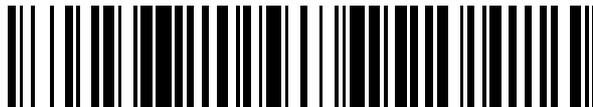


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 595 410**

51 Int. Cl.:

**C12Q 1/68** (2006.01)

**G01N 33/53** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **07.01.2010 PCT/US2010/020397**

87 Fecha y número de publicación internacional: **15.07.2010 WO10080933**

96 Fecha de presentación y número de la solicitud europea: **07.01.2010 E 10729525 (5)**

97 Fecha y número de publicación de la concesión europea: **08.06.2016 EP 2382331**

54 Título: **Biomarcadores de cáncer**

30 Prioridad:

**07.01.2009 US 143077 P**

**19.05.2009 US 179650 P**

**10.06.2009 US 185901 P**

**11.09.2009 US 241748 P**

**30.10.2009 US 256443 P**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**29.12.2016**

73 Titular/es:

**MYRIAD GENETICS, INC. (100.0%)**

**320 Wakara Way**

**Salt Lake City, UT 84108, US**

72 Inventor/es:

**STONE, STEVE;**

**GUTIN, ALEXANDER;**

**WAGNER, SUSANNE y**

**REID, JULIA**

74 Agente/Representante:

**ARIAS SANZ, Juan**

ES 2 595 410 T3

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

**DESCRIPCIÓN**

Biomarcadores de cáncer

**CAMPO DE LA INVENCION**

5 La invención se refiere, en general, a una clasificación molecular del cáncer de próstata mediante procedimientos, a medios y al uso de marcadores moleculares.

**ANTECEDENTES DE LA INVENCION**

10 El cáncer es un problema de salud pública importante, que causa aproximadamente el 25% de todas las muertes en Estados Unidos. Aunque se han diseñado muchos tratamientos para varios tipos de cáncer, estos tratamientos a menudo varían en cuanto a la gravedad de los efectos secundarios. Es útil para los médicos saber la agresividad del cáncer de un paciente con el fin de determinar el grado de agresividad con el que hay que tratar el cáncer.

15 Por ejemplo, la mayoría de los pacientes con cáncer de próstata asintomático en estadio temprano son tratados con prostatectomía radical o radioterapia y, opcionalmente, con terapia adyuvante (por ejemplo, hormonas o quimioterapia), todos los cuales tienen efectos secundarios graves. No obstante, para muchos de estos pacientes, estos tratamientos y sus efectos secundarios y costes asociados son innecesarias porque el cáncer en estos pacientes no es agresivo (es decir, crece lentamente y es poco probable que cause una mortalidad o morbilidad significativa durante el transcurso de la vida del paciente). En otros pacientes, el cáncer es virulento (es decir, es más probable que recurra) y el tratamiento agresivo es necesario para salvar la vida del paciente.

20 Se han diseñado algunas herramientas para ayudar a los médicos a decidir qué pacientes necesitan un tratamiento agresivo y cuáles no. De hecho, actualmente se están usando varios parámetros clínicos para este fin en varios tipos de cáncer diferentes. En el cáncer de próstata, por ejemplo, dichos parámetros clínicos incluyen el antígeno específico de la próstata en suero (PSA), el grado de Gleason, el estadio patológico y los márgenes quirúrgicos. En los últimos años, los parámetros clínicos se han vuelto más útiles a través de su incorporación en nomogramas postoperatorios de varias variables continuas que calculan la probabilidad que tiene el paciente de sufrir progresión / recurrencia del cáncer. Véase, por ejemplo, Kattan *et al.*, J. CLIN. Oncol. (1999) 17:1499-1507; Stephenson *et al.*, J. CLIN. Oncol. (2005) 23:7005-7012). No obstante, a pesar de estos avances, muchos pacientes reciben tratamientos contra el cáncer indebidos y todavía hay una gran necesidad de nuevas y mejores herramientas para predecir la recurrencia del cáncer.

**SUMARIO DE LA INVENCION**

30 La presente invención se basa en parte en la determinación de que los genes cuya expresión sigue de cerca el ciclo celular ("genes del ciclo celular" o "GCC") son genes particularmente decisivos para clasificar ciertos tipos de cáncer. De hecho, los inventores han descubierto que los GCC proporcionan la mayoría si no todo el poder predictivo en muchas firmas de pronóstico. Por ejemplo, los paneles de los GCC de acuerdo con la presente divulgación pueden predecir la agresividad del cáncer de próstata, vejiga, cerebro, mama y pulmón.

En particular, la presente invención proporciona lo siguiente:

- 35 1. Un procedimiento de clasificación del cáncer de próstata que comprende determinar el nivel de expresión de un panel de genes que comprende todos los genes del ciclo celular enumerados en uno cualquiera de los paneles C a G, en el que el aumento de expresión de dichos genes del ciclo celular indica un mal pronóstico.
- 40 2. Un procedimiento de clasificación del cáncer de próstata que comprende determinar al menos un parámetro clínico para un paciente con cáncer y determinar el nivel de expresión de todos los genes del ciclo celular enumerados en uno cualquiera de los paneles C a G en una muestra obtenida del paciente, en el que un parámetro clínico asociado con recurrencia o un aumento de la expresión de dichos genes del ciclo celular indica un mal pronóstico.
- 45 3. El procedimiento del apartado 2, que comprende además la determinación de dos o más parámetros clínicos, en el que dichos dos o más parámetros clínicos forman o son parte de un nomograma que comprende cualquiera los parámetros clínicos del nomograma de Kattan-Stephenson o los parámetros clínicos del índice pronóstico de Nottingham.
- 50 4. El procedimiento del apartado 1, que comprende además la determinación del nivel de expresión de PTEN en una muestra obtenida del paciente, en el que el estado de PTEN negativa o bajo indica un mal pronóstico.
5. El procedimiento del apartado 2, que comprende además la determinación del nivel de expresión de PTEN en una muestra obtenida del paciente, en el que el estado de PTEN negativa o bajo indica un mal pronóstico.
6. Un procedimiento de determinar un curso de tratamiento para un paciente con cáncer de próstata que comprende la determinación del nivel de expresión de un panel de genes que comprende todos los genes del ciclo celular enumerados en uno cualquiera de los paneles C a G en una muestra obtenida del paciente, y (a)

recomendar, prescribir o iniciar el tratamiento activo si dicha muestra ha incrementado la expresión de dichos genes del ciclo celular, o (b) recomendar, prescribir o iniciar una espera vigilante si dicha muestra no ha aumentado la expresión de dichos genes del ciclo celular.

5 7. Un procedimiento para determinar si un paciente de cáncer de próstata tiene un aumento de la probabilidad de recurrencia, que comprende:

determinar el nivel de expresión de un panel de genes que comprende todos los genes del ciclo celular enumerados en uno cualquiera de los paneles C a G en una muestra de un paciente;

(b) acceder a la información obtenida en (a) en un medio legible por ordenador;

10 consultar esta información para determinar si una muestra obtenida del paciente muestra un aumento de la expresión de dichos genes del ciclo celular; y

(d) dar salida a una indicación de que el paciente tiene un aumento de la probabilidad de recurrencia si la muestra presenta una mayor expresión de dichos genes del ciclo celular.

15 8. Un kit que comprende una pluralidad de oligonucleótidos para la determinación del nivel de expresión de un panel de genes, en el que dicho panel comprende todos los genes en uno cualquiera de los paneles C a G, y en el que el kit consiste en reactivos para la determinación de la expresión de no más de 2.500 años.

9. Uso de una pluralidad de oligonucleótidos para la determinación del nivel de expresión de un panel de genes, en el que dicho panel comprende todos los genes enumerados en uno cualquiera de los paneles C a G para determinar el pronóstico o la probabilidad de recurrencia del cáncer de próstata en una muestra obtenida de un paciente.

20 Además, la presente divulgación describe un procedimiento de clasificación del cáncer, que comprende determinar el estado de un panel de genes que comprende al menos dos GCC, en el que un estado anormal indica una clasificación de cáncer particular. La situación que se va a determinar son los niveles de expresión génica. El estado anormal indica que el cáncer es agresivo, que puede incluir que tiene un mal pronóstico, un aumento de la probabilidad de recurrencia o progresión, etc. También se describe un procedimiento para determinar el pronóstico del cáncer de un paciente que comprende la determinación del nivel de expresión de un panel de genes que  
25 comprende al menos dos GCC, en el que la expresión elevada indica un aumento de la probabilidad de recurrencia o progresión del cáncer o la necesidad de un tratamiento agresivo.

30 El número de GCC analizados puede variar dependiendo de, por ejemplo, las restricciones técnicas, las consideraciones de costes, la clasificación que se está realizando, el cáncer que se está analizando, el nivel deseado de capacidad de predicción, etc. Sin embargo, se ha determinado que determinado número mínimo de CCG proporciona predicciones sólidas y significativamente mejoradas. Por tanto, de acuerdo con la presente divulgación, se determina el estado de un panel que comprende al menos 5 GCC. El panel comprende al menos 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50 o más GCC.

El panel puede comprender al menos 10, 15, 20 o más GCC.

35 Como se divulga en el presente documento, los GCC pueden comprender al menos una determinada proporción del panel de genes. El panel puede comprender al menos 25%, 30%, 40%, 50%, 60%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% o 99% GCC.

40 En algunas realizaciones, los GCC son todos los genes enumerados en uno cualquiera de los paneles C a G. De acuerdo con la presente divulgación, el panel puede comprender al menos 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 110, 120 o más de los genes enumerados en cualquiera de la tabla 1 o paneles A a G.

45 Se ha descubierto que los GCC son particularmente predictivos en ciertos tipos de cáncer. Por ejemplo, se ha determinado que un panel de GCC es exacto en la predicción de la recurrencia en el cáncer de próstata. Además, los GCC pueden determinar el pronóstico de cáncer de vejiga, cerebro, mama y pulmón. Por tanto, la divulgación describe un procedimiento que comprende determinar el estado de un panel de genes que comprende al menos dos GCC, en el que un estado anormal indica un mal pronóstico.

50 El panel puede comprender al menos 2 genes seleccionados del grupo de genes en al menos uno de los paneles A a G. El panel puede comprender al menos 10 genes seleccionados del grupo de genes en al menos uno de los paneles A a G. El panel puede comprender al menos 15 genes seleccionados del grupo de genes en al menos uno de los paneles A a G. En algunas realizaciones, el panel comprende todos los genes en al menos uno de los paneles A a G. La divulgación también describe un procedimiento para determinar el pronóstico del cáncer de vejiga, que comprende determinar el estado de un panel de genes que comprende al menos dos GCC (por ejemplo, al menos dos de los genes en cualquiera de los paneles B, C y F), en el que un estado anormal indica un mal pronóstico. La divulgación también describe un procedimiento para determinar el pronóstico del cáncer de cerebro,

que comprende determinar el estado de un panel de genes que comprende al menos dos GCC (por ejemplo, al menos dos de los genes en cualquiera de los paneles B, C y F), en el que un estado anormal indica un mal pronóstico. La divulgación describe adicionalmente un procedimiento para determinar el pronóstico del cáncer de mama, que comprende determinar el estado de un panel de genes que comprende al menos dos GCC (por ejemplo, al menos dos de los genes en cualquiera de los paneles B, C y F), en el que un estado anormal indica un mal pronóstico. La divulgación también describe un procedimiento para determinar el pronóstico del cáncer de pulmón, que comprende determinar el estado de un panel de genes que comprende al menos dos GCC (por ejemplo, al menos dos de los genes en cualquiera de los paneles B, C y F), en el que un estado anormal indica un mal pronóstico.

Se ha descubierto además que el estado de CCG se añade de forma sinérgica a los parámetros clínicos (por ejemplo, nomogramas clínicos) en el pronóstico del cáncer. En el caso del cáncer de próstata, se ha descubierto que un alto nivel de la expresión génica de uno cualquiera de los genes en cualquiera de los paneles A a F se asocia con un mayor riesgo de recurrencia o progresión del cáncer en pacientes cuya puntuación en el nomograma clínico indica un riesgo relativamente bajo de recurrencia o progresión. Véanse, por ejemplo, los ejemplos 3 y 5. En el caso del cáncer de cerebro, se ha descubierto que un alto nivel de la expresión génica de uno cualquiera de los genes en cualquiera de los paneles B, C o F se asocia con un mayor riesgo de recurrencia o progresión del cáncer independiente del grado tumoral. En el caso del cáncer de pulmón, se ha descubierto que un alto nivel de la expresión génica de uno cualquiera de los genes en cualquiera de los paneles B, C o F se asocia con un mayor riesgo de recurrencia o progresión del cáncer independiente del estadio tumoral. En el caso del cáncer de mama, se ha descubierto que un alto nivel de la expresión génica de uno cualquiera de los genes en cualquiera de los paneles B, C o F se asocia con un mayor riesgo de recurrencia o progresión del cáncer independiente de la puntuación en el nomograma del índice pronóstico de Nottingham (NPI). Por tanto, debido a que la evaluación de los niveles de expresión de CCG puede detectar un aumento de riesgo no detectado usando parámetros clínicos solo, la invención proporciona generalmente procedimientos que combinan la evaluación de al menos un parámetro clínico con la evaluación del estado de al menos un GCC.

Se divulga un procedimiento diagnóstico in vitro que comprende determinar al menos un parámetro clínico para un paciente con cáncer y determinar el estado de al menos un GCC en una muestra obtenida del paciente. Se puede determinar el estado de una pluralidad de GCC (por ejemplo, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50 o más). Un estado anormal puede indicar un aumento de la probabilidad de recurrencia o progresión. El paciente puede tener cáncer de próstata. El paciente puede tener cáncer de pulmón.

A menudo, el parámetro clínico es, al menos de un modo algo independiente, predictivo de recurrencia o progresión y la adición del estado de GCC mejora la capacidad de predicción. GCC de ejemplo incluyen los de los paneles A a G. Se puede evaluar más de un parámetro clínico junto con el nivel de expresión de al menos un CCG. Se puede evaluar más de un parámetro clínico y la combinación de parámetros, junto con el nivel de expresión de al menos un GCC, predice la probabilidad de recurrencia o progresión del cáncer.

Los parámetros clínicos a menudo se combinan e incorporan en nomogramas para predecir el riesgo de recurrencia o progresión del cáncer.

Por tanto, se describe un procedimiento diagnóstico in vitro que comprende determinar la puntuación de un nomograma clínico para un paciente con cáncer y determinar el estado de al menos un GCC en una muestra obtenida del paciente. A menudo, los nomogramas clínicos están diseñados de tal manera que una puntuación alta se correlaciona con un aumento del riesgo de recurrencia. Del mismo modo, un estado de GCC elevado (por ejemplo, aumento de la expresión o actividad) se correlaciona con un mayor riesgo. Por tanto, se divulga un procedimiento para determinar si un paciente con cáncer tiene un aumento de la probabilidad de recurrencia o progresión, que comprende determinar una puntuación del nomograma clínico para el paciente y determinar el estado de al menos un GCC en una muestra obtenida del paciente, en el que una puntuación más alta en el nomograma y / o un estado de GCC elevado indican que el paciente tiene un aumento de la probabilidad de recurrencia o progresión. El cáncer puede ser cáncer de próstata. El cáncer puede ser cáncer de mama.

En un aspecto, la invención proporciona un procedimiento de diagnóstico in vitro que comprende determinar al menos un parámetro clínico (o una puntuación en el nomograma) para un paciente y determinar el estado de un panel de GCC. De acuerdo con la presente divulgación, el número de GCC analizados puede variar dependiendo de, por ejemplo, las restricciones técnicas, las consideraciones de costes, la clasificación que se está realizando, el cáncer que se está analizando, el nivel deseado de capacidad de predicción, etc. Como ejemplo, la divulgación describe un procedimiento de diagnóstico in vitro que comprende determinar la puntuación en un nomograma para un paciente y determinar el nivel de expresión de dos o más de los genes enumerados en la tabla 1 o los paneles A a G. Se puede utilizar cualquier combinación de GCC (incluyendo cualquiera de los enumerados en la Tabla 1 o los paneles A a G) para poner en práctica la divulgación. En una realización, se analizan todos los GCC enumerados en los paneles A a G.

También se ha descubierto que el estado de *PTEN* predice el cáncer de próstata agresivo. El estado de *PTEN* se suma a los dos parámetros clínicos y a los GCC. El estado de *PTEN* combinado con el estado de GCC es un predictor mejor que un nomograma de próstata clínico. Debido a que la evaluación del estado de *PTEN* puede

detectar un mayor riesgo no detectado usando parámetros clínicos (por ejemplo, nomograma clínica) o los niveles de expresión del GCC solo o en combinación, algunos aspectos de la invención se refieren a procedimientos que combinan un parámetro clínico o nivel de expresión de GCC con una evaluación de la situación de PTEN.

5 Por lo tanto, un aspecto de la invención proporciona un procedimiento de diagnóstico in vitro que comprende determinar el estado de PTEN y determinar el estado de un panel de GCC en una muestra obtenida de un paciente (y, opcionalmente, determinar uno o más parámetros clínicos para el paciente, por ejemplo, nomograma clínico). Se pueden utilizar diferentes combinaciones de técnicas para determinar el estado de los diversos marcadores. Por ejemplo, el estado de PTEN se determina mediante técnicas de inmunohistoquímica (IHQ), mientras que el estado de la pluralidad de GCC se determina mediante reacción en cadena de la polimerasa cuantitativa (PCRc), por ejemplo, TaqMan™.

10 Debido a que se ha demostrado que el estado de PTEN y de GCC se correlaciona de forma independiente con el riesgo de recurrencia o progresión del cáncer de próstata, se divulga un procedimiento de determinación de la probabilidad que tiene un paciente de sufrir recurrencia que comprende determinar el estado de PTEN en una muestra obtenida del paciente, en el que un estado bajo o negativo de PTEN indica que el paciente tiene un aumento de la probabilidad de recurrencia. Debido a que el estado de PTEN y de GCC también son aditivos en cuanto a la predicción del riesgo de recurrencia o progresión del cáncer de próstata, se divulga un procedimiento de determinación de la probabilidad que tiene un paciente de sufrir recurrencia que comprende determinar el estado de PTEN en una muestra obtenida del paciente y determinar el estado de una pluralidad de GCC en una muestra obtenida del paciente, en el que el estado bajo o negativo de PTEN y / o el estado elevado de GCC indican que el paciente tiene un aumento de la probabilidad de recurrencia.

15 El estado de PTEN puede añadir valor predictivo a los parámetros clínicos en la predicción de la recurrencia de próstata. Por lo tanto, otro aspecto de la invención proporciona un procedimiento de diagnóstico in vitro que comprende determinar el estado de PTEN y determinar al menos un parámetro clínico para un paciente de cáncer de próstata. A menudo, el parámetro clínico es, al menos de un modo algo independiente, predictivo de recurrencia o progresión y la adición del estado de PTEN mejora el poder predictivo. Se divulga un procedimiento para determinar si un paciente con cáncer tiene un aumento de la probabilidad de recurrencia, que comprende determinar el estado de PTEN en una muestra obtenida del paciente y determinar la puntuación en un nomograma clínico para el paciente, en el que un estado bajo o negativo de PTEN y una puntuación alta en el nomograma indican que el paciente tiene un aumento de la probabilidad de recurrencia. Se divulga un procedimiento para determinar si un paciente con cáncer tiene un aumento de la probabilidad de recurrencia, que comprende determinar el estado de PTEN en una muestra obtenida del paciente y determinar la puntuación en un nomograma clínico para el paciente, en el que un estado bajo o negativo de PTEN y una puntuación alta en el nomograma indican que el paciente tiene un aumento de la probabilidad de recurrencia.

20 En los ejemplos siguientes se ha demostrado que los tres marcadores se pueden combinar para producir un poder predictivo aún mayor. Por tanto, se divulga un procedimiento in vitro, que comprende determinar el estado de PTEN, determinar el estado de al menos un GCC y determinar al menos un parámetro clínico.

25 El estado de un gen incluye cualquier característica física, química o genética de ese gen. Los ejemplos incluyen el nivel de expresión, el nivel de actividad, las mutaciones, el número de copias, etc. En el contexto de los GCC para predecir el riesgo de recurrencia o progresión, el nivel de expresión y el nivel de actividad son a menudo las características más útiles para evaluar a la hora de determinar el estado de GCC. La determinación del nivel de expresión de un gen (por ejemplo, un GCC o PTEN) puede comprender al menos uno de:

- (a) determinar la cantidad de polipéptido codificada por el gen;
- (b) determinar la cantidad de polipéptido codificada por el gen;
- (c) determinar el nivel de expresión y / o de actividad de cualquier ARNmi que regula el gen.

30 La determinación del nivel de actividad puede comprender, por ejemplo, determinar el nivel de actividad enzimática de un polipéptido codificado por un gen; determinar el nivel de interacción entre un polipéptido y un asociado de unión; etcétera.

35 Se divulgan composiciones, micromatrices y kits para su uso en los procedimientos anteriores. Tales composiciones incluyen, pero no se limitan a, sondas de ácido nucleico que se hibridan con PTEN o un GCC (o con cualquier ácido nucleico codificado por las mismas o complementario de las mismas); cebadores de ácido nucleico y parejas de cebadores adecuados para amplificar todo o una parte de PTEN o un GCC o cualquier ácido nucleico codificado por los mismos; anticuerpos que se unen inmunológicamente a un polipéptido codificado por PTEN o un GCC; conjuntos de sondas que comprenden una pluralidad de dichas sondas de ácidos nucleicos, cebadores de ácidos nucleicos, anticuerpos y / o polipéptidos; micromatrices que comprenden cualquiera de estos; kits que comprenden cualquiera de estos; etc. También se divulgan procedimientos informáticos, sistemas, software y / o módulos para su uso en los procedimientos anteriores.

A menos que se defina de otro modo, todos los términos técnicos y científicos usados en el presente documento

5 tienen el mismo significado que entiende comúnmente un experto en la técnica a la que pertenece la presente invención. Aunque se pueden usar procedimientos y materiales similares o equivalentes a los descritos en el presente documento en la práctica o la prueba de la presente invención, los procedimientos y materiales adecuados se describen a continuación. En caso de conflicto, la presente memoria descriptiva, incluyendo las definiciones, se controlará.

Otras características y ventajas de la invención serán evidentes a partir de la siguiente descripción detallada y a partir de las reivindicaciones.

#### BREVE DESCRIPCIÓN DE LOS DIBUJOS

10 La figura 1 es una ilustración de la capacidad de predicción sobre un nomograma para paneles de GCC de diferentes tamaños.

La figura 2 es una ilustración de los GCC que predicen el tiempo hasta la recurrencia.

La figura 3 es una ilustración de un nomograma que predice el tiempo hasta la recurrencia.

La figura 4 es una ilustración de la recurrencia no solapante predicha por el nomograma y una firma de GCC.

15 La figura 5 es una ilustración del tiempo hasta la recurrencia de varias poblaciones de pacientes definidas por el nomograma y / o el estado de GCC.

La figura 6 es una ilustración de un ejemplo de un sistema útil en ciertos aspectos y realizaciones de la invención.

La Figura 7 es un diagrama de flujo que ilustra un ejemplo de un procedimiento implementado por ordenador de la invención.

20 La figura 8 un diagrama de dispersión de comparación de los parámetros clínicos y la puntuación de GCC como predictores de la recurrencia del ejemplo 5.

La figura 9 ilustra, a partir del ejemplo 5, el umbral de GCC derivado del análisis de la cohorte de formación para el conjunto de datos de validación, con la puntuación de la firma de GCC subdividiendo con eficacia a los pacientes identificados como de bajo riesgo utilizando los parámetros clínicos en pacientes con tasas de recurrencia muy bajas y un mayor riesgo de recurrencia.

25 La figura 10 ilustra la tasa de recurrencia predicha frente a la puntuación de GCC para los pacientes en la cohorte de validación del ejemplo 5.

La figura 11 ilustra la tasa de recurrencia predicha frente a la puntuación de GCC para los pacientes en la cohorte de validación del ejemplo 5.

30 La figura 12 ilustra la distribución de la puntuación del riesgo clínico en 443 pacientes estudiados en el ejemplo 5. La línea vertical oscura representa el umbral elegido por medio de KM para dividir a los pacientes de bajo y alto riesgo y se utiliza a lo largo de este estudio.

#### DESCRIPCIÓN DETALLADA DE LA INVENCION

35 La presente invención se basa en parte en el descubrimiento de que los genes cuya expresión sigue de cerca el ciclo celular ("genes del ciclo celular" o "GCC") son genes particularmente decisivos para clasificar ciertos tipos de cáncer.

40 "Gen del ciclo celular" y "GCC" en el presente documento se refieren a un gen cuyo nivel de expresión sigue de cerca la progresión de la célula a través del ciclo celular. Véase, por ejemplo, Whitfield y col., Mol. BIOL. Cell 2002, 13:1977-2000. Más específicamente, los GCC muestran aumentos y disminuciones periódicos de la expresión que coinciden con ciertas fases del ciclo celular, por ejemplo, STKL 5 y PLK muestran una expresión máxima en G2/M. *Id.* A menudo, los GCC tienen una función relacionada con el ciclo celular reconocida y clara, por ejemplo, en la síntesis o reparación del de ADN, en la condensación de cromosomas, en la división celular, etc. Sin embargo, algunos GCC tienen niveles de expresión que hacen un seguimiento del ciclo celular sin tener un papel directo obvio en el ciclo celular, por ejemplo, UBE2S codifica una enzima de conjugación con ubiquitina, sin embargo, su expresión sigue de cerca la del ciclo celular. Así, un GCC divulgado en la presente invención no necesita tener una  
45 función reconocida en el ciclo celular. Los GCC de ejemplo se enumeran en las Tablas 1, 2, 3 y 4.

50 Si un gen particular es un GCC puede determinarse mediante cualquier técnica conocida en la materia, incluyendo la que se enseña en Whitfield y col., Mol. Biol. Cell 2002, 13:1977-2000. Por ejemplo, una muestra de células, por ejemplo, células HeLa, se puede sincronizar de tal manera que todas progresen a través de las diferentes fases del ciclo celular al mismo tiempo. Generalmente esto se hace mediante la detención de las células en cada fase, por ejemplo, las células pueden detenerse en la fase S mediante el uso de un bloque doble de timidina o en la mitosis con un bloque de timidina-nocodazol. Véase, por ejemplo, Whitfield y col., Mol. CELL. Biol. (2000) 20:4188-4198). El

ARN se extrae de las células después de la detención en cada fase y la expresión génica se cuantifica usando cualquier técnica adecuada, por ejemplo, micromatriz de expresión (en todo el genoma o genes específicos de interés), PCR™ cuantitativa en tiempo real (RTQ-PCR). Finalmente, se aplica un análisis estadístico (por ejemplo, transformada de Fourier) para determinar qué genes muestran expresión máxima durante determinadas fases del ciclo celular. Los genes pueden clasificarse de acuerdo a una puntuación de la periodicidad que describe cómo de cerca sigue la expresión del gen al ciclo celular, por ejemplo, una puntuación alta indica un gen que sigue muy de cerca al ciclo celular. Finalmente, los genes cuya puntuación de periodicidad supera un nivel umbral definido (véase Whitfield y col., Mol. BIOL. CELL (2002) 13:1977-2000) se pueden designar CCG. En la tabla 1 se proporciona una lista grande, pero no exhaustiva, de los ácidos nucleicos asociados con los GCC (por ejemplo, genes, EST, clones de ADNc, etc.). Véase Whitfield y col., Mol. Biol. Cell (2002) 13:1977-2000. Todos los GCC en la Tabla 2 a continuación forman un panel de GCC ("Panel A") útiles en los procedimientos de la divulgación.

Tabla 2

Símbolo del gen	ID del gen Entrez	ID del ensayo ABI	N.º de acceso en RefSeq
<i>APOBEC3B</i> *	9582	Hs00358981_m1	NM_004900.3
<i>ASF1B</i> *	55723	Hs00216780_m1	NM_018154.2
<i>ASPM</i> *	259266	Hs00411505_m1	NM_018136.4
<i>ATAD2</i> *	29028	Hs00204205_m1	NM_014109.3
<i>BIRC5</i> *	332	Hs00153353_m1; Hs03043576_m1	NM_001012271.1; NM_001012270.1; NM_001168.2
<i>BLM</i> *	641	Hs00172060_m1	NM_000057.2
<i>BUB1</i>	699	Hs00177821_m1	NM_004336.3
<i>BUB1B</i> *	701	Hs01084828_m1	NM_001211.5
<i>C12orf48</i> *	55010	Hs00215575_m1	NM_017915.2
<i>C18orf24</i> *	220134	Hs00536843_m1	NM_145060.3;

			NM_001039535.2
<i>Clorf135*</i>	79000	Hs00225211_m1	NM_024037.1
<i>C21orf45*</i>	54069	Hs00219050_m1	NM_018944.2
<i>CCDC99*</i>	54908	Hs00215019_m1	NM_017785.4
<i>CCNA2*</i>	890	Hs00153138_m1	NM_001237.3
<i>CCNB1*</i>	891	Hs00259126_m1	NM_031966.2
<i>CCNB2*</i>	9133	Hs00270424_m1	NM_004701.2
<i>CCNE1*</i>	898	Hs01026536_m1	NM_001238.1; NM_057182.1
<i>CDC2*</i>	983	Hs00364293_m1	NM_033379.3; NM_001130829.1; NM_001786.3
<i>CDC20*</i>	991	Hs03004916_g1	NM_001255.2
<i>CDC45L*</i>	8318	Hs00185895_m1	NM_003504.3
<i>CDC6*</i>	990	Hs00154374_m1	NM_001254.3
<i>CDC43*</i>	83461	Hs00229905_m1	NM_031299.4
<i>CDC48*</i>	55143	Hs00983655_m1	NM_018101.2
<i>CDKN3*</i>	1033	Hs00193192_m1	NM_001130851.1; NM_005192.3
<i>CDT1*</i>	81620	Hs00368864_m1	NM_030928.3
<i>CENPA</i>	1058	Hs00156455_m1	NM_001042426.1; NM_001809.3
<i>CENPE*</i>	1062	Hs00156507_m1	NM_001813.2
<i>CENPF*</i>	1063	Hs00193201_m1	NM_016343.3
<i>CENPI*</i>	2491	Hs00198791_m1	NM_006733.2
<i>CENPM*</i>	79019	Hs00608780_m1	NM_024053.3
<i>CENPN*</i>	55839	Hs00218401_m1	NM_018455.4; NM_001100624.1; NM_001100625.1
<i>CEP55*</i>	55165	Hs00216688_m1	NM_018131.4; NM_001127182.1
<i>CHEK1*</i>	1111	Hs00967506_m1	NM_001114121.1; NM_001114122.1; NM_001274.4

<i>CKAP2*</i>	26586	Hs00217068_m1	NM_018204.3; NM_001098525.1
<i>CKS1B*</i>	1163	Hs01029137_g1	NM_001826.2
<i>CKS2*</i>	1164	Hs01048812_g1	NM_001827.1
<i>CTPS*</i>	1503	Hs01041851_m1	NM_001905.2
<i>CTSL2*</i>	1515	Hs00952036_m1	NM_001333.2
<i>DBF4*</i>	10926	Hs00272696_m1	NM_006716.3
<i>DDX39*</i>	10212	Hs00271794_m1	NM_005804.2
<i>DLGAP5/DLG7*</i>	9787	Hs00207323_m1	NM_014750.3
<i>DONSON*</i>	29980	Hs00375083_m1	NM_017613.2
<i>DSN1*</i>	79980	Hs00227760_m1	NM_024918.2
<i>DTL*</i>	51514	Hs00978565_m1	NM_016448.2
<i>E2F8*</i>	79733	Hs00226635_m1	NM_024680.2
<i>ECT2*</i>	1894	Hs00216455_m1	NM_018098.4
<i>ESPL1*</i>	9700	Hs00202246_m1	NM_012291.4
<i>EXO1*</i>	9156	Hs00243513_m1	NM_130398.2; NM_003686.3; NM_006027.3
<i>EZH2*</i>	2146	Hs00544830_m1	NM_152998.1; NM_004456.3
<i>FANCI*</i>	55215	Hs00289551_m1	NM_018193.2; NM_001113378.1
<i>FBXO5*</i>	26271	Hs03070834_m1	NM_001142522.1; NM_012177.3
<i>FOXM1*</i>	2305	Hs01073586_m1	NM_202003.1; NM_202002.1; NM_021953.2
<i>GINS1*</i>	9837	Hs00221421_m1	NM_021067.3
<i>GMPS*</i>	8833	Hs00269500_m1	NM_003875.2
<i>GPSM2*</i>	29899	Hs00203271_m1	NM_013296.4
<i>GTSE1*</i>	51512	Hs00212681_m1	NM_016426.5
<i>H2AFX*</i>	3014	Hs00266783_s1	NM_002105.2
<i>HMMR*</i>	3161	Hs00234864_m1	NM_001142556.1; NM_001142557.1; NM_012484.2; NM_012485.2

<i>HNI*</i>	51155	Hs00602957_m1	NM_001002033.1; NM_001002032.1; NM_016185.2
<i>KIAA0101*</i>	9768	Hs00207134_m1	NM_014736.4
<i>KIF11*</i>	3832	Hs00189698_m1	NM_004523.3
<i>KIF15*</i>	56992	Hs00173349_m1	NM_020242.2
<i>KIF18A*</i>	81930	Hs01015428_m1	NM_031217.3
<i>KIF20A*</i>	10112	Hs00993573_m1	NM_005733.2
<i>KIF20B/MPHOSPH1*</i>	9585	Hs01027505_m1	NM_016195.2
<i>KIF23*</i>	9493	Hs00370852_m1	NM_138555.1; NM_004856.4
<i>KIF2C*</i>	11004	Hs00199232_m1	NM_006845.3
<i>KIF4A*</i>	24137	Hs01020169_m1	NM_012310.3
<i>KIFC1*</i>	3833	Hs00954801_m1	NM_002263.3
<i>KPNA2</i>	3838	Hs00818252_g1	NM_002266.2
<i>LMNB2*</i>	84823	Hs00383326_m1	NM_032737.2
<i>MAD2L1</i>	4085	Hs01554513_g1	NM_002358.3
<i>MCAM*</i>	4162	Hs00174838_m1	NM_006500.2
<i>MCM10*</i>	55388	Hs00960349_m1	NM_018518.3; NM_182751.1
<i>MCM2*</i>	4171	Hs00170472_m1	NM_004526.2
<i>MCM4*</i>	4173	Hs00381539_m1	NM_005914.2; NM_182746.1
<i>MCM6*</i>	4175	Hs00195504_m1	NM_005915.4
<i>MCM7*</i>	4176	Hs01097212_m1	NM_005916.3; NM_182776.1
<i>MELK</i>	9833	Hs00207681_m1	NM_014791.2
<i>MKI67*</i>	4288	Hs00606991_m1	NM_002417.3
<i>MYBL2*</i>	4605	Hs00231158_m1	NM_002466.2
<i>NCAPD2*</i>	9918	Hs00274505_m1	NM_014865.3
<i>NCAPG*</i>	64151	Hs00254617_m1	NM_022346.3
<i>NCAPG2*</i>	54892	Hs00375141_m1	NM_017760.5
<i>NCAPH*</i>	23397	Hs01010752_m1	NM_015341.3
<i>NDC80*</i>	10403	Hs00196101_m1	NM_006101.2

ES 2 595 410 T3

<i>NEK2*</i>	4751	Hs00601227_mH	NM_002497.2
<i>NUSAP1*</i>	51203	Hs01006195_m1	NM_018454.6; NM_001129897.1; NM_016359.3
<i>OIP5*</i>	11339	Hs00299079_m1	NM_007280.1
<i>ORC6L*</i>	23594	Hs00204876_m1	NM_014321.2
<i>PAICS*</i>	10606	Hs00272390_m1	NM_001079524.1; NM_001079525.1; NM_006452.3
<i>PBK*</i>	55872	Hs00218544_m1	NM_018492.2
<i>PCNA*</i>	5111	Hs00427214_g1	NM_182649.1; NM_002592.2
<i>PDSS1*</i>	23590	Hs00372008_m1	NM_014317.3
<i>PLK1*</i>	5347	Hs00153444_m1	NM_005030.3
<i>PLK4*</i>	10733	Hs00179514_m1	NM_014264.3
<i>POLE2*</i>	5427	Hs00160277_m1	NM_002692.2
<i>PRC1*</i>	9055	Hs00187740_m1	NM_199413.1; NM_199414.1; NM_003981.2
<i>PSMA7*</i>	5688	Hs00895424_m1	NM_002792.2
<i>PSRC1*</i>	84722	Hs00364137_m1	NM_032636.6; NM_001005290.2; NM_001032290.1; NM_001032291.1
<i>PTTG1*</i>	9232	Hs00851754_u1	NM_004219.2
<i>RACGAP1*</i>	29127	Hs00374747_m1	NM_013277.3
<i>RAD51*</i>	5888	Hs00153418_m1	NM_133487.2; NM_002875.3
<i>RAD51AP1*</i>	10635	Hs01548891_m1	NM_001130862.1; NM_006479.4
<i>RAD54B*</i>	25788	Hs00610716_m1	NM_012415.2
<i>RAD54L*</i>	8438	Hs00269177_m1	NM_001142548.1;

			NM_003579.3
<i>RFC2*</i>	5982	Hs00945948_m1	NM_181471.1; NM_002914.3
<i>RFC4*</i>	5984	Hs00427469_m1	NM_181573.2; NM_002916.3
<i>RFC5*</i>	5985	Hs00738859_m1	NM_181578.2; NM_001130112.1; NM_001130113.1; NM_007370.4
<i>RNASEH2A*</i>	10535	Hs00197370_m1	NM_006397.2
<i>RRM2*</i>	6241	Hs00357247_g1	NM_001034.2
<i>SHCBP1*</i>	79801	Hs00226915_m1	NM_024745.4
<i>SMC2*</i>	10592	Hs00197593_m1	NM_001042550.1; NM_001042551.1; NM_006444.2
<i>SPAG5*</i>	10615	Hs00197708_m1	NM_006461.3
<i>SPC25*</i>	57405	Hs00221100_m1	NM_020675.3
<i>STIL*</i>	6491	Hs00161700_m1	NM_001048166.1; NM_003035.2
<i>STMN1*</i>	3925	Hs00606370_m1; Hs01033129_m1	NM_005563.3; NM_203399.1
<i>TACC3*</i>	10460	Hs00170751_m1	NM_006342.1
<i>TIMELESS*</i>	8914	Hs01086966_m1	NM_003920.2
<i>TK1*</i>	7083	Hs01062125_m1	NM_003258.4
<i>TOP2A*</i>	7153	Hs00172214_m1	NM_001067.2
<i>TPX2*</i>	22974	Hs00201616_m1	NM_012112.4
<i>TRIP13*</i>	9319	Hs01020073_m1	NM_004237.2
<i>TTK*</i>	7272	Hs00177412_m1	NM_003318.3
<i>TUBA1C*</i>	84790	Hs00733770_m1	NM_032704.3
<i>TYMS*</i>	7298	Hs00426591_m1	NM_001071.2
<i>UBE2C</i>	11065	Hs00964100_g1	NM_181799.1; NM_181800.1; NM_181801.1; NM_181802.1; NM_181803.1; NM_007019.2

<i>UBE2S</i>	27338	Hs00819350_m1	NM_014501.2
<i>VRK1*</i>	7443	Hs00177470_m1	NM_003384.2
<i>ZWILCH*</i>	55055	Hs01555249_m1	NM_017975.3; NR_003105.1
<i>ZWINT*</i>	11130	Hs00199952_m1	NM_032997.2; NM_001005413.1; NM_007057.3

\* subconjunto de 124 genes de los GCC útiles en la divulgación ("Panel B"). ID del ensayo ABI significa el número de ID del catálogo para el ensayo de expresión génica disponible comercialmente en Applied Biosystems Inc. (Foster City, CA) para el gen en concreto.

5 Aunque no se desea quedar ligado a teoría alguna, se cree que las muestras de tejidos que contienen una concentración relativamente alta de células que proliferan rápidamente mostrarán, en el agregado, una expresión relativamente más alta de los GCC. Se piensa que esto se debe a que algunos GCC están regulados por aumento durante la proliferación celular (es decir, durante ciertas fases del ciclo celular). A su vez, los tumores que comprenden una concentración alta de células que proliferan rápidamente tienden a ser más agresivos y los  
10 pacientes que tienen estos tumores tienden a tener un peor pronóstico. De hecho, los inventores han determinado que los GCC proporcionan la mayoría si no todo el poder predictivo en la mayoría de las firmas de pronóstico del cáncer. Por ejemplo, los paneles que tienen GCC pueden predecir la agresividad del cáncer de próstata, de cerebro, de mama y de pulmón.

15 Por lo tanto, un aspecto de la divulgación describe un procedimiento de clasificación de cáncer, que comprende determinar el estado de un panel de genes que comprende al menos dos GCC, en una muestra de tejido o de células, en particular una muestra de tumor, de un paciente, en el que un estado anormal indica un cáncer de clasificación negativa del cáncer. Como se usa en el presente documento, "determinar el estado" de un gen se refiere a la determinación de la presencia, ausencia o extensión / nivel de algunas características físicas, químicas o genéticas del gen o su producto o productos de expresión. Tales características incluyen, pero no se limitan a, los  
20 niveles de expresión, los niveles de actividad, las mutaciones, el número de copias, el estado de metilación, etc.

25 En el contexto de los GCC tal como se utiliza para determinar el riesgo de recurrencia o progresión del cáncer o la necesidad de un tratamiento agresivo, las características particularmente útiles incluyen los niveles de expresión (por ejemplo, los niveles de ARNm o proteínas) y de actividad. Las características pueden analizarse directamente (por ejemplo, mediante el análisis del nivel de expresión de un GCC ) o determinarse indirectamente (por ejemplo, analizando el nivel de un gen o genes cuyo nivel de expresión se correlaciona con el nivel de expresión del GCC). Se divulga un procedimiento de clasificación de cáncer que comprende la determinación del nivel de expresión, en particular el nivel de ARNm de un panel de genes que comprende al menos dos GCC, en una muestra de tumor, en el que la expresión elevada indica una clasificación de cáncer negativa, o un aumento del riesgo de recurrencia o progresión del cáncer o la necesidad de un tratamiento agresivo.

30 Los expertos en la técnica están familiarizados con diversas técnicas para determinar el estado de un gen o proteína en una muestra de tejido o de células, incluyendo, pero sin limitaciones, el análisis de micromatrices (por ejemplo, para analizar la expresión del ARNm o microARN, el número de copias , etc.), PCR cuantitativa ("QRT-PCR", por ejemplo, TaqMan <sup>TM</sup>), inmunoanálisis (por ejemplo, ELISA, inmunohistoquímica en tiempo real), etc. El nivel de actividad de un polipéptido codificado por un gen se puede utilizar del mismo modo que el nivel de expresión del gen o polipéptido. A menudo, los niveles de actividad más altos indican mayores niveles de expresión y, al mismo tiempo, niveles de actividad más bajos indican niveles de expresión bajos. Se divulga cualquiera de los procedimientos tratados anteriormente, en el que el nivel de actividad de un polipéptido codificado por el GCC se determina en lugar de o además del nivel de expresión del GCC. Los expertos en la técnica están familiarizados con técnicas para medir la actividad de varias de estas proteínas, incluyendo las codificadas por los genes enumerados en las tablas 1, 2, 3 y 4. Los procedimientos se pueden poner en práctica independientemente de la técnica particular utilizada.

45 "Estado anormal" significa un estado de marcador en una muestra particular que difiere de la situación general que se encuentra las muestras medias (por ejemplo, muestras sanas o muestras enfermas medias). Los ejemplos incluyen mutado, elevado, disminuido, presente, ausente, etc. Un "estado elevado" significa que una o más de las características anteriores (por ejemplo, el nivel de expresión o de ARNm) es superior a los niveles normales.

Generalmente esto significa un aumento de la característica (por ejemplo, el nivel de expresión o de ARNm) en comparación con un valor índice. Por el contrario, un "estado bajo" significa que una o más de las características anteriores (por ejemplo, el nivel de expresión génica o de ARNm) es inferior a los niveles normales. Generalmente esto significa una disminución de la característica (por ejemplo, la expresión) en comparación con un valor índice. En este contexto, un "estado negativo" significa generalmente que la característica está ausente o es indetectable. Por ejemplo, el estado de PTEN es negativo si el ácido nucleico y / o la proteína PTEN está ausente o no es detectable en una muestra. Sin embargo, el estado de PTEN negativo también incluye una mutación o la reducción del número de copias en PTEN.

Los expertos en la técnica apreciarán cómo obtener y utilizar un valor de índice en los procedimientos de la invención. Por ejemplo, el valor índice puede representar los niveles de expresión génica que se encuentran en una muestra normal obtenida del paciente de interés, en cuyo caso un nivel de expresión en la muestra de tumor significativamente mayor que este valor índice podría indicar, por ejemplo, un mal pronóstico o aumento de probabilidad de recurrencia del cáncer o la necesidad de un tratamiento agresivo.

Como alternativa, el valor índice puede representar el nivel promedio de expresión de un conjunto de individuos de una población diversa de cáncer o un subconjunto de la población. Por ejemplo, se puede determinar el nivel promedio de expresión de un gen o panel de genes en un muestreo aleatorio de pacientes con cáncer (por ejemplo, cáncer de próstata, vejiga, cerebro, mama o de pulmón). Este nivel promedio de expresión puede denominarse el "valor índice de umbral", teniendo pacientes que tienen expresión de GCC mayor que este valor esperado un peor pronóstico que los que tienen una expresión menor que este valor.

Como alternativa, el valor índice puede representar el nivel promedio de expresión de un marcador génico particular en una pluralidad de pacientes de formación (por ejemplo, pacientes con cáncer de próstata), con resultados similares cuyos datos clínicos y de seguimiento están disponibles y son suficientes para definir y clasificar los pacientes por evolución de la enfermedad, por ejemplo, la recurrencia o el pronóstico. Véanse, por ejemplo, los ejemplos más adelante. Por ejemplo, un "valor índice de buen pronóstico" se puede generar a partir de una pluralidad de pacientes de formación con cáncer caracterizados por tener un "resultado bueno", por ejemplo, aquellos que no han tenido recurrencia del cáncer a los cinco años (o diez años o más) después del tratamiento inicial, o que no han tenido progresión del cáncer cinco años (o diez años o más) después del diagnóstico inicial. Por ejemplo, un "valor índice de mal pronóstico" se puede generar a partir de una pluralidad de pacientes de formación con cáncer caracterizados por tener un "resultado malo", por ejemplo, aquellos que no han tenido recurrencia del cáncer en un plazo de cinco años (o diez años etc.) después del tratamiento inicial o que no han sufrido progresión del cáncer en un plazo de cinco años (o diez años etc.) después del diagnóstico inicial. Por lo tanto, un valor índice de buen pronóstico de un gen particular puede representar el nivel promedio de expresión del gen en particular en pacientes que tienen un "buen resultado", mientras que un valor índice de mal pronóstico de un gen en particular representa el nivel promedio de expresión del gen en particular en pacientes que tienen un "mal resultado".

De acuerdo con la presente divulgación, los procedimientos pueden comprender la determinación de la expresión de uno o más GCC y, si esta expresión está "aumentada", el paciente tiene un mal pronóstico. En el contexto de la invención, expresión "aumentada" de un GCC significa que el nivel de expresión del paciente está elevado sobre un valor índice normal o un índice umbral (por ejemplo, al menos una cierta cantidad umbral) o más cerca del "valor índice de mal pronóstico" que el "valor índice de buen pronóstico".

De este modo, cuando el nivel determinado de expresión de un gen marcador relevante está más cerca del valor índice de buen pronóstico del gen que del valor índice de mal pronóstico del gen, se puede concluir que es más probable que el paciente tenga un buen pronóstico, es decir, una probabilidad baja (o no aumentada) de recurrencia del cáncer. Por otro lado, cuando el nivel determinado de expresión de un marcador génico relevante está más cerca del valor índice de mal pronóstico del gen que del valor índice de buen pronóstico del gen, se puede concluir que es más probable que el paciente tenga un mal pronóstico, es decir, una probabilidad aumentada de recurrencia del cáncer.

Como alternativa, los valores índice se pueden determinar de esta manera: Con el fin de asignar a los pacientes a grupos de riesgo, se establecerá un valor umbral para la media del ciclo celular. El valor umbral óptimo se selecciona en base a la curva característica de funcionamiento del receptor (ROC), que representa gráficamente la sensibilidad frente a (1 - especificidad). Por cada incremento de la media del ciclo celular se calculan la sensibilidad y la especificidad del ensayo usando este valor como umbral. El umbral real será el valor que optimiza estas métricas de acuerdo con los requisitos de los expertos (por ejemplo, qué grado de sensibilidad o especificidad se desea, etc.). El ejemplo 5 demuestra la determinación de un valor umbral determinado y validado experimentalmente.

El nivel de expresión de un gen particular (en particular GCC) se puede obtener como el número absoluto de copias de ARNm de una muestra de tejido o célula tumoral en particular. El nivel de expresión de un gen particular (en particular, GCC) se puede obtener mediante la normalización de la cantidad de un producto de la expresión de un gen de interés en particular frente a la cantidad de expresión del producto génico de normalización (por ejemplo, uno o más genes domésticos). La normalización puede realizarse para generar un valor índice o simplemente para ayudar en la reducción del ruido de fondo a la hora de determinar el nivel de expresión del gen de interés. Por ejemplo, en la determinación del nivel de expresión de un gen relevante, la cantidad de un producto de expresión del

- gen (por ejemplo, ARNm, ADNc, proteína) se mide dentro de una o más células, en particular células tumorales, y se normaliza frente a la cantidad del producto o productos de expresión de un gen de normalización, o un conjunto de genes de normalización, dentro de las mismas una o más células, para obtener el nivel de expresión del gen marcador relevante. Por ejemplo, cuando se usa un solo gen como gen de normalización, un gen de mantenimiento cuya expresión se determina que es independiente del resultado / pronóstico de cáncer de próstata o no varía entre las células normales y cancerosas de la próstata. También se puede utilizar un conjunto de tales genes domésticos en el análisis de la expresión génica para proporcionar un conjunto de genes de normalización combinado. Los genes domésticos son bien conocidos en la técnica, incluyendo como ejemplos, pero sin limitaciones a los mismos, GUSB (glucuronidasa, beta), HMBS (hidroximetilbilano sintasa), SDHA (complejo succinato deshidrogenasa, subunidad A, flavoproteínas), UBC (ubiquitina C) y YWHAZ (proteína de activación de tirosina 3-monooxigenasa / triptófano 5-monooxigenasa, polipéptido zeta). Cuando un conjunto de genes de normalización combinado se utiliza en la normalización, la cantidad de la expresión génica de estos genes de normalización se puede promediar, combinar mediante adiciones consecutivas o mediante un algoritmo definido. También se pueden utilizar como genes de normalización genes que no sean genes domésticos.
- 15 Otros ejemplos de genes domésticos particularmente útiles para su uso en los procedimientos de la invención incluyen los enumerados en la Tabla A a continuación.

**Tabla A**

Símbolo del gen	ID del gen Entrez	ID del ensayo de Applied Biosystems	N.º de acceso en RefSeq
CLTC*	1213	Hs00191535_m1	NM_004859.3
GUSB	2990	Hs99999908_m1	NM_000181.2
HMBS	3145	Hs00609297_m1	NM_000190.3
MMADHC*	27249	Hs00739517_g1	NM_015702.2
MRFAP1*	93621	Hs00738144_g1	NM_033296.1
PPP2CA*	5515	Hs00427259_m1	NM_002715.2
PSMA1*	5682	Hs00267631_m1	
PSMC1*	5700	Hs02386942_g1	NM_002802.2
RPL13A*	23521	Hs03043885_g1	NM_012423.2
RPL37*	6167	Hs02340038_g1	NM_000997.4
RPL38*	6169	Hs00605263_g1	NM_000999.3
RPL4*	6124	Hs03044647_g1	NM_000968.2
RPL8*	6132	Hs00361285_g1	NM_033301.1; NM_000973.3
RPS29*	6235	Hs03004310_g1	NM_001030001.1; NM_001032.3
SDHA	6389	Hs00188166_m1	NM_004168.2
SLC25A3*	6515	Hs00358082_m1	NM_213611.1; NM_002635.2; NM_005888.2
TXNL1*	9352	Hs00355488_m1	NR_024546.1; NM_004786.2
UBA52*	7311	Hs03004332_g1	NM_001033930.1; NM_003333.3
UBC	7316	Hs00824723_m1	NM_021009.4
YWHAZ	7534	Hs00237047_m1	NM_003406.3

\* Subconjunto de los genes domésticos utilizado en, por ejemplo, el ejemplo 5.

- Los paneles de los GCC (por ejemplo, 2, 3, 4, 5, 6, 7, 8, 9 o 10 o más GCC) pueden predecir con precisión el pronóstico, como se muestra en el ejemplo 3. Los expertos en la materia están familiarizados con diversas formas de determinar la expresión de un panel de genes (es decir, una pluralidad de genes). Se puede determinar la expresión de un panel de genes determinando el nivel promedio de expresión (normalizado o absoluto) de todos los genes del panel en una muestra obtenida de un paciente en particular (ya sea a través de la muestra o en un subconjunto de células de la muestra o en una única célula). El aumento de expresión en este contexto significará que la expresión promedio es mayor que el nivel promedio de expresión de estos genes en pacientes normales (o más alto que algún valor índice que se ha determinado que representa el nivel promedio de expresión en una población de referencia, tal como los pacientes con el mismo cáncer). Como alternativa, se puede determinar la expresión de un panel de genes determinando el nivel promedio de expresión (normalizado o absoluto) de al menos un determinado número (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30 o más) o al menos una determinada proporción (por ejemplo, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 99%, 100%) de los genes en el panel. Como alternativa, se puede determinar la expresión de un panel de genes mediante la determinación del número de copias absoluto del

ARNm (o proteína) de todos los genes en el panel y el total o el promedio de estos en los genes.

Como se usa en el presente documento, "clasificación de un cáncer" y "clasificación de cáncer" se refieren a la determinación de una o más características clínicamente relevantes de un cáncer y / o la determinación de un pronóstico concreto de un paciente que sufre dicho cáncer. Por tanto, "clasificación de un cáncer" incluye, pero no se limita a: (i) evaluar el potencial metastásico, el potencial de metástasis en órganos específicos, el riesgo de recurrencia y / o la evolución del tumor; (ii) evaluar el estadio del tumor; (iii) determinar el pronóstico del paciente en ausencia de tratamiento del cáncer; (iv) determinar el pronóstico de la respuesta del paciente (por ejemplo, la reducción del tumor o la supervivencia libre de progresión) al tratamiento (por ejemplo, quimioterapia, radioterapia, cirugía para extirpar tumores, etc.); (v) el diagnóstico de la respuesta real del paciente al tratamiento actual y/o pasado; (vi) determinar un curso preferido de tratamiento para el paciente; (vii) el pronóstico de recaída del paciente después del tratamiento (ya sea el tratamiento en general o algún tratamiento en particular); (viii) el pronóstico de la esperanza de vida del paciente (por ejemplo, el pronóstico para la supervivencia global), etc.

De este modo, una "clasificación negativa" significa una característica clínica desfavorable del cáncer (por ejemplo, un mal pronóstico). Los ejemplos incluyen (i) un aumento del potencial metastásico, el potencial de metástasis en órganos específicos, y / o el riesgo de recurrencia; (ii) un estadio avanzado del tumor; (iii) un pronóstico malo del paciente en ausencia de tratamiento del cáncer; (iv) un mal pronóstico de la respuesta al paciente (por ejemplo, la reducción del tamaño del tumor o la supervivencia libre de progresión) a un tratamiento en particular (por ejemplo, quimioterapia, radioterapia, cirugía para extirpar tumor, etc.); (v) un mal pronóstico para la recaída del paciente después del tratamiento (ya sea un tratamiento en general o algún tratamiento en particular); (vi) un mal pronóstico de la esperanza de vida del paciente (por ejemplo, el pronóstico para la supervivencia global), etc. Un parámetro clínico asociado a recurrencia (o una puntuación más alta en el nomograma) y un aumento de la expresión de un GCC pueden indicar una clasificación negativa en el cáncer (por ejemplo, aumento de la probabilidad de recurrencia o progresión).

Como se ha tratado anteriormente, se piensa que la expresión elevada de GCC acompaña a las células cancerosas en proliferación rápida (y por lo tanto más agresiva). Dicho cáncer en un paciente que a menudo significará que el paciente tiene un aumento de la probabilidad de recurrencia después del tratamiento (por ejemplo, las células cancerosas no muertas ni eliminadas por el tratamiento crecerán rápidamente de nuevo). Tal cáncer también puede significar que el paciente tiene un aumento de la probabilidad de progresión del cáncer a una progresión más rápida (por ejemplo, las células que proliferan rápidamente harán que cualquier tumor crezca rápidamente, aumente su virulencia y / o produzca metástasis). Un cáncer de este tipo también puede significar que el paciente puede requerir un tratamiento relativamente más agresivo. Por tanto, se divulga un procedimiento de clasificación del cáncer, que comprende determinar el estado de un panel de genes que comprende al menos dos GCC, en el que un estado anormal indica un aumento de la probabilidad de recurrencia o progresión. Como se ha tratado anteriormente, el estado que se va a determinar pueden ser los niveles de expresión génica. Por tanto, se divulga un procedimiento para determinar el pronóstico del cáncer de un paciente que comprende la determinación del nivel de expresión de un panel de genes que comprende al menos dos GCC, en el que la expresión elevada indica un aumento de la probabilidad de recurrencia o progresión del cáncer.

"Recurrencia" y "progresión" son términos bien conocidos en la materia y se usan en el presente documento de acuerdo con sus significados conocidos. Como un ejemplo, el significado de "progresión" puede ser dependiente del tipo de cáncer y progresión del cáncer de pulmón significa algo diferente de la progresión en el cáncer de próstata. Sin embargo, dentro de cada tipo y subtipo de cáncer, los expertos en la materia entienden claramente "progresión". Como se usa en el presente documento, un paciente tiene un "aumento de la probabilidad" de alguna característica clínica o el resultado (por ejemplo, recurrencia o progresión) si la probabilidad de que el paciente tenga la característica o el resultado supera a alguna probabilidad o valor de referencia. La probabilidad de referencia puede ser la probabilidad de la característica o resultado en la población general de pacientes relevantes. Por ejemplo, si la probabilidad de recurrencia en la población general de cáncer de próstata es X % y se ha determinado mediante los procedimientos de la presente invención que un paciente en particular tiene una probabilidad de recurrencia de Y %, y si  $Y > X$ , el paciente tiene un "aumento de la probabilidad" de recurrencia. Como alternativa, como se ha tratado anteriormente, puede determinarse un valor de umbral o de referencia y la probabilidad de recurrencia de un paciente en particular puede compararse con dicho umbral o referencia. Debido a que la predicción de la recurrencia y la predicción de la progresión son intentos pronósticos, en el presente documento "predicción del pronóstico" a menudo se utiliza para hacer referencia a uno o a ambos. En estos casos, un "mal pronóstico" se referirá, generalmente, a un aumento de la probabilidad de recurrencia, la progresión, o ambos.

Como se muestra en el Ejemplo 3, los GCC individuales pueden predecir el pronóstico bastante bien. Por tanto, la invención proporciona un procedimiento para predecir el pronóstico que comprende determinar la expresión de todos los GCC enumerados en uno cualquiera de los paneles C a G.

El ejemplo 3 también muestra que los paneles de los GCC (por ejemplo, 2, 3, 4, 5 o 6 GCC) pueden predecir el pronóstico con exactitud. Por tanto, se describe un procedimiento de clasificación de un cáncer que comprende determinar el estado de un panel (es decir, una pluralidad) de los genes que comprende una pluralidad de GCC. Por ejemplo, el aumento de expresión en un panel de genes puede referirse al nivel promedio de expresión de todos los genes del panel en un paciente particular que es más alto que el nivel promedio de expresión de estos genes en

pacientes normales (o mayor que un valor índice que se ha determinado que representa el nivel promedio de expresión normal). Como alternativa, una expresión aumentada en un panel de genes puede hacer referencia a una expresión aumentada en al menos un determinado número (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30 o más) o al menos una determinada proporción (por ejemplo, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 99%, 100%) de los genes en el panel en comparación con el nivel promedio de expresión normal.

El panel puede comprender al menos 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 70, 80, 90, 100, 200 o más GCC. El panel puede comprender al menos 10, 15, 20 o más GCC. El panel puede comprender entre 5 y 100 GCC, entre 7 y 40 GCC, entre 5 y 25 GCC, entre 10 y 20 GCC o entre 10 y 15 GCC. Los GCC pueden comprender al menos una cierta proporción del panel. El panel puede comprender al menos el 25%, 30%, 40%, 50%, 60%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% o 99% de GCC. El panel puede comprender al menos 10, 15, 20, 25, 30, 35, 40, 45, 50, 70, 80, 90, 100, 200 o más GCC, y dichos GCC constituyen al menos el 50%, 60%, 70%, 75%, 80%, 85%, ; 90%, 95%, 96%, 97%, 98%, o 99% o más del número total de genes en el panel. Los GCC se pueden elegir del grupo que consiste en los genes de la Tabla 1 y de los paneles A a G. El panel puede comprender al menos 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 25, 30, o más de los genes en cualquiera de la Tabla 1 y de los paneles A a G. Se divulga un procedimiento de predicción del pronóstico que comprende determinar el estado de los GCC en los paneles a A a G, en el que el estado anormal indica un mal pronóstico.

La expresión elevada puede indicar un aumento de la probabilidad de recurrencia o progresión. Por tanto, se divulga un procedimiento de predicción del riesgo de recurrencia o progresión del cáncer en un paciente que comprende determinar el estado de un panel de genes, en el que el panel comprende entre aproximadamente 10 y aproximadamente 15 GCC, los GCC constituyen al menos el 90% del panel, y un estado elevado de los GCC indica un aumento de la probabilidad de recurrencia o progresión.

Se han evaluado varios paneles de GCC (tabla 2, citada anteriormente, y tablas 3 y 4, más adelante) según su capacidad para predecir el pronóstico en varios tipos de cáncer diferentes. Los resultados de estos estudios se describen en los Ejemplos 1 a 6 a continuación.

**Tabla 3: "Panel C" Evaluado en el ejemplo 1 a 4**

Símbolo del gen	ID del gen Entrez	Símbolo del gen	ID del gen Entrez	Símbolo del gen	ID del gen Entrez
<i>AURKA</i>	6790	<i>DTL*</i>	51514	<i>PTTG1*</i>	9232
<i>BUB1*</i>	699	<i>FOXM1*</i>	2305	<i>RRM2*</i>	6241
<i>CCNB1*</i>	891	<i>HMMR*</i>	3161	<i>TIMELESS*</i>	8914
<i>CCNB2*</i>	9133	<i>KIF23*</i>	9493	<i>TPX2*</i>	22974
<i>CDC2*</i>	983	<i>KPNA2</i>	3838	<i>TRIP13*</i>	9319
<i>CDC20*</i>	991	<i>MAD2L1*</i>	4085	<i>TTK*</i>	7272
<i>CDC45L*</i>	8318	<i>MELK</i>	9833	<i>UBE2C</i>	11065
<i>CDC48*</i>	55143	<i>MYBL2*</i>	4605	<i>UBE2S*</i>	27338
<i>CENPA</i>	1058	<i>NUSAP1*</i>	51203	<i>ZWINT*</i>	11130
<i>CKS2*</i>	1164	<i>PBK*</i>	55872		
<i>DLG7*</i>	9787	<i>PRC1*</i>	9055		

\*Estos genes se utilizaron como un panel de subconjunto de 26 genes ("Panel D") en la rama de validación del experimento descrito en el ejemplo 2.

**Tabla 4: "Panel E"**

Nombre	GeneID	Nombre	GeneID	Nombre	GeneID
<i>ASF1B*</i>	55723	<i>CENPM*</i>	79019	<i>ORC6L*</i>	23594
<i>ASPM*</i>	259266	<i>CEP55*</i>	55165	<i>PBK*</i>	55872
<i>BIRC5*</i>	332	<i>DLGAP5*</i>	9787	<i>PLK1*</i>	5347
<i>BUB 1B*</i>	701	<i>DTL*</i>	51514	<i>PRC1*</i>	9055
<i>C18orf24*</i>	220134	<i>FOXM1*</i>	2305	<i>PTTG1*</i>	9232
<i>CDC2*</i>	983	<i>KIAA0101*</i>	9768	<i>RAD51*</i>	5888
<i>CDC20*</i>	991	<i>KIF11*</i>	3832	<i>RAD54L*</i>	8438

<i>CDCA3*</i>	83461	<i>KIF20A*</i>	10112	<i>RRM2*</i>	6241
<i>CDCA8*</i>	55143	<i>KIF4A</i>	24137	<i>TK1*</i>	7083
<i>CDKN3*</i>	1033	<i>MCM10*</i>	55388	<i>TOP2A*</i>	7153
<i>CENPF*</i>	1063	<i>NUSAP1*</i>	51203		

\*Estos genes se utilizaron como un panel de subconjunto de 31 genes ("Panel F") en el experimento descrito en el ejemplo 5.

Se ha determinado que la elección de GCC individuales para un panel a menudo puede ser relativamente arbitraria. En otras palabras, se ha encontrado que la mayoría de los GCC son sustitutos muy buenos uno de otro. Una forma de evaluar si los GCC concretos servirán bien en los procedimientos y las composiciones de la divulgación es mediante la evaluación de su correlación con la expresión media de los GCC (por ejemplo, todos los GCC conocidos, un conjunto específico de GCC, etc.). Se espera que los GCC que se correlacionan bien con la media funcionen bien en los ensayos divulgados en el presente documento, por ejemplo, debido a que estos reducirán el ruido en el ensayo. Una clasificación de GCC seleccionados en función de su correlación con la expresión media de GCC se da en la Tabla 23.

En las firmas de GCC, los GCC particulares analizados no suelen ser tan importantes como el número total de GCC. El número de GCC analizados puede variar dependiendo de muchos factores, por ejemplo, las restricciones técnicas, las consideraciones de costes, la clasificación que se está realizando, el cáncer que se está analizando, el nivel deseado de capacidad de predicción, etc. El incremento del número de GCC analizados en un panel divulgado en el presente documento es, como cuestión general, ventajoso porque, por ejemplo, un mayor número de ARNm para analizar significa menos "ruido" causado por valores atípicos y menos probabilidad de un error del ensayo que desecha el poder predictivo global del ensayo. Sin embargo, el coste y otras consideraciones generalmente limitan este número y es deseable encontrar el número óptimo de GCC para una firma.

Se ha descubierto que el poder predictivo de una firma de GCC a menudo deja de aumentar de manera significativa más allá de un cierto número de GCC (véase la figura 1, ejemplo 1). Más específicamente, el número óptimo de GCC en una firma ( $n_o$ ) se puede encontrar siempre que lo siguiente sea cierto

$$(P_{n+1} - P_n) < C_o,$$

en la que P es la capacidad de predicción (es decir,  $P_n$  es la capacidad de predicción de una firma con n genes y  $P_{n+1}$  es la capacidad de predicción de una firma con n genes más uno) y  $C_o$  es una constante de optimización. La capacidad de predicción se puede definir de muchas maneras conocidas por los expertos en la técnica incluyendo, pero sin limitaciones, el valor de p de la firma. El técnico puede elegir la  $C_o$  según sus limitaciones específicas. Por ejemplo, si el coste no es un factor crucial y se desean niveles extremadamente altos de sensibilidad y especificidad, la  $C_o$  se puede fijar a un valor muy bajo de manera que solo se desechen incrementos triviales de la capacidad de predicción. Por otro lado, si el coste es decisivo y se pueden aceptar niveles moderados de sensibilidad y especificidad, la  $C_o$  se puede fijarse a un nivel mayor de modo que solo aumentos significativos de la capacidad de predicción garanticen un aumento del número de genes en la firma.

Como alternativa se puede representar un gráfico de la capacidad de predicción como una función del número de genes (como en la figura 1) y se toma la segunda derivada de este gráfico. El punto en el que la segunda derivada disminuye a un valor predeterminado ( $C_o'$ ) puede ser el número óptimo de genes en la firma.

Los ejemplos 1 y 3 y la figura 1 ilustran la determinación empírica de números óptimos de GCC en los paneles de GCC de los paneles. Los subconjuntos seleccionados al azar de los 31 GCC enumerados en la Tabla 3 se analizaron como firmas de GCC distintas y se determinó la capacidad de predicción (es decir, valor de p) para cada uno. Dado que la FIG. Como muestra la figura 1, los valores dejaron de mejorar significativamente entre aproximadamente 10 y aproximadamente 15 GCC, lo que indica que un número óptimo de GCC en un panel de pronóstico es de aproximadamente 10 a aproximadamente 15. De acuerdo con la presente divulgación se describe un procedimiento para predecir el pronóstico en un paciente con cáncer de próstata que comprende determinar el estado de un panel de genes, en el que el panel comprende entre aproximadamente 10 y aproximadamente 15 GCC y un estado elevado de los GCC indica un mal pronóstico. El panel puede comprender entre aproximadamente 10 y aproximadamente 15 GCC y los GCC constituyen al menos el 90% del panel. El panel puede comprender los GCC más uno o más marcadores adicionales que aumentan significativamente la capacidad de predicción del panel (es decir, mejoran significativamente la capacidad de predicción con respecto a si el panel consistió solamente de los GCC). Se puede utilizar cualquier otra combinación de GCC (incluyendo cualquiera de los enumerados en la Tabla 1 o los paneles A a G) de acuerdo con la presente divulgación.

Se ha descubierto que los GCC son particularmente predictivos en ciertos tipos de cáncer. Por ejemplo, se ha determinado que los paneles de GCC son exactos en la predicción de la recurrencia en el cáncer de próstata (ejemplos 1 a 5). Además, como se divulga en el presente documento, los GCC pueden determinar el pronóstico el cáncer de vejiga, de cerebro, de mama y de pulmón, tal como se resume en el ejemplo 6 y en las tablas 21 y 22 a continuación.

Por tanto, la divulgación describe un procedimiento que comprende determinar el estado de un panel de genes que comprende al menos dos GCC, en el que un estado anormal indica un mal pronóstico. El panel puede comprender al menos 2 genes seleccionados del grupo de genes en al menos uno de los paneles A a G. El panel puede comprender al menos 10 genes seleccionados del grupo de genes en al menos uno de los paneles A a G. El panel puede comprender al menos 15 genes seleccionados del grupo de genes en al menos uno de los paneles A a G. En algunas realizaciones, el panel comprende todos los genes enumerados en uno cualquiera de los paneles A a G. La divulgación también describe un procedimiento para determinar el pronóstico del cáncer de vejiga, que comprende determinar el estado de un panel de genes que comprende al menos dos GCC (por ejemplo, al menos dos de los genes en cualquiera de los paneles B, C y F), en el que un estado anormal indica un mal pronóstico. La divulgación también describe un procedimiento para determinar el pronóstico del cáncer de cerebro, que comprende determinar el estado de un panel de genes que comprende al menos dos GCC (por ejemplo, al menos dos de los genes en cualquiera de los paneles B, C y F), en el que un estado anormal indica un mal pronóstico. La divulgación describe adicionalmente un procedimiento para determinar el pronóstico del cáncer de mama, que comprende determinar el estado de un panel de genes que comprende al menos dos GCC (por ejemplo, al menos dos de los genes en cualquiera de los paneles B, C y F), en el que un estado anormal indica un mal pronóstico. La divulgación también describe un procedimiento para determinar el pronóstico del cáncer de pulmón, que comprende determinar el estado de un panel de genes que comprende al menos dos GCC (por ejemplo, al menos dos de los genes en cualquiera de los paneles B, C y F), en el que un estado anormal indica un mal pronóstico.

El panel puede comprender al menos 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50 o más GCC. El panel puede comprender entre 5 y 100 GCC, entre 7 y 40 GCC, entre 5 y 25 GCC, entre 10 y 20 GCC o entre 10 y 15 GCC. Los GCC pueden comprender al menos una cierta proporción del panel. Por tanto, el panel puede comprender al menos el 25%, 30%, 40%, 50%, 60%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% o 99% de GCC. Los GCC se pueden elegir del grupo que consiste en los genes enumerados en la Tabla 1 y de los paneles A a G. El panel puede comprender al menos 2 genes seleccionados del grupo de genes en cualquiera de los paneles A a G. El panel puede comprender al menos 10 genes seleccionados del grupo de genes en cualquiera de los paneles A a G. El panel puede comprender al menos 15 genes seleccionados del grupo de genes en cualquiera de los paneles A a G. En algunas realizaciones, el panel comprende todos los genes enumerados en cualquiera de los paneles C a G.

Se ha descubierto además que el estado de CCG se añade de forma sinérgica a los parámetros clínicos en el pronóstico del cáncer. En el caso de cáncer de próstata, por ejemplo, se ha descubierto que un alto nivel de la expresión génica de uno cualquiera de los genes en cualquiera de los paneles C a F se asocia con un mayor riesgo de recurrencia o progresión del cáncer de próstata en pacientes cuya puntuación en el nomograma clínico indica un riesgo relativamente bajo de recurrencia o progresión. Por tanto, debido a que la evaluación de los niveles de expresión de CCG puede detectar un aumento de riesgo no detectado usando parámetros clínicos solo, se divulgan procedimientos que combinan la evaluación de al menos un parámetro clínico con la evaluación del estado de al menos un GCC.

Como muestra el ejemplo 3, incluso los GCC individuales se suman a los parámetros clínicos para predecir la recurrencia del cáncer. Por tanto, se divulga un procedimiento diagnóstico in vitro que comprende determinar al menos un parámetro clínico para un paciente con cáncer y determinar el estado de al menos un GCC en una muestra obtenida del paciente. Sin embargo, la evaluación del estado de diversos GCC mejora la capacidad de predicción aún más (también mostrado en el Ejemplo 1). Por tanto, se puede determinar el estado de una pluralidad de GCC (por ejemplo, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50 o más). Un estado anormal puede indicar un aumento de la probabilidad de recurrencia o progresión. El paciente puede tener cáncer de próstata. El paciente puede tener cáncer de pulmón. A menudo, el parámetro clínico es, al menos de un modo algo independientemente, predictivo de recurrencia o progresión y la adición del estado de GCC mejora la capacidad de predicción. Tal como se usa en el presente documento, "parámetro clínico" y "medida clínica" se refieren a características de la enfermedad o del paciente que se aplican normalmente para evaluar la evolución de la enfermedad y / o predecir el resultado. Ejemplos de cáncer incluyen generalmente el estadio del tumor, el grado del tumor, el estado de los ganglios linfáticos, la histología, el estado funcional, el tipo de cirugía, los márgenes quirúrgicos, el tipo de tratamiento y la edad de inicio. En el cáncer de próstata, los clínicos a menudo usan los niveles de PSA en sangre precirugía, el estadio (definido por el tamaño del tumor y la evidencia de metástasis) y la puntuación de Gleason (similar al concepto de grado). Después de la intervención quirúrgica, los parámetros clínicos importantes en el cáncer de próstata incluyen el margen y el estado de los ganglios linfáticos. En el cáncer de mama, los clínicos a menudo usan el tamaño de la lesión índice en cm, la invasión, el número de ganglios afectados y el grado.

A menudo, determinados parámetros clínicos se correlacionan con un carácter de la enfermedad concreta. Por ejemplo, en el cáncer en general, así como en tipos específicos de cáncer, ciertos parámetros clínicos se correlacionan con, por ejemplo, la probabilidad de recurrencia o metástasis, el pronóstico para la supervivencia para una cierta cantidad de tiempo, la probabilidad de respuesta al tratamiento en general o a un tratamiento específico, etc. En el cáncer de próstata, algunos parámetros clínicos son tales que su estado (presencia, ausencia, nivel, etc.) se asocia con un aumento de la probabilidad de recurrencia. Entre los ejemplos de tales parámetros asociados a recurrencia (algunos, pero no todos los cuales son específicos del cáncer de próstata) se incluyen niveles altos de PSA (por ejemplo, mayores que 4 ng / ml), puntuación alta de Gleason, tamaño grande del tumor, evidencia de metástasis, el estadio avanzado del tumor, grado nuclear, afectación de los ganglios linfáticos, edad temprana de inicio. Otros tipos de cáncer pueden tener diferentes parámetros correlacionados con la probabilidad de

recurrencia o progresión y el estado de GCC, como una medida de la actividad proliferativa, se añade a estos parámetros en la predicción del pronóstico en estos tipos de cáncer. Como se usa en el presente documento, "parámetro clínico asociado con recurrencia" tiene su significado convencional para cada cáncer específico, con el que los expertos en la materia están muy familiarizados. De hecho, los expertos en la materia están familiarizados con diversos parámetros clínicos asociados con recurrencia además de los enumerados en el presente documento.

A menudo, un médico evaluará más de un parámetro clínico en un paciente y hará una evaluación más completa de los caracteres patológicos de interés. El ejemplo 5 muestra cómo el estado de GCC puede añadir a una agrupación particular de parámetros clínicos utilizados para determinar el riesgo de recurrencia en el cáncer de próstata. Los parámetros clínicos en el Ejemplo 5 incluyen variables binarias para la enfermedad limitada al órgano y una puntuación de Gleason de menos de o igual a 6, y una variable continua para el PSA logarítmico (Tabla 14). Este modelo incluye todos los parámetros clínicos incorporados en el nomograma post-RP (es decir, nomograma de Kattan -Stephenson), excepto para el Año de la RP y los dos componentes de la puntuación de Gleason. Por tanto, se evalúan al menos dos parámetros clínicos (por ejemplo, dos de los parámetros enumerados anteriormente) junto con el nivel de expresión de al menos un GCC.

Una forma en que los médicos utilizan parámetros clínicos individuales, pero más a menudo diversos, es con la ayuda de nomogramas. En el contexto clínico, los nomogramas son representaciones (a menudo visuales) de una correlación entre uno o más parámetros y uno o más caracteres de los pacientes o la enfermedad. Un ejemplo de un nomograma clínico prevalente utilizado en la determinación de la probabilidad de recurrencia en el paciente de cáncer de próstata se describe en Kattan y col., *CLIN. Oncol.* (1999) 17:1499-1507, y se actualiza en Stephenson *et al.*, *J. CLIN. Oncol.* (2005) 23:7005-7012 ("Kattan-Stephenson nomogram"). Este nomograma evalúa a un paciente mediante la asignación de un valor en puntos a cada uno de diversos parámetros clínicos (año de RP, márgenes quirúrgicos, extensión extracapsular, invasión de vesículas seminales, afectación de los ganglios linfáticos, puntuación primaria de Gleason, puntuación secundaria de Gleason y nivel de PSA preoperatorio), sumando el total de los puntos para un paciente en una puntuación del nomograma y, a continuación, prediciendo la probabilidad del paciente de estar sin recurrencia estar a varios intervalos de tiempo (hasta 10 años) según la puntuación del nomograma. Un ejemplo de un nomograma clínico prevalente usado para determinar el pronóstico de un paciente de cáncer de mama para la supervivencia es el Índice pronóstico de Nottingham (NPI). Véase, por ejemplo, Galea y col., *BREAST CANCER RES. & TREAT.* (1992) 22:207-19).

Se ha descubierto que la determinación del estado de un GCC en una muestra obtenida de un paciente con cáncer de próstata, junto con la puntuación del nomograma de Kattan-Stephenson del paciente, es un mejor predictor de la supervivencia libre de recurrencia a 10 años que la puntuación del nomograma solo. Véanse, por ejemplo, los ejemplos 2 y 5, más adelante. Específicamente, la adición del estado de GCC al nomograma de Kattan-Stephenson detecta pacientes en riesgo significativamente mayor de recurrencia que el nomograma por sí solo no lo hace. La Tabla 3 anterior proporciona un panel de ejemplo de 31 GCC (Panel C) y un panel de subconjunto de 26 GCC (Panel D, mostrado con \*) determinados en el Ejemplo 2 para mostrar sinergia predictiva con el nomograma de Kattan-Stephenson en el pronóstico del cáncer de próstata. También se ha descubierto que la determinación del estado de un GCC en una muestra obtenida de un paciente de cáncer de mama, junto con la puntuación de NPI del paciente, es un predictor pronóstico mejor que la puntuación NPI sola. Véase, por ejemplo, el Ejemplo 6, a continuación. Específicamente, la adición del estado de GCC al nomograma de NPI detecta pacientes en riesgo significativamente mayor de recurrencia que el nomograma por sí solo no lo hace. En el Ejemplo 2 se determinó que los paneles B, C y D mostraban sinergia predictiva con el nomograma de NPI en el pronóstico del cáncer de mama.

Por tanto, se divulga un procedimiento diagnóstico *in vitro* que comprende determinar la puntuación de un nomograma clínico (por ejemplo, puntuación del nomograma de Kattan-Stephenson o de NPI) para un paciente con cáncer y determinar el estado de al menos un GCC en una muestra obtenida del paciente. El ejemplo 3 ilustra la determinación empírica de la capacidad de predicción de los GCC individuales y de varios paneles de GCC de tamaño variable sobre el nomograma de Kattan-Stephenson. Los subconjuntos seleccionados al azar de los 31 GCC enumerados en la Tabla 3 se analizaron como firmas de GCC distintas y se determinó la capacidad de predicción (es decir, valor de p) para cada uno. Como se muestra en la figura 1, las firmas de GCC de 2, 3, 4, 5, 6, 10, 15, 20, 25 y 26 genes se añaden cada uno a la capacidad de predicción al nomograma. Por lo tanto, se divulga un procedimiento para determinar si un paciente de cáncer de próstata tiene un aumento de la probabilidad de recurrencia, que comprende determinar el estado de un panel de genes que comprende al menos 2, 3, 4, 5, 6, 10, 15, 20, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 40, 45, 50, 60, 70, 80, 90, o 100 o más GCC, en el que un estado elevado (por ejemplo, aumento de la expresión) de los GCC indica un aumento de la probabilidad de recurrencia. El procedimiento puede comprender además la determinación de una puntuación de nomograma clínico del paciente. La divulgación describe adicionalmente un procedimiento para determinar si un paciente de cáncer de mama tiene un aumento de la probabilidad de recurrencia, que comprende determinar el estado de un panel de genes que comprende al menos 2, 3, 4, 5, 6, 10, 15, 20, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 40, 45, 50, 60, 70, 80, 90, o 100 o más GCC, en el que un estado elevado (por ejemplo, aumento de la expresión) de los GCC indica un aumento de la probabilidad de recurrencia. El procedimiento puede comprender además la determinación de una puntuación de nomograma clínico del paciente.

A menudo, los nomogramas clínicos para cáncer están diseñados de tal manera que un valor en concreto (por ejemplo, puntuación alta) se correlaciona con un aumento del riesgo de recurrencia. Un estado de GCC elevado (por

ejemplo, aumento de la expresión o actividad) también se correlaciona con un mayor riesgo. Por tanto, se divulga un procedimiento para determinar si un paciente con cáncer tiene un aumento de la probabilidad de recurrencia o progresión, que comprende determinar una puntuación del nomograma clínico para el paciente y determinar el estado de al menos un GCC en una muestra obtenida del paciente, en el que una puntuación más alta en el nomograma y / o un estado de GCC elevado indican que el paciente tiene un aumento de la probabilidad de recurrencia o progresión. El cáncer puede ser cáncer de próstata. El cáncer puede ser cáncer de pulmón.

Esta evaluación se puede realizar antes de la prostatectomía radical (por ejemplo, utilizando una muestra de biopsia de próstata), aunque también se puede realizar después (por ejemplo, utilizando la muestra de próstata resecada). Se obtiene una muestra de una o más células de un paciente de cáncer de próstata antes de o después del tratamiento para análisis. El tratamiento del cáncer de próstata aplicado actualmente en la técnica incluye, por ejemplo, prostatectomía, radioterapia, terapia hormonal (por ejemplo, usando antagonistas de la GnRH, agonistas de la GnRH, antiandrógenos), quimioterapia y ultrasonidos enfocados de alta intensidad. Se pueden obtener una o más células de tumor de próstata a partir de tejido de cáncer de próstata de un paciente de cáncer de próstata durante la biopsia o la prostatectomía y se pueden utilizar para el análisis.

La presente invención también se basa en el descubrimiento de que el estado de PTEN predice el cáncer de próstata agresivo. El estado de PTEN se suma a ambos parámetros clínicos (por ejemplo, nomograma de Kattan-Stephenson) y GCC (por ejemplo, los genes de la Tabla 1 o de los paneles A a G). Como se describe con más detalle en el Ejemplo 4 a continuación, el estado de PTEN se determinó en 191 muestras de pacientes con cáncer de próstata acompañado de los datos del historial clínico y los datos de la firma de GCC. Se encontró que el estado de PTEN negativo era un predictor significativo del riesgo de recurrencia (valor de p de 0,031). El PTEN siguió siendo un predictor significativo de recurrencia después de ajustar por parámetros clínicos posquirúrgicos y la firma GCC mostrada en la Tabla 3 (valor de p 0,026). Además, y más importante, la combinación de PTEN y la firma de GCC parece ser un mejor predictor de recurrencia que los parámetros clínicos posquirúrgicos (valor de p 0,0002).

Debido a que PTEN es un predictor independiente de recurrencia del cáncer de próstata, se divulga un procedimiento de predicción de la probabilidad que tiene un paciente de recurrencia del cáncer de próstata que comprende determinar el estado de PTEN en una muestra del paciente, en el que un estado bajo o negativo de PTEN indica que el paciente tiene una alta probabilidad de recurrencia. El estado de PTEN se puede determinar mediante cualquier técnica conocida en la técnica, incluyendo, pero no limitadas a, las tratadas en el presente documento.

Debido a que el PTEN se añade al estado de GCC en la predicción de la recurrencia del cáncer de próstata, otro aspecto de la invención proporciona un procedimiento in vitro que comprende determinar el estado de PTEN y determinar el estado de una pluralidad de GCC como se define en las reivindicaciones en una muestra obtenida de un paciente. Se pueden utilizar diferentes combinaciones de técnicas para determinar el estado de los diversos marcadores. Por ejemplo, el estado de PTEN se determina mediante técnicas de inmunohistoquímica (IHQ), mientras que el estado de la pluralidad de GCC se determina mediante reacción en cadena de la polimerasa cuantitativa (PCRc), por ejemplo, TaqMan™. También se divulga un procedimiento de determinación de la probabilidad de recurrencia de un paciente de cáncer de próstata que comprende determinar el estado de PTEN en una muestra obtenida del paciente del paciente, determinar el estado de una pluralidad de GCC en una muestra obtenida del paciente, en el que un estado de PTEN bajo o negativo y / o el estado elevado de GCC indican que el paciente tiene un aumento de la probabilidad de recurrencia.

Debido a que el estado de PTEN añade valor predictivo a los parámetros clínicos en la predicción de la recurrencia de próstata, otro aspecto más de la invención proporciona un procedimiento in vitro que comprende determinar el estado de PTEN y determinar al menos un parámetro clínico para un paciente de cáncer. A menudo, el parámetro clínico es, al menos de un modo algo independiente, predictivo de recurrencia o progresión y la adición del estado de PTEN mejora el poder predictivo. Se divulga un procedimiento para determinar si un paciente con cáncer tiene un aumento de la probabilidad de recurrencia, que comprende determinar el estado de PTEN en una muestra obtenida del paciente y determinar la puntuación en un nomograma clínico para el paciente, en el que un estado bajo o negativo de PTEN y/o una puntuación alta en el nomograma indican que el paciente tiene un aumento de la probabilidad de recurrencia.

Debido a que los tres marcadores anteriores son aditivos, se divulga un procedimiento para determinar si un paciente con cáncer tiene un aumento de la probabilidad de recurrencia, que comprende determinar el estado de PTEN en una muestra obtenida del paciente y determinar la puntuación en un nomograma clínico para el paciente, en el que un estado bajo o negativo de PTEN y una puntuación alta en el nomograma indican que el paciente tiene un aumento de la probabilidad de recurrencia.

Los resultados de todos los análisis de acuerdo con la invención a menudo se comunicarán a los médicos, los asesores genéticos y / o los pacientes (u otras partes interesadas, tales como los investigadores) en una forma transmisible que pueda comunicarse o transmitirse a cualquiera de las partes anteriores Tal forma puede variar y puede ser tangible o intangible. Los resultados pueden incluirse en los enunciados descriptivos, diagramas, fotografías, gráficos, imágenes o cualquier otra forma visual. Por ejemplo, los gráficos que muestran la expresión o el nivel de actividad o la información sobre la variación de la secuencia de varios genes se pueden utilizar en la

explicación de los resultados. Los diagramas que muestran dicha información para el gen o genes diana adicionales también son útiles en la indicación de algunos resultados de las pruebas. Las declaraciones y las formas visuales se pueden grabar en un medio tangible, tales como papeles, medios legibles por ordenador, tales como disquetes, discos compactos, etc., o en un medio inmaterial, por ejemplo, un medio electrónico en forma de correo electrónico o página web en internet o una intranet. Además, los resultados también se pueden grabar en forma de sonido y transmitirse a través de cualquier medio adecuado, por ejemplo, líneas de cable analógico o digital, cables de fibra óptica, etc., a través de teléfono, fax, teléfono móvil inalámbrico, teléfono de Internet y similares.

De este modo, la información y los datos sobre un resultado del ensayo se pueden producir en cualquier parte del mundo y transmitir a una ubicación diferente. Como ejemplo ilustrativo, cuando un ensayo del nivel de expresión, el nivel de actividad o la secuenciación (o genotipado) se lleva a cabo fuera de los Estados Unidos, se pueden generar la información y los datos sobre un resultado del ensayo, fundir en una forma transmisible como se ha descrito anteriormente y, después, importar a Estados Unidos. En consecuencia, la presente divulgación también abarca un procedimiento para producir una forma transmisible de información sobre al menos uno de (a) el nivel de expresión o (b) el nivel de actividad para al menos una muestra del paciente. El procedimiento comprende las etapas de (1) determinar al menos uno de (a) o (b) anteriores de acuerdo con los procedimientos de la presente invención; y (2) incluir el resultado de la etapa de determinación en una forma transmisible. La forma transmisible es el producto de un procedimiento de este tipo.

Las técnicas para el análisis de dichos datos de expresión, actividad, y / o secuencia (de hecho, cualquier dato obtenido de acuerdo con la invención) a menudo se implementará utilizando hardware, software o una combinación de los mismos en uno o más sistemas informáticos u otros sistemas de procesamiento capaces de efectuar tal análisis.

La función del análisis basado en ordenador se puede implementar en cualquier idioma y / o navegadores adecuados. Por ejemplo, se puede implementar con el lenguaje C y, preferiblemente, utilizando lenguajes de programación de alto nivel orientados a objetos, tales como Visual Basic, SmallTalk, C ++, y similares. La aplicación se puede escribir para adaptarse a entornos como el entorno de Microsoft Windows™, incluyendo Windows™ 98, Windows™ 2000, Windows NT™, y similares. Además, la aplicación también puede escribirse para el entorno Macintosh™, SUN™, UNIX o LINUX. Además, las etapas funcionales también pueden implementarse utilizando un lenguaje de programación universal o independiente de la plataforma. Ejemplos de tales lenguajes de programación multiplataforma incluyen, pero no se limitan a, lenguaje de marcado de hipertexto ((HTML), JAVA™, JavaScript™, lenguaje de programación Flash, lenguaje Common Gateway Interface / estructurado de consultas (CGI / SQL), lenguaje del informe de extracción práctica (PERL), AppleScript™ y otros lenguajes de script del sistema, lenguaje de programación / lenguaje de consulta estructurada (PL / SQL), y similares. Se pueden usar buscadores habilitados para Java™— o JavaScript™, tales como HotJava™, Microsoft™ Explorer™, o Netscape™. Cuando se utilizan las páginas web de contenido activo, pueden incluir applets de Java o controles ActiveX u otras tecnologías de contenido activo.

La función de análisis también se puede incluir en los productos del programa de ordenador y se utilizan en los sistemas descritos anteriormente u otros sistemas por ordenador o basados en Internet. Por consiguiente, se divulga un producto de programa de ordenador que comprende un medio utilizable por ordenador que tiene códigos o instrucciones de programa informático contenidos en el mismo para permitir que un procesador lleve a cabo el análisis del estado del gen. Estas instrucciones del programa informático pueden cargarse en un ordenador u otro aparato programable para producir una máquina, de un modo tal que las instrucciones que se ejecutan en el ordenador u otro aparato programable crean medios para implementar las funciones o etapas descritas anteriormente. Estas instrucciones del programa informático también se pueden almacenar en una memoria o medio legible por ordenador que puede dirigir la función de un ordenador u otro aparato programable de una manera particular, de un modo tal que las instrucciones almacenadas en la memoria o medio legible por ordenador producen un artículo de fabricación, incluyendo medios de instrucción que implementan el análisis. Las instrucciones del programa informático también pueden cargarse en un ordenador u otro aparato programable para provocar una serie de etapas operacionales que se van a realizar en el ordenador u otro aparato programable para producir un proceso implementado por ordenador de modo que las instrucciones que se ejecutan en el ordenador u otro aparato programable proporcionen etapas para implementar las funciones o etapas descritas anteriormente.

También se divulga un sistema para determinar si un paciente tiene un aumento de la probabilidad de recurrencia. En términos generales, el sistema comprende (1) medios de ordenador para recibir, almacenar y / o recuperar datos del estado génico de un paciente (por ejemplo, el nivel de expresión, el nivel de actividad, las variantes) y, opcionalmente, datos de parámetros clínicos (por ejemplo, puntuación de Gleason, puntuación del nomograma); (2) medios de ordenador para consultar los datos de este paciente; (3) medios informáticos que permiten concluir si existe un aumento de la probabilidad de recurrencia en base a los datos de este paciente; y (4) medios para la exposición / presentación de esta conclusión.

Los medios para exponer la conclusión pueden comprender un medio informático para informar a un profesional de la salud sobre la conclusión.

Un ejemplo de tal sistema es el sistema informático [600] ilustrado en la figura 6. El sistema informático [600] puede

5 incluir al menos un módulo de entrada [630] para introducir los datos del paciente en el sistema informático [600]. El sistema informático [600] puede incluir al menos un módulo de salida [624] para indicar si un paciente tiene una probabilidad mayor o menor de respuesta y / o indicar tratamientos sugeridos determinados por el sistema informático [600]. El sistema informático [600] puede incluir al menos un módulo de memoria [606] en comunicación con el al menos un módulo de entrada [630] y el al menos un módulo de salida [624].

10 El al menos un módulo de memoria [606] puede incluir, por ejemplo, una unidad de almacenamiento extraíble [608], que puede estar en varias formas, incluyendo, pero sin limitaciones, una unidad de cinta magnética, una unidad de disco flexible, una unidad de CD, una unidad de DVD, una unidad de disco óptico, etc. La unidad de almacenamiento extraíble [608] puede ser compatible con una unidad de almacenamiento extraíble [610] de modo que se pueda leer y / o escribir en la unidad de almacenamiento extraíble [610]. La unidad de almacenamiento extraíble [610] puede incluir un medio de almacenamiento utilizable por ordenador que tiene almacenado en el mismo código de programa legibles por ordenador o instrucciones y / o datos legibles por ordenador. Por ejemplo, la unidad de almacenamiento extraíble [610] puede almacenar los datos del paciente. En la técnica se conocen ejemplos de unidad de almacenamiento extraíble [610], incluyendo, pero sin limitaciones, disquetes, cintas magnéticas, discos ópticos, y similares. El al menos un módulo de memoria [606] también puede incluir una unidad de disco duro [612], que se puede utilizar para almacenar códigos de programa o instrucciones legibles por ordenador y / o datos legibles por ordenador.

20 Además, como se muestra en la Figura 1, el al menos un módulo de memoria [606] puede incluir además una interfaz [614] y una unidad de almacenamiento extraíble [616] que es compatible con la interfaz [614], de forma que el software, los códigos o las instrucciones legibles por ordenador pueden transferirse desde la unidad de almacenamiento extraíble [616] al sistema informático [600]. Entre los ejemplos de pares de interfaz [614] y unidad de almacenamiento extraíble [616] se incluyen, por ejemplo, chips de memoria extraíble (por ejemplo, EPROM o PROM) y tomas de corriente asociado con los mismos, cartuchos de programa e interfaz de cartucho, y similares. El sistema informático [600] también puede incluir un módulo secundario de memoria [618], tal como una memoria de acceso aleatorio (RAM).

30 El sistema informático [600] puede incluir al menos un módulo de procesador [602]. Se debe entender que el al menos un módulo de procesador [602] puede consistir en cualquier número de dispositivos. El al menos un módulo de procesador [602] puede incluir un dispositivo de procesamiento de datos, tal como un microprocesador o microcontrolador o una unidad de procesamiento central. El al menos un módulo de procesador [602] puede incluir otro dispositivo lógico, tal como un procesador DMA (acceso directo a la memoria), un dispositivo procesador de comunicación integrado, un dispositivo VLSI (integración a escala muy grande) personalizado o un dispositivo ASIC (circuito integrado específico de aplicación) dispositivo. Además, el al menos un módulo procesador [602] puede incluir cualquier otro tipo de circuito analógico o digital que está diseñado para realizar las funciones de procesamiento descritas en el presente documento.

35 Como se muestra en la figura 6, en el sistema informático [600], el al menos un módulo de memoria [606], el al menos un módulo procesador [602] y el módulo de memoria secundaria [618] están todos operativamente unidos a través de infraestructura de comunicación [620], que puede ser un bus de comunicaciones, la placa del sistema, la barra transversal, etc.). A través de la infraestructura de comunicación [620], los códigos o instrucciones de programa de ordenador o los datos legibles por ordenador pueden ser transferidos e intercambiados. La interfaz de entrada [626] se puede conectar operativamente el al menos un módulo de entrada [626] a la infraestructura de comunicación [620]. Asimismo, la interfaz de salida [622] se puede conectar operativamente el al menos un módulo de salida [624] a la infraestructura de comunicación [620].

45 El al menos un módulo de entrada [630] puede incluir, por ejemplo, un teclado, ratón, pantalla táctil, escáner y otros dispositivos de entrada conocidos en la técnica. El al menos un módulo de salida [624] puede incluir, por ejemplo, una pantalla de visualización, tal como un monitor de ordenador, monitor de televisión, o la pantalla táctil del al menos un módulo de entrada [630]; una impresora y altavoces de audio. El sistema informático [600] también puede incluir, módem, puertos de comunicación, tarjetas de red, tales como tarjetas de Ethernet, y dispositivos desarrollados recientemente para el acceso a intranet o a Internet.

50 El al menos un módulo de memoria [606] se puede configurar para almacenar los datos del paciente introducidos a través del al menos un módulo de entrada [630] y se puede procesar mediante el al menos un módulo procesador [602]. Los datos relevantes del paciente en el presente documento pueden incluir el nivel de expresión, el nivel de actividad, el número de copias y / o la información de la secuencia de PTEN y / o un GCC. Los datos relevantes del paciente en el presente documento también pueden incluir parámetros clínicos relevantes para la enfermedad del paciente. Cualesquiera otros datos del paciente que podrían ser de utilidad para un médico en la toma de decisiones/recomendaciones sobre el tratamiento también pueden introducirse en el sistema, incluyendo, pero sin limitaciones, datos sobre la edad, el sexo y la raza / origen étnico y el estilo de vida, tal como información sobre la dieta. Otros tipos posibles de los datos sobre el paciente incluyen síntomas que se están experimentando actualmente o se han experimentado previamente, el historial de enfermedades del paciente, los medicamentos y los procedimientos médicos.

60 El al menos un módulo de memoria [606] puede incluir un procedimiento implementado por ordenador almacenado

en el mismo. El al menos un módulo procesador [602] se puede utilizar para ejecutar software o códigos de instrucciones legibles por ordenador del procedimiento implementado por ordenador. El procedimiento implementado por ordenador puede estar configurado para, en base a los datos del paciente, indicar si el paciente tiene un aumento de la probabilidad de recurrencia, progresión o respuesta a un tratamiento en particular, generar una lista de posibles tratamientos, etc.

El procedimiento implementado por ordenador puede estar configurado para identificar a un paciente que tiene o no tiene un aumento de la probabilidad de recurrencia o progresión. Por ejemplo, el procedimiento implementado por ordenador puede estar configurado para informar a un médico de que un paciente en particular tiene un aumento de la probabilidad de recurrencia. Como alternativa o adicionalmente, el procedimiento implementado por ordenador puede estar configurado para sugerir en realidad un curso de tratamiento determinado basado en las respuestas a varias preguntas o los resultados de las mismas.

La figura 7 ilustra un procedimiento implementado por ordenador [700] que puede implementarse con el sistema informático [600] de la divulgación. El procedimiento [700] comienza con una de tres preguntas ([710], [711], [712]), ya sea de forma secuencial o sustancialmente de forma simultánea. Si la respuesta a cualquiera de estas preguntas o el resultado de las mismas es "Sí" [720], el procedimiento concluye [730] que el paciente tiene un aumento de la probabilidad de recurrencia. Si la respuesta a cualquiera de todas preguntas o el resultado de las mismas es "No" [721], el procedimiento concluye [731] que el paciente no tiene un aumento de la probabilidad de recurrencia. A continuación, el procedimiento [700] puede proceder con más preguntas, hacer una recomendación de tratamiento en particular ([740], [741]), o simplemente terminar.

Cuando las preguntas se realizan de forma secuencial, se pueden realizar en el orden sugerido por la figura 7, o en cualquier otro orden. Si se hacen o no las preguntas posteriores también puede depender de los resultados / respuestas a las preguntas anteriores. En el procedimiento ilustrado en la figura 7, por ejemplo, el procedimiento puede preguntar acerca de los parámetros clínicos [712] primero y, si el paciente tiene uno o más parámetros clínicos que identifican al paciente como en mayor riesgo de recurrencia, el procedimiento concluye tales [730] u opcionalmente confirma preguntado estado de GCC, mientras que si el paciente no tiene tales parámetros clínicos, el procedimiento procede a preguntar sobre el estado de GCC [711]. Opcionalmente, si el estado de GCC no es elevado, el procedimiento puede continuar y preguntar sobre el estado de PTEN [710]. Como se ha mencionado anteriormente, el orden anterior de las preguntas puede modificarse. Una respuesta de "sí" a una pregunta (por ejemplo, [712]) plantea una o más de las preguntas restantes para confirmar que el paciente tiene un mayor riesgo de recurrencia.

El procedimiento implementado por ordenador de la divulgación [700] puede finalizarse de forma abierta. En otras palabras, la primera etapa aparente [710, 711, y/o 712] en la figura 7 puede realmente formar parte de un proceso más largo y, dentro de este proceso más largo, no necesitar ser la primera etapa / pregunta. Se pueden añadir también etapas adicionales a los procedimientos básicos descritos anteriormente. Estas etapas adicionales incluyen, pero no se limitan a, informar a un profesional de la salud (o al propio paciente) de la conclusión alcanzada; combinar la conclusión alcanzada por el procedimiento ilustrado [700] con otros hechos o conclusiones para llegar a una conclusión adicional o perfeccionada con respecto al diagnóstico, pronóstico, tratamiento etc. del paciente; hacer una recomendación para el tratamiento (por ejemplo, "el paciente debería / no debería someterse a prostatectomía radical"); preguntas adicionales sobre biomarcadores adicionales, parámetros clínicos, u otra información útil sobre el paciente (por ejemplo, la edad en el momento del diagnóstico, la salud general del paciente, etc.).

En cuanto al procedimiento implementado por ordenador anterior [700], las respuestas a las preguntas se pueden determinar mediante el procedimiento que establece una búsqueda en los datos de los pacientes para la respuesta. Por ejemplo, para responder a las preguntas respectivas [710, 711, 712], se puede hacer una búsqueda en los datos del paciente del estado de PTEN (por ejemplo, IHQ de PTEN o detección de mutaciones), el estado de GCC (por ejemplo, los datos del nivel de expresión de GCC) o los parámetros clínicos (por ejemplo, puntuación de Gleason, puntuación del nomograma, etc.). Si dicha comparación no se ha realizado todavía, el procedimiento puede comparar estos datos con alguna referencia con el fin de determinar si el paciente tiene un estado anormal (por ejemplo, elevado, bajo, negativo). Adicional o alternativamente, el procedimiento puede presentar una o más de las preguntas [710, 711, 712] a un usuario (por ejemplo, un médico) del sistema informático [100]. Por ejemplo, las preguntas [710, 711, 712] se pueden presentar a través de un módulo de salida [624]. El usuario puede responder "Sí" o "No" a través de un módulo de entrada [630]. A continuación, el procedimiento puede proceder en base a la respuesta recibida. Del mismo modo, las conclusiones [730, 731] se pueden presentar a un usuario del procedimiento implementado por ordenador a través de un módulo de salida [624].

También se divulga un procedimiento que comprende: acceder a la información sobre el estado de GCC de un paciente, los parámetros clínicos y / o el estado de PTEN almacenado en un medio legible por ordenador; consultar esta información para determinar al menos uno de si una muestra obtenida del paciente muestra un aumento de la expresión de al menos un GCC, si el paciente tiene un parámetro clínico asociado a la recurrencia y / o si el paciente tiene un estado de PTEN bajo / negativo; la salida de [o visualización] del estado de expresión de GCC de la muestra, el estado del parámetro clínico asociado a recurrencia del paciente, y / o el estado de PTEN de la muestra. Como se usa en el presente documento en el contexto de procedimientos implementados por ordenador de la

divulgación, "mostrar" significa comunicar cualquier información por cualquier medio sensorial. Entre los ejemplos se incluyen, pero no se limitan a, pantallas visuales, por ejemplo, en una pantalla de ordenador o en una hoja de papel impreso en la orden del ordenador, y muestras auditivas, por ejemplo, expresión auditiva generada por ordenador o grabada del genotipo de un paciente.

5 Como se ha tratado con detalle anteriormente, los parámetros clínicos asociados con recurrencia o el estado de PTEN combinados con el estado de GCC elevado indican un aumento significativo de la probabilidad de recurrencia. Por tanto, se divulga un procedimiento implementado por ordenador para determinar si un paciente tiene un aumento de la probabilidad de recurrencia que comprende acceder a la información sobre el estado de PTEN de un paciente (por ejemplo, de una muestra de tumor obtenida del paciente) o los parámetros clínicos y el estado de GCC (por ejemplo, a partir de una muestra de tumor obtenida del paciente) almacenado en un medio legible por ordenador; consultar esta información para determinar al menos uno de si el paciente tiene un estado de PTEN bajo / negativo o si el paciente tiene un parámetro clínico asociado a recurrencia; consultar esta información para determinar si una muestra obtenida del paciente muestra un aumento de la expresión de al menos un GCC; exponer (o presentar) una indicación de que el paciente tiene un aumento de la probabilidad de recurrencia si el paciente tiene un estado de PTEN bajo / negativo o un parámetro clínico asociado a recurrencia y la muestra presenta un aumento de la expresión de al menos un GCC. El procedimiento puede comprender además mostrar PTEN, parámetros clínicos (o sus valores) y / o los GCC y su estado (incluyendo, por ejemplo, los niveles de expresión), opcionalmente junto con una indicación de si el estado de PTEN o GCC y / o el parámetro clínico indica un aumento de la probabilidad del riesgo.

20 La práctica de la presente divulgación también puede emplear procedimientos, software y sistemas de biología convencionales. Los productos de software informático de la divulgación típicamente incluyen medios legibles por ordenador que tienen instrucciones ejecutables por ordenador para realizar las etapas lógicas del procedimiento de la invención. Los medios legibles por ordenador adecuados incluyen disquetes, CD-ROM / DVD / DVD-ROM, unidad de disco duro, memoria flash, ROM / RAM, cintas magnéticas, etc. Los procedimientos básicos de biología computacional se describen en, por ejemplo, Setubal y col., INTRODUCTION TO COMPUTATIONAL BIOLOGY METHODS (PWS Publishing Company, Boston, 1997); Salzberg et.al. (Ed.), COMPUTATIONAL METHODS IN MOLECULAR BIOLOGY, (Elsevier, Amsterdam, 1998); Rashidi & Buehler, BIOINFORMATICS BASICS: APPLICATION IN BIOLOGICAL SCIENCE AND MEDICINE (CRC Press, Londres, 2000); y Ouelette & Bzevanis, BIOINFORMATICS: A PRACTICAL GUIDE FOR ANALYSIS OF GENE AND PROTEINS (Wiley & Sons, Inc., 2ª ed., 2001); véase también la patente de Estados Unidos n.º 6.420.108.

35 La presente divulgación también puede hacer uso de diversos productos de programas informáticos y software para diversos propósitos, tales como el diseño de la sonda, la gestión de datos, análisis, y el funcionamiento del instrumento. Véanse las patentes de EE. UU. n.ºs 5.593.839, 5.795.716, 5.733.729, 5.974.164, 6.066.454, 6.090.555, 6.185.561, 6.188.783, 6.223.127, 6.229.911 y 6.308.170. Adicionalmente, la presente divulgación puede incluir procedimientos para proporcionar información genética a través de redes, tales como internet, como se muestra en los documentos de Estados Unidos con n. de serie 10/197.621 (publicación de Estados Unidos n.º 20030097222); 10/063,559 (publicación de Estados Unidos n.º 20020183936); 10/065,856 (publicación de Estados Unidos n.º 20030100995); 10/065,868 (publicación de Estados Unidos n.º 20030120432); 10/423,403 (publicación de Estados Unidos n.º 20040049354).

40 También se divulgan procedimientos de tratamiento de un paciente con cáncer que comprenden la obtención de información sobre el estado de GCC (por ejemplo, los GCC en la Tabla 1 o los paneles A a G), y recomendar, prescribir o administrar un tratamiento para el paciente de cáncer basado en el estado de GCC. El procedimiento puede incluir además la obtención de información de los parámetros clínicos y / o la obtención de información del estado de PTEN de una muestra del paciente y tratar al paciente con un tratamiento en particular en función de la información sobre el estado de GCC, el parámetro clínico y / o el estado de PTEN. Por ejemplo, la divulgación describe un procedimiento de tratamiento de un cáncer en un paciente, que comprende:

- (1) determinar el estado de al menos un GCC;
- (2) determinar el estado de al menos un parámetro clínico;
- (3) determinar el estado de PTEN en una muestra obtenida del paciente; y
- 50 4) recomendar, prescribir o administrar o bien
  - (a) un tratamiento activo (incluyendo uno agresivo) si el paciente tiene al menos uno de aumento de la expresión del GCC, parámetro clínico asociado a recurrencia o estado de PTEN bajo / negativo, o bien
  - (b) un tratamiento pasivo (o menos agresivo) si el paciente no tiene ninguno de de aumento de la expresión del GCC, parámetro clínico asociado a recurrencia o estado de PTEN bajo / negativo.

55 Si un tratamiento es agresivo o no generalmente dependerá del tipo de cáncer, la edad del paciente, etc. Por ejemplo, en el cáncer de mama, la quimioterapia adyuvante es un tratamiento agresivo frecuente administrado para complementar los tratamientos estándar menos agresivos de cirugía y terapia hormonal. Los expertos en la técnica

están familiarizados con varios otros tratamientos agresivos y menos agresivos para cada tipo de cáncer. "Tratamiento activo" en el cáncer de próstata es bien entendido por los expertos en la técnica y, como se usa en el presente documento, tiene el significado convencional en la técnica. En términos generales, el tratamiento activo en el cáncer de próstata es cualquier otra cosa que no sea "espera vigilante". El tratamiento activo aplicado actualmente en la técnica del tratamiento del cáncer de próstata aplicado actualmente en la técnica incluye, por ejemplo, prostatectomía, radioterapia, terapia hormonal (por ejemplo, usando antagonistas de la GnRH, agonistas de la GnRH, antiandrógenos), quimioterapia y ultrasonidos enfocados de alta intensidad ("HIFU") etc. Cada opción de tratamiento conlleva ciertos riesgos, así como efectos secundarios de gravedad variable, por ejemplo, impotencia, incontinencia urinaria, etc. Por lo tanto, es habitual que los médicos, dependiendo de la edad y el estado de salud general del hombre diagnosticado con cáncer de próstata, recomienden un régimen de la "espera vigilante".

La "espera vigilante", también llamada "vigilancia activa", también tiene su significado convencional en la técnica. Esto generalmente significa la observación y seguimiento regular sin tratamiento invasivo. La espera vigilante se utiliza a veces, por ejemplo, cuando se descubre un cáncer de próstata de crecimiento lento en estadio temprano en un hombre mayor.

La espera vigilante también puede sugerirse cuando los riesgos de la cirugía, la radioterapia o la terapia hormonal superan a los posibles beneficios. Otros tratamientos pueden iniciarse si se desarrollan síntomas o si hay señales de que el crecimiento del cáncer se está acelerando (por ejemplo, el rápido aumento del PSA, el aumento de la puntuación de Gleason en la biopsia de repetición, etc.).

Aunque los hombres que eligen la espera vigilante evitan los riesgos de la cirugía y de la radiación, la espera vigilante conlleva sus propios riesgos, por ejemplo, aumento del riesgo de metástasis. Para los hombres más jóvenes, un ensayo de vigilancia activa no puede significar evitar el tratamiento por completo, pero puede permitir razonablemente un retraso de unos pocos años o más, tiempo durante el cual el impacto sobre la calidad de vida del tratamiento activo se puede evitar. Los datos publicados hasta la fecha sugieren que hombres seleccionados cuidadosamente no se perderán un margen de curación con este enfoque. Los problemas de salud adicionales que se desarrollan con la edad durante el periodo de observación también pueden hacer más difícil que se sometan a cirugía y radioterapia. Por lo tanto, es clínicamente importante determinar cuidadosamente qué pacientes de cáncer de próstata son buenos candidatos para la espera vigilante y qué pacientes deben recibir tratamiento activo.

Por lo tanto, la divulgación describe un procedimiento de tratamiento de un paciente con cáncer de próstata o de proporcionar orientación al tratamiento de un paciente. En este procedimiento, se determina el estado de al menos un GCC (por ejemplo, los de la Tabla 1 o los paneles A a G), al menos un parámetro clínico asociado con recurrencia y / o el estado de PTEN y (a) se recomienda, inicia o continúa un tratamiento activo si una muestra del paciente tiene un estado elevado de al menos un GCC, el paciente tiene al menos un parámetro clínico asociado con recurrencia y / o un estado de PTEN bajo / negativo, o (b) se recomienda, inicia o continúa la espera vigilante si el paciente no tiene ni un estado elevado de al menos un GCC, un parámetro clínico asociado a recurrencia ni un estado de PTEN bajo / negativo. El estado de GCC, el parámetro o parámetros clínicos y el estado de PTEN pueden indicar no solo que se recomienda el tratamiento activo, sino que es preferible un tratamiento activo particular para el paciente (incluyendo tratamientos relativamente agresivos, tales como, por ejemplo, PR y / o terapia adyuvante).

En general, la terapia adyuvante (por ejemplo, quimioterapia, radioterapia, HIFU, terapia hormonal, etc. después de la prostatectomía o radioterapia) no es el estándar de atención en el cáncer de próstata. No obstante, según la presente invención, los médicos pueden ser capaces de determinar qué pacientes de cáncer de próstata tienen una enfermedad particularmente agresiva y, por lo tanto, deben recibir terapia adyuvante. Así, en una realización, la divulgación proporciona un procedimiento de tratamiento de un paciente (por ejemplo, un paciente con cáncer de próstata) que comprende determinar el estado de al menos un GCC (por ejemplo, los de la Tabla 1 o los paneles A a G), el estado de al menos un parámetro clínico asociado a recurrencia y / o el estado de PTEN, y la iniciación de la terapia adyuvante después de prostatectomía o de radioterapia si una muestra del paciente tiene un estado elevado de al menos un GCC, el paciente tiene al menos un parámetro clínico asociado a recurrencia y / o el paciente tiene un estado de PTEN bajo / negativo.

También se divulgan composiciones para su uso en los procedimientos anteriores. Tales composiciones incluyen, pero no se limitan a, sondas de ácido nucleico que se hibridan con PTEN o un GCC (o con cualquier ácido nucleico codificado por las mismas o complementario de las mismas); cebadores de ácido nucleico y parejas de cebadores adecuados para amplificar todo o una parte de PTEN o un GCC o cualquiera ácido nucleico codificado por los mismos; anticuerpos que se unen inmunológicamente a un polipéptido codificado por PTEN o un GCC; conjuntos de sondas que comprenden una pluralidad de dichas sondas de ácidos nucleicos, cebadores de ácidos nucleicos, anticuerpos, y / o polipéptidos; micromatrices que comprenden cualquiera de estos; kits que comprenden cualquiera de estos; etc. También se divulgan procedimientos informáticos, sistemas, software y / o módulos para su uso en los procedimientos anteriores.

También se divulga una sonda que comprende un oligonucleótido aislado capaz de hibridarse selectivamente a un PTEN o al menos uno de los genes en la Tabla 1 o los paneles A a G. Los términos "sonda" y "oligonucleótido" (también "oligo"), cuando se utilizan en el contexto de ácidos nucleicos, se refieren de forma intercambiable a un fragmento o secuencia de ácido nucleico relativamente corto. La divulgación también describe cebadores útiles en

los procedimientos de la invención. Los "cebadores" son sondas capaces, en las condiciones adecuadas y con los reactivos complementarios adecuados, de amplificar selectivamente un ácido nucleico diana (por ejemplo, un gen diana). En el contexto de los ácidos nucleicos, "sonda" se usa en el presente documento para abarcar "cebador", ya que, generalmente, los cebadores pueden también servir como sondas.

5 La sonda puede ser, en general, de cualquier tamaño / longitud adecuada. La sonda tiene una longitud de aproximadamente 8 a 200, 15 a 150, 15 a 100, 15 a 75, 15 a 60, o 20 a 55 bases de longitud. Pueden marcarse con marcadores detectables con cualquier marcador de detección adecuado, incluyendo, pero no limitado a, isótopos radiactivos, fluoróforos, biotina, enzimas (por ejemplo, fosfatasa alcalina), sustratos de enzimas, ligandos y anticuerpos, etc. Véase Jablonski y col., NUCLEIC ACIDS RES. (1986) 14:6115-6128; Nguyen y col., BIOTECHNIQUES (1992) 13:116-123; Rigby y col., J. MOL. BIOL. (1977) 113:237-251. MOL. BIOL. (1977) 113:237-251). De hecho, las sondas pueden modificarse de cualquier manera convencional para diversas aplicaciones de biología molecular. Las técnicas para producir y usar tales sondas de oligonucleótidos son convencionales en la técnica.

15 Las sondas de acuerdo con la invención se pueden usar en las técnicas de hibridación / amplificación / detección descritas anteriormente. Por lo tanto, la divulgación comprende conjuntos de sonda adecuados para su uso en una micromatriz en la detección, amplificación y / o cuantificación de PTEN y / o una pluralidad de GCC. Los conjuntos de sondas pueden tener una cierta proporción de sus sondas dirigidas a GCC, por ejemplo, un conjunto de sondas que consisten en el 10%, 20%, 30%, 40%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% o el 100% de sondas específicas para GCC. El conjunto de sonda puede comprender sondas dirigido a al menos 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 40, 45, 50, 60, 70, 80, 90, 100, 125, 150, 175, 200, 250, 300, 350, 400, 450, 500, 600, 700 u 800 más, o todos, de los genes de la Tabla 1 o los paneles A a G. Tal sonda puede incorporarse en matrices de alta densidad que comprenden 5.000, 10.000, 20.000, 50.000, 100.000, 200.000, 300.000, 400.000, 500.000, 600.000, 700.000, 800.000, 900.000 o 1.000.000 o más sondas diferentes. Los conjuntos de sondas pueden comprender cebadores (por ejemplo, pares de cebadores) para amplificar ácidos nucleicos que comprenden al menos una porción de PTEN o de uno o más de los GCC en la Tabla 1 o los paneles A a G.

25 En otro aspecto de la presente invención, se proporciona un kit para poner en práctica el pronóstico de la presente invención. El kit puede incluir un vehículo para los diversos componentes del kit. El vehículo puede ser un recipiente o soporte, en forma de, por ejemplo, una bolsa, caja, tubo, gradilla, y está opcionalmente compartimentado. El vehículo puede definir un confinamiento cerrado a efectos de seguridad durante el envío y almacenamiento. El kit divulgado en el presente documento puede incluir varios componentes útiles en la determinación del estado de uno o más GCC y uno o más marcadores de genes domésticos, utilizando las técnicas de detección descritas anteriormente. Por ejemplo, el kit puede incluir muchos oligonucleótidos que se hibridan específicamente en condiciones de alta rigurosidad con ARNm o ADNc de los genes en la Tabla 1 o los paneles A a G. Tales oligonucleótidos pueden usarse como cebadores de PCR en reacciones de RT-PCR, o sondas de hibridación.

35 El kit puede comprender reactivos (por ejemplo, sondas, cebadores y o anticuerpos) para la determinación del nivel de expresión de un panel de genes, en el que dicho panel comprende al menos el 25%, 30%, 40%, 50%, 60%, 75%, 80%, 90%, 95%, 99% o el 100% de GCC (por ejemplo, GCC de la Tabla 1 o de cualquiera de los paneles A a G). El kit puede consistir en reactivos (por ejemplo, sondas, cebadores y o anticuerpos) para determinar el nivel de expresión de no más de 2.500 genes, en el que al menos 5, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100, 120, 150, 200, 250, o más de estos genes son GCC (por ejemplo, GCC de la Tabla 1 o de cualquiera de los paneles A a G).

45 Los oligonucleótidos en el kit de detección pueden marcarse con cualquier marcador de detección adecuado, incluyendo, pero no limitado a, isótopos radiactivos, fluoróforos, biotina, enzimas (por ejemplo, fosfatasa alcalina), sustratos de enzimas, ligandos y anticuerpos, etc. Véase Jablonski *et al.*, *Nucleic Acids Res.*, 14:6115-6128 (1986); Nguyen y col., *Biotechniques*, 13:116-123 (1992); Rigby y col., *J. Mol. Biol.*, 113:237-251 (1977)]. Alternativamente, los oligonucleótidos incluidos en el kit no están marcados y, en su lugar, en el kit se proporcionan uno o más marcadores para que los usuarios puedan marcar los oligonucleótidos en el momento de su uso.

50 El kit de detección contiene uno o más anticuerpos selectivamente inmunorreactivos con una o más proteínas codificadas por PTEN o uno o más GCC u, opcionalmente, cualesquiera marcadores adicionales. Entre los ejemplos se incluyen anticuerpos que se unen inmunológicamente a PTEN o una proteína codificada por un gen en la Tabla 1 o los paneles A a G. Los procedimientos para la producción y el uso de tales anticuerpos se han descrito anteriormente con detalle.

55 Varios otros componentes útiles en las técnicas de detección también pueden incluirse en el kit de detección descrito en el presente documento. Entre los ejemplos de tales componentes se incluyen, pero no se limitan a, Taq polimerasa, desoxirribonucleótidos, desoxirribonucleótidos, otros cebadores adecuados para la amplificación de una secuencia de ADN diana, RNasa A, y similares. Además, el kit de detección puede incluir instrucciones sobre cómo utilizar el kit para poner en práctica del procedimiento de pronóstico de la presente invención usando muestras humanas.

#### EJEMPLO 1:

La siguiente firma de gen del ciclo celular (GCC) se analizó para la predicción del tiempo hasta la recurrencia química después de una prostatectomía radical.

<b>Firma de 31 GCC de recurrencia de próstata</b>		
<i>AURKA</i>	<i>DTL</i>	<i>PTTG1</i>
<i>BUB1</i>	<i>FOXM1</i>	<i>RRM2</i>
<i>CCNB1</i>	<i>HMMR</i>	<i>TIMELESS</i>
<i>CCNB2</i>	<i>KIF23</i>	<i>TPX2</i>
<i>CDC2</i>	<i>KPNA2</i>	<i>TRIP13</i>
<i>CDC20</i>	<i>MAD2L1</i>	<i>TTK</i>
<i>CDC45L</i>	<i>MELK</i>	<i>UBE2C</i>
<i>CDCA8</i>	<i>MYBL2</i>	<i>UBE2S</i>
<i>CENPA</i>	<i>NUSAP1</i>	<i>ZWINT</i>
<i>CKS2</i>	<i>PBK</i>	
<i>DLG7</i>	<i>PRC1</i>	

5 La expresión media del ARNm para los 31 GCC anteriores se analizó en 440 muestras de FFPE de tumor de próstata utilizando un modelo de riesgos proporcionales de Cox en Splus 7.1 (Insightful, Inc., Seattle WA). El valor de p para la prueba de razón de probabilidades fue  $3,98 \times 10^{-5}$ .

10 La media de la expresión de GCC es sólida con respecto a la medición de errores y la variación individual entre los genes. Con el fin de determinar el número óptimo de genes del ciclo celular para la firma, la capacidad de predicción de la media se analizó para conjuntos seleccionados al azar de 1 a 30 de los GCC mencionados anteriormente. Esta simulación mostró que hay un número umbral de GCC en un panel que proporciona una capacidad de predicción mejorada significativamente.

#### **EJEMPLO 2:**

15 En un análisis univariado se encontró que un conjunto de 31 GCC (Tabla 3) era un predictor significativo de la recurrencia bioquímica (valor de  $p = 1,8 \times 10^{10^{-9}}$ ) después de PR en pacientes con cáncer de próstata. Esta firma se evaluó adicionalmente para determinar si se añadió a un nomograma clínico establecido para la recurrencia del cáncer de próstata (el nomograma de Kattan-Stephenson). En resumen, el nomograma fue un predictor altamente significativo de recurrencia (valor de  $p = 1,6 \times 10^{-10}$ ) y, después de ajustar para el nomograma, la firma de GCC fue un predictor significativo de recurrencia bioquímica (valor de  $p = 4,8 \times 10^{-5}$ , Tabla 6).

#### Pacientes y procedimientos

20 Se realizó un seguimiento de ochocientos cuatro pacientes RP consecutivos durante una media de 9,5 años. Las características de los pacientes y los resultados del tratamiento de toda la cohorte se han comunicado anteriormente (Swanson y col., Urol Oncol. (2007) 25:110-114). Se disponía de bloques y/o cortes de tejido de la evaluación patológica final con suficiente tejido para el análisis de 430 pacientes. La cohorte se dividió al azar en 212 pacientes utilizados para un entrenamiento y 199 muestras de pacientes como un conjunto de validación.

25 Expresión Génica (procedimientos estadísticos): La asociación entre la recurrencia bioquímica y la expresión de GCC se evaluó utilizando modelos de PH de Cox para el tiempo hasta la recurrencia. Todos los valores de p indicados en este estudio se obtuvieron a partir de una prueba de razón de probabilidades comparando el modelo nulo con el modelo que contiene la variable del ensayo. Un conjunto de 31 GCC (Tabla 3, anterior) se seleccionó al azar. Los ensayos se utilizaron para generar los datos de expresión de 212 pacientes en el conjunto de  
30 entrenamiento. Todos los datos de expresión se generaron por triplicado. Los datos de expresión se combinaron en una firma calculando el nivel medio de expresión de 26 GCC.

La asociación entre la recurrencia bioquímica y la expresión de GCC se evaluó utilizando modelos de PH de Cox para el tiempo hasta la recurrencia.

35 Preparación de la muestra y diseño del estudio: Se aisló ARN de secciones de tumor FFPE obtenidas de 411 pacientes de cáncer de próstata tratados con RP. Se utilizaron secciones de tumor de 10  $\mu\text{m}$  de espesor representativas para aislar el ARN. Cuando fue necesario, se utilizó una macrodissección o microdissección de la muestra guiada por el patólogo para enriquecer el tejido del tumor antes del aislamiento del ARN. Ninguna de las muestras en la cohorte de validación se microdisseccionó. Antes de cualquier análisis, la cohorte se dividió en 212  
40 pacientes para la caracterización inicial de la firma ("conjunto de entrenamiento") y 199 pacientes para su validación. Las características clínicas de la cohorte de entrenamiento y validación se enumeran en la Tabla 5.

Tabla 5

	Entrenamiento	Validación	valor de p valor	Estadística
Edad en años en RP, media (sd)	67,3 (5,9)	66,8 (5,8)	0,355	prueba t
Origen étnico (% de raza no blanca)	2,80%	7,60%	0,042	Exacta de Fisher
Procedimiento de disección (% 1 cm)	24%	0%	NA	NA
Recurrencia (%)	71/212 (33,5%)	72/199 (36,2%)	0,605	Exacta de Fisher
Días hasta la recurrencia, mediana	910	822	0,463	prueba t
Días hasta el seguimiento, mediana	3373	3387	0,173	prueba t
PSA previo a la cirugía (mediana)	7,3	6,8	0,163	Prueba t de log
Vesícula seminal	23/212 (10,8%)	28/199 (14,1%)	0,37	Exacta de Fisher
Vejiga	12/212 (5,7%)	17/199 (8,5%)	0,335	Exacta de Fisher
Ganglios linfáticos	8/212 (3,8%)	10/199 (5,0%)	0,632	Exacta de Fisher
Capsular	100/212 (47,2%)	101/199 (50,8%)	0,49	Exacta de Fisher
A través de la cápsula	59/212 (27,8%)	66/199 (33,2%)	0,283	Exacta de Fisher
Márgenes positivos	43/212 (20,3%)	57/199 (28,6%)	0,051	Exacta de Fisher
Puntuación de Gleason post-PR > 6	80/212 (37,7%)	66/199 (33,2%)	0,354	Exacta de Fisher
Nomograma post-PR, media (sp)	137 (19,5)	138 (23,0)	0,424	prueba t

Resultados

5 La firma de la expresión de GCC (Tabla 3, citada anteriormente) fue predictiva de recurrencia de la enfermedad en un análisis univariado (valor de  $p= 1,8 \times 10^{-9}$ , Tabla 6). La distribución de la puntuación de la firma estaba sesgada hacia valores más altos (expresión baja). La mediana del valor de la puntuación de la firma se usó para dividir la cohorte de entrenamiento en dos grupos que contienen muestras, ya sea con expresión alta o baja de GCC. La supervivencia frente al tiempo para ambos grupos se muestra en la figura 2.

10 También se evaluó la capacidad de predicción de la firma de GCC después de considerar las variables clínicas incluidas típicamente en un nomograma posquirúrgico (el nomograma de Kattan-Stephenson). El nomograma fue un predictor altamente significativo de recurrencia (valor de  $p 1,6 \times 10^{-10}$ ). Después de ajustar por el nomograma, la firma de GCC fue un predictor significativo de recurrencia bioquímica (figura 3) en la cohorte de descubrimiento (valor de  $p$  de 0,03) y en la cohorte de validación clínica (valor de  $p 4,8 \times 10^{-5}$ ).

Tabla 6

	N	Co-variables	Valor de p de la media de CCG*	Cociente de riesgos de la recurrencia
ENTRENAMIENTO (31 GCC)	212	ninguno	0,00404	1,24
	204	Nomograma posquirúrgico	0,03320	1,16
VALIDACIÓN (conjunto de 26 GCC)	199	ninguno	$1,8 \times 10^{-9}$	2,68
	197	Nomograma posquirúrgico	$4,8 \times 10^{-5}$	1,94

\* Media de la expresión de genes del ciclo celular con imputación de valores perdidos, prueba de razón de

probabilidades para el modelo de riesgos proporcionales de Cox.

Para ayudar a comprender la interacción entre el nomograma y la firma de expresión de GCC, se generó un gráfico de dispersión que compara estos predictores (figura 4) (las estrellas de color gris claro representan a los pacientes cuyo cáncer recurrió, mientras que las estrellas negras representan a los pacientes cuyo cáncer no lo hizo). El análisis del gráfico de dispersión mediante medias KM dividió las muestras en tres grupos basados únicamente en la puntuación del nomograma. Posteriormente, se descubrió que los grupos se basaban en parámetros clínicos bien entendidos. Los pacientes en el grupo de puntuación más baja (116/117) tenían enfermedad limitada al órgano. Los pacientes en el grupo de puntuación media (48/60) tenían al menos un parámetro posquirúrgico que se sabe que se asocia con un mal resultado (es decir, enfermedad a través de la cápsula, enfermedad con ganglios linfáticos positivos, y / o enfermedad con vesículas seminales positivas) y PSA prequirúrgico bajo (<10 ng / ml) Los pacientes en el grupo de puntuación más alta tenían al menos un parámetro posquirúrgico desfavorable y niveles altos de PSA prequirúrgico. A continuación, los pacientes en los grupos de puntuación baja y media se dividieron por la media de la puntuación de GCC. Los resultados para los pacientes en el grupo de mayor puntuación se predicen adecuadamente mediante el nomograma y, por lo tanto, no se dividieron adicionalmente. Como resultado, el gráfico de dispersión define cinco grupos de pacientes con tasas de recurrencia de la enfermedad del 2%, 40% (para dos grupos), 65% y 80% (Tabla 7). La tasa de recurrencia de los cinco grupos en función del tiempo se muestra en la figura 5.

Tabla 7

Nomograma post-PR			
Puntuación de GCC	Bajo	Medio	Alta
Bajo	1/62 (1,6%)	13/31 (41,9%)	16/20 (80%)
Alta	21/55 (38,2%)	19/29 (65,5%)	

El gráfico de dispersión mostrado en la figura 4 sugiere que hay una interacción no lineal entre la firma de GCC y el nomograma posquirúrgico. Es decir, la firma de GCC es un mejor predictor en pacientes con puntuaciones bajas en el nomograma. Por lo tanto, el estudio analizó la evidencia estadística de una interacción entre estas variables en un modelo multivariado para predecir la recurrencia de la enfermedad (Tabla 8). Se observó evidencia significativa de una interacción favorable tanto en los estudios de entrenamiento como de validación. Incluyendo el término de interacción en el modelo mejorado de manera espectacular, la importancia pronóstica de la firma de GCC después de ajustar para el nomograma (valores de p 0,0015 en la cohorte de entrenamiento y  $1,2 \times 10^{-8}$  en la cohorte de validación).

Tabla 8

Resumen estadístico						
Cohorte	N	Variables independientes	Covariables	Valor de p de la interacción	Valor de p de la variable	Cociente de riesgos de la recurrencia
Entrenamiento	204	Nomograma	ninguno	NA	$1,6 \times 10^{-10}$	
	212	Firma de GCC	ninguno	NA	0,004	1,24
	204	Firma de GCC	Nomograma	0,021	0,0015	
Validación	197	Nomograma	ninguno	NA	$7,7 \times 10^{-13}$	
	199	Firma de GCC	ninguno	NA	$1,8 \times 10^{-9}$	2,68
	197	Firma de GCC	Nomograma	0,0001	$1,28 \times 10^{-8}$	

### 30 EJEMPLO 3

El objetivo del siguiente objetivo fue determinar el número óptimo de GCC que se va a incluir en la firma. Como se ha mencionado anteriormente, los niveles de expresión de GCC se correlacionan entre sí de modo que era posible que la medición de un pequeño número de genes fuera suficiente para predecir el resultado de la enfermedad. De hecho, GCC individuales del conjunto de 31 genes en la Tabla 3 (panel C) suman considerablemente al nomograma de Kattan-Stephenson, como se muestra en la Tabla 9a continuación (después del ajuste por el nomograma y un término de interacción entre el nomograma y la expresión de GCC).

Tabla 9

Gen	Valor de p*	Gen	Valor de p*	Gen	Valor de p*
<i>NUSAP1</i>	$2,8 \times 10^{-7}$	<i>BUB1</i>	$8,3 \times 10^{-5}$	<i>KPNA2</i>	$2,0 \times 10^{-2}$
<i>DLG7</i>	$5,9 \times 10^{-7}$	<i>PBK</i>	$1,2 \times 10^{-4}$	<i>UBE2C</i>	$2,2 \times 10^{-2}$
<i>CDC2</i>	$6,0 \times 10^{-7}$	<i>TTK</i>	$3,2 \times 10^{-4}$	<i>MELK</i>	$2,5 \times 10^{-2}$
<i>FOXM1</i>	$1,1 \times 10^{-6}$	<i>CDC45L</i>	$7,7 \times 10^{-4}$	<i>CENPA</i>	$2,9 \times 10^{-2}$
<i>MYBL2</i>	$1,1 \times 10^{-6}$	<i>PRC1</i>	$1,2 \times 10^{-3}$	<i>CKS2</i>	$5,7 \times 10^{-2}$
<i>CDCA8</i>	$3,3 \times 10^{-6}$	<i>DTL</i>	$1,4 \times 10^{-3}$	<i>MAD2L1</i>	$1,7 \times 10^{-1}$
<i>CDC20</i>	$3,8 \times 10^{-6}$	<i>CCNB1</i>	$1,5 \times 10^{-3}$	<i>UBE2S</i>	$2,0 \times 10^{-1}$
<i>RRM2</i>	$7,2 \times 10^{-6}$	<i>TPX2</i>	$1,9 \times 10^{-3}$	<i>AURKA</i>	$4,8 \times 10^{-1}$
<i>PTTG1</i>	$1,8 \times 10^{-5}$	<i>ZWINT</i>	$9,3 \times 10^{-3}$	<i>TIMELESS</i>	$4,8 \times 10^{-1}$
<i>CCNB2</i>	$5,2 \times 10^{-5}$	<i>KIF23</i>	$1,1 \times 10^{-2}$		
<i>HMMR</i>	$5,2 \times 10^{-5}$	<i>TRIP13</i>	$1,7 \times 10^{-2}$		

\* valor de p para la prueba de razón de probabilidades del modelo de CoxPH completo (puntuación del nomograma post-RP + expresión del ciclo celular + nomograma: ciclo celular) frente a reducido (solo la puntuación del nomograma post-RP) del tiempo hasta la recurrencia.

- 5 Para evaluar cómo funcionaron los subconjuntos más pequeños del conjunto de GCC más grandes (es decir, paneles de GCC más pequeños), el estudio también comparó lo bien que la firma predijo el resultado como una función del número de GCC incluido en la firma (Fig. 1). El tiempo hasta la recurrencia después de la cirugía de próstata retrocedió en la media de GCC ajustada por la puntuación del nomograma post-PR. Los datos consisten en ensayos de TLDA expresados como deltaCT por 199 muestras de tumores de próstata FFPE y 26 GCC y se
- 10 analizaron mediante un modelo multivariado de CoxPH. Los valores de p son para la prueba de razón de probabilidades del modelo completo (nomograma + media del ciclo celular, incluyendo la interacción) frente al reducido (nomograma solamente). Como se muestra en la Tabla 10 a continuación y en la figura 1, las firmas de GCC pequeñas (por ejemplo, 2, 3, 4, 5, 6 GCCS, etc.) suman de manera significativa al nomograma de Kattan-Stephenson:

15

Tabla 10

N.º de GCC	Media del log10 (valor de p)*
1	-3,579
2	-4,279
3	-5,049
4	-5,473
5	-5,877
6	-6,228

\* Para 1000 subconjuntos extraídos al azar, tamaño de 1 a 6, de los genes del ciclo celular.

#### EJEMPLO 4

- 20 El objetivo de este experimento era evaluar la asociación entre las mutaciones de PTEN y la recurrencia bioquímica en pacientes con cáncer de próstata después de la prostatectomía radical. Se encontró que las mutaciones somáticas en PTEN estaban asociadas significativamente con la recurrencia y, de forma importante, añadieron información pronóstica más allá del nomograma clínico establecida para la recurrencia del cáncer de próstata (nomograma de Kattan-Stephenson) y la puntuación de la firma GCC (que se describe en los Ejemplos 1 y 2, citados anteriormente).

#### Pacientes y procedimientos

- 25 Se realizó un seguimiento de ochocientos cuatro pacientes RP consecutivos durante una media de 9,5 años. Las características de los pacientes y los resultados del tratamiento de toda la cohorte se han comunicado anteriormente (Swanson y col., *Oncol.* (2007) 25:110-114). Se disponía de bloques y/o cortes de tejido de la evaluación patológica final con suficiente tejido para el análisis de 430 pacientes. De estos, se seleccionó a 191 para la detección de mutaciones en PTEN en base a la cantidad de tumor disponible.
- 30 El ADN genómico se aisló de las muestras tumorales FFPE para la detección de mutaciones de PTEN utilizando el kit QIAamp Tissue DNA FFPE (Qiagen, Valencia, CA) de acuerdo con el protocolo del kit. Los cortes de FFPE se tiñeron primero con hematoxilina y eosina, y fueron analizados por un patólogo para identificar la región del tumor.

Después de la desparafinación, el tejido tumoral se cortó de los cortes con una hoja de afeitar. Para unas pocas muestras, se ayudó a la disección mediante microscopia de captura láser (LCM), debido a la dispersión de las células tumorales.

- 5 Las mutaciones se detectaron mediante el diseño de cebadores de secuenciación para interrogar a la secuencia genómica de PTEN. Los cebadores contenían colas M13 directas e inversas para facilitar la secuenciación. Después de la amplificación, se determinó la secuencia de ADN en una Mega BASE 4500 (GE healthcare) usando la química de colorante-cebador como se describe en Frank y col., J. Clin. Oncol. (2002) 20:1480-1490). Debido a las dificultades técnicas asociadas con la secuenciación de ADN procedente de material de FFPE, cada mutación se detectó mediante al menos dos reacciones de amplificación y secuenciación independientes.
- 10 Procedimientos estadísticos: A menos que se especifique lo contrario, la asociación entre la recurrencia bioquímica y las mutaciones de PTEN se evaluó utilizando modelos de PH de Cox para el tiempo hasta la recurrencia. Los valores de p resultantes se obtuvieron a partir de una prueba de razón de probabilidades comparando el modelo nulo con el modelo que contiene la variable del ensayo. En este ejemplo (Ejemplo 4), la firma de GCC derivó de 26 GCC (Panel D de la Tabla 2, citada anteriormente). Todos los datos de expresión se generaron por triplicado. Los datos de expresión se combinaron en una firma calculando el nivel medio de expresión de 26 GCC. Los datos clínicos fueron las variables incluidas en el nomograma de Kattan-Stephenson.
- 15

### Resultados

- 20 Las mutaciones de PTEN se encontraron en 13 individuos (13/191). En este subgrupo de 191 pacientes, PTEN fue un importante predictor de la recurrencia bioquímica (valor de  $p = 0,031$ ). La tasa de recurrencia en los portadores de la mutación fue del 69% (9/13) en comparación con el 36% (64/178) de los pacientes no mutantes. La diferencia en la tasa de recurrencia es también significativa mediante la prueba exacta de Fisher (valor de  $p = 0,034$ ). En el subgrupo de pacientes con datos de parámetros clínicos, la puntuación de la firma de GCC y las mutaciones de PTEN, el estado de PTEN fue un predictor significativo de la recurrencia bioquímica después de ajustar por ambos parámetros clínicos y la firma de GCC (valor de  $p$  de 0,024). Por último, la combinación de mutación de PTEN con la firma GCC fue un mejor predictor del resultado después de ajustar por parámetros clínicos que el uso de la firma de GCC después de ajustar por parámetros clínicos (valor de  $p = 0,0002$  para la combinación en comparación con 0,0028 para GCC solamente). Estos resultados muestran que las mutaciones de PTEN proporcionan información acerca de la probabilidad de recurrencia que no está correlacionada con ni los parámetros clínicos ni la firma de GCC y que el uso de los tres parámetros para evaluar el riesgo de recurrencia proporciona una estimación más precisa de la probabilidad de recurrencia de lo que era posible anteriormente.
- 25
- 30

### **EJEMPLO 5**

Este ejemplo describe estudios adicionales para validar y perfeccionar algunas formas de realización de las firmas de GCC de la invención.

### Pacientes y procedimientos

- 35 Se realizó un seguimiento de ochocientos cuatro pacientes de prostatectomía radical consecutivos durante una media de 9,5 años. La mediana de edad fue de 67 años. El estadio clínico fue T1 34%, T2 66% y T3 < 1%. La mediana del PSA preoperatorio fue de 6,6 ng/ml / ml con 72% <10 ng / ml y 28% > 10 ng / ml. Los especímenes se firmaron y los parámetros clínicos se registraron como positivos para margen uretral o del cuello de la vejiga, invasión de la cápsula, extensión a través de la cápsula, márgenes positivos y afectación de las vesículas seminales.
- 40 La recurrencia bioquímica se definió como un PSA > 0,3 ng/ml / ml. Para este estudio, los inventores tuvieron acceso a los datos clínicos de 690 pacientes. Se disponía de bloques y/o cortes de tejido de la evaluación patológica final con suficiente tejido para el análisis de 442 pacientes. La cohorte se dividió en 195 pacientes de una cohorte de entrenamiento y 247 pacientes para validación.

### Selección de genes

- 45 Se realizaron ensayos de 126 GCC y 47 genes HK (domésticos) frente a 96 muestras e FFPE de tumor de próstata anónimas obtenidas comercialmente, sin resultado ni otros datos clínicos. La hipótesis de trabajo era que los ensayos medirían con distintos grados de precisión del mismo fenómeno subyacente (proliferación del ciclo celular en el tumor para los GCC, y concentración de la muestra para los genes HK). Los ensayos se clasificaron mediante el coeficiente de correlación de Pearson entre el gen individual y la media de todos los genes candidatos, que es la mejor estimación disponible de la actividad biológica. Los resultados para la correlación de cada uno de los 126 GCC con la media se muestran en la Tabla 23. Sin incluir GCC con una expresión promedio baja o ensayos que produjeron fallos de la muestra, aproximadamente la mitad de los GCC tenían correlaciones de menos de 0,58, y una cuarta parte de los genes HK tenían correlaciones de menos de 0,95. Estos ensayos se interpretaron como que no reflejaban el fenómeno subyacente y se eliminaron, dejando un subconjunto de 56 GCC (Panel G) y 36 genes candidatos HK (Tablas 11 y 12). Los coeficientes de correlación se volvieron a calcular en este subconjunto y la selección final se hizo a partir de la lista clasificada.
- 50
- 55

Tabla 11: Lista completa de GCC evaluados ("Panel G")

Símbolo del gen	Correlación con la media de GCC	Símbolo del gen	Correlación con la media de GCC	Símbolo del gen	Correlación con la media de GCC
<i>FOXM1*</i>	0,908	<i>C18orf24*</i>	0,817	<i>FANCI</i>	0,702
<i>CDC20</i>	0,907	<i>RAD54L</i>	0,816	<i>KIF15</i>	0,701
<i>CDKN3</i>	0,9	<i>PTTG1</i>	0,814	<i>PLK4</i>	0,688
<i>CDC2</i>	0,899	<i>KIF4A</i>	0,814	<i>APOBEC3B</i>	0,67
<i>K1F11</i>	0,898	<i>CDCA3</i>	0,811	<i>NCAPG</i>	0,667
<i>KIAA0101</i>	0,89	<i>MCM10</i>	0,802	<i>TRIP13</i>	0,653
<i>NUSAP1</i>	0,887	<i>PRC1</i>	0,79	<i>K1F23</i>	0,652
<i>CENPF</i>	0,882	<i>DTL</i>	0,788	<i>NCAPH</i>	0,649
<i>ASPM</i>	0,879	<i>CEP55</i>	0,787	<i>TYMS</i>	0,648
<i>BUB1B</i>	0,879	<i>RAD51</i>	0,783	<i>GINS1</i>	0,639
<i>RRM2</i>	0,876	<i>CENPM</i>	0,781	<i>STMN1</i>	0,63
<i>DLGAP5</i>	0,875	<i>CDCA8</i>	0,774	<i>ZWINT</i>	0,621
<i>BIRC5</i>	0,864	<i>OIP5</i>	0,773	<i>BLM</i>	0,62
<i>KIF20A</i>	0,86	<i>SHCBP1</i>	0,762	<i>TTK</i>	0,62
<i>PLK1</i>	0,86	<i>ORC6L</i>	0,736	<i>CDC6</i>	0,619
<i>TOP2A</i>	0,851	<i>CCNB1</i>	0,727	<i>KIF2C</i>	0,596
<i>TK1</i>	0,837	<i>CHEK1</i>	0,723	<i>RAD51AP1</i>	0,567
<i>PBK</i>	0,831	<i>TACC3</i>	0,722	<i>NCAPG2</i>	0,535
<i>ASFIB</i>	0,827	<i>MCM4</i>	0,703		

Tabla 12: Lista de 15 genes domésticos (HK)

Símbolo del gen	Correlación con la media de HK
<i>RPL38</i>	0,989
<i>UBA52</i>	0,986
<i>PSMC1</i>	0,985
<i>RPL4</i>	0,984
<i>RPL37</i>	0,983
<i>RPS29</i>	0,983
<i>SLC25A3</i>	0,982
<i>CLTC</i>	0,981
<i>TXNL1</i>	0,98
<i>PSMA1</i>	0,98
<i>RPL8</i>	0,98
<i>MMADHC</i>	0,979
<i>RPL13A;</i> <i>LOC728658</i>	0,979
<i>PPP2CA</i>	0,978
<i>MRFAP1</i>	0,978

Expresión génica

- El ARN total se extrajo a partir de secciones de tumor FFPE de 5 µm representativas. Las muestras se desparafinizaron utilizando un baño de xileno y posteriormente se hidrataron en serie graduada de baños de etanol. Después, la región del tumor se diseccionó del corte usando una cuchilla de afeitar de acuerdo con las instrucciones

del patólogo. Alternativamente, la región del tumor se diseccionó directamente en un tubo Eppendorf y la parafina se retiró usando xileno y se lavó con etanol. Después, las muestras se trataron durante la noche con digestión con proteinasa K a 55 °C. El ARN total se extrajo utilizando RNeasy FFPE o miRNeasy (Qiagen) como describe el fabricante (siendo la única excepción la digestión extensa con proteinasa K descrita anteriormente). El ARN total aislado se trató con DNasa I (Sigma) antes de la síntesis de ADNc. Posteriormente, se empleó el kit High– capacity cDNA Archive (Applied Biosystems) para convertir el ARN total en ADNc de cadena sencilla como describe el fabricante. Se requirió un mínimo de 200 ng de ARN para la reacción de RT.

Antes de la medición de los niveles de expresión, el ADNc se preamplificó con una reacción combinada que contenía ensayos TaqMan™. Las condiciones de la reacción de preamplificación fueron: 14 ciclos de 95 °C durante 15 segundos y 60 °C durante 4 minutos. El primer ciclo fue modificado para incluir una incubación de 10 minutos a 95 °C. La reacción de amplificación se diluyó a 1:20 utilizando el tampón IX TE antes de la carga en las matrices de baja densidad TaqMan™ (TLDA, Applied Biosystems) para medir la expresión génica.

#### Puntuación de GCC

La puntuación de GCC se calcula a partir de la expresión del ARN de 31 GCC (Grupo F) normalizada para 15 genes domésticos (HK). Los números relativos de GCC (31) y de los genes domésticos (15) se optimizaron con el fin de reducir al mínimo la varianza de la puntuación de GCC. La puntuación de GCC es la media no ponderada de los valores de CT para la expresión de GCC, normalizados por la media no ponderada de los genes HK para que los valores más altos indiquen una mayor expresión. Una unidad es equivalente a un cambio de dos veces en la expresión. Los valores perdidos se imputaron utilizando la expresión media de cada gen determinado en el conjunto de entrenamiento utilizando solo muestras de buena calidad. Las puntuaciones de GCC se centraron en el valor medio, determinado de nuevo en el conjunto de entrenamiento.

Se realizó un experimento de dilución en cuatro de las muestras de próstata comerciales para estimar el error de medición de la puntuación de GCC (se = 0,10) y el efecto de los valores perdidos. Se encontró que la puntuación de GCC se mantenía estable a medida que la concentración disminuía hasta el punto de 5 fallos de un total de 31 GCC. Sobre la base de este resultado, a las muestras con más de 4 valores perdidos no se les asignó una puntuación de GCC.

El umbral de puntuación de GCC para la determinación del riesgo bajo se basó en la puntuación de GCC más bajo de recurrencias en el conjunto de entrenamiento. A continuación, el umbral se ajustó hacia abajo por 1 desviación estándar con el fin de optimizar el valor predictivo negativo de la prueba.

#### Modelo de riesgo clínico

Se usó un modelo de riesgos proporcionales de Cox para resumir los datos de los parámetros clínicos disponibles y estimar el riesgo clínico previo de recurrencia bioquímica para cada paciente. El conjunto de datos consistió en 195 casos del conjunto de entrenamiento y otros 248 casos con información de parámetros clínicos pero muestra insuficiente para medir la expresión de ARN. Se realizaron pruebas univariadas con los parámetros clínicos que se sabe que están asociados con el resultado (véase la Tabla 13 a continuación). Se excluyó del modelo a los parámetros no significativos. Se creó una variable compuesta para la enfermedad limitada al órgano, con la invasión definida como márgenes quirúrgicos, extensión extracapsular o afectación de cualquiera de las vesículas seminales, márgenes del cuello de la vejiga / uretra o ganglios linfáticos. La variable compuesta para la enfermedad limitada al órgano resultó más significativa en el modelo que cualquiera de sus cinco componentes, algunos de los cuales estaban correlacionados entre sí o no eran frecuentes. El ajuste del modelo se realizó utilizando los criterios AIC para covariables postoperatorias.

**Tabla 13: Análisis univariado de los parámetros clínicos y asociación con la recurrencia bioquímica**

Variable clínica	Valor de p	N.º de apariciones	Total	Frecuencia
VEJIGA URINARIA	0,0002	36	443	0,081
CAPSULAR	1,1x10 <sup>-9</sup>	194	443	0,438
ORIGEN ÉTNICO (BLANCOS)	0,6741	416	439	0,948
GANGLIOS LINFÁTICOS	0,0009	33	443	0,074
MARG.POS	6,1x10 <sup>-11</sup>	83	443	0,187
GRADO PATOGENICO	6,7x 10 <sup>-16</sup>	NA	443	NA
GRADO PATOLÓGICO	2,4x10 <sup>-11</sup>	NA	443	NA

ETAPA PATOLÓGICA	$3,1 \times 10^{-15}$	NA	443	NA
PRE.PSA.LOG10	$6,6 \times 10^{-12}$	NA	443	NA
VES. SEMINAL	$3,0 \times 10^{-8}$	56	443	0,126
CIRUGÍA. AÑO	0,0803	NA	443	NA
A TRAVÉS DE LA CÁPSULA	$1,3 \times 10^{-10}$	114	443	0,257
* Valor de p de Cox PH para la prueba de la razón de probabilidades				

5 El modelo final (es decir, nomograma) tiene variables binarias para la enfermedad limitada al órgano y una puntuación de Gleason de menos de o igual a 6, y una variable continua para el PSA logarítmico (Tabla 14). Este modelo incluye todos los parámetros clínicos incorporados en el nomograma post-RP (es decir, nomograma de Kattan -Stephenson), excepto para el Año de la RP y los dos componentes de la puntuación de Gleason. La distribución del riesgo clínico anterior muestra tres ganglios distintos (Figura 8). El agrupamiento por media K con 3 centros se utilizó para establecer el umbral para el grupo de bajo riesgo, que comprende aproximadamente el 50% de la muestra.

**Tabla 14: Modelo clínico**

Parámetro clínico	Coefficiente	HR	Valor p
Enfermedad limitada al órgano	-0,827	0,44	$3,4 \times 10^{-6}$
Puntuación de Gleason $\leq 6$	-0,8734	0,42	$4,6 \times 10^{-7}$
log PSA	0,6678	1,95	$2,4 \times 10^{-4}$
* Valor de p de Cox PH para la prueba de la razón de probabilidades			

10

#### Análisis estadístico

15 Los parámetros clínicos se compararon entre los conjuntos de entrenamiento y validación utilizando la prueba t de Student para variables continuas y la prueba exacta de Fisher para los parámetros categóricos. El riesgo clínico previo de los pacientes para recurrencia bioquímica después de la cirugía se estimó mediante una puntuación del nomograma post-PR que resume 7 covariables. La agrupación de la media K de la puntuación del nomograma se utilizó para clasificar a los pacientes como de riesgo clínico previo bajo o alto. Los datos de expresión se expresaron como la CT (el ciclo de PCR en el que la intensidad de la fluorescencia supera un umbral predeterminado) de cada GCC normalizado por la media de los 15 genes domésticos (Tabla 12 anterior).

20 Las muestras de mala calidad se excluyeron del análisis para eliminar las muestras de mala calidad o lecturas dudosas sin comprometer la integridad de la firma al excluir inadvertidamente muestras con baja expresión de GCC. En consecuencia, los umbrales para la limpieza o el filtrado de los datos se establecieron de forma conservadora. Los niveles de expresión medios de los genes HK para cada muestra, que eran más altos que los de los de GCC, se utilizaron para identificar muestras de mala calidad. Se utilizaron métricas técnicas para la eficiencia de amplificación y las desviaciones estándar excesivamente altas de los replicados se utilizaron para identificar las mediciones de CT no fiables. No hay fracasos de los genes HK y se permitió no más de 1 fallo de 3 repeticiones para los GCC.

25 La asociación entre la recurrencia bioquímica y la expresión DE GCC después de ajustar por riesgo clínico predicho por parámetros clínicos se evaluó utilizando un modelo de riesgos proporcionales de Cox para el tiempo hasta la recurrencia. La suposición de riesgos proporcionales sin dependencia del tiempo se analizó para el modelo completo de la firma de GCC más la puntuación clínica del parámetro clínico binario con un término de interacción y para la firma de GCC sólo en los subgrupos de riesgo clínico. No fue significativa ni en el grupo de entrenamiento ni en el de validación, lo que indica que no hay evidencia para la dependencia del tiempo. Todos los valores de p notificados se obtuvieron a partir de una prueba de razón de probabilidades comparando el modelo reducido o nulo con el modelo que contiene la variable del ensayo. Los gráficos Kaplan-Meier se utilizan para mostrar las probabilidades de supervivencia estimadas para subgrupos de pacientes; sin embargo, los valores de p son de la prueba de razón de probabilidades Cox para los valores continuos de la variable. Los análisis estadísticos se realizaron en S + Versión 30 8.1 para Linux 0.1 (TIBCO Spotfire) o R 2.9.0 (<http://www.r-project.org>).

#### Resultados

35 Se aisló ARN de secciones de tumor FFPE obtenidas de 442 pacientes de cáncer de próstata tratados con RP. La cohorte se dividió en 195 pacientes para la caracterización inicial de la firma ("conjunto de entrenamiento") y 247 pacientes para su validación. Los parámetros clínicos de la cohorte de entrenamiento y validación se enumeran en la

40

Tabla 15. No hubo diferencias significativas después de ajustar por múltiples comparaciones.

**Tabla 15: Parámetros clínicos de entrenamiento y validación en cohortes de paciente**

Parámetro clínico	Entrenamiento	Validación	Valor p	Análisis estadístico
Edad en años en la PR, media (sd)	67,5 (6,2)	66,8 (5,6)	0,204	prueba t
Origen étnico (% de raza no blanca)	3,10%	7,30%	0,058	Exacta de Fisher
	(2 negros, 3 hispanos, 1 otros)	(10 negros, 7 hispanos, 1 otros)		
Recurrencia	73/195 (37,4%)	90/247 (36,4%)	0,843	Exacta de Fisher
Días hasta la recurrencia, mediana	839	736	0,308	prueba t
Días hasta el seguimiento, mediana	3300	3332	0,556	prueba t
mediana de PSA precirugía de PR	7,4	6,4	0,022	Prueba t de log
Vesículas seminales	23/195 (11,8%)	33/247 (13,4%)	0,668	Exacta de Fisher
Margen del cuello de la vejiga/uretra	12/195 (6,2%)	16/247 (6,5%)	1	Exacta de Fisher
Ganglios linfáticos	8/195 (4,1%)	12/247 (4,9%)	0,819	Exacta de Fisher
Penetración capsular	104/195 (53,3%)	115/247 (46,6%)	0,18	Exacta de Fisher
A través de la cápsula	66/195 (33,8%)	73/247 (29,6%)	0,354	Exacta de Fisher
Márgenes positivos	51/195 (26,2%)	61/247 (24,7%)	0,742	Exacta de Fisher
Puntuación de Gleason post-PR > 7	114/195 (58,5%)	166/247 (67,2%)	0,06	Exacta de Fisher
Enfermedad limitada al órgano	108/195 (55,4%)	156/247 (63,2%)	0,118	Exacta de Fisher
SSP a 10 años (IC del 95%)	61% (52%, 69%)	67% (60 %, 73%)	0,905	Prueba del orden log

5 Para analizar la firma de GCC para este estudio, los inventores analizaron 126 GCC en ARN derivado de 96 tumores de próstata (Tabla 11). Las muestras tumorales fueron anónimas y no se asociaron con los datos clínicos. A partir de este conjunto de genes, se seleccionaron 31 genes (Panel F) para su inclusión en la firma (Tabla 16). Se seleccionaron los genes en función de su rendimiento técnico y por lo bien que cada gen se correlacionó con el nivel de expresión media de todo el conjunto de GCC en las 96 muestras anónimas.

**Tabla 16: Firma de GCC del conjunto de entrenamiento (Panel F)**

Símbolo	GeneID	Símbolo	GeneID	Símbolo	GeneID
<i>ASF1B</i>	55723	<i>CENPM</i>	79019	<i>ORC6L</i>	23594
<i>ASPM</i>	259266	<i>CEP55</i>	55165	<i>PBK</i>	55872
<i>BIRC5</i>	332	<i>DLGAP5</i>	9787	<i>PLK1</i>	5347
<i>BUB1B</i>	701	<i>DTL</i>	51514	<i>PRC1</i>	9055
<i>C18orf24</i>	220134	<i>FOXM1</i>	2305	<i>PTTG1</i>	9232

<i>CDC2</i>	983	<i>KIAA0101</i>	9768	<i>RAD51</i>	5888
<i>CDC20</i>	991	<i>KIF11</i>	3832	<i>RAD54L</i>	8438
<i>CDCA3</i>	83461	<i>KIF20A</i>	10112	<i>RRM2</i>	6241
<i>CDCA8</i>	55143	<i>MCM10</i>	55388	<i>TK1</i>	7083
<i>CDKN3</i>	1033	<i>NUSAP1</i>	51203	<i>TOP2A</i>	7153
<i>CENPF</i>	1063				

Para evaluar la utilidad pronóstica de la firma DE GCC, se generaron datos de expresión en 195 pacientes en el conjunto de entrenamiento. Dado que se correlacionaron los niveles de expresión de genes individuales, se combinaron en una puntuación de firma mediante el cálculo de la expresión media para todo el conjunto de 31 genes (Grupo F), normalizados por 15 domésticos (Tabla 12). La distribución de la puntuación de GCC se centró en cero y cada unidad de puntuación corresponde a un cambio de 2 veces en el nivel de expresión. Las muestras de mala calidad se identificaron mediante la observación de expresión baja de los genes domésticos o un número inaceptable de fallos de GCC, y se excluyeron del análisis. Después de aplicar las reglas de exclusión de los inventores, hubo 140 muestras disponibles para el análisis. La asociación entre la recurrencia bioquímica y la expresión de GCC se evaluó utilizando modelos de PH de Cox para el tiempo hasta la recurrencia. Un valor alto de expresión de GCC fue predictivo de recurrencia de la enfermedad en un análisis univariado (valor de  $p=0,01$ , Tabla 17).

A continuación, se evaluó la utilidad pronóstica de la firma de GCC después de considerar los parámetros clínicos conocidos por estar asociados con la recurrencia después de la PR. Para representar las medidas clínicas en el análisis de los inventores, se creó un modelo / nomograma que incluía el valor preoperatorio de PSA, la puntuación de Gleason y la evidencia de enfermedad fuera de la próstata (es decir, cualquiera de extensión extracapsular o patología posquirúrgica positiva en los ganglios linfáticos, los márgenes, el cuello de la vejiga, el margen uretral o las vesículas seminales). El modelo se ha optimizado en 443 pacientes (Tablas 13 y 14), incluyendo todos los pacientes de los que se disponía de datos clínicos, pero que no estaban en el conjunto de validación, y fue un predictor altamente significativo de recurrencia en la cohorte de entrenamiento (valor de  $2,5 \times 10^{-11}$ ). La distribución de las puntuaciones del modelo clínico contenía varios modos (Figura 8), separando de los grupos de pacientes de riesgo alto y bajo. Por lo tanto, la puntuación se utilizó posteriormente como una variable binaria (riesgo alto y bajo). El grupo de bajo riesgo se correlacionó un conjunto consistente de parámetros clínicos. En concreto, la gran mayoría (215/218) tenía enfermedad limitada al órgano y una puntuación de Gleason  $<7$ . Además, el 80% tenía niveles bajos de PSA prequirúrgico ( $<10$  ng / ml). Los pacientes en el grupo de riesgo alto (N = 225) fueron más heterogéneos, pero tendieron a tener características clínicas conocidas por estar asociadas con un mal resultado (por ejemplo, Gleason  $> 6$  y / o enfermedad a través de la cápsula).

El análisis multivariado de conjunto de entrenamiento que incorpora el modelo clínico binario de los inventores mostró evidencia de una interacción no lineal entre la firma de expresión y los parámetros clínicos (Tabla 17). Para ayudar a comprender la naturaleza de esta interacción, se generó un gráfico de dispersión comparando estos predictores (Figura 8). Como se desprende de la figura, se demostró que la puntuación de GCC era útil para evaluar el riesgo de recurrencia en pacientes definidos como de bajo riesgo mediante los parámetros clínicos. De hecho, incluso después de ajustar el modelo clínico en los pacientes de riesgo bajo, la firma de GCC es un fuerte predictor de recurrencia bioquímica (valor de  $p=0,0071$ ).

**Tabla 17: Resumen estadístico**

	Subpoblación basada en el modelo clínico modelo	31 genes entrenamiento N= 195			31 genes validación N= 247		
		Valor de p del efecto principal	Valor de p de la interacción	n	Valor de p del efecto principal	Valor de p de la interacción	n
Puntuación de GCC		0,01		140	$5,8 \times 10^{-8}$		218
riesgo clínico binario (bajo vs. alto)		$5,1 \times 10^{-6}$		133	$1,1 \times 10^{-10}$		215
Puntuación de GCC ajustada para riesgo binario riesgo + interacción		0,018	0,032	133	$8,3 \times 10^{-7}$	0,026	215

Puntuación de GCC solo	riesgo bajo	0,0038	-	54	$7,5 \times 10^{-5}$	112
Clin. Puntuación del riesgo	Riesgo bajo	0,22		54	0,044	112
Puntuación de GCC ajustada para riesgo clin. (riesgo clin. frente al riesgo clin + CCG)	Riesgo bajo	0,0071		54	0,00019	112
Puntuación de GCC	Riesgo alto	0,48		79	$5,8 \times 10^{-4}$	103
Clin. Puntuación del riesgo	Riesgo alto	$2,8 \times 10^{-6}$		79	0,0076	103
Puntuación de GCC ajustada para riesgo clin. (riesgo clin. frente al riesgo clin + CCG)	Riesgo alto	0,51		79	0,0026	103

Los inventores usaron sus datos de entrenamiento en el gráfico de dispersión para establecer una puntuación umbral optimizada de -0,16 para la firma de GCC (la puntuación media de GCC es cero). La Figura 12 muestra este umbral aplicado a los 443 pacientes estudiados en este ejemplo. El cuarenta por ciento de los pacientes de bajo riesgo cae por debajo de este umbral y se seleccionó de modo que no hubiera recurrencias 10 años después de la PR (es decir, valor predictivo negativo (VPN) del 100%). Como resultado de establecer los valores umbral, tanto para el modelo clínico como para la puntuación de GCC, el gráfico de dispersión se dividió en cuatro secciones con tasas de recurrencia del 0% (GCC bajo) y 26% (GCC alto) para los pacientes de bajo riesgo; y del 60% (GCC bajo) y del 50% para los pacientes de riesgo alto.

5 A continuación, se generaron datos de expresión de GCC en 247 pacientes la cohorte de validación de los inventores. Treinta y dos muestras se eliminaron del análisis adicional de acuerdo con las reglas de exclusión desarrolladas en la cohorte de entrenamiento. El panel F fue un predictor significativo de recurrencia bioquímica en un análisis univariado (valor de  $p = 5,8 \times 10^{-8}$ , Tabla 17). Después de ajustar para el modelo clínico binario, la firma de GCC fue altamente predictiva de recurrencia en la cohorte de validación (valor de  $p = 8,3 \times 10^{-7}$ ), y al igual que en el conjunto de entrenamiento, no hubo evidencia significativa de una interacción no lineal entre las variables. La firma GCC fue informativa en todo el espectro de riesgo clínicamente definido (Tabla 17). En cuanto a la validación de los resultados de entrenamiento, el valor de  $p$  para la asociación entre la recurrencia y la firma de GCC en los pacientes de riesgo bajo fue de  $1,9 \times 10^{-4}$ .

20 Los inventores aplicaron el umbral de GCC derivado de sus análisis de la cohorte de entrenamiento en su conjunto de datos de validación (Figura 9). Los pacientes de riesgo bajo con puntuaciones de GCC por debajo del umbral presentaron una tasa de recurrencia a los 10 años del 5% (equivalente al VPN validado de 0,95). En general, la combinación de puntuación de GCC y parámetros clínicos divide la cohorte en cuatro grupos con tasas de recurrencia a los 10 años predichas del 5%, 22%, 36% y 70% (Tabla 18). La tasa de recurrencia pronosticada frente a la puntuación de GCC para los pacientes en la cohorte de validación se muestra en las figuras 10 y 11.

25 **Tabla 18: Resumen de las tasas de recurrencia en la cohorte de validación definida por el riesgo clínico y la puntuación de GCC**

Riesgo clínico	Puntuación de GCC	Tasa de recurrencia a los 10 años	
		Estimación de Kaplan–Meier	n
baja	baja	0,05	39
baja	alta	0,22	73
alta	baja	0,36	27
alta	alta	0,7	76

30 Los inventores analizaron su umbral validado frente a diversas definiciones de pacientes de bajo riesgo (Tabla 19). La puntuación de la firma fue un indicador pronóstico significativo en diversas definiciones clínicas de bajo riesgo, y dependiendo de la definición, generó una tasa de recurrencia a los 10 años pronosticada de 0,05 a 0,10.

Tabla 19: VPN de la firma de GCC en otras definiciones de pacientes de bajo riesgo

Definición clínica de riesgo bajo	CCP baja*		Valor p
	Tasa de recurrencia a 10 años pronosticada**	n	
Enfermedad limitada al órgano y puntuación de Gleason <7 y PSA <10	0,05	39	9,4x10 <sup>-4</sup>
Enfermedad limitada al órgano y puntuación de Gleason <7	0,08	40	5,8x10 <sup>-3</sup>
Enfermedad limitada al órgano y puntuación de Gleason <8 y PSA <10	0,07	42	8,7x10 <sup>-4</sup>
Enfermedad limitada al órgano y puntuación de Gleason <8	0,1	43	4,1x10 <sup>-3</sup>
Enfermedad limitada al órgano solamente	0,1	44	2,4x10 <sup>-3</sup>
*– definido por el umbral validado			
**_Estimaciones de Kaplan–Meier			
***– para la diferencia entre las estimaciones de KM para el riesgo bajo y alto ajustado por la varianza de Greenwood.			

### Comentarios

5 Los inventores han desarrollado y validado una firma molecular pronóstica para el cáncer de próstata. La firma se basa en la medición de los niveles de expresión del ARNm de los genes del ciclo celular (GCC). Por definición, la expresión de los GCC está regulada como una función de la etapa del ciclo celular. Es decir, se activan en etapas específicas del ciclo celular, por lo que las células en crecimiento activo tienen mayores niveles de expresión de los GCC que las células quiescentes. Es de suponer que este hecho subyace a la capacidad de la firma para predecir la progresión del cáncer. Sin desear quedar ligado a teoría alguna, se piensa que mediante la medición de los niveles de expresión de GCC, los inventores están midiendo indirectamente la tasa de crecimiento y la agresividad inherente del tumor, lo que, en última instancia, afecta a la probabilidad de recurrencia del cáncer de próstata después de la prostatectomía.

15 Existe una distinción importante entre este estudio y muchos otros que han intentado generar firmas moleculares de pronóstico. A menudo, los estudios similares comienzan con un número muy grande de biomarcadores candidatos (a veces superiores a miles de genes) que se evalúan después para determinar la asociación con un fenotipo clínico de interés. Este enfoque puede, a veces, sufrir múltiples pruebas inherentes que pueden hacer que la importancia de la firma derivada sea incierta. En el presente documento se ha analizado una sola hipótesis: Los GCC serían pronósticos en el cáncer de próstata (de hecho, se seleccionaron los genes sobre la base de su correlación con la expresión de GCC, no basados en la asociación con la recurrencia). Y puesto que la expresión de GCC se correlaciona, se combinaron los datos de expresión en una firma predictiva mediante la determinación del valor de la expresión media de todos los genes en la firma. La simplicidad de este enfoque, biológica y computacionalmente, apoya la opinión de que la reivindicación central de este estudio es probable que sea muy sólido y se reproduce en estudios posteriores.

25 La firma de GCC (Panel F) es un factor predictivo independiente y aumenta significativamente la capacidad de predicción de los parámetros clínicos normalmente empleados para predecir la recurrencia de la enfermedad después de la cirugía. Esto es cierto tanto en las cohortes de los inventores tanto de entrenamiento como de validación.

30 La firma es de utilidad inmediata para definir el riesgo de los pacientes que presentan parámetros clínicos de bajo riesgo. En este sentido, esencialmente se definió el riesgo bajo como una puntuación de Gleason <7, PSA <10 y enfermedad limitada al órgano. La puntuación de la firma de GCC subdivide de forma eficaz el grupo de riesgo bajo en pacientes con tasas de recurrencia muy bajas (5%), y un mayor riesgo de recurrencia (22%) (Figura 9 y Tabla 18). Este es el efecto más espectacular de la firma molecular, que redefine con precisión el riesgo de los pacientes previamente definidos como de riesgo bajo en base a los parámetros clínicos. Es de destacar que en esta subpoblación de pacientes (es decir, los pacientes definidos como de riesgo bajo en base a parámetros clínicos), los parámetros clínicos no son particularmente pronósticos (véase la Tabla 17). Por tanto, como una prueba de diagnóstica, la firma podría ser útil para un gran número de pacientes. En este estudio, casi el 60% de la cohorte se caracterizó por ser de riesgo bajo y se espera que el 40% de de ellos tengan puntuaciones bajas de GCC. Por lo tanto, la firma de GCC puede predecir la enfermedad indolente en una cuarta parte de los pacientes que han sido previamente identificados como de riesgo alto (y, por lo tanto, identificados como candidatos para la prostatectomía

radical). Por último, los datos de validación, en particular, sugieren que la firma de GCC puede ser útil para definir el riesgo en todos los pacientes. En concreto, ayudó a dividir a los pacientes definidos como de riesgo alto de acuerdo con los parámetros clínicos en aquellos con tasas de recurrencia del 30% y del 70% (Tabla 18).

5 La combinación de los parámetros clínicos y firma de GCC permite a los médicos predecir con mayor precisión el riesgo de fracaso quirúrgico, y por lo tanto, identificar el curso apropiado de la intervención terapéutica. Como han demostrado los inventores, la firma mejora considerablemente la predicción de recurrencia para los pacientes que presentan parámetros clínicos generales de enfermedad no agresiva (Tabla 19). Dentro de este subgrupo clínico, los pacientes con puntuaciones bajas de GCC se beneficiarían de la certeza absoluta de que no está indicado ningún tratamiento adicional. Por el contrario, el grupo de GCC alto puede justificar la intervención inmediata. Los pacientes con parámetros clínicos posquirúrgicos desfavorables se benefician de la radioterapia adyuvante. Por lo tanto la firma de GCC debe predecir la eficacia de la radioterapia adyuvante para los pacientes con características clínicas de riesgo bajo y puntuaciones altas de GCC. En la cohorte de validación, los pacientes con puntuaciones altas de GCC y enfermedad más allá de la próstata tienen una tasa de recurrencia del 70%, que debería identificar claramente a los pacientes que son buenos candidatos para la radioterapia adyuvante. Así, la combinación de parámetros clínicos y firma de GCC conduce claramente a definir con mayor precisión el riesgo del paciente, lo que debería permitir una evaluación más inteligente de la necesidad de tratamiento adicional.

#### EJEMPLO 6

Algunos de los paneles de GCC descritos en el presente documento se evaluaron después según su capacidad para pronosticar cánceres adicionales. Se encontró que los paneles C, D, y F eran pronósticos en diversos grados en el cáncer de vejiga, de cerebro, de mama, y de pulmón.

#### Procedimientos

La expresión génica y los datos del paciente se obtuvieron a partir de los siguientes conjuntos de datos disponibles a consulta pública: GSE7390 (Desmedt y col., Clin. Cancer Res. (2007) 13:3207–14; PMID 17545524); GSE11121 (Schmidt y col., Cancer Res. (2008) 68:5405–13; PMID 18593943); GSE8894 (Son y col.; no publicación); Shedden (Shedden y col., Nature Med. (2008) 14:822; PMID 18641660); GSE4412 (Freije y col., cancer Res. (2004) 64:6503–10; PMID 15374961); GSE4271 (Phillips y col., CANCER CELL (2006) 9:157-73; PMID 16530701); GSE5287 (Als y col., CLIN. Cancer Res. (2007) 13:4407–14; PMID 17671123). Cada uno de estos conjuntos de datos tiene una descripción detallada asociada de los procedimientos experimentales utilizados en la recopilación de los datos del paciente y de expresión. Las micromatrices de expresión usados para generar cada conjunto de datos se resumen a continuación en la Tabla 20.

**Tabla 20**

Conjunto de datos	Matriz
GSE7390	Affymetrix U133 A
GSE11121	Affymetrix U133 A
GSE8894	Affymetrix U133 plus 2.0
Shedden	Affymetrix U133 A
GSE4412	Affymetrix U133 A y B
GSE4271	Affymetrix U133 A y B
GSE5287	Affymetrix U133 A

Los datos de expresión de cada uno de los genes en los paneles C, D y F se obtuvieron de estas bases de datos y el nivel de expresión medio para cada panel se determinó para cada paciente, cuyo resultado clínico era conocido (por ejemplo, recurrencia, progresión, supervivencia sin progresión, supervivencia global, etc.). La puntuación de GCC es una expresión media de los genes en un panel. Si un gen está representado por más de un conjunto de sondas en la matriz, la expresión génica es una expresión promedio de todos los conjuntos de sondas que representan el gen. La asociación entre la puntuación de GCC y la supervivencia o la recurrencia de la enfermedad se evaluó mediante el modelo de riesgos proporcionales de Cox univariado y multivariado. El análisis multivariado se realizó cuando los parámetros clínicos relevantes (grado en el cáncer de cerebro, estadio en el cáncer de pulmón, NPI en el cáncer de mama) estaban disponibles.

#### Resultados

Como se muestra en la Tabla 21 a continuación, cada panel, en el análisis univariado, fue un factor pronóstico en cada uno de los cánceres analizados.

Tabla 21

Tipo de cáncer	Conjunto de datos	Valor de p		
		Panel C	Panel F	Panel B
Cáncer de mama positivo para ER	GSE7390	2,4x10 <sup>-3</sup>	2,3x10 <sup>-3</sup>	4,3x10 <sup>-3</sup>
Cáncer de mama positivo para ER	GSE11121	1,6x10 <sup>-5</sup>	8,7x10 <sup>-6</sup>	1,5x10 <sup>-5</sup>
Adenocarcinoma de pulmón	GSE8894	2,4x10 <sup>-3</sup>	2,5x10 <sup>-3</sup>	5,6x10 <sup>-3</sup>
Adenocarcinoma de pulmón	Shedden	1,3x10 <sup>-7</sup>	2,6x10 <sup>-7</sup>	2,6x10 <sup>-7</sup>
Cáncer cerebral	GSE4412	3,6x10 <sup>-5</sup>	2,6x10 <sup>-5</sup>	9,4x10 <sup>-5</sup>
Cáncer cerebral	GSE4271	1,3x10 <sup>-3</sup>	1,4x10 <sup>-3</sup>	2,8x10 <sup>-4</sup>
Cáncer de vejiga	GSE5287	6,4x10 <sup>-2</sup>	5,4x10 <sup>-2</sup>	8,6x10 <sup>-2</sup>

Como se muestra en la Tabla 22 a continuación, cada panel fue también pronóstico en el análisis multivariado cuando se combinó con al menos un parámetro clínico (o nomograma).

5

Tabla 22

Tipo de cáncer	Conjunto de datos	Valor de p			Variable clínica adicional/nomograma
		Panel C	Panel F	Panel B	
Cáncer cerebral	GSE4271	0,022	0,017	0,0065	Grado
Adenocarcinoma de pulmón	Shedden	1x10 <sup>-6</sup>	2,1x10 <sup>-6</sup>	1,4x10 <sup>-6</sup>	estadio
Cáncer de mama positivo para ER	GSE7390	0,0077	0,0064	0,011	Índice pronóstico de Nottingham (NPI)
Cáncer de mama positivo para ER	GSE11121	0,0041	0,0027	0,0045	NPI

La Tabla 23 a continuación proporciona una clasificación de GCC seleccionados en función de su correlación con la expresión media de GCC.

Tabla 23

Símbolo del gen	Correl. p/Media	Símbolo del gen	Correl. p/Media	Símbolo del gen	Correl. p/Media
<i>TPX2</i>	0,931	<i>PBK</i>	0,805	<i>KIF18A</i>	0,6987
<i>CCNB2</i>	0,9287	<i>ESPL1</i>	0,805	<i>DONSON</i>	0,688
<i>KIF4A</i>	0,9163	<i>MKI67</i>	0,7993	<i>MCM4</i>	0,686
<i>KIF2C</i>	0,9147	<i>SPAG5</i>	0,7993	<i>RAD54B</i>	0,679
<i>BIRC5</i>	0,9077	<i>MCM10</i>	0,7963	<i>RNASEH2A</i>	0,6733
<i>BIRC5</i>	0,9077	<i>MCM6</i>	0,7957	<i>TUBA1C</i>	0,6697
<i>RACGAP1</i>	0,9073	<i>OIP5</i>	0,7943	<i>C18orf24</i>	0,6697
<i>CDC2</i>	0,906	<i>CDC45L</i>	0,7937	<i>SMC2</i>	0,6697
<i>PRC1</i>	0,9053	<i>KIF23</i>	0,7927	<i>CENPI</i>	0,6697
<i>DLGAP5/DLG7</i>	0,9033	<i>EZH2</i>	0,789	<i>GMPS</i>	0,6683
<i>CEP55</i>	0,903	<i>SPC25</i>	0,7887	<i>DDX39</i>	0,6673
<i>CCNB1</i>	0,9	<i>STIL</i>	0,7843	<i>POLE2</i>	0,6583
<i>TOP2A</i>	0,8967	<i>CENPN</i>	0,783	<i>APOBEC3B</i>	0,6513
<i>CDC20</i>	0,8953	<i>GTSE1</i>	0,7793	<i>RFC2</i>	0,648
<i>KIF20A</i>	0,8927	<i>RAD51</i>	0,779	<i>PSMA7</i>	0,6473
<i>BUB1B</i>	0,8927	<i>CDCA3</i>	0,7783	<i>MPHOSPH1/kif20b</i>	0,6457
<i>CDKN3</i>	0,8887	<i>TACC3</i>	0,778	<i>CDT1</i>	0,645
<i>NUSAP1</i>	0,8873	<i>PLK4</i>	0,7753	<i>H2AFX</i>	0,6387
<i>CCNA2</i>	0,8853	<i>ASF1B</i>	0,7733	<i>ORC6L</i>	0,634
<i>KIF11</i>	0,8723	<i>DTL</i>	0,769	<i>Clorf135</i>	0,6333
<i>CDCA8</i>	0,8713	<i>CHEK1</i>	0,7673	<i>PSRC1</i>	0,633
<i>NCAPG</i>	0,8707	<i>NCAPG2</i>	0,7667	<i>VRK1</i>	0,6323
<i>ASPM</i>	0,8703	<i>PLKI</i>	0,7657	<i>CKAP2</i>	0,6307
<i>FOXM1</i>	0,87	<i>TIMELESS</i>	0,762	<i>CCDC99</i>	0,6303

ES 2 595 410 T3

<i>NEK2</i>	0,869	<i>E2F8</i>	0,7587	<i>CCNE1</i>	0,6283
<i>ZWINT</i>	0,8683	<i>EXO1</i>	0,758	<i>LMNB2</i>	0,625
<i>PTTG1</i>	0,8647	<i>ECT2</i>	0,744	<i>GPSM2</i>	0,625
<i>RRM2</i>	0,8557	<i>STMN1</i>	0,737	<i>PAICS</i>	0,6243
<i>TTK</i>	0,8483	<i>STMN1</i>	0,737	<i>MCAM</i>	0,6227
<i>TRIP13</i>	0,841	<i>RFC4</i>	0,737	<i>DSN1</i>	0,622
<i>GINS1</i>	0,841	<i>CDC6</i>	0,7363	<i>NCAPD2</i>	0,6213
<i>CENPF</i>	0,8397	<i>CENPM</i>	0,7267	<i>RAD54L</i>	0,6213
<i>HMMR</i>	0,8367	<i>MYBL2</i>	0,725	<i>PDSS1</i>	0,6203
<i>NCAPH</i>	0,8353	<i>SHCBP1</i>	0,723	<i>HN1</i>	0,62
<i>NDC80</i>	0,8313	<i>ATAD2</i>	0,723	<i>C21orf45</i>	0,6193
<i>KIF15</i>	0,8307	<i>KIFCI</i>	0,7183	<i>CTSL2</i>	0,619
<i>CENPE</i>	0,8287	<i>DBF4</i>	0,718	<i>CTPS</i>	0,6183
<i>TYMS</i>	0,8283	<i>CKSIB</i>	0,712	<i>MCM7</i>	0,618
<i>KIAA0101</i>	0,8203	<i>PCNA</i>	0,7103	<i>ZWILCH</i>	0,618
<i>FANCI</i>	0,813	<i>FBX05</i>	0,7053	<i>RFC5</i>	0,6177
<i>RAD51AP1</i>	0,8107	<i>C12orf48</i>	0,7027		
<i>CKS2</i>	0,81	<i>TK1</i>	0,7017		
<i>MCM2</i>	0,8063	<i>BLM</i>	0,701		

La Tabla 1 a continuación proporciona una lista grande, aunque no exhaustiva, de los GCC.

**Tabla 1**

	<b>Gen (nombre y/o símbolo) o número (EST, clon de ADNc o acceso)</b>
1	Serina/treonina cinasa 15 Hs.48915 R11407
2	PLK: cinasa similar a polo (Drosophila) Hs.77597 AA629262
3	UBCH10: proteína portadora de ubiquitina E2-C Hs.93002 AA430504
4	MAPK13: Proteína cinasa activada por mitógeno 13 Hs.178695 AA157499 p38delta ARNm = Proteína cinasa activada por estrés 4
5	CDC2: ciclo de división celular 2, G1 a S y G2 a M Hs.184572 AA598974
6	TOP2A: topoisomerasa (ADN) II alfa (170 kD) Hs.156346 AA504348
7	CENPE: Proteína del centrómero E (312 kD) Hs.75573 AA402431 CENP-E = motor del cinetocoro putativo que se acumula justo antes
8	TOP2A: topoisomerasa (ADN) II alfa (170 kD) Hs. 156346 AA026682
9	KPNA2: carioferina alfa 2 (cohorte RAG 1, importina alfa 1) Hs.159557 AA676460
10	FLJ10468: proteína hipotética FLJ10468 Hs.48855 N63744
11	CCNF: ciclina F Hs.1973 AA676797
12	DKFZp762E1312: proteína hipotética DKFZp762E1312 Hs.104859 T66935
13	CKS2: Proteína asociada con CDC2 CKS2 Hs.83758 AA292964
14	C20ORF1: marco de lectura abierto 1 del cromosoma 20 Hs.9329 H73329
15	BUB1: gemación no inhibida por bencimidazoles 1 (homólogo de levadura) Hs.98658 AA430092 BUB1 = proteína ser/thr cinasa del punto de verificación mitótico putativo
16	TOP2A: **topoisomerasa (ADN) II alfa (170 kD) Hs.156346 AI734240
17	CKS2: Proteína asociada con CDC2 CKS1 Hs.83758 AA010065 ckshs2 = homólogo de Cks1 = proteína asociada a p34Cdc28/Cdc2
18	ARL6IP: Proteína de interacción 6 similar al factor de ribosilación de ADP Hs.75249 H20558
19	L2DTL: Proteína L2DTL Hs.126774 R06900
20	STK15: **serina/treonina cinasa 15 Hs.48915 H63492 cinasa relacionada con aurora/IPL1
21	E2-EPF: Proteína portadora de ubiquitina Hs.174070 AA464019
22	UBCH10: proteína portadora de ubiquitina E2-C Hs.93002 AA430504
23	KNSL5: similar a cinesina 5 (proteína 1 similar a cinesina mitótica) Hs.270845 AA452513 proteína 1 similar a cinesina mitótica
24	CENPF: proteína F del centrómero (350/400 kD, mitosis) Hs.77204 AA701455

ES 2 595 410 T3

25	CCNA2: ciclina A2 Hs.85137 AA608568 Ciclina A
26	CDC2: ciclo 2 de la división celular, G1 a S y G2 a M Hs.184572 AA278152 CDC2 = homólogo de proteína 2 de control de la división celular = proteína kin P34
27	HMMR: **Receptor de la motilidad mediada por hialuronano (RHAMM) Hs.72550 AA171715
28	KIAA0008: Producto génico de KIAA0008 Hs.77695 AA262211
29	HSPC145: Proteína HSPC145 Hs.18349 R22949
30	FLJ20510: proteína hipotética FLJ20510 Hs.6844 N53214
31	ARNm de Homo sapiens NUF2R, cds completa Hs.234545 AA421171:
32	FLJ20510: proteína hipotética Hs.13525 T87341
33	P37NB: Proteína de repetición (LRR) rica en leucina de 37 kDa Hs.155545 AA423870
34	CDC20:
35	CCNE1: ciclina E1 Hs.9700 T54121
36	ESTs: Hs.221754 R84407
37	FLJ11252: proteína hipotética FLJ11252 Hs.23495 N30185
38	LOC51203: clon HQ0310 PRO0310p1 Hs.279905 AA620485
39	LOC51203: proteína hipotética FLJ10491 Hs.274283 AA425404
40	KNSL1-: similar a cinesina 1 Hs.8878 AA504625
41	CENPA: proteína del centrómero A (17 kD) Hs.1594 AI369629
42	Homo sapiens, clon IMAGE: 2823731, ARNm, cds parcial Hs.70704 R96941:
43	CDC6: homólogo de CDC6 (ciclo de división celular 6, S. cerevisiae) Hs.69563 H59203
44	ARNm del homólogo de la ADN helicasa de Homo sapiens (PIF1), cds parcial Hs.112160 AA464521:
45	ESTs: Hs.48480 AA135809
46	TSN: translina Hs.75066 AA460927
47	KPNA2: carioferina alfa 2 (cohorte RAG 1, importina alfa 1) Hs.159557 AA676460
48	RRM2: polipéptido M2 de ribonucleótido reductasa Hs.75319 AA187351
49	ESTs: Hs.14119 AA204830
50	CCNB1: ciclina B1 Hs.23960 R25788
51	GTSE1: Expresada en la fase G-2 y S 1 Hs.122552 AI369284
52	C20ORF1: marco de lectura abierto 1 del cromosoma 20 Hs.9329 AA936183
53	TACC3: proteína 3 ácida transformante que contiene hélice superenrollada Hs.104019 AA279990 JkR1 ARNm regulado a la baja tras la activación de células T
54	E2F1: Factor 1 de transcripción E2F Hs.96055 H61303
55	BUB1B: gemación no inhibida por bencimidazoles 1 (homólogo de levaduras), Hs.36708 AA488324
56	ESTs.; Débilmente similar a CGHU7L precursor de la cadena del colágeno alfa 1(III) [H. sapiens] Hs.19322 AA088457
57	KIAA0074: Proteína KIAA0074 Hs.1192 N54344
58	MPHOSPH1: Fosfoproteína 1 de la fase M Hs.240 AA282935
59	ANLN: anilina (homólogo de Drosophila Scraps), proteína de unión a la actina Hs.62180 R12261
60	BIRC5: survivina 5 que contiene repetición IAP baculoviral Hs.1578 AA460685 Survivina = inhibidor de la apoptosis = proteasa de células efectoras EPR-1
61	PTTG1: transformante 1 del tumor hipofisario Hs.252587 AA430032
62	KIAA0159: proteína 1 asociada con SMC relacionada con la condensación de los cromosomas Hs.5719 AA668256
63	ESTs.; Débilmente similar a la proteína OS-4 [H. sapiens] Hs.18714 W93120
64	HMMR: receptor de la motilidad mediada por hialuronano (RHAMM) Hs.72550 R10284
65	DKFZp762E1312: proteína hipotética DKFZp762E1312 Hs.104859 AA936181
66	CKAP2: proteína 2 asociada al citoesqueleto Hs.24641 T52152
67	RAMP: Proteína asociada a la matriz nuclear regulada por RA

ES 2 595 410 T3

68	SMAP: proteína coactivadora del receptor de hormonas tiroideas Hs.5464 AA481555
69	FLJ22624: proteína hipotética FLJ22624 Hs. 166425 AA488791
70	CKS1: Proteína asociada con CDC2 CKS1 Hs.77550 N48162
71	NEK2: cinasa 2 relacionada con NIMA (nunca en el gen a de la mitosis) Hs.153704 W93379
72	MK167: antígeno identificado por el anticuerpo monoclonal Ki-67
73	TTK: Proteína cinasa TTK Hs.169840 AI337292
74	VEGFC: factor de crecimiento endotelial vascular C Hs.79141 H07899 proteína relacionada con el factor de crecimiento endotelial vascular VRP
75	CDKN3: Inhibidor 3 de la cinasa dependiente de ciclina (fosfatasa de especificad doble asociada con CDK2) Hs 84113 AA284072 CIP2 = Cdi1 = fosfatasa KAP1= G1/S gen del ciclo celular
76	ARNm de NUF2R de Homo sapiens, cds completa Hs.234545 R92435:
77	ADNc de Homo sapiens FLJ10325 fis, clon NT2RM2000569 Hs.245342 AA235662:
78	HSPC145: Proteína HSPC145 Hs.18349 AA628867
79	HSU54999: Proteína LGN Hs.278338 W92010
80	FLJ20333: proteína hipotética FLJ20333 Hs.79828 R27552
81	KNSL2: 2 similar a cinesina Hs.20830 N69491
82	ESTs: Hs.133294 AI053446
83	**ESTs: Hs.41294 H95819
84	SMTN: esmotelina Hs.149098 AA449234
85	FLJ23311: proteína hipotética FLJ23311 Hs.94292 N73916
86	FLJ23311: factor de transcripción cadena arriba 1 Hs.247842 AA719022
87	LOC51203: clon HQ0310 PRO0310p1 Hs.279905 AA779949
88	ADH4: alcohol deshidrogenasa 4 (clase II), polipéptido pi Hs.1219 AA007395
89	ESTs: Hs.186579 AA960844
90	CCNB2: ciclina B2 Hs.194698 AA774665
91	Homo sapiens, Similar al clúster rico en genes, gen C8, clon MGC: 2577, ARNm, cds completa Hs.30114 AA634371:
92	ESTs: Hs.99480 AA485454
93	ARNm de Homo sapiens IRE1b para proteína cinasa/ribonucleasa IRE1 beta, cds completa Hs.114905 AA088442:
94	PCNA: antígeno nuclear de células en proliferación Hs.78996 AA450264 PCNA = antígeno nuclear de células en proliferación
95	AA075920:
96	GTSE1: Expresada en la fase G-2 y S 1 Hs.122552 AA449474
97	CKS1: Proteína asociada con CDC2 CKS1 Hs.77550 AA278629
98	CDC25B: ciclo de división celular 25B Hs.153752 AA448659 cdc25B = fosfatasa 2 inductora de la fase M
99	ESTs.; Débilmente similar al producto de proteína sin nombre [H. sapiens] Hs.99807 AA489023 desconocida UG Hs.99807 ESTs sc_id384
100	PCNA: antígeno nuclear de células en proliferación Hs.78996 H05891
101	LTBP3: **proteína 3 de unión al factor de crecimiento beta transformante latente Hs.289019 R60197
102	ARNm de Homo sapiens; ADNc DKFZp434D0818 (del clon DKFZp434D0818) Hs.5855 N95578:
103	ESTs: Hs.126714 AA919126
104	CIT: citron (serina/treonina cinasa 21 de interacción con rho) Hs.15767 H10788
105	LBR: receptor de lamina B Hs.152931 AA099136
106	E2F1: Factor 1 de transcripción E2F Hs.96055 AA424949
107	AA699928:
108	CDKN2C: inhibidor 2C de cinasa dependiente de ciclina (p18, inhibe CDK4) Hs.4854 N72115 p18-INK6 = Inhibidor 6 de cinasa dependiente de ciclina
109	STK12: serina/treonina cinasa 12 Hs.180655 H81023 ARK2 = cinasa 2 relacionada con aurora

ES 2 595 410 T3

110	ESTs: Hs.111471 AA682533
111	ESTs: Hs.44269 AA465090
112	MCM4: Deficiente en mantenimiento de minicromosoma ( <i>S. cerevisiae</i> ) 4 Hs.154443 AA485983
113	PMSCL1: **autoantígeno 1 de polimiositis/esclerodermia (75 kD) Hs.91728 AA458994 Ciclina A
114	MKI67: antígeno identificado por el anticuerpo monoclonal Hs.80976 AA425973 Ki67 (de tipo largo)
115	ESTs: Hs. 133294 AI144063
116	CDC25B: ciclo de división celular 25B Hs.153752 AA448659 cdc25B = fosfatasa 2 inductora de la fase M
117	FOXN1: caja en horquilla M1 Hs.239 AA129552 MPP2 = fosfoproteína 2 de fase M putativa
118	FLJ11029: proteína hipotética FLJ11029 Hs.274448 AI124082
119	H2AFX: Familia de la histona H2A, miembro X 147097 H95392
120	FLJ20333: proteína hipotética FLJ20333 Hs.79828 AA147792
121	SLC17A2: familia del portador de solutos 17 (fosfato sódico), miembro 2 Hs.19710 H60423
122	ARNm de Homo sapiens IRE1b para proteína cinasa/ribonucleasa IRE1 beta, cds completa Hs.114905 AA102368:
123	ESTs: Hs.163921 AA573689
124	MCM5: Deficiente en mantenimiento de minicromosoma ( <i>S. cerevisiae</i> ) 5 (ciclo de división celular 46) Hs.77171 AA283961
125	CDKN1B: inhibidor de cinasa dependiente de ciclina 1B (p27, Kip1) Hs.238990 AA630082
126	AA779865:
127	PTTG1: transformante 1 del tumor hipofisario Hs.252587 AI362866
128	RAD21: homólogo de RAD21 ( <i>S. pombe</i> ) Hs.81848 AA683102
129	ADNc de Homo sapiens FLJ10325 fis, clon NT2RM2000569 Hs.245342 AA430511:
130	NEK2: Cinasa 2 relacionada con NIMA (nunca en el gen a de la mitosis) Hs.153704 AA682321
131	FLJ20101: proteína NUDE1 de interacción con LIS1, homólogo de rata Hs.263925 N79612
132	FZR1: proteína Fzr1 Hs.268384 AA621026
133	ESTs: Hs.120605 AI220472
134	KIAA0855: golgina-67 Hs.182982 AA098902
135	SRD5A1: esteroide-5-alfa-reductasa, polipéptido alfa 1 (3-oxo-5 alfa-esteroide delta 4-deshidrogenasa alfa 1) Hs.552 H16833
136	RAD51.: homólogo de RAD51 ( <i>S. cerevisiae</i> ) (homólogo de RecA de <i>E coli</i> ) Hs.23044 N70010
137	KNL2: similar a cinesina 2 Hs.20830 R11542
138	KIAA0097: Producto génico de KIAA0097 Hs.76989 AA598942
139	TUBB: tubulina, polipéptido beta Hs.179661 AA427899
140	HEC: altamente expresada en cáncer, repeticiones de héptadas ricas en leucina Hs.58169 W72679
141	TROAP: proteína asociada con trofinina (tastina) Hs.171955 H94949
142	ESTs: Hs.49047 N64737
143	ESTs: Hs.15091 AA678348
144	ESTs: Hs.133431 AI061169
145	KIAA0042: Producto génico de KIAA0042 Hs.3104 AA477501
146	FZR1: proteína Fzr1 Hs.268384 AA862886
147	FEN1: endonucleasa 1 específica de estructura flap Hs.4756 AA620553
148	CKS1: Proteína asociada con CDC2 CKS1 Hs.77550 AA459292 cks1 = homólogo de Cks1 = proteína asociada a p34Cdc28/Cdc2
149	ESTs: Hs.193379 N57936
150	CASP8AP2: Proteína 2 asociada con CASP8 Hs. 122843 H50582
151	BIRC2: 2 que contiene repetición IAP baculoviral Hs.289107 R19628 c-IAP1 = MIHB = homólogo B de IAP
152	CKAP2: proteína 2 asociada al citoesqueleto Hs.24641 AA504130

ES 2 595 410 T3

153	HLA-DRA: complejo mayor de histocompatibilidad, clase II, DR alfa Hs.76807 R47979
154	HBP: Proteína de unión en horquilla, histona Hs.75257 AA629558
155	FLJ10483: proteína hipotética FLJ10483 Hs.6877 H12254
156	CASP3: caspasa 3, cisteína proteasa relacionada con la apoptosis Hs.74552 R14760 CASPASE- 3 = CPP32 isoforma alfa= yama = cisteína proteasa
157	**ESTs,: débilmente similar a la proteína que es inmunorreactiva con anticuerpos policlonales anti-PTH [H. sapiens] Hs.301486 AA088258
158	HMG2: proteína 2 del grupo de alta movilidad (histona no cromosómica) Hs.80684 AA019203
159	PRO2000: Proteína PRO2000 Hs.46677 H58234
160	FLJ20333: proteína hipotética FLJ20333 Hs.79828 T48760
161	T56726:
162	TIMP1: inhibidor tisular de la metaloproteínasa 1 (actividad potenciadora de eritroide, inhibidor de colagenasa) Hs.5831 H80214
163	ESTs: Hs.102004 R94281
164	FLJ10858: proteína hipotética FLJ10858 Hs.134403 AA677552
165	ADNc de Homo sapiens FLJ11883 fis, clon HEMBA1007178 Hs.157148 N62451:
166	RFC4: factor de replicación C (activador 1) 4 (37 kD) Hs.35120 N93924 factor de replicación C
167	PRO2000: Proteína PRO2000 Hs.46677 N47113
168	ECT2: oncogén de secuencia 2 transformante de células epiteliales Hs.132808 AI031571
169	ESTs: Hs.165909 AA629538
170	PCF11: Homólogo de PCF11p Hs.123654 AA053411
171	BIRC3: que contiene repeticiones IAP baculoviral 3 Hs.127799 H48533 c-IAP2 = MIHC = homólogo C de IAP = proteína del complejo de señalización TNFR2-TRAF
172	EST,: débilmente similar a dJ45P21,2 [H. sapiens] Hs.326451 AA931528
173	KIAA0952: Proteína KIAA0952 Hs.7935 AA454989
174	KIF5B: miembro de la familia de cinesinas 5B Hs. 149436 AA608707
175	DKFZP566C134: Proteína DKFZP566C134 Hs.20237 N39306
176	ANLN: anilina (homólogo de Drosophila Scraps), proteína de unión a la actina Hs.62180 R17092
177	ORC1L: complejo de reconocimiento de origen, similar a subunidad 1 (homólogo de levadura) Hs.17908 H51719
178	ESTs: Hs.14139 T77757
179	IFIT1: proteína inducida por interferón con repeticiones tetratricopéptido 1 Hs.20315 AA074989
180	MGC5338: proteína hipotética MGC5338 Hs.99598 AA463627
181	COPEB: proteína de unión al elemento promotor del núcleo Hs.285313 AA013481
182	UK114: proteína inhibidora de la traducción p14,5 Hs.18426 N72715
183	ESTs: Hs.265592 H67282
184	HMG4: proteína 4 del grupo de alta movilidad (histona no cromosómica) Hs.19114 AA670197
185	MDS025: proteína hipotética MDS025 Hs. 154938 AI225067
186	DKFZP564A122: Proteína DKFZP564A122 Hs.187991 N53236
187	TSC22: proteína estimulada por factor de crecimiento transformante beta TSC-22 Hs.114360 AA664389
188	AAAS: aladina Hs.125262 AA916726
189	PLAG1: **gen de adenoma pleomórfico 1 Hs.14968 AA418251
190	FLJ23293: **proteína hipotética FLJ23293 similar a la proteína 2 de interacción con ARL-6 Hs.31236 R91583
191	H11: proteína quinasa H11; proteína similar a proteína pequeña del estrés HSP22 Hs.111676 AA010110
192	POLD3: polimerasa (dirigida por ADN), delta 3 Hs.82502 AA504204
193	SERPINB3: inhibidor de la serina (o cisteína) proteínasa, clado B (ovoalbúmina), miembro 3 Hs.227948 AA292860
194	DNAJB1: homólogo de DnaJ (Hsp40), subfamilia B, miembro 1 Hs.82646 AA435948

ES 2 595 410 T3

195	ESTs: Hs.99480 AA458886
196	BUB3: BUB3 (gemación no inhibida por bencimidazoles 3, levadura) homólogo Hs.40323 AA405690
197	TUBB2: tubulina, beta, 2 Hs.251653 AI000256
198	ARNm de Proteína SNC73 de Homo sapiens (SNC73) ARNm, cds completa Hs.293441 H28469:
199	BUB3: homólogo de BUB3 (gemación no inhibida por por bencimidazoles 3, levadura) Hs.40323 H38804
200	FLJ20699: proteína hipotética FLJ20699 Hs.15125 AA459420
201	KIAA0013: Producto génico KIAA0013 Hs.172652 N63575
202	ESTs: Hs.20575 N20305
203	CDC25C: ciclo de división celular 25C Hs.656 W95000 cdc25C = fosfatasa 3 inductora de la fase M
204	ELJ11186: proteína hipotética FLJ11186 Hs.89278 AA394225
205	TOPK: Cinasa de unión a PDZ; proteína cinasa originada en células T Hs.104741 AA448898
206	KIAA0165: polos del huso adicionales, S. cerevisiae, homólogo de Hs.153479 AA948058
207	LOC51659: Proteína HSPC037 Hs.108196 AA961752
208	ESTs: Hs.10338 AA436456
209	SUCLG2: succinato-CoA ligasa, formador de ADP, subunidad beta Hs.247309 AA465233
210	ZNF265: proteína de dedo de cinc 265 Hs.194718 AA452256
211	SKP2: proteína 2 asociada con la cinasa de fase S (p45) Hs.23348 R22188
212	NS1-BP: proteína de unión a NS1 Hs.197298 AA486796
213	C21ORF50: marco de lectura abierto 50 del cromosoma 21 Hs.4055 AA416628
214	BIRC2: 2 que contiene repetición de IAP baculoviral Hs.289107 AA702174
215	BIRC3: que contiene repeticiones de IAP baculoviral Hs.127799 H48533 c-IAP2 = MIHC = homólogo C de IAP = proteína del complejo de señalización TNFR2-TRAF
216	INDO: indoleamina-pirrol 2,3 dioxigenasa Hs.840 AA478279
217	DEEPEST: proteína relacionada con hélice superenrollada del huso mitótico Hs.16244 T97349
218	ESTs: Hs.105826 AA534321
219	C20ORF1: marco de lectura abierto 1 del cromosoma 20 Hs.9329 AI654707
220	ADNc de Homo sapiens: FLJ21869 fis, clon HEP02442 Hs.28465 R63929:
221	RGS3: regulador de la señalización de la proteína G 3 Hs.82294 AI369623
222	ARNm de DC29 de Homo sapiens, cds completa Hs.85573 AA186460:
223	MCM6: deficiente en mantenimiento de minicromosoma (mis5, S. pombe) 6 Hs.155462 AA663995
224	NPAT: proteína nuclear, locus de ataxia-telangiectasia Hs.89385 AA284172 NPAT = E14 = gen en el locus ATM
225	KNSL6: similar a cinesina 6 (cinesina asociada al centrómero mitótico) Hs.69360 AA400450
226	HN1: expresado hematológica y neurológicamente 1 Hs.109706 AA459865
227	TUB A3: Tubulina, alfa, específica de cerebro Hs.272897 AA865469
228	ESTs: Hs.221197 N55457
229	KIAA0175: Producto génico KIAA0175 Hs.184339 AA903137
230	CLASPIN: homólogo de Xenopus Claspin Hs.175613 AA857804
231	CTNNA1: **catenina (proteína asociada a cadherina), alfa1 (102 kD) Hs.178452 AA026631
232	ESTs: Hs.221962 AA229644
233	SMC4L1: SMC4 de tipo (mantenimiento estructural de los cromosomas 4, levadura) 1 Hs.50758 AA452095
234	ICBP90: factor de transcripción Hs.108106 AA026356
235	EXO1: exonucleasa 1 Hs.47504 AA703000
236	ARNm de factor 1 asociado con TRAF4 de Homo sapiens, cds parcial Hs. 181466 T84975:
237	ESTs: Hs.186814 AA700879
238	FLJ11269: proteína hipotética FLJ11269 Hs.25245 R37817

ES 2 595 410 T3

239	SFPQ: factor de corte y empalme rico en prolina / glutamina (asociado a la proteína de unión al tracto de polipirimidina) Hs.180610 AA425258
240	ZF: factor de transcripción de unión a HCF Zhangfei Hs.29417 AA164474
241	TUBA2:tubulina, alfa 2 Hs.98102 AA626698
242	ARNm de Homo sapiens; ADNc DKFZp434M0435 (del clon DKFZp434D0818) DKFZp434M0435) Hs.25700 N94435:
243	FLJ20530: **proteína hipotética FLJ20530 Hs.279521 AA425442
244	BTEB1: proteína 1 de unión al elemento básico de la transcripción Hs.150557 N80235
245	BTEB1: geminina Hs.234896 H51100
246	D21S2056E: secuencia expresada del segmento de ADN en el cromosoma 21 (único) 2056 Hs.110757 AI362799
247	HDAC3: histona desacetilasa 3 Hs.279789 H88540
248	USP1: proteasa específica de ubiquitina 1 Hs.35086 AA099033
249	C21ORF50: marco de lectura abierto 50 del cromosoma 21 Hs.4055 AA135912
250	FLJ13046: **proteína hipotética FLJ13046 similar a la exportina 4 Hs.117102 T95333
251	ESTs: Hs.181059 AA912032
252	FLJ22009: proteína hipotética FLJ22009 Hs.123253 AA401234
253	ESTs: Hs.62711 AA056377
254	RAD51C: homólogo de RAD21 (S. cerevisiae) C Hs.11393 R37145 RAD51C = proteína relacionada con Rad51 de recombinación / reparación
255	ESTs: Hs.268919 H53508
256	ADNc de Homo sapiens FLJ11381 fis, clon HEMBA1000501 Hs.127797 AA885096:
257	SAP30: Polipéptido asociado con sin3, 30 kD Hs.20985 AA126982
258 -	H4FG: familia de histonas H4 , miembro G Hs.46423 AA868008
259	TUBA1: tubulina, alfa 1 (específica de los testículos) Hs.75318 AA180742 tubulina alfa-4
260	DHFR: dihidrofolato reductasa Hs.83765 R00884 DHFR = dihidrofolato reductasa
261	DHFR: dihidrofolato reductasa Hs.83765 N52980
262	MGC5528: proteína hipotética MGC5528 Hs.315167 AA934904
263	NNMT: nicotinamida N-metiltransferasa Hs.76669 T72089
264	TUBB: tubulina, polipéptido beta Hs.179661 AI672565
265	HSPA1L: similar a la proteína 1 del choque térmico de 70 kD Hs.80288 H17513 HSP70-HOM = proteína 1 del choque térmico de 70 KD
266	TUBA1: **tubulina, alfa 1 (específica de los testículos) Hs.75318 R36063
267	PRO1073: **proteína PRO1073 Hs.6975 AA176999 CIP4 = proteína 4 de interacción con Cdc42
268	POLD3: polimerasa (dirigida por ADN), delta 3 Hs.82502 AI017254
269	ESTs.: Moderadamente similar a la proteína hipotética T50635 DKFZp762L0311,1 [H. sapiens] Hs.47378 N38809
270	Proteína DKFZP564A122 Hs.187991 N57723
271	LRRFIP1: **proteína 1 de interacción con repetición rica en leucina (en FLII) Hs.326159 T84633
272	ESTs: Hs.55468 AA165312
273	ESTs:-Hs.31444-H16772
274	AFAP: proteína asociada al filamento de actina Hs.80306 R69355
275	CXCR4: quimiocina (motivo C—X—C), receptor 4 (fusina) Hs.89414 T62491 receptor 4 de quimiocinas CXC = fusina = receptor del neuropéptido Y = L3
276	MSH2: ****homólogo 2 de mutS (E. coli) (cáncer de colon, sin poliposis de tipo 1) Hs.78934 AA679697
277	ESTs: Hs.48474 N62074
278	AA677337:
279	ESTs.: Moderadamente similar a CADENA BETA-2 DE LA TUBULINA HUMANA TBB2 [H. sapiens] Hs.23189 AA629908

ES 2 595 410 T3

280	HP1-BP74: HP1-BP74 Hs.142442 H79795
281	FLJ20101: Proteína NUDE1 de interacción con LIS1, homólogo de rata Hs.263925 AA459394
282	ARNm de Homo sapiens; ADNc DKFZp434D1428 (del clon DKFZp434D1428); cds completa Hs.321775 AA431268:
283	ESTs: Hs.265592 AA992658
284	ESTs:
285	DDX11: Polipéptido de caja DEAD/H (Asp-Glu-Ala-Asp/His) 11 (helicasa de tipo CHL1 de S. cerevisiae) Hs.27424 AA402879
286	CDC27: ciclo de división celular 27 Hs.172405 T81764
287	ARGBP2: Proteína de interacción con Arg/Abl ArgBP2 Hs.278626 N89738
288	DKFZP564A122: Proteína DKFZP564A122 Hs.187991 AA025807
289	OPN3: opsina 3 (encefalopsina) Hs.279926 AA150060
290	DKFZP566C134: Proteína DKFZP566C134 Hs.20237 AA456319
291	KIAA0855: golgina-67 Hs.182982 H15101
292	PIN: dineína, citoplasmática, polipéptido ligero Hs.5120 AA644679
293	ESTs.; débilmente similar al PRECURSOR DE LA PROTEÍNA 1 RELACIONADA CON LA LIPASA PANCREÁTICA HUMANA LIP1 [H. sapiens] Hs.68864 AA088857
294	HDAC3: histona desacetilasa 3 Hs.279789 AA973283
295	DONSON: vecino cadena abajo de SON Hs.17834 AA417895
296	LOC51053: geminina Hs.234896 AA447662
297	FLJ10545: proteína hipotética FLJ10545 Hs.88663 AA460110
298	MAD2L1: similar a MAD2 (deficiente en la detención mitótica, levadura, homólogo) 1 Hs.79078 AA481076 homólogo de la proteína de control de retroalimentación mitótica Madp2
299	TASR2: Proteína 2 de serina-arginina asociada con TLS Hs.3530 H11042
300	MCM6: deficiente en mantenimiento de minicromosoma (mis5, S. pombe) 6 Hs.155462 N57722
301	CIT: citron (serina/treonina cinasa 21 de interacción con rho) Hs.15767 W69425
302	**ESTs: Hs.205066 AA284803
303	ICAM1: molécula 1 de adhesión intercelular (CD54), receptor de rinovirus humano Hs.168383 R77293 CD54 = ICAM-1
304	KIAA0855: golgina-67 Hs.182982 AA456818
305	ESTs.; débilmente similar a p150 putativa [H.sapiens] Hs.300070 R10422
306	DEEPEST: proteína relacionada con hélice superenrollada del huso mitótico Hs. 16244 AI652290
307	MCM2: deficiente en mantenimiento de minicromosoma (S. cerevisiae) 2 (mitotina) Hs.57101 AA454572
308	ADNc de Homo sapiens: FLJ22272 fis, clon HRC03192 Hs.50740 AA495943:
309	WISP1: **proteína WNT1 vía de señalización inducible 1 Hs.194680 T54850
310	KIAA0855: golgina-67 Hs.182982 AA280248
311	TEM8: marcador endotelial del tumor 8 Hs.8966 H58644
312	BITE: proteína de unión a p10 Hs.42315 H96392
313	RAN: RAN, miembro de familia del oncogén RAS Hs.10842 AA456636
314	EZH2: homólogo del potenciador de zeste (Drosophila) 2 Hs.77256 AA428252
315	MCM4: Deficiente en mantenimiento de minicromosoma (S. cerevisiae) 4 Hs.154443 W74071
316	DKFZp434J0310: proteína hipotética Hs.278408 AA279657 EST Desconocido UG Hs.23595 ESTs sc_id6950
317	PPP1R10: proteína fosfatasa 1, subunidad reguladora 10 Hs.106019 AA071526
318	H11: proteína quinasa H11; proteína similar a proteína pequeña del estrés HSP22 Hs.111676 H57493
319	ESTs.; débilmente similar a la proteína KIAA1074 [H.sapiens] Hs.200483 AA463220
320	ESTs.; débilmente similar a la SECUENCIA SX DE LA SUBFAMILIA DE ALU HUMANA ALU8 ADVERTENCIA DE CONTAMINACIÓN [H.sapiens] Hs.193452 N72576
321	AA775033:

ES 2 595 410 T3

322	LOC51004: proteína CGI-10 Hs.12239 AA677920
323	ESTs: Hs. 150028 AI292036
324	MCM6: deficiente en mantenimiento de minicromosoma (mis5, S. pombe) 6 Hs.155462 AA976533
325	ESTs,: Moderadamente similar a la proteína hipotética T50635 DKFZp762L0311,1 [H. sapiens] Hs.47378 AA406348
326	UCP4: proteína de desacoplamiento 4 Hs.40510 H60279
327	MSH5: homólogo de mutS (E. coli) 5 Hs.112193 AA621155
328	ROCK1: proteína quinasa que contiene hélice superenrollada asociada con Rho 1 Hs.17820 AA872143
329	KIAA0855: golgina-67 Hs.182982 AA694481
330	AA705332:
331	CDC27: ciclo de división celular 27 Hs.172405 N47994
332	DONSON: vecino cadena abajo de SON Hs.17834 AI732249
333	SH3GL2: similar a GRB2 de dominio SH3 2 Hs.75149 R12817
334	PRC 1: proteína 1 reguladora de la citocinesis Hs.5101 AA449336
335	ESTs,: débilmente similar al producto de proteína sin nombre [H. sapiens] Hs.99807 AA417744 desconocida UG Hs.119424 EST s sc_id2235
336	Humano: secuencia de ARNm del clon 23719 Hs.80305 AA425722
337	ARNm de Homo sapiens; ADNc de DKFZp564O2364 (a partir del clon DKFZp564O2364) Hs.28893 W90240:
338	ESTs,: débilmente similar al PRECURSOR DE LA PROTEÍNA 1 RELACIONADA CON LA LIPASA PANCREÁTICA LIP1_HUMANA [H. sapiens] Hs.68864 AA132858
339	TUB A3: tubulina, alfa, específica de cerebro Hs.272897 AA864642
340	AI283530:
341	ESTs: Hs.302878 R92512
342	PPP1R10: proteína fosfatasa 1, subunidad reguladora 10 Hs.106019 T75485
343	SFRS5: factor de corte y empalme, rico en arginina / serina 5 Hs.166975 R73672
344	SFRS5: factor de corte y empalme, rico en arginina / serina 3 Hs.166975 AA598400
345	PRIM1: primasa, polipéptido 1 (49 kDa) primasa de ADN AA025937 Hs.82741 (subunidad p48)
346	FLJ20333: proteína hipotética FLJ20333 Hs.79828 H66982
347	HSPA8: proteína 8 del choque térmico de 70 kD Hs.180414 AA620511
348	C4A: componente del complemento 4A Hs.170250 AA664406
349	DKC1: disqueratosis congénita 1, disquerina Hs.4747 AA052960
350	HP1-BP74: HP1-BP74 Hs.I42442 T84669
351	ETV4: gen 4 de la variante ets (proteína de unión potenciadora de E1 A, E1AF) Hs.77711 AA010400 E1 A-F=proteína de unión potenciadora de E1 A = variante de translocación ETS
352	ADNc de Homo sapiens: FLJ23037 fis, clon LNG02036, muy similar a ARNm de homólogo de proteína mad de Homo sapiens HSU68019 (Hmad-3) Hs.288261 W42414 Smad3 = hMAD-3 = homólogo de Madres en Contra de la Decapentaplegia (M):
353	KIAA0952: Proteína KIAA0952 Hs.7935 AA679150
354	STK9: serina/treonina cinasa 9 Hs.48915 N80713
355	NXF1: **factor de exportación de ARN nuclear 1 Hs.323502 R01238
356	FLJ12892: proteína hipotética FLJ12892 Hs.17731 AA449357
357	UNG: uracil-ADN glicosilasa Hs.78853 H15111
358	STK17B: **serina / treonina cinasa 17b (inductora de apoptosis) ) Hs.120996 AA419485
359	YWHAH: proteína de activación de la tirosina 3-monooxigenasa/triptófano 5-monooxigenasa, polipéptido eta Hs.75544 N69107
360	FLJ13154: proteína hipotética FLJ13154 Hs.25303 AA923560
361	LOC51116: proteína CGI-91 Hs.20776 AA459419
362	SSXT: sarcoma sinovial, traslocado al cromosoma X Hs.153221 N59206

ES 2 595 410 T3

363	K1AA0978: proteína K1AA0978 Hs.3686 AA485878
364	EST: Hs.147907 AI223432
365	FLJ23468: proteína hipotética FLJ23468 Hs.38178 AA431741
366	FLJ10339: **proteína hipotética FLJ10339 Hs.203963 N95450
367	BMP2:-proteína morfogenética ósea 2 Hs.73853 AA011061
368	PIR51: RAD51-proteína de interacción Hs.24596 AI214426
369	FLJ20364: proteína hipotética FLJ20364 Hs.32471 AA676296
370	EIE4A2: **factor de iniciación de la traducción eucariótica 4A, isoforma 2 Hs.173912 H54751
371	ESTs.: débilmente similar a la PROTEÍNA PORTADORA DE CARINITINA/ACILCARNITINA MITOCONDRIAL MCAT_HUMANA [H.sapiens] Hs.27769 AA469975
372	FLJ11323: proteína hipotética FLJ11323 Hs.25625 AA775600
373	DKFZP564D0764: proteína DKFZP564D0764 Hs.26799 AA460732
374	CTL2: CTL2 gen Hs.105509 AA454710
375	ESTs: Hs.293419 AA775845
376	IFIT1: proteína inducida por interferón con repeticiones de tetratricopéptido 1 Hs.20315 AA489640 proteína inducida por interferón de 56-kDa
377	RBBP8: proteína 8 de unión a retinoblastoma Hs.29287 H23021
378	**Homo sapiens, secuencia del ARNm del clon 25061 Hs.179397 R38894:
379	Ser humano: secuencia de ADN del clon RP3-383J4 en el cromosoma 1q24,1-24,3 contiene parte de un gen que codifica un motivo kelch que contiene proteína, parte de un nuevo gen que codifica una proteína similar a la aspartil-ARNt sy Hs.117305 N29457
380	FLJ12888: proteína hipotética FLJ12888 Hs.284137 N68390
381	ESTs.: débilmente similar a la SUBUNIDAD 8 DEL FACTOR 3 DE INICIACIÓN DE LA TRADUCCIÓN EUCARIÓTICA HUMANO IF38 [H.sapiens] Hs.222088 AI139629
382	ESTs: Hs.241101 AA133590
383	H4FI: familia de histonas H4 , miembro I Hs.46423 AI218900
384	SP38: proteína de unión a la zona pelúcida Hs.99875 AA400474
385	GABPB1: subunidad beta 1 factor de transcripción de la proteína de unión a GA, subunidad beta 1 (53kD) Hs.78915 H91651
386	LCHN: proteína LCHN Hs.12461 AA029330
387	DKFZP564D0462: proteína hipotética DKFZp564D0462 Hs.44197 N32904
388	LENG8: nuevo gen 8 codificado en el clúster del receptor de leucocitos (LRC) Hs.306121 AA464698
389	HIF1A: factor 1 inducible por hipoxia, subunidad alfa (factor básico de transcripción de la hélice-bucle-hélice) Hs.197540 AA598526
390	ESTs: Hs.93714 R09201
391	FLJ23468: proteína hipotética FLJ23468 Hs.38178 AA454949
392	DKFZP566C134: proteína DKFZP566C134 Hs.20237 AA448164
393	PPP3CA: proteína fosfatasa 3 (anteriormente 2B), subunidad catalítica, isoforma alfa (calcineurina A alfa) Hs.272458 W60310
394	HMGE: proteína cochaperona similar a GrpEHs.151903 H55907
395	CDK7: cinasa dependiente de ciclina 7 (homólogo de la cinasa de activación de cdk de Xenopus MO15 ) Hs.184298 R22624 CAK = = cdk7 NRTALRE = = SDK cinasa de activación de CDK: CDK7
396	ABCC5: **** casete de unión a ATP, subfamilia C (CFTR / MRP), miembro 5 Hs.108660 AA186613
397	AA477707:
398	**ESTs: Hs.15607 R92899
399	LOC57209: proteína de dedo de cinc de tipo Kruppel Hs.25275 N50827
400	FLJ20101: proteína NUDE1 de interacción con LIS1, homólogo de rata Hs.263925 R87716
401	KNSL4: similar a cinesina 4 Hs.119324 AA430503
402	E2F5: E2F factor de transcripción 5, de unión a p130 Hs.2331 AA455521 E2F-5 = factor de transcripción de unión a pRB

ES 2 595 410 T3

403	TMPO: timopoyetina Hs.11355 T63980
404	POLQ: polimerasa (dirigida por ADN), theta Hs.241517 AI057325
405	TGIF: Factor de interacción con TG (homeobox de la familia TALE) Hs.90077 H51705
406	TRIP13: interaccionador con el receptor de la hormona tiroidea 13 Hs.6566 AA630784
407	GAS6: 6 específica de la detención del crecimiento Hs.78501 AA461110
408	HN1: 1 de expresión hematológica y neurológica Hs.109706 AA035429
409	BARD1: dominio RING asociado a BRCA1 1 Hs.54089 AA558464
410	DHFR: dihidrofolato reductasa Hs.83765 AA424790
411	AA490946:
412	ESTs: Hs.130435 AA167114
413	HSPA8: proteína del choque térmico de 70 kD 8 Hs.180414 AA629567
414	RRM2: polipéptido M2 de ribonucleótido reductasa Hs.75319 AA826373
415	FLJ20036: proteína hipotética FLJ20036 Hs.32922 H59114
416	COPEB: proteína de unión al elemento promotor del núcleo Hs.285313 AA055584 CPBP = CBA1 = proteína de unión al ADN
417	FLJ 10604: proteína hipotética FLJ10604 Hs.26516 N72697
418	ESTs: débilmente similar a ADNc EST yk415c12,5 viene de este gen [C. elegans] Hs.108824 H97880
419	UBE2D3: **enzima de conjugación de ubiquitina E2D 3 (homóloga a la levadura UBC4 / 5) Hs.118797 AA017199
420	FLJ10890: **proteína hipotética FLJ 10890 Hs.17283 AA004210
421	ESTs: Hs.214410 AA579336
422	OLR1: receptor de lipoproteína (similar a lectina) de baja densidad oxidada 1 Hs.77729 AA682386
423	FLJ13231: proteína hipotética FLJ13231 Hs.156148 W92787
424	EST:Hs.323101 W40398
425	ESTs: débilmente similar a R06F6.5b [C.elegans] Hs.180591 N59330
426	ADNc de Homo sapiens: FLJ23285 fis, clon HEP09071 Hs.90424 N26163:
427	Clon del ADNc del inserto de longitud completa del ARNm de Homo sapiens EUROIMAGE 42408 Hs.284123 AA211446:
428	NFKB1: factor nuclear del potenciador del gen del polipéptido ligero kappa en las células B 1 (p105) p105 Hs.83428 AA451716 NFKB1 = NF-kappaB = p50
429	LOC58486: proteína similar a la transposasa BusterI derivada de transposón Hs.25726 AA630256
430	ADNc de Homo sapiens-FLJ40976-fis, clon PLACE1001399 Hs.296323 AA424756:
431	KIAA0182: proteína KIAA0182 Hs.75909 A1023801
432	RANGAP1: proteína activadora 1 de Ran GTPasa Hs.183800 AA991855
433	P-KMYT1: cinasa inhibidora de cdc2 específica de tirosina y treonina asociada a la membrana y Cinasa Myt1 Hs.77783
434	HSPA8: proteína 8 del choque térmico de 70 kD Hs.180414 H64096
435	LUC7A: proteína sobreexpresada asociada a la resistencia al cisplatino Hs.3688 AA411969
436	RRM1: polipéptido M1 de ribonucleótido reductasa Hs.2934 AA633549
437	SET07: PR/SET proteína 7 que contiene dominio PR/SET Hs.111988 AA421470
438	**ESTs,: débilmente similar a la SECUENCIA J DE LA SUBFAMILIA DE ALU ALU1 HUMANA ADVERTENCIA DE CONTAMINACIÓN [H.sapiens] Hs.193452 W96179
439	Homo sapiens, secuencia del ARNm del clon 25058 Hs.179397 R38894:
440	ESTs,: débilmente similar a la proteína KIAA0973 [H.sapiens] Hs.14014 AA780791
441	EST: Hs.105298 AA489813
442	CTCF: Factor de unión a CCCTC (proteína de dedo de cinc) Hs.57419 H89996
443	HRB: Proteína de unión a HIV-1 Rev Hs.171545 AA485958
444	**ESTs: Hs.294083 AA447679

ES 2 595 410 T3

445	KIAA0878: Proteína KIAA0878 Hs.188006 AA599094
446	ESTs.; Débilmente similar a ALUB HUMAN !!!! ADVERTENCIA DE ALU CLASE B !!! [H.sapiens] Hs.180552 AA481283
447	OGT: N-acetilglucosamina ligada a O (GlcNAc) transferasa (UDP-N- acetilglucosamina: polipéptido-N-acetilglucosaminil transferasa) Hs. 100293 AA425229
448	ARNm de Homo sapiens para la proteína KIAA1700, cds parcial Hs.20281 N40952:
449	Ser humano: Secuencia de ADN del clon RP1-187J11 en el cromosoma 6q1 1.1-22.33. Contiene el gen para una nueva proteína similar a las proteínas predichas de S. pombe y S. cerevisiae, el gen para una nueva proteína similar Hs.72325 AA159962
450	KIAA1265: proteína KIAA1265 Hs.24936 AA479302
451	H1F0: familia de histonas HI, miembro 0 Hs.226117 H57830
452	ARGBP2: proteína de interacción con Arg / Abl ArgBP2 Hs.278626 H02525
453	ODF2: fibra densa exterior de colas de los espermatozoides 2 Hs.129055 AA149882
454	CD97: Antígeno CD97 Hs.3107 AI651871
455	BMI1: **homólogo de oncogén vírico de la leucemia murina (bmi-1) Hs.431 AA193573
456	POLG: polimerasa (dirigida por ADN), gamma Hs.80961 AA188629
457	XPR1: receptor de retrovirus xenotrópico y politrópico Hs.227656 AA453474
458	ESTs: Hs.128096 AA971179
459	DNAJB1: Homólogo de DnaJ (Hsp40), subfamilia B, miembro 1 Hs.82646 AA481022
460	ARL4: similar al factor de ADP-ribosilación similar 4 Hs.201672 AI142552
461	SFRS5: factor de corte y empalme, rico en arginina / serina 5 Hs.166975 AA598965
462	ESTs: Hs.25933 R11605
463	RIG-I: ARN helicasa Hs.145612 AA126958
464	FLJ-10339: proteína hipotética FLJ10339 Hs.203963 AA628231
465	DR1: regulador a la baja de la transcripción 1, unión a TBP (cofactor 2 negativo) Hs.16697 AA043503
766	Homo-sapiens similar a la proteína hipotética FLJ20093T clon MGC:1076, ARNm, cds completa Hs.298998 AA703249:
467	HSPC163: proteína HSPC163 Hs.108854 H98963
468	DKFZP564A122: proteína DKFZP564A122 Hs.187991 R27345
469	FLJ10128: autoantígeno uveal con dominios de hélice superenrollada y repeticiones de anquirina Hs.49753 T47624
470	DSCR1: Gen 1 de la región crítica del síndrome de Down Hs.184222 AA629707
471	FLJ10342: proteína hipotética FLJ10342 Hs.101514 AA490935
472	ARNm de Homo sapiens; ADNc DKFZp586N1323 (del clon DKFZp586N1323) Hs.24064 R26176:
473	ESTs: Hs.4983 H59921
474	ESTs.; Débilmente similar a ALUB HUMAN !!!! ADVERTENCIA DE ALU CLASE B !!! [H.sapiens] Hs.117949 H91167
475	CDC45L: Similar a CDC45 (ciclo de división celular 45, S. cerevisiae, homólogo) Hs.114311 AA700904
476	STAT5B: transductor de señales y activador de la transcripción 5B Hs.244613 AA280647 STAT5A/5B
477	ADNc de Homo sapiens FLJ14028 fis, clon HEMBA1003838 Hs.281434 AA454682:
478	KIAA1524: proteína KIAA1524 Hs.151343 AI248987
479	CTSD: catepsina D (lisosomal aspartil proteasa) Hs.79572 AA485373
480	Homo sapiens, similar a la proteína hipotética FLJ20093, clon MGC:1076, ARNm, cds completa Hs.298998 AA682274:
481	GTPBP2: GTP proteína de unión 2 Hs.13011 T67069
482	LOC51003: proteína CGI-125 Hs.27289 AA485945
483	VCL: vinculina Hs.75350 AA486727
484	KIF5B: miembro de la familia de cinesinas 5B Hs.149436 AA046613
485	CDC25A: ciclo de división celular 25A Hs.1634 AA071514

ES 2 595 410 T3

486	LOC51141: proteína 2 inducida por la insulina Hs.7089 AA045308
487	**ESTs,: moderadamente similares a CALD HUMAN CALDESMON [H.sapiens] Hs.117774 H48508
488	TBX3-iso: proteína TBX3-iso Hs.267182 T48941
489	KLAA0176: Proteína KLAA0176 Hs.4935 R44371
490	PRKAR1 A: proteína cinasa, dependiente de AMPc, reguladora, tipo I, alfa (extintor específico de tejido 1) Hs.183037 N25969 PKA-R1 alfa=cinasa dependiente de AMPc de tipo I-alfa-cata
491	ESTs: Hs.268991 H77818
492	ESTs,: débilmente similar a A53028 isopentenil-difosfato Delta-isomerasa [H.sapiens] Hs.9?70 R17362
493	ESTs,: débilmente similar a la proteína hipotética B34087 [H.sapiens] Hs.120946 H50656
494	TRN2: carioferina beta 2b, transportina Hs.278378 R08897
495	LMNA: lamina A/C Hs.77886 AA489582
496	NFE2L2: similar al factor nuclear (derivado de eritroide 2) 2 Hs.155396 AA629687
497	DKFZp762L0311: proteína hipotética DKFZp762L0311 Hs.16520 AA486418
498	ESTs,: débilmente similar a la proteína gigante S71752 p619 [H.sapiens] Hs.14870 T96829
499	ARNm de Homo sapiens; ADNc DKFZp434A1315 (del clon DKFZp434A1315); cds completa Hs.298312 AA991355:
500	E2IG4:proteína hipotética, inducida por estradiol Hs.8361 R13844
501	RANGAP1: proteína 1 activadora de Ran GTPasa Hs.183800 AA485734
502	H1F0: familia de histonas H1, miembro 0 Hs.226117 W69399
503	KIAA0239: proteína KIAA0239 Hs.9729 AA454740
504	ESTs,: débilmente similar a la SECUENCIA SQ DE LA SUBFAMILIA DE ALU ALU7 HUMANA ADVERTENCIA DE CONTAMINACIÓN [H.sapiens] Hs.68647 R96804
505	PRO0650: proteína PRO0650 Hs.177258 N54333
506	DNAJB9: homólogo de DnaJ (Hsp40), subfamilia B, miembro 9 Hs.6790 AA045792
507	ADNc de Homo sapiens: FLJ21971 fis, clon HEP05790 Hs.71331 AA774678:
508	LOC56996: **proteína de interacción con el cotransportador de cloruro catiónico Hs.119178 AA037466
509	AP3D1: Complejo 3 de la proteína relacionada con adaptador, subunidad delta 1 Hs.75056 AA630776
510	SGK: cinasa regulada por suero/glucocorticoides Hs.159640 AA486082 sgk= proteína cinasa de serina/treonina transcripcional putativa
511	HSPC148: proteína hipotética Hs.42743 R23666
512	MRPL19: proteína ribosómica mitocondrial L19 Hs.75574 AA521243 KIAA0104
513	AA455102:
514	ESTs: Hs.150325 AI278813
515	**ESTs: Hs.40527 AA029844
516	HSPC145: Proteína HSPC145 18349 AI271431
517	KIAA0170: producto génico de KIAA0170 Hs.277585 H68789
518	FLJ11127: proteína hipotética Hs.91165 T98200
519	KIAA0182: proteína KIAA0182 Hs.75909 H05099
520	FLJ2315I: proteína hipotética FLJ23151 Hs.137260 AA284259
521	AMD1: S-adenosimetionina descarboxilasa 1 Hs.262476 AA425692
522	FLJ10342: ** proteína hipotética FLJ10342 Hs.101514 AA934516
523	SPS: SELENOFOSFATO SINTETASA ; proteína donante de selenio humana Hs.124027 AA486372
524	KIAA1586: proteína KIAA1586 Hs.180663 AA779733
525	ICBP90: factor de transcripción Hs.108106 AA908902
526	ADNc de Homo sapiens: FLJ21971 fis, clon HEP05790 Hs.71331 AI002036:
527	ABCC2: casete de unión a ATP, subfamilia C (CFTR / MRP), miembro 2 Hs.193852 R91502
528	ARHGDIB: inhibidor beta de la disociación de GDP Rho (GDI) Hs.83656 AA487426 LyGDI=inhibidor de la disociación de GDP Rho 2=RHO GDI 2

ES 2 595 410 T3

529	RAD53: proteína cinasa Chk2 Hs.146379 AT653182
530	R96880:
531	TNFAIP3: factor de necrosis tumoral, proteína alfa inducida 3 Hs.211600 AA433807
532	ESTs: Hs.26979 H23469
533	AOC2: amina oxidasa, que contiene cobre 2 (específica de retina) Hs.143102 N50959
534	ARNm de Homo sapiens; ADNc DKFZp586N1323 (del clon DKFZp586N1323) Hs.24064 R30941:
535	AA452872:
536	ESTs: Hs.124169 R58970
537	ACYP1: acilfosfatasa 1, de tipo de eritrocitos (común) Hs.18573 W78754
538	SIL: locus de interrupción de TALI (SCL) Hs.323032 AA704809
539	AA016234:
540	ARNm de Homo sapiens; ADNc DKFZp566PI 124 (del clon DKFZp566P1124) Hs.321022 N50895:
541	KIAA1067: proteína KIAA1067 Hs.325530 AA099138
542	SMC4L1: similar a SMC4 (mantenimiento estructural de los cromosomas 4, levadura) 1 Hs.50758 AA452095
543	ESTs: Hs.29074 R70174
544	SNK: cinasa inducible por suero Hs.3838 AA460152
545	FANCG: anemia de Fanconi, grupo de complementación G Hs.8047 AA427484
546	ADNc de Homo sapiens: FLJ21531 fis, clon COL06036 Hs.102941 N95440:
547	ARNm de Homo sapiens; ADNc DKFZp547B086 (del clon DKFZp547B086) Hs.36606 N48700:
548	C10RF2: marco de lectura abierto 2 del cromosoma 1 Hs.19554 H11464 cote 1=ORF en el locus de la glucocerebrosidasa
549	HTF9C: pequeños fragmentos de HpalI 9C Hs.63609 H17888
550	ATF4: factor de transcripción de activación 4 (elemento potenciador de respuesta a tax B67) Hs.181243 AA600217
551	ESTs: Hs.101014 AA194941
552	CDC25A: ciclo de división celular 25A Hs.1634 AA913262
553	TOPK: PDZ de unión a cinasa; proteína cinasa originada en células T Hs. 104741 AI002631
554	ASIP: proteína de señalización agouti (de ratón) Hs.37006 AI220203
555	DKFZP564F013: **proteína hipotética DKFZp564F013 Hs.128653 R14908
556	ZNF265: proteína de dedo de cinc 265 Hs.194718 N66014
557	SLC30A1: familia portadora de solutos 30 (transportador de cinc), miembro 1 Hs.55610 AA195463
558	ESTs: Hs.28462 R63922
559	ESTs: Hs.114055 R27431
560	IL6: interleucina-6 (interferón, beta 2) Hs.93913 N98591 IL-6
561	H3F3B: histona H3, familia 3B (H3.3B) Hs.180877 AA608514
562	ESTs: Hs.81263 W81524
563	ADNc de Homo sapiens: FLJ23538 fis, clon LNG08010, muy similar al ARNm del clon epsilon/beta de la región MEN1 de BETA2 humana Hs.240443 AA400234:
564	AMD1: S-adenosilmetionina descarboxilasa 1 Hs.262476 R82299
565	MAP3K2: proteína cinasa activada por mitógeno 2 Hs.28827 AA447971
566	NET1: gen 1 transformante de células neuroepiteliales Hs.25155 R24543
567	CHAF1A: subunidad A (p150) del factor 1 de ensamblaje de cromatina Hs.79018 AA704459
568	MGC5585: proteína hipotética MGC5585 Hs.5152 H50655
569	KIAA1598: proteína KIAA1598 Hs.23740 H17868
570	PNN: pinina, proteína asociada a desmosomas Hs.44499 W86139

ES 2 595 410 T3

571	ESTs: Hs.238797 N70848
572	ESTs,: Débilmente similar a ALUB_HUMANA !!!! ADVERTENCIA DE ALU CLASE B!!! [H.sapiens] Hs.180552 AA600192
573	PDGFA: polipéptido alfa del factor de crecimiento derivado de plaquetas Hs.37040 AA701502
574	Homo sapiens, ARNm del clon FLC0675 PRO2870, cds completa Hs.306117 AA443127:
575	ESTs: Hs.143375 AA001841
576	TUBB: tubulina, polipéptido beta Hs.179661 H37989
577	MSH2: homólogo 2 de mutS (E. coli) (cáncer de colon, no poliposis de tipo 1) Hs.78934 AA219060 MSH2=homólogo de mutS para la reparación de apareamientos erróneos en el ADN
578	TOPBP1: proteína de unión II a la topoisomerasa (ADN) Hs.91417 R97785
579	KIAA0869: proteína KIAA0869 Hs.21543 R43798
580	H4FH: miembro H de la familia de histonas H4 Hs.93758 AA702781
581	FLJ23293: proteína hipotética FLJ23293 similar a la proteína 2 de interacción con ARL-6 Hs.31236 AA629027
582	**ADNc de Homo sapiens: FLJ23538 fis, clon LNG08010, muy similar al ARNm del clon epsilon/beta de la región MEN1 de BETA2 humana Hs.240443 AA400234:
583	KIAA0978: proteína KIAA0978 Hs.3686 N64780
584	KIAA1547: proteína KIAA1547 Hs.31305 AA057737
585	DKFZP761C169: proteína hipotética DKFZp761C169 Hs.71252 AA608709
586	WS-3: nueva proteína que contiene RGD Hs.39913 AA449975
587	FRZB: proteína relacionada con Frizzled Hs.153684 H87275
588	BRCA1: cáncer de mama 1, inicio temprano Hs. 194143 H90415 BRCA1=Mutada en el cáncer de mama y de ovarios
589	ESTs: Hs.4983 H22936
590	HSPC150: proteína HSPC150 similar a la enzima de conjugación con ubiquitina Hs.5199 AA460431
591	ARNm de Homo sapiens para la proteína KIAA1712, cds parcial Hs.29798 H54592:
592	FLJ11186: proteína hipotética FLJ11186 Hs.89278 AA504111 Desconocida UG Hs.89278 ESTs
593	ESTs,: Débilmente similar al producto proteico sin nombre [H.sapiens] Hs.118338 R25481
594	APEXL2: proteína similar a la endonucleasa apurínica/apirimidinica (nucleasa APEX) 2 Hs.154149 A1674393
595	CDR2: proteína relacionada con la degeneración del cerebelo (62 kD) Hs.75124 AA074613
596	ESTs: Hs.69662 AA459724
597	PSCD2L: homología a pleckstrina, similar a los dominios 2 Sec7 y de hélice superenrollada Hs.8517 AA464957
598	CRK: homólogo del oncogén CT10 del virus del sarcoma aviar v-crK Hs.306088 H75530
599	CCNE2: ciclina E2 Hs.30464 AA520999 Desconocida UG Hs.30464 ciclina E2
600	LOC51240: proteína hipotética Hs.7870 AA988037
601	FLJ11259: proteína hipotética FLJ11259 Hs.184465 AA485877
602	PTP4A1: proteína tirosina fosfatasa de tipo IVA, miembro 1 Hs.227777 AA482193
603	ADNc de Homo sapiens: FLJ22355 fis, clon HRC06344 Hs.288283 AA026375:
604	Ser humano: secuencia del ARNm del clon 23719 Hs.80305 H43437
605	ARNm del clon de Homo sapiens FLC0675 PRO2870, cds completa Hs.306117 AA485453:
606	MSE55: proteína constituyente del suero Hs.148101 H73234
607	CFLAR: regulador de la apoptosis similar a CASP8 y FADD Hs. 195175 AA453766
608	ADNc de Homo sapiens: FLJ22844 fis, clon KAIA5181 Hs.296322 AA975103:

ES 2 595 410 T3

<b>609</b>	Ser humano: secuencia de ADN del clon RP11-371L19 en el cromosoma 20 contiene dos genes nuevos, el gen para una nueva proteína similar a la proteína ribosómica 40S S10 (RPS10), ESTs, STSs, GSSs y cinco islas CpG Hs. 19002 R00846
<b>610</b>	ESTs: Hs.60054 R26390
<b>611</b>	ESTs.: débilmente similar a la SECUENCIA SQ DE LA SUBFAMILIA DE ALU ALU7 HUMANA ADVERTENCIA DE CONTAMINACIÓN [H.sapiens] Hs.325158 AA032084
<b>612</b>	FLJ10980: proteína hipotética FLJ10980 Hs.29716 N45467
<b>613</b>	IFIT1: **proteína inducida por interferón con repeticiones de tetratricopéptido 1 Hs.20315 AA157787
<b>614</b>	ESTs: Hs.21734 AA429809
<b>615</b>	DKFZP434C245: proteína DKFZP434C245 Hs.59461 AA705518
<b>616</b>	RNPS1: proteína de unión a ARN S1, dominio rico en serina Hs.75104 AA496837
<b>617</b>	FLJ13639: proteína hipotética FLJ13639 Hs.101821 AA131681
<b>618</b>	PCF11: homólogo de PCF11p Hs.123654 W73749
<b>619</b>	EIF4G3: factor 4 gamma de inicio de la traducción eucariótica, 3 Hs.25732 N92469
<b>620</b>	ADNc de Homo sapiens: FLJ21971 fis, clon HEP05790 Hs.71331 AA130595:
<b>621</b>	STAT1: transductor de señal y activador de la transcripción 1, 91kD Hs.21486 AA079495
<b>622</b>	BIRC3: que contiene repetición de IAP baculoviral 3 Hs.127799 R07870
<b>623</b>	HP1-BP74: HP1-BP74 Hs.142442 N20589
<b>624</b>	HSPC228: proteína hipotética Hs.267288 AI734268
<b>625</b>	KIAA0675: producto génico de KIAA0675 Hs.165662 AA454867
<b>626</b>	AMD1: S-adenosilmetionina descarboxilasa 1 Hs.262476 AA504772
<b>627</b>	EST: HS.149338-AI249089
<b>628</b>	PWP1: fosfoproteína nuclear similar a PWP1 de S. cerevisiae Hs.172589 AA485992
<b>629</b>	AI336973:
<b>630</b>	DUSP4: fosfatasa de especificidad dual 4 Hs.2359 AA444049
<b>631</b>	FLJ12788: proteína hipotética FLJ12788 Hs.20242 AA497041
<b>632</b>	HSPC150: proteína HSPC150 similar a la enzima de conjugación con ubiquitina Hs.5199 AA985450
<b>633</b>	FLJ11729: proteína hipotética FLJ11729 Hs.286212 W15533
<b>634</b>	KLF4: factor similar a Kruppel 4 (intestino) Hs.7934 H45668
<b>635</b>	FLJ11058: proteína hipotética FLJ11058 Hs.180817 N63911
<b>636</b>	FLJ23468: proteína hipotética FLJ23468 Hs.38178 AA460299
<b>637</b>	ESTs: Hs. 115315 AI278336
<b>638</b>	EBI3: gen inducido por el virus de Epstein-Barr 3 Hs.185705 AA425028 EBI3=receptor de citocina
<b>639</b>	ESTs: Hs.293797 N63988
<b>640</b>	MGAT2: manosil (alfa-1,6-)-glicoproteína beta- 1,2-N- acetilglucosaminiltransferasa Hs.172195 AA485653
<b>641</b>	H2BFQ: familia de histonas H2B. miembro Q Hs.2178 AA456298
<b>642</b>	NMB: neuromedina B Hs.83321 A1650675
<b>643</b>	SSR3: receptor de la secuencia señal, gamma (proteína gamma asociada con translocón) Hs.28707 AA453486
<b>644</b>	HSPC196: proteína hipotética Hs.239938 R78498
<b>645</b>	EST: Hs.44522 N33610
<b>646</b>	BRF1: factor 1 de respuesta a butirato (factor 1 de respuesta a EGF 1) Hs.85155 AA723035
<b>647</b>	MAN1A2: mannosidasa, alfa, clase 1A, miembro 2 Hs.239114 H97940
<b>648</b>	KIAA1201: proteína KIAA1201 Hs.251278 AA427719

ES 2 595 410 T3

649	NUCKS: similar a la caseína cinasa ubicua nuclear de rata 2 Hs.118064 AA158345
650	MAGEF1: proteína MAGEF1 Hs.306123 AA425302
651	Ser humano: Cromosoma 16 BAC, clon CIT987SK-A-362G6 Hs.6349 N75498
652	R40377:
653	AP3M2: complejo 3 de la proteína relacionada con adaptador, subunidad mu 2 Hs.77770 R14443
654	ESTs,: Débilmente similar a la proteína relacionada con la transcriptasa inversa 1207289A [H.sapiens] Hs.272135 AA705010
655	ARNm de Homo sapiens para la proteína FLJ00116, cds parcial Hs.72363 AA159893:
656	E1F4E: factor de iniciación de la traducción eucariótica 4E Hs.79306 AA193254
657	ARNm de Homo sapiens de proteína hipotética (gen TR2/D15) Hs.180545 N47285:
658	ESTs: Hs.99542 AA461474
659	CTNND1: catenina (proteína asociada a cadherina), delta 1 Hs.166011 AA024656
660	ESTs: Hs.188554 R75884
661	ZNF217: proteína de dedo de cinc 217 Hs.155040 R81830
662	FLJ12892: proteína hipotética-FLJ12892 Hs.17731 AI243595 -
663	ETV5: gen variante 5 de ets (molécula relacionada con ets) Hs.43697 AA460265
664	EST: Hs.251574 T54821
665	RPS25: proteína ribosómica S25 Hs.113029 T98662
666	CNN2: calponina 2 Hs.169718 AA284568
667	ESTs,: Débilmente similar a plakofilina 2b [H.sapiens] Hs.12705 AA485365
668	PAPPA: proteína A de plasma asociada con el embarazo Hs.75874 AA609463
669	TFF3: factor trébol 3 (intestinal) Hs.82961 N74131
670	AI204264:
671	DJ328E19.C1.1: proteína hipotética Hs.218329 AA486041
672	ME3: enzima málica 3, dependiente de NADP(+), mitocondrial Hs.2838 AA779401
673	ESTs,: Débilmente similar a la PROTEÍNA SENSIBLE A LA TRANSFORMACIÓN IEF5_HUMANA IEF SSP 3521 [H.sapiens] Hs.43213 AA490554
674	FLJ13181: proteína hipotética FLJ13181 Hs.301526 AA057266
675	KIAA1547: KIAA1547 proteína Hs.31305 AA136692
676	ZNF281: proteína de dedo de cinc 281 Hs.59757 N47468
677	ADNc de Homo sapiens: FLJ23260 fis, clon COL05804, muy similar a la secuencia del ARNm 23652 del clon humano HSU90911 Hs.13996 AA463961:
678	ESTs: Hs.25933 AA411392
679	NCBP1: proteína de unión al capuchón nuclear, subunidad 1, 80 kD Hs.89563 AA278749 proteína de unión al capuchón nuclear
680	H2BFL: Familia de la histona H2B, miembro L Hs.239884 H70774
681	DKFZP564A122: proteína DKFZP564A122 Hs.187991 H66150
682	NASP: proteína de esperma autoantigénico nuclear (unión a histona) Hs.243886 AA644128
683	**ESTs,: Débilmente similar a la proteína KIAA0822 [H.sapiens] Hs.98368 AA422008
684	MAP2K6: proteína cinasa activada por mitógeno 6 Hs.118825 H07920
685	ESTs: Hs.158357 AA865842
686	GADD45A: detención del crecimiento e inducible por daño de ADN, alfa Hs.80409 AA147214 GADD45 alfa=proteína de detención del crecimiento e inducible por daño de ADN
687	DHFR: dihidrofolato reductasa Hs.83765 AA488803
688	AA151930:

ES 2 595 410 T3

<b>689</b>	ARNm de Homo sapiens; ADNc DKFZp434P116 (del clon DKFZp434P116); cds completa Hs.103378 AA431133:
<b>690</b>	ARNm de Homo sapiens; ADNc DKFZp564D156 (del clon DKFZp564D156) Hs.9927 T55704:
<b>691</b>	ESTs: Hs.32204 R93719
<b>692</b>	PRPSAP1: proteína asociada a la fosforibosil pirofosfato sintetasa 1 Hs.77498 R20005
<b>693</b>	ZNF42: proteína 42 dedo de cinc (de respuesta a ácido retinoico específico de mielóide) Hs.169832 AA987906
<b>694</b>	**ESTs: Hs.43712 N25936
<b>695</b>	RUNX1: factor 1 de transcripción relacionado con runt (oncogén 1 de la leucemia mielógena aguda; lma 1) Hs.129914 AA146826
<b>696</b>	ARNm de Homo sapiens; ADNc DXEZp547C244 (del clon DKFZp547C244) Hs.9460 T64452:
<b>697</b>	TYMS: timidilato sintetasa Hs.82962 AA663310
<b>698</b>	MGC5528: proteína hipotética MGC5528 Hs.315167 AA843451
<b>699</b>	ESTs: Hs.268685 R22952
<b>700</b>	SFPQ: factor de corte y empalme rico en prolina/glutamina (asociado a la proteína de unión al tracto, polipirimidina) Hs.180610 AA418910
<b>701</b>	ESTs: Hs. 155105 AI221390
<b>702</b>	FLJ10624: proteína hipotética FLJ10624 Hs.306000 AA489592
<b>703</b>	TRIP8: interactor 8 con el receptor de hormona tiroidea Hs.6685 AA425205
<b>704</b>	DNAJB6: Homólogo de DnaJ (Hsp40), subfamilia B, miembro 6 Hs.181195 AA496105
<b>705</b>	ESTs: Hs.18331 T98244
<b>706</b>	RBM14: Proteína del motivo de unión a ARN 14 Hs.11170 AA421233
<b>707</b>	SCYA2: citocina A2 inducible pequeña (proteína 1 quimiotáctica de monocitos, homóloga a Sig-je de ratón) Hs.303649 AA425102 MCP-1=MCAF=citocina A2 inducible pequeña =JE=quimiocina
<b>708</b>	MGC4161: proteína hipotética MGC4161 Hs.177688 AI224867
<b>709</b>	TUBB2: tubulina, beta, 2 Hs.251653 AA888148
<b>710</b>	FLJ20280: proteína hipotética FLJ20280 Hs.270134 N74086
<b>711</b>	TERA: Proteína TERA Hs.180780 AA465096
<b>712</b>	CPS1: **carbamoil-fosfato sintetasa 1, mitocondrial Hs.50966 N68399
<b>713</b>	KIAA0802: proteína KIAA0802 Hs.27657 W55875
<b>714</b>	FYN: oncogén FYN relacionado con SRC, FGR, YES Hs.169370 N22980
<b>715:</b>	ARNm de Homo sapiens de PRO2751, cds completa Hs.283978 H12784:
<b>716</b>	CLTH: gen de leucemia mielóide de ensambaje de clatrina Hs.7885 AA441930
<b>717</b>	CHMP1.5: proteína CHMP1.5 Hs.42733 W85875
<b>718</b>	SMARCB1: regulador de la cromatina dependiente de actina, asociado a la matriz, relacionado con SWI/SNF, subfamilia b, miembro 1 Hs.159971 AA446018
<b>719</b>	AA487823: SRF=factor de transcripción de unión al elemento de respuesta en suero c-fos
<b>720</b>	**ESTs: Hs.130741 AA608725
<b>721</b>	ADNc de Homo sapiens FLJ10976 fis, clon PLACE1001399 Hs.296323 R36085:
<b>722</b>	FLJ20036: proteína hipotética FLJ20036 Hs.32922 N91145
<b>723</b>	C11ORF5: marco de lectura abierto 5 del cromosoma 11 Hs. 121025 AA776702
<b>724</b>	AF3P21: proteína SH3 Hs. 102929 N94372
<b>725</b>	LOC54104: proteína hipotética Hs.12871 H05934
<b>726</b>	DF: componente D del complemento (adipsina) Hs.155597 AA233549

ES 2 595 410 T3

<b>727</b>	CEP4: Proteína efectora 4 de Cdc42; enlazador de Rho GTPasas 4 Hs.3903 AA449061
<b>728</b>	KIF5Bt: miembro 5B de la familia de quinesinas Hs.149436 AA644218
<b>729</b>	MGC5627: proteína hipotética MGC5627 Hs.237971 H02336
<b>730</b>	G3BP: proteína de unión al dominio SH3 de la proteína de activación de Ras-GTPasa Hs.220689 AA449834
<b>731</b>	ESTs: Hs.293987 AA229758
<b>732</b>	ESTs: Hs.36828 AA194796
<b>733</b>	ARNm de Homo sapiens para la proteína FLJ00101, cds parcial Hs.221600 W92262:
<b>734</b>	ADNc de Homo sapiens: FLJ21288 fis, clon COL01927 Hs.6019 R07184:
<b>735</b>	ESTs,: Débilmente similar a la proteína relacionada con la transcriptasa inversa 1207289A [H.sapiens] Hs.250594 H86813
<b>736</b>	ADNc de Homo sapiens FLJ11941 fis, clon HEMBB1000649 Hs.124106 AI301573:
<b>737</b>	ESTs: Hs.24908 H77726
<b>738</b>	TOB2: transductor de ERBB2, 2 Hs.4994 AA486088
<b>739</b>	ESTs: Hs.143900 AI193212
<b>740</b>	ARNm PRO2870 del clon de Homo sapiens FLC0675, cds completa Hs.306117 H16589:
<b>741</b>	ESTs,: Débilmente similar a la proteína KIAA0638 [H.sapiens] Hs.296288 T83657
<b>742</b>	FLJ20039: proteína hipotética FLJ20039 Hs.267448 AA448268
<b>743</b>	RPA2: proteína de replicación A2 (32kD) Hs.79411 R13557
<b>744</b>	GAS1: específica de la detención del crecimiento 1 Hs.65029 AA025819
<b>745</b>	Ser humano: Secuencia de ADN del clon 967N21 en el cromosoma 20p12.3-13. Contiene el gen de CHGB para la cromogranina B (secretogranina 1, SCG1), un pseudogene similar a la parte de K1AA0172, el gen para una nueva proteína Hs.88959 R56678
<b>746</b>	ESTs: Hs.21175 AI341642
<b>747</b>	LBC: oncogén de la crisis de blastos linfoides Hs.301946 AA135716
<b>748</b>	ESTs: Hs.194595 R06761
<b>749</b>	MGC4707: proteína hipotética MGC4707 Hs.291003 R14653
<b>750</b>	ZNF183: proteína de dedo de cinc 183 (dedo ANILLO, tipo C3HC4) Hs.64794 AA132766
<b>751</b>	RAD18: proteína de reparación posreplicación hRAD18p Hs.21320 R59197
<b>752</b>	EIF4EBP2: **proteína de unión al factor de iniciación de la traducción eucariótica 4E 2 Hs.278712 H15159
<b>753</b>	**ARNm de Homo sapiens; ADNc DKFZp586M0723 (del clon DKFZp586M0723) Hs.27860 AA446650:
<b>754</b>	ORC3L: similar al complejo de reconocimiento del origen, subunidad 3 (homólogo de levadura) Hs.74420 H00257
<b>755</b>	CDK7: cinasa 7 dependiente de ciclina (homólogo de cinasa de activación de MO15 cdk de Xenopus) Hs.184298 AI311067
<b>756</b>	USP10: proteasa específica de ubiquitina 10 Hs.78829 AA455233
<b>757</b>	K1AA0733: proteína de unión a TAK1 2; proteína KIAA0733 Hs.109727 AA931658
<b>758</b>	R89286:
<b>759</b>	ALDH4: aldehído deshidrogenasa 4 (glutamato gamma-semialdehído deshidrogenasa; pirrolin-5-carboxilato deshidrogenasa) Hs.77448 AA181378
<b>760</b>	IDN3: proteína IDN3 Hs.225767 N62911
<b>761</b>	ESTs: Hs.50180 H48143
<b>762</b>	M1G2: inducible por mitógeno 2 Hs.75260 H29252
<b>763</b>	KIAA0856: proteína KIAA0856 Hs.13264 R12847
<b>764</b>	EST: Hs.47763 N54162
<b>765</b>	ARNm de Homo sapiens; ADNc DKFZp547C244 (del clon DKFZp547C244) Hs.9460 AA447553:

ES 2 595 410 T3

<b>766</b>	KIAA0855: golgina-67 Hs.182982 AA775625
<b>767</b>	ESTs,: Débilmente similar a la nucleolina JH0148 de rata [R.norvegicus] Hs.30120 R54659
<b>768</b>	FLJ22313: proteína hipotética FLJ22313 Hs.30211 H52061
<b>769</b>	ESTs: Hs.71818 AI028074
<b>770</b>	K1AA0618: producto génico de KIAA0618 Hs.295112 AA455506
<b>771</b>	ESTs: Hs.59413 W93056
<b>772</b>	ESTs: Hs. 165607 AA992090
<b>773</b>	UBAP: proteína asociada a ubiquitina Hs.75425 AA446016
<b>774</b>	HAN11: proteína de repetición de WD Hs. 176600 AA725641
<b>775</b>	USP16: proteasa específica de ubiquitina 16 Hs.99819 AA489619
<b>776</b>	ESTs: Hs.67776 AA464963
<b>777</b>	SM-20: similar a la proteína SM-20 de músculo liso de rata Hs.6523 H56028
<b>778</b>	CCNG2: ciclina G2 Hs.79069 AA489647
<b>779</b>	ARNm de Homo sapiens; ADNc DKFZp566P1124 (del clon DKFZp566P1124) Hs.321022 N62953:
<b>780</b>	FLJ20094: proteína hipotética FLJ20094 Hs.29700 N95490
<b>781</b>	LOC51174: delta-tubulina Hs.270847 W33133
<b>782</b>	ARNm de Homo sapiens; ADNc DKFZp43411820 (del clon DKFZp43411820); cds parcial Hs.14235 N52394:
<b>783</b>	FANCA: anemia de Fanconi, grupo de complementación A Hs.284153 AA644129
<b>784</b>	P5-1: ORF de la región del MHC de clase I Hs.1845 T58146
<b>785</b>	DNA2L: similar a DNA2 (helicasa de replicación del ADN, levadura, homólogo) Hs. 194665 AA974495 KIAA0083
<b>786</b>	LOC51578: **proteína de la glándula suprarrenal AD-004 Hs.279586 AA150301
<b>787</b>	ESTs: Hs.326417-AA913304
<b>788</b>	CDKN2D: inhibidor 2D de la cinasa dependiente de ciclina (p19, inhibe CDK4) Hs.29656 R77517 p19-INK4D= inhibidor D de la cinasa 4 dependiente de ciclina
<b>789</b>	FABP1: proteína 1 de unión a ácidos grasos, hígado Hs.5241 AA682392
<b>790</b>	TERA: proteína TERA Hs.180780 AA906997
<b>791</b>	ESTs: Hs.145383 AI253072
<b>792</b>	SLC7A5: vehículo de solutos de la familia 7 (transportador de aminoácidos catiónicos, sistema y+), miembro 5 Hs.184601 AA419176
<b>793</b>	AXL: tirosina cinasa del receptor de AXL -Hs.83341 H15336 axl=ufo= receptor de la tirosina cinasa
<b>794</b>	LOC57190: selenoproteína N Hs.8518 AA284276
<b>795</b>	ESTs: Hs.99037 AA443948
<b>796</b>	STCH: chaperona de la proteína 70 del estrés, asociada a microsomas, 60kD Hs.288799 H85311
<b>797</b>	ESTs: Hs.88523 AA278591 Desconocida UG Hs.88523 ESTs
<b>798</b>	ESD: **esterasa D/formilglutación hidrolasa Hs.82193 AA250931
<b>799</b>	ESTs: Hs.122444 R31021
<b>800</b>	ESTs: Hs.283127 AI291262
<b>801</b>	K1AA0480: **producto génico KIAA0480 Hs.92200 H91332
<b>802</b>	HP1-BP74: HP1-BP74 Hs.142442 AA598791
<b>803</b>	**ESTs,: Moderadamente similar a la SECUENCIA DE LA SUBFAMILIA J DE ALU HUMANA ALU1 ADVERTENCIA DE CONTAMINACIÓN [H.sapiens] Hs.144662 AA987667
<b>804</b>	TTF2: factor de terminación de la transcripción, ARN polimerasa II Hs.142157 AI023603
<b>805</b>	ESTs: Hs.13740 T70541

ES 2 595 410 T3

<b>806</b>	DJ37E16.5: proteína hipotética dJ37E16.5 Hs.5790 AA400021
<b>807</b>	CDH24: similar a cadherina 24 Hs.155912 AI732266
<b>808</b>	DJ465N24.2.1: **proteína hipotética dJ465N24.2.1 Hs.8084 AA932375
<b>809</b>	ESTs,: Débilmente similar a la proteína S57447 HPBRIL-7 [H.sapiens] Hs.16346 AA410490
<b>810</b>	ADNc de Homo sapiens: FLJ23285 fis, clon HEP09071 Hs.90424 AI005038:
<b>811</b>	KRAS2: homólogo el oncogén del virus del sarcoma de rata 2 v-Ki-ras2 Kirsten Hs.184050 N95249
<b>812</b>	FLJ20038: proteína hipotética FLJ20038 Hs.72071 H96090
<b>813</b>	ESTs,: Débilmente similar a la SECUENCIA DE LA SUBFAMILIA SB2 DE ALU HUMANA ALU4 ADVERTENCIA DE CONTAMINACIÓN [H.sapiens] Hs.28848 AA486607
<b>814</b>	H2AFN: Familia de la histona H2A, miembro N Hs.134999 AI095013
<b>815</b>	RERE: repeticiones del dipéptido de arginina-ácido glutámico (RE) Hs.194369 AA490249
<b>816</b>	USP1: proteasa específica de ubiquitina 1 Hs.35086 T55607
<b>817</b>	TIP47: proteína de selección de cargo (proteína de unión al receptor de manosa 6 fosfato) Hs.140452 AA416787
<b>818</b>	KIAA0135: proteína KIAA0135 Hs.79337 AA427740 KIAA0135=relacionada con cinasa pim-1
<b>819</b>	ESTs: Hs.214410 T95273
<b>820</b>	PPP1R2: proteína fosfatasa 1, subunidad 2 reguladora (inhibidora) Hs.267819 N52605
<b>821</b>	ADNc de Homo sapiens: FLJ21210 fis, clon COL00479 Hs.325093 AA978323:
<b>822</b>	CSNK2A2: caseína cinasa 2, polipéptido alfa prime Hs.82201 AA054996
<b>823</b>	HSRTSBETA: proteína rTS beta Hs.180433 N66132
<b>824</b>	FLJ13110: proteína hipotética FLJ13110 Hs.7358 AA431233
<b>825</b>	ESTs: Hs.238797 N30704
<b>826</b>	FYN: oncogén FYN relacionado con SRC, FGR, YES Hs.169370 N35086
<b>827</b>	RBM8A: proteína del motivo de unión a ARN 8A Hs.65648 AA448402
<b>828</b>	ESTs: Hs.21906 AA608546
<b>829</b>	ESTs: Hs.128081 AA971042
<b>830</b>	PP591: proteína hipotética PP591 Hs.118666 AA626336
<b>831</b>	N63866:
<b>832</b>	HM74: receptor putativo de quimiocina; proteína de unión a GTP Hs.137555 R02739
<b>833</b>	MIDI: línea media 1 (síndrome de Opitz/BBB) Hs.27695 AA598640
<b>834</b>	KIAA1586: proteína KIAA1586 Hs.180663 AA938639
<b>835</b>	ARNm de Homo sapiens clon CDABP0014 Hs.92679 AA443139:
<b>836</b>	HSU79274: proteína predicha por el clon 23733 Hs.150555 AA451900
<b>837</b>	AOC3: amina oxidasa, que contiene cobre 3 (proteína de adhesión vascular 1) Hs.198241 AA036974
<b>838</b>	AA548037:
<b>839</b>	FLJ10154: proteína hipotética FLJ10154 Hs.179972 AA457133
<b>840</b>	THBS1: trombospondina 1 Hs.87409 AA464532
<b>841</b>	DNAJB6: homólogo de DnaJ (Hsp40), subfamilia B, miembro 6 Hs.181195 AA431203
<b>842</b>	KIAA1547: proteína KIAA1547 Hs.31305 AI216623

ES 2 595 410 T3

<b>843</b>	GATA2: proteína de unión a GATA 2 Hs.760 R32405
<b>844</b>	ESTs: Hs.176950 R82522
<b>845</b>	KIAA1018: proteína KIAA1018 Hs.5400 AA156859
<b>846</b>	B4GALT1: **UDP-Gal: betaGlcNAc beta 1,4- galactosiltransferasa, polipéptido 1 Hs.198248 AA043795
<b>847</b>	HMGCR: 3-hidroxi-3-metilglutaril-Coenzima A reductasa Hs.11899 AA779417
<b>848</b>	ESTs,: Débilmente similar a la proteína 1819485A CENP-E [H.sapiens] Hs.167652 H94466
<b>849</b>	ESTs: Hs.294088 AA971073
<b>850</b>	KIAA1637: coactivador independiente de AF-2 (CIA); proteína KIAA1637 Hs.288140 AA918007
<b>851</b>	HSPC196: proteína hipotética Hs.239938 H66023
<b>852</b>	DR1: regulador a la baja de la transcripción 1, unión a TBP (cofactor 2 negativo) Hs.16697 AA132007
<b>853</b>	CG1I: proteína putativa de interacción con ciclina G1 Hs.10028 AA486444
<b>854</b>	IGSF4: superfamilia de las inmunoglobulinas, miembro 4 Hs.70337 AA487505
<b>855</b>	ESTs: Hs. 179309 AA664350
<b>856</b>	HSPC163: proteína de repetición de HSPC163 Hs. 108854 AA053139
<b>857</b>	FLJ12788: proteína hipotética FLJ12788 Hs.20242 AI061317
<b>858</b>	FEM1B: homólogo b de FEM-1 (C.elegans) Hs.6048 H82273
<b>859</b>	FXR1: retraso mental por cromosoma X frágil, homólogo autosómico 1 Hs.82712 N62761
<b>860</b>	NCOA3: coactivador 3 del receptor nuclear Hs.225977 AA156793
<b>861</b>	H2BFB: familia de la histona H2B, miembro B Hs.180779 N33927
<b>862</b>	ESTs: Hs.23830 AA460601
<b>863</b>	CDK7: cinasa 7 dependiente de ciclina (homólogo de la cinasa activadora de MO15 cdk de Xenopus) Hs.184298 AA03I961-CAK=cdk7=NRTALRE=sdk= cinasa activadora de CDK
<b>864</b>	FLJ20259: proteína hipotética FLJ20259 Hs.9956 T55949
<b>865</b>	ADNc de Homo sapiens FLJ20678 fis, clon KAIA4163 Hs.143601 T95823:
<b>866</b>	RPS19: proteína ribosómica S19 Hs.298262 T72208
<b>867</b>	ARNm de Homo sapiens; ADNc DKFZp434M0420 (del clon DKFZp434M0420) Hs.326048 AA443976:
<b>868</b>	TP53: proteína tumoral p53 (síndrome de Li-Fraumeni) Hs.1846 R39356 p53
<b>869</b>	FBI1: Inductor del VIH-1 de la proteína de unión a transcritos cortos Hs.104640 R06252
<b>870</b>	GOT1: glutámico-oxaloacético transaminasa 1, soluble (aspartato aminotransferasa 1) Hs.597 H22855
<b>871</b>	FLJ21434: proteína hipotética FLJ21434 Hs.298503 AA680129
<b>872</b>	DNMT2: ADN (citosina-5-)-metiltransferasa 2 Hs.97681 R95731
<b>873</b>	ESTs: Hs.55272 W02785
<b>874</b>	H2BFQ: Familia de la histona H2B, miembro Q Hs.2178 AA010223
<b>875</b>	NFIC: factor nuclear I/C (factor de transcripción de unión a CCAAT) Hs.184771 N20996
<b>876</b>	NPTX1: pentraxina 1 neuronal Hs.84154 H22445
<b>877</b>	TLOC1: proteína de translocación 1 Hs.8146 AA450205
<b>878</b>	MGC5302: proteína 58 residente en el retículo endoplásmico; proteína hipotética MGC5302 Hs.44970 N39195

ES 2 595 410 T3

879	ACTR2: ARP2 (proteína 2 relacionada con actina, levadura) homólogo Hs.42915 AA032090
880	AI287555:
881	ABCA7: casete de unión a ATP, subfamilia A (ABC1), miembro 7 Hs.134514 AI668632
882	COL7A1: colágeno, de tipo VII, alfa 1 (epidermoliosis ampollosa, distrófica, dominante y recesiva) Hs.1640 AA598507
883	RFC2: factor de replicación C (activador 1) 2 (40kD) Hs.139226 AA663472
884	FLJ22583: proteína hipotética FLJ22583 Hs.287700 AA135836
885	**ESTs,: Débilmente similar al ORF2 [M.musculus] Hs.172208 AI820570
886	ESTs: Hs.21667 R15709
887	RBBP4: proteína de unión al retinoblastoma 4 Hs.16003 AA705035
888	ARNm de Homo sapiens; ADNc DKFZp434J1027 (del clon DKJFZp434J1027); cds parcial Hs.22908 R20166:
889	ESTs: Hs.166539 AI080987
890	NKTR: secuencia de reconocimiento de tumor-citolítica natural Hs.241493 AA279666 proteína de reconocimiento de tumor-NK=proteína relacionada con ciclofilina
891	MUC1: mucina 1, transmembrana Hs.89603 AA486365
892	AP4B1: complejo 4 de la proteína relacionada con el adaptador, subunidad beta 1 Hs.28298 AA481045
893	ESTs: Hs.94943 AA452165
894	MITF: factor de transcripción asociado a microftalmia Hs.166017 N66177
895	ESTs: HS.483299 AA286914 Desconocida UG Hs.183299 ESTs sc_id2032
896	BAG3: atanogén asociado con BCL2 3 Hs.15259 AI269958
897	INSR: receptor de la insulina Hs.89695 AA001106
898	TRIP: proteína de interacción con TRAF Hs.21254 AA186426
899	EST: Hs.307975 R22182
900	**ADNc de Homo sapiens: FLJ23037 fis, clon LNG02036, muy similar al ARNm del homólogo de la proteína mad de Homo sapiens HSU68019 (hMAD-3) Hs.288261 W72201:
901	HLA-DNA: complejo mayor de histocompatibilidad , clase II, DN alfa Hs.11135 AA702254 complejo mayor de histocompatibilidad, clase II, DN alfa
902	FLJ10392: **proteína hipotética FLJ10392 Hs.20887 AI261305
903	MPHOSPH1: **fosfoproteína de fase M 1 Hs.240N63752
904	STAG1: antígeno estromal 1 Hs.286148 R36160
905	USP1: proteasa específica de ubiquitina 1 Hs.35086 T55607
906	ESTs,: Débilmente similar a la SECUENCIA DE LA SUBFAMILIA SB2 ALU DE ALU4_HUMANO ADVERTENCIA DE CONTAMINACIÓN [H.sapiens] Hs.28848 AA486607
907	PA26: proteína nuclear PA26 regulada por p53 Hs.14125 AA447661
908	ESTs,: Débilmente similar a la proteína de dedo de cinc [H.sapiens] Hs.71243 N92478
909	SH3PX1: proteína SH3PX1 1ue contiene los dominios SH3 y PX Hs.7905 R69163
910	**ADNc de Homo sapiens: FLJ22554 fis, clon HSI01092 Hs.93842 H58317:
911	RPS25: proteína ribosómica S25 Hs.113029 AA779404
912	ESTs,: Débilmente similar a la región V-I de la cadena kappa de Ig [H.sapiens] Hs.5890 N34799 fra-2=antígeno 2 relacionado con fos
913	TXNRD1: tioredoxina reductasa 1 Hs.13046 AA453335 tioredoxina reductasa

ES 2 595 410 T3

<b>914</b>	**ESTs: Hs.184378 N77828
<b>915</b>	GCSH: proteína H del sistema de escisión de glicina (vehículo de aminometilo) Hs.77631 R71327
<b>916</b>	ADNc de Homo sapiens FLJ11904 fis, clon HEMBB1000048 Hs.285519 AA447098:
<b>917</b>	NCOA3: coactivador del receptor nuclear 3 Hs.225977 H51992 AIBI= amplificado en el cáncer de mama=TRAM-1 =RAC3=ACTR=CAGH 16=nucl
<b>918</b>	FLJ20159: proteína hipotética FLJ20159 Hs.288809 R33122
<b>919</b>	IL7R: receptor de la interleucina 7 Hs.237868 AA487121
<b>920</b>	RAB23: RAB23, miembro de la familia de oncogenes RAS Hs.94769 AA134569
<b>921</b>	ESTs: Hs.132493 AA923168
<b>922</b>	ESTs: Hs.87507 AA236015
<b>923</b>	SHC1: SHC proteína transformante 1 (que contiene el dominio 2 de homología a Src) Hs.81972 R52960
<b>924</b>	KIAA1321: proteína KIAA1321 Hs.24336 W37999
<b>925</b>	GLI: homólogo del oncogén asociado a glioma (proteína de dedo de cinc) Hs.2693 AI373071
<b>926</b>	ESTs: Hs.183299 AA291137 Desconocido UG Hs.183299 ESTs sc_id2032
<b>927</b>	GPRK6: cinasa 6 del receptor acoplado a proteína G Hs.76297 AA291284
<b>928</b>	ESTs: Hs.93704 AA702684
<b>929</b>	CAPS: caleifosina Hs.26685 AA858390
<b>930</b>	ADNc de Homo sapiens FLJ10976 fis, clon PLACE 1001399 Hs.296323 R27711:
<b>931</b>	C6: componente del complemento 6 Hs.1282 N59396
<b>932</b>	UBE2D3: enzima de conjugación con ubiquitina E2D 3 (homólogo de UBC4/5 de levadura) Hs.118797 AA465196
<b>933</b>	DDX8: Polipéptido de caja DEAD/H (Asp-Glu-Ala-Asp/His) 8 (ARN helicasa) Hs.171872 AA465387 ARN helicasa (HRH1)
<b>934</b>	DKFZP434B168: proteína DKFZP434B168 Hs.48604 N62684
<b>935</b>	FLJ10512: proteína hipotética FLJ10512 Hs.93581 T39933
<b>936</b>	ARNm de Homo sapiens; ADNc DKFZp564F093 (del clon DKFZp564F093) Hs.18724 W87709:
<b>937</b>	F8A: factor de coagulación asociado a VUI (transcrito intrónico) Hs.83363 AA463924
<b>938</b>	HSU53209: transformador-2 alfa (htra-2 alfa) Hs.24937 AA465172
<b>939</b>	UBQLN2: ubiquilina 2 Hs.4552 R43580
<b>940</b>	EIF2C2: factor 2C de iniciación de la traducción eucariótica, 2 Hs.193053 N93082
<b>941</b>	ARNm de Homo sapiens para la proteína FLJ00012, cds parcial Hs.21051 HI7645:
<b>942</b>	KIAA0841: proteína KIAA0841 Hs.7426 R20299
<b>943</b>	KCNAB2: canal de potasio dependiente de voltaje, subfamilia relacionada con shaker, miembro beta 2 Hs.298184 H14383
<b>944</b>	KIAA1637: coactivador independiente de AF-2 (CIA); proteína KIAA1637 Hs.288140 AA521358
<b>945</b>	ESTs: Hs.27379 H17455
<b>946</b>	FLJ11323: proteína hipotética FLJ11323 Hs.25625 R49707
<b>947</b>	SSP29: proteína ácida rica en leucinas Hs.84264 AA489201
<b>948</b>	ESTs: Hs.69280 AA486011

ES 2 595 410 T3

949	ADAMTS1: similar a desintegrina y metaloproteasa (de tipo reopolisina) con motivo 1 de tipo trombospondina, 1 Hs.8230 AA057170
950	ESTs: Hs.43466 N23889
951	MLLT4: leucemia de linaje mieloides/linfoides o mixto (homólogo de trithorax (Drosophila)); translocado a, 4 Hs.100469 AA010818
952	ESTs: Hs.271034 AA406581
953	LMNB1: lámina B1 Hs.89497 AA983462
954	ADNc de Homo sapiens FLJ13547 fis, clon PLACE1007053 Hs.7984 AA629264:
955	PTMS: paratimosina Hs.171814 R10451
956	H2AFL: familia de la histona H2A, miembro L Hs.28777 AI268551
957	FLJ21603: proteína hipotética FLJ21603 Hs.129691 R72794
958	FLJ13287: proteína hipotética FLJ13287 Hs.53263 AA621725
959	CXCR4: quimiocina (motivo C-X-C), receptor 4 (fusina) Hs.89414 AA479357
960	INSM1: 1 asociado a insulinooma Hs.89584 R38640
961	FREQ: homólogo de frecuencia (Drosophila) Hs.301760 H16821
962	LOG58486: proteína similar a transposasa de Buster derivada de transposón Hsr25726 AA868020
963	SMARCD1: Regulador de cromatina dependiente de actina, asociado a matriz, relacionado con SWI/SNF, subfamilia d, miembro 1 Hs.79335 H91691
964	ESTs: Hs.242998 T96522
965	INADL: proteína de dominio de PDZ (similar a inaD de Drosophila) Hs.321197 AA005153
966	ESTs.: Débilmente similar a p150 putativa [H.sapiens] Hs.37751 AA436174
967	MGC5338: proteína hipotética MGC5338 Hs.99598 H50550
968	W85890:
969	NUCKS: similar a la caseína cinasa 2 ubicua nuclear de rata Hs.118064 AI053436
970	Secuencia del ARNm del clon 25110 de Homo sapiens Hs.27262 HI8031:
971	A1333214:
972	GAS41: secuencia 41 amplificada en el glioma Hs.4029 T62072
973	LOC51170: deshidrogenasa/reductasa de cadena corta de la retina retSDR2 Hs.12150 N79745
974	H2BFG: **familia de la histona H2B, miembro G Hs.182137 R98472
975	ABCC1: **casete de unión a ATP, subfamilia C (CFTR/MRP), miembro 1 Hs.89433 AA424804
976	EFNA1: efrina-A1 Hs.1624 AA857015
977	ARNm de Homo sapiens; ADNc DKFZp434A1014 (del clon DKFZp434A1014); cds parcial Hs.278531 H00596:
978	PPP2CA: proteína fosfatasa 2 (antes 2 A), subunidad catalítica, isoforma alfa Hs.91773 AA599092
979	ESTs.: Débilmente similar al producto génico sin nombre [H.sapiens] Hs.118338 W85843
980	ADNc de Homo sapiens FLJ11643 fis, clon HEMBA1004366 Hs.111496 AA598803:
981	ESTs.: Moderadamente similar a ALUE HUMANA !!!! ADVERTENCIA DE ALU CLASE E!! [H.sapiens] Hs.125407 AA878944
982	ESTs.: Moderadamente similar a la SECUENCIA DE LA SUBFAMILIA J DE ALU1_HUMANA ALU ADVERTENCIA DE CONTAMINACIÓN [H.sapiens] Hs.144662 AI191290
983	KIAA0916: proteína KIAA0916 Hs.151411 R91388

ES 2 595 410 T3

<b>984</b>	CDC25A: ciclo de división celular 25ª Hs.1634 R09062
<b>985</b>	PRIM2A: primasa, polipéptido 2A (58kD) Hs.74519 R61073
<b>986</b>	DSP: desmoplaquina (DPI, DPII) Hs.74316 H90899
<b>987</b>	KIAA010I: producto génico de KIAA0101 Hs.81892 W68219
<b>988</b>	ESTs,: Débilmente similar a p150 putativa [H.sapiens] Hs.268026 AA411454
<b>989</b>	ESTs: Hs. 18140 T97707
<b>990</b>	H2AFL: Familia de la histona H2A, miembro L Hs.28777 AA457566
<b>991</b>	ARNm de Homo sapiens para la proteína KIAA1700, cds parcial Hs.20281 H00287:
<b>992</b>	STAG3: antígeno estromal 3 Hs.20132 AA453028
<b>993</b>	ZNF207: proteína de dedo de cinc 207 Hs.62112 N59119
<b>994</b>	BMP6: proteína morfogenética ósea 6 Hs.285671 AA424833
<b>995</b>	ESTs,: Moderadamente similar a la sertolina [R.norvegicus] Hs.91192 H60690
<b>996</b>	LOC51064:**homólogo de la subunidad 13 de la glutatión-S-transferasa Hs.279952 W88497
<b>997</b>	NUCKS: similar a la caseína cinasa 2 ubicua nuclear de rata Hs.118064 AA927182
<b>998</b>	ESTs,: Débilmente similar a la proteína hipotética T00370 KIAA0659 [H.sapiens] Hs.131899 W93155
<b>999</b>	FLJ13057: proteína hipotética FLJ13057 similar a la de células germinales less Hs.243122 R23254
<b>1000</b>	ESTs: Hs.144796 AI219737
<b>1001</b>	FLJ10511: proteína hipotética FLJ10511 Hs.106768 R25877
<b>1002</b>	DKFZP564A122: proteína DKFZP564A122 Hs.187991 N31577
<b>1003</b>	ODF2: fibra densa externa de colas de espermatozoides 2 Hs. 129055 AA400407
<b>1004</b>	AMY2A: amilasa, alfa 2A; pancreática Hs.278399 R64129
<b>1005</b>	**ESTs,: Débilmente similar a la placofilina 2b [H.sapiens] Hs.12705 N91589
<b>1006</b>	CYP1B1: citocromo P450, subfamilia I (inducible con dioxina), polipéptido 1 (glaucoma 3 primario infantil) Hs.154654 AA029776
<b>1007</b>	CAPN7: calpaína 7 Hs.7145 N46420
<b>1008</b>	FLJ20069: proteína hipotética FLJ20069 Hs.273294 AA229966
<b>1009</b>	FLJ10618: proteína hipotética FLJ10618 Hs.42484 AA478847
<b>1010</b>	KIAA1637: **coactivador independiente de AF-2 (CIA); proteína KIAA1637 Hs.288140 AA452531
<b>1011</b>	FLJ20004: **proteína hipotética FLJ20004 Hs.17311 AA487895
<b>1012</b>	FLJ12892: proteína hipotética FLJ12892 Hs.17731 AA670363
<b>1013</b>	PLU-1: motivo putativo de unión a ADN/cromatina Hs.143323 AA464869
<b>1014</b>	**ESTs: Hs.36828 AA418448
<b>1015</b>	KIAA0586: producto génico de KIAA0586 Hs.77724 AA905278
<b>1016</b>	MTHFD2: metilentetrahidrofolato deshidrogenasa (dependiente de NAD+), metilentetrahidrofolato ciclohidrolasa Hs.154672 AA480994
<b>1017</b>	BRF1: **factor 1 de respuesta al butirato (factor 1 de respuesta a EGF) Hs.85155 AA424743
<b>1018</b>	TFAP2A: factor de transcripción AP-2 alfa (proteína 2 alfa de unión al potenciador de la activación) Hs.18387 R38044

ES 2 595 410 T3

<b>1019</b>	VIL2: villina 2 (ezrina) Hs.155191 AA411440
<b>1020</b>	SDC1: syndecan 1 Hs.82109 AA074511
<b>1021</b>	RNTRE: relacionada con el extremo N de tre Hs.278526 AA281057
<b>1022</b>	HSPC207: proteína hipotética Hs.75798 H99997
<b>1023</b>	FLJ22376: proteína hipotética FLJ22376 Hs.29341 AI199155
<b>1024</b>	RNF10: proteína de dedo anular 10 Hs.5094 H73586
<b>1025</b>	PNN: pinina, proteína asociada al desmosoma Hs.44499 AA707321
<b>1026</b>	FLJ20516: proteína hipotética FLJ20516 Hs.70811 AA122393
<b>1027</b>	RPL13A: proteína ribosómica LI3a Hs.119122 AI254200
<b>1028</b>	H2BFB: familia de la histona H2B, miembro B Hs.I80779 AA885642
<b>1029</b>	OGT: N-acetilglucosamina (GlcNAc) transferasa unida a O (UDP-N- acetilglucosamina:polipéptido-N-acetilglucosaminil transferasa) Hs.100293 R13317
<b>1030</b>	KIAA0155: producto génico de KIAA0155 Hs.173288 AA433684
<b>1031</b>	ILF2: factor 2 de unión al potenciador de la interleucina, 45kD Hs.75117 H95638
<b>1032</b>	ARNm de Homo sapiens; ADNc DKFZp586l1518 (del clon DKFZp58611518) Hs.21739 AA287917:
<b>1033</b>	PKNOX1: homeocaja 1 de PBX/knotted 1 Hs.158225 AI350546
<b>1034</b>	KMO: **cinurenina 3-monooxigenasa (cinurenina 3-hidroxilasa) Hs.107318 AA044326
<b>1035</b>	VCAM1: molécula 1 de adhesión celular vascular Hs.109225 H16591 CD106=VCAM-1
<b>1036</b>	N54811:
<b>1037</b>	KIAA0618: producto génico de KIAA0618 Hs.295112 H81940
<b>1038</b>	MAFG: familia oncogénica del fibrosarcoma musculoaponeurótico v-maf (aviar), proteína G Hs.252229 N21609 MafG= factor de transcripción de la cremallera de leucina-básica
<b>1039</b>	MATN2: matrilina 2 Hs.19368 AA071473
<b>1040</b>	HOXB4: homeocaja B4 Hs.126666 AA918749
<b>1041</b>	FLJ10466: proteína hipotética FLJ10466 Hs.121073 AA453607
<b>1042</b>	FLJ22557: proteína hipotética FLJ22557 Hs.106101 AA127879
<b>1043</b>	EST: Hs.149260 AI247680
<b>1044</b>	KIAA0677: producto génico de KIAA0677 Hs.155983 AA026751
<b>1045</b>	EST: Hs.104123 AA197344
<b>1046</b>	UCP4: proteína de desacoplamiento 4 Hs.40510 H94680
<b>1047</b>	EST: Hs.144224 N93807
<b>1048</b>	GATA2: proteína 2 de unión a GATA Hs.760 H00625 proteína 2 de unión a GATA
<b>1049</b>	ESTs: Hs.14743 H61082
<b>1050</b>	EST: Hs.116174 AA626786
<b>1051</b>	ITGB3: integrina, beta 3 (glicoproteína IIIa de plaquetas, antígeno CD61) Hs.87149 AA666269
<b>1052</b>	FLJ23399: proteína hipotética FLJ23399 Hs.299883 R19895
<b>1053</b>	ESTs: Hs.21734 N72976

ES 2 595 410 T3

<b>1054</b>	FLJ20425: proteína hipotética FLJ20425 Hs.71040 AA424566
<b>1055</b>	CUL4A: cullina 4A Hs.I83874 AA598836
<b>1056</b>	PTP4A1: proteína tirosina fosfatasa de tipo IVA, miembro 1 Hs.227777 proteína R61007 tirosina fosfatasa PTPCAAX1 (hPTPCAAXI)
<b>1057</b>	ESTs: Hs.7913 N35592
<b>1058</b>	GROI: oncogén GROI (actividad estimulante del crecimiento del melanoma, alfa) Hs.789 W46900
<b>1059</b>	ESTs,: Moderadamente similar a la NRD2 convertasa [H.sapiens] Hs.309734 H78796
<b>1060</b>	FLJ10826: proteína hipotética FLJ10826 Hs.24809 AA486738
<b>1061</b>	TOM34: translocasa de la membrana mitocondrial externa 34 Hs.76927 AA457118
<b>1062</b>	H2AFL: familia de la histona H2A, miembro L Hs.28777 AA452933
<b>1063</b>	D10S170: **segmento de ADN,copia única, sonda pH4 (secuencia transformante, tiroidea-1, Hs.315591 N35493
<b>1064</b>	SCYA2: citocina inducible pequeña A2 (proteína quimiotáctica monocítica 1, homóloga a Sig-je de ratón) Hs.303649 T77816 MCP-1=MCAF=citocina inducible pequeña A2=JE=quimiocina
<b>1065</b>	FLJ10688: proteína hipotética FLJ10688 Hs.118793 AA465358
<b>1066</b>	PTD017: proteína PTD017 s.274417 AA160498
<b>1067</b>	KIAA0026: gen X relacionado con MORF Hs.173714 AA676604
<b>1068</b>	BMP2: proteína morfogenética ósea 2 Hs.73853 AA489383
<b>1069</b>	MNT: proteína de unión a MAX Hs.25497 AA455508
<b>1070</b>	KIAA1170: proteína K1AA1170 Hs.268044 H80507
<b>1071</b>	CRYBA1: cristalina, beta A1 Hs.46275 AA487614
<b>1072</b>	KATNA1: subunidad A 1 de katanina p60 (que contiene ATPasa) Hs.289099 AA609740
<b>1073</b>	ADNc de Homo sapiens FLJ20796 fis, clon COL00301 Hs.113994 N53458:
<b>1074</b>	CEP4: proteína efectora de Cdc42 4; enlazador de Rho GTPasas 4 Hs.3903 W32509
<b>1075</b>	ESTs: Hs.117261 AA682521
<b>1076</b>	CYP1B1: citocromo P450, subfamilia 1 (inducible por dioxina), polipéptido 1 (glaucoma 3, primario infantil) Hs.154654 AA040872
<b>1077</b>	ALTE: elemento de transposón similar a Ac Hs.9933 AA630498
<b>1078</b>	RAD51: homólogo de RAD51 (S. cerevisiae) (homólogo de RecA de E. coli) Hs.23044 AA873056
<b>1079</b>	MAN1A2: mannosidasa, alfa, clase 1 A, miembro 2 Hs.239114 R78501
<b>1080</b>	H53763:
<b>1081</b>	MET: proto-oncogén met(receptor del factor de crecimiento de los hepatocitos) Hs.285754 AA410591
<b>1082</b>	DYRK1 A: cinasa 1A regulada por (Y)-fosforilación de tirosina de especificidad dual Hs.75842 AA676749
<b>1083</b>	ARHGAP8: **proteína 8 de activación de Rho GTPasa Hs.102336 AA037410
<b>1084</b>	LMO4: dominio LIM únicamente 4 Hs.3844 H27986
<b>1085</b>	ADCY6: adenilato ciclase 6 Hs. 12373 AA148044
<b>1086</b>	EST: Hs.135448 AI078552
<b>1087</b>	NCOA3: coactivador 3 del receptor nuclear Hs.225977 W46433
<b>1088</b>	DNAJB4: homólogo de DnaJ (Hsp40), subfamilia B, miembro 4 Hs.41693 AA081471

ES 2 595 410 T3

<b>1089</b>	NAB1: Proteína de unión a NGFI-A 1 (proteína de unión 1 a ERG1) Hs.107474 AA486027
<b>1090</b>	ESTs.: Débilmente similar a la proteína hipotética T08663 DKFZp547G0910.1 [H.sapiens] Hs.172084 N63646
<b>1091</b>	KIAA0735: producto génico de KIAA0735; homólogo de la proteína 2B de la vesícula sináptica Hs.8071 R56082
<b>1092</b>	GNB1: proteína de unión al nucleótido guanina (proteína G), beta polipéptido 1 Hs.215595 AA487912
<b>1093</b>	ARNm de Homo sapiens para la proteína KJAA1716, cds parcial Hs.21446 R49763:
<b>1094</b>	KINESINA: CADENA PESADA
<b>1095</b>	CCND1: ciclina-D4-(PRAD1: adenomatosis paratiroidea 1) Hs.82932 AA487486 Ciclina D1=BCL1=PRAD1=Translocada en la leucemia de las células del manto
<b>1096</b>	ESTs: Hs.106129 R56716
<b>1097</b>	AA431931:
<b>1098</b>	PSEN1: presenilina 1 (enfermedad de Alzheimer 3) Hs.3260 AA403083
<b>1099</b>	ESTs: Hs.193804 AA010918
<b>1100</b>	DKFZp762P2111: proteína hipotética DKFZp762P2111 Hs.14217 AA429586
<b>1101</b>	KIAA1350: proteína KIAA1350 Hs.101799 W37627
<b>1102</b>	FLJ20847: proteína hipotética FLJ20847 Hs.13479 HI6996
<b>1103</b>	HDCMA18P: proteína HDCMA18P Hs.278635 N64387
<b>1104</b>	FLJ12890: proteína hipotética FLJ12890 Hs.43299 N62475
<b>1105</b>	ESTs: Hs.127453 AA973625
<b>1106</b>	BAIAP2: proteína 2 asociada con BAI1 Hs.7936 R60328
<b>1107</b>	ESTs: Hs.317584 AA191424
<b>1108</b>	DKFZP434J046: proteína DKFZP434J046 Hs.116244 AI024401
<b>1109</b>	ESTs: Hs.114055 AA701352
<b>1110</b>	ESTs: Hs.44380 N93122
<b>1111</b>	ESTs: Hs.20142 AA625570
<b>1112</b>	UBL3: similar a ubiquitina 3 Hs.173091 T82438
<b>1113</b>	H2AFL: Familia de la histona H2A, miembro L Hs.147097 N50797
<b>1114</b>	SUCLG2: **succinato-CoA ligasa, formador de ADP, subunidad beta Hs.247309 N68557
<b>1115</b>	ZWINT: ZW10 de interacción Hs.42650 AA706968
<b>1116</b>	FLJ10583: proteína hipotética FLJ10583 Hs.105633 R00425
<b>1117</b>	FLJ20552: proteína hipotética FLJ20552 Hs.69554 AA463982
<b>1118</b>	FADD: Dominio de muerte por vía asociada a Fas (TNFRSF6) Hs.86131 AA430751 FADD=MORT
<b>1119</b>	SFRS7: factor de corte y empalme, rico en arginina / serina 7 Hs.184167 AA418813
<b>1120</b>	RAD54L similar a RAD54 (S.cerevisiae) Hs.66718 AI372035
<b>1121</b>	MYLE: proteína MYLE Hs.11902 T68845
<b>1122</b>	LOC51334: proteína de células madre mesenquimales DSC54 Hs.157461 R63841
<b>1123</b>	PRIM2A: primasa, polipéptido 2A (58kD) Hs.74519 AA434404
<b>1124</b>	KIAA0056: proteína KIAA0056 Hs.13421 AA430545

ES 2 595 410 T3

1125	ESTs.: moderadamente similar a la SECUENCIA SQ DE LA SUBFAMILIA DE ALU ALU7 HUMANA ADVERTENCIA DE CONTAMINACIÓN [H.sapiens] Hs.82590 N53024
1126	ESTs: Hs.117269 AA705050
1127	NSAP1: proteína 1 asociada con NS1 Hs.155489 AA186327
1128	CEACAM5: molécula de adhesión celular 5 relacionada con el antígeno carcinoembrionario Hs.220529 AA130547
1129	FLJ11021: proteína hipotética FLJ11021 similar al factor de corte y empalme, rico en arginina/serina 4 Hs.81648 AA291183 desconocida UG Hs.202583 ESTs, débilmente similar al rico en arginina
1130	FOSL1: antígeno 1 similar a FOS Hs.283565 T82817 fra-1=antígeno 1 relacionado con fos
1131	U3-55K: proteína de 55-kDa asociada a U3 snoRNP Hs.153768 AA465355
1132	DNAJC6: Homólogo de DnaJ (Hsp40), subfamilia B, miembro 6 Hs. 44896 AA455940
1133	KIAA1382: transportador de aminoácidos 2 Hs.298275 R27255 Similar a la proteína transportadora
1134	PCAF: factor asociado con p300/CBP Hs.199061 N74637 P/CAF= factor asociado con p300/CBP
1135	ESTs: Hs.130460 AA927252
1136	ESTs: Hs.112570 AI014667
1137	FLJ10209: proteína hipotética FLJ10209 Hs.260150 AA454626
1138	ESTs: Hs.99014 AA485679
1139	ESTs: Hs.99621 AA464707
1140	ADNc de Homo sapiens FLJ14028 fis, clon HEMBB1000048 Hs. 285519 N74617:
1141	AA928536:
1142	SQSTM1: **secuestrosoma 1 Hs.182248 AA931964
1143	**ADNc de Homo sapiens FLJ13700 fis, clon PLACE2000216, altamente similar a la CADENA DE LA ESPECTRINA BETA, CEREBRO Hs.324648 AA018591:
1144	SLC22A3: familia portadora de solutos 22 (transportador de monoamina extraneuronal), miembro 3 Hs.81086 AA460012
1145	FLJ22557: proteína hipotética FLJ22557 Hs.106101 H00595
1146	FLJ20539: proteína hipotética FLJ20539 Hs.118552 R36152
1147	AA991624:
1148	TRAP 150: proteína asociada al receptor de hormona tiroidea, subunidad de 150 kDa Hs.108319 W85832
1149	ESTs: Hs.221847 R91557
1150	TCFL1: similar al factor de transcripción 1 Hs.2430 AA443950
1151	ESTs.: altamente similar a la variante de corte y empalme de oxitocinasa 1 [H.sapiens] Hs.203271 AA487918
1152	PLAB: factor de diferenciación de la próstata Hs.296638 AA450062
1153	RBM14: proteína motivo de unión ARN 14 Hs.11170 AA417283
1154	EGFL5: dominio similar a EGF, múltiple 5 Hs.5599 W67981
1155	H2AFO: Familia de la histona H2A, miembro O 795 AA047260
1156	ESTs.: débilmente similar a A46661 leucotrieno B4 omega-hidroxilasa [H.sapiens] Hs.169001 N45556
1157	W78784:
1158	TOP3A: topoisomerasa (DNA) III alfa Hs.91175 N21546
1159	W73732: factor-1 de la célula huésped=VP 16 proteína de interacción con el transactivador

ES 2 595 410 T3

<b>1160</b>	CYP1B1: citocromo P450, subfamilia I (inducible por dioxina), polipéptido 1 (glaucoma 3, infantil primario) Hs.154654 AA448157 CitocromoP450 IB1 (inducible por dioxina)
<b>1161</b>	ESTs: Hs.135276 AI092102
<b>1162</b>	RHEB2: homólogo Ras enriquecido en el cerebro 2 Hs.279903 AA482117
<b>1163</b>	ESTs,: altamente similar a EF-9-[M-musculus] Hs.8366 H94467
<b>1164</b>	POLA: polimerasa (dirigida por ADN), alfa Hs.267289 AA707650
<b>1165</b>	KIAA1008: proteína KIAA1008 Hs.323346 AA863115
<b>1166</b>	PIK3CD: fosfoinosítido-3-cinasa, polipéptido delta catalítico Hs.162808 AA281652
<b>1167</b>	T53625:
<b>1168</b>	**ARNm de Homo sapiens; ADNc DKFZp566PI 114 (del clon DKFZp434A1114) Hs.326292 AA417274:
<b>1169</b>	ESTs: Hs.26744 H16988
<b>1170</b>	FLJ13912: proteína hipotética FLJ13912 Hs.47125 W74133
<b>1171</b>	ARNm de Homo sapiens; ADNc DKFZp762B195 (del clon DKFZp762B195) Hs.284158 AA625574:
<b>1172</b>	SSA2: antígeno del síndrome de Sjogren A2 (60kD, autoantígeno de ribonucleoproteína SS-A/Ro) Hs.554 AA010351
<b>1173</b>	BK1048E9.5: proteína hipotética bK1048E9.5 Hs.6657 N68512
<b>1174</b>	TOP1: topoisomerasa (ADN) I Hs.317 AA232856 Topoisomerasa I
<b>1175</b>	ESTs: Hs.15386 HI8472
<b>1176</b>	KPNB1: carioferina (importina) beta 1 Hs.180446 AA121732
<b>1177</b>	MGC861: proteína hipotética MGC861 Hs.208912 N69694
<b>1178</b>	PMS2L8: **similar a 2 de segregación posmeiótica 8 Hs.323954 T62577
<b>1179</b>	TSC22: * proteína estimulada por factor de crecimiento transformante beta TSC-22 Hs.114360 R16390
<b>1180</b>	C80RF1: marco de lectura abierto 1 del cromosoma 8 Hs.4055 AA278836
<b>1181</b>	ESTs: Hs.129165 AA989211
<b>1182</b>	DMTF: factor de transcripción similar a Myb de unión a ciclina D 1 Hs.5671 AA129860
<b>1183</b>	CDC7L1: CDC7 similar a (ciclo de división celular 7, S. Cerevisiae, homólogo)1 Hs.28853 N62245 cinasa relacionada con Cdc7
<b>1184</b>	LOC51700: citocromo b5 reductasa b5R.2 Hs.22142 AA425316
<b>1185</b>	FLNA: filamina A, alfa (proteína de unión a actina 280) Hs.115464 AA598978
<b>1186</b>	FLJ20257: proteína hipotética FLJ20257 Hs.178011 H78675
<b>1187</b>	ADNc de Homo sapiens FLJ13604 fis, clon PLACE1010401 Hs.23193 AA406599:
<b>1188</b>	ESTs: Hs.205227 R73480
<b>1189</b>	SCYB14: subfamilia B de citocinas pequeñas inducibles (Cys-X-Cys), miembro 14 (BRAK) Hs.24395 AA953842
<b>1190</b>	MAPK8IP2: **proteína de interacción con la proteína cinasa activada por mitógeno 2 Hs.80545 AA418293
<b>1191</b>	ZNF42: proteína de dedo de cinc 42 (respondedor a ácido retinoico específico de mieloides) Hs.169832 AA932642
<b>1192</b>	ESTs: Hs.127054 AA862450
<b>1193</b>	NUDT4: Motivo 4 de tipo nudix (resto X unido al nucleósido difosfato) Hs.92381 AA425630
<b>1194</b>	ADNc de Homo sapiens FLJ10632fis, clon NT2RP2005637 Hs.202596 H82421:
<b>1195</b>	LOC51042: proteína de dedo de cinc Hs.102419 AA033532
<b>1196</b>	NUMA1: proteína 1 del aparato mitótico nuclear Hs.301512 AA679293
<b>1197</b>	ESTs,: altamente similar a A56429 I-proteína relacionada con kappa-B [H.sapiens] Hs.144614 AA293771

ES 2 595 410 T3

1198	ESTs: Hs.127703 AA947258
1199	ADNc de Homo sapiens FLJ14214 fis, clon NT2RP3003576 Hs.321236 AA903913:
1200	NFKBIA: inhibidor del factor nuclear del potenciador del gen del polipéptido ligero kappa en células B, alfa Hs.81328 W55872 I kB alfa
1201	ESTs: Hs.120029 AA707598
1202	ESTs,: Moderadamente similar a la cadena A A, Glucosamina-6-fosfato desaminasa isomerasa humana a 1,75 A [H.sapiens] Hs.21398 AA172012
1203	NFIA: factor nuclear I/A Hs.I73933 AI912047
1204	RECQL4: proteína similar a RecQ 4 Hs.31442 AA620446
1205	**ESTs,: débilmente similar a la SECUENCIA J DE LA SUBFAMILIA DE ALU HUMANA ALU1 ADVERTENCIA DE CONTAMINACIÓN [H.sapiens] Hs.318894 R96212
1206	ADNc de Homo sapiens: FLJ21686 fis, clon COL09379 Hs.20787 R11371:
1207	LOC57168: similar a la aspartato beta hidroxilasa (ASPH) Hs.184390 HI7272
1208	ESTs: Hs.26096 R54109
1209	ARNm de proteína 6 relacionada con OSBP de Homo sapiens, cds completa Hs.318775 AA680281:
1210	APACD: proteína de unión a ATP asociada con diferenciación celular Hs.153884 N80741
1211	VIM: **vimentina Hs.297753 AI668662
1212	ADNc de Homo sapiens FLJ13618 fis, clon PLACE1010925 Hs.17448 AA427980:
1213	NR3C1: subfamilia 3 del receptor nuclear, grupo C, miembro 1 Hs.75772 N30428 receptor de glucocorticoides
1214	ADNc de Homo sapiens: FLJ21814 fis, clon HEP01068 Hs.289008 R12808:
1215	BRD7: que contiene bromodominio 7 Hs.279762 AA488428
1216	MAP3K8: **proteína cinasa activada por mitógeno 8 Hs.248 W42450
1217	ESTs: Hs.23213 H29336
1218	ESTs: Hs.122444 AA939019
1219	TUSP: proteína de la superfamilia tubby Hs.102237 H78234
1220	KIAA1117: proteína KJAA1117 Hs.278398 H01516
1221	Ser humano: ARNm del clon 137308, cds parcial Hs.322149 H91303
1222	ESTs: Hs.130214 AA456631
1223	RAB3A: RAN, miembro de familia del oncogén RAS Hs.27744 HI4230
1224	AA598795: Proteína fosfatasa 2 (antes 2A), subunidad B reguladora (P
1225	H2BFC: Familia de las histonas H2B, miembro C Hs.137594 AI340654
1226	CFLAR: regulador de la apoptosis similar a CASP8 y FADD Hs.195175 N94588
1227	CD24: Antígeno CD24 (antígeno 4 del clúster de carcinoma de pulmón de célula pequeña) Hs.286124 H59915
1228	EST: Hs.48532 N62402
1229	CCRK: cinasa relacionada con el ciclo celular Hs.26322 H17616
1230	HECH: proteína 1 similar a la heterocromatina 1 Hs.278554 AI139106
1231	-DKFZp5470-j46- proteína hipotética-DKFZp5470146-Hsr91246-T80848
1232	ESTs: Hs.71574 AA135328

ES 2 595 410 T3

<b>1233</b>	HLXB9: homeocaja HB9 Hs.37035 AI459915
<b>1234</b>	AA600222:
<b>1235</b>	SPINK5: inhibidor de serina proteasa, de tipo Kazal, 5 Hs.5476 W92134
<b>1236</b>	RNUT1: ARN, transportador U 1 Hs.21577 AA447799
<b>1237</b>	ADNc de Homo sapiens: FLJ23013 fis, clon LNG00740 Hs.13075 AA464543:
<b>1238</b>	K1AA0063: KIAA0063 producto génico Hs.3094 T82263
<b>1239</b>	DYRK2: cinasa 2 regulada por la fosforilación de tirosina (Y) de especificidad dual Hs.173135 R63622
<b>1240</b>	R94947:
<b>1241</b>	ADNc de Homo sapiens FLJ14337 fis, clon PLACE4000494 Hs.180187 AA004903:
<b>1242</b>	FLJ20624: proteína hipotética FLJ20624 Hs.52256 AA431909
<b>1243</b>	ESTs: Hs.43838 R38261
<b>1244</b>	FLJ23053: proteína hipotética FLJ23053 Hs.94037 R25654
<b>1245</b>	MGC11266: proteína hipotética MGC11266 Hs.293943 AA400456
<b>1246</b>	ESTs,: Moderadamente similar a la SECUENCIA SX DE LA SUBFAMILIA DE ALU HUMANA ALU1 ADVERTENCIA DE CONTAMINACIÓN [H.sapiens] Hs.34174 AA126603
<b>1247</b>	PLAUR: activador del plasminógeno, receptor de la uroquinasa Hs.179657 AA147962
<b>1248</b>	TSG101: gen de susceptibilidad tumoral 101 Hs.118910 AA670215
<b>1249</b>	HCNGP: proteína reguladora de la transcripción Hs.27299 AA457232
<b>1250</b>	KIAA0978: proteína K1AA0978 Hs. 3686 AA857017
<b>1251</b>	ESTs: Hs.61708 AA033867
<b>1252</b>	ESTs: Hs. 120734 AA827482
<b>1253</b>	ESTs: Hs.5909 AA972654
<b>1254</b>	CDH24: similar a cadherina 24 Hs.155912 AI668564
<b>1255</b>	CCND1: ciclina D1 (PRAD1: adenomatosis paratiroidea 1) Hs.82932 T77237
<b>1256</b>	ESTs: Hs.43148 AA284775
<b>1257</b>	ESTs: Hs.222566 T50982
<b>1258</b>	ESTs: Hs. 194125-N52822
<b>1259</b>	EST: Hs.I54621 AI138644
<b>1260</b>	MAN1A2: manosidasa, alfa, clase 1A, miembro 2 Hs.239114 R22905
<b>1261</b>	MAN2A2: manosidasa, alfa, clase 2A, miembro 2 Hs. 295605 AA454175
<b>1262</b>	Secuencia de ADN humano del clon 967N21 en el cromosoma 20p12.3-13. Contiene el gen de CHGB para la cromogranina B (secretogranina 1, SCG1), un pseudogen similar a la parte de K1AA0172, el gen para una nueva proteína Hs.88959 R56678
<b>1263</b>	ESTs,: Altamente similar a la PROTEÍNA KV3.4 DE LOS CANALES DE POTASIO DEPENDIENTES DE VOLTAJE CIKG [H.sapiens] Hs.106486 H11376
<b>1264</b>	ARNm de HT023de Homo sapiens, cds completa Hs.237225 AA169496:
<b>1265</b>	FLJ10339: **proteína hipotética FLJ10339 Hs.203963 H72354
<b>1266</b>	N66278:
<b>1267</b>	ESTs: Hs.6195 AA454745

ES 2 595 410 T3

<b>1268</b>	KIAA1404: proteína KIAA1404 Hs.200317 W72798
<b>1269</b>	PMAIP1: proteína 1 inducida por forbol-12-miristato-13-acetato 1 Hs.96 AA458838 APR=gen de respuesta inmediata-temprana= respondedor a PMA derivado de ATL
<b>1270</b>	G3BP: Proteína de unión al dominio SH3 de la proteína de activación de Ras-GTPasa Hs.220689 AA598628
<b>1271</b>	ADNc de Homo sapiens: FLJ22807 fis, clon KAIA2887 Hs.261734 R26854:
<b>1272</b>	Homo sapiens, clon IMAGE:3535294, ARNm, cds parcial Hs.80449 T57359:
<b>1273</b>	CDC16: CDC6 (ciclo de división celular 16, S. cerevisiae, homólogo) Hs.1592 AA410559
<b>1274</b>	FGA: **fibrinógeno, alfa-polipéptido A Hs.90765 AA026626
<b>1275</b>	ESTs: Hs.33446 N53560
<b>1276</b>	ADNc de Homo sapiens FLJ14175 fis, clon NT2RP2002979 Hs.288613 AA054704:1:
<b>1277</b>	ESTs: Hs.44243 AA011390
<b>1278</b>	Clon del ADNc del inserto de longitud completa del ARNm de Homo sapiens EUROIMAGE 42408 Hs.284123 R61732:
<b>1279</b>	ESTs: Hs.53455 AA454165
<b>1280</b>	FLJ11264: proteína hipotética FLJ11264 Hs.11260 AI219094
<b>1281</b>	MBD4: proteína 4 del dominio de unión a metil-CpG Hs.35947 AA010492
<b>1282</b>	FLJ11305: proteína hipotética FLJ11305 Hs.7049 N94612
<b>1283</b>	Homo sapiens, similar al producto génico CG5057, clon MGC:5309,ARNm, cds completa Hs.13885 AA460004:
<b>1284</b>	ARHB: familia génica del homólogo de ras, miembro B Hs.204354 H88963
<b>1285</b>	ITPR3: receptor de 1,4,5 Inositol trifosfato, tipo 3, Hs.77515 AA865667
<b>1286</b>	HMG20B: grupo de alta movilidad 20B Hs.32317 AA775743
<b>1287</b>	ESTs: Hs.146276 AI214204
<b>1288</b>	PTPN9: proteína tirosina fosfatasa, no receptora de tipo 9 Hs. 147663 AA434420
<b>1289</b>	ARNm del clon de Homo Sapiens FLB9213 PR02474 , cds completa Hs.21321 AA486770:
<b>1290</b>	H21107:
<b>1291</b>	HSPC157: proteína HSPC157 Hs.279842 N20480
<b>1292</b>	ARNm de Homo sapiens; ADNc de DKFZp56402363 (a partir del clon DKFZp56402363) Hs.321403 AA406332:
<b>1293</b>	ESTs: Hs. 150623 AA693532
<b>1294</b>	EST: Hs.188697 AA199733
<b>1295</b>	CLECSF2: lectina de tipo C (dominio de reconocimiento de carbohidratos, dependiente de calcio), miembro 2 de la superfamilia (inducida por activación) Hs.85201 H11732 AICL= lectina de tipo C inducida por activación
<b>1296</b>	ITPR1: receptor de 1,4,5 Inositol trifosfato, tipo 1, Hs.198443 AA035450
<b>1297</b>	CHML: similar a coroideremia (proteína 2 escort Rab) Hs.170129 R91881
<b>1298</b>	CDC42:ciclo de división celular 42 (proteína de unión a GTP, 25kD) Hs. 146409 AA668681
<b>1299</b>	FKBP5: **proteína de unión a FK506 5 Hs.7557 AA872767

Todas las publicaciones y solicitudes de patentes mencionadas en la memoria descriptiva son indicativas del nivel de los expertos en la técnica a la que pertenece la presente invención.

La mera mención de las publicaciones y solicitudes de patente no constituye necesariamente una admisión de que son una técnica anterior a la presente solicitud.

Aunque se ha descrito la invención anterior con algo de detalle a modo de ilustración y ejemplo por fines de claridad de entendimiento, será obvio que se pueden practicar determinados cambios y modificaciones dentro del alcance de las reivindicaciones adjuntas.

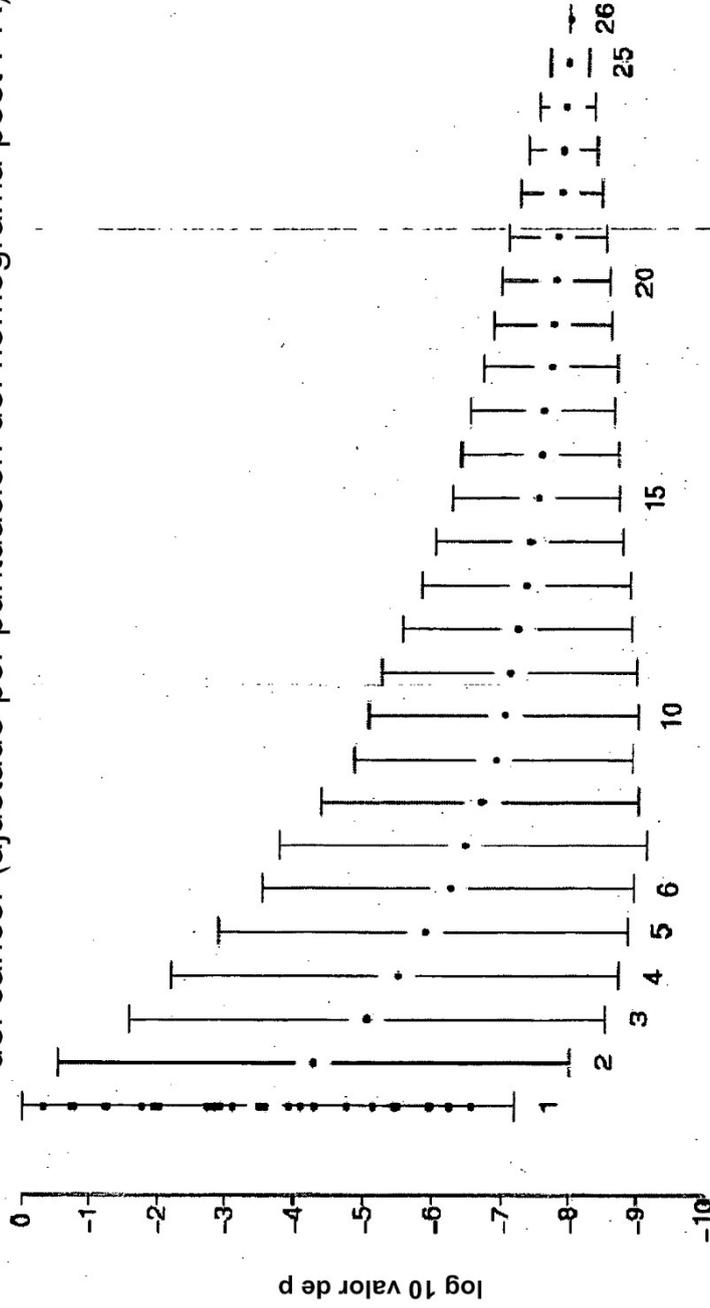
5

**REIVINDICACIONES**

1. Un procedimiento de clasificación del cáncer de próstata que comprende determinar el nivel de expresión de un panel de genes que comprende todos los genes del ciclo celular enumerados en uno cualquiera de los paneles C a G, en el que el aumento de expresión de dichos genes del ciclo celular indica un mal pronóstico.
- 5 2. Un procedimiento de clasificación del cáncer de próstata que comprende determinar al menos un parámetro clínico para un paciente con cáncer y determinar el nivel de expresión de todos los genes del ciclo celular enumerados en uno cualquiera de los paneles C a G en una muestra obtenida del paciente, en el que un parámetro clínico asociado con recurrencia o un aumento de la expresión de dichos genes del ciclo celular indica un mal pronóstico.
- 10 3. El procedimiento de la reivindicación 2, que comprende además la determinación de dos o más parámetros clínicos, en el que dichos dos o más parámetros clínicos forman o son parte de un nomograma que comprende cualquiera los parámetros clínicos del nomograma de Kattan-Stephenson o los parámetros clínicos del índice pronóstico de Nottingham.
- 15 4. El procedimiento de la reivindicación 1, que comprende además la determinación del nivel de expresión de PTEN en una muestra obtenida del paciente, en el que un estado de PTEN negativo o bajo indica un mal pronóstico.
5. El procedimiento de la reivindicación 2, que comprende además la determinación del nivel de expresión de PTEN en una muestra obtenida del paciente, en el que un estado de PTEN negativo o bajo indica un mal pronóstico.
- 20 6. Un procedimiento de determinación de un curso de tratamiento para un paciente con cáncer de próstata que comprende la determinación del nivel de expresión de un panel de genes que comprende todos los genes del ciclo celular enumerados en uno cualquiera de los paneles C a G en una muestra obtenida del paciente, y (a) recomendar, prescribir o iniciar el tratamiento activo si dicha muestra ha incrementado la expresión de dichos genes del ciclo celular, o (b) recomendar, prescribir o iniciar una espera vigilante si dicha muestra no ha incrementado la expresión de dichos genes del ciclo celular.
- 25 7. Un procedimiento para determinar si un paciente con cáncer de próstata tiene un aumento de la probabilidad de recurrencia, que comprende:
- (a) determinar el nivel de expresión de un panel de genes que comprende todos los genes del ciclo celular enumerados en uno cualquiera de los paneles C a G en una muestra de un paciente;
- (b) acceder a la información obtenida en (a) en un medio legible por ordenador;
- 30 (c) consultar esta información para determinar si una muestra obtenida del paciente muestra un aumento de la expresión de dichos genes del ciclo celular; y
- (d) dar salida a una indicación de que el paciente tiene un aumento de la probabilidad de recurrencia si la muestra presenta un aumento de la expresión de dichos genes del ciclo celular.
- 35 8. Un kit que comprende una pluralidad de oligonucleótidos para la determinación del nivel de expresión de un panel de genes, en el que dicho panel comprende todos los genes en uno cualquiera de los paneles C a G, y en el que el kit consiste en reactivos para la determinación de la expresión de no más de 2.500 genes.
9. Uso de una pluralidad de oligonucleótidos para la determinación del nivel de expresión de un panel de genes, en el que dicho panel comprende todos los genes enumerados en uno cualquiera de los paneles C a G para determinar el pronóstico o la probabilidad de recurrencia del cáncer de próstata en una muestra obtenida de un paciente.

40

valores p obtenidos de subpoblaciones aleatorias de 26 genes de ciclo celular usados para predecir la recurrencia química del cáncer (ajustado por puntuación del nomograma post-PR)



Análisis multivariable de CoxPH del tiempo hasta la recurrencia tras cirugía prostática

100 muestras de FFPE

Expresión de ARNm de TLDA

1000 simulaciones aleatorias del dibujo de 26 CCG

Figura 1

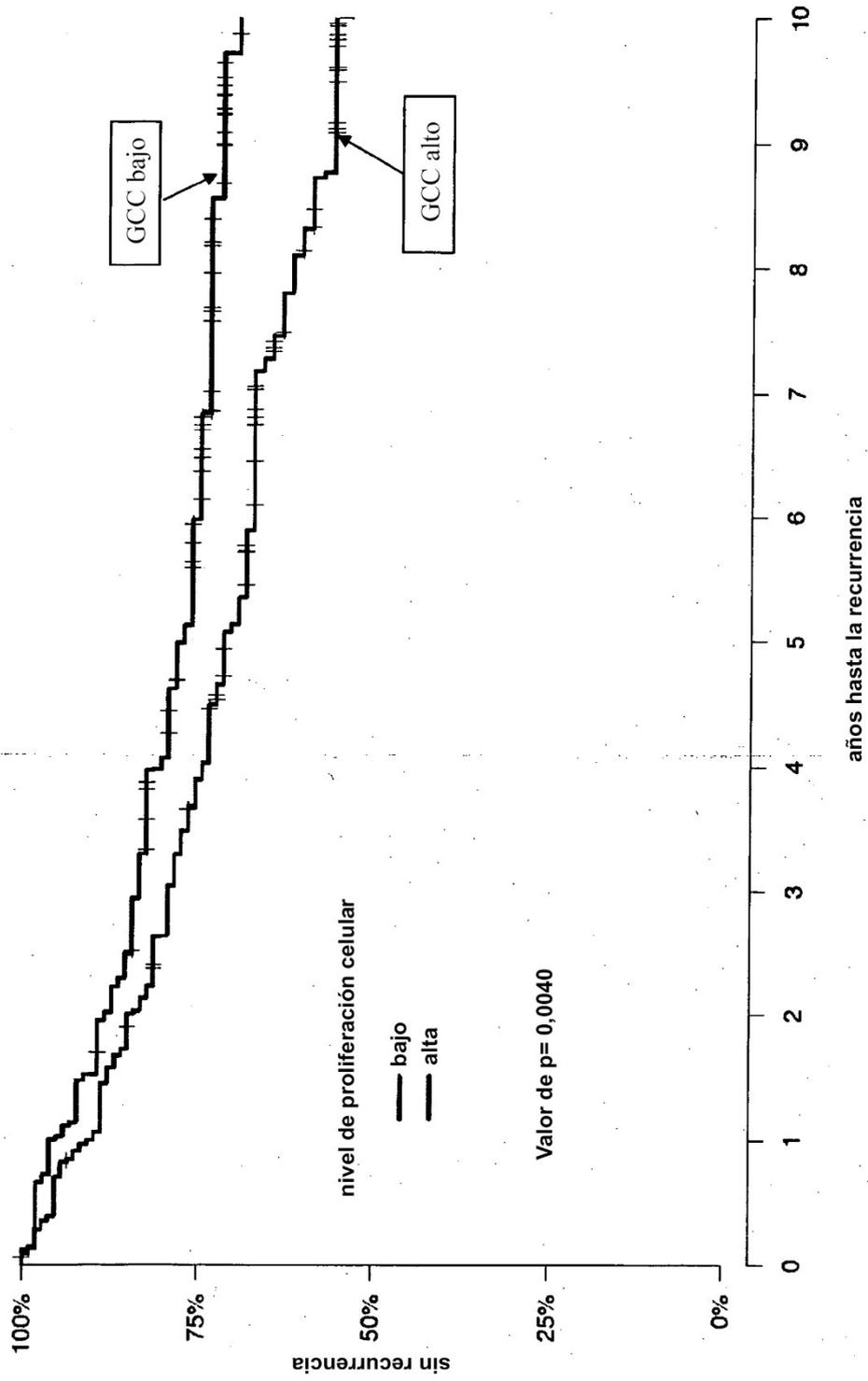


Figura 2

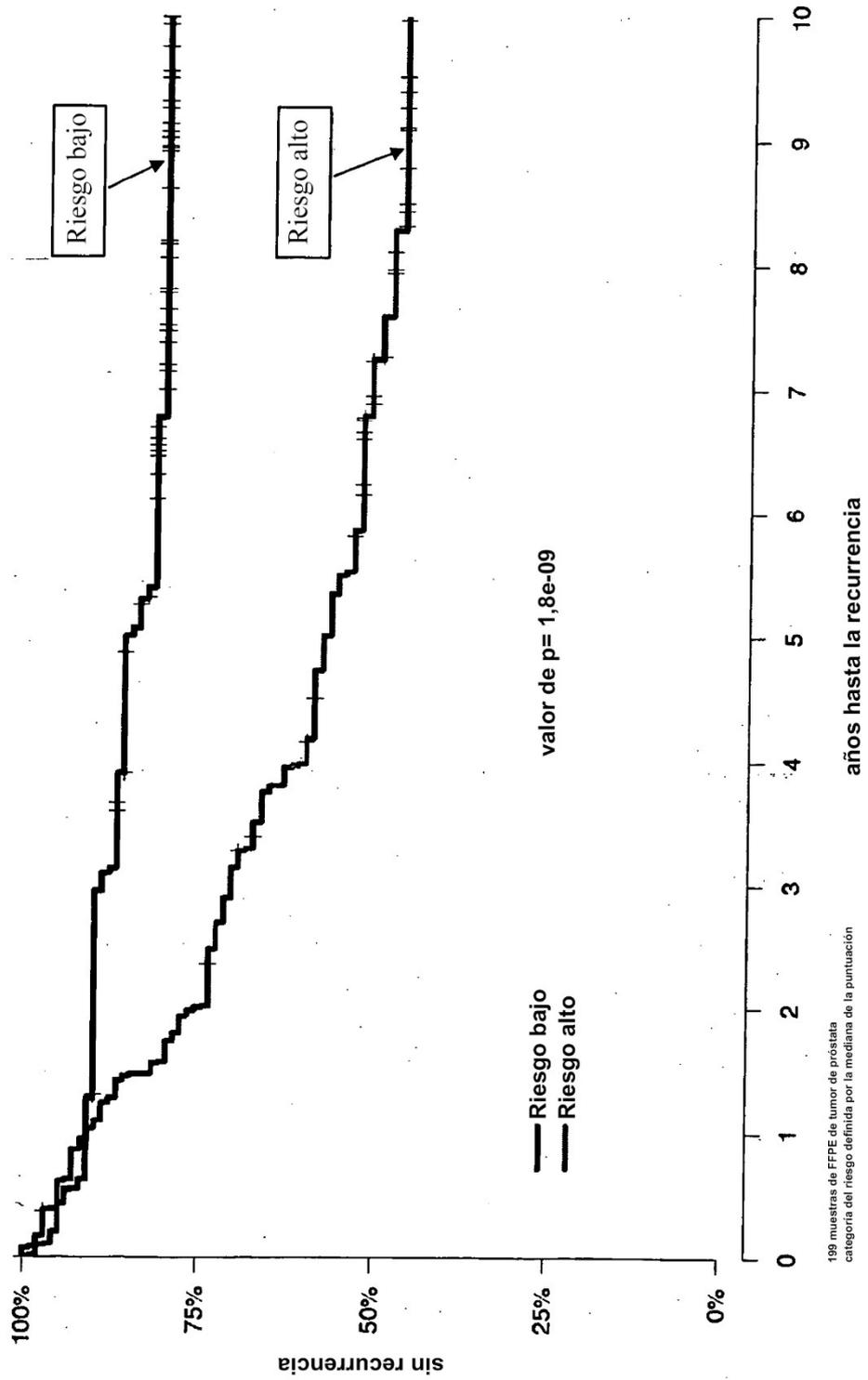
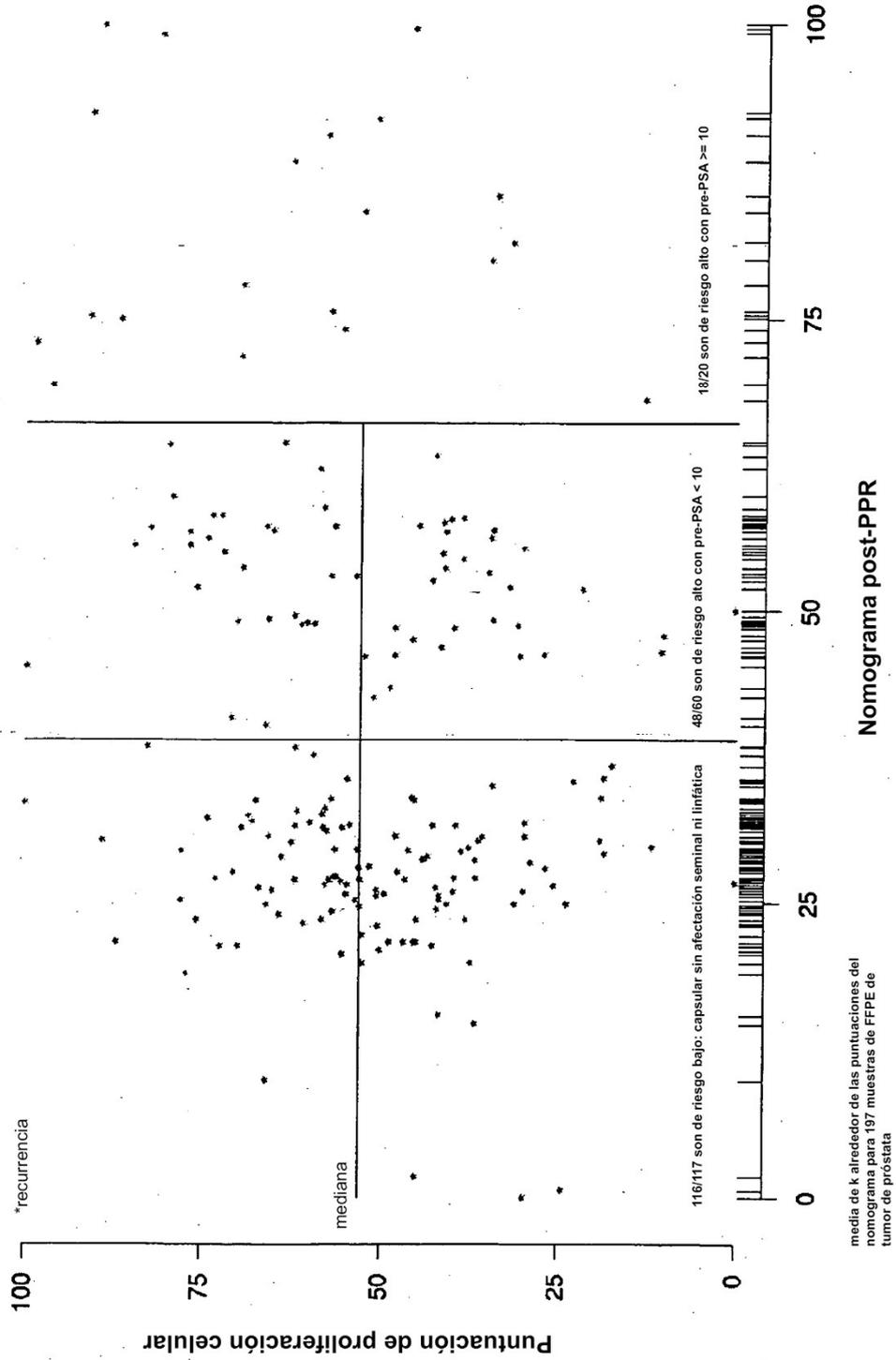


Figura 3



media de k alrededor de las puntuaciones del nomograma para 197 muestras de FFPE de tumor de próstata

Nomograma post-PPR

Figura 4

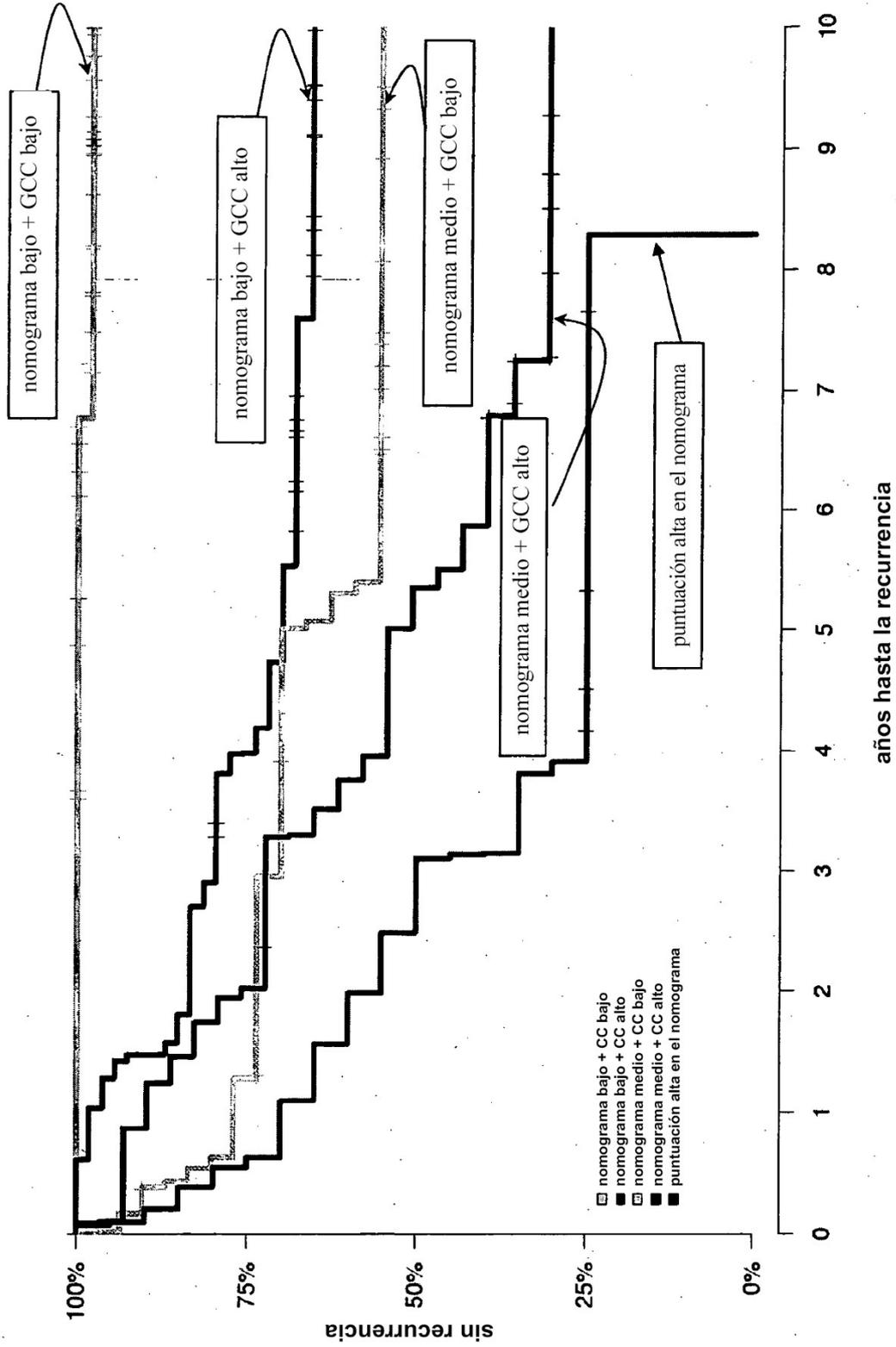


Figura 5

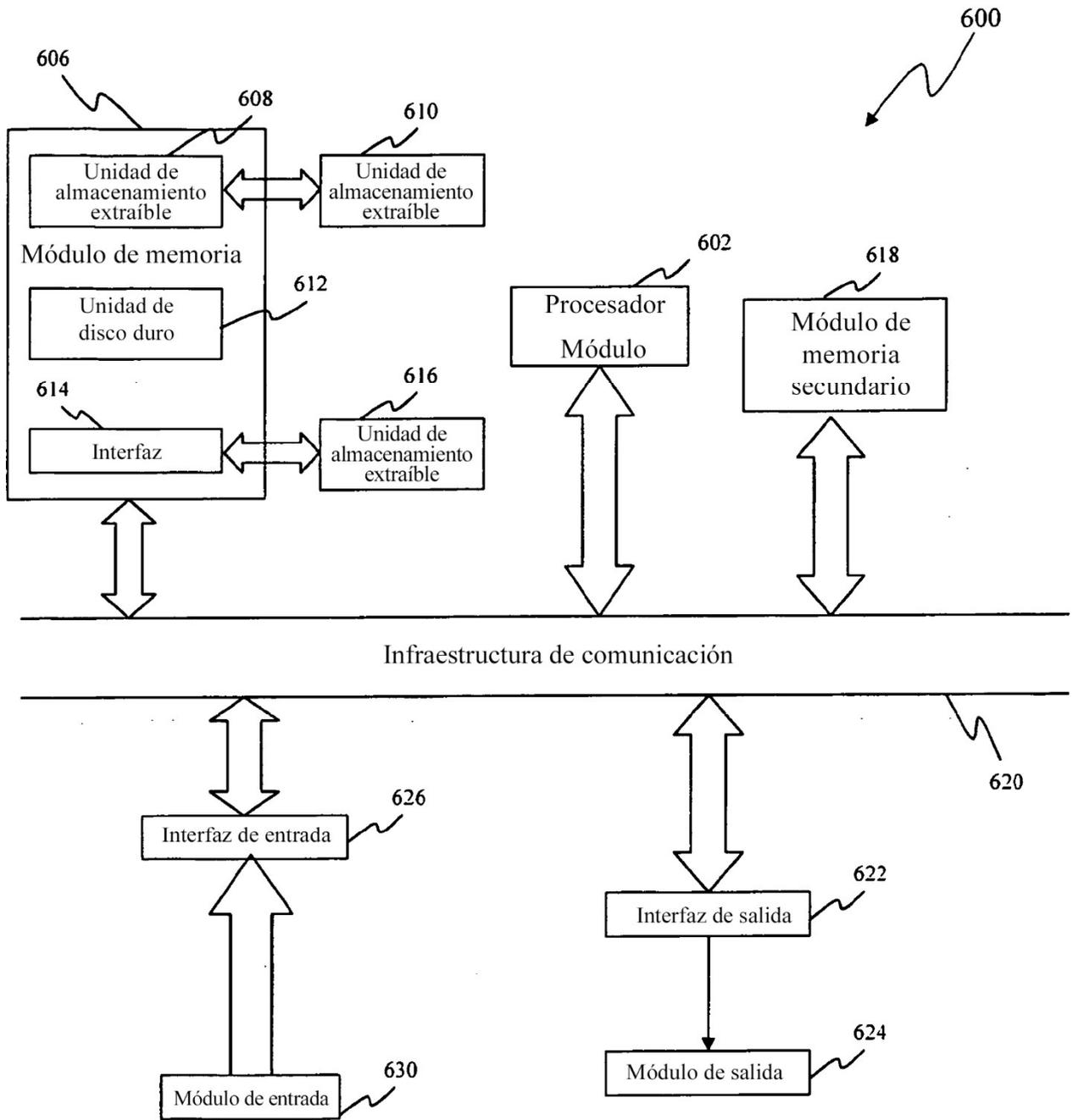


Figura 6

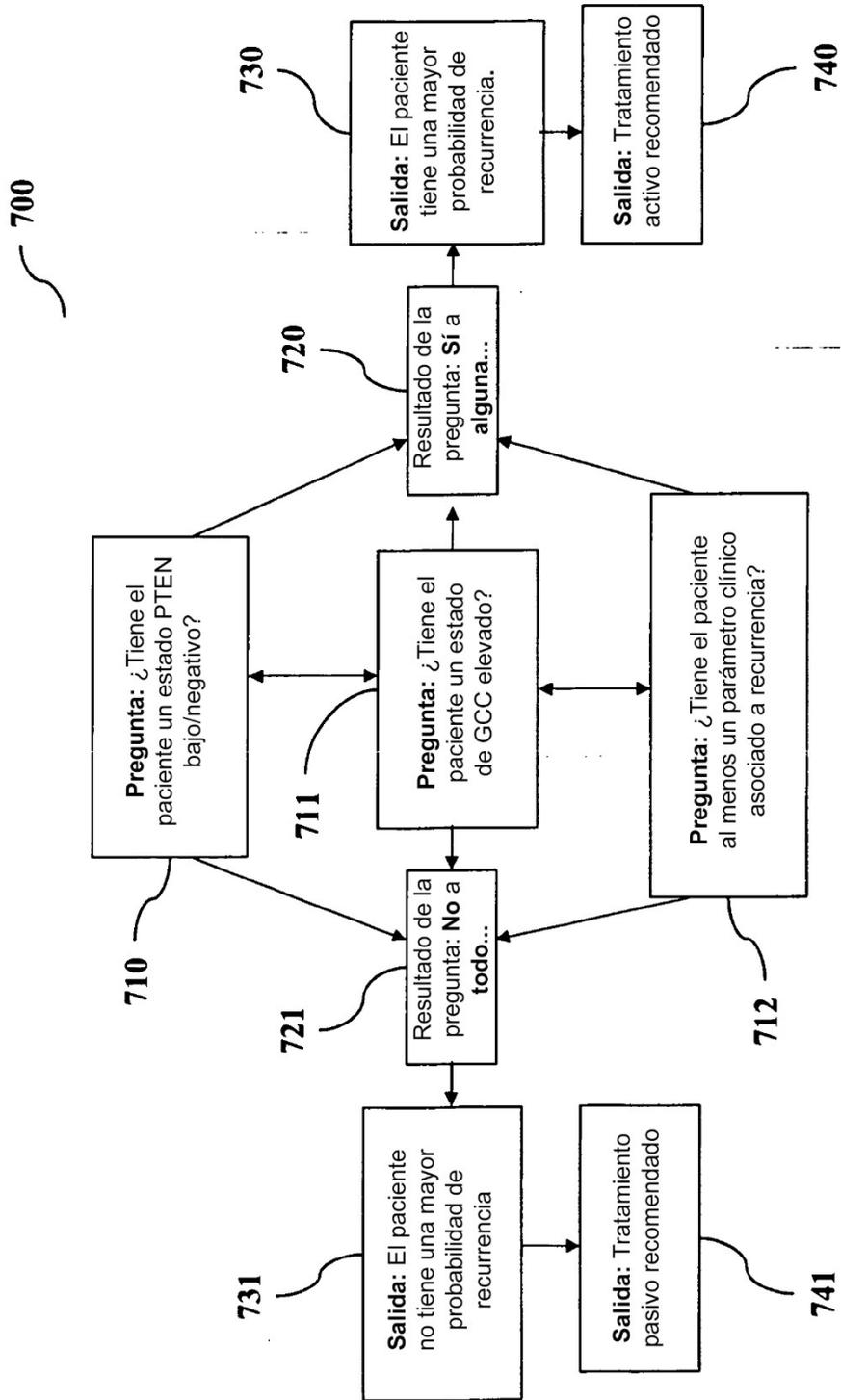


Figura 7

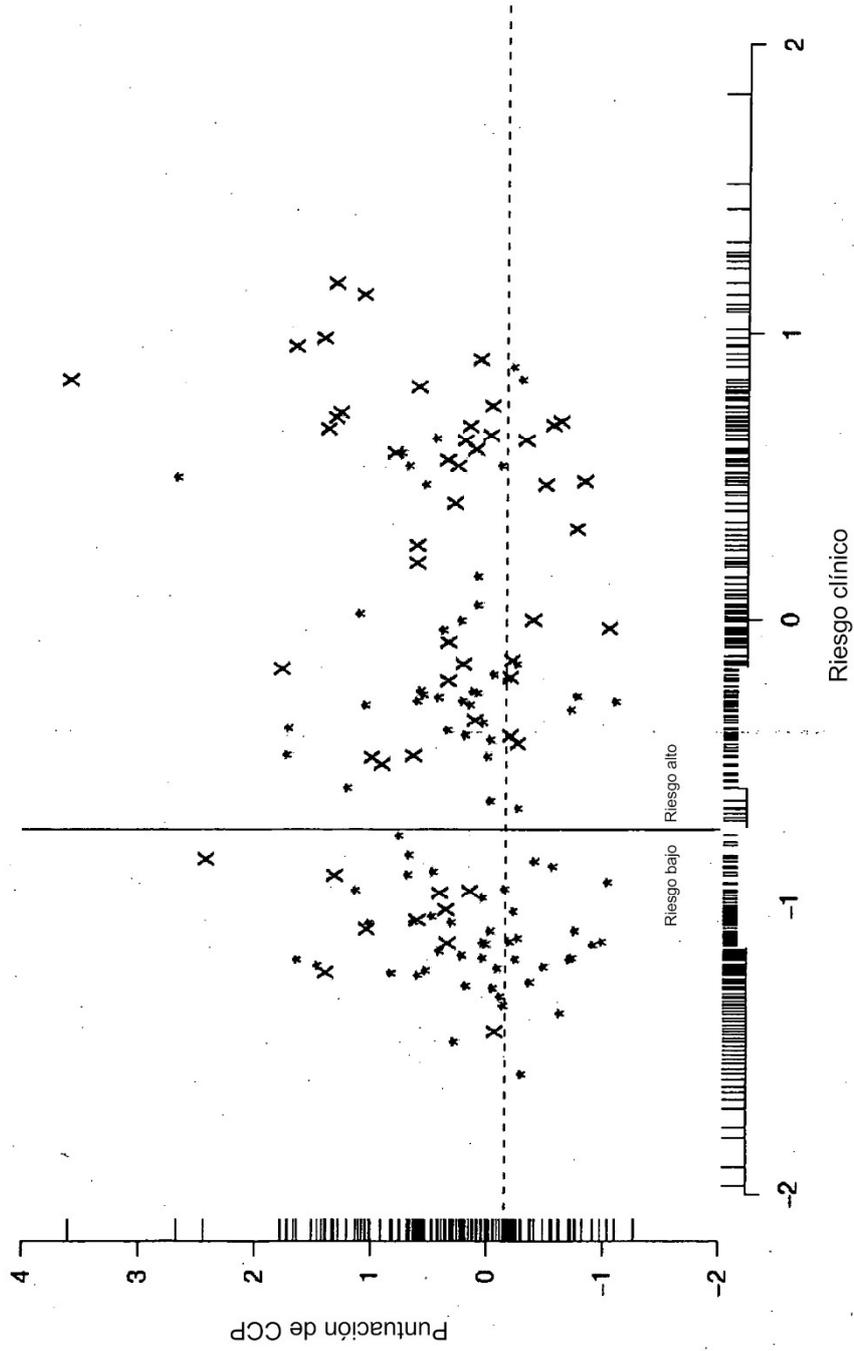


Figura 8

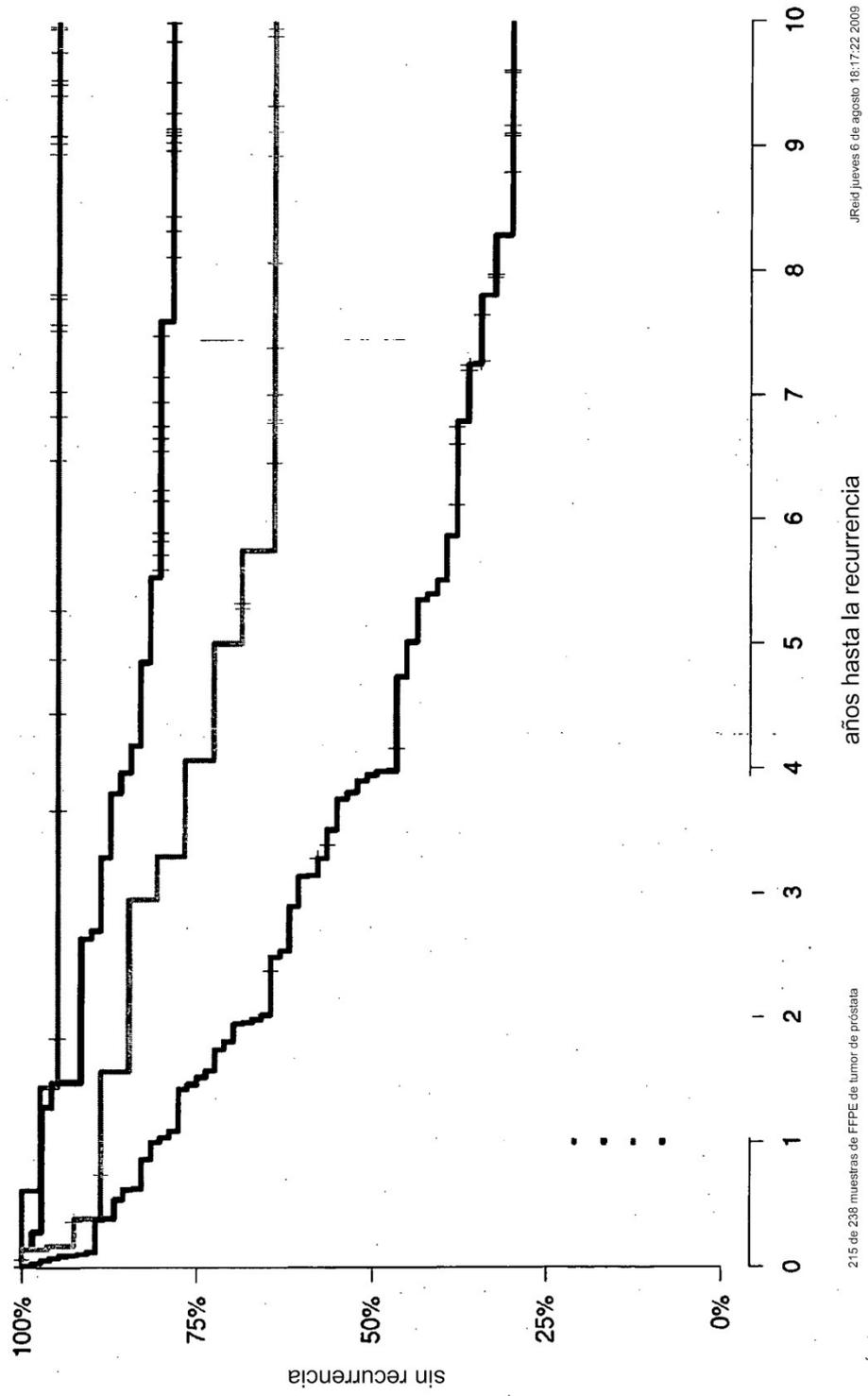


Figura 9

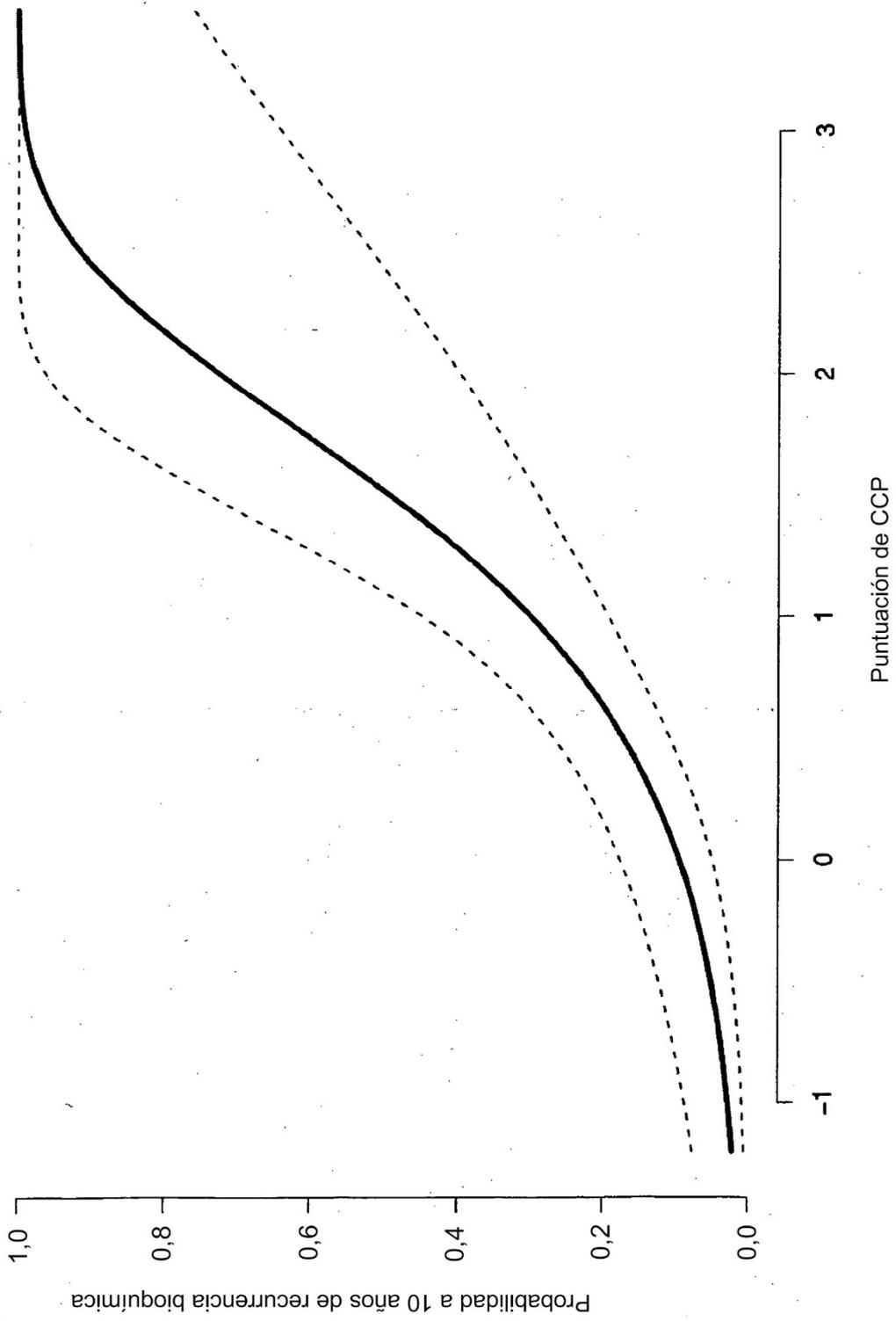


Figura 10

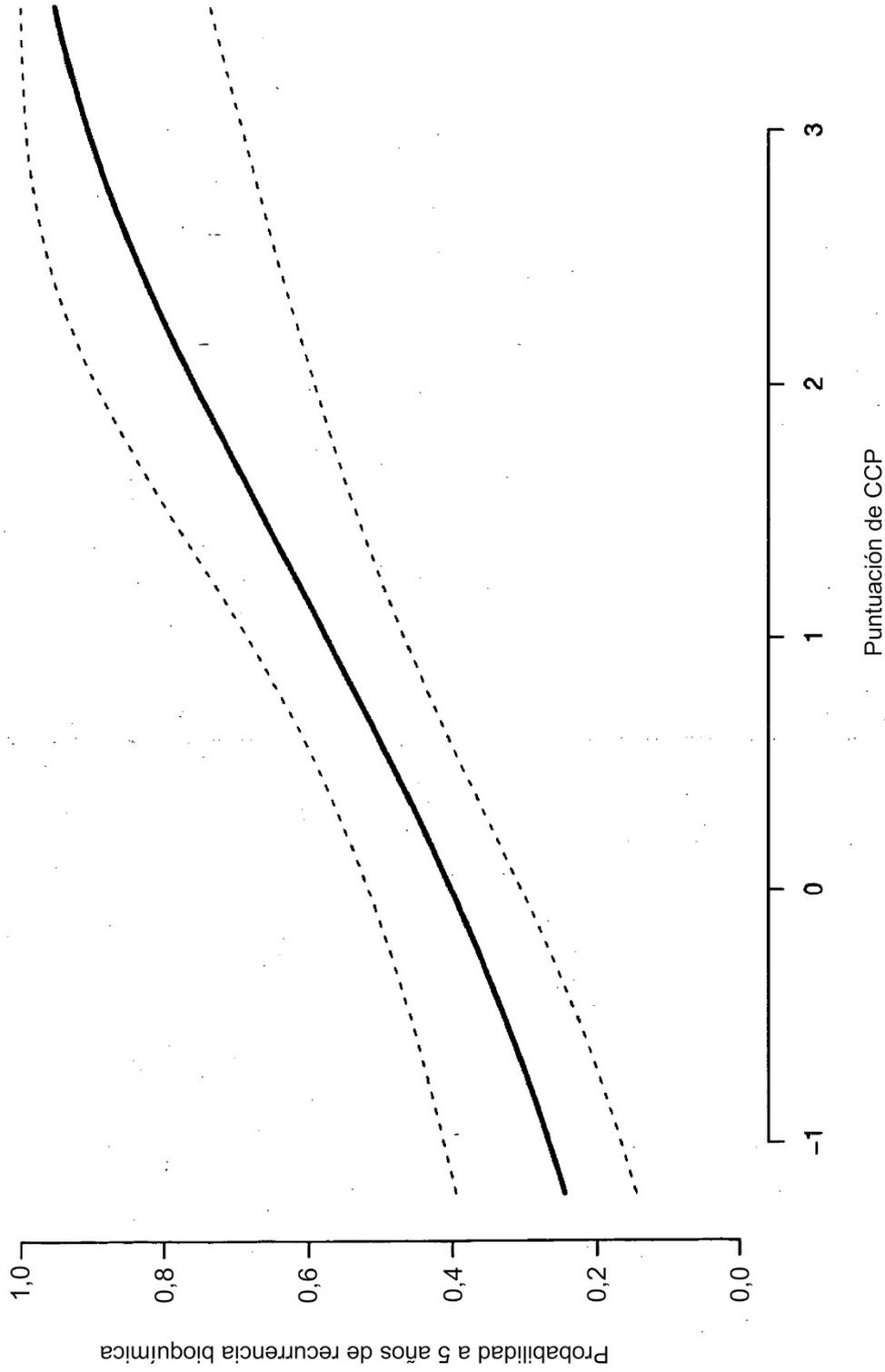


Figura 11

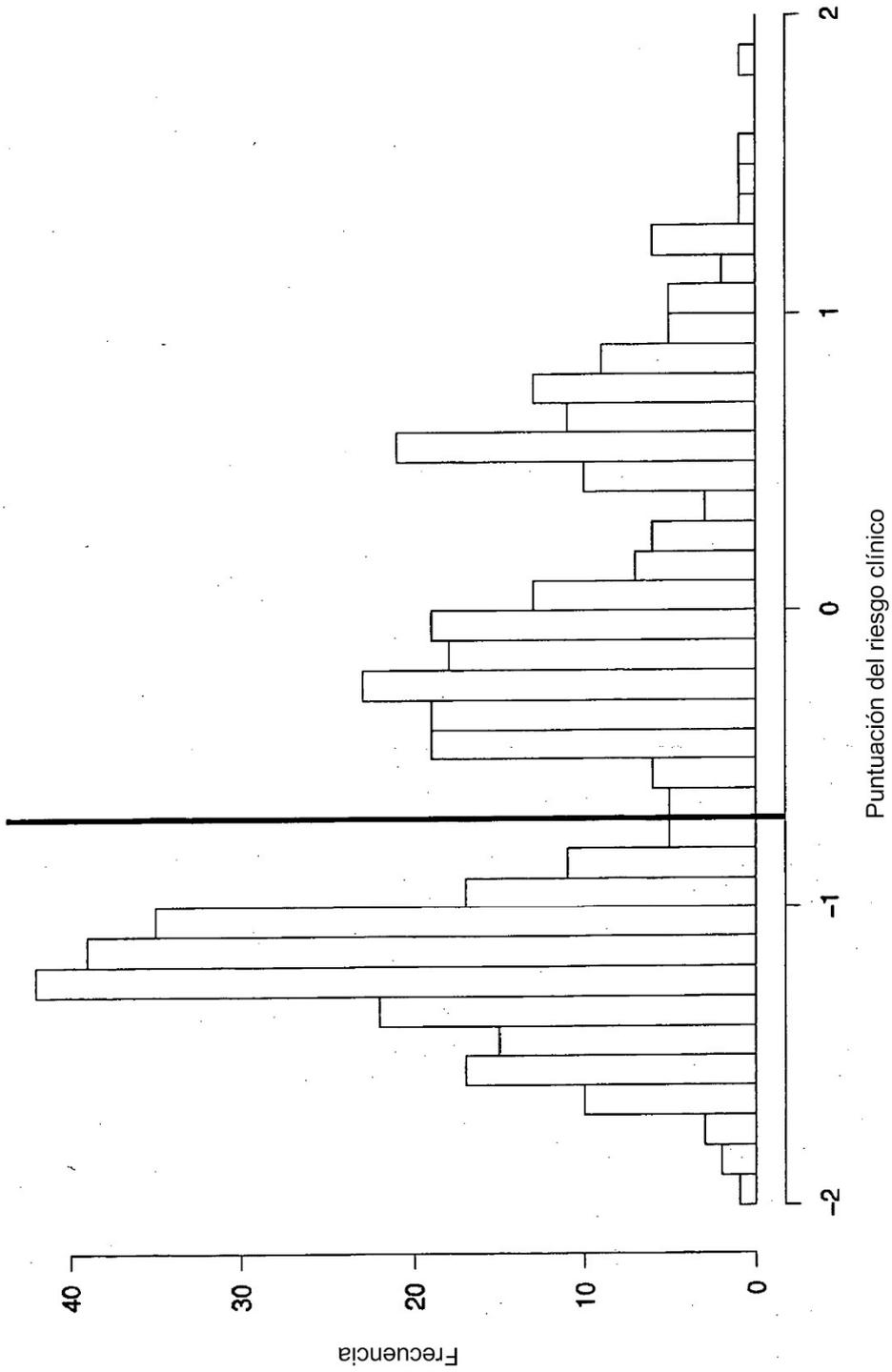


Figura 12

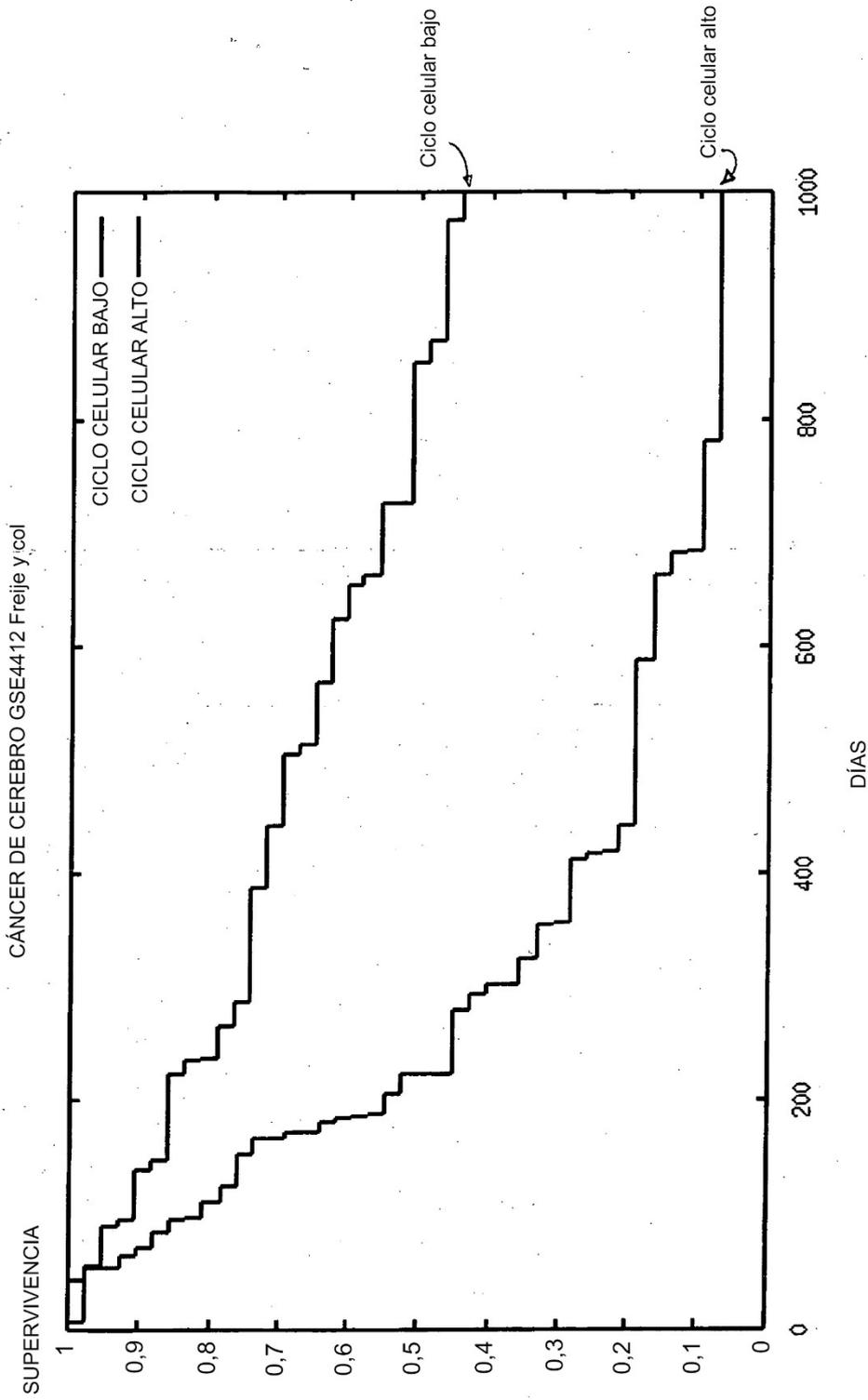


Figura 13

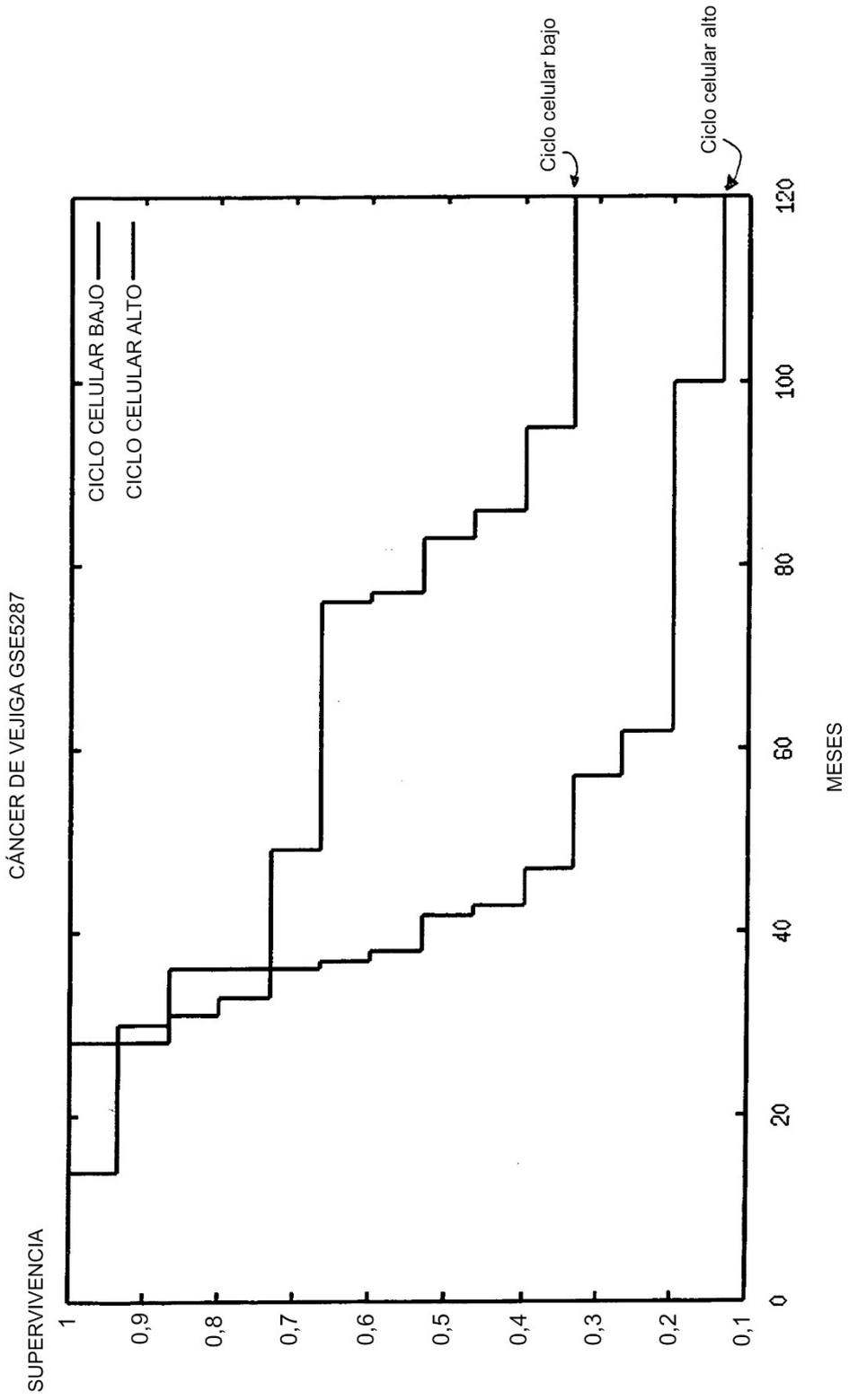


Figura 14

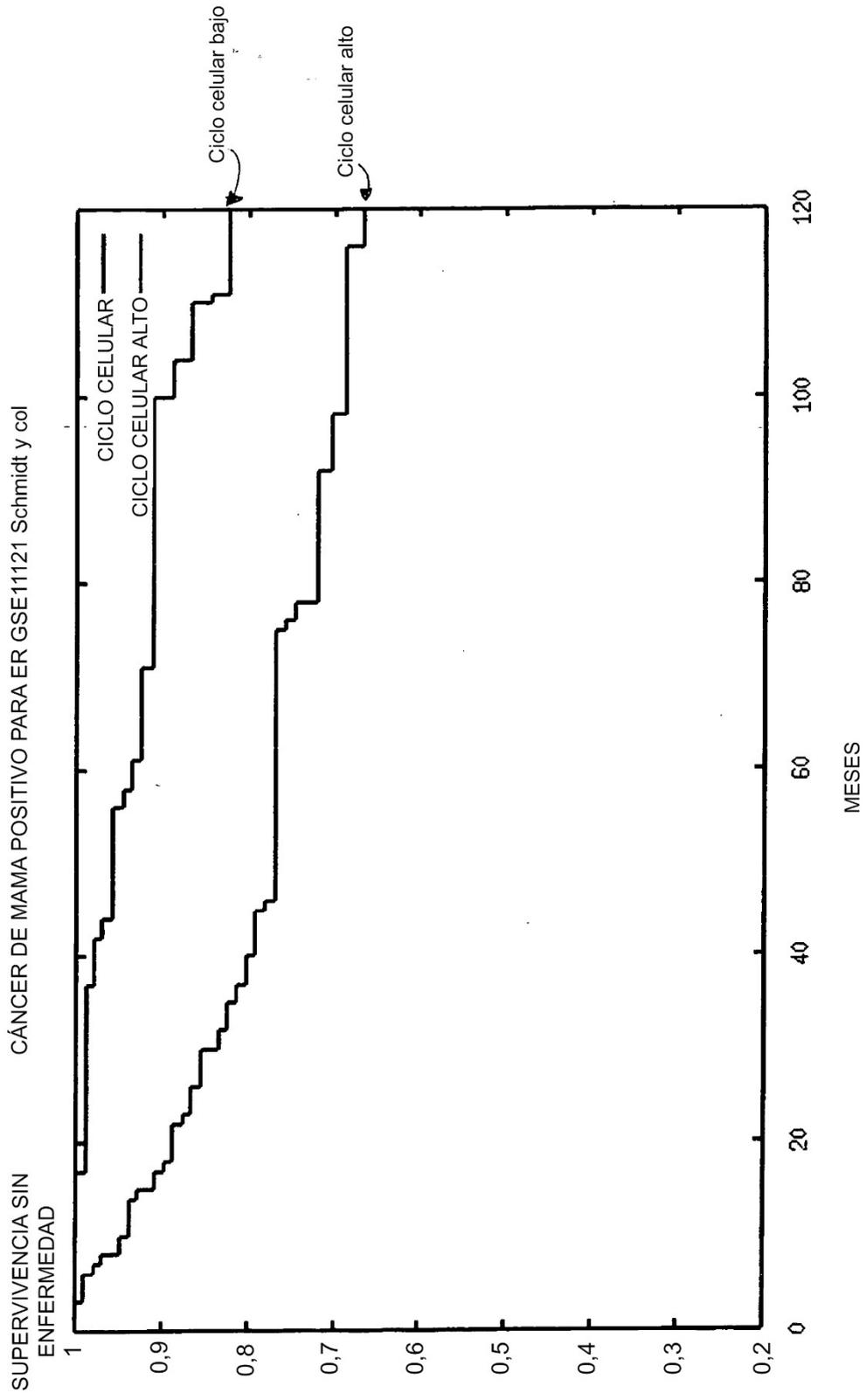


Figura 15

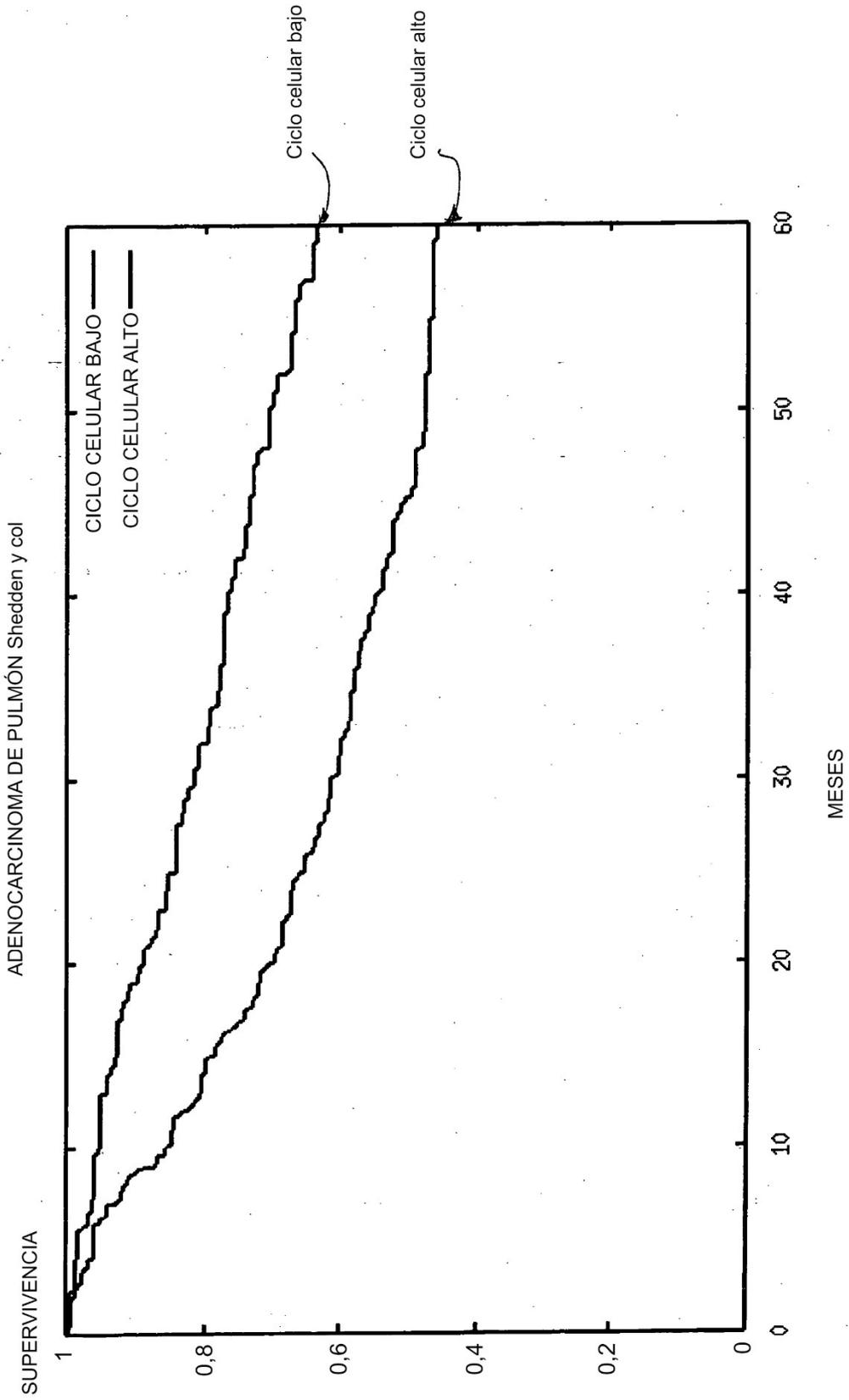


Figura 16