

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 762 403**

51 Int. Cl.:

**G01N 33/574** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **26.04.2013 PCT/NL2013/050316**

87 Fecha y número de publicación internacional: **31.10.2013 WO13162368**

96 Fecha de presentación y número de la solicitud europea: **26.04.2013 E 13720130 (7)**

97 Fecha y número de publicación de la concesión europea: **09.10.2019 EP 2841947**

54 Título: **Biomarcadores**

30 Prioridad:

**26.04.2012 NL 2008707**  
**08.02.2013 NL 2010276**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**25.05.2020**

73 Titular/es:

**STICHTING VUMC (100.0%)**  
**De Boelelaan 1117**  
**1081 HV Amsterdam, NL**

72 Inventor/es:

**BOSCH, LINDA JANNA WILLEMEN;**  
**DE WIT, MEIKE;**  
**PINTO MORAIS DE CARVALHO, BEATRIZ;**  
**FIJNEMAN, REMONDUS JOHANNES ADRIAAN;**  
**MEIJER, GERRIT ALBERT;**  
**JIMENEZ, CORNELIA RAMONA;**  
**PIERSMA, SANDER ROGIER;**  
**PHAM, VIET THANG y**  
**OUDGENOEG, GIDEON**

74 Agente/Representante:

**SÁEZ MAESO, Ana**

**ES 2 762 403 T3**

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

## DESCRIPCIÓN

Biomarcadores

## 5 Campo de la invención

La presente invención se refiere a la identificación de biomarcadores de precursores de cáncer colorrectal, es decir, adenomas de alto riesgo, y su uso en métodos de detección, así como a matrices para uso en métodos de detección de adenoma colorrectal avanzado.

10

## Antecedentes de la invención

El cáncer colorrectal (CRC) es un problema de salud significativo. Es el tercer cáncer más común en todo el mundo, con más de 1.200.000 casos nuevos cada año, con un desenlace fatal para casi la mitad de los pacientes (Karsa LV, et al., Best Pract Res Clin Gastroenterol 2010; 24: 381-96). Debido a que el CRC se desarrolla durante muchos años, existe una excelente oportunidad para detectar la enfermedad en una etapa temprana, curable o incluso premaligna. Esto se puede lograr mediante la detección de individuos asintomáticos. Sin embargo, las pruebas de detección disponibles actualmente son engorrosas o conllevan un riesgo de complicaciones y un tratamiento excesivo como la colonoscopia.

20

La prueba inmunoquímica de sangre oculta en heces (iFOBT), también conocida como prueba inmunoquímica fecal (FIT), es una prueba ampliamente utilizada que detecta pequeñas trazas de sangre en las heces, derivadas de lesiones como tumores que sangran en el colon. La FIT es un método no invasivo que utiliza un anticuerpo dirigido contra la hemoglobina. Sin embargo, no todos los tumores de colon sangran y, por lo tanto, la FIT tiene una sensibilidad que deja margen para mejoramiento (Duffy MJ, et al., Int J Cancer 2011; 128: 3-11; van Veen W, Mali WP. [Colorrectal cancer screening: advice from the Health Council of the Netherlands]. Ned Tijdschr Geneesk 2009; 153:A1441). Los marcadores adicionales que detectan otras características tumorales además del sangrado en las heces podrían aumentar esta sensibilidad y la posibilidad de identificar un tumor de colon. Los cambios moleculares resultantes del proceso neoplásico incluyen cambios en la expresión de proteínas. Las proteínas para las cuales se incrementa la expresión tienen el potencial de servir como biomarcadores informativos con alto desempeño diagnóstico. El documento US 5.552.292 divulga la detección del cáncer colorrectal mediante el uso de proteínas como la mieloperoxidasa como biomarcador en las heces.

25

30

Se han realizado varios estudios para identificar proteínas en heces y sangre que pueden usarse para la detección temprana de CRC y adenomas colorrectales de alto riesgo, pero la mayoría de estas proteínas no se han validado en un entorno de detección o han fallado en mejorar las pruebas actuales para detección temprana, tal como el antígeno carcinoembrionario (CEA) o la calprotectina (Bosch LJ, et al., Molecular tests for colorectal cancer screening. Clin Colorrectal Cancer 2011; 10: 8-23). Por lo tanto, sigue existiendo la necesidad de otros marcadores proteicos específicos tumorales para mejorar las posibilidades de detección no invasivo disponibles actualmente para CRC y adenomas de alto riesgo.

40

Los avances tecnológicos recientes en espectrometría de masas pueden impulsar el descubrimiento de nuevos marcadores de proteínas (de Wit M, et al., Gut 2011; Jimenez CR, et al., J Proteomics 2010; 73: 1873-95). Los enfoques basados en la espectrometría de masas en tándem ahora tienen el poder de analizar muestras de proteínas complejas y detectar proteínas en bajas concentraciones (Cox J, Mann M. Annu Rev Biochem 2011; 80: 273-99). Aunque la mayoría de los estudios de descubrimiento de biomarcadores se han realizado utilizando material de tejido y/o líneas celulares seguido por la validación en heces o sangre, la composición química de la muestra biológica utilizada para la detección puede afectar significativamente la naturaleza de los biomarcadores que pueden identificarse. Esto se aplica especialmente a las muestras de heces, en las que el pH bajo, la actividad de proteasa y glicosidasa de bacterias, enzimas y otras sustancias pueden alterar la detección específica de estos marcadores (Young GP, Bosch LJW. Curr Colorrectal Cancer Rep 2011; 7: 62 -70). Por lo tanto, la medición de moléculas directamente en la muestra biológica que se puede usar para la detección, es decir, las heces, podría ser un enfoque valioso para proporcionarnos biomarcadores confiables que sean estables en el entorno fecal. Un estudio reciente de Ang et al., ha mostrado la viabilidad del descubrimiento de biomarcadores de proteínas en muestras de heces humanas mediante espectrometría de masas (Ang CS, Nice EC. J Proteome Res 2010; 9: 4346-55).

45

50

55

La presente invención tiene como objetivo superar algunos o todos los problemas asociados con la técnica anterior.

## 60 Sumario de la invención

60

De acuerdo con un primer aspecto de la invención, se proporciona un método para la detección de la presencia de adenoma colorrectal avanzado, comprendiendo el método: cribar una muestra de heces obtenida de un individuo para dos o más biomarcadores que comprenden mieloperoxidasa (MPO) y uno o más biomarcadores seleccionados del grupo que consiste en: azurocidina 1 (AZU1), catepsina G (CTSG), defensina, alfa 1 (DEFA1), elastasa, expresada en neutrófilos (ELANE), integrina, alfa M (subunidad del receptor 3 del componente 3 del complemento) (ITGAM), subunidad 1 de la proteína de unión a la caperuza nuclear, 80 kDa (NCBP1), proteinasa 3 (PRTN3), proteína S100 de

65

- 5 unión al calcio A9 (S100A9), inhibidor de la serpina peptidasa, clado B (ovoalbúmina), miembro 10 (SERPINB10), anexina A1 (ANXA1), repetición de IAP baculoviral que contiene 6 residuos (BIRC6), componente 2 del complejo de mantenimiento de minicromosomas (MCM2), quiescina Q6 sulfhidril oxidasa 1 (QSOX1), anexina A2 (ANXA2), inhibidor de la serpina peptidasa, clado B (ovoalbúmina), miembro 5 (SERPINB5), en las que la presencia o aumento de la expresión de MPO y uno o más biomarcadores, en relación con una muestra de control, es indicativo de que el individuo está en riesgo de padecer o padece un adenoma colorrectal avanzado. En una segunda realización, el primer aspecto de la invención comprende la detección de dos o más de los biomarcadores adicionales.
- 10 En una tercera realización, el primer aspecto de la invención comprende la detección de cuatro o más de los biomarcadores adicionales.
- En una cuarta realización, el primer aspecto de la invención comprende la detección de los dieciséis biomarcadores.
- 15 En una quinta realización, el primer aspecto de la invención, la muestra de heces se criba para los dos o más biomarcadores usando espectrometría de masas dirigida.
- En una sexta realización, el primer aspecto de la invención, la muestra de heces se criba para los dos o más biomarcadores usando agentes de unión capaces de unirse a los dos o más biomarcadores.
- 20 En una séptima realización, el primer aspecto de la invención, los agentes de unión son anticuerpos o fragmentos de los mismos.
- En la octava realización, dichos anticuerpos o fragmentos de los mismos se seleccionan del grupo que consiste en scFv, Fab y un dominio de unión de una molécula de inmunoglobulina.
- 25 En una novena realización, el primer aspecto de la invención, la detección se realiza usando una matriz, opcionalmente en la que la matriz es una matriz con base en perlas o una matriz basada en una superficie.
- En una décima realización, el primer aspecto de la invención, la muestra de control comprende una muestra biológica de un individuo que se sabe que está libre de adenoma colorrectal avanzado.
- 30 En la undécima realización, el primer aspecto de la invención, la muestra biológica también se analiza mediante la prueba inmunoquímica fecal.
- 35 En un segundo aspecto, la invención se refiere al uso de dos o más biomarcadores que comprenden mieloperoxidasa (MPO), y uno o más biomarcadores seleccionados del grupo que consiste en: azurocidina 1 (AZU1), catepsina G (CTSG), defensina, alfa 1 (DEFA1), elastasa, expresada en neutrófilos (ELANE), integrina, alfa M (subunidad del receptor 3 del componente 3 del complemento) (ITGAM), subunidad 1 de la proteína de unión a la caperuza nuclear, 80 kDa (NCBP1), proteinasa 3 (PRTN3), proteína S100 de unión al calcio A9 (S100A9), inhibidor de la serpina peptidasa, clado B (ovoalbúmina), miembro 10 (SERPINB10), anexina A1 (ANXA1), repetición de IAP baculoviral que contiene 6 residuos (BIRC6), componente 2 del complejo de mantenimiento de minicromosomas (MCM2), quiescina Q6 sulfhidril oxidasa 1 (QSOX1), anexina A2 (ANXA2), inhibidor de la serpina peptidasa, clado B (ovoalbúmina), miembro 5 (SERPINB5) como biomarcador en muestras de heces para diagnosticar o predecir la presencia de adenoma colorrectal avanzado en un individuo.
- 40
- 45 De acuerdo con un tercer aspecto, la invención se refiere al uso de una matriz para determinar si un individuo está en riesgo de padecer o padece adenoma colorrectal avanzado de acuerdo con el método del primer aspecto de la invención, la matriz comprende uno o más agentes de unión como se define en una cualquiera de las realizaciones 6 a 8 mencionadas anteriormente.
- 50 En un cuarto aspecto, la invención se refiere al uso de mieloperoxidasa (MPO) como biomarcador en muestras de heces para detectar la presencia de adenoma colorrectal avanzado.
- Una o más características descritas para cualquier aspecto de la presente invención o realizaciones preferidas o ejemplos de las mismas, descritas en el presente documento, pueden usarse junto con una o más características descritas para cualquier otro aspecto de la presente invención o realizaciones preferidas o ejemplos de los mismos descritos en el presente documento. El hecho de que una característica solo pueda describirse en relación con un aspecto, realización o ejemplo no limita su relevancia a solo ese aspecto, realización o ejemplo si es técnicamente relevante para uno o más aspectos, realizaciones o ejemplos.
- 55
- 60 Descripción detallada de la invención
- Cáncer colorrectal
- 65 El tipo de célula de cáncer de colon más común es el adenocarcinoma, que representa el 95% de los casos. Otros tipos más raros incluyen el linfoma y el carcinoma de células escamosas. El adenocarcinoma colorrectal surge de

lesiones precursoras llamadas adenomas, de las cuales solo una minoría progresa a cáncer. Los adenomas que progresan a cáncer se denominan adenomas de alto riesgo.

5 Los marcadores de proteínas tienen un gran potencial para aplicarse en la detección de CRC con base en heces, porque pueden medirse en pequeños volúmenes de muestra con ensayos simples y relativamente baratos, de los cuales el FIT ampliamente utilizado es un excelente ejemplo (Bosch LJ, et al., Molecular tests for colorectal cancer screening. Clin Colorectal Cancer 2011; 10: 8-23; Young GP, Bosch LJW. Curr Colorectal Cancer Rep 2011; 7: 62-70; Oort FA, et al., Aliment Pharmacol Ther 2010; 31: 432-9).

10 El presente estudio ha identificado nuevos biomarcadores de proteínas mediante la aplicación de proteómica profunda a muestras de heces. De un total de 830 proteínas humanas detectadas, 134 se enriquecieron significativamente en muestras de heces de pacientes con CRC en comparación con las muestras de heces de control, de las cuales varias mostraron un mayor poder discriminatorio que la hemoglobina y/o complementariedad a la hemoglobina.

15 El enfoque de medir moléculas directamente en las heces es importante para revelar biomarcadores que son estables en el entorno fecal y detectables en el ambiente de moléculas bacterianas y relacionadas con alimentos.

20 La presente descripción se usa ventajosamente para la detección del cáncer colorrectal, es decir, el adenocarcinoma que se encuentra en el colon. Sin embargo, los métodos de la divulgación no deben considerarse como limitados únicamente a la detección de adenocarcinomas de colon. Más bien, los métodos de la invención también son útiles en la detección de adenomas colónicos avanzados o de alto riesgo, permitiendo así la identificación de un individuo en riesgo de desarrollar cáncer colorrectal debido a la presencia de un adenoma avanzado o de alto riesgo.

25 Las referencias en el presente documento para la detección del cáncer colorrectal pueden incluir la detección de adenomas colónicos avanzados y adenomas de alto riesgo, así como adenocarcinoma de colon.

También se espera que los biomarcadores identificados por la presente divulgación también puedan encontrar aplicación para el diagnóstico de adenocarcinomas presentes más arriba en el tracto gastrointestinal.

30 Por lo tanto, la presente divulgación también puede proporcionar un método para la detección de enfermedades gastrointestinales o cáncer gastrointestinal, comprendiendo el método: la detección de una muestra biológica, por ejemplo una muestra de heces, de un individuo para uno o más biomarcadores seleccionados del grupo definido en la Tabla 1, en la que la presencia o aumento de la expresión de uno o más biomarcadores en relación con una muestra de control es indicativo de que el paciente está en riesgo de sufrir o padece una enfermedad gastrointestinal o cáncer gastrointestinal.

#### Muestra para detección

40 La muestra para la detección puede incluir líneas celulares, biopsias, sangre completa, suero sanguíneo, esputo, heces, orina, líquido sinovial, líquido de heridas, líquido cefalorraquídeo, tejido de ojos, intestino, riñón, cerebro, piel, corazón, próstata, pulmón, mama, hígado, músculo o tejido conectivo, dicho tejido opcionalmente incrustado en parafina, portaobjetos de objetos histológicos y todas las combinaciones posibles de los mismos.

45 De acuerdo con la invención, la muestra biológica son heces. La muestra puede prepararse por cualquier método convencional para extraer proteínas de una muestra biológica. Un ejemplo del método se puede encontrar en Ang CS, Nice EC. J Proteome Res. 2010; 9: 4346-55.

#### Biomarcadores

50 La presente descripción proporciona un conjunto de biomarcadores que pueden detectarse directamente a partir de una muestra de heces y que han demostrado ser indicadores fiables de la presencia de adenomas o adenocarcinomas de colon avanzado en un individuo.

55 Los biomarcadores identificados se enumeran en la Tabla 1 y/o la Tabla 6. En una realización, los métodos de la invención detectan más de un biomarcador del grupo definido en la Tabla 1, por ejemplo dos, tres, cuatro, cinco, seis, siete, ocho, nueve, diez de los biomarcadores del grupo definido en la Tabla 1 y/o la Tabla 6. En una realización alternativa, los métodos de la invención detectan más de diez de los biomarcadores del grupo definido en la Tabla 1, por ejemplo, once, doce, trece, catorce, quince, dieciséis, diecisiete, dieciocho, diecinueve, veinte, treinta, cuarenta, cincuenta, sesenta, setenta, ochenta, noventa, cien de los biomarcadores del grupo definido en la Tabla 1 y/o la Tabla 6.

60 Por lo tanto, los métodos de la divulgación para detectar la presencia o aumento de la expresión de uno o más de: componente C4B del complemento (grupo sanguíneo Chido) 2 (C4A/C4B); glutámico-oxaloacético transaminasa 2, mitocondrial (aspartato aminotransferasa 2) (GOT2); glucosa-6-fosfato isomerasa (GPI); transcetolasa (TKT); N-acilaminoacil-péptido hidrolasa (APEH); grupo de histonas 1, H4c (HIST4H4 (incluye otros)); proteína5 de unión a ácidos grasos (asociada a psoriasis) (FABP5); hexosaminidasa B (polipéptido beta) (HEXB); molécula de adhesión a

células epiteliales (EPCAM); que recorre NME1-NME2 (NME1-NME2); superóxido dismutasa 2, mitocondrial (SOD2); factor de alargamiento de traducción Tu, mitocondrial (TUFM); glutatión sintetasa (GSS); anexina A2 (ANXA2); ATP sintasa, transporte de H<sup>+</sup>, complejo mitocondrial F1, polipéptido beta (ATP5B); proteína de choque térmico de 10 kDa (chaperonina 10) (HSPE1); glioxalasa I (GLO1); grupo de histonas 2, H2be (HIST2H2BE (incluye otros)); proteína S100 de unión a calcio A4 (S100A4); proteína S100 de unión a calcio A11 (S100A11); latexina (LXN); miembro 11 de deshidrogenasa/reductasa (familia SDR) (DHRS11); N-acetilglucosaminidasa, alfa (NAGLU); translina (TSN); subunidad del proteasoma (prosoma, macropaina) alfa tipo 4 (PSMA4); subunidad de proteasoma (prosoma, macropaina) alfa tipo-6 (PSMA6); sustrato 1 de toxina botulínica C3 relacionada con ras (familia rho, proteína pequeña Rac1 de unión a GTP) (RAC1); adenosilhomocisteinasa (AHCY); fucosidasa, alfa-L-1, tejido (FUCA1); proteína S100 de unión a calcio P (S100P); subunidad de proteasoma (prosoma, macropaina) beta tipo-2 (PSMB2); X-proil aminopeptidasa (aminopeptidasa P) 1 (XPNPEP1); queratina 18 (KRT18); subunidad 1 de la proteína de unión a la caperuza nuclear de 80 kDa (NCBP1); manosidasa, alfa, clase 2B, miembro 1 (MAN2B1); proteína S100 de unión a calcio A6 (S100A6); proteína que contiene valosina (VCP); quinolinato fosforibosiltransferasa (QPRT); complejo mayor de histocompatibilidad, clase I, B (HLA-B); fosfoglicerato mutasa 1 (cerebro) (PGAM1); ectonucleótido pirofosfatasa/ fosfodiesterasa 3 (ENPP3); inhibidor de la serpina peptidasa, clado B (ovoalbúmina), miembro 10 (SERPINB10); mieloperoxidasa (MPO); creatina quinasa, mitocondrial 1B (CKMT1A/CKMT1B); proteinasa 3 (PRTN3); elastasa, expresada en neutrófilos (ELANE); dedo de zinc 1 tipo CW de la familia MORC (MORC1); ubiquitina B (UBB); fosfolipasa A2, grupo IIA (plaquetas, líquido sinovial) (PLA2G2A); anhidrasa carbónica IV (CA4); factor de alargamiento G, mitocondrial 2 (GFM2); proteína S100 de unión a calcio A7 (S100A7); proteína que aumenta la permeabilidad bactericida (BPI); colágeno, tipo VI, alfa 5 (COL6A5); homeobox 8 de LIM (LHX8); proteína 3 secretora rica en cisteína (CRISP3); azurocidina (AZU1); hemicentina 1 (HMCN1); transglutaminasa 3 (polipéptido E, proteína-glutamina gamma-glutamyltransferasa) (TGM3); proteína quinasa alfa de unión a CDC42 (similar a DMPK) (CDC42BPA); catepsina G (CTSG); resistina (RETN); metilmalonil CoA mutasa (MUT); que contiene repetición armadillo, 4 ligado a X (ARMCX4); integrina alfa-M (subunidad del receptor 3 del componente 3 del complemento) (ITGAM); canal de calcio, dependiente del voltaje, subunidad alfa-1E tipo R (CACNA1E); invasión y metástasis 2 de linfoma de células T (TIAM2); homólogo A defectuoso en la regulación del ciclo celular de histona HIR (S. cerevisiae) (HIRA); miembro 2 de la familia dopey (DOPEY2); proteína 3 de unión a integrina beta 1 (ITGB1BP3); canal de sodio, controlado por voltaje, tipo VII, alfa (SCN7A); Rab3C, miembro de la familia de oncogenes RAS (RAB3C); marco de lectura abierta 79 del cromosoma 9 (C9orf79); factor nuclear de células T activadas, citoplasmático, dependiente de calcineurina 4 (NFATC4); UDP-glucosa glicoproteína glucosiltransferasa 2 (UGGT2); cornulina (CRNN); proteína similar a quelina/cordina (KCP); molécula CD1E (CD1E); dominio de hélice superenrollada que contiene 18 residuos (CCDC18); leucotrieno A-4 hidrolasa (LTA4H); albúmina (ALB); alfa-2-macroglobulina (A2M); componente 3 del complemento (C3); hemoglobina, beta (HBB); transferrina (TF); hemoglobina, alfa 1 (HBA1/HBA2); lactotransferrina (LTF); ceruloplasmina (ferroxidasa) (CP); catalasa (CAT); componente específico del grupo (proteína de unión a la vitamina D) (GC); inhibidor de la serpina peptidasa, clado C (antitrombina), miembro 1 (SERPINC1); cadena gamma de fibrinógeno (FGG); proteína S100 de unión a calcio A8 (S100A8); ferritina, polipéptido ligero (FTL); actina beta (ACTB); fibronectina 1 (FN1); defensina, alfa 1 (DEFA1 (incluye otros)); inhibidor de la serpina peptidasa, clado G (inhibidor de C1), miembro 1 (SERPING1); proteína 4 de unión a retinol, plasma (RBP4); peroxirredoxina 2 (PRDX2); cadena alfa de fibrinógeno (FGA); inhibidor de la serpina peptidasa, clado F (antiplasmina alfa-2, factor derivado del epitelio pigmentario), miembro 2 (SERPINF2); anhidrasa carbónica II (CA2); orosomucoide 1 (ORM1/ORM2); lactato deshidrogenasa A (LDHA); vitronectina (VTN); quinínogeno-1 (KNG1); actina, alfa, músculo cardíaco 1 (ACTC1); alfa-2-glicoproteína 1 rica en leucina (LRG1); gamma-glutamyl hidrolasa (conjugasa, folilpoligammaglutamil hidrolasa) (GGH); enolasa 1, (alfa) (ENO1); profilina 1 (PFN1); inhibidor de la serpina peptidasa, clado A (antiproteinasa alfa-1, antitripsina), miembro 7 (SERPINA7); alfa-1-microglobulina/precursor de bikunina (AMBP); lamina A/C (LMNA); apolipoproteína D (APOD); proteína de interacción con el receptor de la hormona tiroidea 11 (TRIP11); proteína de unión del componente 4 del complemento, alfa (C4BPA); tropomiosina 4 (TPM4); filamina A, alfa (FLNA); haptoglobina (HP); hemopexina (HPX); hemoglobina, delta (HBD); cadena beta de fibrinógeno (FGB); proteína S100 de unión a calcio A9 (S100A9); componente 5 del complemento (C5); familia 26 portadora de soluto, miembro 3 (SLC26A3); componente 9 del complemento (C9); componente amiloide P, suero (APCS); glicoproteína alfa-1-B (A1BG); complemento tipo C3 (LOC100133511); inhibidor H4 inter-alfa (globulina) (glicoproteína sensible a caliceína en plasma) (ITI4); componente C8 del complemento, polipéptido alfa (C8A); inhibidor H1 inter-alfa (globulina) (ITI1); acil-CoA deshidrogenasa, cadena muy larga (ACADVL); ADNc FLJ60317, muy similar a aminoacilasa-1 (ACY1); proteína de 35 residuos que contiene un dominio de repetición de anquirina (ANKRD35); repetición de IAP baculoviral que contiene 6 residuos (BIRC6); bleomicina hidrolasa (BLMH); antígeno 12 de células estromales de médula ósea (BST1); proteína hipotética LOC643677 (C13orf40); citidina desaminasa (CDA); quitinasa 1 (quitotriosidasa) (CHIT1); catepsina C (CTSC); catepsina S (CTSS); isoforma 2 del dedicado de la proteína 4 de citocinesis (DOCK4); glutatión reductasa (GSR); dominio hect (homólogo al terminal carboxilo E6-AP (UBE3A)) y el dominio 1 similar a RCC1 (CHC1) (RLD) (HERC1); dominio hect y RLD 2 (HERC2); complejo mayor de histocompatibilidad, clase II, DR beta 5 (HLA-DRB5); isocitrato deshidrogenasa 1 (NADP +), soluble (IDH1); inhibidor inter-alfa (globulina) H2 (ITI2); proteína no caracterizada KIAA1797 (KIAA1797); lisozima C (LYZ); nebulina (NEB); quinasa 10 relacionada con NIMA (nunca en el gen a de mitosis) (NEK10); peptidasa D (PEPD); quiescina Q6 sulfhidril oxidasa 1 (QSOX1); ribonucleasa T2 (RNASET2); inhibidor de la serpina peptidasa, clado A (antiproteinasa alfa-1, antitripsina), miembro 1 (SERPINA1); inhibidor de la serpina peptidasa, clado A (antiproteinasa alfa-1, antitripsina), miembro 3 (SERPINA3); inhibidor de la serpina peptidasa, clado B (ovoalbúmina), miembro 3 (SERPINB3); dominio SET que contiene 2 residuos (SETD2); similar a shugoshina 2 (SGOL2); ácido siálico acetiltransferasa (SIAE); envoltura nuclear 1, que contiene repetición de espectrina (SYNE1); transaldolasa 1 (TALDO1); receptor del gusto tipo 2 miembro 42 (TAS2R42); trifosfato isomerasa 1

(TPI1); vinculina (VCL); proteína 16 de membrana de gránulos de zimógeno (ZG16); proteína hipotética LOC79887 (PLBD1); isoforma 1 de la subunidad A de repetición de anquirina reguladora de la proteína serina/treonina fosfatasa 6 (ANKRD28); cistatina-C (CST3); D-dopacromo descarboxilasa (DDT); proteína 1 asociada a sinapsis (SYAP1); subunidad alfa tipo 2 del proteasoma (PSMA2); homólogo de SUB1 (*S. cerevisiae*) (SUB1); glicoproteína 3 asociada a microfibrillas (MFAP3); catepsina D (CTSD); subunidad de proteasoma (prosoma, macropaina), tipo beta, 1 (PSMB1); subunidad de proteasoma (prosoma, macropaina), tipo beta, 5 (PSMB5); ADNc FLJ61112, muy similar a la proteína KCTD15 que contiene el dominio BTB/POZ (KCTD15); prolil-4-hidroxilasa, polipéptido beta (P4HB); glutatión peroxidasa 1 (GPX1); inhibidor de la serpina peptidasa, clado B (ovoalbúmina), miembro 5 (SERPINB5); isoforma 1 de colágeno, tipo IV, proteína de unión alfa 3 (antígeno de Goodpasture) (COL4A3BP); subunidad de proteasoma (prosoma, macropaina), tipo beta, 6 (PSMB6); queratina 20 (KRT20); subunidad pequeña 1 de calpaina (CAPNS1); peroxirredoxina 3 (PRDX3); miembro 2 de la familia NACC, que contiene el dominio BEN y BTB (POZ) (NACC2); inhibidor 2 de disociación de Rho GDP (ARHGDI2); factor inhibidor de migración de macrófagos (MIF); proteína 6 de unión a Ran (RANBP6); homólogo 3 de spinster (*Drosophila*) (SPNS3); componente 2 del complejo de mantenimiento de minicromosomas (MCM2); fumarilacetoacetasa (FAH); proteína 8 de choque térmico de 70 kDa (HSPA8); proteína 1 señal, unida a membrana, abundante en el cerebro (BASP1); Aminotransferasa de aminoácidos de cadena ramificada (BCAT2); moesina (MSN); inhibidor de la serpina peptidasa, clado B (ovoalbúmina), miembro 8 (SERPINB8); glucosa-6-fosfato deshidrogenasa (G6PD); isoforma 1 de la proteína UPF0557 C10orf119 (C10orf119); prosaposina (PSAP); factor 1 gamma de alargamiento de traducción eucariota (EEF1G); cuatro y medio dominios 1 de LIM (FHL1); carboxipeptidasa, tipo vitelogénica (CPVL); familia tipo tirosina ligasa de tubulina, miembro 3 (TTLL3); IPI: IPI00942608.1 | ENSEMBL: ENSP0000 (no mapeada; proteína de 26 kDa); subfamilia 2 de la proteína BstNI rica en prolina (PRB1/PRB2); protocadherina 8 (PCDH8); proteína 1 tipo alfa-2-macroglobulina (A2ML1); guanina desaminasa (GDA); lipocalina 1 (LCN1); histona H1.4 (HIST1H1E); IPI: IPI00937064.1 | REFSEQ: XP\_002342720 (ZAN); ribonucleoproteína nuclear heterogénea A2/B1 (HNRNPA2B1); aminopeptidasa 2 del retículo endoplasmático (ERAP2); 14-3-3 proteína zeta/delta (YWHAZ); receptor 39 acoplado a proteína G (GPR39); similar a proteína KIAA1783 (proteína KIAA1783); proteína de unión a apolipoproteína A-I (APOA1BP); dominio de pleckstrina y Sec7 que contiene 2 residuos (PSD2); prolilcarboxipeptidasa (angiotensinasa C) (PRCP); cadena alfa-1C de tubulina (TUBA1C); proteína 5 similar a calmodulina (CALML5); homólogo de proteína 3 relacionada con actina ARP3 (levadura) (ACTR3); miosina, cadena ligera 6, álcali, músculo liso y no muscular (MYL6); fosfoproteína estimulada por vasodilatador (VASP); homólogo de proteína 2 relacionada con actina ARP2 (levadura) (ACTR2); factor reumatoide (RF-IP18); fosfoglicerato quinasa 1 (PGK1); miembro F1 de la familia 35 portadora de soluto (SLC35F1); miembro F1 de la familia 35 portadora de soluto (SLC35F1); fosfatasa alcalina, hígado/hueso/riñón (ALPL); I tropomiosina 3 (TPM3); hexoquinasa 3 (HK3); vimentina (VIM); anexina A1 (ANXA1); IPI: IPI00930073.1 | TREMBL: B2R853 (KRT6C); queratina, citoesquelética 6C tipo II (KRT6C); miosina, cadena pesada 13, músculo esquelético (MYH13); progresión del ciclo celular 1 (CCPG1); proteína hipotética (H-INV); canal de calcio, dependiente del voltaje, tipo L, subunidad alfa 1D (CACNA1D); dominio LY6/PLAUR que contiene 5 residuos (LYPD5); dominio aarF que contiene quinasa 2 (ADCK2); miosina-1c (MYO1C); proteína 2 de unión a la proteína precursora de amiloide beta (cola citoplasmática) (APPBP2); integrina alfa 2b (glicoproteína plaquetaria IIb del complejo IIb/IIIa, antígeno CD41) (ITGA2B); tubulina, beta 6 (TUBB6); tipo sinaptotagmina 4 (SYTL4); aquaporina 4 (AQP4); ciclo 42 de división celular (proteína de unión a GTP, 25 kDa) (CDC42); miosina, cadena ligera 12B, reguladora (MYL12B); proteína tipo L-Myc-2 (LOC100293553); RAP1B, miembro de la familia de oncogenes RAS (RAP1B); glicoproteína IX (plaquetas) (GP9); dextrina (DSTN); componente 1 del complemento, subcomponente q, cadena C (C1QC); sustrato 8 de la vía del receptor del factor de crecimiento epidérmico (EPS8); fosfatasa 3 de doble especificidad (DUSP3); familia de genes homólogos de ras, miembro A (RHOA); miosina, cadena ligera 9, reguladora (MYL9); peptidilprolil isomerasa A (ciclofilina A) (PPIA); cofilina 1 (CFL1); y/o lactotransferrina (LTF), colágeno, tipo XII, alfa 1 (COL12A1), agrina (AGRN), proteína de unión a +MYB (P160) 1a (MYBBP1A), proteína asociada al dominio de transformación/transcripción (TRRAP), anexina A6 (ANXA6), proteína 5 asociada al citoesqueleto (CKAP5), componente 5 del complejo de mantenimiento de minicromosomas (MCM5), importina 4 (IPO4), tipo neurobeachina 2 (NBEAL2), componente 4 del complejo de mantenimiento de minicromosomas (MCM4), 2'-5'-oligoadenilato sintetasa 3, 100 kDa (OAS3), componente 3 del complejo de mantenimiento de minicromosomas (MCM3), subunidad 1 de la enzima E1 activadora de NEDD8 (NAE1), motivo tripartita que contiene 28 residuos (TRIM28), fusionado en sarcoma (FUS), fenilalanil-tARN sintetasa, subunidad alfa (FARSA), antígeno de diferenciación nuclear de células mieloides (MNDA), supresor del homólogo Ty 16 (*S. cerevisiae*) (SUPT16H), polipéptido 5 de la caja DEAD (Asp-Glu-Ala-Asp) (DDX5), tenascina C (TNC), homólogo 7 de importación nuclear (*S. cerevisiae*) (NIP7), proteína 4 de unión a ADN helicasa del cromodominio (CHD4), regulador de la condensación cromosómica 2 (RCC2), ADN (citosina-5-)-metiltransferasa 1 (DNMT1), exportina 4 (XPO4), TCP1 que contiene chaperonina, subunidad 5 (épsilon) (CCT5), factor 9 de empalme rico en serina/arginina (SRSF9), espectrina, beta, no eritrocítica 2 (SPTBN2), inhibidor 1 de la metalopeptidasa TIMP (TIMP1), nidógeno 1 (NID1), ribonucleótido reductasa M1 (RRM1), factor 4 gamma de iniciación de la traducción eucariota, 1 (EIF4G1), componente del complejo 4 oligomérico de Golgi (COG4), polimerasa (dirigida al ADN), delta 1, subunidad catalítica de 125 kDa (POLD1), factor 3b de empalme, subunidad 2, 145 kDa (SF3B2), componente 2 exosómico (EXOSC2), componente 6 del complejo de mantenimiento de minicromosomas (MCM6), plastina 3 (PLS3), aldolasa B, fructosa-bisfosfato (ALDOB), homólogo de SMG1, fosfatidilinositol 3-quinasa relacionada con la quinasa (*C. elegans*) (SMG1), transición 1 de fase G1 a S (GSPT1), proteína reguladora de empalme tipo KH (KHSRP), polipéptido 21 de la caja DEAD (Asp-Glu-Ala-Asp) (DDX21), proteína de transferencia de fosfatidilinositol, beta (PITPNB), homólogo de acuario (ratón) (AQR), ribonucleoproteína nuclear heterogénea tipo D (HNRPDL), anexina A3 (ANXA3), procesamiento del precursor 1, subunidad de ribonucleasa P/MRP (*S. cerevisiae*) (POP1), mantenimiento estructural de los cromosomas 2 (SMC2), dineína, citoplasmática 1, cadena intermedia ligera 2 (DYNC1LI2),

peptidilprolil isomerasa D (PPID), clasificación de proteínas vacuolares 37 homólogo B (*S. cerevisiae*) (VPS37B), adrenérgico, beta, receptor quinasa 1 (ADRBK1), homólogo de control mitótico DIS3 (*S. cerevisiae*) (DIS3), polipéptido A de polimerasa I (ARN), 194 kDa (POLR1A), complejo t 1 (TCP1), placofilina 3 (PKP3), familia del dominio de ribonucleoproteína La, miembro 1B (LARP1B), poli (ADP-ribosa) polimerasa 1 (PARP1), molécula CD46, proteína reguladora del complemento (CD46), quinasa 2 activada por proteína p21 (Cdc42/Rac) (PAK2), casete de unión a ATP, subfamilia E (OABP), miembro 1 (ABCE1), peptidasa 14 específica de ubiquitina (tARN-guanina transglucosilasa) (USP14), TCP1 que contiene chaperonina, subunidad 3 (gamma) (CCT3), proteína 1 activadora de Ran GTPasa (RANGAP1), desoxitimidilato quinasa (timidilato quinasa) (DTYMK), N-miristoiltransferasa 1 (NMT1), tipo dinamina 1 (DNM1L), proteína 2 transmembrana inducida por interferón (1-8D) (IFITM2), miembro 1 de la familia de fermitina (FERMT1), cofactor D de plegado de tubulina (TBCD), factor 10 de empalme rico en serina/arginina (LOC100505793/SRSF10), quinasa tipo STE20 (SLK), mucina 5AC, moco oligomérico/formador de gel (MUC5AC/MUC5B), metionil-tARN sintetasa (MARS), homólogo SMEK 1, supresor de mek1 (*Dictyostelium*) (SMEK1), caja 2 del grupo de alta movilidad (HMGB2), dominio que no contiene POU, de unión a octámero (NONO), factor de crecimiento transformante, inducido por beta, 68 kDa (TGFB1), fibulina 2 (FBLN2), proteína de unión a lipoproteína de alta densidad (HDLBP), colágeno, tipo IV, alfa 2 (COL4A2), copina 1 (CPNE1), N(alfa)-acetiltransferasa 50, subunidad catalítica NatE (NAA50), homólogo de LSM7, pequeño ARN nuclear U6 asociado (*S. cerevisiae*) (LSM7), proteína 1 de reconocimiento específica de la estructura (SSRP1), importina 8 (IPO8), yippee tipo 5 (*Drosophila*) (YPEL5), fosfoglucomutasa 3 (PGM3), proteína 40 de dedo anular (RNF40), mantenimiento estructural de cromosomas 3 (SMC3), familia derivada de islotes de regeneración, miembro 4 (REG4), factor 3a de empalme, subunidad 3, 60 kDa (SF3A3), trombospondina 1 (THBS1), TCP1 que contiene chaperonina, subunidad 6A (zeta 1) (CCT6A), homólogo del factor 8 de procesamiento previo al ARNm PRP8 (*S. cerevisiae*) (PRPF8), simplequina (SYMPK), proteína 1 de unión al elemento alejado secuencia arriba (FUSE) (FUBP1), factor 1 auxiliar de ARN nuclear pequeño U2 (U2AF1), huntingtina (HTT), factor 5B de iniciación de la traducción eucariota (EIF5B), proteína de esperma autoantigénica nuclear (unión a histona) (NASP), ribonucleoproteína K nuclear heterogénea (HNRNPK), proteína 1 de unión a la caja Y (YBX1), anexina A11 (ANXA11), tipo proteína RecQ (ADN helicasa tipo Q1) (RECQL), cortactina (CTTN), tubulina, beta 3 (TUBB3), tipo pirrolina-5-carboxilato reductasa (PYCRL), periplaquina (PPL), fosfoglucomutasa 1 tipo 2 (PGM2L1), marco 49 de lectura abierta del cromosoma 17 (C17orf49), homólogo 4 de cambio de ARNm (*S. cerevisiae*) (MRTO4), metiltransferasa tipo 1 (METTL1), antígeno de carcinoma de células escamosas reconocido por las células T 3 (SART3), proteína S100 de unión a calcio A13 (S100A13), tipo aminopeptidasa 1 (NPEPL1), quinasa 1 dependiente de ciclina (CDK1), componente n-recognina 1 de proteína ligasa ubiquitina E3 (UBR1), proteína 18 activadora de Rho GTPasa (ARHGAP18), partícula de reconocimiento de señal de 14 kDa (proteína de unión al ARN homóloga de Alu) (SRP14), péptido antimicrobiano catelicidina (CAMP), factor de empalme rico en prolina/glutamina (SFPQ), activador de proteína RAS p21 (proteína activadora de GTPasa) 1 (RASA1), proteína activadora de Ral GTPasa, subunidad beta (no catalítica) (RALGAPB), laminina, beta 1 (LAMB1), subunidad 2 de proteína activadora de GTPasa RAB3 (no catalítica) (RAB3GAP2), TCP1 que contiene chaperonina, subunidad 8 (theta) (CCT8), ribonucleoproteína nuclear heterogénea tipo L (HNRPLL), proteína 1 de unión a RAN (RANBP1), cinetocoro asociado 1 (KNTC1), disqueratosis congénita 1, disquerina (DKC1), caseína quinasa 2, polipéptido alfa 1 (CSNK2A1), dominio CAP-GLY que contiene la proteína enlazadora 1 (CLIP1), TCP1 que contiene chaperonina, subunidad 2 (beta) (CCT2), familia tipo tirosina ligasa de tubulina, miembro 12 (TTLL12), ataxia telangiectasia mutada (ATM), factor de empalme 3a, subunidad 1, 120 kDa (SF3A1), proteína S20 ribosómica (RPS20), enzima E2O de conjugación de ubiquitina (UBE2O), región promotora translocada (para el oncogén MET activado) (TPR), BRCA2 y proteína de interacción con CDKN1A (BCCIP), proteína 5 asociada a gem (órgano nuclear) (GEMIN5), subunidad ribonucleasa P/MRP de 30 kDa (RPP30), pérdida de heterocigosidad, 12, región 1 cromosómica (LOH12CR1), proteína 2 de unión a sintaxina (STXBP2), enzima E2H que se conjuga con ubiquitina (UBE2H), homólogo B de la proteína 2 que interactúa con disco DIP2 (*Drosophila*) (DIP2B), RAP1, estimulador 1 de disociación de GTP-GDP (RAP1GDS1), ribonucleoproteína M nuclear heterogénea (HNRNPM), dominio 7 de LIM (LMO7), proteína 25 del motivo de unión a ARN (RBM25), familia de la aldehído deshidrogenasa 7, miembro A1 (ALDH7A1), factor 1 específico de escisión y poliadenilación, 160 kDa (CPSF1), calponina 2 (CNN2), TCP1 que contiene chaperonina, subunidad 7 (eta) (CCT7), lisil-tARN sintetasa (KARS), UDP-N-acteilglucosamina pirofosforilasa 1 (UAP1), proteína de choque térmico de 70 kDa tipo 4 (HSPA4L), proteína de 138 kDa (proteína de 138 kDa), thimet oligopeptidasa 1 (THOP1), glutarredoxina 3 (GLRX3), fosfoglicerato deshidrogenasa (PHGDH), homólogo de CDV3 (ratón) (CDV3), mantenimiento estructural de cromosomas 4 (SMC4), proteína 3 del motivo de unión a ARN (RNP1, RRM) (RBM3), factor de crecimiento derivado de hepatoma (HDGF), ribonucleoproteína nuclear heterogénea U (factor A de unión al andamio) (HNRNPU), proteína 1 de unión al receptor nuclear (NRBP1), polipéptido B, de polimerasa (ARN) I, de 128 kDa (POLR1B), proteína fosfatasa 5, subunidad catalítica (PPP5C), glucosa-6-fosfato deshidrogenasa (G6PD), arginasa, hígado (ARG1), 3-hidroxi-3-metilglutaril-CoA sintasa 1 (soluble) (HMGCS1), enzima 2 activadora del modificador similar a la ubiquitina (UBA2), KIAA1033 (KIAA1033), anexina A4 (ANXA4), polipéptido 17 de la caja DEAD (Asp-Glu-Ala-Asp) (DDX17), familia de fosfoproteína 32 nuclear ácida (rica en leucina), miembro E (ANP32E), glucosamina (UDP-N-acetil)-2-epimerasa/N-acetilmannosamina quinasa (GNE), KIAA0368 (KIAA0368), homólogo B de clasificación de proteínas vacuolar 4 (*S. cerevisiae*) (VPS4B), proteína de replicación A1, 70 kDa (RPA1), factor 2 de iniciación de la traducción eucariota, subunidad 1 alfa, 35 kDa (EIF2S1), factor 3 de iniciación de la traducción eucariota, subunidad J (EIF3J), supresor del homólogo Ty 6 (*S. cerevisiae*) (SUPT6H), proteína 1 de choque térmico de 105 kDa/110 kDa (HSPH1), exportina 5 (XPO5), factor A de alargamiento de la transcripción (SII), 1 (TCEA1), antígeno B del síndrome de Sjogren (autoantígeno La) (SSB), proteína 1 de unión a AE (AEBP1), LIM y dominios 1 ricos en cisteína (LMCD1), factor 3 de unión al potenciador de interleuquina, 90 kDa (ILF3), dominio 61 de repetición de WD (WDR61), N(alfa)-acetiltransferasa 15, subunidad auxiliar NatA (NAA15), factor 4 de empalme rico en serina/arginina (SRSF4), proteína

20 de dedo anular (RNF20), lactamasa, beta 2 (LACTB2), homólogo de ribonucleoproteína NHP2 (levadura) (NHP2), marco de lectura abierto 28 del cromosoma 17 (C17orf28), CTP sintasa II (CTPS2), homólogo 1 de fascina, proteína de agrupamiento de actina (*Strongylocentrotus purpuratus*) (FSCN1), tARN nucleotidil transferasa, adición de CCA, 1 (TRNT1), proteína 1 rica en glutamina/lisina reguladora de empalme (SREK1), antígeno 1 estromal (STAG1), proteína de unión a oxisterol (OSBP), desoxiuridina trifosfatasa (DUT), dominio de hélice superenrollada que contiene 25 residuos (CCDC25), oncogén DEK (DEK), dominio de hélice superenrollada que contiene 72 residuos (CCDC72), polipéptido E de polimerasa (ARN) II (dirigido por ADN), 25 kDa (POLR2E), fosfoserina fosfatasa (PSPH), mantenimiento estructural de los cromosomas 1A (SMC1A), polipéptido 23 de la caja DEAD (Asp-Glu-Ala-Asp) (DDX23), homólogo de tARN metiltransferasa 11-2 (*S. cerevisiae*) (TRMT112), subunidad 2 homóloga fotomorfogénica constitutiva de COP9 (*Arabidopsis*) (COPS2), muerte celular programada 5 (PDCD5), quinasa 2 dependiente de ciclina (CDK2), subunidad 26S de proteasoma (prosoma, macropaina), sin ATPasa, 3 (PSMD3), proteína 2 de unión a RAN (RANBP2), proteína 1 de unión a ARNm SERPINE1 (SERBP1), N-acetilglucosamina transferasa (GlcNAc) enlazada a O (UDP-N-acetilglucosamina: polipéptido-N-acetilglucosaminil transferasa) (OGT), complejo sin SMC condensina I, subunidad D2 (NCAPD2), regulador de cromatina dependiente de actina, relacionado con SWI/SNF, asociado a la matriz, subfamilia c, miembro 2 (SMARCC2), homólogo de ribonucleoproteína NOP10 (levadura) (NOP10), inhibidor 5 de apoptosis (API5), tráfico de vesículas sensibles a LPS, que contiene beige y ancla (LRBA), receptor de partículas de reconocimiento de señal (proteína de acoplamiento) (SRPR), receptor de transferrina (p90, CD71) (TFRC), cremallera básica de leucina y dominios 2 W2 (BZW2), ribonucleasa, familia 3 de RNasa A (RNASE3), inhibidor de unión a diazepam (modulador del receptor GABA, proteína de unión a acil-CoA) (DBI), proteína 4 de unión a FK506, 59 kDa (FKBP4), marco 130 de lectura abierto del cromosoma 6 (C6orf130), cofactor de BRCA1 (COBRA1), endonucleasa 1 específica de la estructura de aleta (FEN1), glucano (1,4-alfa-), enzima ramificadora 1 (GBE1), polipéptido B de ribonucleoproteína nuclear pequeña (SNRPB2), cofactor NSFL1 (p97) (p47) (NSFL1C), acil-CoA tioesterasa 7 (ACOT7), familia del dominio NOP2/Sun, miembro 2 (NSUN2), TCP1 que contiene chaperonina, subunidad 4 (delta) (CCT4), peptidasa 6 relacionada con calicreína (KLK6), glutaminil-péptido ciclotransferasa (QPCT), atanógeno 6 asociado a BCL2 (BAG6), factor 3 de iniciación de la traducción eucariota, subunidad C (EIF3C/EIF3CL), ATPasa, que transporta H<sup>+</sup>, lisosomal de 56/58 kDa, V1 subunidad B2 (ATP6V1B2), matriz de metalopeptidasa 8 (colagenasa de neutrófilos) (MMP8), subunidad 26S del proteasoma (prosoma, macropaina), ATPasa, 5 (PSMC5), regulador de retroalimentación de GTP ciclohidrolasa I (GCHFR), poli(A) polimerasa alfa (PAPOLA), tipo hipocalcina 1 (HPCAL1), proteína 1 de unión a la proteína (dominio SH3) activadora de GTPasa (G3BP1), polipéptido A de polimerasa (ARN) III (dirigido por ADN), 155 kDa (POLR3A), actividad 2 viralicídica superasesina tipo 2 (*S. cerevisiae*) (SKIV2L2), polipéptido A de polimerasa (ARN) II (dirigido por ADN), 220 kDa (POLR2A), colágeno, tipo I, alfa 2 (COL1A2), fibrilarina (FBL), glutamil-prolil-tARN sintetasa (EPRS), ELAV (embrionaria letal, visión anormal, *Drosophila*) tipo 1 (antígeno Hu R) (ELAVL1), subunidad 2 de proteína de unión a la caperuza nuclear, 20 kDa (NCBP2), control general de GCN1 de la síntesis 1 de aminoácidos tipo 1 (levadura) (GCN1L1), histona acetiltransferasa 1 (HAT1), antígeno estromal 2 (STAG2), sorbitol deshidrogenasa (SORD), REX2, homólogo de ARN exonucleasa 2 (*S. cerevisiae*) (REXO2), ribonucleoproteína F nuclear heterogénea (HNRNPF), timopoyetina (TMPO) peptidasa 24 específica de ubiquitina (USP24), KIAA1967 (KIAA1967), componente 1 del complemento, subcomponente r (C1R), anexina A7 (ANXA7), tipo RuvB 2 (*E. coli*) (RUVBL2), acireductona dioxigenasa 1 (ADI1), factor 4A3 de iniciación de la traducción eucariota (EIF4A3), ribonucleoproteína U nuclear heterogénea tipo 2 (HNRNPUL2), n-recognina 4 componente de la proteína ligasa E3 ubiquitina (UBR4), regulador de cromatina dependiente de actina, asociado a la matriz, relacionado a SWI/SNF, subfamilia a, miembro 2 (SMARCA2), citocromo b5 reductasa 2 (CYB5R2), factor 3b de empalme, subunidad 3, 130 kDa (SF3B3), supresor 1 de la ruta de la proteína G (GPS1), detención mitótica deficiente MAD2 tipo 1 (levadura) (MAD2L1), fosfolipasa C, gamma 1 (PLCG1), factor 4H de iniciación de la traducción eucariota (EIF4H), factor 2 auxiliar de ARN nuclear pequeño U2 (U2AF2), proteína 1 de unión a tanquirasa 1, 182 kDa (TNKS1BP1), transglutaminasa 2 (polipéptido C, proteína-glutamina-gamma-glutamiltransferasa) (TGM2), ribonucleoproteína L nuclear heterogénea (HNRNPL), inositol polifosfato-5-fosfatasa, 145 kDa (INPP5D), anexina A10 (ANXA10) homólogo de BUD31 (*S. cerevisiae*) (BUD31), proteína de transferencia de fosfatidilinositol, alfa (PITPNA), leucil-tARN sintetasa (LARS), nicotinamida N-metiltransferasa (NNMT), subunidad 26S de proteasoma (prosoma, macropaina), no ATPasa, 12 (PSMD12), homólogo del Oncogene CT10 del virus de sarcoma de v-crk (aviar) (CRK), proteoglicano 2, médula ósea (activador de células asesinas naturales, proteína básica principal del gránulo de eosinófilos) (PRG2), versicano (VCAN), exportina, tARN (receptor de exportación nuclear para los tARN) (XPOT), homólogo de la proteína nucleolar EMG1 (*S. cerevisiae*) (EMG1), marco 73 de lectura abierto del cromosoma 11 (C11orf73), transportina 1 (TNPO1), proteína 2 de unión beta del factor de crecimiento transformante latente (LTBP2), dominio de choque frío que contiene E1, unión a ARN (CSDE1), sulfirredoxina 1 (SRXN1), componente 1 de paraspeckle (PSPC1), proteína S3A ribosómica (RPS3A), modificador tipo ubiquitina ISG15 (ISG15), polipéptido B de polimerasa (ARN) II (dirigido por ADN), 140 kDa (POLR2B), factor de transcripción general III (GTF2I), proteína 2 de cromosoma sin histona NHP2 tipo 1 (*S. cerevisiae*) (NHP2L1), subunidad 26S del proteasoma (prosoma, macropaina), sin ATPasa, 10 (PSMD10), transductor de señal y activador de la transcripción 1, 91 kDa (STAT1), dominio de unión de GTP del factor de alargamiento Tu que contiene 1 residuo (EFTUD1), subunidad 23 del complejo mediador (MED23), factor 2C de iniciación de la traducción eucariota, 2 (EIF2C2), proteína 4B del motivo de unión de ARN (RBM4B), KIAA0664 (KIAA0664) factor de unión al núcleo, subunidad beta (CBFB), proteína de unión a poli(A), citoplasmática 1 (PABPC1), nicotinamida fosforibosiltransferasa (NAMPT), proteína 2 de unión a ácido retinoico celular (CRABP2), interactuador 12 del receptor de la hormona tiroidea (TRIP12), homólogo de DnaJ (Hsp40), subfamilia C, miembro 9 (DNAJC9), dominio de transferencia de lípidos relacionado con StAR

(START) que contiene 10 residuos (STARD10), proteína 213 de dedo anular (RNF213), factor 2B de iniciación de la traducción eucariota, subunidad 5 épsilon, 82 kDa (EIF2B5), homólogo 2 de bolA (*E. coli*) (BOLA2/BOLA2B), antígeno 5 expresado en meningioma (hialuronidasa) (MGEA5), Rab geranilgeraniltransferasa, subunidad alfa (RABGGTA), dominio de descarboxilasa dependiente de piridoxal que contiene 1 residuo (PDXDC1), componente 8 del exoma (EXOSC8), fosfoserina aminotransferasa 1 (PSAT1), factor 6 de iniciación de la traducción eucariota (EIF6), marco 13 de lectura abierto del cromosoma 16 (C16orf13), transductor de señal y activador de la transcripción 3 (factor de respuesta de fase aguda) (STAT3), proteína 1 de matriz extracelular similar a fibulina que contiene EGF (EFEMP1), defensina, alfa 6, específica de la célula de Paneth (DEFA6), pirofosfatasa (inorgánica) 1 (PPA1), glutatión peroxidasa 2 (gastrointestinal) (GPX2), homólogo D de unc-13 (*C. elegans*) (UNC13D), proteína tirosina fosfatasa, sin receptor tipo 6 (PTPN6), miosina XVIIIa (MYO18A), fibulina 1 (FBLN1), proteína ribosómica L19 (RPL19), homólogo diáfano 1 (*Drosophila*) (DIAPH1), homólogo de reparación por escisión del nucleótido MMS19 (*S. cerevisiae*) (MMS19), motivo 21 tipo nudix (fracción X ligada a nucleósido difosfato) (NUDT21), factor 3b de empalme, subunidad 5, 10 kDa (SF3B5), ribonucleoproteína que contiene el dominio SAP (SARNP), factor 2 de intercambio del nucleótido guanina del factor de ribosilación de ADP (inhibido por brefeldina A) (ARFGEF2), proteína 1 de unión a guanilato, inducible por interferón (GBP1), dominio HECT, UBA y WWE que contiene 1 residuo (HUWE1), procesamiento del precursor 7, subunidad de ribonucleasa P/MRP (*S. cerevisiae*) (POP7), factor de iniciación de la traducción eucariota 2B, subunidad 3 gamma, 58 kDa (EIF2B3), N(alfa)-acetiltransferasa 10, subunidad catalítica NatA (NAA10), homólogo de DnaJ (Hsp40), subfamilia B, miembro 1 (DNAJB1), alfa quinasa 2 del factor 2 de iniciación de la traducción eucariota (EIF2AK2), dominio de unión a GTP del factor Tu de alargamiento que contiene 2 residuos (EFTUD2), grancalcina, proteína de unión a calcio con mano EF (GCA), molécula de adhesión celular neuronal (NRCAM), polipéptido 16 de caja DEAH (Asp-Glu-Ala-His) (DHX16), que contiene la pequeña repetición del tetra-tricopéptido rico en glutamina (TPR), alfa (SGTA), serina/treonina quinasa 10 (STK10), supresor de smu-1 del homólogo de mec-8 y unc-52 (*C. elegans*) (SMU1), homólogo del factor 4 de procesamiento de pre-ARNm PRP4 (levadura) (PRPF4), ribonucleoproteína nuclear heterogénea D (proteína 1 de unión a ARN del elemento rico en AU, 37 kDa) (HNRNPD), sintasa CTP (CTPS), factor de estimulación de escisión, 3' pre-ARN, subunidad 3, 77 kDa (CSTF3), factor de estimulación de escisión, 3' pre-ARN, subunidad 1, 50 kDa (CSTF1), ribonucleoproteína nuclear heterogénea 1 tipo U (HNRNPUL1), proteína ribosómica S5 (RPS5), proteína tirosina fosfatasa, tipo 11 sin receptor (PTPN11), ladinina 1 (LAD1), componente del complejo de Golgi oligomérico 2 (COG2), culina 2 (CUL2), proteína ribosómica S17 (RPS17/RPS17L), subunidad 26S de proteasoma (prosome, macropaína), sin ATPasa, 4 (PSMD4), anexina A1 (ANXA1), anexina A2 (ANXA2), ribonucleótido formiltransferasa 5-aminoimidazol-4-carboxamida/IMP ciclohidrolasa (ATIC), azurocidina 1 (AZU1), repetición de IAP baculoviral que contiene 6 residuos (BIRC6), quitinasa 1 tipo 3 (glicoproteína 39 de cartílago) (CHI3L1), carbamoil-fosfato sintasa 1, mitocondrial (CPS1), catepsina G (CTSG), defensina, alfa 1 (DEFA1 (incluye otros)), eliminada en tumores malignos de cerebro 1 (DMBT1), elastasa, expresada en neutrófilos (ELANE), integrina, alfa M (subunidad del receptor 3 del componente 3 del complemento) (ITGAM), lipocalina 2 (LCN2), lectina, proteína de unión a 3, soluble, que se une a galactósido (LGALS3BP), componente 2 del complejo de mantenimiento de minicromosomas (MCM2), metalopeptidasa 9 de matriz (gelatinasa B, gelatinasa de 92 kDa, colagenasa tipo IV de 92 kDa) (MMP9), mieloperoxidasa (MPO), mucina 5AC, moco oligomérico/formador de gel (MUC5AC/MUC5B), subunidad 1 de proteína de unión a la caperuza nuclear, 80 kDa (NCBP1), neurofibromina 1 (NF1), olfactomedina 4 (OLFM4), PDS5, regulador del mantenimiento de la cohesión, homólogo A (*S. cerevisiae*) (PDS5A), proteína 1 de reconocimiento de peptidoglucano (PGLYRP1), proteinasa 3 (PRTN3), quiescina Q6 sulfhidril oxidasa 1 (QSOX1), regeneración 1 alfa derivada de islotes (REG1A), proteína S100 de unión al calcio A9 (S100A9), inhibidor de la serpina peptidasa, clado B (ovoalbúmina), miembro 10 (SERPINB10), inhibidor de la serpina peptidasa, clado B (ovoalbúmina), miembro 5 (SERPINB5), homólogo A de unc-45 (*C. elegans*) (UNC45A).

45 En una realización, los métodos de la divulgación detectan uno o más biomarcadores, cuya presencia en una muestra es indicativa de que el individuo está en riesgo de sufrir o está sufriendo de cáncer colorrectal.

50 En una realización alternativa, los métodos de la divulgación detectan uno o más biomarcadores, cuya expresión aumentada en una muestra en relación con una muestra de control es indicativo de que el individuo está en riesgo de sufrir o está sufriendo de cáncer colorrectal.

55 En una realización preferida, los métodos de la divulgación detectan uno o más biomarcadores del grupo definido en la Tabla 2. Los biomarcadores proporcionados en la Tabla 2 representan un subconjunto preferido de los definidos en la Tabla 1, cuya expresión diferencial es relativa a las muestras de control se considera estadísticamente significativa. Por lo tanto, este grupo de biomarcadores representa un panel de biomarcadores a partir del cual se puede seleccionar cualquier número de biomarcadores para la detección.

60 Por lo tanto, en una realización, los métodos de la divulgación pueden detectar más de un biomarcador del grupo definido en la Tabla 2, por ejemplo dos, tres, cuatro, cinco, seis, siete, ocho, nueve, diez de los biomarcadores del grupo definido en la Tabla 2. En una realización alternativa, los métodos de la divulgación detectan más de diez de los biomarcadores del grupo definido en la Tabla 2, por ejemplo, once, doce, trece, catorce, quince, dieciséis, diecisiete, dieciocho, diecinueve, veinte, treinta, cuarenta, cincuenta, sesenta, setenta, ochenta, noventa, cien de los biomarcadores del grupo definido en la Tabla 2. Por lo tanto, los métodos de la divulgación detectan la presencia o aumento de la expresión de uno o más de: componente C4B del complemento (grupo sanguíneo Chido) 2 (C4A/C4B); glutámico-oxaloacético transaminasa 2, mitocondrial (aspartato aminotransferasa 2) (GOT2); glucosa-6-fosfato isomerasa (GPI); transcetolasa (TKT); N-acilaminoacil-péptido hidrolasa (APEH); grupo de histonas 1, H4c (HIST4H4

(incluye otros)); proteína 5 de unión a ácidos grasos (asociada a psoriasis) (FABP5); hexosaminidasa B (polipéptido beta) (HEXB); molécula de adhesión celular epitelial (EPCAM); nucleósido difosfato quinasa (NME1-NME2); superóxido dismutasa 2, mitocondrial (SOD2); factor de alargamiento de traducción Tu, mitocondrial (TUFM); glutatión sintetasa (GSS); anexina A2 (ANXA2); ATP sintasa, transporte de H<sup>+</sup>, complejo mitocondrial F1, polipéptido beta (ATP5B); proteína de choque térmico de 10 kDa (chaperonina 10) (HSPE1); glioxalasa I (GLO1); grupo de histonas 2, H2be (HIST2H2BE (incluye otros)); proteína S100 de unión a calcio A4 (S100A4); proteína S100 de unión a calcio A11 (S100A11); latexina (LXN); miembro 11 de deshidrogenasa/reductasa (familia SDR) (DHRS11); N-acetilglucosiltransferasa, alfa (NAGLU); translina (TSN); subunidad mayor del proteasoma (prosoma, macropaína) alfa tipo 4 (PSMA4); subunidad de proteasoma (prosoma, macropaína) alfa tipo-6 (PSMA6); sustrato 1 de toxina botulínica C3 relacionada con ras (familia rho, proteína pequeña Rac1 de unión a GTP) (RAC1); adenosilhomocisteinasa (AHCY); fucosidasa, alfa-L-1, tejido (FUCA1); proteína S100 de unión a calcio P (S100P); subunidad de proteasoma (prosoma, macropaína) beta tipo-2 (PSMB2); X-prolil aminopeptidasa (aminopeptidasa P) 1 (XPNPEP1); queratina 18 (KRT18); subunidad 1 de la proteína de unión a la caperuza nuclear de 80 kDa (NCBP1); manosidasa, alfa, clase 2B, miembro 1 (MAN2B1); proteína S100 de unión a calcio A6 (S100A6); proteína que contiene valosina (VCP); quinolinato fosforibosiltransferasa (QPRT); complejo mayor de histocompatibilidad, clase I, B (HLA-B); fosfoglicerato mutasa 1 (cerebro) (PGAM1); ectonucleótido pirofosfatasa/fosfodiesterasa 3 (ENPP3); inhibidor de la serpina peptidasa, clado B (ovoalbúmina), miembro 10 (SERPINB10); mieloperoxidasa (MPO); creatina quinasa, mitocondrial 1B (CKMT1A/CKMT1B); proteinasa 3 (PRTN3); elastasa, expresada en neutrófilos (ELANE); dedo de zinc 1 tipo CW de la familia MORC (MORC1); ubiquitina B (UBB); fosfolipasa A2, grupo IIA (plaquetas, líquido sinovial) (PLA2G2A); anhidrasa carbónica IV (CA4); factor de alargamiento G, mitocondrial 2 (GFM2); proteína S100 de unión a calcio A7 (S100A7); proteína que aumenta la permeabilidad bactericida (BPI); colágeno, tipo VI, alfa 5 (COL6A5); homeobox 8 de LIM (LHX8); proteína 3 secretora rica en cisteína (CRISP3); azurocidina (AZU1); hemicentina 1, (HMCN1); transglutaminasa 3 (polipéptido E, proteína-glutamina gamma-glutamyltransferasa) (TGM3); proteína quinasa alfa de unión a CDC42 (similar a DMPK) (CDC42BPA); catepsina G (CTSG); resistina (RETN); metilmalonil CoA mutasa (MUT); que contiene repetición armadillo, 4 ligado a X (ARMCX4); integrina alfa-M (ITGAM); canal de calcio, dependiente del voltaje, subunidad alfa-1E tipo R (CACNA1E); invasión y metástasis 2 de linfoma de células T (TIAM2); homólogo A defectuoso en la regulación del ciclo celular de histona HIR (S. cerevisiae) (HIRA); miembro 2 de la familia dopey (DOPEY2); proteína 3 de unión a integrina beta 1 (ITGB1BP3); canal de sodio, controlado por voltaje, tipo VII, alfa (SCN7A); Rab3C, miembro de la familia de oncogenes RAS (RAB3C); marco de lectura abierta 79 del cromosoma 9 (C9orf79); factor nuclear de células T activadas, dependiente de calcineurina 4 (NFATC4); UDP-glucosa glicoproteína glucosiltransferasa 2 (UGGT2); cornulina (CRNN); proteína similar a quelina/cordina (KCP); molécula CD1E (CD1E); dominio de hélice superenrollada que contiene 18 residuos (CCDC18); leucotrieno A-4 hidrolasa (LTA4H); albúmina (ALB); alfa-2-macroglobulina (A2M); componente 3 del complemento (C3); hemoglobina, beta (HBB); transferrina (TF); hemoglobina, alfa 1 (HBA1/HBA2); lactotransferrina (LTF); ceruloplasmina (ferroxidasa) (CP); catalasa (CAT); componente específico del grupo (proteína de unión a la vitamina D) (GC); inhibidor de la serpina peptidasa, clado C (antitrombina), miembro 1 (SERPINC1); cadena gamma de fibrinógeno (FGG); proteína S100 de unión a calcio A8 (S100A8); ferritina, polipéptido ligero (FTL); actina, beta (ACTB); fibronectina 1 (FN1); defensina, alfa 1 (DEFA1 (incluye otros)); inhibidor de la serpina peptidasa, clado G (inhibidor de C1), miembro 1 (SERPING1); proteína 4 de unión a retinol, plasma (RBP4); peroxirredoxina 2 (PRDX2); cadena alfa de fibrinógeno (FGA); inhibidor de la serpina peptidasa, clado F (antiplasmina alfa-2, factor derivado del epitelio pigmentario), miembro 2 (SERPINF2); anhidrasa carbónica II (CA2); orosomucoide 1 (ORM1/ORM2); lactato deshidrogenasa A (LDHA); vitronectina (VTN); quinínogeno-1 (KNG1); actina, alfa, músculo cardíaco 1 (ACTC1); alfa-2-glicoproteína 1 rica en leucina (LRG1); gamma-glutamyl hidrolasa (conjugasa, folilpoligammaglutamil hidrolasa) (GGH); enolasa 1, (alfa) (ENO1); profilina 1 (PFN1); inhibidor de la serpina peptidasa, clado A (antiproteinasas alfa-1, antitripsina), miembro 7 (SERPINA7); alfa-1-microglobulina/precursor de bikunina (AMBP); lamina A/C (LMNA); apolipoproteína D (APOD); proteína de interacción con el receptor de la hormona tiroidea 11 (TRIP11); proteína de unión del componente 4 del complemento, alfa (C4BPA); tropomiosina 4 (TPM4); filamina A, alfa (FLNA); haptoglobina (HP); hemopexina (HPX); hemoglobina, delta (HBD); cadena beta de fibrinógeno (FGB); proteína S100 de unión a calcio A9 (S100A9); componente 5 del complemento (C5); familia 26 portadora de soluto, miembro 3 (SLC26A3); componente 9 del complemento (C9); componente amiloide P, suero (APCS); glicoproteína alfa-1-B (A1BG); complemento tipo C3 (LOC100133511); inhibidor H4 inter-alfa (globulina) (glicoproteína sensible a calicreína en plasma) (ITI4); componente C8 del complemento, polipéptido alfa (C8A); inhibidor inter-alfa (globulina) H1 (ITI1).

En una realización, el método de la divulgación detecta uno o más biomarcadores capaces de diagnosticar o predecir CRC en un individuo que dio negativo en la prueba inmunoquímica fecal.

Por lo tanto, los métodos de la divulgación pueden detectar uno o más biomarcadores seleccionados del grupo definido en la Tabla 3. Los biomarcadores proporcionados en la Tabla 3 representan un subconjunto preferido de los definidos en la Tabla 1, que se ha encontrado que están presentes en niveles significativamente más altos en muestras de CRC que resultaron negativas de la prueba inmunoquímica fecal. Por lo tanto, el grupo de la Tabla 3 representa un panel de biomarcadores a partir del cual se puede seleccionar cualquier número de biomarcadores para la detección.

Por lo tanto, en una realización, los métodos de la divulgación pueden detectar más de un biomarcador del grupo definido en la Tabla 3, por ejemplo, dos, tres, cuatro, cinco, seis, siete, ocho, nueve, diez de los biomarcadores del grupo definido en la Tabla 3. En una realización alternativa, los métodos de la divulgación detectan más de diez de los biomarcadores del grupo definido en la Tabla 3, por ejemplo, once, doce, trece, catorce, quince, dieciséis, diecisiete,

dieciocho, diecinueve, veinte, treinta, cuarenta, cincuenta, sesenta, setenta, ochenta, noventa, cien de los biomarcadores del grupo definido en la Tabla 3. Por lo tanto, los métodos de divulgación pueden detectar la presencia o el aumento de expresión de uno o más de: componente C4B del complemento (grupo sanguíneo Chido) 2 (C4A/C4B); glutámico-oxaloacético transaminasa 2, mitocondrial (aspartato aminotransferasa 2) (GOT2); glucosa-6-fosfato isomerasa (GPI); transcetolasa (TKT); N-acilaminoacil-péptido hidrolasa (APEH); hexosaminidasa B (polipéptido beta) (HEXB); molécula de adhesión celular epitelial (EPCAM); lectura de NME1-NME2 (NME1-NME2); factor de alargamiento de traducción Tu, mitocondrial (TUFM); glutatión sintetasa (GSS); glioxalasa I (GLO1); latexina (LXN); subunidad del proteasoma (prosome, macropaina) alfa tipo 4 (PSMA4); fucosidasa, alfa-L-1, tejido (FUCA1); queratina 18 (KRT18); subunidad 1 de la proteína de unión a la caperuza nuclear de 80 kDa (NCBP1); manosidasa, alfa, clase 2B, miembro 1 (MAN2B1); proteína S100 de unión a calcio A6 (S100A6); complejo mayor de histocompatibilidad, clase I, B (HLA-B); ectonucleótido pirofosfatasa/fosfodiesterasa 3 (ENPP3); inhibidor de la serpin peptidasa, clado B (ovoalbúmina), miembro 10 (SERPINB10); mieloperoxidasa (MPO); proteinasa 3 (PRTN3); fosfolipasa A2, grupo IIA (plaquetas, líquido sinovial) (PLA2G2A); anhidrasa carbónica IV (CA4); factor de alargamiento G, mitocondrial 2 (GFM2); proteína S100 de unión a calcio A7 (S100A7); proteína que aumenta la permeabilidad bactericida (BPI); colágeno, tipo VI, alfa 5 (COL6A5); homeobox 8 de LIM (LHX8); azurocidina (AZU1); hemicentina 1 (HMCN1); metilmalonil CoA mutasa (MUT); que contiene repetición armadillo, 4 ligado a X (ARMCX4); integrina alfa-M (ITGAM); canal de calcio, dependiente del voltaje, subunidad alfa-1E tipo R (CACNA1E); invasión y metástasis 2 de linfoma de células T (TIAM2); homólogo A defectuoso en la regulación del ciclo celular de histona HIR (*S. cerevisiae*) (HIRA); miembro 2 de la familia dopey (DOPEY2); canal de sodio, controlado por voltaje, tipo VII, alfa (SCN7A); marco de lectura abierta 79 del cromosoma 9 (C9orf79); UDP-glucosa glicoproteína glucosiltransferasa 2 (UGGT2); cornulina (CRNN); dominio de hélice superenrollada que contiene 18 residuos (CCDC18); alfa-2-macroglobulina (A2M); componente 3 del complemento (C3); hemoglobina, beta (HBB); transferrina (TF); hemoglobina, alfa 1 (HBA1/HBA2); lactotransferrina (LTF); ceruloplasmina (ferroxidasa) (CP); catalasa (CAT); cadena gamma de fibrinógeno (FGG); proteína S100 de unión a calcio A8 (S100A8); ferritina, polipéptido ligero (FTL); fibronectina 1 (FN1); defensina, alfa 1 (DEFA1 (incluye otros)); inhibidor de la serpin peptidasa, clado G (inhibidor de C1), miembro 1 (SERPING1); proteína 4 de unión a retinol, plasma (RBP4); peroxirredoxina 2 (PRDX2); inhibidor de la serpin peptidasa, clado F (antiplasmina alfa-2, factor derivado del epitelio pigmentario), miembro 2 (SERPINF2); lactato deshidrogenasa (LDHA); vitronectina (VTN); quinínogeno-1 (KNG1); alfa-2-glicoproteína 1 rica en leucina (LRG1); gamma-glutamil hidrolasa (conjugasa, foliopoligammaglutamil hidrolasa) (GGH); inhibidor de la serpin peptidasa, clado A (antiproteinasa alfa-1, antitripsina), miembro 7 (SERPINA7); alfa-1-microglobulina/precursor de bikunina (AMBP); interactuador 11 receptor de la hormona tiroidea (TRIP11); proteína de unión al componente 4 del complemento, alfa (C4BPA); filamina A, alfa (FLNA); hemopexina (HPX); proteína S100 de unión a calcio A9 (S100A9); componente 5 del complemento (C5); familia 26 portadora de soluto, miembro 3 (SLC26A3); componente 9 del complemento (C9); componente amiloide P, suero (APCS); complemento tipo C3 (LOC100133511); inhibidor de la serpin peptidasa, clado A (antiproteinasa alfa-1, antitripsina), miembro 3 (SERPINA3); catepsina S (CTSS); citidina desaminasa (CDA); similar a shugoshina 2 (SGOL2); inhibidor de la serpin peptidasa, clado B (ovoalbúmina), miembro 3 (SERPINB3); dominio hect y RLD 2 (HERC2); triosafosfato isomerasa 1 (TPI1); isoforma 1 de colágeno, tipo IV, proteína de unión alfa 3 (antígeno de Goodpasture) (COL4A3BP); ribonucleasa T2 (RNASET2); glutatión reductasa (GSR); homólogo 3 de spinster (*Drosophila*) (SPNS3); ADNc FLJ60317, muy similar a aminoacilasa-1 (ACY1); inhibidor de la serpin peptidasa, clado B (ovoalbúmina), miembro 8 (SERPINB8); aminotransferasa de aminoácidos de cadena ramificada (BCAT2); ácido siálico acetiltransferasa (SIAE); peptidasa D (PEPD); complejo mayor de histocompatibilidad, clase II, DR beta 5 (HLA-DRB5); dominio SET que contiene 2 residuos (SETD2); dominio hect (homólogo al terminal carboxilo E6-AP (UBE3A)) y el dominio 1 similar a RCC1 (CHC1) (RLD) (HERC1); isocitrato deshidrogenasa 1 (NADP+), soluble (IDH1); catepsina C (CTSC); inhibidor de la serpin peptidasa, clado A (antiproteinasa alfa-1, antitripsina), miembro 1 (SERPINA1); envoltura nuclear 1, que contiene repetición de espectrina (SYNE1); proteína 35 que contiene dominio de repetición de anquirina (ANKRD35); quinasa 10 relacionada con NIMA (nunca en el gen a de mitosis) (NEK10); inhibidor H2 inter-alfa (globulina) (ITH2); acil-CoA deshidrogenasa, cadena muy larga (ACADVL); nebulina (NEB); proteína 16 de membrana de gránulos de zimógeno (ZG16); vinculina (VCL); isoforma 2 del dedicator de la proteína 4 de citoquinesis (DOCK4); proteína hipotética LOC643677 (C13orf40); proteína no caracterizada KIAA1797 (KIAA1797); repetición de IAP baculoviral que contiene 6 residuos (BIRC6); transaldolasa 1 (TALDO1); receptor del gusto tipo 2 miembro 42 (TAS2R42); quitinasa 1 (quitotriosidasa) (CHIT1); quiescina Q6 sulfhidril oxidasa 1 (QSOX1); antígeno 12 de células estromales de médula ósea (BST1); bleomicina hidrolasa (BLMH); lisozima C (LYZ).

En una realización, el método de la divulgación detecta uno o más biomarcadores que tienen un poder discriminatorio más alto que la hemoglobina. En otras palabras, el uno o más biomarcadores tienen una sensibilidad y especificidad más altas que la hemoglobina para detectar la presencia de adenoma avanzado o adenocarcinoma en una muestra.

Por lo tanto, en una realización, el método de la divulgación detecta uno o más biomarcadores, por ejemplo dos, tres, cuatro, cinco, seis, siete, ocho, nueve, diez de los biomarcadores seleccionados del grupo que consiste en: proteína S100 de unión al calcio A8 (S100A8), componente C4B del complemento (grupo sanguíneo Chido) 2 (C4A/C4B), transferrina (TF), alfa-2-macroglobulina (A2M), proteína S100 de unión al calcio A9 (S100A9), proteinasa 3 (PRTN3), azurocidina (AZU1), lactotransferrina (LTF), hemopexina (HPX) y defensina, alfa 1 (DEFA1).

Por lo tanto, en una realización, el método de la divulgación detecta uno o más biomarcadores, por ejemplo dos, tres, cuatro, cinco, seis, siete, ocho, nueve, diez de los biomarcadores seleccionados del grupo que consiste en: anexina A2 (ANXA2), azurocidina 1 (AZU1), catepsina G (CTSG), defensina, alfa 1 (DEFA1 incluye otros), elastasa, expresada

en neutrófilos (ELANE), integrina, alfa M (subunidad del receptor 3 del componente 3 del complemento) (ITGAM), mieloperoxidasa (MPO), subunidad 1 de la proteína de unión a la caperuza nuclear de 80 kDa (NCBP1), proteinasa 3 (PRTN3), proteína S100 de unión al calcio A9 (S100A9), inhibidor de la serpina peptidasa, clado B (ovoalbúmina), miembro 10 (SERPINB10), anexina A1 (ANXA1), repetición de IAP baculoviral que contiene 6 residuos (BIRC6), componente 2 del complejo de mantenimiento de minicromosomas (MCM2), quiescina Q6 sulfhidroxidasa 1 (QSOX1), inhibidor de la serpina peptidasa, clado B (ovoalbúmina), miembro 5 (SERPINB5).

En una realización, el método de la divulgación detecta uno o más biomarcadores, cuya presencia en una muestra es indicativa de que el individuo está en riesgo de sufrir o está sufriendo de cáncer colorrectal.

Por lo tanto, en una realización, el método de la divulgación detecta uno o más biomarcadores, por ejemplo dos, tres, cuatro, cinco, seis, siete, ocho, nueve, diez de los biomarcadores seleccionados del grupo definido en la Tabla 4. Los biomarcadores proporcionados en la Tabla 4 representan un subconjunto preferido de los definidos en la Tabla 1, que se ha encontrado que se expresan solo en muestras de CRC, no en muestras de control. Por lo tanto, el grupo de la Tabla 4 representa un panel de biomarcadores a partir del cual se puede seleccionar cualquier número de biomarcadores para la detección.

En una realización, el método de la divulgación detecta uno o más biomarcadores, cuya presencia en una muestra es indicativa de que el individuo está en riesgo de sufrir o está sufriendo de cáncer colorrectal.

Por lo tanto, en una realización, el método de la divulgación detecta uno o más biomarcadores, por ejemplo dos, tres, cuatro, cinco, seis, siete, ocho, nueve, diez de los biomarcadores seleccionados del grupo definido en la Tabla 5. Los biomarcadores proporcionados en la Tabla 5 representan un subconjunto preferido de los definidos en la Tabla 1, que se identificaron en un conjunto de verificación independiente de muestras de heces.

Por lo tanto, en una realización adicional, el método de la divulgación detecta uno o más biomarcadores, por ejemplo dos, tres, cuatro, cinco, seis, siete, ocho, nueve, diez de los biomarcadores seleccionados del grupo que consiste en: glicoproteína alfa-1-B (A1BG), alfa-2-macroglobulina (A2M), actina, beta (ACTB), actina, alfa, músculo cardíaco 1 (ACTC1), albúmina (ALB), componente amiloide P, suero (APCS), N-acilaminoacil-péptido hidrolasa (APEH), apolipoproteína D (APOD), azurocidina 1 (AZU1), componente 3 del complemento (C3), componente 4B del complemento (grupo sanguíneo Chido) (C4A/C4B), componente 5 del complemento (C5), anhidrasa carbónica II (CA2), anhidrasa carbónica IV (CA4), catalasa (CAT), ceruloplasmina (ferroxidasa) (CP), catepsina G (CTSG), defensina, alfa 1 (DEFA1 (incluye otros)), elastasa, expresada en neutrófilos (ELANE), enolasa 1, (alfa) (ENO1), filamina A, alfa (FLNA), fibronectina 1 (FN1), ferritina, polipéptido ligero (FTL), fucosidasa, alfa-L-1, tejido (FUCA1), gamma-glutamyl hidrolasa (conjugasa, folilpoligammaglutamil hidrolasa) (GGH), glutámico-oxaloacético transaminasa 2, mitocondrial (aspartato aminotransferasa 2) (GOT2), glucosa-6-fosfato isomerasa (GPI), glutatión sintetasa (GSS), hemoglobina, alfa 1 (HBA1/HBA2), hemoglobina, beta (HBB), hemoglobina, delta (HBD), hexosaminidasa B (polipéptido beta) (HEXB), grupo de histonas 1, H4c (HIST4H4 (incluye otros)), complejo mayor de histocompatibilidad, clase I, B (HLA-B), haptoglobina (HP), hemopexina (HPX), proteína 1 de choque térmico de 10kDa (chaperonina 10) (HSP10), lactato deshidrogenasa A (LDHA), complemento tipo C3 (LOC100133511), alfa-2-glicoproteína 1 rica en leucina (LRG1), lactotransferrina (LTF), latexina (LXN), manosidasa, alfa, clase 2B, miembro 1 (MAN2B1), mieloperoxidasa (MPO), N-acetilglucosaminidasa, alfa (NAGLU), orosomucoide 1 (ORM1/ORM2), profilina 1 (PFN1), fosfolipasa A2, grupo IIA (plaquetas, líquido sinovial) (PLA2G2A), proteinasa 3 (PRTN3), subunidad del proteasoma (prosoma, macropaína), tipo alfa, 4 (PSMA4), subunidad del proteasoma (prosoma, macropaína), tipo alfa, 6 (PSMA6), subunidad del proteasoma (prosoma, macropaína), tipo beta, 2 (PSMB2), proteína 4 de unión a retinol, plasma (RBP4), resistina (RETN), proteína S100 de unión a calcio A8 (S100A8), proteína S100 de unión a calcio A9 (S100A9), inhibidor de la serpina peptidasa, clado B (ovoalbúmina), miembro 10 (SERPINB10), inhibidor de la serpina peptidasa, clado C (antitrombina), miembro 1 (SERPINC1), inhibidor de la serpina peptidasa, clado F (antiplasmina, alfa-2, factor derivado del epitelio pigmentario), miembro 2 (SERPINF2), inhibidor de la serpina peptidasa, clado G (inhibidor de C1), miembro 1 (SERPING1), superóxido dismutasa 2, mitocondrial (SOD2), transferrina (TF), transcetolasa (TKT), ubiquitina B (UBB), no mapeado por la invención (IPI00884078), no mapeado por la invención (proteína de 26 kDa, IPI00942608), alfa-2-macroglobulina tipo 1 (A2ML1), acil-CoA deshidrogenasa, cadena muy larga (ACADVL), no mapeado por la invención (ACY1), fosfatasa alcalina, hígado/hueso/riñón (ALPL), bleomicina hidrolasa (BLMH), antígeno 1 de células estromales de médula ósea (BST1), canal de calcio, dependiente del voltaje, tipo L, subunidad alfa 1D (CACNA1D), creatina quinasa, mitocondrial 1B (CKMT1A/CKMT1B), catepsina S (CTSS), fumarilacetoacetato hidrolasa (fumarilacetoacetasa) (FAH), guanina desaminasa (GDA) glutatión reductasa (GSR), ribonucleoproteína nuclear heterogénea A2/B1 (HNRNPA2B1), proteína de choque térmico de 70 kDa (HSPA8), queratina 6C (KRT6C), queratina 6C (KRT6C), lipocalina 1 (prealbúmina de lágrima) (LCN1), lisozima, (LYZ), componente 2 del complejo de mantenimiento de minicromosomas (MCM2), peptidasa D (PEPD), dominio de fosfolipasa B que contiene 1 residuo (PLBD1), subunidad del proteasoma (prosoma, macropaína), tipo alfa, 2 (PSMA2), subunidad del proteasoma (prosoma, macropaína), tipo beta, 1 (PSMB1), subunidad del proteasoma (prosoma, macropaína), tipo beta, 6 (PSMB6), quiescina Q6 sulfhidroxidasa 1 (QSOX1), ribonucleasa T2 (RNASET2), inhibidor de la serpina peptidasa, clado A (antiproteinasa alfa-1, antitripsina), miembro 1 (SERPINA1), inhibidor de la serpina peptidasa, clado A (antiproteinasa alfa-1, antitripsina), miembro 3 (SERPINA3), inhibidor de la serpina peptidasa, clado B (ovoalbúmina), miembro 3 (SERPINB3), ácido siálico acetiltransferasa (SIAE), trifosfato isomerasa 1 (TPI1), homólogo de proteína de gránulos de zimógeno 16 (rata) (ZG16).

En una realización, el método de la divulgación detecta uno o más biomarcadores, cuya presencia en una muestra es indicativa de que el individuo está en riesgo de sufrir o está sufriendo de cáncer colorrectal, cuyos biomarcadores se han seleccionado con base en la correlación con secretomas de tejido tumoral de colon.

5 Por lo tanto, en una realización, el método de la divulgación detecta uno o más biomarcadores, por ejemplo dos, tres, cuatro, cinco, seis, siete, ocho, nueve, diez de los biomarcadores seleccionados del grupo definido en la Tabla 7. Los biomarcadores proporcionados en la Tabla 7 representan un subconjunto preferido de aquellos definidos en la Tabla 6, que se ha encontrado que se expresan tanto en la emisión de tumores como en las muestras de heces. Por lo tanto, 10 el grupo de la Tabla 7 representa un panel de biomarcadores a partir del cual se puede seleccionar cualquier número de biomarcadores para la detección. Por lo tanto, en una realización adicional, el método de la divulgación detecta uno o más biomarcadores, por ejemplo dos, tres, cuatro, cinco, seis, siete, ocho, nueve, diez de los biomarcadores seleccionados del grupo que consiste en: anexina A2 (ANXA2), azurocidina 1 (AZU1), cathepsina G (CTSG), defensina, alfa 1 (DEFA1 (incluye otros)), elastasa, expresada en neutrófilos (ELANE), integrina, alfa M (subunidad del receptor 3 del componente 3 del complemento) (ITGAM), mieloperoxidasa (MPO), subunidad 1 de la proteína de unión a la caperuza nuclear de 80 kDa (NCBP1), proteinasa 3 (PRTN3), proteína S100 de unión al calcio A9 (S100A9), inhibidor de la serpin peptidasa, clado B (ovoalbúmina), miembro 10 (SERPINB10), anexina A1 (ANXA1), repetición de IAP baculoviral que contiene 6 residuos (BIRC6), componente 2 del complejo de mantenimiento de minicromosomas (MCM2), quiescina Q6 sulfhidroxidasa 1 (QSX1), inhibidor de la serpin peptidasa, clado B (ovoalbúmina), miembro 5 (SERPINB5).

Preferiblemente, el método de la divulgación tiene una precisión de al menos 75%, por ejemplo una precisión del 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100%. Preferiblemente, el método de la invención tiene una sensibilidad de al menos 75%, por ejemplo una sensibilidad del 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100%.

Preferiblemente, el método de la invención tiene una especificidad de al menos 75%, por ejemplo una especificidad del 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100%.

Por "precisión" se entiende la proporción de resultados correctos de un diagnóstico, por "sensibilidad" se entiende la proporción de todos los diagnósticos positivos que se clasifican correctamente como positivos, y por "especificidad" se entiende la proporción de todos los diagnósticos negativos que se clasifican correctamente como negativos. Los métodos de la invención pueden verificarse mediante colonoscopia posterior.

#### Métodos de detección

En una realización, la muestra biológica, por ejemplo una muestra de heces, puede detectarse mediante uno o más biomarcadores usando espectrometría de masas (dirigida).

El descubrimiento y la detección de marcadores de proteínas por LC-MS/MS y la detección de marcadores de proteínas basados en hipótesis por SRM-MS se han aplicado antes en muestras de heces humanas y murinas, respectivamente (Ang CS, Nice EC. J Proteome Res 2010; 9: 4346-55; Ang CS, et al., Electrophoresis 2011; 32: 1926-38; Ang CS, et al., J Chromatogr A 2010; 1217: 3330-40).

A pesar de la complejidad del material de heces, la presente divulgación ha detectado con éxito varias de estas proteínas de heces humanas previamente informadas. Además, el número de muestras incluidas y las proteínas identificadas correspondientes en el estudio actual supera el de estudios anteriores. Por lo tanto, los datos presentados aquí son la visión general más grande del proteoma de heces humano hasta la fecha.

Aunque se ha demostrado que la detección de muestras mediante las técnicas de espectrometría de masas descritas en el presente documento es eficaz, un método de detección preferido comprende un ensayo de detección basado en anticuerpos. La prueba inmunoquímica fecal (FIT) comprende un ensayo de detección basado en anticuerpos y, por lo tanto, un ensayo de detección basado en anticuerpos para uno o más de los biomarcadores identificados en la presente solicitud proporciona una detección complementaria que puede incorporarse fácilmente en el ensayo FIT existente.

Por lo tanto, en una realización de los métodos de la divulgación, la muestra biológica se analiza para uno o más biomarcadores usando un agente de unión capaz de unirse a uno o más biomarcadores.

Los agentes de unión (también denominados moléculas de unión) se pueden seleccionar de una biblioteca, en función de su capacidad para unirse a un motivo dado, como se describe a continuación.

En una realización, el primer agente de unión es un anticuerpo o un fragmento del mismo. Por lo tanto, un fragmento puede contener uno o más de los dominios pesados variables ( $V_H$ ) o ligeros variables ( $V_L$ ). Por ejemplo, el término

fragmento de anticuerpo incluye moléculas similares a Fab (Better et al., Science 1988; 240, 1041); moléculas Fv (Skerra et al., Science 1988; 240, 1038); moléculas de Fv de cadena sencilla (ScFv) donde los dominios asociados V<sub>H</sub> y V<sub>L</sub> están unidos a través de un oligopéptido flexible (Bird et al., Science 1988; 242,423; Huston et al., Proc. Natl. Acad. Sci. USA 1988; 85, 5879) y anticuerpos de un solo dominio (dAb) que comprenden dominios V aislados (Ward et al., Nature 1989; 341,544).

El término "variante de anticuerpo" incluye cualquier anticuerpo sintético, anticuerpo recombinante o híbridos de anticuerpo, tal como, pero sin limitación, una molécula de anticuerpo de cadena sencilla producida por presentación en fagos de regiones variable y/o constante de cadena ligera y/o pesada de inmunoglobulina u otra molécula inmunointeractiva capaz de unirse a un antígeno en un formato de inmunoensayo conocido por los expertos en la técnica.

Se puede encontrar una revisión general de las técnicas implicadas en la síntesis de fragmentos de anticuerpos que retienen sus sitios de unión específicos en Winter & Milstein Nature 1991; 349, 293-299.

En una realización, el anticuerpo o fragmento del mismo es un anticuerpo recombinante o fragmento del mismo (tal como un scFv). Por "moléculas de ScFv" se entiende moléculas en las que los dominios asociados V<sub>H</sub> y V<sub>L</sub> están unidos a través de un oligopéptido flexible.

Las ventajas de usar fragmentos de anticuerpos, en lugar de anticuerpos completos, son múltiples. Se eliminan las funciones efectoras de los anticuerpos completos, tales como la unión del complemento. Los fragmentos de anticuerpos Fab, Fv, ScFv y dAb pueden expresarse todos y secretarse a partir de *E. coli*, lo que permite la producción fácil de grandes cantidades de dichos fragmentos.

Los anticuerpos completos y los fragmentos F(ab')<sub>2</sub> son "bivalentes". Por "bivalente" se entiende que dichos anticuerpos y los fragmentos F(ab')<sub>2</sub> tienen dos sitios de combinación de antígeno. En contraste, los fragmentos Fab, Fv, ScFv y dAb son monovalentes y tienen un solo sitio de combinación de antígeno.

Los anticuerpos pueden ser monoclonales o policlonales. Los anticuerpos adecuados pueden prepararse mediante técnicas conocidas y no necesitan discusión adicional.

Además o alternativamente, el agente de unión puede ser un aptámero. Los aptámeros adecuados pueden prepararse mediante técnicas conocidas y no necesitan más discusión.

En una realización, los uno o más biomarcadores en la muestra de prueba se marcan con una fracción detectable. En una realización, los uno o más biomarcadores en la muestra de control se marcan con una fracción detectable (que puede ser la misma o diferente de la fracción detectable usada para marcar la muestra de prueba).

Una "fracción detectable" es una que puede detectarse y determinarse la cantidad relativa y/o la ubicación de la fracción (por ejemplo, la ubicación en una matriz).

Las fracciones detectables son bien conocidas en la técnica. Una fracción detectable puede ser una fracción fluorescente y/o luminiscente y/o quimioluminiscente que, cuando se expone a condiciones específicas, puede detectarse. Por ejemplo, una fracción fluorescente puede necesitar exponerse a radiación (es decir, luz) a una longitud de onda e intensidad específicas para provocar la excitación de la fracción fluorescente, lo que le permite emitir fluorescencia detectable a una longitud de onda específica que puede detectarse.

Como alternativa, la fracción detectable puede ser una enzima que es capaz de convertir un sustrato (preferiblemente indetectable) en un producto detectable que se puede visualizar y/o detectar. Los ejemplos de enzimas adecuadas se analizan con más detalle a continuación en relación con, por ejemplo, ensayos de ELISA.

Como alternativa, la fracción detectable puede ser un marcador radiactivo, que puede incorporarse mediante métodos bien conocidos en la técnica.

## Matrices

En una realización, los métodos de la presente invención pueden llevarse a cabo en una matriz.

Las matrices por sí mismas ya son bien conocidas en la técnica. Típicamente, están formadas por una estructura lineal o bidimensional que tiene regiones separadas (es decir, discretas) ("puntos"), cada una con un área finita, formada en la superficie de un soporte sólido. Una matriz también puede ser una estructura de perlas donde cada perla puede identificarse mediante un código molecular o código de color o identificarse en un flujo continuo. El análisis también se puede realizar secuencialmente cuando la muestra se pasa sobre una serie de puntos, cada uno de los cuales adsorbe la clase de moléculas de la solución.

El soporte sólido es típicamente vidrio o un polímero, siendo los polímeros usados más comúnmente celulosa,

5 poliacrilamida, nylon, poliestireno, cloruro de polivinilo o polipropileno. Los soportes sólidos pueden estar en forma de tubos, perlas, discos, chips de silicio, microplacas, membrana de difluoruro de polivinilideno (PVDF), membrana de nitrocelulosa, membrana de nylon, otra membrana porosa, membrana no porosa (por ejemplo, plástico, polímero, perspex, silicio, entre otros), una pluralidad de pernos poliméricos, o una pluralidad de pozos de microtitulación, o cualquier otra superficie adecuada para inmovilizar proteínas, anticuerpos y otras moléculas adecuadas y/o realizar un inmunoensayo.

10 Los procesos de unión son bien conocidos en la técnica y generalmente consisten en unión covalente por entrecruzamiento o adsorción física de una molécula de proteína, anticuerpo o similar al soporte sólido. Mediante el uso de técnicas bien conocidas, tales como la impresión por contacto o sin contacto, el enmascaramiento o la fotolitografía, se puede definir la ubicación de cada punto.

15 Una vez que se han identificado y aislado las moléculas de unión adecuadas (discutidas anteriormente), la persona experta puede fabricar una matriz usando métodos bien conocidos en la técnica de la biología molecular.

En una realización, la detección puede comprender el uso de un ensayo que comprende un segundo agente de unión capaz de unirse a uno o más biomarcadores, teniendo el segundo agente de unión una fracción detectable.

20 En una realización, el segundo agente de unión es un anticuerpo o un fragmento del mismo (por ejemplo, como se describió anteriormente en relación con el primer agente de unión).

25 Típicamente, el ensayo es un ELISA (ensayo de inmunosorbente ligado a enzimas) que generalmente implica el uso de enzimas que dan un producto de reacción coloreado, generalmente en ensayos de fase sólida. Las enzimas como la peroxidasa de rábano picante y la fosfatasa se han empleado ampliamente. Una forma de amplificar la reacción de fosfatasa es usar NADP como sustrato para generar NAD que ahora actúa como una coenzima para un segundo sistema enzimático. La pirofosfatasa de *Escherichia coli* proporciona un buen conjugado porque la enzima no está presente en los tejidos, es estable y proporciona un buen color de reacción. También se pueden utilizar sistemas quimioluminiscentes basados en enzimas como la luciferasa.

30 La conjugación con la vitamina biotina también se utiliza, ya que esto puede detectarse fácilmente por su reacción con avidina o estreptavidina unida a enzimas a las que se une con gran especificidad y afinidad.

35 Los expertos en la materia apreciarán que existe un grado de fluidez en la composición de biomarcadores de las firmas de la invención. Por lo tanto, diferentes combinaciones de los biomarcadores pueden ser igualmente útiles en el diagnóstico, pronóstico y/o caracterización del cáncer colorrectal. De esta manera, cada biomarcador (ya sea solo o en combinación con uno o más de otros biomarcadores) contribuye a la firma.

Compuestos y métodos para tratar el CRC

40 La identificación de los biomarcadores como se define en la Tabla 1 y/o la Tabla 6 permite no solo la detección de adenomas de colon avanzados y adenocarcinomas de colon (cáncer colorrectal), sino que también permite métodos para tratar cánceres colorrectales y también proporciona compuestos para su uso en métodos de tratamiento de cánceres colorrectales.

45 Aunque el diagnóstico precoz del cáncer colorrectal a menudo permite la extirpación quirúrgica curativa del tumor, el diagnóstico posterior puede dar lugar a un tratamiento (quimio)terapéutico. Los agentes terapéuticos utilizados para tratar el cáncer colorrectal incluyen anticuerpos monoclonales, inhibidores de moléculas pequeñas y agentes quimioterapéuticos.

50 Los anticuerpos monoclonales terapéuticos típicos incluyen, pero sin limitación, bevacizumab, cetuximab o panitumumab. Los inhibidores típicos de moléculas pequeñas incluyen, entre otros, erlotinib, sorafenib o alisertib. Los agentes quimioterapéuticos típicos incluyen, pero sin limitación, 5-FU, capecitabina, irinotecano oxaliplatino o leucovorina o cualquier combinación de los mismos. Se pueden usar terapias combinadas de, por ejemplo, un anticuerpo monoclonal terapéutico y un inhibidor de molécula pequeña. Por lo tanto, se contempla cualquier combinación de dos o más de un anticuerpo monoclonal, un inhibidor de molécula pequeña y un agente quimioterapéutico.

Kit para realizar el método

60 El kit para realizar el método de acuerdo con la invención puede seleccionarse de cualquier ensayo y aparatos y equipos de procesamiento de datos adecuados.

La selección adecuada estará dentro de la capacidad de los expertos en la materia, y en el presente documento no es necesaria una descripción adicional.

65 "Que comprende"

El término "que comprende" y los términos relacionados en este documento deben interpretarse como que abarca "que consiste esencialmente en" y "que consiste en", estas dos expresiones son intercambiables con "que comprende" en todas las definiciones y discusión en esta patente para especificar extensiones alternativas de exclusión de elementos no especificados adicionales a los elementos mencionados.

5 El término "que comprende" y expresiones relacionadas significa "que incluye" y, por lo tanto, deja abierta la opción de incluir elementos no especificados, ya sean esenciales o no. El término "que consiste esencialmente en" y expresiones relacionadas permite la presencia de elementos que no afectan materialmente las características básicas y novedosas o funcionales de esa realización de la invención. El término "que consiste en" y expresiones relacionadas  
10 significa "que consiste solamente en" y, por lo tanto, excluye cualquier elemento no mencionado en esa descripción de la realización.

Breve descripción de las Figuras

15 La invención se describirá ahora adicionalmente mediante el siguiente ejemplo con referencia a las figuras adjuntas. El ejemplo se proporciona solo a modo de ejemplo, sin ninguna limitación prevista del alcance de la invención. La Figura 1A muestra el número de proteínas humanas detectadas en muestras de heces de 12 pacientes con CRC y de 10 pacientes de control. El diagrama de Venn muestra el número de proteínas detectadas en CRC o muestras de heces de control y su superposición;

20 La Figura 1B muestra la ubicación subcelular de las proteínas humanas identificadas. El gráfico circular muestra el porcentaje de diferentes ubicaciones subcelulares de las 830 proteínas humanas que se identificaron en 22 muestras de heces. Las proporciones fueron similares para pacientes con CRC y pacientes de control;

La Figura 2A muestra un diagrama de dispersión de cuantificaciones de proteínas de hemoglobina alfa medida por LC-MS/MS en comparación con la cuantificación de hemoglobina por FIT de muestras individuales de heces;

25 La Figura 2B muestra un diagrama de dispersión de las cuantificaciones de proteínas de la hemoglobina beta medida por LC-MS/MS en comparación con la cuantificación de hemoglobina por FIT de muestras individuales de heces;

La Figura 2C muestra un diagrama de dispersión de cuantificaciones de proteínas de hemoglobina alfa y beta medidas por LC-MS/MS de muestras individuales de heces;

30 La Figura 2D muestra un diagrama de dispersión de cuantificaciones de proteínas de S100A8 y S100A9 medidas por LC-MS/MS de muestras individuales de heces. Los círculos rellenos representan muestras de heces de sujetos con CRC (n = 12), las cruces representan muestras de heces de sujetos sin neoplasia de colon (n = 10); y

La Figura 3 muestra un diagrama de dispersión de cuantificaciones de proteínas del componente 4B del complemento medido por LC-MS/MS en relación con los valores de FIT de muestras individuales de heces. La línea punteada muestra el corte de 75 ng/mL para FIT. Un valor de FIT inferior a 75 ng/mL se considera un resultado negativo de la prueba (Van Veen, NTG, 2009). Los círculos rellenos representan muestras de heces de sujetos con CRC (n = 12), las cruces representan muestras de heces de sujetos sin neoplasia de colon (n = 10).

35 Las Figuras 4 y 5 muestran gráficos de barras con la media de las áreas bajo las curvas de todas las transiciones de cada péptido en análisis por triplicado. Se representan de uno a cinco péptidos diferentes por proteína. Los péptidos se midieron por triplicado en agrupaciones de muestras de heces de controles (N, n = 5), pacientes con adenoma no avanzado (A, n = 5), pacientes con adenoma avanzado (AA, n = 5) y pacientes con CRC (CRC, n = 5). Las barras de error representan desviaciones estándar.

Tablas

45 Las siguientes tablas enumeran los biomarcadores de acuerdo con la divulgación, y pueden continuar durante varias páginas.

Tabla 1: Biomarcadores indicativos de cáncer colorrectal.

Número de acceso	Símbolo del gen	Nombre del gen
IPI00892604	C4A/C4B	componente C4B del complemento (grupo sanguíneo Chido)
IPI00018206	GOT2	glutámico-oxaloacético transaminasa 2, mitocondrial (aspartato aminotransferasa 2)
IPI00027497	GPI	glucosa-6-fosfato isomerasa
IPI00643920	TKT	transcetolasa
IPI00337741	APEH	N-acilaminoacil-péptido hidrolasa
IPI00453473	HIST4H4 (incluye otros)	grupo de histonas 1, H4c
IPI00007797	FABP5	proteína 5 de unión a ácidos grasos (asociada a psoriasis)
IPI00012585	HEXB	hexosaminidasa B (beta polipéptido)
IPI00296215	EPCAM	molécula de adhesión a células epiteliales
IPI00604590	NME1-NME2	lectura de NME1-NME2
IPI00022314	SOD2	superóxido dismutasa 2, mitocondrial

ES 2 762 403 T3

IPI00027107	TUFM	factor de alargamiento de traducción Tu, mitocondrial,
IPI00010706	GSS	glutación sintetasa
IPI00418169	ANXA2	anexina A2
IPI00303476	ATP5B	ATP sintasa, transporte de H+, complejo F1 mitocondrial, polipéptido beta
IPI00220362	HSPE1	proteína de choque térmico de 10 kDa (chaperonina 10)
IPI00220766	GLO1	glioxalasa I
IPI00003935	HIST2H2BE (incluye otros)	grupo de histonas 2, H2be
IPI00032313	S100A4	proteína S100 de unión a calcio A4
IPI00013895	S100A11	proteína S100 de unión a calcio A11
IPI00106687	LXN	latexina
IPI00034280	DHRS11	miembro 11 de deshidrogenasa/reductasa (familia SDR)
IPI00008787	NAGLU	N-acetilglucosaminidasa, alfa
IPI00018768	TSN	translina
IPI00299155	PSMA4	subunidad del proteasoma (prosome, macropaína), tipo alfa, 4
IPI00029623	PSMA6	subunidad del proteasoma (prosome, macropaína), tipo alfa, 6
IPI00010271	RAC1	sustrato 1 de toxina botulínica C3 relacionada con ras (familia rho, proteína pequeña Rac1 de unión a GTP)
IPI00012007	AHCY	adenosilhomocisteinasa
IPI00385751	FUCA1	fucosidasa, alfa-L-1, tejido
IPI00017526	S100P	proteína S100 de unión a calcio P
IPI00028006	PSMB2	subunidad del proteasoma (prosome, macropaína), tipo alfa, 2
IPI00793375	XPNPEP1	X-prolil aminopeptidasa (aminopeptidasa P) 1, soluble
IPI00554788	KRT18	queratina 18
IPI00019380	NCBP1	subunidad 1 de proteína de unión a la caperuza nuclear, 80 kDa
IPI00012989	MAN2B1	manosidasa, alfa, clase 2B, miembro 1
IPI00027463	S100A6	proteína S100 de unión a calcio A6
IPI00022774	VCP	proteína que contiene valosina
IPI00300086	QPRT	quinolinato fosforibosiltransferasa
IPI00472073	HLA-B	complejo mayor de histocompatibilidad, clase I, B
IPI00453476	PGAM1	fosfoglicerato mutasa 1 (cerebro)
IPI00020999	ENPP3	ectonucleótido pirofosfatasa/fosfodiesterasa 3
IPI00010304	SERPINF10	inhibidor de la serpin peptidasa, clado B (ovoalbúmina), miembro 10
IPI00007244	MPO	mieloperoxidasa
IPI00877726	CKMT1A/ CKMT1B	creatina quinasa, mitocondrial 1B
IPI00027409	PRTN3	proteínasa 3
IPI00027769	ELANE	elastasa, expresada en neutrófilos
IPI00004362	MORC1	dedo de zinc 1 tipo CW de la familia MORC
IPI00719280	UBB	ubiquitina B
IPI00026962	PLA2G2A	fosfolipasa A2, grupo IIA (plaquetas, líquido sinovial)
IPI00027466	CA4	anhidrasa carbónica IV
IPI00071703	GFM2	factor de alargamiento G, mitocondrial 2
IPI00219806	S100A7	proteína S100 de unión a calcio A7
IPI00827847	BPI	proteína bactericida/que aumenta la permeabilidad
IPI00176125	COL6A5	colágeno, tipo VI, alfa 5
IPI00470355	LHX8	homeobox 8 de LIM
IPI00942117	CRISP3	proteína 3 secretora rica en cisteína
IPI00022246	AZU1	azurocidina 1
IPI00045512	HMCN1	hemicentina 1
IPI00300376	TGM3	transglutaminasa 3 (polipéptido E, proteína-glutamina-gamma-glutamyltransferasa)
IPI00640468	CDC42BPA	proteína quinasa alfa de unión a CDC42 (similar a DMPK)
IPI00028064	CTSG	catepsina G
IPI00006988	RETN	resistina
IPI00024934	MUT	metilmalonil CoA mutasa

ES 2 762 403 T3

IPI00552983	ARMCX4	que contiene repetición armadillo, 4 ligado a X
IPI00217987	ITGAM	integrina alfa M (subunidad del receptor 3 del componente 3 del complemento)
IPI00165045	CACNA1E	canal de calcio, dependiente del voltaje, tipo R, subunidad alfa 1E
IPI00018363	TIAM2	invasión y metástasis 2 de linfoma de células T
IPI00217560	HIRA	homólogo A defectuoso en la regulación del ciclo celular de histona HIR (S. cerevisiae)
IPI00294653	DOPEY2	miembro 2 de la familia dopey
IPI00004401	ITGB1BP3	proteína 3 de unión a integrina beta 1
IPI00300117	SCN7A	canal de sodio, controlado por voltaje, tipo VII, alfa
IPI00061114	RAB3C	RAB3C, miembro de la familia de oncogenes RAS
IPI00168442	C9orf79	marco de lectura abierto 79 del cromosoma 9
IPI00385215	NFATC4	factor nuclear de células T activadas, dependiente de calcineurina 4
IPI00024467	UGGT2	UDP-glucosa glicoproteína glucosiltransferasa 2
IPI00297056	CRNN	cornulina
IPI00914663	KCP	proteína similar a quelina/cordina
IPI00418592	CD1E	molécula CD1e
IPI00642206	CCDC18	dominio de hélice superenrollada que contiene 18 residuos
IPI00514090	LTA4H	leucotrieno A4 hidrolasa
IPI00745872	ALB	albúmina
IPI00478003	A2M	alfa-2-macroglobulina
IPI00783987	C3	componente 3 del complemento
IPI00654755	HBB	hemoglobina, beta
IPI00022463	TF	transferrina
IPI00410714	HBA1/HBA2	hemoglobina alfa 1
IPI00298860	LTF	lactotransferrina
IPI00017601	CP	ceruloplasmina (ferroxidasa)
IPI00465436	CAT	catalasa
IPI00555812	GC	componente específico del grupo (proteína de unión a la vitamina D)
IPI00032179	SERPINC1	inhibidor de la serpina peptidasa, clado C (antitrombina), miembro 1
IPI00219713	FGG	cadena gamma de fibrinógeno
IPI00007047	S100A8	proteína S100 de unión a calcio A8
IPI00375676	FTL	ferritina, polipéptido ligero
IPI00021439	ACTB	actina, beta
IPI00022418	FN1	fibronectina 1
IPI00005721	DEFA1(incluye otros)	defensina, alfa 1
IPI00291866	SERPING1	inhibidor de la serpina peptidasa, clado G (inhibidor de C1), miembro 1
IPI00022420	RBP4	proteína 4 de unión a retinol, plasma
IPI00027350	PRDX2	peroxirredoxina 2
IPI00021885	FGA	cadena alfa de fibrinógeno
IPI00029863	SERPINF2	inhibidor de la serpina peptidasa, clado F (antiplasmina alfa-2, factor derivado del epitelio pigmentario), miembro 2
IPI00218414	CA2	anhidrasa carbónica II
IPI00020091	ORM1/ORM2	orosomucoide 1
IPI00217966	LDHA	lactato deshidrogenasa A
IPI00298971	VTN	vitronectina
IPI00215894	KNG1	quininógeno 1
IPI00023006	ACTC1	actina, alfa, músculo cardíaco 1
IPI00022417	LRG1	alfa-2-glicoproteína 1 rica en leucina
IPI00023728	GGH	gamma-glutamil hidrolasa (conjugasa, folilpoligammaglutamil hidrolasa)
IPI00465248	ENO1	enolasa 1, (alfa)
IPI00216691	PFN1	profilina 1
IPI00292946	SERPINA7	inhibidor de la serpina peptidasa, clado A (antiproteinasa alfa-1, antitripsina), miembro 7
IPI00022426	AMBP	precursor de alfa-1-microglobulina/bikunina
IPI00021405	LMNA	lamina A/C

ES 2 762 403 T3

IPI00006662	APOD	apolipoproteína D
IPI00003515	TRIP11	interactuador 11 del receptor de la hormona tiroidea
IPI00021727	C4BPA	proteína de unión al componente 4 del complemento, alfa
IP100010779	TPM4	tropomiosina 4
IPI00302592	FLNA	filamina A, alfa
IPI00641737	HP	haptoglobina
IPI00022488	HPX	hemopexina
IPI00473011	HBD	hemoglobina, delta
IPI00298497	FGB	cadena beta de fibrinógeno
IPI00027462	S100A9	proteína S100 de unión a calcio A9
IPI00032291	C5	componente 5 del complemento
IPI00031036	SLC26A3	familia 26 portadora de soluto, miembro 3
IPI00022395	C9	componente 9 del complemento
IPI00022391	APCS	componente amiloide P, suero
IPI00022895	A1BG	glicoproteína alfa-1-B
IPI00887739	LOC100133511	complemento tipo C3
IPI00218192	ITIH4	inhibidor H4 inter-alfa (globulina) (glicoproteína sensible a calicreína en plasma)
IPI00011252	C8A	componente 8 del complemento, polipéptido alfa
IPI00292530	ITIH1	inhibidor H1 inter-alfa (globulina)
IPI00937735	ACADVL	acil-CoA deshidrogenasa, cadena muy larga
IPI00009268	ACY1	ADNc FLJ60317, muy similar a aminoacilasa-1
IPI00166331	ANKRD35	proteína 35 que contiene el dominio de repetición de anquirina
IPI00299635	BIRC6	repetición de IAP baculoviral que contiene 6 residuos
IPI00219575	BLMH	bleomicina hidrolasa
IPI00026240	BST1	antígeno 1 de célula estromal de médula ósea
IPI00929313	C13orf40	proteína hipotética LOC643677
IPI00027983	CDA	citidina desaminasa
IPI00852865	CHIT1	quitinasa 1 (quitotriosidasa)
IPI00022810	CTSC	dipeptidil-peptidasa 1
IPI00299150	CTSS	catepsina S
IPI00006024	DOCK4	isoforma 2 del dedicator de la proteína 4 de citoquinas
IPI00016862	GSR	glutaciona reductasa
IPI00941901	HERC1	dominio hect (homólogo al terminal carboxilo E6-AP (UBE3A)) y el dominio 1 similar a RCC1 (CHC1) (RLD)
IPI00005826	HERC2	dominio hect y RLD 2
IPI00746667	HLA-DRB5	complejo mayor de histocompatibilidad, clase II, DR beta 5
IPI00027223	IDH1	isocitrato deshidrogenasa 1 (NADP+), citoplásmica
IPI00305461	ITIH2	inhibidor H2 inter-alfa (globulina)
IPI00748360	KIAA1797	proteína no caracterizada KIAA1797
IPI00019038	LYZ	lisozima C
IPI00303335	NEB	nebulina
IPI00065454	NEK10	quinasa 10 relacionada con NIMA (nunca en el gen a de mitosis) (NEK10);
IPI00257882	PEPD	peptidasa D
IPI00003590	QSOX1	quiescina Q6 sulfhidril oxidasa 1
IPI00939604	RNAS2	ribonucleasa T2
IPI00553177	SERPINA1	inhibidor de la serpina peptidasa, clado A (alfa-1 antiproteinasa, antitripsina), miembro 1
IPI00550991	SERPINA3	inhibidor de la serpina peptidasa, clado A (alfa-1 antiproteinasa, antitripsina), miembro 3
IPI00022204	SERPINA3	inhibidor de la serpina peptidasa, clado B (ovoalbúmina), miembro 3
IPI00307733	SETD2	dominio SET que contiene 2 residuos
IPI00218013	SGOL2	similar a shugoshina 2
IPI00010949	SIAE	ácido siálico acetiltransferasa
IPI00386444	SYNE1	envoltura nuclear 1, que contiene repetición de espectrina
IPI00744692	TALDO1	transaldolasa 1

ES 2 762 403 T3

IPI00375371	TAS2R42	receptor del gusto tipo 2 miembro 42
IPI00465028	TP11	triosafosfato isomerasa 1
IPI00291175	VCL	vinculina
IPI00029647	ZG16	proteína 16 de membrana de gránulos de zimógeno
IPI00016255	PLBD1	proteína hipotética LOC79887
IPI00477505	ANKRD28	isoforma 1 de la subunidad A de repetición de anquirina reguladora de la proteína serina/treonina fosfatasa 6
IPI00032293	CST3	cistatina-C
IPI00293867	DDT	D-dopacromo descarboxilasa
IPI00059242	SYAP1	proteína 1 asociada a sinapsis
IPI00219622	PSMA2	subunidad alfa tipo-2 del proteasoma
IPI00221222	SUB1	homólogo de SUB1 (S. cerevisiae)
IPI00022791	MFAP3	glicoproteína 3 asociada a microfibrillas
IPI00011229	CTSD	catepsina D
IPI00025019	PSMB1	subunidad del proteasoma (prosoma, macropaína), tipo beta, 1
IPI00479306	PSMB5	subunidad del proteasoma (prosoma, macropaína), tipo beta, 5
IPI00234793	KCTD15	ADNc FLJ61112, muy similar a la proteína KCTD15 que contiene el dominio BTB/POZ
IPI00010796	P4HB	prolil-4-hidroxilasa, polipéptido beta
IPI00927606	GPX1	glutación peroxidasa 1
IPI00783625	SERPINB5	inhibidor de la serpina peptidasa, clado B (ovoalbúmina), miembro 5
IPI00024701	COL4A3BP	proteína de unión a colágeno, tipo IV, alfa 3 (antígeno de Goodpasture)
IPI00000811	PSMB6	subunidad del proteasoma (prosoma, macropaína), tipo beta, 6
IPI00021298	KRT20	queratina 20
IPI00025084	CAPNS1	subunidad pequeña 1 de calpaína
IPI00024919	PRDX3	peroxirredoxina 3
IPI00059930	NACC2	miembro 2 de la familia NACC, que contiene el dominio BEN y BTB (POZ)
IPI00003817	ARHGDI1	inhibidor 2 de disociación de Rho GDP
IPI00293276	MIF	factor inhibidor de migración de macrófagos
IPI00514622	RANBP6	proteína 6 de unión a Ran
IPI00333141	SPNS3	homólogo 3 de spinster (Drosophila)
IPI00922181	MCM2	componente 2 del complejo de mantenimiento de minicromosomas
IPI00031708	FAH	fumarilacetoacetasa
IPI00003865	HSPA8	proteína 8 de choque térmico de 70 kDa
IPI00299024	BASP1	proteína 1 señal, unida a membrana, abundante en el cerebro
IPI00181135	BCAT2	aminotransferasa de aminoácidos de cadena ramificada
IPI00219365	MSN	moesina
IPI00032134	SERPINB8	inhibidor de la serpina peptidasa, clado B (ovoalbúmina), miembro 8
IPI00853547	G6PD	isoforma a de la glucosa-6-fosfato deshidrogenasa
IPI00478758	C10orf119	isoforma 1 de la proteína UPF0557 C10orf119
IPI00873020	PSAP	prosaposina
IPI00000875	EEF1G	factor 1 gamma de alargamiento de traducción eucariota
IPI00014398	FHL1	cuatro y medio dominios 1 de LIM
IPI00301395	CPVL	carboxipeptidasa, tipo vitelogénica
IPI00790262	TTLL3	familia tipo tirosina ligasa de tubulina, miembro 3
IPI00942608	proteína de 26 kDa	IPI:IPI00942608.1 ENSEMBL:ENSP0000
IPI00023038	PRB1/PRB2	subfamilia 2 de la proteína BstNI rica en prolina
IPI00001895	PCDH8	protocadherina 8
IPI00419215	A2ML1	proteína 1 tipo alfa-2-macroglobulina
IPI00644409	GDA	guanina desaminasa
IPI00009650	LCN1	lipocalina 1
IPI00217467	HIST1H1E	histona H1.4
IPI00937064	ZAN	IPI:IPI00937064.1 REFSEQ:XP_002342720
IPI00396378	HNRNPA2B1	ribonucleoproteína nuclear heterogénea A2/B1
IPI00465261	ERAP2	aminopeptidasa 2 del retículo endoplasmático

ES 2 762 403 T3

IPI00021263	YWHAZ	14-3-3 proteína zeta/delta
IPI00011241	GPR39	receptor 39 acoplado a proteína G
IPI00786880	KIAA1783	similar a proteína KIAA1783
IPI00168479	APOA1 BP	proteína de unión a apolipoproteína A-I
IPI00031084	PSD2	dominio de pleckstrina y Sec7 que contiene 2 residuos
IPI00001593	PRCP	prolilcarboxipeptidasa (angiotensinasa C)
IPI00218343	TUBA1C	cadena alfa-1C de tubulina
IPI00021536	CALML5	proteína 5 similar a calmodulina
IPI00028091	ACTR3	homólogo de proteína 3 relacionada con actina ARP3 (levadura)
IPI00796366	MYL6	miosina, cadena ligera 6, álcali, músculo liso y no muscular
IPI00301058	VASP	fosfoproteína estimulada por vasodilatador
IPI00470573	ACTR2	homólogo de proteína 2 relacionada con actina ARP2 (levadura)
IPI00884078	RF-IP18	factor reumatoide RF-IP18
IPI00169383	PGK1	fosfoglicerato quinasa 1
IPI00299619	SLC35F1	miembro F1 de la familia 35 portadora de soluto
IPI00419916	ALPL	fosfatasa alcalina, hígado/hueso/riñón
IPI00218319	TPM3	tropomiosina 3
IPI00005118	HK3	hexoquinasa 3
IPI00418471	VIM	vimentina
IPI00218918	ANXA1	anexina A1
IPI00930073	KRT6C	IPI:IPI00930073.1  TREMBL:B2R853
IPI00299145	KRT6C	queratina, citoesquelética 6C tipo II
IPI00007858	MYH13	miosina, cadena pesada 13, músculo esquelético
IPI00297235	CCPG1	progresión del ciclo celular 1
IPI00927726	H-INV	proteína hipotética
IPI00009008	CACNA1D	canal de calcio, dependiente del voltaje, tipo L, subunidad alfa 1D
IPI00304554	LYPD5	dominio LY6/PLAUR que contiene 5 residuos
IPI00815807	ADCK2	dominio aarF que contiene quinasa 2
IPI00743335	MYO1C	miosina-1c
IPI00007393	APPBP2	proteína 2 de unión a la proteína precursora de amiloide beta (cola citoplasmática)
IPI00295976	ITGA2B	integrina, alfa 2b (glicoproteína plaquetaria IIb del complejo IIb/IIIa, antígeno CD41)
IPI00641706	TUBB6	tubulina, beta 6
IPI00060201	SYTL4	tipo sinaptotagmina 4
IPI00908634	AQP4	aquaporina 4
IPI00016786	CDC42	ciclo 42 de división celular (proteína de unión a GTP, 25 kDa)
IPI00033494	MYL12B	miosina, cadena ligera 12B, reguladora
IPI00409756	LOC100293553	proteína tipo L-Myc-2
IPI00015148	RAP1B	RAP1B, miembro de la familia de oncogenes RAS
IPI00027502	GP9	glicoproteína IX (plaqueta)
IPI00473014	DSTN	dextrina
IPI00022394	C1QC	componente 1 del complemento, subcomponente q, cadena C
IPI00290337	EPS8	sustrato 8 de la vía del receptor del factor de crecimiento epidérmico
IPI00018671	DUSP3	fosfatasa 3 de doble especificidad
IPI00478231	RHOA	familia de genes homólogos de ras, miembro A
IPI00220278	MYL9	miosina, cadena ligera 9, reguladora
IPI00419585	PPIA	peptidilprolil isomerasa A (ciclofilina A)
IPI00012011	CFL1	cofilina 1

Tabla 2 - Un subconjunto de biomarcadores de la Tabla 1 que se expresan significativamente de manera diferencial en todas las muestras de CRC en relación con una muestra de control

Número de Acceso	Símbolo del gen	Nombre del gen	Control de las veces que cambia vs CRC	Control de valores P vs CRC	Identificado en no. de CRC (n=12)	Identificado en no. de controles (n=10)	Detectado en plasma	Identificado en líneas celulares de cáncer	media de recuento espectral por CRC positivo	media de recuento espectral por control positivo
IP100892604	C4A/C4B	componente C4B del complemento (grupo sanguíneo Chido)	~	2,45E-06	10	0	no	si	11	na
IP100018206	GOT2	glutámico-oxaloacético transaminasa 2, mitocondrial (aspartato aminotransferasa 2)	~	4,93E-03	5	0	no	si	9	na
IP100027497	GPI	glucosa-6-fosfato isomerasa	49,74	5,44E-03	7	1	no	si	7	1
IP100643920	TKT	transcetolasa	24,13	1,71E-02	6	1	no	si	6	1
IP100337741	APEH	N-acilaminoacil-péptido hidrolasa	12,45	4,72E-04	9	2	no	si	6	2
IP100453473	HIST4H4 (incluye otros)	grupo de histonas 1, H4c	2,72	4,42E-02	9	4	no	si	5	2
IP100007797	FABP5	proteína 5 de unión a ácidos grasos (asociada a psoriasis)	~	3,48E-02	3	0	no	si	5	na
IP100012585	HEXB	hexosaminidasa B (polipéptido beta)	14,77	1,35E-02	6	1	no	si	5	2
IP100296215	EPCAM	molécula de adhesión celular epitelial	~	3,56E-02	3	0	no	si	4	na
IP100604590	NME1-NME2	lectura de NME1-NME2	~	1,34E-02	4	0	no	si	4	na
IP100022314	SOD2	superóxido dismutasa 2, mitocondrial	~	6,04E-03	5	0	no	si	4	na
IP100027107	TUFM	factor de alargamiento de traducción Tu, mitocondrial	2,02	7,75E-03	10	2	no	si	3	5
IP100010706	GSS	glutatión sintetasa	~	1,08E-02	4	0	no	si	3	na
IP100418169	ANXA2	anexina A2	~	4,47E-03	5	0	no	si	3	na
IP100303476	ATP5B	ATP sintasa, transporte de H <sup>+</sup> , complejo F1 mitocondrial, polipéptido beta	7,02	2,62E-02	6	1	no	si	3	1
IP100220362	HSPE1	proteína de choque térmico de 10 kDa (chaperonina 10)	~	9,57E-03	4	0	no	si	3	na
IP100220766	GLO1	glioxalasa I	~	5,24E-04	7	0	no	si	2	na
IP100003935	HIST2H2BE (incluye otros)	Grupo de histonas 2, H2be	9,25	3,83E-02	5	1	no	si	2	1

(continuación)

Número de Acceso	Símbolo del gen	Nombre del gen	Control de las veces que cambia vs CRC	Control de valores P vs CRC	Identificado en no. de CRC (n=12)	Identificado en no. de controles (n=10)	Detectado en plasma	Identificado en líneas celulares de cáncer	media de recuento espectral por CRC positivo	media de recuento espectral por control positivo
IP100032313	S100A4	proteína S100 de unión a calcio A4	~	1,40E-02	4	0	no	si	2	na
IP100013895	S100A11	proteína S100 de unión a calcio A11	~	3,55E-03	5	0	no	si	2	na
IP100106687	LXN	latexina	4,85	1,98E-02	7	2	no	si	2	2
IP100034280	DHRS11	miembro 11 de deshidrogenasa/reductasa (familia SDR)	7,23	6,98E-03	8	2	no	si	2	1
IP100008787	NAGLU	N-acetilglucosaminidasa, alfa	11,00	4,44E-03	7	1	no	si	2	1
IP100018768	TSN	translina	~	1,45E-02	4	0	no	si	2	na
IP100299155	PSMA4	subunidad del proteasoma (prosoma, macropaina) alfa tipo 4	~	1,48E-02	4	0	no	si	2	na
IP100029623	PSMA6	subunidad de proteasoma (prosoma, macropaina) alfa tipo-6	~	3,02E-02	3	0	no	si	2	na
IP100010271	RAC1	sustrato 1 de toxina botulínica C3 relacionada con ras (familia rho, proteína pequeña Rac1 de unión a GTP)	~	4,46E-02	3	0	no	si	2	na
IP100012007	AHCY	adenosilhomocisteinasa	~	1,68E-03	6	0	no	si	2	na
IP100385751	FUCA1	fucosidasa, alfa-L-1, tejido	14,06	2,67E-02	6	1	no	si	2	1
IP100017526	S100P	proteína S100 de unión a calcio P	~	3,57E-03	5	0	no	si	2	na
IP100028006	PSMB2	subunidad de proteasoma (prosoma, macropaina) beta tipo-2	~	3,77E-03	5	0	no	si	2	na
IP100793375	XPNPEP1	X-prolil aminopeptidasa (aminopeptidasa P) 1, soluble	~	3,35E-02	3	0	no	si	2	na
IP100554788	KRT18	queratina 18	~	3,79E-02	3	0	no	si	2	na
IP100019380	NCBP1	subunidad 1 de proteína de unión a la caperuza nuclear, 80 kDa	~	4,02E-02	3	0	no	si	2	na
IP100012989	MAN2B1	manosidasa, alfa, clase 2B, miembro 1	7,64	2,11E-02	7	2	no	si	1	1
IP100027463	S100A6	proteína S100 de unión a calcio A6	4,33	7,50E-03	10	3	no	si	1	1

(continuación)

Número de Acceso	Símbolo del gen	Nombre del gen	Control de las veces que cambia vs CRC	Control de valores P vs CRC	Identificado en no. de CRC (n=12)	Identificado en no. de controles (n=10)	Detectado en plasma	Identificado en líneas celulares de cáncer	media de recuento espectral por CRC positivo	media de recuento espectral por control positivo
IP100022774	VCP	proteína que contiene valosina	~	3,16E-02	3	0	no	sí	1	na
IP100300086	QPR1	quinolinato fosforibosiltransferasa	~	1,33E-02	4	0	no	sí	1	na
IP100472073	HLA-B	complejo mayor de histocompatibilidad, clase I, B	~	5,15E-03	5	0	no	sí	1	na
IP100453476	PGAM1	fosfoglicerato mutasa 1 (cerebro)	~	2,06E-02	4	0	no	sí	1	na
IP100020999	ENPP3	ectonucleótido pirofosfatasa/fosfodiesterasa 3	2,77	6,30E-04	10	2	no	no	11	14
IP100010304	SERP1NB10	inhibidor de la serpin peptidasa, clado B (ovoalbúmina), miembro 10	32,16	2,84E-02	6	1	no	no	10	4
IP100007244	MPO	mieloperoxidasa	22,73	5,30E-04	10	3	no	no	10	2
IP100877726	CKMT1A/CKMT1B	creatina quinasa, mitocondrial 1B	2,08	2,15E-03	11	4	no	no	7	7
IP100027409	PRTN3	proteinasa 3	93,49	1,90E-05	10	1	no	no	5	1
IP100027769	ELANE	elastasa, expresada en neutrófilos	23,48	1,47E-02	6	1	no	no	5	1
IP100004362	MORC1	dedo de zinc 1 tipo CW de la familia MORC	6,22	3,95E-02	5	1	no	no	4	1
IP100719280	UBB	ubiquitina B	3,45	1,49E-02	11	6	no	no	4	2
IP100026962	PLA2G2A	fosfolipasa A2, grupo IIA (plaquetas, líquido sinovial)	6,13	4,30E-03	7	1	no	no	3	1
IP100027466	CA4	anhidrasa carbónica IV	4,02	3,87E-03	10	2	no	no	3	4
IP100071703	GFM2	factor de alargamiento G, mitocondrial 2	4,03	2,72E-02	10	5	no	no	3	1
IP100219806	S100A7	proteína S100 de unión a calcio A7	3,15	4,75E-02	7	4	no	no	3	1
IP100827847	BPI	proteína que aumenta la permeabilidad bactericida	~	1,61E-02	4	0	no	no	3	na
IP100176125	COL6A5	colágeno, tipo VI, alfa 5	2,74	2,13E-02	8	2	no	no	3	2
IP100470355	LHX8	homeobox 8 de LIM	~	1,30E-02	4	0	no	no	3	na
IP100942117	CRISP3	proteína 3 secretora rica en cisteína	~	1,45E-02	4	0	no	no	2	na
IP100022246	AZU1	azurocidina 1	~	2,27E-05	8	0	no	no	2	na

(continuación)

Número de Acceso	Símbolo del gen	Nombre del gen	Control de las veces que cambia vs CRC	Control de valores P vs CRC	Identificado en no. de CRC (n=12)	Identificado en no. de controles (n=10)	Detectado en plasma	Identificado en líneas celulares de cáncer	media de recuento espectral por CRC positivo	media de recuento espectral por control positivo
IP100045512	HMCN1	hemicentina 1	8,30	6,80E-03	8	2	no	no	2	1
IP100300376	TGM3	transglutaminasa 3 (polipéptido E, proteína-glutamina-gamma-glutamiltransferasa)	~	3,67E-02	3	0	no	no	2	na
IP100640468	CDC42BPA	proteína quinasa alfa de unión a CDC42 (similar a DMPK)	1,88	4,92E-02	8	3	no	no	2	1
IP100028064	CTSG	cathepsina G	~	1,03E-02	4	0	no	no	2	na
IP100006988	RETN	resistina	~	3,74E-02	3	0	no	no	2	na
IP100024934	MUT	metilmalonil CoA mutasa	11,42	7,07E-03	8	2	no	no	2	1
IP100552983	ARMCX4	que contiene repetición armadillo, 4 ligado a X	12,86	1,65E-02	7	1	no	no	2	2
IP100217987	ITGAM	integrina alfa M (subunidad del receptor 3 del componente 3 del complemento)	~	1,59E-02	4	0	no	no	2	na
IP100165045	CACNA1E	canal de calcio, dependiente del voltaje, tipo R, subunidad alfa-1E	~	7,77E-04	7	0	no	no	1	na
IP100018363	TIAM2	invasión y metástasis 2 de linfoma de células T	18,20	4,58E-03	7	1	no	no	1	1
IP100217560	HIRA	homólogo A defectuoso en la regulación del ciclo celular de histona HIR ( <i>S. cerevisiae</i> )	~	8,36E-03	5	0	no	no	1	na
IP100294653	DOPEY2	miembro 2 de la familia dopey	13,89	2,46E-02	6	1	no	no	1	1
IP100004401	ITGB1BP3	proteína 3 de unión a integrina beta 1	~	3,38E-02	3	0	no	no	1	na
IP100300117	SCN7A	canal de sodio, controlado por voltaje, tipo VII, alfa	~	3,38E-02	3	0	no	no	1	na
IP100061114	RAB3C	Rab3C, miembro de la familia de oncogenes RAS	~	3,41E-02	3	0	no	no	1	na
IP100168442	C9orf79	marco de lectura abierto 79 del cromosoma 9	~	3,89E-02	3	0	no	no	1	na
IP100385215	NFATC4	factor nuclear de células T activadas, dependiente de calcineurina 4	~	3,95E-02	3	0	no	no	1	na
IP100024467	UGGT2	UDP-glucosa glicoproteína glucosiltransferasa 2	~	4,02E-02	3	0	no	no	1	na

(continuación)

Número de Acceso	Símbolo del gen	Nombre del gen	Control de las veces que cambia vs CRC	Control de valores P vs CRC	Identificado en no. de CRC (n=12)	Identificado en no. de controles (n=10)	Detectado en plasma	Identificado en líneas celulares de cáncer	media de recuento espectral por CRC positivo	media de recuento espectral por control positivo
IP100297056	CRNN	cornulina	~	4,04E-02	3	0	no	no	1	na
IP100914663	KCP	proteína similar a quielina/cordina	~	9,68E-03	4	0	no	no	1	na
IP100418592	CD1E	molécula CD1e	~	1,46E-02	4	0	no	no	1	na
IP100642206	CCDC18	dominio de hélice superenrollada que contiene 18 residuos	~	4,58E-03	5	0	no	no	1	na
IP100514090	LTA4H	leucotrieno A-4 hidrolasa	~	4,50E-02	3	0	no	no	1	na
IP100745872	ALB	albúmina	4,71	1,41E-02	12	10	sí	sí	576	124
IP100478003	A2M	alfa-2-macroglobulina	34,08	9,01E-06	12	6	sí	sí	141	11
IP100783987	C3	componente 3 del complemento	43,72	2,21E-04	12	6	sí	sí	119	5
IP100654755	HBB	hemoglobina, beta	9,87	5,89E-05	12	7	sí	sí	108	14
IP100022463	TF	transferrina	1713,68	7,84E-06	11	1	sí	sí	96	1
IP100410714	HBA1/HBA2	hemoglobina, alfa 1	44,44	1,77E-04	12	5	sí	sí	68	2
IP100298860	LTF	lactotransferrina	59,22	3,14E-05	10	2	sí	sí	43	6
IP100017601	CP	ceruloplasmina (ferroxidasa)	94,27	1,28E-04	11	3	sí	sí	41	2
IP100465436	CAT	catalasa	77,03	1,66E-03	10	3	sí	sí	24	1
IP100555812	GC	componente específico del grupo (proteína de unión a la vitamina D)	~	3,43E-02	3	0	sí	sí	20	na
IP100032179	SERPINC1	inhibidor de la serpin peptidasa, clado C (antitrombina), miembro 1	2,17	3,20E-02	12	10	sí	sí	20	11
IP100219713	FGG	cadena gamma de fibrinógeno	~	1,11E-03	6	0	sí	sí	20	na
IP100007047	S100A8	proteína S100 de unión a calcio A8	3,87	1,12E-06	12	9	sí	sí	17	5
IP100375676	FTL	ferritina, polipéptido ligero	2,51	2,92E-02	12	10	sí	sí	12	5
IP100021439	ACTB	actina, beta	2,44	1,07E-02	10	3	sí	sí	12	11
IP100022418	FN1	fibronectina 1	17,62	1,31E-02	10	4	sí	sí	11	2

(continuación)

Número de Acceso	Símbolo del gen	Nombre del gen	Control de las veces que cambia vs CRC	Control de valores P vs CRC	Identificado en no. de CRC (n=12)	Identificado en no. de controles (n=10)	Detectado en plasma	Identificado en líneas celulares de cáncer	media de recuento espectral por CRC positivo	media de recuento espectral por control positivo
IP100005721	DEFA1 (incluye otros)	defensina, alfa 1	4,16	4,47E-05	12	9	si	si	8	2
IP100291866	SERPING1	inhibidor de la serpina peptidasa, clado G (inhibidor de C1), miembro 1	~	2,88E-04	7	0	si	si	8	na
IP100022420	RBP4	proteína 4 de unión a retinol, plasma	4,47	1,20E-02	8	3	si	si	8	6
IP100027350	PRDX2	peroxirredoxina 2	36,98	4,57E-03	7	1	si	si	6	1
IP100021885	FGA	cadena alfa de fibrinógeno	21,98	3,30E-02	6	2	si	si	6	1
IP100029863	SERPINF2	inhibidor de la serpina peptidasa, clado F (antiplasmina alfa-2, factor derivado del epitelio pigmentario), miembro 2	8,61	8,05E-04	9	1	si	si	4	8
IP100218414	CA2	anhidrasa carbónica II	2,87	1,12E-02	7	1	si	si	4	10
IP100020091	ORM1/ORM2	orosomucoide 1	4,11	2,56E-02	9	3	si	si	4	3
IP100217966	LDHA	lactato deshidrogenasa A	~	1,85E-03	6	0	si	si	4	na
IP100298971	VTN	vitronectina	~	1,15E-03	6	0	si	si	4	na
IP100215894	KNG1	quininógeno 1	~	3,33E-04	7	0	si	si	4	na
IP100023006	ACTC1	actina, alfa, músculo cardíaco 1	3,17	2,33E-03	10	3	si	si	4	2
IP100022417	LRG1	alfa-2-glicoproteína 1 rica en leucina	31,63	4,06E-03	7	1	si	si	3	1
IP100023728	GGH	gamma-glutamyl hidrolasa (conjugasa, folilipoligammaglutamil hidrolasa)	6,57	1,73E-02	7	2	si	si	3	3
IP100465248	ENO1	enolasa 1, (alfa)	5,38	3,17E-02	7	2	si	si	3	1
IP100216691	PFN1	profilina 1	~	4,15E-02	3	0	si	si	3	na
IP100292946	SERPINA7	inhibidor de la serpina peptidasa, clado A, (antiproteinasa alfa-1, antitripsina), miembro 7	~	6,51E-04	7	0	si	si	3	na
IP100022426	AMBIP	precursor de alfa-1-microglobulina/bikunina	~	1,75E-03	6	0	si	si	3	na
IP100021405	LMNA	lamina A/C	~	1,47E-02	4	0	si	si	3	na
IP100006662	APOD	apolipoproteína D	3,38	4,74E-03	8	1	si	si	2	4

(continuación)

Número de Acceso	Símbolo del gen	Nombre del gen	Control de las veces que cambia vs CRC	Control de valores P vs CRC	Identificado en no. de CRC (n=12)	Identificado en no. de controles (n=10)	Detectado en plasma	Identificado en líneas celulares de cáncer	media de recuento espectral por CRC positivo	media de recuento espectral por control positivo
IP100003515	TRIP11	interactuador 11 del receptor de la hormona tiroidea	7,65	3,20E-03	9	1	si	si	2	2
IP100021727	C4BPA	proteína de unión al componente 4 del complemento, alfa	~	1,23E-02	4	0	si	si	2	na
IP100010779	TPM4	tropomiosina 4	~	4,12E-02	3	0	si	si	2	na
IP100302592	FLNA	filamina A, alfa	~	2,45E-03	6	0	si	si	1	na
IP100641737	HP	haptoglobina	18,82	2,89E-04	11	6	si	no	89	6
IP100022488	HPX	hemopexina	~	4,42E-05	8	0	si	no	23	na
IP100473011	HBD	hemoglobina, delta	~	6,31E-05	8	0	si	no	15	na
IP100298497	FGB	cadena beta de fibronectina	~	1,18E-02	4	0	si	no	11	na
IP100027462	S100A9	proteína S100 de unión a calcio A9	4,41	1,82E-05	12	8	si	no	10	3
IP100032291	C5	componente 5 del complemento	33,44	9,34E-03	8	2	si	no	8	1
IP100031036	SLC26A3	familia 26 portadora de soluto, miembro 3	64,79	1,18E-02	7	1	si	no	6	1
IP100022395	C9	componente 9 del complemento	12,69	4,19E-04	11	5	si	no	6	1
IP100022391	APCS	componente amiloide P, suero	5,75	2,53E-03	10	3	si	no	6	5
IP100022895	A1BG	glicoproteína alfa-1-B	~	1,31E-02	4	0	si	no	6	na
IP100887739	LOC100133511	complemento tipo C3	~	4,38E-03	5	0	si	no	5	na
IP100218192	ITIH4	inhibidor H4 inter-alfa (globulina) (glicoproteína sensible a calcireina en plasma)	~	3,64E-02	3	0	si	no	5	na
IP100011252	C8A	componente 8 del complemento, polipéptido alfa	~	3,26E-02	3	0	si	no	3	na
IP100292530	ITIH1	inhibidor H1 inter-alfa (globulina)	~	1,05E-02	4	0	si	no	1	na

ES 2 762 403 T3

Tabla 3 - Un subconjunto de biomarcadores de la Tabla 1 que se expresan diferencialmente en muestras CRC negativas para FIT en relación con una muestra de control

Número de acceso	Símbolo del gen	Nombre del gen
IPI00022426	AMBP	precursor de alfa-1-microglobulina/bikunina
IPI00022246	AZU1	azurocidina 1
IPI00827847	BPI	proteína que aumenta la permeabilidad bactericida
IPI00892604	C4A/C4B	componente C4B del complemento (grupo sanguíneo Chido)
IPI00021727	C4BPA	proteína de unión al componente 4 del complemento, alfa
IPI00168442	C9orf79	marco de lectura abierto 79 del cromosoma 9
IPI00165045	CACNA1E	canal de calcio, dependiente del voltaje, tipo R, subunidad alfa 1E
IPI00642206	CCDC18	dominio de hélice superenrollada que contiene 18 residuos
IPI00297056	CRNN	cornulina
IPI00296215	EPCAM	molécula de adhesión celular epitelial
IPI00219713	FGG	cadena gamma de fibrinógeno
IPI00302592	FLNA	filamina A, alfa
IPI00220766	GLO1	glioxalasa I
IPI00018206	GOT2	glutámico-oxaloacético transaminasa 2, mitocondrial (aspartato aminotransferasa 2)
IPI00010706	GSS	glutatión sintetasa
IPI00217560	HIRA	homólogo A defectuoso en la regulación del ciclo celular de histona HIR (S. cerevisiae)
IPI00472073	HLA-B	complejo mayor de histocompatibilidad, clase I, B
IPI00022488	HPX	hemopexina
IPI00217987	ITGAM	integrina alfa M (subunidad del receptor 3 del componente 3 del complemento)
IPI00215894	KNG1	quininógeno 1
IPI00554788	KRT18	queratina 18
IPI00217966	LDHA	lactato deshidrogenasa A
IPI00470355	LHX8	homeobox 8 de LIM
IPI00887739	LOC100133511	complemento tipo C3
IPI00019380	NCBP1	subunidad 1 de la proteína de unión a la caperuza nuclear, 80 kDa
IPI00604590	NME1-NME2	lectura de NME1-NME2
IPI00299155	PSMA4	subunidad del proteasoma (prosome, macropaína), tipo alfa, 4
IPI00300117	SCN7A	canal de sodio, controlado por voltaje, tipo VII, alfa
IPI00292946	SERPINA7	inhibidor de la serpina peptidasa, clado A (antiproteinasas alfa-1, antitripsina), miembro 7
IPI00291866	SERPING1	inhibidor de la serpina peptidasa, clado G (inhibidor de C1), miembro 1
IPI00024467	UGGT2	UDP-glucosa glicoproteína glucosiltransferasa 2
IPI00298971	VTN	vitronectina
IPI00478003	A2M	alfa-2-macroglobulina
IPI00022391	APCS	componente amiloide P, suero
IPI00337741	APEH	N-acilaminoacil-péptido hidrolasa
IPI00552983	ARMCX4	que contiene repetición armadillo, 4 ligado a X
IPI00783987	C3	componente 3 del complemento
IPI00032291	C5	componente 5 del complemento
IPI00022395	C9	componente 9 del complemento
IPI00027466	CA4	anhidrasa carbónica IV
IPI00465436	CAT	catalasa
IPI00176125	COL6A5	colágeno, tipo VI, alfa 5
IPI00017601	CP	ceruloplasmina (ferroxidasa)
IPI00005721	DEFA1(incluye otros)	defensina, alfa 1
IPI00294653	DOPEY2	miembro 2 de la familia dopey
IPI00020999	ENPP3	ectonucleótido pirofosfatasa/fosfodiesterasa 3
IPI00022418	FN1	fibronectina 1
IPI00375676	FTL	ferritina, polipéptido ligero
IPI00385751	FUCA1	fucosidasa, alfa-L- 1, tejido

ES 2 762 403 T3

IPI00071703	GFM2	factor de alargamiento G, mitocondrial 2
IPI00023728	GGH	gamma-glutamil hidrolasa (conjugasa, folilpoligammaglutamil hidrolasa)
IPI00027497	GPI	glucosa-6-fosfato isomerasa
IPI00410714	HBA1/HBA2	hemoglobina alfa 1
IPI00654755	HBB	hemoglobina, beta
IPI00012585	HEXB	hexosaminidasa B (polipéptido beta)
IPI00045512	HMCN1	hemicentina 1
IPI00022417	LRG1	alfa-2-glicoproteína 1 rica en leucina
IPI00298860	LTF	lactotransferrina
IPI00106687	LXN	latexina
IPI00012989	MAN2B1	manosidasa, alfa, clase 2B, miembro 1
IPI00007244	MPO	mieloperoxidasa
IPI00024934	MUT	metilmalonil CoA mutasa
IPI00026962	PLA2G2A	fosfolipasa A2, grupo IIA (plaquetas, líquido sinovial)
IPI00027350	PRDX2	peroxirredoxina 2
IPI00027409	PRTN3	proteínasa 3
IPI00022420	RBP4	proteína 4 de unión a retinol, plasma
IPI00027463	S100A6	proteína S100 de unión a calcio A6
IPI00219806	S100A7	proteína S100 de unión a calcio A7
IPI00007047	S100A8	proteína S100 de unión a calcio A8
IPI00027462	S100A9	proteína S100 de unión a calcio A9
IPI00010304	SERPINB10	inhibidor de la serpina peptidasa, clado B (ovoalbúmina), miembro 10
IPI00029863	SERPINF2	inhibidor de la serpina peptidasa, clado F (antiplasmina alfa-2, factor derivado del epitelio pigmentario), miembro 2
IPI00031036	SLC26A3	familia 26 portadora de soluto, miembro 3
IPI00022463	TF	transferrina
IPI00018363	TIAM2	Invasión y metástasis de linfoma de células T 2
IPI00643920	TKT	transcetolasa
IPI00003515	TRIP11	interactuador 11 del receptor de la hormona tiroidea
IPI00027107	TUFM	factor de alargamiento de traducción Tu, mitocondrial
IPI00181135	BCAT2	aminotransferasa de aminoácidos de cadena ramificada
IPI00024701	COL4A3BP	proteína de unión a colágeno, tipo IV, alfa 3 (antígeno de Goodpasture)
IPI00032134	SERPINB8	inhibidor de la serpina peptidasa, clado B (ovoalbúmina), miembro 8
IPI00333141	SPNS3	homólogo 3 de spinster (Drosophila)
IPI00937735	ACADVL	acil-CoA deshidrogenasa, cadena muy larga
IPI00009268	ACY1	no mapeado por la invención
IPI00166331	ANKRD35	Dominio 35 de repetición de anquirina
IPI00299635	BIRC6	repetición de IAP baculoviral que contiene 6 residuos
IPI00219575	BLMH	bleomicina hidrolasa
IPI00026240	BST1	antígeno 1 de célula estromal de médula ósea
IPI00929313	C13orf40	marco de lectura abierto 40 del cromosoma 13
IPI00027983	CDA	citidina desaminasa
IPI00852865	CHIT1	quitinasa 1 (quitotriosidasa)
IPI00022810	CTSC	catepsina C
IPI00299150	CTSS	catepsina S
IPI00006024	DOCK4	dedicador de citoquinesis 4
IPI00016862	GSR	glutaciona reductasa
IPI00941901	HERC1	dominio hect (homólogo al terminal carboxilo E6-AP (UBE3A)) y el dominio 1 similar a RCC1 (CHC1) (RLD)
IPI00005826	HERC2	dominio hect y RLD 2
IPI00746667	HLA-DRB5	complejo mayor de histocompatibilidad, clase II, DR beta 5
IPI00027223	IDH1	isocitrato deshidrogenasa 1 (NADP+), soluble
IPI00305461	ITIH2	inhibidor H2 inter-alfa (globulina)
IPI00748360	KIAA1797	KIAA1797
IPI00019038	LYZ	lisozima

ES 2 762 403 T3

IPI00303335	NEB	nebulina
IPI00065454	NEK10	quinasa 10 relacionada con NIMA (nunca en el gen a de mitosis)
IPI00257882	PEPD	peptidasa D
IPI00003590	QSOX1	quiescina Q6 sulfhidril oxidasa 1 1
IPI00939604	RNASET2	ribonucleasa T2
IPI00553177	SERPINA1	inhibidor de la serpina peptidasa, clado A (alfa-1 antiproteinasa, antitripsina), miembro 1
IPI00550991	SERPINA3	inhibidor de la serpina peptidasa, clado A (alfa-1 antiproteinasa, antitripsina), miembro 3
IPI00022204	SERPINB3	inhibidor de la serpina peptidasa, clado B (ovoalbúmina), miembro 3
IPI00307733	SETD2	dominio SET que contiene 2 residuos
IPI00218013	SGOL2	similar a shugoshina 2
IPI00010949	SIAE	ácido siálico acetiltransferasa
IPI00386444	SYNE1	envoltura nuclear 1, que contiene repetición de espectrina
IPI00744692	TALDO1	transaldolasa 1
IPI00375371	TAS2R42	receptor del gusto tipo 2 miembro 42
IPI00465028	TPI1	triosafosfato isomerasa 1
IPI00291175	VCL	vinculina
IPI00029647	ZG16	proteína 16 de gránulos de zimógeno

Tabla 4 - Un subconjunto de biomarcadores de la Tabla 1 que se encontró que se expresan solamente en muestras de CRC y no en muestras de control

Número de acceso	Símbolo del gen	Nombre del gen
IPI00892604	C4A/C4B	componente C4B del complemento (grupo sanguíneo Chido)
IPI00018206	GOT2	glutámico-oxaloacético transaminasa 2, mitocondrial (aspartato aminotransferasa 2)
IPI00007797	FABP5	proteína 5 de unión a ácidos grasos (asociada a psoriasis)
IPI00296215	EPCAM	molécula de adhesión celular epitelial
IPI00604590	NME1-NME2	lectura de NME1-NME2
IPI00022314	SOD2	superóxido dismutasa 2, mitocondrial
IPI00010706	GSS	glutathione sintetasa
IPI00418169	ANXA2	anexina A2
IPI00220362	HSPE1	proteína de choque térmico de 10 kDa (chaperonina 10)
IPI00220766	GLO1	glioxalasa I
IPI00032313	S100A4	proteína S100 de unión a calcio A4
IPI00013895	S100A11	proteína S100 de unión a calcio A11
IPI00018768	TSN	translina
IPI00299155	PSMA4	subunidad del proteasoma (prosome, macropaina), tipo alfa, 4
IPI00029623	PSMA6	subunidad de proteasoma (prosome, macropaina) alfa tipo-6
IPI00010271	RAC1	sustrato 1 de toxina botulínica C3 relacionada con ras (familia rho, proteína pequeña Rac1 de unión a GTP)
IPI00012007	AHCY	adenosilhomocisteinasa
IPI00017526	S100P	proteína S100 de unión a calcio P
IPI00028006	PSMB2	subunidad de proteasoma (prosome, macropaina) beta tipo-2
IPI00793375	XPNPEP1	X-prolil aminopeptidasa (aminopeptidasa P) 1, soluble
IPI00554788	KRT18	queratina 18
IPI00019380	NCBP1	subunidad 1 de la proteína de unión a la caperuza nuclear, 80 kDa
IPI00022774	VCP	proteína que contiene valosina
IPI00300086	QPRT	quinolinato fosforibosiltransferasa
IPI00472073	HLA-B	complejo mayor de histocompatibilidad, clase I, B
IPI00453476	PGAM1	fosfoglicerato mutasa 1 (cerebro)
IPI00827847	BPI	proteína que aumenta la permeabilidad bactericida
IPI00470355	LHX8	homeobox 8 de LIM
IPI00942117	CRISP3	proteína 3 secretora rica en cisteína
IPI00022246	AZU1	azurocidina 1

ES 2 762 403 T3

IPI00300376	TGM3	transglutaminasa 3 (polipéptido E, proteína-glutamina-gamma-glutamiltransferasa)
IPI00028064	CTSG	catepsina G
IPI00006988	RETN	resistina
IPI00217987	ITGAM	integrina alfa M (subunidad del receptor 3 del componente 3 del complemento)
IPI00165045	CACNA1E	canal de calcio, dependiente del voltaje, tipo R, subunidad alfa 1E
IPI00217560	HIRA	homólogo A defectuoso en la regulación del ciclo celular de histona HIR (S. cerevisiae)
IPI00004401	ITGB1 BP3	proteína 3 de unión a integrina beta 1
IPI00300117	SCN7A	canal de sodio, controlado por voltaje, tipo VII, alfa
IPI00061114	RAB3C	RAB3C, miembro de la familia de oncogenes RAS
IPI00168442	C9orf79	marco de lectura abierto 79 del cromosoma 9
IPI00385215	NFATC4	factor nuclear de células T activadas, dependiente de calcineurina 4
IPI00024467	UGGT2	UDP-glucosa glicoproteína glucosiltransferasa 2
IPI00297056	CRNN	cornulina
IPI00914663	KCP	proteína similar a quelina/cordina
IPI00418592	CD1E	molécula CD1e
IPI00642206	CCDC18	dominio de hélice superenrollada que contiene 18 residuos
IPI00514090	LTA4H	leucotrieno A4 hidrolasa
IPI00555812	GC	componente específico del grupo (proteína de unión a la vitamina D)
IPI00219713	FGG	cadena gamma de fibrinógeno
IPI00291866	SERPING1	inhibidor de la serpina peptidasa, clado G (inhibidor de C1), miembro 1
IPI00217966	LDHA	lactato deshidrogenasa A
IPI00298971	VTN	vitronectina
IPI00215894	KNG1	quininógeno 1
IPI00216691	PFN1	profilina 1
IPI00292946	SERPINA7	inhibidor de la serpina peptidasa, clado A (antiproteinasa alfa-1, antitripsina), miembro 7
IPI00022426	AMBP	precursor de alfa-1-microglobulina/bikunina
IPI00021405	LMNA	lamina A/C
IPI00021727	C4BPA	proteína de unión al componente 4 del complemento, alfa
IP100010779	TPM4	tropomiosina 4
IPI00302592	FLNA	filamina A, alfa
IPI00022488	HPX	hemopexina
IPI00473011	HBD	hemoglobina, delta
IPI00298497	FGB	cadena beta de fibrinógeno
IPI00022895	A1BG	glicoproteína alfa-1-B
IPI00887739	LOC100133511	complemento tipo C3
IPI00218192	ITIH4	inhibidor H4 inter-alfa (globulina) (glicoproteína sensible a caliceína en plasma)
IPI00011252	C8A	componente 8 del complemento, polipéptido alfa
IPI00292530	ITIH1	inhibidor H1 inter-alfa (globulina)
IPI00016255	PLBD1	proteína hipotética LOC79887
IPI00477505	ANKRD28	isoforma 1 de la subunidad A de repetición de anquirina reguladora de la proteína serina/treonina fosfatasa 6
IPI00032293	CST3	cistatina-C
IPI00293867	DDT	D-dopacromo descarboxilasa
IPI00059242	SYAP1	proteína 1 asociada a sinapsis
IPI00219622	PSMA2	subunidad alfa tipo-2 del proteasoma
IPI00221222	SUB1	homólogo de SUB1 (S. cerevisiae)
IPI00022791	MFAP3	glicoproteína 3 asociada a microfibrillas
IPI00011229	CTSD	catepsina D
IPI00025019	PSMB1	Subunidad de proteasoma (prosoma, macropaina), tipo beta, 1
IPI00479306	PSMB5	subunidad del proteasoma (prosoma, macropaina), tipo beta, 5
IPI00234793	KCTD15	ADNc FLJ61112, muy similar a la proteína que contiene el dominio BTB/POZ KCTD15

ES 2 762 403 T3

IPI00010796	P4HB	prolil-4-hidroxilasa, polipéptido beta
IPI00927606	GPX1	glutaciona peroxidasa 1
IPI00783625	SERPINB5	inhibidor de la serpina peptidasa, clado B (ovoalbúmina), miembro 5
IPI00024701	COL4A3BP	proteína de unión a colágeno, tipo IV, alfa 3 (antígeno de Goodpasture)
IPI00000811	PSMB6	subunidad del proteasoma (prosoma, macropaina), tipo beta, 6
IPI00021298	KRT20	queratina 20
IPI00025084	CAPNS1	subunidad pequeña 1 de calpaína
IPI00024919	PRDX3	peroxirredoxina 3
IPI00059930	NACC2	miembro 2 de la familia NACC, que contiene el dominio BEN y BTB (POZ)
IPI00003817	ARHGDI B	inhibidor 2 de disociación de Rho GDP
IPI00293276	MIF	factor inhibidor de migración de macrófagos
IPI00514622	RANBP6	proteína 6 de unión a Ran
IPI00333141	SPNS3	isoforma 1 del homólogo 3 de la proteína spinster
IPI00922181	MCM2	componente 2 del complejo de mantenimiento de minicromosomas
IPI00031708	FAH	fumarilacetoacetasa
IPI00003865	HSPA8	proteína 8 de choque térmico de 70 kDa
IPI00299024	BASP1	proteína 1 señal, unida a membrana, abundante en el cerebro
IPI00181135	BCAT2	aminotransferasa de aminoácidos de cadena ramificada
IPI00219365	MSN	moesina
IPI00032134	SERPINB8	serpina B8
IPI00853547	G6PD	glucosa-6-fosfato deshidrogenasa
IPI00478758	C10orf119	isoforma 1 de la proteína UPF0557 C10orf119
IPI00873020	PSAP	prosaposina
IPI00000875	EEF1G	factor 1 gamma de alargamiento de traducción eucariota
IPI00014398	FHL1	cuatro y medio dominios 1 de LIM
IPI00301395	CPVL	carboxipeptidasa, tipo vitelogénica
IPI00790262	TTLL3	familia tipo tirosina ligasa de tubulina, miembro 3
IPI00942608	proteína de 26 kDa	IPI:IPI00942608.1 ENSEMBL:ENSP0000
IPI00023038	PRB1/PRB2	subfamilia 2 de la proteína BstNI rica en prolina
IPI00001895	PCDH8	protocadherina 8
IPI00419215	A2ML1	proteína 1 tipo alfa-2-macroglobulina
IPI00644409	GDA	guanina desaminasa
IPI00009650	LCN1	lipocalina 1
IPI00217467	HIST1H1E	histona H1.4
IPI00937064	ZAN	IPI:IPI00937064.1 REFSEQ:XP_002342720
IPI00396378	HNRNPA2B1	ribonucleoproteína nuclear heterogénea A2/B1
IPI00465261	ERAP2	aminopeptidasa 2 del retículo endoplasmático
IPI00021263	YWHAZ	14-3-3 proteína zeta/delta
IPI00011241	GPR39	receptor 39 acoplado a proteína G
IPI00786880	KIAA1783	similar a proteína KIAA1783
IPI00168479	APOA1 BP	proteína de unión a apolipoproteína A-I
IPI00031084	PSD2	dominio de pleckstrina y Sec7 que contiene 2 residuos
IPI00001593	PRCP	prolilcarboxipeptidasa (angiotensinasa C)
IPI00218343	TUBA1C	cadena alfa-1C de tubulina
IPI00021536	CALML5	proteína 5 similar a calmodulina
IPI00028091	ACTR3	homólogo de proteína 3 relacionada con actina ARP3 (levadura)
IPI00796366	MYL6	miosina, cadena ligera 6, álcali, músculo liso y no muscular
IPI00003817	ARHGDI B	inhibidor 2 de disociación de Rho GDP
IPI00293276	MIF	factor inhibidor de migración de macrófagos
IPI00514622	RANBP6	proteína 6 de unión a Ran
IPI00333141	SPNS3	isoforma 1 del homólogo 3 de la proteína spinster
IPI00922181	MCM2	componente 2 del complejo de mantenimiento de minicromosomas
IPI00031708	FAH	fumarilacetoacetasa
IPI00003865	HSPA8	proteína 8 de choque térmico de 70 kDa
IPI00299024	BASP1	proteína 1 señal, unida a membrana, abundante en el cerebro

ES 2 762 403 T3

IPI00181135	BCAT2	aminotransferasa de aminoácidos de cadena ramificada
IPI00219365	MSN	moesina
IPI00032134	SERPINB8	serpina B8
IPI00853547	G6PD	glucosa-6-fosfato deshidrogenasa
IPI00478758	C10orf119	isoforma 1 de la proteína UPF0557 C10orf119
IPI00873020	PSAP	prosaposina
IPI00000875	EEF1G	factor 1 gamma de alargamiento de traducción eucariota
IPI00014398	FHL1	cuatro y medio dominios 1 de LIM
IPI00301395	CPVL	carboxipeptidasa, tipo vitelogénica
IPI00790262	TTLL3	familia tipo tirosina ligasa de tubulina, miembro 3
IPI00942608	proteína de 26 kDa	IPI:IPI00942608.1 ENSEMBL:ENSP0000
IPI00023038	PRB1/PRB2	subfamilia 2 de la proteína BstNI rica en prolina
IPI00001895	PCDH8	protocadherina 8
IPI00419215	A2ML1	proteína 1 tipo alfa-2-macroglobulina
IPI00644409	GDA	guanina desaminasa
IPI00009650	LCN1	lipocalina 1
IPI00217467	HIST1H1E	histona H1.4
IPI00937064	ZAN	IPI:IPI00937064.1 REFSEQ:XP_002342720
IPI00396378	HNRNPA2B1	ribonucleoproteína nuclear heterogénea A2/B1
IPI00465261	ERAP2	aminopeptidasa 2 del retículo endoplasmático
IPI00021263	YWHAZ	14-3-3 proteína zeta/delta
IPI00011241	GPR39	receptor 39 acoplado a proteína G
IPI00786880	KIAA1783	similar a proteína KIAA1783
IPI00168479	APOA1 BP	proteína de unión a apolipoproteína A-I
IPI00031084	PSD2	dominio de pleckstrina y Sec7 que contiene 2 residuos
IPI00001593	PRCP	prolilcarboxipeptidasa (angiotensinasa C)
IPI00218343	TUBA1C	cadena alfa-1C de tubulina
IPI00021536	CALML5	proteína 5 similar a calmodulina
IPI00028091	ACTR3	homólogo de proteína 3 relacionada con actina ARP3 (levadura)
IPI00796366	MYL6	miosina, cadena ligera 6, álcali, músculo liso y no muscular
IPI00301058	VASP	fosfoproteína estimulada por vasodilatador
IPI00470573	ACTR2	homólogo de proteína 2 relacionada con actina ARP2 (levadura)
IPI00884078	RF-IP18	factor reumatoide RF-IP18
IPI00169383	PGK1	fosfoglicerato quinasa 1
IPI00299619	SLC35F1	miembro F1 de la familia 35 portadora de soluto
IPI00419916	ALPL	fosfatasa alcalina, hígado/hueso/riñón
IPI00218319	TPM3	tropomiosina 3
IPI00005118	HK3	hexoquinasa 3
IPI00418471	VIM	vimentina
IPI00218918	ANXA1	anexina A1
IPI00930073	KRT6C	IPI:IPI00930073.1 TREMBL:B2R853
IPI00299145	KRT6C	queratina, citoesquelética 6C tipo II
IPI00007858	MYH13	miosina, cadena pesada 13, músculo esquelético
IPI00297235	CCPG1	progresión del ciclo celular 1
IPI00927726	H-INV	proteína hipotética
IPI00009008	CACNA1D	canal de calcio, dependiente del voltaje, tipo L, subunidad alfa 1D
IPI00304554	LYPD5	dominio LY6/PLAUR que contiene 5 residuos
IPI00815807	ADCK2	dominio aarF que contiene quinasa 2
IPI00743335	MYO1C	miosina-1c
IPI00007393	APPBP2	proteína 2 de unión a la proteína precursora de amiloide beta (cola citoplasmática)
IPI00295976	ITGA2B	integrina, alfa 2b (glicoproteína plaquetaria IIb del complejo IIb/IIIa, antígeno CD41)
IPI00641706	TUBB6	tubulina, beta 6
IPI00060201	SYTL4	tipo sinaptotagmina 4

ES 2 762 403 T3

IPI00908634	AQP4	aquaporina 4
IPI00016786	CDC42	ciclo 42 de división celular (proteína de unión a GTP, 25 kDa)
IPI00033494	MYL12B	miosina, cadena ligera 12B, reguladora
IPI00409756	LOC100293553	proteína tipo L-Myc-2
IPI00015148	RAP1B	RAP1B, miembro de la familia de oncogenes RAS
IPI00027502	GP9	glicoproteína IX (plaqueta)
IPI00473014	DSTN	dextrina
IPI00022394	C1QC	componente 1 del complemento, subcomponente q, cadena C
IPI00290337	EPS8	sustrato 8 de la vía del receptor del factor de crecimiento epidérmico
IPI00018671	DUSP3	fosfatasa 3 de doble especificidad
IPI00478231	RHOA	familia de genes homólogos de ras, miembro A
IPI00220278	MYL9	miosina, cadena ligera 9, reguladora
IPI00419585	PPIA	peptidilprolil isomerasa A (ciclofilina A)
IPI00012011	CFL1	cofilina 1

Tabla 5: Un subconjunto de biomarcadores de la tabla 1 que se identificaron en un conjunto independiente (verificación) de muestras de heces

Número de acceso	Símbolo del gen	Nombre del gen
IPI00022895	A1BG	glicoproteína alfa-1-B
IPI00478003	A2M	alfa-2-macroglobulina
IPI00021439	ACTB	actina, beta
IPI00023006	ACTC1	actina, alfa, músculo cardíaco 1
IPI00745872	ALB	albúmina
IPI00022391	APCS	componente amiloide P, suero
IPI00337741	APEH	N-acilaminoacil-péptido hidrolasa
IPI00006662	APOD	apolipoproteína D
IPI00022246	AZU1	azurocidina 1
IPI00783987	C3	componente 3 del complemento
IPI00892604	C4A/C4B	componente C4B del complemento (grupo sanguíneo Chido)
IPI00032291	C5	componente 5 del complemento
IPI00218414	CA2	anhidrasa carbónica II
IPI00027466	CA4	anhidrasa carbónica IV
IPI00465436	CAT	catalasa
IPI00017601	CP	ceruloplasmina (ferroxidasa)
IPI00028064	CTSG	catepsina G
IPI00005721	DEFA1 (incluye otros)	defensina, alfa 1
IPI00027769	ELANE	elastasa, expresada en neutrófilos
IPI00465248	ENO1	enolasa 1, (alfa)
IPI00302592	FLNA	filamina A, alfa
IPI00022418	FN1	fibronectina 1
IPI00375676	FTL	ferritina, polipéptido ligero
IPI00385751	FUCA1	fucosidasa, alfa-L- 1, tejido
IPI00023728	GGH	gamma-glutamil hidrolasa (conjugasa, folilpoligammaglutamil hidrolasa)
IPI00018206	GOT2	glutámico-oxaloacético transaminasa 2, mitocondrial (aspartato aminotransferasa 2)
IPI00027497	GPI	glucosa-6-fosfato isomerasa
IPI00010706	GSS	glutación sintetasa
IPI00410714	HBA1/HBA2	hemoglobina alfa 1
IPI00654755	HBB	hemoglobina, beta
IPI00473011	HBD	hemoglobina, delta
IPI00012585	HEXB	hexosaminidasa B (polipéptido beta)
IPI00453473	HIST4H4 (incluye otros)	grupo de histonas 1, H4c
IPI00472073	HLA-B	complejo mayor de histocompatibilidad, clase I, B
IPI00641737	HP	haptoglobina

ES 2 762 403 T3

IPI00022488	HPX	hemopexina
IPI00220362	HSPE1	proteína de choque térmico de 10 kDa (chaperonina 10)
IPI00217966	LDHA	lactato deshidrogenasa A
IPI00887739	LOC100133511	complemento tipo C3
IPI00022417	LRG1	alfa-2-glicoproteína 1 rica en leucina
IPI00298860	LTF	lactotransferrina
IPI00106687	LXN	latexina
IPI00012989	MAN2B1	manosidasa, alfa, clase 2B, miembro 1
IPI00007244	MPO	mieloperoxidasa
IPI00008787	NAGLU	N-acetilglucosaminidasa, alfa
IPI00020091	ORM1/ORM2	orosomucoide 1
IPI00216691	PFN1	profilina 1
IPI00026962	PLA2G2A	fosfolipasa A2, grupo IIA (plaquetas, líquido sinovial)
IPI00027409	PRTN3	proteínasa 3
IPI00299155	PSMA4	subunidad del proteasoma (prosoma, macropaína), tipo alfa, 4
IPI00029623	PSMA6	subunidad de proteasoma (prosoma, macropaína) alfa tipo-6
IPI00028006	PSMB2	subunidad de proteasoma (prosoma, macropaína) beta tipo-2
IPI00022420	RBP4	proteína 4 de unión a retinol, plasma
IPI00006988	RETN	resistina
IPI00007047	S100A8	proteína S100 de unión a calcio A8
IPI00027462	S100A9	proteína S100 de unión a calcio A9
IPI00010304	SERPINB10	inhibidor de la serpina peptidasa, clado B (ovoalbúmina), miembro 10
IPI00032179	SERPINC1	inhibidor de la serpina peptidasa, clado C (antitrombina), miembro 1
IPI00029863	SERPINF2	inhibidor de la serpina peptidasa, clado F (antiplasmina alfa-2, factor derivado del epitelio pigmentario), miembro 2
IPI00291866	SERPING1	inhibidor de la serpina peptidasa, clado G (inhibidor de C1), miembro 1
IPI00022314	SOD2	superóxido dismutasa 2, mitocondrial
IPI00022463	TF	transferrina
IPI00643920	TKT	transcetolasa
IPI00719280	UBB	ubiquitina B
IPI00884078		no mapeado por la invención
IPI00942608	proteína de 26 kDa	no mapeado por la invención
IPI00419215	A2ML1	alfa-2-macroglobulina tipo 1
IPI00937735	ACADVL	acil-CoA deshidrogenasa, cadena muy larga
IPI00009268	ACY1	no mapeado por la invención
IPI00419916	ALPL	fosfatasa alcalina, hígado/hueso/riñón
IPI00219575	BLMH	bleomicina hidrolasa
IPI00026240	BST1	antígeno 1 de célula estromal de médula ósea
IPI00009008	CACNA1D	canal de calcio, dependiente del voltaje, tipo L, subunidad alfa 1D
IPI00877726	CKMT1A/ CKMT1 B	creatina quinasa, mitocondrial 1B
IPI00299150	CTSS	catepsina S
IPI00031708	FAH	fumarilacetoacetato hidrolasa (fumarilacetoacetasa)
IPI00644409	GDA	guanina desaminasa
IPI00016862	GSR	glutathion reductasa
IPI00396378	HNRNPA2B1	ribonucleoproteína nuclear heterogénea A2/B1
IPI00003865	HSPA8	proteína 8 de choque térmico de 70 kDa
IPI00299145	KRT6C	queratina 6C
IPI00930073	KRT6C	queratina 6C
IPI00009650	LCN1	lipocalina 1 (prealbúmina de lágrima)
IPI00019038	LYZ	lisozima
IPI00922181	MCM2	componente 2 del complejo de mantenimiento de minicromosomas
IPI00257882	PEPD	peptidasa D
IPI00016255	PLBD1	dominio de fosfolipasa B que contiene 1 residuo
IPI00219622	PSMA2	subunidad del proteasoma (prosome, macropaína), tipo alfa, 2
IPI00025019	PSMB1	subunidad de proteasoma (prosoma, macropaína), tipo beta, 1

ES 2 762 403 T3

IPI00000811	PSMB6	subunidad del proteasoma (prosoma, macropaina), tipo beta, 6
IPI00003590	QSOX1	quiescina Q6 sulfhidril oxidasa 1 1
IPI00939604	RNASET2	ribonucleasa T2
IPI00553177	SERPINA1	inhibidor de la serpina peptidasa, clado A (alfa-1 antiproteinasa, antitripsina), miembro 1
IPI00550991	SERPINA3	inhibidor de la serpina peptidasa, clado A (alfa-1 antiproteinasa, antitripsina), miembro 3
IPI00022204	SERPINB3	inhibidor de la serpina peptidasa, clado B (ovoalbúmina), miembro 3
IPI00010949	SIAE	ácido siálico acetilesterasa
IPI00465028	TPI1	triosafosfato isomerasa 1
IPI00029647	ZG16	homólogo de proteína 16 de gránulos de zimógeno (rata)

Tabla 6: Proteínas que son significativamente más abundantes en secretomas de tumor de colon en comparación con secretomas de colon normal

Número de acceso	Símbolo del gen	Nombre del gen
IPI00903112	LTF	lactotransferrina
IPI00329573	COL12A1	colágeno, tipo XII, alfa 1
IPI00374563	AGRN	agrina
IPI00005024	MYBBP1A	proteína de unión a MYB (P160) 1a
IPI00925034	TRRAP	proteína asociada al dominio de transformación/transcripción
IPI00221226	ANXA6	anexina A6
IPI00028275	CKAP5	proteína 5 asociada al citoesqueleto
IPI00018350	MCM5	componente 5 del complejo de mantenimiento de minicromosomas
IPI00156374	IPO4	importina 4
IPI00894235	NBEAL2	tipo neurobeachina 2
IPI00795318	MCM4	componente 4 del complejo de mantenimiento de minicromosomas
IPI00002405	OAS3	2'-5'-oligoadenilato sintetasa 3, 100 kDa
IPI00013214	MCM3	componente 3 del complejo de mantenimiento de minicromosomas
IPI00018968	NAE1	subunidad 1 de la enzima E1 activadora de NEDD8
IPI00438229	TRIM28	Motivo tripartita que contiene 28 residuos
IPI00221354	FUS	fusionado en sarcoma
IPI00031820	FARSA	fenilalanil-tARN sintetasa, subunidad alfa
IPI00013163	MNDA	antígeno de diferenciación nuclear de células mieloides
IPI00026970	SUPT16H	supresor del homólogo Ty 16 (S. cerevisiae)
IPI00017617	DDX5	polipéptido 5 de la caja DEAD (Asp-Glu-Ala-Asp)
IPI00031008	TNC	tenascina C
IPI00007175	NIP7	homólogo 7 de importación nuclear (S. cerevisiae)
IPI00000846	CHD4	proteína 4 de unión a ADN helicasa del cromodominio
IPI00465044	RCC2	regulador de la condensación cromosómica 2
IPI00031519	DNMT1	ADN (citosina-5-)-metiltransferasa 1
IPI00028357	XPO4	exportina 4
IPI00010720	CCT5	TCP1 que contiene chaperonina, subunidad 5 (epsilon)
IPI00012340	SRSF9	factor 9 de empalme rico en serina/arginina
IPI00012645	SPTBN2	espectrina, beta, no eritrocítica 2
IPI00032292	TIMP1	inhibidor 1 de la metalopeptidasa TIMP
IPI00026944	NID1	nidógeno 1
IPI00013871	RRM1	ribonucleótido reductasa M1
IPI00386533	EIF4G1	factor 4 gamma de iniciación de la traducción eucariota, 1
IPI00149849	COG4	componente del complejo 4 oligomérico de Golgi
IPI00002894	POLD1	polimerasa (dirigida al ADN), delta 1, subunidad catalítica de 125 kDa
IPI00221106	SF3B2	factor 3b de empalme, subunidad 2, 145 kDa
IPI00015905	EXOSC2	componente 2 exosómico
IPI00031517	MCM6	componente 6 del complejo de mantenimiento de minicromosomas
IPI00216694	PLS3	plastina 3
IPI00218407	ALDOB	aldolasa B, fructosa-bisfosfato

ES 2 762 403 T3

IPI00556369	SMG1	homólogo de SMG1, fosfatidilinositol 3-quinasa relacionada con la quinasa (C. elegans)
IPI00909083	GSPT1	transición 1 de fase G1 a S
IPI00479786	KHSRP	proteína reguladora de empalme tipo KH
IPI00015953	DDX21	polipéptido 21 de la caja DEAD (Asp-Glu-Ala-Asp)
IPI00334907	PITPNB	proteína de transferencia de fosfatidilinositol, beta
IPI00297572	AQR	homólogo de acuario (ratón)
IPI00011274	HNRPDL	ribonucleoproteína nuclear heterogénea tipo D
IPI00024095	ANXA3	anexina A3
IPI00293331	POP1	procesamiento del precursor 1, subunidad de ribonucleasa P/MRP (S. cerevisiae)
IPI00007927	SMC2	mantenimiento estructural de los cromosomas 2
IPI00011592	DYNC1LI2	dineína, citoplasmática 1, cadena intermedia ligera 2
IPI00003927	PPID	peptidilprolil isomerasa D
IPI00002926	VPS37B	clasificación de proteínas vacuolares 37 homólogo B (S. cerevisiae)
IPI00012497	ADRBK1	adrenérgico, beta, receptor quinasa 1
IPI00746351	DIS3	homólogo de control mitótico DIS3 (S. cerevisiae)
IPI00031960	POLR1A	polipéptido A de polimerasa I (ARN), 194 kDa
IPI00290566	TCP1	complejo t 1
IPI00026952	PKP3	placofilina 3
IPI00856120	LARP1B	familia del dominio de ribonucleoproteína La, miembro 1B
IPI00449049	PARP1	poli (ADP-ribosa) polimerasa 1
IPI00012035	CD46	molécula CD46, proteína reguladora del complemento
IPI00419979	PAK2	quinasa 2 activada por proteína p21 (Cdc42/Rac)
IPI00303207	ABCE1	casete de unión a ATP, subfamilia E (OABP), miembro 1
IPI00872664	USP14	peptidasa 14 específica de ubiquitina (tARN-guanina transglucosilasa)
IPI00290770	CCT3	CP1 que contiene chaperonina, subunidad 3 (gamma)
IPI00294879	RANGAP1	proteína 1 activadora de Ran GTPasa
IPI00013862	DTYMK	desoxitimidilato quinasa (timidilato quinasa)
IPI00329692	NMT1	N-miristoiltransferasa 1
IPI00037283	DNM1L	dinamina tipo 1
IPI00008922	IFITM2	proteína 2 transmembrana inducida por interferón (1-8D)
IPI00304754	FERMT1	miembro 1 de la familia de fermitina
IPI00396203	TBCD	cofactor D de plegado de tubulina
IPI00009071	LOC100505793/SRSF10	factor 10 de empalme rico en serina/arginina
IPI00022827	SLK	quinasa tipo STE20
IPI00918002	MUC5AC/MUC5B	mucina 5AC, moco oligomérico/formador de gel
IPI00008240	MARS	metionil-tARN sintetasa
IPI00217013	SMEK1	homólogo SMEK 1, supresor de mek1 (Dictyostelium)
IPI00219097	HMGB2	caja 2 del grupo de alta movilidad
IPI00304596	NONO	dominio que no contiene POU, de unión a octámero
IPI00018219	TGFBI	factor de crecimiento transformante, inducido por beta, 68 kDa
IPI00023824	FBLN2	fibulina 2
IPI00022228	HDLBP	proteína de unión a lipoproteína de alta densidad
IPI00306322	COL4A2	colágeno, tipo IV, alfa 2
IPI00018452	CPNE1	copina I
IPI00018627	NAA50	N(alfa)-acetiltransferasa 50, subunidad catalítica NatE
IPI00007163	LSM7	homólogo de LSM7, pequeño ARN nuclear U6 asociado (S. cerevisiae)
IPI00005154	SSRP1	proteína 1 de reconocimiento específica de la estructura
IPI00007401	IPO8	importina 8
IPI00429538	YPEL5	yippee tipo 5 (Drosophila)
IPI00030116	PGM3	fosfoglucomutasa 3
IPI00162563	RNF40	proteína 40 de dedo anular
IPI00219420	SMC3	mantenimiento estructural de cromosomas 3
IPI00012501	REG4	familia derivada de islotes de regeneración, miembro 4
IPI00029764	SF3A3	factor 3a de empalme, subunidad 3, 60 kDa

ES 2 762 403 T3

IPI00296099	THBS1	trombospondina 1
IPI00027626	CCT6A	TCP1 que contiene chaperonina, subunidad 6A (zeta 1)
IPI00007928	PRPF8	homólogo del factor 8 de procesamiento previo al ARNm PRP8 (S. cerevisiae)
IPI00023344	SYMPK	simplequina
IPI00375441	FUBP1	proteína 1 de unión al elemento alejado secuencia arriba (FUSE)
IPI00005613	U2AF1	factor 1 auxiliar de ARN nuclear pequeño U2
IPI00002335	HTT	huntingtina
IPI00299254	EIF5B	factor 5B de iniciación de la traducción eucariota
IPI00179953	NASP	proteína de esperma autoantigénica nuclear (unión a histona)
IPI00514561	HNRNPK	ribonucleoproteína K nuclear heterogénea
IPI00031812	YBX1	proteína 1 de unión a la caja Y
IPI00414320	ANXA11	anexina A11
IPI00178431	RECQL	tipo proteína RecQ (ADN helicasa tipo Q1)
IPI00029601	CTTN	cortactina
IPI00013683	TUBB3	tubulina, beta 3
IPI00646105	PYCRL	tipo pirrolina-5-carboxilato reductasa
IPI00298057	PPL	periplaquina
IPI00173346	PGM2L1	fosfoglucomutasa 1 tipo 2
IPI00373869	C17orf49	marco 49 de lectura abierto del cromosoma 17
IPI00106491	MRTO4	homólogo 4 de cambio de ARNm (S. cerevisiae)
IPI00290184	METTL1	metiltransferasa tipo 1
IPI00006025	SART3	antígeno de carcinoma de células escamosas reconocido por las células T 3
IPI00016179	S100A13	proteína S100 de unión a calcio A13
IPI00100292	NPEPL1	tipo aminopeptidasa 1
IPI00026689	CDK1	quinasa 1 dependiente de ciclina
IPI00217405	UBR1	componente n-recognina 1 de proteína ligasa ubiquitina E3
IPI00296353	ARHGAP18	proteína 18 activadora de Rho GTPasa
IPI00293434	SRP14	partícula de reconocimiento de señal de 14 kDa (proteína de unión al ARN homóloga de Alu)
IPI00292532	CAMP	péptido antimicrobiano catelicidina
IPI00010740	SFPQ	factor de empalme rico en prolina/glutamina
IPI00026262	RASA1	activador de proteína RAS p21 (proteína activadora de GTPasa) 1
IPI00409601	RALGAPB	proteína activadora de Ral GTPasa, subunidad beta (no catalítica)
IPI00013976	LAMB1	laminina, beta 1
IPI00554590	RAB3GAP2	subunidad 2 de proteína activadora de GTPasa RAB3 (no catalítica)
IPI00302925	CCT8	TCP1 que contiene chaperonina, subunidad 8 (theta)
IPI00103247	HNRPLL	ribonucleoproteína nuclear heterogénea tipo L
IPI00414127	RANBP1	proteína 1 de unión a RAN
IPI00001458	KNTC1	cinetocoro asociado 1
IPI00221394	DKC1	disqueratosis congénita 1, disquerina
IPI00016613	CSNK2A1	caseína quinasa 2, polipéptido alfa 1
IPI00013455	CLIP1	dominio CAP-GLY que contiene la proteína enlazadora 1
IPI00297779	CCT2	TCP1 que contiene chaperonina, subunidad 2 (beta)
IPI00029048	TTL12	familia tipo tirosina ligasa de tubulina, miembro 12
IPI00298306	ATM	ataxia telangiectasia mutada
IPI00017451	SF3A1	factor de empalme 3a, subunidad 1, 120 kDa
IPI00012493	RPS20	proteína S20 ribosómica
IPI00783378	UBE20	enzima conjugadora de ubiquitina E2O
IPI00742682	TPR	región promotora translocada (para el oncogén MET activado)
IPI00002203	BCCIP	BRCA2 y proteína de interacción con CDKN1A
IPI00291783	GEMIN5	proteína 5 asociada a gem (orgánulo nuclear)
IPI00019196	RPP30	subunidad ribonucleasa P/MRP de 30 kDa
IPI00056386	LOH12CR1	pérdida de heterocigosidad, 12, región 1 cromosómica
IPI00019971	STXBP2	proteína 2 de unión a syntaxina
IPI00020965	UBE2H	enzima E2H que se conjuga con ubiquitina

ES 2 762 403 T3

IPI00465045	DIP2B	homólogo B de la proteína 2 que interactúa con disco DIP2 (Drosophila)
IPI00607591	RAP1GDS1	RAP1, estimulador 1 de disociación de GTP-GDP
IPI00171903	HNRNPM	ribonucleoproteína M nuclear heterogénea
IPI00291802	LMO7	dominio 7 de LIM
IPI00004273	RBM25	proteína 25 del motivo de unión a ARN
IPI00221234	ALDH7A1	familia de la aldehído deshidrogenasa 7, miembro A1
IPI00026219	CPSF1	factor 1 específico de escisión y poliadenilación, 160 kDa
IPI00015262	CNN2	calponina 2
IPI00018465	CCT7	TCP1 que contiene chaperonina, subunidad 7 (eta)
IPI00014238	KARS	lisil-tARN sintetasa
IPI00000684	UAP1	UDP-N-acteilglucosamina pirofosforilasa 1
IPI00295485	HSPA4L	proteína de choque térmico 70 kDa tipo 4
IPI00926528	proteína 138 kDa	proteína 138 kDa
IPI00549189	THOP1	thimet oligopeptidasa 1
IPI00008552	GLRX3	glutarredoxina 3
IPI00011200	PHGDH	fosfoglicerato deshidrogenasa
IPI00014197	CDV3	homólogo de CDV3 (ratón)
IPI00411559	SMC4	mantenimiento estructural de cromosomas 4
IPI00024320	RBM3	proteína 3 del motivo de unión a ARN (RNP1, RRM)
IPI00020956	HDGF	factor de crecimiento derivado de hepatoma
IPI00479217	HNRNPU	ribonucleoproteína nuclear heterogénea U (factor A de unión a Scaffold)
IPI00604756	NRBP1	proteína 1 de unión al receptor nuclear
IPI00418797	POLR1B	polipéptido B, de polimerasa (ARN) I, de 128 kDa
IPI00019812	PPP5C	proteína fosfatasa 5, subunidad catalítica
IPI00216008	G6PD	glucosa-6-fosfato deshidrogenasa
IPI00291560	ARG1	arginasa, hígado
IPI00008475	HMGCS1	3-hidroxi-3-metilglutaril-CoA sintasa 1 (soluble)
IPI00023234	UBA2	Enzima 2 que activa al modificador tipo ubiquitina
IPI00298991	KIAA1033	KIAA1033
IPI00793199	ANXA4	anexina A4
IPI00889541	DDX17	polipéptido 17 de la caja DEAD (Asp-Glu-Ala-Asp)
IPI00165393	ANP32E	familia de fosfoproteína 32 nuclear ácida (rica en leucina), miembro E
IPI00896504	GNE	glucosamina (UDP-N-acetil)-2-epimerasa/N-acetilmannosamina quinasa
IPI00157790	KIAA0368	KIAA0368
IPI00182728	VPS4B	homólogo B de clasificación de proteínas vacuolar 4 (S. cerevisiae)
IPI00020127	RPA1	proteína de replicación A1, 70 kDa
IPI00219678	EIF2S1	factor 2 de iniciación de la traducción eucariota, subunidad 1 alfa, 35 kDa
IPI00290461	EIF3J	factor 3 de iniciación de la traducción eucariota, subunidad J
IPI00784161	SUPT6H	supresor del homólogo de Ty 6 (S. cerevisiae)
IPI00218993	HSPH1	proteína 1 de choque térmico de 105 kDa/110 kDa
IPI00640703	XPO5	exportina 5
IPI00333215	TCEA1	factor A de alargamiento de la transcripción (SII), 1
IPI00009032	SSB	antígeno B del síndrome de Sjogren (autoantígeno La)
IPI00745313	AEBP1	proteína 1 de unión a AE
IPI00303258	LMCD1	LIM y dominios 1 ricos en cisteína
IPI00219330	ILF3	factor 3 de unión al potenciador de interleuquina, 90 kDa
IPI00019269	WDR61	dominio 61 de repetición de WD
IPI00386189	NAA15	N(alfa)-acetiltransferasa 15, subunidad auxiliar NatA
IPI00000015	SRSF4	factor 4 de empalme rico en serina/arginina
IPI00251559	RNF20	proteína 20 de dedo anular
IPI00006952	LACTB2	lactamasa, beta 2
IPI00041325	NHP2	homólogo de ribonucleoproteína NHP2 (levadura)
IPI00607714	C17orf28	marco de lectura abierto 28 del cromosoma 17
IPI00645702	CTPS2	CTP sintasa II
IPI00163187	FSCN1	homólogo 1 de fascina, proteína de agrupamiento de actina (Strongylocentrotus)

		purpuratus)
IPI00289807	TRNT1	tARN nucleotidil transferasa, adición de CCA, 1
IPI00375462	SREK1	proteína 1 rica en glutamina/lisina reguladora de empalme
IPI00025158	STAG1	antígeno 1 estromal
IPI00024971	OSBP	proteína de unión a oxisterol
IPI00375015	DUT	desoxiuridina trifosfatasa
IPI00396174	CCDC25	dominio de hélice superenrollada que contiene 25 residuos
IPI00020021	DEK	oncogén DEK
IPI00006378	CCDC72	dominio de hélice superenrollada que contiene 72 residuos
IPI00291093	POLR2E	polipéptido E de polimerasa (ARN) II (dirigido por ADN), 25 kDa
IPI00019178	PSPH	fosfoserina fosfatasa
IPI00291939	SMC1A	mantenimiento estructural de los cromosomas 1A
IPI00006725	DDX23	polipéptido 23 de la caja DEAD (Asp-Glu-Ala-Asp)
IPI00009010	TRMT112	homólogo de tARN metiltransferasa 11-2 (S. cerevisiae)
IPI00018813	COPS2	subunidad 2 homóloga fotomorfogénica constitutiva de COP9 (Arabidopsis)
IPI00023640	PDCD5	muerte celular programada 5
IPI00031681	CDK2	quinasa 2 dependiente de ciclina
IPI00011603	PSMD3	subunidad 26S de proteasoma (prosoma, macropaina), sin ATPasa, 3
IPI00221325	RANBP2	proteína 2 de unión a RAN
IPI00410693	SERBP1	proteína 1 de unión a ARNm SERPINE1
IPI00005780	OGT	N-acetilglucosamina transferasa (GlcNAc) enlazada a O (UDP-N-acetilglucosamina: polipéptido-N-acetilglucosaminil transferasa)
IPI00299524	NCAPD2	complejo sin SMC condensina I, subunidad D2
IPI00150057	SMARCC2	regulador de cromatina dependiente de actina, relacionado con SWI/SNF, asociado a la matriz, subfamilia c, miembro 2
IPI00032853	NOP10	homólogo de ribonucleoproteína NOP10 (levadura)
IPI00002255	LRBA	tráfico de vesículas sensibles a LPS, que contiene beige y ancla
IPI00554742	API5	inhibidor 5 de apoptosis
IPI00385267	SRPR	receptor de partículas de reconocimiento de señal (proteína de acoplamiento)
IPI00022462	TFRC	receptor de transferrina (p90, CD71)
IPI00022305	BZW2	cremallera básica de leucina y dominios 2 W2
IPI00025427	RNASE3	ribonucleasa, familia 3 de RNasa A
IPI00010182	DBI	inhibidor de unión a diazepam (modulador del receptor GABA, proteína de unión a acil-CoA)
IPI00219005	FKBP4	proteína 4 de unión a FK506, 59 kDa
IPI00184871	C6orf130	marco 130 de lectura abierto del cromosoma 6
IPI00103483	COBRA1	cofactor de BRCA1
IPI00026215	FEN1	endonucleasa 1 específica de la estructura de aleta
IPI00296635	GBE1	glucano (1,4-alfa-), enzima ramificadora 1
IPI00029267	SNRPB2	polipéptido B de ribonucleoproteína nuclear pequeña
IPI00100197	NSFL1C	cofactor NSFL1 (p97) (p47)
IPI00010415	ACOT7	acil-CoA tioesterasa 7
IPI00306369	NSUN2	familia del dominio NOP2/Sun, miembro 2
IPI00302927	CCT4	TCP1 que contiene chaperonina, subunidad 4 (delta)
IPI00023845	KLK6	peptidasa 6 relacionada con calicreín
IPI00003919	QPCT	glutaminil-péptido ciclotransferasa
IPI00465128	BAG6	atanógeno 6 asociado a BCL2
IPI00016910	EIF3C/EIF3CL	factor 3 de iniciación de la traducción eucariota, subunidad C
IPI00007812	ATP6V1B2	ATPasa, que transporta H <sup>+</sup> , lisosomal de 56/58 kDa, V1 subunidad B2
IPI00027846	MMP8	matriz de metalopeptidasa 8 (colagenasa de neutrófilos)
IPI00023919	PSMC5	subunidad 26S del proteasoma (prosoma, macropaina), ATPasa, 5
IPI00217253	GCHFR	regulador de retroalimentación de GTP ciclohidrolasa I
IPI00384028	PAPOLA	poli(A) polimerasa alfa
IPI00219344	HPCAL1	tipo hipocalcina 1
IPI00012442	G3BP1	proteína 1 de unión a la proteína (dominio SH3) activadora de GTPasa

ES 2 762 403 T3

IPI00024163	POLR3A	polipéptido A de polimerasa (ARN) III (dirigido por ADN), 155 kDa
IPI00647217	SKIV2L2	actividad 2 viralicídica superasesina tipo 2 ( <i>S. cerevisiae</i> )
IPI00031627	POLR2A	polipéptido A de polimerasa (ARN) III (dirigido por ADN), 220 kDa
IPI00873137	COL1A2	colágeno, tipo I, alfa 2
IPI00025039	FBL	fibrilarina
IPI00923606	EPRS	glutamil-prolil-tARN sintetasa
IPI00301936	ELAVL1	ELAV (embrionaria letal, visión anormal, <i>Drosophila</i> ) tipo 1 (antígeno Hu R)
IPI00183500	NCBP2	subunidad 2 de proteína de unión a la caperuza nuclear, 20 kDa
IPI00001159	GCN1L1	control general de GCN1 de la síntesis 1 de aminoácidos tipo 1 (levadura)
IPI00024719	HAT1	histona acetiltransferasa 1
IPI00470883	STAG2	antígeno estromal 2
IPI00216057	SORD	sorbitol deshidrogenasa
IPI00032830	REXO2	REX2, homólogo de ARN exonucleasa 2 ( <i>S. cerevisiae</i> )
IPI00003881	HNRNPF	ribonucleoproteína F nuclear heterogénea
IPI00216230	TMPO	timopoyetina
IPI00902614	USP24	peptidasa 24 específica de ubiquitina
IPI00182757	KIAA1967	KIAA1967
IPI00296165	C1R	componente 1 del complemento, subcomponente r
IPI00002460	ANXA7	anexina A7
IPI00009104	RUVBL2	tipo RuvB 2 ( <i>E. coli</i> )
IPI00651738	ADI1	acireductona dioxigenasa 1
IPI00009328	EIF4A3	factor 4A3 de iniciación de la traducción eucariota
IPI00456887	HNRNPUL2	ribonucleoproteína U nuclear heterogénea tipo 2
IPI00646605	UBR4	n-recognina 4 componente de la proteína ligasa E3 ubiquitina
IPI00386718	SMARCA2	regulador de cromatina dependiente de actina, asociado a la matriz, relacionado a SWI/SNF, subfamilia a, miembro 2
IPI00008234	CYB5R2	citocromo b5 reductasa 2
IPI00300371	SF3B3	factor 3b de empalme, subunidad 3, 130 kDa
IPI00156282	GPS1	supresor 1 de la ruta de la proteína G
IPI00012369	MAD2L1	detención mitótica deficiente MAD2 tipo 1
IPI00016736	PLCG1	fosfolipasa C, gamma 1
IPI00014263	EIF4H	factor 4H de iniciación de la traducción eucariota
IPI00031556	U2AF2	factor 2 auxiliar de ARN nuclear pequeño U2
IPI00304589	TNKS1BP1	proteína 1 de unión a tanquirasa 1, 182 kDa
IPI00294578	TGM2	transglutaminasa 2 (polipéptido C, proteína-glutamina-gamma-glutamyltransferasa)
IPI00027834	HNRNPL	ribonucleoproteína L nuclear heterogénea
IPI00329213	INPP5D	inositol polifosfato-5-fosfatasa, 145 kDa
IPI00296528	ANXA10	anexina A10
IPI00013180	BUD31	homólogo de BUD31 ( <i>S. cerevisiae</i> )
IPI00216048	PITPNA	proteína de transferencia de fosfatidilinositol, alfa
IPI00103994	LARS	leucil-tARN sintetasa
IPI00027681	NNMT	nicotinamida N-metiltransferasa
IPI00185374	PSMD12	subunidad 26S de proteasoma (prosome, macropoína), no ATPasa, 12
IPI00004838	CRK	homólogo del Oncogene CT10 del virus de sarcoma de v-crk (aviar)
IPI00847535	PRG2	proteoglicano 2, médula ósea (activador de células asesinas naturales, proteína básica principal del gránulo de eosinófilos)
IPI00009802	VCAN	versicano
IPI00306290	XPOT	exportina, tARN (receptor de exportación nuclear para los tARN)
IPI00025347	EMG1	homólogo de la proteína nucleolar EMG1 ( <i>S. cerevisiae</i> )
IPI00410091	C11orf73	marco 73 de lectura abierto del cromosoma 11
IPI00024364	TNP01	transportina 1
IPI00292150	LTBP2	proteína 2 de unión beta del factor de crecimiento transformante latente
IPI00470891	CSDE1	dominio de choque frío que contiene E1, unión a ARN
IPI00168554	SRXN1	sulfirredoxina 1

ES 2 762 403 T3

IPI00103525	PSPC1	componente 1 de paraspeckle
IPI00419880	RPS3A	proteína S3A ribosómica
IPI00375631	ISG15	ISG15 modificador tipo ubiquitina
IPI00027808	POLR2B	polipéptido B de polimerasa (ARN) II (dirigido por ADN), 140 kDa
IPI00054042	GTF21	factor de transcripción general Ili
IPI00026167	NHP2L1	proteína 2 de cromosoma sin histona NHP2 tipo 1 (S. cerevisiae)
IPI00003565	PSMD10	subunidad 26S del proteasoma (prosome, macropáina), sin ATPasa, 10
IPI00030781	STAT1	transductor de señal y activador de la transcripción 1, 91 kDa
IPI00293026	EFTUD1	dominio de unión de GTP del factor de alargamiento Tu que contiene 1 residuo
IPI00646226	MED23	subunidad 23 del complejo mediador
IPI00745433	EIF2C2	factor 2C de iniciación de la traducción eucariota, 2
IPI00027838	RBM4B	proteína 4B del motivo de unión de ARN
IPI00024425	KIAA0664	KIAA0664
IPI00024871	CBFB	factor de unión al núcleo, subunidad beta
IPI00008524	PABPC1	proteína de unión a poli(A), citoplasmática 1
IPI00018873	NAMPT	nicotinamida fosforibosiltransferasa
IPI00216088	CRABP2	proteína 2 de unión a ácido retinoico celular
IPI00032342	TRIP12	interactuador 12 receptor de la hormona tiroidea
IPI00154975	DNAJC9	homólogo de DnaJ (Hsp40), subfamilia C, miembro 9
IPI00186008	STARD10	dominio de transferencia de lípidos relacionado con StAR (START) que contiene 10 residuos
IPI00828098	RNF213	proteína 213 de dedo anular
IPI00011898	EIF2B5	factor 2B de iniciación de la traducción eucariota, subunidad 5 épsilon, 82 kDa
IPI00301434	BOLA2/BOLA2B	homólogo 2 de bolA (E. coli)
IPI00181391	MGEA5	antígeno 5 expresado en meningioma (hialuronidasa)
IPI00022664	RABGGTA	Rab geranilgeraniltransferasa, subunidad alfa
IPI00384689	PDXDC1	dominio de descarboxilasa dependiente de piridoxal que contiene 1
IPI00552920	EXOSC8	componente 8 del exoma
IPI00001734	PSAT1	fosfoserina aminotransferasa 1
IPI00010105	EIF6	factor 6 de iniciación de la traducción eucariota
IPI00745105	C16orf13	marco 13 de lectura abierto del cromosoma 16
IPI00306436	STAT3	transductor de señal y activador de la transcripción 3 (factor de respuesta de fase aguda)
IPI00220815	EFEMP1	proteína 1 de matriz extracelular similar a fibulina que contiene EGF
IPI00008301	DEFA6	defensina, alfa 6, específica de la célula de Paneth
IPI00015018	PPA1	pirofosfatasa (inorgánica) 1
IPI00298176	GPX2	glutación peroxidasa 2 (gastrointestinal)
IPI00456635	UNC13D	homólogo D de unc-13 (C. elegans)
IPI00218604	PTPN6	proteína tirosina fosfatasa, sin receptor tipo 6
IPI00760846	MYO18A	miosina XVIIIa
IPI00296534	FBLN1	fibulina 1
IPI00025329	RPL19	proteína ribosómica L19
IPI00030876	DIAPH1	homólogo diáfano 1 (Drosophila)
IPI00154451	MMS19	homólogo de reparación por escisión del nucleótido MMS19 (S. cerevisiae)
IPI00646917	NUDT21	motivo 21 tipo nudix (fracción X ligada a nucleósido difosfato)
IPI00010404	SF3B5	factor 3b de empalme, subunidad 5, 10 kDa
IPI00014938	SARNP	ribonucleoproteína que contiene el dominio SAP
IPI00002186	ARFGEF2	factor 2 de intercambio del nucleótido guanina del factor de ribosilación de ADP (inhibido por brefeldina A)
IPI00028564	GBP1	proteína 1 de unión a guanilato, inducible por interferón
IPI00179298	HUWE1	dominio HECT, UBA y WWE que contiene 1 residuo
IPI00027142	POP7	procesamiento del precursor 7, subunidad de ribonucleasa P/MRP (S. cerevisiae)
IPI00006504	EIF2B3	factor de iniciación de la traducción eucariota 2B, subunidad 3 gamma, 58 kDa
IPI00013184	NAA10	N(alfa)-acetiltransferasa 10, subunidad catalítica NatA
IPI00015947	DNAJB1	homólogo de DnaJ (Hsp40), subfamilia B, miembro 1

ES 2 762 403 T3

IPI00019463	EIF2AK2	alfa quinasa 2 del factor 2 de iniciación de la traducción eucariota
IPI00003519	EFTUD2	dominio de unión a GTP del factor Tu de alargamiento que contiene 2 residuos
IPI00004524	GCA	grancalcina, proteína de unión a calcio con mano EF
IPI00333776	NRCAM	molécula de adhesión celular neuronal
IPI00552073	DHX16	polipéptido 16 de caja DEAH (Asp-Glu-Ala-His)
IPI00013949	SGTA	que contiene la pequeña repetición del tetratricopéptido rico en glutamina (TPR), alfa
IPI00304742	STK10	serina/treonina quinasa 10
IPI00305833	SMU1	supresor de smu-1 del homólogo de mec-8 y unc-52 (C. elegans)
IPI00150269	PRPF4	homólogo del factor 4 de procesamiento de pre-ARNm PRP4 (levadura)
IPI00028888	HNRNPD	ribonucleoproteína nuclear heterogénea D (proteína 1 de unión a ARN del elemento rico en AU, 37 kDa)
IPI00290142	CTPS	CTP sintasa
IPI00015195	CSTF3	factor de estimulación de escisión, 3' pre-ARN, subunidad 3, 77 kDa
IPI00011528	CSTF1	factor de estimulación de escisión, 3' pre-ARN, subunidad 1, 50 kDa
IPI00013070	HNRNPUL1	heterogeneous nuclear ribonucleoprotein U-like 1
IPI00008433	RPS5	proteína ribosómica S5
IPI00298347	PTPN11	proteína tirosina fosfatasa, tipo 11 sin receptor
IPI00514234	LAD1	ladinina 1
IPI00000057	COG2	componente del complejo de Golgi oligomérico 2
IPI00014311	CUL2	culina 2
IPI00221093	RPS17/RPS17L	proteína ribosómica S17
IPI00022694	PSMD4	subunidad 26S de proteasoma (prosome, macropaina), sin ATPasa, 4
IPI00218918	ANXA1	anexina A1
IPI00418169	ANXA2	anexina A2
IPI00289499	ATIC	ribonucleótido formiltransferasa 5-aminoimidazol-4-carboxamida/IMP ciclohidrolasa
IPI00022246	AZU1	azurocidina 1
IPI00299635	BIRC6	repetición de IAP baculoviral que contiene 6 residuos
IPI00002147	CHI3L1	quitinasa 1 tipo 3 (glicoproteína 39 de cartílago)
IPI00011062	CPS1	carbamoil-fosfato sintasa 1, mitocondrial
IPI00028064	CTSG	catepsina G
IPI00005721	DEFA1 (incluidos otros)	defensina, alfa 1
IPI00099110	DMBT1	eliminada en tumores malignos de cerebro 1
IPI00027769	ELANE	elastasa, expresada en neutrófilos
IPI00217987	ITGAM	integrina alfa M (subunidad del receptor 3 del componente 3 del complemento)
IPI00299547	LCN2	lipocalina 2
IPI00023673	LGALS3BP	lectina, proteína de unión a 3, soluble, que se une a galactósido
IPI00922181	MCM2	componente 2 del complejo de mantenimiento de minicromosomas
IPI00027509	MMP9	metalopeptidasa 9 de matriz (gelatinasa B, gelatinasa de 92 kDa, colagenasa tipo IV de 92 kDa)
IPI00007244	MPO	mieloperoxidasa
IPI00103397	MUC5AC/MUC5B	mucina 5AC, moco oligomérico/formador de gel
IPI00019380	NCBP1	subunidad 1 de la proteína de unión a la caperuza nuclear, 80 kDa
IPI00299512	NF1	neurofibromina 1
IPI00022255	OLFM4	olfactomedina 4
IPI00303063	PDS5A	PDS5, regulador del mantenimiento de la cohesión, homólogo A (S. cerevisiae)
IPI00021085	PGLYRP1	proteína 1 de reconocimiento de peptidoglucano
IPI00027409	PRTN3	proteínasa 3
IPI00003590	QSOX1	quiescina Q6 sulfhidril oxidasa 1 1
IPI00009027	REG1A	regeneración 1 alfa derivada de islotes
IPI00027462	S100A9	proteína S100 de unión a calcio A9
IPI00010304	SERPINB10	inhibidor de la serpina peptidasa, clado B (ovoalbúmina), miembro 10
IPI00783625	SERPINB5	inhibidor de la serpina peptidasa, clado B (ovoalbúmina), miembro 5
IPI00072534	UNC45A	homólogo A de unc-45 (C. elegans)

Tabla 7: Un subconjunto de biomarcadores de la Tabla 1 que también fueron significativamente más secretados por las muestras de tumor de colon que por el colon normal (Tabla 6).

Acceso	Símbolo del gen	Nombre del gen
IPI00418169	ANXA2	anexina A2
IPI00022246	AZU1	azurocidina 1
IPI00028064	CTSG	catepsina G
IPI00005721	DEFA1 (incluye otros)	defensina, alfa 1
IPI00027769	ELANE	elastasa, expresada en neutrófilos
IPI00217987	ITGAM	integrina alfa M (subunidad del receptor 3 del componente 3 del complemento)
IPI00007244	MPO	mieloperoxidasa
IPI00019380	NCBP1	subunidad 1 de la proteína de unión a la caperuza nuclear, 80 kDa
IPI00027409	PRTN3	proteínasa 3
IPI00027462	S100A9	proteína S100 de unión a calcio A9
IPI00010304	SERPINB10	inhibidor de la serpina peptidasa, clado B (ovoalbúmina), miembro 10
IPI00218918	ANXA1	anexina A1
IPI00299635	BIRC6	repetición de IAP baculoviral que contiene 6 residuos
IPI00922181	MCM2	componente 2 del complejo de mantenimiento de minicromosomas
IPI00003590	QSOX1	quiescina Q6 sulfhidril oxidasa 1 1
IPI00783625	SERPINB5	inhibidor de la serpina peptidasa, clado B (ovoalbúmina), miembro 5

5 La invención se describe adicionalmente mediante los siguientes ejemplos ilustrativos. Estos no están destinados a limitar el alcance de la invención.

Ejemplo 1

10 Materiales y métodos

Muestras de heces y extracción de proteínas.

15 Se obtuvo el consentimiento informado por escrito de todos los sujetos que proporcionaron muestras de heces y este estudio fue aprobado por el Comité de Ética Médica del VU University Medical Center, Los Países Bajos.

20 Se recogieron muestras parciales de heces de sujetos de referencia que se sometieron a una colonoscopia entre noviembre de 2003 y junio de 2006 en el VU University Medical Center en Ámsterdam, Los Países Bajos. Se recogieron muestras parciales de heces antes de la colonoscopia o después del diagnóstico en la colonoscopia y antes de la resección quirúrgica de sus tumores (Véase la Tabla 8 para las características del paciente).

Conjunto independiente (verificación) de muestras de heces

25 Se recogieron muestras de heces enteras homogeneizadas de sujetos remitidos y sometidos a colonoscopia entre julio de 2009 y abril de 2011 en el VU University Medical Center en Ámsterdam, Los Países Bajos. Se recogieron muestras de heces enteras antes de la colonoscopia (Véase la Tabla 1 para las características del paciente). Los participantes del estudio agregaron inmediatamente tampón de estabilización a las muestras de heces después de la defecación, las muestras se procesaron en el laboratorio con una proporción final de heces: tampón p/v de 1:7 en 72 horas, y se almacenaron a -80 °C hasta su uso. Los extractos de proteínas se prepararon como se describió anteriormente usando una muestra de heces homogeneizada de 2 mL como material de partida. Se mezclaron cantidades iguales de extractos de proteínas y se trataron adicionalmente como una muestra única. Se compusieron cuatro grupos de diferentes categorías, es decir, controles (n = 5) e individuos con adenomas no avanzados (n = 5), adenomas avanzados (n = 5) y CRC (n = 5) (Véase la Tabla 8 para las características del paciente).

35 Tabla 8. Características clinicopatológicas de los pacientes

A. Conjunto de descubrimiento							
Número del paciente <sup>1</sup>	Edad	Género	Categoría	Tamaño del tumor (mm)	Estadio de Dukes	Localización del tumor	FIT <sup>2</sup> positivo
1	74	f	CRC	40	C	ciego	+
2	85	m	CRC	35	A	lado derecho	-
3	86	m	CRC	30	A	Descendiente/sigmoideo	-
4	82	f	CRC	59	B	lado derecho	-
5	60	m	CRC	15	B	rectosigmoideo	+
6	85	m	CRC	60	B	lado derecho	+

7	83	m	CRC	70	B	flexura hepática	-
8	71	f	CRC	35	A	sigmoideo	+
9	63	f	CRC	50	C	rectosigmoideo	+
10	59	m	CRC	52	C	ciego	+
11	84	f	CRC	50	C	ileocecal	+
12	86	m	CRC	68	C	ascendente	+
13	69	m	control	na	na	na	-
14	75	m	control	na	na	na	-
15	68	f	control	na	na	na	-
16	73	f	control	na	na	na	-
17	75	f	control	na	na	na	-
18	67	f	control	na	na	na	-
19	62	m	control	na	na	na	-
20	60	f	control	na	na	na	-
21	60	m	control	na	na	na	-
22	58	m	control	na	na	na	-

## B. Conjunto de verificación

Número del paciente <sup>1</sup>	Edad	Género	Categoría	Tamaño de tumor (mm)	Estadio de Dukes	Localización del tumor	FIT <sup>2</sup> positivo
23	43	f	control	na	na	na	-
24	59	f	control	na	na	na	-
25	44	f	control	na	na	na	-
26	48	m	control	na	na	na	-
27	78	m	control	na	na	na	-
28	78	m	A	2	na	sigmoideo	-
29	68	m	A	3	na	ascendente	NA
30	55	f	A	2	na	sigmoideo	NA
31	72	m	A	8	na	sigmoideo	-
32	66	f	A	7	na	sigmoideo	+
33	84	f	AA	8	na	sigmoideo	NA
34	78	f	AA	10	na	sigmoideo	-
35	87	m	AA	2	na	descendente	-
36	78	f	AA	10	na	ascendente	+
37	78	m	AA	4	na	sigmoideo	+
38	89	f	CRC	30	C	rectosigmoideo	NA
39	58	m	CRC	50	B	recto	NA
40	71	m	CRC	35	C	ascendente/flexura hepática	+
41	72	m	CRC	40	A	sigmoideo	-
42	70	f	CRC	15	C	ascendente	-

<sup>a</sup> Esta columna representa un código numérico de anonimización

<sup>b</sup> Según la clasificación modificada de Astler Coller (Compton y Greene, CA Cancer J Clin 2004)

<sup>c</sup> corte > 75 ng/μL (van Veen, Ned tijdschriftGeneeskde, 2009)

5 En la recolección, las muestras de heces se almacenaron inmediatamente a 4 °C y se transfirieron a -20 °C en 36 horas. Las muestras de heces se descongelaron y después de realizar FIT (prueba inmunoquímica fecal; FIT), se tomó aproximadamente 1 g de muestra de cada muestra de heces para la extracción de proteínas. Para este estudio, se seleccionaron muestras de heces de 10 sujetos sin neoplasia de colon y 4 muestras de heces de pacientes con CRC con una puntuación FIT negativa junto a 8 muestras de heces de pacientes con CRC con una puntuación FIT positiva.

15 Las proteínas de las heces se extrajeron como se describió anteriormente (Ang CS, Nice EC. Targeted in-gel MRM: a hypothesis driven approach for colorectal cancer biomarker discovery in human feces. J Proteome Res 2010; 9: 4346-55) con pocas adaptaciones. En resumen, las muestras se homogeneizaron en un exceso de volumen de PBS de dos veces mediante agitación tipo vórtice y se centrifugaron a 4 °C durante 15 minutos a 16.000 g. Los sobrenadantes se centrifugaron una vez más a 4 °C durante 10 minutos a velocidad máxima. Después del último ciclo de centrifugado, los sobrenadantes se limpiaron de las partículas restantes mediante filtración a través de un filtro de PVDF de 0,22 μm (Millipore, Billerica, MA, EE. UU.). Finalmente, las muestras se concentraron hasta aproximadamente 200 μL usando un filtro de corte de 3 kDa (Amicon Ultra, Millipore Corporation, Billerica, MA, EE. UU.).

20 Electroforesis en gel 1D-SDS y procesamiento de muestras para análisis proteómico

Se cargaron cantidades iguales de proteína aproximadamente 30 μg) y se separaron en geles de gradiente prefabricados al 4-12% (Invitrogen, Carlsbad, EE. UU.), Se fijaron en etanol al 50% que contenía ácido fosfórico al 3%

lavado y teñido durante una noche con Coomassie R-250. Después de la tinción, los geles se lavaron en agua MilliQ y se almacenaron a 4 °C hasta el procesamiento para la digestión en gel. Cada carril se cortó en 10 bandas individuales iguales y cada banda se procesó adicionalmente en péptidos tripticos como se describió anteriormente (Albrethsen J, et al., Mol Cell Proteomics 2010; 9: 988-1005; Piersma SR, et al., J Proteome Res 2010; 9: 1913-22).

Para muestras de una colección de heces independiente (experimento de verificación), se cargaron cantidades iguales de las muestras (20 µL) en un gel SDS-PAGE al 12,5% y se corrieron en el gel hasta que las proteínas entraron en el gel de operación. Luego el gel se fijó y se tiñó como se describió anteriormente. Las muestras se cortaron del gel como una banda única y se procesaron adicionalmente en péptidos tripticos como se describió anteriormente (Albrethsen J, et al., Mol Cell Proteomics 2010; 9: 988-1005; Piersma SR, et al., J Proteome Res 2010; 9: 1913-22). Los péptidos se extrajeron y el volumen de las fracciones peptídicas desaladas se redujo a 50 µL en una centrifuga de vacío.

#### Análisis proteómico por nanoLC-MS/MS

Los péptidos se separaron mediante un sistema de nanoLC Ultimate 3000 (Dionex LC-Packings, Ámsterdam, Los Países Bajos) equipado con una columna de sílice fundida de 20 cm x 75 µm de diámetro interno empaquetada a medida con 3 µm de 120 Å ReproSil Pur C18 aqua (Dr Maisch GMBH, Ammerbuch-Entringen, Alemania) como se describió anteriormente (Piersma SR, et al., J Proteome Res 2010; 9: 1913-22). Los péptidos fueron atrapados en un cartucho Pepmap C18 de 5 mm x 300 µm de diámetro interno (Dionex LC-Packings, Amsterdam, Países Bajos) y se separaron a 300 nL/min en un gradiente de 60 min. Los espectros de masas de péptidos intactos y los espectros de fragmentación (5 principales) se adquirieron en un espectrómetro de masas híbrido LTQ-FT (Thermo Fisher, Bremen, Alemania). La exclusión dinámica se aplicó con un recuento repetido de 1 y un tiempo de exclusión de 30 segundos. Los péptidos de las cuatro muestras agrupadas (experimento de verificación) se separaron por triplicado en una columna Reprosil C18 Aqua de 75 µm x 50 cm empaquetada a medida (1,9 µm, 120 Å) en un gradiente de 240 minutos (acetronitrilo al 5-32% + ácido acético al 0,5% a 300 nL/min) usando un nanoLC de alta presión U3000 RSLC (Dionex). Los péptidos de elución se midieron en línea mediante un espectrómetro de masas Q Exactive (ThermoFisher Scientific) que funciona en modo de adquisición dependiente de datos. Los péptidos se ionizaron usando un emisor de acero inoxidable a un potencial de +2 kV (ThermoScientific). Los iones peptídicos intactos se detectaron a una resolución de 35000 y los iones fragmentados a una resolución de 17500; el rango de masa de MS fue de 350-1500 Da. La configuración de AGC Target para MS fue 3E6 cargas y para MS/MS 2E5 cargas. Los péptidos se seleccionaron para la fragmentación de HCD en una proporción de relleno inferior del 1% y una ventana de aislamiento de cuadrupolo de 1,5 Da, los péptidos se fragmentaron a una energía de colisión normalizada de 30 eV.

#### Análisis de SRM

El análisis de LC-SRM se realizó en los cuatro grupos del conjunto de verificación. La separación cromatográfica de los péptidos se realizó mediante un sistema Ultimate 3000 RSLC Nano LC (Dionex) equipado con columnas nano-LC empaquetadas a medida que consisten en sílice fundida de 20 cm x 75 µm de diámetro interno empacada a medida con ReproSil Pur C18 aqua de 3 µm 100 Å (Dr Maisch GMBH, Ammerbuch-Entringen, Alemania) como se describió anteriormente<sup>16</sup>. Las muestras se analizaron por triplicado en un instrumento QTRAP 5500 (AB SCIEX, Foster City, CA) operado en modo SRM positivo y equipado con una fuente de nano-electroaspersión con voltaje aplicado de 2,2 kV y una temperatura del calefactor capilar de 225 °C. El modo SRM programado comprendía los siguientes parámetros: ventana de detección de SRM de 900 segundos, tiempo de escaneo objetivo de 3,0 s, gas de cortina de 15, gas de fuente de iones 1 de 25, potencial de reducción de 100, potencial de entrada de 10. Se fijó una resolución Q1 para la unidad y una resolución Q3 para la unidad. La pausa entre intervalos de masa se estableció en 1 ms. Los potenciales de salida de la celda de colisión se establecieron en 36 para todas las transiciones.

#### Desarrollo del ensayo de SRM

Se desarrolló un ensayo de SRM para las proteínas objetivo utilizando el software MRMPilot<sup>MR</sup> versión 2.1 (AB SCIEX, Concord, ON, Canadá). Para cada proteína incluida, se seleccionaron 5 péptidos y se adquirieron a través de JPT Peptide Technologies (Berlín, Alemania) como 'SpikeTides' no purificados. Para cada proteína, se inyectó 1 µL de una mezcla de los 5 péptidos seleccionados a una concentración de aproximadamente 50 fmol por péptido. Para cada uno de los péptidos, el software MRMPilot<sup>MR</sup> se configuró para generar hasta 20 transiciones de SRM teóricamente posibles, cada una de las cuales consiste en el m/z calculado del ión precursor (en cualquier estado de carga predicho por el software) en combinación con los iones del fragmento predicho para cada precursor predicho. Se determinaron los péptidos/transiciones de mayor respuesta a una energía de colisión óptima calculada teóricamente, así como la identidad del péptido a través de MS/MS activada por SRM. Cada análisis de verificación se configuró para detectar 100 de todas las transiciones teóricamente predichas y su energía de colisión óptima teóricamente predicha correspondiente a los 5 péptidos evaluados para cada proteína candidata. La información combinada de cada experimento de adquisición dependiente de información de SRM se utilizó para realizar búsquedas de Mascot contra la base de datos de proteínas Swiss-prot humana y el software MultiQuant<sup>MR</sup> versión 2.1 (AB SCIEX). El software MultiQuant<sup>MR</sup> también se utilizó para generar archivos de métodos para la verificación de péptidos y la optimización de energía de colisión e integrar los resultados de todos los ciclos de optimización. MRMPilot se usó para programar tres transiciones en el tiempo de retención óptimo encontrado experimentalmente y el encontrado experimentalmente para cada péptido. El ensayo final contenía 114 transiciones programadas, 3 para cada péptido (1-5 péptidos para

cada una de las proteínas candidatas) y 6 para 2 péptidos de control externo. Los análisis de SRM en cada grupo se llevaron a cabo por triplicado.

#### Análisis de datos de SRM

5 El área bajo la curva (AUC) y el tiempo de retención de cada transición se determinaron usando el software Multiquant (AB SCIEX). Las transiciones obtenidas fueron verificadas por dos características; en primer lugar, se realizó una verificación del tiempo de retención para revelar falsos positivos o falsos negativos en la asignación de las transiciones. Si el tiempo de retención de tres transiciones pertenecientes a 1 péptido fue > 3 segundos, las tres transiciones se inspeccionaron manualmente para evaluar la exactitud de la asignación de las tres transiciones por el software Multiquant. Si esto ocurrió en más de una de las réplicas técnicas, entonces el péptido para esa muestra no se tuvo en cuenta. En segundo lugar, se verificó la consistencia de la relación entre dos áreas bajo la curva de dos transiciones (Transición 1/ Transición 2, Transición 2/ Transición 3 y Transición 1/ Transición 3) para cada péptido en el conjunto de CRC. Se calculó un porcentaje de desviación estándar relativa (RSD) de estas relaciones en el análisis por triplicado. Si la transición causada por el porcentaje de RSD para dos relaciones era > 20%, las transiciones se inspeccionaron manualmente y si solo una transición parecía ser incorrecta, las dos transiciones que resultaron en < 20% de RSD sobre todos los análisis se seleccionaron para un análisis posterior. Nuevamente, si las tres relaciones de transición tenían un porcentaje de RSD superior a 20, entonces el péptido para esa muestra no se tuvo en cuenta.

#### 20 Búsqueda de bases de datos y análisis estadístico

Se realizaron búsquedas en los espectros de MS/MS contra la base de datos de IPI humana 3.62 (83.947 entradas) usando Sequest (versión 27, rev 12). Se usó Scaffold versión 3.00 (Proteomesoftware, Portland, OR) para organizar los datos y validar las identificaciones de péptidos usando PeptideProphet (probabilidad > 95%) y ProteinProphet (probabilidad de > 99% con 2 péptidos o más en una muestra) (Keller A, et al., Anal Chem 2002; 74: 5383-92; Nesvizhskii AI, et al., Anal Chem 2003; 75: 4646-58). Los programas de software SecretomeP y SignalP se usaron para la predicción de la secreción no clásica y la presencia de un péptido señal (Bendtsen JD, et al., Protein Eng Des Sel 2004; 17: 349-56; Bendtsen JD, et al., J Mol Biol 2004 ; 340: 783-95). Se buscaron los espectros de MS/MS de Q Exactive contra la base de datos IPI humana 3.68 usando MaxQuant 1.2.2.5. La desviación de masa máxima permitida se estableció en 20 ppm para MS y MS/MS. La metilación de cisteína carboxamida se estableció como modificación fija y la oxidación de metionina y la acetilación de la proteína N-terminal se establecieron como modificaciones variables. Las identificaciones se filtraron al 1% de FDR tanto a nivel de péptido como de proteína. Para cada proteína, el número de espectros de MS/MS asignados se sumó a las 10 fracciones y se exportó a Excel. Los recuentos normalizados se calcularon dividiendo los recuentos por proteína por la suma de todos los recuentos por muestra y la multiplicación por la suma promedio en todas las muestras. El análisis diferencial de proteínas presentes en muestras de heces de pacientes con CRC versus muestras de heces de controles sanos se realizó mediante la prueba beta-binomial. La prueba beta-binomial toma en cuenta la variación dentro de la muestra y la variación entre muestras en un solo modelo estadístico (Pham TV, et al., Bioinformatics 2010; 26: 363-9).

#### 40 Análisis de prueba inmunoquímica fecal

Las muestras FIT (OC-sensor®, Eiken Chemical Co., Tokio, Japón) se procesaron con el analizador de mesa OC sensor MICRO (Eiken Chemical) y se analizaron de acuerdo con las instrucciones del fabricante. Se usó un corte de 75 ng/mL para determinar un resultado positivo de la prueba (van Veen W, Mali WP. [Colorectal cancer screening: advice from the Health Council of the Netherlands]. Ned Tijdschr Geneesk 2009; 153:A1441).

#### Resultados

##### Identificación de proteínas humanas en muestras de heces

50 En total, se identificaron 830 proteínas humanas en al menos una de las 22 muestras de heces. De estas, 624 proteínas se identificaron tanto en muestras de heces de CRC como de control, 164 proteínas se detectaron solo en muestras de heces de CRC y 42 proteínas se detectaron solo en muestras de heces de control (véase la Figura 1A). En promedio, el número de proteínas identificadas fue consistente en todas las muestras y no difirió significativamente entre muestras de CRC y de control (326 y 296 proteínas, respectivamente).

La anotación primaria de la localización subcelular de las proteínas fue principalmente el citoplasma (35%) y el espacio extracelular (16%) y no difirió entre muestras de CRC y de control (véase Figura 1B). Cuando se examinaron las secuencias de proteínas para detectar señales de secreción con las herramientas de software, signal P que indica los péptidos señal y el secretoma P que indica la secreción no clásica, se predijo que el 38% de las proteínas tenían un péptido señal y se predijo que el 64% de la mayoría de las proteínas era una proteína secretada. Estos datos indican que un alto porcentaje de proteínas humanas en muestras de heces consiste en proteínas secretadas, por ejemplo, en comparación con solo el 13% de las proteínas con un péptido señal en un lisado de la línea celular (Piersma SR, et al., J Proteome Res 2010; 9: 1913-22).

##### 65 Verificación de la cuantificación de proteínas por LC-MS/MS

La prueba inmunoquímica fecal (FIT) se usa en muchos países como una prueba no invasiva para la detección temprana de CRC, y se basa en la detección de hemoglobina. Para verificar los resultados obtenidos por LC-MS/MS, los recuentos espectrales de hemoglobina se compararon con los valores de FIT (ng/mL) determinados en las mismas muestras de heces. La proteína de hemoglobina adulta es un heterodímero que consta de dos cadenas  $\alpha$  y dos cadenas  $\beta$  (Schechter AN. *Blood* 2008; 112: 3927-38). Como se puede ver en las Figuras 2A y 2B, los datos de LC-MS/MS en las cadenas de hemoglobina  $\alpha$  y  $\beta$  se correlacionaron significativamente con los valores de FIT (ambos  $p = 0,04$ ). Además, se observó una correlación muy fuerte entre los recuentos espectrales de la cadena  $\alpha$  y  $\beta$  (véase la Figura 2C; correlación de Pearson de 0,98,  $p = 1,1 * 10^{-14}$ ).

Otra proteína que se ha detectado con frecuencia en las heces y asociada al CRC es la calprotectina (Bosch LJ, et al., *Molecular tests for colorectal cancer screening. Clin Colorectal Cancer* 2011; 10: 8-23). La calprotectina también es un heterodímero, que consiste en las subunidades S100A8 y S100A9 (Yui S, et al., *Biol Pharm Bull* 2003; 26: 753-60). Se observó una fuerte correlación entre los recuentos espectrales de S100A8 y S100A9 (vease la Figura 2D; correlación de Pearson de 0,92,  $p = 1,2 * 10^{-9}$ ), validando aún más los datos de LC-MS/MS obtenidos de las heces.

Estos resultados en hemoglobina y calprotectina muestran que LC-MS/MS en muestras de heces proporciona una cuantificación robusta de proteínas, y es un enfoque válido para el descubrimiento de marcadores de proteínas.

## Origen de las proteínas de las heces humanas

Los mecanismos de cómo los biomarcadores derivados de tumores terminan en heces se pueden dividir ampliamente en marcadores filtrados, marcadores secretados y marcadores exfoliados (Osborn NK, Ahlquist DA. *Gastroenterology* 2005; 128: 192-206). La hemoglobina es un ejemplo típico de un marcador filtrado de vasos sanguíneos alterados en una lesión neoplásica. Los marcadores secretados y exfoliados pueden derivarse de las células epiteliales que recubren el lumen colorrectal, pero también pueden originarse en las células del estroma circundante, tal como las células inmunes. Un análisis de superposición con un conjunto de datos de proteínas plasmáticas obtenido previamente (datos no publicados) junto con las bases de datos Human Proteome Organisation (HUPO), Human Plasma Proteome Project (HPPP) públicamente disponibles ([www.hupo.org/research/hppp](http://www.hupo.org/research/hppp); Omenn GS, et al., *Proteomics* 2005; 5: 3226-45; States DJ, et al., *Nat Biotechnol* 2006; 24: 333-8) (lista altamente confiable) reveló que el 21,6% de las 830 proteínas identificadas posiblemente se originan en la sangre. Para estimar cuántas de las 830 proteínas humanas identificadas se originan a partir de células epiteliales, se realizó un análisis de superposición con proteínas detectadas en líneas celulares de CRC (datos no publicados). Casi la mitad de estas proteínas también se identificaron en las líneas celulares de CRC que sugieren un origen epitelial.

## Proteínas de heces que discriminan a los pacientes con CRC de los sujetos de control

El análisis de grupos jerárquicos no supervisados de proteínas humanas identificadas en todas las muestras de heces de CRC y control revelaron dos grupos. El grupo uno agrupó nueve muestras de heces de CRC y el grupo dos agrupó las diez muestras de heces de control y tres muestras de heces de CRC (datos no mostrados). Por lo tanto, las muestras de heces de pacientes con CRC muestran un patrón específico de expresión de proteínas en comparación con las heces de los sujetos de control. Por lo tanto, este patrón de expresión de proteínas se puede aplicar para discriminar la mayoría de las muestras de heces de CRC de las muestras de heces de control sin una selección adicional de proteínas específicas. El conjunto completo de datos de proteínas de heces se analizó con estadísticas beta binomiales para identificar posibles proteínas asociadas al cáncer (Pham TV, et al., *Bioinformatics* 2010; 26: 363-9). De las 830 proteínas de heces humanas, se detectaron de manera diferencial 221 proteínas, de las cuales los niveles de 134 proteínas fueron significativamente más altos en muestras de heces de CRC en comparación con las de control ( $p < 0,05$ ; Tabla 2), mientras que 87 proteínas fueron significativamente más bajas en las muestras de heces de control en comparación con líneas de CRC.

El análisis de agrupamiento basado en las 221 proteínas detectadas diferencialmente en muestras de heces de CRC y de control reveló también dos agrupamientos (datos no mostrados). Nuevamente, las nueve muestras de heces de CRC agrupadas en un grupo, y las mismas tres muestras de heces de CRC mencionadas anteriormente agrupadas junto con las muestras de heces de control en el segundo grupo.

Dentro del segundo grupo, las tres muestras de heces de CRC se agruparon con una muestra de control. El patrón de expresión de proteínas de estas tres muestras de heces de CRC contenía aspectos tanto de las muestras de heces de control como de las otras nueve muestras de heces de CRC, lo que indica que una selección de estas proteínas podría discriminar estas muestras de CRC de los controles.

Además, tres de cada cuatro pacientes con CRC negativo para FIT se agruparon junto con pacientes con CRC positivo para FIT. Por lo tanto, este grupo de proteínas contiene potenciales biomarcadores candidatos que complementan el FIT.

## Verificación por nanoLC-MS/MS de semi-alto rendimiento

Para verificar estos resultados iniciales, se crearon cuatro grupos a partir de un conjunto independiente de 20 muestras de heces. Además de los casos de CRC y los controles negativos, también se incluyeron muestras de pacientes con adenomas. De hecho, dado que la mayoría de los adenomas no progresan a cáncer, con el propósito de la detección de CRC, solo la detección de adenomas con alto riesgo de progresión además del CRC es importante (Levin B et al., Gastroenterology 2008; 134: 1570-1595). Por lo tanto, se incluyeron grupos de muestras de heces de pacientes con adenoma no avanzado, así como adenomas avanzados (generalmente considerados de mayor riesgo de progresión). En estas muestras de verificación, se identificaron 414 proteínas humanas. De las 830 proteínas humanas identificadas en el conjunto de descubrimiento, 331 (40%) también se detectaron en el conjunto de verificación (Véase la Tabla 5). De las 134 proteínas humanas significativamente enriquecidas en muestras de heces de CRC, 63 (47%) también se detectaron en el grupo de verificación. Las proteínas detectadas tanto en el grupo de descubrimiento como en el de verificación representaron proteínas con una abundancia significativamente mayor (recuento espectral promedio de 15) en comparación con las proteínas que no se encontraron en el grupo de verificación (recuento espectral promedio de 4) ( $p = 3,5 \cdot 10^{-27}$ ).

Prueba del concepto de biomarcadores de proteína candidatos para la validación mediante MS dirigida

De las 134 proteínas significativamente enriquecidas en muestras de heces de CRC, un subconjunto de proteínas con un posible origen epitelial (es decir, detectadas en líneas celulares de CRC), que se enriquecieron significativamente en pacientes con CRC negativo para FIT en comparación con los controles, y se recuperaron en el grupo de verificación en los grupos de adenoma avanzado y/o CRC ( $n = 29$ ) se seleccionaron para una validación adicional. Los primeros resultados para 13 de estas proteínas confirmaron los resultados obtenidos por LC-MS/MS en el grupo de verificación. Estos incluyeron 6 proteínas detectadas en grupos de CRC y/o de adenoma avanzado mientras no se detectaron en controles o adenomas (por ejemplo, C4B y glucosa-6-fosfato isomerasa (GPI) (Figura 4); y 7 proteínas detectadas en niveles bajos en los controles y adenomas mientras está más presente en adenomas avanzados y CRC, por ejemplo, PRTN3 y A2M (Figura 4).

Proteínas que complementan FIT para detectar CRC

Los nuevos biomarcadores para la detección temprana de CRC deberían funcionar significativamente mejor que FIT, o deberían complementar FIT, para aumentar la precisión del diagnóstico. Como se esperaba de su correlación con FIT, la hemoglobina  $\alpha$  y  $\beta$  estaban significativamente más abundantemente presentes en muestras de heces de CRC en comparación con las de control. Los datos revelaron varias proteínas cuyos niveles fueron significativamente más altos en las muestras de heces de CRC en comparación con las muestras de heces de control con un poder discriminativo más alto que la hemoglobina. Es interesante ver si estas proteínas detectan las mismas muestras de CRC que la hemoglobina, o si detectan otras muestras de CRC. La Figura 3 muestra un ejemplo de una de estas proteínas (C4A/C4B) que detecta a todos los pacientes con CRC negativo para FIT, sin detectar el sujeto de control, alcanzando así una sensibilidad y especificidad del 100% en este entorno. Sin embargo, no solo las proteínas con un poder discriminativo más alto que la hemoglobina tienen potencial para complementar FIT.

Por esta razón, se investigaron proteínas que mostraron niveles significativamente más altos en las muestras de heces de CRC negativas para FIT en comparación con todos los controles. De hecho, de las 134 proteínas con niveles significativamente más altos en muestras de heces de CRC versus control, alrededor del 90% también mostró niveles significativamente más altos en las muestras de heces de CRC negativas para FIT. Esto indica que estos candidatos tienen un alto potencial para ser de valor agregado a la detección actual de hemoglobina.

El presente estudio ha realizado un análisis proteómico en profundidad en muestras de heces humanas, proporcionando una lista de 134 proteínas humanas que se han enriquecido significativamente en muestras de heces de pacientes con CRC, de las cuales varias mostraron un poder discriminativo muy significativo. Por lo tanto, se han identificado marcadores discriminativos para mejorar las pruebas FIT actuales.

Ejemplo 2

Generación de un conjunto de datos de biomarcadores a partir de tejido de CRC y tejido de colon normal emparejado con el paciente, para la detección en biofluidos.

Material y métodos

Pacientes

Se incluyeron en este estudio un total de cuatro pacientes que se sometieron a resección quirúrgica en el centro médico de la Universidad VU (Ámsterdam, Los Países Bajos). La recopilación, el almacenamiento y el uso de tejidos y datos de pacientes se realizaron de acuerdo con el Código para el uso secundario adecuado de tejidos humanos en Los Países Bajos (Societies DFOBS. [Http://www.federa.org/](http://www.federa.org/)). Un patólogo inspeccionó todas las muestras para obtener información sobre el tamaño del tumor, el tumor y el estadio ganglionar, el grado de diferenciación, la diferenciación mucínica. Para una visión general de las características clinicopatológicas, Véase la Tabla 1.

## Manejo de tejidos y colección de secretomas de tejidos

La recolección de secretoma tisular se realizó como se describió anteriormente (Celis JE. Et al., Mol. Cell Proteomics 2004; 3: 327-4). En resumen, después de la resección quirúrgica, la muestra se transfirió inmediatamente al departamento de patología, donde un patólogo extirpó una parte representativa del tumor y la mucosa del colon normal adyacente (cerca de un margen de resección no afectado). Estas piezas de tejido se cortaron en cubos de aproximadamente 1 mm<sup>3</sup> y se enjuagaron en PBS para eliminar la sangre y las partículas de heces. Posteriormente, las partículas de tejido se incubaron en 100 µL de PBS durante 1 hora a 37 °C. Después de esta incubación, las muestras se centrifugaron brevemente (2000 rpm a 4 °C durante 2 minutos) y el sobrenadante se transfirió a un nuevo tubo eppendorf. Los sobrenadantes se centrifugaron a la velocidad máxima (13.200 rpm a 4 °C durante 20 minutos) para eliminar todas las células y residuos restantes. Además, las fracciones solubles, denominadas 'secretomas de tejido', se almacenaron a -80 °C hasta su uso posterior. Los tejidos se procesaron mediante fijación estándar con formalina y se incluyeron en parafina para evaluación histológica (figura complementaria 4). El estado de inestabilidad de microsátélites (MSI) se determinó utilizando el sistema de análisis MSI versión 1.2 (Promega Corporation, Madison, EE. UU.) como se describió anteriormente (de Wit M. et al., Gut 2012; 61: 855-64).

## Electroforesis en gel y preparación de muestras para análisis de proteómica.

Las concentraciones de proteína se determinaron usando el ensayo de proteína BCA (Pierce, Thermo Fisher Scientific, Rockford, EE. UU.). Veinte µg de proteínas se separaron por electroforesis en gel usando un gel SDS-PAGE con gradiente 4-12% 1D prefabricado (Invitrogen, Carlsbad, EE. UU.). Para imágenes en gel, véase figura complementaria 2a-c. Los geles se fijaron en etanol al 50% que contenía ácido fosfórico al 3% y se tiñeron con Coomassie R-250. Los geles se lavaron en agua MilliQ y se almacenaron a 4 °C hasta el procesamiento para la digestión en gel. Cada carril se cortó en 10 bandas de gel y se procesó adicionalmente en péptidos tripticos como se describió anteriormente (Piersma SR, et al., J. Proteome Res. 2010; 9: 1913-22). El volumen de las fracciones peptídicas desaladas se redujo a 50 µL en una centrifuga de vacío.

## Análisis de proteómica por nanoLC-MS/MS

Los péptidos se separaron mediante un sistema de nanoLC Ultimate 3000 (Dionex LC-Packings, Ámsterdam, Los Países Bajos) equipado con una columna de sílice fundida de 20 cm x 75 µm de diámetro interno empaquetada a medida con 3 µm de 120 Å ReproSil Pur C18 aqua (Dr Maisch GMBH, Ammerbuch-Entringen, Alemania). Después de la inyección, los péptidos se atraparon a razón de 6 µL/min en una columna trampa de 1 cm x 100 µm de diámetro interno con 5 µm de 120 Å ReproSil C18aqua al 2% de tampón B (tampón A: 0,05% de ácido fórmico en MQ; tampón B: 80% acetonitrilo + 0,05% de ácido fórmico en MQ) y se separó a razón de 300 nL/min en un gradiente de tampón B al 10-40% en 60 min. Los péptidos de elución se ionizaron a 1,7 kV en una fuente de nanoaspersión basada en Chip Nanomate Triversa usando un acoplador LC Triversa (Advion, Ithaca, NJ). Los espectros de masas de péptidos intactos y los espectros de fragmentación se adquirieron en un espectrómetro de masas híbrido LTQ-FT (Thermo Fisher, Bremen, Alemania). Las masas intactas se midieron a una resolución de 50.000 en la celda de ICR usando un valor objetivo de 1 x 10<sup>6</sup> cargas. Paralelamente, después de un escaneo previo de FT, las 5 señales peptídicas principales (estados de carga 2+ y superiores) se enviaron a MS/MS en la trampa de iones lineal (ancho de aislamiento de 3 amu, 30 ms de activación, 35% de energía de activación normalizada, Valor Q de 0,25 y un umbral de 5000 recuentos). La exclusión dinámica se aplicó con un recuento repetido de 1 y un tiempo de exclusión de 30 segundos.

## Búsqueda en bases de datos

Para identificar proteínas de los datos adquiridos, los espectros MS/MS se buscaron en la base de datos IPI humana v3.59 (80.128 entradas) usando Sequest (versión 27, rev 12), que es parte del paquete de análisis de datos BioWorks 3.3 (Thermo Fisher, San José, CA). Tras la búsqueda en la base de datos, los archivos DTA y OUT se importaron a Scaffold Scaffold\_2\_06\_01 (Proteome software, Portland, OR). Se usó Scaffold para organizar los datos de la banda de gel y para validar las identificaciones de péptidos usando el algoritmo de Peptide Prophet (Nesvizhskii AI, et al., Anal. Chem. 2003; 75: 4646-58). Solo se retuvieron las identificaciones con una probabilidad > 95%. Posteriormente, se aplicó el algoritmo Protein Prophet (Keller A, et al., Anal. Chem. 2002; 74: 5383-92) y se retuvieron las identificaciones de proteínas con una probabilidad de > 99% con 2 péptidos o más en al menos una de las muestras. Se agruparon las proteínas que contenían péptidos similares y que no podían diferenciarse basándose únicamente en el análisis MS/MS.

## Minería de datos y análisis estadístico

Para cada proteína identificada, el número de espectros asignados se exportó a Excel. El análisis diferencial de las muestras se realizó mediante una prueba beta-binomial pareada como se describió anteriormente (Pham TV et al., Expert Rev. Mol. Diagn. 2012; 12: 343-59). (Szkarczyk D et al., Nucleic Acids Res. 2011; 39: D561-8). Esta modificación dio como resultado una prueba pareada, teniendo en cuenta el origen de estas muestras, por ejemplo, comparando firmas de proteínas entre dos tejidos derivados del mismo paciente. Se recuperó información proteica general adicional usando Ingenuity Pathway Analysis (Ingenuity® Systems, www.ingenuity.com). Las interacciones proteína-proteína conocidas y predichas se investigaron utilizando STRING versión 9.0 (www.string-db.org)

(Szkłarczyk D, et al., *Nucleic Acids Res.* 2011; 39: D561-8). Para los análisis de ontología de grupos y genes se utilizó la plataforma Cytoscape para el análisis de redes ([www.cytoscape.org](http://www.cytoscape.org)) (Smoot ME et al., *Bioinformatics* 2011; 27: 431-2), utilizando los complementos ClusterONE versión 0.93 para la agrupación y BINGO versión 2.44 para el análisis de procesos biológicos dentro de las redes obtenidas con base en GO annotations Maere S et al., *Bioinformatics* 2005; 21: 3448-9 y Nepusz T et al., *Nat Meth* 2012; 9: 471-2). Además, los dominios transmembrana y las secuencias de péptidos señal se investigaron usando el servidor secretomeP 2.0 (Bendtsen JD et al., *Protein Eng. Des. Sel.* 2004; 17: 349-56).

## Resultados

### Identificación de proteínas secretadas en secretomas de tejido de CRC y normal

Se recogieron secretomas de tejido de cuatro CRC y se combinaron con tejido de colon normal adyacente como se describió anteriormente por Celis et al., (Celis JE et al., *Mol. Cell Proteomics* 2004; 3: 327-44) y se procesaron para espectrometría de masas. Se encontró que el tumor del paciente 4 era MSI y los tumores de los pacientes 1-3 eran microsatélites estables (detalles del paciente en la tabla 8). Se identificaron un total de 2703 proteínas no redundantes en las secreciones tumorales de los cuatro secretomas de tejido de CRC y sus contrapartes normales. En promedio, se identificaron 1986 proteínas por muestra, que van desde 1264 a 2292 proteínas.

Del total de 2703 proteínas, 2366 se identificaron en los secretomas tumorales así como en los secretomes de tejido normal, 283 solo en los secretomes tumorales y 54 solo en los secretomes de tejido normal. El número de proteínas identificadas fue significativamente mayor en los secretomas de CRC que en los secretomas de tejido normal (2198 y 1775 en promedio, respectivamente,  $p = 0,03$ ), mientras que el número de identificaciones únicas en las muestras no difirió significativamente (figura complementaria 2b y c,  $P = 0,2$ ). En total, se identificaron 961 de 2420 (40%) de proteínas en los cuatro secretomas de tejido normal y 1627 de 2649 (61%) en los cuatro secretomas de cáncer.

### Proteínas enriquecidas en los secretomas de cáncer

Para obtener una visión general del conjunto de datos, se realizó una agrupación jerárquica no supervisada utilizando los datos del recuento espectral normalizado de todas las 2703 proteínas identificadas. Se identificaron dos grupos principales; uno que contiene la muestra de tejido normal del paciente 3 y el otro que contiene todas las demás muestras, esto podría explicarse por la menor cantidad de proteínas identificadas en esta muestra en comparación con las otras ( $N = 1264$ ). En el segundo grupo nuevamente se formaron dos subgrupos; uno que contiene las otras tres muestras de tejido normal y el otro todas las muestras de cáncer. Dentro del grupo de cáncer, las muestras de los pacientes 1-3 formaron un grupo separado de la muestra del paciente 4, lo que puede explicarse por los diferentes antecedentes moleculares de esta muestra, es decir, este último tumor era MSI mientras que los tumores de los pacientes 1-3 eran microsatélites estables (tabla 8). En general, el análisis de conglomerados no supervisado indicó que el patrón de expresión de proteínas era más similar entre las cuatro muestras de tumor que entre el tumor emparejado y los tejidos normales del mismo paciente. Los biomarcadores potenciales deben enriquecerse en los secretomas tumorales en comparación con los secretomas de tejido normal. Para identificar tales proteínas, el conjunto completo de datos de proteínas se analizó adicionalmente con una prueba estadística para muestras emparejadas, es decir, teniendo en cuenta la relación entre muestras de tejido normal y cancerosas de pacientes individuales, para identificar proteínas asociadas a CRC (Pham TV et al., *Expert Rev. Mol. Diagn.* 2012; 12: 343-59). De las 2703 proteínas, 522 proteínas estaban significativamente diferencialmente presentes ( $P < 0,05$ ), 409 de las cuales eran más abundantes en secretomas tumorales en comparación con secretomas de tejido normal, lo que representa las proteínas de mayor valor probable para la discriminación entre pacientes con CRC y controles normales (Véase la tabla 6).

### Análisis de superposición de proteínas enriquecidas en los secretomas de cáncer y las proteínas de heces.

Los biomarcadores de proteínas que pueden detectarse en sangre o heces son de interés ya que podrían aplicarse en un entorno clínico estándar junto con los marcadores usados habitualmente tales como CEA y CA19-9 o hemoglobina. La medición de moléculas directamente en la muestra biológica que se utilizará para la detección, es decir, las heces, puede producir biomarcadores que son estables y se pueden detectar de manera confiable en las muestras de heces. Un análisis de superposición con un conjunto de datos de proteínas de heces previamente obtenido (datos no publicados) reveló que 383 de las 2703 proteínas identificadas son posiblemente detectables en las heces. De las 409 proteínas significativamente más abundantes en las secretomas de CRC en comparación con los secretomas normales, 16 proteínas de estas también se detectaron en las heces (Véase la tabla 7).

### Aplicabilidad industrial

La presente descripción proporciona así un conjunto de biomarcadores para la detección del cáncer colorrectal, o la susceptibilidad al cáncer colorrectal. Los biomarcadores permiten un diagnóstico precoz de manera que es posible la extirpación quirúrgica de un tumor en etapa temprana antes de la metástasis.

Lo anterior describe ampliamente la presente invención sin limitación a realizaciones particulares. Las variaciones y

modificaciones serán fácilmente evidentes para los expertos en la materia. El alcance de la invención está definido por las siguientes reivindicaciones.

**REIVINDICACIONES**

1. Un método para detectar la presencia de adenoma colorrectal avanzado, comprendiendo el método:
  - 5 examinar una muestra de heces obtenida de un individuo para dos o más biomarcadores que comprenden mieloperoxidasa (MPO) y uno o más biomarcadores seleccionados del grupo que consiste en: azurocidina 1 (AZU1), catepsina G (CTSG), defensina, alfa 1 (DEFA1), elastasa, expresada en neutrófilos (ELANE), integrina, alfa M (subunidad del receptor 3 del componente 3 del complemento) (ITGAM), subunidad 1 de la proteína de unión a la caperuza nuclear 80 kDa (NCBP1), proteinasa 3 (PRTN3), proteína S100 de unión al calcio A9 (S100A9), inhibidor de la serpina peptidasa, clado B (ovoalbúmina), miembro 10 (SERPINB10), anexina A1 (ANXA1), repetición de IAP baculoviral que contiene 6 residuos (BIRC6), componente 2 del complejo de mantenimiento de minicromosomas (MCM2), quiescina Q6 sulfhidril oxidasa 1 (QSOX1), anexina A2 (ANXA2), inhibidor de la serpina peptidasa, clado B (ovoalbúmina), miembro 5 (SERPINB5), en el que la presencia o aumento de la expresión de MPO y los uno o más biomarcadores, en relación con una muestra de control, es indicativo de que el individuo está en riesgo de sufrir o está sufriendo de adenoma colorrectal avanzado.
  - 10
  - 15
  2. El método de la reivindicación 1, que comprende la detección de dos o más de los biomarcadores adicionales.
  3. El método de una cualquiera de las reivindicaciones anteriores, que comprende la detección de cuatro o más de los biomarcadores adicionales.
  - 20
  4. El método de una cualquiera de las reivindicaciones anteriores, que comprende la detección de los dieciséis biomarcadores.
  - 25
  5. El método de una cualquiera de las reivindicaciones anteriores, en el que la muestra de heces se analiza para los dos o más biomarcadores usando espectrometría de masas dirigida.
  6. El método de una cualquiera de las reivindicaciones 1 a 4, en el que la muestra de heces se analiza para los dos o más biomarcadores usando agentes de unión capaces de unirse a los dos o más biomarcadores.
  - 30
  7. El método de la reivindicación 6, en el que los agentes de unión son anticuerpos o fragmentos de los mismos.
  8. El método de la reivindicación 7, en el que los anticuerpos o fragmentos de los mismos se seleccionan del grupo que consiste en scFv, Fab y un dominio de unión de una molécula de inmunoglobulina.
  - 35
  9. El método de una cualquiera de las reivindicaciones 1-4, en el que la detección se realiza usando una matriz, opcionalmente en la que la matriz es una matriz basada en perlas o una matriz basada en una superficie.
  - 40
  10. El método de una cualquiera de las reivindicaciones anteriores, en el que la muestra de control comprende una muestra biológica de un individuo que se sabe que está libre de adenoma colorrectal avanzado.
  11. El método de una cualquiera de las reivindicaciones anteriores, en el que la muestra biológica también se analiza mediante la prueba inmunquímica fecal.
  - 45
  12. Uso de dos o más biomarcadores que comprenden mieloperoxidasa (MPO), y uno o más biomarcadores seleccionados del grupo que consiste en: azurocidina 1 (AZU1), catepsina G (CTSG), defensina, alfa 1 (DEFA1), elastasa, expresada en neutrófilos (ELANE), integrina, alfa M (subunidad del receptor 3 del componente 3 del complemento) (ITGAM), subunidad 1 de la proteína de unión a la caperuza nuclear 80 kDa (NCBP1), proteinasa 3 (PRTN3), proteína S100 de unión al calcio A9 (S100A9), inhibidor de la serpina peptidasa, clado B (ovoalbúmina), miembro 10 (SERPINB10), anexina A1 (ANXA1), repetición de IAP baculoviral que contiene 6 residuos (BIRC6), componente 2 del complejo de mantenimiento de minicromosomas (MCM2), quiescina Q6 sulfhidril oxidasa 1 (QSOX1), anexina A2 (ANXA2), inhibidor de la serpina peptidasa, clado B (ovoalbúmina), miembro 5 (SERPINB5) como biomarcador en muestras de heces para diagnosticar o predecir la presencia de adenoma colorrectal avanzado en un individuo.
  - 50
  - 55
  13. Uso de una matriz para determinar si un individuo está en riesgo de sufrir o está sufriendo adenoma colorrectal avanzado de acuerdo con el método de cualquiera de las reivindicaciones 1-12, comprendiendo la matriz uno o más agentes de unión como se define en cualquiera de las reivindicaciones 6 a 8.
  - 60
  14. Uso de mieloperoxidasa (MPO) como biomarcador en muestras de heces para detectar la presencia de adenoma colorrectal avanzado.

Figura 1A

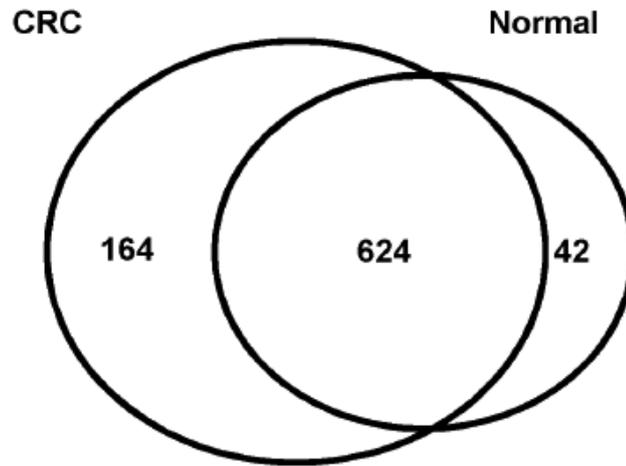


Figura 1B

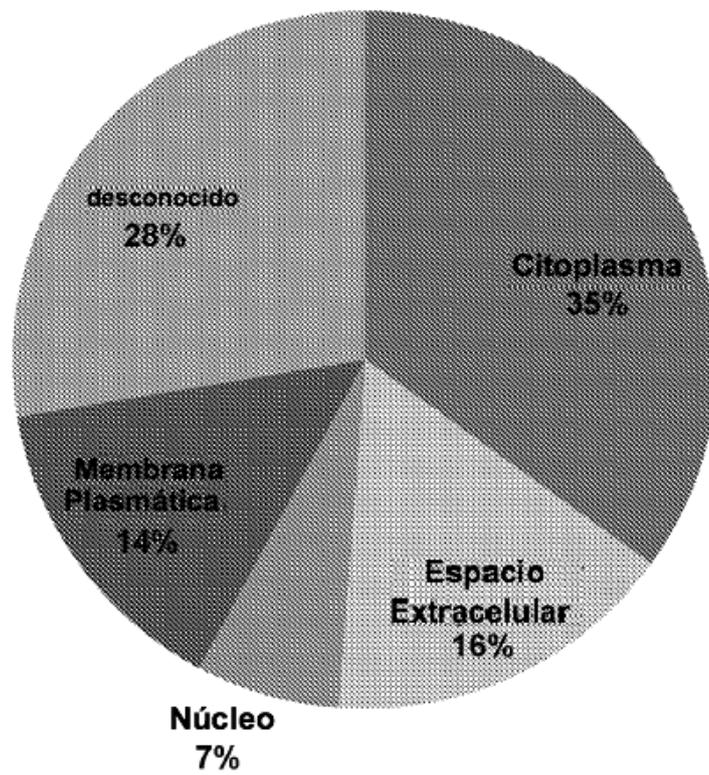


Figura 2

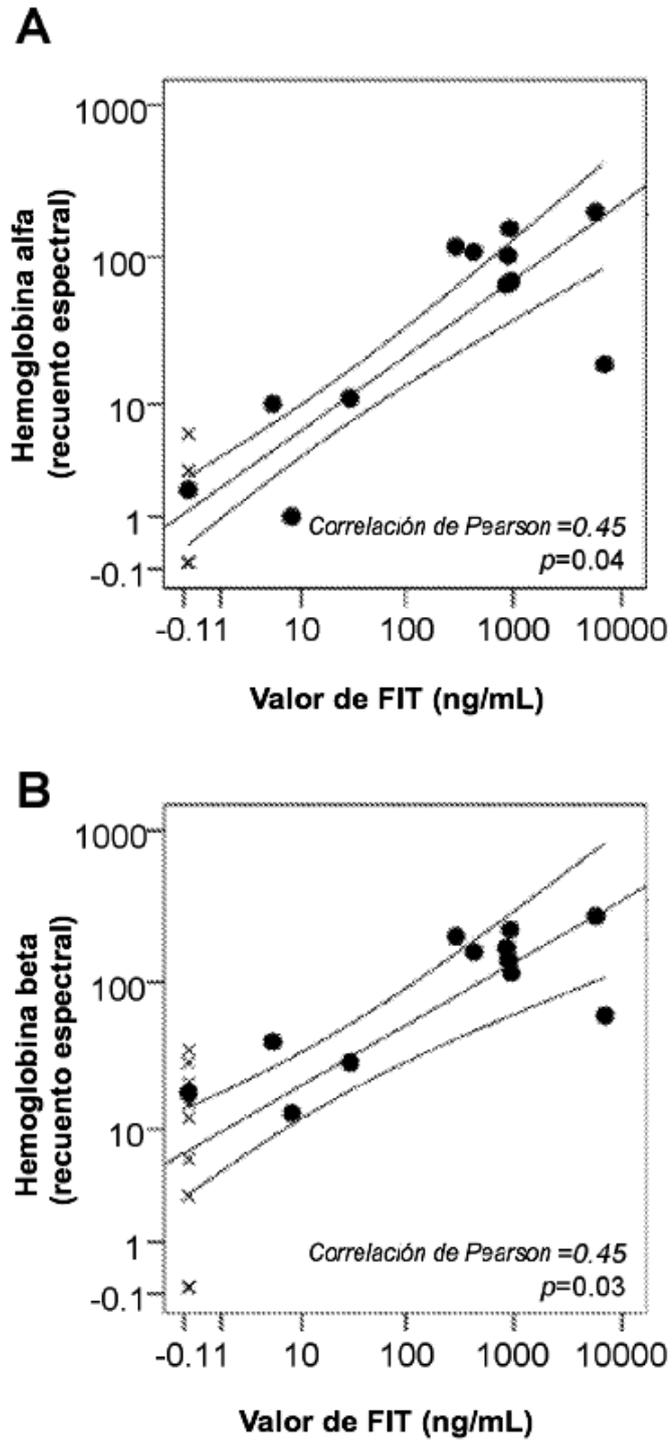


Figura 2 (continuación)

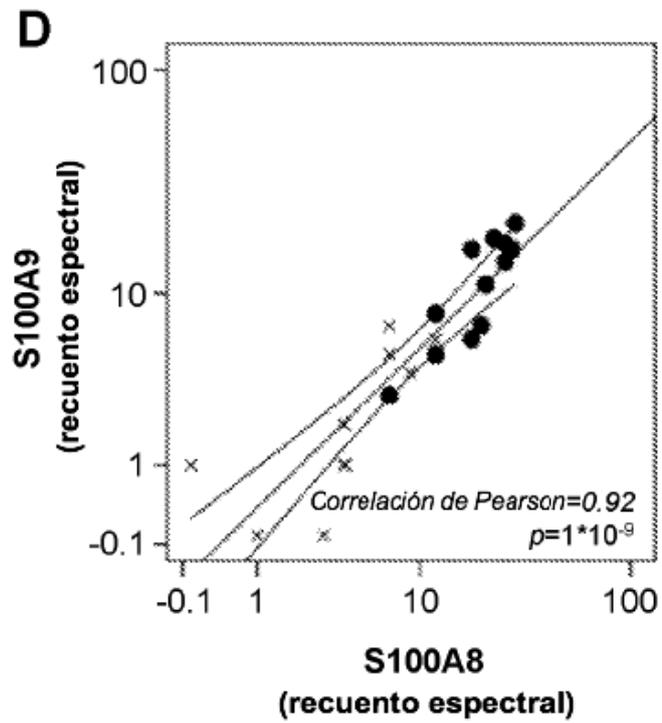
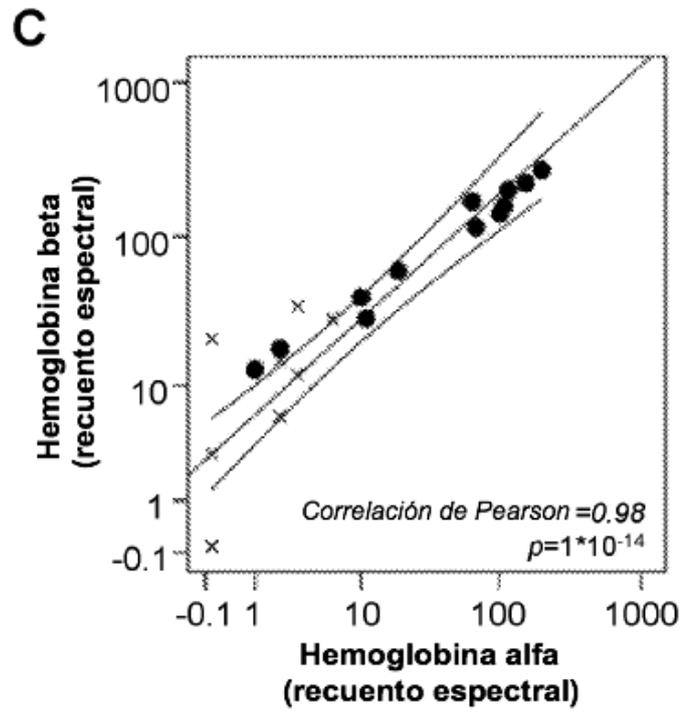


Figura 3

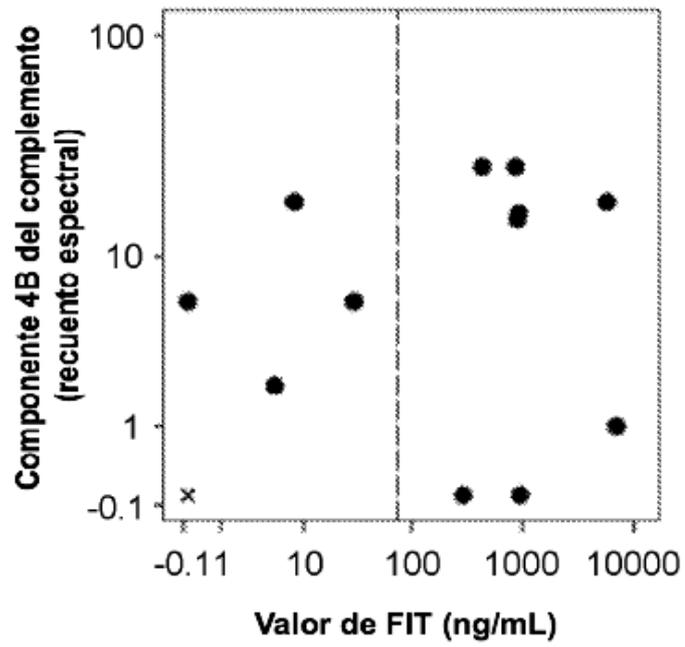


Figura 4

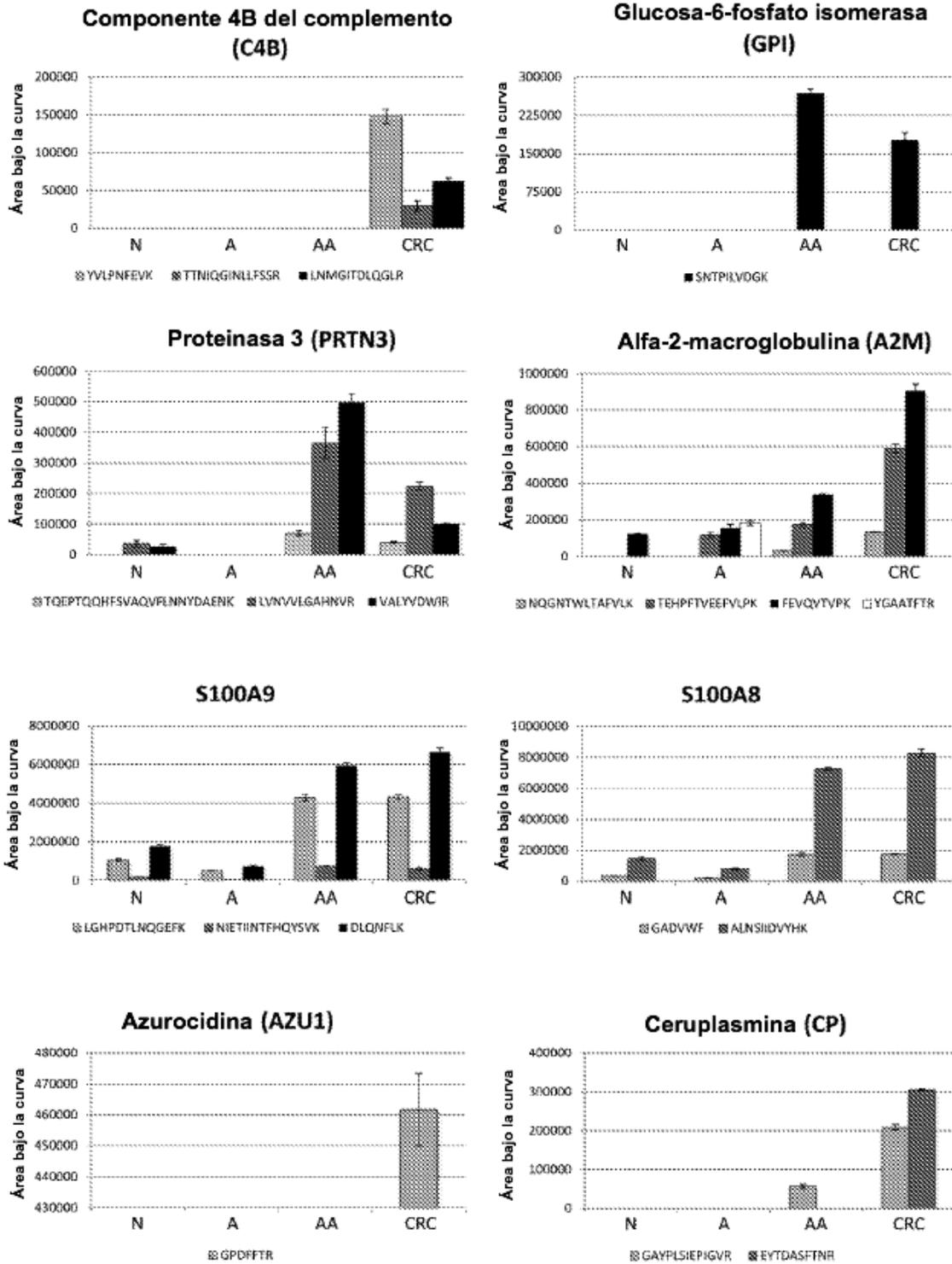


Figura 5

