

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 819 055**

51 Int. Cl.:

**G01N 33/574** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **13.05.2011 PCT/ES2011/070346**

87 Fecha y número de publicación internacional: **17.11.2011 WO11141612**

96 Fecha de presentación y número de la solicitud europea: **13.05.2011 E 11780257 (9)**

97 Fecha y número de publicación de la concesión europea: **17.06.2020 EP 2570813**

54 Título: **Método para el diagnóstico/pronóstico de cáncer colorrectal**

30 Prioridad:

**13.05.2010 ES 201030708**

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

**14.04.2021**

73 Titular/es:

**CONSEJO SUPERIOR DE INVESTIGACIONES  
CIENTÍFICAS (CSIC) (100.0%)**

**Serrano 117  
28006 Madrid, ES**

72 Inventor/es:

**CASAL ÁLVAREZ, JOSÉ IGNACIO;  
BARDERAS MANCHADO, RODRIGO y  
BABEL, INGRID HENRIETTE SUZANNE**

74 Agente/Representante:

**ARIAS SANZ, Juan**

ES 2 819 055 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

Método para el diagnóstico/pronóstico de cáncer colorrectal

5 **Campo de la invención**

La presente invención se encuadra dentro del campo de la biomedicina. Específicamente, se refiere a un método para detectar un autoanticuerpo en una muestra y a un kit para llevar a cabo dicho método.

10 **Antecedentes de la invención**

El cáncer colorrectal (CCR) es el segundo cáncer más prevalente en el mundo occidental. El desarrollo de la enfermedad se produce durante décadas e implica múltiples sucesos genéticos. A pesar de que el CCR es uno de los tumores sólidos mejor caracterizados desde un punto de vista genético, sigue siendo una de las principales causas de mortalidad en países desarrollados por el diagnóstico tardío de los pacientes debido al tiempo de espera que transcurre para realizar ciertas pruebas diagnósticas, tales como colonoscopia.

15  
20 Hoy en día, existen pocas proteínas que hayan sido descritas como biomarcadores eficaces de CCR (antígeno carcinoembrionario (CEA), CA19.9 y CA125) (Crawford et al. 2003. Journal of surgical oncology 84 (4), 239-248; Duffy et al. 2007 Eur J Cancer 43 (9), 1348-1360) y no son lo suficientemente específicas como para llevar a cabo cribados clínicos con vistas a detectar CCR (Locker et al. 2006. J Clin Oncol 24 (33), 5313-5327).

25 Los análisis proteómicos están siendo activamente utilizados para la identificación de nuevos biomarcadores. En diferentes estudios proteómicos previos se han identificado, mediante el uso de micromatrices de anticuerpos y 2D-DIGE, proteínas diferencialmente expresadas en tejido de CCR, incluyendo isoformas y modificaciones postraduccionales responsables de modificaciones en rutas de señalización (Alfonso et al. 2005. Proteomics 5 (10), 2602-2611; Kopf et al. 2005. Proteomics 5(9), 2412-2416; Madoz-Gurpide et al. 2007. Mol Cell Proteomics 6 (12), 2150-2164; Alfonso et al. 2008. Journal of proteome research 7 (10), 4247-4255). Estos dos enfoques han permitido la identificación de una amplia colección de potenciales marcadores tumorales de tejido de CCR que actualmente están siendo investigados.

35 Sin embargo, la implementación de métodos diagnósticos no invasivos y más simples que permitan la detección temprana del CCR debería basarse en la identificación de proteínas o anticuerpos detectables en suero o plasma (Hanash et al. 2008. Nature 452 (7187), 571-579; Hudson et al. 2007. Proceedings of the National Academy of Sciences of the United States of America 104 (44), 17494-17499). La existencia de una respuesta inmunitaria a cáncer en humanos se ha puesto de manifiesto por la presencia de autoanticuerpos en el suero de pacientes con cáncer. Así, diferentes proteínas humanas (autoantígenos) se pueden ver afectadas antes o durante la formación del tumor pudiendo producir una respuesta inmunitaria una vez liberadas (Hudson et al. 2007. Proceedings of the National Academy of Sciences of the United States of America 104 (44), 17494-17499; Wang et al. 2005. The New England journal of medicine 353 (12), 1224-1235; Sreekumar et al. 2004. J Natl Cancer Inst 96 (11), 834-843). Dichos autoanticuerpos se pueden detectar en estadios tempranos de la enfermedad e incluso antes de que el cáncer pueda ser detectado mediante otras técnicas indicando su elevado potencial como biomarcadores de la enfermedad. Estas proteínas tumorales pueden verse afectadas por mutaciones puntuales, tener un plegamiento anómalo, sobreexpresarse, glucosilarse de manera anómala, estar truncadas, o bien sufrir una degradación anómala como es el caso de p53, HER2, NY-ESO1 o MUC1, respectivamente (Chen et al. 1997. Proceedings of the National Academy of Sciences of the United States of America 94 (5), 1914-1918; Schubert et al. 2000. Nature 404 (6779), 770-774; Ulanet et al. 2003. Proceedings of the National Academy of Sciences of the United States of America 100 (21), 12361-12366). De hecho, autoantígenos asociados a tumores (AAT) han sido previamente caracterizados en CCR utilizando distintos enfoques (Scanlan et al. 1998. International Journal of cancer 76 (5), 652-658). Varios autores han descrito algunos paneles de AAT como biomarcadores de CCR, entre los que se encuentra la proteína STK4/MST1 (Tan et al. 2009. Journal 276: 6880-6904; Babel et al. 2009. Molecular and Cellular Proteomics 8: 2382-2395; documento WO 2010/136629). El documento WO2009/008496 divulga que hSULF1 se expresa mucho en cáncer de colon y es un antígeno que se puede usar para la producción de un anticuerpo anti-hSULF1 que inhibe la proliferación celular de una célula cancerosa. No obstante, la validez diagnóstica de los autoanticuerpos asociados con CCR identificados hasta la fecha requiere aún una validación independiente para su uso generalizado en el diagnóstico/pronóstico del CCR.

60 Existe, por tanto, una necesidad de disponer de biomarcadores que permitan el diagnóstico de CCR, su clasificación en los diferentes estadios de la evolución tumoral, el pronóstico de la evolución de la enfermedad, la evaluación de su respuesta a un determinado tratamiento y la detección de la recurrencia o la diseminación de CCR, mediante un método sencillo, eficaz y no invasivo.

**Breve descripción de la invención**

La presente divulgación se refiere a un método para la obtención de datos útiles para el diagnóstico, pronóstico o seguimiento de la evolución de cáncer colorrectal (CCR), a un método de diagnóstico de CCR, a un método de pronóstico de CCR y a un kit para llevar a cabo dichos métodos.

5 La presente divulgación proporciona, por tanto, una respuesta a la necesidad de disponer de biomarcadores que permitan el diagnóstico del CCR, su clasificación en los diferentes estadios de la evolución tumoral, el pronóstico de la evolución de la enfermedad, la evaluación de su respuesta a un determinado tratamiento y la detección de la recurrencia o la diseminación de CCR, mediante un método sencillo, eficaz y no invasivo.

10 La sangre es normalmente el fluido biológico óptimo en que se basan los métodos no invasivos para el cribado masivo con fines diagnósticos de grandes poblaciones de pacientes. Por un lado, el suero y el plasma son fáciles de obtener y, por otra parte, la circulación sanguínea facilita el contacto de la sangre con todos los tejidos del cuerpo humano, incluyendo, en el caso de pacientes con cáncer, el contacto con tejido tumoral y sus antígenos representativos. La liberación de estos antígenos asociados al tumor probablemente ocurre a muy baja concentración en plasma y probablemente sufren proteólisis en un corto periodo de tiempo. Por el contrario, los anticuerpos son moléculas muy estables que han sido usados durante años en diferentes inmunoensayos clínicos, lo cual facilita la estandarización de los ensayos. El uso de autoanticuerpos es también beneficioso ya que el sistema inmunitario amplifica la respuesta, facilitando su identificación y cuantificación.

20 En la presente invención se han empleado micromatrices de fagos (en ocasiones identificadas en esta descripción como "fagos-péptidos") para identificar autoanticuerpos presentes en el suero de pacientes con CCR en distintos estadios.

25 Se seleccionaron 6 fagos que contenían las secuencias homólogas a las proteínas NHSL1, GRN, MST1, SULF1, SREBF2 y GTF2i. La combinación de las proteínas recombinantes MST1 y SULF1 con las otras 4 secuencias de fagos permitió predecir la enfermedad con un 72% de sensibilidad y un 87% de especificidad, con un área bajo la curva (AUC, por sus siglas en inglés) de 0,83. Si además se tiene en cuenta la edad del paciente, el AUC es de 0,91. Además, estos marcadores permiten agrupar los resultados discriminando, no sólo los individuos enfermos, sino también los distintos estadios de la enfermedad. La detección de este panel de autoanticuerpos en suero es, por tanto, un método sencillo y no invasivo para el diagnóstico/pronóstico de CCR.

En un primer aspecto, la invención se refiere a un método para detectar un autoanticuerpo en una muestra, que comprende:

35 a) poner en contacto una muestra de un sujeto con una entidad de captura de anticuerpos (ECA), en donde dicha ECA es una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 4 o un péptido o proteína que tiene al menos el 90% de identidad de secuencia con respecto a SEQ ID NO: 4 y que es reconocido por un autoanticuerpo contra SEQ ID NO: 4;

40 y  
b) detectar la formación de un complejo entre un autoanticuerpo y dicha ECA,

en donde la detección de dicho complejo autoanticuerpo-ECA es indicativa de la presencia de dicho autoanticuerpo en dicha muestra, en donde la muestra es sangre, plasma o suero.

45 En un segundo aspecto, la invención se refiere a una entidad de captura de anticuerpos (ECA) que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 4, en donde dicha ECA no es la proteína sulfatasa-1 (SULF1).

En un tercer aspecto, la invención se refiere a una composición que comprende una entidad de captura de anticuerpos (ECA) según la invención.

50 En un cuarto aspecto, la invención se refiere a un kit que comprende una composición según la invención.

En un quinto aspecto, la invención se refiere a un uso in vitro de un kit según la invención para detectar un anticuerpo en una muestra, o para detectar un autoanticuerpo en un sujeto que se sospecha tiene cáncer colorrectal (CCR), o para diagnosticar si un sujeto tiene CCR, o para determinar el riesgo de que un sujeto desarrolle CCR, o para seguir la evolución del CCR en un sujeto, o para evaluar la eficacia de un tratamiento contra CCR, o para predecir la supervivencia de un sujeto que tiene CCR.

### Breve descripción de las figuras

60 La figura 1 muestra la respuesta de los autoanticuerpos a los seis fagos específicos. Intensidad de la señal de cada fago con los sueros de CCR y los sueros control. Los resultados muestran los datos después de la normalización y en una escala de unidades arbitrarias (u.a.).

65 La figura 2 muestra el análisis de competición entre los péptidos de los fagos y sus respectivas proteínas homólogas. A. Se realizó un ensayo de competición de ELISA entre los fagos que presentaban péptidos con

- homología a SULF1 y MST1 y sus respectivas proteínas recombinantes. Se utilizó GST como control negativo de inhibición. Cantidades crecientes de las proteínas recombinantes se preincubaron con el suero de pacientes y se ensayó su unión al fago respectivo mediante ELISA (Barras verticales: gris claro, proteína recombinante; gris oscuro, GST). Se usó la proteína EBNA1 como control para demostrar que la inhibición era específica de la proteína usada como inhibidor y que no se introdujo ningún error en el ensayo (datos no mostrados). B. Localización esquemática de los péptidos con homología a SULF1 y MST1 dentro de las proteínas recombinantes. La posición del péptido en la proteína aparece marcada en la figura. Las líneas verticales corresponden a potenciales sitios de fosforilación. Los aminoácidos diferentes entre la secuencia del fago y la proteína aparecen en minúsculas.
- La figura 3 muestra el análisis de la expresión de SULF1, MST1, GTF2i, NHSL1, GRN y SREBF2 en líneas celulares y tejido de pacientes con CCR. A. Los niveles de expresión génica de las proteínas cuyo péptido se presenta en los fagos T7 se determinó mediante metaanálisis utilizando la base de datos Oncomine. Los valores p se indican en la figura. Se encontraron niveles relativos de expresión génica para NHSL1, SREBF2, GTF2i, SULF1, MST1 y GRN. B. Inmunodetección en membrana de SULF1 y MST1 en líneas celulares de cáncer colorrectal en comparación con líneas celulares control y tejidos emparejados de CCR pertenecientes a estadios I, II y III. Se usó un anticuerpo anti-tubulina como control de carga. C, Datos de la micromatriz de tejidos para GTF2i y GRN obtenidos del Atlas de Proteínas Humanas (Human Protein Atlas WebPage).
- La figura 4 muestra los valores predictivos de las proteínas MST1 y SULF1. A. Media de los valores de absorbancia obtenidos con los sueros de CCR y los controles mediante ELISA indirecto. Los puntos representan el valor individual para cada suero. Las barras de error representan el valor de desviación estándar. Las imágenes de los geles de poliacrilamida corresponden a las proteínas recombinantes usadas en los ensayos de ELISA. B. Ambas proteínas eran capaces de discriminar sueros control de sueros de pacientes con CCR con valores  $p < 0,0001$  y  $0,0006$ , para MST1 y SULF1, respectivamente. El AUC para MST1 fue de  $0,75$  (IC del 95% =  $0,647-0,829$ ) con una sensibilidad y una especificidad del 60,0% y el 82,6 %, respectivamente, utilizando como punto de corte  $0,63$ . Para SULF1 el AUC fue de  $0,72$  (IC del 95% =  $0,617-0,805$ ), con una sensibilidad y una especificidad del 68% y el 67,4%, respectivamente, utilizando como punto de corte  $0,36$ . C y D, Curvas ROC para CEA y la combinación de 4 fagos, 2 proteínas y la variable edad, que arrojaron valores de AUC de  $0,81$  y  $0,89$ , respectivamente.
- La figura 5 muestra el análisis de supervivencia utilizando los autoanticuerpos contra las proteínas MST1 y NHSL1. Las curvas de supervivencia Kaplan-Meier se calcularon utilizando un conjunto independiente de 95 sueros de pacientes con CCR para analizar el efecto de la presencia de autoanticuerpos en la supervivencia absoluta de los pacientes con CCR.
- La figura 6 muestra la validación de la combinación de cuatro fagos con las proteínas MST1 y SULF1 en el diagnóstico de cáncer colorrectal. Comportamiento de la combinación de los fagos MST1 y NHSL1 GTF2i, NHSL1, GRN y SREBF2 y las proteínas MST1 y NHSL1 en el ensayo de validación. A. Comportamiento de muestras CCR frente controles sanos. B. Comportamiento de muestras CCR frente a sueros de referencia. C. Comportamiento de sueros sanos frente a sueros de tumores. D. Diagrama de puntos que muestra la probabilidad individual de ser clasificado como paciente CCR para cada uno de los sujetos con diferentes patologías. La mayoría de las muestras fueron clasificadas por debajo del valor de  $0,5$  de probabilidad (línea vertical discontinua).

## Descripción detallada de la invención

### Definiciones

- Para facilitar su comprensión, a continuación, se indica el significado de algunos términos y expresiones tal como se utilizan en la presente descripción.
- El término “anticuerpo”, tal como se utiliza en el presente documento, se refiere a moléculas de inmunoglobulinas y porciones inmunológicamente activas de moléculas de inmunoglobulinas, es decir, moléculas que contienen un sitio de fijación de antígeno que se une específicamente (inmunorreacciona) con un antígeno, tal como, por ejemplo, una proteína. Hay 5 isotipos o clases principales de inmunoglobulinas: inmunoglobulina M (IgM), inmunoglobulina D (IgD), inmunoglobulina G (IgG), inmunoglobulina A (IgA) e inmunoglobulina E (IgE).
- El término “autoanticuerpo”, tal como aquí se utiliza, se aplica a un anticuerpo que reacciona contra un antígeno presente en el propio organismo de un sujeto, incluso si la reacción ocurre sólo *in vitro*, y tanto si causa efectos patológicos *in vivo* como si no los produce.
- Tal como se usa en el presente documento, el término “cáncer colorrectal” o “CCR”, también llamado cáncer de colon, incluye cualquier tipo de neoplasias del colon, recto y apéndice, así como cualquier subtipo histológico que aparece típicamente en el cáncer de colon, por ejemplo, células de carcinoma transicional, células de carcinoma escamoso y adenocarcinoma, cualquier subtipo clínico, por ejemplo, superficial, músculo invasivo o cáncer de enfermedad metastásica, o cualquier estadio TNM incluyendo tumores T0-T4, N0-N2 y M0-M1. Los pacientes pueden ser clasificados en diferentes grupos con respecto al estadio del tumor. La clasificación del cáncer de colon es una estimación de la penetración de un cáncer particular. Se lleva a cabo con fines de investigación, diagnóstico



y para determinar el mejor método de tratamiento. El sistema para la clasificación del cáncer colorrectal depende del nivel de la invasión local, del grado de nodos linfáticos implicados y de si existe metástasis distante. El sistema más común de clasificación es el sistema TNM (para tumores/ganglios/metástasis), del "American Joint Committee on Cancer" (AJCC). El sistema TNM asigna un número basado en tres categorías. "T" indica el grado de invasión de la pared intestinal, "N" el grado de afectación de ganglios linfáticos y "M" el grado de metástasis. El estadio más amplio del cáncer es normalmente citado como un número I, II, III, IV derivado del valor TNM agrupado por el pronóstico, un número más alto indica un cáncer más avanzado y un peor pronóstico. En la Tabla 1 se indican detalles de la clasificación.

Tabla 1  
Sistema TNM para la clasificación de CCR

Estadio AJCC	Estadio TNM	Criterios de los estadios TNM para CCR
Estadio 0	Tis N0 M0	Tis: El tumor confinado a la mucosa; <i>cáncer-in-situ</i>
Estadio I	T1 N0 M0	T1: El tumor invade la mucosa
Estadio I	T2 N0 M0	T2: El tumor invade la musculatura propia
Estadio II-A	T3 N0 M0	T3: El tumor invade la capa subserosa o más allá (otros órganos no implicados)
Estadio II-B	T4 N0 M0	T4: El tumor invade órganos adyacentes o perfora el peritoneo visceral
Estadio III-A	T1-2 N1 M0	N1: Metástasis de 1 a 3 ganglios linfáticos regionales. T1 o T2.
Estadio III-B	T3-4 N1 M0	N1: Metástasis de 1 a 3 ganglios linfáticos regionales. T3 o T4.
Estadio III-C	cualquier T, N2 M0	N2: Metástasis de 4 o más ganglios linfáticos regionales. Cualquier T.
Estadio IV	cualquier T, cualquier N, M1	M1: Presencia de metástasis distante. Cualquier T, cualquier N.

El término "entidad capturadora de anticuerpo" (ECA), tal como se utiliza en el presente documento, se refiere a una entidad macromolecular que se une específicamente a un anticuerpo (o autoanticuerpo). En una divulgación particular, dicha ECA comprende un péptido o una proteína que se une específicamente a un anticuerpo (o autoanticuerpo). Dicho péptido puede estar inmovilizado sobre un soporte o bien expuesto sobre la superficie de un fago. En una divulgación particular preferida, dicha ECA es un péptido, una proteína o un fago en cuya superficie se expone dicho péptido o dicha proteína. Dicha ECA puede estar, si se desea, inmovilizada sobre un soporte sólido.

El término "muestra", tal como aquí se utiliza, se refiere, pero no se limita, a tejidos y/o fluidos biológicos de un sujeto, obtenidos mediante cualquier método conocido por un experto en la materia que sirva para llevar a cabo cualquiera de los métodos proporcionados por la presente divulgación; es decir, dicha muestra biológica debe ser una muestra susceptible de contener anticuerpos, por ejemplo, autoanticuerpos contra las proteínas SULF1, MST1, etc., o contra las ECA que comprenden las secuencias de aminoácidos mostradas en las SEQ ID NO: 1-6, o variantes de las mismas que contienen epítopos reconocibles por autoanticuerpos, etc. A modo ilustrativo, no limitante, dicha muestra biológica puede ser una muestra de sangre, orina, saliva, suero, plasma, un frotis bucal o buco-faríngeo, un espécimen quirúrgico, un espécimen obtenido de una biopsia o autopsia, etc. En una divulgación particular, dicha muestra es un fluido biológico. En una divulgación preferida, para la detección de autoanticuerpos, la muestra del sujeto es sangre, plasma o suero sanguíneo. En otra divulgación particular, dicha muestra es una muestra de tejido. En una divulgación preferida, para la cuantificación del nivel de la proteína SULF1, dicha muestra es, preferentemente, una muestra de tejido colorrectal, o muestra de tejido tumoral, etc., obtenida por métodos convencionales, por ejemplo, mediante una biopsia, una resección, etc.

El término "proteína SULF1", según se usa en el presente documento, incluye la proteína SULF1 y variantes de la misma; en una divulgación particular, dicha proteína es la proteína cuyo número de acceso en la base de datos NCBI (versión de 1 de mayo de 2011) es EAW86954.1 y su secuencia de aminoácidos es SEQ ID NO: 10.

El término "proteína NHSL1", según se usa en el presente documento, incluye la proteína NHSL1 y variantes de la misma; en una divulgación particular, dicha proteína es la proteína cuyo número de acceso en la base de datos NCBI (versión de 1 de mayo de 2011) es NP\_001137532.1 y su secuencia de aminoácidos es SEQ ID NO: 7.

El término "proteína GRN", según se usa en el presente documento, incluye la proteína GRN y variantes de la misma; en una divulgación particular, dicha proteína es la proteína cuyo número de acceso en la base de datos NCBI (versión de 1 de mayo de 2011) es 2JYT y su secuencia de aminoácidos es SEQ ID NO: 8.

El término "proteína MST1", según se usa en el presente documento, incluye la proteína MST1 y variantes de la misma; en una divulgación particular, dicha proteína es la proteína cuyo número de acceso en la base de datos NCBI (versión de 1 de mayo de 2011) es AAA83254.1 y su secuencia de aminoácidos es SEQ ID NO: 9.

El término “proteína SREBF2”, según se usa en el presente documento, incluye la proteína SREBF2 y variantes de la misma; en una divulgación particular, dicha proteína es la proteína cuyo número de acceso en la base de datos NCBI (versión de 1 de mayo de 2011) es NP\_004590.2 y su secuencia de aminoácidos es SEQ ID NO: 11.

- 5 El término “proteína GTF2i”, según se usa en el presente documento, incluye la proteína GTF2i y variantes de la misma; en una divulgación particular, dicha proteína es la proteína cuyo número de acceso en la base de datos NCBI (versión de 1 de mayo de 2011) es NP\_001157108.1 y su secuencia de aminoácidos es SEQ ID NO: 12.

10 El término “variante”, tal como aquí se utiliza, se refiere a una proteína o péptido sustancialmente homólogo a otra proteína o péptido, por ejemplo, a los péptidos cuyas secuencias de aminoácidos se muestran en SEQ ID NO:1 a 6, respecto a las proteínas SULF1, MST1, NHSL1, GRN, SREBF2 o GTF2i, etc. En general, una variante incluye adiciones, deleciones o sustituciones de uno o más aminoácidos. El experto en la materia entenderá que las secuencias de aminoácidos a las que se hace referencia en esta descripción pueden estar modificadas químicamente, por ejemplo, mediante modificaciones químicas que son fisiológicamente relevantes, tales como, fosforilaciones, acetilaciones, glucosilaciones o metilaciones. Dichas variantes, según la presente divulgación, son reconocidas por autoanticuerpos contra la proteína o péptido en cuestión. Variantes de dichos péptidos o proteínas incluyen péptidos o proteínas que muestran al menos un 25%, al menos un 40%, al menos un 60%, al menos un 70%, al menos un 80%, al menos un 90%, al menos un 95%, al menos un 96%, al menos un 97%, al menos un 98% o al menos un 99% de identidad de secuencia con respecto a unas secuencias aminoacídicas de péptidos o proteínas determinadas. El grado de identidad entre dos secuencias de aminoácidos puede determinarse por métodos convencionales, por ejemplo, mediante algoritmos estándar de alineamiento de secuencias conocidos en el estado de la técnica, tales como, por ejemplo, BLAST (Altschul S.F. *et al.* Basic local alignment search tool. J Mol Biol. 1990; 215(3):403-10).

25 Método de detección de autoanticuerpos en sujetos con CCR

Se divulga un método para detectar un autoanticuerpo en un sujeto sospechoso de tener cáncer colorrectal (CCR), en adelante primer método de la divulgación, que comprende:

- 30 a) poner en contacto una muestra de dicho sujeto con una entidad capturadora de anticuerpo (ECA), en donde dicha ECA se selecciona del grupo que consiste en:
- (i) una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 4 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo;
  - 35 (ii) una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 2 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo;
  - (iii) una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 6 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo;
  - 40 (iv) una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 3 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo, en donde dicha ECA no es la proteína MST1;
  - (v) una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 5 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo;
  - 45 (vi) una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 1 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo; y
  - (vii) cualquier combinación de dichas ECAs (i)-(vi); y
- b) detectar la formación de un complejo autoanticuerpo-ECA,

50 en donde la detección de dicho complejo autoanticuerpo-ECA es indicativa de la presencia de dicho autoanticuerpo en dicho sujeto.

En general, la muestra será una muestra biológica susceptible de contener anticuerpos, procedente de un sujeto, y puede ser obtenida por métodos convencionales, conocidos por los expertos en la materia, dependiendo de la naturaleza de la muestra. En una divulgación particular, dicha muestra biológica es una muestra de sangre, suero o plasma, que puede obtenerse por cualquier método convencional, por ejemplo, mediante una extracción de sangre, etc. La sangre es normalmente el fluido biológico óptimo a utilizar en métodos no invasivos para el cribado masivo con fines diagnósticos de grandes poblaciones de sujetos. Por un lado, el suero y el plasma son fáciles de obtener, y, por otro lado, la circulación sanguínea facilita el contacto de la sangre con todos los tejidos del cuerpo humano, incluyendo en el caso de pacientes con cáncer el contacto con tejido tumoral y sus antígenos representativos.

El primer método divulgado en el presente documento comprende poner en contacto una muestra de un sujeto sospechoso de padecer CCR con una ECA seleccionada de entre las ECAs (i) a (vi), indicadas en la Tabla 2, y sus combinaciones [etapa a)], en condiciones que permiten la formación de un complejo autoanticuerpo-ECA.

65

Tabla 2

## ECA

- 5 (i) ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 4 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo.
- (ii) ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 2 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo.
- 10 (iii) ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 6 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo.
- (iv) ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 3 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo; en donde dicha ECA no es la proteína MST1.
- 15 (v) ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 5 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo.
- 20 (vi) ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 1 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo.

25 Las condiciones adecuadas para que tenga lugar la formación del complejo autoanticuerpo-ECA son conocidas por los expertos en la materia. Si la muestra contiene autoanticuerpos contra a dichas ECAs, entonces se formará el correspondiente complejo autoanticuerpo-ECA; en caso contrario, no se formará dicho complejo.

Aunque dichas ECAs podrían estar juntas en un mismo medio, en la práctica, resulta ventajoso que dichas ECAs estén separadas entre sí. Las ECAs pueden estar en solución o suspensión en un medio apropiado, o, alternativamente, pueden estar depositadas o apoyadas sobre un soporte [por ejemplo, una placa de microtitulación, perlas (magnéticas o no magnéticas), columnas, matrices, membranas, etc.]. Estos materiales se pueden utilizar en las formas convenientes, tales como películas, hojas, placas, etc., o pueden utilizarse para recubrir portadores inertes (por ejemplo, papel, cristal, películas de plástico, etc.). En una divulgación particular, la muestra que se va a analizar se pone en contacto con dichas ECAs, separadas entre sí, y depositadas sobre un soporte adecuado.

35 La detección de dichos autoanticuerpos contra las ECAs mencionadas puede llevarse a cabo por métodos convencionales conocidos por los expertos en la materia. En una divulgación particular, la detección de dichos autoanticuerpos se lleva a cabo mediante un inmunoensayo; ejemplos ilustrativos, no limitantes, de inmunoensayos conocidos en el estado de la técnica incluyen inmunotransferencia, enzimoimmunoanálisis de adsorción (ELISA), inmunoensayo lineal (LIA), radioinmunoensayo (RIA), inmunofluorescencia (IF), inmunohistoquímica (IHQ), micromatrices de proteínas, micromatrices de fagos, etc. El experto en la materia también entenderá que pueden utilizarse otros métodos basados, por ejemplo, en técnicas electroforéticas o cromatográficas, para detectar dichos autoanticuerpos.

45 En una divulgación particular, la detección de autoanticuerpos contra una o más ECAs se realiza mediante un ELISA. La técnica ELISA se basa en la premisa de que un inmunorreactivo (por ejemplo, un antígeno o un anticuerpo) se inmoviliza sobre un soporte sólido, y, a continuación, ese sistema se pone en contacto con una fase fluida que contiene el reactivo complementario que puede estar unido a un compuesto marcador. Existen diferentes tipos de ELISA, por ejemplo, ELISA directo, ELISA indirecto o ELISA sándwich.

50 La detección de autoanticuerpos contra una o más ECA(s) mediante un ELISA, por ejemplo, mediante un ELISA indirecto, comprende, en general, las siguientes etapas: (a) recubrir un soporte sólido con una o más ECAs, preferentemente separadas entre sí; (b) incubar el soporte recubierto de la etapa (a) con una muestra, tal como una muestra biológica del sujeto a estudiar, en condiciones que permitan la formación de un complejo autoanticuerpo-ECA; y (c) añadir un anticuerpo secundario, que reconoce al autoanticuerpo contra la(s) ECA(s), conjugado o unido a un compuesto marcador.

55 En otra divulgación particular, la detección de autoanticuerpos contra una o más ECAs se realiza mediante una micromatriz de proteínas. Una micromatriz de proteínas consiste en una colección de proteínas inmovilizadas sobre un soporte sólido en una disposición regular y prefijada. Existen varios factores importantes a tener en cuenta en el diseño de micromatrices de proteínas, entre los que se encuentran, por ejemplo, la naturaleza del soporte sobre el que inmovilizar las proteínas (o fragmentos apropiados de las mismas), la técnica de inmovilización de las proteínas, el formato de micromatriz, el agente de captura empleado o el método de detección a emplear. Son conocidos en el estado de la técnica diferentes formatos, soportes y técnicas que pueden ser empleados para la realización de este primer método.

65 La detección de autoanticuerpos contra una o más ECAs mediante una micromatriz de proteínas comprende, en general, las siguientes etapas: (a) recubrir un soporte sólido con dicha(s) ECA(s), preferentemente separadas entre

5 sí; (b) incubar el soporte recubierto de la etapa (a) con una muestra, tal como una muestra biológica del sujeto a estudiar, en condiciones que permitan la formación de un inmunocomplejo del autoanticuerpo contra la(s) ECA(s) presente(s) en dicha muestra con los determinantes antigénicos correspondientes presentes en dichas ECAs; y (c) añadir un anticuerpo secundario, que reconoce al autoanticuerpo contra la(s) ECA(s), conjugado o unido a un compuesto marcador.

10 En otra divulgación particular, la detección de autoanticuerpos contra una o más ECAs se realiza mediante una micromatriz de fagos. Una micromatriz de fagos consiste en una colección de péptidos expuestos en la superficie de fagos. En una divulgación particular, dichos péptidos están fusionados a la proteína 10B de la cápside de fagos T7. Dichos fagos se inmovilizan sobre un soporte sólido en una disposición regular y prefijada. Existen varios factores importantes a tener en cuenta en el diseño de micromatrices de fagos tales como, por ejemplo, la naturaleza del soporte sobre el cual inmovilizar los fagos, la técnica de inmovilización, el formato de micromatriz o el método de detección a emplear. Son conocidos en el estado de la técnica diferentes formatos, soportes y técnicas que pueden ser empleados para la realización de esta divulgación preferida del método divulgado en el presente documento.

15 La detección de autoanticuerpos contra una o más ECA(s) mediante una micromatriz de fagos comprende, en general, las siguientes etapas: (a) recubrir un soporte sólido con un lisado de fagos, por ejemplo, fagos T7, que presentan expuesta sobre su superficie un péptido o una secuencia de aminoácidos susceptible de ser reconocida por un autoanticuerpo; (b) incubar el soporte recubierto de la etapa (a) con una muestra, tal como una muestra biológica del sujeto a estudiar, en condiciones que permitan la formación de un complejo autoanticuerpo-ECA; y (c) añadir un anticuerpo secundario, que reconoce al autoanticuerpo contra la(s) ECA(s), conjugado o unido a un compuesto marcador. En una divulgación concreta, dicha micromatriz de fagos comprende un fago seleccionado del grupo que consiste en:

- 25 (i) un fago que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 4 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo, en donde dicha secuencia de aminoácidos está expuesta sobre la superficie del fago;
- 30 (ii) un fago que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 2 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo, en donde dicha secuencia de aminoácidos está expuesta sobre la superficie del fago;
- (iii) un fago que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 6 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo, en donde dicha secuencia de aminoácidos está expuesta sobre la superficie del fago;
- 35 (iv) un fago que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 3 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo, en donde dicha secuencia de aminoácidos está expuesta sobre la superficie del fago;
- (v) un fago que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 5 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo, en donde dicha secuencia de aminoácidos está expuesta sobre la superficie del fago;
- 40 (vi) un fago que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 1 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo, en donde dicha secuencia de aminoácidos está expuesta sobre la superficie del fago; y
- (vii) cualquiera de las combinaciones de (i) a (vi).

45 El marcador unido al anticuerpo secundario al que se hace referencia en estas técnicas, es un compuesto capaz de dar lugar a una señal cromogénica, fluorogénica, radiactiva y/o quimioluminiscente que permita la detección, identificación y, opcionalmente, cuantificación de la cantidad del autoanticuerpo contra la(s) ECA(s) presente(s) en la muestra analizada. En una divulgación particular, dicho compuesto marcador se selecciona del grupo que consiste en radioisótopos, enzimas, fluoróforos o cualquier molécula susceptible de ser conjugada con otra molécula o detectada y/o cuantificada de forma directa. Este compuesto marcador puede unirse al autoanticuerpo directamente, o a través de otro compuesto. Ejemplos ilustrativos, no limitantes, de dichos compuestos marcadores que se unen directamente al autoanticuerpo incluyen enzimas, tales como la fosfatasa alcalina, la peroxidasa, etc., isótopos radiactivos, tales como  $^{32}\text{P}$ ,  $^{35}\text{S}$ , etc., fluorocromos, tales como fluoresceína, etc., o partículas metálicas, para su detección directa mediante colorimetría, autorradiografía, fluorimetría, o metalografía, respectivamente.

55 La detección de los autoanticuerpos puede llevarse a cabo aplicando una única técnica o bien puede llevarse a cabo aplicando una combinación de dos o más técnicas; a modo ilustrativo, se pueden detectar unos autoanticuerpos mediante un ELISA y otros mediante una micromatriz de proteínas, o bien unos mediante un ELISA y otros mediante una micromatriz de fagos, o bien unos mediante una micromatriz de proteínas y otros mediante una micromatriz de fagos, etc.

60 En una divulgación particular, la muestra a analizar se pone en contacto con una única ECA seleccionada del grupo de ECAs (i)-(vi) mostradas en la Tabla 2, y sus combinaciones, en condiciones que permiten la formación de un complejo autoanticuerpo-ECA con el fin de identificar autoanticuerpos contra dicha ECA. En otra divulgación particular, dicha muestra biológica se pone en contacto con dos o más de dichas ECAs susceptibles de ser reconocidas por dichos autoanticuerpos, separadas entre sí, opcionalmente depositadas sobre un soporte



adecuado, con el fin de identificar autoanticuerpos contra dichas ECAs.

En una divulgación particular, el primer método divulgado en el presente documento comprende la detección de un autoanticuerpo contra una ECA que comprende la secuencia de aminoácidos mostrada en la SEQ ID NO: 4 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo, y, además, la detección de un autoanticuerpo seleccionado del grupo que consiste en: (i) un autoanticuerpo contra una ECA que comprende la secuencia de aminoácidos mostrada en la SEQ ID NO: 2 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo; (ii) un autoanticuerpo contra una ECA que comprende la secuencia de aminoácidos mostrada en la SEQ ID NO: 6 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo; (iii) un autoanticuerpo contra una ECA que comprende la secuencia de aminoácidos mostrada en la SEQ ID NO: 3 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo; (iv) un autoanticuerpo contra una ECA que comprende la secuencia de aminoácidos mostrada en la SEQ ID NO: 5 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo; (v) un autoanticuerpo contra una ECA que comprende la secuencia de aminoácidos mostrada en la SEQ ID NO: 1 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo; y (vi) cualquier combinación de autoanticuerpos (i) a (v). En una forma más concreta de dicha divulgación particular, dicho primer método de la divulgación comprende la detección de un autoanticuerpo contra una ECA que comprende la secuencia de aminoácidos mostrada en la SEQ ID NO: 4 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo, y, además, la detección de 1, 2, 3, 4, o 5 cualesquiera de dichos autoanticuerpos (i) a (v) previamente indicados.

En una divulgación concreta, dicha ECA es la proteína SULF1 o una variante o fragmento de la misma que contiene un epítipo reconocible por un autoanticuerpo, mientras que, en otra divulgación concreta, dicha ECA es un fago que comprende la secuencia de aminoácidos mostrada en la SEQ ID NO: 4, o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo, en donde dicha secuencia de aminoácidos está expuesta sobre la superficie del fago.

En otra divulgación particular, el primer método de la divulgación comprende la detección de un autoanticuerpo contra una ECA que comprende la secuencia de aminoácidos mostrada en la SEQ ID NO: 1 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo, y, además, la detección de un autoanticuerpo seleccionado del grupo que consiste en: (i') un autoanticuerpo contra una ECA que comprende la secuencia de aminoácidos mostrada en la SEQ ID NO: 4 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo; (ii') un autoanticuerpo contra una ECA que comprende la secuencia de aminoácidos mostrada en la SEQ ID NO: 2 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo; (iii') un autoanticuerpo contra una ECA que comprende la secuencia de aminoácidos mostrada en la SEQ ID NO: 6 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo; (iv') un autoanticuerpo contra una ECA que comprende la secuencia de aminoácidos mostrada en la SEQ ID NO: 3 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo; (v') un autoanticuerpo contra una ECA que comprende la secuencia de aminoácidos mostrada en la SEQ ID NO: 5 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo; y (vi') cualquier combinación de autoanticuerpos (i') a (v'). En una forma más concreta, dicho primer método de la divulgación comprende la detección de un autoanticuerpo contra una ECA que comprende la secuencia de aminoácidos mostrada en la SEQ ID NO: 1 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo, y, además, la detección de 1, 2, 3, 4, o 5 cualesquiera de dichos autoanticuerpos (i') a (v') previamente indicados.

La etapa b) del método de detección de autoanticuerpos de la divulgación comprende detectar la formación de un complejo autoanticuerpo-ECA. Esta etapa puede llevarse a cabo por métodos convencionales, conocidos por los expertos en la materia, para la detección de la formación de complejos de anticuerpo-antígeno (en este caso, autoanticuerpo-ECA).

En una divulgación particular, a modo ilustrativo, no limitante, para la detección de dicho complejo, puede añadirse un conjugado que comprende un anticuerpo que reconoce al autoanticuerpo y un marcador (anticuerpo secundario marcado) en condiciones que permiten la formación de un complejo (autoanticuerpo-ECA)-anticuerpo/marcador y detectar la formación de dicho complejo. Si la muestra biológica contiene autoanticuerpos contra una o más de dichas ECAs, entonces se habrá formado previamente el complejo autoanticuerpo-ECA, con lo que, cuando se pone en contacto dicho complejo con dicho conjugado que comprende el anticuerpo y el marcador, en las condiciones adecuadas, se forma el complejo (autoanticuerpo-ECA)-anticuerpo/marcador, que será visualizado mediante la técnica apropiada dependiendo del marcador utilizado, tal como se menciona más adelante; mientras que, en caso contrario, es decir, cuando la muestra biológica no contiene autoanticuerpos contra dicha(s) ECA(s) entonces no se formará dicho complejo (autoanticuerpo-ECA)-anticuerpo/marcador. Las condiciones adecuadas para que tenga lugar la formación de este último complejo son conocidas por los expertos en la materia.

Prácticamente cualquier reactivo indicador que permita detectar dicho complejo (autoanticuerpo-ECA)-anticuerpo/marcador puede ser utilizado en la puesta en práctica de la presente divulgación. A modo ilustrativo, no limitante, dicho marcador puede ser una enzima que cataliza una reacción detectable (por ejemplo, peroxidasa, glucosidasa, fosfatasa alcalina, glucosa-6-fosfato deshidrogenasa,  $\beta$ -galactosidasa,  $\beta$ -glucosidasa,  $\beta$ -glucuronidasa, etc.), un compuesto que genera una señal cuando forma parte de dicho complejo (por ejemplo, un compuesto



fluorescente o fluoróforo, tal como fluoresceína, rodamina, etc.; un compuesto (químico)luminiscente, tal como un dioxetano, un acridinio, un fenantridinio, rutenio, luminol, etc.), un elemento radiactivo (por ejemplo, azufre, yodo, etc.), etc. En una divulgación particular, dicho marcador es una peroxidasa. La selección de un marcador particular no es crítica, siempre y cuando sea capaz de producir una señal por sí mismo o conjuntamente con unas o más sustancias adicionales. De este modo, el complejo (autoanticuerpo-ECA)-anticuerpo/marcador formado puede ser detectado o visualizado por cualquier técnica apropiada, dependiendo del marcador elegido, conocida por los expertos en la materia, utilizando los dispositivos apropiados, por ejemplo, mediante técnicas basadas en métodos colorimétricos, fluorimétricos, (químico)luminiscentes, radiactivos, etc., todas ellas conocidas por los expertos en la materia.

El conjugado que comprende dicho anticuerpo que reconoce dicho autoanticuerpo y dicho marcador puede obtenerse por métodos convencionales conocidos por los técnicos en la materia.

A modo ilustrativo, cuando el marcador es una enzima, la detección del complejo en cuestión puede llevarse a cabo poniendo en contacto dicho complejo con un sustrato apropiado y, opcionalmente, con los activadores y/o agentes de amplificación enzimáticos apropiados. Ejemplos ilustrativos, no limitativos, de dichos sustratos incluyen:

- Para la fosfatasa alcalina:  
Cromogénico: sustratos basados en fosfato de p-nitrofenilo (p-NPP), fosfato de 5-bromo-4-cloro-3-indolilo /nitroazul tetrazolio (BCIP/NPT), etc.  
Fluorogénico: fosfato de 4-metilumbelifenilo (4-MUP), 2-(5'-cloro-2'-fosforiloxifenil)-6-cloro-4-(3H)-quinazolinona (CPPCQ), difosfato de 3,6-fluoresceína (3,6-FDP), etc.
- Para peroxidadas:  
Cromogénico: sustratos basados en ácido 2,2-azino-bis(3-etilbenzotiazolin-6-sulfónico) (ABTS), o-fenilendiamina (OPT), 3,3',5,5'-tetrametilbenzidina (TMB), o-dianisidina, ácido 5-aminosalicílico, ácido 3-dimetilaminobenzoico (DMAB) y 3-metil-2-benzotiazolinhidrazona (MBTH), 3-amino-9-etilcarbazol (AEC) y tetracloruro de 3,3'-diaminobenzidina (DAB), etc.  
Fluorogénico: ácido 4-hidroxi-3-metoxifenilacético, fenoxazinas reducidas y benzotiazinas reducidas, incluyendo el reactivo Amplex® Red, Amplex UltraRed, dihidroxantenos reducidos, etc.
- Para glucosidasas:  
Cromogénico: sustratos basados en o-nitrofenil-β-D-galactósido (o-NPG), p-nitrofenil-β-D-galactósido y 4-metilumbelifenil-β-D-galactósido (MUG) para β-D-galactosidasa, etc.  
Fluorogénico: resorufin β-D-galactopiranosido, fluoresceín digalactósido (FDG), fluoresceín diglucurónido, 4-metilumbeliferil beta-D-galactopiranosido, carboxiumbeliferil beta-D-galactopiranosido, cumarin beta-D-galactopiranosidos fluorados, etc.

Mediante la puesta en práctica del primer método de la divulgación, es posible detectar y obtener autoanticuerpos contra las ECAs indicadas en la Tabla 2. Adicionalmente, si se desea, se podría determinar el nivel o cantidad (cuantificar) de dichos autoanticuerpos contra dichas ECAs presentes en la muestra en estudio ya que la señal generada por algunos marcadores (por ejemplo, enzimas, etc.) es proporcional a la cantidad de autoanticuerpo presente en dicha muestra.

Opcionalmente, si se desea, se puede aislar el complejo autoanticuerpo-ECA mediante técnicas convencionales, por ejemplo, mediante el empleo de técnicas de inmunoprecipitación, etc., y, posteriormente, secuenciar el autoanticuerpo que se une a la ECA mediante el empleo de métodos de proteómica estándar convencionales, por ejemplo, mediante la determinación de la huella peptídica o análisis MS/MS (Vikas Dhingraa, et al. 2005. *International Journal of Pharmaceutics* 299 (1-2):1-18; Hanash SM et al. *Nature*. 3 Abr. 2008;452(7187):571-9).

Según el primer método de la divulgación, la detección del complejo autoanticuerpo-ECA es indicativa de la presencia del correspondiente autoanticuerpo (o autoanticuerpos) específico(s) contra dicha(s) ECA(s) en la muestra analizada y, por tanto, en el sujeto analizado.

En una divulgación particular, la formación de dicho complejo autoanticuerpo-ECA en dicha muestra se puede correlacionar con un diagnóstico de CCR en el sujeto del que procede la muestra analizada, o con el pronóstico de dicha enfermedad, o con el seguimiento de la evolución de dicha enfermedad. En el sentido utilizado en esta descripción, el término "correlacionar" se refiere a comparar la presencia o cantidad del indicador en un sujeto (por ejemplo, un sujeto sospechoso de padecer CCR) con su presencia o cantidad en sujetos que padecen dicha enfermedad (CCR), o que están predispuestos a desarrollarla, o en sujetos libres de dicha enfermedad.

El término "diagnóstico", tal como aquí se utiliza, se refiere, en general, al proceso por el cual se identifica una enfermedad, entidad nosológica, síndrome, o cualquier estado de salud-enfermedad. En particular, el término "diagnóstico de cáncer colorrectal (o CCR)" se refiere a la capacidad de identificar o detectar la presencia de CCR; esta detección, tal y como la entiende un experto en la materia, no pretende ser correcta en un 100% de las muestras analizadas. Sin embargo, requiere que una cantidad estadísticamente significativa de las muestras analizadas sean clasificadas correctamente. La cantidad que es estadísticamente significativa puede ser establecida por un experto en la materia mediante el uso de diferentes herramientas estadísticas; ejemplos ilustrativos, no

limitantes, de dichas herramientas estadísticas incluyen la determinación de intervalos de confianza, la determinación del valor p, la prueba de la t de Student o las funciones discriminantes de Fisher, etc. (véase, por ejemplo, Dowdy y Wearden, *Statistics for Research*, John Wiley & Sons, Nueva York 1983). Preferiblemente, los intervalos de confianza son al menos del 90%, al menos del 95%, al menos del 97%, al menos del 98% o al menos del 99%. Preferiblemente, el valor p es menor de 0,1, de 0,05, de 0,01, de 0,005 o de 0,0001. Ventajosamente, las enseñanzas de la presente divulgación permiten detectar correctamente CCR en al menos el 50%, preferentemente en al menos el 60%, más preferentemente en al menos el 70%, aún más preferentemente en al menos el 80%, o todavía más preferentemente en al menos el 90% de los sujetos de un determinado grupo o población analizada.

En otra divulgación particular, el primer método divulgado en el presente documento comprende comparar la inmunorreactividad de la muestra analizada con la inmunorreactividad de una segunda muestra procedente del mismo sujeto en un periodo de tiempo posterior. De este modo, es posible evaluar la evolución de la enfermedad o bien evaluar la eficacia del tratamiento en caso de que dicha segunda muestra haya sido obtenida después de que el sujeto haya sido tratado para CCR. En el sentido utilizado en esta descripción, el término "inmunorreactividad" se refiere a la presencia o nivel de unión de un anticuerpo o anticuerpos en una muestra a uno o más antígenos diana, por ejemplo, las ECAs de la Tabla 2. Un "patrón de inmunorreactividad" se refiere a un perfil de unión de anticuerpos en una muestra (autoanticuerpos) a una pluralidad de antígenos diana (por ejemplo, las ECAs de la Tabla 2).

En otra divulgación, el primer método divulgado en el presente documento comprende, además, analizar la presencia uno o más marcadores adicionales de CCR, por ejemplo, CEA o autoanticuerpos contra las proteínas Pim1, SRC, MAPKAPK3, FGFR4, STK4 y/o ACVR2B.

#### Método de detección de autoanticuerpos

Se divulga también un método para detectar un autoanticuerpo en una muestra, que comprende:

(a) poner en contacto una muestra con una entidad capturadora de anticuerpo (ECA), en donde dicha ECA se selecciona del grupo que consiste en:

- (i) una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 4 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo;
- (ii) una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 2 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo;
- (iii) una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 6 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo;
- (iv) una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 3 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo, en donde dicha ECA no es la proteína MST1;
- (v) una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 5 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo;
- (vi) una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 1 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo; y
- (vii) cualquier combinación de dichas ECAs (i)-(vi); y

(b) detectar la formación de un complejo autoanticuerpo-ECA,

en donde la detección de dicho complejo autoanticuerpo-ECA es indicativa de la presencia de dicho autoanticuerpo en dicha muestra. La invención se refiere a un método para detectar un autoanticuerpo en una muestra, método de la invención, que comprende:

a) poner en contacto una muestra de un sujeto con una entidad de captura de anticuerpo (ECA), en donde dicha ECA es una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 4 o un péptido o proteína que tiene al menos el 90% de identidad de secuencia con respecto a SEQ ID NO: 4 y que está reconocida por un autoanticuerpo contra SEQ ID NO: 4;

y  
b) detectar la formación de un complejo entre un autoanticuerpo y dicha ECA,

en donde la detección de dicho complejo autoanticuerpo-ECA es indicativa de la presencia de dicho autoanticuerpo en dicha muestra, en donde la muestra es sangre, plasma o suero.

En general, las características de dicha ECA a la que se hace referencia en el método de la invención son las mismas que las características de la ECA a las que se hace referencia en el primer método de la divulgación. Las técnicas para detectar los autoanticuerpos según el método de la invención son las mismas que las mencionadas en relación con el primer método de la divulgación, por lo que se incorporan aquí mediante referencia.

Las divulgaciones particulares del primer método divulgado en el presente documento, así como las definiciones de los términos utilizados son aplicables también al método de la invención, por lo que se incorporan en este método de la invención por referencia.

5 Este método de la invención permite correlacionar los resultados obtenidos con aquellas patologías en las que se generan respuestas inmunitarias con la consiguiente producción de autoanticuerpos. Ejemplos ilustrativos, no limitantes, de dichas patologías incluyen algunos tipos de cáncer, por ejemplo, CCR, carcinoma hepatocelular (Imai, H et al. Intervirology 35:73-85), cáncer de mama, cáncer de próstata (Wang X et al. N Engl J Med. 2005; 353(12):1224-35), cáncer de pulmón, etc. y enfermedades autoinmunitarias.

10 En una realización particular del método de la invención, la muestra es una muestra de un sujeto que se sospecha tiene cáncer colorrectal (CCR). Más en particular, el método de la invención comprende además correlacionar la formación del complejo autoanticuerpo-ECA en la muestra del sujeto con un diagnóstico de CCR o comprende además comparar la inmunorreactividad de la muestra con la inmunorreactividad de una segunda muestra del mismo sujeto en un periodo de tiempo posterior. En particular, dicha segunda muestra del sujeto se ha obtenido después de que dicho sujeto haya sido tratado para CCR.

15 En una realización particular, el método de la invención comprende la detección de un autoanticuerpo contra una ECA que comprende la secuencia de aminoácidos mostrada en la SEQ ID NO: 4 o una proteína o péptido que tiene al menos el 90% de identidad de secuencia con respecto a SEQ ID NO:4 que contiene un epítipo reconocible por un autoanticuerpo contra SEQ ID NO: 4, y, además, la detección de un autoanticuerpo seleccionado del grupo que consiste en:

- 25 (i) un autoanticuerpo contra una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 2 o una proteína o péptido que tiene al menos el 90% de identidad de secuencia con respecto a SEQ ID NO:2 que contiene un epítipo reconocible por un autoanticuerpo contra SEQ ID NO: 2;
- (ii) un autoanticuerpo contra una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 6 o una proteína o péptido que tiene al menos el 90% de identidad de secuencia con respecto a SEQ ID NO:6 que contiene un epítipo reconocible por un autoanticuerpo contra SEQ ID NO:6;
- 30 (iii) un autoanticuerpo contra una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 3 o una proteína o péptido que tiene al menos el 90% de identidad de secuencia con respecto a SEQ ID NO:3 que contiene un epítipo reconocible por un autoanticuerpo contra SEQ ID NO:3;
- (iv) un autoanticuerpo contra una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 5 o una proteína o péptido que tiene al menos el 90% de identidad de secuencia con respecto a SEQ ID NO:5 que contiene un epítipo reconocible por un autoanticuerpo contra SEQ ID NO:5;
- 35 (v) un autoanticuerpo contra una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 1 o una proteína o péptido que tiene al menos el 90% de identidad de secuencia con respecto a SEQ ID NO:1 que contiene un epítipo reconocible por un autoanticuerpo contra SEQ ID NO:1; y
- 40 (vi) cualquier combinación de autoanticuerpos (i) a (v).

En otra realización preferida del método de la invención, dicha ECA es sulfatasa 1 (SULF1) o una proteína o péptido que tiene al menos el 90% de identidad de secuencia con respecto a SULF1 que contiene un epítipo reconocible contra SULF1,

45 o alternativamente,

dicha ECA es un fago que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 4 o una proteína o péptido que tiene al menos el 90% de identidad de secuencia con respecto a SEQ ID NO:4 que contiene un epítipo reconocible por un autoanticuerpo contra SEQ ID NO:4, en donde dicha secuencia de aminoácidos está expuesta en la superficie del fago.

50 Se divulga un método que comprende la detección de un autoanticuerpo contra una ECA que comprende la secuencia de aminoácidos mostrada en la SEQ ID NO: 1 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo, y, además, la detección de un autoanticuerpo seleccionado del grupo que consiste en:

- (i) un autoanticuerpo contra una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 4 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo;
- 60 (ii) un autoanticuerpo contra una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 2 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo;
- (iii) un autoanticuerpo contra una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 6 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo;
- (iv) un autoanticuerpo contra una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 3 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo;
- 65 (v) un autoanticuerpo contra una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 5 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo; y

(vi) cualquier combinación de autoanticuerpos (i) a (v).

Entidad capturadora de anticuerpos (ECA)

5 También se divulga una entidad capturadora de anticuerpos (ECA), seleccionada del grupo que consiste en:

- (i) una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 4 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo, en donde dicha ECA no es la proteína SULF1;
- 10 (ii) una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 2 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo, en donde dicha ECA no es la proteína GRN;
- (iii) una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 6 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo, en donde dicha ECA no es la proteína
- 15 GTF2i; y
- (iv) una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 3 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo, en donde dicha ECA no es la proteína MST1;
- (v) una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 5 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo, en donde dicha ECA no es la proteína
- 20 SREBF2;
- (vi) una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 1 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo, en donde dicha ECA no es la proteína NHSL1; y
- 25 (vii) cualquier combinación de ECAs (i) a (vi).

Una ECA, tal como se ha definido previamente, es una entidad macromolecular, por ejemplo, un péptido, una proteína o un fago, que se une específicamente a un anticuerpo (o autoanticuerpo). En una divulgación particular, dicha ECA comprende un péptido o una proteína que se une específicamente a un anticuerpo (o autoanticuerpo). Dicho péptido puede estar inmovilizado sobre un soporte o bien expuesto sobre la superficie de un fago. En una divulgación particular preferida, dicha ECA es un péptido, una proteína o un fago en cuya superficie se expone dicho péptido o dicha proteína. Dicha ECA puede estar, si se desea, inmovilizada sobre un soporte sólido.

La invención se refiere a una entidad de captura de anticuerpos (ECA) que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 4, en donde dicha ECA no es la proteína sulfatasa-1 (SULF1). En una realización preferida de la entidad de captura de anticuerpos (ECA) según la invención, dicha ECA es un fago que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 4, en donde dicha secuencia de aminoácidos está expuesta en la superficie del fago.

40 En una divulgación particular, la ECA se selecciona del grupo que consiste en:

- (i) un fago que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 1 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo, en donde dicha secuencia de aminoácidos está expuesta sobre la superficie del fago;
- 45 (ii) un fago que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 2 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo, en donde dicha secuencia de aminoácidos está expuesta sobre la superficie del fago;
- (iii) un fago que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 3 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo, en donde dicha secuencia de aminoácidos está expuesta sobre la superficie del fago;
- 50 (iv) un fago que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 4 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo, en donde dicha secuencia de aminoácidos está expuesta sobre la superficie del fago;
- (v) un fago que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 5 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo, en donde dicha secuencia de aminoácidos está expuesta sobre la superficie del fago;
- 55 (vi) un fago que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 6 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo, en donde dicha secuencia de aminoácidos está expuesta sobre la superficie del fago; y
- 60 (vii) cualquiera de las combinaciones de (i) a (vi).

Dichos fagos pueden ser obtenidos por métodos convencionales, conocidos por los expertos en la materia, y, de forma más concreta, mediante el procedimiento descrito en el Ejemplo 1.

65 Las ECAs divulgadas en el presente documento pueden ser utilizadas en la detección de anticuerpos o autoanticuerpos contra dichas ECAs en una muestra, en particular, contra las secuencias aminoácidas



identificadas como SEQ ID NO: 1-6 presentes en dichas ECAs, y la presencia de dichos autoanticuerpos en dicha muestra puede ser correlacionada con el diagnóstico, pronóstico, seguimiento de la evolución, o eficacia del tratamiento, de una enfermedad, para la que dichos autoanticuerpos sean marcadores, por ejemplo, CCR.

## 5 Composición de la invención

También se divulga la composición 1 de la divulgación, que comprende una ECA divulgada en el documento. Como se ha indicado previamente, en una divulgación particular, dicha ECA puede ser un péptido, una proteína o un fago.

10 En una divulgación particular, la composición 1 de la divulgación comprende, al menos, una ECA divulgada en el presente documento. En otra divulgación particular, la composición 1 de la divulgación comprende al menos 2 ECAs de la divulgación, por ejemplo, 2, 3, 4, 5 o incluso las 6 ECAs de la divulgación.

15 En otra divulgación particular, la composición 1 comprende una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 4 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo, y, además, al menos una ECA seleccionada del grupo que consiste en:

- (i) una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 2 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo;
- 20 (ii) una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 6 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo;
- (iii) una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 3 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo;
- (iv) una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 5 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo;
- 25 (v) una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 1 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo; y
- (vi) cualquier combinación de ECAs (i) a (v).

30 En otra divulgación particular, la composición 1 de la divulgación comprende una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 1 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo, y, además, al menos una ECA seleccionada del grupo que consiste en:

- (i) una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 2 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo;
- 35 (ii) una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 3 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo;
- (iii) una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 4 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo;
- 40 (iv) una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 5 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo;
- (v) una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 6 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo; y
- (vi) cualquier combinación de ECAs (i) a (v).

45 En otra divulgación particular, la composición 1 de la divulgación comprende, al menos, una ECA divulgada en el presente documento y, al menos, una proteína seleccionada del grupo que consiste en la proteína SULF1 o una variante de la misma, la proteína MST1 o una variante de la misma, y sus combinaciones. El experto en la materia apreciará que es posible usar proteínas SULF1 o MST1 de distintas especies; no obstante, en una divulgación preferida, la composición 1 de la divulgación incluye proteínas SULF1 o MST1 de origen humano, tal como la proteína SULF1 humana, cuyo número de acceso en la base de datos NCBI (versión de 1 de mayo de 2011) es EAW86954.1 y su secuencia de aminoácidos es SEQ ID NO: 10, y la proteína MST1 humana cuyo número de acceso en la base de datos NCBI (versión de 1 de mayo de 2011) es AAA83254.1 y su secuencia de aminoácidos es SEQ ID NO: 9. El término "variante" ya ha sido definido previamente en el apartado "Definiciones".

55 En otra divulgación particular, la composición 1 de la divulgación comprende:

a) una ECA seleccionada del grupo que consiste en:

- 60 (i) una ECA que comprende la secuencia de aminoácidos mostrada en la SEQ ID NO: 4 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo,
- (ii) una ECA que comprende la secuencia de aminoácidos mostrada en la SEQ ID NO: 2 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo;
- 65 (iii) una ECA que comprende la secuencia de aminoácidos mostrada en la SEQ ID NO: 6 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo;



- (iv) una ECA que comprende la secuencia de aminoácidos mostrada en la SEQ ID NO: 3 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo;
- (v) una ECA que comprende la secuencia de aminoácidos mostrada en la SEQ ID NO: 5 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo;
- (vi) una ECA que comprende la secuencia de aminoácidos mostrada en la SEQ ID NO: 1 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo; y
- (vii) cualquier combinación de ECAs (i) a (vi); y

b) una proteína seleccionada del grupo formado por la proteína SULF1 o una variante de la misma, la proteína MST1 o una variante de la misma, y sus combinaciones.

En una divulgación preferida, dicha composición 1 de la divulgación comprende:

a) una ECA seleccionada del grupo que consiste en:

- (i) una ECA que comprende la secuencia de aminoácidos mostrada en la SEQ ID NO: 2 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo,
- (ii) una ECA que comprende la secuencia de aminoácidos mostrada en la SEQ ID NO: 6 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo;
- (iii) una ECA que comprende la secuencia de aminoácidos mostrada en la SEQ ID NO: 5 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo;
- (iv) una ECA que comprende la secuencia de aminoácidos mostrada en la SEQ ID NO: 1 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo; y
- (v) cualquier combinación de ECAs (i) a (iv); y

b) una proteína seleccionada del grupo formado por la proteína SULF1 o una variante de la misma, la proteína MST1 o una variante de la misma, y sus combinaciones.

La composición de la invención comprende una entidad de captura de anticuerpo (ECA) según la invención, composición 1 de la invención. La composición 1 de la divulgación ha proporcionado unos buenos resultados en el diagnóstico de CCR [Ejemplo 4].

En una realización preferida, la composición de la invención comprende además al menos una ECA según la invención, y al menos una proteína seleccionada del grupo que consiste en la proteína sulfatasa 1 (SULF1) o una proteína o péptido que tiene al menos el 90% de identidad de secuencia con respecto a SULF1 y reconocible por un autoanticuerpo contra SULF1, la proteína MST1 o una proteína o péptido que tiene al menos el 90% de identidad de secuencia con respecto a MST1 y reconocible por un autoanticuerpo contra MST1, y sus combinaciones.

También se divulga la composición 2 de la divulgación, que comprende la proteína SULF1 o una variante de la misma y la proteína MST1 o una variante de la misma. Las características de dichas proteínas SULF1 y MST1, y de sus variantes, han sido mencionadas previamente.

En una divulgación particular, la composición 2 de la divulgación comprende, además, al menos, una ECA de la divulgación. En otra divulgación particular, la composición 2 de la divulgación comprende al menos 2 ECAs de la divulgación.

En otra divulgación particular, dicha composición 2 de la divulgación comprende:

- a) la proteína SULF1 o una variante de la misma;
- b) la proteína MST1 o una variante de la misma; y
- c) una ECA seleccionada del grupo que consiste en:

- (i) una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 2 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo,
- (ii) una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 6 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo;
- (iii) una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 5 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo;
- (iv) una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 1 o una variante de la misma que contiene un epítipo reconocible por un autoanticuerpo; y
- (v) cualquier combinación de ECAs (i) a (iv).

La invención también se refiere a una composición, composición 2 de la invención, que comprende proteína sulfatasa 1 (SULF1) o una proteína o péptido que tiene al menos el 90% de identidad de secuencia con respecto a SULF1 y reconocible por un autoanticuerpo contra SULF1 y proteína MST1 o una proteína o péptido que tiene al

menos el 90% de identidad de secuencia con respecto a MST1 y reconocible por un autoanticuerpo contra MST1 y una entidad de captura de anticuerpo (ECA) según la invención.

5 En una realización particular, tanto la composición 1 de la invención como la composición 2 de la invención están apoyadas sobre un soporte sólido.

Kit de la invención y aplicaciones

10 En otro aspecto, la invención se relaciona con un kit, en adelante kit 1 de la invención, que comprende una composición de la invención. En una realización particular, dicha composición de la invención es la composición 1 de la invención. En otra realización particular, dicha composición de la invención es la composición 2 de la invención.

En otro aspecto, la invención se refiere al uso in vitro del kit 1 de la invención para:

- 15
- detectar un anticuerpo en una muestra,
  - detectar un autoanticuerpo en un sujeto sospechoso de padecer cáncer colorrectal (CCR),
  - diagnosticar si un sujeto padece CCR,
  - determinar el riesgo de un sujeto a desarrollar CCR,
  - seguir la evolución de CCR en un sujeto,

20

  - evaluar la eficacia de un tratamiento contra CCR, o
  - predecir la supervivencia de un sujeto que padece CCR.

Para dichas aplicaciones, el kit 1 de la invención incluirá los reactivos necesarios para detectar autoanticuerpos contra, al menos, una ECA de la Tabla 2.

25 En una realización particular, el kit 1 de la invención comprende una ECA de la invención. En otra realización particular, el kit de la invención comprende la composición 1 de la invención. En otra realización particular, el kit de la invención comprende la composición 2 de la invención. Se divulga un kit 1 que comprende una proteína seleccionada del grupo formado por la proteína SULF1 o una variante de la misma, la proteína MST1 o una variante de la misma, y sus combinaciones.

30

En otro aspecto, la invención se refiere a un kit, en adelante kit 2 de la invención, que comprende un reactivo para detectar la proteína SULF1 o una proteína o péptido que tiene al menos el 90% de identidad de secuencia con respecto a SULF1 y reconocible por un autoanticuerpo contra SULF1. En una realización particular, dicho kit 2 de la invención comprende, además, un reactivo para detectar la proteína MST1 o una proteína o péptido que tiene al menos el 90% de identidad de secuencia con respecto a MST1 y reconocible por un autoanticuerpo contra MST1. En otra realización particular, el kit 2 de la invención comprende una ECA seleccionada del grupo de ECAs mencionadas en la Tabla 2.

35

40 El kit 2 de la invención puede ser utilizado en las mismas aplicaciones que el kit 1 de la invención.

Tanto el kit 1 de la invención como el kit 2 de la invención pueden contener, además, todos aquellos reactivos necesarios para detectar la cantidad de autoanticuerpos contra las ECAs definidas anteriormente, o contra las proteínas SULF1, MST1 o sus variantes, como, por ejemplo, pero sin limitarse, a

- 45
- anticuerpos secundarios marcados con un marcador que reconocen específicamente los complejos autoanticuerpo-ECA;
  - sustratos para los marcadores presentes en dichos anticuerpos secundarios marcados; y
  - controles positivos y/o negativos.

50

Asimismo, dichos kits 1 y 2 de la invención pueden, además, incluir, sin ningún tipo de limitación, tampones, agentes para prevenir la contaminación, inhibidores de la degradación de proteínas, etc. Por otro lado, el kit de la invención puede incluir todos los soportes y recipientes necesarios para su puesta en práctica y optimización. Preferiblemente, el kit comprende además instrucciones para su utilización.

55

Método de diagnóstico de CCR

Adicionalmente, los autores de la presente invención han encontrado que la sobreexpresión de la proteína SULF1 se correlaciona con CCR, tal como se muestra en la Figura 3.

60

Por tanto, también se divulga un método para diagnosticar si un sujeto padece cáncer colorrectal (CCR), que comprende determinar el nivel de la proteína SULF1 en una muestra de dicho sujeto, en donde si el nivel de dicha proteína SULF1 es mayor que el nivel de la proteína SULF1 de una muestra de referencia, es indicativo de que el sujeto padece CCR.

65

Para la puesta en práctica de este método divulgado en el presente documento, la muestra puede ser, preferentemente, una muestra de tejido, tal como una muestra de tejido de colon o tumoral.

El término “diagnóstico” ya ha sido definido previamente.

Los métodos para la determinación del nivel (concentración) de una proteína son bien conocidos por un experto en la materia e incluyen numerosas alternativas. Prácticamente cualquier método que permita la determinación del nivel (cuantificación) de la proteína SULF1 puede ser utilizado en la puesta en práctica del método divulgado en el presente documento.

En una divulgación particular, el nivel de la proteína SULF1 es cuantificado mediante un método convencional que permite detectar y cuantificar dicha proteína en una muestra a estudiar, tal como una muestra de un sujeto. La determinación del nivel de dicha proteína SULF1 se puede realizar, a modo ilustrativo, no limitante, mediante un inmunoensayo, por ejemplo, ELISA, etc., mediante resonancia magnética nuclear (RMN) o mediante cualquier otra técnica apropiada conocida en el estado de la técnica. En una divulgación preferida, la determinación del nivel de proteínas se realiza mediante un inmunoensayo. En una realización particular preferida, dicho inmunoensayo es una inmunotransferencia (transferencia Western o inmunodetección en membrana). Para ello, brevemente, se obtiene un extracto de proteínas de una muestra biológica aislada de un sujeto y se separa la proteína mediante electroforesis en un medio soporte capaz de retenerla. Una vez separadas las proteínas se transfieren a un soporte diferente o membrana donde pueden ser detectadas mediante el uso de anticuerpos específicos que reconocen a la proteína en cuestión (SULF1). Dicha membrana se hibrida con un primer anticuerpo específico (o anticuerpo primario) que reconoce la proteína SULF1. A continuación, la membrana se hibrida con un segundo anticuerpo (o anticuerpo secundario) que reconoce de manera específica al anticuerpo primario y que se encuentra conjugado o unido con un compuesto marcador. En una divulgación alternativa, el anticuerpo que reconoce a la proteína SULF1 está conjugado o unido a un compuesto marcador, y no es necesario el uso de un anticuerpo secundario. Se conocen diferentes formatos, soportes y técnicas que pueden ser empleados para la realización de esta divulgación preferida del método divulgado en el presente documento.

En otra divulgación particular preferida, el inmunoensayo comprende un ensayo inmunohistoquímico. Las técnicas de inmunohistoquímica permiten la identificación, en muestras tisulares o citológicas, de determinantes antigénicos característicos. El análisis mediante inmunohistoquímica (IHQ) se realiza sobre cortes de tejido, ya sea congelado o incluido en parafina, procedente de una muestra biológica aislada de un sujeto. Estos cortes se hibridan con un anticuerpo específico o anticuerpo primario que reconoce a anticuerpos específicos que reconocen a la proteína SULF1. A continuación, los cortes se hibridan con un anticuerpo secundario capaz de reconocer de manera específica el anticuerpo primario y que se encuentra conjugado o unido con un compuesto marcador. En una divulgación alternativa, el anticuerpo que reconoce a la proteína SULF1 está conjugado o unido a un compuesto marcador, y no es necesario el uso de un anticuerpo secundario.

A modo ilustrativo, no limitante, el “nivel de la proteína SULF1”, se refiere, pero no se limita, a la cantidad cuantificable, semicuantificable, o relativa de dicha proteína SULF1, así como a cualquier otro valor o parámetro relacionado con dicha proteína o que pueda derivarse de la misma. Dichos valores o parámetros comprenden valores de intensidad de la señal obtenidos a partir de cualquiera de las propiedades físicas o químicas de dicha proteína obtenidos bien mediante medida directa, por ejemplo, valores de intensidad de espectroscopía de masas, resonancia magnética nuclear, etc., o bien mediante medida indirecta, por ejemplo, mediante cualquiera de los sistemas de medida descritos en este documento, por ejemplo, mediante la medida obtenida de un componente secundario o un sistema de medida biológica (por ejemplo, la medida de respuestas celulares, ligandos, “etiquetas” o productos de reacción enzimática). El nivel de la proteína SULF1, determinado en una muestra, tal como una muestra biológica procedente del sujeto sometido a estudio, se dice que es “mayor” que el nivel de referencia de dicha proteína SULF1 cuando, según la divulgación, el nivel de dicha proteína en la muestra a analizar es de, al menos 1,5 veces, 5 veces, 10 veces, 20 veces, 30 veces, 40 veces, 50 veces, 60 veces, 70 veces, 80 veces, 90 veces, 100 veces o incluso más, con respecto al nivel de referencia de dicha proteína.

El término “nivel de referencia”, tal como aquí se utiliza, se refiere, en general, al nivel de un producto, por ejemplo, la proteína SULF1, presente en sujetos control. En una divulgación particular, dichos sujetos control son sujetos que no padecen una determinada enfermedad (por ejemplo, CCR), mientras que en otra divulgación particular, dicho sujeto control es el propio sujeto en estudio, lo que resulta particularmente útil para evaluar el seguimiento de una enfermedad (por ejemplo, CCR) o para evaluar la eficacia de un tratamiento contra dicha enfermedad (por ejemplo, CCR), etc., para lo cual el nivel de referencia de un producto dado puede ser el nivel de dicho producto determinado en una muestra del mismo sujeto en estudio pero tomada unos días, semanas, meses o incluso años antes con el fin de evaluar el seguimiento de la enfermedad, o bien tomada un antes de, por ejemplo, la aplicación al sujeto de un tratamiento contra dicha enfermedad con el fin de evaluar su eficacia.

Debido a la variabilidad que pueda producirse entre los diferentes sujetos en cuanto a la producción de proteína SULF1, el nivel de referencia podría obtenerse a partir de un conjunto de muestras procedente de una población de sujetos sanos (por ejemplo, sujetos que no padecen CCR) y calculando el nivel medio del producto en cuestión (proteína SULF1) en dicha población de sujetos sanos.

El nivel de referencia de un determinado producto, por ejemplo, la proteína SULF1, puede ser determinado a partir de una muestra de referencia que puede ser analizada, por ejemplo, simultánea o consecutivamente, junto con la muestra a analizar (muestra problema). Generalmente, el nivel de referencia puede derivarse de los límites de distribución normal de una cantidad fisiológica encontrada en una población de sujetos control. Dicha cantidad fisiológica puede ser determinada por varias técnicas bien conocidas, tal como se describe en esta descripción.

Dicho nivel de referencia, según la presente divulgación, permite discriminar la presencia de CCR y, por tanto, puede ser utilizado en el diagnóstico, pronóstico o seguimiento de la evolución de CCR.

Los marcadores y métodos proporcionados por la presente divulgación son adecuados para diagnosticar CCR, así como para predecir el desarrollo de CCR, seguir la evolución de CCR y/o evaluar la eficacia del tratamiento administrado a un sujeto que padece CCR.

Ensayos realizados por los inventores han puesto de manifiesto que los fagos identificados como MST1, SULF1, NHSL1, SREBF2, GRN y GTF2i permiten discriminar entre sueros tumorales (CCR) y sueros controles por ELISA. Aunque la sensibilidad era relativamente baja para los fagos individuales, oscilando entre el 46% y el 60%, la especificidad era mayor, entre el 50% y el 73,9%, y, además, la combinación de los 6 fagos como predictor de CCR proporcionó un AUC de 0,82 con una sensibilidad y una especificidad del 70% y el 73,9% [Ejemplo 4].

Por otra parte, la combinación de las proteínas recombinantes SULF1 y MST1 junto con los fagos NHSL1, GRN, SREBF2 y GTF2i permitió predecir la enfermedad (CCR) con una sensibilidad del 72% y una especificidad del 87%, con un AUC de 0,83. Si además se tiene en cuenta la edad del paciente, el AUC es de 0,91. Además, estos marcadores permiten agrupar los resultados discriminando, no sólo los individuos enfermos, sino también los distintos estadios de la enfermedad. Asimismo, los autoanticuerpos contra MST1 y NHSL1 están asociados con el pronóstico clínico de los pacientes con CCR. Por tanto, la detección de este panel de autoanticuerpos en el suero es, por tanto, un método sencillo y no invasivo para el diagnóstico/pronóstico del CCR.

Adicionalmente, se ha observado que la sobreexpresión de la proteína SULF1 está correlacionada con el diagnóstico de CCR [Figura 3].

## Ejemplos

A continuación, se ilustrará la invención mediante ensayos realizados por los inventores, que muestran claramente la especificidad y eficacia del método para el diagnóstico/pronóstico de CCR basado en la detección en suero de anticuerpos contra los autoantígenos tumorales descritos.

### Ejemplo 1

#### Análisis de sueros de pacientes con CCR con micromatrices impresas con fagos T7

##### I. MATERIALES Y MÉTODOS

##### Sueros de CCR y de referencia

Los sueros usados en los análisis de micromatrices y de supervivencia se obtuvieron de pacientes del Hospital Universitario de Bellvitge, Instituto Catalán de Oncología de Barcelona, Hospital Puerta de Hierro de Madrid y Hospital de Cabueñes de Gijón, tras obtener el consentimiento escrito de todos los pacientes incluidos en el estudio.

Para la selección de las genotecas de fagos T7 específicos de CCR se usaron 3 sueros de pacientes con CCR de estadio B de Duke, 3 de estadio C y 6 de estadio D (3 con metástasis en hígado y 3 con metástasis en pulmón). Para el cribado mediante micromatrices se seleccionaron 15 sueros de pacientes con CCR de diferentes estadios, con una edad media de 66,3 años (intervalo de edad 54-82) y 15 sueros de individuos control con similar edad media y la misma proporción de sexos que los pacientes con CCR.

Para el análisis de supervivencia, se ensayó otro panel de 95 sueros de CCR con más de 10 años de seguimiento. La mediana de la edad era 66,2 años (intervalo entre 23-90 años). Los datos clínicos de todos los pacientes se recogen en la Tabla 3. Todas las muestras se manejaron de forma anónima de acuerdo a las normas éticas y jurídicas del Consejo Superior de Investigaciones Científicas (CSIC).

Tabla 3. Información clínico-patológica de los pacientes cuyo suero fue utilizado para la identificación y validación de los autoanticuerpos.

Pacientes con CCR (n)	160	Donantes sanos (n)	61
Media de edad (años)	67,7	Media de edad (años)	61,7
Intervalo de edad (años)	23-91	Intervalo de edad (años)	34-89
Sexo:		Sexo:	
Hombre	65,6%	Hombre	60,7%
Mujer	34,4%	Mujer	39,3%
Estadio de Duke:			
I	40,6%		
II	19,4%		
III	15,6%		
IV	24,4%		
Pronóstico:			
Muerto	33,1%		
Vivo	60%		
Desconocido	6,9%		
Tiempo medio de supervivencia (meses)	57,5		

- 5 Para la validación, se usó un grupo independiente de sueros; 50 sueros de CCR representativos de todos los estadios de Duke (A-D), 46 sueros controles, 10 pacientes asintomáticos con antecedentes familiares, 2 con pólipos hiperplásicos, 2 con colitis ulcerosa y 43 sueros de otros tipos de cáncer (vejiga, mama, pulmón, páncreas y estómago).
- 10 Los sueros utilizados se procesaron de la misma manera en los diferentes hospitales; las muestras de sangre se dejaron a temperatura ambiente durante al menos 30 min (y un máximo de 60 min) para permitir la formación del coágulo. Posteriormente, las muestras se centrifugaron a 3.000 g a 4°C durante 10 min. Los sueros se congelaron y almacenaron a -80°C hasta su utilización.
- 15 Síntesis de la genoteca de ADNc en fagos T7 y rondas de selección
- El ARN total de 3 tejidos tumorales de pacientes con CCR en estadios de Duke A (muestras A) y 3 de estadio C (muestras C), se aisló mediante el reactivo Trizol (Invitrogen). Se utilizaron 4 microgramos de cada ARN, mezclando por separado las muestras A y las muestras C, para la síntesis del ADNc. Las genotecas de ADNc de CCR en fagos T7 se construyeron utilizando el sistema de síntesis y clonación de ADNc OrientExpress (Novagen) según las instrucciones del fabricante. El ADNc se sintetizó mediante RT-PCR usando el oligonucleótido oligo(dT). La clonación se realizó indistintamente en los vectores T7Select 415-1 y T7Select 10-3b que difieren en el tamaño de inserto que admiten. El crecimiento de los fagos se obtuvo en las cepas BL21 y BLT 5403 de *E. coli*, respectivamente. Las cuatro genotecas T7 se titularon haciendo diluciones seriadas de los fagos T7 en placas Petri.
- 20 Los tamaños de las genotecas fueron superiores a 10<sup>6</sup> ufp/ml en todos los casos. Las rondas de selección de los fagos se realizaron usando por un lado la combinación de las 2 genotecas construidas con el vector T7Select 415-1 y por otro la combinación de las 2 genotecas construidas con el vector T7Select 10-3b.
- 25 En primer lugar, se realizó una selección negativa. Para ello, se incubaron partículas magnéticas acopladas a proteína A/G (Invitrogen) con una mezcla de 8 sueros controles (120 µl de mezcla de sueros controles, diluido 1:50, a 4° C toda la noche) para unir las IgG de los sujetos control. Posteriormente, se incubaron los fagos con dichas partículas magnéticas para eliminar aquellos fagos unidos a las IgG de los sueros control. En segundo lugar se incubaron 4 mezclas de sueros (mezcla estadio B: de 3 pacientes con CCR de estadio de Duke B, mezcla estadio C: de 3 pacientes de estadio de Duke C, mezcla estadio D-H: de 3 pacientes con estadio de Duke D y metástasis a hígado y mezcla estadio D-P: de 3 pacientes con estadio de Duke D y metástasis a pulmón) con partículas magnéticas acopladas a proteína A/G para enriquecer las genotecas de fagos en fagos específicos de CCR. Los fagos no retenidos en la selección negativa se incubaron con las partículas magnéticas previamente incubadas con las diferentes mezclas de sueros de pacientes con CCR. Los fagos unidos a dichas partículas magnéticas se eluyeron con 100 µl de SDS al 1% y se amplificaron en BLT5406 o BL21 de *E. coli*. En total se hicieron 4 rondas de selección para enriquecer las genotecas de fagos T7 en fagos específicos de CCR. Finalmente, se amplificaron clones individuales de las 8 selecciones que se imprimieron en micromatrices de nitrocelulosa.
- 30
- 35
- 40



Micromatrices de fagos

Tras la amplificación de los fagos monoclonales, los lisados de bacterias se centrifugaron y los sobrenadantes que contenían los fagos se diluyeron 1:2 en PBS con Tween 20 al 0,1% (PBST) para su impresión por duplicado en micromatrices de nitrocelulosa (Whatman/Schleicher & Schuell's) con el robot OmniGrid Spotter (GeneMachines, San Carlos, CA). Los controles negativos empleados en la impresión fueron BSA (Sigma Aldrich), tampón de impresión o huecos vacíos. La proteína T7 e IgG humana (Sigma-Aldrich) se imprimieron como controles positivos para verificar la calidad de la matriz.

Treinta sueros (15 de pacientes con CCR y 15 de individuos normales) se incubaron con las micromatrices de fagos como se describió previamente (Chatterjee, M *et al.* 2006. Cancer Res. 66:1181-1190). Brevemente, los portaobjetos fueron equilibrados con PBS a temperatura ambiente durante 5 min y bloqueados con leche desnatada al 3% en PBS (MPBS 3%) durante 1h a temperatura ambiente con agitación. Posteriormente, las matrices se incubaron con 6,6 µl de suero humano (dilución 1:300), 120 µg de lisado de *E. coli* y 0,3 µg de anticuerpo monoclonal anti-etiqueta T7 (Novagen) diluidos en 2 ml de MPBS 3% durante 90 min a temperatura ambiente. Los portaobjetos se lavaron 3 veces con PBST durante 10 min para eliminar la unión inespecífica y se incubaron con un anticuerpo anti-IgG humana conjugado a AlexaFluor 647 (Invitrogen) diluido 1:2.000 y un anticuerpo anti-IgG de ratón conjugado a AlexaFluor 555 (Invitrogen) diluido 1:40.000 en MPBS 3% para detectar los anticuerpos humanos unidos a los fagos T7 y los fagos T7, respectivamente. Posteriormente, las micromatrices se lavaron 3 veces con PBST, una vez con PBS y se secaron mediante centrifugación a 1.200 rpm durante 3 min. Finalmente, los portaobjetos se escanearon en un ScanArrayTM5000 (Packard BioChip Technologies). Para la cuantificación de la intensidad de los puntos se utilizó el programa de análisis de imagen Genepix Pro 7 (Axon Laboratories).

Análisis estadístico

Los datos de micromatrices se normalizaron y procesaron usando las aplicaciones Asterias (<http://asterias.bioinfo.cnio.es/>), interfaz para el uso de paquetes informáticos, Limma y marrayNorm de Bioconductor. Tras aplicar una corrección de ruido de fondo y la normalización global de loess (<http://dnmad.bioinfo.cnio.es/>), los datos se procesaron para filtrar los valores que faltaban o los puntos con una varianza demasiado alta, para combinar duplicados y luego obtener un valor único para cada fago transformado en logaritmo en base 2 (<http://prep.bioinfo.cnio.es/>). Los grupos de pacientes con CCR e individuos sanos se compararon realizando una prueba de la t con el programa Pomelo II (<http://pomelo2.bioinfo.cnio.es/>), donde los valores p se obtuvieron mediante 200000 permutaciones. El programa Pomelo II generó un gráfico mostrando los fagos con un valor de tasa de falsos resultados positivos (FDR) por debajo de 0,15 y un valor p no ajustado por debajo de 0,05.

Los análisis de remuestreo se ajustaron mediante un modelo de regresión logística donde se modeló la probabilidad de ser tumoral contra la de ser normal como una función de las variables (fagos y proteínas). En el modelo también se incluyeron la edad y el sexo de los pacientes para corregir los posibles efectos de estas variables. Para evaluar la capacidad predictiva de los modelos se calculó el área debajo de la curva ROC (AUC). El AUC calculado directamente con el modelo original y el conjunto completo de datos está sesgado hacia los valores altos. Así, para obtener un AUC corregido, no sesgado hacia valores altos, que dé una estimación del AUC que se podría obtener con una validación futura independiente se utilizó el remuestreo con 1.000 muestras replicadas (Efron B. J. Am. Stat Assoc. 1983;78:316-331). Los modelos se ajustaron utilizando la genoteca de diseño de Harrell (Harrell F. Springer. 2001) con el sistema de computación estadístico R (Team RDC, 2009).

## II. RESULTADOS

El ARN de tejido de CCR de 6 pacientes (tres con estadio de Duke A y tres con estadio de Duke C) se utilizó para construir genotecas de fagos T7 que contenían fragmentos de ADNc en 2 vectores (T7Select 415-1 o T7select 10-3b). Tras realizar la selección de fagos específicos de CCR se obtuvieron 8 genotecas distintas enriquecidas en fagos específicos de tumor, dependiendo del vector y de la mezcla de sueros (B, C, H y P) usados durante la selección. Un total de 1.536 fagos individuales se amplificaron (192 fagos individuales de cada selección) y se imprimieron en duplicado en matrices de nitrocelulosa. Como control de la cantidad de fago impreso en la matriz se utilizó un anticuerpo anti-T7 que permitió observar la presencia de una señal homogénea en la matriz. Con el fin de determinar la reproducibilidad intra e inter matriz, se representó la intensidad de los 2 puntos correspondientes al mismo fago dentro de una misma matriz y entre dos matrices diferentes. Se determinó que la reproducibilidad intra e inter matriz era buena con unos valores de R2 de 0,9703 y 0,9091, respectivamente.

Para evaluar la respuesta inmune en pacientes con CCR, se incubaron las matrices con 30 sueros (15 de pacientes con diferentes estadios de CCR y 15 de controles sanos). Tras la cuantificación de las imágenes y la normalización de los datos, se comparó la señal de los sueros tumorales con los sueros sanos usando una prueba de la t con 200.000 permutaciones. 128 fagos mostraron reactividades diferentes entre los 2 grupos, con un FDR < 0,22. Entre ellos, 78 fagos mostraron una reactividad aumentada en CCR mientras que 50 tenían una reactividad disminuida en sueros con CCR. La representación del análisis supervisado de los 45 fagos con el FDR más bajo (<0.15) mostró una clara separación entre los pacientes con CCR y los individuos sanos.

**Ejemplo 2**Identificación de las secuencias presentadas en fagos

5

## I. MATERIALES Y MÉTODOS

Secuenciación y análisis de las secuencias internas mediante BLASTp

10 El ADN insertado en el genoma del fago se amplificó mediante PCR usando el cebador directo T7\_up2: 5'-TGCTAAGGACAAACGTTATCGG-3' (SEQ ID NO:13) y el cebador inverso T7\_down2: 5'-TTGATACCGGACGTTAC-3' (SEQ ID NO:14). Los productos de PCR se precipitaron con etanol y se secuenciaron directamente con el cebador directo T7\_up2.

15 Con el fin de encontrar homología de secuencia para cada péptido desplegado en la superficie de los fagos seleccionados se realizó una búsqueda en la base de datos NCBI con el software BLASTp.

Proteínas, anticuerpos y líneas celulares

20 Las proteínas recombinantes humanas MST1/STK4 y SULF1 se expresaron en *E. coli*. El ADNc de MST1 se subclonó en el vector pET28a (Novagen). El ADNc de SULF1 se clonó en el vector pDONR221 y posteriormente en el vector de expresión pDEST17. Las 2 proteínas de fusión 6xHis-MST1 y 6xHis-SULF1 se expresaron en la cepa BL21 (DE3) de *E. coli* y se purificaron a homogeneidad mediante cromatografía de afinidad en una columna HisTrap (GE Healthcare). Por último, las proteínas se dializaron frente PBS y se concentraron. La proteína EBNA1 utilizada como control positivo en ensayos de ELISA se compró a la empresa Tebu-Bio.

30 Los anticuerpos contra MST1/STK4, SULF1 y tubulina usados en la inmunodetección en membrana se compraron a las compañías Atlas antibodies, Santa Cruz Biotechnology y Sigma, respectivamente. El anticuerpo anti-IgG de conejo TrueBlot conjugado a la peroxidasa se compró a la compañía eBioscience y los anticuerpos anti-IgG de ratón y anti-IgG humana conjugado a peroxidasa a DakoCytomation.

35 Las líneas celulares de cáncer colorrectal RKO, Caco2, Hct15, Hct116, Colo320, SW480, SW48, KM12C, KM12SM, HT29, Colo205 y las líneas celulares de referencia (HEK293 y MOLT4) se cultivaron según protocolos establecidos. Los linfocitos de sangre periférica (LSP) y los monocitos se aislaron de un donante sano.

Análisis por inmunotransferencia

40 La preparación de los extractos de tejido emparejados y de líneas celulares se realizó según el siguiente protocolo. Las células y los tejidos se lavaron 2 veces con PBS antes de su lisado con 500 µl de SDS al 0,5% con inhibidores de proteasas (Roche Applied Science). La concentración de los extractos se determinó mediante el kit 2D-Quant (GE Healthcare) después de clarificar la muestra mediante centrifugación a 12.000 g durante 15 min a 4°C.

45 25 µg de extracto proteico se separaron en un gel SDS-PAGE al 10% y se transfirieron a membranas de nitrocelulosa (Hybond-C Extra) según protocolos establecidos (Babel *et al.* Mol. Cell Proteomics 2009; 8:2382-95). La membrana se bloqueó con MPBS 3% y se incubó toda la noche a 4°C con los anticuerpos contra MST1 (dilución 1:1.000), SULF1 (dilución 1:3.000) o tubulina (dilución 1:5.000). La inmunodetección se realizó usando un anticuerpo anti-IgG de ratón conjugado a HRP (dilución 1:5.000) o un anticuerpo anti-IgG de conejo conjugado a HRP (dilución 1:5.000). Finalmente, la unión del anticuerpo se detectó utilizando ECL (GE Healthcare) o SuperSignal Femto (Pierce).

ELISA

55 Las placas de ELISA para capturar fagos T7 (Novagen) se bloquearon 2 h a 37°C con MPBS 3% y se incubaron toda la noche con 100 µl del lisado bacteriano de los fagos diluido en MPBS 3%. Tras lavar 3 veces con PBST, las placas se bloquearon con MPBS 3% 1h a 37°C y se incubaron con 100 µl de suero humano (dilución 1:50 en MPBS 3%) 1 h a 37°C. Después de 3 lavados adicionales, se añadió el anticuerpo anti-IgG humana marcado con peroxidasa (1:3.000 en MPBS 3%) durante 2 h a temperatura ambiente. La señal se detectó con el sustrato 3,3',5,5'-tetrametilbenzidina (Sigma) durante 10 min, parando la reacción con HCl 1 M y midiendo la señal a 450 nm.

60 El ensayo de competición entre los péptidos presentados en la superficie de los fagos y las proteínas recombinantes se realizó usando las placas de captura de fago T7 (Novagen) siguiendo el protocolo anterior, exceptuando que los sueros humanos se preincubaron toda la noche a 4°C con diluciones seriadas de las proteínas MST1, SULF1 o GST. Los sueros así preincubados se ensayaron contra EBNA1 en placas de ELISA (Maxisorp, Nunc) como control positivo para verificar que la competición para las IgG entre el fago y su respectiva proteína recombinante era específica.

65

Los ELISA con las proteínas MST1, SULF1 y EBNA1 se realizaron como se ha descrito previamente (Babel *et al.* Mol. Cell Proteomics 2009; 8:2382-95). La concentración de CEA en los sueros de los pacientes con CCR y los sueros control se determinó mediante un ensayo inmunológico específico siguiendo las recomendaciones del fabricante (MP Biomedicals).

5

#### Análisis estadístico

Los datos de ELISA para cada marcador individual (fago o proteína de longitud completa) se evaluaron calculando una curva ROC (curva de eficacia diagnóstica). El área debajo de la curva correspondiente (AUC) se calculó usando el programa JMP7 (SAS). La media y la desviación estándar de los resultados de inmunohistoquímica se calcularon utilizando el programa Microsoft Office Excel 2007. La prueba de la t de Student unilateral se llevó a cabo utilizando los resultados de inmunohistoquímica, asumiendo que las varianzas no iguales para determinar las medias de los grupos normales y tumorales eran significativamente diferentes entre ellos.

## 15 II. RESULTADOS

Se obtuvieron 43 secuencias de aminoácidos únicas fusionadas a la proteína 10B de la cápside del fago T7 entre los 78 fagos que mostraban una reactividad aumentada en sueros de pacientes con CCR.

20 De los 43 fagos únicos, se seleccionaron para verificar los resultados aquellos que contenían entre 8 y 20 residuos con una elevada homología con secuencias de proteínas conocidas, los que aparecían un mayor número de veces con la misma secuencia de aminoácidos y los que tenían un valor p bajo. Se identificaron secuencias homólogas a las proteínas MST1/STK4, SULF1, NHSL1, SREBF2, GRN y GTF2i. Todas ellas tenían una señal en las micromatrices significativamente mayor con el suero de pacientes con CCR que con los sueros control (Figura 1).  
25 Los fagos fueron identificados por el nombre de la proteína con la que se identificaron secuencias homólogas. La proteína MST1/STK4 se identificó previamente como un antígeno asociado a tumor en CCR usando micromatrices de proteínas comerciales (Babel *et al.* Mol. Cell Proteomics 2009;8:2382-95) y el gen SULF1 se describió como sobre expresado en un análisis transcriptómico de CCR (Madoz-Gurpide *et al.* Mol Cell Proteomics, 2006;5:1471-83).

30

Con el fin de confirmar que los péptidos presentados en los fagos cuya secuencia homóloga pertenecía a las proteínas SULF1 y MST1, se realizó un ensayo de competición de IgG entre los fagos y las proteínas recombinantes humanas SULF1 y MST1. La unión de las inmunoglobulinas presentes en los sueros humanos a los 2 fagos se inhibió de una manera dependiente de la dosis con las proteínas recombinantes MST1 y SULF1 (Figura 2A). GST no afectó a la unión de las IgG a los fagos (control negativo). Como control de inhibición específica, se observó que la unión de los anticuerpos de los pacientes contra la proteína EBNA no se vio afectada por la incubación con las proteínas MST1 o SULF1.

35

Por otra parte, se determinó que las secuencias de los péptidos presentados en los fagos se localizaban en la región C-terminal de MST1 y en el N-terminal de SULF1 (Figura 2B).

40

Todos estos resultados confirman que los péptidos presentados corresponden a epítopos inmunodominantes de las proteínas MST1 y SULF1.

### 45 **Ejemplo 3**

#### Las proteínas identificadas están sobreexpresadas en cáncer colorrectal

Los antígenos asociados a tumores reconocidos por autoanticuerpos se encuentran generalmente sobreexpresados en líneas celulares y en tejidos tumorales. Se realizó un metaanálisis de los niveles de expresión del ARNm de las proteínas homólogas correspondientes a los 6 fagos seleccionados [MST1/STK4, SULF1, NHSL1, SREBF2, GRN y GTF2i] con la base de datos de micromatrices OncoPrint (Rhodes *et al.* Neoplasia 2004;6:1-6) (Figura 3A). Se encontró que SULF1 era el gen más sobreexpresado en cáncer de colon, seguido por GTF2i, MST1, GRN, NHSL1 y SREBF2. Por otra parte, se realizó una inmunodetección en membrana con los anticuerpos contra MST1 y SULF1 usando 11 líneas celulares de CCR y tejidos tumorales de pacientes con CCR representando los diferentes estadios de evolución de la enfermedad (Figura 3B). Se encontró que las proteínas MST1 y SULF1 estaban expresadas en la mayor parte de las líneas celulares de cáncer de colon. La mayor expresión de SULF1 se observó en líneas celulares metastásicas (SW48, HT29 y COLO205) y en tejido tumoral de CCR de estadios tardíos.

55

Los patrones de expresión celular de las proteínas seleccionadas se caracterizaron mediante inmunohistoquímica (TMA) usando tumores de CCR independientes dispuestos en micromatrices o mediante metaanálisis de datos de micromatrices de tejidos obtenidos del atlas humano de proteínas (Berglund *et al.* Mol Cell Proteomics. 2008;7:2019-27) (Figura 3C). En todos los casos, se detectó una expresión más abundante de la proteína estudiada en tejidos tumorales.

65

Por tanto, existe una buena correlación entre la presencia de autoanticuerpos, la abundancia de las proteínas y la expresión génica.

#### Ejemplo 4

##### Validación del predictor formado por péptidos presentados en fagos y sus proteínas homólogas

Se utilizó un conjunto independiente de 96 muestras séricas (50 de cáncer colorrectal con 19 muestras de estadios tempranos (A+B) y 46 controles sanos) para la validación de los resultados. Se ensayaron los fagos MST1, SULF1, NHSL1, SREBF2, GRN y GTF2i para su capacidad para discriminar entre sueros tumorales y sueros controles por ELISA. Se construyeron las curvas ROC para cada uno de los marcadores con los resultados de ELISA. Mientras que la sensibilidad era relativamente baja para los fagos individuales, oscilando entre el 46% y el 60%, la especificidad era mayor, entre el 50% y el 73,9%. Para investigar si diferentes combinaciones de fagos tenían una mayor precisión en discriminar individuos sanos de pacientes con cáncer, se ajustaron los datos a una curva logística, realizando regresiones lineales y produciendo diferentes modelos con diferentes combinaciones de fagos. Así, el resultado de la combinación de los 6 fagos como predictor de CCR da un AUC de 0,82 con una sensibilidad y una especificidad del 70% y el 73,9%, respectivamente (Tabla 4).

Tabla 4. Datos de las curvas ROC obtenidas de los valores de ELISA de la validación tanto de los fagos individuales como de los fagos combinados.

Péptido presentado en fago	Especificidad (%)	Sensibilidad (%)	AUC
SULF1	73,9	50,0	0,63
NHSL1	50,0	56,0	0,59
MST1	71,7	46,0	0,58
GTF2i	52,2	60,0	0,57
SREBF2	69,6	54,0	0,61
GRN	50,0	58,0	0,53
Combinación de los 6 fagos	73,9	70,0	0,82

El siguiente paso consistió en ver si el reemplazamiento de los fagos por sus proteínas recombinantes MST1 y SULF1 mejoraría el poder discriminante del modelo (Figura 4). Los resultados confirmaron una mejora significativa de la predicción utilizando las proteínas recombinantes, con AUC de 0,71 y 0,74 para las proteínas SULF1 y MST1 frente a 0,63 y 0,58 de los respectivos fagos (Tabla 2). Combinando las dos proteínas (SULF1 y MST1) y los cuatro fagos (NHSL1, SREBF2, GRN y GTF2i), el AUC aumentó hasta 0,86 con una sensibilidad del 82,6% y una especificidad del 70% (Figura 6A). Los valores de CEA eran menores (AUC 0,81) y combinados con el resto de las predicciones apenas mejoraban el modelo (AUC 0,89). Además, en la etapa de validación, se realizaron diferentes estimaciones de AUC para comparar no sólo CRC frente a sano, sino también CRC frente a los sueros de referencia y sano frente a otros tumores (Figura 6). El resultado más relevante fue la capacidad del modelo de discriminar no sólo CRC de sueros sanos (AUC 0,86) (Figura 6A), sino también CRC de todos los sueros de referencia, que incluía otras patologías relacionadas con colon (AUC 0,85) (Figura 6B). Notablemente, el panel parecía no discriminar adecuadamente controles sanos de otros tumores (AUC 0,63) (Figura 6C). Además, el panel parecía discriminar significativamente controles sanos de pacientes asintomáticos con historia familiar de CRC (AUC 0,78).

##### Análisis de remuestreo

También se realizó un remuestreo para obtener las AUC corregidas. El modelo inicial incluía términos lineales para todos los fagos y proteínas, junto con otras dos variables: sexo y edad de los pacientes. Con este modelo, el valor del AUC corregido fue 0,83.

Este modelo probablemente era más complejo de lo necesario. Por eso, se realizó una selección de variables con el criterio de información de Akaike como punto final. El modelo final únicamente retuvo 3 proteínas (GRN, MST1 y SULF1), además de la edad de los pacientes (Tabla 5). Sin embargo, para evitar una sobreestimación de la capacidad predictiva del modelo, se obtuvieron valores corregidos estimados de AUC, mediante remuestreo del proceso completo de la selección de variables (es decir, por cada muestra de remuestreo, se realizó el modelo completo con 8 variables y se utilizó el criterio de información de Akaike). El AUC corregido fue 0,84. El remuestreo también proporcionó información sobre la estabilidad del proceso de selección; la mayoría de los modelos de remuestreo contenían cuatro, cinco, seis o siete variables. Algunas de las variables aparecían en la mayoría de los modelos; el fago GRN en 976, la proteína SULF1 en 954, la edad en 952 y la proteína MST1 en 833.

Además, se utilizó este modelo para predecir la probabilidad de ser CRC del grupo de 57 sueros que comprendían diversas patologías. Se generó una gráfica de puntos (Figura 6D), en donde se muestra la probabilidad individual para cada sujeto. Se observó una gran variabilidad en probabilidad dentro de cada grupo, pero la mediana se encontraba bastante por debajo de 0,5, indicando una probabilidad baja de padecer CRC.

Tabla 5. Modelo final con remuestreo después del modelo de selección.

	Estimado	Error Estd.	Valor z	Pr(> z )
(Intercepto*)	-5,92318	2,15625	-2,747	0,00601**
GRN-fago	-6,28345	2,91418	-2,156	0,03107*
NHSL1- fago	6,40976	2,8367	2,26	0,02385*
GTF2B- fago	-9,00788	2,63827	-3,414	0,00064***
SREBF2- fago	10,31184	3,30888	3,116	0,00183**
MST1- fago	3,53576	1,32002	2,679	0,00739**
SULF1- fago	7,26445	2,97056	2,445	0,01447*
Sexo	-0,98593	0,65076	-1,515	0,12976
Edad	0,05674	0,0232	2,446	0,01444*

\*El intercepto es el valor de  $\log(p/(1-p))$ , donde  $p$  es la probabilidad de ser tumoral, cuando el valor de las otras variables en el modelo es 0. Estimado, coeficiente estimado (pendiente); Error Estd., el error estándar de la variable especificada; valor z o estadística Wald, que es exactamente igual al coeficiente estimado dividido por su error estándar;  $Pr(>|z|)$ , valor p de la prueba de Wald para ese coeficiente específico comparando el valor z frente al estándar normal; \*, grado de significancia.

Posteriormente, se ensayó el predictor de acuerdo con el estadio del paciente, usando el modelo con los 6 marcadores (4 fagos + 2 proteínas) más la edad de los pacientes. El AUC corregido usando el remuestreo fue 0,786 para los estadios A+B; 0,857 para el estadio C; y 0,849 para el estadio D. Si se aplica el mismo test con los valores de CEA, los valores corregidos de AUC eran 0,742 para los estadios A+B, 0,770 para el estadio C y 0,973 para el estadio D. Estos resultados indican una clara superioridad del modelo predictor para el diagnóstico de CCR en estadios A, B y C, siendo CEA superior para el estadio D, tal como se esperaba.

### Ejemplo comparativo 5

#### Los autoanticuerpos contra MST1 y NHSL1 están asociados con el pronóstico clínico de los pacientes con CCR

Después de haber demostrado que el panel predictivo era capaz de identificar los pacientes con tumores tanto en estadios tempranos como en avanzados, se procedió a investigar el potencial pronóstico de los autoanticuerpos analizando su asociación con la supervivencia absoluta de los pacientes.

La supervivencia se estimó mediante el método estándar de Kaplan-Meier. El valor pronóstico de los anticuerpos candidatos se evaluó mediante los modelos de riesgo proporcional de Cox estratificado por estadio. Los valores p se obtuvieron de las pruebas de cociente de probabilidad. Únicamente los anticuerpos encontrados como significativos para el diagnóstico se evaluaron para el pronóstico en un proceso multivariable paso a paso.

Los pacientes con bajos niveles de anticuerpos contra MST1 mostraban una supervivencia acumulativa menor que aquellos con niveles elevados de autoanticuerpos ( $p=0,08$ ). En contraste, pacientes con títulos más altos de anticuerpos contra NHSL1 estaban asociados con una supervivencia menor ( $p=0,06$ ). La combinación de ambos efectos mejora el valor estadístico de la predicción (0,032) (Figura 5) y confirma la predicción de supervivencia. Colectivamente, estos datos indican una correlación entre la presencia de anticuerpos contra MST1 y NHSL1 con el pronóstico clínico de pacientes de cáncer colorrectal.

Este valor pronóstico era más importante para estadios tempranos de la enfermedad con un índice de riesgo de 5,1; aunque no era significativo ( $p=0,12$ ) porque sólo se observaron 8 sucesos en los 47 pacientes de estadios I-II. Además, se observó un efecto similar en el análisis de la supervivencia sin enfermedad cuando los pacientes se restringían a los márgenes de resección quirúrgica libre de tumor.

#### Lista de secuencias

<110> CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS

<120> MÉTODO DE DIAGNÓSTICO/PRONÓSTICO DEL CÁNCER COLORRECTAL

<130> P6917PC00

<150> ES202030708

<151> 13-05-2010

<160> 14



<170> PatentIn versión 3.5

<210> 1  
<211> 7  
<212> PRT  
<213> Homo sapiens

<400> 1

Ser Arg Ala Thr Met Pro Gly  
1 5

<210> 2  
<211> 14  
<212> PRT  
<213> Homo sapiens

<400> 2

Ser Leu Gly Gln Arg Ser Gln Lys Leu Ala Ala Ala Leu Glu  
1 5 10

<210> 3  
<211> 8  
<212> PRT  
<213> Homo sapiens

<400> 3

Tyr Asn Ser Lys Arg Gln Pro Asn  
1 5

<210> 4  
<211> 8  
<212> PRT  
<213> Homo sapiens

<400> 4

Glu Asn Lys Thr Arg Lys Ile Met  
1 5

<210> 5  
<211> 10  
<212> PRT  
<213> Homo sapiens

<400> 5

ES 2 819 055 T3

Arg Asp Arg Ala Thr Ala Ile Gln Pro Gly  
1 5 10

<210> 6  
<211> 5  
<212> PRT  
<213> Homo sapiens

<400> 6

Ser Ser Lys Ser Pro  
1 5

<210> 7  
<211> 1606  
<212> PRT  
<213> Homo sapiens

<400> 7

Met Val Val Phe Ile Asn Ala Lys Ile Lys Ser Leu Ile Lys Leu Phe  
1 5 10 15

Lys Lys Lys Thr Val Ser Asn Leu Asp Glu Glu Ser Arg Trp Thr Val  
20 25 30

His Tyr Thr Ala Pro Trp His Gln Gln Glu Asn Val Phe Leu Pro Thr  
35 40 45

Thr Arg Pro Pro Cys Val Glu Asp Leu His Arg Gln Ala Lys Leu Asn  
50 55 60

Leu Lys Ser Val Leu Arg Glu Cys Asp Lys Leu Arg His Asp Gly Tyr  
65 70 75 80

Arg Ser Ser Gln Tyr Tyr Ser Gln Gly Pro Thr Phe Ala Ala Asn Ala  
85 90 95

Ser Pro Phe Cys Asp Asp Tyr Gln Asp Glu Asp Glu Glu Thr Asp Gln  
100 105 110

Lys Cys Ser Leu Ser Ser Ser Glu Glu Glu Arg Phe Ile Ser Ile Arg  
115 120 125

Arg Pro Lys Thr Pro Ala Ser Ser Asp Phe Ser Asp Leu Asn Thr Gln  
130 135 140

Thr Asn Trp Thr Lys Ser Leu Pro Leu Pro Thr Pro Glu Glu Lys Met  
145 150 155 160

Arg Gln Gln Ala Gln Thr Val Gln Ala Asp Val Val Pro Ile Asn Ile  
165 170 175

ES 2 819 055 T3

Thr Gly Glu Asn Phe Asp Arg Gln Ala Ser Leu Arg Arg Ser Leu Ile  
 180 185 190  
 Tyr Thr Asp Thr Leu Val Arg Arg Pro Lys Lys Val Lys Arg Arg Lys  
 195 200 205  
 Thr Ile Thr Gly Val Pro Asp Asn Ile Gln Lys Glu Leu Ala Ser Gly  
 210 215 220  
 Thr Gly Gln Asp Asp Ala Asp Gly His Ser Val Tyr Thr Pro Asp His  
 225 230 235 240  
 Tyr Ser Thr Leu Gly Arg Phe Asn Ser Cys Arg Ser Ala Gly Gln Arg  
 245 250 255  
 Ser Glu Thr Arg Asp Ser Ser Cys Gln Thr Glu Asp Val Lys Val Val  
 260 265 270  
 Pro Pro Ser Met Arg Arg Ile Arg Ala Gln Lys Gly Gln Gly Ile Ala  
 275 280 285  
 Ala Gln Met Gly His Phe Ser Gly Ser Ser Gly Asn Met Ser Val Leu  
 290 295 300  
 Ser Asp Ser Ala Gly Ile Val Phe Pro Ser Arg Leu Asp Ser Asp Ala  
 305 310 315 320  
 Gly Phe His Ser Leu Pro Arg Ser Gly Ala Arg Ala Asn Ile Gln Ser  
 325 330 335  
 Leu Glu Pro Arg Leu Gly Ala Leu Gly Pro Ala Gly Asp Met Asn Gly  
 340 345 350  
 Thr Phe Leu Tyr Gln Arg Gly His Pro Gln Ala Asp Glu Asn Leu Gly  
 355 360 365  
 His Leu Gly Gly Ala Ser Gly Thr Gly Thr Leu Leu Arg Pro Lys Ser  
 370 375 380  
 Gln Glu Leu Arg His Phe Glu Ser Glu Asn Ile Met Ser Pro Ala Cys  
 385 390 395 400  
 Val Val Ser Pro His Ala Thr Tyr Ser Thr Ser Ile Ile Pro Asn Ala  
 405 410 415  
 Thr Leu Ser Ser Ser Ser Glu Val Ile Ala Ile Pro Thr Ala Gln Ser  
 420 425 430

ES 2 819 055 T3

Ala Gly Gln Arg Glu Ser Lys Ser Ser Gly Ser Ser His Ala Arg Ile  
 435 440 445

Lys Ser Arg Asp His Leu Ile Ser Arg His Ala Val Lys Gly Asp Pro  
 450 455 460

Gln Ser Pro Gly Arg His Trp Asn Glu Gly His Ala Thr Ile Leu Ser  
 465 470 475 480

Gln Asp Leu Asp Pro His Ser Pro Gly Glu Pro Ala Leu Leu Ser Leu  
 485 490 495

Cys Asp Ser Ala Val Pro Leu Asn Ala Pro Ala Asn Arg Glu Asn Gly  
 500 505 510

Ser Gln Ala Met Pro Tyr Asn Cys Arg Asn Asn Leu Ala Phe Pro Ala  
 515 520 525

His Pro Gln Asp Val Asp Gly Lys Ser Glu Ser Ser Tyr Ser Gly Gly  
 530 535 540

Gly Gly His Ser Ser Ser Glu Pro Trp Glu Tyr Lys Ser Ser Gly Asn  
 545 550 555 560

Gly Arg Ala Ser Pro Leu Lys Pro His Leu Ala Thr Pro Gly Tyr Ser  
 565 570 575

Thr Pro Thr Ser Asn Met Ser Ser Cys Ser Leu Asp Gln Thr Ser Asn  
 580 585 590

Lys Glu Asp Ala Gly Ser Leu Tyr Ser Glu Asp His Asp Gly Tyr Cys  
 595 600 605

Ala Ser Val His Thr Asp Ser Gly His Gly Ser Gly Asn Leu Cys Asn  
 610 615 620

Ser Ser Asp Gly Phe Gly Asn Pro Arg His Ser Val Ile Asn Val Phe  
 625 630 635 640

Val Gly Arg Ala Gln Lys Asn Gln Gly Asp Arg Ser Asn Tyr Gln Asp  
 645 650 655

Lys Ser Leu Ser Arg Asn Ile Ser Leu Lys Lys Ala Lys Lys Pro Pro  
 660 665 670

Leu Pro Pro Ser Arg Thr Asp Ser Leu Arg Arg Ile Pro Lys Lys Ser  
 675 680 685

ES 2 819 055 T3

Ser Gln Cys Asn Gly Gln Val Leu Asn Glu Ser Leu Ile Ala Thr Leu  
690 695 700

Gln His Ser Leu Gln Leu Ser Leu Pro Gly Lys Ser Gly Ser Ser Pro  
705 710 715 720

Ser Gln Ser Pro Cys Ser Asp Leu Glu Glu Pro Trp Leu Pro Arg Ser  
725 730 735

Arg Ser Gln Ser Thr Val Ser Ala Gly Ser Ser Met Thr Ser Ala Thr  
740 745 750

Thr Pro Asn Val Tyr Ser Leu Cys Gly Ala Thr Pro Ser Gln Ser Asp  
755 760 765

Thr Ser Ser Val Lys Ser Glu Tyr Thr Asp Pro Trp Gly Tyr Tyr Ile  
770 775 780

Asp Tyr Thr Gly Met Gln Glu Asp Pro Gly Asn Pro Ala Gly Gly Cys  
785 790 795 800

Ser Thr Ser Ser Gly Val Pro Thr Gly Asn Gly Pro Val Arg His Val  
805 810 815

Gln Glu Gly Ser Arg Ala Thr Met Pro Gln Val Pro Gly Gly Ser Val  
820 825 830

Lys Pro Lys Ile Met Ser Pro Glu Lys Ser His Arg Val Ile Ser Pro  
835 840 845

Ser Ser Gly Tyr Ser Ser Gln Ser Asn Thr Pro Thr Ala Leu Thr Pro  
850 855 860

Val Pro Val Phe Leu Lys Ser Val Ser Pro Ala Asn Gly Lys Gly Lys  
865 870 875 880

Pro Lys Pro Lys Val Pro Glu Arg Lys Ser Ser Leu Ile Ser Ser Val  
885 890 895

Ser Ile Ser Ser Ser Ser Thr Ser Leu Ser Ser Ser Thr Ser Thr Glu  
900 905 910

Gly Ser Gly Thr Met Lys Lys Leu Asp Pro Ala Val Gly Ser Pro Pro  
915 920 925

Ala Pro Pro Pro Pro Pro Val Pro Ser Pro Pro Phe Pro Cys Pro Ala  
930 935 940

Asp Arg Ser Pro Phe Leu Pro Pro Pro Pro Pro Val Thr Asp Cys Ser





ES 2 819 055 T3

Val Val Pro Pro Pro Gln Lys Asp Phe Ala Val Glu Pro Ala Glu  
1205 1210 1215

Asn Val Ser Glu Ala Leu Arg Ala Val Pro Ser Pro Thr Thr Gly  
1220 1225 1230

Glu Glu Gly Ser Val His Ser Arg Glu Ala Lys Glu Ser Ser Ala  
1235 1240 1245

Ala Gln Ala Gly Ser His Ala Thr His Pro Gly Thr Ser Val Leu  
1250 1255 1260

Glu Gly Gly Ala Ala Gly Ser Met Ser Pro Ser Arg Val Glu Ala  
1265 1270 1275

Asn Val Pro Met Val Gln Pro Asp Val Ser Pro Ala Pro Lys Gln  
1280 1285 1290

Glu Glu Pro Ala Glu Asn Ser Ala Asp Thr Gly Gly Asp Gly Glu  
1295 1300 1305

Ser Cys Leu Ser Gln Gln Asp Gly Ala Ala Gly Val Pro Glu Thr  
1310 1315 1320

Asn Ala Ala Gly Ser Ser Ser Glu Ala Cys Asp Phe Leu Lys Glu  
1325 1330 1335

Asp Gly Asn Asp Glu Val Met Thr Pro Ser Arg Pro Arg Thr Thr  
1340 1345 1350

Glu Asp Leu Phe Ala Ala Ile His Arg Ser Lys Arg Lys Val Leu  
1355 1360 1365

Gly Arg Arg Asp Ser Asp Asp Asp His Ser Arg Asn His Ser Pro  
1370 1375 1380

Ser Pro Pro Val Thr Pro Thr Gly Ala Ala Pro Ser Leu Ala Ser  
1385 1390 1395

Pro Lys Gln Val Gly Ser Ile Gln Arg Ser Ile Arg Lys Ser Ser  
1400 1405 1410

Thr Ser Ser Asp Asn Phe Lys Ala Leu Leu Leu Lys Lys Gly Ser  
1415 1420 1425

Arg Ser Asp Thr Ser Ala Arg Met Ser Ala Ala Glu Met Leu Lys  
1430 1435 1440

ES 2 819 055 T3

Asn Thr Asp Pro Arg Phe Gln Arg Ser Arg Ser Glu Pro Ser Pro  
1445 1450 1455

Asp Ala Pro Glu Ser Pro Ser Ser Cys Ser Pro Ser Lys Asn Arg  
1460 1465 1470

Arg Ala Gln Glu Glu Trp Ala Lys Asn Glu Gly Leu Met Pro Arg  
1475 1480 1485

Ser Leu Ser Phe Ser Gly Pro Arg Tyr Gly Arg Ser Arg Thr Pro  
1490 1495 1500

Pro Ser Ala Ala Ser Ser Arg Tyr Ser Met Arg Asn Arg Ile Gln  
1505 1510 1515

Ser Ser Pro Met Thr Val Ile Ser Glu Gly Glu Gly Glu Ala Val  
1520 1525 1530

Glu Pro Val Asp Ser Ile Ala Arg Gly Ala Leu Gly Ala Ala Glu  
1535 1540 1545

Gly Cys Ser Leu Asp Gly Leu Ala Arg Glu Glu Met Asp Glu Gly  
1550 1555 1560

Gly Leu Leu Cys Gly Glu Gly Pro Ala Ala Ser Leu Gln Pro Gln  
1565 1570 1575

Ala Pro Gly Pro Val Asp Gly Thr Ala Ser Ala Glu Gly Arg Glu  
1580 1585 1590

Pro Ser Pro Gln Cys Gly Gly Ser Leu Ser Glu Glu Ser  
1595 1600 1605

<210> 8  
<211> 69  
<212> PRT  
<213> Homo sapiens

<400> 8

Ala Met Val Pro Cys Asp Asn Val Ser Ser Cys Pro Ser Ser Asp Thr  
1 5 10 15

Cys Cys Gln Leu Thr Ser Gly Glu Trp Gly Cys Cys Pro Ile Pro Glu  
20 25 30

Ala Val Cys Cys Ser Asp His Gln His Cys Cys Pro Gln Gly Tyr Thr  
35 40 45

Cys Val Ala Glu Gly Gln Cys Gln Lys Leu Ala Ala Ala Leu Glu His



ES 2 819 055 T3

Ser Leu Gly Ile Thr Ala Ile Glu Met Ala Glu Gly Lys Arg Pro Tyr  
 210 215 220

Ala Asp Ile His Pro Met Arg Ala Ile Phe Met Ile Pro Thr Asn Pro  
 225 230 235 240

Pro Pro Thr Phe Arg Lys Pro Glu Leu Trp Ser Asp Asn Phe Thr Asp  
 245 250 255

Phe Val Lys Gln Cys Leu Val Lys Ser Pro Glu Gln Arg Ala Thr Ala  
 260 265 270

Thr Gln Leu Leu Gln His Pro Phe Val Arg Ser Ala Lys Gly Val Ser  
 275 280 285

Ile Leu Arg Asp Leu Ile Asn Glu Ala Met Asp Val Lys Leu Lys Arg  
 290 295 300

Gln Glu Ser Gln Gln Arg Glu Met Asp Gln Asp Asp Glu Glu Asn Ser  
 305 310 315 320

Glu Glu Asp Glu Met Asp Ser Gly Thr Met Val Arg Ala Val Gly Asp  
 325 330 335

Glu Met Gly Thr Val Arg Val Ala Ser Thr Met Thr Asp Gly Ala Asn  
 340 345 350

Thr Met Ile Glu His Asp Asp Thr Leu Pro Ser Gln Leu Gly Thr Met  
 355 360 365

Val Ile Asn Ala Glu Asp Glu Glu Glu Glu Gly Thr Met Lys Arg Arg  
 370 375 380

Asp Glu Thr Met Gln Pro Ala Lys Pro Ser Phe Leu Glu Tyr Phe Glu  
 385 390 395 400

Gln Lys Glu Lys Glu Asn Gln Ile Asn Ser Phe Gly Lys Ser Val Pro  
 405 410 415

Gly Pro Leu Lys Asn Ser Ser Asp Trp Lys Ile Pro Gln Asp Gly Asp  
 420 425 430

Tyr Glu Phe Leu Lys Ser Trp Thr Val Glu Asp Leu Gln Lys Arg Leu  
 435 440 445

Leu Ala Leu Asp Pro Met Met Glu Gln Glu Ile Glu Glu Ile Arg Gln  
 450 455 460



ES 2 819 055 T3

Lys Tyr Gln Ser Lys Arg Gln Pro Ile Leu Asp Ala Ile Glu Ala Lys  
 465 470 475 480

Lys Arg Arg Gln Gln Asn Phe  
 485

<210> 10  
 <211> 882  
 <212> PRT  
 <213> Homo sapiens  
 <400> 10

Met Lys Tyr Ser Cys Cys Ala Leu Val Leu Ala Val Leu Gly Thr Glu  
 1 5 10 15

Leu Leu Gly Ser Leu Cys Ser Thr Val Arg Ser Pro Arg Phe Arg Gly  
 20 25 30

Arg Ile Gln Gln Glu Arg Lys Asn Ile Arg Pro Asn Ile Ile Leu Val  
 35 40 45

Leu Thr Asp Asp Gln Asp Val Glu Leu Gly Ser Leu Gln Val Met Asn  
 50 55 60

Lys Thr Arg Lys Ile Met Glu His Gly Gly Ala Thr Phe Ile Asn Ala  
 65 70 75 80

Phe Val Thr Thr Pro Met Cys Cys Pro Ser Arg Ser Ser Met Leu Thr  
 85 90 95

Gly Lys Tyr Val His Asn His Asn Val Tyr Thr Asn Asn Glu Asn Cys  
 100 105 110

Ser Ser Pro Ser Trp Gln Ala Met His Glu Pro Arg Thr Phe Ala Val  
 115 120 125

Tyr Leu Asn Asn Thr Gly Tyr Arg Thr Ala Phe Phe Gly Lys Tyr Leu  
 130 135 140

Asn Glu Tyr Asn Gly Ser Tyr Ile Pro Pro Gly Trp Arg Glu Trp Leu  
 145 150 155 160

Gly Leu Ile Lys Asn Ser Arg Phe Tyr Asn Tyr Thr Val Cys Arg Asn  
 165 170 175

Gly Ile Lys Glu Lys His Gly Phe Asp Tyr Ala Lys Asp Tyr Phe Thr  
 180 185 190

Asp Leu Ile Thr Asn Glu Ser Ile Asn Tyr Phe Lys Met Ser Lys Arg  
 195 200 205

ES 2 819 055 T3

Met Tyr Pro His Arg Pro Val Met Met Val Ile Ser His Ala Ala Pro  
 210 215 220

His Gly Pro Glu Asp Ser Ala Pro Gln Phe Ser Lys Leu Tyr Pro Asn  
 225 230 235 240

Ala Ser Gln His Ile Thr Pro Ser Tyr Asn Tyr Ala Pro Asn Met Asp  
 245 250 255

Lys His Trp Ile Met Gln Tyr Thr Gly Pro Met Leu Pro Ile His Met  
 260 265 270

Glu Phe Thr Asn Ile Leu Gln Arg Lys Arg Leu Gln Thr Leu Met Ser  
 275 280 285

Val Asp Asp Ser Val Glu Arg Leu Tyr Asn Met Leu Val Glu Thr Gly  
 290 295 300

Glu Leu Glu Asn Thr Tyr Ile Ile Tyr Thr Ala Asp His Gly Tyr His  
 305 310 315 320

Ile Gly Gln Phe Gly Leu Val Lys Gly Lys Ser Met Pro Tyr Asp Phe  
 325 330 335

Asp Ile Arg Val Pro Phe Phe Ile Arg Gly Pro Ser Val Glu Pro Gly  
 340 345 350

Ser Ile Val Pro Gln Ile Val Leu Asn Ile Asp Leu Ala Pro Thr Ile  
 355 360 365

Leu Asp Ile Ala Gly Leu Asp Thr Pro Pro Asp Val Asp Gly Lys Ser  
 370 375 380

Val Leu Lys Leu Leu Asp Pro Glu Lys Pro Gly Asn Arg Phe Arg Thr  
 385 390 395 400

Asn Lys Lys Ala Lys Ile Trp Arg Asp Thr Phe Leu Val Glu Arg Gly  
 405 410 415

Lys Phe Leu Arg Lys Lys Glu Glu Ser Ser Lys Asn Ile Gln Gln Ser  
 420 425 430

Asn His Leu Pro Lys Tyr Glu Arg Val Lys Glu Leu Cys Gln Gln Ala  
 435 440 445

Arg Tyr Gln Thr Ala Cys Glu Gln Pro Gly Gln Lys Trp Gln Cys Ile  
 450 455 460

ES 2 819 055 T3

Glu Asp Thr Ser Gly Lys Leu Arg Ile His Lys Cys Lys Gly Pro Ser  
 465 470 475 480  
 Asp Leu Leu Thr Val Arg Gln Ser Thr Arg Asn Leu Tyr Ala Arg Gly  
 485 490 495  
 Phe His Asp Lys Asp Lys Glu Cys Ser Cys Arg Glu Ser Gly Tyr Arg  
 500 505 510  
 Ala Ser Arg Ser Gln Arg Lys Ser Gln Arg Gln Phe Leu Arg Asn Gln  
 515 520 525  
 Gly Thr Pro Lys Tyr Lys Pro Arg Phe Val His Thr Arg Gln Thr Arg  
 530 535 540  
 Ser Leu Ser Val Glu Phe Glu Gly Glu Ile Tyr Asp Ile Asn Leu Glu  
 545 550 555 560  
 Glu Glu Glu Glu Leu Gln Val Leu Gln Pro Arg Asn Ile Ala Lys Arg  
 565 570 575  
 His Asp Glu Gly His Lys Gly Pro Arg Asp Leu Gln Ala Ser Ser Gly  
 580 585 590  
 Gly Asn Arg Gly Arg Met Leu Ala Asp Ser Ser Asn Ala Val Gly Pro  
 595 600 605  
 Pro Thr Thr Val Arg Val Thr His Lys Cys Phe Ile Leu Pro Asn Asp  
 610 615 620  
 Ser Ile His Cys Glu Arg Glu Leu Tyr Gln Ser Ala Arg Ala Trp Lys  
 625 630 635 640  
 Asp His Lys Ala Tyr Ile Asp Lys Glu Ile Glu Ala Leu Gln Asp Lys  
 645 650 655  
 Ile Lys Asn Leu Arg Glu Val Arg Gly His Leu Lys Arg Arg Lys Pro  
 660 665 670  
 Glu Glu Cys Ser Cys Ser Lys Gln Ser Tyr Tyr Asn Lys Glu Lys Gly  
 675 680 685  
 Val Lys Lys Gln Glu Lys Leu Lys Ser His Leu His Pro Phe Lys Glu  
 690 695 700  
 Ala Ala Gln Glu Val Asp Ser Lys Leu Gln Leu Phe Lys Glu Asn Asn  
 705 710 715 720

ES 2 819 055 T3

Arg Arg Arg Lys Lys Glu Arg Lys Glu Lys Arg Arg Gln Arg Lys Gly  
725 730 735

Glu Glu Cys Ser Leu Pro Gly Leu Thr Cys Phe Thr His Asp Asn Asn  
740 745 750

His Trp Gln Thr Ala Pro Phe Trp Asn Leu Gly Ser Phe Cys Ala Cys  
755 760 765

Thr Ser Ser Asn Asn Asn Thr Tyr Trp Cys Leu Arg Thr Val Asn Glu  
770 775 780

Thr His Asn Phe Leu Phe Cys Glu Phe Ala Thr Gly Phe Leu Glu Tyr  
785 790 795 800

Phe Asp Met Asn Thr Asp Pro Tyr Gln Leu Thr Asn Thr Val His Thr  
805 810 815

Val Glu Arg Gly Ile Leu Asn Gln Leu His Val Gln Leu Met Glu Leu  
820 825 830

Arg Ser Cys Gln Gly Tyr Lys Gln Cys Asn Pro Arg Pro Lys Asn Leu  
835 840 845

Asp Val Gly Asn Lys Asp Gly Gly Ser Tyr Asp Arg Thr Val Met Gly  
850 855 860

Trp Met Gly Arg Leu Ile Ser Pro Val Ser Leu Gln Thr Ser Thr Gly  
865 870 875 880

Lys Ala

<210> 11  
<211> 1141  
<212> PRT  
<213> Homo sapiens

<400> 11

Met Asp Asp Ser Gly Glu Leu Gly Gly Leu Glu Thr Met Glu Thr Leu  
1 5 10 15

Thr Glu Leu Gly Asp Glu Leu Thr Leu Gly Asp Ile Asp Glu Met Leu  
20 25 30

Gln Phe Val Ser Asn Gln Val Gly Glu Phe Pro Asp Leu Phe Ser Glu  
35 40 45

Gln Leu Cys Ser Ser Phe Pro Gly Ser Gly Gly Ser Gly Ser Ser Ser  
50 55 60

ES 2 819 055 T3

Gly Ser Ser Gly Ser Ser Ser Ser Ser Ser Asn Gly Arg Gly Ser Ser  
 65 70 75 80  
  
 Ser Gly Ala Val Asp Pro Ser Val Gln Arg Ser Phe Thr Gln Val Thr  
 85 90 95  
  
 Leu Pro Ser Phe Ser Pro Ser Ala Ala Ser Pro Gln Ala Pro Thr Leu  
 100 105 110  
  
 Gln Val Lys Val Ser Pro Thr Ser Val Pro Thr Thr Pro Arg Ala Thr  
 115 120 125  
  
 Pro Ile Leu Gln Pro Arg Pro Gln Pro Gln Pro Gln Pro Gln Thr Gln  
 130 135 140  
  
 Leu Gln Gln Gln Thr Val Met Ile Thr Pro Thr Phe Ser Thr Thr Pro  
 145 150 155 160  
  
 Gln Thr Arg Ile Ile Gln Gln Pro Leu Ile Tyr Gln Asn Ala Ala Thr  
 165 170 175  
  
 Ser Phe Gln Val Leu Gln Pro Gln Val Gln Ser Leu Val Thr Ser Ser  
 180 185 190  
  
 Gln Val Gln Pro Val Thr Ile Gln Gln Gln Val Gln Thr Val Gln Ala  
 195 200 205  
  
 Gln Arg Val Leu Thr Gln Thr Ala Asn Gly Thr Leu Gln Thr Leu Ala  
 210 215 220  
  
 Pro Ala Thr Val Gln Thr Val Ala Ala Pro Gln Val Gln Gln Val Pro  
 225 230 235 240  
  
 Val Leu Val Gln Pro Gln Ile Ile Lys Thr Asp Ser Leu Val Leu Thr  
 245 250 255  
  
 Thr Leu Lys Thr Asp Gly Ser Pro Val Met Ala Ala Val Gln Asn Pro  
 260 265 270  
  
 Ala Leu Thr Ala Leu Thr Thr Pro Ile Gln Thr Ala Ala Leu Gln Val  
 275 280 285  
  
 Pro Thr Leu Val Gly Ser Ser Gly Thr Ile Leu Thr Thr Met Pro Val  
 290 295 300  
  
 Met Met Gly Gln Glu Lys Val Pro Ile Lys Gln Val Pro Gly Gly Val  
 305 310 315 320



ES 2 819 055 T3

Lys Gln Leu Glu Pro Pro Lys Glu Gly Glu Arg Arg Thr Thr His Asn  
 325 330 335

Ile Ile Glu Lys Arg Tyr Arg Ser Ser Ile Asn Asp Lys Ile Ile Glu  
 340 345 350

Leu Lys Asp Leu Val Met Gly Thr Asp Ala Lys Met His Lys Ser Gly  
 355 360 365

Val Leu Arg Lys Ala Ile Asp Tyr Ile Lys Tyr Leu Gln Gln Val Asn  
 370 375 380

His Lys Leu Arg Gln Glu Asn Met Val Leu Lys Leu Ala Asn Gln Lys  
 385 390 395 400

Asn Lys Leu Leu Lys Gly Ile Asp Leu Gly Ser Leu Val Asp Asn Glu  
 405 410 415

Val Asp Leu Lys Ile Glu Asp Phe Asn Gln Asn Val Leu Leu Met Ser  
 420 425 430

Pro Pro Ala Ser Asp Ser Gly Ser Gln Ala Gly Phe Ser Pro Tyr Ser  
 435 440 445

Ile Asp Ser Glu Pro Gly Ser Pro Leu Leu Asp Asp Ala Lys Val Lys  
 450 455 460

Asp Glu Pro Asp Ser Pro Pro Val Ala Leu Gly Met Val Asp Arg Ser  
 465 470 475 480

Arg Ile Leu Leu Cys Val Leu Thr Phe Leu Cys Leu Ser Phe Asn Pro  
 485 490 495

Leu Thr Ser Leu Leu Gln Trp Gly Gly Ala His Asp Ser Asp Gln His  
 500 505 510

Pro His Ser Gly Ser Gly Arg Ser Val Leu Ser Phe Glu Ser Gly Ser  
 515 520 525

Gly Gly Trp Phe Asp Trp Met Met Pro Thr Leu Leu Leu Trp Leu Val  
 530 535 540

Asn Gly Val Ile Val Leu Ser Val Phe Val Lys Leu Leu Val His Gly  
 545 550 555 560

Glu Pro Val Ile Arg Pro His Ser Arg Ser Ser Val Thr Phe Trp Arg  
 565 570 575

ES 2 819 055 T3

His Arg Lys Gln Ala Asp Leu Asp Leu Ala Arg Gly Asp Phe Ala Ala  
 580 585 590  
 Ala Ala Gly Asn Leu Gln Thr Cys Leu Ala Val Leu Gly Arg Ala Leu  
 595 600 605  
 Pro Thr Ser Arg Leu Asp Leu Ala Cys Ser Leu Ser Trp Asn Val Ile  
 610 615 620  
 Arg Tyr Ser Leu Gln Lys Leu Arg Leu Val Arg Trp Leu Leu Lys Lys  
 625 630 635 640  
 Val Phe Gln Cys Arg Arg Ala Thr Pro Ala Thr Glu Ala Gly Phe Glu  
 645 650 655  
 Asp Glu Ala Lys Thr Ser Ala Arg Asp Ala Ala Leu Ala Tyr His Arg  
 660 665 670  
 Leu His Gln Leu His Ile Thr Gly Lys Leu Pro Ala Gly Ser Ala Cys  
 675 680 685  
 Ser Asp Val His Met Ala Leu Cys Ala Val Asn Leu Ala Glu Cys Ala  
 690 695 700  
 Glu Glu Lys Ile Pro Pro Ser Thr Leu Val Glu Ile His Leu Thr Ala  
 705 710 715 720  
 Ala Met Gly Leu Lys Thr Arg Cys Gly Gly Lys Leu Gly Phe Leu Ala  
 725 730 735  
 Ser Tyr Phe Leu Ser Arg Ala Gln Ser Leu Cys Gly Pro Glu His Ser  
 740 745 750  
 Ala Val Pro Asp Ser Leu Arg Trp Leu Cys His Pro Leu Gly Gln Lys  
 755 760 765  
 Phe Phe Met Glu Arg Ser Trp Ser Val Lys Ser Ala Ala Lys Glu Ser  
 770 775 780  
 Leu Tyr Cys Ala Gln Arg Asn Pro Ala Asp Pro Ile Ala Gln Val His  
 785 790 795 800  
 Gln Ala Phe Cys Lys Asn Leu Leu Glu Arg Ala Ile Glu Ser Leu Val  
 805 810 815  
 Lys Pro Gln Ala Lys Lys Lys Ala Gly Asp Gln Glu Glu Glu Ser Cys  
 820 825 830  
 Glu Phe Ser Ser Ala Leu Glu Tyr Leu Lys Leu Leu His Ser Phe Val

ES 2 819 055 T3

835	840	845																			
Asp	Ser	Val	Gly	Val	Met	Ser	Pro	Pro	Leu	Ser	Arg	Ser	Ser	Val	Leu						
850						855					860										
Lys	Ser	Ala	Leu	Gly	Pro	Asp	Ile	Ile	Cys	Arg	Trp	Trp	Thr	Ser	Ala						
865					870					875					880						
Ile	Thr	Val	Ala	Ile	Ser	Trp	Leu	Gln	Gly	Asp	Asp	Ala	Ala	Val	Arg						
				885					890					895							
Ser	His	Phe	Thr	Lys	Val	Glu	Arg	Ile	Pro	Lys	Ala	Leu	Glu	Val	Thr						
			900					905					910								
Glu	Ser	Pro	Leu	Val	Lys	Ala	Ile	Phe	His	Ala	Cys	Arg	Ala	Met	His						
		915					920					925									
Ala	Ser	Leu	Pro	Gly	Lys	Ala	Asp	Gly	Gln	Gln	Ser	Ser	Phe	Cys	His						
		930				935					940										
Cys	Glu	Arg	Ala	Ser	Gly	His	Leu	Trp	Ser	Ser	Leu	Asn	Val	Ser	Gly						
945					950					955					960						
Ala	Thr	Ser	Asp	Pro	Ala	Leu	Asn	His	Val	Val	Gln	Leu	Leu	Thr	Cys						
				965					970					975							
Asp	Leu	Leu	Leu	Ser	Leu	Arg	Thr	Ala	Leu	Trp	Gln	Lys	Gln	Ala	Ser						
			980					985					990								
Ala	Ser	Gln	Ala	Val	Gly	Glu	Thr	Tyr	His	Ala	Ser	Gly	Ala	Glu	Leu						
		995					1000					1005									
Ala	Gly	Phe	Gln	Arg	Asp	Leu	Gly	Ser	Leu	Arg	Arg	Leu	Ala	His							
	1010					1015					1020										
Ser	Phe	Arg	Pro	Ala	Tyr	Arg	Lys	Val	Phe	Leu	His	Glu	Ala	Thr							
	1025					1030					1035										
Val	Arg	Leu	Met	Ala	Gly	Ala	Ser	Pro	Thr	Arg	Thr	His	Gln	Leu							
	1040					1045					1050										
Leu	Glu	His	Ser	Leu	Arg	Arg	Arg	Thr	Thr	Gln	Ser	Thr	Lys	His							
	1055					1060					1065										
Gly	Glu	Val	Asp	Ala	Trp	Pro	Gly	Gln	Arg	Glu	Arg	Ala	Thr	Ala							
	1070					1075					1080										
Ile	Leu	Leu	Ala	Cys	Arg	His	Leu	Pro	Leu	Ser	Phe	Leu	Ser	Ser							
	1085					1090					1095										

ES 2 819 055 T3

Pro Gly Gln Arg Ala Val Leu Leu Ala Glu Ala Ala Arg Thr Leu  
 1100 1105 1110

Glu Lys Val Gly Asp Arg Arg Ser Cys Asn Asp Cys Gln Gln Met  
 1115 1120 1125

Ile Val Lys Leu Gly Gly Gly Thr Ala Ile Ala Ala Ser  
 1130 1135 1140

<210> 12  
 <211> 976  
 <212> PRT  
 <213> Homo sapiens  
 <400> 12

Met Ala Gln Val Ala Met Ser Thr Leu Pro Val Glu Asp Glu Glu Ser  
 1 5 10 15

Ser Glu Ser Arg Met Val Val Thr Phe Leu Met Ser Ala Leu Glu Ser  
 20 25 30

Met Cys Lys Glu Leu Ala Lys Ser Lys Ala Glu Val Ala Cys Ile Ala  
 35 40 45

Val Tyr Glu Thr Asp Val Phe Val Val Gly Thr Glu Arg Gly Arg Ala  
 50 55 60

Phe Val Asn Thr Arg Lys Asp Phe Gln Lys Asp Phe Val Lys Tyr Cys  
 65 70 75 80

Val Glu Glu Glu Glu Lys Ala Ala Glu Met His Lys Met Lys Ser Thr  
 85 90 95

Thr Gln Ala Asn Arg Met Ser Val Asp Ala Val Glu Ile Glu Thr Leu  
 100 105 110

Arg Lys Thr Val Glu Asp Tyr Phe Cys Phe Cys Tyr Gly Lys Ala Leu  
 115 120 125

Gly Lys Ser Thr Val Val Pro Val Pro Tyr Glu Lys Met Leu Arg Asp  
 130 135 140

Gln Ser Ala Val Val Val Gln Gly Leu Pro Glu Gly Val Ala Phe Lys  
 145 150 155 160

His Pro Glu Asn Tyr Asp Leu Ala Thr Leu Lys Trp Ile Leu Glu Asn  
 165 170 175

ES 2 819 055 T3

Lys Ala Gly Ile Ser Phe Ile Ile Lys Arg Pro Phe Leu Glu Pro Lys  
180 185 190

Lys His Val Gly Gly Arg Val Met Val Thr Asp Ala Asp Arg Ser Ile  
195 200 205

Leu Ser Pro Gly Gly Ser Cys Gly Pro Ile Lys Val Lys Thr Glu Pro  
210 215 220

Thr Glu Asp Ser Gly Ile Ser Leu Glu Met Ala Ala Val Thr Val Lys  
225 230 235 240

Glu Glu Ser Glu Asp Pro Asp Tyr Tyr Gln Tyr Asn Ile Gln Gly Pro  
245 250 255

Ser Glu Thr Asp Asp Val Asp Glu Lys Gln Pro Leu Ser Lys Pro Leu  
260 265 270

Gln Gly Ser His His Ser Ser Glu Gly Asn Glu Gly Thr Glu Met Glu  
275 280 285

Val Pro Ala Glu Asp Asp Asp Tyr Ser Pro Pro Ser Lys Arg Pro Lys  
290 295 300

Ala Asn Glu Leu Pro Gln Pro Pro Val Pro Glu Pro Ala Asn Ala Gly  
305 310 315 320

Lys Arg Lys Val Arg Glu Phe Asn Phe Glu Lys Trp Asn Ala Arg Ile  
325 330 335

Thr Asp Leu Arg Lys Gln Val Glu Glu Leu Phe Glu Arg Lys Tyr Ala  
340 345 350

Gln Ala Ile Lys Ala Lys Gly Pro Val Thr Ile Pro Tyr Pro Leu Phe  
355 360 365

Gln Ser His Val Glu Asp Leu Tyr Val Glu Gly Leu Pro Glu Gly Ile  
370 375 380

Pro Phe Arg Arg Pro Ser Thr Tyr Gly Ile Pro Arg Leu Glu Arg Ile  
385 390 395 400

Leu Leu Ala Lys Glu Arg Ile Arg Phe Val Ile Lys Lys His Glu Leu  
405 410 415

Leu Asn Ser Thr Arg Glu Asp Leu Gln Leu Asp Lys Pro Ala Ser Gly  
420 425 430

Val Lys Glu Glu Trp Tyr Ala Arg Ile Thr Lys Leu Arg Lys Met Val



ES 2 819 055 T3

435	440	445																			
Asp	Gln	Leu	Phe	Cys	Lys	Lys	Phe	Ala	Glu	Ala	Leu	Gly	Ser	Thr	Glu						
450						455					460										
Ala	Lys	Ala	Val	Pro	Tyr	Gln	Lys	Phe	Glu	Ala	His	Pro	Asn	Asp	Leu						
465					470					475					480						
Tyr	Val	Glu	Gly	Leu	Pro	Glu	Asn	Ile	Pro	Phe	Arg	Ser	Pro	Ser	Trp						
				485					490					495							
Tyr	Gly	Ile	Pro	Arg	Leu	Glu	Lys	Ile	Ile	Gln	Val	Gly	Asn	Arg	Ile						
			500					505					510								
Lys	Phe	Val	Ile	Lys	Arg	Pro	Glu	Leu	Leu	Thr	His	Ser	Thr	Thr	Glu						
		515					520					525									
Val	Thr	Gln	Pro	Arg	Thr	Asn	Thr	Pro	Val	Lys	Glu	Asp	Trp	Asn	Val						
	530					535					540										
Arg	Ile	Thr	Lys	Leu	Arg	Lys	Gln	Val	Glu	Glu	Ile	Phe	Asn	Leu	Lys						
545					550					555					560						
Phe	Ala	Gln	Ala	Leu	Gly	Leu	Thr	Glu	Ala	Val	Lys	Val	Pro	Tyr	Pro						
				565					570					575							
Val	Phe	Glu	Ser	Asn	Pro	Glu	Phe	Leu	Tyr	Val	Glu	Gly	Leu	Pro	Glu						
			580					585					590								
Gly	Ile	Pro	Phe	Arg	Ser	Pro	Thr	Trp	Phe	Gly	Ile	Pro	Arg	Leu	Glu						
		595					600						605								
Arg	Ile	Val	Arg	Gly	Ser	Asn	Lys	Ile	Lys	Phe	Val	Val	Lys	Lys	Pro						
610						615					620										
Glu	Leu	Val	Ile	Ser	Tyr	Leu	Pro	Pro	Gly	Met	Ala	Ser	Lys	Ile	Asn						
625					630					635					640						
Thr	Lys	Ala	Leu	Gln	Ser	Pro	Lys	Arg	Pro	Arg	Ser	Pro	Gly	Ser	Asn						
				645					650					655							
Ser	Lys	Val	Pro	Glu	Ile	Glu	Val	Thr	Val	Glu	Gly	Pro	Asn	Asn	Asn						
			660					665					670								
Asn	Pro	Gln	Thr	Ser	Ala	Val	Arg	Thr	Pro	Thr	Gln	Thr	Asn	Gly	Ser						
		675					680					685									
Asn	Val	Pro	Phe	Lys	Pro	Arg	Gly	Arg	Glu	Phe	Ser	Phe	Glu	Ala	Trp						
	690					695					700										

ES 2 819 055 T3

Asn Ala Lys Ile Thr Asp Leu Lys Gln Lys Val Glu Asn Leu Phe Asn  
 705 710 715 720  
 Glu Lys Cys Gly Glu Ala Leu Gly Leu Lys Gln Ala Val Lys Val Pro  
 725 730 735  
 Phe Ala Leu Phe Glu Ser Phe Pro Glu Asp Phe Tyr Val Glu Gly Leu  
 740 745 750  
 Pro Glu Gly Val Pro Phe Arg Arg Pro Ser Thr Phe Gly Ile Pro Arg  
 755 760 765  
 Leu Glu Lys Ile Leu Arg Asn Lys Ala Lys Ile Lys Phe Ile Ile Lys  
 770 775 780  
 Lys Pro Glu Met Phe Glu Thr Ala Ile Lys Glu Ser Thr Ser Ser Lys  
 785 790 795 800  
 Ser Pro Pro Arg Lys Ile Asn Ser Ser Pro Asn Val Asn Thr Thr Ala  
 805 810 815  
 Ser Gly Val Glu Asp Leu Asn Ile Ile Gln Val Thr Ile Pro Asp Asp  
 820 825 830  
 Asp Asn Glu Arg Leu Ser Lys Val Glu Lys Ala Arg Gln Leu Arg Glu  
 835 840 845  
 Gln Val Asn Asp Leu Phe Ser Arg Lys Phe Gly Glu Ala Ile Gly Met  
 850 855 860  
 Gly Phe Pro Val Lys Val Pro Tyr Arg Lys Ile Thr Ile Asn Pro Gly  
 865 870 875 880  
 Cys Val Val Val Asp Gly Met Pro Pro Gly Val Ser Phe Lys Ala Pro  
 885 890 895  
 Ser Tyr Leu Glu Ile Ser Ser Met Arg Arg Ile Leu Asp Ser Ala Glu  
 900 905 910  
 Phe Ile Lys Phe Thr Val Ile Arg Pro Phe Pro Gly Leu Val Ile Asn  
 915 920 925  
 Asn Gln Leu Val Asp Gln Ser Glu Ser Glu Gly Pro Val Ile Gln Glu  
 930 935 940  
 Ser Ala Glu Pro Ser Gln Leu Glu Val Pro Ala Thr Glu Glu Ile Lys  
 945 950 955 960  
 Glu Thr Asp Gly Ser Ser Gln Ile Lys Gln Glu Pro Asp Pro Thr Trp  
 965 970 975

ES 2 819 055 T3

<211> 21  
<212> ADN  
<213> Artificial

5 <220>  
<223> Cebador directo T7\_up2

<400> 13  
tgctaaggac aacgttatcg g 21

10 <210> 14  
<211> 18  
<212> ADN  
<213> Artificial

15 <220>  
<223> Cebador inverso T7\_down2

<400> 14  
ttgataccgg acgttcac 18

20

## REIVINDICACIONES

1. Un método para detectar un autoanticuerpo en una muestra, que comprende:

- 5 a) poner en contacto una muestra de un sujeto con una entidad capturadora de anticuerpo (ECA), en donde dicha ECA es una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 4 o un péptido o proteína que tiene al menos el 90% de identidad de secuencia con respecto a SEQ ID NO: 4 y que es reconocida por un autoanticuerpo contra SEQ ID NO: 4;
- 10 y  
b) detectar la formación de un complejo autoanticuerpo-ECA,

en donde la detección de dicho complejo autoanticuerpo-ECA es indicativa de la presencia de dicho autoanticuerpo en dicha muestra, en donde la muestra es sangre, plasma o suero.

15 2. Un método según la reivindicación 1, en donde la muestra es una muestra de un sujeto que se sospecha tiene cáncer colorrectal (CCR).

3. El método según la reivindicación 1 o 2, que comprende la detección de un autoanticuerpo contra una ECA que comprende la secuencia de aminoácidos mostrada en la SEQ ID NO: 4 o una proteína o péptido que tiene al menos el 90% de identidad de secuencia con respecto a SEQ ID NO: 4 que tiene un epítipo que es reconocible por un autoanticuerpo contra SEQ ID NO: 4 y, además, la detección de un autoanticuerpo seleccionado del grupo que consiste en:

- 25 (i) un autoanticuerpo contra una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 2 o una proteína o péptido que tiene al menos el 90% de identidad de secuencia con respecto a SEQ ID NO: 2 que contiene un epítipo que es reconocible por un autoanticuerpo contra SEQ ID NO: 2;
- 30 (ii) un autoanticuerpo contra una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 6 o una proteína o péptido que tiene al menos el 90% de identidad de secuencia con respecto a SEQ ID NO: 6 que contiene un epítipo que es reconocible por un autoanticuerpo contra SEQ ID NO:6;
- (iii) un autoanticuerpo contra una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 3 o una proteína o péptido que tiene al menos el 90% de identidad de secuencia con respecto a SEQ ID NO: 3 que contiene un epítipo que es reconocible por un autoanticuerpo contra SEQ ID NO:3;
- 35 (iv) un autoanticuerpo contra una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 5 o una proteína o péptido que tiene al menos el 90% de identidad de secuencia con respecto a SEQ ID NO: 5 que contiene un epítipo que es reconocible por un autoanticuerpo contra SEQ ID NO:5;
- 40 (v) un autoanticuerpo contra una ECA que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 1 o una proteína o péptido que tiene al menos el 90% de identidad de secuencia con respecto a SEQ ID NO: 1 que contiene un epítipo que es reconocible por un autoanticuerpo contra SEQ ID NO:1; y
- 45 (vi) cualquier combinación de autoanticuerpos (i) a (v).

4. El método según la reivindicación 1 o 2, en donde dicha ECA es sulfatasa 1 (SULF1) o una proteína o péptido que tiene al menos el 90% de identidad de secuencia con respecto a SULF1 que contiene un epítipo reconocible por un autoanticuerpo contra SULF1,

50 o alternativamente

en donde dicha ECA es un fago que comprende la secuencia de aminoácidos en SEQ ID NO: 4 o una proteína o péptido que tiene al menos el 90% de identidad de secuencia con respecto a SEQ ID NO: 4 que contiene un epítipo reconocible por autoanticuerpo contra SEQ ID NO: 4, en donde dicha secuencia de aminoácidos está expuesta en la superficie del fago.

55 5. El método según la reivindicación 2, que comprende además correlacionar la formación del complejo autoanticuerpo-ECA en la muestra del sujeto con un diagnóstico de CCR.

60 6. El método según la reivindicación 2, que comprende además comparar la inmunorreactividad de la muestra con la inmunorreactividad de una segunda muestra del mismo sujeto en un periodo de tiempo posterior.

7. El método según la reivindicación 6, en donde dicha segunda muestra del sujeto se ha obtenido después de que dicho sujeto se haya tratado para CCR.

65

8. Una entidad capturadora de anticuerpos (ECA) que comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 4, en donde dicha ECA no es la proteína sulfatasa-1 (SULF1).
- 5 9. La entidad capturadora de anticuerpos (ECA) según la reivindicación 8, en donde dicha ECA es un fago que comprende la secuencia de aminoácidos de SEQ ID NO: 4, en donde dicha secuencia de aminoácidos está expuesta en la superficie del fago.
- 10 10. Una composición que comprende una entidad capturadora de anticuerpo (ECA) según la reivindicación 8.
- 10 11. La composición según la reivindicación 10, que comprende, al menos, una ECA según la reivindicación 8, y, al menos, una proteína seleccionada del grupo que consiste en la proteína sulfatasa 1 (SULF1) o una proteína o péptido que tiene al menos el 90% de identidad de secuencia con respecto a SULF1 y reconocible por un autoanticuerpo contra SULF1, la proteína MST1 o una proteína o péptido que tiene al menos el 90% de identidad de secuencia con respecto a MST1 y reconocible por un autoanticuerpo contra MST1, y sus combinaciones.
- 15 12. Una composición que comprende la proteína sulfatasa 1 (SULF1) o una proteína o péptido que tiene al menos el 90% de identidad de secuencia con respecto a SULF1 y reconocible por un autoanticuerpo contra SULF1 y la proteína MST1 o una proteína o péptido que tiene al menos el 90% de identidad de secuencia con respecto a MST1 y reconocible por un autoanticuerpo contra MST1 y una entidad capturadora de anticuerpo (ECA) según la reivindicación 8.
- 20 13. Un kit que comprende una composición según cualquiera de las reivindicaciones 10 a 12.
- 25 14. Un uso in vitro de un kit según la reivindicación 13 para detectar un anticuerpo en una muestra, o para detectar un autoanticuerpo en un sujeto que se sospecha tiene cáncer colorrectal (CCR), o para diagnosticar si un sujeto tiene CCR, o para determinar el riesgo de que un sujeto desarrolle CCR, o para seguir la evolución de CCR en un sujeto, o para evaluar la eficacia de un tratamiento contra CCR, o para predecir la supervivencia de un sujeto que tiene CCR.
- 30

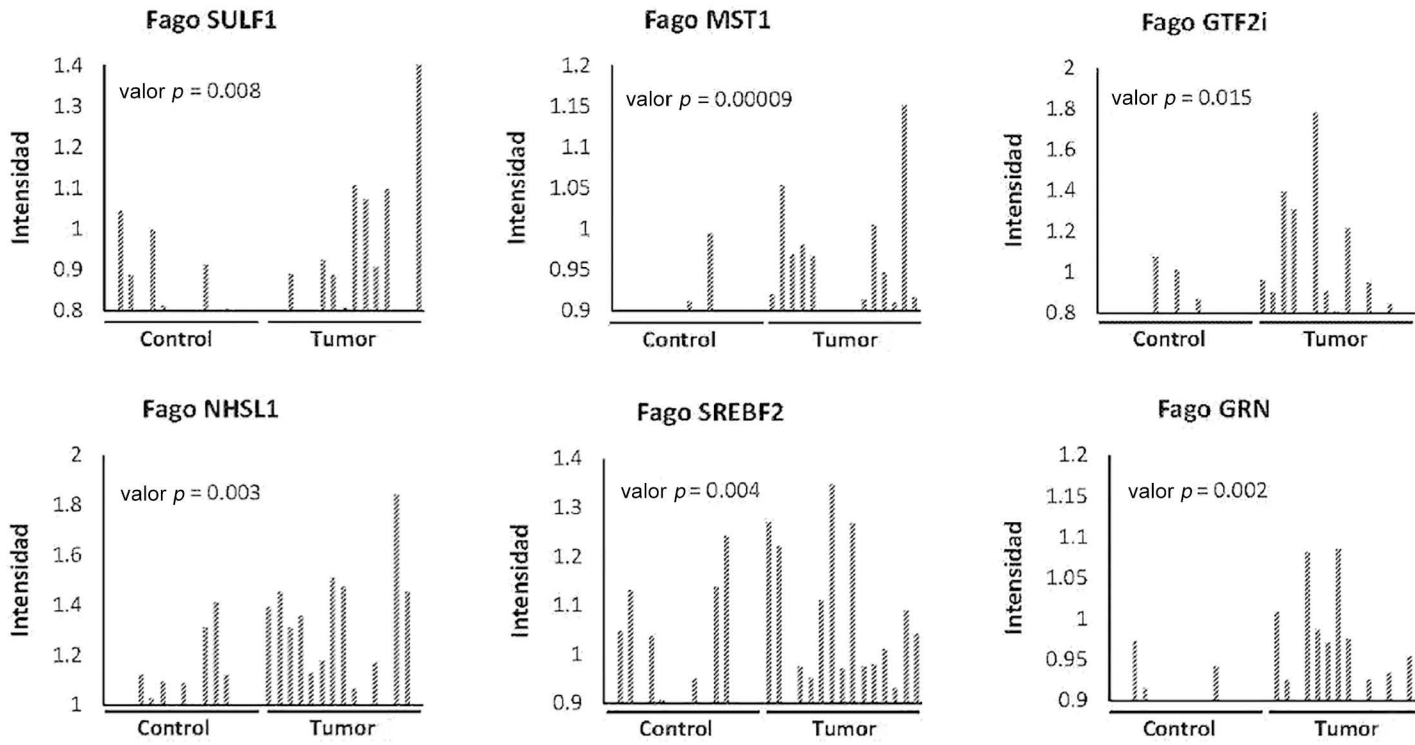


Figura 1

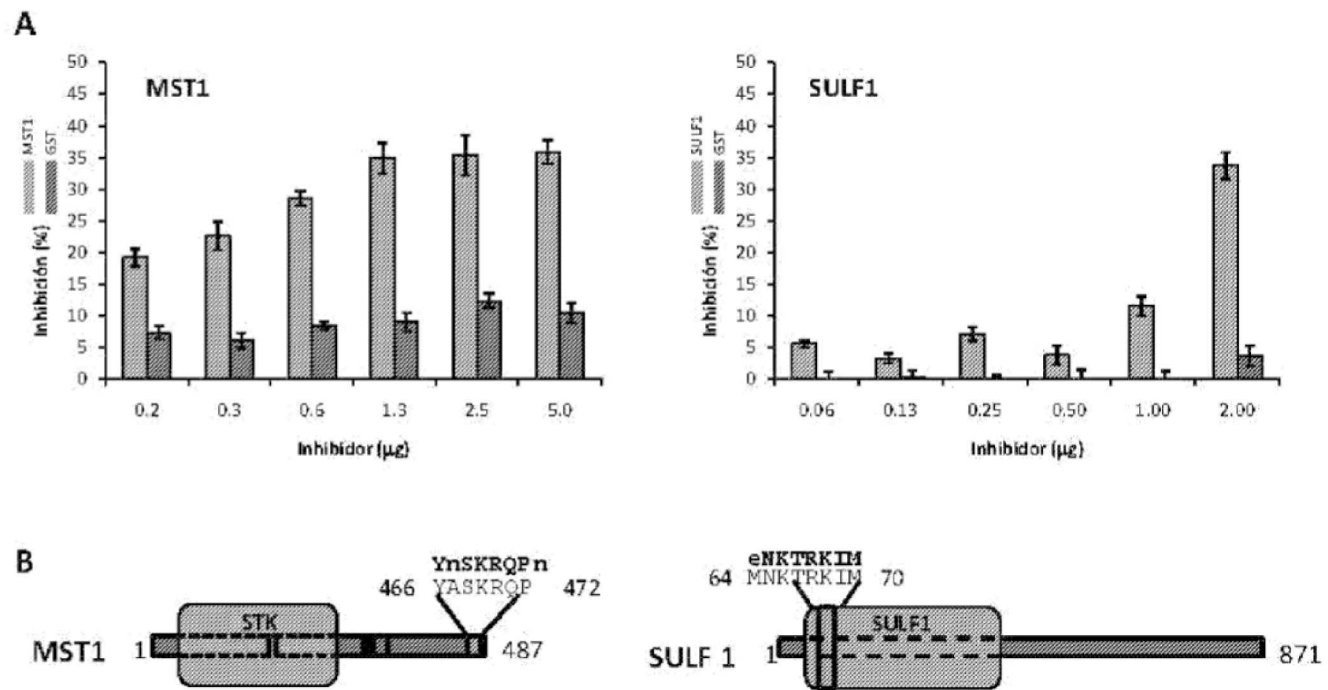


Figura 2



Figura 3A

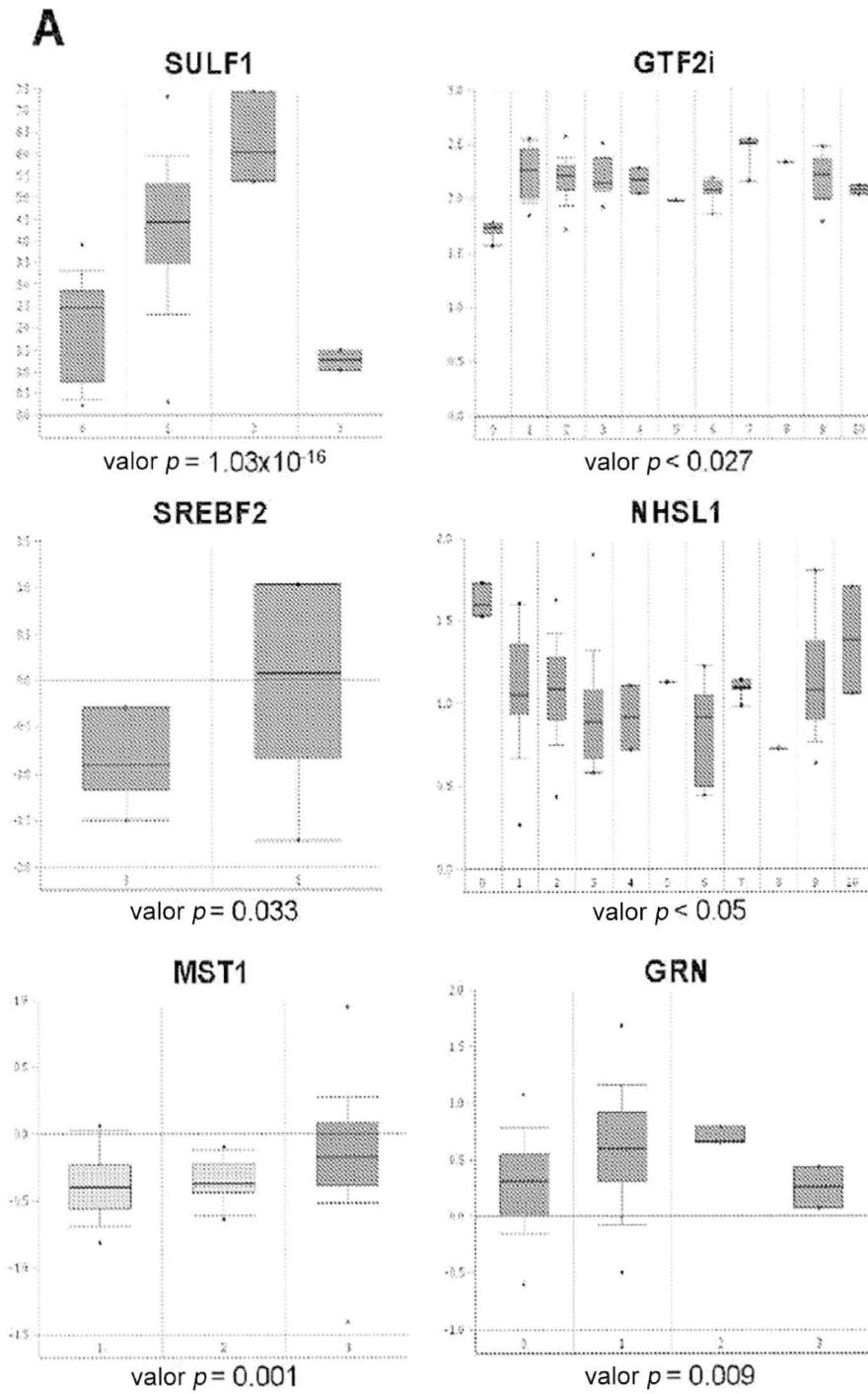


Figura 3B

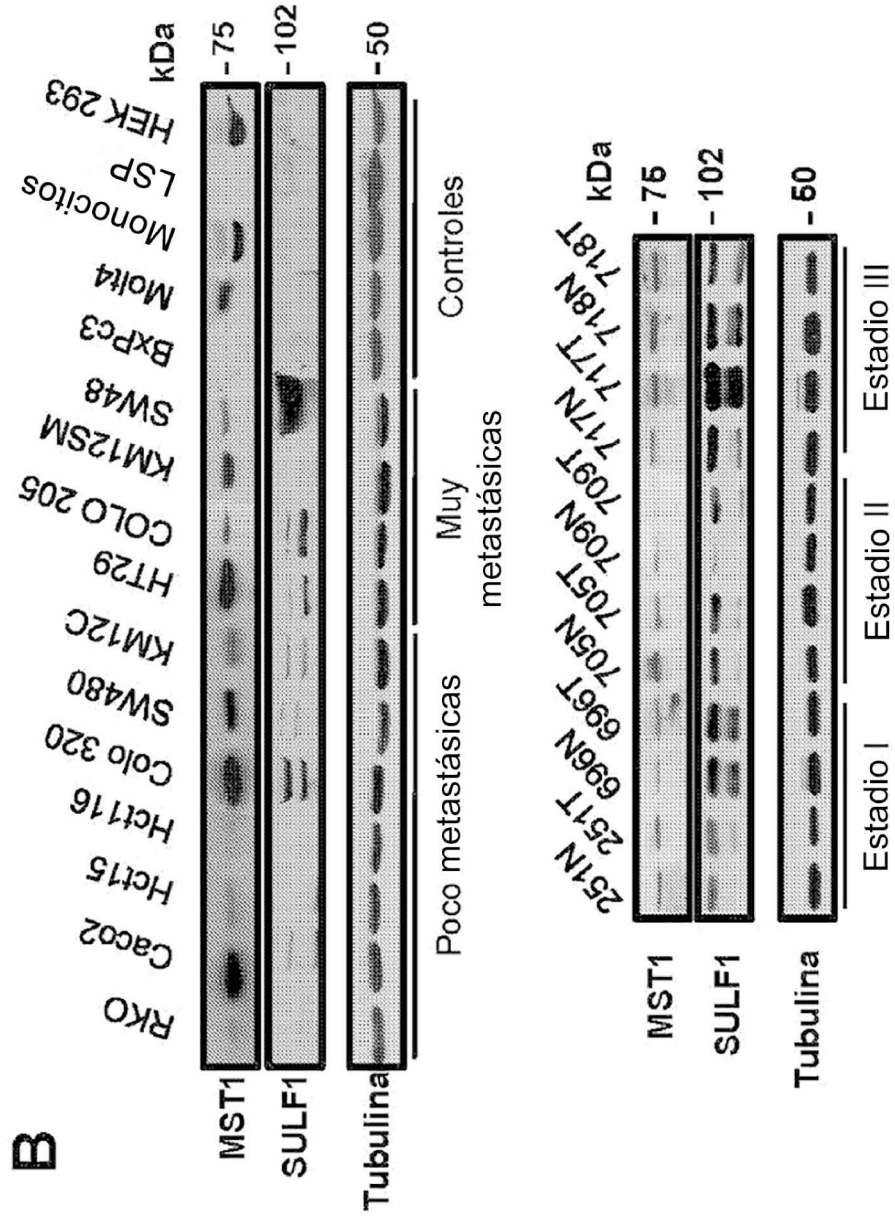
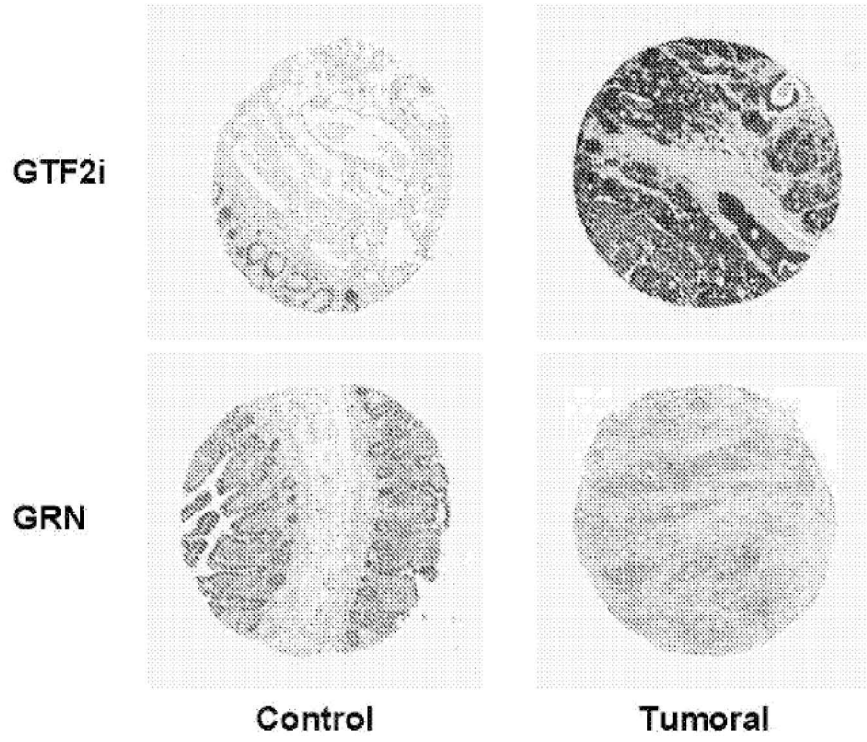


Figura 3C

**C**



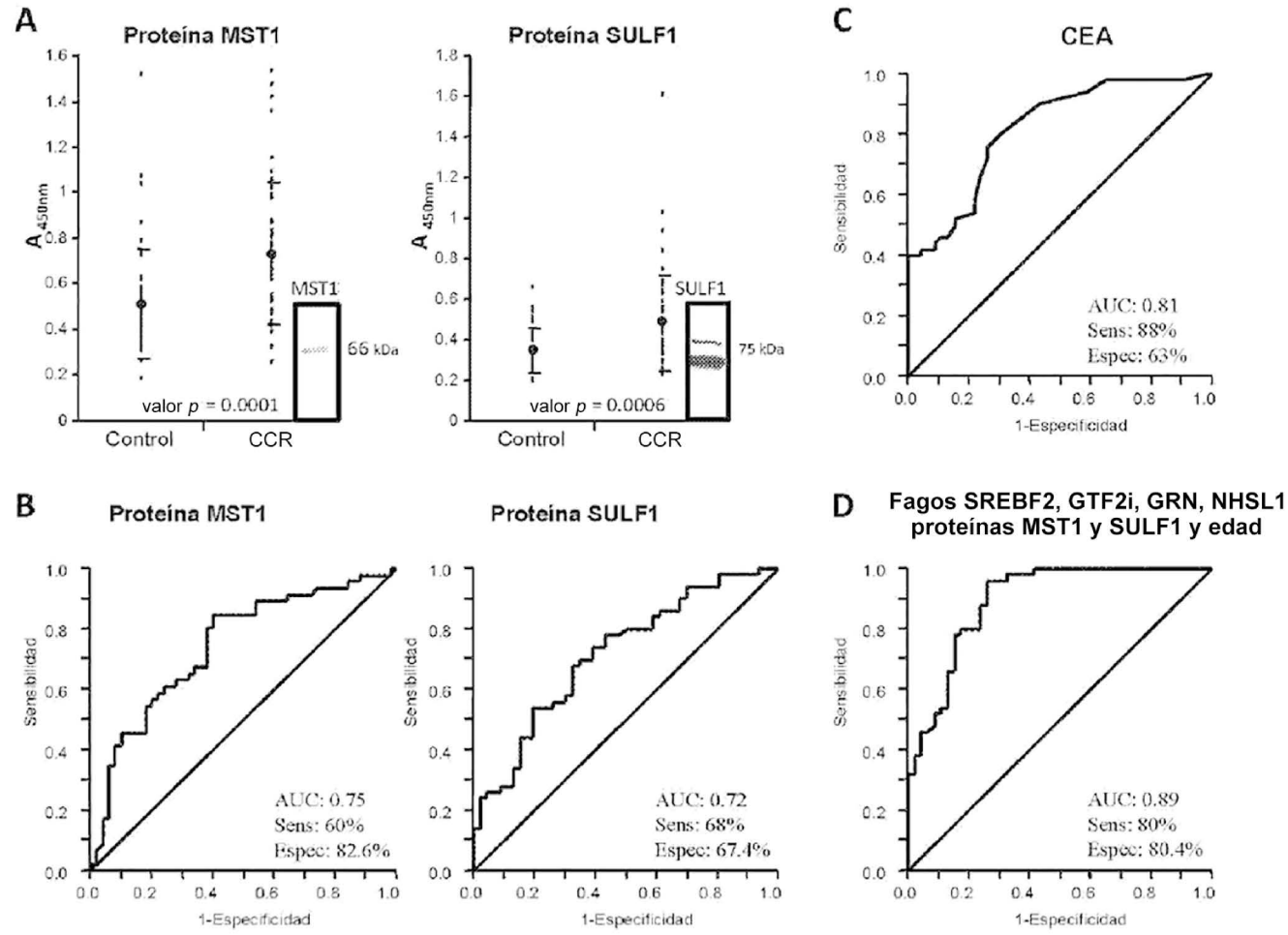


Figura 4

Figura 5

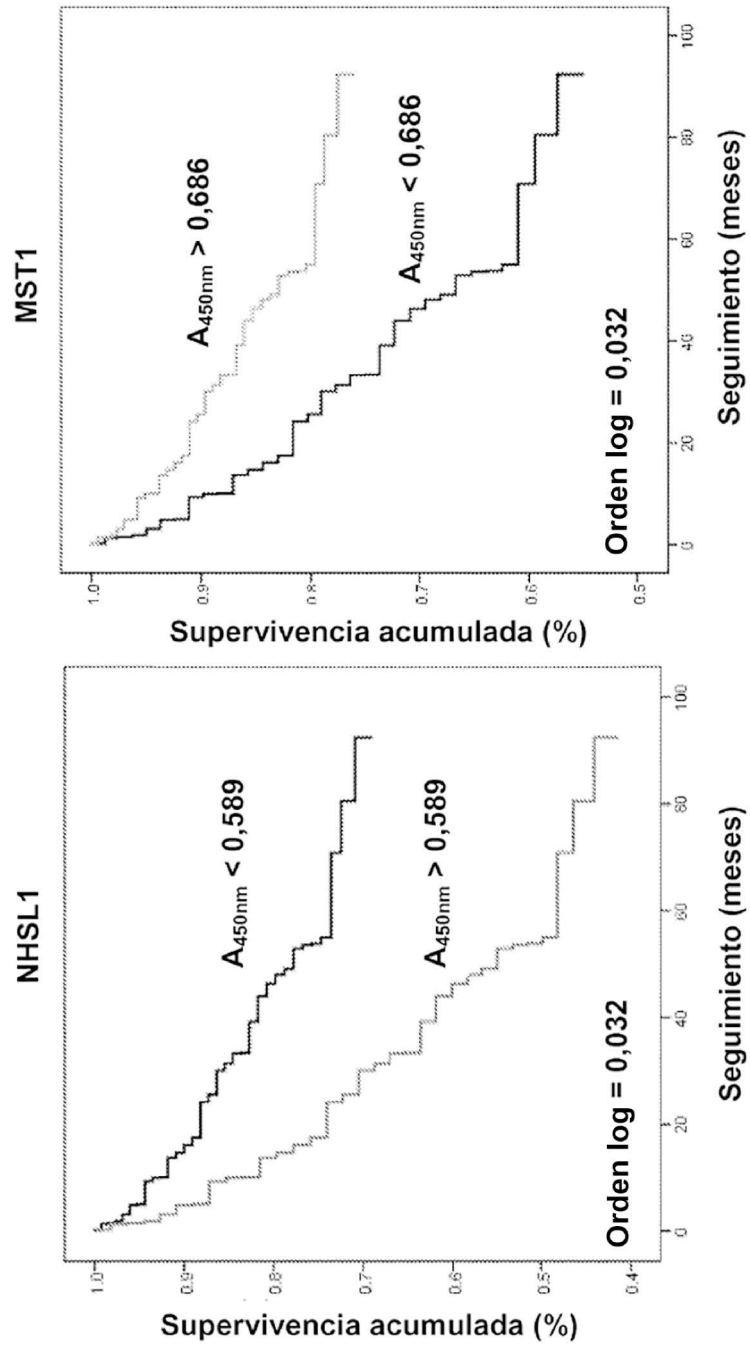
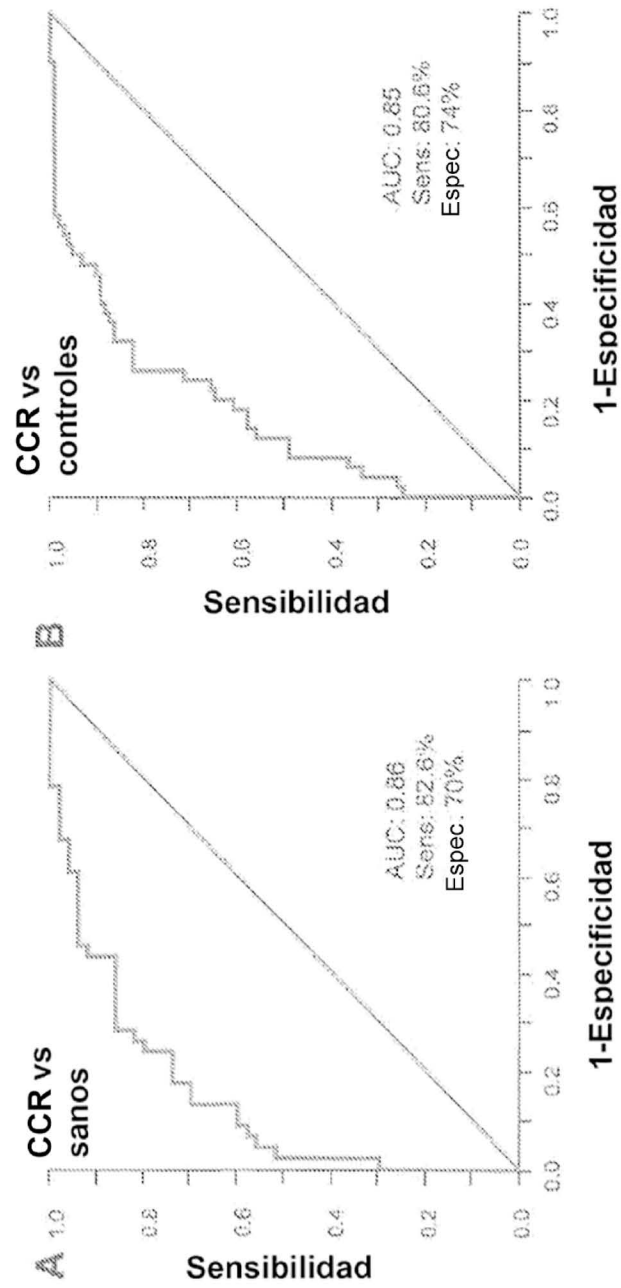


Figura 6



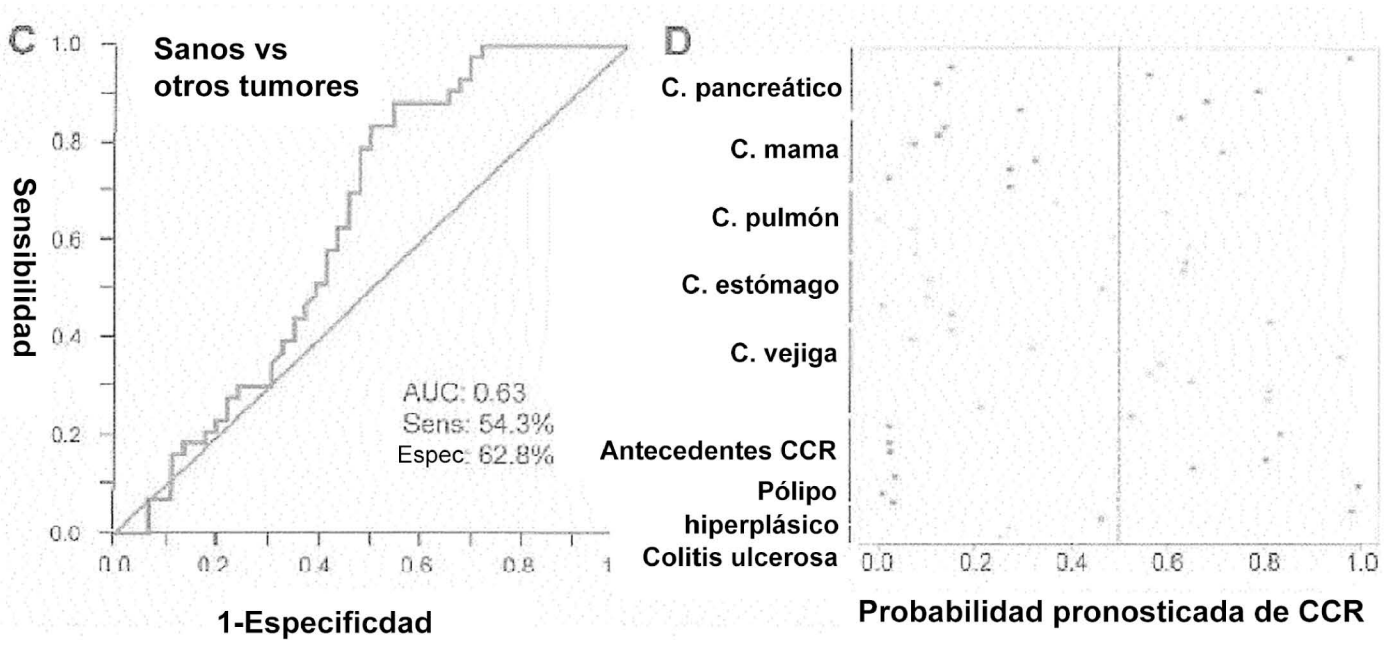


Figura 6 (cont.)