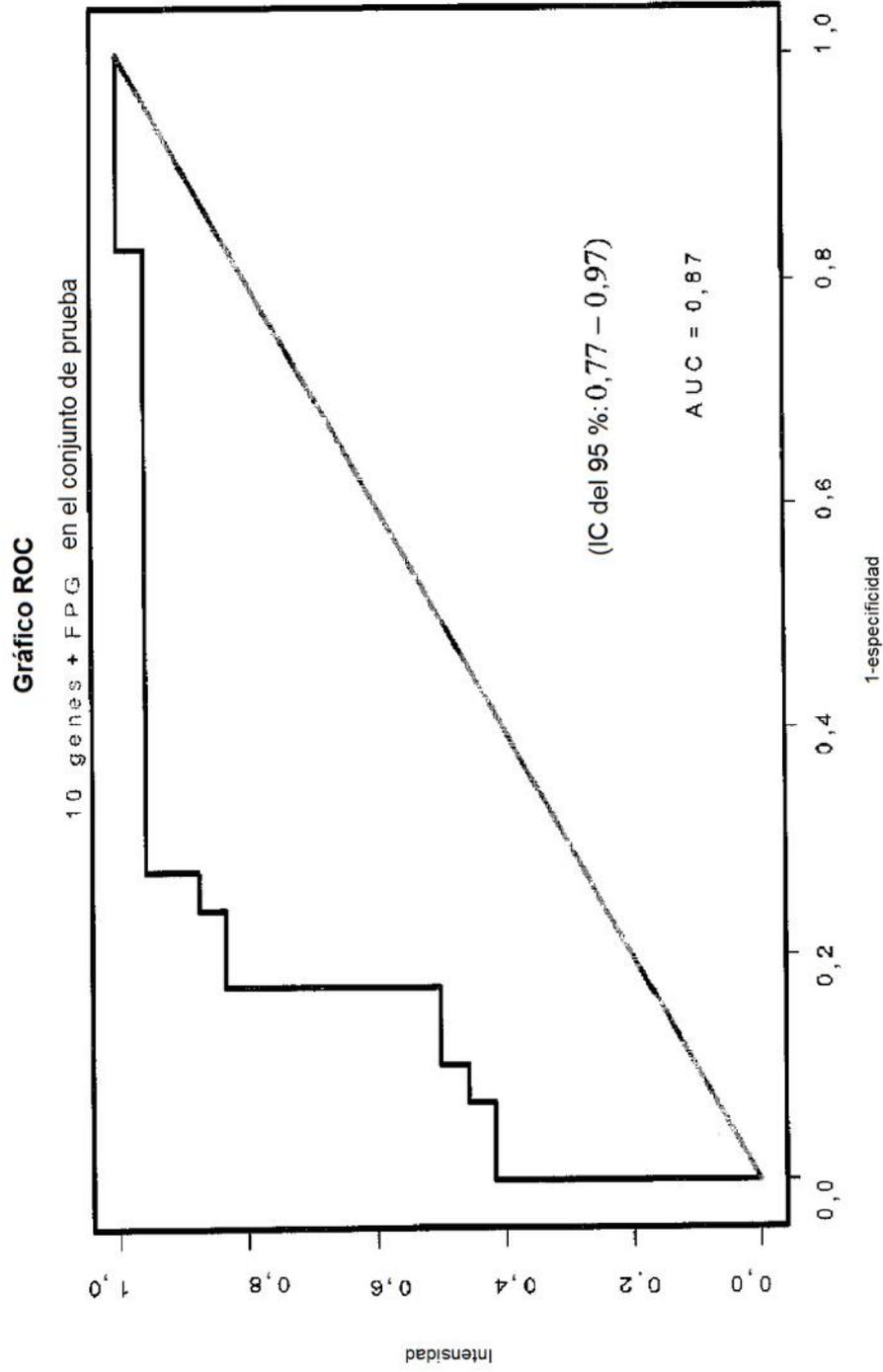


FIGURA 5B:

FIGURA 5C:

