

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 614 494**

51 Int. Cl.:

C12Q 1/68 (2006.01)

C07H 21/04 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **07.11.2007 PCT/US2007/023404**

87 Fecha y número de publicación internacional: **15.05.2008 WO08057545**

96 Fecha de presentación y número de la solicitud europea: **07.11.2007 E 07861764 (4)**

97 Fecha y número de publicación de la concesión europea: **04.01.2017 EP 2087139**

54 Título: **Identificación basada en establecimiento de perfiles de expresión génica de firmas genómicas de mieloma múltiple de alto riesgo**

30 Prioridad:

07.11.2006 US 857456 P

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

31.05.2017

73 Titular/es:

**THE BOARD OF TRUSTEES OF THE UNIVERSITY OF ARKANSAS (100.0%)
2404 North University Avenue
Little Rock, AR 72207-3608, US**

72 Inventor/es:

**SHAUGHNESSY, JOHN D.;
ZHAN, FENGHUANG;
BARLOGIE, BART y
BURLINGTON, BART E.**

74 Agente/Representante:

DE ELZABURU MÁRQUEZ, Alberto

ES 2 614 494 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

Identificación basada en establecimiento de perfiles de expresión génica de firmas genómicas de mieloma múltiple de alto riesgo

Antecedentes de la invención

5 Campo de la invención

La presente invención se refiere en general al campo de la investigación del cáncer. Más específicamente, la presente invención se refiere a la utilización de perfiles de expresión génica para identificar firmas genómicas específicas para el mieloma múltiple de alto riesgo útil para predecir el resultado clínico y la supervivencia.

Descripción de la técnica relacionada

10 El mieloma múltiple (MM), una neoplasia maligna de las células plasmáticas diferenciadas terminalmente que anidan en y se expanden a la médula ósea, se caracteriza por una gran heterogeneidad en los resultados siguientes a las terapias convencionales de dosis elevadas. Si bien, se conocen muchas de las lesiones genéticas y moleculares asociadas con el inicio de la enfermedad, las lesiones que promueven un curso clínico agresivo no se han podido concretar.

15 Todos los mielomas pueden dividirse a grandes rasgos en enfermedad hiperdiploide y no hiperdiploide [1-4]. La hiperploidía, típicamente asociada con las trisomías de los cromosomas 3, 5, 9, 11, 15, 19 y 21, está presente en aproximadamente 60% de los pacientes [5]. El agrupamiento no supervisado y la factorización de una matriz no negativa de datos de hibridación genómica comparativa de una matriz de oligonucleótidos de alta resolución (aCGH) han revelado que el mieloma hiperdiploide puede ser segregado adicionalmente en dos grupos, uno que exhibe trisomías de los cromosomas impares descritos anteriormente y otro que exhibe, además, las ganancias de los cromosomas 1q y 7, la delección del cromosoma 13 y la ausencia de la trisomía 11 [6]. El mieloma no hiperdiploide también se puede dividir en dos grupos, uno que se caracteriza por la amplificación de alto nivel del cromosoma 1q y las deleciones de los cromosomas 1p y 13, y otro que se caracteriza por la ausencia de anomalías en el cromosoma 1 pero que alberga deleciones de los cromosomas 8 y 13 [6]. Además, la activación transcripcional de *CCND1*, *CCND3*, *MAF*, *MAFB*, o *FGFR3/MMSET* (como resultado de las translocaciones que afectan al locus de la cadena pesada de inmunoglobulina en el cromosoma 14q32) es típica del mieloma no hiperdiploide y está presente en aproximadamente 40% de los casos [5,7,8].

20 Utilizando un agrupamiento jerárquico sin supervisión de los patrones de expresión génica global, se definieron recientemente subgrupos que presentaban fuertes correlaciones con la hiperdiploidía y las translocaciones recurrentes y se validó la existencia de siete mielomas [9]. Se identificaron dos entidades de alto riesgo, una que revelaba la expresión en exceso de genes de proliferación y derivada de casos que evolucionaban desde las otras seis clases, mientras que la otra estaba definida por t(4;14)(p16;q32) [9].

25 Las ganancias del brazo largo del cromosoma 1 (1q) son una de las anomalías genéticas más comunes en el mieloma [10]. Las duplicaciones en tándem y las duplicaciones segmentarias de salto de banda del cromosoma 1q, resultantes de la descondensación de la heterocromatina pericentromérica, están frecuentemente asociadas con el progreso de la enfermedad [11-13]. Utilizando la aCGH sobre ADN aislado de células plasmáticas derivadas de pacientes con mieloma latente, Rosinol y sus colegas demostraron que el riesgo de conversión en enfermedad manifiesta estaba vinculado con las ganancias de 1q21 y la pérdida del cromosoma 13 [14]. Estos descubrimientos fueron confirmados mediante el uso del análisis de hibridación in situ con fluorescencia (FISH) de la interfase. Además, se demostró que las ganancias de 1q21 obtenidas en el mieloma sintomático estaban vinculadas a una supervivencia inferior y se amplificaban adicionalmente en la recaída de la enfermedad [15].

30 El documento WO 2005/116259 describe una lista de 70 genes cuya expresión puede ser analizada para determinar un pronóstico de supervivencia después del tratamiento para el cáncer.

35 Shaughnessy et al. [Biology of Blood and Marrow Transplantation 12:77-80 (2006)] describen que la expresión de *CKS1B* está vinculada con el resultado de supervivencia tras el tratamiento con bortezomib y talidomida.

40 Zhan et al. [Blood 2006 108:6 2020-2028] describen el uso de análisis de perfiles de genes utilizando un pequeño número de genes para predecir la supervivencia en el mieloma múltiple en pacientes tratados con melfalán [9].

Compendio de la invención

45 La presente invención está dirigida a un método de establecimiento de perfiles de la expresión génica para identificar firmas genómicas vinculadas a un pronóstico de supervivencia en un sujeto con mieloma múltiple sintomático o mieloma múltiple, que comprende la determinación de los niveles de expresión de genes que consisten en *KIF14*, *SLC19A1*, *CKS1B*, *YWHAZ*, *MPHOSPH1*, *TMPO*, *NADK*, *LARS2*, *TBRG4*, *AIM2*, *ASPM*, *AHCYL1*, *CTBS*, *MCLC*, *LTBP1* y genes detectables con ácidos nucleicos que tienen los números de referencia de Affymetrix 242488_at y

1557277-9_at en una micromatriz Affymetrix U133PLW2.0 en una muestra de ácido nucleico de células plasmáticas aisladas del sujeto, y la determinación de la presencia de niveles anormales de expresión, realizando de ese modo un pronóstico del sujeto.

Breve descripción de los dibujos

5 **Las Figuras 1A-1D** muestran patrones de expresión de genes que distinguen los grupos de riesgo en la cohorte de formación. La **Figura 1A** muestra los mapas de calor de los 70 genes que ilustran patrones de expresión notablemente similares entre los 351 pacientes recién diagnosticados utilizados para identificar los 70 genes. Las barras de color rojo por encima de las columnas de pacientes denotan casos con muertes relacionadas con la enfermedad. Los 51 genes en las filas designadas por la barra de color rojo a la izquierda (filas superiores, regulados al alza), identifican los pacientes en el cuartil superior de expresión con alto riesgo de muerte temprana relacionada con la enfermedad. Las líneas de los 19 genes designados por la barra de color verde (regulados a la baja), identifican los pacientes en el cuartil inferior de expresión con alto riesgo de muerte temprana relacionada con la enfermedad. La **Figura 1B** muestra las frecuencias de la cohorte de formación para las diferencias de muestras entre las razones de la media de la expresión en \log_2 de los 51 genes regulados al alza/19 genes regulados a la baja. Esta razón de la expresión auto-normalizadora tiene una marcada distribución bimodal, en consonancia con el análisis de expresión diferencial de rangos del cuartil superior/inferior, que fue diseñado para detectar los genes que definen un único grupo de alto riesgo (13,1%) con una distribución de la expresión extrema. Interpretada como una razón de la regulación al alza/a la baja en la escala \log_2 , los valores más altos se asocian con un peor resultado. La línea vertical muestra el punto de corte de alto riesgo vs. bajo riesgo para la proporción a escala \log_2 determinada por el agrupamiento de K-medias: también se muestra el porcentaje de muestras por debajo y por encima del punto de corte. Las **Figuras 1C-1D** demuestran que las estimaciones de Kaplan-Meier de la supervivencia libre de eventos (**Figura 1C**) y global (**Figura 1D**) en el mieloma de bajo riesgo (color azul) y el mieloma de alto riesgo (color rojo) mostraron probabilidades de supervivencia actuarial inferiores a los 5 años libre de eventos (18% v 60%, $P < ,0001$; HR = 4,51) y supervivencia global (28% v 78%, $P < ,0001$, HR = 5,16) en 13,1% de pacientes con la firma de alto riesgo.

La **Figura 2** muestra el análisis de agrupamiento ("Clustergram") de la expresión génica de 70 genes de alto riesgo en las células plasmáticas de 22 sujetos sanos (NPC), 14 sujetos con gammapatía monoclonal de significado incierto ("MGUS por sus siglas en inglés"), 351 pacientes con mieloma recién diagnosticado (MM) y 42 líneas celulares de mieloma múltiple humano ("HMCL por sus siglas en inglés"). Cada fila representa un gen y cada columna una muestra. Los genes se ordenan de arriba a abajo basándose en el rango de la Tabla 1. El color rojo para un gen indica la expresión por encima de la mediana y el color azul por debajo de la mediana. Las muestras dentro de los grupos de riesgo de mieloma fueron ordenadas de manera que el riesgo pronosticado aumenta continuamente de izquierda a derecha.

Las **Figuras 3A-3C** muestran la distribución de grupos de riesgo y los análisis de supervivencia en la cohorte de ensayo. La **Figura 3A** muestra las frecuencias de las cohortes de ensayo para la razón de la media del \log_2 de los genes regulados al alza/a la baja. El punto de corte para el alto riesgo se determinó por el agrupamiento independiente de la razón del \log_2 . Los conjuntos de formación y validación tienen una distribución similar para este resumen de expresión de los 70 genes, incluyendo puntos de corte similares para el alto riesgo y proporciones similares agrupadas en el grupo de alto riesgo. Las **Figuras 3B-3C** muestran las estimaciones de Kaplan-Meier de la supervivencia libre de eventos (**Figuras 3B**) y global (**Figuras 3C**) entre los grupos de riesgo molecular en la cohorte de ensayo.

Las **Figuras 4A-4B** muestran que la puntuación de riesgo de 70 genes en el momento del diagnóstico y la recaída pronostica la supervivencia después de la recaída. La **Figura 4A** muestra la puntuación de riesgo de 70 genes en muestras pareadas de diagnóstico (color azul) y recaída (color rojo) de 51 casos de la cohorte de formación. La puntuación de riesgo de la expresión génica se indica a la izquierda. Los pares de muestras, se ordenan de izquierda a derecha, basándose en la puntuación de referencia más baja. La **Figura 4B** muestra gráficos de Kaplan-Meier de la supervivencia después de la recaída de los tres grupos definidos por un bajo riesgo, tanto en el diagnóstico como en la recaída (Bajo-Bajo), bajo riesgo en el momento del diagnóstico y alto riesgo en la recaída (Bajo-Alto) y alto riesgo en ambos puntos temporales (Alto-Alto).

Las **Figuras 5A-5B** muestran la supervivencia libre de eventos y global en grupos de riesgo definidos por el modelo de 17 genes en el conjunto de ensayo. Los 181 casos de MM recién diagnosticado fueron pronosticados en los grupos de alto riesgo (16,6%) y de bajo riesgo (83,4%). Las estimaciones de Kaplan-Meier de supervivencia en el mieloma de bajo riesgo y de alto riesgo mostraron probabilidades de supervivencia actuarial a los 2 años libre de eventos (**Figura 5A**) de 88% para el alto riesgo (color rojo) frente a 50% para la supervivencia de bajo riesgo (color azul) ($P < ,0001$) y global (**Figura 5B**) de 91% para el alto riesgo (color rojo) frente a 54% para el bajo riesgo (color azul) ($P < ,0001$).

La **Figura 6** muestra la relación entre el alto y el bajo riesgo definida por el modelo supervisado de 70 genes y el clasificador no supervisado del subgrupo 7 (9). Los datos se presentan como una vista de diagrama de barras apiladas del número de casos de alto riesgo (color rojo) y de bajo riesgo (color azul) en cada uno de los 7 subtipos, incluyendo el grupo de casos con la denominada firma mieloides (MY) (extremo izquierdo).

Las **Figuras 7A-7C** relacionaban el alto riesgo definido por el modelo de 70 genes con las características moleculares. La **Figura 7A** muestra un gráfico de dispersión del índice de proliferación basado en la expresión génica (eje x) por la puntuación de riesgo de 70 genes en 351 casos de la cohorte de formación. Se indican los casos de bajo riesgo (color azul) y los casos de alto riesgo (color rojo) definidos por el modelo de 70 genes (véase el texto). Las dos variables muestran un alto grado de correlación ($r = 0,73$; $P < 0,001$). Para evaluar la influencia de las dos variables en el resultado, la población se dividió en 4 subgrupos utilizando un IP del punto de corte de 5 y un punto de corte de alto riesgo de ,66. Los grupos están definidos por la intersección de las dos líneas de color verde. El cuadrante superior izquierdo contiene los casos con un IP alto/bajo riesgo, el cuadrante superior derecho los casos con un IP alto/alto riesgo, el cuadrante inferior izquierdo contiene los casos con IP bajo/bajo riesgo y el cuadrante inferior derecho contiene los casos con IP bajo/alto riesgo. La línea representa la tendencia lineal en los datos. La **Figura 7B** muestra gráficos de Kaplan-Meier de las estimaciones de supervivencia global de los cuatro grupos definidos en la **Figura 7A**, que no revelan ningún impacto de IP dentro de los grupos de riesgo. La **Figura 7C** muestra los gráficos de Kaplan-Meier de las estimaciones de supervivencia global de mieloma positivo para t(4;14) en relación con la designación de la puntuación de alto riesgo de 70 genes de la muestra dada, que muestra el profundo impacto de las puntuaciones de alto y bajo riesgo.

Las **Figuras 8A-8D** muestran que el modelo de expresión de 17 genes se puede utilizar para pronosticar el resultado de recaída de la enfermedad tratada con un único agente Bortezomib. La **Figura 8A** muestra la predicción de grupos de riesgo a partir de datos de U2 y U4 y los modelos de 17 genes y 16 genes. El análisis de predicción del grupo de riesgo se realizó en 144 casos recién diagnosticados de mieloma. Hubo un alto grado de correlación entre los dos modelos. La **Figura 8B** muestra el análisis de supervivencia de Kaplan-Meier en el conjunto de datos Millenium. El modelo de 16 genes se aplicó a 156 pacientes con recaída de mieloma tratados en la prueba APEX [31]. Se pronosticaron 13,5% y 86,5% como de alto riesgo y de bajo riesgo, respectivamente. Las estimaciones de supervivencia global de probabilidades actuariales al año fueron de 74% para la enfermedad de bajo riesgo (color azul) frente a 32% para la de alto riesgo (color rojo) ($P = 0,0014$; $HR = 2,52$). La **Figura 8C** muestra la supervivencia global en la cohorte con bortezomib. Las estimaciones de Kaplan-Meier de la supervivencia en el mieloma de bajo riesgo y de alto riesgo mostraron probabilidades actuariales al año de SG de 79% para la enfermedad de bajo riesgo frente a 36% para la de alto riesgo ($P = 0,0495$; $HR = 2,34$). La **Figura 8D** muestra la supervivencia global en la cohorte con dexametasona. Las estimaciones de Kaplan-Meier de la supervivencia en el mieloma de bajo riesgo y de alto riesgo mostraron probabilidades actuariales al año de SG de 64% para la enfermedad de bajo riesgo en comparación con 23% para la de alto riesgo ($p = 0,0174$; $HR = 2,55$). Los análisis de la prueba APEX completa (669 pacientes) revelaron una SG superior con bortezomib frente a dex, a pesar del cruce de pacientes (30 frente a 20 meses; $P = 0,027$; seguimiento de la mediana de 22 meses, se produjeron 44% de eventos) [32].

Las **Figuras 9A-9B** muestran el análisis de supervivencia global en los 144 casos recién diagnosticados tanto con UA como con U2. La **Figura 9A** muestra el modelo derivado de U2 de 17 genes aplicado a 144 pacientes recién diagnosticados con mieloma. Las estimaciones de Kaplan-Meier de la supervivencia global revelaron diferencias significativas en las probabilidades de supervivencia actuarial a los 5 años (50% vs 68%, $P = 0,0046$; $HR = 2,40$) en el mieloma de bajo riesgo color (azul) y el mieloma de alto riesgo (color rojo), respectivamente. La **Figura 9B** muestra las estimaciones de Kaplan-Meier de SG en los 144 pacientes basándose en el riesgo definido por el modelo derivado de UA de 16 genes. Éste reveló probabilidades de supervivencia actuarial a los 5 años en 46 vs 70%, $P = 0,0026$; $HZ = 2,58$ en el riesgo alto frente al riesgo bajo, respectivamente.

Descripción detallada de la invención

La variabilidad de la supervivencia de los pacientes con mieloma múltiple no está bien representada con los parámetros de laboratorio actuales, tales como los niveles de beta-2-microglobulina y albúmina empleados en el sistema de estadificación ISS [23]. La enfermedad de alto riesgo de novo puede ser fundamentalmente diferente del mieloma con resistencia adquirida a fármacos y un curso clínico agresivo después de recaídas recurrentes.

El marcado aumento en la frecuencia de designación de alto riesgo de 13% en el momento del diagnóstico con respecto a 76% en la recaída proporciona una evidencia molecular de la evolución de la enfermedad que influye en el resultado después de la recaída. Se puede desarrollar un fenotipo de mieloma agresivo, ya sea de novo o adquirido, a través de un mecanismo similar. Con el perfeccionamiento adicional de este modelo comentado en la presente memoria, la presente invención contempla el desarrollo de herramientas para la evaluación cuantitativa de los riesgos durante todo el curso de la gestión terapéutica.

Además de su relevancia clínica, los descubrimientos presentados en la presente memoria también pueden arrojar luz sobre los mecanismos moleculares subyacentes que impulsan la progresión de la enfermedad. Una característica notable de la firma de alto riesgo fue la significativa representación en exceso de los genes del cromosoma 1: casi 50% de los 19 genes poco expresados y 30% de los 51 genes expresados en exceso derivaban del cromosoma 1p y 1q, respectivamente. El predominio de los genes derivados del cromosoma 1q en la puntuación de alto riesgo está de acuerdo con un informe reciente que muestra que la progresión de la enfermedad está asociada no sólo con un aumento en el número de copias, sino también con el porcentaje de células con amplificación de 1q21 [15]. La firma de alto riesgo basada en la expresión de genes definida en la presente memoria también es notablemente compatible con una clase de enfermedad definida por un perfil de aCGH de alta resolución, que se caracteriza por la amplificación de alto nivel de 1q21 y la supresión de 1p13 [6]. Tomados en conjunto, estos

datos sugieren que las alteraciones en este cromosoma, a través de modificaciones genéticas y/o epigenéticas, pueden jugar un papel importante en la evolución de la enfermedad, proporcionando una ventaja en el crecimiento y/o la supervivencia.

5 Utilizando una combinación de aCGH de alta resolución y establecimiento de perfiles de micromatrices, se identificaron 47 regiones mínimas comunes (MCR) de ganancia genómica en todo el genoma de mieloma y 207 genes que se cartografiaban en estas MCR cuya expresión aumenta con el aumento en el número de copias [6]. Cuando la expresión de estos genes sensibles al número de copias se comparó entre las clases de alto y bajo riesgo definidas por el modelo de 70 genes, se encontró que los genes que se cartografiaban en las MCR en 1q21, 1q22 y 1q43-q44 eran significativamente expresados en exceso en la enfermedad de alto riesgo.

10 Aunque los genes del cromosoma 1 están implicados como actores clave en la progresión de la enfermedad, la residencia de los otros 4 genes, *FABP5*, *YWHAZ*, *EXOSC4*, y *EIFC2*, en la región 8q21-8q24 significa que las ganancias de 8q también pueden contribuir a la enfermedad de alto riesgo. Estos genes, abarcan las MCR recientemente definidas de ganancia/amplificación en 8q24.12-8q24.13 y 8q24.2-8q24.3 [6]. Curiosamente, la expresión de *MYC*, que se cartografía en una MCR en 8q24, no estaba ligada a la supervivencia en el estudio actual.

15 La delección en el cromosoma 13q14 es un importante predictor de la supervivencia en pacientes con mieloma tratados en pruebas de trasplante en tándem [24]. Es de destacar que la pérdida de expresión de un único gen cartografiado en el cromosoma 13q14, *RFP2*, previamente identificado como un gen supresor tumoral candidato en B-LLC con una homología significativa con *BRCA1* [25] fue una vez más vinculada a la escasa supervivencia en este análisis. También se encontró que *RFP2* mostraba una expresión sensible al número de copias en el mieloma [6]. La
20 amplificación frecuente del cromosoma 1 en muchos tipos de cáncer en etapas avanzadas, incluyendo 1q21 en el linfoma no Hodgkiniano, tumor de Wilms, sarcoma de Ewing, cáncer de mama y de ovario [12,26-30] garantiza estudios para determinar que el modelo de expresión del gen descrito aquí tiene relevancia pronóstica en otros tipos de cáncer. Aunque la presente invención ha utilizado el establecimiento de perfiles de expresión génica para identificar firmas genómicas de mieloma múltiple, el método y el kit proporcionados en la presente memoria pueden ser utilizados para identificar los mismos genes o diferentes genes que son predictivos de los resultados en otros
25 tipos de cáncer. Por la misma razón, los genes identificados en la presente memoria pueden ser predictivos de los resultados en otros tipos de cáncer.

A través del análisis discriminante multivariante, de los 70 genes originales, se pudieron utilizar 17 conjuntos de sondas para detectar el mieloma de alto riesgo. Por lo tanto, la presente invención contempla el desarrollo y la
30 validación de un análisis basado en RT-PCR cuantitativa que combina estos genes asociados con la estadificación/riesgo con genes vinculados con el subtipo molecular/etiología identificados en la clasificación molecular sin supervisión. La evaluación de los niveles de expresión de estos genes puede proporcionar un ensayo de pronóstico de base molecular simple y poderoso que eliminaría la necesidad de someter a ensayo muchas de las variables convencionales actualmente en uso con implicaciones pronósticas limitadas desprovistas también de
35 dianas idóneas para fármacos. El uso de una metodología basada en la PCR no solamente reduciría drásticamente el tiempo y el esfuerzo invertidos en los análisis basados en la hibridación de fluorescencia in situ sino que también reducirían notablemente la cantidad de tejido necesaria para el análisis. Si estas firmas de genes son únicas para las células tumorales de mieloma, semejante ensayo puede ser útil después del tratamiento para evaluar la enfermedad residual mínima, posiblemente utilizando sangre periférica como fuente de muestra.

40 Además, la presente invención también aplica el modelo de 17 genes para predecir el resultado de recaída de la enfermedad tratada con agente único, Bortezomib. Como este modelo fue desarrollado originalmente para los pacientes con mieloma tratados con quimioterapia con múltiples agentes seguida de HDT apoyada por trasplante y mantenimiento, el resultado discutido en la presente memoria fue algo inesperado. Estos datos sugieren fuertemente que los patrones de expresión de genes relacionados con los resultados son similares en pacientes con MM en
45 recaída y recién diagnosticados. El modelo de expresión de genes de alto riesgo también parecía ser independiente de la modalidad terapéutica específica, sugiriendo que identifica los pacientes que son generalmente insensibles a los fármacos o combinaciones de fármacos actualmente utilizados. Esta interpretación es apoyada por la observación de que el clasificador de supervivencia previamente validado desarrollado con un subconjunto de pacientes con mieloma en recaída tratados con bortezomib [34] también identificaba pacientes de alto riesgo
50 tratados con Terapia Total 2 (datos no mostrados).

Más allá de las diferencias en la plataforma de micromatrices, otras varias características del diseño del estudio distinguen estos 2 conjuntos de datos clinicogenómicos de MM. Notablemente, las muestras Millennium se obtuvieron en múltiples centros en oposición a un solo centro en el estudio UAMS. UAMS purificó células plasmáticas por la
55 selección de esferas inmunomagnéticas basada en CD138, mientras que los 96 centros clínicos participantes en el estudio Millennium realizaron cada uno un método de selección negativa antes de enviar los especímenes de tumor congelados.

Dos implicaciones importantes se derivan de estas observaciones. En primer lugar, como los patrones de expresión génica variados a menudo representan diferentes estados biológicos subyacentes de tejidos normales [35] y transformados [35, 36, 37], parece probable que la firma de alto riesgo esté relacionada con un fenotipo biológico de
60 resistencia a los fármacos y/o rápida recaída en el mieloma múltiple. Por consiguiente, este fenotipo de mieloma

merece estudios adicionales con el fin de caracterizar mejor las vías más relevantes e identificar las oportunidades terapéuticas. Los conjuntos de datos de expresión génica relativamente grandes empleados aquí proporcionan una vía para definir con más detalle estos tipos de tumores. En segundo lugar, si bien siguen existiendo algunos obstáculos para la aplicación clínica de rutina de la estratificación de alto riesgo, este trabajo pone de manifiesto que un subconjunto específico de pacientes con mieloma continúa recibiendo un beneficio mínimo de las terapias actuales. Un método práctico para identificar estos pacientes debe mejorar notablemente la atención al paciente. Para los pacientes que se pronostica que tienen un resultado favorable, podrían estar indicados los esfuerzos para minimizar la toxicidad de la terapia convencional, mientras que aquellos que se pronostica que tiene un mal resultado, independientemente de la terapia actual utilizada, pueden ser considerados para la administración temprana de regímenes experimentales. La presente invención contempla la determinación de si este modelo GEP de tumor de alto riesgo podría ser implementado clínicamente y si sería relevante para otros regímenes de primera línea, incluyendo aquellos que someten a ensayo nuevas combinaciones de inhibidores del proteasoma y/o IMiD con agentes anti-mieloma convencionales y HDT. Por último, los análisis anteriores de la respuesta a la terapia sugirieron que un clasificador de la respuesta a bortezomib era específico para la terapia con bortezomib como agente único [34]; se necesitan análisis adicionales de estos 2 conjuntos de datos para aclarar las rutas biológicas asociadas a la actividad de las tres terapias, así como las rutas vinculadas específicamente a la Terapia Total, el bortezomib como agente único o la dexametasona como agente único.

En una realización de la presente invención, se proporciona un método de elaboración de perfiles de la expresión génica para identificar firmas genómicas vinculadas a un pronóstico de supervivencia en un sujeto con mieloma múltiple sintomático o mieloma múltiple, que comprende la determinación de los niveles de expresión de genes que consisten en *KIF14*, *SLC19A1*, *CKS1B*, *YWHAZ*, *MPHOSPH1*, *TMPO*, *NADK*, *LARS2*, *TBRG4*, *AIM2*, *ASPM*, *AHCYL1*, *CTBS*, *MCLC*, *LTBP1* y genes detectables con ácidos nucleicos que tienen los números de referencia de Affymetrix 242488_at y 1557277-9_at en una micromatriz Affymetrix U133PLW2.0 y una muestra de ácido nucleico de células plasmáticas aisladas del sujeto, y la determinación de la presencia de niveles anormales de expresión, elaborando de este modo un pronóstico para el sujeto.

Semejante método puede comprender adicionalmente la aplicación de un análisis discriminante por pasos multivariante (MSDA) a través de las firmas genómicas, donde la aplicación identifica los 17 genes vinculados a al menos uno de supervivencia o capacidad de discriminación de enfermedad de alto riesgo y de bajo riesgo.

Adicionalmente, una razón media elevada de expresión obtenida utilizando este método puede ser indicativa de una firma genómica asociada con la enfermedad de alto riesgo. Semejante firma genómica de la enfermedad de alto riesgo se puede correlacionar con una menor duración de las remisiones completas, carencia de eventos, muerte relacionada con la enfermedad temprana o una combinación de las mismas. Adicionalmente, un individuo que lleva la firma genómica de la enfermedad de alto riesgo puede ser seleccionado para las pruebas de prevención secundaria. Alternativamente, una razón media baja de expresión puede ser indicativa de una firma genómica asociada con una enfermedad de bajo riesgo. Tal firma genómica de enfermedad de bajo riesgo se puede correlacionar con una mayor duración de las remisiones completas, mayor supervivencia, un buen pronóstico o una combinación de los mismos.

Además, el método descrito en la presente memoria puede pronosticar el resultado clínico y la supervivencia de un individuo, puede ser eficaz en la selección del tratamiento para un individuo que padece una enfermedad, puede pronosticar el riesgo de recaída después del tratamiento y la supervivencia de un individuo, puede correlacionar la clasificación molecular de una enfermedad con la firma genómica que define los grupos de riesgo, o una combinación de los mismos. La clasificación molecular puede ser CD1 y se puede correlacionar con la firma genómica de mieloma múltiple de alto riesgo. La clasificación CD1 puede comprender un aumento de la expresión de las firmas MMSET, MAF/MAFB, PROLIFERATION o una combinación de las mismas. Por otra parte, la clasificación molecular puede ser CD2 y se puede correlacionar con la firma de mieloma múltiple de bajo riesgo. La clasificación CD2 puede comprender las translocaciones HYPERDIPLOIDY, LOW BONE DISEASE, CCND1/CCND3, la expresión de CD20 o una combinación de los mismos. Adicionalmente, el tipo de enfermedad cuya firma genómica se identifica utilizando semejante método puede incluir, pero no se limita a mieloma múltiple sintomático o mieloma múltiple.

Según se utiliza en la presente memoria, el término "un", "una" o "uno" puede significar uno o más. Según se utiliza en la presente memoria en las reivindicaciones, cuando se utilizan junto con la palabra "que comprende", las palabras "un", "una" o "uno" pueden significar uno o más de uno. Según se utiliza en la presente memoria "otro" u "otra" pueden significar al menos un segundo o más del mismo o diferente elemento de la reivindicación o componentes de la misma.

Los siguientes ejemplos se proporcionan con el propósito de ilustrar diversas realizaciones de la invención y no están destinados a limitar la presente invención de ninguna manera.

Ejemplo 1

Pacientes

Se obtuvieron células plasmáticas purificadas de sujetos sanos normales y de pacientes con gammapatía monoclonal de significado incierto (MGUS) y con mieloma manifiesto que requería tratamiento. Las características de los pacientes de los grupos de formación (n = 351) y de validación (n = 181) se han descrito anteriormente⁹. De 351 casos en el grupo de formación, 51 también tenían muestras tomadas en la recaída. Ambos protocolos utilizaron regímenes de inducción, seguidos de autotrasplantes en tándem basados en melfalán, quimioterapia de consolidación y tratamiento de mantenimiento.

Ejemplo 2

Elaboración de perfiles de expresión génica

Las purificaciones de células plasmáticas y GEP, utilizando la micromatriz Affymetrix U133Plus2.0, se llevaron a cabo como se ha descrito anteriormente [9,16]. Los datos de las micromatrices y los datos de resultados en los 532 pacientes utilizados en este estudio han sido depositados en el NIH Gene Expression Omnibus con el número de acceso GSE2658.

Ejemplo 3

Análisis estadísticos y de micromatrices

Se procesaron previamente micromatrices U133Plus2.0 Affymetrix con el soporte lógico GCOS1.1 y se normalizaron mediante escalamiento GCOS1.1 convencional. Las pruebas de rango logarítmico para la asociación univariante con la supervivencia relacionada con la enfermedad se realizaron para cada uno de los 54.675 resúmenes "Señal". En concreto, se realizaron pruebas de rango logarítmico para el cuartil 1 (Q1) vs. cuartiles 2-4 (Q2-Q4) y el cuartil 4 (Q4) vs. cuartiles 1-3 (Q1-Q3) con el fin de identificar los genes de pronóstico expresados en exceso y poco expresados, respectivamente. Se aplicó un corte para la tasa de falsos descubrimiento de 2,5% a cada lista de valores P de rangos logarítmicos [17] obteniéndose 19 conjuntos de sondas poco expresados y 51 expresados en exceso. Los dendrogramas del mapa de calor de columnas se calcularon con agrupamiento jerárquico utilizando distancias de correlación de Pearson entre la expresión en la escala \log_2 de pares de pacientes. Las ramas del dendrograma de columnas se ordenaron de izquierda a derecha basándose en la diferencia de cada paciente entre el \log_2 promedio de objetos y obtiene los fines y ventajas mencionados, así como los objetos, fines y ventajas inherentes a la presente memoria.

Ejemplo 1

Pacientes

Se obtuvieron células plasmáticas purificadas de sujetos sanos normales y de pacientes con gammapatía monoclonal de significado incierto (MGUS) y con mieloma manifiesto que requería tratamiento. Las características de los pacientes de los grupos de formación (n = 351) y de validación (n = 181) se han descrito anteriormente⁹. De 351 casos en el grupo de formación, 51 también tenían muestras tomadas en la recaída. Ambos protocolos utilizaron regímenes de inducción, seguido de autotrasplantes en tándem basados en melfalán, quimioterapia de consolidación y tratamiento de mantenimiento.

Ejemplo 2

Elaboración de perfiles de expresión génica

Las purificaciones de células plasmáticas y GEP, utilizando la micromatriz Affymetrix U133Plus2.0, se llevaron a cabo como se ha descrito anteriormente [9,16]. Los datos de las micromatrices y los datos de resultados en los 532 pacientes utilizados en este estudio han sido depositados en el NIH Gene Expression Omnibus con el número de acceso GSE2658.

Ejemplo 3

Análisis estadísticos y de micromatrices

Se procesaron previamente micromatrices U133Plus2.0 Affymetrix con el soporte lógico GCOS1.1 y se normalizaron mediante escalamiento GCOS1.1 convencional. Las pruebas de rango logarítmico para la asociación univariante con la supervivencia relacionada con la enfermedad se realizaron para cada uno de los 54.675 resúmenes "Señal". En concreto, se realizaron pruebas de rango logarítmico para el cuartil 1 (Q1) vs. cuartiles 2-4 (Q2-Q4) y el cuartil 4 (Q4) vs. cuartiles 1-3 (Q1-Q3) con el fin de identificar los genes de pronóstico expresados en exceso y poco expresados, respectivamente. Se aplicó un corte para la tasa de falsos descubrimiento de 2,5% a cada lista de valores P de rangos logarítmicos [17] obteniéndose 19 conjuntos de sondas poco expresadas y 51 expresadas en exceso. Los dendrogramas del mapa de calor de columnas se calcularon con agrupamiento jerárquico utilizando distancias de correlación de Pearson entre la expresión en la escala \log_2 de pares de pacientes. Las ramas del dendrograma de columnas se ordenaron de izquierda a derecha basándose en la diferencia de cada paciente entre la expresión en la escala \log_2 promedio de los 51 genes regulados al alza y los 19 regulados a la baja: esta diferencia se interpreta como una razón media (es decir, la media geométrica) regulada al alza/a la baja en la escala \log_2 . Este resumen

simple, univariante del perfil de expresión de 70 genes para cada paciente puede mejorar la robustez de los efectos residuales de la matriz (es decir, después del procesamiento MAS5.0) que aumentan o disminuyen los 70 genes multiplicativamente, y también es independiente del factor de escala MAS5.0. La ponderación de la expresión por razones de riesgo, no normalizadas o normalizadas (es decir, estadística de Wald), no mejora esta puntuación, y el diseño fue utilizar ausencia de supervisión de la supervivencia global (SG) o la supervivencia libre de eventos (SLE) más allá de las pruebas de rango logarítmico de gen a gen. La razón media regulada al alza/a la baja en \log_2 se agrupó a continuación utilizando las K-medias en 3 grupos para separar la pequeña moda del extremo derecho en el histograma: los dos grupos con razones medias al alza/a la baja inferiores se combinaron. La moda del único extremo en la razón de expresión media al alza/a la baja es compatible con las pruebas de rango logarítmico del cuartil extremo utilizadas en el análisis de la expresión diferencial, aunque los histogramas y el lado derecho de los mapas de calor sugieren que el grupo de pacientes extremo es menor de 25%, más cercano a 13%.

Obsérvese que diferentes algoritmos de agrupamiento y números de grupos generan grupos de alta razón media entre 12% y 29% de los pacientes: se eligieron K-medias (con $K = 3$) ya que era mejor (es decir, entre los algoritmos simples para la razón \log_2 univariante) para separar la pequeña moda del lado derecho de la distribución más grande. Cualquier captación de corte univariante entre 10% y 30% de los pacientes es significativa para la SG en el conjunto de formación de 351 pacientes. En el conjunto de validación de 181 pacientes, se llevó a cabo un agrupamiento de K-medias independientemente para producir un corte independiente para la razón \log_2 alta versus baja. La aplicación del corte del conjunto de formación en el conjunto de validación proporciona un error de clasificación validado independientemente de 1,7% (es decir, 3 pacientes en el conjunto de validación de bajo riesgo se clasifican como de alto riesgo). Se presentó una validación temprana basándose de una cohorte independiente tratada bajo un nuevo protocolo con el fin de ilustrar y proporcionar fuerte evidencia de apoyo para la asociación de la razón media de 70 genes regulados al alza/a la baja con la supervivencia global. El punto de corte de alto riesgo para la razón media debe estar asociado con la supervivencia en general en pacientes recién diagnosticados, independientemente del protocolo, por lo que la diferencia de protocolo para el conjunto de validación refuerza la evidencia en lugar de debilitarla. La razón media también puede estar asociada con los resultados en pacientes previamente tratados, se requerirían sin embargo nuevos puntos de corte para la razón para definir un grupo de alto riesgo. Una advertencia importante es que los 70 genes no son particularmente adecuados para explicar el resultado entre los dos tercios inferiores de pacientes (clasificados por la razón media): esto es compatible con los escrutinios de rangos logarítmicos originales que aglutinaron 75% de los pacientes en un solo grupo para las pruebas de rango logarítmico Q1 y Q4: estos genes identifican las células plasmáticas de mieloma más agresivas, mediante diseño.

Para determinar la ubicación exacta del mapa del genoma y el orden de los conjuntos de sondas en la micromatriz Affymetrix U133Plus2.0, se desarrolló un soporte lógico para consultar automáticamente el motor de búsqueda NCBI (<http://www.ncbi.nlm.nih.gov/entrez>) para todos los sitios de inicio y final de los genes. La ubicación de cada conjunto de sondas se comparó con su correspondiente gen o punto de inicio del transcrito y se alineó del telómero del brazo p al telómero del brazo q. De esta manera se adjudicó una posición exacta en el cromosoma a más de 98% (53.581 de 54.675) conjuntos de sondas. El soporte lógico utilizado para la cartografía se puede encontrar en (<http://lambertlab.uams.edu/software>).

Las distribuciones de la supervivencia libre de eventos, global y la duración de la remisión completa (fecha de inicio de la respuesta completa) se calcularon utilizando el método de Kaplan-Meier, [18] y se utilizaron estadísticas de rangos log para probar su igualdad en todos los grupos [19]. Se utilizaron pruebas de la Chi-cuadrado y pruebas exactas de Fisher para probar la independencia de las categorías. Se calcularon análisis de riesgos proporcionales multivariantes, ajustados los efectos de los predictores y las proporciones de heterogeneidad observadas explicadas por los predictores combinados, es decir, R^2 [20]. La Tabla 5 resume un análisis de regresión lineal multivariante de la razón de regulación al alza/a la baja en la escala \log_2 . Para este análisis se utilizó el paquete estadístico R versión 2.0.1 [21].

Se utilizó un análisis discriminante lineal múltiple por pasos (MSDA) con el criterio de la lambda de Wilk, [22] para seleccionar un subconjunto de los 70 genes igualmente capaces de diferenciar MM de alto riesgo y de bajo riesgo. El MSDA seleccionó la siguiente ecuación: Puntuación discriminante = $200\ 638_s_at \times 0,283 - 1\ 557\ 277_a_at \times 0,296 - 200\ 850_s_at \times 0,208 + 201\ 897_s_at \times 0,314 - 202\ 729_s_at \times 0,287 + 203\ 432_at \times 0,251 + 204\ 016_at \times 0,193 + 205\ 235_s_at \times 0,269 + 206\ 364_at \times 0,375 + 206\ 513_at \times 0,158 + 211\ 576_s_at \times 0,316 + 213\ 607_at \times 0,232 - 213\ 628_at \times 0,251 - 218\ 924_s_at \times 0,230 - 219\ 918_s_at \times 0,402 + 220\ 789_s_at \times 0,191 + 242\ 488_at \times 0,148$ (en donde las variables representan el valor de Affymetrix para la sonda concreta). El valor de corte fue de 1,5, de tal manera que los valores inferiores a 1,5 indicaron que la muestra pertenecía al grupo de bajo riesgo y los valores superiores a 1,5 indicaron que la muestra pertenecía al grupo de MM de alto riesgo. Se llevaron a cabo ambas selecciones de variables de avance y retroceso. La elección para introducir o eliminar variables se basó en la minimización de la variabilidad dentro del grupo con respecto a la variabilidad total a través de todas las muestras.

Ejemplo 4

Los patrones de expresión génica son un predictor independiente de supervivencia en el mieloma

Para identificar una firma molecular distintiva de mieloma de alto riesgo, la muerte relacionada con la enfermedad temprana se correlacionó con los extremos de la expresión génica. Los niveles de expresión génica a partir de los

datos de las micromatrices sobre las células plasmáticas seleccionadas con CD138 de 351 pacientes recién diagnosticados se dividieron en cuartiles, y se utilizaron pruebas de rango log para identificar 70 genes que estaban vinculados a la corta supervivencia: 51 tenían alta expresión (cuartil 4, Q4) y 19 tenían baja expresión (cuartil 1, Q1) (**Tabla 1**), cuyos niveles de expresión se representan en un colorgrama (**Figura 1A**). Cabe destacar la regulación al alza de los 51 genes y regulación a la baja de los 19 genes simultáneas entre los pacientes en el lado derecho. Se calculó la diferencia entre el promedio de la expresión a escala \log_2 de Q4 y Q1 para cada paciente. Este resumen de expresión no supervisada es interpretable como una razón de expresión media regulada al alza vs. a la baja a escala \log_2 (referida como una puntuación de riesgo). Su distribución de frecuencias revela un grupo distinto que tiene elevadas razones de regulación al alza/a la baja en \log_2 (**Figura 1B**). Este es precisamente el tipo de grupo de expresión extrema para cuyo escrutinio fueron diseñadas las pruebas de rango logarítmico Q1 y Q4, aunque tanto el gráfico de frecuencia como el mapa de calor sugieren que el tamaño del grupo es menor de 25%. El agrupamiento de K-medias sin supervisión de la razón \log_2 estimó su proporción en 13,4%. Este grupo mostró supervivencia libre de eventos significativamente más pobre (**Figura 1C**, $P < 0,001$) con una razón de riesgo no ajustada (HR) de 4,51 y también supervivencia global inferior (**Figura 1D**, $P < 0,001$), con una HR ajustada de 5,16. Se esperan asociaciones significativas para la cohorte de formación, en las que se descubrieron los 70 genes, y son referidas con fines ilustrativos.

El resultado de muerte temprana relacionada con enfermedad se escogió específicamente con el propósito de identificar los genes diana en el mieloma agresivo y, por consiguiente, sólo estuvieron disponibles 24 muertes para las pruebas de rangos logarítmicos utilizadas para el descubrimiento de genes en la cohorte original de 351 pacientes. El agrupamiento supervisado con los 70 genes se aplicó a las células plasmáticas de 22 donantes sanos, 14 casos de MGUS, 351 pacientes de la cohorte de formación y 38 líneas celulares de mieloma humano. Los resultados revelaron que el grupo de bajo riesgo de mieloma tenía un patrón similar al de MGUS y las células plasmáticas normales, mientras que el grupo de alto riesgo exhibe un patrón similar al de líneas celulares de mieloma humano (**Figura 2**).

A continuación, se examinó la asociación de la firma de expresión con la supervivencia global en una cohorte de ensayo independiente de 181 pacientes. De hecho, un agrupamiento no supervisado, independiente de la razón de expresión regulada al alza/a la baja a escala \log_2 identificó un subconjunto proporcionalmente similar de pacientes que mostraban una desregulación extrema (12,2%, **Figura 3A**). Un resultado similar de distribución de supervivencia y razón de riesgo se encontró tanto en la supervivencia libre de eventos (HR = 3,41, $P = 0,002$, **Figura 3B**) como en la supervivencia global (HR = 4,75, $P < 0,001$, **Figura 3C**) como se observa en la cohorte de formación. La ausencia de una puntuación de alto riesgo identificó un subconjunto favorable de pacientes con una remisión completa continua a los 5 años de 60%, frente a una tasa a los 3 años de sólo 20% en aquellos con una puntuación de alto riesgo (datos no mostrados).

Para evaluar adicionalmente la validez de los agrupamientos con respecto a las características clínicas, se analizaron las correlaciones de varios parámetros clínicos entre los subgrupos de bajo y de alto riesgo en los conjuntos tanto de formación (**Tabla 2**) como de ensayo (**Tabla 3**). Se observó una notable similitud de la distribución del rasgo clínico en los grupos de riesgo en las cohortes de formación y de ensayo: los mayores niveles séricos de β_2 -microglobulina, proteína C reactiva, creatinina y lactato deshidrogenasa (LDH), así como la delección del cromosoma 13 definida mediante FISH y las anomalías citogenéticas en metafase fueron todos significativamente más frecuentes en el grupo de alto riesgo de los conjuntos de formación y de ensayo ($P < 0,05$). Del mismo modo, el subgrupo *CCND1* clínicamente más benigno predominó en el subgrupo de bajo riesgo y el subgrupo *MMSET/FGFR3* de la cohorte de alto riesgo, como se representa para el conjunto de formación en la **Tabla 2** y para el conjunto de ensayo en la **Tabla 3**.

En un análisis multivariante de las variables asociadas con la supervivencia global y libre de eventos, el predictor de la razón de regulación al alza/a la baja alto (puntuación de alto riesgo) conservó su significación después del ajuste de las variables genéticas y clínicas competitivas (incluso incluyendo el Sistema Internacional de Estadificación) tanto en el conjunto de formación (**Tabla 4**: HR = 4,1; $P < 0,001$) como en el conjunto de ensayo (datos no mostrados, $P = 0,025$). Es importante destacar que la puntuación de alto riesgo también fue el único parámetro de referencia independiente que afectó negativamente a la duración de respuesta completa (razón de riesgo, 3,07; $P < 0,001$). Este fuerte resultado pronóstico de la puntuación de riesgo derivada de GEP puede explicarse en parte por su fuerte asociación con las variables de pronóstico clínico conocidas, como lo demuestra un análisis multivariante con la razón de regulación al alza/a la baja como resultado (**Tabla 5**). Si bien las variables de la Tabla 5 pueden servir como sustitutos temporales, parciales para un ensayo de GEP ampliamente disponible, la Tabla 4 sugiere que un análisis de este tipo, combinado con translocaciones de alto riesgo (también cuantificables a través de GEP), tiene el potencial de proporcionar un ensayo de pronóstico sencillo poderoso para el mieloma.

Tabla 1: Lista de genes que comprenden la firma de alto riesgo de 70 genes

Rango (Q4)	Cromosoma	Conjunto sondas Affymetrix	Símbolo
1	8q21.13	202345_s_at	<i>FABP5</i>
2	Xp22.12	1555864_s_at	<i>PDHA1</i>
3	5p15.33	204033_at	<i>TRIP13</i>
4	11q22	206513_at	<i>AIM2</i>
5	2p24.1	1555274_a_at	<i>SELI</i>
6	21q22.3	211576_s_at	<i>SLC19A1</i>
7	3p21.3	204016_at	<i>LARS2</i>
8	17q43	1565951_s_at	<i>OPN3</i>
9	1q31.3	219918_s_at	<i>ASPM</i>
10	12q15	201947_s_at	<i>CCT2</i>
11	16p13.3	213535_s_at	<i>UBE2I</i>
12	20q13.31	204092_s_at	<i>STK6</i>
13	1p36.33	213607_x_at	<i>FLJ13052</i>
14	Xq12	208117_s_at	<i>LAS1L</i>
15	17q25	210334_x_at	<i>BIRC5</i>
16	3q27	204023_at	<i>RFC4</i>
17	1q21.2	201897_s_at	<i>CKS1B</i>
18	19q13.12	216194_s_at	<i>CKAP1</i>
19	1p11	225834_at	<i>MGC57827</i>
20	19q13.12	238952_x_at	<i>DKFZp779O175</i>
21	17p13.3	200634_at	<i>PFN1</i>
22	19p13.2	208901_s_at	<i>ILF3</i>
23	1q22	206332_s_at	<i>IFI16</i>
24	7p13	220789_s_at	<i>TBRG4</i>
25	10p11.23	218947_s_at	<i>PAPD1</i>
26	8q24	213310_at	<i>EIF2C2</i>
27	3q12.1	224523_s_at	<i>MGC4308</i>
28	1p36.13	201231_s_at	<i>ENO1</i>
29	18q12.1	217901_at	<i>DSG2</i>
30	6q22	226936_at	<i>C6orf173</i>

ES 2 614 494 T3

31	8q24.3	58696_at	<i>EXOSC4</i>
32	1q23.3	200916_at	<i>TAGLN2</i>
33	3q21	201614_s_at	<i>RUVBL1</i>
34	16p11.2	200966_x_at	<i>ALDOA</i>
35	2p25.1	225082_at	<i>CPSF3</i>
36	1q43	242488_at	<i>NA</i>
37	3q12.3	243011_at	<i>MGC15606</i>
38	22q13.1	201105_at	<i>LGALS1</i>
39	3p25.3	224200_s_at	<i>RAD18</i>
40	20p11	222417_s_at	<i>SNX5</i>
41	1q21.2	210460_s_at	<i>PSMD4</i>
42	12q24.3	200750_s_at	<i>RAN</i>
43	1q32.1	206364_at	<i>KIF14</i>
44	7p15.2	201091_s_at	<i>CBX3</i>
45	12q22	203432_at	<i>TMPO</i>
46	17q24.2	221970_s_at	<i>DKFZP586L0724</i>
47	11p15.4	212533_at	<i>WEE1</i>
48	3p12	213194_at	<i>ROBO1</i>
49	5q33.1	244686_at	<i>TCOF1</i>
50	8q23.1	200638_s_at	<i>YWHAZ</i>
51	10q23.31	205235_s_at	<i>MPHOSPH1</i>

Problema (ID)	Cromosoma	Conjunto Sondas Affymetrix	Símbolo Gen
1	9q31.3	201921_at	<i>GNG10</i>
2	11p13	227278_at	<i>NA</i>
3	Xp22.3	209740_s_at	<i>PNPLA4</i>
4	20q11.21	227547_at	<i>NA</i>
5	10q25.1	225582_at	<i>KIAA1754</i>
6	1p13.2	200850_s_at	<i>AHCYL1</i>
7	1p13.3	213628_at	<i>MCLC</i>
8	11p22	209717_at	<i>EVI5</i>
9	1p13.3	222495_at	<i>AD-020</i>

10	6p21.31	1557277_at	NA
11	1p22.1	1554736_at	PARG1
12	1p22	218924_s_at	CTBS
13	9p13.2	226954_at	UBE2R2
14	1p34	202838_at	FUCA1
15	13q14	230192_at	RFP2
16	12q13.11	48106_at	FLJ20489
17	11q13.1	237964_at	NA
18	2p22.3	202729_s_at	LTBP1
19	1p13.1	212435_at	TRIM33

Tabla 2: Correlación de parámetros clínicos con grupos de riesgo en la cohorte de formación (n=351)

Característica	% Bajo Riesgo	% Alto Riesgo	P
Edad \geq 65 años	20	20	0,856
Albúmina \geq 3,5 g/dL	13	35	0,001
β_2 -microglobulina			
< 3,5 mg/L	62	42	0,005
\geq 3,5 y < 5,5 mg/L	20	20	
\geq 5,5 mg/L	19	40	
Proteína C reactiva \geq 4 mg/L	51	62	0,235
LDH $>$ 190 U/L	30	59	< 0,001
del 13 definida por FISH de la interfase	31	49	0,031
Anomalías citogenéticas	26	70	< 0,001
Translocaciones basadas en GEP			
CCND1	20	0	< 0,001
MMSET	12	28	
MMSE/MAFB	3	9	
Sin pico	65	63	

Tabla 3

Correlación de parámetros clínicos con los grupos de riesgo en la cohorte de formación (n=181)

Característica	% Bajo riesgo	% Alto riesgo	P
Edad \geq 65 años	30	23	0,692
Albumina \geq 3,5 g/dL	17	32	0,163
β_2 -microglobulina			
< 3,5 mg/L	57	32	0,005
\geq 3,5 y < 5,5 mg/L	23	18	
\geq 5 mg/L	19	50	
Proteína C reactiva \geq 4 mg/L	44	59	0,271
LDH \geq 190 IU/L	18	59	0,000
Anomalías citogenéticas	27	77	0,000
Translocaciones basadas en GEP			
CCND1	14	0	< 0,001
MYC/IRF	12	23	
MAF/MAFB	7	36	
Sin pico	67	41	

Tabla 4: Análisis multivariante de supervivencia libre de eventos y global en la cohorte de formación

Predictores significativos *	Supervivencia libre eventos			Supervivencia global	
	%	HR	P	HR	P
Regulado al alza/a la baja de alto riesgo	13	3,24	< 0,001	4,09	< 0,001
Razón de expresión (escala Log ₂)*					
Beta-2-microglobulina \geq 3,5 mg/L	20	1,72	0,001	--	--
Beta-2-microglobulina < 5,5 mg/L					
Beta-2-microglobulina \geq 5,5 mg/L	21	2,01		--	--
LDH \geq 190 IU/L	34	--	--	1,92	0,004
del13 definida por FISH en interfase	33	1,63	0,007	--	--
Translocaciones de alto riesgo definidas por GEP [†]	18	1,97	0,001	1,85	0,0120
Eventos / Muertes	138			87	
R²	0,324			0,288	

[†] 26 de 351 pacientes perdían del13 definida por FISH.

[‡] Predictores con P > 0,05 para ambos resultados: Edad \geq 65, anomalías citogenéticas en la metafase, albúmina \leq 3,5 g/dL y proteína C reactiva \geq 4. Los guiones indican la ausencia de significación para uno u otro resultado.

* log₂ promedio (expresión) de los 51 genes Q4 menos el log₂ promedio (expresión) de los 19 genes Q1 (es decir la

razón de la escala log₂ de la media geométrica de los genes regulados al alza versus regulados a la baja). El alto riesgo por el agrupamiento de K-medias es $\geq 0,66$ (es decir, una razón de 1,58). Se combinan picos de MMSET/FGFR3 (14,1%) con picos de MAF/MAFB (3,7%). El bajo riesgo incluye picos de CCND1 (16,9%) y ausencia de picos (65,3%). Las categorías cerradas funcionan mejor como categorías de pronóstico debido a la similitud en la distribución de resultados para los subgrupos dentro de las categorías de Alto y Bajo riesgo y el pequeño tamaño del subgrupo MAF/MAFB.

Tabla 5: Análisis multivariante de multiplicidad de cambio en la razón de expresión regulada al alza/a la baja N=250 [†]

Predictores significativos	Multiplicidad		
	%	Cambio	Valor P
amp1q21 definido por FISH en interfase	43	0,316	< 0,001
Anomalías citogenéticas	30	0,353	<< 0,0001
aumento CCND1 o CCND3	20	-0,248	0,008
Aumento MAF/MAFB	4	0,430	0,030
Aumento MMSET/FGFR3	14	-0,297	0,005
LDH ≥ 190 U/L	31	0,332	< 0,001
Albúmina $\leq 3,5$ g/dL	18	0,249	0,014
R ²	0,324		

[†] 98 de los 351 pacientes habían perdido amp1q21 por medio de FISH y 3 más habían perdido albúmina.

[‡] Predictores con P > 0,05 para ambos resultados: Edad ≥ 65 , Beta-2-microglobulina ($\geq 3,5$, $\geq 5,5$), Proteína C reactiva (≥ 4).

Ejemplo 5

El modelo de expresión génica pronostica el riesgo posterior de recaída y supervivencia

Cuando se aplicó el modelo de riesgo de 70 genes a muestras de recaída de 51 de los 351 pacientes del conjunto de formación, 39 (76%) mostraron una puntuación de alto riesgo (Figura 4A). En un análisis pareado de muestras de referencia y de recaída, los 25 pacientes con denominación de bajo riesgo, tanto en el diagnóstico como en la recaída tuvieron una supervivencia post-recaída superior, seguido de 11 pacientes con designación de bajo riesgo en el momento del diagnóstico y de alto riesgo en la recaída y 13 pacientes que mostraron una designación de alto riesgo en ambos tiempos de observación (Figura 4B). Hubo sólo 2 casos con alto riesgo en el momento del diagnóstico y bajo riesgo en la recaída.

Ejemplo 6

Los genes del cromosoma 1 están representados en exceso en el modelo de alto riesgo

Para determinar si la firma de alto riesgo de 70 genes puede reflejar ganancias o pérdidas específicas de ADN genómico en el MM de alto riesgo, se compararon las posiciones en el mapa de los 70 genes que comprendían la firma de riesgo de expresión génica (Tabla 6). Si bien representaban sólo 10% de los genes en la micromatriz, 21 (30%) de los 70 genes de alto riesgo se cartografiaron en el cromosoma 1 (P <,0001): 9 de 19 (47%) genes del cuartil 1 se cartografiaron en 1p cartografiándose 5 en 1p13; 12 de 51 (24%) genes del cuartil 4 se cartografiaron en el cromosoma 1, 9 residían en 1q mientras que los 4 de 1p se cartografiaron en el extremo telomérico y las regiones centroméricas del brazo p. Estos datos sugieren que la ganancia de material de ADN en 1q y la pérdida de 1p son determinantes significativos de alto riesgo en el MM.

Tabla 6: Distribución en el cromosoma de todos los conjuntos de sondas cartografiados sobre la micromatriz U133Plus2,0 y los 70 genes de la firma de alto riesgo

Cromo soma	U133 Plus2.0		Q1		Q4		Combi- nados		P*
	Núm. Genes	%	Núm. Genes	%	Núm. Genes	%	Núm. Genes	%	
1	5.379	10	9	47,4	12	23,5	21	30	< 0,0001
2	3.958	7,3	1	5,3	2	3,9	3	4,3	
3	3.275	6,1	0	0	7	13,7	7	10	
4	2.314	4,3	0	0	0	0	0	0	
5	2.615	4,8	0	0	2	3,9	2	2,9	
6	2.956	5,5	1	5,3	1	2	2	2,9	
7	2.769	5,1	0	0	2	3,9	2	2,9	
8	2.014	3,7	0	0	4	7,8	4	5,7	
9	2.139	4	2	10,5	0	0	2	2,9	
10	2.192	4,1	1	5,3	2	3,9	3	4,3	
11	2.889	5,4	1	5,3	1	2	2	2,9	
12	2.739	5,1	1	5,3	3	5,9	4	5,7	
13	1.250	2,3	1	5,3	0	0	1	1,4	
14	1.793	3,3	0	0	0	0	0	0	
15	1.805	3,3	0	0	0	0	0	0	
16	2.084	3,9	0	0	2	3,9	2	2,9	
17	2.843	5,3	0	0	3	5,9	3	4,3	
18	966	1,8	0	0	1	2	1	1,4	
19	2.839	5,3	0	0	3	5,9	3	4,3	
20	1.487	2,8	1	5,3	2	3,9	3	4,3	
21	662	1,2	0	0	1	2	1	1,4	
22	1.275	2,3	0	0	1	2	1	1,4	
X	1.691	3,1	1	5,3	2	3,9	3	4,3	
Y	107	0,2	0	0	0	0	0	0	
	53.991		19		51		70		
Control Affy	62								
Desconocido	622								
Total	54.675								

* Se utilizó una prueba exacta para las proporciones binomiales para comparar la proporción de los conjuntos de sondas retenidos que se cartografiaron en el cromosoma 1 con la proporción para la matriz completa.

Ejemplo 7

Un modelo de 17 genes puede sustituir el modelo de 70 genes

Habiendo demostrado que el alto riesgo está probablemente relacionado con alteraciones genómicas del

5 cromosoma 1, se determinó un conjunto mínimo de genes capaces de discriminar el mieloma de alto riesgo y de bajo riesgo. Aplicando un análisis discriminante por pasos multivariante ("MSDA por sus siglas en inglés") de los 70 genes de alto riesgo asociados a través de los casos de alto riesgo (N = 46) y de bajo riesgo (N = 305) definidos por el modelo de 70 genes en el conjunto de formación, 17 genes fueron identificados en la función discriminante lineal resultante (**Tabla 7**). Cabe destacar que 3 de los 5 (60%) genes Q1 y 5 de los 12 (45%) genes Q4 en el modelo se cartografiaron en 1p y 1q, respectivamente. A continuación, el modelo de 17 genes se aplicó al grupo de formación y se pronosticó, con 97,7% de precisión, basándose la clase correcta en la clasificación de alto/bajo riesgo del modelo de 70 genes (**Tabla 8A**). Se realizó un análisis de validación cruzada donde las muestras se eliminaron de una en una del conjunto de muestras, y se volvió a calcular el modelo predictivo sin esa muestra. A continuación, se utilizó el modelo para clasificar la observación eliminada. En este enfoque de validación cruzada, la precisión de la predicción fue de 96,9%. A continuación, se aplicó el modelo de 17 genes al conjunto de ensayo de 181 pacientes recién diagnosticados que recibían el segundo protocolo UARK 03-033. El modelo MSDA clasificó de nuevo correctamente 150 de 159 (94,3%) muestras de bajo riesgo y 21 de 22 (95,5%) de alto riesgo (**Tabla 8B**). Las estimaciones de Kaplan-Meier de la supervivencia global de los grupos de alto riesgo y de bajo riesgo fueron similares ya fueran definidas por el modelo de 17 genes (**Figura 5**) o el modelo de 70 genes (**Figura 3D**).

Tabla 7: 17 Genes definidos por MSDA ordenados por su puntuación

Conjunto sondas Affymetrix	Símbolo gen	Localización cromosoma	Puntuación MSDA	Cuartil 70 genes
206364_at	<i>KIF14</i>	1q32.1	0,38	Q4
201576_s_at	<i>SLC19A1</i>	21q22.3	0,32	Q4
201897_s_at	<i>CKS1B</i>	1q21.2	0,31	Q4
200638_s_at	<i>YWHAZ</i>	8q23.1	0,28	Q4
205235_s_at	<i>MPHOSPH1</i>	10q23.31	0,27	Q4
203492_at	<i>TMPO</i>	12q22	0,25	Q4
213607_x_at	<i>NADK</i>	1p36.21	0,23	Q4
204016_at	<i>LARS2</i>	3p21.3	0,19	Q4
220789_s_at	<i>TBRG4</i>	7p14-p13	0,19	Q4
206519_at	<i>AIM2</i>	1q22	0,16	Q4
242488_at	<i>NA</i>	1q43	0,15	Q4
219918_s_at	<i>ASPM</i>	1q31	-0,40	Q4
200850_s_at	<i>AHCYL1</i>	1p13.2	-0,21	Q1
218924_s_at	<i>CTBS</i>	1p22	-0,23	Q1
213628_at	<i>MCLC</i>	1p13.3	-0,25	Q1
202729_s_at	<i>LTBP1</i>	2p22-p21	-0,29	Q1
1557277_a_at	<i>NA</i>	6p21	-0,30	Q1

Tabla 8A: Matriz de confusión de predicción del riesgo en el conjunto de formación utilizando el modelo de 17 genes

Grupo de Riesgo de 70 Genes	Grupo de Riesgo de 17 Genes		
	Total	Bajo	Alto
Bajo	305	298	7
Alto	46	1	45

Tabla. 8B Matriz de confusión de predicción del riesgo en el conjunto de formación utilizando el modelo de 17 genes

Grupo de Riesgo de 70 Genes	Grupo de Riesgo de 17 Genes		
	Total	Bajo Riesgo	Alto Riesgo
Bajo	159	150	9
Alto	22	1	21

Ejemplo 8

- 5 Mieloma de alto riesgo definido por modelo de 70 genes relacionado con subgrupos moleculares definidos mediante análisis de agrupamiento jerárquico no supervisado

El modelo de alto riesgo identificado fue examinado en el contexto de una clasificación molecular definida previamente [9]. La designación de enfermedad de alto riesgo se refería a todas las clases de mieloma a excepción de tipo CD-2 caracterizado por picos *CCND1* o *CCND3* y expresión de *CD20* y *VPREB3* (**Figura 6**). A pesar de una fuerte correlación entre la firma de alto riesgo y el subgrupo de Proliferación (PR) (**Figura 6**), la presencia de casos atípicos sugiere que la firma de alto riesgo no sólo refleja la proliferación de células tumorales, sino que puede abarcar también otras características de la enfermedad que confieren corta supervivencia tal como resistencia a los fármacos. El análisis de los 351 casos de formación de acuerdo con un punto de corte de alto riesgo de 70 genes de .66 y un índice de proliferación (IP) de 5 (**Figura 7A**) revelaron que las designaciones de alto y bajo IP no lograron identificar subgrupos con diferente supervivencia entre los grupos de bajo riesgo y de alto riesgo (**Figura 7B**). Cuando se aplicó a los 50 pacientes con t(4;14)(p16;q32), la puntuación de riesgo de 70 genes separó de nuevo subgrupos de bajo y alto riesgo ($P < 0,001$) (**Figura 7C**).

Ejemplo 9

- 20 Aplicación del modelo de 17 genes para pronosticar el resultado de recaída de la enfermedad tratada con agente único, Bortezomib

Para investigar si el modelo de 17 genes podría pronosticar el alto riesgo en el conjunto de datos Millenium, se reconstruyó el modelo de 17 genes (U-2) derivado de U133Plus2.0 utilizando los datos U133AB (U-AB). En pocas palabras, se ha descrito la micromatriz Affymetrix U133Plus2.0 (U2) (Affymetrix, Santa Clara, CA) en células plasmáticas seleccionadas mediante CD138 a partir de 351 casos nuevos diagnosticados de mieloma tratados con la terapia de alta dosis y apoyo de células madre (acceso GEO GSE2658) [33]. También se analizó exactamente la misma muestra de ARN de las primeras 144 de las 351 descritas anteriormente en la micromatriz U133A/B (UA) (Affymetrix, Santa Clara, CA) y estos datos se han depositado en GEO (GSE8991). Se han descrito los datos de UA sobre 156 pacientes con mieloma múltiple recurrente tratados con bortezomib o dexametasona en una prueba en fase III^{2,3} (Pendiente del número de acceso GEO).

- 30 Se utilizó un análisis discriminante lineal múltiple por pasos (MSDA) con el criterio lambda de Wilks utilizando datos derivados de U2 para definir el modelo de 17 genes predictivo de la enfermedad de alto y bajo riesgo [33]. Las distribuciones de supervivencia se presentaron con el uso del método de Kaplan-Meier y se compararon con la prueba de rango logarítmico. Las pruebas estadísticas se realizaron con el paquete estadístico SPSS 12.0 (SPSS, Chicago, IL).

35 De los 17 genes identificados utilizando la plataforma U2, 16 estaban en la micromatriz UA (**Tabla 9**). El modelo de análisis discriminante por pasos multivariante (MSDA) que se utiliza para desarrollar el modelo basado en U2 de 17 genes se aplicó a continuación a la intensidad de la señal de los 16 genes de UA en los 144 casos que tienen datos de ambos tipos de matriz. Mediante la correlación de las puntuaciones de riesgo resultantes derivadas de los modelos derivados de U2 y UA, este análisis reveló una fuerte correlación (**Figura 8A**; $r = 0,89$; $P < 0,001$) y demostró que una puntuación superior a 1,6 era de alto riesgo en ambos modelos. Utilizando 1,6 como punto de corte en ambos modelos, una matriz de confusión mostró una concordancia de 96,5% entre el alto y bajo riesgo definidos en U2 y UA (**Tabla 10**) de modo que sólo 5 pacientes difirieron en su análisis de supervivencia de Kaplan-Meier de asignación de riesgo (la versión UA del clasificador reasignó 2 pacientes de alto riesgo y 3 de bajo riesgo) de los grupos de alto y bajo riesgo definidos por el modelo U2 revelando una diferencia significativa en la supervivencia general entre los dos grupos con una razón de riesgo no ajustada (HR) de 2,49. (**Figura 9A**). Del mismo modo, el análisis de Kaplan-Meier de los grupos de riesgo definidos por UA reveló una diferencia significativa en la supervivencia entre los dos grupos ($P = 0,0026$) con un alto riesgo (HR) asociado de 2,58 (**Figura 9B**). Estos

resultados indican la capacidad de reconstruir el modelo de riesgo utilizando la plataforma UA.

Después, la versión UA se aplicó a continuación al conjunto de datos Millenium de 156 casos de recaída de enfermedad tratados con bortezomib o dexametasona. El modelo de alto riesgo definió 13,5% de los 156 pacientes APEX como de alto riesgo mediante el uso del modelo de 16 genes. Estos pacientes tenían tiempos de supervivencia significativamente más cortos que el resto de pacientes (**Figura 8B**, $P = 0,0014$, $HR = 2,5$). Este resultado indica que el modelo de alto riesgo fue relevante en el mieloma múltiple de recaída, así como en la enfermedad recién diagnosticada. Este resultado también sugiere que el modelo es relevante para los pacientes tratados con terapias de un solo agente, tales como bortezomib y dexametasona. Por lo tanto, cada uno de los brazos de tratamiento APEX (80 y 76 pacientes tratados con bortezomib y dexametasona, respectivamente) se examinó por separado para determinar si el modelo se veía afectado de manera significativa por el agente terapéutico específico. Como se muestra en las **Figuras 8C y 8D**, el modelo de alto riesgo identificó con éxito los pacientes con alto riesgo de muerte en ambos grupos ($P = 0,049$, $HR = 2,3$ y $P = 0,017$, $HR = 2,5$, respectivamente).

Tabla 9

Lista de genes que comprenden la firma de alto riesgo de 16 genes.		
Cromosoma	Conjunto de Sondas Affymetrix	Símbolo del Gen
1q22	206513_at	<i>AIM2</i>
21q22.3	211576_s_at	<i>SLC19A1</i>
3p21.3	204016_at	<i>LARS2</i>
1q31	219918_s_at	<i>ASPM</i>
1p36.33-p36.21	213607_x_at	<i>FLJ13052</i>
1q21.2	201897_s_at	<i>CKS1B</i>
7p14-p13	220789_s_at	<i>TBRG4</i>
1q43	242488_at	<i>NA</i>
1pter-q31.3	206364_at	<i>KIF14</i>
12q22	203432_at	<i>TMPO</i>
8q23.1	200638_s_at	<i>YWHAZ</i>
10q23.31	205235_s_at	<i>MPHOSPH1</i>
1p13.2	200850_s_at	<i>AHCYL1</i>
1p13.3	213628_at	<i>MCLC</i>
1p22	218924_s_at	<i>CTBS</i>
2p22-p21	202729_s_at	<i>LTBP1</i>

Tabla 10

Matriz de confusión de designaciones de subgrupos mediante los modelos de 17 y 16 genes en los 144 casos.				
Modelo	17 Genes	Alto Riesgo (N = 26)	Bajo riesgo (N = 118)	Exactitud (%)
	Alto riesgo	24	3	
16 Genes	Riesgo bajo	2	115	96,5

En la presente memoria se citan las siguientes referencias:

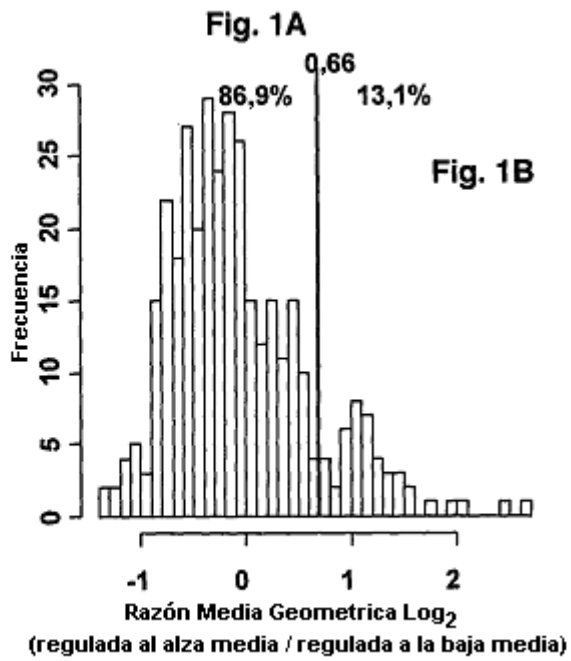
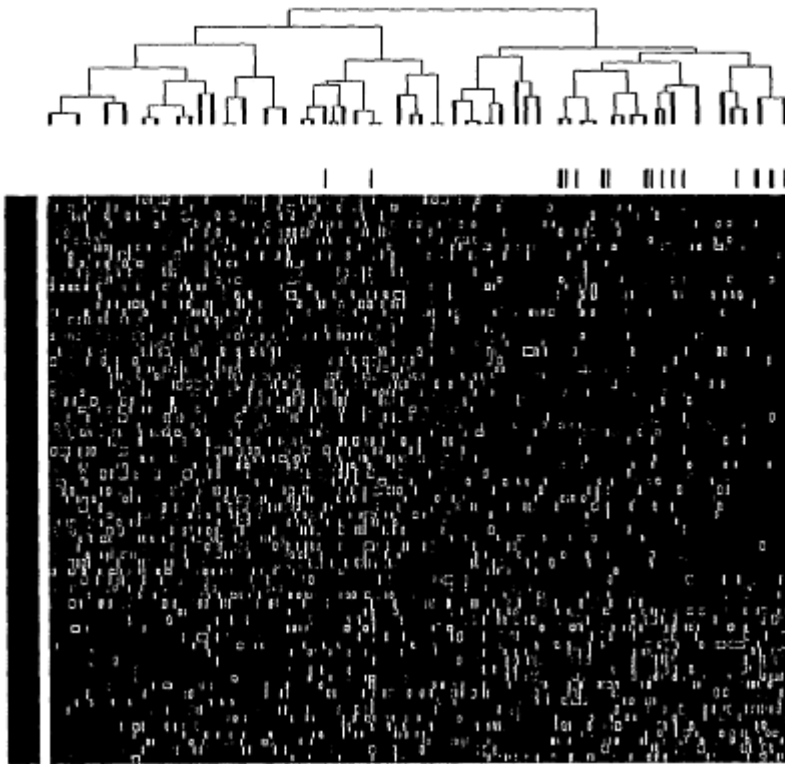
1. Smadja NV et al. *Leukemia*. 1998;12:960-969.
- 5 2. Wuilleme S et al. *Leukemia*. 2002;19:275-278.
3. Cremer FW et al. *Genes Chromosomes Cancer*. 2005;44:194-203.
4. Gutierrez NC et al. *Blood*. 2004;104:2661-2666.
5. Fonseca R et al. *Cancer Res*. 2004;64:1546-1558.
6. Carrasco D et al. *Cancer Cell*. 2006;9:313-325.
- 10 7. Shaughnessy J, Barlogie, B. *Immunol. Rev*. 2003;94:140-163.
8. Kuehl WM, Bergsagel PL *Nature Rev. Cancer*. 2002;2:175-187.
9. Zhan F et al. *Blood*. 2006;108:2020-2028.
10. Avet-Loiseau et al. *Genes Chromosomes Cancer*. 1997;19:124-133.
11. Sawyer JR et al. *Blood*. 1998;91:1732-1741.
- 15 12. Le Baccon et al. *Genes Chromosomes Cancer*. 2001;32:250-64.
13. Sawyer JR, et al. *Genes Chromosomes Cancer*. 2005;42:95-106.
14. Rosinol L, Carrio A, Blade J, et al. *Br J Haematol*. 2005;130:729-732.
15. Hanamura et al. *Blood*. 2006;16: Epub antes de imprenta.
16. Zhan F, et al. *Blood*. 2002; 99:1745-1757.
- 20 17. Storey JD, Tibshirani R. *Proc Natl Acad Sci*. 2003;100:9440-9445.
18. Kaplan EL, Meier P, *J Am Stat Assoc* 1958; 53:457-481,
19. Mantel N. *Cancer Chemother Rep* 1966; 50:163-170.
20. O'Quigley J, Xu R, Stare J. *Stat Med*. 2005; 24:479-489.
21. R. Development Core Team. Viena, Austria. 2004; (ISBM 3-900051-07-0, URL <http://www.R-project.org>).
- 25 22. Rao CR., Wiley, New York (1973).
23. Greipp P et al. *J Clin Oncol*. 2005;23:3412-20.
24. Shaughnessy, J. et al. *Blood*. 2003;101:3849-3856.
25. Kapanadze et al. *FEBS Lett*. 1998;426:266-270.
26. Itoyama et al. *Genes Chromosomes Cancer*. 2002;35:318-328.
- 30 27. Lu YJ et al. *Lancet* 2002;360:385-386.

28. Hattinger et al. Br J Cancer. 2002;86:1763-1769.
29. Cheng KW, et al. Nat Med. 2004; 10:1251-2156
30. Zudaire I et al. Histopathology. 2002;40:547-555.
31. Richardson P et al. N Engl J Med 2005; 16; 352: 2487-2498.
- 5 32. Richardson P et al. Blood 9 Agosto 2007; [epub antes de imprenta].
33. Shaughnessy JD et al. Blood 2007; 109: 2276-2284.
34. Mulligan G et al. Blood 2007; 109: 3177-3188.
35. Shaffer AL et al. Immunity 2001; 15: 375-385.
36. Ferrando AA et al. Cancer Cell 2002; 1: 75-87.
- 10 37. Ross ME et al. Blood 2004; 104: 3679-3687.

REIVINDICACIONES

1. Un método de elaboración de perfiles de expresión génica para identificar firmas genómicas vinculadas a un pronóstico de supervivencia en un sujeto con mieloma múltiple sintomático o mieloma múltiple, que comprende:
- 5 determinar los niveles de expresión de genes que consisten en *KIF14*, *SLC19A1*, *CKS1B*, *YWHAZ*, *MPHOSPH1*, *TMPO*, *NADK*, *LARS2*, *TBRG4*, *AIM2*, *ASPM*, *AHCYL1*, *CTBS*, *MCLC*, *LTBP1* y genes detectables con ácidos nucleicos que tienen los números de referencia de Affymetrix 242488_at y 1557277_a_at en una micromatriz Affymetrix U133Plus2.0, en una muestra de ácido nucleico a partir de células plasmáticas aisladas del sujeto; y
- determinar la presencia de niveles de expresión anormales, llevando a cabo de este modo la prognosis del sujeto.
- 10 2. El método de la reivindicación 1, en donde el perfil de expresión anormal es evaluado mediante una puntuación discriminante a partir del análisis discriminante por pasos multivariante (MSDA).
3. El método de la reivindicación 2, en donde la puntuación discriminante se calcula con la fórmula: $200638_s_at \times 0,283 - 1557277_a_at \times 0,296 - 200850_s_at \times 0,208 + 201897_s_at \times 0,314 - 202729_s_at \times 0,287 + 203432_at \times 0,251 + 204016_at \times 0,193 + 205235_s_at \times 0,269 + 206364_at \times 0,375 + 206513_at \times 0,158 + 211576_s_at \times 0,316 + 213607_at \times 0,232 - 213628_at \times 0,251 - 218924_s_at \times 0,230 - 219918_s_at \times 0,402 + 220789_s_at \times 0,191 + 242488_at \times 0,148$, donde los conjuntos de sondas 1557277_a_at y 242488_at son parte de una micromatriz Affymetrix U133Plus2.0, y, en donde las variables representan el valor para la sonda concreta y una puntuación por encima de un umbral predeterminado indica una prognosis de alto riesgo.
- 15 4. El método de acuerdo con una cualquiera de las reivindicaciones 1 a 3, caracterizado porque una alta razón media de expresión es indicativa de una firma genómica asociada con enfermedad de alto riesgo y una baja razón media de expresión es indicativa de una firma genómica asociada con una enfermedad de bajo riesgo.
- 20 5. El método de acuerdo con la reivindicación 4, caracterizado porque la firma genómica de la enfermedad de alto riesgo se correlaciona con la menor duración de la remisión completa, la ausencia de eventos, la muerte temprana relacionada con la enfermedad o una de sus combinaciones y dicha firma genómica de enfermedad de bajo riesgo se correlaciona con una mayor duración de la remisión completa, una supervivencia más larga, un buen pronóstico o una de sus combinaciones.
- 25 6. El método de acuerdo con una cualquiera de las reivindicaciones 1 a 5, caracterizado porque el método pronostica el resultado clínico y la supervivencia de un individuo, es eficaz en la selección de tratamiento para un individuo que padece una enfermedad, pronostica el riesgo de recaída después del tratamiento y la supervivencia de un individuo, correlaciona la clasificación molecular de una enfermedad con la firma genómica que define los grupos de riesgo, o una de sus combinaciones.
- 30 7. El método de acuerdo con la reivindicación 6, caracterizado porque la clasificación molecular es CD1 y se correlaciona con la firma genómica de alto riesgo de mieloma múltiple.
8. El método de acuerdo con la reivindicación 7, caracterizado porque la clasificación CD1 comprende el aumento de expresión de firmas MMSET, MAF/MAFB, PROLIFERATION o una de sus combinaciones.
- 35 9. El método de acuerdo con la reivindicación 6, caracterizado porque la clasificación molecular es CD2 y se correlaciona con la firma genómica de bajo riesgo de mieloma múltiple.
10. El método de acuerdo con la reivindicación 9, caracterizado porque la clasificación CD2 comprende translocaciones HYPERDIPLOIDY, LOW BONE DISEASE, CCND1/CCND3, expresión de CD20 o una de sus combinaciones.

40



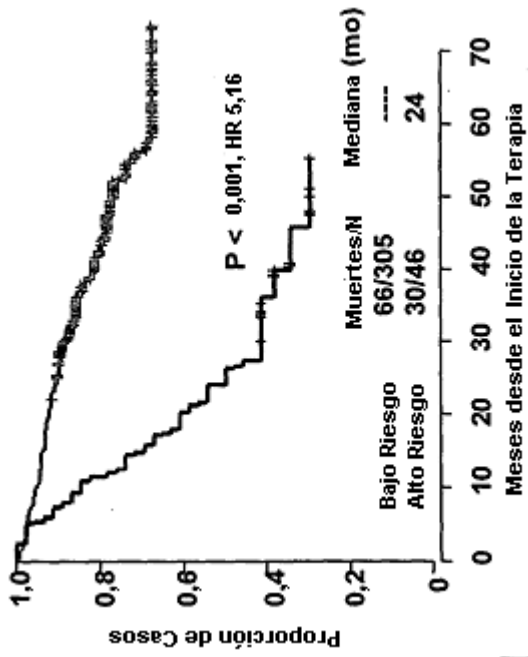


Fig. 1D

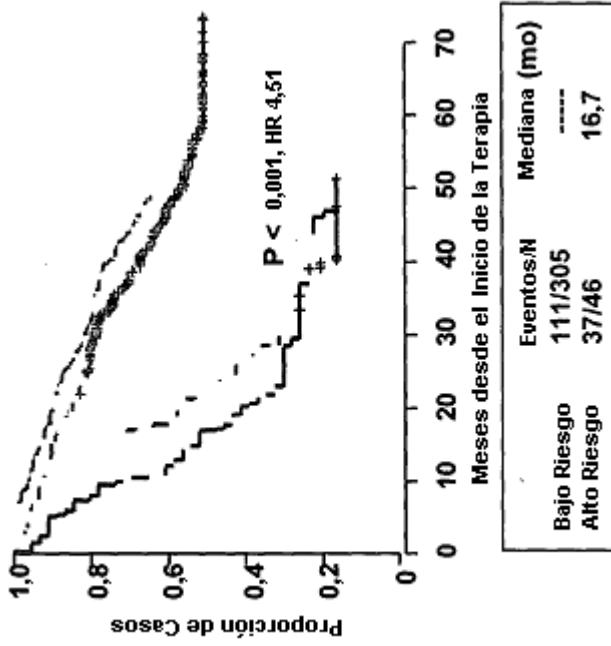


Fig. 1C

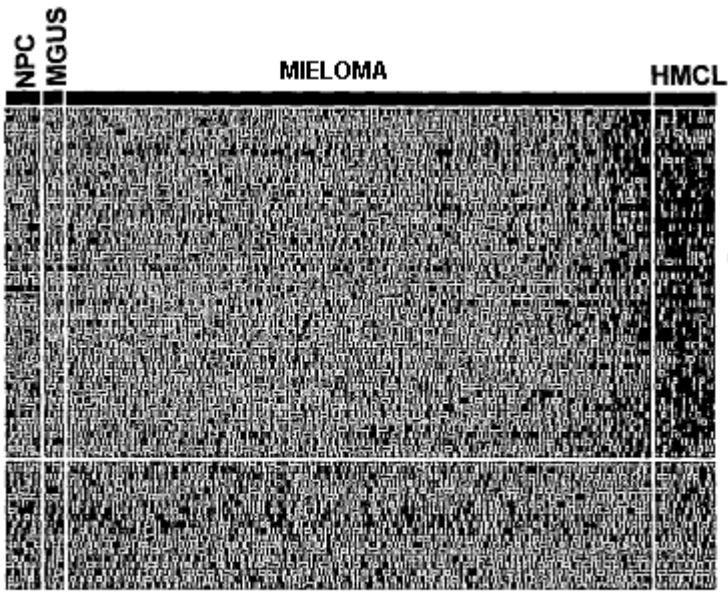
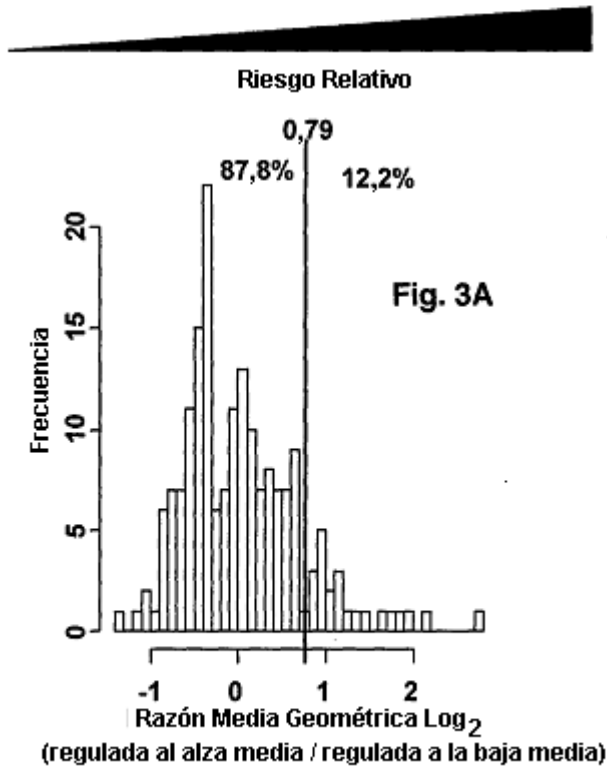


Fig. 2



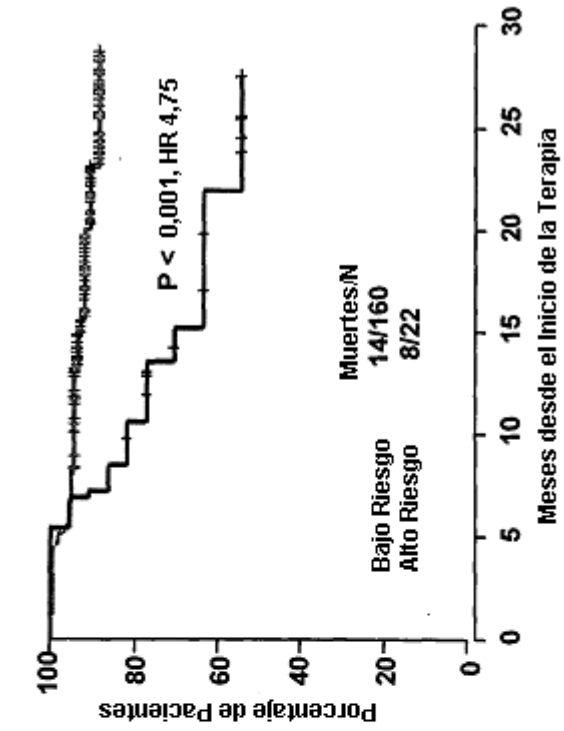


Fig. 3B

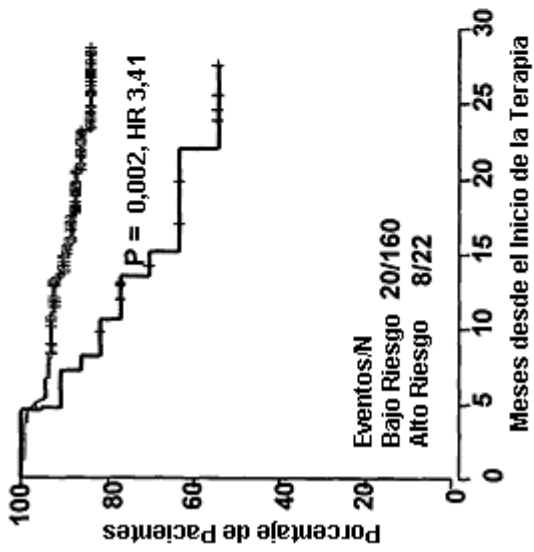
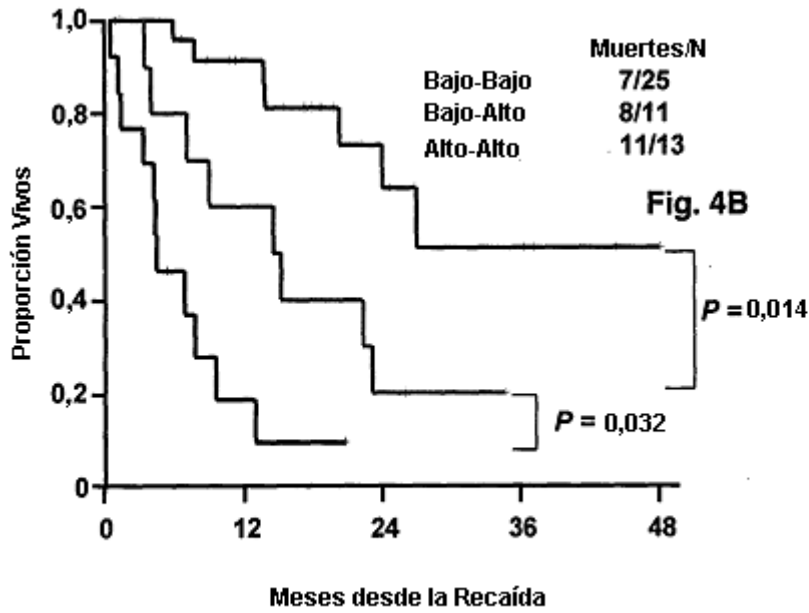
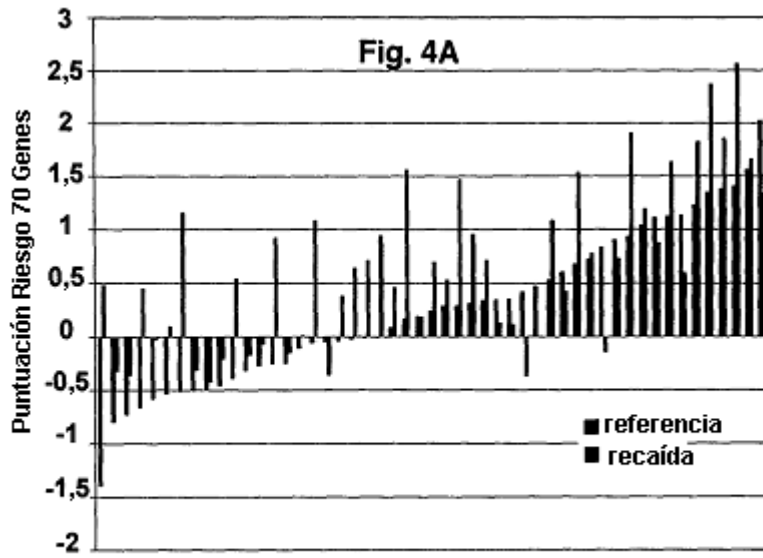
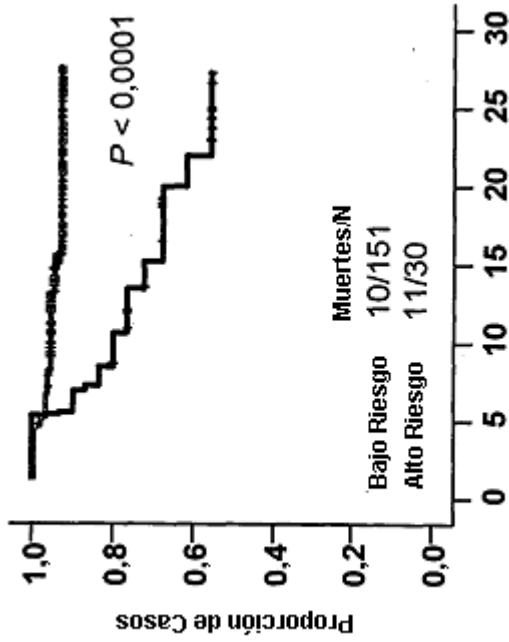


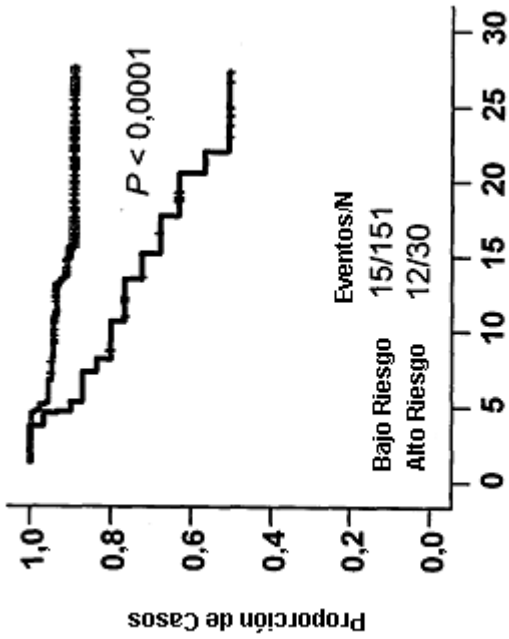
Fig. 3C





Meses desde el Inicio de la Terapia

Fig. 5B



Meses desde el Inicio de la Terapia

Fig. 5A

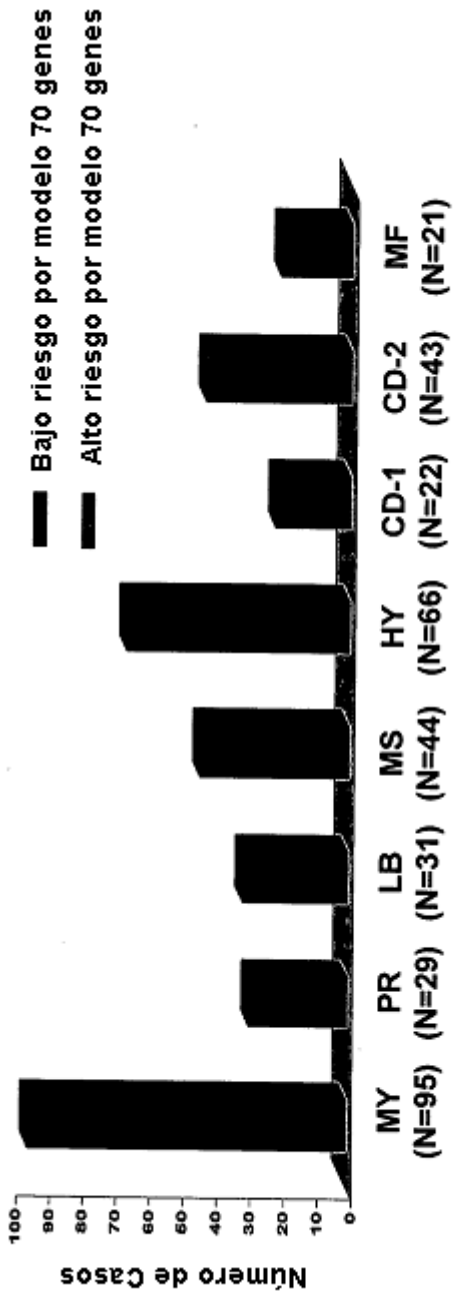
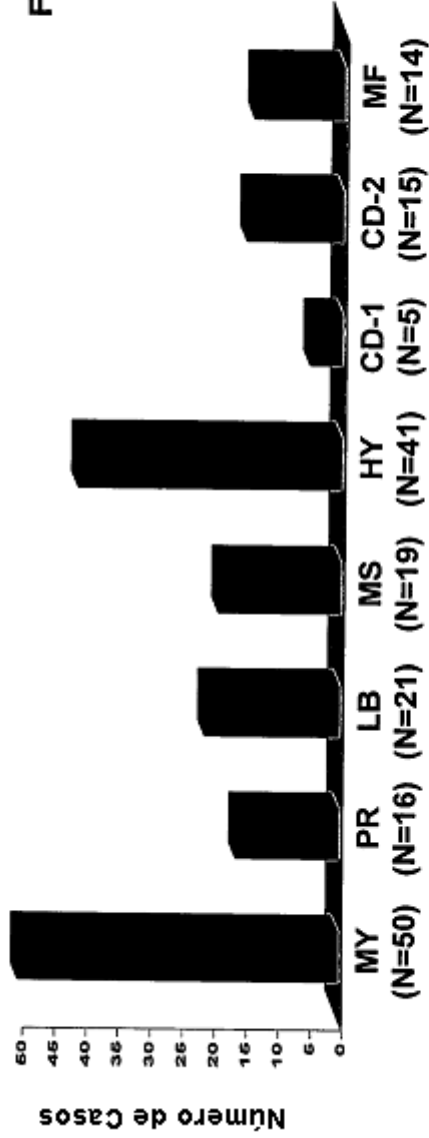
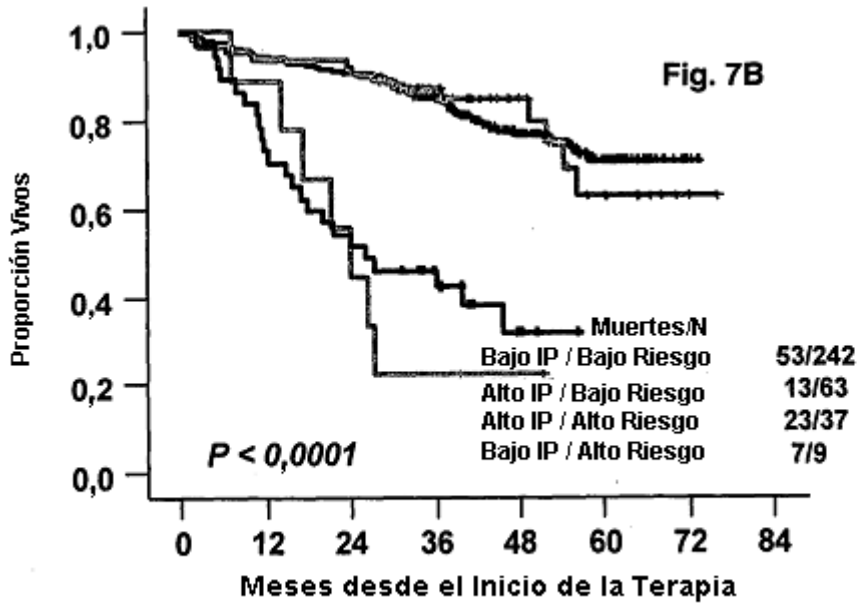
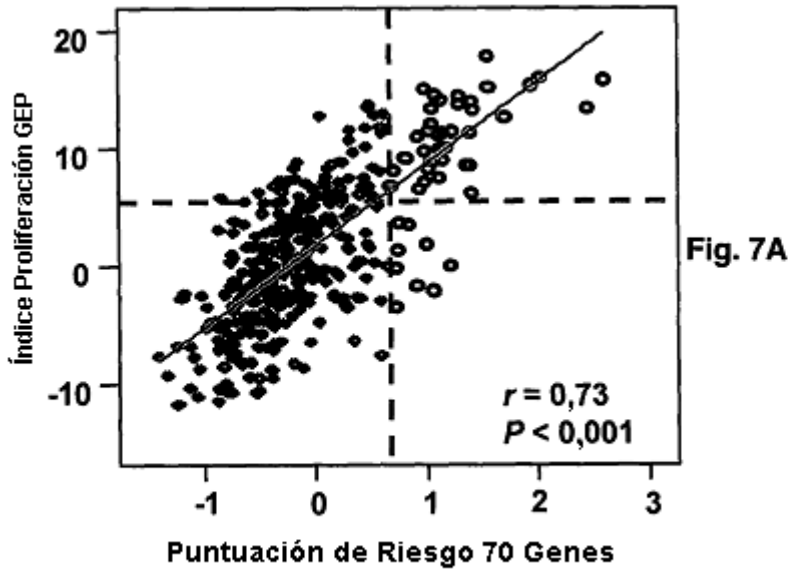


Fig. 6





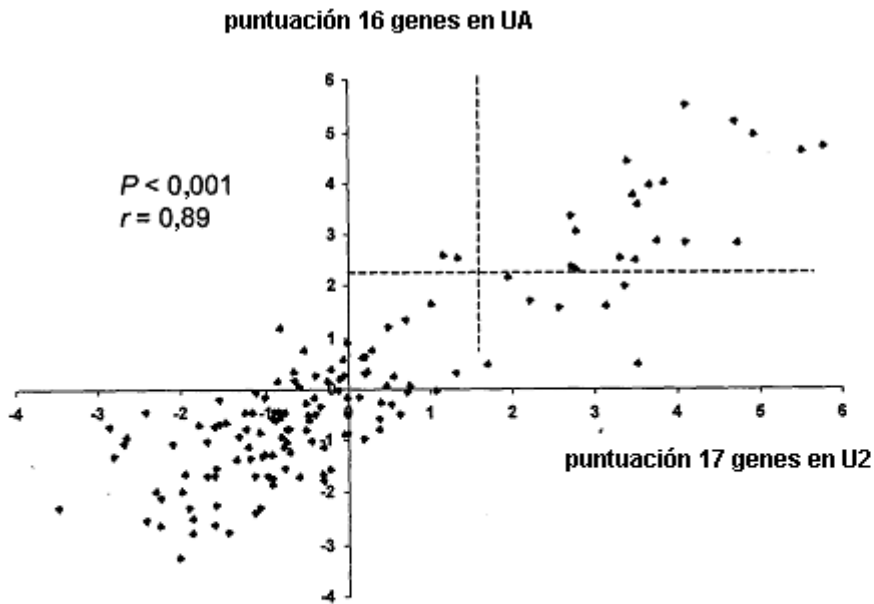
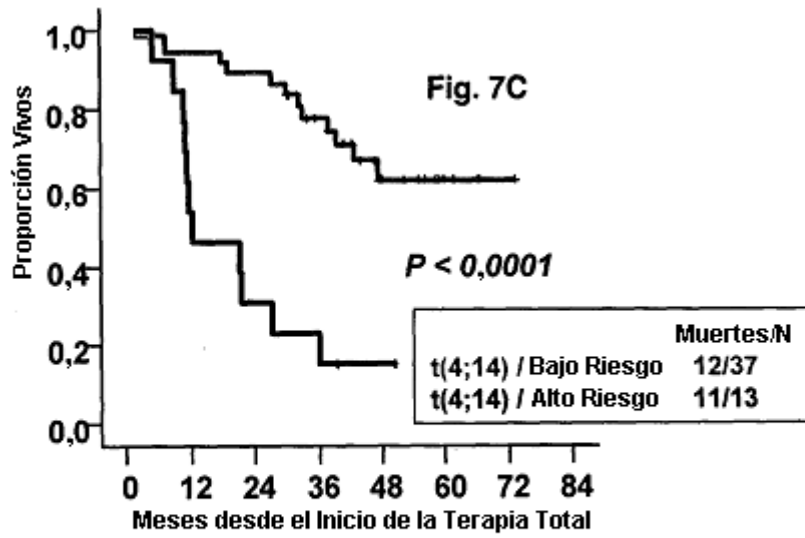
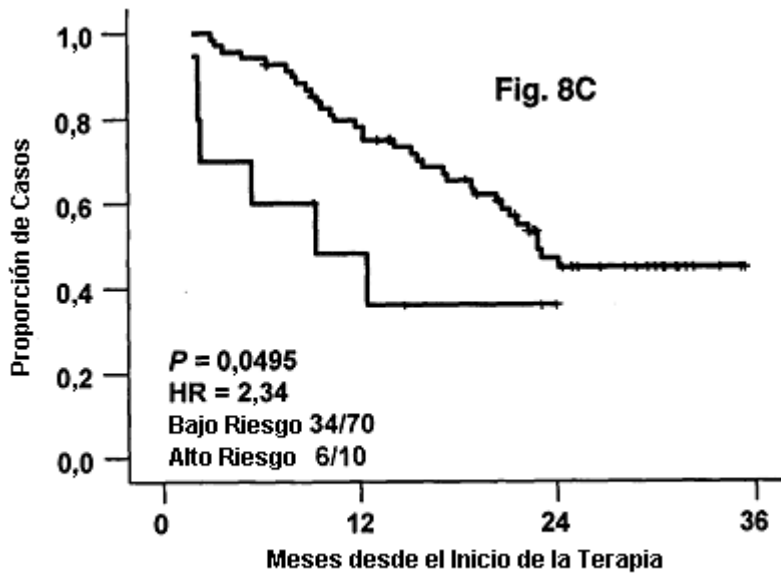
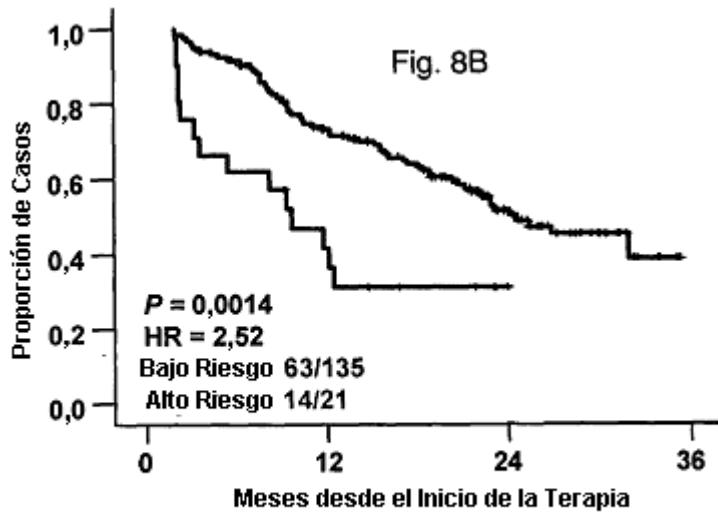


Fig. 8A



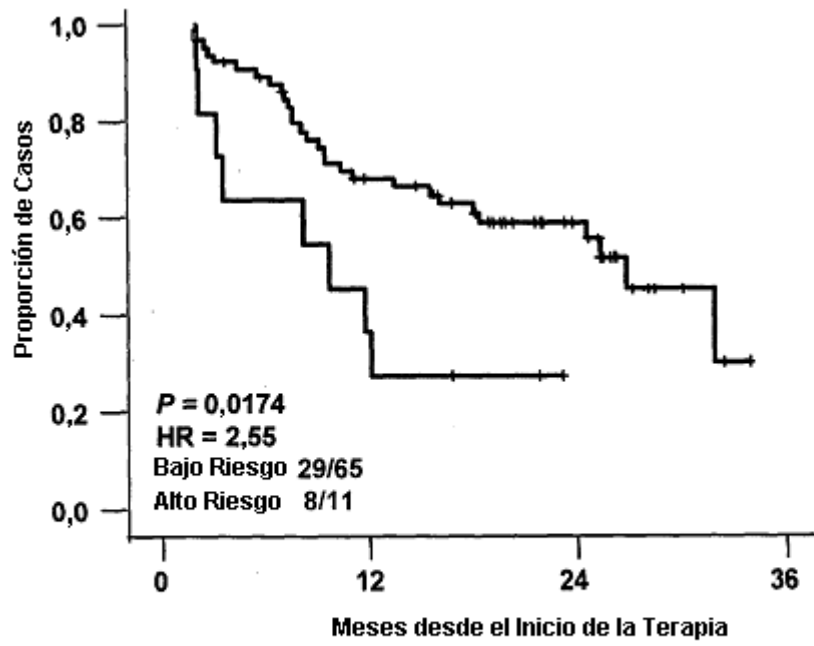


Fig. 8D

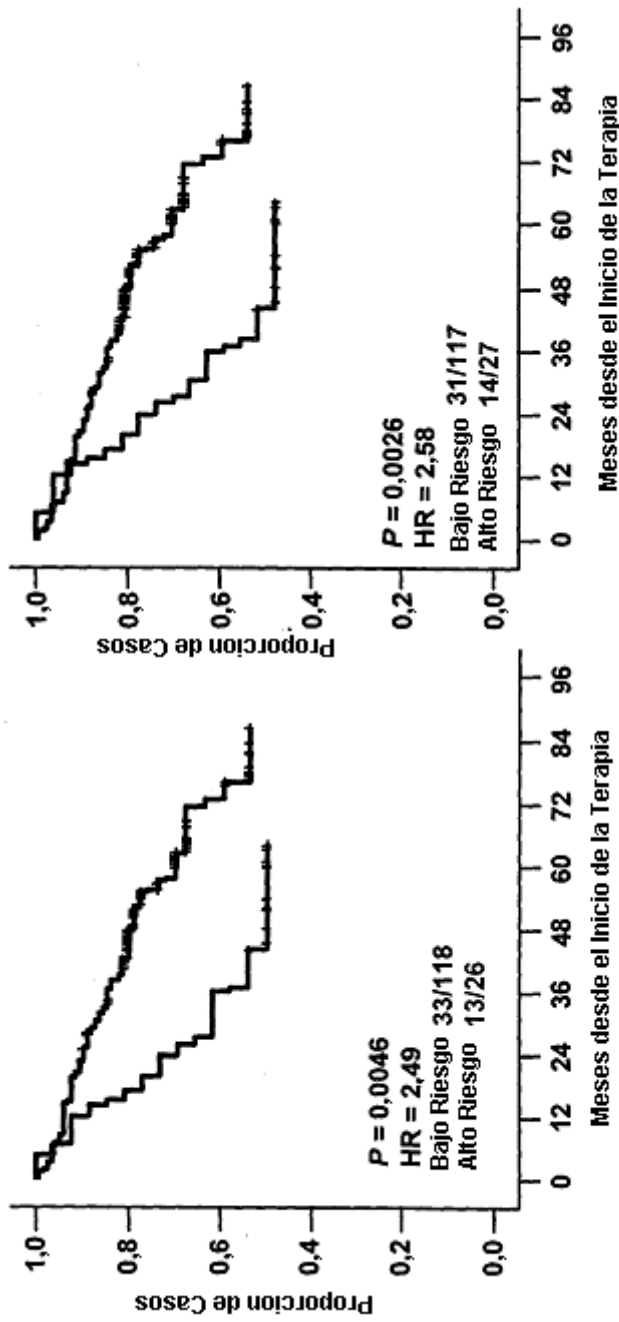


Fig. 9A

Fig. 9B