

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 991 436**

51 Int. Cl.:

C12Q 1/6886 (2008.01)

G01N 33/574 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **07.01.2010** **E 18209293 (2)**

97 Fecha y número de publicación de la concesión europea: **02.10.2024** **EP 3524697**

54 Título: **Biomarcadores de cáncer**

30 Prioridad:

07.01.2009 US 14307709 P

19.05.2009 US 17965009 P

10.06.2009 US 18590109 P

11.09.2009 US 24174809 P

30.10.2009 US 25644309 P

45 Fecha de publicación y mención en BOPI de la
traducción de la patente:
03.12.2024

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, JUTTA

74 Agente/Representante:

QUIRÓS MARÍN, María

ES 2 991 436 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

5 Campo de la invención

La invención se refiere de manera general a una clasificación molecular de enfermedad y de manera particular a marcadores moleculares para el cáncer y a métodos de uso de los mismos.

10 Antecedentes de la invención

El cáncer es uno de los principales problemas de salud pública, que representa aproximadamente el 25 % de todas las muertes en los Estados Unidos. Aunque se han concebido muchos tratamientos para diversos cánceres, estos tratamientos a menudo varían en cuanto a la gravedad de los efectos secundarios. Resulta útil que los médicos sepan cómo de agresivo es el cáncer de un paciente con el fin de determinar la agresividad con la que tratar el cáncer.

Por ejemplo, a la mayoría de los pacientes con cáncer de próstata asintomático de estadio temprano se les trata con prostatectomía radical o radioterapia y opcionalmente terapia complementaria (por ejemplo, terapia hormonal o quimioterapia), teniendo todas ellas graves efectos secundarios. Sin embargo, para muchos de estos pacientes, estos tratamientos y sus efectos secundarios y costes asociados son innecesarios porque el cáncer en estos pacientes no es agresivo (es decir, crece lentamente y es improbable que provoque mortalidad o morbilidad significativa durante la vida del paciente). En otros pacientes el cáncer es virulento (es decir, es más probable que reaparezca) y es necesario un tratamiento agresivo para salvar la vida del paciente.

Se han concebido algunas herramientas para ayudar a los médicos a decidir qué pacientes necesitan un tratamiento agresivo y cuáles no. De hecho, en la actualidad hay varios parámetros clínicos en uso para este propósito en diversos cánceres diferentes. En el cáncer de próstata, por ejemplo, tales parámetros clínicos incluyen antígeno prostático específico (PSA) en suero, grado de Gleason, estadio patológico, y márgenes quirúrgicos. En los últimos años, se ha hecho que los parámetros clínicos sean más útiles mediante su incorporación en nomogramas posoperatorios de múltiples variables continuas que calculan la probabilidad de que un paciente presente progresión/recidiva 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. Sin embargo, a pesar de estos avances, a muchos pacientes se les administran tratamientos contra el cáncer inadecuados y todavía existe la gran necesidad de herramientas nuevas y mejoradas para predecir la recidiva del cáncer.

Bibikova *et al* (Genomics, vol. 89, n.º 6, 12 de mayo de 2007, páginas 666-672) divulgaron firmas de expresión génica que se correlacionan con la puntuación de Gleason en pacientes con cáncer de próstata. Se identificó un panel de 16 genes, cuyos niveles de expresión se usaron para derivar una puntuación de expresión que se correlaciona con la recaída de la enfermedad.

Sumario de la invención

La presente invención es tal como se expone en las reivindicaciones. Se basa en la determinación de que los genes cuya expresión sigue de cerca al ciclo celular ("genes del ciclo celular" o "CCG") son genes particularmente poderosos para clasificar ciertos cánceres. De hecho, los inventores han descubierto que los CCG proporcionan la mayor parte de, si no todo, el poder predictivo en muchas firmas pronósticas. Por ejemplo, en el presente documento se muestra que los paneles de CCG son capaces de predecir la agresividad del cáncer de próstata, vejiga, cerebro, mama y pulmón.

Por tanto, un aspecto de la invención proporciona un método para clasificar el cáncer de próstata que comprende determinar el estado de un panel de genes que comprende al menos diez CCG de cualquiera de los paneles C-G, en el que un estado anómalo indica una clasificación de cáncer particular. En algunas realizaciones, el estado que va a determinarse son los niveles de expresión génica. En algunas realizaciones, el estado anómalo indica que el cáncer es agresivo, que puede incluir tener un mal pronóstico, una probabilidad aumentada de recidiva o progresión, etc. En algunas realizaciones, la invención proporciona un método para determinar el pronóstico del cáncer de próstata de un paciente que comprende determinar el nivel de expresión de un panel de genes que comprende al menos diez CCG de cualquiera de los paneles C-G, en el que una expresión elevada indica una probabilidad aumentada de recidiva o progresión del cáncer o la necesidad de un tratamiento agresivo.

El número de CCG sometidos a ensayo puede variar dependiendo de, por ejemplo, restricciones técnicas, consideraciones de costes, la clasificación que esté realizándose, el cáncer que esté sometiéndose a prueba, el nivel deseado de poder predictivo, etc. Sin embargo, se ha determinado que ciertos números mínimos de CCG proporcionan predicciones robustas y significativamente mejoradas. En algunas realizaciones, el panel comprende al menos 10, 15, 20, 25, 30, 35, 40, 45, 50 o más CCG. En algunas realizaciones, el panel comprende al menos 10, 15, 20 o más CCG.

En algunas realizaciones, los CCG comprenden al menos una cierta proporción del panel de genes. Por tanto, en algunas realizaciones, el panel comprende al menos el 25 %, el 30 %, el 40 %, preferiblemente al menos el 50 %, el 60 %, el 70 %, más preferiblemente al menos el 75 %, el 80 %, el 85 %, y lo más preferiblemente el 90 %, el 95 %, el 96 %, el 97 %, el 98 % o el 99 % de CCG.

En algunas realizaciones, los CCG se eligen del grupo que consiste en los genes en los paneles C a G. En algunas realizaciones, el panel comprende al menos 10 o más de los genes enumerados en cualquiera de los paneles C a G.

Se ha descubierto que los CCG son particularmente predictivos en ciertos cánceres. Por ejemplo, se ha determinado que un panel de CCG es preciso a la hora de predecir la recidiva en el cáncer de próstata. Además, los CCG pueden determinar el pronóstico en cáncer de vejiga, cerebro, mama y pulmón. Por tanto, la invención proporciona un método que comprende determinar el estado de un panel de genes que comprende al menos diez CCG, en el que un estado anómalo indica un mal pronóstico. En algunas realizaciones, el panel comprende al menos 10 genes elegidos del grupo de genes en al menos uno de los paneles C a G. En algunas realizaciones, el panel comprende al menos 15 genes elegidos del grupo de genes en al menos uno de los paneles A a G. En algunas realizaciones, el panel comprende la totalidad de los genes en al menos uno de los paneles A a G.

Se ha descubierto además que el estado de CCG se añade sinérgicamente a los parámetros clínicos (por ejemplo, nomogramas clínicos) a la hora de pronosticar el cáncer. En el caso de cáncer de próstata, se ha descubierto que un nivel alto de expresión génica de uno cualquiera de los genes en cualquiera de los paneles A a F está asociado con un riesgo aumentado de recidiva o progresión del cáncer en pacientes cuya puntuación de nomograma clínico indica un riesgo relativamente bajo de recidiva o progresión. Véanse, por ejemplo, los ejemplos 3 y 5. En el caso de cáncer de cerebro, se ha descubierto que un nivel alto de expresión génica de uno cualquiera de los genes en cualquiera de los paneles B, C o F está asociado con un riesgo aumentado de recidiva o progresión del cáncer independientemente del grado tumoral. En el caso de cáncer de pulmón, se ha descubierto que un nivel alto de expresión génica de uno cualquiera de los genes en cualquiera de los paneles B, C o F está asociado con un riesgo aumentado de recidiva o progresión del cáncer independientemente del estadio tumoral. En el caso de cáncer de mama, se ha descubierto que un nivel alto de expresión génica de uno cualquiera de los genes en cualquiera de los paneles B, C o F está asociado con un riesgo aumentado de recidiva o progresión del cáncer independientemente de la puntuación de nomograma del índice pronóstico de Nottingham (NPI). Por tanto, dado que la evaluación de los niveles de expresión de CCG puede detectar un riesgo aumentado no detectado usando los parámetros clínicos solos, la invención proporciona de manera general métodos que combinan la evaluación de al menos un parámetro clínico con la evaluación del estado de al menos un CCG.

Por tanto, un aspecto de la invención proporciona un método de diagnóstico *in vitro* que comprende determinar al menos un parámetro clínico para un paciente con cáncer de próstata y determinar el estado de al menos diez CCG en una muestra obtenida del paciente. En algunas realizaciones, se determina el estado de una pluralidad de CCG (por ejemplo, 10, 15, 20, 25, 30, 40, 50 o más). En algunas realizaciones, el estado anómalo indica una probabilidad aumentada de recidiva o progresión. A menudo, el parámetro clínico es al menos algo independientemente predictivo de recidiva o progresión y la adición del estado de CCG mejora el poder predictivo. Los CCG a modo de ejemplo incluyen aquellos en los paneles A a G. En algunas realizaciones, se evalúa más de un parámetro clínico junto con el nivel de expresión de al menos diez CCG. En algunas realizaciones, se evalúa más de un parámetro clínico, y la combinación de parámetros, junto con el nivel de expresión de al menos diez CCG, predice la probabilidad de recidiva o progresión del cáncer.

A menudo, los parámetros clínicos se combinan e incorporan en nomogramas para predecir el riesgo de recidiva o progresión del cáncer. Por tanto, otro aspecto de la invención proporciona un método de diagnóstico *in vitro* que comprende determinar una puntuación de nomograma clínico para un paciente con cáncer y determinar el estado de al menos diez CCG de cualquiera de los paneles C-G en una muestra obtenida del paciente. A menudo, los nomogramas clínicos están diseñados de manera que una puntuación alta se correlaciona con un riesgo aumentado de recidiva. De manera similar, un estado de CCG elevado (por ejemplo, expresión o actividad aumentada) se correlaciona con un riesgo aumentado. Por tanto, en algunas realizaciones, la invención proporciona un método para determinar si un paciente con cáncer de próstata tiene una probabilidad aumentada de recidiva o progresión que comprende determinar una puntuación de nomograma clínico para el paciente y determinar el estado de al menos diez CCG de cualquiera de los paneles C-G en una muestra obtenida del paciente, en el que una puntuación alta de nomograma y/o un estado de CCG elevado indican que el paciente tiene una probabilidad aumentada de recidiva o progresión.

En un aspecto, la invención proporciona un método de diagnóstico *in vitro* que comprende determinar al menos un parámetro clínico (o una puntuación de nomograma) para un paciente y determinar el estado de un panel de CCG que comprende al menos diez CCG de cualquiera de los paneles C-G. El número de CCG sometidos a ensayo puede variar dependiendo de, por ejemplo, restricciones técnicas, consideraciones de costes, la clasificación que esté realizándose, el cáncer que esté sometiéndose a prueba, el nivel deseado de poder predictivo, etc. Como un ejemplo, la invención proporciona un método de diagnóstico *in vitro* que comprende determinar una puntuación de nomograma para un paciente y determinar el nivel de expresión de diez o más de los genes enumerados en los

paneles C a G. Puede usarse cualquier combinación de CCG (incluyendo cualquiera de los enumerados en la tabla 1 o los paneles A a G) para poner en práctica la invención. En una realización, se someten a ensayo todos los CCG enumerados en la tabla 1 o 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 añade tanto a los parámetros clínicos como a los CCG. El estado de *PTEN* combinado con el estado de CCG es un mejor factor predisponente que un nomograma clínico de próstata. Dado que la evaluación del estado de *PTEN* puede detectar un riesgo aumentado no detectado usando los parámetros clínicos (por ejemplo, nomograma clínico) o los niveles de expresión de CCG solos o en combinación, algunos métodos combinan un parámetro clínico o nivel de expresión de CCG con una evaluación del estado de *PTEN*.

En los ejemplos a continuación se ha demostrado que pueden combinarse los tres marcadores para proporcionar un poder predictivo incluso mayor.

El estado de un gen incluye cualquier característica física, química o genética de ese gen. Los ejemplos incluyen nivel de expresión, nivel de actividad, mutaciones, número de copias, etc. En el contexto de CCG para predecir el riesgo de recidiva o progresión, el nivel de expresión y el nivel de actividad a menudo son las características más útiles a evaluar a la hora de determinar el estado de un CCG. En algunas realizaciones, determinar el nivel de expresión de un gen comprende al menos uno de:

(a) determinar la cantidad de polipéptido codificado por el gen; o

(b) determinar la cantidad de ARN codificado por el gen.

Breve descripción de los dibujos

La figura 1 es una ilustración del poder predictivo a lo largo del nomograma para paneles de CCG de diferentes tamaños.

La figura 2 es una ilustración de CCG que predicen el tiempo hasta la recidiva.

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

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

La figura 5 es una ilustración del tiempo hasta la recidiva para varias poblaciones de pacientes definidas por nomograma y/o estado de CCG.

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

La figura 7 es un diagrama de flujo que ilustra un ejemplo de un método implementado por ordenador.

La figura 8 es un diagrama de dispersión que compara los parámetros clínicos y la puntuación de CCG como factores predisponentes de recidiva del ejemplo 5.

La figura 9 ilustra, a partir del ejemplo 5, el umbral de CCG derivado del análisis de la cohorte de entrenamiento con respecto al conjunto de datos de validación, subdividiendo eficazmente la puntuación de firma de CCG a pacientes identificados como con riesgo bajo usando parámetros clínicos en pacientes con tasas de recidiva muy bajas y con un mayor riesgo de recidiva.

La figura 10 ilustra la tasa de recidiva predicha frente a la puntuación de CCG para pacientes en la cohorte de validación del ejemplo 5.

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

La figura 12 ilustra la distribución de puntuación de riesgo clínico en 443 pacientes estudiados en el ejemplo 5. La línea vertical oscura representa el umbral elegido mediante medias de KM para dividir a pacientes con riesgo bajo y alto y usado a lo largo de este estudio.

Descripción detallada de la invención

La presente invención se basa en parte en el descubrimiento de que los genes cuya expresión sigue de cerca al ciclo celular ("genes del ciclo celular" o "CCG") son genes particularmente poderosos para clasificar ciertos cánceres.

"Gen del ciclo celular" y "CCG" en el presente documento se refieren a un gen cuyo nivel de expresión sigue de

cerca a la progresión de la célula a través del ciclo celular. Véase, por ejemplo, Whitfield *et al.*, MOL. BIOL. CELL (2002) 13:1977-2000. Más específicamente, los CCG muestran aumentos y disminuciones periódicos en la expresión que coinciden con ciertas fases del ciclo celular, por ejemplo, *STK15* y *PLK* muestran una expresión máxima en G2/M. *Id.* A menudo, los CCG tienen una función relacionada con el ciclo celular clara y reconocida, por ejemplo, en la síntesis o reparación del ADN, en la condensación de cromosomas, en la división celular, etc. Sin embargo, algunos CCG tienen niveles de expresión que siguen el ciclo celular sin tener ningún papel directo obvio en el ciclo celular, por ejemplo, *UBE2S* codifica para una enzima que se conjuga con ubiquitina, aunque su expresión sigue de cerca al ciclo celular. Por tanto, no es necesario que un CCG según la presente invención tenga un papel reconocido en el ciclo celular. Se enumeran CCG a modo de ejemplo en las tablas 1, 2, 3 y 4.

Si un gen particular es un CCG puede determinarse mediante cualquier técnica conocida en la técnica, incluyendo aquella enseñada en Whitfield *et al.*, MOL. BIOL. CELL (2002) 13:1977-2000. Por ejemplo, una muestra de células, por ejemplo, células HeLa, pueden sincronizarse de manera que todas ellas avancen a través de las diferentes fases del ciclo celular al mismo tiempo. Generalmente, esto se realiza deteniendo a las células en cada fase, por ejemplo, las células pueden detenerse en la fase S usando un bloque de timidina doble o en la mitosis con un bloque de timidina-nocodazol. Véase, por ejemplo, Whitfield *et al.*, MOL. CELL. BIOL. (2000) 20:4188-4198. Se extrae el ARN de las células después de la detención en cada fase y se cuantifica la expresión génica usando cualquier técnica adecuada, por ejemplo, micromatriz de expresión (de genoma completo o genes de interés específicos), PCR[™] cuantitativa en tiempo real (RTQ-PCR). Finalmente, se aplica análisis estadístico (por ejemplo, transformada de Fourier) para determinar qué genes muestran una expresión máxima durante fases particulares del ciclo celular. Los genes pueden clasificarse según una puntuación de periodicidad que describe cómo de cerca sigue la expresión génica al ciclo celular, por ejemplo, una puntuación alta indica un gen que sigue muy de cerca al ciclo celular. Finalmente, esos genes cuya puntuación de periodicidad supera un nivel de umbral definido (véase Whitfield *et al.*, MOL. BIOL. CELL (2002) 13:1977-2000) pueden designarse como CCG. En la tabla 1 se proporciona una lista larga, pero no exhaustiva, de ácidos nucleicos asociados con CCG (por ejemplo, genes, EST, clones de ADNc, etc.). Véase Whitfield *et al.*, MOL. BIOL. CELL (2002) 13:1977-2000. Todos los CCG en la tabla 2 a continuación forman un panel de CCG ("panel A")

Tabla 2

Símbolo de gen	ID de gen de Entrez	ID de ensayo de ABI	N.ºs de registro de 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_17915.2
<i>C18orf24</i> *	220134	Hs00536843_m1	NM_145060.3; NM_001039535.2
<i>C1orf135</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

ES 2 991 436 T3

<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>CDCA3*</i>	83461	Hs00229905_m1	NM_031299.4
<i>CDCA8*</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

ES 2 991 436 T3

<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>HN1*</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>KIF25*</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/IMPHOSPH1*</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
<i>NEK2*</i>	4751	Hs00601227_mH	NM_002497.2

ES 2 991 436 T3

<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 CCG útiles en la invención ("panel B"). ID de ensayo de ABI significa el número de ID de catálogo para el ensayo de expresión génica disponible comercialmente de Applied Biosystems Inc. (Foster City, CA) para el gen particular.

- 5 Aunque no se desea estar limitado por ninguna teoría, se piensa que las muestras de tejido que contienen una concentración relativamente alta de células que proliferan rápidamente mostrarán, en total, una expresión relativamente mayor de los CCG. Se piensa que esto es debido a que algunos CCG están regulados por incremento durante la proliferación celular (es decir, durante ciertas fases del ciclo celular). A su vez, los tumores que comprenden una alta concentración de células que proliferan rápidamente tienden a ser más agresivos y los
- 10 pacientes que padecen tales tumores tienden a tener un peor pronóstico. De hecho, los inventores han determinado que los CCG proporcionan la mayor parte de, si no todo, el poder predictivo en la mayoría de las firmas pronósticas de cáncer. Por ejemplo, los paneles que tienen CCG pueden predecir la agresividad del cáncer de próstata, cerebro, mama y pulmón.
- 15 Por tanto, un aspecto de la invención proporciona un método para clasificar el cáncer de próstata que comprende determinar el estado de un panel de genes que comprende al menos diez CCG de uno cualquiera de los paneles C-G, en una muestra de tejido o células, particularmente una muestra de tumor, de un paciente, en el que un estado anómalo indica una clasificación de cáncer negativa. Tal como se usa en el presente documento, "determinar el estado" de un gen se refiere a determinar la presencia, la ausencia o el grado/nivel de alguna característica física,
- 20 química o genética del gen o su(s) producto(s) de expresión. La presente invención hace uso de la determinación de la expresión media de un panel de genes tal como se recita en las reivindicaciones.

En el contexto de los CCG que se usan para determinar el riesgo de recidiva o progresión del cáncer o la necesidad de un tratamiento agresivo, las características particularmente útiles incluyen los niveles de expresión (por ejemplo, niveles de ARNm o proteína) y los niveles de actividad. Las características pueden someterse a ensayo directamente (por ejemplo, sometiendo a ensayo el nivel de expresión de un CCG) o determinarse indirectamente (por ejemplo, sometiendo a ensayo el nivel de un gen o genes cuyo nivel de expresión se correlaciona con el nivel de expresión del CCG). Por tanto, algunas realizaciones de la invención proporcionan un método para clasificar el

25 cáncer de próstata que comprende determinar el nivel de expresión, particularmente el nivel de ARNm, de un panel de genes que comprende al menos diez CCG de uno cualquiera de los paneles C-G, en una muestra de tumor, en el que una expresión elevada indica una clasificación de cáncer negativa, o un riesgo aumentado de recidiva 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 una

proteína en una muestra de tejido o células, incluyendo, pero sin limitarse a, análisis de micromatrices (por ejemplo, para someter a ensayo la expresión de ARNm o microARN, el número de copias, etc.), PCR[™] cuantitativa en tiempo real ("qRT-PCR[™]", por ejemplo, TaqMan[™]), inmunoanálisis (por ejemplo, ELISA, inmunohistoquímica), etc.

Un "estado anómalo" significa que el estado de un marcador en una muestra particular difiere del estado generalmente hallado en muestras regulares (por ejemplo, muestras sanas o muestras enfermas regulares). Los ejemplos incluyen mutado, elevado, disminuido, presente, ausente, etc. Un "estado elevado" significa que una o más de las características anteriores (por ejemplo, nivel de expresión o ARNm) es mayor que los niveles normales. Generalmente esto significa un aumento en la característica (por ejemplo, nivel de expresión o ARNm) en comparación con un valor indicador. A la inversa un "estado bajo" significa que una o más de las características anteriores (por ejemplo, nivel de expresión génica o ARNm) es menor que los niveles normales. Generalmente esto significa una disminución en la característica (por ejemplo, expresión) en comparación con un valor indicador. En este contexto, un "estado negativo" generalmente significa que la característica está ausente o es indetectable. Por ejemplo, el estado de *PTEN* es negativo si el ácido nucleico y/o la proteína de *PTEN* está ausente o es indetectable en una muestra. Sin embargo, el estado de *PTEN* negativo también incluye una mutación o reducción del número de copias en *PTEN*.

Los expertos en la técnica apreciarán cómo obtener y usar un valor indicador en los métodos de la invención. Por ejemplo, el valor indicador puede representar los niveles de expresión génica hallados en una muestra normal obtenida del paciente de interés, en cuyo caso un nivel de expresión en la muestra de tumor significativamente mayor a este valor indicador indicaría, por ejemplo, un mal pronóstico o una probabilidad aumentada de recidiva del cáncer o la necesidad de un tratamiento agresivo.

Alternativamente, el valor indicador puede representar el nivel de expresión promedio de un conjunto de individuos de una población de cáncer diversa o un subconjunto de la población. Por ejemplo, puede determinarse el nivel de expresión promedio de un gen o panel de genes en un muestreo aleatorio de pacientes con cáncer de próstata. Este nivel de expresión promedio puede denominarse "valor indicador de umbral", esperándose que los pacientes que tienen una expresión de CCG mayor a este valor tengan un peor pronóstico que los que tienen una expresión menor a este valor.

Alternativamente, el valor indicador puede representar el nivel de expresión promedio de un marcador génico particular en una pluralidad de pacientes de entrenamiento (por ejemplo, pacientes con cáncer de próstata) con desenlaces similares cuyos datos clínicos y de seguimiento están disponibles y son suficientes para definir y categorizar a los pacientes por desenlace de enfermedad, por ejemplo, recidiva o pronóstico. Véanse, por ejemplo, los ejemplos a continuación. Por ejemplo, un "valor indicador de buen pronóstico" puede generarse a partir de una pluralidad de pacientes con cáncer de entrenamiento caracterizados como que tienen un "buen desenlace", por ejemplo, aquellos que no han tenido recidiva del cáncer cinco años (o diez años o más) después del tratamiento inicial, o aquellos que no han tenido progresión en su cáncer cinco años (o diez años o más) después del diagnóstico inicial. Un "valor indicador de mal pronóstico" puede generarse a partir de una pluralidad de pacientes con cáncer de entrenamiento definidos como que tienen un "mal desenlace", por ejemplo, aquellos que han tenido recidiva del cáncer en el plazo de cinco años (o diez años, etc.) después del tratamiento inicial, o aquellos que han tenido progresión en su cáncer en el plazo de cinco años (o diez años, etc.) después del diagnóstico inicial. Por tanto, un valor indicador de buen pronóstico de un gen particular puede representar el nivel de expresión promedio del gen particular en pacientes que tienen un "buen desenlace", mientras que un valor indicador de mal pronóstico de un gen particular representa el nivel de expresión promedio del gen particular en pacientes que tienen un "mal desenlace".

En algunas realizaciones de la invención, los métodos comprenden determinar la expresión de diez o más CCG, y si esta expresión está "aumentada", el paciente tiene un mal pronóstico. En el contexto de la invención, expresión "aumentada" de un CCG significa que el nivel de expresión del paciente o bien está elevado por encima de un valor indicador normal o indicador de umbral (por ejemplo, en al menos cierta cantidad de umbral) o bien es más próximo al "valor indicador de mal pronóstico" que al "valor indicador de buen pronóstico".

Por tanto, cuando el nivel determinado de expresión de un marcador génico relevante es más próximo al valor indicador de buen pronóstico del gen que al valor indicador de mal pronóstico del gen, entonces puede concluirse que es más probable que el paciente tenga un buen pronóstico, es decir, una probabilidad baja (o no aumentada) de recidiva del cáncer. Por otro lado, si el nivel determinado de expresión de un marcador génico relevante es más próximo al valor indicador de mal pronóstico del gen que al valor indicador de buen pronóstico del gen, entonces puede concluirse que es más probable que el paciente tenga un mal pronóstico, es decir, una probabilidad aumentada de recidiva del cáncer.

Alternativamente, los valores indicadores pueden determinarse de la siguiente manera: con el fin de asignar a los pacientes a grupos de riesgo, se establecerá un valor de umbral para la media de ciclo celular. El valor de umbral óptimo se selecciona basándose en la curva característica operativa del receptor (ROC), que representa gráficamente sensibilidad frente a (1-especificidad). Para cada incremento de la media de ciclo celular, se calculan la sensibilidad y la especificidad de la prueba usando ese valor como umbral. El umbral real será el valor que optimiza

estos parámetros según requisitos de los expertos (por ejemplo, qué grado de sensibilidad o especificidad se desea, etc.). El ejemplo 5 demuestra la determinación de un valor de umbral determinado y validado experimentalmente.

En algunas realizaciones, el nivel de expresión de un gen particular (particularmente CCG) se obtiene como número absoluto de copias de ARNm de una célula o muestra de tejido tumoral particular. En otras realizaciones, el nivel de expresión de un gen particular (particularmente CCG) se obtiene normalizando la cantidad de un producto de expresión de un gen particular de interés frente a la cantidad de expresión de un producto génico de normalización (por ejemplo, uno o más genes de mantenimiento). La normalización puede realizarse para generar un valor indicador o simplemente para ayudar a reducir el ruido de fondo cuando se determina el nivel de expresión del gen de interés. En una realización, por ejemplo, a la hora de determinar el nivel de expresión de un gen relevante según la presente invención, se mide la cantidad de un producto de expresión del gen (por ejemplo, ARNm, ADNc, proteína) dentro de una o más células, particularmente células tumorales, y se normaliza frente a la cantidad de producto(s) de expresión de un gen de normalización, o un conjunto de genes de normalización, dentro de la misma una o más células, para obtener el nivel de expresión del gen marcador relevante. Por ejemplo, cuando se usa un único gen como gen de normalización, puede usarse un gen de mantenimiento cuya expresión se determina que es independiente del desenlace/pronóstico del cáncer de próstata o no varía entre células normales y células de cáncer de próstata. También puede usarse un conjunto de tales genes de mantenimiento en el análisis de la expresión génica para proporcionar un conjunto de genes de normalización combinados. Los genes de mantenimiento se conocen bien en la técnica, con ejemplos que incluyen, pero no se limitan a, *GUSB* (glucuronidasa, beta), *HMBS* (hidroximetilbilano sintasa), *SDHA* (complejo de succinato deshidrogenasa, subunidad A, flavoproteína), *UBC* (ubiquitina C) y *YWHAZ* (proteína de activación de tirosina 3-monooxigenasa/triptófano 5-monooxigenasa, polipéptido zeta). Cuando se usa un conjunto de genes de normalización combinados en la normalización, puede promediarse la cantidad de expresión génica de tales genes de normalización, combinándose conjuntamente mediante adiciones directas o mediante un algoritmo definido. Como genes de normalización también pueden usarse genes distintos de los genes de mantenimiento. Otros ejemplos de genes de mantenimiento particularmente útiles incluyen los enumerados en la tabla A a continuación.

Tabla A

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

YWHAZ	7534	Hs00237047_m1	NM_003406.3
-------	------	---------------	-------------

* Subconjunto de genes de mantenimiento usados, por ejemplo, en el ejemplo 5.

Los paneles de CCG (por ejemplo, 2, 3, 4, 5, 6, 7, 8, 9 ó 10 o más CCG) pueden predecir con precisión el pronóstico, tal como se muestra en el ejemplo 3. Los expertos en la técnica están familiarizados con diversos modos de determinar la expresión de un panel de genes (es decir, una pluralidad de genes). Puede determinarse la expresión de un panel de genes determinando el nivel de expresión promedio (normalizado o absoluto) de la totalidad de genes de panel en una muestra obtenida de un paciente particular (o bien en toda la muestra o bien en un subconjunto de células de la muestra o bien en una única célula). Expresión aumentada en este contexto significará que la expresión promedio es mayor que el nivel de expresión promedio de estos genes en pacientes normales (o mayor que algún valor indicador que se ha determinado que representa el nivel de expresión promedio en una población de referencia tal como pacientes con el mismo cáncer). Alternativamente, puede determinarse la expresión de un panel de genes determinando el nivel de expresión promedio (normalizado o absoluto) de al menos un cierto número (por ejemplo, 10, 15, 20, 25, 30 o más) de los genes en el panel. Alternativamente, puede determinarse la expresión de un panel de genes determinando el número de copias absoluto de ARNm (o proteína) de todos los genes en el panel y sumando o promediando los mismos a lo largo de los genes.

Tal como se usa en el presente documento, “clasificar un cáncer” y “clasificación de cáncer” se refieren a determinar una o más características clínicamente relevantes de un cáncer y/o determinar un pronóstico particular de un paciente que tiene dicho cáncer. Por tanto “clasificar un cáncer” incluye, pero no se limita a: (i) evaluar el potencial metastásico, el potencial para metastatizarse a órganos específicos, el riesgo de recidiva y/o la evolución del tumor; (ii) evaluar el estadio tumoral; (iii) determinar el pronóstico del paciente en ausencia de tratamiento del cáncer; (iv) determinar el pronóstico de respuesta del paciente (por ejemplo, reducción del volumen tumoral o supervivencia libre de progresión) al tratamiento (por ejemplo, quimioterapia, radioterapia, cirugía para extirpar el tumor, etc.); (v) diagnóstico de la respuesta real del paciente al tratamiento actual y/o anterior; (vi) determinar un transcurso de tratamiento preferido para el paciente; (vii) pronóstico de la recaída del paciente después del tratamiento (o bien tratamiento en general o bien algún tratamiento particular); (viii) pronóstico de la esperanza de vida del paciente (por ejemplo, pronóstico de la supervivencia global), etc.

Por tanto, una “clasificación negativa” significa una característica clínica desfavorable del cáncer (por ejemplo, un mal pronóstico). Los ejemplos incluyen (i) un potencial metastásico, potencial para metastatizarse a órganos específicos y/o riesgo de recidiva aumentados; (ii) un estadio tumoral avanzado; (iii) un mal pronóstico del paciente en ausencia de tratamiento del cáncer; (iv) un mal pronóstico de la respuesta del paciente (por ejemplo, reducción del volumen tumoral o supervivencia libre de progresión) a un tratamiento particular (por ejemplo, quimioterapia, radioterapia, cirugía para extirpar el tumor, etc.); (v) un mal pronóstico de la recaída del paciente después del tratamiento (o bien tratamiento en general o bien algún tratamiento particular); (vi) un mal pronóstico de la esperanza de vida del paciente (por ejemplo, pronóstico de la supervivencia global), etc. En algunas realizaciones, un parámetro clínico asociado a recidiva (o una alta puntuación de nomograma) y una expresión aumentada de un CCG indican una clasificación negativa en el cáncer (por ejemplo, probabilidad aumentada de recidiva o progresión).

Tal como se comentó anteriormente, se piensa que una expresión de CCG elevada acompaña a células cancerosas que proliferan rápidamente (y, por tanto, más agresivas). Un cáncer de este tipo en un paciente a menudo significará que el paciente tiene una probabilidad aumentada de recidiva después del tratamiento (por ejemplo, las células cancerosas no destruidas o eliminadas mediante el tratamiento crecerán rápidamente de nuevo). Un cáncer de este tipo también puede significar que el paciente tiene una probabilidad aumentada de progresión del cáncer para una progresión más rápida (por ejemplo, las células que proliferan rápidamente harán que cualquier tumor crezca rápidamente, aumente su virulencia y/o se metastatice). Un cáncer de este tipo también puede significar que el paciente puede requerir un tratamiento relativamente más agresivo. Por tanto, en algunas realizaciones, la invención proporciona un método para clasificar el cáncer de próstata que comprende determinar el estado de un panel de genes que comprende al menos diez CCG de uno cualquiera de los paneles C-G, en el que un estado anómalo indica una probabilidad aumentada de recidiva o progresión. Tal como se comentó anteriormente, en algunas realizaciones, el estado que va a determinarse son los niveles de expresión génica. Por tanto, en algunas realizaciones, la invención proporciona un método para determinar el pronóstico del cáncer de próstata de un paciente que comprende determinar el nivel de expresión de un panel de genes que comprende al menos diez CCG de uno cualquiera de los paneles C-G, en el que una expresión elevada indica una probabilidad aumentada de recidiva o progresión del cáncer.

“Recidiva” y “progresión” son términos bien conocidos en la técnica y se usan en el presente documento según sus significados conocidos. Como un ejemplo, el significado de “progresión” puede ser dependiente del tipo de cáncer, significando progresión en cáncer de pulmón algo diferente de progresión en cáncer de próstata. Sin embargo, dentro de cada tipo y subtipo de cáncer, los expertos en la técnica entienden claramente “progresión”. Tal como se usa en el presente documento, un paciente tiene una “probabilidad aumentada” de alguna característica o algún desenlace clínico (por ejemplo, recidiva o progresión) si la probabilidad de que el paciente tenga la característica o el desenlace supera alguna probabilidad o algún valor de referencia. La probabilidad de referencia puede ser la probabilidad de la característica o el desenlace a lo largo de la población general de pacientes relevantes. Por

ejemplo, si la probabilidad de recidiva en la población general de cáncer de próstata es del X % y se ha determinado que un paciente particular tiene una probabilidad de recidiva del Y %, y si $Y > X$, entonces el paciente tiene una “probabilidad aumentada” de recidiva. Alternativamente, tal como se comentó anteriormente, puede determinarse un valor de umbral o referencia y puede compararse la probabilidad de recidiva de un paciente particular con ese umbral o esa referencia. Dado que predecir la recidiva y predecir la progresión son actividades pronósticas, “predecir el pronóstico” a menudo se usará en el presente documento para referirse a cualquiera o a ambos. En estos casos, un “mal pronóstico” generalmente se referirá a una probabilidad aumentada de recidiva, progresión, o ambas.

Tal como se muestra en el ejemplo 3, los CCG individuales pueden predecir el pronóstico bastante bien.

El ejemplo 3 también muestra que los paneles de CCG (por ejemplo, 2, 3, 4, 5 ó 6 CCG) pueden predecir con precisión el pronóstico. Por tanto, en algunos aspectos, la invención proporciona un método para clasificar el cáncer de próstata que comprende determinar el estado de un panel (es decir, una pluralidad) de genes que comprende una pluralidad de CCG. Por ejemplo, expresión aumentada en un panel de genes puede referirse al hecho de que el nivel de expresión promedio de todos los genes de panel en un paciente particular es mayor que el nivel de expresión promedio de estos genes en pacientes normales (o mayor que algún valor indicador que se ha determinado que representa el nivel de expresión promedio normal). Alternativamente, expresión aumentada en un panel de genes puede referirse a expresión aumentada en al menos un cierto número (por ejemplo, 10, 15, 20, 25, 30 o más) de los genes en el panel en comparación con el nivel de expresión promedio normal.

En algunas realizaciones, los CCG se eligen del grupo que consiste en los genes en los paneles C a G. En algunas realizaciones, el panel comprende al menos 10, 11, 12, 13, 14, 15, 20, 25, 30 o más de los genes en cualquiera de los paneles C a G. En algunas realizaciones, la invención proporciona un método para predecir el pronóstico que comprende determinar el estado de los CCG en los paneles C a G, en el que un estado anómalo indica un mal pronóstico.

En algunas de estas realizaciones, expresión elevada indica una probabilidad aumentada de recidiva o progresión. Por tanto, en una realización preferida, la invención proporciona un método para predecir el riesgo de recidiva o progresión del cáncer de próstata en un paciente que comprende determinar el estado de un panel de genes, tal como se define en las reivindicaciones, y un estado elevado para los CCG indica una probabilidad aumentada de recidiva o progresión.

Se han evaluado varios paneles de CCG (tabla 2 anteriormente y tablas 3 y 4 a continuación) para determinar su capacidad para predecir el pronóstico en varios cánceres diferentes. Los resultados de estos estudios se describen en los ejemplos 1 a 6 a continuación.

Tabla 3: “Panel C” evaluado en los ejemplos 1 a 4

Símbolo de gen	ID de gen de Entrez	Símbolo de gen	ID de gen de Entrez	Símbolo de gen	ID de gen de 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>CDCA8*</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 usaron como 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	ID de gen	Nombre	ID de gen	Nombre	ID de gen
<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>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 usaron como panel de subconjunto de 31 genes ("panel F") en el experimento descrito en el ejemplo 5.

Se ha determinado que la elección de los CCG individuales para un panel a menudo puede ser relativamente arbitraria. Dicho de otro modo, se ha hallado que la mayoría de los CCG son muy buenos sustitutos entre sí. Un modo para evaluar si los CCG particulares servirán bien en los métodos y las composiciones de la invención es evaluando su correlación con la expresión media de los CCG (por ejemplo, todos los CCG conocidos, un conjunto específico de CCG, etc.). Se espera que esos CCG que se correlacionan particularmente bien con la media funcionen bien, por ejemplo, porque estos reducirán el ruido en el ensayo. En la tabla 23 se proporciona una clasificación de CCG seleccionados según su correlación con la expresión de CCG media.

En las firmas de CCG, los CCG particulares sometidos a ensayo a menudo no es tan importante como el número total de CCG. El número de CCG sometidos a ensayo puede variar dependiendo de muchos factores, por ejemplo, restricciones técnicas, consideraciones de costes, la clasificación que esté realizándose, el cáncer que esté sometiéndose a prueba, el nivel deseado de poder predictivo, etc. Aumentar el número de CCG sometidos a ensayo en un panel según la invención es, como cuestión general, ventajoso porque, por ejemplo, un mayor grupo de ARNm que van a someterse a ensayo significa menos "ruido" provocado por valores atípicos y menor oportunidad de que un error de ensayo altere el poder predictivo global de la prueba. Sin embargo, el coste y otras consideraciones limitarán generalmente este número y es deseable hallar el número óptimo de CCG para una firma.

Se ha descubierto que el poder predictivo de una firma de CCG a menudo deja de aumentar significativamente más allá de un cierto número de CCG (véase la figura 1; ejemplo 1). Más específicamente, el número óptimo de CCG en una firma (n_0) puede hallarse siempre que sea cierto lo siguiente

$$(P_{n+1} - P_n) < C_0,$$

donde P es el poder predictivo (es decir, P_n es el poder predictivo de una firma con n genes y P_{n+1} es el poder predictivo de una firma con n genes más uno) y C_0 es alguna constante de optimización. El poder predictivo puede definirse de muchos modos conocidos por los expertos en la técnica, incluyendo, pero sin limitarse a, el valor de p de la firma. C_0 puede elegirla el experto basándose en sus restricciones específicas. Por ejemplo, si el coste no es un factor crítico y se desean niveles extremadamente altos de sensibilidad y especificidad, puede establecerse una C_0 muy baja de manera que sólo se menosprecian aumentos triviales en el poder predictivo. Por otro lado, si el coste es decisivo y son aceptables niveles moderados de sensibilidad y especificidad, puede establecerse una C_0 más alta de manera que sólo aumentos significativos en el poder predictivo justifican el aumento del número de genes en la firma.

Alternativamente, puede representarse gráficamente un gráfico de poder predictivo en función del número de genes (como en la figura 1) y tomarse la segunda derivada de esta representación gráfica. El punto en el que la segunda derivada disminuye hasta cierto valor predeterminado (C_0') puede ser el número óptimo de genes en la firma.

Los ejemplos 1 y 3 y la figura 1 ilustran la determinación empírica de los números óptimos de CCG en paneles de CCG. Se sometieron a prueba subconjuntos seleccionados aleatoriamente de los 31 CCG enumerados en la tabla 3 como firmas de CCG distintas y se determinó el poder predictivo (es decir, el valor de p) para cada uno. Tal como muestra la figura 1, los valores de p dejaron de mejorar significativamente entre aproximadamente 10 y aproximadamente 15 CCG, lo que indica que un número óptimo de CCG en un panel de pronóstico es de desde aproximadamente 10 hasta aproximadamente 15.

Se ha descubierto que los CCG son particularmente predictivos en ciertos cánceres. Por ejemplo, se ha determinado que los paneles de CCG son precisos a la hora de predecir la recidiva en el cáncer de próstata (ejemplos 1 a 5). Además, los CCG pueden determinar el pronóstico en cánceres de vejiga, cerebro, mama y pulmón, tal como se

resume en el ejemplo 6 y las tablas 21 y 22 a continuación.

Por tanto, la invención proporciona un método que comprende determinar el estado de un panel de genes que comprende al menos 10 genes elegidos del grupo de genes en uno cualquiera de los paneles C a G. En algunas realizaciones, el panel comprende al menos 15 genes elegidos del grupo de genes en uno cualquiera de los paneles C a G. En algunas realizaciones, el panel comprende la totalidad de los genes en uno cualquiera de los paneles C a G.

Se ha descubierto además que el estado de CCG se añade sinérgicamente a los parámetros clínicos a la hora de pronosticar el cáncer. En el caso de cáncer de próstata, por ejemplo, se ha descubierto que un nivel alto de expresión génica de uno cualquiera de los genes en los paneles C a F está asociado con un riesgo aumentado de recidiva o progresión del cáncer de próstata en pacientes cuya puntuación de nomograma clínico indica un riesgo relativamente bajo de recidiva o progresión. Por tanto, dado que la evaluación de los niveles de expresión de CCG puede detectar un riesgo aumentado no detectado usando los parámetros clínicos solos, la invención proporciona de manera general métodos que combinan evaluar al menos un parámetro clínico con evaluar el estado de al menos diez CCG de uno cualquiera de los paneles C-G.

Tal como muestra el ejemplo 3, incluso los CCG individuales se añaden a los parámetros clínicos a la hora de predecir la recidiva del cáncer. Por tanto, un aspecto de la invención proporciona un método de diagnóstico *in vitro* que comprende determinar al menos un parámetro clínico para un paciente con cáncer de próstata y determinar el estado de al menos diez CCG de uno cualquiera de los paneles C-G en una muestra obtenida del paciente. Sin embargo, evaluar el estado de múltiples CCG mejora incluso más el poder predictivo (tal como se muestra en el ejemplo 1). Por tanto, en algunas realizaciones, se determina el estado de una pluralidad de CCG (por ejemplo, 10, 15, 20, 25, 30, 40, 50 o más). En algunas realizaciones, el estado anómalo indica una probabilidad aumentada de recidiva o progresión. A menudo, el parámetro clínico es al menos algo independientemente predictivo de la recidiva o progresión, y la adición del estado de CCG mejora el poder predictivo. 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 el transcurso de la enfermedad y/o predecir el desenlace. Los ejemplos en cáncer generalmente incluyen estadio tumoral, grado tumoral, estado de los ganglios linfáticos, histología, estado de rendimiento, tipo de cirugía, márgenes quirúrgicos, tipo de tratamiento, y edad de aparición. En el cáncer de próstata, los médicos a menudo usan los niveles de PSA en sangre previos a la cirugía, el estadio (definido por el tamaño del tumor y evidencias 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 médicos a menudo usan el tamaño de la lesión indicadora en cm, la invasión, el número de ganglios implicados, y el grado.

A menudo, determinados parámetros clínicos se correlacionan con un carácter de enfermedad particular. Por ejemplo, en el cáncer generalmente así como en cánceres específicos, determinados parámetros clínicos se correlacionan con, por ejemplo, la probabilidad de recidiva o metástasis, el pronóstico de la supervivencia durante un determinado periodo de tiempo, la probabilidad de respuesta al tratamiento generalmente o a un tratamiento específico, etc. En el cáncer de próstata, algunos parámetros clínicos son de tal manera que su estado (presencia, ausencia, nivel, etc.) está asociado con la probabilidad aumentada de recidiva. Los ejemplos de tales parámetros asociados a la recidiva (algunos de los cuales, pero no todos, son específicos del cáncer de próstata) incluyen altos niveles de PSA (por ejemplo, mayores de 4 ng/ml), alta puntuación de Gleason, tamaño tumoral grande, evidencia de metástasis, estadio tumoral avanzado, grado nuclear, afectación ganglionar, edad temprana de aparición. Otros tipos de cáncer pueden tener parámetros diferentes correlacionados con la probabilidad de recidiva o progresión, y el estado de CCG, como medida de actividad proliferativa, se añade a estos parámetros para predecir el pronóstico en estos cánceres. Tal como se usa en el presente documento, "parámetro clínico asociado a recidiva" tiene su significado convencional para cada cáncer específico, con lo que los expertos en la técnica están bastante familiarizados. De hecho, los expertos en la técnica están familiarizados con diversos parámetros clínicos asociados a la recidiva más allá de los enumerados en el presente documento.

A menudo, un médico evaluará más de un parámetro clínico en un paciente y realizará una evaluación más completa para los caracteres de enfermedad de interés. El ejemplo 5 muestra cómo puede añadirse el estado de CCG a una agrupación particular de parámetros clínicos usados para determinar el riesgo de recidiva en el cáncer de próstata. Los parámetros clínicos en el ejemplo 5 incluyen variables binarias para la enfermedad confinada en órgano y puntuación de Gleason menor 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 posterior a PR (es decir, nomograma de Kattan-Stephenson) excepto para el año de PR y los dos componentes de la puntuación de Gleason. Por tanto, en algunas realizaciones, se evalúan al menos dos parámetros clínicos (por ejemplo, dos de los parámetros enumerados anteriormente) junto con el nivel de expresión de al menos diez CCG.

Un modo en el que los médicos utilizan parámetros clínicos individuales, pero más a menudo múltiples, es con la ayuda de nomogramas. En el ámbito clínico, los nomogramas son representaciones (a menudo visuales) de una correlación entre uno o más parámetros y uno o más caracteres de paciente o de enfermedad. Un ejemplo de un nomograma clínico prevalente usado para determinar la probabilidad de recidiva de un paciente con cáncer de

próstata se describe en Kattan *et al.*, J. CLIN. ONCOL. (1999) 17:1499-1507, y se actualiza en Stephenson *et al.*, J. CLIN. ONCOL. (2005) 23:7005-7012 ("nomograma de Kattan-Stephenson"). Este nomograma evalúa un paciente asignando un valor de punto a cada uno de varios parámetros clínicos (año de PR, márgenes quirúrgicos, extensión extracapsular, invasión de vesículas seminales, afectación ganglionar, puntuación primaria de Gleason, puntuación secundaria de Gleason y nivel de PSA preoperatorio), sumando los puntos para un paciente en una puntuación de nomograma, y luego prediciendo la probabilidad del paciente de estar libre de recidiva a intervalos de tiempo variables (hasta 10 años) basándose en esta puntuación de nomograma. Un ejemplo de un nomograma prevalente usado para determinar el pronóstico de la supervivencia de un paciente con cáncer de mama es el índice pronóstico de Nottingham (NPI). Véase, por ejemplo, Galea *et al.*, BREAST CANCER RES. & TREAT. (1992) 22:207-19.

Se ha descubierto que determinar el estado de un CCG en una muestra obtenida de un paciente con cáncer de próstata, junto con la puntuación de nomograma de Kattan-Stephenson del paciente, es un mejor factor predisponente de la supervivencia libre de recidiva a los 10 años que la puntuación de nomograma sola. Véanse, por ejemplo, los ejemplos 2 y 5 a continuación. Específicamente, añadir el estado de CCG al nomograma de Kattan-Stephenson detecta pacientes con riesgo significativamente aumentado de recidiva que el nomograma por sí solo no detecta. La tabla 3 anterior proporciona un panel a modo de ejemplo de 31 CCG (panel C) y un panel de subconjuntos de 26 CCG (panel D, mostrado con *) determinados en el ejemplo 2 para mostrar la sinergia predictiva con el nomograma de Kattan-Stephenson en el pronóstico de cáncer de próstata. También se ha descubierto que determinar el estado de un CCG en una muestra obtenida de un paciente con cáncer de mama, junto con la puntuación de NPI del paciente, es un mejor factor predisponente de pronóstico que la puntuación de NPI sola. Véase, por ejemplo, el ejemplo 6 a continuación. Específicamente, añadir el estado de CCG al nomograma de NPI detecta pacientes con riesgo significativamente aumentado de recidiva que el nomograma por sí solo no detecta. Se determinó en el ejemplo 2 que los paneles B, C y D muestran una sinergia predictiva con el nomograma de NPI en el pronóstico de cáncer de mama.

Por tanto, otro aspecto de la invención proporciona un método *in vitro* que comprende determinar una puntuación de nomograma clínico (por ejemplo, puntuación de nomograma de Kattan-Stephenson o de NPI) para un paciente con cáncer de próstata y determinar el estado de al menos diez CCG de uno cualquiera de los paneles C-G en una muestra obtenida del paciente. El ejemplo 3 ilustra la determinación empírica del poder predictivo de CCG individuales y de varios paneles de CCG de tamaño variable a lo largo del nomograma de Kattan-Stephenson. Se sometieron a prueba subconjuntos seleccionados aleatoriamente de los 31 CCG enumerados en la tabla 3 como firmas de CCG distintas y se determinó el poder predictivo (es decir, el valor de p) para cada uno. Tal como muestra la figura 1, las firmas de CCG de 2, 3, 4, 5, 6, 10, 15, 20, 25 y 26 genes añaden, cada una, poder predictivo al nomograma. Por tanto, la invención proporciona un método para determinar si un paciente con cáncer de próstata tiene una probabilidad aumentada de recidiva que comprende determinar el estado de un panel de genes que comprende al menos 10, 15, 20, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 40, 45, 50, 60, 70, 80, 90 ó 100 o más CCG, en el que un estado elevado (por ejemplo, expresión aumentada) para los CCG indica una probabilidad aumentada de recidiva. En algunas realizaciones, el método comprende además determinar una puntuación de nomograma clínico del paciente.

A menudo, los nomogramas clínicos para el cáncer se diseñan de manera que un valor particular (por ejemplo, puntuación alta) se correlaciona con un riesgo aumentado de recidiva. Un estado de CCG elevado (por ejemplo, expresión o actividad aumentada) también se correlaciona con un riesgo aumentado. Por tanto, en algunas realizaciones, la invención proporciona un método para determinar si un paciente con cáncer de próstata tiene una probabilidad aumentada de recidiva o progresión que comprende determinar una puntuación de nomograma clínico para el paciente y determinar el estado de al menos diez CCG de uno cualquiera de los paneles C-G en una muestra obtenida del paciente, en el que una puntuación de nomograma alta y/o un estado de CCG elevado indican que el paciente tiene una probabilidad aumentada de recidiva o progresión.

En algunas realizaciones, esta evaluación se realiza antes de la prostatectomía radical (por ejemplo, usando una muestra de biopsia de próstata), mientras que en algunas realizaciones se realiza después (por ejemplo, usando la muestra de próstata resecada). En algunas realizaciones, una muestra de una o más células se obtiene de un paciente con cáncer de próstata antes o después del tratamiento para su análisis según la presente invención. El tratamiento de cáncer de próstata aplicado actualmente en la técnica incluye, por ejemplo, prostatectomía, radioterapia, terapia hormonal (por ejemplo, usando antagonistas de GnRH, agonistas de GnRH, antiandrógenos), quimioterapia, y ultrasonidos enfocados de alta intensidad. En algunas realizaciones, una o más células tumorales de próstata de tejido de cáncer de próstata se obtienen de un paciente con cáncer de próstata durante la biopsia o prostatectomía y se usan para su análisis en el método de la presente invención.

A menudo, los resultados de cualquier análisis según la invención se comunicarán a los médicos, asesores genéticos y/o pacientes (u otras partes interesadas tales como investigadores) en una forma transmisible que puede comunicarse o transmitirse a cualquiera de las partes anteriores. Una forma de este tipo puede variar y puede ser tangible o intangible. Los resultados pueden representarse en declaraciones descriptivas, diagramas, fotografías, esquemas, imágenes o cualquier otra forma visual. Por ejemplo, para explicar los resultados pueden usarse gráficos que muestran el nivel de expresión o de actividad o la información de variación de secuencia para diversos genes.

Los diagramas que muestran tal información para gen(es) diana adicional(es) también son útiles para indicar algunos resultados de prueba. Las declaraciones y formas visuales pueden grabarse en un medio tangible tal como papeles, medios legibles por ordenador tales como discos flexibles, discos compactos, etc., o en un medio intangible, por ejemplo, un medio electrónico en forma de correo electrónico o sitio web en Internet o Intranet. Además, los resultados también pueden grabarse en forma de sonido y transmitirse a través de cualquier medio adecuado, por ejemplo, líneas de cable analógicas o digitales, cables de fibra óptica, etc., a través del teléfono, facsímil, teléfono móvil inalámbrico, teléfono por Internet, y similares.

Por tanto, la información y los datos sobre un resultado de prueba pueden producirse en cualquier parte del mundo y transmitirse a una ubicación diferente. Como un ejemplo ilustrativo, cuando se lleva a cabo un ensayo de nivel de expresión, nivel de actividad o secuenciación (o genotipado) fuera de los Estados Unidos, la información y los datos sobre un resultado de prueba pueden generarse, emitirse en una forma transmisible tal como se describió anteriormente, y luego importarse a los Estados Unidos. Por consiguiente, la presente divulgación también abarca un método para producir una forma transmisible de información sobre al menos uno de (a) nivel de expresión para al menos una muestra de paciente. El método comprende las etapas de (1) determinar (a) anterior según los métodos 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 método de este tipo.

A menudo se implementarán técnicas para analizar tales datos de expresión, actividad y/o secuencia (de hecho, cualquier dato obtenido según la invención) usando 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 de análisis basada en ordenador puede implementarse en cualquier lenguaje y/o navegador adecuado. Por ejemplo, puede implementarse con lenguaje C y preferiblemente usando lenguajes de programación de alto nivel orientados a objetos tales como Visual Basic, Small Talk, C++, y similares. La aplicación puede escribirse para adecuarse a entornos tales como el entorno de Microsoft Windows™ incluyendo Windows™ 98, Windows™ 2000, Windows™ NT, y similares. Además, la aplicación también puede escribirse para el entorno de Macintosh™, SUN™, UNIX o LINUX. Además, las etapas funcionales también pueden implementarse usando un lenguaje de programación universal o independiente de plataforma. Los ejemplos de tales lenguajes de programación de múltiples plataformas incluyen, pero no se limitan a, lenguaje de marcas de hipertexto (HTML), JAVA™, JavaScript™, lenguaje de programación Flash, interfaz de entrada común/lenguaje de consulta estructurado (CGI/SQL), lenguaje práctico de extracción e informe (PERL), AppleScript™ y otros lenguajes de guiones de sistema, lenguaje de programación/lenguaje de consulta estructurado (PL/SQL), y similares. Pueden usarse navegadores habilitados para Java™ o JavaScript™ tales como HotJava™, Microsoft™ Explorer™ o Netscape™. Cuando se usan páginas web de contenido activo, pueden incluir miniprogramas (*applets*) de Java™ o controles de ActiveX™ u otras tecnologías de contenido activo.

La función de análisis también puede incluirse en productos de programa informático y usarse en los sistemas descritos anteriormente u otros sistemas basados en ordenador o en Internet. Por consiguiente, otro aspecto se refiere a un producto de programa informático que comprende un medio utilizable con ordenador que tiene códigos o instrucciones de programa legibles por ordenador incluidos en el mismo para permitir que un procesador lleve a cabo el análisis del estado de genes. Estas instrucciones de programa informático pueden cargarse en un ordenador u otro aparato programable para producir una máquina, de manera 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 de programa informático también pueden almacenarse en un medio o una memoria legible por ordenador que puede dirigir un ordenador u otro aparato programable a funcionar de una manera particular, de manera que las instrucciones almacenadas en el medio o la memoria legible por ordenador producen un artículo de fabricación que incluye un medio de instrucción que implementa el análisis. Las instrucciones de programa informático también pueden cargarse en un ordenador u otro aparato programable para hacer que se realicen una serie de etapas de operación en el ordenador u otro aparato programable para producir un procedimiento implementado por ordenador de manera que las instrucciones que se ejecutan en el ordenador u otro aparato programable proporcionan etapas para implementar las funciones o etapas descritas anteriormente.

Por tanto, un aspecto proporciona un sistema para determinar si un paciente tiene una probabilidad aumentada de recidiva. En términos generales, el sistema comprende (1) un medio informático para recibir, almacenar y/o recuperar datos del estado de genes de un paciente (por ejemplo, nivel de expresión, nivel de actividad, variantes) y opcionalmente datos de parámetros clínicos (por ejemplo, puntuación de Gleason, puntuación de nomograma); (2) un medio informático para consultar estos datos de paciente; (3) un medio informático para concluir si existe una probabilidad aumentada de recidiva basándose en estos datos de paciente; y (4) un medio informático para emitir/visualizar esta conclusión. En algunos aspectos, este medio para emitir la conclusión puede comprender un medio informático para informar a un profesional sanitario sobre la conclusión.

Un ejemplo de un sistema de este tipo es el sistema [600] informático ilustrado en la figura 6. El sistema [600] informático puede incluir al menos un módulo [630] de entrada para introducir datos de paciente en el sistema [600] informático. El sistema [600] informático puede incluir al menos un módulo [624] de salida para indicar si un paciente tiene una probabilidad aumentada o disminuida de respuesta y/o indicar tratamientos sugeridos determinados por el

sistema [600] informático. El sistema [600] informático puede incluir al menos un módulo [606] de memoria en comunicación con el al menos un módulo [630] de entrada y el al menos un módulo [624] de salida.

El al menos un módulo [606] de memoria puede incluir, por ejemplo, un dispositivo [608] de almacenamiento extraíble que puede estar en diversas formas, incluyendo, pero sin limitarse a, un dispositivo de banda magnética, un dispositivo de disco flexible, un dispositivo de VCD, un dispositivo de DVD, un dispositivo de disco óptico, etc. El dispositivo [608] de almacenamiento extraíble puede ser compatible con una unidad [610] de almacenamiento extraíble de manera que puede leer a partir de y/o escribir en la unidad [610] de almacenamiento extraíble. La unidad [610] de almacenamiento extraíble puede incluir un medio de almacenamiento utilizable con ordenador que tiene almacenados en el mismo códigos o instrucciones de programa legibles por ordenador y/o datos legibles por ordenador. Por ejemplo, la unidad [610] de almacenamiento extraíble puede almacenar datos de paciente. En la técnica se conocen bien ejemplos de unidades [610] de almacenamiento extraíbles, incluyendo, pero sin limitarse a, discos flexibles, bandas magnéticas, discos ópticos, y similares. El al menos un módulo [606] de memoria también puede incluir un dispositivo [612] de disco duro que puede usarse para almacenar códigos o instrucciones de programa legibles por ordenador y/o datos legibles por ordenador.

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

El sistema [600] informático puede incluir al menos un módulo [602] de procesador. Debe entenderse que el al menos un módulo [602] de procesador puede consistir en cualquier número de dispositivos. El al menos un módulo [602] de procesador 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 [602] de procesador puede incluir otro dispositivo lógico tal como un procesador de DMA (acceso directo a memoria), un dispositivo de procesador de comunicación integrada, un dispositivo de VLSI (integración a escala muy grande) personalizado o un dispositivo de ASIC (circuito integrado específico de aplicación). Además, el al menos un módulo [602] de procesador puede incluir cualquier otro tipo de circuitos analógicos o digitales que están diseñados para realizar las funciones de procesamiento descritas en el presente documento.

Tal como se muestra en la figura 6, en el sistema [600] informático, el al menos un módulo [606] de memoria, el al menos un módulo [602] de procesador y el módulo [618] de memoria secundario están todos ellos unidos operativamente entre sí a través de una infraestructura [620] de comunicación, que puede ser un bus de comunicaciones, un tablero de sistema, una barra cruzada, etc. A través de la infraestructura [620] de comunicación, pueden transferirse e intercambiarse códigos o instrucciones del programa informático o datos legibles por ordenador. La interfaz [626] de entrada puede conectar de manera operable el al menos un módulo [626] de entrada a la infraestructura [620] de comunicación. Del mismo modo, la interfaz [622] de salida puede conectar de manera operable el al menos un módulo [624] de salida a la infraestructura [620] de comunicación.

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

El al menos un módulo [606] de memoria puede estar configurado para almacenar datos de paciente introducidos a través del al menos un módulo [630] de entrada y procesados a través del al menos un módulo [602] de procesador. Los datos de paciente pueden incluir el nivel de expresión, el nivel de actividad, el número de copias y/o la información de secuencia para *PTEN* y/o un CCG. Los datos de paciente también pueden incluir parámetros clínicos relevantes para la enfermedad del paciente. También pueden introducirse al sistema cualesquiera otros datos de paciente que un médico podría encontrar útiles al tomar decisiones/hacer recomendaciones de tratamiento, incluyendo, pero sin limitarse a, la edad, el sexo y la raza/etnia y datos del estilo de vida tales como información sobre la dieta. Otros posibles tipos de datos de paciente incluyen síntomas experimentados en la actualidad o anteriormente, antecedentes de enfermedades del paciente, medicamentos, y procedimientos médicos.

El al menos un módulo [606] de memoria puede incluir un método implementado por ordenador almacenado en el mismo. El al menos un módulo [602] de procesador puede usarse para ejecutar software o códigos de instrucción legibles por ordenador del método implementado por ordenador. El método implementado por ordenador puede estar configurado, basándose en los datos de paciente, para indicar si el paciente tiene una probabilidad aumentada de recidiva, progresión o respuesta a cualquier tratamiento particular, para generar una lista de posibles

tratamientos, etc.

El método implementado por ordenador puede estar configurado para identificar si un paciente tiene o no una probabilidad aumentada de recidiva o progresión. Por ejemplo, el método implementado por ordenador puede estar configurado para informar a un médico que un paciente particular tiene una probabilidad aumentada de recidiva. De manera alternativa o adicional, el método implementado por ordenador puede estar configurado para sugerir realmente un transcurso de tratamiento particular basándose en respuestas a/resultados para diversas consultas.

La figura 7 ilustra un método [700] implementado por ordenador que puede implementarse con el sistema [600] informático. El método [700] comienza con una de tres consultas ([710], [711], [712]), o bien de manera secuencial o bien de manera sustancialmente simultánea. Si la respuesta a/resultado para cualquiera de estas consultas es "Sí" [720], el método concluye [730] que el paciente tiene una probabilidad aumentada de recidiva. Si la respuesta a/resultado para todas estas consultas es "No" [721], el método concluye [731] que el paciente no tiene una probabilidad aumentada de recidiva. El método [700] puede proceder entonces con más consultas, hacer una recomendación ([740], [741]) de tratamiento particular, o simplemente terminar.

Cuando las consultas se realizan de manera secuencial, pueden realizarse en el orden sugerido por la figura 7 o en cualquier otro orden. Si se realizan consultas posteriores también puede depender de los resultados/respuestas para consultas previas. En algunas realizaciones del método ilustrado en la figura 7, por ejemplo, el método pregunta en primer lugar sobre parámetros clínicos [712], y si el paciente tiene uno o más parámetros clínicos que identifican que el paciente tiene un riesgo aumentado de recidiva, entonces el método lo concluye [730] u opcionalmente lo confirma consultando el estado de CCG, mientras que si el paciente no tiene tales parámetros clínicos, entonces el método procede a preguntar sobre el estado de CCG [711]. Opcionalmente, si el estado de CCG no está elevado, entonces el método puede continuar preguntando sobre el estado de *PTEN* [710]. Tal como se mencionó anteriormente, el orden previo de consultas puede modificarse. En algunas realizaciones, una respuesta de "Sí" a una consulta (por ejemplo, [712]) da lugar a una o más de las consultas restantes para confirmar que el paciente tiene un riesgo aumentado de recidiva.

En algunos aspectos, el método [700] implementado por ordenador es abierto. Dicho de otro modo, la primera etapa [710, 711 y/o 712] aparente en la figura 7 en realidad puede formar parte de un procedimiento más largo y, dentro de este procedimiento más largo, no es necesario que sea la primera etapa/consulta. También pueden añadirse etapas adicionales a los métodos principales comentados anteriormente. Estas etapas adicionales incluyen, pero no se limitan a, informar a un profesional sanitario (o al propio paciente) sobre la conclusión a la que se llega; combinar la conclusión a la que se llega mediante el método [700] ilustrado con otros hechos u otras conclusiones para llegar a cierta conclusión adicional o refinada sobre el diagnóstico, pronóstico, tratamiento, etc., del paciente; hacer una recomendación para el tratamiento (por ejemplo, "el paciente debe/no debe someterse a prostatectomía radical"); consultas adicionales sobre biomarcadores adicionales, parámetros clínicos, u otra información útil del paciente (por ejemplo, edad en el momento del diagnóstico, salud general del paciente, etc.).

Con respecto al método [700] implementado por ordenador anterior, las respuestas a las consultas pueden determinarse mediante el método que establece una búsqueda de la respuesta en los datos de paciente. Por ejemplo, para responder a las consultas [710, 711, 712] respectivas, puede realizarse una búsqueda en los datos de paciente para el estado de *PTEN* (por ejemplo, cribado de mutaciones o IHC de *PTEN*), el estado de CCG (por ejemplo, datos de nivel de expresión de CCG), o parámetros clínicos (por ejemplo, puntuación de Gleason, puntuación de nomograma, etc.). Si aún no se ha realizado una comparación de este tipo, el método puede comparar estos datos con cierta referencia con el fin de determinar si el paciente tiene un estado anómalo (por ejemplo, elevado, bajo, negativo). De manera adicional o alternativa, el método puede presentar una o más de las consultas [710, 711, 712] a un usuario (por ejemplo, un médico) del sistema [100] informático. Por ejemplo, las consultas [710, 711, 712] pueden presentarse a través de un módulo [624] de salida. El usuario puede responder entonces "Sí" o "No" a través de un módulo [630] de entrada. El método puede proceder entonces basándose en la respuesta recibida. Del mismo modo, las conclusiones [730, 731] pueden presentarse a un usuario del método implementado por ordenador a través de un módulo [624] de salida.

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

La presente invención también puede hacer uso de diversos productos de programa informático y software para una

variedad de propósitos, tales como diseño de sondas, gestión de datos, análisis, y operación instrumental. Véanse las patentes estadounidenses 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. Además, se muestran métodos para proporcionar información genética sobre redes tales como Internet en las patentes estadounidenses con n.ºs de serie 10/197.621 (publicación estadounidense n.º 20030097222); 10/063.559 (publicación estadounidense n.º 20020183936), 10/065.856 (publicación estadounidense n.º 20030100995); 10/065.868 (publicación estadounidense n.º 20030120432); 10/423.403 (publicación estadounidense n.º 20040049354).

En un aspecto, la presente invención proporciona un método para determinar un transcurso de tratamiento para un paciente con cáncer de próstata tal como se define en la reivindicación 11.

Si un tratamiento es agresivo o no dependerá generalmente del tipo de cáncer, la edad del paciente, etc. Por ejemplo, en el cáncer de mama, la quimioterapia complementaria es un tratamiento agresivo habitual proporcionado para complementar los tratamientos habituales menos agresivos de cirugía y terapia hormonal. Los expertos en la técnica están familiarizados con otros diversos tratamientos agresivos y menos agresivos para cada tipo de cáncer. Los expertos en la técnica entienden bien "tratamiento activo" en el cáncer de próstata y, tal como se usa en el presente documento, tiene el significado convencional en la técnica. En términos generales, tratamiento activo en el cáncer de próstata es cualquier otra cosa distinta de "conducta expectante". El tratamiento activo aplicado actualmente en la técnica del tratamiento de cáncer de próstata incluye, por ejemplo, prostatectomía, radioterapia, terapia hormonal (por ejemplo, antagonistas de GnRH, agonistas de GnRH, antiandrógenos), quimioterapia, 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 tanto, es habitual que los médicos, dependiendo de la edad y la salud general del varón diagnosticado con cáncer de próstata, recomienden un régimen de "conducta expectante".

"Conducta expectante", también denominada "vigilancia activa", también tiene su significado convencional en la técnica. De manera general, esto significa observación y monitorización regular sin tratamiento invasivo. Conducta expectante se usa algunas veces, por ejemplo, cuando se halla un cáncer de próstata de estadio temprano y de crecimiento lento en un varón más anciano. La conducta expectante también puede sugerirse cuando los riesgos de cirugía, radioterapia o terapia hormonal superan los posibles beneficios. Pueden iniciarse otros tratamientos si se desarrollan síntomas, o si existen signos de que el crecimiento del cáncer está acelerándose (por ejemplo, aumento rápido de PSA, aumento de la puntuación de Gleason en biopsia repetida, etc.).

Aunque los varones que eligen conducta expectante evitan los riesgos de cirugía y radiación, la conducta expectante conlleva sus propios riesgos, por ejemplo, riesgo aumentado de metástasis. Para varones más jóvenes, un ensayo de vigilancia activa puede no significar evitar el tratamiento completamente, sino que puede permitir de manera razonable un retraso de algunos años o más, tiempo durante el cual puede evitarse el impacto del tratamiento activo sobre la calidad de vida. Los datos publicados hasta la fecha sugieren que los varones cuidadosamente seleccionados no se perderán un margen para la cura con este enfoque. Los problemas de salud adicionales que se desarrollan con el paso de la edad durante el periodo de observación también pueden hacer que sea más difícil someterse a cirugía y radioterapia. Por tanto, es clínicamente importante determinar cuidadosamente qué pacientes con cáncer de próstata son buenos candidatos para conducta expectante y qué pacientes deben recibir tratamiento activo.

Por tanto, se divulga un método para tratar a un paciente con cáncer de próstata o proporcionar pautas para el tratamiento de un paciente. En este método, se determina el estado de al menos un CCG (por ejemplo, aquellos en la tabla 1 o en los paneles A a G), al menos un parámetro clínico asociado a recidiva, y/o el estado de *PTEN*, y (a) se recomienda, inicia o continúa el tratamiento activo si una muestra del paciente tiene un estado elevado para al menos un CCG, el paciente tiene al menos un parámetro clínico asociado a recidiva, y/o un estado de *PTEN* bajo/negativo, o (b) se recomienda/inicia/continúa la conducta expectante si el paciente no tiene ninguno de estado elevado para al menos un CCG, parámetro clínico asociado a recidiva, y estado de *PTEN* bajo/negativo. En ciertos aspectos, el estado de CCG, el/los parámetro(s) clínico(s) y el estado de *PTEN* pueden indicar no sólo qué tratamiento activo se recomienda, sino qué tratamiento activo particular es preferible para el paciente (incluyendo tratamientos relativamente agresivos tales como, por ejemplo, PR y/o terapia complementaria).

En general, la terapia complementaria (por ejemplo, quimioterapia, radioterapia, HIFU, terapia hormonal, etc., después de la prostatectomía o radioterapia) no es el tratamiento habitual en el cáncer de próstata. Sin embargo, según la presente invención, los médicos pueden ser capaces de determinar qué pacientes con cáncer de próstata tienen una enfermedad particularmente agresiva y, por tanto, deben recibir terapia complementaria. Se divulga un método para tratar a un paciente (por ejemplo, un paciente con cáncer de próstata) que comprende determinar el estado de al menos un CCG (por ejemplo, aquellos en la tabla 1 o en los paneles A a G), el estado de al menos un parámetro clínico asociado a recidiva y/o el estado de *PTEN* e iniciar la terapia complementaria después de la prostatectomía o radioterapia si una muestra del paciente tiene un estado elevado para al menos un CCG, el paciente tiene al menos un parámetro clínico asociado a recidiva y/o el paciente tiene un estado de *PTEN* bajo/negativo.

Además, se divulgan composiciones para su uso en los métodos anteriores. Tales composiciones incluyen, pero no se limitan a, sondas de ácido nucleico que se hibridan con *PTEN* o un CCG (o con cualquier ácido nucleico codificado por los mismos o complementario a los mismos); cebadores y pares de cebadores de ácido nucleico adecuados para amplificar la totalidad o una porción de *PTEN* o un CCG o cualquier ácido nucleico codificado por los mismos; anticuerpos que se unen inmunológicamente a un polipéptido codificado por *PTEN* o un CCG; conjuntos de sondas que comprenden una pluralidad de dichas sondas de ácido nucleico, cebadores de ácido nucleico, anticuerpos y/o polipéptidos; micromatrices que comprenden cualesquiera de éstos; kits que comprenden cualesquiera de éstos; etc.

La divulgación proporciona una sonda que comprende un oligonucleótido aislado capaz de hibridarse selectivamente con *PTEN* o al menos uno de los genes en la tabla 1 o en los paneles A a G. Los términos "sonda" y "oligonucleótido" (también "oligo"), cuando se usan en el contexto de ácidos nucleicos, se refieren indistintamente a una secuencia o un fragmento de ácido nucleico relativamente corto. La divulgación también proporciona cebadores útiles en los métodos de la invención. En las condiciones adecuadas y con los reactivos acompañantes adecuados, los "cebadores" son sondas capaces de amplificar selectivamente un ácido nucleico diana (por ejemplo, un gen diana). En el contexto de ácidos nucleicos, "sonda" se usa en el presente documento para abarcar "cebador" puesto que los cebadores generalmente también pueden servir como sondas.

La sonda puede tener generalmente cualquier longitud/tamaño adecuado. En algunos aspectos, la sonda tiene una longitud de aproximadamente 8 a 200, de 15 a 150, de 15 a 100, de 15 a 75, de 15 a 60, o de 20 a 55 bases de longitud. Pueden marcarse con marcadores detectables con cualquier marcador de detección adecuado, incluyendo, pero sin limitarse a, isótopos radiactivos, fluoróforos, biotina, enzimas (por ejemplo, fosfatasa alcalina), sustratos enzimáticos, ligandos y anticuerpos, etc. Véanse Jablonski *et al.*, NUCLEIC ACIDS RES. (1986) 14:6115-6128; Nguyen *et al.*, BIOTECHNIQUES (1992) 13:116-123; Rigby *et al.*, J. 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 oligonucleotídicas son convencionales en la técnica.

Las sondas pueden usarse en las técnicas de hibridación/amplificación/detección comentadas anteriormente. Por tanto, algunos aspectos comprenden conjuntos de sondas adecuados para su uso en una micromatriz para detectar, amplificar y/o cuantificar *PTEN* y/o una pluralidad de CCG. En algunos aspectos, los conjuntos de sondas tienen una cierta proporción de sus sondas dirigidas a los CCG, por ejemplo, un conjunto de sondas que consiste en el 10 %, el 20 %, el 30 %, el 40 %, el 50 %, el 55 %, el 60 %, el 65 %, el 70 %, el 75 %, el 80 %, el 85 %, el 90 %, el 95 %, el 96 %, el 97 %, el 98 %, el 99 % o el 100 % de sondas específicas de los CCG. En algunos aspectos, el conjunto de sondas comprende sondas dirigidas 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 o más, o la totalidad, de los genes en la tabla 1 o en los paneles A a G. Tales conjuntos de sondas pueden 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 ó 1.000.000 o más de sondas diferentes. En otros aspectos, los conjuntos de sondas comprenden 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 CCG en la tabla 1 o en los paneles A a G.

En otro aspecto, se proporciona un kit para poner en práctica el pronóstico de la presente invención. El kit puede incluir un portador para los diversos componentes del kit. El portador puede ser un recipiente o soporte, en forma de, por ejemplo, bolsa, caja, tubo, rejilla, y opcionalmente está compartimentalizado. El portador puede definir un confinamiento cerrado con propósitos de seguridad durante el envío y el almacenamiento. El kit incluye diversos componentes útiles para determinar el estado de uno o más CCG y uno o más marcadores de genes de mantenimiento, usando las técnicas de detección comentadas anteriormente. Por ejemplo, el kit puede incluir oligonucleótidos que se hibridan específicamente bajo alta rigurosidad con ARNm o ADNc de los genes en la tabla 1 o en los paneles A a G. Tales oligonucleótidos pueden usarse como cebadores de PCR en reacciones de RT-PCR, o como sondas de hibridación. En algunos aspectos, el kit comprende reactivos (por ejemplo, sondas, cebadores y/o anticuerpos) para determinar el nivel de expresión de un panel de genes, en el que dicho panel comprende al menos el 25 %, el 30 %, el 40 %, el 50 %, el 60 %, el 75 %, el 80 %, el 90 %, el 95 %, el 99 % o el 100 % de los CCG (por ejemplo, los CCG en la tabla 1 o en cualquiera de los paneles A a G). En algunos aspectos, el kit consiste en reactivos (por ejemplo, sondas, cebadores y/o anticuerpos) para determinar el nivel de expresión de no más de 2500 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 CCG (por ejemplo, los CCG en la tabla 1 o en cualquiera de los paneles A a G).

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

En otro aspecto, el kit de detección contiene uno o más anticuerpos selectivamente inmunorreactivos con una o más

proteínas codificadas por *PTEN* o uno o más CCG u opcionalmente cualquier marcador adicional. Los ejemplos incluyen anticuerpos que se unen inmunológicamente a *PTEN* o a una proteína codificada por un gen en la tabla 1 o en los paneles A a G. Los métodos para producir y usar tales anticuerpos se han descrito anteriormente con detalle.

En el kit de detección también pueden incluirse otros diversos componentes útiles en las técnicas de detección. Los ejemplos de tales componentes incluyen, pero no se limitan a, polimerasa Taq, desoxirribonucleótidos, didesoxirribonucleótidos, otros cebadores adecuados para la amplificación de una secuencia de ADN diana, ARNasa A, y similares. Además, el kit de detección incluye preferiblemente instrucciones para usar el kit para poner en práctica el método de pronóstico de la presente invención usando muestras humanas.

EJEMPLO 1

Se sometió a prueba la siguiente firma de genes del ciclo celular (CCG) para predecir el tiempo hasta la recidiva química después de la prostatectomía radical.

Firma de recidiva de próstata de 31 CCG		
<i>AURKA</i>	<i>DTL</i>	<i>PTTG1</i>
<i>BUB1</i>	<i>FOXO1</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>	

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

La media de la expresión de CCG es robusta al error de medición y a la variación individual entre genes. Con el fin de determinar el número óptimo de genes del ciclo celular para la firma, se sometió a prueba el poder predictivo de la media para conjuntos seleccionados aleatoriamente de desde 1 hasta 30 de los CCG enumerados anteriormente. Esta simulación demostró que existe un número umbral de CCG en un panel que proporciona un poder predictivo significativamente mejorado.

EJEMPLO 2

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

Pacientes y métodos

Se realizó un seguimiento de ochocientos cuatro pacientes sometidos a PR consecutivos durante una mediana de 9,5 años. Las características de los pacientes y los desenlaces del tratamiento de toda la cohorte se han notificado anteriormente (Swanson *et al.*, UROL. ONCOL. (2007) 25:110-114). Se disponía de portaobjetos y/o bloques de tejido de la evaluación patológica final con suficiente tejido para el análisis de 430 pacientes. La cohorte se dividió aleatoriamente en 212 pacientes utilizados para un conjunto entrenamiento y 199 muestras de pacientes como conjunto de validación.

Expresión génica (métodos estadísticos): la asociación entre la recidiva bioquímica y la expresión de CCG se evaluó usando modelos de RP de Cox para el tiempo hasta la recidiva. Todos los valores de p notificados en este estudio se derivaron de una prueba de razón de probabilidades que comparaba el modelo nulo con el modelo que contenía

la variable de prueba. Se seleccionó aleatoriamente un conjunto de 31 CCG (tabla 3, véase anteriormente). Los ensayos se usaron para generar datos de expresión de 212 pacientes en el conjunto de entrenamiento. Todos los datos de expresión se generaron por triplicado. Los datos de expresión se combinaron en una firma calculando el nivel de expresión medio para 26 CCG. La asociación entre la recidiva bioquímica y la expresión de CCG se evaluó usando modelos de RP de Cox para el tiempo hasta la recidiva.

Preparación de las muestras y diseño del estudio: se aisló ARN de cortes FFPE de tumores derivados de 411 pacientes con cáncer de próstata tratados con PR. Se usaron cortes tumorales representativos de 10 µm de grosor para aislar el ARN. Cuando fue necesario, se usó una macro o microdissección de la muestra guiada por un patólogo para enriquecer el tejido tumoral antes del aislamiento del ARN. Ninguna de las muestras de la cohorte de validación se sometió a microdissección. Antes de cualquier análisis, la cohorte se dividió en 212 pacientes para la caracterización inicial de la firma ("conjunto de entrenamiento") y 199 pacientes para la validación. Las características clínicas de la cohorte de entrenamiento y validación se enumeran en la tabla 5.

Tabla 5

	Entrenamiento	Validación	Valor de p	Estadística
Edad en años en la PR, media (d.e.)	67,3 (5,9)	66,8 (5,8)	0,355	Prueba de la t
Etnia (% de raza no blanca)	2,80 %	7,60 %	0,042	Prueba exacta de Fisher
Método de disección (% de lcm)	24 %	0 %	NA	NA
Recidiva (%)	71/212 (33,5 %)	72/199 (36,2 %)	0,605	Prueba exacta de Fisher
Días hasta la recidiva, mediana	910	822	0,463	Prueba de la t
Días hasta el seguimiento, mediana	3373	3387	0,173	Prueba de la t
PSA previo a la cirugía (mediana)	7,3	6,8	0,163	Prueba de la t de log
Vesícula seminal	23/212 (10,8 %)	28/199 (14,1 %)	0,37	Prueba exacta de Fisher
Vejiga	12/212 (5,7 %)	17/199 (8,5 %)	0,335	Prueba exacta de Fisher
Ganglio linfático	8/212 (3,8 %)	10/199 (5,0 %)	0,632	Prueba exacta de Fisher
Capsular	100/212 (47,2 %)	101/199 (50,8 %)	0,49	Prueba exacta de Fisher
A través de la cápsula	59/212 (27,8 %)	66/199 (33,2 %)	0,283	Prueba exacta de Fisher
Márgenes positivos	43/212 (20,3 %)	57/199 (28,6 %)	0,051	Prueba exacta de Fisher
Puntuación de Gleason posterior a PR > 6	80/212 (37,7 %)	66/199 (33,2 %)	0,354	Prueba exacta de Fisher
Nomograma posterior a PR, media (d.e.)	137 (19,5)	138 (23,0)	0,424	Prueba de la t

Resultados

La firma de expresión de CCG (tabla 3, véase anteriormente) fue predictiva de la recidiva de la enfermedad en un análisis univariante (valor de $p = 1,8 \times 10^{-9}$, tabla 6). La distribución de la puntuación de firma estaba sesgada hacia valores más altos (menor expresión). El valor de la mediana de la puntuación de firma se usó para dividir la cohorte de entrenamiento en dos grupos que contenían muestras con expresión de CCG alta o baja. En la figura 2 se muestra la supervivencia frente al tiempo para ambos grupos.

También se evaluó el poder predictivo de la firma de CCG después de tener en cuenta las variables clínicas que se incluyen normalmente en un nomograma posquirúrgico (el nomograma de Kattan-Stephenson). El nomograma fue

un factor predisponente altamente significativo de la recidiva (valor de p de $1,6 \times 10^{-10}$). Después de ajustar para el nomograma, la firma de CCG fue un factor predisponente significativo de la recidiva bioquímica (figura 3) en la cohorte de descubrimiento (valor de p de 0,03) y en la cohorte de validación clínica (valor de p de $4,8 \times 10^{-5}$).

5

Tabla 6

	N	Covariables	Valor de p medio de CCG*	Razón de riesgos de recidiva
ENTRENAMIENTO (31 CCG)	212	ninguna	0,00404	1,24
	204	nomograma posterior a la cirugía	0,03320	1,16
VALIDACIÓN (subconjunto de 26 CCG)	199	ninguna	$1,8 \times 10^{-9}$	2,68
	197	nomograma posterior a la cirugía	$4,8 \times 10^{-5}$	1,94

* Media de la expresión de genes del ciclo celular con imputación de valores faltantes, prueba de razón de probabilidades para el modelo de riesgos proporcionales de Cox.

- 10 Para ayudar a entender la interacción entre el nomograma y la firma de expresión de CCG, se generó un diagrama de dispersión que comparaba estos factores predisponentes (figura 4) (las estrellas de color gris claro representan pacientes cuyo cáncer recidivó, mientras que las estrellas de color negro representan pacientes cuyo cáncer no recidivó). El análisis del diagrama de dispersión mediante medias de KM dividió las muestras en tres agrupaciones basándose únicamente en la puntuación de nomograma. Posteriormente, se descubrió que las agrupaciones se basaban en parámetros clínicos bien entendidos. Los pacientes en la agrupación de puntuación más baja (116/117) presentaban una enfermedad confinada en órgano. Los pacientes en la agrupación de puntuación media (48/60) tenían al menos un parámetro posquirúrgico que se sabe que está asociado con un mal desenlace (es decir, enfermedad a través de la cápsula, ganglios linfáticos positivos para la enfermedad y/o vesículas seminales positivas para la enfermedad) y un PSA prequirúrgico bajo (<10 ng/ml). Los pacientes en la agrupación de puntuación más alta tenían al menos un parámetro posquirúrgico desfavorable y un PSA prequirúrgico alto. A continuación, los pacientes en las agrupaciones de puntuaciones baja y media se dividieron por la media de la puntuación de CCG. Los desenlaces para los pacientes en la agrupación de puntuación más alta se predicen adecuadamente mediante el nomograma y, por tanto, no se dividieron adicionalmente. Como resultado, el diagrama de dispersión define cinco grupos de pacientes con tasas de recidiva de la enfermedad del 2 %, del 40 % (para dos grupos), del 65 % y del 80 % (tabla 7). En la figura 5 se muestra la tasa de recidiva de los cinco grupos frente al tiempo.

Tabla 7

	Nomograma posterior a PR		
Puntuación de CCG	<i>Bajo</i>	<i>Medio</i>	<i>Alto</i>
<i>Baja</i>	1/62 (1,6 %)	13/31 (41,9 %)	16/20 (80 %)
<i>Alta</i>	21/55 (38,2 %)	19/29 (65,5 %)	

- 30 El diagrama de dispersión mostrado en la figura 4 sugiere que existe una interacción no lineal entre la firma de CCG y el nomograma posquirúrgico. Es decir, la firma de CCG es un mejor factor predisponente en pacientes con puntuaciones de nomograma bajas. Por tanto, el estudio buscó evidencias estadísticas de una interacción entre estas variables en un modelo multivariante para predecir la recidiva de la enfermedad (tabla 8). Hubo evidencias significativas de una interacción favorable tanto en los estudios de entrenamiento como en los de validación. La inclusión del término de interacción en el modelo mejoró drásticamente la significación pronóstica de la firma de CCG después de ajustar para el nomograma (valores de p de 0,0015 en la cohorte de entrenamiento y de $1,2 \times 10^{-8}$ en la cohorte de validación).

40

Tabla 8

Resumen estadístico						
Cohorte	N	Variables independientes	Covariables	Valor de p de interacción	Valor de p de variable	Razón de riesgos de recidiva
Entrenamiento	204	nomograma	ninguna	NA	$1,6 \times 10^{-10}$	
	212	firma de CCG	ninguna	NA	0,004	1,24
	204	firma de CCG	nomograma	0,021	0,0015	

Validación	197	nomograma	ninguna	NA	$7,7 \times 10^{-13}$	
	199	firma de CCG	ninguna	NA	$1,8 \times 10^{-9}$	2,68
	197	firma de CCG	nomograma	0,0001	$1,28 \times 10^{-8}$	

EJEMPLO 3

El siguiente estudio tenía como objetivo determinar el número óptimo de CCG que debían incluirse en la firma. Tal como se mencionó anteriormente, los niveles de expresión de CCG están correlacionados entre sí, por lo que era posible que la medición de un pequeño número de genes fuera suficiente para predecir el desenlace de la enfermedad. De hecho, los CCG individuales del conjunto de 31 genes de la tabla 3 (panel C) se añaden significativamente al nomograma de Kattan-Stephenson, tal como se muestra en la tabla 9 a continuación (después de ajustar para el nomograma y un término de interacción entre el nomograma y la expresión de CCG):

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>FOX M1</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 RP de Cox completo (puntuación de nomograma posterior a PR + expresión de ciclo celular + nomograma:ciclo celular) frente a reducido (puntuación de nomograma posterior a PR sola) del tiempo hasta la recidiva.

Para evaluar el rendimiento de subconjuntos más pequeños del conjunto de CCG más grande (es decir, paneles de CCG más pequeños), el estudio también comparó lo bien que la firma predice el desenlace en función del número de CCG incluidos en la firma (figura 1). El tiempo hasta la recidiva química después de la cirugía de próstata se regresionó sobre la media de CCG ajustada por la puntuación de nomograma posterior a PR. Los datos consisten en ensayos de TLDA expresados como deltaCT para 199 muestras FFPE de tumores de próstata y 26 CCG y se analizaron mediante un modelo de RP de Cox multivariante. Los valores de p corresponden a la prueba de razón de probabilidades del modelo completo (nomograma + media del ciclo celular, incluida la interacción) frente al modelo reducido (sólo nomograma). Tal como se muestra en la tabla 10 a continuación y en la figura 1, las firmas de CCG pequeñas (por ejemplo, 2, 3, 4, 5, 6 CCG, etc.) se añaden significativamente al nomograma de Kattan-Stephenson:

Tabla 10

# de CCG	Media de 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 aleatoriamente, de tamaño 1 a 6, de genes del ciclo celular.

EJEMPLO 4

El objetivo de este experimento fue evaluar la asociación entre las mutaciones de *PTEN* y la recidiva bioquímica en pacientes con cáncer de próstata después de la prostatectomía radical. Se encontró que las mutaciones somáticas de *PTEN* estaban significativamente asociadas con la recidiva y, lo que es más importante, añadían información pronóstica más allá del nomograma clínico establecido para la recidiva del cáncer de próstata (el nomograma de Kattan-Stephenson) y la puntuación de firma de CCG (descrita en los ejemplos 1 y 2, véanse anteriormente).

Pacientes y métodos

Se realizó un seguimiento de ochocientos cuatro pacientes sometidos a PR consecutivos durante una mediana de 9,5 años. Las características de los pacientes y los desenlaces del tratamiento de toda la cohorte se han notificado anteriormente (Swanson *et al.*, UROL. ONCOL. (2007) 25:110-114). Se disponía de portaobjetos y/o bloques de tejido de la evaluación patológica final con suficiente tejido para el análisis de 430 pacientes. De ellos, se seleccionaron 191 para el cribado de mutaciones de *PTEN* en función de la cantidad de tumor disponible.

Se aisló ADN genómico de las muestras FFPE de tumores para el cribado de mutaciones de *PTEN* usando el kit de tejido FFPE para ADN QIAamp (Qiagen, Valencia, CA) según el protocolo del kit. Los portaobjetos FFPE se tiñeron en primer lugar con hematoxilina y eosina y se examinaron por un patólogo para identificar la región tumoral. Después de la desparafinización, se recortó el tejido tumoral de los portaobjetos con una cuchilla. En algunas muestras, la disección se facilitó mediante microscopía de captura láser (LCM), debido a la dispersión de las células tumorales.

Las mutaciones se detectaron diseñando cebadores de secuenciación para interrogar la secuencia genómica de *PTEN*. Los cebadores contenían colas M13 directas e inversas para facilitar la secuenciación. Después de la amplificación, la secuencia de ADN se determinó en un equipo Mega BASE 4500 (GE Healthcare) usando la química de colorante-cebador tal como se describe en Frank *et al.*, J. CLIN. ONCOL. (2002) 20:1480-1490. Debido a las dificultades técnicas asociadas con la secuenciación del ADN derivado de material FFPE, cada mutación se detectó mediante al menos dos reacciones de amplificación y secuenciación independientes.

Métodos estadísticos: a menos que se especifique lo contrario, la asociación entre la recidiva bioquímica y las mutaciones de *PTEN* se evaluó usando modelos de RP de Cox para el tiempo hasta la recidiva. Los valores de *p* resultantes se derivaron de una prueba de razón de probabilidades que comparaba el modelo nulo con el modelo que contenía la variable de prueba. En este ejemplo (ejemplo 4), la firma de CCG se derivó de 26 CCG (panel D de la tabla 2, véase anteriormente). Todos los datos de expresión se generaron por triplicado. Los datos de expresión se combinaron en una firma calculando el nivel de expresión medio para 26 CCG. Los datos clínicos fueron las variables incluidas en el nomograma de Kattan-Stephenson.

Resultados

Se encontraron mutaciones de *PTEN* en 13 individuos (13/191). En este subconjunto de 191 pacientes, *PTEN* fue un factor predisponente significativo de recidiva bioquímica (valor de *p* = 0,031). La tasa de recidiva en portadores de mutaciones fue del 69 % (9/13) en comparación con el 36 % (64/178) en pacientes sin mutaciones. La diferencia en la tasa de recidiva también es significativa usando una prueba exacta de Fisher (valor de *p* = 0,034). En el subconjunto de pacientes con datos de parámetros clínicos, puntuación de firma de CCG y mutaciones de *PTEN*, el estado de *PTEN* fue un factor predisponente significativo de recidiva bioquímica después de ajustar tanto para los parámetros clínicos como para la firma de CCG (valor de *p* de 0,024). Finalmente, la combinación de la mutación de *PTEN* con la firma de CCG fue un mejor factor predisponente del desenlace después de ajustar para los parámetros clínicos que el uso de la firma de CCG después de ajustar para los parámetros clínicos (valor de *p* = 0,0002 para la combinación en comparación con 0,0028 para CCG solamente). Estos resultados muestran que las mutaciones de *PTEN* proporcionan información sobre la probabilidad de recidiva que no está correlacionada con los parámetros clínicos ni con la firma de CCG, y que el uso de los tres parámetros para evaluar el riesgo de recidiva proporciona una estimación más precisa de la probabilidad de recidiva de lo que era posible anteriormente.

EJEMPLO 5

Este ejemplo describe estudios adicionales para validar y refinar las firmas de CCG.

Pacientes y métodos

Se realizó un seguimiento de ochocientos cuatro pacientes sometidos a prostatectomía radical consecutivos durante una mediana 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 de PSA preoperatorio fue de 6,6 ng/ml, con un 72 % < 10 ng/ml y un 28 % > 10 ng/ml. Se entintaron las muestras y se registraron los parámetros clínicos en cuanto a margen del cuello vesical o de la uretra positivo, invasión de la cápsula, extensión a través de la cápsula, márgenes positivos y afectación de las vesículas seminales. La recidiva bioquímica se definió como un PSA >0,3 ng/ml. Para este estudio se tenía acceso a datos clínicos de 690 pacientes. Se disponía de portaobjetos y/o bloques de tejido de la evaluación patológica final con suficiente tejido para el análisis de 442 pacientes. La cohorte se dividió en 195 pacientes para una cohorte de entrenamiento y

247 pacientes para la validación.

Selección de genes

- 5 Se realizaron ensayos de 126 CCG y 47 genes HK (de mantenimiento) contra 96 muestras FFPE de tumores de próstata anónimas obtenidas comercialmente, sin desenlaces ni otros datos clínicos. La hipótesis de trabajo era que los ensayos medirían con distintos grados de precisión el mismo fenómeno subyacente (proliferación del ciclo celular dentro del tumor para los CCG y concentración de la muestra para los genes HK). Los ensayos se clasificaron según 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 de la correlación de cada uno de los 126 CCG con la media se presentan en la tabla 23. Sin incluir los CCG con una expresión promedio baja, o los ensayos que produjeron fallos en las muestras, aproximadamente la mitad de los CCG tuvieron correlaciones inferiores a 0,58, y una cuarta parte de los genes HK tuvieron correlaciones inferiores a 0,95. Se interpretó que estos ensayos no reflejaban el fenómeno subyacente y se eliminaron, dejando un subconjunto de 56 CCG (panel G) y 36 genes HK candidatos (tablas 11 y 12). Se volvieron a calcular los coeficientes de correlación en este subconjunto, y se realizó la selección final a partir de la lista clasificada.

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

Símbolo de gen	Correlación con la media de CCG	Símbolo de gen	Correlación con la media de CCG	Símbolo de gen	Correlación con la media de CCG
<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>KIF11</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>KIF23</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>GIN51</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>ASF1B</i>	0,827	<i>MCM4</i>	0,703		

20

Tabla 12: Lista de 15 genes de mantenimiento (HK)

Símbolo de 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

Se extrajo ARN total de cortes FFPE de tumores representativos de 5 µm de grosor. Las muestras se desparafinizaron usando un baño de xileno y posteriormente se hidrataron en series graduales de baños de etanol. Después de eso, la región tumoral se diseccionó del portaobjetos con una cuchilla siguiendo las instrucciones del patólogo. Alternativamente, la región tumoral se diseccionó directamente en un tubo Eppendorf y se eliminó la parafina 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 usando RNeasy FFPE o miRNeasy (Qiagen) tal como describe el fabricante (con la única excepción de la digestión prolongada con proteinasa K descrita anteriormente). El ARN total aislado se trató con ADNasa I (Sigma) antes de la síntesis de ADNc. Posteriormente, se empleó el kit de archivo de ADNc de alta capacidad (Applied Biosystems) para convertir el ARN total en ADNc monocatenario, tal como describe el fabricante. Se necesitó un mínimo de 200 ng de ARN para la reacción de RT.

Antes de medir los niveles de expresión, el ADNc se amplificó previamente con una reacción conjunta que contenía ensayos TaqMan. Las condiciones de la reacción de amplificación previa fueron: 14 ciclos de 95 °C durante 15 s y 60 °C durante 4 minutos. El primer ciclo se modificó para incluir una incubación de 10 minutos a 95 °C. La reacción de amplificación se diluyó 1:20 usando el tampón TE 1X antes de su carga en las matrices de baja densidad TaqMan™ (TLDA, Applied Biosystems) para medir la expresión génica.

Puntuación de CCG

La puntuación de CCG se calcula a partir de la expresión de ARN de 31 CCG (panel F) normalizada por 15 genes de mantenimiento (HK). Los números relativos de CCG (31) y genes HK (15) se optimizaron con el fin de minimizar la varianza de la puntuación de CCG. La puntuación de CCG es la media no ponderada de los valores CT para la expresión de CCG, normalizada por la media no ponderada de los genes HK, de modo que los valores más altos indican una mayor expresión. Una unidad equivale a un cambio de dos veces en la expresión. Los valores faltantes se imputaron usando la expresión media para cada gen determinada en el conjunto de entrenamiento usando únicamente muestras de buena calidad. Las puntuaciones de CCG se centraron en el valor medio, también determinado en el conjunto de entrenamiento.

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

El umbral de puntuación de CCG para determinar un riesgo bajo se basó en la puntuación de CCG más baja de recidivas en el conjunto de entrenamiento. Luego, el umbral se ajustó a la baja en 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 parámetros clínicos disponibles y estimar el riesgo clínico previo de recidiva 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 univariantes de los parámetros clínicos que se sabe que están asociados con el desenlace (véase la tabla 13 a continuación). Los parámetros no significativos se excluyeron del modelo. Se creó una variable compuesta para la enfermedad confinada en órgano, con invasión definida como márgenes quirúrgicos, extensión extracapsular o afectación de cualquiera de vesículas seminales, márgenes del cuello vesical/uretra o ganglios linfáticos. La variable compuesta para la enfermedad confinada en órgano resultó más significativa en el modelo que cualquiera de sus cinco componentes, algunos de los cuales

estaban correlacionados unos con otros o no eran prevalentes. El ajuste del modelo se realizó utilizando los criterios AIC para las covariables posoperatorias.

Tabla 13: Análisis univariante de los parámetros clínicos y asociación con la recidiva bioquímica

Variable clínica	Valor de p*	# de apariciones	Total	Frecuencia
VEJIGA	0,0002	36	443	0,081
CAPSULAR	$1,1 \times 10^{-9}$	194	443	0,438
ETNIA (RAZA BLANCA)	0,6741	416	439	0,948
GANGL.LINF.	0,0009	33	443	0,074
MARG.POS	$6,1 \times 10^{-11}$	83	443	0,187
PATHGLEA	$6,7 \times 10^{-16}$	NA	443	NA
PATHGRAD	$2,4 \times 10^{-11}$	NA	443	NA
PATHSTAG	$3,1 \times 10^{-15}$	NA	443	NA
PRE.PSA.LOG10	$6,2 \times 10^{-12}$	NA	443	NA
VES.SEM.	$3,0 \times 10^{-8}$	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 RP de Cox para la prueba de razón de probabilidades				

El modelo final (es decir, el nomograma) tiene variables binarias para la enfermedad confinada en órgano y la puntuación de Gleason menor que 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 posterior a PR (es decir, el nomograma de Kattan-Stephenson), excepto el año de PR y los dos componentes de la puntuación de Gleason. La distribución del riesgo clínico previo muestra tres ganglios distintos (figura 8). Se usó la agrupación de *k*-medias con 3 centros para establecer el umbral de la agrupación de bajo riesgo, que comprende aproximadamente el 50 % de la muestra.

Tabla 14: Modelo clínico

Parámetro clínico	Coefficiente	HR	Valor de p*
enfermedad confinada en órgano	-0,827	0,44	$3,4 \times 10^{-6}$
puntuación de Gleason ≤ 6	-0,8734	0,42	$4,2 \times 10^{-7}$
PSA logarítmico	0,6678	1,95	$2,0 \times 10^{-4}$
* Valor de p de RP de Cox para la prueba de razón de probabilidades			

Análisis estadístico

Los parámetros clínicos se compararon entre los conjuntos de entrenamiento y de validación usando la prueba de la *t* de Student para los parámetros continuos y la prueba exacta de Fisher para los parámetros categóricos. El riesgo clínico previo de recidiva bioquímica de los pacientes después de la cirugía se estimó mediante una puntuación de nomograma posterior a PR que resumía 7 covariables. Se usó la agrupación de *k*-medias de la puntuación de nomograma para clasificar a los pacientes como de bajo o alto riesgo clínico previo. Los datos de expresión se expresaron como CT (el ciclo de PCR en el que la intensidad de fluorescencia supera un umbral predeterminado) de cada CCG normalizado por la media de los 15 genes de mantenimiento (tabla 12 anterior).

Las muestras de mala calidad se excluyeron del análisis para eliminar las muestras de mala calidad o las lecturas dudosas sin comprometer la integridad de la firma al excluir inadvertidamente muestras con baja expresión de CCG. Por consiguiente, los umbrales de limpieza o filtrado de los datos se establecieron de manera conservadora. Los niveles medios de expresión de los genes HK de cada muestra, que eran superiores a los de los CCG, se usaron para identificar las muestras de mala calidad. Las métricas técnicas de la eficiencia de amplificación y las desviaciones estándar excesivamente altas de las réplicas se usaron para identificar mediciones de CT poco fiables. No se permitieron fallos de genes HK, ni más de 1 fallo de 3 réplicas para los CCG.

La asociación entre la recidiva bioquímica y la expresión de CCG después de ajustar para el riesgo clínico predicho por los parámetros clínicos se evaluó usando un modelo de riesgos proporcionales de Cox para el tiempo hasta la

recidiva. La suposición de ausencia de dependencia temporal de los riesgos proporcionales se sometió a prueba para el modelo completo de la firma de CCG más la puntuación binaria de parámetros clínicos con un término de interacción, y para la firma de CCG sólo en los subconjuntos de riesgo clínico. No fue significativo ni en el entrenamiento ni en la validación, lo que indica que no hay evidencias de dependencia temporal. Todos los valores de p notificados proceden de una prueba de razón de probabilidades que compara el modelo reducido o nulo con el modelo que contiene la variable de prueba. Los diagramas de Kaplan-Meier se usan para mostrar las probabilidades de supervivencia estimadas para subconjuntos de pacientes; sin embargo, los valores de p proceden de la prueba de razón de probabilidades de Cox para los valores continuos de la variable. Todos los análisis estadísticos se realizaron en S+ versión 8.1.1 para Linux (TIBCO Spotfire) o R 2.9.0 (<http://www.r-project.org>).

Resultados

Se aisló ARN de cortes FFPE de tumores derivados de 442 pacientes con cáncer de próstata tratados con PR. La cohorte se dividió en 195 pacientes para la caracterización inicial de la firma ("conjunto de entrenamiento") y 247 pacientes para la validación. Los parámetros clínicos de la cohorte de entrenamiento y de validación se enumeran en la tabla 15. No hubo diferencias significativas después de ajustar para comparaciones múltiples.

Tabla 15: Parámetros clínicos de las cohortes de pacientes de entrenamiento y de validación

Parámetro clínico	Entrenamiento	Validación	Valor de p	Análisis estadístico
Edad en años en la PR, media (d.e.)	67,5 (6,2)	66,8 (5,6)	0,204	Prueba de la t
Etnia (% de raza no blanca)	3,10 % (2 de raza negra, 3 de raza hispánica, 1 de otra raza)	7,30 % (10 de raza negra, 7 de raza hispánica, 1 de otra raza)	0,058	Prueba exacta de Fisher
Recidiva	73/195 (37,4 %)	90/247 (36,4 %)	0,843	Prueba exacta de Fisher
Días hasta la recidiva, mediana	839	736	0,308	Prueba de la t
Días hasta el seguimiento, mediana	3300	3332	0,556	Prueba de la t
PSA previo a la cirugía de PR, mediana	7,4	6,4	0,022	Prueba de la t de log
Vesículas seminales	23/195 (11,8 %)	33/247 (13,4 %)	0,668	Prueba exacta de Fisher
Margen del cuello vesical/uretra	12/195 (6,2 %)	16/247 (6,5 %)	1	Prueba exacta de Fisher
Ganglios linfáticos	8/195 (4,1 %)	12/247 (4,9 %)	0,819	Prueba exacta de Fisher
Penetración capsular	104/195 (53,3 %)	115/247 (46,6 %)	0,18	Prueba exacta de Fisher
A través de la cápsula	66/195 (33,8 %)	73/247 (29,6 %)	0,354	Prueba exacta de Fisher
Márgenes positivos	51/195 (26,2 %)	61/247 (24,7 %)	0,742	Prueba exacta de Fisher
Puntuación de Gleason posterior a PR < 7	114/195 (58,5 %)	166/247 (67,2 %)	0,06	Prueba exacta de Fisher
Enfermedad confinada en órgano	108/195 (55,4 %)	156/247 (63,2 %)	0,118	Prueba exacta de Fisher
PFP a los 10 años (IC del 95 %)	61 % (52 %, 69 %)	67 % (60 %, 73 %)	0,905	Prueba de rangos logarítmicos

Para analizar la firma de CCG para este estudio, se sometieron a prueba 126 CCG en ARN derivado de 96 tumores de próstata (tabla 11). Las muestras tumorales eran anónimas y no estaban asociadas con datos clínicos. De este conjunto de genes, se seleccionaron 31 genes (panel F) para su inclusión en la presente firma (tabla 16). Los genes se seleccionaron en función de su rendimiento técnico y de lo bien que se correlacionaba cada gen con el nivel de

expresión medio de todo el conjunto de CCG en las 96 muestras anónimas.

Tabla 16: Firma de CCG del conjunto de entrenamiento (panel F)

Símbolo	ID de gen	Símbolo	ID de gen	Símbolo	ID de gen
<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 CCG, se generaron datos de expresión en 195 pacientes en el conjunto de entrenamiento. Puesto que los niveles de expresión génica individuales estaban correlacionados, se combinaron en una puntuación de firma calculando la expresión media para el conjunto completo de 31 genes (panel F), normalizada por 15 genes de mantenimiento (tabla 12). La distribución de la puntuación de CCG 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 observando una baja expresión de los genes de mantenimiento o un número inaceptable de fallos de CCG, y se excluyeron del análisis. Después de aplicar las reglas de exclusión, quedaban 140 muestras disponibles para el análisis. La asociación entre la recidiva bioquímica y la expresión de CCG se evaluó usando modelos de RP de Cox para el tiempo hasta la recidiva. Un valor alto de expresión de CCG fue predictivo de recidiva de la enfermedad en un análisis univariante (valor de $p = 0,01$, tabla 17).

A continuación, se evaluó la utilidad pronóstica de la firma de CCG después de tener en cuenta los parámetros clínicos que se sabe que están asociados con la recidiva después de la PR. Para tener en cuenta las medidas clínicas en el presente análisis, se creó un modelo/nomograma que incluía el PSA preoperatorio, la puntuación de Gleason y evidencias de enfermedad fuera de la próstata (es decir, cualquiera de extensión extracapsular o patología posquirúrgica positiva en ganglios linfáticos, márgenes, cuello vesical, margen de la uretra o vesículas seminales). El modelo se optimizó en 443 pacientes (tablas 13 y 14), incluyendo todos los pacientes de los que se tenían datos clínicos pero que no estaban en el conjunto de validación, y fue un factor predisponente altamente significativo de recidiva en la cohorte de entrenamiento (valor de $p = 2,5 \times 10^{-11}$). La distribución de las puntuaciones del modelo clínico contenía varios modos (figura 8), que separaban los grupos de pacientes de alto y bajo riesgo. Por tanto, la puntuación se usó posteriormente como variable binaria (alto o bajo riesgo). La agrupación de bajo riesgo se correlacionó con un conjunto coherente de parámetros clínicos. Específicamente, la gran mayoría (215/218) tenía enfermedad confinada en órgano y puntuación de Gleason < 7 . Además, el 80 % tenía un PSA prequirúrgico bajo (< 10 ng/ml). Los pacientes de la agrupación de alto riesgo ($N=225$) eran más heterogéneos, pero tendían a presentar características clínicas que se sabe que están asociadas con un mal desenlace (por ejemplo, Gleason > 6 y/o enfermedad a través de la cápsula).

El análisis multivariante del conjunto de entrenamiento que incorporaba el presente modelo clínico binario mostró evidencias de una interacción no lineal entre la firma de expresión y los parámetros clínicos (tabla 17). Para ayudar a entender la naturaleza de esta interacción, se generó un diagrama de dispersión comparando estos factores predisponentes (figura 8). Tal como resulta evidente a partir de la figura, la puntuación de CCG resultó útil para evaluar el riesgo de recidiva en pacientes definidos como de bajo riesgo por los parámetros clínicos. De hecho, incluso después de ajustar para el modelo clínico dentro de los pacientes de bajo riesgo, la firma de CCG fue un fuerte factor predisponente de recidiva bioquímica (valor de $p = 0,0071$).

Tabla 17: Resumen estadístico

		Entrenamiento de 31 genes N=195	Validación de 31 genes N=247

	Subconjunto basado en el modelo clínico	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 CCG		0,01		140	$5,8 \times 10^{-8}$		218
Riesgo clínico binario (bajo frente a alto)		$5,1 \times 10^{-6}$		133	$1,1 \times 10^{-10}$		215
Puntuación de CCG ajustada para riesgo clínico binario + interacción		0,018	0,032	133	$8,3 \times 10^{-7}$	0,026	215
Puntuación de CCG sólo	bajo riesgo	0,0038		54	$7,5 \times 10^{-5}$		112
Puntuación de riesgo clínico	bajo riesgo	0,22		54	0,044		112
Puntuación de CCG ajustada para riesgo clínico (riesgo clínico frente a riesgo clínico + CCG)	bajo riesgo	0,0071		54	0,00019		112
Puntuación de CCG	alto riesgo	0,48		79	$5,8 \times 10^{-4}$		103
Puntuación de riesgo clínico	alto riesgo	$2,8 \times 10^{-6}$		79	0,0076		103
Puntuación de CCG ajustada para riesgo clínico (riesgo clínico frente a riesgo clínico + CCG)	alto riesgo	0,51		79	0,0026		103

Se usaron los presentes datos de entrenamiento en el diagrama de dispersión para establecer una puntuación de umbral optimizada de -0,16 para la firma de CCG (la puntuación de CCG media 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 se sitúan por debajo de este umbral, y se seleccionó de forma que no hubiera recidivas a los 10 años después de la PR (es decir, un valor predictivo negativo [VPN] del 100 %). Como resultado del establecimiento de valores de umbral tanto para el modelo clínico como para la puntuación de CCG, el diagrama de dispersión se dividió en cuatro secciones con tasas de recidiva del 0 % (CCG bajo) y del 26 % (CCG alto) para los pacientes de bajo riesgo; y del 60 % (CCG bajo) y del 50 % para los pacientes de alto riesgo.

A continuación, se generaron datos de expresión de CCG en 247 pacientes en la presente cohorte de validación. Se eliminaron treinta y dos muestras de los análisis adicionales según las reglas de exclusión desarrolladas en la cohorte de entrenamiento. El panel F fue un factor predisponente significativo de recidiva bioquímica en un análisis univariante (valor de $p = 5,8 \times 10^{-8}$, tabla 17). Después de ajustar para el modelo clínico binario, la firma de CCG fue altamente predictiva de recidiva en la cohorte de validación (valor de p de $8,3 \times 10^{-7}$), y como en el conjunto de entrenamiento, hubo evidencias significativas de una interacción no lineal entre variables. La firma de CCG fue informativa en todo el espectro de riesgo definido clínicamente (tabla 17). En cuanto a la validación de los resultados de entrenamiento, el valor de p para la asociación entre recidiva y firma de CCG en pacientes de bajo riesgo fue de $1,9 \times 10^{-4}$.

Se aplicó el umbral de CCG derivado del presente análisis de la cohorte de entrenamiento al presente conjunto de datos de validación (figura 9). Los pacientes de bajo riesgo con puntuaciones de CCG por debajo del umbral tenían una tasa de recidiva predicha a los 10 años del 5 % (equivalente a un VPN validado de 0,95). En general, la combinación de la puntuación de CCG y los parámetros clínicos dividió la cohorte en cuatro grupos con tasas de recidiva predichas a los 10 años del 5 %, del 22 %, del 36 % y del 70 % (tabla 18). En las figuras 10 y 11 se muestra la tasa de recidiva predicha frente a la puntuación de CCG para los pacientes de la cohorte de validación.

Tabla 18: Resumen de tasas de recidiva en la cohorte de validación definida por riesgo clínico y puntuación de CCG

Riesgo clínico	Puntuación de CCG	Tasa de recidiva a los 10 años	
		Estimación de Kaplan-Meier	n
bajo	baja	0,05	39
bajo	alta	0,22	73
alto	baja	0,36	27
alto	alta	0,7	76

Se sometió a prueba el presente umbral validado frente a diversas definiciones de pacientes de bajo riesgo (tabla 19). La puntuación de firma fue un indicador pronóstico significativo en una variedad de definiciones clínicas de bajo riesgo y, dependiendo de la definición, generó una tasa de recidiva predicha a los 10 años de 0,05 a 0,10.

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

Definición clínica de riesgo bajo	CCP baja*		Valor de p***
	Recidiva predicha a los 10 años**	n	
Enfermedad confinada en órgano y puntuación de Gleason < 7 y PSA < 10	0,05	39	$9,4 \times 10^{-4}$
Enfermedad confinada en órgano y puntuación de Gleason < 7	0,08	40	$5,8 \times 10^{-3}$
Enfermedad confinada en órgano y puntuación de Gleason < 8 y PSA < 10	0,07	42	$8,7 \times 10^{-4}$
Enfermedad confinada en órgano y puntuación de Gleason < 8	0,1	43	$4,1 \times 10^{-3}$
Enfermedad confinada en órgano sólo	0,1	44	$2,4 \times 10^{-3}$
* - definido por umbral validado			
** - estimaciones de Kaplan-Meier			
*** - para la diferencia entre las estimaciones de KM para riesgo bajo y alto ajustadas por la varianza de Greenwood.			

5 Comentario

Los presentes 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 de ARNm de los genes del ciclo celular (CCG). Por definición, la expresión de los CCG se regula en función de la fase del ciclo celular. Es decir, se activan en fases específicas del ciclo celular, de modo que las células en crecimiento activo tienen mayores niveles de expresión de CCG que las células quiescentes. Supuestamente, este hecho subyace a la capacidad de la firma para predecir la progresión del cáncer. Sin querer limitarse a la teoría, se cree que al medir los niveles de expresión de CCG están midiéndose indirectamente la tasa de crecimiento y la agresividad inherente del tumor, lo que en última instancia repercute en la probabilidad de recidiva del cáncer de próstata después de la prostatectomía.

Existe una distinción importante entre este estudio y muchos otros que han intentado generar firmas moleculares pronósticas. A menudo, los estudios similares comienzan con un gran número de biomarcadores candidatos (que a veces superan los 1.000 genes) que luego se evalúan en busca de su asociación con un fenotipo clínico de interés. En ocasiones, este enfoque puede adolecer de pruebas múltiples inherentes, lo que puede hacer que la significación de la firma derivada sea incierta. En este caso se ha sometido a prueba una única hipótesis: el CCG sería pronóstico en el cáncer de próstata (de hecho, se seleccionaron los genes en función de su correlación con la expresión de CCG, no en función de su asociación con la recidiva). Y puesto que la expresión de CCG está correlacionada, se combinaron los datos de expresión en una firma predictiva determinando el valor medio de expresión de todos los genes de la firma. La simplicidad de este enfoque, desde el punto de vista biológico y computacional, respalda la opinión de que es probable que la afirmación principal de este estudio sea muy sólida y se reproduzca en estudios posteriores.

La firma de CCG (panel F) es independientemente predictiva y se añade significativamente al poder predictivo de los parámetros clínicos normalmente empleados para predecir la recidiva de la enfermedad después de la cirugía. Esto es cierto tanto en la cohorte de entrenamiento como en la de validación.

La firma es inmediatamente útil para definir el riesgo de los pacientes que presentan parámetros clínicos de bajo riesgo. En este caso, se definió esencialmente el bajo riesgo como Gleason <7, PSA<10 y enfermedad confinada en órgano. La puntuación de firma de CCG subdivide eficazmente el grupo de bajo riesgo en pacientes con tasas de recidiva muy bajas (5 %) y un mayor riesgo de recidiva (22 %) (figura 9 y tabla 18). Este es el efecto más drástico de la firma molecular, redefinir con precisión el riesgo de los pacientes previamente definidos como de bajo riesgo basándose en parámetros clínicos. Cabe destacar que, dentro de esta subpoblación de pacientes (es decir, pacientes definidos como de bajo riesgo basándose en parámetros clínicos), los parámetros clínicos no son particularmente pronósticos (véase la tabla 17). Por tanto, como prueba diagnóstica, la firma podría ser útil para un gran número de pacientes. En este estudio, casi el 60 % de la cohorte se caracterizó como de bajo riesgo y se esperaba que el 40 % de ellos tuvieran puntuaciones de CCG bajas. Por tanto, la firma de CCG puede predecir una enfermedad indolente en una cuarta parte de los pacientes que han sido identificados previamente como de alto riesgo (y, por tanto, identificados como candidatos a prostatectomía radical). Finalmente, los datos de validación en particular sugieren que la firma de CCG puede ser útil para definir el riesgo en todos los pacientes. Específicamente,

ayudó a dividir a los pacientes definidos como de alto riesgo según los parámetros clínicos en aquellos con tasas de recidiva del 30 % y del 70 % (tabla 18).

La combinación de los parámetros clínicos y la firma de CCG permite a los médicos predecir con mayor precisión el riesgo de fallo quirúrgico y, por tanto, identificar el transcurso apropiado de intervención terapéutica. Tal como se demostró, la firma mejora drásticamente la predicción de recidiva en pacientes que presentan parámetros clínicos generales de enfermedad no agresiva (tabla 19). Dentro de este subgrupo clínico, los pacientes con puntuaciones de CCG bajas se beneficiarían de la seguridad absoluta de que no está indicado ningún tratamiento adicional. Por el contrario, el grupo con CCG alto puede requerir una intervención inmediata. Los pacientes con parámetros clínicos posquirúrgicos desfavorables se benefician de la radioterapia complementaria. Por tanto, la firma de CCG debería predecir la eficacia de la radiación complementaria para pacientes con características clínicas de bajo riesgo y puntuaciones de CCG altas. En la cohorte de validación, los pacientes con puntuaciones de CCG altas y enfermedad más allá de la próstata tienen una tasa de recidiva del 70 %, lo que debería identificar claramente a los pacientes que son buenos candidatos para la radiación complementaria. Por tanto, la combinación de los parámetros clínicos y la firma de CCG conduce claramente a una definición más precisa del riesgo del paciente, lo que debería permitir una evaluación más inteligente de la necesidad de tratamiento adicional.

EJEMPLO 6

Algunos de los paneles de CCG descritos en el presente documento se evaluaron adicionalmente para determinar su capacidad de pronóstico en otros cánceres. Se encontró que los paneles C, D y F fueron pronósticos en diversos grados en el cáncer de vejiga, cerebro, mama y pulmón.

Métodos

Los datos de expresión génica y de pacientes se obtuvieron de los siguientes conjuntos de datos disponibles públicamente: GSE7390 (Desmedt *et al.*, CLIN. CANCER RES. (2007) 13:3207-14; PMID 17545524); GSE11121 (Schmidt *et al.*, CANCER RES. (2008) 68:5405-13; PMID 18593943); GSE8894 (Son *et al.*; sin publicar); Shedden (Shedden *et al.*, NATURE MED. (2008) 14:822; PMID 18641660); GSE4412 (Freije *et al.*, CANCER RES. (2004) 64:6503-10; PMID 15374961); GSE4271 (Phillips *et al.*, CANCER CELL (2006) 9:157-73; PMID 16530701); GSE5287 (Als *et al.*, CLIN. CANCER RES. (2007) 13:4407-14; PMID 17671123). Cada uno de estos conjuntos de datos tiene asociada una descripción detallada de los procedimientos experimentales usados en la recopilación de datos de expresión y de pacientes. Las micromatrices de expresión usadas 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 recopilaron a partir de estos conjuntos de datos y se determinó el nivel de expresión medio de cada panel para cada paciente, cuyo desenlace clínico se conocía (por ejemplo, recidiva, progresión, supervivencia libre de progresión, supervivencia global, etc.). La puntuación de CCG es una expresión promedio de los genes de 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 al gen. La asociación entre la puntuación de CCG y la supervivencia o la recidiva de la enfermedad se sometió a prueba usando un modelo de riesgos proporcionales de Cox univariante y multivariante. El análisis multivariante se realizó cuando se disponía de parámetros clínicos relevantes (grado en el cáncer de cerebro, estadio en el cáncer de pulmón, NPI en el cáncer de mama).

Resultados

Tal como se muestra en la tabla 21 a continuación, cada panel, en el análisis univariante, 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	$2,4 \times 10^{-3}$	$2,3 \times 10^{-3}$	$4,3 \times 10^{-3}$
Cáncer de mama positivo para ER	GSE11121	$1,2 \times 10^{-5}$	$8,7 \times 10^{-6}$	$1,5 \times 10^{-5}$
Adenocarcinoma de pulmón	GSE8894	$2,0 \times 10^{-3}$	$2,5 \times 10^{-3}$	$5,6 \times 10^{-3}$
Adenocarcinoma de pulmón	Shedden	$1,3 \times 10^{-7}$	$2,6 \times 10^{-7}$	$2,2 \times 10^{-7}$
Cáncer de cerebro	GSE4412	$3,2 \times 10^{-5}$	$2,2 \times 10^{-5}$	$9,0 \times 10^{-5}$
Cáncer de cerebro	GSE4271	$1,3 \times 10^{-3}$	$1,0 \times 10^{-3}$	$2,8 \times 10^{-4}$
Cáncer de vejiga	GSE5287	$6,4 \times 10^{-2}$	$5,0 \times 10^{-2}$	$8,6 \times 10^{-2}$

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

Tabla 22

Tipo de cáncer	Conjunto de datos	Valor de p			Nomograma/variable clínica adicional
		Panel C	Panel F	Panel B	
Cáncer de cerebro	GSE4271	0,022	0,017	0,0065	grado
Adenocarcinoma de pulmón	Shedden	1×10^{-6}	$2,1 \times 10^{-6}$	$1,4 \times 10^{-6}$	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 CCG seleccionados según su correlación con la expresión de CCG media.

Tabla 23

Símbolo de gen	Correl. con la media	Símbolo de gen	Correl. con la media	Símbolo de gen	Correl. con la 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>C1orf135</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>PLK1</i>	0,7657	<i>CKAP2</i>	0,6307
<i>FOXM1</i>	0,87	<i>TIMELESS</i>	0,762	<i>CCDC99</i>	0,6303
<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>GIN51</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>KIFC1</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>CKS1B</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>FBXO5</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 larga, pero no exhaustiva, de CCG.

Tabla 1

5

	Gen (nombre y/o símbolo) o número (EST, clon de ADNc, o registro)
1	STK15: serina/treonina cinasa 15 Hs.48915 R11407
2	PLK: cinasa similar a polo (<i>Drosophila</i>) Hs.77597 AA629262
3	UBCH10: proteína portadora de ubiquitina E2-C Hs.93002 AA430504
4	MAPK13: proteína cinasa activada por mitógenos 13 Hs.178695 AA157499 p38delta ARNm=proteína cinasa activada por estrés 4
5	CDC2: ciclo de división celular 2, de G1 a S y de G2 a M Hs.184572 AA598974
6	TOP2A: topoisomerasa (ADN) II alfa (170 kD) Hs.156346 AA504348
7	CENPE: proteína centromérica E (312 kD) Hs.75573 AA402431 CENP-E=supuesto motor de cinetocoro que se acumula justo antes
8	TOP2A: topoisomerasa (ADN) II alfa (170 kD) Hs.156346 AA026682
9	KPNA2: carioferina alfa 2 (cohorte 1 de RAG, importina alfa 1) Hs.159557 AA676460

ES 2 991 436 T3

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 a 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 levaduras) Hs.98658 AA430092 BUB1=supuesta proteína ser/thr cinasa de punto de control mitótico
16	TOP2A: **topoisomerasa (ADN) II alfa (170 kD) Hs.156346 A1734240
17	CKS2: proteína asociada a CDC2 CKS1 Hs.83758 AA010065 ckshs2=homólogo de Cks1=proteína asociada a Cdc2/p34Cdc28
18	ARL6IP: proteína de interacción 6 similar a 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 IPL1/aurora
21	E2-EPF: proteína portadora de ubiquitina Hs.174070 AA464019
22	UBCH10: proteína portadora de ubiquitina E2-C Hs.93002 R80790
23	KNLSL5: similar a cinesina 5 (proteína similar a cinesina mitótica 1) Hs.270845 AA452513 proteína similar a cinesina mitótica 1
24	CENPF: proteína centromérica F (350/400 kD, mitosina) Hs.77204 AA701455
25	CCNA2: ciclina A2 Hs.85137 AA608568 ciclina A
26	CDC2: ciclo de división celular 2, de G1 a S y de G2 a M Hs.184572 AA278152 CDC2=homólogo de proteína de control de división celular 2=proteína cinasa P34
27	HMMR: **receptor de motilidad mediado 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 NUF2R de <i>Homo sapiens</i> , cds completa Hs.234545 AA421171:
32	HSPC216: proteína hipotética Hs.13525 T87341
33	P37NB: proteína de repeticiones ricas en leucina de 37 kDa (LRR) 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	FLJ10491: proteína hipotética FLJ10491 Hs.274283 AA425404
40	KNLSL1: similar a cinesina 1 Hs.8878 AA504625
41	CENPA: proteína centromérica A (17 kD) Hs.1594 A1369629
42	<i>Homo sapiens</i> , clon IMAGE:2823731, ARNm, cds parcial Hs.70704 R96941:
43	CDC6: homólogo de CDC6 (ciclo de división celular 6, <i>S. cerevisiae</i>) Hs.69563 H59203
44	ARNm de homólogo de ADN helicasa de <i>Homo sapiens</i> (PIF1), cds parcial Hs.112160 AA464521:
45	ESTs: Hs.48480 AA135809
46	TSN: translina Hs.75066 AA460927
47	KPNA2: carioferina alfa 2 (cohorte 1 de RAG, importina alfa 1) Hs.159557 AA489087
48	RRM2: polipéptido de ribonucleótido reductasa M2 Hs.75319 AA187351
49	ESTs: Hs.14119 AA204830

ES 2 991 436 T3

50	CCNB1: ciclina B1 Hs.23960 R25788
51	GTSE1: expresada en 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 ácida que contiene hélice superenrollada transformante 3 Hs.104019 AA279990 ARNm de JkR1 regulado por disminución tras activación de células T
54	E2F1: factor de transcripción de E2F 1 Hs.96055 H61303
55	BUB1B: gemación no inhibida por bencimidazoles 1 (homólogo de levaduras), beta Hs.36708 AA488324
56	ESTs: débilmente similar a precursor de cadena de CGHU7L del colágeno alfa 1(III) [<i>H. sapiens</i>] Hs.19322 AA088457
57	KIAA0074: proteína KIAA0074 Hs.1192 N54344
58	MPHOSPH1: fosfoproteína de fase M 1 Hs.240 AA282935
59	ANLN: anilina (homólogo de restos de <i>Drosophila</i>), proteína de unión a actina Hs.62180 R12261
60	BIRC5: proteína baculoviral que contiene repeticiones IAP 5 (survivina) Hs.1578 AA460685 survivina=inhibidor de apoptosis=proteasa de células efectoras EPR-1
61	PTTG1: transformante de tumor de hipófisis 1 Hs.252587 AA430032
62	KIAA0159: proteína asociada a SMC relacionada con la condensación de cromosomas 1 Hs.5719 AA668256
63	ESTs: débilmente similar a proteína OS-4 [<i>H. sapiens</i>] Hs.18714 W93120
64	HMMR: receptor de motilidad mediado por hialuronano (RHAMM) Hs.72550 R10284
65	DKFZp762E1312: proteína hipotética DKFZp762E1312 Hs.104859 AA936181
66	CKAP2: proteína asociada al citoesqueleto 2 Hs.24641 T52152
67	RAMP: proteína asociada a la matriz nuclear regulada por RA
68	SMAP: proteína coactivadora del receptor de la hormona tiroidea Hs.5464 AA481555
69	FLJ22624: proteína hipotética FLJ22624 Hs.166425 AA488791
70	CKS1: proteína asociada a CDC2 CKS1 Hs.77550 N48162
71	NEK2: cinasa relacionada con NIMA (nunca en gen a de mitosis) 2 Hs.153704 W93379
72	MK167: antígeno identificado por 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 factor de crecimiento endotelial vascular VRP
75	CDKN3: inhibidor de cinasa dependiente de ciclina 3 (fosfatasa de doble especificidad asociada a CDK2) Hs.84113 AA284072 CIP2=Cdi1=fosfatasa KAP1=gen del ciclo celular G1/S
76	ARNm de NUF2R de <i>Homo sapiens</i> , cds completa Hs.234545 R92435:
77	ADNc FLJ10325 fis de <i>Homo sapiens</i> , 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	KNLS2: similar a cinesina 2 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	USF1: factor de transcripción aguas 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

ES 2 991 436 T3

89	ESTs: Hs.186579 AA960844
90	CCNB2: ciclina B2 Hs.194698 AA774665
91	<i>Homo sapiens</i> , similar a agrupación ricas en genes, gen C8, clon MGC:2577, ARNm, cds completa Hs.30114 AA634371:
92	ESTs: Hs.99480 AA485454
93	ARNm de IRE1b de <i>Homo sapiens</i> 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: expresado en fase G-2 y S 1 Hs.122552 AA449474
97	CKS1: proteína asociada a CDC2 CKS1 Hs.77550 AA278629
98	CDC25B: proteína 25B de ciclo de división celular Hs.153752 AA448659 cdc25B=fosfatasa inductora de fase M 2
99	ESTs: débilmente similar a producto de proteína sin nombre [<i>H. sapiens</i>] Hs.99807 AA489023 UG desconocido Hs.99807 ESTs sc_id384
100	PCNA: antígeno nuclear de células en proliferación Hs.78996 H05891
101	LTBP3: **proteína de unión a factor de crecimiento transformante latente beta 3 Hs.289019 R60197
102	ARNm de <i>Homo sapiens</i> ; ADNc DKFZp434D0818 (del clon DKFZp434D0818) Hs.5855 N95578:
103	ESTs: Hs.126714 AA919126
104	CIT: citrón (serina/treonina cinasa de interacción con rho, 21) Hs.15767 H10788
105	LBR: receptor de lámina B Hs.152931 AA099136
106	E2F1: factor de transcripción de E2F 1 Hs.96055 AA424949
107	AA699928:
108	CDKN2C: inhibidor de cinasa dependiente de ciclinas 2C (p18, inhibe CDK4) Hs.4854 N72115 p18-INK6=inhibidor de cinasa dependiente de ciclinas 6
109	STK12: serina/treonina cinasa 12 Hs.180655 H81023 ARK2=cinasa relacionada con aurora 2
110	ESTs: Hs.111471 AA682533
111	ESTs: Hs.44269 AA465090
112	MCM4: proteína deficiente en el mantenimiento de minicromosomas (<i>S. cerevisiae</i>) 4 Hs.154443 AA485983
113	PMSCL1: **autoantígeno de polimiositis/esclerodermia 1 (75 kD) Hs.91728 AA458994 ciclina A
114	MKI67: antígeno identificado por anticuerpo monoclonal Ki-67 Hs.80976 AA425973 Ki67 (tipo largo)
115	ESTs: Hs.133294 AI144063
116	CDC25B: ciclo de división celular 25B Hs.153752 H14343 cdc25B=fosfatasa inductora de fase M 2
117	FOX M1: caja de cabeza de horquilla M1 Hs.239 AA129552 MPP2=supuesta fosfoproteína de fase M 2
118	FLJ11029: proteína hipotética FLJ11029 Hs.274448 AI124082
119	H2AFX: familia de histonas H2A, miembro X Hs.147097 H95392
120	FLJ20333: proteína hipotética FLJ20333 Hs.79828 AA147792
121	SLC17A2: familia de portadores de soluto 17 (fosfato de sodio), miembro 2 Hs.19710 H60423
122	ARNm de IRE1b de <i>Homo sapiens</i> para proteína cinasa/ribonucleasa IRE1 beta, cds completa Hs.114905 AA102368:
123	ESTs: Hs.163921 AA573689
124	MCM5: proteína deficiente en el mantenimiento de minicromosomas (<i>S. cerevisiae</i>) 5 (ciclo de división celular 46) Hs.77171 AA283961
125	CDKN1B: inhibidor de cinasa dependiente de ciclinas 1B (p27, Kip1) Hs.238990 AA630082

ES 2 991 436 T3

126	AA779865:
127	PTTG1: transformante de tumor de hipófisis 1 Hs.252587 AI362866
128	RAD21: homólogo de RAD21 (<i>S. pombe</i>) Hs.81848 AA683102
129	ADNc FLJ10325 fis de <i>Homo sapiens</i> , clon NT2RM2000569 Hs.245342 AA430511:
130	NEK2: cinasa relacionada con NIMA (nunca en gen a de mitosis) 2 Hs.153704 AA682321
131	FLJ20101: proteína de interacción con LIS1 NUDE1, 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	KNLS2: 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, rico en repeticiones de heptadas de leucina Hs.58169 W72679
141	TROAP: proteína asociada a 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 específica de estructura Flap 1 Hs.4756 AA620553
148	CKS1: proteína asociada a CDC2 CKS1 Hs.77550 AA459292 ckshs1=homólogo de Cks1=proteína asociada a Cdc2/p34Cdc28
149	ESTs: Hs.193379 N57936
150	CASP8AP2: proteína asociada a CASP8 2 Hs.122843 H50582
151	BIRC2: proteína baculoviral que contiene repeticiones de IAP 2 Hs.289107 R19628 c-IAP1=MIHB=homólogo B de IAP
152	CKAP2: proteína asociada al citoesqueleto 2 Hs.24641 AA504130
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 CASPASA-3=isoforma alfa de CPP32=yama=cisteína proteasa
157	**ESTs: débilmente similar a proteína que es inmunorreactiva con anticuerpos policlonales anti-PTH [<i>H. sapiens</i>] Hs.301486 AA088258
158	HMG2: proteína (cromosómica distinta de histona) con grupos de alta movilidad 2 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 metaloproteínasa 1 (actividad potenciadora de eritroides, inhibidor de colagenasa) Hs.5831 H80214
163	ESTs: Hs.102004 R94281
164	FLJ10858: proteína hipotética FLJ10858 Hs.134403 AA677552

ES 2 991 436 T3

165	ADNc FLJ11883 fis de <i>Homo sapiens</i> , 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 transformante de células epiteliales 2 Hs.132808 AI031571
169	ESTs: Hs.165909 AA629538
170	PCF11: homólogo de PCF11p Hs.123654 AA053411
171	BIRC3: proteína baculoviral que contiene repeticiones de IAP 3 Hs.127799 H48533 c-IAP2=MIHC=homólogo C de IAP=prot. de complejo de señalización de TNFR2-TRAF
172	EST: débilmente similar a dJ45P21.2 [<i>H. sapiens</i>] Hs.326451 AA931528
173	KIAA0952: proteína KIAA0952 Hs.7935 AA454989
174	KIF5B: familia de cinesina, miembro 5B Hs.149436 AA608707
175	DKFZP566C134: proteína DKFZP566C134 Hs.20237 N39306
176	ANLN: anilina (homólogo de restos de <i>Drosophila</i>), proteína de unión a actina Hs.62180 R17092
177	ORC1L: complejo de reconocimiento de origen, similar a subunidad 1 (homólogo de levaduras) Hs.17908 H51719
178	ESTs: Hs.14139 T77757
179	IFIT1: proteína inducida por interferón con repeticiones tetratricopeptídicas 1 Hs.20315 AA074989
180	MGC5338: proteína hipotética MGC5338 Hs.99598 AA463627
181	COPEB: proteína de unión a elemento promotor central Hs.285313 AA013481
182	UK114: proteína inhibidora de la traducción P14.5 Hs.18426 N72715
183	ESTs: Hs.265592 H67282
184	HMG4: proteína (cromosómica distinta de histona) con grupos de alta movilidad 4 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 pleiomórfico 1 Hs.14968 AA418251
190	FLJ23293: **proteína hipotética FLJ23293 similar a proteína de interacción con ARL-6 2 Hs.31236 R91583
191	H11: proteína cinasa H11; proteína similar a proteína de estrés pequeña HSP22 Hs.111676 AA010110
192	POLD3: polimerasa (dirigida por ADN), delta 3 Hs.82502 AA504204
193	SERPINB3: inhibidor de serina (o cisteína) proteinasa, clado B (ovoalbúmina), miembro 3 Hs.227948 AA292860
194	DNAJB1: homólogo de DnaJ (Hsp40), subfamilia B, miembro 1 Hs.82646 AA435948
195	ESTs: Hs.99480 AA458886
196	BUB3: homólogo de BUB3 (gemación no inhibida por bencimidazoles 3, levadura) Hs.40323 AA405690
197	TUBB2: tubulina, beta, 2 Hs.251653 AI000256
198	ARNm de proteína SNC73 (SNC73) de <i>Homo sapiens</i> , cds completa Hs.293441 H28469:
199	BUB3: homólogo de BUB3 (gemación no inhibida por bencimidazoles 3, levadura) Hs.40323 H38804
200	FLJ20699: proteína hipotética FLJ20699 Hs.15125 AA459420
201	KIAA0013: producto génico de KIAA0013 Hs.172652 N63575
202	ESTs: Hs.20575 N20305
203	CDC25C: ciclo de división celular 25C Hs.656 W95000 cdc25C=fosfatasa inductora de fase M 3
204	FLJ11186: proteína hipotética FLJ11186 Hs.89278 AA394225

ES 2 991 436 T3

205	TOPK: cinasa de unión a PDZ; proteína cinasa originada por células T Hs.104741 AA448898
206	KIAA0165: polos adicionales del husillo, <i>S. cerevisiae</i> , homólogo de Hs.153479 AA948058
207	LOC51659: proteína HSPC037 Hs.108196 AA961752
208	ESTs: Hs.10338 AA436456
209	SUCLG2: succinato-CoA ligasa, formador de GDP, subunidad beta Hs.247309 AA465233
210	ZNF265: proteína de dedos de zinc 265 Hs.194718 AA452256
211	SKP2: proteína asociada a cinasa de fase S 2 (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: proteína baculoviral que contiene repeticiones de IAP 2 Hs.289107 AA702174
215	BIRC3: proteína baculoviral que contiene repeticiones de IAP 3 Hs.127799 AA002125 c-IAP2=MIHC=homólogo C de IAP=prot. de complejo de señalización de TNFR2-TRAF
216	INDO: indoleamina-pirrol-2,3-dioxigenasa Hs.840 AA478279
217	DEEPEST: proteína relacionada con la hélice superenrollada del husillo 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 <i>Homo sapiens</i> : FLJ21869 fis, clon HEP02442 Hs.28465 R63929:
221	RGS3: regulador de señalización de proteína G 3 Hs.82294 AI369623
222	ARNm de DC29 de <i>Homo sapiens</i> , cds completa Hs.85573 AA186460:
223	MCM6: proteína deficiente en el mantenimiento de minicromosomas (mis5, <i>S. pombe</i>) 6 Hs.155462 AA663995
224	NPAT: proteína nuclear, locus de ataxia-telangiectasia Hs.89385 AA284172 NPAT=E14=gen en el locus de ATM
225	KNSL6: similar a cinesina 6 (cinesina asociada con centrómeros mitóticos) Hs.69360 AA400450
226	HN1: de expresión hematológica y neurológica 1 Hs.109706 AA459865
227	TUBA3: tubulina, alfa, específica de cerebro Hs.272897 AA865469
228	ESTs: Hs.221197 N55457
229	KIAA0175: producto génico de KIAA0175 Hs.184339 AA903137
230	CLASPIN: homólogo de claspina de <i>Xenopus</i> Hs.175613 AA857804
231	CTNNA1: **catenina (proteína asociada a cadherina), alfa 1 (102 kD) Hs.178452 AA026631
232	ESTs: Hs.221962 AA229644
233	SMC4L1: similar a SMC4 (mantenimiento estructural de 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 asociado a TRAF4 de <i>Homo sapiens</i> 1, cds parcial Hs.181466 T84975:
237	ESTs: Hs.186814 AA700879
238	FLJ11269: proteína hipotética FLJ11269 Hs.25245 R-37817
239	SFPQ: factor de corte y empalme, rico en prolina/glutamina (asociado a 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 <i>Homo sapiens</i> ; ADNc DKFZp434M0435 (del clon DKFZp434M0435) Hs.25700 N94435:
243	FLJ20530: **proteína hipotética FLJ20530 Hs.279521 AA425442

ES 2 991 436 T3

244	BTEB1: proteína de unión a elemento de transcripción básico 1 Hs.150557 N80235
245	LOC51053: geminina Hs.234896 H51100
246	D21S2056E: segmento de ADN en la secuencia expresada 2056 del cromosoma 21 (único) 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 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 C de RAD51 (<i>S. cerevisiae</i>) Hs.11393 R37145 RAD51C=proteína relacionada con Rad51 de recombinación/repación
255	ESTs: Hs.268919 H53508
256	ADNc FLJ11381 fis de <i>Homo sapiens</i> , clon HEMBA1000501 Hs.127797 AA885096:
257	SAP30: polipéptido asociado a sin3, 30 kD Hs.20985 AA126982
258	H4FG: familia de histonas H4, miembro G Hs.46423 AA868008
259	TUBA1: tubulina, alfa 1 (específica de testículo) Hs.75318 AA180742 tubulin-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 proteína de 70 kD de choque térmico 1 Hs.80288 H17513 HSP70-HOM=proteína de 70 kD de choque térmico 1
266	TUBA1: **tubulina, alfa 1 (específica de testículo) Hs.75318 R36063
267	PRO1073: **proteína PRO1073 Hs.6975 AA176999 CIP4=proteína de interacción con Cdc42 4
268	POLD3: polimerasa (dirigida por ADN), delta 3 Hs.82502 AI017254
269	ESTs: moderadamente similar a proteína hipotética T50635, DKFZp762L0311.1, [<i>H. sapiens</i>] Hs.47378 N38809
270	DKFZP564A122: proteína DKFZP564A122 Hs.187991 N57723
271	LRRFIP1: **proteína de interacción con repeticiones ricas en leucina (en FLII) 1 Hs.326159 T84633
272	ESTs: Hs.55468 AA165312
273	ESTs: Hs.31444 H16772
274	AFAP: proteína asociada a filamentos de actina Hs.80306 R69355
275	CXCR4: quimiocina (motivo C-X-C), receptor 4 (fusina) Hs.89414 T62491 receptor de quimiocina CXCR4=fusina=receptor Y de neuropéptido=L3
276	MSH2: **homólogo 2 de mutS (<i>E. coli</i>) (cáncer de colon, tipo 1 distinto de poliposis) Hs.78934 AA679697
277	ESTs: Hs.48474 N62074
278	AA677337:
279	ESTs: moderadamente similar a TBB2, CADENA BETA-2 DE TUBULINA HUMANA [<i>H. sapiens</i>] Hs.23189 AA629908
280	HP1-BP74: HP1-BP74 Hs.142442 H79795
281	FLJ20101: proteína de interacción con LIS1 NUDE1, homólogo de rata Hs.263925 AA459394

ES 2 991 436 T3

282	ARNm de <i>Homo sapiens</i> ; ADNc DKFZp434D1428 (del clon DKFZp434D1428); cds completa Hs.321775 AA431268:
283	ESTs: Hs.265592 AA992658
284	ESTs:
285	DDX11: polipéptido 11 de caja DEAD/H (Asp-Glu-Ala-Asp/His) (helicasa similar a CHL1 de <i>S. cerevisiae</i>) Hs.27424 AA402879
286	CDC27: ciclo de división celular 27 Hs.72405 T81764
287	ARGBP2: proteína de interacción con Arg/Ab1 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, citoplásmico, polipéptido ligero Hs.5120 AA644679
293	ESTs: débilmente similar a LIP1, PRECURSOR DE PROTEÍNA 1 RELACIONADA CON LIPASA PANCREÁTICA HUMANA [<i>H. sapiens</i>] Hs.68864 AA088857
294	HDAC3: histona desacetilasa 3 Hs.279789 AA973283
295	DONSON: vecino aguas 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 detención mitótica, levadura, homólogo) 1 Hs.79078 AA481076 homólogo de proteína de control de retroalimentación mitótica Madp2
299	TASR2: proteína de serina-arginina asociada a TLS 2 Hs.3530 H11042
300	MCM6: proteína deficiente en el mantenimiento de minicromosomas (mis5, <i>S. pombe</i>) 6 Hs.155462 N57722
301	CIT: citrón (serina/treonina cinasa 21 de interacción con rho) Hs.15767 W69425
302	**ESTs: Hs.205066 AA284803
303	ICAM1: molécula de adhesión intercelular 1 (CD54), receptor de rinovirus humano Hs.168383 R77293 CD54=ICAM-1
304	KIAA0855: golgina-67 Hs.182982 AA456818
305	ESTs: débilmente similar a supuesta p150 [<i>H. sapiens</i>] Hs.300070 R10422
306	DEEPEST: proteína relacionada con la hélice superenrollada del husillo mitótico Hs.16244 A1652290
307	MCM2: proteína deficiente en el mantenimiento de minicromosomas (<i>S. cerevisiae</i>) 2 (mitotina) Hs.57101 AA454572
308	ADNc de <i>Homo sapiens</i> : FLJ22272 fis, clon HRC03192 Hs.50740 AA495943:
309	WISP1: **proteína WNT1 de la ruta de señalización inducible 1 Hs.194680 T54850
310	KIAA0855: golgina-67 Hs.182982 AA280248
311	TEM8: marcador endotelial tumoral 8 Hs.8966 H58644
312	BITE: proteína de unión a p10 Hs.42315 H96392
313	RAN: RAN, miembro de la familia de oncogenes de RAS Hs.10842 AA456636
314	EZH2: homólogo de potenciador de zeste (<i>Drosophila</i>) 2 Hs.77256 AA428252
315	MCM4: proteína deficiente en el mantenimiento de minicromosomas (<i>S. cerevisiae</i>) 4 Hs.154443 W74071
316	DKFZp434J0310: proteína hipotética Hs.278408 AA279657 UG desconocido Hs.23595 ESTs sc_id6950
317	PPP1R10: proteína fosfatasa 1, subunidad reguladora 10 Hs.106019 AA071526
318	H11: proteína cinasa H11; proteína similar a proteína de estrés pequeña HSP22 Hs.111676 H57493

ES 2 991 436 T3

319	ESTs: débilmente similar a proteína KIAA1074 [<i>H. sapiens</i>] Hs.200483 AA463220
320	ESTs: débilmente similar a ALU8, ENTRADA DE ADVERTENCIA DE CONTAMINACIÓN DE SECUENCIA SX DE SUBFAMILIA DE ALU HUMANA [<i>H. sapiens</i>] Hs.226414 N72576
321	AA775033:
322	LOC51004: proteína CGI-10 Hs.12239 AA677920
323	ESTs: Hs.150028 AI292036
324	MCM6: proteína deficiente en el mantenimiento de minicromosomas (mis5, <i>S. pombe</i>) 6 Hs.155462 AA976533
325	ESTs: moderadamente similar a proteína hipotética T50635, DKFZp762L0311.1, [<i>H. sapiens</i>] Hs.47378 AA406348
326	UCP4: proteína de desacoplamiento 4 Hs.40510 H60279
327	MSH5: homólogo 5 de mutS (<i>E. coli</i>) Hs.112193 AA621155
328	ROCK1: proteína cinasa asociada a Rho que contiene hélice superenrollada 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 aguas abajo de SON Hs.17834 AI732249
333	SH3GL2: similar a GRB2 de dominio SH3 2 Hs.75149 R12817
334	PRC1: proteína reguladora de citocinesis 1 Hs.5101 AA449336
335	ESTs: débilmente similar a producto de proteína sin nombre [<i>H. sapiens</i>] Hs.99807 AA417744 UG desconocido Hs.119424 ESTs sc_id2235
336	Humano: secuencia de ARNm del clon 23719 Hs.80305 AA425722
337	ARNm de <i>Homo sapiens</i> ; ADNc DKFZp564O2364 (del clon DKFZp564O2364) Hs.28893 W90240:
338	ESTs: débilmente similar a LIP1, PRECURSOR DE PROTEÍNA 1 RELACIONADA CON LIPASA PANCREÁTICA HUMANA [<i>H. sapiens</i>] Hs.68864 AA132858
339	TUBA3: 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	SFRS3: factor de corte y empalme, rico en arginina/serina 3 Hs.167460 AA598400
345	PRIM1: primasa, polipéptido 1 (49 kD) Hs.82741 AA025937 ADN primasa (subunidad p48)
346	FLJ20333: proteína hipotética FLJ20333 Hs.79828 H66982
347	HSPA8: proteína de 70 kD de choque térmico 8 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.142442 T84669
351	ETV4: gen 4 de variante de ets (proteína de unión a potenciador E1A, E1AF) Hs.77711 AA010400 E1A-F=proteína de unión a potenciador E1A=variante de translocación de ETS
352	ADNc de <i>Homo sapiens</i> : FLJ23037 fis, clon LNG02036, altamente similar a HSU68019, ARNm de homólogo de proteína mad de <i>Homo sapiens</i> (hMAD-3) Hs.288261 W42414 Smad3=hMAD-3=homólogo de madres contra decapentapléjico (M:
353	KIAA0952: proteína KIAA0952 Hs.7935 AA679150
354	STK9: serina/treonina cinasa 9 Hs.50905 N80713
355	NXF1: **factor de exportación de ARN nuclear 1 Hs.323502 R01238

ES 2 991 436 T3

356	FLJ12892: proteína hipotética FLJ12892 Hs.17731 AA449357
357	UNG: uracil-ADN glicosilasa Hs.78853 H15111
358	STK17B: **serina/treonina cinasa 17b (de inducción de apoptosis) Hs.120996 AA419485
359	YWHAH: proteína de activación de 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, translocada en el cromosoma X Hs.153221 N59206
363	KIAA0978: proteína KIAA0978 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: proteína de interacción con RAD51 Hs.24596 AI214426
369	FLJ20364: proteína hipotética FLJ20364 Hs.32471 AA676296
370	EIF4A2: **factor de inicio de traducción eucariota 4A, isoforma 2 Hs.173912 H54751
371	ESTs: débilmente similar a MCAT, PROTEÍNA PORTADORA DE CARNITINA/ACILCARNITINA MITOCONDRIAL HUMANA [<i>H. sapiens</i>] Hs.27769 AA469975
372	FLJ11323: proteína hipotética FLJ11323 Hs.25625 AA775600
373	DKFZP564D0764: proteína DKFZP564D0764 Hs.26799 AA460732
374	CTL2: gen de CTL2 Hs.105509 AA454710
375	ESTs: Hs.293419 AA775845
376	IFIT1: proteína inducida por interferón con repeticiones tetratricopeptídicas 1 Hs.20315 AA489640 proteína de 56 KDa inducida por interferón
377	RBBP8: proteína de unión a retinoblastoma 8 Hs.29287 H23021
378	**Secuencia de ARNm del clon 25061 de <i>Homo sapiens</i> Hs.183475 R38944:
379	Humano: secuencia de ADN del clon RP3-383J4 en el cromosoma 1q24.1-24.3 que contiene parte de un gen que codifica para una proteína que contiene motivos kelch, parte de un nuevo gen que codifica para una proteína similar a aspartil-TRNA sy Hs.117305 N29457
380	FLJ12888: proteína hipotética FLJ12888 Hs.284137 N68390
381	ESTs: débilmente similar a IF38, SUBUNIDAD 8 DEL FACTOR DE INICIO DE TRADUCCIÓN EUCARIOTA HUMANO 3 [<i>H. sapiens</i>] Hs.222088 AI139629
382	ESTs: Hs.241101 AA133590
383	H4FI: familia de histonas H4, miembro I Hs.143080 AI218900
384	SP38: proteína de unión a la zona pelúcida Hs.99875 AA400474
385	GABPB1: factor de transcripción de proteína de unión a GA, subunidad beta 1 (53 kD) Hs.78915 1191651
386	LCHN: proteína LCHN Hs.12461 AA029330
387	DKFZP564D0462: proteína hipotética DKFZp564D0462 Hs.44197 N32904
388	LENG8: nuevo gen codificado por agrupación del receptor de leucocitos (LRC) 8 Hs.306121 AA464698
389	HIF1A: factor inducible por hipoxia 1, subunidad alfa (factor de transcripción de hélice-bucle-hélice básico) 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 GrpE Hs.151903 H55907
395	CDK7: cinasa dependiente de ciclinas 7 (homólogo de cinasa de activación de cdk MO15 de <i>Xenopus</i>) Hs.184298 R22624 CAK=cdk7=NRTALRE=sdk=cinasa de activación de CDK
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 dedos de zinc de tipo Kruppel Hs.25275 N50827
400	FLJ20101: proteína de interacción con LIS1 NUDE1, homólogo de rata Hs.263925 R87716
401	KNLSL4: similar a cinesina 4 Hs.119324 AA430503
402	E2F5: factor 5 de transcripción de E2F, de unión a p130 Hs.2331 AA455521 E2F-5=factor de transcripción de unión a pRB
403	TMPO: timopoyetina Hs.11355 T63980
404	POLQ: polimerasa (dirigida por ADN), theta Hs.241517 AI057325
405	TGIF: factor de interacción con TG (homeosecuencia de la familia de TALE) Hs.90077 H51705
406	TRIP13: interactor del receptor de la hormona tiroidea 13 Hs.6566 AA630784
407	GAS6: proteína específica de la detención del crecimiento 6 Hs.78501 AA461110
408	HN1: de expresión hematológica y neurológica 1 Hs.109706 AA035429
409	BARD1: dominio 1 de RING asociado a BRCA1 Hs.54089 AA558464
410	DHFR: dihidrofolato reductasa Hs.83765 AA424790
411	AA490946:
412	ESTs: Hs.130435 AA167114
413	HSPA8: proteína de 70 kD de choque térmico 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 a elemento promotor central Hs.285313 AA055584 CPBP=CBA1=proteína de unión a ADN
417	FLJ10604: proteína hipotética FLJ10604 Hs.26516 N72697
418	ESTs: débilmente similar a ADNc EST yk415c12.5 procede de este gen [<i>C. elegans</i>] Hs.108824 H97880
419	UBE2D3: **enzima de conjugación con ubiquitina E2D 3 (homóloga a UBC4/5 de levadura) Hs.118797 AA017199
420	FLJ10890: **proteína hipotética FLJ10890 Hs.17283 AA004210
421	ESTs: Hs.214410 AA579336
422	OLR1: receptor de lipoproteína de baja densidad oxidada (similar a lectina) 1 Hs.77729 AA682386
423	FLJ13231: proteína hipotética FLJ13231 Hs.156148 W92787
424	EST: Hs.323101 W40398
425	ESTs: débilmente similar a R06F6.5b [<i>C. elegans</i>] Hs.180591 N59330
426	ADNc de <i>Homo sapiens</i> : FLJ23285 fis, clon HEP09071 Hs.90424 N26163:
427	ADNc de inserto de longitud completa de ARNm de <i>Homo sapiens</i> , clon EUROIMAGE 42408 Hs.284123 AA211446:
428	NFKB1: factor nuclear de potenciador de genes de polipéptido ligero kappa en células B 1 (p105) Hs.83428 AA451716 NFKB1=NF-kappaB p105=p50
429	LOC58486: proteína similar a transposasa Buster1 derivada de transposones Hs.25726 AA630256
430	ADNc FLJ10976 fis de <i>Homo sapiens</i> , clon PLACE1001399 Hs.296323 AA424756:

ES 2 991 436 T3

431	KIAA0182: proteína KIAA0182 Hs.75909 AI023801
432	RANGAP1: proteína de activación de GTPasa Ran 1 Hs.183800 AA991855
433	PKMYT1: cinasa inhibidora de cdc2 específica de tirosina y treonina asociada a la membrana Hs.77783 AA478066 cinasa Myt1
434	HSPA8: proteína de 70 kD de choque térmico 8 Hs.180414 H64096
435	LUC7A: proteína sobreexpresada asociada a resistencia a cisplatino Hs.3688 AA411969
436	RRM1: polipéptido M1 de ribonucleótido reductasa Hs.2934 AA633549
437	SET07: proteína que contiene dominios PR/SET 7 Hs.111988 AA421470
438	**ESTs: débilmente similar a ALU1, ENTRADA DE ADVERTENCIA DE CONTAMINACIÓN DE SECUENCIA J DE LA SUBFAMILIA ALU HUMANA [<i>H. sapiens</i>] Hs.193452 W96179
439	Secuencia de ARNm del clon 25058 de <i>Homo sapiens</i> Hs.179397 R38894:
440	ESTs: débilmente similar a proteína KIAA0973 [<i>H. sapiens</i>] Hs.14014 AA780791
441	EST: Hs.105298 AA489813
442	CTCF: factor de unión a CCCTC (proteína de dedos de zinc) Hs.57419 H89996
443	HRB: proteína de unión a Rev de VIH-1 Hs.171545 AA485958
444	**ESTs: Hs.294083 AA447679
445	KIAA0878: proteína KIAA0878 Hs.188006 AA599094
446	ESTs: débilmente similar a ALUB !!!! ENTRADA DE ADVERTENCIA DE CLASE B DE ALU HUMANA!!! [<i>H. sapiens</i>] Hs.180552 AA481283
447	OGT: N-acetilglucosamina (GlcNAc) transferasa unida a O (UDP-N-acetilglucosamina:polipéptido-N-acetilglucosaminiltransferasa) Hs.100293 AA425229
448	ARNm de <i>Homo sapiens</i> para proteína KIAA1700, cds parcial Hs.20281 N40952:
449	Humano: secuencia de ADN del clon RP1-187J11 en el cromosoma 6q11.1-22.33. Contiene el gen para una nueva proteína similar a proteínas predichas de <i>S. pombe</i> y <i>S. cerevisiae</i> , el gen para una nueva proteína similar Hs.72325 AA159962
450	KIAA1265: proteína KIAA1265 Hs.24936 AA479302
451	H1F0: familia de histonas H1, 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 espermatozoides 2 Hs.129055 AA149882
454	CD97: antígeno CD97 Hs.3107 AI651871
455	BMI1: **homólogo de oncogén viral de 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 a factor de ribosilación de ADP 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	FLJ10339: proteína hipotética FLJ10339 Hs.203963 AA628231
465	DR1: regulador por disminución de la transcripción 1, de unión a TBP (cofactor negativo 2) Hs.16697 AA043503
466	<i>Homo sapiens</i> , similar a proteína hipotética FLJ20093, clon MGC:1076, ARNm, cds completa Hs.298998 AA703249:
467	HSPC163: proteína HSPC163 Hs.108854 H98963

ES 2 991 436 T3

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 región crítica de síndrome de Down Hs.184222 AA629707
471	FLJ10342: proteína hipotética FLJ10342 Hs.101514 AA490935
472	ARNm de <i>Homo sapiens</i> ; ADNc DKFZp586N1323 (del clon DKFZp586N1323) Hs.24064 R26176:
473	ESTs: Hs.4983 H59921
474	ESTs: débilmente similar a ALUB, !!!! ENTRADA DE ADVERTENCIA DE CLASE B DE ALU HUMANA!!! [<i>H. sapiens</i>] Hs.117949 H91167
475	CDC45L: similar a CDC45 (ciclo de división celular 45, <i>S. cerevisiae</i> , homólogo) Hs.114311 AA700904
476	STAT5B: transductor de señales y activador de la transcripción 5B Hs.244613 AA280647 STAT5A/5B
477	ADNc FLJ14028 fis de <i>Homo sapiens</i> , clon HEMBA1003838 Hs.281434 AA454682:
478	KIAA1524: proteína KIAA1524 Hs.151343 AI248987
479	CTSD: catepsina D (aspartil proteasa lisosómica) Hs.79572 AA485373
480	<i>Homo sapiens</i> , similar a proteína hipotética FLJ20093, clon MGC:1076, ARNm, cds completa Hs.298998 AA682274:
481	GTPBP2: proteína de unión a GTP 2 Hs.13011 T67069
482	LOC51003: proteína CGI-125 Hs.27289 AA485945
483	VCL: vinculina Hs.75350 AA486727
484	KIF5B: miembro 5B de la familia de cinesinas Hs.149436 AA046613
485	CDC25A: ciclo de división celular 25A Hs.1634 AA071514
486	LOC51141: proteína inducida por insulina 2 Hs.7089 AA045308
487	**ESTs: moderadamente similar a CALD, CALDESMONA HUMANA [<i>H. sapiens</i>] Hs.117774 H48508
488	TBX3-iso: proteína TBX3-iso Hs.267182 T48941
489	KIAA0176: proteína KIAA0176 Hs.4935 R44371
490	PRKAR1A: proteína cinasa, dependiente de cAMP, reguladora, tipo I, alfa (extinguidor específico de tejido 1) Hs.183037 N25969 PKA-R1 alfa=proteína cinasa dependiente de cAMP de tipo I-alfa-cata
491	ESTs: Hs.268991 H77818
492	ESTs: débilmente similar a isopentenil-difosfato delta-isomerasa A53028 [<i>H. sapiens</i>] Hs.9270 R17362
493	ESTs: débilmente similar a proteína hipotética B34087 [<i>H. sapiens</i>] Hs.120946 H50656
494	TRN2: carioferina beta 2b, transportina Hs.278378 R08897
495	LMNA: lamina A/C Hs.77886 AA489582
496	NFE2L2: similar a factor nuclear 2 (derivado de eritroides 2) Hs.155396 AA629687
497	DKFZp762L0311: proteína hipotética DKFZp762L0311 Hs.16520 AA486418
498	ESTs: débilmente similar a S71752, proteína gigante p619 [<i>H. sapiens</i>] Hs.14870 T96829
499	ARNm de <i>Homo sapiens</i> ; ADNc DKFZp434A1315 (del clon DKFZp434A1315); cds completa Hs.298312 AA991355:
500	E2IG4: proteína hipotética, inducida por estradiol Hs.8361 R13844
501	RANGAP1: proteína de activación de GTPasa Ran 1 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 ALU7, ENTRADA DE ADVERTENCIA DE CONTAMINACIÓN DE SECUENCIA SQ DE LA SUBFAMILIA DE ALU HUMANA [<i>H. sapiens</i>] Hs.68647 R96804
505	PRO0650: proteína PRO0650 Hs.177258 N54333
506	DNAJB9: homólogo de DnaJ (Hsp40), subfamilia B, miembro 9 Hs.6790 AA045792

ES 2 991 436 T3

507	ADNc de <i>Homo sapiens</i> : FLJ21971 fis, clon HEP05790 Hs.71331 AA774678:
508	LOC56996: **proteína de interacción con cotransportador de cationes-cloruro Hs.119178 AA037466
509	AP3D1: complejo 3 de proteína relacionada con adaptador, subunidad delta 1 Hs.75056 AA630776
510	SGK: cinasa regulada por suero/glucocorticoides Hs.159640 AA486082 sgk=supuesta serina/treonina proteína cinasa transcripcional
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 Hs.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	FLJ23151: proteína hipotética FLJ23151 Hs.137260 AA284259
521	AMD1: S-adenosilmetionina descarboxilasa 1 Hs.262476 AA425692
522	FLJ10342: **proteína hipotética FLJ10342 Hs.101514 AA934516
523	SPS: SELENOFOSFATO SINTETASA; proteína donadora de selenio humana Hs.124027 AA486372
524	KIAA1586: proteína KIAA1586 Hs.180663 AA779733
525	ICBP90: factor de transcripción Hs.108106 AA908902
526	ADNc de <i>Homo sapiens</i> : 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 de disociación de Rho GDP (GDI) beta Hs.83656 AA487426 LyGDI=inhibidor de disociación de Rho GDP 2=RHO GDI 2
529	RAD53: proteína cinasa Chk2 Hs.146329 AI653182
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 <i>Homo sapiens</i> ; ADNc DKFZp586N1323 (del clon DKFZp586N1323) Hs.24064 R30941:
535	AA452872:
536	ESTs: Hs.124169 R58970
537	ACYP1: acilfosfatasa 1, de tipo eritrocitos (habitual) Hs.18573 W78754
538	SIL: locus interruptor de TAL1 (SCL) Hs.323032 AA704809
539	AA016234:
540	ARNm de <i>Homo sapiens</i> ; ADNc DKFZp566P1124 (del clon DKFZp566P1124) Hs.321022 N50895:
541	KIAA1067: proteína KIAA1067 Hs.325530 AA099138
542	SMC4L1: similar a SMC4 (mantenimiento estructural de cromosomas 4, levadura) 1 Hs.50758 AA283006
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 <i>Homo sapiens</i> : FLJ21531 fis, clon COL06036 Hs.102941 N95440:
547	ARNm de <i>Homo sapiens</i> ; ADNc DKFZp547B086 (del clon DKFZp547B086) Hs.36606 N48700:

ES 2 991 436 T3

548	C1ORF2: marco de lectura abierto 2 del cromosoma 1 Hs.19554 H11464 cote1=ORF en el locus de glucocerebrosidasa
549	HTF9C: locus de diminutos fragmentos de HpalI 9C Hs.63609 H17888
550	ATF4: factor de transcripción activante 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: cinasa de unión a PDZ; proteína cinasa originada por células T Hs.104741 AI002631
554	ASIP: proteína de señalización de agutí (ratón) Hs.37006 AI220203
555	DKFZP564F013: **proteína hipotética DKFZp564F013 Hs.128653 R14908
556	ZNF265: proteína de dedos de zinc 265 Hs.194718 N66014
557	SLC30A1: familia de portadores de soluto 30 (transportador de zinc), 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 <i>Homo sapiens</i> : FLJ23538 fis, clon LNG08010, altamente similar a ARNm del clon épsilon/beta de la región de MEN1 humana BETA2 Hs.240443 AA400234:
564	AMD1: S-adenosilmetionina descarboxilasa 1 Hs.262476 R82299
565	MAP3K2: proteína cinasa activada por mitógenos 2 Hs.28827 AA447971
566	NET1: gen transformante de células neuroepiteliales 1 Hs.25155 R24543
567	CHAF1A: factor de ensamblaje de cromatina 1, subunidad A (p150) 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
571	ESTs: Hs.238797 N70848
572	ESTs: débilmente similar a ALUB, !!!! ENTRADA DE ADVERTENCIA DE CLASE B HUMANA!!! [<i>H. sapiens</i>] Hs.180552 AA600192
573	PDGFA: polipéptido alfa de factor de crecimiento derivado de plaquetas Hs.37040 AA701502
574	ARNm del clon FLC0675 PRO2870 de <i>Homo sapiens</i> , cds completa Hs.306117 AA443127:
575	ESTs: Hs.143375 AA001841
576	TUBB: tubulina, polipéptido beta Hs.179661 H37989
577	MSH2: homólogo 2 de mutS (<i>E. coli</i>) (cáncer de colon, tipo 1 distinto de poliposis) Hs.78934 AA219060 MSH2=homólogo de mutS de reparación de apareamiento erróneo de ADN
578	TOPBP1: proteína de unión a topoisomerasa (ADN) II Hs.91417 R97785
579	KIAA0869: proteína KIAA0869 Hs.21543 R43798
580	H4FH: familia de histonas H4, miembro H Hs.93758 AA702781
581	FLJ23293: proteína hipotética FLJ23293 similar a proteína de interacción con ARL-6 2 Hs.31236 AA629027
582	**ADNc de <i>Homo sapiens</i> : FLJ23538 fis, clon LNG08010, altamente similar a ARNm del clon épsilon/beta de la región de MEN1 humana BETA2 Hs.240443 AA053165:
583	KIAA0978: proteína KIAA0978 Hs.3686 N64780
584	KIAA1547: proteína KIAA1547 Hs.31305 AA057737
585	DKFZP761C169: proteína hipotética DKFZp761C169 Hs.71252 AA608709

ES 2 991 436 T3

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, aparición temprana Hs.194143 H90415 BRCA1=mutado en cáncer de mama y de ovario
589	ESTs: Hs.4983 H22936
590	HSPC150: proteína HSPC150 similar a enzima de conjugación con ubiquitina Hs.5199 AA460431
591	ARNm de <i>Homo sapiens</i> para proteína KIAA1712, cds parcial Hs.29798 H54592:
592	FLJ11186: proteína hipotética FLJ11186 Hs.89278 AA504111 UG desconocido Hs.89278 ESTs
593	ESTs: débilmente similar a producto de proteína sin nombre [<i>H. sapiens</i>] Hs.118338 R25481
594	APEXL2: proteína similar a endonucleasa apurínica/apirimidínica (nucleasa APEX) 2 Hs.154149 AI674393
595	CDR2: proteína relacionada con la degeneración cerebelosa (62 kD) Hs.75124 AA074613
596	ESTs: Hs.69662 AA459724
597	PSCD2L: homólogo de pleckstrina, similar a dominios de Sec7 y de hélice superenrollada 2 Hs.8517 AA464957
598	CRK: homólogo de oncogén de CT10 del virus de sarcoma aviar v-crk Hs.306088 H75530
599	CCNE2: ciclina E2 Hs.30464 AA520999 UG desconocido 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 <i>Homo sapiens</i> : FLJ22355 fis, clon HRC06344 Hs.288283 AA026375:
604	Humano: secuencia de ARNm del clon 23719 Hs.80305 H43437
605	ARNm del clon FLC0675 PRO2870 de <i>Homo sapiens</i> , cds completa Hs.306117 AA485453:
606	MSE55: proteína constituyente del suero Hs.148101 H73234
607	CFLAR: regulador de apoptosis similar a CASP8 y FADD Hs.195175 AA453766
608	ADNc de <i>Homo sapiens</i> : FLJ22844 fis, clon KAIA5181 Hs.296322 AA975103:
609	Humano: secuencia de ADN del clon RP11-371L19 en el cromosoma 20 que contiene dos nuevos genes, el gen para una nueva proteína similar a proteína ribosómica 40S S10 (RPS10), ESTs, STSs, GSSs y cinco islas de CpG Hs.19002 R00846
610	ESTs: Hs.60054 R26390
611	ESTs: débilmente similar a ALU7, ENTRADA DE ADVERTENCIA DE CONTAMINACIÓN DE SECUENCIA SQ DE LA SUBFAMILIA DE ALU HUMANA [<i>H. sapiens</i>] Hs.325158 AA032084
612	FLJ10980: proteína hipotética FLJ10980 Hs.29716 N45467
613	IFIT1: **proteína inducida por interferón con repeticiones tetratricopeptídicas 1 Hs.20315 AA157787
614	ESTs: Hs.21734 AA429809
615	DKFZP434C245: proteína DKFZP434C245 Hs.59461 AA705518
616	RNPS1: proteína de unión a ARN S1, dominio rico en serina Hs.75104 AA496837
617	FLJ13639: proteína hipotética FLJ13639 Hs.101821 AA131681
618	PCF11: homólogo de PCF11p Hs.123654 W73749
619	EIF4G3: factor de inicio de traducción eucariota 4 gamma, 3 Hs.25732 N92469
620	ADNc de <i>Homo sapiens</i> : FLJ21971 fis, clon HEP05790 Hs.71331 AA130595:
621	STAT1: transductor de señales y activador de la transcripción 1, 91 kD Hs.21486 AA079495
622	BIRC3: proteína baculoviral que contiene repeticiones de IAP 3 Hs.127799 R07870
623	HP1-BP74: HP1-BP74 Hs.142442 N20589
624	HSPC228: proteína hipotética Hs.267288 AI734268

625	KIAA0675: producto génico de KIAA0675 Hs.165662 AA454867
626	AMD1: S-adenosilmetionina descarboxilasa 1 Hs.262476 AA504772
627	EST: Hs.149338 AI249089
628	PWP1: fosfoproteína nuclear similar a PWP1 de <i>S. cerevisiae</i> Hs.172589 AA485992
629	AI336973:
630	DUSP4: fosfatasa de doble especificidad 4 Hs.2359 AA444049
631	FLJ12788: proteína hipotética FLJ12788 Hs.20242 AA497041
632	HSPC150: proteína HSPC150 similar a enzima de conjugación con ubiquitina Hs.5199 AA985450
633	FLJ11729: proteína hipotética FLJ11729 Hs.286212 W15533
634	KLF4: factor similar a Kruppel 4 (intestino) Hs.7934 H45668
635	FLJ11058: proteína hipotética FLJ11058 Hs.180817 N63911
636	FLJ23468: proteína hipotética FLJ23468 Hs.38178 AA460299
637	ESTs: Hs.115315 AI278336
638	EBI3: gen 3 inducido por virus de Epstein-Barr Hs.185705 AA425028 EBI3=receptor de citocinas
639	ESTs: Hs.293797 N63988
640	MGAT2: manosil (alfa-1,6-)-glicoproteína beta-1,2-N-acetilglucosaminiltransferasa Hs.172195 AA485653
641	H2BFQ: familia de histonas H2B, miembro Q Hs.2178 AA456298
642	NMB: neuromedina B Hs.83321 AI650675
643	SSR3: receptor de secuencia señal, gamma (proteína asociada a translocos gamma) Hs.28707 AA453486
644	HSPC196: proteína hipotética Hs.239938 R78498
645	EST: Hs.44522 N33610
646	BRF1: factor de respuesta a butirato 1 (factor de respuesta a EGF 1) Hs.85155 AA723035
647	MAN1A2: manosidasa, alfa, clase 1A, miembro 2 Hs.239114 H97940
648	KIAA1201: proteína KIAA1201 Hs.251278 AA427719
649	NUCKS: similar a caseína cinasa ubicua nuclear de rata 2 Hs.118064 AA158345
650	MAGEF1: proteína MAGEF1 Hs.306123 AA425302
651	Humano: cromosoma 16 de BAC, clon CIT987SK-A-362G6 Hs.6349 N75498
652	R40377:
653	AP3M2: complejo 3 de proteína relacionada con adaptador, subunidad mu 2 Hs.77770 R14443
654	ESTs: débilmente similar a proteína relacionada con transcriptasa inversa 1207289A [<i>H. sapiens</i>] Hs.272135 AA705010
655	ARNm de <i>Homo sapiens</i> para proteína FLJ00116, cds parcial Hs.72363 AA159893:
656	EIF4E: factor de inicio de la traducción eucariota 4E Hs.79306 AA193254
657	ARNm de <i>Homo sapiens</i> para proteína hipotética (gen de 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 dedos de zinc 217 Hs.155040 R81830
662	FLJ12892: proteína hipotética FLJ12892 Hs.17731 AI243595
663	ETV5: gen 5 de variante 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 placofilina 2b [<i>H. sapiens</i>] Hs.12705 AA485365
668	PAPPA: proteína plasmática asociada al embarazo A Hs.75874 AA609463
669	TFF3: factor de 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 IEFs, PROTEÍNA SENSIBLE A TRANSFORMACIÓN HUMANA IEF SSP 3521 [<i>H. sapiens</i>] Hs.43213 AA490554
674	FLJ13181: proteína hipotética FLJ13181 Hs.301526 AA057266
675	KIAA1547: proteína KIAA1547 Hs.31305 AA136692
676	ZNF281: proteína de dedos de zinc 281 Hs.59757 N47468
677	ADNc de <i>Homo sapiens</i> : FLJ23260 fis, clon COL05804, altamente similar a secuencia de ARNm del clon 23652 humano HSU90911 Hs. 13996 AA463961:
678	ESTs: Hs.25933 AA411392
679	NCBP1: subunidad 1 de proteína de unión a caperuza nuclear, 80 kD Hs.89563 AA278749 proteína de unión a caperuza nuclear
680	H2BFL: familia de histonas H2B, miembro L Hs.239884 H70774
681	DKFZP564A122: proteína DKFZP564A122 Hs.187991 H66150
682	NASP: proteína espermática autoantigénica nuclear (de unión a histonas) Hs.243886 AA644128
683	**ESTs: débilmente similar a proteína KIAA0822 [<i>H. sapiens</i>] Hs.98368 AA422008
684	MAP2K6: proteína cinasa activada por mitógenos 6 Hs.118825 H07920
685	ESTs: Hs.158357 AA865842
686	GADD45A: detención del crecimiento e inducible por daño al ADN, alfa Hs.80409 AA147214 GADD45 alfa=proteína de detención del crecimiento e inducible por daño al ADN
687	DHFR: dihidrofolato reductasa Hs.83765 AA488803
688	AA151930:
689	ARNm de <i>Homo sapiens</i> ; ADNc DKFZp434P116 (del clon DKFZp434P116); cds completa Hs.103378 AA431133:
690	ARNm de <i>Homo sapiens</i> ; ADNc DKFZp564D156 (del clon DKFZp564D156) Hs.9927 T55704:
691	ESTs: Hs.32204 R93719
692	PRPSAP1: proteína asociada a fosforribosil pirofosfato sintetasa 1 Hs.77498 R20005
693	ZNF42: proteína de dedos de zinc 42 (sensible a ácido retinoico específico de mieloides) Hs.169832 AA987906
694	.**ESTs: Hs.43712 N25936
695	RUNX1: factor de transcripción relacionado con Runt 1 (leucemia mieloides aguda 1; oncogén de aml1) Hs.129914 AA146826
696	ARNm de <i>Homo sapiens</i> ; ADNc DKFZp547C244 (del clon DKFZp547C244) Hs.9460 T64452:
697	TYMS: timidilato sintetasa Hs.82962 AA663310
698	MGC5528: proteína hipotética MGC5528 Hs.315167 AA843451
699	ESTs: Hs.268685 R22952
700	SFPQ: factor de corte y empalme, rico en prolina/glutamina (asociado a proteína de unión al tracto de polipirimidina) Hs.180610 AA418910
701	ESTs: Hs.155105 AI221390
702	FLJ10624: proteína hipotética FLJ10624 Hs.306000 AA489592

ES 2 991 436 T3

703	TRIP8: interactor del receptor de la hormona tiroidea 8 Hs.6685 AA425205
704	DNAJB6: homólogo de DnaJ (Hsp40), subfamilia B, miembro 6 Hs.181195 AA496105
705	ESTs: Hs.18331 T98244
706	RBM14: proteína de motivos de unión a ARN 14 Hs.11170 AA421233
707	SCYA2: citocina inducible pequeña A2 (proteína quimiotáctica de monocitos 1, homóloga a Sig-je de ratón) Hs.303649 AA425102 MCP-1=MCAF=citocina inducible pequeña A2=JE=quimiocina
708	MGC4161: proteína hipotética MGC4161 Hs.177688 AI224867
709	TUBB2: tubulina, beta, 2 Hs.251653 AA888148
710	FLJ20280: proteína hipotética FLJ20280 Hs.270134 N74086
711	TERA: proteína TERA Hs.180780 AA465096
712	CPS1: **carbamoil-fosfato sintetasa 1, mitocondrial Hs.50966 N68399
713	KIAA0802: proteína KIAA0802 Hs.27657 W55875
714	FYN: oncogén de FYN relacionado con SRC, FGR, YES Hs.169370 N22980
715	ARNm de PRO2751 de <i>Homo sapiens</i> , cds completa Hs.283978 H12784:
716	CLTH: gen de leucemia mieloide linfóide de ensamblaje de clatrina Hs.7885 AA441930
717	CHMP1.5: proteína CHMP1.5 Hs.42733 W85875
718	SMARCB1: regulador de la cromatina dependiente de actina, relacionado con SWI/SNF, asociado a matriz, subfamilia b, miembro 1 Hs.159971 AA446018
719	AA487823: SRF= factor de transcripción de unión al elemento de respuesta sérico de c-fos
720	**ESTs: Hs.130741 AA608725
721	ADNc FLJ10976 fis de <i>Homo sapiens</i> , clon PLACE1001399 Hs. 296323 R36085:
722	FLJ20036: proteína hipotética FLJ20036 Hs.32922 N91145
723	C11ORF5: marco de lectura abierto 5 del cromosoma 11 Hs.121025 AA776702
724	AF3P21: proteína SH3 Hs.102929 N94372
725	LOC54104: proteína hipotética Hs.12871 H05934
726	DF: componente D del complemento (adipsina) Hs.155597 AA233549
727	CEP4: proteína efectora de Cdc42 4; grupo de unión de GTPasas Rho 4 Hs.3903 AA449061
728	KIF5B: miembro 5B de familia de cinesinas Hs.149436 AA644218
729	MGC5627: proteína hipotética MGC5627 Hs.237971 H02336
730	G3BP: proteína de unión a dominio SH3 de proteína de activación de Ras-GTPasa Hs.220689 AA449834
731	ESTs: Hs.293987 AA229758
732	ESTs: Hs.36828 AA194796
733	ARNm de <i>Homo sapiens</i> para proteína FLJ00101, cds parcial Hs.221600 W92262:
734	ADNc de <i>Homo sapiens</i> : FLJ21288 fis, clon COL01927 Hs.6019 R07184:
735	ESTs: débilmente similar a proteína relacionada con transcriptasa inversa 1207289A [<i>H. sapiens</i>] Hs.250594 H86813
736	ADNc FLJ11941 fis de <i>Homo sapiens</i> , clon HEMBB1000649 Hs.124106 AI301573:
737	ESTs: Hs.24908 H77726
738	TOB2: transductor de ERBB2, 2 Hs.4994 AA486088
739	ESTs: Hs.143900 AI193212
740	ARNm del clon FLC0675 PRO2870 de <i>Homo sapiens</i> , cds completa Hs.306117 H16589:
741	ESTs: débilmente similar a proteína KIAA0638 [<i>H. sapiens</i>] Hs.296288 T83657
742	FLJ20039: proteína hipotética FLJ20039 Hs.267448 AA448268

ES 2 991 436 T3

743	RPA2: proteína de replicación A2 (32 kD) Hs.79411 R13557
744	GAS1: específica de detención del crecimiento 1 Hs.65029 AA025819
745	Humano: secuencia de ADN del clon 967N21 en el cromosoma 201p12.3-13. Contiene el gen de CHGB para cromogranina B (secretogranina 1, SCG1), un pseudogén similar a parte de KIAA0172, el gen para una nueva proteína Hs.88959 R56678
746	ESTs: Hs.21175 AI341642
747	LBC: oncogén de crisis blástica linfóide Hs.301946 AA135716
748	ESTs: Hs.194595 R06761
749	MGC4707: proteína hipotética MGC4707 Hs.291003 R14653
750	ZNF183: proteína de dedos de zinc 183 (dedos de RING, tipo C3HC4) Hs.64794 AA132766
751	RAD18: proteína de reparación posterior a la replicación hRAD18p Hs.21320 R59197
752	EIF4EBP2: **proteína de unión a factor 4E de inicio de la traducción eucariota 2 Hs.278712 H15159
753	**ARNm de <i>Homo sapiens</i> ; ADNc DKFZp586M0723 (del clon DKFZp586M0723) Hs.27860 AA446650:
754	ORC3L: similar a complejo de reconocimiento de origen, subunidad 3 (homólogo de levaduras) Hs.74420 H99257
755	CDK7: cinasa dependiente de ciclinas 7 (homólogo de cinasa de activación de cdk MO15 de <i>Xenopus</i>) Hs.184298 AI311067
756	USP10: proteasa específica de ubiquitina 10 Hs.78829 AA455233
757	KIAA0733: proteína de unión a TAK1 2; proteína KIAA0733 Hs.109727 AA931658
758	R89286:
759	ALDH4: aldehído deshidrogenasa 4 (glutamato gamma-semialdehído deshidrogenasa; pirrolin-5-carboxilato deshidrogenasa) Hs.77448 AA181378
760	IDN3: proteína IDN3 Hs.225767 N62911
761	ESTs: Hs.50180 H48143
762	MIG2: inducible por mitógenos 2 Hs.75260 H29252
763	KIAA0856: proteína KIAA0856 Hs.13264 R12847
764	EST: Hs.47763 N54162
765	ARNm de <i>Homo sapiens</i> ; ADNc DKFZp547C244 (del clon DKFZp547C244) Hs.9460 AA447553:
766	KIAA0855: golgina-67 Hs.182982 AA775625
767	ESTs: débilmente similar a nucleolina JH0148 - rata [<i>R. norvegicus</i>] Hs.30120 R54659
768	FLJ22313: proteína hipotética FLJ22313 Hs.30211 H52061
769	ESTs: Hs.71818 AI028074
770	KIAA0618: producto génico de KIAA0618 Hs.295112 AA455506
771	ESTs: Hs.59413 W93056
772	ESTs: Hs.165607 AA992090
773	UBAP: proteína asociada a ubiquitina Hs.75425 AA446016
774	HAN11: proteína de repeticiones de WD Hs.176600 AA725641
775	USP16: proteasa específica de ubiquitina 16 Hs.99819 AA489619
776	ESTs: Hs.67776 AA464963
777	SM-20: similar a proteína de músculo liso de rata SM-20 Hs.6523 H56028
778	CCNG2: ciclina G2 Hs.79069 AA489647
779	ARNm de <i>Homo sapiens</i> ; ADNc DKFZp566P1124 (del clon DKFZp566P1124) Hs.321022 N62953:
780	FLJ20094: proteína hipotética FLJ20094 Hs.29700 N95490
781	LOC51174: delta-tubulina Hs.270847 W33133

ES 2 991 436 T3

782	ARNm de <i>Homo sapiens</i> ; ADNc DKFZp434I1820 (del clon DKFZp434I1820); cds parcial Hs.14235 N52394:
783	FANCA: anemia de Fanconi, grupo de complementación A Hs.284153 AA644129
784	P5-1: ORF de región de MHC de clase 1 Hs.1845 T58146
785	DNA2L: similar a DNA2 (helicasa de replicación de ADN, levadura, homólogo) Hs.194665 AA974495 KIAA0083
786	LOC51578: **proteína de glándula suprarrenal AD-004 Hs.279586 AA150301
787	ESTs: Hs.326417 AA913304
788	CDKN2D: inhibidor de cinasa dependiente de ciclinas 2D (p19, inhibe CDK4) Hs.29656 R77517 p19-INK4D=inhibidor D de cinasa dependiente de ciclinas 4
789	FABP1: proteína de unión a ácidos grasos 1, hígado Hs.5241 AA682392
790	TERA: proteína TERA Hs.180780 AA906997
791	ESTs: Hs.145383 AI253072
792	SLC7A5: familia de portadores de soluto 7 (transportador de aminoácidos catiónicos, sistema y+), miembro 5 Hs.184601 AA419176
793	AXL: tirosina cinasa receptora AXL Hs.83341 H15336 axl=ufo=tirosina cinasa receptora
794	LOC57190: selenoproteína N Hs.8518 AA284276
795	ESTs: Hs.99037 AA443948
796	STCH: proteína de estrés 70, chaperona, asociada a microsomas, 60 kD Hs.288799 H85311
797	ESTs: Hs.88523 AA278591 UG desconocido Hs.88523 ESTs
798	ESD: **esterasa D/formilglutación hidrolasa Hs.82193 AA250931
799	ESTs: Hs.122444 R31021
800	ESTs: Hs.283127 AI291262
801	KIAA0480: **producto génico de KIAA0480 Hs.92200 H91332
802	HP1-BP74: HP1-BP74 Hs.142442 AA598791
803	**ESTs: moderadamente similar a ALU1, ENTRADA DE ADVERTENCIA DE CONTAMINACIÓN DE SECUENCIA DE SUBFAMILIA J DE ALU HUMANA [<i>H. sapiens</i>] Hs.144662 AA987667
804	TTF2: factor de terminación de transcripción, ARN polimerasa II Hs.142157 AI023603
805	ESTs: Hs.13740 T70541
806	DJ37E16.5: proteína hipotética dJ37E16.5 Hs.5790 AA400021
807	CDH24: similar a cadherina 24 Hs.155912 AI732266
808	DJ465N24.2.1: **proteína hipotética dJ465N24.2.1 Hs.8084 AA932375
809	ESTs: débilmente similar a proteína S57447 HPBR11-7 [<i>H. sapiens</i>] Hs.16346 AA410490
810	ADNc de <i>Homo sapiens</i> : FLJ23285 fis, clon HEP09071 Hs.90424 AI005038:
811	KRAS2: homólogo de oncogén viral de sarcoma de rata de Kirsten v-Ki-ras2 2 Hs.184050 N95249
812	FLJ20038: proteína hipotética FLJ20038 Hs.72071 H96090
813	ESTs: débilmente similar a ALU4, ENTRADA DE ADVERTENCIA DE CONTAMINACIÓN DE SECUENCIA DE SUBFAMILIA SB2 DE ALU HUMANA [<i>H. sapiens</i>] Hs.28848 AA486607
814	H2AFN: familia de histonas H2A, miembro N Hs.134999 AI095013
815	RERE: repeticiones dipeptídicas de arginina-ácido glutámico (RE) Hs.194369 AA490249
816	USP1: proteasa específica de ubiquitina 1 Hs.35086 T55607
817	TIP47: proteína de selección de carga (proteína de unión a receptor de manosa-6-fosfato) Hs.140452 AA416787
818	KIAA0135: proteína KIAA0135 Hs.79337 AA427740 KIAA0135=relacionada con cinasa pim-1
819	ESTs: Hs.214410 T95273

ES 2 991 436 T3

820	PPP1R2: proteína fosfatasa 1, subunidad 2 reguladora (inhibidor) Hs.267819 N52605
821	ADNc de <i>Homo sapiens</i> : FLJ21210 fis, clon COL00479 Hs.325093 AA978323:
822	CSNK2A2: caseína cinasa 2, polipéptido principal alfa Hs.82201 AA054996
823	HSRTSBETA: proteína rTS beta Hs.180433 N66132
824	FLJ13110: proteína hipotética FLJ13110 Hs.7358 AA431233
825	ESTs: Hs.238797 N30704
826	FMN: oncogén de FYN relacionado con SRC, FGR, YES Hs.169370 N35086
827	RBM8A: proteína de motivos de unión a ARN 8A Hs.65648 AA448402
828	ESTs: Hs.21906 AA608546
829	ESTs: Hs.128081 AA971042
830	PP591: proteína hipotética PP591 Hs.118666 AA626336
831	N63866:
832	HM74: supuesto receptor de quimiocinas; proteína de unión a GTP Hs.137555 R02739
833	MID1: línea media 1 (síndrome de Opitz/BBB) Hs.27695 AA598640
834	KIAA1586: proteína KIAA1586 Hs.180663 AA938639
835	Secuencia de ARNm del clon CDABP0014 de <i>Homo sapiens</i> Hs.92679 AA443139:
836	HSU79274: proteína predicha por el clon 23733 Hs.150555 AA451900
837	AOC3: amina oxidasa, que contiene cobre 3 (proteína de adhesión vascular 1) Hs.198241 AA036974
838	AA548037:
839	FLJ10154: proteína hipotética FLJ10154 Hs.179972 AA457133
840	THBS1: trombospondina 1 Hs.87409 AA464532
841	DNAJB6: homólogo de DnaJ (Hsp40), subfamilia B, miembro 6 Hs.181195 AA431203
842	KIAA1547: proteína KIAA1547 Hs.31305 AI216623
843	GATA2: proteína de unión a GATA 2 Hs.760 R32405
844	ESTs: Hs.176950 R82522
845	KIAA1018: proteína KIAA1018 Hs.5400 AA156859
846	B4GALT1: **UDP-Gal:betaGlcNAc beta 1,4-galactosiltransferasa, polipéptido 1 Hs.198248 AA043795
847	HMGCR: 3-hidroxi-3-metilglutaril-coenzima A reductasa Hs.11899 AA779417
848	ESTs: débilmente similar a proteína 1819485A CENP-E [<i>H. sapiens</i>] Hs.167652 H94466
849	ESTs: Hs.294088 AA971073
850	KIAA1637: coactivador independiente de AF-2 (CIA); proteína KIAA1637 Hs.288140 AA918007
851	HSPC196: proteína hipotética Hs.239938 H66023
852	DR1: regulador por disminución de la transcripción 1, de unión a TBP (cofactor negativo 2) Hs.16697 AA132007
853	CG1I: supuesta proteína de interacción con ciclina G1 Hs.10028 AA486444
854	IGSF4: superfamilia de inmunoglobulinas, miembro 4 Hs.70337 AA487505
855	ESTs: Hs.179309 AA664350
856	HSPC163: proteína HSPC163 Hs.108854 AA053139
857	FLJ12788: proteína hipotética FLJ12788 Hs.20242 AI061317
858	FEM1B: homólogo b de FEM-1 (<i>C. elegans</i>) Hs.6048 H82273
859	FXR1: retraso mental asociado al cromosoma X frágil, homólogo autosómico 1 Hs.82712 N62761
860	NCOA3: coactivador del receptor nuclear 3 Hs.225977 AA156793
861	H2BFB: familia de histonas H2B, miembro B Hs.180779 N33927

ES 2 991 436 T3

862	ESTs: Hs.23830 AA460601
863	CDK7: cinasa dependiente de ciclinas 7 (homólogo de cinasa de activación de cdk MO15 de <i>Xenopus</i>) Hs.184298 AA031961 CAK=cdk7=NRTALRE=sdk=cinasa de activación de CDK
864	FLJ20259: proteína hipotética FLJ20259 Hs.9956 T55949
865	ADNc FLJ20678 fis de <i>Homo sapiens</i> , clon KAIA4163 Hs.143601 T95823:
866	RPS19: proteína ribosómica S19 Hs.298262 T72208
867	ARNm de <i>Homo sapiens</i> ; ADNc DKFZp434M0420 (del clon DKFZp434M0420) Hs.326048 AA443976:
868	TP53: proteína tumoral p53 (síndrome de Li-Fraumeni) Hs.1846 R39356 p53
869	FBI1: inductor del VIH-1 de proteína de unión a transcritos cortos Hs.104640 R06252
870	GOT1: transaminasa glutámica-oxaloacética 1, soluble (aspartato aminotransferasa 1) Hs.597 H22855
871	FLJ21434: proteína hipotética FLJ21434 Hs.298503 AA680129
872	DNMT2: ADN (citosina-5-)-metiltransferasa 2 Hs.97681 R95731
873	ESTs: Hs.55272 W02785
874	H2BFQ: familia de histonas H2B, miembro Q Hs.2178 AA010223
875	NFIC: factor nuclear I/C (factor de transcripción de unión a CCAAT) Hs.184771 N20996
876	NPTX1: pentraxina neuronal 1 Hs.84154 H22445
877	TLOC1: proteína de translocación 1 Hs.8146 AA450205
878	MGC5302: proteína residente en el retículo endoplásmico 58; proteína hipotética MGC5302 Hs.44970 N39195
879	ACTR2: homólogo de ARP2 (proteína relacionada con actina 2, levadura) Hs.42915 AA032090
880	AI287555:
881	ABCA7: casete de unión a ATP, subfamilia A (ABC1), miembro 7 Hs.134514 AI668632
882	COL7A1: colágeno, tipo VII, alfa I (epidermólisis ampollosa, distrófica, dominante y recesiva) Hs.1640 AA598507
883	RFC2: factor de replicación C (activador 1) 2 (40 kD) Hs.139226 AA663472
884	FLJ22583: proteína hipotética FLJ22583 Hs.287700 AA135836
885	**ESTs: débilmente similar a ORF2 [<i>M. musculus</i>] Hs.172208 AI820570
886	ESTs: Hs.21667 R15709
887	RBBP4: proteína de unión a retinoblastoma 4 Hs.16003 AA705035
888	ARNm de <i>Homo sapiens</i> ; ADNc DKFZp434J1027 (del clon DKFZp434J1027); cds parcial Hs.22908 R20166:
889	ESTs: Hs.166539 AI080987
890	NKTR: secuencia de reconocimiento de tumor de linfocitos citolíticos naturales Hs.241493 AA279666 proteína de reconocimiento de tumor de NK=proteína relacionada con ciclofilina
891	MUC1: mucina 1, transmembrana Hs.89603 AA486365
892	AP4B1: complejo 4 de proteína relacionada con 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.183299 AA286914 UG desconocido Hs.183299 ESTs sc_id2032
896	BAG3: atanogén asociado a BCL2 3 Hs.15259 AI269958
897	INSR: receptor de insulina Hs.89695 AA001106
898	TRIP: proteína de interacción con TRAF Hs.21254 AA186426
899	EST: Hs.307975 R22182
900	**ADNc de <i>Homo sapiens</i> : FLJ23037 fis, clon LNG02036, altamente similar a ARNm de homólogo de proteína mad de <i>Homo sapiens</i> (hMAD-3) HSU68019 Hs.288261 W72201:

901	HLA-ADN: complejo principal de histocompatibilidad, clase II, DN alfa Hs.11135 AA702254 complejo principal de histocompatibilidad, clase II, DN alfa
902	FLJ10392: **proteína hipotética FLJ10392 Hs.20887 AI261305
903	MPHOSPH1: **fosfoproteína de fase M 1 Hs.240 N63752
904	STAG1: antígeno estromal 1 Hs.286148 R36160
905	USP1: proteasa específica de ubiquitina 1 Hs.35086 AA970066
906	ESTs: moderadamente similar a ALU4, ENTRADA DE ADVERTENCIA DE CONTAMINACIÓN DE SECUENCIA DE SUBFAMILIA SB2 DE ALU HUMANA [<i>H. sapiens</i>] Hs.181315 AA448251
907	PA26: proteína nuclear PA26 regulada por p53 Hs.14125 AA447661
908	ESTs: débilmente similar a proteína de dedos de zinc [<i>H. sapiens</i>] Hs.71243 N92478
909	SH3PX1: proteína que contiene dominios SH3 y PX SH3PX1 Hs.7905 R69163
910	**ADNc de <i>Homo sapiens</i> : FLJ22554 fis, clon HSI01092 Hs.93842 H58317:
911	RPS25: proteína ribosómica S25 Hs.113029 AA779404
912	ESTs: débilmente similar a región V-I de cadena kappa de Ig A49134 [<i>H. sapiens</i>] Hs.5890 N34799 fra-2=antígeno relacionado con fos 2
913	TXNRD1: tiorredoxina reductasa 1 Hs.13046 AA453335 tiorredoxina reductasa
914	**ESTs: Hs.184378 N77828
915	GCSH: proteína de sistema de escisión de glicina H (portador de aminometilo) Hs.77631 R71327
916	ADNc FLJ11904 fis de <i>Homo sapiens</i> , clon HEMBB1000048 Hs.285519 AA447098:
917	NCOA3: coactivador del receptor nuclear 3 Hs.225977 H51992 AIB1=amplificado en cáncer de mama=TRAM-1=RAC3=ACTR=CAGH16=nucl
918	FLJ20159: proteína hipotética FLJ20159 Hs.288809 R33122
919	IL7R: receptor de interleucina 7 Hs.237868 AA487121
920	RAB23: RAB23, miembro de familia de oncogenes RAS Hs.94769 AA134569
921	ESTs: Hs.132493 AA923168
922	ESTs: Hs.87507 AA236015
923	SHC1: proteína transformante de SHC (que contiene dominio de homología de Src 2) 1 Hs.81972 R52960
924	KIAA1321: proteína KIAA1321 Hs.24336 W37999
925	GLI: homólogo de oncogén asociado a glioma (proteína de dedos de zinc) Hs.2693 AI373071
926	ESTs: Hs.183299 AA291137 UG desconocido Hs.183299 ESTs sc_id2032
927	GPRK6: cinasa receptora acoplada a proteína G 6 Hs.76297 AA291284
928	ESTs: Hs.93704 AA702684
929	CAPS: calcifosina Hs.26685 AA858390
930	ADNc FLJ10976 fis de <i>Homo sapiens</i> , clon PLACE1001399 Hs.296323 R27711:
931	C6: componente 6 del complemento Hs.1282 N59396
932	UBE2D3: enzima de conjugación con ubiquitina E2D 3 (homóloga a UBC4/5 de levadura) Hs.118797 AA465196
933	DDX8: polipéptido de 8 caja DEAD/H (Asp-Glu-Ala-Asp/His) (ARN helicasa) Hs.171872 AA465387 ARN helicasa (HRH1)
934	DKFZP434B168: proteína DKFZP434B168 Hs.48604 N62684
935	FLJ10512: proteína hipotética FLJ10512 Hs.93581 T39933
936	ARNm de <i>Homo sapiens</i> ; ADNc DKFZp564F093 (del clon DKFZp564F093) Hs.18724 W87709:
937	F8A: asociado a factor coagulación VIII (transcrito intrónico) Hs.83363 AA463924
938	HSU53209: transformador-2 alfa (htra-2 alfa) Hs.24937 AA465172

939	UBQLN2: ubiquilina 2 Hs.4552 R43580
940	EIF2C2: factor de inicio de traducción eucariota 2C, 2 Hs.193053 N93082
941	ARNm de <i>Homo sapiens</i> para proteína FLJ00012, cds parcial Hs.21051 H17645:
942	KIAA0841: proteína KIAA0841 Hs.7426 R20299
943	KCNAB2: subfamilia de canales de potasio dependientes de voltaje, relacionados con agitador, miembro beta 2 Hs.298184 H14383
944	KIAA1637: coactivador independiente de AF-2 (CIA); proteína KIAA1637 Hs.288140 AA521358
945	ESTs: Hs.27379 H17455
946	FLJ11323: proteína hipotética FLJ11323 Hs.25625 R49707
947	SSP29: proteína ácida rica en leucinas Hs.84264 AA489201
948	ESTs: Hs.69280 AA486011
949	ADAMTS1: similar a desintegrina y metaloproteasa (tipo repolisina) con motivo de trombospondina tipo 1, 1 Hs.8230 AA057170
950	ESTs: Hs.43466 N23889
951	MLLT4: leucemia mieloide/linfoide o de linaje mixto (homólogo de Trithorax (<i>Drosophila</i>)); translocada, 4 Hs.100469 AA010818
952	ESTs: Hs.271034 AA406581
953	LMNB1: lámina B1 Hs.89497 AA983462
954	ADNc FLJ13547 fis de <i>Homo sapiens</i> , clon PLACE1007053 Hs.7984 AA629264:
955	PTMS: paratimosina Hs.171814 R10451
956	H2AFL: familia de histonas H2A, miembro 1 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: asociado a insulinooma 1 Hs.89584 R38640
961	FREQ: homólogo de frecuenina (<i>Drosophila</i>) Hs.301760 H16821
962	LOC58486: proteína similar a transposasa Buster1 derivada de transposones Hs.25726 AA868020
963	SMARCD1: regulador de la cromatina dependiente de actina, relacionado con SWI/SNF, asociado a matriz, subfamilia d, miembro 1 Hs.79335 H91691
964	ESTs: Hs.242998 T96522
965	INADL: proteína de dominio PDZ (similar a inaD de <i>Drosophila</i>) Hs.321197 AA005153
966	ESTs: débilmente similar a supuesta p150 [<i>H. sapiens</i>] Hs.37751 AA436174
967	MGC5338: proteína hipotética MGC5338 Hs.99598 H50550
968	W85890:
969	NUCKS: similar a caseína cinasa ubicua nuclear de rata 2 Hs.118064 AI053436
970	Secuencia de ARNm del clon 25110 de <i>Homo sapiens</i> Hs.27262 H18031:
971	AI333214:
972	GAS41: secuencia amplificada en glioma 41 Hs.4029 T62072
973	LOC51170: deshidrogenasa/reductasa de cadena corta retiniana retSDR2 Hs.12150 N79745
974	H2BFG: **familia de histonas 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 <i>Homo sapiens</i> ; ADNc DKFZp434A1014 (del clon DKFZp434A1014); cds parcial Hs.278531 H00596:

ES 2 991 436 T3

978	PPP2CA: proteína fosfatasa 2 (anteriormente 2A), subunidad catalítica, isoforma alfa Hs.91773 AA599092
979	ESTs: débilmente similar a producto de proteína sin nombre [<i>H. sapiens</i>] Hs.118338 W85843
980	ADNc FLJ11643 fis de <i>Homo sapiens</i> , clon HEMBA1004366 Hs.111496 AA598803:
981	ESTs: moderadamente similar a ALUE, !!!! ENTRADA DE ADVERTENCIA DE CLASE E DE ALU HUMANA!!!! [<i>H. sapiens</i>] Hs.125407 AA878944
982	ESTs: moderadamente similar a ALU1, ENTRADA DE ADVERTENCIA DE CONTAMINACIÓN DE SECUENCIA DE SUBFAMILIA J DE ALU HUMANA [<i>H. sapiens</i>] Hs.144662 AI191290
983	KIAA0916: proteína KIAA0916 Hs.151411 R91388
984	CDC25A: ciclo de división celular 25A Hs.1634 R09062
985	PRIM2A: primasa, polipéptido 2A (58 kD) Hs.74519 R61073
986	DSP: desmoplaquina (DPI, DPII) Hs.74316 H90899
987	KIAA0101: producto génico de KIAA0101 Hs.81892 W68219
988	ESTs: débilmente similar a supuesta p150 [<i>H. sapiens</i>] Hs.268026 AA411454
989	ESTs: Hs.18140 T97707
990	H2AFL: familia de histonas H2A, miembro L Hs.28777 AA457566
991	ARNm de <i>Homo sapiens</i> para proteína KIAA1700, cds parcial Hs.20281 H00287:
992	STAG3: antígeno estromal 3 Hs.20132 AA453028
993	ZNF207: proteína de dedos de zinc 207 Hs.62112 N59119
994	BMP6: proteína morfogenética ósea 6 Hs.285671 AA424833
995	ESTs: moderadamente similar a sertolina [<i>R. norvegicus</i>] Hs.91192 H60690
996	LOC51064: **homólogo de subunidad 13 de glutatión S-transferasa Hs.279952 W88497
997	NUCKS: similar a caseína cinasa ubicua nuclear de rata 2 Hs.118064 AA927182
998	ESTs: débilmente similar a T00370, proteína hipotética KIAA0659 [<i>H. sapiens</i>] Hs.131899 W93155
999	FLJ13057: proteína hipotética FLJ13057 similar a GCL (<i>germ cell-less</i>) Hs.243122 R23254
1000	ESTs: Hs.144796 AI219737
1001	FLJ10511: proteína hipotética FLJ10511 Hs.106768 R25877
1002	DKFZP564A122: proteína DKFZP564A122 Hs.187991 N31577
1003	ODF2: fibra densa exterior de colas de espermatozoides 2 Hs.129055 AA400407
1004	AMY2A: amilasa, alfa 2A; pancreática Hs.278399 R64129
1005	**ESTs: débilmente similar a placofilina 2b [<i>H. sapiens</i>] Hs.12705 N91589
1006	CYP1B1: citocromo P450, subfamilia I (inducible por dioxina), polipéptido 1 (glaucoma 3, infantil primario) Hs.154654 AA029776
1007	CAPN7: calpaína 7 Hs.7145 N46420
1008	FLJ20069: proteína hipotética FLJ20069 Hs.273294 AA229966
1009	FLJ10618: proteína hipotética FLJ10618 Hs.42484 AA478847
1010	KIAA1637: **coactivador independiente de AF-2 (CIA); proteína KIAA1637 Hs.288140 AA452531
1011	FLJ20004: **proteína hipotética FLJ20004 Hs.17311 AA487895
1012	FLJ12892: proteína hipotética FLJ12892 Hs.17731 AA670363
1013	PLU-1: supuesto motivo de unión a ADN/cromatina Hs.143323 AA464869
1014	**ESTs: Hs.36828 AA418448
1015	KIAA0586: producto génico de KIAA0586 Hs.77724 AA905278
1016	MTHFD2: metilen-tetrahidrofolato deshidrogenasa (dependiente de NAD+), meteniltetrahidrofolato ciclohidrolasa Hs.154672 AA480994

ES 2 991 436 T3

1017	BRF1: **factor de respuesta a butirato 1 (factor de a respuesta EGF 1) Hs.85155 AA424743
1018	TFAP2A: factor de transcripción AP-2 alfa (proteína de unión a potenciador de activación 2 alfa) Hs.18387 R38044
1019	VIL2: vilina 2 (ezrina) Hs.155191 AA411440
1020	SDC1: sindecano 1 Hs.82109 AA074511
1021	RNTRE: relacionado con el extremo N-terminal de tre Hs.278526 AA281057
1022	HSPC207: proteína hipotética Hs.75798 H99997
1023	FLJ22376: proteína hipotética FLJ22376 Hs.29341 AI199155
1024	RNF10: proteína de dedos de Ring 10 Hs.5094 H73586
1025	PNN: pinina, proteína asociada a desmosomas Hs.44499 AA707321
1026	FLJ20516: proteína hipotética FLJ20516 Hs.70811 AA122393
1027	RPL13A: proteína ribosómica L13a Hs.119122 AI254200
1028	H2BFB: familia de histonas H2B, miembro B Hs.180779 AA885642
1029	OGT: N-acetilglucosamina (GlcNAc) transferasa unida a O (UDP-N-acetilglucosamina:polipéptido-N-acetilglucosaminiltransferasa) Hs.100293 R13317
1030	KIAA0155: producto génico de KIAA0155 Hs.173288 AA133684
1031	ILF2: factor de unión a potenciador de interleucina 2, 45 kD Hs.75117 H95638
1032	ARNm de <i>Homo sapiens</i> ; ADNc DKFZp586I1518 (del clon DKFZp586I1518) Hs.21739 AA287917:
1033	PKNOX1: homeosecuencia 1 de PBX/knotted 1 Hs.158225 AI350546
1034	KMO: **quinurenina 3-monooxigenasa (quinurenina 3-hidroxilasa) Hs.107318 AA044326
1035	VCAM1: molécula de adhesión celular vascular 1 Hs.109225 H16591 CD106=VCAM-1
1036	N54811:
1037	KIAA0618: producto génico de KIAA0618 Hs.295112 H81940
1038	MAFG: familia de oncogenes de fibrosarcoma musculoaponeurótico v-maf (aviar), proteína G Hs.252229 N21609 MafG=factor de transcripción de cremallera de leucinas básica
1039	MATN2: matrilina 2 Hs.19368 AA071473
1040	HOXB4: homeosecuencia B4 Hs.126666 AA918749
1041	FLJ10466: proteína hipotética FLJ10466 Hs.121073 AA453607
1042	FLJ22557: proteína hipotética FLJ22557 Hs.106101 AA127879
1043	EST: Hs.149260 AI247680
1044	KIAA0677: producto génico de KIAA0677 Hs.155983 AA026751
1045	EST: Hs.104123 AA197344
1046	UCP4: proteína de desacoplamiento 4 Hs.40510 H94680
1047	EST: Hs.144224 N93807
1048	GATA2: proteína de unión a GATA 2 Hs.760 H00625 proteína de unión a GATA 2
1049	ESTs: Hs.14743 H61082
1050	EST: Hs.116174 AA626786
1051	ITGB3: integrina, beta 3 (glicoproteína de plaquetas IIIa, antígeno CD61) Hs.87149 AA666269
1052	FLJ23399: proteína hipotética FLJ23399 Hs.299883 R19895
1053	ESTs: Hs.21734 N72976
1054	FLJ20425: proteína hipotética FLJ20425 Hs.71040 AA424566
1055	CUL4A: culina 4A Hs.183874 AA598836
1056	PTP4A1: proteína tirosina fosfatasa de tipo IVA, miembro 1 Hs.227777 R61007 proteína tirosina fosfatasa PTPCAAX1 (hPTPCAAX1)

ES 2 991 436 T3

1057	ESTs: Hs.7913 N35592
1058	GRO1: oncogén de GRO1 (actividad estimulante de crecimiento de melanoma, alfa) Hs.789 W46900
1059	ESTs: moderadamente similar a convertasa NRD2 [<i>H. sapiens</i>] Hs.309734 H78796
1060	FLJ10826: proteína hipotética FLJ10826 Hs.24809 AA486738
1061	TOM34: translocasa de membrana mitocondrial externa 34 Hs.76927 AA457118
1062	H2AFL: familia de histonas H2A, miembro L Hs.28777 AA452933
1063	D10S170: **segmento de ADN, copia única, sonda pH4 (secuencia transformante, tiroide-1, Hs.315591 N35493
1064	SCYA2: citocina inducible pequeña A2 (proteína quimiotáctica de monocitos 1, homóloga a Sig-je de ratón) Hs.303649 T77816 MCP-1=MCAF=citocina inducible pequeña A2=JE=quimiocina
1065	FLJ10688: proteína hipotética FLJ10688 Hs.118793 AA465358
1066	PTD017: proteína PTD017 Hs.274417 AA160498
1067	KIAA0026: gen relacionado con MORF X Hs.173714 AA676604
1068	BMP2: proteína morfogenética ósea 2 Hs.73853 AA489383
1069	MNT: proteína de unión a MAX Hs.25497 AA455508
1070	KIAA1170: proteína KIAA1170 Hs.268044 H80507
1071	CRYBA1: cristalina, beta A1 Hs.46275 AA487614
1072	KATNA1: subunidad A de katanina p60 (que contiene ATPasa) 1 Hs.289099 AA609740
1073	ADNc FLJ20796 fis de <i>Homo sapiens</i> , clon COL00301 Hs.113994 N53458:
1074	CEP4: proteína efectora de Cdc42 4; grupo de unión de GTPasas Rho 4 Hs.3903 W32509
1075	ESTs: Hs.117261 AA682521
1076	CYP1B1: citocromo P450, subfamilia I (inducible por dioxina), polipéptido 1 (glaucoma 3, infantil primario) Hs.154654 AA040872
1077	ALTE: elemento de transposón similar a Ac Hs.9933 AA630498
1078	RAD51: homólogo de RAD51 (<i>S. cerevisiae</i>) (homólogo de RecA de <i>E coli</i>) Hs.23044 AA873056
1079	MAN1A2: manosidasa, alfa, clase 1A, miembro 2 Hs.239114 R78501
1080	H53763.
1081	MET: proto-oncogén de met (receptor de factor de crecimiento de hepatocitos) Hs.285754 AA410591
1082	DYRK1A: cinasa regulada por fosforilación de tirosina (Y) de doble especificidad 1A Hs.75842 AA676749
1083	ARHGAP8: **proteína de activación de GTPasa Rho 8 Hs.102336 AA037410
1084	LMO4: único dominio de LIM 4 Hs.3844 H27986
1085	ADCY6: adenilato ciclase 6 Hs.12373 AA148044
1086	EST: Hs.135448 AI078552
1087	NCOA3: coactivador del receptor nuclear 3 Hs.225977 W46433
1088	DNAJB4: homólogo de DnaJ (Hsp40), subfamilia B, miembro 4 Hs.41693 AA081471
1089	NAB1: proteína de unión a NGFI-A 1 (proteína de unión a ERG1 1) Hs.107474 AA486027
1090	ESTs: débilmente similar a proteína hipotética T08663 DKFZp547G0910.1 [<i>H. sapiens</i>] Hs.172084 N63646
1091	KIAA0735: producto génico de KIAA0735; homólogo 2B de proteína de vesículas sinápticas Hs.8071 R56082
1092	GNB1: proteína de unión a nucleótidos de guanina (proteína G), polipéptido beta 1 Hs.215595 AA487912
1093	ARNm de <i>Homo sapiens</i> para proteína KIAA1716, cds parcial Hs.21446 R49763:
1094	CINESINA: CADENA PESADA

ES 2 991 436 T3

1095	CCND1: ciclina D1 (PRAD1: adenomatosis paratiroidea 1) Hs.82932 AA487486 ciclina D1=BCL1=PRAD1=translocada en leucemia de células del manto
1096	ESTs: Hs.106129 R56716
1097	AA431931.
1098	PSEN1: presenilina 1 (enfermedad de Alzheimer 3) Hs.3260 AA403083
1099	ESTs: Hs.193804 AA010918
1100	DKFZp762P2111: proteína hipotética DKFZp762P2111 Hs.14217 AA429586
1101	KIAA1350: proteína KIAA1350 Hs.101799 W37627
1102	FLJ20847: proteína hipotética FLJ20847 Hs.13479 H16996
1103	HDCMA18P: proteína HDCMA18P Hs.278635 N64387
1104	FLJ12890: proteína hipotética FLJ12890 Hs.43299 N62475
1105	ESTs: Hs.127453 AA973625
1106	BAIAP2: proteína asociada a BAI1 2 Hs.7936 R60328
1107	ESTs: Hs.317584 AA191424
1108	DKFZP434J046: proteína DKFZP434J046 Hs.116244 AI024401
1109	ESTs: Hs.114055 AA701352
1110	ESTs: Hs.44380 N93122
1111	ESTs: Hs.20142 AA625570
1112	UBL3: similar a ubiquitina 3 Hs.173091 T82438
1113	H2AFL: familia de histonas H2A, miembro L Hs.28777 N50797
1114	SUCLG2: **succinato-CoA ligasa, de formación de GDP, subunidad beta Hs.247309 N68557
1115	ZWINT: interactor de ZW10 Hs.42650 AA706968
1116	FLJ10583: proteína hipotética FLJ10583 Hs.105633 R00425
1117	FLJ20552: proteína hipotética FLJ20552 Hs.69554 AA463982
1118	FADD: asociado a Fas (TNFRSF6) a través de dominio de muerte Hs.86131 AA430751 FADD=MORT
1119	SFRS7: factor de corte y empalme, rico en arginina/serina 7 (35 kD) Hs.184167 AA418813
1120	RAD54L: similar a RAD54 (<i>S. cerevisiae</i>) Hs.66718 AI372035
1121	MYLE: proteína MYLE Hs.11902 T68845
1122	LOC51334: proteína de células madre mesenquimatosas DSC54 Hs.157461 R63841
1123	PRIM2A: primasa, polipéptido 2A (58 kD) Hs.74519 AA434404
1124	KIAA0056: proteína KIAA0056 Hs.13421 AA430545
1125	ESTs: moderadamente similar a ALU7, ENTRADA DE ADVERTENCIA DE CONTAMINACIÓN DE SECUENCIA DE SUBFAMILIA SQ DE ALU HUMANA [<i>H. sapiens</i>] Hs.82590 N53024
1126	ESTs: Hs.17269 AA705050
1127	NSAP1: proteína asociada a NS1 1 Hs.155489 AA186327
1128	CEACAM5: molécula de adhesión celular relacionada con antígeno carcinoembrionario 5 Hs.220529 AA130547
1129	FLJ11021: proteína hipotética FLJ11021 similar a factor de corte y empalme, rico en arginina/serina 4 Hs.81648 AA291183 UG desconocido Hs.202583 ESTs, débilmente similar a rico en arginina
1130	FOSL1: antígeno similar a FOS 1 Hs.283565 T82817 fra-1=antígeno relacionado con fos 1
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 proteína transportadora
1134	PCAF: factor asociado a p300/CBP Hs.199061 N74637 P/CAF=factor asociado a p300/CBP

ES 2 991 436 T3

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 FLJ11904 fis de <i>Homo sapiens</i> , clon HEMBB1000048 Hs.285519 N74617:
1141	AA928536:
1142	SQSTM1: **secuestrosoma 1 Hs.182248 AA931964
1143	**ADNc FLJ13700 fis de <i>Homo sapiens</i> , clon PLACE2000216, altamente similar a CADENA BETA DE ESPECTRINA, CEREBRO Hs.324648 AA018591:
1144	SLC22A3: familia 22 de portador de soluto 22 (transportador de monoaminas extraneuronales), miembro 3 Hs.81086 AA460012
1145	FLJ22557: proteína hipotética FLJ22557 Hs.106101 H00595
1146	FLJ20539: proteína hipotética FLJ20539 Hs.118552 R36152
1147	AA991624:
1148	TRAP150: proteína asociada a receptor de hormona tiroidea, subunidad de 150 kDa Hs.108319 W85832
1149	ESTs: Hs.221847 R91557
1150	TCFL1: similar a factor de transcripción 1 Hs.2430 AA443950
1151	ESTs: altamente similar a variante 1 de corte y empalme de oxitocinasa [<i>H. sapiens</i>] Hs.203271 AA487918
1152	PLAB: factor de diferenciación de próstata Hs.296638 AA450062
1153	RBM14: proteína de motivos de unión a ARN 14 Hs.1170 AA417283
1154	EGFL5: dominio similar a EGF, múltiple 5 Hs.5599 W67981
1155	H2AFO: familia de histonas H2A, miembro O Hs.795 AA047260
1156	ESTs: débilmente similar a leucotrieno B4 omega-hidroxilasa A46661 [<i>H. sapiens</i>] Hs.169001 N45556
1157	W78784:
1158	TOP3A: topoisomerasa (ADN) III alfa Hs.91175 N21546
1159	W73732: factor de células huésped 1=proteína de interacción con transactivador de VP16
1160	CYP1B1: citocromo P450, subfamilia I (inducible por dioxina), polipéptido 1 (glaucoma 3, infantil primario) Hs.154654 AA448157 citocromo P450 IB1 (inducible por dioxina)
1161	ESTs: Hs.135276 AI092102
1162	RHEB2: homólogo de Ras enriquecido en cerebro 2 Hs.279903 AA482117
1163	ESTs: altamente similar a EF-9 [<i>M. musculus</i>] Hs.8366 H94467
1164	POLA: polimerasa (dirigida por ADN), alfa Hs.267289 AA707650
1165	KIAA1008: proteína KIAA1008 Hs.323346 AA863115
1166	PIK3CD: fosfoinositida-3-cinasa, catalítica, polipéptido delta Hs.162808 AA281652
1167	T53625:
1168	**ARNm de <i>Homo sapiens</i> ; ADNc DKFZp434A1114 (del clon DKFZp434A1114) Hs.326292 AA417274:
1169	ESTs: Hs.26744 H16988
1170	FLJ13912: proteína hipotética FLJ13912 Hs.47125 W74133
1171	ARNm de <i>Homo sapiens</i> ; ADNc DKFZp762B195 (del clon DKFZp762B195) Hs.284158 AA625574:
1172	SSA2: antígeno de síndrome de Sjögren A2 (60 kD, autoantígeno de ribonucleoproteína SS-A/Ro) Hs.554 AA010351
1173	BK1048E9.5: proteína hipotética bK1048E9.5 Hs.6657 N68512

ES 2 991 436 T3

1174	TOP1: topoisomerasa (ADN) I Hs.317 AA232856 topoisomerasa I
1175	ESTs: Hs.15386 H18472
1176	KPNB1: carioferina (importina) beta 1 Hs.180446 AA121732
1177	MGC861: proteína hipotética MGC861 Hs.208912 N69694
1178	PMS2L8: **proteína 8 similar a proteína 2 de segregación posmeiótica aumentada Hs.323954 T62577
1179	TSC22: **proteína estimulada por factor de crecimiento transformante beta TSC-22 Hs.114360 R16390
1180	C8ORF1: marco de lectura abierto 1 del cromosoma 8 Hs.40539 AA278836
1181	ESTs: Hs.129165 AA989211
1182	DMTF: factor de transcripción similar a Myb de unión a ciclina D 1 Hs.5671 AA129860
1183	CDC7L1: similar a CDC7 (ciclo de división celular 7, <i>S. cerevisiae</i> , homólogo) 1 Hs.28853 N62245 cinasa relacionada con Cdc7
1184	LOC51700: citocromo b5 reductasa b5R.2 Hs.22142 AA425316
1185	FLNA: filamina A, alfa (proteína de unión a actina 280) Hs.195464 AA598978
1186	FLJ20257: proteína hipotética FLJ20257 Hs.178011 H78675
1187	ADNc FLJ13604 fis de <i>Homo sapiens</i> , clon PLACE1010401 Hs.23193 AA406599:
1188	ESTs: Hs.205227 R73480
1189	SCYB14: subfamilia B de citocinas inducibles pequeñas (Cys-X-Cys), miembro 14 (BRAK) Hs.24395 AA953842
1190	MAPK8IP2: **proteína 2 de interacción con proteína cinasa activada por mitógenos 8 Hs.80545 AA418293
1191	ZNF42: proteína de dedos de zinc 42 (sensible a ácido retinoico específico de mieloides) Hs.169832 AA932642
1192	ESTs: Hs.127054 AA862450
1193	NUDT4: motivo 4 de tipo Nudix (resto X unido a nucleósido difosfato) Hs.92381 AA425630
1194	ADNc FLJ10632 fis de <i>Homo sapiens</i> , clon NT2RP2005637 Hs.202596 H82421:
1195	LOC51042: proteína de dedos de zinc Hs.102419 AA033532
1196	NUMA1: proteína de aparato mitótico nuclear 1 Hs.301512 AA679293
1197	ESTs: altamente similar a proteína I relacionada con kappa B A56429 [<i>H. sapiens</i>] Hs.144614 AA293771
1198	ESTs: Hs.127703 AA947258
1199	ADNc FLJ14214 fis de <i>Homo sapiens</i> , clon NT2RP3003576 Hs.321236 AA903913:
1200	NFKBIA: factor nuclear de potenciador de genes de polipéptido ligero kappa en inhibidor de células B, alfa Hs.81328 W55872 Ikb alfa
1201	ESTs: Hs.120029 AA707598
1202	ESTs: moderadamente similar a cadena A, glucosamina-6-fosfato desaminasa isomerasa humana a 1,75 A [<i>H. sapiens</i>] Hs.21398 AA172012
1203	NFIA: factor nuclear I/A Hs.173933 AI912047
1204	RECQL4: similar a proteína RecQ 4 Hs.31442 AA620446
1205	*ESTs: débilmente similar a ALU1, ENTRADA DE ADVERTENCIA DE CONTAMINACIÓN DE SECUENCIA DE SUBFAMILIA J DE ALU HUMANA [<i>H. sapiens</i>] Hs.318894 R96212
1206	ADNc de <i>Homo sapiens</i> : FLJ21686 fis, clon COL09379 Hs.20787 R11371:
1207	LOC57168: similar a aspartato beta hidroxilasa (ASPH) Hs.184390 H17272
1208	ESTs: Hs.26096 R54109
1209	ARNm de proteína relacionada con OSBP de <i>Homo sapiens</i> 6, cds completa Hs.318775 AA680281:
1210	APACD: proteína de unión a ATP asociada con la diferenciación celular Hs.153884 N80741

ES 2 991 436 T3

1211	VIM: **vimentina Hs.297753 AI668662
1212	ADNc FLJ13618 fis de <i>Homo sapiens</i> , clon PLACE1010925 Hs.17448 AA427980:
1213	NR3C1: subfamilia 3 de receptor nuclear, grupo C, miembro 1 Hs.75772 N30428 receptor de glucocorticoides
1214	ADNc de <i>Homo sapiens</i> : FLJ21814 fis, clon HEP01068 Hs.289008 R12808:
1215	BRD7: proteína que contiene bromodominio 7 Hs.279762 AA488428
1216	MAP3K8: **proteína cinasa activada por mitógenos 8 Hs.248 W42450
1217	ESTs: Hs.23213 H29336
1218	ESTs: Hs.122444 AA939019
1219	TUSP: proteína de superfamilia de Tubby Hs.102237 H78234
1220	KIAA1117: proteína KIAA1117 Hs.278398 H01516
1221	Humano: ARNm del clon 137308, cds parcial Hs.322149 H91303
1222	ESTs: Hs.130214 AA456631
1223	RAB3A: RAB3A, miembro de la familia de oncogenes de RAS Hs.27744 H14230
1224	AA598795: proteína fosfatasa 2 (anteriormente 2A), subunidad reguladora B (P
1225	H2BFC: familia de histonas H2B, miembro C Hs.137594 AI340654
1226	CFLAR: regulador de apoptosis similar a CASP8 y FADD Hs.195175 N94588
1227	CD24: antígeno CD24 (antígeno de agrupación 4 de carcinoma de pulmón de células pequeñas) Hs.286124 H59915
1228	EST: Hs.48532 N62402
1229	CCRK: cinasa relacionada con el ciclo celular Hs.26322 H17616
1230	HECH: proteína similar a heterocromatina 1 Hs.278554 AI139106
1231	DKFZp547O146: proteína hipotética DKFZp547O146 Hs.91246 T80848
1232	ESTs: Hs.71574 AA135328
1233	HLXB9: homeosecuencia HB9 Hs.37035 AI459915
1234	AA600222:
1235	SPINK5: inhibidor de serina proteasa, tipo Kazal, 5 Hs.5476 W92134
1236	RNUT1: ARN, transportador de U 1 Hs.21577 AA447799
1237	ADNc de <i>Homo sapiens</i> : FLJ23013 fis, clon LNG00740 Hs.13075 AA464543:
1238	KIAA0063: producto génico de KIAA0063 Hs.3094 T82263
1239	DYRK2: cinasa regulada por fosforilación de tirosina (Y) de doble especificidad 2 Hs.173135 R63622
1240	R94947:
1241	ADNc FLJ14337 fis de <i>Homo sapiens</i> , clon PLACE4000494 Hs.180187 AA004903:
1242	FLJ20624: proteína hipotética FLJ20624 Hs.52256 AA431909
1243	ESTs: Hs.43838 R38261
1244	FLJ23053: proteína hipotética FLJ23053 Hs.94037 R25654
1245	MGC11266: proteína hipotética MGC11266 Hs.293943 AA400456
1246	ESTs: moderadamente similar a ALU8, ENTRADA DE ADVERTENCIA DE CONTAMINACIÓN DE SECUENCIA DE SUBFAMILIA SX DE ALU HUMANA [<i>H. sapiens</i>] Hs.34174 AA126603
1247	PLAUR: activador de plasminógeno, receptor de urocinasa Hs.179657 AA147962
1248	TSG101: gen de susceptibilidad tumoral 101 Hs.118910 AA670215
1249	HCNGP: proteína reguladora transcripcional Hs.27299 AA457232
1250	KIAA0978: proteína KIAA0978 Hs.3686 AA857017
1251	ESTs: Hs.61708 AA033867

1252	ESTs: Hs.120734 AA827482
1253	ESTs: Hs.5909 AA972654
1254	CDH24: similar a cadherina 24 Hs.155912 AI668564
1255	CCND1: ciclina D1 (PRAD1: adenomatosis paratiroidea 1) Hs.82932 T77237
1256	ESTs: Hs.43148 AA284775
1257	ESTs: Hs.222566 T50982
1258	ESTs: Hs.194125 N52822
1259	EST: Hs.154621 AI138644
1260	MAN1A2: manosidasa, alfa, clase 1A, miembro 2 Hs.239114 R22905
1261	MAN2A2: manosidasa, alfa, clase 2A, miembro 2 Hs.295605 AA454175
1262	Secuencia de ADN humano del clon 967N21 en el cromosoma 20p12.3-13. Contiene el gen de CHGB para cromogranina B (secretogranina 1, SCG1), un pseudogén similar a parte de KIAA0172, el gen para una nueva proteína Hs.88959 W94690
1263	ESTs: altamente similar a CIKG, PROTEÍNA DE CANALES DE POTASIO DEPENDIENTES DE VOLTAJE HUMANA KV3.4 [<i>H. sapiens</i>] Hs.106486 H11376
1264	ARNm de HT023 de <i>Homo sapiens</i> , cds completa Hs.237225 AA169496:
1265	FLJ10339: **proteína hipotética FLJ10339 Hs.203963 H72354
1266	N66278:
1267	ESTs: Hs.6195 AA454745
1268	KIAA1404: proteína KIAA1404 Hs.200317 W72798
1269	PMAIP1: proteína inducida por forbol-12-miristato-13-acetato 1 Hs.96 AA458838 APR=gen de respuesta temprana inmediata=sensible a PMA derivado de ATL
1270	G3BP: proteína de unión a dominio SH3 de proteína de activación de Ras-GTPasa Hs.220689 AA598628
1271	ADNc de <i>Homo sapiens</i> : FLJ22807 fis, clon KAIA2887 Hs.261734 R26854:
1272	<i>Homo sapiens</i> , clon IMAGE:3535294, ARNm, cds parcial Hs.80449 T57359:
1273	CDC16: CDC16 (ciclo de división celular 16, <i>S. cerevisiae</i> , homólogo) Hs.1592 AA410559
1274	FGA: **fibrinógeno, polipéptido A alfa Hs.90765 AA026626
1275	ESTs: Hs.33446 N53560
1276	ADNc FLJ14175 fis de <i>Homo sapiens</i> , clon NT2RP2002979 Hs.288613 AA054704:
1277	ESTs: Hs.44243 AA011390
1278	Clon de ADNc de inserto de longitud completa de ARNm de <i>Homo sapiens</i> EUROIMAGE 42408 Hs.284123 R61732:
1279	ESTs: Hs.53455 AA454165
1280	FLJ11264: proteína hipotética FLJ11264 Hs.11260 AI219094
1281	MBD4: proteína de dominio de unión a metil-CpG 4 Hs.35947 AA010492
1282	FLJ11305: proteína hipotética FLJ11305 Hs.7049 N94612
1283	<i>Homo sapiens</i> , similar a producto génico de CG5057, clon MGC:5309, ARNm, cds completa Hs.13885 AA460004:
1284	ARHB: familia de genes de homólogo de ras, miembro B Hs.204354 H88963
1285	ITPR3: receptor de inositol 1,4,5-trifosfato, tipo 3 Hs.77515 AA865667
1286	HMG20B: grupo de alta movilidad 20B Hs.32317 AA775743
1287	ESTs: Hs.146276 AI214204
1288	PTPN9: proteína tirosina fosfatasa, tipo no receptora 9 Hs.147663 AA434420
1289	ARNm del clon FLB9213 PRO2474 de <i>Homo sapiens</i> , cds completa Hs.21321 AA486770:

ES 2 991 436 T3

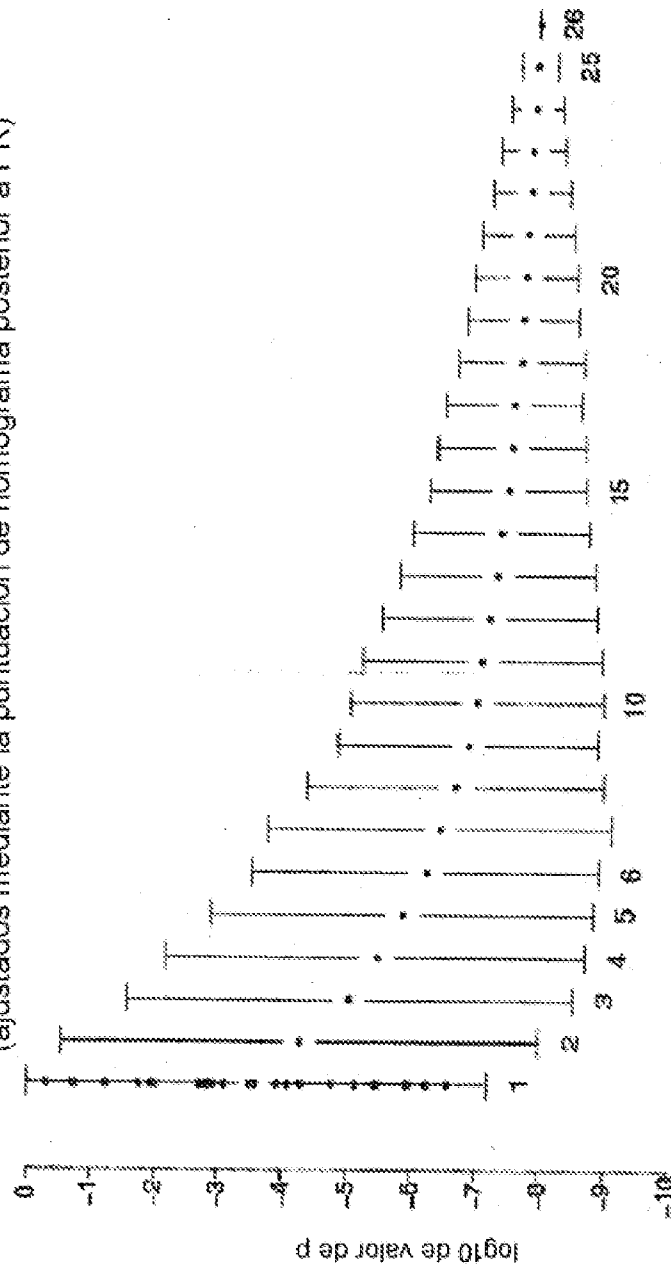
1290	H21107:
1291	HSPC157: proteína HSPC157 Hs.279842 N20480
1292	ARNm de <i>Homo sapiens</i> ; ADNc DKFZp564O2363 (del clon DKFZp564O2363) Hs.321403 AA406332:
1293	ESTs: Hs.150623 AA693532
1294	EST: Hs.188697 AA199733
1295	CLECSF2: lectina de tipo C (dependiente de calcio, dominio de reconocimiento de hidratos de carbono), miembro 2 de superfamilia (inducido por activación) Hs.85201 H11732 AICL=lectina de tipo C inducida por activación
1296	ITPR1: receptor de inositol 1,4,5-trifosfato, tipo 1 Hs.198443 AA035450
1297	CHML: similar a coroideremia (proteína de escolta de Rab 2) Hs.170129 R91881
1298	CDC42: ciclo de división celular 42 (proteína de unión a GTP, 25 kD) Hs.146409 AA668681
1299	FKBP5: **proteína de unión a FK506 5 Hs.7557 AA872767

REIVINDICACIONES

1. Método *in vitro* para clasificar el cáncer de próstata en un paciente, comprendiendo el método:
 - 5 (a) determinar, en una muestra del paciente, la expresión media de un panel de genes que comprende al menos diez genes del ciclo celular seleccionados de uno cualquiera de los paneles C-G; y
 - (b) comparar la expresión media con el nivel de expresión medio de los mismos genes en una población de referencia de pacientes con el mismo cáncer,
 - 10 en el que una elevación de la expresión media con respecto a la población de referencia indica un mal pronóstico.
2. Método según la reivindicación 1, que comprende además determinar al menos un parámetro clínico asociado a riesgo para el paciente.
3. Método según la reivindicación 2, en el que la presencia de al menos un parámetro clínico asociado a riesgo y una expresión media elevada para dichos genes del ciclo celular indica que dicho paciente tiene una probabilidad aumentada de recidiva.
4. Método según una cualquiera de las reivindicaciones 2 a 3, que comprende determinar al menos dos o más parámetros clínicos que forman o son parte de un nomograma.
5. Método según la reivindicación 4, en el que dicho nomograma comprende los parámetros clínicos del nomograma de Kattan-Stephenson.
6. Método según una cualquiera de las reivindicaciones 2 a 5, en el que dicho parámetro clínico se elige del grupo que consiste en estadio tumoral, grado tumoral, tamaño tumoral, estado de los ganglios linfáticos, histología, estado de rendimiento, tipo de cirugía, márgenes quirúrgicos, tipo de tratamiento, edad de aparición, niveles de PSA, y puntuación de Gleason.
7. Método según una cualquiera de las reivindicaciones 1 a 5, en el que el paciente presenta parámetros clínicos de enfermedad no agresiva o de riesgo bajo, o en el que el paciente se ha tratado previamente para el cáncer.
8. Método según una cualquiera de las reivindicaciones 1 a 7, en el que dicho panel de genes comprende al menos diez genes del ciclo celular en el panel F.
9. Método según una cualquiera de las reivindicaciones 1 a 8, en el que determinar el estado del panel de genes comprende al menos uno de:
 - (a) determinar el nivel de un ácido nucleico codificado por dichos genes del ciclo celular; o
 - (b) determinar el nivel de un polipéptido codificado por dichos genes del ciclo celular.
10. Método según una cualquiera de las reivindicaciones 1 a 9, en el que dicha expresión media elevada de los genes del ciclo celular comprende al menos uno de:
 - (a) un nivel de expresión aumentado de un ácido nucleico codificado por dichos genes del ciclo celular; o
 - (b) un nivel de expresión aumentado de un polipéptido codificado por dichos genes del ciclo celular.
11. Método para determinar un transcurso de tratamiento para un paciente con cáncer de próstata, que comprende:
 - (a) determinar al menos un parámetro clínico asociado a riesgo para dicho paciente;
 - (b) determinar, en una muestra del paciente, la expresión media de un panel de genes que comprende al menos diez genes del ciclo celular seleccionados de uno cualquiera de los paneles C-G; y
 - (c) comparar la expresión media con el nivel de expresión medio de los mismos genes en una población de referencia de pacientes con el mismo cáncer,
 - en el que una elevación de la expresión media con respecto a la población de referencia indica un mal pronóstico, y;

- (a) recomendar tratamiento activo si dicho paciente tiene al menos un parámetro clínico asociado a riesgo y una expresión media elevada para dichos genes del ciclo celular, o
- 5 (b) recomendar conducta expectante si dicho paciente no tiene ni un parámetro clínico asociado a riesgo ni una expresión media elevada para dichos genes del ciclo celular.
12. Método para determinar si un paciente con cáncer de próstata tiene una probabilidad aumentada de recidiva, que comprende:
- 10 (a) determinar, en una muestra del paciente, la expresión media de un panel de genes que comprende al menos diez genes del ciclo celular seleccionados de uno cualquiera de los paneles C-G;
- (b) acceder a la información obtenida en (a) en un medio legible por ordenador;
- 15 (c) consultar esta información para determinar si la muestra muestra una expresión media aumentada de los genes del ciclo celular en comparación con el nivel de expresión medio de los mismos genes en una población de referencia de pacientes con el mismo cáncer; y
- 20 (d) emitir una indicación de que el paciente tiene una probabilidad aumentada de recidiva si la muestra muestra una expresión media aumentada de los genes del ciclo celular.
13. Uso de una pluralidad de oligonucleótidos para determinar, en una muestra de un paciente, el nivel de expresión de un panel de genes que comprende al menos diez genes del ciclo celular seleccionados de uno cualquiera de los paneles C-G, para determinar el pronóstico o la probabilidad de recidiva del cáncer de
- 25 próstata en el paciente.

Valores de p obtenidos de subconjuntos aleatorios de 26 genes del ciclo celular
usados para predecir la recidiva química del cáncer
(ajustados mediante la puntuación de nomograma posterior a PR)



Análisis de RP de Cox multivariante del tiempo hasta la recidiva tras la cirugía de próstata
100 muestras FFPE
Expresión de ARNm mediante TLDA
1000 simulaciones aleatorias extraídas de 26 CCC

Figura 1

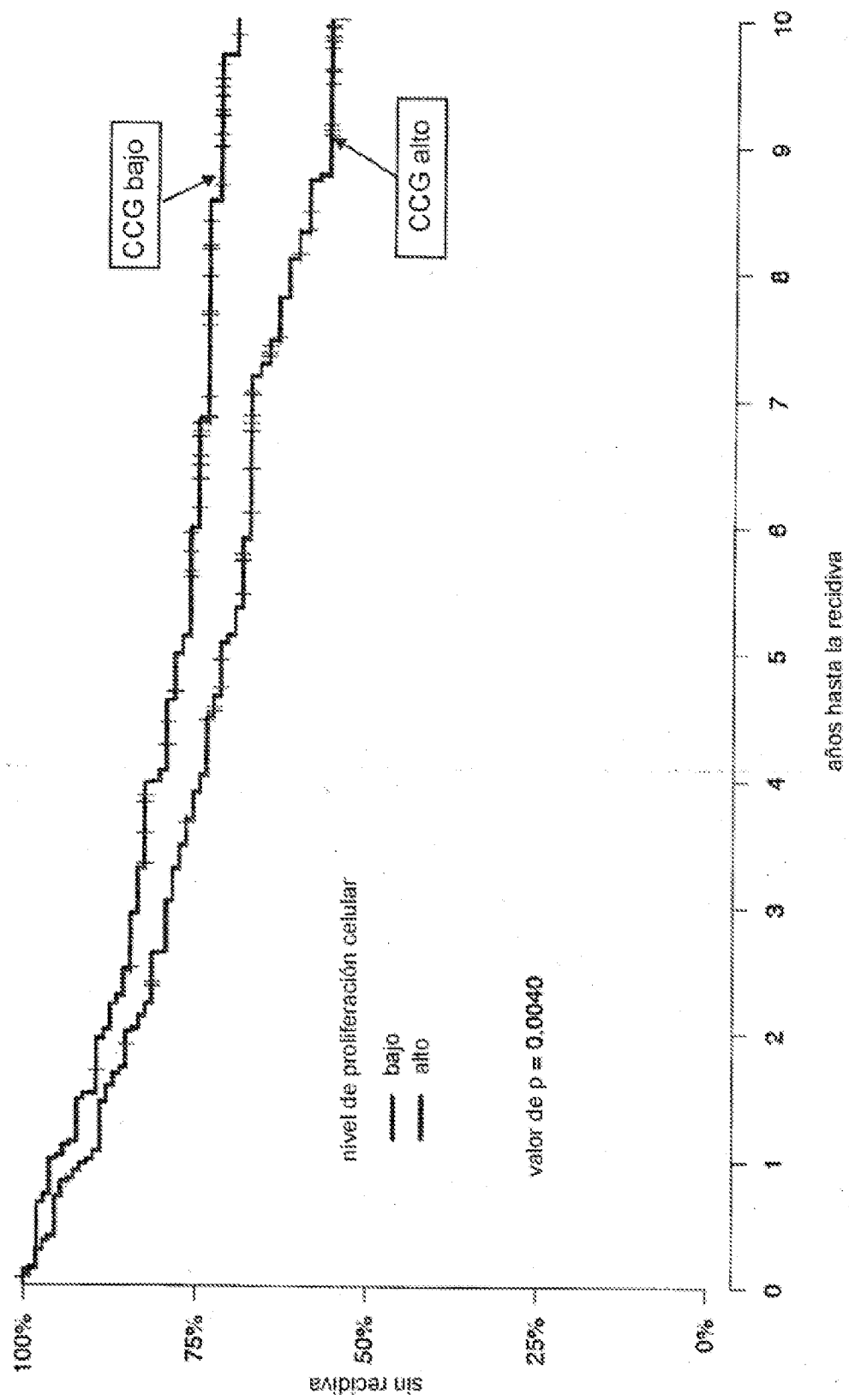


Figura 2

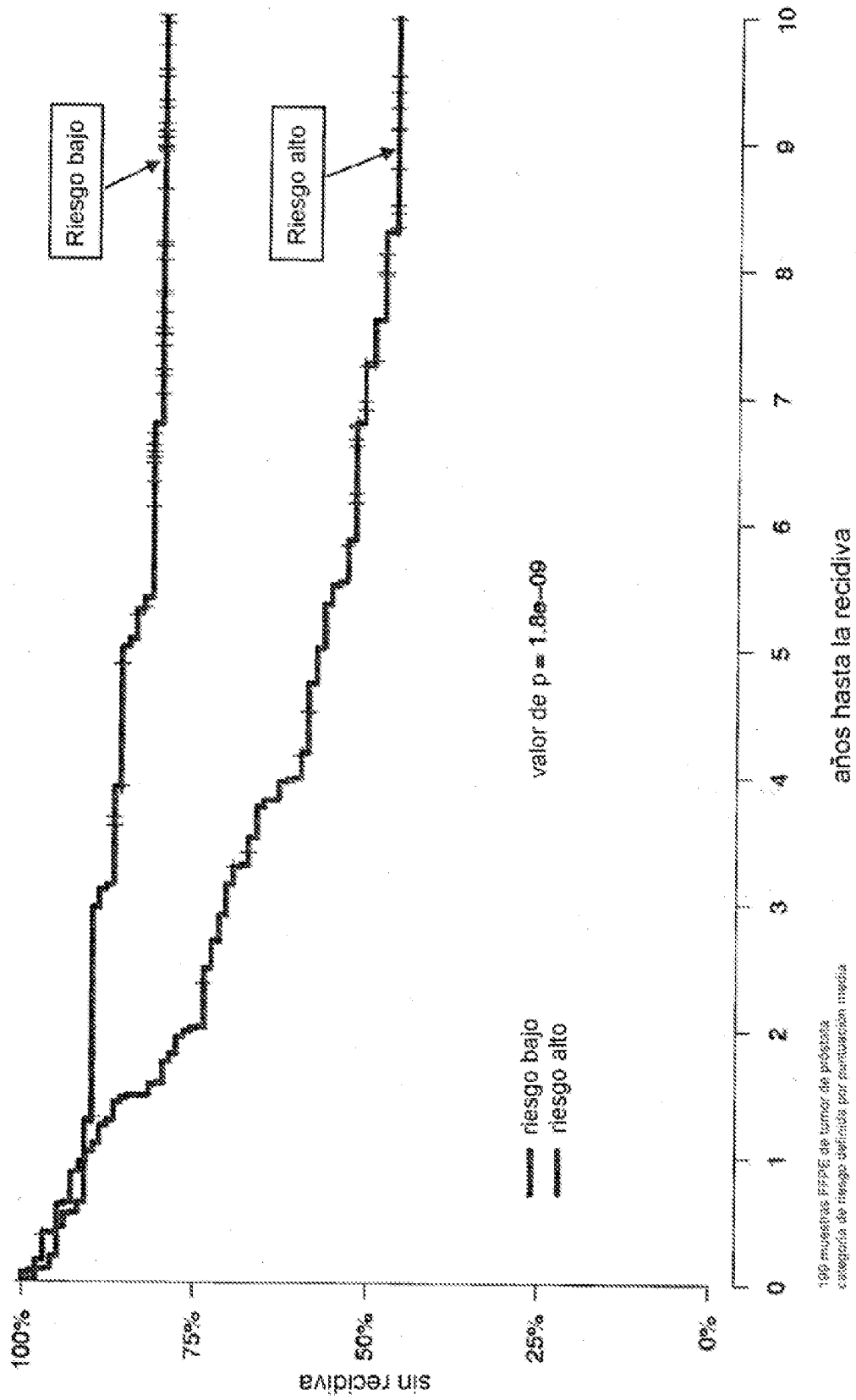


Figura 3

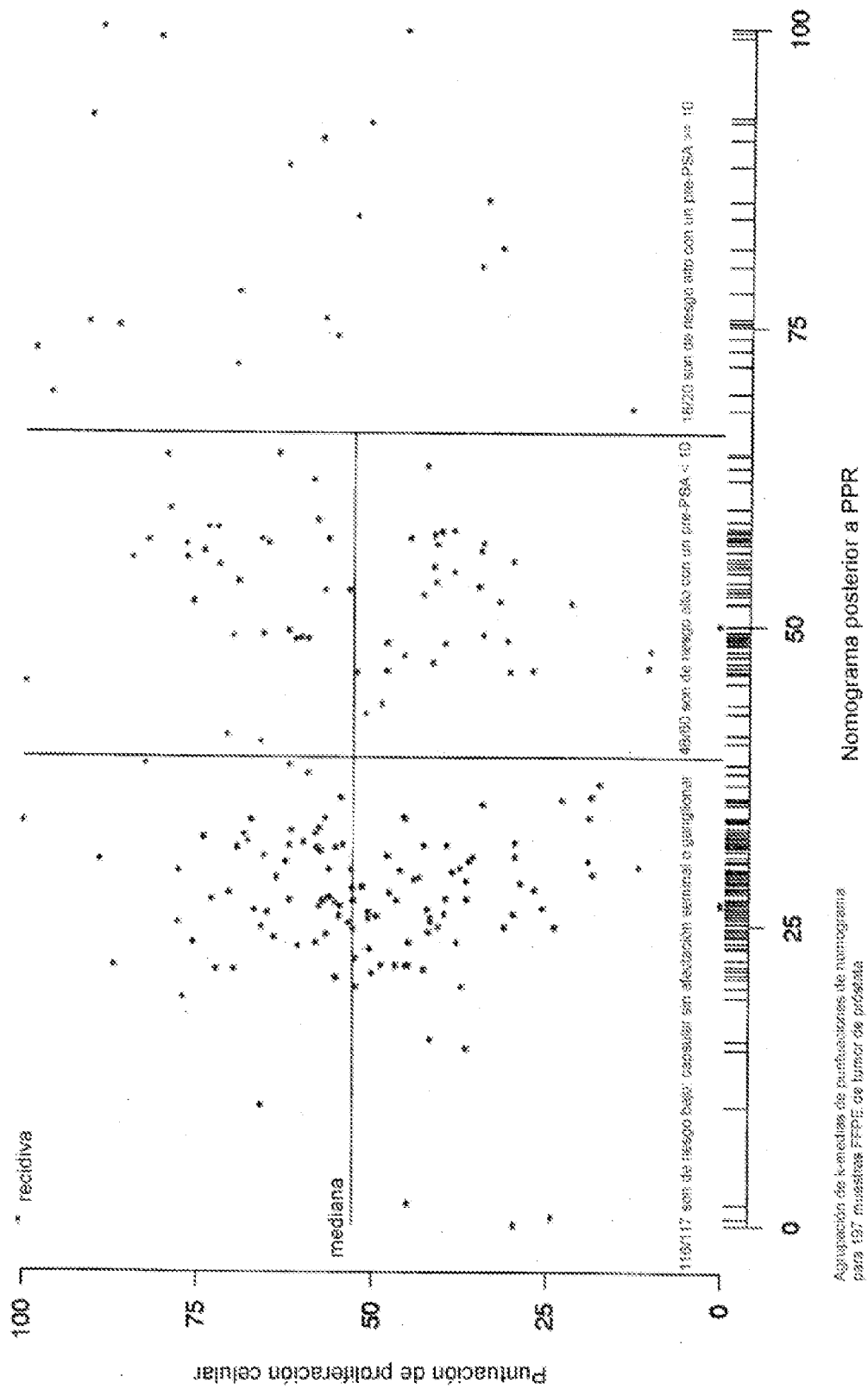


Figura 4

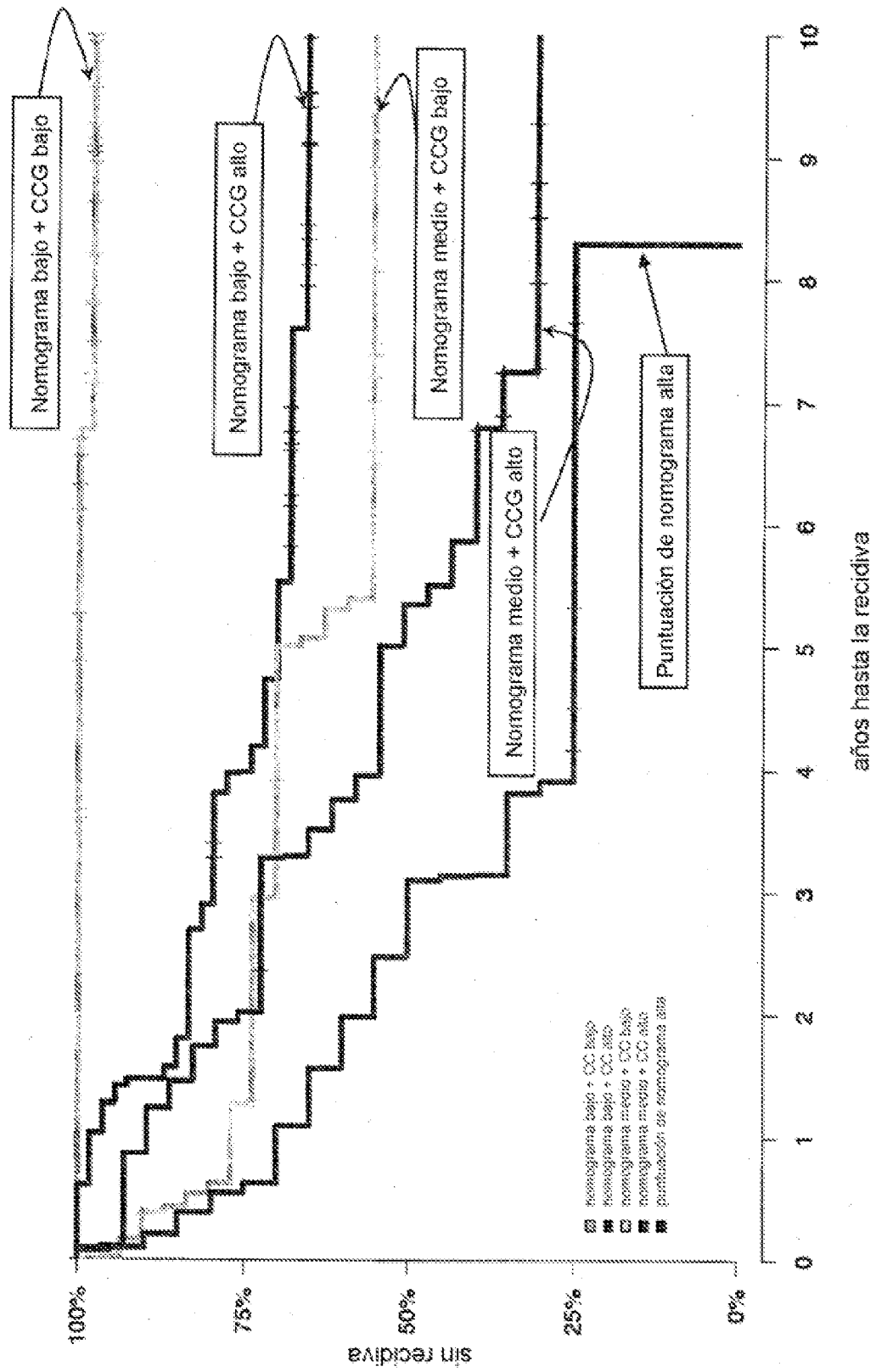


Figura 5

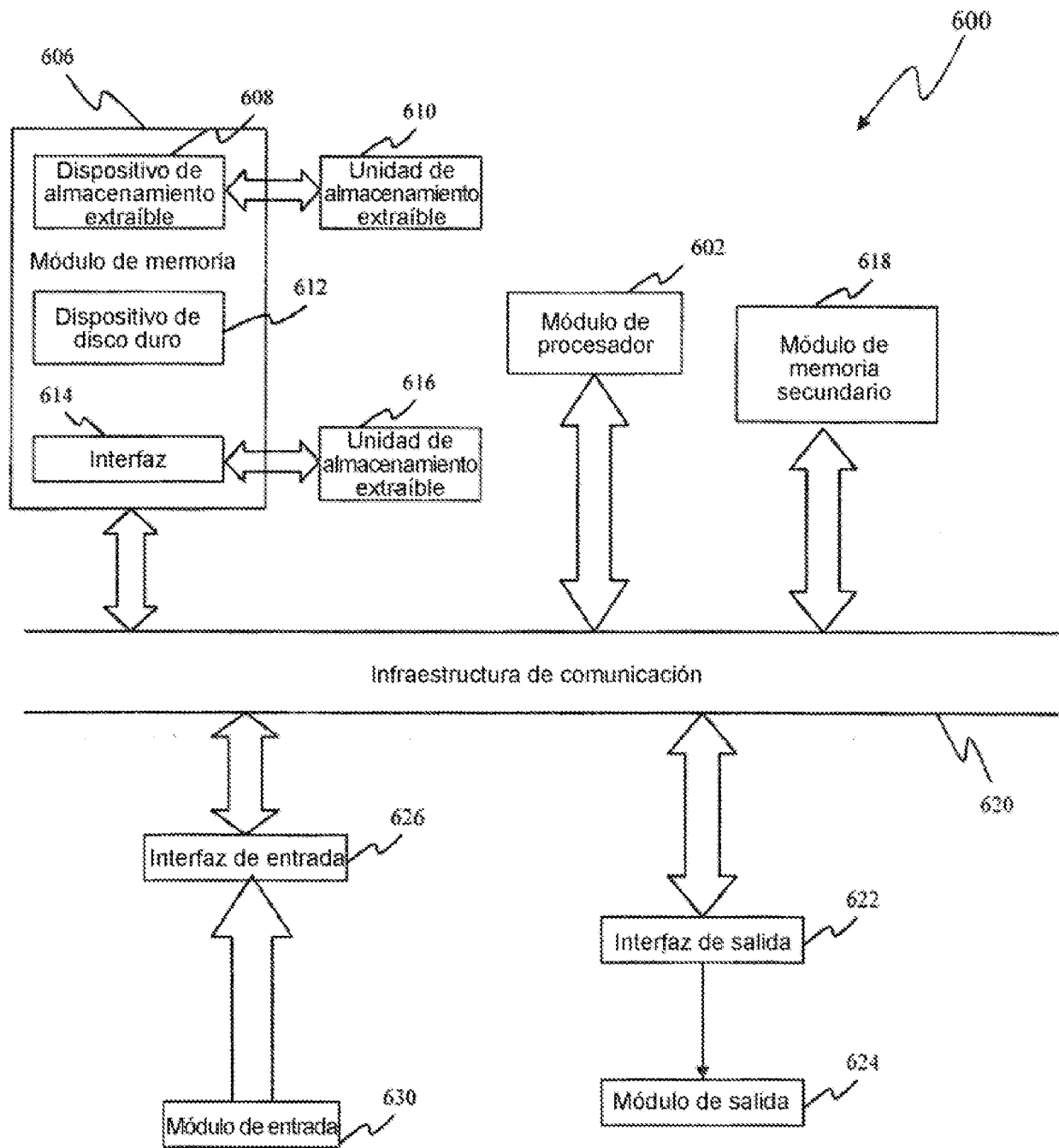


Figura 6

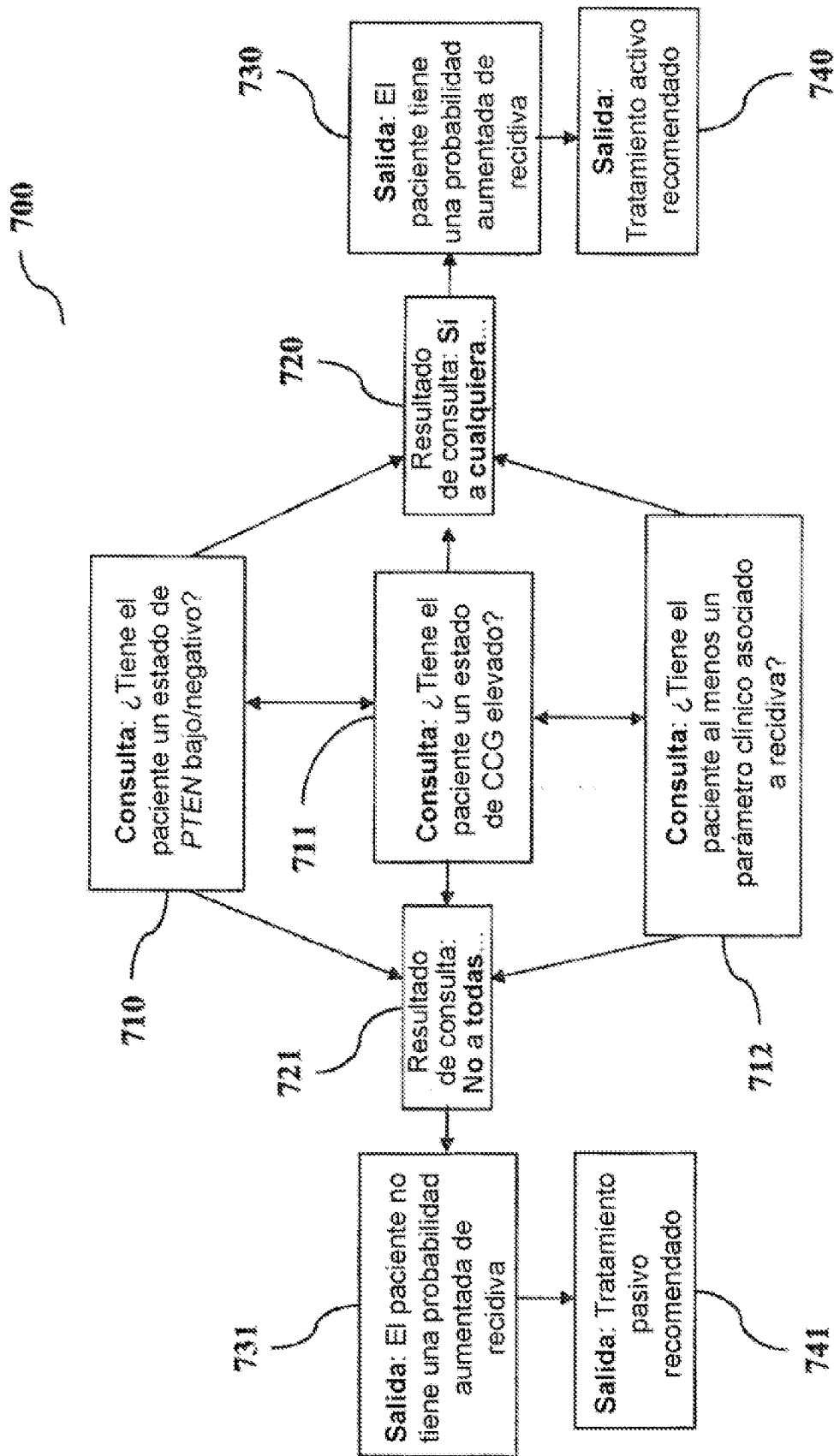


Figura 7

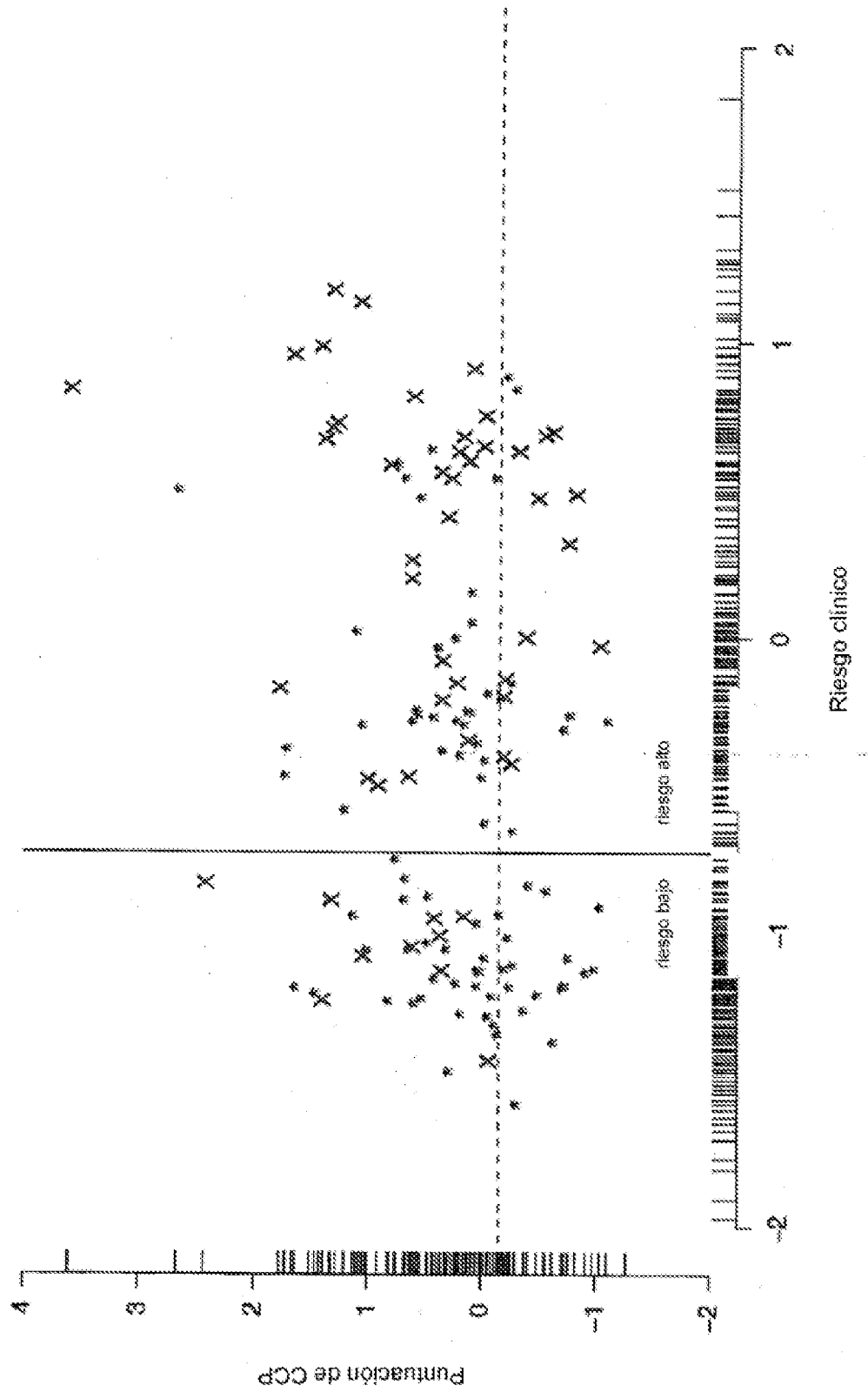


Figura 8

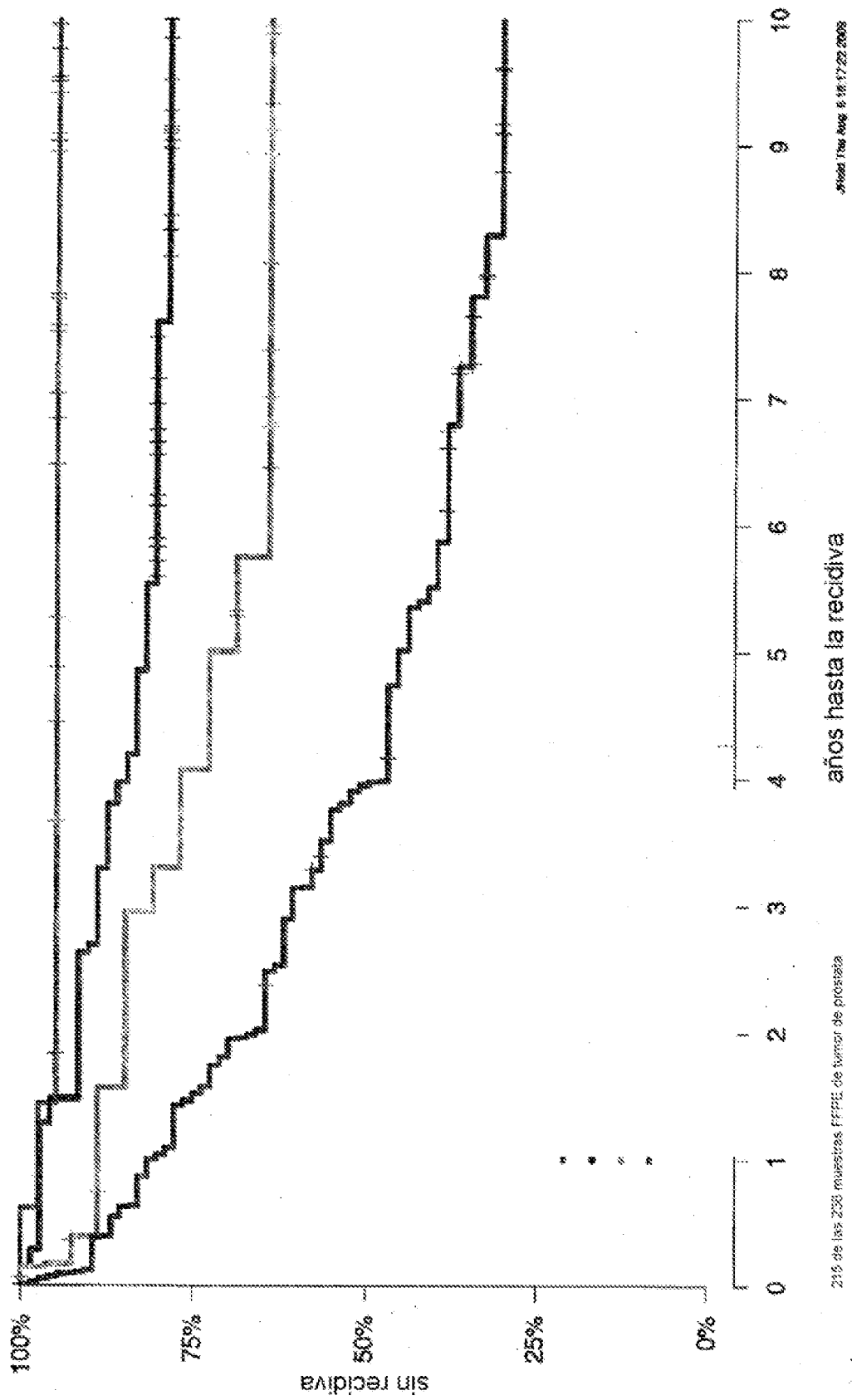


Figura 9

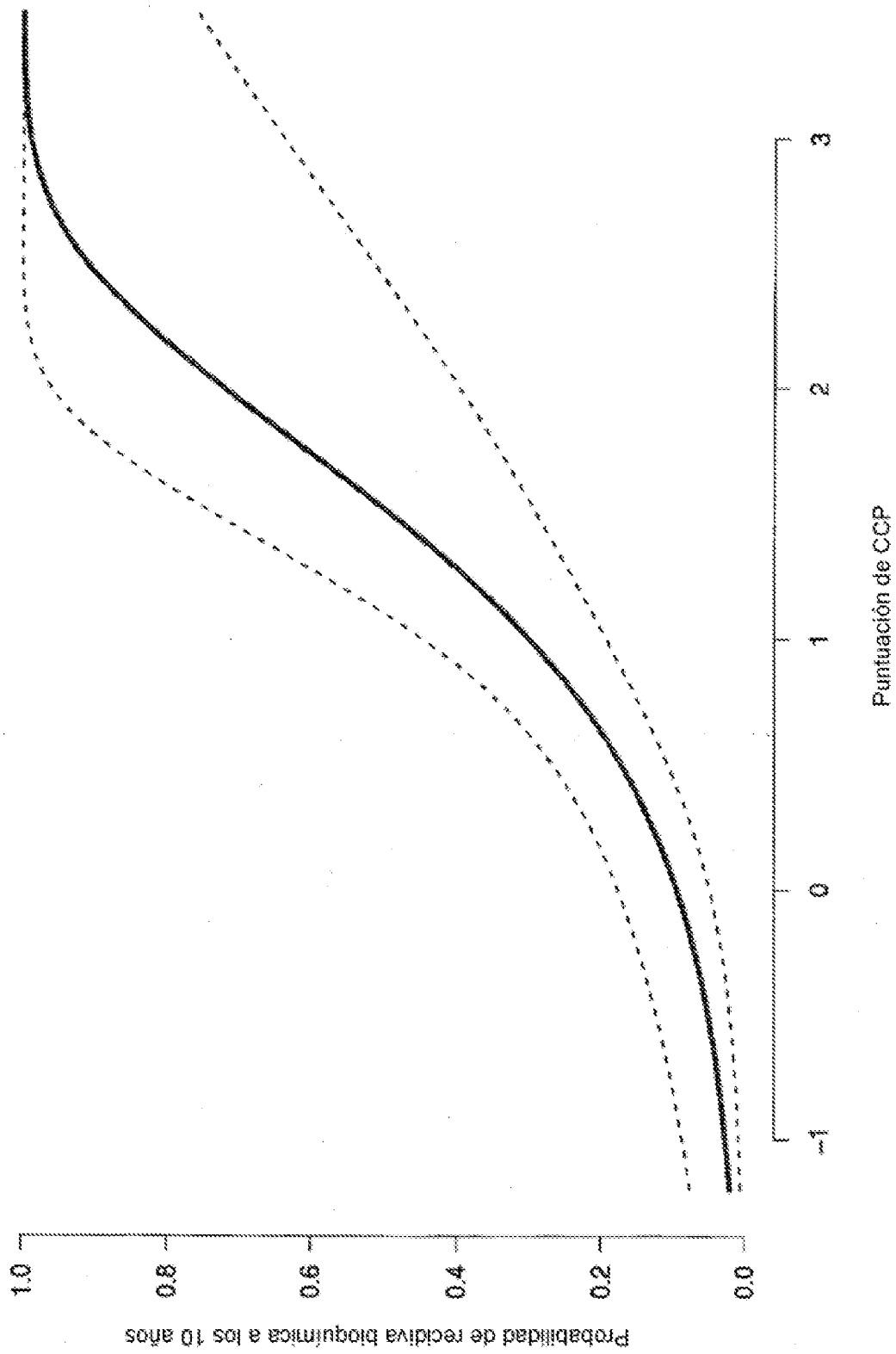


Figura 10

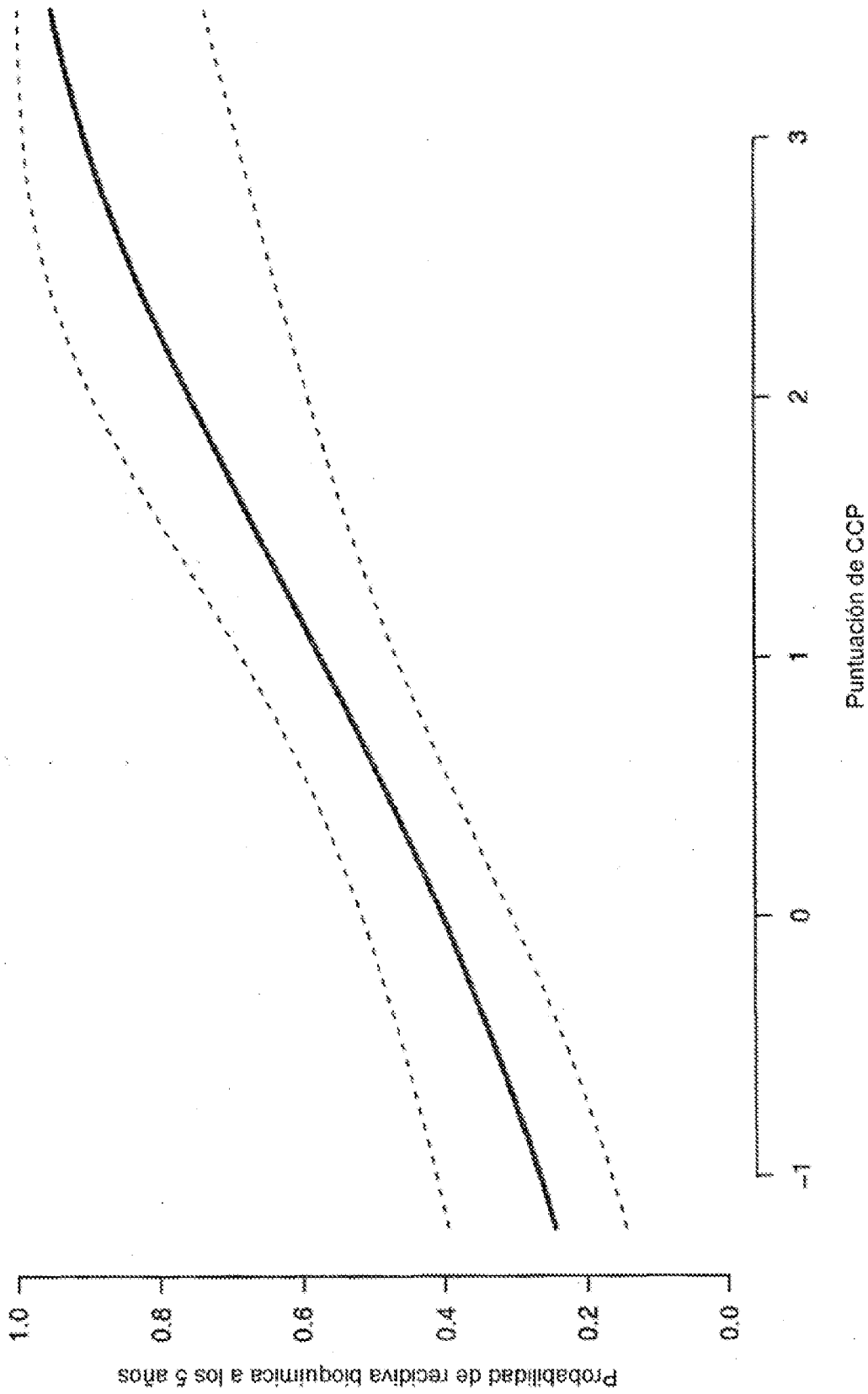


Figura 11

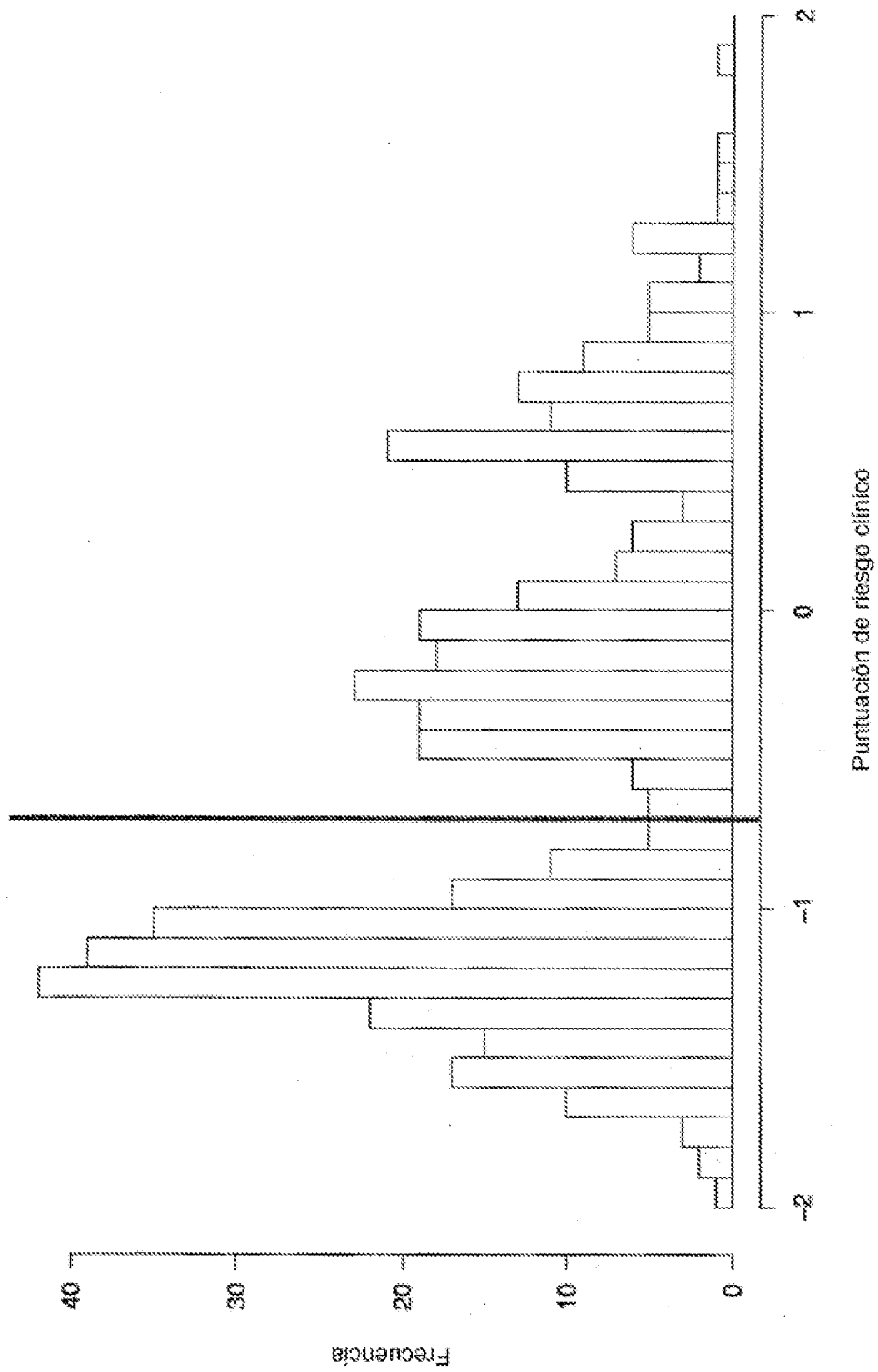


Figura 12

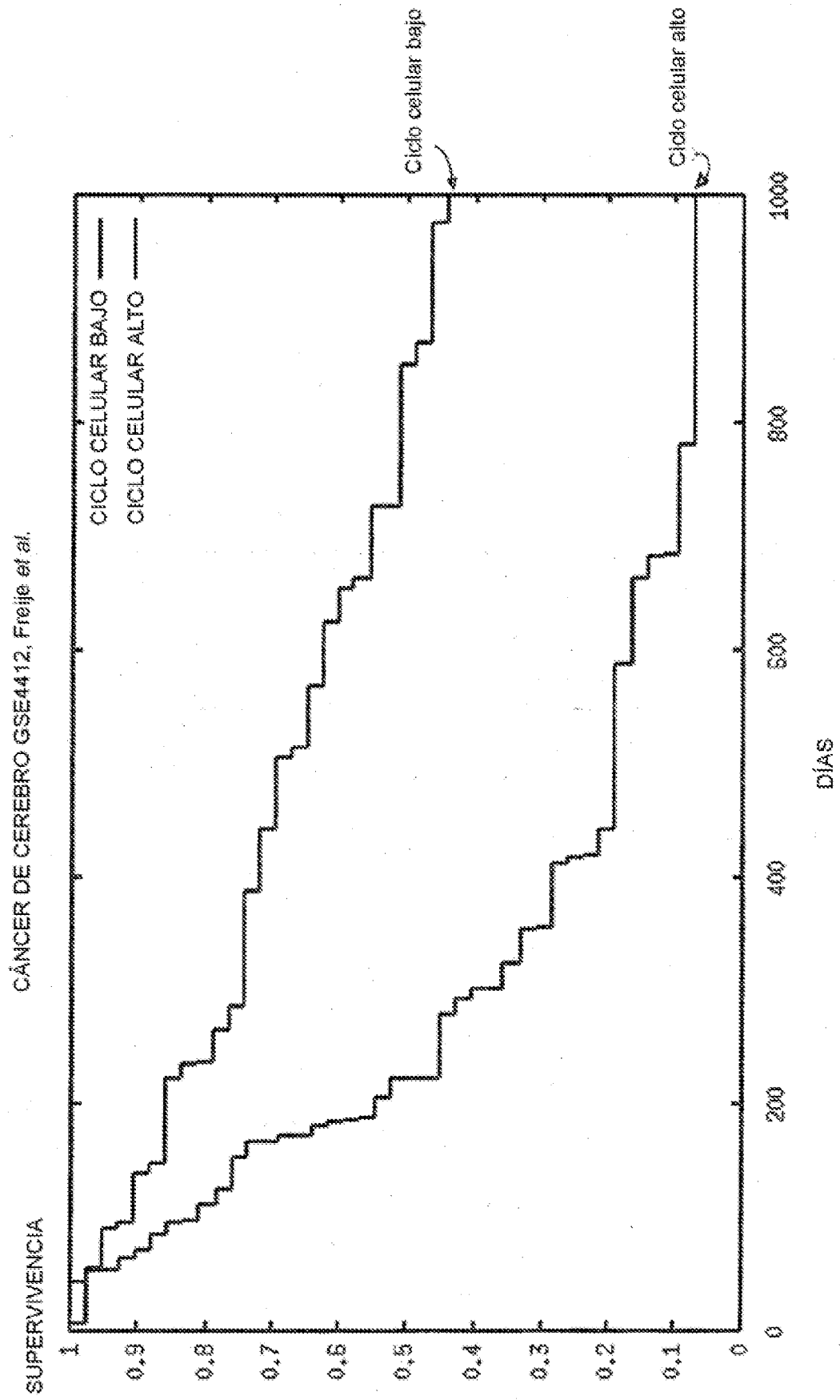


Figura 13

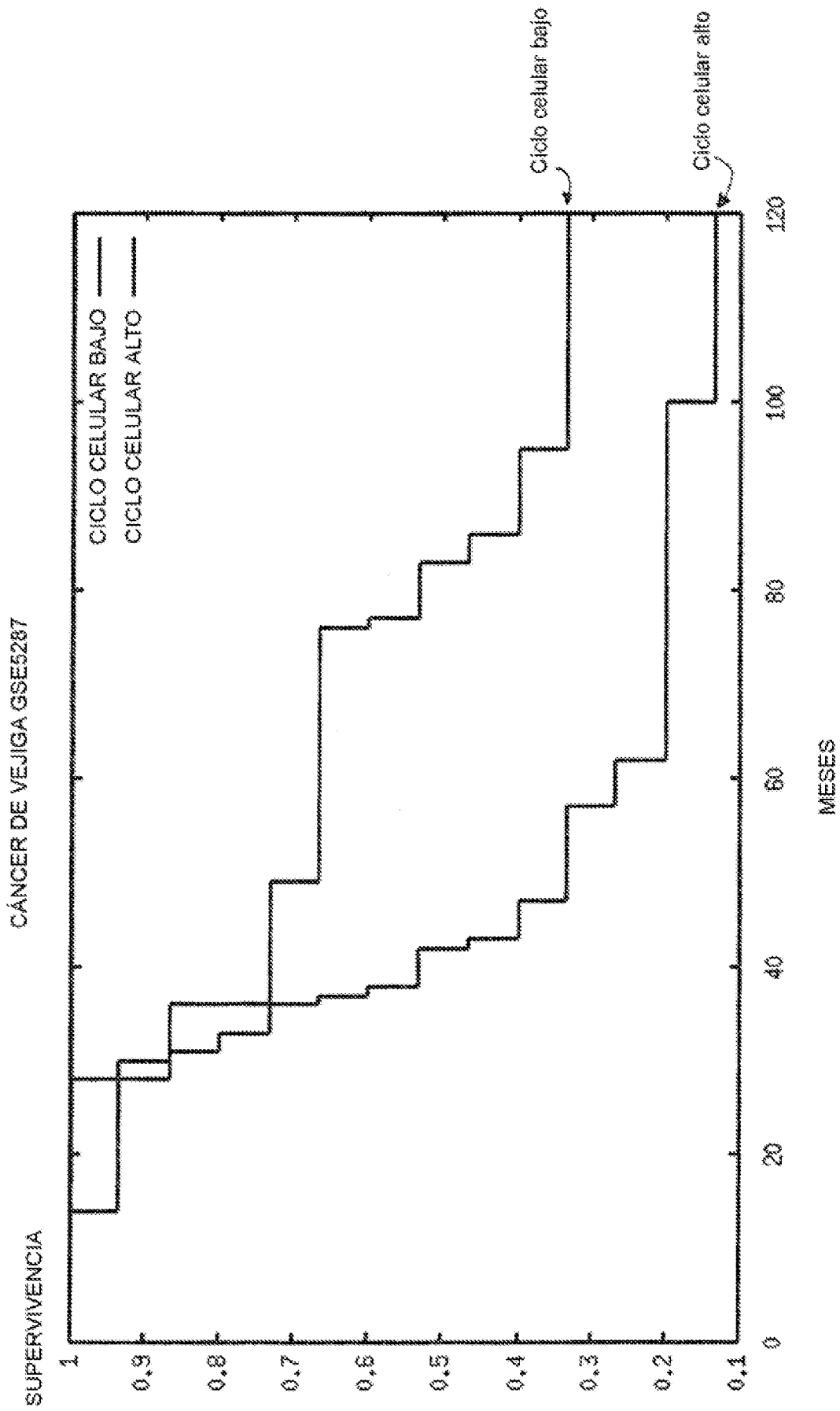


Figura 14

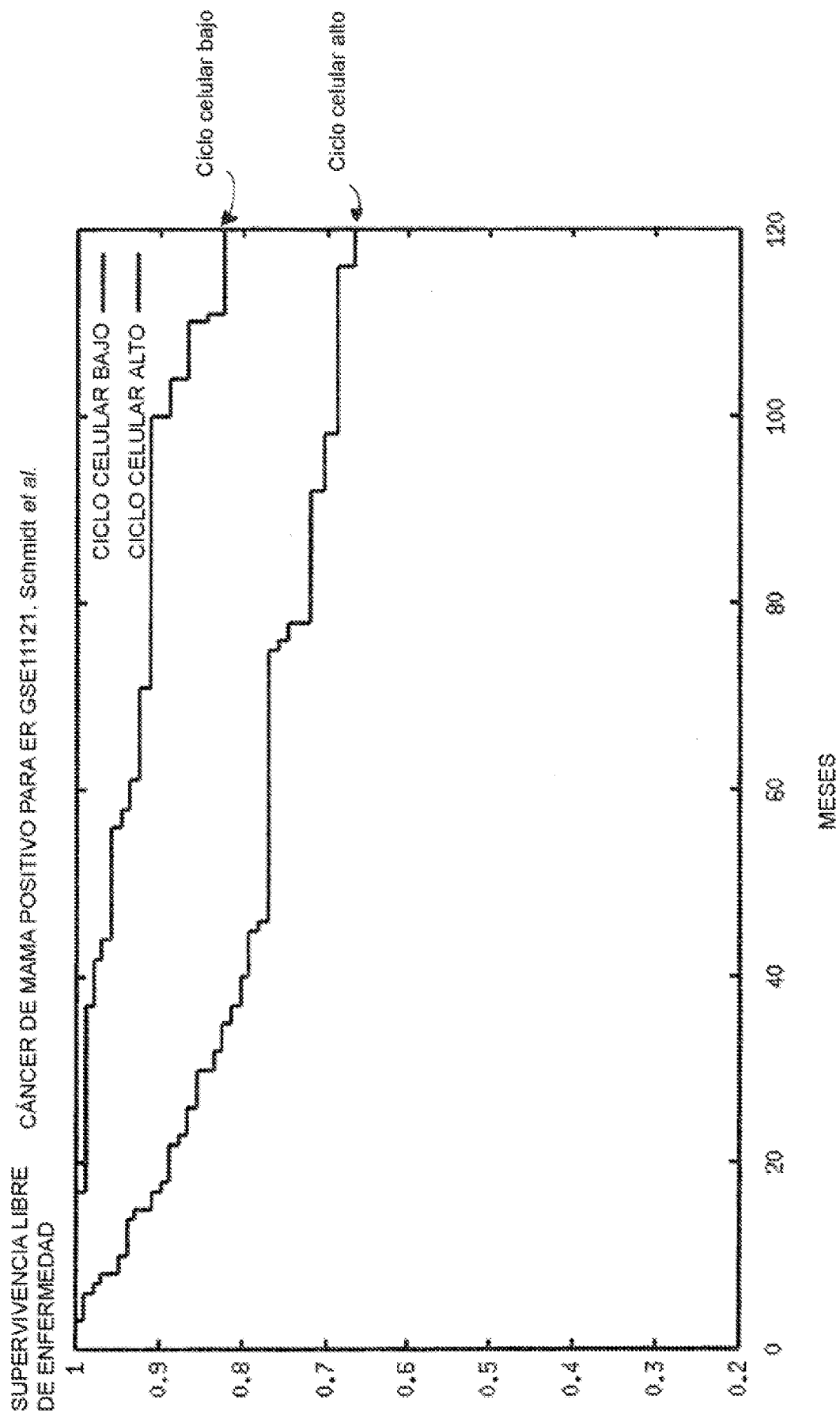


Figura 15

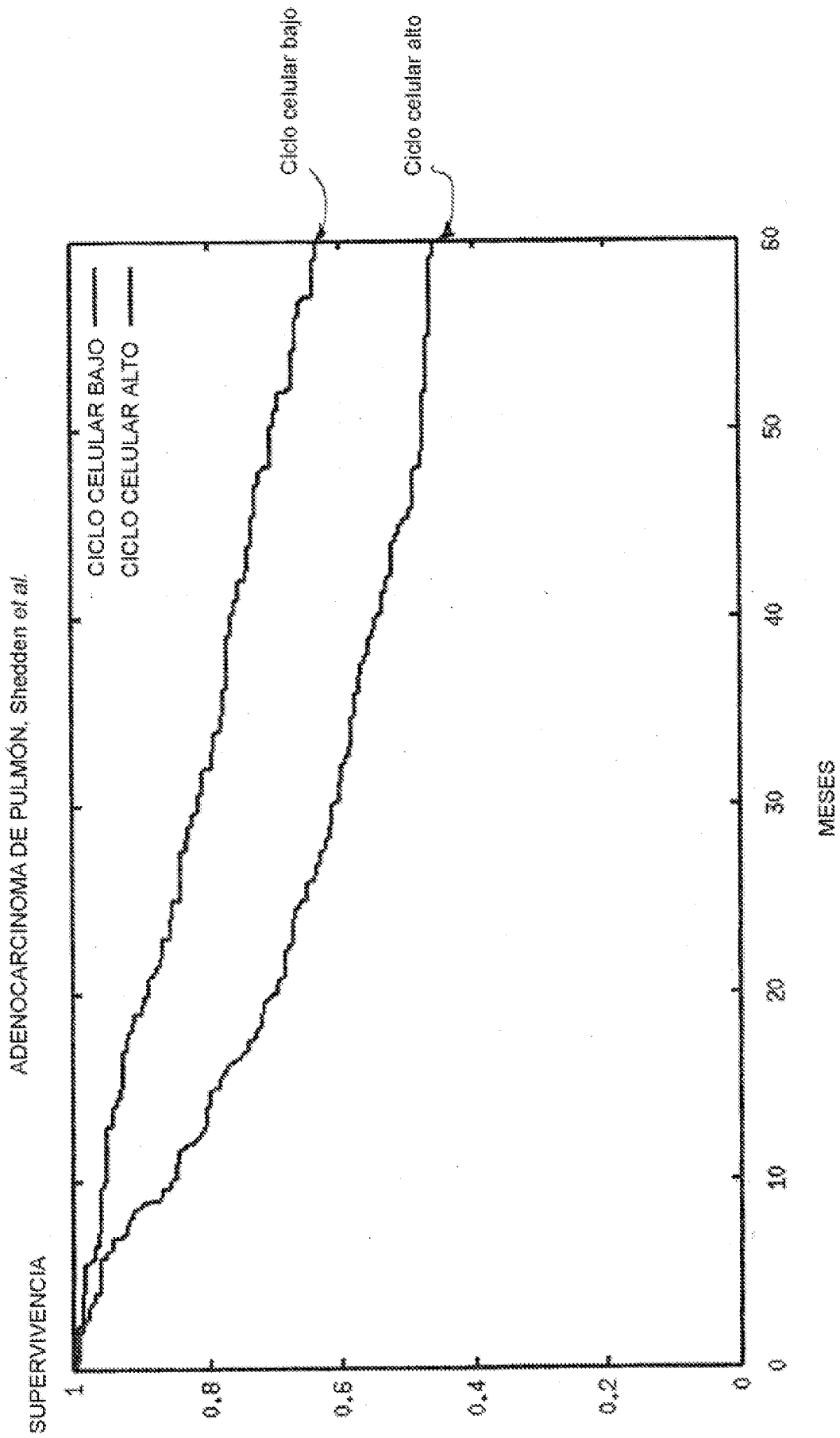


Figura 16