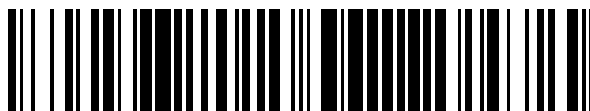


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 552 937**

51 Int. Cl.:

**G01N 33/574** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **05.06.2009 E 09759517 (7)**

97 Fecha y número de publicación de la concesión europea: **12.08.2015 EP 2294420**

54 Título: **Factor de predicción de supervivencia para linfoma difuso de células B grandes**

30 Prioridad:

**06.06.2008 US 59678 P**

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

**03.12.2015**

73 Titular/es:

**THE UNITED STATES OF AMERICA, AS REPRESENTED BY THE SECRETARY, DEPT. OF HEALTH AND HUMAN SERVICES (9.1%)  
Office Of Technology Transfer National Institutes of Health 6011 Executive Boulevard Suite 325, MSC 7660  
Bethesda, MD 20892-7660, US;  
BRITISH COLUMBIA CANCER AGENCY BRANCH (9.1%);  
UNIVERSITY OF ROCHESTER (9.1%);  
OREGON HEALTH & SCIENCE UNIVERSITY (9.1%);  
ARIZONA BOARD OF REGENTS ON BEHALF OF THE UNIVERSITY OF ARIZONA (9.1%);  
QUEEN MARY AND WESTFIELD COLLEGE, UNIVERSITY OF LONDON (9.1%);  
BOARD OF REGENTS OF THE UNIVERSITY OF NEBRASKA (9.1%);  
UNIVERSITAT DE BARCELONA HOSPITAL CLINIC (9.1%);  
OSLO UNIVERSITY HOSPITAL HF (9.1%);  
JULIUS-MAXIMILIANS-UNIVERSITY OF WUERZBURG (9.1%) y  
HOSPITAL CLINIC (9.1%)**

72 Inventor/es:

**STAUDT, LOUIS M.;  
RIMSZA, LISA;  
LISTER, ANDREW T.;  
WEISENBURGER, DENNIS;  
DELABIE, JAN;  
SMELAND, ERLAND B.;  
HOLTE, HARALD;  
KVALOY, STEIN;  
BRAZIEL, RITA M.;  
FISHER, RICHARD I.;  
JARES, PEDRO;  
LÓPEZ-GUILLERMO, ARMANDO;  
GUERRI, ELÍAS CAMPO;**

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).

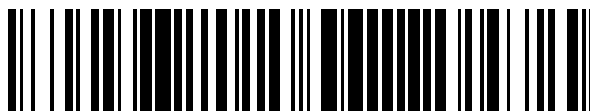
ES 2 552 937 T3

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 552 937**

96 Fecha de presentación y número de la solicitud europea: **05.06.2009 E 09759517 (7)**

97 Fecha y número de publicación de la concesión europea: **12.08.2015 EP 2294420**

72 Inventor/es:

**JAFFE, ELAINE S.;**  
**LENZ, GEORG;**  
**WYNDHAM, H.WILSON.;**  
**WRIGHT, GEORGE;**  
**DAVE, SANDEEP S.;**  
**GASCOYNE, RANDY D.;**  
**CONNORS, JOSEPH M.;**  
**MÜLLER-HERMELINK, HANS-KONRAD;**  
**ROSENWALD, ANDREAS;**  
**OTT, GERMAN y**  
**WING, C CHAN**

74 Agente/Representante:

**MILTENYI, Peter**

**ES 2 552 937 T3**

---

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

---

## DESCRIPCIÓN

Factor de predicción de supervivencia para linfoma difuso de células B grandes

## Antecedentes de la invención

- 5 El tratamiento de referencia actual para tratar el linfoma difuso de células B grandes (DLBCL) incluye regímenes quimioterápicos basados en antraciclina tales como CHOP en combinación con la administración del anticuerpo monoclonal anti-CD20, rituximab. Este régimen de combinación (R-CHOP) puede curar aproximadamente al 60% de los pacientes y ha mejorado la supervivencia global de los pacientes con DLBCL en un 10-15% (Coiffier *et al.*, N. Engl. J. Med., 346: 235-42 (2002)). No obstante, se desconoce la base molecular de la respuesta o resistencia a esta terapia.
- 10 El DLBCL es una enfermedad molecularmente heterogénea (Staudt *et al.*, Adv. Immunol., 87: 163-208 (2005)), y diferentes subtipos moleculares de DLBCL pueden tener pronósticos muy diferentes tras el tratamiento. Por ejemplo, la obtención de perfiles de expresión génica ha identificado dos subtipos moleculares de DLBCL que son biológica y clínicamente distintos (Rosenwald *et al.*, N. Engl. J. Med., 346: 1937-47 (2002); Alizadeh *et al.*, Nature, 403: 503-11 (2000)). El subtipo de DLBCL de tipo de células B de centro germinal (GCB) surge probablemente a partir de células B de centro germinal normales, mientras que el subtipo de DLBCL de tipo de células B activadas (ABC) puede surgir a partir de una célula B de centro posgerminal que se bloquea durante la diferenciación plasmacítica. Muchos mecanismos oncogénicos distinguen estos subtipos: Los DLBCL GCB tienen translocaciones recurrentes t(14,18), mientras que los DLBCL ABC tienen trisomía 3 recurrente y delección del locus INK4a/ARF así como activación constitutiva de la ruta de señalización anti-apoptótica de NFκB (Rosenwald *et al.*, N. Engl. J. Med., 346: 1937-47 (2002); Bea *et al.*, Blood, 106: 3183-90 (2005); Tagawa *et al.*, Blood, 106: 1770-77 (2005); Davis *et al.*, J. Exp. Med., 194:1861-74 (2001); Ngo *et al.*, Nature, 441: 106-10 (2006); Lenz *et al.*, Science, 319: 1676-79 (2008)). Cuando se trata con quimioterapia de tipo CHOP, las tasas de supervivencia global de los pacientes con DLBCL GCB y DLBCL ABC fueron del 60% y el 30%, respectivamente (Wright *et al.*, Proc. Nat'l. Acad. Sci. USA, 100: 9991-96 (2003)). Por tanto, el pronóstico para los diferentes DLBCL puede variar ampliamente.
- 15 Un enfoque analítico separado identificó cuatro firmas de expresión génica que reflejan distintos atributos tumorales de DLBCL y que se asociaron con distintos perfiles de supervivencia en pacientes con DLBCL tratados con CHOP (Rosenwald *et al.*, N. Engl. J. Med., 346: 1937-47 (2002)). Una firma de "células B de centro germinal" (GCB) se asoció con un pronóstico favorable y estaba relacionada con la distinción entre DLBCL GCB y ABC. La firma de "proliferación" se asoció con un pronóstico adverso e incluyó MYC y sus genes diana. La firma de "MHC de clase II" estaba silenciada en las células malignas en un subconjunto de casos de DLBCL, un acontecimiento que se asoció con supervivencia inferior (Rosenwald *et al.*, N. Engl. J. Med., 346: 1937-47 (2002); Rimsza *et al.*, Blood, 103: 4251-58 (2004)). Una cuarta firma de pronóstico, denominada firma de "ganglio linfático" se asoció con pronóstico favorable e incluyó componentes de la matriz extracelular, lo que sugiere que refleja la naturaleza de las células no malignas infiltrantes de tumor. Estas firmas predijeron la supervivencia de una forma estadísticamente independiente, lo que indica que múltiples variables biológicas dictan la respuesta a la quimioterapia con CHOP en DLBCL.
- 20 Los informes han sugerido que el beneficio de la inmunoterapia con rituximab podría estar limitado a determinados subtipos moleculares de DLBCL. La alta expresión de BCL-2 o la baja expresión de BCL-6 se asoció con supervivencia inferior con terapia con CHOP. Sin embargo, esta distinción desapareció con la terapia con R-CHOP (Mounier *et al.*, Blood, 101: 4279-84 (2003); Winter *et al.*, Blood, 107: 4207-13 (2006)). También se ha usado inmunohistoquímica para distinguir los DLBCL con un centro germinal frente a fenotipo de centro posgerminal. Aunque tales fenotipos inmunohistoquímicos fueron significativos desde el punto de vista del pronóstico en casos tratados con CHOP, no fueron de valor pronóstico para casos tratados con R-CHOP (Nyman *et al.*, Blood, 109: 4930-35 (2007)).
- 25 El documento WO2005/024043 da a conocer un método para predecir la supervivencia en un sujeto con linfoma difuso de células B grandes (DLBCL), que comprende las etapas de:
- a). obtener una biopsia de dicho sujeto, y aislar ARN de las muestras;
  - b). obtener datos de expresión génica para un conjunto de genes en dicha muestra de biopsia. El nivel de expresión génica de los genes de un centro germinal de genes de células B comprende una firma de expresión génica de GCB, el nivel de expresión génica de genes de un ganglio linfático comprende una firma de expresión génica de ganglio linfático, el nivel de expresión génica de genes de proliferación comprende una firma de expresión génica de proliferación y el nivel de expresión génica de genes de actividad de MHC de clase II comprende una firma de expresión génica de MHC de clase II;
  - c). obtener el promedio del nivel de expresión génica de los genes de dichas firmas para obtener dichos valores de firma de expresión génica respectivos; y
  - d). calcular una puntuación de factor de predicción de supervivencia usando una ecuación, en el que una puntuación superior de factor de predicción de supervivencia está asociada con peor supervivencia.

El documento WO2005/024043 no menciona casos tratados con R-CHOP.

Por consiguiente, existe la necesidad de nuevos métodos para distinguir entre subtipos de DLBCL que sean significativos desde el punto de vista del pronóstico para pacientes tratados con R-CHOP.

### Breve resumen de la invención

5 La invención proporciona métodos y matrices relacionadas con un factor de predicción de supervivencia basado en la expresión génica para pacientes con DLBCL, incluyendo pacientes tratados con el tratamiento de referencia actual, que incluye quimioterapia y la administración de rituximab.

La invención proporciona un método de predicción del resultado de supervivencia basándose en una o más muestras de linfoma difuso de células B grandes (DLBCL). El método incluye aislar productos de expresión génica de una o más muestras de biopsia de DLBCL de un sujeto tratado con un régimen terapéutico que incluye quimioterapia y rituximab (R-CHOP). El método comprende además obtener un perfil de expresión génica a partir de los productos de expresión génica, en el que el perfil de expresión comprende un nivel de expresión para cada uno de al menos 31 de los genes enumerados en la tabla 1 para la firma de expresión génica de células B de centro germinal (GCB) y al menos 226 de los genes enumerados en la tabla 1 para la firma de expresión génica de estromal-1. El método también incluye determinar un valor de firma de GCB y un valor de firma de estromal-1 a partir del perfil de expresión génica, en el que el valor de firma de GCB es un promedio de los niveles de expresión de los al menos 31 genes enumerados en la tabla 1 para la firma de expresión génica de GCB y el valor de firma de estromal-1 es un promedio de los niveles de expresión de los al menos 226 genes enumerados en la tabla 1 para la firma de expresión génica de estromal-1. Se calcula una puntuación de factor de predicción de supervivencia que es igual a  $A - [(x) \cdot (\text{el valor de firma de GCB})] - [(y) \cdot (\text{el valor de firma de estromal-1})]$ , en el que A es un término de desviación, (x) es un factor de escala de entre 0,200 y 0,625, e (y) es un factor de escala de entre 0,800 y 1,250. El método incluye además determinar el resultado de supervivencia para el sujeto basándose en la puntuación de factor de predicción de supervivencia calculada para cualquier valor igual fijo de A y para cualquier valor igual fijo de (x) e (y), en el que una puntuación inferior de factor de predicción de supervivencia indica un resultado de supervivencia más favorable y una puntuación superior de factor de predicción de supervivencia indica un resultado de supervivencia menos favorable para el sujeto.

La invención también proporciona un método de predicción del resultado de supervivencia basándose en una o más muestras de linfoma difuso de células B grandes (DLBCL). El método comprende aislar productos de expresión génica de una o más muestras de biopsia de DLBCL de un sujeto tratado con un régimen terapéutico que incluye quimioterapia y rituximab (R-CHOP). El método incluye además obtener un perfil de expresión génica a partir de los productos de expresión génica, en el que el perfil de expresión génica comprende un nivel de expresión para cada uno de al menos 31 de los genes enumerados en la tabla 1 para la firma de expresión génica de células B de centro germinal (GCB), al menos 226 de los genes enumerados en la tabla 1 para la firma de expresión génica de estromal-1 y al menos 58 de los genes enumerados en la tabla 1 para la firma de expresión génica de estromal-2. El método también comprende determinar un valor de firma de GCB, un valor de firma de estromal-1 y un valor de firma de estromal-2 a partir del perfil de expresión génica, en el que el valor de firma de GCB es un promedio de los niveles de expresión de los al menos 31 genes enumerados en la tabla 1 para la firma de expresión génica de GCB y el valor de firma de estromal-1 es un promedio de los niveles de expresión de los al menos 226 genes enumerados en la tabla 1 para la firma de expresión génica de estromal-1 y el valor de firma de estromal-2 es un promedio de los niveles de expresión de los al menos 58 genes enumerados en la tabla 1 para la firma de expresión génica de estromal-2. Se calcula una puntuación de factor de predicción de supervivencia usando la ecuación: puntuación de factor de predicción de supervivencia =  $A - [(x) \cdot (\text{el valor de firma de GCB})] - [(y) \cdot (\text{el valor de firma de estromal-1})] + [(z) \cdot (\text{el valor de firma de estromal-2})]$ , en la que A es un término de desviación y (x) es un factor de escala de entre 0,200 y 0,625, (y) es un factor de escala de entre 0,800 y 1,250 y (z) es un factor de escala de entre 0,450 y 0,900. El método comprende además calcular la probabilidad de un resultado de supervivencia para el sujeto más allá de una cantidad de tiempo t tras el tratamiento para DLBCL, en el que la probabilidad del resultado de supervivencia del sujeto  $P(SO)$  se calcula usando la ecuación:  $P(SO) = SO_0(t)^{\exp(s) \cdot \text{puntuación de factor de predicción de supervivencia}}$ , en la que  $SO_0(t)$  es la probabilidad del resultado de supervivencia, que corresponde al valor de tiempo más largo menor que t en una curva de resultado de supervivencia, y en la que (s) es un factor de escala de entre 0,970 y 0,980.

50 La invención proporciona además una matriz seleccionado como diana que comprende al menos una sonda o conjunto de sondas para cada gen en (a) una firma de expresión génica de células B de centro germinal (GCB) que comprende al menos 31 de los genes de firma de GCB enumerados en la tabla 1, (b) una firma de expresión génica de estromal-1 que comprende al menos 226 de los genes de firma de estromal-1 enumerados en la tabla 1 y (c) una firma de expresión génica de estromal-2 que comprende al menos 58 de los genes de firma de estromal-2 enumerados en la tabla 1, en la que la matriz comprende sondas para menos de 10.000 genes.

### Breve descripción de los dibujos

La figura 1A es una gráfica de estimación de Kaplan-Meier que representa la probabilidad de supervivencia libre de progresión frente al tiempo (en años) de pacientes con DLBCL GCB y DLBCL ABC. La gráfica indica que los pacientes con GCB tienen una probabilidad más favorable, es decir, superior de tasa de supervivencia libre de

progresión que los pacientes con ABC durante al menos cinco años tras la terapia con R-CHOP.

5 La figura 1B es una gráfica de estimación de Kaplan-Meier que representa la probabilidad de supervivencia global frente al tiempo (en años) de pacientes con DLBCL GCB y DLBCL ABC. La gráfica indica que los pacientes con GCB tienen una probabilidad más favorable, es decir, superior de supervivencia global que los pacientes con ABC durante al menos cinco años tras la terapia con R-CHOP.

10 La figura 1C es una serie de cuatro gráficas de estimación de Kaplan-Meier que representan las probabilidades de supervivencia global frente al tiempo (en años) en pacientes con DLBCL. Cada una de las cuatro gráficas correlaciona la probabilidad de supervivencia global con la firma de expresión génica de ganglio linfático/estromal-1, células B de centro germinal, proliferación o MHC de clase II, respectivamente. Además, en cada gráfica, se usó la expresión promedio de los genes de firma en cada muestra de biopsia para clasificar los casos y dividir la cohorte en grupos de cuartil tal como se indica.

15 La figura 2A es un par de gráficas de estimación de Kaplan-Meier que representan la probabilidad de supervivencia libre de progresión y la probabilidad de supervivencia global, tal como se indica, frente al tiempo (en años) entre pacientes con DLBCL tratados con R-CHOP. Las muestras de pacientes se clasificaron según un modelo de dos variables creado usando las firmas de células B de centro germinal (GCB) y estromal-1 y se dividieron entre los grupos de cuartil.

20 La figura 2B es un par de gráficas de Kaplan-Meier que representan la probabilidad de supervivencia libre de progresión y la probabilidad de supervivencia global, tal como se indica, frente al tiempo (en años) entre pacientes con DLBCL tratados con R-CHOP. Las muestras de pacientes se clasificaron según una puntuación de factor de predicción de supervivencia derivada de un modelo que incorpora las firmas de células B de centro germinal, estromal-1 y estromal-2 y se dividieron entre los grupos de cuartil.

25 La figura 2C es una serie de tres gráficas de estimación de Kaplan-Meier que representan la probabilidad de supervivencia global frente al tiempo (en años) entre pacientes con DLBCL tratados con R-CHOP en los grupos de riesgo de IPI bajo, intermedio o alto indicados. Se estratificaron las muestras de pacientes según la misma puntuación de factor de predicción de supervivencia usada en la figura 2B, excepto porque se fusionaron los cuartiles primero y segundo, y se fusionaron los cuartiles tercero y cuarto.

30 La figura 3 representa los niveles de expresión de los genes de firma de células GCB, estromal-1 y estromal-2 indicados en muestras de biopsia de DLBCL ABC, GCB y sin clasificar. Se representan los niveles de expresión génica relativa según la escala mostrada. En la parte inferior se muestran los promedios de firma para cada paciente. También se muestra la puntuación estromal, que es el componente del modelo de supervivencia al que contribuye la diferencia entre los promedios de firma de estromal-2 y estromal-1. Se muestra la puntuación de factor de predicción de supervivencia para cada paciente y se usó para ordenar las clases, tras agrupar en las categorías de DLBCL ABC, DLBCL GCB y sin clasificar.

35 La figura 4A representa la expresión génica relativa de firmas de estromal-1, estromal-2 y células B de centro germinal en subpoblaciones malignas CD19+ y no malignas CD19- de células aisladas de tres muestras de biopsia de pacientes con DLBCL. Los genes de firma estromal-1 y estromal-2 se expresaban más en las células no malignas, mientras que los genes de firma de células B de centro germinal se expresaban más en las células malignas. Se representan las razones de log2 de niveles de expresión génica en la subpoblación CD19- con respecto a los de las subpoblaciones CD19+ según la escala mostrada.

40 La figura 4B representa los resultados del análisis de enriquecimiento génico que compara la firma génica de estromal-1 con las firmas de mesénquima-1 y mesénquima-2 (de células de origen mesenquimal normales), con una firma de monocitos expresada más en monocitos de sangre normales que en células B, T y NK de la sangre, y en una firma de células pan-T expresadas más en células T de la sangre que en células B, células NK y monocitos de la sangre. Aunque se observó una relación entre la firma de estromal-1 y las firmas de mesénquima-1, mesénquima-2 de monocitos, no se observó ninguna relación entre la firma de estromal-1 y una firma de células pan-T expresada más en células T de la sangre que en células B, células NK y monocitos de la sangre. Los niveles de expresión génica relativa se expresan según la escala mostrada.

50 La figura 5A es una gráfica de estimación de Kaplan-Meier que representa la probabilidad de supervivencia global frente al tiempo (en años) en casos de DLBCL segregados según los niveles de expresión de proteína SPARC, tal como se indica.

La figura 5B es un par de imágenes que muestran la identificación de vasos sanguíneos tumorales mediante análisis de inmunohistoquímica de células endoteliales CD34+ en biopsias de DLBCL representativas que tienen densidad de vasos sanguíneos baja o alta (objetos CD34+/ $\mu\text{M}^2$ ), tal como se indica.

55 La figura 5C es una gráfica que representa la correlación entre la densidad de vasos sanguíneos tumorales y la puntuación estromal en biopsias de DLBCL analizadas.

La figura 6A es una gráfica de estimación de Kaplan-Meier que representa la probabilidad de supervivencia global

frente al tiempo (en años) para pacientes con DLBCL tratados con “LLMPP CHOP” tras la terapia. La gráfica indica que en esta cohorte los pacientes con DLBCL GCB muestran una supervivencia global significativamente superior en comparación con pacientes con DLBCL ABC tras la terapia con CHOP.

5 La figura 6B es una gráfica de estimación de Kaplan-Meier que representa la probabilidad de supervivencia global frente al tiempo (en años) para pacientes con DLBCL tratados con “MMMLNP CHOP” tras la terapia. En esta cohorte, los pacientes con DLBCL GCB muestran una supervivencia global significativamente superior en comparación con pacientes con DLBCL ABC tras la terapia con CHOP.

10 La figura 7 es un conjunto de cuatro gráficas de estimación de Kaplan-Meier que representan la probabilidad de supervivencia global frente al tiempo (en años) en una cohorte de “MMMLNP CHOP”. Cada uno de las cuatro gráficas correlaciona la probabilidad de supervivencia global con la firma de expresión génica de ganglio linfático/estromal-1, células B de centro germinal, proliferación o MHC de clase II, respectivamente. Además, en cada gráfica, se usó la expresión promedio de los genes de firma en cada muestra de biopsia para clasificar los casos y dividir la cohorte en grupos de cuartil tal como se indica.

15 La figura 8A es una gráfica de estimación de Kaplan-Meier que representa la probabilidad de supervivencia global frente al tiempo (en años) en una cohorte de “LLMPP CHOP”, que se dividió según los niveles de expresión de firma de MHC de clase II. Los pacientes con expresión baja de firma de MHC de clase II tienen una supervivencia global significativamente inferior en comparación con pacientes con expresión normal de MHC de clase II.

20 La figura 8B es una gráfica de estimación de Kaplan-Meier que representa la probabilidad de supervivencia global frente al tiempo (en años) en una cohorte de “MMMLNP CHOP”, que se dividió según los niveles de expresión de firma de MHC de clase II. Los pacientes con expresión baja de firma de MHC de clase II tienen supervivencia global significativamente inferior en comparación con pacientes con expresión normal de MHC de clase II.

25 La figura 8C es una gráfica de estimación de Kaplan-Meier que representa la probabilidad de supervivencia global frente al tiempo (en años) en una cohorte de “LLMPP R-CHOP”, que se dividió según los niveles de expresión de firma de MHC de clase II. No hubo diferencia significativa en la supervivencia global de pacientes con expresión baja de firma de MHC de clase II en comparación con pacientes con expresión normal de MHC de clase II.

La figura 9A es un par de gráficas de estimación de Kaplan-Meier que representan las probabilidades de supervivencia libre de progresión o supervivencia global, tal como se indica, frente al tiempo (en años) entre pacientes agrupados en cuartiles según un modelo de expresión génica que consiste en firma de estromal-1, firma de GCB y firma 122 tras la terapia con R-CHOP.

30 La figura 9B es un par de gráficas de estimación de Kaplan-Meier que representan las probabilidades de supervivencia global frente al tiempo (en años) entre pacientes de la cohorte de “MMMLNP CHOP” agrupados en cuartiles según un modelo de expresión génica que consiste en o bien firma de estromal-1 y firma de GCB o bien firma de estromal-1, GCB y firma de 122, tal como se indica, tras la terapia con CHOP.

35 La figura 9C es una gráfica de estimación de Kaplan-Meier que representa las probabilidades de supervivencia global frente al tiempo (en años) entre pacientes de la cohorte de “MMMLNP CHOP” agrupados en cuartiles según un modelo de expresión génica que consiste en firma de estromal-1, firma de GCB y firma de estromal-2 tras la terapia con CHOP.

40 La figura 10A es una gráfica de estimación de Kaplan-Meier que representa la supervivencia global entre pacientes de grupo de riesgo de índice de pronóstico internacional (IPI) revisado bajo estratificados según la puntuación de factor de predicción de resultado basado en la expresión génica. Tras agrupar a los pacientes en cuartiles según la puntuación de factor de predicción de resultado basado en la expresión génica, se fusionaron los cuartiles 1 y 2 (puntuación de modelo baja), y se fusionaron los cuartiles 3 y 4 (puntuación de modelo alta).

45 La figura 10B es una gráfica de estimación de Kaplan-Meier que representa la supervivencia global entre pacientes de grupo de riesgo de índice de pronóstico internacional (IPI) revisado intermedio estratificados según el factor de predicción de resultado basado en la expresión génica. Tras agrupar a los pacientes en cuartiles según la puntuación de factor de predicción de resultado basado en la expresión génica, se fusionaron los cuartiles 1 y 2 (puntuación de modelo baja), y se fusionaron los cuartiles 3 y 4 (puntuación de modelo alta).

50 La figura 10C es una gráfica de estimación de Kaplan-Meier que representa la supervivencia global entre pacientes de grupo de riesgo de índice de pronóstico internacional (IPI) revisado alto estratificados según el factor de predicción de resultado basado en la expresión génica. Tras agrupar a los pacientes en cuartiles según la puntuación de factor de predicción de resultado basado en la expresión génica, se fusionaron los cuartiles 1 y 2 (puntuación de modelo baja), y se fusionaron los cuartiles 3 y 4 (puntuación de modelo alta).

La figura 11 representa la expresión génica de firma mesenquimal-1 normal y firma mesenquimal-2 normal en diversos tejidos normales.

55 **Descripción detallada de la invención**

La invención proporciona un factor de predicción de supervivencia basado en la expresión génica para pacientes con DLBCL, incluyendo aquellos pacientes que reciben el tratamiento de referencia actual, R-CHOP. El factor de predicción de supervivencia puede usarse para determinar la probabilidad relativa de un resultado de supervivencia en un sujeto específico. El factor de predicción de supervivencia también puede usarse para predecir; es decir, determinar la probabilidad esperada de que se produzca un resultado de supervivencia en un periodo definido tras el tratamiento para DLBCL. Tal información de pronóstico puede ser muy útil tanto para el paciente como para el médico. Los pacientes con puntuaciones de factor de predicción de supervivencia que indican un resultado inferior con terapia con R-CHOP podrían ser candidatos a un régimen terapéutico diferente si, por ejemplo, experimentan recidiva a partir del tratamiento con R-CHOP. El factor de predicción de supervivencia también puede usarse en el diseño de estudios clínicos y análisis de datos clínicos para proporcionar un estudio cuantitativo de los tipos de pacientes con DLBCL de los que se reunieron datos clínicos. El factor de predicción puede usarse para mejorar una o más comparaciones entre datos de fuentes diferentes (por ejemplo, de ensayos clínicos diferentes), permitiendo comparaciones con respecto a características de pacientes, que se manifiestan en los niveles de expresión génica que determinan y, por tanto, se incorporan en el factor de predicción. Además, la invención proporciona información que puede ser muy valiosa para un paciente con DLBCL, puesto que el paciente puede verse inclinado a organizar su vida de manera muy diferente, dependiendo de si el paciente tiene una probabilidad de supervivencia alta o baja y/o permanece libre de progresión durante un periodo de tiempo tras el tratamiento.

En el presente documento se usan las siguientes abreviaturas: ABC, linfoma difuso de células B grandes de tipo de células B activadas; CHOP, ciclofosfamida, doxorubicina, vincristina y prednisona; IC, intervalo de confianza; COP, ciclofosfamida, vincristina y prednisona; DLBCL, linfoma difuso de células B grandes; DOD, muerto por enfermedad; ECOG, Eastern Cooperative Oncology Group (grupo oncológico cooperativo del este); FACS, clasificación celular activada por fluorescencia; FH, hiperplasia folicular; FISH, hibridación *in situ* por fluorescencia; FL, linfoma folicular; GC, centro germinal; GCB, linfoma difuso de células B grandes de tipo de células B de centro germinal; IPI, índice de pronóstico internacional; LPC, linfoma linfoplasmácítico; MHC, complejo mayor de histocompatibilidad; NA, no disponible o no aplicable; NK, linfocito citolítico natural; PCR, reacción en cadena de la polimerasa; RQ-PCR, PCR cuantitativa en tiempo real; RT-PCR, reacción en cadena de la polimerasa con transcriptasa inversa; SAGE, análisis en serie de la expresión génica; WHO, Organización Mundial de la Salud.

El término "R-CHOP" tal como se usa en el presente documento se refiere generalmente a cualquier régimen terapéutico que incluye quimioterapia y la administración de rituximab. Por consiguiente, aunque el término puede referirse a la terapia de combinación con rituximab que incluye un régimen de CHOP de ciclofosfamida, doxorubicina, vincristina y prednisona, el término R-CHOP también puede referirse a terapia que incluye rituximab en combinación con un régimen quimioterápico distinto de CHOP.

El término "datos de expresión génica" así como "nivel de expresión génica" tal como se usan en el presente documento se refiere a información en relación con el nivel de expresión relativo o absoluto de un gen o conjunto de genes en una célula o grupo de células. El nivel de expresión de un gen puede determinarse basándose en el nivel de ARN, tal como ARNm, codificado por el gen. Alternativamente, el nivel de expresión puede determinarse basándose en el nivel de un polipéptido o fragmento del mismo codificado por el gen. Los datos de expresión génica pueden adquirirse para una célula individual, o para un grupo de células tal como una muestra de tumor o biopsia. Los datos de expresión génica y los niveles de expresión génica pueden almacenarse en medios legibles por ordenador, por ejemplo, el medio legible por ordenador usado conjuntamente con un dispositivo de lectura de chip o micromatriz. Tales datos de expresión génica pueden manipularse para generar firmas de expresión génica.

El término "micromatriz" "matriz," o "chip" se refiere a una pluralidad de sondas de ácido nucleico acopladas a la superficie de un sustrato en ubicaciones conocidas diferentes. El sustrato es preferiblemente sólido. Se han descrito micromatrices de manera general en la técnica en, por ejemplo, las patentes estadounidenses n.ºs 5.143.854 (Pirrung), 5.424.186 (Fodor), 5.445.934 (Fodor), 5.677.195 (Winkler), 5.744.305 (Fodor), 5.800.992 (Fodor) y 6.040.193 (Winkler) y Fodor *et al.*, Science, 251:767-777 (1991).

El término "firma de expresión génica" o "firma" tal como se usa en el presente documento se refiere a un grupo de genes que se expresan de manera coordinada. Los genes que constituyen esta firma pueden expresarse en un linaje celular, fase de diferenciación específicos, o durante una respuesta biológica particular. Los genes pueden reflejar aspectos biológicos de los tumores en los que se expresan, tales como la célula de origen del cáncer, la naturaleza de las células no malignas en la biopsia y los mecanismos oncogénicos responsables del cáncer (Shaffer *et al.*, Immunity, 15: 375-385 (2001)). Los ejemplos de firmas de expresión génica incluyen firmas de ganglio linfático, proliferación (Rosenwald *et al.*, New Engl. J. Med., 346: 1937-1947 (2002)), MHC de clase II, DLBCL ABC alto, diferenciación de células B, células T, macrófagos, respuesta inmunitaria-1 y respuesta inmunitaria-2 (publicación de solicitud de patente estadounidense n.º 2007/0105136 (Staudt)).

El término "valor de firma" tal como se usa en el presente documento corresponde a una combinación matemática de mediciones de niveles de expresión de los genes en una firma de expresión génica. Un valor de firma a modo de ejemplo es un promedio de firma que corresponde al promedio o media de los niveles de expresión individuales en una firma de expresión génica.

El término "puntuación de factor de predicción de supervivencia" tal como se usa en el presente documento se

refiere a una puntuación generada por un modelo de múltiples variables usado para predecir la supervivencia basándose en la expresión génica. Se prevé que un sujeto con una puntuación superior de factor de predicción de supervivencia tendrá una supervivencia inferior que un sujeto con una puntuación inferior de factor de predicción de supervivencia.

- 5 El término “supervivencia” o “supervivencia global” tal como se usa en el presente documento puede referirse a la probabilidad o posibilidad de que un sujeto sobreviva durante un periodo de tiempo particular. Alternativamente, puede referirse al periodo probable de supervivencia para un sujeto, tal como la media o la mediana esperada del tiempo de supervivencia para un sujeto con un patrón de expresión génica particular.

- 10 El término “supervivencia libre de progresión” tal como se usa en el presente documento puede referirse a la probabilidad o posibilidad de que un sujeto sobreviva sin progresión o empeoramiento significativo de la enfermedad durante un periodo de tiempo particular. Alternativamente, puede referirse al periodo probable de supervivencia para un sujeto sin progresión o empeoramiento significativo de la enfermedad, tal como la media o la mediana esperada del tiempo de supervivencia para un sujeto con un patrón de expresión génica particular sin progresión o empeoramiento significativo de la enfermedad.

- 15 El término “resultado de supervivencia” tal como se usa en el presente documento puede referirse a supervivencia, supervivencia global o supervivencia libre de progresión.

El término “factor de escala” tal como se usa en el presente documento se refiere a un factor que relaciona el cambio en la expresión génica con el pronóstico. Un ejemplo de un factor de escala es un factor obtenido maximizando las probabilidades parciales del modelo de riesgos proporcionales de Cox.

- 20 Las firmas de expresión génica, los valores de firma, las puntuaciones de factor de predicción de supervivencia, las puntuaciones estromales, las curvas de estimación de supervivencia y las probabilidades de supervivencia dadas a conocer en el presente documento pueden almacenarse en formato codificado digitalmente en medios legibles por ordenador, por ejemplo, medios legibles por ordenador usados conjuntamente con dispositivos de lectura de chip o micromatriz o medios legibles por ordenador usados para almacenar datos de pacientes durante el tratamiento para DLBCL. Tales medios y dispositivos especializados que los usan, por ejemplo, para aplicaciones de diagnóstico y clínicas, se conocen en la técnica.

- La invención proporciona un método para predecir un resultado de supervivencia en un sujeto al que se le diagnostica DLBCL usando datos de expresión génica. Tales datos pueden reunirse usando cualquier método eficaz de cuantificación de la expresión génica. Por ejemplo, los datos de expresión génica pueden medirse o estimarse usando una o más micromatrices. Las micromatrices pueden ser de cualquier tipo eficaz, incluyendo, pero sin limitarse a, a base de ácidos nucleicos o a base de anticuerpos. La expresión génica también puede medirse mediante una variedad de otras técnicas, incluyendo, pero sin limitarse a, PCR, RT-PCR cuantitativa, PCR en tiempo real, amplificación de ARN, hibridación *in situ*, inmunohistoquímica, inmunocitoquímica, FACS, análisis en serie de la expresión génica (SAGE) (Velculescu *et al.*, *Science*, 270: 484-87 (1995)), hibridación por transferencia de tipo Northern o hibridación por inmunotransferencia de tipo Western.

- Las micromatrices de ácidos nucleicos comprenden generalmente sondas de ácidos nucleicos derivadas de genes individuales y colocadas en una matriz ordenada o sobre un soporte. Este soporte puede ser, por ejemplo, un portaobjetos de vidrio, una membrana de nailon o una oblea de silicio. Los patrones de expresión génica en una muestra se obtienen hibridando la micromatriz con el producto de expresión génica de la muestra. Este producto de expresión génica puede ser, por ejemplo, ARNm celular total, ARNr o ADNc obtenido mediante transcripción inversa del ARN celular total. El producto de expresión génica de una muestra se marca con un marcador radiactivo, fluorescente u otro marcador para permitir la detección. Tras la hibridación, se lava la micromatriz y se detecta la hibridación del producto de expresión génica para cada sonda de ácido nucleico en la micromatriz y se cuantifica usando un dispositivo de detección tal como un sistema de detección y cuantificación de la radiactividad o microscopio confocal de barrido.

- Hay dos clases amplias de micromatrices: alineamientos de ADNc y de oligonucleótidos. Los alineamientos de ADNc consisten en cientos o miles de sondas de ADNc inmovilizadas sobre un soporte sólido. Estas sondas de ADNc tienen habitualmente un tamaño de 100 nucleótidos o más. Hay dos diseños usados comúnmente para alineamientos de ADNc. El primero es el alineamiento de filtro de nitrocelulosa, que se prepara generalmente mediante disposición en manchas robótica de fragmentos de ADN purificados o lisados de bacterias que contienen clones de ADNc sobre un filtro de nitrocelulosa (Southern *et al.*, *Genomics*, 13: 1008-17 (1992); Southern *et al.*, *Nucl Acids Res* 22: 1368-73 (1994); Gress *et al.*, *Oncogene*, 13: 1819-30 (1996); Pietu *et al.*, *Genome Res.*, 6: 492-503 (1996)). Los otros alineamientos de ADNc usados comúnmente se fabrican mediante disposición en manchas robótica de fragmentos de PCR de clones de ADNc sobre portaobjetos de vidrio de microscopio (Skena *et al.*, *Science*, 270: 467-70 (1995); DeRisi *et al.*, *Nature Genet.*, 14: 457-60 (1996); Skena *et al.*, *Proc. Nat'l. Acad. Sci. USA*, 93: 10614-19 (1996); Shalon *et al.*, *Genome Res.*, 6: 639-45 (1996); DeRisi *et al.*, *Science*, 278: 680-86 (1997); Heller *et al.*, *Proc. Nat'l. Acad. Sci. USA*, 94: 2150-55 (1997); Lashkari *et al.*, *Proc. Nat'l. Acad. Sci. USA*, 94: 13057-62 (1997)). Estas micromatrices de ADNc se hibridan simultáneamente con dos sondas de ADNc fluorescentes, marcada cada una con un colorante fluorescente diferente (normalmente Cy3 o Cy5). En este formato, la expresión



de ARNm relativa en dos muestras se compara directamente para cada gen en la micromatriz. Los alineamientos de oligonucleótidos difieren de los alineamientos de ADNc en que las sondas tienen oligonucleótidos de 20 a 25 meros. Los alineamientos de oligonucleótidos se producen generalmente mediante síntesis de oligonucleótidos *in situ* conjuntamente con técnicas de enmascaramiento fotolitográfico (Pease *et al.*, Proc. Nat'l. Acad. Sci. USA, 91: 5022-26 (1994); Lipshutz *et al.*, Biotechniques 19: 442-47 (1995); Chee *et al.*, Science, 274: 610-14 (1996); Lockhart *et al.*, Nature Biotechnol., 14: 1675-80 (1996); Wodicka *et al.*, Nature Biotechnol., 15: 1359-6714 (1997)). El soporte sólido para los alineamientos de oligonucleótidos es normalmente una superficie de vidrio o de silicio.

Se han descrito métodos y técnicas aplicables a la síntesis y el uso de alineamientos, por ejemplo, en las patentes estadounidenses n.ºs 5.143.854 (Pirrung), 5.242.974 (Holmes), 5.252.743 (Barrett), 5.324.633 (Fodor), 5.384.261 (Winkler), 5.424.186 (Fodor), 5.445.934 (Fodor), 5.451.683 (Barrett), 5.482.867 (Barrett), 5.491.074 (Aldwin), 5.527.681 (Holmes), 5.550.215 (Holmes), 5.571.639 (Hubbell), 5.578.832 (Trulson), 5.593.839 (Hubbell), 5.599.695 (Pease), 5.624.711 (Sundberg), 5.631.734 (Stern), 5.795.716 (Chee), 5.831.070 (Pease), 5.837.832 (Chee), 5.856.101 (Hubbell), 5.858.659 (Sapolsky), 5.936.324 (Montagu), 5.968.740 (Fodor), 5.974.164 (Chee), 5.981.185 (Matson), 5.981.956 (Stern), 6.025.601 (Trulson), 6.033.860 (Lockhart), 6.040.193 (Winkler), 6.090.555 (Fiekowsky) y 6.410.229 (Lockhart) y la publicación de solicitud de patente estadounidense n.º 2003/0104411 (Fodor).

Los micromatrices pueden producirse generalmente usando una variedad de técnicas, tales como métodos de síntesis mecánica o dirigida por luz que incorporan una combinación de métodos fotolitográficos y métodos de síntesis en fase sólida. Se describen técnicas para la síntesis de micromatrices usando métodos de síntesis mecánica en, por ejemplo, las patentes estadounidenses n.ºs 5.384.261 (Winkler) y 6.040.193 (Winkler). Aunque se prefiere una superficie de alineamiento plana, la micromatriz puede fabricarse sobre una superficie de prácticamente cualquier forma, o incluso sobre una multiplicidad de superficies. Las micromatrices pueden ser ácidos nucleicos sobre perlas, geles, superficies poliméricas, fibras tales como fibras ópticas, vidrio o cualquier otro sustrato apropiado. Véanse, por ejemplo, las patentes estadounidenses n.ºs 5.708.153 (Dower), 5.770.358 (Dower), 5.789.162 (Dower), 5.800.992 (Fodor) y 6.040.193 (Winkler).

Las micromatrices pueden empaquetarse de tal manera que se permite el uso diagnóstico, o pueden ser dispositivos de todo incluido. Véanse, por ejemplo, las patentes estadounidenses n.ºs 5.856.174 (Lipshutz) y 5.922.591 (Anderson).

Se dispone comercialmente de micromatrices dirigidas a una variedad de fines de Affymetrix (Santa Clara, CA). Por ejemplo, estas micromatrices pueden usarse para genotipado y monitorización de la expresión génica.

Los datos de expresión génica pueden usarse para identificar genes que se regulan de manera coordinada. Los genes que codifican para componentes del mismo complejo de proteínas de múltiples subunidades a menudo se regulan de manera coordinada. También se observa regulación coordinada entre genes cuyos productos funcionan en un programa de diferenciación común o en la misma ruta de respuesta fisiológica. La reciente aplicación de la obtención de perfiles de expresión génica al sistema inmunitario ha mostrado que la diferenciación y activación de linfocitos van acompañadas por cambios paralelos en la expresión entre cientos de genes. Pueden usarse bases de datos de expresión génica para interpretar los cambios patológicos en la expresión génica que acompañan a autoinmunidad, deficiencias inmunitarias, cánceres de células inmunitarias y de respuestas inmunitarias normales.

El examen y la interpretación de grandes cuerpos de datos relativos de expresión génica es una tarea formidable. Esta tarea se facilita en gran medida por algoritmos diseñados para organizar los datos de forma que destacan características sistemáticas, y mediante herramientas de visualización que representan la expresión diferencial de cada gen como intensidades y matices de color variables (Eisen *et al.*, Proc. Nat'l. Acad. Sci. USA, 95: 14863-68 (1998)). El desarrollo de micromatrices, que pueden generar cantidades masivas de datos de expresión en un solo experimento, ha aumentando en gran medida la necesidad de métodos más rápidos y más eficaces de análisis de los conjuntos de datos de expresión a gran escala. Para utilizar eficazmente los datos de expresión génica de micromatrices para la predicción de supervivencia en pacientes con DLBCL, existe la necesidad de desarrollar nuevos algoritmos que puedan identificar información importante y convertirla en un formato más manejable. Además, las micromatrices usadas para generar estos datos pueden estar optimizadas para incorporar conjuntos de sondas que son útiles para la predicción del resultado de supervivencia.

El análisis matemático de los datos de expresión génica es una ciencia que evoluciona rápidamente basada en una ciencia matemática rica de reconocimiento de patrones desarrollada en otros contextos (Kohonen, Self-Organizing Maps, Springer Press (Berlín 1997)). Puede usarse el análisis matemático de los datos de expresión génica, por ejemplo, para identificar grupos de genes que se regulan de manera coordinada dentro de un sistema biológico, para reconocer e interpretar similitudes entre muestras biológicas basándose en similitudes en patrones de expresión génica, y/o para reconocer e interpretar aquellas características de un patrón de expresión génica que están relacionadas con procesos o fenotipos biológicos distintos.

El análisis matemático de los datos de expresión génica a menudo comienza estableciendo el patrón de expresión para cada gen sobre un alineamiento a través de un número (n) de muestras experimentales. El patrón de expresión de cada gen puede representarse por un punto en un espacio n-dimensional, especificándose cada coordenada mediante una medición de expresión en una de las n muestras (Eisen *et al.*, Proc. Nat'l. Acad. Sci. USA, 95: 14863-

68 (1998)). Entonces puede aplicarse un algoritmo de agrupamiento que usa métricas de distancia para localizar agrupaciones de genes en este espacio n-dimensional. Estas agrupaciones indican genes con patrones similares de variación en la expresión a lo largo de una serie de experimentos. Los métodos de agrupamiento que se han aplicado a datos de micromatrices en el pasado incluyen agrupamiento jerárquico (Eisen *et al.*, citado anteriormente), mapas de auto-organización (SOM) (Tamayo *et al.*, Proc. Nat'l. Acad. Sci. USA, 96: 2907-12 (1999)), k-medias (Tavazoie *et al.*, Nature Genet., 22: 281-85 (1999)) y apareamiento determinístico (Alon *et al.*, Proc. Nat'l. Acad. Sci. USA, 96: 6745-50 (1999)).

Puede requerirse una variedad de algoritmos diferentes, destacando cada uno características de orden distinto de los datos, para recoger el conocimiento biológico máximo de un conjunto de muestras (Alizadeh *et al.*, J. Clin. Immunol., 18: 373-79 (1998)). Uno de tales algoritmos, el agrupamiento jerárquico, comienza determinando los coeficientes de correlación de expresión génica para cada par de los n genes estudiados. Los genes con coeficientes de correlación de expresión génica similares se agrupan uno junto al otro de forma jerárquica. Generalmente, genes con patrones de expresión similares en un conjunto de condiciones particulares pueden codificar para productos proteicos con papeles relacionados en la adaptación fisiológica a esas condiciones. Los genes novedosos de función desconocida que se agrupan con un gran grupo de genes relacionados funcionalmente participan posiblemente en procesos biológicos similares o relacionados. Asimismo, otros métodos de agrupamiento mencionados en el presente documento también pueden agrupar entre sí genes que codifican para proteínas con función biológica relacionada.

En tales métodos de agrupamiento, los genes que se agrupan entre sí reflejan una función biológica particular, y se denominan firmas de expresión génica (Shaffer *et al.*, Immunity 15: 375-85 (2001)). Un tipo general de firma de expresión génica incluye genes que se expresan de manera característica en un tipo de célula particular o en una fase de diferenciación o activación celular particular. Otro tipo general de firma de expresión génica incluye genes que se regulan en su expresión por un proceso biológico particular tal como proliferación, o por la actividad de un factor de transcripción o ruta de señalización particular.

El patrón de expresión génica en una muestra biológica puede proporcionar una imagen molecular particular y accesible de su estado e identidad funcionales (DeRisi *et al.*, Science, 278: 680-86 (1997); Cho *et al.*, Mol. Cell., 2: 65-73 (1998); Chu *et al.*, Science, 282: 699-705 (1998); Holstege *et al.*, Cell., 95: 717-728 (1998); Spellman *et al.*, Mol. Biol. Cell, 9: 3273-97 (1998)). Cada célula transduce variaciones en su entorno, estado interno y estado de desarrollo en variaciones reconocibles y medidas fácilmente en sus patrones de expresión génica. Por tanto, es probable que dos muestras diferentes con patrones de expresión génica relacionados sean biológica y funcionalmente similares entre sí. Por tanto, una firma de expresión génica específica en una muestra puede proporcionar conocimientos biológicos importantes en su composición celular y la función de diversas rutas intracelulares dentro de esas células.

Las bases de datos de firmas de expresión génica han demostrado ser útiles para elucidar los patrones expresión génica complejos de diversos cánceres. Por ejemplo, el patrón de expresión de genes en la firma de células B de centro germinal en una biopsia de linfoma indica que el linfoma incluye células derivadas de la fase de diferenciación de centro germinal. En la misma biopsia de linfoma, la expresión de genes de la firma de células T puede usarse para estimar el grado de infiltración del tumor por las células T del huésped, mientras que la expresión de genes de la firma de proliferación puede usarse para cuantificar la tasa de proliferación de células tumorales. De esta manera, las firmas de expresión génica proporcionan un "resumen ejecutivo" de las propiedades biológicas de una muestra de tumor. Las firmas de expresión génica también pueden ser útiles para interpretar los resultados de un análisis supervisado de datos de expresión génica. Un análisis supervisado genera una lista de genes con patrones de expresión que se correlacionan con la supervivencia. Las firmas de expresión génica pueden ser útiles para asignar estos genes "de predicción" a categorías funcionales. En la construcción de un modelo de múltiples variables de supervivencia basado en datos de expresión génica, esta categorización funcional ayuda a limitar la inclusión de múltiples genes en el modelo que mide el mismo aspecto de la biología tumoral.

El siguiente enfoque se utilizó para crear modelos de predicción de supervivencia para el DLBCL de la invención. Se usaron perfiles de expresión génica para crear modelos de múltiples variables para predecir la supervivencia. Los métodos para crear estos modelos estaban "supervisados" porque usaban datos clínicos para guiar la selección de genes que iban a usarse en la clasificación de pronóstico. El método identificó genes con patrones de expresión que se correlacionaban con la duración de la supervivencia global tras quimioterapia. Generalmente, el procedimiento para identificar el modelo de múltiples variables para predecir la supervivencia incluía las siguientes etapas:

1. Se identificaron genes que tenían patrones de expresión asociados utilizando una sola variable con un resultado clínico particular usando un modelo de riesgos proporcionales de Cox. Generalmente, un valor de p de una sola variable <0,01 se considera el punto de corte para significación (sin embargo, puede usarse otro criterio). Estos genes se denominaron genes "de predicción".

2. Dentro de un conjunto de genes de predicción, se identificaron firmas de expresión génica.

3. Para cada firma de expresión génica asociada significativamente con supervivencia, se usó la expresión promedio de cada gen componente dentro de esta firma para generar un valor de firma de expresión génica.

4. Se construyó un modelo de Cox de múltiples variables de resultado clínico usando los valores de firma de expresión génica.

5. Se añadieron genes adicionales al modelo, que se añadieron a la potencia estadística del modelo.

5 El modelo de la invención genera una puntuación de factor de predicción de supervivencia, estando asociada una puntuación superior con un resultado clínico peor. El modelo resultante puede usarse por separado para predecir un resultado de supervivencia. Alternativamente, el modelo puede usarse conjuntamente con uno o más de otros modelos, dados a conocer en el presente documento o en otras referencias, para predecir un resultado de supervivencia.

10 La presente invención da a conocer varias firmas de expresión génica relacionadas con el resultado clínico de pacientes con DLBCL. Las firmas se identificaron usando los datos clínicos y métodos descritos más adelante en los ejemplos 1 y 2. Tres de estas firmas de expresión génica son la firma de células B de centro germinal (GCB), la firma de estromal-1 y la firma de estromal-2. Cada gen componente de estas firmas se identifica en la tabla 1 según su número de registro de GenBank, su ID de gen asignada por Entrez Gene, un símbolo de gen común y un título de gen descriptivo. La tabla 1 también proporciona la ID de conjunto de sondas de Affymetrix, que pueden usarse (por ejemplo, en la micromatriz de Affymetrix U133+ (Affymetrix, Santa Clara, CA)) para determinar el nivel de expresión génica para el gen indicado. La lista de secuencias legible por ordenador presentada con el presente documento incluye una secuencia de fragmento representativa (de aproximadamente 100 pb o mayor) para cada secuencia diana genómica enumerada en la tabla 1, seguido por la secuencia para cada sonda en el conjunto de sondas de Affymetrix correspondiente enumerado en la tabla 1.

20 Tabla 1

| Firma | N.º de registro de GenBank | ID de gen de Entrez | Símbolo del gen | Título del gen  | ID de conjunto de sondas de Affymetrix |
|-------|----------------------------|---------------------|-----------------|---|--|
| GCB   | NM_052932                  | 114908              | TMEM123         | proteína transmembrana 123  | 211967_at                              |
| GCB   | NM_001014380               | 84056               | KATNAL1         | proteína 1 similar a subunidad A de katanina p60  | 227713_at                              |
| GCB   | NM_004665                  | 8875                | VNN2            | vanina 2  | 205922_at                              |
| GCB   | NM_004760                  | 9263                | STK17A          | serina/treonina cinasa 17a (inductora de apoptosis)   | 202693_s_at                            |
| GCB   | CR590554                   |                     |                 | Clon de ADNc de longitud completa CS0DF007YJ21 de cerebro fetal de <i>Homo sapiens</i> (ser humano) | 228464_at                              |
| GCB   | NM_017599                  | 55591               | VEZT            | vezatina, proteína transmembrana de uniones adherentes  | 223089_at                              |
| GCB   | NM_018351                  | 55785               | FGD6            | Proteína 6 que contiene los dominios FYVE, RhoGEF y PH  | 1555136_at                             |
| GCB   | NM_001007075               | 51088               | KLHL5           | proteína 5 similar a kelch ( <i>Drosophila</i> )  | 226001_at                              |
| GCB   | NM_004845                  | 9468                | PCYT1B          | fosfato citidililtransferasa 1, colina, beta  | 228959_at                              |
| GCB   | AK026881                   |                     |                 | ADNc: FLJ23228 fis, clon CAE06654   | 226799_at                              |
| GCB   | NM_018440                  | 55824               | PAG1            | fosfoproteína asociada con microdominios de glicosfingolípido 1                                     | 225626_at                              |
| GCB   | NM_004965                  | 3150                | HMG1            | dominio 1 de unión a nucleosoma de grupo de alta movilidad  | 200944_s_at                            |
| GCB   | NM_001706                  | 604                 | BCL6            | linfoma 6/LLC de células B (proteína de dedos de cinc 51)   | 228758_at                              |
| GCB   | NM_020747                  | 57507               | ZNF608          | proteína de dedos de cinc 608   | 229817_at                              |
| GCB   | NM_001001695               | 400941              | FLJ42418        | proteína FLJ42418   | 231455_at                              |
| GCB   | NM_015055                  | 23075               | SWAP70          | proteína SWAP-70  | 209306_s_at                            |
| GCB   | NM_005607                  | 5747                | PTK2            | PTK2 proteína tirosina cinasa 2   | 208820_at                              |
| GCB   | XM_027236                  | 23508               | TTC9            | dominio 9 de repetición de  | 213172_at                              |

| Firma      | N.º de registro de GenBank | ID de gen de Entrez | Símbolo del gen | Título del gen   | ID de conjunto de sondas de Affymetrix |
|------------|----------------------------|---------------------|-----------------|--|--|
|            |                            |                     |                 | tetratricopéptido  |  |
| GCB        | BQ213652                   | 440864              | LOC440864       | gen hipotético soportado por BC040724  | 1569034_a_at                           |
| GCB        | NM_005574                  | 4005                | LMO2            | dominio LIM sólo 2 (proteína 1 similar a rombotina)  | 204249_s_at                            |
| GCB        | NM_014667                  | 9686                | VGLL4           | proteína 4 similar a vestigial ( <i>Drosophila</i> )   | 212399_s_at                            |
| GCB        | NM_002221                  | 3707                | ITPKB           | inositol 1,4,5-trifosfato 3-cinasa B   | 203723_at                              |
| GCB        | NM_000902                  | 4311                | MME             | metalo-endopeptidasa de membrana (endopeptidasa neutra, encefalinasa)                        | 203434_s_at                            |
| GCB        | NM_012446                  | 23635               | SSBP2           | proteína 2 de unión a ADN monocatenario  | 203787_at                              |
| GCB        | NM_024613                  | 79666               | PLEKHF2         | miembro 2 de la familia F (con dominio FYVE) que contiene dominio de homología a pleckstrina | 222699_s_at                            |
| GCB        | AV705976                   |                     |                 | Locus transcrito   | 204681_s_at                            |
| GCB        | NM_012108                  | 26228               | BRDG1           | proteína 1 de señalización posterior de BCR 1  | 220059_at                              |
| GCB        | NM_014397                  | 10783               | NEK6            | cinasa 6 relacionada con NIMA (gen a nunca en mitosis)                                       | 223158_s_at                            |
| GCB        | NM_018981                  | 54431               | DNAJC10         | homólogo de DnaJ (Hsp40), subfamilia C, miembro 10   | 225174_at                              |
| GCB        | NM_001379                  | 1786                | DNMT1           | ADN (citosina-5-)-metiltransferasa 1   | 227684_at                              |
| GCB        | NM_006152                  | 4033                | LRMP            | proteína de membrana restringida a linaje linfoide   | 35974_at                               |
| GCB        | NM_024701                  | 79754               | ASB13           | proteína 13 que contiene repetición de anquirina y caja SOCS                                 | 218862_at                              |
| GCB        | NM_006085                  | 10380               | BPNT1           | 3'(2'), 5'-bisfosfato nucleotidasa 1   | 232103_at                              |
| GCB        | NM_023009                  | 65108               | MARCKSL1        | proteína 1 similar a MARCKS  | 200644_at                              |
| GCB        | NM_033121                  | 88455               | ANKRD13A        | dominio 13A de repetición de anquirina   | 224810_s_at                            |
| GCB        | NM_015187                  | 23231               | KIAA0746        | proteína KIAA0746  | 235353_at                              |
| GCB        | NM_175739                  | 327657              | SERPINA9        | inhibidor de serpina peptidasa, clado A (alfa-1 antiproteinasa, antitripsina), miembro 9     | 1553499_s_at                           |
| GCB        | NM_001012391               | 400509              | RUNDC2B         | proteína 2B que contiene dominio RUN   | 1554413_s_at                           |
| GCB        | XM_034274                  | 4603                | MYBL1           | proteína 1 similar a homólogo (aviar) de oncogén de mieloblastosis viral v-myb 1             | 213906_at                              |
| Estromal-1 | NM_024579                  | 79630               | C1orf54         | marco de lectura abierto 54 del cromosoma 1  | 219506_at                              |
| Estromal-1 | NM_001645                  | 341                 | APOC1           | apolipoproteína C-1  | 213553_x_at                            |
| Estromal-1 | NM_001562                  | 3606                | IL18            | interleucina 18 (factor de inducción de interferón-gamma)                                    | 206295_at                              |
| Estromal-1 | NM_014479                  | 27299               | ADAMDEC1        | decisina 1; similar a ADAM   | 206134_at                              |
| Estromal-1 | NM_003465                  | 1118                | CHIT1           | quitinasa 1 (quitotriosidasa)  | 208168_s_at                            |
| Estromal-1 | NM_000954                  | 5730                | PTGDS           | prostaglandina D2 sintasa de 21 kDa (cerebro)  | 211748_x_at                            |
| Estromal-1 | NM_001056                  | 6819                | SULT1C1         | familia de sulfotransferasa, citosólico, 1C, miembro 1                                       | 211470_s_at                            |
| Estromal-1 | NM_018000                  | 55686               | MREG            | melanorregulina  | 219648_at                              |

| Firma      | N.º de registro de GenBank | ID de gen de Entrez | Símbolo del gen | Título del gen  | ID de conjunto de sondas de Affymetrix |
|------------|----------------------------|---------------------|-----------------|---|--|
| Estromal-1 | NM_001018058               | 22797               | TFEC            | factor de transcripción EC  | 206715_at                              |
| Estromal-1 | NM_000239                  | 4069                | LYZ             | lizozima (amiloidosis renal)  | 213975_s_at                            |
| Estromal-1 | NM_006834                  | 10981               | RAB32           | RAB32, miembro de la familia del oncogén RAS  | 204214_s_at                            |
| Estromal-1 | NM_000416                  | 3459                | IFNGR1          | receptor 1 de interferón gamma  | 202727_s_at                            |
| Estromal-1 | NM_004666                  | 8876                | VNN1            | vanina 1  | 205844_at                              |
| Estromal-1 | NM_031491                  | 83758               | RBP5            | proteína 5 de unión a retinol, celular  | 223820_at                              |
| Estromal-1 | NM_001276                  | 1116                | CHI3L1          | proteína 1 similar a quitinasa 3 (glicoproteína-39 de cartílago)                                    | 209396_s_at                            |
| Estromal-1 | NM_138434                  | 113763              | C7orf29         | marco de lectura abierto 29 del cromosoma 7   | 227598_at                              |
| Estromal-1 | NM_001005340               | 10457               | GPMB            | glicoproteína (transmembrana) nmb   | 201141_at                              |
| Estromal-1 | NM_002294                  | 3920                | LAMP2           | proteína 2 de membrana asociada al lisosoma   | 203041_s_at                            |
| Estromal-1 | NM_002888                  | 5918                | RARRES1         | agente de respuesta 1 (inducido por tazaroteno) al receptor de ácido retinoico                      | 221872_at                              |
| Estromal-1 | NM_172248                  | 1438                | CSF2RA          | receptor del factor 2 estimulante de colonias, alfa, de baja afinidad (granulocitos-macrófagos)     | 210340_s_at                            |
| Estromal-1 | NM_018344                  | 55315               | SLC29A3         | familia 29 de portador de solutos (transportadores de nucleósidos), miembro 3                       | 219344_at                              |
| Estromal-1 | NM_032413                  | 84419               | C15orf48        | marco de lectura abierto 48 del cromosoma 15  | 223484_at                              |
| Estromal-1 | NM_001001851               | 80760               | ITIH5           | inhibidor H5 inter-alfa (globulina)   | 1553243_at                             |
| Estromal-1 | NM_000211                  | 3689                | ITGB2           | integrina, beta 2 (subunidad del receptor 3 de los componentes del complemento 3 y 4)               | 1555349_a_at                           |
| Estromal-1 | NM_005213                  | 1475                | CSTA            | cistatina A (estefina A)  | 204971_at                              |
| Estromal-1 | NM_003874                  | 8832                | CD84            | molécula de CD84  | 205988_at                              |
| Estromal-1 | NM_000228                  | 3914                | LAMB3           | laminina, beta 3  | 209270_at                              |
| Estromal-1 | NM_005651                  | 6999                | TDO2            | triptófano 2,3-dioxigenasa  | 205943_at                              |
| Estromal-1 | NM_001005266               | 283651              | C15orf21        | marco de lectura abierto 21 del cromosoma 15  | 242649_x_at                            |
| Estromal-1 | AV659177                   |                     |                 | Locus transcrito  | 230391_at                              |
| Estromal-1 | NM_001747                  | 822                 | CAPG            | proteína con protección (filamento de actina), similar a gelsolina                                  | 201850_at                              |
| Estromal-1 | NM_000784                  | 1593                | CYP27A1         | citocromo P450, familia 27, subfamilia A, polipéptido 1   | 203979_at                              |
| Estromal-1 | NM_052998                  | 113451              | ADC             | arginina descarboxilasa   | 228000_at                              |
| Estromal-1 | NM_016240                  | 51435               | SCARA3          | clase A de receptor eliminador, miembro 3   | 219416_at                              |
| Estromal-1 | Z74615                     |                     | COL1A1          | Colágeno, tipo I, alfa 1  | 217430_x_at                            |
| Estromal-1 | NM_052947                  | 115701              | ALPK2           | alfa-cinasa 2   | 228367_at                              |
| Estromal-1 | NM_021136                  | 6252                | RTN1            | reticulón 1   | 210222_s_at                            |
| Estromal-1 | AL049370                   |                     |                 | Clon de ADNc de longitud completa CL0BB018ZE07 de neuroblastoma de <i>Homo sapiens</i> (ser humano) | 213100_at                              |
| Estromal-1 | NM_006042                  | 9955                | HS3ST3A1        | heparán sulfato (glucosamina) 3-O-sulfotransferasa 3A1  | 219985_at                              |

| Firma      | N.º de registro de GenBank | ID de gen de Entrez | Símbolo del gen | Título del gen  | ID de conjunto de sondas de Affymetrix |
|------------|----------------------------|---------------------|-----------------|---|--|
| Estromal-1 | NM_000041                  | 348                 | APOE            | apolipoproteína E   | 203382_s_at                            |
| Estromal-1 | NM_004994                  | 4318                | MMP9            | metalopectidasa de la matriz 9 (gelatinasa B, gelatinasa de 92 kDa, colagenasa de tipo IV de 92kDa)   | 203936_s_at                            |
| Estromal-1 | NM_001831                  | 1191                | CLU             | clusterina  | 222043_at                              |
| Estromal-1 | NM_002305                  | 3956                | LGALS1          | lectina, de unión a galactósido, soluble, 1 (galectina 1)   | 201105_at                              |
| Estromal-1 | NM_032024                  | 83938               | C10orf11        | marco de lectura abierto 11 del cromosoma 10  | 223703_at                              |
| Estromal-1 | NM_001025201               | 1123                | CHN1            | quimerina (quimaerina) 1  | 212624_s_at                            |
| Estromal-1 | NM_003489                  | 8204                | NRIP1           | proteína 1 de interacción con receptor nuclear  | 202599_s_at                            |
| Estromal-1 | NM_032646                  | 94015               | TTYH2           | homólogo 2 de tweety ( <i>Drosophila</i> )  | 223741_s_at                            |
| Estromal-1 | NM_001312                  | 1397                | CRIP2           | proteína 2 rica en cisteína   | 208978_at                              |
| Estromal-1 | NM_023075                  | 65258               | MPPE1           | metalofosfoesterasa 1   | 213924_at                              |
| Estromal-1 | NM_004364                  | 1050                | CEBPA           | proteína de unión a CCAAT/potenciador (C/EBP), alfa   | 204039_at                              |
| Estromal-1 | NM_000248                  | 4286                | MITF            | factor de transcripción asociado a microftalmia   | 207233_s_at                            |
| Estromal-1 | NM_002185                  | 3575                | IL7R            | receptor de interleucina 7  | 226218_at                              |
| Estromal-1 | NM_021638                  | 60312               | AFAP            | proteína asociada a filamentos de actina  | 203563_at                              |
| Estromal-1 | NM_003786                  | 8714                | ABCC3           | casete de unión a ATP, subfamilia C (CFTR/MRP), miembro 3   | 208161_s_at                            |
| Estromal-1 |                            | 730351              | LOC730351       | proteína hipotética LOC730351   | 229407_at                              |
| Estromal-1 | NM_012153                  | 26298               | EHF             | factor homólogo a ets   | 225645_at                              |
| Estromal-1 | NM_004887                  | 9547                | CXCL14          | ligando 14 de quimiocina (motivo C-X-C)   | 222484_s_at                            |
| Estromal-1 | NM_002030                  | 2359                | FPRL2           | proteína 2 similar a receptor de péptido de formilo   | 230422_at                              |
| Estromal-1 | NM_001321                  | 1466                | CSRP2           | proteína 2 rica en cisteína y glicina   | 207030_s_at                            |
| Estromal-1 | NM_001945                  | 1839                | HBEGF           | factor de crecimiento similar a EGF de unión a heparina   | 203821_at                              |
| Estromal-1 | NM_031412                  | 23710               | GABARAPL1       | proteína 1 similar a proteína asociada a receptor de GABA(A)  | 208869_s_at                            |
| Estromal-1 | NM_006022                  | 8848                | TSC22D1         | familia de dominio TSC22, miembro 1   | 215111_s_at                            |
| Estromal-1 | NM_016174                  | 51148               | CEECAM1         | molécula 1 de adhesión a células endoteliales cerebrales  | 224794_s_at                            |
| Estromal-1 | NM_015103                  | 23129               | PLXND1          | plexina D1  | 212235_at                              |
| Estromal-1 | NM_003270                  | 7105                | TSPAN6          | tetraespanina 6   | 209109_s_at                            |
| Estromal-1 | NM_000887                  | 3687                | ITGAX           | integrina, alfa X (subunidad del receptor 4 del componente del complemento 3)                         | 210184_at                              |
| Estromal-1 | NM_001864                  | 1346                | COX7A1          | polipéptido 1 (músculo) de subunidad VIIa de citocromo oxidasa  | 204570_at                              |
| Estromal-1 | CR599008                   |                     | GPR157          | clon de ADNc de longitud completa CS0DJ007YL22 de células T (línea celular Jurkat) Cot 10 normalizado | 227970_at                              |

| Firma      | N.º de registro de GenBank | ID de gen de Entrez | Símbolo del gen | Título del gen  | ID de conjunto de sondas de Affymetrix |
|------------|----------------------------|---------------------|-----------------|---|--|
|            |                            |                     |                 | de <i>Homo sapiens</i> (ser humano)   |  |
| Estromal-1 | NM_198580                  | 376497              | SLC27A1         | familia 27 de portadores de solutos (transportador de ácidos grasos), miembro 1 | 226728_at                              |
| Estromal-1 | NM_025106                  | 80176               | SPSB1           | proteína 1 que contiene dominio receptor de splA/rianodina y caja SOCS          | 226075_at                              |
| Estromal-1 | NM_020130                  | 56892               | C8orf4          | marco de lectura abierto 4 del cromosoma 8                                      | 218541_s_at                            |
| Estromal-1 | NM_173833                  | 286133              | SCARA5          | clase A de receptor eliminador, miembro 5 (supuesto)                            | 229839_at                              |
| Estromal-1 | NM_007223                  | 11245               | GPR176          | receptor 176 acoplado a proteínas G   | 227846_at                              |
| Estromal-1 | NM_013437                  | 29967               | LRP12           | proteína 12 relacionada con lipoproteína de baja densidad                       | 219631_at                              |
| Estromal-1 | NM_007332                  | 8989                | TRPA1           | canal catiónico de potencial de receptor transitorio, subfamilia A, miembro 1   | 228438_at                              |
| Estromal-1 | NM_152744                  | 221935              | SDK1            | homólogo 1 de sidekick (pollo)  | 229912_at                              |
| Estromal-1 | NM_001409                  | 1953                | MEGF6           | dominios 6 similares a EGF múltiple   | 226869_at                              |
| Estromal-1 | NM_012082                  | 23414               | ZFPM2           | proteína de dedos de cinc, multitypo2   | 219778_at                              |
| Estromal-1 | NM_080430                  | 140606              | SELM            | selenoproteína M  | 226051_at                              |
| Estromal-1 | NM_030971                  | 81855               | SFXN3           | sideroflexina 3   | 217226_s_at                            |
| Estromal-1 | NM_003246                  | 7057                | THBS1           | trombospondina 1  | 201109_s_at                            |
| Estromal-1 | NM_003882                  | 8840                | WISP1           | proteína 1 de la ruta de señalización inducible de WNT1                         | 235821_at                              |
| Estromal-1 | NM_005202                  | 1296                | COL8A2          | colágeno, tipo VIII, alfa 2   | 221900_at                              |
| Estromal-1 | NM_003711                  | 8611                | PPAP2A          | ácido fosfatídico fosfatasa tipo 2A   | 210946_at                              |
| Estromal-1 | NM_004995                  | 4323                | MMP14           | metalopeptidasa de la matriz 14 (insertada en la membrana)                      | 202828_s_at                            |
| Estromal-1 | NM_001005336               | 1759                | DNM1            | dinamina 1  | 215116_s_at                            |
| Estromal-1 | NM_153717                  | 2121                | EVC             | síndrome de Ellis van Creveld   | 219432_at                              |
| Estromal-1 | NM_173462                  | 89932               | PAPLN           | papilina, glicoproteína sulfatada similar a proteoglicano                       | 226435_at                              |
| Estromal-1 | XM_496707                  | 441027              | FLJ12993        | LOC441027 hipotética  | 229623_at                              |
| Estromal-1 | NM_001839                  | 1266                | CNN3            | calponina 3, ácida  | 228297_at                              |
| Estromal-1 | NM_015429                  | 25890               | ABI3BP          | proteína de unión a la familia génica ABI, miembro 3 (NESH)                     | 223395_at                              |
| Estromal-1 | NM_002840                  | 5792                | PTPRF           | proteína tirosina fosfatasa, tipo receptor, F                                   | 200636_s_at                            |
| Estromal-1 | NM_001001522               | 6876                | TAGLN           | transgelina   | 1555724_s_at                           |
| Estromal-1 | NM_017637                  | 54796               | BNC2            | basonuclina 2   | 229942_at                              |
| Estromal-1 | NM_003391                  | 7472                | WNT2            | miembro 2 de la familia de sitio de integración de MMTV de tipo wingless        | 205648_at                              |
| Estromal-1 | NM_015461                  | 25925               | ZNF521          | proteína de dedos de cinc 521   | 226677_at                              |
| Estromal-1 | NM_006475                  | 10631               | POSTN           | periostina, factor específico de osteoblastos                                   | 210809_s_at                            |
| Estromal-1 | NM_005418                  | 6764                | ST5             | supresión de tumorigenicidad  | 202440_s_at                            |

| Firma      | N.º de registro de GenBank | ID de gen de Entrez | Símbolo del gen | Título del gen   | ID de conjunto de sondas de Affymetrix |
|------------|----------------------------|---------------------|-----------------|--|--|
|            |                            |                     |                 | 5  |  |
| Estromal-1 | NM_005203                  | 1305                | COL13A1         | colágeno, tipo XIII, alfa 1  | 211343_s_at                            |
| Estromal-1 | NM_000681                  | 150                 | ADRA2A          | receptor adrenérgico, alfa-2A  | 209669_at                              |
| Estromal-1 | NM_006622                  | 10769               | PLK2            | cinasa 2 similar a polo ( <i>Drosophila</i> )  | 201939_at                              |
| Estromal-1 | AL528626                   |                     |                 | Clon de ADNc de longitud completa CS0DD001YA12 de neuroblastoma Cot 50 normalizado de <i>Homo sapiens</i> (ser humano) | 228573_at                              |
| Estromal-1 | AF180519                   | 23766               | GABARAPL3       | proteína 3 similar a proteína asociada a receptores de GABA(A)   | 211458_s_at                            |
| Estromal-1 | NM_024723                  | 79778               | MICALL2         | proteína 2 similar a MICAL   | 219332_at                              |
| Estromal-1 | NM_057177                  | 117583              | PARD3B          | homólogo B 3 defectuoso en el reparto de par-3 ( <i>C. elegans</i> )   | 228411_at                              |
| Estromal-1 | NM_004949                  | 1824                | DSC2            | desmocolina 2  | 226817_at                              |
| Estromal-1 | NM_032784                  | 84870               | RSPO3           | homólogo 3 de R-espondina ( <i>Xenopus laevis</i> )  | 228186_s_at                            |
| Estromal-1 | NM_007039                  | 11099               | PTPN21          | proteína tirosina fosfatasa, tipo 21 no receptora  | 226380_at                              |
| Estromal-1 | NM_031935                  | 83872               | HMCN1           | hemicentina 1  | 235944_at                              |
| Estromal-1 | AK022877                   |                     |                 | ARNm de la región Cri-du-chat del clon TUA8  | 213169_at                              |
| Estromal-1 | AK127644                   |                     |                 | ADNc FLJ45742 fis, clon KIDNE2016327   | 236297_at                              |
| Estromal-1 | AK056963                   |                     |                 | clon de ADNc de inserto de longitud completa ZE03F06   | 226282_at                              |
| Estromal-1 | NM_000899                  | 4254                | KITLG           | ligando de KIT   | 226534_at                              |
| Estromal-1 | NM_002387                  | 4163                | MCC             | mutada en cánceres colorrectales   | 226225_at                              |
| Estromal-1 | NM_198270                  | 4810                | NHS             | síndrome de Nance-Horan (cataratas congénitas y anomalías dentales)  | 228933_at                              |
| Estromal-1 | NM_183376                  | 91947               | ARRDC4          | proteína 4 que contiene dominio arrestina  | 225283_at                              |
| Estromal-1 | NM_000216                  | 3730                | KAL1            | secuencia 1 de síndrome de Kallmann  | 205206_at                              |
| Estromal-1 | NM_001008224               | 55075               | UACA            | autoantígeno uveal con dominios superenrollados y repeticiones de anquirina  | 223279_s_at                            |
| Estromal-1 | NM_133493                  | 135228              | CD109           | molécula CD109   | 226545_at                              |
| Estromal-1 | NM_005545                  | 3671                | ISLR            | superfamilia de inmunoglobulinas que contiene repetición rica en leucinas  | 207191_s_at                            |
| Estromal-1 | NM_014365                  | 26353               | HSPB8           | proteína 8 de 22 kDa de choque térmico   | 221667_s_at                            |
| Estromal-1 | NM_014476                  | 27295               | PDLIM3          | dominio 3 de PDZ y LIM   | 209621_s_at                            |
| Estromal-1 | NM_020962                  | 57722               | NOPE            | ortólogo probable de vecino de ratón de Punc E11   | 227870_at                              |
| Estromal-1 | NM_018357                  | 55323               | LARP6           | familia de dominios de ribonucleoproteína La, miembro 6  | 218651_s_at                            |
| Estromal-1 | NM_012323                  | 23764               | MAFF            | homólogo F (aviar) de oncogén de fibrosarcoma musculoaponeurótico v-maf  | 36711_at                               |
| Estromal-1 | NM_003713                  | 8613                | PPAP2B          | ácido fosfatídico fosfatasa tipo 2B  | 212230_at                              |
| Estromal-1 | NM_023016                  | 65124               | ANKRD57         | dominio 57 de repetición de  | 227034_at                              |



| Firma      | N.º de registro de GenBank | ID de gen de Entrez | Símbolo del gen | Título del gen   | ID de conjunto de sondas de Affymetrix |
|------------|----------------------------|---------------------|-----------------|--|--|
|            |                            |                     |                 | anquirina  |  |
| Estromal-1 | NM_032777                  | 25960               | GPR124          | receptor 124 acoplado a proteínas G  | 65718_at                               |
| Estromal-1 | NM_001554                  | 3491                | CYR61           | inductor angiogénico, rico en cisteínas, 61  | 201289_at                              |
| Estromal-1 | NM_145117                  | 89797               | NAV2            | navegador neuronal 2   | 218330_s_at                            |
| Estromal-1 | NM_001002292               | 79971               | GPR177          | receptor 177 acoplado a proteínas G  | 228950_s_at                            |
| Estromal-1 | NM_001401                  | 1902                | EDG2            | receptor acoplado a proteínas G de ácido lisofosfatídico, diferenciación endotelial, 2 | 204036_at                              |
| Estromal-1 | NM_198282                  | 340061              | TMEM173         | proteína transmembrana 173   | 224929_at                              |
| Estromal-1 | NM_014934                  | 22873               | DZIP1           | proteína 1 de interacción con DAZ  | 204556_s_at                            |
| Estromal-1 | NM_001901                  | 1490                | CTGF            | factor de crecimiento de tejido conjuntivo   | 209101_at                              |
| Estromal-1 | NM_024600                  | 79652               | C16orf30        | marco de lectura abierto 30 del cromosoma 16   | 219315_s_at                            |
| Estromal-1 | NM_138370                  | 91461               | LOC91461        | proteína hipotética BC007901   | 225380_at                              |
| Estromal-1 | NM_014632                  | 9645                | MICAL2          | proteína 2 que contiene dominio monooxigenasa asociado a microtúbulos, calponina y LIM | 212472_at                              |
| Estromal-1 | NM_032866                  | 84952               | CGNL1           | proteína 1 similar a cingulina   | 225817_at                              |
| Estromal-1 | NM_003687                  | 8572                | PDLIM4          | dominio 4 de PDZ y LIM   | 211564_s_at                            |
| Estromal-1 | BM544548                   |                     |                 | Locus transcrito   | 236179_at                              |
| Estromal-1 | NM_001856                  | 1307                | COL16A1         | colágeno, tipo XVI, alfa 1   | 204345_at                              |
| Estromal-1 | XM_087386                  | 57493               | HEG1            | homólogo 1 de HEG 1 (pez cebra)  | 213069_at                              |
| Estromal-1 | NM_003887                  | 8853                | DDEF2           | factor 2 de potenciación del desarrollo y la diferenciación                            | 2064144_s_at                           |
| Estromal-1 | NM_002844                  | 5796                | PTPRK           | proteína tirosina fosfatasa, tipo receptor, K  | 203038_at                              |
| Estromal-1 | NM_022138                  | 64094               | SMOC2           | proteína 2 de unión a calcio modular relacionada con SPARC                             | 223235_s_at                            |
| Estromal-1 | NM_001006624               | 10630               | PDPN            | podoplanina  | 204879_at                              |
| Estromal-1 | NM_003174                  | 6840                | SVIL            | supervilina  | 202565_s_at                            |
| Stromat-1  | NM_002845                  | 5797                | PTPRM           | proteína tirosina fosfatasa, tipo receptor, M  | 1555579_s_at                           |
| Estromal-1 | NM_002889                  | 5919                | RARRES2         | agente de respuesta 2 (inducido por tazaroteno) al receptor de ácido retinoico         | 209496_at                              |
| Estromal-1 | NM_006094                  | 10395               | DLC1            | proteína 1 delecionada en cáncer de hígado   | 210762_s_at                            |
| Estromal-1 | NM_022463                  | 64359               | NXN             | nucleorredoxina  | 219489_s_at                            |
| Estromal-1 | AK027294                   |                     |                 | ADNc FLJ14388 fis, clon HEMBA1002716   | 229802_at                              |
| Estromal-1 | NM_005711                  | 10085               | EDIL3           | repeticiones similares a EGF y dominios 3 similares a discoidina I                     | 225275_at                              |
| Estromal-1 | NM_000177                  | 2934                | GSN             | gelsolina (amiloidosis, tipo finés)  | 200696_s_at                            |
| Estromal-1 | NM_016639                  | 51330               | TNFRSF12A       | superfamilia de receptores de factor de necrosis tumoral, miembro 12A                  | 218368_s_at                            |
| Estromal-1 | NM_004460                  | 2191                | FAP             | proteína de activación fibroblastos, alfa  | 209955_s_at                            |
| Estromal-1 | NM_000064                  | 718                 | C3              | componente 3 del complemento   | 217767_at                              |

| Firma      | N.º de registro de GenBank | ID de gen de Entrez | Símbolo del gen | Título del gen   | ID de conjunto de sondas de Affymetrix |
|------------|----------------------------|---------------------|-----------------|--|--|
| Estromal-1 | NM_016206                  | 389136              | VGLL3           | proteína 3 similar a vestigial ( <i>Drosophila</i> )                                       | 227399_at                              |
| Estromal-1 | NM_004339                  | 754                 | PTTG1 IP        | proteína de interacción 1 de transformación de tumor hipofisario                           | 200677_at                              |
| Estromal-1 | NM_003255                  | 7077                | TIMP2           | inhibidor 2 de metalopeptidasa TIMP  | 224560_at                              |
| Estromal-1 | NM_002998                  | 6383                | SDC2            | sindecano 2 (heparán sulfato proteoglicano 1, asociado a superficie celular, fibroglicano) | 212158_at                              |
| Estromal-1 | NM_012223                  | 4430                | MYO1B           | miosina IB   | 212364_at                              |
| Estromal-1 | NM_020650                  | 57333               | RCN3            | reticulocalbina 3, dominio de unión a calcio de mano EF                                    | 61734_at                               |
| Estromal-1 | AL573464                   |                     |                 | Locus transcrito   | 229554_at                              |
| Estromal-1 | AK001903                   |                     |                 | ADNc FLJ11041 fis, clon PLACE1004405   | 227140_at                              |
| Estromal-1 | NM_005928                  | 4240                | MFGE8           | proteína de factor 8 de EGF de glóbulos de la grasa de la leche                            | 210605_s_at                            |
| Estromal-1 | NM_000943                  | 5480                | PPIC            | peptidilproil isomerasa C (ciclofilina C)  | 204518_s_at                            |
| Estromal-1 | NM_001008397               | 493869              | LOC493869       | similar a ADNc RIKEN 2310016C16  | 227628_at                              |
| Estromal-1 | AK025431                   | 768211              | RELL1           | proteína 1 similar a receptor expresado en tejidos linfoides                               | 226430_at                              |
| Estromal-1 | NM_000297                  | 5311                | PKD2            | proteína 2 de enfermedad de riñón poliquistico (autosómica dominante)                      | 203688_at                              |
| Estromal-1 | NM_002975                  | 6320                | CLEC11A         | familia 11 de dominios de lectina de tipo C, miembro A                                     | 211709_s_at                            |
| Estromal-1 | NM_001920                  | 1634                | DCN             | decorina   | 211813_x_at                            |
| Estromal-1 | NM_001723                  | 667                 | DST             | distonina  | 215016_x_at                            |
| Estromal-1 | CR749529                   |                     |                 | ARNm; ADNc DKFZp686118116 (del clon DKFZp686118116)  | 227554_at                              |
| Estromal-1 | NM_000165                  | 2697                | GJA1            | proteína de unión gap, alfa 1, 43 kDa (conexina 43)  | 201667_at                              |
| Estromal-1 | NM_012104                  | 23621               | BACE1           | enzima 1 de escisión de APP de sitio beta  | 217904_s_at                            |
| Estromal-1 | NM_001957                  | 1909                | EDNRA           | receptor de endotelina tipo A  | 204464_s_at                            |
| Estromal-1 | NM_138455                  | 115908              | CTHRC1          | proteína 1 que contiene repetición de triple hélice de colágeno                            | 225681_at                              |
| Estromal-1 | NM_001331                  | 1500                | CTNND1          | catenina (proteína asociada a cadherina), delta 1  | 208407_s_at                            |
| Estromal-1 | NM_001613                  | 59                  | ACTA2           | actina, alfa 2, músculo liso, aorta  | 200974_at                              |
| Estromal-1 | NM_002192                  | 3624                | INHBA           | inhibina, beta A (activina A, polipéptido alfa de activina AB)                             | 210511_s_at                            |
| Estromal-1 | NM_000935                  | 5352                | PLOD2           | procolágeno-lisina, 2-oxoglutarato 5-dioxigenasa 2   | 202620_s_at                            |
| Estromal-1 | NM_015170                  | 23213               | SULF1           | sulfatasa 1  | 212354_at                              |
| Estromal-1 | NM_006039                  | 9902                | MRC2            | receptor de manosa, C tipo 2   | 37408_at                               |
| Estromal-1 | NM_005261                  | 2669                | GEM             | proteína de unión a GTP sobreexpresada en músculo esquelético                              | 204472_at                              |
| Estromal-1 | NM_001008707               | 2009                | EML1            | proteína 1 similar a proteína asociada a microtúbulos de equinodermos                      | 204797_s_at                            |

| Firma      | N.º de registro de GenBank | ID de gen de Entrez | Símbolo del gen | Título del gen  | ID de conjunto de sondas de Affymetrix |
|------------|----------------------------|---------------------|-----------------|---|--|
| Estromal-1 | NM_001031679               | 253827              | MSRB3           | metionina sulfóxido reductasa B3  | 225782_at                              |
| Estromal-1 | NM_001004125               | 286319              | TUSC1           | proteína 1 candidata a supresor de tumores  | 227388_at                              |
| Estromal-1 | NM_005965                  | 4638                | MYLK            | miosina, cinasa de cadena ligera  | 202555_s_at                            |
| Estromal-1 | NM_016205                  | 56034               | PDGFC           | factor de crecimiento C derivado de plaquetas   | 218718_at                              |
| Estromal-1 | NM_015976                  | 51375               | SNX7            | nexina 7 de clasificación   | 205573_s_at                            |
| Estromal-1 | NM_130830                  | 131578              | LRRC15          | proteína 15 que contiene repetición rica en leucinas  | 213909_at                              |
| Estromal-1 | NM_002026                  | 2335                | FN1             | fibronectina 1  | 212464_s_at                            |
| Estromal-1 | NM_006855                  | 11015               | KDEL3           | receptor 3 de retención de proteínas en el retículo endoplasmático KDEL (Lys-Asp-Glu-Leu)   | 204017_at                              |
| Estromal-1 | NM_002292                  | 3913                | LAMB2           | laminina, beta 2 (laminina S)   | 216264_s_at                            |
| Estromal-1 | NM_002658                  | 5328                | PLAU            | activador de plasminógeno, urocinasa  | 205479_s_at                            |
| Estromal-1 | NM_005529                  | 3339                | HSPG2           | heparán sulfato proteoglicano 2 (perlecano)   | 201655_s_at                            |
| Estromal-1 | NM_001235                  | 871                 | SERPINH1        | inhibidor de serpina peptidasa, clado H (proteína 47 de choque térmico), miembro 1, (proteína 1 de unión a colágeno 1)  | 207714_s_at                            |
| Estromal-1 | AJ318805                   | CDNA                | CDNA            | ADNc FLJ44429 fis, clon UTERU2015653  | 227061_at                              |
| Estromal-1 | NM_000396                  | 1513                | CTSK            | catepsina K   | 202450_s_at                            |
| Estromal-1 | NM_031302                  | 83468               | GLT8D2          | proteína 2 que contiene dominio glicosiltransferasa 8   | 227070_at                              |
| Estromal-1 | NM_080821                  | 116151              | C20orf108       | marco de lectura abierto 108 del cromosoma 20   | 224690_at                              |
| Estromal-1 | NM_002345                  | 4060                | LUM             | lumicano  | 201744_s_at                            |
| Estromal-1 | NM_005110                  | 9945                | GFPT2           | glutamina-fructosa-6-fosfato transaminasa 2   | 205100_at                              |
| Estromal-1 | NM_002941                  | 6091                | ROBO1           | roundabout, receptor de guiado de axones, homólogo 1 ( <i>Drosophila</i> )  | 213194_at                              |
| Estromal-1 | NM_005429                  | 7424                | VEGFC           | factor de crecimiento C endotelial vascular   | 209946_at                              |
| Estromal-1 | NM_002213                  | 3693                | ITGB5           | integrina, beta 5   | 201125_s_at                            |
| Estromal-1 | XM_051017                  | 23363               | OBSL1           | proteína 1 similar a obscurina  | 212775_at                              |
| Estromal-1 | NM_181724                  | 338773              | TMEM119         | proteína transmembrana 119  | 227300_at                              |
| Estromal-1 | NM_003474                  | 8038                | ADAM12          | domino 12 de metalopeptidasa ADAM (meltrina alfa)   | 213790_at                              |
| Estromal-1 | NM_018222                  | 55742               | PARVA           | parvina, alfa   | 217890_s_at                            |
| Estromal-1 | NM_006478                  | 10634               | GAS2L1          | proteína 1 similar a proteína 2 específica de detención del crecimiento   | 31874_at                               |
| Estromal-1 | NM_000093                  | 1289                | COL5A1          | colágeno, tipo V, alfa 1  | 212489_at                              |
| Estromal-1 | NM_006288                  | 7070                | THY1            | antígeno de superficie celular Thy-1  | 208851_s_at                            |
| Estromal-1 | CD357685                   |                     | TIMP2           | Locus transcrito, fuertemente similar a XP_511714.1 similar a precursor de inhibidor 2 de metaloproteinasas (TIMP-2) (inhibidor tisular de metaloproteinasas-2) (CSC-21 K) [ <i>Pan troglodytes</i> ] | 231579_s_at                            |

| Firma      | N.º de registro de GenBank | ID de gen de Entrez | Símbolo del gen | Título del gen  | ID de conjunto de sondas de Affymetrix |
|------------|----------------------------|---------------------|-----------------|---|--|
| Estromal-1 | NM_003247                  | 7058                | THBS2           | trombospondina 2  | 203083_at                              |
| Estromal-1 | NM_000088                  | 1277                | COL1A1          | colágeno, tipo I, alfa 1  | 1556499_s_at                           |
| Estromal-1 | NM_006832                  | 10979               | PLEKHC1         | miembro 1 que contiene dominio de homología a pleckstrina, familia C (con dominio FERM) | 209210_s_at                            |
| Estromal-1 | NM_021961                  | 7003                | TEAD1           | miembro 1 de familia de dominios TEA (factor de potenciador transcripcional de SV40)    | 224955_at                              |
| Estromal-1 | AK128814                   |                     |                 | ADNc FLJ25106 fis, clon CBR01467  | 213675_at                              |
| Estromal-1 | NM_153367                  | 219654              | C10orf56        | marco de lectura abierto 56 del cromosoma 10  | 212423_at                              |
| Estromal-1 | AK092048                   |                     |                 | ARNm; ADNc DKFZp313C0240 (del clon DKFZp313C0240)                                       | 227623_at                              |
| Estromal-1 | NM_005245                  | 2195                | FAT             | homólogo 1 de supresor de tumores FAT ( <i>Drosophila</i> )                             | 201579_at                              |
| Estromal-1 | NM_001129                  | 165                 | AEBP1           | proteína 1 de unión a AE  | 201792_at                              |
| Estromal-1 | NM_002403                  | 4237                | MFAP2           | proteína 2 asociada a microfibrillas  | 203417_at                              |
| Estromal-1 | NM_004342                  | 800                 | CALD1           | caldesmona 1  | 201616_s_at                            |
| Estromal-1 | NM_005576                  | 4016                | LOXL1           | proteína 1 similar a lisil oxidasa  | 203570_at                              |
| Estromal-1 | NM_199511                  | 151887              | CCDC80          | proteína 80 que contiene dominio superenrollado   | 225242_s_at                            |
| Estromal-1 | NM_012098                  | 23452               | ANGPTL2         | proteína 2 similar a angiopoyetina  | 213001_at                              |
| Estromal-1 | NM_002210                  | 3685                | ITGAV           | integrina, alfa V (receptor de vitronectina, polipéptido alfa, antígeno CD51)           | 202351_at                              |
| Estromal-1 | NM_000366                  | 7168                | TPM1            | tropomiosina 1 (alfa)   | 210986_s_at                            |
| Estromal-1 | NM_198474                  | 283298              | OLFML1          | proteína 1 similar a olfactomedina  | 217525_at                              |
| Estromal-1 | NM_001424                  | 2013                | EMP2            | proteína 2 de membrana epitelial  | 225078_at                              |
| Estromal-1 | NM_032575                  | 84662               | GLIS2           | dedo de cinc 2 de la familia GLIS   | 223378_at                              |
| Estromal-1 | NM_007173                  | 11098               | PRSS23          | proteasa, serina, 23  | 226279_at                              |
| Estromal-1 | NM_001015880               | 9060                | PAPSS2          | 3'-fosfoadenosina 5'-fosfosulfato sintasa 2   | 203060_s_at                            |
| Estromal-1 | NM_015645                  | 114902              | C1QTNF5         | proteína 5 relacionada con C1q y factor de necrosis tumoral                             | 223499_at                              |
| Estromal-1 | AK130049                   |                     |                 | ADNc FLJ26539 fis, clon KDN09310  | 213429_at                              |
| Estromal-1 | NM_001849                  | 1292                | COL6A2          | colágeno, tipo VI, alfa 2   | 209156_s_at                            |
| Estromal-1 | NM_001014796               | 4921                | DDR2            | familia de receptor de dominio de discoidina, miembro 2                                 | 225442_at                              |
| Estromal-1 | NM_015463                  | 25927               | C2orf32         | marco de lectura abierto 32 del cromosoma 2   | 226751_at                              |
| Estromal-1 | AK055628                   |                     | ADAM12          | ADNc FLJ31066 fis, clon HSYRA2001153  | 226777_at                              |
| Estromal-1 | NM_014799                  | 9843                | HEPH            | hefaestina  | 203903_s_at                            |
| Estromal-1 | NM_004385                  | 1462                | CSPG2           | condroitin sulfato proteoglicano 2 (versicano)  | 221731_x_at                            |
| Estromal-1 | NM_152330                  | 122786              | FRMD6           | proteína 6 que contiene dominio FERM 6  | 225481_at                              |
| Estromal-1 | BQ917964                   |                     | PPP4R2          | Locus transcrito  | 235733_at                              |

| Firma      | N.º de registro de GenBank | ID de gen de Entrez | Símbolo del gen | Título del gen  | ID de conjunto de sondas de Affymetrix |
|------------|----------------------------|---------------------|-----------------|---|--|
| Estromal-1 | NM_002615                  | 5176                | SERPINF1        | inhibidor de serpina peptidasa, clado F (alfa-2 antiplasmina, factor derivado de epitelio pigmentario), miembro 1 | 202283_at                              |
| Estromal-1 | NM_032348                  | 54587               | MXRA8           | proteína 8 asociada a remodelación de la matriz   | 213422_s_at                            |
| Estromal-1 | NM_006106                  | 10413               | YAP1            | proteína 1 asociada a Yes, 65kDa  | 224894_at                              |
| Estromal-1 | NM_020182                  | 56937               | TMEPAI          | ARN inducido por andrógenos de próstata, transmembrana  | 222449_at                              |
| Estromal-1 | CB999028                   |                     |                 | Locus transcrito  | 226834_at                              |
| Estromal-1 | NM_001711                  | 633                 | BGN             | biglicano   | 201261_x_at                            |
| Estromal-1 | NM_006902                  | 5396                | PRRX1           | caja homeótica 1 relacionada con paired   | 226695_at                              |
| Estromal-1 | NM_000428                  | 4053                | LTBP2           | proteína 2 de unión a factor de crecimiento transformante beta latente  | 204682_at                              |
| Estromal-1 | NM_004369                  | 1293                | COL6A3          | colágeno, tipo VI, alfa 3   | 201438_at                              |
| Estromal-1 | NM_000393                  | 1290                | COL5A2          | colágeno, tipo V, alfa 2  | 221730_at                              |
| Estromal-1 | NM_015419                  | 25878               | MXRA5           | proteína 5 asociada a remodelación de la matriz   | 209596_at                              |
| Estromal-1 | NM_001102                  | 87                  | ACTN1           | actinina, alfa 1  | 208637_x_at                            |
| Estromal-1 | NM_000877                  | 3554                | IL1R1           | receptor de interleucina 1, tipo I  | 202948_at                              |
| Estromal-1 | NM_015927                  | 7041                | TGFB111         | transcrito 1 inducido por factor de crecimiento transformante beta 1  | 209651_at                              |
| Estromal-1 | NM_032772                  | 84858               | ZNF503          | proteína de dedos de cinc 503   | 227195_at                              |
| Estromal-1 | NM_020440                  | 5738                | PTGFRN          | regulador negativo de receptor de prostaglandina F2   | 224937_at                              |
| Estromal-1 | NM_000138                  | 2200                | FBN1            | fibrilina 1   | 202765_s_at                            |
| Estromal-1 | NM_031442                  | 83604               | TMEM47          | proteína transmembrana 47   | 209656_s_at                            |
| Estromal-1 | NM_001734                  | 716                 | C1S             | componente 1 del complemento, subcomponente s   | 208747_s_at                            |
| Estromal-1 | NM_002290                  | 3910                | LAMA4           | laminina, alfa 4  | 202202_s_at                            |
| Estromal-1 | CN312045                   |                     | PPP4R2          | Locus transcrito, débilmente similar a NP_001013658.1 proteína LOC387873 [ <i>Homo sapiens</i> ]                  | 222288_at                              |
| Estromal-1 | NM_000089                  | 1278                | COL1A2          | colágeno, tipo I, alfa 2  | 202403_s_at                            |
| Estromal-1 | NM_004530                  | 4313                | MMP2            | metalopeptidasa de la matriz 2 (gelatinasa A, gelatinasa de 72 kDa, colagenasa de 72 kDa tipo IV)                 | 201069_at                              |
| Estromal-1 | NM_001387                  | 1809                | DPYSL3          | proteína 3 similar a dihidropirimidinas   | 201431_s_at                            |
| Estromal-1 | NM_138389                  | 92689               | FAM114A1        | familia con similitud de secuencia 114, miembro A1  | 213455_at                              |
| Estromal-1 | NM_006670                  | 7162                | TPBG            | glicoproteína de trofoblastos   | 203476_at                              |
| Estromal-1 | NM_000304                  | 5376                | PMP22           | proteína 22 de mielina periférica   | 210139_s_at                            |
| Estromal-1 | NM_002775                  | 5654                | HTRA1           | HtrA serina peptidasa 1   | 201185_at                              |
| Estromal-1 | NM_002593                  | 5118                | PCOLCE          | potenciador de procolágeno C-endopeptidasa  | 202465_at                              |
| Estromal-1 | NM_003118                  | 6678                | SPARC           | proteína secretada, ácida, rica en cisteína   | 212667_at                              |

| Firma      | N.º de registro de GenBank | ID de gen de Entrez | Símbolo del gen | Título del gen   | ID de conjunto de sondas de Affymetrix |
|------------|----------------------------|---------------------|-----------------|--|--|
|            |                            |                     |                 | (osteonectina)   |  |
| Estromal-1 | NM_007085                  | 11167               | FSTL1           | proteína 1 similar a folistatina   | 208782_at                              |
| Estromal-1 | NM_001080393               | 727936              |                 | proteína 4 que contiene dominio de glicosil-transferasa 8 predicho   | 235371_at                              |
| Estromal-1 | NM_018153                  | 84168               | ANTXR1          | receptor 1 de toxina antrácica   | 224694_at                              |
| Estromal-1 | NM_001733                  | 715                 | C1R             | componente 1 del complemento, subcomponente r  | 212067_s_at                            |
| Estromal-1 | NM_001797                  | 1009                | CDH11           | cadherina 11, tipo 2, OB-cadherina (osteoblasto)   | 207173_x_at                            |
| Estromal-1 | NM_016938                  | 30008               | EFEMP2          | proteína 2 de la matriz extracelular similar a fibulina que contiene EGF   | 209356_x_at                            |
| Estromal-2 | NM_014601                  | 30846               | EHD2            | proteína 2 que contiene dominio EH   | 45297_at                               |
| Estromal-2 | NM_017789                  | 54910               | SEMA4C          | dominio sema, dominio de inmunoglobulina (Ig), dominio transmembrana (TM) y dominio citoplasmático corto (semaforina) 4C | 46665_at                               |
| Estromal-2 | NM_000484                  | 351                 | APP             | proteína precursora de amiloide beta (A4) (peptidasa nexina-II, enfermedad de Alzheimer)                                 | 200602_at                              |
| Estromal-2 | NM_004684                  | 8404                | SPARCL1         | proteína 1 similar a SPARC (mast9, hevina)   | 200795_at                              |
| Estromal-2 | NM_002291                  | 3912                | LAMB1           | laminina, beta 1   | 201505_at                              |
| Estromal-2 | NM_000210                  | 3655                | ITGA6           | integrina, alfa 6  | 201656_at                              |
| Estromal-2 | NM_000552                  | 7450                | VWF             | factor de von Willebrand   | 202112_at                              |
| Estromal-2 | NM_001233                  | 858                 | CAV2            | caveolina 2  | 203323_at                              |
| Estromal-2 | NM_006404                  | 10544               | PROCR           | receptor de proteína C, endotelial (EP.CR)   | 203650_at                              |
| Estromal-2 | NM_000609                  | 6387                | CXCL12          | ligando 12 de quimiocina (motivo C-X-C) (factor 1 derivado de células del estroma)                                       | 203666_at                              |
| Estromal-2 | NM_002253                  | 3791                | KDR             | receptor de dominio de inserto cinasa (una tirosina cinasa receptor de tipo III)   | 203934_at                              |
| Estromal-2 | NM_001442                  | 2167                | FABP4           | proteína 4 de unión a ácidos grasos, adipocito   | 203980_at                              |
| Estromal-2 | NM_016315                  | 51454               | GULP1           | GULP, proteína 1 que contiene dominio PTB adaptador de internamiento   | 204237_at                              |
| Estromal-2 | NM_006307                  | 8406                | SRPX            | proteína que contiene repetición sushi, ligada al cromosoma X  | 204955_at                              |
| Estromal-2 | NM_000163                  | 2690                | GHR             | receptor de hormona de crecimiento   | 205498_at                              |
| Estromal-2 | NM_000950                  | 5638                | PRRG1           | Gla (ácido G-carboxiglutámico) 1 rico en prolina   | 205618_at                              |
| Estromal-2 | NM_002666                  | 5346                | PLIN            | perlipina  | 205913_at                              |
| Estromal-2 | NM_000459                  | 7010                | TEK             | TEK tirosina cinasa, endotelial (malformaciones venosas, cutáneas y mucosas múltiples)                                   | 206702_at                              |
| Estromal-2 | NM_004797                  | 9370                | ADIPOQ          | proteína que contiene dominio adiponectina, C1Q y colágeno   | 207175_at                              |

| Firma      | N.º de registro de GenBank | ID de gen de Entrez | Símbolo del gen | Título del gen   | ID de conjunto de sondas de Affymetrix |
|------------|----------------------------|---------------------|-----------------|--|--|
| Estromal-2 | NM_000442                  | 5175                | PECAM1          | molécula de adhesión a células endoteliales/plaquetas (antígeno CD31)              | 208981_at                              |
| Estromal-2 | NM_198098                  | 358                 | AQP1            | acuaporina 1 (grupo sanguíneo Colton)  | 209047_at                              |
| Estromal-2 | NM_021005                  | 7026                | NR2F2           | subfamilia 2 de receptores nucleares, grupo F, miembro 2                           | 209120_at                              |
| Estromal-2 | NM_014220                  | 4071                | TM4SF1          | miembro 1 de la familia transmembrana 4 L seis                                     | 209386_at                              |
| Estromal-2 | NM_001001549               | 2887                | GRB10           | proteína 10 unida a receptor de factor de crecimiento                              | 209409_at                              |
| Estromal-2 | NM_006108                  | 10418               | SPON1           | espondina 1, proteína de la matriz extracelular                                    | 209436_at                              |
| Estromal-2 | NM_001003679               | 3953                | LEPR            | receptor de leptina  | 209894_at                              |
| Estromal-2 | NM_000599                  | 3488                | IGFBP5          | proteína 5 de unión a factor de crecimiento de tipo insulina                       | 211959_at                              |
| Estromal-2 | NM_001753                  | 857                 | CAV1            | caveolina 1, proteína de caveolas, 22 kDa  | 212097_at                              |
| Estromal-2 | NM_005841                  | 10252               | SPRY1           | homólogo 1 de sprouty, antagonista de la señalización de FGF ( <i>Drosophila</i> ) | 212558_at                              |
| Estromal-2 | NM_015345                  | 23500               | DAAM2           | activador asociado a dishevelled de la morfogénesis 2                              | 212793_at                              |
| Estromal-2 | NM_015234                  | 221395              | GPR116          | receptor 116 acoplado a proteínas G  | 212950_at                              |
| Estromal-2 | NM_006108                  | 10418               | SPON1           | espondina 1, proteína de la matriz extracelular                                    | 213993_at                              |
| Estromal-2 | NM_016215                  | 51162               | EGFL7           | dominio similar a EGF, múltiple 7  | 218825_at                              |
| Estromal-2 | NM_022481                  | 64411               | CENTD3          | centaurina, delta 3  | 218950_at                              |
| Estromal-2 | XM_371262                  | 64123               | ELTD1           | proteína 1 que contiene EGF, latrofilina y siete dominios transmembrana            | 219134_at                              |
| Estromal-2 | NM_016563                  | 51285               | RASL12          | familia 12, similar a RAS  | 219167_at                              |
| Estromal-2 | NM_006094                  | 10395               | DLC1            | proteína 1 delecionada en cáncer 1   | 224822_at                              |
| Estromal-2 | NM_019035                  | 54510               | PCDH18          | protocadherina 18  | 225975_at                              |
| Estromal-2 | NM_019055                  | 54538               | ROBO4           | homólogo 4 de roundabout, magic roundabout ( <i>Drosophila</i> )                   | 226028_at                              |
| Estromal-2 | NM_002207                  | 3680                | ITGA9           | integrina, alfa 9  | 227297_at                              |
| Estromal-2 | XM_93060B                  | 641700              | ECSM2           | molécula 2 específica de células endoteliales                                      | 227779_at                              |
| Estromal-2 | XM_037493                  | 85358               | SHANK3          | dominios 3 de SH3 y repeticiones de anquirina múltiples                            | 227923_at                              |
| Estromal-2 | NM_052954                  | 116159              | CYYR1           | proteína 1 rica en cisteína/tirosina   | 228665_at                              |
| Estromal-2 | NM_002837                  | 5787                | PTPRB           | proteína tirosina fosfatasa, tipo receptor, B                                      | 230250_at                              |
| Estromal-2 | NM_019558                  | 3234                | HOXD8           | caja homeótica D8  | 231906_at                              |
| Estromal-2 | NM_001442                  | 2167                | FABP4           | proteína 4 de unión a ácidos grasos, adipocito                                     | 235978_at                              |
| Estromal-2 | NM_024756                  | 79812               | MMRN2           | multimerina 2  | 236262_at                              |
| Estromal-2 | BQ897248                   |                     |                 | Locus transcrito   | 242680_at                              |
| Estromal-2 | NM_020663                  | 57381               | RHOJ            | familia génica homóloga a  | 243481_at                              |

| Firma      | N.º de registro de GenBank | ID de gen de Entrez | Símbolo del gen | Título del gen  | ID de conjunto de sondas de Affymetrix |
|------------|----------------------------|---------------------|-----------------|---|--|
|            |                            |                     |                 | ras, miembro J  |  |
| Estromal-2 | AK091419                   |                     |                 | ADNc FLJ34100 fis, clon FCBBF3007597  | 1558397_at                             |
| Estromal-2 | NM_015719                  | 50509               | COL5A3          | colágeno, tipo V, alfa 3  | 52255_s_at                             |
| Estromal-2 | NM_012072                  | 22918               | CD93            | molécula CD93   | 202878_s_at                            |
| Estromal-2 | NM_000300                  | 5320                | PLA2G2A         | fosfolipasa A2, grupo IIA (plaquetas, líquido sinovial)   | 203649_s_at                            |
| Estromal-2 | NM_019105                  | 7148                | TNXB            | tenascina XB  | 206093_x_at                            |
| Estromal-2 | NM_030754                  | 6289                | SAA2            | amilioide sérico A2   | 208607_s_at                            |
| Estromal-2 | NM_019105                  | 7148                | TNXB            | tenascina XB  | 208609_s_at                            |
| Estromal-2 | NM_014220                  | 4071                | TM4SF1          | miembro 1 de familia transmembrana 4 L seis   | 209367_s_at                            |
| Estromal-2 | NM_000668                  | 125                 | ADH1 B          | alcohol deshidrogenasa IB (clase I), polipéptido beta   | 209612_s_at                            |
| Estromal-2 | NM_000668                  | 125                 | ADH1B           | alcohol deshidrogenasa IB (clase I), polipéptido beta   | 209613_s_at                            |
| Estromal-2 | NM_001354                  | 1646                | AKR1C2          | familia 1 de aldo-ceto reductasa, miembro C2 (dihidrodiol-deshidrogenasa 2; proteína de unión a ácidos biliares; 3-alfa hidroxisteroide deshidrogenasa, tipo III) | 209699_x_at                            |
| Estromal-2 | NM_001032281               | 7035                | TFPI            | inhibidor de la ruta de factor tisular (inhibidor de la coagulación asociado a lipoproteínas)   | 210664_s_at                            |
| Estromal-2 | NM_001001924               | 57509               | MTUS1           | supresor 1 de tumores mitocondriales  | 212096_s_at                            |
| Estromal-2 | NM_019105                  | 7148                | TNXB            | tenascina XB  | 213451_x_at                            |
| Estromal-2 | NM_004449                  | 2078                | ERG             | homólogo (aviar) de oncogén E26 de virus de eritroblastosis v-ets   | 213541_s_at                            |
| Estromal-2 | NM_018407                  | 55353               | LAPTM4B         | proteína transmembrana 4 beta asociada al lisosoma  | 214039_s_at                            |
| Estromal-2 | NM_000331                  | 6288                | SAA1            | amilioide sérico A1   | 214456_x_at                            |
| Estromal-2 | NM_019105                  | 7148                | TNXB            | tenascina XB  | 216333_x_at                            |
| Estromal-2 | NM_001034954               | 10580               | SORBS1          | proteína 1 que contiene dominio sorbina y SH3   | 218087_s_at                            |
| Estromal-2 | NM_017734                  | 54873               | PALMD           | palmdelfina   | 218736_s_at                            |
| Estromal-2 | NM_024756                  | 79812               | MMRN2           | multimerina 2   | 219091_s_at                            |
| Estromal-2 | NM_006744                  | 5950                | RBP4            | proteína 4 de unión a retinol, plasma   | 219140_s_at                            |
| Estromal-2 | NM_001034954               | 10580               | SORBS1          | proteína 1 que contiene dominio sorbina y SH3   | 222513_s_at                            |

Se generaron factores de predicción de supervivencia de DLBCL de la invención usando datos de expresión y los métodos descritos en los ejemplos 1 y 2, a continuación. El primer factor de predicción de supervivencia de dos variables incorpora las firmas de expresión génica de GCB y estromal-1. El ajuste del modelo de riesgos proporcionales de Cox a los datos de expresión génica obtenidos a partir de estas dos firmas dio como resultado una puntuación de factor de predicción de supervivencia de modelo de dos variables calculada usando la siguiente ecuación generalizada:

$$5 \quad \text{Puntuación de factor de predicción de supervivencia de DLBCL de dos variables} = A - [(x) \cdot (\text{valor de firma de GCB})] - [(y) \cdot (\text{valor de firma de estromal-1})]$$

10 En esta ecuación, A es un término de desviación, mientras que (x) e (y) son factores de escala. El valor de firma de GCB y el valor de firma de estromal-1 pueden corresponder al promedio de los niveles de expresión de todos los genes en la firma de GCB y la firma de estromal-1, respectivamente. Una puntuación inferior de factor de predicción de supervivencia indica un resultado de supervivencia más favorable, y una puntuación superior de factor de predicción de supervivencia indica un resultado de supervivencia menos favorable para el sujeto.



El factor de predicción de supervivencia de dos variables se refinó para dar un factor de predicción de supervivencia de múltiples variables que incorpora las firmas de expresión génica de GCB, estromal-1 y estromal-2. El ajuste del modelo de riesgos proporcionales de Cox a los datos de expresión génica obtenidos a partir de estas tres firmas dio como resultado una puntuación de factor de predicción de supervivencia de modelo de múltiples variables calculado usando la siguiente ecuación generalizada:

Puntuación de factor de predicción de supervivencia de DLBCL de múltiples variables general =  $A - [(x) \cdot (\text{valor de firma de GCB})] - [(y) \cdot (\text{valor de firma de estromal-1})] + [(z) \cdot (\text{valor de firma de estromal-2})]$ .

En esta ecuación, A es un término de desviación, mientras que (x), (y) y (z) son factores de escala. El valor de firma de GCB, el valor de firma de estromal-1 y el valor de firma de estromal-2 pueden corresponder al promedio de los niveles de expresión de todos los genes en la firma de GCB, la firma de estromal-1 y la firma de estromal-2, respectivamente. Una puntuación inferior de factor de predicción de supervivencia indica un resultado de supervivencia más favorable y una puntuación superior de factor de predicción de supervivencia indica un resultado de supervivencia menos favorable para el sujeto.

En una realización, la invención proporciona la siguiente ecuación de factor de predicción de supervivencia de múltiples variables:

Puntuación de factor de predicción de supervivencia de DLBCL de múltiples variables =  $8,11 - [0,419 \cdot (\text{valor de firma de GCB})] - [1,015 \cdot (\text{valor de firma de estromal-1})] + [0,675 \cdot (\text{valor de firma de estromal-2})]$ .

En esta ecuación, una puntuación inferior de factor de predicción de supervivencia indica un resultado de supervivencia más favorable, y una puntuación superior de factor de predicción de supervivencia indica un peor resultado de supervivencia para el sujeto.

En otras realizaciones de la ecuación de puntuación de factor de predicción de supervivencia de DLBCL de múltiples variables, el término de desviación (A) o (8,11) puede variarse sin afectar a la utilidad de la ecuación en la predicción del resultado clínico. También pueden variarse los factores de escala (x), (y) y (z), individualmente o en combinación. Por ejemplo, el factor de escala (x) puede ser de desde aproximadamente 0,200 o más, desde aproximadamente 0,225 o más, desde aproximadamente 0,250 o más, desde aproximadamente 0,275 o más, desde aproximadamente 0,300, desde aproximadamente 0,325 o más, desde aproximadamente 0,350 o más, desde aproximadamente 0,375 o más, o desde aproximadamente 0,400 o más. Alternativamente, o además, el factor de escala (x) puede ser de aproximadamente 0,625 o menos, aproximadamente 0,600 o menos, aproximadamente 0,575 o menos, aproximadamente 0,550 o menos, aproximadamente 0,525 o menos, aproximadamente 0,500 o menos, aproximadamente 0,475 o menos, aproximadamente 0,450 o menos, o aproximadamente 0,425 o menos. Por tanto, el factor de escala (z) puede ser uno que está limitado por dos cualesquiera de los puntos de extremo anteriores. Por ejemplo, el factor de escala (x) puede ser un valor de desde 0,200-0,625, desde 0,350-0,550, desde 0,350-0,475 o desde 0,400-0,425. De manera similar, el factor de escala (y) puede ser de desde aproximadamente 0,800 o más, desde aproximadamente 0,825 o más, desde aproximadamente 0,850 o más, desde aproximadamente 0,875 o más, desde aproximadamente 0,900 o más, desde aproximadamente 0,925 o más, desde aproximadamente 0,950 o más, desde aproximadamente 0,975 o más, o desde aproximadamente 1,000 o más. Alternativamente, o además, el factor de escala (y) puede ser, por ejemplo, de aproximadamente 1,250 o menos, por ejemplo, aproximadamente 1,225 o menos, aproximadamente 1,200, aproximadamente 1,175 o menos, aproximadamente 1,150 o menos, aproximadamente 1,125 o menos, aproximadamente 1,100 o menos, aproximadamente 1,075 o menos, aproximadamente 1,050 o menos, o aproximadamente 1,025 o menos. Por tanto, el factor de escala (y) puede ser uno que está limitado por dos cualesquiera de los puntos de extremo anteriores. Por ejemplo, el factor de escala (y) puede ser un valor de desde 0,800-1,250, un valor de desde 0,950-1,1025, un valor de desde 0,950-1,200 o un valor de desde 1,000-1,025. También de manera similar, el factor de escala (z) puede ser de desde aproximadamente 0,450 o más, aproximadamente 0,475 o más, aproximadamente 0,500 o más, aproximadamente 0,525 o más, aproximadamente 0,550 o más, aproximadamente 0,575 o más, aproximadamente 0,600 o más, aproximadamente 0,625 o más, o aproximadamente 0,650 o más. Alternativamente, o además, el factor de escala (z) puede ser, por ejemplo, de aproximadamente 0,900 o menos, por ejemplo, aproximadamente 0,875 o menos, aproximadamente 0,850, aproximadamente 0,825 o menos, aproximadamente 0,800 o menos, aproximadamente 0,775 o menos, aproximadamente 0,750 o menos, o aproximadamente 0,725 o menos. Por tanto, el factor de escala (z) puede ser uno que está limitado por dos cualesquiera de los puntos de extremo anteriores. Por ejemplo, el factor de escala (z) puede ser un valor de desde 0,450-0,900, cualquier valor de desde 0,650-0,725, cualquier valor de desde 0,625-0,775 o cualquier valor de desde 0,650-0,700.

Además, la invención incluye cualquier conjunto de factores de escala (x), (y) y (z) conjuntamente en la puntuación de factor de predicción de supervivencia de DLBCL de múltiples variables general que crea una función que está relacionada de manera monótona con una ecuación de puntuación de factor de predicción de supervivencia de DLBCL de múltiples variables que usa cualquier combinación de los valores de factor de escala (x), (y) y (z) especificados anteriormente.

En algunas realizaciones de la invención, puede calcularse una puntuación de factor de predicción de supervivencia usando menos de la totalidad de los componentes génicos de la firma de GCB, la firma de estromal-1 y/o la firma de

estromal-2 enumerados en la tabla 1. Por ejemplo, las ecuaciones de predicción de supervivencia dadas a conocer en el presente documento pueden calcularse usando combinaciones matemáticas de las expresiones del 98% (38), el 95% (37), el 93% (36) o el 90% (35) de los genes enumerados en la tabla 1 para la firma de GCB, aproximadamente el 99% (aproximadamente 280), aproximadamente el 98% (aproximadamente 277), el 97% (aproximadamente 275), aproximadamente el 96% (aproximadamente 272), aproximadamente el 95% (aproximadamente 270), aproximadamente el 94% (aproximadamente 266), aproximadamente el 93% (aproximadamente 263), aproximadamente el 92% (aproximadamente 260), aproximadamente el 91% (aproximadamente 257) o aproximadamente el 90% (aproximadamente 255) de los genes enumerados en la tabla 1 para la firma de estromal-1, y/o el 99% (71), el 97% (70), el 96% (69), el 95% (68) el 93% (67), el 92% (66) o el 90% (65) de los genes enumerados en la tabla 1 para la firma de estromal-2 (en lugar de usar todos los genes correspondientes a una firma génica en la tabla 1 para calcular el valor de firma de GCB, el valor de firma de estromal-1 y/o el valor de firma de estromal-2, respectivamente). En otras realizaciones, las ecuaciones de predicción de supervivencia dadas a conocer en el presente documento pueden calcularse usando combinaciones matemáticas de las expresiones del 88% (34 genes), el 85% (33 genes), el 82% (32 genes), el 80% (31 genes) de los genes enumerados en la tabla 1 para la firma de GCB, aproximadamente el 89% (aproximadamente 252), aproximadamente el 88% (aproximadamente 249), aproximadamente el 87% (aproximadamente 246), aproximadamente el 86% (aproximadamente 243), aproximadamente el 85% (aproximadamente 241), aproximadamente el 84% (aproximadamente 238), aproximadamente el 83% (aproximadamente 235), aproximadamente el 82% (aproximadamente 232), aproximadamente el 81% (aproximadamente 229) o aproximadamente el 80% (aproximadamente 226) de los genes enumerados en la tabla 1 para la firma de estromal-1, y/o el 89% (64), el 88% (63), el 86% (62), el 85% (61), el 83% (60), el 82% (59) o el 80% (58) de los genes enumerados en la tabla 1 para la firma de estromal-2 (en lugar de usar todos los genes correspondientes a una firma de gen en la tabla 1 para calcular el valor de firma de GCB, el valor de firma de estromal-1 y/o el valor de firma de estromal-2, respectivamente).

La invención también proporciona un método para usar una puntuación de factor de predicción de supervivencia de DLBCL para predecir la probabilidad de un resultado de supervivencia más allá de una cantidad de tiempo t tras el tratamiento para DLBCL. El método incluye calcular la probabilidad de un resultado de supervivencia para un sujeto usando la siguiente ecuación general:

$$P(SO) = SO_0(t)^{\text{(exp((S)*(puntuación de factor de predicción de supervivencia))}}$$

En esta ecuación, P(SO) es la probabilidad del resultado de supervivencia del sujeto más allá del tiempo t tras el tratamiento para DLBCL,  $SO_0(t)$  es la probabilidad de resultado de supervivencia, que corresponde al valor de tiempo más grande menor que t en una curva de resultado de supervivencia, y (s) es un factor de escala. El tratamiento para DLBCL puede incluir quimioterapia y la administración de rituximab. Puede calcularse una curva de supervivencia usando métodos estadísticos, tales como el modelo de riesgos proporcionales de Cox. Se expone información adicional referente a curvas de resultado de supervivencia en Lawless, *Statistical Models and Methods for Lifetime Data*, John Wiley and Sons (Nueva York 1982) y Kalbfleisch *et al.*, *Biometrika*, 60: 267-79 (1973).

En una realización, el método de la invención incluye calcular la probabilidad de supervivencia global para un sujeto más allá de una cantidad de tiempo t tras el tratamiento para DLBCL. El método incluye calcular la probabilidad de un resultado de supervivencia para un sujeto usando la siguiente ecuación general:

$$P(SO) = SO_0(t)^{\text{(exp(puntuación de factor de predicción de supervivencia))}}$$

En la ecuación, P(OS) es la probabilidad de supervivencia global del sujeto más allá del tiempo t tras el tratamiento para DLBCL,  $SO_0(t)$  es la probabilidad de resultado de supervivencia de curva, que corresponde al valor de tiempo más largo en una curva de supervivencia que es menor que t, y el factor de escala de la ecuación general (s) = 1. El tratamiento para DLBCL puede incluir quimioterapia sola o en combinación con la administración de rituximab (R-CHOP).

En otra realización, el método de la invención incluye calcular la probabilidad de supervivencia libre de progresión para un sujeto más allá de una cantidad de tiempo t tras el tratamiento para DLBCL. El método incluye calcular la probabilidad de un resultado de supervivencia para un sujeto usando la siguiente ecuación general:

$$P(PFS) = SO_0(t)^{\text{(exp(0,976*(puntuación de factor de predicción de supervivencia))}}$$

En esta ecuación, P(PFS) es la probabilidad de supervivencia libre de progresión del sujeto más allá del tiempo t tras el tratamiento para DLBCL,  $SO_0(t)$  es la probabilidad de resultado de supervivencia libre de progresión de curva, que corresponde al valor de tiempo más largo en una curva de supervivencia que es menor que t, y el factor de escala de la ecuación general (s) = 0,976. El tratamiento para DLBCL puede incluir quimioterapia sola o en combinación con la administración de rituximab (R-CHOP).

Las ecuaciones anteriores para P(OS) y P(PFS) se generaron maximizando las probabilidades parciales del modelo de riesgos proporcionales de Cox dentro de los datos de LLMP CHOP descritos a continuación en los ejemplos 1 y 2. Se consideraron modelos de riesgos proporcionales de Cox variables individuales separados para la supervivencia global P(OS) y para la supervivencia libre de progresión P(PFS) basándose en esta formulación de

puntuación del modelo. Se generaron factores de escala variables individuales (1,0 para supervivencia global y 0,997 para supervivencia libre de progresión) para cada modelo mediante la maximización de las probabilidades parciales dentro de los pacientes con R-CHOP descritos a continuación en los ejemplos 1 y 2.

5 En otras realizaciones, el factor de escala en la P(PFS) anterior puede variarse de manera que (en lugar de 0,976) el factor de escala (s) sea un valor entre 0,970 y 0,980, por ejemplo 0,971, 0,972, 0,973, 0,973, 0,974, 0,975, 0,977, 0,978 y 0,979.

10 La invención proporciona además un alineamiento seleccionado como diana que puede usarse para detectar los niveles de expresión de todos o casi todos los genes en una firma de expresión génica de células B de centro germinal (GCB), una firma de expresión génica de estromal-1 y/o una firma de expresión génica de estromal-2. Un alineamiento seleccionado como diana, tal como se usa en el presente documento, es un alineamiento dirigido a un conjunto limitado de genes y por tanto difiere de un alineamiento de genoma completo. El alineamiento seleccionado como diana de la invención puede incluir sondas para menos de 10.000 genes, menos de 8.000 genes, menos de 7.000 genes, menos de 6.000 genes, menos de 5.000 genes o menos de 4.000 genes. Generalmente, el  
15 alineamiento seleccionado como diana incluye sondas para al menos el 80% de los genes en una firma de expresión génica de células B de centro germinal (GCB), una firma de expresión génica de estromal-1 y/o una firma de expresión génica de estromal-2. Los alineamientos seleccionados como diana de la invención pueden usarse, por ejemplo, para detectar niveles de expresión para su uso en los métodos descritos en el presente documento.

20 La invención proporciona un alineamiento seleccionado como diana que incluye sondas para todos los genes en la firma de expresión génica de estromal-1. La invención también proporciona un alineamiento seleccionado como diana que incluye sondas para todos los genes en la firma de expresión génica de estromal-2. Adicionalmente, la invención proporciona un alineamiento seleccionado como diana que incluye sondas para todos los genes en la firma de expresión génica de estromal-1 y todos los genes en la firma de expresión génica de estromal-2. Además, la invención proporciona un alineamiento seleccionado como diana que incluye sondas para todos los genes en la firma de expresión génica de estromal-1, todos los genes en la firma de expresión génica de estromal-2 y todos los  
25 genes en la firma de GCB.

30 En determinadas realizaciones, los alineamientos de la invención pueden incluir el 98% (38), el 95% (37), el 93% (36) o el 90% (35) de los genes enumerados en la tabla 1 para la firma de GCB, aproximadamente el 99% (aproximadamente 280), aproximadamente el 98% (aproximadamente 277), el 97% (aproximadamente 275), aproximadamente el 96% (aproximadamente 272), aproximadamente el 95% (aproximadamente 270), aproximadamente el 94% (aproximadamente 266), aproximadamente el 93% (aproximadamente 263), aproximadamente el 92% (aproximadamente 260), aproximadamente el 91% (aproximadamente 257), o aproximadamente el 90% (aproximadamente 255) de los genes enumerados en la tabla 1 para la firma de estromal-1, y/o el 99% (71), el 97% (70), el 96% (69), el 95% (68) el 93% (67), el 92% (66) o el 90% (65) de los genes  
35 enumerados en la tabla 1 para la firma de estromal-2 (en lugar de todos los genes enumerados en la tabla 1 para el promedio de firma de GCB, el promedio de firma de estromal-1 y/o el promedio de firma de estromal-2, respectivamente). En determinadas realizaciones, los alineamientos de la invención pueden incluir el 88% (34 genes), el 85% (33 genes), el 82% (32 genes), el 80% (31 genes) de los genes enumerados en la tabla 1 para la firma de GCB, aproximadamente el 89% (aproximadamente 252), aproximadamente el 88% (aproximadamente 249), aproximadamente el 87% (aproximadamente 246), aproximadamente el 86% (aproximadamente 243), aproximadamente el 85% (aproximadamente 241), aproximadamente el 84% (aproximadamente 238), aproximadamente el 83% (aproximadamente 235), aproximadamente el 82% (aproximadamente 232), aproximadamente el 81% (aproximadamente 229) o aproximadamente el 80% (aproximadamente 226) de los genes  
40 enumerados en la tabla 1 para la firma de estromal-1, y/o el 89% (64), el 88% (63), el 86% (62), el 85% (61), el 83% (60), el 82% (59) o el 80% (58) de los genes enumerados en la tabla 1 para la firma de estromal-2 (en lugar de todos los genes enumerados en la tabla 1 para el promedio de firma de GCB, el promedio de firma de estromal-1 y/o el promedio de firma de estromal-2, respectivamente).  
45

Los siguientes ejemplos ilustran adicionalmente la invención.

#### EJEMPLO 1

50 Este ejemplo demuestra que se encontraron diferencias significativas entre los resultados de supervivencia para pacientes con DLBCL ABC y DLBCL GCB tratados con R-CHOP y que el resultado de supervivencia se correlacionaba con tres firmas de expresión génica de pronóstico.

55 Se obtuvieron datos clínicos y muestras de biopsias tumorales antes del tratamiento de 414 pacientes con DLBCL *de novo* tratados en 10 instituciones en América del Norte y Europa y se estudiaron según un protocolo aprobado por la Junta de Revisión Institucional del National Cancer Institute. Los pacientes incluidos en una "cohorte de LLMP CHOP" de 181 pacientes se trataron con combinaciones a base de antraciclina, lo más a menudo ciclofosfamida, doxorubicina, vincristina y prednisona (CHOP) o regímenes similares, tal como se describió anteriormente (Rosenwald *et al.*, N. Engl. J. Med., 346: 1937-47 (2002)). Los 233 pacientes restantes constituyeron una cohorte de R-CHOP que recibió quimioterapia similar más rituximab. La mediana del seguimiento en la cohorte de R-CHOP fue de 2,1 años (2,8 años para los supervivientes). Un panel de hematopatólogos expertos confirmó el diagnóstico de

DLBCL usando los criterios actuales de la OMS. En la tabla 2 se describen características clínicas adicionales de los pacientes para la cohorte de R-CHOP. Un análisis adicional usó una segunda cohorte de “MMMLNP CHOP” de 177 pacientes estudiados por el Molecular Mechanisms of Non-Hodgkin’s Lymphoma Network Project (Hummel *et al.*, N. Engl. J. Med., 354: 2419-30 (2006)).

5 Tabla 2: Características clínicas de pacientes con DLBCL tratados con R-CHOP

| Característica   | % del total (N=233) | % de DLBCL de tipo células B de centro germinal (N=107) | % de DLBCL de tipo células B activadas (N=93) | % de DLBCL no clasificado (N=33) | Valor de P |
|--|---------------------|---|---|----------------------------------|------------|
| Edad > 60 años   | 52                  | 47  | 63  | 39                               | 0,02       |
| Estadio de Ann Arbor > II                                      | 54                  | 48  | 62  | 50                               | 0,06       |
| Lactato deshidrogenasa > 1x Normal                             | 48                  | 43  | 58  | 41                               | 0,06       |
| N.º de sitios extraganglionares >1                             | 15                  | 14  | 15  | 14                               | 0,8        |
| Estado funcional del Eastern Cooperative Oncology Group (ECOG) | 25                  | 17  | 33  | 27                               | 0,02       |
| Puntuación del índice de pronóstico internacional (IPI)        |                     |   |   |                                  | <0,001     |
| 0 ó 1  | 41                  | 55  | 21  | 50                               |            |
| 2 ó 3  | 46                  | 33  | 63  | 38                               |            |
| 4 ó 5  | 13                  | 12  | 15  | 12                               |            |
| Puntuación de IPI revisada                                     |                     |   |   |                                  | <0,001     |
| 0  | 19                  | 27  | 5   | 28                               |            |
| 1 ó 2  | 56                  | 52  | 64  | 48                               |            |
| 3-5  | 25                  | 21  | 31  | 24                               |            |

Se realizó la obtención de perfiles de expresión génica usando micromatrices de Affymetrix U133+ 2.0. Los datos de la obtención de perfiles de expresión génica están disponibles a través del sitio web del National Center for Biotechnology Information tal como se describe en Lenz *et al.*, New Engl. J. Med, 359: 2313-23 (2008), en la página 2314. Se normalizaron todos los datos de alineamientos de expresión génica usando el software MAS 5.0, y se transformaron mediante log2. Para explicar las diferencias técnicas en el procesamiento de micromatrices entre los datos de la cohorte de R-CHOP y los datos de la cohorte de LLMPP CHOP, se ajustaron los valores de expresión de cada gen en los datos de la cohorte de R-CHOP de modo que su mediana coincidiera con la mediana de los datos de LLMPP CHOP.

La identificación de firmas de expresión génica y el desarrollo de un modelo de factores de predicción de supervivencia se basaron únicamente en los datos del conjunto de entrenamiento de LLMPP CHOP. No se realizaron análisis de subgrupos o análisis de supervivencia previos con los conjuntos de prueba (cohortes de MMMLNP CHOP y R-CHOP). Se usó el modelo de Cox para identificar genes asociados con supervivencia en el conjunto de entrenamiento de CHOP y para construir modelos de supervivencia de múltiples variables. Los modelos y sus coeficientes de escala asociados se fijaron basándose en el conjunto de entrenamiento de CHOP y se evaluaron entonces en los conjuntos de prueba. Se calcularon los valores de P de los efectos de supervivencia de variables continuas tales como expresión génica o expresión de firma con la prueba de razón de probabilidades de Cox. Se calculó la significación de los efectos de supervivencia basados variables diferenciadas tales como el subtipo de linfoma o el índice de pronóstico internacional (IPI) usando la prueba de rangos logarítmicos. Los valores de P de validación presentados son unilaterales en la dirección observada en el conjunto de entrenamiento. El resto de los otros valores de P eran bilaterales. Se estimaron curvas de supervivencia usando el método de Kaplan-Meier.

Todos los aspectos de la identificación de firmas de expresión génica y el desarrollo de un modelo de factores de predicción de supervivencia se basaron únicamente en los datos del conjunto de entrenamiento de CHOP. No se realizaron análisis de subgrupos o análisis de supervivencia previos con los conjuntos de prueba (cohortes de MMMLNP CHOP y R-CHOP). Se usó el modelo de Cox para identificar genes asociados con supervivencia en el conjunto de entrenamiento de CHOP y para construir modelos de supervivencia de múltiples variables. Los modelos y sus factores de escala asociados se fijaron basándose en el conjunto de entrenamiento de CHOP y se evaluaron entonces en los conjuntos de prueba.

Puesto que los subtipos de ABC y GCB DLBCL tienen distintas tasas de supervivencia global con quimioterapia con CHOP (Rosenwald *et al.*, N. Engl. J. Med., 346: 1937-47 (2002); Alizadeh *et al.*, Nature, 403:503-11(2000); Hummel *et al.*, N. Engl. J. Med., 354:2419-30 (2006); Monti, Blood, 105:1851-61(2005)), se sometió a prueba si esta distinción

sigue siendo significativa para el pronóstico entre pacientes tratados con R-CHOP (Coiffier *et al.*, N. Engl. J. Med., 346: 235-42 (2002)). Se determinaron los perfiles de expresión génica para muestras de biopsia antes del tratamiento de un “conjunto de entrenamiento” de 181 pacientes tratados con quimioterapia con CHOP o similar a CHOP solo y de un “conjunto de prueba” de 233 pacientes tratados con R-CHOP. Los pacientes en estas dos cohortes eran comparables con respecto al intervalo de edad y distribución de las variables clínicas de pronóstico que constituyen el índice de pronóstico internacional (IPI) (tabla 2). En la cohorte de R-CHOP, los pacientes con DLBCL GCB tenían mejores tasas de supervivencia que aquéllos con DLBCL ABC. Específicamente, los pacientes con DLBCL GCB y DLBCL ABC tratados con R-CHOP tenían tasas de supervivencia global a los 3 años del 84% y el 56%, respectivamente, y tasas de supervivencia libre de progresión a los 3 años del 74% y el 40%, respectivamente (figuras 1A y 1B). En el conjunto de entrenamiento de CHOP, y en una segunda cohorte de “MMMLMP” CHOP (Hummel *et al.*, citado anteriormente), las tasas de supervivencia global para DLBCL ABC y DLBCL GCB fueron menores que en la cohorte de R-CHOP (figura 6). El análisis de múltiples variables indicó que el beneficio relativo (es decir, el cambio en el resultado de supervivencia) debido a la terapia con R-CHOP (en comparación con CHOP) no era significativamente diferente entre DLBCL ABC y GCB.

Se ha mostrado anteriormente que cuatro firmas de expresión génica tienen significación en el pronóstico en pacientes con DLBCL tratados con CHOP (Rosenwald *et al.*, citado anteriormente). De éstas, la firma de GCB y la firma de ganglio linfático estaban asociadas con supervivencia favorable, y la firma de proliferación estaba asociada con supervivencia inferior dentro del conjunto de entrenamiento de CHOP, en la cohorte de MMMLNP CHOP (véanse los paneles de firmas correspondientes en la figura 7), y en la cohorte de R-CHOP (véanse los paneles de firmas correspondientes en la figura 1C). Por tanto, las diferencias biológicas entre tumores de DLBCL reflejadas por estas tres firmas siguen siendo importantes para el pronóstico en pacientes tratados con rituximab, aun cuando el tratamiento con rituximab mejoró en general la supervivencia en DLBCL.

Las cuarta firma de expresión génica restante, la firma de MHC de clase II, que estaba asociada con supervivencia en el conjunto de entrenamiento de CHOP cuando se trataba como una variable continua, no estaba asociada con supervivencia en la cohorte de R-CHOP (véase también el panel de firmas de MHC de clase II en la figura 1C). Además, tumores con expresión de “resultados atípicos” extremadamente baja de esta firma estaban asociados con supervivencia inferior en ambas cohortes de CHOP (véanse las figuras 8A y 8B), pero no en la cohorte de R-CHOP (véase la figura 8C).

Los resultados anteriores indican que la inmunoterapia con rituximab combinada con quimioterapia (R-CHOP) es beneficiosa para los subtipos tanto ABC como GCB de DLBCL y que las firmas de expresión génica que predecían la supervivencia en el contexto de la quimioterapia con CHOP mantenían su poder de pronóstico entre pacientes tratados con R-CHOP.

Los resultados anteriores también indican que la variación biológica entre tumores de DLBCL, tal como se mide mediante las firmas de expresión génica, tiene una relación constante con la respuesta terapéutica independientemente del régimen de tratamiento usado. Hay una diferencia sorprendente en la supervivencia libre de progresión a los 3 años entre pacientes con DLBCL ABC y pacientes con DLBCL GCB tratados con R-CHOP (el 40% frente al 74%). Esta diferencia se debe probablemente a diferencias genéticas y biológicas entre estos subtipos de DLBCL (Staudt *et al.*, Adv. Immunol., 87: 163-208 (2005)).

Por tanto, futuros ensayos clínicos en DLBCL deben incorporar métodos cuantitativos para discernir estas diferencias biológicas de modo que las cohortes de pacientes en diferentes ensayos puedan compararse y las respuestas al tratamiento puedan relacionarse con fenotipos tumorales definidos.

## EJEMPLO 2

Este ejemplo demuestra el desarrollo de firmas de supervivencia de GCB, estromal-1 y estromal-2 y un modelo de múltiples variables de supervivencia para DLBCL tratado con R-CHOP.

A menos que se indique otra cosa, las cohortes de pacientes y los métodos de análisis de la expresión génica son tal como se describen en el ejemplo 1.

En los datos de la cohorte de LLMP CHOP, se identificaron 936 genes como asociados con mal pronóstico  $p < 0,01$  (unilateral). Para genes que tienen múltiples conjuntos de sondas de alineamientos asociados con supervivencia, sólo se usó el conjunto de sondas con la asociación más fuerte con supervivencia. Entonces se agruparon los valores de expresión de los conjuntos de sondas en los datos de la cohorte de LLMP CHOP. Se identificó la mayor agrupación con una correlación promedio de  $>0,6$  y que contenía *myc* como la firma de supervivencia de proliferación. Se identificaron 1396 genes como asociados con resultado favorable. Se identificó la mayor agrupación con una correlación promedio de  $>0,6$  y que contenía *BCL6* como firma de supervivencia de células B de centro germinal (GCB). Se identificó una agrupación con una correlación promedio de  $>0,6$  y que contenía *FNI* como firma de supervivencia de estromal-1, mientras que se identificó otra agrupación con una correlación promedio de  $>0,6$  que contenía *HLADRA* como firma de supervivencia de MHC de clase II. Entonces se calculó el promedio de los niveles de expresión de genes dentro de cada firma para crear un “promedio de firma” para cada muestra de biopsia. Para el conjunto de datos de MMMLNP CHOP, se calculó el promedio para los elementos de alineamientos

representados en la micromatriz de Affymetrix U133A.

De las cuatro firmas o agrupaciones de pronóstico, se usaron dos firmas, la firma de estromal-1 y de GCB para crear el mejor modelo de supervivencia de dos variables. Ni la firma de proliferación ni la de MHC de clase II se añadieron al valor de pronóstico de este modelo de dos variables. Este modelo de dos variables funcionó bien en la cohorte de MMMLNP CHOP (figura 9A) y en la cohorte de R-CHOP (figura 2A).

Se usó el conjunto de entrenamiento de CHOP para descubrir y refinar firmas que se añadieron a la significación de pronóstico de este modelo de dos variables, y se sometieron a prueba los modelos de múltiples variables resultantes en la cohorte de R-CHOP. Se identificó que 563 genes se añadían al modelo en la dirección de pronóstico adverso. Estos genes se agruparon mediante agrupamiento jerárquico, y se identificaron tres agrupaciones de más de 10 genes con una correlación promedio de  $>0,6$ . Además, se identificaron 542 genes que se añadieron al modelo de firma de estromal-1 y de GCB en la dirección de pronóstico favorable. Se agruparon estos genes, y se identificaron dos agrupaciones de más de 10 genes con una correlación promedio de  $>0,6$ . Se determinaron los promedios de las firmas para estas agrupaciones, y se formaron tres modelos de variables que contenían la firma de estromal-1 y de GCB y se formó cada uno de los promedios de agrupaciones en los conjuntos de datos de MMMLNP CHOP y R-CHOP. De los cinco promedios de agrupaciones, se encontró que dos añadían significación estadística ( $p<0,02$ ) en los datos de MMMLNP CHOP en comparación con un modelo que contenía las firmas de estromal-1 y de GCB solo. En cambio, en los datos de R-CHOP, se encontró que tres de cinco promedios de agrupaciones añadían significación ( $p<0,02$ ) al modelo de dos variables. Uno de estos promedios de agrupaciones se añadió significativamente al modelo de dos variables en los datos de tanto MMMLNP CHOP como R-CHOP. También se encontró que esta firma, designada firma 122, se añadía a la firma de estromal-1 y de GCB mucho más significativamente que cualquiera de las otras cuatro firmas en los datos de LLMP CHOP y, por tanto, se mantuvo para su análisis adicional.

La firma 122 se añadía significativamente al modelo de dos variables tanto en la cohorte de MMMLNP CHOP ( $p=0,011$ ) como en la cohorte de R-CHOP ( $p=0,001$ ) (figuras 9B y 9C). Esta firma 122 se correlacionaba positivamente con la firma de estromal-1, aunque estaba asociada con supervivencia adversa cuando se añadía al modelo de dos variables. Para refinar adicionalmente el modelo, se identificaron genes que estaban más correlacionados con la firma 122 que con la firma de estromal-1 ( $p<0,02$ ). Se organizaron estos genes mediante agrupamiento jerárquico, y se observaron tres conjuntos de genes correlacionados ( $r>0,6$ ). Una de estas agrupaciones, la firma de estromal-2, se añadió a la significación del modelo de dos variables tanto en la cohorte de MMMLNP CHOP ( $p=0,002$ ) como en la cohorte de R-CHOP ( $p<0,001$ ) (figuras 2B y 9D).

Se formó un modelo de supervivencia de múltiples variables ajustando un modelo de Cox con las firmas de GCB, de estromal-1 y de estromal-2 a los datos de la cohorte de LLMP CHOP mostrados en la tabla 3. Entonces se evaluó este modelo de múltiples variables final con sus coeficientes de escala asociados en los conjuntos de datos de las cohortes de MMMLNP CHOP y R-CHOP. Se usaron las puntuaciones de factores de predicción de supervivencia del modelo final para dividir la cohorte de R-CHOP en grupos de cuartil con tasas de supervivencia global a los 3 años del 89%, el 82%, el 74% y el 48%, y tasas de supervivencia libre de progresión a los 3 años del 84%, el 69%, el 61% y el 33% (figura 2B). Las puntuaciones de factores de predicción de supervivencia del modelo final se ilustran en la figura 3 junto con las firmas de tres componentes y genes representativos de cada firma.

TABLA 3

| Paciente | Tiempo hasta la muerte o el último seguimiento (años) | Estado en el último seguimiento (1=muerto, 0=vivo) | Tiempo hasta la muerte, progresión o último seguimiento (años) | Estado en el último seguimiento (1=con progresión o muerto, 0=sin progresión) | Promedio de firma de centro germinal | Promedio de firma de estromal-1 | Promedio de firma de estromal-2 | Puntuación del modelo |
|----------|---|--|--|---|--------------------------------------|---------------------------------|---------------------------------|-----------------------|
| 2        | 2,75  | 0  | 2,75   | 0   | 9,238                                | 8,778                           | 7,475                           | 0,376                 |
| 3        | 2,67  | 0  | 2,67   | 0   | 9,942                                | 8,227                           | 7,102                           | 0,387                 |
| 5        | 1,27  | 1  | 0,72   | 1   | 8,859                                | 9,033                           | 8,716                           | 1,113                 |
| 21       | 2,39  | 0  | 2,40   | 0   | 10,573                               | 8,519                           | 6,959                           | -0,270                |
| 22       | 2,38  | 0  | 2,38   | 0   | 8,737                                | 8,686                           | 7,598                           | 0,761                 |
| 23       | 2,52  | 0  | 2,52   | 0   | 10,694                               | 10,322                          | 8,817                           | -0,897                |
| 24       | 5,11  | 0  | 5,11   | 0   | 11,376                               | 7,854                           | 7,598                           | 0,500                 |
| 26       | 4,01  | 0  | 4,01   | 0   | 9,829                                | 9,956                           | 8,507                           | -0,372                |
| 28       | 3,96  | 0  | 3,96   | 0   | 10,957                               | 9,277                           | 8,248                           | -0,330                |
| 41       | 0,52  | 1  | 0,52   | 1   | 9,273                                | 9,437                           | 8,202                           | 0,183                 |
| 47       | 1,53  | 1  | 0,77   | 1   | 9,548                                | 8,802                           | 8,061                           | 0,617                 |
| 48       | 0,37  | 1  | 0,12   | 1   | 8,660                                | 8,279                           | 6,891                           | 0,729                 |
| 49       | 2,37  | 0  | 2,35   | 1   | 10,915                               | 8,988                           | 6,847                           | -0,965                |
| 53       | 3,89  | 0  | 2,23   | 1   | 9,530                                | 9,792                           | 9,693                           | 0,721                 |
| 61       | 0,90  | 1  | 0,46   | 1   | 8,649                                | 8,038                           | 8,104                           | 1,798                 |
| 65       | 4,04  | 0  | 4,04   | 0   | 10,744                               | 9,330                           | 7,930                           | -0,508                |
| 66       | 4,04  | 0  | 4,04   | 0   | 10,714                               | 10,016                          | 7,536                           | -1,459                |
| 95       | 0,62  | 1  | 0,44   | 1   | 9,244                                | 9,197                           | 8,105                           | 0,373                 |
| 96       | 5,37  | 0  | 5,37   | 0   | 10,107                               | 8,723                           | 7,608                           | 0,157                 |
| 97       | 5,07  | 0  | 5,07   | 0   | 9,777                                | 9,192                           | 7,359                           | -0,349                |
| 98       | 0,94  | 1  | 0,59   | 1   | 8,794                                | 7,711                           | 7,367                           | 1,571                 |
| 99       | 0,40  | 1  | 0,40   | 1   | 9,024                                | 9,272                           | 9,160                           | 1,101                 |
| 103      | 0,03  | 1  | 0,02   | 1   | 8,883                                | 8,190                           | 7,742                           | 1,301                 |
| 104      | 3,76  | 0  | 3,76   | 0   | 9,785                                | 9,866                           | 7,929                           | -0,652                |
| 106      | 2,95  | 0  | 2,95   | 0   | 10,585                               | 7,797                           | 6,824                           | 0,367                 |
| 107      | 2,94  | 0  | 2,94   | 0   | 11,535                               | 8,358                           | 6,660                           | -0,711                |
| 108      | 2,73  | 0  | 2,73   | 0   | 9,653                                | 8,495                           | 7,550                           | 0,539                 |
| 109      | 0,16  | 1  | 0,11   | 1   | 9,301                                | 9,376                           | 7,994                           | 0,092                 |
| 110      | 2,46  | 0  | 2,46   | 0   | 10,254                               | 8,980                           | 7,324                           | -0,357                |
| 111      | 2,44  | 0  | 2,44   | 0   | 10,137                               | 10,691                          | 8,948                           | -0,949                |
| 113      | 2,12  | 0  | 2,12   | 0   | 10,746                               | 8,555                           | 6,942                           | -0,390                |
| 114      | 1,98  | 0  | 0,88   | 1   | 8,562                                | 8,159                           | 7,120                           | 1,047                 |
| 115      | 1,92  | 0  | 1,92   | 0   | 10,313                               | 9,385                           | 8,157                           | -0,231                |
| 118      | 1,64  | 0  | 1,64   | 0   | 10,209                               | 10,194                          | 8,231                           | -0,959                |
| 119      | 1,60  | 0  | 1,60   | 0   | 11,059                               | 8,852                           | 7,479                           | -0,461                |
| 1087     | 0,05  | 1  | 0,05   | 1   | 8,756                                | 8,491                           | 7,949                           | 1,188                 |

| Paciente | Tiempo hasta la muerte o el último seguimiento (años) | Estado en el último seguimiento (1=muerto, 0=vivo) | Tiempo hasta la muerte, progresión o último seguimiento (años) | Estado en el último seguimiento (1=con progresión o muerto, 0=sin progresión) | Promedio de firma de centro germinal | Promedio de firma de estromal-1 | Promedio de firma de estromal-2 | Puntuación del modelo |
|----------|---|--|--|---|--------------------------------------|---------------------------------|---------------------------------|-----------------------|
| 1089     | 5,12  | 0  | 1,27   | 1   | 9,863                                | 9,135                           | 8,034                           | 0,129                 |
| 1091     | 5,15  | 0  | 5,15   | 0   | 10,454                               | 9,918                           | 8,742                           | -0,437                |
| 1092     | 5,06  | 0  | 5,07   | 0   | 9,452                                | 9,467                           | 8,912                           | 0,556                 |
| 1093     | 3,83  | 1  | 1,62   | 1   | 9,915                                | 9,138                           | 7,747                           | -0,090                |
| 1096     | 4,02  | 0  | 4,02   | 0   | 8,887                                | 9,236                           | 7,795                           | 0,274                 |
| 1097     | 1,26  | 1  | 1,08   | 1   | 11,219                               | 9,234                           | 8,321                           | -0,347                |
| 1098     | 3,53  | 0  | 3,53   | 0   | 9,117                                | 9,236                           | 7,655                           | 0,082                 |
| 1099     | 3,07  | 0  | 0,91   | 1   | 9,284                                | 8,798                           | 7,741                           | 0,515                 |
| 1101     | 5,64  | 0  | 5,64   | 0   | 9,803                                | 9,466                           | 8,156                           | -0,101                |
| 1108     | 3,30  | 0  | 3,30   | 0   | 9,195                                | 10,456                          | 9,065                           | -0,237                |
| 1109     | 3,78  | 0  | 3,78   | 0   | 11,008                               | 10,051                          | 8,273                           | -1,120                |
| 1164     | 0,19  | 1  | 0,16   | 1   | 9,242                                | 10,307                          | 10,548                          | 0,896                 |
| 1167     | 1,49  | 1  | 0,45   | 1   | 9,809                                | 9,105                           | 8,784                           | 0,687                 |
| 1168     | 0,42  | 1  | 0,30   | 1   | 8,718                                | 8,368                           | 7,149                           | 0,790                 |
| 1169     | 1,71  | 1  | 1,22   | 1   | 11,512                               | 8,108                           | 7,507                           | 0,125                 |
| 1172     | 2,82  | 0  | 2,82   | 0   | 11,137                               | 8,871                           | 8,153                           | -0,057                |
| 1173     | 0,87  | 1  | 0,79   | 1   | 11,324                               | 9,914                           | 8,514                           | -0,950                |
| 1175     | 1,06  | 1  | 0,56   | 1   | 9,107                                | 10,310                          | 9,063                           | -0,053                |
| 1179     | 2,53  | 0  | 2,53   | 0   | 9,506                                | 9,437                           | 8,461                           | 0,260                 |
| 1181     | 1,72  | 0  | 1,72   | 0   | 10,688                               | 9,018                           | 7,647                           | -0,360                |
| 1184     | 4,74  | 0  | 2,97   | 1   | 10,812                               | 8,979                           | 7,922                           | -0,187                |
| 1185     | 3,71  | 0  | 3,71   | 0   | 10,431                               | 8,397                           | 7,317                           | 0,156                 |
| 1186     | 3,43  | 0  | 3,43   | 0   | 8,688                                | 8,944                           | 8,552                           | 1,164                 |
| 1187     | 5,23  | 0  | 5,23   | 0   | 10,072                               | 10,192                          | 8,667                           | -0,604                |
| 1189     | 5,13  | 0  | 5,13   | 0   | 10,109                               | 9,212                           | 7,967                           | -0,097                |
| 1190     | 3,66  | 0  | 3,66   | 0   | 10,713                               | 10,409                          | 8,910                           | -0,930                |
| 1192     | 0,16  | 1  | 0,16   | 1   | 8,825                                | 9,903                           | 8,061                           | -0,199                |
| 1195     | 4,36  | 0  | 4,36   | 0   | 11,539                               | 7,567                           | 6,873                           | 0,234                 |
| 1197     | 3,13  | 0  | 3,13   | 0   | 10,287                               | 10,365                          | 9,549                           | -0,275                |
| 1200     | 0,31  | 1  | 0,31   | 1   | 9,432                                | 8,950                           | 9,805                           | 1,692                 |
| 1206     | 6,51  | 0  | 6,51   | 0   | 10,410                               | 9,946                           | 8,925                           | -0,323                |
| 1211     | 6,25  | 0  | 6,25   | 0   | 11,596                               | 7,908                           | 6,524                           | -0,372                |
| 1215     | 5,35  | 0  | 5,35   | 0   | 10,504                               | 9,061                           | 7,550                           | -0,392                |
| 1216     | 0,46  | 1  | 0,29   | 1   | 10,017                               | 9,010                           | 7,794                           | 0,028                 |
| 1219     | 0,51  | 1  | 0,51   | 1   | 10,614                               | 10,014                          | 8,619                           | -0,683                |
| 1220     | 2,24  | 1  | 2,25   | 1   | 8,850                                | 9,400                           | 8,036                           | 0,286                 |
| 1221     | 3,94  | 0  | 3,95   | 0   | 8,777                                | 7,489                           | 6,672                           | 1,334                 |
| 1222     | 3,53  | 0  | 3,53   | 0   | 10,463                               | 9,310                           | 7,019                           | -0,986                |



| Paciente | Tiempo hasta la muerte o el último seguimiento (años) | Estado en el último seguimiento (1=muerto, 0=vivo) | Tiempo hasta la muerte, progresión o último seguimiento (años) | Estado en el último seguimiento (1=con progresión o muerto, 0=sin progresión) | Promedio de firma de centro germinal | Promedio de firma de estromal-1 | Promedio de firma de estromal-2 | Puntuación del modelo |
|----------|---|--|--|---|--------------------------------------|---------------------------------|---------------------------------|-----------------------|
| 1224     | 3,22  | 0  | 2,11   | 1   | 9,751                                | 9,505                           | 8,453                           | 0,082                 |
| 1225     | 2,95  | 0  | 2,95   | 0   | 8,613                                | 8,313                           | 7,668                           | 1,240                 |
| 1226     | 0,08  | 1  | 0,08   | 1   | 9,229                                | 8,851                           | 7,950                           | 0,625                 |
| 1228     | 2,78  | 0  | 0,99   | 1   | 11,532                               | 8,261                           | 6,932                           | -0,428                |
| 1230     | 0,59  | 1  | 0,54   | 1   | 9,369                                | 6,951                           | 6,956                           | 1,825                 |
| 1231     | 1,41  | 0  | 1,41   | 0   | 10,248                               | 8,788                           | 8,011                           | 0,303                 |
| 1232     | 2,49  | 0  | 0,68   | 1   | 10,362                               | 8,528                           | 7,975                           | 0,495                 |
| 1233     | 2,50  | 0  | 2,50   | 0   | 9,239                                | 10,581                          | 8,470                           | -0,784                |
| 1236     | 2,56  | 0  | 2,56   | 0   | 9,156                                | 10,000                          | 7,805                           | -0,608                |
| 1238     | 0,16  | 1  | 0,16   | 1   | 9,488                                | 9,055                           | 8,256                           | 0,517                 |
| 1239     | 2,24  | 0  | 2,24   | 0   | 8,886                                | 8,978                           | 7,838                           | 0,564                 |
| 1240     | 1,48  | 0  | 1,48   | 0   | 10,474                               | 9,073                           | 7,702                           | -0,288                |
| 1241     | 1,41  | 1  | 1,17   | 1   | 9,044                                | 9,054                           | 7,451                           | 0,160                 |
| 1251     | 2,72  | 0  | 2,72   | 0   | 8,410                                | 8,687                           | 7,082                           | 0,549                 |
| 1252     | 0,01  | 1  | 0,01   | 1   | 11,167                               | 8,070                           | 7,358                           | 0,206                 |
| 1255     | 5,17  | 0  | 5,17   | 0   | 9,501                                | 9,411                           | 7,887                           | -0,099                |
| 1271     | 4,72  | 0  | 4,73   | 0   | 10,718                               | 8,452                           | 7,060                           | -0,194                |
| 1272     | 5,68  | 0  | 5,68   | 0   | 9,161                                | 9,080                           | 7,668                           | 0,231                 |
| 1275     | 1,89  | 1  | 1,48   | 1   | 9,257                                | 8,559                           | 8,607                           | 1,354                 |
| 1277     | 5,06  | 0  | 5,07   | 0   | 11,091                               | 9,938                           | 8,274                           | -1,038                |
| 1279     | 4,87  | 0  | 4,87   | 0   | 9,309                                | 10,085                          | 9,676                           | 0,504                 |
| 1281     | 3,36  | 0  | no disponible (n/a)  | n/a   | 9,535                                | 9,969                           | 9,090                           | 0,132                 |
| 1284     | 3,51  | 0  | 3,51   | 0   | 10,922                               | 9,680                           | 8,481                           | -0,567                |
| 1288     | 1,54  | 0  | n/a  | n/a   | 9,430                                | 8,896                           | 8,037                           | 0,554                 |
| 1289     | 0,03  | 1  | 0,03   | 1   | 8,915                                | 9,052                           | 8,002                           | 0,589                 |
| 1290     | 5,23  | 0  | 5,23   | 0   | 10,432                               | 10,426                          | 8,154                           | -1,340                |
| 1291     | 0,04  | 1  | 0,04   | 1   | 11,319                               | 8,246                           | 7,323                           | -0,059                |
| 1292     | 0,10  | 1  | 0,10   | 1   | 8,667                                | 8,764                           | 8,110                           | 1,058                 |
| 1293     | 4,81  | 0  | 4,81   | 0   | 11,116                               | 9,842                           | 8,083                           | -1,081                |
| 1294     | 0,53  | 1  | 0,53   | 1   | 10,138                               | 10,181                          | 8,501                           | -0,733                |
| 1295     | 5,16  | 0  | 5,17   | 0   | 9,445                                | 9,694                           | 7,739                           | -0,463                |
| 1296     | 4,79  | 0  | 4,79   | 0   | 10,228                               | 9,064                           | 8,852                           | 0,600                 |
| 1297     | 4,24  | 0  | 4,24   | 0   | 9,524                                | 7,990                           | 7,008                           | 0,740                 |
| 1298     | 4,56  | 0  | 4,56   | 0   | 9,022                                | 9,000                           | 7,695                           | 0,389                 |
| 1331     | 3,29  | 0  | 3,29   | 0   | 11,004                               | 9,488                           | 8,289                           | -0,536                |
| 1334     | 2,87  | 0  | 2,87   | 0   | 11,434                               | 9,509                           | 8,109                           | -0,859                |
| 1335     | 1,38  | 1  | 0,90   | 1   | 9,586                                | 8,545                           | 7,423                           | 0,431                 |
| 1336     | 2,44  | 0  | 2,44   | 0   | 10,844                               | 9,704                           | 7,706                           | -1,082                |

| Paciente | Tiempo hasta la muerte o el último seguimiento (años) | Estado en el último seguimiento (1=muerto, 0=vivo) | Tiempo hasta la muerte, progresión o último seguimiento (años) | Estado en el último seguimiento (1=con progresión o muerto, 0=sin progresión) | Promedio de firma de centro germinal | Promedio de firma de estromal-1 | Promedio de firma de estromal-2 | Puntuación del modelo |
|----------|---|--|--|---|--------------------------------------|---------------------------------|---------------------------------|-----------------------|
| 1337     | 0,02  | 1  | 0,02   | 1   | 8,521                                | 7,788                           | 7,860                           | 1,941                 |
| 1449     | 1,62  | 0  | 1,62   | 0   | 9,604                                | 8,463                           | 8,030                           | 0,917                 |
| 1450     | 1,30  | 0  | 0,53   | 1   | 8,571                                | 8,112                           | 7,241                           | 1,173                 |
| 1451     | 1,84  | 0  | 1,85   | 0   | 10,637                               | 9,205                           | 7,759                           | -0,452                |
| 1453     | 1,71  | 0  | 1,71   | 0   | 10,964                               | 9,089                           | 8,226                           | -0,157                |
| 1454     | 0,62  | 0  | 0,62   | 0   | 11,106                               | 8,514                           | 7,604                           | -0,052                |
| 1553     | 2,93  | 0  | 1,92   | 1   | 8,975                                | 9,284                           | 7,475                           | -0,029                |
| 1612     | 5,37  | 0  | 5,37   | 0   | 10,526                               | 9,471                           | 7,809                           | -0,643                |
| 1613     | 5,81  | 0  | n/a  | n/a   | 10,868                               | 9,695                           | 7,730                           | -1,067                |
| 1614     | 4,36  | 1  | 4,36   | 1   | 10,358                               | 9,226                           | 8,765                           | 0,322                 |
| 1617     | 0,52  | 0  | 0,52   | 0   | 10,332                               | 8,723                           | 7,180                           | -0,227                |
| 1618     | 1,70  | 0  | 0,98   | 1   | 11,233                               | 8,956                           | 7,852                           | -0,387                |
| 1619     | 0,25  | 1  | 0,25   | 1   | 8,646                                | 8,028                           | 7,123                           | 1,146                 |
| 1620     | 2,17  | 0  | 2,17   | 0   | 11,647                               | 8,385                           | 7,343                           | -0,325                |
| 1623     | 2,80  | 0  | 2,80   | 0   | 9,611                                | 9,484                           | 8,249                           | 0,024                 |
| 1626     | 1,76  | 0  | 1,76   | 0   | 11,236                               | 9,495                           | 8,108                           | -0,763                |
| 1628     | 3,13  | 0  | 1,23   | 1   | 8,714                                | 7,972                           | 7,149                           | 1,192                 |
| 1645     | 2,85  | 0  | 2,85   | 0   | 10,146                               | 9,476                           | 8,914                           | 0,258                 |
| 1647     | 2,79  | 0  | 2,80,  | 0   | 10,485                               | 10,495                          | 8,707                           | -1,058                |
| 1650     | 0,75  | 1  | 0,75   | 1   | 8,830                                | 7,346                           | 6,486                           | 1,333                 |
| 1651     | 1,66  | 0  | 1,66   | 0   | 9,190                                | 7,949                           | 6,829                           | 0,801                 |
| 1652     | 1,64  | 0  | n/a  | n/a   | 8,798                                | 8,943                           | 8,331                           | 0,969                 |
| 1702     | 1,05  | 0  | 1,05   | 1   | 9,008                                | 8,217                           | 8,078                           | 1,447                 |
| 1703     | 0,70  | 1  | 0,70   | 1   | 9,499                                | 8,637                           | 7,790                           | 0,621                 |
| 1704     | 3,14  | 0  | 3,14   | 0   | 9,908                                | 9,231                           | 7,503                           | -0,347                |
| 1705     | 3,94  | 0  | 3,94   | 0   | 8,933                                | 8,445                           | 8,187                           | 1,321                 |
| 1707     | 2,80  | 0  | 2,80   | 0   | 10,610                               | 9,348                           | 7,872                           | -0,510                |
| 1742     | 3,27  | 0  | n/a  | n/a   | 10,033                               | 8,715                           | 7,412                           | 0,063                 |
| 1746     | 1,91  | 0  | 1,55   | 1   | 9,249                                | 8,705                           | 8,205                           | 0,937                 |
| 1747     | 1,48  | 0  | 1,48   | 0   | 10,162                               | 8,866                           | 7,602                           | -0,016                |
| 1756     | 3,47  | 0  | 3,47   | 0   | 10,815                               | 9,638                           | 7,248                           | -1,312                |
| 1761     | 0,23  | 1  | 0,23   | 1   | 9,842                                | 10,192                          | 8,664                           | -0,511                |
| 1762     | 5,20  | 0  | 5,20   | 0   | 10,583                               | 9,333                           | 7,445                           | -0,772                |
| 1763     | 5,51  | 0  | 5,51   | 0   | 8,917                                | 8,925                           | 8,084                           | 0,771                 |
| 1766     | 1,59  | 0  | 1,59   | 0   | 10,919                               | 10,037                          | 8,389                           | -0,990                |
| 1782     | 1,09  | 0  | 1,09   | 0   | 10,753                               | 9,600                           | 8,332                           | -0,516                |
| 1788     | 0,39  | 1  | 0,24   | 1   | 10,364                               | 8,738                           | 8,914                           | 0,915                 |
| 1861     | 0,56  | 1  | 0,19   | 1   | 9,728                                | 8,604                           | 7,594                           | 0,427                 |

| Paciente | Tiempo hasta la muerte o el último seguimiento (años) | Estado en el último seguimiento (1=muerto, 0=vivo) | Tiempo hasta la muerte, progresión o último seguimiento (años) | Estado en el último seguimiento (1=con progresión o muerto, 0=sin progresión) | Promedio de firma de centro germinal | Promedio de firma de estromal-1 | Promedio de firma de estromal-2 | Puntuación del modelo |
|----------|---|--|--|---|--------------------------------------|---------------------------------|---------------------------------|-----------------------|
| 1867     | 1,17  | 1  | 0,38   | 1   | 8,903                                | 11,501                          | 10,559                          | -0,166                |
| 1916     | 1,41  | 0  | n/a  | n/a   | 9,295                                | 11,197                          | 11,508                          | 0,619                 |
| 1920     | 1,32  | 0  | 1,32   | 0   | 10,165                               | 9,630                           | 8,789                           | 0,009                 |
| 1927     | 1,53  | 0  | 1,53   | 0   | 9,195                                | 10,261                          | 9,791                           | 0,451                 |
| 1928     | 0,72  | 0  | 0,72   | 0   | 9,769                                | 8,510                           | 7,330                           | 0,328                 |
| 1939     | 0,47  | 1  | 0,47   | 1   | 9,097                                | 9,363                           | 7,647                           | -0,043                |
| 2002     | 1,29  | 0  | 1,30   | 0   | 9,469                                | 9,542                           | 8,600                           | 0,262                 |
| 2006     | 1,23  | 0  | 1,23   | 0   | 10,434                               | 8,223                           | 7,162                           | 0,227                 |
| 2067     | 2,18  | 0  | 2,18   | 0   | 10,244                               | 11,186                          | 9,391                           | -1,197                |
| 2070     | 0,31  | 0  | 0,12   | 1   | 10,486                               | 10,680                          | 10,353                          | -0,135                |
| 2162     | 0,38  | 1  | 0,38   | 1   | 10,934                               | 10,020                          | 7,960                           | -1,268                |
| 2270     | 1,59  | 0  | 1,59   | 0   | 10,117                               | 9,904                           | 8,506                           | -0,440                |
| 2271     | 1,60  | 0  | 1,60   | 0   | 8,995                                | 9,349                           | 8,261                           | 0,428                 |
| 2274     | 0,41  | 0  | 0,41   | 0   | 8,863                                | 7,623                           | 7,222                           | 1,533                 |
| 2283     | 1,19  | 0  | 1,19   | 0   | 10,501                               | 8,361                           | 6,741                           | -0,226                |
| 2291     | 0,87  | 1  | 0,85   | 1   | 10,732                               | 10,184                          | 9,436                           | -0,353                |
| 2299     | 0,93  | 0  | 0,93   | 0   | 10,661                               | 9,905                           | 8,189                           | -0,883                |
| 2301     | 0,61  | 0  | 0,61   | 0   | 9,852                                | 9,903                           | 8,352                           | -0,432                |
| 2306     | 0,68  | 0  | 0,68   | 0   | 8,586                                | 8,759                           | 8,191                           | 1,151                 |
| 2309     | 0,43  | 0  | 0,43   | 0   | 10,839                               | 7,671                           | 6,860                           | 0,413                 |
| 2311     | 0,80  | 0  | 0,80   | 0   | 10,901                               | 7,797                           | 6,912                           | 0,294                 |
| 2318     | 0,99  | 0  | 0,99   | 0   | 10,283                               | 9,403                           | 8,655                           | 0,100                 |
| 2321     | 0,82  | 0  | 0,82   | 0   | 9,691                                | 8,956                           | 7,404                           | -0,044                |
| 2411     | 0,67  | 0  | 0,67   | 0   | 8,986                                | 8,383                           | 7,854                           | 1,137                 |
| 2415     | 0,62  | 0  | 0,62   | 0   | 9,296                                | 10,509                          | 9,551                           | -0,005                |
| 2444     | 3,99  | 0  | 3,99   | 0   | 10,154                               | 9,871                           | 9,026                           | -0,071                |
| 2445     | 3,36  | 0  | 3,36   | 0   | 8,788                                | 8,184                           | 7,964                           | 1,497                 |
| 2479     | 0,51  | 0  | 0,51   | 0   | 11,151                               | 9,023                           | 8,199                           | -0,186                |
| 2482     | 4,54  | 0  | 4,54   | 0   | 10,373                               | 9,847                           | 8,208                           | -0,691                |
| 2483     | 3,89  | 1  | 3,89   | 1   | 9,241                                | 8,902                           | 7,742                           | 0,428                 |
| 2484     | 2,69  | 1  | 1,90   | 1   | 10,279                               | 9,619                           | 8,312                           | -0,349                |
| 2485     | 4,43  | 0  | 4,43   | 0   | 9,957                                | 9,665                           | 8,439                           | -0,378                |
| 2486     | 4,37  | 0  | n/a  | n/a   | 10,698                               | 10,203                          | 8,041                           | -1,301                |
| 2487     | 4,34  | 0  | 4,34   | 0   | 11,227                               | 9,909                           | 8,260                           | -1,076                |
| 2488     | 4,20  | 0  | 4,21   | 0   | 9,510                                | 8,709                           | 7,615                           | 0,426                 |
| 2490     | 4,02  | 0  | 4,02   | 0   | 10,510                               | 10,961                          | 8,956                           | -1,374                |
| 2491     | 0,50  | 1  | 0,25   | 1   | 9,047                                | 8,554                           | 7,624                           | 0,784                 |
| 2492     | 3,96  | 0  | 3,96   | 0   | 9,904                                | 10,901                          | 9,140                           | -0,935                |

| Paciente | Tiempo hasta la muerte o el último seguimiento (años) | Estado en el último seguimiento (1=muerto, 0=vivo) | Tiempo hasta la muerte, progresión o último seguimiento (años) | Estado en el último seguimiento (1=con progresión o muerto, 0=sin progresión) | Promedio de firma de centro germinal | Promedio de firma de estromal-1 | Promedio de firma de estromal-2 | Puntuación del modelo |
|----------|---|--|--|---|--------------------------------------|---------------------------------|---------------------------------|-----------------------|
| 2497     | 3,44  | 0  | 3,44   | 0   | 9,221                                | 9,438                           | 8,065                           | 0,111                 |
| 2498     | 3,37  | 0  | 3,37   | 0   | 9,318                                | 9,427                           | 8,003                           | 0,040                 |
| 2500     | 3,31  | 0  | 3,31   | 0   | 11,014                               | 9,406                           | 7,375                           | -1,074                |
| 2501     | 3,28  | 0  | n/a  | n/a   | 8,822                                | 8,551                           | 7,750                           | 0,966                 |
| 2503     | 2,99  | 0  | 2,99   | 0   | 8,301                                | 7,967                           | 6,929                           | 1,222                 |
| 2504     | 2,78  | 0  | 2,78   | 0   | 10,145                               | 8,004                           | 7,017                           | 0,472                 |
| 2505     | 2,76  | 0  | 2,76   | 0   | 11,036                               | 8,442                           | 7,136                           | -0,266                |
| 2507     | 0,86  | 1  | 0,54   | 1   | 9,737                                | 9,475                           | 8,988                           | 0,480                 |
| 2508     | 2,58  | 0  | 2,58   | 0   | 8,678                                | 9,389                           | 8,230                           | 0,498                 |
| 2509     | 0,96  | 1  | 0,76   | 1   | 8,895                                | 10,441                          | 9,088                           | -0,081                |
| 2511     | 1,55  | 1  | 1,06   | 1   | 9,225                                | 9,267                           | 9,191                           | 1,042                 |
| 2512     | 2,45  | 0  | 2,45   | 0   | 11,047                               | 10,465                          | 9,337                           | -0,838                |
| 2513     | 0,61  | 1  | 0,61   | 1   | 10,855                               | 10,378                          | 8,395                           | -1,305                |
| 2514     | 2,18  | 0  | 2,18   | 0   | 10,477                               | 9,832                           | 7,498                           | -1,198                |
| 2515     | 2,13  | 0  | 2,13   | 0   | 9,295                                | 10,519                          | 9,788                           | 0,145                 |
| 2516     | 2,07  | 0  | 2,07   | 0   | 10,575                               | 10,592                          | 8,642                           | -1,238                |
| 2517     | 2,04  | 0  | 0,76   | 1   | 9,385                                | 9,163                           | 8,328                           | 0,498                 |
| 2584     | 0,68  | 0  | 0,68   | 0   | 10,759                               | 9,356                           | 8,135                           | -0,404                |
| 2599     | 4,05  | 0  | 4,05   | 0   | 10,629                               | 9,158                           | 7,724                           | -0,425                |
| 2600     | 1,01  | 1  | 0,54   | 1   | 9,785                                | 8,619                           | 7,291                           | 0,184                 |
| 2601     | 1,22  | 1  | 0,88   | 1   | 9,385                                | 8,044                           | 7,178                           | 0,859                 |
| 2603     | 4,43  | 0  | 4,43   | 0   | 9,582                                | 10,707                          | 9,803                           | -0,156                |
| 2604     | 0,84  | 0  | 0,36   | 1   | 9,844                                | 10,511                          | 8,382                           | -1,026                |
| 2609     | 8,89  | 0  | 2,55   | 1   | 8,981                                | 8,775                           | 7,506                           | 0,507                 |
| 2610     | 0,74  | 0  | 0,74   | 0   | 10,793                               | 8,964                           | 7,421                           | -0,502                |
| 2611     | 0,66  | 0  | 0,66   | 0   | 10,353                               | 10,233                          | 9,032                           | -0,518                |
| 2612     | 1,17  | 1  | 1,13   | 1   | 10,290                               | 9,028                           | 8,287                           | 0,230                 |
| 2613     | 1,66  | 0  | 1,66   | 0   | 10,997                               | 9,089                           | 7,749                           | -0,493                |
| 2614     | 0,21  | 1  | 0,21   | 1   | 8,768                                | 7,850                           | 7,100                           | 1,261                 |
| 2615     | 0,48  | 0  | 0,48   | 0   | 11,359                               | 9,470                           | 7,647                           | -1,100                |
| 2639     | 10,29   | 0  | 10,30  | 0   | 11,085                               | 10,385                          | 8,003                           | -1,674                |
| 2641     | 1,38  | 0  | 1,38   | 0   | 9,199                                | 8,818                           | 7,340                           | 0,259                 |
| 2642     | 3,67  | 0  | 3,67   | 0   | 10,731                               | 8,777                           | 7,167                           | -0,458                |
| 2643     | 5,49  | 0  | 5,49   | 0   | 10,236                               | 10,578                          | 8,473                           | -1,197                |
| 2645     | 0,19  | 0  | n/a  | n/a   | 11,130                               | 9,997                           | 8,254                           | -1,129                |
| 2646     | 0,18  | 1  | 0,18   | 1   | 8,893                                | 7,648                           | 6,871                           | 1,260                 |
| 2648     | 0,25  | 0  | 0,25   | 0   | 8,855                                | 7,745                           | 7,060                           | 1,303                 |
| 2649     | 2,13  | 0  | 2,13   | 0   | 9,688                                | 10,354                          | 9,885                           | 0,214                 |

| Paciente | Tiempo hasta la muerte o el último seguimiento (años) | Estado en el último seguimiento (1=muerto, 0=vivo) | Tiempo hasta la muerte, progresión o último seguimiento (años) | Estado en el último seguimiento (1=con progresión o muerto, 0=sin progresión) | Promedio de firma de centro germinal | Promedio de firma de estromal-1 | Promedio de firma de estromal-2 | Puntuación del modelo |
|----------|---|--|--|---|--------------------------------------|---------------------------------|---------------------------------|-----------------------|
| 2650     | 2,43  | 0  | n/a  | n/a   | 10,007                               | 10,052                          | 8,861                           | -0,305                |
| 2651     | 1,61  | 0  | n/a  | n/a   | 10,660                               | 9,452                           | 7,831                           | -0,665                |
| 2652     | 1,84  | 0  | 1,84   | 0   | 11,378                               | 9,247                           | 7,684                           | -0,856                |
| 2653     | 1,88  | 0  | 1,88   | 0   | 11,182                               | 9,638                           | 7,781                           | -1,106                |
| 2654     | 1,43  | 0  | 1,43   | 0   | 8,791                                | 9,395                           | 8,905                           | 0,902                 |
| 2813     | 3,97  | 0  | 3,97   | 0   | 10,701                               | 9,366                           | 8,258                           | -0,306                |
| 2814     | 0,81  | 1  | 0,70   | 1   | 10,561                               | 9,176                           | 9,275                           | 0,632                 |

5 El índice de pronóstico internacional (IPI), que se basa en 5 variables clínicas, predice la supervivencia en pacientes tanto tratados con CHOP como tratados con R-CHOP (Shipp *et al.*, N. Engl. J. Med., 329:987-94 (1993); Sehn *et al.*, Blood, 109: 1857-61 (2007)). El modelo de supervivencia basado en la expresión génica de la invención conservó su significación de pronóstico entre pacientes tratados con R-CHOP segregados según el IPI en grupos de riesgo alto, intermedio y bajo, tanto tal como se definió originalmente (Shipp *et al.*, citado anteriormente) ( $p < 0,001$ ) (figura 2C) como tal como se modificó recientemente para DLBCL tratado con R-CHOP (Sehn *et al.*, citado anteriormente) ( $p < 0,001$ ) (figura 10).

10 Los resultados anteriores indican que el modelo de múltiples variables basado en la expresión génica puede usarse para identificar grandes disparidades en la supervivencia entre pacientes con diferentes perfiles de firmas genes de DLBCL. Por tanto, se usaron las puntuaciones de factores de predicción de supervivencia para dividir a los pacientes en grupos de cuartil menos y más favorables que tienen tasas de supervivencia libre de progresión a los 3 años del 33% y el 84%, respectivamente. Dada su independencia estadística del IPI, el factor de predicción de supervivencia basado en la expresión génica proporciona una vista complementaria de la variación de DLBCL que puede considerarse cuando se analizan datos de ensayos clínicos de DLBCL. Adicionalmente, los resultados anteriores indican que los perfiles de expresión génica del genoma completo conjuntamente con el modelo de supervivencia descrito en el presente documento pueden usarse para proporcionar predicciones óptimas de resultados de supervivencia esperados para sujetos que padecen DLBCL.

### EJEMPLO 3

20 Este ejemplo demuestra el uso de una puntuación de factor de predicción de supervivencia para predecir la probabilidad de resultados de supervivencia global y libre de progresión en un periodo de tiempo t tras el tratamiento con R-CHOP según la invención.

25 Se aísla ARN de la biopsia de DLBCL de un paciente y se hibrida con un alineamiento U133+ de Affymetrix (Santa Clara, CA). Se escanea el alineamiento, y se aplica el algoritmo MAS 5.0 para obtener valores de señal normalizados a una intensidad objetivo de 500. Los valores de señal se transforman mediante log2 en valores de intensidad. Para genes de interés con múltiples conjuntos de sondas, se calcula el promedio del valor de intensidad de los múltiples conjuntos de sondas para obtener un único valor de intensidad para cada gen. Se calcula el promedio de los valores de intensidad únicos de genes en la firma de GCB para obtener un promedio de firma de GCB de 9,2. Se calcula el promedio de los valores de intensidad únicos en la firma de estromal-1 para obtener un promedio de firma de estromal-1 de 8,5. Se calcula el promedio de los valores de intensidad únicos de genes en la firma de estromal-2 para obtener un promedio de firma de estromal-2 de 7,2.

30 Se calcula la puntuación de factor de predicción de supervivencia del paciente usando la siguiente ecuación  $8,11 - [0,419 * (\text{promedio de firma de GCB}) - [1,015 * (\text{promedio de firma de estromal-1})] + [0,675 * (\text{promedio de firma de estromal-2})]$ , de manera que la puntuación de factor de predicción de supervivencia =  $8,11 - [0,419 * (9,2)] - [1,015 * (8,5)] + [0,675 * (7,2)] = 0,389$

35 La tabla 4 incluye valores de una curva de supervivencia libre de progresión generada usando funciones de riesgo iniciales calculadas a partir de los datos de pacientes con R-CHOP descritos en la tabla 3. Se generó la curva según los métodos de Kalbfleisch y Prentice, Biometrika, 60: 267-279 (1973), que implican maximizar la probabilidad total, suponiendo que los coeficientes de escala verdaderos eran iguales a las estimaciones previas. En la tabla 4,  $F_0(t)$  es la probabilidad de supervivencia libre de progresión para cada periodo de tiempo indicado tras el tratamiento con R-CHOP (t-R-CHOP).

Tabla 4

| t-RCHOP (años) | $F_0(t)$ |
|----------------|----------|
| 0,000          | 1,000    |
| 0,008          | 0,997    |
| 0,016          | 0,993    |
| 0,025          | 0,990    |
| 0,030          | 0,987    |
| 0,036          | 0,983    |
| 0,049          | 0,980    |
| 0,082          | 0,977    |
| 0,096          | 0,973    |
| 0,107          | 0,970    |
| 0,118          | 0,967    |
| 0,120          | 0,963    |
| 0,156          | 0,960    |
| 0,156          | 0,956    |
| 0,159          | 0,953    |
| 0,178          | 0,950    |

ES 2 552 937 T3

| <b>t-RCHOP (años)</b> | <b>F<sub>0</sub>(t)</b> |
|-----------------------|-------------------------|
| 0,192                 | 0,946                   |
| 0,211                 | 0,943                   |
| 0,233                 | 0,939                   |
| 0,241                 | 0,936                   |
| 0,246                 | 0,932                   |
| 0,252                 | 0,928                   |
| 0,290                 | 0,925                   |
| 0,298                 | 0,921                   |
| 0,307                 | 0,918                   |
| 0,364                 | 0,914                   |
| 0,381                 | 0,910                   |
| 0,381                 | 0,907                   |
| 0,400                 | 0,903                   |
| 0,441                 | 0,899                   |
| 0,446                 | 0,895                   |
| 0,463                 | 0,891                   |
| 0,468                 | 0,887                   |
| 0,515                 | 0,884                   |
| 0,517                 | 0,880                   |
| 0,531                 | 0,876                   |
| 0,534                 | 0,872                   |
| 0,537                 | 0,868                   |
| 0,537                 | 0,864                   |
| 0,539                 | 0,860                   |
| 0,561                 | 0,856                   |
| 0,586                 | 0,852                   |
| 0,611                 | 0,848                   |
| 0,679                 | 0,843                   |
| 0,698                 | 0,839                   |
| 0,698                 | 0,834                   |
| 0,720                 | 0,830                   |
| 0,747                 | 0,826                   |
| 0,756                 | 0,821                   |
| 0,761                 | 0,816                   |
| 0,767                 | 0,812                   |
| 0,786                 | 0,807                   |
| 0,849                 | 0,803                   |
| 0,879                 | 0,798                   |
| 0,884                 | 0,793                   |
| 0,898                 | 0,789                   |
| 0,912                 | 0,784                   |
| 0,977                 | 0,779                   |
| 0,986                 | 0,774                   |
| 1,046                 | 0,770                   |
| 1,057                 | 0,765                   |
| 1,076                 | 0,760                   |
| 1,128                 | 0,755                   |
| 1,166                 | 0,750                   |
| 1,216                 | 0,745                   |
| 1,227                 | 0,740                   |
| 1,270                 | 0,735                   |
| 1,481                 | 0,729                   |
| 1,547                 | 0,724                   |
| 1,624                 | 0,718                   |
| 1,900                 | 0,711                   |
| 1,919                 | 0,705                   |
| 2,105                 | 0,699                   |
| 2,231                 | 0,692                   |
| 2,245                 | 0,685                   |
| 2,352                 | 0,678                   |
| 2,546                 | 0,671                   |
| 2,968                 | 0,662                   |

| t-RCHOP (años) | F <sub>0</sub> (t) |
|----------------|--------------------|
| 3,890          | 0,648              |
| 4,364          | 0,623              |

Se calcula la probabilidad de supervivencia libre de progresión a los 2 años del paciente usando la ecuación:  $P(\text{PFS}) = F_0(t)^{\exp(0.976 \cdot \text{puntuación de factor de predicción de supervivencia})}$ , en la que  $F_0(t)$  es el valor de  $F_0(t)$  que corresponde al valor de tiempo más largo menor de 2 años en la curva de supervivencia libre de progresión. En la tabla 4, el valor de tiempo más largo menor de 2 es 1,919, y el valor de  $PF_0(t)$  correspondiente es 0,705. Por consiguiente, la probabilidad de supervivencia libre de progresión a los 2 años del paciente  $P(\text{PFS}) = 0,705^{\exp(0.976 \cdot \text{puntuación de factor de predicción de supervivencia})} = 0,705^{1.462} = 0,600$  o de aproximadamente el 60%.

5

La tabla 5 incluye valores de una curva de supervivencia global generada usando funciones de riesgo iniciales calculadas a partir de los datos de pacientes con R-CHOP en la tabla 3. Se preparó la curva según el método de Kalbfleisch y Prentice, *Biometrika*, 60: 267-279 (1973), que implica maximizar la probabilidad total, suponiendo que los coeficientes de escala verdaderos eran iguales a las estimaciones. En la tabla 5,  $OS_0(t)$  es la probabilidad de supervivencia global para cada periodo de tiempo indicado tras el tratamiento con R-CHOP (t-R-CHOP).

10

Tabla 5

| t-RCHOP (años) | OS <sub>0</sub> (t) |
|----------------|---------------------|
| 0,000          | 1,000               |
| 0,008          | 0,997               |
| 0,016          | 0,994               |
| 0,030          | 0,991               |
| 0,033          | 0,988               |
| 0,036          | 0,984               |
| 0,049          | 0,981               |
| 0,082          | 0,978               |
| 0,096          | 0,975               |
| 0,156          | 0,972               |
| 0,156          | 0,969               |
| 0,159          | 0,965               |
| 0,178          | 0,962               |
| 0,192          | 0,959               |
| 0,211          | 0,956               |
| 0,233          | 0,952               |
| 0,246          | 0,949               |
| 0,307          | 0,946               |
| 0,367          | 0,942               |
| 0,380          | 0,939               |
| 0,386          | 0,935               |
| 0,402          | 0,932               |
| 0,416          | 0,928               |
| 0,463          | 0,925               |
| 0,468          | 0,921               |
| 0,504          | 0,918               |
| 0,515          | 0,914               |
| 0,517          | 0,910               |
| 0,531          | 0,907               |
| 0,556          | 0,903               |
| 0,586          | 0,900               |
| 0,610          | 0,896               |
| 0,619          | 0,892               |
| 0,698          | 0,888               |
| 0,747          | 0,885               |
| 0,807          | 0,881               |
| 0,862          | 0,877               |
| 0,868          | 0,873               |
| 0,873          | 0,869               |
| 0,895          | 0,864               |
| 0,944          | 0,860               |
| 0,963          | 0,856               |
| 1,010          | 0,852               |
| 1,057          | 0,848               |



| t-RCHOP (años) | OS <sub>0</sub> (t) |
|----------------|---------------------|
| 1,169          | 0,843               |
| 1,169          | 0,839               |
| 1,215          | 0,835               |
| 1,262          | 0,830               |
| 1,273          | 0,826               |
| 1,382          | 0,821               |
| 1,412          | 0,817               |
| 1,492          | 0,812               |
| 1,527          | 0,807               |
| 1,552          | 0,802               |
| 1,708          | 0,796               |
| 1,889          | 0,791               |
| 2,244          | 0,784               |
| 2,693          | 0,777               |
| 3,826          | 0,763               |
| 3,889          | 0,749               |
| 4,363          | 0,724               |

Se calcula la probabilidad de supervivencia global a los 2 años del paciente usando la ecuación:  $P(OS) = OS_0(t)^{\text{(exp(puntuación de factor de predicción de supervivencia))}}$ , en la que  $OS_0(t)$  es el valor que corresponde al valor de tiempo más largo en la curva de supervivencia global que es menor de 2 años. En la tabla 5, el valor de tiempo más largo menor de 2 es 1,889, y el valor de  $OS_0(t)$  correspondiente es 0,791. Por consiguiente, la probabilidad de supervivencia global a los 2 años del paciente es  $P(PFS) = 0,791^{\text{(exp(0,389))}} = 0,791^{1,4476} = 0,707$  o el 70,7%.

#### EJEMPLO 4

Este ejemplo demuestra la base biológica para las firmas de pronóstico de DLBCL.

A menos que se indique otra cosa, en los ejemplos 1 y 2 se describen cohortes y métodos de análisis de la expresión génica. Además, se separaron suspensiones celulares de tres biopsias mediante citometría de flujo en una subpoblación maligna CD19+ y una subpoblación no maligna CD19-. Se realizó la obtención de perfiles de expresión génica tras dos rondas de amplificación lineal a partir de ARN total (Dave *et al.*, N. Engl. J. Med., 351: 2159-69 (2004)). Tras la normalización con MAS5.0, se seleccionaron genes que tenían un valor de señal de log2 mayor de 7 en las fracciones o bien CD19+ o bien CD19- en al menos dos de las muestras clasificadas.

Para evaluar si las la firmas de expresión génica en el modelo de supervivencia final del ejemplo 2 se derivaban de las células de linfoma malignas o del microentorno del huésped, se fraccionaron tres muestras de biopsia de DLBCL en células malignas CD19+ y células no malignas CD19- mediante clasificación celular. La mayoría de los genes de firma de células B de centro germinal se expresaban más en la fracción maligna, mientras que los genes de las firmas de estromal-1 y estromal-2 se expresaban más en la fracción estromal no maligna (figura 4A), de ahí su nombre. Puesto que estas dos firmas eran sinérgicas en la predicción de la supervivencia, se combinaron en una "puntuación estromal" (figura 3), cuyos altos valores estaban asociados con resultado adverso.

La firma de células B de centro germinal se refiere a la distinción entre los subtipos de DLBCL ABC y GCB (figura 3). En cambio, los genes que definen la firma de estromal-1 codifican para componentes de la matriz extracelular, incluyendo fibronectina, osteonectina, diversas isoformas de colágeno y laminina, y el factor antiangiogénico trombospondina (figura 3 y tabla 1). Esta firma también codifica para modificadores de la síntesis de colágeno (*LOXL1*, *SERPINH1*), proteínas que remodelan la matriz extracelular (*MMP2*, *MMP9*, *MMP14*, *PLAU*, *TIMP2*), y *CTGF*, una proteína secretada que puede iniciar respuestas fibróticas (Frazier *et al.*, J. Invest. Dermatol., 107(3): 404-11 (1996)). Además, la firma de estromal-1 incluye genes expresados de manera característica en células del linaje monocítico, tales como *CEBPA* y *CSF2RA*.

La firma de estromal-1 está relacionada significativamente con varias firmas de expresión génica curadas anteriormente (Shaffer *et al.*, Immunol. Rev., 210: 67-85 (2006)) basándose en análisis de enriquecimiento de conjuntos de genes (Subramanian *et al.*, Proc. Nat'l. Acad. Sci. USA, 102(43): 15545-50 (2005)). Dos de estas firmas incluyen genes que se expresan de manera coordinada en tejidos mesenquimales normales pero no en subconjuntos hematopoyéticos, muchos de los cuales codifican para proteínas de la matriz extracelular (tasa de falso descubrimiento (FDR) < 0,001) (figuras 4B y 11) (Su *et al.*, Proc. Nat'l. Acad. Sci. USA, 101: 6062-7 (2004)). También se enriqueció una firma de "monocitos", compuesta por genes que se expresan más en monocitos de la sangre CD14+ que en células B, células T o células NK (FDR=0,014) (figura 4B). En cambio, una firma de células pan-T no estaba relacionada con la firma de estromal-1 (figura 4B). Estos hallazgos sugieren que una alta expresión de la firma de estromal-1 identifica tumores con deposición vigorosa de matriz extracelular e infiltración por células del linaje monocítico.

En este sentido, el producto génico de la firma de estromal-1, fibronectina, se localizada de manera destacada

mediante inmunohistoquímica en hebras fibrosas que discurren entre las células malignas en muestras de biopsia de DLBCL, lo que concuerda con su papel en la formación de la matriz extracelular. En cambio, los productos de proteína de los otros tres genes de estromal-1 (*MP9*, *SPARC* y *CTGF*) se localizaban principalmente en células histiocíticas que se infiltraban en las biopsias de DLBCL. Mediante inmunofluorescencia, *SPARC* y *CTGF* se localizaban conjuntamente con *CD68*, que es un marcador para células del linaje monocítico. Tal como se esperaba para un producto génico de estromal-1, los niveles de proteína *SPARC* estaban asociados con supervivencia global favorable (figura 5A).

La firma de estromal-1 incluye genes que se expresan de manera coordinada en muchos tejidos mesenquimales normales, la mayoría de los cuales codifican para proteínas que forman o modifican la matriz extracelular. La localización de fibronectina en hebras fibrosas insinuada entre las células de linfoma malignas sugiere que la firma de estromal-1 refleja la naturaleza fibrótica de muchos tumores de DLBCL. Esta reacción fibrótica puede relacionarse con otro componente de la firma de estromal-1, *CTGF*, que participa en muchas respuestas y enfermedades fibróticas, y que promueve el crecimiento tumoral y la metástasis de cánceres epiteliales (Shi-Wen *et al.*, *Cytokine Growth Factor Rev.*, 19: 133-44 (2008)).

Los resultados anteriores también indican que la firma de estromal-1 refleja una reacción del huésped rica en monocitos al linfoma que está asociada con la deposición abundante de matriz extracelular. Tumores con alta expresión de la firma de estromal-1 estaban infiltrados por células del linaje mielóide, que incluyen células que se han implicado en la patogénesis cánceres epiteliales, incluyendo macrófagos asociados a tumores, células supresoras derivadas del linaje mielóide y monocitos que expresan *Tie2* (revisado en Wels *et al.*, *Genes Dev.*, 22: 559-74 (2008)). En modelos animales, estas células del linaje mielóide promueven la invasión de células tumorales secretando metaloproteinasas de la matriz tales como *MMP9*, suprimen respuestas inmunitarias de células T e inician la angiogénesis.

Varios genes de firma de estromal-2 codifican para marcadores bien conocidos de células endoteliales. Estos incluyen factor de von-Willebrand (*VWF*) y *CD31* (*PECAM1*), así como otros genes expresados específicamente en el endotelio tales como *EGFL7*, *MMRN2*, *GPR116* y *SPARCL* (tabla 1). Esta firma también incluye genes que codifican para reguladores clave de la angiogénesis, tales como, por ejemplo, *KDR* (receptor de *VEGF* 2); *Grb10*, que media en la señalización de *KDR*; integrina alfa 9, que potencia la señalización de *VEGF*; *TEK*, la tirosina cinasa receptora para la citocina angiopoyetina; *ROBO4*, una molécula de guiado molecular específica del endotelio que regula la angiogénesis; y *ERG*, un factor de transcripción requerido para la formación de tubos endoteliales. Los genes de firma de estromal-2 *CAV1*, *CA V2* y *EHD2* codifican para componentes de caveolas, que son estructuras especializadas de la membrana plasmática que son abundantes en células endoteliales y se requieren para la angiogénesis (Frank *et al.*, *Arterioscler. Thromb. Vasc. Biol.*, 23: 1161-8 (2003); Woodman *et al.*, *Am. J. Pathol.*, 162: 2059-68 (2003)). Aunque la firma de estromal-2 incluye un gran número de genes expresados en células endoteliales, otros genes se expresan exclusivamente en adipocitos, incluyendo *ADIPOQ*, *FABP4*, *RBP4* y *PLIN*.

Se realizaron pruebas cuantitativas para determinar si la expresión de la firma de estromal-2 en relación con la firma de estromal-1 (es decir, alta puntuación estromal) está relacionada con alta densidad de vasos sanguíneos tumorales, dada la conexión entre muchos genes de firma de estromal-2 y angiogénesis. Más específicamente, se restaron los promedios de firma de estromal-1 del promedio de firma de estromal-2 para obtener de ese modo una puntuación estromal para cada biopsia. Las pruebas mostraron una medida cuantitativa de la densidad de vasos sanguíneos que se correlacionaba significativamente con la puntuación estromal ( $r=0,483$ ,  $p=0,019$ ) (véanse las figuras 5B y 5C), de manera que densidades de vasos sanguíneos superiores se correlacionaban con puntuaciones estromales superiores.

Por tanto, las firmas de expresión génica de estromal-1 y estromal-2 reflejan el carácter de las células no malignas en tumores de DLBCL, y la firma de estromal-2 puede representar un "cambio angiogénico" en el que la progresión de una lesión hiperplásica a un tumor completamente maligno va acompañada por la formación de nuevos vasos sanguíneos (Hanahan *et al.*, *Cell*, 86: 353-64 (1996)). Tumores de DLBCL con alta expresión relativa de la firma de estromal-2 estaban asociados con un aumento de la densidad de vasos sanguíneos tumorales y supervivencia adversa. La infiltración de macrófagos significativa en algunos tumores de DLBCL puede predisponer a angiogénesis puesto que, en modelos experimentales, se acumulan macrófagos asociados a tumores antes del cambio angiogénico y se requieren para que se produzca el cambio (Lin *et al.*, *Cancer Res.*, 66: 11238-46 (2006)). Adicionalmente, *CXCL12* (*SDF-1*), un componente de la firma de estromal-2, es una quimiocina secretada o bien por fibroblastos o bien por células endoteliales que puede promover la angiogénesis reclutando células precursoras endoteliales *CXCR4+* de la médula ósea (Orimo *et al.*, *Cell*, 121: 335-48 (2005)). Además, un antagonista de la angiogénesis, trombospondina-2 (Kazerounian *et al.*, *Cell Mol. Life Sci.*, 65: 700-12 (2008)), es un componente de la firma de estromal-1, lo que puede explicar por qué muchos tumores con expresión relativa baja de esta firma tenían una densidad elevada de vasos sanguíneos. Además, la expresión de genes asociados a adipocitos en tumores de DLBCL con alta expresión de la firma de estromal-2 puede desempeñar un papel en la angiogénesis puesto que algunas células en el tejido adiposo pueden tener el potencial para diferenciarse en células endoteliales (Planat-Benard *et al.*, *Circulation*, 109: 656-63 (2004)). Alternativamente, la expresión de genes asociados a tejido adiposo puede reflejar el reclutamiento de células madre mesenquimales derivadas de la médula ósea, que se dirigen eficazmente a tumores (Karnoub *et al.*, *Nature*, 449: 557-63 (2007)) y pueden estabilizar vasos sanguíneos recién formados (Au *et al.*, *Blood*, 111: 4551-4558 (2008)).

Los resultados anteriores indican que las firmas de genes de estromal-1 y estromal-2 pueden usarse para generar una puntuación estromal que se correlaciona con un aumento de la densidad de vasos sanguíneos. Por tanto, la puntuación estromal puede usarse para determinar si un paciente con DLBCL es probable que se beneficie de la administración de la terapia antiangiogénica (sola, o conjuntamente con otro régimen terapéutico para DLBCL).

5 El uso de los términos “un” y “una” y “el/la” y referentes similares en el contexto de la descripción de la invención (especialmente en el contexto de las siguientes reivindicaciones) debe interpretarse que cubre tanto el singular como el plural, a menos que se indique otra cosa en el presente documento o se contradiga claramente por el contexto. Los términos “que comprende”, “que tiene”, “que incluye” y “que contiene” deben interpretarse como términos de extremos abiertos (es decir, que significan “que incluye, pero sin limitarse a,”) a menos que se indique otra cosa. La  
10 mención de intervalos de valores en el presente documento pretende meramente servir como método abreviado para referirse individualmente a cada valor separado que se encuentra dentro del intervalo, a menos que se indique otra cosa en el presente documento, y cada valor separado se incorpora en la memoria descriptiva como si se mencionase individualmente en el presente documento. Todos los métodos descritos en el presente documento pueden realizarse en cualquier orden adecuado a menos que se indique otra cosa en el presente documento o por lo  
15 demás se contradiga claramente por el contexto. El uso de todos y cada uno de los ejemplos, o expresiones a modo de ejemplo (por ejemplo, “tal como”) proporcionados en el presente documento, pretende meramente esclarecer mejor la invención. Ninguna expresión en la memoria descriptiva debe interpretarse que indica que cualquier elemento no reivindicado sea esencial para la práctica de la invención.

20 En el presente documento se describen realizaciones preferidas de esta invención, incluyendo el mejor modo conocido por los inventores para llevar a cabo la invención. Pueden resultar evidentes variaciones de esas realizaciones preferidas a los expertos habituales en la técnica tras la lectura de la descripción anterior. Los inventores esperan que los expertos en la técnica empleen tales variaciones según sea apropiado.

**REIVINDICACIONES**

1. Método de predicción del resultado de supervivencia basándose en una o más muestras de linfoma difuso de células B grandes (DLBCL), método que comprende:
  - 5 a) aislar productos de expresión génica de una o más muestras de biopsia de DLBCL de un sujeto tratado con un régimen terapéutico que incluye quimioterapia y rituximab (R-CHOP);
  - b) obtener un perfil de expresión génica a partir de los productos de expresión génica, en el que el perfil de expresión comprende un nivel de expresión para cada uno de al menos 31 de los genes enumerados en la tabla 1 para la firma de expresión génica de células B de centro germinal (GCB) y al menos 226 de los genes enumerados en la tabla 1 para la firma de expresión génica estromal-1;
  - 10 c) determinar un valor de firma de GCB y un valor de firma estromal-1 a partir del perfil de expresión génica, en el que el valor de firma de GCB es un promedio de los niveles de expresión de los al menos 31 genes enumerados en la tabla 1 para la firma de expresión génica de GCB y el valor de firma de estromal-1 es un promedio de los niveles de expresión de los al menos 226 genes enumerados en la tabla 1 para la firma de expresión génica estromal-1;
  - 15 d) calcular una puntuación de factor de predicción de supervivencia =  $A - [(x) \cdot (\text{el valor de firma de GCB})] - [(y) \cdot (\text{el valor de firma de estromal-1})]$ , en la que A es un término de desviación, (x) es un factor de escala de entre 0,200 y 0,625, e (y) es un factor de escala de entre 0,800 y 1,250, y
  - 20 e) determinar el resultado de supervivencia para el sujeto basándose en la puntuación de factor de predicción de supervivencia calculada en la etapa d) para cualquier valor igual fijo de A y para cualquier valor igual fijo de (x) e (y), en el que una puntuación inferior de factor de predicción de supervivencia indica un resultado de supervivencia más favorable y una puntuación superior de factor de predicción de supervivencia indica un resultado de supervivencia menos favorable para el sujeto.
2. Método según la reivindicación 1, en el que el método comprende:
  - 25 b) obtener un perfil de expresión génica a partir de los productos de expresión génica, en el que el perfil de expresión comprende un nivel de expresión para cada uno de al menos 31 de los genes enumerados en la tabla 1 para la firma de expresión de GCB, al menos 226 de los genes enumerados en la tabla 1 para la firma de expresión génica estromal-1 y al menos 58 de los genes enumerados en la tabla 1 para la firma de expresión génica estromal-2;
  - 30 c) determinar un valor de firma de GCB, un valor de firma estromal-1 y un valor de firma génica estromal-2 a partir del perfil de expresión génica, en el que el valor de firma de GCB es un promedio de los niveles de expresión de los al menos 31 genes enumerados en la tabla 1 para la firma de expresión génica de GCB y el valor de firma estromal-1 es un promedio de los niveles de expresión de los al menos 226 genes enumerados en la tabla 1 para la firma de expresión génica de estromal-1 y el valor de firma estromal-2 es un promedio de los niveles de expresión de los al menos 58 genes enumerados en la tabla 1 para la firma de expresión génica estromal-2; y
  - 35 d) calcular una puntuación de factor de predicción de supervivencia =  $A - [(x) \cdot (\text{el valor de firma de GCB})] - [(y) \cdot (\text{el valor de firma de estromal-1})] + [(z) \cdot (\text{el valor de firma de estromal-2})]$ , en la que A es un término de desviación, (x) es un factor de escala de entre 0,200 y 0,625, (y) es un factor de escala de entre 0,800 y 1,250 y (z) es un factor de escala de entre 0,450 y 0,900, y
  - 40 e) determinar el resultado de supervivencia para el sujeto basándose en la puntuación de factor de predicción de supervivencia calculada en la etapa d) para cualquier valor igual fijo de A y para cualquier valor igual fijo de (x), (y) y (z), en el que una puntuación inferior de factor de predicción de supervivencia indica un resultado de supervivencia más favorable y una puntuación superior de factor de predicción de supervivencia indica un resultado de supervivencia menos favorable para el sujeto.
- 45 3. Método según la reivindicación 2, en el que el método comprende
  - d) calcular una puntuación de factor de predicción de supervivencia usando la ecuación: puntuación de factor de predicción de supervivencia =  $8,11 - [0,419 \cdot (\text{el valor de firma de GCB})] - [1,015 \cdot (\text{el valor de firma de estromal-1})] + [0,675 \cdot (\text{el valor de firma de estromal-2})]$
- 50 4. Método de predicción del resultado de supervivencia basándose en una o más muestras de linfoma difuso de células B grandes (DLBCL), método que comprende:
  - a) aislar productos de expresión génica de una o más muestras de biopsia de DLBCL de un sujeto tratado con un régimen terapéutico que incluye quimioterapia y Rituximab (R-CHOP);
  - b) obtener un perfil de expresión génica a partir de los productos de expresión génica, en el que el perfil de

expresión génica comprende un nivel de expresión para cada uno de al menos 31 de los genes enumerados en la tabla 1 para la firma de expresión génica de células B de centro germinal (GCB), al menos 226 de los genes enumerados en la tabla 1 para la firma de expresión génica estromal-1 y al menos 58 de los genes enumerados en la tabla 1 para la firma de expresión génica estromal-2;

5 c) determinar un valor de firma de GCB, un valor de firma estromal-1 y un valor de firma estromal-2 a partir del perfil de expresión génica, en el que el valor de firma de GCB es un promedio de los niveles de expresión de los al menos 31 genes enumerados en la tabla 1 para la firma de expresión génica de GCB y el valor de firma estromal-1 es un promedio de los niveles de expresión de los al menos 226 genes enumerados en la tabla 1 para la firma de expresión génica de estromal-1 y el valor de firma estromal-2 es un promedio de los niveles de expresión de los al menos 58 genes enumerados en la tabla 1 para la firma de expresión génica estromal-2;

10 d) calcular una puntuación de factor de predicción de supervivencia usando la ecuación: puntuación de factor de predicción de supervivencia =  $A - [(x) \cdot (\text{el valor de firma de GCB})] - [(y) \cdot (\text{el valor de firma de estromal-1})] + [(z) \cdot (\text{el valor de firma de estromal-2})]$ , en la que A es un término de desviación y (x) es un factor de escala de entre 0,200 y 0,625, (y) es un factor de escala de entre 0,800 y 1,250 y (z) es un factor de escala de entre 0,450 y 0,900; y

15 e) calcular la probabilidad de un resultado de supervivencia para el sujeto más allá de una cantidad de tiempo t tras el tratamiento para DLBCL, en el que la probabilidad del resultado de supervivencia del sujeto P(SO) se calcula usando la ecuación:  $P(SO) = SO_0(t)^{\exp((s) \cdot \text{puntuación de factor de predicción de supervivencia})}$ , en la que SO<sub>0</sub>(t) es la probabilidad del resultado de supervivencia, que corresponde al valor de tiempo más largo menor que t en una curva de resultado de supervivencia, y en la que (s) es un factor de escala de entre 0,970 y 0,980.

5. Método según la reivindicación 4, en el que el resultado de supervivencia es supervivencia global y el método comprende:

25 d) calcular una puntuación de factor de predicción de supervivencia usando la ecuación: puntuación de factor de predicción de supervivencia =  $[0,419 \cdot (\text{el valor de firma de GCB})] - [1,015 \cdot (\text{el valor de firma estromal-1})] + [0,675 \cdot (\text{el valor de firma estromal-2})]$ ; y

30 e) calcular la probabilidad de supervivencia global tras el tiempo t para el sujeto, en el que la probabilidad de supervivencia global del sujeto P(OS) se calcula usando la ecuación:  $P(OS) = OS_0(t)^{\exp(\text{puntuación de factor de predicción de supervivencia})}$ , en la que OS<sub>0</sub>(t) es la probabilidad de supervivencia global, que corresponde al valor de tiempo más largo menor que t en una curva de supervivencia global.

6. Método según la reivindicación 4, en el que el resultado de supervivencia es supervivencia libre de progresión y el método comprende:

35 d) calcular una puntuación de factor de predicción de supervivencia usando la ecuación: puntuación de factor de predicción de supervivencia =  $8,11 - [0,419 \cdot (\text{el valor de firma de GCB})] - [1,015 \cdot (\text{el valor de firma estromal-1})] + [0,675 \cdot (\text{el valor de firma estromal-2})]$ ; y

40 e) calcular la probabilidad de supervivencia libre de progresión tras el tiempo t para el sujeto, en el que la probabilidad de supervivencia libre de progresión del sujeto P(PFS) se calcula usando la ecuación  $P(PFS) = F0(t)^{\exp(0,976 \cdot \text{puntuación de factor de predicción de supervivencia})}$ , en la que F0(t) es la probabilidad de supervivencia libre de progresión, que corresponde al tiempo más largo menor que t en una curva de supervivencia, y en la que en el que el factor de escala (s) = 0,976.

7. Método según una cualquiera de las reivindicaciones 4-6, en el que la probabilidad de un resultado de supervivencia es la probabilidad de un resultado de supervivencia para el sujeto más allá de una cantidad de tiempo t tras el tratamiento para DLBCL que incluye quimioterapia y la administración de Rituximab.

45 8. Método según una cualquiera de las reivindicaciones 4-7, en el que el método incluye además proporcionar al sujeto la probabilidad calculada del resultado de supervivencia tras el tiempo t.

9. Matriz diana que comprende al menos una sonda o conjunto de sondas para cada gen en (a) una firma de expresión génica de células B de centro germinal (GCB) que comprende al menos 31 de los genes de firma de GCB enumerados en la tabla 1, (b) una firma de expresión génica de estromal-1 que comprende al menos 226 de los genes de firma de estromal-1 enumerados en la tabla 1 y (c) una firma de expresión génica de estromal-2 que comprende al menos 58 de los genes de firma de estromal-2 enumerados en la tabla 1, en el que el alineamiento comprende sondas para menos de 10.000 genes.

50 10. Matriz diana según la reivindicación 9, en el que la firma de expresión de GCB comprende al menos 37 de los genes de firma de GCB enumerados en la tabla 1, en el que la firma de expresión génica estromal-1 comprende al menos 270 de los genes de firma estromal-1 enumerados en la tabla 1, y en el que la firma

55

de expresión génica estromal-2 comprende al menos 68 de los genes de firma estromal-2 enumerados en la tabla 1.

- 5
11. Método según la reivindicación 1, en el que el valor de firma de GCB es un promedio de los niveles de expresión de al menos 37 genes enumerados en la tabla 1 para la firma de GCB y el valor de firma estromal-1 es un promedio de los niveles de expresión de al menos 270 genes enumerados en la tabla 1 para la firma estromal-1.
- 10
12. Método según la reivindicación 1, en el que el valor de firma de GCB es un promedio de los niveles de expresión de todos los genes enumerados en la tabla 1 para la firma de GCB y el valor de firma estromal-1 es un promedio de los niveles de expresión de todos los genes enumerados en la tabla 1 para la firma estromal-1.
- 15
13. Método según una cualquiera de las reivindicaciones 2-8, en el que el valor de firma de GCB es un promedio de los niveles de expresión de al menos 37 genes enumerados en la tabla 1 para la firma de GCB, el valor de firma estromal-1 es un promedio de los niveles de expresión de al menos 270 genes enumerados en la tabla 1 para la firma estromal-1 y el valor de firma estromal-2 es un promedio de los niveles de expresión de al menos 68 genes enumerados en la tabla 1 para el valor de firma estromal-2.
- 20
14. Método según una cualquiera de las reivindicaciones 2-8, en el que el valor de firma de GCB es un promedio de los niveles de expresión de todos los genes enumerados en la tabla 1 para la firma de GCB, el valor de firma estromal-1 es un promedio de los niveles de expresión de todos los genes enumerados en la tabla 1 para la firma estromal-1 y el valor de firma estromal-2 es un promedio de los niveles de expresión de todos los genes enumerados en la tabla 1 para el valor de firma estromal-2.

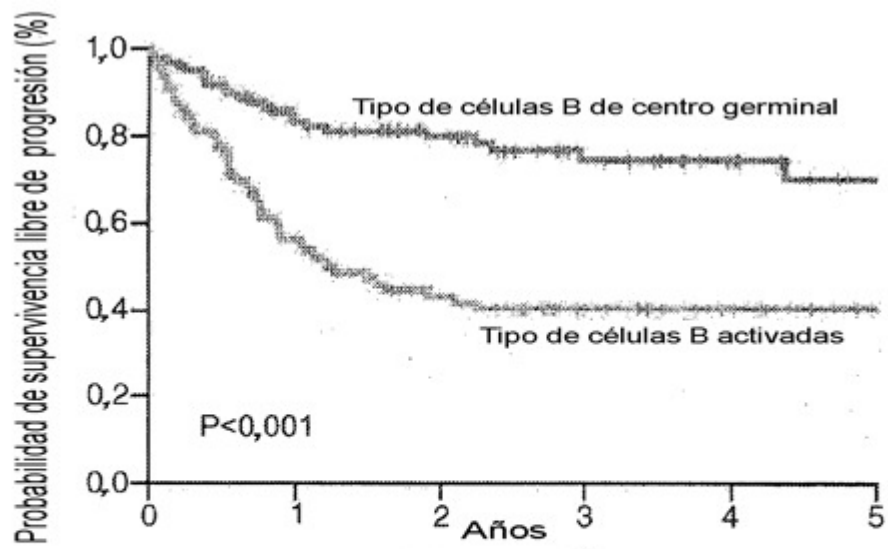


Figura 1A

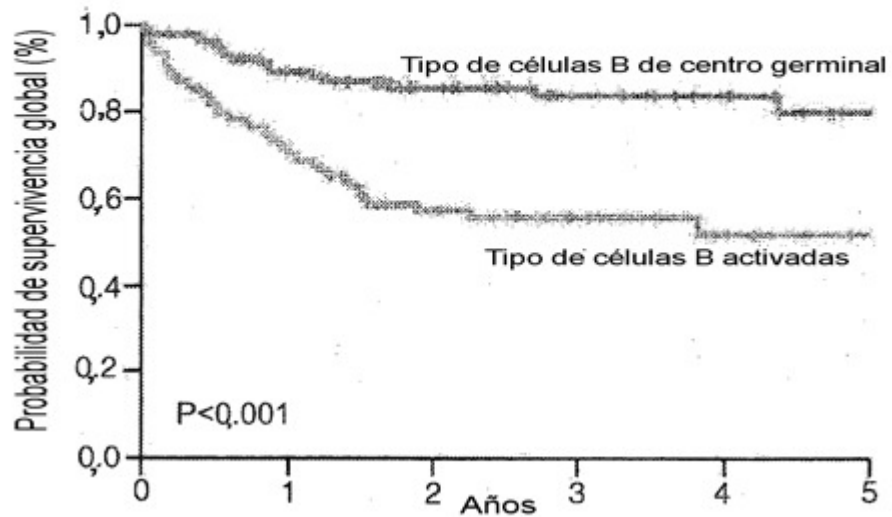


Figura 1B

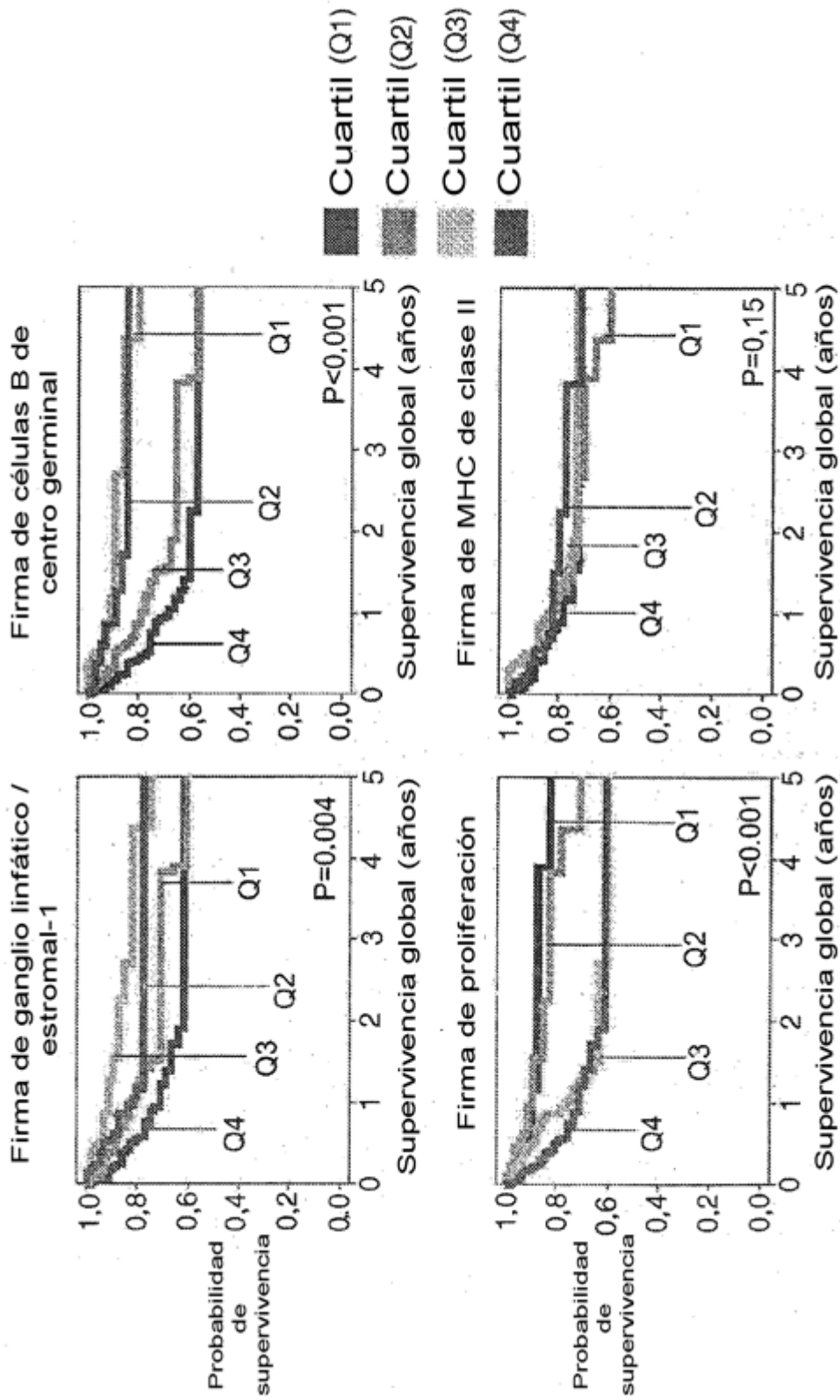


Figura 1C



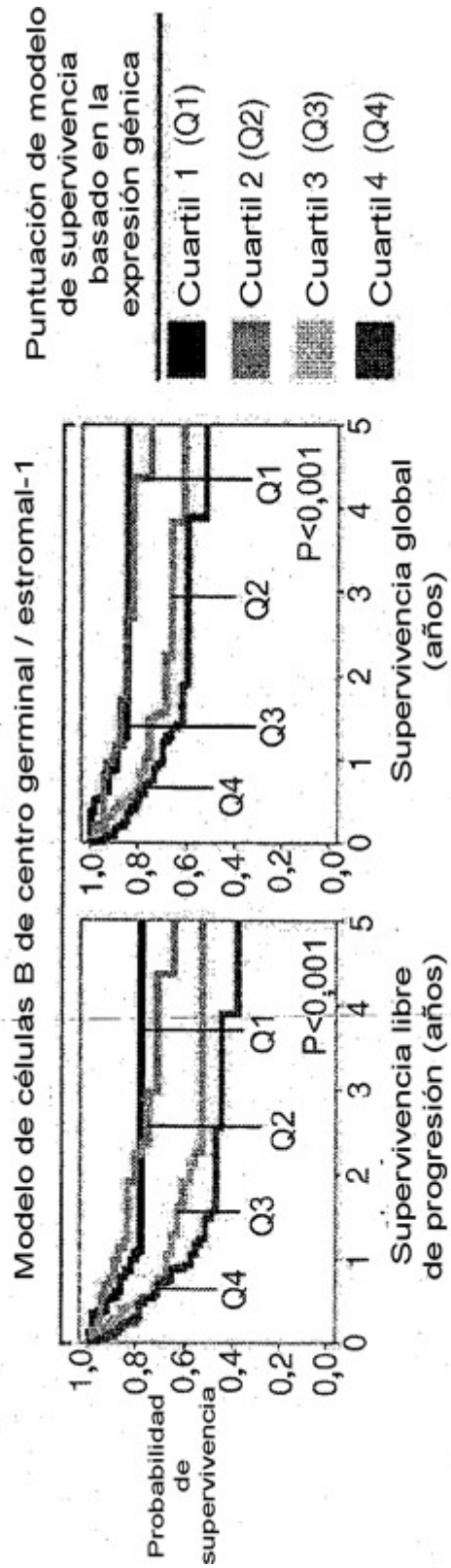


Figura 2A

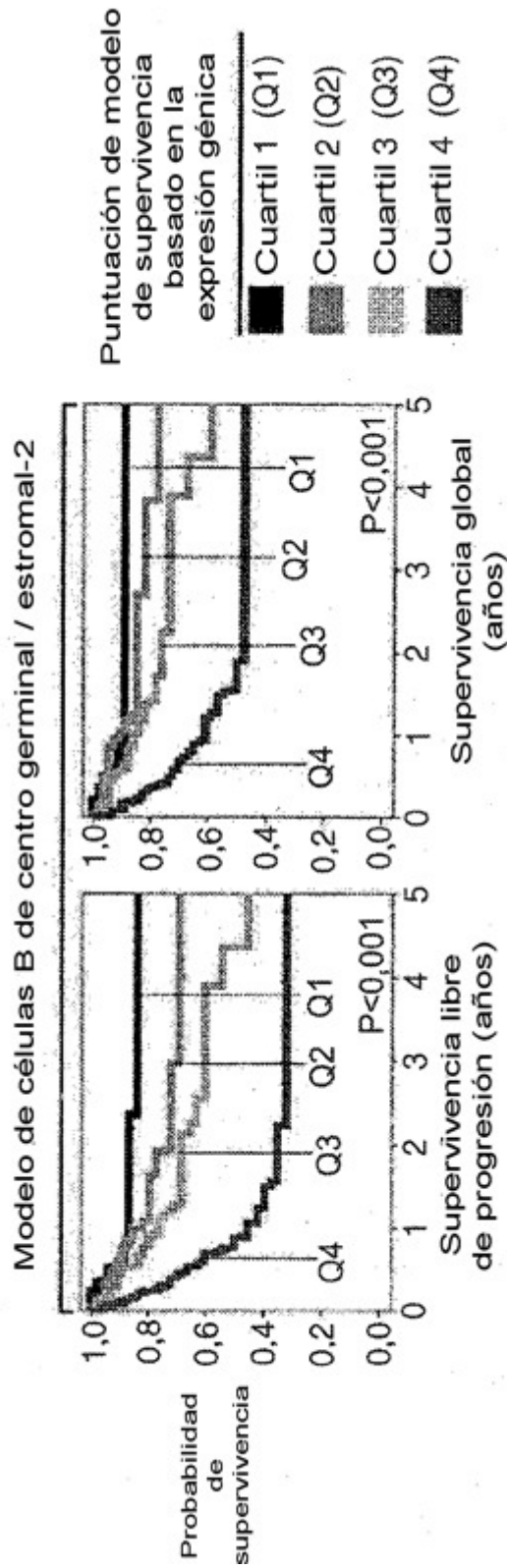


Figura 2B

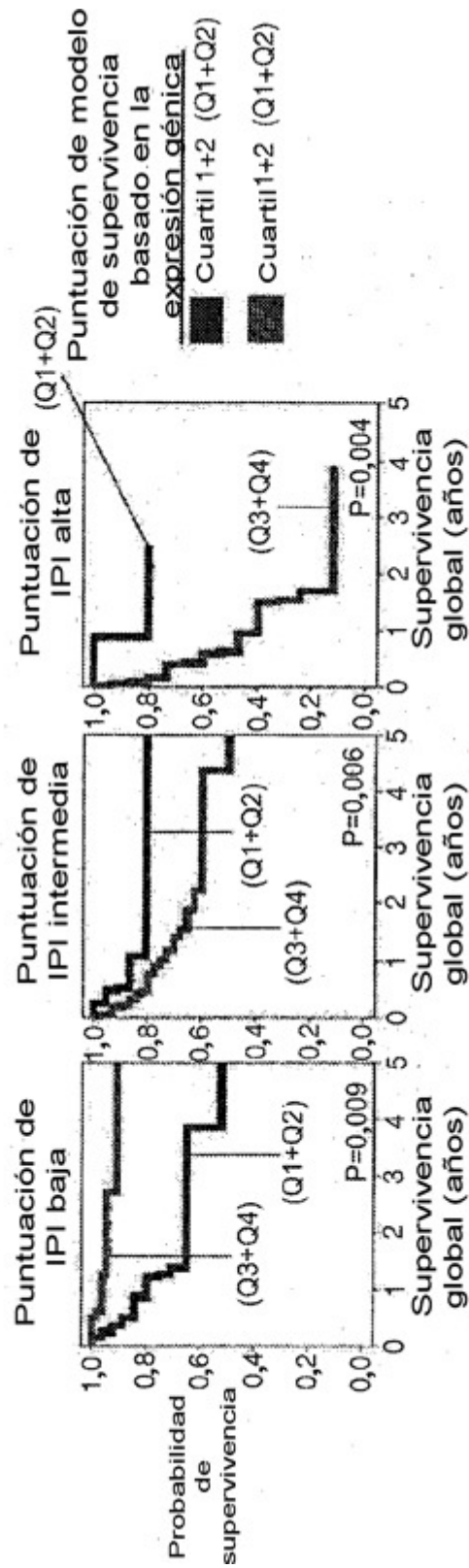
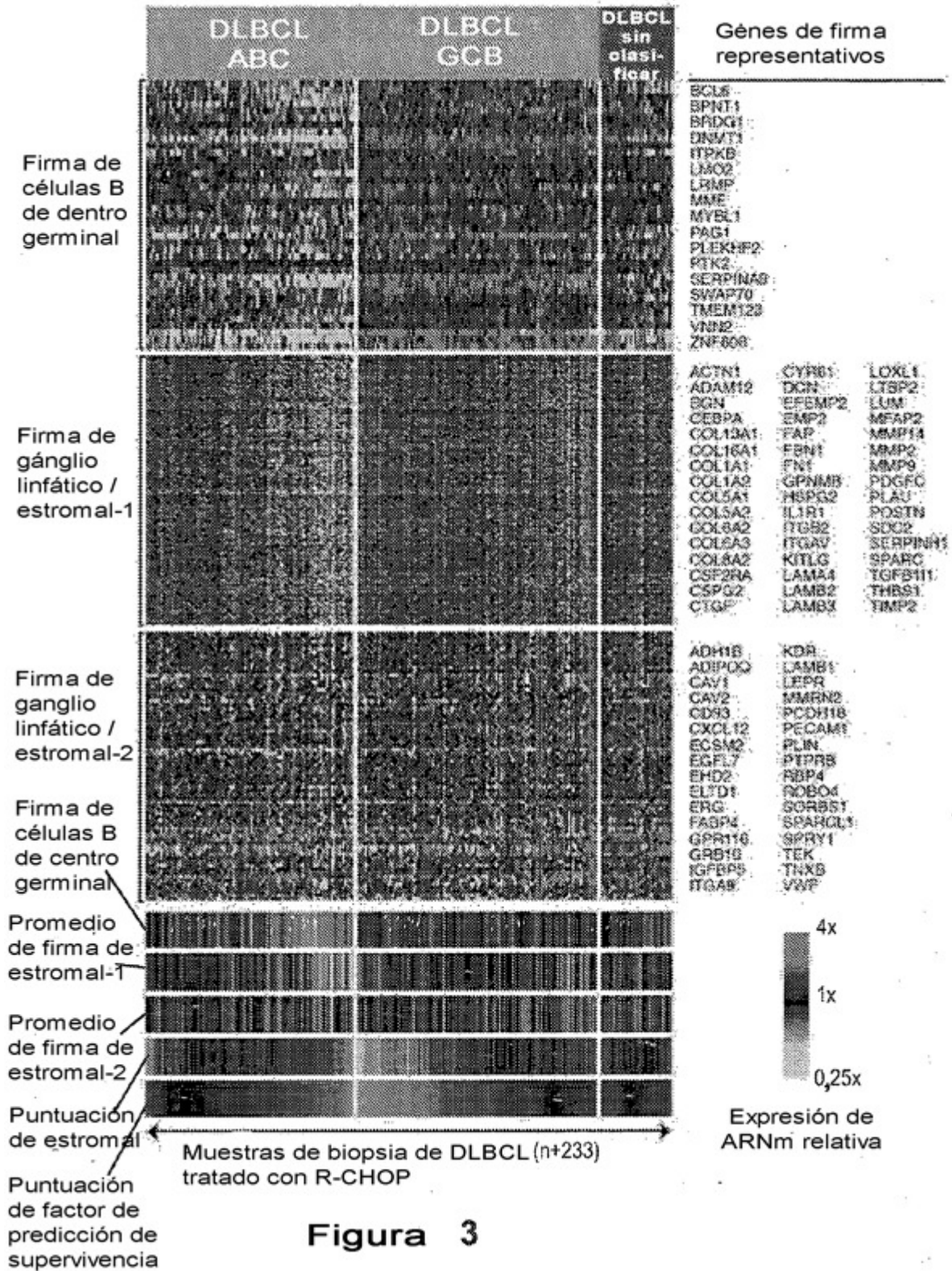


Figura 2C



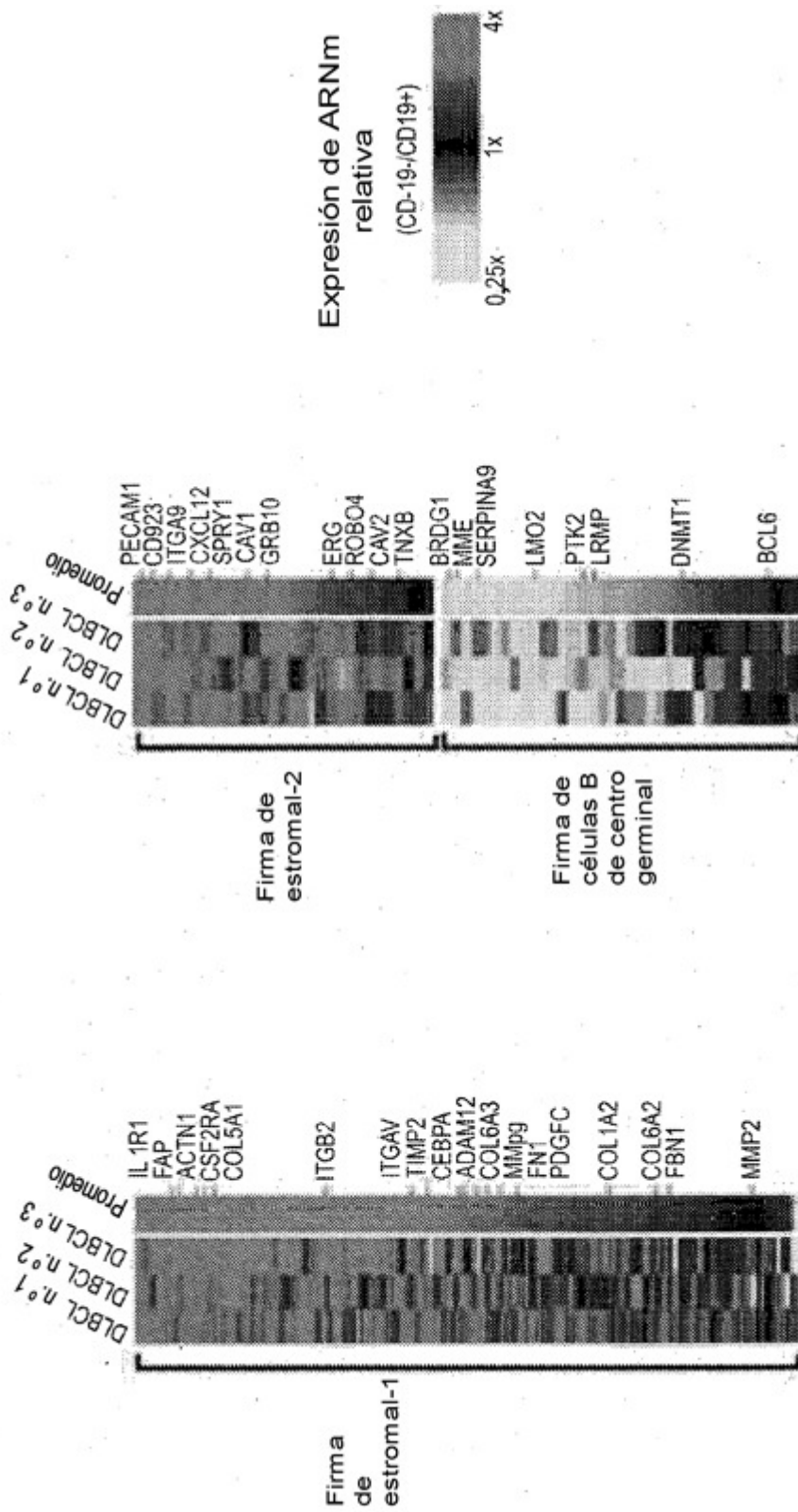


Figura 4A

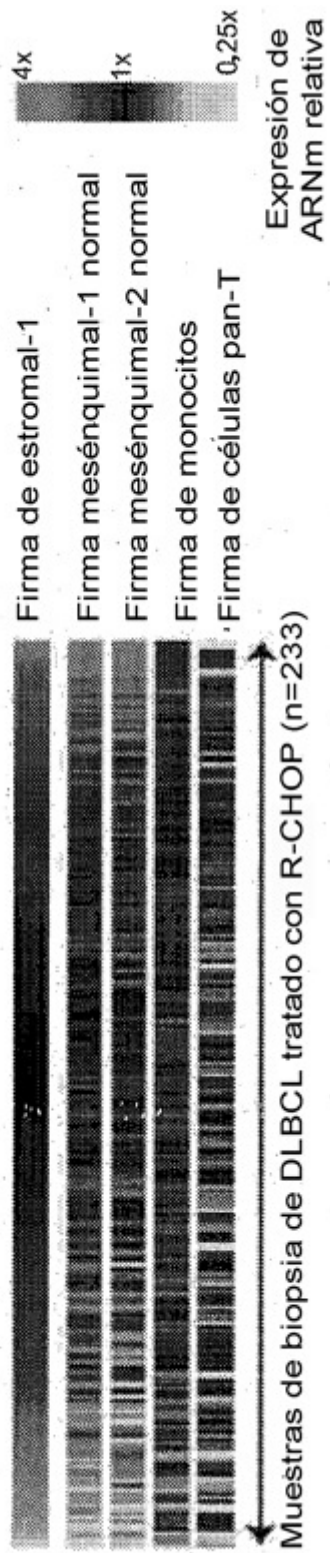


Figura 4B

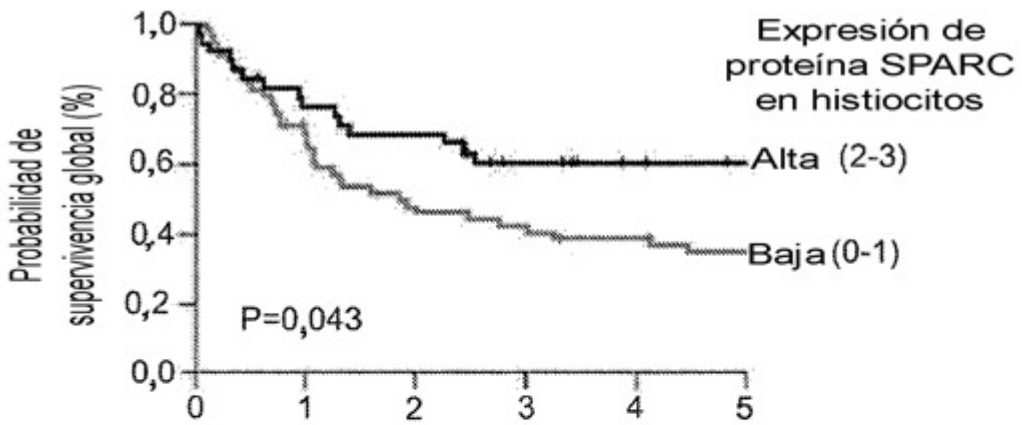


Figura 5A

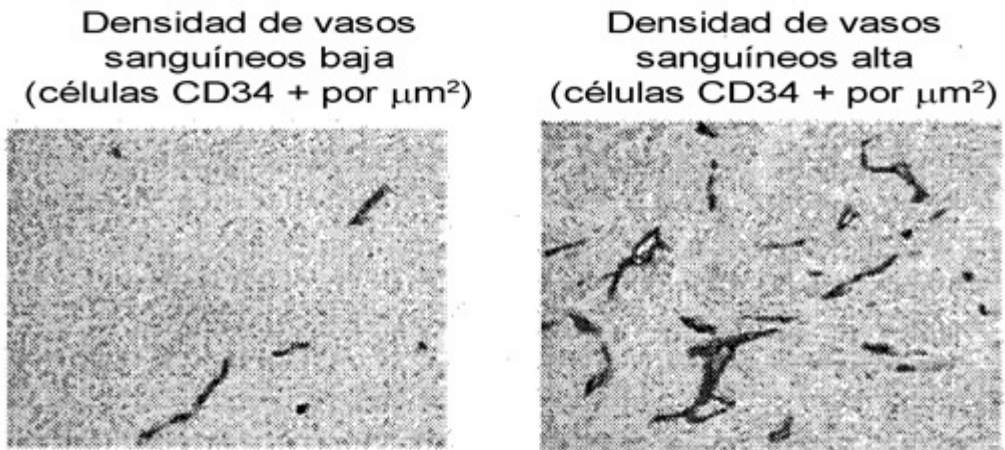


Figura 5B

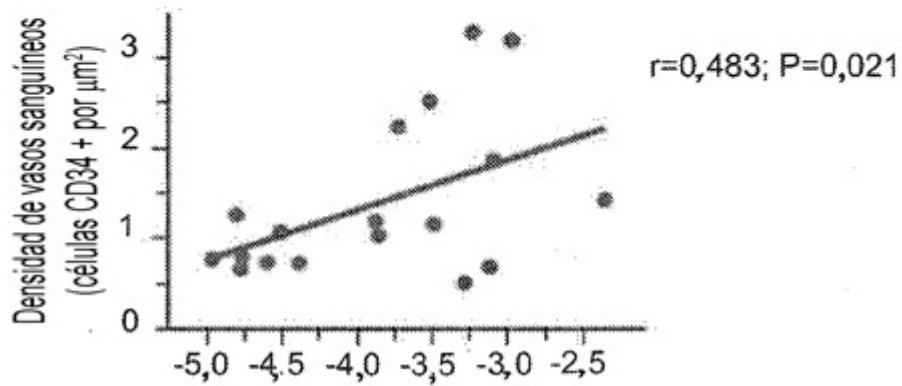


Figura 5C

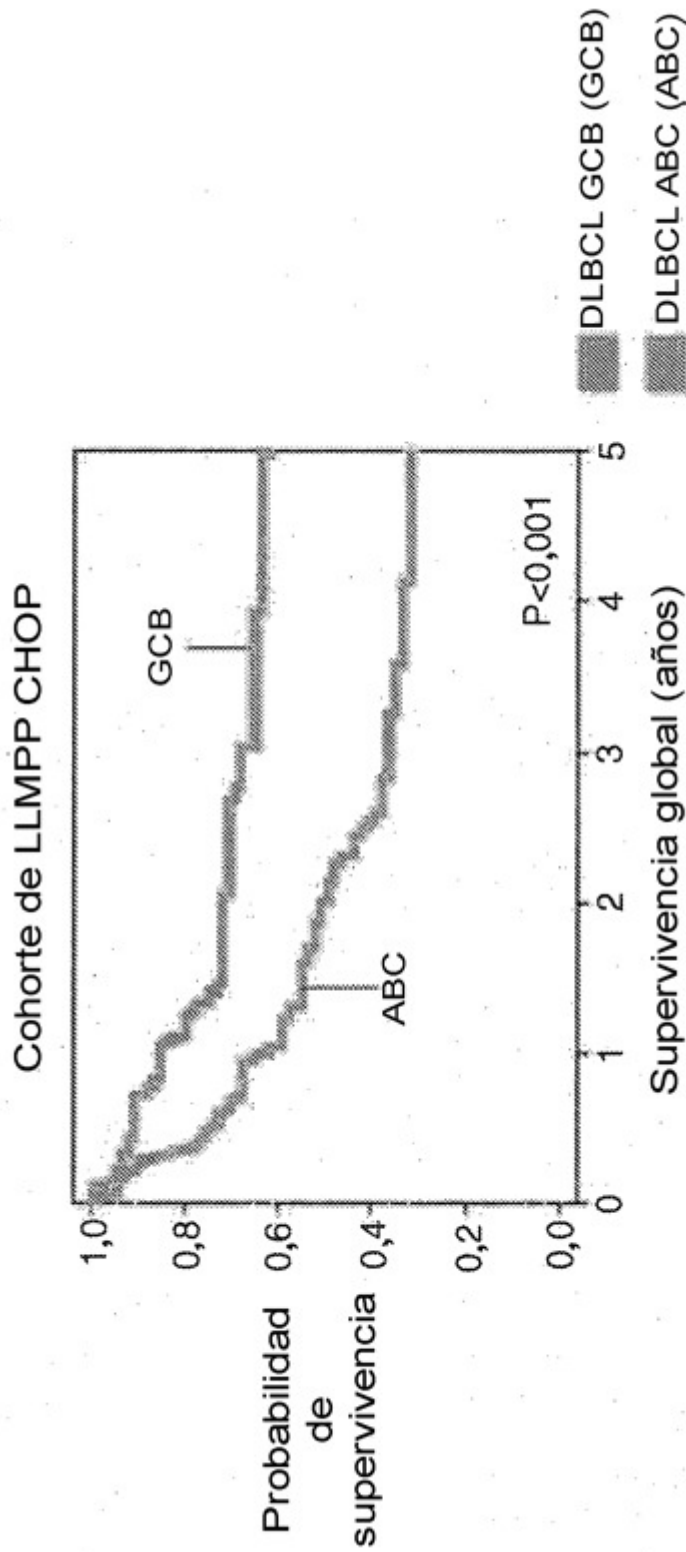


Figura 6A



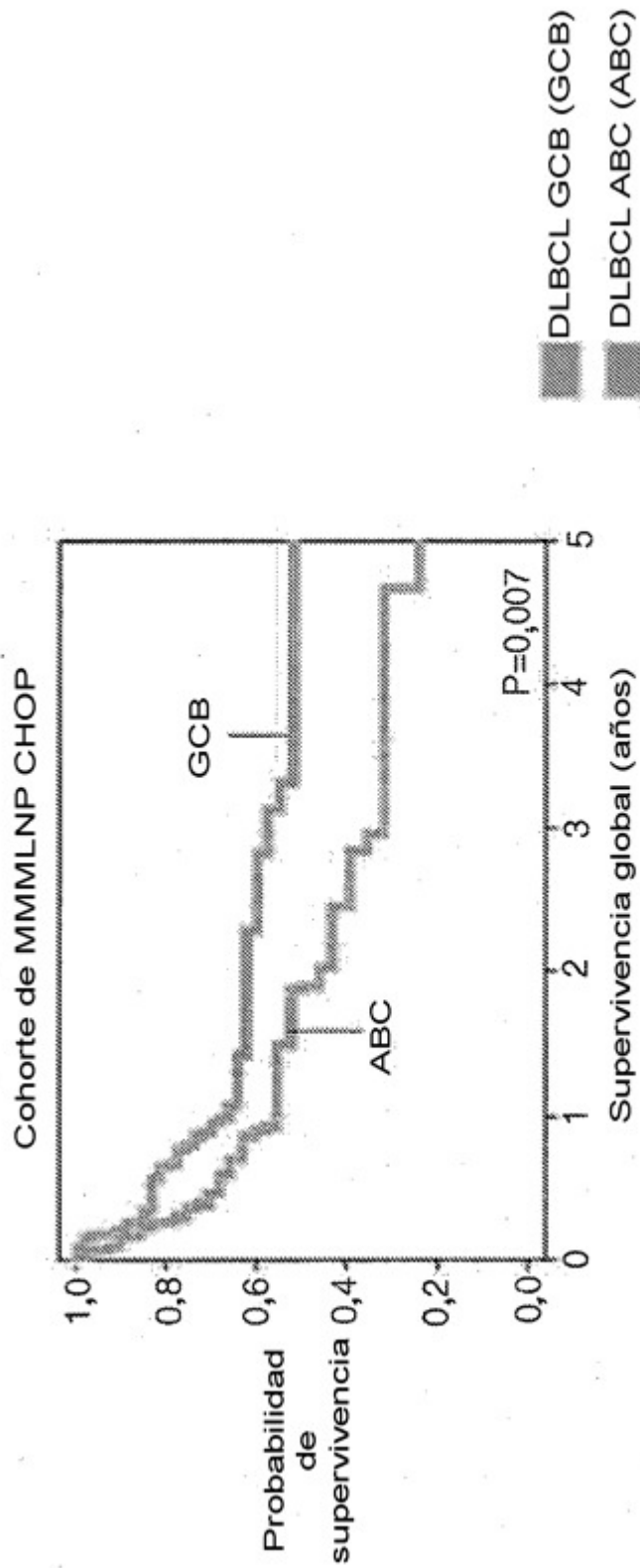


Figura 6B

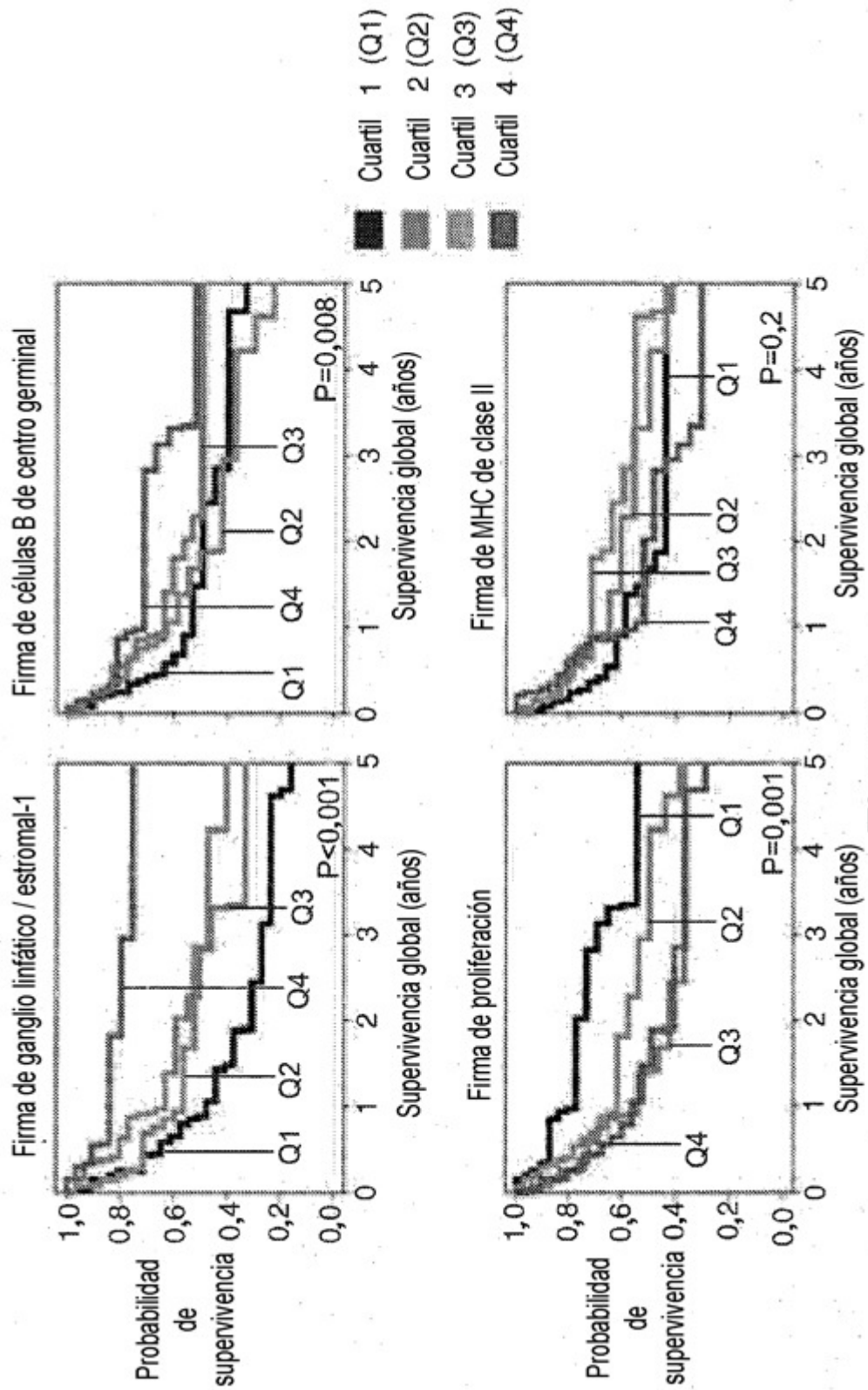


Figura 7

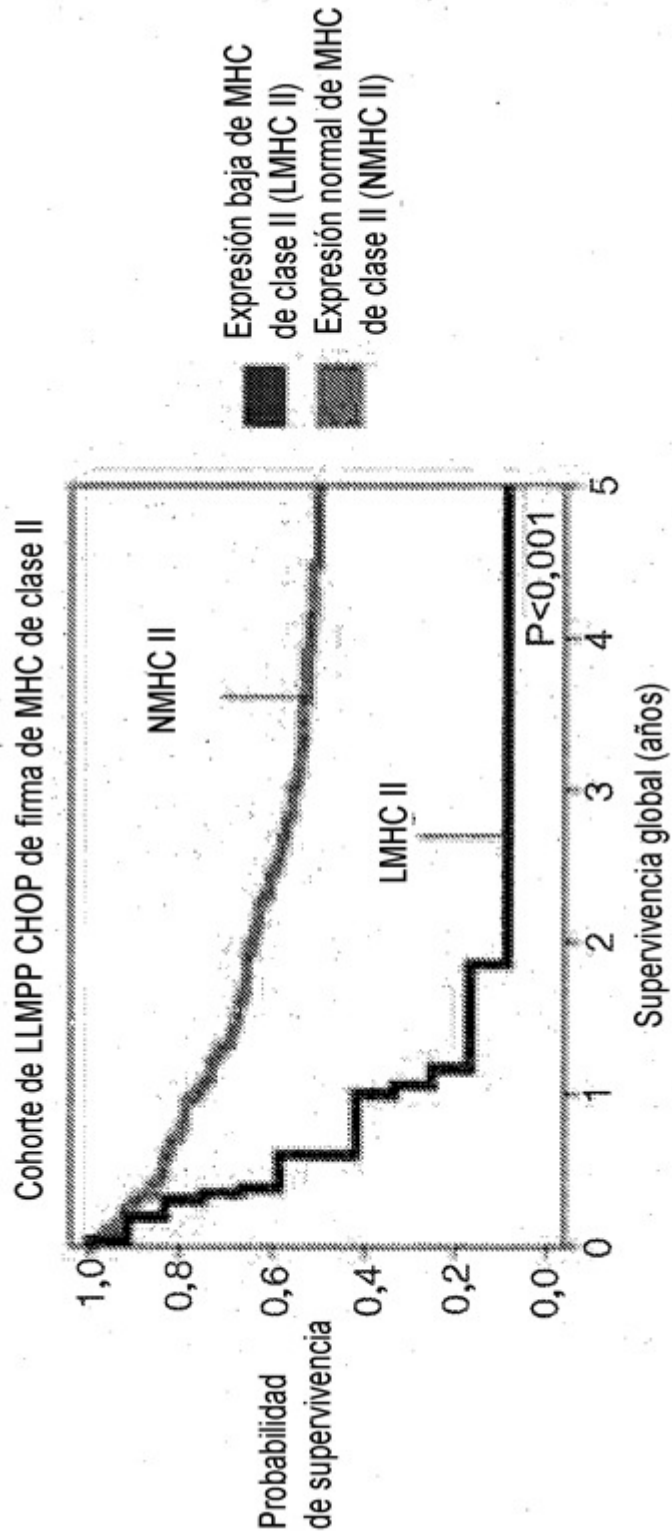


Figura 8A

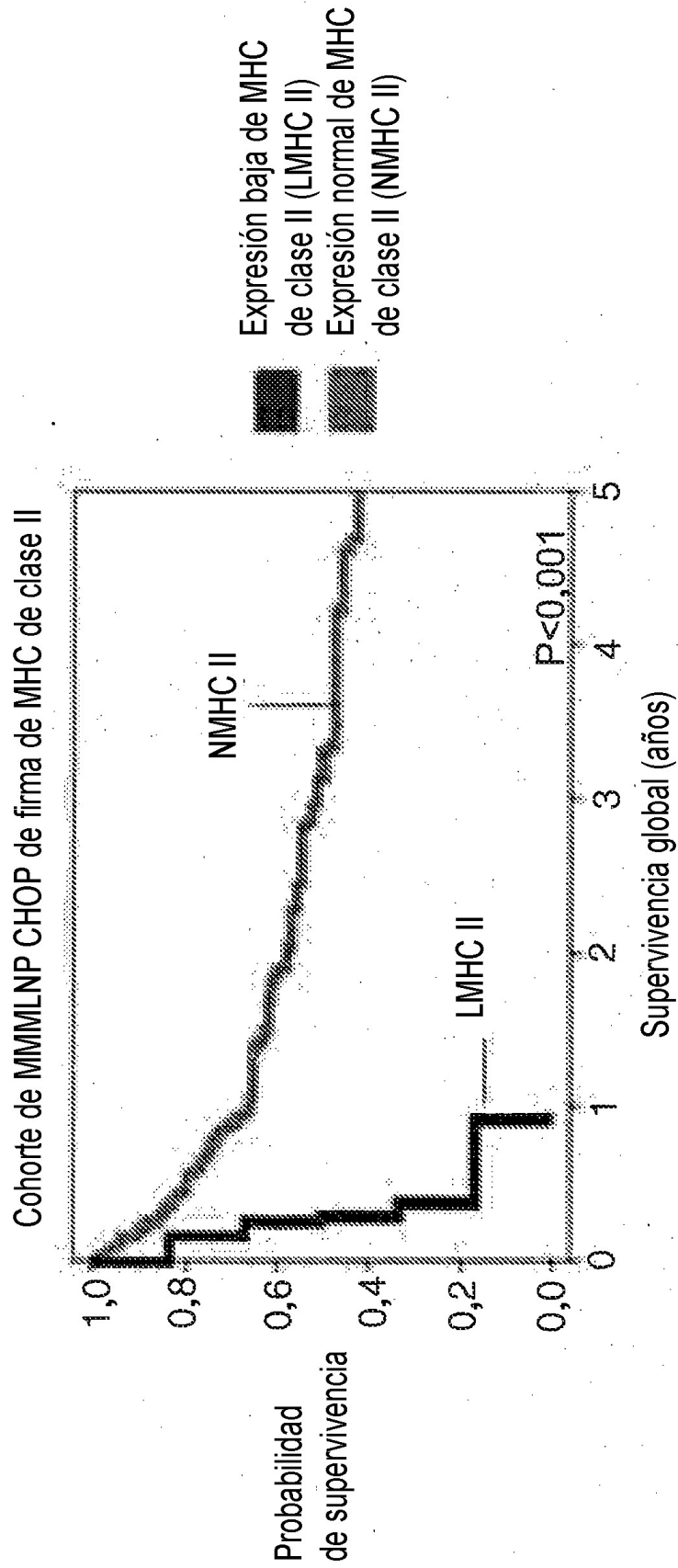


Figura 8B

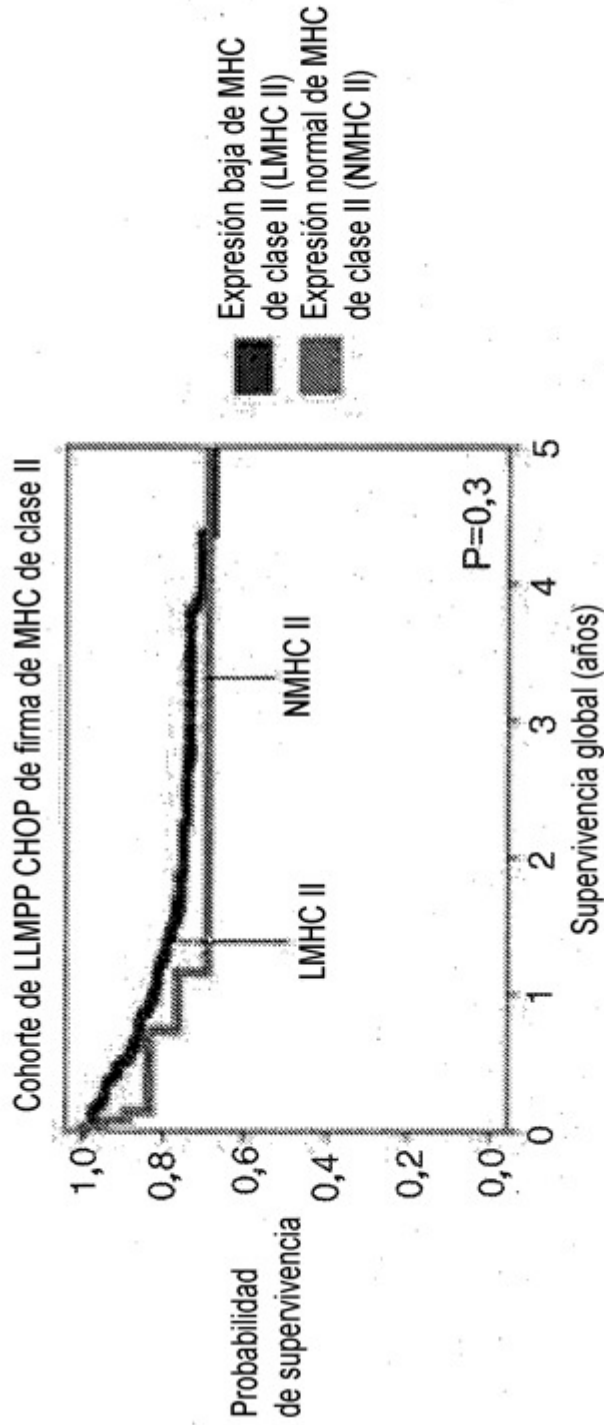


Figura 8C

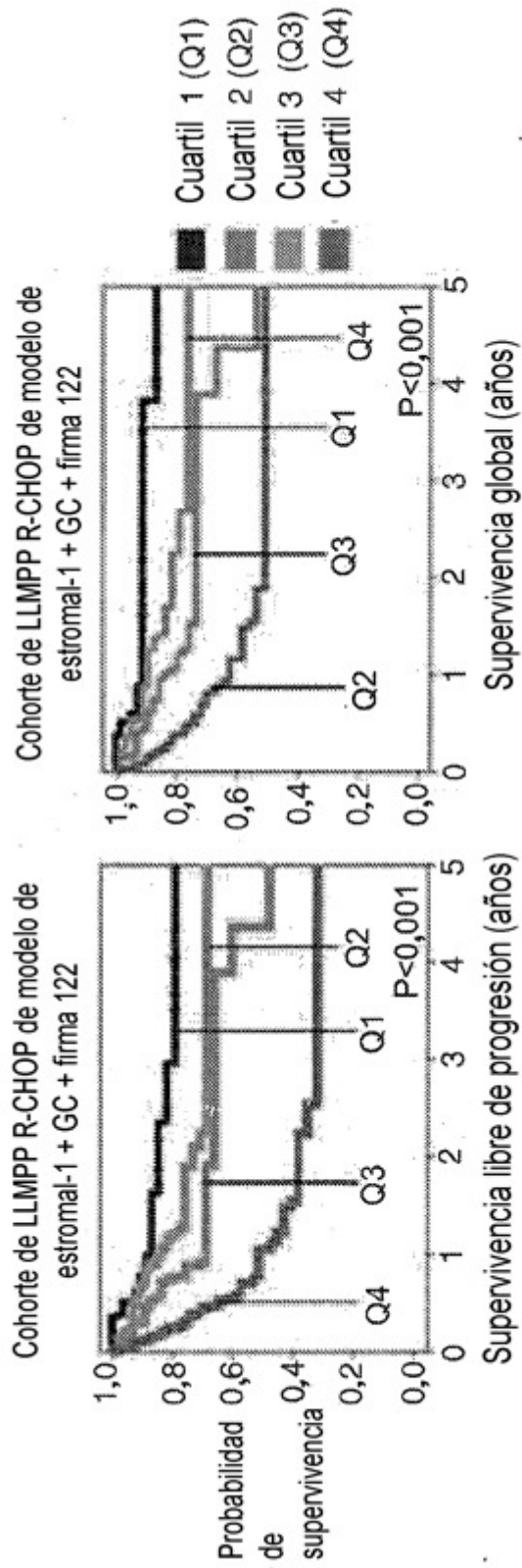


Figura 9A

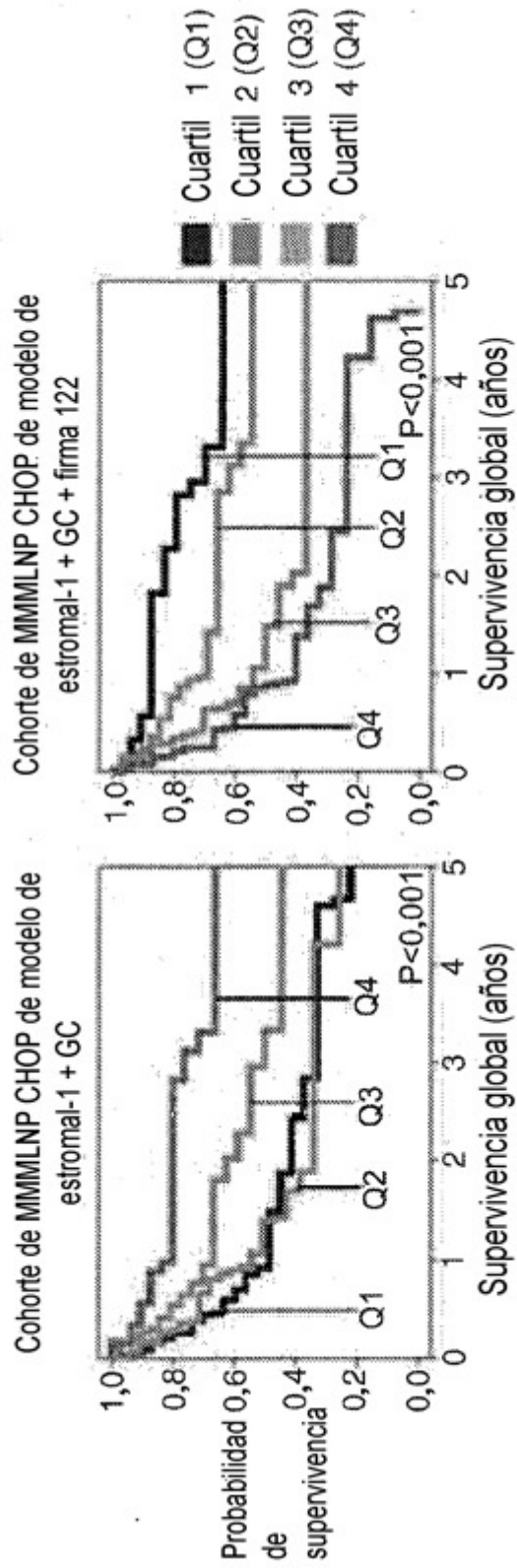


Figura 9B

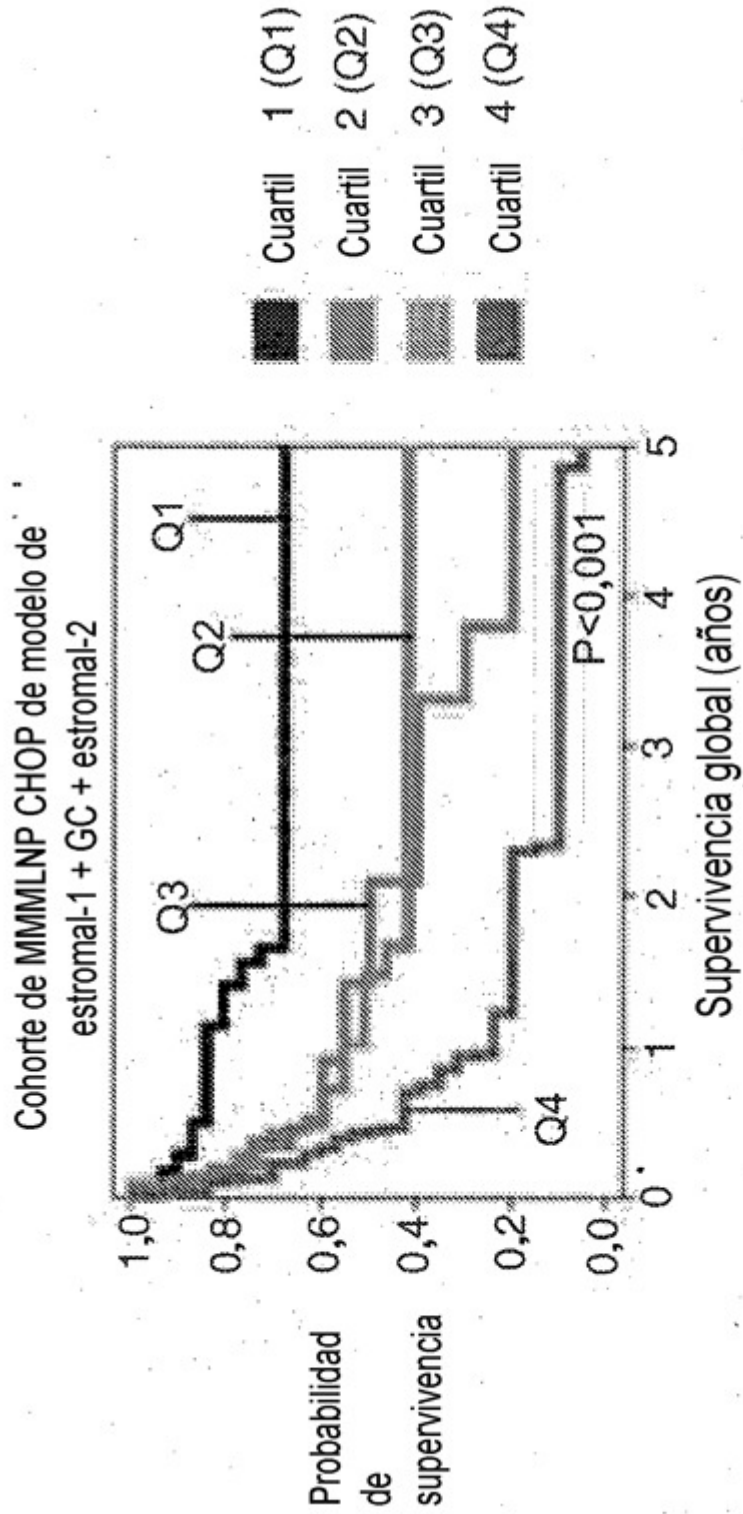


Figura 9C



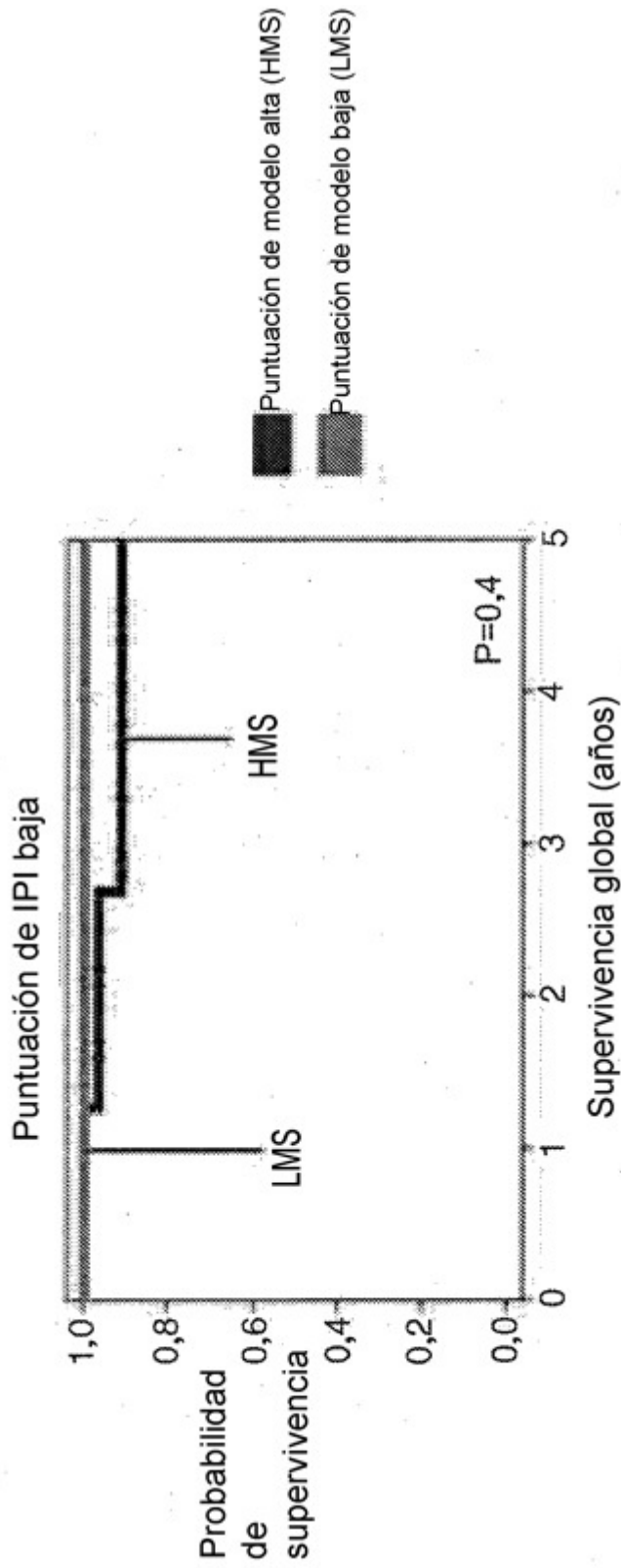


Figura 10A

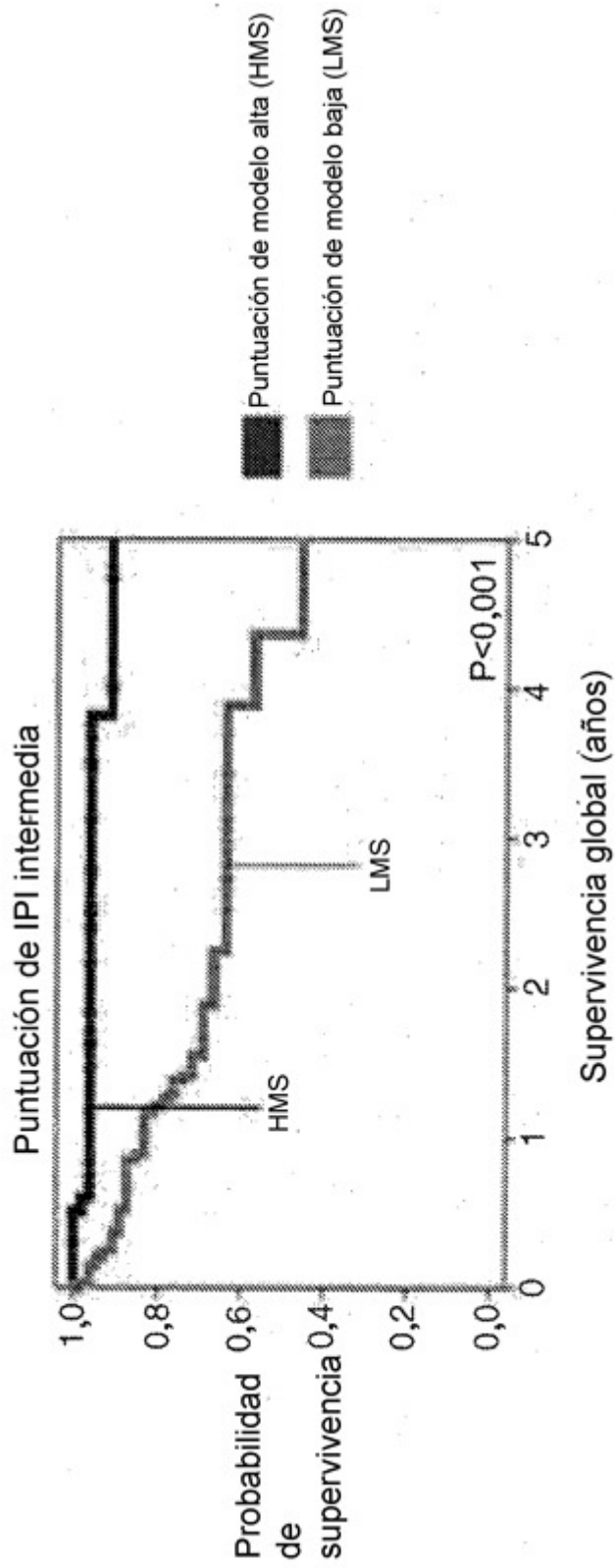


Figura 10B

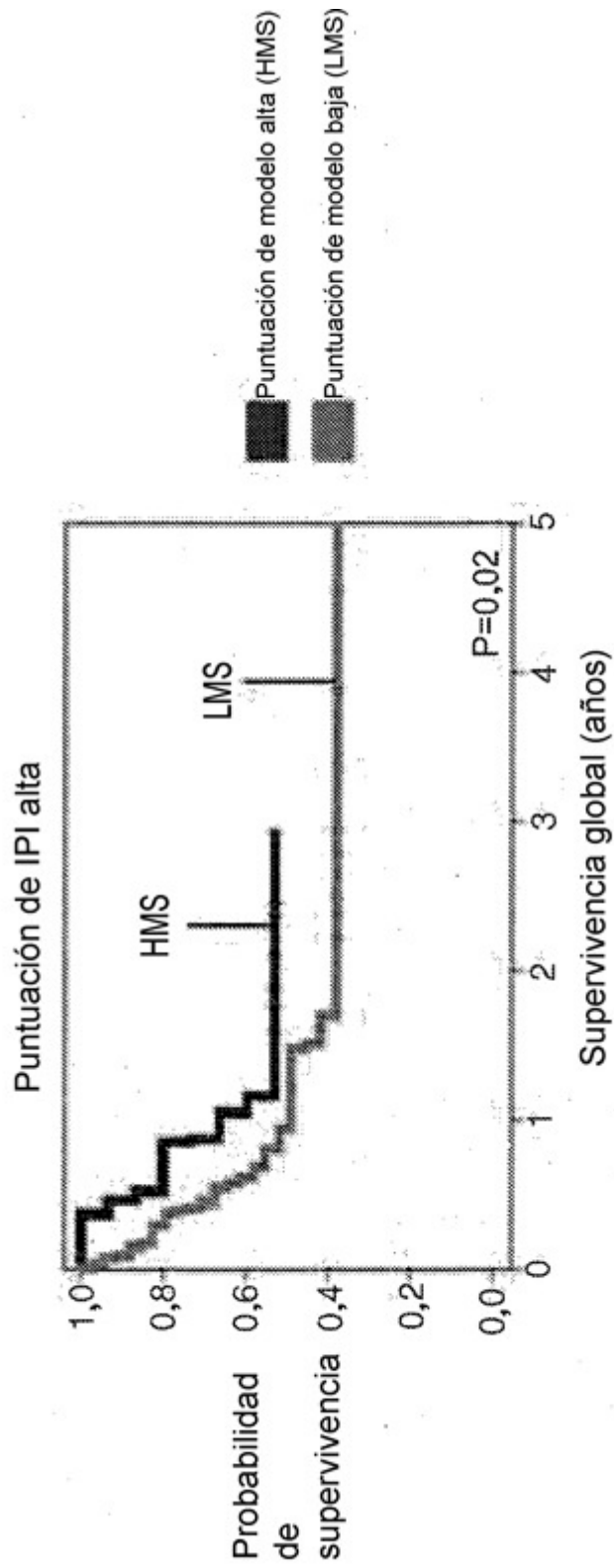


Figura 10C

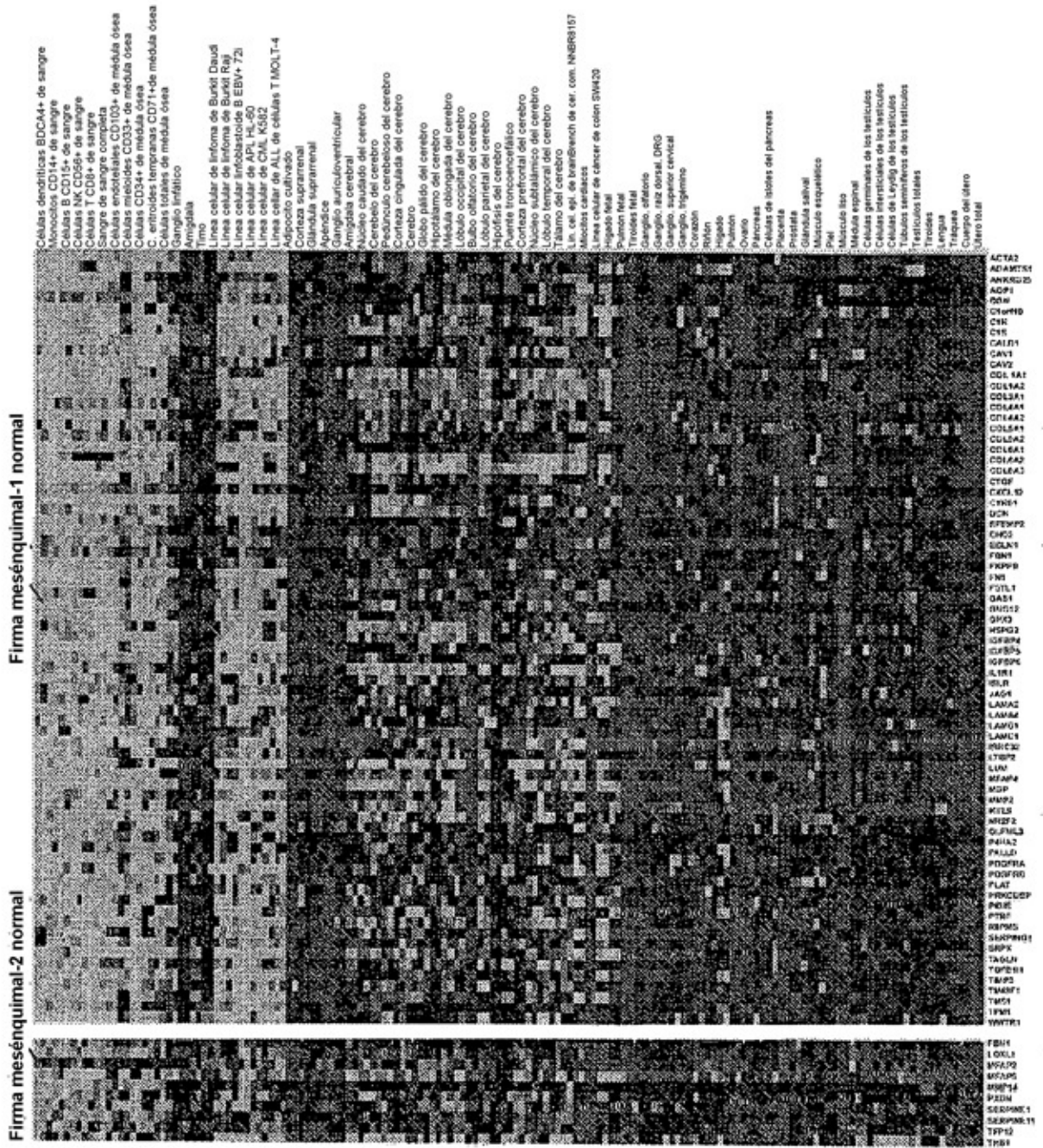


Figura 11